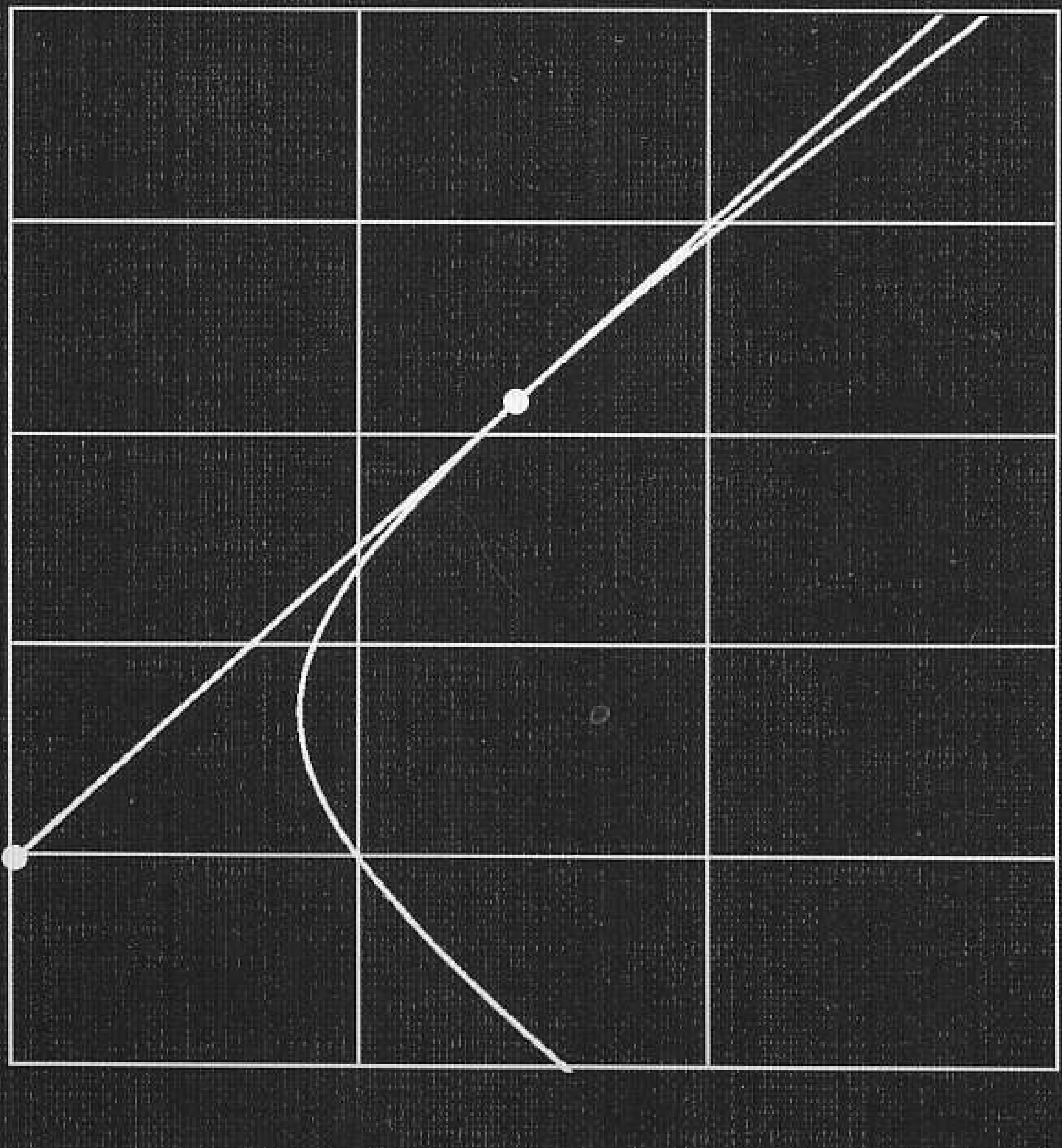
П.К. Катышев
Я.Р. Магнус
А.А. Пересецкий
С.В. Головань

# Сборник задач к начальному курсу эконометрики

$$\hat{\beta}_{RE} = W\hat{\beta}_{B} + (I_k - W)\hat{\beta}_{FE}$$



Д Е Л О

П.К. Катышев, Я.Р. Магнус, А.А. Пересецкий, С.В. Головань

# Сборник задач к начальному курсу ЭКОНОМЕТРИКИ

Допущено Министерством образования
и науки Российской Федерации
в качестве учебного пособия для студентов
высших учебных заведений,
обучающихся по специальности
"Математические методы в экономике"

Академия народного хозяйства
при Правительстве Российской Федерации

Москва
Издательство "ДЕЛО"
2007

УДК 330.43(075.8)
ББК 65в6я73
К29

**Катышев П.К., Магнус Я.Р., Пересецкий А.А., Головань С.В.**

К29 Сборник задач к начальному курсу эконометрики: Учеб. пособие. – 4-е изд., перераб. и доп. — М.: Дело, 2007. – 368 с.

ISBN 978-5-7749-0459-4

Книга содержит подробные решения задач и упражнений по эконометрике из 7, 8-го изданий учебника: *Магнус Я.Р., Катышев П.К., Пересецкий А.А.* Эконометрика. Начальный курс (М.: Дело, 2005, 2007). В книге приведены также условия задач, поэтому она может использоваться не только в комплекте с указанным базовым учебником, но также с любым другим учебником по эконометрике. Книга является первым изданием подобного типа на русском языке.

В новом издании значительно увеличилось количество задач, появились главы, соответствующие новым главам в учебнике, в частности, «Панельные данные». Включены упражнения с реальными данными, доступными для читателя на web-сайте книги.

В первую очередь книга будет полезна студентам и аспирантам, изучающим эконометрику, а также преподавателям эконометрики и специалистам по прикладной экономике и финансам.

**УДК 330.43(075.8)**
**ББК 65в6я73**

ISBN 978-5-7749-0459-4

# Оглавление

Глава 1. Предисловие . . . . . . . . . . . . . . . . . . . 4

Глава 2. Модель парной регрессии 6

Глава 3. Модель множественной регрессии . . . . . . . . . . 35

Глава 4. Различные аспекты множественной регрессии . . . . 104

Глава 5. Некоторые обобщения множественной регрессии . . . . 156

Глава 6. Гетероскедастичность и корреляция по времени . . . . . . 170

Глава 7. Прогнозирование в регрессионных моделях 224

Глава 8. Инструментальные переменные 230

Глава 9. Системы регрессионных уравнений . . . . . . . . . 236

Глава 10. Метод максимального правдоподобия в моделях регрессии 246

Глава 11. Временные ряды 269

Глава 12. Дискретные зависимые переменные и цензурированные выборки . . . . . . . . . . . . . . . . . . . . . . . 302

Глава 13. Панельные данные . . . . . . . . . . . . . . . . 337

Глава 14. Предварительное тестирование: введение . . . . . 354

Глава 15. Эконометрика финансовых рынков . . . . . . . . 359

# Глава 1

# Предисловие

Четвертое издание задачника содержит решения задач, включенных в книгу Я. Р. Магнуса, П. К. Катышева, А. А. Пересецкого «Эконометрика. Начальный курс» (изд. 7, 8-е. М.: Дело, 2005, 2007).

Количество задач в четвертом издании по сравнению с третьим изданием задачника увеличилось примерно в 2 раза. Особенностью этого издания является наличие компьютерных упражнений, данные для которых доступны в Интернете по адресу `http://econometrics.nes.ru/mkp/`.

Данный сборник содержит как относительно простые задачи, которые доступны обучающимся по программе бакалавриата, так и более сложные задачи, которые целесообразно решать при обучении по магистерской программе. Решение отдельных задач содержит изложение тех теоретических результатов, которые недостаточно подробно рассмотрены в нашем учебнике.

Для облегчения изложения мы приводим полностью условия задач, что позволяет использовать данный задачник в комплекте с любым учебником эконометрики. Однако при описании решений часто даются ссылки на формулы или разделы нашей книги, поэтому представляется целесообразным использовать данное руководство совместно с указанной книгой. Мы надеемся, что это руководство будет полезно для преподавателей курса эконометрики, студентов и аспирантов, а также для специалистов по прикладной экономике и финансам.

При подготовке данного руководства огромную помощь нам оказали несколько поколений студентов Российской экономической школы, которым мы выражаем искреннюю благодарность.

Мы благодарим профессоров Тилбургского университета Артура ван Соста и Бертрана Меленберга за плодотворные обсуждения материалов книги, а также Дмитрия Данилова за помощь в составлении решений задач главы 14.

Мы в огромном долгу перед Надеждой Викторовной Андриановой, которая кропотливо правила все издания нашей книги.

Мы благодарны Нидерландской организации научных исследований (NWO) за финансовую поддержку подготовки второго издания данной книги.

Наконец, мы считаем своим приятным долгом поблагодарить Центр экономических исследований Тилбургского университета (Нидерланды) за возможность стажировки и научных визитов, финансовую поддержку, что в значительной степени способствовало появлению нового издания.

Конечно, все оставшиеся незамеченными неточности — наши.

Тилбург/Москва, декабрь 2004 г.

# Глава 2

# Модель парной регрессии

## Задача 2.1

Наблюдения 16 пар $(X, Y)$ дали следующие результаты:

$$\sum Y^2 = 526, \qquad \sum X^2 = 657, \qquad \sum XY = 492,$$
$$\sum Y = 64, \qquad \sum X = 96.$$

Оцените регрессию $Y_t = \alpha + \beta X_t + \varepsilon_t$ и проверьте гипотезу, что коэффициент $\beta$ равен 1.0.

## Решение

1) Оценка параметров регрессии. Из формул (2.4а), (2.4б)[1] получаем:

$$\widehat{\beta} = \frac{n\sum X_t Y_t - \left(\sum X_t\right)\left(\sum Y_t\right)}{n\sum X_t^2 - \left(\sum X_t\right)^2} = \frac{16 \cdot 492 - 96 \cdot 64}{16 \cdot 657 - 96^2} = \frac{4}{3} \approx 1.33,$$

$$\widehat{\alpha} = \frac{1}{n}\sum Y_t - \frac{1}{n}\sum X_t \widehat{\beta} = \overline{Y} - \overline{X}\,\widehat{\beta} = \frac{64}{16} - \frac{96}{16} \cdot \frac{4}{3} = -4.$$

Следовательно, МНК-оценки для $\beta$ и $\alpha$ таковы: $\widehat{\beta} = 4/3$, $\widehat{\alpha} = -4$. Сумма квадратов остатков:

$$\begin{aligned}\sum e_t^2 &= \sum(Y_t - (\widehat{\alpha} + \widehat{\beta}X_t))^2 = \sum\left(Y_t^2 - 2Y_t(\widehat{\alpha} + \widehat{\beta}X_t) + (\widehat{\alpha} + \widehat{\beta}X_t)^2\right) \\ &= \sum Y_t^2 - 2\widehat{\alpha}\sum Y_t - 2\widehat{\beta}\sum Y_t X_t + n\widehat{\alpha}^2 + 2\widehat{\alpha}\widehat{\beta}\sum X_t + \widehat{\beta}^2\sum X_t^2 \\ &= 126.\end{aligned}$$

---

[1] Здесь и далее ссылки на формулы, разделы и страницы книги *Я. Р. Магнуса, П. К. Катышева, А. А. Пересецкого* «Эконометрика. Начальный курс» (7-е изд. М.: Дело, 2005).

По формуле (2.15) оценка дисперсии ошибок равна:

$$\widehat{\sigma}^2 = s^2 = \frac{1}{n-2}\sum e_t^2 = \frac{126}{14} = 9.$$

2) Проверка гипотезы $\mathrm{H}_0$: $\beta = 1$. Оценка дисперсии $\widehat{\beta}$ (см. (2.16)) равна:

$$s_{\widehat{\beta}}^2 = \frac{s^2}{\sum x_t^2} = \frac{s^2}{\sum(X_t - \bar{X})^2} = \frac{9}{657 - 96^2/16} = \frac{1}{9}.$$

Для проверки гипотезы $\mathrm{H}_0$ вычислим статистику (2.23)

$$t = \frac{\widehat{\beta} - \beta_0}{s_{\widehat{\beta}}} = \frac{4/3 - 1}{1/3} = 1.$$

2.5%-ная точка $t$-распределения $t_{0.025}(14)$ равна 2.145, следовательно, $|t| = 1 < t_c = 2.145$, поэтому гипотеза $\mathrm{H}_0$: $\beta = 1$ не отвергается на 5%-ном уровне значимости.

## Задача 2.2

Покажите, что $\widehat{\beta} = r_{XY}\dfrac{s_Y}{s_X}$, где $r_{XY}$ — выборочный коэффициент корреляции между $X$ и $Y$, а $s_X$, $s_Y$ — стандартные отклонения $X$ и $Y$ соответственно.

## Решение

Из определения выборочного коэффициента корреляции (приложение МС, п. 6) и из равенства (2.4а) получаем:

$$r_{XY}\frac{s_Y}{s_X} = \frac{c_{XY}}{s_X s_Y}\frac{s_Y}{s_X} = \frac{c_{XY}}{s_X^2} = \frac{(\sum x_t y_t)/(n-1)}{(\sum x_t^2)/(n-1)} = \frac{\mathrm{Cov}(X,Y)}{\mathrm{Var}(X)} = \widehat{\beta}.$$

## Задача 2.3

Пусть $\widehat{\beta}$ — оценка коэффициента наклона в регрессии $Y$ на $X$, а $\widehat{\gamma}$ — оценка коэффициента наклона в регрессии $X$ на $Y$. Покажите, что $\widehat{\beta} = 1/\widehat{\gamma}$ тогда и только тогда, когда $R^2 = 1$.

## Решение

Задача решается проще при переходе к уравнениям в отклонениях. Коэффициенты $\widehat{\beta}$ и $\widehat{\gamma}$ (см. (2.6)) записываются следующим образом: $\widehat{\beta} = \dfrac{\sum x_t y_t}{\sum x_t^2}$

и $\widehat{\gamma} = \dfrac{\sum x_t y_t}{\sum y_t^2}$. Получаем $\widehat{\beta}\widehat{\gamma} = \dfrac{(\sum x_t y_t)^2}{\sum x_t^2 \sum y_t^2}$. Так как $\widehat{y}_t = \widehat{Y}_t - \overline{Y} = \widehat{\alpha} + \widehat{\beta} X_t - \widehat{\alpha} - \widehat{\beta}\overline{X} = \widehat{\beta} x_t$, то из (2.31) получаем:

$$R^2 = \frac{\widehat{\boldsymbol{y}}_*' \widehat{\boldsymbol{y}}_*}{\boldsymbol{y}_*' \boldsymbol{y}_*} = \frac{\widehat{\beta}^2 \sum x_t^2}{\sum y_t^2} = \left(\frac{\sum x_t y_t}{\sum x_t^2}\right)^2 \frac{\sum x_t^2}{\sum y_t^2} = \frac{\left(\sum x_t y_t\right)^2}{\sum x_t^2 \sum y_t^2},$$

откуда $R^2 = \widehat{\beta}\widehat{\gamma}$. Таким образом, условие $R^2 = 1$ эквивалентно условию $\widehat{\beta} = 1/\widehat{\gamma}$.

Геометрический смысл доказанного утверждения заключается в том, что произведение угловых коэффициентов (тангенсов углов наклона соответствующих регрессионных прямых) равно 1 тогда и только тогда, когда можно осуществить точную подгонку, т. е. когда вектор $\boldsymbol{y}$ лежит в плоскости, образованной единичным вектором и вектором наблюдений $\boldsymbol{x}$. Другими словами, все точки наблюдений на плоскости $(X, Y)$ лежат на прямой линии.

## Задача 2.4

Рассмотрим модель $Y_t = \alpha + \beta X_t^\gamma + \varepsilon_t$, где ошибки являются независимыми одинаково распределенными нормальными случайными величинами. Почему для оценивания параметров нельзя применять метод наименьших квадратов? Выведите уравнение для оценок максимального правдоподобия.

## Решение

При попытке непосредственно применить метод наименьших квадратов мы получаем следующую задачу:

$$F(\alpha, \beta, \gamma) = \sum_{t=1}^{n} \left(Y_t - (\alpha + \beta X_t^\gamma)\right)^2 \to \min.$$

Заметим, что задача не является квадратичной по параметрам $\alpha, \beta, \gamma$. А значит, условия первого порядка не являются линейными. Следовательно, метод наименьших квадратов применять нельзя. Выпишем функцию правдоподобия:

$$\begin{aligned} L(Y_1, \ldots, Y_n, \alpha, \beta, \gamma, \sigma^2) &= p(Y_1, \ldots, Y_n \mid X_1, \ldots, X_n, \alpha, \beta, \gamma, \sigma^2) \\ &= \prod_{t=1}^{n} p(Y_t) = \frac{1}{(2\pi)^{n/2} (\sigma^2)^{n/2}} \exp\Bigl(-\frac{1}{2\sigma^2} \sum_{t=1}^{n} (Y_t - (\alpha + \beta X_t^\gamma))^2\Bigr). \end{aligned}$$

Ее логарифм равен

$$\ln L(Y_1, \ldots, Y_n, \alpha, \beta, \gamma, \sigma^2) = -\frac{n}{2}\ln(2\pi) - \frac{n}{2}\ln(\sigma^2) - \frac{1}{2\sigma^2} \sum_{t=1}^{n} (Y_t - \alpha - \beta X_t^\gamma)^2.$$

Необходимые условия экстремума имеют вид:

$$\frac{\partial \ln L}{\partial \alpha} = \frac{1}{\sigma^2}\sum_{t=1}^{n}(Y_t - \alpha - \beta X_t^{\gamma}) = 0,$$

$$\frac{\partial \ln L}{\partial \beta} = \frac{1}{\sigma^2}\sum_{t=1}^{n}X_t^{\gamma}(Y_t - \alpha - \beta X_t^{\gamma}) = 0,$$

$$\frac{\partial \ln L}{\partial \gamma} = \frac{1}{\sigma^2}\sum_{t=1}^{n}\beta X_t^{\gamma}\ln X_t(Y_t - \alpha - \beta X_t^{\gamma}) = 0,$$

$$\frac{\partial \ln L}{\partial \sigma^2} = -\frac{n}{2}\frac{1}{\sigma^2} + \frac{1}{2\sigma^4}\sum_{t=1}^{n}(Y_t - \alpha - \beta X_t^{\gamma})^2 = 0.$$

Это и есть система уравнений для оценок максимального правдоподобия $\widehat{\alpha}_{\mathrm{ML}}$, $\widehat{\beta}_{\mathrm{ML}}$, $\widehat{\gamma}_{\mathrm{ML}}$ и $\widehat{\sigma}^2_{\mathrm{ML}}$.

## Задача 2.5

Могут ли следующие уравнения быть преобразованы в уравнения, линейные по параметрам?

а) $Y_i = \alpha \cdot \exp(\beta X_i) \cdot \varepsilon_i$,

б) $Y_i = \alpha \cdot \exp(-\beta X_i) + \varepsilon_i$,

в) $Y_i = \exp(\alpha + \beta X_i + \varepsilon_i)$,

г) $Y_i = \alpha \,/\, (\beta - X_i) + \varepsilon_i$.

## Решение

б), г) Эти уравнения не могут быть преобразованы в уравнения, линейные по параметрам.

а) Да, может. Прологарифмировав обе части уравнения а), получим

$$\ln(Y_i) = \ln(\alpha) + \beta X_i + \ln(\varepsilon_i).$$

Введем следующие обозначения: $Z_i = \ln(Y_i)$, $\gamma = \ln(\alpha)$, $u_i = \ln(\varepsilon_i)$.

Полученное уравнение $Z_i = \gamma + \beta X_i + u_i$ является линейным по параметрам. (Конечно, здесь предполагается, что случайная величина $\varepsilon_i$ принимает положительные значения.)

в) Как и в пункте а), уравнение можно преобразовать, взяв логарифм левой и правой частей:

$$\ln(Y_i) = \alpha + \beta X_i + \varepsilon_i.$$

С помощью замены $Z_i = \ln(Y_i)$ уравнение может быть приведено к линейному по параметрам: $Z_i = \alpha + \beta X_i + \varepsilon_i$.

## Задача 2.6

Зависимая переменная в регрессии $Y_t = \alpha + \beta X_t + \varepsilon_t$ разбивается на две компоненты: $Y_t = Y_{1t} + Y_{2t}$. Рассмотрим две регрессии для компонент: $Y_{1t} = \alpha_1 + \beta_1 X_t + \varepsilon_{1t}$ и $Y_{2t} = \alpha_2 + \beta_2 X_t + \varepsilon_{2t}$. Докажите следующие соотношения для МНК-оценок параметров трех регрессий: $\widehat{\alpha} = \widehat{\alpha}_1 + \widehat{\alpha}_2$, $\widehat{\beta} = \widehat{\beta}_1 + \widehat{\beta}_2$.

## Решение

Утверждение следует из того, что оценки (2.4а), (2.4б) для коэффициентов регрессии линейны по $Y$:

$$\widehat{\beta} = \frac{n\sum X_t(Y_{1t}+Y_{2t}) - \left(\sum X_t\right)\left(\sum (Y_{1t}+Y_{2t})\right)}{n\sum X_t^2 - \left(\sum X_t\right)^2}$$

$$= \frac{n\sum X_tY_{1t} - \left(\sum X_t\right)\left(\sum Y_{1t}\right)}{n\sum X_t^2 - \left(\sum X_t\right)^2} + \frac{n\sum X_tY_{2t} - \left(\sum X_t\right)\left(\sum Y_{2t}\right)}{n\sum X_t^2 - \left(\sum X_t\right)^2} = \widehat{\beta}_1 + \widehat{\beta}_2.$$

$$\widehat{\alpha} = \frac{1}{n}\sum (Y_{1t}+Y_{2t}) - \frac{1}{n}\sum X_t\,(\widehat{\beta}_1 + \widehat{\beta}_2)$$

$$= \frac{1}{n}\sum Y_{1t} - \frac{1}{n}\sum X_t\,\widehat{\beta}_1 + \frac{1}{n}\sum Y_{2t} - \frac{1}{n}\sum X_t\,\widehat{\beta}_2 = \widehat{\alpha}_1 + \widehat{\alpha}_2.$$

## Задача 2.7

Уравнение $Y_i = \alpha + \beta X_i + \varepsilon_i$ оценивается методом наименьших квадратов. Остатки регрессии равны $e_i$, $y_i = Y_i - \overline{Y}$, $x_i = X_i - \overline{X}$, $\widehat{y}_i = \widehat{Y}_i - \overline{Y}$ — отклонения от средних. Докажите, что следующие меры качества подгонки совпадают:

а) $(\sum x_iy_i)^2 / (\sum x_i^2 \sum y_i^2)$, б) $\widehat{\beta}(\sum x_iy_i)/(\sum y_i^2)$,

в) $(\sum \widehat{y}_iy_i)^2 / (\sum \widehat{y}_i^2 \sum y_i^2)$, г) $1 - \sum e_i^2/\sum y_i^2$.

## Решение

Докажем равенство а) = б): $\dfrac{(\sum x_iy_i)^2}{\sum x_i^2 \sum y_i^2} = \dfrac{\widehat{\beta}\sum x_iy_i}{\sum y_i^2}$. Так как $\sum y_i^2 > 0$, то это равенство эквивалентно тому, что $\widehat{\beta} = \dfrac{(\sum x_iy_i)}{\sum x_i^2}$, что и является МНК-оценкой параметра $\beta$ в отклонениях.

Равенство в) = а) доказываем, подставляя вместо $\widehat{y}_i^2$ его выражение $\widehat{y}_i^2 = \widehat{\beta}^2 x_i^2$ (см. решение задачи 2.3):

$$\frac{(\sum \widehat{y}_iy_i)^2}{\sum \widehat{y}_i^2 \sum y_i^2} = \frac{\widehat{\beta}^2(\sum x_iy_i)^2}{\widehat{\beta}^2\sum x_i^2 \sum y_i^2} = \frac{(\sum x_iy_i)^2}{\sum x_i^2 \sum y_i^2}.$$

Докажем, наконец, равенство г) = а). Из определения следует, что $e_i = y_i - \widehat{\beta}x_i$, поэтому

$$1 - \frac{\sum e_i^2}{\sum y_i^2} = \frac{\sum y_i^2 - \sum (y_i - \widehat{\beta}x_i)^2}{\sum y_i^2} = \frac{2\widehat{\beta}\sum x_i y_i - \widehat{\beta}^2 \sum x_i^2}{\sum y_i^2}$$

$$= 2\frac{(\sum x_i y_i)^2}{\sum x_i^2 \sum y_i^2} - \frac{(\sum x_i y_i)^2}{\sum x_i^2 \sum y_i^2} = \frac{(\sum x_i y_i)^2}{\sum x_i^2 \sum y_i^2}.$$

## Задача 2.8

Выведите непосредственно формулу для оценки коэффициента наклона в регрессии без свободного члена, т. е. найдите оценку параметра $\beta$ в регрессии $Y_t = \beta X_t + \varepsilon_t$ минимизацией суммы квадратов отклонений $\sum (Y_t - \widehat{Y}_t)^2$.

## Решение

Задача состоит в отыскании минимума функции

$$F(\beta) = \sum (Y_t - \beta X_t)^2.$$

Запишем необходимое условие экстремума (аналогично (2.2)):

$$\frac{\partial}{\partial \beta} \sum (Y_t - \beta X_t)^2 = 0 \quad \text{или} \quad \sum (Y_t - \beta X_t) X_t = 0.$$

Решая это уравнение относительно $\beta$, получаем МНК-оценку

$$\widehat{\beta} = \frac{\sum X_t Y_t}{\sum X_t^2}.$$

## Задача 2.9

Для наблюдений

| $t$ | $Y_t$ | $X_t$ | $t$ | $Y_t$ | $X_t$ |
|---|---|---|---|---|---|
| 1 | 70 | 5 | 6 | 35 | 22 |
| 2 | 65 | 11 | 7 | 40 | 25 |
| 3 | 55 | 15 | 8 | 30 | 27 |
| 4 | 60 | 17 | 9 | 25 | 30 |
| 5 | 50 | 20 | 10 | 32 | 35 |

вычислите следующие величины:

а) коэффициент детерминации $R^2$ в регрессии $Y_t$ на $X_t$ при наличии свободного члена;

б) коэффициент детерминации $R^2$ в регрессии $Y_t$ на $X_t$ при отсутствии свободного члена;

в) коэффициент детерминации $R^2$ в регрессии $y_t$ на $x_t$ при наличии свободного члена, где $y_t$ и $x_t$ — отклонения переменных $Y_t$ и $X_t$ от их средних значений;

г) коэффициент детерминации $R^2$ в регрессии $y_t$ на $x_t$ при отсутствии свободного члена.

## Решение

а) Уравнение регрессии $Y$ на $X$ при наличии свободного члена имеет вид $Y_t = \alpha + \beta X_t + \varepsilon_t$. Оценки $\widehat{\beta}$ и $\widehat{\alpha}$ вычисляем по формулам (2.4а) и (2.4б) соответственно:

$$\widehat{\beta} = \frac{n\sum X_tY_t - \left(\sum X_t\right)\left(\sum Y_t\right)}{n\sum X_t^2 - \left(\sum X_t\right)^2} = \frac{10\cdot 8360 - 207\cdot 462}{10\cdot 5023 - 207^2} \approx -1.63,$$

$$\widehat{\alpha} = \overline{Y} - \widehat{\beta}\overline{X} = 46.2 - (-1.63)\cdot 20.7 = 79.95.$$

Итак, $\widehat{Y}_t = 79.95 - 1.63X_t$. Теперь у нас есть все данные, чтобы вычислить $R^2$ по формуле

$$R^2 = \frac{\mathrm{RSS}}{\mathrm{TSS}} = \frac{\sum(\widehat{Y}_t - \overline{Y})^2}{\sum(Y_t - \overline{Y})^2} = \frac{1962.0}{2279.6} = 0.8607.$$

б) Уравнение регрессии $Y$ на $X$ при отсутствии свободного члена записывается в виде $Y_t = \beta X_t + \varepsilon_t$. Оценку $\widehat{\beta}$ можно вычислить по формуле (см. задача 2.8):

$$\widehat{\beta} = \frac{\sum Y_tX_t}{\sum X_t^2} = \frac{8360}{5023} \approx 1.66.$$

В случае отсутствия константы равенство (2.26)

$$\mathrm{TSS} = \mathrm{ESS} + \mathrm{RSS},$$

вообще говоря, не выполняется и корректно определить коэффициент детерминации (2.27) невозможно. Вычислим $R^2$ в этом случае двумя возможными способами:

$$R^2_{(1)} = \frac{\mathrm{RSS}}{\mathrm{TSS}} = \frac{3424.7}{2279.6} = 1.50,$$

$$R^2_{(2)} = 1 - \frac{\mathrm{ESS}}{\mathrm{TSS}} = 1 - \frac{9710.1}{2279.6} = -3.26.$$

Отметим, что в отсутствие константы значения $R^2$ не обязаны лежать в интервале между 0 и 1, что у нас и получилось.

в), г) Уравнение регрессии $y$ на $x$ сводится к рассмотренному в пункте а), так как взаимное расположение точек и регрессионной прямой геометрически совпадает во всех этих трех случаях. Таким образом, $R^2 = 0.8607$. В случае в) регрессионная прямая проходит через начало координат и, следовательно, случаи г) и в) совпадают.

## Задача 2.10

Предположим, что модель

$$Y_t = \alpha + \beta X_t + \varepsilon_t, \quad t = 1, \ldots, n$$

удовлетворяет условиям классической регрессии. Рассматривается следующая оценка коэффициента $\beta$:

$$\widetilde{\beta} = \frac{1}{n} \sum_{t=1}^{n} \frac{Y_t - \overline{Y}}{X_t - \overline{X}}.$$

а) Является ли оценка $\widetilde{\beta}$ несмещенной? Является ли она линейной?

б) Вычислите дисперсию оценки $\widetilde{\beta}$.

в) Проверьте теорему Гаусса–Маркова, сравнив полученную дисперсию оценки $\widetilde{\beta}$ с дисперсией МНК-оценки $\sigma^2 / \sum_{t=1}^{n}(X_t - \overline{X})^2$.

## Решение

а) Вычислим математическое ожидание оценки $\widetilde{\beta}$:

$$\begin{aligned}
\mathrm{E}(\widetilde{\beta}) &= \frac{1}{n} \sum_{t=1}^{n} \frac{\mathrm{E}(Y_t) - \mathrm{E}(\overline{Y})}{X_t - \overline{X}} = \frac{1}{n} \sum_{t=1}^{n} \frac{\mathrm{E}(\alpha + \beta X_t + \varepsilon_t) - \mathrm{E}(\alpha + \beta \overline{X} + \overline{\varepsilon})}{X_t - \overline{X}} \\
&= \frac{1}{n} \sum_{t=1}^{n} \frac{(\alpha + \beta X_t) - (\alpha + \beta \overline{X})}{X_t - \overline{X}} = \frac{1}{n} \sum_{t=1}^{n} \frac{\beta (X_t - \overline{X})}{X_t - \overline{X}} = \beta.
\end{aligned}$$

Следовательно, оценка $\widetilde{\beta}$ является несмещенной.

Представим $\widetilde{\beta}$ в следующем виде:

$$\begin{aligned}
\widetilde{\beta} &= \frac{1}{n} \sum_{t=1}^{n} \frac{Y_t - \overline{Y}}{X_t - \overline{X}} = \frac{1}{n} \sum_{t=1}^{n} \frac{Y_t}{x_t} - \frac{1}{n} \sum_{s=1}^{n} \frac{\overline{Y}}{x_s} \\
&= \frac{1}{n} \sum_{t=1}^{n} \frac{1}{x_t} Y_t - \frac{1}{n^2} \left( \sum_{t=1}^{n} Y_t \right) \left( \sum_{s=1}^{n} \frac{1}{x_s} \right) \\
&= \sum_{t=1}^{n} \frac{1}{n} \left( \frac{1}{x_t} - \frac{1}{n} \sum_{s=1}^{n} \frac{1}{x_s} \right) Y_t = \sum_{t=1}^{n} c_t Y_t.
\end{aligned}$$

Так как переменные $x_t$ являются неслучайными, то оценка $\widetilde{\beta}$ является линейной.

б) Вычислим дисперсию оценки $\widetilde{\beta}$. При этом воспользуемся тем, что $Y_t$ — некоррелированные случайные величины с дисперсиями $\mathrm{V}(Y_t) = \sigma^2$.

$$\mathrm{V}(\widetilde{\beta}) = \mathrm{V}\left(\sum_{t=1}^{n} c_t Y_t\right) = \sum_{t=1}^{n} c_t^2 \mathrm{V}(Y_t) = \sigma^2 \sum_{t=1}^{n} c_t^2 = \sigma^2 \sum_{t=1}^{n} \frac{1}{n^2}\left(\frac{1}{x_t} - \frac{1}{n}\sum_{s=1}^{n}\frac{1}{x_s}\right)^2.$$

в) Для непосредственной проверки теоремы Гаусса–Маркова воспользуемся неравенством Коши–Буняковского.

$$\begin{aligned}
\frac{\mathrm{V}(\widetilde{\beta})}{\mathrm{V}(\widehat{\beta}_{\mathrm{OLS}})} &= \left(\sigma^2 \frac{1}{n^2}\sum_{t=1}^{n}\left(\frac{1}{x_t} - \frac{1}{n}\sum_{s=1}^{n}\frac{1}{x_s}\right)^2\right)\left(\frac{1}{\sigma^2}\sum_{t=1}^{n} x_t^2\right) \\
&= \frac{1}{n^2}\left(\sum_{t=1}^{n}\left(\frac{1}{x_t} - \frac{1}{n}\sum_{s=1}^{n}\frac{1}{x_s}\right)^2\right)\left(\sum_{t=1}^{n} x_t^2\right) \\
&\geqslant \frac{1}{n^2}\left(\sum_{t=1}^{n}\left(\frac{1}{x_t} - \frac{1}{n}\sum_{s=1}^{n}\frac{1}{x_s}\right)x_t\right)^2 = \frac{1}{n^2}\left(\sum_{t=1}^{n}\left(1 - \frac{1}{n}\sum_{s=1}^{n}\frac{x_t}{x_s}\right)\right)^2 \\
&= \frac{1}{n^2}\left(n - \frac{1}{n}\sum_{s=1}^{n}\left(\frac{1}{x_s}\sum_{t=1}^{n} x_t\right)\right)^2 = \frac{1}{n^2}n^2 = 1.
\end{aligned}$$

Здесь мы также использовали тождество $\sum_{t=1}^{n} x_t = 0$. Таким образом, мы получили, что $\mathrm{V}(\widetilde{\beta}) \geqslant \mathrm{V}(\widehat{\beta}_{\mathrm{OLS}})$.

## Задача 2.11

Приведите пример набора данных $(X_t, Y_t)$, для которого решение задачи поиска параметров $\alpha$, $\beta$, минимизирующих функционал

$$F = \sum_{t=1}^{n} |Y_t - (\alpha + \beta X_t)|,$$

не единственно.

## Решение

Пусть $n = 3$ и имеется набор данных: $(X_1, Y_1) = (0, 0)$, $(X_2, Y_2) = (0, 2)$, $(X_3, Y_3) = (\sqrt{3}, 1)$ (равносторонний треугольник со стороной 2). Тогда функционал равен

$$F = \sum_{t=1}^{3} |Y_t - (\alpha + \beta X_t)| = |\alpha| + |2 - \alpha| + |1 - \alpha - \sqrt{3}\beta|,$$

и нетрудно проверить, что $F_{\min} = 2$, и этот минимум достигается при любых значениях параметров $\alpha$ и $\beta$, удовлетворяющих соотношению $\alpha + \sqrt{3}\beta = 1$, $\alpha \in [0, 2]$ (т.е. когда прямая $Y = \alpha + \beta X$ проходит через точку $(\sqrt{3}, 1)$ и пересекает ось $Y$ между точками 1 и 2), в частности, например, при $\alpha = 0$, $\beta = 1/\sqrt{3}$ (прямая, проходящая через первую и третью вершины треугольника) и при $\alpha = 2$, $\beta = -1/\sqrt{3}$ (прямая, проходящая через вторую и третью вершины). То есть задача поиска параметров $\alpha$ и $\beta$, минимизирующих функционал $F$, имеет бесконечно много решений.

## Задача 2.12

Рассмотрим модель регрессии на константу

$$Y_t = \alpha + \varepsilon_t, \quad t = 1, \ldots, n.$$

а) Найдите оценки метода наименьших квадратов для $\alpha$ и $\sigma^2$.

б) Найдите дисперсию оценки $\widehat{\alpha}$.

в) Покажите, что статистика $\dfrac{\widehat{\alpha} - \alpha}{s_{\widehat{\alpha}}}$ имеет распределение $t(n-1)$.

г) Чему равен коэффициент детерминации $R^2$?

## Решение

а) Задача состоит в отыскании минимума суммы квадратов остатков (см. (2.1))

$$F(\alpha) = \sum (Y_t - \alpha)^2.$$

Запишем необходимое условие экстремума (аналогично (2.2)):

$$\frac{\partial}{\partial \alpha} \sum (Y_t - \alpha)^2 = 0 \qquad \text{или} \qquad \sum (Y_t - \alpha) = 0.$$

Решая это уравнение относительно $\alpha$, получаем МНК-оценку

$$\widehat{\alpha} = \frac{1}{n} \sum Y_t = \overline{Y}.$$

Для того чтобы найти оценку дисперсии ошибок $\sigma^2$, рассмотрим сумму квадратов остатков

$$\sum e_t^2 = \sum (Y_t - \widehat{\alpha})^2 = \sum (Y_t - \overline{Y})^2.$$

Из курса статистики известно, что

$$\mathrm{E}\left(\sum e_t^2\right) = (n-1)\sigma^2.$$

В самом деле,

$$\begin{aligned}\mathrm{E}\left(\sum(Y_t-\overline{Y})^2\right)&=\mathrm{E}\left(\sum(\varepsilon_t-\overline{\varepsilon})^2\right)=\mathrm{E}\left(\sum(\varepsilon_t^2+\overline{\varepsilon}^2-2\varepsilon_t\overline{\varepsilon})\right)\\&=\mathrm{E}\left(\sum\varepsilon_t^2+n\overline{\varepsilon}^2-2\sum\varepsilon_t\overline{\varepsilon}\right)=\mathrm{E}\left(\sum\varepsilon_t^2+n\overline{\varepsilon}^2-2n\overline{\varepsilon}^2\right)\\&=\mathrm{E}\left(\sum\varepsilon_t^2-n\overline{\varepsilon}^2\right)=n\sigma^2-n\frac{\sigma^2}{n}=(n-1)\sigma^2.\end{aligned}$$

Отсюда получаем, что

$$\widehat{\sigma}^2=s^2=\frac{1}{(n-1)}\sum e_t^2=\frac{1}{(n-1)}\sum(Y_t-\overline{Y})^2$$

является несмещенной оценкой $\sigma^2$.

б) $\mathrm{V}(\widehat{\alpha})=\mathrm{V}\left(\frac{1}{n}\sum Y_t\right)=\frac{1}{n^2}n\mathrm{V}(Y_t)=\frac{1}{n}\sigma^2.$

в) Пусть выполняется дополнительное условие нормальности ошибок $\varepsilon_t\sim N(0,\sigma^2)$. Оценка $\widehat{\alpha}$ тогда также распределена нормально: $\widehat{\alpha}\sim N(0,\sigma^2/n)$. Из приложения МС, п. 4, N12 следует, что

$$\frac{(\overline{Y}-\alpha)\sqrt{n}}{\sqrt{\frac{1}{n-1}\sum(Y_t-\overline{Y})^2}}\sim t(n-1).$$

Поскольку $V(\widehat{\alpha})=s^2_{\widehat{\alpha}}=\dfrac{s^2}{n}$, получаем

$$\frac{(\overline{Y}-\alpha)\sqrt{n}}{\sqrt{\frac{1}{n-1}\sum(Y_t-\overline{Y})^2}}=\frac{(\overline{Y}-\alpha)}{s/\sqrt{n}}=\frac{\widehat{\alpha}-\alpha}{s_{\widehat{\alpha}}}\sim t(n-1).$$

г) Вычислим значение $R^2$ по формуле (2.27)

$$R^2=1-\frac{\mathrm{ESS}}{\mathrm{TSS}}=1-\frac{\sum(Y_t-\widehat{\alpha})^2}{\sum(Y_t-\overline{Y})^2}=1-\frac{\sum(Y_t-\overline{Y})^2}{\sum(Y_t-\overline{Y})^2}=1-1=0.$$

## Задача 2.13

Рассмотрим модель регрессии без константы

$$Y_t=\beta X_t+\varepsilon_t,\quad t=1,\ldots,n.$$

а) Найдите оценки метода наименьших квадратов для $\beta$ и $\sigma^2$.

б) Найдите дисперсию оценки $\widehat{\beta}$.

в) Покажите, что статистика $\dfrac{\widehat{\beta}-\beta}{s_{\widehat{\beta}}}$ имеет распределение $t(n-1)$.

г) Приведите примеры данных, для которых: значение коэффициента $R^2$, рассчитанное по формуле $R^2 = \mathrm{RSS}/\mathrm{TSS}$, отличается от значения $R^2$, рассчитанного по формуле $R^2 = 1 - \mathrm{ESS}/\mathrm{TSS}$; значение коэффициента $R^2$, рассчитанное по формуле $R^2 = \mathrm{RSS}/\mathrm{TSS}$, больше 1; значение $R^2$, рассчитанное по формуле $R^2 = 1 - \mathrm{ESS}/\mathrm{TSS}$, меньше 0.

## Решение

а) В задаче 2.8 мы получили формулу для МНК-оценки в регрессии без константы:

$$\widehat{\beta} = \frac{\sum X_t Y_t}{\sum X_t^2}.$$

Остатки регрессии равны

$$e_t = Y_t - \widehat{\beta} X_t = Y_t - X_t \frac{\sum X_s Y_s}{\sum X_s^2}.$$

Запишем это в векторной форме:

$$\boldsymbol{e} = \left(\boldsymbol{y} - \boldsymbol{x}\frac{\boldsymbol{x}'\boldsymbol{y}}{\boldsymbol{x}'\boldsymbol{x}}\right) = \left(\boldsymbol{I} - \frac{\boldsymbol{x}\boldsymbol{x}'}{\boldsymbol{x}'\boldsymbol{x}}\right)\boldsymbol{y} = \boldsymbol{A}\boldsymbol{y}.$$

Нетрудно проверить, что матрица $\boldsymbol{A}$ идемпотентная:

$$\boldsymbol{A}^2 = \left(\boldsymbol{I} - \frac{\boldsymbol{x}\boldsymbol{x}'}{\boldsymbol{x}'\boldsymbol{x}}\right)^2 = \boldsymbol{I} - 2\frac{\boldsymbol{x}\boldsymbol{x}'}{\boldsymbol{x}'\boldsymbol{x}} + \frac{\boldsymbol{x}\boldsymbol{x}'\boldsymbol{x}\boldsymbol{x}'}{(\boldsymbol{x}'\boldsymbol{x})^2} = \boldsymbol{I} - \frac{\boldsymbol{x}\boldsymbol{x}'}{\boldsymbol{x}'\boldsymbol{x}} = \boldsymbol{A}.$$

Очевидно также, что $\boldsymbol{A}' = \boldsymbol{A}$. Поскольку выполняется равенство

$$\boldsymbol{A}\boldsymbol{x} = \left(\boldsymbol{I} - \frac{\boldsymbol{x}\boldsymbol{x}'}{\boldsymbol{x}'\boldsymbol{x}}\right)\boldsymbol{x} = \boldsymbol{x} - \frac{\boldsymbol{x}\boldsymbol{x}'\boldsymbol{x}}{\boldsymbol{x}'\boldsymbol{x}} = \boldsymbol{x} - \boldsymbol{x} = \boldsymbol{0},$$

получаем

$$\boldsymbol{e} = \boldsymbol{A}\boldsymbol{y} = \boldsymbol{A}(\beta\boldsymbol{x} + \boldsymbol{\varepsilon}) = \beta\boldsymbol{A}\boldsymbol{x} + \boldsymbol{A}\boldsymbol{\varepsilon} = \boldsymbol{A}\boldsymbol{\varepsilon}.$$

Вычислим математическое ожидание суммы квадратов остатков регрессии:

$$\begin{aligned}
\mathrm{E}\left(\sum e_t^2\right) &= \mathrm{E}(\boldsymbol{e}'\boldsymbol{e}) = \mathrm{E}(\mathrm{tr}(\boldsymbol{e}'\boldsymbol{e})) = \mathrm{E}(\mathrm{tr}(\boldsymbol{e}\boldsymbol{e}')) = \mathrm{E}(\mathrm{tr}((\boldsymbol{A}\boldsymbol{\varepsilon})(\boldsymbol{A}\boldsymbol{\varepsilon})')) \\
&= \mathrm{E}(\mathrm{tr}(\boldsymbol{A}\boldsymbol{\varepsilon}\boldsymbol{\varepsilon}'\boldsymbol{A}')) = \mathrm{tr}(\boldsymbol{A}\mathrm{E}(\boldsymbol{\varepsilon}\boldsymbol{\varepsilon}')\boldsymbol{A}') = \mathrm{tr}(\boldsymbol{A}\mathrm{V}(\boldsymbol{\varepsilon})\boldsymbol{A}') \\
&= \sigma^2\,\mathrm{tr}(\boldsymbol{A}\boldsymbol{A}') = \sigma^2\,\mathrm{tr}(\boldsymbol{A}^2) = \sigma^2\,\mathrm{tr}(\boldsymbol{A}) \\
&= \sigma^2\,\mathrm{tr}\left(\boldsymbol{I}_n - \frac{\boldsymbol{x}\boldsymbol{x}'}{\boldsymbol{x}'\boldsymbol{x}}\right) = \sigma^2\left(\mathrm{tr}\,\boldsymbol{I}_n - \frac{1}{\boldsymbol{x}'\boldsymbol{x}}\mathrm{tr}(\boldsymbol{x}\boldsymbol{x}')\right) \\
&= \sigma^2\left(n - \frac{1}{\boldsymbol{x}'\boldsymbol{x}}\mathrm{tr}(\boldsymbol{x}'\boldsymbol{x})\right) = \sigma^2(n-1).
\end{aligned}$$

Отсюда получаем, что

$$\widehat{\sigma}^2 = s^2 = \frac{1}{n-1}\sum e_t^2$$

есть несмещенная оценка $\sigma^2$.

б) Дисперсия оценки $\widehat{\beta}$ равна:

$$\mathrm{V}(\widehat{\beta}) = \mathrm{V}\left(\frac{\boldsymbol{x}'\boldsymbol{y}}{\boldsymbol{x}'\boldsymbol{x}}\right) = \frac{1}{(\boldsymbol{x}'\boldsymbol{x})^2}\mathrm{V}(\boldsymbol{x}'\boldsymbol{y}) = \frac{1}{(\boldsymbol{x}'\boldsymbol{x})^2}\boldsymbol{x}'\mathrm{V}(\boldsymbol{y})\boldsymbol{x}$$
$$= \sigma^2\frac{\boldsymbol{x}'\boldsymbol{x}}{(\boldsymbol{x}'\boldsymbol{x})^2} = \frac{\sigma^2}{\boldsymbol{x}'\boldsymbol{x}} = \frac{\sigma^2}{\sum X_t^2}.$$

в) Предположим теперь, что ошибки имеют нормальное распределение, $\varepsilon \sim N(\mathbf{0}, \sigma^2\boldsymbol{I}_n)$.

В силу идемпотентности матрицы $A$, используя утверждение N8 (приложение МС, п. 4), получаем:

$$\frac{1}{\sigma^2}\sum e_t^2 = \frac{1}{\sigma^2}\boldsymbol{e}'\boldsymbol{e} = \frac{1}{\sigma^2}\boldsymbol{\varepsilon}'\boldsymbol{A}\boldsymbol{\varepsilon} = \left(\frac{\varepsilon}{\sigma}\right)'\boldsymbol{A}\left(\frac{\varepsilon}{\sigma}\right) \sim \chi^2(\mathrm{rank}(\boldsymbol{A})).$$

Из вычислений в пункте а) мы имеем: $\mathrm{rank}\,\boldsymbol{A} = \mathrm{tr}\,\boldsymbol{A} = n-1$, поэтому

$$(n-1)\frac{s^2}{\sigma^2} = \frac{1}{\sigma^2}\sum e_t^2 \sim \chi^2(n-1).$$

Оценка $\widehat{\beta}$ имеет нормальное распределение (как линейная функция от нормального вектора):

$$\widehat{\beta} \sim N\left(\beta, \frac{\sigma^2}{\sum X_t^2}\right) \quad \text{или} \quad \widehat{\beta} - \beta \sim N\left(0, \frac{\sigma^2}{\sum X_t^2}\right).$$

Таким образом, искомая статистика имеет вид:

$$\frac{\widehat{\beta}-\beta}{s_{\widehat{\beta}}} = \frac{\widehat{\beta}-\beta}{s\Big/\sqrt{\sum X_t^2}} = \frac{(\widehat{\beta}-\beta)\Big/\left(\sigma\Big/\sqrt{\sum X_t^2}\right)}{s/\sigma} \sim \frac{N(0,1)}{\sqrt{\frac{1}{n-1}\chi^2(n-1)}}.$$

Для того чтобы доказать, что эта статистика имеет распределение Стьюдента $t(n-1)$, осталось показать независимость числителя и знаменателя. Так как знаменатель есть функция от остатков $e_t$, то достаточно доказать независимость оценки $\widehat{\beta}$ и вектора остатков $\boldsymbol{e}$. Вектор остатков распределен нормально (как линейная функция от нормального вектора), поэтому достаточно показать, что $\widehat{\beta}$ и $\boldsymbol{e}$ некоррелированы:

$$\widehat{\beta}-\beta=\frac{\boldsymbol{x}'(\beta\boldsymbol{x}+\boldsymbol{\varepsilon})}{\boldsymbol{x}'\boldsymbol{x}}=\beta+\frac{\boldsymbol{x}'\boldsymbol{\varepsilon}}{\boldsymbol{x}'\boldsymbol{x}}-\beta=\frac{\boldsymbol{x}'\boldsymbol{\varepsilon}}{\boldsymbol{x}'\boldsymbol{x}},$$

$$\operatorname{Cov}(\widehat{\beta}-\beta,\boldsymbol{e})=\operatorname{Cov}\left(\frac{\boldsymbol{x}'\boldsymbol{\varepsilon}}{\boldsymbol{x}'\boldsymbol{x}},A\boldsymbol{\varepsilon}\right)=\mathrm{E}\left(\frac{\boldsymbol{x}'\boldsymbol{\varepsilon}}{\boldsymbol{x}'\boldsymbol{x}}\boldsymbol{\varepsilon}'A\right)=\frac{1}{\boldsymbol{x}'\boldsymbol{x}}\boldsymbol{x}'\mathrm{V}(\boldsymbol{\varepsilon})\left(I-\frac{\boldsymbol{x}\boldsymbol{x}'}{\boldsymbol{x}'\boldsymbol{x}}\right)$$
$$=\frac{\sigma^2}{\boldsymbol{x}'\boldsymbol{x}}\left(\boldsymbol{x}'-\frac{\boldsymbol{x}'\boldsymbol{x}\boldsymbol{x}'}{\boldsymbol{x}'\boldsymbol{x}}\right)=\frac{\sigma^2}{\boldsymbol{x}'\boldsymbol{x}}(\boldsymbol{x}'-\boldsymbol{x}')=\boldsymbol{0}.$$

г) Рассмотрим следующие данные: $n=2$, $(X_1,Y_1)=(0,2)$, $(X_2,Y_2)=(1,1)$. Очевидно, регрессия без константы имеет вид $\widehat{Y}=X$, $\widehat{\beta}=1$, т. е. линия регрессии проходит через вторую точку и начало координат. Получаем $\mathrm{TSS}=0.5$, $\mathrm{ESS}=2^2+0=4$, $\mathrm{RSS}=(0-1.5)^2+(1-1.5)^2=2.5$ и

$$\frac{\mathrm{RSS}}{\mathrm{TSS}}=5,\qquad 1-\frac{\mathrm{ESS}}{\mathrm{TSS}}=1-8=-7.$$

## Задача 2.14

Менеджер новой чебуречной не уверен в правильности выбранной цены на чебуреки, поэтому в течение 12 недель он варьирует цену и записывает количество проданных чебуреков. Полученные данные приведены в табл. 2.1 ($t$ — номер недели, $q_t$ — количество проданных чебуреков, $p_t$ — цена одного чебурека (руб.)).

**Таблица 2.1**

| $t$ | $p_t$ | $q_t$ | $t$ | $p_t$ | $q_t$ |
|---|---|---|---|---|---|
| 1 | 12.3 | 795 | 7 | 12.8 | 714 |
| 2 | 11.5 | 915 | 8 | 9.9 | 1180 |
| 3 | 11.0 | 965 | 9 | 12.2 | 851 |
| 4 | 12.0 | 892 | 10 | 12.5 | 779 |
| 5 | 13.5 | 585 | 11 | 13.0 | 625 |
| 6 | 12.5 | 644 | 12 | 10.5 | 1001 |

а) Оцените параметры модели

$$\ln q_t=\alpha+\beta\ln p_t+\varepsilon_t,\quad t=1,\ldots,12.$$

б) Используя полученные оценки коэффициентов, найдите оптимальную в смысле максимума выручки от продаж цену чебурека.

## Решение

а) Оценка параметров модели с применением эконометрического пакета дает следующий результат.

Dependent Variable: $\ln(q)$

| Variable | Coefficient | Std. Error | $t$-Statistic | Probability |
|---|---|---|---|---|
| const | 12.027 | 0.6322 | 19.0236 | 0.0000 |
| $\ln(p)$ | −2.149 | 0.2549 | −8.4315 | 0.0000 |
| $R^2$ | 0.8767 | | | |

б) Выручка от продаж $F(p)$ равна произведению количества проданных чебуреков на цену:

$$F(p) = pe^{12.03-2.149\ln p} = e^{12.03}p^{-1.149}.$$

Из вида функции $F(p)$ видно, что она монотонно убывает при возрастании $p$. Откуда следует, что чем меньше цена, тем больше выручка. Однако следует учесть, что наши наблюдения сделаны на интервале цен [9.9, 13.5] и наши выводы недостоверны при значениях $p$, значительно выходящих за рамки этого диапазона.

Чтобы все-таки попытаться ответить на поставленный вопрос, рассмотрим линейную модель:

$$q_t = \alpha + \beta p_t + \varepsilon_t, \quad t = 1, \ldots, 12.$$

Оценка параметров модели с применением эконометрического пакета дает следующий результат.

Dependent Variable: $q$

| Variable | Coefficient | Std. Error | $t$-Statistic | Probability |
|---|---|---|---|---|
| const | 2717.70 | 174.92 | 15.537 | 0.0000 |
| $p$ | −157.734 | 14.555 | −10.837 | 0.0000 |
| $R^2$ | 0.92154 | | | |

Выручка от продаж, рассчитанная по последней модели, равна:

$$F(p) = p(2717.70 - 157.734p) = 2717.70p - 157.734p^2.$$

Оптимальная цена находится из условия $F'(p) = 0$, откуда

$$p_{\text{opt}} = \frac{2717.70}{2 \cdot 157.734} = 8.6.$$

Качественная рекомендация та же, что и в первой модели — стоит подумать о снижении цены.

## Задача 2.15

Пусть $\widehat{\sigma}^2_{\mathrm{ML}} = \sum e_t^2/n$ и $\widehat{\sigma}^2_{\mathrm{OLS}} = \sum e_t^2/(n-2)$ — оценки методов максимального правдоподобия и наименьших квадратов для дисперсии ошибок $\sigma^2$ в классической модели парной регрессии $Y_t = \beta_1 + \beta_2 X_t + \varepsilon_t$, $t = 1, \ldots, n$, $\varepsilon_t \sim N(0, \sigma^2)$.

а) Найдите дисперсию и среднеквадратичное отклонение $(\mathrm{MSE}(\widehat{\theta}) = \mathrm{E}((\widehat{\theta} - \theta)^2))$ каждой из двух оценок.

б) Какая из двух оценок обладает наименьшей дисперсией? Наименьшим среднеквадратичным отклонением?

## Решение

а) Обозначим через $\xi$ сумму остатков регрессии: $\xi = \sum e_t^2$. Известно (см. п. 2.5), что при условии нормальности остатков

$$\frac{1}{\sigma^2}\xi \sim \chi^2(n-2).$$

Поэтому (приложение МС, п. 3)

$$\mathrm{E}(\xi) = (n-2)\sigma^2 \quad \text{и} \quad \mathrm{V}(\xi) = 2(n-2)\sigma^4.$$

Отсюда получаем:

$$\widehat{\sigma}^2_{\mathrm{ML}} = \frac{1}{n}\xi,$$

$$\mathrm{E}(\widehat{\sigma}^2_{\mathrm{ML}}) = \frac{1}{n}\mathrm{E}\xi = \frac{n-2}{n}\sigma^2,$$

$$\mathrm{V}(\widehat{\sigma}^2_{\mathrm{ML}}) = \frac{1}{n^2}\mathrm{V}(\xi) = \frac{2(n-2)}{n^2}\sigma^4;$$

$$\widehat{\sigma}^2_{\mathrm{OLS}} = \frac{1}{n-2}\xi,$$

$$\mathrm{E}(\widehat{\sigma}^2_{\mathrm{OLS}}) = \frac{1}{n-2}\mathrm{E}\xi = \frac{n-2}{n-2}\sigma^2 = \sigma^2,$$

$$\mathrm{V}(\widehat{\sigma}^2_{\mathrm{OLS}}) = \frac{1}{(n-2)^2}\mathrm{V}(\xi) = \frac{2(n-2)}{(n-2)^2}\sigma^4 = \frac{2}{n-2}\sigma^4.$$

Для вычисления среднеквадратичных отклонений используем формулу

$$\mathrm{MSE}(\widehat{\theta}) = \mathrm{E}((\widehat{\theta} - \theta)^2) = \mathrm{E}\left(((\widehat{\theta} - \mathrm{E}\widehat{\theta}) + (\mathrm{E}\widehat{\theta} - \theta))^2\right) = \mathrm{V}(\widehat{\theta}) + (\mathrm{E}\widehat{\theta} - \theta)^2.$$

Среднеквадратичные отклонения равны соответственно

$$\mathrm{MSE}(\widehat{\sigma}^2_{\mathrm{ML}}) = \left(\frac{2(n-2)}{n^2} + \left(\frac{n-2}{n} - 1\right)^2\right)\sigma^4 = \frac{2}{n}\sigma^4,$$

$$\mathrm{MSE}(\widehat{\sigma}^2_{\mathrm{OLS}}) = \left(\frac{2}{n-2} - 0^2\right)\sigma^4 = \frac{2}{n-2}\sigma^4.$$

б) Из предыдущих вычислений следует, что

$$\mathrm{V}(\widehat{\sigma}^2_{\mathrm{ML}}) \leqslant \mathrm{V}(\widehat{\sigma}^2_{\mathrm{OLS}}) \quad \text{и} \quad \mathrm{MSE}(\widehat{\sigma}^2_{\mathrm{ML}}) \leqslant \mathrm{MSE}(\widehat{\sigma}^2_{\mathrm{OLS}}).$$

## Задача 2.16

Так называемая кривая Филлипса описывает связь темпа роста зарплаты и уровня безработицы. А именно,

$$\delta w_t = \beta_1 + \beta_2 \frac{1}{u_t} + \varepsilon_t,$$

где $w_t$ — уровень заработной платы, $\delta w_t = 100(w_t - w_{t-1})/w_{t-1}$ — темп роста зарплаты (в процентах) и $u_t$ — процент безработных в год $t$. Теория предполагает, что $\beta_1 < 0$ и $\beta_2 > 0$.

**Таблица 2.2**

| Год $t$ | $w_t$ | $u_t$ | Год $t$ | $w_t$ | $u_t$ |
|---|---|---|---|---|---|
| 1 | 1.62 | 1.0 | 10 | 2.66 | 1.8 |
| 2 | 1.65 | 1.4 | 11 | 2.73 | 1.9 |
| 3 | 1.79 | 1.1 | 12 | 2.80 | 1.5 |
| 4 | 1.94 | 1.5 | 13 | 2.92 | 1.4 |
| 5 | 2.03 | 1.5 | 14 | 3.02 | 1.8 |
| 6 | 2.12 | 1.2 | 15 | 3.13 | 1.1 |
| 7 | 2.26 | 1.0 | 16 | 3.28 | 1.5 |
| 8 | 2.44 | 1.1 | 17 | 3.43 | 1.3 |
| 9 | 2.57 | 1.3 | 18 | 3.58 | 1.4 |

Используя данные для некоторой страны из табл. 2.2:

а) Найдите оценки коэффициентов уравнения и проверьте наличие значимой связи между $\delta w$ и $u$.

б) Найдите «естественный уровень безработицы», т.е. такой уровень безработицы, при котором $\delta w = 0$.

в) Когда изменения в уровне безработицы оказывали наибольшее (наименьшее) влияние на темп изменения зарплаты?

г) Найдите 95%-доверительные интервалы для $\beta_1$ и $\beta_2$.

## Решение

а) Оценка параметров модели с применением эконометрического пакета дает следующий результат.

Dependent Variable: $\delta w$

| Variable | Coefficient | Std. Error | $t$-Statistic | Probability |
|---|---|---|---|---|
| const | −1.034 | 2.3716 | −0.4362 | 0.6689 |
| $1/u$ | 7.896 | 3.1626 | 2.4966 | 0.0247 |
| $R^2$ | 0.294 | | | |

Заметим, что коэффициент при $1/u$ статистически значим, а свободный член не отличается достоверно от нуля. Тем не менее уравнение показывает, что при увеличении уровня безработицы $u$ темп роста заработной платы уменьшается. Знаки обоих коэффициентов находятся в соответствии с теорией.

б) Для того чтобы найти «естественный уровень безработицы», решим уравнение

$$\delta w = -1.034 + 7.896\frac{1}{u} = 0.$$

Получаем $u_0 = 7.896/1.034 = 7.63$.

в) В силу полученного уравнения

$$\frac{\partial \delta w}{\partial u} = -7.896\frac{1}{u^2}.$$

Таким образом, изменения в уровне безработицы оказывали наибольшее (наименьшее) влияние на темп изменения зарплаты тогда, когда уровень безработицы был наименьшим (наибольшим), т. е. соответственно в 1-й или 7-й (11-й) год.

г) В силу полученных результатов

$$\beta_1 = -1.034 \pm t_{0.025}(16)\cdot 2.372 = -1.034 \pm 2.12\cdot 2.372 = -1.03 \pm 5.03,$$
$$\beta_2 = 7.896 \pm t_{0.025}(16)\cdot 3.163 = 7.896 \pm 2.12\cdot 3.163 = 7.90 \pm 6.70.$$

## Задача 2.17

В табл. 2.3 представлены расходы на агрегированное потребление $Y$ и агрегированный располагаемый доход $X$ в некоторой национальной экономике в течение 12 лет — с 1986 по 1997 г.

а) Изобразите графически зависимость $Y$ от $X$ и определите, есть ли приближенная линейная зависимость $Y$ от $X$.

б) Вычислите парную регрессию агрегированного потребления $Y$ на $X$ по данным, представленным в табл. 2.3.

в) Вычислите $s^2, s^2_{\widehat{\alpha}}, s^2_{\widehat{\beta}}$.

**Таблица 2.3**

| Год | $t$ | $Y_t$ | $X_t$ | Год | $t$ | $Y_t$ | $X_t$ |
|---|---|---|---|---|---|---|---|
| 1986 | 1 | 152 | 170 | 1992 | 7 | 177 | 200 |
| 1987 | 2 | 159 | 179 | 1993 | 8 | 179 | 207 |
| 1988 | 3 | 162 | 187 | 1994 | 9 | 184 | 215 |
| 1989 | 4 | 165 | 189 | 1995 | 10 | 186 | 216 |
| 1990 | 5 | 170 | 193 | 1996 | 11 | 190 | 220 |
| 1991 | 6 | 172 | 199 | 1997 | 12 | 191 | 225 |

## Решение

а) Данные представлены на графике на рис. 2.1.

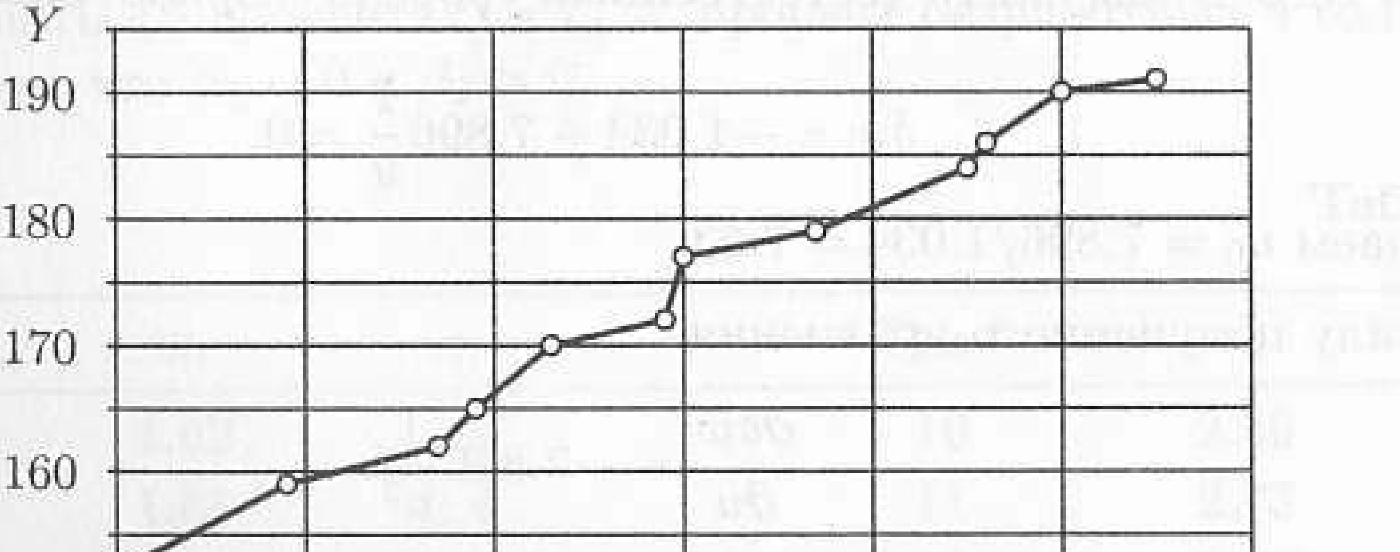



**Рис. 2.1**

Из графика видно, что данные можно приблизить линейной функцией.

б) Оценка параметров модели с применением эконометрического пакета дает следующий результат (см. таблицу 2.4).

в) Из результатов компьютерного расчета, приведенного в п. б), получаем:

$$s^2 = 1.534^2 = 2.354,$$

$$s^2_{\widehat{\alpha}} = 5.413^2 = 29.29,$$

$$s^2_{\widehat{\beta}} = 0.02697^2 = 0.000727.$$

Таблица 2.4

Dependent Variable: $Y$

| Variable | Coefficient | Std. Error | $t$-Statistic | Probability |
|---|---|---|---|---|
| const | 26.698 | 5.41248 | 4.933 | 0.0006 |
| $X$ | 0.7361 | 0.02697 | 27.29 | 0.0000 |
| $R^2$ | 0.9868 | | | |
| S.E. of regression | 1.534 | | | |

Конечно, все эти результаты можно получить и вручную. Приведем вычисления.

$$n = 12, \qquad \overline{X} = 200,$$
$$\sum X = 2400, \qquad \overline{Y} = 173.92,$$
$$\sum Y = 2087, \qquad \sum Y^2 = 364741,$$
$$\sum X^2 = 483236, \qquad \sum XY = 419782.$$

Далее по формулам из гл. 2 получаем:

$$\widehat{\beta} = \frac{\sum XY - n\overline{XY}}{\sum X^2 - n\overline{X}^2} = \frac{419782 - 12 \cdot 200 \cdot 173.92}{483236 - 12 \cdot 200^2} = 0.736,$$

$$\widehat{\alpha} = \overline{Y} - \widehat{\beta}\overline{X} = 173.92 - 0.736 \cdot 200 = 26.70,$$

$$s^2 = \frac{1}{n-2}\sum(e^2) = \frac{1}{12-2}23.54 = 2.354,$$

$$s^2_{\widehat{\alpha}} = s^2 \frac{\sum X^2}{n(\sum X^2 - n\overline{X}^2)} = 2.354 \frac{483236}{12\,(483236 - 12 \cdot 200^2)} = 29.29,$$

$$s^2_{\widehat{\beta}} = s^2 \frac{1}{\sum X^2 - n\overline{X}^2} = 2.354 \frac{1}{483236 - 12 \cdot 200^2} = 0.000727.$$

## Задача 2.18

Рассмотрим регрессию, построенную в задаче 2.17.

а) Сформулируйте нулевую (основную) и альтернативную гипотезы при проверке статистической значимости коэффициентов регрессии.

б) Какое распределение имеют оценки $\widehat{\alpha}$ и $\widehat{\beta}$?

в) Какое распределение используется при проверке статистической значимости $\alpha$ и $\beta$?

г) Чему равно число степеней свободы?

д) Проверьте на 5%-ном уровне значимость коэффициентов $\alpha$ и $\beta$.

е) Постройте 95%-ный доверительный интервал для коэффициентов $\alpha$ и $\beta$ в регрессии задачи 2.17.

ж) Вычислите коэффициент детерминации, используя равенства $R^2 = \text{RSS}/\text{TSS}$ и $R^2 = 1 - \text{ESS}/\text{TSS}$.

## Решение

а) $\mathrm{H}_0 : \beta = 0$, $\mathrm{H}_1 : \beta \neq 0$; $\mathrm{H}_0 : \alpha = 0$, $\mathrm{H}_1 : \alpha \neq 0$.

б) При условии нормально распределенных ошибок регрессии оценки $\widehat{\alpha}$ и $\widehat{\beta}$ имеют нормальное распределение.

в), г) Распределение Стьюдента с $n - 2 = 10$ степенями свободы.

д) 2.5%-ная точка распределения Стьюдента с $n - 2 = 10$ степенями свободы равна 2.228. Поскольку обе $t$-статистики, приведенные в результатах регрессии (4.933, 27.29), превосходят по абсолютной величине это значение, то обе нулевые гипотезы отвергаются, т. е. оба коэффициента статистически достоверно отличаются от нуля на 5%-ном уровне значимости. (К этому же выводу можно прийти, заметив, что оба $P$-значения меньше 0.05.)

е) Доверительные интервалы следующие:

$$\alpha = \widehat{\alpha} \pm t_{0.025}(n-2)s_{\widehat{\alpha}} = 26.69 \pm 2.228 \cdot 5.41 = 26.69 \pm 12.05;$$
$$\beta = \widehat{\beta} \pm t_{0.025}(n-2)s_{\widehat{\beta}} = 0.736 \pm 2.228 \cdot 0.0269 = 0.736 \pm 0.060.$$

ж) Мы уже нашли ранее $\mathrm{ESS} = 23.54$. Вычислим TSS и RSS.

$$\mathrm{TSS} = \sum (Y_t - \overline{Y})^2 = 1776.9,$$
$$\mathrm{RSS} = \sum (\widehat{Y}_t - \overline{Y})^2 = 1753.37,$$
$$\frac{\mathrm{RSS}}{\mathrm{TSS}} = \frac{1753.37}{1776.9} = 0.98675,$$
$$1 - \frac{\mathrm{ESS}}{\mathrm{TSS}} = 1 - \frac{23.54}{1776.9} = 0.98675.$$

## Задача 2.19

Дана модель парной регрессии $Y_t = \alpha + \beta X_t + \varepsilon_t$, $t = 1, \ldots, n$, для которой выполнены стандартные условия классической линейной модели. Известно, что $n = 2m$. Все множество наблюдений $(Y_t, X_t)$ разбито на две группы $a$ и $b$ по $m$ наблюдений в каждой группе. Обозначим $\overline{X}_a, \overline{X}_b, \overline{Y}_a, \overline{Y}_b$ выборочные средние наблюдений $X, Y$ по группам $a, b$ соответственно. В качестве оценки параметра $\beta$ берется величина $\widetilde{\beta} = (\overline{Y}_a - \overline{Y}_b)/(\overline{X}_a - \overline{X}_b)$.

а) Найдите $\mathrm{E}(\widetilde{\beta})$ и $\mathrm{V}(\widetilde{\beta})$.

б) Каким должно быть разбиение наблюдений на группы $a$ и $b$, чтобы дисперсия $\mathrm{V}(\widetilde{\beta})$ была минимальной?

## Решение

а) Найдем математическое ожидание и дисперсию оценки $\widetilde{\beta}$. В вычислениях мы используем независимость средних $\overline{Y}_a$ и $\overline{Y}_b$, а также свойства выборочного среднего:

$$\mathrm{E}(\overline{X}) = \mathrm{E}(X), \qquad \mathrm{V}(\overline{X}) = \frac{1}{n}\mathrm{V}(X).$$

Используя равенство $\mathrm{E}(\overline{Y}) = \alpha + \beta\overline{X}$, получаем:

$$\mathrm{E}(\widetilde{\beta}) = \frac{\alpha + \beta\overline{X}_a - \alpha - \beta\overline{X}_b}{\overline{X}_a - \overline{X}_b} = \beta,$$

$$\mathrm{V}(\widetilde{\beta}) = \frac{\mathrm{V}(\overline{Y}_a - \overline{Y}_b)}{(\overline{X}_a - \overline{X}_b)^2} = \frac{\frac{1}{m}\sigma^2 + \frac{1}{m}\sigma^2}{(\overline{X}_a - \overline{X}_b)^2} = \frac{2\sigma^2}{m}\frac{1}{(\overline{X}_a - \overline{X}_b)^2}.$$

б) Таким образом, дисперсия $\mathrm{V}(\widetilde{\beta})$ тем меньше, чем больше разница между средними $|\overline{X}_a - \overline{X}_b|$. Очевидно, что если упорядочить все $2m$ наблюдений по возрастанию $X$ и включить первые $m$ наблюдений в группу $a$, а остальные — в группу $b$, то мы получим разбиение с наибольшим возможным значением $|\overline{X}_a - \overline{X}_b|$, а следовательно, и с наименьшим значением дисперсии оценки $\widetilde{\beta}$.

## Задача 2.20

Пусть $Y_t = \beta X_t + \varepsilon_t$, $t = 1, \ldots, n$, где $\mathrm{E}(\varepsilon_t) = 0$, и матрица ковариаций вектора $\varepsilon$ известна. При каких условиях оценки

$$\frac{1}{n}\sum_{t=1}^{n}\frac{Y_t}{X_t} \quad \text{и} \quad \frac{\sum_{t=1}^{n} Y_t}{\sum_{t=1}^{n} X_t}$$

являются наилучшими среди несмещенных линейных оценок параметра $\beta$?

## Решение

а) Рассмотрим первую оценку. Разделим исходное уравнение на $X_t$

$$\frac{Y_t}{X_t} = \beta + \frac{\varepsilon_t}{X_t}.$$

Введем обозначения $Z_t = Y_t/X_t$ и $u_t = \varepsilon_t/X_t$, тогда наше уравнение примет вид регрессии на константу, рассмотренной в задаче 2.12:

$$Z_t = \beta + u_t.$$

Из решения задачи 2.12 следует, что МНК-оценка коэффициента $\beta$ имеет вид:

$$\widehat{\beta} = \overline{Z} = \frac{1}{n}\sum Z_t = \frac{1}{n}\sum\frac{Y_t}{X_t},$$

т. е. совпадает с первой оценкой. Из теоремы Гаусса–Маркова следует тогда, что эта оценка является оптимальной, если $V(\boldsymbol{u}) = \sigma^2 \boldsymbol{I}_n$. Или, другими словами,

$$V(\varepsilon_t) = V(X_t u_t) = X_t^2 V(u_t) = \sigma^2 X_t^2.$$

Итак, если дисперсия ошибок $V(\varepsilon_t)$ пропорциональна $X_t^2$, то первая оценка является наилучшей среди всех несмещенных линейных оценок параметра $\beta$.

б) Предположим для простоты, что $X_t > 0$. Разделим исходное уравнение на $\sqrt{X_t}$ и введем обозначения

$$W_t = \frac{Y_t}{\sqrt{X_t}}, \quad Z_t = \frac{X_t}{\sqrt{X_t}} = \sqrt{X_t}, \quad u_t = \frac{\varepsilon_t}{\sqrt{X_t}}.$$

Получим:

$$\frac{Y_t}{\sqrt{X_t}} = \beta \frac{X_t}{\sqrt{X_t}} + \frac{\varepsilon_t}{\sqrt{X_t}},$$

или

$$W_t = \beta Z_t + u_t.$$

Из задачи 2.8 следует, что МНК-оценка коэффициента $\beta$ имеет вид:

$$\widehat{\beta} = \frac{\sum Z_t W_t}{\sum Z_t^2} = \frac{\sum \sqrt{X_t} \dfrac{Y_t}{\sqrt{X_t}}}{\sum (\sqrt{X_t})^2} = \frac{\sum Y_t}{\sum X_t},$$

т. е. совпадает со второй оценкой. Из теоремы Гаусса–Маркова следует тогда, что эта оценка является оптимальной, если $V(\boldsymbol{u}) = \sigma^2 \boldsymbol{I}_n$. Или, другими словами,

$$V(\varepsilon_t) = V(\sqrt{X_t} u_t) = X_t V(u_t) = \sigma^2 X_t.$$

Итак, если дисперсия ошибок $V(\varepsilon_t)$ пропорциональна $X_t$, то вторая оценка является наилучшей среди всех несмещенных линейных оценок параметра $\beta$.

Конечно, решение задачи выглядит несколько вычурно. Однако мы советуем вернуться к ней после изучения гл. 5 и 6.

## Задача 2.21

Проведены две регрессии:

$$Y_t = \alpha + \beta X_t + \varepsilon_t \quad \text{и} \quad Y_t = \alpha' + \beta' x_t + \varepsilon'_t, \quad t = 1, \ldots, T,$$

где $x_t = X_t - \overline{X}$.

а) По известным МНК-оценкам $\widehat{\alpha}$, $\widehat{\beta}$ параметров $\alpha$, $\beta$ в первой регрессии найдите МНК-оценки $\widehat{\alpha}'$, $\widehat{\beta}'$ параметров $\alpha'$, $\beta'$ во второй регрессии.

б) Найдите $\mathrm{Cov}(\widehat{\alpha}', \widehat{\beta}')$.

## Решение

а) Поскольку $\bar{x} = 0$, то по стандартной формуле для МНК-оценки получаем:

$$\widehat{\beta}' = \frac{\sum(x_t - \bar{x})(Y_t - \overline{Y})}{\sum(x_t - \bar{x})^2} = \frac{\sum(X_t - \overline{X})(Y_t - \overline{Y})}{\sum(X_t - \overline{X})^2} = \widehat{\beta},$$

$$\widehat{\alpha}' = \overline{Y} - \widehat{\beta}'\bar{x} = \overline{Y} - \widehat{\beta}' \cdot 0 = (\overline{Y} - \widehat{\beta X}) + \widehat{\beta X} = \widehat{\alpha} + \widehat{\beta X}.$$

б) Аналогично,

$$\operatorname{Cov}(\widehat{\alpha}', \widehat{\beta}') = -\frac{\bar{x}}{\sum(x_t - \bar{x})^2}\sigma^2 = 0.$$

## Задача 2.22

В таблице 2.5 приведены ежегодные значения денежной массы и национального дохода некоторой гипотетической страны (все величины выражены в миллиардах кварков (название национальной валюты)).

Таблица 2.5

| Год | Денежная масса | Нац. доход | Год | Денежная масса | Нац. доход |
|---|---|---|---|---|---|
| 1981 | 2.0 | 5.0 | 1986 | 4.0 | 7.7 |
| 1982 | 2.5 | 5.5 | 1987 | 4.2 | 8.4 |
| 1983 | 3.2 | 6.0 | 1988 | 4.6 | 9.0 |
| 1984 | 3.6 | 7.0 | 1989 | 4.8 | 9.7 |
| 1985 | 3.3 | 7.2 | 1990 | 5.0 | 10.0 |

а) Проведите регрессию национального дохода ($Y$) на денежную массу ($X$) и константу.

б) Постройте 95%-ный доверительный интервал для оцениваемых параметров. Можете ли вы отвергнуть гипотезу $\beta = 0$? $\beta = 1$?

## Решение

а) $\overline{X} = 3.72$; $\overline{Y} = 7.55$; $\sum(X_t - \overline{X})^2 = 8.796$; $\sum(X_t - \overline{X})(Y_t - \overline{Y}) = 15.09$; $\sum(Y_t - \overline{Y})^2 = 27.005$. Отсюда получаем:

$$\widehat{\beta} = \frac{\sum(X_t - \overline{X})(Y_t - \overline{Y})}{\sum(X_t - \overline{X})^2} = \frac{15.09}{8.796} = 1.715,$$

$$\widehat{\alpha} = \overline{Y} - \widehat{\beta}\overline{X} = 7.55 - 1.715 \cdot 3.72 = 1.168,$$

$$Y_t = 1.168 + 1.715X_t.$$

б) Сумма квадратов остатков равна

$$\sum e_t^2 = \sum (Y_t - \widehat{\alpha} - \widehat{\beta} X_t)^2 = 1.117.$$

Соответственно,

$$s^2 = \frac{1}{n-2} \sum e_t^2 = \frac{1.117}{8} = 0.1397, \quad s_{\widehat{\beta}}^2 = \frac{s^2}{\sum (X_t - \overline{X})^2} = \frac{0.1397}{8.796} = 0.01588,$$

и $s_{\widehat{\beta}} = 0.126$. 95%-ный доверительный интервал для $\beta$ равен

$$\beta = \widehat{\beta} \pm t_{0.025}(n-2) s_{\widehat{\beta}} = 1.715 \pm 2.306 \cdot 0.126, \quad \text{или} \quad (1.42,\ 2.01).$$

Обе гипотезы отвергаются, поскольку 0 и 1 не принадлежат полученному интервалу.

К тому же результату можно прийти, рассчитав соответствующие $t$-статистики:

$$t_1 = \frac{\widehat{\beta}}{s_{\widehat{\beta}}} = 13.61, \quad t_2 = \frac{\widehat{\beta} - 1}{s_{\widehat{\beta}}} = 5.68.$$

Так как $|t_1|$ и $|t_2|$ превосходят $t_{0.025}(8) = 2.306$, то обе гипотезы отвергаются.

## Задача 2.23

Два исследователя, работая независимо друг от друга, изучают одну и ту же регрессионную модель

$$Y_t = \alpha + \beta X_t + \varepsilon_t,$$

для которой выполнены все условия классической модели. В таблице 2.6 приведены результаты, полученные ими на основе независимых выборок.

**Таблица 2.6**

| Выборка I | Выборка II |
|---|---|
| $n = 20$ | $n = 20$ |
| $\sum X_t = 100$ | $\sum X_t = 200$ |
| $\sum X_t^2 = 600$ | $\sum X_t^2 = 2400$ |
| $\sum Y_t = 500$ | $\sum Y_t = 500$ |
| $\widehat{\beta}_{\text{I}} = 2$ | $\widehat{\beta}_{\text{II}} = 2.5$ |

Узнав о работе друг друга, они решают вывести единую оценку параметра $\beta$. Первый исследователь предлагает взять

$$\widetilde{\beta} = \frac{1}{2}(\widehat{\beta}_{\text{I}} + \widehat{\beta}_{\text{II}}).$$

Второй исследователь считает, что весовые коэффициенты первой и второй оценок выбраны неэффективно, и можно построить несмещенную оценку с меньшей дисперсией. Научный руководитель этих исследователей утверждает, что он знает способ еще улучшить общую оценку.

а) Какую оценку предлагает использовать второй исследователь?

б) Какую оценку предлагает использовать научный руководитель?

Оцените улучшение точности оценок пп. а), б) по сравнению с оценкой $\widetilde{\beta}$.

## Решение

Считаем, что исследователи использовали МНК-оценки параметра $\beta$.

а) Напишем оценку $\widetilde{\beta}$ в общем виде: $\widetilde{\beta}_\theta = \theta\widehat{\beta}_{\mathrm{I}} + (1-\theta)\widehat{\beta}_{\mathrm{II}}, \quad 0 \leqslant \theta \leqslant 1$, как взвешенную сумму исходных оценок. Оценка также является несмещенной и ее дисперсия в силу независимости выборок равна

$$\mathrm{V}(\widetilde{\beta}_\theta) = \theta^2\mathrm{V}(\widehat{\beta}_{\mathrm{I}}) + (1-\theta)^2\mathrm{V}(\widehat{\beta}_{\mathrm{II}}) = \theta^2\frac{\sigma^2}{\sum_{\mathrm{I}}(X_t-\overline{X})^2} + (1-\theta)^2\frac{\sigma^2}{\sum_{\mathrm{II}}(X_t-\overline{X})^2}$$

$$= \sigma^2\left(\frac{\theta^2}{\sum_{\mathrm{I}} X_t^2 - \frac{1}{n_{\mathrm{I}}}\sum_{\mathrm{I}} X_t} + \frac{(1-\theta)^2}{\sum_{\mathrm{II}} X_t^2 - \frac{1}{n_{\mathrm{II}}}\sum_{\mathrm{I}} X_t}\right)$$

$$= \sigma^2\left(\frac{\theta^2}{600 - \frac{100^2}{20}} + \frac{(1-\theta)^2}{2400 - \frac{200^2}{20}}\right) = \sigma^2\left(\frac{\theta^2}{100} + \frac{(1-\theta)^2}{400}\right).$$

Минимизируя последнее выражение по $0 \leqslant \theta \leqslant 1$, получим $\theta^* = 1/5$ — оптимальный выбор параметра взвешивания, а оптимальная взвешенная оценка равна

$$\widetilde{\beta}^* = \frac{1}{5}\widehat{\beta}_{\mathrm{I}} + \frac{4}{5}\widehat{\beta}_{\mathrm{II}} = 2.4.$$

б) По теореме Гаусса–Маркова оптимальной линейной несмещенной оценкой $\beta$ является

$$\widehat{\beta} = \frac{n\sum_{\mathrm{I,II}} X_tY_t - \left(\sum_{\mathrm{I,II}} X_t\right)\left(\sum_{\mathrm{I,II}} Y_t\right)}{n\sum_{\mathrm{I,II}} X_t^2 - \left(\sum_{\mathrm{I,II}} X_t\right)^2}.$$

Значения $\sum_{\mathrm{I}} X_tY_t$ и $\sum_{\mathrm{II}} X_tY_t$ можно найти из аналогичных формул для первой и второй выборок. Таким образом, в выражении для $\widehat{\beta}$ все известно, и в итоге можно получить $\widehat{\beta} = 1.6$.

Рассчитаем дисперсии оценок:

$$\mathrm{V}(\widetilde{\beta}) = \mathrm{V}(\widetilde{\beta}_{0.5}) = \sigma^2\left(\frac{0.25}{100} + \frac{0.25}{400}\right) = \frac{\sigma^2}{320},$$

$$\mathrm{V}(\widetilde{\beta}^*) = \mathrm{V}(\widetilde{\beta}_{0.2}) = \sigma^2\left(\frac{0.04}{100} + \frac{0.64}{400}\right) = \frac{\sigma^2}{500},$$

$$\mathrm{V}(\widehat{\beta}) = \frac{\sigma^2}{\sum_{\mathrm{I,II}} X_t^2 - \frac{1}{n}\left(\sum_{\mathrm{I,II}} X_t\right)^2} = \frac{\sigma^2}{600 + 2400 - \frac{1}{40}(100+200)^2} = \frac{\sigma^2}{750}.$$

Отсюда $\mathrm{V}(\widetilde{\beta}^*) = 0.640\,\mathrm{V}(\widetilde{\beta})$ и $\mathrm{V}(\widehat{\beta}) = 0.427\,\mathrm{V}(\widetilde{\beta})$.

## Задача 2.24

Предположим, что модель $Y_t = \alpha + \beta X_t + \varepsilon_t$, $t = 1, \ldots, n$, удовлетворяет условиям классической регрессии. Пусть $\widehat{\alpha}$, $\widehat{\beta}$ — оценки метода наименьших квадратов. Оценка $\widetilde{\beta}$ получена по методу наименьших квадратов при дополнительном (вообще говоря, неверном) предположении, что $\alpha = 0$.

а) Найдите МНК-оценку $\widetilde{\beta}$. При каких условиях она является несмещенной оценкой параметра $\beta$?

б) Найдите дисперсию оценки $\widetilde{\beta}$, сравните ее с дисперсией оценки $\widehat{\beta}$.

в) Обсудите, какую из двух оценок лучше использовать.

## Решение

а) Из решения задачи 2.8 получаем МНК-оценку параметра $\beta$ при дополнительном предположении, что $\alpha = 0$:

$$\widetilde{\beta} = \frac{\sum X_t Y_t}{\sum X_t^2}.$$

Математическое ожидание оценки $\widetilde{\beta}$ равно

$$\mathrm{E}\widetilde{\beta} = \mathrm{E}\frac{\sum X_t Y_t}{\sum X_t^2} = \frac{\sum X_t \mathrm{E} Y_t}{\sum X_t^2} = \frac{\sum X_t(\alpha + \beta X_t)}{\sum X_t^2} = \beta + \alpha\frac{\sum X_t}{\sum X_t^2}.$$

Таким образом, смещение оценки равно 0 в двух случаях:

1) $\alpha = 0$ (т.е. константа отсутствует в истинном уравнении),

2) $\sum X_t = 0$ (менее очевидный случай – регрессор ортогонален константе).

б) $$\mathrm{V}(\widetilde{\beta}) = \mathrm{V}\left(\frac{\sum X_t Y_t}{\sum X_t^2}\right) = \frac{\mathrm{V}(\sum X_t Y_t)}{(\sum X_t^2)^2} = \frac{\sum X_t \mathrm{V}(Y_t)}{(\sum X_t^2)^2} = \frac{\sum X_t^2 \sigma^2}{(\sum X_t^2)^2} = \frac{\sigma^2}{\sum X_t^2}.$$

Поскольку $\sum X_t^2 \leqslant \sum X_t^2 - n\overline{X}^2 = \sum x_t^2$, то

$$\mathrm{V}(\widehat{\beta}) = \frac{\sigma^2}{\sum x_t^2} = \frac{\sigma^2}{\sum X_t^2 - n\overline{X}^2} \geqslant \frac{\sigma^2}{\sum X_t^2} = \mathrm{V}(\widetilde{\beta}).$$

в) Оценка $\widetilde{\beta}$ смещенная, но зато обладает меньшей (по сравнению с оценкой $\widehat{\beta}$) дисперсией. Используя среднеквадратичное отклонение, как критерий сравнения оценок, получаем, что оценка $\widetilde{\beta}$ предпочтительнее оценки $\widehat{\beta}$ при условии

$$\mathrm{MSE}(\widetilde{\beta}) < \mathrm{MSE}(\widehat{\beta}), \quad \text{или}$$

$$\frac{\sigma^2}{\sum X_t^2} + \left(\alpha\frac{\sum X_t}{\sum X_t^2}\right)^2 < \frac{\sigma^2}{\sum X_t^2 - n\overline{X}^2}, \quad \text{или}$$

$$n\alpha^2 < \sigma^2\frac{\sum X_t^2}{\sum X_t^2 - n\overline{X}^2}.$$

## Задача 2.25

Рассмотрим модель парной регрессии $Y_t = \alpha + \beta X_t + \varepsilon_t$. Пусть $Z_t = X_t^2$. Рассмотрим следующую оценку параметра $\beta$:

$$\widetilde{\beta} = \frac{\sum_{t=1}^n (Z_t - \overline{Z})Y_t}{\sum_{t=1}^n (Z_t - \overline{Z})X_t}.$$

а) Покажите, что оценка $\widetilde{\beta}$ несмещенная.

б) Найдите дисперсию оценки $\widetilde{\beta}$.

в) Не повторяя доказательство теоремы Гаусса–Маркова, непосредственно проверьте, что $\mathrm{V}(\widetilde{\beta}) \geqslant \mathrm{V}(\widehat{\beta}_{\mathrm{OLS}})$.

## Решение

а) Используя тождество $\sum(Z_t - \overline{Z}) = 0$, получаем:

$$\mathrm{E}(\widetilde{\beta}) = \mathrm{E}\left(\frac{\sum(Z_t - \overline{Z})Y_t}{\sum(Z_t - \overline{Z})X_t}\right) = \frac{\sum(Z_t - \overline{Z})\mathrm{E}(Y_t)}{\sum(Z_t - \overline{Z})X_t} = \frac{\sum(Z_t - \overline{Z})(\alpha + \beta X_t)}{\sum(Z_t - \overline{Z})X_t}$$
$$= \frac{\sum(Z_t - \overline{Z})(\alpha + \beta X_t)}{\sum(Z_t - \overline{Z})X_t} = \frac{\alpha\sum(Z_t - \overline{Z}) + \beta\sum(Z_t - \overline{Z})X_t}{\sum(Z_t - \overline{Z})X_t} = \beta.$$

б) $$\mathrm{V}(\widetilde{\beta}) = \mathrm{V}\left(\frac{\sum(Z_t - \overline{Z})Y_t}{\sum(Z_t - \overline{Z})X_t}\right) = \frac{\sum(Z_t - \overline{Z})^2\mathrm{V}(Y_t)}{\left(\sum(Z_t - \overline{Z})X_t\right)^2} = \sigma^2\frac{\sum(Z_t - \overline{Z})^2}{\left(\sum(Z_t - \overline{Z})X_t\right)^2}.$$

в) Обозначим $z_t = Z_t - \overline{Z}$, имеем $X_t = \overline{X} + x_t$. Необходимо показать, что

$$\sigma^2\frac{\sum z_t^2}{\left(\sum z_t(\overline{X} + x_t)\right)^2} \geqslant \sigma^2\frac{1}{\sum x_t^2}.$$

Поскольку $\sum z_t = 0$, получаем эквивалентное неравенство:

$$\frac{\sum z_t^2}{\left(\sum z_t x_t\right)^2} \geqslant \frac{1}{\sum x_t^2}, \quad \text{или}$$

$$\left(\sum z_t x_t\right)^2 \leqslant \sum x_t^2 \sum z_t^2.$$

Заметим, что последнее неравенство является неравенством Коши–Буняковского.

# Глава 3

# Модель множественной регрессии

## Задача 3.1

Рассмотрим уравнения:

$$\ln y_t = \beta_1 + \beta_2 \ln w_t + \beta_3 s_t + \varepsilon_t,$$
$$\ln(y_t/w_t) = \gamma_1 + \gamma_2 \ln w_t + \gamma_3 s_t + \varepsilon_t, \qquad t = 1, \ldots, n,$$

где $y_t$ — годовой доход $t$-го индивидуума, $w_t$ — число его рабочих недель в году, $s_t$ — полное число лет, потраченных им на образование.

а) Покажите, что для соответствующих МНК-оценок выполнены соотношения: $\widehat{\gamma}_1 = \widehat{\beta}_1$, $\widehat{\gamma}_3 = \widehat{\beta}_3$, $\widehat{\gamma}_2 = \widehat{\beta}_2 - 1$.

б) Покажите, что остатки этих регрессий совпадают.

в) При каких условиях коэффициент детерминации $R^2$ в первой регрессии будет больше коэффициента детерминации второй регрессии? Что при этом можно сказать о качестве подгонки?

## Решение

а) Воспользуемся формулой (3.4) оценки коэффициентов регрессии. Обозначим $\boldsymbol{X} = [\boldsymbol{\imath} \;\; \boldsymbol{x}_2 \;\; \boldsymbol{x}_3]$ матрицу регрессоров, где $\boldsymbol{\imath} = (1, \ldots, 1)'$, $\boldsymbol{x}_2 = (\ln w_1, \ldots, \ln w_n)'$, $\boldsymbol{x}_3 = (s_1, \ldots, s_n)'$. Пусть $\boldsymbol{z}$ и $\boldsymbol{r}$ — объясняемые переменные для первой и второй регрессий соответственно. Как нетрудно заметить, $\boldsymbol{r} = \boldsymbol{z} - \boldsymbol{x}_2$. Поэтому

$$\widehat{\boldsymbol{\gamma}} = (\boldsymbol{X}'\boldsymbol{X})^{-1}\boldsymbol{X}'\boldsymbol{r} = (\boldsymbol{X}'\boldsymbol{X})^{-1}\boldsymbol{X}'(\boldsymbol{z} - \boldsymbol{x}_2) = \widehat{\boldsymbol{\beta}} - (\boldsymbol{X}'\boldsymbol{X})^{-1}\boldsymbol{X}'\boldsymbol{x}_2.$$

Так как $(\boldsymbol{X}'\boldsymbol{X})^{-1}\boldsymbol{X}'\boldsymbol{X} = \boldsymbol{I}$ (единичная матрица), то $(\boldsymbol{X}'\boldsymbol{X})^{-1}\boldsymbol{X}'\boldsymbol{x}_2$ есть второй столбец матрицы $\boldsymbol{I}$, т.е. $(\boldsymbol{X}'\boldsymbol{X})^{-1}\boldsymbol{X}'\boldsymbol{x}_2 = (0, 1, 0)'$. Итак,

$$\widehat{\boldsymbol{\gamma}} = \widehat{\boldsymbol{\beta}} - (0, 1, 0)', \quad \text{или} \quad \widehat{\gamma}_1 = \widehat{\beta}_1,\ \widehat{\gamma}_2 = \widehat{\beta}_2 - 1,\ \widehat{\gamma}_3 = \widehat{\beta}_3.$$

б) Если $\boldsymbol{e}$ и $\boldsymbol{u}$ — остатки первой и второй регрессий соответственно, то

$$\begin{aligned}\boldsymbol{e} &= \boldsymbol{z} - \boldsymbol{X}\widehat{\boldsymbol{\beta}} = \boldsymbol{r} + \boldsymbol{x}_2 - \boldsymbol{X}(\widehat{\boldsymbol{\gamma}} - (0, 1, 0)') = \boldsymbol{r} + \boldsymbol{x}_2 - \boldsymbol{X}\widehat{\boldsymbol{\gamma}} - \boldsymbol{x}_2 \\ &= \boldsymbol{r} - \boldsymbol{X}\widehat{\boldsymbol{\gamma}} = \boldsymbol{u}.\end{aligned}$$

в) Пусть

$$R_1^2 = 1 - \frac{\boldsymbol{e}'\boldsymbol{e}}{\sum(z_t - \overline{z})^2} = 1 - \frac{\boldsymbol{e}'\boldsymbol{e}}{(\boldsymbol{z} - \overline{z}\boldsymbol{\imath})'(\boldsymbol{z} - \overline{z}\boldsymbol{\imath})},$$
$$R_2^2 = 1 - \frac{\boldsymbol{u}'\boldsymbol{u}}{\sum(r_t - \overline{r})^2} = 1 - \frac{\boldsymbol{u}'\boldsymbol{u}}{(\boldsymbol{r} - \overline{r}\boldsymbol{\imath})'(\boldsymbol{r} - \overline{r}\boldsymbol{\imath})}$$

— коэффициенты детерминации для первой и второй регрессий соответственно. Поэтому $R_1^2 > R_2^2$ в точности тогда, когда

$$(\boldsymbol{z} - \overline{z}\boldsymbol{\imath})'(\boldsymbol{z} - \overline{z}\boldsymbol{\imath}) > (\boldsymbol{r} - \overline{r}\boldsymbol{\imath})'(\boldsymbol{r} - \overline{r}\boldsymbol{\imath}). \qquad (*)$$

Имеем

$$\begin{aligned}(\boldsymbol{r} - \overline{r}\boldsymbol{\imath})'(\boldsymbol{r} - \overline{r}\boldsymbol{\imath}) &= ((\boldsymbol{z} - \overline{z}\boldsymbol{\imath}) - (\boldsymbol{x}_2 - \overline{x}_2\boldsymbol{\imath}))'((\boldsymbol{z} - \overline{z}\boldsymbol{\imath}) - (\boldsymbol{x}_2 - \overline{x}_2\boldsymbol{\imath})) \\ &= (\boldsymbol{z} - \overline{z}\boldsymbol{\imath})'(\boldsymbol{z} - \overline{z}\boldsymbol{\imath}) + (\boldsymbol{x}_2 - \overline{x}_2\boldsymbol{\imath})'(\boldsymbol{x}_2 - \overline{x}_2\boldsymbol{\imath}) \\ &\quad - 2(\boldsymbol{x}_2 - \overline{x}_2\boldsymbol{\imath})'(\boldsymbol{z} - \overline{z}\boldsymbol{\imath}) \\ &= (\boldsymbol{z} - \overline{z}\boldsymbol{\imath})'(\boldsymbol{z} - \overline{z}\boldsymbol{\imath}) - (\boldsymbol{x}_2 - \overline{x}_2\boldsymbol{\imath})'((2\boldsymbol{z} - 2\overline{z}\boldsymbol{\imath}) - (\boldsymbol{x}_2 - \overline{x}_2\boldsymbol{\imath})) \\ &= (\boldsymbol{z} - \overline{z}\boldsymbol{\imath})'(\boldsymbol{z} - \overline{z}\boldsymbol{\imath}) - (\boldsymbol{x}_2 - \overline{x}_2\boldsymbol{\imath})'(2\boldsymbol{z} - (\boldsymbol{x}_2 - \overline{x}_2\boldsymbol{\imath})),\end{aligned}$$

так как $(\boldsymbol{x}_2 - \overline{x}_2\boldsymbol{\imath})'\overline{z}\boldsymbol{\imath} = \overline{z}(\boldsymbol{x}_2'\boldsymbol{\imath}) - \overline{x}_2\overline{z}(\boldsymbol{\imath}'\boldsymbol{\imath}) = \overline{z}(n\overline{x}_2) - \overline{x}_2\overline{z}n = 0$. Значит, $(*)$ выполняется тогда и только тогда, когда

$$(\boldsymbol{x}_2 - \overline{x}_2\boldsymbol{\imath})'(2\boldsymbol{z} - (\boldsymbol{x}_2 - \overline{x}_2\boldsymbol{\imath})) > 0.$$

Обозначим $\boldsymbol{l} = \boldsymbol{x}_2 - \overline{x}_2\boldsymbol{\imath}$.

Условие $(*)$ переписывается в виде $\boldsymbol{l}'(2\boldsymbol{z} - \boldsymbol{l}) > 0$, т.е. угол между векторами $\boldsymbol{l}$ и $2\boldsymbol{z} - \boldsymbol{l}$ (в $R^n$) — острый.

Однако обеим регрессиям соответствует одна и та же геометрическая картинка и экономически содержательная ситуация, поэтому качество подгонки одинаково. А коэффициенты детерминации $R^2$ не совпадают только потому, что зависимость сформулирована в разных координатах.

## Задача 3.2

Покажите, что в регрессии $y_i$ на прогнозные значения $\widehat{y}_i$ и константу свободный член равен 0, а угловой коэффициент равен 1.

## Решение

Проще всего решить эту задачу, используя геометрическую интерпретацию регрессионной модели. Напомним, что вектор $\widehat{\boldsymbol{y}} = \boldsymbol{X}\widehat{\boldsymbol{\beta}}$ есть ортогональная проекция вектора $\boldsymbol{y}$ на $k$-мерную гиперплоскость $P$ — линейную оболочку векторов $\boldsymbol{\imath}, \boldsymbol{x}_2, \ldots, \boldsymbol{x}_k$ (здесь $\boldsymbol{\imath} = (1, 1, \ldots, 1)'$). Пусть $\boldsymbol{e} = \boldsymbol{y} - \widehat{\boldsymbol{y}}$ — вектор остатков. Обозначим через $T$ двумерное подпространство, порожденное векторами $\boldsymbol{\imath}$ и $\widehat{\boldsymbol{y}}$. Так как вектор остатков $\boldsymbol{e}$ ортогонален $P$, а плоскость $T$ является подпространством $P$, то $\boldsymbol{e}$ ортогонален $T$. Поскольку $\boldsymbol{y} = \widehat{\boldsymbol{y}} + \boldsymbol{e}$, то отсюда следует, что ортогональная проекция $\widetilde{\boldsymbol{y}}$ вектора $\boldsymbol{y}$ на $T$ совпадает с $\widehat{\boldsymbol{y}}$: $\widetilde{\boldsymbol{y}} = \widehat{\boldsymbol{y}}$, т. е. $\widetilde{\boldsymbol{y}} = \boldsymbol{0} + 1 \cdot \widehat{\boldsymbol{y}}$, что и требуется.

## Задача 3.3

Дано регрессионное уравнение $y_t = \beta x_t + \varepsilon_t$, $t = 1, \ldots, T$. Ошибки $\varepsilon_t$ — независимые одинаково распределенные нормальные величины. Мы хотим проверить гипотезу, что после наблюдения с номером $n$ значение параметра $\beta$ изменилось. Сумма квадратов остатков с ограничением $\mathrm{ESS_R}$ получается из регрессии $y$ на $x$ по всем $T$ наблюдениям. Для нахождения $\mathrm{ESS_{UR}}$ используются две разные процедуры: 1) оцениваем суммы квадратов остатков регрессий по двум подпериодам $\mathrm{ESS_1}$ и $\mathrm{ESS_2}$, затем их складываем: $\mathrm{ESS_{UR}} = \mathrm{ESS_1} + \mathrm{ESS_2}$; 2) переписываем уравнение в виде

$$y_t = \beta_1 x_t d_{t1} + \beta_2 x_t d_{t2} + \varepsilon_t,$$

где

$$d_{t1} = \begin{cases} 1, & t = 1, \ldots, n, \\ 0, & t = n+1, \ldots, T, \end{cases} \qquad d_{t2} = \begin{cases} 0, & t = 1, \ldots, n, \\ 1, & t = n+1, \ldots, T. \end{cases}$$

Далее мы получаем $\mathrm{ESS_{UR}}$ как сумму квадратов остатков этой регрессии по всем $T$ наблюдениям.

Докажите, что эти две процедуры дают одинаковые значения $F$-статистик.

Покажите также, что тот же результат может быть получен, если регрессия без ограничений записана в виде $y_t = \beta_1 x_t + \delta x_t d_{t2} + \varepsilon_t$, где $\delta = \beta_2 - \beta_1$.

## Решение

В случае 2) метод наименьших квадратов сводится к нахождению минимума функции

$$
\begin{aligned}
F(\beta_1, \beta_2) &= \sum_{t=1}^{T} (y_t - \beta_1 x_t d_{t1} - \beta_2 x_t d_{t2})^2 \\
&= \sum_{t=1}^{n} (y_t - \beta_1 x_t)^2 + \sum_{t=n+1}^{T} (y_t - \beta_2 x_t)^2 = F_1(\beta_1) + F_2(\beta_2),
\end{aligned}
$$

где $F_1(\beta_1)$, $F_2(\beta_2)$ — как раз те функции, которые надо минимизировать при применении метода наименьших квадратов отдельно по первому и второму подпериодам.

Таким образом, первый и второй способы вычисления $\mathrm{ESS_{UR}}$ совпадают, а следовательно, и совпадают значения $F$-статистик (см. формулу (3.50)).

Уравнение $y_t = \beta_1 x_t d_{t1} + \beta_2 x_t d_{t2} + \varepsilon_t$ может быть переписано в следующем виде:

$$
\begin{aligned}
y_t &= \beta_1 x_t (1 - d_{t2}) + \beta_2 x_t d_{t2} + \varepsilon_t = \beta_1 x_t + (\beta_2 - \beta_1) x_t d_{t2} + \varepsilon_t \\
&= \beta_1 x_t + \delta x_t d_{t2} + \varepsilon_t,
\end{aligned}
$$

где $\delta = \beta_2 - \beta_1$ (здесь мы использовали тождество $d_{t1} + d_{t2} \equiv 1$). Таким образом, третий способ совпадает со вторым.

## Задача 3.4

Регрессия зависимой переменной $y$ на три независимые переменные на основе $n = 30$ наблюдений дала следующие результаты:

| | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| $y$ | $=$ | 25.1 | $+$ | $1.2x_1$ | $+$ | $1.0x_2$ | $-$ | $0.50x_3$ |
| Стандартные ошибки | | (2.1) | | (1.5) | | (1.3) | | (0.060) |
| $t$-значения | | (11.9) | | (  ) | | (  ) | | (  ) |
| 95%-ные доверительные границы | | ($\pm$4.3) | | (  ) | | (  ) | | (  ) |

а) Заполните пропуски.

б) Истинны или ложны следующие утверждения (если ложны, исправьте их):

1) Оценка коэффициента при $x_1$ есть 1.2. Другие исследователи могут собрать другие данные и построить другие оценки этого коэффициента. Распределение этих оценок сосредоточено вокруг истинного значения 1.2. Поэтому оценка называется несмещенной.

2) Если есть априорная уверенность в том, что $x_1$ не влияет на $y$, то представляется разумным отвергнуть нулевую гипотезу $\mathrm{H_0}$: $\beta_1 = 0$ на 5%-ном уровне значимости.

3) Если есть априорная уверенность в том, что $x_2$ влияет на $y$, то представляется более разумным использовать оценку 1.0, чем принимать нулевую гипотезу $\mathrm{H_0}$: $\beta_2 = 0$.

## Решение

а) Стандартные ошибки в модели множественной регрессии $y_t = \beta_0 + \beta_1 x_{t1} + \beta_2 x_{t2} + \beta_3 x_{t3} + \varepsilon_t$, $t = 1, \ldots, n$ — это оценки $\widehat{\sigma}_{\widehat{\beta}_i}$ среднеквадратичных отклонений $\sigma_{\widehat{\beta}_i}$, а $t$-значения — это отношения $\widehat{\beta}_i / \widehat{\sigma}_{\widehat{\beta}_i}$, которые распределены по закону Стьюдента с $n-4$ степенями свободы, если верна гипотеза $\mathrm{H}_0\colon \beta_i = 0$. Для получения 95%-ных доверительных границ необходимо умножить $\widehat{\sigma}_{\widehat{\beta}_i}$ на 2.5%-ную точку $t_{0.025}(26) = 2.056$ распределения Стьюдента. После заполнения таблица выглядит так:

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| $y$ | $=$ 25.1 | $+$ | $1.2x_1$ | $+$ | $1.0x_2$ | $-$ | $0.5x_3$ |
| Стандартные ошибки | (2.1) | | (1.5) | | (1.3) | | (0.060) |
| $t$-значения | (11.9) | | (0.8) | | (0.77) | | (−8.33) |
| 95%-ные доверительные границы | (±4.3) | | (±3.08) | | (±2.67) | | (±0.12) |

б) 1) Утверждение ложно. Оценка $\widehat{\beta}_1$ является несмещенной, если ее среднее значение совпадает с истинным значением коэффициента $\beta_1$. То есть другие исследователи действительно могут получить другие значения оценки $\widehat{\beta}_1$, но их распределение не обязано быть сосредоточенным вокруг значения 1.2.

2) Утверждение ложно. Если есть априорная уверенность, что $x_1$ не влияет на $y$ и, кроме того, оценка $\widehat{\beta}_1$ не является значимой, то скорее представляется возможным принять нулевую гипотезу $\mathrm{H}_0\colon \beta_1 = 0$.

3) Утверждение истинно. Несмотря на то, что оценка $\widehat{\beta}_2$ не является значимой, но априорно (например, из общей теории) известно, что $x_2$ влияет на $y$, то представляется разумным не принимать нулевую гипотезу $\mathrm{H}_0\colon \beta_2 = 0$. Возможно, оценка не является значимой из-за недостаточного количества наблюдений.

## Задача 3.5

Бюджетное обследование пяти случайно выбранных семей дало следующие результаты (в тыс. руб.):

| Семья | Накопления, $S$ | Доход, $Y$ | Имущество, $W$ |
|---|---|---|---|
| 1 | 3.0 | 40 | 60 |
| 2 | 6.0 | 55 | 36 |
| 3 | 5.0 | 45 | 36 |
| 4 | 3.5 | 30 | 15 |
| 5 | 1.5 | 30 | 90 |

а) Оцените регрессию $S$ на $Y$ и $W$.

б) Спрогнозируйте накопления семьи, имеющей доход 40 тыс. руб. и имущество стоимостью 25 тыс. руб.

в) Предположим, что доход семьи возрос на 10 тыс. руб., в то время как стоимость имущества не изменилась. Оцените, как возрастут ее накопления.

г) Оцените, как возрастут накопления семьи, если ее доход вырос на 5 тыс. руб., а стоимость имущества увеличилась на 15 тыс. руб.

д) Найдите сумму квадратов остатков и постройте оценку дисперсии регрессии.

## Решение

Запишем регрессию $s_t = \beta_1 + \beta_2 y_t + \beta_3 w_t + \varepsilon_t$ в матричной форме $\boldsymbol{s} = \boldsymbol{X\beta} + \boldsymbol{\varepsilon}$, где $\boldsymbol{X} = [\boldsymbol{\imath} \quad \boldsymbol{y} \quad \boldsymbol{w}]$, $\boldsymbol{\beta} = (\beta_1, \beta_2, \beta_3)'$. Имеем

$$\boldsymbol{s} = \begin{bmatrix} 3.0 \\ 6.0 \\ 5.0 \\ 3.5 \\ 1.5 \end{bmatrix}, \quad \boldsymbol{X} = \begin{bmatrix} 1 & 40 & 60 \\ 1 & 55 & 36 \\ 1 & 45 & 36 \\ 1 & 30 & 15 \\ 1 & 30 & 90 \end{bmatrix}.$$

В дальнейших вычислениях нам понадобятся следующие матрицы:

$$\boldsymbol{X'X} = \begin{bmatrix} 5 & 200 & 237 \\ 200 & 8450 & 9150 \\ 237 & 9150 & 14517 \end{bmatrix}, \quad \boldsymbol{X's} = \begin{bmatrix} 19.0 \\ 825.0 \\ 763.5 \end{bmatrix},$$

$$(\boldsymbol{X'X})^{-1} = 10^{-3} \begin{bmatrix} 5691.635 & -107.392 & -25.231 \\ -107.392 & 2.399 & 0.241 \\ -25.231 & 0.241 & 0.329 \end{bmatrix}.$$

а) Согласно формуле (3.4)

$$\widehat{\boldsymbol{\beta}} = (\boldsymbol{X'X})^{-1}\boldsymbol{X's} = \begin{bmatrix} 0.279 \\ 0.123 \\ -0.029 \end{bmatrix}.$$

б) $\widehat{s} = \widehat{\beta}_1 + \widehat{\beta}_2 y + \widehat{\beta}_3 w = 0.279 + 0.123 \cdot 40 + (-0.029) \cdot 25 = 4.47.$

в) $\Delta s = \widehat{\beta}_2 \Delta y = 10 \cdot 0.123 = 1.23.$

г) $\Delta s = \widehat{\beta}_2 \Delta y + \widehat{\beta}_3 \Delta w = 0.123 \cdot 5 + (-0.029) \cdot 15 = 0.18.$

д) Согласно формулам (3.6) и (3.19) имеем:

$$\sum e_t^2 = \boldsymbol{e'e} = \boldsymbol{s's} - \widehat{\boldsymbol{\beta}}'\boldsymbol{X's} = 84.5 - 84.219 = 0.281,$$

$$s^2 = \widehat{\sigma}^2 = \boldsymbol{e'e}/(n-k) = 0.281/(5-3) = 0.141.$$

## Задача 3.6

Рассмотрим регрессию $s_t = \beta_1 + \beta_2 y_t + \beta_3 w_t + \varepsilon_t$ из предыдущей задачи 3.5.

а) Постройте 95%-ное доверительное множество для

1) $\beta_2$ и $\beta_3$;
2) $\beta_2$;
3) $\beta_3$;
4) $\beta_1$ и $\beta_2$.

б) Проверьте на 5%-ном уровне значимости следующие гипотезы:

1) $\beta_2 = 0$ и $\beta_3 = 0$;
2) $\beta_3 = 0$ (стоимость имущества несущественна);
3) $\beta_2 = 0$ (величина дохода несущественна);
4) $\beta_2 = 1$ (таким мог быть ответ вашего коллеги на вопрос о зависимости накопления от дохода);
5) $\beta_2 = 1.57$ (такое значение коэффициента $\beta_2$ могло быть с высокой степенью надежности установлено для другой страны, и вас интересует вопрос, верно ли это для вашей страны);
6) $\beta_2 = -5\beta_3$ (т. е. эффект дохода противоположен эффекту богатства в фиксированной пропорции).

в) Пусть некоторая семья имеет доход $Y = 30$ тыс. руб. и имущество стоимостью $W = 52.5$ тыс. руб.

1) Чему равна прогнозная величина ее накоплений?
2) В каком смысле эта семья может рассматриваться как средняя между семьями 4 и 5 (задача 3.5)? Почему прогнозная величина ее накоплений не есть среднее между 3.5 и 1.5 тыс. руб.?
3) Постройте 95%-ный доверительный интервал для прогнозной величины накоплений этой семьи.

## Решение

а) Здесь и в дальнейшем будем применять результаты, полученные при решении задачи 3.5.

1) Воспользуемся $F$-статистикой (3.40) при

$$\boldsymbol{H} = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} = [\boldsymbol{0}\ \ \boldsymbol{I}_2]$$

и вычислениями (3.41). Имеем

$$\boldsymbol{H}(\boldsymbol{X}'\boldsymbol{X})^{-1}\boldsymbol{H}' = 10^{-3}\begin{bmatrix} 2.399 & 0.241 \\ 0.241 & 0.329 \end{bmatrix},$$

$$\left(\boldsymbol{H}(\boldsymbol{X}'\boldsymbol{X})^{-1}\boldsymbol{H}'\right)^{-1} = \begin{bmatrix} 450.00 & -330.00 \\ -330.00 & 3283.17 \end{bmatrix}.$$

Таким образом, (см. (3.41)) имеем:

$$\begin{aligned}F &= \frac{(\widehat{\beta}-\beta)'\boldsymbol{H}'\left(\boldsymbol{H}(\boldsymbol{X}'\boldsymbol{X})^{-1}\boldsymbol{H}'\right)^{-1}\boldsymbol{H}(\widehat{\beta}-\beta)/q}{\boldsymbol{e}'\boldsymbol{e}/(n-k)} \\ &= \left(450(0.123-\beta_2)^2+2(-330)(0.123-\beta_2)(-0.029-\beta_3)\right. \\ &\quad \left.+3283.17(-0.029-\beta_3)^2\right)\big)/2/\left(0.281/(5-3)\right) \\ &= 1600.7(0.123-\beta_2)^2+2347.7(0.123-\beta_2)(0.029+\beta_3) \\ &\quad +11678.9(0.029+\beta_3)^2.\end{aligned}$$

Из таблиц находим: $F_{0.05}(2,2)=19.0$, т. е. 95%-ная доверительная область для $(\beta_2,\beta_3)$ описывается неравенством

$$\begin{aligned}&1600.7(0.123-\beta_2)^2+2347.7(0.123-\beta_2)(0.029+\beta_3) \\ &\quad +11678.9(0.029+\beta_3)^2<19.0\end{aligned}$$

или

$$\begin{aligned}&84.25(\beta_2-0.123)^2-123.57(\beta_2-0.123)(\beta_3+0.029) \\ &\quad +614.58(\beta_3+0.029)^2<1. \qquad (*)\end{aligned}$$

2), 3) Из (3.34) получаем, что (для $i=2,3$) интервал $[\widehat{\beta}_i-t_c\widehat{\sigma}_{\widehat{\beta}_i};\ \widehat{\beta}_i+t_c\widehat{\sigma}_{\widehat{\beta}_i}]$ является 95%-ным доверительным интервалом для коэффициента $\beta_i$, где $t_c=4.303$ — 2.5%-ная точка распределения Стьюдента с $n-k=5-3=2$ степенями свободы. Имеем

$$\begin{aligned}\widehat{\beta}_2-t_c\widehat{\sigma}_{\widehat{\beta}_2}&=0.123-4.303\sqrt{0.141\cdot 0.0024}\approx 0.044,\\ \widehat{\beta}_2+t_c\widehat{\sigma}_{\widehat{\beta}_2}&=0.12+4.303\sqrt{0.141\cdot 0.0024}\approx 0.202,\\ \widehat{\beta}_3-t_c\widehat{\sigma}_{\widehat{\beta}_3}&=-0.029-4.303\sqrt{0.141\cdot 0.000033}\approx -0.059,\\ \widehat{\beta}_3+t_c\widehat{\sigma}_{\widehat{\beta}_3}&=-0.029+4.303\sqrt{0.141\cdot 0.000033}\approx -0.00016.\end{aligned}$$

Таким образом, получаем 95%-ные доверительные интервалы

$$\text{для } \beta_2: \ (0.044; 0.202); \quad \text{для } \beta_3: \ (-0.059; -0.00016).$$

4) Аналогично 1), но в этом случае

$$\boldsymbol{H}=\begin{bmatrix}1&0&0\\0&1&0\end{bmatrix}=\left[\boldsymbol{I}_2\ \ \boldsymbol{0}\right].$$

При этом вычисления (3.41) переписываются в виде

$$
\begin{aligned}
F &= \frac{1/q}{e'e/(n-k)}(\widehat{\beta}-\beta)'\begin{bmatrix} I_2 \\ 0 \end{bmatrix}\left(\begin{bmatrix} I_2 & 0 \end{bmatrix}(X'X)^{-1}\begin{bmatrix} I_2 \\ 0 \end{bmatrix}\right)^{-1}\begin{bmatrix} I_2 & 0 \end{bmatrix}(\widehat{\beta}-\beta) \\
&= \frac{0.5}{0.141}\begin{bmatrix} 0.279-\beta_1 & 0.123-\beta_2 \end{bmatrix} 10^3 \begin{bmatrix} 5691.63 & -107.39 \\ -107.39 & 2.40 \end{bmatrix}^{-1}\begin{bmatrix} 0.279-\beta_1 \\ 0.123-\beta_2 \end{bmatrix} \\
&= 3.56\begin{bmatrix} 0.279-\beta_1 & 0.123-\beta_2 \end{bmatrix}\begin{bmatrix} 1.13 & 50.63 \\ 50.63 & 2683.17 \end{bmatrix}\begin{bmatrix} 0.279-\beta_1 \\ 0.123-\beta_2 \end{bmatrix} \\
&= 3.56(1.13(0.279-\beta_1)^2 + 101.26(0.279-\beta_1)(0.123-\beta_2) \\
&\qquad + 2683.17(0.123-\beta_2)^2) \\
&= 4.02(0.279-\beta_1)^2 + 360.13(0.279-\beta_1)(0.123-\beta_2) \\
&\qquad + 9543.16(0.123-\beta_2)^2.
\end{aligned}
$$

Таким образом, 95%-ное доверительное множество для $(\beta_1, \beta_2)$ имеет вид

$$
\begin{aligned}
4.02(0.279-\beta_1)^2 + 360.13(0.279-\beta_1)(0.123-\beta_2)& \\
+ 9543.16(0.123-\beta_2)^2 < 19.0&
\end{aligned}
$$

или

$$
\begin{aligned}
0.21(0.279-\beta_1)^2 + 18.95(0.279-\beta_1)(0.123-\beta_2)& \\
+ 502.27(0.123-\beta_2)^2 < 1.&
\end{aligned}
$$

б) Воспользуемся результатами пункта а).

1) Легко заметить, что точка $\beta_2 = \beta_3 = 0$ не удовлетворяет соотношению $(*)$:

$$84.25(-0.123)^2 - 123.57(-0.123)0.029 + 614.58 \cdot 0.029^2 = 2.23 > 1.$$

Следовательно, гипотезу $H_0$: $\beta_2 = \beta_3 = 0$ отвергаем.

2) Отвергаем, так как $\beta_3 = 0 \notin (-0.059; -0.00016)$.

3) Отвергаем, так как $\beta_2 = 0 \notin (0.044; 0.202)$.

4) Отвергаем, так как $\beta_2 = 1 \notin (0.044; 0.202)$.

5) Отвергаем, так как $\beta_2 = 1.57 \notin (0.044; 0.202)$.

6) Гипотеза $H_0$: $\beta_2 = -5\beta_3$ — частный случай общей линейной гипотезы, рассмотренный в гл. 3, где $c' = (0, 1, 5)$, $\theta = 0$. Вычислим $t$-статистику (3.47):

$$
\begin{aligned}
t &= \frac{c'\widehat{\beta}-\theta}{\sqrt{s^2 c'(X'X)^{-1}c}} \approx \frac{0.12 + 5(-0.029)}{\sqrt{0.141 \cdot 10^{-3}(2.399 + 2\cdot 5\cdot 0.241 + 25\cdot 0.329)}} \\
&\approx -0.60, \qquad |t| = 0.60 < t_{0.025}(2) = 4.303.
\end{aligned}
$$

Значит, данную гипотезу не отвергаем.

в) Вновь воспользуемся полученными выше результатами.

1) Пусть $\boldsymbol{x}_6 = (1,\, 30,\, 52.5)'$ — значения независимых переменных для шестой семьи. Тогда

$$\widehat{s}_6 = \boldsymbol{x}_6'\widehat{\boldsymbol{\beta}} = 0.279 + 30 \cdot 0.123 + 52.5(-0.029) \approx 2.45.$$

2) Если обозначить $\boldsymbol{x}_4 = (1,\, 30,\, 15)'$, $\boldsymbol{x}_5 = (1,\, 30,\, 90)'$ значения независимых переменных (регрессоров) для 4-й и 5-й семей соответственно, то для шестой семьи, о которой говорится в этом пункте, соответствующий набор независимых переменных есть $\boldsymbol{x}_6 = (\boldsymbol{x}_4 + \boldsymbol{x}_5)/2$. Если бы подгонка была идеальной, то тогда действительно имело бы место равенство $\widehat{s}_6 = (s_4 + s_5)/2 = 2.5$, но поскольку, как правило, этого нет, то и $\widehat{s}_6 \neq (s_4 + s_5)/2$.

3) Предполагается, что для рассматриваемой семьи справедлива та же модель, что и для предыдущих пяти, т. е. $s_6 = \boldsymbol{x}_6'\boldsymbol{\beta} + \varepsilon_6$. Тогда $s_6 - \widehat{s}_6 = \boldsymbol{x}_6'(\boldsymbol{\beta} - \widehat{\boldsymbol{\beta}}) + \varepsilon_6$. Из независимости $\widehat{\boldsymbol{\beta}}$ и $\varepsilon_6$ следует, что

$$\mathrm{E}(s_6 - \widehat{s}_6)^2 = \sigma^2\left(1 + \boldsymbol{x}_6'(\boldsymbol{X}'\boldsymbol{X})^{-1}\boldsymbol{x}_6\right).$$

Обозначим

$$\delta = \sqrt{s^2\left(1 + \boldsymbol{x}_6'(\boldsymbol{X}'\boldsymbol{X})^{-1}\boldsymbol{x}_6\right)},$$

где $s^2$ — оценка дисперсии $\sigma^2$, полученная в задаче 3.5 д). Используя те же аргументы, что и в разд. 3.5, получаем, что величина $(s_6 - \widehat{s}_6)/\delta$ имеет распределение Стьюдента с $n - k = 2$ степенями свободы. Поэтому 95%-ный доверительный интервал для $s_6$ есть $(\widehat{s}_6 - t_{0.025}(2)\delta,\ \widehat{s}_6 + t_{0.025}(2)\delta)$. Из таблиц $t_{0.025}(2) = 4.303$. Проводя вычисления (с использованием результатов задачи 3.5), получаем доверительный интервал $(0.52;\ 4.38)$.

## Задача 3.7

Всегда ли доверительный интервал для $\beta_1 + \beta_2$ шире каждого из доверительных интервалов для $\beta_1$ и $\beta_2$? Если да, то почему?

### Решение

Нет. Действительно, пусть $s^2$ — стандартная оценка дисперсии $\sigma^2$ и матрица $s^2(\boldsymbol{X}'\boldsymbol{X})^{-1}$ с элементами $\widehat{\sigma}_{ij}$ — несмещенная оценка ковариационной матрицы $\mathrm{V}(\widehat{\boldsymbol{\beta}}_{\mathrm{OLS}})$. Тогда аналогично (3.34) имеем

$$t = \frac{\widehat{\beta}_1 + \widehat{\beta}_2 - (\beta_1 + \beta_2)}{\sqrt{2\widehat{\sigma}_{12} + \widehat{\sigma}_{11} + \widehat{\sigma}_{22}}} \sim t(n-k),$$

и

$$\left(\widehat{\beta}_1 + \widehat{\beta}_2 - t_c\sqrt{2\widehat{\sigma}_{12} + \widehat{\sigma}_{11} + \widehat{\sigma}_{22}};\ \widehat{\beta}_1 + \widehat{\beta}_2 + t_c\sqrt{2\widehat{\sigma}_{12} + \widehat{\sigma}_{11} + \widehat{\sigma}_{22}}\right)$$

является 95%-ным доверительным интервалом для суммы $\beta_1 + \beta_2$, где $t_c = t_{0.025}(n-k)$. Его ширина равна, очевидно, $2t_c\sqrt{2\widehat{\sigma}_{12} + \widehat{\sigma}_{11} + \widehat{\sigma}_{22}}$, в то время как длины 95%-ных доверительных интервалов для $\beta_1$ и $\beta_2$ равны $2t_c\sqrt{\widehat{\sigma}_{11}}$ и $2t_c\sqrt{\widehat{\sigma}_{22}}$ соответственно. Нетрудно понять, что если $\widehat{\beta}_1$, $\widehat{\beta}_2$ достаточно сильно отрицательно коррелированы (коэффициент корреляции близок к $-1$), то длина доверительного интервала для $\beta_1 + \beta_2$ будет меньше, чем длина каждого из доверительных интервалов для $\beta_1$ и $\beta_2$. Достаточно, например, взять

$$\boldsymbol{X} = \begin{bmatrix} 1 & -1 \\ 2 & 2.1 \\ 1 & 0 \end{bmatrix}.$$

Тогда непосредственные вычисления показывают, что длины доверительных интервалов для $\beta_1 + \beta_2$, $\beta_1$ и $\beta_2$ относятся как $1 : 1.04 : 1.09$.

## Задача 3.8

В этой задаче изучается влияние преобразований зависимой и независимых переменных на МНК-оценки.

а) Что произойдет с МНК-оценками в парной регрессии $y$ на $x$, если добавить константу к каждому наблюдению $y$? к каждому наблюдению $x$? Что произойдет с МНК-оценками в множественной регрессии $y$ на $x_1$ и $x_2$, если добавить константу $c_1$ к каждому наблюдению $x_1$ и другую константу $c_2$ к каждому наблюдению $x_2$?

б) Что произойдет с МНК-оценками в множественной регрессии $y$ на $x_1$ и $x_2$, если переменные $x_1$ и $x_2$ заменить их отклонениями от средних значений?

в) Что произойдет с МНК-оценками в множественной регрессии, если умножить зависимую переменную $y$ на константу? если на константу умножить какой-либо регрессор?

## Решение

а) Будем использовать обозначения главы 3 для формул парной регрессии, записанных в отклонениях от средних значений.

1) Добавление константы к каждому наблюдению $y_t$.

Рассмотрим парную регрессию $y_t = \alpha + \beta x_t + \varepsilon_t$. Заменим $y_t$ на $z_t = y_t + \lambda$. Новая регрессия имеет вид $z_t = \gamma + \delta x_t + u_t$.

В соответствии с (2.6) МНК-оценки для $\beta$ и $\delta$ равны

$$\widehat{\beta} = \frac{\sum x_{*t} y_{*t}}{\sum x_{*t}^2}, \quad \widehat{\delta} = \frac{\sum x_{*t} z_{*t}}{\sum x_{*t}^2}, \quad x_{*t} = x_t - \overline{x}, \quad y_{*t} = y_t - \overline{y}, \quad z_{*t} = z_t - \overline{z}.$$

Заметим, что

$$\overline{z} = \frac{1}{n}\sum z_t = \frac{1}{n}\sum (y_t + \lambda) = \overline{y} + \lambda, \qquad z_{*t} = z_t - \overline{z} = y_{*t}.$$

Следовательно,
$$\widehat{\delta} = \frac{\sum x_{*t} z_{*t}}{\sum x_{*t}^2} = \frac{\sum x_{*t} y_{*t}}{\sum x_{*t}^2} = \widehat{\beta}.$$

Из (2.4б) получаем: $\widehat{\gamma} = \overline{z} - \widehat{\delta}\overline{x} = \overline{y} + \lambda - \widehat{\beta}\overline{x} = \widehat{\alpha} + \lambda$.

Итак, новые МНК-оценки: $\widehat{\delta} = \widehat{\beta}$, $\widehat{\gamma} = \widehat{\alpha} + \lambda$.

2) Добавление константы к каждому наблюдению $x_t$.

Рассмотрим регрессию $y_t = \alpha + \beta x_t + \varepsilon_t$. Заменим теперь $x_t$ на $z_t = x_t + \lambda$, тогда $y_t = \gamma + \delta z_t + u_t$.

В силу (2.6) МНК-оценки для $\beta$ и $\gamma$ равны соответственно
$$\widehat{\beta} = \frac{\sum x_{*t} y_{*t}}{\sum x_{*t}^2},\ \widehat{\delta} = \frac{\sum z_{*t} y_{*t}}{\sum z_{*t}^2},\ x_{*t} = x_t - \overline{x},\ y_{*t} = y_t - \overline{y},\ z_{*t} = z_t - \overline{z}.$$

Заметим, что
$$\overline{z} = \frac{1}{n}\sum z_t = \frac{1}{n}\sum (x_t + \lambda) = \overline{x} + \lambda,\ z_{*t} = z_t - \overline{z} = x_{*t}.$$

Следовательно,
$$\widehat{\delta} = \frac{\sum z_{*t} y_{*t}}{\sum z_{*t}^2} = \frac{\sum x_{*t} y_{*t}}{\sum x_{*t}^2} = \widehat{\beta}.$$

Из (2.4б) получаем: $\widehat{\gamma} = \overline{y} - \widehat{\delta}\overline{z} = \overline{y} - \widehat{\beta}(\overline{x} + \lambda) = \widehat{\alpha} - \widehat{\beta}\lambda$.

Итак, новые МНК-оценки: $\widehat{\delta} = \widehat{\beta}$, $\widehat{\gamma} = \widehat{\alpha} - \lambda\widehat{\beta}$.

3) Добавление констант $c_1$, $c_2$ в регрессии с двумя независимыми переменными.

В матричной записи модель имеет вид $\boldsymbol{y} = \boldsymbol{X}\boldsymbol{\beta} + \boldsymbol{\varepsilon}$, где
$$\boldsymbol{X} = \begin{bmatrix} 1 & x_{11} & x_{12} \\ \vdots & \vdots & \vdots \\ 1 & x_{n1} & x_{n2} \end{bmatrix} \text{— матрица } n \times 3, \quad \boldsymbol{y} = \begin{bmatrix} y_1 \\ \vdots \\ y_n \end{bmatrix}, \quad \boldsymbol{\beta} = \begin{bmatrix} \beta_1 \\ \beta_2 \\ \beta_3 \end{bmatrix}.$$

Заменим теперь $\boldsymbol{X}$ на $\boldsymbol{Z} = \boldsymbol{X} + \boldsymbol{C} = \boldsymbol{X}\boldsymbol{A}$, где
$$\boldsymbol{C} = \begin{bmatrix} 0 & c_1 & c_2 \\ \vdots & \vdots & \vdots \\ 0 & c_1 & c_2 \end{bmatrix}, \quad \boldsymbol{A} = \begin{bmatrix} 1 & c_1 & c_2 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

— матрицы размерности $n \times 3$ и $3 \times 3$ соответственно.

Построим регрессию $\boldsymbol{y}$ на $\boldsymbol{Z}$: $\boldsymbol{y} = \boldsymbol{Z}\boldsymbol{\gamma} + \boldsymbol{u}$. В соответствии с формулой (3.4)
$$\widehat{\boldsymbol{\beta}}_{\text{OLS}} = (\boldsymbol{X}'\boldsymbol{X})^{-1}\boldsymbol{X}'\boldsymbol{y}, \quad \widehat{\boldsymbol{\gamma}}_{\text{OLS}} = (\boldsymbol{Z}'\boldsymbol{Z})^{-1}\boldsymbol{Z}'\boldsymbol{y}.$$

Подставляя $\boldsymbol{Z} = \boldsymbol{X}\boldsymbol{A}$ в выражение для $\widehat{\boldsymbol{\gamma}}_{\text{OLS}}$, получаем:
$$\widehat{\boldsymbol{\gamma}}_{\text{OLS}} = ((\boldsymbol{X}\boldsymbol{A})'(\boldsymbol{X}\boldsymbol{A}))^{-1}(\boldsymbol{X}\boldsymbol{A})'\boldsymbol{y} = (\boldsymbol{A}'\boldsymbol{X}'\boldsymbol{X}\boldsymbol{A})^{-1}\boldsymbol{A}'\boldsymbol{X}'\boldsymbol{y}$$
$$= \boldsymbol{A}^{-1}(\boldsymbol{X}'\boldsymbol{X})^{-1}(\boldsymbol{A}')^{-1}\boldsymbol{A}'\boldsymbol{X}'\boldsymbol{y} = \boldsymbol{A}^{-1}(\boldsymbol{X}'\boldsymbol{X})^{-1}\boldsymbol{X}'\boldsymbol{y} = \boldsymbol{A}^{-1}\widehat{\boldsymbol{\beta}}_{\text{OLS}}.$$

Имеем

$$A^{-1} = \begin{bmatrix} 1 & -c_1 & -c_2 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}.$$

Окончательно получаем следующее выражение для новой МНК-оценки:

$$\widehat{\gamma}_{\text{OLS}} = A^{-1}\widehat{\beta}_{\text{OLS}} = \begin{bmatrix} 1 & -c_1 & -c_2 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} \widehat{\beta}_1 \\ \widehat{\beta}_2 \\ \widehat{\beta}_3 \end{bmatrix} = \begin{bmatrix} \widehat{\beta}_1 - c_1\widehat{\beta}_2 - c_2\widehat{\beta}_3 \\ \widehat{\beta}_2 \\ \widehat{\beta}_3 \end{bmatrix}$$

б) Замена переменных $X$ их отклонениями.

Теперь вместо матрицы $X$ возьмем матрицу $Z = X + B$, где

$$X = \begin{bmatrix} 1 & x_{11} & x_{12} \\ \vdots & \vdots & \vdots \\ 1 & x_{n1} & x_{n2} \end{bmatrix}, \quad B = \begin{bmatrix} 0 & -\overline{x}_1 & -\overline{x}_2 \\ \vdots & \vdots & \vdots \\ 0 & -\overline{x}_1 & -\overline{x}_2 \end{bmatrix}$$

— матрицы размера $n \times 3$, причем $\overline{x}_1 = 1/n \sum x_{t1}$, $\overline{x}_2 = 1/n \sum x_{t2}$ — средние значения регрессоров. Аналогично п. а 3) данной задачи имеем

$$X + B = XD, \quad \text{где} \ D = \begin{bmatrix} 1 & -\overline{x}_1 & -\overline{x}_2 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}.$$

Рассматривая регрессии $y = X\beta + \varepsilon$, $y = Z\gamma + u$, получаем (аналогично п. а 3)):

$$\widehat{\gamma}_{\text{OLS}} = D^{-1}\widehat{\beta}_{\text{OLS}}, \quad D^{-1} = \begin{bmatrix} 1 & \overline{x}_1 & \overline{x}_2 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}.$$

Окончательно выражение для новой МНК-оценки имеет вид

$$\widehat{\gamma}_{\text{OLS}} = D^{-1}\widehat{\beta}_{\text{OLS}} = \begin{bmatrix} \widehat{\beta}_1 + \overline{x}_1\widehat{\beta}_2 + \overline{x}_2\widehat{\beta}_3 \\ \widehat{\beta}_2 \\ \widehat{\beta}_3 \end{bmatrix}.$$

в) Воспользуемся основными формулами множественной регрессии.

1) Умножение зависимой переменной на константу.

Рассматривая регрессии $y = X\beta + \varepsilon$, $z = X\gamma + u$, где $z = cy$, и пользуясь формулой (3.4), получаем $\widehat{\gamma}_{\text{OLS}} = (X'X)^{-1}X'z = (X'X)^{-1}X'cy = c\widehat{\beta}_{\text{OLS}}$.

Итак, новая МНК-оценка есть $\widehat{\gamma}_{\text{OLS}} = c\widehat{\beta}_{\text{OLS}}$.

2) Умножение на константу одного из регрессоров.

Без ограничения общности можно считать, что на константу умножается первый регрессор. Тогда матрицу

$$X = \begin{bmatrix} x_{11} & \cdots & x_{1k} \\ \dots & \dots & \dots \\ x_{n1} & \cdots & x_{nk} \end{bmatrix} \quad \text{заменяем на матрицу} \quad Z = \begin{bmatrix} cx_{11} & x_{12} & \cdots & x_{1k} \\ \dots & \dots & \dots & \dots \\ cx_{n1} & x_{n2} & \cdots & x_{nk} \end{bmatrix},$$

где $\boldsymbol{X}$, $\boldsymbol{Z}$ — $n \times k$ матрицы. Нетрудно проверить, что $\boldsymbol{Z} = \boldsymbol{XC}$, где

$$\boldsymbol{C} = \begin{bmatrix} c & 0 & \dots & 0 \\ 0 & 1 & \dots & 0 \\ 0 & 0 & \dots & 0 \\ \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & 1 \end{bmatrix}$$

— $k \times k$ матрица.

Рассматривая регрессии $\boldsymbol{y} = \boldsymbol{X\beta} + \varepsilon$, $\boldsymbol{y} = \boldsymbol{Z\gamma} + \boldsymbol{u}$, получаем (аналогично п. а 3)):

$$\widehat{\gamma}_{\mathrm{OLS}} = \boldsymbol{C}^{-1}\widehat{\boldsymbol{\beta}}_{\mathrm{OLS}}, \quad \boldsymbol{C}^{-1} = \begin{bmatrix} 1/c & 0 & \dots & 0 \\ 0 & 1 & \dots & 0 \\ 0 & 0 & \dots & 0 \\ \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & 1 \end{bmatrix}.$$

Таким образом, новая МНК-оценка есть

$$\widehat{\gamma}_{\mathrm{OLS}} = \left[\frac{1}{c}\widehat{\beta}_1, \ \widehat{\beta}_2, \ \dots, \ \widehat{\beta}_k\right]'.$$

## Задача 3.9

Рассмотрим оценку вида $\widetilde{\boldsymbol{\beta}} = ((\boldsymbol{X}'\boldsymbol{X})^{-1} + \gamma \boldsymbol{I})\,\boldsymbol{X}'\boldsymbol{y}$ для вектора коэффициентов регрессионного уравнения $\boldsymbol{y} = \boldsymbol{X\beta} + \varepsilon$. ($\boldsymbol{I}$ — единичная $k \times k$ матрица.)

а) Найдите математическое ожидание, матрицу ковариаций и матрицу среднеквадратичных отклонений оценки $\widetilde{\boldsymbol{\beta}}$ $(\mathrm{MSE}(\widehat{\boldsymbol{\theta}}) = \mathrm{E}((\widehat{\boldsymbol{\theta}} - \boldsymbol{\theta})(\widehat{\boldsymbol{\theta}} - \boldsymbol{\theta})'))$.

б) Можно ли найти $\gamma$ такое, что оценка $\widetilde{\boldsymbol{\beta}}$ более эффективна, чем оценка метода наименьших квадратов $\widehat{\boldsymbol{\beta}}$ (т. е. для всех $i = 1, \dots, k$, $\mathrm{MSE}(\widetilde{\beta}_i) < \mathrm{MSE}(\widehat{\beta}_i)$)?

## Решение

Преобразуем выражение для $\widetilde{\boldsymbol{\beta}}$:

$$\begin{aligned} \widetilde{\boldsymbol{\beta}} &= ((\boldsymbol{X}'\boldsymbol{X})^{-1} + \gamma\boldsymbol{I})\,\boldsymbol{X}'\boldsymbol{y} = ((\boldsymbol{X}'\boldsymbol{X})^{-1} + \gamma\boldsymbol{I})\,\boldsymbol{X}'(\boldsymbol{X\beta} + \varepsilon) \\ &= \boldsymbol{\beta} + \gamma\boldsymbol{X}'\boldsymbol{X\beta} + (\boldsymbol{X}'\boldsymbol{X})^{-1}\boldsymbol{X}'\varepsilon + \gamma\boldsymbol{X}'\varepsilon = \boldsymbol{\beta} + \gamma\boldsymbol{X}'\boldsymbol{X\beta} + \boldsymbol{C}\varepsilon, \end{aligned}$$

где $\boldsymbol{C} = (\boldsymbol{X}'\boldsymbol{X})^{-1}\boldsymbol{X}' + \gamma\boldsymbol{X}'$.

а) Так как $\mathrm{E}(\varepsilon) = \boldsymbol{0}$, то $\mathrm{E}(\widetilde{\boldsymbol{\beta}}) = \boldsymbol{\beta} + \gamma\boldsymbol{X}'\boldsymbol{X\beta}$. Далее, поскольку вектор $\boldsymbol{\beta} + \gamma\boldsymbol{X}'\boldsymbol{X\beta}$ является неслучайным, то

$$\begin{aligned} \mathrm{V}(\widetilde{\boldsymbol{\beta}}) &= \mathrm{V}(\boldsymbol{C}\varepsilon) = \sigma^2\boldsymbol{CC}' = \sigma^2\left((\boldsymbol{X}'\boldsymbol{X})^{-1} + \gamma\boldsymbol{I}\right)\boldsymbol{X}'\boldsymbol{X}\left((\boldsymbol{X}'\boldsymbol{X})^{-1} + \gamma\boldsymbol{I}\right) \\ &= \sigma^2\left((\boldsymbol{X}'\boldsymbol{X})^{-1} + 2\gamma\boldsymbol{I} + \gamma^2\boldsymbol{X}'\boldsymbol{X}\right) = \mathrm{V}(\widehat{\boldsymbol{\beta}}_{\mathrm{OLS}}) + \sigma^2(2\gamma\boldsymbol{I} + \gamma^2\boldsymbol{X}'\boldsymbol{X}). \end{aligned}$$

Для матрицы среднеквадратичных отклонений получаем следующее выражение:

$$\begin{aligned}\mathrm{MSE}(\widetilde{\beta}) &= \mathrm{E}(\gamma X'X\beta + C\varepsilon)(\gamma X'X\beta + C\varepsilon)' = \gamma^2 X'X\beta\beta'X'X + \sigma^2 CC' \\ &= \gamma^2 X'X\beta\beta'X'X + \sigma^2\left(((X'X)^{-1}X' + \gamma X')((X'X)^{-1}X' + \gamma X')'\right) \\ &= \gamma^2(X'X\beta\beta'X'X + \sigma^2 X'X) + 2\sigma^2\gamma I + \sigma^2(X'X)^{-1} \\ &= \gamma^2 A + 2\sigma^2\gamma I + \mathrm{MSE}(\widehat{\beta}_{\mathrm{OLS}}),\end{aligned}$$

где $A = X'X\beta\beta'X'X + \sigma^2 X'X$.

б) Имеем $\mathrm{MSE}(\widetilde{\beta}_i) = \gamma^2 a_{ii} + 2\sigma^2\gamma + \mathrm{MSE}(\widehat{\beta}_{i,\mathrm{OLS}})$, где $a_{ii}$ — $i$-й диагональный элемент матрицы $A$. Рассматривая $\mathrm{MSE}(\widetilde{\beta}_i)$ как функцию $\gamma$ и обозначая для краткости $\mathrm{MSE}(\widetilde{\beta}_i) = h_i(\gamma)$, видим, что $h_i(0) = \mathrm{MSE}(\widehat{\beta}_{i,\mathrm{OLS}})$ и $h_i'(0) = 2\sigma^2$, т. е. функция $h_i$ возрастает в точке $\gamma = 0$ и, следовательно, существует такое $\gamma < 0$, что $h_i(\gamma) < h_i(0)$ для всех $i = 1, \ldots, k$.

## Задача 3.10

Рассмотрим оценку вида $\widetilde{\beta} = (X'X + rD)^{-1}X'y$ (ридж-регрессия (*ridge regression*)) для вектора коэффициентов регрессионного уравнения $y = X\beta + \varepsilon$, где $D$ — диагональная $k \times k$ матрица, состоящая из диагональных элементов матрицы $X'X$.

а) Найдите математическое ожидание, матрицу ковариаций и матрицу среднеквадратичных отклонений оценки $\widetilde{\beta}$ ($\mathrm{MSE}(\widehat{\theta}) = \mathrm{E}((\widehat{\theta} - \theta)(\widehat{\theta} - \theta)')$).

б) Покажите, что существует $r > 0$ такое, что $\mathrm{V}(\widetilde{\beta}) < \mathrm{V}(\widehat{\beta})$, где $\widehat{\beta}$ — оценка метода наименьших квадратов.

в) Можно ли найти такое $r > 0$, что оценка $\widetilde{\beta}$ более эффективна, чем оценка метода наименьших квадратов $\widehat{\beta}$ (т.е. для всех $i = 1, \ldots, k$, $\mathrm{MSE}(\widetilde{\beta}_i) < \mathrm{MSE}(\widehat{\beta}_i)$)?

## Решение

а) Преобразуем выражение для $\widetilde{\beta}$:

$$\begin{aligned}\widetilde{\beta} &= (X'X + rD)^{-1}X'y = (X'X + rD)^{-1}X'X\beta + (X'X + rD)^{-1}X'\varepsilon \\ &= (X'X + rD)^{-1}(X'X + rD - rD)\beta + (X'X + rD)^{-1}X'\varepsilon \\ &= \beta - r(X'X + rD)^{-1}D\beta + (X'X + rD)^{-1}X'\varepsilon \\ &= \beta - r(X'X + rD)^{-1}D\beta + C\varepsilon,\end{aligned}$$

где $C = (X'X + rD)^{-1}X'$. Так как $\mathrm{E}(\varepsilon) = 0$, то

$$\mathrm{E}(\widetilde{\beta}) = \beta - r(X'X + rD)^{-1}D\beta.$$

Вычислим матрицу ковариаций. Поскольку вектор $\boldsymbol{\beta} - r(\boldsymbol{X}'\boldsymbol{X} + r\boldsymbol{D})^{-1}\boldsymbol{D}\boldsymbol{\beta}$ не случайный, то

$$\mathrm{V}(\widetilde{\boldsymbol{\beta}}) = \mathrm{V}(\boldsymbol{C}\boldsymbol{\varepsilon}) = \sigma^2\boldsymbol{C}\boldsymbol{C}' = \sigma^2(\boldsymbol{X}'\boldsymbol{X} + r\boldsymbol{D})^{-1}\boldsymbol{X}'\boldsymbol{X}(\boldsymbol{X}'\boldsymbol{X} + r\boldsymbol{D})^{-1}.$$

Для матрицы среднеквадратичных отклонений получаем следующее выражение:

$$\begin{aligned}\mathrm{MSE}(\widetilde{\boldsymbol{\beta}}) &= \mathrm{E}((\widetilde{\boldsymbol{\beta}} - \boldsymbol{\beta})(\widetilde{\boldsymbol{\beta}} - \boldsymbol{\beta})') \\ &= \mathrm{E}[(-r(\boldsymbol{X}'\boldsymbol{X} + r\boldsymbol{D})^{-1}\boldsymbol{D}\boldsymbol{\beta} + (\boldsymbol{X}'\boldsymbol{X} + r\boldsymbol{D})^{-1}\boldsymbol{X}'\boldsymbol{\varepsilon}) \\ &\quad \times (-r(\boldsymbol{X}'\boldsymbol{X} + r\boldsymbol{D})^{-1}\boldsymbol{D}\boldsymbol{\beta} + (\boldsymbol{X}'\boldsymbol{X} + r\boldsymbol{D})^{-1}\boldsymbol{X}'\boldsymbol{\varepsilon})'] \\ &= r^2(\boldsymbol{X}'\boldsymbol{X} + r\boldsymbol{D})^{-1}\boldsymbol{D}\boldsymbol{\beta}\boldsymbol{\beta}'\boldsymbol{D}(\boldsymbol{X}'\boldsymbol{X} + r\boldsymbol{D})^{-1} \\ &\quad + \sigma^2(\boldsymbol{X}'\boldsymbol{X} + r\boldsymbol{D})^{-1}\boldsymbol{X}'\boldsymbol{X}(\boldsymbol{X}'\boldsymbol{X} + r\boldsymbol{D})^{-1}.\end{aligned}$$

б) Так как оценка $\widehat{\boldsymbol{\beta}}$ несмещенная, то $\mathrm{V}(\widehat{\boldsymbol{\beta}}) = \mathrm{MSE}(\widehat{\boldsymbol{\beta}})$. Из а) следует, что

$$\mathrm{MSE}(\widetilde{\boldsymbol{\beta}}) = r^2(\boldsymbol{X}'\boldsymbol{X} + r\boldsymbol{D})^{-1}\boldsymbol{D}\boldsymbol{\beta}\boldsymbol{\beta}'\boldsymbol{D}(\boldsymbol{X}'\boldsymbol{X} + r\boldsymbol{D})^{-1} + \mathrm{V}(\widetilde{\boldsymbol{\beta}}),$$

откуда, учитывая положительную определенность матрицы

$$r^2(\boldsymbol{X}'\boldsymbol{X} + r\boldsymbol{D})^{-1}\boldsymbol{D}\boldsymbol{\beta}\boldsymbol{\beta}'\boldsymbol{D}(\boldsymbol{X}'\boldsymbol{X} + r\boldsymbol{D})^{-1}$$

(см. приведенную ниже лемму), получаем

$$\mathrm{V}(\widetilde{\boldsymbol{\beta}}) \leqslant \mathrm{MSE}(\widetilde{\boldsymbol{\beta}}).$$

Таким образом, б) вытекает из в).

в) Рассмотрим разность матриц среднеквадратичных отклонений:

$$\begin{aligned}&\mathrm{MSE}(\widetilde{\boldsymbol{\beta}}) - \mathrm{MSE}(\widehat{\boldsymbol{\beta}}) \\ &\quad = (\boldsymbol{X}'\boldsymbol{X} + r\boldsymbol{D})^{-1}\Big(r^2\boldsymbol{D}\boldsymbol{\beta}\boldsymbol{\beta}'\boldsymbol{D} + \sigma^2(\boldsymbol{X}'\boldsymbol{X})\Big)(\boldsymbol{X}'\boldsymbol{X} + r\boldsymbol{D})^{-1} \\ &\qquad - \sigma^2(\boldsymbol{X}'\boldsymbol{X})^{-1} \\ &\quad = (\boldsymbol{X}'\boldsymbol{X} + r\boldsymbol{D})^{-1}\Big(r^2\boldsymbol{D}\boldsymbol{\beta}\boldsymbol{\beta}'\boldsymbol{D} + \sigma^2(\boldsymbol{X}'\boldsymbol{X}) \\ &\qquad - \sigma^2(\boldsymbol{X}'\boldsymbol{X} + r\boldsymbol{D})(\boldsymbol{X}'\boldsymbol{X})^{-1}(\boldsymbol{X}'\boldsymbol{X} + r\boldsymbol{D})\Big)(\boldsymbol{X}'\boldsymbol{X} + r\boldsymbol{D})^{-1} \\ &\quad = (\boldsymbol{X}'\boldsymbol{X} + r\boldsymbol{D})^{-1}(-2r\sigma^2\boldsymbol{D} + r^2\boldsymbol{H})(\boldsymbol{X}'\boldsymbol{X} + r\boldsymbol{D})^{-1},\end{aligned}$$

где $\boldsymbol{H} = \boldsymbol{D}\boldsymbol{\beta}\boldsymbol{\beta}'\boldsymbol{D} - \sigma^2\boldsymbol{D}(\boldsymbol{X}'\boldsymbol{X})^{-1}\boldsymbol{D}$. Матрица $\boldsymbol{D}$ диагональная с положительными элементами на диагонали (квадраты длин векторов-столбцов матрицы $\boldsymbol{X}$), следовательно, положительно определенная. Матрица $\boldsymbol{H}$ симметричная. Нетрудно понять, что можно найти такое $r_0$, что при $r < r_0$ матрица $-2r\sigma^2\boldsymbol{D} + r^2\boldsymbol{H}$ отрицательно определена. Отсюда следует, что и матрица

$$\mathrm{MSE}(\widetilde{\boldsymbol{\beta}}) - \mathrm{MSE}(\widehat{\boldsymbol{\beta}}) = (\boldsymbol{X}'\boldsymbol{X} + r\boldsymbol{D})^{-1}(-2r\sigma^2\boldsymbol{D} + r^2\boldsymbol{H})(\boldsymbol{X}'\boldsymbol{X} + r\boldsymbol{D})^{-1}$$

отрицательно определена, что и требовалось показать.

(*Лемма.* Пусть $k \times k$ симметричная матрица $\boldsymbol{A}$ невырождена. Тогда $k \times k$ матрица $\boldsymbol{H}$ положительно определена тогда и только тогда, когда положительно определена матрица $\boldsymbol{AHA}$.)

## Задача 3.11

После финансового кризиса спрос на чебуреки (см. задачу 2.14) упал, и менеджер был вынужден тратить часть средств на рекламу. Для изучения зависимости объема продаж от цены и расходов на рекламу менеджер использует следующую модель:

$$q_t = \beta_1 + \beta_2 p_t + \beta_3 a_t + \beta_4 a_t^2 + \varepsilon_t.$$

В таблице 3.1 приведены данные наблюдений за 20 недель ($t$ — номер недели, $q_t$ — количество проданных чебуреков, $p_t$ — цена одного чебурека (руб.), $a_t$ — затраты на рекламу (100 руб.)).

**Таблица 3.1**

| $t$ | $q_t$ | $p_t$ | $a_t$ | $t$ | $q_t$ | $p_t$ | $a_t$ |
|---|---|---|---|---|---|---|---|
| 1 | 525 | 5.92 | 4.79 | 11 | 407 | 6.67 | 5.19 |
| 2 | 567 | 6.50 | 3.61 | 12 | 608 | 6.92 | 3.27 |
| 3 | 396 | 6.54 | 5.49 | 13 | 399 | 6.97 | 4.69 |
| 4 | 726 | 6.11 | 2.78 | 14 | 631 | 6.59 | 3.79 |
| 5 | 265 | 6.62 | 5.74 | 15 | 545 | 6.50 | 4.29 |
| 6 | 615 | 5.15 | 1.34 | 16 | 512 | 6.86 | 2.71 |
| 7 | 370 | 5.02 | 5.81 | 17 | 845 | 5.09 | 2.21 |
| 8 | 789 | 5.02 | 3.39 | 18 | 571 | 6.08 | 3.09 |
| 9 | 513 | 6.77 | 3.74 | 19 | 539 | 6.36 | 4.65 |
| 10 | 661 | 5.57 | 3.59 | 20 | 620 | 6.22 | 1.97 |

Используя данные таблицы 3.1, ответьте на следующие вопросы:

а) Отклик количества проданных чебуреков на изменение цены измеряется коэффициентом $\beta_2 = \partial q/\partial p$. Аналогично, $\partial q/\partial a = \beta_3 + 2\beta_4 a$. Какие знаки $\beta_2$, $\beta_3$, $\beta_4$ вы ожидаете получить?

б) Найдите оценки коэффициентов регрессии и их стандартные ошибки. Соответствуют ли знаки оценок вашим ожиданиям?

в) Пусть себестоимость производства одного чебурека равна 2 руб. Тогда чистый доход за неделю задается формулой $profit = pq - 2q - 100a$.

г) Найдите оптимальную цену при расходах на рекламу, равных 280 руб.

д) Найдите оптимальный уровень расходов на рекламу при цене чебурека, равной 6 руб.

е) Помогите менеджеру найти оптимальное решение (максимизирующее чистый доход).

ж) Найдите 95%-ные доверительные интервалы для $\beta_2$, $\beta_3$, $\beta_4$. Проверьте значимость влияния цены, а также расходов на рекламу на количество проданных чебуреков.

## Решение

а) Естественно предполагать, что рост цены на чебуреки приводит к уменьшению спроса на них, т. е. ожидаемая оценка коэффициента $\beta_2$ отрицательная. Увеличение расходов на рекламу должно приводить к росту спроса на них, при этом разумно считать, что при большом уровне расходов каждый последующий рубль, вложенный в рекламу, должен давать меньшую отдачу, чем тот же рубль при более низком уровне расходов на рекламу. Иными словами, разумно предполагать, что для расходов на рекламу действует «эффект насыщения». Таким образом, ожидаемый знак оценки $\beta_3$ — «+», оценки $\beta_4$ — «−».

б) Оценка параметров модели с применением эконометрического пакета дает следующий результат.

Dependent Variable: $q$

| Variable | Coefficient | Std. Error | $t$-Statistic | Probability |
|---|---|---|---|---|
| const | 957.335 | 129.218 | 7.409 | 0.0000 |
| $p$ | −110.153 | 21.327 | −5.165 | 0.0001 |
| $a$ | 255.529 | 62.225 | 4.107 | 0.0008 |
| $a^2$ | −43.376 | 8.020 | −5.408 | 0.0001 |
| $R^2$ | 0.8778 | | | |

Как мы видим, все коэффициенты значимы (на 5%-ном уровне) и их знаки соответствуют здравому смыслу и экономической интуиции.

в), г) С помощью непосредственных вычислений получаем:

$$\frac{\partial(profit)}{\partial p} = \beta_1 + 2p\beta_2 + \beta_3 a + \beta_4 a^2 - 2\beta_2.$$

Приравнивая эту производную нулю и учитывая, что $\beta_2 < 0$, получаем следующее значение цены, при которой максимизируется средний доход при фиксированных расходах на рекламу $a$:

$$p_{\text{opt}} = \frac{2\beta_2 - \beta_1 - \beta_3 a - \beta_4 a^2}{2\beta_2}.$$

Подставляя вместо коэффициентов $\beta_i$ их оценки $\widehat{\beta}_i$ и полагая затраты на рекламу $a = 280$ руб., получаем оценку оптимальной цены:

$$\widehat{p}_{\text{opt}} = \frac{2(-110.153) - 957.335 - 255.529 \cdot 2.8 - (-43.376)2.8^2}{2(-110.153)} = 7.05 \text{ руб.}$$

д) Аналогично предыдущему получаем:

$$\frac{\partial(profit)}{\partial a} = p(\beta_3 + 2\beta_4\, a) - 2(\beta_3 + 2\beta_4\, a) - 100.$$

Приравнивая эту производную нулю и учитывая, что $\beta_4 < 0$, получаем следующее значение расходов на рекламу, при которых максимизируется средний доход при фиксированной цене чебурека $p$:

$$a_{\text{opt}} = \frac{100 + 2\beta_3 - p\beta_3}{2\beta_4(p-2)}.$$

Подставляя вместо коэффициентов $\beta_i$ их оценки $\widehat{\beta}_i$ и полагая $p = 6$ руб., получаем оценку оптимальных расходов на рекламу:

$$\widehat{a}_{\text{opt}} = \frac{100 + 2 \cdot 255.529 - 6 \cdot 255.529}{2(-43.376)(6-2)} = 266 \text{ руб.}$$

е) Рассмотрим чистый доход как функцию двух переменных: $profit = h(p, a)$. С помощью непосредственных вычислений получаем:

$$\frac{\partial h}{\partial p} = \beta_1 + 2\beta_2 p + \beta_3 a + \beta_4 a^2 - 2\beta_2,$$
$$\frac{\partial h}{\partial a} = \beta_3 p + 2\beta_4 ap - 2\beta_3 - 4\beta_4 a - 100.$$

Приравнивая эти производные нулю, получаем систему уравнений относительно $p$ и $a$:

$$\begin{cases} \beta_1 + 2\beta_2 p + \beta_3 a + \beta_4 a^2 - 2\beta_2 = 0 \\ \beta_3 p + 2\beta_4 ap - 2\beta_3 - 4\beta_4 a - 100 = 0. \end{cases} \qquad (*)$$

Выражая из первого уравнения $p$ через $a$ и подставляя это выражение во второе уравнение, получаем относительно $a$ кубическое уравнение

$$2\beta_4^2 a^3 + 3\beta_3\beta_4 a^2 + (\beta_3^2 + 2\beta_1\beta_4 + 4\beta_1\beta_4)a + \beta_1\beta_3 + 2\beta_2\beta_3 + 200\beta_2 = 0.$$

Заменяя в этом уравнении коэффициенты $\beta_i$ их оценками $\widehat{\beta}_1 = 957.335$, $\widehat{\beta}_2 = -110.153$, $\widehat{\beta}_3 = 255.529$, $\widehat{\beta}_4 = -43.376$ и решая уравнение с помощью какого-либо численного метода (например, используя встроенную процедуру Solver программы Microsoft Excel), получаем три решения системы (*):

$$\begin{cases} \widehat{a}_1 = 2.72 \\ \widehat{p}_1 = 7.04 \\ \widehat{q}_1 = 556 \end{cases}, \quad \begin{cases} \widehat{a}_2 = 8.12 \\ \widehat{p}_2 = 1.78 \\ \widehat{q}_2 = -24.5 \end{cases}, \quad \begin{cases} \widehat{a}_3 = -2.00 \\ \widehat{p}_3 = 2.32 \\ \widehat{q}_3 = 17.2 \end{cases}.$$

Ясно, что второе и третье решения не имеют экономического смысла.

Покажем, что точка $(\widehat{p}_1, \widehat{a}_1)$ является точкой максимума функции $h(p, a)$. Вычислим вторые производные функции $h$:

$$\frac{\partial^2 h}{\partial p^2} = 2\beta_2, \quad \frac{\partial^2 h}{\partial a^2} = 2\beta_4 p - 4\beta_4, \quad \frac{\partial^2 h}{\partial p \partial a} = \beta_3 + 2a\beta_4.$$

Для оценки матрицы вторых производных $\boldsymbol{D}^2$ и ее определителя в точке $(\widehat{p}_1, \widehat{a}_1)$ получаем следующее выражение:

$$\boldsymbol{D}^2(\widehat{p}_1, \widehat{a}_1) = \begin{bmatrix} -220.31 & 19.83 \\ 19.83 & -437.53 \end{bmatrix}, \quad \det \boldsymbol{D}^2(\widehat{p}_1, \widehat{a}_1) = 95996,$$

откуда сразу вытекает, что матрица $\boldsymbol{D}^2(\widehat{a}_1, \widehat{p}_1)$ отрицательно определена и, следовательно, точка $(\widehat{a}_1, \widehat{p}_1)$ — точка максимума функции $h$.

Окончательно получаем, что при цене одного чебурека 7.04 руб. и расходах на рекламу 272 руб. средняя величина чистой прибыли максимальна.

ж) Из результатов оценивания уравнения для количества проданных чебуреков (см. п. б)) следует, что все коэффициенты значимы на 5%-ном уровне, поскольку для каждой оценки соответствующее $P$-значение (Probability) меньше 0.05.

Число степеней свободы модели есть $n - k = 20 - 4 = 16$, таким образом, доверительные интервалы для коэффициентов $\beta_i$ суть $[\widehat{\beta}_i - t_c \widehat{\sigma}_{\widehat{\beta}_i}, \widehat{\beta}_i + t_c \widehat{\sigma}_{\widehat{\beta}_i}]$, $i = 2, 3, 4$, где, как обычно, $\widehat{\sigma}_{\widehat{\beta}_i}$ — оценки стандартных ошибок (Std. Error), а $t_c = t_{0.025}(16) = 2.12$. Используя результаты п. б), получаем:

$$\beta_2 \in (-155.37;\ -64.94),$$
$$\beta_3 \in (123.62;\ 387.44),$$
$$\beta_4 \in (-60.38;\ -26.37).$$

## Задача 3.12

В кейнсианской теории спрос на деньги зависит от доходов и процентных ставок. Рассмотрим следующую модель:

$$m_t = \beta_1 + \beta_2 y_t + \beta_3 i_t + \varepsilon_t, \qquad (*)$$

где $m_t$ — агрегат денежной массы M1 (млрд. долл.), $y_t$ — валовой национальный продукт (ВНП) (млрд. долл.), $i_t$ — процентные ставки по 6-месячным государственным облигациям США (6-*month US Treasury Bills*, %). В таблице 3.2 представлены данные по этим переменным за период 1960–1983 гг. по экономике США.

**Таблица 3.2**

| Год | $y_t$ | $m_t$ | $i_t$ | Год | $y_t$ | $m_t$ | $i_t$ |
|---|---|---|---|---|---|---|---|
| 1960 | 506.5 | 141.8 | 3.247 | 1972 | 1185.9 | 251.9 | 4.466 |
| 1961 | 524.6 | 146.5 | 2.605 | 1973 | 1326.4 | 265.8 | 7.178 |
| 1962 | 565.0 | 149.2 | 2.908 | 1974 | 1434.2 | 277.5 | 7.926 |
| 1963 | 596.7 | 154.7 | 3.253 | 1975 | 1549.2 | 291.1 | 6.122 |
| 1964 | 637.7 | 161.8 | 3.686 | 1976 | 1718.0 | 310.4 | 5.266 |
| 1965 | 691.1 | 169.5 | 4.055 | 1977 | 1918.3 | 335.5 | 5.510 |
| 1966 | 756.0 | 173.7 | 5.082 | 1978 | 2163.9 | 363.2 | 7.572 |
| 1967 | 799.6 | 185.1 | 4.630 | 1979 | 2417.8 | 389.0 | 10.017 |
| 1968 | 873.4 | 199.4 | 5.470 | 1980 | 2631.7 | 414.1 | 11.374 |
| 1969 | 944.0 | 205.8 | 6.853 | 1981 | 2954.1 | 440.6 | 13.776 |
| 1970 | 992.7 | 216.5 | 6.562 | 1982 | 3073.0 | 478.2 | 11.084 |
| 1971 | 1077.6 | 230.7 | 4.511 | 1983 | 3309.5 | 521.1 | 8.750 |

Источник: *Economic Report of the President*, Department of Commerce, Bureau of Economic Analysis.

а) Найдите оценки коэффициентов регрессии (∗). Интерпретируйте знаки коэффициентов.

б) Рассчитайте прогноз спроса на деньги при значениях:

$$(1)\ y = 1000,\ i = 10 \quad \text{и} \quad (2)\ y = 2500,\ i = 5.$$

в) Рассчитайте эластичность спроса на деньги $m$ по доходам $y$ и по процентным ставкам $(\partial \ln m/\partial \ln y,\ \partial \ln m/\partial \ln i)$ в двух точках (1) и (2) из б). Сравните результаты.

г) Рассмотрим модель

$$\ln m_t = \beta_1 + \beta_2 \ln y_t + \beta_3 \ln i_t + \varepsilon_t. \qquad (**)$$

Повторите б) и в) и сравните результаты, полученные по разным моделям. Сравните модели (∗) и (∗∗). Какая из них вам представляется более предпочтительной?

## Решение

а) Оценка параметров модели (∗) с применением эконометрического пакета дает следующий результат.

Dependent Variable: $m$

| Variable | Coefficient | Std. Error | $t$-Statistic | Probability |
|---|---|---|---|---|
| const | 89.777 | 4.103 | 21.877 | 0.0000 |
| $y$ | 0.136 | 0.004 | 34.333 | 0.0000 |
| $i$ | −2.577 | 1.189 | −2.167 | 0.0419 |
| $R^2$ | 0.9953 | | | |

Все коэффициенты значимы на 5%-ном уровне. Знаки оценок согласуются с экономической теорией и здравым смыслом. Неравенство $\beta_2 > 0$ означает, что увеличение ВНП при фиксированных процентных ставках приводит в среднем к возрастанию денежной массы M1. В то же время неравенство $\beta_3 < 0$ означает, что при увеличении процентных ставок по государственным ценным бумагам спрос на них возрастает, и это приводит к «связыванию» наличности и уменьшению денежной массы M1. Свободный член показывает денежную массу M1 при $i = y = 0$. Однако такие значения ВНП и процентных ставок лежат вдалеке от наших данных. Поэтому величина свободного члена не имеет хорошей интерпретации и не так важна. Эластичность денежной массы по ВНП положительна и выше в точке (2); эластичность денежной массы по процентным ставкам отрицательна и выше по абсолютной величине в точке (1).

б) В соответствии с моделью $(*)$ $\widehat{m} = \widehat{\beta}_1 + \widehat{\beta}_2 y + \widehat{\beta}_3 i$. Используя результаты п. а), получаем:

$$\widehat{m}_{(1)} = \widehat{\beta}_1 + \widehat{\beta}_2 y_{(1)} + \widehat{\beta}_3 i_{(1)} = 89.777 + 0.136 \cdot 1000 + (-2.577) \cdot 10 = 199.95,$$

$$\widehat{m}_{(2)} = \widehat{\beta}_1 + \widehat{\beta}_2 y_{(2)} + \widehat{\beta}_3 i_{(2)} = 89.777 + 0.136 \cdot 2500 + (-2.577) \cdot 5 = 416.74.$$

в) По определению

$$e_{my} = \frac{\partial(\ln m)}{\partial(\ln y)} = \frac{\partial m}{\partial y}\frac{y}{m} = \beta_2 \frac{y}{m},$$

$$e_{mi} = \frac{\partial(\ln m)}{\partial(\ln i)} = \frac{\partial m}{\partial i}\frac{i}{m} = \beta_3 \frac{i}{m}.$$

Поэтому

$$\widehat{e}_{my}^{(1)} = 0.136(1000/199.95) = 0.680,$$

$$\widehat{e}_{mi}^{(1)} = (-2.577)(10/199.95) = -0.129,$$

$$\widehat{e}_{my}^{(2)} = 0.136(2500/416.74) = 0.815,$$

$$\widehat{e}_{mi}^{(2)} = (-2.577)(5/416.74) = -0.031.$$

г) Оценка параметров модели $(**)$ с применением эконометрического пакета дает следующий результат.

Dependent Variable: $\ln m$

| Variable | Coefficient | Std. Error | $t$-Statistic | Probability |
|---|---|---|---|---|
| const | 0.575 | 0.082 | 7.032 | 0.0000 |
| $\ln y$ | 0.709 | 0.016 | 44.406 | 0.0000 |
| $\ln i$ | −0.053 | 0.021 | −2.492 | 0.0211 |
| $R^2$ | 0.9977 | | | |

Теперь в соответствии с моделью (**) $\widehat{m} = e^{\widehat{\beta}_1} y^{\widehat{\beta}_2} i^{\widehat{\beta}_3}$. Поэтому

$$\widehat{m}_{(1)} = e^{0.575} \cdot 1000^{0.709} \cdot 10^{-0.053} = 210.71,$$
$$\widehat{m}_{(2)} = e^{0.575} \cdot 2500^{0.709} \cdot 5^{-0.053} = 418.73.$$

Из (**) следует, что эластичности спроса на деньги по $y$ и $i$ постоянны и равны соответствующим коэффициентам $\beta_2$ и $\beta_3$. Таким образом,

$$\widehat{e}_{my}^{(1)} = \widehat{e}_{my}^{(2)} = 0.709,$$
$$\widehat{e}_{mi}^{(1)} = \widehat{e}_{mi}^{(2)} = -0.053.$$

Сравнивая модели (*) и (**) с точки зрения качества соответствующих регрессий, видим, что во второй значения $t$-статистик выше, чем в первой, однако существенной разницы в значимости коэффициентов нет: в обеих регрессиях все коэффициенты статистически значимы на 5%-ном уровне. Коэффициенты детерминации в обеих моделях близки к 1, однако по этому показателю их сравнивать нельзя, так как у них разные зависимые переменные. Кроме того, для регрессий одних макроэкономических временных рядов на другие вообще характерны высокие значения коэффициентов $R^2$ из-за наличия общего временного тренда. С точки зрения экономической теории модель (**) выглядит более предпочтительной, поскольку в широком диапазоне значений агрегат денежной массы M1 проявляет постоянную эластичность по ВНП и процентной ставке, что говорит в пользу логарифмической модели (**).

## Задача 3.13

Рассмотрим классическую модель линейной регрессии $\boldsymbol{y} = \boldsymbol{X\beta} + \boldsymbol{\varepsilon}$ с ограничением $\boldsymbol{H\beta} = \boldsymbol{r}$ на вектор коэффициентов.

а) Покажите, что оценка метода наименьших квадратов при наличии ограничения, $\widehat{\boldsymbol{\beta}}_{\mathrm{R}}$, получающаяся из решения соответствующей задачи минимизации, следующим образом выражается через обычную оценку метода наименьших квадратов без учета ограничения, $\widehat{\boldsymbol{\beta}}_{\mathrm{UR}}$:

$$\widehat{\boldsymbol{\beta}}_{\mathrm{R}} = \widehat{\boldsymbol{\beta}}_{\mathrm{UR}} + (\boldsymbol{X}'\boldsymbol{X})^{-1}\boldsymbol{H}'\left(\boldsymbol{H}(\boldsymbol{X}'\boldsymbol{X})^{-1}\boldsymbol{H}'\right)^{-1}(\boldsymbol{r} - \boldsymbol{H}\widehat{\boldsymbol{\beta}}_{\mathrm{UR}}).$$

б) Покажите, что

$$(H\widehat{\beta}_{\text{UR}} - r)' \left(H(X'X)^{-1}H'\right)^{-1} (H\widehat{\beta}_{\text{UR}} - r) = e_{\text{R}}'e_{\text{R}} - e_{\text{UR}}'e_{\text{UR}},$$

где $e_{\text{R}} = y - X\widehat{\beta}_{\text{R}}$, $e_{\text{UR}} = y - X\widehat{\beta}_{\text{UR}}$ — векторы остатков в регрессиях с ограничениями и без ограничений соответственно.

## Решение

а) Мы предполагаем, что $q \times k$ $(q \leqslant k)$ матрица $H$ имеет полный ранг $q$.

Оценка $\widehat{\beta}_{\text{R}}$ является по определению решением задачи

$$\begin{aligned}(y - X\beta)'(y - X\beta) &\to \min,\\ H\beta &= r.\end{aligned}$$

Построим функцию Лагранжа

$$L(\beta, l) = (y - X\beta)'(y - X\beta) - 2l'(H\beta - r),$$

где $l = (l_1, \ldots, l_q)'$ — вектор множителей Лагранжа, и запишем необходимое условие минимума:

$$\begin{aligned}\frac{\partial L}{\partial \beta} &= -2X'y + 2X'X\beta - 2H'l = 0,\\ H\beta &= r.\end{aligned}$$

Из первого уравнения следует, что

$$(X'X)\beta = X'y + H'l,$$

откуда вытекает, что

$$\beta = (X'X)^{-1}X'y + (X'X)^{-1}H'l = \widehat{\beta}_{\text{UR}} + (X'X)^{-1}H'l. \qquad (*)$$

Умножая последнее равенство слева на матрицу $H$ и учитывая ограничение, получаем:

$$r = H\widehat{\beta}_{\text{UR}} + H(X'X)^{-1}H'l,$$

откуда

$$l = \left(H(X'X)^{-1}H'\right)^{-1}(r - H\widehat{\beta}_{\text{UR}}).$$

Подставляя это выражение для $l$ в $(*)$, получаем:

$$\beta = \widehat{\beta}_{\text{R}} = \widehat{\beta}_{\text{UR}} + (X'X)^{-1}H'\left(H(X'X)^{-1}H'\right)^{-1}(r - H\widehat{\beta}_{\text{UR}}), \qquad (**)$$

что и требуется.

б) Из соотношения (**) следует, что

$$e_{\mathrm{R}} = y - X\widehat{\beta}_{\mathrm{R}} = e_{\mathrm{UR}} - X(X'X)^{-1}H'\left(H(X'X)^{-1}H'\right)^{-1}(r - H\widehat{\beta}_{\mathrm{UR}})$$
$$= e_{\mathrm{UR}} - m,$$

где $m = X(X'X)^{-1}H'\left(H(X'X)^{-1}H'\right)^{-1}(r - H\widehat{\beta}_{\mathrm{UR}})$. Тогда

$$e'_{\mathrm{R}}e_{\mathrm{R}} = e'_{\mathrm{UR}}e_{\mathrm{UR}} - 2e'_{\mathrm{UR}}m + m'm.$$

По свойствам МНК вектор $e_{\mathrm{UR}}$ ортогонален каждому столбцу матрицы $X$, следовательно, $e'_{\mathrm{UR}}m = 0$. Таким образом,

$$\begin{aligned} e'_{\mathrm{R}}e_{\mathrm{R}} - e'_{\mathrm{UR}}e_{\mathrm{UR}} &= m'm \\ &= (r - H\widehat{\beta}_{\mathrm{UR}})'\left(H(X'X)^{-1}H'\right)^{-1}H(X'X)^{-1}X'X \\ &\quad\times (X'X)^{-1}H'\left(H(X'X)^{-1}H'\right)^{-1}(r - H\widehat{\beta}_{\mathrm{UR}}) \\ &= (r - H\widehat{\beta}_{\mathrm{UR}})'\left(H(X'X)^{-1}H'\right)^{-1}(r - H\widehat{\beta}_{\mathrm{UR}}), \end{aligned}$$

что и требуется.

## Задача 3.14

Оценивание модели $y_t = \beta_1 + \beta_2 x_{t2} + \beta_3 x_{t3} + \beta_4 x_{t4} + \varepsilon_t$ методом наименьших квадратов по 26 наблюдениям дало следующие результаты:

$$y_t = 2 + \underset{(1.9)}{3.5}\,x_{t2} - \underset{(-2.2)}{0.7}\,x_{t3} + \underset{(1.5)}{2.0}\,x_{t4} + e_t, \qquad R^2 = 0.882$$

(в скобках даны значения $t$-статистик).

Оценивание той же модели при ограничении $\beta_2 = \beta_4$ дало следующие результаты:

$$y_t = 1.5 + \underset{(2.7)}{3.0}\,(x_{t2} + x_{t4}) - \underset{(-2.4)}{0.6}\,x_{t3} + u_t, \qquad R^2 = 0.876.$$

а) Проверьте значимость вектора $\beta' = (\beta_2, \beta_3, \beta_4)$ в регрессии без ограничений.

б) Проверьте ограничение $\beta_2 = \beta_4$.

## Решение

а) Проверка значимости вектора $\beta' = (\beta_2, \beta_3, \beta_4)$ — это проверка гипотезы $\mathrm{H}_0\colon \beta_2 = 0,\ \beta_3 = 0,\ \beta_4 = 0$. Для ее тестирования воспользуемся статистикой (3.36)

$$F = \frac{R^2}{1 - R^2} \cdot \frac{n-k}{k-1} = \frac{0.882}{1 - 0.882} \cdot \frac{26-4}{4-1} = 54.81.$$

Если верна гипотеза $\mathrm{H}_0$, то статистика $F$ имеет распределение Фишера $F(k-1, n-k)$. Из таблиц находим $F_{0.05}(3,22) = 3.05$. Так как $F > 3.05$, то гипотеза $\mathrm{H}_0$ отвергается на 5%-ном уровне значимости.

б) Вторая регрессия представляет собой исходную регрессию с ограничением $\beta_2 = \beta_4$, поэтому для тестирования гипотезы $\mathrm{H}_0$: $\beta_2 = \beta_4$ можно воспользоваться статистикой (3.45):

$$F = \frac{(R^2_{\mathrm{UR}} - R^2_{\mathrm{R}})/q}{(1 - R^2_{\mathrm{UR}})/(n-k)},$$

где $R^2_{\mathrm{UR}} = 0.882$, $R^2_{\mathrm{R}} = 0.876$ — коэффициенты детерминации регрессий без ограничения и с ограничением соответственно, $n-k = 22$, $q = 1$. Имеем

$$F = \frac{0.882 - 0.876}{1 - 0.882} \cdot \frac{22}{1} = 1.12.$$

Если гипотеза $\mathrm{H}_0$ верна, то статистика $F$ имеет распределение Фишера $F(q, n-k)$. Из таблиц находим $F_{0.05}(1,22) = 4.30$. Так как $F < 4.30$, то гипотеза $\mathrm{H}_0$ не отвергается на 5%-ном уровне значимости.

## Задача 3.15

В таблице 3.3 представлены реальный доход на душу населения $y$ (тыс. долл.), процент рабочей силы, занятой в сельском хозяйстве, $x_1$ и средний уровень образования населения в возрасте после 25 лет $x_2$ (число лет, проведенных в учебных заведениях) для 15 развитых стран в 1983 г.

**Таблица 3.3**

| Страна | $y$ | $x_1$ | $x_2$ | Страна | $y$ | $x_1$ | $x_2$ |
|---|---|---|---|---|---|---|---|
| 1 | 7 | 8 | 9 | 9 | 10 | 6 | 12 |
| 2 | 9 | 9 | 13 | 10 | 11 | 7 | 14 |
| 3 | 9 | 7 | 11 | 11 | 11 | 6 | 11 |
| 4 | 8 | 6 | 11 | 12 | 12 | 4 | 15 |
| 5 | 8 | 10 | 12 | 13 | 9 | 8 | 15 |
| 6 | 14 | 4 | 16 | 14 | 10 | 5 | 10 |
| 7 | 9 | 5 | 11 | 15 | 12 | 8 | 13 |
| 8 | 8 | 5 | 11 | | | | |

а) Проведите множественную регрессию $y$ на константу, $x_1$ и $x_2$ и проинтерпретируйте полученные результаты.

б) Определите $s^2$, $s^2_{\widehat{\beta}_1}$ и $s^2_{\widehat{\beta}_2}$.

в) Почему, как правило, константа $\beta_0$ не играет существенной роли при рассмотрении регрессии?

г) Постройте 95%-ные доверительные интервалы для коэффициентов $\beta_1, \beta_2$ и вычислите коэффициент детерминации $R^2$ и скорректированный коэффициент детерминации $R^2_{\text{adj}}$.

д) Проверьте на 5%-ном уровне значимость коэффициентов $\beta_1, \beta_2$.

## Решение

а) Оценивание множественной регрессии переменной $y$ на константу и переменные $x_1$, $x_2$ с применением эконометрического пакета дает следующий результат.

Dependent Variable: $y$

| Variable | Coefficient | Std. Error | $t$-Statistic | Probability |
|---|---|---|---|---|
| const | 5.084 | 2.546 | 1.997 | 0.0690 |
| $x_1$ | −0.409 | 0.188 | −2.178 | 0.0501 |
| $x_2$ | 0.602 | 0.168 | 3.579 | 0.0038 |
| $R^2$ | 0.6216 | | | |

Коэффициент $\beta_1$ при переменной $x_1$ (процент рабочей силы, занятой в сельском хозяйстве) значим практически на 5%-ном уровне, а коэффициент $\beta_2$ при переменной $x_2$ (уровень образования) значим даже на 0.5%-ном уровне. Среднедушевой доход ниже для более «аграрных» и выше для более «образованных» стран. При этом уровень образования оказывает большее влияние на среднедушевой доход и более значим, чем уровень аграрности.

б) С помощью непосредственных вычислений получаем:

$$e'e = 19.07, \quad s^2 = \frac{e'e}{n-k} = \frac{19.07}{15-3} = 1.59,$$

$$X'X = \begin{bmatrix} 15 & 98 & 184 \\ 98 & 686 & 1196 \\ 184 & 1196 & 2314 \end{bmatrix}, \quad (X'X)^{-1} = \begin{bmatrix} 4.078 & -0.174 & -0.234 \\ -0.174 & 0.022 & -0.002 \\ -0.234 & -0.002 & 0.018 \end{bmatrix}.$$

Отсюда

$$s^2_{\hat{\beta}_1} = 1.59 \cdot 0.022 = 0.035, \quad s^2_{\hat{\beta}_2} = 1.59 \cdot 0.018 = 0.028$$

(ср. с результатами п. а)).

в) В регрессионных моделях константа, как правило, играет вспомогательную роль. Ее вводят в регрессии для гибкости модели, величина ее оценки обычно не имеет реального экономического смысла. Например, если рассматривается логарифмическая модель, то величина константы зависит от основания логарифма, в то время как оценки коэффициентов при других независимых переменных не изменяются при изменении основания логарифмов.

г) Имеем
$$\beta_i = \widehat{\beta}_i \pm t_c s_{\widehat{\beta}_i}, \quad i = 1, 2,$$
где $t_c = t_{0.025}(n-k) = t_{0.025}(12) = 2.179$. Таким образом,
$$\beta_1 = -0.409 \pm 2.179 \cdot 0.188, \quad \beta_1 \in [-0.818; 0.000],$$
$$\beta_2 = \phantom{-}0.602 \pm 2.179 \cdot 0.168, \quad \beta_2 \in [\phantom{-}0.236; 0.368].$$

$R^2 = 0.6216$ (см. таблицу к п. а),
$R^2_{\text{adj}} = 1 - (1-R^2)\frac{n-1}{n-k} = 1 - (1 - 0.6216)\frac{15-1}{15-3} = 0.5585.$

д) Как уже отмечалось в п. а), коэффициент $\beta_1$ значим практически на 5%-ном уровне, а коэффициент $\beta_1$ значим даже на 0.5%-ном уровне и, следовательно, значим и на 5%-ном уровне.

## Задача 3.16

Вместо того чтобы оценивать параметры $\boldsymbol{\beta}_1, \boldsymbol{\beta}_2$ в модели
$$\boldsymbol{y} = \boldsymbol{X}_1\boldsymbol{\beta}_1 + \boldsymbol{X}_2\boldsymbol{\beta}_2 + \boldsymbol{\varepsilon}, \tag{$*$}$$
($\boldsymbol{X}_1, \boldsymbol{X}_2$ — $n \times k_1$, $n \times k_2$ матрицы соответственно, $\boldsymbol{\beta}_1, \boldsymbol{\beta}_2$ — векторы размерности $k_1, k_2$ соответственно), строятся МНК-оценки этих параметров исходя из модели
$$\boldsymbol{y} = \boldsymbol{X}_1^*\boldsymbol{\beta}_1 + \boldsymbol{X}_2\boldsymbol{\beta}_2 + \boldsymbol{\varepsilon}^*, \tag{$**$}$$
где $\boldsymbol{X}_1^*$ — матрица остатков, полученных в результате регрессии каждого столбца матрицы $\boldsymbol{X}_1$ на $\boldsymbol{X}_2$.

а) Покажите, что полученная таким образом оценка вектора $\boldsymbol{\beta}_2$ совпадает с оценкой, полученной в результате регрессии $\boldsymbol{y}$ только на $\boldsymbol{X}_2$.

б) Найдите смещение оценки вектора $\boldsymbol{\beta}_2$.

в) Покажите, что МНК-оценки вектора $\boldsymbol{\beta}_1$, построенные по моделям ($*$) и ($**$), совпадают.

## Решение

а) Из определения матрицы $\boldsymbol{X}_1^*$ непосредственно следует, что каждый ее столбец ортогонален каждому столбцу матрицы $\boldsymbol{X}_2$, что эквивалентно равенству $\boldsymbol{X}_2'\boldsymbol{X}_1^* = \mathbf{0}$. Обозначим $\widehat{\boldsymbol{\beta}}^* = (\widehat{\boldsymbol{\beta}}_1^*, \widehat{\boldsymbol{\beta}}_2^*)'$ МНК-оценку вектора $\boldsymbol{\beta} = (\boldsymbol{\beta}_1, \boldsymbol{\beta}_2)'$ в регрессии ($**$). По правилам действия с блочными матрицами получаем:

$$\widehat{\beta}^* = \left(\begin{bmatrix}X_1^* & X_2\end{bmatrix}'\begin{bmatrix}X_1^* & X_2\end{bmatrix}\right)^{-1}\begin{bmatrix}X_1^* & X_2\end{bmatrix}'y$$

$$= \left(\begin{bmatrix}X_1^{*'} \\ X_2'\end{bmatrix}\begin{bmatrix}X_1^* & X_2\end{bmatrix}\right)^{-1}\begin{bmatrix}X_1^{*'} \\ X_2'\end{bmatrix}y = \begin{bmatrix}X_1^{*'}X_1^* & X_1^{*'}X_2 \\ X_2'X_1^* & X_2'X_2\end{bmatrix}^{-1}\begin{bmatrix}X_1^{*'}y \\ X_2'y\end{bmatrix}$$

$$= \begin{bmatrix}X_1^{*'}X_1^* & 0 \\ 0 & X_2'X_2\end{bmatrix}^{-1}\begin{bmatrix}X_1^{*'}y \\ X_2'y\end{bmatrix} = \begin{bmatrix}(X_1^{*'}X_1^*)^{-1}X_1^{*'}y \\ (X_2'X_2)^{-1}X_2'y\end{bmatrix}.$$

Таким образом, $\widehat{\beta}_2^* = (X_2'X_2)^{-1}X_2'y$, что совпадает с МНК-оценкой вектора $\beta_2$ в регрессии $y$ только на $X_2$.

б) Имеем:

$$\widehat{\beta}_2^* = (X_2'X_2)^{-1}X_2'y = (X_2'X_2)^{-1}X_2'(X_1\beta_1 + X_2\beta_2 + \varepsilon)$$
$$= \beta_2 + (X_2'X_2)^{-1}X_2'X_1\beta_1 + (X_2'X_2)^{-1}X_2'\varepsilon.$$

Отсюда

$$\mathrm{E}(\widehat{\beta}_2^*) = \beta_2 + (X_2'X_2)^{-1}X_2'X_1\beta_1,$$

т.е. смещение оценки $\widehat{\beta}_2^*$ равно $(X_2'X_2)^{-1}X_2'X_1\beta_1$.

в) Совпадение МНК-оценок вектора $\beta_1$ в регрессиях (*) и (**) можно установить с помощью непосредственных вычислений, используя правила обращения блочных матриц. Можно, однако, избежать громоздких выкладок, если прибегнуть к геометрической интерпретации метода наименьших квадратов.

Напомним, что МНК-оценки компонент вектора $\beta = (\beta_1, \beta_2)'$ в регрессии (*) получаются как коэффициенты разложения ортогональной проекции $\widehat{y}$ вектора $y$ в евклидовом пространстве $R^n$ на подпространство, порожденное столбцами матрицы $X = [X_1\ X_2]$:

$$\widehat{y} = X_1\widehat{\beta}_1 + X_2\widehat{\beta}_2.$$

Нетрудно установить, что подпространства, порожденные столбцами матриц $[X_1\ X_2]$ и $[X_1^*\ X_2]$, совпадают и, следовательно, совпадают ортогональные проекции на них вектора $y$. Иными словами,

$$X_1\widehat{\beta}_1 + X_2\widehat{\beta}_2 = X_1^*\widehat{\beta}_1^* + X_2\widehat{\beta}_2^*.$$

Далее, так как столбцы матрицы $X_1^*$ представляют собой остатки регрессий столбцов матрицы $X_1$ на матрицу $X_2$, то

$$X_1^* = X_1 - X_2(X_2'X_2)^{-1}X_2'X_1.$$

Поэтому

$$X_1\widehat{\beta}_1 + X_2\widehat{\beta}_2 = X_1\widehat{\beta}_1^* + X_2(\widehat{\beta}_2^* - (X_2'X_2)^{-1}X_2'X_1\widehat{\beta}_1^*) = X_1\widehat{\beta}_1^* + X_2\widehat{\gamma}^*,$$

где $\widehat{\gamma}^* = \widehat{\beta}_2^* - (X_2'X_2)^{-1}X_2'X_1\widehat{\beta}_1^*$. Так как матрица $X$ имеет по предположению полный ранг, то ее столбцы линейно независимы, и поэтому из последнего соотношения вытекают следующие равенства:

$$\widehat{\beta}_1 = \widehat{\beta}_1^*,$$
$$\widehat{\beta}_2 = \widehat{\gamma}^* = \widehat{\beta}_2^* - (X_2'X_2)^{-1}X_2'X_1\widehat{\beta}_1^*,$$

что и требуется.

Заметим также, что из этих равенств следует, что

$$\widehat{\beta}_2^* = \widehat{\beta}_2 + (X_2'X_2)^{-1}X_2'X_1\widehat{\beta}_1,$$

что согласуется с п. б).

## Задача 3.17

Строится регрессия $n \times 1$ вектора $y$ на $n \times k$ матрицу регрессоров $X$ и вычисляется коэффициент детерминации $R_1^2$. Затем к матрице $X$ добавляется дополнительный $(k+1)$-й столбец, проводится регрессия $y$ на новую матрицу и вычисляется коэффициент детерминации $R_2^2$. При каких условиях $R_1^2 = R_2^2$?

## Решение

Обозначим через $X_1$ старую матрицу $X_1 = X$ и пусть $X_2 = [X\ z]$ — дополненная матрица регрессоров $X$. $n \times 1$ вектор $z$ является дополнительным $(k+1)$-м столбцом в матрице $X_2$. Новое уравнение регрессии может быть записано в виде:

$$y = X\beta + z\gamma + \varepsilon.$$

По определению (см. (3.27))

$$R_i^2 = 1 - \frac{e_i'e_i}{y'y}, \quad i = 1, 2,$$

где $e_i$, остатки соответствующих регрессий, можно представить в таком виде (см. п. 3.3): $e_i = M_i y$, $M_i = I - X_i(X_i'X_i)^{-1}X_i'$. Условие $R_1^2 = R_2^2$ эквивалентно условию $e_1'e_1 = e_2'e_2$ или $y'(M_1'M_1 - M_2'M_2)y = 0$. Поскольку матрицы $M_i$ идемпотентные, это условие эквивалентно следующему: $y'(M_1 - M_2)y = 0$. Поскольку (см. (3.11)) $I - M_1 = X(X'X)^{-1}X'$, то

$$I - M_2 = X_2(X_2'X_2)^{-1}X_2' = [X\quad z]\begin{bmatrix} X'X & X'z \\ z'X & z'z \end{bmatrix}^{-1}\begin{bmatrix} X' \\ z' \end{bmatrix}.$$

Матрица, обратная матрице $X_2'X_2$, вычисляется по следующей формуле (см. (ЛА.17)):

$$(X_2'X_2)^{-1} = \begin{bmatrix} Q(I + X'z(z'M_1z)^{-1}z'XQ) & -QX'z(z'M_1z)^{-1} \\ -(z'M_1z)^{-1}z'XQ & (z'M_1z)^{-1} \end{bmatrix},$$

где $Q = (X'X)^{-1}$. Отсюда получаем:

$$I - M_2 = I - M_1 + M_1 z(z'M_1 z)^{-1} z'M_1,$$

следовательно,

$$y'(M_1 - M_2)y = y'M_1 z(z'M_1 z)^{-1} z'M_1 y = \frac{(z'M_1 y)^2}{z'M_1 z}.$$

Поскольку $\widehat{\gamma} = (z'M_1 z)^{-1} z'M_1 y$, то условие $y'(M_1 - M_2)y = 0$ выполняется тогда и только тогда, когда $\widehat{\gamma} = 0$ в новом (дополненном) уравнении регрессии. Или, другими словами, вектор $z$ ортогонален вектору остатков первой регрессии $e_1$.

Этот результат можно интерпретировать следующим образом. Новый регрессор $z$ не дает вклад в прогноз $y$ и, следовательно, не изменяет значение $R^2$.

## Задача 3.18

Рассматривается стандартная линейная модель множественной регрессии $y = X\beta + \varepsilon$, где $X$ — $n \times k$ матрица ранга $k$.

а) Пусть $G$ — $k \times m$ матрица, имеющая ранг $m < k$, и пусть множество $L = \{\beta\colon \beta = G\gamma$ для некоторого $\gamma\}$. Постройте тест для проверки гипотезы $\mathrm{H}_0 : \beta \in L$ против альтернативы $\mathrm{H}_1 : \beta \notin L$.

б) Пусть матрица $X$ разбита на две матрицы: $X = [X_1 \ X_2]$, где $X_1$ — $n \times k_1$ матрица, $X_2$ — $n \times k_2$ матрица, и пусть $q_1 = X_1 r_1$, $q_2 = X_2 r_2$, где $r_1, r_2$ — известные векторы. Рассматривается новая модель $y = \alpha_1 q_1 + \alpha_2 q_2$, где $\alpha_1, \alpha_2$ — скалярные параметры.

Каким образом, используя результаты а), можно проверить, является ли новая модель приемлемой?

## Решение

а) Если $\beta \in L$, то существует вектор $\gamma$ такой, что $\beta = G\gamma$. Рассмотрим систему линейных уравнений $\beta = G\gamma$ относительно неизвестного вектора $\gamma$. Необходимым и достаточным условием существования решения этой системы является следующее условие:

$$G(G'G)^{-1}G'\beta = \beta,$$

поскольку столбцы матрицы $G$ линейно независимы (ранг матрицы равен $m$). Докажем это утверждение.

1) Пусть $G\gamma = \beta$, тогда $G'G\gamma = G'\beta$. Матрица $G'G$ невырожденная, поэтому $\gamma = (G'G)^{-1}G'\beta$. Домножив на $G$, получаем $\beta = G\gamma = G(G'G)^{-1}G'\beta$.

2) Докажем обратное. Пусть $G(G'G)^{-1}G'\beta = \beta$. Очевидно, что в качестве $\gamma$ можно взять $\gamma = (G'G)^{-1}G'\beta$.

Полученное условие представимости $\beta$ в виде $\beta = G\gamma$ можно записать в следующем виде:

$$M_G\beta = 0, \qquad M_G = I_k - G(G'G)^{-1}G'.$$

Матрица $M_G$ является идемпотентной $k \times k$ матрицей ранга $k - m$. Поэтому ее можно представить в виде произведения $M_G = R'R$, где столбцами $k \times (k-m)$ матрицы $R'$ являются собственные векторы матрицы $M_G$, соответствующие собственному числу, равному единице. Поскольку ранг матрицы $R$ равен $k - m$, условие $R'R\beta = 0$ эквивалентно условию $R\beta = 0$. И таким образом мы приходим к проблеме тестирования системы линейных ограничений на коэффициенты регрессии. Для этой задачи мы можем применить, например, $F$-тест (3.45).

б) Запишем правую часть уравнения новой модели в следующем виде:

$$\begin{aligned}\alpha_1 q_1 + \alpha_2 q_2 &= \alpha_1 X_1 r_1 + \alpha_2 X_2 r_2 \\ &= \begin{bmatrix} X_1 & X_2 \end{bmatrix} \begin{bmatrix} \alpha_1 r_1 \\ \alpha_2 r_2 \end{bmatrix} = \begin{bmatrix} X_1 & X_2 \end{bmatrix} \begin{bmatrix} r_1 & 0 \\ 0 & r_2 \end{bmatrix} \begin{bmatrix} \alpha_1 \\ \alpha_2 \end{bmatrix}.\end{aligned}$$

Введем следующие обозначения:

$$G = \begin{bmatrix} r_1 & 0 \\ 0 & r_2 \end{bmatrix}, \qquad \gamma = \begin{bmatrix} \alpha_1 \\ \alpha_2 \end{bmatrix}.$$

В этих обозначениях новая модель верна в том случае, если $G\gamma = \beta$, и можно применить результат пункта а) для проверки справедливости этой гипотезы.

## Задача 3.19

Покажите, что при добавлении в модель регрессора скорректированный коэффициент детерминации $R^2_{\text{adj}}$ увеличивается тогда и только тогда, когда $t$-статистика оценки коэффициента при этом регрессоре по модулю превосходит единицу.

## Решение

а) Квадрат $t$-статистики оценки коэффициента совпадает с $F$-статистикой для проверки гипотезы о равенстве этого коэффициента нулю, которая в свою очередь равна (см. (3.45)):

$$F = \frac{R^2_{\text{UR}} - R^2_{\text{R}}}{(1 - R^2_{\text{UR}})/(n-k)},$$

где $R^2_{\text{UR}}$ и $R^2_{\text{R}}$ — коэффициенты детерминации для «длинной» и «короткой» регрессий соответственно, а $k$ — число регрессоров в «длинной» регрессии.

Скорректированный $R^2$ равен (см. (3.28)):

$$R^2_{\text{adj}} = 1 - \frac{\boldsymbol{e}'\boldsymbol{e}/(n-k)}{\boldsymbol{y}'_*\boldsymbol{y}_*/(n-1)} = 1 - (1-R^2)\frac{n-1}{n-k}.$$

Таким образом, увеличение $R^2_{\text{adj}}$ при добавлении регрессора эквивалентно следующему условию:

$$\begin{aligned}
0 &< \left(1-(1-R^2_{\text{UR}})\frac{n-1}{n-k}\right) - \left(1-(1-R^2_{\text{R}})\frac{n-1}{n-k+1}\right)\\
&= (1-R^2_{\text{R}})\frac{n-1}{n-k+1} - (1-R^2_{\text{UR}})\frac{n-1}{n-k}\\
&= \frac{(n-1)((1-R^2_{\text{R}})(n-k)-(1-R^2_{\text{UR}})(n-k+1))}{(n-k+1)(n-k)}\\
&= \frac{n-1}{(n-k+1)(n-k)}((n-k+1)R^2_{\text{UR}} - (n-k)R^2_{\text{R}} - 1)\\
&= \frac{n-1}{(n-k+1)(n-k)}((n-k)(R^2_{\text{UR}}-R^2_{\text{R}}) - (1-R^2_{\text{UR}}))\\
&= \frac{(n-1)(1-R^2_{\text{UR}})}{(n-k+1)(n-k)}\left(\frac{R^2_{\text{UR}}-R^2_{\text{R}}}{(1-R^2_{\text{UR}})/(n-k)} - 1\right)\\
&= \frac{(n-1)(1-R^2_{\text{UR}})}{(n-k+1)(n-k)}(F-1).
\end{aligned}$$

То есть $F = t^2 > 1$, что эквивалентно тому, что $t$-статистика оценки коэффициента при этом регрессоре по модулю превосходит единицу.

## Задача 3.20

Оценивание четырех регрессионных моделей на основании 40 наблюдений дало следующие результаты:

$$\begin{aligned}
w &= \underset{(5.0)}{20} + \underset{(0.09)}{0.8}\,age + \underset{(1.31)}{3.7}\,edu, && R^2 = 0.40,\\
\ln w &= \underset{(3.0)}{3.2} + \underset{(0.009)}{0.10}\ln age + \underset{(0.03)}{0.19}\ln edu, && R^2 = 0.71,\\
w &= \underset{(0.3)}{20} + \underset{(0.09)}{0.6}\,age + \underset{(0.12)}{0.4}\,exp, && R^2 = 0.59,\\
w &= \underset{(0.4)}{2.05} + \underset{(0.19)}{0.5}\,age + \underset{(0.35)}{0.6}\,edu + \underset{(0.13)}{0.2}\,exp, && R^2 = 0.63
\end{aligned}$$

(в скобках указаны стандартные ошибки), где $w$ — зарплата работника (в рублях в час), $age$ — его возраст (в годах), $edu$ — уровень образования (число лет, проведенных в учебных заведениях), $exp$ — стаж работы.

а) Сравните эти четыре регрессии с точки зрения их качества и прогностической силы.

б) Дайте интерпретацию коэффициентов при переменных $age$ и $\ln age$ в первом и втором уравнениях соответственно.

## Решение

а) Регрессия может считаться качественной только в том случае, когда она не содержит незначимых коэффициентов. В этом смысле четвертая регрессия плоха, так как коэффициенты при регрессорах $edu$ и $exp$ являются незначимыми на 5%-ном уровне ($t$-статистики равны $0.6/0.35 = 1.714$ и $0.2/0.13 = 1.538$ соответственно). У остальных регрессий подобных проблем нет (незначимость константы во второй регрессии не должна вводить в заблуждение).

Сравнивать по прогностической силе вторую регрессию со всеми остальными нельзя, так как различаются зависимые переменные. Тот результат, что четвертая регрессия имеет больший $R^2$ по сравнению с первой и третьей, частично является результатом увеличения числа регрессоров, поэтому для сравнения четвертой регрессии с первой и третьей следует применять, например, $R^2_{\text{adj}} = 1-(1-R^2)(n-1)/(n-k)$. Получаем, что для первой регрессии $R^2_{\text{adj}} = 0.37$, для третьей $R^2_{\text{adj}} = 0.57$, для четвертой $R^2_{\text{adj}} = 0.6$. Таким образом, и по этому критерию четвертая регрессия более информативная, чем первая и третья. При сравнении первой и третьей регрессий обладающей наибольшей прогностической силой следует признать третью.

б) В первой регрессии коэффициент при $age$ означает, что при увеличении возраста на один год при прочих равных условиях зарплата в среднем возрастает на 0.8 руб. в час. Во второй регрессии коэффициент при $\ln age$ показывает, что при увеличении возраста на 1% при прочих равных условиях зарплата в среднем увеличивается на 0.10% (эластичность заработной платы по возрасту).

## Задача 3.21

Рассмотрим три модели

$$y_t = \frac{\alpha}{t} + \varepsilon_t,$$

$$y_t = \alpha g^t + \varepsilon_t,$$

$$y_t = \alpha + \frac{\beta}{t^2} + \varepsilon_t,$$

где $t = 1, \dots, T$, $\varepsilon \sim N(0, \sigma^2 \boldsymbol{I})$. Во втором уравнении $g$ — известная константа.

а) Покажите, что МНК-оценка параметра $\alpha$ в первом уравнении не может быть состоятельна. Верно ли то же самое для второго уравнения?

б) Являются ли состоятельными МНК-оценки параметров $\alpha$ и $\beta$ в третьем уравнении?

*Указание.* $\sum_{t=1}^{\infty} t^{-2} = \pi^2/6, \quad \sum_{t=1}^{\infty} t^{-4} = \pi^4/90.$

## Решение

а) Рассмотрим первую модель. Это регрессия на одну переменную (без константы), поэтому МНК-оценка записывается следующим образом:

$$\widehat{\alpha} = \left(\sum_{t=1}^{T} \frac{y_t}{t}\right) \Bigg/ \left(\sum_{t=1}^{T} \frac{1}{t^2}\right).$$

Ее дисперсия равна

$$\mathrm{V}(\widehat{\alpha}) = \left(\sum_{t=1}^{T} \frac{\mathrm{V}(y_t)}{t^2}\right) \Bigg/ \left(\sum_{t=1}^{T} \frac{1}{t^2}\right)^2 = \sigma^2 \Bigg/ \left(\sum_{t=1}^{T} \frac{1}{t^2}\right).$$

При $T \to \infty$ имеем: $\mathrm{V}(\widehat{\alpha}) \to 6\sigma^2/\pi^2 > 0$. Отсюда делаем вывод, что оценка $\widehat{\alpha}$ не является состоятельной.

Рассмотрим вторую модель. Если $g = 1$, то это регрессия на константу, и МНК-оценка является состоятельной. Если $g \neq 1$, то это регрессия на одну переменную (без константы), поэтому МНК-оценка записывается следующим образом:

$$\widehat{\alpha} = \frac{\sum_{t=1}^{T} g^t y_t}{\sum_{t=1}^{T} g^{2t}} = \frac{\sum_{t=1}^{T} g^t y_t}{(g^{2T+2}-1)/(g^2-1)}.$$

Ее дисперсия равна

$$\mathrm{V}(\widehat{\alpha}) = \frac{\sum_{t=1}^{T} g^{2t}\mathrm{V}(y_t)}{((g^{2T+2}-1)/(g^2-1))^2} = \frac{\sigma^2}{(g^{2T+2}-1)/(g^2-1)}.$$

Если $T \to \infty$, имеем два случая: при $|g| < 1$ дисперсия $\mathrm{V}(\widehat{\alpha}) \to (1-g^2)\sigma^2 > 0$. Отсюда делаем вывод, что оценка $\widehat{\alpha}$ не является состоятельной. Второй случай — $|g| > 1$. Тогда $\mathrm{V}(\widehat{\alpha}) \to 0$, и оценка $\widehat{\alpha}$ (например, в силу неравенства Чебышева) является состоятельной.

б) Рассмотрим третью модель. Это парная регрессия, так что МНК-оценки параметров равны

$$\widehat{\beta} = \frac{\sum_{t=1}^{T}(x_t - \overline{x})(y_t - \overline{y})}{\sum_{t=1}^{T}(x_t - \overline{x})^2}, \quad \widehat{\alpha} = \overline{y} - \widehat{\beta}\overline{x},$$

где $x_t = 1/t^2$. Обе оценки несмещенные, их дисперсии равны (см. (2.11) и (2.13)):

$$\mathrm{V}(\widehat{\beta}) = \frac{\sigma^2}{\sum_{t=1}^{T}(x_t - \overline{x})^2}, \quad \mathrm{V}(\widehat{\alpha}) = \frac{\sigma^2 \sum_{t=1}^{T} x^2}{T\sum_{t=1}^{T}(x_t - \overline{x})^2}.$$

Рассмотрим $\mathrm{V}(\widehat{\beta})$. Заметим, что $\sum_{t=1}^{T}(x_t-\overline{x})^2=\sum_{t=1}^{T}x_t^2-T\overline{x}^2$. Пр $T\to\infty$ имеем:

$$\lim_{T\to\infty}\sum_{t=1}^{T}x_t^2=\sum_{t=1}^{T}\frac{1}{t^4}=\frac{\pi^4}{90},$$

$$T\overline{x}^2=\frac{1}{T}\left(\sum_{t=1}^{T}x\right)^2=\frac{1}{T}\left(\sum_{t=1}^{T}\frac{1}{t^2}\right)^2\leqslant\frac{1}{T}\left(\frac{\pi^2}{6}\right)^2\to 0.$$

Следовательно,

$$\lim_{T\to\infty}\sum_{t=1}^{T}(x_t-\overline{x})^2=\frac{\pi^4}{90}.$$

Таким образом, $\mathrm{V}(\widehat{\beta})$ не стремится к нулю при $T\to\infty$. Отсюда делаем вывод, что оценка $\widehat{\beta}$ не является состоятельной.

Для оценки $\widehat{\alpha}$, использовав предыдущие выкладки, получаем, что

$$\lim_{T\to\infty}\mathrm{V}(\widehat{\alpha})=\lim_{T\to\infty}\frac{\sigma^2\sum_{t=1}^{T}x^2}{T\sum_{t=1}^{T}(x_t-\overline{x})^2}=0.$$

Отсюда следует, что оценка $\widehat{\alpha}$ состоятельная.

*Замечание.* Из того, что $\mathrm{V}(\widehat{\alpha})\not\to 0$, вообще говоря, не следует, что оценка несостоятельная. Более строгое рассуждение требует, однако, некоторой техники.

Например, для первого уравнения $\widehat{\alpha}$ можно представить в виде:

$$\widehat{\alpha}=\left(\sum_{t=1}^{T}\frac{1}{t}\left(\frac{\alpha}{t}+\varepsilon_t\right)\right)\Bigg/\left(\sum_{t=1}^{T}\frac{1}{t^2}\right)=\alpha+\left(\sum_{t=1}^{T}\frac{\varepsilon_t}{t}\right)\Bigg/\left(\sum_{t=1}^{T}\frac{1}{t^2}\right)=$$
$$=\alpha+\left(\sum_{t=1}^{T}a_t\varepsilon_t\right)\Bigg/\left(\sum_{t=1}^{T}\frac{1}{t^2}\right).$$

Поскольку знаменатель $\sum_{t=1}^{T}1/t^2$ сходится, то достаточно исследовать сходимость числителя.

*Лемма.* Рассмотрим ряд

$$\sum_{t=0}^{\infty}a_t\varepsilon_t,$$

где $\varepsilon_t$ — независимые, одинаково распределенные нормальные случайные величины с нулевым средним, и ряд $\sum_{t=0}^{\infty}a_t^2$ сходится к положительному числу. Тогда этот ряд сходится по вероятности к некоторой нормальной случайной величине $\eta$, причем $\eta$ не является постоянной.

*Схема доказательства.* Последовательность частичных сумм этого ряда является фундаментальной в пространстве случайных величин, для которых существует второй момент. Это гильбертово пространство, следовательно, ряд сходится к некоторой случайной величине $\eta$ (и дисперсия частичных сумм стремится к $\mathrm{V}(\eta) = \mathrm{V}(\varepsilon_t)\sum_{t=0}^{\infty} a_t^2 > 0$). Значит, он и по вероятности тоже сходится к этой же величине $\eta$, и $\eta$ не является постоянной.

Из леммы следует, что оценки $\widehat{\alpha}$ в первой и второй моделях и $\widehat{\beta}$ в третьей модели не могут сходиться к $\alpha$ и $\beta$ соответственно.

## Задача 3.22

Пусть истинная модель, $y_t = \beta_1 + \beta_2 x_{t2} + \beta_3 x_{t3} + \beta_4 x_{t4} + \varepsilon_t$, удовлетворяет условиям теоремы Гаусса–Маркова. Оценки $\widetilde{\beta}_1$, $\widetilde{\beta}_2$, $\widetilde{\beta}_3$ являются МНК-оценками в регрессии $\boldsymbol{y}$ на $\boldsymbol{x}_2$ и $\boldsymbol{x}_3$. Покажите, что

$$\mathrm{E}\widetilde{\beta}_2 = \beta_2 + \beta_4 \frac{\sum_{t=1}^{n} r_{t2} x_{t4}}{\sum_{t=1}^{n} r_{t2}^2},$$

где $r_{t2}$ — МНК-остатки в регрессии $\boldsymbol{x}_2$ на $\boldsymbol{x}_3$.
*Указание.* Покажите сначала, что МНК-оценка коэффициента $\beta_2$ в уравнении $y_t = \beta_1 + \beta_2 x_{t2} + \beta_3 x_{t3} + \varepsilon_t$ представляется в виде

$$\widetilde{\beta}_2 = \frac{\sum_{t=1}^{n} r_{t2} y_t}{\sum_{t=1}^{n} r_{t2}^2}.$$

## Решение

Докажем, что

$$\widetilde{\beta}_2 = \frac{\sum_{t=1}^{n} r_{t2} y_t}{\sum_{t=1}^{n} r_{t2}^2}.$$

Для этого заметим, что $x_{t2} = \widehat{\alpha}_1 + \widehat{\alpha}_3 x_{t3} + r_{t2}$, где $\boldsymbol{r}_2$ ортогонально $\boldsymbol{\imath}$ и $\boldsymbol{x}_3$. Таким образом, линейные оболочки векторов $\boldsymbol{\imath}$, $\boldsymbol{x}_2$, $\boldsymbol{x}_3$ и $\boldsymbol{\imath}$, $\boldsymbol{r}_2$, $\boldsymbol{x}_3$ совпадают.

Рассмотрим (любой) вектор $\boldsymbol{z} \in \langle \boldsymbol{\imath}, \boldsymbol{x}_2, \boldsymbol{x}_3 \rangle$, $\boldsymbol{z} = a\boldsymbol{\imath} + b\boldsymbol{x}_2 + c\boldsymbol{x}_3 = a\boldsymbol{\imath} + b(\widehat{\alpha}_1\boldsymbol{\imath} + \widehat{\alpha}_3 x_{t3} + r_{t2}) + c\boldsymbol{x}_3 = (a + b\widehat{\alpha}_1)\boldsymbol{\imath} + b\boldsymbol{r}_2 + (c + b\widehat{\alpha}_3)\boldsymbol{x}_3$, т. е. в разложении $\boldsymbol{z}$ по векторам $\boldsymbol{\imath}$, $\boldsymbol{r}_2$, $\boldsymbol{x}_3$ коэффициент при $\boldsymbol{r}_2$ совпадает с коэффициентом при $\boldsymbol{x}_2$ в разложении по $\boldsymbol{\imath}$, $\boldsymbol{x}_2$, $\boldsymbol{x}_3$. Следовательно, МНК-оценки коэффициента $\beta_2$ в регрессиях $\boldsymbol{y}$ на $\boldsymbol{\imath}$, $\boldsymbol{x}_2$, $\boldsymbol{x}_3$ и $\boldsymbol{\imath}$, $\boldsymbol{r}_2$, $\boldsymbol{x}_3$ совпадают.

С другой стороны, так как $\boldsymbol{r}_2$ ортогонален $\boldsymbol{\imath}$ и $\boldsymbol{x}_3$, МНК-оценка коэффициента $\beta_2$ в регрессии $\boldsymbol{y}$ на $\boldsymbol{\imath}$, $\boldsymbol{r}_2$, $\boldsymbol{x}_3$ совпадает с МНК-оценкой коэффициента $\beta_2$ в регрессии $\boldsymbol{y}$ на $\boldsymbol{r}_2$ (без константы). Отсюда следует соотношение

$$\widetilde{\beta}_2 = \frac{\sum_{t=1}^{n} r_{t2} y_t}{\sum_{t=1}^{n} r_{t2}^2}.$$

Теперь докажем основное утверждение задачи.

$$\begin{aligned}
\mathrm{E}\widetilde{\beta}_2 &= \mathrm{E}\frac{\sum_{t=1}^n r_{t2}y_t}{\sum_{t=1}^n r_{t2}^2} = \frac{\sum_{t=1}^n r_{t2}\mathrm{E}y_t}{\sum_{t=1}^n r_{t2}^2} \\
&= \frac{\sum_{t=1}^n r_{t2}\mathrm{E}(\beta_1+\beta_2x_{t2}+\beta_3x_{t3}+\beta_4x_{t4}+\varepsilon_t)}{\sum_{t=1}^n r_{t2}^2} \\
&= \frac{\sum_{t=1}^n r_{t2}(\beta_1+\beta_2x_{t2}+\beta_3x_{t3}+\beta_4x_{t4})}{\sum_{t=1}^n r_{t2}^2} \\
&= \frac{\sum_{t=1}^n r_{t2}(\beta_1+\beta_2(\widehat{\alpha}_1+\widehat{\alpha}_3x_{t3}+r_{t2})+\beta_3x_{t3}+\beta_4x_{t4})}{\sum_{t=1}^n r_{t2}^2} \\
&= \frac{\sum_{t=1}^n r_{t2}((\beta_1+\beta_2\widehat{\alpha}_1)+\beta_2r_{t2}+(\beta_3+\beta_2\widehat{\alpha}_3)x_{t3}+\beta_4x_{t4})}{\sum_{t=1}^n r_{t2}^2} \\
&= \beta_2\frac{\sum_{t=1}^n r_{t2}r_{t2}}{\sum_{t=1}^n r_{t2}^2}+\beta_4\frac{\sum_{t=1}^n r_{t2}x_{t4}}{\sum_{t=1}^n r_{t2}^2} = \beta_2+\beta_4\frac{\sum_{t=1}^n r_{t2}x_{t4}}{\sum_{t=1}^n r_{t2}^2}.
\end{aligned}$$

Здесь мы опять воспользовались тем, что $r_2$ ортогонален $\imath$ и $x_3$. Утверждение доказано.

## Задача 3.23

Рассматривается классическая линейная нормальная модель $y = X\beta + \varepsilon$, $\mathrm{V}(\varepsilon) = 2I$, причем известно, что

$$X'X = \begin{bmatrix} 5 & 2 \\ 2 & 4 \end{bmatrix} \quad \text{и} \quad \widehat{\beta}_1 = 3,\ \widehat{\beta}_2 = 2.$$

а) Постройте 95%-ный доверительный интервал для $\theta = \beta_1 + \beta_2$.

б) Постройте 95%-ную доверительную область для вектора $\begin{bmatrix} \beta_1 \\ \beta_2 \end{bmatrix}$.

## Решение

Согласно условию имеем $\sigma^2 = 2$, и прямые вычисления дают

$$(X'X)^{-1} = \begin{bmatrix} 0.25 & -0.125 \\ -0.125 & 0.3125 \end{bmatrix}.$$

а) Имеем $\mathrm{Cov}(\widehat{\beta}_1, \widehat{\beta}_2) = -0.125\sigma^2 = -0.125 \cdot 2 = -0.25$. Таким образом, $\mathrm{V}(\widehat{\beta}_1 + \widehat{\beta}_2) = 0.5 + 0.625 - 0.5 = 0.625$ и $\widehat{\beta}_1 + \widehat{\beta}_2 \sim N(\beta_1 + \beta_2, 0.625)$. Поэтому 95%-ный доверительный интервал для $\beta_1 + \beta_2$ такой: $\beta_1 + \beta_2 \in [\widehat{\beta}_1 + \widehat{\beta}_2 - 1.96 \cdot \sqrt{0.625}, \widehat{\beta}_1 + \widehat{\beta}_2 + 1.96 \cdot \sqrt{0.625}]$, или $\beta_1 + \beta_2 \in [3.450, 6.550]$.

б) Поскольку

$$\begin{bmatrix}\widehat{\beta}_1\\ \widehat{\beta}_2\end{bmatrix} \sim N\left(\begin{bmatrix}\beta_1\\ \beta_2\end{bmatrix}, \begin{bmatrix}0.5 & -0.25\\ -0.25 & 0.625\end{bmatrix}\right),$$

то согласно известной лемме (см. приложение МС, раздел 4, N9) имеем:

$$\begin{bmatrix}\widehat{\beta}_1-\beta_1 & \widehat{\beta}_2-\beta_2\end{bmatrix}\begin{bmatrix}2.5 & 1\\ 1 & 2\end{bmatrix}\begin{bmatrix}\widehat{\beta}_1-\beta_1\\ \widehat{\beta}_2-\beta_2\end{bmatrix} \sim \chi^2(2).$$

(Здесь стоит отметить, что число степеней свободы равно 2, так как $\sigma^2$ известна.) Учитывая, что 5%-ная точка $\chi^2$-распределения с двумя степенями свободы равна $\chi^2_{0.05}(2) = 5.99$, для 95%-ной доверительной области получаем внутренность следующего эллипса:

$$2.5(\beta_1-3)^2 + 2(\beta_1-3)(\beta_2-2) + 2(\beta_2-2)^2 < 5.99.$$

На рис. 3.1 графически представлены две области: внутренность эллипса

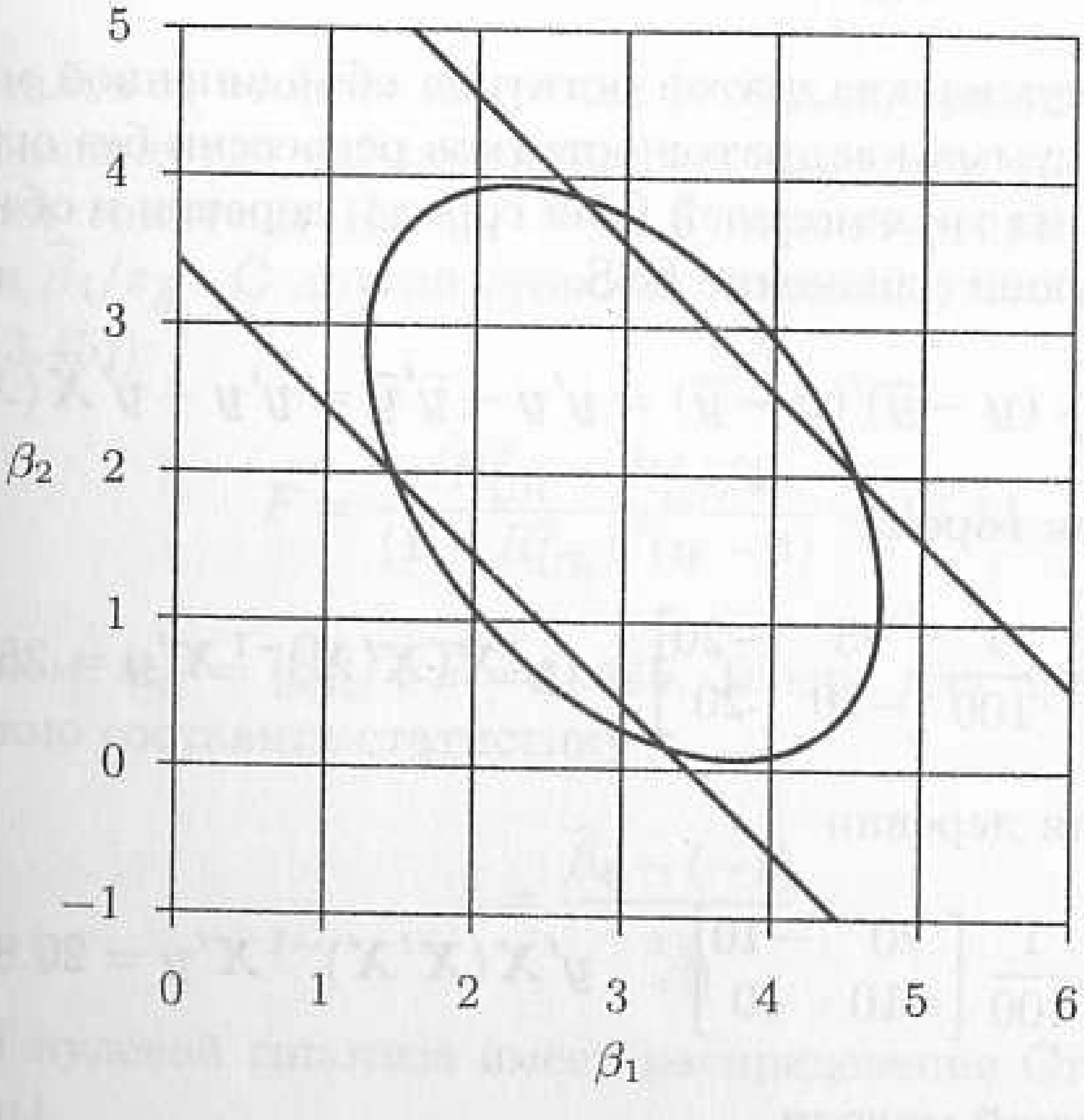



**Рис. 3.1**

из пункта б) и полоса между двумя прямыми ($\beta_1+\beta_2 \in [3.450, 6.550]$). Вероятность попадания параметров $(\beta_1, \beta_2)$ в каждую из этих двух областей равна 95%.

## Задача 3.24

Для города и для деревни рассматриваются две модели парной регрессии. Двадцать наблюдений для города дали следующие результаты:

$$\boldsymbol{X}'\boldsymbol{X} = \begin{bmatrix} 20 & 20 \\ 20 & 25 \end{bmatrix}, \quad \boldsymbol{X}'\boldsymbol{y} = \begin{bmatrix} 10 \\ 20 \end{bmatrix}, \quad \boldsymbol{y}'\boldsymbol{y} = 30,$$

а десять наблюдений для деревни —

$$\boldsymbol{X}'\boldsymbol{X} = \begin{bmatrix} 10 & 10 \\ 10 & 20 \end{bmatrix}, \quad \boldsymbol{X}'\boldsymbol{y} = \begin{bmatrix} 8 \\ 20 \end{bmatrix}, \quad \boldsymbol{y}'\boldsymbol{y} = 24.$$

На 5%-ном уровне значимости проверьте гипотезу о том, что эти две модели совпадают.

## Решение

Для тестирования гипотезы совпадения моделей воспользуемся тестом Чоу (см. (3.50)):

$$F = \frac{(\mathrm{ESS_R} - \mathrm{ESS_{UR}})/2}{\mathrm{ESS_{UR}}/(20 + 10 - 2 \cdot 2)} \sim F(2, 20 + 10 - 2 \cdot 2).$$

Здесь $\mathrm{ESS_R}$ — сумма квадратов остатков объединенной модели, $\mathrm{ESS_{UR}} = \mathrm{ESS_1} + \mathrm{ESS_2}$ — сумма квадратов остатков регрессии без ограничения.

Для каждой из трех моделей (для города, деревни и объединенной) найдем соответствующее значение ESS.

$$\mathrm{ESS} = \boldsymbol{e}'\boldsymbol{e} = (\boldsymbol{y} - \widehat{\boldsymbol{y}})'(\boldsymbol{y} - \widehat{\boldsymbol{y}}) = \boldsymbol{y}'\boldsymbol{y} - \widehat{\boldsymbol{y}}'\widehat{\boldsymbol{y}} = \boldsymbol{y}'\boldsymbol{y} - \boldsymbol{y}'\boldsymbol{X}(\boldsymbol{X}'\boldsymbol{X})^{-1}\boldsymbol{X}'\boldsymbol{y}.$$

В модели для города

$$(\boldsymbol{X}'\boldsymbol{X})^{-1} = \frac{1}{100}\begin{bmatrix} 25 & -20 \\ -20 & 20 \end{bmatrix}, \quad \boldsymbol{y}'\boldsymbol{X}(\boldsymbol{X}'\boldsymbol{X})^{-1}\boldsymbol{X}'\boldsymbol{y} = 25, \quad \mathrm{ESS} = 5.$$

В модели для деревни

$$(\boldsymbol{X}'\boldsymbol{X})^{-1} = \frac{1}{100}\begin{bmatrix} 20 & -10 \\ -10 & 10 \end{bmatrix}, \quad \boldsymbol{y}'\boldsymbol{X}(\boldsymbol{X}'\boldsymbol{X})^{-1}\boldsymbol{X}'\boldsymbol{y} = 20.8, \quad \mathrm{ESS} = 3.2.$$

В объединенной модели

$$\boldsymbol{X}'\boldsymbol{X} = \begin{bmatrix} 30 & 30 \\ 30 & 45 \end{bmatrix}, \quad \boldsymbol{X}'\boldsymbol{y} = \begin{bmatrix} 18 \\ 40 \end{bmatrix}, \quad \boldsymbol{y}'\boldsymbol{y} = 54,$$

$$(\boldsymbol{X}'\boldsymbol{X})^{-1} = \frac{1}{450}\begin{bmatrix} 45 & -30 \\ -30 & 30 \end{bmatrix}, \quad \boldsymbol{y}'\boldsymbol{X}(\boldsymbol{X}'\boldsymbol{X})^{-1}\boldsymbol{X}'\boldsymbol{y} = 39.067, \quad \mathrm{ESS} = 14.933.$$

Искомая $F$-статистика равна

$$F = \frac{(14.933 - (5 + 3.2))/2}{(5 + 3.2)/26} = 10.675.$$

Если выполняется нулевая гипотеза, то $F$ подчиняется $F$-распределению с параметрами $(2, 26)$. При этом ее $p$-значение равно 0.0004, что меньше 0.05. Следовательно, гипотеза о совпадении моделей отвергается на 5%-ном уровне значимости.

## Задача 3.25

Проведены две регрессии ежеквартальных данных со второго квартала 1990 г. по третий квартал 2001 г. Они имеют следующий вид:

$$\widehat{y} = 40 + 0.3x_2 + 0.8x_3 - 1.8x_4, \qquad R^2 = 0.82,$$
$$\widehat{y} = 60 + 0.5x_2 + 0.6x_3, \qquad R^2 = 0.75.$$

Для первой регрессии проверьте (на 5%-ном уровне значимости) гипотезу $\mathrm{H}_0$: $\beta_4 = -1$.

## Решение

Число кварталов со второго квартала 1990 г. по третий квартал 2001 г. равно 46. Это количество наблюдений. Как известно, значение $F$-статистики для проверки гипотезы $\mathrm{H}_0$: $\beta_4 = 0$ в первой регрессии равно квадрату $t$-статистики $\widehat{\beta}_4/s_{\widehat{\beta}_4}$. С другой стороны, известно, что эта же $F$-статистика равна (см. (3.45)):

$$F = \frac{(R^2_{\mathrm{UR}} - R^2_{\mathrm{R}})/1}{(1 - R^2_{\mathrm{UR}})/(n-4)} = 16.33.$$

Таким образом, $s_{\widehat{\beta}_4} = |\widehat{\beta}_4|/\sqrt{F} = 0.445$. Теперь для проверки гипотезы $\mathrm{H}_0$: $\beta_4 = -1$ можно составить статистику

$$t = \frac{\widehat{\beta}_4 - (-1)}{s_{\widehat{\beta}_4}},$$

которая при нулевой гипотезе имеет распределение Стьюдента с 42 степенями свободы.

$$t = \frac{-1.8 - (-1)}{0.445} = -1.80, \quad \text{двустороннее } p\text{-значение равно } 0.079 > 0.05.$$

Следовательно, мы не можем отвергнуть гипотезу $\mathrm{H}_0$: $\beta_4 = -1$ на 5%-ном уровне значимости.

## Задача 3.26

Известно, что процесс, порождающий данные (истинная модель), описывается классической линейной моделью регрессии $\boldsymbol{y} = \boldsymbol{X}\beta + \varepsilon$. Оценка $\widehat{\beta}_R$ получается регрессией $\boldsymbol{y}$ на $\boldsymbol{X}$ (МНК-оценка) при ограничении $\boldsymbol{H}\beta = \boldsymbol{r}$. Найдите матрицу ковариаций $\mathrm{V}(\widehat{\beta}_R)$ и сравните ее с матрицей ковариаций $\mathrm{V}(\widehat{\beta})$ — МНК-оценки в регрессии без ограничений. Как полученный вами результат соотносится с теоремой Гаусса–Маркова?

## Решение

Без ограничения общности можно считать, что матрица $\boldsymbol{H}$ имеет независимые строки, и, значит, матрица $Q = \boldsymbol{H}(\boldsymbol{X}'\boldsymbol{X})^{-1}\boldsymbol{H}'$ обратима (и положительно определена, так как $(\boldsymbol{X}'\boldsymbol{X})^{-1}$ положительно определена, и для любого $\boldsymbol{u} \neq 0$ имеем $\boldsymbol{v} = \boldsymbol{H}'\boldsymbol{u} \neq 0$, и $\boldsymbol{v}(\boldsymbol{X}'\boldsymbol{X})^{-1}\boldsymbol{v} > 0$). Решая задачу поиска минимума функции при наличии ограничений методом Лагранжа

$$(\boldsymbol{y} - \boldsymbol{X}\beta)'(\boldsymbol{y} - \boldsymbol{X}\beta) \to \min \quad \text{при } \boldsymbol{H}\beta = \boldsymbol{r},$$

получаем (см. решение задачи 3.13):

$$\begin{aligned}\widehat{\beta}_{\mathrm{R}} &= \widehat{\beta}_{\mathrm{UR}} + (\boldsymbol{X}'\boldsymbol{X})^{-1}\boldsymbol{H}'Q^{-1}(\boldsymbol{r} - \boldsymbol{H}\widehat{\beta}_{\mathrm{UR}}) \\ &= (I - (\boldsymbol{X}'\boldsymbol{X})^{-1}\boldsymbol{H}'Q^{-1}\boldsymbol{H})\widehat{\beta}_{\mathrm{UR}} + (\boldsymbol{X}'\boldsymbol{X})^{-1}\boldsymbol{H}'Q^{-1}\boldsymbol{r}.\end{aligned}$$

Отсюда

$$\begin{aligned}\mathrm{V}(\widehat{\beta}_{\mathrm{R}}) &= (I - (\boldsymbol{X}'\boldsymbol{X})^{-1}\boldsymbol{H}'Q^{-1}\boldsymbol{H})\mathrm{V}(\widehat{\beta}_{\mathrm{UR}})(I - (\boldsymbol{X}'\boldsymbol{X})^{-1}\boldsymbol{H}'Q^{-1}\boldsymbol{H})' \\ &= \sigma^2(I - (\boldsymbol{X}'\boldsymbol{X})^{-1}\boldsymbol{H}'Q^{-1}\boldsymbol{H})(\boldsymbol{X}'\boldsymbol{X})^{-1}(I - (\boldsymbol{X}'\boldsymbol{X})^{-1}\boldsymbol{H}'Q^{-1}\boldsymbol{H})' \\ &= \sigma^2(I - (\boldsymbol{X}'\boldsymbol{X})^{-1}\boldsymbol{H}'Q^{-1}\boldsymbol{H})(\boldsymbol{X}'\boldsymbol{X})^{-1}(I - \boldsymbol{H}'Q^{-1}\boldsymbol{H}(\boldsymbol{X}'\boldsymbol{X})^{-1}) \\ &= \sigma^2(\boldsymbol{X}'\boldsymbol{X})^{-1} - 2\sigma^2(\boldsymbol{X}'\boldsymbol{X})^{-1}\boldsymbol{H}'Q^{-1}\boldsymbol{H}(\boldsymbol{X}'\boldsymbol{X})^{-1} \\ &\quad + \sigma^2(\boldsymbol{X}'\boldsymbol{X})^{-1}\boldsymbol{H}'Q^{-1}\boldsymbol{H}(\boldsymbol{X}'\boldsymbol{X})^{-1}\boldsymbol{H}'Q^{-1}\boldsymbol{H}(\boldsymbol{X}'\boldsymbol{X})^{-1} \\ &= \mathrm{V}(\widehat{\beta}_{\mathrm{UR}}) - \sigma^2(\boldsymbol{X}'\boldsymbol{X})^{-1}\boldsymbol{H}'Q^{-1}\boldsymbol{H}(\boldsymbol{X}'\boldsymbol{X})^{-1}.\end{aligned}$$

Симметричная матрица $(\boldsymbol{X}'\boldsymbol{X})^{-1}\boldsymbol{H}'Q^{-1}\boldsymbol{H}(\boldsymbol{X}'\boldsymbol{X})^{-1}$ является неотрицательно определенной как положительно определенная, умноженная слева на $\boldsymbol{B} = (\boldsymbol{X}'\boldsymbol{X})^{-1}\boldsymbol{H}'$ и справа на $\boldsymbol{B}' = \boldsymbol{H}(\boldsymbol{X}'\boldsymbol{X})^{-1}$. Значит, $\mathrm{V}(\widehat{\beta}_{\mathrm{R}}) \leqslant \mathrm{V}(\widehat{\beta}_{\mathrm{UR}})$. Полученный результат не противоречит теореме Гаусса–Маркова, так как оценка $\widehat{\beta}_{\mathrm{R}}$ является смещенной.

## Задача 3.27

При каких условиях добавление в уравнение еще одного регрессора не изменяет коэффициент детерминации?

## Решение

См. решение задачи 3.17.

## Задача 3.28

Оценивание производственной функции по методу наименьших квадратов дало следующие результаты:

$$\ln Q = 1.37 + \underset{(0.257)}{0.632} \ln K + \underset{(0.219)}{0.452} \ln L, \quad R^2 = 0.98, \ \widehat{\mathrm{Cov}}(\widehat{\beta}_K, \widehat{\beta}_L) = 0.055$$

(в скобках даны стандартные ошибки). Проверьте гипотезы:

а) эластичности по труду и капиталу совпадают;

б) выполнено свойство постоянства отдачи на масштаб.

*Замечание.* В задаче не указано число наблюдений. Будут ли ваши выводы зависеть от этого числа?

## Решение

а) Формально, гипотезу можно представить в таком виде: $\mathrm{H}_0\colon \beta_K - \beta_L = 0$. Оценка дисперсии случайной величины $\widehat{\beta}_K - \widehat{\beta}_L$ есть

$$\begin{aligned}\widehat{\mathrm{V}}(\widehat{\beta}_K - \widehat{\beta}_L) &= \widehat{\mathrm{V}}(\widehat{\beta}_K) + \widehat{\mathrm{V}}(\widehat{\beta}_L) - 2\widehat{\mathrm{Cov}}(\widehat{\beta}_K, \widehat{\beta}_L) \\ &= 0.257^2 + 0.219^2 - 2 \cdot 0.055 = 0.004.\end{aligned}$$

Для проверки гипотезы запишем $t$-статистику

$$t = \frac{(\widehat{\beta}_K - \widehat{\beta}_L) - 0}{\sqrt{\widehat{\mathrm{V}}(\widehat{\beta}_K - \widehat{\beta}_L)}} = \frac{0.632 - 0.452}{\sqrt{0.004}} = 2.846.$$

Если нулевая гипотеза справедлива, то $t$ подчиняется распределению Стьюдента с $n-3$ степенями свободы.

При этом для $n < 6$ оказывается, что 2.5%-ная точка $t$-распределения оказывается правее полученного значения 2.846 ($t_{0.025}(3) = 3.182$, $t_{0.025}(4) = 2.776$). Следовательно, выводы зависят от числа наблюдений. При $n < 6$ нулевую гипотезу о равенстве эластичностей отвергнуть не удается (недостаточно наблюдений). Если $n > 6$, то нулевая гипотеза отвергается на 5%-ном уровне значимости.

б) Формально, гипотезу можно представить в следующем виде:

$$\mathrm{H}_0\colon \ \beta_K + \beta_L = 1.$$

Оценка дисперсии случайной величины $\widehat{\beta}_K + \widehat{\beta}_L$ есть

$$\begin{aligned}\widehat{\mathrm{V}}(\widehat{\beta}_K + \widehat{\beta}_L) &= \widehat{\mathrm{V}}(\widehat{\beta}_K) + \widehat{\mathrm{V}}(\widehat{\beta}_L) + 2\widehat{\mathrm{Cov}}(\widehat{\beta}_K, \widehat{\beta}_L) \\ &= 0.257^2 + 0.219^2 + 2 \cdot 0.055 = 0.224.\end{aligned}$$

Для проверки гипотезы запишем $t$-статистику

$$t = \frac{(\widehat{\beta}_K + \widehat{\beta}_L) - 1}{\sqrt{\widehat{\mathrm{V}}(\widehat{\beta}_K + \widehat{\beta}_L)}} = \frac{0.632 + 0.452 - 1}{\sqrt{0.224}} = 0.177.$$

Если нулевая гипотеза справедлива, то $t$ подчиняется распределению Стьюдента с $n-3$ степенями свободы. В этом случае выводы не зависят от числа наблюдений, так как 2.5%-ная точка $t$-распределения $t_{0.025}(n-3)$ при всех $n$ превосходит 0.177. Поэтому нулевая гипотеза о постоянстве отдачи на масштаб не отвергается на 5%-ном уровне значимости.

## Задача 3.29

Рассматривается стандартная линейная регрессионная модель

$$y_t = \alpha + \beta x_t + \delta w_t + \theta z_t + \varepsilon_t.$$

а) Какую регрессию следует осуществить, чтобы учесть (истинную) информацию, что $\beta = 2\delta$?

б) Будет ли коэффициент детерминации $R^2$ этой регрессии (п. а)) больше, меньше или равен $R^2$ исходной регрессии?

в) Будут ли оценки параметра $\theta$ в исходной модели и в п. а) несмещенными?

г) Будет ли дисперсия этой оценки (п. а)) больше, меньше или равна дисперсии оценки $\theta$ в исходной регрессии. Объясните на содержательном уровне.

## Решение

а) Подставим соотношение $\beta = 2\delta$ в уравнение регрессии. Получим:

$$y_t = \alpha + 2\delta x_t + \delta w_t + \theta z_t + \varepsilon_t = \alpha + \delta(2x_t + w_t) + \theta z_t + \varepsilon_t.$$

Это уравнение и следует оценивать.

б) Так как линейная оболочка векторов $\boldsymbol{\imath}$, $2\boldsymbol{x} + \boldsymbol{w}$, $\boldsymbol{z}$ содержится в линейной оболочке векторов $\boldsymbol{\imath}$, $\boldsymbol{x}$, $\boldsymbol{w}$, $\boldsymbol{z}$, то коэффициент детерминации при такой замене переменных уменьшится (вообще говоря, не увеличится).

в) Так как информация о том, что $\beta = 2\delta$, истинная, то оценки всех параметров являются несмещенными как в исходной модели (фактически это модель с включенной несущественной переменной), так и в модели п. а). Для формального доказательства воспользуемся утверждением задачи 3.13. Оценка с ограничением имеет вид

$$\widehat{\boldsymbol{\beta}}_{\mathrm{R}} = \widehat{\boldsymbol{\beta}}_{\mathrm{UR}} + (\boldsymbol{X}'\boldsymbol{X})^{-1}\boldsymbol{H}'\left(\boldsymbol{H}(\boldsymbol{X}'\boldsymbol{X})^{-1}\boldsymbol{H}'\right)^{-1}(\boldsymbol{r} - \boldsymbol{H}\widehat{\boldsymbol{\beta}}_{\mathrm{UR}}),$$

где $\boldsymbol{X} = \begin{bmatrix}\boldsymbol{\imath} & \boldsymbol{x} & \boldsymbol{w} & \boldsymbol{z}\end{bmatrix}$, $\boldsymbol{\beta} = \begin{bmatrix}\alpha & \beta & \delta & \theta\end{bmatrix}'$, $\boldsymbol{H} = \begin{bmatrix}0 & 2 & -1 & 0\end{bmatrix}$, $\boldsymbol{r} = 0$.

Поскольку $\boldsymbol{H}\boldsymbol{\beta} = \boldsymbol{r}$, то

$$\begin{aligned}\mathrm{E}(\widehat{\boldsymbol{\beta}}_{\mathrm{R}}) &= \mathrm{E}(\widehat{\boldsymbol{\beta}}_{\mathrm{UR}}) + (\boldsymbol{X}'\boldsymbol{X})^{-1}\boldsymbol{H}'\left(\boldsymbol{H}(\boldsymbol{X}'\boldsymbol{X})^{-1}\boldsymbol{H}'\right)^{-1}(\boldsymbol{r} - \boldsymbol{H}\mathrm{E}(\widehat{\boldsymbol{\beta}}_{\mathrm{UR}}))\\ &= \boldsymbol{\beta} + (\boldsymbol{X}'\boldsymbol{X})^{-1}\boldsymbol{H}'\left(\boldsymbol{H}(\boldsymbol{X}'\boldsymbol{X})^{-1}\boldsymbol{H}'\right)^{-1}(\boldsymbol{r} - \boldsymbol{H}\boldsymbol{\beta})\\ &= \boldsymbol{\beta} + (\boldsymbol{X}'\boldsymbol{X})^{-1}\boldsymbol{H}'\left(\boldsymbol{H}(\boldsymbol{X}'\boldsymbol{X})^{-1}\boldsymbol{H}'\right)^{-1}\cdot 0 = \boldsymbol{\beta}.\end{aligned}$$

г) Дисперсия оценки $\theta$ модели из п. а) будет меньше, чем в исходной модели. На содержательном уровне это объясняется тем, что в п. а) учтена дополнительная информация о том, что $\beta = 2\delta$, что позволяет более точно оценить параметры модели. Формальное доказательство см. в решении задачи 3.26.

## Задача 3.30

В файле `gnovgorod.xls` содержатся данные по стоимости квартир в Новгороде.

а) Постройте и оцените минимальную модель, с помощью которой вы сможете оценить параметр $r$, равный относительному приросту стоимости квартиры при добавлении к ней комнаты площадью 18 кв.м.

б) Найдите 95%-ный доверительный интервал для $r$.

в) Помогает ли включение в модель дополнительных факторов более точно оценить параметр $r$?

г) Можете ли вы предложить модель, в которой параметр $r$ был бы одним из коэффициентов? Изменяется ли при этом способе оценивания доверительный интервал?

д) Зависит ли параметр $r$ от количества комнат в квартире? Почему?

## Решение

а) Рассмотрим, например, следующую спецификацию модели — регрессию логарифма цены на константу, число комнат, жилую площадь, площадь кухни и площадь вспомогательных помещений.

$$\ln price = \beta_1 + \beta_2 numroom + \beta_3 sliv + \beta_4 skit + \beta_5(sall - sliv - skit).$$

Результат оценивания см. в таблице 3.4. При добавлении одной комнаты площадью 18 кв.м относительное изменение цены в среднем равно

$$\frac{P_1 - P_0}{P_0} = e^{\beta_2 + 18\beta_3} - 1 = r.$$

Таким образом, чтобы получить оценку для $r$, достаточно оценить параметр $p = \beta_2 + 18\beta_3$ и в качестве $\widehat{r}$ взять $e^{\widehat{p}} - 1$.

Из этой модели получаем оценку $\widehat{p} = 0.148239 + 18 \cdot 0.007187 = 0.278$ и соответствующую оценку $\widehat{r} = e^{\widehat{p}} - 1 = 0.320$.

**Таблица 3.4**

Dependent Variable: $\ln price$

| Variable | Coefficient | Std. Error | $t$-Statistic | Probability |
|---|---|---|---|---|
| const | 8.544078 | 0.034848 | 245.1802 | 0.0000 |
| *numroom* | 0.148239 | 0.021639 | 6.850671 | 0.0000 |
| *sliv* | 0.007187 | 0.001506 | 4.772626 | 0.0000 |
| *skit* | 0.058311 | 0.005330 | 10.94114 | 0.0000 |
| *sall − sliv − skit* | 0.008734 | 0.001887 | 4.628108 | 0.0000 |
| $R^2$ | 0.778412 | | | |

б) Оценка ковариации оценок $\widehat{\beta}_2$ и $\widehat{\beta}_3$ равна $-2.815 \cdot 10^{-5}$.

Оценка дисперсии оценки $\widehat{p}$ равна:

$$\begin{aligned}\widehat{\mathrm{V}}(\widehat{p}) &= \widehat{\mathrm{V}}(\widehat{\beta}_2) + 2 \cdot 18 \cdot \widehat{\mathrm{Cov}}(\widehat{\beta}_2, \widehat{\beta}_3) + 18^2 \cdot \widehat{\mathrm{V}}(\widehat{\beta}_3)\\ &= 4.682 \cdot 10^{-4} - 2 \cdot 18 \cdot 2.815 \cdot 10^{-5} + 18^2 \cdot 2.268 \cdot 10^{-6}\\ &= 0.0001897 = 0.01377^2.\end{aligned}$$

Следовательно, 95%-ный доверительный интервал для $p$ равен

$$(\widehat{p} - 0.01377 \cdot t_{0.025}(462),\ \widehat{p} + 0.01377 \cdot t_{0.025}(462)) = (0.251, 0.305)$$

(данные содержат 467 наблюдений), и доверительный интервал для $r$ равен (в силу монотонности функции $e^x - 1$):

$$(e^{0.251} - 1,\ e^{0.305} - 1) = (0.285, 0.356).$$

г) Собственно параметр $r$ ввести в регрессию сложно. Но можно так переопределить регрессоры, что $p$ будет входить в регрессию в качестве коэффициента. Пусть $p = \beta_2 + 18\beta_3$ или $\beta_2 = p - 18\beta_3$. Подставим в модель и получим:

$$\begin{aligned}\ln price &= \beta_1 + \beta_2 numroom + \beta_3 sliv + \beta_4 skit + \beta_5(sall - sliv - skit) + \varepsilon\\ &= \beta_1 + p \cdot numroom + \beta_3(sliv - 18 \cdot numroom) + \beta_4 skit\\ &\quad + \beta_5(sall - sliv - skit).\end{aligned}$$

Результаты оценивания этого уравнения приведены в таблице 3.5.

Так как это по сути та же модель, что и в п. а), то ни численное значение оценок $\widehat{p}, \widehat{r}$, ни доверительные интервалы не изменились.

в) Попробуем включить в модель другие доступные параметры, от которых зависит цена квартиры. Это может не уменьшить стандартное отклонение параметра, но возможно, устранит смещение (см. таблицу 3.6).

**Таблица 3.5**

Dependent Variable: $\ln price$

| Variable | Coefficient | Std. Error | $t$-Statistic | Probability |
|---|---|---|---|---|
| const | 8.544078 | 0.034848 | 245.1802 | 0.0000 |
| $numroom$ | **0.277606** | **0.013773** | 20.15527 | 0.0000 |
| $sliv - 18 \cdot numroom$ | 0.007187 | 0.001506 | 4.772626 | 0.0000 |
| $skit$ | 0.058311 | 0.005330 | 10.94114 | 0.0000 |
| $sall - sliv - skit$ | 0.008734 | 0.001887 | 4.628108 | 0.0000 |
| $R^2$ | 0.778412 | | | |

**Таблица 3.6**

Dependent Variable: $\ln price$

| Variable | Coefficient | Std. Error | $t$-Statistic | Probability |
|---|---|---|---|---|
| const | 8.567838 | 0.035650 | 240.3319 | 0.0000 |
| $numroom$ | **0.296536** | **0.013501** | 21.96340 | 0.0000 |
| $sliv - 18 \cdot numroom$ | 0.009403 | 0.001548 | 6.072588 | 0.0000 |
| $skit$ | 0.048308 | 0.005355 | 9.020974 | 0.0000 |
| $sall - sliv - skit$ | 0.006445 | 0.001821 | 3.539112 | 0.0004 |
| $lodzhia$ | 0.098493 | 0.015719 | 6.265850 | 0.0000 |
| $firstfloor$ | −0.052270 | 0.021576 | −2.422571 | 0.0158 |
| $walltype$ | 0.050798 | 0.017056 | 2.978285 | 0.0031 |
| $R^2$ | 0.801303 | | | |

Здесь $lodzhia$, $firstfloor$, $walltype$ — переменные, показывающие наличие лоджии, то, что квартира находится на первом этаже, тип дома (кирпичный или панельный) соответственно.

Доверительные интервалы для $p$ немного сместились вправо.

$$p \in (0.2965 - 0.0135 \cdot t_{0.025}(462),\ 0.2965 + 0.0135 \cdot t_{0.025}(462)) = (0.270, 0.323),$$
$$r \in (e^{0.270} - 1,\ e^{0.323} - 1) = (0.310, 0.381).$$

д) Приведенные выше модели не позволяют определить зависимость $r$ от числа комнат. Но, скорее всего, зависимость должна быть убывающей по следующим причинам: 1) доля цены одной комнаты в цене квартиры больше в маленькой квартире, чем в большой, 2) сама цена еще одной комнаты меньше при переходе, скажем, от 4- к 5-комнатной квартире, чем при переходе от 1- к 2-комнатной квартире.

Для оценки зависимости $r$ от числа комнат можно, например, добавить в модель $numroom^2$ (см. таблицу 3.7).

Таблица 3.7

Dependent Variable: $\ln price$

| Variable | Coefficient | Std. Error | $t$-Statistic | Probability |
|---|---|---|---|---|
| const | 8.286998 | 0.046282 | 179.0529 | 0.0000 |
| $numroom$ | **0.549826** | 0.031772 | 17.30533 | 0.0000 |
| $numroom^2$ | **−0.053479** | 0.006165 | −8.674727 | 0.0000 |
| $sliv - 18 \cdot numroom$ | 0.011363 | 0.001454 | 7.813097 | 0.0000 |
| $skit$ | 0.049557 | 0.004970 | 9.970645 | 0.0000 |
| $sall - sliv - skit$ | 0.010139 | 0.001742 | 5.819358 | 0.0000 |
| $lodzhia$ | 0.085992 | 0.014655 | 5.867851 | 0.0000 |
| $firstfloor$ | −0.043811 | 0.020042 | −2.185999 | 0.0293 |
| $walltype$ | 0.053308 | 0.015827 | 3.368234 | 0.0008 |
| $R^2$ | 0.829342 | | | |

Таким образом, получилась следующая зависимость $r$ от числа комнат:

$$\widehat{r} = e^{0.5498 - 2 \cdot 0.0535 \cdot numroom} = e^{0.5498 - 0.1070 \cdot numroom}.$$

## Задача 3.31

В примере рассматриваются данные по стоимости квартир в Москве, собранные студентами первого курса РЭШ осенью 1997 г. Описание переменных содержится в таблице 3.8.

Данные находятся в файле `flat98s.xls`

а) Постройте модель стоимости квартиры (или стоимости квадратного метра жилой площади квартиры) в зависимости от имеющихся факторов.

б) Проверьте гипотезу, что модели для 1, 2, 3–4-комнатных квартир различаются между собой, т. е. гипотезу, что рынок распадается на рынки однокомнатных, двухкомнатных и трех-четырехкомнатных квартир.

## Решение

а) При построении модели руководствуемся критериями простоты, значимости коэффициентов, прогнозной силы.

Для начала рассмотрим описательные статистики всех переменных (см. таблицу 3.9). При этом добавим к уже имеющимся переменную $dopsp = totsp - livsp - kitsp$. Эта переменная имеет смысл площади дополнительных помещений (ванная, туалет, коридор и т. п.). Всего у нас доступно 3226 наблюдений, но все переменные заданы только в 3009 из них (для них и рассматриваем описательные статистики).

Таблица 3.8

| Переменная | Описание |
|---|---|
| *price* | цена квартиры, тыс. долл. |
| *rooms* | количество жилых комнат |
| *totsp* | общая площадь, кв. м |
| *livsp* | жилая площадь, кв. м |
| *kitsp* | площадь кухни, кв. м |
| *dist* | расстояние до центра, км |
| *metrdist* | расстояние до ближайшей станции метро, мин |
| *walk* | 1, если пешком от метро, 0 — иначе |
| *brick* | 1, если дом кирпичный, 0 — иначе |
| *tel* | 1, если есть телефон, 0 — иначе |
| *bal* | 1, если есть балкон или лоджия, 0 — иначе |
| *floor* | 0, если квартира находится на первом или последнем этаже, 1 — иначе |

Таблица 3.9

| | *price* | *rooms* | *totsp* | *livsp* | *kitsp* | *dopsp* | *dist* |
|---|---|---|---|---|---|---|---|
| Mean | 53.20 | 1.79 | 49.67 | 29.97 | 8.28 | 11.41 | 9.32 |
| Median | 42.00 | 1.00 | 40.00 | 22.70 | 8.50 | 10.00 | 9.60 |
| Maximum | 902.00 | 4.00 | 180.00 | 119.00 | 30.00 | 58.00 | 21.88 |
| Minimum | 18.00 | 1.00 | 18.00 | 10.00 | 0.00 | −6.10 | 0.40 |
| Std. Dev. | 35.11 | 0.91 | 19.72 | 14.16 | 2.14 | 5.95 | 4.53 |

| | *metrdist* | *bal* | *brick* | *floor* | *tel* | *walk* |
|---|---|---|---|---|---|---|
| Mean | 8.66 | 0.72 | 0.36 | 0.70 | 0.86 | 0.73 |
| Median | 10.00 | 1.00 | 0.00 | 1.00 | 1.00 | 1.00 |
| Maximum | 25.00 | 1.00 | 1.00 | 3.50 | 1.00 | 1.00 |
| Minimum | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 |
| Std. Dev. | 3.63 | 0.45 | 0.48 | 0.46 | 0.33 | 0.44 |

При внимательном рассмотрении статистик привлекают внимание два обстоятельства. Во-первых, минимальное значение переменной *dopsp* отрицательное, что выглядит неправдоподобно. Кроме того, минимальное значение переменной *kitsp* равно нулю (надо полагать, существуют квартиры без кухни, но тогда их цена должна устанавливаться по другим правилам, и эти наблюдения следует исключить из рассмотрения). Кроме того, исключим также наблюдения, в которых площадь кухни очень мала (считаем, что кухня должна иметь площадь не меньше 5 квадратных метров).

Все остальные переменные выглядят более правдоподобно.

После того, как неправдоподобные наблюдения были отброшены, осталось 2982 наблюдения, для которых определены все переменные.

При построении модели будем отталкиваться от вопроса, на который нам надо ответить в п. б). Для этого введем три бинарные переменные $r1$, $r2$, $r34$, равные 1, если квартира одно-, двух-, трех- или четырехкомнатная соответственно.

В результате мы остановились на двух следующих моделях цены и цены квадратного метра квартиры (см. таблицы 3.10 и 3.11):

**Таблица 3.10**

Dependent Variable: *price*

| Variable | Coefficient | Std. Error | $t$-Statistic | Probability |
|---|---|---|---|---|
| $r1$ | −28.73402 | 2.932344 | −9.798994 | 0.0000 |
| $r2$ | −39.32051 | 3.211052 | −12.24537 | 0.0000 |
| $r34$ | −46.55833 | 4.114623 | −11.31533 | 0.0000 |
| *livsp* | 1.423778 | 0.073509 | 19.36887 | 0.0000 |
| *kitsp* | 2.606294 | 0.215474 | 12.09562 | 0.0000 |
| *dopsp* | 1.949937 | 0.098770 | 19.74221 | 0.0000 |
| *dist* | −1.087874 | 0.097145 | −11.19850 | 0.0000 |
| *metrdist* | −0.504183 | 0.110331 | −4.569734 | 0.0000 |
| *walk* | 4.000565 | 0.900171 | 4.444229 | 0.0000 |
| *bal* | 3.159488 | 0.864566 | 3.654423 | 0.0003 |
| *brick* | 4.167797 | 0.914909 | 4.555422 | 0.0000 |
| *floor* | 4.836079 | 0.837665 | 5.773285 | 0.0000 |
| *tel* | 7.763834 | 1.189693 | 6.525912 | 0.0000 |
| $R^2$ | 0.654407 | | | |

Обе модели получились довольно простыми, все коэффициенты (кроме переменной *livsp* во второй модели) значимы, т. е. выбрасывать переменные нет смысла. При добавлении других переменных (например, перекрестных членов, квадратов уже включенных) прогнозные свойства модели практически не меняются.

Коэффициент при *livsp* в модели для цены квадратного метра общей площади оказался незначимым. Коэффициенты при *kitsp* и *dopsp* положительные, что означает, что квадратный метр кухни и комнаты стоит дороже, чем квадратный метр комнаты. Это согласуется с результатами, полученными из первой модели. Действительно, если проверять гипотезу о равенстве коэффициентов, например при *dopsp* и *livsp*, то получим, что значение $F$-статистики равно 14.41, и гипотеза уверенно отвергается ($p$-значение 0.0001).

**Таблица 3.11**

Dependent Variable: *price/totsp*

| Variable | Coefficient | Std. Error | $t$-Statistic | Probability |
|---|---|---|---|---|
| $r1$ | 0.755563 | 0.032439 | 23.29216 | 0.0000 |
| $r2$ | 0.693431 | 0.035522 | 19.52134 | 0.0000 |
| $r34$ | 0.705593 | 0.045517 | 15.50166 | 0.0000 |
| *livsp* | −0.000885 | 0.000813 | −1.087822 | 0.2768 |
| *kitsp* | 0.026837 | 0.002384 | 11.25880 | 0.0000 |
| *dopsp* | 0.008412 | 0.001093 | 7.698996 | 0.0000 |
| *dist* | −0.018849 | 0.001075 | −17.53981 | 0.0000 |
| *metrdist* | −0.008131 | 0.001221 | −6.661624 | 0.0000 |
| *walk* | 0.079115 | 0.009958 | 7.944920 | 0.0000 |
| *bal* | 0.055470 | 0.009564 | 5.799846 | 0.0000 |
| *brick* | 0.100093 | 0.010121 | 9.889583 | 0.0000 |
| *floor* | 0.085143 | 0.009267 | 9.188260 | 0.0000 |
| *tel* | 0.087055 | 0.013161 | 6.614744 | 0.0000 |
| $R^2$ | 0.342542 | | | |

б) Проверим гипотезу совпадения рынков для двух построенных моделей. Для этого сначала протестируем гипотезу равенства равенства коэффициентов при переменных $r1$, $r2$ и $r3$. В модели для цены квартиры $F$-статистика равна 38.42602 ($p$-значение 0.0000), в модели для цены квадратного метра квартиры $F$-статистика равна 10.58067 ($p$-значение 0.0000), и гипотеза о совпадении рынков квартир с разным числом комнат уверенно отвергается.

Попробуем теперь сравнить рынки попарно. Результаты проведения тестов приведены в таблице 3.12.

**Таблица 3.12**

| | Модель 1 | | Модель 2 | |
|---|---|---|---|---|
| $\mathrm{H}_0$ | $F$-статистика | $p$-значение | $F$-статистика | $p$-значение |
| $r1 = r2$ | 68.55784 | 0.0000 | 19.29737 | 0.0000 |
| $r2 = r3$ | 17.50125 | 0.0000 | 0.403843 | 0.5252 |
| $r1 = r3$ | 59.44194 | 0.0000 | 3.817615 | 0.0508 |

Как видим, модель цены квартиры уверенно разбивает рынок на три категории по числу комнат в квартире. В то же время модель цены квадратного метра позволяет разбить рынок только на две группы (однокомнатные квартиры и остальные квартиры), причем различия в ценообразовании однокомнатных и трех-, четырехкомнатных удается установить не слишком уверенно.

Чем можно объяснить обнаруженное различие? Одно из объяснений может быть таким: фиксированные издержки при строительстве квартиры в удельном выражении выше для небольших (однокомнатных квартир), поэтому квадратный метр в них должен стоить больше. Это рассуждение, однако, не годится для объяснения того, что цена квадратного метра увеличивается при переходе от двухкомнатных к трех-четырехкомнатным квартирам. Однако это отличие статистически незначимо. (Если разделить трех- и четырехкомнатные квартиры на две категории, то последовательность цен квадратного метра общей площади станет монотонной, впрочем, отличия между 2-, 3-, 4-комнатными квартирами также незначимы).

## Задача 3.32

(автор — Arthur van Soest, Tilburg University)

*Введение.* Рассматриваемые здесь упражнения в значительной мере опираются на статью (Mankiw et al., 1992)[1] и направлены на проверку полученных там результатов (в первую очередь, с точки зрения здравого экономического смысла). В цитированной статье изучается расширенный вариант модели экономического роста Солоу. Основным объектом, изучаемым в модели Солоу, является удельная величина валового внутреннего продукта (ВВП) в стационарном состоянии. Таким образом, модель объясняет различие в уровне благосостояния разных стран в долговременном плане. Обобщение модели Солоу в работе (Mankiw et al., 1992) состоит в том, что в отличие от первоначальной модели здесь допускаются инвестиции не только в физический, но и в человеческий капитал. Приведенный там эмпирический анализ основан на межстрановых данных, взятых из работы (Summers, Heston, 1988)[2]. Мы также будем использовать эти данные.

*Обобщенная модель Солоу.* Дадим краткое описание обобщенной модели Солоу, предложенной в работе (Mankiw et al., 1992). Исходная модель Солоу изложена во многих учебниках по макроэкономике (см., например, (Romer, 2001)[3]). Предполагается, что в каждый момент времени $t$ производство задается производственной функцией Кобба–Дугласа с постоянной отдачей на масштаб:

$$Y_t = K_t^{\alpha} H_t^{\beta} (A_t L_t)^{1-\alpha-\beta},$$

где $Y$ — выпуск, $K$ и $H$ — объем физического и человеческого капитала, соответственно, $L$ — труд, а переменная $A$ описывает уровень технологии.

[1]Mankiw N. G., Romer D. and Weil D. N. (1992). A Contribution to the empirics of economic growth. *Quarterly Journal of Economics*, v. 107(2), pp. 407–437.

[2]Summers R. and Heston A. (1988). A new set of international comparisons of real product and price levels estimates for 130 countries, 1950–85. *Review of Income and Wealth*, v. 34, pp. 1–26.

[3]Romer D. (2001). *Advanced Macroeconomics*. 2nd edition. McGraw-Hill, New York.

Предположение о постоянстве отдачи на масштаб позволяет оперировать с удельными величинами (на единицу эффективного труда):

$$y = \frac{Y}{AL},\ k = \frac{K}{AL},\ h = \frac{H}{AL}.$$

Будем также считать, что выполнены следующие условия:

- фиксированные доли $s_k$, $s_h$ суммарного выпуска $Y$ инвестируются в физический и человеческий капитал, соответственно;
- $L_t = L_0 e^{nt}$, где $n$ — скорость роста населения;
- $A_t = A_0 e^{gt}$, где $g$ — скорость роста технологического уровня;
- интенсивность амортизации $\delta$ одинакова для физического и человеческого капитала.

Из этих предположений вытекает, что эволюция капитала описывается следующей системой дифференциальных уравнений:

$$\dot{k}_t = s_k y_t - (n+g+\delta)k_t,$$

$$\dot{h}_t = s_h y_t - (n+g+\delta)h_t.$$

Стационарное состояние характеризуется условиями $\dot{k}_t = \dot{h}_t = 0$.

3.32.1. Покажите, что в стационарном состоянии выполнено равенство

$$\ln y_t = \frac{1}{1-\alpha-\beta}(\alpha \ln s_k + \beta \ln s_h - (\alpha+\beta)\ln(n+g+\delta)). \quad (1)$$

Равенство (1) устанавливает соотношение (в стационарном состоянии) между благосостоянием страны, скоростью роста ее населения и интенсивностью инвестиций в физический и человеческий капитал. Следствием этого соотношения является то, что и в долговременном плане можно ожидать сохранение различия в уровне благосостояния разных стран.

Модель также позволяет описать траекторию сходимости к стационарному состоянию. Пусть $y^*$ — значение $y_t$ в стационарном состоянии. Тогда можно показать, что имеет место следующее приближенное соотношение:

$$\frac{d \ln y_t}{dt} = \lambda(\ln y^* - \ln y_t),$$

где $\lambda = (n+g+\delta)(1-\alpha-\beta)$. Решая это уравнение, получаем

$$\ln y_t = (1-e^{-\lambda t})\ln y^* + e^{-\lambda t}\ln y_0. \quad (2)$$

3.32.2. Покажите, что из (1) и (2) вытекает следующее уравнение для траектории сходимости:

$$\ln \frac{y_t}{y_0} = (1-e^{-\lambda t}) \times \left\{ \frac{1}{1-\alpha-\beta}(\alpha \ln s_k + \beta \ln s_h - (\alpha+\beta)\ln(n+g+\delta)) - \ln y_0 \right\}. \quad (3)$$

*Данные.* Используются данные, извлеченные из архива журнала Journal of Applied Econometrics. Они соответствуют работе (Durlauf, Johnson 1995). Начало этим исследованиям положила работа (Summers, Heston 1988). Единицей наблюдения является страна, даны результаты наблюдений 121 страны. Используются переменные, перечисленные в таблице 3.13.

**Таблица 3.13**

| Переменная | Описание |
|---|---|
| *num* | номер страны в базе данных Summers, Heston (1988) |
| *noil* | 1 для страны, не добывающей нефть, 0 — для добывающей |
| *inter* | 1 для страны с хорошим качеством данных, 0 — в противном случае |
| *oecd* | 1 для страны, входящей в Организацию экономического сотрудничества и развития, 0 — в противном случае |
| *gdp*60 | ВВП на душу населения в 1960 г. (долл.) |
| *gdp*85 | ВВП на душу населения в 1985 г. (долл.) |
| *gdpgro* | средний рост ВВП на душу населения с 1960 по 1985 г. (%) |
| *popgro* | средний рост работоспособного населения с 1960 по 1985 г. (%) |
| *iony* | средняя доля инвестиций (включая государственные) в общем объеме ВВП с 1960 по 1985 г. (%) |
| *sch* | средняя доля населения, продолжающего получать образование одновременно с работой с 1960 по 1985 г. (%) |
| *lit* | доля людей среди населения старше 15 лет, умеющих читать и писать в 1960 г. |

Все данные, за исключением *lit*, взяты из работы (Mankiw et al., 1992); переменная *lit* взята из доклада Всемирного банка. Данные содержатся в файле `growth.xls`. Список стран приведен в приложении к работе (Mankiw et al., 1992).

3.32.3. а) Вычислите суммарные статистики всех переменных. Проверьте, имеют ли смысл ваши результаты.

б) Вычислите корреляционную матрицу всех переменных. Дайте интерпретацию наиболее важных результатов. Соответствуют ли они тому, что вы ожидали?

*Анализ стационарного состояния.* Если предположить, что в 1985 г. страны достигли стационарного состояния, то мы можем использовать достигнутый в 1985 г. уровень ВВП для оценивания уравнения (1). Поскольку мы используем данные, относящиеся к одному и тому же году, то индекс $t$ можно опустить. Уравнение (1) переписывается в следующем виде:

$$\ln gdp85 = \pi_0 + \pi_1 \ln s_k + \pi_2 \ln s_h + \pi_3 \ln(n + g + \delta), \qquad (4)$$

где $\pi_0 = \ln A_0 + gt$ — постоянный член. При оценивании уравнения (4) представляется разумным в качестве $s_k$ использовать переменную *iony*, а в качестве $s_h$ переменную *sch*. Мы не наблюдаем величины $g$ и $\delta$, поэтому будем считать, как в работе (Mankiw et al., 1992), что $g = 2\%$ и $\delta = 3\%$. В качестве $n$ берется переменная *popgro*.

3.32.4. а) Оцените уравнение (4), используя данные по всем странам, за исключением тех, для которых пропущены наблюдения какой-либо переменной.

б) Исходная модель Солоу не включает человеческий капитал. Оцените уравнение (4), удалив переменную $\ln sch$. Сравните с результатом, полученным в п. а). В чем состоит основное различие? Объясните это различие, используя также результаты упражнения 3.32.3.

3.32.5. а) Структурная форма (1) накладывает некоторое линейное ограничение на параметры $\pi_1, \pi_2, \pi_3$ приведенной формы. Что это за ограничение?

б) Протестируйте (на 5%-ном уровне значимости) выполнимость этого ограничения.

в) Оцените вновь уравнение (4), используя это ограничение. Сравните ваш результат с результатом, полученным в упражнении 3.32.4 а).

г) Выразите структурные параметры $\alpha$ и $\beta$ через $\pi_1, \pi_2$ и постройте, таким образом, их оценки.

3.32.6. а) Добавьте в регрессионное уравнение упражнения 3.32.4 а) фиктивные переменные *noil* и *oecd* и проверьте их значимость.

б) Проверьте, является линейная спецификация (4) разумной, добавляя квадраты независимых переменных и перекрестные члены.

3.32.7. Согласно «золотому правилу накопления капитала», доли инвестиций $s_k$, $s_h$ должны выбираться таким образом, чтобы в стационарном состоянии величина $c = (1 - s_k - s_h)y$ была максимальна.

а) Найдите теоретические оптимальные значения величин $s_k$, $s_h$.

б) Используя оценки, полученные в упражнении 3.32.5, проверьте, удовлетворяют ли в среднем инвестиции в физический капитал «золотому правилу».

*Рост ВВП.* Уравнение (3) служит основой эмпирического анализа роста ВВП в период с 1960 ($t = 0$) по 1985 г. ($t = 25$). Заметим, что уравнение (3) можно переписать следующим образом:

$$\ln \frac{gdp85}{gdp60} = \pi_0 + \pi_1 \ln s_k + \pi_2 \ln s_h + \pi_3 \ln(n + g + \delta) + \pi_4 \ln gdp60. \qquad (5)$$

При оценивании этого уравнения будем использовать те же предположения, что и в предыдущих разделах. Так, например, $g = 2\%$, $\delta = 3\%$ и т. д.

3.32.8. а) Оцените уравнение (5), интерпретируйте результат.

б) Исходная модель Солоу не включает человеческий капитал. Оцените уравнение (5), удалив переменную $\ln sch$. Сравните результат с тем, что получен в п. а). Объясните разницу.

3.32.9. а) Структурная форма (3) накладывает некоторое линейное ограничение на параметры $\pi_1, \pi_2, \pi_3$ приведенной формы. Что это за ограничение?

б) Протестируйте (на 5%-ном уровне значимости) выполнимость этого ограничения.

в) Оцените вновь уравнение (5), используя это ограничение. Сравните ваш результат с результатом, полученным в упражнении 3.32.8 а).

г) Используя результат п. в), постройте оценки структурных параметров $\lambda, \alpha, \beta$. Проинтерпретируйте результаты. Сравните ваши оценки параметров $\alpha, \beta$ с оценками, полученными в предыдущих упражнениях.

3.32.10. а) Добавьте фиктивные переменные *noil* и *oecd* в уравнение упражнения 3.32.8 а) и проверьте их значимость.

б) Проверьте, является линейная спецификация (5) разумной, добавляя квадраты независимых переменных и перекрестные члены.

3.32.11. а) Оцените уравнение (5) отдельно для стран — членов OECD и для стран, не входящих в OECD, и проинтерпретируйте результаты.

б) Проверьте, совпадают ли коэффициенты уравнения (5) (за исключением свободного члена) для стран — членов OECD и для стран, не входящих в OECD.

3.32.12. Выберите наилучшее, с вашей точки зрения, уравнение и постройте 95%-ный доверительный интервал для скорости сходимости $\lambda$. Проинтерпретируйте результат.

## Решение

3.32.1. Стационарное состояние характеризуется соотношениями

$$s_k y_t = (n+g+\delta)k_t,$$
$$s_h y_t = (n+g+\delta)h_t.$$

Прологарифмировав оба уравнения и сложив их с коэффициентами $\alpha$ и $\beta$, получим:

$$\begin{aligned}\alpha \ln s_k + \beta \ln s_h + (\alpha+\beta)\ln y_t &= \alpha \ln k_t + \beta \ln h_t + (\alpha+\beta)\ln(n+g+\delta) \\ &= \ln y_t + (\alpha+\beta)\ln(n+g+\delta)\end{aligned}$$

(Последнее равенство справедливо в силу того, что $Y = K^{\alpha}H^{\beta}(AL)^{1-\alpha-\beta}$, а значит, $y = k^{\alpha}h^{\beta}$ и $\ln y = \alpha \ln k + \beta \ln h$). Отсюда немедленно следует (1).

3.32.2. Вычитая $\ln y_0$ из левой и правой частей равенства (2) и подставляя выражение для $y^*$ из (1), получим (3):

$$\ln \frac{y_t}{y_0} = (1 - e^{-\lambda t})(\ln y^* - \ln y_0)$$
$$= (1 - e^{-\lambda t}) \left\{ \frac{1}{1-\alpha-\beta} (\alpha \ln s_k + \beta \ln s_h - (\alpha+\beta)\ln(n+g+\delta)) - \ln y_0 \right\}.$$

3.32.3. а) Суммарные статистики показаны в таблице 3.14. Из таблицы мы можем заключить, что 19% стран добывают нефть, что 18% стран входят в OECD, в 1985 г. средний ВВП на душу населения вырос по сравнению с 1960 г. В то же время максимум среднедушевого ВВП уменьшился, т. е. страны подровнялись (как и предсказывает модель Солоу). Темп роста населения варьирует довольно сильно, от отрицательных значений до 9.2% в год. Никаких явных ошибок в данных просмотр суммарных статистик не выявляет.

**Таблица 3.14**

| | *noil* | *oecd* | *gdp60* | *gdp85* | *gdpgro* | *iony* | *sch* | *lit* |
|---|---|---|---|---|---|---|---|---|
| Mean | 0.81 | 0.18 | 3681.8 | 5683.6 | 4.09 | 18.16 | 5.53 | 48.17 |
| Median | 1.00 | 0.00 | 1962.0 | 3484.5 | 3.90 | 17.70 | 4.95 | 39.00 |
| Maximum | 1.00 | 1.00 | 77881.0 | 25635.0 | 9.20 | 36.90 | 12.10 | 100.00 |
| Minimum | 0.00 | 0.00 | 383.0 | 412.0 | −0.90 | 4.10 | 0.40 | 1.00 |
| Std. Dev. | 0.39 | 0.39 | 7492.9 | 5688.7 | 1.89 | 7.85 | 3.53 | 35.35 |

б) Корреляционная матрица приведена в таблице 3.15. Прокомментируем наиболее интересные результаты.

**Таблица 3.15**

| | *noil* | *oecd* | *gdp60* | *gdp85* | *gdpgro* | *iony* | *sch* | *lit* |
|---|---|---|---|---|---|---|---|---|
| *noil* | 1.000 | 0.108 | −0.498 | −0.248 | −0.077 | 0.084 | −0.071 | 0.166 |
| *oecd* | 0.108 | 1.000 | 0.193 | 0.711 | −0.031 | 0.570 | 0.560 | 0.658 |
| *gdp60* | −0.498 | 0.193 | 1.000 | 0.629 | −0.112 | 0.088 | 0.348 | 0.253 |
| *gdp85* | −0.248 | 0.711 | 0.629 | 1.000 | 0.161 | 0.595 | 0.728 | 0.726 |
| *gdpgro* | −0.077 | −0.031 | −0.112 | 0.161 | 1.000 | 0.370 | 0.280 | 0.138 |
| *iony* | 0.084 | 0.570 | 0.088 | 0.595 | 0.370 | 1.000 | 0.632 | 0.636 |
| *sch* | −0.071 | 0.560 | 0.348 | 0.728 | 0.280 | 0.632 | 1.000 | 0.812 |
| *lit* | 0.166 | 0.658 | 0.253 | 0.726 | 0.138 | 0.636 | 0.812 | 1.000 |

Так, *gdp*85 положительно коррелирует с *iony* ($r = 0.595$), в то же время коэффициент корреляции между *iony* и *gdp*60 равен всего 0.088 (это вполне объяснимо с точки зрения модели Солоу, если считать, что страны к 1985 г. уже приблизились к стационарному состоянию, а в 1960 г. еще были далеко от него, причем «с разных сторон» от него). Величина $gdp85$ также высоко коррелирована с $sch$ ($r = 0.728$), $lit$ ($r = 0.726$). Из других результатов отметим положительную корреляцию между $iony$ и $sch$ ($r = 0.632$).

3.32.4. а) Результат оценивания уравнения (4) приведен в таблице 3.16.

**Таблица 3.16**

Dependent Variable: $\ln gdp85$

| Variable | Coefficient | Std. Error | $t$-Statistic | Probability |
|---|---|---|---|---|
| const | 6.381717 | 1.063682 | 5.999648 | 0.0000 |
| $\ln iony$ | 0.702930 | 0.155419 | 4.522795 | 0.0000 |
| $\ln sch$ | 0.685665 | 0.081615 | 8.401229 | 0.0000 |
| $\ln(popgro + 2 + 3)$ | −0.580552 | 0.452453 | −1.283122 | 0.2024 |
| $R^2$ | 0.689678 | | | |
| $R^2_{\text{adj}}$ | 0.680551 | | | |

Все коэффициенты значимы на 5%-ном уровне, за исключением коэффициента при $\ln(popgro + 2 + 3)$, который удалять не следует, так как его наличие продиктовано теорией.

б) Удалим переменную $\ln sch$ (человеческий капитал) из регрессии, как было в исходной модели Солоу (см. таблицу 3.17). Стоит отметить, что по сравнению с моделью из таблицы 3.16 модель значительно ухудшилась с точки зрения прогноза ($R^2$ значительно уменьшился). Кроме того, в этой модели коэффициент $\pi_1 = \alpha/(1 - \alpha)$. Откуда получаем, что $\widehat{\alpha} = \widehat{\pi}_1/(1 + \widehat{\pi}_1) = 1.478/2.478 = 0.596$, что значительно отличается от эмпирически полученных оценок ($\approx 0.3$, поиски исследований, посвященных изменению доли капитала в экономиках разных стран, можно начать с работы (Mankiw et al., 1992)).

**Таблица 3.17**

Dependent Variable: $\ln gdp85$

| Variable | Coefficient | Std. Error | $t$-Statistic | Probability |
|---|---|---|---|---|
| const | 4.929000 | 1.354773 | 3.638248 | 0.0004 |
| $\ln iony$ | 1.477996 | 0.163085 | 9.062717 | 0.0000 |
| $\ln(popgro + 2 + 3)$ | −0.457349 | 0.583666 | −0.783581 | 0.4351 |
| $R^2$ | 0.476729 | | | |
| $R^2_{\text{adj}}$ | 0.466666 | | | |

Такое большое положительное смещение получилось из-за того, что мы исключили из регрессии существенную переменную ($\ln sch$), которая положительно коррелирована с $\ln iony$.

3.32.5. а) Коэффициенты в уравнении (1) при переменных $\ln(n+g+\delta)$, $\ln s_k$, $\ln s_h$ равны $\pi_3 = -(\alpha+\beta)/(1-\alpha-\beta)$, $\pi_1 = \alpha/(1-\alpha-\beta)$, $\pi_2 = \beta/(1-\alpha-\beta)$ соответственно. Поэтому сумма всех трех коэффициентов $\pi_1+\pi_2+\pi_3 = 0$.

б) Протестируем гипотезу $\mathrm{H}_0$: $\pi_1+\pi_2+\pi_3 = 0$ с помощью $F$-теста. Получаем, что значение $F$-статистики равно 2.564, $p$-значение равно 0.1124.

Нулевая гипотеза не отвергается. Следовательно, можно сказать, что ограничение, накладываемое теорией на коэффициенты регрессии, выполняется.

в) Оценим уравнение (4) с ограничением $\pi_3 = -(\pi_1+\pi_2)$:

$$\ln gdp85 = \pi_0 + \pi_1 \ln iony + \pi_2 \ln sch - (\pi_1+\pi_2)\ln(popgro+2+3).$$

Результаты оценивания см. в таблице 3.18.

**Таблица 3.18**

Dependent Variable: $\ln gdp85$

| Coefficient | Value | Std. Error | $t$-Statistic | Probability |
|---|---|---|---|---|
| $\pi_0$ | 8.066284 | 0.158421 | 50.91678 | 0.0000 |
| $\pi_1$ | 0.593827 | 0.140745 | 4.219157 | 0.0001 |
| $\pi_2$ | 0.690057 | 0.082186 | 8.396294 | 0.0000 |
| $R^2$ | 0.681876 | | | |
| $R^2_{\mathrm{adj}}$ | 0.675699 | | | |

Результаты не сильно отличаются от результатов п. 3.32.4 а). Несколько уменьшилось значение оценки параметра $\pi_1$. Дисперсии оценок также несколько уменьшились, как и $R^2_{\mathrm{adj}}$.

г) Несложные арифметические вычисления приводят к результату:

$$\alpha = \frac{\pi_1}{1+\pi_1+\pi_2}, \quad \beta = \frac{\pi_2}{1+\pi_1+\pi_2}.$$

Подставив вместо $\pi_1$ и $\pi_2$ их оценки, получим:

$$\widehat{\alpha} = \frac{\widehat{\pi}_1}{1+\widehat{\pi}_1+\widehat{\pi}_2} = \frac{0.594}{1+0.594+0.690} = 0.260,$$

$$\widehat{\beta} = \frac{\widehat{\pi}_2}{1+\widehat{\pi}_1+\widehat{\pi}_2} = \frac{0.690}{1+0.594+0.690} = 0.302.$$

Полученные оценки согласуются с другими эмпирическими исследованиями.

3.32.6. а) Добавим в регрессию переменные *noil* и *oecd*. Результаты оценивания представлены в таблице 3.19.

**Таблица 3.19**

Dependent Variable: $\ln gdp85$

| Variable | Coefficient | Std. Error | $t$-Statistic | Probability |
|---|---|---|---|---|
| const | 5.218731 | 1.144819 | 4.558564 | 0.0000 |
| $\ln iony$ | 0.513899 | 0.127919 | 4.017379 | 0.0001 |
| $\ln sch$ | 0.586696 | 0.069818 | 8.403204 | 0.0000 |
| $\ln(popgro + 2 + 3)$ | 0.694190 | 0.516964 | 1.342819 | 0.1824 |
| $noil$ | −0.985646 | 0.188933 | −5.216904 | 0.0000 |
| $oecd$ | 1.010208 | 0.193752 | 5.213914 | 0.0000 |
| $R^2$ | 0.803149 | | | |

Как видно из таблицы, существенно увеличился $R^2$, коэффициенты при новых переменных значимы.

б) Добавляя в регрессию квадраты независимых переменных и перекрестные произведения, улучшить модель из предыдущего пункта не удается (см. таблицу 3.20). Прогнозные свойства модели улучшаются незначительно, а

**Таблица 3.20**

Dependent Variable: $\ln gdp85$

| Variable | Coefficient | Std. Error | $t$-Statistic | Probability |
|---|---|---|---|---|
| const | 4.926086 | 1.783862 | 2.761472 | 0.0069 |
| $\ln iony$ | 0.962252 | 1.174769 | 0.819099 | 0.4147 |
| $\ln^2 iony$ | −0.142147 | 0.248302 | −0.572475 | 0.5683 |
| $\ln sch$ | −0.221084 | 0.461593 | −0.478958 | 0.6330 |
| $\ln^2 sch$ | 0.049805 | 0.070996 | 0.701518 | 0.4847 |
| $\ln iony \cdot \ln sch$ | 0.278729 | 0.197513 | 1.411188 | 0.1614 |
| $\ln(popgro + 2 + 3)$ | 0.744993 | 0.512243 | 1.454374 | 0.1491 |
| $noil$ | −1.042483 | 0.188905 | −5.518553 | 0.0000 |
| $oecd$ | 0.864686 | 0.202116 | 4.278170 | 0.0000 |
| $R^2$ | 0.815298 | | | |

значимость коэффициентов пропадает. Значит, линейная спецификация (4) является удовлетворительной, и отдача физического и человеческого капитала не зависит от размеров экономики.

3.32.7. а) Найдем максимум потребления в стационарном состоянии (удобнее искать максимум логарифма потребления).

$$\begin{aligned}\ln c &= \ln(1 - s_k - s_h) + \ln y \\ &= \ln(1 - s_k - s_h) + \frac{1}{1-\alpha-\beta}(\alpha \ln s_k + \beta \ln s_h - (\alpha+\beta)\ln(n+g+\delta)).\end{aligned}$$

Выпишем условия первого порядка:

$$\frac{\partial \ln c}{\partial s_k} = -\frac{1}{1 - s_k - s_h} + \frac{\alpha}{s_k(1-\alpha-\beta)} = 0,$$
$$\frac{\partial \ln c}{\partial s_h} = -\frac{1}{1 - s_k - s_h} + \frac{\beta}{s_h(1-\alpha-\beta)} = 0.$$

Откуда получаем решение: $s_k = \alpha$, $s_h = \beta$.

б) Среднее значение $iony$ равно 18.16% (или 0.1816, см. упражнение 3.32.1). Оценка параметра $\alpha$, полученная в упражнении 3.32.5, равна $\widehat{\alpha} = 0.260$. То есть получается, что инвестируется меньше, чем следовало бы в стационарном состоянии. Проверим, значимое ли это отличие. Параметр $\alpha$ выражается через параметры регрессии следующим образом:

$$\alpha = \frac{\pi_1}{1+\pi_1+\pi_2}, \quad \frac{\partial \alpha}{\partial \pi} = \left[\frac{1+\pi_2}{(1+\pi_1+\pi_2)^2} \quad -\frac{\pi_1}{(1+\pi_1+\pi_2)^2}\right].$$

В точке $(\widehat{\pi}_1, \widehat{\pi}_2)$ производная равна

$$\left[\frac{1+0.690}{(1+0.593+0.690)^2} \quad -\frac{0.593}{(1+0.593+0.690)^2}\right] = [0.324 \quad -0.114].$$

Оценку ковариационной матрицы $\widehat{\pi}_1$ и $\widehat{\pi}_2$ возьмем из регрессии. Тогда оценка дисперсии $\widehat{\alpha}$ равна

$$\begin{aligned}\widehat{\mathrm{V}}(\widehat{\alpha}) &= \frac{\partial \alpha}{\partial \pi}\bigg|_{(\widehat{\pi}_1, \widehat{\pi}_2)} \widehat{\mathrm{V}}(\widehat{\pi}) \frac{\partial \alpha}{\partial \pi'}\bigg|_{(\widehat{\pi}_1, \widehat{\pi}_2)} \\ &= [0.324 \quad -0.114] \begin{bmatrix} 0.0198 & -0.00729 \\ -0.00729 & 0.00676 \end{bmatrix} \begin{bmatrix} 0.324 \\ -0.114 \end{bmatrix} = 0.002705.\end{aligned}$$

Доверительный интервал для $\alpha$:

$$\alpha \in (0.260 - 1.96\sqrt{0.002705},\ 0.260 + 1.96\sqrt{0.002705}) = (0.158, 0.362).$$

Интервал накрывает значение 0.1816, следовательно, отличие среднего уровня доли инвестиций в ВВП от уровня «золотого правила» статистически незначимо.

3.32.8. а) Оценивая уравнение (5), получаем результаты, приведенные в таблице 3.21.

**Таблица 3.21**

Dependent Variable: $\ln(gdp85/gdp60)$

| Variable | Coefficient | Std. Error | $t$-Statistic | Probability |
|---|---|---|---|---|
| const | 1.903409 | 0.687715 | 2.767729 | 0.0067 |
| $\ln iony$ | 0.552855 | 0.087700 | 6.303923 | 0.0000 |
| $\ln sch$ | 0.216450 | 0.059284 | 3.651085 | 0.0004 |
| $\ln(popgro+2+3)$ | −0.506675 | 0.252321 | −2.008060 | 0.0474 |
| $\ln gdp60$ | −0.297321 | 0.049973 | −5.949603 | 0.0000 |
| $R^2$ | 0.515954 | | | |

Получилась модель, которая соответствует теоретической спецификации. Знаки «правильные», коэффициент при переменной $\ln gdp60$ отрицательный и по абсолютному значению меньше единицы (что означает сходимость к стационарному состоянию, а не расходимость).

б) Если удалить переменную $\ln sch$ из регрессии, то точно так же, как и в упражнении 3.32.4 б), уменьшается $R^2$, и резко увеличивается показатель отдачи физического капитала (см. таблицу 3.22).

**Таблица 3.22**

Dependent Variable: $\ln(gdp85/gdp60)$

| Variable | Coefficient | Std. Error | $t$-Statistic | Probability |
|---|---|---|---|---|
| const | 0.778015 | 0.653765 | 1.190053 | 0.2368 |
| $\ln iony$ | 0.693223 | 0.083682 | 8.284027 | 0.0000 |
| $\ln(popgro+2+3)$ | −0.424770 | 0.265172 | −1.601868 | 0.1123 |
| $\ln(gdp60)$ | −0.182964 | 0.041637 | −4.394231 | 0.0000 |
| $R^2$ | 0.449671 | | | |

3.32.9. а) Аналогично упражнению 3.32.5, получаем, что коэффициенты $\pi_1$, $\pi_2$, $\pi_3$ связаны соотношением $\pi_1+\pi_2+\pi_3=0$.

б) Протестируем гипотезу $H_0$: $\pi_1+\pi_2+\pi_3=0$ с помощью $F$-теста. Получаем, что значение $F$-статистики равно 0.868, $p$-значение равно 0.354.

Нулевая гипотеза не отвергается. Следовательно, можно сказать, что ограничение, накладываемое теорией на коэффициенты регрессии, выполняется.

в) Чтобы учесть ограничение, проведем следующую регрессию:

$$\ln \frac{gdp85}{gdp60} = \pi_0 + \pi_1(\ln s_k - \ln(n+g+\delta)) + \pi_2(\ln s_h - \ln(n+g+\delta)) + \pi_4 \ln gdp60.$$

Результаты представлены в таблице 3.23.

**Таблица 3.23**

Dependent Variable: $\ln(gdp85/gdp60)$

| Variable | Coefficient | Std. Error | $t$-Statistic | Probability |
|---|---|---|---|---|
| const | 2.418306 | 0.409082 | 5.911542 | 0.0000 |
| $\ln iony - \ln(popgro+2+3)$ | 0.516437 | 0.078457 | 6.582447 | 0.0000 |
| $\ln sch - \ln(popgro+2+3)$ | 0.215641 | 0.059238 | 3.640230 | 0.0004 |
| $\ln gdp60$ | −0.293187 | 0.049743 | −5.894023 | 0.0000 |
| $R^2$ | 0.511710 | | | |

Модель с точки зрения точности подгонки практически не изменилась. При этом немного уменьшилась оценка $\widehat{\pi}_1$.

г) Действуя аналогично упражнению 3.32.5 г), получаем, что

$$\lambda = -\ln(1+\pi_4)/t, \quad \alpha = \frac{\pi_1}{\pi_1+\pi_2-\pi_4}, \quad \beta = \frac{\pi_2}{\pi_1+\pi_2-\pi_4}.$$

Подставим в качестве оценок $\pi_1$, $\pi_2$ и $\pi_4$ оценки из модели предыдущего пункта, получим ($t = 25$ — число лет между 1960 и 1985 гг.):

$$\widehat{\lambda} = 0.01388, \quad \widehat{\alpha} = 0.5037, \quad \widehat{\beta} = 0.2103.$$

Получилось, что значение оценки $\widehat{\alpha}$ заметно увеличилось, $\widehat{\beta}$ уменьшилось.

3.32.10. а) Добавим в модель из упражнения 3.32.9 в) переменные $noil$ и $oecd$ (см. таблицу 3.24). Как видим, переменная $noil$ значимая, переменная $oecd$ незначимая.

Коэффициент при переменной $noil$ значимый отрицательный, это имеет ту же интерпретацию, что и раньше: в странах, добывающих нефть, средний доход на душу населения в стационарном состоянии выше, чем в остальных странах (при одинаковом вложении в физический и человеческий капитал).

б) Попытки ввести квадратичные члены в регрессию не приводят к улучшению модели. Например, при добавлении квадратичных и перекрестных членов в модель из упражнения 3.32.8 а) получается регрессия, представленная в таблице 3.25.

**Таблица 3.24**

Dependent Variable: $\ln(gdp85/gdp60)$

| Variable | Coefficient | Std. Error | $t$-Statistic | Probability |
|---|---|---|---|---|
| const | 3.284511 | 0.496377 | 6.616971 | 0.0000 |
| $\ln iony - \ln(popgro + 2 + 3)$ | 0.504614 | 0.084082 | 6.001461 | 0.0000 |
| $\ln sch - \ln(popgro + 2 + 3)$ | 0.236528 | 0.057690 | 4.100016 | 0.0001 |
| $\ln gdp60$ | −0.353130 | 0.053697 | −6.576388 | 0.0000 |
| $noil$ | −0.438120 | 0.147506 | −2.970178 | 0.0037 |
| $oecd$ | 0.123944 | 0.111970 | 1.106934 | 0.2710 |
| $R^2$ | 0.552974 | | | |

**Таблица 3.25**

Dependent Variable: $\ln(gdp85/gdp60)$

| Variable | Coefficient | Std. Error | $t$-Statistic | Probability |
|---|---|---|---|---|
| const | 2.779838 | 1.141940 | 2.434312 | 0.0168 |
| $\ln iony$ | 0.307298 | 0.650325 | 0.472529 | 0.6376 |
| $\ln^2 iony$ | 0.041652 | 0.124459 | 0.334661 | 0.7386 |
| $\ln sch$ | 0.192573 | 0.082431 | 2.336180 | 0.0216 |
| $\ln^2 sch$ | 0.032097 | 0.039365 | 0.815380 | 0.4169 |
| $\ln(popgro + 2 + 3)$ | −0.257705 | 0.354600 | −0.726749 | 0.4692 |
| $\ln gdp60$ | −0.390155 | 0.058379 | −6.683204 | 0.0000 |
| $noil$ | −0.436043 | 0.150610 | −2.895180 | 0.0047 |
| $oecd$ | 0.226884 | 0.149376 | 1.518872 | 0.1321 |
| $R^2$ | 0.565356 | | | |

3.32.11. а) Для удобства сравнения коэффициентов оценим две разные модели в одном уравнении (см. таблицу 3.26).

Оценки коэффициентов различаются для стран — членов OECD и стран, не входящих в OECD, но ввиду того, что стран — членов OECD всего 22, эти различия ненадежны.

б) Протестируем гипотезу о равенстве коэффициентов при соответствующих переменных для стран — членов OECD и стран — нечленов OECD (т. е. о равенстве коэффициентов при переменных $oecd \cdot \ln iony$ и $(1 - oecd) \cdot \ln iony$, $oecd \cdot \ln sch$ и $(1 - oecd) \cdot \ln sch$, $oecd \cdot \ln(popgro + 2 + 3)$ и $(1 - oecd) \cdot \ln(popgro + 2 + 3)$, $oecd \cdot \ln gdp60$ и $(1 - oecd) \cdot \ln gdp60$). $F$-статистика равна 0.369 и подчиняется распределению Фишера с параметрами $(4, 94)$, $p$-значение равно 0.830.

**Таблица 3.26**

Dependent Variable: $\ln(gdp85/gdp60)$

| Variable | Coefficient | Std. Error | $t$-Statistic | Probability |
|---|---|---|---|---|
| $oecd$ | 4.154971 | 2.317804 | 1.792633 | 0.0762 |
| $oecd \cdot \ln iony$ | 0.331793 | 0.402721 | 0.823877 | 0.4121 |
| $oecd \cdot \ln sch$ | 0.227699 | 0.336812 | 0.676041 | 0.5007 |
| $oecd \cdot \ln(popgro + 2 + 3)$ | −0.863414 | 0.784378 | −1.100762 | 0.2738 |
| $oecd \cdot \ln gdp60$ | −0.397690 | 0.162964 | −2.440363 | 0.0165 |
| $(1 - oecd)$ | 1.102103 | 0.828098 | 1.330885 | 0.1864 |
| $(1 - oecd) \cdot \ln iony$ | 0.539122 | 0.092180 | 5.848570 | 0.0000 |
| $(1 - oecd) \cdot \ln sch$ | 0.212730 | 0.061449 | 3.461898 | 0.0008 |
| $(1 - oecd) \cdot \ln(popgro + 2 + 3)$ | 0.043841 | 0.419781 | 0.104438 | 0.9170 |
| $(1 - oecd) \cdot \ln gdp60$ | −0.336243 | 0.061835 | −5.437731 | 0.0000 |
| $R^2$ | 0.533136 | | | |

Таким образом, гипотеза о совпадении моделей роста для стран — членов OECD и стран — не членов OECD не отвергается.

3.32.12. Рассмотрим, например, модель из табл. 3.27 (компромисс между простотой, точностью подгонки и теоретическими ограничениями):

**Таблица 3.27**

Dependent Variable: $\ln(gdp85/gdp60)$

| Variable | Coefficient | Std. Error | $t$-Statistic | Probability |
|---|---|---|---|---|
| const | 3.080358 | 0.461366 | 6.676606 | 0.0000 |
| $\ln iony - \ln(popgro + 2 + 3)$ | 0.543383 | 0.076527 | 7.100529 | 0.0000 |
| $\ln sch - \ln(popgro + 2 + 3)$ | 0.235029 | 0.057739 | 4.070524 | 0.0001 |
| $\ln gdp60$ | −0.331410 | 0.050040 | −6.622918 | 0.0000 |
| $noil$ | −0.402689 | 0.144155 | −2.793437 | 0.0063 |
| $R^2$ | 0.547385 | | | |

Оценка параметра $\lambda$ выражается через коэффициенты модели следующим образом:

$$\widehat{\lambda} = -\frac{1}{t}\ln(1 + \widehat{\pi}_4),$$

где $\widehat{\pi}_4$ — оценка коэффициента при $\ln gdp60$, $t = 25$ — число лет между 1960 и 1985 гг.

Оценку дисперсии $\widehat{\lambda}$ можно получить из того, что $\widehat{\pi}_4$ является асимптотически нормальной величиной:

$$\sqrt{n}(\widehat{\pi}_4 - \pi_4) \xrightarrow{d} N(0, V).$$

В этом случае, если $\lambda = f(\pi_4)$ и $\widehat{\lambda} = f(\widehat{\pi}_4)$, то

$$\sqrt{n}(\widehat{\lambda} - \lambda) \xrightarrow{d} N(0, (f'(\pi_4))^2 V).$$

В нашем случае

$$f(\pi_4) = -\frac{1}{t}\ln(1 + \widehat{\pi}_4), \quad \text{и} \quad f'(\pi_4) = -\frac{1}{t(1 + \widehat{\pi}_4)}.$$

Оценка $\widehat{\lambda}$ равна 0.0161. Оценка дисперсии оценки $\widehat{\lambda}$ равна

$$\widehat{\mathrm{V}}(\widehat{\lambda}) = \frac{\widehat{\mathrm{V}}(\widehat{\pi}_4)}{t^2(1 + \widehat{\pi}_4)^2} = \frac{0.002504}{25^2 \cdot 0.66859^2} = 8.9626 \cdot 10^{-6}.$$

А значит, доверительный интервал можно построить следующим образом:

$$\lambda \in \left(0.0161 - \sqrt{8.9626 \cdot 10^{-6}}, 0.0161 + \sqrt{8.9626 \cdot 10^{-6}}\right) = (0.0131, 0.019).$$

Таким образом, скорость сходимости к стационарному состоянию невысокая.

## Задача 3.33

Рассматривается информация о стоимости коттеджей в Московской области по Киевскому направлению (по данным строительной компании «Стройсервис», осень 1997 г.).

Данные находятся в файле `villa.xls`. Переменные описаны в таблице 3.28.

**Таблица 3.28**

| Переменная | Описание |
|---|---|
| $n$ | номер по порядку |
| $price$ | цена в тыс. долл. |
| $dist$ | расстояние от кольцевой автодороги в км |
| $house$ | площадь дома в кв. м |
| $area$ | площадь участка в сотках |

Подберите функциональную форму зависимости цены коттеджа от его параметров, учитывая такие факторы, как $t$-статистики и коэффициент детерминации $R^2$.

## Решение

Сначала исследуем описательные статистики представленных переменных (см. таблицу 3.29).

**Таблица 3.29**

| | *price* | *dist* | *house* | *area* |
|---|---|---|---|---|
| Mean | 78.25 | 44.05 | 192.24 | 13.75 |
| Median | 46.00 | 30.00 | 160.00 | 14.00 |
| Maximum | 320.00 | 105.00 | 600.00 | 40.00 |
| Minimum | 5.00 | 0.50 | 22.00 | 5.00 |
| Std. Dev. | 84.343 | 28.80 | 151.85 | 6.91 |

Как видно из таблицы, в данных представлены довольно разные коттеджи — от маленьких дачных домиков до больших загородных домов. Площадь участка тоже варьирует в широких пределах. Все статистики выглядят правдоподобно.

Корреляционная матрица (см. таблицу 3.30) показывает, что, как и ожидалось, цена положительно коррелирована с площадью участка и площадью дома и отрицательно коррелирована с расстоянием до кольцевой автодороги.

**Таблица 3.30**

| | *price* | *dist* | *house* | *area* |
|---|---|---|---|---|
| *price* | 1.000 | −0.547 | 0.674 | 0.577 |
| *dist* | −0.547 | 1.000 | −0.541 | −0.160 |
| *house* | 0.674 | −0.541 | 1.000 | 0.603 |
| *area* | 0.577 | −0.160 | 0.603 | 1.000 |

Рассмотрим теперь четыре диаграммы рассеивания для предварительного отбора функциональной формы. На первой (см. рис. 3.2) изображена связь переменных *house* и *price* (линейная модель), на второй (см. рис. 3.3) — $\ln house$ и *price*, на третьей (см. рис. 3.4) — *house* и $\ln price$ (полулогарифмическая модель) и на четвертой (см. рис. 3.5) — $\ln house$ и $\ln price$ (логарифмическая модель).

Как видим из графиков, логарифмическая модель лучше подходит для описания связи между площадью дома и его ценой.

Посмотрим теперь на формально построенные четыре модели. Линейная (см. таблицу 3.31 а) и модель зависимости цены от логарифмов независимых переменных (см. таблицу 3.31 в) нехороши тем, что коэффициенты при переменной *house* в обеих моделях оказались незначимыми (на 5%-ном

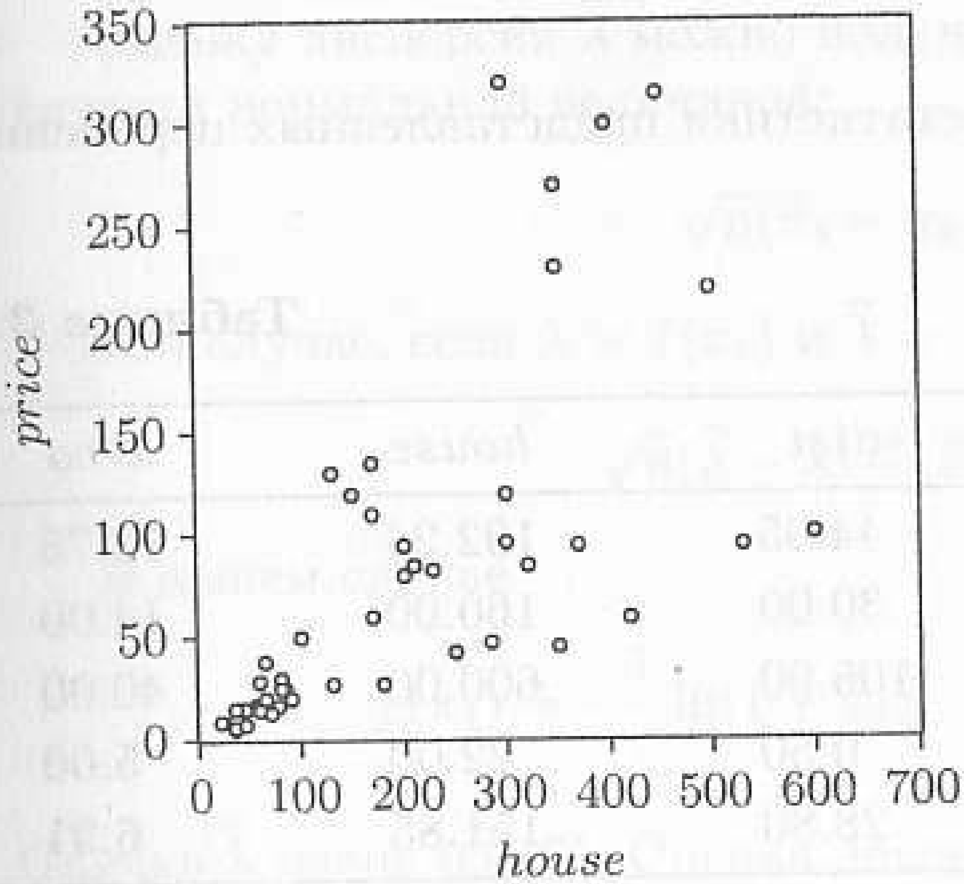



**Рис. 3.2**

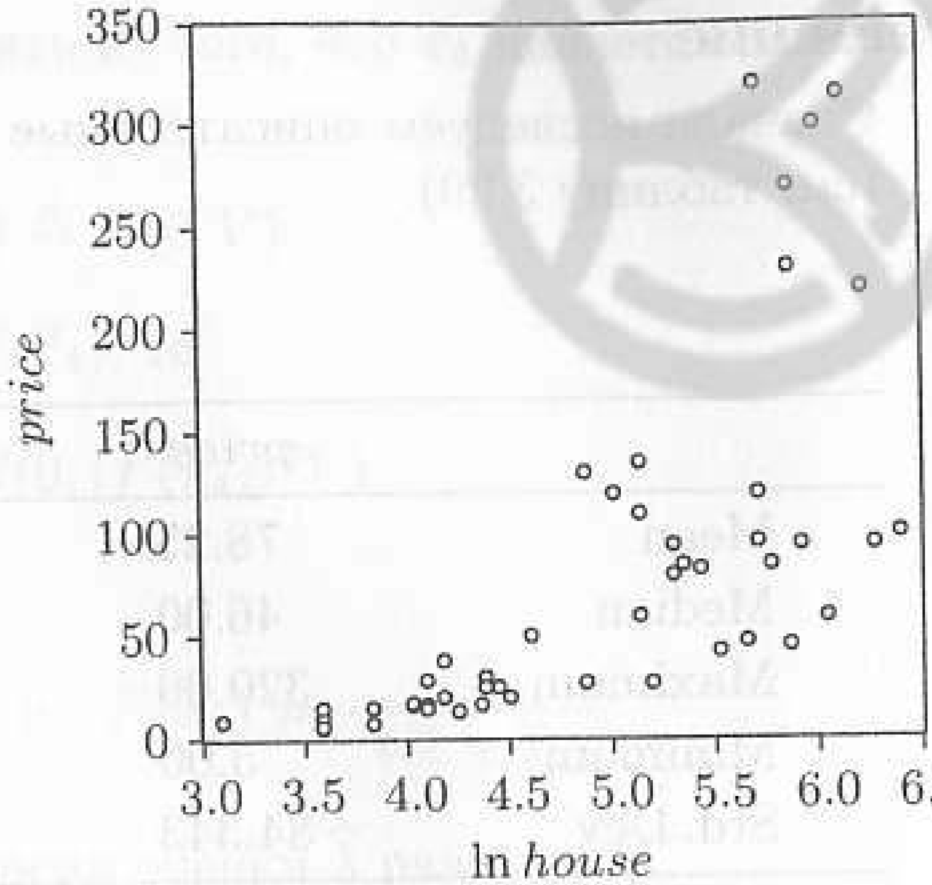



**Рис. 3.3**

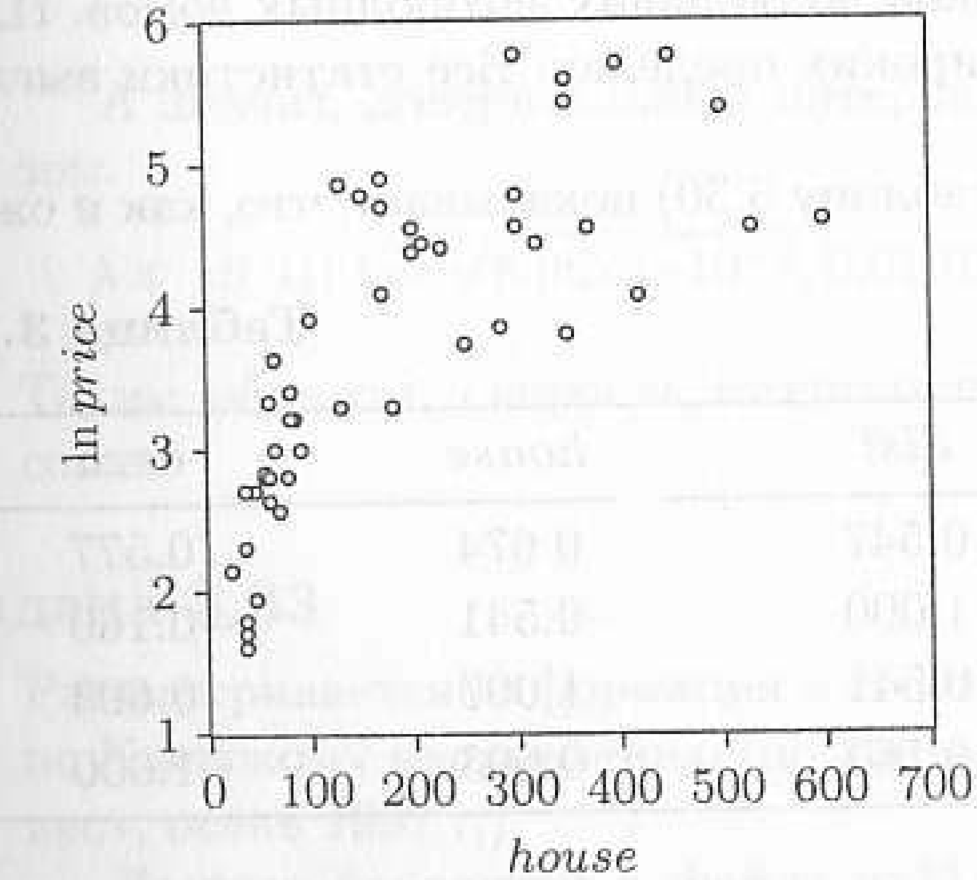



**Рис. 3.4**

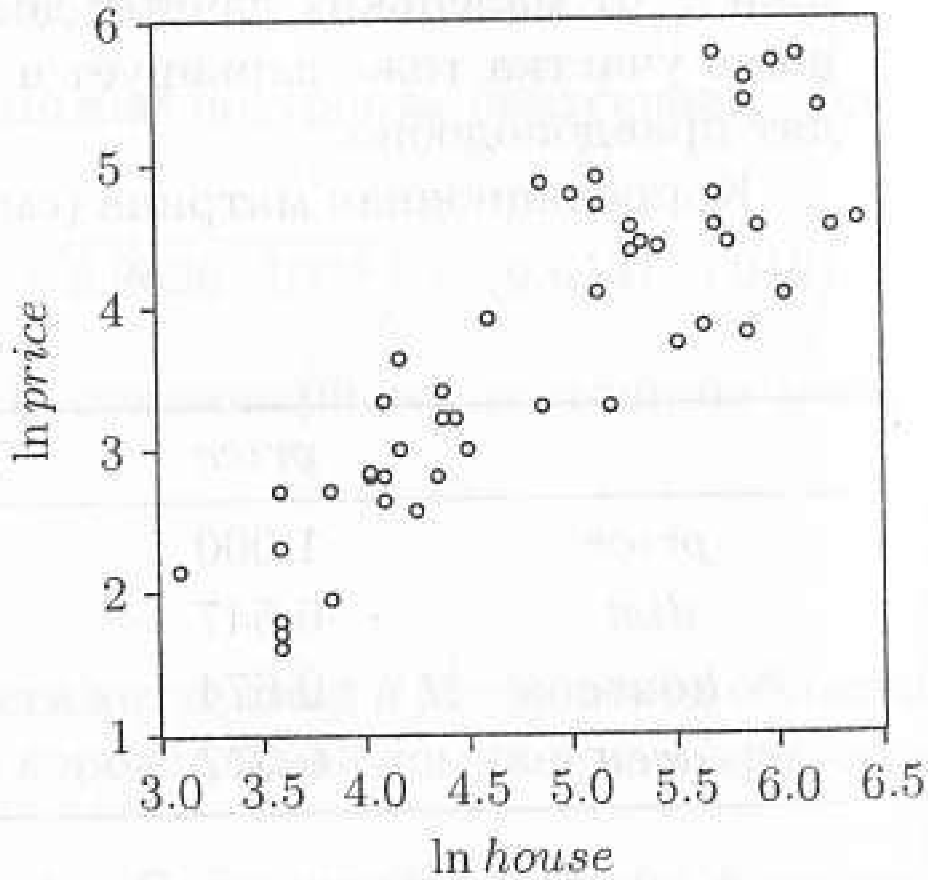



**Рис. 3.5**

уровне). Очень странно выглядит модель, в которой цена коттеджа не зависит от площади дома. Полулогарифмическая модель (см. таблицу 3.31 б) лишена этого недостатка, но если сравнивать ее с логарифмической (см. таблицу 3.31 г), то можно заметить, что переход к логарифмической модели значительно увеличивает $R^2$. (Заметим, что в логарифмической модели переменная *dist* входит в правую часть уравнения линейно. Это продиктовано улучшением прогнозных свойств модели.) Следовательно, логарифмическая модель выглядит предпочтительнее в данной задаче.

В этой модели коэффициенты при $\ln house$ и $\ln area$ имеют смысл эластичностей цены по соответствующей переменной. Так, увеличение площади дома на 1% приводит (в среднем и при прочих равных) к увеличению цены коттеджа на 0.67%. Коэффициент при *dist* интерпретируется иначе. Увеличение расстояния от кольцевой автодороги на 1 км приводит (в среднем и

Таблица 3.31

| | а) Dependent Var.: *price* | | | б) Dependent Var.: ln *price* | | |
|---|---|---|---|---|---|---|
| Variable | Coeff. | *t*-Stat. | Prob. | Coeff. | *t*-Stat. | Prob. |
| const | 33.385 | 1.2836 | 0.206 | 3.4519 | 12.732 | 0.000 |
| *house* | 0.1513 | 1.8693 | 0.068 | 0.0027 | 3.1953 | 0.003 |
| *area* | 4.3612 | 2.8795 | 0.006 | 0.0452 | 2.8599 | 0.006 |
| *dist* | −1.0033 | −2.9102 | 0.006 | −0.0183 | −5.0947 | 0.000 |
| $R^2$ | 0.5772 | | | 0.7507 | | |

| | в) Dependent Var.: *price* | | | г) Dependent Var.: ln *price* | | |
|---|---|---|---|---|---|---|
| Variable | Coeff. | *t*-Stat. | Prob. | Coeff. | *t*-Stat. | Prob. |
| const | −56.116 | −0.7863 | 0.436 | −0.2747 | −0.5005 | 0.619 |
| ln *house* | 21.459 | 1.7951 | 0.079 | 0.6672 | 5.5457 | 0.000 |
| ln *area* | 70.786 | 3.7222 | 0.001 | 0.5688 | 3.1474 | 0.003 |
| *dist* | – | – | – | −0.0145 | −4.6292 | 0.000 |
| ln *dist* | −42.248 | −4.2180 | 0.000 | – | – | – |
| $R^2$ | 0.6576 | | | 0.8430 | | |

при прочих равных) к тому, что цена умножается на $e^{-0.0145}$ или уменьшается приблизительно на 1.4%.

# Глава 4

# Различные аспекты множественной регрессии

## Задача 4.1

С помощью бинарных переменных напишите уравнение, соответствующее наличию двух структурных изменений в моменты времени $t_0$ и $t_1$ (предполагается, что $t_0 < t_1$).

## Решение

Пусть $y$ — зависимая переменная и пусть независимых переменных только две: $x$ и константа. Предположим, что $x$ и $y$ представлены в виде временных рядов $\{(x_t, y_t),\ t = 1, \ldots, n\}$. Из некоторых соображений исследователь считает, что имели место два структурных изменения: в моменты времени $t_0$ и $t_1$, $t_0 < t_1$.

Чтобы оценить такую модель, введем две бинарные переменные $R$ и $S$, полагая $R_t = 0$, если $t \leqslant t_0$, $R_t = 1$, если $t > t_0$; $S_t = 0$, если $t \leqslant t_1$, $S_t = 1$, если $t > t_1$, и запишем следующее регрессионное уравнение:

$$y_t = \beta_0 + \beta_1 x_t + \beta_2 (x_t - x_{t_0}) R_t + \beta_3 (x_t - x_{t_1}) S_t + \varepsilon_t.$$

Регрессионная линия, соответствующая этому уравнению, имеет коэффициент наклона $\beta_1$ для $t \leqslant t_0$, $\beta_1 + \beta_2$ для $t_0 < t \leqslant t_1$ и $\beta_1 + \beta_2 + \beta_3$ для $t > t_1$. Разрыва в точках $x_{t_0}$ и $x_{t_1}$ не происходит.

Таким образом, для проверки предположения об отсутствии структурного изменения в момент времени $t_0$ необходимо протестировать гипотезу $\mathrm{H}_0$: $\beta_2 = 0$, в момент времени $t_1$ — гипотезу $\mathrm{H}_0$: $\beta_3 = 0$ и, наконец, чтобы проверить предположение об отсутствии и того, и другого структурного сдвига, надо протестировать гипотезу $\mathrm{H}_0$: $\beta_2 = \beta_3 = 0$.

## Задача 4.2

Докажите равенство (4.8):

$$r(y, x_1 \mid x_2) = \frac{r(y, x_1) - r(y, x_2) r(x_1, x_2)}{\sqrt{1 - r^2(x_1, x_2)}\sqrt{1 - r^2(y, x_2)}}.$$

## Решение

(Все обозначения указаны в разд. 4.3 книги и разд. 6 приложения МС.) Согласно определению,

$$r(y, x_1 \mid x_2) = r\left(e_y, e_{x_1}\right) = \frac{\operatorname{Cov}\left(e_y, e_{x_1}\right)}{\sqrt{\operatorname{Var}(e_y)}\sqrt{\operatorname{Var}(e_{x_1})}}. \qquad (*)$$

Распишем числитель $(*)$ в виде

$$\begin{aligned}\operatorname{Cov}\left(e_y, e_{x_1}\right) &= \operatorname{Cov}(y - \widehat{y}, x_1 - \widehat{x}_1) = \operatorname{Cov}(y, x_1) \\ &\quad - \operatorname{Cov}(y, \widehat{x}_1) - \operatorname{Cov}(\widehat{y}, x_1) + \operatorname{Cov}(\widehat{y}, \widehat{x}_1) \\ &= \operatorname{Cov}(y, x_1) - \widehat{\gamma}_2 \operatorname{Cov}(y, x_2) - \widehat{\alpha}_2 \operatorname{Cov}(x_2, x_1) + \widehat{\alpha}_2 \widehat{\gamma}_2 \operatorname{Var}(x_2).\end{aligned}$$

Замечаем, что коэффициенты равны

$$\widehat{\alpha}_2 = \frac{\operatorname{Cov}(y, x_2)}{\operatorname{Var}(x_2)} \quad \text{и} \quad \widehat{\gamma}_2 = \frac{\operatorname{Cov}(x_1, x_2)}{\operatorname{Var}(x_2)},$$

и подставляем их значения в полученное выражение:

$$\begin{aligned}\operatorname{Cov}\left(e_y, e_{x_1}\right) &= \operatorname{Cov}(y, x_1) - \frac{\operatorname{Cov}(x_1, x_2)}{\operatorname{Var}(x_2)} \operatorname{Cov}(y, x_2) \\ &\quad - \frac{\operatorname{Cov}(y, x_2)}{\operatorname{Var}(x_2)} \operatorname{Cov}(x_2, x_1) \\ &\quad + \frac{\operatorname{Cov}(y, x_2)}{\operatorname{Var}(x_2)} \frac{\operatorname{Cov}(x_1, x_2)}{\operatorname{Var}(x_2)} \operatorname{Var}(x_2) \\ &= \operatorname{Cov}(y, x_1) - \frac{\operatorname{Cov}(x_1, x_2)}{\operatorname{Var}(x_2)} \operatorname{Cov}(y, x_2).\end{aligned}$$

Далее, найдем знаменатель $(*)$:

$$\begin{aligned}\operatorname{Var}(e_y) &= \operatorname{Var}(y - \widehat{y}) \\ &= \operatorname{Var}(y) - 2\frac{\operatorname{Cov}(y, x_2)}{\operatorname{Var}(x_2)} \operatorname{Cov}(y, x_2) + \left(\frac{\operatorname{Cov}(y, x_2)}{\operatorname{Var}(x_2)}\right)^2 \operatorname{Var}(x_2) \\ &= \operatorname{Var}(y) - \frac{\operatorname{Cov}^2(y, x_2)}{\operatorname{Var}(x_2)};\end{aligned}$$

$$\operatorname{Var}(e_{x_1}) = \operatorname{Var}(x_1) - \frac{\operatorname{Cov}^2(x_1, x_2)}{\operatorname{Var}(x_2)}.$$

Окончательно имеем:

$$r(y,x_1 \mid x_2) = \frac{\operatorname{Cov}(y,x_1) - \dfrac{\operatorname{Cov}(x_1,x_2)\operatorname{Cov}(y,x_2)}{\operatorname{Var}(x_2)}}{\sqrt{\operatorname{Var}(x_1) - \dfrac{\operatorname{Cov}^2(x_1,x_2)}{\operatorname{Var}(x_2)}}\sqrt{\operatorname{Var}(y) - \dfrac{\operatorname{Cov}^2(y,x_2)}{\operatorname{Var}(x_2)}}}$$

$$= \frac{\dfrac{\operatorname{Cov}(y,x_1)}{\sqrt{\operatorname{Var}(x_1)\operatorname{Var}(y)}} - \dfrac{\operatorname{Cov}(x_1,x_2)\operatorname{Cov}(y,x_2)}{\sqrt{\operatorname{Var}(x_1)\operatorname{Var}(y)}\sqrt{\operatorname{Var}(x_2)\operatorname{Var}(x_2)}}}{\sqrt{\dfrac{\operatorname{Var}(x_1)}{\operatorname{Var}(x_1)} - \dfrac{\operatorname{Cov}^2(x_1,x_2)}{\operatorname{Var}(x_1)\operatorname{Var}(x_2)}}\sqrt{\dfrac{\operatorname{Var}(y)}{\operatorname{Var}(y)} - \dfrac{\operatorname{Cov}^2(y,x_2)}{\operatorname{Var}(y)\operatorname{Var}(x_2)}}}$$

$$= \frac{r(y,x_1) - r(y,x_2)r(x_1,x_2)}{\sqrt{1-r^2(x_1,x_2)}\sqrt{1-r^2(y,x_2)}},$$

что и требовалось доказать.

## Задача 4.3

Докажите эквивалентность формул (4.19)

$$\widehat{\beta} = (X'X)^{-1}X'y - (X'X)^{-1}X'Z(Z'MZ)^{-1}Z'My,$$

где $M = I - X(X'X)^{-1}X'$, и (4.20)

$$\widehat{\beta} = (X'M_ZX)^{-1}X'M_Zy, \text{ где } M_Z = I - Z(Z'Z)^{-1}Z',$$

для оценки по модели (4.12б) $y = X\beta + Z\gamma + \varepsilon$.

(*Указание.* Первое выражение получается из блочного представления матриц; второе получается после замены $X$ и $y$ на их остатки при регрессии на $Z$.)

## Решение

Пусть $(\widehat{\beta},\widehat{\gamma})'$ — МНК-оценка вектора параметров $(\beta,\gamma)'$ в модели (4.12б). Тогда

$$\begin{bmatrix}\widehat{\beta}\\ \widehat{\gamma}\end{bmatrix} = ([X \quad Z]'\,[X \quad Z])^{-1}[X \quad Z]'\,y = \begin{bmatrix}X'X & X'Z\\ Z'X & Z'Z\end{bmatrix}^{-1}\begin{bmatrix}X'y\\ Z'y\end{bmatrix}.$$

Найдем матрицу

$$\begin{bmatrix}X'X & X'Z\\ Z'X & Z'Z\end{bmatrix}^{-1} = \begin{bmatrix}A & B\\ B' & C\end{bmatrix}.$$

Из формулы (ЛА.17) получаем:

$$
\begin{aligned}
C &= (Z'Z - Z'X(X'X)^{-1}X'Z)^{-1} = (Z'(I - X(X'X)^{-1}X')Z)^{-1} \\
&= (Z'MZ)^{-1}, \quad \text{где} \quad M = I - X(X'X)^{-1}X', \\
B &= -(X'X)^{-1}(X'Z)(Z'MZ)^{-1}.
\end{aligned}
$$

Из равенства $A(X'X) + B(Z'X) = I$ получаем:

$$
\begin{aligned}
A &= (I - B(Z'X))(X'X)^{-1} \\
&= \left(I + (X'X)^{-1}(X'Z)(Z'MZ)^{-1}(Z'X)\right)(X'X)^{-1}.
\end{aligned}
$$

Тогда (4.19) получается из

$$
\begin{aligned}
\widehat{\beta} &= (AX' + BZ')y \\
&= ((X'X)^{-1}X' + (X'X)^{-1}(X'Z)(Z'MZ)^{-1}(Z'X)(X'X)^{-1}X' \\
&\quad - (X'X)^{-1}(X'Z)(Z'MZ)^{-1}Z')y \\
&= (X'X)^{-1}X'y - (X'X)^{-1}(X'Z)(Z'MZ)^{-1}Z'(I - X(X'X)^{-1}X')y \\
&= (X'X)^{-1}X'y - (X'X)^{-1}(X'Z)(Z'MZ)^{-1}Z'My.
\end{aligned}
$$

Докажем (4.20). Остатки регрессии $y$ и $X$ на $Z$ равны соответственно $e_X = M_Z X$ и $e_y = M_Z y$, где $M_Z = I - Z(Z'Z)^{-1}Z'$. Применим к оцененному уравнению регрессии $y = X\widehat{\beta} + Z\widehat{\gamma} + e$ оператор проектирования на подпространство, ортогональное $Z$: $M_Z y = M_Z X\widehat{\beta} + e$ (мы используем равенство $M_Z Z = \mathbf{0}$ и $M_Z e = e$). Отсюда

$$
\widehat{\beta} = ((M_Z X)'(M_Z X))^{-1}(M_Z X)'M_Z y = (X'M_Z X)^{-1}X'M_Z y.
$$

## Задача 4.4

Предположим, что вы оцениваете линейную функцию потребления $c_t = \alpha + \beta y_t + \varepsilon_t$ среди $n$ индивидуумов. Как учесть возможный сдвиг этой функции при переходе от городского к сельскому потребителю, если вы считаете, что предельная (маржинальная) склонность к потреблению постоянна, в то время как средняя склонность к потреблению может меняться? Как проверить гипотезу о том, что предельные склонности к потреблению индивидуумов с доходом выше и ниже уровня $y^*$ отличаются?

## Решение

Если функция потребления линейна и имеет вид $c_t = \alpha + \beta y_t + \varepsilon_t$, предельная склонность к потреблению, т. е. прирост потребления при увеличении дохода на единицу, характеризуется коэффициентом $\beta (= \partial c/\partial y)$. Средняя склонность к потреблению есть доля дохода, идущая на потребление, т. е. $c/y$, она определяется не только коэффициентом $\beta$, но и $\alpha$. Таким образом, чтобы учесть тот факт, что предельная склонность к потреблению ($\beta$) постоянна, в то время как средняя склонность к потреблению может меняться (т. е. может меняться $\alpha$) при переходе от городского к сельскому потребителю, введем фиктивную переменную $r$, полагая $r_t = 0$, если наблюдение $t$ соответствует городскому населению, $r_t = 1$ — сельскому. Получим модель

$$c_t = \alpha + \beta y_t + \gamma r_t + \varepsilon_t.$$

Проверка данного утверждения сводится к проверке гипотезы $\mathrm{H}_0$: $\gamma = 0$.

Введем фиктивную переменную $s$: $s_t = 0$, если $y_t \leqslant y^*$, и $s_t = 1$, если $y_t > y^*$. Для проверки гипотезы о том, что предельные склонности к потреблению индивидуумов с доходом выше и ниже уровня $y^*$ отличаются, может быть использована модель:

$$c_t = \alpha + \beta y_t + \gamma(y_t - y^*)s_t + \varepsilon_t.$$

Проверка утверждения сводится к проверке гипотезы $\mathrm{H}_0$: $\gamma = 0$. Действительно, в этой модели предельная склонность к потреблению равна $\beta$, если доход не превосходит $y^*$, и равна $\beta + \gamma$ в противном случае.

## Задача 4.5

Рассмотрим регрессию

$$y_t = \beta_1 + \beta_2 d_t + \varepsilon_t, \quad t = 1, \ldots, n,$$

где $d$ — некоторая фиктивная переменная. Пусть $\overline{y}_0$ — среднее значение переменной $y$ по $n_0$ наблюдениям, для которых $d = 0$, и $\overline{y}_1$ — среднее значение переменной $y$ по $n_1$ наблюдениям, для которых $d = 1$ ($n_0 + n_1 = n$). Найдите $\mathrm{V}(\widehat{\beta}_1)$, $\mathrm{V}(\widehat{\beta}_2)$.

## Решение

Без ограничения общности можно считать, что наблюдения упорядочены так, что $d_t = 0$, $t = 1, \ldots, n_0$; $d_t = 1$, $t = n_0 + 1, \ldots, n_0 + n_1$. По формуле для оценки вектора $\boldsymbol{\beta}$ по методу наименьших квадратов имеем $\boldsymbol{\beta} = (\boldsymbol{X}'\boldsymbol{X})^{-1}\boldsymbol{X}'\boldsymbol{y}$, где

$$\boldsymbol{X} = [\imath \quad d] = \begin{bmatrix} 1 & 0 \\ \ldots & \ldots \\ 1 & 0 \\ 1 & 1 \\ \ldots & \ldots \\ 1 & 1 \end{bmatrix}$$

(во втором столбце первые $n_0$ компонент равны нулю, остальные $n_1$ — единице). Тогда

$$X'X = \begin{bmatrix} n & n_1 \\ n_1 & n_1 \end{bmatrix}, \quad (X'X)^{-1} = \frac{1}{n_0 n_1}\begin{bmatrix} n_1 & -n_1 \\ -n_1 & n \end{bmatrix},$$

$$X'y = \begin{bmatrix} \overline{y}n \\ \overline{y}_1 n_1 \end{bmatrix} = \begin{bmatrix} \overline{y}_0 n_0 + \overline{y}_1 n_1 \\ \overline{y}_1 n_1 \end{bmatrix}.$$

Следовательно,

$$\widehat{\beta} = \begin{bmatrix} \widehat{\beta}_1 \\ \widehat{\beta}_2 \end{bmatrix} = \begin{bmatrix} \overline{y}_0 \\ \overline{y}_1 - \overline{y}_0 \end{bmatrix}.$$

Пользуясь некоррелированностью ошибок для разных наблюдений (а значит, и некоррелированностью $y_s$, $y_t$ при $s \neq t$), получаем:

$$\mathrm{V}(\widehat{\beta}_1) = \mathrm{V}(\overline{y}_0) = \frac{\sum_{t=1}^{n_0} \mathrm{V}(y_t)}{n_0^2}.$$

Как известно, $\mathrm{V}(y_t) = \sigma^2,\ \forall t = 1, \ldots, n$. Значит, $\mathrm{V}(\widehat{\beta}_1) = \sigma^2/n_0$.

Аналогично, $\mathrm{V}(\overline{y}_1) = \sigma^2/n_1$. Снова используя некоррелированность ошибок, получим $\mathrm{V}(\widehat{\beta}_2) = \mathrm{V}(\overline{y}_1 - \overline{y}_0) = \mathrm{V}(\overline{y}_1) + \mathrm{V}(\overline{y}_0)$.

Таким образом, окончательный ответ:

$$\mathrm{V}(\widehat{\beta}_1) = \frac{\sigma^2}{n_0}, \quad \mathrm{V}(\widehat{\beta}_2) = \frac{n\sigma^2}{n_0 n_1}.$$

## Задача 4.6

На основе квартальных данных с 1971 по 1976 г. с помощью метода наименьших квадратов получено следующее уравнение:

$$y_t = \underset{(2.14)}{1.12} - \underset{(0.0034)}{0.0098}x_{t1} - \underset{(3.42)}{5.62}x_{t2} + \underset{(0.009)}{0.044}x_{t3},$$

в скобках указаны стандартные ошибки, RSS = 110.32, ESS = 21.43.

а) Проверьте значимость каждого из коэффициентов.

б) Найдите коэффициент детерминации.

в) Протестируйте значимость регрессии в целом.

г) Когда в уравнение были добавлены три фиктивные переменные, соответствующие трем первым кварталам года, величина RSS выросла до 118.20. Проверьте гипотезу о наличии сезонности, сформулировав необходимые предположения о виде этой сезонности.

д) Для той же исходной модели были раздельно проведены две регрессии на основе данных: 1-й квартал 1971 г. — 1-й квартал 1975 г. и 2-й квартал 1975 г. — 4-й квартал 1976 г. соответственно и получены следующие значения сумм квадратов остатков: $\mathrm{ESS}_1 = 12.25$, $\mathrm{ESS}_2 = 2.32$. Проверьте гипотезу о том, что между 1-м и 2-м кварталами 1975 г. произошло структурное изменение.

## Решение

а) Заметим, что $n = 24$ (6 лет по 4 квартала), $k = 4$. Для проверки значимости каждого коэффициента тестируем гипотезы $\mathrm{H}_0$: $\beta_i = 0$, $i = 1, 2, 3, 4$ (на 5%-ном уровне). Вычисляем $t$-статистику $t_i = \widehat{\beta}_i/\widehat{\sigma}_{\widehat{\beta}_i}$ и сравниваем ее с $t_{0.025}(20) = 2.086$:

$$|t_1| = \frac{1.12}{2.14} = 0.523 < 2.086,$$

$$|t_2| = \frac{0.0098}{0.0034} = 2.882 > 2.086,$$

$$|t_3| = \frac{5.62}{3.42} = 1.643 < 2.086,$$

$$|t_4| = \frac{0.044}{0.009} = 4.889 > 2.086.$$

Значит, статистически незначимыми (на 5%-ном уровне) можем считать первый и третий коэффициенты, статистически значимыми — второй и четвертый.

б) Имеем

$$R^2 = \frac{\mathrm{RSS}}{\mathrm{TSS}} = \frac{\mathrm{RSS}}{\mathrm{ESS} + \mathrm{RSS}} = \frac{110.32}{131.75} = 0.837.$$

в) Тестируем гипотезу $\mathrm{H}_0$: $\beta_2 = \beta_3 = \beta_4 = 0$. Воспользуемся статистикой (3.36):

$$F = \frac{R^2}{1 - R^2} \cdot \frac{n-k}{k-1} = \frac{\mathrm{RSS}}{\mathrm{ESS}} \cdot \frac{n-k}{k-1} = \frac{110.32}{21.43} \cdot \frac{20}{3} = 34.319 > 3.098 = F_{0.05}(3, 20).$$

Значит, регрессию в целом считаем статистически значимой (на 5%-ном уровне).

г) В условии сказано, что в уравнение добавлено три фиктивные сезонные переменные. Тем самым предполагается, что возможны сезонные колебания среднего значения $y$, а маржинальные характеристики не меняются. Иными словами, теперь рассматривается уравнение

$$y_t = \beta_1 + \beta_2 x_{t1} + \beta_3 x_{t2} + \beta_4 x_{t3} + \beta_5 d_{t1} + \beta_6 d_{t2} + \beta_7 d_{t3} + \varepsilon_t.$$

Для этого уравнения тестируем гипотезу $\mathrm{H}_0$: $\beta_5 = \beta_6 = \beta_7 = 0$ (отсутствие сезонности). Воспользуемся статистикой (3.44):

$$F = \frac{(\mathrm{ESS_R} - \mathrm{ESS_{UR}})/q}{\mathrm{ESS_{UR}}/(n-k)}.$$

Здесь $q = 3$, $k = 7$, $n - k = 17$, $\mathrm{RSS_{UR}} = 118.20$, $\mathrm{RSS_R} = 110.32$. Поскольку $\mathrm{TSS_R} = \mathrm{TSS_{UR}} = \mathrm{TSS}$, то

$$\mathrm{ESS_R} - \mathrm{ESS_{UR}} = \mathrm{RSS_{UR}} - \mathrm{RSS_R} = 118.20 - 110.32 = 7.88,$$
$$\mathrm{ESS_{UR}} = \mathrm{TSS} - \mathrm{RSS_{UR}} = 131.75 - 118.20 = 13.55,$$
$$F = \frac{7.88/3}{13.55/17} = 3.295 > 3.197 = F_{0.05}(3,\,17).$$

Гипотезу об отсутствии сезонности, таким образом, отвергаем (на 5%-ном уровне значимости).

д) Для проверки гипотезы об отсутствии структурного изменения между первым и вторым кварталами 1975 г. воспользуемся тестом Чоу (3.50):

$$F = \frac{(\mathrm{ESS_R} - \mathrm{ESS_{UR}})/k}{\mathrm{ESS_{UR}}/(n+m-2k)}.$$

Здесь $\mathrm{ESS_R} = 21.43$, $\mathrm{ESS_{UR}} = \mathrm{ESS_1} + \mathrm{ESS_2} = 12.25 + 2.32 = 14.57$, $n + m = 24$, $k = 4$,

$$F = \frac{(21.43 - 14.57)/4}{14.57/16} = 1.883 < 3.007 = F_{0.05}(4,\,16).$$

Следовательно, гипотеза об отсутствии структурного изменения не отвергается на 5%-ном уровне значимости.

## Задача 4.7

Процесс, порождающий данные (истинная модель), описывается соотношениями

$$y_t = \beta_1 x_{t1} + \beta_2 x_{t2} + \varepsilon_t,$$
$$\mathrm{E}(\varepsilon_t) = 0, \quad \mathrm{E}(\varepsilon_t^2) = \sigma^2, \quad \mathrm{E}(\varepsilon_t \varepsilon_s) = 0, \quad t \neq s, \;\; t = 1, \ldots, n.$$

Обозначим через $\widehat{\beta}_1$ МНК-оценку параметра $\beta_1$ в этой регрессии, а через $\widehat{\beta}_1^*$ — МНК-оценку параметра $\beta_1$ в регрессии $y$ только на $x_1$.

а) Покажите, что

$$\frac{\mathrm{MSE}(\widehat{\beta}_1^*)}{\mathrm{MSE}(\widehat{\beta}_1)} = 1 + r_{12}^2(t_2^2 - 1),$$

где

$$r_{12}^2 = \frac{(\sum x_1 x_2)^2}{\sum x_1^2 \sum x_2^2} = \frac{(\boldsymbol{x}_1' \boldsymbol{x}_2)^2}{\boldsymbol{x}_1' \boldsymbol{x}_1 \boldsymbol{x}_2' \boldsymbol{x}_2}, \qquad t_2^2 = \frac{\beta_2^2}{\mathrm{V}(\widehat{\beta}_2)}.$$

б) Рассмотрим смесь оценок $\widetilde{\beta}_1 = \lambda\widehat{\beta}_1 + (1-\lambda)\widehat{\beta}_1^*$. При каком значении $\lambda$ величина $\mathrm{MSE}(\widetilde{\beta}_1)$ минимальна?

## Решение

а) По определению $\mathrm{MSE}(\widehat{\beta}_1) = \mathrm{E}(\widehat{\beta}_1 - \beta_1)^2 = \mathrm{V}(\widehat{\beta}_1)$ (в силу несмещенности оценки $\widehat{\beta}_1$). Из формулы (4.18) следует, что

$$\mathrm{V}(\widehat{\beta}_1) = \frac{\sigma^2}{\boldsymbol{x}_1'\boldsymbol{x}_1(1-r_{12}^2)}.$$

Далее,

$$\mathrm{MSE}(\widehat{\beta}_1^*) = \mathrm{E}\left((\widehat{\beta}_1^* - \mathrm{E}(\widehat{\beta}_1^*)) + (\mathrm{E}(\widehat{\beta}_1^*) - \beta_1)\right)^2 = \frac{\sigma^2}{\boldsymbol{x}_1'\boldsymbol{x}_1} + \left(\frac{\boldsymbol{x}_1'\boldsymbol{x}_2}{\boldsymbol{x}_1'\boldsymbol{x}_1}\beta_2\right)^2$$

(в силу (4.18), (4.14) и того факта, что $\mathrm{E}(\widehat{\beta}_1^*) - \beta_1$ есть неслучайная величина).

Следовательно,

$$\begin{aligned}\frac{\mathrm{MSE}(\widehat{\beta}_1^*)}{\mathrm{MSE}(\widehat{\beta}_1)} &= \frac{\dfrac{\sigma^2}{\boldsymbol{x}_1'\boldsymbol{x}_1} + \left(\dfrac{\boldsymbol{x}_1'\boldsymbol{x}_2}{\boldsymbol{x}_1'\boldsymbol{x}_1}\beta_2\right)^2}{\dfrac{\sigma^2}{\boldsymbol{x}_1'\boldsymbol{x}_1}}(1-r_{12}^2) \\ &= (1-r_{12}^2)\left(1 + \frac{\beta_2^2\boldsymbol{x}_2'\boldsymbol{x}_2(1-r_{12}^2)(\boldsymbol{x}_1'\boldsymbol{x}_2)^2}{\sigma^2\boldsymbol{x}_2'\boldsymbol{x}_2\boldsymbol{x}_1'\boldsymbol{x}_1(1-r_{12}^2)}\right) \\ &= (1-r_{12}^2)\left(1 + t_2^2\frac{r_{12}^2}{1-r_{12}^2}\right) = 1 + r_{12}^2(t_2^2 - 1).\end{aligned}$$

Мы здесь использовали формулу (4.18), согласно которой

$$\mathrm{V}(\widehat{\beta}_2) = \frac{\sigma^2}{\boldsymbol{x}_2'\boldsymbol{x}_2(1-r_{12}^2)}.$$

б) По определению среднеквадратичной ошибки получаем:

$$\begin{aligned}\mathrm{MSE}(\widetilde{\beta}_1) &= \mathrm{MSE}(\lambda\widehat{\beta}_1 + (1-\lambda)\widehat{\beta}_1^*) = \mathrm{E}\big(\lambda(\widehat{\beta}_1 - \beta_1) + (1-\lambda)(\widehat{\beta}_1^* - \beta_1)\big)^2 \\ &= \lambda^2\mathrm{MSE}(\widehat{\beta}_1) + (1-\lambda)^2\mathrm{MSE}(\widehat{\beta}_1^*) + 2\lambda(1-\lambda)\mathrm{E}\big((\widehat{\beta}_1 - \beta_1)(\widehat{\beta}_1^* - \beta_1)\big).\end{aligned}$$

Вычислим последнее слагаемое. Обозначим

$$\boldsymbol{x}_1 = (x_{11}, \ldots, x_{n1})',\ \boldsymbol{x}_2 = (x_{12}, \ldots, x_{n2})',$$
$$\varepsilon = (\varepsilon_1, \ldots, \varepsilon_n)',\ \Delta = (\boldsymbol{x}_1'\boldsymbol{x}_1)(\boldsymbol{x}_2'\boldsymbol{x}_2) - (\boldsymbol{x}_1'\boldsymbol{x}_2)^2.$$

Тогда

$$\widehat{\beta}_1^* - \beta_1 = \frac{x_1'x_2\beta_2 + x_1'\varepsilon}{x_1'x_1}, \qquad \widehat{\beta}_1 - \beta_1 = \frac{x_2'x_2x_1'\varepsilon - x_1'x_2x_2'\varepsilon}{\Delta},$$

$$\begin{aligned}\mathrm{E}\left((\widehat{\beta}_1 - \beta_1)(\widehat{\beta}_1^* - \beta_1)\right) &= \mathrm{E}\left(\frac{x_1'x_2\beta_2 + x_1'\varepsilon}{x_1'x_1} \cdot \frac{x_2'x_2x_1' - x_1'x_2x_2'}{\Delta}\varepsilon\right) \\ &= \mathrm{E}\left(\frac{x_2'x_2(x_1'\varepsilon)^2 - x_1'x_2(x_1'\varepsilon)(x_2'\varepsilon)}{x_1'x_1\Delta}\right) \\ &= \frac{\sigma^2x_1'x_1x_2'x_2 - \sigma^2(x_1'x_2)^2}{x_1'x_1\Delta} = \frac{\sigma^2}{x_1'x_1}.\end{aligned}$$

Отсюда, учитывая результаты п. а), получаем:

$$\begin{aligned}\mathrm{MSE}(\widetilde{\beta}_1) &= \lambda^2\mathrm{MSE}(\widehat{\beta}_1) + (1-\lambda)^2\mathrm{MSE}(\widehat{\beta}_1^*) + 2\lambda(1-\lambda)\frac{\sigma^2}{x_1'x_1} \\ &= \mathrm{MSE}(\widehat{\beta}_1)\left(\lambda^2 + (1-\lambda)^2(1 + r_{12}^2(t_2^2 - 1)) + 2\lambda(1-\lambda)(1 - r_{12}^2)\right) \\ &= \mathrm{MSE}(\widehat{\beta}_1)f(\lambda).\end{aligned}$$

Таким образом, задача сводится к минимизации квадратичной функции

$$\begin{aligned}f(\lambda) &= \lambda^2 + (1-\lambda)^2(1 + r_{12}^2(t_2^2 - 1)) + 2\lambda(1-\lambda)(1 - r_{12}^2) \\ &= r_{12}^2(t_2^2 + 1)\lambda^2 - 2r_{12}^2t_2^2\lambda + 1 + r_{12}^2(t_2^2 - 1),\end{aligned}$$

которая достигает минимума при $\lambda_{\min} = \dfrac{t_2^2}{1 + t_2^2}$.

## Задача 4.8

Процесс, порождающий данные (истинная модель), описывается соотношениями:

$$\begin{aligned}&y_t = \beta x_t + \gamma z_t + \varepsilon_t, \\ &\mathrm{E}(\varepsilon_t) = 0, \quad \mathrm{E}(\varepsilon_t^2) = \sigma_\varepsilon^2, \quad \mathrm{E}(\varepsilon_t\varepsilon_s) = 0, \quad t \neq s, \quad t = 1, \ldots, n.\end{aligned}$$

Переменная $x$ наблюдается с ошибками, т. е. в регрессии могут быть использованы лишь величины $w_t = x_t + u_t$, при этом предполагается, что ошибки $u$ удовлетворяют условиям $\mathrm{E}(u_t) = 0$, $\mathrm{E}(u_t^2) = \sigma_u^2$, $\mathrm{E}(u_tu_s) = 0$, $t \neq s$, $\mathrm{E}(u_t\varepsilon_s) = 0$, $\forall\, s, t$.

Проводятся две регрессии: первая — $y$ на $z$; вторая — $y$ на $z$ и $w$.

Покажите, что смещение оценки параметра $\gamma$ во второй регрессии меньше, чем в первой.

## Решение

Обозначим $\boldsymbol{x}, \boldsymbol{y}, \boldsymbol{w}, \boldsymbol{z}$ векторы, составленные из наблюдений соответствующих величин. Пусть также $\widehat{\gamma}_1, \widehat{\gamma}_2$ — оценки параметра $\gamma$, полученные в первой и второй регрессиях соответственно. Тогда

$$\widehat{\gamma}_1 = \frac{\boldsymbol{z}'\boldsymbol{y}}{\boldsymbol{z}'\boldsymbol{z}} = \frac{\boldsymbol{z}'(\boldsymbol{x}\beta + \boldsymbol{z}\gamma + \boldsymbol{\varepsilon})}{\boldsymbol{z}'\boldsymbol{z}} = \gamma + \beta\frac{\boldsymbol{z}'\boldsymbol{x}}{\boldsymbol{z}'\boldsymbol{z}} + \frac{\boldsymbol{z}'\boldsymbol{\varepsilon}}{\boldsymbol{z}'\boldsymbol{z}} \quad \text{и} \quad \mathrm{E}(\widehat{\gamma}_1) - \gamma = \beta\frac{\boldsymbol{z}'\boldsymbol{x}}{\boldsymbol{z}'\boldsymbol{z}}.$$

Далее,

$$\begin{bmatrix} \widehat{\beta}_2 \\ \widehat{\gamma}_2 \end{bmatrix} = \begin{bmatrix} \boldsymbol{w}'\boldsymbol{w} & \boldsymbol{w}'\boldsymbol{z} \\ \boldsymbol{z}'\boldsymbol{w} & \boldsymbol{z}'\boldsymbol{z} \end{bmatrix}^{-1} \begin{bmatrix} \boldsymbol{w}' \\ \boldsymbol{z}' \end{bmatrix} \boldsymbol{y},$$

откуда

$$\begin{aligned} \widehat{\gamma}_2 &= \frac{\boldsymbol{w}'\boldsymbol{w}\boldsymbol{z}' - \boldsymbol{w}'\boldsymbol{z}\boldsymbol{w}'}{\boldsymbol{w}'\boldsymbol{w}\boldsymbol{z}'\boldsymbol{z} - (\boldsymbol{z}'\boldsymbol{w})^2}\boldsymbol{y} = \frac{\boldsymbol{w}'\boldsymbol{w}\boldsymbol{z}' - \boldsymbol{w}'\boldsymbol{z}\boldsymbol{w}'}{\boldsymbol{w}'\boldsymbol{w}\boldsymbol{z}'\boldsymbol{z} - (\boldsymbol{z}'\boldsymbol{w})^2}(\boldsymbol{x}\beta + \boldsymbol{z}\gamma + \boldsymbol{\varepsilon}) \\ &= \gamma + \frac{\boldsymbol{w}'\boldsymbol{w}\boldsymbol{z}'\boldsymbol{x} - \boldsymbol{w}'\boldsymbol{z}\boldsymbol{w}'\boldsymbol{x}}{\boldsymbol{w}'\boldsymbol{w}\boldsymbol{z}'\boldsymbol{z} - (\boldsymbol{z}'\boldsymbol{w})^2}\beta + \frac{\boldsymbol{w}'\boldsymbol{w}\boldsymbol{z}' - \boldsymbol{w}'\boldsymbol{z}\boldsymbol{w}'}{\boldsymbol{w}'\boldsymbol{w}\boldsymbol{z}'\boldsymbol{z} - (\boldsymbol{z}'\boldsymbol{w})^2}\boldsymbol{\varepsilon}. \end{aligned}$$

Поскольку $\boldsymbol{u}$ и $\boldsymbol{\varepsilon}$ независимы, то математическое ожидание третьего слагаемого в последней формуле равно 0. Таким образом,

$$\mathrm{E}(\widehat{\gamma}_2) - \gamma = \beta\mathrm{E}\left(\frac{\boldsymbol{w}'\boldsymbol{w}\boldsymbol{z}'\boldsymbol{x} - \boldsymbol{w}'\boldsymbol{z}\boldsymbol{w}'\boldsymbol{x}}{\boldsymbol{w}'\boldsymbol{w}\boldsymbol{z}'\boldsymbol{z} - (\boldsymbol{z}'\boldsymbol{w})^2}\right).$$

Числитель и знаменатель дроби, стоящей под знаком математического ожидания, являются случайными величинами, поэтому не удается получить эффективную формулу для смещений в этом случае, чтобы можно было провести сравнение с первым случаем. Поэтому найдем асимптотическое (при $n \to \infty$) смещение, где $n$ — число наблюдений. Без ограничения общности можно считать, что $\overline{x} = \overline{z} = 0$. Имеем

$$\frac{\boldsymbol{w}'\boldsymbol{w}\boldsymbol{z}'\boldsymbol{x} - \boldsymbol{w}'\boldsymbol{z}\boldsymbol{w}'\boldsymbol{x}}{\boldsymbol{w}'\boldsymbol{w}\boldsymbol{z}'\boldsymbol{z} - (\boldsymbol{z}'\boldsymbol{w})^2} = \frac{\dfrac{\boldsymbol{w}'\boldsymbol{w}}{n}\cdot\dfrac{\boldsymbol{z}'\boldsymbol{x}}{n} - \dfrac{\boldsymbol{w}'\boldsymbol{z}}{n}\cdot\dfrac{\boldsymbol{w}'\boldsymbol{x}}{n}}{\dfrac{\boldsymbol{w}'\boldsymbol{w}}{n}\cdot\dfrac{\boldsymbol{z}'\boldsymbol{z}}{n} - \left(\dfrac{\boldsymbol{z}'\boldsymbol{w}}{n}\right)^2}.$$

Будем предполагать, что существуют пределы

$$p\lim_{n\to\infty}\frac{\boldsymbol{z}'\boldsymbol{x}}{n} = \mathrm{Cov}(x, z), \quad p\lim_{n\to\infty}\frac{\boldsymbol{x}'\boldsymbol{x}}{n} = \mathrm{Var}(x), \quad p\lim_{n\to\infty}\frac{\boldsymbol{z}'\boldsymbol{z}}{n} = \mathrm{Var}(z)$$

и компоненты векторов $\boldsymbol{x}$, $\boldsymbol{z}$ равномерно ограничены. Тогда по закону больших чисел

$$p\lim_{n\to\infty}\frac{\boldsymbol{w}'\boldsymbol{w}}{n} = p\lim_{n\to\infty}\frac{(\boldsymbol{x}+\boldsymbol{u})'(\boldsymbol{x}+\boldsymbol{u})}{n} = \mathrm{Var}(x) + \sigma_u^2,$$

$$p\lim_{n\to\infty}\frac{\boldsymbol{w}'\boldsymbol{x}}{n} = \mathrm{Var}(x), \quad p\lim_{n\to\infty}\frac{\boldsymbol{w}'\boldsymbol{z}}{n} = \mathrm{Cov}(x, z).$$

Отсюда

$$p\lim_{n\to\infty}(\gamma - \mathrm{E}(\widehat{\gamma}_2)) = -\beta\frac{(\mathrm{Var}(x)+\sigma_u^2)\,\mathrm{Cov}(x,z) - \mathrm{Cov}(x,z)\,\mathrm{Var}(x)}{(\mathrm{Var}(x)+\sigma_u^2)\,\mathrm{Var}(z) - \mathrm{Cov}^2(x,z)} = -\beta\frac{\sigma_u^2\,\mathrm{Cov}(x,z)}{\mathrm{Var}(x)\,\mathrm{Var}(z) - \mathrm{Cov}^2(x,z) + \sigma_u^2\,\mathrm{Var}(z)}.$$

В силу неравенства Коши-Буняковского $\mathrm{Var}(x)\,\mathrm{Var}(z) - \mathrm{Cov}^2(x,z) \geqslant 0$, поэтому (асимптотически)

$$|\gamma - \mathrm{E}(\widehat{\gamma}_2)| \leqslant |\beta|\frac{\sigma_u^2\,\mathrm{Cov}(x,z)}{\sigma_u^2\,\mathrm{Var}(z)} = |\gamma - \mathrm{E}(\widehat{\gamma}_1)|.$$

## Задача 4.9

Процесс, порождающий данные (истинная модель), описывается соотношениями:

$$y_t = \beta_1 x_{t1} + \beta_2 x_{t2} + \varepsilon_t,$$
$$\mathrm{E}(\varepsilon_t) = 0,\quad \mathrm{E}(\varepsilon_t^2) = \sigma^2,\quad \mathrm{E}(\varepsilon_t\varepsilon_s) = 0,\quad t \neq s,\quad t = 1,\ldots,n.$$

Проводится регрессия $\boldsymbol{y}$ на $\boldsymbol{x}_1$ и стандартным образом через остатки этой регрессии оценивается дисперсия $\sigma^2$. Покажите, что полученная оценка смещена вверх.

## Решение

Регрессия $\boldsymbol{y}$ на $\boldsymbol{x}_1$ записывается как $y_t = \beta_1 x_{t1} + u_t$. Используя стандартную формулу, получаем оценку

$$\widehat{\beta}_1 = \frac{\boldsymbol{x}_1'\boldsymbol{y}}{\boldsymbol{x}_1'\boldsymbol{x}_1}.$$

Применив далее формулу (4.15), получим:

$$\mathrm{E}(\widehat{\sigma}^2) = \sigma^2 + \frac{1}{n-1}\beta_2^2\boldsymbol{x}_2'\left(\boldsymbol{I} - \frac{\boldsymbol{x}_1\boldsymbol{x}_1'}{\boldsymbol{x}_1'\boldsymbol{x}_1}\right)\boldsymbol{x}_2 = \sigma^2 + \frac{1}{n-1}\beta_2^2\frac{(\boldsymbol{x}_1'\boldsymbol{x}_1)(\boldsymbol{x}_2'\boldsymbol{x}_2) - (\boldsymbol{x}_1'\boldsymbol{x}_2)^2}{\boldsymbol{x}_1'\boldsymbol{x}_1} \geqslant \sigma^2$$

в силу неравенства Коши–Буняковского.

## Задача 4.10

Предположим, что некоторые ежегодные данные удовлетворяют соотношениям:

$$y_t = \alpha + \beta_1 x_t + \beta_2 t + \varepsilon_t \quad \text{(истинная модель)},$$

причем выполнены все условия классической регрессии. Однако оценивается «неправильная» модель без временного тренда

$$y_t = a + b_1 x_t + u_t.$$

а) Какие из условий классической регрессии не выполнены для модели без временного тренда?

б) Будет ли равна нулю сумма остатков для этой регрессии? Как это связано с ошибочным предположением, что $\mathrm{E}(u_t) = 0$?

в) Предположим, что коэффициент $\beta_2$ положителен и нарисован график остатков $\widehat{u}_t$ регрессии $y_t = a + b_1 x_t + \widehat{u}_t$ как функция времени. Как должен выглядеть этот график?

## Решение

а) Проверим выполнение условий классической регрессии.

Итак, рассматривается модель $y_t = \alpha + \beta_1 x_t + u_t$, $t = 1, \ldots, n$. Выясним свойства ошибок $u_t$.

1)$\mathrm{E}(u_t) = \mathrm{E}(y_t - (\alpha + \beta_1 x_t)) = \mathrm{E}(\beta_2 t + \varepsilon_t) = \beta_2 t \neq 0$. Таким образом, условие равенства нулю математических ожиданий ошибок *не выполняется*.

2)$\mathrm{V}(u_t) = \mathrm{V}(y_t - (\alpha + \beta_1 x_t)) = \mathrm{V}(\beta_2 t + \varepsilon_t) = \mathrm{V}(\varepsilon_t) = \sigma^2$ — дисперсия ошибок не зависит от $t$.

3) $\mathrm{Cov}(u_t, u_s) = \mathrm{Cov}(\beta_2 t + \varepsilon_t, \beta_2 s + \varepsilon_s) = \mathrm{Cov}(\varepsilon_t, \varepsilon_s) = 0$ при $t \neq s$ — ошибки некоррелированы.

б) Одно из уравнений метода наименьших квадратов для данной задачи таково: $\sum e_t = \sum_{t=1}^{n} (y_t - \alpha - \beta_1 x_t) = 0$, т.е. сумма остатков равна нулю. Это верно для любой регрессии, если среди регрессоров есть константа. Поскольку сумма остатков может быть не равна нулю, несмотря на то, что $\mathrm{E}(u_t) = 0$ (в случае классической модели и отсутствия константы среди регрессоров), то мы делаем вывод, что равенство суммы остатков нулю не связано с предположением, что $\mathrm{E}(u_t) = 0$.

в) Однозначного ответа на этот вопрос дать нельзя. На рисунках приведены примеры различных графиков остатков, полученных из таких моделей.

В первом примере (рис. 4.1) использовалась модель

$$y_t = \alpha + \beta_2 t + \varepsilon_t, \ \ t = 1, \ldots, n \ (x_t = 0).$$

Во втором примере (рис. 4.2) использовалась модель

$$y_t = \alpha + \beta_1 t^2 + \beta_2 t + \varepsilon_t, \ \ t = 1, \ldots, n \ (x_t = t^2).$$

В обоих случаях проводилась регрессия без учета временного тренда $\beta_2 t$. Поведение остатков различно.

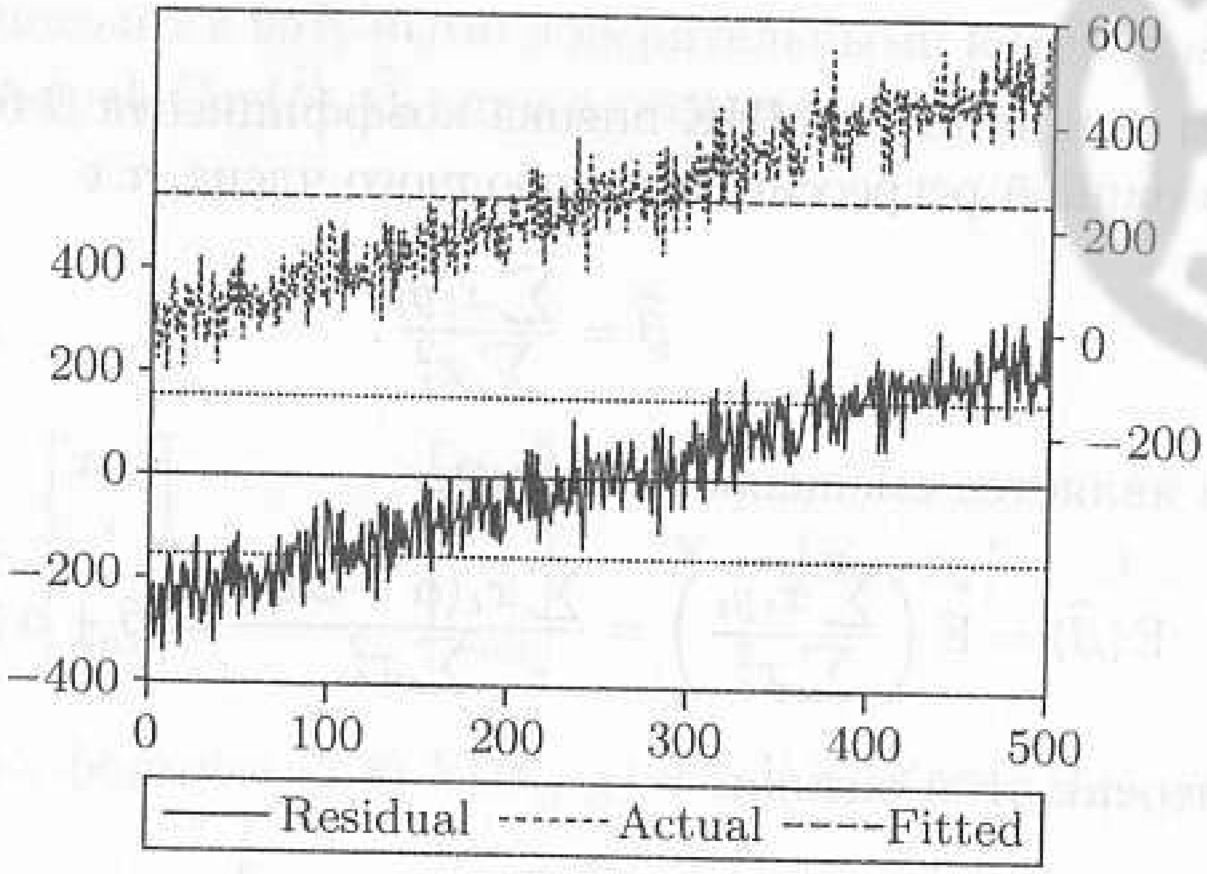



**Рис. 4.1**

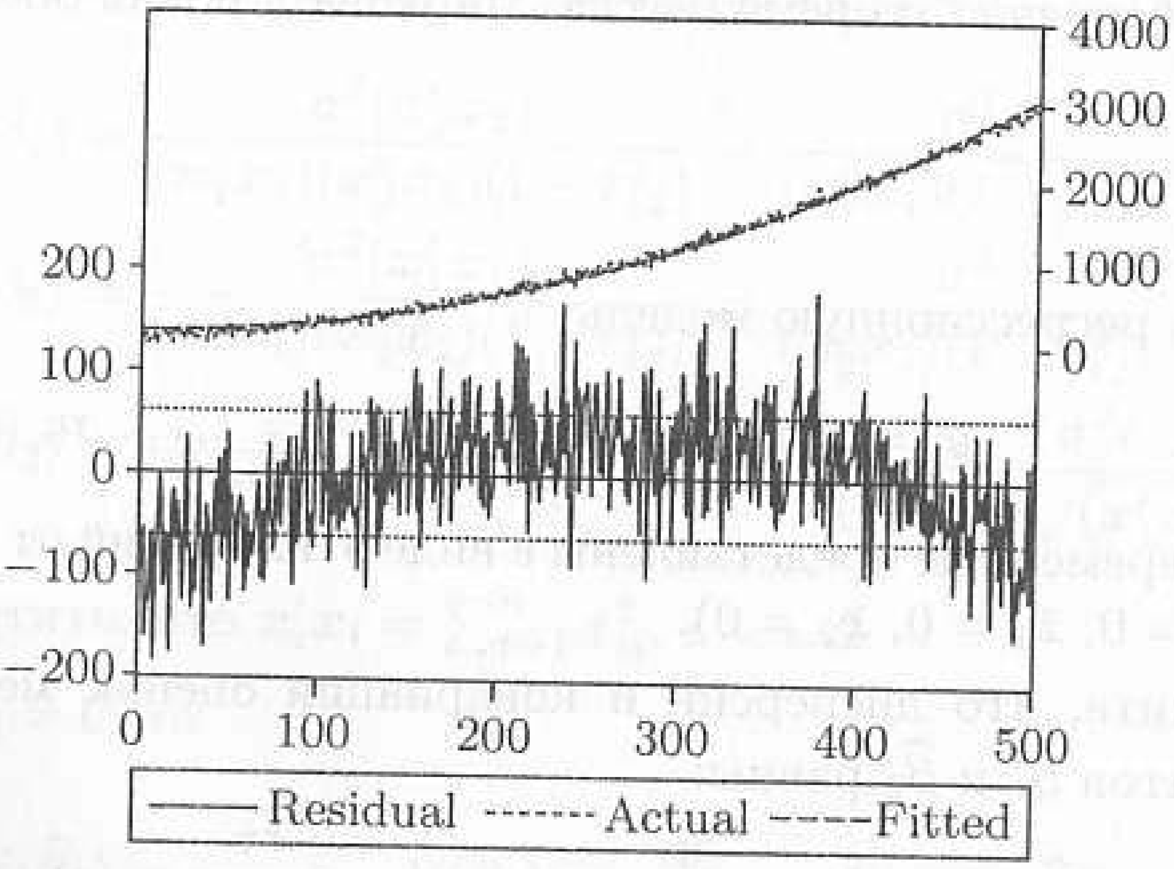



**Рис. 4.2**

## Задача 4.11

Дана стандартная модель парной регрессии

$$y_t = \alpha + \beta x_t + \varepsilon_t, \quad t = 1, \ldots, n.$$

а) Чему равна МНК-оценка коэффициента $\beta$ при ограничении $\alpha = 0$?

б) Чему равна дисперсия оценки в а)? Покажите, что она меньше, чем $\sigma^2 / \sum_{t=1}^{n}(x_t - \overline{x})^2$ — дисперсия МНК-оценки $\beta$ в регрессии без ограничения. Противоречит ли это теореме Гаусса–Маркова?

## Решение

а) При этом ограничении МНК-оценка коэффициента $\beta$ будет точно такой же, как и в парной регрессии без свободного члена, т. е.

$$\widehat{\beta} = \frac{\sum x_t y_t}{\sum x_t^2}.$$

Эта оценка является смещенной, так как

$$\mathrm{E}(\widehat{\beta}) = \mathrm{E}\left(\frac{\sum x_t y_t}{\sum x_t^2}\right) = \frac{\sum x_t(\alpha + \beta x_t)}{\sum x_t^2} = \beta + \alpha\frac{\sum x_t}{\sum x_t^2}.$$

б) Дисперсия этой оценки:

$$\mathrm{V}(\widehat{\beta}) = \mathrm{V}\left(\frac{\sum x_t y_t}{\sum x_t^2}\right) = \frac{\sigma^2}{\sum x_t^2} \leqslant \frac{\sigma^2}{\sum x_t^2 - n\overline{x}^2} = \mathrm{V}(\widehat{\beta}_{\mathrm{OLS}}).$$

Это не противоречит теореме Гаусса–Маркова, так как оценка является смещенной.

## Задача 4.12

Рассмотрим регрессионную модель

$$y_t = \beta_1 x_{t1} + \beta_2 x_{t2} + \varepsilon_t, \quad t = 1, \ldots, n,$$

в которой переменные представлены в виде отклонений от выборочных средних (т. е. $\bar{y} = 0$, $\bar{x}_1 = 0$, $\bar{x}_2 = 0$).

а) Покажите, что дисперсии и ковариация оценок метода наименьших квадратов $\widehat{\beta}_1$ и $\widehat{\beta}_2$ равны:

$$\mathrm{V}(\widehat{\beta}_1) = \frac{\sigma^2}{\sum_{t=1}^n x_{t1}^2(1 - r_{12}^2)}, \quad \mathrm{V}(\widehat{\beta}_2) = \frac{\sigma^2}{\sum_{t=1}^n x_{t2}^2(1 - r_{12}^2)},$$

$$\mathrm{Cov}(\widehat{\beta}_1, \widehat{\beta}_2) = \frac{-\sigma^2 r_{12}}{(1 - r_{12}^2)\sqrt{\sum_{t=1}^n x_{t1}^2 \sum_{t=1}^n x_{t2}^2}},$$

где

$$r_{12} = \frac{\sum_{t=1}^n x_{t1}x_{t2}}{\sqrt{\sum_{t=1}^n x_{t1}^2 \sum_{t=1}^n x_{t2}^2}}$$

— выборочный коэффициент корреляции между $x_1$ и $x_2$.

б) Чему равны дисперсии и ковариация в случае $r_{12} = 0$? Как это связано с проблемой мультиколлинеарности?

в) Постройте график отношения $\mathrm{V}(\widehat{\beta}_1)$ к значению $\mathrm{V}(\widehat{\beta}_1)$, полученному в б), в диапазоне $0 < r_{12} < 1$. Как этот график связан с проблемой мультиколлинеарности?

г) Что происходит с 95%-ными доверительными интервалами для $\beta_1$ и $\beta_2$ и ковариацией $\mathrm{Cov}(\widehat{\beta}_1, \widehat{\beta}_2)$ при возрастании $r_{12}$ в диапазоне $0 < r_{12} < 1$?

## Решение

а) Обозначим

$$x_1 = \begin{bmatrix} x_{11} \\ \vdots \\ x_{n1} \end{bmatrix}, \quad x_2 = \begin{bmatrix} x_{12} \\ \vdots \\ x_{n2} \end{bmatrix}, \quad X = \begin{bmatrix} x_1 & x_2 \end{bmatrix}, \quad \beta = \begin{bmatrix} \beta_1 \\ \beta_2 \end{bmatrix}.$$

Тогда согласно формуле (3.8) $\mathrm{V}(\widehat{\beta}_{\mathrm{OLS}}) = \sigma^2 (X'X)^{-1}$. Нетрудно проверить, что

$$(X'X)^{-1} = \frac{1}{\Delta} \begin{bmatrix} x_2'x_2 & -x_2'x_1 \\ -x_1'x_2 & x_1'x_1 \end{bmatrix},$$

где $\Delta = (x_1'x_1)(x_2'x_2) - (x_1'x_2)^2 = (x_1'x_1)(x_2'x_2)(1 - r_{12}^2)$. Таким образом,

$$\mathrm{V}(\widehat{\beta}_1) = \frac{\sigma^2 (x_2'x_2)}{(x_1'x_1)(x_2'x_2)(1 - r_{12}^2)} = \frac{\sigma^2}{(x_1'x_1)(1 - r_{12}^2)},$$

$$\mathrm{V}(\widehat{\beta}_2) = \frac{\sigma^2 (x_1'x_1)}{(x_1'x_1)(x_2'x_2)(1 - r_{12}^2)} = \frac{\sigma^2}{(x_2'x_2)(1 - r_{12}^2)},$$

$$\mathrm{Cov}(\widehat{\beta}_1, \widehat{\beta}_2) = -\frac{\sigma^2 (x_1'x_2)}{(x_1'x_1)(x_2'x_2)(1-r_{12}^2)} = -\frac{\sigma^2 r_{12}}{(1-r_{12}^2)\sqrt{(x_1'x_1)(x_2'x_2)}}.$$

Остается заметить, что $x_i'x_i = \sum_{t=1}^{n} x_{ti}^2$, $i = 1, 2$.

б) Если $r_{12} = 0$, то

$$\mathrm{V}(\widehat{\beta}_1) = \frac{\sigma^2}{x_1'x_1}, \quad \mathrm{V}(\widehat{\beta}_2) = \frac{\sigma^2}{x_2'x_2}, \quad \mathrm{Cov}(\widehat{\beta}_1, \widehat{\beta}_2) = 0,$$

и легко понять, что это — минимальные значения дисперсий оценок. В этом случае, как нетрудно видеть, исходная регрессия равносильна двум раздельным регрессиям $y$ на $x_1$ и $x_2$ соответственно.

Этот пример иллюстрирует тот факт, что точность МНК-оценок тем выше, чем меньше корреляция между регрессорами.

в) Как легко проверить, требуемое отношение дисперсий равно $1/(1 - r_{12}^2)$.

г) Поскольку длина 95%-ного доверительного интервала для $\beta_i$, $i = 1, 2$, пропорциональна $\sqrt{\mathrm{V}(\widehat{\beta}_i)}$, то с ростом $r_{12}$ от 0 до 1 длина 95%-ного доверительного интервала неограниченно возрастает. Абсолютная величина ковариации $\mathrm{Cov}(\widehat{\beta}_1, \widehat{\beta}_2)$ также неограниченно возрастает.

## Задача 4.13

Некоторая фирма занимается продажей молока. В таблице 4.1 представлены объемы ежемесячных продаж $Q$ (тыс. л) по различным ценам $P$ (руб. за 1 л). Во время пятого, шестого и седьмого месяцев на одном из предприятий фирмы происходила забастовка. С помощью регрессий $Q$ на $P$ определите:

а) произошел ли сдвиг свободного члена (константы) во время забастовки по сравнению с обычным режимом;

б) произошел ли сдвиг как константы, так и коэффициента наклона при $P$.

**Таблица 4.1**

| Месяц | $Q$ | $P$ | Месяц | $Q$ | $P$ |
|---|---|---|---|---|---|
| 1 | 98 | 10.0 | 8 | 113 | 13.0 |
| 2 | 100 | 11.0 | 9 | 116 | 13.0 |
| 3 | 103 | 12.5 | 10 | 118 | 13.8 |
| 4 | 105 | 12.5 | 11 | 121 | 14.2 |
| 5 | 80 | 14.6 | 12 | 123 | 14.4 |
| 6 | 87 | 14.6 | 13 | 126 | 15.0 |
| 7 | 94 | 14.9 | 14 | 128 | 16.1 |

## Решение

Регрессия $Q$ на $P$, проведенная по всем наблюдениям, включающим пятый, шестой и седьмой месяцы, дает следующие результаты (см. таблицу 4.2).

**Таблица 4.2**

Dependent Variable: $Q$

| Variable | Coefficient | Std. Error | $t$-Statistic | Probability |
|---|---|---|---|---|
| const | 74.819 | 34.281 | 2.183 | 0.0497 |
| $P$ | 2.450 | 2.514 | 0.975 | 0.3490 |
| $R^2$ | 0.0733 | | | |

Качество регрессии плохое, так как значимость коэффициентов низкая и коэффициент детерминации близок к нулю. Если исключить наблюдения, относящиеся к забастовке (пятый, шестой и седьмой месяцы), то результаты такой же регрессии выглядят иначе (см. таблицу 4.3).

В этой регрессии все коэффициенты значимы и коэффициент детерминации превышает 0.9. Поэтому один из возможных подходов к решению задачи состоит в исследовании регрессии, из которой исключены наблюдения, относящиеся к забастовке.

**Таблица 4.3**

Dependent Variable: $Q$

| Variable | Coefficient | Std. Error | $t$-Statistic | Probability |
|---|---|---|---|---|
| const | 36.617 | 8.229 | 4.450 | 0.0016 |
| $P$ | 5.830 | 0.617 | 9.445 | 0.0000 |
| $R^2$ | 0.9084 | | | |

а) Введем фиктивную переменную $D$, принимающую значение 0 для наблюдений, относящихся к месяцам 1–4, и принимающую значение 1 для наблюдений, относящихся к месяцам 8–14 (наблюдения для 5-го, 6-го и 7-го месяцев исключены).

Регрессия $Q$ на $P$ и $D$ дает следующие результаты (см. таблицу 4.4).

**Таблица 4.4**

Dependent Variable: $Q$

| Variable | Coefficient | Std. Error | $t$-Statistic | Probability |
|---|---|---|---|---|
| const | 57.236 | 6.458 | 8.863 | 0.0000 |
| $P$ | 3.849 | 0.555 | 6.930 | 0.0001 |
| $D$ | 8.769 | 1.925 | 4.554 | 0.0019 |
| $R^2$ | 0.9745 | | | |

Видим, что коэффициент при переменной $D$ значим на 0.2%-ном уровне, что позволяет говорить о том, что после забастовки в регрессии произошел сдвиг константы.

б) Для ответа на вопрос этого пункта рассмотрим регрессию $Q$ на $P$, $D$ и $D \cdot P$ (см. таблицу 4.5).

**Таблица 4.5**

Dependent Variable: $Q$

| Variable | Coefficient | Std. Error | $t$-Statistic | Probability |
|---|---|---|---|---|
| const | 73.389 | 7.932 | 9.253 | 0.0000 |
| $P$ | 2.444 | 0.687 | 3.559 | 0.0092 |
| $D$ | −19.645 | 11.003 | −1.781 | 0.1182 |
| $D \cdot P$ | 2.667 | 0.873 | 2.598 | 0.0355 |
| $R^2$ | 0.9745 | | | |

Видим, что коэффициент $\beta_3$ (при переменной $D$) незначим на 5%-ном уровне, однако проверка гипотезы $H_0$: $\beta_3 = \beta_4 = 0$ с помощью соответствующего $F$-теста позволяет уверенно (на 0.1%-ном уровне) ее отвергнуть. Иными словами, в рамках данного подхода можно утверждать, что после забастовки произошло изменение как константы, так и коэффициента при $P$.

## Задача 4.14

В таблице 4.6 представлены совокупный объем внутренних инвестиций $y$ и валовой внутренний продукт США (млрд. долл.) за период с 1939 по 1954 г.

**Таблица 4.6.** Инвестиции и ВВП США

| Год | $y$ | $x$ | Год | $y$ | $x$ |
|---|---|---|---|---|---|
| 1939 | 9.3 | 90.8 | 1947 | 34.0 | 232.8 |
| 1940 | 13.1 | 100.0 | 1948 | 45.9 | 259.1 |
| 1941 | 17.9 | 124.9 | 1949 | 35.3 | 258.0 |
| 1942 | 9.9 | 158.3 | 1950 | 53.8 | 286.2 |
| 1943 | 5.8 | 192.0 | 1951 | 59.2 | 330.2 |
| 1944 | 7.2 | 210.5 | 1952 | 52.1 | 347.2 |
| 1945 | 10.6 | 212.3 | 1953 | 53.3 | 366.1 |
| 1946 | 30.7 | 209.3 | 1954 | 52.7 | 366.3 |

Источник: *D.Salvatore. Statistics and Econometrics*, McGraw-Hill, 1982.

Напишите и оцените уравнения, позволяющие ответить на вопрос, изменилась ли зависимость инвестиций от валового внутреннего продукта во время войны (1942–1945 гг.) по сравнению с мирным временем.

## Решение

Визуальный анализ графика показывает, что если исключить военные годы, то прослеживается линейная зависимость $y$ от $x$. Наблюдения, относящиеся к военному периоду (1942–1945 гг.), представляют собой явные «выбросы». Регрессия $y$ на $x$ по всем наблюдениям дает результаты, представленные в таблице 4.7.

**Таблица 4.7**

Dependent Variable: $y$

| Variable | Coefficient | Std. Error | $t$-Statistic | Probability |
|---|---|---|---|---|
| const | −14.362 | 7.543 | −1.904 | 0.0777 |
| $x$ | 0.192 | 0.030 | 6.368 | 0.0000 |
| $R^2$ | 0.7434 | | | |

Та же регрессия, но проведенная по наблюдениям без военного периода, представлена в таблице 4.8.

Сравнение этих двух регрессий подтверждает тот факт, что наблюдения, относящиеся к военному периоду, не очень хорошо укладываются в общую картину: во второй регрессии выше значение $t$-статистики коэффициента при $x$ и существенно выше коэффициент детерминации.

**Таблица 4.8**

Dependent Variable: $y$

| Variable | Coefficient | Std. Error | $t$-Statistic | Probability |
|---|---|---|---|---|
| const | −3.335 | 4.086 | −0.816 | 0.4334 |
| $x$ | 0.167 | 0.015 | 10.872 | 0.0000 |
| $R^2$ | 0.9220 | | | |

Можно более строго обосновать различие моделей для мирного и военного времени, например, с помощью теста Чоу. Разбивая все наблюдения на две группы — мирное время (1939–1941 гг., 1946–1954 гг.) и военное время (1942–1945 гг.) — и применяя к исходной регрессии $y_t = \beta_1 + \beta_2 x_t + \varepsilon_t$ тест Чоу, получаем значение соответствующей $F$-статистики 27.608 и $P$-значение 0.0000, что позволяет уверенно отвергнуть гипотезу о совпадении моделей для мирного и военного времени.

Этот же вывод можно получить, не применяя тест Чоу, а используя фиктивную переменную, позволяющую разделить мирное и военное время. Введем переменную $D$, которая принимает значение 1 для наблюдений, относящихся к военному времени (1942–1945 гг.), и 0 — к мирному времени (1939–1941 гг., 1946–1954 гг.). Тогда гипотезу о совпадении моделей для мирного и военного времени можно тестировать следующим образом: провести регрессию $y$ на константу, $x$, $D$ и произведение $x$ на $D$, а затем проверить совместную значимость двух последних коэффициентов. Ниже приведены результаты регрессии (см. таблицу 4.9).

**Таблица 4.9**

Dependent Variable: $y$

| Variable | Coefficient | Std. Error | $t$-Statistic | Probability |
|---|---|---|---|---|
| const | −3.335 | 3.836 | −0.870 | 0.4016 |
| $x$ | 0.167 | 0.014 | 11.583 | 0.0000 |
| $D$ | 14.579 | 21.763 | 0.669 | 0.5162 |
| $D \cdot x$ | −0.182 | 0.111 | −1.638 | 0.1273 |
| $R^2$ | 0.9541 | | | |

Применяя $F$-тест для проверки гипотезы $\mathrm{H}_0$: $\beta_3 = \beta_4 = 0$, получаем значение соответствующей $F$-статистики 27.608, что (как и следовало ожидать) совпадает со значением $F$-статистики, полученной в тесте Чоу. Отметим также, что в последней регрессии каждый из коэффициентов $\beta_3$ и $\beta_4$ незначим, в то время как совместно они являются значимыми, что позволяет уверенно отвергнуть гипотезу $\mathrm{H}_0$.

## Задача 4.15

В таблице 4.10 представлены квартальные данные об объемах продаж и доходах текстильных корпораций США с первого квартала 1974 г. по третий квартал 1979 г. Введите сезонные фиктивные переменные и с помощью регрессии дохода на объем продаж исследуйте наличие или отсутствие сезонных колебаний.

**Таблица 4.10**

| Год | Квартал | Объем продаж | Доход |
|---|---|---|---|
| 1974 | I | 242.0 | 13.5 |
| | II | 269.4 | 16.3 |
| | III | 272.1 | 15.5 |
| | IV | 277.0 | 13.4 |
| 1975 | I | 247.1 | 9.3 |
| | II | 265.8 | 12.4 |
| | III | 271.0 | 13.2 |
| | IV | 281.3 | 14.2 |
| 1976 | I | 284.2 | 14.8 |
| | II | 307.6 | 18.1 |
| | III | 301.6 | 16.0 |
| | IV | 309.8 | 15.6 |
| 1977 | I | 311.5 | 15.6 |
| | II | 338.6 | 19.7 |
| | III | 331.7 | 16.7 |
| | IV | 346.2 | 18.4 |
| 1978 | I | 340.2 | 16.0 |
| | II | 377.5 | 22.1 |
| | III | 376.9 | 20.4 |
| | IV | 401.8 | 22.6 |
| 1979 | I | 406.2 | 22.6 |
| | II | 436.4 | 26.8 |
| | III | 437.5 | 24.8 |

Источник: *D.Salvatore. Statistics and Econometrics*, McGraw-Hill, 1982.

## Решение

Визуальный просмотр графика зависимости дохода от объемов продаж показывает целесообразность использования линейной регрессионной модели. Обозначим через $VOL$ объем продаж, а через $INC$ — доход фирм и оценим (с помощью эконометрического пакета) регрессию $INC = \beta_1 + \beta_2 VOL$ (см. таблицу 4.11).

**Таблица 4.11**

Dependent Variable: $INC$

| Variable | Coefficient | Std. Error | $t$-Statistic | Probability |
|---|---|---|---|---|
| const | −5.056 | 1.682 | −3.006 | 0.0067 |
| $VOL$ | 0.069 | 0.005 | 13.503 | 0.0000 |
| $R^2$ | 0.8967 | | | |

Полученные результаты подтверждают предварительный вывод: при относительно небольшом числе наблюдений значимость коэффициентов высокая.

Для ответа на поставленный вопрос введем фиктивные квартальные переменные $QRTi_t$: $QRTi_t = 1$, если наблюдение $t$ относится к $i$-му кварталу, $QRTi_t = 0$ в противном случае ($i = 1, 2, 3, 4$). Оценим теперь регрессию $INC = \beta_1 + \beta_2 VOL + \gamma_1 QRT1 + \gamma_2 QRT2 + \gamma_3 QRT3$ (переменную $QRT4$ не включаем в регрессию, чтобы избежать ситуации *«dummy trap»*, поскольку $QRT1 + QRT2 + QRT3 + QRT4 \equiv 1$) (см. таблицу 4.12).

**Таблица 4.12**

Dependent Variable: $INC$

| Variable | Coefficient | Std. Error | $t$-Statistic | Probability |
|---|---|---|---|---|
| const | −4.874 | 1.568 | −3.109 | 0.0061 |
| $VOL$ | 0.067 | 0.005 | 14.791 | 0.0000 |
| $QRT1$ | −0.329 | 0.749 | −0.440 | 0.6655 |
| $QRT2$ | 1.767 | 0.746 | 2.368 | 0.0293 |
| $QRT3$ | 0.350 | 0.746 | 0.470 | 0.6443 |
| $R^2$ | 0.9331 | | | |

Видим, что значимо (на 3%-ном уровне) отличается от нуля лишь коэффициент $\gamma_2$ при переменной $QRT2$. Соответствующий $F$-тест показывает, что коэффициенты $\gamma_1$, $\gamma_3$ совместно незначимы. Таким образом, статистически значимое среднее отклонение дохода происходит лишь во втором квартале. Оценка регрессии $INC$ на $VOL$ и $QRT2$ представлена в таблице 4.13.

**Таблица 4.13**

Dependent Variable: $INC$

| Variable | Coefficient | Std. Error | $t$-Statistic | Probability |
|---|---|---|---|---|
| const | −5.096 | 1.421 | −3.585 | 0.0018 |
| $VOL$ | 0.068 | 0.004 | 15.610 | 0.0000 |
| $QRT2$ | 1.750 | 0.570 | 3.069 | 0.0061 |
| $R^2$ | 0.9298 | | | |

Поставим теперь вопрос: оказывает ли влияние второй квартал на коэффициент $\beta_2$ при переменной $VOL$? Для этого оценим регрессию

$$INC = \beta_1 + \beta_2 VOL + \gamma_2 QRT2 + \delta_2 QRT2 \cdot VOL$$

(см. таблицу 4.14).

Таблица 4.14

Dependent Variable: $INC$

| Variable | Coefficient | Std. Error | $t$-Statistic | Probability |
|---|---|---|---|---|
| const | $-4.501$ | 1.701 | $-2.647$ | 0.0159 |
| $VOL$ | 0.066 | 0.005 | 12.611 | 0.0000 |
| $QRT2$ | $-0.359$ | 3.250 | $-0.110$ | 0.9137 |
| $QRT2 \cdot VOL$ | 0.006 | 0.010 | 0.659 | 0.5179 |
| $R^2$ | 0.9313 | | | |

Видим, что каждый из коэффициентов $\gamma_2$, $\delta_2$ незначим, однако проверка гипотезы $\mathrm{H}_0$: $\gamma_2 = 0$, $\delta_2 = 0$ дает значение соответствующей $F$-статистики 4.720 и $P$-значение 0.0207, что позволяет отвергнуть гипотезу $\mathrm{H}_0$ на 2%-ном уровне значимости. Если провести регрессию $INC$ только на $VOL$ и $QET2 \cdot VOL$, то коэффициент при $QET2 \cdot VOL$ оказывается значимым на 1%-ном уровне. Поэтому можно утверждать, что второй квартал оказывает влияние как на константу, так и на коэффициент при $VOL$ в регрессии $INC$ на $VOL$.

*Замечание.* Интересно отметить, что если оценить регрессию

$$\begin{aligned} INC = \beta_1 + \beta_2 VOL + \gamma_1 QRT1 + \gamma_2 QRT2 + \gamma_3 QRT3 \\ + \delta_1 QRT1 \cdot VOL + \delta_2 QRT2 \cdot VOL + \delta_3 QRT3 \cdot VOL \end{aligned}$$

и провести тест на совместную значимость всех коэффициентов $\gamma_i$, $\delta_i$ при переменных, содержащих $QRTi$, $i = 1, 2, 3$, то они окажутся совместно незначимыми. Оставляем читателю объяснение этого обстоятельства. Также рекомендуем читателю решить исходную задачу, рассматривая регрессию $\ln INC$ на $\ln VOL$.

## Задача 4.16

Таблица 4.15 содержит данные об объеме импорта $y$ (млрд. долл.), валовом национальном продукте $x_1$ (млрд. долл.) и индексе потребительских цен $x_2$ в США за период с 1964 по 1979 г.

а) Вычислите выборочный коэффициент корреляции между $x_1$ и $x_2$.

б) Оцените регрессию $y$ на константу и $x_1$.

в) Оцените регрессию $y$ на константу и $x_2$.

г) Оцените регрессию $y$ на константу, $x_1$ и $x_2$.

**Таблица 4.15**

| Год | $y$ | $x_1$ | $x_2$ | Год | $y$ | $x_1$ | $x_2$ |
|---|---|---|---|---|---|---|---|
| 1964 | 28.4 | 635.7 | 92.9 | 1972 | 75.9 | 1171.1 | 125.3 |
| 1965 | 32.0 | 688.1 | 94.5 | 1973 | 94.4 | 1306.6 | 133.1 |
| 1966 | 37.7 | 753.0 | 97.2 | 1974 | 131.9 | 1412.9 | 147.7 |
| 1967 | 40.6 | 796.3 | 100.0 | 1975 | 126.9 | 1528.8 | 161.2 |
| 1968 | 47.7 | 868.5 | 104.2 | 1976 | 155.4 | 1702.2 | 170.5 |
| 1969 | 52.9 | 935.5 | 109.8 | 1977 | 185.8 | 1899.5 | 181.5 |
| 1970 | 58.5 | 982.4 | 116.3 | 1978 | 217.5 | 2127.6 | 195.4 |
| 1971 | 64.0 | 1063.4 | 121.3 | 1979 | 260.9 | 2368.5 | 217.4 |

Источник: *D.Salvatore. Statistics and Econometrics*, McGraw-Hill, 1982.

Как можно проинтерпретировать полученные результаты? Можно ли ограничиться только одной из регрессий б) или в)?

## Решение

а) Непосредственными вычислениями получаем, что выборочный коэффициент корреляции между $x_1$ и $x_2$ равен 0.997. Такое высокое значение может привести к мультиколлинеарности в регрессиях, содержащих одновременно $x_1$ и $x_2$ в качестве объясняющих переменных.

б), в) В таблицах 4.16 и 4.17 приведены результаты регрессий $y$ на $x_1$ и $x_2$ (и константу).

**Таблица 4.16**

Dependent Variable: $y$

| Variable | Coefficient | Std. Error | $t$-Statistic | Probability |
|---|---|---|---|---|
| const | −69.027 | 5.750 | −12.004 | 0.0000 |
| $x_1$ | 0.134 | 0.004 | 31.866 | 0.0000 |
| $R^2$ | 0.9864 | | | |

**Таблица 4.17**

Dependent Variable: $y$

| Variable | Coefficient | Std. Error | $t$-Statistic | Probability |
|---|---|---|---|---|
| const | −146.516 | 8.335 | −17.579 | 0.0000 |
| $x_2$ | 1.824 | 0.060 | 30.795 | 0.0000 |
| $R^2$ | 0.9855 | | | |

Видим, что качество обеих регрессий высокое, иными словами, каждая из переменных $x_1$ и $x_2$ хорошо объясняет изменение $y$.

г) Регрессия $y$ на $x_1$, $x_2$ и константу дает результаты, представленные в таблице 4.18.

**Таблица 4.18**

Dependent Variable: $y$

| Variable | Coefficient | Std. Error | $t$-Statistic | Probability |
|---|---|---|---|---|
| const | −101.489 | 33.080 | −3.068 | 0.0090 |
| $x_1$ | 0.079 | 0.560 | 1.403 | 0.1839 |
| $x_2$ | 0.759 | 0.761 | 0.996 | 0.3372 |
| $R^2$ | 0.9874 | | | |

Как и следовало ожидать, значимость коэффициентов резко упала, что является следствием высокой мультиколлинеарности между регрессорами.

Если необходимо прогнозировать значение $y$, то для этой цели можно успешно использовать любую из регрессий б) или в), поскольку высокая значимость коэффициентов и большое значение коэффициента детерминации $R^2$ обеспечивают высокую прогностическую силу этих моделей. Однако если необходимо оценить влияние каждого из факторов $x_1$ и $x_2$ на $y$, то приходится использовать регрессию г), хотя точность оценок будет невелика.

Рекомендуем читателю исследовать логарифмическую модель.

## Задача 4.17

В таблице 4.19 представлены выпуск $Q$, трудозатраты $L$ и капиталовложения $K$ 15 фирм некоторой отрасли.

**Таблица 4.19**

| Фирма | $Q$ | $L$ | $K$ | Фирма | $Q$ | $L$ | $K$ |
|---|---|---|---|---|---|---|---|
| 1 | 2350 | 2334 | 1570 | 9 | 2550 | 2446 | 1880 |
| 2 | 2470 | 2425 | 1850 | 10 | 2450 | 2403 | 1790 |
| 3 | 2110 | 2230 | 1150 | 11 | 2290 | 2301 | 1480 |
| 4 | 2560 | 2463 | 1940 | 12 | 2160 | 2253 | 1240 |
| 5 | 2650 | 2565 | 2450 | 13 | 2400 | 2367 | 1660 |
| 6 | 2240 | 2278 | 1340 | 14 | 2490 | 2430 | 1850 |
| 7 | 2430 | 2380 | 1700 | 15 | 2590 | 2470 | 2000 |
| 8 | 2530 | 2437 | 1860 | | | | |

а) Оцените по этим данным производственную функцию Кобба–Дугласа $Q = \alpha L^{\beta_1} K^{\beta_2}$, вычислите коэффициент детерминации, скорректированный коэффициент детерминации и выборочный коэффициент корреляции между $\ln L$ и $\ln K$.

б) Проведите регрессию $\ln Q$ только на $\ln K$.

Как можно проинтерпретировать полученные результаты?

## Решение

а) Прологарифмируем выражение для производственной функции Кобба–Дугласа: $\ln Q = \ln \alpha + \beta_1 \ln L + \beta_2 \ln K$. Оценка этой регрессии дает следующие результаты:

Dependent Variable: $\ln Q$

| Variable | Coefficient | Std. Error | $t$-Statistic | Probability |
|---|---|---|---|---|
| const | 0.501 | 4.480 | 0.112 | 0.9129 |
| $\ln L$ | 0.758 | 0.707 | 1.071 | 0.3052 |
| $\ln K$ | 0.188 | 0.139 | 1.356 | 0.2001 |
| $R^2$ | 0.9689 | | | |

Вычисляя скорректированный коэффициент детерминации по формуле

$$R^2_{\text{adj}} = 1 - (1 - R^2)\frac{n-1}{n-k}$$

(см. разд. 3.4 гл. 3), получаем $R^2_{\text{adj}} = 0.9637$.

Выборочный коэффициент корреляции между $\ln L$ и $\ln K$ равен 0.992, что и является причиной незначимости коэффициентов $\beta_1$, $\beta_2$ (аналогично предыдущей задаче 4.16).

б) Регрессия $\ln Q$ на $\ln K$ и константу дает следующие результаты:

Dependent Variable: $\ln Q$

| Variable | Coefficient | Std. Error | $t$-Statistic | Probability |
|---|---|---|---|---|
| const | 5.297 | 0.130 | 40.778 | 0.0000 |
| $\ln K$ | 0.335 | 0.017 | 19.192 | 0.0000 |
| $R^2$ | 0.9659 | | | |

Видим, что в этой регрессии коэффициенты значимы и коэффициент детерминации незначительно отличается от коэффициента детерминации первой регрессии, как и следовало ожидать.

Так же, как и в предыдущей задаче 4.16, регрессия $\ln Q$ только на $\ln K$ (или на $\ln L$) может применяться в целях получения прогноза, но, например, для определения эластичностей выпуска по труду и капиталу необходимо использовать модель регрессии на обе переменные, что, как мы видим, дает оценки этих эластичностей с большими ошибками вследствие мультиколлинеарности. Один из возможных способов борьбы с ней предлагается в следующей задаче 4.18.

## Задача 4.18

Можно ли преодолеть проблему мультиколлинеарности, возникающую в задаче 4.17, если известно, что производственная функция обладает постоянной отдачей на масштаб ($\beta_1 + \beta_2 = 1$)?

## Решение

Тестируя для регрессии $\ln Q = \ln\alpha + \beta_1 \ln L + \beta_2 \ln K$ п. а) предыдущей задачи 4.17 гипотезу $\mathrm{H}_0 : \beta_1 + \beta_2 = 1$ (постоянство отдачи на масштаб), получаем значение $F$-статистики 0.009 и $P$-значение 0.9255, что позволяет уверенно не отвергать эту гипотезу. Подставляя $\beta_2 = 1 - \beta_1$ в исходное уравнение, после элементарных преобразований получаем:

$$\ln\frac{Q}{K} = \ln\alpha + \beta_1 \ln\frac{L}{K}.$$

Оценивая эту регрессию, получаем следующие результаты:

Dependent Variable: $\ln Q/K$

| Variable | Coefficient | Std. Error | $t$-Statistic | Probability |
|---|---|---|---|---|
| const | 0.073 | 0.008 | 9.261 | 0.0000 |
| $\ln L/K$ | 0.825 | 0.021 | 39.812 | 0.0000 |
| $R^2$ | 0.9919 | | | |

Качество регрессии хорошее. Таким образом, введение ограничения $\beta_1 + \beta_2 = 1$ позволило в данном случае решить проблему мультиколлинеарности.

## Задача 4.19

Пусть $\boldsymbol{y} = \boldsymbol{X\beta} + \boldsymbol{\varepsilon}$ — стандартная модель множественной регрессии и $\widehat{\boldsymbol{\beta}}$ — МНК-оценка вектора коэффициентов $\boldsymbol{\beta}$.

а) Покажите, каким образом можно использовать $\widehat{\boldsymbol{\beta}}$ для получения более эффективной оценки параметров $\boldsymbol{\beta}$, если известно, что $\boldsymbol{\beta}$ удовлетворяет линейному ограничению $\boldsymbol{R\beta} = \boldsymbol{r}$.

б) Для модели $y_t = \alpha + \beta x_{t1} + \gamma x_{t2} + \varepsilon_t$ по $n = 100$ наблюдениям получены следующие данные (матрица сумм произведений соответствующих переменных):

| | $y-\overline{y}$ | $x_1-\overline{x}_1$ | $x_2-\overline{x}_2$ |
|---|---|---|---|
| $y-\overline{y}$ | 2000 | 100 | 90 |
| $x_1-\overline{x}_1$ | 100 | 10 | 5 |
| $x_2-\overline{x}_2$ | 90 | 5 | 5 |

Проверьте гипотезу $\mathrm{H}_0$: $5\beta = \gamma$ против альтернативы $\mathrm{H}_1$: $5\beta \neq \gamma$.

## Решение

а) Если известно, что вектор параметров $\boldsymbol{\beta}$ удовлетворяет ограничению $\boldsymbol{R\beta} = \boldsymbol{r}$, то естественно пытаться улучшить МНК-оценку $\widehat{\boldsymbol{\beta}}$, решая задачу минимизации суммы квадратов отклонений при этом ограничении:

$$(\boldsymbol{y} - \boldsymbol{X\beta})'(\boldsymbol{y} - \boldsymbol{X\beta}) \to \min,$$
$$\boldsymbol{R\beta} = \boldsymbol{r}.$$

Эта задача была решена в упражнении 3.13:

$$\widehat{\boldsymbol{\beta}}_{\mathrm{R}} = \widehat{\boldsymbol{\beta}}_{\mathrm{UR}} + (\boldsymbol{X}'\boldsymbol{X})^{-1}\boldsymbol{R}'\left(\boldsymbol{R}(\boldsymbol{X}'\boldsymbol{X})^{-1}\boldsymbol{R}'\right)^{-1}(\boldsymbol{r} - \boldsymbol{R}\widehat{\boldsymbol{\beta}}_{\mathrm{UR}}),$$

где $\widehat{\boldsymbol{\beta}}_{\mathrm{UR}} = \widehat{\boldsymbol{\beta}}$. (Напомним, что индексы «R» и «UR» означают «с ограничением» (*restricted*) и «без ограничения» (*unrestricted*) соответственно.) Покажем, что оценка $\widehat{\boldsymbol{\beta}}_{\mathrm{R}}$ является более точной, чем оценка $\widehat{\boldsymbol{\beta}}$. Действительно, поскольку $\mathrm{E}(\widehat{\boldsymbol{\beta}}) = \boldsymbol{\beta}$ (несмещенность), то

$$\mathrm{E}(\widehat{\boldsymbol{\beta}}_{\mathrm{R}}) = \boldsymbol{\beta} + (\boldsymbol{X}'\boldsymbol{X})^{-1}\boldsymbol{R}'\left(\boldsymbol{R}(\boldsymbol{X}'\boldsymbol{X})^{-1}\boldsymbol{R}'\right)^{-1}(\boldsymbol{r} - \boldsymbol{R}\mathrm{E}(\widehat{\boldsymbol{\beta}}_{\mathrm{UR}})) = \boldsymbol{\beta}$$

в силу ограничения, т. е. оценка $\widehat{\boldsymbol{\beta}}_{\mathrm{R}}$ является несмещенной. Перепишем выражение для $\widehat{\boldsymbol{\beta}}_{\mathrm{R}}$:

$$\widehat{\boldsymbol{\beta}}_{\mathrm{R}} = \left(\boldsymbol{I} - (\boldsymbol{X}'\boldsymbol{X})^{-1}\boldsymbol{R}'\left(\boldsymbol{R}(\boldsymbol{X}'\boldsymbol{X})^{-1}\boldsymbol{R}'\right)^{-1}\boldsymbol{R}\right)\widehat{\boldsymbol{\beta}} + (\boldsymbol{X}'\boldsymbol{X})^{-1}\boldsymbol{R}'\left(\boldsymbol{R}(\boldsymbol{X}'\boldsymbol{X})^{-1}\boldsymbol{R}'\right)^{-1}\boldsymbol{r},$$

заметим, что последнее слагаемое в этом представлении — неслучайный вектор, и обозначим для краткости

$$\boldsymbol{D} = \boldsymbol{I} - (\boldsymbol{X}'\boldsymbol{X})^{-1}\boldsymbol{R}'\left(\boldsymbol{R}(\boldsymbol{X}'\boldsymbol{X})^{-1}\boldsymbol{R}'\right)^{-1}\boldsymbol{R}.$$

Тогда из свойств матрицы ковариаций следует, что

$$\mathrm{V}(\widehat{\boldsymbol{\beta}}_{\mathrm{R}}) = \boldsymbol{D}\mathrm{V}(\widehat{\boldsymbol{\beta}})\boldsymbol{D}'.$$

Учитывая, что $\mathrm{V}(\widehat{\boldsymbol{\beta}}) = \sigma^2(\boldsymbol{X}'\boldsymbol{X})^{-1}$, и проводя элементарные вычисления, получаем:

$$\mathrm{V}(\widehat{\boldsymbol{\beta}}) - \mathrm{V}(\widehat{\boldsymbol{\beta}}_{\mathrm{R}}) = \sigma^2(\boldsymbol{X}'\boldsymbol{X})^{-1}\boldsymbol{R}'(\boldsymbol{R}(\boldsymbol{X}'\boldsymbol{X})^{-1}\boldsymbol{R}')^{-1}\boldsymbol{R}(\boldsymbol{X}'\boldsymbol{X})^{-1}.$$

Матрица $\left(\boldsymbol{R}(\boldsymbol{X}'\boldsymbol{X})^{-1}\boldsymbol{R}'\right)^{-1}$ является симметричной и положительно определенной (это следует из того, что $\sigma^2\boldsymbol{R}(\boldsymbol{X}'\boldsymbol{X})^{-1}\boldsymbol{R}'$ — ковариационная матрица случайного вектора $\widehat{\boldsymbol{\beta}}$), поэтому (см. разд. 5.2 главы 5) существует такая матрица $\boldsymbol{P}$, что $\left(\boldsymbol{R}(\boldsymbol{X}'\boldsymbol{X})^{-1}\boldsymbol{R}'\right)^{-1} = \boldsymbol{P}'\boldsymbol{P}$. Наконец, обозначая $\boldsymbol{A} = (\boldsymbol{X}'\boldsymbol{X})^{-1}\boldsymbol{R}'$, получаем:

$$\mathrm{V}(\widehat{\boldsymbol{\beta}}) - \mathrm{V}(\widehat{\boldsymbol{\beta}}_{\mathrm{R}}) = \sigma^2\boldsymbol{AP}'\boldsymbol{PA}' = \sigma^2(\boldsymbol{AP}')(\boldsymbol{AP}')' \geqslant \boldsymbol{0}.$$

б) Установим следующий простой факт: МНК-оценки параметров $\beta$ и $\gamma$ в регрессии $y_t = \alpha + \beta x_{t1} + \gamma x_{t2} + \varepsilon_t$ (т. е. в регрессии $\boldsymbol{y}$ на $\boldsymbol{\imath}$, $\boldsymbol{x}_1$, $\boldsymbol{x}_2$ совпадают с МНК-оценками в регрессии $\boldsymbol{y}_* = \boldsymbol{y} - \overline{y}\boldsymbol{\imath}$ на $\boldsymbol{x}_{*1} = \boldsymbol{x}_1 - \overline{x}_1\boldsymbol{\imath}$, $\boldsymbol{x}_{*2} = \boldsymbol{x}_2 - \overline{x}_2\boldsymbol{\imath}$. Действительно, пусть $\widehat{\alpha}, \widehat{\beta}, \widehat{\gamma}$ — МНК-оценки, $\boldsymbol{e}$ — вектор остатков в исходной регрессии. Тогда

$$y_t = \widehat{\alpha} + \widehat{\beta} x_{t1} + \widehat{\gamma} x_{t2} + e_t,$$

откуда

$$\overline{y} = \widehat{\alpha} + \widehat{\beta}\overline{x}_1 + \widehat{\gamma}\overline{x}_2,$$

поскольку $\overline{e} = 0$. Вычитая почленно два последних равенства, получаем:

$$y_t - \overline{y} = \widehat{\beta}(x_{t1} - \overline{x}_1) + \widehat{\gamma}(x_{t2} - \overline{x}_2) + e_t,$$

что эквивалентно представлению

$$\boldsymbol{y}_* = \boldsymbol{x}_{*1}\widehat{\beta} + \boldsymbol{x}_{*2}\widehat{\gamma} + \boldsymbol{e}.$$

Отсюда следует требуемое утверждение, так как вектор $\boldsymbol{e}$ ортогонален каждому вектору $\boldsymbol{x}_{*1}$, $\boldsymbol{x}_{*2}$.

Таким образом, если обозначить $\widehat{\boldsymbol{\delta}} = (\widehat{\beta}, \widehat{\gamma})'$, то

$$\widehat{\boldsymbol{\delta}} = (\boldsymbol{X}'\boldsymbol{X})^{-1}\boldsymbol{X}'\boldsymbol{y}_*, \quad \text{где} \quad \boldsymbol{X}'\boldsymbol{X} = \begin{bmatrix} 10 & 5 \\ 5 & 5 \end{bmatrix}, \quad \boldsymbol{X}'\boldsymbol{y}_* = \begin{bmatrix} 100 \\ 90 \end{bmatrix}.$$

Проводя элементарные вычисления, получаем:

$$(\boldsymbol{X}'\boldsymbol{X})^{-1} = \begin{bmatrix} 0.2 & -0.2 \\ -0.2 & 0.4 \end{bmatrix}, \quad \widehat{\beta} = 2, \quad \widehat{\gamma} = 16.$$

Найдем сумму квадратов остатков $\boldsymbol{e}'\boldsymbol{e}$. Имеем

$$\begin{aligned} \boldsymbol{e}'\boldsymbol{e} &= \boldsymbol{e}'(\boldsymbol{y}_* - \boldsymbol{X}\widehat{\boldsymbol{\delta}}) = \boldsymbol{e}'\boldsymbol{y}_* = (\boldsymbol{y}_* - \boldsymbol{X}\widehat{\boldsymbol{\delta}})'\boldsymbol{y}_* = \boldsymbol{y}_*'\boldsymbol{y}_* - \widehat{\boldsymbol{\delta}}'\boldsymbol{X}'\boldsymbol{y}_* \\ &= 2000 - (2 \cdot 100 + 16 \cdot 90) = 2000 - 1640 = 360. \end{aligned}$$

Число степеней свободы равно 97, поэтому $\widehat{\sigma}^2 = 360/97 = 3.711$. Таким образом,

$$\widehat{\mathrm{V}}(\widehat{\boldsymbol{\delta}}) = \begin{bmatrix} 0.742 & -0.742 \\ -0.742 & 1.484 \end{bmatrix}.$$

Проверим гипотезу $\mathrm{H}_0$: $5\beta = \gamma$ с помощью статистики

$$t = \frac{5\widehat{\beta} - \widehat{\gamma}}{\widehat{\sigma}_{5\widehat{\beta} - \widehat{\gamma}}}$$

(см. разд. 3.5 гл. 3). Имеем

$$\begin{aligned} \widehat{\sigma}^2_{5\widehat{\beta} - \widehat{\gamma}} &= 25 \cdot 0.742 + 1.484 - 10 \cdot (-0.742) = 27.464, \\ \widehat{\sigma}_{5\widehat{\beta} - \widehat{\gamma}} &= 5.241, \\ t &= (-6)/5.241 = -1.145. \end{aligned}$$

Поскольку $|t| < 1.985 = t_{0.95}(97)$, то на 5%-ном уровне значимости гипотеза $H_0$ не отвергается.

## Задача 4.20

С помощью обычного метода наименьших квадратов получены две спецификации модели: $y = \widehat{\alpha} + \widehat{\beta}x + e$ и $y = \widehat{\alpha}^* + \widehat{\beta}^* x + \widehat{\gamma}^* z + u$, где $e$, $u$ — остатки соответствующих регрессий. Объясните, при каких обстоятельствах выполнены следующие условия:

а) $\widehat{\beta} = \widehat{\beta}^*$;

б) $\sum u_i^2 \leqslant \sum e_i^2$;

в) оценка $\widehat{\beta}$ статистически значима на 5%-ном уровне, а оценка $\widehat{\beta}^*$ незначима;

г) оценка $\widehat{\beta}^*$ статистически значима на 5%-ном уровне, а оценка $\widehat{\beta}$ незначима.

## Решение

Эта задача связана с задачей 3.17. Введем следующие обозначения: $X_1 = [\imath \;\; x]$ и $X_2 = [\imath \;\; x \;\; z] = [X_1 \;\; z]$; $M_i = I - X_i(X_i'X_i)^{-1}X_i'$, $i = 1, 2$. Через $\widehat{b} = (\widehat{\alpha}, \widehat{\beta})'$ и $\widehat{b}^* = (\widehat{\alpha}^*, \widehat{\beta}^*, \widehat{\gamma}^*)'$ обозначим векторы оценок коэффициентов в короткой и длинной регрессиях. В приведенных обозначениях имеем:

$$\widehat{b} = (X_1'X_1)^{-1}X_1'y, \quad \widehat{b}^* = (X_2'X_2)^{-1}X_2'y.$$

Матрица, обратная матрице

$$X_2'X_2 = \begin{bmatrix} X_1'X_1 & X_1'z \\ z'X_1 & z'z \end{bmatrix},$$

вычисляется по формуле (см. задачу 3.17):

$$(X_2'X_2)^{-1} = \begin{bmatrix} Q + QX_1'z(z'M_1z)^{-1}z'X_1Q & -QX_1'z(z'M_1z)^{-1} \\ -(z'M_1z)^{-1}z'X_1Q & (z'M_1z)^{-1} \end{bmatrix},$$

где $Q = (X_1'X_1)^{-1}$. Отсюда получаем:

$$\begin{aligned} \begin{bmatrix} \widehat{\alpha}^* \\ \widehat{\beta}^* \end{bmatrix} &= (X_1'X_1)^{-1}\left(I + X_1'z(z'M_1z)^{-1}z'X_1(X_1'X_1)^{-1}\right)X_1'y \\ &\quad - (X_1'X_1)^{-1}X_1'z(z'M_1z)^{-1}z'y \\ &= \widehat{b} + \frac{(X_1'X_1)^{-1}\left(X_1'zz'X_1(X_1'X_1)^{-1}X_1'y - X_1'z(z'y)\right)}{z'M_1z} \\ &= \widehat{b} - \frac{z'M_1y}{z'M_1z}(X_1'X_1)^{-1}X_1'z = \widehat{b} - \widehat{\gamma}^* q, \end{aligned}$$

$$\widehat{\gamma}^* = \frac{z'M_1y}{z'M_1z}, \quad q = (X_1'X_1)^{-1}X_1'z.$$

Отсюда получаем, что

$$\widehat{\beta}^* = \widehat{\beta} - q_2\widehat{\gamma}^*, \quad q_2 = \frac{\sum_{i=1}^n (x_i - \overline{x})(z_i - \overline{z})}{\sum_{i=1}^n (x_i - \overline{x})^2}.$$

а) $\widehat{\beta}^* = \widehat{\beta}$ тогда и только тогда, когда $q_2 = 0$ или $\widehat{\gamma}^* = 0$, т.е. тогда и только тогда, когда $\sum_{i=1}^n (x_i - \overline{x})(z_i - \overline{z}) = 0$ или $\boldsymbol{z}'\boldsymbol{M}_1\boldsymbol{y} = 0$. (Заметим, что $\boldsymbol{z}'\boldsymbol{M}_1\boldsymbol{z} \neq 0$, так как в противном случае столбцы матрицы $\boldsymbol{X}_2$ были бы линейно зависимы.) Заметим, что если $\boldsymbol{x}$ и $\boldsymbol{z}$ некоррелированы, то добавление $\boldsymbol{z}$ в число регрессоров не изменяет оценку параметра $\beta$. Поэтому условие а) выполнено, если $\boldsymbol{x}$ и $\boldsymbol{z}$ некоррелированы.

б) Необходимо показать, что $\boldsymbol{u}'\boldsymbol{u} - \boldsymbol{e}'\boldsymbol{e} \leqslant 0$. Поскольку $\boldsymbol{u} = \boldsymbol{M}_2\boldsymbol{y}$ и $\boldsymbol{e} = \boldsymbol{M}_1\boldsymbol{y}$, то из решения задачи 3.17 следует, что

$$\begin{aligned}\boldsymbol{u}'\boldsymbol{u} - \boldsymbol{e}'\boldsymbol{e} &= \boldsymbol{y}'(\boldsymbol{M}_2 - \boldsymbol{M}_1)\boldsymbol{y} = -\boldsymbol{y}'\boldsymbol{M}_1\boldsymbol{z}(\boldsymbol{z}'\boldsymbol{M}_1\boldsymbol{z})^{-1}\boldsymbol{z}'\boldsymbol{M}_1\boldsymbol{y} \\ &= -(\boldsymbol{z}'\boldsymbol{M}_1\boldsymbol{z})(\widehat{\gamma}^*)^2 \leqslant 0,\end{aligned}$$

причем равенство выполняется в том и только том случае, когда $\widehat{\gamma}^* = 0$, или, иначе, тогда и только тогда, когда $\boldsymbol{z}'\boldsymbol{M}_1\boldsymbol{y} = 0$.

Таким образом, б) верно, поскольку добавление независимой переменной не может увеличить (а в общем случае — уменьшает) сумму квадратов остатков.

в) и г) Обозначим через $s^2$ оценку дисперсии $\sigma^2$ в короткой регрессии, а через $s_*^2$ — оценку дисперсии $\sigma^2$ в длинной регрессии. Заметим, что

$$\mathrm{V}(\widehat{\beta}) = \frac{\sigma^2}{\sum_i (x_i - \overline{x})^2} = \sigma^2\omega_1^2$$

и

$$\mathrm{V}(\widehat{\beta}^*) = \frac{\sigma^2}{\sum_i (x_i - \overline{x})^2} + \frac{\sigma^2 q_2^2}{\boldsymbol{z}'\boldsymbol{M}_1\boldsymbol{z}} = \sigma^2(\omega_1^2 + \omega_2^2).$$

Пусть $t$ и $t^*$ обозначают $t$-статистики оценки коэффициента $\beta$ в соответственно короткой и длинной регрессии. Тогда

$$t = \frac{\widehat{\beta}}{s\omega_1}, \quad t^* = \frac{\widehat{\beta}^*}{\sqrt{s_*^2(\omega_1^2 + \omega_2^2)}}.$$

Ситуация в) встречается, когда выполняются неравенства $|t| > c$ и $|t^*| \leqslant c$ для $c = 1.96$. Другими словами, это происходит в том случае, когда

$$\frac{(\widehat{\beta}^*)^2}{s_*^2(\omega_1^2 + \omega_2^2)} \leqslant c^2 < \frac{\widehat{\beta}^2}{s^2\omega_1^2}.$$

Таким образом, в) может случиться, если регрессор $\boldsymbol{z}$ сильно коррелирован с $\boldsymbol{x}$.

Ситуация г) встречается, когда выполняется неравенство

$$\frac{\widehat{\beta}^2}{s^2\omega_1^2} \leqslant c^2 < \frac{(\widehat{\beta}^*)^2}{s_*^2(\omega_1^2+\omega_2^2)}.$$

Это может произойти в том случае, когда регрессор $z$ слабо коррелирован с $x$, но хорошо объясняет изменение зависимой переменной, и добавление $z$ в число регрессоров значительно уменьшает вариацию остатков. В качестве предельного случая рассмотрим ситуацию, когда $z$ и $x$ некоррелированы. Тогда $\widehat{\beta}^* = \widehat{\beta}$, но, возможно, оценка дисперсии $\sigma^2$ значительно уменьшается при включении $z$. В этом случае $s_*^2$ значительно больше, чем $s^2$, и утверждение г) справедливо.

## Задача 4.21

Дана стандартная модель множественной регрессии $y = X\beta + \varepsilon$.

а) Выразите матрицу ковариаций МНК-оценки вектора $\beta$ в терминах собственных значений и собственных векторов матрицы $X'X$.

б) Объясните, как соотносится результат а) с проблемой мультиколлинеарности.

## Решение

а) Матрица $X'X$ симметрична и положительно определена, поэтому существует ортонормированная система $c_1, \ldots, c_k$ ее собственных векторов с собственными значениями $\lambda_i > 0$, $i = 1, \ldots, k$ (см. разд. 14 приложения ЛА). При этом $\Lambda = O^{-1}(X'X)O$, где $O$ — ортогональная матрица ($O^{-1} = O'$), столбцы которой составлены из координат векторов $c_i$: $O = [c_1, \ldots, c_k]$, а $\Lambda$ — диагональная матрица, на главной диагонали которой стоят числа $\lambda_i$, $i = 1, \ldots, k$ (см. (ЛА.11)). Отсюда следует, что $(X'X)^{-1} = O\Lambda^{-1}O'$. Окончательно получаем:

$$\mathrm{V}(\widehat{\beta}) = \sigma^2(X'X)^{-1} = \sigma^2 O\Lambda^{-1}O'.$$

б) Из последнего равенства следует, что

$$\left(\mathrm{V}(\widehat{\beta})\right)_{ij} = \sigma^2 \sum_{t=1}^{k} \frac{c_{it}c_{jt}}{\lambda_t},$$

где $c_{ij}$ — $(i, j)$-й элемент матрицы $O$ ($i$-я координата вектора $c_j$). В частности,

$$\mathrm{V}(\widehat{\beta}_i) = \sigma^2 \sum_{t=1}^{k} \frac{c_{it}^2}{\lambda_t}.$$

Пусть $\lambda_{\min}$ — минимальное собственное число матрицы $X'X$, причем $\lambda_{\min} = \lambda_s$. Так как столбец с номером $s$ матрицы $O$ является собственным вектором матрицы $X'X$, имеющим длину 1, то среди его элементов обязательно найдется такой элемент $c_{rs}$, что $c_{rs}^2 \geqslant (1/k)$. Из последнего равенства следует, что

$$\mathrm{V}(\widehat{\beta}_r) \geqslant \sigma^2 \frac{c_{rs}^2}{\lambda_{\min}} \geqslant \frac{\sigma^2}{k\lambda_{\min}}.$$

В случае мультиколлинеарности матрица $X'X$, говоря нестрого, близка к вырожденной. Это означает, что число $\lambda_{\min}$ близко к нулю. Полученный результат показывает, что какие-то из МНК-оценок обязательно будут иметь большие ошибки.

## Задача 4.22

Дана модель множественной регрессии $y = X_1\beta_1 + X_2\beta_2 + \varepsilon$, где $\beta_1$, $\beta_2$ — векторы размерности $k_1$, $k_2$ соответственно. Предположим, что у вектора $\beta_1$ есть несмещенная оценка $\widehat{\beta}_1$, некоррелированная с $\varepsilon$, с известной ковариационной матрицей $V_1$.

а) Вычислите ковариационную матрицу оценки вектора $\beta_2$, получаемой регрессией $y - X_1\widehat{\beta}_1$ на $X_2$.

б) Сравните ковариационную матрицу, полученную в а), с ковариационной матрицей МНК-оценки вектора $\beta_2$ в исходной модели.

в) Можете ли вы предложить более эффективную оценку вектора $\beta_2$, чем те, что получены в а) и б)?

г) Как полученные результаты связаны с проблемой мультиколлинеарности?

## Решение

а) Нам известно, что $\mathrm{E}(\widehat{\beta}_1) = \widehat{\beta}$, $\mathrm{V}(\widehat{\beta}_1) = V_1$ и $\varepsilon$ и $\widehat{\beta}_1$ некоррелированы, т.е. $\mathrm{Cov}(\widehat{\beta}_1, \varepsilon) = \mathrm{E}(\widehat{\beta}_1\varepsilon') = \mathbf{0}$. Обозначим через $y_2$ разность $y_2 = y - X_1\widehat{\beta}_1$. Регрессия $y - X_1\widehat{\beta}_1$ на $X_2$ может быть записана в виде

$$y_2 = X_2\beta_2 + u,$$

где $u = \varepsilon - X_1(\widehat{\beta}_1 - \beta_1)$. Математическое ожидание вектора ошибок $u$ равно $\mathbf{0}$, а матрица ковариаций равна матрице $\mathrm{V}(u) = \sigma^2 I_n + X_1 V_1 X_1'$, вообще говоря, не являющейся скалярной. МНК-оценка вектора коэффициентов $\beta_2$ в этой новой регрессии может быть записана в виде

$$\widehat{\beta}_2^{(2)} = (X_2'X_2)^{-1}X_2'y_2 = \beta_2 + (X_2'X_2)^{-1}X_2'u$$

или

$$\widehat{\beta}_2^{(2)} - \beta_2 = (X_2'X_2)^{-1}X_2'u = (X_2'X_2)^{-1}\Big(X_2'\varepsilon - X_2'X_1(\widehat{\beta}_1 - \beta_1)\Big).$$

Дисперсия оценки $\widehat{\beta}_2^{(2)}$ равна

$$\mathrm{V}(\widehat{\beta}_2^{(2)}) = \sigma^2(X_2'X_2)^{-1} + (X_2'X_2)^{-1}X_2'X_1V_1X_1'X_2(X_2'X_2)^{-1}.$$

б) МНК-оценка вектора $\beta_2$ в регрессии

$$y = X_1\beta_1 + X_2\beta_2 + \varepsilon$$

может быть записана в виде

$$\widehat{\beta}_2^{(1)} = (X_2'M_1X_2)^{-1}X_2'M_1y = \beta_2 + (X_2'M_1X_2)^{-1}X_2'M_1\varepsilon.$$

Отсюда получаем:

$$\widehat{\beta}_2^{(1)} - \beta_2 = (X_2'M_1X_2)^{-1}X_2'M_1\varepsilon$$

и

$$\mathrm{V}(\widehat{\beta}_2^{(1)}) = \sigma^2(X_2'M_1X_2)^{-1} = \sigma^2(X_2'(I_n - X_1(X_1'X_1)^{-1}X_1')X_2)^{-1}.$$

Обе оценки $\widehat{\beta}_2^{(1)}$ и $\widehat{\beta}_2^{(2)}$ являются несмещенными. От конкретного вида матрицы $V_1$ зависит, какая из оценок имеет меньшую дисперсию.

в) Рассмотрим полученные выше оценки. Их можно представить в виде

$$\begin{aligned}\widehat{\beta}_2^{(2)} &= (X_2'X_2)^{-1}X_2'(y - X_1\widehat{\beta}_1),\\ \widehat{\beta}_2^{(1)} &= (X_2'M_1X_2)^{-1}X_2'M_1y.\end{aligned}$$

Обе оценки несмещенные. Рассмотрим семейство оценок вида

$$\widehat{\beta}_2 = A(y - X_1\widehat{\beta}_1).$$

Оценка $\widehat{\beta}_2$ является несмещенной в том и только том случае, когда выполняется условие $AX_2 = I$, которое мы будем в дальнейшем предполагать выполненным. Две оценки, рассмотренные выше, являются частными случаями последней при специальном выборе матрицы $A$:

$$\begin{aligned}A^{(2)} &= (X_2'X_2)^{-1}X_2' & &(\text{для } \widehat{\beta}_2^{(2)}),\\ A^{(1)} &= (X_2'M_1X_2)^{-1}X_2'M_1 & &(\text{для } \widehat{\beta}_2^{(1)}).\end{aligned}$$

Рассмотрим семейство матриц $A$ вида

$$A = (X_2'WX_2)^{-1}X_2'W,$$

где $\boldsymbol{W}$ – симметричная неотрицательно определенная матрица. (При $\boldsymbol{W} = \boldsymbol{I}$ и $\boldsymbol{W} = \boldsymbol{M}_1$ получаем две предыдущие оценки.) Поскольку $\boldsymbol{AX}_2 = \boldsymbol{I}$, то оценка $\widehat{\boldsymbol{\beta}}_2$ несмещенная, $\mathrm{E}(\widehat{\boldsymbol{\beta}}_2) = \boldsymbol{\beta}_2$. Матрица ковариаций равна

$$\mathrm{V}(\widehat{\boldsymbol{\beta}}_2) = (\boldsymbol{X}_2'\boldsymbol{W}\boldsymbol{X}_2)^{-1}\boldsymbol{X}_2'\boldsymbol{W}(\sigma^2\boldsymbol{I} + \boldsymbol{X}_1\boldsymbol{V}_1\boldsymbol{X}_1')\boldsymbol{W}\boldsymbol{X}_2(\boldsymbol{X}_2'\boldsymbol{W}\boldsymbol{X}_2)^{-1}.$$

Выберем теперь «оптимальную» матрицу $\boldsymbol{W}$, минимизирующую матрицу ковариаций $\mathrm{V}(\widehat{\boldsymbol{\beta}}_2)$. Покажем, что решением этой задачи является матрица

$$\boldsymbol{W} = (\sigma^2\boldsymbol{I} + \boldsymbol{X}_1\boldsymbol{V}_1\boldsymbol{X}_1')^{-1},$$

при этом

$$\mathrm{V}(\widehat{\boldsymbol{\beta}}_2) = (\boldsymbol{X}_2'\boldsymbol{W}\boldsymbol{X}_2)^{-1}.$$

Для доказательства напомним, что для регрессионной модели

$$\boldsymbol{y} = \boldsymbol{X}\boldsymbol{\beta} + \boldsymbol{u}, \quad \mathrm{E}(\boldsymbol{u}) = \boldsymbol{0}, \quad \mathrm{V}(\boldsymbol{u}) = \boldsymbol{\Omega},$$

«наилучшей» линейной несмещенной оценкой параметра $\boldsymbol{\beta}$ является оценка обобщенного метода наименьших квадратов $\widehat{\boldsymbol{\beta}}_{\mathrm{GLS}} = (\boldsymbol{X}'\boldsymbol{\Omega}^{-1}\boldsymbol{X})^{-1}\boldsymbol{X}'\boldsymbol{\Omega}^{-1}\boldsymbol{y}$, матрица ковариаций которой равна $(\boldsymbol{X}'\boldsymbol{\Omega}^{-1}\boldsymbol{X})^{-1}$. ковариационная матрица обычной оценки метода наименьших квадратов $\widehat{\boldsymbol{\beta}}_{\mathrm{OLS}} = (\boldsymbol{X}'\boldsymbol{X})^{-1}\boldsymbol{X}'\boldsymbol{y}$, равна $(\boldsymbol{X}'\boldsymbol{X})^{-1}\boldsymbol{X}'\boldsymbol{\Omega}\boldsymbol{X}(\boldsymbol{X}'\boldsymbol{X})^{-1}$. Отсюда вытекает:

*Лемма.* Для всякой положительно определенной матрицы $\boldsymbol{\Omega}$ и матрицы полного ранга $\boldsymbol{X}$ справедливо следующее неравенство:

$$(\boldsymbol{X}'\boldsymbol{X})^{-1}\boldsymbol{X}'\boldsymbol{\Omega}\boldsymbol{X}(\boldsymbol{X}'\boldsymbol{X})^{-1} \geqslant (\boldsymbol{X}'\boldsymbol{\Omega}^{-1}\boldsymbol{X})^{-1}.$$

Применим теперь лемму к матрицам

$$\boldsymbol{X} = \boldsymbol{W}^{1/2}\boldsymbol{X}_2 \quad \text{и} \quad \boldsymbol{\Omega} = \boldsymbol{W}^{1/2}(\sigma^2\boldsymbol{I} + \boldsymbol{X}_1\boldsymbol{V}_1\boldsymbol{X}_1')\boldsymbol{W}^{1/2}.$$

Получаем следующее неравенство:

$$(\boldsymbol{X}_2'\boldsymbol{W}\boldsymbol{X}_2)^{-1}\boldsymbol{X}_2'\boldsymbol{W}(\sigma^2\boldsymbol{I} + \boldsymbol{X}_1\boldsymbol{V}_1\boldsymbol{X}_1')\boldsymbol{W}\boldsymbol{X}_2(\boldsymbol{X}_2'\boldsymbol{W}\boldsymbol{X}_2)^{-1} \geqslant (\boldsymbol{X}_2'(\sigma^2\boldsymbol{I} + \boldsymbol{X}_1\boldsymbol{V}_1\boldsymbol{X}_1')^{-1}\boldsymbol{X}_2)^{-1},$$

справедливое для любой симметричной неотрицательно определенной матрицы $\boldsymbol{W}$, причем равенство достигается при

$$\boldsymbol{W} = (\sigma^2\boldsymbol{I} + \boldsymbol{X}_1\boldsymbol{V}_1\boldsymbol{X}_1')^{-1}.$$

Таким образом, оптимальной оценкой параметра $\boldsymbol{\beta}$ является оценка

$$\widehat{\boldsymbol{\beta}}_2 = (\boldsymbol{X}_2'(\sigma^2\boldsymbol{I} + \boldsymbol{X}_1\boldsymbol{V}_1\boldsymbol{X}_1')^{-1}\boldsymbol{X}_2)^{-1}\boldsymbol{X}_2'(\sigma^2\boldsymbol{I} + \boldsymbol{X}_1\boldsymbol{V}_1\boldsymbol{X}_1')^{-1}(\boldsymbol{y} - \boldsymbol{X}_1\widehat{\boldsymbol{\beta}}_1).$$

Эта оценка очевидно является более эффективной оценкой вектора $\boldsymbol{\beta}_2$, чем оценки, полученные в п. а) и б).

г) Из результата п. в) следует, что дисперсия «оптимальной» оценки $\widehat{\beta}_2$, полученной с использованием оценки $\widehat{\beta}_1$, независимой с $\varepsilon$, всегда меньше дисперсии оценки метода наименьших квадратов $\widehat{\beta}_{2(\text{OLS})}$. Таким образом, если в задаче присутствует мультиколлинеарность в $\boldsymbol{X}_1$, то дисперсия МНК-оценки $\mathrm{V}(\widehat{\beta}_{2(\text{OLS})}) = \sigma^2(\boldsymbol{X}_2'\boldsymbol{M}_1\boldsymbol{X}_2)^{-1}$ может быть очень велика, однако, используя «внешнюю» оценку $\widehat{\beta}_1$, независимую с ошибками $\varepsilon$, можно получить более точную оценку $\widehat{\beta}_2$. (В качестве предельного случая рассмотрите точную мультиколлинеарность в $\boldsymbol{X}_1$.)

## Задача 4.23

Для проверки гипотезы о том, что удельный выпуск $Q/L$ в металлургической промышленности зависит от уровня зарплаты $W$, на основе межстрановых наблюдений была получена регрессия

$$\ln\frac{Q}{L} = 0.374 + \underset{(0.049)}{0.805}W + e, \qquad R^2 = 0.929$$

(в скобках указана стандартная ошибка).

а) Проверьте гипотезу.

б) Было высказано предположение, что приведенное выше уравнение содержит ошибки спецификации, поскольку оно не учитывает разницу в эффективности между странами, которая оказывает влияние на удельный выпуск и положительно коррелирована с зарплатой. Как это предположение повлияет на ваш вывод?

## Решение

а) Проверка указанного предположения в данном случае сводится к тестированию гипотезы $\mathrm{H}_0$: $\beta_2 = 0$ (используются обозначения гл. 3). Нам не дано число наблюдений $n$, однако известно (см. разд. 3.5 гл. 3), что квадрат $t$-статистики для коэффициента $\beta_2$ должен в данном случае совпадать со статистикой (3.35) при $k = 2$:

$$t_2^2 = \left(\frac{\widehat{\beta}_2}{\widehat{\sigma}_{\widehat{\beta}_2}}\right)^2 = \frac{R^2}{1-R^2}(n-2).$$

Разрешая это соотношение относительно $n$, получаем:

$$n = 2 + \frac{t^2(1-R^2)}{R^2}.$$

Подставляя в эту формулу $t^2 = (0.805/0.049)^2 = 16.429^2 = 269.900$, $R^2 = 0.929$, получаем $n = 23$. Сравнивая $t = 16.429$ с $t_{0.025}(21) = 2.080$, приходим к выводу, что гипотеза $H_0$ уверенно отвергается. Иными словами, в рамках рассматриваемой модели удельный выпуск зависит от зарплаты.

б) Включение в модель новой объясняющей переменной (в данном случае — какой-либо количественной характеристики, описывающей уровень эффективности), положительно коррелированной как с зарплатой, так и с удельным выпуском, повлечет уменьшение оценки коэффициента при $W$, что может привести к его незначимости.

## Задача 4.24

Модель

$$y_t = \beta_1 + \beta_2 d_{t2} + \beta_3 d_{t3} + \beta_4 d_{t4} + \beta_5 x_t + \varepsilon_t$$

оценивается с помощью обычного метода наименьших квадратов на основе ежеквартальных наблюдений, где $d_{ti}$, $i = 2, 3, 4$ — фиктивные переменные для соответствующих кварталов, т. е.

$$\begin{aligned} d_{t2} &= 1, \quad \text{если } t \text{ — второй квартал,} && = 0 \text{ в остальных случаях,} \\ d_{t3} &= 1, \quad \text{если } t \text{ — третий квартал,} && = 0 \text{ в остальных случаях,} \\ d_{t4} &= 1, \quad \text{если } t \text{ — четвертый квартал,} && = 0 \text{ в остальных случаях.} \end{aligned}$$

а) Почему в модель не включена переменная $d_{t1}$?

б) Покажите, что оценка $\widehat{\beta}_5$ совпадает с МНК-оценкой коэффициента $\beta$ в регрессии $y_t^* = \alpha + \beta x_t^* + u_t$, где $y_t^*$ — остатки регрессии $y_t$ на $d_{t2}, d_{t3}, d_{t4}$ и константу, а $x_t^*$ — остатки регрессии $x_t$ на $d_{t2}, d_{t3}, d_{t4}$ и константу.

## Решение

а) Включение в модель переменной приводит к точной мультиколлинеарности, так как при каждом $t$ выполнено равенство

$$d_{t1} + d_{t2} + d_{t3} + d_{t4} = 1.$$

б) Обозначим через $\boldsymbol{Z} = (\imath\, d_2\, d_3\, d_4)$ матрицу, составленную из первых четырех регрессоров, а через $\boldsymbol{x}$, $\boldsymbol{y}$ — векторы $(x_1, \ldots, x_n)'$ и $(y_1, \ldots, y_n)'$. Тогда выполняются следующие равенства:

$$\begin{aligned} \boldsymbol{y} &= \boldsymbol{Z}\widehat{\beta} + \widehat{\beta}_5 \boldsymbol{x} + \boldsymbol{e}, \\ \boldsymbol{y} &= \boldsymbol{Z}\widehat{\gamma} + \boldsymbol{y}^*, \\ \boldsymbol{x} &= \boldsymbol{Z}\widehat{\delta} + \boldsymbol{x}^*, \end{aligned}$$

где $\widehat{\beta}$, $\widehat{\gamma}$, $\widehat{\delta}$ — векторы МНК-оценок соответствующих коэффициентов, а $\boldsymbol{e}$, $\boldsymbol{y}^*$, $\boldsymbol{x}^*$ — векторы остатков регрессий. Подставив выражения для $\boldsymbol{y}^*$, $\boldsymbol{x}^*$ из второго и третьего равенств в первое, получаем:

$$\boldsymbol{y}^* - \widehat{\beta}_5 \boldsymbol{x}^* - \boldsymbol{e} = \boldsymbol{Z}(\widehat{\beta} + \widehat{\beta}_5 \widehat{\delta} - \widehat{\gamma}).$$

Заметим, что по свойству МНК-оценок векторы $\boldsymbol{e}$, $\boldsymbol{y}^*$, $\boldsymbol{x}^*$ ортогональны столбцам матрицы $\boldsymbol{Z}$. Значит, в левой и правой частях последнего равенства стоят ортогональные векторы, а следовательно, они оба равны нулю. Получаем равенство:

$$\boldsymbol{y}^* = \widehat{\beta}_5\boldsymbol{x}^* + \boldsymbol{e}.$$

Поскольку вектор $\boldsymbol{e}$ ортогонален вектору $\boldsymbol{x}$ и столбцам матрицы $\boldsymbol{Z}$, то он ортогонален вектору $\boldsymbol{x}^* = \boldsymbol{x} - \boldsymbol{Z}\widehat{\boldsymbol{\delta}}$, являющемуся линейной комбинацией $\boldsymbol{x}$ и столбцов матрицы $\boldsymbol{Z}$. Это означает, что $\widehat{\beta}_5$ есть МНК-оценка коэффициента регрессии $\boldsymbol{y}^*$ на $\boldsymbol{x}^*$, а так как остатки $\boldsymbol{x}^*$ и $\boldsymbol{y}^*$ ортогональны столбцам матрицы $\boldsymbol{Z}$, первый из которых является константой, то $\sum x^* = \sum y^* = 0$ и регрессия $\boldsymbol{y}^*$ на $\boldsymbol{x}^*$ совпадает с регрессией $\boldsymbol{y}^*$ на $\boldsymbol{x}^*$ и константу.

## Задача 4.25

В программе исследований $k$ разных удобрений, предназначенных для повышения урожайности лимонных бананов, использованы в опытах на $n = n_1 + \dots + n_k$ опытных участках. Удобрение номер $s$ $(s = 1, \dots, k)$ использовалось на $n_s$ опытных участках. Для изучения влияния удобрений использовалась регрессионная модель

$$y_t = \beta_1 d_{1t} + \dots + \beta_k d_{kt}, \quad t = 1, \dots, n.$$

Здесь $\boldsymbol{y}$ — урожайность, $\boldsymbol{d}_s$ — фиктивная переменная, равная 1 для участка номер $s$ и 0 в других случаях. Известны выборочные средние $\overline{y}_s$ и стандартные отклонения $s_s$ для $s = 1, \dots, k$.

$$\overline{y}_s = \frac{1}{n_s}\sum_{d_{st}=1} y_t, \quad s_s^2 = \frac{1}{n_s - 1}\sum_{d_{st}=1}(y_t - \overline{y}_s)^2.$$

Выразить через известные величины $F$-статистику для тестирования нулевой гипотезы о равном влиянии всех удобрений $(\beta_1 = \dots = \beta_k)$.

## Решение

Для проверки гипотезы $\mathrm{H}_0\colon \beta_1 = \dots = \beta_k$ воспользуемся обычной $F$-статистикой

$$F = \frac{(\mathrm{ESS_R} - \mathrm{ESS_{UR}})/k}{\mathrm{ESS_{UR}}/(n-k)}.$$

Ясно, что в регрессии с ограничением $\beta_1 = \dots = \beta_k = \beta$ МНК-оценка параметра $\beta$ и сумма квадратов остатков $\mathrm{ESS_R}$ есть, соответственно,

$$\widehat{\beta} = \overline{\overline{y}} = \frac{1}{n}\sum_{t=1}^{n} y_t = \frac{1}{n}\sum_{s=1}^{k} n_s\overline{y}_s, \quad \mathrm{ESS_R} = \sum_{t=1}^{n}(y_t - \overline{\overline{y}})^2.$$

Покажем, что в регрессии без ограничения МНК-оценками параметров $\beta_s$ являются $\widehat{\beta}_s = \overline{y}_s$, $s = 1, \ldots, k$ (что на содержательном уровне почти очевидно). Упорядочим наблюдения таким образом, чтобы вначале шли $n_1$ наблюдений, относящихся к первому удобрению, затем $n_2$ наблюдений, относящихся ко второму удобрению, и т.д. Тогда матрица $\boldsymbol{X}$ регрессии без ограничения имеет следующий вид:

$$\boldsymbol{X} = \begin{bmatrix} \imath_{n_1} & \mathbf{0} & \cdots & \mathbf{0} \\ \mathbf{0} & \imath_{n_2} & \cdots & \mathbf{0} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{0} & \mathbf{0} & \cdots & \imath_{n_k} \end{bmatrix},$$

где $\imath_{n_s}$ — вектор из единиц размерности $n_s$. Очевидно, что

$$(\boldsymbol{X}'\boldsymbol{X})^{-1} = \begin{bmatrix} 1/n_1 & 0 & \cdots & 0 \\ 0 & 1/n_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 1/n_k \end{bmatrix}, \quad \boldsymbol{X}'y = \begin{bmatrix} \sum_{d_{1t}=1} y_t \\ \sum_{d_{2t}=1} y_t \\ \vdots \\ \sum_{d_{kt}=1} y_t \end{bmatrix},$$

откуда и вытекает требуемое утверждение. Отсюда непосредственно следует, что

$$\mathrm{ESS}_{\mathrm{UR}} = \sum_{s=1}^{k} \sum_{d_{st}=1} (y_t - \overline{y}_s)^2 = \sum_{s=1}^{k} (n_s - 1) s_s^2.$$

Таким образом,

$$\mathrm{ESS}_{\mathrm{R}} - \mathrm{ESS}_{\mathrm{UR}} = \sum_{t=1}^{n} (y_t - \overline{\overline{y}})^2 - \sum_{s=1}^{k} \sum_{d_{st}=1} (y_t - \overline{y}_s)^2$$

$$= \sum_{t=1}^{n} y_t^2 - n\overline{\overline{y}}^2 - \sum_{t=1}^{n} y_t^2 + \sum_{s=1}^{k} n_s \overline{y}_s^2 = \sum_{s=1}^{k} n_s (\overline{y}_s - \overline{\overline{y}})^2.$$

Окончательно получаем:

$$F = \frac{\sum_{s=1}^{k} n_s (\overline{y}_s - \overline{\overline{y}})^2}{\sum_{s=1}^{k} (n_s - 1) s_s^2} \cdot \frac{n-k}{k}.$$

## Задача 4.26

Вы в скором времени планируете поступить на должность политическог[о] аналитика на некую телевизионную станцию. Все телевизионные компани[и] уделяют большое внимание освещению выборов в Конгресс (США), а ва[-] ших знаний по этому вопросу недостаточно, чтобы получить эту работ[у]

Поэтому вы решили рассмотреть несколько регрессионных моделей, чтобы подкрепить свое мнение относительно выборов в Конгресс 1996 г.

Вам понадобится три модели. Каждая из них пытается объяснить различия в проценте голосов, отданных Республиканской партии, среди всех 50 штатов. То есть все модели включают 50 наблюдений, каждое соответствует одному штату. У вас также есть четыре типа объясняющих переменных:

1) уровень безработицы для каждого штата;
2) региональные фиктивные переменные, показывающие, что штат находится на северо-востоке, юге, среднем западе или на западе;
3) фиктивная переменная, показывающая, что Альберт Гор (вице-президент, демократ) появлялся в этом штате, агитируя за кандидатов в Конгресс;
4) перекрестные произведения региональных фиктивных переменных и фиктивной переменной Гора.

Три ваших модели отличаются только набором объясняющих переменных:

модель I содержит переменные 1) и 2);
модель II содержит переменные 1), 2) и 3);
модель III содержит переменные 1) и 4).

а) Запишите уравнение регрессии для каждой из моделей. Это можно сделать разными способами, используйте формулировку, которая вам больше нравится.

б) Укажите, как бы вы тестировали с помощью этих моделей следующие гипотезы (если вы хотите предложить $F$-тест, укажите регрессию с ограничениями и без ограничений):

1) появление Гора не оказывает влияния на процент голосов, отданных республиканцам;
2) вся страна голосует одинаково, без различий по региональному признаку;
3) северо-восток и средний запад («пояс холода») голосуют одинаково;
4) «пояс холода» голосует одинаково, «солнечный пояс» (юг и запад) голосует одинаково, но между этими поясами может быть разница;
5) появление Гора приводит к одному и тому же эффекту для всех регионов.

## Решение

Обозначим *unemp* уровень безработицы в штате, фиктивные переменные для северо-востока, юга, среднего запада и запада обозначим, соответственно, $NE$, $S$, $MW$, $W$. Фиктивную переменную, соответствующую появлению

Альберта Гора, обозначим *Gore*, и, наконец, процент голосов, набранных в штате членами Республиканской партии, будем обозначать *resp*.

а) Рассмотрим простейшие линейные модели.

Модель I:

$$resp = c + \alpha_u unemp + \alpha_{NE} NE + \alpha_S S + \alpha_{MW} MW + \alpha_W W.$$

Модель II:

$$resp = c + \beta_u unemp + \beta_{NE} NE + \beta_S S + \beta_{MW} MW + \beta_W W + \beta_G Gore.$$

Модель III:

$$resp = c + \gamma_u unemp + \gamma_{NEG} NE \cdot Gore + \gamma_{SG} S \cdot Gore + \gamma_{MWG} MW \cdot Gore + \gamma_{WG} W \cdot Gore.$$

б) 1) Гипотеза: появление Гора не оказывает влияния на процент голосов, отданных республиканцам.

Модель I. Не дает возможности тестировать гипотезу.

Модель II. $H_0: \beta_G = 0$.

Модель III. Не дает возможности тестировать гипотезу.

2) Гипотеза: вся страна голосует одинаково, без различий по региональному признаку.

Модель I. $H_0: \alpha_{NE} = \alpha_S = \alpha_{MW} = \alpha_W$.

Модель II. $H_0: \beta_{NE} = \beta_S = \beta_{MW} = \beta_W$.

Модель III. Не дает возможности тестировать гипотезу.

3) Гипотеза: северо-восток и средний запад («пояс холода») голосуют одинаково.

Модель I. $H_0: \alpha_{NE} = \alpha_{MW}$.

Модель II. $H_0: \beta_{NE} = \beta_{MW}$.

Модель III. Не дает возможности тестировать гипотезу.

4) Гипотеза: «пояс холода» голосует одинаково, «солнечный пояс» (юг и запад) голосует одинаково, но между этими поясами может быть разница.

Модель I. $H_0: \alpha_{NE} = \alpha_{MW}, \quad \alpha_S = \alpha_W$.

Модель II. $H_0: \beta_{NE} = \beta_{MW}, \quad \beta_S = \beta_W$.

Модель III. Не дает возможности тестировать гипотезу.

5) Гипотеза: появление Гора приводит к одному и тому же эффекту для всех регионов.

Модель I. Не дает возможности тестировать гипотезу.

Модель II. Не дает возможности тестировать гипотезу.

Модель III. $H_0: \gamma_{NEG} = \gamma_{SG} = \gamma_{MWG} = \gamma_{WG}$.

Оставляем читателю в виде простой самостоятельной работы построить регрессии с ограничением и без ограничения для построения соответствующих $F$-тестов.

*Замечание.* Относительно самого последнего предложения тестировать гипотезу об одинаковом для всех регионов эффекте появления Гора с помощью модели III могут возникнуть вполне резонные возражения: до появления Гора могли существовать региональные различия в отношении к республиканцам, однако появление Гора привело к тому, что эти различия исчезли и в результате нулевая гипотеза не отвергается, хотя реально влияние Гора могло быть разным в зависимости от региона. Однако по условию задачи мы ограничены рамками модели III, а в ней по существу постулируется отсутствие региональных различий. Иными словами, модель III можно сделать более гибкой, например, включив в нее дополнительно региональные фиктивные переменные. Тогда гипотеза $\mathrm{H}_0\colon \gamma_{NEG} = \gamma_{SG} = \gamma_{MWG} = \gamma_{WG}$ будет более адекватно выражать идею об одинаковом влиянии в регионах появления вице-президента США.

## Задача 4.27[1]

В данной задаче исследуется вопрос о том, до какой степени $\widehat{\beta}$, МНК-оценка вектора коэффициентов в регрессии $(*)$ по $n$ наблюдениям,

$$\boldsymbol{y} = \boldsymbol{X}\beta + \varepsilon \qquad (*)$$

зависит от одного наблюдения. Как сильно изменится оценка, если наблюдение номер $t$ удалить из данных?

а) Пусть регрессоры $\boldsymbol{X}$ разбиты на 2 группы: $\boldsymbol{X} = \begin{bmatrix}\boldsymbol{X}_1 & \boldsymbol{X}_2\end{bmatrix}$. Обозначим через $\boldsymbol{P}_1$, $\boldsymbol{P}_2$ и $\boldsymbol{P}$ матрицы операторов ортогонального проектирования на подпространства, порожденные регрессорами $\boldsymbol{X}_1$, $\boldsymbol{X}_2$ и $\boldsymbol{X}$, соответственно. Через $\boldsymbol{M}_1$, $\boldsymbol{M}_2$ и $\boldsymbol{M}$ обозначим матрицы операторов ортогонального проектирования на ортогональные дополнения к подпространствам, порожденным регрессорами $\boldsymbol{X}_1$, $\boldsymbol{X}_2$ и $\boldsymbol{X}$, соответственно.

Докажите следующую теорему. В регрессиях

$$\boldsymbol{y} = \boldsymbol{X}_1\beta_1 + \boldsymbol{X}_2\beta_2 + \boldsymbol{u} \quad \text{и} \quad \boldsymbol{M}_1\boldsymbol{y} = \boldsymbol{M}_1\boldsymbol{X}_2\beta_2 + \boldsymbol{v}$$

совпадают 1) МНК-оценки вектора $\beta_2$ и 2) остатки регрессий.

б) Обозначим через $\boldsymbol{u}_t$ вектор размера $n\times 1$, у которого компонента номер $t$ равна 1, а все остальные компоненты равны 0. Пусть $\boldsymbol{P}_{\boldsymbol{X}}$, $\boldsymbol{P}_t$ и $\boldsymbol{P}$ — матрицы операторов ортогонального проектирования на подпространства, порожденные регрессорами $\boldsymbol{X}$, $\boldsymbol{u}_t$ и $\{\boldsymbol{X}, \boldsymbol{u}_t\}$, соответственно, а матрицы $\boldsymbol{M}_{\boldsymbol{X}}$, $\boldsymbol{M}_t$ и $\boldsymbol{M}$ есть операторы проектирования на соответствующие ортогональные дополнения.

Покажите, что оценка $\widehat{\beta}^{(t)}$ вектора $\beta$ в уравнении $(*)$, полученная по $n-1$ наблюдениям (наблюдение номер $t$ удалено из исходных данных),

[1] В 6-м издании учебника эта задача совпадает с задачей 4.6 и заменена в 7-м издании на новую задачу.

совпадает с оценкой вектора $\beta$ в уравнении (**), полученной по всему набору из $n$ наблюдений:

$$y = X\beta + \alpha u_t + \eta. \qquad (**)$$

в) Покажите, что

$$X(\widehat{\beta}^{(t)} - \widehat{\beta}) = -\widehat{\alpha} P_X u_t.$$

г) Покажите, что

$$\widehat{\beta}^{(t)} - \widehat{\beta} = -\frac{1}{1-h_t}(X'X)^{-1} x_t e_t,$$

где $x_t'$ — $t$-я строка матрицы $X$, $e_t$ — $t$-я компонента вектора остатков регрессии (*), а $h_t$ — $t$-й диагональный элемент матрицы $P_X$. Обсудите полученную формулу.

д) Покажите, что 1) значения $h_t$ лежат в интервале $(0, 1)$; 2) среднее значение $h_t$ равно $k/n$.

е) Для уравнения

$$\ln w_t = \beta_1 + \beta_2 sex_t + \beta_3 age_t + \beta_4 edu_t + \varepsilon_t$$

постройте диаграмму рассеивания $(age_t, h_t)$ (данные содержатся в файле wages.xls).

## Решение

а) Пусть $\widehat{\beta}_1$ и $\widehat{\beta}_2$ — МНК-оценки коэффициентов уравнения

$$y = X_1\beta_1 + X_2\beta_2 + u,$$

а $\widehat{u}$ — вектор остатков. Имеем

$$y = X_1\widehat{\beta}_1 + X_2\widehat{\beta}_2 + \widehat{u}.$$

Поскольку $P + X = I$, $PX_1 = X_1$, $PX_2 = X_2$ и $P\widehat{u} = 0$, получаем:

$$y = Py + My = X_1\widehat{\beta}_1 + X_2\widehat{\beta}_2 + My.$$

Умножив слева на $M_1$, учитывая, что $M_1X_1 = 0$ и $M_1M = M$, получаем:

$$M_1 y = M_1 X_2 \widehat{\beta}_2 + My. \qquad (1)$$

Далее, после умножения слева на $X_2'$, учитывая, что вектор $My$ ортогонален регрессорам $X_2$, имеем

$$X_2' M_1 y = X_2' M_1 X_2 \widehat{\beta}_2.$$

Из последнего равенства получаем:

$$\widehat{\beta}_2 = (X_2' M_1 X_2)^{-1} X_2' M_1 y,$$

что совпадает со стандартной формулой для МНК-оценок регрессии

$$M_1 y = M_1 X_2 \beta_2 + v.$$

Из (1) получаем, что остатки $My$ в регрессии (*) совпадают с остатками в регрессии (**). Теорема доказана.

б) Умножим уравнение (**) на $M_t$. По теореме из а) полученное уравнение

$$M_t y = M_t X\beta + w \tag{2}$$

имеет МНК-оценку вектора коэффициентов, совпадающую с оценкой, получаемой из (**). Поскольку $M_t$ — ортогональная проекция вдоль вектора $u_t$, то

$$M_t = I - u_t(u_t' u_t)^{-1} u_t' = I - u_t u_t'$$

и

$$M_t y = y - u_t u_t' y = y - y_t u_t,$$

т.е. вектор $M_t y$ имеет те же компоненты, что и вектор $y$, за исключением $t$-й компоненты, замененной на 0. Аналогично, матрица $M_t X$ совпадает с матрицей $X$, за исключением $t$-й строки, замененной нулями.

Поскольку в формуле для МНК-оценки вектора коэффициентов присутствуют только скалярные произведения переменных, то МНК-оценка уравнения (2) по $n$ наблюдениям (совпадающая с оценкой из (**)) является оценкой, полученной из уравнения (*) по $n-1$ наблюдениям (когда $t$-е наблюдение удалено).

в) Оценивая уравнение (**) с помощью МНК, получаем

$$y = X\widehat{\beta}^{(t)} + \widehat{\alpha} u_t + My.$$

Поскольку $P_X M = 0$, то, умножая на $P_X$, получаем:

$$P_X y = X\widehat{\beta}^{(t)} + \widehat{\alpha} P_X u_t.$$

Подставляя $P_X y = X\widehat{\beta}$ в предыдущее уравнение, получаем:

$$X(\widehat{\beta}^{(t)} - \widehat{\beta}) = -\widehat{\alpha} P_X u_t. \tag{3}$$

г) Снова воспользуемся теоремой из а) для того, чтобы найти оценку $\widehat{\alpha}$. Умножив (**) на $M_X$, получим:

$$M_X y = \alpha M_X u_t + M_X \eta.$$

Отсюда, используя идемпотентность матрицы $M_X$, получаем:

$$\widehat{\alpha} = \frac{(M_X y)'(M_X u_t)}{(M_X u_t)'(M_X u_t)} = \frac{y' M_X u_t}{u_t' M_X u_t}.$$

В последнем выражении числитель равен остатку $e_t$ в $t$-м наблюдении в регрессии $(*)$, а знаменатель равен $t$-му диагональному элементу матрицы $M_X$, или $1-h_t$, где $h_t$ — $t$-й диагональный элемент матрицы $P_X = I - M_X$.

Умножим уравнение (3) слева на $(X'X)^{-1}X'$:

$$\widehat{\beta}^{(t)} - \widehat{\beta} = -\widehat{\alpha}(X'X)^{-1}X'P_X u_t.$$

Имеем $X'P_X = (P_X'X)' = (P_X X)' = X'$. Кроме того, вектор $X'u_t$ равен $x_t e_t$, так как умножение на $u_t$ выделяет столбец $x_t$ матрицы $X'$. Таким образом, получаем:

$$\widehat{\beta}^{(t)} - \widehat{\beta} = -\frac{1}{1-h_t}(X'X)^{-1}x_t e_t.$$

Итак, если одно или оба из значений $e_t$, $h_t$ велики, то и вклад наблюдения $t$ также велик. Такие наблюдения можно назвать «влиятельными». Остаток $e_t$ связан со значением $y_t$, а величина $h_t$ зависит только от значений независимых переменных. Наблюдения с большими значениями $h_t$ потенциально являются «влиятельными» и называются *leverage points*.

д) Величину $h_t$ можно представить в виде

$$h_t = u_t' P_X u_t = \|P_X u_t\|^2.$$

Отсюда сразу следует, что $h_t \geqslant 0$, а поскольку $P_X$ является проектором, то $\|P_X u_t\|^2 \leqslant \|v u_t\|^2 = 1$, поэтому $h_t \leqslant 1$.

Вычислим среднее значение $h_t$:

$$\sum_{t=1}^{n} h_t = \operatorname{tr}(P_X) = \operatorname{tr}(X(X'X)^{-1}X') = \operatorname{tr}((X'X)^{-1}X'X) = \operatorname{tr}(I_k) = k,$$

откуда

$$\frac{1}{n}\sum_{t=1}^{n} h_t = \frac{k}{n}.$$

е) Искомая диаграмма рассеивания представлена на рис. 4.3.

Мы видим, что потенциально наиболее «влиятельными» являются наблюдения с возрастом $age \approx 55 \div 60$.

## Задача 4.28

Ниже приведены результаты регрессии $W$ — зарплаты менеджера фирмы, на объем ее продаж $S$ и доход $P$ (число наблюдений $n = 102$, в скобках даны

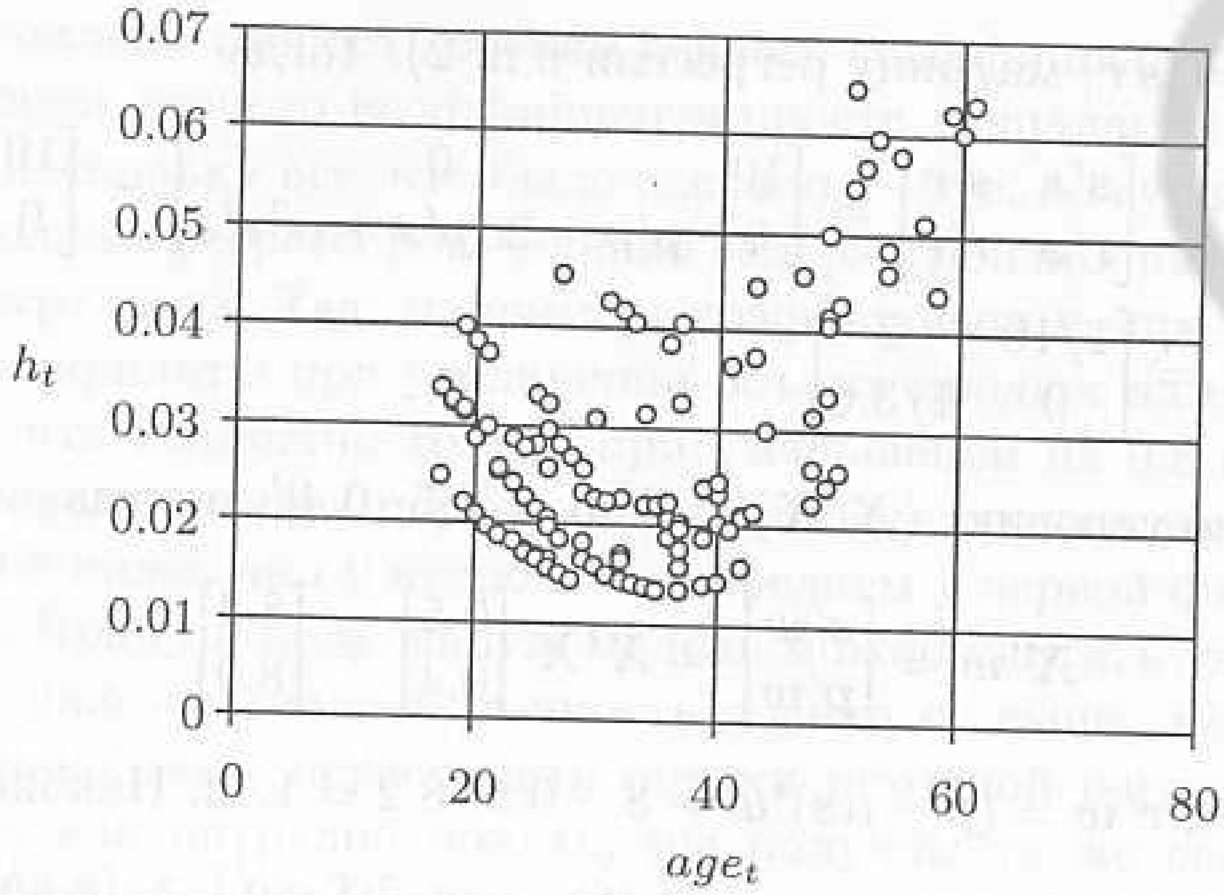



**Рис. 4.3**

стандартные ошибки):

$$W = \underset{(0.83)}{0.50}\,S + \underset{(0.83)}{0.40}\,P, \quad e'e = 250, \quad X'X = \begin{bmatrix} 10 & 8 \\ 8 & 10 \end{bmatrix}$$

(для удобства все переменные представлены в отклонениях от средних). Ввиду большой зависимости между объемами продаж и доходом возникает проблема мультиколлинеарности, что не позволяет точно оценить соответствующие параметры. Для решения этой проблемы было предложено действовать следующим образом.

1) Провести регрессию $P$ на $S$ и получить остатки $r$.

2) Провести регрессию $W$ на $S$ и $r$.

Обозначим результат последней регрессии $W = c_1 S + c_2 r$.

а) Вычислите $c_1$, $c_2$.

б) Дайте оценку предложенному методу как способу борьбы с мультиколлинеарностью.

в) Дайте оценку предложенному методу как способу получения более точных оценок исходных параметров.

## Решение

а) Приведем сначала чисто формальное решение этой части задачи. Будем считать, что $w, s, p,$ — это векторы-столбцы размерности $n = 102$, соответствующие переменным $W$, $S$, $P$, а $e$ и $r$ — векторы остатков. Тогда результатом шага 1) является вектор $r = p - \alpha s$, где

$$\alpha = \frac{p's}{s's} = \frac{8}{10} = 0.8.$$

Обозначим $Z = [s\,r]$ матрицу регрессии в п. 2). Тогда

$$Z'Z = \begin{bmatrix} s's & s'r \\ r's & r'r \end{bmatrix} = \begin{bmatrix} 10 & 0 \\ 0 & p'p - 2\alpha p's + \alpha^2 s's \end{bmatrix} = \begin{bmatrix} 10 & 0 \\ 0 & 3.6 \end{bmatrix},$$

$$(Z'Z)^{-1} = \begin{bmatrix} 1/10 & 0 \\ 0 & 1/3.6 \end{bmatrix}.$$

Далее, согласно условию, $(X'X)^{-1}X'w = [0.5,\ 0.4]'$, откуда

$$X'w = \begin{bmatrix} s'w \\ p'w \end{bmatrix} = X'X \begin{bmatrix} 0.5 \\ 0.4 \end{bmatrix} = \begin{bmatrix} 8.2 \\ 8.0 \end{bmatrix}.$$

Следовательно, $r'w = (p - \alpha s)'w = 8 - 0.8 \cdot 8.2 = 1.44$. Наконец,

$$\begin{bmatrix} c_1 \\ c_2 \end{bmatrix} = (Z'Z)^{-1}Z'w = \begin{bmatrix} 1/10 & 0 \\ 0 & 1/3.6 \end{bmatrix} \begin{bmatrix} 8.2 \\ 1.44 \end{bmatrix} = \begin{bmatrix} 0.82 \\ 0.4 \end{bmatrix},$$

т. е.

$$c_1 = 0.82, \quad c_2 = 0.4. \tag{$*$}$$

Обозначим вектор остатков последней регрессии через $u$. Поскольку в исходной регрессии и в регрессии п. 2) проекции вектора $w$ осуществляются на одно и то же двумерное подпространство, то векторы остатков этих регрессий совпадают: $e = u$. Поэтому $e'e = u'u = 250$, $s^2 = \widehat{\omega}^2 = 2.5$ и

$$s_{c_1} = \sqrt{\frac{2.5}{10}} = 0.5, \quad s_{c_2} = \sqrt{\frac{2.5}{3.6}} = 0.83, \tag{$**$}$$

где $s^2$, $\widehat{\omega}^2$ — оценки дисперсий исходной и последней регрессий, соответственно.

Менее формально этот же результат можно получить так. Обозначим $b_1 = 0.5$, $b_2 = 0.4$ — оценки коэффициентов исходной регрессии. Используя известные формулы, получаем:

$$s^2 = 2.5, \quad s_{b_1}^2 = s_{b_2}^2 = 0.694, \quad \widehat{\mathrm{Cov}}(b_1, b_2) = -0.556.$$

Имеем $w = b_1 s + b_2 p + e$, $p = \alpha s + r$, где $\alpha = 0.8$, и $w = c_1 s + c_2 r + u$. Так как остатки $e$ и $u$ совпадают и проекция вектора $w$ на двумерное подпространство, порожденное $s$ и $p$, единственна, то

$$b_1 s + b_2 p = c_1 s + c_2 r = (c_1 - \alpha c_2)s + c_2 p.$$

Следовательно,

$$c_1 = b_1 + \alpha b_2 = 0.82, \quad c_2 = b_2 = 0.4,$$
$$s_{c_1}^2 = s_{b_1}^2 + \alpha^2 s_{b_2}^2 + 2\alpha, \quad \widehat{\mathrm{Cov}}(b_1, b_2) = 0.25, \quad s_{c_2}^2 = s_{b_2}^2 = 0.694,$$

что совпадает с $(*)$ и $(**)$.

б) Формально точность оценки первого коэффициента в новой регрессии выше (оценки второго коэффициента просто совпадают). Этого следовало ожидать, поскольку все, что было сделано, — это, по существу, ортогонализация исходных регрессоров. Однако содержательная интерпретация новой модели затруднена. Так, например, коэффициент $c_1$ представляет среднее изменение зарплаты при увеличении объема продаж на единицу при неизменном $r$, что возможно только при уменьшении на 0.8 величины дохода. Иными словами, если у первой фирмы объем продаж на единицу выше, а доход на 0.8 ниже, чем у второй, то в среднем у первой фирмы зарплата на 0.82 выше. Вряд ли полученную модель можно считать удачной.

в) Как уже отмечалось, точность оценки $c_1$ выше, чем точность оценки $b_1$. Однако, если пересчитать оценки исходной регрессии через новые оценки, то, как нетрудно понять, мы получим те же самые $b_1, b_2$ с теми же самыми ошибками. В общем случае это выглядит следующим образом. Пусть есть невырожденное линейное преобразование регрессоров $\boldsymbol{A}$ ($k \times k$ матрица), и пусть $\boldsymbol{y} = \boldsymbol{X}\widehat{\boldsymbol{\beta}} + \boldsymbol{e}$ и $\boldsymbol{y} = \widetilde{\boldsymbol{X}}\widehat{\boldsymbol{\gamma}} + \widetilde{\boldsymbol{e}}$ — соответствующие регрессии, где $\widetilde{\boldsymbol{X}} = \boldsymbol{X}\boldsymbol{A}$. Если взять теперь оценку $\widetilde{\boldsymbol{\beta}} = \boldsymbol{A}\widehat{\boldsymbol{\gamma}}$, то нетрудно проверить, что $\widetilde{\boldsymbol{\beta}} = \widehat{\boldsymbol{\beta}}$. Действительно, в силу того, что матрица $\boldsymbol{A}$ невырожденная, $k$-мерные пространства, порожденные регрессорами $\boldsymbol{X}$ и $\widetilde{\boldsymbol{X}}$, совпадают, а потому совпадают проекции на них вектора $\boldsymbol{y}$: $\boldsymbol{X}\widehat{\boldsymbol{\beta}} = \widetilde{\boldsymbol{X}}\widehat{\boldsymbol{\gamma}}$ и векторы остатков: $\boldsymbol{e} = \widetilde{\boldsymbol{e}}$. Таким образом, $\boldsymbol{X}\widehat{\boldsymbol{\beta}} = \widetilde{\boldsymbol{X}}\widehat{\boldsymbol{\gamma}} = \boldsymbol{X}\boldsymbol{A}\widehat{\boldsymbol{\gamma}}$. Так как столбцы матрицы $\boldsymbol{X}$ линейно независимы, то $\widehat{\boldsymbol{\beta}} = \boldsymbol{A}\widehat{\boldsymbol{\gamma}}$, т. е. $\widehat{\boldsymbol{\beta}} = \widetilde{\boldsymbol{\beta}}$.

Таким образом, мы не достигаем никакого выигрыша в эффективности оценок исходных параметров.

## Задача 4.29

Файл `usa_import.xls` содержит данные об объеме импорта $imp$ (млрд. долл.), валовом национальном продукте $gdp$ (млрд. долл.) и индексе потребительских цен $cpi$ в США за период с 1964 по 1979 г.

а) Вычислите выборочный коэффициент корреляции между $gdp$ и $cpi$.

б) Оцените регрессию $imp$ на константу и $gdp$.

в) Оцените регрессию $imp$ на константу и $cpi$.

г) Оцените регрессию $imp$ на константу, $gdp$ и $cpi$.

Как можно интерпретировать полученные результаты? Можно ли ограничиться только одной из регрессий б) или в)?

## Решение

а) Выборочный коэффициент корреляции $gdp$ и $cpi$ равен $r(gdp, cpi) = 0.997$.

Ниже приведены результаты соответствующих регрессий.

б) См. таблицу 4.20.

в) См. таблицу 4.21.

Таблица 4.20

Dependent Variable: *imp*

| Variable | Coefficient | Std. Error | $t$-Statistic | Probability |
|---|---|---|---|---|
| const | −69.027 | 5.750 | −12.004 | 0.0000 |
| *gdp* | 0.134 | 0.004 | 31.867 | 0.0000 |
| $R^2$ | 0.9864 | | | |

Таблица 4.21

Dependent Variable: *imp*

| Variable | Coefficient | Std. Error | $t$-Statistic | Probability |
|---|---|---|---|---|
| const | −146.516 | 8.335 | −17.579 | 0.0000 |
| *cpi* | 1.824 | 0.059 | 30.795 | 0.0000 |
| $R^2$ | 0.9855 | | | |

Таблица 4.22

Dependent Variable: *imp*

| Variable | Coefficient | Std. Error | $t$-Statistic | Probability |
|---|---|---|---|---|
| const | −101.489 | 33.080 | −3.068 | 0.0090 |
| *gdp* | 0.079 | 0.056 | 1.403 | 0.1839 |
| *cpi* | 0.759 | 0.761 | 0.996 | 0.3372 |
| $R^2$ | 0.9874 | | | |

г) См. таблицу 4.22.

В данном случае мы имеем дело с типичным проявлением мультиколлинеарности: первые две (раздельные) регрессии высоко значимы, имеют большие значения коэффициента детерминации. В то же время совместная регрессия приводит к незначимым на любом разумном уровне коэффициентам при обеих объясняющих переменных. Этот результат следовало ожидать, поскольку выборочный коэффициент корреляции между *gdp* и *cpi* очень близок к единице. Совместная модель не позволяет достоверно оценить связь между объемом импорта и ВНП и индексом цен. Регрессии с одной объясняющей переменной дают некоторое представление о связи импорта с каждым из факторов, однако для прогнозных целей предпочтительнее модель г).

Имеет смысл обратить внимание на высокие значения выборочных коэффициентов корреляции всех переменных со временем: $r(gdp, year) = 0.967$, $r(cpi, year) = 0.963$, $r(imp, year) = 0.937$. Возникает естественный вопрос, сохранится ли значимая связь объема импорта с каждым из объясняющих факторов при учете влияния фактора времени (см. Пример рынка валютных фьючерсов, разд. 4.3 гл. 4). В данном случае частные коэффициенты корреляции между переменными при исключении времени остаются достаточно высокими: $r(imp, gdp \mid year) = 0.977$, $r(imp, cpi \mid year) = 0.958$. Значимыми остаются первые две регрессии при включении в правую часть уравнения времени (см. таблицы 4.23 и 4.24).

**Таблица 4.23**

Dependent Variable: *imp*

| Variable | Coefficient | Std. Error | $t$-Statistic | Probability |
|---|---|---|---|---|
| const | 10458.810 | 2373.086 | 4.404 | 0.0007 |
| *gdp* | 0.181 | 0.011 | 16.687 | 0.0000 |
| *year* | −5.370 | 1.210 | −4.436 | 0.0007 |
| $R^2$ | 0.9874 | | | |

**Таблица 4.24**

Dependent Variable: *imp*

| Variable | Coefficient | Std. Error | $t$-Statistic | Probability |
|---|---|---|---|---|
| const | 7325.813 | 3017.964 | 2.427 | 0.0305 |
| *cpi* | 2.271 | 0.188 | 12.102 | 0.0000 |
| *year* | −3.821 | 1.543 | −2.476 | 0.0278 |
| $R^2$ | 0.9874 | | | |

Эти результаты дают основание утверждать, что существует значимая связь между объемом импорта и ВНП и индексом цен.

*Замечание.* Забегая немного вперед, следует сказать, что к полученным здесь результатам и выводам следует относиться весьма осторожно. В нашем случае все переменные являются нестационарными временными рядами, поэтому высокая значимость рассмотренных моделей может оказаться проявлением так называемой мнимой регрессии (подробнее см. гл. 11). В пользу наличия зависимости объема импорта от ВВП и индекса цен говорит тот факт, что регрессии первой разности $\Delta imp$ на первые разности $\Delta gdp$ и $\Delta cpi$ являются высоко значимыми.

## Задача 4.30

Построение модели цены колготок в московских оптовых торговых фирмах, осень 1997 г. Данные содержатся в файле `tights.xls` (всего 74 наблюдения). Описание переменных содержится в таблице 4.25.

а) Постройте уравнения зависимости цены колготок от их плотности, состава и производителя. Подберите наиболее подходящую форму модели. Какие проблемы с данными вы при этом встретили?

Таблица 4.25

| Переменная | Описание |
|---|---|
| $n$ | номер по порядку |
| $price$ | цена колготок (руб.) |
| $den$ | плотность (DEN) |
| $polyamid$ | содержание полиамида (%) |
| $lykra$ | содержание лайкры (%) |
| $cotton$ | содержание хлопка (%) |
| $wool$ | содержание шерсти (%) |
| $firm$ | фирма-производитель: 0 — Levante, 1 — Golden Lady |

б) С помощью построенной модели ответьте на вопросы: верно ли, что цены колготок двух фирм-производителей различаются статистически достоверно? Какая из фирм устанавливает более высокие цены?

## Решение

а) Если оценить линейную модель, включив в правую часть константу и все переменные, содержащиеся в списке данных, то большинство коэффициентов будут незначимыми. Показательным является тот факт, что метод наименьших квадратов выдает какие-то оценки. Это означает, что регрессоры линейно независимы, хотя ожидается, что сумма процентов лайкры, полиамида, шерсти и хлопка для каждого наблюдения должна быть равна 100. Создав переменную $sum = lykra + polyamid + cotton + wool$, убеждаемся, что для двух наблюдений ее значение меньше чем 100. Эти наблюдения целесообразно удалить, поскольку либо они содержат еще какую-то компоненту, либо они просто ошибочны. Просматривая далее данные, видим, что только для двух наблюдений значение переменной $wool$ отличается от нуля. Трудно ожидать, что в этом случае можно оценить влияние шерсти на цену колготок. Эти наблюдения также целесообразно удалить. Наконец, замечаем, что для большинства наблюдений значение переменной $cotton$ не превышает 3%, и лишь для трех наблюдений ее значение не меньше 50%, что свидетельствует о том, что эти колготки существенно отличаются от всех остальных, и их также целесообразно удалить из списка наблюдений. Итак, рассматриваем те наблюдения, для которых выполнены следующие условия: $sum = 100$, $wool = 0$, $cotton < 50$. Ясно, что из трех оставшихся компонент $(polyamid, cotton, lykra)$ в правую часть линейной модели нужно включать только две (мы, как обычно, включаем в регрессию константу). Ниже приведены результаты оценивания линейной модели (см. таблицу 4.26).

Все переменные, за исключением $firm$, значимы на 5%-ном уровне. Переменная $firm$ значима на 7%-ном уровне.

**Таблица 4.26**

Dependent Variable: *price*

| Variable | Coefficient | Std. Error | $t$-Statistic | Probability |
|---|---|---|---|---|
| const | 2728.327 | 2112.034 | 1.292 | 0.2012 |
| *cotton* | 2203.028 | 521.882 | 4.221 | 0.0001 |
| *lykra* | 329.3805 | 102.727 | 3.206 | 0.0021 |
| *den* | 166.2228 | 35.769 | 4.647 | 0.0000 |
| *firm* | 2566.686 | 1373.875 | 1.868 | 0.0665 |
| $R^2$ | 0.5007 | | | |

Приведем результаты оценивания полулогарифмической модели, у которой зависимой переменной является $\ln price$, а объясняющие переменные такие же, как и в линейной модели (см. таблицу 4.27).

**Таблица 4.27**

Dependent Variable: $\ln price$

| Variable | Coefficient | Std. Error | $t$-Statistic | Probability |
|---|---|---|---|---|
| const | 8.446 | 0.132 | 64.171 | 0.0000 |
| *cotton* | 0.149 | 0.033 | 4.581 | 0.0000 |
| *lykra* | 0.0353 | 0.006 | 5.511 | 0.0000 |
| *den* | 0.012 | 0.002 | 5.197 | 0.0000 |
| *firm* | 0.287 | 0.086 | 3.348 | 0.0014 |
| $R^2$ | 0.6086 | | | |

В этой регрессии все переменные значимы на 5%-ном уровне.

б) Согласно линейной модели колготки фирмы Golden Lady в среднем на 2567 руб. дороже колготок фирмы Levante, имеющих тот же состав и плотность. Согласно полулогарифмической модели колготки фирмы Golden Lady в среднем на 33.1% дороже колготок фирмы Levante. В линейной модели различие значимо на 7%-ном уровне, в полулогарифмической модели различие значимо на любом разумном уровне.

# Глава 5

# Некоторые обобщения множественной регрессии

## Задача 5.1

Проверьте несмещенность оценки (5.4):

$$\widehat{\beta}_{\text{GLS}} = (X'\Omega^{-1}X)^{-1}X'\Omega^{-1}y.$$

## Решение

Чтобы проверить несмещенность оценки $\widehat{\beta}_{\text{GLS}}$, вычислим ее математическое ожидание:

$$\begin{aligned}\mathrm{E}(\widehat{\beta}_{\text{GLS}}) &= \mathrm{E}((X'\Omega^{-1}X)^{-1}X'\Omega^{-1}y) = (X'\Omega^{-1}X)^{-1}X'\Omega^{-1}\mathrm{E}(y)\\ &= (X'\Omega^{-1}X)^{-1}X'\Omega^{-1}X\beta = \beta.\end{aligned}$$

## Задача 5.2

Проверьте равенство (5.8):

$$\mathrm{V}(\widehat{\beta}_{\text{GLS}}) = (X'\Omega^{-1}X)^{-1}.$$

## Решение

Обозначим $A = (X'\Omega^{-1}X)^{-1}X'\Omega^{-1}$. Тогда $\widehat{\beta}_{\text{GLS}} = Ay$ и, значит,

$$\begin{aligned}\mathrm{V}(\widehat{\beta}_{\text{GLS}}) &= \mathrm{V}(Ay) = A\mathrm{V}(y)A' = A\Omega A'\\ &= [(X'\Omega^{-1}X)^{-1}X'\Omega^{-1}]\Omega[\Omega^{-1}X(X'\Omega^{-1}X)^{-1}]\\ &= (X'\Omega^{-1}X)^{-1}X'(\Omega^{-1}\Omega)\Omega^{-1}X(X'\Omega^{-1}X)^{-1}\\ &= (X'\Omega^{-1}X)^{-1}(X'\Omega^{-1}X)(X'\Omega^{-1}X)^{-1} = (X'\Omega^{-1}X)^{-1}.\end{aligned}$$

## Задача 5.3

Докажите, что $\operatorname{Cov}(\widehat{\beta}_{\text{OLS}}, \widehat{\beta}_{\text{GLS}}) = \operatorname{V}(\widehat{\beta}_{\text{GLS}})$.

## Решение

Напомним, что $\widehat{\beta}_{\text{GLS}} = (X'\Omega^{-1}X)^{-1}X'\Omega^{-1}y$. Обозначим

$$A = (X'\Omega^{-1}X)^{-1}X'\Omega^{-1},$$
$$B = (X'X)^{-1}X'.$$

Далее,

$$\begin{aligned}\operatorname{Cov}(\widehat{\beta}_{\text{OLS}}, \widehat{\beta}_{\text{GLS}}) &= \operatorname{Cov}(By, Ay) = B\operatorname{V}(y)A' \\ &= (X'X)^{-1}X'\Omega\Omega^{-1}X(X'\Omega^{-1}X)^{-1} \\ &= (X'\Omega^{-1}X)^{-1} = \operatorname{V}(\widehat{\beta}_{\text{GLS}}).\end{aligned}$$

Равенство доказано.

## Задача 5.4

Согласно результатам п. 3.2 для классической регрессионной модели

$$\operatorname{Cov}(\widehat{y}_t, e_t) = 0, t = 1, \ldots, n,$$

где $\widehat{y} = (\widehat{y}_1, \ldots, \widehat{y}_n)' = X\widehat{\beta}_{\text{OLS}}$ — прогнозное значение $y$, $e = (e_1, \ldots, e_n)' = y - \widehat{y}$ — вектор остатков. Сохраняется ли это свойство для обобщенной регрессионной модели (5.3):

$$y = X\beta + \varepsilon,$$

т.е. верно ли, что $\operatorname{Cov}(\widehat{y}, e) = 0$, где $\widehat{y} = X\widehat{\beta}_{\text{OLS}}$ и $e = y - \widehat{y}$?

## Решение

Поскольку

$$\widehat{\beta}_{\text{OLS}} = (X'X)^{-1}X'y,$$

то

$$\widehat{y} = X(X'X)^{-1}X'y = Ny, \qquad e = (I - N)y,$$

где $N$ есть матрица оператора ортогонального проектирования на линейную оболочку, порожденную регрессорами. Тогда

$$\operatorname{Cov}(\widehat{y}, e) = \operatorname{Cov}(Ny, (I - N)y) = N\operatorname{Cov}(y, y)(I - N) = N\Omega(I - N)$$

(мы воспользовались здесь формулой (МС.9) и симметричностью матрицы $N$). В классической линейной регрессии $\Omega = \sigma^2 I$, поэтому равенство $\operatorname{Cov}(\widehat{y}, e) = 0$ сводится к утверждению, что $N^2 = N$.

В случае обобщенной модели равенство $\mathrm{Cov}(\widehat{\boldsymbol{y}}, \boldsymbol{e}) = \mathbf{0}$, вообще говоря, не выполняется. Простой контрпример:

$$\boldsymbol{X} = \begin{bmatrix} 1 \\ 0 \end{bmatrix}, \quad \boldsymbol{\Omega} = \begin{bmatrix} 1 & 1/2 \\ 1/2 & 1 \end{bmatrix}, \quad \boldsymbol{N} = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}, \quad \boldsymbol{N\Omega}(\boldsymbol{I} - \boldsymbol{N}) = \begin{bmatrix} 0 & 1/2 \\ 0 & 0 \end{bmatrix}.$$

## Задача 5.5

Докажите, что если в (5.3):

$$\boldsymbol{y} = \boldsymbol{X\beta} + \boldsymbol{\varepsilon}$$

вектор ошибок $\boldsymbol{\varepsilon}$ имеет многомерное нормальное распределение, то $\widehat{\boldsymbol{\beta}}_{\mathrm{GLS}} = \widehat{\boldsymbol{\beta}}_{\mathrm{ML}}$.

## Решение

Согласно (5.4),

$$\widehat{\boldsymbol{\beta}}_{\mathrm{GLS}} = (\boldsymbol{X}'\boldsymbol{\Omega}^{-1}\boldsymbol{X})^{-1}\boldsymbol{X}'\boldsymbol{\Omega}^{-1}\boldsymbol{y}.$$

Кроме того, в силу условия задачи $\boldsymbol{y} \sim N(\boldsymbol{X\beta}, \boldsymbol{\Omega})$.

Запишем функцию правдоподобия:

$$L(\boldsymbol{y}) = (2\pi)^{-n/2}(\det \boldsymbol{\Omega})^{-1/2} \exp\left(-\frac{1}{2}(\boldsymbol{y} - \boldsymbol{X\beta})'\boldsymbol{\Omega}^{-1}(\boldsymbol{y} - \boldsymbol{X\beta})\right),$$

где $n$ — число наблюдений. Очевидно, что $L(\boldsymbol{y})$ максимальна, когда минимально выражение:

$$g(\boldsymbol{\beta}) = (\boldsymbol{y} - \boldsymbol{X\beta})'\boldsymbol{\Omega}^{-1}(\boldsymbol{y} - \boldsymbol{X\beta}).$$

Таким образом, $g(\boldsymbol{\beta}_{\mathrm{ML}}) = \min g(\boldsymbol{\beta})$. Далее, как показано в доказательстве теоремы Айткена, оценка $\widehat{\boldsymbol{\beta}}_{\mathrm{GLS}}$ получается минимизацией по $\boldsymbol{\beta}$ суммы квадратов отклонений

$$f(\boldsymbol{\beta}) = \boldsymbol{e}^{*\prime}\boldsymbol{e}^*$$

для преобразованного уравнения (5.6):

$$\boldsymbol{y}^* = \boldsymbol{X}^*\boldsymbol{\beta} + \boldsymbol{\varepsilon}^*,$$

где $\boldsymbol{y}^* = \boldsymbol{Py}$, $\boldsymbol{X}^* = \boldsymbol{PX}$, $\boldsymbol{\varepsilon}^* = \boldsymbol{P\varepsilon}$, $\boldsymbol{P}'\boldsymbol{P} = \boldsymbol{\Omega}^{-1}$. Но

$$\begin{aligned} f(\boldsymbol{\beta}) &= (\boldsymbol{y}^* - \boldsymbol{X}^*\boldsymbol{\beta})'(\boldsymbol{y}^* - \boldsymbol{X}^*\boldsymbol{\beta}) = (\boldsymbol{y} - \boldsymbol{X\beta})'\boldsymbol{P}'\boldsymbol{P}(\boldsymbol{y} - \boldsymbol{X\beta}) \\ &= (\boldsymbol{y} - \boldsymbol{X\beta})'\boldsymbol{\Omega}^{-1}(\boldsymbol{y} - \boldsymbol{X\beta}). \end{aligned}$$

Функции $g(\boldsymbol{\beta})$ и $f(\boldsymbol{\beta})$ совпадают, и, следовательно, $\widehat{\boldsymbol{\beta}}_{\mathrm{GLS}} = \widehat{\boldsymbol{\beta}}_{\mathrm{ML}}$.

## Задача 5.6

Рассмотрим уравнение регрессии:

$$y_t = \beta + \varepsilon_t, \quad t = 1, \ldots, n.$$

Пусть ошибки регрессии удовлетворяют следующим условиям:

$$\mathrm{E}(\varepsilon_t) = 0; \quad \mathrm{Cov}(\varepsilon_t, \varepsilon_s) = 0,\ t \neq s; \quad \mathrm{V}(\varepsilon_t) = \sigma^2 x_t,\ x_t > 0.$$

а) Найдите оценку метода наименьших квадратов $\widehat{\beta}$ и ее дисперсию.

б) Предложите несмещенную оценку, обладающую меньшей дисперсией, чем оценка метода наименьших квадратов. Получите дисперсию этой оценки и сравните ее с дисперсией оценки метода наименьших квадратов. Интерпретируйте результат.

## Решение

а) Известно (см. задачу 2.12), что

$$\widehat{\beta}_{\mathrm{OLS}} = \widehat{\beta} = \overline{y} = \frac{1}{n}(y_1 + \ldots + y_n).$$

Из условия следует, что $\mathrm{E}(y_t) = \beta$, $\mathrm{V}(y_t) = \sigma^2 x_t$ и $\mathrm{Cov}(y_t, y_s) = 0$, при $t \neq s$. Поэтому

$$\mathrm{E}(\widehat{\beta}) = \beta$$

(т.е. оценка $\widehat{\beta}$ несмещенная) и

$$\mathrm{V}(\widehat{\beta}) = \frac{\sigma^2}{n^2}\sum_{t=1}^{n} x_t.$$

б) Обобщенный метод наименьших квадратов в данном случае сводится к взвешенному методу наименьших квадратов с весами $1/\sqrt{x_t}$:

$$\frac{y_t}{\sqrt{x_t}} = \beta\frac{1}{\sqrt{x_t}} + \frac{\varepsilon_t}{\sqrt{x_t}}.$$

Здесь ошибки $u_t = \varepsilon_t/\sqrt{x_t}$ уже удовлетворяют условию гомоскедастичности: $\mathrm{V}(u_t) = \sigma^2$. Применяя к этому уравнению МНК, получаем:

$$\widehat{\beta}_{\mathrm{GLS}} = \left(\sum_{t=1}^{n}\frac{y_t}{x_t}\right)\Bigg/\left(\sum_{t=1}^{n}\frac{1}{x_t}\right),$$

$$\mathrm{V}(\widehat{\beta}_{\mathrm{GLS}}) = \left(\sum_{t=1}^{n}\frac{\sigma^2 x_t}{x_t^2}\right)\Bigg/\left(\sum_{t=1}^{n}\frac{1}{x_t}\right)^2 = \sigma^2\Bigg/\left(\sum_{t=1}^{n}\frac{1}{x_t}\right).$$

Неравенство $V(\widehat{\beta}) \geqslant V(\widehat{\beta}_{\mathrm{GLS}})$ эквивалентно неравенству

$$\left(\sum_{t=1}^{n} x_t\right)\left(\sum_{t=1}^{n} \frac{1}{x_t}\right) \geqslant n^2,$$

которое, в свою очередь, вытекает непосредственно из неравенства Коши–Буняковского.

## Задача 5.7

Рассмотрим следующую регрессионную модель, в которой $2n$ наблюдений разбиты на две равные группы по $n$ наблюдений в каждой:

$$\boldsymbol{y} = \boldsymbol{X\beta} + \varepsilon, \qquad \mathrm{E}(\varepsilon) = \mathbf{0}; \quad \mathrm{Cov}(\varepsilon_t, \varepsilon_s) = 0,\ t \neq s;$$
$$\mathrm{V}(\varepsilon_t) = \sigma_1^2,\ t = 1, \ldots, n; \quad \mathrm{V}(\varepsilon_t) = \sigma_2^2,\ t = n+1, \ldots, 2n.$$

Введем естественное разбиение матриц на блоки:

$$\boldsymbol{y} = \begin{bmatrix} \boldsymbol{y}_1 \\ \boldsymbol{y}_2 \end{bmatrix}, \quad \boldsymbol{X} = \begin{bmatrix} \boldsymbol{X}_1 \\ \boldsymbol{X}_2 \end{bmatrix}, \quad \varepsilon = \begin{bmatrix} \varepsilon_1 \\ \varepsilon_2 \end{bmatrix}.$$

(Здесь $\boldsymbol{y}_1$, $\boldsymbol{y}_2$, $\varepsilon_1$, $\varepsilon_2$ — $n \times 1$ векторы, $\boldsymbol{X}_1$, $\boldsymbol{X}_2$ — $n \times k$ матрицы.)

а) Пусть $\widehat{\beta}_1$, $\widehat{\beta}_2$ и $\widehat{\beta}$ — оценки метода наименьших квадратов вектора коэффициентов $\beta$ по первой группе наблюдений, по второй группе наблюдений и по всем $2n$ наблюдениям соответственно. Покажите, что $\widehat{\beta}$ есть «взвешенное среднее» оценок $\widehat{\beta}_1$ и $\widehat{\beta}_2$, в том смысле, что $\widehat{\beta} = \boldsymbol{L}_1\widehat{\beta}_1 + \boldsymbol{L}_2\widehat{\beta}_2$, где $\boldsymbol{L}_1$ и $\boldsymbol{L}_2$ — $k \times k$ матрицы такие, что $\boldsymbol{L}_1 + \boldsymbol{L}_2 = \boldsymbol{I}_k$.

б) Выведите следующие формулы для оценки обобщенного метода наименьших квадратов:

$$\widehat{\beta}_{\mathrm{GLS}} = \left(\frac{\boldsymbol{X}'_1\boldsymbol{X}_1}{\sigma_1^2} + \frac{\boldsymbol{X}'_2\boldsymbol{X}_2}{\sigma_2^2}\right)^{-1}\left(\frac{\boldsymbol{X}'_1\boldsymbol{y}_1}{\sigma_1^2} + \frac{\boldsymbol{X}'_2\boldsymbol{y}_2}{\sigma_2^2}\right),$$

$$\mathrm{V}(\widehat{\beta}_{\mathrm{GLS}}) = \left(\frac{\boldsymbol{X}'_1\boldsymbol{X}_1}{\sigma_1^2} + \frac{\boldsymbol{X}'_2\boldsymbol{X}_2}{\sigma_2^2}\right)^{-1}.$$

в) Покажите, что $\widehat{\beta}_{\mathrm{GLS}}$ также является «взвешенным средним» оценок $\widehat{\beta}_1$ и $\widehat{\beta}_2$, в том смысле, что существуют $k \times k$ матрицы $\Lambda_1$ и $\Lambda_2$ такие, что $\widehat{\beta}_{\mathrm{GLS}} = \Lambda_1\widehat{\beta}_1 + \Lambda_2\widehat{\beta}_2$, $\Lambda_1 + \Lambda_2 = \boldsymbol{I}_k$.

## Решение

а) Используя МНК и правила вычислений с блочными матрицами, получаем:

$$
\begin{aligned}
\widehat{\beta}_1 &= (X_1'X_1)^{-1}X_1'y_1,\\
\widehat{\beta}_2 &= (X_2'X_2)^{-1}X_2'y_2,\\
\widehat{\beta} &= (X'X)^{-1}X'y = \left(\begin{bmatrix} X_1' & X_2' \end{bmatrix}\begin{bmatrix} X_1 \\ X_2 \end{bmatrix}\right)^{-1}\begin{bmatrix} X_1' & X_2' \end{bmatrix}\begin{bmatrix} y_1 \\ y_2 \end{bmatrix}\\
&= (X_1'X_1 + X_2'X_2)^{-1}(X_1'y_1 + X_2'y_2).
\end{aligned}
$$

Из формул для $\widehat{\beta}_i$ получаем $X_i'y_i = X_i'X_i\widehat{\beta}_i$, $i = 1, 2$. Подставляя в предыдущее уравнение, получаем:

$$
\widehat{\beta} = (X_1'X_1 + X_2'X_2)^{-1}(X_1'X_1\widehat{\beta}_1 + X_2'X_2\widehat{\beta}_2) = L_1\widehat{\beta}_1 + L_2\widehat{\beta}_2,
$$

где

$$
\begin{aligned}
L_1 &= (X_1'X_1 + X_2'X_2)^{-1}X_1'X_1,\\
L_2 &= (X_1'X_1 + X_2'X_2)^{-1}X_2'X_2.
\end{aligned}
$$

Очевидно, что $L_1 + L_2 = I$.

б) Пусть $\Omega$ — матрица ковариаций вектора ошибок $\varepsilon$. Тогда

$$
\Omega = \begin{bmatrix} \sigma_1^2 I_n & 0 \\ 0 & \sigma_2^2 I_n \end{bmatrix}, \qquad \Omega^{-1} = \begin{bmatrix} \frac{1}{\sigma_1^2} I_n & 0 \\ 0 & \frac{1}{\sigma_2^2} I_n \end{bmatrix}.
$$

По формуле (5.4) для оценки обобщенного метода наименьших квадратов получаем:

$$
\begin{aligned}
\widehat{\beta}_{\mathrm{GLS}} &= (X'\Omega^{-1}X)^{-1}X'\Omega^{-1}y\\
&= \left(\begin{bmatrix} X_1' & X_2' \end{bmatrix}\begin{bmatrix} \frac{1}{\sigma_1^2} I_n & 0 \\ 0 & \frac{1}{\sigma_2^2} I_n \end{bmatrix}\begin{bmatrix} X_1 \\ X_2 \end{bmatrix}\right)^{-1}\begin{bmatrix} X_1' & X_2' \end{bmatrix}\begin{bmatrix} \frac{1}{\sigma_1^2} I_n & 0 \\ 0 & \frac{1}{\sigma_2^2} I_n \end{bmatrix}\begin{bmatrix} y_1 \\ y_2 \end{bmatrix}\\
&= \left(\frac{X_1'X_1}{\sigma_1^2} + \frac{X_2'X_2}{\sigma_2^2}\right)^{-1}\left(\frac{X_1'y_1}{\sigma_1^2} + \frac{X_2'y_2}{\sigma_2^2}\right).
\end{aligned}
$$

Далее, в силу (5.8)

$$
\mathrm{V}(\widehat{\beta}_{\mathrm{GLS}}) = (X'\Omega^{-1}X)^{-1} = \left(\frac{X_1'X_1}{\sigma_1^2} + \frac{X_2'X_2}{\sigma_2^2}\right)^{-1}.
$$

в) Аналогично п. а) нетрудно получить, что

$$
\widehat{\beta}_{\mathrm{GLS}} = \Lambda_1\widehat{\beta}_1 + \Lambda_2\widehat{\beta}_2,
$$

где

$$\Lambda_1 = \left(\frac{X_1'X_1}{\sigma_1^2} + \frac{X_2'X_2}{\sigma_2^2}\right)^{-1} \frac{X_1'X_1}{\sigma_1^2},$$

$$\Lambda_2 = \left(\frac{X_1'X_1}{\sigma_1^2} + \frac{X_2'X_2}{\sigma_2^2}\right)^{-1} \frac{X_2'X_2}{\sigma_2^2}.$$

Очевидно, $\Lambda_1 + \Lambda_2 = I_k$.

## Задача 5.8

Рассмотрим модель из задачи 5.7. Опишите процедуру доступного обобщенного метода наименьших квадратов в применении к этой модели.

## Решение

Согласно формуле (5.16) оценка доступного обобщенного метода наименьших квадратов имеет вид:

$$\widehat{\beta}_{\text{FGLS}} = (X'\Omega(\widehat{\theta})^{-1}X)^{-1}X'\Omega(\widehat{\theta})^{-1}y,$$

где $\widehat{\theta}$ — состоятельная оценка вектора параметров $\theta$. В нашем случае

$$\theta = \begin{bmatrix}\sigma_1^2\\ \sigma_2^2\end{bmatrix}, \qquad \Omega(\theta) = \begin{bmatrix}\sigma_1^2 I_n & 0\\ 0 & \sigma_2^2 I_n\end{bmatrix}.$$

Проведем оценку уравнения регрессии $y = \beta X + \varepsilon$ по первым $n$ и оставшимся $n$ наблюдениям. Обозначим через

$$e_1 = y_1 - X_1(X_1'X_1)^{-1}X_1'y_1,$$
$$e_2 = y_2 - X_2(X_2'X_2)^{-1}X_2'y_2$$

соответствующие векторы остатков. Поскольку в каждом из двух случаев выполнены условия классической регрессионной модели, то оценки дисперсий ошибок

$$\widehat{\sigma}_i^2 = s_i^2 = \frac{e_i'e_i}{n-k}, \quad i = 1, 2,$$

являются состоятельными. Поэтому в соответствии с формулами, полученными в задаче 5.7, оценку доступного обобщенного метода наименьших квадратов получаем по формуле

$$\widehat{\beta}_{\text{FGLS}} = \left(\begin{bmatrix}X_1\\ X_2\end{bmatrix}' \begin{bmatrix}s_1^2 I_n & 0\\ 0 & s_2^2 I_n\end{bmatrix}^{-1} \begin{bmatrix}X_1\\ X_2\end{bmatrix}\right)^{-1} \begin{bmatrix}X_1\\ X_2\end{bmatrix}' \begin{bmatrix}s_1^2 I_n & 0\\ 0 & s_2^2 I_n\end{bmatrix}^{-1} \begin{bmatrix}y_1\\ y_2\end{bmatrix}$$
$$= \left(\frac{X_1'X_1}{s_1^2} + \frac{X_2'X_2}{s_2^2}\right)^{-1} \left(\frac{X_1'y_1}{s_1^2} + \frac{X_2'y_2}{s_2^2}\right).$$

## Задача 5.9

В этой задаче мы покажем, что если $\boldsymbol{\Omega} = \boldsymbol{\Omega}(\theta)$ и $\widehat{\theta}$ — состоятельная оценка $\theta$, то, вообще говоря, оценка доступного обобщенного метода наименьших квадратов не будет иметь то же асимптотическое распределение, что и оценка обобщенного метода наименьших квадратов.

Пусть

$$\boldsymbol{x} = \begin{bmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{bmatrix}, \quad \boldsymbol{\Omega}^{-1}(\theta) = \begin{bmatrix} 1 & 0 & \dots & 0 \\ 0 & \theta & \dots & 0 \\ \vdots & \vdots & \ddots & 0 \\ 0 & \dots & 0 & \theta^{n-1} \end{bmatrix} \quad \text{и} \quad \widehat{\theta} = \theta + \frac{1}{n}.$$

а) Покажите, что $p\lim(\widehat{\theta}) = \theta$.

б) При $\theta = 1$ покажите, что

$$\frac{1}{n}\boldsymbol{x}'\boldsymbol{\Omega}^{-1}\boldsymbol{x} = 1 \quad \text{и} \quad \frac{\boldsymbol{x}'\boldsymbol{\Omega}^{-1}\boldsymbol{\varepsilon}}{\sqrt{n}} \sim N(0, \sigma^2).$$

в) Пусть $\widehat{\beta}(\theta) = (\boldsymbol{x}'\boldsymbol{\Omega}^{-1}(\theta)\boldsymbol{x})^{-1}\boldsymbol{x}'\boldsymbol{\Omega}^{-1}(\theta)\boldsymbol{y}$. Покажите, что при $\theta = 1$

$$\sqrt{n}(\widehat{\beta}(\theta) - \beta) \sim N(0, \sigma^2).$$

г) С другой стороны, покажите, что при $\theta = 1$

$$\frac{\boldsymbol{x}'\boldsymbol{\Omega}^{-1}(\widehat{\theta})\boldsymbol{x}}{n} \to e - 1$$

и

$$\frac{\boldsymbol{x}'\boldsymbol{\Omega}^{-1}(\widehat{\theta})\boldsymbol{\varepsilon}}{\sqrt{n}} \xrightarrow{d} N\left(0, \sigma^2\frac{e^2-1}{2}\right).$$

д) Следовательно, при $\theta = 1$

$$\sqrt{n}(\widehat{\beta}(\widehat{\theta}) - \beta) \xrightarrow{d} N\left(0, \sigma^2\frac{e+1}{2(e-1)}\right)$$

(напомним, что символом $\xrightarrow{d}$ обозначается сходимость по распределению (см. приложение МС, п. 5)).

е) Выведите отсюда, что когда значение $\theta$ равно 1, асимптотическое распределение $\widehat{\beta}(\widehat{\theta})$ (оценки доступного обобщенного метода наименьших квадратов) *не совпадает* с асимптотическим распределением $\widehat{\beta}(\theta)$ (оценки обобщенного метода наименьших квадратов) вопреки тому, что $\widehat{\theta}$ является состоятельной оценкой $\theta$.

## Решение

а) Величина $\widehat{\theta} = \theta + 1/n$ неслучайна и сходится к $\theta$ при $n \to \infty$. Следовательно, оценка $\widehat{\theta}$ состоятельная.

б) При $\theta = 1$ матрица $\boldsymbol{\Omega}^{-1}(\theta)$ единичная. Отсюда получаем:

$$\frac{1}{n}\boldsymbol{x}'\boldsymbol{\Omega}^{-1}\boldsymbol{x} = \frac{1}{n}\begin{bmatrix}1 & \dots & 1\end{bmatrix}\begin{bmatrix}1 & \dots & 0\\ \vdots & \ddots & \vdots\\ 0 & \dots & 1\end{bmatrix}\begin{bmatrix}1\\ \vdots\\ 1\end{bmatrix} = \frac{1}{n}n = 1.$$

Заметим, что

$$\boldsymbol{x}'\boldsymbol{\Omega}^{-1} = \begin{bmatrix}1 & \dots & 1\end{bmatrix}\begin{bmatrix}1 & \dots & 0\\ \vdots & \ddots & \vdots\\ 0 & \dots & 1\end{bmatrix} = \begin{bmatrix}1 & \dots & 1\end{bmatrix},$$

следовательно,

$$\frac{\boldsymbol{x}'\boldsymbol{\Omega}^{-1}\boldsymbol{\varepsilon}}{\sqrt{n}} = \frac{\varepsilon_1 + \ldots + \varepsilon_n}{\sqrt{n}} \sim N(0, \sigma^2).$$

Здесь мы воспользовались тем, что вектор $\boldsymbol{\varepsilon} \sim N(\boldsymbol{0}, \sigma^2\boldsymbol{I})$.

в) Пусть $\widehat{\beta}(\theta) = (\boldsymbol{x}'\boldsymbol{\Omega}^{-1}(\theta)\boldsymbol{x})^{-1}\boldsymbol{x}'\boldsymbol{\Omega}^{-1}(\theta)\boldsymbol{y}$. Тогда при $\theta = 1$

$$\sqrt{n}(\widehat{\beta}(\theta) - \beta) = \sqrt{n}((\boldsymbol{x}'\boldsymbol{\Omega}^{-1}(\theta)\boldsymbol{x})^{-1}\boldsymbol{x}'\boldsymbol{\Omega}^{-1}(\theta)\boldsymbol{y} - \beta) = \sqrt{n}\left(\frac{1}{n}\boldsymbol{x}'\boldsymbol{\Omega}^{-1}(\theta)\boldsymbol{y} - \beta\right)$$

$$= \sqrt{n}\left(\frac{1}{n}\boldsymbol{x}'\boldsymbol{\Omega}^{-1}(\theta)(\beta\boldsymbol{x} + \boldsymbol{\varepsilon}) - \beta\right) = \frac{\boldsymbol{x}'\boldsymbol{\Omega}^{-1}(\theta)\boldsymbol{\varepsilon}}{\sqrt{n}} \sim N(0, \sigma^2).$$

(Здесь мы воспользовались утверждением из п. б).)

г) При $\theta = 1$ имеем:

$$\frac{\boldsymbol{x}'\boldsymbol{\Omega}^{-1}(\widehat{\theta})\boldsymbol{x}}{n} = \frac{1}{n}\begin{bmatrix}1 & 1 & \dots & 1\end{bmatrix}\begin{bmatrix}1 & 0 & \dots & 0\\ 0 & 1+\frac{1}{n} & \dots & 0\\ \vdots & \vdots & \ddots & \vdots\\ 0 & 0 & \dots & (1+\frac{1}{n})^{n-1}\end{bmatrix}\begin{bmatrix}1\\ 1\\ \vdots\\ 1\end{bmatrix}$$

$$= \frac{1}{n}\sum_{k=0}^{n-1}\left(1+\frac{1}{n}\right)^k = \frac{1}{n}\frac{\left(1+\frac{1}{n}\right)^n - 1}{(1+\frac{1}{n}) - 1}$$

$$= \left(1+\frac{1}{n}\right)^n - 1 \longrightarrow e - 1 \quad \text{при} \quad n \to \infty.$$

Рассмотрим случайную величину

$$\xi_n = \frac{\boldsymbol{x}'\boldsymbol{\Omega}^{-1}(\widehat{\theta})\boldsymbol{\varepsilon}}{\sigma\sqrt{n}} = \frac{1}{\sqrt{n}}\sum_{k=1}^{n}\left(1+\frac{1}{n}\right)^{k-1}\frac{\varepsilon_k}{\sigma}.$$

Ее характеристическая функция равна

$$\phi_{\xi_n}(\lambda) = \mathrm{E}\exp(i\lambda\xi_n) = \mathrm{E}\exp\left\{\frac{i\lambda}{\sqrt{n}}\sum_{k=1}^{n}\left(1+\frac{1}{n}\right)^{k-1}\frac{\varepsilon_k}{\sigma}\right\}$$

$$= \prod_{k=1}^{n}\phi_{\varepsilon_k/\sigma}\left(\frac{\lambda}{\sqrt{n}}\left(1+\frac{1}{n}\right)^{k-1}\right) = \prod_{k=1}^{n}\exp\left\{-\frac{1}{2}\left(\frac{\lambda}{\sqrt{n}}\left(1+\frac{1}{n}\right)^{k-1}\right)^2\right\}$$

$$= \exp\left\{-\frac{\lambda^2}{2n}\sum_{k=1}^{n}\left(\left(1+\frac{1}{n}\right)^2\right)^{k-1}\right\} = \exp\left\{-\frac{\lambda^2}{2n}\frac{(1+1/n)^{2n}-1}{(1+1/n)^2-1}\right\}$$

$$= \exp\left\{-\frac{\lambda^2}{2}\frac{(1+1/n)^{2n}-1}{2+1/n}\right\} \longrightarrow \exp\left\{-\frac{\lambda^2}{2}\frac{e^2-1}{2}\right\} \quad \text{при} \quad n\to\infty.$$

Предел совпадает с характеристической функцией нормальной случайной величины $N(0, \frac{1}{2}(e^2-1))$. Следовательно,

$$\sigma\xi_n \xrightarrow{d} N\left(0, \sigma^2\frac{e^2-1}{2}\right) \quad \text{при} \quad n\to\infty.$$

д) Пусть $\widehat{\beta}(\widehat{\theta}) = (\boldsymbol{x}'\boldsymbol{\Omega}^{-1}(\widehat{\theta})\boldsymbol{x})^{-1}\boldsymbol{x}'\boldsymbol{\Omega}^{-1}(\widehat{\theta})\boldsymbol{y}$. Тогда при $\theta = 1$ получаем:

$$\begin{aligned}\sqrt{n}(\widehat{\beta}(\widehat{\theta})-\beta) &= \sqrt{n}\left((\boldsymbol{x}'\boldsymbol{\Omega}^{-1}(\widehat{\theta})\boldsymbol{x})^{-1}\boldsymbol{x}'\boldsymbol{\Omega}^{-1}(\widehat{\theta})\boldsymbol{y}-\beta\right)\\ &= \sqrt{n}\left((\boldsymbol{x}'\boldsymbol{\Omega}^{-1}(\widehat{\theta})\boldsymbol{x})^{-1}\boldsymbol{x}'\boldsymbol{\Omega}^{-1}(\widehat{\theta})(\boldsymbol{x}\beta+\boldsymbol{\varepsilon})-\beta\right)\\ &= \sqrt{n}\left((\boldsymbol{x}'\boldsymbol{\Omega}^{-1}(\widehat{\theta})\boldsymbol{x})^{-1}\boldsymbol{x}'\boldsymbol{\Omega}^{-1}(\widehat{\theta})\boldsymbol{x}\beta-\beta\right)\\ &\quad+\sqrt{n}(\boldsymbol{x}'\boldsymbol{\Omega}^{-1}(\widehat{\theta})\boldsymbol{x})^{-1}\boldsymbol{x}'\boldsymbol{\Omega}^{-1}(\widehat{\theta})\boldsymbol{\varepsilon}\\ &= n(\boldsymbol{x}'\boldsymbol{\Omega}^{-1}(\widehat{\theta})\boldsymbol{x})^{-1}\sigma\xi_n.\end{aligned}$$

Воспользовавшись утверждениями, доказанными в п. г), получим, что

$$n(\boldsymbol{x}'\boldsymbol{\Omega}^{-1}(\widehat{\theta})\boldsymbol{x})^{-1} \longrightarrow \frac{1}{e-1},$$

и, следовательно,

$$\sqrt{n}(\widehat{\beta}(\widehat{\theta})-\beta) \xrightarrow{d} \frac{1}{e-1}N\left(0,\sigma^2\frac{e^2-1}{2}\right) = N\left(0,\sigma^2\frac{e+1}{2(e-1)}\right).$$

е) Утверждение задачи является прямым следствием того, что

$$\sqrt{n}(\widehat{\beta}(\theta)-\beta) \sim N(0,\sigma^2)$$

и

$$\sqrt{n}(\widehat{\beta}(\widehat{\theta})-\beta) \xrightarrow{d} N\left(0,\sigma^2\frac{e+1}{2(e-1)}\right).$$

## Задача 5.10

Дана обобщенная линейная регрессионная модель $\boldsymbol{y} = \boldsymbol{X}\boldsymbol{\beta} + \boldsymbol{\varepsilon}$, $\mathrm{E}\boldsymbol{\varepsilon} = 0$, $\mathrm{V}(\boldsymbol{\varepsilon}) = \boldsymbol{\Omega}$. Пусть $\widehat{\boldsymbol{\beta}}$ — оценка вектора $\boldsymbol{\beta}$ с помощью обычного метода наименьших квадратов, и пусть $\widehat{\boldsymbol{y}} = \boldsymbol{X}\widehat{\boldsymbol{\beta}}$.

а) Вычислите $\mathrm{V}(\widehat{\boldsymbol{y}})$.
б) Вычислите $\mathrm{V}(\boldsymbol{e}) = \mathrm{V}(\boldsymbol{y} - \widehat{\boldsymbol{y}})$.
в) Покажите, что в общем случае $\boldsymbol{e}$ и $\widehat{\boldsymbol{y}}$ коррелированы.

## Решение

а) Матрица ковариаций вектора $\widehat{\boldsymbol{y}}$ равна

$$\begin{aligned}\mathrm{V}(\widehat{\boldsymbol{y}}) &= \mathrm{V}(\boldsymbol{X}\widehat{\boldsymbol{\beta}}) = \mathrm{V}(\boldsymbol{X}(\boldsymbol{X}'\boldsymbol{X})^{-1}\boldsymbol{X}'\boldsymbol{y}) = \boldsymbol{X}(\boldsymbol{X}'\boldsymbol{X})^{-1}\boldsymbol{X}'\mathrm{V}(\boldsymbol{y})\boldsymbol{X}'(\boldsymbol{X}'\boldsymbol{X})^{-1}\boldsymbol{X}' \\ &= \boldsymbol{X}(\boldsymbol{X}'\boldsymbol{X})^{-1}\boldsymbol{X}'\Omega\boldsymbol{X}'(\boldsymbol{X}'\boldsymbol{X})^{-1}\boldsymbol{X}'.\end{aligned}$$

б) Матрица ковариаций вектора остатков равна

$$\begin{aligned}\mathrm{V}(\boldsymbol{e}) &= \mathrm{V}(\boldsymbol{y} - \widehat{\boldsymbol{y}}) = \mathrm{V}((\boldsymbol{I} - \boldsymbol{X}(\boldsymbol{X}'\boldsymbol{X})^{-1}\boldsymbol{X}')\boldsymbol{y}) \\ &= (\boldsymbol{I} - \boldsymbol{X}(\boldsymbol{X}'\boldsymbol{X})^{-1}\boldsymbol{X}')\mathrm{V}(\boldsymbol{y})(\boldsymbol{I} - \boldsymbol{X}(\boldsymbol{X}'\boldsymbol{X})^{-1}\boldsymbol{X}') \\ &= (\boldsymbol{I} - \boldsymbol{X}(\boldsymbol{X}'\boldsymbol{X})^{-1}\boldsymbol{X}')\Omega(\boldsymbol{I} - \boldsymbol{X}(\boldsymbol{X}'\boldsymbol{X})^{-1}\boldsymbol{X}').\end{aligned}$$

в) Вычислим ковариацию векторов $\boldsymbol{e}$ и $\widehat{\boldsymbol{y}}$:

$$\begin{aligned}\mathrm{Cov}(\boldsymbol{e}, \widehat{\boldsymbol{y}}) &= \mathrm{Cov}(\boldsymbol{y} - \widehat{\boldsymbol{y}}, \widehat{\boldsymbol{y}}) = \mathrm{Cov}((\boldsymbol{I} - \boldsymbol{X}(\boldsymbol{X}'\boldsymbol{X})^{-1}\boldsymbol{X}')\boldsymbol{y}, \boldsymbol{X}(\boldsymbol{X}'\boldsymbol{X})^{-1}\boldsymbol{X}'\boldsymbol{y}) \\ &= (\boldsymbol{I} - \boldsymbol{X}(\boldsymbol{X}'\boldsymbol{X})^{-1}\boldsymbol{X}')\mathrm{V}(\boldsymbol{y})\boldsymbol{X}(\boldsymbol{X}'\boldsymbol{X})^{-1}\boldsymbol{X}' \\ &= (\boldsymbol{I} - \boldsymbol{X}(\boldsymbol{X}'\boldsymbol{X})^{-1}\boldsymbol{X}')\Omega\boldsymbol{X}(\boldsymbol{X}'\boldsymbol{X})^{-1}\boldsymbol{X}'.\end{aligned}$$

В общем случае это выражение не равно 0.

## Задача 5.11

Как известно (см. задачу 3.26), для классической линейной модели $\boldsymbol{y} = \boldsymbol{X}\boldsymbol{\beta} + \boldsymbol{\varepsilon}$, $\mathrm{E}\boldsymbol{\varepsilon} = 0$, $\mathrm{V}(\boldsymbol{\varepsilon}) = \sigma^2\boldsymbol{I}$ выполнено неравенство $\mathrm{V}(\widehat{\boldsymbol{\beta}}_{\mathrm{R}}) \leqslant \mathrm{V}(\widehat{\boldsymbol{\beta}})$, где $\widehat{\boldsymbol{\beta}}$ — МНК-оценка вектора $\boldsymbol{\beta}$, а $\widehat{\boldsymbol{\beta}}_R$ — оценка, получаемая регрессией $\boldsymbol{y}$ на $\boldsymbol{X}$ при линейном ограничении $\boldsymbol{H}\boldsymbol{\beta} = \boldsymbol{r}$. Сохраняется ли это неравенство (для тех же оценок), если модель обобщенная, т. е. $\boldsymbol{y} = \boldsymbol{X}\boldsymbol{\beta} + \boldsymbol{\varepsilon}$, $\mathrm{E}\boldsymbol{\varepsilon} = 0$, $\mathrm{V}(\boldsymbol{\varepsilon}) = \Omega$?

## Решение

Неравенство не сохраняется. Приведем пример.

Рассмотрим модель парной регрессии $y_t = \beta_1 + \beta_2 x_t + \varepsilon_t$. Пусть есть два наблюдения с $x_1 = 0$, $x_2 = 2$ и дисперсии ошибок равны соответственно $\sigma_1^2 = 1$, $\sigma_2^2 = 4$ (ошибки некоррелированы). В силу формулы (2.12) оценка коэффициента $\beta_1$ равна

$$\widehat{\beta}_1 = \sum_{t=1}^{n} \left(\frac{1}{n} - \overline{x}w_t\right) y_t, \quad \text{где} \quad w_t = \frac{x_{*t}}{\sum_s x_{*s}^2}, \quad \text{и} \quad x_{*t} = x_t - \overline{x}.$$

Соответственно, дисперсия этой оценки равна

$$\mathrm{V}(\widehat{\beta}_1) = \sum_{t=1}^{n}\left(\frac{1}{n} - \overline{x}w_t\right)\mathrm{V}(y_t) = \sum_{t=1}^{n}\left(\frac{1}{n} - \overline{x}w_t\right)\sigma_t^2.$$

В нашем примере $n = 2$, $\overline{x} = 1$, $x_{*1} = -1$, $x_{*1} = 1$, $\sum_s x_{*s}^2 = 2$, $w_1 = -0.5$, $w_1 = 0.5$ и

$$\mathrm{V}(\widehat{\beta}_1) = (0.5 - 1\cdot(-0.5))\cdot 1 + (0.5 - 1\cdot 0.5)\cdot 4 = 1.$$

Рассмотрим регрессию с ограничением $\beta_2 = 0$. Оценка метода наименьших квадратов в регрессии с ограничением свободного члена равна

$$\widehat{\beta}_{\mathrm{R}1} = 0.5\cdot(y_1 + y_2).$$

Дисперсия этой оценки соответственно равна

$$\mathrm{V}(\widehat{\beta}_{\mathrm{R}1}) = 0.5^2\cdot(\sigma_1^2 + \sigma_2^2) = 0.25\cdot(1+4) = 1.25 > 1 = \mathrm{V}(\widehat{\beta}_1).$$

Таким образом, в приведенном примере неравенство $\mathrm{V}(\widehat{\beta}_{\mathrm{R}}) \leqslant \mathrm{V}(\widehat{\beta})$ не выполняется.

## Задача 5.12

Пусть $\widehat{\beta}_{\mathrm{GLS}} = (\boldsymbol{X}'\boldsymbol{\Omega}^{-1}\boldsymbol{X})^{-1}\boldsymbol{X}'\boldsymbol{\Omega}^{-1}\boldsymbol{y}$ — оценка, полученная с помощью обобщенного метода наименьших квадратов в обобщенной модели $\boldsymbol{y} = \boldsymbol{X}\boldsymbol{\beta} + \boldsymbol{\varepsilon}$, $\mathrm{E}\boldsymbol{\varepsilon} = 0$, $\mathrm{V}(\boldsymbol{\varepsilon}) = \boldsymbol{\Omega}$. Определим коэффициент детерминации

$$R^2_{\mathrm{GLS}} = 1 - \frac{\boldsymbol{e}'\boldsymbol{e}}{\sum_{t=1}^{n}(y_t - \overline{y})^2}, \quad \text{где} \quad \boldsymbol{e} = \boldsymbol{y} - \boldsymbol{X}\widehat{\boldsymbol{\beta}}_{\mathrm{GLS}}.$$

Обладает ли этот коэффициент привычными свойствами коэффициента детерминации в классической линейной модели? В частности, верно ли, что $R^2_{\mathrm{GLS}}$ лежит в интервале $[0, 1]$?

## Решение

При оценивании уравнения при помощи обобщенного метода наименьших квадратов минимизируется взвешенная сумма квадратов остатков (точнее — $\boldsymbol{e}'\boldsymbol{\Omega}^{-1}\boldsymbol{e}$). Поэтому при добавлении еще одного регрессора выражение $\boldsymbol{e}'\boldsymbol{\Omega}^{-1}\boldsymbol{e}$ уменьшается. Отсюда, конечно, не следует, что и сумма квадратов остатков $\boldsymbol{e}'\boldsymbol{e}$ также уменьшается. Соответственно, $R^2_{\mathrm{GLS}}$ может уменьшиться.

(*Пример.* Пусть $\boldsymbol{e}'\boldsymbol{\Omega}^{-1}\boldsymbol{e} = 4e_1^2 + e_2^2$, тогда при переходе от вектора $\boldsymbol{e} = [1 \;\; 1]'$ к вектору $\boldsymbol{e} = [0.5 \;\; 1.5]'$ выражение $\boldsymbol{e}'\boldsymbol{\Omega}^{-1}\boldsymbol{e}$ убывает от $4\cdot 1^2 + 1^2 = 5$ до $4\cdot 0.5^2 + 1.5^2 = 3.25$, а $\boldsymbol{e}'\boldsymbol{e}$ возрастает от $1^2 + 1^2 = 2$ до $0.5^2 + 1.5^2 = 2.5$.)

Свойство $R^2_{\mathrm{GLS}} \leqslant 1$ сохраняется, поскольку $\boldsymbol{e}'\boldsymbol{e}$ и $\sum_{t=1}^{n}(y_t - \overline{y})^2$ не отрицательны. Однако $R^2_{\mathrm{GLS}}$ может принимать отрицательные значения.

Как известно, оценка обобщенного метода наименьших квадратов получается как оценка метода наименьших квадратов для преобразованного уравнения $\boldsymbol{Py} = \boldsymbol{PX\beta} + \boldsymbol{P\varepsilon}$, где $P$ такая матрица, что $\boldsymbol{\Omega}^{-1} = \boldsymbol{P}'\boldsymbol{P}$. Поскольку преобразованное уравнение не содержит константу, то не обязательно $\boldsymbol{e}'\boldsymbol{e}$ не превосходит $\sum_{t=1}^{n}(y_t - \overline{y})^2$. Рассмотрим следующий пример:

$$y_t = \beta + \varepsilon_t, \quad t = 1, 2, \quad \mathrm{V}(\varepsilon_t) = \sigma_t^2.$$

Пусть $y_1 = 1$, $y_2 = -1$, $\sigma_1 = 2$, $\sigma_2 = 2/\sqrt{3}$. Получаем:

$$\begin{aligned} \widehat{\beta}_{\mathrm{GLS}} &= (\boldsymbol{X}'\boldsymbol{\Omega}^{-1}\boldsymbol{X})^{-1}\boldsymbol{X}'\boldsymbol{\Omega}^{-1}\boldsymbol{y} \\ &= \left( \begin{bmatrix} 1 & 1 \end{bmatrix} \begin{bmatrix} 1/4 & 0 \\ 0 & 3/4 \end{bmatrix} \begin{bmatrix} 1 \\ 1 \end{bmatrix} \right)^{-1} \begin{bmatrix} 1 & 1 \end{bmatrix} \begin{bmatrix} 1/4 & 0 \\ 0 & 3/4 \end{bmatrix} \begin{bmatrix} 1 \\ -1 \end{bmatrix} = -\frac{1}{2}. \end{aligned}$$

Вычислим $\boldsymbol{e}'\boldsymbol{e}$ и $\sum_{t=1}^{n}(y_t - \overline{y})^2$:

$$e_1 = y_1 - \widehat{\beta}_{\mathrm{GLS}} = 1 - \left(-\frac{1}{2}\right) = \frac{3}{2},$$

$$e_2 = y_2 - \widehat{\beta}_{\mathrm{GLS}} = -1 - \left(-\frac{1}{2}\right) = -\frac{1}{2},$$

$$\boldsymbol{e}'\boldsymbol{e} = e_1^2 + e_2^2 = \frac{9}{4} + \frac{1}{4} = \frac{5}{2} = 2.5,$$

$$\sum_{t=1}^{n}(y_t - \overline{y})^2 = y_1^2 + y_2^2 = 2, \qquad \text{поскольку} \quad \overline{y} = 0.$$

Таким образом, получаем

$$R_{\mathrm{GLS}}^2 = 1 - \frac{\boldsymbol{e}'\boldsymbol{e}}{\sum_{t=1}^{n}(y_t - \overline{y})^2} = 1 - \frac{2.5}{2} = -0.25 < 0.$$

## Задача 5.13

Пусть в уравнении $y_t = \boldsymbol{x}_t'\boldsymbol{\beta} + \varepsilon_t$, $t = 1, \ldots, n$, ошибки удовлетворяют уравнению авторегрессии первого порядка $\varepsilon_t = \rho\varepsilon_{t-1} + u_t$, $u_t \sim iid(0, \sigma_u^2)$. Пусть $\boldsymbol{\Omega} = \mathrm{V}(\boldsymbol{\varepsilon})$. Найти матрицу $\boldsymbol{P}$ такую, что $\boldsymbol{\Omega}^{-1} = \boldsymbol{P}'\boldsymbol{P}$. Покажите, как выглядит преобразованное уравнение $\boldsymbol{Py} = \boldsymbol{PX\beta} + \boldsymbol{P\varepsilon}$, которое используется для вычисления оценок обобщенного метода наименьших квадратов.

## Решение

Вычислим матрицу $\boldsymbol{\Omega}$.

$$\mathrm{V}(\varepsilon_t) = \mathrm{V}(\rho\varepsilon_{t-1} + u_t) = \rho^2\mathrm{V}(\varepsilon_{t-1}) + \sigma_u^2.$$

Здесь использовалось то, что $\varepsilon_{t-1}$ содержит только $u_{t-1}, u_{t-2}, \ldots$, и потому некоррелирована с $u_t$. Предполагая стационарность ряда $\varepsilon_t$, получаем $\sigma_\varepsilon^2 = \sigma_u^2/(1-\rho^2)$. Далее вычислим ковариации:

$$\mathrm{Cov}(\varepsilon_t, \varepsilon_{t-1}) = \mathrm{Cov}(\rho\varepsilon_{t-1} + u_t, \varepsilon_{t-1}) = \rho\sigma_\varepsilon^2,$$

$$\mathrm{Cov}(\varepsilon_t, \varepsilon_{t-2}) = \mathrm{Cov}(\rho\varepsilon_{t-1} + u_t, \varepsilon_{t-2}) = \rho\,\mathrm{Cov}(\rho\varepsilon_{t-1}, \varepsilon_{t-2}) = \rho^2\sigma_\varepsilon^2,$$

$$\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots$$

Получаем

$$\boldsymbol{\Omega} = \mathrm{V}(\varepsilon) = \frac{\sigma_u^2}{1-\rho^2}\begin{bmatrix} 1 & \rho & \rho^2 & \cdots & \rho^{n-1} \\ \rho & 1 & \rho & \cdots & \rho^{n-2} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \rho^{n-1} & \rho^{n-2} & \rho^{n-3} & \cdots & 1 \end{bmatrix}.$$

После нескольких попыток (например, рассматривая случаи $n = 2$, $n = 3$) можно понять, что

$$\boldsymbol{\Omega}^{-1} = \frac{1}{\sigma_u^2}\begin{bmatrix} 1 & -\rho & 0 & \cdots & 0 & 0 & 0 \\ -\rho & 1+\rho^2 & -\rho & \cdots & 0 & 0 & 0 \\ 0 & -\rho & 1+\rho^2 & \cdots & 0 & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & -\rho & 1+\rho^2 & -\rho \\ 0 & 0 & 0 & \cdots & 0 & -\rho & 1 \end{bmatrix},$$

а затем проверить это прямым перемножением матриц. Далее проверяем, что в качестве матрицы $\boldsymbol{P}$ можно взять матрицу

$$\boldsymbol{P} = \begin{bmatrix} \sqrt{1-\rho^2} & 0 & 0 & \cdots & 0 & 0 & 0 \\ -\rho & 1 & 0 & \cdots & 0 & 0 & 0 \\ 0 & -\rho & 1 & \cdots & 0 & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & -\rho & 1 & 0 \\ 0 & 0 & 0 & \cdots & 0 & -\rho & 1 \end{bmatrix}.$$

(Заметим, что $\sigma_u^2\boldsymbol{P}'\boldsymbol{P} = \boldsymbol{\Omega}^{-1}$ и $\mathrm{V}(\boldsymbol{P}\varepsilon) = \sigma_u^2\boldsymbol{I}$.)

Преобразованные уравнения $\boldsymbol{P}\boldsymbol{y} = \boldsymbol{P}\boldsymbol{X}\beta + \boldsymbol{P}\varepsilon$ выглядят следующим образом:

$$\sqrt{1-\rho^2}\,y_1 = \sqrt{1-\rho^2}\,x_t'\beta + \sqrt{1-\rho^2}\,\varepsilon_1,$$
$$y_t - \rho y_{t-1} = (x_t - \rho x_{t-1})'\beta + (\varepsilon_t - \rho\varepsilon_{t-1}), \quad t \geqslant 2.$$

# Глава 6

# Гетероскедастичность и корреляция по времени

## Задача 6.1

Проверьте непосредственно, что для парной регрессии (п. 2.3) с гетероскедастичностью

$$y_t = \alpha + \beta x_t + \varepsilon_t, \quad t = 1, \ldots, n;$$
$$\mathrm{E}(\varepsilon_t) = 0, \quad \mathrm{V}(\varepsilon_t) = \sigma_t^2, \quad \mathrm{E}(\varepsilon_t \varepsilon_s) = 0, \ s \neq t,$$

дисперсия оценки параметра $\beta$, полученная с помощью метода взвешенных наименьших квадратов, меньше дисперсии МНК-оценки.

## Решение

Приведем три возможных решения этой задачи.

*Решение* 1. Рассмотрим модель парной регрессии с гетероскедастичностью:

$$y_t = \alpha + \beta x_t + \varepsilon_t, \quad t = 1, \ldots, n;$$
$$\mathrm{E}(\varepsilon_t) = 0, \quad \mathrm{V}(\varepsilon_t) = \sigma_t^2, \quad \mathrm{E}(\varepsilon_t \varepsilon_s) = 0, \ s \neq t.$$

Из формул (2.4а), (2.6) в «новых» обозначениях, принятых в книге начиная с гл. 3, имеем:

$$\widehat{\beta}_{\mathrm{OLS}} = \frac{\sum x_{*t} y_{*t}}{\sum x_{*t}^2} = \frac{\sum x_{*t} y_t}{\sum x_{*t}^2}.$$

Напомним, что через $x_{*t}$ и $y_{*t}$ обозначены отклонения от средних:

$$x_{*t} = x_t - \overline{x}, \qquad y_{*t} = y_t - \overline{y}.$$

Отсюда получаем дисперсию МНК-оценки

$$\mathrm{V}(\widehat{\beta}_{\mathrm{OLS}}) = \frac{\sum x_{*t}^2 \mathrm{V}(y_t)}{(\sum x_{*t}^2)^2} = \frac{\sum x_{*t}^2 \sigma_t^2}{(\sum x_{*t}^2)^2}.$$

Далее, используя формулу (5.8), получаем дисперсию оценки обобщенного метода наименьших квадратов:

$$\mathrm{V}(\widehat{\beta}_{\mathrm{GLS}}) = \left(\sum \frac{1}{\sigma_t^2}\right) \Big/ \left(\sum \frac{1}{\sigma_t^2} \sum \frac{x_t^2}{\sigma_t^2} - \left(\sum \frac{x_t}{\sigma_t^2}\right)^2\right).$$

В силу того что уравнение

$$y_t = \alpha + \beta x_t + \varepsilon_t$$

может быть преобразовано к виду:

$$y_t = (\alpha + \beta \overline{x}) + \beta x_{*t} + \varepsilon_t,$$

выражение для дисперсии оценки обобщенного метода наименьших квадратов можно также записать в следующем виде:

$$\mathrm{V}(\widehat{\beta}_{\mathrm{GLS}}) = \left(\sum \frac{1}{\sigma_t^2}\right) \Big/ \left(\sum \frac{1}{\sigma_t^2} \sum \frac{x_{*t}^2}{\sigma_t^2} - \left(\sum \frac{x_{*t}}{\sigma_t^2}\right)^2\right).$$

Для доказательства справедливости требуемого неравенства нам понадобится следующая

*Лемма.* Справедливо следующее неравенство:

$$\sum a_t^2 \sum b_t^2 \sum c_t^2 + 2 \sum a_t b_t \sum a_t c_t \sum b_t c_t$$
$$- \sum a_t^2 \left(\sum b_t c_t\right)^2 - \sum b_t^2 \left(\sum a_t c_t\right)^2 - \sum c_t^2 \left(\sum a_t b_t\right)^2 \geqslant 0.$$

*Доказательство.* Рассмотрим симметричную матрицу

$$\boldsymbol{A} = \begin{bmatrix} \sum a_t^2 & \sum a_t b_t & \sum a_t c_t \\ \sum a_t b_t & \sum b_t^2 & \sum b_t c_t \\ \sum a_t c_t & \sum b_t c_t & \sum c_t^2 \end{bmatrix}.$$

Для любого вектора $\boldsymbol{\alpha} = (\alpha, \beta, \gamma)'$

$$\begin{aligned} \boldsymbol{\alpha}' \boldsymbol{A} \boldsymbol{\alpha} &= \begin{bmatrix} \alpha & \beta & \gamma \end{bmatrix} \begin{bmatrix} \sum a_t^2 & \sum a_t b_t & \sum a_t c_t \\ \sum a_t b_t & \sum b_t^2 & \sum b_t c_t \\ \sum a_t c_t & \sum b_t c_t & \sum c_t^2 \end{bmatrix} \begin{bmatrix} \alpha \\ \beta \\ \gamma \end{bmatrix} \\ &= \alpha^2 \sum a_t^2 + \beta^2 \sum b_t^2 + \gamma^2 \sum c_t^2 \\ &\qquad + 2\alpha\beta \sum a_t b_t + 2\alpha\gamma \sum a_t c_t + 2\beta\gamma \sum b_t c_t \\ &= \sum (\alpha a_t + \beta b_t + \gamma c_t)^2 \geqslant 0. \end{aligned}$$

Следовательно, матрица $\boldsymbol{A}$ неотрицательно определена. Значит, ее определитель неотрицателен, что и доказывает лемму:

$$\det \boldsymbol{A} = \sum a_t^2 \sum b_t^2 \sum c_t^2 + 2\sum a_t b_t \sum a_t c_t \sum b_t c_t \\ - \sum a_t^2 \left(\sum b_t c_t\right)^2 - \sum b_t^2 \left(\sum a_t c_t\right)^2 - \sum c_t^2 \left(\sum a_t b_t\right)^2 \geqslant 0.$$

Заметим, что если $\sum b_t c_t = 0$, то неравенство принимает вид:

$$\sum a_t^2 \sum b_t^2 \sum c_t^2 - \sum b_t^2 \left(\sum a_t c_t\right)^2 - \sum c_t^2 \left(\sum a_t b_t\right)^2 \geqslant 0.$$

Разделим $\mathrm{V}(\widehat{\beta}_{\mathrm{OLS}})$ на $\mathrm{V}(\widehat{\beta}_{\mathrm{GLS}})$, обозначим $a_t = x_{*t}/\sigma_t$, $b_t = x_{*t}\sigma_t$, $c_t = 1/\sigma_t$, заметим, что $\sum b_t c_t = \sum x_{*t}\sigma_t/\sigma_t = \sum x_{*t} = 0$, и воспользуемся доказанной леммой:

$$\frac{\mathrm{V}(\widehat{\beta}_{\mathrm{OLS}})}{\mathrm{V}(\widehat{\beta}_{\mathrm{GLS}})} - 1 = \frac{\sum x_{*t}^2\sigma_t^2}{(\sum x_{*t}^2)^2} \cdot \frac{\sum \frac{1}{\sigma_t^2}\sum \frac{x_{*t}^2}{\sigma_t^2} - \left(\sum \frac{x_{*t}}{\sigma_t^2}\right)^2}{\sum \frac{1}{\sigma_t^2}} - 1$$

$$= \frac{\sum a_t^2 \sum b_t^2 \sum c_t^2 - \sum b_t^2 (\sum a_t c_t)^2 - \sum c_t^2 (\sum a_t b_t)^2}{\sum c_t^2 (\sum a_t b_t)^2} \geqslant 0.$$

*Решение* 2. Воспользуемся матричными обозначениями:

$$\begin{aligned} \widehat{\beta}_{\mathrm{OLS}} - \widehat{\beta}_{\mathrm{GLS}} &= (\boldsymbol{X}'\boldsymbol{X})^{-1}\boldsymbol{X}'\boldsymbol{y} - (\boldsymbol{X}'\boldsymbol{\Omega}^{-1}\boldsymbol{X})^{-1}\boldsymbol{X}'\boldsymbol{\Omega}^{-1}\boldsymbol{y} \\ &= (\boldsymbol{X}'\boldsymbol{X})^{-1}\boldsymbol{X}'\varepsilon - (\boldsymbol{X}'\boldsymbol{\Omega}^{-1}\boldsymbol{X})^{-1}\boldsymbol{X}'\boldsymbol{\Omega}^{-1}\varepsilon \\ &= ((\boldsymbol{X}'\boldsymbol{X})^{-1}\boldsymbol{X}' - (\boldsymbol{X}'\boldsymbol{\Omega}^{-1}\boldsymbol{X})^{-1}\boldsymbol{X}'\boldsymbol{\Omega}^{-1})\varepsilon. \end{aligned}$$

Вычислим матрицу ковариаций разности оценок (здесь мы используем обозначение $\mathrm{V}(\varepsilon) = \boldsymbol{\Omega}$ и равенство (MC.9): $\mathrm{V}(\boldsymbol{A}\varepsilon) = \boldsymbol{A}\mathrm{V}(\varepsilon)\boldsymbol{A}'$):

$$\begin{aligned} \mathrm{V}(\widehat{\beta}_{\mathrm{OLS}} - \widehat{\beta}_{\mathrm{GLS}}) &= \mathrm{V}\Big(\big((\boldsymbol{X}'\boldsymbol{X})^{-1}\boldsymbol{X}' - (\boldsymbol{X}'\boldsymbol{\Omega}^{-1}\boldsymbol{X})^{-1}\boldsymbol{X}'\boldsymbol{\Omega}^{-1}\big)\varepsilon\Big) \\ &= (\boldsymbol{X}'\boldsymbol{X})^{-1}\boldsymbol{X}'\boldsymbol{\Omega}\boldsymbol{X}(\boldsymbol{X}'\boldsymbol{X})^{-1} \\ &\quad - (\boldsymbol{X}'\boldsymbol{X})^{-1}\boldsymbol{X}'\boldsymbol{\Omega}\boldsymbol{\Omega}^{-1}\boldsymbol{X}(\boldsymbol{X}'\boldsymbol{\Omega}^{-1}\boldsymbol{X})^{-1} \\ &\quad - (\boldsymbol{X}'\boldsymbol{\Omega}^{-1}\boldsymbol{X})^{-1}\boldsymbol{X}'\boldsymbol{\Omega}^{-1}\boldsymbol{\Omega}\boldsymbol{X}(\boldsymbol{X}'\boldsymbol{X})^{-1} \\ &\quad + (\boldsymbol{X}'\boldsymbol{\Omega}^{-1}\boldsymbol{X})^{-1}\boldsymbol{X}'\boldsymbol{\Omega}^{-1}\boldsymbol{\Omega}\boldsymbol{\Omega}^{-1}\boldsymbol{X}(\boldsymbol{X}'\boldsymbol{\Omega}^{-1}\boldsymbol{X})^{-1} \\ &= (\boldsymbol{X}'\boldsymbol{X})^{-1}\boldsymbol{X}'\boldsymbol{\Omega}\boldsymbol{X}(\boldsymbol{X}'\boldsymbol{X})^{-1} \\ &\quad - (\boldsymbol{X}'\boldsymbol{\Omega}^{-1}\boldsymbol{X})^{-1} - (\boldsymbol{X}'\boldsymbol{\Omega}^{-1}\boldsymbol{X})^{-1} + (\boldsymbol{X}'\boldsymbol{\Omega}^{-1}\boldsymbol{X})^{-1} \\ &= (\boldsymbol{X}'\boldsymbol{X})^{-1}\boldsymbol{X}'\boldsymbol{\Omega}\boldsymbol{X}(\boldsymbol{X}'\boldsymbol{X})^{-1} - (\boldsymbol{X}'\boldsymbol{\Omega}^{-1}\boldsymbol{X})^{-1} \\ &= \mathrm{V}(\widehat{\beta}_{\mathrm{OLS}}) - \mathrm{V}(\widehat{\beta}_{\mathrm{GLS}}). \end{aligned}$$

Поскольку матрица ковариаций неотрицательно определена, то

$$V(\widehat{\beta}_{\mathrm{OLS}} - \widehat{\beta}_{\mathrm{GLS}}) \geqslant 0,$$

а следовательно,

$$V(\widehat{\beta}_{\mathrm{OLS}}) - V(\widehat{\beta}_{\mathrm{GLS}}) \geqslant 0.$$

Отсюда и следует искомое неравенство

$$V(\widehat{\beta}_{\mathrm{OLS}}) - V(\widehat{\beta}_{\mathrm{GLS}}) \geqslant 0.$$

*Примечание.* Заметим, что аналогичная идея используется в тесте Хаусмана (см. гл. 8).

*Решение* 3. Так как $\boldsymbol{X} = \begin{bmatrix}\boldsymbol{\imath} & \boldsymbol{x}\end{bmatrix}$ и

$$\boldsymbol{\Omega} = \begin{bmatrix} \sigma_1^2 & \dots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \dots & \sigma_n^2 \end{bmatrix},$$

то

$$(\boldsymbol{X}'\boldsymbol{\Omega}^{-1}\boldsymbol{X})^{-1} = \frac{1}{\Delta}\begin{bmatrix} \sum \frac{x_t^2}{\sigma_t^2} & -\sum \frac{x_t}{\sigma_t^2} \\ -\sum \frac{x_t}{\sigma_t^2} & \sum \frac{1}{\sigma_t^2} \end{bmatrix},$$

где

$$\Delta = \sum \frac{1}{\sigma_t^2} \sum \frac{x_t^2}{\sigma_t^2} - \left(\sum \frac{x_t}{\sigma_t^2}\right)^2.$$

Рассмотрим теперь следующую матрицу $\boldsymbol{P}$:

$$\boldsymbol{P} = \boldsymbol{\Omega}^{-1/2}\boldsymbol{X}(\boldsymbol{X}'\boldsymbol{\Omega}^{-1}\boldsymbol{X})^{-1}\boldsymbol{X}'\boldsymbol{\Omega}^{-1/2}.$$

Нетрудно проверить, что эта матрица идемпотентная (т. е. она симметричная и является проектором). Тогда по свойству идемпотентной матрицы получаем, что для любого вектора $\boldsymbol{z}$ выполняется неравенство: $\boldsymbol{z}'\boldsymbol{P}\boldsymbol{z} \leqslant \boldsymbol{z}'\boldsymbol{z}$.

Возьмем теперь в качестве $\boldsymbol{z}$ вектор с компонентами $z_t = \sigma_t x_{*t}$. Поскольку

$$\boldsymbol{X}'\boldsymbol{\Omega}^{-1/2}\boldsymbol{z} = \begin{bmatrix}\boldsymbol{\imath}' \\ \boldsymbol{x}'\end{bmatrix}\begin{bmatrix} x_{*1} \\ \vdots \\ x_{*n}\end{bmatrix} = \sum x_{*t}^2 \begin{bmatrix}0 \\ 1\end{bmatrix},$$

то

$$\boldsymbol{z}'\boldsymbol{P}\boldsymbol{z} = \frac{\left(\sum x_{*t}^2\right)^2}{\Delta}\begin{bmatrix}0 & 1\end{bmatrix}\begin{bmatrix} \sum \frac{x_t^2}{\sigma_t^2} & -\sum \frac{x_t}{\sigma_t^2} \\ -\sum \frac{x_t}{\sigma_t^2} & \sum \frac{1}{\sigma_t^2} \end{bmatrix}\begin{bmatrix}0 \\ 1\end{bmatrix}$$

$$= \frac{\left(\sum x_{*t}^2\right)^2 \sum \frac{1}{\sigma_t^2}}{\Delta} \leqslant \boldsymbol{z}'\boldsymbol{z} = \sum \sigma_t^2 x_{*t}^2.$$

Разделив последнее неравенство на $\left(\sum x_{*t}^2\right)^2$, получаем требуемое неравенство.

*Замечание*. Установим свойство идемпотентных матриц, которое мы использовали в доказательстве. Легко проверить, что если матрица $\boldsymbol{P}$ идемпотентная, то матрица $\boldsymbol{I}-\boldsymbol{P}$ также является идемпотентной. Поскольку идемпотентная матрица может иметь собственные числа, равные только 1 или 0, то она неотрицательно определена. Отсюда для любого вектора $\boldsymbol{z}$ верно:

$$\boldsymbol{z}'(\boldsymbol{I}-\boldsymbol{P})\boldsymbol{z} \geqslant 0, \quad \text{или} \quad \boldsymbol{z}'\boldsymbol{I}\boldsymbol{z}-\boldsymbol{z}'\boldsymbol{P}\boldsymbol{z} \geqslant 0, \quad \text{или} \quad \boldsymbol{z}'\boldsymbol{z} \geqslant \boldsymbol{z}'\boldsymbol{P}\boldsymbol{z},$$

что и требовалось доказать.

## Задача 6.2

Процесс, порождающий данные, описывается уравнением

$$\begin{aligned} &y_t = \beta x_t + \varepsilon_t, \\ &\mathrm{E}\,\varepsilon_t = 0, \quad \mathrm{E}\,(\varepsilon_t^2) = \sigma^2, \quad \mathrm{E}\,(\varepsilon_t\varepsilon_s) = 0, \qquad t \neq s, \quad t = 1, \ldots, n. \end{aligned}$$

Экспериментатор не имеет доступа к исходным данным, а может использовать лишь «групповые» данные. А именно, значения независимой переменной упорядочиваются по величине ($x_1 < x_2 < \cdots < x_n$), вычисляются средние значения в первой группе из $n_1$ наблюдений

$$\overline{x}_1 = \frac{1}{n_1}\sum_{t=1}^{n_1} x_t, \quad \overline{y}_1 = \frac{1}{n_1}\sum_{t=1}^{n_1} y_t,$$

во второй группе — из $n_2$ наблюдений

$$\overline{x}_2 = \frac{1}{n_2}\sum_{t=n_1+1}^{n_1+n_2} x_t, \quad \overline{y}_2 = \frac{1}{n_2}\sum_{t=n_1+1}^{n_1+n_2} y_t$$

и т.д. Всего есть $J$ групп наблюдений, $j$-я группа имеет объем $n_j$. Параметр $\beta$ оценивается с помощью регрессии $\overline{y}_j$ на $\overline{x}_j$, $j = 1, \ldots, J$. Вычислите среднее значение и дисперсию оценки. Оцените потерю эффективности в результате такой группировки данных.

### Решение

Обозначим через $N_i = n_1 + \ldots + n_{i-1}$ при $i = 2, \ldots, J+1$, и $N_1 = 0$. Воспользовавшись определением $\overline{x}_i$ и $\overline{y}_i$, получаем:

$$\begin{aligned} \overline{y}_i &= \frac{1}{n_i}\sum_{t=N_i+1}^{N_{i+1}} y_t = \frac{1}{n_i}\sum_{t=N_i+1}^{N_{i+1}} (\beta x_t + \varepsilon_t) = \beta\overline{x}_i + \frac{1}{n_i}\sum_{t=N_i+1}^{N_{i+1}} \varepsilon_t \\ &= \beta\overline{x}_i + \overline{\varepsilon}_i, \qquad i = 1, \ldots, J. \end{aligned}$$

При этом

$$\mathrm{E}(\overline{\varepsilon}_i) = \frac{1}{n_i} \sum_{t=N_i+1}^{N_{i+1}} \mathrm{E}\varepsilon_t = 0, \qquad \mathrm{V}(\overline{\varepsilon}_i) = \mathrm{V}\left(\frac{1}{n_i} \sum_{t=N_i+1}^{N_{i+1}} \varepsilon_t\right) = \frac{\sigma^2}{n_i}$$

и

$$\mathrm{Cov}(\overline{\varepsilon}_i, \overline{\varepsilon}_j) = 0, \quad i \neq j.$$

Таким образом, матрица ковариаций ошибок выглядит следующим образом:

$$\sigma^2 \mathbf{\Omega} = \sigma^2 \begin{bmatrix} 1/n_1 & 0 & \cdots & 0 \\ 0 & 1/n_2 & \cdots & 0 \\ \vdots & \vdots & & 0 \\ 0 & 0 & \cdots & 1/n_J \end{bmatrix}.$$

Применив обобщенный метод наименьших квадратов, получим оценку

$$\widehat{\beta}_{\mathrm{GLS}} = \frac{\sum_{i=1}^{J} n_i \overline{x}_i \overline{y}_i}{\sum_{i=1}^{J} n_i \overline{x}_i^2}.$$

Вычислим ее среднее и дисперсию.

$$\mathrm{E}\widehat{\beta}_{\mathrm{GLS}} = \frac{\sum_{i=1}^{J} n_i \overline{x}_i \mathrm{E}\overline{y}_i}{\sum_{i=1}^{J} n_i \overline{x}_i^2} = \beta + \frac{\sum_{i=1}^{J} n_i \overline{x}_i \mathrm{E}\overline{\varepsilon}_i}{\sum_{i=1}^{J} n_i \overline{x}_i^2} = \beta.$$

Таким образом, оценка $\widehat{\beta}_{\mathrm{GLS}}$ является несмещенной.

$$\mathrm{V}(\widehat{\beta}_{\mathrm{GLS}}) = \frac{\sum_{i=1}^{J} n_i^2 \overline{x}_i^2 \mathrm{V}(\overline{y}_i)}{\left(\sum_{i=1}^{J} n_i \overline{x}_i^2\right)^2} = \frac{\sum_{i=1}^{J} n_i^2 \overline{x}_i^2 \frac{\sigma^2}{n_i}}{\left(\sum_{i=1}^{J} n_i \overline{x}_i^2\right)^2} = \frac{\sigma^2}{\sum_{i=1}^{J} n_i \overline{x}_i^2}.$$

Имея доступ к исходным данным, мы можем получить оценку

$$\widehat{\beta}_{\mathrm{OLS}} = \frac{\sum_{t=1}^{n} x_t y_t}{\sum_{t=1}^{n} x_t^2}.$$

Эта оценка также является несмещенной ($\mathrm{E}\widehat{\beta}_{\mathrm{OLS}} = \beta$) и ее дисперсия равна

$$\mathrm{V}(\widehat{\beta}_{\mathrm{OLS}}) = \frac{\sigma^2}{\sum\limits_{t=1}^{n} x_t^2}.$$

Используя равенство

$$\sum_{t=N_i+1}^{N_{i+1}} x_t^2 - n_i \overline{x}_i^2 = \sum_{t=N_i+1}^{N_{i+1}} (x_t - \overline{x}_i)^2 \geqslant 0,$$

получаем, что

$$\mathrm{V}(\widehat{\beta}_{\mathrm{GLS}}) = \frac{\sigma^2}{\sum\limits_{i=1}^{J} n_i \overline{x}_i^2} \geqslant \frac{\sigma^2}{\sum\limits_{i=1}^{J} \sum\limits_{t=N_i+1}^{N_{i+1}} x_t^2} = \frac{\sigma^2}{\sum\limits_{t=1}^{n} x_t^2} = \mathrm{V}(\widehat{\beta}_{\mathrm{OLS}}).$$

Таким образом, мы убедились, что действительно произошла потеря эффективности.

Отношение дисперсий оценок можно представить в следующем виде:

$$\frac{\mathrm{V}(\widehat{\beta}_{\mathrm{OLS}})}{\mathrm{V}(\widehat{\beta}_{\mathrm{GLS}})} = \frac{\sigma^2}{\sum\limits_{t=1}^{n} x_t^2} \Bigg/ \frac{\sigma^2}{\sum\limits_{i=1}^{J} n_i \overline{x}_i^2} = \frac{\sum\limits_{i=1}^{J} n_i \overline{x}_i^2}{\sum\limits_{t=1}^{n} x_t^2} = \frac{\sum\limits_{i=1}^{J} n_i \overline{x}_i^2 - \sum\limits_{t=1}^{n} x_t^2 + \sum\limits_{t=1}^{n} x_t^2}{\sum\limits_{t=1}^{n} x_t^2}$$

$$= 1 - \frac{\sum\limits_{t=1}^{n} x_t^2 - \sum\limits_{i=1}^{J} n_i \overline{x}_i^2}{\sum\limits_{t=1}^{n} x_t^2} = 1 - \frac{\sum\limits_{i=1}^{J} \sum\limits_{t=N_i+1}^{N_{i+1}} (x_t - \overline{x}_i)^2}{\sum\limits_{t=1}^{n} x_t^2}.$$

Отсюда следует, что потери эффективности не происходит только в том случае, когда в каждой группе все наблюдения совпадают.

## Задача 6.3

Рассмотрим уравнение $y_t = \alpha + \beta x_t + \varepsilon_t$, где ошибки $\varepsilon_t$ порождаются авторегрессионным процессом второго порядка:

$$\varepsilon_t = \rho_1 \varepsilon_{t-1} + \rho_2 \varepsilon_{t-2} + u_t.$$

Предложите обобщение итеративной процедуры Кохрейна–Оркатта для оценивания параметров этой модели.

## Решение

Может быть предложена следующая версия итеративной процедуры Кохрейна–Оркатта для оценивания параметров указанной модели:

1) Применяется МНК к исходной системе

$$y_t = \alpha + \beta x_t + \varepsilon_t$$

и находится вектор остатков $\boldsymbol{e} = (e_1, \ldots, e_n)'$.

2) В качестве приближенного значения вектора $\boldsymbol{\rho} = (\rho_1, \rho_2)'$ берется его МНК-оценка $\widehat{\rho}_{\text{OLS}} = (\widehat{\rho}_1, \widehat{\rho}_2)'$ в модели

$$e_t = \rho_1 e_{t-1} + \rho_2 e_{t-2} + u_t.$$

3) Рассматривается преобразованное уравнение

$$y_t - \rho_1 y_{t-1} - \rho_2 y_{t-2} = \alpha(1 - \rho_1 - \rho_2) + \beta(x_t - \rho_1 x_{t-1} - \rho_2 x_{t-2}) + u_t.$$

Здесь ошибки уже удовлетворяют условиям классической регрессионной модели. После подстановки в него $\widehat{\rho}_1$ и $\widehat{\rho}_2$ вместо $\rho_1$ и $\rho_2$ соответственно находятся МНК-оценки $\widehat{\alpha}$ и $\widehat{\beta}$.

4) Строится новый вектор остатков $e_t = y_t - \widehat{\alpha} - \widehat{\beta} x_t$.

5) Процедура повторяется, начиная с п. 2), до тех пор, пока не будет достигнута требуемая точность.

## Задача 6.4

Рассмотрим модель $y_t = \beta x_t + \varepsilon_t$, где

$$\varepsilon_t = \rho\varepsilon_{t-1} + u_t, \quad \mathrm{E}\, u_t = 0, \quad \mathrm{E}\, u_t^2 = \sigma^2, \quad \mathrm{E}\, u_t u_s = 0, \qquad t \neq s.$$

Предлагается оценивать параметр $\beta$ с помощью регрессии первой разности $\Delta y_t = y_t - y_{t-1}$ на $\Delta x_t$.

а) Покажите, что эта оценка является линейной и несмещенной.

б) Вычислите дисперсию оценки и покажите, что стандартная оценка этой дисперсии смещена.

## Решение

Будем считать, что $t = 0, \ldots, n$, причем $\varepsilon_0 = u_0$. Тогда, очевидно, $\varepsilon_t = \sum_{s=0}^{t} \rho^{t-s} u_s$. Обозначая

$$\varepsilon_t^* = \varepsilon_t - \varepsilon_{t-1} = (\rho - 1)\varepsilon_{t-1} + u_t,$$

получаем:

$$\Delta y_t = \beta \Delta x_t + \varepsilon_t^* = \beta \Delta x_t + (\rho - 1)\varepsilon_{t-1} + u_t.$$

Для сокращения записей введем обозначения

$$z_t = \Delta y_t = y_t - y_{t-1}, \quad w_t = \Delta x_t = x_t - x_{t-1}, \quad t = 1, \ldots, n,$$

в которых модель в первых разностях записывается так:

$$z_t = \beta w_t + \varepsilon_t^*, \quad t = 1, \ldots, n.$$

Заметим, что $\mathrm{E}(\varepsilon_t^*) = 0$ при любом $t$, так как $\boldsymbol{\varepsilon}^*$ есть линейная функция $\boldsymbol{u}$.

а) Рассмотрим МНК-оценку

$$\widehat{\beta} = \frac{\sum_{t=1}^n w_t z_t}{\sum_{t=1}^n w_t^2}.$$

Очевидно, эта оценка является линейной по $y_t$ и несмещенной:

$$\mathrm{E}(\widehat{\beta}) = \frac{\sum_{t=1}^n w_t \mathrm{E}(z_t)}{\sum_{t=1}^n w_t^2} = \frac{\sum_{t=1}^n w_t \beta w_t}{\sum_{t=1}^n w_t^2} = \beta.$$

б) Введем следующие обозначения. Пусть

$$\boldsymbol{\varepsilon},\ \boldsymbol{u},\ \begin{bmatrix} \varepsilon_0 \\ \boldsymbol{\varepsilon}^* \end{bmatrix} \text{ — векторы размерности } (n+1).$$

(Соответственно $\boldsymbol{\varepsilon}^*$ — вектор размерности $n$.) Нетрудно проверить, что из условия следует, что эти векторы удовлетворяют следующим соотношениям:

$$(\boldsymbol{I} - \rho \boldsymbol{K})\boldsymbol{\varepsilon} = \boldsymbol{u}, \quad \begin{bmatrix} \varepsilon_0 \\ \boldsymbol{\varepsilon}^* \end{bmatrix} = (\boldsymbol{I} - \boldsymbol{K})\boldsymbol{\varepsilon},$$

где через $\boldsymbol{K}$ обозначена $(n+1) \times (n+1)$ матрица, у которой единственные отличные от нуля элементы находятся под главной диагональю:

$$\boldsymbol{K} = \begin{bmatrix} 0 & 0 & 0 & \ldots & 0 \\ 1 & 0 & 0 & \ldots & 0 \\ 0 & 1 & 0 & \ldots & 0 \\ \multicolumn{5}{c}{\dots\dots\dots\dots\dots} \\ 0 & 0 & \ldots & 1 & 0 \end{bmatrix}.$$

Вычислим $\boldsymbol{\Omega}$ — матрицу коварияций вектора $\boldsymbol{\varepsilon}^*$.

$$\begin{bmatrix} \varepsilon_0 \\ \boldsymbol{\varepsilon}^* \end{bmatrix} = (\boldsymbol{I} - \boldsymbol{K})(\boldsymbol{I} - \rho \boldsymbol{K})^{-1}\boldsymbol{u} = \boldsymbol{A}\boldsymbol{u},$$

$$\mathrm{V}\left(\begin{bmatrix} \varepsilon_0 \\ \boldsymbol{\varepsilon}^* \end{bmatrix}\right) = \mathrm{V}(\boldsymbol{A}\boldsymbol{u}) = \sigma^2 \boldsymbol{A}\boldsymbol{A}' = \begin{bmatrix} \omega_{00} & \boldsymbol{\omega}_{10}' \\ \boldsymbol{\omega}_{10} & \boldsymbol{\Omega} \end{bmatrix}.$$

Для доказательства смещенности стандартной оценки дисперсии оценки $\widehat{\beta}$ выпишем соответствующие формулы (см. п. 5.2):

$$\mathrm{V}(\widehat{\beta}) = (w'w)^{-1}w'\Omega w(w'w)^{-1} = \frac{w'\Omega w}{(w'w)^2},$$

$$\widehat{\mathrm{V}}(\widehat{\beta}) = \widehat{\sigma}^2(w'w)^{-1} = \frac{\widehat{\sigma}^2}{w'w},$$

$$\mathrm{E}(\widehat{\mathrm{V}}(\widehat{\beta})) = \frac{1}{n-1}\,\mathrm{tr}(M\Omega)(w'w)^{-1} = \frac{\mathrm{tr}(M\Omega)}{(n-1)w'w},$$

где $M = I - w(w'w)^{-1}w' = I - \frac{1}{w'w}ww'$.

Хотя понятно, что в общем случае оценка смещена, мы все же приведем пример. Пусть $n = 2$ и

$$x = (1\ 0\ 0)', \quad \text{тогда } w = (-1\ 0)', \quad M = \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}.$$

Отсюда получаем $w'w = 1$ и тогда

$$\mathrm{V}(\widehat{\beta}) = \omega_{11}, \qquad \mathrm{E}(\widehat{\mathrm{V}}(\widehat{\beta})) = \omega_{22}.$$

Из полученного выше выражения для $\Omega$ для данного случая получаем:

$$\omega_{11} = 1 + (\rho - 1)^2,$$
$$\omega_{22} = 1 + (\rho - 1)^2 + \rho^2(\rho - 1)^2.$$

Для $\rho \neq 0$ и $\rho \neq 1$ эти два выражения не совпадают, т. е.

$$\mathrm{E}(\widehat{\mathrm{V}}(\widehat{\beta})) \neq \mathrm{V}(\widehat{\beta}).$$

## Задача 6.5

Предположим, что для системы $y_t = \alpha + \beta x_t + \varepsilon_t$, $t = 1, \ldots, n$, выполнены все предположения классической нормальной модели, за одним исключением: дисперсии ошибок удовлетворяют соотношениям

$$\sigma_t^2 = \mu + \delta x_t.$$

Предложите двухшаговую процедуру оценивания параметров $\alpha$ и $\beta$.

## Решение

Рассмотрим следующую процедуру:

1) Оцениваем уравнение $y_t = \alpha + \beta x_t + \varepsilon_t$ с помощью обычного МНК. Получаем остатки регрессии $e_t$.

2) Оцениваем уравнение $e_t^2 = \mu + \delta x_t + u_t$, и берем в качестве оценок дисперсий $\widehat{\sigma}_t^2$ величины $\widehat{\sigma}_t^2 = \widehat{\mu} + \widehat{\delta} x_t$.

3) Применяем метод взвешенных наименьших квадратов к исходной модели $y_t = \alpha + \beta x_t + \varepsilon_t$, т. е. строим регрессию

$$\frac{y_t}{\widehat{\sigma}_t} = \alpha \frac{1}{\widehat{\sigma}_t} + \beta \frac{x_t}{\widehat{\sigma}_t} + \frac{\varepsilon_t}{\widehat{\sigma}_t},$$

по которой оцениваем $\alpha$ и $\beta$ при помощи МНК.

Полученные оценки будут состоятельными.

## Задача 6.6

Рассмотрим модель, связывающую количество вакансий $w_t$ и уровень безработицы $u_t$:

$$\ln w_t = \beta_1 + \beta_2 \ln u_t + \varepsilon_t.$$

Ошибки $\varepsilon_t$ независимы и нормально распределены $N(0, \sigma_\varepsilon^2)$.

а) Используя (искусственные) данные из таблицы 6.1, найдите МНК-оценки параметров $\beta_1$ и $\beta_2$, а также 95%-ный доверительный интервал для $\beta_2$.

б) Вычислите статистику Дарбина–Уотсона. Что ее значение говорит об исходном предположении об ошибках $\varepsilon_t$? Что можно сказать о доверительном интервале, найденном в п. а)?

в) Оцените модель заново, используя модель автокорреляции первого порядка для ошибок регрессии. Найдите 95%-ный доверительный интервал для $\beta_2$. Сравните результат с интервалом, полученным в п. а).

**Таблица 6.1**

| $t$ | $w_t$ | $u_t$ | $t$ | $w_t$ | $u_t$ | $t$ | $w_t$ | $u_t$ |
|---|---|---|---|---|---|---|---|---|
| 1 | 1.73 | 8.65 | 9 | 5.06 | 2.87 | 17 | 3.15 | 4.72 |
| 2 | 1.94 | 4.82 | 10 | 2.81 | 5.29 | 18 | 1.92 | 7.45 |
| 3 | 3.05 | 2.67 | 11 | 4.43 | 3.31 | 19 | 2.26 | 6.21 |
| 4 | 4.17 | 2.67 | 12 | 3.19 | 5.44 | 20 | 6.18 | 2.64 |
| 5 | 2.52 | 2.58 | 13 | 2.23 | 6.80 | 21 | 2.07 | 8.55 |
| 6 | 1.71 | 8.07 | 14 | 2.06 | 8.25 | 22 | 8.39 | 2.60 |
| 7 | 1.95 | 8.83 | 15 | 3.33 | 3.44 | 23 | 2.75 | 6.25 |
| 8 | 2.57 | 5.54 | 16 | 2.12 | 7.80 | 24 | 6.10 | 2.70 |

## Решение

а) Оценка параметров модели с применением эконометрического пакета дает следующий результат:

Dependent Variable: $\ln w$

| Variable | Coefficient | Std. Error | $t$-Statistic | Probability |
|---|---|---|---|---|
| const | 2.2997 | 0.1860 | 12.366 | 0.0000 |
| $\ln u$ | −0.7791 | 0.1133 | −6.8746 | 0.0000 |
| $R^2$ | 0.6824 | | | |
| $DW$ | 1.1026 | | | |

Мы получили $\widehat{\beta}_1 = 2.30$, $\widehat{\beta}_2 = -0.78$. 95%-ный доверительный интервал для $\beta_2$ находим по формуле

$$\beta_2 = \widehat{\beta}_2 \pm t_{0.025}(24-2) \cdot s_{\widehat{\beta}_2} = -0.7791 \pm 2.074 \cdot 0.1133,$$

или $(-1.014, -0.544)$.

б) Статистика Дарбина–Уотсона значительно меньше 2, что может свидетельствовать о наличии положительной автокорреляции первого порядка ошибок регрессии. Это нарушает условие отсутствия автокорреляции ошибок, которое лежит в основе стандартной модели регрессии, и соответственно стандартные ошибки, приведенные в п. а), рассчитаны по неверным формулам. Это может привести к тому, что полученный доверительный интервал будет неверным.

в) Оценка параметров модели с включением в нее модели автокорреляции первого порядка для ошибок регрессии с применением эконометрического пакета дает следующий результат:

Dependent Variable: $\ln w$

| Variable | Coefficient | Std. Error | $t$-Statistic | Probability |
|---|---|---|---|---|
| const | 2.3949 | 0.1717 | 13.9452 | 0.0000 |
| $\ln u$ | −0.8312 | 0.0954 | −8.7144 | 0.0000 |
| AR(1) | 0.4676 | 0.2038 | 2.2944 | 0.0327 |
| $R^2$ | 0.7291 | | | |
| $DW$ | 2.1617 | | | |

Получаем 95%-ный доверительный интервал для $\beta_2$: $(-1.03, -0.633)$, т. е. результат, почти не отличающийся в случае данной конкретной выборки от полученного в п. а). Интервал немного уже. Статистика Дарбина–Уотсона близка к 2, что говорит об отсутствии автокорреляции ошибок в модели, и сама первая оценка меньше второй.

## Задача 6.7

В таблице 6.2 представлены данные о потребительских расходах $C$ и располагаемом доходе $Y_d$ тридцати семей (долл.).

а) Проведите регрессию $C$ на $Y_d$ и проверьте наличие или отсутствие гетероскедастичности.

б) Если в п. а) выявлена гетероскедастичность, осуществите коррекцию на гетероскедастичность.

**Таблица 6.2**

| Потребление | | | Доход |
|---|---|---|---|
| 10700 | 10900 | 11200 | 12000 |
| 11400 | 11700 | 12100 | 13000 |
| 12300 | 12600 | 13200 | 14000 |
| 13000 | 13300 | 13600 | 15000 |
| 13800 | 14000 | 14200 | 16000 |
| 14400 | 14900 | 15300 | 17000 |
| 15000 | 15700 | 16400 | 18000 |
| 15900 | 16500 | 16900 | 19000 |
| 16900 | 17500 | 18100 | 20000 |
| 17200 | 17800 | 18500 | 21000 |

## Решение

а) Оценка параметров модели с применением эконометрического пакета дает следующий результат.

Dependent Variable: $C$

| Variable | Coefficient | Std. Error | $t$-Statistic | Probability |
|---|---|---|---|---|
| const | 1580.0 | 447.56 | 3.530 | 0.0015 |
| $Y$ | 0.7830 | 0.0267 | 29.30 | 0.0000 |
| $R^2$ | 0.9684 | | | |

График зависимости квадратов остатков регрессии от независимой переменной $Y$ (рис. 6.1) имеет вид, заставляющий предполагать гетероскедастичность.

Проведем тест Голдфелда–Куандта. Упорядочим наблюдения по возрастанию $Y$. В регрессиях по первым 12 наблюдениям и последним 12 наблюдениям получаем суммы квадратов остатков 1046000 и 3344000 соответственно. Получаем значение $F$-статистики $F = 3344000/1046000 = 3.20$, большее, чем критическое значение статистики Фишера $F_{0.95}(10, 10) = 2.98$, что указывает на наличие гетероскедастичности.

Проведем тест Бреуша–Пагана. Проведем первую регрессию и построим оценку

$$\widehat{\sigma}^2 = (1/n) \sum e_i^2 = 4949727/30 = 164957.6.$$

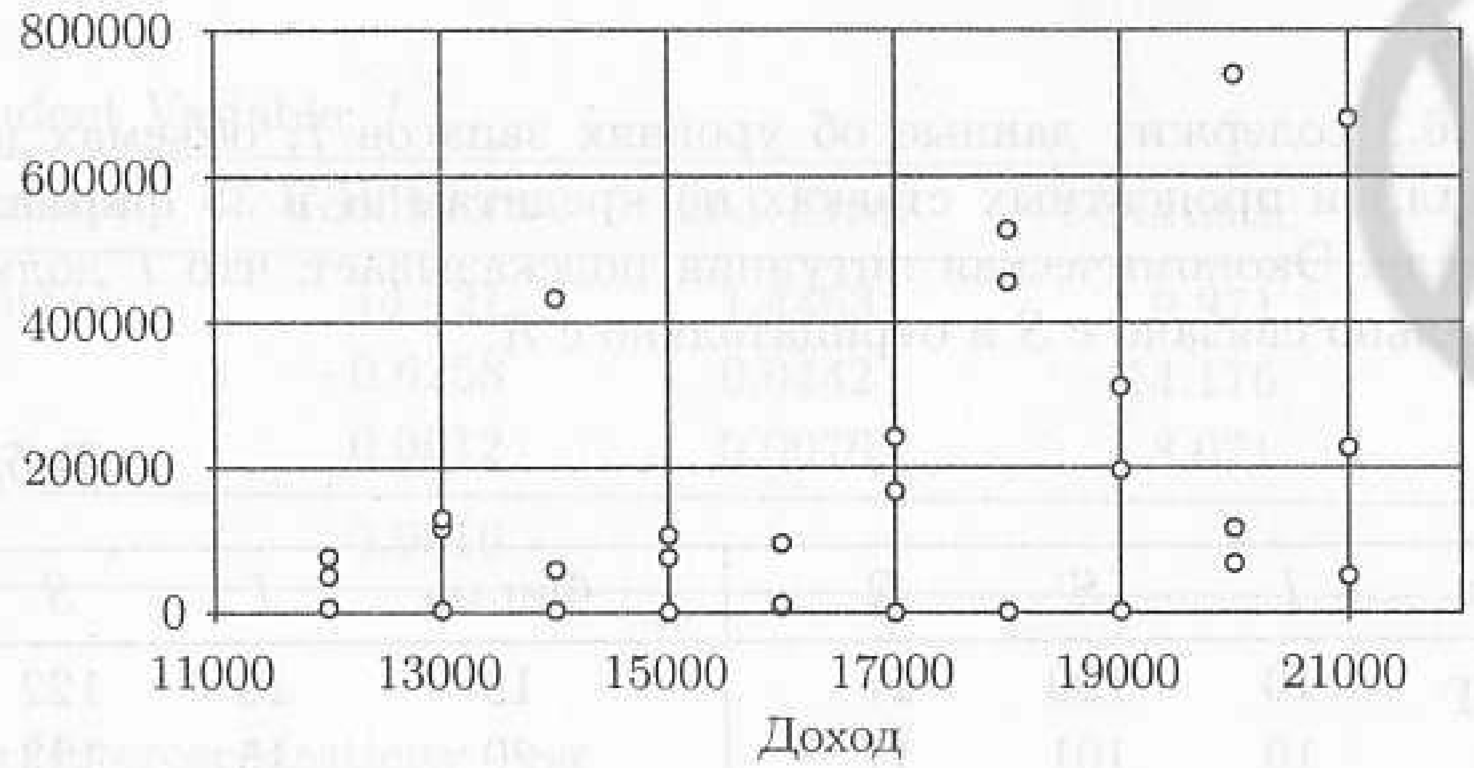



**Рис. 6.1**

Далее оценим регрессию $e_t^2/\widehat{\sigma}^2$ на константу и переменную $Y$. Получим статистику RSS/2 = 4.11. При нулевой гипотезе отсутствия гетероскедастичности эта статистика имеет распределение $\chi^2(1)$, 5%-ная точка которого равна 3.84. Таким образом, гипотеза гомоскедастичности опять отвергается.

б) Проведем двухшаговую процедуру коррекции гетероскедастичности, а именно оценим исходное уравнение при помощи взвешенного метода наименьших квадратов, взяв в качестве весов $1/Y$. Получим следующий результат:

Dependent Variable: $C/Y$

| Variable | Coefficient | Std. Error | $t$-Statistic | Probability |
|---|---|---|---|---|
| $1/Y$ | 1557.0 | 392.28 | 3.969 | 0.0005 |
| const | 0.7844 | 0.0249 | 31.46 | 0.0000 |
| $R^2$ | 0.3600 | | | |
| Unweighted $R^2$ | 0.9684 | | | |

Несмотря на то, что в данном примере оценки почти совпадают, следует помнить, что в случае гетероскедастичности оценки дисперсий МНК-оценок коэффициентов являются смещенными.

Тест Уайта показывает отсутствие гетероскедастичности в последней модели:

White Heteroscedasticity Test

| | | | |
|---|---|---|---|
| F-statistic | 0.5182 | Probability | 0.6014 |
| Obs*R-squared | 1.1090 | Probability | 0.5744 |

## Задача 6.8

Таблица 6.3 содержит данные об уровнях запасов $I$, объемах продаж $S$ (млн. долл.) и процентных ставках по кредитам $R$ в 35 фирмах некоторой отрасли. Экономическая интуиция подсказывает, что $I$ должно быть положительно связано с $S$ и отрицательно с $R$.

Таблица 6.3

| Фирма | $I$ | $S$ | $R$ | Фирма | $I$ | $S$ | $R$ |
|---|---|---|---|---|---|---|---|
| 1 | 10 | 100 | 17 | 19 | 15 | 122 | 11 |
| 2 | 10 | 101 | 17 | 20 | 15 | 123 | 11 |
| 3 | 10 | 103 | 17 | 21 | 15 | 125 | 11 |
| 4 | 11 | 105 | 16 | 22 | 16 | 128 | 10 |
| 5 | 11 | 106 | 16 | 23 | 16 | 128 | 10 |
| 6 | 11 | 106 | 16 | 24 | 16 | 131 | 10 |
| 7 | 12 | 108 | 15 | 25 | 17 | 133 | 10 |
| 8 | 12 | 109 | 15 | 26 | 17 | 134 | 9 |
| 9 | 12 | 111 | 14 | 27 | 17 | 135 | 9 |
| 10 | 12 | 111 | 14 | 28 | 17 | 136 | 9 |
| 11 | 12 | 112 | 14 | 29 | 18 | 139 | 8 |
| 12 | 13 | 113 | 14 | 30 | 18 | 143 | 8 |
| 13 | 13 | 114 | 13 | 31 | 19 | 147 | 8 |
| 14 | 13 | 114 | 13 | 32 | 19 | 151 | 8 |
| 15 | 14 | 116 | 12 | 33 | 19 | 157 | 8 |
| 16 | 14 | 117 | 12 | 34 | 20 | 163 | 7 |
| 17 | 14 | 118 | 12 | 35 | 20 | 171 | 7 |
| 18 | 15 | 120 | 11 | | | | |

а) Проведите регрессию $I$ на $S$ и $R$ и тест на гетероскедастичность.

б) Если в п. а) выявлена гетероскедастичность, осуществите коррекцию на гетероскедастичность, предполагая, что дисперсия ошибки пропорциональна $S^2$.

## Решение

а) Оценка параметров модели с применением эконометрического пакета дает следующий результат (см. таблицу 6.4).

Проведем тест Уайта на гетероскедастичность. Получим результат, представленный в таблице 6.5.

Результат теста указывает на гетероскедастичность, если принять 10%-ный уровень значимости теста.

**Таблица 6.4**

Dependent Variable: $I$

| Variable | Coefficient | Std. Error | $t$-Statistic | Probability |
|---|---|---|---|---|
| const | 14.421 | 1.4463 | 9.971 | 0.0000 |
| $R$ | −0.6258 | 0.0442 | −14.176 | 0.0000 |
| $S$ | 0.0612 | 0.0076 | 8.021 | 0.0000 |
| $R^2$ | 0.9916 | | | |

**Таблица 6.5**

White Heteroscedasticity Test

| | | | |
|---|---|---|---|
| F-statistic | 2.2408 | Probability | 0.0882 |
| Obs*R-squared | 8.0515 | Probability | 0.0897 |

б) Проведем оценку исходного уравнения взвешенным методом наименьших квадратов с весами $1/S_t$ (см. таблицу 6.6).

**Таблица 6.6**

Dependent Variable: $I$
Weighting series: $1/S$

| Variable | Coefficient | Std. Error | $t$-Statistic | Probability |
|---|---|---|---|---|
| const | 13.469 | 1.5990 | 8.423 | 0.0000 |
| $R$ | −0.5987 | 0.0458 | −13.062 | 0.0000 |
| $S$ | 0.0663 | 0.0087 | 7.599 | 0.0000 |
| $R^2$ | 0.9447 | | | |
| *Unweighted* $R^2$ | 0.9915 | | | |

Проведем тест Уайта на гетероскедастичность (см. таблицу 6.7).

**Таблица 6.7**

White Heteroscedasticity Test

| | | | |
|---|---|---|---|
| F-statistic | 2.4986 | Probability | 0.0636 |
| Obs*R-squared | 8.7464 | Probability | 0.0678 |

Результат теста по-прежнему указывает на гетероскедастичность при 10%-ном уровне значимости теста.

Попробуем более сложную (экспоненциальную) модель гетероскедастичности:

$$\sigma_t^2 = \exp(\gamma_0 + \gamma_1 R_t + \gamma_2 S_t).$$

Оценим регрессию логарифма квадратов остатков $e_t^2$ исходного уравнения на $\ln R_t$ и $\ln S_t$ (см. таблицу 6.8).

**Таблица 6.8**

Dependent Variable: $\ln e^2$
Weighting series: $1/S$

| Variable | Coefficient | Std. Error | $t$-Statistic | Probability |
|---|---|---|---|---|
| const | −241.94 | 121.66 | −1.989 | 0.0553 |
| $\ln R$ | 21.46 | 10.36 | 2.072 | 0.0464 |
| $\ln S$ | 38.43 | 20.14 | 1.909 | 0.0653 |
| $R^2$ | 0.1263 | | | |

Коэффициенты значимы на 10%-ном уровне. Оценим теперь исходное уравнение взвешенным методом наименьших квадратов с весовыми коэффициентами, равными $1/\exp(\widehat{\ln e^2})^{1/2}$, где $\widehat{\ln e^2}$ — прогнозные значения последней вспомогательной регрессии (см. таблицу 6.9).

**Таблица 6.9**

Dependent Variable: $I$
Weighting series: $1/\exp(\widehat{\ln e^2})^{1/2}$

| Variable | Coefficient | Std. Error | $t$-Statistic | Probability |
|---|---|---|---|---|
| const | 13.805 | 1.6119 | 8.564 | 0.0000 |
| $R$ | −0.6021 | 0.0491 | −12.640 | 0.0000 |
| $S$ | 0.0656 | 0.0088 | 7.476 | 0.0000 |
| $R^2$ | 0.9997 | | | |
| *Unweighted* $R^2$ | 0.9911 | | | |

Проведем тест Уайта на гетероскедастичность (см. таблицу 6.10).

**Таблица 6.10**

White Heteroscedasticity Test

| | | | |
|---|---|---|---|
| F-statistic | 1.4633 | Probability | 0.2380 |
| Obs*R-squared | 5.7138 | Probability | 0.2216 |

Результат теста на этот раз указывает на отсутствие гетероскедастичности на 10%-ном уровне значимости.

## Задача 6.9

В таблице 6.11 приведены данные об объеме импорта $M$ и ВНП США (млрд. долл.) за период с 1960 по 1979 г.

а) Проведите регрессию $M$ на ВНП и на 5%-ном уровне значимости протестируйте гипотезу об отсутствии автокорреляции ошибок.

б) Если в п. а) гипотеза отвергается, проведите коррекцию на автокорреляцию.

**Таблица 6.11**

| Год | $M$ | ВНП | Год | $M$ | ВНП |
|---|---|---|---|---|---|
| 1960 | 23.2 | 506.0 | 1970 | 58.5 | 982.4 |
| 1961 | 23.1 | 523.3 | 1971 | 64.0 | 1063.4 |
| 1962 | 25.2 | 563.8 | 1972 | 75.9 | 1171.1 |
| 1963 | 26.4 | 594.7 | 1973 | 94.4 | 1306.6 |
| 1964 | 28.4 | 635.7 | 1974 | 131.9 | 1424.9 |
| 1965 | 32.0 | 688.1 | 1975 | 126.9 | 1528.1 |
| 1966 | 37.7 | 753.0 | 1976 | 155.4 | 1702.2 |
| 1967 | 40.6 | 796.3 | 1977 | 185.8 | 1899.5 |
| 1968 | 47.7 | 868.5 | 1978 | 217.5 | 2127.6 |
| 1969 | 52.9 | 935.5 | 1979 | 260.9 | 2368.5 |

Источник: *D.Salvatore. Statistics and Econometrics*, McGraw-Hill, 1982.

## Решение

а) Построив графики $M$ и ВНП, мы видим особенное поведение $M$ в 1974 г. (Наиболее вероятное объяснение — энергетический кризис.) Это дает нам основание включить в регрессию фиктивную переменную $d1974$, равную 1 в 1974 г. и 0 в остальные годы. Таким образом, мы оцениваем регрессию:

$$m_t = \beta_1 + \beta_2 \text{ВНП}_t + \beta_3 d1974_t + \varepsilon_t.$$

Результаты МНК-оценивания приведены ниже (таблица 6.12).

**Таблица 6.12**

Dependent Variable: $m$

| Variable | Coefficient | Std. Error | $t$-Statistic | Probability |
|---|---|---|---|---|
| const | −56.0818 | 5.4556 | −10.280 | 0.0000 |
| ВНП | 0.1257 | 0.0044 | 28.501 | 0.0000 |
| $d1974$ | 8.8341 | 11.008 | 0.8025 | 0.4333 |
| $R^2$ | 0.9800 | | | |
| $DW$ | 0.3240 | | | |

Полученное значение статистики Дарбина–Уотсона DW = 0.32 значительно ниже 5% нижней границы, примерно равной 0.97. Это говорит о наличии положительной автокорреляции ошибок регрессии.

б) При наличии автокорреляции ошибок оценки дисперсий коэффициентов регрессии, вообще говоря, занижены. Поэтому заново оценим регрессию с поправкой на автокорреляцию ошибок первого порядка. Получим следующий результат (см. таблицу 6.13).

**Таблица 6.13**

Dependent Variable: $m$

| Variable | Coefficient | Std. Error | $t$-Statistic | Probability |
|---|---|---|---|---|
| const | 2869.2 | 68657. | 0.0418 | 0.9672 |
| ВНП | 0.1835 | 0.0211 | 8.7101 | 0.0000 |
| $d1974$ | 19.859 | 1.7447 | 11.382 | 0.0000 |
| AR(1) | 1.0018 | 0.0431 | 23.220 | 0.0000 |
| $R^2$ | 0.9990 | | | |
| $DW$ | 2.1127 | | | |

Теперь статистика DW близка к 2, однако эту величину невозможно интерпретировать, так как в EViews модель оценивается с помощью нелинейного МНК для уравнения, содержащего в правой части лагированное значение зависимой переменной. Статистика Льюнга–Бокса для остатков этой регрессии не позволяет отвергнуть нулевую гипотезу отсутствия автокорреляции ошибок в оцениваемом уравнении. Коэффициент при ВНП (оценивание которого и было нашей задачей) теперь имеет стандартное отклонение больше, чем его слишком оптимистичная оценка в первой регрессии.

Коэффициент $\rho$, однако, близок к 1. Это говорит о возможном наличии единичного корня. Связанная с этим проблема подробно обсуждается в гл. 11. Сейчас заметим только, что в данной ситуации было бы более правильно оценивать уравнение в первых разностях:

$$\Delta m_t = \beta_1 + \beta_2 \Delta \text{ВНП}_t + \beta_3 d1974_t + \varepsilon_t.$$

## Задача 6.10

Рассмотрим модель $y_t = \beta x_t + \varepsilon_t$, $t = 1, 2, \ldots, n$, где $\beta$, $x_t$ — скаляры, и предположим, что

$$\lim_{n\to\infty} (1/n) \sum_{t=1}^{n} x_t^2 = \sigma_x^2, \quad \lim_{n\to\infty} (1/n) \sum_{t=s+1}^{n} x_t x_{t-s} = \alpha^s \sigma_x^2,$$

где $|\alpha| \leqslant 1$ и $\mathrm{E}(\varepsilon_t) = 0$, $\mathrm{E}(\varepsilon_t \varepsilon_{t-s}) = \rho^s \sigma_\varepsilon^2$ ($|\rho| < 1$) при всех $s$.

а) Вычислите асимптотическую дисперсию МНК-оценки параметра $\beta$, т. е. $\lim_{n\to\infty} n\mathrm{V}(\widehat{\beta}_{\mathrm{OLS}})$.

б) Вычислите асимптотическую дисперсию оценки параметра $\beta$, полученной с помощью обобщенного метода наименьших квадратов, и покажите, что при $\alpha = \rho$ асимптотическая эффективность МНК-оценки равна $(1-\rho^2)/(1+\rho^2)$.

## Решение

а) Имеем

$$\widehat{\beta}_{\text{OLS}} = \frac{\sum_{t=1}^{n} x_t y_t}{\sum_{t=1}^{n} x_t^2} = \beta + \frac{\sum_{t=1}^{n} x_t\varepsilon_t}{\sum_{t=1}^{n} x_t^2},$$

$$n\mathrm{V}(\widehat{\beta}_{\text{OLS}}) = n\frac{\mathrm{E}\left(\sum_{t=1}^{n} x_t\varepsilon_t\right)^2}{\left(\sum_{t=1}^{n} x_t^2\right)^2} = \frac{\frac{1}{n}\mathrm{E}\left(\sum_{t=1}^{n} x_t\varepsilon_t\right)^2}{\left(\frac{1}{n}\sum_{t=1}^{n} x_t^2\right)^2}.$$

Числитель последней дроби может быть представлен следующим образом:

$$\frac{1}{n}\mathrm{E}\Big(\sum_{t=1}^{n} x_t\varepsilon_t\Big)^2 = \sigma_\varepsilon^2\frac{1}{n}\Big(\sum_{t=1}^{n} x_t^2 + 2\rho\sum_{t=2}^{n} x_t x_{t-1} + 2\rho^2\sum_{t=3}^{n} x_t x_{t-2} + \ldots$$

$$+\, 2\rho^s\sum_{t=s+1}^{n} x_t x_{t-s} + \ldots + 2\rho^{n-1}x_1x_n\Big).$$

Из леммы (см. ниже) следует, что при $n\to\infty$ это выражение сходится к

$$\sigma_\varepsilon^2\sigma_x^2(1 + 2\rho\alpha + 2\rho^2\alpha^2 + \ldots) = \sigma_\varepsilon^2\sigma_x^2\left(1 + \frac{2\rho\alpha}{1-\rho\alpha}\right) = \sigma_\varepsilon^2\sigma_x^2\frac{1+\rho\alpha}{1-\rho\alpha}.$$

Отсюда получаем:

$$\lim_{n\to\infty} n\mathrm{V}(\widehat{\beta}_{\text{OLS}}) = \frac{\lim\limits_{n\to\infty}\dfrac{1}{n}\Big(\sum\limits_{t=1}^{n} x_t\varepsilon_t\Big)^2}{\lim\limits_{n\to\infty}\Big(\dfrac{1}{n}\sum\limits_{t=1}^{n} x_t\varepsilon_t\Big)^2} = \frac{\sigma_\varepsilon^2\sigma_x^2\dfrac{1+\rho\alpha}{1-\rho\alpha}}{(\sigma_x^2)^2} = \frac{\sigma_\varepsilon^2}{\sigma_x^2}\frac{1+\rho\alpha}{1-\rho\alpha}.$$

*Лемма.* Пусть $|\rho| < 1$, $|\alpha| \leqslant 1$ и

$$\lim_{n\to\infty}\frac{1}{n}\sum_{t=s+1}^{n} x_t x_{t-s} = \alpha^s\sigma_x^2, \quad s\geqslant 0.$$

Тогда существует предел

$$\lim_{n\to\infty}\frac{1}{n}\Big(\rho\sum_{t=2}^{n} x_t x_{t-1} + \rho^2\sum_{t=3}^{n} x_t x_{t-2} + \ldots + \rho^s\sum_{t=s+1}^{n} x_t x_{t-s} + \ldots\Big)$$

$$= \sigma_x^2\frac{\rho\alpha}{1-\rho\alpha}.$$

*Доказательство.* Обозначим

$$f_{n,s} = \begin{cases} \frac{1}{n}\sum_{t=s+1}^{n} x_t x_{t-s}, & 0\leqslant s\leqslant n-1,\\ 0, & s\geqslant n\end{cases}$$

и
$$S_n = \sum_{s=1}^{n-1} \rho^s f_{n,s} = \sum_{s=1}^{\infty} \rho^s f_{n,s}.$$

В этих обозначениях утверждение леммы записывается следующим образом:
$$\lim_{n\to\infty} S_n = \sigma_x^2 \frac{\rho\alpha}{1-\rho\alpha}.$$

В силу неравенства Коши–Буняковского имеем:
$$|f_{n,s}| = \frac{1}{n}\Big|\sum_{t=s+1}^{n} x_t x_{t-s}\Big| \leqslant \frac{1}{n}\Big(\sum_{t=s+1}^{n} x_t^2\Big)^{1/2}\Big(\sum_{t=s+1}^{n} x_{t-s}^2\Big)^{1/2} \leqslant \frac{1}{n}\sum_{t=1}^{n} x_t^2 = f_{n,0}.$$

По условию леммы $\lim_{n\to\infty} f_{n,s} = \alpha^s \sigma_x^2$, в частности, существует предел $\lim_{n\to\infty} f_{n,0} = \sigma_x^2$, поэтому последовательность $f_{n,0}$ ограничена: существует число $M$ такое, что для любого $n$ выполняется $f_{n,0} < M$. Получаем тогда, что все $f_{n,s}$ ограничены:
$$|f_{n,s}| \leqslant f_{n,0} < M.$$

Зафиксируем теперь некоторое $\varepsilon > 0$. Из того, что $|\rho| < 1$, следует, что существует $n_0$ такое, что $|\rho|^{n_0} < \varepsilon$.

Разобьем $S_n$ на два слагаемых:
$$S_n = \sum_{s=1}^{n_0-1} \rho^s f_{n,s} + \sum_{s=n_0}^{\infty} \rho^s f_{n,s}.$$

Второе слагаемое ограничено по модулю:
$$\Big|\sum_{s=n_0}^{\infty} \rho^s f_{n,s}\Big| \leqslant \sum_{s=n_0}^{\infty} |\rho|^s M = M\frac{|\rho|^{n_0}}{1-|\rho|} < \frac{M}{1-|\rho|}\varepsilon.$$

Запишем правую часть предела из условия леммы как
$$\sigma_x^2 \frac{\rho\alpha}{1-\rho\alpha} = \sum_{s=1}^{\infty} \sigma_x^2 \rho^{s-1}\alpha^{s-1}.$$

Получаем:
$$\begin{aligned}\Big|S_n - \sigma_x^2\frac{\rho\alpha}{1-\rho\alpha}\Big| &= \Big|\sum_{s=1}^{n_0-1} \rho^s (f_{n,s} - \alpha^s\sigma_x^2) + \sum_{s=n_0}^{\infty} \rho^s (f_{n,s} - \alpha^s\sigma_x^2)\Big| \\ &\leqslant \sum_{s=1}^{n_0-1} |\rho^s (f_{n,s} - \alpha^s\sigma_x^2)| + \sum_{s=n_0}^{\infty} \rho^s |f_{n.s}| + \sum_{s=n_0}^{\infty} \rho^s\alpha^s\sigma_x^2 \\ &\leqslant \sum_{s=1}^{n_0-1} |\rho^s (f_{n,s} - \alpha^s\sigma_x^2)| + \frac{M}{1-|\rho|}\varepsilon + \sigma_x^2\frac{|\alpha|^{n_0}}{1-|\rho\alpha|}\varepsilon.\end{aligned}$$

Поскольку $|\alpha| \leqslant 1$, то

$$\sigma_x^2 \frac{|\alpha|^{n_0}}{1-|\rho\alpha|}\varepsilon \leqslant \frac{\sigma_x^2}{1-|\rho\alpha|}\varepsilon.$$

Из того, что существуют пределы $\lim_{n\to\infty} f_{n,s} = \alpha^s\sigma_x^2$, следует, что существует $n_1 = n_1(\varepsilon)$ такое, что при всех $n > n_1$ и $1 \leqslant s \leqslant n_0 - 1$ выполняется неравенство $|f_{n,s} - \alpha^s\sigma_x^2| < \varepsilon$. Поэтому при $n > n_1$

$$\begin{aligned}\Big|S_n - \sigma_x^2\frac{\rho\alpha}{1-\rho\alpha}\Big| &< \sum_{s=1}^{n_0-1}|\rho|^s\varepsilon + \frac{M}{1-|\rho|}\varepsilon + \frac{\sigma_x^2}{1-|\rho\alpha|}\varepsilon \\ &\leqslant \Big(\frac{1}{1-|\rho|} + \frac{M}{1-|\rho|} + \frac{\sigma_x^2}{1-|\rho\alpha|}\Big)\varepsilon.\end{aligned}$$

Последнее выражение стремится к 0 при $\varepsilon \to 0$. Лемма доказана.

б) Рассмотрим случайные величины $u_t = \varepsilon_t - \rho\varepsilon_{t-1}$. Математическое ожидание и дисперсия $u_t$ равны:

$$\begin{aligned}\mathrm{E}(u_t) &= \mathrm{E}(\varepsilon_t) - \rho\mathrm{E}(\varepsilon_{t-1}) = 0,\\ \mathrm{V}(u_t) &= \mathrm{E}(u_t^2) = \mathrm{E}(\varepsilon_t - \rho\varepsilon_{t-1})^2 = \mathrm{E}(\varepsilon_t^2) - 2\rho\mathrm{E}(\varepsilon_t\varepsilon_{t-1}) + \rho^2\mathrm{E}(\varepsilon_{t-1}^2)\\ &= \sigma_\varepsilon^2 - 2\rho\rho\sigma_\varepsilon^2 + \rho^2\sigma_\varepsilon^2 = (1-\rho^2)\sigma_\varepsilon^2.\end{aligned}$$

Кроме того, $u_t$ и $u_s$ некоррелированы при $t \neq s$. Докажем это. Пусть для определенности $t > s$. Тогда, несколько раз используя условие $\mathrm{E}(\varepsilon_t\varepsilon_s) = \rho^{|t-s|}\sigma_\varepsilon^2$, получаем:

$$\begin{aligned}\mathrm{E}(u_t u_s) &= \mathrm{E}(\varepsilon_t - \rho\varepsilon_{t-1})(\varepsilon_s - \rho\varepsilon_{s-1})\\ &= \mathrm{E}(\varepsilon_t\varepsilon_s - \rho\varepsilon_{t-1}\varepsilon_s - \rho\varepsilon_t\varepsilon_{s-1} + \rho^2\varepsilon_{t-1}\varepsilon_{s-1})\\ &= \sigma_\varepsilon^2(\rho^{t-s} - \rho\rho^{t-s-1} - \rho\rho^{t-s+1} + \rho^2\rho^{t-s})\\ &= \sigma_\varepsilon^2(\rho^{t-s} - \rho^{t-s} - \rho^{t-s+2} + \rho^{t-s+2}) = 0.\end{aligned}$$

Применим к исходной системе преобразование (6.11)

$$y_t - \rho y_{t-1} = \beta(x_t - \rho x_{t-1}) + \varepsilon_t - \rho\varepsilon_{t-1}, \quad t = 2,\ldots,n.$$

В преобразованной системе ошибки $u_t = \varepsilon_t - \rho\varepsilon_{t-1}$ удовлетворяют условиям классической регрессионной модели. Оценка метода наименьших квадратов для преобразованной модели асимптотически эквивалентна оценке обобщенного метода наименьших квадратов для исходной модели. (Совпадение не полное, так как мы здесь не принимаем во внимание уравнение (6.12) для $t = 1$.) Будем, однако, с этой оговоркой обозначать МНК-оценку для преобразованной модели как $\beta_{\mathrm{GLS}}$. Получаем:

$$\beta_{\mathrm{GLS}} = \frac{\sum_{t=2}^n (x_t - \rho x_{t-1})(y_t - \rho y_{t-1})}{\sum_{t=2}^n (x_t - \rho x_{t-1})^2}.$$

Как показано выше, слагаемые в числителе некоррелированы, поэтому

$$\mathrm{V}(\beta_{\mathrm{GLS}}) = \frac{\sum_{t=2}^{n}(x_t - \rho x_{t-1})^2(1-\rho^2)\sigma_\varepsilon^2}{\left(\sum_{t=2}^{n}(x_t - \rho x_{t-1})^2\right)^2} = \frac{(1-\rho^2)\sigma_\varepsilon^2}{\sum_{t=2}^{n}(x_t - \rho x_{t-1})^2},$$

и тогда

$$\lim_{n\to\infty} n\mathrm{V}(\beta_{\mathrm{GLS}}) = \lim_{n\to\infty} \frac{(1-\rho^2)\sigma_\varepsilon^2}{\frac{1}{n}\sum_{t=2}^{n}(x_t - \rho x_{t-1})^2} = \frac{(1-\rho^2)\sigma_\varepsilon^2}{\sigma_x^2 - 2\rho\alpha\sigma_x^2 + \rho^2\sigma_x^2}.$$

При $\rho = \alpha$

$$\lim_{n\to\infty} n\mathrm{V}(\beta_{\mathrm{GLS}}) = \frac{\sigma_\varepsilon^2}{\sigma_x^2}.$$

Используя результат, полученный в п. а), получаем, что асимптотическая эффективность МНК-оценки равна

$$\lim_{n\to\infty} \frac{\mathrm{V}(\widehat{\beta}_{\mathrm{GLS}})}{\mathrm{V}(\widehat{\beta}_{\mathrm{OLS}})} = \frac{1-\rho^2}{1+\rho^2}.$$

## Задача 6.11

Рассмотрим модель $y_t = \beta x_t + \varepsilon_t$, $t = 1, \ldots, n$, где $\mathrm{E}(\varepsilon_t) = 0$, $\mathrm{E}(\varepsilon_t^2) = a x_t^2$, $\mathrm{E}(\varepsilon_t \varepsilon_s) = 0$ при $t \neq s$ и $\sum_{t=1}^{n} x_t^2 = n$.

а) Покажите, что МНК-оценка $\widehat{\beta}$ параметра $\beta$ является несмещенной, но неэффективной.

б) Покажите, что стандартная оценка дисперсии $\widehat{\beta}$ смещена вниз по отношению к истинной дисперсии $\widehat{\beta}$.

## Решение

а) МНК-оценка параметра $\beta$ равна:

$$\widehat{\beta} = \frac{\sum_{t=1}^{n} x_t y_t}{\sum_{t=1}^{n} x_t^2} = \frac{\sum_{t=1}^{n} x_t y_t}{n}.$$

Так как $\mathrm{E}(y_t) = \beta x_t$, $\mathrm{V}(y_t) = \mathrm{V}(\varepsilon_t) = a x_t^2$, $t = 1, \ldots, n$, а $y_t$ и $y_s$ некоррелированы при $t \neq s$, то $\mathrm{E}(\widehat{\beta}) = \beta$ (оценка несмещенная) и

$$\mathrm{V}(\widehat{\beta}) = \frac{a\sum_{t=1}^{n} x_t^4}{n^2}.$$

Применяя обобщенный метод наименьших квадратов, получаем оценку

$$\widehat{\beta}_{\mathrm{GLS}} = \frac{1}{n}\sum_{t=1}^{n} \frac{y_t}{x_t},$$

для которой $\mathrm{E}(\widehat{\beta}_{\mathrm{GLS}}) = \beta$ и $\mathrm{V}(\widehat{\beta}_{\mathrm{GLS}}) = a/n$. Из неравенства Коши–Буняковского следует, что $\left(\sum_{t=1}^n x_t^2\right)^2 \leqslant n\sum_{t=1}^n x_t^4$, т.е. $\sum_{t=1}^n x_t^4 \geqslant n$. Поэтому (как и следовало ожидать) $\mathrm{V}(\widehat{\beta}) \geqslant \mathrm{V}(\widehat{\beta}_{\mathrm{GLS}})$, т.е. оценка $\widehat{\beta}$ неэффективна.

б) Как известно, стандартной оценкой дисперсии $\widehat{\beta}$ является величина

$$s_{\widehat{\beta}}^2 = \frac{\widehat{\sigma}^2}{\sum_{t=1}^n x_t^2} = \frac{\widehat{\sigma}^2}{n}, \quad \text{где} \quad \widehat{\sigma}^2 = \frac{\sum_{t=1}^n e_t^2}{n-1}.$$

(Здесь $e_t = y_t - \widehat{\beta}x_t$, $t = 1, \ldots, n$ — остатки регрессии).

Имеем

$$\mathrm{ESS} = \sum_{t=1}^n e_t^2 = \sum_{t=1}^n e_t(y_t - \widehat{\beta}x_t) = \sum_{t=1}^n (y_t - \widehat{\beta}x_t)y_t = \sum_{t=1}^n y_t^2 - \frac{\left(\sum_{t=1}^n x_t y_t\right)^2}{\sum_{t=1}^n x_t^2}.$$

Заметим, что

$$\mathrm{E}(y_t^2) = \beta^2 x_t^2 + a x_t^2 \quad \text{и} \quad \mathrm{E}(y_t y_s) = \beta^2 x_t x_s \quad \text{при} \quad t \neq s.$$

Получаем

$$\begin{aligned}
\mathrm{E}(\mathrm{ESS}) &= (\beta^2 + a)n - \frac{\sum_{t=1}^n x_t^4(\beta^2 + a) + \sum_{t\neq s}\beta^2 x_t^2 x_s^2}{n} \\
&= (\beta^2 + a)n - \frac{\beta^2\left(\sum_{t=1}^n x_t^4 + \sum_{t\neq s} x_t^2 x_s^2\right) + a\sum_{t=1}^n x_t^4}{n} \\
&= (\beta^2 + a)n - \frac{\beta^2\left(\sum_{t=1}^n x_t^2\right)^2 + a\sum_{t=1}^n x_t^4}{n} \\
&= a\left(n - \frac{\sum_{t=1}^n x_t^4}{n}\right).
\end{aligned}$$

Таким образом,

$$\begin{aligned}
\mathrm{E}(\widehat{\sigma}^2) &= \frac{\mathrm{E}(\mathrm{ESS})}{n-1} = \frac{a}{n-1}\left(n - \frac{\sum_{t=1}^n x_t^4}{n}\right) \quad \text{и} \\
\mathrm{E}(s_{\widehat{\beta}}^2) &= \frac{a}{n(n-1)}\left(n - \frac{\sum_{t=1}^n x_t^4}{n}\right).
\end{aligned}$$

Как было отмечено в п. а), $\sum_{t=1}^n x_t^4 \geqslant n$, поэтому

$$\mathrm{E}(s_{\widehat{\beta}}^2) \leqslant \frac{a}{n(n-1)}(n-1) = \frac{a}{n}, \quad \text{а} \quad \mathrm{V}(\widehat{\beta}) = \frac{a\sum_{t=1}^n x_t^4}{n^2} \geqslant \frac{a}{n},$$

т.е. оценка $s_{\widehat{\beta}}^2$ смещена вниз.

## Задача 6.12

Уравнение $y_t = \beta_1 + \beta_2 x_t + \varepsilon_t$ оценивается по следующим наблюдениям (см. таблицу 6.14):

**Таблица 6.14**

| | | | | | |
|---|---|---|---|---|---|
| $x$ | 4 | 1 | 5 | 8 | 2 |
| $y$ | 6 | 3 | 12 | 15 | 4 |

Известна функциональная форма гетероскедастичности: $V(\varepsilon_t) = \sigma_t^2 = \sigma^2 x_t^2$. Вычислить оценки обобщенного метода наименьших квадратов параметров $\beta_1$, $\beta_2$ и их стандартные отклонения.

## Решение

Так как в данной модели присутствует только гетероскедастичность, то оценка обобщенного МНК совпадает с оценкой взвешенного МНК (см. (6.2)). Разделив обе части уравнения на $x_t$, получим взвешенное уравнение:

$$\frac{y_t}{x_t} = \beta_1 \frac{1}{x_t} + \beta_2 + u_t, \quad u_t = \frac{\varepsilon_t}{x_t}, \quad \mathrm{E}u_t = 0, \quad \mathrm{V}(u_t) = \sigma^2.$$

Оценка обобщенного МНК в исходном уравнении совпадает с МНК-оценкой $\widehat{\beta}_1$, $\widehat{\beta}_2$ в этом уравнении:

$$\begin{bmatrix} \widehat{\beta}_1 \\ \widehat{\beta}_2 \end{bmatrix} = \widehat{\boldsymbol{\beta}} = (\boldsymbol{X}'\boldsymbol{X})^{-1}\boldsymbol{X}'\boldsymbol{y},$$

где

$$\boldsymbol{X}'\boldsymbol{X} = \begin{bmatrix} \sum \frac{1}{x_t^2} & \sum \frac{1}{x_t} \\ \sum \frac{1}{x_t} & n \end{bmatrix} = \begin{bmatrix} 1.368 & 2.075 \\ 2.075 & 5 \end{bmatrix}, \quad \boldsymbol{X}'\boldsymbol{y} = \begin{bmatrix} \sum \frac{y_t}{x_t^2} \\ \sum \frac{y_t}{x_t} \end{bmatrix} = \begin{bmatrix} 5.089 \\ 10.775 \end{bmatrix}.$$

Таким образом, оценка обобщенного МНК для исходной модели равна

$$\begin{bmatrix} \widehat{\beta}_1 \\ \widehat{\beta}_2 \end{bmatrix} = \begin{bmatrix} 1.368 & 2.075 \\ 2.075 & 5 \end{bmatrix}^{-1} \begin{bmatrix} 5.089 \\ 10.775 \end{bmatrix} = \begin{bmatrix} 1.218 \\ 1.649 \end{bmatrix}.$$

Оценка ковариационной матрицы $\widehat{\mathrm{V}}(\widehat{\boldsymbol{\beta}})$ равна

$$\widehat{\mathrm{V}}(\widehat{\boldsymbol{\beta}}) = \widehat{\sigma}^2(\boldsymbol{X}'\boldsymbol{X})^{-1}.$$

Оценка дисперсии величины $u$ может быть вычислена по формуле

$$\widehat{\sigma}^2 = \frac{1}{n-2}\sum\left(\frac{y_t}{x_t} - \widehat{\beta}_1\frac{1}{x_t} - \widehat{\beta}_2\right)^2 = \frac{0.552808}{3} = 0.18427.$$

Таким образом, оценка ковариационной матрицы равна

$$\widehat{\mathrm{V}}(\widehat{\beta}) = 0.18427 \cdot \begin{bmatrix} 1.368 & 2.075 \\ 2.075 & 5 \end{bmatrix}^{-1} = \begin{bmatrix} 0.363 & -0.151 \\ -0.151 & 0.099 \end{bmatrix}.$$

Искомые стандартные отклонения получаются, если взять квадратные корни из диагональных элементов ковариационной матрицы:

$$\widehat{\sigma}_{\widehat{\beta}_1} = \sqrt{0.363} = 0.602, \quad \widehat{\sigma}_{\widehat{\beta}_2} = \sqrt{0.099} = 0.315.$$

## Задача 6.13

В модели $y_t = \beta x_t + \varepsilon_t$ ($\beta$, $x_t$ — скаляры и $x_t > 0$) ошибки $\varepsilon_t$ образуют авторегрессию первого порядка: $\varepsilon_t = \rho\varepsilon_{t-1} + u_t$, $0 < \rho < 1$. Покажите, что стандартная оценка дисперсии $\varepsilon_t$, полученная с помощью обычного метода наименьших квадратов, смещена вниз.

## Решение

Найдем сначала дисперсию МНК-оценки параметра $\beta$. Сама оценка равна

$$\widehat{\beta} = \frac{\sum_t x_t y_t}{\sum_t x_t^2} = \frac{\sum_t x_t(\beta x_t + \varepsilon_t)}{\sum_t x_t^2} = \beta + \frac{\sum_t x_t \varepsilon_t}{\sum_t x_t^2}.$$

Эта оценка несмещенная, поэтому ее дисперсия равна

$$\mathrm{V}(\widehat{\beta}) = \mathrm{E}(\widehat{\beta} - \beta)^2 = \mathrm{E}\left(\frac{\sum_t x_t \varepsilon_t}{\sum_t x_t^2}\right)^2 = \frac{\mathrm{E}\left(\sum_t x_t \varepsilon_t\right)^2}{\left(\sum_t x_t^2\right)^2}$$

$$= \frac{\sum_{t=1}^n \sum_{s=1}^n x_t x_s \mathrm{E}(\varepsilon_t \varepsilon_s)}{\left(\sum_{t=1}^n x_t^2\right)^2} = \sigma^2 \frac{\sum_{t=1}^n \sum_{s=1}^n x_t x_s \rho^{|t-s|}}{\left(\sum_{t=1}^n x_t^2\right)^2}.$$

Последнее равенство справедливо, так как из-за авторегрессионной природы ошибок $\mathrm{E}\varepsilon_t\varepsilon_s = \sigma^2\rho^{|t-s|}$ (см. (6.10)). В условиях задачи ($0 < \rho < 1$, $x_t > 0$) имеем:

$$\mathrm{V}(\widehat{\beta}) = \sigma^2 \frac{\sum_{t=1}^n x_t^2 + \sum_{t=1}^n \sum_{s=1, s\neq t}^n x_t x_s \rho^{|t-s|}}{\left(\sum_{t=1}^n x_t^2\right)^2} > \frac{\sigma^2}{\sum x_t^2}.$$

Стандартная оценка параметра $\sigma^2$ (дисперсии $\varepsilon_t$) равна

$$s^2 = \frac{\sum_t (y_t - \widehat{\beta} x_t)^2}{n-1}$$

Найдем ее математическое ожидание.

$$
\begin{aligned}
\mathrm{E}s^2 &= \frac{1}{n-1}\mathrm{E}\sum_t (y_t - \widehat{\beta}x_t)^2 = \frac{1}{n-1}\mathrm{E}\sum_t (\beta x_t + \varepsilon_t - \widehat{\beta}x_t)^2 \\
&= \frac{1}{n-1}\mathrm{E}\sum_t ((\beta - \widehat{\beta})x_t + \varepsilon_t)^2 \\
&= \frac{1}{n-1}\mathrm{E}\sum_t \left((\beta - \widehat{\beta})^2 x_t^2 + 2(\beta - \widehat{\beta})x_t\varepsilon_t + \varepsilon_t^2\right) \\
&= \frac{1}{n-1}\left(\mathrm{E}(\beta - \widehat{\beta})^2\sum_t x_t^2 + 2\sum_t x_t \mathrm{E}((\beta - \widehat{\beta})\varepsilon_t) + \sum_t \mathrm{E}\varepsilon_t^2\right) \\
&= \frac{1}{n-1}\left(\mathrm{V}(\widehat{\beta})\sum_t x_t^2 + 2\sum_t x_t\beta\mathrm{E}\varepsilon_t - 2\sum_t x_t\mathrm{E}(\widehat{\beta}\varepsilon_t) + n\sigma^2\right) \\
&= \frac{1}{n-1}\left(\mathrm{V}(\widehat{\beta})\sum_t x_t^2 - 2\sum_t x_t\mathrm{E}\varepsilon_t\left(\beta + \frac{\sum_s x_s\varepsilon_s}{\sum_s x_s^2}\right) + n\sigma^2\right) \\
&= \frac{1}{n-1}\left(\mathrm{V}(\widehat{\beta})\sum_t x_t^2 - 2\sum_t x_t\mathrm{E}\left(\varepsilon_t\frac{\sum_s x_s\varepsilon_s}{\sum_s x_s^2}\right) + n\sigma^2\right) \\
&= \frac{1}{n-1}\left(\mathrm{V}(\widehat{\beta})\sum_t x_t^2 - 2\frac{\sum_t\sum_s x_t x_s \mathrm{E}(\varepsilon_t\varepsilon_s)}{\sum x_s^2} + n\sigma^2\right) \\
&= \frac{1}{n-1}\left(\mathrm{V}(\widehat{\beta})\sum_t x_t^2 - 2\mathrm{V}(\widehat{\beta})\sum_t x_t^2 + n\sigma^2\right) \\
&= \frac{1}{n-1}\left(n\sigma^2 - \mathrm{V}(\widehat{\beta})\sum_t x_t^2\right) < \frac{1}{n-1}(n\sigma^2 - \sigma^2) = \sigma^2.
\end{aligned}
$$

(В последней строке мы воспользовались выведенным ранее неравенством.)

Таким образом, математическое ожидание оценки $s^2$ оказалось меньше, чем $\sigma^2$, что и означает «смещение вниз».

## Задача 6.14

Расходы домашних хозяйств в Нидерландах (см. продолжение в упражнении 12.16).

*Введение.* Традиционной задачей эмпирических исследований в микроэкономике является оценивание кривых Энгеля. Эрнст Энгель установил, что при увеличении дохода семьи доля расходов на питание уменьшается (закон Энгеля). В современных микроэкономических терминах это означает, что эластичность расходов на питание по доходу меньше единицы. (При этом говорят также, что еда является необходимым товаром, а не предметом роскоши.) Зависимость расходов на приобретение некоторого вида товара от доходов называется кривой Энгеля. В настоящее время принято, как правило, вместо дохода рассматривать полные расходы.

В этой серии упражнений мы будем изучать ежегодные расходы домашних хозяйств (*household*) на питание в зависимости от полных ежегодных расходов и некоторых других переменных на основании данных по расходам семей в Нидерландах.

*Данные.* Используются данные, полученные из архива журнала *Journal of Applied Econometrics* (`expend.xls`). Для наших целей мы взяли данные о годовых расходах на питание, отдых и другие товары за период с октября 1986 г. по сентябрь 1987 г. (427 наблюдений). Список переменных содержится в таблице 6.15.

**Таблица 6.15**

| Переменная | Описание |
|---|---|
| $f3$ | расходы на питание одной семьи с октября 1986 г. по сентябрь 1987 г. в голландских гульденах |
| $v3$ | расходы на отдых с октября 1986 г. по сентябрь 1987 г. в голландских гульденах |
| $tot3$ | полные расходы с октября 1986 г. по сентябрь 1987 г. в голландских гульденах |
| $prov$ | провинция |
| $reg$ | регион |
| $scl$ | социальный класс (1 — нижний класс, ..., 5 — верхний класс) |
| $nahm$ | число членов семьи старше 11 лет |
| $durb$ | степень урбанизации (1 — маленькая деревня, ..., 13 — большой город) |
| $nch06$ | число детей младше 6 лет |
| $nch711$ | число детей от 7 до 11 лет |
| $nch1217$ | число детей от 12 до 17 лет |
| $nch18$ | число детей старше 18 лет |

6.14.1. а) Вычислите суммарные статистики всех переменных. Проверьте, имеют ли смысл ваши результаты.

б) Вычислите корреляционную матрицу переменных $f3$, $v3$, $tot3$ и $nahm$. Проинтерпретируйте результат. Соответствует ли он вашим ожиданиям?

*Расходы на питание.* Здесь мы изучим линейную модель для объяснения логарифма расходов на питание: $lf3 = \ln(f3)$. Мы рассмотрим также некоторую модель для объяснения доли расходов на питание.

6.14.2. а) Проведите парную регрессию $lf3$ на $ltot3 = \ln(tot3)$. С ее помощью постройте двусторонний 95%-ный доверительный интервал для эластичности по доходу.

б) В регрессии п. а) не принимается во внимание размер семьи. Объясните, почему это может привести к смещенности оценки. Объясните, почему таким образом вы, возможно, переоцениваете эластичность по доходу.

6.14.3. а) Проведите регрессию $lf3$ на $ltot3$, $nahm$, $nch06$, $nch711$ и константу.

б) С ее помощью постройте 95%-ный двусторонний доверительный интервал для эластичности по доходу. Сравните ваш результат с результатом упражнения 6.14.2 а).

в) Проинтерпретируйте эффект включения в регрессию состава семьи.

г) Проверьте, отличается ли влияние детей в возрасте до 6 лет от влияния детей в возрасте от 7 до 11 лет.

д) Проверьте, зависит ли каким-нибудь образом влияние наличия детей от их возраста.

6.14.4. а) Добавьте переменную $ltot3s = ltot3 \cdot ltot3$ в правую часть регрессионного уравнения упражнения 6.14.3 и проведите новую регрессию. Является ли переменная $ltot3s$ значимой? Что она означает?

б) Воспользуйтесь полученными результатами для оценивания эластичности при различных уровнях дохода $tot3$.

в) Воспользуйтесь полученными результатами и/или дополнительными вычислениями для построения доверительного интервала для эластичности при $tot3 = 36000$.

6.14.5. а) Постройте фиктивные переменные для каждой из 13 провинций. Проведите регрессию $lf3$ на переменные, включенные в упражнении 6.14.2, и на двенадцать из тринадцати фиктивных переменных. Почему не следует включать все 13 фиктивных переменных? Проверьте (на 5%-ном уровне значимости) совместную значимость эффекта провинции.

б) Проведите аналогичную процедуру с заменой провинций на социальные классы.

6.14.6. а) Воспользуйтесь стратегией «от общего к частному» для построения наиболее подходящей модели, объясняющей величину $lf3$. Проинтерпретируйте ваш результат. Используйте выбранную вами модель в последующих упражнениях.

б) Оцените эластичность по доходу для различных уровней переменной $tot3$. Является еда необходимым товаром или предметом роскоши?

в) Постройте доверительный интервал для эластичности по доходу при $tot3 = 36000$.

г) Постройте двусторонний 95%-ный доверительный интервал для прогнозного значения расходов на питание бездетной семьи из двух человек, принадлежащей среднему классу и проживающей в Амстердаме, если ее общие расходы составляют 50000 гульденов.

д) Нарисуйте график зависимости остатков регрессии от $ltot3$. Какой вывод вы можете сделать относительно предположения о независимости ошибок и регрессора $ltot3$?

6.14.7. а) Постройте переменную $sf3 = 100 \cdot f3/tot3$ (доля расходов на питание в общем бюджете, в %). Следуя процедуре упражнения 6.14.6, постройте наиболее подходящую модель для объяснения $sf3$.

б) Используя результаты п. а), оцените эластичность по доходу для различных значений $tot3$. Сравните с результатами, полученными в упражнении 6.14.6.

в) Нарисуйте график зависимости остатков регрессии от $ltot3$. Какой вывод вы можете сделать?

6.14.8. Долю расходов на питание в общем бюджете можно рассматривать как отрицательный показатель благосостояния: более высокое значение этой доли соответствует более низкому благосостоянию семьи. Тогда «стоимость детей» можно измерить, ответив на следующий вопрос: какой дополнительный доход требуется семье с каждым новым ребенком, чтобы остаться на том же уровне благосостояния, т. е. чтобы иметь ту же долю расходов на питание в общем бюджете?

а) Воспользуйтесь результатами, полученными в упражнении 6.14.7, чтобы оценить стоимость одного ребенка в возрасте 12 лет, беря в качестве отправной точки бездетную семью из двух человек.

б) Вычислите стоимость (первого) ребенка в каждой возрастной группе. Сравните с п. а). Проинтерпретируйте результаты.

*Расходы на питание и гетероскедастичность.* В предыдущих упражнениях мы предполагали, что ошибки *гомоскедастичны.* Сейчас мы попытаемся ответить на вопрос, является ли эта гипотеза приемлемой. В тех случаях, когда ее целесообразно отвергнуть, мы будем исследовать модель с учетом *гетероскедастичности.*

6.14.9. В качестве отправной точки используется модель упражнения 6.14.6.

а) Вычислите остатки $e$ в этой модели и постройте переменную $\ln(e^2)$.

б) Проведите регрессию $\ln(e^2)$ на все независимые переменные модели упражнения 6.14.6.

в) Проверьте совместную значимость всех переменных в п. б).

г) Объясните, почему этот тест можно рассматривать как тест на наличие экспоненциальной гетероскедастичности ($\sigma^2 = \exp(x'\alpha)$), и проинтерпретируйте результат.

6.14.10. В качестве отправной точки используйте модель упражнения 6.14.7.

а) С помощью теста Бреуша–Пагана проверьте гипотезу о гетероскедастичности вида $\sigma = f(\boldsymbol{x}'\boldsymbol{\alpha})$ с неизвестной функцией $f$.

б) С помощью теста Голдфилда–Квандта проверьте гипотезу о наличии гетероскедастичности типа «$\sigma$ возрастает с ростом $ltot3$».

в) Проверьте гипотезу о наличии экспоненциальной гетероскедастичности, следуя той же схеме, что и в упражнении 6.14.9.

6.14.11. В качестве отправной точки используйте модель упражнения 6.14.7, но теперь допускается наличие экспоненциальной гетероскедастичности, т. е. $\sigma = \exp(\boldsymbol{z}'\boldsymbol{\alpha})$.

а) Основываясь на результатах упражнения 6.14.10 в), выберите в качестве $\boldsymbol{z}$ подходящий подвектор вектора $\boldsymbol{x}$.

б) Оцените $\boldsymbol{\alpha}$.

в) Примените метод взвешенных наименьших квадратов с весовыми коэффициентами $\exp(-\boldsymbol{z}'\widehat{\boldsymbol{\alpha}})$.

г) Сравните результаты с результатами упражнения 6.14.7.

## Решение

6.14.1. а) Выборочные статистики всех переменных приведены в таблице 6.16. Просмотр таблицы не выявляет никаких странностей. Данные вполне правдоподобные.

Таблица 6.16

| | $f3$ | $v3$ | $tot3$ | $prov$ | $reg$ | $scl$ |
|---|---|---|---|---|---|---|
| Mean | 6890.882 | 1488.268 | 35762.81 | 7.281030 | 3.377049 | 3.163934 |
| Median | 6838.431 | 614.3000 | 31847.12 | 7.000000 | 4.000000 | 3.000000 |
| Maximum | 17270.75 | 24147.05 | 124564.2 | 13.00000 | 6.000000 | 5.000000 |
| Minimum | 883.9856 | 0.000000 | 8728.860 | 1.000000 | 1.000000 | 1.000000 |
| Std. Dev. | 2745.748 | 2405.462 | 18053.31 | 3.028145 | 1.426028 | 1.096863 |

| | $nahm$ | $durb$ | $nch06$ | $nch711$ | $nch1217$ | $nch18$ |
|---|---|---|---|---|---|---|
| Mean | 2.185012 | 8.264637 | 0.278689 | 0.255269 | 0.231850 | 1.121780 |
| Median | 2.000000 | 8.000000 | 0.000000 | 0.000000 | 0.000000 | 1.000000 |
| Maximum | 5.000000 | 13.00000 | 3.000000 | 3.000000 | 2.000000 | 4.000000 |
| Minimum | 1.000000 | 3.000000 | 0.000000 | 0.000000 | 0.000000 | 0.000000 |
| Std. Dev. | 0.888614 | 2.954703 | 0.631408 | 0.579735 | 0.548251 | 1.045522 |

б) Корреляционная матрица приведена в таблице 6.17. Как видим из табли-

**Таблица 6.17**

| | $f3$ | $v3$ | $tot3$ | $nahm$ |
|---|---|---|---|---|
| $f3$ | 1.000000 | 0.086730 | 0.507776 | 0.519597 |
| $v3$ | 0.086730 | 1.000000 | 0.393359 | 0.145177 |
| $tot3$ | 0.507776 | 0.393359 | 1.000000 | 0.334865 |
| $nahm$ | 0.519597 | 0.145177 | 0.334865 | 1.000000 |

цы, расходы на питание положительно коррелированы как с общими расходами (читай: доходами), так и с числом членов семьи старше 11 лет. В то же время расходы на отдых коррелированы с числом членов семьи старше 11 лет слабее, чем с общими расходами (это может, например, объясняться тем, что большая семья реже ездит отдыхать, потому что это получается слишком дорого). Полученная матрица корреляций вполне соответствует ожиданиям.

6.14.2. а) Построим парную регрессию логарифма расходов на питание на логарифм общих расходов (модель постоянной эластичности расходов на питание по общим расходам). Результаты представлены в таблице 6.18. Число

**Таблица 6.18**

Dependent Variable: $lf3$

| Variable | Coefficient | Std. Error | $t$-Statistic | Probability |
|---|---|---|---|---|
| const | 3.046584 | 0.410839 | 7.415526 | 0.0000 |
| $ltot3$ | 0.549213 | 0.039579 | 13.87632 | 0.0000 |
| $R^2$ | 0.311799 | | | |

наблюдений равно 475. При таком большом числе наблюдений можно считать, что распределение величины $(\widehat{\beta}-\beta)/s_{\widehat{\beta}}$ является стандартным нормальным, и использовать процентные точки стандартного нормального распределения.

Обозначим эластичность расходов на питание по доходу (общим расходам) через $\epsilon$. Тогда 95%-ный доверительный интервал для нее, т. е. для коэффициента при $ltot3$, такой:

$$\epsilon \in (0.5492 - 1.96 \cdot 0.0396, 0.5492 + 1.96 \cdot 0.0396) = (0.4716, 0.6268).$$

б) Действительно, так как доход положительно коррелирован с числом членов семьи, то переменная $ltot3$ «берет на себя» эффект увеличения размера семьи. Соответственно, эластичность расходов на питание по доходу переоценивается.

6.14.3. а) Построим регрессию $lf3$ на $ltot3$, $nahm$, $nch06$, $nch711$ (см. таблицу 6.19). Как видим, оценка эластичности стала заметно ниже.

**Таблица 6.19**

Dependent Variable: $lf3$

| Variable | Coefficient | Std. Error | $t$-Statistic | Probability |
|---|---|---|---|---|
| const | 4.495370 | 0.377578 | 11.90579 | 0.0000 |
| $ltot3$ | 0.359512 | 0.037962 | 9.470237 | 0.0000 |
| $nahm$ | 0.205586 | 0.020471 | 10.04269 | 0.0000 |
| $nch06$ | 0.139467 | 0.026697 | 5.224150 | 0.0000 |
| $nch711$ | 0.118017 | 0.029411 | 4.012623 | 0.0001 |
| $R^2$ | 0.484869 | | | |

б) Доверительный интервал строим аналогично упражнению 6.14.2 а):

$$\epsilon \in (0.3595 - 1.96 \cdot 0.0380, 0.3595 + 1.96 \cdot 0.0380) = (0.2851, 0.4339).$$

Сравнивая его с интервалом из упражнения 6.14.2 а), видим, что он сместился влево (причем он даже не перекрывается с предыдущим интервалом). Значит, подозрения о переоценке эластичности подтвердились.

в) Интерпретируем коэффициенты при переменных $nahm$, $nch06$, $nch711$.

- Увеличение числа членов семьи старше 11 лет на один ведет при прочих равных к увеличению расходов на питание в среднем в $e^{0.206} = 1.228$ раза (примерно на 23%).
- Увеличение числа детей младше 7 лет на один ведет при прочих равных к увеличению расходов на питание в среднем в $e^{0.139} = 1.149$ раза (примерно на 15%).
- Увеличение числа детей от 7 до 11 лет на один ведет при прочих равных к увеличению расходов на питание в среднем в $e^{0.118} = 1.125$ раза (примерно на 12.5%).

г) Чтобы проверить, одинаково ли влияние дополнительного ребенка младше 7 лет и от 7 до 11 лет на расходы семьи на питание, протестируем гипотезу о равенстве коэффициентов при переменных $nch06$ и $nch711$. При этом $F$-статистика равна 0.2768 и $p$-значение равно 0.5991, т. е. гипотеза не отвергается на 5%-ном уровне.

д) См. п. г).

6.14.4. а) Построим регрессию

$$lf3 = \beta_1 + \beta_2 ltot3 + \beta_3 ltot3s + \beta_4 nahm + \beta_5 nch06 + \beta_6 nch711.$$

**Таблица 6.20**

Dependent Variable: $lf3$

| Variable | Coefficient | Std. Error | $t$-Statistic | Probability |
|---|---|---|---|---|
| const | −12.61347 | 5.542773 | −2.275661 | 0.0234 |
| $ltot3$ | 3.667213 | 1.069821 | 3.427877 | 0.0007 |
| $ltot3s$ | −0.159380 | 0.051517 | −3.093737 | 0.0021 |
| $nahm$ | 0.200661 | 0.020329 | 9.870714 | 0.0000 |
| $nch06$ | 0.127341 | 0.026719 | 4.766007 | 0.0000 |
| $nch711$ | 0.112573 | 0.029170 | 3.859190 | 0.0001 |
| $R^2$ | 0.496320 | | | |

Результаты оценивания представлены в таблице 6.20. Как видим, переменная $ltot3s$ значима, и коэффициент при ней отрицательный. Это означает, что с ростом дохода семьи эластичность расходов на питание по доходу уменьшается.

б) В модели из п. а) численное значение оценки эластичности вычисляется по формуле:

$$\epsilon(ltot3) = \widehat{\beta}_2 + 2\widehat{\beta}_3 ltot3 = 3.667 - 0.318 \cdot ltot3$$

(частная производная по $ltot3$). Оценка дисперсии $\epsilon(ltot3)$ равна

$$\begin{aligned}\widehat{\mathrm{V}}(\epsilon(ltot3)) &= \widehat{\mathrm{V}}(\widehat{\beta}_2 + 2\widehat{\beta}_3 ltot3) = \widehat{\mathrm{V}}(\widehat{\beta}_2) + 4\widehat{\mathrm{Cov}}(\widehat{\beta}_2, \widehat{\beta}_3) ltot3 + 4\widehat{\mathrm{V}}(\widehat{\beta}_3) ltot3^2 \\ &= 1.14452 - 0.22032 \cdot ltot3 + 0.010616 \cdot ltot3^2.\end{aligned}$$

в) Подставим в формулы из предыдущего пункта вместо $ltot3$ величину $\ln 36000 = 10.491$. Получим следующее значение оценки эластичности:

$$\epsilon(10.491) = 3.667 - 0.318 \cdot 10.491 = 0.323,$$

дисперсии оценки эластичности:

$$\widehat{\mathrm{V}}(\epsilon(10.491)) = 1.14452 - 0.22032 \cdot 10.491 + 0.010616 \cdot 10.491^2 = 0.00155.$$

Следовательно, 95%-ный доверительный интервал для эластичности равен:

$$(0.323 - 1.96\sqrt{0.00155},\ 0.323 + 1.96\sqrt{0.00155}) = (0.246, 0.400).$$

6.14.5. а) Построим фиктивные переменные для каждой из 13 провинций и проведем регрессию с этими переменными в правой части (см. таблицу 6.21). Все 13 фиктивных переменных при этом включать нельзя из-за «dummy trap» (линейной зависимости регрессоров). Протестируем гипотезу

**Таблица 6.21**

Dependent Variable: $lf3$

| Variable | Coefficient | Std. Error | $t$-Statistic | Probability |
|---|---|---|---|---|
| const | 4.476391 | 0.388914 | 11.50997 | 0.0000 |
| $ltot3$ | 0.345717 | 0.038703 | 8.932480 | 0.0000 |
| $nahm$ | 0.206892 | 0.020723 | 9.983744 | 0.0000 |
| $nch06$ | 0.134469 | 0.027058 | 4.969708 | 0.0000 |
| $nch711$ | 0.112473 | 0.029790 | 3.775594 | 0.0002 |
| $prov2$ | 0.124373 | 0.110888 | 1.121609 | 0.2627 |
| $prov3$ | 0.145834 | 0.141523 | 1.030464 | 0.3034 |
| $prov4$ | 0.213162 | 0.105372 | 2.022940 | 0.0437 |
| $prov5$ | 0.186409 | 0.094758 | 1.967200 | 0.0498 |
| $prov6$ | 0.052354 | 0.103375 | 0.506443 | 0.6128 |
| $prov7$ | 0.214454 | 0.096641 | 2.219090 | 0.0270 |
| $prov8$ | 0.148695 | 0.094592 | 1.571968 | 0.1167 |
| $prov9$ | 0.300002 | 0.121748 | 2.464129 | 0.0141 |
| $prov10$ | 0.195675 | 0.092138 | 2.123705 | 0.0343 |
| $prov11$ | 0.141827 | 0.108845 | 1.303013 | 0.1933 |
| $prov12$ | 0.103903 | 0.102658 | 1.012130 | 0.3121 |
| $prov13$ | 0.282735 | 0.216256 | 1.307412 | 0.1918 |
| $R^2$ | 0.501716 | | | |

о равенстве коэффициентов при переменных $prov2$–$prov13$ нулю (это соответствует незначимости эффекта провинции). Значение $F$-статистики равно 1.155 и $p$-значение равно 0.314. Следовательно, гипотеза незначимости эффекта провинции не отвергается на 5%-ном уровне.

б) Теперь добавим в регрессию четыре из пяти социальных классов (см. таблицу 6.22).

Протестируем гипотезу о незначимости эффекта социального класса: $F$-статистика равна 2.403, $p$-значение равно 0.049. Таким образом, гипотеза о незначимости эффекта социального класса отвергается на 5%-ном уровне (но не слишком уверенно). Если протестировать гипотезу о незначимости эффекта всех социальных классов, кроме второго, то $F$-статистика равна 1.239 ($p$-значение 0.295). Это вместе со значимостью коэффициента при $scl2$ (при выброшенных остальных фиктивных переменных для социальных классов) дает возможность включить в модель фактор $scl2$.

6.14.6. а) Мы остановились на следующей модели (см. таблицу 6.23):

$$lf3 = \beta_1 + \beta_2 ltot3 + \beta_3 ltot3s + \beta_4(nahm - nch1217 - nch18) + \beta_5 nch06 + \beta_6 nch711 + \beta_7 nch1217 + \beta_8 nch18 + \beta_9 scl2 + \beta_{10} durb.$$

**Таблица 6.22**

Dependent Variable: $lf3$

| Variable | Coefficient | Std. Error | $t$-Statistic | Probability |
|---|---|---|---|---|
| const | 4.216024 | 0.403484 | 10.44906 | 0.0000 |
| $ltot3$ | 0.393412 | 0.041536 | 9.471664 | 0.0000 |
| $nahm$ | 0.194458 | 0.021183 | 9.179704 | 0.0000 |
| $nch06$ | 0.140867 | 0.026749 | 5.266350 | 0.0000 |
| $nch711$ | 0.120583 | 0.029383 | 4.103834 | 0.0000 |
| $scl1$ | −0.123351 | 0.091095 | −1.354080 | 0.1764 |
| $scl2$ | −0.131591 | 0.079701 | −1.651070 | 0.0995 |
| $scl3$ | 0.003100 | 0.078065 | 0.039710 | 0.9683 |
| $scl4$ | −0.026105 | 0.072535 | −0.359899 | 0.7191 |
| $R^2$ | 0.496447 | | | |

**Таблица 6.23**

Dependent Variable: $lf3$

| Variable | Coefficient | Std. Error | $t$-Statistic | Probability |
|---|---|---|---|---|
| const | −11.71148 | 5.472015 | −2.140251 | 0.0329 |
| $ltot3$ | 3.505387 | 1.056682 | 3.317353 | 0.0010 |
| $ltot3s$ | −0.151284 | 0.050896 | −2.972390 | 0.0031 |
| $nahm - nch1217 - nch18$ | 0.236386 | 0.032645 | 7.241047 | 0.0000 |
| $nch06$ | 0.114336 | 0.026690 | 4.283892 | 0.0000 |
| $nch711$ | 0.115598 | 0.028876 | 4.003314 | 0.0001 |
| $nch1217$ | 0.132814 | 0.031602 | 4.202755 | 0.0000 |
| $nch18$ | 0.225595 | 0.029826 | 7.563756 | 0.0000 |
| $scl2$ | −0.098154 | 0.041486 | −2.365925 | 0.0184 |
| $durb$ | −0.013885 | 0.005681 | −2.443996 | 0.0149 |
| $R^2$ | 0.515323 | | | |
| $\widehat{\sigma}$ | 0.334866 | | | |

Переменная $nahm - nch1217 - nch18$ (число взрослых членов семьи) использована, чтобы избежать мультиколлинеарности, потому что дети от 12 лет и старше включены в $nahm$. Кроме того, введение этой переменной облегчает интерпретацию коэффициентов.

Проинтерпретируем некоторые коэффициенты:

– Увеличение числа взрослых членов семьи на один ведет при прочих равных к увеличению расходов на питание в среднем в $e^{0.236} = 1.266$ раза (на 26.6%).

– Увеличение числа детей 18 лет или старше на один ведет при прочих равных к увеличению расходов на питание в среднем в $e^{0.226} = 1.254$ раза (на 25.4%).

– Семья, принадлежащая второму социальному классу, при прочих равных тратит на питание в среднем в $e^{-0.0982} = 0.906$ раза больше (на самом деле на 9.4% меньше), чем семья из другого социального класса.

– Увеличение степени урбанизации на единицу ведет при прочих равных к изменению расходов на питание в среднем в $e^{-0.0139} = 0.986$ раза (уменьшает на 1.4%).

б, в) Эластичность расходов на питание по доходу (общим расходам) переменная и убывает по $ltot3$. Вычислим 95%-ные доверительные интервалы для трех разных значений $ltot3$, в том числе для $ltot3 = \ln 36000$. Вычисления аналогичны упражнению 6.14.4 в), поэтому приводим только результаты (см. таблицу 6.24). Как видим, эластичность расходов на питание по доходу не превосходит единицу, следовательно, питание является необходимым товаром.

**Таблица 6.24**

| $tot3$ | $ltot3$ | Эластичность | Дисперсия | Доверительный интервал |
|---|---|---|---|---|
| 20000 | 9.903 | 0.509 | 0.00377 | (0.389, 0.629) |
| 36000 | 10.491 | 0.331 | 0.00162 | (0.252, 0.410) |
| 50000 | 10.819 | 0.232 | 0.00354 | (0.115, 0.348) |

г) Для этой семьи переменные принимают такие значения: $ltot3 = \ln 50000$, $ltot3s = \ln^2 50000$, $nahm - nch1217 - nch18 = 2$, $nch06 = 0$, $nch711 = 0$, $nch1217 = 0$, $nch18 = 0$, $scl2 = 0$, $durb = 13$. Численное значение прогноза расходов на питание вычисляется по формуле:

$$\widehat{lf3} = \widehat{\beta}_1 + \widehat{\beta}_2 \ln 50000 + \widehat{\beta}_3 \ln^2 50000 + \widehat{\beta}_4 \cdot 2 + \widehat{\beta}_{10} \cdot 13 = 8.798.$$

Это соответствует $\widehat{f3} = e^{8.798} = 6620$. Оценка дисперсии $\widehat{lf3}$ равна

$$\widehat{\mathrm{V}}(\widehat{lf3}) = \widehat{\mathrm{V}}(\widehat{\beta}_1 + \widehat{\beta}_2 \ln 50000 + \widehat{\beta}_3 \ln^2 50000 + \widehat{\beta}_4 \cdot 2 + \widehat{\beta}_{10} \cdot 13) = 0.002077$$

Оценка среднего квадрата ошибки прогноза равна

$$\widehat{\mathrm{V}}(\widehat{lf3}) + \widehat{\sigma}^2 = 0.002077 + 0.112135 = 0.114212$$

(подробнее о прогнозировании см. гл. 7).

Доверительный интервал для $lf3$ равен

$$(\widehat{lf3} - 1.96\sqrt{0.114212},\ \widehat{lf3} + 1.96\sqrt{0.114212}) = (8.135, 9.460).$$

Монотонность экспоненты позволяет построить и доверительный интервал для $f3$:

$$(e^{8.135}, e^{9.460}) = (3413, 12838).$$

д) Диаграмма рассеивания остатков регрессии против $ltot3$ приведена на рис. 6.2. Как видим, никакой зависимости остатков от дохода не наблюда-

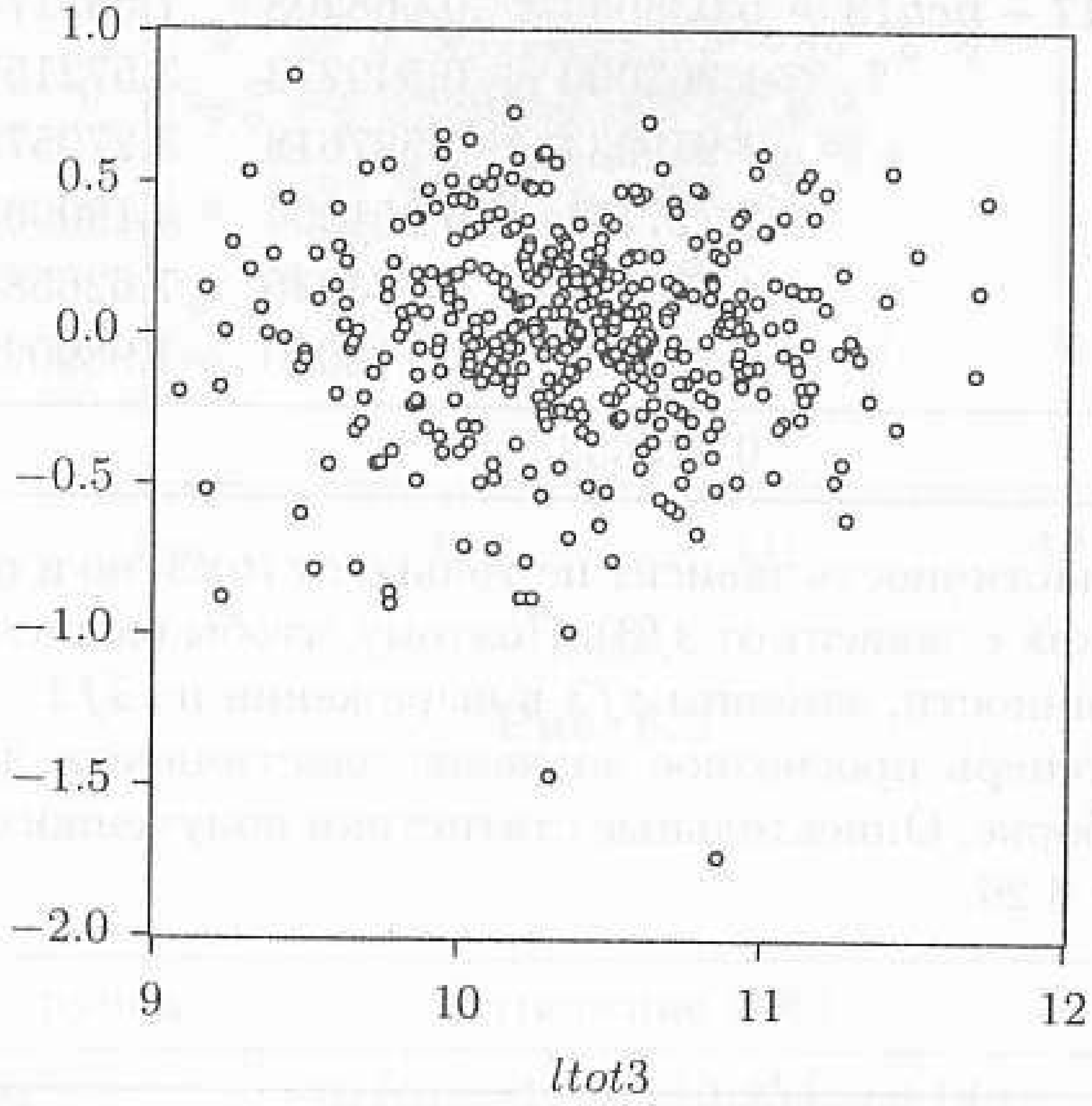



**Рис. 6.2**

ется. Значит, можно сделать вывод о том, что дисперсия ошибок не зависит от $ltot3$.

6.14.7. а) Например, можно остановиться на следующей модели:

$$sf3 = \beta_1 + \beta_2 ltot3 + \beta_3(nahm - nch1217 - nch18) + \beta_4 nch06 + \beta_5 nch711 + \beta_6 nch1217 + \beta_7 nch18 + \beta_8 durb$$

(см. таблицу 6.25). При моделировании $sf3$ ни социальный класс, ни провинции не дают значимого эффекта.

б) Сначала выразим искомую эластичность через параметры модели. Обозначим $f = f3$, $t = tot3$, $s = sf3 = 100 \cdot f3/tot3 = 100f/t$. Имеем:

$$\epsilon = \frac{df}{dt}\frac{t}{f} = \frac{d(st/100)}{dt}\frac{100}{s} = \frac{d(st)}{dt}\frac{1}{s} = \left(s + \frac{ds}{dt}t\right)\frac{1}{s} = \left(s + \frac{d(\beta_1 + \beta_2 \ln t + \cdots)}{dt}t\right)\frac{1}{s} = \left(s + \frac{\beta_2}{t}t\right)\frac{1}{s} = 1 + \frac{\beta_2}{s}.$$

**Таблица 6.25**

Dependent Variable: $sf3$

| Variable | Coefficient | Std. Error | $t$-Statistic | Probability |
|---|---|---|---|---|
| const | 149.8204 | 7.915423 | 18.92765 | 0.0000 |
| $ltot3$ | −13.25445 | 0.778751 | −17.02014 | 0.0000 |
| $nahm - nch1217 - nch18$ | 5.033063 | 0.668208 | 7.532175 | 0.0000 |
| $nch06$ | 1.962090 | 0.549273 | 3.572157 | 0.0004 |
| $nch711$ | 1.959934 | 0.597618 | 3.279578 | 0.0011 |
| $nch1217$ | 2.727751 | 0.651004 | 4.190065 | 0.0000 |
| $nch18$ | 4.675374 | 0.613036 | 7.626584 | 0.0000 |
| $durb$ | −0.234804 | 0.118050 | −1.989022 | 0.0473 |
| $R^2$ | 0.425653 | | | |

Заметим, что эластичность зависит не только от $ltot3$, но и от остальных переменных (так как $\epsilon$ зависит от $sf3$). Поэтому, чтобы вычислить прогнозные значения эластичности, заменим $sf3$ в выражении на $\widehat{s}f3$.

Вычислим теперь прогнозное значение эластичности для каждого наблюдения в выборке. Описательные статистики полученного ряда представлены в таблице 6.26.

**Таблица 6.26**

| | Mean | Median | Maximum | Minimum | Std. Dev. |
|---|---|---|---|---|---|
| $\epsilon$ | 0.241416 | 0.363721 | 0.802258 | −6.769912 | 0.504989 |

Получилось, что эластичность всегда меньше единицы, откуда следует, что питание является необходимым товаром. При этом для 31 наблюдения оказалось, что эластичность меньше нуля. Это отражает ограниченность модели.

в) Диаграмма рассеивания остатков регрессии против $ltot3$ приведена на рис. 6.3. Как видим, наблюдается явное уменьшение разброса остатков с ростом $ltot3$.

6.14.8. а) Дополнительный ребенок 12 лет приводит к тому, что при прочих равных доля расходов на питание возрастает на величину $\beta_6$. Чтобы компенсировать это увеличение, достаточно изменить $ltot3$ на $-\beta_6/\beta_1$, т. е. $ltot3$ должна увеличиться на $2.728/13.254 = 0.206$ или доход должен увеличиться на $(e^{0.206} - 1) \cdot 100\% = 22.9\%$.

б) Аналогично п. а), для разных возрастных групп получились такие результаты (см. таблицу 6.27).

Таким образом, чем старше ребенок, тем дороже он обходится.

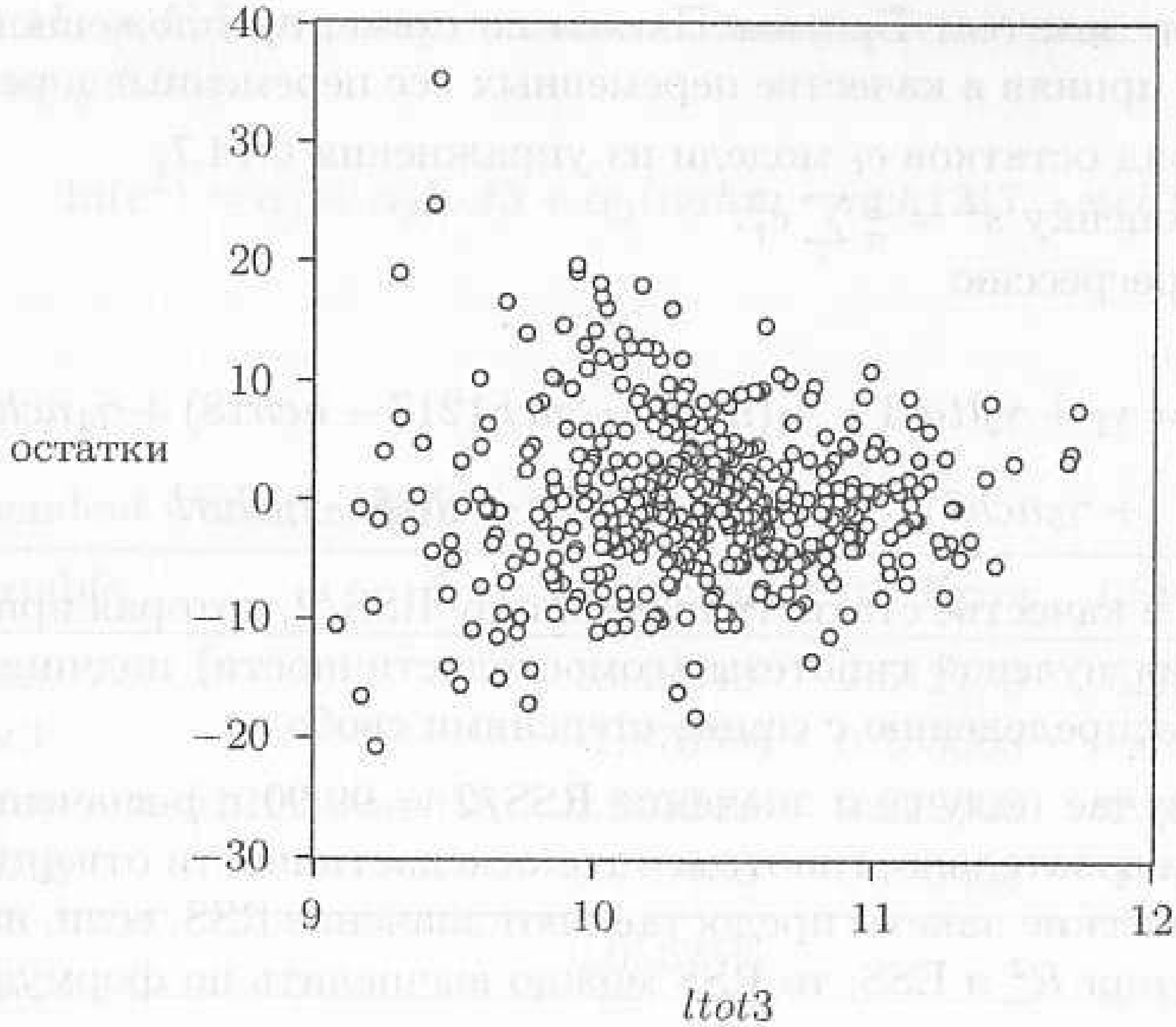



**Рис. 6.3**

**Таблица 6.27**

| Возрастная группа | Изменение $ltot3$ | Прирост дохода |
|---|---|---|
| 0–6 лет | $-\beta_4/\beta_1 = 1.962/13.254 = 0.148$ | 16.0% |
| 7–11 лет | $-\beta_5/\beta_1 = 1.960/13.254 = 0.148$ | 16.0% |
| 12–17 лет | $-\beta_6/\beta_1 = 2.728/13.254 = 0.206$ | 22.9% |
| от 18 лет | $-\beta_7/\beta_1 = 4.675/13.254 = 0.353$ | 42.3% |

6.14.9. а, б, в) Проводимая регрессия $\ln(e^2)$ на все независимые переменные модели, построенной в упражнении 6.14.6, является вспомогательной, поэтому ее результаты не приводятся. При проверке на совместную значимость всех коэффициентов (кроме константы) получается следующая $F$-статистика: $F = 1.9613$ ($p$-значение 0.0424). Следовательно, гипотеза о равенстве нулю всех коэффициентов отвергается на 5%-ном уровне значимости.

г) Эта процедура похожа на тест Уайта. Только в тесте Уайта не предполагается никакая конкретная функциональная форма гетероскедастичности, и производится регрессия квадратов остатков на все независимые переменные и их квадраты. В нашем случае мы пытаемся выявить именно экспоненциальную зависимость квадратов остатков от независимых переменных.

Следовательно, можно сделать вывод о наличии экспоненциальной гетероскедастичности.

6.14.10. а) Проведем тест Бреуша–Пагана по схеме, предложенной в разд. 6.1 на стр. 160, приняв в качестве переменных все переменные в регрессии:

а) получим ряд остатков $e_t$ модели из упражнения 6.14.7;

б) построим оценку $s^2 = \frac{1}{n}\sum e_t^2$;

в) проведем регрессию

$$\frac{e_t^2}{s^2} = \gamma_1 + \gamma_2 ltot3 + \gamma_3(nahm - nch1217 - nch18) + \gamma_4 nch06 + \gamma_5 nch711 + \gamma_6 nch1217 + \gamma_7 nch18 + \gamma_8 durb$$

и возьмем в качестве статистики величину RSS/2, которая при условии выполнения нулевой гипотезы (гомоскедастичности) подчиняется хи-квадрат распределению с семью степенями свободы.

В нашем случае получаем значение RSS/2 = 99.90 с $p$-значением, равным 0.0000. Следовательно, гипотеза о гомоскедастичности отвергается. Не все эконометрические пакеты предоставляют значение RSS, если, например, известны значения $R^2$ и ESS, то RSS можно вычислить по формуле

$$\mathrm{RSS} = \frac{R^2}{1 - R^2}\mathrm{ESS}.$$

б) Проведем тест Голдфелда–Куандта по схеме, предложенной в разд. 6.1 на стр. 160:

а) упорядочим данные по убыванию $ltot3$ и исключим средние 107 (примерно четвертую часть) наблюдений;

б) проведем две регрессии (с теми же переменными, что и в упражнении 6.14.7) по первым 160 и по последним 160 наблюдениям и построим соответствующие суммы квадратов остатков $e_1$ и $e_2$;

в) построим статистику $F = e_1'e_1/e_2'e_2$, которая при условии гомоскедастичности подчиняется распределению Фишера с параметрами $(160 - 8, 160 - 8)$ (здесь 8 — число регрессоров).

В нашем случае получилось значение $F = 0.2848$ с $p$-значением 1.0000. Следовательно, гипотеза о гомоскедастичности не отвергается в пользу гипотезы о том, что $\sigma_t$ монотонно возрастает по $ltot3$.

Если же проверять гипотезу о том, что $\sigma_t$ убывает по $ltot3$, то соответствующее значение $F$-статистики равно 3.5110 ($p$-значение 0.0000). Значит, $\sigma_t$ убывает по $ltot3$.

в) Проведем регрессию логарифмов квадратов остатков $\ln(e^2)$ на все независимые переменные модели, построенной в упражнении 6.14.7, и проверим совместную значимость всех коэффициентов (кроме константы). Получается следующая $F$-статистика: $F = 4.3836$ ($p$-значение 0.0001). Следовательно, гипотеза о гомоскедастичности отвергается на 5%-ном уровне значимости в пользу гипотезы об экспоненциальной гетероскедастичности.

6.14.11. а, б) Например, можно остановиться на такой модели для $\ln(e^2)$ (см. таблицу 6.28):

$$\ln(e^2) = \alpha_1 + \alpha_2 ltot3 + \alpha_3(nahm - nch1217 - nch18) + \alpha_4 nch18.$$

**Таблица 6.28**

Dependent Variable: $\ln(e^2)$

| Variable | Coefficient | Std. Error | $t$-Statistic | Probability |
|---|---|---|---|---|
| const | 13.48349 | 2.512720 | 5.366093 | 0.0000 |
| $ltot3$ | −1.170594 | 0.253068 | −4.625610 | 0.0000 |
| $nahm - nch1217 - nch18$ | 0.600203 | 0.222700 | 2.695117 | 0.0073 |
| $nch18$ | 0.410436 | 0.203053 | 2.021327 | 0.0439 |
| $R^2$ | 0.056256 | | | |

Заметим, что прогнозная сила, так же как и значимость коэффициентов, в этой регрессии не существенна.

в) Создадим переменную $sgm$, равную

$$sgm = \exp(\widehat{\alpha}_1) + \widehat{\alpha}_2 ltot3 + \widehat{\alpha}_3(nahm - nch1217 - nch18) + \widehat{\alpha}_4 nch18.$$

Применим метод взвешенных наименьших квадратов с весами $1/sgm$. Получим следующие результаты (см. таблицу 6.29):

**Таблица 6.29**

Dependent Variable: $sf3$
Weighting series: $1/sgm$

| Variable | Coefficient | Std. Error | $t$-Statistic | Probability |
|---|---|---|---|---|
| const | 136.9761 | 6.619474 | 20.69290 | 0.0000 |
| $ltot3$ | −11.97640 | 0.631139 | −18.97586 | 0.0000 |
| $nahm - nch1217 - nch18$ | 4.530232 | 0.625193 | 7.246137 | 0.0000 |
| $nch06$ | 1.802321 | 0.455736 | 3.954749 | 0.0001 |
| $nch711$ | 1.622363 | 0.457108 | 3.549193 | 0.0004 |
| $nch1217$ | 2.424626 | 0.498812 | 4.860801 | 0.0000 |
| $nch18$ | 4.337045 | 0.563352 | 7.698647 | 0.0000 |
| $durb$ | −0.181426 | 0.104296 | −1.739539 | 0.0827 |
| $R^2$ | 0.420824 | | | |

г) По сравнению с моделью из упражнения 6.14.7 все коэффициенты немного уменьшились по абсолютной величине. Знаки остались прежними.

## Задача 6.15

Рассматриваются следующие данные из газеты «Из рук в руки» за период с декабря 1996 г. по сентябрь 1997 г., касающиеся стоимости однокомнатных квартир в юго-западной части Москвы. Данные содержатся в файле `room1.xls`. В таблице 6.30 приведено описание переменных.

**Таблица 6.30**

| Переменная | Описание |
|---|---|
| $n$ | номер по порядку |
| $distc$ | удаленность от центра, км |
| $distm$ | удаленность от метро, мин |
| $totsq$ | общая площадь квартиры, кв. м |
| $kitsq$ | площадь кухни, кв. м |
| $livsq$ | площадь комнаты, кв. м |
| $floor$ | этаж, 0 — первый или последний, 1 — нет |
| $cat$ | категория дома, 1 — кирпичный, 0 — нет |
| $date$ | дата рекламного объявления |
| $price$ | цена квартиры, тыс. долл. |

а) Найдите среднее, стандартное отклонение и другие выборочные статистики переменных. Найдите коэффициенты корреляции переменных с ценой квартиры. Соответствуют ли полученные значения экономической интуиции?

б) Исследуйте значимость влияния различных факторов на цену квартиры. (Вы можете брать в качестве зависимой переменной цену квартиры, цену квадратного метра общей площади или их логарифмы.)

в) Есть ли существенная зависимость цены квартиры от расстояния до центра? От расстояния до метро? Как интерпретировать результаты?

г) Подберите модель, которая наилучшим способом прогнозирует цену квартиры по имеющимся данным. Проверьте наличие гетероскедастичности.

## Решение

а) Описательные статистики всех переменных приведены в таблице 6.31.

Как видим из таблицы, все статистики выглядят правдоподобно. Введем дополнительную переменную $dopsq = totsq - livsq - kitsq$ — площадь дополнительных помещений (ванной, коридора и т. п.). Ее минимальное значение равно 3.20, что тоже служит сигналом об отсутствии проблем с данными.

**Таблица 6.31**

| | *price* | *totsq* | *livsq* | *kitsq* | *distc* | *distm* | *cat* | *floor* |
|---|---|---|---|---|---|---|---|---|
| Mean | 39.65 | 35.40 | 19.61 | 7.93 | 10.55 | 8.28 | 0.29 | 0.80 |
| Median | 37.00 | 34.00 | 19.00 | 8.00 | 11.40 | 7.00 | 0.00 | 1.00 |
| Maximum | 75.00 | 54.00 | 34.00 | 13.00 | 14.80 | 15.00 | 1.00 | 1.00 |
| Minimum | 28.00 | 28.00 | 14.00 | 5.20 | 4.00 | 1.00 | 0.00 | 0.00 |
| Std. Dev. | 10.04 | 5.15 | 3.16 | 1.72 | 3.62 | 4.16 | 0.46 | 0.41 |

Корреляции цены квартиры с остальными переменными приведены в таблице 6.32.

**Таблица 6.32**

| | *totsq* | *livsq* | *kitsq* | *dopsq* | *distc* | *distm* | *cat* | *floor* |
|---|---|---|---|---|---|---|---|---|
| *price* | 0.743 | 0.562 | 0.518 | 0.405 | −0.440 | −0.012 | 0.599 | 0.141 |

Цена положительно коррелирована со всеми переменными, кроме расстояния до центра (отрицательно) и расстояния до метро (незначимо). Это вполне соответствует интуиции.

б, в, г) Например, можно остановиться на линейной спецификации (см. таблицу 6.33).

**Таблица 6.33**

Dependent Variable: *price*
White Heteroskedasticity-Consistent Standard Errors

| Variable | Coefficient | Std. Error | *t*-Statistic | Probability |
|---|---|---|---|---|
| const | −12.58312 | 5.639750 | −2.231148 | 0.0293 |
| *livsq* | 0.871627 | 0.249282 | 3.496551 | 0.0009 |
| *kitsq* | 2.214960 | 0.323748 | 6.841622 | 0.0000 |
| *dopsq* | 0.923971 | 0.201641 | 4.582249 | 0.0000 |
| $1/distc$ | 38.92379 | 18.02178 | 2.159820 | 0.0347 |
| *cat* | 7.350850 | 2.686298 | 2.736424 | 0.0081 |
| *floor* | 4.840677 | 1.288817 | 3.755908 | 0.0004 |
| $R^2$ | 0.791983 | | | |

Плюсы этой модели в том, что она простая, коэффициенты при всех переменных (кроме $1/distc$) имеют ясную интерпретацию. Кроме того, эта модель оказалась одной из лучших при прогнозировании. Еще один плюс — эта модель оказалась одной из немногих, позволяющих выявить зависимость цены квартиры от расстояния до центра. (Зависимость от расстояния до метро выявить не удалось, во всех моделях коэффициент был незначимый.)

Протестируем модель на гетероскедастичность с помощью теста Уайта (по схеме, приведенной в разд. 6.1 на стр. 159). Получим, что значение статистики $nR^2$ равно 48.00 ($p$-значение равно 0.0037). Значит, гипотеза о гомоскедастичности отвергается. Поэтому в модели используются стандартные ошибки в форме Уайта, которые являются состоятельными оценками стандартных ошибок коэффициентов в условиях гетероскедастичности.

Таким образом, цена квартиры зависит от следующих величин: $livsq$, $kitsq$, $dopsq$, $1/distc$, $cat$, $floor$. Проинтерпретируем некоторые коэффициенты:

– коэффициент 0.872 при $livsq$ означает, что увеличение площади комнаты (квартиры рассматриваются однокомнатные) на 1 кв.м при прочих равных приводит к увеличению цены в среднем на 872 долл.;
– коэффициент 38.924 при $1/distc$ означает, что при увеличении расстояния до центра на $\delta$ км при прочих равных приводит к уменьшению цены в среднем на $(38.924/distc^2)\delta$ долл.; этот коэффициент значим, следовательно, цена квартиры зависит от расстояния до центра;
– коэффициент 7.351 при $cat$ означает, что при прочих равных квартира в кирпичном доме стоит в среднем на 7.351 тыс. долл. дороже, чем в панельном.

## Задача 6.16

Выборка состоит из 70 объявлений о продаже двухкомнатных квартир из газеты «Недвижимость» за сентябрь 1997 г. Были отобраны квартиры в окраинных районах Москвы (новостройки). Данные находятся в файле `room2.xls`, таблица 6.34 содержит описание переменных.

а) Найдите среднее, стандартное отклонение и другие выборочные статистики переменных. Найдите коэффициенты корреляции переменных с ценой квартиры. Соответствуют ли полученные значения экономической интуиции?

б) Исследуйте значимость влияния различных факторов на цену квартиры. (Вы можете брать в качестве зависимой переменной цену квартиры, цену квадратного метра общей площади или их логарифмы.)

в) Есть ли существенная зависимость цены квартиры от расстояния до метро? От наличия телефона? лифта? Как интерпретировать результаты?

г) Что «стоит дороже»: квадратный метр кухни, коридора или комнаты?

д) Подберите модель, которая наилучшим способом прогнозирует цену квартиры по имеющимся данным. Проверьте наличие гетероскедастичности.

**Таблица 6.34**

| Переменная | Описание |
|---|---|
| *n* | номер по порядку |
| *price* | цена квартиры, тыс. долл. |
| *totsq* | общая площадь квартиры, кв. м |
| *livsq* | жилая площадь квартиры, кв. м |
| *kitsq* | площадь кухни, кв. м |
| *distm* | расстояние пешком до метро, мин |
| *floor* | этаж, 0 — первый или последний, 1 — нет |
| *cat* | категория дома, 1 — кирпичный, 0 — нет |
| *tel* | телефон, 1 — есть, 0 — нет |
| *lift* | лифт, 1 — есть, 0 — нет |
| *balc* | балкон, 1 — есть, 0 — нет |

## Решение

а) Описательные статистики всех переменных приведены в таблице 6.35.

**Таблица 6.35**

| | *price* | *totsq* | *livsq* | *kitsq* | *dopsq* | *distm* |
|---|---|---|---|---|---|---|
| Mean | 40.664 | 48.219 | 29.886 | 7.583 | 10.750 | 19.071 |
| Median | 39.250 | 45.400 | 30.000 | 6.700 | 10.100 | 15.000 |
| Maximum | 58.000 | 67.500 | 37.400 | 13.800 | 22.000 | 60.000 |
| Minimum | 31.000 | 37.600 | 23.100 | 5.300 | 4.000 | 2.000 |
| Std. Dev. | 6.992 | 6.379 | 2.927 | 1.871 | 3.128 | 15.052 |

| | *balc* | *cat* | *floor* | *lift* | *tel* |
|---|---|---|---|---|---|
| Mean | 0.800 | 0.143 | 0.729 | 0.771 | 0.914 |
| Median | 1.000 | 0.000 | 1.000 | 1.000 | 1.000 |
| Maximum | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 |
| Minimum | 0.000 | 0.000 | 0.000 | 0.000 | 0.000 |
| Std. Dev. | 0.403 | 0.352 | 0.448 | 0.423 | 0.282 |

Как видим из таблицы, все статистики выглядят правдоподобно.

Корреляции цены квартиры с остальными переменными приведены в таблице 6.36.

Цена положительно коррелирована с *totsq*, *livsq*, *kitsq*, *dopsq*, *floor* (что соответствует интуиции), отрицательно коррелирована с *tel* (что противоречит интуици) и незначимо коррелирована с остальными переменными (что-

Таблица 6.36

| | *totsq* | *livsq* | *kitsq* | *dopsq* | *distm* | *balc* | *cat* | *floor* | *lift* | *tel* |
|---|---|---|---|---|---|---|---|---|---|---|
| *price* | 0.827 | 0.611 | 0.689 | 0.703 | −0.029 | 0.192 | −0.066 | 0.315 | 0.182 | −0.243 |

бы проверить значимость парной корреляции, достаточно оценить парную регрессию и протестировать значимость коэффициента наклона).

б, д) После подбора разных вариантов (в качестве критериев отбора можно использовать простоту (интерпретируемость), значимость коэффициентов, прогнозные свойства (например, $R^2$)) мы остановились на двух моделях (одна, для цены квартиры, представлена в таблице 6.37, другая, для цены квадратного метра общей площади, представлена в таблице 6.38).

Таблица 6.37

Dependent Variable: *price*

| Variable | Coefficient | Std. Error | *t*-Statistic | Probability |
|---|---|---|---|---|
| const | −13.68154 | 4.374040 | −3.127895 | 0.0027 |
| *livsq* | 0.911457 | 0.143015 | 6.373134 | 0.0000 |
| *kitsq* | 1.784517 | 0.282172 | 6.324216 | 0.0000 |
| *dopsq* | 0.784280 | 0.147585 | 5.314081 | 0.0000 |
| *distm* | −0.140563 | 0.026718 | −5.260973 | 0.0000 |
| *cat* | 3.402139 | 1.046388 | 3.251318 | 0.0019 |
| *floor* | 2.563213 | 0.832541 | 3.078782 | 0.0031 |
| *lift* | 2.913062 | 0.932395 | 3.124280 | 0.0027 |
| *tel* | 3.525764 | 1.403592 | 2.511959 | 0.0147 |
| $R^2$ | 0.855493 | | | |

Проверим наличие гетероскедастичности в этих моделях с помощью теста Уайта (см. таблицу 6.39).

Гипотеза о гомоскедастичности не отвергается на 5%-ном уровне значимости в обеих моделях. Следовательно, вычисление ошибок в форме Уайта не нужно.

в) Обе модели показывают зависимость цены квартиры или квадратного метра от расстояния до метро, наличия телефона, лифта (соответствующие коэффициенты значимы, кроме того, они имеют «правильные знаки»).

Проинтерпретируем некоторые коэффициенты: коэффициент при *distm* означает, что при увеличении расстояния пешком до метро на 1 минуту при прочих равных цена квартиры уменьшается в среднем на 141 долл. (соответственно, для другой модели цена квадратного метра уменьшится на 2.79 долл.); коэффициент при *tel* означает, что при прочих равных квартира с телефоном в среднем стоит дороже квартиры без телефона на 3526

**Таблица 6.38**

Dependent Variable: *price/totsq*

| Variable | Coefficient | Std. Error | $t$-Statistic | Probability |
|---|---|---|---|---|
| const | 0.599761 | 0.050119 | 11.96684 | 0.0000 |
| *kitsq* | 0.018297 | 0.004825 | 3.792270 | 0.0003 |
| *distm* | −0.002785 | 0.000539 | −5.164805 | 0.0000 |
| *cat* | 0.074866 | 0.021205 | 3.530666 | 0.0008 |
| *floor* | 0.046743 | 0.016410 | 2.848506 | 0.0059 |
| *lift* | 0.061334 | 0.018123 | 3.384202 | 0.0012 |
| *tel* | 0.070884 | 0.027593 | 2.568909 | 0.0126 |
| $R^2$ | 0.527775 | | | |

**Таблица 6.39**

White Heteroscedasticity Test

| | | | | |
|---|---|---|---|---|
| Модель для *price* | Obs*R-squared | 45.449 | Probability | 0.1206 |
| Модель для *price/totsq* | Obs*R-squared | 20.279 | Probability | 0.5036 |

долл. (соответственно, квадратный метр стоит дороже на 7.09 долл.). Сумма в 3526 долл. выглядит неправдоподобно (срочная установка телефона в Москве стоит дешевле). Видимо, переменная *tel* отражает еще какие-то факторы, не присутствующие среди данных переменных.

г) Сравним цену квадратного метра кухни, комнаты и коридора с помощью модели для цены квартиры (см. таблицу 6.37). Коэффициент, скажем, при *livsq* интерпретируется как величина, на которую в среднем при прочих равных увеличивается цена квартиры при увеличении жилой площади на 1 кв.м. То есть коэффициент при *livsq* — это и есть «цена квадратного метра комнаты». Аналогично для площади кухни и коридора. Результаты попарного сравнения коэффициентов приведены в таблице 6.40.

**Таблица 6.40**

| Нулевая гипотеза | $F$-статистика | $p$-значение |
|---|---|---|
| $\beta_{livsq} = \beta_{kitsq}$ | 6.090 | 0.0164 |
| $\beta_{livsq} = \beta_{dopsq}$ | 0.297 | 0.5875 |
| $\beta_{kitsq} = \beta_{dopsq}$ | 7.717 | 0.0073 |

Гипотеза о равенстве цены квадратного метра комнаты и коридора не отвергается, в то время как цена квадратного метра кухни получилась больше, чем комнаты и коридора (гипотезы о совпадении цен отвергаются на 5%-ном уровне).

Таким образом, квадратный метр кухни «стоит» дороже.

## Задача 6.17

(Arthur van Soest, Tilburg University) Файл `wages.xls` содержит данные о 75 мужчинах и 75 женщинах, работавших на полную ставку (не менее 4 дней в неделю в 1987 г.). Данные получены на основании опроса. В таблице 6.41 приведено описание переменных.

**Таблица 6.41**

| Переменная | Описание |
|---|---|
| $w$ | зарплата, гульденов/час до вычета налогов (1987 г.) |
| $age$ | возраст, лет |
| $sex$ | пол, 1 — для мужчин, 2 — для женщин |
| $edu$ | уровень образования,<br>1 — начальная школа или менее;<br>2 — низшее ремесленное;<br>3 — среднее;<br>4 — высшее ремесленное;<br>5 — университет |

*Вопросы для обсуждения*

– Верно ли, что зарплата мужчин выше, чем зарплата женщин? Если да, то может ли это быть объяснено разницей в возрасте или образовании?
– Какова отдача от образования?
– Одинакова ли зависимость зарплаты от возраста для мужчин и женщин?

а) Вычислите описательные статистики. Постройте матрицу корреляций.
б) Создайте переменную $s = sex - 1$. Обсудите регрессию $w = \beta_0 + \beta_1 s + \beta_2 age$. Проделайте тест на гетероскедастичность. Получите оценку Уайта стандартных отклонений коэффициентов в МНК-оценивании. Проделайте двухшаговую процедуру коррекции на гетероскедастичность.
в) Обсудите регрессию $w = \beta_0 + \beta_1 s + \beta_2 age + \beta_3 edu$. Что можно сказать о коэффициенте при $s$ в этой и предыдущей регрессиях? Насколько реалистична эта модель?
г) Обсудите регрессию $w = \beta_0 + \beta_1 s + \beta_2 age + \beta_3 edu + \beta_5 age^2$. Что будет, если взять полулогарифмическую модель? При каком возрасте зарплата наибольшая? Зависит ли этот возраст от уровня образования? Как интерпретировать коэффициент при $s$ в предыдущих регрессиях?

## Решение

а) Описательные статистики всех переменных приведены в таблице 6.42.

**Таблица 6.42**

| | $w$ | $age$ | $sex$ | $edu$ |
|---|---|---|---|---|
| Mean | 21.784 | 33.113 | 1.500 | 2.767 |
| Median | 20.300 | 31.000 | 1.500 | 3.000 |
| Maximum | 81.670 | 61.000 | 2.000 | 5.000 |
| Minimum | 5.920 | 17.000 | 1.000 | 1.000 |
| Std. Dev. | 10.282 | 10.964 | 0.502 | 1.149 |

Для переменной $edu$ приведем также распределение по уровням (см. таблицу 6.43).

**Таблица 6.43**

| Значение | 1 | 2 | 3 | 4 | 5 |
|---|---|---|---|---|---|
| Количество наблюдений | 24 | 36 | 53 | 25 | 12 |
| Процент наблюдений | 16.00 | 24.00 | 35.33 | 16.67 | 8.00 |

Приведем также таблицу 6.44, содержащую суммарные статистики отдельно по выборкам мужчин ($sex = 1$) и женщин ($sex = 2$).

**Таблица 6.44**

| | Для мужчин | | | Для женщин | | |
|---|---|---|---|---|---|---|
| | $w$ | $age$ | $edu$ | $w$ | $age$ | $edu$ |
| Mean | 24.901 | 35.613 | 2.840 | 18.667 | 30.613 | 2.693 |
| Median | 22.410 | 34.000 | 3.000 | 18.040 | 27.000 | 3.000 |
| Maximum | 81.670 | 61.000 | 5.000 | 48.300 | 59.000 | 5.000 |
| Minimum | 9.240 | 18.000 | 1.000 | 5.920 | 17.000 | 1.000 |
| Std. Dev. | 11.716 | 10.797 | 1.231 | 7.477 | 10.621 | 1.065 |

Как видим, средняя зарплата у женщин меньше, чем у мужчин (на 25%). Но однозначный вывод о дискриминации женщин делать рано. Ведь средний возраст попавших в выборку женщин тоже меньше. И средний уровень образования у женщин в выборке тоже ниже, чем у мужчин.

Матрица корреляций переменных приведена в таблице 6.45.

Она показывает, что зарплата положительно коррелирует с возрастом и уровнем образования (что не противоречит интуиции).

**Таблица 6.45**

| | $w$ | $age$ | $sex$ | $edu$ |
|---|---|---|---|---|
| $w$ | 1.000 | 0.603 | −0.304 | 0.494 |
| $age$ | 0.603 | 1.000 | −0.229 | 0.255 |
| $sex$ | −0.304 | −0.229 | 1.000 | −0.064 |
| $edu$ | 0.494 | 0.255 | −0.064 | 1.000 |

б) Результаты регрессии $w = \beta_0 + \beta_1 s + \beta_2 age$ приведены в таблице 6.46.

**Таблица 6.46**

Dependent Variable: $w$
White Heteroskedasticity-Consistent Standard Errors

| Variable | Coefficient | Std. Error | $t$-Statistic | Probability |
|---|---|---|---|---|
| const | 6.112605 | 2.394150 | 2.553142 | 0.0117 |
| $s$ | −3.596522 | 1.353477 | −2.657246 | 0.0087 |
| $age$ | 0.527576 | 0.061927 | 8.519288 | 0.0000 |
| $R^2$ | 0.392480 | | | |

Протестируем модель на гетероскедастичность с помощью теста Уайта. Получим значение статистики $nR^2 = 29.547$, $p$-значение 0.0000. Значит, гипотеза о гомоскедастичности отвергается, и следует вычислять стандартные ошибки по формулам Уайта.

Эта модель показывает, что зарплата положительно зависит от возраста (при увеличении возраста на 1 год зарплата в среднем увеличивается на 0.526 гульдена в час). И кроме того, модель показывает, что зарплаты мужчины и женщины одного возраста различаются в среднем на 3.597 гульдена в час (что составляет 14.4% от средней зарплаты мужчин). Таким образом, первоначальное предположение о дискриминации в 25% оказалось неправильным (не учтены существенные переменные).

Осуществим двухшаговую процедуру коррекции гетероскедастичности. Для этого возьмем ряд из логарифмов квадратов остатков и проведем вспомогательную регрессию. Например, можно взять модель

$$\ln(e^2) = \alpha_1 + \alpha_2 s + \alpha_3 age + \alpha_4 age^2 + \alpha_5 edu.$$

После этого возьмем прогнозные значения $\widehat{\ln(e^2)}$ из этой модели, построим ряд $sgm = \exp(\widehat{\ln(e^2)}/2)$ и используем ряд $1/sgm$ в качестве взвешивающего. Результат представлен в таблице 6.47.

Тест Уайта теперь не отвергает гипотезу о гомоскедастичности ($nR^2 = 1.559$, $p$-значение 0.8162).

**Таблица 6.47**

Dependent Variable: $w$
White Heteroskedasticity-Consistent Standard Errors

| Variable | Coefficient | Std. Error | $t$-Statistic | Probability |
|---|---|---|---|---|
| const | 6.459675 | 2.093099 | 3.086178 | 0.0024 |
| $s$ | −2.839380 | 1.023035 | −2.775449 | 0.0062 |
| $age$ | 0.502847 | 0.059733 | 8.418261 | 0.0000 |
| $R^2$ | 0.389891 | | | |

в) Результаты регрессии $w = \beta_0 + \beta_1 s + \beta_2 age + \beta_3 edu$ приведены в таблице 6.48.

**Таблица 6.48**

Dependent Variable: $w$
White Heteroskedasticity-Consistent Standard Errors

| Variable | Coefficient | Std. Error | $t$-Statistic | Probability |
|---|---|---|---|---|
| const | −0.035849 | 3.085344 | −0.011619 | 0.9907 |
| $s$ | −3.551143 | 1.121399 | −3.166707 | 0.0019 |
| $age$ | 0.441475 | 0.077447 | 5.700352 | 0.0000 |
| $edu$ | 3.244638 | 0.561057 | 5.783077 | 0.0000 |
| $R^2$ | 0.515500 | | | |

Модель показывает, что заработная плата зависит не только от возраста и пола, но и от уровня образования. При прочих равных увеличение уровня образования на 1 приводит в среднем к увеличению зарплаты на 3.245 гульдена в час. Коэффициент при $s$ еще чуть уменьшился по абсолютной величине, но не существенно.

Эта модель предполагает, что зарплата монотонно возрастает вместе с возрастом. В каком-то диапазоне возрастов это скорее всего верно, но в нашей выборке присутствуют и пожилые люди, зарплата которых может быть ниже, чем в молодости. Поэтому модель не слишком реалистична.

в) Добавим в модель регрессор $age^2$. Если коэффициент при этой переменной окажется отрицательным, то это подтвердит гипотезу о том, что зарплата не является монотонной в зависимости от возраста. Результаты оценки модели см. в таблице 6.49

Как видно из таблицы, коэффициент при $age^2$ положительный (!) и незначимый. Значит, гипотезу о немонотонности зарплаты в зависимости от возраста подтвердить не удалось.

Попробуем теперь исследовать полулогарифмическую модель $\ln w = \beta_0 + \beta_1 s + \beta_2 age + \beta_3 edu + \beta_4 age^2$. Она имеет ясную интерпретацию, ведь все-

**Таблица 6.49**

Dependent Variable: $w$
White Heteroskedasticity-Consistent Standard Errors

| Variable | Coefficient | Std. Error | $t$-Statistic | Probability |
|---|---|---|---|---|
| const | 4.328495 | 9.396166 | 0.460666 | 0.6457 |
| $s$ | −3.682554 | 1.207619 | −3.049434 | 0.0027 |
| $age$ | 0.170743 | 0.664889 | 0.256799 | 0.7977 |
| $edu$ | 3.346771 | 0.735142 | 4.552548 | 0.0000 |
| $age^2$ | 0.003605 | 0.009570 | 0.376712 | 0.7069 |
| $R^2$ | 0.517340 | | | |

возможные прибавки к зарплате, премии, повышение зарплаты и т. п. чаще всего выплачиваются в процентах от основной зарплаты, а не в абсолютных единицах (см. таблицу 6.50).

**Таблица 6.50**

Dependent Variable: $w$
White Heteroskedasticity-Consistent Standard Errors

| Variable | Coefficient | Std. Error | $t$-Statistic | Probability |
|---|---|---|---|---|
| const | 1.276962 | 0.240828 | 5.302377 | 0.0000 |
| $s$ | −0.144716 | 0.046675 | −3.100473 | 0.0023 |
| $age$ | 0.065958 | 0.014020 | 4.704545 | 0.0000 |
| $edu$ | 0.124410 | 0.020956 | 5.936671 | 0.0000 |
| $age^2$ | −0.000613 | 0.000184 | −3.322503 | 0.0011 |
| $R^2$ | 0.609760 | | | |

Как видим, теперь коэффициент при $age^2$ отрицательный и значимый. Это позволяет найти возраст, при котором зарплата наибольшая:

$$age_{\max} = -\frac{\beta_2}{2\beta_4} = \frac{0.065958}{2 \cdot 0.000613} = 53.82.$$

Модель не позволяет показать, как возраст, при котором зарплата максимальна, зависит от уровня образования. Чтобы ответить на этот вопрос, введем в модель параметр $edu \cdot age$ (при этом коэффициент при $edu$ становится незначимым, поэтому удалим его). Получается модель, результаты которой показаны в таблице 6.51.

Теперь возраст, при котором зарплата максимальная, зависит от уровня образования, причем зависит положительно, что не противоречит интуиции:

$$age_{\max} = -\frac{\beta_2 + \beta_3 edu}{2\beta_4} = \frac{0.070856 + 0.003235 \cdot edu}{2 \cdot 0.000789} = 44.88 + 2.05 \cdot edu,$$

Таблица 6.51

Dependent Variable: $w$
White Heteroskedasticity-Consistent Standard Errors

| Variable | Coefficient | Std. Error | $t$-Statistic | Probability |
|---|---|---|---|---|
| const | 1.364076 | 0.242050 | 5.635525 | 0.0000 |
| $s$ | −0.139162 | 0.046595 | −2.986619 | 0.0033 |
| $age$ | 0.070856 | 0.013787 | 5.139384 | 0.0000 |
| $edu \cdot age$ | 0.003235 | 0.000541 | 5.976540 | 0.0000 |
| $age^2$ | −0.000789 | 0.000179 | −4.406767 | 0.0000 |
| $R^2$ | 0.610785 | | | |

и изменяется от 46.92 у человека с начальным образованием до 55.13 у человека с университетским образованием.

Проинтерпретируем теперь коэффициент при переменной $s$. Он означает, что при прочих равных (возрасте и уровне образования) зарплата женщины на $(1 - e^{-0.1392}) \cdot 100\% = 13.0\%$ меньше, чем зарплата мужчины. То есть дискриминация женщин все-таки существует, хотя ее величина и меньше, чем могло показаться вначале.

Какова отдача от образования? Из модели 6.50 следует, что повышение уровня образования на 1 (при прочих равных) приводит к увеличению зарплаты в среднем на $(e^{0.124} - 1) \cdot 100\% = 13.2\%$. В модели 6.51 эффект от образования более сложный. Здесь повышение уровня образования на 1 (при прочих равных) приводит к увеличению зарплаты в среднем на $(e^{0.00324 age} - 1) \cdot 100\%$ (эффект увеличивается с возрастом, т. е. в 20-летнем возрасте дополнительный уровень образования дает прибавку в 6.7%, а в 40-летнем — 13.8%).

Одинакова ли зависимость зарплаты от возраста для мужчин и женщин? При попытке добавить в модель перекрестные члены $s \cdot age$ или $s \cdot age^2$ выявить значимую разницу в поведении заработной платы не удалось. Таким образом, единственное различие в зарплате мужчин и женщин то, что зарплата женщин несколько ниже, чем у мужчин. Поведение зарплаты с возрастом одинаковое.

# Глава 7

# Прогнозирование в регрессионных моделях

## Задача 7.1

Проверьте формулу (7.8):

$$\mathrm{E}(\widehat{y} - y_{n+1})^2 = \sigma^2(1 + \boldsymbol{x}'_{n+1}(\boldsymbol{X}'\boldsymbol{X})^{-1}\boldsymbol{x}_{n+1})$$

для среднеквадратичной ошибки прогноза.

## Решение

Ошибка прогноза $e$ равна

$$e = \widehat{y} - y_{n+1} = \boldsymbol{x}'_{n+1}\widehat{\boldsymbol{\beta}} - (\boldsymbol{x}'_{n+1}\boldsymbol{\beta} + \varepsilon_{n+1}).$$

Поскольку оценка $\widehat{\boldsymbol{\beta}}$ несмещенная и $\mathrm{E}(\varepsilon_{n+1}) = 0$, то математическое ожидание ошибки прогноза равно 0:

$$\mathrm{E}(e) = \mathrm{E}(\boldsymbol{x}'_{n+1}\widehat{\boldsymbol{\beta}}) - \boldsymbol{x}'_{n+1}\boldsymbol{\beta} - \mathrm{E}(\varepsilon_{n+1}) = \boldsymbol{x}'_{n+1}\boldsymbol{\beta} - \boldsymbol{x}'_{n+1}\boldsymbol{\beta} = 0.$$

Отсюда получаем:

$$\begin{aligned}\mathrm{E}(e^2) = \mathrm{V}(e) &= \mathrm{V}(\boldsymbol{x}'_{n+1}\widehat{\boldsymbol{\beta}} - \boldsymbol{x}'_{n+1}\boldsymbol{\beta} - \varepsilon_{n+1}) = \mathrm{V}(\boldsymbol{x}'_{n+1}\widehat{\boldsymbol{\beta}} - \varepsilon_{n+1}) \\ &= \mathrm{V}(\boldsymbol{x}'_{n+1}\widehat{\boldsymbol{\beta}}) + \mathrm{V}(\varepsilon_{n+1}) - 2\,\mathrm{Cov}(\boldsymbol{x}'_{n+1}\widehat{\boldsymbol{\beta}}, \varepsilon_{n+1}).\end{aligned}$$

Последнее слагаемое равно 0, так как $\varepsilon_{n+1}$ и $\widehat{\boldsymbol{\beta}}$ некоррелированы. Таким образом,

$$\begin{aligned}\mathrm{E}(e^2) &= \mathrm{V}(\boldsymbol{x}'_{n+1}\widehat{\boldsymbol{\beta}}) + \mathrm{V}(\varepsilon_{n+1}) = \boldsymbol{x}'_{n+1}\mathrm{V}(\widehat{\boldsymbol{\beta}})\boldsymbol{x}_{n+1} + \sigma^2 \\ &= \sigma^2\boldsymbol{x}'_{n+1}(\boldsymbol{X}'\boldsymbol{X})^{-1}\boldsymbol{x}_{n+1} + \sigma^2 = \sigma^2(1 + \boldsymbol{x}'_{n+1}(\boldsymbol{X}'\boldsymbol{X})^{-1}\boldsymbol{x}_{n+1}),\end{aligned}$$

что и требовалось показать.

## Задача 7.2

Докажите равенство (7.9):

$$\mathrm{E}(\widehat{y}-y_{n+1})^2=\sigma^2\left(1+\frac{1}{n}+\frac{(x_{n+1}-\overline{x})^2}{\sum(x_t-\overline{x})^2}\right)$$

для среднеквадратичной ошибки прогноза в случае парной регрессии.

## Решение

Из решения задачи 7.1 следует, что достаточно показать, что

$$\boldsymbol{x}'_{n+1}(\boldsymbol{X}'\boldsymbol{X})^{-1}\boldsymbol{x}_{n+1}=\frac{1}{n}+\frac{(x_{n+1}-\overline{x})^2}{\sum(x_t-\overline{x})^2}.$$

В случае парной регрессии $k=2$ и

$$\boldsymbol{X}=\begin{bmatrix}1 & x_1\\ \ldots & \ldots\\ 1 & x_n\end{bmatrix},\qquad \boldsymbol{x}_{n+1}=\begin{bmatrix}1\\ x_{n+1}\end{bmatrix}.$$

Тогда

$$\begin{aligned}\boldsymbol{x}'_{n+1}(\boldsymbol{X}'\boldsymbol{X})^{-1}\boldsymbol{x}_{n+1}&=\begin{bmatrix}1 & x_{n+1}\end{bmatrix}\begin{bmatrix}n & n\overline{x}\\ n\overline{x} & \sum x_t^2\end{bmatrix}^{-1}\begin{bmatrix}1\\ x_{n+1}\end{bmatrix}\\
&=\frac{1}{n\sum x_t^2-n^2\overline{x}^2}\begin{bmatrix}1 & x_{n+1}\end{bmatrix}\begin{bmatrix}\sum x_t^2 & -n\overline{x}\\ -n\overline{x} & n\end{bmatrix}\begin{bmatrix}1\\ x_{n+1}\end{bmatrix}\\
&=\frac{1}{n(\sum x_t^2-n\overline{x}^2)}\left(\sum x_t^2+nx_{n+1}^2-2n\overline{x}x_{n+1}\right)\\
&=\frac{\sum x_t^2-n\overline{x}^2+n\overline{x}^2+nx_{n+1}^2-2n\overline{x}x_{n+1}}{n(\sum x_t^2-n\overline{x}^2)}\\
&=\frac{(\sum x_t^2-n\overline{x}^2)+n(x_{n+1}-\overline{x})^2}{n(\sum x_t^2-n\overline{x}^2)}=\frac{1}{n}+\frac{(x_{n+1}-\overline{x})^2}{\sum(x_t-\overline{x})^2}.\end{aligned}$$

## Задача 7.3

Имеется $\boldsymbol{y}=\boldsymbol{X}\boldsymbol{\beta}+\boldsymbol{\varepsilon}$ — классическая регрессионная модель ($\boldsymbol{y}$ — $n\times1$ вектор, $\boldsymbol{X}$ — $n\times k$ матрица, $\boldsymbol{\varepsilon}$ — $n\times1$ вектор ошибок, $\boldsymbol{\beta}$ — $k\times1$ вектор коэффициентов, $\mathrm{E}\boldsymbol{\varepsilon}=\boldsymbol{0}$, $\mathrm{V}(\boldsymbol{\varepsilon})=\sigma^2\boldsymbol{I}$). Пусть $\boldsymbol{x}_{n+1}=(x_{n+1,1},\ldots,x_{n+1,k})'$ — дополнительное наблюдение независимых переменных и $y_{n+1}=\boldsymbol{x}'_{n+1}\boldsymbol{\beta}+\varepsilon_{n+1}$.

Покажите, что если матрица $\boldsymbol{X}$ содержит константу, то ошибка прогноза минимальна, если каждое $x_{n+1,j}$ равно среднему $j$-го столбца матрицы $\boldsymbol{X}$.

## Решение

Как следует из формулы (7.8) (см. также задачу 7.1), надо найти вектор $\boldsymbol{x}_{n+1}$, минимизирующий квадратичную форму $F = \boldsymbol{x}'_{n+1}(\boldsymbol{X}'\boldsymbol{X})^{-1}\boldsymbol{x}_{n+1}$. Введем следующие обозначения (считая, что константа является первым регрессором):

$$\boldsymbol{x}_{n+1} = \begin{bmatrix} 1 \\ \widetilde{\boldsymbol{x}}_{n+1} \end{bmatrix}, \quad \boldsymbol{X} = \begin{bmatrix} \boldsymbol{\imath} & \widetilde{\boldsymbol{X}} \end{bmatrix}, \quad \boldsymbol{\imath} = \begin{bmatrix} 1 & \dots & 1 \end{bmatrix}',$$

где $\widetilde{\boldsymbol{x}}_{n+1}$ — $(k-1)\times 1$ вектор, $\boldsymbol{\imath}$ — $n \times 1$ вектор, $\widetilde{\boldsymbol{X}}$ — $n \times (k-1)$ матрица. Квадратичная форма $F$ в этих обозначениях имеет вид

$$F = \boldsymbol{x}'_{n+1}(\boldsymbol{X}'\boldsymbol{X})^{-1}\boldsymbol{x}_{n+1} = \begin{bmatrix} 1 & \widetilde{\boldsymbol{x}}'_{n+1} \end{bmatrix} \left( \begin{bmatrix} \boldsymbol{\imath}' \\ \widetilde{\boldsymbol{X}}' \end{bmatrix} \begin{bmatrix} \boldsymbol{\imath} & \widetilde{\boldsymbol{X}} \end{bmatrix} \right)^{-1} \begin{bmatrix} 1 \\ \widetilde{\boldsymbol{x}}_{n+1} \end{bmatrix}$$

$$= \begin{bmatrix} 1 & \widetilde{\boldsymbol{x}}'_{n+1} \end{bmatrix} \begin{bmatrix} \boldsymbol{\imath}'\boldsymbol{\imath} & \boldsymbol{\imath}'\widetilde{\boldsymbol{X}} \\ \widetilde{\boldsymbol{X}}'\boldsymbol{\imath} & \widetilde{\boldsymbol{X}}'\widetilde{\boldsymbol{X}} \end{bmatrix}^{-1} \begin{bmatrix} 1 \\ \widetilde{\boldsymbol{x}}_{n+1} \end{bmatrix}.$$

Введем обозначение для матрицы, обратной $\boldsymbol{X}'\boldsymbol{X}$:

$$(\boldsymbol{X}'\boldsymbol{X})^{-1} = \begin{bmatrix} \boldsymbol{\imath}'\boldsymbol{\imath} & \boldsymbol{\imath}'\widetilde{\boldsymbol{X}} \\ \widetilde{\boldsymbol{X}}'\boldsymbol{\imath} & \widetilde{\boldsymbol{X}}'\widetilde{\boldsymbol{X}} \end{bmatrix}^{-1} = \begin{bmatrix} \alpha & \boldsymbol{a}' \\ \boldsymbol{a} & \boldsymbol{A} \end{bmatrix}.$$

В этих обозначениях

$$F = \begin{bmatrix} 1 & \widetilde{\boldsymbol{x}}'_{n+1} \end{bmatrix} \begin{bmatrix} \alpha & \boldsymbol{a}' \\ \boldsymbol{a} & \boldsymbol{A} \end{bmatrix} \begin{bmatrix} 1 \\ \widetilde{\boldsymbol{x}}_{n+1} \end{bmatrix} = \alpha + 2\widetilde{\boldsymbol{x}}'_{n+1}\boldsymbol{a} + \widetilde{\boldsymbol{x}}'_{n+1}\boldsymbol{A}\widetilde{\boldsymbol{x}}_{n+1}.$$

Запишем условия экстремума первого порядка (см. (ЛА.22), (ЛА.24)):

$$\frac{\partial F}{\partial \widetilde{\boldsymbol{x}}'_{n+1}} = 2\boldsymbol{a}' + 2\widetilde{\boldsymbol{x}}'_{n+1}\boldsymbol{A} = 0.$$

Получаем $\widetilde{\boldsymbol{x}}_{n+1} = -\boldsymbol{A}^{-1}\boldsymbol{a}$. Из формул для обращения блочной матрицы (ЛА.17), (ЛА.18) получаем

$$\widetilde{\boldsymbol{x}}_{n+1} = -\boldsymbol{A}^{-1}\boldsymbol{a} = -(-\widetilde{\boldsymbol{X}}'\boldsymbol{\imath})(\boldsymbol{\imath}'\boldsymbol{\imath})^{-1} = \frac{1}{n}\widetilde{\boldsymbol{X}}'\boldsymbol{\imath},$$

т. е. $j$-я компонента вектора $\widetilde{\boldsymbol{x}}_{n+1}$ равна среднему $j$-й колонки матрицы $\widetilde{\boldsymbol{X}}$.

## Задача 7.4

Для модели парной регрессии $y_t = \alpha + \beta x_t + \varepsilon_t$, $t = 1, \ldots, 10$, известно, что

$$\sum y_t = 8, \sum x_t = 40, \sum y_t^2 = 26, \sum x_t^2 = 200, \sum y_t x_t = 20$$

(всюду суммирование от 1 до 10). Для некоторого наблюдения $s$ дано $x_s = 10$. Предполагая, что наблюдение $s$ удовлетворяет исходной модели,

а) вычислите наилучший линейный несмещенный прогноз величины $y_s$;

б) оцените стандартную ошибку прогноза.

## Решение

а) Имеем:

$$\widehat{\beta} = \frac{\sum_{t=1}^n x_t y_t - n\overline{x}\overline{y}}{\sum_{t=1}^n x_t^2 - n\overline{x}^2} = \frac{20 - 10 \cdot 4 \cdot 0.8}{200 - 10 \cdot 4^2} = -0.3,$$

$$\widehat{\alpha} = \overline{y} - \widehat{\beta}\overline{x} = 0.8 + 0.3 \cdot 4 = 2.$$

В соответствии с теорией (см. (7.3)) получаем:

$$\widehat{y}_s = \widehat{\alpha} + \widehat{\beta} x_s = 2 - 0.3 \cdot 10 = -1.$$

б) Найдем сумму квадратов остатков ESS в исходной регрессии. Учитывая, что $\sum_{t=1}^n e_t = 0$ и $\sum_{t=1}^n x_t e_t = 0$, получаем:

$$\begin{aligned} \mathrm{ESS} &= \sum e_t^2 = \sum e_t(y_t - \widehat{\alpha} - \widehat{\beta} x_t) = \sum e_t y_t = \sum (y_t - \widehat{\alpha} - \widehat{\beta} x_t) y_t \\ &= \sum y_t^2 - \widehat{\alpha} \sum y_t - \widehat{\beta} \sum x_t y_t = 26 - 2 \cdot 8 + 0.3 \cdot 20 = 16. \end{aligned}$$

Оценку дисперсии ошибок получаем по формуле

$$\widehat{\sigma}^2 = \frac{\mathrm{ESS}}{n-2} = \frac{16}{8} = 2.$$

Согласно формуле (7.9) оценка среднеквадратичной ошибки прогноза есть величина

$$\delta = \sigma^2\left(1 + \frac{1}{n} + \frac{(x_s - \overline{x})^2}{\sum (x_t - \overline{x})^2}\right) = 2\left(1 + \frac{1}{10} + \frac{(10-4)^2}{200 - 10 \cdot 4^2}\right) = 4.$$

Отсюда получаем, что оценка стандартной ошибки прогноза равна 2.

## Задача 7.5

Стандартная линейная модель $\boldsymbol{y} = \boldsymbol{X}\beta + \varepsilon$, где $\boldsymbol{y}$ — $n \times 1$ вектор, $\boldsymbol{X}$ — $n \times k$ матрица, оценивается обычным методом наименьших квадратов. Имеется дополнительное наблюдение $y_0$, $\boldsymbol{x}_0' = (x_{0\,1}, \ldots, x_{0\,k})$.

С помощью какой статистики можно ответить на вопрос: удовлетворяет ли это наблюдение исходной модели?

## Решение

Одно из возможных решений основано на том факте, что если дополнительное наблюдение $y_0$ удовлетворяет исходной модели, то статистика

$$t = \frac{\widehat{y}_0 - y_0}{\delta}$$

имеет распределение Стьюдента с $n - k$ степенями свободы. Здесь $\widehat{y}_0$ — прогноз для $y_0$, а $\delta$ — оценка стандартной ошибки прогноза,

$$\delta = \sqrt{s^2(1 + \boldsymbol{x}_0'(\boldsymbol{X}'\boldsymbol{X})^{-1}\boldsymbol{x}_0)}.$$

Задавая уровень значимости $\alpha$ и вычислив статистику $t$, мы отвергаем нулевую гипотезу о том, что дополнительное наблюдение удовлетворяет исходной модели, в том случае, когда $|t| > t_{1-\alpha/2}(n-k)$, и не отвергаем эту гипотезу, если $|t| \leqslant t_{1-\alpha/2}(n-k)$. (Здесь $t_{1-\alpha/2}(n-k)$ — двусторонняя $(1-\alpha/2)$-квантиль распределения Стьюдента с $n - k$ степенями свободы.)

## Задача 7.6

Дана регрессионная модель $y_t = \alpha + \beta x_t + \varepsilon_t$, $t = 1, \ldots, n$. Предположим, что параметр $\beta$ известен. Предложите способ прогноза величины $y_{n+1}$ (для заданного $x_{n+1}$) и найдите дисперсию ошибки прогноза.

## Решение

Рассмотрим переменную $z_t = y_t - \beta x_t$. Для нее справедлива модель:

$$z_t = \alpha + \varepsilon_t, \quad t = 1, \ldots, n.$$

Согласно теореме из п. 7.1 главы 7, величина $\widehat{\alpha} = \overline{z}$ является наилучшим (в смысле среднеквадратичного отклонения) линейным прогнозом для $z_{n+1}$. Значит, величина $\widehat{y}_{n+1} = \widehat{\alpha} + \beta x_{n+1}$ может служить прогнозом для $y_{n+1}$.

Докажем несмещенность прогноза.

$$\mathrm{E}(y_{n+1}-\widehat{y}_{n+1})=\mathrm{E}(\alpha+\beta x_{n+1}+\varepsilon_{n+1}-(\widehat{\alpha}+\beta x_{n+1}))=\mathrm{E}(\alpha-\widehat{\alpha})+\mathrm{E}\varepsilon_{n+1}=0,$$

так как $\widehat{\alpha}$ — несмещенная МНК-оценка параметра $\alpha$.

Найдем теперь дисперсию ошибки прогноза $\widehat{y}_{n+1}$.

$$\mathrm{V}(y_{n+1}-\widehat{y}_{n+1})=\mathrm{V}(\alpha+\beta x_{n+1}+\varepsilon_{n+1}-(\widehat{\alpha}+\beta x_{n+1}))=\mathrm{V}(\alpha-\widehat{\alpha}+\varepsilon_{n+1})$$

(так как $\widehat{\alpha}$ строится по наблюдениям с 1 по $n$ и, следовательно, $\widehat{\alpha}$ и $\varepsilon_{n+1}$ независимые)

$$=\mathrm{V}(\alpha-\widehat{\alpha})+\mathrm{V}(\varepsilon_{n+1})=\mathrm{V}(\widehat{\alpha})+\mathrm{V}(\varepsilon_{n+1})$$

$$=\mathrm{V}(\overline{z})+\sigma^2=\frac{1}{n}\mathrm{V}(z)+\sigma^2=\frac{\sigma^2}{n}+\sigma^2=\frac{n+1}{n}\sigma^2.$$

# Глава 8

# Инструментальные переменные

## Задача 8.1

Проверьте формулу (8.5)

$$\widehat{\beta}_{\text{IV}} = (\widehat{X}'\widehat{X})^{-1}\widehat{X}'y = \left(X'Z(Z'Z)^{-1}Z'X\right)^{-1}X'Z(Z'Z)^{-1}Z'y$$

для оценки, полученной двухшаговым методом наименьших квадратов.

## Решение

Имеем $\widehat{X} = Z(Z'Z)^{-1}Z'X$. Тогда

$$\begin{aligned}\widehat{\beta}_{\text{IV}} = (\widehat{X}'\widehat{X})^{-1}\widehat{X}'y &= (X'Z(Z'Z)^{-1}Z'Z(Z'Z)^{-1}Z'X)^{-1}X'Z(Z'Z)^{-1}Z'y\\ &= \left(X'Z(Z'Z)^{-1}Z'X\right)^{-1}X'Z(Z'Z)^{-1}Z'y.\end{aligned}$$

## Задача 8.2

Докажите, что при $m = k$ оценка (8.5)

$$\widehat{\beta}_{\text{IV}} = \left(X'Z(Z'Z)^{-1}Z'X\right)^{-1}X'Z(Z'Z)^{-1}Z'y$$

совпадает с оценкой (8.2)

$$\widehat{\beta}_{\text{IV}} = (Z'X)^{-1}Z'y.$$

## Решение

При $m = k$ матрица $Z'X$ является квадратной невырожденной матрицей. Поэтому согласно (8.5) имеем

$$\widehat{\beta}_{\mathrm{IV}} = (Z'X)^{-1}(Z'Z)(X'Z)^{-1}X'Z(Z'Z)^{-1}Z'y = (Z'X)^{-1}Z'y.$$

## Задача 8.3

Найдите $\mathrm{V}(\widehat{\beta}_{\mathrm{IV}})$ для оценок (8.2):

$$\widehat{\beta}_{\mathrm{IV}} = (Z'X)^{-1}Z'y$$

и (8.5):

$$\widehat{\beta}_{\mathrm{IV}} = \left(X'Z(Z'Z)^{-1}Z'X\right)^{-1}X'Z(Z'Z)^{-1}Z'y.$$

## Решение

Рассмотрим сначала случай, когда $Z$ и $X$ — неслучайные матрицы. Тогда

$$\begin{aligned}\mathrm{V}(\widehat{\beta}_{\mathrm{IV}}) &= \mathrm{V}\Big(\big(X'Z(Z'Z)^{-1}Z'X\big)^{-1}X'Z(Z'Z)^{-1}Z'y\Big)\\ &= \big(X'Z(Z'Z)^{-1}Z'X\big)^{-1}X'Z(Z'Z)^{-1}Z'\mathrm{V}(y)\\ &\quad\times Z(Z'Z)^{-1}Z'X\big(X'Z(Z'Z)^{-1}Z'X\big)^{-1}\\ &= \sigma^2\big(X'Z(Z'Z)^{-1}Z'X\big)^{-1}\\ &\quad\times X'Z(Z'Z)^{-1}Z'Z(Z'Z)^{-1}Z'X\big(X'Z(Z'Z)^{-1}Z'X\big)^{-1}\\ &= \sigma^2\big(X'Z(Z'Z)^{-1}Z'X\big)^{-1}.\end{aligned}$$

Этот случай, однако, не имеет содержательного смысла, так как если регрессоры $X$ не случайные, то сама проблема коррелированности регрессоров и ошибок, для решения которой используется метод инструментальных переменных, отсутствует.

Рассмотрим действительно важный случай, когда $X$ и $Z$ случайные. Однако в этом случае мы можем вычислить только асимптотическую матрицу ковариаций оценок.

Сделаем следующие стандартные для метода IV предположения:

1) Существует $p\lim(1/n)Z'\varepsilon = 0$.
2) Существует $p\lim(1/n)Z'X$ — конечная матрица максимального ранга $k$.
3) Существует $p\lim(1/n)Z'Z$ — конечная положительно определенная матрица.

Рассмотрим общий случай. Обозначим $P = Z(Z'Z)^{-1}Z'$, тогда

$$\widehat{\beta}_{\mathrm{IV}} = (X'PX)^{-1}X'Py = \beta + (X'PX)^{-1}X'P\varepsilon.$$

Заметим сначала, что из 1)–3) вытекает, что следующий предел существует и равен положительно определенной матрице:

$$p\lim\left(\frac{1}{n}\boldsymbol{X}'\boldsymbol{P}\boldsymbol{X}\right) = p\lim\left(\frac{1}{n}\boldsymbol{X}'\boldsymbol{Z}\right)p\lim\left(\frac{1}{n}\boldsymbol{Z}'\boldsymbol{Z}\right)^{-1}p\lim\left(\frac{1}{n}\boldsymbol{Z}'\boldsymbol{X}\right).$$

Также имеем

$$p\lim\left(\frac{1}{n}\boldsymbol{X}'\boldsymbol{P}\boldsymbol{\varepsilon}\right) = p\lim\left(\frac{1}{n}\boldsymbol{X}'\boldsymbol{Z}\right)p\lim\left(\frac{1}{n}\boldsymbol{Z}'\boldsymbol{Z}\right)^{-1}p\lim\left(\frac{1}{n}\boldsymbol{Z}'\boldsymbol{\varepsilon}\right) = \mathbf{0}.$$

Отсюда получаем:

$$\begin{aligned} p\lim(\widehat{\boldsymbol{\beta}}_{\mathrm{IV}}) &= \boldsymbol{\beta} + p\lim((\boldsymbol{X}'\boldsymbol{P}\boldsymbol{X})^{-1}\boldsymbol{X}'\boldsymbol{P}\boldsymbol{\varepsilon}) \\ &= \boldsymbol{\beta} + \left(p\lim\frac{1}{n}\boldsymbol{X}'\boldsymbol{P}\boldsymbol{X}\right)^{-1}p\lim\left(\frac{1}{n}\boldsymbol{X}'\boldsymbol{P}\boldsymbol{\varepsilon}\right) = \boldsymbol{\beta}. \end{aligned}$$

Таким образом, мы доказали, что оценка $\widehat{\boldsymbol{\beta}}_{\mathrm{IV}}$ является состоятельной.

Предположим дополнительно, что $n^{-1/2}\boldsymbol{Z}'\boldsymbol{\varepsilon}$ — последовательность, удовлетворяющая условиям центральной предельной теоремы. (Это верно, например, в том случае, если инструменты $\boldsymbol{Z}$ не случайные.) Тогда оценка метода инструментальных переменных $\widehat{\boldsymbol{\beta}}_{\mathrm{IV}}$ имеет асимптотически нормальное распределение, и мы получаем, что вектор

$$\begin{aligned} \sqrt{n}\,(\widehat{\boldsymbol{\beta}}_{\mathrm{IV}} - \boldsymbol{\beta}) &= \left(\frac{1}{n}\boldsymbol{X}'\boldsymbol{P}\boldsymbol{X}\right)^{-1}\left(\frac{1}{\sqrt{n}}\boldsymbol{X}'\boldsymbol{P}\boldsymbol{\varepsilon}\right) \\ &= \left(\frac{1}{n}\boldsymbol{X}'\boldsymbol{P}\boldsymbol{X}\right)^{-1}\left(\frac{1}{n}\boldsymbol{X}'\boldsymbol{Z}\right)\left(\frac{1}{n}\boldsymbol{Z}'\boldsymbol{Z}\right)^{-1}\left(\frac{1}{\sqrt{n}}\boldsymbol{Z}'\boldsymbol{\varepsilon}\right) \end{aligned}$$

имеет асимптотическое нормальное распределение с асимптотической матрицей ковариаций

$$\sigma^2\,p\lim\left(\frac{1}{n}\boldsymbol{X}'\boldsymbol{P}\boldsymbol{X}\right)^{-1}.$$

Таким образом, мы показали, что

$$\sqrt{n}\,(\widehat{\boldsymbol{\beta}}_{\mathrm{IV}} - \boldsymbol{\beta})$$

имеет асимптотическое нормальное распределение с асимптотической матрицей ковариаций

$$\sigma^2\,p\lim\left(\frac{1}{n}\boldsymbol{X}'\boldsymbol{Z}(\boldsymbol{Z}'\boldsymbol{Z})^{-1}\boldsymbol{Z}'\boldsymbol{X}\right)^{-1}.$$

В частном случае (8.2), когда матрица $\boldsymbol{Z}$ имеет ту же размерность, что и матрица $\boldsymbol{X}$, имеет место равенство

$$(\boldsymbol{X}'\boldsymbol{P}\boldsymbol{X})^{-1} = (\boldsymbol{Z}'\boldsymbol{X})^{-1}\boldsymbol{Z}'\boldsymbol{Z}(\boldsymbol{X}'\boldsymbol{Z})^{-1},$$

так как в этом случае матрица $Z'X$ квадратная, невырожденная. Тогда асимптотическая матрица коварнаций имеет вид

$$\sigma^2 \operatorname{plim}\left(\frac{1}{n}Z'X\right)^{-1} \operatorname{plim}\left(\frac{1}{n}Z'Z\right) \operatorname{plim}\left(\frac{1}{n}X'Z\right)^{-1}.$$

## Задача 8.4

Пусть мы оцениваем регрессионное уравнение

$$y_t = \beta_1 + \beta_2 x_t + \varepsilon_t, \quad t = 1, \ldots, n$$

с помощью метода инструментальных переменных, используя переменную $z_t$ как инструмент для $x_t$.

Покажите, что оценки коэффициентов имеют вид

$$\widehat{\beta}_{1\mathrm{IV}} = \overline{y} - \widehat{\beta}_{2\mathrm{IV}}\overline{x}, \qquad \widehat{\beta}_{2\mathrm{IV}} = \frac{\sum(z_t - \overline{z})(y_t - \overline{y})}{\sum(z_t - \overline{z})(x_t - \overline{x})}$$

и являются решениями системы уравнений

$$n\widehat{\beta}_{1\mathrm{IV}} + \left(\sum x_t\right)\widehat{\beta}_{2\mathrm{IV}} = \sum y_t,$$
$$\left(\sum z_t\right)\widehat{\beta}_{1\mathrm{IV}} + \left(\sum z_t x_t\right)\widehat{\beta}_{2\mathrm{IV}} = \sum z_t y_t.$$

## Решение

Будем использовать обозначения $X = [\imath \;\; x]$ для матрицы регрессоров и $Z = [\imath \;\; z]$ для матрицы инструментальных переменных. Вычислим оценку метода инструментальных переменных:

$$\widehat{\beta}_{\mathrm{IV}} = \begin{bmatrix}\widehat{\beta}_{1\mathrm{IV}}\\ \widehat{\beta}_{2\mathrm{IV}}\end{bmatrix} = (Z'X)^{-1}Z'y = \left(\begin{bmatrix}\imath'\\ z'\end{bmatrix}\begin{bmatrix}\imath & x\end{bmatrix}\right)^{-1}\begin{bmatrix}\imath'\\ z'\end{bmatrix}y$$
$$= \begin{bmatrix}\imath'\imath & \imath'x\\ z'\imath & z'x\end{bmatrix}^{-1}\begin{bmatrix}\imath'y\\ z'y\end{bmatrix} = \frac{1}{nz'x - \imath'x\imath'z}\begin{bmatrix}z'x & -\imath'x\\ -\imath'z & n\end{bmatrix}\begin{bmatrix}\imath'y\\ z'y\end{bmatrix}.$$

Перемножая, получаем:

$$\widehat{\beta}_{2\mathrm{IV}} = \frac{nz'y - \imath'z\imath'y}{nz'x - \imath'x\imath'z} = \frac{nz'y - n\overline{z}\imath'y}{nz'x - n\overline{z}\imath'x} = \frac{(z - \overline{z}\imath)'y}{(z - \overline{z}\imath)'x}$$
$$= \frac{(z - \overline{z}\imath)'(y - \overline{y}\imath)}{(z - \overline{z}\imath)'(x - \overline{x}\imath)} = \frac{\sum(z_t - \overline{z})(y_t - \overline{y})}{\sum(z_t - \overline{z})(x_t - \overline{x})}.$$

(Мы использовали тождество $(z_t - \overline{z})'\imath = 0$.) Теперь вычислим $\widehat{\beta}_{1\mathrm{IV}}$:

$$\widehat{\beta}_{1\mathrm{IV}} = \frac{z'x\imath'y - \imath'xz'y}{nz'x - \imath'x\imath'z} = \frac{nz'x\overline{y} - nz'y\overline{x}}{nz'x - n^2\overline{z}\overline{x}} = \frac{(z'x\overline{y} - n\overline{z}\overline{x}\overline{y}) - (z'y\overline{x} - n\overline{z}\overline{x}\overline{y})}{z'x - n\overline{z}\overline{x}}$$
$$= \frac{(z'x - n\overline{z}\overline{x})\overline{y} - (z'y - n\overline{z}\overline{y})\overline{x}}{z'x - n\overline{z}\overline{x}} = \overline{y} - \widehat{\beta}_{2\mathrm{IV}}\overline{x}.$$

Рассмотрим вторую часть задачи. Оценка $\widehat{\beta}_{\mathrm{IV}}$ удовлетворяет уравнению $\widehat{\beta}_{\mathrm{IV}} = (Z'X)^{-1}Z'y$, а следовательно, и уравнению

$$(Z'X)\widehat{\beta}_{\mathrm{IV}} = Z'y.$$

Запишем последнее уравнение в координатах:

$$\begin{bmatrix} n & \imath'x \\ z'\imath & z'x \end{bmatrix} \begin{bmatrix} \widehat{\beta}_{1\mathrm{IV}} \\ \widehat{\beta}_{2\mathrm{IV}} \end{bmatrix} = \begin{bmatrix} \imath'y \\ z'y \end{bmatrix},$$

или

$$n\widehat{\beta}_{1\mathrm{IV}} + \imath'x\widehat{\beta}_{2\mathrm{IV}} = \imath'y,$$

$$\imath'z\widehat{\beta}_{1\mathrm{IV}} + z'x\widehat{\beta}_{2\mathrm{IV}} = z'y,$$

что совпадает с системой, указанной в условии задачи.

## Задача 8.5

Рассмотрим модель (8.3)

$$y = X\beta + \varepsilon, \quad \mathrm{V}(\varepsilon) = \sigma^2 I,$$

в которой регрессоры $x_{tp}$ коррелированы с ошибками $\varepsilon_t$. Пусть $Z$ — некоторая матрица. Преобразуем исходное уравнение, умножив его слева на $Z'$:

$$Z'y = Z'X\beta + Z'\varepsilon. \qquad (*)$$

Покажите, что оценка обобщенного метода наименьших квадратов (5.4) для вектора коэффициентов уравнения $(*)$ равна

$$\widehat{\beta}_{\mathrm{GLS}} = \left(X'Z(Z'Z)^{-1}Z'X\right)^{-1}X'Z(Z'Z)^{-1}Z'y.$$

Сравните результат с формулой (8.5) для оценки метода инструментальных переменных.

## Решение

Введем обозначения: $\widetilde{X} = Z'X$, $\widetilde{y} = Z'y$ и $\widetilde{\varepsilon} = Z'\varepsilon$. Уравнение $(*)$ в этих обозначениях имеет вид:

$$\widetilde{y} = \widetilde{X}\beta + \widetilde{\varepsilon}, \qquad \mathrm{V}(\widetilde{\varepsilon}) = \mathrm{V}(Z'\varepsilon) = Z'\mathrm{V}(\varepsilon)Z = \sigma^2 Z'Z = \Omega.$$

Применяя формулу (5.4) для оценок вектора коэффициентов уравнения $(*)$ обобщенным методом наименьших квадратов, получаем следующее выражение:

$$\begin{aligned} \widehat{\beta}_{\mathrm{GLS}} &= \left(\widetilde{X}'\Omega^{-1}\widetilde{X}\right)^{-1}\widetilde{X}'\Omega^{-1}\widetilde{y} \\ &= \left((Z'X)'(Z'Z)^{-1}(Z'X)\right)^{-1}(Z'X)(Z'Z)^{-1}Z'y \\ &= \left(X'Z(Z'Z)^{-1}Z'X\right)^{-1}X'Z(Z'Z)^{-1}Z'y, \end{aligned}$$

которое совпадает с формулой (8.5) для оценки метода инструментальных переменных.

## Задача 8.6

Пусть переменные $y_t^*$, $z_t^*$ связаны (точным) уравнением

$$y_t^* = \beta_1 + \beta_2 z_t^*.$$

Однако вместо точных значений мы наблюдаем измеренные (с ошибками измерений) значения $y_t = y_t^* + u_t$ и $z_t = z_t^* + v_t$, где $u_t \sim iid(0, \sigma_u^2)$, $v_t \sim iid(0, \sigma_v^2)$, ошибки $u_t$ и $v_s$ независимы при всех $t$ и $s$. Мы оцениваем методом наименьших квадратов уравнение

$$y_t = \beta_1 + \beta_2 z_t + \varepsilon_t.$$

а) Удовлетворяют ли ошибки в данном уравнении условиям стандартной линейной модели?

б) Найти $\mathrm{Cov}(z_t, \varepsilon_t)$.

в) Найти $p\lim \widehat{\beta}_2$.

## Решение

а, б) Подставив выражения для $y_t^*$, $z_t^*$ в исходное уравнение, получаем:

$$y_t - u_t = \beta_1 + \beta_2(z_t - v_t), \quad \text{или}$$
$$y_t = \beta_1 + \beta_2 z_t + (u_t - \beta_2 v_t).$$

Таким образом, $\varepsilon_t = u_t - \beta_2 v_t$. В силу условия задачи получаем, что ошибки независимы при всех $t$ и дисперсия ошибок постоянна

$$\mathrm{V}(\varepsilon_t) = \mathrm{V}(u_t - \beta_2 v_t) = \sigma_u^2 + \beta_2^2 \sigma_v^2.$$

Поскольку рассматривается модель со стохастическими регрессорами, то в понятие «стандартная модель» входит также некоррелированность ошибок с регрессорами. Вычислим ковариацию $z_t$ и $\varepsilon_t$.

$$\mathrm{Cov}(z_t, \varepsilon_t) = \mathrm{Cov}(z_t^* + v_t, u_t - \beta_2 v_t) = -\beta_2 \sigma_v^2 \neq 0.$$

в) Вычислим предел

$$p\lim \widehat{\beta}_2 = \frac{\mathrm{Cov}(z, y)}{\mathrm{V}(z)} = \frac{\mathrm{Cov}(z^* + v, y^* + u)}{\mathrm{V}(z + v)} = \frac{\mathrm{Cov}(z^* + v, \beta_1 + \beta_2 z_t^* + u)}{\mathrm{V}(z + v)}$$
$$= \frac{\beta_2 \sigma_{z^*}^2}{\sigma_{z^*}^2 + \sigma_v^2} = \beta_2 - \frac{\beta_2 \sigma_v^2}{\sigma_{z^*}^2 + \sigma_v^2}.$$

Таким образом, МНК-оценка параметра $\beta_2$ является асимптотически смещенной и несостоятельной.

# Глава 9

# Системы регрессионных уравнений

## Задача 9.1

Рассмотрим следующую модель:

$$\begin{cases} C_t = \alpha + \beta Y_t + \varepsilon_{1t}, \\ Y_t = C_t + I_t + G_t, \\ I_t = \gamma + \delta Y_t + \varepsilon_{2t}. \end{cases}$$

Эндогенные переменные — $C_t$, $Y_t$, $I_t$, экзогенная переменная — $G_t$. Напишите эту модель в матричной форме и найдите ее приведенную форму. Сколько ограничений накладывается на шесть коэффициентов приведенной формы модели и каковы эти ограничения? Покажите, что при заданных значениях коэффициентов приведенной формы можно единственным образом получить значения коэффициентов $\alpha$, $\beta$, $\gamma$ и $\delta$, т.е. при заданной матрице $\mathbf{\Pi}$ уравнение $\mathbf{B}\mathbf{\Pi} + \mathbf{\Gamma} = \mathbf{0}$ имеет единственное решение относительно $\mathbf{B}$ и $\mathbf{\Gamma}$.

## Решение

Введем обозначения:

$$z_t = \begin{bmatrix} C_t \\ I_t \\ Y_t \end{bmatrix}, \quad w_t = \begin{bmatrix} 1 \\ G_t \end{bmatrix}, \quad \varepsilon_t = \begin{bmatrix} \varepsilon_{1t} \\ \varepsilon_{2t} \\ 0 \end{bmatrix}.$$

Тогда модель может быть записана в матричной форме:

$$\mathbf{B} z_t + \mathbf{\Gamma} w_t = \varepsilon_t,$$

где

$$\mathbf{B} = \begin{bmatrix} 1 & 0 & -\beta \\ 0 & 1 & -\delta \\ -1 & -1 & 1 \end{bmatrix}, \quad \mathbf{\Gamma} = \begin{bmatrix} -\alpha & 0 \\ -\gamma & 0 \\ 0 & -1 \end{bmatrix}.$$

Приведенная форма записывается в виде:

$$z_t = -\mathbf{B}^{-1}\mathbf{\Gamma} w_t + \mathbf{B}^{-1}\varepsilon_t = \mathbf{\Pi} w_t + u_t,$$

где

$$\mathbf{B}^{-1} = \frac{1}{1-\beta-\delta}\begin{bmatrix} 1-\delta & \beta & \beta \\ \delta & 1-\beta & \delta \\ 1 & 1 & 1 \end{bmatrix},$$

$$\mathbf{\Pi} = \frac{1}{1-\beta-\delta}\begin{bmatrix} \alpha-\delta\alpha+\beta\gamma & \beta \\ \gamma-\beta\gamma+\alpha\delta & \delta \\ \alpha+\gamma & 1 \end{bmatrix} = \begin{bmatrix} \pi_C^0 & \pi_C^1 \\ \pi_I^0 & \pi_I^1 \\ \pi_Y^0 & \pi_Y^1 \end{bmatrix}. \qquad (*)$$

На шесть коэффициентов приведенной формы наложены два ограничения:

$$\begin{cases} \pi_Y^0 = \pi_C^0 + \pi_I^0, \\ \pi_Y^1 = \pi_C^1 + \pi_I^1 + 1. \end{cases} \qquad (**)$$

Покажем, что по заданным коэффициентам приведенной формы (элементам матрицы $\mathbf{\Pi}$), удовлетворяющим ограничениям $(**)$, можно единственным образом восстановить коэффициенты $\alpha$, $\beta$, $\gamma$ и $\delta$, т.е. уравнение $\mathbf{B\Pi} + \mathbf{\Gamma} = \mathbf{0}$ всегда имеет единственное решение.

В самом деле, из $(*)$ получаем (если $\pi_Y^1 \neq 0$):

$$\begin{cases} \delta = \dfrac{\pi_I^1}{\pi_Y^1}, \quad \beta = \dfrac{\pi_C^1}{\pi_Y^1}, \\ \alpha + \gamma = \dfrac{\pi_Y^0}{\pi_Y^1}, \\ \alpha(1-\delta) + \gamma\beta = \dfrac{\pi_C^0}{\pi_Y^1}. \end{cases}$$

Подставив $\delta$ и $\beta$ в последние два уравнения, получим систему относительно $\gamma$ и $\alpha$:

$$\begin{cases} \alpha + \gamma = \dfrac{\pi_Y^0}{\pi_Y^1}, \\ \alpha\left(1 - \dfrac{\pi_I^1}{\pi_Y^1}\right) + \gamma\dfrac{\pi_C^1}{\pi_Y^1} = \dfrac{\pi_C^0}{\pi_Y^1}. \end{cases}$$

Подставляя $\gamma = \dfrac{\pi_Y^0}{\pi_Y^1} - \alpha$ во второе уравнение, получаем:

$$\alpha\left(1 - \frac{\pi_I^1}{\pi_Y^1}\right) + \left(\frac{\pi_Y^0}{\pi_Y^1} - \alpha\right)\frac{\pi_C^1}{\pi_Y^1} = \frac{\pi_C^0}{\pi_Y^1}.$$

Отсюда

$$\alpha(\pi_Y^1 - \pi_I^1 - \pi_C^1) = \pi_C^0 - \frac{\pi_Y^0 \pi_C^1}{\pi_Y^1}$$

или, в силу (**),

$$\alpha = \pi_C^0 - \frac{\pi_Y^0 \pi_C^1}{\pi_Y^1}, \quad \gamma = \frac{\pi_Y^0}{\pi_Y^1} - \pi_C^0 + \frac{\pi_Y^0 \pi_C^1}{\pi_Y^1}, \quad \delta = \frac{\pi_I^1}{\pi_Y^1}, \quad \beta = \frac{\pi_C^1}{\pi_Y^1}.$$

## Задача 9.2

Рассмотрим проблему идентифицируемости каждого из уравнений в следующей модели:

$$\begin{cases} P_t + \beta_{12} W_t + \gamma_{11} Q_t + \gamma_{13} P_{t-1} = \varepsilon_{1t}, \\ \beta_{21} P_t + W_t + \beta_{23} N_t + \gamma_{22} S_t + \gamma_{24} W_{t-1} = \varepsilon_{2t}, \\ + \beta_{32} W_t + N_t + \gamma_{32} S_t + \gamma_{33} P_{t-1} + \gamma_{34} W_{t-1} = \varepsilon_{3t}, \end{cases}$$

где $P_t$, $W_t$, $N_t$ — индекс цен, зарплата, профсоюзный взнос соответственно (эндогенные переменные), а $Q_t$ и $S_t$ — производительность труда и количество забастовок (экзогенные переменные). Как выглядят порядковое и ранговое условия, если известно, что:

а) $\gamma_{11} = 0$,

б) $\beta_{21} = \gamma_{22} = 0$,

в) $\gamma_{33} = 0$?

## Решение

При исследовании идентифицируемости уравнений в системе одновременных уравнений при наличии простейших ограничений (равенство нулю того или иного коэффициента) можно воспользоваться следующим практическим приемом, суть которого может быть продемонстрирована на данном примере.

В нашем случае мы имеем три эндогенные переменные — $P_t$, $W_t$, $N_t$, две экзогенные — $Q_t$, $S_t$ и две лагированные эндогенные — $P_{t-1}$, $W_{t-1}$. Представим исходную систему в виде следующей таблицы, в ячейках которой стоят коэффициенты при соответствующей переменной в соответствующем уравнении:

| | $P_t$ | $W_t$ | $N_t$ | $Q_t$ | $S_t$ | $P_{t-1}$ | $W_{t-1}$ |
|---|---|---|---|---|---|---|---|
| 1-е уравнение | 1 | $\beta_{12}$ | 0 | $\gamma_{11}$ | 0 | $\gamma_{13}$ | 0 |
| 2-е уравнение | $\beta_{21}$ | 1 | $\beta_{23}$ | 0 | $\gamma_{22}$ | 0 | $\gamma_{24}$ |
| 3-е уравнение | 0 | $\beta_{32}$ | 1 | 0 | $\gamma_{32}$ | $\gamma_{33}$ | $\gamma_{34}$ |

Тогда выполнение порядкового условия эквивалентно тому, что в каждом уравнении число нулей не меньше числа уравнений минус 1 (в нашем случае — 2). Отсюда следует, что для каждого уравнения порядковое условие выполнено даже без дополнительных ограничений а), б), в). Проверка выполнения рангового условия для любого уравнения осуществляется так. Надо взять какой-либо нулевой коэффициент этого уравнения, выписать весь соответствующий столбец таблицы (исключая этот нулевой коэффициент), повторить эту операцию для всех нулевых коэффициентов уравнения и получить матрицу, число строк которой будет на единицу меньше числа уравнений, а число столбцов не меньше, чем число уравнений минус 1, в силу выполнения порядкового условия. Тогда выполнение рангового условия эквивалентно тому, что построенная матрица имеет полный ранг (т. е. число уравнений минус 1).

В нашем случае соответствующие матрицы таковы:

$$\text{1-е уравнение} — \begin{bmatrix} \beta_{23} & \gamma_{22} & \gamma_{24} \\ 1 & \gamma_{32} & \gamma_{34} \end{bmatrix},$$

$$\text{2-е уравнение} — \begin{bmatrix} \gamma_{11} & \gamma_{13} \\ 0 & \gamma_{33} \end{bmatrix},$$

$$\text{3-е уравнение} — \begin{bmatrix} 1 & \gamma_{11} \\ \beta_{21} & 0 \end{bmatrix}.$$

Тогда получаем:

а) если $\gamma_{11} = 0$, то первое уравнение идентифицируемо, а второе и третье нет;

б) если $\beta_{21} = \gamma_{22} = 0$, то первое и второе уравнения идентифицируемы, а третье нет;

в) если $\gamma_{33} = 0$, то первое и третье уравнения идентифицируемы, а второе нет.

## Задача 9.3

Опишите процедуру оценивания каждого из уравнений следующей системы:

$$\begin{cases} y_{1t} + \beta_{12}y_{2t} + \gamma_{11} + \gamma_{12}x_{2t} = \varepsilon_{1t}, \\ y_{2t} + \gamma_{21} + \gamma_{23}x_{3t} = \varepsilon_{2t}, \\ \beta_{32}y_{2t} + y_{3t} + \gamma_{31} + \gamma_{33}x_{3t} = \varepsilon_{3t}. \end{cases}$$

## Решение

Как и в задаче 9.2, составим таблицу.

| | $y_{1t}$ | $y_{2t}$ | $y_{3t}$ | const | $x_{2t}$ | $x_{3t}$ |
|---|---|---|---|---|---|---|
| 1-е уравнение | 1 | $\beta_{12}$ | 0 | $\gamma_{11}$ | $\gamma_{12}$ | 0 |
| 2-е уравнение | 0 | 1 | 0 | $\gamma_{21}$ | 0 | $\gamma_{23}$ |
| 3-е уравнение | 0 | $\beta_{32}$ | 1 | $\gamma_{31}$ | 0 | $\gamma_{33}$ |

Применяя процедуру, описанную в задаче 9.2, получаем, что первое и второе уравнения идентифицируемы, а третье нет, несмотря на то, что для каждого уравнения выполнены порядковые условия. Согласно двухшаговому методу наименьших квадратов, для оценивания первого уравнения надо осуществить регрессию $y_{2t}$ на $x_{2t}$, $x_{3t}$ и константу, получить прогнозные значения $\widehat{y}_{2t}$, а затем провести регрессию $y_{1t}$ на $\widehat{y}_{2t}$, $x_{2t}$, $x_{3t}$ и константу. Для оценивания второго уравнения достаточно провести регрессию $y_{2t}$ на $x_{3t}$ и константу.

## Задача 9.4

Рассматривается следующая система уравнений:

$$\begin{cases} y_{1t} = \gamma_{10} + \beta_{12}y_{2t} + \beta_{13}y_{3t} + \gamma_{11}x_{1t} + \gamma_{12}x_{2t} + \varepsilon_{1t}, \\ y_{2t} = \gamma_{20} + \beta_{21}y_{1t} + \gamma_{21}x_{1t} + \varepsilon_{2t}, \\ y_{3t} = \gamma_{30} + \beta_{31}y_{1t} + \beta_{32}y_{2t} + \gamma_{31}x_{1t} + \gamma_{33}x_{3t} + \varepsilon_{3t}. \end{cases}$$

Идентифицируемо ли каждое из уравнений системы? Что получится, если применить к первому уравнению двухшаговый метод наименьших квадратов?

## Решение

Составим таблицу (ср. задача 9.2).

| | $y_{1t}$ | $y_{2t}$ | $y_{3t}$ | const | $x_{1t}$ | $x_{2t}$ | $x_{3t}$ |
|---|---|---|---|---|---|---|---|
| 1-е уравнение | $-1$ | $\beta_{12}$ | $\beta_{13}$ | $\gamma_{10}$ | $\gamma_{11}$ | $\gamma_{12}$ | 0 |
| 2-е уравнение | $\beta_{21}$ | $-1$ | 0 | $\gamma_{20}$ | $\gamma_{21}$ | 0 | 0 |
| 3-е уравнение | $\beta_{31}$ | $\beta_{32}$ | $-1$ | $\gamma_{30}$ | $\gamma_{31}$ | 0 | $\gamma_{33}$ |

Для первого и третьего уравнений не выполнены порядковые условия, следовательно, они не идентифицируемы. Для второго уравнения выполнено как порядковое, так и ранговое условие, поэтому оно идентифицируемо.

Применение к первому уравнению двухшагового метода наименьших квадратов приведет к системе с точной коллинеарностью.

## Задача 9.5

Задана система одновременных уравнений ($y_1$, $y_2$, $y_3$ — эндогенные переменные).

$$\begin{cases} y_{1t} = \gamma_{10} + \beta_{12}y_{2t} + \gamma_{11}x_{1t} + \varepsilon_{1t}, \\ y_{2t} = \gamma_{20} + \beta_{23}y_{3t} + \gamma_{21}x_{1t} + \gamma_{23}x_{3t} + \varepsilon_{2t}, \\ y_{3t} = \beta_{31}y_{1t} + \beta_{32}y_{2t} + \gamma_{31}x_{1t} + \gamma_{32}x_{2t} + \gamma_{33}x_{3t} + \varepsilon_{3t}. \end{cases}$$

а) Для каждого из трех уравнений определите, выполняются ли порядковые и ранговые условия идентифицируемости.

б) Повторите п. а) при дополнительном ограничении: $\gamma_{32} = 0$.

в) Повторите п. а) при дополнительном ограничении: $\gamma_{32} = 1$.

г) Повторите п. а) при дополнительном ограничении: $\gamma_{32} = \gamma_{33}$.

## Решение

Составим таблицу (ср. задача 9.2).

| | $y_{1t}$ | $y_{2t}$ | $y_{3t}$ | const | $x_{1t}$ | $x_{2t}$ | $x_{3t}$ |
|---|---|---|---|---|---|---|---|
| 1-е уравнение | $-1$ | $\beta_{12}$ | 0 | $\gamma_{10}$ | $\gamma_{11}$ | 0 | 0 |
| 2-е уравнение | 0 | $-1$ | $\beta_{23}$ | $\gamma_{20}$ | $\gamma_{21}$ | 0 | $\gamma_{23}$ |
| 3-е уравнение | $\beta_{31}$ | $\beta_{32}$ | $-1$ | 0 | $\gamma_{31}$ | $\gamma_{32}$ | $\gamma_{33}$ |

а) Как видим, для первого и второго уравнений порядковые условия выполняются (число нулей в строке не меньше 2). Для третьего уравнения порядковое (а значит, и ранговое) условие не выполняется.

Матрицы для проверки рангового условия для первого и второго уравнений следующие:

$$\text{1-е уравнение} - \begin{bmatrix} \beta_{23} & 0 & \gamma_{23} \\ -1 & \gamma_{32} & \gamma_{33} \end{bmatrix},$$

$$\text{2-е уравнение} - \begin{bmatrix} -1 & 0 \\ \beta_{31} & \gamma_{33} \end{bmatrix}.$$

Ранговые условия выполняются для обоих уравнений, следовательно, первое и второе уравнения идентифицируемы.

б) Если положить $\gamma_{32} = 0$, то порядковые условия для первого и второго уравнений не изменятся, а для третьего уравнения порядковое условие становится выполненным.

Ранговое условие для первого уравнения выполняется, для второго и третьего нет. Следовательно, только первое уравнение идентифицируемо.

в) Если положить $\gamma_{32} = 1$, то для проверки ранговых условий в первом и во втором уравнениях достаточно подставить $\gamma_{32} = 1$ в матрицы из п. а). Ранговые условия выполняются, значит, первое и второе уравнения идентифицируемы.

Чтобы проверить идентифицируемость третьего уравнения, сделаем замену переменной: положим $z_{3t} = y_{3t} - x_{2t}$. Тогда таблица примет вид:

| | $y_{1t}$ | $y_{2t}$ | $z_{3t}$ | const | $x_{1t}$ | $x_{2t}$ | $x_{3t}$ |
|---|---|---|---|---|---|---|---|
| 1-е уравнение | $-1$ | $\beta_{12}$ | 0 | $\gamma_{10}$ | $\gamma_{11}$ | 0 | 0 |
| 2-е уравнение | 0 | $-1$ | $\beta_{23}$ | $\gamma_{20}$ | $\gamma_{21}$ | $\beta_{23}$ | $\gamma_{23}$ |
| 3-е уравнение | $\beta_{31}$ | $\beta_{32}$ | $-1$ | 0 | $\gamma_{31}$ | 0 | $\gamma_{33}$ |

Матрица для проверки рангового условия такова:

$$\text{3-е уравнение} - \begin{bmatrix} \gamma_{10} & 0 \\ \gamma_{20} & \beta_{23} \end{bmatrix}.$$

Ее ранг равен 2, следовательно, третье уравнение идентифицируемо.

г) Если положить $\gamma_{32} = \gamma_{33}$, то аналогично п. в) подставим это условие в матрицы из п. а) и убедимся, что первое и второе уравнения остались идентифицируемыми.

Чтобы проверить идентифицируемость третьего уравнения, сделаем замену переменной: положим $z_{2t} = x_{3t} + x_{2t}$. Тогда таблица примет вид:

| | $y_{1t}$ | $y_{2t}$ | $y_{3t}$ | const | $x_{1t}$ | $z_{2t}$ | $x_{3t}$ |
|---|---|---|---|---|---|---|---|
| 1-е уравнение | $-1$ | $\beta_{12}$ | 0 | $\gamma_{10}$ | $\gamma_{11}$ | 0 | 0 |
| 2-е уравнение | 0 | $-1$ | $\beta_{23}$ | $\gamma_{20}$ | $\gamma_{21}$ | 0 | $\gamma_{23}$ |
| 3-е уравнение | $\beta_{31}$ | $\beta_{32}$ | $-1$ | 0 | $\gamma_{31}$ | $\gamma_{32}$ | 0 |

Матрица для проверки рангового условия такова:

$$\text{3-е уравнение} - \begin{bmatrix} \gamma_{10} & 0 \\ \gamma_{20} & \gamma_{23} \end{bmatrix}.$$

Ее ранг равен 2, следовательно, третье уравнение идентифицируемо.

## Задача 9.6

Рассматривается модель, состоящая из двух внешне не связанных уравнений (SUR):

$$\begin{cases} y_{t1} = \beta_1 + \varepsilon_{t1}, \\ y_{t2} = \beta_2 x_t + \varepsilon_{t2}. \end{cases}$$

По 50 наблюдениям (по каждому уравнению) получены следующие результаты: $\sum x_t = 100$, $\sum x_t^2 = 600$, $\sum x_t y_{t1} = 60$, $\sum x_t y_{t2} = 50$, $\sum y_{t1} = 150$, $\sum y_{t1}^2 = 500$, $\sum y_{t1} y_{t2} = 150$, $\sum y_{t2} = 50$, $\sum y_{t2}^2 = 90$.

а) Напишите формулу для GLS-оценки параметров $\beta_1$, $\beta_2$.

б) Найдите OLS-оценку этих параметров.

в) Найдите SUR (FGLS)-оценку этих параметров и оцените матрицы ковариаций этих оценок.

## Решение

а) Обозначим ковариационную матрицу вектора $[\varepsilon_{t1},\ \varepsilon_{t2}]'$ через

$$\boldsymbol{\Sigma} = \begin{bmatrix} \sigma_{11} & \sigma_{12} \\ \sigma_{12} & \sigma_{22} \end{bmatrix}.$$

Модель SUR, следуя обозначениям из разд. 9.1, можно записать в виде

$$\boldsymbol{y} = \boldsymbol{X\beta} + \boldsymbol{\varepsilon},$$

где

$$\boldsymbol{y} = \begin{bmatrix} \boldsymbol{y}_1 \\ \boldsymbol{y}_2 \end{bmatrix},\ \boldsymbol{X} = \begin{bmatrix} \boldsymbol{X}_1 & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{X}_2 \end{bmatrix},\ \boldsymbol{\beta} = \begin{bmatrix} \beta_1 \\ \beta_2 \end{bmatrix},\ \boldsymbol{\varepsilon} = \begin{bmatrix} \boldsymbol{\varepsilon}_1 \\ \boldsymbol{\varepsilon}_2 \end{bmatrix},\ \mathrm{E}(\boldsymbol{\varepsilon\varepsilon}') = \boldsymbol{\Omega} = \boldsymbol{\Sigma} \otimes \boldsymbol{I}_n.$$

Оценка $\widehat{\beta}_{\mathrm{GLS}}$ равна:

$$\widehat{\beta}_{\mathrm{GLS}} = (\boldsymbol{X}'\boldsymbol{\Omega}^{-1}\boldsymbol{X})^{-1}\boldsymbol{X}'\boldsymbol{\Omega}^{-1}\boldsymbol{y}.$$

Матрица $\boldsymbol{\Omega}^{-1}$ равна:

$$\boldsymbol{\Omega}^{-1} = \boldsymbol{\Sigma}^{-1} \otimes \boldsymbol{I}_n = \frac{1}{\sigma_{11}\sigma_{22} - \sigma_{12}^2} \begin{bmatrix} \sigma_{22} & -\sigma_{12} \\ -\sigma_{12} & \sigma_{11} \end{bmatrix} \otimes \boldsymbol{I}_n.$$

Матрица $\boldsymbol{X}'\boldsymbol{\Omega}^{-1}\boldsymbol{X}$ равна:

$$\boldsymbol{X}'\boldsymbol{\Omega}^{-1}\boldsymbol{X} = \frac{1}{\sigma_{11}\sigma_{22} - \sigma_{12}^2} \begin{bmatrix} \sigma_{22}n & -\sigma_{12}\sum x_t \\ -\sigma_{12}\sum x_t & \sigma_{22}\sum x_t^2 \end{bmatrix}.$$

Вектор $\boldsymbol{X}'\boldsymbol{\Omega}^{-1}\boldsymbol{y}$ равен:

$$\boldsymbol{X}'\boldsymbol{\Omega}^{-1}\boldsymbol{y} = \frac{1}{\sigma_{11}\sigma_{22} - \sigma_{12}^2} \begin{bmatrix} \sigma_{22}\sum y_{t1} - \sigma_{12}\sum y_{t2} \\ -\sigma_{12}\sum x_t y_{t1} + \sigma_{11}\sum x_t y_{t2} \end{bmatrix}.$$

Матрица $(\boldsymbol{X}'\boldsymbol{\Omega}^{-1}\boldsymbol{X})^{-1}$ равна:

$$(\boldsymbol{X}'\boldsymbol{\Omega}^{-1}\boldsymbol{X})^{-1} = \frac{\sigma_{11}\sigma_{22} - \sigma_{12}^2}{\sigma_{11}\sigma_{22}n\sum x_t^2 - \sigma_{11}^2\left(\sum x_t\right)^2} \begin{bmatrix} \sigma_{11}\sum x_t^2 & \sigma_{12}\sum x_t \\ \sigma_{12}\sum x_t & \sigma_{22}n \end{bmatrix}.$$

Оценка $\widehat{\beta}_{\mathrm{GLS}}$ равна:

$$\widehat{\beta}_1 = \frac{1}{\sigma_{11}\sigma_{22}n\sum x_t^2 - \sigma_{11}^2\left(\sum x_t\right)^2}\Big(\sigma_{11}\sigma_{22}\sum x_t^2\sum y_{t1} - \sigma_{11}\sigma_{12}\sum x_t^2\sum y_{t2}$$

$$- \sigma_{12}^2\sum x_t\sum x_t y_{t1} + \sigma_{11}\sigma_{12}\sum x_t\sum x_t y_{t2}\Big),$$

$$\widehat{\beta}_2 = \frac{1}{\sigma_{11}\sigma_{22}n\sum x_t^2 - \sigma_{11}^2\left(\sum x_t\right)^2}\Big(\sigma_{12}\sigma_{22}\sum x_t\sum y_{t1} - \sigma_{12}^2\sum x_t\sum y_{t2}$$

$$- \sigma_{12}\sigma_{22}n\sum x_t y_{t1} + \sigma_{11}\sigma_{22}n\sum x_t y_{t2}\Big).$$

б) МНК-оценки параметров $\beta_1$ и $\beta_2$ получаются оцениванием двух уравнений отдельно. Поэтому

$$\widehat{\beta}_1 = \frac{1}{n}\sum y_{t1} = \frac{150}{50} = 3, \quad \widehat{\beta}_2 = \frac{\sum x_t y_{t2}}{\sum x_t^2} = \frac{50}{600} = 0.0833.$$

в) Сначала оценим параметры матрицы $\mathbf{\Sigma}$ и оценим ковариационную матрицу оценки $\widehat{\beta}_{\mathrm{OLS}}$.

$$\widehat{\sigma}_{11} = \frac{1}{n}\sum e_{t1}^2 = \frac{1}{n}\sum\left(y_{t1} - \frac{\sum y_{t1}}{n}\right)^2 = \frac{\sum y_{t1}^2}{n} - \frac{\left(\sum y_{t1}\right)^2}{n^2}$$
$$= \frac{500}{50} - \frac{150^2}{50^2} = 1,$$
$$\widehat{\sigma}_{22} = \frac{1}{n}\sum e_{t2}^2 = \frac{1}{n}\sum\left(y_{t2} - \frac{\sum x_t y_{t2}}{\sum x_t^2}x_t\right)^2 = \frac{\sum y_{t2}^2}{n} - \frac{\left(\sum x_t y_{t2}\right)^2}{n\sum x_t^2}$$
$$= \frac{50}{50} - \frac{(50)^2}{50\cdot 600} = 0.917,$$
$$\widehat{\sigma}_{12} = \frac{1}{n}\sum e_{t1}e_{t2} = \frac{1}{n}\sum\left(y_{t1} - \frac{\sum y_{t1}}{n}\right)\left(y_{t2} - \frac{\sum x_t y_{t2}}{\sum x_t^2}x_t\right)$$
$$= \frac{\sum y_{t1}y_{t2}}{n} - \frac{\sum y_{t1}\sum y_{t2}}{n^2} - \frac{\sum x_t y_{t1}\sum x_t y_{t2}}{n\sum x_t^2} + \frac{\sum y_{t1}\sum x_t y_{t2}\sum x_t}{n^2\sum x_t^2}$$
$$= \frac{170}{50} - \frac{150\cdot 50}{50^2} - \frac{200\cdot 50}{50\cdot 600} + \frac{150\cdot 50\cdot 100}{50^2\cdot 600} = 0.567.$$

Ковариационная матрица оценки $\widehat{\beta}_{\mathrm{OLS}}$ равна:

$$\mathrm{V}(\widehat{\beta}_{\mathrm{OLS}}) = (\boldsymbol{X}'\boldsymbol{X})^{-1}(\boldsymbol{X}'\boldsymbol{\Omega}\boldsymbol{X})(\boldsymbol{X}'\boldsymbol{X})^{-1}$$
$$= \begin{bmatrix}\frac{1}{n} & 0\\ 0 & \frac{1}{\sum x_t^2}\end{bmatrix}\begin{bmatrix}\sigma_{11}n & \sigma_{12}\sum x_t\\ \sigma_{12}\sum x_t & \sigma_{22}\sum x_t^2\end{bmatrix}\begin{bmatrix}\frac{1}{n} & 0\\ 0 & \frac{1}{\sum x_t^2}\end{bmatrix}$$
$$= \frac{1}{n\sum x_t^2}\begin{bmatrix}\sigma_{11}\sum x_t^2 & \sigma_{12}\sum x_t\\ \sigma_{12} & \sigma_{22}n\end{bmatrix}.$$

Подставим в эту формулу оценки $\widehat{\sigma}_{ij}$ и получим:

$$\widehat{\mathrm{V}}(\widehat{\beta}_1) = \frac{1\cdot 600}{50\cdot 600} = 0.02,$$
$$\widehat{\mathrm{V}}(\widehat{\beta}_2) = \frac{0.917\cdot 50}{50\cdot 600} = 0.00153,$$
$$\widehat{\mathrm{Cov}}(\widehat{\beta}_1, \widehat{\beta}_2) = \frac{0.567\cdot 100}{50\cdot 600} = 0.00189.$$

Подставив оценки $\widehat{\sigma}_{ij}$ в формулу для оценки $\widehat{\beta}_{\text{GLS}}$, получим значения оценок $\widehat{\beta}_{FGLS}$:

$$\widehat{\beta}_1 = 2.549, \quad \widehat{\beta}_2 = 0.135.$$

Ковариационная матрица оценки $\widehat{\beta}_{\text{GLS}}$ равна

$$\mathrm{V}(\widehat{\boldsymbol{\beta}}_{\text{GLS}}) = (\boldsymbol{X}\boldsymbol{\Omega}^{-1}\boldsymbol{X})^{-1}.$$

Подставив оценки $\widehat{\sigma}_{ij}$ в формулу для $(\boldsymbol{X}\boldsymbol{\Omega}^{-1}\boldsymbol{X})^{-1}$, получаем:

$$\widehat{\mathrm{V}}(\widehat{\beta}_1) = 0.0147, \quad \widehat{\mathrm{V}}(\widehat{\beta}_2) = 0.00112, \quad \widehat{\mathrm{Cov}}(\widehat{\beta}_1, \widehat{\beta}_2) = 0.00139.$$

# Глава 10

# Метод максимального правдоподобия в моделях регрессии

### Задача 10.1

Рассмотрим выборку размера 10 из пуассоновского распределения с параметром $\theta$: 1, 4, 3, 2, 3, 0, 1, 1, 0, 5.

а) Вычислите оценку максимального правдоподобия для $\theta$.

б) Покажите графически, что функция правдоподобия и ее логарифм достигают максимума в одной и той же точке $\theta$.

### Решение

а) Значение плотности распределения для значения $y_i$ из выборки равно

$$h(y_i, \theta) = \frac{e^{-\theta}\theta^{y_i}}{y_i!}.$$

Отсюда получаем функцию правдоподобия:

$$L(\theta) = \prod_{i=1}^{10} h(y_i, \theta) = \frac{e^{-10\theta}\,\theta^{\sum y_i}}{\prod y_i!} = \frac{e^{-10\theta}\,\theta^{20}}{207360}.$$

Логарифм функции правдоподобия равен

$$\ln L(\theta) = -10\theta + 20\ln\theta - \ln 207360.$$

Для того чтобы найти максимум функции правдоподобия, приравняем нулю производную:

$$\frac{d\ln L(\theta)}{d\theta} = -10 + \frac{20}{\theta} = 0,$$

откуда получаем оценку максимального правдоподобия

$$\widehat{\theta} = 2.$$

б) См. рис. 10.1.

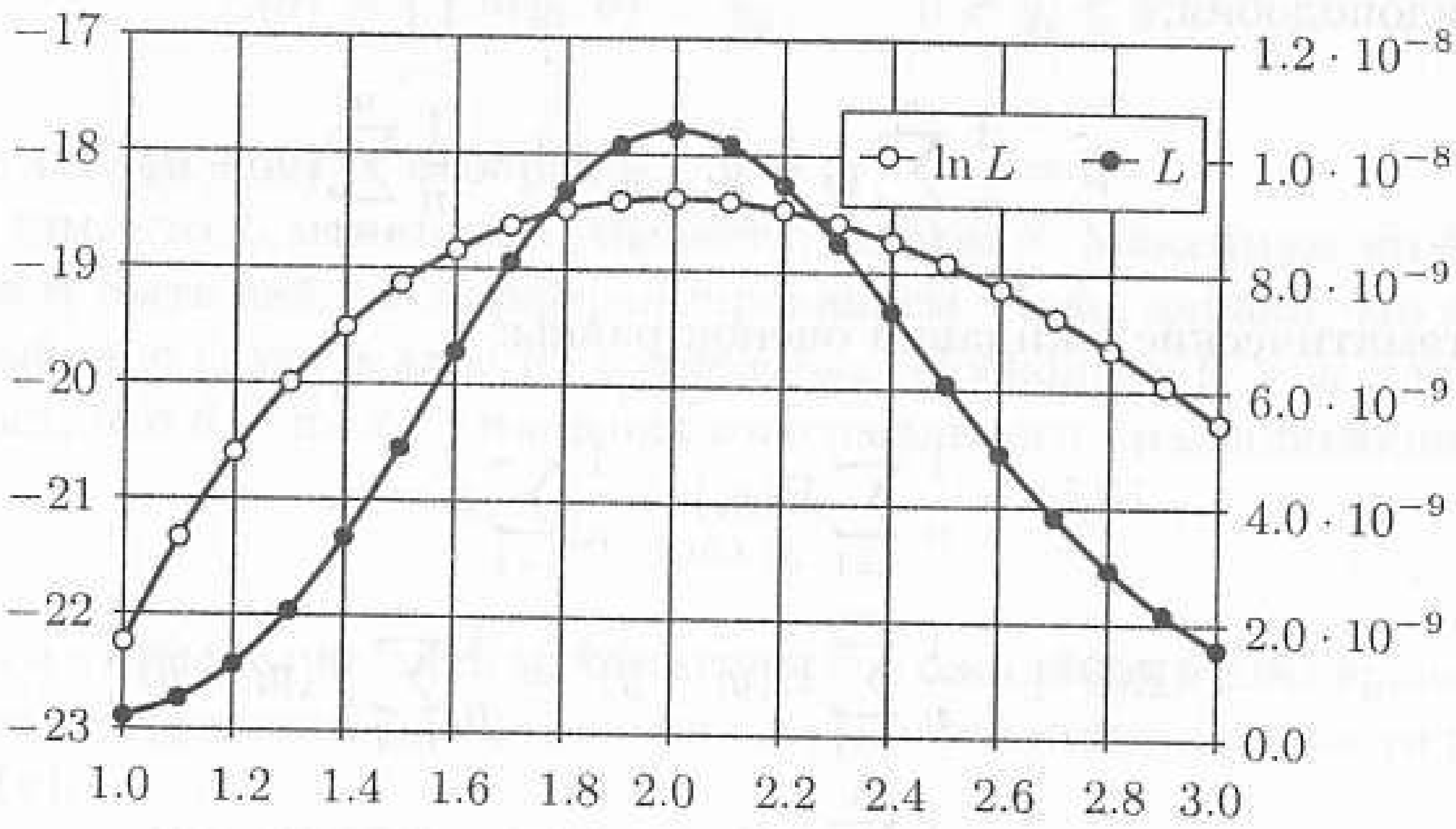



**Рис. 10.1**

## Задача 10.2

Дана выборка размера $n$ из нормального распределения $N(\mu, \sigma^2)$. Запишите логарифмическую функцию правдоподобия и найдите ML-оценки параметров $\mu$ и $\sigma^2$. Найдите смещения этих оценок.

## Решение

Значение плотности распределения для значения $y_i$ из выборки равно

$$h(y_i; \mu, \sigma^2) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left\{-\frac{(y_i - \mu)^2}{2\sigma^2}\right\}.$$

Функция правдоподобия равна

$$L(\mu, \sigma^2) = \prod h(y_i; \mu, \sigma^2) = (2\pi)^{-n/2}(\sigma^2)^{-n/2} \exp\left\{-\frac{1}{2\sigma^2}\sum (y_i - \mu)^2\right\},$$

ее логарифм соответственно равен

$$\ln L(\mu, \sigma^2) = -\frac{n}{2}\ln 2\pi - \frac{n}{2}\ln\sigma^2 - \frac{1}{2\sigma^2}\sum (y_i - \mu)^2.$$

Чтобы найти максимум функции правдоподобия, приравняем нулю частные производные по $\mu$ и $\sigma^2$:

$$\frac{\partial \ln L}{\partial \mu} = \frac{1}{\sigma^2}\sum(y_i - \mu) = 0,$$
$$\frac{\partial \ln L}{\partial \sigma^2} = -\frac{n}{2}\frac{1}{\sigma^2} + \frac{1}{2\sigma^4}\sum(y_i - \mu)^2 = 0.$$

Решив полученную систему уравнений, получаем оценки максимального правдоподобия:

$$\widehat{\mu} = \frac{1}{n}\sum_{i=1}^{n} y_i \equiv \overline{y}, \qquad \widehat{\sigma}^2 = \frac{1}{n}\sum_{i=1}^{n}(y_i - \overline{y})^2.$$

Математические ожидания оценок равны:

$$\mathrm{E}(\widehat{\mu}) = \frac{1}{n}\sum_{i=1}^{n}\mathrm{E}(y_i) = \frac{1}{n}\sum_{i=1}^{n}\mu = \mu,$$
$$\mathrm{E}(\widehat{\sigma}^2) = \frac{1}{n}\sum_{i=1}^{n}\mathrm{E}(y_i - \overline{y})^2 = \frac{1}{n}\sum_{i=1}^{n}\mathrm{V}(y_i - \overline{y})$$
$$= \frac{1}{n}\sum_{i=1}^{n}(\mathrm{V}(y_i) + \mathrm{V}(\overline{y}) - 2\,\mathrm{Cov}(y_i, \overline{y}))$$
$$= \frac{1}{n}\sum_{i=1}^{n}\left(\sigma^2 + \frac{1}{n}\sigma^2 - 2\frac{1}{n}\sigma^2\right) = \frac{n-1}{n}\sigma^2.$$

Смещения оценок равны соответственно:

$$\mathrm{bias}(\widehat{\mu}) = \mathrm{E}(\widehat{\mu}) - \mu = 0, \qquad \mathrm{bias}(\widehat{\sigma}^2) = \mathrm{E}(\widehat{\sigma}^2) - \sigma^2 = -\frac{1}{n}\sigma^2.$$

## Задача 10.3

Пусть $y_1, \ldots, y_n$ — выборка из распределения с плотностью $h(y, \theta) = 1/\theta$, если $0 < x \leqslant \theta$, и $h(y, \theta) = 0$ — в остальных случаях $(0 < \theta < \infty)$. Покажите, что $\widehat{\theta} = \max y_i$ является оценкой максимального правдоподобия, и найдите ее смещение.

## Решение

По условию задачи случайные величины $y_1, y_2, \ldots, y_n$ независимы и равномерно распределены на интервале $(0, \theta]$. (Заметим, что поскольку распределение непрерывное, мы вольны включать или исключать крайние точки по соображениям удобства. Здесь мы позволяем $y$ принимать значение $\theta$, чтобы избежать обсуждения о различии максимума и супремума.) Функция распределения случайной величины $y$ имеет вид

$$F_y(z) = \frac{z}{\theta}, \quad 0 < z \leqslant \theta.$$

($F_y(z) = 0$ при $z \leqslant 0$ и $F_y(z) = 1$ при $z > \theta$.) Функция правдоподобия равна

$$L(\theta) = \prod_{i=1}^{n} h(y_i, \theta) = \frac{1}{\theta^n}, \qquad 0 < y_i \leqslant \theta,$$

и равна нулю при других наборах $y_i$.

Мы видим, что $L$ монотонно убывает с ростом $\theta$. Максимум этой функции не может быть найден дифференцированием. Ясно, однако, что $\theta$ должно быть выбрано самое малое из возможных. Поскольку $\theta \geqslant y_i$ для всех $i$, это означает, что $\theta \geqslant \max y_i$, и оценка максимального правдоподобия равна

$$\widehat{\theta} = \max y_i.$$

Для того чтобы вычислить математическое ожидание и дисперсию оценки $\widehat{\theta}$, найдем ее функцию распределения $F(z)$ и функцию плотности распределения $f(z)$.

$$F(z) = \mathrm{P}(\widehat{\theta} \leqslant z) = \mathrm{P}(\max y_i \leqslant z) = \prod_{i=1}^{n} \mathrm{P}(y_i \leqslant z) = \left(\frac{z}{\theta}\right)^n = \theta^{-n} z^n.$$

(Мы использовали здесь независимость $y_i$.) Функция плотности распределения находится дифференцированием:

$$f(z) = F'(z) = n\theta^{-n} z^{n-1}.$$

Отсюда получаем:

$$\mathrm{E}(\widehat{\theta}) = \mathrm{E}(\max y_i) = \int_0^\theta z f(z) dz = n\theta^{-n} \int_0^\theta z\, z^{n-1} dz = \theta \frac{n}{n+1},$$

$$\mathrm{E}(\widehat{\theta}^2) = \int_0^\theta z^2 f(z) dz = n\theta^{-n} \int_0^\theta z^2 z^{n-1} dz = \theta^2 \frac{n}{n+2},$$

$$\mathrm{V}(\widehat{\theta}) = \mathrm{E}(\widehat{\theta}^2) - (\mathrm{E}\widehat{\theta})^2 = \theta^2 \left[\frac{n}{n+2} - \left(\frac{n}{n+1}\right)^2\right] = \frac{n\theta^2}{(n+1)^2(n+2)}.$$

Таким образом, смещение оценки $\widehat{\theta}$ равно

$$\mathrm{bias}(\widehat{\theta}) = \mathrm{E}(\widehat{\theta}) - \theta = \theta \frac{n}{n+1} - \theta = -\theta \frac{1}{n+1}.$$

Для больших $n$ имеем

$$\mathrm{V}(\theta) \approx \frac{\theta^2}{n^2}.$$

## Задача 10.4

Выведите оценки максимального правдоподобия для параметров $\boldsymbol{\mu}$ и $\boldsymbol{\Omega}$ многомерного нормального распределения по выборке размера $n$.

## Решение

Итак, у нас есть $n$ независимых одинаково распределенных нормальных случайных $k\times 1$ векторов $\boldsymbol{y}_i \sim N(\boldsymbol{\mu}, \boldsymbol{\Omega})$. Каждый из них имеет распределение с плотностью

$$h(\boldsymbol{y}) = (2\pi)^{-k/2}|\boldsymbol{\Omega}|^{-1/2}\exp\left\{-\frac{1}{2}(\boldsymbol{y}-\boldsymbol{\mu})'\boldsymbol{\Omega}^{-1}(\boldsymbol{y}-\boldsymbol{\mu})\right\}.$$

Здесь $\boldsymbol{\mu}$ — $k\times 1$ вектор, а $\boldsymbol{\Omega}$ — $k\times k$ симметричная положительно определенная матрица. Таким образом, общее количество неизвестных параметров равно $k + k(k+1)/2 = k(k+3)/2$.

С учетом того, что $|\boldsymbol{\Omega}^{-1}| = |\boldsymbol{\Omega}|^{-1}$, мы можем записать логарифм функции правдоподобия как

$$\begin{aligned}\ln L(\boldsymbol{\mu}, \boldsymbol{\Omega}) &= \ln \prod_{i=1}^{n} h(\boldsymbol{y}_i) \\ &= -\frac{kn}{2}\ln 2\pi + \frac{n}{2}\ln|\boldsymbol{\Omega}^{-1}| - \frac{1}{2}\sum_{i=1}^{n}(\boldsymbol{y}_i-\boldsymbol{\mu})'\boldsymbol{\Omega}^{-1}(\boldsymbol{y}_i-\boldsymbol{\mu}).\end{aligned}$$

Нам здесь удобно рассматривать $\ln L$ как функцию $\boldsymbol{\mu}$ и $\boldsymbol{\Omega}^{-1}$. Приравняем нулю частные производные функции $\ln L$ по $\boldsymbol{\mu}$ и $\boldsymbol{\Omega}^{-1}$ (при этом мы будем использовать две леммы, сформулированные в конце решения задачи).

Используя формулу (ЛА.24), получаем производную по $\boldsymbol{\mu}$:

$$\frac{\partial \ln L}{\partial \boldsymbol{\mu}} = -\frac{1}{2}\sum_{i=1}^{n} 2\,\boldsymbol{\Omega}^{-1}(\boldsymbol{y}_i-\boldsymbol{\mu}) = 0.$$

Отсюда

$$\boldsymbol{\Omega}^{-1}\sum_{i=1}^{n}(\boldsymbol{y}_i-\boldsymbol{\mu}) = 0$$

и

$$\widehat{\boldsymbol{\mu}} = \frac{1}{n}\sum_{i=1}^{n}\boldsymbol{y}_i = \overline{\boldsymbol{y}}.$$

Используя лемму 1 и лемму 2, получаем производную по $\Omega^{-1}$:

$$\frac{\partial \ln L}{\partial \Omega^{-1}} = \frac{n}{2}\Omega - \frac{1}{2}\sum_{i=1}^{n}(y_i - \mu)(y_i - \mu)'.$$

Оценка максимального правдоподобия матрицы ковариаций равна соответственно:

$$\widehat{\Omega} = \frac{1}{n}\sum_{i=1}^{n}(y_i - \overline{y})(y_i - \overline{y})'.$$

*Лемма* 1. Пусть $A$ — $k \times k$ матрица, а $x$ — $k \times 1$ вектор. Тогда

$$\frac{\partial x'Ax}{\partial A} = xx'.$$

*Доказательство.* Вычислим производную $x'Ax$ по $a_{ij}$:

$$\frac{\partial x'Ax}{\partial a_{ij}} = \frac{\partial}{\partial a_{ij}}\sum x_t a_{ts} x_s = x_i x_j = (xx')_{ij}.$$

*Лемма* 2. Пусть $A$ — симметричная невырожденная $k \times k$ матрица, а $|A|$ — ее определитель. Тогда

$$\frac{\partial \ln|A|}{\partial A} = A^{-1}.$$

*Доказательство.* Из разложения определителя по строке (приложение ЛА, п. 9)

$$|A| = \sum_{j=1}^{n} a_{ij}(-1)^{i+j}|M_{ij}|$$

($M_{ij}$ – миноры матрицы), получаем:

$$\frac{\partial|A|}{\partial a_{ij}} = (-1)^{i+j}|M_{ij}|.$$

Формула для элементов $a^{ij}$ обратной матрицы $A^{-1}$ имеет вид

$$a^{ij} = \frac{(-1)^{i+j}|M_{ji}|}{|A|}.$$

Комбинируя две последние формулы, получаем

$$\frac{\partial \ln|A|}{\partial a_{ij}} = a^{ji}, \quad \text{или} \quad \frac{\partial \ln|A|}{\partial A} = (A^{-1})'.$$

Учитывая симметричность матрицы $A$, получаем утверждение леммы.

## Задача 10.5

Пусть $L_n(\theta)$ — функция правдоподобия. Докажите, что

$$\mathrm{E}\left(\frac{\partial \ln L_n(\theta)}{\partial \theta}\right) = 0, \qquad \mathrm{V}\left(\frac{\partial \ln L_n(\theta)}{\partial \theta}\right) = \mathcal{F}_n(\theta).$$

## Решение

Напомним, что функция правдоподобия $L_n(\theta)$ (будем для сокращения выкладок обозначать ее в дальнейшем просто $L$) является также функцией плотности совместного распределения наблюдений. Поэтому $\int L\,dy = 1$ для всех $\theta$, и, следовательно, все частные производные $\int L\,dy$ по $\theta$ равны 0. В частности,

$$\mathbf{0} = \frac{\partial}{\partial \theta}\int L\,dy = \int \frac{\partial}{\partial \theta} L\,dy = \int \left(\frac{\partial}{\partial \theta} L\right)\frac{1}{L} L\,dy = \mathrm{E}\left(\frac{\partial \ln L}{\partial \theta}\right).$$

(Здесь мы предполагаем, что можно дифференцировать по $\theta$ под знаком интеграла по $y$, что верно при выполнении некоторых «условий регулярности».) Первая часть задачи доказана.

Для доказательства второй части проведем сначала следующие выкладки:

$$\begin{aligned}\frac{\partial^2 \ln L}{\partial \theta\,\partial \theta'} &= \frac{\partial}{\partial \theta}\left(\frac{\partial \ln L}{\partial \theta'}\right) = \frac{\partial}{\partial \theta}\left(\frac{1}{L}\frac{\partial L}{\partial \theta'}\right) = \frac{\partial}{\partial \theta}\left(\frac{1}{L}\right)\frac{\partial L}{\partial \theta'} + \frac{1}{L}\frac{\partial^2 L}{\partial \theta\,\partial \theta'} \\ &= -\frac{1}{L^2}\frac{\partial L}{\partial \theta}\frac{\partial L}{\partial \theta'} + \frac{1}{L}\frac{\partial^2 L}{\partial \theta\,\partial \theta'} = -\frac{\partial \ln L}{\partial \theta}\frac{\partial \ln L}{\partial \theta'} + \frac{1}{L}\frac{\partial^2 L}{\partial \theta\,\partial \theta'}.\end{aligned}$$

Отсюда получаем тождество

$$\frac{1}{L}\frac{\partial^2 L}{\partial \theta\,\partial \theta'} = \frac{\partial^2 \ln L}{\partial \theta\,\partial \theta'} + \left(\frac{\partial \ln L}{\partial \theta}\right)\left(\frac{\partial \ln L}{\partial \theta}\right)'. \tag{$*$}$$

Вторые частные производные $\int L\,dy$ по $\theta$ также равны 0.

$$\begin{aligned}\mathbf{0} &= \frac{\partial^2}{\partial \theta\,\partial \theta'}\int L\,dy = \int \frac{\partial^2}{\partial \theta\,\partial \theta'} L\,dy = \int \frac{1}{L}\frac{\partial^2 L}{\partial \theta\,\partial \theta'} L\,dy = \mathrm{E}\left(\frac{1}{L}\frac{\partial^2 L}{\partial \theta\,\partial \theta'}\right) \\ &\quad \text{(используем тождество } (*)\text{)} \\ &= \mathrm{E}\left(\frac{\partial^2 \ln L}{\partial \theta\,\partial \theta'}\right) + \mathrm{E}\left(\left(\frac{\partial \ln L}{\partial \theta}\right)\left(\frac{\partial \ln L}{\partial \theta}\right)'\right).\end{aligned}$$

В последнем выражении первое слагаемое равно $-\mathcal{F}_n(\theta)$ по определению (см. (10.6), (10.7)), а второе равно

$$\mathrm{E}\left(\left(\frac{\partial \ln L}{\partial \theta}\right)\left(\frac{\partial \ln L}{\partial \theta}\right)'\right) = \mathrm{V}\left(\frac{\partial \ln L}{\partial \theta}\right),$$

поскольку, как следует из первой части задачи,

$$\mathrm{E}\left(\frac{\partial \ln L}{\partial \theta}\right) = \mathbf{0}.$$

Таким образом, получаем

$$\mathcal{F}_n(\theta) = \mathrm{V}\left(\frac{\partial \ln L}{\partial \theta}\right).$$

## Задача 10.6

Пусть $y_1, \ldots, y_n$ — независимые, одинаково распределенные случайные величины, равномерно распределенные на интервале $(\theta, 2\theta)$. Покажите, что:

а) оценка максимального правдоподобия есть $\widehat{\theta} = \max y_i/2$;

б) $\widehat{\theta}$ является смещенной, но асимптотически несмещенной;

в) $\mathrm{V}(\widehat{\theta})$ асимптотически равна $\theta^2/(4n^2)$.

## Решение

а) По условию задачи случайные величины $y_1, y_2, \ldots, y_n$ независимы и равномерно распределены на интервале $[\theta, 2\theta]$. (Заметим, что поскольку распределение непрерывное, мы вольны включать или исключать крайние точки интервала по соображениям удобства.) Функция распределения случайной величины $y$ имеет вид

$$F_y(z) = \frac{z-\theta}{\theta}, \quad \theta \leqslant z \leqslant 2\theta.$$

($F_y(z) = 0$ при $z < \theta$ и $F_y(z) = 1$ при $z > 2\theta$.) Функция правдоподобия равна

$$L(\theta) = \prod_{i=1}^{n} h(y_i, \theta) = \frac{1}{\theta^n}, \quad \theta \leqslant y_i \leqslant 2\theta,$$

и равна нулю при других наборах $y_i$.

Мы видим, что $L$ монотонно убывает с ростом $\theta$. Максимум этой функции не может быть найден дифференцированием. Ясно, однако, что $\theta$ должно быть выбрано самое малое из возможных.

Поскольку $y_i/2 \leqslant \theta \leqslant y_i$ для всех $i$, это означает, что

$$\frac{1}{2}\max y_i \leqslant \theta \leqslant \min y_i,$$

и оценка максимального правдоподобия равна

$$\widehat{\theta} = \frac{1}{2}\max y_i.$$

б) Для того чтобы вычислить математическое ожидание и дисперсию оценки $\widehat{\theta}$, найдем ее функцию распределения $F(z)$ и функцию плотности распределения $f(z)$.

$$F(z) = \mathrm{P}(\widehat{\theta} \leqslant z) = \mathrm{P}(2\widehat{\theta} \leqslant 2z) = \mathrm{P}(\max y_i \leqslant 2z) = \prod_{i=1}^{n} \mathrm{P}(y_i \leqslant 2z)$$

$$= F_y(2z)^n = \left(\frac{2z-\theta}{\theta}\right)^n = \theta^{-n}(2z-\theta)^n, \qquad \theta/2 \leqslant z \leqslant \theta.$$

(Мы использовали здесь независимость $y_i$.) Функция плотности распределения находится дифференцированием:

$$f(z) = F'(z) = 2n\theta^{-n}(2z-\theta)^{n-1}, \qquad \theta/2 \leqslant z \leqslant \theta.$$

Отсюда получаем:

$$\mathrm{E}(\widehat{\theta}) = \mathrm{E}\left(\frac{1}{2}\max y_i\right) = \int_{-\infty}^{\infty} zf(z)dz = 2n\theta^{-n}\int_{\theta/2}^{\theta} z(2z-\theta)^{n-1}dz$$

(замена переменной: $t = 2z - \theta$)

$$= \frac{1}{2}n\theta^{-n}\int_0^{\theta}(t+\theta)t^{n-1}dt = \frac{1}{2}n\theta^{-n}\left(\int_0^{\theta} t^n dt + \theta\int_0^{\theta} t^{n-1}dt\right)$$

$$= \frac{1}{2}\theta\left(\frac{n}{n+1}+1\right) = \theta\,\frac{2n+1}{2n+2} = \theta\,\frac{n+1/2}{n+1} = \theta - \theta\,\frac{1}{2n+2}.$$

Отсюда мы видим, что смещение оценки $\widehat{\theta}$ равно

$$\mathrm{bias}(\widehat{\theta}) = \mathrm{E}(\widehat{\theta}) - \theta = -\theta\,\frac{1}{2n+2},$$

откуда следует, что оценка $\widehat{\theta}$ является смещенной, но асимптотически несмещенной.

в) Вычислим теперь дисперсию оценки $\widehat{\theta}$.

$$\begin{aligned}
\mathrm{E}(\widehat{\theta}^2) &= \int_{-\infty}^{\infty} z^2 f(z)dz = 2n\theta^{-n}\int_{\theta/2}^{\theta} z^2(2z-\theta)^{n-1}dz \\
&\quad\text{(замена переменной: } t = 2z-\theta) \\
&= \frac{1}{4}n\theta^{-n}\int_0^{\theta}(t+\theta)^2 t^{n-1}dt \\
&= \frac{1}{4}n\theta^{-n}\left(\int_0^{\theta} t^{n+1}dt + 2\theta\int_0^{\theta} t^n dt + \theta^2\int_0^{\theta} t^{n-1}dt\right) \\
&= \frac{1}{4}\theta^2\left(\frac{n}{n+2}+\frac{2n}{n+1}+1\right) = \theta^2\frac{4n^2+8n+2}{4(n+1)(n+2)}, \\
\mathrm{V}(\widehat{\theta}) &= \mathrm{E}(\widehat{\theta}^2) - (\mathrm{E}\widehat{\theta})^2 = \theta^2\left(\frac{4n^2+8n+2}{4(n+1)(n+2)} - \frac{(2n+1)^2}{4(n+1)^2}\right) \\
&= \theta^2\frac{n}{4(n+1)^2(n+2)}.
\end{aligned}$$

Из последнего равенства следует, что $\mathrm{V}(\widehat{\theta})$ асимптотически равна $\theta^2/(4n^2)$.

## Задача 10.7

(Продолжение задачи 10.6) Рассмотрим альтернативную оценку:

$$\widetilde{\theta} = (\min y_i + 2\max y_i)/5.$$

а) Покажите, что $\mathrm{V}(\widetilde{\theta})$ асимптотически равна $\theta^2/(5n^2)$.

б) Покажите, что $\widetilde{\theta}$ более эффективна, чем $\widehat{\theta}$.

в) Противоречит ли это асимптотической эффективности оценки максимального правдоподобия?

## Решение

а) В этой задаче нам потребуется функция совместного распределения минимума и максимума набора независимых одинаково распределенных случайных величин. Обозначим

$$V_n = \min y_i, \qquad W_n = \max y_i.$$

Функция совместного распределения $V_n$ и $W_n$ равна

$$\begin{aligned}
F(v,w) &= \mathrm{P}(V_n \leqslant v,\, W_n \leqslant w) = \mathrm{P}(W_n \leqslant w) - \mathrm{P}(V_n > v,\, W_n \leqslant w) \\
&= (\mathrm{P}(y_i \leqslant w))^n - (\mathrm{P}(v < y_i \leqslant w))^n \\
&= \left(\frac{w-\theta}{\theta}\right)^n - \left(\frac{w-v}{\theta}\right)^n = \theta^{-n}\left((w-\theta)^n - (w-v)^n\right).
\end{aligned}$$

Плотность совместного распределения находится дифференцированием:

$$f(v,w) = \frac{\partial^2 F(v,w)}{\partial v\,\partial w} = n(n-1)\theta^{-n}(w-v)^{n-2}.$$

Плотности распределений $V_n$ и $W_n$ вычисляются следующим способом:

$$f_v(v) = \int_v^{2\theta} f(v,w)\,dw = n\theta^{-n}(2\theta - v)^{n-1},$$
$$f_w(w) = \int_\theta^{w} f(v,w)\,dv = n\theta^{-n}(w - \theta)^{n-1}.$$

(Все вычисления аналогичны вычислениям из задачи 10.6.)

Вычислим математическое ожидание и дисперсию $V_n$.

$$\mathrm{E}(V_n) = \int_{-\infty}^{\infty} v f_v(v)\,dv = n\theta^{-n}\int_\theta^{2\theta} v(2\theta - v)^{n-1}\,dv$$

(замена переменной: $t = 2\theta - v$)

$$= n\theta^{-n}\int_0^\theta (2\theta - t)t^{n-1}\,dt = n\theta^{-n}\left(2\theta\frac{\theta^n}{n} - \frac{\theta^{n+1}}{n+1}\right)$$
$$= \theta\,\frac{n+2}{n+1},$$

$$\mathrm{E}(V_n^2) = \int_{-\infty}^{\infty} v^2 f_v(v)\,dv = n\theta^{-n}\int_\theta^{2\theta} v^2(2\theta - v)^{n-1}\,dv$$

(замена переменной: $t = 2\theta - v$)

$$= n\theta^{-n}\int_0^\theta (2\theta - t)^2 t^{n-1}\,dt$$
$$= n\theta^2\left(4 - 4\frac{n}{n+1} + \frac{n}{n+2}\right) = \theta^2\frac{n^2+5n+8}{(n+1)(n+2)},$$
$$\mathrm{V}(V_n) = \mathrm{E}(V_n^2) - (\mathrm{E}V_n)^2 = \theta^2\frac{n^2+5n+8}{(n+1)(n+2)} - \theta^2\frac{(n+2)^2}{(n+1)^2}$$
$$= \theta^2\frac{n}{(n+1)^2(n+2)}.$$

Аналогично можно вычислить математическое ожидание и дисперсию $W_n$, впрочем, фактически это уже было сделано в задаче 10.6.

$$\mathrm{E}(W_n) = \theta\,\frac{2n+1}{n+1}, \qquad \mathrm{V}(W_n) = \theta^2\frac{n}{(n+1)^2(n+2)}.$$

Осталось еще вычислить ковариацию $V_n$ и $W_n$. Вычислим сначала математическое ожидание произведения $V_n$ и $W_n$:

$$
\begin{aligned}
\mathrm{E}(V_n W_n) &= \int_\theta^{2\theta} dw \int_\theta^w vwf(v,w)\,dv \\
&= n(n-1)\theta^{-n} \int_\theta^{2\theta} w\,dw \int_\theta^w v(w-v)^{n-2}\,dv \\
&\qquad \text{(замена переменной: } z = w - v\text{)} \\
&= n(n-1)\theta^{-n} \int_\theta^{2\theta} w\,dw \int_0^{w-\theta} (w-z)z^{n-2}\,dz \\
&= n(n-1)\theta^{-n} \int_\theta^{2\theta} w\,dw \left( w\frac{1}{n-1}(w-\theta)^{n-1} - \frac{1}{n}(w-\theta)^n \right) \\
&\qquad \text{(замена переменной: } z = w - \theta\text{)} \\
&= \frac{n(n-1)}{\theta^n}\left( \frac{1}{n-1}\int_0^\theta (\theta+z)^2 z^{n-1}\,dz - \frac{1}{n}\int_0^\theta (\theta+z)z^n\,dz \right) \\
&= \theta^2 \left( n\left(\frac{1}{n} + \frac{2}{n+1} + \frac{1}{n+2}\right) - (n-1)\left(\frac{1}{n+1} + \frac{1}{n+2}\right) \right) \\
&= \theta^2\,\frac{2n+5}{n+2}.
\end{aligned}
$$

Ковариация $V_n$ и $W_n$ равна:

$$
\begin{aligned}
\mathrm{Cov}(V_n, W_n) &= \mathrm{E}(V_n W_n) - \mathrm{E}(V_n)\mathrm{E}(W_n) \\
&= \theta^2\,\frac{2n+5}{n+2} - \theta^2\,\frac{n+2}{n+1}\,\frac{2n+1}{n+1} = \theta^2\,\frac{1}{(n+1)^2(n+2)}.
\end{aligned}
$$

Рассмотрим теперь оценку $\widetilde{\theta} = \lambda V_n + \mu W_n$, более общего вида, чем оценка в условии задачи. Из предыдущих результатов следует, что при $n \to \infty$

$$
\mathrm{E}(\widetilde{\theta}) \to (\lambda + 2\mu)\theta.
$$

Наложим ограничение $\lambda + 2\mu = 1$, при котором $\widetilde{\theta}$ является асимптотически несмещенной оценкой. Дисперсия этой оценки, альтернативной к оценке максимального правдоподобия, равна:

$$
\begin{aligned}
\mathrm{V}(\widetilde{\theta}) &= \mathrm{V}\left((1-2\mu)V_n + \mu W_n\right) \\
&= (1-2\mu)^2\mathrm{V}(V_n) + \mu^2\mathrm{V}(V_n) + 2\mu(1-2\mu)\,\mathrm{Cov}(V_n, W_n).
\end{aligned}
$$

Дифференцируя это выражение по $\mu$ и приравняв производную нулю, получаем «оптимальное» значение $\mu$, при котором дисперсия оценки $\widetilde{\theta}$ минимальна:

$$
\mu = \frac{2\mathrm{V}(V_n) - \mathrm{Cov}(V_n, W_n)}{4\mathrm{V}(V_n) + \mathrm{V}(W_n) - 4\,\mathrm{Cov}(V_n, W_n)} = \frac{2n-1}{5n-4}.
$$

Отсюда видно, что наименьшая дисперсия (асимптотически) достигается при $\mu = 2/5$, т.е. когда оценка $\widetilde{\theta}$ совпадает с указанной в условии задачи. При этом сама эта дисперсия (асимптотически) равна

$$\mathrm{V}(\widetilde{\theta}) \approx \frac{\theta^2}{5n^2}.$$

б) Утверждение вытекает из полученной выше формулы и формулы, полученной в задаче 10.6:

$$\mathrm{V}(\widetilde{\theta}) \approx \frac{\theta^2}{5n^2} < \frac{\theta^2}{4n^2} \approx \mathrm{V}(\widehat{\theta}).$$

в) Кажущийся парадокс состоит в том, что оценка $\widetilde{\theta}$ более эффективна, чем оценка максимального правдоподобия $\widehat{\theta}$, которая, как мы знаем, асимптотически наиболее эффективна. Объяснение состоит в том, что в данном случае равномерного распределения случайной величины $y$ нарушено одно из условий «регулярности», лежащее в основе теоремы об асимптотической эффективности оценки максимального правдоподобия. (А именно: «область возможных значений исследуемой случайной величины $y$, в которой $h(y,\theta) \neq 0$, *не зависит от* $\theta$».)

## Задача 10.8

Дана линейная модель $\boldsymbol{y} = \boldsymbol{X\beta} + \boldsymbol{u}$, $\boldsymbol{u} \sim N(\boldsymbol{0}, \sigma^2 \boldsymbol{I}_n)$. Покажите, что

$$\mathrm{LM} = n(\mathrm{ESS_R} - \mathrm{ESS_{UR}})/\mathrm{ESS_R}, \qquad \mathrm{LR} = n\ln(\mathrm{ESS_R}/\mathrm{ESS_{UR}})$$

и

$$\mathrm{W} = n(\mathrm{ESS_R} - \mathrm{ESS_{UR}})/\mathrm{ESS_{UR}}.$$

Покажите, что выполняются неравенства $\mathrm{LM} \leqslant \mathrm{LR} \leqslant \mathrm{W}$.

## Решение

В решении этой задачи мы существенно используем результаты и обозначения из гл. 16 учебника. Напомним, что $\widehat{\boldsymbol{\beta}}$ и $\widetilde{\boldsymbol{\beta}}$ обозначают оценки коэффициентов в регрессии без ограничения и в регрессии с ограничением, остатки регрессий равны соответственно

$$\widehat{\boldsymbol{u}} = \boldsymbol{y} - \boldsymbol{X}\widehat{\boldsymbol{\beta}}, \qquad \widetilde{\boldsymbol{u}} = \boldsymbol{y} - \boldsymbol{X}\widetilde{\boldsymbol{\beta}}.$$

Суммы квадратов остатков и ML-оценки дисперсии ошибок равны

$$\mathrm{ESS_{UR}} = \widehat{\boldsymbol{u}}'\widehat{\boldsymbol{u}}, \quad \mathrm{ESS_R} = \widetilde{\boldsymbol{u}}'\widetilde{\boldsymbol{u}}, \quad \widehat{\sigma}^2 = \widehat{\boldsymbol{u}}'\widehat{\boldsymbol{u}}/n, \quad \widetilde{\sigma}^2 = \widetilde{\boldsymbol{u}}'\widetilde{\boldsymbol{u}}/n.$$

Поскольку $\boldsymbol{\Omega} = \sigma^2 \boldsymbol{I}$, выражения для $\widehat{\boldsymbol{\beta}}$ и $\widetilde{\boldsymbol{\beta}}$ не зависят от $\sigma^2$:

$$\widehat{\boldsymbol{\beta}} = (\boldsymbol{X}'\boldsymbol{X})^{-1}\boldsymbol{X}'\boldsymbol{y}, \qquad \widetilde{\boldsymbol{\beta}} = \widehat{\boldsymbol{\beta}} - (\boldsymbol{X}'\boldsymbol{X})^{-1}\boldsymbol{R}'\boldsymbol{P}^{-1}(\boldsymbol{R}\widehat{\boldsymbol{\beta}} - \boldsymbol{r}).$$

Здесь $P = R(X'X)^{-1}R'$, а ограничение на коэффициенты имеет вид $R\beta = r$.

Из формул (10.32) и (10.40) получаем:

$$\ln L(\widehat{\beta}, \widehat{\sigma}^2) = \text{constant } -(n/2)\ln \widehat{\sigma}^2,$$

$$\ln L(\widetilde{\beta}, \widetilde{\sigma}^2) = \text{constant } -(n/2)\ln \widetilde{\sigma}^2,$$

а из (10.48):

$$\widetilde{u}'\widetilde{u} - \widehat{u}'\widehat{u} = (\widehat{\beta} - \widetilde{\beta})'X'X(\widehat{\beta} - \widetilde{\beta}).$$

Поскольку

$$\widehat{\beta} - \widetilde{\beta} = (X'X)^{-1}R'P^{-1}(R\widehat{\beta} - r),$$

то

$$(R\widehat{\beta} - r)'(R(X'X)^{-1}R')^{-1}(R\widehat{\beta} - r) = \widetilde{u}'\widetilde{u} - \widehat{u}'\widehat{u}.$$

Так как $X'\widetilde{u} = R'P^{-1}(R\widehat{\beta} - r)$, то, умножив вектор $X'\widetilde{u}$ слева на $(X'X)^{-1/2}$ и вычислив его скалярный квадрат, получаем:

$$(R\widehat{\beta} - r)'(R(X'X)^{-1}R')^{-1}(R\widehat{\beta} - r) = \widetilde{u}'X(X'X)^{-1}X'\widetilde{u}.$$

Требуемые формулы для трех статистик получаются тогда из их определения (10.52)–(10.54).

Обозначим $x = \widetilde{\sigma}^2/\widehat{\sigma}^2$. Тогда

$$\mathrm{W} = n\frac{\widetilde{u}'\widetilde{u} - \widehat{u}'\widehat{u}}{\widehat{u}'\widehat{u}} = n(x-1),$$

$$\mathrm{LM} = n\frac{\widetilde{u}'\widetilde{u} - \widehat{u}'\widehat{u}}{\widetilde{u}'\widetilde{u}} = n\left(1 - \frac{1}{x}\right),$$

$$\mathrm{LR} = n\ln\frac{\widetilde{u}'\widetilde{u}}{\widehat{u}'\widehat{u}} = n\ln x.$$

Неравенства $\mathrm{LM} \leqslant \mathrm{LR} \leqslant \mathrm{W}$ в этих обозначениях выглядят как

$$1 - \frac{1}{x} \leqslant \ln x \leqslant x - 1,$$

и то, что они верны для всех $x > 0$, очевидно.

## Задача 10.9

Представим стандартную линейную модель в следующем виде:

$$\begin{bmatrix} y_1 \\ y_2 \\ y_3 \end{bmatrix} = \begin{bmatrix} X_1 \\ X_2 \\ X_3 \end{bmatrix}\beta + \begin{bmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \varepsilon_3 \end{bmatrix},$$

где $\boldsymbol{y}_i, \varepsilon_i$ — $n_i \times 1$ векторы, $\boldsymbol{X}_i$ — $n_i \times k$ матрицы, $\boldsymbol{\beta}$ — $k \times 1$ вектор, векторы $\varepsilon_i$ имеют нормальное распределение $N\left(\mathbf{0}, \sigma^2 \boldsymbol{I}_{n_i}\right)$ и независимы. Вектор $\boldsymbol{y}_3$ представляет собой пропущенные наблюдения зависимой переменной, а матрица $\boldsymbol{X}_2$ — пропущенные наблюдения независимых переменных.

Вычислите следующие оценки вектора $\boldsymbol{\beta}$ и сравните их свойства:

а) МНК-оценка только по полным наблюдениям $\boldsymbol{y}_1, \boldsymbol{X}_1$;

б) МНК-оценка при замене матрицы $\boldsymbol{X}_2$ на нулевую и исключении наблюдений $\boldsymbol{y}_3, \boldsymbol{X}_3$;

в) МНК-оценка по всей модели при замене $\boldsymbol{y}_3, \boldsymbol{X}_2$ на соответственно нулевой вектор и нулевую матрицу;

г) оценка максимального правдоподобия, предполагая $\boldsymbol{y}_3, \boldsymbol{X}_2$ неизвестными параметрами наряду с $\boldsymbol{\beta}$.

## Решение

Обозначим оценки из п. а), б), в) и г) соответственно $\widehat{\boldsymbol{\beta}}_a, \widehat{\boldsymbol{\beta}}_b, \widehat{\boldsymbol{\beta}}_c$ и $\widehat{\boldsymbol{\beta}}_d$. Первые три оценки рассчитываются по стандартной формуле:

$$\widehat{\boldsymbol{\beta}}_a = (\boldsymbol{X}_1'\boldsymbol{X}_1)^{-1}\boldsymbol{X}_1'\boldsymbol{y}_1,$$

$$\widehat{\boldsymbol{\beta}}_b = \left(\begin{bmatrix}\boldsymbol{X}_1\\ \mathbf{0}\end{bmatrix}'\begin{bmatrix}\boldsymbol{X}_1\\ \mathbf{0}\end{bmatrix}\right)^{-1}\begin{bmatrix}\boldsymbol{X}_1\\ \mathbf{0}\end{bmatrix}'\begin{bmatrix}\boldsymbol{y}_1\\ \boldsymbol{y}_2\end{bmatrix} = (\boldsymbol{X}_1'\boldsymbol{X}_1)^{-1}\boldsymbol{X}_1'\boldsymbol{y}_1 \equiv \widehat{\boldsymbol{\beta}}_a,$$

$$\widehat{\boldsymbol{\beta}}_c = \left(\begin{bmatrix}\boldsymbol{X}_1\\ \mathbf{0}\\ \boldsymbol{X}_3\end{bmatrix}'\begin{bmatrix}\boldsymbol{X}_1\\ \mathbf{0}\\ \boldsymbol{X}_3\end{bmatrix}\right)^{-1}\begin{bmatrix}\boldsymbol{X}_1\\ \mathbf{0}\\ \boldsymbol{X}_3\end{bmatrix}'\begin{bmatrix}\boldsymbol{y}_1\\ \boldsymbol{y}_2\\ \mathbf{0}\end{bmatrix} = (\boldsymbol{X}_1'\boldsymbol{X}_1 + \boldsymbol{X}_3'\boldsymbol{X}_3)^{-1}\boldsymbol{X}_1'\boldsymbol{y}_1.$$

Оценка $\widehat{\boldsymbol{\beta}}_d$ получается максимизацией (логарифмической) функции правдоподобия

$$\ln L = -\frac{n}{2}\ln 2\pi - \frac{n}{2}\ln\sigma^2 - \frac{1}{2\sigma^2}\sum e_t^2, \qquad \text{где}$$

$$\sum e_t^2 = (\boldsymbol{y}_1 - \boldsymbol{X}_1\boldsymbol{\beta})'(\boldsymbol{y}_1 - \boldsymbol{X}_1\boldsymbol{\beta}) + (\boldsymbol{y}_2 - \boldsymbol{X}_2\boldsymbol{\beta})'(\boldsymbol{y}_2 - \boldsymbol{X}_2\boldsymbol{\beta}) + (\boldsymbol{y}_3 - \boldsymbol{X}_3\boldsymbol{\beta})'(\boldsymbol{y}_3 - \boldsymbol{X}_3\boldsymbol{\beta})$$

по $\boldsymbol{\beta}$, $\boldsymbol{X}_2$, $\boldsymbol{y}_3$ и $\sigma^2$. Из вида этой функции видно, что можно отдельно минимизировать $\sum e_t^2$ по $\boldsymbol{\beta}$, $\boldsymbol{X}_2$, $\boldsymbol{y}_3$, а затем найти $\sigma^2$. Все три слагаемых в $\sum e_t^2$ неотрицательны. Третье слагаемое всегда можно обратить в 0, выбрав $\boldsymbol{y}_3 = \boldsymbol{X}_3\boldsymbol{\beta}$. Второе слагаемое также можно обратить в 0 подходящим выбором $\boldsymbol{X}_2$. (Рассмотрим геометрическую интерпретацию $\boldsymbol{X}_2$ как линейного отображения пространства $R^k$ в $R^{n_2}$. Очевидно, всегда существует отображение, переводящее вектор $\boldsymbol{\beta}$ в вектор $\boldsymbol{y}_2$.) Таким образом, минимизация $\sum e_t^2$ сводится к минимизации первого слагаемого, а следовательно, ответом является стандартная МНК-оценка по первым $n_1$ наблюдениям:

$$\widehat{\boldsymbol{\beta}}_d = (\boldsymbol{X}_1'\boldsymbol{X}_1)^{-1}\boldsymbol{X}_1'\boldsymbol{y}_1 \equiv \widehat{\boldsymbol{\beta}}_a.$$

Таким образом, осталось только сравнить свойства оценок $\widehat{\beta}_a$ и $\widehat{\beta}_c$. Их математические ожидания и матрицы ковариаций легко вычисляются:

$$\begin{aligned}
&\mathrm{E}(\widehat{\beta}_a) = \beta, \quad \mathrm{V}(\widehat{\beta}_a) = \sigma^2(X_1'X_1)^{-1};\\
&\mathrm{E}(\widehat{\beta}_c) = (X_1'X_1 + X_3'X_3)^{-1}X_1'X_1\beta,\\
&\mathrm{V}(\widehat{\beta}_c) = \sigma^2(X_1'X_1 + X_3'X_3)^{-1}X_1'X_1(X_1'X_1 + X_3'X_3)^{-1}.
\end{aligned}$$

Таким образом, оценка $\widehat{\beta}_a$ является несмещенной, а смещение оценки $\widehat{\beta}_c$ равно

$$\mathrm{bias}(\widehat{\beta}_c) = \mathrm{E}(\widehat{\beta}_c) - \beta = -(X_1'X_1 + X_3'X_3)^{-1}X_3'X_3\beta.$$

Для того чтобы упростить дальнейшие выкладки, введем следующие обозначения:

$$A = X_1'X_1, \qquad C = X_3'X_3.$$

Очевидно, что матрицы $A$, $C$, $A+C$, а также им обратные симметричны и положительно определены. В этих обозначениях

$$\mathrm{V}(\widehat{\beta}_a) = \sigma^2 A^{-1}, \qquad \mathrm{V}(\widehat{\beta}_c) = \sigma^2(A+C)^{-1}A(A+C)^{-1}.$$

Рассмотрим разность

$$\begin{aligned}
\mathrm{V}(\widehat{\beta}_a) - \mathrm{V}(\widehat{\beta}_c) &= \sigma^2(A^{-1} - (A+C)^{-1}A(A+C)^{-1})\\
&= \sigma^2(A+C)^{-1}\left((A+C)A^{-1}(A+C) - A\right)(A+C)^{-1}\\
&= \sigma^2(A+C)^{-1}(2C + CA^{-1}C)(A+C)^{-1} \geqslant \mathbf{0}.
\end{aligned}$$

Таким образом, дисперсия оценки $\widehat{\beta}_c$ меньше, чем дисперсия оценки $\widehat{\beta}_a$.

Попытаемся сравнить матрицы среднеквадратичных ошибок двух оценок:

$$\begin{aligned}
\mathrm{MSE}(\widehat{\beta}_a) &= \mathrm{V}(\widehat{\beta}_a) = \sigma^2 A^{-1},\\
\mathrm{MSE}(\widehat{\beta}_c) &= \mathrm{V}(\widehat{\beta}_c) + \mathrm{bias}(\widehat{\beta}_c)\,\mathrm{bias}(\widehat{\beta}_c)'\\
&= \mathrm{V}(\widehat{\beta}_c) + (A+C)^{-1}C\beta\beta'C(A+C)^{-1}.
\end{aligned}$$

Вычислим разность

$$\begin{aligned}
&\mathrm{MSE}(\widehat{\beta}_a) - \mathrm{MSE}(\widehat{\beta}_c)\\
&\quad = \mathrm{V}(\widehat{\beta}_a) - \mathrm{V}(\widehat{\beta}_c) - (A+C)^{-1}C\beta\beta'C(A+C)^{-1}\\
&\quad = \sigma^2(A+C)^{-1}(2C + CA^{-1}C)(A+C)^{-1}\\
&\qquad - (A+C)^{-1}C\beta\beta'C(A+C)^{-1}\\
&\quad = (A+C)^{-1}C\left(\sigma^2(2C^{-1}+A^{-1}) - \beta\beta'\right)C(A+C)^{-1}.
\end{aligned}$$

Таким образом, разность $\mathrm{MSE}(\widehat{\beta}_a) - \mathrm{MSE}(\widehat{\beta}_c)$ может быть положительно или отрицательно определена одновременно с матрицей

$$\sigma^2(2C^{-1}+A^{-1}) - \beta\beta',$$

про которую ничего определенного сказать нельзя.

## Задача 10.10

Известно, что в модели множественной регрессии $y = X\beta + \varepsilon$ имеется гетероскедастичность, причем

$$\begin{aligned}
\mathrm{V}(\varepsilon_t) &= \sigma_1^2, & t &= 1, \ldots, n_1, \\
\mathrm{V}(\varepsilon_t) &= \sigma_2^2, & t &= n_1+1, \ldots, n_1+n_2 \ (n = n_1+n_2), \\
\mathrm{E}(\varepsilon_t\varepsilon_s) &= 0, & t &\neq s.
\end{aligned}$$

В предположении нормальности вектора ошибок постройте тест отношения правдоподобия (LR-test) для проверки гипотезы $\mathrm{H}_0\colon \sigma_1^2 = \sigma_2^2$.

## Решение

Проверочная статистика теста отношения правдоподобия есть

$$\mathrm{LR} = -2(\ln L(\widetilde{\beta}, \widetilde{\sigma}^2) - \ln L(\widehat{\beta}, \widehat{\sigma}_1^2, \widehat{\sigma}_2^2)),$$

где $\widetilde{\beta}$, $\widetilde{\sigma}^2$ — оценки в задаче с ограничением ($\sigma_1^2 = \sigma_2^2 = \sigma$), а $\widehat{\beta}$, $\widehat{\sigma}_1^2$, $\widehat{\sigma}_2^2$ — оценки в задаче без ограничения.

Найдем выражение для $\ln L(\widetilde{\beta}, \widetilde{\sigma}^2)$. Функция правдоподобия имеет следующий вид (см. (10.19)):

$$\ln L(\beta, \sigma^2) = -\frac{n}{2}\ln 2\pi - \frac{n}{2}\ln\sigma^2 - \frac{1}{2\sigma^2}\sum_{t=1}^{n}(y_t - x_t'\beta)^2.$$

Обозначим $\widetilde{e} = y - X\widetilde{\beta}$ и $\widetilde{\sigma}^2 = \widetilde{e}'\widetilde{e}/n$ вектор остатков и оценку максимального правдоподобия дисперсии ошибок в регрессии с ограничением. Получаем:

$$\begin{aligned}
\ln L(\widetilde{\beta}, \widetilde{\sigma}^2) &= -\frac{n}{2}\ln 2\pi - \frac{n}{2}\ln\widetilde{\sigma}^2 - \frac{1}{2}\frac{\widetilde{e}'\widetilde{e}}{\widetilde{\sigma}^2} = -\frac{n}{2}\ln 2\pi - \frac{n}{2}\ln\widetilde{\sigma}^2 - \frac{n\widetilde{\sigma}^2}{2\widetilde{\sigma}} \\
&= -\frac{n}{2}\ln 2\pi - \frac{n}{2}\ln\widetilde{\sigma}^2 - \frac{n}{2}.
\end{aligned}$$

В задаче без ограничения логарифмическая функция правдоподобия равна

$$\begin{aligned}
\ln L(\beta, \sigma_1^2, \sigma_2^2) = &-\frac{n}{2}\ln 2\pi - \frac{n_1}{2}\ln\sigma_1^2 - \frac{n_2}{2}\ln\sigma_2^2 \\
&- \frac{1}{2\sigma_1^2}\sum_{t=1}^{n_1}(y_t - x_t'\beta)^2 - \frac{1}{2\sigma_2^2}\sum_{t=n_1+1}^{n_1+n_2}(y_t - x_t'\beta)^2.
\end{aligned}$$

Дифференцируя логарифмическую функцию правдоподобия по $\sigma_1^2$ и $\sigma_2^2$, получаем:

$$\widehat{\sigma}_1^2 = \frac{1}{n_1}\sum_{t=1}^{n_1}(y_t - x_t'\beta)^2,$$

$$\widehat{\sigma}_2^2 = \frac{1}{n_2}\sum_{t=n_1+1}^{n_1+n_2}(y_t - x_t'\beta)^2.$$

Значение логарифмической функции правдоподобия для задачи без ограничения в точке максимума равно

$$\ln L(\widehat{\beta}, \widehat{\sigma}_1^2, \widehat{\sigma}_2^2) = -\frac{n}{2}\ln 2\pi - \frac{n_1}{2}\ln\widehat{\sigma}_1^2 - \frac{n_2}{2}\ln\widehat{\sigma}_2^2 - \frac{1}{2}\sum_{t=1}^{n_1}\frac{e_t^2}{\widehat{\sigma}_1^2} - \frac{1}{2}\sum_{t=n_1+1}^{n_1+n_2}\frac{e_t^2}{\widehat{\sigma}_2^2}$$

$$= -\frac{n}{2}\ln 2\pi - \frac{n_1}{2}\ln\widehat{\sigma}_1^2 - \frac{n_2}{2}\ln\widehat{\sigma}_2^2 - \frac{n_1\widehat{\sigma}_1^2}{2\widehat{\sigma}_1^2} - \frac{n_2\widehat{\sigma}_2^2}{2\widehat{\sigma}_2^2}$$

$$= -\frac{n}{2}\ln 2\pi - \frac{n_1}{2}\ln\widehat{\sigma}_1^2 - \frac{n_2}{2}\ln\widehat{\sigma}_2^2 - \frac{n}{2}.$$

Здесь $e_t$ — остатки в регрессии без ограничения. Оценки дисперсий равны, соответственно,

$$\widehat{\sigma}_1^2 = \frac{1}{n_1}\sum_{t=1}^{n_1} e_t^2 \quad \text{и} \quad \widehat{\sigma}_2^2 = \frac{1}{n_2}\sum_{t=n_1+1}^{n_1+n_2} e_t^2.$$

Отсюда вычисляем значение тестовой статистики:

$$\mathrm{LR} = -2(\ln L(\widetilde{\beta}, \widetilde{\sigma}^2) - \ln L(\widehat{\beta}, \widehat{\sigma}_1^2, \widehat{\sigma}_2^2))$$

$$= -2\frac{n}{2}\ln 2\pi - 2\frac{n}{2}\ln\widetilde{\sigma}^2 - 2\frac{n}{2} + 2\frac{n}{2}\ln 2\pi + 2\frac{n_1}{2}\ln\widehat{\sigma}_1^2 + 2\frac{n_2}{2}\ln\widehat{\sigma}_2^2 + 2\frac{n}{2}$$

$$= -n\ln\widetilde{\sigma}^2 + n_1\ln\widehat{\sigma}_1^2 + n_2\ln\widehat{\sigma}_2^2.$$

## Задача 10.11

Дана линейная модель $y = X\beta + \varepsilon$, где $y$ — $n \times 1$ вектор, $\beta$ — $k \times 1$ вектор, $\varepsilon$ — $n \times 1$ вектор, $X$ — $n \times k$ матрица, $\varepsilon \sim N(0, \Omega)$ и

$$\Omega = \begin{bmatrix} \sigma_1^2 I_{n_1} & 0 & \dots & 0 \\ 0 & \sigma_2^2 I_{n_2} & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \sigma_r^2 I_{n_r} \end{bmatrix}, \quad n_1 + n_2 + \dots + n_r = n$$

(групповая гетероскедастичность). Как выглядит LR-тест (тест отношения правдоподобия) для проверки гипотезы $\mathrm{H}_0$: $\sigma_1^2 = \sigma_2^2 = \dots = \sigma_r^2$?
*Указание.* Получите ответ в терминах ML-оценок дисперсий $\sigma^2$.

## Решение

Решение задачи дословно повторяет решение предыдущей задачи 10.10. Соответствующая статистика имеет вид

$$\mathrm{LR} = -n \ln \widehat{\sigma}^2 + n_1 \ln \widehat{\sigma}_1^2 + \cdots + n_2 \ln \widehat{\sigma}_2^2.$$

## Задача 10.12

Пусть $p$ — вероятность выпадения орла при бросании монеты. Из $n = 100$ испытаний $x = 42$ раза выпал орел и 58 — решка. Тестируйте на 5%-ном уровне значимости гипотезу $\mathrm{H}_0\colon p = 0.5$:

а) при помощи теста Вальда (W);

б) при помощи теста множителей Лагранжа (LM);

в) при помощи теста отношения правдоподобия (LR).

## Решение

Логарифмическая функция правдоподобия для данной задачи равна

$$l = \ln L = x \ln p + (n - x) \ln(1 - p).$$

Уравнение правдоподобия:

$$\frac{\partial l}{\partial p} = \frac{x}{p} - \frac{n-x}{1-p} = \frac{x - np}{p(1-p)} = 0,$$

откуда получаем оценку максимального правдоподобия $\widehat{p} = \dfrac{x}{n} = \dfrac{42}{100} = 0.42.$

При построении тестов нам еще понадобится количество информации Фишера для данной задачи:

$$I_n(p) = -\mathrm{E}\left(\frac{\partial^2 l}{\partial p^2}\right) = \mathrm{E}\left(\frac{x}{p^2} + \frac{n-x}{(1-p)^2}\right) = \frac{\mathrm{E}x}{p^2} + \frac{n - \mathrm{E}x}{(1-p)^2}$$
$$= \frac{np}{p^2} + \frac{n(1-p)}{(1-p)^2} = \frac{n}{p(1-p)}.$$

Кроме того, заметим, что все три статистики при условии выполнения нулевой гипотезы имеют асимптотическое распределение $\chi^2$ с одной степенью свободы. И $p$-значения будут вычислены исходя из этого.

а) Статистика Вальда строится из соображений, что величина $\widehat{p} - p_0$ при выполнении нулевой гипотезы должна быть мала. Величина $\widehat{p}$ является асимптотически нормальной с асимптотической дисперсией, равной

$I_1^{-1}(p) = p(1-p)$. Взяв в качестве оценки $I_1^{-1}(p)$ величину $I_1^{-1}(\widehat{p}) = \widehat{p}(1-\widehat{p})$, получим статистику Вальда:

$$\mathrm{W} = (\widehat{p} - p_0)' \frac{n}{\widehat{p}(1-\widehat{p})}(\widehat{p} - p_0) = \frac{n(\widehat{p} - p_0)^2}{\widehat{p}(1-\widehat{p})} = \frac{100 \cdot (0.42 - 0.5)^2}{0.42(1-0.42)} = 2.627.$$

$P$-значение статистики W равно 0.105. Следовательно, нулевая гипотеза не отвергается на 5%-ном уровне значимости.

б) Статистика теста множителей Лагранжа строится из соображений, что производная логарифмической функции правдоподобия в точке $p_0$ при выполнении нулевой гипотезы должна быть близка к нулю. При этом в качестве оценки $I_n(p)$ берется $I_n(p_0)$. Имеем:

$$\begin{aligned}\mathrm{LM} &= \left(\frac{x - np_0}{p_0(1-p_0)}\right)' I_n^{-1}(p_0) \left(\frac{x - np_0}{p_0(1-p_0)}\right) = n^2 \frac{(\widehat{p} - p_0)^2}{p_0(1-p_0)} \frac{p_0(1-p_0)}{n} \\ &= \frac{n(\widehat{p} - p_0)^2}{p_0(1-p_0)} = \frac{100 \cdot (0.42 - 0.5)^2}{0.5(1-0.5)} = 2.56.\end{aligned}$$

$P$-значение статистики W равно 0.110. Следовательно, нулевая гипотеза не отвергается на 5%-ном уровне значимости.

в) Статистика теста отношения правдоподобия строится из соображений, что значения логарифмической функции правдоподобия в точках $\widehat{p}$ и $p_0$ при условии нулевой гипотезы должны быть близки. Имеем:

$$\begin{aligned}\mathrm{LR} &= -2(l(p_0) - l(\widehat{p})) \\ &= -2((x \ln p_0 + (n-x)\ln(1-p_0)) - (x \ln \widehat{p} + (n-x)\ln(1-\widehat{p}))) \\ &= -2n((\widehat{p} \ln p_0 + (1-\widehat{p})\ln(1-p_0)) - (\widehat{p} \ln \widehat{p} + (1-\widehat{p})\ln(1-\widehat{p}))) \\ &= 2n\left(\widehat{p} \ln \frac{\widehat{p}}{p_0} + (1-\widehat{p}) \ln \frac{1-\widehat{p}}{1-p_0}\right) \\ &= 2 \cdot 100 \cdot \left(0.42 \ln \frac{0.42}{0.5} + (1-0.42) \ln 1 - 0.421 - 0.5\right) = 2.571.\end{aligned}$$

$P$-значение статистики W равно 0.109. Следовательно, нулевая гипотеза не отвергается на 5%-ном уровне значимости.

Заметим, что в данном примере выполняется следующее неравенство между тремя статистиками: LM < LR < W. Однако, в отличие от такого же неравенства при проверке линейного ограничения в линейной модели, это неравенство оказалось справедливым случайно (посмотрите, например, какое соотношение будет между тремя статистиками, если $\widehat{p} = 0.5$, $p_0 = 0.42$).

## Задача 10.13

Имеется 80 наблюдений пуассоновской случайной величины $X$. Их среднее значение равно $\overline{x} = 1.7$. Тестируйте на 5%-ном уровне значимости гипотезу $\mathrm{H}_0: \lambda = 2.0$:

а) при помощи теста Вальда (W);

б) при помощи теста множителей Лагранжа (LM);

в) при помощи теста отношения правдоподобия (LR).

## Решение

Случайная величина $X$ подчиняется распределению Пуассона с параметром $\lambda$. Следовательно,

$$\mathrm{P}(X = k) = \frac{\lambda^k}{k!}e^{-\lambda},$$

и функция правдоподобия равна

$$L = \prod_{t=1}^{n} \frac{\lambda^{x_t}}{x_t!}e^{-\lambda},$$

а логарифмическая функция правдоподобия равна

$$l = \ln L = -\sum_{t=1}^{n}\ln(x_t!) + \sum_{t=1}^{n} x_t \ln\lambda - n\lambda = -\sum_{t=1}^{n}\ln(x_t!) + n\overline{x}\ln\lambda - n\lambda.$$

Уравнение правдоподобия:

$$\frac{\partial l}{\partial p} = \frac{n\overline{x}}{\lambda} - n = n\frac{\overline{x} - \lambda}{\lambda} = 0,$$

откуда получаем оценку максимального правдоподобия $\widehat{\lambda} = \overline{x} = 1.7$.

При построении тестов нам еще понадобится количество информации Фишера для данной задачи:

$$I_n(p) = -\mathrm{E}\left(\frac{\partial^2 l}{\partial p^2}\right) = \mathrm{E}\left(\frac{n\overline{x}}{\lambda^2}\right) = \frac{n\lambda}{\lambda^2} = \frac{n}{\lambda}$$

(здесь мы воспользовались тем, что $\mathrm{E}X = \lambda$).

Далее действуем аналогично задаче 10.12. После построения статистик $p$-значения вычисляются, исходя из того, что при выполнении нулевой гипотезы все три статистики имеют асимптотическое распределение $\chi^2$ с одной степенью свободы.

а) Статистика Вальда равна

$$\mathrm{W} = (\widehat{\lambda} - \lambda_0)' I_n(\widehat{\lambda})(\widehat{\lambda} - \lambda_0) = \frac{n(\widehat{\lambda} - \lambda_0)^2}{\widehat{\lambda}} = \frac{80 \cdot (1.7 - 2.0)^2}{1.7} = 4.235.$$

$P$-значение статистики W равно 0.040. Следовательно, нулевая гипотеза отвергается на 5%-ном уровне значимости.

б) Статистика теста множителей Лагранжа равна

$$\mathrm{LM} = \left(n\frac{\overline{x}-\lambda_0}{\lambda_0}\right)' I_n^{-1}(p_0)\left(n\frac{\overline{x}-\lambda_0}{\lambda_0}\right) = n^2\frac{(\widehat{\lambda}-\lambda_0)^2}{\lambda_0^2}\frac{\lambda_0}{n}$$

$$= \frac{n(\widehat{\lambda}-\lambda_0)^2}{\lambda_0} = \frac{80\cdot(1.7-2.0)^2}{2.0} = 3.6.$$

$P$-значение статистики W равно 0.058. Следовательно, нулевая гипотеза не отвергается на 5%-ном уровне значимости.

в) Статистика теста отношения правдоподобия равна

$$\begin{aligned}\mathrm{LR} &= -2(l(\lambda_0) - l(\widehat{\lambda}))\\ &= -2((n\overline{x}\ln\lambda_0 - n\lambda_0) - (n\overline{x}\ln\lambda_0 - n\lambda_0))\\ &= 2n\left(\widehat{\lambda}\ln\frac{\widehat{\lambda}}{\lambda_0} + (\lambda_0 - \widehat{\lambda})\right)\\ &= 2\cdot 80\cdot\left(1.7\ln\frac{1.7}{2.0} + (2.0-1.7)\right) = 3.795.\end{aligned}$$

$P$-значение статистики W равно 0.051. Следовательно, нулевая гипотеза не отвергается на 5%-ном уровне значимости.

В этом случае получилось, что разные тесты приводят к разным выводам. Кроме того, здесь, как и в задаче 10.12, тоже выполняется неравенство $\mathrm{LM} < \mathrm{LR} < \mathrm{W}$, и тоже случайно (рассмотрите случай, когда $\widehat{\lambda} = 2.0$, $\lambda_0 = 1.7$).

## Задача 10.14

Пусть имеется следующая модель:

$$y_t = \begin{cases} 0, & y_t^* < 0, \\ y_t^* & y_t^* \geqslant 0, \end{cases} \quad \text{где} \quad y_t^* = \varepsilon_t - \gamma, \quad \varepsilon_t \sim iidU(0,\theta).$$

(Ошибки $\varepsilon_t$ независимы и имеют равномерное распределение.) Имеется $n$ наблюдений величины $y_t$, величины $y_t^*$ и $\varepsilon_t$ не наблюдаемы.

Найдите оценки максимального правдоподобия параметров $\gamma$ и $\theta$. (Известно, что $\theta > \gamma > 0$.)

## Решение

Величина $y_t^*$ равномерно распределена на отрезке $[-\gamma, \theta-\gamma]$, ее плотность распределения в этом интервале равна $\theta^{-1}$. Вероятность того, что $y_t = 0$ равна $\mathrm{P}(y_t = 0) = \gamma/\theta$. Отсюда получаем функцию правдоподобия:

$$L(\theta,\gamma) = \prod_{y_t=0}\frac{\gamma}{\theta}\prod_{y_t\geqslant 0}\frac{1}{\theta}I_{[0,\theta-\gamma]}(y_t) = \prod_{y_t=0}\frac{\gamma}{\theta}\prod_{0<y_t<\theta-\gamma}\frac{1}{\theta}\prod_{y_t>\theta-\gamma} 0.$$

Обозначим через $n_0$ количество наблюдений с $y_t = 0$, а через $n_1$ количество наблюдений с $y_t > 0$; $n = n_0 + n_1$. Максимальное значение $y_t$ обозначим через $M = \max_t\{y_t\}$. Задача максимизации функции правдоподобия выглядит следующим образом:

$$\left(\frac{\gamma}{\theta}\right)^{n_0} \frac{1}{\theta^{n_1}} \to \max \quad \text{при условии} \quad y_t < \theta - \gamma, \quad t = 1, \ldots, n,$$

или

$$\frac{\gamma^{n_0}}{\theta^n} \to \max \quad \text{при условии} \quad \theta - \gamma > M.$$

Поскольку $\theta - \gamma \geqslant M$, то для некоторого положительного $\delta$ имеем $\theta = M + \gamma + \delta$, тогда, в терминах $\gamma$ и $\delta$, задача выглядит следующим образом:

$$\frac{\gamma^{n_0}}{(M + \gamma + \delta)^n} \to \max \quad \text{при условии} \quad \gamma, \delta \geqslant 0.$$

Из полученного выражения видно, что при фиксированном $\gamma$ максимум достигается при $\delta = 0$. Остается найти $\gamma$ из условия максимума

$$\frac{\gamma^{n_0}}{(M + \gamma)^n} \quad \text{или} \quad \ln \frac{\gamma^{n_0}}{(M + \gamma)^n} = n_0 \ln \gamma - n \ln(M + \gamma).$$

Приравнивая производную по $\gamma$ нулю, получаем:

$$\frac{n_0}{\gamma} - \frac{n}{M + \gamma} = 0 \quad \text{или} \quad \widehat{\gamma} = \frac{n_0 M}{n - n_0} = \frac{n_0}{n - n_0} \max_t\{y_t\}.$$

Итак, получаем:

$$\widehat{\gamma}_{\mathrm{ML}} = \frac{n_0}{n - n_0} \max_t\{y_t\},$$
$$\widehat{\theta}_{\mathrm{ML}} = \frac{n}{n - n_0} \max_t\{y_t\}.$$

# Глава 11

# Временные ряды

## Задача 11.1

Покажите, что выражение для суммарного влияния в модели (11.2)

$$y_t = \beta_1 + \beta_2 x_t + \beta_3 y_{t-1} + \varepsilon_t$$

может быть получено как изменение (отклик) эндогенной переменной при единичном приращении экзогенной переменной в стационарном состоянии (т. е. когда $y_t = y_{t+1} = \cdots = \overline{y}$ и $x_t = x_{t+1} = \cdots = \overline{x}$).

## Решение

Предположим, что $x_t$ зафиксировано постоянным на некотором уровне $\overline{x}$. Отсюда следует, что если выполнено условие устойчивости $|\beta_3| < 1$, то $y_t$ сходится к постоянной $\overline{y}$, определяемой условием:

$$\overline{y} = \beta_1 + \beta_2 \overline{x} + \beta_3 \overline{y}. \qquad (*)$$

Точнее, рассмотрим последовательность $\mathrm{E}y_t$, которая в силу $\mathrm{E}\varepsilon_t = 0$ удовлетворяет уравнению

$$\mathrm{E}y_t = \beta_1 + \beta_2 \overline{x} + \beta_3 \mathrm{E}y_{t-1}.$$

Вычитая два последних уравнения одно из другого, получаем:

$$\mathrm{E}y_t - \overline{y} = \beta_3(\mathrm{E}y_{t-1} - \overline{y}).$$

Из условия устойчивости $|\beta_3| < 1$ следует, что $\mathrm{E}y_t - \overline{y} \to 0$, или $\mathrm{E}y_t \to \overline{y}$. Из условия $(*)$ следует, что

$$\overline{y} = \frac{\beta_1}{1-\beta_3} + \frac{\beta_2}{1-\beta_3}\overline{x}.$$

Отсюда получаем, что увеличение $\overline{x}$ на 1 приводит к увеличению $\overline{y}$ на величину $\beta_2/(1-\beta_3)$, равную суммарному влиянию.

## Задача 11.2

Покажите, что суммарное влияние $x$ на $y$ в модели (11.3)

$$A(L)y_t = \delta + B(L)x_t + \varepsilon_t; \qquad t = 1, \ldots, n,$$
$$A(L) = 1 - \alpha_1 L - \cdots - \alpha_p L^p,$$
$$B(L) = \beta_0 + \beta_1 L + \cdots + \beta_q L^q,$$

равно $B(1)/A(1)$.

## Решение

Рассмотрим стационарное состояние (см. решение задачи 11.1): $y_t \equiv \overline{y}$ и $x_t \equiv \overline{x}$. Очевидно, в стационарном состоянии $L\overline{y} \equiv \overline{y}$, или $A(L)\,\overline{y} = A(1)\,\overline{y}$. Таким образом,

$$A(1)\,\overline{y} = \delta + B(1)\,\overline{x}, \quad \text{или} \quad \overline{y} = \frac{\delta}{A(1)} + \frac{B(1)}{A(1)}\,\overline{x}.$$

Отсюда получаем, что увеличение $\overline{x}$ на 1 приводит к увеличению $\overline{y}$ на величину $B(1)/A(1)$, равную суммарному влиянию.

## Задача 11.3

Покажите, что уравнение (11.3)

$$A(L)y_t = \delta + B(L)x_t + \varepsilon_t; \qquad t = 1, \ldots, n,$$
$$A(L) = 1 - \alpha_1 L - \cdots - \alpha_p L^p,$$
$$B(L) = \beta_0 + \beta_1 L + \cdots + \beta_q L^q,$$

устойчиво, если выполнено условие: все корни многочлена $A(z) = 1 - \alpha_1 z - \cdots - \alpha_p z^p$ лежат вне единичной окружности.

## Решение

Заметим сначала, что в разложении многочлена $A(L)$ на множители

$$A(L) = (1 - \lambda_1 L)(1 - \lambda_2 L) \cdots (1 - \lambda_p L)$$

числа $1/\lambda_i$ являются корнями многочлена $A(z)$. Поэтому по условию задачи все $\lambda_i$ лежат *внутри* единичного круга: $|\lambda_i| < 1$, $i = 1, \ldots, p$.

Теперь, как и в решении задачи 11.1, предположим, что $x_t$ зафиксировано постоянным на некотором уровне $\overline{x}$. Если уравнение устойчиво, $y_t$ сходится к постоянной $\overline{y}$, определяемой условием:

$$A(L)\overline{y} = \delta + B(L)\overline{x}, \quad \text{или} \quad A(1)\overline{y} = \delta + B(1)\overline{x}.$$

Точнее, рассмотрим последовательность $\mathrm{E}y_t$, которая в силу $\mathrm{E}\varepsilon_t = 0$ удовлетворяет уравнению

$$A(L)\mathrm{E}y_t = \delta + B(L)\overline{x}.$$

Вычитая два последних уравнения одно из другого, получаем:

$$A(L)(\mathrm{E}y_t - \overline{y}) = 0.$$

Обозначим через $w_t = \mathrm{E}y_t - \overline{y}$, тогда последовательность $w_t$ удовлетворяет разностному уравнению $A(L)w_t = 0$. Как известно, общее решение этого уравнения может быть записано в виде

$$w_t = c_1\lambda_1^t + c_2\lambda_2^t + \ldots + c_p\lambda_p^t,$$

где $c_i$ — некоторые постоянные.

Отсюда следует, что $w_t$ сходится к 0 при $t \to \infty$ (или, что то же самое, $\mathrm{E}y_t$ сходится к $\overline{y}$) тогда и только тогда, когда все $\lambda_i$ лежат внутри единичного круга.

## Задача 11.4

Выведите формулы, выражающие переменные $\widetilde{x}_{0t}, \ldots, \widetilde{x}_{rt}$ в уравнении (11.7)

$$y_t = \delta + \gamma_0\widetilde{x}_{0t} + \cdots + \gamma_r\widetilde{x}_{rt} + \varepsilon_t, \qquad t = 1, \ldots, n,$$

через переменные $x_t, x_{t-1}, \ldots, x_{t-q}$ из уравнения (11.5)

$$y_t = \delta + \beta_0 x_t + \beta_1 x_{t-1} + \cdots + \beta_q x_{t-q} + \varepsilon_t, \qquad t = 1, \ldots, n,$$

при аппроксимации $\beta_i$ полиномом (11.6) степени $r \leqslant q$:

$$\beta_i = \gamma_0 + \gamma_1 i + \cdots + \gamma_r i^r, \qquad r \leqslant q.$$

## Решение

Преобразуем уравнение (11.5), подставив в него выражения для $\beta_i$:

$$\begin{aligned}
y_t &= \delta + \beta_0 x_t + \beta_1 x_{t-1} + \cdots + \beta_q x_{t-q} + \varepsilon_t \\
&= \delta + (\gamma_0)x_t \\
&\qquad + (\gamma_0 + \gamma_1 + \gamma_2 + \cdots + \gamma_r)x_{t-1} \\
&\qquad + (\gamma_0 + 2\gamma_1 + 2^2\gamma_2 + \cdots + 2^r\gamma_r)x_{t-2} \\
&\qquad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\
&\qquad + (\gamma_0 + q\gamma_1 + q^2\gamma_2 + \cdots + q^r\gamma_r)x_{t-q} + \varepsilon_t \\
&= \delta + \gamma_0(x_t + x_{t-1} + x_{t-2} + \cdots + x_{t-q}) \\
&\qquad + \gamma_1(x_{t-1} + 2x_{t-2} + 3x_{t-3} + \cdots + qx_{t-q}) \\
&\qquad + \gamma_2(x_{t-1} + 2^2 x_{t-2} + 3^2 x_{t-3} + \cdots + q^2 x_{t-q}) \\
&\qquad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\
&\qquad + \gamma_r(x_{t-1} + 2^r x_{t-2} + 3^r x_{t-3} + \cdots + q^r x_{t-q}) + \varepsilon_t \\
&= \delta + \gamma_0\widetilde{x}_{0t} + \gamma_1\widetilde{x}_{1t} + \gamma_2\widetilde{x}_{2t} + \cdots + \gamma_r\widetilde{x}_{rt} + \varepsilon_t.
\end{aligned}$$

Приравнивая коэффициенты при $\gamma_i$, получаем:

$$\widetilde{x}_{0t} = x_t + x_{t-1} + x_{t-2} + \cdots + x_{t-q},$$
$$\widetilde{x}_{kt} = x_{t-1} + 2^k x_{t-2} + 3^k x_{t-3} + \cdots + q^k x_{t-q}, \quad k = 1, 2, \ldots, r.$$

## Задача 11.5

Выведите формулу для дисперсии (11.12) $\mathrm{V}(y_t) = \sigma^2/(1-\beta^2)$ для AR(1) процесса

$$y_t = \beta y_{t-1} + \varepsilon_t, \quad \varepsilon_t \sim iid(0, \sigma^2), \quad t = 1, \ldots, n, \quad |\beta| < 1.$$

## Решение

Из уравнения AR(1) процесса ясно, что $y_{t-1}$ содержит информацию только от значений «возмущений» $\varepsilon_s$ в предыдущие моменты времени $s < t$. Поэтому $y_{t-1}$ и $\varepsilon_t$ некоррелированы. Обозначив $\sigma_t^2 = \mathrm{V}(y_t)$ и взяв дисперсию от обеих частей уравнения, получаем:

$$\sigma_t^2 = \beta^2 \sigma_{t-1}^2 + \sigma^2, \qquad t = 1, 2, \ldots$$

Выпишем несколько первых равенств:

$$\sigma_1^2 = \beta^2 \sigma_0^2 + \sigma^2,$$
$$\sigma_2^2 = \beta^2 \sigma_1^2 + \sigma^2,$$
$$\sigma_3^2 = \beta^2 \sigma_2^2 + \sigma^2, \ldots$$

После нескольких подстановок получим:

$$\sigma_t^2 = \beta^{2t}\sigma_0^2 + \sigma^2(\beta^{2t-2} + \beta^{2t-4} + \ldots + \beta^2 + 1)$$
$$= \beta^{2t}\sigma_0^2 + \frac{1-\beta^{2t}}{1-\beta^2}\sigma^2 = \frac{\sigma^2}{1-\beta^2} + \beta^{2t}\left(\sigma_0^2 - \frac{\sigma^2}{1-\beta^2}\right).$$

Отсюда видно, что $\sigma_t^2 \to \sigma^2/(1-\beta^2)$ при возрастании $t$, так как $|\beta| < 1$. Это означает, что влияние начального наблюдения $y_0$ со временем убывает и процесс сходится к стационарному. Если положить дисперсию начального наблюдения $\sigma_0^2$ равной $\sigma^2/(1-\beta^2)$, то получаем стационарный процесс с дисперсией $\sigma_t^2 \equiv \sigma^2/(1-\beta^2)$.

## Задача 11.6

Покажите, что для случайных величин, удовлетворяющих уравнению (11.11):

$$y_t = \beta y_{t-1} + \varepsilon_t, \quad \varepsilon_t \sim iid(0, \sigma^2), \quad t = 1, \ldots, n,$$

при условии существования момента $\mathrm{E}(\varepsilon_t^4)$ выполнено:

а) $p\lim(1/n)\sum y_{t-1}\varepsilon_t = 0$.

б) $p\lim(1/n)\sum y_{t-1}^2 = \sigma^2/(1-\beta^2)$.

## Решение

Приведем три решения задачи.

*Первое решение.* Сформулируем две леммы, которые будут использованы при решении задачи.

*Лемма* 1. Пусть $x_n$, $n = 1, 2, \ldots$, есть последовательность случайных величин, такая, что $\mathrm{E}(x_n) \to \mu$ и $\mathrm{V}(x_n) \to 0$ при $n \to \infty$, тогда существует предел $p\lim x_n = \mu$.

*Доказательство.* См. приложение МС, п. 5.

*Лемма* 2. Пусть $\{\varepsilon_t\} \sim \mathrm{iid}(0, \sigma^2)$ — независимые одинаково распределенные случайные величины со средним 0 и (конечной) дисперсией $\sigma^2$ и пусть существует $\mathrm{E}(\varepsilon_t^4)$. Обозначим через $\kappa = \mathrm{E}(\varepsilon_t/\sigma)^4 - 3$ коэффициент эксцесса распределения. Пусть $\boldsymbol{\varepsilon} = (\varepsilon_1, \varepsilon_2, \ldots, \varepsilon_n)'$ — случайный $n \times 1$ вектор, а $\boldsymbol{A}$ — симметричная $n \times n$ матрица. Тогда

$$\mathrm{E}(\boldsymbol{\varepsilon}'\boldsymbol{A}\boldsymbol{\varepsilon}) = \sigma^2 \operatorname{tr}(\boldsymbol{A}), \qquad \mathrm{V}(\boldsymbol{\varepsilon}'\boldsymbol{A}\boldsymbol{\varepsilon}) = \sigma^4\left(2\operatorname{tr}(\boldsymbol{A}^2) + \kappa\sum_{t=1}^{n} a_{tt}^2\right).$$

*Доказательство.* Вычислим математические ожидания:

$$\mathrm{E}(\boldsymbol{\varepsilon}'\boldsymbol{A}\boldsymbol{\varepsilon}) = \sum_{st} a_{st}\mathrm{E}(\varepsilon_s\varepsilon_t) = \sum_t a_{tt}\mathrm{E}(\varepsilon_t^2) = \sigma^2 \operatorname{tr}\boldsymbol{A},$$

и

$$\begin{aligned}
\mathrm{E}\left((\boldsymbol{\varepsilon}'\boldsymbol{A}\boldsymbol{\varepsilon})^2\right) &= \sum_{ij}\sum_{st} a_{ij}a_{st}\mathrm{E}(\varepsilon_i\varepsilon_j\varepsilon_s\varepsilon_t) \\
&= \sigma^4\left(\sum_{i\neq s} a_{ii}a_{ss} + 2\sum_{i\neq j} a_{ij}^2\right) + \mathrm{E}(\varepsilon_t^4)\sum_t a_{tt}^2 \\
&= \sigma^4\left((\operatorname{tr}\boldsymbol{A})^2 + 2\operatorname{tr}(\boldsymbol{A}^2) + \kappa\sum_t a_{tt}^2\right).
\end{aligned}$$

(Заметим, что в случае нормального распределения $\varepsilon_t$ коэффициент эксцесса равен 0, и, следовательно, $\mathrm{E}(\boldsymbol{\varepsilon}'\boldsymbol{A}\boldsymbol{\varepsilon}) = \sigma^2\operatorname{tr}\boldsymbol{A}$ и $\mathrm{V}(\boldsymbol{\varepsilon}'\boldsymbol{A}\boldsymbol{\varepsilon}) = 2\sigma^4\operatorname{tr}(\boldsymbol{A}^2)$. В частном случае, если $\sigma^2 = 1$ и $\boldsymbol{A} = I_n$, мы получаем, что $\boldsymbol{\varepsilon}'\boldsymbol{A}\boldsymbol{\varepsilon}$ имеет среднее $n$ и дисперсию $2n$, что согласуется с тем, что в этом случае $\boldsymbol{\varepsilon}'\boldsymbol{A}\boldsymbol{\varepsilon}$ распределено по закону $\chi^2(n)$.)

Рассмотрим теперь уравнение $y_t = \beta y_{t-1} + \varepsilon_t$. После последовательных итераций получим:

$$y_t = \beta^t y_0 + \beta^{t-1}\varepsilon_1 + \beta^{t-2}\varepsilon_2 + \cdots + \varepsilon_t.$$

Для простоты вычислений предположим, что $y_0 = 0$. (На самом деле доказательство лишь незначительно усложняется, если предположить, что $y_0$ является константой или случайной величиной, независимой от $\varepsilon_1, \dots, \varepsilon_n$.)
Перепишем последнее уравнение в векторном виде:

$$y_t = \boldsymbol{a}_t' \boldsymbol{\varepsilon}, \qquad \boldsymbol{a}_t = (\beta^{t-1}, \beta^{t-2}, \dots, \beta, 1, 0, \dots, 0)'.$$

Обозначим через $\boldsymbol{e}_t$ $(n \times 1)$ вектор, у которого координата $t$ равна 1, а все остальные равны 0. Введем также следующие матрицы:

$$\boldsymbol{A} = \frac{1}{2} \sum_t (\boldsymbol{a}_{t-1} \boldsymbol{e}_t' + \boldsymbol{e}_t \boldsymbol{a}_{t-1}'), \qquad \boldsymbol{B} = \sum_t \boldsymbol{a}_{t-1} \boldsymbol{a}_{t-1}'.$$

Выражения из условия задачи можно записать через матрицы следующим способом:

$$\sum_t y_{t-1} \varepsilon_t = \boldsymbol{\varepsilon}' \boldsymbol{A} \boldsymbol{\varepsilon}, \qquad \sum_t y_{t-1}^2 = \boldsymbol{\varepsilon}' \boldsymbol{B} \boldsymbol{\varepsilon}.$$

Теперь вычислим выражения, которые нам понадобятся в дальнейшем:

$$\operatorname{tr} \boldsymbol{A} = \sum_t \boldsymbol{e}_t' \boldsymbol{a}_{t-1} = 0,$$

$$\operatorname{tr}(\boldsymbol{A}^2) = \frac{1}{2} \sum_{st} (\boldsymbol{e}_t' \boldsymbol{a}_{s-1})(\boldsymbol{e}_s' \boldsymbol{a}_{t-1}) + \frac{1}{2} \sum_t \boldsymbol{a}_{t-1}' \boldsymbol{a}_{t-1}$$
$$= \frac{1}{2} \sum_t \boldsymbol{a}_{t-1}' \boldsymbol{a}_{t-1} \approx \frac{n}{2(1-\beta^2)},$$

$$\sum a_{tt}^2 = 0,$$

а также

$$\operatorname{tr} \boldsymbol{B} = \sum_t \boldsymbol{a}_{t-1}' \boldsymbol{a}_{t-1} \approx \frac{n}{1-\beta^2},$$

$$\operatorname{tr}(\boldsymbol{B}^2) = \sum_{st} (\boldsymbol{a}_{s-1}' \boldsymbol{a}_{t-1})^2 \approx \frac{n(1+\beta^2)}{(1-\beta^2)^3},$$

$$\sum b_{tt}^2 \approx \frac{n}{1-\beta^2}.$$

Используя лемму 2 и предыдущие формулы, получаем:

$$\mathrm{E}\left(\frac{1}{n} \sum_t y_{t-1} \varepsilon_t\right) = \frac{1}{n} \mathrm{E}(\boldsymbol{\varepsilon}' \boldsymbol{A} \boldsymbol{\varepsilon}) = \frac{1}{n} \operatorname{tr} \boldsymbol{A} = 0,$$

$$\mathrm{V}\left(\frac{1}{n} \sum_t y_{t-1} \varepsilon_t\right) = \frac{1}{n^2} \mathrm{V}(\boldsymbol{\varepsilon}' \boldsymbol{A} \boldsymbol{\varepsilon}) = \sigma^4 \frac{1}{n^2} \left(2 \operatorname{tr}(\boldsymbol{A}^2) + \kappa \sum_{t=1}^n a_{tt}^2\right) = O\left(\frac{1}{n}\right) \to 0,$$

а также

$$\mathrm{E}\left(\frac{1}{n}\sum_t y_{t-1}^2\right) = \frac{1}{n}\mathrm{E}(\varepsilon' B\varepsilon) = \frac{1}{n}\sigma^2\frac{n}{1-\beta^2} \to \frac{\sigma^2}{1-\beta^2},$$

$$\mathrm{V}\left(\frac{1}{n}\sum_t y_{t-1}^2\right) = \frac{1}{n^2}\mathrm{V}(\varepsilon' B\varepsilon) = \frac{1}{n^2}\sigma^4\left(2\,\mathrm{tr}(B^2) + \kappa\sum_{t=1}^n b_{tt}^2\right) = O(n^{-1}) \to 0.$$

Применяя лемму 1, получаем утверждение задачи.

*Второе решение.* В этом решении будем предполагать стационарность ряда наблюдений $y_t$. Будем использовать следующий факт, доказательство которого можно найти в книге Брокуэлла и Дэвиса (Proposition 7.3.1)[1].

*Теорема.* Пусть $\{\varepsilon_t\} \sim \mathrm{iid}(0, \sigma^2)$, а $\kappa = \mathrm{E}(\varepsilon_t/\sigma)^4 - 3$ — коэффициент эксцесса. Пусть $\{y_t\}$ имеет представление в виде скользящего среднего

$$y_t = \sum_{j=0}^{\infty} \psi_j \varepsilon_{t-j},$$

такое, что $\sum_{j=0}^{\infty} |\psi_j| < \infty$. Тогда для любого $h \geqslant 0$, существует предел

$$\lim_{n\to\infty} n\mathrm{V}\left(n^{-1}\sum_{t=1}^n y_t y_{t-h}\right) = \kappa\gamma_h^2 + \sum_{j=0}^{\infty}\left(\gamma_j^2 + \gamma_{j+h}\gamma_{j-h}\right),$$

где $\gamma_h = \mathrm{Cov}(y_t, y_{t-h})$.

Из этой теоремы следует, что при весьма общих предположениях сумма $n^{-1}\sum_{t=1}^n y_t y_{t-h}$ сходится по вероятности к ее математическому ожиданию $\gamma_h$.

В частном случае AR(1) модели имеем

$$y_t = \sum_{j=0}^{\infty} \beta^j \varepsilon_{t-j},$$

и ковариации равны

$$\gamma_h = \frac{\sigma^2\beta^h}{1-\beta^2}, \qquad h = 0, 1, \ldots$$

Поскольку в данном случае условия теоремы, очевидно, выполнены, получаем:

$$\operatorname{plim} n^{-1}\sum y_t y_{t-1} = \gamma_1 = \frac{\sigma^2\beta}{1-\beta^2},$$

$$\operatorname{plim} n^{-1}\sum y_{t-1}^2 = \gamma_0 = \frac{\sigma^2}{1-\beta^2}.$$

[1] P. J. Brockwell and R. A. Davis (1987). *Time Series: Theory and Methods*, New York: Springer-Verlag.

Учитывая, что $y_t y_{t-1} = \beta y_{t-1}^2 + y_{t-1}\varepsilon_t$, получаем:

$$p\lim n^{-1}\sum y_{t-1}\varepsilon_t = \gamma_1 - \beta\gamma_0 = 0.$$

*Третье решение.* Воспользуемся следующим общим результатом из теории случайных процессов[2].

*Теорема.* Пусть $\{z_t\}$ — стационарный процесс и

$$\lambda(k) = \mathrm{Cov}(z_t, z_{t-k})/\mathrm{V}(z_t)$$

— его корреляционная функция. Тогда, если $\lambda(k) \to 0$ при $k \to \infty$, то процесс $\{z_t\}$ является эргодическим:

$$p\lim_{n\to\infty}\frac{1}{n}\sum_{t=1}^{n} z_t = \mathrm{E}(z_t).$$

а) Нетрудно проверить, что процесс $z_t = y_{t-1}\varepsilon_t$ является стационарным, поскольку стационарным является двумерный процесс $(y_{t-1}, \varepsilon_t)'$. При этом

$$\mathrm{E}(z_t) = \mathrm{E}(y_{t-1}\varepsilon_t) = \mathrm{E}(y_{t-1})\mathrm{E}(\varepsilon_t) = 0$$

в силу независимости $y_{t-1}$ и $\varepsilon_t$ и

$$\mathrm{E}(z_t z_{t-k}) = \mathrm{E}(y_{t-1}\varepsilon_t y_{t-k-1}\varepsilon_{t-k}) = \mathrm{E}(y_{t-1}y_{t-k-1}\varepsilon_{t-k})\mathrm{E}(\varepsilon_t) = 0,$$

так как случайные величины $y_{t-1}y_{t-k-1}\varepsilon_{t-k}$ и $\varepsilon_t$ независимы. Таким образом, выполнены условия теоремы и, следовательно,

$$p\lim\frac{1}{n}\sum_{t=1}^{n} y_{t-1}\varepsilon_t = \mathrm{E}(z_t) = 0.$$

б) Пусть $\delta^2 = \sigma^2/(1-\beta^2)$ — дисперсия случайной величины $y_{t-1}$, введем также обозначение $z_t = y_{t-1}^2$. Очевидно, что процесс $z_t$ является стационарным и $\mathrm{E}(z_t) = \delta^2$. Найдем ковариацию

$$\mathrm{Cov}(z_t, z_{t-k}) = \mathrm{E}(y_{t-1}^2 y_{t-k-1}^2) - \mathrm{E}(y_{t-1}^2)\mathrm{E}(y_{t-k-1}^2) = \mathrm{E}(y_{t-1}^2 y_{t-k-1}^2) - \delta^4.$$

Из представления (11.11) следует, что $y_{t-1} = \beta^k y_{t-k-1} + u_{t,k}$, где $u_{t,k}$ — случайная величина, независимая с $y_{t-k-1}$. Очевидно, что $\mathrm{E}(u_{t,k}) = 0$ и $\mathrm{V}(y_{t-1}) = \beta^{2k}\mathrm{V}(y_{t-k-1}) + \mathrm{V}(u_{t,k})$, откуда $\mathrm{V}(u_{t,k}) = \delta^2(1-\beta^{2k})$. Таким образом,

$$\begin{aligned}
\mathrm{E}(y_{t-1}^2 y_{t-k-1}^2) &= \mathrm{E}((\beta^k y_{t-k-1} + u_{t,k})^2 y_{t-k-1}^2) \\
&= \mathrm{E}(\beta^{2k} y_{t-k-1}^4 + 2\beta^k y_{t-k-1}^3 u_{t,k} + y_{t-k-1}^2 u_{t,k}^2) \\
&= \beta^{2k}\mathrm{E}(y_{t-k-1}^4) + \mathrm{E}(y_{t-k-1}^2)\mathrm{E}(u_{t,k}^2) \\
&= \beta^{2k}\mathrm{E}(y_{t-k-1}^4) + \delta^4(1-\beta^{2k}).
\end{aligned}$$

---

[2] См., например, Справочник по теории вероятностей и математической статистике. М.: Наука, 1985.

Отсюда

$$\operatorname{Cov}(z_t, z_{t-k}) = \beta^{2k}(\operatorname{E}(y_{t-k-1}^4) - \delta^4).$$

Наконец, из представления (11.12) и существования $\operatorname{E}(\varepsilon_t^4)$ следует, что $\operatorname{E}(y_{t-k-1}^4) = \text{const} < \infty$, и, значит, $\operatorname{Cov}(z_t, z_{t-k}) \to 0$ при $k \to \infty$, так как $|\beta| < 1$. Применяя теорему, получаем, что

$$p\lim \frac{1}{n}\sum_{t=1}^{n} y_{t-1}^2 = p\lim \frac{1}{n}\sum_{t=1}^{n} z_t = \operatorname{E}(z_t) = \operatorname{E}(y_{t-1}^2) = \operatorname{V}(y_{t-1}) = \delta^2 = \frac{\sigma^2}{1-\beta^2},$$

что и требовалось доказать.

## Задача 11.7

Докажите формулы (11.16)

$$p\lim \widehat{\beta} = \frac{\beta + \rho}{1 + \beta\rho} \neq \beta$$

и (11.17)

$$p\lim \widehat{\rho} = \frac{\beta\rho(\beta + \rho)}{1 + \beta\rho} \neq \rho$$

для модели (11.15)

$$y_t = \beta y_{t-1} + u_t, \qquad t = 1, \ldots, n,$$
$$u_t = \rho u_{t-1} + \varepsilon_t, \qquad \varepsilon_t \sim iid(0, \sigma^2),$$

при выполнении условий устойчивости $|\beta| < 1$, $|\rho| < 1$.

## Решение

Приведем два решения задачи.

*Первое решение.* Предположим, что $y_0 = u_0 = 0$, и запишем уравнение $y_t = \beta y_{t-1} + u_t$ в виде

$$y_t = \beta^{t-1}u_1 + \beta^{t-2}u_2 + \cdots + u_t = \boldsymbol{a}_t'\boldsymbol{u},$$

где

$$\boldsymbol{a}_t' = (\beta^{t-1}, \beta^{t-2}, \ldots, \beta, 1, 0, \ldots, 0).$$

Как и в решении задачи 11.6, через $\boldsymbol{e}_t$ обозначим $n \times 1$ вектор, у которого координата $t$ равна 1, а все остальные координаты равны 0. Введем также следующие матрицы:

$$\boldsymbol{A} = \frac{1}{2}\sum_t (\boldsymbol{a}_{t-1}\boldsymbol{e}_t' + \boldsymbol{e}_t\boldsymbol{a}_{t-1}'), \qquad \boldsymbol{B} = \sum_t \boldsymbol{a}_{t-1}\boldsymbol{a}_{t-1}'.$$

В этих обозначениях

$$\sum_t y_{t-1}u_t = \boldsymbol{u}'\boldsymbol{A}\boldsymbol{u}, \qquad \sum_t y_{t-1}^2 = \boldsymbol{u}'\boldsymbol{B}\boldsymbol{u}.$$

Через $L$ обозначим матрицу

$$\boldsymbol{L} = \begin{bmatrix} 1 & 0 & 0 & \dots & 0 & 0 \\ -\rho & 1 & 0 & \dots & 0 & 0 \\ 0 & -\rho & 1 & \dots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \dots & 1 & 0 \\ 0 & 0 & 0 & \dots & -\rho & 1 \end{bmatrix}.$$

Тогда $\boldsymbol{L}\boldsymbol{u} = \boldsymbol{\varepsilon}$, а интересующие нас выражения можно представить как

$$\sum_t y_{t-1}u_t = \boldsymbol{\varepsilon}'\boldsymbol{A}^*\boldsymbol{\varepsilon}, \qquad \sum_t y_{t-1}^2 = \boldsymbol{\varepsilon}'\boldsymbol{B}^*\boldsymbol{\varepsilon},$$

где

$$\boldsymbol{A}^* = \boldsymbol{L}'^{-1}\boldsymbol{A}\boldsymbol{L}^{-1}, \qquad \boldsymbol{B}^* = \boldsymbol{L}'^{-1}\boldsymbol{B}\boldsymbol{L}^{-1}.$$

Теми же методами, какими мы пользовались в задаче 11.6, получаем следующие результаты, обобщающие результаты задачи 11.6:

$$n^{-1}\operatorname{tr}\boldsymbol{A}^* \to \frac{\rho}{(1-\rho^2)(1-\beta\rho)}$$

и

$$n^{-1}\operatorname{tr}\boldsymbol{B}^* \to \frac{1+\beta\rho}{(1-\beta\rho)(1-\beta^2)(1-\rho^2)}.$$

Как и в задаче 11.6, можно показать, что дисперсии сумм $n^{-1}\sum y_{t-1}u_t$ и $n^{-1}\sum y_{t-1}^2$ имеют порядок $n^{-1}$. Отсюда получаем:

$$\begin{aligned} p\lim\widehat{\beta} &= \beta + \frac{p\lim n^{-1}\sum y_{t-1}u_t}{p\lim n^{-1}\sum y_{t-1}^2} = \beta + \frac{p\lim n^{-1}\boldsymbol{\varepsilon}'\boldsymbol{A}^*\boldsymbol{\varepsilon}}{p\lim n^{-1}\boldsymbol{\varepsilon}'\boldsymbol{B}^*\boldsymbol{\varepsilon}} \\ &= \beta + \frac{\rho}{(1-\rho^2)(1-\beta\rho)} \Bigg/ \frac{1+\beta\rho}{(1-\beta\rho)(1-\beta^2)(1-\rho^2)} \\ &= \beta + \frac{\rho(1-\beta^2)}{1+\beta\rho} = \frac{\beta+\rho}{1+\beta\rho}. \end{aligned}$$

Таким образом, мы доказали (11.16). Аналогично можно доказать и второе равенство, однако вычисления становятся весьма громоздкими.

*Второе решение.* В этом решении мы используем стационарность процесса. Заметим, что последовательность $\{y_t\}$ является AR(2) процессом:

$$y_t = \phi_1 y_{t-1} + \phi_2 y_{t-2} + \varepsilon_t, \qquad \phi_1 = \beta + \rho, \quad \phi_2 = -\beta\rho, \qquad \text{или}$$
$$(1-\rho L)(1-\beta L)y_t = \varepsilon_t.$$

Выпишем MA($\infty$) представление AR(1)процесса $y_t$:

$$y_t = \sum_{j=0}^{\infty}\psi_j\varepsilon_{t-j}.$$

Из свойств AR(2) процесса следует, что

$$\psi_0 = 1, \qquad \psi_1 = \phi_1, \qquad \psi_j = \phi_1\psi_{j-1} + \phi_2\psi_{j-2} \ \ (j \geqslant 2).$$

Выпишем характеристическое уравнение полученного разностного уравнения

$$z^2 - \phi_1 z - \phi_2 = 0.$$

Его корнями, очевидно, являются

$$z_1 = \beta, \quad \text{и} \quad z_2 = \rho.$$

Отсюда получаем: $\psi_0 = 1$, $\psi_1 = \beta + \rho$, а для $j \geqslant 2$,

$$\psi_j = \begin{cases} (\beta^{j+1} - \rho^{j+1})/(\beta - \rho), & \text{если } \beta \neq \rho, \\ (j+1)\beta^j, & \text{если } \beta = \rho. \end{cases}$$

Вычислим теперь ковариации. Имеем:

$$\begin{aligned} \mathrm{Cov}(y_{t-1}, u_t) &= \mathrm{Cov}(y_{t-1}, \rho u_{t-1} + \varepsilon_t) \\ &= \rho \, \mathrm{Cov}(y_{t-1}, u_{t-1}) + \mathrm{Cov}(y_{t-1}, \varepsilon_t) = \rho \, \mathrm{Cov}(y_t, u_t) \\ &= \rho \, \mathrm{Cov}(\beta y_{t-1} + u_t, u_t) = \beta\rho \, \mathrm{Cov}(y_{t-1}, u_t) + \rho \mathrm{V}(u_t). \end{aligned}$$

(Мы здесь использовали стационарность и тот факт, что $\varepsilon_t$ не коррелировано ни с чем, что предшествует периоду $t$.) Отсюда получаем:

$$\mathrm{Cov}(y_{t-1}, u_t) = \omega^2 \rho (1 - \beta^2), \qquad \omega^2 = \frac{\sigma^2}{(1-\beta^2)(1-\rho^2)(1-\beta\rho)}.$$

Таким образом,

$$\begin{aligned} \mathrm{V}(y_t) &= \beta^2 \mathrm{V}(y_{t-1}) + \mathrm{V}(u_t) + 2\beta \, \mathrm{Cov}(y_{t-1}, u_t) \\ &= \beta^2 \mathrm{V}(y_t) + \frac{\sigma^2}{1-\rho^2} + 2\omega^2 \beta\rho(1-\beta^2), \end{aligned}$$

откуда следует, что

$$\gamma_0 = \mathrm{V}(y_t) = \omega^2(1 + \beta\rho).$$

Ковариация первого порядка вычисляется следующим образом:

$$\gamma_1 = \mathrm{Cov}(y_t, y_{t-1}) = \mathrm{Cov}(\beta y_{t-1} + u_t, y_{t-1}) = \beta\gamma_0 + \mathrm{Cov}(y_{t-1}, u_t) = \omega^2(\beta + \rho).$$

Ковариации удовлетворяют тому же разностному уравнению, что и $\psi_t$:

$$\gamma_j = \phi_1 \gamma_{j-1} + \phi_2 \gamma_{j-2},$$

решением которого при $j \geqslant 2$ является

$$\gamma_j = \begin{cases} \omega^2 \left(\beta^{j+1}(1-\rho^2) - \rho^{j+1}(1-\beta^2)\right)/(\beta - \rho), & \text{если } \beta \neq \rho, \\ \omega^2 \beta^j \left(j(1-\beta^2) + (1+\beta^2)\right), & \text{если } \beta = \rho. \end{cases}$$

В частности,

$$\gamma_2 = \omega^2(\beta^2 + \rho^2 + \beta\rho - \beta^2\rho^2).$$

Нетрудно проверить, что для $y_t$ выполнены условия теоремы из задачи 11.6. Обозначим $S_n(j) = n^{-1}\sum_{t=1}^{n} y_t y_{t-j}$. Из теоремы следует, что существуют пределы

$$p\lim S_n(j) = \gamma_j, \qquad j = 0, 1, 2, \ldots$$

Отсюда получаем:

$$p\lim(\widehat{\beta}) = \frac{\gamma_1}{\gamma_0} = \frac{\beta+\rho}{1+\beta\rho}.$$

Таким образом, равенство (11.16) доказано. Чтобы доказать (11.17), выпишем формулу для оценки $\rho$:

$$\widehat{\rho} = \frac{\sum \widehat{u}_t \widehat{u}_{t-1}}{\sum \widehat{u}_{t-1}^2},$$

где $\widehat{u}_t = y_t - \widehat{\beta} y_{t-1}$. Получаем

$$n^{-1}\sum \widehat{u}_t\widehat{u}_{t-1} = S_n(1) - \widehat{\beta}S_n(2) - \widehat{\beta}S_n(0) + \widehat{\beta}^2 S_n(1),$$

и, следовательно,

$$p\lim\left(n^{-1}\sum \widehat{u}_t\widehat{u}_{t-1}\right) = \gamma_1 - \frac{\gamma_1}{\gamma_0}\gamma_2 - \frac{\gamma_1}{\gamma_0}\gamma_0 + \left(\frac{\gamma_1}{\gamma_0}\right)^2\gamma_1 = \frac{\gamma_1}{\gamma_0^2}(\gamma_1^2 - \gamma_0\gamma_2).$$

Аналогично,

$$n^{-1}\sum \widehat{u}_{t-1}^2 = S_n(0) - 2\widehat{\beta}S_n(1) + \widehat{\beta}^2 S_n(0)$$

и поэтому

$$p\lim\left(n^{-1}\sum \widehat{u}_{t-1}^2\right) = \gamma_0 - 2\frac{\gamma_1}{\gamma_0}\gamma_1 + \left(\frac{\gamma_1}{\gamma_0}\right)^2\gamma_0 = \frac{\gamma_0^2 - \gamma_1^2}{\gamma_0}.$$

Из полученных выражений для двух пределов получаем:

$$p\lim(\widehat{\rho}) = \frac{p\lim\left(n^{-1}\sum \widehat{u}_t\widehat{u}_{t-1}\right)}{p\lim\left(n^{-1}\sum \widehat{u}_{t-1}^2\right)} = \frac{\gamma_1}{\gamma_0}\,\frac{\gamma_1^2 - \gamma_0\gamma_2}{\gamma_0^2 - \gamma_1^2} = \frac{\beta\rho(\beta+\rho)}{1+\beta\rho}.$$

Таким образом, равенство (11.17) доказано.

## Задача 11.8

Докажите, что многочлен

$$A(L) = (1 - \phi_1 L - \phi_2 L^2) = (1 - \lambda_1 L)(1 - \lambda_2 L)$$

обратим тогда и только тогда, когда $|\lambda_1| < 1$ и $|\lambda_2| < 1$.

## Решение

Эта задача в определенном смысле является частным случаем задачи 11.3. Приведем здесь, однако, другие соображения, уточняющие понятие обратимости многочлена от оператора лага, или понятие сходимости степенных рядов от оператора лага.

Рассмотрим многочлены

$$A_{1r}(L) = 1 + \lambda_1 L + \lambda_1^2 L^2 + \ldots + \lambda_1^{r-1} L^{r-1},$$
$$A_{2s}(L) = 1 + \lambda_2 L + \lambda_2^2 L^2 + \ldots + \lambda_2^{s-1} L^{s-1}.$$

Очевидно, что

$$A_{1r}(L)(1 - \lambda_1 L) = 1 - \lambda_1^r L^r, \qquad A_{2s}(L)(1 - \lambda_2 L) = 1 - \lambda_2^s L^s.$$

Рассмотрим теперь ограниченную последовательность $y_t$, в том смысле, что существует число $C$ такое, что для всех $t$ выполняется $|y_t| < C$. Рассмотрим действие оператора $A_{1r}(L)A_{2s}(L)A(L)$ на эту последовательность:

$$A_{1r}(L)A_{2s}(L)A(L)y_t = (1 - \lambda_1^r L^r - \lambda_2^s L^s + \lambda_1^r \lambda_2^s L^{r+s})y_t = y_t + u_t.$$

Очевидно, что из ограниченности $y_t$ и того, что $|\lambda_1| < 1$ и $|\lambda_2| < 1$, следует, что $u_t \to 0$ при $s, r \to \infty$. Таким образом, оператор $A_{1r}(L)A_{2s}(L)$ приблизительно равен оператору, обратному $A(L)$, в указанном выше смысле (по его действию на ограниченную последовательность).

Таким образом, оператором, обратным оператору

$$A(L) = (1 - \lambda_1 L)(1 - \lambda_2 L),$$

естественно считать оператор

$$A(L)^{-1} = \lim_{r,s\to\infty}(1+\lambda_1 L+ \ldots + \lambda_1^{r-1} L^{r-1})(1+\lambda_2 L+ \ldots + \lambda_2^{s-1} L^{s-1})$$
$$= \Big(\sum\nolimits_{i=0}^{\infty} \lambda_1^i L^i\Big)\Big(\sum\nolimits_{j=0}^{\infty} \lambda_2^j L^j\Big).$$

Сходимость ряда степеней оператора сдвига здесь понимается в смысле сходимости результата применения его к любой ограниченной последовательности, т. е. ряд $\sum_{i=0}^{\infty} \phi^i L^i$ называется сходящимся, если сходится ряд $\sum_{i=0}^{\infty} |\phi|^i$.

Мы показали, что если выполняются условия $|\lambda_1| < 1$ и $|\lambda_2| < 1$, то оператор $A(L)$ обратим. Из тех же рассуждений очевидно, что если хотя бы одно $|\lambda_i| \geqslant 1$, то оператор $A(L)$ не является обратимым.

## Задача 11.9

Покажите, что тестирование наличия одного единичного корня в процессе AR($p$) можно свести к тесту ADF, приведя уравнение к виду, аналогичному (11.52)

$$\Delta y_t = (\phi_1 + \phi_2 - 1)y_{t-1} - \phi_2 \Delta y_{t-1} + \varepsilon_t, \quad \varepsilon_t \sim iid(0, \sigma^2)$$

для AR(2) модели.

## Решение

Рассмотрим AR($p$) модель с нулевым средним:

$$y_t = \phi_1 y_{t-1} + \phi_2 y_{t-2} + \cdots + \phi_p y_{t-p} + \varepsilon_t,$$

или
$$A(L)y_t = \varepsilon_t, \quad \text{где} \quad A(L) = 1 - \phi_1 L - \phi_2 L^2 - \ldots - \phi_p L^p.$$

Попытаемся представить эту модель в виде, подходящем для теста Дики–Фуллера.

$$\begin{aligned}
\Delta y_t &= (\phi_1 - 1)y_{t-1} + \phi_2 y_{t-2} + \cdots + \phi_p y_{t-p} + \varepsilon_t \\
&= \pi y_{t-1} + c_1 \Delta y_{t-1} + c_2 \Delta y_{t-2} + \cdots + c_{p-1}\Delta y_{t-p+1} + \varepsilon_t \\
&= \pi y_{t-1} + c_1(y_{t-1} - y_{t-2}) + c_2(y_{t-2} - y_{t-3}) + \cdots \\
&\quad + c_{p-1}(y_{t-p+1} - y_{t-p}) + \varepsilon_t.
\end{aligned}$$

Для того чтобы найти коэффициенты $c_i$, $\pi$, приравняем коэффициенты при соответствующих $y_i$. Получим следующую систему уравнений:

$$\begin{gathered}
-c_{p-1} = \phi_p, \\
c_{p-1} - c_{p-2} = \phi_{p-1}, \\
\ldots\ldots\ldots\ldots\ldots\ldots \\
c_3 - c_2 = \phi_3, \\
c_2 - c_1 = \phi_2, \\
\pi + c_1 = \phi_1 - 1.
\end{gathered}$$

Последовательно решая эту систему, получаем:

$$\begin{aligned}
c_{p-1} &= -\phi_p, \\
c_{p-2} &= -\phi_{p-1} + c_{p-1} = -\phi_{p-1} - \phi_p, \\
c_{p-3} &= -\phi_{p-2} + c_{p-2} = -\phi_{p-2} - \phi_{p-1} - \phi_p, \\
&\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\
c_1 &= -\phi_2 - \phi_3 - \ldots - \phi_p, \\
\pi &= \phi_1 - 1 - c_1 = \phi_1 + \phi_2 + \ldots + \phi_p - 1.
\end{aligned}$$

Таким образом, нам удалось представить AR($p$) модель в виде

$$\Delta y_t = \pi y_{t-1} + c_1 \Delta y_{t-1} + c_2 \Delta y_{t-2} + \cdots + c_{p-1}\Delta y_{t-p+1} + \varepsilon_t,$$

где $\pi = \phi_1 + \phi_2 + \ldots + \phi_p - 1 = A(1)$.

Теперь, тестируя (с помощью расширенного теста Дики–Фуллера) гипотезу $\mathrm{H}_0$: $\pi = 0$, мы тестируем условие $A(1) = 0$, т.е. наличие по крайней мере одного единичного корня.

## Задача 11.10

Вычислите частную автокорреляционную функцию PACF($k$) для AR(1) процесса.

## Решение

По определению PACF(1) $= \rho_1$. Для того чтобы найти PACF(2), выпишем соответствующую систему Юла–Уолкера (см. (11.72), а также задачу 11.14):

$$\rho_1 = \phi_1 + \phi_2\rho_1,$$
$$\rho_2 = \phi_1\rho_1 + \phi_2.$$

С другой стороны, из уравнения AR(1) процесса

$$y_t = \phi_1 y_{t-1} + \varepsilon_t,$$

домножив на $y_{t-1}$ и $y_{t-2}$ соответственно, взяв математическое ожидание и разделив на $\gamma_0$, получаем:

$$\rho_1 = \phi_1,$$
$$\rho_2 = \phi_1\rho_1.$$

Сравнивая с предыдущей системой, получаем, что PACF(2) $= \phi_2 = 0$. Аналогично получаем, что PACF($k$) $= \phi_k = 0$ для всех $k \geqslant 2$.

## Задача 11.11

Покажите, что автокорреляционная функция стационарного процесса AR(2) убывает экспоненциально в случае, когда характеристические корни $\Phi(L) = 1 - \phi_1 L - \phi_2 L^2$ действительны, или изменяется по синусоиде с экспоненциально убывающей амплитудой, когда корни комплексные.

## Решение

Вычислив коэффициенты корреляции $y_{t-1}, y_{t-2}, \ldots$ с обеими частями уравнения AR(2) модели

$$y_t = \phi_1 y_{t-1} + \phi_2 y_{t-2} + \varepsilon_t,$$

получим следующую систему уравнений:

$$\rho_1 = \phi_1 + \phi_2\rho_1,$$
$$\rho_2 = \phi_1\rho_1 + \phi_2,$$
$$\rho_k = \phi_1\rho_{k-1} + \phi_2\rho_{k-2}, \quad k \geqslant 3.$$

Из первых двух уравнений можно выразить первые два значения ACF:

$$\rho_1 = \frac{\phi_1}{1-\phi_2}, \qquad \rho_2 = \phi_2 + \frac{\phi_1^2}{1-\phi_2}.$$

Остальные значения $\rho_k$, $k \geqslant 3$, можно найти, как решения разностного уравнения.

В случае общего положения, когда корни характеристического уравнения

$$z^2 - \phi_1 z - \phi_2 = 0$$

$z_1$, $z_2$ различны, общий вид решения имеет вид:

$$\rho_k = c_1 z_1^k + c_2 z_2^k.$$

Заметим, что

$$\Phi(L) = 1 - \phi_1 L - \phi_2 L^2 = (1 - z_1 L)(1 - z_2 L),$$

а из условия стационарности вытекает, что $|z_1| < 1$ и $|z_2| < 1$.

Рассмотрим случай, когда оба корня $z_1$ и $z_2$ вещественные и пусть, для определенности, $|z_1| > |z_2|$. Тогда $\rho_k = c_1 z_1^k + c_2 z_2^k$ является суммой двух экспоненциально убывающих компонент. Асимптотически $\rho_k \approx c_1 z_1^k$, убывает монотонно при $z_1 > 0$ и «мигает» при $z_1 < 0$.

В вырожденном случае, когда корни характеристического уравнения совпадают, общее решение имеет вид:

$$\rho_k = (c_1 + c_2 k) z_1^k,$$

при этом асимптотически $\rho_k \approx c_2 k z_1^k$, т. е. убывает несколько медленнее, чем в случае различных корней.

В случае если корни комплексные, они являются парой комплексно сопряженных чисел: $z_1 = z$ и $z_2 = \bar{z}_1 = \bar{z}$. Из граничных условий

$$c_1 z + c_2 \bar{z} = \rho_1, \quad c_1 z^2 + c_2 \bar{z}^2 = \rho_2$$

получаем

$$c_1 = \frac{\rho_1 \bar{z}^2 - \rho_2 \bar{z}}{z\bar{z}(\bar{z} - z)}, \qquad c_2 = \frac{-\rho_1 z^2 + \rho_2 z}{z\bar{z}(\bar{z} - z)}.$$

Подставляя полученные значения коэффициентов $c_1$, $c_2$ в формулу для $\rho_k$ и представляя $z$ в тригонометрической форме $z = \rho e^{i\phi}$, получаем:

$$\rho_k = \frac{\rho_1 \bar{z}^2 - \rho_2 \bar{z}}{z\bar{z}(\bar{z} - z)} z^k + \frac{-\rho_1 z^2 + \rho_2 z}{z\bar{z}(\bar{z} - z)} \bar{z}^k = -\rho_1 z\bar{z} \frac{z^{k-2} - \bar{z}^{k-2}}{z - \bar{z}} + \rho_2 \frac{z^{k-1} - \bar{z}^{k-1}}{z - \bar{z}}$$

$$= \rho^{k-2} \left( \frac{-\rho\rho_1 \sin(k-2)\phi + \rho_2 \sin(k-1)\phi}{\sin\phi} \right).$$

Легко видеть, что последнее выражение можно преобразовать к виду

$$\rho_k = C\rho^{k-2} \sin(k\phi - \psi),$$

где константы $C$, $\psi$ являются функциями $\rho_1$, $\rho_2$, $\rho$ и $\phi$, откуда и следует утверждение задачи.

## Задача 11.12

Покажите, что условия (11.76)

$$|\phi_2| < 1, \quad \phi_2 + \phi_1 < 1, \quad \phi_2 - \phi_1 < 1$$

стационарности процесса AR(2) эквивалентны тому, что оба корня характеристического уравнения $\Phi(L) = 1 - \phi_1 L - \phi_2 L^2 = 0$ лежат вне единичной окружности.

## Решение

Условие того, что корни уравнения $1 - \phi_1 L - \phi_2 L^2 = 0$ лежат вне единичной окружности, эквивалентно условию, что корни уравнения $z^2 - \phi_1 z - \phi_2 = 0$ лежат внутри единичной окружности: $|z_1| < 1$, $|z_2| < 1$.

Сначала рассмотрим случай, когда оба корня вещественные:

$$\phi_2 \geqslant -\frac{1}{4}\phi_1^2, \qquad z_{1,2} = \frac{\phi_1 \pm \sqrt{\phi_1^2 + 4\phi_2}}{2}.$$

Условия $|z_1| < 1$, $|z_2| < 1$ приводят к следующей системе неравенств:

$$\begin{cases} -2 < \phi_1 + \sqrt{\phi_1^2 + 4\phi_2} < 2, \\ -2 < \phi_1 - \sqrt{\phi_1^2 + 4\phi_2} < 2, \end{cases}$$

учитывая, что значение квадратного корня неотрицательно, получаем эквивалентную систему:

$$\begin{cases} \sqrt{\phi_1^2 + 4\phi_2} < 2 - \phi_1, \\ \sqrt{\phi_1^2 + 4\phi_2} < 2 + \phi_1. \end{cases}$$

Разобьем систему на два случая:

$$\begin{cases} 0 \leqslant \phi_1, \\ \sqrt{\phi_1^2 + 4\phi_2} < 2 - \phi_1; \end{cases} \qquad \begin{cases} \phi_1 < 0, \\ \sqrt{\phi_1^2 + 4\phi_2} < 2 + \phi_1, \end{cases}$$

или

$$\begin{cases} 0 \leqslant \phi_1 < 2, \\ \phi_1^2 + 4\phi_2 < (2 - \phi_1)^2; \end{cases} \qquad \begin{cases} -2 < \phi_1 < 0, \\ \phi_1^2 + 4\phi_2 < (2 + \phi_1)^2, \end{cases}$$

получаем:

$$\begin{cases} 0 \leqslant \phi_1 < 2, \\ \phi_2 < 1 - \phi_1; \end{cases} \qquad \begin{cases} -2 < \phi_1 < 0, \\ \phi_2 < 1 + \phi_1. \end{cases}$$

С учетом условия вещественности корней получаем следующую систему условий, выделяющую на плоскости $(\phi_1, \phi_2)$ область криволинейного треугольника (см. рис. 11.1), соответствующую паре вещественных корней, меньших 1 по абсолютной величине:

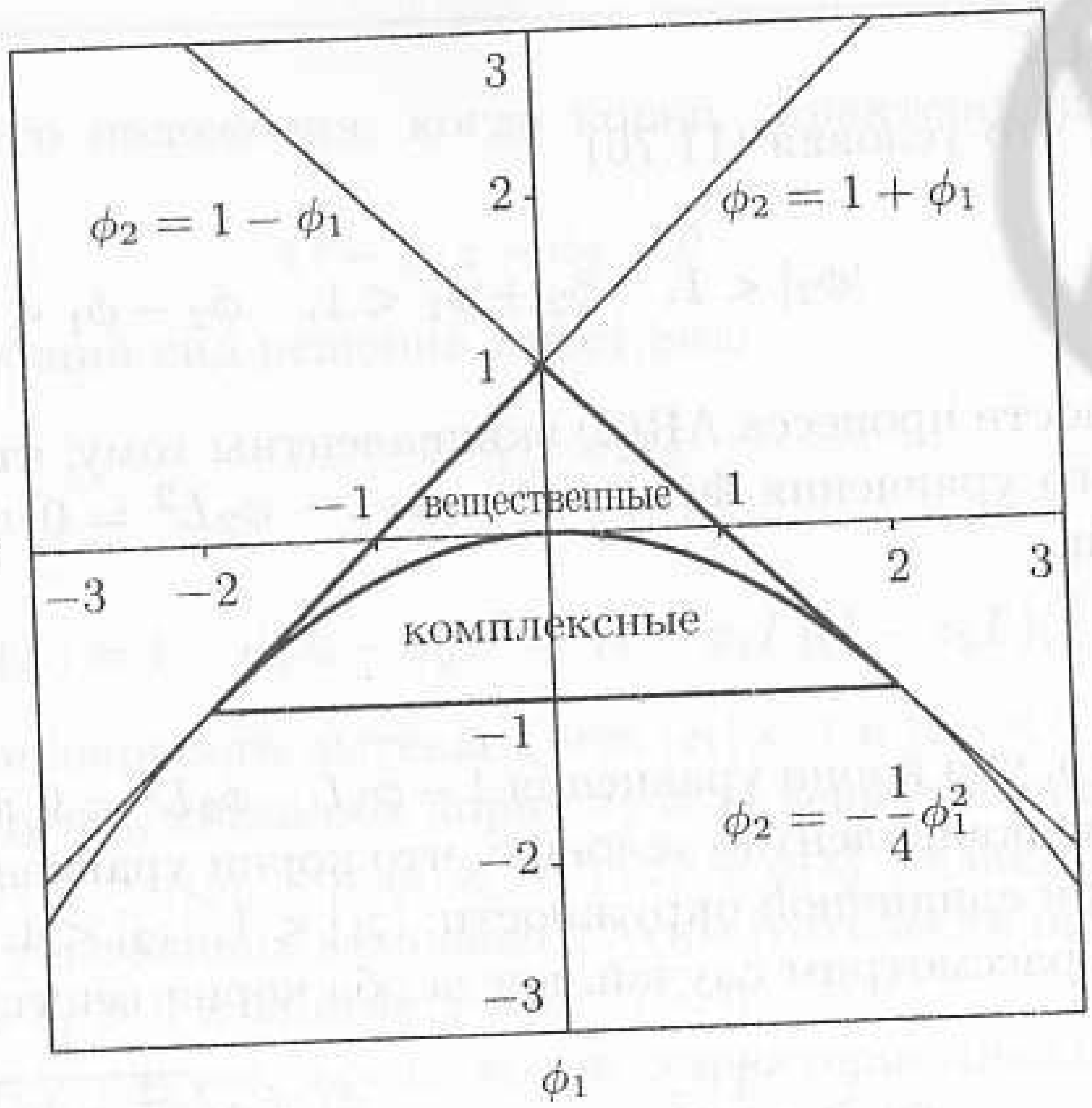



**Рис. 11.1**

$$\begin{cases} \phi_2 \geqslant -\dfrac{1}{4}\phi_1^2, \\ \phi_2 < 1 - \phi_1, \\ \phi_2 < 1 + \phi_1, \\ -2 < \phi_1 < 2. \end{cases}$$

Рассмотрим теперь случай двух сопряженных комплексных корней, $\phi_2 < -\frac{1}{4}\phi_1^2$:

$$z_{1,2} = \frac{\phi_1 \pm i\sqrt{-\phi_1^2 - 4\phi_2}}{2},$$

$$|z_1|^2 = |z_2|^2 = \frac{1}{4}\left(\phi_1^2 + (-\phi_1^2 - 4\phi_2)\right) = -\phi_2.$$

Условие наличия двух комплексных корней, по модулю меньших 1, выделяет на плоскости $(\phi_1, \phi_2)$ область (см. рис. 11.1):

$$\begin{cases} \phi_2 < -\dfrac{1}{4}\phi_1^2, \\ -1 < \phi_2. \end{cases}$$

Объединим две области, полученные для случаев вещественных и комплексных корней, в одну:

$$\begin{cases} \phi_2 < 1 - \phi_1, \\ \phi_2 < 1 + \phi_1, \\ -1 < \phi_2, \end{cases}$$

являющуюся треугольником на плоскости $(\phi_1, \phi_2)$ с вершинами в точках (0,1), (-2,-1), (2,-1), что и требовалось показать.

## Задача 11.13

Вычислите частную автокорреляционную функцию PACF($k$) для AR(2) процесса.

## Решение

По определению, $\text{PACF}(1) = \rho_1$. Для того чтобы найти PACF(2), выпишем соответствующую систему Юла–Уолкера (см. (11.72), а также задачу 11.14):

$$\rho_1 = \phi_1 + \phi_2\rho_1,$$
$$\rho_2 = \phi_1\rho_1 + \phi_2,$$

решая которую получаем

$$\text{PACF}(2) = \phi_2 = \frac{\rho_2 - \rho_1^2}{1 - \rho_1^2}.$$

Для того чтобы найти PACF(3), выпишем соответствующую систему Юла–Уолкера третьего порядка:

$$\begin{aligned} \rho_1 &= \phi_1 + \phi_2\rho_1 + \phi_3\rho_2, \\ \rho_2 &= \phi_1\rho_1 + \phi_2 + \phi_3\rho_1, \\ \rho_3 &= \phi_1\rho_2 + \phi_2\rho_1 + \phi_3. \end{aligned}$$

С другой стороны, из уравнения AR(2) процесса

$$y_t = \phi_1 y_{t-1} + \phi_2 y_{t-2} + \varepsilon_t,$$

умножив на $y_{t-3}$, взяв математическое ожидание и разделив на $\gamma_0$, получаем:

$$\rho_3 = \phi_1\rho_2 + \phi_2\rho_1.$$

Сравнивая с предыдущей системой, получаем, что $\text{PACF}(3) = \phi_3 = 0$. Аналогично получаем, что $\text{PACF}(k) = \phi_k = 0$ для всех $k \geqslant 3$.

## Задача 11.14

Примените процедуру вычисления выборочного частного коэффициента корреляции (см. п. 4.3) для стационарного ряда $y_t$. Покажите, что $k$-е значение выборочной частной автокорреляционной функции PACF($k$) вычисляется как МНК-оценка последнего коэффициента $\beta_k$ в AR($k$) регрессионном уравнении:

$$y_t = \beta_0 + \beta_1 y_{t-1} + \beta_2 y_{t-2} + \cdots + \beta_k y_{t-k} + \varepsilon_t.$$

## Решение

Рассмотрим параллельно случаи теоретической и выборочной частной автокорреляционной функции PACF($k$). Сначала рассмотрим линейные пространства и скалярные произведения, участвующие в определении теоретических и выборочных корреляций и регрессий.

*Пространство случайных величин.* Пусть $x_0, x_1, \ldots, x_{k-1}, x_k, \ldots$ — последовательность случайных величин, являющаяся слабо стационарным процессом. Везде в этой задаче для упрощения вычислений мы будем предполагать, что имеем дело с процессом с нулевым средним: $\mathrm{E}(x_s) = 0$; будем также предполагать случай общего положения: любой конечный набор случайных величин $x_i$ является линейно независимым. Рассмотрим $(k+1)$-мерное векторное пространство $L$, порожденное линейными комбинациями случайных величин $x_0, x_1, \ldots, x_k$, $L = \{\sum_{i=0}^{k} c_i x_i\}$, и его $(k-1)$-мерное подпространство (коразмерности 2) $\pi = \{\sum_{i=1}^{k-1} c_i x_i\}$. В качестве скалярного произведения в $L$ возьмем ковариацию: $(x, y) = \mathrm{Cov}(x, y)$. (Мы сохраним здесь скалярные обозначения для векторо́в из $L$, чтобы подчеркнуть, что они являются скалярными случайными величинами.)

*Пространство выборочных значений.* Пусть $\{y_t\}$ — реализация слабо стационарного процесса. Рассмотрим векторы $x_t = (y_t, y_{t+1}, \ldots, y_{t+n-1})'$, $t = 0, 1, \ldots, k-1, k$, принадлежащие стандартному $n$-мерному евклидову пространству $R^n$ со скалярным произведением $(u, v) = (1/n)\sum_{i=1}^{n} u_i v_i = (1/n)u'v$. (Отметим, что мы используем в этой задаче обозначение $x_t$ для векторов, чтобы сохранить единообразные обозначения со случаем пространства скалярных случайных величин.) Рассмотрим $(k+1)$-мерное пространство $L = \{\sum_{i=0}^{k} c_i x_i\}$, и его $(k-1)$-мерное подпространство (коразмерности 2) $\pi = \{\sum_{i=1}^{k-1} c_i x_i\}$.

*PACF(k). Первое определение.* Обозначим через $P$ оператор ортогонального проектирования на $\pi$. Рассмотрим следующие разложения векторов:

$$x_0 = \widehat{x}_0 + \widehat{e}_0, \qquad x_k = \widehat{x}_k + \widehat{e}_k.$$

Здесь $\widehat{x}_0 = Px_0$ и $\widehat{x}_k = Px_k$. По построению, векторы $\widehat{e}_0, \widehat{e}_k$ ортогональны $\pi$. Определим $\mathrm{PACF}_1$ как:

$$\mathrm{PACF}_1(k) = \mathrm{Corr}\,(\widehat{e}_0, \widehat{e}_k) = \frac{(\widehat{e}_0, \widehat{e}_k)}{\sqrt{(\widehat{e}_0, \widehat{e}_0)}\sqrt{(\widehat{e}_k, \widehat{e}_k)}}. \tag{1}$$

*PACF(k). Второе определение.* Рассмотрим подпространство

$$\widetilde{L} = \{\textstyle\sum_{i=1}^{k} c_i x_i\}, \quad \widetilde{L} \subset L.$$

Обозначим через $\widetilde{P}$ оператор ортогонального проектирования на $\widetilde{L}$. Рассмотрим следующие разложения:

$$x_0 = \widetilde{x}_0 + \gamma x_k + \widetilde{e}_0. \tag{2}$$

Здесь первые два слагаемых являются проекцией вектора $x_0$ на $\widetilde{L}$: $\widetilde{P}x_0 = \widetilde{x}_0 + \gamma x_k$, $\widetilde{x}_0 \in \pi$, а вектор $\widetilde{e}_0$ ортогонален $\widetilde{L}$. ($\widetilde{e}_0$ ортогонален как $\pi$, так и $x_k$.) Определим $\mathrm{PACF}_2$ как:

$$\mathrm{PACF}_2(k) = \gamma. \tag{3}$$

Применим оператор $P$ к уравнению (2):

$$\widehat{x}_0 = \widetilde{x}_0 + \gamma \widehat{x}_k, \tag{4}$$

и вычтем (4) из (2):

$$\widehat{e}_0 = \gamma \widehat{e}_k + \widetilde{e}_0. \tag{5}$$

Возьмем теперь скалярное произведение обеих частей равенства (5) с $\widehat{e}_k$:

$$(\widehat{e}_0, \widehat{e}_k) = \gamma(\widehat{e}_k, \widehat{e}_k),$$

получим:

$$\mathrm{PACF}_2(k) = \gamma = \frac{(\widehat{e}_0, \widehat{e}_k)}{(\widehat{e}_k, \widehat{e}_k)}. \tag{6}$$

*Эквивалентность двух определений теоретической PACF(k).* Из (1) и (6) получаем

$$\mathrm{PACF}_1(k) = \mathrm{PACF}_2(k) \frac{\sqrt{(\widehat{e}_k, \widehat{e}_k)}}{\sqrt{(\widehat{e}_0, \widehat{e}_0)}}, \tag{7}$$

или

$$\mathrm{PACF}_1(k) = \mathrm{PACF}_2(k) \frac{\sqrt{\mathrm{V}(\widehat{e}_k)}}{\sqrt{\mathrm{V}(\widehat{e}_0)}},$$

поэтому, чтобы доказать эквивалентность двух определений теоретической PACF, нам достаточно проверить равенство

$$\mathrm{V}(\widehat{e}_0) = \mathrm{V}(\widehat{e}_k). \tag{8}$$

Очевидно,

$$\mathrm{V}(\widehat{e}_0) = \mathrm{V}(x_0) - \mathrm{V}(\widehat{x}_0), \qquad \mathrm{V}(\widehat{e}_k) = \mathrm{V}(x_k) - \mathrm{V}(\widehat{x}_k).$$

(С геометрической точки зрения это — теорема Пифагора.) Таким образом, необходимо проверить равенство

$$\mathrm{V}(\widehat{x}_0) = \mathrm{V}(\widehat{x}_k). \tag{9}$$

Введем следующие обозначения: $\boldsymbol{x} = (x_1, x_2, \ldots, x_{k-1})'$ — $(k-1 \times 1)$ случайный вектор, $\mathrm{V}(\boldsymbol{x})$ — $(k-1 \times k-1)$ матрица. $\mathrm{V}(\boldsymbol{x})_{ij} = \gamma_{|i-j|}$, где $\gamma_s = \mathrm{Cov}(x_t, x_{t-s})$, и $\mathrm{Cov}(\boldsymbol{x}, x_0) = (\gamma_1, \gamma_2, \ldots, \gamma_{k-1})'$ — $(k-1 \times 1)$ случайный вектор.

Используя стационарность процесса $\{x_t\}$, получаем:

$$\mathrm{V}(\widehat{x}_0) = \mathrm{V}\left(\boldsymbol{x}'\mathrm{V}(\boldsymbol{x})^{-1}\,\mathrm{Cov}(\boldsymbol{x}, x_0)\right) = \mathrm{Cov}(\boldsymbol{x}, x_0)'\mathrm{V}(\boldsymbol{x})^{-1}\,\mathrm{Cov}(\boldsymbol{x}, x_0)$$

$$= \begin{bmatrix} \gamma_1 & \gamma_2 & \cdots & \gamma_{k-1} \end{bmatrix} \mathrm{V}(\boldsymbol{x})^{-1} \begin{bmatrix} \gamma_1 \\ \gamma_2 \\ \vdots \\ \gamma_{k-1} \end{bmatrix},$$

$$\mathrm{V}(\widehat{x}_k) = \begin{bmatrix} \gamma_{k-1} & \gamma_{k-2} & \cdots & \gamma_1 \end{bmatrix} \mathrm{V}(\boldsymbol{x})^{-1} \begin{bmatrix} \gamma_{k-1} \\ \gamma_{k-2} \\ \vdots \\ \gamma_1 \end{bmatrix}.$$

Отсюда

$$\mathrm{V}(\widehat{x}_0) = \mathrm{V}(\widehat{x}_k) = \sum_{st} \gamma_t \gamma_s \gamma'_{st}.$$

(Здесь $\gamma'_{st} = (\mathrm{V}(x)^{-1})_{st}$. Симметричная матрица $\boldsymbol{A} = \mathrm{V}(\boldsymbol{x})$ инвариантна относительно перенумерации строк и столбцов в обратном порядке: $a_{i,j} \equiv a_{k-i,k-j}$, очевидно, этим же свойством обладает и обратная матрица: $\boldsymbol{A}^{-1} = \mathrm{V}(\boldsymbol{x})^{-1}$.)

Таким образом, мы получили, что $\mathrm{PACF}_1 \equiv \mathrm{PACF}_2$.

*Способ вычисления теоретической PACF(k).* Таким образом, теоретический PACF$(k)$ можно вычислить как коэффициент при $x_k$ в разложении по векторам $x_1, x_2, \ldots, x_k$ ортогональной проекции $\widetilde{P}x_0$ вектора $x_0$ на подпространство $\widetilde{L}$. Пусть

$$x_0 = \sum_{i=1}^{k} x_i \phi_i + \widetilde{e}_0 = \boldsymbol{x}'\boldsymbol{\phi} + \widetilde{e}_0.$$

Здесь $\boldsymbol{x} = (x_1, x_2, \ldots, x_k)'$, $\boldsymbol{\phi} = (\phi_1, \phi_2, \ldots, \phi_k)'$. Поскольку вектор $\widetilde{e}_0$ ортогонален $\widetilde{L}$, получаем:

$$\mathrm{Cov}(\boldsymbol{x}, x_0) = \mathrm{Cov}(\boldsymbol{x}, \boldsymbol{x}'\boldsymbol{\phi}), \quad \text{или} \quad \mathrm{Cov}(\boldsymbol{x}, x_0) = \mathrm{V}(\boldsymbol{x})\boldsymbol{\phi}.$$

Разделив последнее уравнение на $\gamma_0$ и обозначив через $\rho_s = \gamma_s/\gamma_0$ коэффициенты корреляции, получим уравнения Юла–Уолкера (11.72):

$$\begin{bmatrix} 1 & \rho_1 & \rho_2 & \cdots & \rho_{k-1} \\ \rho_1 & 1 & \rho_1 & \cdots & \rho_{k-2} \\ \rho_2 & \rho_1 & 1 & \cdots & \rho_{k-3} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \rho_{k-1} & \rho_{k-2} & \rho_{k-3} & \cdots & 1 \end{bmatrix} \begin{bmatrix} \phi_1 \\ \phi_2 \\ \phi_3 \\ \vdots \\ \phi_k \end{bmatrix} = \begin{bmatrix} \rho_1 \\ \rho_2 \\ \rho_3 \\ \vdots \\ \rho_k \end{bmatrix}.$$

Решая эту систему уравнений, мы находим $\mathrm{PACF}(k) = \phi_k$.

*Выборочная PACF(k).* В случае пространства выборочных значений, конечно, условие

$$(\widehat{e}_0, \widehat{e}_k) = \gamma(\widehat{e}_k, \widehat{e}_k)$$

выполняется лишь приблизительно, и соответственно (см. (6)) два определения PACF$(k)$ совпадают лишь асимптотически. На практике из соображений простоты вычислений берут второе определение, а именно PACF$(k)$ вычисляется как МНК-оценка последнего коэффициента $\beta_k$ в AR$(k)$ регрессионном уравнении:

$$y_t = \beta_0 + \beta_1 y_{t-1} + \beta_2 y_{t-2} + \cdots + \beta_k y_{t-k} + \varepsilon_t.$$

## Задача 11.15

Покажите, что для MA($q$) процесса ACF($k$)=0 при $k > q$.

## Решение

Из определения MA($q$) процесса (11.77) имеем:

$$y_t = \delta + \varepsilon_t - \theta_1\varepsilon_{t-1} - \cdots - \theta_q\varepsilon_{t-q},$$
$$y_{t-k} = \delta + \varepsilon_{t-k} - \theta_1\varepsilon_{t-k-1} - \cdots - \theta_q\varepsilon_{t-k-q}.$$

Очевидно, что при $k > q$ выражения для $y_t$ и $y_{t-k}$ не содержат общих $\varepsilon_s$. Поэтому, так как $\varepsilon_s$ и $\varepsilon_t$ при $s \neq t$ являются некоррелированными, имеем: $\mathrm{ACF}(k) = r(y_t, y_{t-k}) = 0$.

## Задача 11.16

Сформулируйте условия обратимости MA(2) процесса

$$y_t = \delta + \varepsilon_t - \theta_1\varepsilon_{t-1} - \theta_2\varepsilon_{t-2}.$$

## Решение

Рассмотрим квадратный многочлен $\Theta(L) = 1 - \theta_1 L - \theta_2 L^2$, соответствующий MA(2) процессу $y_t = \delta + \Theta(L)\varepsilon_t$. Для обратимости процесса, т. е. представления его в AR($\infty$) виде, необходимо существование обратного оператора $\Theta(L)^{-1}$:

$$\Theta(L)^{-1}y_t = \Theta(L)^{-1}\delta + \varepsilon_t = \frac{\delta}{1-\theta_1-\theta_2} + \varepsilon_t.$$

Для обратимости оператора необходимо и достаточно, чтобы все корни многочлена $1 - \theta_1 L - \theta_2 L^2$ лежали вне единичной окружности. В задаче 11.12 были найдены условия обратимости аналогичного многочлена $\Phi(L) = 1 - \phi_1 L - \phi_2 L^2 = 0$.

Соответственно ответом является система неравенств (ср. (11.76)):

$$\begin{cases} \theta_2 < 1 - \theta_1, \\ \theta_2 < 1 + \theta_1, \\ -1 < \theta_2. \end{cases}$$

## Задача 11.17

Покажите, что для AR(1) процесса (11.67) в виде $y_t - \mu = \phi_1(y_{t-1} - \mu) + \varepsilon_t$ прогноз на $s$ шагов вперед вычисляется по формуле $\widehat{y}_{n+s} = \mu + \phi_1^s(y_n - \mu)$, а дисперсия ошибки прогноза равна $\mathrm{V}(e_{n+s}) = \left(1 + \phi_1^2 + \phi_1^4 + \cdots + \phi_1^{2s-2}\right)\sigma^2$.

## Решение

Согласно (11.68) для AR(1) процесса $y_t = \delta + \phi_1 y_{t-1} + \varepsilon_t$ имеем:

$$\mu = \mathrm{E}(y_t) = \frac{\delta}{1-\phi_1}, \quad \mathrm{V}(y_t) = \gamma_0 = \frac{\sigma^2}{1-\phi_1^2}.$$

Обозначим $z_t = y_t - \mu$, тогда AR(1) процесс имеет вид: $z_t = \phi_1 z_{t-1} + \varepsilon_t$. Итерируя эту формулу, получаем:

$$\begin{aligned} z_{n+s} &= \phi_1 z_{n+s-1} + \varepsilon_{n+s} \\ &= \phi_1^2 z_{n+s-2} + \phi_1 \varepsilon_{n+s-1} + \varepsilon_{n+s} \\ &\dots \\ &= \phi_1^s z_n + \phi_1^{s-1}\varepsilon_{n+1} + \phi_1^{s-2}\varepsilon_{n+2} + \dots + \phi_1 \varepsilon_{n+s-1} + \varepsilon_{n+s}. \end{aligned}$$

Наилучший прогноз при информации $I_n = \{z_n, z_{n-1}, \dots\}$, доступной в момент $n$, равен:

$$\widehat{z}_{n+s} = \mathrm{E}(z_{n+s}|I_n) = \phi_1^s z_n.$$

Прогноз $\widehat{y}_{n+s}$ равен:

$$\widehat{y}_{n+s} = \mu + \widehat{z}_{n+s} = \mu + \phi_1^s (y_n - \mu).$$

Соответственно ошибка прогноза равна

$$\begin{aligned} e_{n+s} &= y_{n+s} - \widehat{y}_{n+s} = z_{n+s} - \widehat{z}_{n+s} \\ &= \phi_1^{s-1}\varepsilon_{n+1} + \phi_1^{s-2}\varepsilon_{n+2} + \dots + \phi_1 \varepsilon_{n+s-1} + \varepsilon_{n+s}, \end{aligned}$$

а в силу некоррелированности $\varepsilon_s$ и $\varepsilon_t$ при $s \neq t$ ее дисперсия равна

$$\mathrm{V}(e_{n+s}) = \left(1 + \phi_1^2 + \phi_1^4 + \dots + \phi_1^{2s-2}\right)\sigma^2.$$

## Задача 11.18

Покажите, что для MA(1) процесса (11.78) $y_t = \delta + \varepsilon_t - \theta_1 \varepsilon_{t-1}$ прогноз на $s$ шагов вперед вычисляется по формулам $\widehat{y}_{n+1} = \delta - \theta_1 \varepsilon_n$, $\widehat{y}_{n+s} = \delta$, $s \geqslant 2$, а дисперсия ошибки прогноза равна $\mathrm{V}(e_{n+s}) = \left(1 + \theta_1^2\right)\sigma^2 = \mathrm{V}(y_t)$.

## Решение

Наилучший прогноз на один шаг при информации $I_n = \{y_n, y_{n-1}, \dots\}$, доступной в момент $n$, равен:

$$\begin{aligned} \widehat{y}_{n+1} &= \mathrm{E}(y_{n+1}|I_n) = \mathrm{E}(\delta + \varepsilon_{n+1} - \theta_1 \varepsilon_n | I_n) = \delta - \theta_1 \varepsilon_n, \\ \mathrm{V}(e_{n+1}) &= \mathrm{V}(y_{n+1} - \widehat{y}_{n+1}) = \mathrm{V}(\delta + \varepsilon_{n+1} - \theta_1 \varepsilon_n - \delta + \theta_1 \varepsilon_n) \\ &= \mathrm{V}(\varepsilon_{n+1}) = \sigma^2. \end{aligned}$$

Для прогноза на $s \geqslant 2$ шагов получаем:

$$\begin{aligned}\widehat{y}_{n+s} &= \mathrm{E}(y_{n+s}|I_n) = \mathrm{E}(\delta + \varepsilon_{n+s} - \theta_1\varepsilon_{n+s-1}|I_n) = \delta,\\ \mathrm{V}(e_{n+s}) &= \mathrm{V}(y_{n+s} - \widehat{y}_{n+s}) = \mathrm{V}(\delta + \varepsilon_{n+s} - \theta_1\varepsilon_{n+s-1} - \delta)\\ &= \mathrm{V}(\varepsilon_{n+s} - \theta_1\varepsilon_{n+s-1}) = (1+\theta_1^2)\sigma^2.\end{aligned}$$

## Задача 11.19

Для ARMA(1,1) модели $y_t = \delta + \phi_1 y_{t-1} + \varepsilon_t - \theta_1\varepsilon_{t-1}$ выведите формулы (11.104):

$$(\widehat{y}_{n+s} - \mu) = \phi_1^s(y_n - \mu) - \phi_1^{s-1}\theta_1\varepsilon_n,$$

и (11.105):

$$\lim_{s\to\infty} \mathrm{V}(e_{n+s}) = \sigma^2\left(\frac{1 - 2\phi_1\theta_1 + \theta_1^2}{1-\phi_1^2}\right).$$

## Решение

Нетрудно вычислить математическое ожидание стационарного процесса ARMA(1,1):

$$\mathrm{E}(y_t) = \mu = \frac{\delta}{1-\phi_1}.$$

Переходя к центрированной переменной $z_t = y_t - \mu$, получаем (ср. (11.99)):

$$z_t = \phi_1 z_{t-1} + \varepsilon_t - \theta_1\varepsilon_{t-1}.$$

Докажем первую формулу (11.104) индукцией по $s$ — дальности горизонта прогнозирования.

1) $s = 1$. Наилучшим прогнозом является условное математическое ожидание (см. задачи 11.17, 11.18):

$$\widehat{z}_{n+1} = \mathrm{E}(\phi_1 z_n + \varepsilon_{n+1} - \theta_1\varepsilon_n|I_n) = \phi_1 z_n - \theta_1\varepsilon_n.$$

2) Пусть теперь формула верна для некоторого $s \geqslant 1$:

$$\widehat{z}_{n+s} = \phi_1^s z_n - \phi_1^{s-1}\theta_1\varepsilon_n.$$

Докажем ее для $s+1$:

$$\begin{aligned}\widehat{z}_{n+s+1} &= \mathrm{E}(\phi_1 z_{n+s} + \varepsilon_{n+s+1} - \theta_1\varepsilon_{n+s}|I_n)\\ &= \phi_1\widehat{z}_{n+s} = \phi_1(\phi_1^s z_n - \phi_1^{s-1}\theta_1\varepsilon_n)\\ &= \phi_1^{s+1} z_n - \phi_1^s\theta_1\varepsilon_n,\end{aligned}$$

что и требовалось доказать.

Рассмотрим теперь ошибку прогноза:

$$\begin{aligned} e_{n+s} &= y_{n+s} - \widehat{y}_{n+s} = z_{n+s} - \widehat{z}_{n+s} \\ &= (\phi_1 z_{n+s-1} + \varepsilon_{n+s} - \theta_1 \varepsilon_{n+s-1}) - (\phi_1 \widehat{z}_{n+s-1}) \\ &= \phi_1 e_{n+s-1} + \varepsilon_{n+s} - \theta_1 \varepsilon_{n+s-1}. \end{aligned}$$

Дисперсия ошибки прогноза равна:

$$\begin{aligned} \mathrm{V}(e_{n+s}) &= \mathrm{V}(\phi_1 e_{n+s-1} + \varepsilon_{n+s} - \theta_1 \varepsilon_{n+s-1}) \\ &= \sigma^2 + \phi_1^2 \mathrm{V}(e_{n+s-1}) + \sigma^2\theta_1^2 - 2\theta_1\phi_1 \operatorname{Cov}(e_{n+s-1}, \varepsilon_{n+s-1}). \end{aligned}$$

(Мы здесь воспользовались некоррелированностью $\varepsilon_{n+s}$ с ошибкой прогноза $e_{n+s-1}$, которая содержит только предыдущие возмущения $\varepsilon_k$, $k < n+s$.)

Последнее слагаемое легко вычисляется:

$$\begin{aligned} \operatorname{Cov}(e_{n+s-1}, \varepsilon_{n+s-1}) &= \operatorname{Cov}(z_{n+s-1} - \widehat{z}_{n+s-1}, \varepsilon_{n+s-1}) \\ &= \operatorname{Cov}(z_{n+s-1}, \varepsilon_{n+s-1}) = \operatorname{Cov}(\varepsilon_{n+s-1}, \varepsilon_{n+s-1}) = \sigma^2. \end{aligned}$$

Отсюда получаем рекуррентную формулу для дисперсии:

$$\begin{aligned} \mathrm{V}(e_{n+s}) &= \sigma^2 + \phi_1^2 \mathrm{V}(e_{n+s-1}) + \sigma^2\theta_1^2 - 2\theta_1\phi_1\sigma^2 \\ &= (1 - 2\phi_1\theta_1 + \theta_1^2)\sigma^2 + \phi_1^2 \mathrm{V}(e_{n+s-1}). \end{aligned}$$

Вычислим дисперсию прогноза на один шаг:

$$\begin{aligned} \mathrm{V}(e_{n+1}) &= \mathrm{V}(z_{n+1} - \widehat{z}_{n+1}) \\ &= \mathrm{V}(\phi_1 z_n + \varepsilon_{n+1} - \theta_1\varepsilon_n - (\phi_1 z_n - \theta_1 \varepsilon_n)) = \mathrm{V}(\varepsilon_{n+1}) = \sigma^2. \end{aligned}$$

Для сокращения выкладок введем обозначение

$$c = (1 - 2\phi_1\theta_1 + \theta_1^2)\sigma^2.$$

Производя последовательные подстановки, получаем:

$$\begin{aligned} \mathrm{V}(e_{n+1}) &= \sigma^2, \\ \mathrm{V}(e_{n+2}) &= c + \phi_1^2 \mathrm{V}(e_{n+1}) = c + \phi_1^2\sigma^2, \\ \mathrm{V}(e_{n+3}) &= c + \phi_1^2 \mathrm{V}(e_{n+2}) = c + \phi_1^2(c + \phi_1^2\sigma^2) = c + \phi_1^2 c + \phi_1^4\sigma^2, \\ &\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \end{aligned}$$

$$\begin{aligned} \mathrm{V}(e_{n+s}) &= c + \phi_1^2 c + \phi_1^4 c + \cdots + \phi_1^{2s-4} c + \phi_1^{2s-2}\sigma^2 \\ &= \frac{1 - \phi_1^{2s-2}}{1 - \phi_1^2} c + \phi_1^{2s-2}\sigma^2 \\ &= \frac{1 - \phi_1^{2s-2}}{1 - \phi_1^2} (1 - 2\phi_1\theta_1 + \theta_1^2)\sigma^2 + \phi_1^{2s-2}\sigma^2. \end{aligned}$$

Переходя к пределу и учитывая, что $|\phi_1| < 1$, получаем:

$$\lim_{s\to\infty} \mathrm{V}(e_{n+s}) = \frac{1}{1-\phi_1^2}(1 - 2\phi_1\theta_1 + \theta_1^2)\sigma^2,$$

что и требовалось доказать.

## Задача 11.20

Вычислите автокорреляционную функцию ACF($k$) и частную автокорреляционную функцию PACF($k$) для MA(2) процесса.

## Решение

Уравнение MA(2) процесса имеет вид:

$$y_t = \delta + \varepsilon_t - \theta_1\varepsilon_{t-1} - \theta_q\varepsilon_{t-2}, \quad \varepsilon_t \sim iid(0, \sigma^2).$$

Без ограничения общности можно считать $\delta = 0$. В силу задачи 11.15 имеем ACF($k$) = 0, при $k = 3, 4, \ldots$. Вычислим первые два значения ACF. Для этого сначала вычислим дисперсию и ковариации:

$$\begin{aligned}
\gamma_0 &= \mathrm{V}(\varepsilon_t - \theta_1\varepsilon_{t-1} - \theta_2\varepsilon_{t-2}) = \sigma^2(1 + \theta_1^2 + \theta_2^2),\\
\gamma_1 &= \mathrm{Cov}(\varepsilon_t - \theta_1\varepsilon_{t-1} - \theta_2\varepsilon_{t-2}, \varepsilon_{t-1} - \theta_1\varepsilon_{t-2} - \theta_2\varepsilon_{t-3}) = \sigma^2(-\theta_1 + \theta_1\theta_2),\\
\gamma_2 &= \mathrm{Cov}(\varepsilon_t - \theta_1\varepsilon_{t-1} - \theta_2\varepsilon_{t-2}, \varepsilon_{t-2} - \theta_1\varepsilon_{t-3} - \theta_2\varepsilon_{t-4}) = \sigma^2(-\theta_2).
\end{aligned}$$

Отсюда получаем значения ACF:

$$\mathrm{ACF}(1) = \frac{\gamma_1}{\gamma_0} = \frac{-\theta_1 + \theta_1\theta_2}{1 + \theta_1^2 + \theta_2^2},$$

$$\mathrm{ACF}(2) = \frac{\gamma_2}{\gamma_0} = \frac{-\theta_2}{1 + \theta_1^2 + \theta_2^2}.$$

Значения частной автокорреляционной функции PACF($k$) находятся из системы уравнений Юла–Уолкера (11.72), см. также решение задачи 11.14. Найдем три первых значения. По определению, PACF(1) = ACF(1) = $\rho_1$. PACF(2) равно значению $\phi_2$, найденному из системы

$$\begin{bmatrix} 1 & \rho_1 \\ \rho_1 & 1 \end{bmatrix} \begin{bmatrix} \phi_1 \\ \phi_2 \end{bmatrix} = \begin{bmatrix} \rho_1 \\ \rho_2 \end{bmatrix}.$$

Получаем:

$$\mathrm{PACF}(2) = \phi_2 = \frac{-\rho_1^2 + \rho_2}{1 - \rho_1^2}.$$

PACF(3) равно значению $\phi_3$, найденному из системы

$$\begin{bmatrix} 1 & \rho_1 & \rho_2 \\ \rho_1 & 1 & \rho_1 \\ \rho_2 & \rho_1 & 1 \end{bmatrix} \begin{bmatrix} \phi_1 \\ \phi_2 \\ \phi_3 \end{bmatrix} = \begin{bmatrix} \rho_1 \\ \rho_2 \\ 0 \end{bmatrix}.$$

Получаем

$$\text{PACF}(3) = \phi_3 = \frac{\rho_1\rho_2^2 + \rho_1^3 - 2\rho_1\rho_2}{1 + 2\rho_1^2\rho_2 - \rho_2^2 - 2\rho_1^2}.$$

В общем случае, для того, чтобы найти PACF($k$) надо найти величину $\phi_k$ из следующей системы $k$ уравнений относительно $k$ неизвестных $\phi_1, \ldots, \phi_k$

$$\begin{bmatrix} 1 & \rho_1 & \rho_2 & 0 & & \cdots & 0 \\ \rho_1 & 1 & \rho_1 & \rho_2 & \ddots & & \vdots \\ \rho_2 & \rho_1 & \ddots & \ddots & \ddots & \ddots & \\ 0 & \rho_2 & \ddots & & & & 0 \\ & \ddots & \ddots & & & \ddots & \rho_2 \\ \vdots & & \ddots & \ddots & \ddots & \ddots & \rho_1 \\ 0 & \cdots & & 0 & \rho_2 & \rho_1 & 1 \end{bmatrix} \begin{bmatrix} \phi_1 \\ \phi_2 \\ \phi_3 \\ \vdots \\ \phi_k \end{bmatrix} = \begin{bmatrix} \rho_1 \\ \rho_2 \\ 0 \\ \vdots \\ 0 \end{bmatrix}.$$

Решение этой задачи достаточно громоздко и оставляется читателю.

## Задача 11.21

Покажите, что остатки при оценивании методом наименьших квадратов уравнений

$$y_t = \alpha y_{t-1} + \beta x_t + \varepsilon_t,$$
$$\Delta y_t = \gamma y_{t-1} + \beta x_t + \varepsilon_t,$$

совпадают. (Здесь $\Delta y_t = y_t - y_{t-1}$.)

## Решение

Пусть $\widehat{\boldsymbol{y}}$ — вектор прогнозных значений для первого уравнения, т. е. ортогональная проекция вектора $\boldsymbol{y}$ на векторы $\boldsymbol{y}_{-1}$ и $\boldsymbol{x}$ (вектор $\boldsymbol{y}_{-1}$ состоит из компонент $y_{t-1}$). Тогда проекция вектора $\boldsymbol{y} - \boldsymbol{y}_{-1}$ на векторы $\boldsymbol{y}_{-1}$ и $\boldsymbol{x}$ равна, очевидно, $\widehat{\boldsymbol{y}} - \boldsymbol{y}_{-1}$. Получаем, что вектор остатков во второй регрессии равен

$$\boldsymbol{e}_2 = (\boldsymbol{y} - \boldsymbol{y}_{-1}) - (\widehat{\boldsymbol{y}} - \boldsymbol{y}_{-1}) = \boldsymbol{y} - \widehat{\boldsymbol{y}} = \boldsymbol{e}_1,$$

что и требовалось показать.

## Задача 11.22

Пусть $y_t = 1 + 0.4y_{t-1} + 0.3y_{t-2} + u_t$ — AR(2) процесс, где $u_t$ — независимые $N(0,1)$ случайные величины. Вычислите прогнозные значения

$$\mathrm{E}(y_t \mid y_{t-1}), \quad \mathrm{E}(y_t \mid y_{t-1}, y_{t-2}), \quad \mathrm{E}(y_t \mid y_{t-1}, y_{t-2}, y_{t-3}).$$

## Решение

*Решение 1.* Найдем среднее значение $\mu = \mathrm{E}(y_t)$, для этого возьмем математическое ожидание от левой и правой частей уравнения процесса. Получим:

$$\begin{aligned}\mathrm{E}(y_t) &= 1 + 0.4\mathrm{E}(y_{t-1}) + 0.3\mathrm{E}(y_{t-2}),\\ \mu &= 1 + 0.4\mu + 0.3\mu,\\ \mu &= 10/3 = 0.333.\end{aligned}$$

Получаем AR(2) процесс с нулевым средним $z_t = y_t - \mu$, соответственно, удовлетворяющий уравнению

$$z_t = 0.4z_{t-1} + 0.3z_{t-2} + u_t.$$

Вычислим далее несколько первых значений ACF.

$$\gamma_k = \mathrm{Cov}(z_{t-k}, z_t) = \mathrm{Cov}(z_{t-k}, 0.4z_{t-1} + 0.3z_{t-2} + u_t) = 0.4\gamma_{k-1} + 0.4\gamma_{k-2}.$$

Принимая во внимание то, что $\gamma_{-1} = \gamma_1$ получаем:

$$\begin{aligned}\gamma_1 &= 0.4\gamma_0 + 0.3\gamma_1,\\ \gamma_2 &= 0.4\gamma_1 + 0.3\gamma_0,\\ \gamma_3 &= 0.4\gamma_2 + 0.3\gamma_1,\end{aligned}$$

отсюда первые значения ACF соответственно равны

$$\begin{aligned}\rho_1 &= \gamma_1/\gamma_0 = 4/7 &&\approx 0.571,\\ \rho_2 &= \gamma_2/\gamma_0 = 37/70 &&\approx 0.529,\\ \rho_3 &= \gamma_3/\gamma_0 = 67/175 &&\approx 0.383.\end{aligned}$$

Для вычисления наилучшего линейного прогноза применим известную формулу (здесь $\mathrm{E}(z) = 0$, $\mathrm{E}(\boldsymbol{x}) = \mathbf{0}$):

$$\mathrm{E}(z \mid \boldsymbol{x}) = \mathrm{Cov}(z, \boldsymbol{x})\mathrm{V}(\boldsymbol{x})^{-1}\boldsymbol{x}.$$

Получаем:

$$\begin{aligned}\mathrm{E}(z_t \mid z_{t-1}) &= \begin{bmatrix}\gamma_1\end{bmatrix}\begin{bmatrix}\gamma_0\end{bmatrix}^{-1} z_{t-1} = \frac{4}{7}z_{t-1},\\ \mathrm{E}(y_t \mid y_{t-1}) &= \mathrm{E}(\mu + z_t \mid z_{t-1}) = \frac{10}{3} + \frac{4}{7}(z_{t-1}) = \frac{10}{3} + \frac{4}{7}\left(-\frac{10}{3} + y_{t-1}\right)\\ &= \frac{10}{7} + \frac{4}{7}y_{t-1} \approx 1.43 + 0.57y_{t-1}.\end{aligned}$$

Аналогично,

$$\mathrm{E}(z_t \mid z_{t-1}, z_{t-2}) = \begin{bmatrix}\gamma_1 & \gamma_2\end{bmatrix} \begin{bmatrix}\gamma_0 & \gamma_1 \\ \gamma_1 & \gamma_0\end{bmatrix}^{-1} \begin{bmatrix}z_{t-1} \\ z_{t-2}\end{bmatrix} = \begin{bmatrix}\rho_1 & \rho_2\end{bmatrix} \begin{bmatrix}1 & \rho_1 \\ \rho_1 & 1\end{bmatrix}^{-1} \begin{bmatrix}z_{t-1} \\ z_{t-2}\end{bmatrix} =$$

$$= \frac{1}{1-\rho_1^2} \begin{bmatrix}\rho_1 - \rho_1\rho_2 & -\rho_1^2 + \rho_2\end{bmatrix} \begin{bmatrix}z_{t-1} \\ z_{t-2}\end{bmatrix} \approx$$

$$\approx \frac{1}{0.673} \begin{bmatrix}0.269 & 0.855\end{bmatrix} \begin{bmatrix}z_{t-1} \\ z_{t-2}\end{bmatrix} = 0.40 z_{t-1} + 0.30 z_{t-2},$$

$$\mathrm{E}(y_t \mid y_{t-1}, y_{t-2}) = 3.33 + 0.40(-3.33 + y_{t-1}) + 0.30(-3.33 + y_{t-2}) = 1 + 0.4 y_{t-1} + 0.3 y_{t-2}.$$

(Конечно, этот результат очевиден из исходного уравнения.) Продолжая вычисления, получаем:

$$\mathrm{E}(z_t \mid z_{t-1}, z_{t-2}, z_{t-3}) = \begin{bmatrix}\gamma_1 & \gamma_2 & \gamma_3\end{bmatrix} \begin{bmatrix}\gamma_0 & \gamma_1 & \gamma_2 \\ \gamma_1 & \gamma_0 & \gamma_1 \\ \gamma_2 & \gamma_1 & \gamma_0\end{bmatrix}^{-1} \begin{bmatrix}z_{t-1} \\ z_{t-2} \\ z_{t-3}\end{bmatrix} =$$

$$= \begin{bmatrix}\rho_1 & \rho_2 & \rho_2\end{bmatrix} \begin{bmatrix}1 & \rho_1 & \rho_2 \\ \rho_1 & 1 & \rho_1 \\ \rho_2 & \rho_1 & 1\end{bmatrix}^{-1} \begin{bmatrix}z_{t-1} \\ z_{t-2} \\ z_{t-3}\end{bmatrix} =$$

$$= \begin{bmatrix}0.4 & 0.3 & 0\end{bmatrix} \begin{bmatrix}z_{t-1} \\ z_{t-2} \\ z_{t-3}\end{bmatrix} = 0.40 z_{t-1} + 0.30 z_{t-2},$$

$$\mathrm{E}(y_t \mid y_{t-1}, y_{t-2}, y_{t-3}) = 1 + 0.4 y_{t-1} + 0.3 y_{t-2}.$$

*Решение 2.* Наилучший линейный прогноз является ортогональной проекцией в смысле скалярного произведения в пространстве случайных величин с нулевым средним, задаваемого ковариацией. Величина $u_t$ ортогональна к величинам $z_{t-k}$, $k = 1, 2, \ldots$ . Сразу получаем, что ортогональная проекция $z_t = 0.4 z_{t-1} + 0.3 z_{t-2} + u_t$ на пространства, порожденные наборами

$$\{z_{t-1}, z_{t-2}\} \quad \text{и} \quad \{z_{t-1}, z_{t-2}, z_{t-3}\},$$

равна $0.4 z_{t-1} + 0.3 z_{t-2}$.

Из линейной алгебры известно, что проекция вектора $b$ на вектор $a$ равна

$$\frac{(b, a)}{(a, a)} a.$$

Соответственно, проекция $z_t$ на $z_{t-1}$ равна

$$0.4 z_{t-1} + 0.3 \frac{\mathrm{Cov}(z_{t-2}, z_{t-1})}{\mathrm{V}(z_{t-1})} z_{t-1} = (0.4 + 0.3\rho_1) z_{t-1} = \frac{4}{7} z_{t-1}.$$

Отсюда, как и в первом решении, получаются проекции в терминах $y_t$.

## Задача 11.23

Дана линейная модель

$$y_t = \beta_1 x_t + \beta_2 y_{t-1} + u_t, \quad u_t = \rho u_{t-1} + \varepsilon_t,$$

где $0 < \rho < 1$ и $\varepsilon$ — гауссовский белый шум. Проводятся две регрессии: для исходных величин и их разностей, т. е.

$$y_t = \beta_1 x_t + \beta_2 y_{t-1} + u_t, \tag{$*$}$$

$$\Delta y_t = \beta_1 \Delta x_t + \beta_2 \Delta y_{t-1} + v_t, \tag{$**$}$$

где $v_t = (\rho - 1)u_{t-1} + \varepsilon_t$.

а) Покажите, что в обеих регрессиях $(*)$ и $(**)$ МНК-оценки вектора $\beta = [\beta_1\ \beta_2]'$ будут смещенными и несостоятельными.

б) Покажите, что смещение в регрессии $(*)$ не снижается до нуля, когда $\rho \to 1$.

в) Предложите оценку вектора $\beta$ с помощью инструментальных переменных и покажите, что она состоятельна.

## Решение

а) Будем считать, что процессы $x$ и $y$ являются стационарными, и предположим для простоты, что $\mathrm{E}(y_t) = \mathrm{E}(x_t) = 0$. Кроме того, естественно предполагать, что переменная $x$ является экзогенной, в частности, не коррелированной с шумом $u$. Будем использовать следующие обозначения:

$$\sigma_{xx} = \mathrm{Cov}(x_t, x_t), \quad \sigma_{yy} = \mathrm{Cov}(y_t, y_t), \quad \sigma_{xy} = \mathrm{Cov}(x_t, y_{t-1}),$$
$$\sigma_{yu} = \mathrm{Cov}(y_t, u_t), \quad \sigma_{uu} = \mathrm{Cov}(u_t, u_t), \quad \sigma_{\varepsilon\varepsilon} = \mathrm{Cov}(\varepsilon_t, \varepsilon_t).$$

Имеем

$$\sigma_{uu} = \frac{\sigma_{\varepsilon\varepsilon}}{1 - \rho^2}.$$

Проведем некоторые предварительные вычисления.

$$\begin{aligned}
\mathrm{Cov}(x_t, y_t) &= \mathrm{Cov}(x_t, \beta_1 x_t + \beta_2 y_{t-1} + u_t) = \beta_1 \sigma_{xx} + \beta_2 \sigma_{xy}, \\
\mathrm{Cov}(y_{t-1}, y_t) &= \mathrm{Cov}(y_{t-1}, \beta_1 x_t + \beta_2 y_{t-1} + u_t) \\
&= \beta_1 \sigma_{xy} + \beta_2 \sigma_{yy} + \mathrm{Cov}(y_{t-1}, \rho u_{t-1} + \varepsilon_t), \\
&= \beta_1 \sigma_{xy} + \beta_2 \sigma_{yy} + \rho \sigma_{yu}
\end{aligned}$$

(мы воспользовались тем, что $\mathrm{Cov}(x_t, u_t) = 0$ и $\mathrm{Cov}(y_{t-1}, \varepsilon_t) = 0$).

МНК-оценка коэффициентов в первом уравнении равна

$$\widehat{\beta} = \begin{bmatrix} \widehat{\beta}_1 \\ \widehat{\beta}_2 \end{bmatrix} = \begin{bmatrix} \sum x_t^2 & \sum x_t y_{t-1} \\ \sum x_t y_{t-1} & \sum y_{t-1}^2 \end{bmatrix}^{-1} \begin{bmatrix} \sum x_t y_t \\ \sum y_{t-1} y_t \end{bmatrix}$$

$$= \frac{1}{\sum x_t^2 \sum y_{t-1}^2 - (\sum x_t y_{t-1})^2} \begin{bmatrix} \sum y_{t-1}^2 & -\sum x_t y_{t-1} \\ -\sum x_t y_{t-1} & \sum x_t^2 \end{bmatrix} \begin{bmatrix} \sum x_t y_t \\ \sum y_{t-1} y_t \end{bmatrix},$$

$$\widehat{\beta}_1 = \frac{\sum y_{t-1}^2 \sum x_t y_t - \sum x_t y_{t-1} \sum y_{t-1} y_t}{\sum x_t^2 \sum y_{t-1}^2 - (\sum x_t y_{t-1})^2},$$

$$\widehat{\beta}_2 = \frac{\sum x_t^2 \sum y_{t-1} y_t - \sum x_t y_{t-1} \sum x_t y_t}{\sum x_t^2 \sum y_{t-1}^2 - (\sum x_t y_{t-1})^2}.$$

Разделив числитель и знаменатель на $n^2$ и переходя к пределу, получаем:

$$p\lim \widehat{\beta}_1 = \frac{\sigma_{yy}(\beta_1 \sigma_{xx} + \beta_2 \sigma_{xy}) - \sigma_{xy}(\beta_1 \sigma_{xy} + \beta_2 \sigma_{yy} + \rho \sigma_{yu})}{\sigma_{xx}\sigma_{yy} - \sigma_{xy}^2}$$

$$= \beta_1 - \rho \frac{\sigma_{xy}\sigma_{yu}}{\sigma_{xx}\sigma_{yy} - \sigma_{xy}^2},$$

$$p\lim \widehat{\beta}_2 = \frac{-\sigma_{xy}(\beta_1 \sigma_{xx} + \beta_2 \sigma_{xy}) + \sigma_{xx}(\beta_1 \sigma_{xy} + \beta_2 \sigma_{yy} + \rho \sigma_{yu})}{\sigma_{xx}\sigma_{yy} - \sigma_{xy}^2}$$

$$= \beta_2 + \rho \frac{\sigma_{xx}\sigma_{yu}}{\sigma_{xx}\sigma_{yy} - \sigma_{xy}^2}.$$

При этом

$$\sigma_{yu} = \mathrm{Cov}(u_t, y_t) = \mathrm{E}(\beta_2 y_{t-1} u_t) + \sigma_{uu} = \mathrm{E}(\beta_2 y_{t-1}(\rho u_{t-1} + \varepsilon_t)) + \sigma_{uu}$$
$$= \beta_2 \rho \sigma_{yu} + \sigma_{uu},$$

откуда

$$\sigma_{yu} = \frac{\sigma_{uu}}{1 - \rho\beta_2}.$$

Асимптотическое значение смещений оценок в общем случае отличается от нуля и, как и следовало ожидать, исчезает при $\rho \to 0$. Поэтому оценки не состоятельные и смещенные (по крайней мере, начиная с некоторого $n$).

Рассмотрим теперь второе уравнение. Проводя такие же вычисления, как и раньше, получаем следующие равенства (в обозначениях, аналогичных приведенным выше):

$$p\lim \widehat{\beta}_1 = \beta_1 - (\rho - 1)\frac{\sigma_{\Delta x \Delta y}\sigma_{\Delta y u}}{\sigma_{\Delta x \Delta x}\sigma_{\Delta y \Delta y} - \sigma_{\Delta x \Delta y}^2},$$

$$p\lim \widehat{\beta}_2 = \beta_2 + (\rho - 1)\frac{\sigma_{\Delta x \Delta x}\sigma_{\Delta y u}}{\sigma_{\Delta x \Delta x}\sigma_{\Delta y \Delta y} - \sigma_{\Delta x \Delta y}^2}.$$

б) Наличие асимптотического смещения оценок в регрессии (*) при $\rho \to 1$ непосредственно вытекает из соотношений, полученных в п. а). В то же время асимптотическое смещение в регрессии (**) при $\rho \to 1$ исчезает.

в) Воспользуемся методом инструментальных переменных IV (гл. 8.1). В качестве инструментов возьмем $x_t$ и $x_{t-1}$. Таким образом, в данном случае матрица исходных регрессоров $\boldsymbol{X}$ и матрица инструментальных переменных $\boldsymbol{Z}$ есть, соответственно,

$$\boldsymbol{X} = \begin{bmatrix} x_1 & y_0 \\ \vdots & \vdots \\ x_n & y_{n-1} \end{bmatrix}, \quad \boldsymbol{Z} = \begin{bmatrix} x_1 & x_0 \\ \vdots & \vdots \\ x_n & x_{n-1} \end{bmatrix}.$$

Как известно, для состоятельности IV-оценок достаточно, чтобы последовательность матриц $(1/n)\boldsymbol{Z}'\boldsymbol{X}$ сходилась к невырожденной матрице. Имеем

$$\frac{1}{n}\boldsymbol{Z}'\boldsymbol{X} = \frac{1}{n}\begin{bmatrix} \sum x_t^2 & \sum x_t y_{t-1} \\ \sum x_t x_{t-1} & \sum x_{t-1} y_{t-1} \end{bmatrix} \to \text{(при } n \to \infty)$$

$$\boldsymbol{A} = \begin{bmatrix} \mathrm{Cov}(x_t, x_t) & \mathrm{Cov}(x_t, y_{t-1}) \\ \mathrm{Cov}(x_t, x_{t-1}) & \mathrm{Cov}(x_{t-1}, y_{t-1}) \end{bmatrix} = \begin{bmatrix} \sigma_{xx} & \sigma_{xy} \\ \sigma_{xx_{-1}} & \beta_1\sigma_{xx} + \beta_2\sigma_{xy} \end{bmatrix},$$

где по определению $\sigma_{xx_{-1}} = \mathrm{Cov}(x_t, x_{t-1})$. Имеем

$$\det(\boldsymbol{A}) = \sigma_{xx}(\beta_1\sigma_{xx} + \beta_2\sigma_{xy}) - \sigma_{xx_{-1}}\sigma_{xy},$$

что в общем случае при $\beta_1 \neq 0$ не равно нулю. Если же $\beta_1 = 0$, то, очевидно, $\sigma_{xy} = 0$ и $\det(\boldsymbol{A}) = 0$, и переменные $\boldsymbol{Z}$ нельзя использовать в качестве инструментальных переменных. Впрочем, это ясно и из содержательных соображений: при $\beta_1 = 0$ переменные $x$ и $y$ между собой никак не связаны, и переменные $\boldsymbol{Z}$ не могут быть инструментами для $y$. Итак, получаем, что IV-оценки параметров $\beta_1$ и $\beta_2$ при $\beta_1 \neq 0$ являются состоятельными.

# Глава 12

# Дискретные зависимые переменные и цензурированные выборки

## Задача 12.1

Покажите, что если среди регрессоров линейной модели вероятности или *logit*-модели есть константа, то среднее значение прогнозных вероятностей равно доле единиц во всей выборке зависимой переменной.

## Решение

1) Рассмотрим линейную модель $y_t = \boldsymbol{x}_t' \boldsymbol{\beta} + \varepsilon_t$, или

$$\mathrm{E}(y_t) = \mathrm{P}(y_t = 1) = \boldsymbol{x}_t' \boldsymbol{\beta}.$$

Вектор прогнозных значений вероятностей $\widehat{\boldsymbol{p}}$ равен:

$$\widehat{\boldsymbol{p}} = \boldsymbol{X}(\boldsymbol{X}'\boldsymbol{X})^{-1}\boldsymbol{X}'\boldsymbol{y}.$$

Вычитая из среднего значения прогнозных вероятностей долю единиц во всей выборке $y_t$, получаем:

$$n\left(\frac{1}{n}\sum \widehat{p}_t - \frac{1}{n}\sum y_t\right) = -\boldsymbol{\imath}'\left(\boldsymbol{I} - \boldsymbol{X}(\boldsymbol{X}'\boldsymbol{X})^{-1}\boldsymbol{X}'\right)\boldsymbol{y} = -\boldsymbol{\imath}'\boldsymbol{M}\boldsymbol{y}.$$

Поскольку $\boldsymbol{M}$ является проектором на подпространство, ортогональное подпространству, порожденному столбцами матрицы $\boldsymbol{X}$, которое, по условию, содержит вектор $\boldsymbol{\imath}$, то $\boldsymbol{\imath}'\boldsymbol{M} = \boldsymbol{0}$, откуда следует первая часть задачи.

2) Рассмотрим теперь *logit*-модель

$$\mathrm{P}(y_t = 1) = \Lambda(\boldsymbol{x}_t'\boldsymbol{\beta}), \qquad \Lambda(u) = \frac{e^u}{1+e^u}.$$

Логарифмическая функция правдоподобия равна (ср. (12.9))

$$\ln L = \sum_t \left(y_t \ln \Lambda(\boldsymbol{x}_t'\boldsymbol{\beta}) + (1-y_t)\ln(1-\Lambda(\boldsymbol{x}_t'\boldsymbol{\beta}))\right).$$

Дифференцируя равенство по вектору $\boldsymbol{\beta}$ (приложение ЛА, п. 19), получаем векторное уравнение правдоподобия

$$\frac{\partial l}{\partial \boldsymbol{\beta}} = \sum_t \left( y_t \frac{\Lambda'(\boldsymbol{x}_t'\boldsymbol{\beta})}{\Lambda(\boldsymbol{x}_t'\boldsymbol{\beta})} - (1-y_t)\frac{\Lambda'(\boldsymbol{x}_t'\boldsymbol{\beta})}{1-\Lambda(\boldsymbol{x}_t'\boldsymbol{\beta})} \right) \boldsymbol{x}_t = \mathbf{0}.$$

Применяя тождество $\Lambda'(u) = \Lambda(u)(1-\Lambda(u))$, получаем

$$\begin{aligned}\frac{\partial l}{\partial \boldsymbol{\beta}} &= \sum_t \left(y_t(1-\Lambda(\boldsymbol{x}_t'\boldsymbol{\beta})) - (1-y_t)\Lambda(\boldsymbol{x}_t'\boldsymbol{\beta})\right)\boldsymbol{x}_t \\ &= \sum_t (y_t - \Lambda(\boldsymbol{x}_t'\boldsymbol{\beta}))\,\boldsymbol{x}_t = \sum_t (y_t - \widehat{p}_t)\,\boldsymbol{x}_t = \mathbf{0}.\end{aligned}$$

Отсюда следует, что

$$(\boldsymbol{y} - \widehat{\boldsymbol{p}})'\boldsymbol{X} = \mathbf{0},$$

и так как по условию константа включена в число регрессоров, то

$$0 = (\boldsymbol{y} - \widehat{\boldsymbol{p}})'\boldsymbol{\imath} = n\left(\frac{1}{n}\sum y_t - \frac{1}{n}\sum p_t\right),$$

что и требовалось показать.

## Задача 12.2

Докажите равенства (12.20)

$$\mathrm{E}(Z \mid Z > a) = \mu + \sigma\lambda(c), \qquad \mathrm{V}(Z \mid Z > a) = \sigma^2(1-\gamma(c)),$$

где

$$c = (a-\mu)/\sigma, \quad \lambda(c) = \frac{\phi(c)}{1-\Phi(c)}, \quad \gamma(c) = \lambda(c)(\lambda(c) - c),$$

а случайная величина $Z$ имеет нормальное распределение $N(\mu, \sigma^2)$. $\phi(x)$ и $\Phi(x)$ — плотность и функция распределения стандартного нормального распределения.

(*Указание.* Воспользуйтесь равенством $\phi'(u) = -u\phi(u)$.)

## Решение

Представим $Z$ в виде $Z = \mu + \sigma U$, где $U \sim N(0,1)$. Обозначим через $f(z)$ и $F(z)$ плотность и функция распределения случайной величины $Z$. Легко видеть, что

$$F(z) = \mathrm{P}(Z \leqslant z) = \mathrm{P}(\mu + \sigma U \leqslant z) = \mathrm{P}\left(U \leqslant \frac{z-\mu}{\sigma}\right) = \Phi\left(\frac{z-\mu}{\sigma}\right),$$

$$f(z) = F(z)' = \frac{1}{\sigma}\phi\left(\frac{z-\mu}{\sigma}\right).$$

Условная плотность может быть найдена по формуле (12.19):

$$f(z \mid Z > a) = \frac{f(z)}{1 - F(a)} = \frac{1}{\sigma}\frac{\phi\left(\frac{z-\mu}{\sigma}\right)}{1-\Phi(c)}, \quad z \geqslant a.$$

Очевидно, $\int_a^\infty f(z \mid Z > a)dz = 1$.

Вычислим условное математическое ожидание:

$$\mathrm{E}(Z \mid Z > a) = \int_a^\infty z\, f(z \mid Z > a)dz =$$

$$\text{(замена переменной: } z = \mu + \sigma t)$$

$$= \frac{1}{1-\Phi(c)}\int_c^\infty (\mu + \sigma t)\phi(t)dt$$

$$= \frac{1}{1-\Phi(c)}\left(\mu\int_c^\infty \phi(t)dt + \sigma\int_c^\infty t\phi(t)dt\right)$$

$$= \mu + \sigma\frac{1}{1-\Phi(c)}\left(-\int_c^\infty \phi'(t)dt\right) = \mu + \sigma\frac{\phi(c)}{1-\Phi(c)}$$

$$= \mu + \sigma\lambda(c).$$

(Мы здесь использовали тождество $\phi'(u) = -u\phi(u)$.)

Вычислим теперь условное математическое ожидание $Z^2$:

$$\mathrm{E}(Z^2 \mid Z > a) = \int_a^\infty z^2\, f(z \mid Z > a)dz =$$

$$\text{(замена переменной: } z = \mu + \sigma t)$$

$$= \frac{1}{1-\Phi(c)}\int_c^\infty (\mu + \sigma t)^2\phi(t)dt$$

$$= \frac{1}{1-\Phi(c)}\left(\mu^2\int_c^\infty \phi(t)dt + 2\mu\sigma\int_c^\infty t\phi(t)dt + \sigma^2\int_c^\infty t^2\phi(t)dt\right)$$

$$= \mu^2 + 2\mu\sigma\lambda(c) + \frac{\sigma^2}{1-\Phi(c)}\left(-t\phi(t)\Big|_c^\infty + \int_c^\infty \phi(t)dt\right)$$

$$= \mu^2 + 2\mu\sigma\lambda(c) + c\sigma^2\lambda(c) + \sigma^2.$$

Теперь вычислим условную дисперсию:

$$\begin{aligned}
\mathrm{V}(Z \mid Z > a) &= \mathrm{E}(Z^2 \mid Z > a) - (\mathrm{E}(Z \mid Z > a))^2 \\
&= \mu^2 + 2\mu\sigma\lambda(c) + c\sigma^2\lambda(c) + \sigma^2 - (\mu + \sigma\lambda(c))^2 \\
&= \sigma^2(1 + c\lambda(c) - \lambda(c)^2) = \sigma^2\,(1 - \lambda(c)(\lambda(c) - c)).
\end{aligned}$$

## Задача 12.3

Докажите равенство (12.33)

$$\frac{\partial \mathrm{E}(y)}{\partial \boldsymbol{x}} = \Phi\left(\frac{\boldsymbol{x}'\boldsymbol{\beta}}{\sigma}\right)\boldsymbol{\beta}$$

для предельного эффекта объясняющих факторов в *tobit*-модели.

## Решение

Используя в вычислениях формулу (12.31) для $\mathrm{E}(y_t)$ и правила дифференцирования по векторному аргументу (опуская индекс $t$), получим:

$$\begin{aligned}
\frac{\partial \mathrm{E}(y)}{\partial \boldsymbol{x}} &= \frac{\partial}{\partial \boldsymbol{x}}\left[\Phi\left(\frac{\boldsymbol{x}'\boldsymbol{\beta}}{\sigma}\right)\boldsymbol{x}'\boldsymbol{\beta} + \sigma\phi\left(\frac{\boldsymbol{x}'\boldsymbol{\beta}}{\sigma}\right)\right] \\
&= \phi\left(\frac{\boldsymbol{x}'\boldsymbol{\beta}}{\sigma}\right)\frac{1}{\sigma}\boldsymbol{\beta}\,\boldsymbol{x}'\boldsymbol{\beta} + \Phi\left(\frac{\boldsymbol{x}'\boldsymbol{\beta}}{\sigma}\right)\boldsymbol{\beta} + \sigma\left(-\frac{\boldsymbol{x}'\boldsymbol{\beta}}{\sigma}\right)\phi\left(\frac{\boldsymbol{x}'\boldsymbol{\beta}}{\sigma}\right)\frac{1}{\sigma}\boldsymbol{\beta} \\
&= \Phi\left(\frac{\boldsymbol{x}'\boldsymbol{\beta}}{\sigma}\right)\boldsymbol{\beta}.
\end{aligned}$$

## Задача 12.4

Проверьте справедливость представления (12.35):

$$\frac{\partial \mathrm{E}(y)}{\partial \boldsymbol{x}} = \mathrm{P}(y^* > 0)\frac{\partial \mathrm{E}(y^* \mid y^* > 0)}{\partial \boldsymbol{x}} + \frac{\partial \mathrm{P}(y^* > 0)}{\partial \boldsymbol{x}}\mathrm{E}(y^* \mid y^* > 0)$$

для предельного эффекта объясняющих факторов в *tobit*-модели.

## Решение

Ненаблюдаемая величина $y^*$ удовлетворяет регрессионному уравнению

$$y_t^* = \boldsymbol{x}_t'\boldsymbol{\beta} + \varepsilon_t.$$

Введем обозначения:

$$z = \frac{\boldsymbol{x}'\boldsymbol{\beta}}{\sigma}, \qquad r = \frac{\phi(z)}{\Phi(z)}$$

и вычислим компоненты требуемого разложения:

$$\begin{aligned}
\mathrm{P}(y^*>0) &= \mathrm{P}(\varepsilon > -\boldsymbol{x}'\boldsymbol{\beta}) = \mathrm{P}\left(\frac{\varepsilon}{\sigma} > -\frac{\boldsymbol{x}'\boldsymbol{\beta}}{\sigma}\right) = 1-\Phi\left(-\frac{\boldsymbol{x}'\boldsymbol{\beta}}{\sigma}\right) \\
&= \Phi\left(\frac{\boldsymbol{x}'\boldsymbol{\beta}}{\sigma}\right) = \Phi(z), \\
\frac{\partial \mathrm{P}(y^*>0)}{\partial \boldsymbol{x}} &= \phi\left(\frac{\boldsymbol{x}'\boldsymbol{\beta}}{\sigma}\right)\frac{\boldsymbol{\beta}}{\sigma} = \phi(z)\frac{\boldsymbol{\beta}}{\sigma}.
\end{aligned}$$

Применяя формулу для условного математического ожидания (см. (12.20), а также см. задачу 12.2), получаем:

$$\begin{aligned}
\mathrm{E}(y^* \mid y^*>0) &= \boldsymbol{x}'\boldsymbol{\beta} + \sigma\frac{\phi\left(-\frac{\boldsymbol{x}'\boldsymbol{\beta}}{\sigma}\right)}{1-\Phi\left(-\frac{\boldsymbol{x}'\boldsymbol{\beta}}{\sigma}\right)} = \boldsymbol{x}'\boldsymbol{\beta} + \sigma\frac{\phi\left(\frac{\boldsymbol{x}'\boldsymbol{\beta}}{\sigma}\right)}{\Phi\left(\frac{\boldsymbol{x}'\boldsymbol{\beta}}{\sigma}\right)} = \sigma(z+r), \\
\frac{\partial \mathrm{E}(y^* \mid y^*>0)}{\partial \boldsymbol{x}} &= \boldsymbol{\beta} - \sigma\frac{\frac{\boldsymbol{x}'\boldsymbol{\beta}}{\sigma}\phi\left(\frac{\boldsymbol{x}'\boldsymbol{\beta}}{\sigma}\right)}{\Phi\left(\frac{\boldsymbol{x}'\boldsymbol{\beta}}{\sigma}\right)}\frac{\boldsymbol{\beta}}{\sigma} - \sigma\phi\left(\frac{\boldsymbol{x}'\boldsymbol{\beta}}{\sigma}\right)\frac{\phi\left(\frac{\boldsymbol{x}'\boldsymbol{\beta}}{\sigma}\right)}{\Phi\left(\frac{\boldsymbol{x}'\boldsymbol{\beta}}{\sigma}\right)^2}\frac{\boldsymbol{\beta}}{\sigma} \\
&= \boldsymbol{\beta} - zr\boldsymbol{\beta} - r^2\boldsymbol{\beta} = (1-zr-r^2)\boldsymbol{\beta}.
\end{aligned}$$

Собирая все вычисленные компоненты вместе, получаем:

$$\begin{aligned}
\frac{\partial \mathrm{E}(y)}{\partial \boldsymbol{x}} &= \mathrm{P}(y^*>0)\frac{\partial \mathrm{E}(y^* \mid y^*>0)}{\partial \boldsymbol{x}} + \frac{\partial \mathrm{P}(y^*>0)}{\partial \boldsymbol{x}}\mathrm{E}(y^* \mid y^*>0) \\
&= \left(\Phi(z)(1-zr-r^2) + \phi(z)(z+r)\right)\boldsymbol{\beta} = \Phi\left(\frac{\boldsymbol{x}'\boldsymbol{\beta}}{\sigma}\right)\boldsymbol{\beta}.
\end{aligned}$$

Последнее выражение совпадает с формулой (12.33) (см. задачу 12.3) для предельного эффекта объясняющих факторов в *tobit*-модели.

## Задача 12.5

Рассмотрим модель бинарного выбора $\mathrm{P}(y_t=1) = F(\alpha+\beta d_t)$, где $d$ — фиктивная переменная (принимающая значения 0 или 1). Ниже представлены результаты 100 наблюдений:

| | $y=0$ | $y=1$ |
|---|---|---|
| $d=0$ | 20 | 32 |
| $d=1$ | 36 | 12 |

а) Оцените параметры $\alpha$, $\beta$, используя *logit*-модель. Проверьте гипотезу $\mathrm{H}_0$: $\beta=0$.

б) Повторите п. а) для *probit*-модели. Изменяются ли ваши выводы?

## Решение

Рассмотрим модель $\mathrm{P}(y_t = 1) = F(\alpha + \beta d_t)$. Функция правдоподобия равна (см. (12.8)):

$$L = F(\alpha)^{32} F(\alpha+\beta)^{12} (1 - F(\alpha))^{20} (1 - F(\alpha+\beta))^{36}.$$

Обозначим $\gamma = \alpha + \beta$. Логарифмическая функция правдоподобия равна:

$$\ln L = 32 \ln F(\alpha) + 12 \ln F(\gamma) + 20 \ln(1 - F(\alpha)) + 36 \ln(1 - F(\gamma)).$$

Оценка максимального правдоподобия является решением системы уравнений:

$$\frac{\partial \ln L}{\partial \alpha} = 32 \frac{f(\alpha)}{F(\alpha)} - 20 \frac{f(\alpha)}{1 - F(\alpha)} = 0,$$
$$\frac{\partial \ln L}{\partial \gamma} = 12 \frac{f(\gamma)}{F(\gamma)} - 36 \frac{f(\gamma)}{1 - F(\gamma)} = 0.$$

Решением является

$$F(\alpha) = 8/13, \qquad F(\gamma) = 1/4.$$

Для *probit*-модели $F(z) = \Phi(z)$ получаем:

$$\widehat{\alpha} = 0.2934, \qquad \widehat{\beta} = -0.9679.$$

Для *logit*-модели $F(z) = \Lambda(z)$ получаем:

$$\widehat{\alpha} = 0.4700, \qquad \widehat{\beta} = -1.5686.$$

Проверим теперь гипотезу $\beta = 0$ при помощи теста отношения правдоподобия, теста Вальда и теста множителей Лагранжа.

Сначала применим тест отношения правдоподобия (10.54):

$$\mathrm{LR} = -2(\ln \widetilde{L} - \ln \widehat{L}).$$

Значение функции правдоподобия для регрессии без ограничения равно

$$\ln \widehat{L} = 32 \ln(8/13) + 12 \ln(1/4) + 20 \ln(5/13) + 36 \ln(3/4) = -61.6386.$$

Оценим модель с ограничением $\beta = 0$ или $\gamma = \alpha$:

$$L = F(\alpha)^{44} (1 - F(\alpha))^{56},$$
$$\ln L = 44 \ln F(\alpha) + 56 \ln(1 - F(\alpha))$$
$$\frac{\partial \ln L}{\partial \alpha} = 44 \frac{f(\alpha)}{F(\alpha)} - 56 \frac{f(\alpha)}{1 - F(\alpha)} = 0.$$

Откуда получаем $F(\alpha) = 11/25$ и

$$\ln \widetilde{L} = 44\ln(11/25) + 56\ln(14/25) = -68.5930.$$

Получаем $\mathrm{LR} = 13.9088$, что превышает критическое значение $\chi^2_{0.05}(1) = 3.84$, т. е. гипотеза $\beta = 0$ уверенно отвергается. Как видно из вычислений, этот результат не зависит от вида функции распределения $F(z)$ и, следовательно, одинаков для *logit*- и *probit*-моделей.

Перед тем, как рассматривать тесты Вальда и множителей Лагранжа, проделаем предварительные действия.

Найдем информационную матрицу Фишера для *logit*- и *probit*-моделей. Логарифмическая функция правдоподобия (для одного наблюдения) равна

$$\widetilde{l} = y\ln F(\alpha + \beta d) + (1 - y)\ln(1 - F(\alpha + \beta d)).$$

Продифференцируем ее по $\alpha$ и $\beta$. Получим:

$$\begin{aligned}
\frac{\partial \widetilde{l}}{\partial \alpha} &= y\frac{F'(\alpha+\beta d)}{F(\alpha+\beta d)} - (1-y)\frac{F'(\alpha+\beta d)}{1-F(\alpha+\beta d)} \\
&= \frac{F'(\alpha+\beta d)}{F(\alpha+\beta d)(1-F(\alpha+\beta d))}(y - F(\alpha+\beta d)), \\
\frac{\partial \widetilde{l}}{\partial \beta} &= y\frac{F'(\alpha+\beta d)}{F(\alpha+\beta d)}d - (1-y)\frac{F'(\alpha+\beta d)}{1-F(\alpha+\beta d)}d \\
&= \frac{F'(\alpha+\beta d)}{F(\alpha+\beta d)(1-F(\alpha+\beta d))}(y - F(\alpha+\beta d))d.
\end{aligned}$$

Обозначим $\theta = \begin{bmatrix}\alpha & \beta\end{bmatrix}'$. Информационная матрица равна

$$\begin{aligned}
\boldsymbol{I}_1(\theta, d) &= \mathrm{E}\left(\frac{\partial \widetilde{l}}{\partial \theta}\cdot\frac{\partial \widetilde{l}}{\partial \theta'}\right) = \frac{(F'(\alpha+\beta d))^2}{F(\alpha+\beta d)(1-F(\alpha+\beta d))}\begin{bmatrix}1 & d \\ d & d^2\end{bmatrix}, \\
\boldsymbol{I}_n(\theta, \boldsymbol{d}) &= \sum_{t=1}^{n}\boldsymbol{I}_1(\theta, d_t) = \sum_{t=1}^{n}\frac{(F'(\alpha+\beta d_t))^2}{F(\alpha+\beta d_t)(1-F(\alpha+\beta d_t))}\begin{bmatrix}1 & d_t \\ d_t & d_t^2\end{bmatrix} \\
&= \#\{d_t = 0\}\frac{(F'(\alpha))^2}{F(\alpha)(1-F(\alpha))}\begin{bmatrix}1 & 0 \\ 0 & 0\end{bmatrix} \\
&\quad + \#\{d_t = 1\}\frac{(F'(\alpha+\beta))^2}{F(\alpha+\beta)(1-F(\alpha+\beta))}\begin{bmatrix}1 & 1 \\ 1 & 1\end{bmatrix},
\end{aligned}$$

где через $\#\{d_t = 0\}$ и $\#\{d_t = 1\}$ обозначено число наблюдений, для которых $d_t = 0$ и $d_t = 1$ соответственно.

Теперь рассмотрим тест Вальда. Тестируется гипотеза $\mathrm{H}_0\colon c(\alpha,\beta) = 0$, где $c(\alpha,\beta) = \beta$. Статистика Вальда имеет вид

$$\mathrm{W} = \frac{c^2(\widehat{\alpha}_{\mathrm{UR}}, \widehat{\beta}_{\mathrm{UR}})}{\dfrac{\partial c}{\partial[\alpha,\beta]}\boldsymbol{I}_n^{-1}(\widehat{\alpha}_{\mathrm{UR}}, \widehat{\beta}_{\mathrm{UR}}, \boldsymbol{d})\dfrac{\partial c}{\partial[\alpha,\beta]'}} = \frac{\widehat{\beta}^2_{\mathrm{UR}}}{[0,1]\boldsymbol{I}_n^{-1}(\widehat{\alpha}_{\mathrm{UR}}, \widehat{\beta}_{\mathrm{UR}}, \boldsymbol{d})[0,1]'}.$$

Для *logit*-модели $\widehat{\beta}_{\mathrm{UR}} = -1.5686$,

$$\boldsymbol{I}_n(\widehat{\alpha}_{\mathrm{UR}}, \widehat{\beta}_{\mathrm{UR}}, \boldsymbol{d}) = 52 \cdot \frac{8}{13} \cdot \frac{5}{13} \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} + 48 \cdot \frac{1}{4} \cdot \frac{3}{4} \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix} = \begin{bmatrix} 21.3 & 9 \\ 9 & 9 \end{bmatrix}.$$

Следовательно, статистика Вальда

$$\mathrm{W} = \frac{(-1.5686)^2}{\begin{bmatrix} 0 & 1 \end{bmatrix} \begin{bmatrix} 21.3 & 9 \\ 9 & 9 \end{bmatrix}^{-1} \begin{bmatrix} 0 \\ 1 \end{bmatrix}} = 11.92.$$

Сравнивая с критическим значением $\chi^2_{0.05}(1) = 3.84$, делаем вывод о том, что гипотеза $\beta = 0$ отвергается.

Для *probit*-модели $\widehat{\alpha}_{\mathrm{UR}} = 0.2934$, $\widehat{\beta}_{\mathrm{UR}} = -0.9679$, $F'(\widehat{\alpha}_{\mathrm{UR}}) = 0.3814$, $F'(\widehat{\alpha}_{\mathrm{UR}} + \widehat{\beta}_{\mathrm{UR}}) = 0.315$

$$\boldsymbol{I}_n(\widehat{\alpha}_{\mathrm{UR}}, \widehat{\beta}_{\mathrm{UR}}, \boldsymbol{d}) = 52 \cdot 0.3814^2 \frac{13 \cdot 13}{8 \cdot 5} \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} + 48 \cdot 0.315^2 \frac{4 \cdot 4}{1 \cdot 3} \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix} = \begin{bmatrix} 57.36 & 25.40 \\ 25.40 & 25.40 \end{bmatrix}.$$

Следовательно, статистика Вальда

$$\mathrm{W} = \frac{(-0.9679)^2}{\begin{bmatrix} 0 & 1 \end{bmatrix} \begin{bmatrix} 57.36 & 25.40 \\ 25.40 & 25.40 \end{bmatrix}^{-1} \begin{bmatrix} 0 \\ 1 \end{bmatrix}} = 13.599.$$

Сравнивая с критическим значением $\chi^2_{0.05}(1) = 3.84$, делаем вывод о том, что гипотеза $\beta = 0$ отвергается.

И, наконец, рассмотрим тест множителей Лагранжа. Для вычисления статистики LM-теста нам понадобятся значения производных логарифмической функции правдоподобия в точке $(\widehat{\alpha}_{\mathrm{R}}, \widehat{\beta}_{\mathrm{R}}) = (\widehat{\alpha}_{\mathrm{R}}, 0)$. Ясно, что поскольку $\widehat{\alpha}_{\mathrm{R}}$ — точка максимума функции правдоподобия при ограничении $\beta = 0$, то

$$\frac{\partial \ln L}{\partial \alpha}(\widehat{\alpha}_{\mathrm{R}}, 0) = 0.$$

Производная по $\beta$ равна

$$\frac{\partial \ln L}{\partial \beta}(\widehat{\alpha}_{\mathrm{R}}, 0) = 12 \frac{F'(\alpha_{\mathrm{R}})}{F(\alpha_{\mathrm{R}})} - 36 \frac{F'(\alpha_{\mathrm{R}})}{1 - F(\alpha_{\mathrm{R}})} \frac{F'(\alpha_{\mathrm{R}})}{F(\alpha_{\mathrm{R}})(1 - F(\alpha_{\mathrm{R}}))}(12 - 48F(\alpha_{\mathrm{R}})).$$

Для *logit*-модели $\widehat{\alpha}_{\mathrm{R}} = -0.2412$,

$$\frac{\partial \ln L}{\partial \beta}(\widehat{\alpha}_{\mathrm{R}}, 0) = 1 \cdot (12 - 48F(\alpha_{\mathrm{R}})) = 12 - 48 \cdot \frac{11}{25} = -9.12,$$

$$\boldsymbol{I}_n(\widehat{\alpha}_{\mathrm{R}}, \widehat{\beta}_{\mathrm{R}}, \boldsymbol{d}) = 52 \cdot \frac{11}{25} \cdot \frac{14}{25} \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} + 48 \cdot \frac{11}{25} \cdot \frac{14}{25} \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix} = \begin{bmatrix} 24.64 & 11.83 \\ 11.83 & 11.83 \end{bmatrix}.$$

Следовательно, статистика теста множителей Лагранжа равна

$$\mathrm{LM} = \begin{bmatrix} 0 & -9.12 \end{bmatrix} \begin{bmatrix} 24.64 & 11.83 \\ 11.83 & 11.83 \end{bmatrix}^{-1} \begin{bmatrix} 0 \\ -9.12 \end{bmatrix} = 13.524.$$

Сравнивая с критическим значением $\chi^2_{0.05}(1) = 3.84$, делаем вывод о том, что гипотеза $\beta = 0$ отвергается.

Для *probit*-модели $\widehat{\alpha}_{\mathrm{R}} = -0.1510$, $F'(\widehat{\alpha}_{\mathrm{R}}) = 0.3944$,

$$\frac{\partial \ln L}{\partial \beta}(\widehat{\alpha}_{\mathrm{R}}, 0) = \frac{0.3944 \cdot 25^2}{11 \cdot 14} \cdot (12 - 48F(\alpha_{\mathrm{R}})) = 1.60 \cdot (-9.12) = -14.592,$$

$$\begin{aligned} \boldsymbol{I}_n(\widehat{\alpha}_{\mathrm{R}}, \widehat{\beta}_{\mathrm{R}}, \boldsymbol{d}) &= 52 \cdot \frac{0.3944^2 \cdot 25^2}{11 \cdot 14} \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} + 48 \cdot \frac{0.3944^2 \cdot 25^2}{11 \cdot 14} \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix} \\ &= \begin{bmatrix} 63.17 & 30.31 \\ 30.31 & 30.31 \end{bmatrix}. \end{aligned}$$

Следовательно, статистика теста множителей Лагранжа равна

$$\mathrm{LM} = \begin{bmatrix} 0 & -14.592 \end{bmatrix} \begin{bmatrix} 63.17 & 30.31 \\ 30.31 & 30.31 \end{bmatrix}^{-1} \begin{bmatrix} 0 \\ -14.592 \end{bmatrix} = 13.524.$$

Сравнивая с критическим значением $\chi^2_{0.05}(1) = 3.84$, делаем вывод о том, что гипотеза $\beta = 0$ отвергается.

## Задача 12.6

Докажите вогнутость (по $\beta$) логарифмической функции правдоподобия для *logit*-модели.

## Решение

Логарифмическая функция правдоподобия для *logit*-модели имеет вид (см. (12.9)):

$$\ln L = \sum_t \left( y_t \ln \Lambda(\boldsymbol{x}_t' \boldsymbol{\beta}) + (1 - y_t) \ln(1 - \Lambda(\boldsymbol{x}_t' \boldsymbol{\beta})) \right).$$

Вычислим матрицу вторых производных $\ln L$. При этом мы будем использовать тождество $\Lambda(u)' = \Lambda(u)(1 - \Lambda(u))$, правила дифференцирования по векторному аргументу (приложение ЛА, п.19) и сокращенные обозначения $\Lambda_t = \Lambda(\boldsymbol{x}_t' \boldsymbol{\beta})$, $\Lambda_t' = \Lambda'(\boldsymbol{x}_i' \boldsymbol{\beta})$:

$$\frac{\partial \ln L}{\partial \boldsymbol{\beta}'} = \sum_t \left( y_t \frac{\Lambda_t'}{\Lambda_t} - (1 - y_t) \frac{\Lambda_t'}{1 - \Lambda_t} \right) \boldsymbol{x}_t' = \sum_t \left( y_t (1 - \Lambda_t) - (1 - y_t) \Lambda_t \right) \boldsymbol{x}_t',$$

$$\frac{\partial^2 \ln L}{\partial \boldsymbol{\beta}\, \partial \boldsymbol{\beta}'} = \sum_t \left( -y_t \Lambda_t' - (1 - y_t) \Lambda_t' \right) \boldsymbol{x}_t \boldsymbol{x}_t' = -\sum_t \Lambda_t (1 - \Lambda_t) \boldsymbol{x}_t \boldsymbol{x}_t'.$$

Последняя матрица является суммой неотрицательно определенных матриц $x_t x_t'$ с отрицательными коэффициентами, поэтому является неположительно определенной матрицей:

$$\frac{\partial^2 \ln L}{\partial\beta\,\partial\beta'} \leqslant \mathbf{0},$$

что и требовалось доказать.

## Задача 12.7

Рассмотрим простейшую *tobit*-модель: $y_t^* = \alpha + \varepsilon_t$, где $\varepsilon_t \sim N(0, \sigma^2)$, и $y_t = y_t^*$, если $y_t^* > 0$, $y_t = 0$, если $y_t^* \leqslant 0$.

Даны результаты 30 наблюдений переменной $y_t$:

| | | | | | |
|---|---|---|---|---|---|
| 0.768 | 2.911 | 0.461 | 0.000 | 0.678 | 0.000 |
| 0.000 | 1.233 | 1.868 | 0.709 | 0.000 | 0.000 |
| 1.010 | 0.000 | 1.422 | 1.543 | 4.411 | 2.385 |
| 0.000 | 2.487 | 3.469 | 1.778 | 2.931 | 1.283 |
| 0.000 | 0.060 | 3.198 | 5.546 | 1.546 | 4.680 |

а) Вычислите МНК-оценку параметра $\alpha$ по цензурированным наблюдениям $y_t$. Будет ли эта оценка завышена или занижена по сравнению с истинным значением $\alpha$?

б) Повторите п. а) для усеченных на уровне 0 наблюдений $y_t$.

## Решение

а) Известно, что МНК-оценка в регрессии на константу равна среднему значению, поэтому при оценке по цензурированным наблюдениям получаем:

$$\widehat{\alpha} = \overline{y} = \frac{46.377}{30} = 1.546.$$

Конечно, следует ожидать, что эта оценка завышена по сравнению с истинным значением $\alpha$, поскольку при вычислении среднего отрицательные значения $y_t$ заменены нулями. Оценивание *tobit*-модели с использованием эконометрического пакета дает значение $\widehat{\alpha} = 1.22$.

б) Если отбросить цензурированные наблюдения, то завышение становится еще больше

$$\widehat{\alpha} = \overline{y} = \frac{46.377}{22} = 2.108.$$

## Задача 12.8

В таблице приведены 100 наблюдений бинарных переменных $(x, y, z)$.

| Количество наблюдений | $y$ | $x$ | $z$ |
|---|---|---|---|
| 12 | 0 | 0 | 0 |
| 10 | 0 | 0 | 1 |
| 20 | 0 | 1 | 0 |
| 0 | 0 | 1 | 1 |
| 8 | 1 | 0 | 0 |
| 28 | 1 | 0 | 1 |
| 22 | 1 | 1 | 0 |
| 0 | 1 | 1 | 1 |

а) В рамках *logit*-модели $\mathrm{P}(y=1) = \Lambda(\alpha + \beta x + \gamma z)$ тестируйте значимость влияния $z$ на $y$.

б) В рамках линейной модели вероятности $\mathrm{P}(y=1) = \alpha + \beta x + \gamma z$ тестируйте значимость влияния $z$ на $y$.

## Решение

а) Рассмотрим модель $\mathrm{P}(y_t = 1) = F(\alpha + \beta x_t + \gamma z_t)$, где $F(u)$ — некоторая функция распределения вероятностей (при $F(u) = \Lambda(u)$ мы получаем *logit*-модель). В нашем случае функция правдоподобия равна:

$$L = F(\alpha)^8 F(\alpha+\gamma)^{28} F(\alpha+\beta)^{22} (1-F(\alpha))^{12} (1-F(\alpha+\gamma))^{10} (1-F(\alpha+\beta))^{20}.$$

Ввиду инвариантности оценок максимального правдоподобия (см. п. 10.4) мы можем перейти от параметров $\alpha$, $\gamma$ и $\beta$ к параметрам $\alpha$, $\delta = \alpha + \beta$ и $\phi = \alpha + \gamma$. Логарифмическая функция правдоподобия в этих параметрах имеет вид:

$$\begin{aligned}\ln L = {} & 8\ln F(\alpha) + 28\ln F(\phi) + 22\ln F(\delta) \\ & + 12\ln(1-F(\alpha)) + 10\ln(1-F(\phi)) + 20\ln(1-F(\delta)).\end{aligned}$$

Приравняв к нулю частные производные, получим такую систему уравнений:

$$\frac{\partial \ln L}{\partial \alpha} = 8\frac{f(\alpha)}{F(\alpha)} - 12\frac{f(\alpha)}{1-F(\alpha)} = 0,$$

$$\frac{\partial \ln L}{\partial \delta} = 22\frac{f(\delta)}{F(\delta)} - 20\frac{f(\delta)}{1-F(\delta)} = 0,$$

$$\frac{\partial \ln L}{\partial \phi} = 28\frac{f(\phi)}{F(\phi)} - 10\frac{f(\phi)}{1-F(\phi)} = 0.$$

Получаем:

$$F(\alpha) = \frac{2}{5}, \quad F(\delta) = \frac{11}{21}, \quad F(\phi) = \frac{14}{19}.$$

(В частном случае *logit*-модели получаем следующие оценки коэффициентов: $\widehat{\alpha} = -0.4055$, $\widehat{\beta} = 0.5008$, $\widehat{\gamma} = 1.4351$.) Значение функции правдоподобия для модели без ограничения равно

$$\ln \widehat{L} = 8\ln\frac{2}{5} + 28\ln\frac{14}{19} + 22\ln\frac{11}{21} + 12\ln\frac{3}{5} + 10\ln\frac{5}{19} + 20\ln\frac{10}{21} \approx -64.4255.$$

Рассмотрим теперь модель с ограничением $\gamma = 0$ (или $\alpha = \phi$): $\mathrm{P}(y_t = 1) = F(\alpha + \beta x_t)$. Логарифмическая функция правдоподобия для модели с ограничением равна:

$$\ln L = 36\ln F(\alpha) + 22\ln F(\delta) + 22\ln(1 - F(\alpha)) + 20\ln(1 - F(\delta)).$$

Приравняв к нулю частные производные, получим такую систему уравнений:

$$\frac{\partial \ln L}{\partial \alpha} = 36\frac{f(\alpha)}{F(\alpha)} - 22\frac{f(\alpha)}{1 - F(\alpha)} = 0,$$
$$\frac{\partial \ln L}{\partial \delta} = 22\frac{f(\delta)}{F(\delta)} - 20\frac{f(\delta)}{1 - F(\delta)} = 0,$$

решая которую получаем:

$$F(\alpha) = \frac{18}{29}, \qquad F(\delta) = \frac{11}{21}.$$

Значение функции правдоподобия для модели с ограничением равно

$$\ln \widetilde{L} = 36\ln\frac{18}{29} + 22\ln\frac{11}{21} + 22\ln\frac{11}{29} + 20\ln\frac{10}{21} \approx -67.5606.$$

Для проверки гипотезы $\gamma = 0$ применим тест отношения правдоподобия (10.54):

$$\mathrm{LR} = -2(\ln \widetilde{L} - \ln \widehat{L}) = 6.27.$$

Полученное значение превышает критическое значение $\chi^2_{0.05}(1) = 3.84$, т. е. гипотеза $\gamma = 0$ уверенно отвергается на 5%-ном уровне значимости. (Заметим, что этот результат не зависит от выбора функции распределения $F(u)$ и, в частности, верен и для *logit*-модели.)

б) Рассмотрим теперь линейную модель вероятности

$$\mathrm{P}(y = 1) = \alpha + \beta x + \gamma z.$$

Запишем ее в виде

$$y_t = \alpha + \beta x_t + \gamma z_t + \varepsilon_t = x_t'\beta + \varepsilon_t.$$

Поскольку $\varepsilon_t$ принимает только два значения $-\boldsymbol{x}_t'\boldsymbol{\beta}$ и $1-\boldsymbol{x}_t'\boldsymbol{\beta}$, то из условия $\mathrm{E}(\varepsilon_t)=0$ следует, что вероятности этих значений равны соответственно $1-\boldsymbol{x}_t'\boldsymbol{\beta}$ и $-\boldsymbol{x}_t'\boldsymbol{\beta}$. Таким образом, на истинные значения параметров и значения независимых переменных в силу модели наложены условия $0 \leqslant \boldsymbol{x}_t'\boldsymbol{\beta} \leqslant 1$. Тогда функция правдоподобия равна функции правдоподобия из п. а) с $F(u)=u$. Поскольку вычисление значения теста отношения правдоподобия LR в п. а) не зависело от вида функции $F(u)$, то получим тот же результат $\mathrm{LR}=6.27$ и тот же вывод о том, что гипотеза $\gamma=0$ отвергается на 5%-ном уровне значимости.

Заметим, что если мы оцениваем эту модель методом наименьших квадратов, то по стандартным формулам получаем значение $t$-статистики оценки $\widehat{\gamma}$, равное $t_{\widehat{\gamma}}=2.52$. Однако, так как в модели присутствует гетероскедастичность, то этому результату не следует доверять.

## Задача 12.9

Запишите функцию правдоподобия для оценки параметров $(\boldsymbol{\beta},\sigma^2)$ модели

$$y_t^* = \boldsymbol{x}_t'\boldsymbol{\beta}+\varepsilon_t, \qquad (y_t^* \text{— ненаблюдаемая переменная}),$$
$$y_t = \max\{y_t^*, c\}, \qquad \text{где } c\neq 0,\ \mathrm{E}\varepsilon=0,\ \mathrm{V}(\varepsilon)=\sigma^2\boldsymbol{I}.$$

## Решение

Найдем вероятность того, что $y_t=c$:

$$\mathrm{P}\{y_t=c\}=\mathrm{P}\{y_t^*\leqslant c\}=\mathrm{P}\{\boldsymbol{x}_t'\boldsymbol{\beta}+\varepsilon_t\leqslant c\}=\mathrm{P}\{\varepsilon_t\leqslant c-\boldsymbol{x}_t'\boldsymbol{\beta}\}=\Phi\left(\frac{c-\boldsymbol{x}_t'\boldsymbol{\beta}}{\sigma}\right).$$

Далее, так как наблюдения имеют смешанное распределение, то функция правдоподобия имеет вид (ср. с (12.32)):

$$L=\prod_{y_t=c}\Phi\left(\frac{c-\boldsymbol{x}_t'\boldsymbol{\beta}}{\sigma}\right)\prod_{y_t>c}\frac{1}{\sqrt{2\pi}\sigma}\exp\left(-\frac{1}{2\sigma^2}(y_t-\boldsymbol{x}_t'\boldsymbol{\beta})^2\right).$$

## Задача 12.10

Пусть $y_t^*=\beta_1+\beta_2x_{t2}+\beta_3x_{t3}+u_t$, где $u_t\sim N(0,\sigma^2)$ и $u_t$ независимы.

1) Бинарная переменная $d$ определяется следующим образом:

$$d_t=\begin{cases}1, & \text{если } y_t^*>0,\\ 0, & \text{в противном случае.}\end{cases}$$

а) Выпишите вероятность того, что $d_t=1$, как функцию переменных $x_{ti}$.

б) Какие параметры вы можете оценить по наблюдениям $(x_{ti}, d_t)$?

в) Найдите выражение для асимптотической матрицы ковариаций оценок максимального правдоподобия в этом случае.

г) Вы хотите проверить гипотезу, что переменная $x_{t2}$ незначима. Опишите процедуру проверки этой гипотезы с помощью теста отношения правдоподобия.

2) Пусть

$$y_t = \begin{cases} y_t^*, & \text{если } y_t^* > 0, \\ 0, & \text{в противном случае.} \end{cases}$$

а) Какие параметры можно оценить по наблюдениям $(x_t, y_t)$?

б) Найдите выражение для предельного эффекта фактора $x_{t3}$ для $y_t$ и $y_t^*$.

в) Повторите процедуру проверки гипотезы о незначимости $x_{t2}$ из п. 1 г) в данном случае. Какой из этих двух тестов выглядит для вас более предпочтительным?

## Решение

1) а) Найдем вероятность того, что $d_t = 1$:

$$\begin{aligned} \mathrm{P}\{d_t = 1\} &= \mathrm{P}\{y_t^* > 0\} = 1 - \mathrm{P}\{y_t^* \leqslant 0\} = 1 - \mathrm{P}\{\beta_1 + \beta_2 x_{t2} + \beta_3 x_{t3} + u_t \leqslant 0\} \\ &= 1 - \mathrm{P}\{u_t \leqslant -(\beta_1 + \beta_2 x_{t2} + \beta_3 x_{t3})\} \\ &= 1 - \Phi\left(-\frac{\beta_1 + \beta_2 x_{t2} + \beta_3 x_{t3}}{\sigma}\right) = \Phi\left(\frac{\beta_1}{\sigma} + \frac{\beta_2}{\sigma} x_{t2} + \frac{\beta_3}{\sigma} x_{t3}\right). \end{aligned}$$

б) Как видим, вероятность $\mathrm{P}\{d_t = 1\}$ зависит только от отношений $\beta_1/\sigma$, $\beta_2/\sigma$, $\beta_3/\sigma$. Поэтому оценить все параметры $\beta_1$, $\beta_2$, $\beta_3$ и $\sigma$ невозможно. Можно оценить только отношения $\beta_1/\sigma$, $\beta_2/\sigma$, $\beta_3/\sigma$. (Если известны какие-нибудь априорные ограничения на параметры, как, например, $\sigma = 1$ в *probit*-модели, то все остальные параметры можно оценить.)

в) Так как мы можем оценить только отношения $\beta_1/\sigma$, $\beta_2/\sigma$, $\beta_3/\sigma$, то положим $\sigma = 1$. Обозначим $\boldsymbol{x}_t = \begin{bmatrix} x_{t1} & x_{t2} & x_{t3} \end{bmatrix}'$, $\boldsymbol{\beta} = \begin{bmatrix} \beta_1 & \beta_2 & \beta_3 \end{bmatrix}'$ и рассмотрим оценку максимального правдоподобия для $\boldsymbol{\beta}$. Логарифмическая функция правдоподобия для нашей модели равна:

$$\ln L_n = \sum_{t=1}^{n} (d_t \ln \Phi(\boldsymbol{x}_t'\boldsymbol{\beta}) + (1 - d_t) \ln(1 - \Phi(\boldsymbol{x}_t'\boldsymbol{\beta}))). \qquad (*)$$

Асимптотическая ковариационная матрица оценки максимального правдоподобия равна

$$\mathcal{F}^{-1}(\boldsymbol{\beta}), \quad \text{где } \mathcal{F}(\boldsymbol{\beta}) = \lim_{n\to\infty} \frac{1}{n} \mathcal{F}_n(\boldsymbol{\beta}) = -\lim_{n\to\infty} \frac{1}{n} \mathrm{E}\left(\frac{\partial^2 \ln L_n}{\partial \boldsymbol{\beta} \partial \boldsymbol{\beta}'}\right)$$

— асимптотическая информационная матрица (см. 9.8).

Для нашей модели

$$\frac{\partial \ln L_n}{\partial \beta} = \sum_{t=1}^{n} \left( d_t \frac{\phi(x_t'\beta)}{\Phi(x_t'\beta)} - (1-d_t)\frac{\phi(x_t'\beta)}{1-\Phi(x_t'\beta)} \right) x_t.$$

Вторая производная равна

$$\frac{\partial^2 \ln L_n}{\partial \beta \partial \beta'} = \sum_{t=1}^{n} \Bigg( d_t \frac{\phi'(x_t'\beta)\Phi(x_t'\beta) - \phi^2(x_t'\beta)}{\Phi^2(x_t'\beta)} - (1-d_t)\frac{\phi'(x_t'\beta)(1-\Phi(x_t'\beta)) + \phi^2(x_t'\beta)}{(1-\Phi(x_t'\beta))^2} \Bigg) x_t x_t'.$$

Переменная $d_t$ — бернуллиевская случайная величина, принимающая значение 1 с вероятностью $\Phi(x_t'\beta)$. Поэтому матрица $\mathcal{F}_n(\beta)$ равна

$$\begin{aligned}
\mathcal{F}_n(\beta) &= -\mathrm{E}\left(\frac{\partial^2 \ln L_n}{\partial \beta \partial \beta'}\right) \\
&= -\sum_{t=1}^{n} \Bigg( \mathrm{E}(d_t) \frac{\phi'(x_t'\beta)\Phi(x_t'\beta) - \phi^2(x_t'\beta)}{\Phi^2(x_t'\beta)} \\
&\qquad - \mathrm{E}(1-d_t)\frac{\phi'(x_t'\beta)(1-\Phi(x_t'\beta)) + \phi^2(x_t'\beta)}{(1-\Phi(x_t'\beta))^2} \Bigg) x_t x_t' \\
&= -\sum_{t=1}^{n} \Bigg( \Phi(x_t'\beta) \frac{\phi'(x_t'\beta)\Phi(x_t'\beta) - \phi^2(x_t'\beta)}{\Phi^2(x_t'\beta)} \\
&\qquad - (1-\Phi(x_t'\beta))\frac{\phi'(x_t'\beta)(1-\Phi(x_t'\beta)) + \phi^2(x_t'\beta)}{(1-\Phi(x_t'\beta))^2} \Bigg) x_t x_t' \\
&= -\sum_{t=1}^{n} \Bigg( \frac{\phi'(x_t'\beta)\Phi(x_t'\beta) - \phi^2(x_t'\beta)}{\Phi(x_t'\beta)} \\
&\qquad - \frac{\phi'(x_t'\beta)(1-\Phi(x_t'\beta)) + \phi^2(x_t'\beta)}{1-\Phi(x_t'\beta)} \Bigg) x_t x_t' \\
&= \sum_{t=1}^{n} \frac{\phi^2(x_t'\beta)}{\Phi(x_t'\beta)(1-\Phi(x_t'\beta))} x_t x_t'.
\end{aligned}$$

Таким образом, асимптотическая матрица ковариаций оценки максимального правдоподобия вектора $\beta$ равна

$$\left( \lim_{n\to\infty} \frac{1}{n} \mathcal{F}_n(\beta) \right)^{-1} = \left( \lim_{n\to\infty} \frac{1}{n} \sum_{t=1}^{n} \frac{\phi^2(x_t'\beta)}{\Phi(x_t'\beta)(1-\Phi(x_t'\beta))} x_t x_t' \right)^{-1}.$$

Заметим, что в качестве оценки асимптотической ковариационной матрицы оценки $\widehat{\beta}_{\mathrm{ML}}$ можно взять

$$\left( \frac{1}{n} \mathcal{F}_n(\widehat{\beta}_{\mathrm{ML}}) \right)^{-1} = \left( \frac{1}{n} \sum_{t=1}^{n} \frac{\phi^2(x_t'\widehat{\beta}_{\mathrm{ML}})}{\Phi(x_t'\widehat{\beta}_{\mathrm{ML}})(1-\Phi(x_t'\widehat{\beta}_{\mathrm{ML}}))} x_t x_t' \right)^{-1}.$$

г) Для проверки гипотезы $\beta_2 = 0$ (незначимость переменной $x_2$) с помощью теста отношения правдоподобия следует проделать следующие шаги:

1. найти оценку $\widehat{\beta}_{\mathrm{UR}}$, максимизирующую функцию правдоподобия $(*)$ без ограничений;
2. найти оценку $\widehat{\beta}_{\mathrm{R}}$, максимизирующую функцию правдоподобия $(*)$ при ограничении $\beta_2 = 0$;
3. вычислить значение статистики $\mathrm{LR} = -2(L_n(\widehat{\beta}_{\mathrm{R}}) - L_n(\widehat{\beta}_{\mathrm{UR}}))$;
4. так как при условии выполнения нулевой гипотезы статистика LR подчиняется $\chi^2$-распределению с одной степенью свободы, то полученное значение LR надо сравнить с 5%-ной точкой $\chi^2_{0.05}$ этого распределения (тестируем гипотезу на 5%-ном уровне значимости);
5. если $\mathrm{LR} > \chi^2_{0.05}$, то гипотеза о незначимости переменной $x_2$ отвергается, иначе — не отвергается.

2) а) В этом случае модель является моделью с цензурированными наблюдениями (*tobit*-моделью). Значит, можно оценить все четыре параметра $\beta_1$, $\beta_2$, $\beta_3$ и $\sigma$.

б) Для переменной $y_t^*$ справедлива линейная модель; поэтому

$$\mathrm{E}y_t^* = \mathrm{E}(\beta_1 + \beta_2 x_{t2} + \beta_3 x_{t3} + u_t + \varepsilon_t) = \beta_1 + \beta_2 x_{t2} + \beta_3 x_{t3}.$$

Следовательно, предельный эффект фактора $x_{t3}$ для $y_t^*$ равен

$$\frac{\partial \mathrm{E}y_t^*}{\partial x_{t3}} = \beta_3.$$

Математическое ожидание $y_t$ равно (см. (12.31)):

$$\mathrm{E}y_t = \Phi\left(\frac{\beta_1 + \beta_2 x_{t2} + \beta_3 x_{t3}}{\sigma}\right)\cdot(\beta_1 + \beta_2 x_{t2} + \beta_3 x_{t3}) + \sigma\phi\left(\frac{\beta_1 + \beta_2 x_{t2} + \beta_3 x_{t3}}{\sigma}\right).$$

Предельный эффект фактора $x_{t3}$ для $y_t$ равен

$$\begin{aligned}\frac{\partial \mathrm{E}y_t}{\partial x_{t3}} &= \phi\left(\frac{\beta_1 + \beta_2 x_{t2} + \beta_3 x_{t3}}{\sigma}\right)\cdot(\beta_1 + \beta_2 x_{t2} + \beta_3 x_{t3})\cdot\frac{\beta_3}{\sigma} \\ &\quad + \Phi\left(\frac{\beta_1 + \beta_2 x_{t2} + \beta_3 x_{t3}}{\sigma}\right)\cdot\beta_3 + \sigma\phi'\left(\frac{\beta_1 + \beta_2 x_{t2} + \beta_3 x_{t3}}{\sigma}\right)\cdot\frac{\beta_3}{\sigma} \\ &= \Phi\left(\frac{\beta_1 + \beta_2 x_{t2} + \beta_3 x_{t3}}{\sigma}\right)\cdot\beta_3,\end{aligned}$$

в силу того, что $\phi'(x) = -x\phi(x)$.

в) Процедура для теста отношения правдоподобия в этом случае в точности такая же, как и в п. 1г). Отличие в функции правдоподобия, которая теперь имеет вид (см. (12.32)):

$$L_n = \prod_{y_t=0}\left(1-\Phi\left(\frac{x_t'\beta}{\sigma}\right)\right)\prod_{y_t>0}\frac{1}{\sqrt{2\pi}\sigma}\exp\left(-\frac{1}{2\sigma^2}(y_t-x_t'\beta)^2\right).$$

Оба теста (в п. 1г) и в п. 2в)) позволяют проверить гипотезу о незначимости переменной $x_2$. Из них разумно было бы предпочесть тест большей мощности. По-видимому, в данном случае более мощный тест в п. 2в). Интуитивно это следует из того, что при оценивании параметров используется больше информации.

## Задача 12.11

Дано $n = n_1 + n_2 + n_3$ наблюдений переменных $x$ и $y$. Известно, что для $n_1$ наблюдений $y = 1$ и $x = 1$, для $n_2$ наблюдений $y = 0$ и $x = 1$, для $n_3$ наблюдений $y = 0$ и $x = 0$. Покажите, что как для *logit*-, так и для *probit*-модели уравнение правдоподобия не имеет решения.

## Решение

Рассмотрим модель $\mathrm{P}(y_t = 1) = F(\alpha + \beta x_t)$. Функция правдоподобия равна (см. (12.8))

$$L = F(\alpha+\beta)^{n_1}(1-F(\alpha+\beta))^{n_2}(1-F(\alpha))^{n_3}.$$

Обозначим $\gamma = \alpha + \beta$. Логарифмическая функция правдоподобия равна

$$\ln L = n_1 \ln F(\gamma) + n_2 \ln(1-F(\gamma)) + n_3 \ln(1-F(\alpha)).$$

Система уравнений для оценки максимального правдоподобия такова:

$$\frac{\partial \ln L}{\partial \alpha} = -n_3\frac{f(\alpha)}{1-F(\alpha)} = 0,$$

$$\frac{\partial \ln L}{\partial \gamma} = n_1\frac{f(\gamma)}{F(\gamma)} - n_2\frac{f(\gamma)}{1-F(\gamma)} = 0.$$

Как видим, в первом из уравнений левая часть меньше нуля при всех $\alpha$. Значит, система не имеет решений.

## Задача 12.12

Покажите, что логарифмическая функция правдоподобия для *probit*-модели является вогнутой (по $\beta$) функцией.

*Указание.* Покажите, что при любом $x$ выполнено неравенство

$$x + \phi(x)/\Phi(x) > 0,$$

и воспользуйтесь этим фактом.

## Решение

Докажем сначала следующую лемму:

*Лемма.* Пусть $f_t(x)$ — строго вогнутые вещественные функции, $\boldsymbol{x}_1, \ldots, \boldsymbol{x}_n$ — $k$-мерные векторы, такие что матрица

$$\boldsymbol{X} = \begin{bmatrix} \boldsymbol{x}_1' \\ \vdots \\ \boldsymbol{x}_n' \end{bmatrix}$$

имеет ранг $k$. Тогда функция

$$G(\boldsymbol{\beta}) = \sum_{t=1}^{n} f_t(\boldsymbol{x}_t'\boldsymbol{\beta})$$

является строго вогнутой по $\boldsymbol{\beta}$.

*Доказательство.* Рассмотрим матрицу вторых производных

$$\boldsymbol{D} = \frac{\partial^2 G}{\partial\boldsymbol{\beta}\partial\boldsymbol{\beta}'}(\boldsymbol{\beta}) = \sum_{t=1}^{n} f_t''(\boldsymbol{x}_t'\boldsymbol{\beta})\boldsymbol{x}_t\boldsymbol{x}_t'.$$

Докажем, что эта матрица отрицательно определена.

Рассмотрим произвольный $k$-мерный вектор $\boldsymbol{z}$. Тогда, в силу того, что $f_t''(x) < 0$ при любом $x$, имеем

$$\boldsymbol{z}'\boldsymbol{D}\boldsymbol{z} = \sum_{t=1}^{n} f_t''(\boldsymbol{x}_t'\boldsymbol{\beta})\boldsymbol{z}'\boldsymbol{x}_t\boldsymbol{x}_t'\boldsymbol{z} = \sum_{t=1}^{n} f_t''(\boldsymbol{x}_t'\boldsymbol{\beta})(\boldsymbol{x}_t'\boldsymbol{z})^2 \leqslant 0.$$

Предположим теперь, что вектор $\boldsymbol{z}$ такой, что $\boldsymbol{z}'\boldsymbol{D}\boldsymbol{z} = 0$. Значит, $\boldsymbol{x}_t'\boldsymbol{z} = 0$ для любого $t = 1, \ldots, n$, или $\boldsymbol{X}\boldsymbol{z} = 0$. Так как ранг матрицы $\boldsymbol{X}$ равен $k$, то отсюда следует, что $\boldsymbol{z} = 0$. Таким образом, если $\boldsymbol{z} \neq 0$, то $\boldsymbol{z}'\boldsymbol{D}\boldsymbol{z} < 0$, следовательно, матрица $\boldsymbol{D}$ отрицательно определена. Таким образом, функция $G(\boldsymbol{\beta})$ строго вогнутая по $\boldsymbol{\beta}$. Лемма доказана.

Теперь рассмотрим логарифмическую функцию правдоподобия *probit*-модели (см. (12.9)):

$$l = \sum_{t=1}^{n} (y_t \ln\Phi(\boldsymbol{x}_t'\boldsymbol{\beta}) + (1 - y_t)\ln(1 - \Phi(\boldsymbol{x}_t'\boldsymbol{\beta}))) = \sum_{t=1}^{n} f_t(\boldsymbol{x}_t'\boldsymbol{\beta}),$$

где $f_t(x) = y_t \ln\Phi(x) + (1 - y_t)\ln(1 - \Phi(x))$ и $y_t$ может принимать значения 0 или 1.

Если $y_t = 1$, то вторая производная функции $f_t(x)$ равна

$$f_t''(x) = \left(\frac{\phi(x)}{\Phi(x)}\right)' = \frac{\phi'(x)\Phi(x) - \phi^2(x)}{\Phi^2(x)} = -\frac{x\phi(x)\Phi(x) + \phi^2(x)}{\Phi^2(x)}$$

$$= -\frac{x\Phi(x) + \phi(x)}{\Phi^2(x)/\phi(x)}.$$

(Здесь мы воспользовались равенством $\phi'(x) = -x\phi(x)$.)

Знаменатель дроби всегда положительный. Докажем, что и числитель тоже всегда положительный, т. е. $x\Phi(x) + \phi(x) > 0$ при всех $x$. Рассмотрим производную

$$(x\Phi(x) + \phi(x))' = \Phi(x) + x\phi(x) + \phi'(x) = \Phi(x) > 0.$$

При $x \to -\infty$ функция $\Phi(x)$ стремится к нулю быстрее, чем $1/x$. Это можно показать, например, так: при $x < -1$

$$0 < \Phi(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} e^{-t^2/2} dt < \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} e^{t/2} dt = \sqrt{\frac{2}{\pi}} e^{x/2}.$$

Отсюда следует, что $x\Phi(x) \to 0$, а значит, и $x\Phi(x) + \phi(x) \to 0$ при $x \to -\infty$. Отсюда и из того, что функция $x\Phi(x) + \phi(x)$ возрастающая, следует, что $x\Phi(x) + \phi(x) > 0$ при всех $x$.

Таким образом, функция $f_t(x)$ строго вогнутая, если $y_t = 1$. Если $y_t = 0$, то $f_t(x) = \ln(1 - \Phi(x)) = \ln \Phi(-x)$, и функция $f_t(-x)$ строго вогнутая (по доказанному выше). Но тогда и сама функция $f_t(x)$ строго вогнутая.

Применяя доказанную выше лемму, получаем утверждение задачи. А именно, логарифмическая функция правдоподобия *probit*-модели является строго вогнутой по $\beta$ функцией.

## Задача 12.13

Модель бинарного выбора описывается стандартным образом:

$$y_t = \begin{cases} 1, & \text{если } y_t^* > 0, \\ 0, & \text{в противном случае,} \end{cases}$$

где $y_t^* = x_t'\beta + \varepsilon_t$, а ошибки $\varepsilon_t$ имеют распределение Лапласа.

а) Найдите логарифмическую функцию правдоподобия для оценивания вектора $\beta$. Является ли эта функция вогнутой по $\beta$?

б) Предположим, что вы оценили вектор $\beta$, используя *probit*-модель, и эти оценки примерно пропорциональны оценкам, полученным с помощью исходной модели. Чему приблизительно должен быть равен коэффициент пропорциональности?

## Решение

а) Плотность распределения Лапласа равна

$$f(x) = \frac{1}{2} e^{-|x|},$$

функция распределения равна

$$F(x) = \int_{-\infty}^{x} f(t)dt = \int_{-\infty}^{x} \frac{1}{2} e^{-|x|} dt = \begin{cases} \frac{1}{2} e^{x}, & \text{при } x \leqslant 0, \\ 1 - \frac{1}{2} e^{-x}, & \text{при } x \geqslant 0. \end{cases}$$

Значит, логарифмическая функция правдоподобия равна (см. (12.9)):

$$
\begin{aligned}
l &= \sum_{t=1}^{n}(y_t \ln F(x_t'\beta) + (1-y_t)\ln(1-F(x_t'\beta))) \\
&= \sum_{y_t=1} \ln F(x_t'\beta) + \sum_{y_t=0} \ln(1-F(x_t'\beta)) \\
&= \sum_{\substack{y_t=1 \\ x_t'\beta\geqslant 0}} \ln\left(1-\frac{1}{2}e^{-x_t'\beta}\right) + \sum_{\substack{y_t=1 \\ x_t'\beta<0}} \ln\left(\frac{1}{2}e^{x_t'\beta}\right) \\
&\quad + \sum_{\substack{y_t=0 \\ x_t'\beta>0}} \ln\left(\frac{1}{2}e^{-x_t'\beta}\right) + \sum_{\substack{y_t=0 \\ x_t'\beta\leqslant 0}} \ln\left(1-\frac{1}{2}e^{x_t'\beta}\right) \\
&= \sum_{\substack{y_t=1 \\ x_t'\beta\geqslant 0}} \ln\left(1-\frac{1}{2}e^{-x_t'\beta}\right) + \sum_{\substack{y_t=1 \\ x_t'\beta<0}} \left(\ln\frac{1}{2}+x_t'\beta\right) \\
&\quad + \sum_{\substack{y_t=0 \\ x_t'\beta>0}} \left(\ln\frac{1}{2}-x_t'\beta\right) + \sum_{\substack{y_t=0 \\ x_t'\beta\leqslant 0}} \ln\left(1-\frac{1}{2}e^{x_t'\beta}\right).
\end{aligned}
$$

Рассмотрим функцию

$$
g(x) = \begin{cases} \ln\frac{1}{2}+x, & \text{если } x<0, \\ \ln\left(1-\frac{1}{2}e^{-x}\right), & \text{если } x\geqslant 0. \end{cases}
$$

Если $x>0$, то вторая производная

$$
g''(x) = -\frac{\frac{1}{2}e^{-x}}{\left(1-\frac{1}{2}e^{-x}\right)} < 0.
$$

В точке $x=0$ производные

$$
\left(\ln\frac{1}{2}+x\right)' = 1 \quad \text{и} \quad \left(\ln\left(1-\frac{1}{2}e^{-x}\right)\right)' = \frac{\frac{1}{2}e^{-x}}{1-\frac{1}{2}e^{-x}}
$$

совпадают. Следовательно, функция $g(x)$ является вогнутой на множестве всех вещественных чисел, и $g(x_t'\beta)$ является вогнутой на $R^n$.

Логарифмическую функцию правдоподобия можно переписать в виде:

$$
l = \sum_{y_t=1} g(x_t'\beta) + \sum_{y_t=0} g(-x_t'\beta).
$$

Следовательно, она является вогнутой (как сумма вогнутых функций).

Заметим, что строго вогнутой логарифмическая функция правдоподобия, вообще говоря, не является. На некоторых участках она может быть линейной.

б) Если оценки вектора коэффициентов $\beta$ оказались примерно пропорциональными для модели с ошибками, распределенными по Лапласу, и *probit*-модели, то коэффициент пропорциональности примерно равен отношению среднеквадратичных отклонений распределения Лапласа и стандартного нормального распределения. То есть

$$\frac{\widehat{\beta}_{\mathrm{Laplace}}}{\widehat{\beta}_{\mathrm{Probit}}} \approx \frac{\sqrt{2}}{1} = \sqrt{2}.$$

(Здесь через $\widehat{\beta}_{\mathrm{Laplace}}$ обозначена оценка в модели с ошибками, распределенными по Лапласу, $\widehat{\beta}_{\mathrm{Probit}}$ — оценка в *probit*-модели.)

## Задача 12.14

Пусть $y_t^* = x_t'\beta + \varepsilon_t$, где ошибки $\varepsilon_t$ имеют плотность распределения $f(x)$ и

$$y_t = \begin{cases} \alpha_1, & \text{если } y_t^* \leqslant \alpha_1, \\ y_t^*, & \text{если } \alpha_1 < y_t^* < \alpha_2, \\ \alpha_2, & \text{если } y_t^* \geqslant \alpha_2. \end{cases}$$

а) Найдите распределение $y_t$.

б) Найдите логарифмическую функцию правдоподобия для оценивания вектора $\beta$.

в) Найдите $(\partial \mathrm{E} y)/(\partial x)$.

## Решение

а) Найдем функцию распределения $y_t$. При $x < \alpha_1$ функция распределения $F_{y_t}(x) = 0$, при $x \geqslant \alpha_2$ функция распределения $F_{y_t}(x) = 1$.

Пусть $\alpha_1 < y_t < \alpha_2$. Тогда $y_t = y_t^*$, и

$$F_{y_t}(x) = \mathrm{P}\{y_t^* \leqslant x\} = \mathrm{P}\{x_t'\beta + \varepsilon_t \leqslant x\} = \mathrm{P}\{\varepsilon_t \leqslant x - x_t'\beta\} = F(x - x_t'\beta).$$

(Здесь $F(x)$ — функция распределения ошибок $\varepsilon_t$.) Таким образом,

$$F_{y_t}(x) = \begin{cases} 0, & \text{если } x < \alpha_1, \\ F(x - x_t'\beta), & \text{если } \alpha_1 \leqslant x < \alpha_2, \\ 1, & \text{если } x \geqslant \alpha_2. \end{cases}$$

При этом

$$\mathrm{P}\{y_t = \alpha_1\} = F(\alpha_1 - x_t'\beta), \quad \mathrm{P}\{y_t = \alpha_2\} = 1 - F(\alpha_2 - x_t'\beta).$$

б) Действуя аналогично построению функции правдоподобия для *tobit*-модели (см. (12.32)), получаем следующую функцию правдоподобия:

$$L = \prod_{y_t=\alpha_1} F(\alpha_1 - x_t'\beta) \prod_{\alpha_1 < y_t < \alpha_2} f(y_t - x_t'\beta) \prod_{y_t=\alpha_2} (1 - F(\alpha_2 - x_t'\beta)).$$

Логарифмическая функция правдоподобия равна

$$l = \sum_{y_t=\alpha_1} \ln F(\alpha_1 - x_t'\beta) \sum_{\alpha_1 < y_t < \alpha_2} \ln f(y_t - x_t'\beta) \sum_{y_t=\alpha_2} \ln(1 - F(\alpha_2 - x_t'\beta)).$$

в) Математическое ожидание величины $y_t$ равно (ср. с (12.31)):

$$\begin{aligned}
\mathrm{E}(y_t) &= \mathrm{E}(y_t \mid y_t^* \leqslant \alpha_1)\mathrm{P}\{y_t^* \leqslant \alpha_1\} + \mathrm{E}(y_t \mid \alpha_1 < y_t^* < \alpha_2)\mathrm{P}\{\alpha_1 < y_t^* < \alpha_1\} \\
&+ \mathrm{E}(y_t \mid y_t^* \geqslant \alpha_2)\mathrm{P}\{y_t^* \geqslant \alpha_2\} \\
&= \alpha_1 F(\alpha_1 - x_t'\beta) + \int_{\alpha_1}^{\alpha_2} x f(x - x_t'\beta)dx + \alpha_2(1 - F(\alpha_2 - x_t'\beta)) \\
&= \alpha_1 F(\alpha_1 - x_t'\beta) + \int_{\alpha_1}^{\alpha_2} x dF(x - x_t'\beta) + \alpha_2(1 - F(\alpha_2 - x_t'\beta)) \\
&= \alpha_1 F(\alpha_1 - x_t'\beta) + \alpha_2 F(\alpha_2 - x_t'\beta) - \alpha_1 F(\alpha_1 - x_t'\beta) \\
&- \int_{\alpha_1}^{\alpha_2} F(x - x_t'\beta)dx + \alpha_2(1 - F(\alpha_2 - x_t'\beta)) \\
&= -\int_{\alpha_1}^{\alpha_2} F(x - x_t'\beta)dx + \alpha_2.
\end{aligned}$$

Найдем теперь предельный эффект:

$$\frac{\partial \mathrm{E} y_t}{\partial x_t} = -\int_{\alpha_1}^{\alpha_2} f(x - x_t'\beta)(-\beta)dx = (F(\alpha_2 - x_t'\beta) - F(\alpha_1 - x_t'\beta))\beta.$$

## Задача 12.15

Докажите, что в (12.40):

$$\varepsilon_t = \frac{\sigma_{\varepsilon u}}{\sigma_u^2} u_t + \eta_t = \sigma_{\varepsilon u} u_t + \eta_t,$$

где $[\varepsilon_t, u_t]'$ — двумерный нормальный вектор, $\mathrm{E}(\varepsilon_t) = \mathrm{E}(u_t) = 0$, $\mathrm{V}(\varepsilon_t) = \sigma_\varepsilon^2$, $\mathrm{V}(u_t) = \sigma_u^2 = 1$, $\mathrm{Cov}(\varepsilon_t, u_t) = \sigma_{\varepsilon u}$, случайные величины $u_t$ и $\eta_t$ независимы.

## Решение

Рассмотрим ковариацию величин $u_t$ и $\eta_t$:

$$\begin{aligned}
\mathrm{Cov}(u_t, \eta_t) &= \mathrm{E} u_t \eta_t = \mathrm{E}\left(u_t\left(\varepsilon_t - \frac{\sigma_{\varepsilon u}}{\sigma_u^2} u_t\right)\right) = \mathrm{E} u_t \varepsilon_t - \frac{\sigma_{\varepsilon u}}{\sigma_u^2} \mathrm{E} u_t^2 \\
&= \sigma_{\varepsilon u} - \frac{\sigma_{\varepsilon u}}{\sigma_u^2}\sigma_u^2 = 0.
\end{aligned}$$

Получилось, что $u_t$ и $\eta_t$ некоррелированные. Но в силу того, что $u_t$ и $\eta_t$ являются компонентами нормального вектора, из некоррелированности следует и их независимость (см. приложение МС, п. 4, N4).

## Задача 12.16

Расходы домашних хозяйств в Нидерландах (см. начало в упражнении 6.14). *Расходы на отдых и модели с усеченными переменными.* Не все семьи расходуют деньги на отдых. В нашем случае $v3 = 0$ для 22.5% наблюдений. В этом разделе мы сначала рассмотрим модели бинарного выбора для ответа на вопрос, тратит семья какие-нибудь средства на отдых или нет, игнорируя информацию о размерах этих затрат. Мы рассмотрим также *tobit*-модель, в которой явно учитывается смешанный дискретно-непрерывный тип переменной $v3$.

12.16.1. а) Постройте фиктивную переменную $y$, такую что $y = 1$, если $v3 > 0$, и $y = 0$, если $v3 = 0$.

б) Оцените линейную модель вероятности для $y$ (выберите подходящий набор объясняющих переменных среди тех, что использовались в упражнениях 6.14.1–6.14.10). Проинтерпретируйте результаты. Как вычислены стандартные ошибки? Почему?

в) Найдите прогноз для $y$, основываясь на линейной модели вероятности. Находятся ли прогнозные значения в интервале от 0 до 1?

12.16.2. а) Оцените *probit*-модель для $y$. Проинтерпретируйте результаты.

б) Проверьте совместную значимость переменных в модели п. а).

в) Используя стратегию «от общего к частному», постройте подходящую *probit*-модель для $y$. Проинтерпретируйте результаты. Используйте эту модель в последующих упражнениях.

г) Найдите прогноз для $y_i$, т. е. оценки вероятностей событий $y_i = 1$. Вычислите также коэффициент детерминации для этой *probit*-модели.

12.16.3. а) Повторите упражнение 12.16.2, используя *logit*-модель вместо *probit*-модели.

б) Сравните результаты линейной модели вероятности, *probit*-модели и *logit*-модели.

12.16.4. а) Оцените *tobit*-модель для переменной $y = \ln(v3 + 1)$, используя регрессоры, выбранные ранее в упражнении 12.16.2.

б) Сравните результаты линейной модели вероятности, *probit*-модели, *logit*-модели и *tobit*-модели.

# Решение

12.16.1. а, б) При построении линейной вероятностной модели оказалось, что практически все регрессоры незначимые. Оставив значимые на 10%-ном уровне факторы, получаем модель, представленную в таблице 12.1. Как ви-

Dependent Variable: $y$
White Heteroskedasticity-Consistent Standard Errors

**Таблица 12.1**

| Variable | Coefficient | Std. Error | $t$-Statistic | Probability |
|---|---|---|---|---|
| const | −1.921781 | 0.430275 | −4.466405 | 0.0000 |
| $ltot3$ | 0.252804 | 0.039562 | 6.390017 | 0.0000 |
| $nahm - nch1217 - nch18$ | −0.036657 | 0.020965 | −1.748461 | 0.0811 |
| $durb$ | 0.012841 | 0.006871 | 1.868741 | 0.0623 |
| $R^2$ | 0.091485 | | | |

дим, вероятность того, что семья расходует деньги на отдых, положительно зависит от дохода семьи, отрицательно зависит от числа взрослых членов семьи и положительно зависит от степени урбанизации. В модели применяются стандартные ошибки в форме Уайта, так как по построению ошибки в ней являются гетероскедастичными.

в) Выборочные статистики переменной $\widehat{y}$ представлены в таблице 12.2. Как

**Таблица 12.2**

| | Mean | Median | Maximum | Minimum | Std. Dev. |
|---|---|---|---|---|---|
| $\widehat{y}$ | 0.775176 | 0.769459 | 1.159832 | 0.414425 | 0.126417 |

видим, наибольшее значение переменной $\widehat{y}$ больше 1. Следовательно, прогноз не всегда можно интерпретировать как вероятность того, что семья будет расходовать деньги на отдых.

12.16.2. а, б, в) При отборе модели используются значимость коэффициентов, информационные критерии (Акаике и Шварца). Можно остановиться, например, на модели, представленной в таблице 12.3. При этом так же, как и в упражнении 12.16.1, пришлось оставить факторы, значимые на 10%-ном уровне.

Как видим, вероятность того, что семья будет тратить деньги на отдых, положительно зависит от дохода и степени урбанизации и отрицательно зависит от числа взрослых членов семьи (так же, как и в упражнении 12.16.1).

г) Коэффициент детерминации (McFadden $R^2$) для этой $probit$-модели равен 0.089. Выборочные статистики переменной $\widehat{y}_{probit}$ представлены в

**Таблица 12.3**

Dependent Variable: $y$
Method: ML Probit

| Variable | Coefficient | Std. Error | $z$-Statistic | Probability |
|---|---|---|---|---|
| const | $-8.958572$ | 1.678708 | $-5.336588$ | 0.0000 |
| $ltot3$ | 0.918563 | 0.158378 | 5.799833 | 0.0000 |
| $nahm - nch1217 - nch18$ | $-0.128775$ | 0.076024 | $-1.693862$ | 0.0903 |
| $durb$ | 0.044819 | 0.024002 | 1.867325 | 0.0619 |
| Akaike info criterion | 0.989749 | | | |
| Schwarz criterion | 1.027752 | | | |
| McFadden $R^2$ | 0.089014 | | | |

**Таблица 12.4**

| | Mean | Median | Maximum | Minimum | Std. Dev. |
|---|---|---|---|---|---|
| $\widehat{y}_{probit}$ | 0.774654 | 0.790164 | 0.986855 | 0.320371 | 0.128191 |

таблице 12.4. Здесь уже все значения лежат в интервале $(0,1)$, что позволяет интерпретировать их как вероятность того, что семья будет отдыхать.

12.16.3. а) Используем теперь *logit*-модель (см. таблицу 12.5). Знаки коэф-

**Таблица 12.5**

Dependent Variable: $y$
Method: ML Logit

| Variable | Coefficient | Std. Error | $z$-Statistic | Probability |
|---|---|---|---|---|
| const | $-15.65068$ | 2.947123 | $-5.310495$ | 0.0000 |
| $ltot3$ | 1.600157 | 0.280017 | 5.714509 | 0.0000 |
| $nahm - nch1217 - nch18$ | $-0.246143$ | 0.132443 | $-1.858479$ | 0.0631 |
| $durb$ | 0.080120 | 0.041614 | 1.925326 | 0.0542 |
| Akaike info criterion | 0.988510 | | | |
| Schwarz criterion | 1.026512 | | | |
| McFadden $R^2$ | 0.090177 | | | |

фициентов такие же, как и в *probit*-модели. Коэффициенты, которые не были значимыми на 5%-ном уровне, так и остались незначимыми (впрочем, $z$-статистики увеличились по абсолютной величине). Коэффициент детерминации практически не изменился (0.090). Выборочные статистики переменной $\widehat{y}_{logit}$ представлены в таблице 12.6.

б) Качественно все три модели очень похожи. Знаки коэффициентов при одинаковых факторах совпадают. Значимость коэффициентов примерно

Таблица 12.6

| | Mean | Median | Maximum | Minimum | Std. Dev. |
|---|---|---|---|---|---|
| $\widehat{y}_{logit}$ | 0.774654 | 0.790164 | 0.986855 | 0.320371 | 0.128191 |

одинаковая. Так что все три модели можно использовать для качественной оценки влияния включенных факторов на принятие решения о семейном отдыхе.

12.16.4. а) Результаты оценивания *tobit*-модели приведены в таблице 12.7.

Таблица 12.7

Dependent Variable: $\ln(v3+1)$
Method: ML Tobit

| Variable | Coefficient | Std. Error | $z$-Statistic | Probability |
|---|---|---|---|---|
| const | −29.66058 | 4.065465 | −7.295741 | 0.0000 |
| $ltot3$ | 3.203598 | 0.378134 | 8.472128 | 0.0000 |
| $nahm-nch1217-nch18$ | −0.207549 | 0.191153 | −1.085775 | 0.2776 |
| $durb$ | 0.173076 | 0.061070 | 2.834049 | 0.0046 |
| $\widehat{\sigma}$ | 3.568014 | 0.149169 | 23.91921 | 0.0000 |

б) Полученная *tobit*-модель несколько отличается от бинарных моделей, построенных в предыдущих упражнениях. Выявилось, что переменная $nahm-nch1217-nch18$ (число взрослых членов семьи) является незначимой при определении суммы, которую семья тратит на отдых. В то же время оказалось, что зависимость от степени урбанизации высокозначимо положительно влияет на эту сумму (хоть это и Голландия, а город есть город, и хочется из него уехать).

## Задача 12.17

В Великобритании подростки в 16 лет делают важный выбор дальнейшей карьеры. В этом возрасте все они сдают специальный экзамен. Через несколько месяцев после этого они должны решить, продолжать учебу в школе или нет. Те из них, кто решает оставить школу, могут, в свою очередь, работать полный рабочий день или совмещать работу с учебой. В данных упражнениях мы попытаемся выяснить, какие факторы определяют этот выбор.

Здесь используются данные Британского национального опроса (UK National Child and Development Survey). Они содержат информацию о людях, родившихся в Великобритании в марте 1958 г. Детальное описание данных можно найти в статье (Micklewright, 1986)[1]. Данные о респондентах

[1] Micklewright J. (1986). A note on household income data in NCDS 3. *National Child Development Study User Support Group Working Paper 18.*

собирались в разные моменты их жизни. В упражнениях используются данные о юношах и девушках (подвыборка тех, кто живет не в Шотландии). Большинство значений переменных относится к шестнадцатилетнему возрасту. Файл `choice.xls`[2] содержит следующие переменные (см. таблицу 12.8):

**Таблица 12.8**

| Переменная | Описание |
|---|---|
| *at*16 | решение, принятое в 16-летнем возрасте (1 — продолжать учебу, 2 — совмещение учебы и работы, 3 — работать полный рабочий день) |
| *able*7 | результат теста общих способностей, проводящегося в 7-летнем возрасте |
| *loginc* | логарифм дохода семьи (в 16 лет) |
| *ctratio* | число учеников на одного учителя в школе (показатель, отражающий качество школы) |
| *oldsib* | число старших братьев и сестер (в 16 лет) |
| *yngsib* | число младших братьев и сестер (в 16 лет) |
| *etot* | число полученных на выпускных экзаменах высших оценок (экзамен проводится в 16 лет, до принятия решения о продолжении учебы) |
| *female* | пол (1 — для девушек, 0 — для юношей) |

12.17.1. Вычислите описательные статистики переменных как для всей выборки, так и отдельно для девушек и юношей. Постройте гистограммы значений переменной *at*16 отдельно для юношей и девушек. Интерпретируйте результаты.

12.17.2. Сконструируйте переменную *school*, равную 1, если $at16 = 1$, и равную 0 в противном случае. Оцените *logit*-регрессию переменной *school* на все остальные переменные и интерпретируйте результаты. Согласуются ли они с вашими ожиданиями?

12.17.3. Повторите упражнение 12.17.2, используя *probit* вместо *logit*. Сравните полученные результаты с результатами упражнения 12.17.2.

12.17.4. Выберите из *logit*- и *probit*-моделей более подходящую. Проверьте, одинаково ли влияет на принятие решения о продолжении учебы наличие в семье старших и младших братьев и сестер.

---

[2] К исходным данным Британского национального опроса можно получить доступ по адресу `http://www.data-archive.ac.uk/findingData/ncdsNews.asp`

12.17.5. Оцените *logit*- или *probit*-модель раздельно для юношей и для девушек (не забудьте удалить из списка объясняющих переменных переменную $female$). Прокомментируйте различия в результатах оценивания.

12.17.6. Проверьте гипотезу о применимости общей модели для всех подростков против гипотезы о том, что нужно использовать разные модели для юношей и для девушек (используйте тест отношения правдоподобия).

12.17.7. Используя наиболее подходящую, с вашей точки зрения, модель, вычислите прогнозную вероятность продолжить учебу для юноши и девушки со средними по выборке характеристиками. Также вычислите влияние на эту вероятность наличия в семье еще одного младшего брата (сестры).

12.17.8. Рассмотрите тех подростков, для которых $school = 0$, постройте переменную $job$, равную 1, если $at16 = 3$ (полный рабочий день), и 0, если $at16 = 2$ (совмещение работы и учебы). Проделайте упражнения 12.17.1–12.17.7 и проанализируйте, какие факторы влияют на выбор между работой и совмещением работы и учебы.

12.17.9. Используя наиболее подходящие, с вашей точки зрения, модели, вычислите прогнозную вероятность выбрать полный рабочий день ($school = 0$ и $job = 1$) для юноши и девушки со средними по выборке характеристиками. Также вычислите влияние на эту вероятность одного дополнительного младшего брата (сестры).

## Решение

12.17.1. Выборочные статистики переменных для всей выборки (3423 наблюдения), для юношей (1713 наблюдений) и для девушек (1710 наблюдений) представлены в таблицах 12.9, 12.10 и 12.11.

**Таблица 12.9**

| | $at16$ | $able7$ | $loginc$ | $ctratio$ | $oldsib$ | $yngsib$ | $etot$ |
|---|---|---|---|---|---|---|---|
| Mean | 2.069 | 73.934 | 3.863 | 17.119 | 0.427 | 1.204 | 6.128 |
| Median | 2.000 | 80.000 | 3.861 | 17.212 | 0.000 | 1.000 | 7.000 |
| Maximum | 3.000 | 100.000 | 5.106 | 35.286 | 8.000 | 9.000 | 22.000 |
| Minimum | 1.000 | 5.000 | 0.916 | 4.800 | 0.000 | 0.000 | 0.000 |
| Std. Dev. | 0.831 | 20.447 | 0.392 | 1.973 | 0.642 | 1.243 | 3.347 |

Как видим, все статистики для юношей и девушек различаются незначительно.

Гистограммы переменной $at16$ представлены в таблице 12.12. И здесь различия небольшие. Только среди тех, кто решил не продолжать учебу, а по крайней мере совмещать ее с работой, есть отличия — девушки чаще выбирают работать полный рабочий день.

Таблица 12.10

| | *at16* | *able7* | *loginc* | *ctratio* | *oldsib* | *yngsib* | *etot* |
|---|---|---|---|---|---|---|---|
| Mean | 1.990 | 72.351 | 3.865 | 17.026 | 0.429 | 1.214 | 6.029 |
| Median | 2.000 | 80.000 | 3.861 | 17.184 | 0.000 | 1.000 | 7.000 |
| Maximum | 3.000 | 100.000 | 4.959 | 30.000 | 8.000 | 7.000 | 20.000 |
| Minimum | 1.000 | 5.000 | 1.099 | 6.625 | 0.000 | 0.000 | 0.000 |
| Std. Dev. | 0.782 | 21.209 | 0.373 | 2.036 | 0.646 | 1.235 | 3.451 |

Таблица 12.11

| | *at16* | *able7* | *loginc* | *ctratio* | *oldsib* | *yngsib* | *etot* |
|---|---|---|---|---|---|---|---|
| Mean | 2.149 | 75.520 | 3.861 | 17.212 | 0.425 | 1.193 | 6.226 |
| Median | 2.000 | 82.500 | 3.861 | 17.233 | 0.000 | 1.000 | 7.000 |
| Maximum | 3.000 | 100.000 | 5.106 | 35.286 | 4.000 | 9.000 | 22.000 |
| Minimum | 1.000 | 7.500 | 0.916 | 4.800 | 0.000 | 0.000 | 0.000 |
| Std. Dev. | 0.871 | 19.532 | 0.411 | 1.904 | 0.639 | 1.252 | 3.236 |

Таблица 12.12

| Значение | 1 | 2 | 3 |
|---|---|---|---|
| Для юношей | 31.06% | 38.88% | 30.06% |
| Для девушек | 31.58% | 21.99% | 46.43% |

12.17.2. а) Результаты оценивания представлены в таблице 12.13. Как видим, вероятность принять решение остаться в школе зависит положительно от результатов теста общих способностей *able*7, от дохода семьи *loginc*, от числа полученных высших оценок на выпускных экзаменах *etot*. Отрицательно зависит от числа учеников на одного учителя в школе *ctratio*, от числа старших братьев и сестер *yngsib*. Не удалось выявить значимого влияния на решение остаться в школе таких факторов, как пол (*female*) и число младших братьев и сестер (*yngsib*). Результаты не противоречат интуиции. Скажем, отрицательное влияние числа старших братьев и сестер можно объяснить финансовыми соображениями (если кто-то из старших братьев и сестер учится, то на обучение младших остается меньше средств), в то же время младшие братья и сестры при решении старших не имеют такого влияния. Довольно интересно, что такая простая мера качества школы, как число учеников на одного учителя, оказалась значимо влияющей на вероятность остаться в школе.

Таблица 12.13

Dependent Variable: *school*
Method: ML Logit

| Variable | Coefficient | Std. Error | $z$-Statistic | Probability |
|---|---|---|---|---|
| const | −3.795433 | 0.684000 | −5.548876 | 0.0000 |
| *able7* | 0.040774 | 0.003158 | 12.91104 | 0.0000 |
| *loginc* | 0.676020 | 0.117480 | 5.754340 | 0.0000 |
| *ctratio* | −0.242808 | 0.024175 | −10.04365 | 0.0000 |
| *oldsib* | −0.293574 | 0.074398 | −3.945997 | 0.0001 |
| *yngsib* | −0.063416 | 0.037922 | −1.672255 | 0.0945 |
| *etot* | 0.221943 | 0.017021 | 13.03975 | 0.0000 |
| *female* | −0.059071 | 0.085801 | −0.688462 | 0.4912 |
| Loglikelihood | −1645.635 | | | |
| McFadden $R^2$ | 0.226602 | | | |

12.17.3. Результаты оценивания представлены в таблице 12.14. Знаки при

Таблица 12.14

Dependent Variable: *school*
Method: ML Probit

| Variable | Coefficient | Std. Error | $z$-Statistic | Probability |
|---|---|---|---|---|
| const | −2.060609 | 0.383825 | −5.368613 | 0.0000 |
| *able7* | 0.022505 | 0.001681 | 13.38445 | 0.0000 |
| *loginc* | 0.397010 | 0.066449 | 5.974638 | 0.0000 |
| *ctratio* | −0.142332 | 0.013742 | −10.35725 | 0.0000 |
| *oldsib* | −0.177661 | 0.043045 | −4.127283 | 0.0000 |
| *yngsib* | −0.040403 | 0.022099 | −1.828262 | 0.0675 |
| *etot* | 0.124034 | 0.009195 | 13.48882 | 0.0000 |
| *female* | −0.032854 | 0.050439 | −0.651367 | 0.5148 |
| Loglikelihood | −1651.797 | | | |
| McFadden $R^2$ | 0.223706 | | | |

всех переменных такие же, как и в *logit*-модели. Значимость факторов тоже не изменилась.

Построенные модели почти не отличаются, однако *logit*-модель чуть точнее соответствует данным (McFadden $R^2$ чуть больше, или, что то же самое, функция правдоподобия), так что предпочтем ее.

12.17.4. Чтобы проверить, одинаково ли влияет на принятие решения о продолжении учебы наличие старших и младших братьев и сестер, сравним коэффициенты при факторах *oldsib* и *yngsib* (тем самым тестируя гипотезу об одинаковости влияния каждого дополнительного младшего и старшего брата или сестры). Получаем, что $F$-статистика равна 7.945 ($p$-значение 0.005), и гипотеза о том, что старшие и младшие братья и сестры влияют одинаково, отвергается на 5%-ном уровне значимости.

12.17.5. Результаты оценивания *logit*-модели раздельно для юношей и для девушек представлены в таблицах 12.15 (для юношей) и 12.16 (для девушек).

**Таблица 12.15**

Dependent Variable: *school*
Method: ML Logit
Sampling condition: $female = 0$

| Variable | Coefficient | Std. Error | $z$-Statistic | Probability |
|---|---|---|---|---|
| const | −3.065706 | 0.985409 | −3.111100 | 0.0019 |
| *able7* | 0.045425 | 0.004445 | 10.21830 | 0.0000 |
| *loginc* | 0.595668 | 0.177051 | 3.364380 | 0.0008 |
| *ctratio* | −0.293860 | 0.034270 | −8.574772 | 0.0000 |
| *oldsib* | −0.316030 | 0.111492 | −2.834540 | 0.0046 |
| *yngsib* | −0.109785 | 0.056572 | −1.940646 | 0.0523 |
| *etot* | 0.238720 | 0.024122 | 9.896502 | 0.0000 |
| Loglikelihood | −773.3956 | | | |
| McFadden $R^2$ | 0.271271 | | | |

**Таблица 12.16**

Dependent Variable: *school*
Method: ML Logit
Sampling condition: $female = 1$

| Variable | Coefficient | Std. Error | $z$-Statistic | Probability |
|---|---|---|---|---|
| const | −4.559795 | 0.969843 | −4.701582 | 0.0000 |
| *able7* | 0.036308 | 0.004544 | 7.989930 | 0.0000 |
| *loginc* | 0.748078 | 0.158486 | 4.720162 | 0.0000 |
| *ctratio* | −0.191555 | 0.034415 | −5.566000 | 0.0000 |
| *oldsib* | −0.280935 | 0.100070 | −2.807376 | 0.0050 |
| *yngsib* | −0.022338 | 0.050987 | −0.438119 | 0.6613 |
| *etot* | 0.204467 | 0.024207 | 8.446723 | 0.0000 |
| Loglikelihood | −867.4097 | | | |
| McFadden $R^2$ | 0.186638 | | | |

Сравнивая две модели, заметим, что существенных различий между ними нет. Знаки коэффициентов одинаковые. Значимые факторы в обеих моделях совпадают.

12.17.6. Протестируем гипотезу о применимости общей модели (из упражнения 12.17.2) для всех подростков против альтернативной гипотезы о том, что модели должны быть разные (модели из упражнения 12.17.5). Для этого воспользуемся тестом отношения правдоподобия. Составим статистику $\mathrm{LR} = -2(\ln L_{\mathrm{R}} - \ln L_{\mathrm{UR}})$. В нашем случае регрессия с ограничением — это регрессия из упражнения 12.17.2, и $\ln L_{\mathrm{R}} = -1645.635$. Регрессия без ограничения — это две модели раздельно для юношей и девушек. Поскольку множества наблюдений, для которых построены модели, не пересекаются, то логарифмическая функция правдоподобия аддитивна, и $\ln L_{\mathrm{UR}} = \ln L_1 + \ln L_2 = -773.3956 - 867.4097 = -1640.805$. Искомая статистика равна

$$\mathrm{LR} = -2(\ln L_{\mathrm{R}} - \ln L_{\mathrm{UR}}) = -2(-1645.635 + 1640.805) = 9.66.$$

Число ограничений равно 6 (все факторы, кроме постоянной, поскольку в модель с ограничением включена переменная $female$), поэтому при условии выполнения нулевой гипотезы статистика LR подчиняется $\chi^2$-распределению с 6 степенями свободы. Значит, $p$-значение равно 0.140. Следовательно, гипотеза о совпадении моделей не отвергается.

12.17.7. Удалим из модели упражнения 12.17.2 незначимые факторы $yngsib$ и $female$. Получим модель, представленную в таблице 12.17.

Dependent Variable: *school*
Method: ML Logit

**Таблица 12.17**

| Variable | Coefficient | Std. Error | $z$-Statistic | Probability |
|---|---|---|---|---|
| const | −3.883971 | 0.681699 | −5.697485 | 0.0000 |
| *able7* | 0.040696 | 0.003148 | 12.92682 | 0.0000 |
| *loginc* | 0.674617 | 0.117520 | 5.740439 | 0.0000 |
| *ctratio* | −0.243626 | 0.024113 | −10.10339 | 0.0000 |
| *oldsib* | −0.287406 | 0.074063 | −3.880571 | 0.0001 |
| *etot* | 0.223452 | 0.016997 | 13.14641 | 0.0000 |
| Loglikelihood | −1647.274 | | | |
| McFadden $R^2$ | 0.225832 | | | |

Средние по выборке характеристики возьмем из таблиц 12.10 и 12.11. Подставив их в модель, получим следующие значения прогнозной вероятности остаться в школе для юноши и девушки со средними по выборке характеристиками:

$$\mathrm{P}(school = 1 \text{ для юноши}) = 0.2216,$$
$$\mathrm{P}(school = 1 \text{ для девушки}) = 0.2442.$$

Заметим, что число младших братьев и сестер в построенной модели не влияет на вероятность остаться в школе. Поэтому при появлении в семье еще одного младшего брата (сестры) прогнозная вероятность остаться в школе не изменится.

12.17.8. Для оценивания влияния различных факторов на выбор между работой и совмещением работы и учебы тоже рассматриваем *logit*-модель (*probit*-модель не приводим). Кроме того, удалим из модели незначимую переменную *ctratio* (по-видимому, она не оказывает влияния на выбор между работой и совмещением). Результаты оценивания совместной модели (с включенной фиктивной переменной *female*) и моделей отдельно для юношей и для девушек представлены в таблицах 12.18–12.20. Главное отличие

**Таблица 12.18**

Dependent Variable: *job*
Method: ML Logit
Sampling condition: *school* = 0

| Variable | Coefficient | Std. Error | $z$-Statistic | Probability |
|---|---|---|---|---|
| const | 2.210365 | 0.491103 | 4.500816 | 0.0000 |
| *able7* | −0.010737 | 0.002381 | −4.510336 | 0.0000 |
| *loginc* | −0.354837 | 0.124163 | −2.857831 | 0.0043 |
| *oldsib* | 0.189099 | 0.070284 | 2.690496 | 0.0071 |
| *yngsib* | 0.108994 | 0.035710 | 3.052229 | 0.0023 |
| *etot* | −0.125589 | 0.015322 | −8.196362 | 0.0000 |
| *female* | 1.193640 | 0.092118 | 12.95780 | 0.0000 |
| Loglikelihood | −1446.224 | | | |
| McFadden $R^2$ | 0.104169 | | | |

от моделей выбора остаться или нет в школе в том, что теперь коэффициент при переменной *female* положительный и значимый, т. е. если девушка и юноша решили не оставаться в школе, то при прочих равных девушка с большей вероятностью совсем оставит обучение. Кроме того, наличие младших и старших братьев и сестер теперь положительно влияет на вероятность работать полный рабочий день (раньше наличие младших было незначимо).

Значимость переменной *female* означает, что нет нужды проверять формально совпадение отдельных моделей для юношей и девушек. Они не совпадают. Проверим с помощью LR-теста, что существенных различий в коэффициентах при других переменных нет.

Действуя аналогично упражнению 12.17.6, получаем значение LR-статистики:

$$\mathrm{LR} = -2(\ln L_{\mathrm{R}} - \ln L_{\mathrm{UR}}) = -2(-1446.224 + (763.9121 + 680.0228)) = 4.5782.$$

Таблица 12.19

Dependent Variable: *job*
Method: ML Logit
Sampling condition: $school = 0$ AND $female = 0$

| Variable | Coefficient | Std. Error | $z$-Statistic | Probability |
|---|---|---|---|---|
| const | 2.002855 | 0.682492 | 2.934619 | 0.0033 |
| *able7* | −0.008529 | 0.003026 | −2.818999 | 0.0048 |
| *loginc* | −0.316510 | 0.173679 | −1.822382 | 0.0684 |
| *oldsib* | 0.125716 | 0.092978 | 1.352107 | 0.1763 |
| *yngsib* | 0.070364 | 0.047197 | 1.490852 | 0.1360 |
| *etot* | −0.126627 | 0.020262 | −6.249380 | 0.0000 |
| Loglikelihood | −763.9121 | | | |
| McFadden $R^2$ | 0.055648 | | | |

Таблица 12.20

Dependent Variable: *job*
Method: ML Logit
Sampling condition: $school = 0$ AND $female = 1$

| Variable | Coefficient | Std. Error | $z$-Statistic | Probability |
|---|---|---|---|---|
| const | 3.756614 | 0.721343 | 5.207807 | 0.0000 |
| *able7* | −0.014371 | 0.003913 | −3.673163 | 0.0002 |
| *loginc* | −0.402951 | 0.178572 | −2.256510 | 0.0240 |
| *oldsib* | 0.271090 | 0.107808 | 2.514558 | 0.0119 |
| *yngsib* | 0.160820 | 0.055712 | 2.886635 | 0.0039 |
| *etot* | −0.123439 | 0.023498 | −5.253233 | 0.0000 |
| Loglikelihood | −680.0228 | | | |
| McFadden $R^2$ | 0.074343 | | | |

Эта статистика при условии выполнения нулевой гипотезы подчиняется $\chi^2$-распределению с 5 степенями свободы. Значит, $p$-значение равно 0.469. Следовательно, гипотеза о том, что для различения моделей достаточно ввести только переменную $female$, не отвергается.

Теперь вычислим прогнозную вероятность выбрать полный рабочий день для юноши и девушки со средними по выборке характеристиками. Средние характеристики по выборке возьмем из таблиц 12.10 и 12.11. Подставив их в модель (см. таблицу 12.18), получим:

$$\mathrm{P}(job = 1 \text{ для юноши}) = 0.3819,$$
$$\mathrm{P}(job = 1 \text{ для девушки}) = 0.6573.$$

При увеличении числа младших братьев или сестер на 1 вероятности выбрать полный рабочий день для юноши и для девушки станут равны 0.4079 и 0.6815 соответственно.

12.17.9. Представим процедуру выбора полного рабочего дня как последовательность бинарных выборов (подробнее см. разд. 12.1, стр. 293). Сначала происходит выбор в рамках модели из таблицы 12.17, потом если выбрали не продолжать учебу, то происходит выбор в рамках модели из таблицы 12.18.

Тем самым прогнозная вероятность выбрать полный рабочий день равна

$$\mathrm{P}(job = 1, school = 0) = \mathrm{P}(job = 1 \mid school = 0)\mathrm{P}(school = 0).$$

Подставив значения факторов в обе модели, получаем, что для юноши и девушки со средними характеристиками прогнозные вероятности выбрать полный рабочий день равны $0.3819 \cdot (1 - 0.2216) = 0.2973$ и $0.6573 \cdot (1 - 0.2442) = 0.4968$ соответственно. При появлении одного дополнительного младшего брата или сестры прогнозные вероятности выбрать полный рабочий день становятся равны 0.3175 и 0.5151 соответственно.

# Глава 13

# Панельные данные

## Задача 13.1

Докажите равенства (13.17)

$$\Sigma^{-1} = \frac{1}{\sigma_\varepsilon^2}\left[I_T - \frac{\sigma_u^2}{\sigma_\varepsilon^2 + T\sigma_u^2}\imath_T\imath_T'\right]$$

и (13.18)

$$\Sigma^{-1} = \frac{1}{\sigma_\varepsilon^2}\left[\left(I_T - \frac{1}{T}\imath_T\imath_T'\right) + \theta^2\frac{1}{T}\imath_T\imath_T'\right].$$

## Решение

1) Имеем

$$\begin{aligned}\Sigma\Sigma^{-1} &= (\sigma_u^2\imath_T\imath_T' + \sigma_\varepsilon^2 I_T)\frac{1}{\sigma_\varepsilon^2}\left(I_T - \frac{\sigma_u^2}{\sigma_\varepsilon^2 + T\sigma_u^2}\imath_T\imath_T'\right) \\ &= \frac{1}{\sigma_\varepsilon^2}\left(\sigma_\varepsilon^2 I_T + \left(\sigma_u^2 - \frac{T\sigma_u^4}{\sigma_\varepsilon^2 + T\sigma_u^2} - \frac{\sigma_\varepsilon^2\sigma_u^2}{\sigma_\varepsilon^2 + T\sigma_u^2}\right)\imath_T\imath_T'\right) = I_T.\end{aligned}$$

2) Для доказательства формулы (13.18) достаточно проверить равенство

$$\frac{1}{T}\frac{\sigma_\varepsilon^2}{\sigma_\varepsilon^2 + T\sigma_u^2} - \frac{1}{T} = -\frac{\sigma_u^2}{\sigma_\varepsilon^2 + T\sigma_u^2},$$

что делается с помощью элементарной арифметики.

## Задача 13.2

Докажите равенство (13.22)

$$\widehat{\beta}_{\mathrm{RE}} = W\widehat{\beta}_{\mathrm{B}} + (I_k - W)\widehat{\beta}_{\mathrm{W}} = W\widehat{\beta}_{\mathrm{B}} + (I_k - W)\widehat{\beta}_{\mathrm{FE}}.$$

## Решение

(Ср. (Greene, 1997, Ch. 14.3).) Сформулируем предложение, которое вытекает из более общего результата, доказанного в упражнении 4.3.

*Предложение.* Пусть задана классическая линейная модель регрессии со свободным членом

$$y_t = \alpha + x_t'\beta + \varepsilon_t, \quad \mathrm{V}(\varepsilon_t) = \sigma_\varepsilon^2, \quad \mathrm{Cov}(\varepsilon_t, \varepsilon_s) = 0, \quad t \neq s, \quad t, s = 1, \ldots, n.$$

Тогда

$$\widehat{\beta}_{\mathrm{OLS}} = \left(\sum_{t=1}^{n}(x_t - \overline{x})(x_t - \overline{x})'\right)^{-1} \sum_{t=1}^{n}(x_t - \overline{x})(y_t - \overline{y}),$$

$$\mathrm{V}(\widehat{\beta}_{\mathrm{OLS}}) = \sigma_\varepsilon^2 \left(\sum_{t=1}^{n}(x_t - \overline{x})(x_t - \overline{x})'\right)^{-1},$$

где

$$\overline{x} = \frac{1}{n}\sum_{t=1}^{n} x_t, \quad \overline{y} = \frac{1}{n}\sum_{t=1}^{n} y_t.$$

Оценка $\widehat{\beta}_{\mathrm{RE}}$ может быть получена как обычная МНК-оценка преобразованной модели (13.24):

$$y_{it} - (1-\theta)\overline{y}_i = (1-\theta)\mu + (x_{it} - (1-\theta)\overline{x}_i)'\beta + \eta_{it}, \qquad (*)$$
$$i = 1, \ldots, n, \ t = 1, \ldots, T,$$

где

$$\overline{x}_i = \frac{1}{T}\sum_{t=1}^{T} x_{it}, \quad \overline{y}_i = \frac{1}{T}\sum_{t=1}^{T} y_{it}, \quad \theta^2 = \frac{\sigma_\varepsilon^2}{\sigma_\varepsilon^2 + T\sigma_u^2}.$$

В модели $(*)$ ошибки $\eta_{it}$ не коррелированы и имеют постоянную дисперсию $\mathrm{V}(\eta_{it}) = \sigma_\varepsilon^2$.

Из сформулированного выше предложения следует, что МНК-оценка вектора $\beta$ в уравнении $(*)$ может быть получена применением МНК к этому уравнению, записанному в отклонениях от глобальных средних:

$$(y_{it} - \overline{y}_i) + \theta(\overline{y}_i - \overline{y}) = ((x_{it} - \overline{x}_i) + \theta(\overline{x}_i - \overline{x}))'\,\beta + \eta_{it} - \overline{\eta}, \qquad (**)$$

где

$$\overline{x} = \frac{1}{nT}\sum_{i=1}^{n}\sum_{t=1}^{T} x_{it}, \quad \overline{y} = \frac{1}{nT}\sum_{i=1}^{n}\sum_{t=1}^{T} y_{it}.$$

Обозначим

$$\Sigma^w_{xx} = \sum_{i=1}^{n}\sum_{t=1}^{T}(x_{it} - \overline{x}_i)(x_{it} - \overline{x}_i)', \qquad \Sigma^w_{xy} = \sum_{i=1}^{n}\sum_{t=1}^{T}(x_{it} - \overline{x}_i)(y_{it} - \overline{y}_i),$$

$$\Sigma^b_{xx} = \sum_{i=1}^{n} T(\overline{x}_i - \overline{x})(\overline{x}_i - \overline{x})', \qquad \Sigma^b_{xy} = \sum_{i=1}^{n} T(\overline{x}_i - \overline{x})(\overline{y}_i - \overline{y}).$$

Для уравнения $(**)$ соответствующая матрица $X'X$ есть

$$\sum_{i=1}^{n}\sum_{t=1}^{T}\left((x_{it} - \overline{x}_i) + \theta(\overline{x}_i - \overline{x})\right)\left((x_{it} - \overline{x}_i) + \theta(\overline{x}_i - \overline{x})\right)' = \Sigma^w_{xx} + \theta^2\Sigma^b_{xx},$$

а соответствующий вектор $X'y$ –

$$\sum_{i=1}^{n}\sum_{t=1}^{T}\left((x_{it} - \overline{x}_i) + \theta(\overline{x}_i - \overline{x})\right)\left((y_{it} - \overline{y}_i) + \theta(\overline{y}_i - \overline{y})\right) = \Sigma^w_{xy} + \theta^2\Sigma^b_{xy}.$$

Таким образом,

$$\widehat{\beta}_{\mathrm{RE}} = \left(\Sigma^w_{xx} + \theta^2\Sigma^b_{xx}\right)^{-1}\left(\Sigma^w_{xy} + \theta^2\Sigma^b_{xy}\right). \qquad (***)$$

Поскольку

$$\widehat{\beta}_{\mathrm{W}} = \left(\Sigma^w_{xx}\right)^{-1}\Sigma^w_{xy}, \quad \widehat{\beta}_{\mathrm{B}} = \left(\Sigma^b_{xx}\right)^{-1}\Sigma^b_{xy},$$

то

$$\Sigma^w_{xy} = \Sigma^w_{xx}\widehat{\beta}_{\mathrm{W}}, \quad \Sigma^b_{xy} = \Sigma^b_{xx}\widehat{\beta}_{\mathrm{B}}.$$

Подставляя эти выражения в $(***)$, получаем:

$$\widehat{\beta}_{\mathrm{RE}} = \left(\Sigma^w_{xx} + \theta^2\Sigma^b_{xx}\right)^{-1}\left(\Sigma^w_{xx}\widehat{\beta}_{\mathrm{W}} + \theta^2\Sigma^b_{xx}\widehat{\beta}_{\mathrm{B}}\right) = W_1\widehat{\beta}_{\mathrm{W}} + W_2\widehat{\beta}_{\mathrm{B}},$$

где

$$W_1 = \left(\Sigma^w_{xx} + \theta^2\Sigma^b_{xx}\right)^{-1}\Sigma^w_{xx}, \quad W_2 = \left(\Sigma^w_{xx} + \theta^2\Sigma^b_{xx}\right)^{-1}\theta^2\Sigma^b_{xx},$$

и

$$W_1 + W_2 = I.$$

## Задача 13.3

Покажите, что существует $k \times k$ матрица $S$, такая, что

$$\widehat{\beta}_{\mathrm{OLS}} = S\widehat{\beta}_{\mathrm{B}} + (I_k - S)\widehat{\beta}_{\mathrm{W}},$$

где $\widehat{\beta}_{\mathrm{OLS}}$ — МНК-оценка вектора параметров $\beta$ в объединенной регрессии (13.2).

## Решение

Оценка $\widehat{\beta}_{\mathrm{OLS}}$ является частным случаем оценки $\widehat{\beta}_{\mathrm{RE}}$, когда $\sigma_u^2 = 0$, т.е. когда $\theta = 1$. Поэтому из результатов упражнения 13.2 следует, что

$$\boldsymbol{S} = \left(\boldsymbol{\Sigma}_{xx}^w + \boldsymbol{\Sigma}_{xx}^b\right)^{-1} \boldsymbol{\Sigma}_{xx}^b$$

(мы сохранили обозначения упражнения 13.2).

## Задача 13.4

Докажите равенство (13.23)

$$\boldsymbol{P} = \frac{1}{\sigma_\varepsilon}\left(\boldsymbol{I}_T - \frac{1-\theta}{T}\boldsymbol{\imath}_T\boldsymbol{\imath}_T'\right).$$

## Решение

Необходимо показать, что $\boldsymbol{P}'\boldsymbol{P} = \boldsymbol{\Sigma}^{-1}$. Имеем

$$\begin{aligned}\boldsymbol{P}'\boldsymbol{P} &= \frac{1}{\sigma_\varepsilon^2}\left(\boldsymbol{I}_T - \frac{2(1-\theta)}{T}\boldsymbol{\imath}_T\boldsymbol{\imath}_T' + \frac{(1-\theta)^2}{T^2}T\boldsymbol{\imath}_T\boldsymbol{\imath}_T'\right) \\ &= \frac{1}{\sigma_\varepsilon^2}\left(\boldsymbol{I}_T - \frac{1}{T}\boldsymbol{\imath}_T\boldsymbol{\imath}_T' + \frac{\theta^2}{T}\boldsymbol{\imath}_T\boldsymbol{\imath}_T'\right),\end{aligned}$$

что в силу (13.18) совпадает с $\boldsymbol{\Sigma}^{-1}$.

## Задача 13.5

Докажите равенство (13.25)

$$\mathrm{V}(\widehat{\beta}_{\mathrm{RE}}) = \sigma_\varepsilon^2\left(\sum_{i=1}^n\sum_{t=1}^T(\boldsymbol{x}_{it} - \overline{\boldsymbol{x}}_i)(\boldsymbol{x}_{it} - \overline{\boldsymbol{x}}_i)' + \theta^2 T\sum_{i=1}^n(\overline{\boldsymbol{x}}_i - \overline{\boldsymbol{x}})(\overline{\boldsymbol{x}}_i - \overline{\boldsymbol{x}})'\right)^{-1}.$$

## Решение

Требуемое равенство непосредственно следует из решения упражнения 13.2 и сформулированного в этом упражнении предложения.

## Задача 13.6

Докажите, что решение задачи (13.42)

$$\left(\frac{1}{n}\sum_{i=1}^n \boldsymbol{Z}_i'(\Delta\boldsymbol{y}_i - \gamma\Delta\boldsymbol{y}_i(-1))\right)'\boldsymbol{S}\left(\frac{1}{n}\sum_{i=1}^n \boldsymbol{Z}_i'(\Delta\boldsymbol{y}_i - \gamma\Delta\boldsymbol{y}_i(-1))\right) \to \min,$$

задается равенством (13.43)

$$\widehat{\gamma}_{\mathrm{GMM}} = \left(\left(\sum_{i=1}^{n} \Delta y_i'(-1) Z_i\right) S \left(\sum_{i=1}^{n} Z_i' \Delta y_i(-1)\right)\right)^{-1} \times \left(\left(\sum_{i=1}^{n} \Delta y_i'(-1) Z_i\right) S \left(\sum_{i=1}^{n} Z_i' \Delta y_i\right)\right).$$

## Решение

Обозначим для краткости

$$\frac{1}{n}\sum_{i=1}^{n} Z_i' \Delta y_i = a, \quad \frac{1}{n}\sum_{i=1}^{n} Z_i' \Delta y_i(-1)) = b.$$

Тогда левую часть формулы (13.42) можно представить в следующем виде:

$$f(\gamma) = (a - \gamma b)' S (a - \gamma b) = a' S a - 2\gamma b' S a + \gamma^2 b' S b.$$

Минимум этой функции (по $\gamma$) достигается в точке

$$\widehat{\gamma}_{\min} = \frac{b' S a}{b' S b},$$

что совпадает с формулой (13.43).

## Задача 13.7

Постройте функцию правдоподобия для модели (13.47), (13.48)

$$y_{it}^* = x_{it}'\beta + \alpha_i + \varepsilon_{it}, \quad \text{и} \quad y_{it} = \begin{cases} 1, & \text{если } y_{it}^* \geqslant 0, \\ 0, & \text{если } y_{it}^* < 0. \end{cases}$$

## Решение

Предполагая, что ошибки $\varepsilon_{it}$ независимы и имеют одинаковое симметричное распределение, по аналогии с обычными моделями бинарного выбора (см. (12.4) и (12.5)) получаем следующее выражение для функции правдоподобия:

$$L(\alpha_1, \ldots, \alpha_n, \beta) = \prod_{i=1}^{n}\prod_{t=1}^{T} \left[F(x_{it}'\beta + \alpha_i)\right]^{y_{it}} \left[1 - F(x_{it}'\beta + \alpha_i)\right]^{1 - y_{it}},$$

где $F(\cdot)$ — функция распределения ошибок $\varepsilon_{it}$.

## Задача 13.8

Рассмотрим модель
$$y_{it} = \gamma y_{it-1} + \alpha_i + \varepsilon_{it},$$
где ошибки $\varepsilon_{it} \sim iid(0, \sigma^2)$ и независимы с $y_{js}$ при $s < t$. В уравнении внутригрупповой регрессии (13.35)
$$y_{it} - \overline{y}_i = \gamma(y_{it-1} - \overline{y}_{i,-1}) + \varepsilon_{it} - \overline{\varepsilon}_i$$
ошибки коррелированы с регрессорами. Покажите, что МНК-оценка $\widehat{\gamma}$ в последнем уравнении несостоятельна (например, покажите, что при $\gamma = 0$ получаем $p \lim\limits_{n\to\infty} \widehat{\gamma} \neq 0$).

## Решение

Применяя формулу для простейшей регрессии с одной объясняющей переменной, получаем:
$$\widehat{\gamma} = \gamma + \frac{\sum_{i=1}^n \sum_{t=1}^T (y_{it-1} - \overline{y}_{i,-1})(\varepsilon_{it} - \overline{\varepsilon}_i)}{\sum_{i=1}^n \sum_{t=1}^T (y_{it-1} - \overline{y}_{i,-1})^2}.$$

Если $\gamma = 0$, то $y_{it-1} - \overline{y}_{i,-1} = \varepsilon_{it-1} - \overline{\varepsilon}_{i,-1}$. Таким образом, смещение $\Delta\gamma$ оценки $\widehat{\gamma}$ есть
$$\Delta\gamma = \frac{\sum_{i=1}^n \sum_{t=1}^T (\varepsilon_{it-1} - \overline{\varepsilon}_{i,-1})(\varepsilon_{it} - \overline{\varepsilon}_i)}{\sum_{i=1}^n \sum_{t=1}^T (\varepsilon_{it-1} - \overline{\varepsilon}_{i,-1})^2},$$
где
$$\overline{\varepsilon}_{i,-1} = \frac{1}{T}(\varepsilon_0 + \varepsilon_1 + \dots + \varepsilon_{T-1}), \quad \overline{\varepsilon}_i = \frac{1}{T}(\varepsilon_1 + \varepsilon_2 + \dots + \varepsilon_T).$$

Мы предполагаем, что ошибки $\varepsilon_{it}$ независимы и одинаково распределены. Тогда с помощью прямых вычислений получаем:
$$\mathrm{E}(\overline{\varepsilon}_{i,-1}\varepsilon_{it}) = \begin{cases} \dfrac{\sigma_\varepsilon^2}{T}, & \text{если } t \leqslant T-1, \\ 0, & \text{если } t = T, \end{cases} \qquad \mathrm{E}(\overline{\varepsilon}_i\varepsilon_{it-1}) = \begin{cases} \dfrac{\sigma_\varepsilon^2}{T}, & \text{если } t \geqslant 2, \\ 0, & \text{если } t = 1, \end{cases}$$
$$\mathrm{E}(\overline{\varepsilon}_i\overline{\varepsilon}_{i,-1}) = \frac{(T-1)\sigma_\varepsilon^2}{T^2}, \qquad \mathrm{E}(\varepsilon_{it-1} - \overline{\varepsilon}_{i,-1})^2 = \left(1 - \frac{1}{T}\right)\sigma_\varepsilon^2.$$

Обозначим
$$\xi_i = \sum_{t=1}^T (\varepsilon_{it-1} - \overline{\varepsilon}_{i,-1})(\varepsilon_{it} - \overline{\varepsilon}_i), \quad \eta_i = \sum_{t=1}^T (\varepsilon_{it-1} - \overline{\varepsilon}_{i,-1})^2.$$

Случайные величины $\{\xi_i\}$ независимы и одинаково распределены, то же самое справедливо относительно $\{\eta_i\}$. Из предыдущих формул следует, что
$$\mathrm{E}(\xi_i) = -\frac{\sigma_\varepsilon^2(T-1)}{T}, \quad \mathrm{E}(\eta_i) = \sigma_\varepsilon^2(T-1).$$

Следовательно,

$$p\lim_{n\to\infty}\Delta\gamma = p\lim_{n\to\infty}\frac{\sum_{i=1}^{n}\xi_i}{\sum_{i=1}^{n}\eta_i} = p\lim_{n\to\infty}\frac{\frac{1}{n}\sum_{i=1}^{n}\xi_i}{\frac{1}{n}\sum_{i=1}^{n}\eta_i} = -\frac{1}{T}$$

в силу закона больших чисел. Таким образом, оценка $\widehat{\gamma}$ имеет асимптотическое смещение.

## Задача 13.9

Оценка производственной функции российских предприятий топливно-энергетического комплекса.

В файле `fuel.xls` содержатся ежегодные данные об объемах выпуска, трудозатратах, капитальных вложениях российских предприятий топливно-энергетического комплекса за период 1993–2000 гг. (Е. В. Бессонова, ЦЭФИР). В панель включено около 2400 предприятий, панель не сбалансирована (см. таблицу 13.1).

Цель примера — оценить производственную функцию предприятий.

**Таблица 13.1**

| Переменная | Описание |
|---|---|
| *okpo* | номер предприятия по классификации ОКПО |
| *okonh* | код отрасли ОКОНХ |
| *year* | год |
| *rout* | реальный выпуск |
| *emp* | численность работников |
| *wor* | промышленно-производственный персонал |
| *rk* | реальные капиталовложения |

13.9.1. Вычислите описательные статистики основных переменных.

13.9.2. Оцените производственную функцию Кобба–Дугласа с помощью простой полной регрессии. Выполняется ли условие постоянства отдачи на масштаб?

13.9.3. Повторите упражнение 13.9.2 для регрессий с фиксированным и случайным эффектами. Сравните результаты.

13.9.4. Является ли влияние индивидуальных эффектов существенным? Проверьте гипотезы:

- простая регрессия против регрессии с фиксированным эффектом;
- простая регрессия против регрессии со случайным эффектом;
- регрессия со случайным эффектом против регрессии с фиксированным эффектом.

13.9.5. Повторите предыдущие упражнения для более сложной модели производственной функции путем включения квадратичных и перекрестных членов. Выберите наиболее адекватную, с вашей точки зрения, модель.

## Решение

13.9.1. Описательные статистики основных переменных представлены в таблице 13.2.

**Таблица 13.2**

| | *emp* | *wor* | *rout* | *rk* |
|---|---|---|---|---|
| Mean | 1116.51 | 886.14 | 3369.06 | 5488.18 |
| Maximum | 53218 | 46380 | 531620 | 1139876 |
| Minimum | 1 | 0 | −1.88 | 0.004 |
| Std. Dev. | 3045.87 | 2463.59 | 16164.16 | 24403.61 |

Если более детально проанализировать данные, то можно заметить, что распределение всех переменных сильно скошено влево: большинство наблюдений сосредоточено в области средних значений и имеется относительно небольшое число сравнительно больших выбросов.

13.9.2. Рассмотрим производственную функцию Кобба–Дугласа: $Output = A \cdot (Capital)^{\alpha} \cdot (Labor)^{\beta}$, полагая $Output = rout$, $Capital = rk$, $Labor = emp$. Для оценивания эластичностей $\alpha$ и $\beta$ перейдем к логарифмам и проведем простую регрессию $\ln rout$ на $\ln rk$, $\ln emp$ и константу. Результаты представлены в таблице 13.3. Условие постоянства отдачи на масштаб означает

**Таблица 13.3.** Простая регрессия

Dependent Variable: $\ln rout$

| Variable | Coefficient | Std. Error | $t$-Statistic | Probability |
|---|---|---|---|---|
| const | −2.488 | 0.080 | −31.08 | 0.0000 |
| $\ln rk$ | 0.330 | 0.016 | 20.15 | 0.0000 |
| $\ln emp$ | 0.928 | 0.024 | 37.66 | 0.0000 |
| $R^2$ | 0.5804 | | | |

выполнение равенства $\alpha + \beta = 1$. Проверяя эту гипотезу с помощью стандартного $F$-теста, получаем значение соответствующей статистики, равное 359.00, что позволяет уверенно отвергнуть гипотезу о постоянстве отдачи на масштаб.

Заменяя переменную $\ln emp$ на $\ln wor$, получаем несколько иные результаты (см. таблицу13.4). Однако содержательно эти две регрессии мало от-

**Таблица 13.4.** Простая регрессия

Dependent Variable: $\ln rout$

| Variable | Coefficient | Std. Error | $t$-Statistic | Probability |
|---|---|---|---|---|
| const | −2.054 | 0.074 | −27.67 | 0.0000 |
| $\ln rk$ | 0.391 | 0.015 | 24.78 | 0.0000 |
| $\ln wor$ | 0.825 | 0.023 | 35.45 | 0.0000 |
| $R^2$ | 0.5767 | | | |

личаются. Во второй регрессии гипотеза постоянства отдачи на масштаб отвергается столь же уверенно, как и в первой, — значение соответствующей статистики равно 270.40.

13.9.3. Результаты оценивания модели с фиксированным эффектом приведены в таблице 13.5, модели со случайным эффектом — в таблице 13.6. В этой и следующей задачах при оценивании модели с фиксированным эффектом используется нормировка $\sum_{i=1}^{n} \alpha_i = 0$, т. е. сумма индивидуальных эффектов равна нулю, поэтому в регрессии присутствует константа.

**Таблица 13.5.** Фиксированный эффект

Dependent Variable: $\ln rout$

| Variable | Coefficient | Std. Error | $t$-Statistic | Probability |
|---|---|---|---|---|
| const | 0.759 | 0.194 | 3.92 | 0.0001 |
| $\ln rk$ | 0.114 | 0.023 | 4.78 | 0.0000 |
| $\ln emp$ | 0.604 | 0.032 | 18.98 | 0.0000 |
| $R^2$ | within | 0.0781 | | |
| | between | 0.4993 | | |
| | overall | 0.5776 | | |

**Таблица 13.6.** Случайный эффект

Dependent Variable: $\ln rout$

| Variable | Coefficient | Std. Error | $t$-Statistic | Probability |
|---|---|---|---|---|
| const | −1.137 | 0.120 | −9.49 | 0.0000 |
| $\ln rk$ | 0.246 | 0.017 | 13.99 | 0.0000 |
| $\ln emp$ | 0.775 | 0.026 | 30.08 | 0.0000 |
| $R^2$ | within | 0.0770 | | |
| | between | 0.4999 | | |
| | overall | 0.5803 | | |

Можно заметить, что результаты оценивания модели с фиксированным эффектом значительно отличаются как от простой регрессии, так и от регрессии со случайным эффектом. В то же время оценки в простой регрессии и в регрессии со случайным эффектом сравнительно близки. Отметим, что в модели со случайным эффектом гипотеза постоянства отдачи на масштаб не отвергается — значение соответствующей хи-квадрат статистики равно 1.24 с $p$-значением 0.265.

13.9.4. Будем использовать обозначения учебника.

1) Простая регрессия против регрессии с фиксированным эффектом. Проверяем гипотезу $\mathrm{H}_0\colon \alpha_1 = \alpha_2 = \cdots = \alpha_n = 0$ с помощью $F$-теста. Значение соответствующей $F$-статистики равно 17.62 с $p$-значением 0.0000. Гипотеза об отсутствии индивидуальных эффектов уверенно отвергается.

2) Простая регрессия против регрессии со случайным эффектом. Проверяем гипотезу $\mathrm{H}_0 : \sigma_u = 0$ с помощью теста множителей Лагранжа. Значение соответствующей хи-квадрат статистики равно 6505.11, и нулевая гипотеза безоговорочно отвергается.

3) Регрессия со случайным эффектом против регрессии с фиксированным эффектом. Проверяем гипотезу $\mathrm{H}_0\colon \operatorname{Cov}(\alpha_i, \boldsymbol{x}_{jt}) = \mathbf{0}$ с помощью теста Хаусмана. Значение соответствующей хи-квадрат статистики равно 141.00, что позволяет уверенно отвергнуть модель со случайным эффектом в пользу модели с фиксированным эффектом.

13.9.5. Мы приведем здесь лишь один из возможных результатов. Сравнительно адекватной оказалась модель с фиксированным эффектом, включающая дополнительно квадраты логарифмов капиталовложений и трудозатрат (см. таблицу13.7):

**Таблица 13.7.** Фиксированный эффект

Dependent Variable: $\ln rout$

| Variable | Coefficient | Std. Error | $t$-Statistic | Probability |
|---|---|---|---|---|
| const | −2.852 | 0.404 | −7.05 | 0.0000 |
| $\ln rk$ | 0.350 | 0.066 | 5.29 | 0.0000 |
| $\ln emp$ | 1.658 | 0.127 | 13.04 | 0.0000 |
| $\ln^2 rk$ | −0.017 | 0.005 | −3.54 | 0.0001 |
| $\ln^2 emp$ | −0.088 | 0.010 | −8.60 | 0.0000 |
| $R^2$ | within | 0.0938 | | |
| | between | 0.4216 | | |
| | overall | 0.4981 | | |

## Задача 13.10

(Tammo Bijmolt, Erwin Charlier) В этом упражнении модели панельных данных используются для анализа продаж некоторого сорта тунца (обозначенного как A) в нескольких магазинах. Использовались данные о продажах консервированного тунца в 28 магазинах Чикаго в течение 104 недель. Данные для этого примера (описание переменных приведено в таблице 13.8) находятся в файле `brand_a.xls`[1].

**Таблица 13.8**

| Переменная | Описание |
|---|---|
| $sales_a$ | объем продаж тунца сорта $A$ |
| $totsales$ | общий объем продаж магазина за весь рассматриваемый период |
| $regpr_a$ | цена тунца сорта $A$ |
| $actpr_a$ | цена тунца сорта $A$ с учетом скидки |
| $feat_a$ | фиктивная переменная (1 — если в магазине была реклама рыбы сорта $A$, и рыба сорта $A$ не выкладывалась на витрину, 0 — иначе) |
| $displ_a$ | фиктивная переменная (1 — если в магазине рыба сорта $A$ выкладывалась на витрину, и не было рекламы рыбы сорта $A$, 0 — иначе) |
| $ftdpl_a$ | фиктивная переменная (1 — если в магазине рыба сорта $A$ выкладывалась на витрину, и была реклама рыбы сорта $A$, 0 — иначе) |
| $regpr_b$ | цена тунца сорта $B$ ($B$, $C$, $D$ — конкурирующие сорта) |
| $actpr_b$ | цена тунца сорта $B$ с учетом скидки |
| $regpr_c$ | цена тунца сорта $C$ |
| $actpr_c$ | цена тунца сорта $C$ с учетом скидки |
| $regpr_d$ | цена тунца сорта $D$ |
| $actpr_d$ | цена тунца сорта $D$ с учетом скидки |

13.10.1. Исследуйте описательные статистики данных. Проверьте (например, с помощью графиков), как связана зависимая переменная с объясняющими.

13.10.2. Оцените простую (pooled) модель зависимости объема продаж от всех остальных переменных. Ввиду того что розничные цены и цены с учетом скидок сильно коррелированы, для каждого сорта $i$ тунца используйте переменные $regpr_i$ и $discount_i = regpr_i - actpr_i$ ($i = a$, $b$, $c$, $d$).

[1] Оригинальный пример доступен на странице Эрвина Чарлиера по адресу `http://center.uvt.nl/staff/charlier/paneldata.html`

13.10.3. Оцените панельную модель с фиксированными эффектами. Все ли параметры удалось оценить? Если нет, то почему? (В дальнейшем исключите из модели переменную, вызвавшую проблему.)

13.10.4. Приведите оценки стандартных ошибок коэффициентов в модели упражнения 13.10.3, интерпретируйте результаты, сравните с результатами упражнения 13.10.2.

13.10.5. Вычислите межгрупповую (between-group) оценку для модели. Интерпретируйте результаты, сравните их с результатами модели с фиксированными эффектами.

13.10.6. Оцените панельную модель со случайными эффектами. Интерпретируйте результаты и сравните с результатами упражненний 13.10.4 и 13.10.5.

13.10.7. Используя известные вам тесты (тест Хаусмана, LM-тест Бреуша–Пагана), выберите наиболее подходящую модель.

13.10.8. Считая, что издержки продажи $sales_a$ тунца сорта $A$ равны $0.5 \cdot sales_a$, выведите из модели, выбранной в упражнении 13.10.7, оптимальную цену данного сорта рыбы.

## Решение

13.10.1. В таблице 13.9 представлены описательные статистики переменных, относящихся к тунцу сорта A, и цен одного из конкурирующих сортов B.

**Таблица 13.9**

| | $sales_a$ | $regpr_a$ | $actpr_a$ | $feat_a$ | $displ_a$ | $ftdpl_a$ | $regpr_b$ | $actpr_b$ |
|---|---|---|---|---|---|---|---|---|
| Mean | 310.63 | 0.895 | 0.793 | 0.053 | 0.055 | 0.149 | 0.906 | 0.831 |
| Maximum | 5207 | 1.12 | 1.12 | 1 | 1 | 1 | 1.09 | 1.09 |
| Minimum | 7 | 0.49 | 0.4 | 0 | 0 | 0 | 0.56 | 0.39 |
| Std. Dev. | 458.63 | 0.139 | 0.169 | 0.223 | 0.227 | 0.356 | 0.131 | 0.155 |

Графически связь между переменными $actpr_a$ и $sales_a$, $actpr_b$ и $sales_b$ показана на рис. 13.1 и 13.2.

13.10.2. Результаты простой регрессии объема продаж тунца сорта A на все остальные переменные (с учетом замены $actpr_i$ на $discount_i$, $i = a, b, c, d$) представлены в таблице 13.10. Мы включили в число объясняющих переменных фактор $totsales$, хотя с содержательной точки зрения это может вызывать возражения, ведь она содержит в качестве слагаемого зависимую

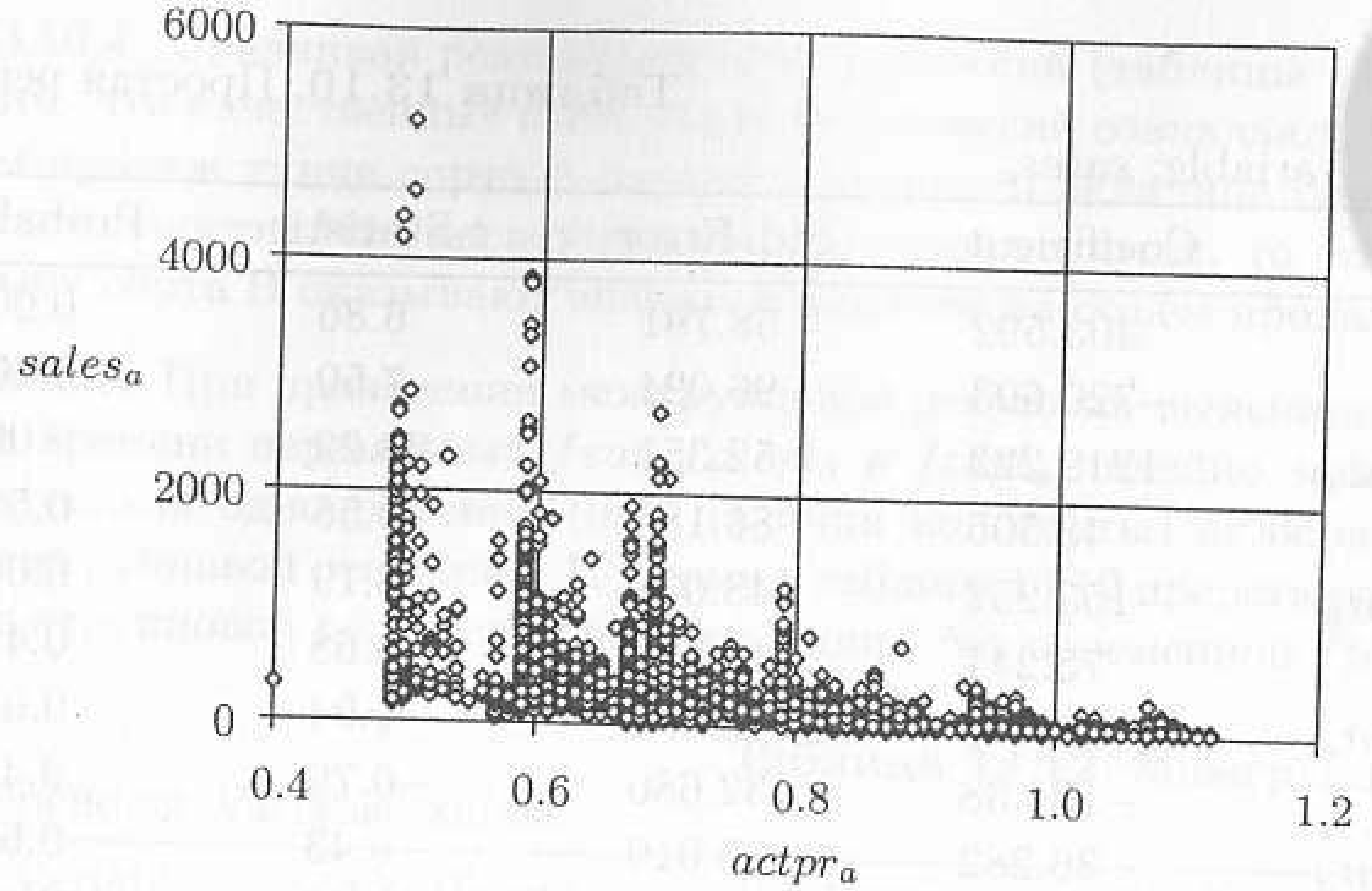



**Рис. 13.1**

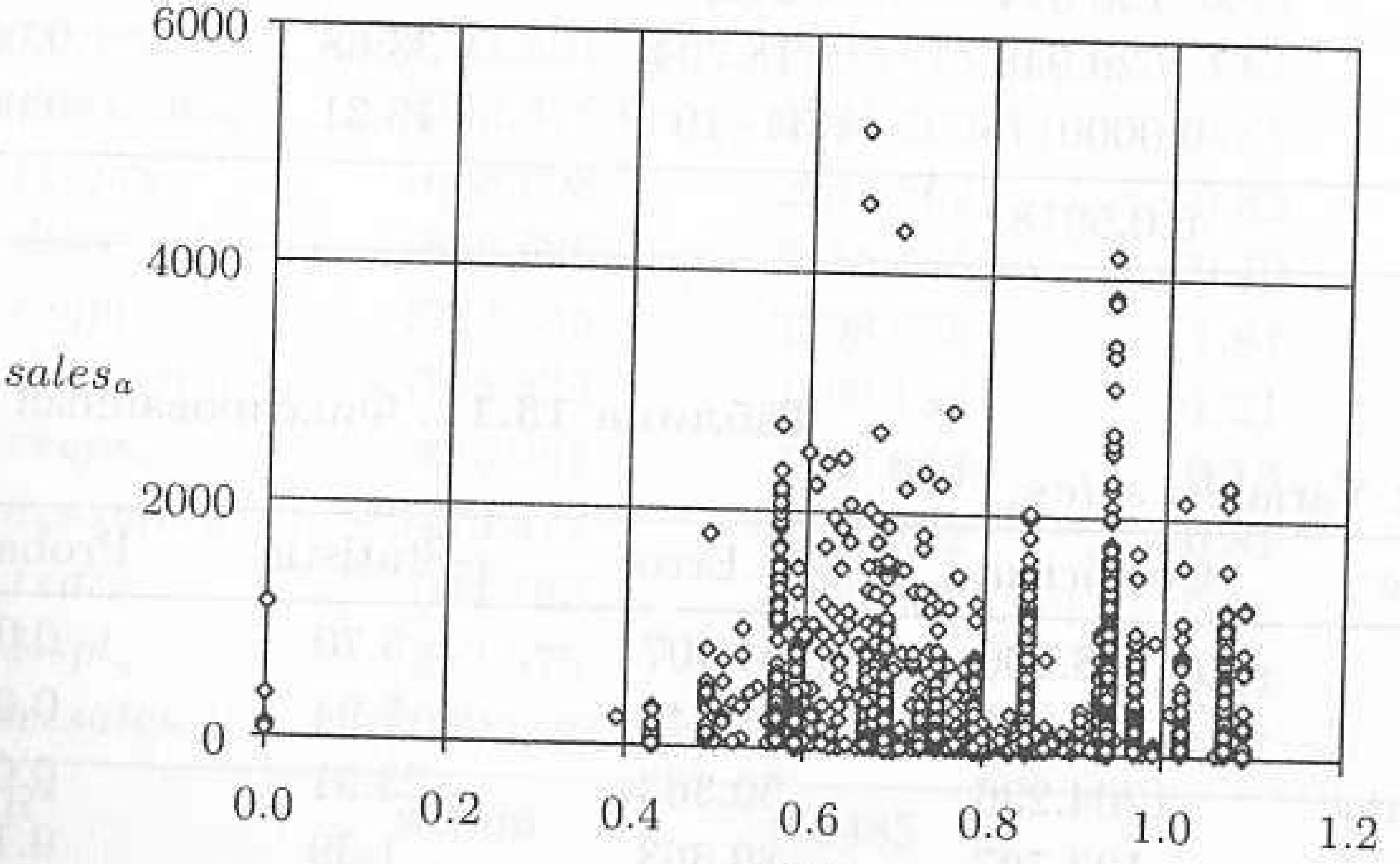



**Рис. 13.2**

переменную $sales_a$. В данном случае это обстоятельство не играет значительной роли, поскольку выборочный коэффициент корреляции между переменными $sales_a$ и $totsales$ равен 0.175. Переменную $totsales$ следует трактовать как фактор, характеризующий величину магазина.

Видим, что ценовые характеристики сортов C и D незначимы, а у сорта B значимо влияет лишь величина скидки. Знаки коэффициентов (при значимых переменных) соответствуют здравому смыслу.

13.10.3. Результаты оценивания модели с фиксированным эффектом приведены в таблице 13.11.

Как и следовало ожидать, пришлось исключить переменную $totsales$, поскольку для каждого магазина эта величина не меняется по времени.

**Таблица 13.10.** Простая регрессия

Dependent Variable: $sales_a$

| Variable | Coefficient | Std. Error | $t$-Statistic | Probability |
|---|---|---|---|---|
| const | 403.592 | 58.791 | 6.86 | 0.000 |
| $regpr_a$ | −720.603 | 96.024 | −7.50 | 0.000 |
| $discount_a$ | 1219.223 | 52.351 | 23.29 | 0.000 |
| $regpr_b$ | 48.505 | 86.186 | 0.56 | 0.574 |
| $discount_b$ | −153.252 | 48.076 | −3.19 | 0.001 |
| $regpr_c$ | 75.241 | 110.991 | 0.68 | 0.498 |
| $discount_c$ | −90.503 | 46.696 | −1.94 | 0.053 |
| $regpr_d$ | −104.38 | 132.680 | −0.79 | 0.432 |
| $discount_d$ | −36.282 | 85.010 | −0.43 | 0.670 |
| $feat_a$ | 115.369 | 26.227 | 4.40 | 0.000 |
| $displ_a$ | 190.524 | 24.578 | 7.75 | 0.000 |
| $ftdpl_a$ | 629.948 | 18.704 | 33.68 | 0.000 |
| $totsales$ | 0.0000178 | $1.34 \cdot 10^{-6}$ | 13.31 | 0.000 |
| $R^2$ | 0.0.5918 | | | |

**Таблица 13.11.** Фиксированный эффект

Dependent Variable: $sales_a$

| Variable | Coefficient | Std. Error | $t$-Statistic | Probability |
|---|---|---|---|---|
| const | 538.300 | 94.407 | 5.70 | 0.000 |
| $regpr_a$ | −581.423 | 97.948 | −5.94 | 0.000 |
| $discount_a$ | 1204.236 | 50.363 | 23.91 | 0.000 |
| $regpr_b$ | 123.797 | 89.363 | 1.39 | 0.166 |
| $discount_b$ | −181.680 | 45.909 | −3.96 | 0.001 |
| $regpr_c$ | −9.233 | 124.624 | −0.07 | 0.941 |
| $discount_c$ | −96.384 | 44.151 | −2.18 | 0.029 |
| $regpr_d$ | −34.719 | 141.210 | −0.25 | 0.806 |
| $discount_d$ | 28.537 | 80.806 | 0.35 | 0.724 |
| $feat_a$ | 116.736 | 24.887 | 4.69 | 0.000 |
| $displ_a$ | 167.413 | 23.507 | 7.12 | 0.000 |
| $ftdpl_a$ | 636.136 | 17.738 | 35.86 | 0.000 |
| $R^2$ | within | 0.6021 | | |
| | between | 0.4272 | | |
| | overall | 0.5645 | | |

13.10.4. Сравнивая результаты двух регрессий (таблицы 13.10 и 13.11), видим, что качественные выводы обеих регрессий совпадают: значимо на объем продаж тунца сорта A влияют его цена и величина скидки, рекламная деятельность. Что касается конкурирующих сортов, то только скидки на рыбу сорта B оказывают значимое влияние на объем продаж.

13.10.5. При проведении межгрупповой регрессии выяснилось, что средние по времени переменных $feat_a$, $displ_a$ и $ftdpl_a$ линейно зависимы, поэтому их нельзя одновременно (при наличии константы) включать в уравнение межгрупповой регрессии. Поэтому в таблице 13.12 представлены результаты межгрупповой регрессии (between group) без переменной $ftdpl_a$.

**Таблица 13.12.** Межгрупповая регрессия

Dependent Variable: $sales_a$

| Variable | Coefficient | Std. Error | $t$-Statistic | Probability |
|---|---|---|---|---|
| const | 421.091 | 950.565 | 0.44 | 0.664 |
| $regpr_a$ | −7241.391 | 2008.478 | −3.61 | 0.002 |
| $discount_a$ | 8392.376 | 2201.460 | 3.81 | 0.002 |
| $regpr_b$ | 1668.908 | 2044.569 | 0.82 | 0.426 |
| $discount_b$ | −866.626 | 2156.196 | −0.40 | 0.693 |
| $regpr_c$ | 4232.836 | 2298.989 | 1.84 | 0.084 |
| $discount_c$ | −3755.339 | 3096.198 | −1.21 | 0.243 |
| $regpr_d$ | 420.087 | 3266.613 | 0.13 | 0.899 |
| $discount_d$ | −3420.477 | 4231.492 | −0.81 | 0.431 |
| $feat_a$ | 594.193 | 1020.065 | 0.58 | 0.568 |
| $displ_a$ | 312.375 | 485.616 | 0.64 | 0.529 |
| $totsales$ | 0.0000239 | $5.23 \cdot 10^{-6}$ | 4.57 | 0.000 |
| $R^2$ | within | 0.3485 | | |
| | between | 0.8202 | | |
| | overall | 0.3430 | | |

Результаты межгрупповой регрессии значительно отличаются от результатов внутригрупповой (с фиксированным эффектом) регрессии. Объясняется это, по-видимому, тем, что число наблюдений в межгрупповой регрессии ($n = 28$) слишком небольшое, чтобы можно было надежно оценить 12 параметров. Поэтому прямое сопоставление межгрупповой и внутригрупповой регрессий в данном случае малоинформативно.

13.10.6. Результаты оценивания модели со случайным эффектом представлены в таблице 13.13 Мы видим, что результаты оценивания моделей с фиксированным и со случайным эффектом близки не только качественно, но и количественно. Можно проверить, что удаление из регрессий незначимых

**Таблица 13.13.** Случайный эффект

Dependent Variable: $sales_a$

| Variable | Coefficient | Std. Error | $t$-Statistic | Probability |
|---|---|---|---|---|
| const | 306.365 | 101.918 | 3.01 | 0.003 |
| $regpr_a$ | −604.542 | 97.093 | −6.23 | 0.000 |
| $discount_a$ | 1202.405 | 50.338 | 23.89 | 0.000 |
| $regpr_b$ | 103.463 | 88.107 | 1.17 | 0.240 |
| $discount_b$ | −177.512 | 45.988 | −3.86 | 0.000 |
| $regpr_c$ | −23.630 | 119.882 | −0.20 | 0.844 |
| $discount_c$ | −95.543 | 44.295 | −2.16 | 0.031 |
| $regpr_d$ | −54.481 | 139.081 | −0.39 | 0.695 |
| $discount_d$ | 21.077 | 81.000 | 0.26 | 0.795 |
| $feat_a$ | 117.565 | 24.943 | 4.71 | 0.000 |
| $displ_a$ | 169.988 | 23.550 | 7.22 | 0.000 |
| $ftdpl_a$ | 635.906 | 17.785 | 35.76 | 0.000 |
| $totsales$ | 0.0000176 | $3.69 \cdot 10^{-6}$ | 4.76 | 0.000 |
| $R^2$ | within | 0.6020 | | |
| | between | 0.4921 | | |
| | overall | 0.5905 | | |

переменных не приводит к существенному изменению оценок коэффициентов при оставшихся переменных.

Поскольку все модели, которые здесь рассматриваются, являются простыми линейными моделями, интерпретация коэффициентов стандартная, и мы рекомендуем читателю проделать это самостоятельно.

13.10.7. Так же, как и в предыдущей задаче, проверим гипотезу «простая регрессия против модели с фиксированным эффектом» с помощью $F$-теста. Значение соответствующей $F$-статистики равно 14.35, что говорит в пользу модели с фиксированным эффектом.

Проверяя гипотезу «простая регрессия против модели со случайным эффектом» с помощью теста множителей Лагранжа (теста Бреуша–Пагана), получаем значение соответствующей хи-квадрат статистики, равное 1730.17. Вновь модель простой регрессии отвергается.

Наконец, воспользуемся тестом Хаусмана для проверки гипотезы «модель со случайным эффектом против модели с фиксированным эффектом». Значение соответствующей хи-квадрат статистики равно 8.96. При нулевой гипотезе проверочная статистика имеет хи-квадрат распределение с 12 степенями свободы. Соответствующее $p$-значение равно 0.706. Таким образом, тест Хаусмана не отвергает гипотезу о наличии случайного индивидуального эффекта.

Следует отметить, что к выводам относительно выбора модели следует относиться достаточно осторожно. Если из рассмотренной модели удалить незначимые переменные и оценить ее как модель с фиксированным эффектом и как модель со случайным эффектом, то вновь результаты оценивания будут очень схожими, однако, применив тест Хаусмана, получим $p$-значение, равное 0.024, что говорит в пользу модели с фиксированным эффектом.

13.10.8. Для определенности остановимся на модели со случайным эффектом, результаты оценивания которой представлены в таблице 13.13. Кроме того, для упрощения выкладок будем считать, что продажа тунца сорта А происходит по регулярной цене $regpr_a$. Итак, согласно принятой модели средний объем продаж есть линейная комбинация переменных, включенных в модель, с коэффициентами, приведенными в таблице 13.13. Выделим слагаемое, относящееся к переменной $regpr_a$, т. е. представим объем продаж в виде

$$sales_a = \beta_1 regpr_a + V,$$

где $V$ — линейная комбинация остальных переменных. Обозначим также для удобства $regpr_a = x$, $sales_a = y$. Тогда чистая прибыль $netprofit$ от продажи объемом $y$ равна

$$\begin{aligned} netprofit = xy - 0.5y = (x-0.5)y = (x-0.5)(\beta_1 x + V) \\ = \beta_1 x^2 + (V - 0.5\beta_1)x - 0.5V. \end{aligned}$$

Поскольку $\beta_1 < 0$, то максимум этой функции достигается при

$$x_{\max} = -\frac{V - 0.5\beta_1}{2\beta_1}. \qquad (*)$$

Таким образом, оптимальная цена зависит от значений остальных переменных. В подобных случаях можно либо найти значение $x_{\max}$ для среднего (по выборке) значения $V$, либо вычислить $x_{\max}$ по формуле $(*)$ для каждого наблюдения и затем взять среднее (по выборке) значение $x_{\max}$. Мы здесь дадим результаты второго подхода, оставляя первый читателю. Итак, проводя необходимые вычисления, получаем, что среднее значение оптимальной регулярной цены тунца сорта А в соответствии с предложенной моделью равно 0.954. Интересно отметить, что среднее значение переменной $regpr_a$ равно 0.895. Полученный результат говорит в пользу адекватности построенной модели.

# Глава 14

# Предварительное тестирование: введение

## Задача 14.1

Покажите, что $\widehat{\beta}_r$ и $\widehat{\theta}$ независимы.[1]

## Решение

Из формул (4.19) (см. также п. 14.2) получаем

$$\widehat{\beta}_r = (X'X)^{-1}Xy \quad \text{и} \quad \widehat{\theta} = (Z'MZ)^{1/2}\widehat{\gamma} = (Z'MZ)^{-1/2}Z'My.$$

Из тождества $MX = 0$ получаем, что ковариация $\widehat{\beta}_r$ и $\widehat{\theta}$ равна 0:

$$\begin{aligned}\operatorname{Cov}(\widehat{\beta}_r, \widehat{\theta}) &= \mathrm{E}\left((X'X)^{-1}X\varepsilon\varepsilon' MZ(Z'MZ)^{-1/2}\right)\\ &= \sigma^2(X'X)^{-1}X'MZ(Z'MZ)^{-1/2} = 0.\end{aligned}$$

Поскольку $\widehat{\beta}_r$ и $\widehat{\theta}$ имеют совместное нормальное распределение, то из некоррелированности следует их независимость.

---

[1]Все задачи данного параграфа ссылаются на формулы, разделы и теоремы главы 14 учебника *Я. Р. Магнус, П. К. Катышев, А. А. Пересецкий* «Эконометрика. Начальный курс» (изд. 7, 8-е. М.: Дело, 2005, 2007).

## Задача 14.2

Докажите теорему 14.1, используя следующие шаги.

а) Пусть $X_* = \begin{bmatrix} X & Z \end{bmatrix}$, $\beta_*' = \begin{bmatrix} \beta' & \gamma' \end{bmatrix}$ и $R = \begin{bmatrix} 0 & S_i' \end{bmatrix}$. Покажите, что МНК-оценка параметра $\beta_*$ в модели $y = X_*\beta_* + \varepsilon$ при ограничении $R\beta_* = 0$ задается следующей формулой:

$$\widehat{\beta}_* = (X_*'X_*)^{-1}X_*'y - (X_*'X_*)^{-1}R'\left(R(X_*'X_*)^{-1}R'\right)^{-1}R(X_*'X_*)^{-1}X_*'y.$$

б) Покажите, что

$$(X_*'X_*)^{-1} = \begin{bmatrix} X'X & X'Z \\ Z'X & Z'Z \end{bmatrix}^{-1} = \begin{bmatrix} (X'X)^{-1} + QQ' & -Q(Z'MZ)^{-1/2} \\ -(Z'MZ)^{-1/2}Q' & (Z'MZ)^{-1} \end{bmatrix}.$$

в) Упростите и получите остальные результаты теоремы 14.1.

## Решение

а) Модель $y = X\beta + Z\gamma + \varepsilon$ с ограничением $S_i'\gamma = 0$ может быть записана как модель $y = X_*\beta_* + \varepsilon$ с ограничением $R\beta_* = 0$. Из формулы (4.19) следует, что МНК-оценка параметра $\beta_*$ равна

$$\widehat{\beta}_* = (X_*'X_*)^{-1}X_*'y - (X_*'X_*)^{-1}R'\left(R(X_*'X_*)^{-1}R'\right)^{-1}R(X_*'X_*)^{-1}X_*'y.$$

б) Непосредственным умножением проверяем следующее равенство:

$$\begin{bmatrix} X'X & X'Z \\ Z'X & Z'Z \end{bmatrix} \begin{bmatrix} (X'X)^{-1} + QQ' & -Q(Z'MZ)^{-1/2} \\ -(Z'MZ)^{-1/2}Q' & (Z'MZ)^{-1} \end{bmatrix} = \begin{bmatrix} A & B \\ C & D \end{bmatrix},$$

где

$$\begin{aligned} A &= I + X'XQQ' - X'Z(Z'MZ)^{-1/2}Q', \\ B &= -X'XQ(Z'MZ)^{-1/2} + X'Z(Z'MZ)^{-1}, \\ C &= Z'X(X'X)^{-1} + Z'XQQ' - Z'Z(Z'MZ)^{-1/2}Q', \\ D &= -Z'XQ(Z'MZ)^{-1/2} + Z'Z(Z'MZ)^{-1}. \end{aligned}$$

Напомним, что матрица $Q = (X'X)^{-1}X'Z(Z'MZ)^{-1/2}$, тогда $X'XQ = X'Z(Z'MZ)^{-1/2}$, откуда следует, что $A = I$ и $B = 0$. Подставляя выражение для $Q$, получаем:

$$\begin{aligned} Z'XQ &= Z'X(X'X)^{-1}X'Z(Z'MZ)^{-1/2} \\ &= Z'(I - M)Z(Z'MZ)^{-1/2} = Z'Z(Z'MZ)^{-1/2} - (Z'MZ)^{1/2}, \end{aligned}$$

откуда следует

$$\begin{aligned} C &= Z'X(X'X)^{-1} + (Z'XQ - Z'Z(Z'MZ)^{-1/2})Q' \\ &= Z'X(X'X)^{-1} - (Z'MZ)^{1/2}Q' = 0 \end{aligned}$$

и $D = I$.

в) Из $X_* = [X \quad Z]$, $\beta'_* = [\beta' \quad \gamma']$ и $R = [0 \quad S'_i]$ получаем:

$$R(X'_*X_*)^{-1}R' = S'_i(Z'MZ)^{-1}S_i,$$

и поэтому МНК-оценка параметров $\beta$ и $\gamma$ в модели $y = X\beta + Z\gamma + \varepsilon$ с ограничением $S'_i\gamma = 0$ задается следующей формулой:

$$\begin{bmatrix}\widehat{\beta}\\ \widehat{\gamma}\end{bmatrix} = \begin{bmatrix}X'X & X'Z\\ Z'X & Z'Z\end{bmatrix}^{-1}\begin{bmatrix}X'y\\ Z'y\end{bmatrix}$$
$$-\begin{bmatrix}X'X & X'Z\\ Z'X & Z'Z\end{bmatrix}^{-1}\begin{bmatrix}0 & 0\\ 0 & S_i(S'_i(Z'MZ)^{-1}S_i)^{-1}S'_i\end{bmatrix}\begin{bmatrix}X'X & X'Z\\ Z'X & Z'Z\end{bmatrix}^{-1}\begin{bmatrix}X'y\\ Z'y\end{bmatrix}.$$

Отсюда получаем, что

$$\begin{aligned}\widehat{\beta}_{(i)} = \widehat{\beta}\\ &= ((X'X)^{-1} + QQ')X'y - Q(Z'MZ)^{-1/2}Z'y\\ &\quad - QP_iQ'X'y + QP_i(Z'MZ)^{-1/2}Z'y\\ &= (X'X)^{-1}X'y + QQ'X'y + QW_iQ'X'y - QW_i(Z'MZ)^{-1/2}Z'y\\ &= \widehat{\beta}_r - QW_i(-Q'X' + (Z'MZ)^{-1/2}Z')y = \widehat{\beta}_r - QW_i\widehat{\theta}\end{aligned}$$

и

$$\begin{aligned}\widehat{\gamma}_{(i)} = \widehat{\gamma}\\ &= -(Z'MZ)^{-1/2}Q'X'y + (Z'MZ)^{-1}Z'y\\ &\quad + (Z'MZ)^{-1/2}P_iQ'X'y - (Z'MZ)^{-1/2}P_i(Z'MZ)^{-1/2}Z'y\\ &= (Z'MZ)^{-1/2}W_i(-Q'X' + (Z'MZ)^{-1/2}Z')y = (Z'MZ)^{-1/2}W_i\widehat{\theta}.\end{aligned}$$

Из полученных формул легко выводятся остальные утверждения теоремы 14.1.

## Задача 14.3

Покажите, что в случае $\eta = 1$ (см. п. 14.8) исследователю безразлично, какую из двух моделей выбрать (с ограничением или без ограничения).

## Решение

Данный вопрос относится к ситуации, когда у нас только один вспомогательный регрессор, т.е. $m = 1$, и модель имеет вид $y = X\beta + \gamma z + \varepsilon$. В регрессии с ограничением получаем:

$$\widehat{\beta}_r = (X'X)^{-1}X'y$$

и, следовательно,

$$\mathrm{E}(\widehat{\beta}_r - \beta) = \gamma(X'X)^{-1}X'z, \quad \mathrm{V}(\widehat{\beta}_r) = \sigma^2(X'X)^{-1},$$

отсюда

$$\begin{aligned}\mathrm{MSE}(\widehat{\beta}_r) &= \sigma^2(X'X)^{-1} + \gamma^2(X'X)^{-1}X'zz'X(X'X)^{-1}\\ &= \sigma^2(X'X)^{-1} + \gamma^2(z'Mz)qq' = \sigma^2((X'X)^{-1} + \eta^2 qq'),\end{aligned}$$

где $q = (z'Mz)^{-1/2}(X'X)^{-1}X'z$ и $\eta = (z'Mz)^{-1/2}(X'X)^{-1}X'z$. Оценка в модели без ограничений равна $\widehat{\beta}_u = \widehat{\beta}_r - \widehat{\theta}q$. Эта оценка несмещенная $\mathrm{E}(\widehat{\beta}_u) = \beta$, величины $\widehat{\beta}_r$ и $\widehat{\theta}$ независимы (см. упражнение 14.1) и $\mathrm{V}(\widehat{\theta}) = \mathrm{V}((z'Mz)^{-1/2}z'My) = \sigma^2$. Получаем:

$$\mathrm{MSE}(\widehat{\beta}_u) = \mathrm{V}(\widehat{\beta}_u) = \mathrm{V}(\widehat{\beta}_r) + \mathrm{V}(\widehat{\theta})qq' = \sigma^2((X'X)^{-1} + qq').$$

Отсюда видно, что $\mathrm{MSE}(\widehat{\beta}_r) = \mathrm{MSE}(\widehat{\beta}_u)$ тогда и только тогда, когда $\eta^2 = 1$. Это как раз тот случай, когда исследователю безразлично, какую из двух моделей (с ограничением или без ограничения) выбрать.

## Задача 14.4

Покажите, что не любой выбор функции $\lambda$ приводит к естественной процедуре.

## Решение

Рассмотрим следующую функцию $\lambda$: $\lambda = 1$, если $\widehat{\eta} \leqslant c$, и 0 в остальных случаях. Тогда в точке $\eta = 0$ для любого $c > 0$ получаем: $\mathrm{MSE}(\lambda\widehat{\eta}) < \mathrm{E}(\lambda)$. Таким образом, процедура не является естественной.

## Задача 14.5

Покажите, что матрица $\mathrm{V}(W\widehat{\eta})$ ограничена при $n \to \infty$.

## Решение

Известно, что $\widehat{\eta}$ имеет нормальное распределение со средним $\eta$ и матрицей ковариаций $I_m$. Также известно, что матрица $W = \sum_i \lambda_i W_i$ является взвешенным средним конечного числа идемпотентных матриц. Отсюда с вероятностью 1 все элементы этой матрицы не превосходят 1, а все ее диагональные элементы (и собственные значения) с вероятностью 1 лежат в интервале $[0, 1]$. При $n \to \infty$ эти свойства не изменяются.

## Задача 14.6

Найдите функцию распределения величины $\lambda\widehat{\eta} - \eta$ при условии $\eta = 0$, т. е. $\gamma = 0$.

## Решение

Необходимо найти функцию распределения случайной величины $\xi = \lambda(\widehat{\eta})\widehat{\eta}$, где $\widehat{\eta} \sim N(0,1)$ и $\lambda(x) = 0$ при $x \leqslant c$, и 1 в противном случае. Распределение не является абсолютно непрерывным и не имеет плотности. Однако нетрудно найти функцию распределения этой случайной величины. По определению функции распределения имеем:

$$F_\xi(y) = \mathrm{P}(\xi \leqslant y) = \begin{cases} \Phi(y), & -\infty < y < -c, \\ \Phi(-c), & -c \leqslant y < 0, \\ \Phi(c), & 0 \leqslant y < c, \\ \Phi(y), & c \leqslant y < \infty, \end{cases}$$

где через $\Phi(y)$ обозначена функция распределения стандартной нормальной случайной величины. Видно, что $F_\xi$ имеет скачок в нуле, и величина скачка равна вероятности выбора модели с ограничением, т. е. равна $\Phi(c) - \Phi(-c)$. График функции распределения для случая $c = 1$ представлен на рис. 14.1.

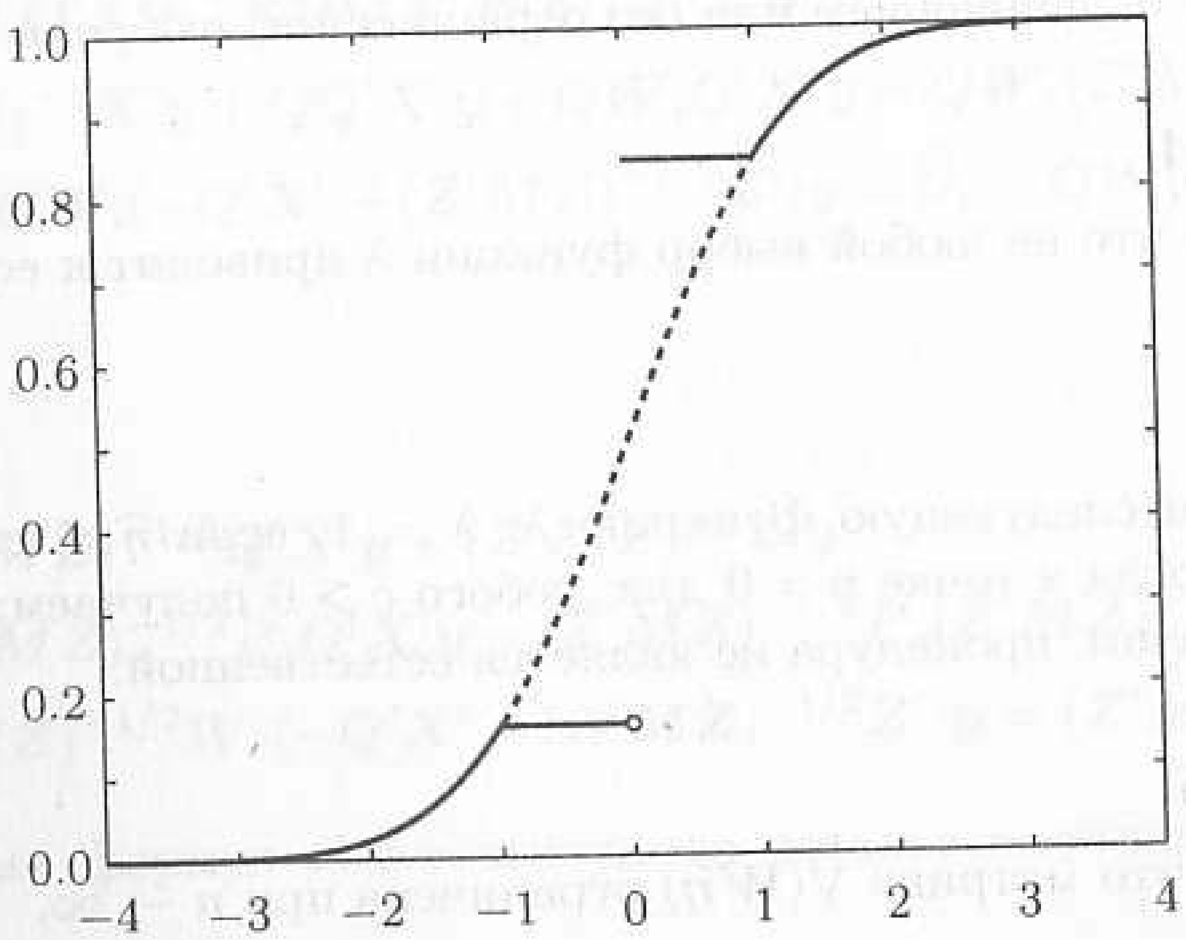



**Рис. 14.1.** График функции распределения величины $\xi$ для $c = 1$ (сплошная линия) и стандартной нормальной функции распределения (пунктирная линия)

# Глава 15

# Эконометрика финансовых рынков

## Задача 15.1

Выведите формулы для параметров эффективного портфеля с наименьшей возможной дисперсией доходности (15.14)

$$\mu_g = \mathrm{E}(R_g) = \frac{B}{C}, \quad \boldsymbol{w}_g = \frac{1}{C}\boldsymbol{\Sigma}^{-1}\boldsymbol{\imath}, \quad \sigma_g^2 = \mathrm{V}(R_g) = \frac{1}{C}$$

и (15.15)

$$\sigma^2(\mu) = \mathrm{V}(R) = \frac{C\mu^2 - 2B\mu + A}{AC - B^2}.$$

## Решение

Вычислим дисперсию доходности эффективного портфеля с ожидаемой доходностью $\mu$ (используем формулу (15.13)) и обозначим $D = AC - B^2$

$$\begin{aligned}
\sigma^2(\mu) &= \boldsymbol{w}'\boldsymbol{\Sigma}\boldsymbol{w} = \\
&= \left(\frac{C\mu - B}{D}\boldsymbol{m} + \frac{A - B\mu}{D}\boldsymbol{\imath}\right)' \boldsymbol{\Sigma}^{-1}\left(\frac{C\mu - B}{D}\boldsymbol{m} + \frac{A - B\mu}{D}\boldsymbol{\imath}\right) \\
&= D^{-2}\left((C\mu - B)^2 A + 2(C\mu - B)(A - B\mu)B + (A - B\mu)^2 C\right) \\
&= D^{-2}\Big(\mu^2(C^2A - 2CB^2 + B^2C) \\
&\quad - 2\mu(CBA - CAB - B^3 + ABC) + (B^2A - 2AB^2 + A^2C)\Big) \\
&= D^{-1}(C\mu^2 - 2B\mu + A).
\end{aligned}$$

Для того, чтобы найти наименьшее $\sigma$, дифференцируем полученное выражение по $\mu$ и приравниваем 0:

$$0 = \frac{d\sigma^2}{d\mu} = D^{-1}(2C\mu - 2B),$$

откуда получаем:

$$\begin{aligned}
\mu_g &= \frac{B}{C},\\
\sigma_g &= \frac{1}{D}\left(C\frac{B^2}{C^2} - 2B\frac{B}{C} + A\right) = \frac{B^2 - 2B^2 + AC}{DC} = \frac{1}{C},\\
\boldsymbol{w}_g &= \boldsymbol{\Sigma}^{-1}\left(\frac{C(B/C) - B}{D}\boldsymbol{m} + \frac{A - B(B/C)}{D}\boldsymbol{\imath}\right)\\
&= \boldsymbol{\Sigma}^{-1}\frac{1}{DC}(AC - B^2)\boldsymbol{\imath} = \frac{1}{C}\boldsymbol{\Sigma}^{-1}\boldsymbol{\imath}.
\end{aligned}$$

## Задача 15.2

Выведите формулу (15.16) для ковариации доходностей двух эффективных портфелей

$$\mathrm{Cov}(R_1, R_2) = \frac{C}{AC - B^2}\left(\mu_1 - \frac{B}{C}\right)\left(\mu_2 - \frac{B}{C}\right) + \frac{1}{C}.$$

## Решение

Используем формулу (15.13) и обозначим, как и в задаче 15.1, $D = AC - B^2$. Введем также обозначения $\eta_1 = \mu_1 - B/C$, $\eta_2 = \mu_2 - B/C$, тогда $C\mu_i - B = C\eta_i$ и $A - B\mu_i = D/C - B\eta_i$. Получаем:

$$\begin{aligned}
\mathrm{Cov}(R_1, R_2) &= \mathrm{Cov}(\boldsymbol{w}_1'\boldsymbol{r}, \boldsymbol{w}_2'\boldsymbol{r}) = \boldsymbol{w}_1'\,\mathrm{Cov}(\boldsymbol{r}, \boldsymbol{r})\boldsymbol{w}_2 = \boldsymbol{w}_1'\boldsymbol{\Sigma}\boldsymbol{w}_2\\
&= \left(\frac{C\mu_1 - B}{D}\boldsymbol{m} + \frac{A - B\mu_1}{D}\boldsymbol{\imath}\right)'\boldsymbol{\Sigma}^{-1}\left(\frac{C\mu_2 - B}{D}\boldsymbol{m} + \frac{A - B\mu_2}{D}\boldsymbol{\imath}\right)\\
&= D^{-2}\left(C\eta_1\boldsymbol{m} + \left(\frac{D}{C} - B\eta_1\right)\boldsymbol{\imath}\right)'\boldsymbol{\Sigma}^{-1}\left(C\eta_2\boldsymbol{m} + \left(\frac{D}{C} - B\eta_2\right)\boldsymbol{\imath}\right)\\
&= D^{-2}\Big(AC^2\eta_1\eta_2 + (D - BC\eta_2)\eta_1 B + (D - BC\eta_1)\eta_2 B\\
&\qquad + C^{-1}(D - BC\eta_1)(D - BC\eta_2)\Big)\\
&= D^{-2}\left((AC^2 - B^2C)\eta_1\eta_2 + C^{-1}D^2\right)\\
&= \frac{C}{D}\eta_1\eta_2 + \frac{1}{C} = \frac{C}{D}\left(\mu_1 - \frac{B}{C}\right)\left(\mu_2 - \frac{B}{C}\right) + \frac{1}{C},
\end{aligned}$$

что и требовалось показать. Заметим, что из последней формулы при $\mu_1 = \mu_2 = \mu$ легко можно получить решение задачи 15.1.

## Задача 15.3

Выведите формулу для оптимального портфеля при наличии безрискового актива, аналогичную формуле (15.13)

$$w = \Sigma^{-1}\left(\frac{C\mu - B}{AC - B^2}m + \frac{A - B\mu}{AC - B^2}\imath\right),$$

повторив шаги (15.8)–(15.12) вывода формулы (15.13), т. е. записав целевую функцию, затем ограничения и решив задачу оптимизации при наличии ограничений методом Лагранжа.

## Решение

Рассмотрим портфель с распределением ресурсов $(w_0, w)$, где $w_0$ — доля средств, вложенная в безрисковый актив, $w_0 + w'\imath = 1$. Доходность такого портфеля равна

$$R = w_0 R^f + w'r.$$

Средняя доходность и дисперсия доходности равны соответственно

$$\mu = \mathrm{E}(R) = w_0 R^f + w'm = (1 - w'\imath)R^f + w'm,$$
$$\sigma^2 = \mathrm{V}(R) = w'\Sigma w.$$

Задача поиска оптимального портфеля с данной средней доходностью формулируется следующим образам:

$$w'\Sigma w \to \min, \quad \text{при условии} \quad (1 - w'\imath)R^f + w'm = \mu.$$

Чтобы решить эту задачу, рассмотрим функцию Лагранжа

$$L(w) = w'\Sigma w - 2\lambda((1 - w'\imath)R^f + w'm - \mu).$$

Условия минимума первого порядка имеют вид

$$\Sigma w - \lambda(-R^f\imath + m) = 0, \quad w = \lambda\Sigma^{-1}(-R^f\imath + m).$$

Подставим это выражение для $w$ в ограничение:

$$\mu = (1 - w'\imath)R^f + w'm,$$
$$\mu = (m - R^f\imath)'w + R^f = \lambda(m - R^f\imath)'\Sigma^{-1}(m - R^f\imath) + R^f.$$

Получаем:

$$\lambda = \frac{\mu - R^f}{(m - R^f\imath)'\Sigma^{-1}(m - R^f\imath)},$$

или

$$w = \frac{\mu - R^f}{(m - R^f\imath)'\Sigma^{-1}(m - R^f\imath)}\Sigma^{-1}(m - R^f\imath),$$

(заметим, что с точностью до нормировки эта формула совпадает с формулой (15.37) для тангенциального портфеля) и

$$\sigma^2 = w'\Sigma w = \left(\frac{\mu - R^f}{(m - R^f \imath)'\Sigma^{-1}(m - R^f \imath)}\right)^2 (m - R^f \imath)'\Sigma^{-1}(m - R^f \imath)$$
$$= \frac{(\mu - R^f)^2}{(m - R^f \imath)'\Sigma^{-1}(m - R^f \imath)}.$$

Отсюда получаем фронт эффективных портфелей — прямая линия на плоскости $(\sigma, \mu)$, проходящая через точку $(0, R^f)$

$$\sigma = (\mu - R^f)\left((m - R^f \imath)'\Sigma^{-1}(m - R^f \imath)\right)^{-1/2} = \gamma^{-1}(\mu - R^f).$$

Угловой коэффициент этой прямой (коэффициент Шарпа) равен

$$\gamma = \left((m - R^f \imath)'\Sigma^{-1}(m - R^f \imath)\right)^{1/2}.$$

## Задача 15.4

Докажите эквивалентность двух определений закона одной цены (*law of one price*), разд. 15.6, стр. 406.

*Определение 1.* $p(aX_1 + bX_2) = ap(X_1) + bp(X_2)$. Другими словами, $p$ — линейная функция на пространстве $\mathcal{X}$.

*Определение 2.* Если есть две выплаты, которые одинаковы при всех возможных состояниях экономики в конце периода, то их сегодняшние цены совпадают.

## Решение

Пусть выполняется определение 1, и пусть $X_1$ и $X_2$ — две выплаты, одинаковые при всех состояниях экономики; $p_1 = p(X_1)$ и $p_1 = p(X_1)$ — их цены. Рассмотрим новую выплату $X_3 = X_1 - X_2$, которая тождественно равна 0 при всех состояниях экономики: $X_3 \equiv 0$. В силу линейности функции $p(X)$ имеем $0 = p(0) = p(X_3) = p(X_1 - X_2) = p(X_1) - p(X_2)$, откуда следует $p_1 = p_2$.

Пусть выполняется определение 2, и функция $p(X)$ непрерывна. Рассмотрим две выплаты $X_1$ и $X_2$. Рассмотрим также выплату $X_3 \equiv X_1 + X_2$. Очевидно, выплата $X_3$ во всех состояниях экономики совпадает с объединением выплат $X_1$ и $X_2$. Отсюда $p(X_3) = p(X_1 + X_2) = p(X_1) + p(X_2)$. Аналогично можно доказать, что $p\left(\frac{m}{n}X\right) = \frac{m}{n}p(X)$, далее, используя непрерывность, получаем $p(aX) = ap(X)$ для всех $a$.

## Задача 15.5

Докажите, что из закона одной цены следует теорема о существовании дисконтирующего множителя (см. разд. 15.6, стр. 407).

## Решение

Рассмотрим линии уровня функции $p(X)$: $\mathcal{X}_a = \{X : X \in \mathcal{X}, p(X) = a\}$. Очевидно, они представляют собой гиперплоскости, параллельные друг другу. Гиперплоскость $\mathcal{X}_0$ содержит выплаты с нулевой ценой (*excess returns*). Возьмем некоторый ненулевой вектор $Z$, ортогональный к $\mathcal{X}_0$. Таким образом, $\mathrm{E}(ZX) = 0$ для всех $X \in \mathcal{X}_0$. Выберем вектор $X_1 \in \mathcal{X}_1$. Все другие векторы из $\mathcal{X}_1$ имеют вид $X = X_1 + X_0$, где $X_0 \in \mathcal{X}_0$. Имеем $\mathrm{E}(ZX) = \mathrm{E}(ZX_1) + \mathrm{E}(ZX_0) = \mathrm{E}(ZX_1)$. Осталось подобрать подходящее кратное $X^*$ вектора $Z$. Возьмем $X^* = Z/\mathrm{E}(ZX_1)$, тогда для всех $X \in \mathcal{X}_1$ имеем $\mathrm{E}(X^*X) = \mathrm{E}(X^*X_1) = \mathrm{E}(ZX_1)/\mathrm{E}(ZX_1) = 1$. Таким образом, равенство $p(X) = \mathrm{E}(X^*X)$ выполняется для всех $X$ из $\mathcal{X}_0$ и $\mathcal{X}_1$. Нетрудно проверить, что это равенство выполняется также и для всех остальных $\mathcal{X}_a$.

## Задача 15.6

В файле `index.xls` содержатся данные по четырем фондовым индексам.

а) Рассчитайте средние месячные доходности, дисперсии и корреляции доходностей индексов. В дальнейшем предположите, что доходности имеют нормальное распределение и выборочные моменты достаточно точно описывают распределение будущих доходностей.

б) Постройте эффективные портфели, состоящие из фондовых индексов США, Великобритании, Японии, с годовой ожидаемой доходностью в 6%, 9%, 12%, 15% и 18%.

в) Покажите, как информация о двух эффективных портфелях с годовой доходностью, например, 12% и 15% может быть использована для построения других эффективных портфелей.

г) Найдите портфель с наименьшей дисперсией.

д) Тот же вопрос, что в п. б), но теперь в случае, когда «короткие продажи» запрещены.

е) Повторите пп. б), в), г), д) для случая портфеля, в который добавлен четвертый актив (фондовый индекс Гонконга).

ж) Повторите предыдущие пункты в предположении, что доступен также безрисковый актив с годовой доходностью 6%.

## Решение

а) Рассчитаем однодневные логарифмические доходности по наблюдениям по рабочим дням (считая, что за выходные ничего не происходит). Поскольку в выборке в среднем около 261 наблюдения в год, то переходим к годичным доходностям, умножая однодневные доходности на множитель 261. Средние и корреляции логарифмических доходностей представлены в таблице 15.1.

Таблица 15.1

| | Коэффициенты корреляции | | | |
|---|---|---|---|---|
| | DJINDUS | FTSE100 | JAPDOWA | HNGKNGI |
| DJINDUS | 1 | 0.383 | 0.065 | 0.116 |
| FTSE100 | 0.383 | 1 | 0.244 | 0.372 |
| JAPDOWA | 0.065 | 0.244 | 1 | 0.379 |
| HNGKNGI | 0.116 | 0.372 | 0.379 | 1 |
| Средние | 15.5% | 9.17% | −6.46% | 7.08% |

б) Воспользуемся формулой (15.13)

$$w = \Sigma^{-1}\left(\frac{C\mu - B}{AC - B^2}m + \frac{A - B\mu}{AC - B^2}\imath\right),$$

где $A = m'\Sigma^{-1}m = 0.004018$, $B = m'\Sigma^{-1}\imath = 0.019568$ и $C = \imath'\Sigma^{-1}\imath = 0.234775$. Результаты см. в таблице 15.2, а также на рис. 15.1 ($\mu$).

Таблица 15.2

| $\mu$ | 0.06 | 0.09 | 0.12 | 0.15 | 0.18 |
|---|---|---|---|---|---|
| | Доли вложений в активы | | | | |
| DJINDUS | 0.311 | 0.438 | 0.565 | 0.693 | 0.820 |
| FTSE100 | 0.360 | 0.374 | 0.387 | 0.400 | 0.413 |
| JAPDOWA | 0.329 | 0.188 | 0.048 | −0.092 | −0.233 |
| $\sigma$ | 2.118 | 2.068 | 2.196 | 2.474 | 2.859 |

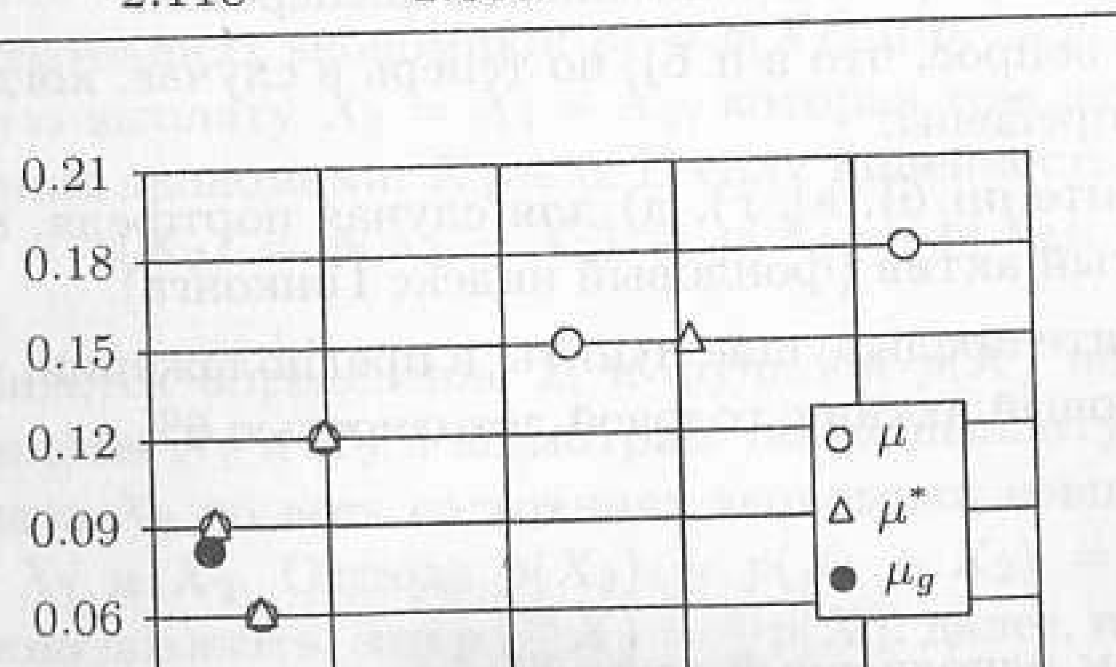



**Рис. 15.1.** $\mu$ — эффективная граница; $\mu^*$ — эффективная граница при запрете на «короткие продажи»; $\mu_g$ — портфель с наименьшей дисперсией

в) По портфелям $w_{0.12}$ и $w_{0.15}$ остальные портфели получаются как следующие линейные комбинации:

$$w_{0.06} = 3w_{0.12} - 2w_{0.15},$$
$$w_{0.09} = 2w_{0.12} - w_{0.15},$$
$$w_{0.18} = -w_{0.12} + 2w_{0.15}.$$

г) Портфель с наименьшей дисперсией доходности находится по формулам (15.14):

$$\mu_g = \frac{B}{C} = 0.0833, \quad \sigma_g^2 = \frac{1}{C} = 4.259, \quad \sigma_g = 2.064,$$
$$w_g = \frac{1}{C}\Sigma^{-1}\imath = (0.410,\ 0.371,\ 0.220)'$$

(см. рис. 15.1 ($\mu_g$)).

д) При условии отсутствия «коротких продаж» необходимо минимизировать дисперсию портфеля $w'\Sigma w$ при условиях $\mu = w'm$ и $w \geqslant 0$ (все компоненты вектора $w$ неотрицательные). Это можно сделать, например, используя Solver в Excel. Результат представлен в таблице 15.3 и на рис. 15.1 ($\mu^*$).

**Таблица 15.3**

| $\mu$ | 0.06 | 0.09 | 0.12 | 0.15 | 0.18 |
|---|---|---|---|---|---|
| | Доли вложений в активы | | | | |
| DJINDUS | 0.311 | 0.438 | 0.565 | 0.921 | – |
| FTSE100 | 0.360 | 0.374 | 0.387 | 0.0788 | – |
| JAPDOWA | 0.329 | 0.188 | 0.048 | 0 | – |
| $\sigma$ | 2.118 | 2.068 | 2.196 | 2.613 | – |

Прочерк в последней колонке означает, что решения соответствующей задачи не существует. В отсутствие коротких продаж невозможно получить портфель с ожидаемой доходностью выше, чем ожидаемая доходность каждого из активов, входящих в портфель.

е) Действуя аналогично пп. а)–д), получаем $A = 0.00415$, $B = 0.01972$ и $C = 0.23495$. При наличии четырех активов получаем оптимальные портфели, представленные в таблице 15.4 и на рис. 15.2 ($\mu$).

По портфелям $w_{0.12}$ и $w_{0.15}$ остальные портфели получаются как следующие линейные комбинации (сравните с п. в)):

$$w_{0.06} = 3w_{0.12} - 2w_{0.15},$$
$$w_{0.09} = 2w_{0.12} - w_{0.15},$$
$$w_{0.18} = -w_{0.12} + 2w_{0.15}.$$

Таблица 15.4

| $\mu$ | 0.06 | 0.09 | 0.12 | 0.15 | 0.18 |
|---|---|---|---|---|---|
| | Доли вложений в активы | | | | |
| DJINDUS | 0.312 | 0.435 | 0.557 | 0.679 | 0.802 |
| FTSE100 | 0.367 | 0.362 | 0.357 | 0.353 | 0.348 |
| JAPDOWA | 0.331 | 0.185 | 0.038 | −0.108 | −0.254 |
| HNGKNGI | −0.010 | 0.019 | 0.047 | 0.076 | 0.104 |
| $\sigma$ | 2.118 | 2.067 | 2.186 | 2.451 | 2.821 |

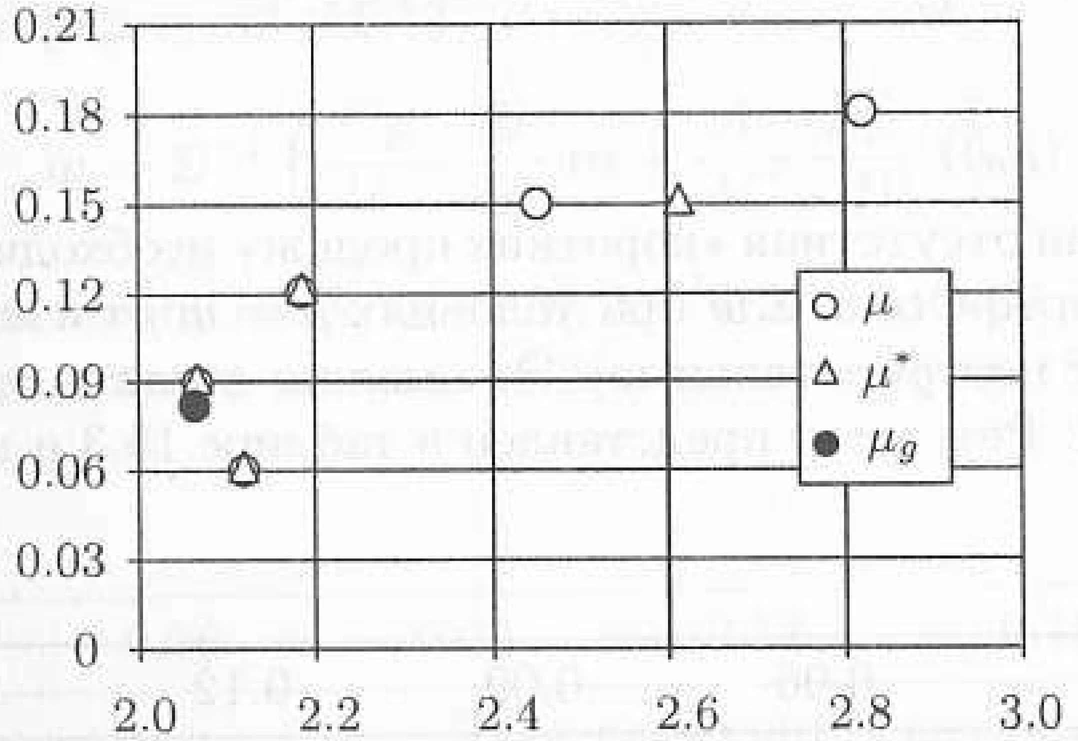



**Рис. 15.2.** Граница эффективных портфелей для 4 активов. $\mu$ — эффективная граница; $\mu^*$ — эффективная граница при запрете на «короткие продажи»; $\mu_g$ — портфель с наименьшей дисперсией

Портфель с наименьшей дисперсией доходности находится по формулам (15.14):

$$\mu_g = \frac{B}{C} = 0.0839, \qquad \sigma_g^2 = \frac{1}{C} = 4.256, \quad \sigma_g = 2.063,$$

$$w_g = \frac{1}{C}\Sigma^{-1}\imath = (0.410,\, 0.363,\, 0.214,\, 0.013)',$$

см. рис. 15.2 ($\mu_g$).

При ограничении на «короткие продажи» получаем оптимальные портфели, представленные в таблице 15.5 и на рис. 15.2 ($\mu^*$).

ж) Рассмотрим случай с 4 индексами и безрисковым активом. Как известно, в этом случае оптимальный портфель является взвесью безрискового актива и тангенциального портфеля. Распределение ресурсов в тангенциальном портфеле находится по формуле (15.37):

$$w_d = \frac{\Sigma^{-1}(m - R^f\imath)}{\imath'\Sigma^{-1}(m - R^f\imath)} = (2.218,\, 0.295,\, -1.948,\, 0.435)'.$$

**Таблица 15.5**

| $\mu$ | 0.06 | 0.09 | 0.12 | 0.15 | 0.18 |
|---|---|---|---|---|---|
| | | Доли вложений в активы | | | |
| DJINDUS | 0.311 | 0.435 | 0.557 | 0.921 | – |
| FTSE100 | 0.360 | 0.362 | 0.357 | 0.079 | – |
| JAPDOWA | 0.329 | 0.185 | 0.038 | 0 | – |
| HNGKNGI | 0 | 0.019 | 0.047 | 0 | – |
| $\sigma$ | 2.118 | 2.067 | 2.186 | 2.613 | – |

Фронт оптимальных портфелей на плоскости $(\sigma, \mu)$ является прямой линией, соединяющей точки, соответствующие безрисковому активу и тангенциальному портфелю. Соответственно, получаем оптимальные портфели, представленные в таблице 15.6 и на рис. 15.3 ($\mu_{rf}$).

**Таблица 15.6**

| $\mu$ | 0.06 | 0.09 | 0.12 | 0.15 | 0.18 |
|---|---|---|---|---|---|
| | | Доли вложений в активы | | | |
| DJINDUS | 0 | 0.142 | 0.285 | 0.427 | 0.569 |
| FTSE100 | 0 | 0.019 | 0.038 | 0.057 | 0.076 |
| JAPDOWA | 0 | −0.125 | −0.250 | −0.375 | −0.500 |
| HNGKNGI | 0 | 0.028 | 0.056 | 0.084 | 0.112 |
| Risk-free | 1 | 0.936 | 0.872 | 0.807 | 0.743 |
| $\sigma$ | 0 | 0.583 | 1.170 | 1.755 | 2.340 |

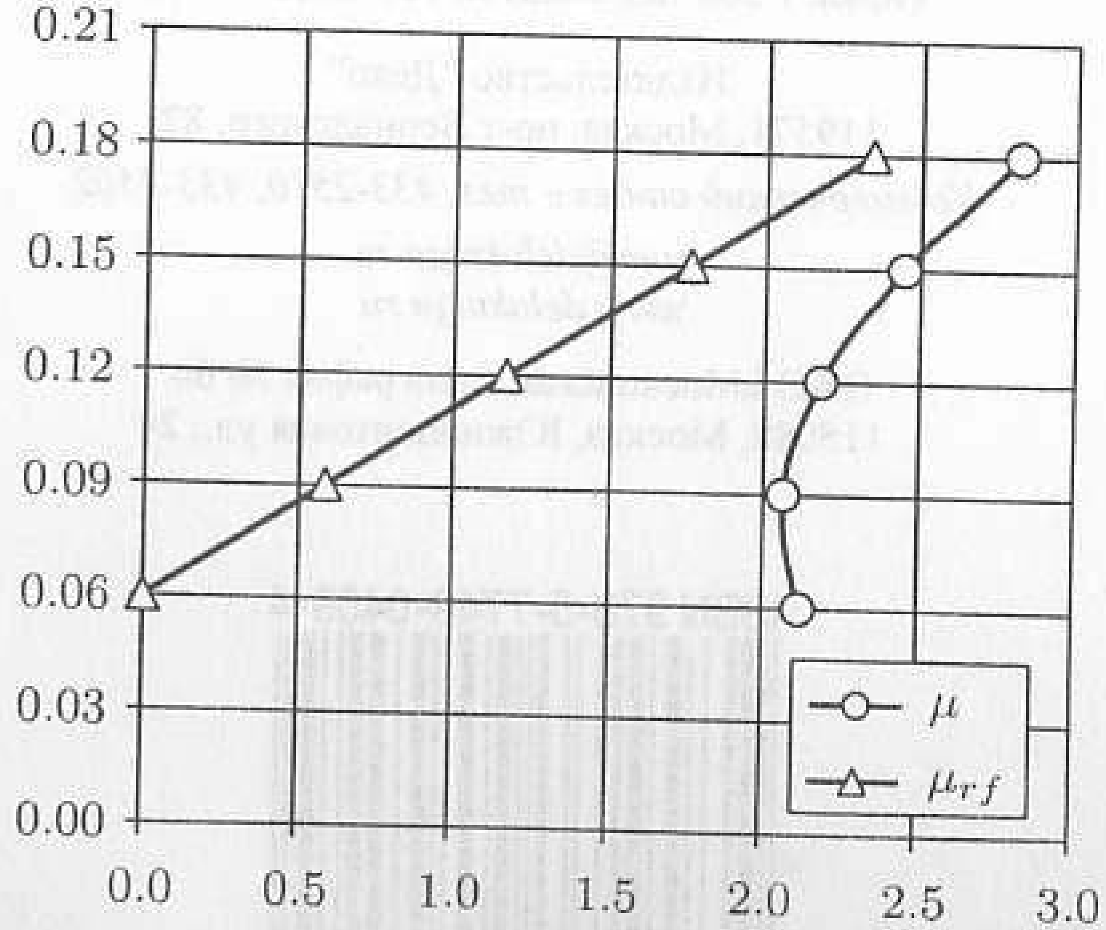



**Рис. 15.3.** Граница эффективных портфелей при отсутствии безрискового актива для ($\mu$) и при наличии безрискового актива ($\mu_{rf}$)

*Учебное пособие*

**Павел Константинович КАТЫШЕВ,**
**Ян Р. МАГНУС,**
**Анатолий Абрамович ПЕРЕСЕЦКИЙ,**
**Сергей Витальевич ГОЛОВАНЬ**

# СБОРНИК ЗАДАЧ К НАЧАЛЬНОМУ КУРСУ ЭКОНОМЕТРИКИ

Главный редактор *Ю.В. Луизо*
Зав. редакцией *Г.Г. Кобякова*
Художник *Н.В. Пьяных*
Компьютерная подготовка
оригинал-макета *С.В. Головань*
Технический редактор *А.Л. Гулина*
Корректоры *Н.В. Андрианова, М.А. Миловидова*

Санитарно-эпидемиологическое заключение
№77.99.02.953.Д009376.10.06 от 12.10.2006 г.

Подписано в печать 14.08.2007. Формат 70x100 $^1/_{16}$. Бумага офсетная.
Гарнитура Таймс. Печать офсетная. Усл. печ. л. 29,9.
Тираж 1 500 экз. Заказ № 951. Изд. № 538.

Издательство "Дело"
119571, Москва, пр-т Вернадского, 82
*Коммерческий отдел – тел. 433-2510, 433-2502*
*com@delokniga.ru*
*www.delokniga.ru*

ОАО «Московская типография № 6»
115088, Москва, Южнопортовая ул., 24

ISBN 978-5-7749-0459-4

9 785774 904594