

Дополнительные задачи по Эконометрике

Артамонов Н.В.

25 декабря 2016 г.

Оглавление

1	Метод наименьших квадратов	5
2	Стандартная линейная модель регрессии	9
3	Спецификация модели регрессии	21
4	Линейная регрессия с гетероскедастичной ошибкой	27
5	Линейная регрессия с автокоррелированной ошибкой	31
6	Системы одновременных уравнений	35
7	Метод инструментальных переменных	39
8	Временные ряды	41
8.1	Модели стационарных временных рядов	41
8.2	Модели стационарных временных рядов. Практика	43
8.3	Модели распределенных лагов	44
8.4	Модели распределенных лагов. Практика	46
8.4.1	TS – ряды	47
8.4.2	DS – ряды	49
8.4.3	DS – ряды. Практика	50
8.4.4	Многомерные модели временных рядов	52
8.4.5	Из пособия	53
9	Модели регрессии с бинарной зависимой переменной	57
9.1	LPM-модель	57
9.2	probit/logit-модель	59

Глава 1

Метод наименьших квадратов

№1. Пусть задано n наблюдений (точек на плоскости) $\{x_i, y_i\}_{i=1}^n$. Для линейной функции $y = \beta_0 + \beta_1 x$

1. применив метод наименьших квадратов выведите систему нормальных уравнений для нахождения параметров (оптимальной) прямой, наименее уклоняющейся от заданных наблюдений (точек на плоскости);
2. выведите формулы для оценок $\hat{\beta}_0$ и $\hat{\beta}_1$ коэффициентов оптимальной прямой;
3. покажите, что для оценок коэффициентов верно

$$\hat{\beta}_1 = \frac{\widehat{\text{cov}}(x, y)}{\widehat{\text{Var}}(x)}, \quad \hat{\beta}_0 = \bar{y} - \hat{\beta}_1 \cdot \bar{x}.$$

№2. Пусть задано n наблюдений (точек на плоскости) $\{x_i, y_i\}_{i=1}^n$. Для линейной функции $y = \beta x$

1. применив метод наименьших квадратов выведите систему нормальных уравнений для нахождения параметров (оптимальной) прямой, наименее уклоняющейся от заданных наблюдений (точек на плоскости);
2. выведите формулы для оценки $\hat{\beta}$ коэффициента оптимальной прямой.

№3. Вывите систему нормальных уравнений для функций

$$y = \beta_0 + \beta_1 x + \beta_2 z, \quad y = \beta_0 + \beta_1 x + \beta_2 x^2$$

№4. Пусть $\hat{\beta}$ есть OLS-оценка коэффициента наклона линейной функции y на x без константы, а $\hat{\gamma}$ – OLS-оценка коэффициента наклона в линейной функции x на y без константы. Верно ли для этих оценок равенство

$$\hat{\gamma} = \frac{1}{\hat{\beta}}?$$

№5. Пусть $\hat{\beta}_1$ есть OLS-оценка коэффициента наклона линейной функции y на x с константой, а $\hat{\gamma}_1$ – OLS-оценка коэффициента наклона линейной функции x на y с константой. Верно ли равенство

$$\hat{\gamma}_1 = \frac{1}{\hat{\beta}_1}?$$

Ответ поясните.

№6. Пусть $\hat{\beta}_1$ есть OLS-оценка коэффициента наклона линейной функции y на x с константой, а $\hat{\gamma}_1$ – OLS-оценка коэффициента наклона линейной функции x на y с константой. Покажите, что

$$\hat{\gamma}_1 = \frac{1}{\hat{\beta}_1} \iff \widehat{\text{corr}}(x, y) = \pm 1.$$

№7. Пусть $\hat{\beta}_0, \hat{\beta}_1$ – OLS-оценки коэффициентов линейной функции y на x , а $\tilde{\beta}_0, \tilde{\beta}_1$ – OLS-оценки коэффициентов линейной функции $(c_1 y)$ на $(c_2 x)$ ($c_1, c_2 \neq 0$). Покажите, что

$$\tilde{\beta}_1 = \frac{c_1}{c_2} \cdot \hat{\beta}_1, \quad \tilde{\beta}_0 = c_1 \hat{\beta}_0.$$

№8. Пусть $\hat{\beta}_0, \hat{\beta}_1$ – OLS-оценки коэффициентов линейной функции y на x , а $\tilde{\beta}_0, \tilde{\beta}_1$ – OLS-оценки коэффициентов линейной функции $(y + c_1)$ на $(x + c_2)$. Покажите, что

$$\tilde{\beta}_1 = \hat{\beta}_1, \quad \tilde{\beta}_0 = \hat{\beta}_0 + c_1 - c_2 \hat{\beta}_1.$$

№9. Пусть $\hat{\beta}_0, \hat{\beta}_1$ – OLS-оценки коэффициентов линейной функции y на x , а $\hat{\gamma}_0, \hat{\gamma}_1$ – OLS-оценки коэффициентов линейной функции $(y + cx)$ на x ($c \neq 0$). Как связаны $\hat{\beta}_0, \hat{\beta}_1$ и $\hat{\gamma}_0, \hat{\gamma}_1$

№10. Пусть $\widehat{\beta}_0, \widehat{\beta}_1$ – OLS-оценки коэффициентов линейной функции y на x , а $\widehat{\gamma}_0, \widehat{\gamma}_1$ – OLS-оценки коэффициентов линейной функции $y - \bar{y}$ на $x - \bar{x}$. Покажите, что $\widehat{\gamma}_0 = 0$ и $\widehat{\gamma}_1 = \widehat{\beta}_1$.

№11. Рассмотрим линейную функцию $y = \beta_0 + \beta_1 x_1 + \beta_2 x_2$. Покажите, что систему нормальных уравнений можно записать в виде

$$\begin{aligned}\beta_0 + \beta_1 \cdot \bar{x}_1 + \beta_2 \cdot \bar{x}_2 &= \bar{y} \\ \widehat{\text{Var}}(x_1)\beta_1 + \widehat{\text{cov}}(x_1, x_2)\beta_2 &= \widehat{\text{cov}}(x_1, y) \\ \widehat{\text{cov}}(x_1, x_2)\beta_1 + \widehat{\text{Var}}(x_2)\beta_2 &= \widehat{\text{cov}}(x_2, y)\end{aligned}$$

Найдите формулы для OLS-оценок $\widehat{\beta}_0$, $\widehat{\beta}_1$ и $\widehat{\beta}_2$.

№12. Рассмотрим линейную функцию $y = \beta_1 x_1 + \beta_2 x_2$.

1. Выведите систему нормальных уравнений для нахождения OLS-оценок коэффициентов.
2. Найдите явные формулы для OLS-оценок коэффициентов.

Глава 2

Стандартная линейная модель регрессии

Для вычисления p -значений используйте следующие функции MS Excel

Простая гипотеза (двустор. альтернатива)	СТЬЮДРАСП($t, k, 2$)
Простая гипотеза (одностор. альтернатива)	СТЬЮДРАСП($t, k, 1$)
Сложная гипотеза	ФРАСП(F, k_1, k_2)

№1. Для выборки

y	2	-1	1	3	-2	0
x_1	1	-1	2	0	2	1
x_2	1	2	-1	2	1	2

1. Напишите матрицы \mathbf{X} и \mathbf{y} для модели регрессии $y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \text{error}$.
2. Запишите в матричном виде систему нормальных уравнений.

№2. Рассмотрим парную модель линейной регрессии без константы

$$y_i = \beta x_i + u_i.$$

Покажите, что для этой модели выполнено равенство

$$\sum_{i=1}^n y_i^2 = \sum_{i=1}^n \hat{y}_i^2 + \sum_{i=1}^n e_i^2,$$

где $\hat{y}_i = \hat{\beta} x_i$ – предсказанные значения и $e_i = y_i - \hat{y}_i$ – остатки.

10 ГЛАВА 2. СТАНДАРТНАЯ ЛИНЕЙНАЯ МОДЕЛЬ РЕГРЕССИИ

№3. Была оценена модель регрессии зависимости стоимости дома *price* (в \$) от площади дома *area* (кв. м), числа комнат *rooms*, площади участка *land* (кв. м), числа ванных комнат *baths* и возраста дома *age*

Зависимая переменная: *price*; объем выборки 321

	const	<i>area</i>	<i>rooms</i>	<i>land</i>	<i>baths</i>	<i>age</i>	<i>age</i> ²
coeff	12396.19	29.90904	1793.846	0.0412428	6818.556	-664.3816	2.358722
Std. Err	15138.56	3.356915	2462.598	0.0433943	3831.339	176.1998	1.131844

$$R^2 = 0.5407 \quad R_{adj}^2 = 0.5319$$

1. Дайте интерпретацию коэффициентов при факторах *area*, *rooms*, *land* и *baths*.
2. Тестируйте значимость влияния числа ванных комнат. Какая альтернатива предпочтительней: односторонняя или двухсторонняя? Ответ поясните.
3. Дайте интерпретацию R^2 .
4. Тестируйте значимость регрессии «в целом».

№4. По четырехнедельным историческим данным была оценена регрессионная модель зависимости спроса на мороженное *consumption* (в пинтах на человека) от среднего недельного дохода домохозяйства *income* (в \$), цены за пинту *price* (в \$) и средней температуры *temp* (по Фаренгейту)

Зависимая переменная: *consumption*; объем выборки 30

	const	<i>income</i>	<i>price</i>	<i>temp</i>
coeff	0.1973149	0.0033078	-1.044413	0.0034584
Std. Err	0.2702161	0.0011714	0.834357	0.0004455

$$R^2 = 0.7890 \quad R_{adj}^2 = 0.6866$$

1. Дайте интерпретацию коэффициентов при факторах *income*, *price* и *temp*.
2. Проверьте значимость каждого из коэффициентов. Какие можно сделать выводы?

3. Дайте интерпретацию R^2 .
4. Тестируйте значимость регрессии «в целом».

№5. Была оценена регрессионная модель зависимости продолжительности недельного сна *sleep* (в мин) от недельной занятости *totwrk* (в мин), возраста *age*, уровня образования *educ*, почасовой оплаты труда *hrwage*, числа детей *yngkid* в возрасте меньше 3 лет и бинарной переменной *male* (гендерный фактор)

Зависимая переменная: *sleep*; объем выборки 532

	const	<i>totwrk</i>	<i>male</i>	<i>age</i>	<i>educ</i>	<i>hrwage</i>	<i>yngkid</i>
coeff	3593.091	-0.1571236	42.53493	1.613842	-10.58972	0.1303132	67.42672
Std. Err	139.4089	0.0208295	41.65473	1.738797	7.140286	5.417981	54.30572

$$R^2 = 0.1154 \quad R^2_{adj} = 0.1053$$

1. Дайте интерпретацию всех коэффициентов.
2. Тестируйте значимость влияния числа детей. Какая альтернатива предпочтительней: односторонняя или двухсторонняя? Ответ поясните.
3. Проверьте значимость влияния гендерного фактора. Какие можно сделать выводы?
4. Дайте интерпретацию R^2 .
5. Тестируйте значимость регрессии «в целом».

№6. Рассмотрим линейную модель регрессии, количественно описывающую зависимость продолжительности недельного сна *sleep* (в мин) от недельной занятости *totwrk* (в мин), возраста *age*, уровня образования *educ*, гендерного фактора *gender*, семейного статуса *marr*, почасовой оплаты *hrwage* и числа детей младше 3 лет *yngkid*

Дайте описание теста для проверки совместной значимости гендерного фактора, семейного статуса и числа детей.

12 ГЛАВА 2. СТАНДАРТНАЯ ЛИНЕЙНАЯ МОДЕЛЬ РЕГРЕССИИ

№7. Была оценена модель регрессии зависимости зарплаты CEO $salary$ (тыс. \$) от объема продаж фирмы $sales$ (млн \$), доходности на собственный капитал roe (в %), доходности акций ros (в %) и фиктивных переменных $indus$ и $finance$ (принадлежит ли фирма индустриальному или финансовому сектору)

Зависимая переменная: $\ln(salary)$; объем выборки 209

	const	$\ln(sales)$	roe	ros	$indus$	$finance$
coeff	4.232828	0.2806496	0.0188655	0.0003348	0.0110863	0.19038
Std. Err	0.3178304	0.0350516	0.0041113	0.0005569	0.0792622	0.0886405

$$R^2 = 0.3006 \quad R_{adj}^2 = 0.2834$$

1. Дайте интерпретацию коэффициентов модели.
2. Проверьте значимость коэффициентов. Какие можно сделать выводы?
3. Дайте интерпретацию R^2 .
4. Проверьте значимость регрессии «в целом» и сформулируйте проверяемую статистическую гипотезу. Как интерпретируется проверяемая статистическая гипотеза?

№8. Была оценена регрессионная модель зависимости издержек (фактор $COST$) энергетических компаний от объемов производства KWH (100 кВт. в ч), географического положения $South$ (бинарная переменная) и от цен на рабочую силу PL , топливо PF и капитал PK .

Зависимая переменная: $\ln(COST)$; объем выборки 23

	const	KWH	$\ln(PL)$	$\ln(PF)$	$\ln(PK)$	$South$
coefficient	2.3	0.002	0.21	0.31	0.34	-0.03
Std. Error	1.45	0.0008	0.07	0.13	0.07	0.0014

$$ESS = 250.4 \quad RSS = 230.6$$

Предполагается, что ошибка модели регрессии удовлетворяет условиям Гаусса–Маркова. Уровень значимости 5%.

- a) Дайте интерпретацию коэффициентов модели.
- b) Тестируйте значимость влияния объемов производства. Какая альтернатива предпочтительней: односторонняя или двухсторонняя? Ответ поясните.
- c) Вычислите p -значение для проверки значимости влияния объемов производства.
- d) Значимо ли эластичность издержек по цене на топливо отличается от 0.25?
- e) Вычислите коэффициент R^2 и дайте его интерпретацию.
- f) Проверьте значимость регрессии «в целом» и сформулируйте проверяемую статистическую гипотезу. Как интерпретируется проверяемая статистическая гипотеза?
- g) Вычислите p -значение для проверки значимости регрессии «в целом».
- h) Напишите спецификацию регрессии с ограничениями для проверки статистической гипотезы $H_0 : \beta_2 = \beta_3 = \beta_4$. Дайте интерпретацию нулевой гипотезе.
- i) Была оценена регрессия с ограничениями и ее остаточная сумма квадратов равна 258.2. Тестируйте гипотезу из предыдущего пункта.
- j) Вычислите p -значение для тестирования гипотезы из пункта h).

№9. Была оценена регрессионная модель зависимости цены дома $Price$ (в \$1000) от его площади $hsize$ (в m^2), площади участка вокруг дома $lsize$ (в m^2), числа ванных комнат $bath$, числа спален bdr и бинарной переменной $poor$ (равна единице, если состояние дома оценивается как плохое)

Зависимая переменная: $Price$; объем выборки 33

	const	$hsize$	$lsize$	$bath$	bdr	$poor$	
coefficient	121.3	0.13	0.001	12.1	0.42	-50.7	ESS = 350.4
Std. Error	12.2	0.012	0.00076	4.13	1.3	19.7	RSS = 533.2

Предполагается, что ошибка модели регрессии удовлетворяет условиям Гаусса–Маркова.

14 ГЛАВА 2. СТАНДАРТНАЯ ЛИНЕЙНАЯ МОДЕЛЬ РЕГРЕССИИ

- a) Дайте интерпретацию коэффициентов модели.
- b) Тестируйте значимость влияния числа спален. Какая альтернатива предпочтительней: односторонняя или двухсторонняя? Ответ поясните.
- c) Вычислите p -значение для проверки значимости влияния числа спален.
- d) Значимо ли предельный эффект от изменения площади дома отличается от 0.2?
- e) Вычислите коэффициент R^2 и дайте его интерпретацию.
- f) Проверьте значимость регрессии «в целом» и сформулируйте проверяемую статистическую гипотезу. Как интерпретируется проверяемая статистическая гипотеза?
- g) Вычислите p -значение для проверки значимости регрессии «в целом».
- h) Напишите спецификацию регрессии с ограничениями для проверки статистической гипотезы $H_0 : \beta_3 = \beta_4$. Дайте интерпретацию нулевой гипотезе.
- i) Была оценена регрессия с ограничениями и ее остаточная сумма квадратов равна 560.3. Тестируйте гипотезу из предыдущего пункта.
- j) Вычислите p -значение для тестирования гипотезы из пункта h).

№10. На основе опроса 25–35 летних была оценена регрессионная модель зависимости почасовой оплаты $Wage$ от уровня образования $educ$, возраста age , членства в профсоюзе $union$ (бинарная переменная), пола $female$ (бинарная переменная) и семейного статуса $married$ (бинарная переменная)

Зависимая переменная: $\ln(Wage)$; объем выборки 26

	const	$educ$	age	$union$	$female$	$married$	$R^2 = 0.58$
coefficient	1.82	0.031	0.018	0.039	-0.04	-0.03	
Std. Error	0.49	0.011	0.006	0.013	0.001	0.04	

Предполагается, что ошибка регрессии удовлетворяет условиям Гаусса–Маркова.

- a) Дайте интерпретацию коэффициентов модели.
- b) Тестируйте значимость влияния возраста. Какая альтернатива предпочтительней: односторонняя или двухсторонняя? Ответ поясните.
- c) Вычислите p -значение для проверки значимости влияния возраста.
- d) Значимо ли отдача от образования отличается от 0.05?
- e) Дайте интерпретацию коэффициента R^2 .
- f) Проверьте значимость регрессии «в целом» и сформулируйте проверяемую статистическую гипотезу. Как интерпретируется проверяемая статистическая гипотеза?
- g) Вычислите p -значение для проверки значимости регрессии «в целом».
- h) Напишите спецификацию регрессии с ограничениями для проверки статистической гипотезы $H_0 : \beta_4 = \beta_5 = 0$. Дайте интерпретацию нулевой гипотезе.
- i) Была оценена регрессия с ограничениями и для нее коэффициент R^2 равен 0.045. Тестируйте гипотезу из предыдущего пункта.
- j) Вычислите p -значение для тестирования гипотезы из пункта h).

№11. В Eviews была оценена производственная функция Кобба – Ду-гласа:

Dependent variable: $\ln(Q)$; Sample: 1 142

	const	$\ln(K)$	$\ln(L)$	$\ln(H)$
coefficient	3.45	0.47	0.12	0.34
Std. Error	1.45	0.23	0.56	0.13
t-Statistic	–	–	–	–
Prob.	0.0210	0.0461	0.8312	0.0116

- a) Дайте интерпретацию коэффициентов модели.
- b) Заполните прочерки.

16 ГЛАВА 2. СТАНДАРТНАЯ ЛИНЕЙНАЯ МОДЕЛЬ РЕГРЕССИИ

- с) Проверьте значимость коэффициентов при уровне значимости 5%.
- d) Запишите оцененное уравнение регрессии.
- e) Дайте экономическую интерпретацию гипотезе $H_0 : \beta_1 + \beta_2 + \beta_3 = 1$.
- f) Пусть F -статистика для гипотезы из предыдущего пункта равна 1.76. Тестируйте гипотезу. Вычислите p -значение.

№12. Для мужчин возрастной группы 28–38 лет оценили модели регрессии, описывающие зависимость почасовой оплаты *wage* от возраста *age*, стажа на текущем месте работы *tenure*, семейного статуса *married* (бинарная переменная), места жительства *urban* (бинарная переменная), количества братьев и сестер *sibs*, результатов IQ-теста (в скобках указаны t -статистики).

Зависимая переменная: $\ln(wage)$; объем выборки 935

	const	<i>age</i>	<i>age</i> ²	<i>tenure</i>	<i>married</i>	<i>urban</i>	<i>sibs</i>	<i>iq</i>	R^2
№1	3.3868 (2.0480)	0.1166 (1.1701)	-0.0015 (-1.0048)	0.0113 (4.4614)	0.1901 (4.7400)	0.1879 (6.8530)	-0.0092 (-1.6526)	0.0082 (9.6327)	0.2085
№2	5.5812 (53.7947)			0.0140 (0.0025)	0.2041 (5.0699)	0.1902 (6.8872)	-0.0106 (-1.8818)	0.0080 (9.2794)	0.1939
№3	5.0396 (29.6252)	0.0166 (3.9984)		0.0113 (4.4663)	0.1897 (4.7295)	0.1893 (6.9086)	-0.0095 (-1.6983)	0.0082 (9.6038)	0.2076
№4	2.7556 (1.6302)	0.1577 (1.5464)	-0.0021 (-1.3791)	0.0107 (4.1183)	0.1801 (4.3819)			0.0089 (10.5712)	0.1656

1. Дайте интерпретацию коэффициентов третьей модели.
2. Проверьте совместную значимость факторов возраст и квадрат возраста.
3. Проверьте совместную значимость факторов *urban* и *sibs*.

№13. Рассмотрим регрессионную модель зависимости оценки за финальный экзамен *exam* от оценки за промежуточный срез *midterm* и времени, отведенного на выполнение экзаменационного задания *time*:

$$exam = \beta_1 midterm + \beta_2 time + \text{error}.$$

Выведите систему нормальных уравнений для нахождения оценок наименьших квадратов коэффициентов модели. Найдите формулы для оценок наименьших квадратов коэффициентов модели.

№14. Исследуется зависимость логарифма зарплаты $\ln(Wage)$ от уровня образования $Educ$, опыта работы $Experience$ и уровня образования родителей $Feduc$ и $Meduc$

$$\widehat{\ln(Wage)} = \beta_0 + \beta_1 Educ + \beta_2 Experience + \beta_3 Experience^2 + \beta_4 Feduc + \beta_5 Meduc, \quad n = 500 \quad R^2 = 0.273$$

- Напишите спецификацию модели регрессии с ограничениями для проверки статистической гипотезы $H_0 : \beta_4 = \beta_5 = 0$.
- Дайте интерпретацию статистической гипотезы из п. а).
- Для модели регрессии с ограничениями из п. а) был вычислен $R_r^2 = 0.262$. Тестируйте нулевую гипотезу при уровне значимости 5%.
- Вычислите скорректированный коэффициент \bar{R}^2 для исходной модели. Сравните обе модели (с ограничениями и без ограничений) по критерию \bar{R}^2 .

№15. По предприятиям энергетической отрасли была оценена производственная функция Кобба–Дугласа (K – капитал, L – труд) по данным за февраль 2006 года.

Зависимая переменная: $\ln(Q)$; объем выборки 30

	const	$\ln(K)$	$\ln(L)$	$R^2 = 0.45$
coefficient	3.45	0.47	0.32	$\text{Prob}(F\text{-stat})=0.00031$
Prob.	0.0210	0.461	0.528	$DW = 3.43$

- Проверьте значимость влияния капитала, значимость влияния труда, совместную значимость влияния труда и капитала.
- Как можно объяснить результаты предыдущего пункта? Дайте не менее двух возможных объяснений.

№16. Для мужчин возрастной группы 28–38 лет был оценена модель регрессии зависимости почасовой оплаты $wage$ от возраста age , стажа на текущем месте работы $tenure$, семейного статуса $married$, места жительства $urban$, количества братьев и сестер $sibs$, и результатов IQ-теста.

Зависимая переменная: $\ln(wage)$; объем выборки 935

18 ГЛАВА 2. СТАНДАРТНАЯ ЛИНЕЙНАЯ МОДЕЛЬ РЕГРЕССИИ

	const	age	age ²	tenure	married	urban	sibs	iq
coeff.	3.3868	0.1166	-0.0015	0.0113	0.1901	0.1879	-0.0092	0.0082
Prob.	0.0408	0.42423	0.3153	0.0000	0.0000	0.0000	0.0987	0.0000

$$R^2 = 0.2085 \quad F\text{-statistic}=34.8769 \quad \text{Prob}(F\text{-stat})=0.0000$$

1. Дайте интерпретацию коэффициентов модели регрессии.
2. Дайте интерпретацию коэффициента R^2 и проверьте значимость регрессии «в целом».
3. Проверьте значимость влияния фактора возраст и фактора квадрат возраста.
4. F-статистика для проверки совместной значимости age и age^2 равна 8.4986. Значимо ли совместное влияние факторов возраст и квадрат возраста?
5. Как можно объяснить результаты, полученные в п.3 и п.4? Дайте развернутый ответ.

№17. Предполагается, что влияние инвестиций на экономический рост (ВВП) различно в развитых и развивающихся странах. Какое эконометрическое исследование надо провести, чтобы проверить этот факт.

№18. Риэлтора интересует, насколько стоимость 1 кв.метра жилой площади в квартирах в монолитных домах отличается в среднем от стоимости 1 кв.метра жилой площади в квартирах в кирпичных домах. Какое эконометрическое исследование он должен провести.

№19. Предприниматель хочет оценить эффективность влияния стоимости капитала и труда на выпуск продукции, а также стоит ли ему в дальнейшем стимулировать труд, или вклад факторов отличается существенно. Какое эконометрическое исследование надо провести, чтобы это проверить.

№20. Считается, что влияние цены на объем продаж фруктов существенно зависит от сезонности. Какое эконометрическое исследование надо провести, чтобы выяснить, меняется ли цена в зависимости от сезона и является ли сезонность вообще значимым фактором.

№21. В модель регрессии были добавлены два новых фактора. Как изменились обобщая, объясненная и остаточная суммы квадратов?

№22. Рассмотрим две модели регрессии

$$y = \beta_0 + \beta_1 x + \text{error}$$
$$y - x = \gamma_0 + \gamma_1 x + \text{error}$$

1. Как связаны OLS-оценки коэффициентов первой и второй регрессии?
2. Будут ли коэффициенты R^2 в первой и во второй регрессии равны?

Глава 3

Спецификация модели регрессии

№1. Исследуется влияние образования $Educ$, пола $Male$ и возраста Age на почасовую оплату труда $Wage$ для возрастной группы 25 – 75 лет. Рассматривается модель регрессии

$$\ln Wage = \beta_0 + \beta_1 Educ + \beta_2 Male + \beta_3 Age + \beta_4 Age^2 + \text{error}$$

С какой целью в модель регрессии был включен квадрат возраста? Какой ожидаемый знак у оценки $\hat{\beta}_4$? Как можно объяснить включение именно *квадрата* возраста (а не, например, третьей степени)?

Как можно объяснить использования в качестве зависимой переменной логарифма зарплаты?

№2. Предложите и экономически обоснуйте модель регрессии для оценки влияния на почасовую оплату $wage$ возраста age , гендерного признака $female$ и недельной занятости $hours$

1. для возрастной группы 30–37 лет
2. для возрастной группы 30–80 лет

Дайте интерпретацию параметров модели

№3. Рассмотрим модель регрессии зависимости цены автомобиля $Price$ от его возраста Age . Какая модель предпочтительней: линейная или полупологарифмическая? Какие преимущества у полупологарифмической модели?

Эконометрист полагает, что скорость падения цены автомобиля с возрастом уменьшается. Как учесть это предположение в модели регрессии?

№4. Предложите и экономически обоснуйте модель регрессии для оценки влияния на зарплату CEO (фактор *Salary*) объема продаж фирмы *sales*, ее рыночной стоимости *mktval*, продолжительности работы в фирме *ceotenure*, возраста *age*. Дайте интерпретацию параметров модели.

№5. Предложите и экономически обоснуйте модель регрессии для оценки влияния на зарплату CEO (фактор *salary*) объема продаж фирмы *sales*, доходности на собственный капитал *roe* (в %) и доходности акций *ros* (также в %) если

1. модель регрессии рассматривается только для фирм с объемом продаж \$200 млн– \$250 млн в год
2. модель регрессии рассматривается для фирм с объемами продаж \$10 млн – \$1 млрд в год

Дайте интерпретацию параметров модели.

№6. Рассматриваются следующие регрессионные модели зависимости зарплаты CEO (*S*) от годового уровня продаж фирмы (*Sales*), дохода на собственный капитал *roe* и доходности акций *ros* ($n = 7340$):

$$\widehat{\ln(S)} = \beta_0 + \beta_1 \ln(Sales) + \beta_2 roe + \beta_3 ros \quad R^2 = 0.2129$$

$$\widehat{S} = \beta_0 + \beta_1 Sales + \beta_2 Sales^2 + \beta_3 roe + \beta_4 roe^2 + \beta_5 ros \quad R^2 = 0.2139$$

Можно ли сравнить эти модели по критерию R^2 или \bar{R}^2 ? Ответ обоснуйте. Если можно, то в пользу какой модели можно сделать выбор?

№7. Для мужчин возрастной группы 28–38 лет оценили модели регрессии, описывающие зависимость почасовой оплаты *wage* от возраста *age*, стажа на текущем месте работы *tenure*, семейного статуса *married* (бинарная переменная), места жительства *urban* (бинарная переменная), количества братьев и сестер *sibs*, результатов IQ-теста (в скобках указаны *t*-статистики).

Зависимая переменная: $\ln(wage)$; объем выборки 935

	const	age	age ²	tenure	married	urban	sibs	iq	R ²
№1	3.3868 (2.0480)	0.1166 (1.1701)	-0.0015 (-1.0048)	0.0113 (4.4614)	0.1901 (4.7400)	0.1879 (6.8530)	-0.0092 (-1.6526)	0.0082 (9.6327)	0.2085
№2	5.5812 (53.7947)			0.0140 (0.0025)	0.2041 (5.0699)	0.1902 (6.8872)	-0.0106 (-1.8818)	0.0080 (9.2794)	0.1939
№3	5.0396 (29.6252)	0.0166 (3.9984)		0.0113 (4.4663)	0.1897 (4.7295)	0.1893 (6.9086)	-0.0095 (-1.6983)	0.0082 (9.6038)	0.2076
№4	2.7556 (1.6302)	0.1577 (1.5464)	-0.0021 (-1.3791)	0.0107 (4.1183)	0.1801 (4.3819)			0.0089 (10.5712)	0.1656

Какая модель регрессии предпочтительней?

№8. Рассмотрим линейную регрессионную модель зависимости зарплаты $Wage$ от уровня школьного образования $Educ$ (в годах), возраста Age , наличия высшего образования $Heduc$ (фиктивная переменная), школьного образования родителей $Feduc, Meduc$ и пола $Male$:

$$\ln(\widehat{Wage}) = \beta_0 + \beta_1 Educ + \beta_2 Age + \beta_3 Heduc + \beta_4 Feduc + \beta_5 Meduc + \beta_6 Male \quad n = 500 \quad R^2 = 0.302$$

- Пусть $\widehat{\beta}_3 = 0.34$. Дайте интерпретацию полученной оценке.
- Пусть $\widehat{\beta}_6 = 0.02$ и $s_6 = 0.002$. Дайте интерпретацию полученной оценке. Значима ли дискриминация в оплате труда между мужчинами и женщинами на 1%-м уровне значимости?
- Для тестирования справедливости линейной спецификации обычно используют RESET-тест. Напишите спецификацию вспомогательной регрессии для применения RESET-теста (с включением вспомогательных переменных до 4-й степени). Сформулируйте проверяемую статистическую гипотезу.
- Пусть во вспомогательной регрессии $R_1^2 = 0.361$. Тестируйте нулевую гипотезу из пункта с) при уровне значимости 1%. Какой можно сделать вывод?
- Можно ли доверять выводам, сделанным в пп. а) и б)?

№9. Рассмотрим регрессионную модель зависимости зарплаты CEO (S) от годового уровня продаж фирмы ($sales$) и дохода на собственный капитал roe :

$$\ln(S) = \beta_0 + \beta_1 \ln(Sales) + \beta_2 roe + error \quad n = 22 \quad R^2 = 0.213$$

Напишите спецификацию вспомогательной регрессии для применения RESET-теста на справедливость данной спецификации (с включением вспомогательных переменных до 4-й степени). Сформулируйте проверяемую статистическую гипотезу. Пусть во вспомогательной регрессии $R_1^2 = 0.421$. Тестируйте нулевую гипотезу при уровне значимости 5%. Какой можно сделать вывод?

№10. Рассмотрим линейную регрессионную модель зависимости продолжительности сна *sleep* (мин в неделю) от недельной занятости *totwrk* (в мин), пола *male*, количества детей *kids* и уровня образования *educ*.

- Опишите процедуру RESET-теста и укажите проверяемую статистическую гипотезу
- Если в RESET-тесте нулевая гипотеза отвергается, то какие возможны корректировки модели (приведите не менее двух)?

№11. Для количественного описания зависимости зарплаты CEO (фактор *salary*) от объема продаж фирмы *sales*, ее рыночной стоимости *mktval* и стажа работы в должности в фирме *ceoten* рассматривается модель регрессии

$$\ln(\text{salary}) = \beta_0 + \beta_1 \ln(\text{sales}) + \beta_2 \ln(\text{mktval}) + \beta_3 \text{ceoten} + \text{error}$$

Дайте описание теста на согласованность выборочных данных с этой спецификацией модели регрессии.

№12. В декабре 2007 были оценены (на одних и тех же данных) две модели регрессии зависимости зарплаты CEO S от годового уровня продаж фирмы *sales*, дохода на собственный капитал *roe* и доходности акций *ros*

Модель №1: зависимая переменная $\ln(S)$, объем выборки 27

	const	$\ln(\text{sales})$	<i>roe</i>	<i>ros</i>
coefficient	2.4	0.14	0.008	0.0013
Std. Error	1.02	0.013	0.0023	0.0003

$$R^2 = 0.101 \quad DW = 0.23$$

Модель №2: зависимая переменная $\ln(S)$, объем выборки 27

	const	$\ln(\text{sales})$	roe	ros	roe^2	ros^2
coefficient	2.31	0.18	0.006	0.0009	-0.00003	-0.0000078
Std. Error	1.02	0.026	0.0024	0.00034	0.00004	0.0000063

$$R^2 = 0.127 \quad DW = 2.03$$

- a) По каким критериям можно сравнить эти модели? Укажите не менее двух и дайте их описание.
- b) Какая модель предпочтительней? Ответ обоснуйте.

№13. Рассматривается регрессионная модель зависимости недельной занятости $hours$ от почасовой оплаты $wage$ и уровня образования $educ$

$$hours = \beta_0 + \beta_1 wage + \beta_2 wage^2 + \beta_3 educ + \text{error}$$

- a) Как можно объяснить включение в модель регрессии квадрата (и именно квадрата) почасовой оплаты? Какой ожидаемый знак у оценки коэффициента $\hat{\beta}_2$? Ответ обоснуйте.
- b) Предположим, что при оценке модели регрессии знак оценки коэффициента при квадрате получился противоположным ожидаемому. Как это можно объяснить? Дайте несколько возможных объяснений (не менее двух).

Глава 4

Линейная регрессия с гетероскедастичной ошибкой

№1. Рассмотрим модель регрессии

$$y_i = \beta x_i + u_i$$

для ошибки которой выполнены условия

$$E(u_i) = 0, \quad \text{cov}(u_i, u_j) = 0 \ (i \neq j), \quad \text{Var}(u_i) = \sigma^2 x_i^2.$$

- Будет ли OLS-оценка коэффициента β несмещенной? Состоятельной? Оптимальной?
- Предложите несмещенную оценку коэффициента β с меньшей дисперсией. Как это соотносится с теоремой Гаусса–Маркова?

№2. Рассмотрим модель регрессии $y_i = \beta_0 + \beta_1 x_i + \varepsilon_i$ для ошибок которой выполнено

$$E\varepsilon_i = 0, \quad \text{Var}(\varepsilon_i) = \sigma^2 \exp(10x_i), \quad \text{cov}(\varepsilon_i, \varepsilon_j) = 0 \ (i \neq j).$$

Пусть $\hat{\beta}_1$ есть OLS-оценка коэффициента β_1 .

1. Будет ли $\hat{\beta}_1$ несмещенной оценкой β_1 ? Состоятельной оценкой?
2. Будет ли $\hat{\beta}_1$ наилучшей оценкой коэффициента β_1 (с минимальной дисперсией)? Если нет, то предложите оценку коэффициента с меньшей дисперсией (более точную оценку).

28 ГЛАВА 4. ЛИНЕЙНАЯ РЕГРЕССИЯ С ГЕТЕРОСКЕДАСТИЧНОЙ ОШИБКОЙ

№3. Рассмотрим модель регрессии $y_i = \beta_0 + \beta_1 x_i + u_i$ ($i = 1, \dots, n$) для ошибок которой выполнены условия (δ_0, δ_1 неизвестны и $\delta_1 \neq 0$)

$$Eu_i = 0 \quad \text{cov}(u_i, u_j) = 0 \quad (i \neq j) \quad \text{Var}(u_i) = \exp(\delta_0 + \delta_1 x_i)$$

- Будут ли оценки наименьших квадратов коэффициентов в этой модели регрессии эффективными? Ответ поясните.
- Если оценки наименьших квадратов не будут эффективными, то как получить более эффективные оценки коэффициентов (при больших выборках)?

№4. Исследуется зависимость цена коттеджа Price от удаленности от МКАД Dist, площади Hsize и площади участка Lsize. Эконометрист полагает, что дисперсия ошибки в линейной модели регрессии зависит от площади дома и площади участка.

- Опишите процедуру теста Breusch–Pagan’a и укажите проверяемую статистическую гипотезу.
- Можно ли применять этот тест при небольших выборках?
- Если тест отвергает нулевую гипотезу, то какие могут быть причины?

№5. По методу наименьших квадратов была оценена регрессионная модель зависимости недельной занятости hours от почасовой оплаты труда wage (в \$100) и уровня образования educ

Зависимая переменная: hours; объем выборки 1321

	const	wage	wage ²	educ
coefficient	23.4	0.84	-0.008	-0.04
Std. Error	1.02	0.013	0.0027	0.023

Статистика теста White’a равна 6.34.

- Можно ли для прогнозирования занятости использовать полученные оценки коэффициентов? Ответ обоснуйте.
- Можно ли проверять значимость влияния уровня образования на занятость с использованием стандартной t -статистики? Ответ поясните.

№6. По методу наименьших квадратов была оценена регрессионная модель зависимости недельной занятости $hours$ от почасовой оплаты труда $wage$ (в \$100) и уровня образования $educ$

Зависимая переменная: $hours$; объем выборки 5321

	const	$wage$	$wage^2$	$educ$	
coefficient	23.4	0.84	-0.008	-0.04	ESS = 380.12
Std. Error	1.02	0.013	0.0027	0.023	RSS = 130.23

Статистика теста Breusch–Pagan'a на зависимость дисперсии ошибки модели регрессии от почасовой оплаты и пола человека равна 8.75.

- Можно ли для прогнозирования занятости использовать полученные оценки коэффициентов? Ответ обоснуйте.
- Можно ли проверять значимость влияния уровня образования на занятость с помощью стандартной t -статистики? Ответ поясните.

№7. Для количественного описания зависимости зарплаты CEO ($salary$) от объема продаж фирмы $sales$, её рыночной стоимости $mktval$ и стажа работы в должности в фирме $ceoten$ была оценена модель регрессии

Зависимая переменная: $\ln(salary)$; объем выборки 177

	const	$\ln(sales)$	$\ln(mktval)$	$ceoten$
coefficient	4.5038	0.1629	0.1092	0.0117
OLS Std. Error	0.2572	0.0392	0.0496	0.0053
HC St. Error	0.2801	0.0382	0.0499	0.0073

- Как можно объяснить выбор такой спецификации модели?
- Проверьте значимость коэффициентов модели с использованием стандартных ошибок коэффициентов метода наименьших квадратов и стандартных ошибок, устойчивых к гетероскедастичности (стандартных ошибок White'a). Какой можно сделать вывод?

№8. Для модели регрессии из предыдущей задачи была оценена вспомогательная регрессия

Зависимая переменная: e^2 ; объем выборки 177

30 ГЛАВА 4. ЛИНЕЙНАЯ РЕГРЕССИЯ С ГЕТЕРОСКЕДАСТИЧНОЙ ОШИБКОЙ

	const	$\ln(sales)$	$\ln(mktval)$	$ceoten$
coefficient	-0.0521	-0.0805	0.0992	0.0188
Std. Error	0.3052	0.0466	0.0588	0.0063
Prob.	0.8647	0.0856	0.0938	0.0033

R-squared=0.0657 F-statistic=4.0537 Prob(F-statistic)=0.0082

1. Какие выводы можно сделать на основе этой регрессии?
2. Можно ли (экономически) интерпретировать эту модель регрессии?

№9. Для количественного описания зависимости зарплаты CEO ($salary$) от объема продаж фирмы $sales$, ее рыночной стоимости $mktval$ и стажа работы в должности в фирме $ceoten$ рассматривается модель регрессии

$$\ln(salary) = \beta_0 + \beta_1 \ln(sales) + \beta_2 \ln(mktval) + \beta_3 ceoten + \text{error}$$

Графический анализ остатков показал, что разброс остатков уменьшается с увеличением фактора $ceoten$ и нет явной зависимости разброса остатков от $\ln(sales)$ и $\ln(mktval)$.

Дайте описание теста на гетероскедастичность ошибки модели регрессии, который можно применить в данной ситуации.

№10. Для пространственных данных рассмотрим модель регрессии

$$y_i = \beta_0 + \beta_1 x_i + \beta_2 x_i^2 + \beta_3 z_i + u_i \quad E(u_i) = 0 \quad \text{Var}(u_i) = \sigma^2 x_i^2 z_i^2.$$

Как получить эффективную оценку коэффициентов β_1 , β_2 и β_3 ? Дайте развернутый ответ.

Глава 5

Линейная регрессия с автокоррелированной ошибкой

№1. Рассмотрим модель регрессии

$$y_t = \beta_0 + \beta_1 x_t + \varepsilon_t$$

для ошибок которой выполнено

$$E\varepsilon_t = 0, \quad \text{Var}(\varepsilon_t) = \sigma^2, \quad \varepsilon_t = 0.15\varepsilon_{t-1} + u_t$$

и u_t удовлетворяет условиям Гаусса–Маркова. Пусть $\hat{\beta}_1$ есть OLS-оценка коэффициента β_1 .

1. Сформулируйте условия Гаусса–Маркова для u_t .
2. Будет ли $\hat{\beta}_1$ несмещенной оценкой β_1 ? Состоятельной оценкой?
3. Будет ли $\hat{\beta}_1$ наилучшей оценкой коэффициента β_1 (с минимальной дисперсией)? Если нет, то предложите оценку коэффициента с меньшей дисперсией (более точную оценку).

№2. Оценивается кривая Филлипса с ожиданиями, связывающая инфляцию с безработицей, безработицей с единичным лагом и с инфляцией с лагом один.

1. Опишите тест на автокорреляцию первого порядка ошибок этой модели регрессии и укажите проверяемую статистическую гипотезу.

32 ГЛАВА 5. ЛИНЕЙНАЯ РЕГРЕССИЯ С АВТОКОРРЕЛИРОВАННОЙ ОШИБКОЙ

2. Можно ли применять этот тест при небольших выборках?
3. Если тест отвергает нулевую гипотезу, то какие могут быть причины?

№3. По квартальным данным за пять с половиной лет по методу наименьших квадратов была оценена функция спроса с учетом сезонности (факторы *Spr*, *Summ*, *Fall* – весна, лето, осень)

Зависимая переменная: $\ln(Q)$

	const	$\ln(P)$	<i>Spr</i>	<i>Summ</i>	<i>Fall</i>	
coefficient	4.4	0.64	-0.08	-0.13	-0.07	$R^2 = 0.13$
Std. Error	1.02	0.013	0.013	0.023	0.03	DW = 0.78

- a) Можно ли для прогнозирования объема продаж использовать полученные оценки коэффициентов модели регрессии? Ответ обоснуйте.
- b) Можно ли для проверки значимости регрессии «в целом» использовать стандартную F -статистику? Ответ обоснуйте.

№4. По четырехнедельным историческим данным была оценена регрессионная модель зависимости спроса на мороженное *consumption* (в пинтах на человека) от среднего недельного дохода домохозяйства *income* (в \$), цены за пинту *price* (в \$) и средней температуры *temp* (по Фаренгейту)

Зависимая переменная: *consumption*; объем выборки 30

	const	<i>income</i>	<i>price</i>	<i>temp</i>
coeff	0.1973149	0.0033078	-1.044413	0.0034584
Std. Err	0.2702161	0.0011714	0.834357	0.0004455

$$R^2 = 0.7890 \quad R_{adj}^2 = 0.6866 \quad DW = 1.021169$$

1. Тестируйте ошибку модели регрессии на автокорреляцию первого порядка.
2. Какие можно сделать выводы?

№5. Для временных рядов рассмотрим модель регрессии (ϵ_t – белый шум)

$$y_t = \beta_0 + \beta_1 x_t + \beta_2 z_t + u_t \quad u_t = -0.25u_{t-1} + \epsilon_t.$$

Как получить эффективную оценку коэффициентов β_1 и β_2 ? Дайте развернутый ответ.

№6. Рассмотрим макроэкономическую кейнсианскую функцию потребления без автономного потребления

$$C_t = \beta_1 Y_t + \beta_2 T_t + \beta_3 r_t + u_t,$$

где C_t – потребление, Y_t – национальный доход, T_t – налоговые поступления, r_t – ставка рефинансирования. Модель оценивается по историческим данным и предполагается, что объясняющие переменные экзогенны.

Дайте описание теста на автокоррелированность первого порядка ошибок этой модели регрессии.

34 ГЛАВА 5. ЛИНЕЙНАЯ РЕГРЕССИЯ С АВТОКОРРЕЛИРОВАННОЙ ОШИБКОЙ

Глава 6

Системы одновременных уравнений

№1. Рассмотрим кейнсианскую модель потребления

$$\begin{aligned}C_t &= \beta_0 + \beta_1 Y_t + u_t \\ Y_t &= C_t + I_t,\end{aligned}\tag{6.1}$$

где C_t – агрегированное потребление, Y_t – национальный доход, I_t – инвестиции. Коэффициент β_1 имеет смысл предельной склонности к потреблению. Предполагается, что u_t (шок потребления) удовлетворяет условиям Гаусса–Маркова и $\text{cov}(I_t, u_t) = 0$.

1. Дайте интерпретацию уравнений модели.
2. Напишите приведенную форму системы нормальных уравнений
3. Укажите алгоритм оценки коэффициентов модели.

№2. Рассмотрим модель спроса–предложения

$$Q_{dt} = \beta_0 + \beta_1 P_t + \mathbf{x}'_t \boldsymbol{\beta} + u_t\tag{6.2}$$

$$Q_{st} = \gamma_0 + \gamma_1 P_t + \mathbf{z}'_t \boldsymbol{\gamma} + w_t\tag{6.3}$$

где $\mathbf{x}'_t = (x_{1t}, \dots, x_{kt})$ – факторы, влияющие на спрос (например, цвет), $\mathbf{z}'_t = (z_{1t}, \dots, z_{lt})$ – факторы, влияющие на предложение, коэффициенты $\boldsymbol{\beta}' = (\beta_2, \dots, \beta_{k+1})$ и $\boldsymbol{\gamma}' = (\gamma_2, \dots, \gamma_{l+1})$, u_t – шоки спроса, w_t – шоки предложения. Предполагаем, что $\text{cov}(u_t, w_t) = 0$ и факторы \mathbf{x} и \mathbf{z} различаются. Рынок находится в равновесии.

1. Выразите Q и P через \mathbf{x} , \mathbf{z} и шоки u и w .
2. Оценивается функция спроса (6.2). Можно ли применить метод наименьших квадратов к первому уравнению модели? Ответ обоснуйте.
3. Предложите метод оценивания функции спроса.
4. В чем отличие модели спроса–предложения (6.2), (6.3) от модели только одной функции спроса

$$Q_{dt} = \beta_0 + \beta_1 P_t + \mathbf{x}'_t \boldsymbol{\beta} + u_t?$$

Дайте экономическую интерпретацию каждой модели.

№3. Рассмотрим эконометрическую модель, связывающую недельную занятость `hours` и месячную зарплату `wage` замужних женщин

$$\begin{aligned} \text{hours} &= \beta_0 + \beta_1 \log(\text{wage}) + \beta_2 \text{educ} + \beta_3 \text{age} + \beta_4 \text{kidslt6} + \beta_5 \text{nwifeinc} + \text{error}_1 \\ \log(\text{wage}) &= \gamma_0 + \gamma_1 \text{hours} + \gamma_2 \text{exper} + \gamma_3 \text{exper}^2 + \gamma_4 \text{educ} + \text{error}_2 \end{aligned}$$

где `educ` – уровень образования, `age` – возраст, `exper` – опыт работы, `kidslt6` – число детей до 6 лет, `nwifeinc` – прочий доход (например, доход мужа).

1. Экономиста интересует оценка коэффициента γ_1 . Дайте интерпретацию этого коэффициента.
2. Дайте экономические обоснования того, что занятость и зарплата в одном уравнении являются зависимой переменной, а в другом объясняющей переменной.
3. Выполнено ли условие идентифицируемости? Ответ обоснуйте.
4. Какой метод следует использовать для оценки коэффициента γ_1 ? Обоснуйте ответ и дайте описание метода оценки коэффициентов.

№4 (Практика). Рассмотрим систему одновременных уравнений из предыдущей задачи:

$$\begin{aligned} \text{hours} &= \beta_0 + \beta_1 \log(\text{wage}) + \beta_2 \text{educ} + \beta_3 \text{age} + \beta_4 \text{kidslt6} + \beta_5 \text{nwifeinc} + \text{error}_1 \\ \log(\text{wage}) &= \gamma_0 + \gamma_1 \text{hours} + \gamma_2 \text{exper} + \gamma_3 \text{exper}^2 + \gamma_4 \text{educ} + \text{error}_2 \end{aligned}$$

Для данных из файла `mroz.gdt`

- Оцените (раздельно) уравнения системы методом наименьших квадратов (OLS). Тестируйте значимость β_1 и γ_1 .
- Оцените каждое уравнение системы используя двухшаговый метод наименьших квадратов (2SLS). Для этого используйте метод инструментальных переменных, выбрав в качестве инструментов **все** экзогенные факторы модели. Сравните полученные оценки коэффициентов с OLS-оценками. Тестируйте значимость β_1 и γ_1 .

Глава 7

Метод инструментальных переменных

№1. Рассмотрим кейнсианскую модель потребления

$$\begin{aligned} C &= \beta_0 + \beta_1 Y + u \\ Y &= C + I, \end{aligned} \tag{7.1}$$

где C – агрегированное потребление, Y – национальный доход, I – инвестиции. Коэффициент β_1 имеет смысл предельной склонности к потреблению. Предполагается, что u (шок потребления) удовлетворяет условиям Гаусса–Маркова и $\text{cov}(I, u) = 0$.

- a) Выразите C и Y через I и u .
- b) Оценивается уравнение (7.1). Имеет ли место проблема эндогенности? Обоснуйте ответ
- c) Если на предыдущий пункт ответ положительный, то будет ли фактор I инструментом для Y ? Ответ обоснуйте.
- d) Если на предыдущие пункты ответы положительны, то найдите IV-оценку коэффициента β_1 (напишите формулу или укажите алгоритм нахождения).

№2. Модель регрессии $\ln Wage = \beta_0 + \beta_1 Educ + \beta_2 Age + \beta_3 Age^2 + u$ описывает зависимость почасовой оплаты от уровня образования и возраста человека. Но ошибка u в этой модели содержит (ненаблюдаемый) фактор

«индивидуальные способности человека», который положительно коррелирует с уровнем образования.

- а) Можно ли для оценки отдачи от образования использовать метод наименьших квадратов? Ответ обоснуйте.
- б) Для учета фактора «индивидуальные способности человека» в модель регрессии была включена объясняющая переменная $iqscores$ (результат IQ-теста), которая, как предполагает эконометрист, приближенно оценивает индивидуальные способности человека (т.е. $iqscores$ есть прокси-переменная). Можно ли для оценки отдачи от образования использовать метод наименьших квадратов для новой модели регрессии (с включенным фактором $iqscores$)? Ответ поясните.

№3. Рассмотрим модель регрессии

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + u, \quad \text{cov}(x_1, u) = 0 \quad \text{cov}(x_2, u) \neq 0.$$

Будет ли оценка наименьших квадратов $\hat{\beta}_1$ коэффициента β_1 состоятельной? Ответ кратко поясните. Пусть z – инструмент для фактора x_2 . Каким условиям должен удовлетворять инструмент? Опишите двухшаговый метод оценки коэффициентов регрессии методом инструментальных переменных.

№4. Рассмотрим модель ADL с автокоррелированными ошибками

$$y_t = \beta_0 + \beta_1 y_{t-1} + \gamma_1 x_t + u_t \quad u_t = \rho u_{t-1} + v_t \quad (|\rho| < 1, \rho \neq 0)$$

где v_t удовлетворяет условиям Гаусса–Маркова. Как автокоррелированность ошибок влияет на свойства оценок наименьших квадратов коэффициентов регрессии?

Глава 8

Временные ряды

8.1 Модели стационарных временных рядов

№1. Постройте автоковариационную и автокорреляционную функцию для модели $MA(1)$, $MA(2)$, $MA(3)$.

№2. Для модели $ARMA(1, 1)$

$$y_t = \beta_0 + \beta_1 y_{t-1} + u_t + \theta_1 u_{t-1} \quad u_t \sim WN$$

1. Вычислите Ey_t и $Var(y_t)$.
2. Найдите представление временного ряда в виде $MA(\infty)$.
3. Вычислите автоковариационную и автокорреляционную функцию.

№3. Для модели $ARMA(1, 2)$

$$y_t = \beta_0 + \beta_1 y_{t-1} + u_t + \theta_1 u_{t-1} + \theta_2 u_{t-2} \quad u_t \sim WN$$

1. Вычислите Ey_t и $Var(y_t)$.
2. Найдите представление временного ряда в виде $MA(\infty)$.
3. Вычислите автоковариационную и автокорреляционную функцию.

№4. Задаёт ли модель

$$x_t = 10 + 1.2x_{t-1} - 0.2x_{t-2} + u_t \quad u_t \sim WN$$

стационарный ряд? Ответ поясните.

№5. По временному ряду длины $n = 110$ были оценены следующие модели ARMA:

1. $y_t = 16 + 0.67y_{t-1} + u_t - 0.3u_{t-1}$, $s^2 = 1.63$;
2. $y_t = 15.1 + 0.6y_{t-1} - 0.1y_{t-2} + u_t$, $s^2 = 1.8$;
3. $y_t = 17.1 + 0.71y_{t-1} + u_t - 0.34u_{t-1} + 0.12u_{t-2}$, $s^2 = 1.86$.

Какую модель вы выберете?

№6. По 100 наблюдениям была оценена модель

$$x_t = 2 + 0.5x_{t-1} - 0.06x_{t-2} + u_t.$$

и проверена её адекватность. Известно, что $x_{100} = 3.1$, $x_{99} = 2.5$, $x_{98} = 2.8$, $x_{97} = 3.6$. Найдите \hat{x}_{102} , \hat{x}_{103} и \hat{x}_{104} .

№7. Временной ряд x_t – первая разность логарифма реального ВВП США с 1984:2 по 2013:2. Для этого ряда были оценены следующие модели (в таблице приведены значения информационных критериев и Q_{LB} -статистики с лагом 4 для остатков):

модель	MA(1)	AR(1)	ARMA(1,1)	AR(2)	MA(2)	AR(3)	ARMA(2,1)
AIC	-870.85	-880.17	-883.55	-885.54	-882.83	-884.10	-884.54
BIC	-862.57	-871.88	-872.50	-874.49	-871.78	-870.29	-870.73
Q_{LB}	21.5127	11.0688	3.68299	1.34523	4.21655	0.963561	0.494841

Какая модель предпочтительней? Ответ обоснуйте.

№8. Временной ряд x_t – первая разность логарифма реального ВВП США с 1984:2 по 2013:2 (квартальные данные). Была оценена модель ARMA(2,1)

	const	x_{t-1}	x_{t-2}	u_{t-1}
коэфф.	0.007	-0.026	0.397	0.378
P-значение	$1.02 \cdot 10^{-10}$	0.9172	0.0006	0.1534

Известны значения временного ряда и остатки модели

t	2012:2	2012:3	2012:4	2013:1	2013:2
x_t	0.003	0.007	0.0004	0.003	0.006
\hat{u}_t	-0.007	0.001	-0.006	-0.002	0.003

Постройте прогноз на четвёртый квартал 2013 года (2013:4).

8.2 Модели стационарных временных рядов. Практика

Для практических задач данные нужно взять с сайта Федерального Банка г. С.-Луиса www.stlouisfed.org

№9. Рассмотрим месячные данные по денежной массе M2 для US с 1990 г. по н.в. Постройте ряд x_t – первая разность логарифма M2.

1. Постройте график x_t . Какие можно сделать выводы?
2. Постройте ACF до лага 20. Какие можно сделать выводы? Подтверждается ли гипотеза что x_t – белый шум?
3. Какую модель ARMA выбрать для ряда x_t ? Проверьте “адекватность” выбранной модели.
4. Сделайте прогноз на 7 шагов по выбранной и оценённой модели.

№10. Рассмотрим квартальные данные по ВВП US с 1985 г. по н.в. Постройте ряд x_t – первая разность логарифма ВВП.

1. Постройте график x_t . Какие можно сделать выводы?
2. Постройте ACF до лага 20. Какие можно сделать выводы? Подтверждается ли гипотеза что x_t – белый шум?
3. Какую модель ARMA выбрать для ряда x_t ? Проверьте “адекватность” выбранной модели.
4. Сделайте прогноз на 4 шага по выбранной и оценённой модели.

№11. Рассмотрим месячные данные по краткосрочной ставке 3-х месячных Treasury Bills с 1990 г. по н.в. Постройте ряд x_t – первая разность ставки.

1. Постройте график x_t . Какие можно сделать выводы?
2. Постройте ACF до лага 20. Какие можно сделать выводы? Подтверждается ли гипотеза что x_t – белый шум?
3. Какую модель ARMA выбрать для ряда x_t ? Проверьте “адекватность” выбранной модели.

4. Сделайте прогноз на 7 шагов по выбранной и оценённой модели.

№12. Рассмотрим месячные данные по долгосрочной ставке 10-ти летних бондов с 1990 г. пл н.в. Постройте ряд x_t – первая разность ставки.

1. Постройте график x_t . Какие можно сделать выводы?
2. Постройте АСФ до лага 20. Какие можно сделать выводы? Подтверждается ли гипотеза что x_t – белый шум?
3. Какую модель ARMA выбрать для ряда x_t ? Проверьте “адекватность” выбранной модели.
4. Сделайте прогноз на 7 шагов по выбранной и оценённой модели.

8.3 Модели распределенных лагов

№13. Рассмотрим модель FDL(2)

$$y_t = \mu + \beta_0 x_t + \beta_1 x_{t-1} + \beta_2 x_{t-2} + \text{error}$$

1. Напишите функцию импульсного отклика (impulse response function) для краткосрочной зависимости.
2. Напишите функцию импульсного отклика (impulse response function) для долгосрочной зависимости.
3. Напишите уравнение долгосрочной зависимости и дайте его интерпретацию.

№14. Рассмотрим модель FDL(4)

$$y_t = \mu + \beta_0 x_t + \beta_1 x_{t-1} + \beta_2 x_{t-2} + \beta_3 x_{t-3} + \gamma_0 z_t + \gamma_1 z_{t-1} + \gamma_2 z_{t-2} + u_t.$$

Сформулируйте условия Гаусса–Маркова для ошибок этой модели регрессии. Найдите долгосрочные мультипликаторы y по x и по z . Напишите уравнение «долгосрочной зависимости» y от x и z .

№15. Рассмотрим модель ADL

$$y_t = \mu + \phi y_{t-1} + \beta x_t + \text{error}$$

1. Когда для этой модели выполнено условие стационарности?
2. Напишите функцию импульсного отклика для краткосрочной зависимости.
3. Напишите функцию импульсного отклика для долгосрочной зависимости.
4. Напишите уравнение долгосрочной зависимости и дайте его интерпретацию.

№16. Рассмотрим модель ADL

$$y_t = \mu + \phi y_{t-1} + \beta_0 x_t + \beta_1 x_{t-1} + \text{error}$$

1. Когда для этой модели выполнено условие стационарности?
2. Напишите функцию импульсного отклика для краткосрочной зависимости.
3. Напишите функцию импульсного отклика для долгосрочной зависимости.
4. Напишите уравнение долгосрочной зависимости и дайте его интерпретацию.

№17. Рассмотрим модель ADL

$$y_t = \mu + \phi_1 y_{t-1} + \phi_2 y_{t-2} + \beta x_t + \text{error}$$

1. Когда для этой модели выполнено условие стационарности?
2. Напишите функцию импульсного отклика для краткосрочной зависимости.
3. Напишите функцию импульсного отклика для долгосрочной зависимости.
4. Напишите уравнение долгосрочной зависимости и дайте его интерпретацию.

№18. Рассмотрим модель ADL(2,2)

$$y_t = 5 + \frac{5}{6}y_{t-1} - \frac{1}{6}y_{t-2} + 2.1x_t + 1.2x_{t-1} + 0.2x_{t-2} + 4.3z_t + 2.3z_{t-1} + u_t$$

для ошибок выполнены условия Гаусса–Маркова.

1. Выполнено ли для этой модели условие стационарности?
2. Если да, то найдите долгосрочные мультипликаторы y по x и z и напишите уравнение «долгосрочной зависимости» y от x и z .

№19. Рассмотрим модель ADL(2,1)

$$y_t = \mu + \phi_1 y_{t-1} + \phi_2 y_{t-2} + \beta_0 x_t + \beta_1 x_{t-1} + u_t,$$

где x_t – стационарный временной ряд и $u_t \sim \text{WN}$.

1. Когда для этой модели выполнено условие стационарности?
2. Напишите уравнение долгосрочной зависимости и дайте его интерпретацию.

8.4 Модели распределенных лагов. Практика

Для практических задач данные нужно взять с сайта Федерального Банка г. С.-Луиса www.stlouisfed.org

№20. Рассмотрим недельные исторические данные для US по M2, краткосрочной 3-х месячной ставке и долгосрочной 10-ти летней ставке с 1990 года по н.в. Пусть $d\ln M2$ – первая разность логарифма M2, $dTB3$ – первая разность краткосрочной ставки, $dTSM10$ – первая разность долгосрочной ставки.

1. Постройте модель FDL зависимости $d\ln M2$ от $dTB3$ и $dTSM10$. Какой порядок шагов выбрать?
2. Дайте интерпретацию коэффициентов оценённой модели.
3. Напишите уравнение долгосрочной зависимости и дайте его интерпретацию.

№21. Рассмотрим квартальные исторические данные для US по ВВП, M2, краткосрочной 3-х месячной ставке и долгосрочной 10-ти летней ставке с 1990 года по н.в. Пусть IGDP – первая разность логарифма ВВП, dLM2 – первая разность логарифма M2, dTB3 – первая разность краткосрочной ставки, dTSM10 – первая разность долгосрочной ставки.

1. Постройте модель FDL зависимости IGDP от dLM2 , dTB3 и dTSM10 . Какой порядок шагов выбрать?
2. Дайте интерпретацию коэффициентов оценённой модели.
3. Напишите уравнение долгосрочной зависимости и дайте его интерпретацию.

8.4.1 TS – ряды

№1. Рассмотрим модель временного ряда

$$x_t = \gamma_0 + \gamma_1 t + \phi x_{t-1} + u_t \quad u_t \sim \text{WN}$$

1. Сформулируйте условия стационарности.
2. Напишите уравнение тренда для временного ряда.

№2. Рассмотрим модель временного ряда

$$x_t = \gamma_0 + \gamma_1 t + \gamma_2 t^2 + \phi x_{t-1} + u_t \quad u_t \sim \text{WN}$$

1. Сформулируйте условия стационарности.
2. Напишите уравнение тренда для временного ряда.

№3. Рассмотрим модель временного ряда

$$x_t = \gamma_0 + \gamma_1 t + \phi_1 x_{t-1} + \phi_2 x_{t-2} + u_t \quad u_t \sim \text{WN}$$

1. Сформулируйте условия стационарности.
2. Напишите уравнение тренда для временного ряда.

№4. Для TS-рядов (v_t – стационарный ряд)

1. $x_t = \beta_0 + \beta_1 t + v_t$;

$$2. x_t = \beta_0 + \beta_1 t + \beta_2 t^2 + v_t$$

Найдите Δx_t и $\Delta^2 x_t$.

№5. Для TS-рядов ($u_t \sim \text{WN}$)

$$1. x_t = \gamma_0 + \gamma_1 t + \phi x_{t-1} + u_t;$$

$$2. x_t = \gamma_0 + \gamma_1 t + \gamma_2 t^2 + \phi x_{t-1} + u_t$$

$$3. x_t = \gamma_0 + \gamma_1 t + \phi_1 x_{t-1} + \phi_2 x_{t-2} + u_t$$

найдите уравнение для Δx_t и $\Delta^2 x_t$.

№6. Рассмотрим модель квадратичного тренда с автокорреляцией первого порядка

$$\begin{aligned} x_t &= \beta_0 + \beta_1 t + \beta_2 t^2 + v_t \\ v_t &= \rho v_{t-1} + u_t \quad u_t \sim \text{WN} \quad |\rho| < 1 \end{aligned}$$

1. Проведите авторегрессионное преобразование для получения эффективной оценки тренда.
2. Как связаны коэффициенты преобразованного уравнения с параметрами тренда?

№7. Для TS-ряда

$$x_t = \gamma_0 + \gamma_1 t + \gamma_2 t^2 + \phi x_{t-1} + u_t + \theta u_{t-1} \quad u_t \sim \text{WN} \quad |\phi| < 1$$

1. Найдите уравнение для первой разности x_t .
2. Найдите уравнение для второй разности x_t .

№8. Рассмотрим модель временного ряда для ВВП

$$\ln \text{GDP}_t = \gamma_0 + \gamma_1 t + \phi_1 \ln \text{GDP}_{t-1} + \phi_2 \ln \text{GDP}_{t-2} + u_t, \quad u_t \sim \text{WN}$$

1. При каких условиях эта модель временного ряда задаёт TS-ряд?
2. Найдите тренд ВВП (для случая TS).

№9. Рассмотрим модель тренда для ВВП (в млн. \$)

$$\widehat{GDP}_t = \underset{(23.4)}{230.4} + \underset{(0.038)}{0.32} t \quad (t = 0, \dots, 20)$$

Дайте интерпретацию коэффициентов модели.

Проверьте значимость тренда при альтернативе о тенденции к росту ВВП. Уровень значимости 5%

№10. Рассмотрим модель тренда для численности населения (в млн. чел.)

$$\widehat{\ln POP}_t = \underset{(0.4)}{2.4} + \underset{(0.01)}{0.023} t \quad (t = 0, \dots, 23)$$

Дайте интерпретацию коэффициентов модели.

Проверьте значимость тренда при альтернативе о тенденции к росту населения. Уровень значимости 1%.

№11. Исследует регрессионная модель влияния численности населения (POP) на ВВП (GDP) на основе временных рядов

$$\ln GDP_t = \beta_0 + \beta_1 \ln POP_t + \text{error}.$$

Однако временные ряды для ВВП и численности населения как правило имеют тенденцию к росту. Как это повлияет на статистические выводы для регрессионной модели?

8.4.2 DS – ряды

№1. Для временного ряда

$$x_t = x_{t-1} + u_t \quad u_t \sim \text{WN}(0, \sigma^2) \quad x_0 = 0$$

найдите $E x_t$ и $\text{Var } x_t$.

№2. Для временного ряда ($\mu \neq 0$)

$$x_t = \mu + x_{t-1} + u_t \quad u_t \sim \text{WN}(0, \sigma^2) \quad x_0 = 0$$

найдите $E x_t$ и $\text{Var } x_t$.

№3. Для временного ряда ($\mu \neq 0$)

$$x_t = \mu + \beta t + x_{t-1} + u_t \quad u_t \sim \text{WN}(0, \sigma^2) \quad x_0 = 0$$

найдите $E x_t$ и $\text{Var } x_t$.

№4. Пусть $x_t \sim \text{ARIMA}(1, 1, 0)$. Запишите представление ряда в виде (нестационарной) модели $\text{ARMA}(2, 0)$. Покажите, что авторегрессионный многочлен модели ARMA имеет единичный корень кратности 1.

№5. Пусть $x_t \sim \text{ARIMA}(1, 2, 0)$. Запишите представление ряда в виде (нестационарной) модели $\text{ARMA}(3, 0)$. Покажите, что авторегрессионный многочлен модели ARMA имеет единичный корень кратности 2.

№6. Пусть $x_t \sim \text{ARIMA}(2, 2, 3)$. Запишите представление ряда в виде (нестационарной) модели ARMA . Покажите, что авторегрессионный многочлен модели ARMA имеет единичный корень и найдите кратность этого корня.

№7. Пусть $x_t \sim \text{ARIMA}(3, 1, 1)$. Запишите представление ряда в виде (нестационарной) модели ARMA . Покажите, что авторегрессионный многочлен модели ARMA имеет единичный корень и найдите кратность этого корня.

№8. Пусть $x_t \sim \text{ARIMA}(1, 2, 2)$.

1. Что означает $x_t \sim \text{ARIMA}(1, 2, 2)$?
2. Запишите представление ряда в виде (нестационарной) модели ARMA .

№9. Рассмотрим модель временно ряда

$$x_t = 2x_{t-1} - x_{t-2} + u_t + 0.5u_{t-1} - 0.1u_{t-2} \quad u_t \sim \text{WN}$$

1. Покажите, что $x_t \sim I(2)$.
2. Найдите порядок модели $\text{ARIMA}(p, k, q)$, соответствующей ряду x_t .

8.4.3 DS – ряды. Практика

Для практических задач данные нужно взять с сайта Федерального Банка г. С.-Луиса www.stlouisfed.org

№10. Рассмотрим месячные данные по денежной массе M2 для US с 1990 г. по н.в. Ряд x_t – логарифм M2.

1. Постройте график x_t . Какие можно сделать выводы?
2. Тестируйте ряда x_t на единичный корень.

3. Какой порядок интегрирования у ряда x_t ?
4. Какую модель ARIMA выбрать для ряда x_t ?

№11. Рассмотрим квартальные данные по ВВП US с 1985 г. по н.в. Постройте ряд x_t – логарифм ВВП.

1. Постройте график x_t . Какие можно сделать выводы?
2. Тестируйте ряда x_t на единичный корень.
3. Какой порядок интегрирования у ряда x_t ?
4. Какую модель ARIMA выбрать для ряда x_t ?

№12. Рассмотрим месячные данные по краткосрочной ставке 3-х месячных Treasury Bills с 1990 г. пл н.в. (ряд x_t)

1. Постройте график x_t . Какие можно сделать выводы?
2. Тестируйте ряда x_t на единичный корень.
3. Какой порядок интегрирования у ряда x_t ?
4. Какую модель ARIMA выбрать для ряда x_t ?

№13. Рассмотрим месячные данные по долгосрочной ставке 10-ти летних бондов с 1990 г. по н.в. (ряд x_t)

1. Постройте график x_t . Какие можно сделать выводы?
2. Тестируйте ряда x_t на единичный корень.
3. Какой порядок интегрирования у ряда x_t ?
4. Какую модель ARIMA выбрать для ряда x_t ?

8.4.4 Многомерные модели временных рядов

VAR и коинтеграция

№1. Рассмотрим модель VAR(1)

$$\mathbf{x}_t = \mathbf{A}\mathbf{x}_{t-1} + \mathbf{u}_t \quad \mathbf{x}_t = \begin{pmatrix} x_t \\ y_t \end{pmatrix} \quad \mathbf{u}_t = \begin{pmatrix} u_t \\ v_t \end{pmatrix},$$

где $u_t \sim \text{WN}(0, \sigma_u^2)$, $v_t \sim \text{WN}(0, \sigma_v^2)$, $\text{cov}(u_t, v_t) = \sigma_{uv}$. Проверьте условие стационарности для матриц

$$\mathbf{A} = \begin{pmatrix} 0.5 & 1 \\ 0 & 0.3 \end{pmatrix} \quad \begin{pmatrix} 0 & 0.5 \\ -0.5 & 0 \end{pmatrix} \quad \begin{pmatrix} 1 & 3 \\ 0 & 0.2 \end{pmatrix} \quad \begin{pmatrix} 0 & 1 \\ 0 & 1 \end{pmatrix} \quad \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix} \quad \begin{pmatrix} 0 & 1 \\ 1 & 1 \end{pmatrix}$$

№2. Рассмотрим модели временных рядов

$$\begin{cases} x_t = x_{t-1} + u_t \\ y_t = y_{t-1} + v_t \end{cases} \quad \begin{cases} x_t = x_{t-1} + u_t \\ y_t = cx_t + v_t \end{cases}$$

$$\begin{cases} x_t = x_{t-1} + u_t \\ y_t = x_t + x_{t-1} + v_t \end{cases} \quad \begin{cases} x_t = 3x_{t-1} - 7y_{t-1} + u_t \\ y_t = x_{t-1} - 2.5y_{t-1} + v_t \end{cases}$$

1. Запишите эти модели в виде моделей VAR.
2. Проверьте условие стационарности.
3. Какие временные ряды коинтегрированы?
 - Если ряды коинтегрированы, то найдите коинтеграционное соотношение, корректирующую ошибку и запишите модель VECM.
 - Если ряды неинтегрированы, то постройте VAR-модель для первых разностей.

№3. Рассмотрим модель VAR(1)

$$\mathbf{x}_t = \mathbf{A}\mathbf{x}_{t-1} + \mathbf{u}_t \quad \mathbf{x}_t = \begin{pmatrix} x_t \\ y_t \\ z_t \end{pmatrix} \quad \mathbf{u}_t = \begin{pmatrix} u_t \\ v_t \\ w_t \end{pmatrix} \sim \text{WN}$$

Проверьте условие стационарности для матриц

$$\mathbf{A} = \begin{pmatrix} 0 & 1 & 3 \\ -1 & 0 & -2 \\ 0 & 0 & 0.5 \end{pmatrix} \quad \begin{pmatrix} 0 & 0 & 0.5 \\ 0.5 & 0 & 0 \\ 0 & 0.5 & 0 \end{pmatrix}$$

№4. Рассмотрим модель VAR(2)

$$\mathbf{x}_t = \mathbf{A}_1 \mathbf{x}_{t-1} + \mathbf{A}_2 \mathbf{x}_{t-2} + \mathbf{u}_t \quad \mathbf{x}_t = \begin{pmatrix} x_t \\ y_t \end{pmatrix} \quad \mathbf{u}_t = \begin{pmatrix} u_t \\ v_t \end{pmatrix} \sim \text{WN}$$

Проверьте условие стационарности для матриц

$$\begin{aligned} 1) \mathbf{A}_1 &= \begin{pmatrix} 2 & 3 \\ 0 & 1 \end{pmatrix} & \mathbf{A}_2 &= \begin{pmatrix} -1 & 3 \\ 0 & -0.25 \end{pmatrix} \\ 2) \mathbf{A}_1 &= \begin{pmatrix} 0 & 0.5 \\ 0.5 & 0 \end{pmatrix} & \mathbf{A}_2 &= \begin{pmatrix} 0 & -0.25 \\ -0.25 & 0 \end{pmatrix} \end{aligned}$$

№5. Рассмотрим модели временных рядов

$$\begin{cases} x_t = 2x_{t-1} - x_{t-2} + u_t \\ y_t = 1.5y_{t-1} - 0.5y_{t-2} + v_t \end{cases} \quad \begin{cases} x_t = 1.5x_{t-1} + y_{t-1} - 0.5x_{t-2} - y_{t-2} + u_t \\ y_t = -x_{t-1} - 0.5y_{t-1} + x_{t-2} + 1.5y_{t-2} + v_t \end{cases}$$

$$\begin{cases} x_t = x_{t-1} + u_t \\ y_t = x_t + x_{t-1} + v_t \\ z_t = x_t + y_{t-1} + w_t \end{cases} \quad \begin{cases} x_t = y_{t-1} + u_t \\ y_t = z_{t-1} + v_t \\ z_t = x_{t-1} + w_t \end{cases}$$

1. Запишите эти модели в виде моделей VAR.
2. Проверьте условие стационарности.
3. Какие временные ряды коинтегрированы?
 - Если ряды коинтегрированы, то найдите коинтеграционное соотношение, скорректируйте ошибку и запишите модель VECM.
 - Если ряды не коинтегрированы, то постройте VAR-модель для первых разностей.

8.4.5 Из пособия

№1. Является ли временной ряд, заданный авторегрессионным разностным уравнением, стационарным?

1. $y_t = 7 + 0.5y_{t-1} + \varepsilon_t$.
2. $y_t = 10 + 0.25y_{t-2} + \varepsilon_t$.

3. $y_t = 10 + y_{t-1} - 0.25y_{t-2} + \varepsilon_t$.
4. $y_t = \frac{3}{2}y_{t-1} - \frac{3}{4}y_{t-2} + \frac{1}{8}y_{t-3} + \varepsilon_t$.
5. $y_t = 3 + 0.4y_{t-1} - 0.04y_{t-2} + \varepsilon_t$.
6. $y_t = 5 - 3y_{t-1} - 3y_{t-2} - y_{t-3} + \varepsilon_t$.
7. $y_t = -2y_{t-1} + 1.25y_{t-2} - 0.25y_{t-3} + \varepsilon_t$.

№2. Написать формулы для прогноза

1. на $l = 3$ шага для процесса $y_t = 5 + 0.5y_{t-1} + \varepsilon_t$.
2. на $l = 4$ шага для процесса $y_t = 2 + 0.25y_{t-2} + \varepsilon_t$.
3. на $l = 3$ шага для процесса $y_t = \frac{1}{27}y_{t-3} + \varepsilon_t$.
4. на $l = 5$ шагов для процесса $y_t = 0.5y_{t-2} - 0.05y_{t-3} + 0.001y_{t-4} + \varepsilon_t$.
5. на $l = 6$ шагов для процесса $y_t = 0.5y_{t-3} + 0.001y_{t-4} + \varepsilon_t$.

№3. Написать в общем виде уравнения Юла – Уолкера для моделей AR(2), AR(3) и AR(4).

№4. По временному ряду длины $n = 60$ были оценены следующие авторегрессионные модели:

1. $\hat{y}_t = 2 + 0.7y_{t-1}, s^2 = 2.1$;
2. $\hat{y}_t = 2.3 + 0.6y_{t-1} - 0.3y_{t-2}, s^2 = 1.9$;
3. $\hat{y}_t = 1.8 + 0.55y_{t-1} - 0.25y_{t-2} + 0.01y_{t-3}, s^2 = 1, 85$.

Какую модель вы выберете?

№5. Для временного ряда были вычислены коэффициенты автокорреляции

$$\rho(1) = 0.7; \quad \rho(2) = 0.4; \quad \rho(3) = -0.2$$

и выборочное среднее значение $\bar{y} = 1.7$. Найти оценки коэффициентов в модели

1. AR(1);

2. AR(2);

3. AR(3).

№6. Для модели временного ряда длины $n = 50$ были оценены несколько моделей и в каждой из моделей вычислены коэффициенты автокорреляции остатков. Исследовать адекватность этих моделей

1. $r_e(1) = 0.001$; $r_e(2) = 0.0006$; $r_e(3) = 0.0002$; $r_e(4) = 0.001$ в модели AR(2).

2. $r_e(1) = 0.04$; $r_e(2) = 0.02$; $r_e(3) = 0.006$ в модели AR(1).

3. $r_e(1) = 0.02$; $r_e(2) = 0.0008$; $r_e(3) = 0.003$; $r_e(4) = 0.001$ в модели AR(3).

№7. Для временного ряда оценить модели AR(1), AR(2) и AR(3). Используя информационные критерии выбрать из этих моделей и проверить ее адекватность.

t	Ряд 1	Ряд 2	Ряд 3	Ряд 4	Ряд 5	Ряд 6
1	-1.98	0.67	4.22	5.84	1.88	19.05
2	0.75	2.57	4.75	5.92	0.78	20.11
3	1.41	-0.70	8.02	4.64	2.87	22.42
4	1.25	4.71	5.27	4.77	-0.26	18.80
5	1.27	-1.79	3.61	4.64	2.50	20.64
6	1.70	4.34	2.65	6.28	0.25	18.36
7	0.48	-0.57	4.99	6.71	0.02	20.72
8	0.20	2.71	7.67	7.47	3.59	21.12
9	-0.28	1.70	6.08	7.58	-0.52	19.90
10	0.25	2.84	3.81	6.47	2.47	19.84
11	2.51	0.47	3.78	5.42	1.40	19.96
12	2.37	1.78	4.63	4.70	-0.81	19.78
13	3.65	2.02	6.45	5.08	5.07	21.11
14	2.81	1.69	5.97	4.23	-4.03	20.06
15	0.51	1.60	6.09	4.03	4.60	21.64
16	1.44	1.18	6.07	2.17	-2.30	21.27
17	0.70	0.17	5.96	3.61	1.54	21.84
18	-0.73	2.39	5.17	5.02	3.52	21.19
19	-1.14	2.34	4.41	5.79	-1.10	20.97
20	0.38	1.45	4.15	5.65	2.60	20.54
21	1.50	3.48	3.48	5.91	2.57	19.47
22	3.33	2.05	4.19	6.56	-1.24	20.02
23	2.15	-0.16	5.51	6.31	4.39	20.19
24	3.95	2.99	4.30	5.99	-1.08	18.67
25	3.10	0.48	6.03	3.93	4.01	21.48
26	0.68	2.34	4.19	2.96	1.11	17.94
27	-0.32	-0.81	3.66	2.98	0.83	19.75
28	1.41	4.44	3.34	1.70	1.85	17.89
29	-0.58	-2.16	4.65	3.26	0.18	19.36
30	1.00	4.77	6.21	6.56	1.38	19.55
31	0.50	0.04	6.37	8.26	0.34	19.88
32	-1.23	4.03	4.84	9.21	-0.21	19.19
33	-3.08	0.49	3.60	7.08	3.05	18.98
34	-2.61	3.07	4.11	4.94	-0.80	19.26
35	-2.89	0.41	7.36	2.90	1.71	21.60

Глава 9

Модели регрессии с бинарной зависимой переменной

9.1 LPM-модель

№1. Рассмотрим LPM модель

$$\widehat{Male} = 0.37 + \underset{[0.011]}{0.02} Educ - \underset{[0.046]}{0.1} City \quad n = 700$$

зависимости пола от уровня образования $Educ$ и места жительства $City$ (в квадратных скобках указаны стандартные ошибки в форме Уайта). Почему при оценке модели были использованы стандартные ошибки в форме Уайта (робастные НС-стандартные ошибки)? Кратко поясните. Проверьте значимость влияния каждого из факторов (уровень значимости 1%). Вычислите предсказанное значение для человека с образованием 15 лет и живущего в городе. Как следует интерпретировать предсказанное значение?

№2. Рассмотрим LPM-модель (male – пол, age – возраст, educ – уровень образования в годах, city – бинарная переменная, отвечающая за место жительства)

Зависимая переменная: male; объем выборки 1320

	const	educ	age	city
coefficient	0.47	0.002	-0.001	-0.81
White's Std. Error	0.02	0.0008	0.0004	0.13

58 ГЛАВА 9. МОДЕЛИ РЕГРЕССИИ С БИНАРНОЙ ЗАВИСИМОЙ ПЕРЕМЕННОЙ

- a) Дайте интерпретацию полученных оценок коэффициентов
- b) Почему почему при оценке модели были использованы стандартные ошибки в форме Уайта? Ответ поясните.
- c) Вычислите предсказанное значение для 35-ти летнего городского жителя с уровнем образования 12 лет и дайте его интерпретацию.
- d) Значимо ли уровень образования влияет на зависимую переменную? Сформулируйте и тестируйте соответствующую статистическую гипотезу

№3. Пусть *approve* – бинарная переменная, равная 1 если кредит заёмщику одобрен. Рассмотрим LPM-модель (*appinc* – доход заёмщика в \$1000, *mortno* – бинарный фактор, равен 1 если нет ипотечной кредитной истории, *unem* – уровень безработицы в отрасли в %, *dep* – количество иждивенцев у заёмщика, *male* – гендерный фактор, *married* – семейный статус, *yjob* – стаж на текущей работе)¹

Зависимая переменная: *approve*; объем выборки: 1971

	const	appinc	mortno	unem	dep	male	married	yjob
коэфф.	0.875	-0.0001	0.074	-0.008	-0.018	0.002	0.047	-0.0003
OLS-ст.ошибка	0.023	0.00008	0.016	0.003	0.007	0.02	0.017	0.007
НС-ст.ошибка	0.024	0.00011	0.015	0.0039	0.0075	0.021	0.0187	0.006

$$R^2 = 0.020483 \quad R^2_{adj} = 0,016990 \quad F - \text{стат} = 5,711436$$

- a) Проверьте значимость влияния дохода заёмщика.
- b) Уровень безработицы в отрасли равен 3.7%. Для неженатого мужчины с ипотечной кредитной истории, с уровнем дохода \$107000, работающего на текущем месте 2 года и без иждивенцев вычислите предсказанное значение и дайте его интерпретацию.
- c) Дайте интерпретацию коэффициента при *married*, *dep* & *unem*.

¹НС – робастные стандартные ошибки, устойчивые к гетероскедастичности

9.2 probit/logit-модель

№4. Рассмотрим probit-модель

$$P(y = 1) = \Phi\left(0.45 + \underset{(0.2)}{0.3x_1} + \underset{(1.2)}{2.1x_2} - \underset{(0.2)}{0.3x_3} + \underset{(0.9)}{1.2x_4}\right)$$

$$\ln \hat{\mathcal{L}} = -1202.3 \quad \ln \hat{\mathcal{L}}_0 = -1382.2 \quad n = 1050$$

Проверьте значимость коэффициентов регрессии. Проверьте значимость регрессии «в целом».

№5. Рассмотрим probit-модель (объем выборки 920)

$$P(\text{Male} = 1) = \Phi\left(0.02 + \underset{(0.009)}{0.013educ} - \underset{(0.046)}{0.1city} - \underset{(0.011)}{0.029age} - \underset{(0.07)}{0.22unem}\right)$$

зависимости пола от уровня образования *educ*, места жительства *city*, возраста *age* и статуса работающий/безработный *unem* (в скобках указаны стандартные ошибки).

- Напишите формулу для функции $\Phi(x)$.
- Проверьте значимость влияния каждого из факторов (уровень значимости 10%).
- Дайте интерпретацию коэффициентов модели.
- Как интерпретируются предсказанные значения в этой модели регрессии?
- Вычислите предельное значение $P(\text{Male} = 1)$ по *educ*.

№6. Рассмотрим *Logit*-модель, оцененную в EViews

Dependent variable: crime86
Sample 1 2725

Variable	Coefficient	Std. Error	z-Statistic	Prob.
C	-0.655477	0.119205	-5.498726	0.0000
pcnv	-0.924175	0.124452	-7.425948	0.0000
avgsen	0.018503	0.035298	0.524185	0.6001
totime	-0.010810	0.028172	-0.383710	0.7012
ptime86	-0.123048	0.031627	-3.890563	0.0001
qemp86	0.066483	0.045187	1.471275	0.1412
inc86	-0.009057	0.001290	-7.021432	0.0000
durat	0.019617	0.010214	1.920591	0.0548
black	0.776393	0.118454	6.554379	0.0000
hispan	0.498160	0.109949	4.530833	0.0000
born60	0.018356	0.094461	0.194319	0.8459

$$\ln \hat{\mathcal{L}} = -1481.458 \quad \ln \hat{\mathcal{L}}_0 = -1608.184$$

- Какие из факторов значимо, а какие незначимо влияют на зависимую переменную? Уровень значимости 10%.
- Вычислите $R^2_{McFadden}$
- Тестируйте значимость регрессии «в целом» при уровне значимости 1%.
- Была оценена *Logit*-модель без факторов black & hispan и получено значение логарифма функции правдоподобия $\ln \hat{\mathcal{L}}_r = -1506.727$. Тестируйте совместную значимость факторов black & hispan при уровне значимости 5%.

№7. Пусть *City* – бинарная переменная, равная 1 если человек живет в городе. Рассмотрим *logit*-модель (*educ* – уровень образования, *age* – возраст, *female* – пол человека)

Зависимая переменная: *City*; объем выборки 1023

	const	educ	age	female
coefficient	0.02	0.004	0.0018	0.04
Std. Error	0.13	0.0013	0.0002	0.023

$$\ln \hat{\mathcal{L}} = -1281 \quad \ln \hat{\mathcal{L}}_0 = -1508$$

Уровень значимости 5%

- a) Запишите уравнение регрессии и формулу для logit-функции.
- b) Дайте интерпретацию коэффициентов регрессии.
- c) Значимо ли влияние уровня образования?
- d) Вычислите $R_{McFadden}^2$
- e) Тестируйте значимость регрессии «в целом» и сформулируйте проверяемую статическую гипотезу.
- f) Как интерпретируются предсказанные значения в этой модели регрессии?
- g) Для мужчины возраста 30 лет и с образованием 15 лет вычислите предсказанное значение и дайте его интерпретацию.
- h) Укажите формулу для предельных значений по educ и по age.
- i) Для мужчин возраста 25 лет и 35 лет и с образованием 15 лет (каждый) вычислите предельные значения по age и дайте их интерпретацию.
- j) Пусть LR-статистика для проверки совместной значимости факторов age и female равна 13.5. Проверьте совместную значимость этих факторов и сформулируйте проверяемую статистическую гипотезу.

62 ГЛАВА 9. МОДЕЛИ РЕГРЕССИИ С БИНАРНОЙ ЗАВИСИМОЙ ПЕРЕМЕННОЙ

№8. Пусть *approve* – бинарная переменная, равная 1 если кредит заёмщику одобрен. Рассмотрим несколько альтернативных logit-моделей (*appinc* – доход заёмщика в \$1000, *mortno* – бинарный фактор, равен 1 если нет ипотечной кредитной истории, *unem* – уровень безработицы в отрасли в %, *dep* – количество иждивенцев у заёмщика, *male* – гендерный фактор, *married* – семейный статус, *yjob* – стаж на текущей работе в годах)

Зависимая переменная: *approve*; объем выборки: 1971

Модель	№1	№2	№3	№4	№5	№6	№7
const	1.766 (0.158)	1.942 (0.199)	1.968 (0.169)	2.299 (0.153)	1.882 (0.160)	2.113 (0.156)	1.899 (0.165)
appinc		-0.001 (0.0007)	-0.001 (0.0007)	-0.0008 (0.0006)		-0.001 (0.0007)	-0.0008 (0.0006)
mortno		0.773 (0.172)	0.770 (0.171)		0.748 (0.171)	0.786 (0.166)	
unem		-0.062 (0.028)	-0.064 (0.029)	-0.066 (0.028)	-0.062 (0.029)	-0.065 (0.028)	
dep	-0.146 (0.064)	-0.162 (0.064)	-0.161 (0.064)		-0.170 (0.064)		
male	-0.018 (0.183)	0.022 (0.186)					0.147 (0.171)
married	0.516 (0.161)	0.427 (0.162)	0.432 (0.154)		0.421 (0.154)		
yjob		-0.003 (0.066)		0.010 (0.062)		0.003 (0.065)	0.017 (0.063)
$\ln \hat{\mathcal{L}}$	-731.83	-717.42	-719.30	-737.07	-720.60	-724.89	-737.48
pseudo- R^2	0.0083	0.0279	0.0277	0.0042	0.0260	0.0207	0.0012
AIC	1471.656	1450.834	1450.593	1482.138	1451.202	1459.778	1482.967
BIC	1494.001	1495.525	1484.153	1504.517	1479.169	1487.752	1505.318

В скобках указаны стандартные ошибки коэффициентов, $\ln \hat{\mathcal{L}}_0 = -740.2157$

- Дайте интерпретацию **значимым** коэффициентам первой регрессии.
- Тестируйте значимость «в целом» регрессии №7.

- c) Тестируйте совместную значимость факторов `mortno`, `dep`, `male` & `married`.
- d) Тестируйте совместную значимость `arrinc`, `male` & `ujob`.
- e) Какая из альтернативных моделей предпочтительней?