МИНИСТЕРСТВО ОБРАЗОВАНИЯ И НАУКИ РОССИЙСКОЙ ФЕДЕРАЦИИ
Государственное образовательное учреждение высшего профессионального образования
УЛЬЯНОВСКИЙ ГОСУДАРСТВЕННЫЙ ТЕХНИЧЕСКИЙ УНИВЕРСИТЕТ

Н. И. Шанченко

ЭКОНОМЕТРИКА

Лабораторный практикум

Учебное пособие
для студентов высших учебных заведений, обучающихся
по специальности и направлению
«Прикладная информатика (в экономике)»

Ульяновск
УлГТУ
2011

УДК 330.43 (075.8)
ББК 65в6я73
Ш 20

Рецензенты:
доктор физико-математических наук, профессор кафедры информационной безопасности и управления УлГУ, А. С. Андреев;
кафедра общепрофессиональных дисициплин УВАУГА

Утверждено редакционно-издательским советом университета в качестве учебного пособия

Шанченко, Н. И.

Ш 20 Эконометрика: лабораторный практикум : учебное пособие / Н. И. Шанченко. – Ульяновск : УлГТУ, 2011. – 117 с.

ISBN 978-5-9795-0780-4

Содержит описание лабораторных работ по дисциплине «Эконометрика», включая теоретический материал, руководства по выполнению лабораторных работ, необходимые статистические таблицы. Предназначено для студентов экономических и информационных специальностей и направлений.

УДК 330.43 (075.8)
ББК 65в6я73

ISBN 978-5-9795-0780-4

Содержание

Введение

Лабораторный практикум предназначен для формирования у обучающихся практических навыков в области построения эконометрических моделей и применения эконометрических методов исследования.

Лабораторный практикум состоит из введения, шести разделов с описаниями лабораторных работ, шести приложений и библиографического списка литературы и интернет-ресурсов.

Лабораторный практикум включает описание шести лабораторных работ:

1. Парный корреляционный анализ: проверка наличия и степени тесноты линейной и нелинейной связи.
2. Парный регрессионный анализ: построение модели в виде парной регрессии и проверка ее качества.
3. Множественный регрессионный анализ: построение модели в виде уравнения множественной регрессии с учетом только значимых факторов и проверка ее качества.
4. Системы эконометрических уравнений: построение модели в виде системы взаимосвязанных эконометрических уравнений.
5. Динамические эконометрические модели: построение модели авторегрессии и оценка ее качества.
6. Модели стохастических процессов: построение модели стохастического процесса на основе данных временного ряда, оценка ее качества и прогнозирование.

Описание каждой лабораторной работы включает:

– краткое изложение теории;

– контрольные вопросы к разделу;

– задачи к разделу с ответами (ответы приведены в скобках после текста задачи);

– задание к лабораторной работе;

– пример выполнения лабораторной работы.

Лабораторный практикум включает следующие приложения:

– Приложение 1. Таблицы исходных данных;

– Приложение 2. Использование возможностей MS Excel для проведения корреляционного и регрессионного анализа;

– Приложение 3. Использование возможностей пакета Matrixer 5.1 для проведения корреляционного и регрессионного анализа;

– Приложение 4. Статистические таблицы;

– Приложение 5. Функции табличных процессоров MS Excel и OpenOffice. org Calc;

– Приложение 6. Пример отчета по лабораторной работе.

Указания к выполнению лабораторных работ

Перед выполнением лабораторных работ студенты должны проработать методические материалы к лабораторной работе и ознакомиться с соответствующими разделами учебного пособия [6] и ответить на контрольные вопросы.

Разделы учебного пособия, которые необходимо проработать перед выполнением лабораторных работ

Лабораторная работа	Разделы учебного пособия [6]
1	1; 2.7
2	1; 2.1-2.4; 2.6
3	1; 3.1; 3.2; 3.4; 3.6; 3.7; 3.8; 3.9; 3.10
4	1; 4.1; 4.2; 4.4
5	7.1; 7.2.3; 7.3
6	6.1.1; 6.2-6.6

В лабораторных работах 1 и 2 студенты должны самостоятельно провести все вычисления по соответствующим формулам при минимальном использовании компьютера (компьютер применяется только для вычисления результатов алгебраических операций и вычисления элементарных функций).

В лабораторных работах 3 – 6 при построении уравнений регрессии и корреляционного анализа компьютер следует максимально использовать. Рекомендуются табличный процессор MS Excel; эконометрический пакет Matrixer 5.1 и статистические пакеты общего назначения STATISTICA, STATGRAPHICS и др.

Отчетность по лабораторной работе

Лабораторная работа сдается в следующей последовательности:

1) ответы на контрольные вопросы к лабораторной работе;
2) защита отчета о лабораторной работе, представляемого в бумажном и электронном виде.

Требования к оформлению результатов

Отчет о лабораторной работе должен содержать разделы:

1) титульный лист;
2) описание задания лабораторной работы;
3) таблица исходных данных;
4) описание результатов выполнения лабораторной работы (по этапам);
5) итоговое изложение полученных результатов.

1. Парная корреляция

1.1. Коэффициенты корреляции

Корреляционный анализ ставит своей целью проверку наличия и тесноты зависимости между переменными без разделения переменных на зависимые и объясняющие. Ответ на эти вопросы дается с помощью вычисления показателей или коэффициентов корреляции.

По аналитическому выражению зависимости подразделяются на линейные относительно фактора x, определяемые соотношением

$$y = a + b \cdot x \qquad (1.1)$$

и нелинейные, к которым относятся все остальные виды зависимостей, например

$$y = a + b_1 \cdot x + b_2 \cdot x^2 + b_3 \cdot x^3, \; y = a + \frac{b}{x}, \quad y = a \cdot b^x \quad \text{и т. д.}$$

Расчет коэффициентов корреляции основывается на использовании данных наблюдений за совместным изменением величин x и y, которые удобно представить в виде таблицы

Таблица 1.1

Данные наблюдений

	x	y
1	x_1	y_1
2	x_2	y_2
…	…	…
n	x_n	y_n

Каждая строка таблицы представляет собой результат одного наблюдения (x_i, y_i) за величинами x и y, проведенного в одних и тех условиях. Либо это значения двух показателей, характеризующие уровни одного и того же изучаемого объекта в различные моменты или периоды времени. Либо это значения двух показателей, характеризующие различные однородные объекты в один и тот же момент или период времени.

Тесноту связи в случае линейной зависимости характеризуют с помощью линейного коэффициента корреляции r_{xy}

$$r_{xy} = \frac{\frac{1}{n}\sum_i (x_i - \bar{x})(y_i - \bar{y})}{\sigma_x \sigma_y} = \frac{\mathrm{cov}(x, y)}{\sigma_x \sigma_y} \qquad (1.2)$$

или

$$r_{xy} = \frac{\overline{yx} - \bar{y} \cdot \bar{x}}{\sigma_x \sigma_y}, \qquad (1.3)$$

где n – количество наблюдений; x_i, y_i – данные наблюдений; $\bar{x}, \bar{y}$ – средние значения переменных x и y; σ_x, σ_y – средние квадратические отклонения переменных x и y

$$\sigma_x = \sqrt{\frac{1}{n}\sum_i (x_i - \bar{x})^2} = \sqrt{= \overline{x^2} - \bar{x}^2}; \quad \sigma_y = \sqrt{\frac{1}{n}\sum_i (y_i - \bar{y})^2} = \sqrt{\overline{y^2} - \bar{y}^2}\,. \quad (1.4)$$

Линейный коэффициент корреляции r_{xy} принимает значения в диапазоне

$$-1 \le r_{xy} \le 1.$$

При $r_{xy} > 0$ связь является прямой, при $r_{xy} < 0$ – обратной.

Чем ближе величина $|r_{xy}|$ к единице, тем теснее линейная связь и тем лучше линейная зависимость согласуется с данными наблюдений. При $|r_{xy}| = 1$ связь становится функциональной, т. е. соотношение $y_i = a + b \cdot x_i$ выполняется для всех наблюдений.

На практике часто применяется градация степени тесноты связи, приведенная в таблице 1.2.

Таблица 1.2

Количественные критерии оценки тесноты связи

Величина модуля коэффициента корреляции $\lvert r_{xy}\rvert$	Характер связи
$\lvert r_{xy}\rvert < 0{,}3$	Практически отсутствует
$0{,}3 \le \lvert r_{xy}\rvert < 0{,}5$	Слабая
$0{,}5 \le \lvert r_{xy}\rvert < 0{,}7$	Умеренная
$0{,}7 \le \lvert r_{xy}\rvert$	Сильная

Тесноту нелинейной связи, задаваемой соотношением $\hat{y} = f(x)$, оценивают с помощью индекса корреляции R

$$R = \sqrt{1 - \frac{\sum_{i=1}^{n} (\hat{y}_i - y_i)^2}{\sum_{i=1}^{n} (y_i - \bar{y})^2}}, \quad (1.5)$$

где n – количество наблюдений; x_i, y_i – данные наблюдений; $\bar{x}, \bar{y}$ – средние значения переменных x и y; $\hat{y}_i$ – расчетные значения переменной y, вычисленные по уравнению связи, т. е. $\hat{y}_i = f(x_i)$.

Индекс корреляции R принимает значения в диапазоне

$$0 \le R \le 1.$$

Чем ближе величина R к единице, тем теснее данная связь, тем лучше зависимость $\hat{y} = f(x)$ согласуется с данными наблюдений. При $R = 1$ связь становится функциональной, т. е. соотношение $\hat{y}_i = f(x_i)$ выполняется для всех наблюдений.

1.2. Оценка значимости коэффициентов корреляции

Значения линейного коэффициента корреляции r_{xy} либо индекса корреляции R, близкие к нулевому, свидетельствуют о незначительности рассматриваемой зависимости между переменными x и y и случайности величины r_{xy} либо R.

Для оценки статистической значимости полученного значения линейного коэффициента корреляции r_{xy} используется t-критерий Стьюдента, согласно которому значение r_x считается статистически значимым, если выполняется условие

$$t_r = \frac{r_{xy}}{\sqrt{\dfrac{1-r_{xy}^2}{n-2}}} > t_{\text{крит}}, \tag{1.6}$$

где n – количество наблюдений; $t_{\text{крит}} = t_{1-\alpha,n-2}$ представляет собой табличное значение t-критерия Стьюдента при уровне значимости α и числе степеней свободы $k = n–2$ (определяется по таблице П4.2).

Под *уровнем значимости* α понимается вероятность отвергнуть верную гипотезу. Обычно уровень значимости принимается равным $\alpha = 0,05$ или $\alpha = 0,01$.

Для оценки статистической значимости полученного значения индекса корреляции R используется F-критерий Фишера, согласно которому значение R считается статистически значимым, если выполняется условие

$$F_r = \frac{R^2}{1-R^2}(n-2) > F_{\text{крит}}, \tag{1.7}$$

где n – количество наблюдений; $F_{\text{крит}}$ представляет собой табличное значение F-критерия Фишера при уровне значимости α и числе степеней свободы $k_1 = 1$, $k_2 = n - 2$ (определяется по таблицам П4.3, П4.4).

1.3. Построение доверительного интервала

Проверка значимости полученного значения линейного коэффициента корреляции r_{xy} ничего не говорит о том, насколько это значение может отличаться от точного значения. Ответ на этот вопрос дает построение доверительного интервала.

Под *доверительным интервалом* понимаются пределы, в которых лежит точное значение определяемого показателя с заданной вероятностью ($P = 1-\alpha$).

Если в границы доверительного интервала попадает ноль, т. е. нижняя граница отрицательна, а верхняя положительна, то значение r_{xy} принимается равным нулю, так как он не может одновременно принимать и положительное, и отрицательное значения.

Для статистически значимого коэффициента корреляции r_{xy} доверительный интервал получают с использованием Z-преобразования Фишера:

$$z = Z(r_{xy}) = \frac{1}{2} \cdot \ln \frac{1+r_{xy}}{1-r_{xy}}. \tag{1.8}$$

Первоначально определяется интервальная оценка для z

$$z \in \left[z' \pm t_{1-\alpha/2} \cdot \sqrt{\frac{1}{n-3}} \right], \tag{1.9}$$

где $t_{1-\alpha/2}$ – квантиль стандартного нормального распределения порядка 1–α/2, $z' = Z\,(r_{xy})$ – значение Z-преобразования Фишера, соответствующее полученному значению коэффициента корреляции r_{xy}.

Граничные значения доверительного интервала (r^-, r^+) для r_{xy} получаются из граничных значений доверительного интервала (z^-, z^+) для z с помощью обратного Z-преобразования Фишера $r_{xy} = Z^{-1}(z)$

$$r^- = Z^{-1}(z^-); \quad r^+ = Z^{-1}(z^+). \tag{1.10}$$

Контрольные вопросы

1. Как вычисляется линейный коэффициент парной корреляции r_{xy}?
2. Как вычисляется индекс корреляции R?
3. Как осуществляется оценка статистической значимости линейного коэффициента парной корреляции r_{xy}?
4. Как осуществляется оценка статистической значимости индекса корреляции R?
5. Что называется уровнем значимости?
6. Как строится доверительный интервал для линейного коэффициента парной корреляции?

Задачи

1. По величине коэффициента линейной корреляции $r_{xy} = 0{,}46$ определить степень тесноты зависимости между признаками x и y. (Слабая).

2. На основе имеющихся исходных данных определить, какая из двух аналитических зависимостей определяет более тесную взаимосвязь:

1) $y = 54{,}1 + 12{,}5 \cdot x$, $\quad r_{xy} = 0{,}56$;

2) $y = 61{,}2 \cdot 1{,}06^x$, $\quad R = 0{,}74$.

(Вторая).

3. Можно ли говорить о наличии линейной зависимости между переменными x и y, если по 52 наблюдениям было получено значение $r_{xy} = 0{,}42$. Ответ дать с вероятностью ошибки 5%. (Можно).

Лабораторная работа №1. Парный корреляционный анализ: проверка наличия и степени тесноты линейной и нелинейной связи

Задание. На основании данных таблицы П1.1 для соответствующего варианта (табл. 1.3):

1. Вычислить линейный коэффициент парной корреляции r_{xy} и индекс корреляции R.

2. Проверить значимость коэффициента парной корреляции r_{xy} и индекса корреляции R при заданном уровне значимости α.

3. Построить доверительный интервал для значимого линейного коэффициента парной корреляции r_{xy}.

Таблица 1.3

Варианты кривых выравнивания к лабораторной работе № 1

Вариант	Графы из табл. П1.1	Уровень значимости	Зависимость
1	1, 2	0,05	$-65+45\cdot\ln x$
2	1, 3	0,025	$3{,}5\cdot x^{0,9}$
3	1, 4	0,01	$5+4\cdot\ln x$
4	1, 5	0,05	$e^{0,55+0,08\cdot x}$
5	1, 6	0,025	$5{,}9\cdot x^{0,6}$
6	1, 7	0,01	$15+85/x$
7	1, 8	0,05	$22\cdot x^{0,4}$
8	2,3	0,025	$-200+70\cdot\ln x$
9	2, 4	0,01	$7\cdot x^{0,24}$
10	2, 5	0,05	$0{,}008\cdot x^{2}$
11	2, 6	0,025	$13\cdot e^{0,013x}$
12	2, 7	0,01	$40-5\cdot\ln x$
13	2, 8	0,05	$75-35\cdot\ln x$
14	3,4	0,025	$1{,}5+4\cdot\ln x$
15	3,5	0,01	$0{,}055\cdot x^{1,6}$
16	3, 6	0,05	$1{,}5\cdot x^{0,8}$
17	3, 7	0,025	$35-4\cdot\ln x$
18	3, 8	0,01	$40-30\cdot\ln x$
19	4,5	0,05	$12\cdot e^{0,07x}$
20	4,6	0,025	$19\cdot e^{0,04x}$
21	4,7	0,01	$35\cdot x^{-0,2}$
22	4,8	0,05	$25\cdot x^{0,4}$
23	5,6	0,025	$9\cdot x^{0,4}$
24	5,7	0,01	$25-1{,}7\cdot\ln x$
25	5,8	0,05	$14+18\cdot\ln x$

Пример выполнения лабораторной работы №1

Исходные данные:

– наблюдаемые значения переменных x и y заданы в таблице 1.4;

– для расчета индекса корреляции R использовать зависимость $\hat{y} = 0{,}09 \cdot x^{1.25}$;

– уровень значимости $\alpha = 0{,}05$.

Таблица 1. 4

Исходные данные

	Области	x	y		Области	x	y
1	Белгородская	113	39	12	Рязанская область	120	34
2	Брянская	124	37	13	Смоленская	125	39
3	Владимирская	124	36	14	Тамбовская	118	37

Окончание табл. 1.4

1	2	3	4	5	6	7	8
4	Воронежская	122	36	15	Тверская	122	35
5	Ивановская	128	26	16	Тульская	133	54
6	Калужская	140	43	17	Ярославская	136	36
7	Костромская	117	31	18	Архангельская	136	35
8	Курская	113	40	19	Вологодская	138	34
9	Липецкая	122	48	20	Калининградская	124	48
10	Московская	139	64	21	Ленинградская	123	30
11	Орловская	126	39	22	Мурманская	149	59

1) Вычисление σ_x, σ_y и r_{xy} (1.3), (1.4). Используя данные таблицы 1.5 получим

Таблица 1.5

Промежуточные результаты расчетов

Номер наблюдения	x	y	x^2	y^2	xy	$\hat{y}$	$(\hat{y}-y)^2$	$(y-\bar{y})^2$
1	113	39	12769	1521	4407	33,16	34,13	1,00
2	124	37	15376	1369	4588	37,24	0,06	9,00
3	124	36	15376	1296	4464	37,24	1,54	16,00
4	122	36	14884	1296	4392	36,49	0,24	16,00
5	128	26	16384	676	3328	38,75	162,52	196,00
6	140	43	19600	1849	6020	43,34	0,12	9,00
7	117	31	13689	961	3627	34,63	13,19	81,00
8	113	40	12769	1600	4520	33,16	46,81	0,00
9	122	48	14884	2304	5856	36,49	132,44	64,00
10	139	64	19321	4096	8896	42,95	442,90	576,00
11	126	39	15876	1521	4914	37,99	1,01	1,00
12	120	34	14400	1156	4080	35,75	3,05	36,00
13	125	39	15625	1521	4875	37,62	1,91	1,00
14	118	37	13924	1369	4366	35,00	3,99	9,00
15	122	35	14884	1225	4270	36,49	2,22	25,00
16	133	54	17689	2916	7182	40,65	178,23	196,00
17	136	36	18496	1296	4896	41,80	33,63	16,00
18	136	35	18496	1225	4760	41,80	46,23	25,00

Окончание таблицы 1. 5

Номер наблюдения	x	y	x^2	y^2	xy	$\hat{y}$	$(\hat{y}-y)^2$	$(y-\bar{y})^2$
19	138	34	19044	1156	4692	42,57	73,42	36,00
20	124	48	15376	2304	5952	37,24	115,76	64,00
21	123	30	15129	900	3690	36,87	47,14	100,00
22	149	59	22201	3481	8791	46,85	147,58	361,00
Сумма	2792	880	356192	37038	112566	844,081	1488,136	1838
Среднее значение	126,91	40	16190,55	1683,545	5116,636	38,367	67,643	83,545

$$\sigma_x = \sqrt{\overline{x^2} - \bar{x}^2} = \sqrt{(16190{,}55 - (126{,}91)^2} = 9{,}199,$$

$$\sigma_y = \sqrt{\overline{y^2} - \bar{y}^2} = \sqrt{(1683{,}545 - 40^2)} = 9{,}140,$$

$$r_{xy} = \frac{\overline{yx} - \bar{y}\cdot\bar{x}}{\sigma_x\sigma_y} = \frac{5116{,}636 - 126{,}91\cdot 40}{9{,}199\cdot 9{,}140} = 0{,}479.$$

Вычисление R (1.5):

$$R = \sqrt{1 - \frac{\sum_{i=1}^{n}(\hat{y}_i - y_i)^2}{\sum_{i=1}^{n}(y_i - \bar{y})^2}} = \sqrt{1 - \frac{1486{,}136}{1838}} = 0{,}436\,.$$

2) Проверка значимости r_{xy} (1.6).

$$t_r = \frac{r_{xy}}{\sqrt{\frac{1 - r_{xy}^2}{n-2}}} = \frac{0{,}479}{\sqrt{\frac{1 - 0{,}479^2}{22-2}}} = 2{,}44.$$

Для определения $t_{\text{крит}}$ может использоваться статистическая функция СТЬЮДРАСПОБР() из MS Excel (рис. 1.1), либо функция TINV() из OpenOffice.org Calc, либо таблица П4.2 из приложения.

При α = 0,05 и степени свободы $k = n–2 = 20–2 = 20$

$t_{\text{крит}} = t_{1-\alpha,n-2}$ = СТЬЮДРАСПОБР(0,05;20) =2,086.

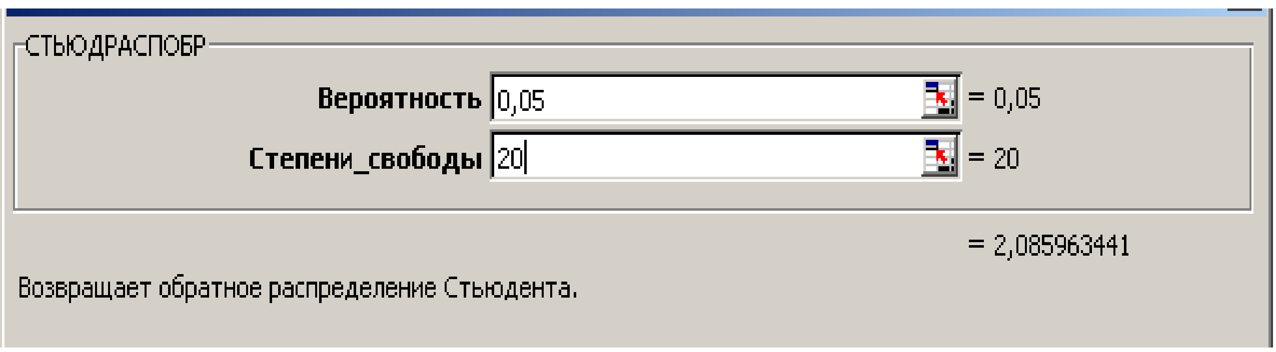


Рис. 1.1. Окно ввода параметров функции СТЬЮДРАСПОБР() MS Excel

Так как

$$t_r = 2{,}44 > t_{1-\alpha,n-2} = 2{,}086,$$

то делаем вывод о *статистической значимости* линейного коэффициента парной корреляции r_{xy}.

Проверка значимости индекса корреляции R (1.7). Значение F-критерия Фишера

$$F_r = \frac{R^2}{1-R^2}(n-2) = \frac{0{,}436^2}{1-0{,}436^2}(22-2) = 4{,}702$$

Для определения $F_{\text{крит}}$ может использоваться статистическая функция FРАСПОБР() из MS Excel (рис. 1.2), либо функция FINV() из OpenOffice.org Calc, либо таблицы П4.3, П4.4 из приложения.

При $\alpha = 0{,}05$ и степенях свободы $k_1 = 1$, $k_2 = n - 2 = 20 - 2 = 20$.

$F_{\text{крит}}$ = FРАСПОБР(0,05;1;20) = 4,35

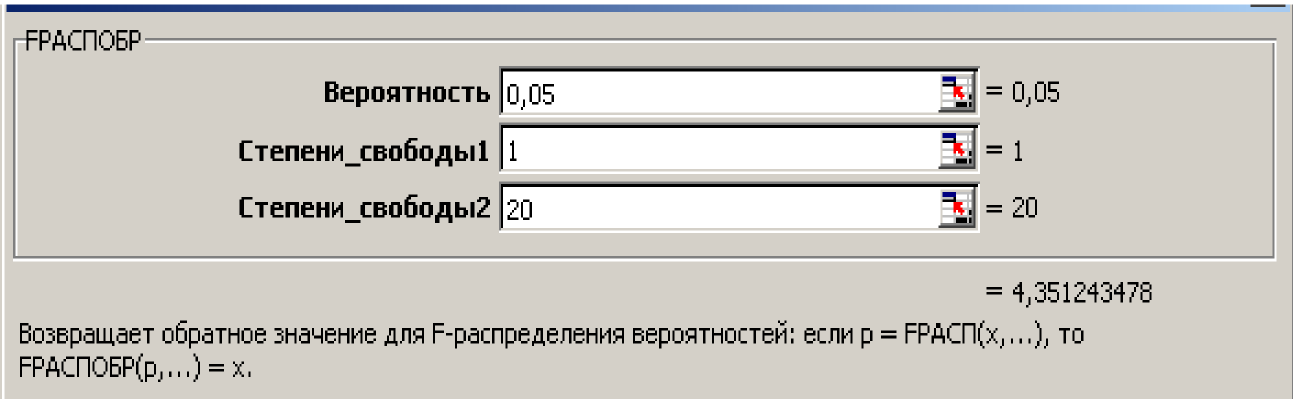


Рис. 1.2. Окно ввода параметров функции FРАСПОБР() MS Excel

Так как

$$F_r = 4{,}702 > F_{\text{крит}} = 4{,}35,$$

то формально с погрешностью 5% индекс корреляции следует считать значимым и следовательно с вероятность 95% нельзя отвергать наличие исследуемой зависимости, но учитывая близость значений $F_r = 4{,}702$ и $F_{\text{крит}} = 4{,}35$ зависимость следует считать практически отсутствующей, нельзя считать адекватной.

3) Построение доверительного интервала для линейного коэффициента корреляции r_{xy} (1.8) – (1.10).

Определим величину z (1.8) Z-преобразования Фишера

$$z = \frac{1}{2}\cdot\ln\frac{1+r_{xy}}{1-r_{xy}} = \frac{1}{2}\cdot\ln\frac{1+0{,}479}{1-0{,}479} = 0{,}522.$$

Для определения $t_{1-\alpha/2}$ – квантиля стандартного нормального распределения порядка $1-\alpha/2 = 1 - 0{,}05/2 = 0{,}975$ может использоваться статистическая функция НОРМСТОБР(0,975) из MS Excel (рис. 1.3), либо функция NORMSINV(0,975) из OpenOffice.org Calc, либо таблицы П4.1 из приложения.

$t_{1-\alpha/2}$ = НОРМСТОБР(0,975) = 1,96.

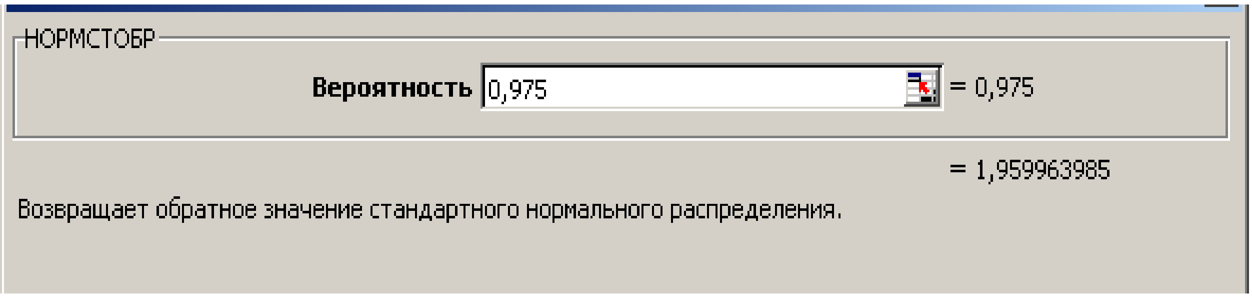


Рис. 1.3. Окно ввода параметров функции НОРМСТОБР() MS Excel

Для получения $t_{1-\alpha/2}$ из таблицы П4.1 нужно использовать соотношение

$$1-\alpha/2 - 0{,}5 = \Phi(t_{1-\alpha/2}),$$

т. е. нужно определить ячейку (клетку) таблицы, содержащую значение $1-\alpha/2 - 0{,}5$ и сложить значение t, соответствующее данной строке с номером столбца, умноженным на 0,01: $t_{1-\alpha/2} = t + N\text{столбца}\cdot 0{,}01$.

Так как $\alpha = 0{,}5$, $1-\alpha/2 - 0{,}5 = 1 - 0{,}05/2 - 0{,}5 = 0{,}475$. Ячейке, содержащей число 0,475, соответствуют $t = 1{,}9$ и Nстолбца = 6, поэтому

$t_{1-\alpha/2} = t + N\text{столбца}\cdot 0{,}01 = 1{,}9 + 0{,}06 = 1{,}96$.

Вычислим $t_{1-\alpha/2}\cdot\sqrt{\dfrac{1}{n-3}} = 1{,}96\cdot\sqrt{\dfrac{1}{22-3}} = 1{,}96\cdot 0{,}2294 = 0{,}450$.

Вычислим границы доверительного интервала (z^-, z^+) для величины z

$$z^- = z' - t_{1-\alpha/2}\cdot\sqrt{\frac{1}{n-3}} = 0{,}522 - 0{,}45 = 0{,}072,$$

$$z^+ = z' + t_{1-\alpha/2}\cdot\sqrt{\frac{1}{n-3}} = 0{,}522 + 0{,}45 = 0{,}972.$$

Определим граничные значения доверительного интервала (r^-, r^+) для r_{xy}. Для определения значения $r = Z^{-1}(z)$ может использоваться статистическая функция ФИШЕРОБР()) из MS Excel (рис. 1.4), либо функция FISHERINV() из OpenOffice.org Calc.

$$r^- = Z^{-1}(z^-) = Z^{-1}(0{,}072) = 0{,}072; \quad r^+ = Z^{-1}(z^+) = Z^{-1}(0{,}972) = 0{,}75.$$

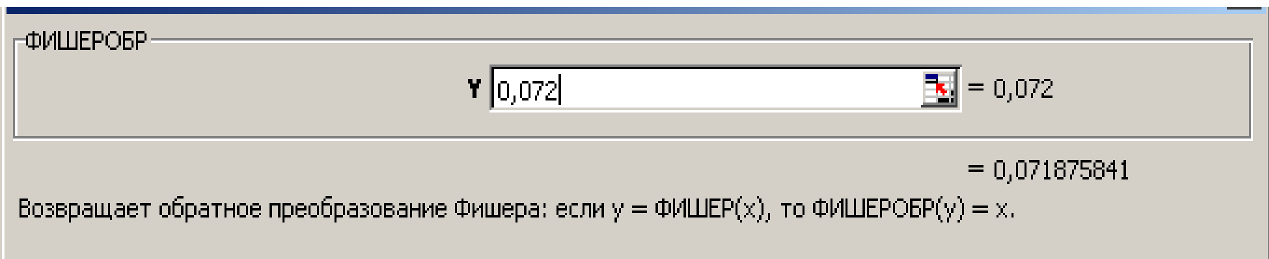


Рис. 1.4. Окно ввода параметров функции ФИШЕРОБР() MS Excel

Искомый доверительный интервал для r_{xy} имеет вид (0,072; 0,75).

Результаты:

1) Линейный коэффициент парной корреляции $r_{xy} = 0{,}479$,

индекс корреляции $R = 0{,}436$.

2) Коэффициенты для r_{xy} и R статистически значимы.

3) Доверительный интервал для r_{xy} – (0,072; 0,75).

2. Парная регрессия

2.1. Понятие регрессии

Парной регрессией называется уравнение связи двух переменных y и x вида

$$y = f(x), \tag{2.1}$$

где y – зависимая переменная (результативный признак); x – независимая, объясняющая переменная (признак-фактор).

Парная регрессия применяется, если имеется доминирующий фактор, который и используется в качестве объясняющей переменной.

Различают линейные и нелинейные относительно фактора x регрессии.

Наиболее часто применяются следующие модели регрессий:

– линейная – $\hat{y} = a + b \cdot x$;
– гиперболическая – $\hat{y} = a + b / x$;
– экспоненциальная – $\hat{y} = e^{a_0 + bx}$;
– модифицированная экспонента – $\hat{y} = K + a_0 \times b^x$;
– параболическая – $\hat{y} = a + b \cdot x + cx^2$;
– показательная – $\hat{y} = a \cdot b^x$;
– степенная – $\hat{y} = a \cdot x^b$;
– логарифмическая – $\hat{y} = a_0 + b \ln x$.

Интерпретация коэффициента b при факторной переменной x в линейной регрессии: коэффициент b показывает, на сколько единиц в среднем изменится величина переменной y при изменении фактора x на 1 единицу измерения.

2.2. Построение уравнения регрессии

Постановка задачи. По имеющимся данным n наблюдений за совместным изменением двух переменных x и y $\{(x_i,y_i),\ i=1,2,...,n\}$ (табл. 1.1) необходимо определить аналитическую зависимость $\hat{y}=f(x)$, наилучшим образом описывающую данные наблюдений.

Построение уравнения регрессии осуществляется в два этапа (предполагает решение двух задач):

– спецификация модели (определение вида аналитической зависимости $\hat{y}=f(x)$);

– оценка параметров (определение численных значений) выбранной модели.

2.2.1. Спецификация модели

Для выбора вида аналитической зависимости применяется три основных метода:

– графический (на основе анализа поля корреляций);

– аналитический, т. е. исходя из теории изучаемой взаимосвязи;

– экспериментальный, т. е. путем сравнения величины остаточной дисперсии $D_{\text{ост}}$ или средней ошибки аппроксимации $\overline{A}$, рассчитанных для различных моделей регрессии (метод перебора).

2.2.2. Оценка параметров линейной модели

Для оценки параметров регрессий, линейных по этим параметрам, используется *метод наименьших квадратов* (МНК). МНК позволяет получить такие оценки параметров, при которых сумма квадратов отклонений фактических значений результативного признака y от теоретических значений $\hat{y}_x$ при тех же значениях фактора x минимальна, т. е.

$$\sum\left(y-\hat{y}_x\right)^2 \to \min.$$

В случае *линейной регрессии* параметры a и b находятся из следующей системы нормальных уравнений метода МНК:

$$\begin{cases} na + b\sum x = \sum y, \\ a\sum x + b\sum x^2 = \sum yx. \end{cases} \tag{2.2}$$

Можно воспользоваться готовыми формулами, которые вытекают из этой системы:

$$a = \bar{y} - b\cdot\bar{x}, \quad b = \frac{\operatorname{cov}(x,y)}{\sigma_x^2} = \frac{\overline{y\cdot x} - \bar{y}\cdot\bar{x}}{\overline{x^2} - \bar{x}^2}. \tag{2.3}$$

2.2.3. Оценка параметров линейных моделей

Нелинейные уравнения регрессии предварительно приводятся к линейному виду с помощью преобразования переменных (обычно рассматриваются модели, для которых такое преобразование возможно), а затем к преобразованным переменным применяется обычный МНК. Для нелинейных уравнений регрессии, приводимых к линейным с помощью преобразования $(x, y) \to (x', y')$, система нормальных уравнений имеет вид (2.2) в преобразованных переменных x', y'. Далее для наиболее часто применяемых нелинейных моделей приводятся линеаризующие преобразования и формулы для расчета параметров.

Гиперболическая регрессия: $y_x = a_0 + a_1 / x$.

Линеаризующее преобразование: $x' = 1/x$; $y' = y$.

Уравнения (2.2) и формулы (2.3) принимают вид

$$\begin{cases} na + b\sum\frac{1}{x} = \sum y, \\ a\sum\frac{1}{x} + b\sum\frac{1}{x^2} = \sum y\frac{1}{x}. \end{cases} \tag{2.4}$$

$$a_0 = \frac{1}{n}\sum y - \frac{1}{n}a_1\sum\frac{1}{x}, \qquad a_1 = \frac{n\sum\frac{y}{x} - \sum\frac{1}{x}\sum y}{n\sum\frac{1}{x^2} - \left(\sum\frac{1}{x}\right)^2}. \tag{2.5}$$

Экспоненциальная регрессия: $y_x = e^{a_0 + a_1 x}$.

Линеаризующее преобразование и расчетные формулы: $x' = x$; $y' = \ln y$.

$$a_0 = \frac{1}{n}\sum \ln y - \frac{1}{n} a_1 \sum x, \qquad a_1 = \frac{n\sum x \ln y - \sum x \sum \ln y}{n\sum x^2 - (\sum x)^2}. \quad (2.6)$$

Модифицированная экспонента: $y_x = K + a_0 \times a_1{}^x$, $(0 < a_1 < 1)$.

Линеаризующее преобразование и расчетные формулы: $x' = x$; $y' = ln\,|y - K|$.

$$\ln|a_0| = \frac{1}{n}\sum \ln|y - K| - \frac{1}{n}\ln a_1 \sum x, \quad a_1 = \frac{n\sum (x \ln|y - K|) - \sum x \sum \ln|y - K|}{n\sum x^2 - (\sum x)^2}. \quad (2.7)$$

Величина предела роста K выбирается предварительно на основе анализа поля корреляций либо из качественных соображений. Параметр a_0 берется со знаком «+», если $y_x > K$ и со знаком «–» в противном случае.

Степенная функция: $y_x = a_0 \times x^{a_1}$, (a_0>0).

Линеаризующее преобразование и расчетные формулы: $x' = ln\,x$; $y' = ln\,y$.

$$\ln a_0 = \frac{1}{n}\sum \ln y - \frac{1}{n} a_1 \sum \ln x, \qquad a_1 = \frac{n\sum (\ln x \cdot \ln y) - \sum \ln x \sum \ln y}{n\sum (\ln x)^2 - (\sum \ln x)^2}. \quad (2.8)$$

Показательная функция: $y_x = a_0 \times a_1^x$.

Линеаризующее преобразование: $x' = x$; $y' = lny$.

$$\ln a_0 = \frac{1}{n}\sum \ln y - \frac{1}{n}\ln a_1 \sum x, \qquad \ln a_1 = \frac{n\sum x \ln y - \sum x \sum \ln y}{n\sum x^2 - (\sum x)^2}. \quad (2.9)$$

Логарифмическая функция: $y_x = a_0 + a_1 \ln x$.

Линеаризующее преобразование: $x' = ln\,x$; $y' = y$.

$$a_0 = \frac{1}{n}\sum y - \frac{1}{n} a_1 \sum \ln x, \qquad a_1 = \frac{n\sum \ln x \cdot y - \sum \ln x \sum y}{n\sum (\ln x)^2 - (\sum \ln x)^2}. \quad (2.10)$$

Парабола второго порядка: $y_x = a_0 + a_1 x + a_2 x^2$.

Линеаризующее преобразование: $x_1' = x$; $x_2' = x^2$; $y' = y$. Этой модели соответствуют две факторные переменные $x_1' = x$; $x_2' = x^2$.

Парабола второго порядка имеет 3 параметра a_0, a_1, a_2, которые определяются из системы трех уравнений

$$\begin{cases} n \cdot a_0 + a_1 \sum x + a_2 \sum x^2 = \sum y, \\ a_0 \sum x + a_1 \sum x^2 + a_2 \sum x^3 = \sum xy, \\ a_0 \sum x^2 + a_1 \sum x^3 + a_2 \sum x^4 = \sum x^2 y. \end{cases} \quad (2.11)$$

2.3. Оценка качества и точности построенной модели регрессии

Качество построенной модели регрессии оценивается с помощью индекса корреляции R (1.5) или *коэффициента детерминации* $R^2 = (R)^2$ вычисляемого как квадрат индекса корреляции. Для линейной регрессии $R^2 = r^2_{xy}$. Коэффициент детерминации R^2 принимает значения в диапазоне

$$0 \le R^2 \le 1.$$

Чем ближе величина R^2 к единице, тем лучше уравнение регрессии $\hat{y} = f(x)$ согласуется с данными наблюдений. При R = 1 зависимость $\hat{y} = f(x)$ становится функциональной, т. е. соотношение $\hat{y}_i = f(x_i)$ выполняется для всех наблюдений.

Коэффициент детерминации R^2 интерпретируется следующим образом. Величина R^2 показывает, какая доля общей дисперсии (вариации) результативного признака y объясняется уравнением регрессии. Например, значение R^2 = 0,75 означает, что уравнение регрессии объясняет 75% общей дисперсии (вариации) результативного признака y.

Точность построенной модели регрессии оценивается с помощью *средней квадратической ошибки* ($\varepsilon_{кв}$)

$$\varepsilon_{кв} = \sqrt{\frac{\sum_{i=1}^{n}(\hat{y}_i - y_i)^2}{n}} \tag{2.12}$$

либо *средней ошибки аппроксимации* ($\overline{A}$), которая представляет собой среднее относительное отклонение расчетных значений от наблюдаемых

$$\overline{A} = \frac{1}{n}\sum_{i=1}^{n}\frac{y - \hat{y}}{y} \cdot 100\,\%. \tag{2.13}$$

Чем выше ниже средняя ошибка аппроксимации $\varepsilon_{кв}$, тем лучше модель регрессии описывает исходные данные.

Построенное уравнение регрессии можно считать удовлетворительным, если значение $\overline{A}$ не превышает 10–12%.

2.4. Оценка значимости уравнения регрессии

Оценка статистической значимости уравнения регрессии в целом осуществляется с помощью F-критерия Фишера по следующему алгоритму:

1) выдвигается нулевая гипотеза H_0 о статистической незначимости уравнения регрессии (или коэффициента детерминации R^2);

2) вычисляется фактическое значение F–критерия $F_{факт}$ и определяется критическое (табличное) значение F–критерия $F_{табл}$;

3) проверяется условие $F_{факт} > F_{табл}$. Если условие выполняется, то нулевая гипотеза H_0 о статистической незначимости уравнения регрессии отвергается и уравнение считается статистически значимым. Если $F_{факт} \le F_{табл}$, то гипотеза H_0 не отклоняется и признается статистическая незначимость или ненадежность уравнения регрессии.

Величина $F_{факт}$ определяется по формуле

$$F_{\hat{o}\grave{a}\hat{e}\grave{o}} = \frac{R^2}{1 - R^2} \cdot \frac{n - p - 1}{p}, \tag{2.14}$$

где n – число единиц совокупности; p – число параметров при факторных переменных. Для парной линейной регрессии $p = 1$.

$F_{табл}$ представляет собой табличное значение F-критерия Фишера при уровне значимости α и числе степеней свободы $k_1 = p$, $k_2 = n - p - 1$.

2.5. Оценка значимости коэффициентов уравнения регрессии

Для оценки статистической значимости коэффициентов линейной регрессии применяется t-критерий Стьюдента:

1) выдвигается нулевая гипотеза H_o о статистической незначимости коэффициента уравнения регрессии;

2) вычисляется фактическое значение t–критерия $t_{факт}$ и определяется критическое (табличное) значение t–критерия $t_{табл}$;

3) проверяется условие $t_{факт} > t_{табл}$. Если условие выполняется, то нулевая гипотеза H_o о статистической незначимости коэффициента уравнения регрессии отвергается и коэффициент уравнения считается статистически значимым. Если $t_{факт} \le t_{табл}$, то гипотеза H_o не отклоняется и признается статистическая незначимость или ненадежность коэффициента уравнения регрессии.

Величины $t_{b,факт}, t_{a,факт}$ определяются по формулам

$$t_{b,факт} = \frac{b}{s_b}; \quad t_{a,факт} = \frac{a}{s_a}, \tag{2.15}$$

где s_a и s_b – стандартные ошибки коэффициентов регрессии (2.16), (2.17).

Величина $t_{крит} = t_{1-\alpha,n-2}$ представляет собой табличное значение t-критерия Стьюдента при уровне значимости α и числе степеней свободы $k = n - p - 1$ (определяется по таблицам).

2.6. Оценка точности коэффициентов уравнения регрессии

Получаемые оценки коэффициентов регрессии зависят от используемой выборки значений переменных x и y и являются случайными величинами. Представление о точности полученных оценок, о том, насколько далеко они могут отклониться от истинных значений коэффициентов, можно получить, используя, так называемые, «стандартные ошибки» коэффициентов регрессии.

Стандартные ошибки коэффициентов регрессии (s_a), (s_b) определяются соотношениями

$$s_b = \sqrt{\frac{\sum_{i=1}^{n}(\hat{y}_i - y_i)^2/(n-2)}{\sum_{i=1}^{n}(x_i - \bar{x})^2}} = \sqrt{\frac{s_{ост}^2}{\sum\left(x - \bar{x}\right)^2}} = \frac{s_{ост}}{\sigma_x\sqrt{n}}, \tag{2.16}$$

$$s_a = \sqrt{\frac{\sum_{i=1}^{n}(\hat{y}_i - y_i)^2}{(n-2)} \cdot \frac{\sum_{i=1}^{n} x_i^2}{n\sum_{i=1}^{n}(x_i - \bar{x})^2}} = \sqrt{s_{ост}^2 \frac{\sum_{i=1}^{n} x_i^2}{n^2\sigma_x^2}} = s_{ост}\frac{\sqrt{\sum_{i=1}^{n} x_i^2}}{n\sigma_x}, \tag{2.17}$$

где $s^2_{ост}$ представляет собой несмещенную оценку остаточной дисперсии

$$s_{ост}^2 = \frac{\sum_{i=1}^{n}(\hat{y}_i - y_i)^2}{(n-2)}. \tag{2.18}$$

Проверка значимости оценок параметров ничего не говорит о том, насколько эти оценки могут отличаться от точных значений. Ответ на этот вопрос дает построение доверительных интервалов.

Под доверительным интервалом понимаются пределы, в которых лежит точное значение определяемого показателя с заданной вероятностью ($P = 1-\alpha$).

Доверительные интервалы для точных значений параметров $\widetilde{a}$ и $\widetilde{b}$ уравнения линейной регрессии определяются соотношениями:

$$a - t_{1-\alpha,n-2} \cdot s_a < \widetilde{a} < a + t_{1-\alpha,n-2} \cdot s_a; \quad b - t_{1-\alpha,n-2} \cdot s_b .< \widetilde{b} < b + t_{1-\alpha,n-2} \cdot s_b \qquad (2.19)$$

Величина $t_{1-\alpha,n-2}$ представляет собой табличное значение t-критерия Стьюдента при уровне значимости α и числе степеней свободы n–2 (п. 2.5).

Если в границы доверительного интервала попадает ноль, т. е. нижняя граница отрицательна, а верхняя положительна, то оцениваемый параметр принимается равным нулю, так как он не может одновременно принимать и положительное, и отрицательное значения.

2.7. Точечный и интервальный прогноз по уравнению линейной регрессии

Точечный прогноз заключается в получении прогнозного значения y_p, которое определяется путем подстановки в уравнение регрессии $\hat{y}_x = a + b \cdot x$ соответствующего (прогнозного) значения x_p

$$y_p = a + b \cdot x_p.$$

Интервальный прогноз заключается в построении доверительного интервала прогноза, т. е. нижней и верхней границ y_{pmin} , y_{pmax} интервала, содержащего точную величину для прогнозного значения y_p ($y_{p\min} < y_p < y_{p\min}$) с заданной вероятностью.

При построении доверительного интервала прогноза используется стандартная ошибка индивидуального значения прогноза s_{y_p}, связанная с дисперсией ошибки прогноза $s^2_{y_p}$ соотношением $s_{y_p} = \sqrt{s^2_{y_p}}$.

$$s_{y_p} = s_{ост} \cdot \sqrt{1 + \frac{1}{n} + \frac{(x_p - \overline{x})^2}{\sum\limits_{i=1}^{n}(x_i - \overline{x})^2}} . \qquad (2.20)$$

Доверительный интервал для индивидуального значения прогноза y_p определяется соотношением:

$$y_p - t_{1-\alpha;n-2} \cdot s_{y_p} \leq \hat{y}_p \leq y_p + t_{1-\alpha;n-2} \cdot s_{y_p} , \qquad (2.21)$$

где величина $t_{1-\alpha,n-2}$ представляет собой табличное значение t-критерия Стьюдента на уровне значимости α при числе степеней свободы n–2.

2.8. Коэффициент эластичности

В экономических исследованиях широкое применение находит такой показатель, как *коэффициент эластичности* (Э), вычисляемый по формуле

$$Э = f'(x)\frac{x}{y}. \quad (2.19)$$

Коэффициент эластичности показывает, на сколько процентов изменится результат y при изменении фактора x на 1% от своего номинального значения. Для линейной регрессии коэффициент эластичности равен

$$Э = b\frac{x}{y} \quad (2.20)$$

и зависит от x, поэтому рассчитывают средний коэффициент эластичности

$$\overline{Э} = f'(\overline{x})\frac{\overline{x}}{\overline{y}} = b\frac{\overline{x}}{\overline{y}}. \quad (2.21)$$

Средний коэффициент эластичности $(\overline{Э})$ показывает, на сколько процентов в среднем по совокупности изменится результат y от своей величины при изменении фактора x на 1% от своего номинального значения.

Контрольные вопросы

1. Что понимается под парной регрессией?
2. Какие задачи решаются при построении уравнения регрессии?
3. Какие методы применяются для выбора вида модели регрессии?
4. Какие функции чаще всего используются для построения уравнения парной регрессии?
5. Какой вид имеет система нормальных уравнений метода наименьших квадратов в случае линейной регрессии?
6. Как вычисляется и что показывает индекс детерминации?
7. Как проверяется значимость уравнения регрессии?
8. Как проверяется значимость коэффициентов уравнения регрессии?
9. Понятие доверительного интервала для коэффициентов регрессии.
10. Понятие точечного и интервального прогноза по уравнению линейной регрессии.
11. Как вычисляются и что показывают коэффициент эластичности Э, средний коэффициент эластичности $\overline{Y}$?

Задачи

1. Из предложенных уравнений регрессии выбрать лучшее, т. е. то, которое дает лучшее приближение к данным наблюдения

$\hat{y}_x = 21{,}5 + 4{,}35 \cdot x$, $R^2 = 0{,}95$,

$\hat{y}_x = 20 + 1{,}115 \cdot \ln(x)$, $R^2 = 0{,}79$.

(Первое)

2. По величине коэффициента детерминации $R^2 = 0{,}56$ определить долю вариации результативного признака, объясненного уравнением регрессии. (56%)

3. Найти критические значения F–критерия и t–критерия по количеству наблюдений и уровню значимости: $n = 50$, $\alpha = 0{,}01$, $m = 1$; $n = 20$, $\alpha = 0{,}05$, $m = 1$, где m – количество факторных переменных. (7,19; 4,41).

4. По величине коэффициента детерминации $R^2 = 0,4$ проверить значимость (α=0,05) уравнения линейной парной регрессии. Число наблюдений n = 50. (Значимо).

5. По заданному уравнению регрессии

$$\hat{y}_x = 21,5 + 4,35 \cdot x,$$

найти средний коэффициент эластичности, если $\bar{x} = 52, \bar{y} = 250$. (0,90).

6. По заданному коэффициенту эластичности Э = 1,5 определить, на сколько изменится y при изменении x на 2 единицы, если до изменения признаки y и x принимали значения $x = 40$, $y = 10$. (0,75).

Лабораторная работа №2. Парный регрессионный анализ: построение модели в виде парной регрессии и проверка ее качества

Задание. На основании данных таблицы П1.2 для соответствующего варианта (табл. 2.1):

1. Построить предложенные в таблице 2.1 уравнения регрессии, включая линейную регрессию, используя формулы (2.3) – (2.11).
2. Вычислить показатели качества и точности для каждого уравнения.
3. Проверить значимость уравнений регрессии при уровнях значимости 0,05 и 0,01.
4. Определить лучшее уравнение регрессии на основе средней ошибки аппроксимации.
5. Проверить значимость коэффициентов линейной регрессии и построить доверительные интервалы для точных значений параметров $\tilde{a}$ и $\tilde{b}$ уравнения линейной регрессии с уровнем значимости 0,05.
6. Построить точечный и интервальный прогноз для значения $x = x_{max}$ по уравнению линейной регрессии с уровнем значимости 0,05.
7. Определить средний коэффициент эластичности по уравнению линейной регрессии.
8. Графически представить результаты моделирования.

Таблица 2.1

Варианты кривых выравнивания к лабораторной работе №2

Вариант	Графы из табл. П1.2 (x, y)	Виды кривых выравнивания					
		Линейная	Степенная	Экспоненциальная	Показательная	Логарифмическая	Гиперболическая
1	1,14	*	*				
2	2,14	*		*			
3	4,14	*			*		
4	6,14	*				*	
5	9,14	*					*
6	11,14	*	*				
7	12,14	*		*			

Окончание таблицы 2.1

Ва-ри-ант	Графы из табл. П1.2	Виды кривых выравнивания					
		Линейная	Степен-ная	Экспо-ненци-альная	Показа-тельная	Логариф-мическая	Гипербо-лическая
8	2,15	*			*		
9	3,15	*				*	
10	7,15	*					*
11	8,15	*	*				
12	12,15	*		*			
13	1,17	*			*		
14	2,17	*				*	
15	4,17	*					*
16	6,17	*	*				
17	9,17	*		*			
18	11,17	*			*		
19	12,17	*				*	
20	1,19	*					*
21	2,19	*	*				
22	4,19	*		*			
23	6,19	*			*		
24	9,19	*				*	
25	11,19	*					*

Пример выполнения лабораторной работы №2

Исходные данные:

- наблюдаемые значения переменных *x* и *y* заданы в таблице 2.2;
- построить две модели: линейную и степенную.

Таблица 2.2

Исходные данные для примера выполнения лабораторной работы №2

	Области	x	y		Области	x	y
1	Белгородская	113	39	12	Рязанская область	120	34
2	Брянская	124	37	13	Смоленская	125	39
3	Владимирская	124	36	14	Тамбовская	118	37
4	Воронежская	122	36	15	Тверская	122	35
5	Ивановская	128	26	16	Тульская	133	54
6	Калужская	140	43	17	Ярославская	136	36
7	Костромская	117	31	18	Архангельская	136	35
8	Курская	113	40	19	Вологодская	138	34
9	Липецкая	122	48	20	Калининградская	124	48
10	Московская	139	64	21	Ленинградская	123	30
11	Орловская	126	39	22	Мурманская	149	59

1) Определим коэффициенты *a* и *b* *линейной регрессии* (2.3), используя результаты промежуточных расчетов, приведенные в таблице 2.3.

Таблица 2.3

Промежуточные результаты расчетов для линейной регрессии

Номер наблюдения	x	y	x^2	y^2	xy	$\hat{y}$	$\lvert(\hat{y}-y)/y\rvert$	$(\hat{y}-y)^2$	$(y-\bar{y})^2$
1	113	39	12769	1521	4407	33,381	0,150	31,57	1,00
2	124	37	15376	1369	4588	38,616	0,007	2,61	9,00
3	124	36	15376	1296	4464	38,616	0,034	6,84	16,00
4	122	36	14884	1296	4392	37,664	0,014	2,77	16,00
5	128	26	16384	676	3328	40,519	0,490	210,81	196,00
6	140	43	19600	1849	6020	46,230	0,008	10,43	9,00
7	117	31	13689	961	3627	35,284	0,117	18,36	81,00
8	113	40	12769	1600	4520	33,381	0,171	43,81	0,00
9	122	48	14884	2304	5856	37,664	0,240	106,84	64,00
10	139	64	19321	4096	8896	45,754	0,329	332,92	576,00
11	126	39	15876	1521	4914	39,567	0,026	0,32	1,00
12	120	34	14400	1156	4080	36,712	0,051	7,36	36,00
13	125	39	15625	1521	4875	39,092	0,035	0,01	1,00
14	118	37	13924	1369	4366	35,760	0,054	1,54	9,00
15	122	35	14884	1225	4270	37,664	0,043	7,10	25,00
16	133	54	17689	2916	7182	42,899	0,247	123,24	196,00
17	136	36	18496	1296	4896	44,326	0,161	69,33	16,00
18	136	35	18496	1225	4760	44,326	0,194	86,98	25,00
19	138	34	19044	1156	4692	45,278	0,252	127,19	36,00
20	124	48	15376	2304	5952	38,616	0,224	88,07	64,00
21	123	30	15129	900	3690	38,140	0,229	66,26	100,00
22	149	59	22201	3481	8791	50,513	0,206	72,04	361,00
Сумма	2792	880	356192	37038	112566	880	3,282	1416,371	1838
Среднее значение	126,91	40	16190,55	1683,545	5116,636	40	0,149	64,381	83,545

$$b=\frac{\overline{y\cdot x}-\bar{y}\cdot\bar{x}}{\overline{x^2}-\bar{x}^2}=\frac{5116{,}636-126{,}91\cdot 40}{16190{,}55-126{,}91^2}=0{,}476,$$

$$a=\bar{y}-b\cdot\bar{x}=40-126{,}91\cdot 0.476=-20{,}39.$$

Уравнение линейной регрессии: $y=-20{,}39+0{,}476\cdot x$.

2) Для построения *степенной модели* $y_x=a_0\times x^{a_1}$ введем новые переменные $x'=\ln x$; $y'=\ln y$, вычислим значения новых переменных и выполним промежуточные расчеты (табл. 2.4).

Таблица 2.4

Промежуточные результаты расчетов для степенной регрессии

Номер наблюдения	*lnx*	*lny*	ln^2x	ln^2y	*lnx·lny*	$\hat{y}$	$\lvert(\hat{y}-y)/y\rvert$	$(\hat{y}-y)^2$	$(y-\bar{y})^2$
1	4,727	3,664	22,348	13,422	17,319	33,850	0,144	26,523	1,000
2	4,820	3,611	23,235	13,039	17,406	38,065	0,044	1,134	9,000
3	4,820	3,584	23,235	12,842	17,274	38,065	0,073	4,265	16,000
4	4,804	3,584	23,079	12,842	17,215	37,291	0,046	1,667	16,000
5	4,852	3,258	23,542	10,615	15,808	39,623	0,558	185,585	196,000
6	4,942	3,761	24,420	14,147	18,587	44,373	0,075	1,884	9,000
7	4,762	3,434	22,678	11,792	16,353	35,371	0,138	19,103	81,000
8	4,727	3,689	22,348	13,608	17,439	33,850	0,165	37,823	0,000
9	4,804	3,871	23,079	14,986	18,597	37,291	0,215	114,681	64,000
10	4,934	4,159	24,349	17,296	20,522	43,973	0,285	401,095	576,000
11	4,836	3,664	23,390	13,422	17,718	38,842	0,015	0,025	1,000
12	4,787	3,526	22,920	12,435	16,882	36,520	0,080	6,353	36,000
13	4,828	3,664	23,313	13,422	17,689	38,453	0,002	0,299	1,000
14	4,771	3,611	22,759	13,039	17,227	35,753	0,034	1,555	9,000
15	4,804	3,555	23,079	12,640	17,080	37,291	0,076	5,249	25,000
16	4,890	3,989	23,916	15,912	19,508	41,588	0,206	154,048	196,000
17	4,913	3,584	24,134	12,842	17,605	42,777	0,231	45,928	16,000
18	4,913	3,555	24,134	12,640	17,466	42,777	0,266	60,483	25,000
19	4,927	3,526	24,278	12,435	17,375	43,573	0,332	91,649	36,000
20	4,820	3,871	23,235	14,986	18,660	38,065	0,196	98,702	64,000
21	4,812	3,401	23,157	11,568	16,367	37,678	0,271	58,947	100,000
22	5,004	4,078	25,039	16,626	20,404	48,007	0,144	120,855	361,000
Сумма	106,50	80,638	515,667	296,556	390,500	863,077	3,596	1437,853	1838,000
Среднее значение	4,841	3,665	23,439	13,480	17,750	39,231	0,163	65,357	83,545

Используя формулы (2.8), получим

$$\ln a_0 = \frac{1}{n}\sum \ln y - \frac{1}{n} a_1 \sum \ln x, \qquad a_1 = \frac{n\sum(\ln x \cdot \ln y) - \sum \ln x \sum \ln y}{n\sum(\ln x)^2 - (\sum \ln x)^2}.$$

Вычислим значения параметров

$$a_1 = \frac{\overline{\ln x \cdot \ln y} + \overline{\ln x} \cdot \overline{\ln y}}{\overline{\ln^2 x} - \overline{\ln x}^2} = \frac{39{,}231 - 4{,}841 \cdot 3{,}665}{23{,}439 - 4{,}841^2} = 1{,}263,$$

$$\ln a_0 = \overline{\ln y} - a_1 \overline{\ln x} = 3{,}665 - 1{,}263 \cdot 4{,}841 = 2{,}451,$$

$$a_0 = \exp(2{,}451) = 0{,}0862.$$

Уравнение степенной регрессии имеет вид: $y = 0{,}0862 \cdot x^{1{,}263}$.

3) Вычисление показателей качества: индекс корреляции R (1.5), коэффициент детерминации R^2, средняя квадратическая ошибка $\varepsilon_{кв}$ (2.12), средняя ошибки аппроксимации $\overline{A}$ (2.13).

Для *линейной* модели (табл. 2.3)

$$R = \sqrt{1 - \frac{\sum_{i=1}^{n}(\hat{y}_i - y_i)^2}{\sum_{i=1}^{n}(y_i - \overline{y})^2}} = \sqrt{1 - \frac{1416{,}371}{1838}} = 0{,}479,$$

$$R^2 = R \cdot R = 0{,}479 \cdot 0{,}479 = 0{,}229,$$

$$\varepsilon_{кв} = \sqrt{\frac{\sum_{i=1}^{n}(\hat{y}_i - y_i)^2}{n}} = \sqrt{\frac{1416{,}371}{22}} = \sqrt{64{,}381} = 8{,}024,$$

$$\overline{A} = \frac{1}{n}\sum_{i=1}^{n}\frac{y - \hat{y}}{y} \cdot 100\,\% = \frac{1}{22} 3{,}282 \cdot 100\% = 14{,}9\%.$$

Для *степенной* модели (табл. 2.4)

$$R = \sqrt{1 - \frac{\sum_{i=1}^{n}(\hat{y}_i - y_i)^2}{\sum_{i=1}^{n}(y_i - \overline{y})^2}} = \sqrt{1 - \frac{1437{,}853}{1838}} = 0{,}467,$$

$$R^2 = R \cdot R = 0{,}467 \cdot 0{,}467 = 0{,}218,$$

$$\varepsilon_{кв} = \sqrt{\frac{\sum_{i=1}^{n}(\hat{y}_i - y_i)^2}{n}} = \sqrt{\frac{1437{,}853}{22}} = \sqrt{65{,}357} = 8{,}084,$$

$$\overline{A} = \frac{1}{n}\sum_{i=1}^{n}\frac{y - \hat{y}}{y} \cdot 100\,\% = \frac{1}{22} 3{,}596 \cdot 100\% = 16{,}3\%.$$

4) Проверка значимости уравнений регрессии (п. 2.4).

Для *линейной* модели (табл. 2.3):

$$F_{факт} = \frac{R^2}{1 - R^2} \cdot \frac{n - m - 1}{m} = \frac{0{,}229}{1 - 0{,}229} \cdot \frac{22 - 1 - 1}{1} = 5{,}95.$$

$F_{крит,0,05}$ = FРАСПОБР(0,05;1;20)=4,35.

$F_{крит,0,01}$ = FРАСПОБР(0,01;1;20)=8,10.

При α = 0,05 линейное уравнение значимо, при α = 0,01 – не значимо.

Для *степенной* модели (табл. 2.4):

$$F_{факт} = \frac{R^2}{1 - R^2} \cdot \frac{n - m - 1}{m} = \frac{0{,}218}{1 - 0{,}218} \cdot \frac{22 - 1 - 1}{1} = 5{,}57.$$

$F_{крит,0,05}$ и $F_{крит,0,01}$ – те же самые.

При α = 0,05 степенное уравнение значимо, при α = 0,01 – не значимо.

5) Определение лучшего уравнения регрессии (по средней ошибке аппроксимации).

Так как $\overline{A}_{лин} = 14,9\% > \overline{A}_{степ} = 16,3\%$, то линейная модель дает меньшую погрешность.

6) Проверка значимости коэффициентов линейной регрессии (п. 2.5).

Определим оценку остаточной дисперсии $s^2_{ост}$, используя данные таблицы 2.3

$$s^2_{ост} = \frac{\sum_{i=1}^{n}(\hat{y}_i - y_i)^2}{(n-2)} = \frac{1416,37}{22-2} = 70,8.$$

Определим стандартные ошибки коэффициентов регрессии (s_a), (s_b), используя значение σ_x, полученное в лабораторной работе №1:

$$s_b = \sqrt{\frac{\sum_{i=1}^{n}(\hat{y}_i - y_i)^2/(n-2)}{\sum_{i=1}^{n}(x_i - \overline{x})^2}} = \frac{\sqrt{s^2_{ост}}}{\sigma_x\sqrt{n}} = \frac{\sqrt{70,8}}{22\cdot 9,199} = 0,0416,$$

$$s_a = \sqrt{\frac{\sum_{i=1}^{n}(\hat{y}_i - y_i)^2}{(n-2)}\cdot\frac{\sum_{i=1}^{n}x_i^2}{n\sum_{i=1}^{n}(x_i - \overline{x})^2}} = \sqrt{s^2_{ост}}\cdot\frac{\sqrt{\sum_{i=1}^{n}x_i^2}}{n\sigma_x} = \sqrt{70,8}\cdot\frac{\sqrt{356192}}{22\cdot 9,199} = 24,81.$$

Вычислим (2.15)

$$t_a = \frac{a}{s_a} = \frac{-20,39}{24,81} = -0,82, \qquad t_b = \frac{b}{s_b} = \frac{0,476}{0,0416} = 11,44.$$

Значение $t_{1-\alpha,n-2}$ определим по таблице П4.2 из приложения.

При $\alpha = 0,05$ и степени свободы $k = n-2 = 20-2 = 20$ получаем $t_{1-\alpha,n-2} = 2,09$.

Так как $|t_a| = 0,82 < t_{1-\alpha,n-2} = 2,09$, то делаем вывод о *незначимости* коэффициента a.

Так как $|t_b| = 11,44 > t_{1-\alpha,n-2} = 2,09$, то делаем вывод о *значимости* коэффициента b.

7) Определение доверительных интервалов для точных значений параметров $\tilde{a}$ и $\tilde{b}$ уравнения линейной регрессии (2.16) – (2.19).

Для точного значения параметра $\tilde{a}$ доверительный интервал определять не нужно, так как значение коэффициента a не значимо.

Доверительный интервал для точного значения параметра $\tilde{b}$

(0,476 – 2,09·0,0416; 0,476 + 2,09·0,0416) = (0,389; 0,563).

8) Построение точечного и интервального прогноза для значения $x = x_{max}$ по уравнению линейной регрессии.

$$x_{max} = 149.$$

Точечный прогноз y_p

$$y_p = -20,39 + 0,476\cdot x_{max} = -20,39 + 0,476\cdot 149 = 50,53.$$

Вычислим

$$s_{y_p} = s_{ост} \cdot \sqrt{1+\frac{1}{n}+\frac{(x_{\max}-\overline{x})^2}{\sum_{i=1}^{n}(x_i-\overline{x})^2}} = \sqrt{70,8} \cdot \sqrt{1+\frac{1}{22}+\frac{(149-126,9)^2}{1861,7}} = 9,62\,.$$

Интервальный прогноз (2.21)

$$((y_p - t_{1-\alpha;n-2} \cdot s_{y_p}\ ;\ y_p + t_{1-\alpha;n-2} \cdot s_{y_p}) =$$

$$= (50,53 - 2,09 \cdot 9,62;\ \ 50,53 + 2,09 \cdot 9,62) = (30,4;\ \ 70,6)$$

8) Определение среднего коэффициента эластичности по уравнению линейной регрессии (2.17).

$$\overline{Y} = b\frac{\overline{x}}{\overline{y}} = 0,476\frac{126,91}{40} = 1,51.$$

Результаты:

1) Уравнение линейной регрессии $y = -20,39 + 0,476 \cdot x$.

2) Уравнение степенной регрессии $y = 0,0862 \cdot x^{1,263}$.

3) Показатели качества и точности:

Для *линейной* модели (табл. 2.3)

$R = 0,479$,

$R^2 = 0,229$,

$\varepsilon_{кв} = 8,024$,

$\overline{A} = 14,9\%$.

Для степенной модели (табл. 2.4)

$R = 0,467$,

$R^2 = 0,218$,

$\varepsilon_{кв} = 8,084$,

$\overline{A} = 16,3\%$.

4) Линейное уравнение значимо при $\alpha = 0,05$ и не значимо при $\alpha = 0,01$.

Степенное уравнение значимо при $\alpha = 0,05$ и не значимо при $\alpha = 0,01$.

5) Линейная модель дает меньшую погрешность.

6) Коэффициент *a незначим*,

коэффициент *b значим*.

7) Доверительный интервал для точного значения параметра $\tilde{b}$ (0,389; 0,563).

8) Точечный прогноз $y_p = 50,53$.

Интервальный прогноз (30,4; 70,6).

9) Средний коэффициент эластичности $\overline{Y} = 1,51$.

10) Графическое представление результатов моделирования (рис. 2.1).

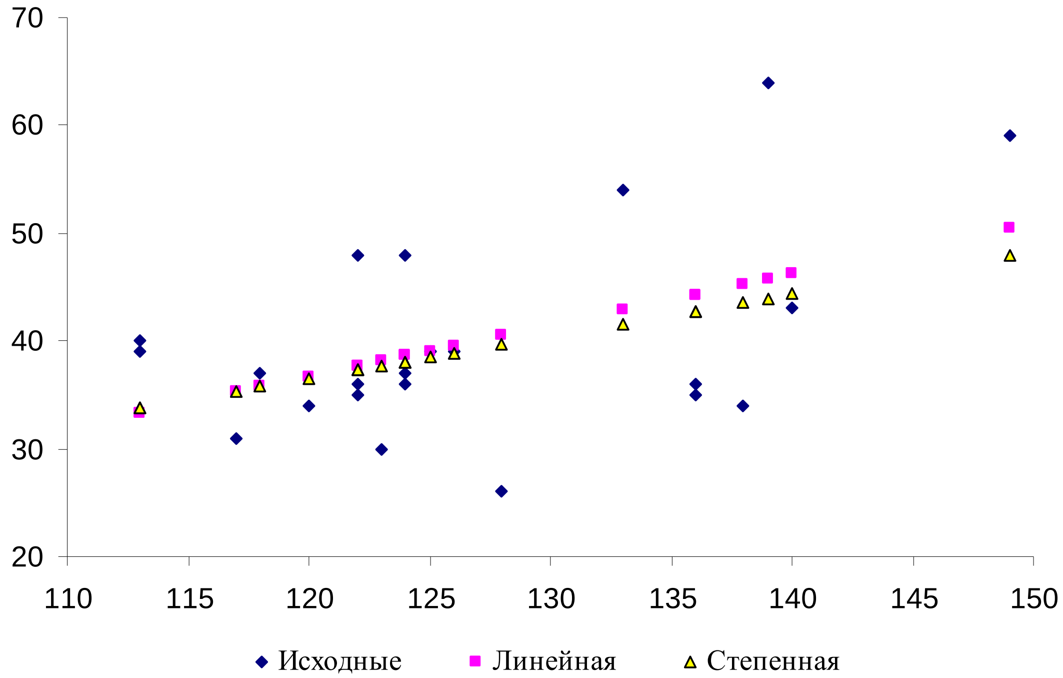


Рис. 2.1. Графическое представление результатов моделирования.

3. Множественная регрессия и корреляция

3.1. Понятие множественной регрессии

Множественная регрессия представляет собой уравнение связи с несколькими независимыми переменными:

$$y = f(x_1, x_2, \ldots, x_p),$$

где y – зависимая переменная (результативный признак); $x_1, x_2, \ldots, x_p$ – независимые переменные (факторы).

Множественная регрессия применяется в ситуациях, когда из множества факторов, влияющих на результативный признак, нельзя выделить один доминирующий фактор и необходимо учитывать влияние нескольких факторов.

Основная цель множественной регрессии – построить модель с большим числом факторов, определив при этом влияние каждого из них в отдельности, а также совокупное их воздействие на моделируемый показатель.

Постановка задачи множественной регрессии. По имеющимся данным n наблюдений за совместным изменением p+1 переменной y *и* x_j и $\{(y_i, x_{j,i});$ $j=1,2,\ldots,p;$ $i=1,2,\ldots,n\}$ (табл. 3.1) необходимо определить аналитическую зависимость $\hat{y} = f(x_1, x_2, \ldots, x_p)$, наилучшим образом описывающую данные наблюдений.

Таблица 3.1

Результаты наблюдений

	y	x_1	x_2	…	x_p
1	y_1	x_{11}	x_{21}	…	x_{p1}
2	y_2	x_{12}	x_{22}	…	x_{p2}
…	…	…	…	…	…
n	y_n	x_{1n}	x_{2n}	…	x_{pn}

Как и в случае парной регрессии, построение уравнения множественной регрессии осуществляется в два этапа:

– спецификация модели;

– оценка параметров выбранной модели.

Спецификация модели включает в себя решение двух задач:

– отбор p факторов x_j, наиболее влияющих на величину y;

– выбор вида уравнения регрессии $\hat{y} = f(x_1, x_2, \ldots, x_p)$;.

3.2. Отбор факторов при построении множественной регрессии

Включение в уравнение множественной регрессии того или иного набора факторов связано, прежде всего, с представлением исследователя о природе взаимосвязи моделируемого показателя с другими экономическими явлениями. Факторы, включаемые во множественную регрессию, должны отвечать следующим требованиям:

1. *Факторы не должны быть взаимно коррелированы* и, тем более, находиться в точной функциональной связи. Если между факторами существует высокая корреляция, то нельзя определить их изолированное влияние на результа-

тивный показатель, и параметры уравнения регрессии оказываются неинтерпретируемыми.

2. Включаемые во множественную регрессию факторы *должны существенно влиять на вариацию независимой переменной*. Т. е. включаемые в модель факторы должны быть статистически значимыми и существенно улучшать показатель качества модели (например, коэффициент детерминации R^2).

Отбор факторов производится на основе качественного теоретико-экономического анализа и обычно осуществляется в две стадии:

– на первой стадии факторы подбираются исходя из сущности проблемы;

– на второй стадии применяются формальные статистические критерии, например, значения t-статистики для соответствующих коэффициентов регрессии.

Наличие высокой корреляции выявляется по значению линейного коэффициента корреляции $r_{x_i x_j}$. Если выполняется условие

$$r_{x_i x_j} \geq 0{,}8\,, \tag{3.1}$$

то факторные переменные x_i, x_j находятся в линейной зависимости между собой, а сами переменные x_i, x_j называются явно коллинеарными.

Значения линейных коэффициентов корреляции $r_{x_i x_j}$ для всевозможных комбинаций переменные x_i, x_j составляют корреляционную матрицу $\{r_{x_i x_j}\}$. Для трех факторов матрица $\{r_{x_i x_j}\}$ принимает вид

$$\left\{r_{x_i x_j}\right\} = \begin{bmatrix} r_{x_1x_1} & r_{x_2x_1} & r_{x_3x_1} \\ r_{x_1x_2} & r_{x_2x_2} & r_{x_3x_2} \\ r_{x_1x_3} & r_{x_2x_3} & r_{x_3x_3} \end{bmatrix}$$

В уравнение регрессии включается только один из коллинеарных факторов, при этом предпочтение отдается тому фактору, который при достаточно тесной связи с результатом имеет наименьшую тесноту связи с другими факторами.

Для преодоления сильной межфакторной корреляции используется ряд подходов:

– *исключение из модели одного или нескольких факторов*;

– *преобразование факторов*, при котором уменьшается корреляция между ними;

– *переход к совмещенным уравнениям регрессии*, т. е. к уравнениям, которые отражают не только влияние факторов, но и их взаимодействие, например
$y = a + b_1 \cdot x_1 + b_2 \cdot x_2 + b_3 \cdot x_3 + b_{12} \cdot x_1 \cdot x_2 + b_{13} \cdot x_1 \cdot x_3 + \varepsilon$,
где члены $b_{12} \cdot x_1 \cdot x_2$, $b_{13} \cdot x_1 \cdot x_3$ выражают взаимодействие факторов.

После исключения коллинеарных факторов осуществляется процедура отбора факторов, наиболее влияющих на изменение результативного признака (факторов, включаемых в регрессию). Наиболее широкое применение получили:

– метод исключения;

– метод включения.

В уравнении регрессии включаются только значимые факторы, что проверяется с помощью критерия Стьюдента.

При отборе факторов рекомендуется, кроме всего прочего, пользоваться следующим правилом: число включаемых факторов должно быть в 6–7 раз меньше объема совокупности, по которой строится регрессия.

3.3. Выбор формы уравнения регрессии

Кроме точности модели для исследователя наиболее важными качествами модели являются простота модели и возможность наглядной интерпретации параметров модели. По этой причине наиболее широко используются линейная и степенная модели.

В уравнении *линейной* множественной регрессии

$$\hat{y}_x = a + b_1 \cdot x_1 + b_2 \cdot x_2 + ... + b_p \cdot x_p \tag{3.2}$$

параметры b_i при x_i называются коэффициентами «чистой» регрессии и интерпретируется следующим образом. Параметры b_i характеризуют среднее изменение результата с изменением соответствующего фактора на единицу при неизмененном значении других факторов, закрепленных на среднем уровне.

В уравнении *степенной* множественной регрессии

$$\hat{y}_x = a \cdot x_1^{b_1} \cdot x_2^{b_2} \cdot ... \cdot x_p^{b_p} \tag{3.3}$$

показатели степеней b_j являются коэффициентами эластичности. Они показывают, на сколько процентов изменяется в среднем результат с изменением соответствующего фактора на 1% при неизменности действия других факторов. Этот вид уравнения регрессии получил наибольшее распространение в производственных функциях, в исследованиях спроса и потребления.

3.4. Оценка параметров уравнения множественной регрессии

Для оценки параметров уравнения множественной регрессии применяют *метод наименьших квадратов* (МНК). Для линейных уравнений регрессии строится система нормальных уравнений, решение которой позволяет получить оценки параметров регрессии. В случае линейной множественной регрессии (3.2) система нормальных уравнений имеет следующий вид:

$$\begin{aligned}
&\sum y = n \cdot a + b_1 \sum x_1 + b_2 \sum x_2 + ... + b_p \sum x_p; \\
&\sum yx_1 = a \sum x_1 + b_1 \sum x_1^2 + b_2 \sum x_2 x_1 + ... + b_p \sum x_p x_1; \\
&\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots \\
&\sum yx_p = a \sum x_p + b_1 \sum x_1 x_p + b_2 \sum x_2 x_p + ... + b_p \sum x_p^2.
\end{aligned} \tag{3.4}$$

3.5. Частные уравнения регрессии

Уравнение линейной множественной регрессии (3.2) позволяет построить, так называемые, частные уравнения регрессии, показывающие зависимость результативного признака от отдельного фактора, при исключении влияния остальных факторов, входящих в уравнение множественной регрессии.

Частные уравнения регрессии получаются из уравнения множественной регрессии (3.2) с помощью замены всех факторов, кроме одного на их средние значения:

$$\begin{aligned} y_{y,x_1} &= a + b_1 \cdot x_1 + b_2 \cdot \bar{x}_2 + b_3 \cdot \bar{x}_3 + ... + b_p \cdot \bar{x}_p + \varepsilon; \\ y_{y,x_2} &= a + b_1 \cdot \bar{x}_1 + b_2 \cdot x_2 + b_3 \cdot \bar{x}_3 + ... + b_p \cdot \bar{x}_p + \varepsilon; \\ &\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\ y_{y,x_p} &= a + b_1 \cdot \bar{x}_1 + b_2 \cdot \bar{x}_2 + ... + b_{p-1} \cdot \bar{x}_{p-1} + b_p \cdot x_p + \varepsilon. \end{aligned} \tag{3.5}$$

Уравнения (3.5) можно представить в виде

$$\begin{aligned} y_{y,x_1} &= A_1 + b_1 \cdot x_1, \\ y_{y,x_2} &= A_2 + b_2 \cdot x_2, \\ &\cdots\cdots\cdots\cdots \\ y_{y,x_p} &= A_p + b_p \cdot x_p, \end{aligned} \tag{3.6}$$

где

$$\begin{aligned} A_1 &= a + b_2 \cdot \bar{x}_2 + b_3 \cdot \bar{x}_3 + ... + b_p \cdot \bar{x}_p; \\ A_2 &= a + b_1 \cdot \bar{x}_1 + b_3 \cdot \bar{x}_3 + ... + b_p \cdot \bar{x}_p; \\ &\cdots\cdots\cdots\cdots\cdots\cdots \\ A_p &= a + b_1\bar{x}_1 + b_2 \cdot \bar{x}_2 + ... + b_{p-1} \cdot \bar{x}_{p-1}. \end{aligned} \tag{3.7)}$$

В отличие от парной регрессии, частные уравнения регрессии характеризуют изолированное влияние фактора на результат, ибо другие факторы закреплены на неизменном уровне. Эффекты влияния других факторов присоединены в них к свободному члену уравнения множественной регрессии.

Частные уравнения регрессии позволяют определить *частные коэффициенты эластичности*

$$Э_{yx_i} = b_i \frac{x_i}{y_{y,x_i}}, \tag{3.8}$$

где b_i – коэффициенты регрессии для фактора x_i в уравнении множественной регрессии; y_{y,x_i} – значение результативного фактора, полученное из частного уравнения регрессии при данном значении фактора x_i.

Средние показатели эластичности можно сравнивать друг с другом и соответственно ранжировать факторы по силе их воздействия на результат.

Средние коэффициенты эластичности для линейной множественной регрессии рассчитываются по формуле

$$\bar{Э}_{yx_i} = b_i \frac{\bar{x}_i}{\bar{y}_{y,x_i}} \tag{3.9}$$

и показывают, на сколько процентов в среднем по совокупности изменится результат y от своей величины при изменении фактора x на 1% от своего значения при неизменных значениях других факторов.

3.6. Множественная корреляция

Коэффициент множественной корреляции характеризует тесноту связи рассматриваемого набора факторов с исследуемым признаком, или, иначе говоря, оценивает тесноту совместного влияния факторов на результат и вычисляется по формуле (1.5).

$$R = R_{yx_1x_2\dots x_p} = \sqrt{1 - \frac{\sum_{i=1}^{n}(\hat{y}_i - y_i)^2}{\sum_{i=1}^{n}(y_i - \bar{y})^2}}, \qquad (3.10)$$

где n – количество наблюдений; x_i, y_i – данные наблюдений; $\bar{x}, \bar{y}$ – средние значения переменных x и y; $\hat{y}_i$ – расчетные значения переменной y, вычисленные по уравнению множественной регрессии, т. е. $\hat{y} = f(x_1, x_2, \dots, x_p)$.

Коэффициент множественной корреляции изменяется от 0 до 1. Чем ближе его значение к 1, тем теснее связь результативного признака со всем набором исследуемых факторов. Величина коэффициента множественной корреляции больше или равна максимальному парному коэффициенту корреляции

$$R_{yx_1x_2\dots x_p} \geq R_{yx_i(\max)} (i = \overline{1, p}).$$

При правильном включении факторов в регрессионный анализ величина индекса множественной корреляции будет существенно отличаться от индекса корреляции парной зависимости.

Квадрат коэффициента множественной корреляции называется *коэффициентом детерминации* и обозначается R^2. Величина коэффициента детерминации используется для оценки качества регрессионной модели. Чем его величина больше, тем лучше данная модель согласуется с данными наблюдений.

Низкое значение коэффициента (индекса) множественной корреляции означает, что либо в регрессионную модель не включены существенные факторы, либо рассматриваемая форма связи не отражает реальные соотношения между переменными, включенными в модель. В этом случае требуются дальнейшие исследования по улучшению качества модели и увеличению ее практической значимости.

3.7. Оценка качества результатов моделирования

Статистическая значимость уравнения множественной регрессии в целом оценивается с помощью F-критерия Фишера (п. 2.4).

Статистическая значимость коэффициентов уравнения множественной регрессии в целом оценивается с помощью t-критерия Стьюдента (п. 2.5).

3.8. Проверка остатков регрессии на гомоскедастичность

Для того чтобы МНК давал надежные оценки параметров линейной регрессии, требуется чтобы дисперсии остатков модели ε

$$y = a + b_1 \cdot x_1 + b_2 \cdot x_2 + \dots + b_p \cdot x_p + \varepsilon \qquad (3.11)$$

для каждого наблюдения были одинаковыми. Остатки, обладающие таким свойством, называются *гомоскедастичными*, а не обладающие – *гетероскедастичными*.

При нарушении гомоскедастичности мы имеем неравенства

$$\sigma^2_{\varepsilon_i} \neq \sigma^2_{\varepsilon_j} \neq \sigma^2, \qquad j \neq i.$$

Для оценки гетероскедастичности можно использовать метод Гольдфельда–Квандта, который проверяет наличие зависимости остатков ε от одной из факторных переменных x_i. Алгоритм применения теста Гольдфельда–Квандта состоит из следующих шагов:

1) исходные данные наблюдений упорядочиваются по мере возрастания выбранной переменной x_i;

2) выделяются первые n_0 и последние n_0 наблюдений и исключаются из рассмотрения $C = n–2n_0$ центральных наблюдений. При этом должно выполняться условие $n_0 > p$, где p – число оцениваемых параметров;

3) для каждой из групп наблюдений оцениваются уравнения регрессии остатков ε по значимым факторам

$$\varepsilon = a + b_1 \cdot x_1 + b_2 \cdot x_2 + ... + b_p \cdot x_p + u\,; \tag{3.12}$$

4) для каждого уравнения определяются остаточные суммы квадратов (S_1) и (S_2) остатков u_i и находится их отношение: $R = \max(S_2, S_1) / \min(S_2, S_1)$.

Если выполняется условие

$$R > F_{табл}, \tag{3.13}$$

где $F_{табл}$ представляет собой табличное значение F-критерия Фишера при уровне значимости α и числе степенях свободы $k_1 = n_0 – p$, $k_2 = n_0 – p$, то предпосылка о *равенстве дисперсий* остаточных величин отвергается с уровнем значимости α.

Чем больше величина R превышает табличное значение критерия $F_{табл}$, тем более нарушена предпосылка о равенстве дисперсий остаточных величин.

Авторами метода рекомендовано для случая одного фактора n=20 принимать C=4, при n=30 принимать C=8, при n=60 принимать C=16.

Контрольные вопросы

1. Что понимается под множественной регрессией?
2. Какие задачи решаются при построении уравнения регрессии?
3. Какие задачи решаются при спецификации модели?
4. Какие требования предъявляются к факторам, включаемым в уравнение регрессии?
5. Что понимается под коллинеарностью факторов?
6. Как проверяется наличие коллинеарности?
7. Какие подходы применяются для преодоления межфакторной корреляции?
8. Какие функции чаще используются для построения уравнения множественной регрессии?
9. По какой формуле вычисляется индекс множественной корреляции?
10. Как вычисляются индекс множественной детерминации?
11. Что означает низкое значение коэффициента множественной корреляции?
12. Как проверяется значимость уравнения регрессии и отдельных коэффициентов?
13. Как строятся частные уравнения регрессии?
14. Как вычисляются средние частные коэффициенты эластичности?
15. Что понимается под гомоскедастичностью ряда остатков?
16. Как проверяется гипотеза о гомоскедастичности ряда остатков?

Задачи

1. Проверить наличие линейной коллинеарности между факторами x, z, t, если корреляционная матрица имеет вид

	x	z	t
x	1		
z	0,35	1	
t	0,56	0,86	1

(Коллинеарны z и t).

2. Дать интерпретацию параметрам уравнения регрессии

$$\hat{y}_x = 21{,}5 + 4{,}35 \cdot x + 2{,}1 \cdot z.$$

(При увеличении x на 1 величина $\hat{y}_x$ изменится на 4,35, а при увеличении z на 1 – на 2,1).

3. Дать интерпретацию параметрам уравнения регрессии

$$\hat{y}_x = 3{,}4 \cdot x^{0,4} \cdot z^{1,1}.$$

(При увеличении x на 1% величина $\hat{y}_x$ изменится на 0,4%, а при увеличении z на 1% – на 1,1%).

4. По заданному уравнению регрессии

$$\hat{y}_x = 20 + 4 \cdot x + 2{,}5 \cdot z$$

построить частные уравнения регрессии, если $\bar{x} = 5$, $\bar{z} = 20$.

($\hat{y}_x = 70 + 4 \cdot x$, $\hat{y}_x = 40 + 2{,}5 \cdot z$).

5. По заданному уравнению регрессии

$$\hat{y}_x = 20 + 4 \cdot x + 2{,}5 \cdot z$$

найти коэффициенты эластичности, если $\bar{x} = 5$, $\bar{z} = 20$, $\bar{y} = 100$. (0,2; 0,5).

6. По величине коэффициента детерминации $R^2 = 0{,}45$ определить долю вариации результативного признака, объясненного уравнением регрессии.

(45%)

7. Из предложенных уравнений выбрать лучшее

$\hat{y}_x = 21{,}5 + 4{,}35 \cdot x - 0{,}2 \cdot x^2 + 2{,}1 \cdot z$, $R^2 = 0{,}456$,

$\hat{y}_x = 3{,}4 \cdot x^{0,4} \cdot z^{1,1}$, $R^2 = 0{,}56$.

(Второе).

8. Найти критические значения F–критерия и t–критерия по количеству наблюдений и уровню значимости: n = 50, α =0,01, m = 2; n = 20, α =0,05, m = 3, где m – количество факторов в уравнении регрессии. (5,09; 2,81).

9. По величине множественного коэффициента корреляции $r_{xy} = 0{,}56$ для уравнения регрессии

$$\hat{y}_x = 21{,}5 + 4{,}35 \cdot x + 2{,}1 \cdot z,$$

проверить его значимость (α =0,05). Число наблюдений n = 25. (Значимо).

Лабораторная работа №3. Множественный регрессионный анализ: построение модели в виде уравнения множественной регрессии с учетом только значимых факторов и проверка ее качества

Задание. На основании данных табл. П1.2 для соответствующего варианта (табл. 3.2):

1) Проверить факторы на наличие коллинеарности. Отобрать неколлинеарные факторы.

2) Построить уравнение линейной множественной регрессии.

3) Определить значения коэффициента множественной корреляции и коэффициента детерминации.

4) Проверить значимость уравнения при заданном уровне значимости.

5) Проверить значимость коэффициентов уравнения при заданном уровне значимости.

6) Построить уравнение линейной множественной регрессии с учетом только значимых факторов.

7) Проверить гипотезу о гомоскедастичности ряда остатков с уровнем значимости $\alpha = 0{,}05$.

8) Построить частные уравнения регрессии.

9) Определить средние частные коэффициенты эластичности.

Указания к решению. При выполнении лабораторной работы использовать возможности надстройки «Анализ данных» табличного процессора MS Excel (для расчета корреляционной матрицы, нахождения уравнений регрессии, нахождения коэффициентов координации и др.), либо программного пакета Matrixer 5.1, либо какого-либо другого статистического или эконометрического программного пакета.

Таблица 3.2

Варианты выполнения лабораторной работы № 3

Варианты	Номер графы для переменной y (табл. П1.2)	Номера граф для переменных-факторов (табл. П1.2)	Уровень значимости α
1	14	1,2,3	0,05
2	15	1,2,3	0,01
3	16	1,2,3	0,05
4	17	1,2,3	0,01
5	18	1,2,3	0,05
6	14	2,3,4	0,01
7	15	2,3,4	0,05
8	17	2,3,4	0,01
9	18	2,3,4	0,05
10	15	3,4,5	0,01
11	19	3,4,5	0,05
12	14	6,7,8	0,01
13	15	6,7,8	0,05
14	16	6,7,8	0,01
15	17	6,7,8	0,05

Окончание таблицы 3.2

Вари-анты	Номер графы для переменной y (табл. П1.2)	Номера граф для перемен-ных-факторов (табл. П1.2)	Уровень значимости α
16	18	6,7,8	0,05
17	14	7,8,9	0,01
18	15	7,8,9	0,05
19	17	7,8,9	0,01
20	18	7,8,9	0,05
21	15	8,9,10	0,01
22	19	8,9,10	0,05
23	16	11,12,13	0,01
24	17	11,12,13	0,05
25	18	11,12,13	0,01

Пример выполнения лабораторной работы №3

Исходные данные:

– данные наблюдений переменных y и x_1, x_2, x_3 даны в таблицы 3.3;

– уровень значимости α = 0,05.

Таблица 3.3

Исходные данные для примера выполнения лабораторной работы №3

	Области	y	x_1	x_2	x_3		Области	y	x_1	x_2	x_3
1	Белгородская	113	10	38	163	113	Рязанская	120	16	38	160
2	Брянская	124	5	37	165	124	Смоленская	125	10	37	167
3	Владимирская	124	10	38	163	124	Тамбовская	118	12	37	163
4	Воронежская	122	13	36	163	122	Тверская	122	8	38	162
5	Ивановская	128	9	37	152	128	Тульская	133	29	39	184
6	Калужская	140	14	39	176	140	Ярославская	136	9	39	167
7	Костромская	117	12	36	155	117	Архангельская	136	91	37	166
8	Курская	113	15	36	164	113	Вологодская	138	14	39	166
9	Липецкая	122	13	38	175	122	Калининградская	124	12	38	176
10	Московская	139	27	38	194	139	Ленинградская	123	11	38	156
11	Орловская	126	8	39	167	126	Мурманская	149	8	39	194
12	Оренбургская	125	17	38	164	125	Астраханская	126	11	38	182
13	Пензенская	124	7	37	175	124	Волгоградская	109	8	37	164
14	Пермская	121	15	37	162	121	Ростовская	120	20	38	170
15	Самарская	123	24	38	168	123	Ульяновская	115	16	37	165

1) *Проверка факторов на наличие коллинеарности* (п. 3.2). *Отбор неколлинеарных факторов*.

Построим корреляционную матрицу, используя функцию «Сервис.Анализ данных.Корреляция» табличного процессора MS Excel (приложение 2).

	y	x_1	x_2	x_3
y	1			
x_1	0,263	1		
x_2	0,605	-0,071	1	
x_3	0,599	0,091	0,471	1

Рис. 3.1 Корреляционная матрица

Из матрицы следует, что $r_{x_1x_2} = -0{,}071$, $r_{x_1x_3} = 0{,}091$, $r_{x_2x_3} = 0{,}471$, следовательно, коллинеарность между факторами отсутствует и нет основания исключать какой-либо фактор из рассмотрения.

Таким образом, далее будет строиться регрессия y по факторам x_1, x_2 и x_3.

2) *Построение уравнения линейной множественной регрессии.*

Для построения уравнения линейной регрессии используем функцию «Сервис.Анализ данных.Регрессия» табличного процессора MS Excel (рис 3.2):

1) вызов функции осуществляется через пункты меню: <Сервис> – <Анализ данных> – <Регрессия>;

2) указываются ячейки, содержащие исходные значения переменных y и x_i (рис. 3.2);

3) если отсутствует свободный член в уравнении регрессии – установить флажок «Константа–ноль» (рис. 3.2);

4) указать место, где будут представлены результаты работы функции (выходной интервал на данном рабочем листе, новый рабочий лист, новая рабочая книга);

5) искомые значения коэффициентов линейного уравнения регрессии (a, b_i) берутся из столбца «Коэффициенты» таблицы результатов регрессии (табл. 3.6).

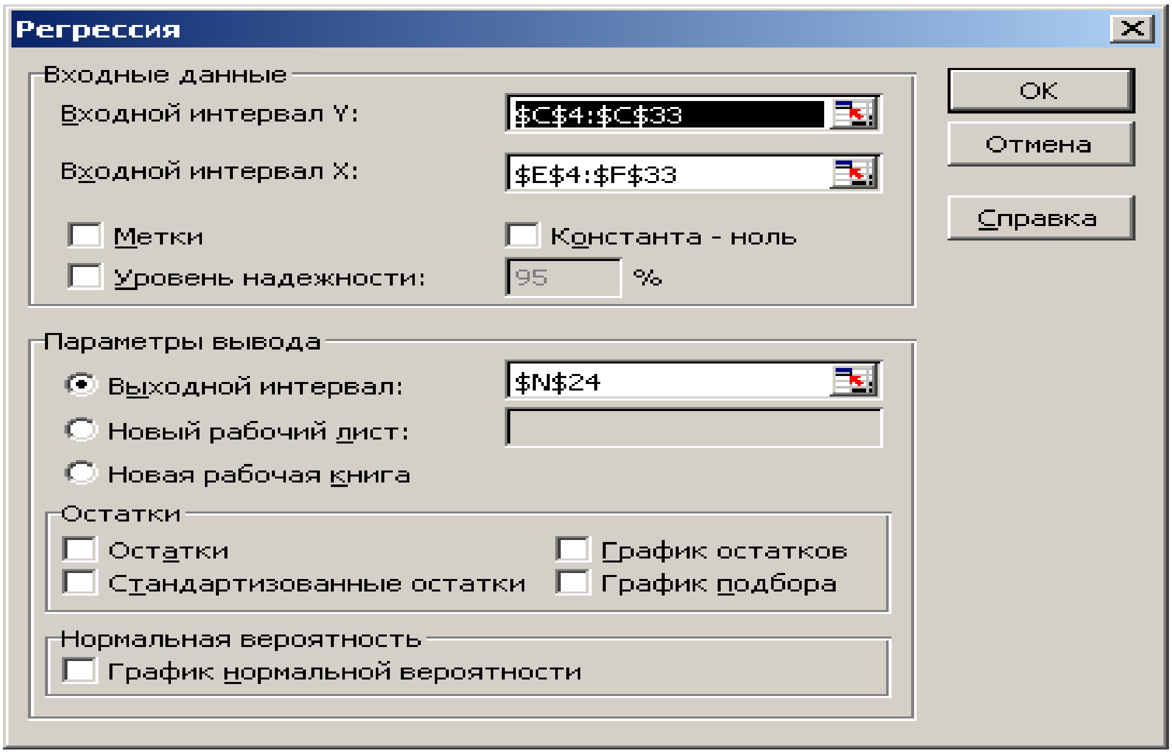


Рис. 3.2. Окно ввода параметров регрессии MS Excel

Результаты работы функции приведены в таблицах 3.4, 3.5, 3.6.

Таблица 3.4

Результаты корреляционного анализа

Множественный *R*	0,748	*Множественный коэффициент корреляции R*
R-квадрат	0,560	*Коэффициент детерминации R^2*
Нормированный *R*-квадрат	0,509	*Модифицированный коэффициент детерминации R*
Стандартная ошибка	6,302	*Стандартная ошибка определения R*
Наблюдения	30	*Число наблюдений*

Таблица 3.5

Результаты дисперсионного анализа

Пояснения	Число степеней свободы *df*	Сумма квадратов отклонений *SS*	Дисперсия на 1 степень свободы *MS*	Статистика Фишера *F*	Уровень значимости *Значимость F*
Регрессия	3	1311,7	437,2	11,011	7,55E-05
Остаток	26	1032,4	39,7		
Итого	29	2344,2			

Таблица 3.6

Результаты регрессионного анализа

Пояснения	Коэффициенты уравнения регрессии	Стандартная ошибка определения коэффициентов	t-статистика	Вероятность ошибки α	Нижние 95%-пределы	Верхние 95%-пределы
Показатели	*Коэффициенты*	Стандартная ошибка	t-статистика	P-Значение	Нижние 95%	Верхние 95%
Y-пересечение	–99,816	48,6093	–2,0534	0,0502	–199,7334	0,1023
Переменная X 1	0,154	0,0775	1,9856	0,0577	–0,0054	0,3131
Переменная X 2	4,459	1,4617	3,0504	0,0052	1,4542	7,4634
Переменная X 3	0,324	0,1337	2,4203	0,0228	0,0488	0,5985

Из таблицы 3.6 следует, что уравнение регрессии имеет вид

$$y = -99{,}816 + 0{,}154 \cdot x_1 + 4{,}459 \cdot x_2 + 0{,}324 \cdot x_3.$$

3) *Определение значений коэффициента множественной корреляции R и коэффициента детерминации R^2.*

Как следует из таблицы 3.2

$$R = 0{,}748; \qquad R^2 = 0{,}560.$$

4) Проверка значимости уравнения регрессии (п. 2.4).

Применим F-критерий Фишера.

Вычислим фактическое значение критерия (2.14)

$$F_{факт} = \frac{R^2}{1-R^2} \cdot \frac{n-m-1}{m} = \frac{0{,}560}{1-0{,}560} \cdot \frac{30-3-1}{3} = 11{,}01.$$

Это же значение $F_{факт}$ можно было взять из таблицы 3.4.

Определим критическое значение критерия $F_{крит}$ F-критерия Фишера, используя функцию MS Excel «FРАСПОБР()»:

- уровень значимости $\alpha = 0{,}05$;
- число степеней свободы $k_1 = m = 3$; $k_2 = n - m - 1 = 30 - 3 - 1 = 26$;
- $F_{крит}$ = FРАСПОБР(0,05; 3; 26) = 2,98.

Так как $F_{факт} = 11{,}01 > F_{крит} = 2{,}28$, то делаем вывод о значимости построенного уравнения регрессии.

Из таблицы 3.5 следует, что уровень значимости уравнения регрессии $\alpha = 7{,}55 \cdot 10^{-5}$, т. е. заведомо ниже требуемого уровня $\alpha = 0{,}05$, т. е. уравнение значимо и при более низком уровне значимости.

5) *Проверка значимости коэффициентов уравнения регрессии* (п. 2.5).

Применим t-критерий Стьюдента. Из таблицы 3.6 следует, что уровни значимости коэффициентов уравнения регрессии имеют значения:

$\alpha_a = 0{,}050$; $\alpha_{b1} = 0{,}058$; $\alpha_{b2} = 0{,}005$; $\alpha_{b3} = 0{,}023$.

Таким образом, оценки параметров a, b_2, b_3 значимы при уровне значимости $\alpha = 0{,}05$, а значение b_1 не значимо при уровне значимости $\alpha = 0{,}05$.

6) *Построение уравнения линейной множественной регрессии* с учетом только значимых факторов.

Значимыми факторами являются x_2, x_3.

Для построения уравнения линейной регрессии используем функцию «Сервис.Анализ данных.Регрессия» табличного процессора MS Excel (рис 3.1).

Задав соответствующие диапазоны данных в окне ввода параметров регрессии (рис. 3.2), получим

Множественный коэффициент корреляции $R = 0{,}702$,

Коэффициент детерминации $R^2 = 0{,}493$,

$F_{факт} = 13{,}12$,

уровень значимости уравнения регрессии $\alpha = 0{,}0001$.

Таблица 3.7

Результаты регрессионного анализа

Показатели	Коэффициенты уравнения регрессии	Стандартная ошибка определения коэффициентов	t-статистика	Вероятность ошибки α	Нижние 95%–пределы	Верхние 95%–пределы
Y-пересечение	−89,520	50,898	−1,759	0,090	−193,954	14,914
Переменная X 2	4,082	1,526	2,674	0,013	0,950	7,214
Переменная X 3	0,361	0,139	2,592	0,015	0,075	0,647

Из таблицы 3.7 следует, что уравнение регрессии имеет вид

$$y = -89{,}520 + 4{,}082 \cdot x_2 + 0{,}361 \cdot x_3.$$

7) *Проверка гипотезы о гомоскедастичности ряда остатков* с уровнем значимости $\alpha = 0{,}05$ (п. 2.9).

Рассчитаем модельные значения переменной по уравнению регрессии и соответствующие остатки (табл. 3.8).

Таблица 3.8

Промежуточные результаты расчета

Наблюдения	y	$x2$	$x3$	Модельные значения $\hat{y} = f(x)$	Отклонения $\varepsilon = \hat{y} - y)$
1	113	38	163	124,49	11,49
2	124	37	165	121,13	-2,87
3	124	38	163	124,49	0,49
4	122	36	163	116,32	-5,68
5	128	37	152	116,43	-11,57
6	140	39	176	133,27	-6,73
7	117	36	155	113,43	-3,57
8	113	36	164	116,69	3,69
9	122	38	175	128,82	6,82
10	139	38	194	135,69	-3,31
11	126	39	167	130,02	4,02
12	125	38	164	124,85	-0,15
13	124	37	175	124,74	0,74
14	121	37	162	120,05	-0,95
15	123	38	168	126,29	3,29
16	120	38	160	123,40	3,40
17	125	37	167	121,85	-3,15
18	118	37	163	120,41	2,41
19	122	38	162	124,13	2,13
20	133	39	184	136,16	3,16
21	136	39	167	130,02	-5,98
22	136	37	166	121,49	-14,51
23	138	39	166	129,65	-8,35
24	124	38	176	129,19	5,19
25	123	38	156	121,96	-1,04
26	149	39	194	139,77	-9,23
27	126	38	182	131,35	5,35
28	109	37	164	120,77	11,77
29	120	38	170	127,02	7,02
30	115	37	165	121,13	6,13

Отсортируем данные таблицы 3.8 по переменным x_2 и x_3 (табл. 3.9).

Таблица 3.9

Результаты сортировки данных таблицы 3.8 по переменным x_2 и x_3

Результаты сортировки по x_2			Результаты сортировки по x_3		
Отклонения $\varepsilon = \hat{y} - y)$	x2	x3	Отклонения $\varepsilon = \hat{y} - y)$	x2	x3
-5,68	36	163	-11,57	37	152
-3,57	36	155	-3,57	36	155
3,69	36	164	-1,04	38	156
-14,51	37	166	3,40	38	160
-11,57	37	152	-0,95	37	162
-3,15	37	167	2,13	38	162
-2,87	37	165	-5,68	36	163
-0,95	37	162	2,41	37	163
0,74	37	175	0,49	38	163
2,41	37	163	11,49	38	163
6,13	37	165	3,69	36	164
11,77	37	164	11,77	37	164
-3,31	38	194	-0,15	38	164
-1,04	38	156	-2,87	37	165
-0,15	38	164	6,13	37	165
0,49	38	163	-14,51	37	166
2,13	38	162	-8,35	39	166
3,29	38	168	-3,15	37	167
3,40	38	160	-5,98	39	167
5,19	38	176	4,02	39	167
5,35	38	182	3,29	38	168
6,82	38	175	7,02	38	170
7,02	38	170	0,74	37	175
11,49	38	163	6,82	38	175
-9,23	39	194	5,19	38	176
-8,35	39	166	-6,73	39	176
-6,73	39	176	5,35	38	182
-5,98	39	167	3,16	39	184
3,16	39	184	-3,31	38	194
4,02	39	167	-9,23	39	194

Темным цветом в таблице 3.9 выделены данные, которые исключаются из дальнейшего рассмотрения.

а) Для проверки гомоскедастичности ряда остатков по переменной x_2 построим два уравнения регрессии $\varepsilon = a + b_2 \cdot x_2 + b_3 \cdot x_3$ по первой и второй группе отсортированных данных по переменной x_2 из таблицы 3.9 и определим остаточные суммы квадратов (S_1) и (S_2) остатков при n_0 = 11 и p = 2:

1) $\varepsilon = 27{,}071 - 2{,}675 \cdot x_2 + 0{,}419 \cdot x_3$, $S_1 = 326{,}42$,

2) $\varepsilon = 436{,}74 - 10{,}70 \cdot x_2 + -0{,}133 \cdot x_3$, $S_2 = 184{,}00$.

Вычислим отношение

$R = \max(S_2, S_1) / \min(S_2, S_1) = 326{,}42/184{,}00 = 1{,}77$

и определим критическое значение F-критерия Фишера $F_{крит}$

$F_{крит}$ = FРАСПОБР(0,05; $n_0 - p$; $n_0 - p$) =

=FРАСПОБР(0,05; 11–2; 11– 2) = FРАСПОБР(0,05; 9; 9) = 3,18.

Так как не выполняется условие

$R = 1{,}77 > F_{крит} = 3{,}18$,

то нет оснований отвергать предположение о постоянстве дисперсии остатков (остатки гомоскедастичны) по переменной x_2.

8) Для проверки гомоскедастичности ряда остатков по переменной x_3 построим два уравнения регрессии $\varepsilon = a + b_2 \cdot x_2 + b_3 \cdot x_3$ по первой и второй группе отсортированных данных по переменной x_3 из таблицы 3.9 и определим остаточные суммы квадратов (S_1) и (S_2) остатков

1) $\varepsilon = -249{,}8 - 2{,}687 \cdot x_2 + 0{,}936 \cdot x_3$, $S_1 = 142{,}2$,

2) $\varepsilon = 154{,}6 - 2{,}456 \cdot x_2 - 0{,}332 \cdot x_3$, $S_2 = 166{,}4$.

Вычислим отношение

$R = \max(S_2, S_1) / \min(S_2, S_1) = 166{,}4/142{,}2 = 1{,}17$.

Критическое значение F-критерия Фишера то же самое $F_{крит} = 3{,}18$.

Так как не выполняется условие

$R = 1{,}17 > F_{крит} = 3{,}18$,

то нет оснований отвергать предположение об постоянстве дисперсии остатков (остатки гомоскедастичны) по переменной x_3.

9) *Построение частных уравнений регрессии* (п. 3.5, формулы (3.6), (3.7)) на основании уравнения

$y = -89{,}520 + 4{,}082 \cdot x_2 + 0{,}361 \cdot x_3$.

Определим средние значения переменных используя функции СРЗНАЧ() табличного процессора MS Excel $\bar{x}_2, \bar{x}_3, \bar{y}$

$\bar{x}_2$ = СРЗНАЧ() = 37,70; $\bar{x}_3 = 168{,}27$; $\bar{y} = 125{,}17$.

Вычислим свободные члены частных уравнений регрессии (3.7)

$$A_2 = a + b_3 \cdot \bar{x}_3 = -89{,}520 + 0{,}361 \cdot 168{,}27 = -28{,}77;$$

$$A_3 = a + b_2 \cdot \bar{x}_2 = -89{,}520 + 4{,}082 \cdot 37{,}70 = 64{,}37.$$

Частные уравнения регрессии

$y_{y,x_2} = -28{,}77 + 4{,}082 \cdot x_2$,

$y_{y,x_3} = 64{,}37 + 0{,}361 \cdot x_3$.

10) *Определение средних частных коэффициентов эластичности* (п. 3.6, формула (3.9))

$$\bar{Y}_{y_{x_2}} = b_2 \frac{\bar{x}_2}{\bar{y}_{y,x_2}} = 4{,}082 \frac{37{,}70}{125{,}17} = 1{,}228;$$

$$\bar{Y}_{y_{x_3}} = b_3 \frac{\bar{x}_3}{\bar{y}_{y,x_3}} = 0{,}361 \frac{168{,}27}{125{,}17} = 0{,}485.$$

Результаты

1) Проверка факторов на наличие коллинеарности показала, что коллинеарность между факторами отсутствует.

2) Уравнение линейной множественной регрессии

$$y = -99{,}816 + 0{,}154 \cdot x_1 + 4{,}459 \cdot x_2 + 0{,}324 \cdot x_3.$$

3) Значения коэффициента множественной корреляции R и коэффициента детерминации R^2

$$R = 0{,}748; \qquad R^2 = 0{,}560.$$

4) Проверка значимости уравнения регрессии.

$$y = -99{,}816 + 0{,}154 \cdot x_1 + 4{,}459 \cdot x_2 + 0{,}324 \cdot x_3.$$

Построенное уравнение регрессии значимо при уровне значимости $\alpha = 0{,}05$.

5) Проверка значимости коэффициентов уравнения регрессии.

Оценки параметров a, b_2, b_3 значимы при уровне значимости $\alpha = 0{,}05$, а значение b_1 не значимо при уровне значимости $\alpha = 0{,}05$.

6) Построение уравнения линейной множественной регрессии с учетом только значимых факторов.

Уравнение регрессии имеет вид

$$y = -89{,}520 + 4{,}082 \cdot x_2 + 0{,}361 \cdot x_3.$$

7) Проверка гипотезы о гомоскедастичности ряда остатков с уровнем значимости $\alpha = 0{,}05$.

Дисперсии остатков постоянны (остатки гомоскедастичны) по переменным x_2 и x_3.

8) Частные уравнения регрессии

$$y_{y,x_2} = -28{,}77 + 4{,}082 \cdot x_2,$$

$$y_{y,x_3} = 64{,}37 + 0{,}361 \cdot x_3.$$

9) Средние частные коэффициенты эластичности

$$\bar{Y}_{y\,x_2} = 1{,}228; \qquad \bar{Y}_{y\,x_3} = 0{,}485.$$

4. Системы эконометрических уравнений

4.1. Структурная и приведенная формы модели

Экономические процессы и явления, как правило, представляют собой сложные системы, характеризующиеся большим количеством параметров и сложными взаимосвязями. При исследовании сложных экономических систем приходится учитывать, как правило, несколько взаимосвязей и, следовательно, вводить в модель несколько зависимых переменных, которые могут влиять друг на друга. В таких ситуациях эконометрические модели строятся в виде систем эконометрических уравнений.

Переменные, используемые в системах эконометрических уравнений, подразделяются на эндогенные и экзогенные.

Эндогенными переменными называются взаимозависимые переменные, которые определяются внутри модели (системы). Число эндогенных переменных, обозначаемых обычно буквой y, равно числу уравнений системы.

Экзогенными переменными называются переменные, которые определяются вне системы. Это независимые переменные, обозначаемые буквой x.

Предопределенными переменными называются экзогенные и лаговые эндогенные переменные системы (характеризуют систему в предыдущие моменты времени $y_{-1}, y_{-2},\ldots$).

В общем случае система эконометрических уравнений с n зависимыми переменными y_i имеет вид

$$\begin{aligned}
&y_1 = b_{12}\cdot y_2 + b_{13}\cdot y_3 + \ldots + b_{1n}\cdot y_n + a_{11}\cdot x_1 + \ldots + a_{1m}\cdot x_m + \varepsilon_1;\\
&y_2 = b_{21}\cdot y_1 + b_{23}\cdot y_3 + \ldots + b_{2n}\cdot y_n + a_{21}\cdot x_1 + \ldots + a_{2m}\cdot x_m + \varepsilon_2;\\
&\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\\
&y_n = b_{n1}\cdot y_1 + b_{n2}\cdot y_2 + \ldots + b_{n,n-1}\cdot y_{n-1} + a_{n1}\cdot x_1 + \ldots + a_{nm}\cdot x_m + \varepsilon_n.
\end{aligned} \tag{4.1}$$

Система (4.1) называется системой *взаимозависимых, одновременных* уравнений, а также *структурной формой модели*, так как она показывает взаимное влияние между всеми переменными модели.

Частными случаями системы (4.1) являются система *независимых* уравнений, в которой каждая зависимая переменная y_i является функцией только предопределенных переменных x_i

$$\begin{aligned}
&y_1 = a_{11}\cdot x_1 + \ldots + a_{1m}\cdot x_m + \varepsilon_1;\\
&y_2 = a_{21}\cdot x_1 + \ldots + a_{2m}\cdot x_m + \varepsilon_2;\\
&\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\\
&y_n = a_{n1}\cdot x_1 + \ldots + a_{nm}\cdot x_m + \varepsilon;
\end{aligned} \tag{4.2}$$

и система *рекурсивных* уравнений

$$
\begin{aligned}
&y_1 = a_{11} \cdot x_1 + \ldots + a_{1m} \cdot x_m + \varepsilon_1; \\
&y_2 = b_{21} \cdot y_1 + a_{21} \cdot x_1 + \ldots + a_{2m} \cdot x_m + \varepsilon_2; \\
&y_3 = b_{31} \cdot y_1 + b_{32} \cdot y_2 + a_{31} \cdot x_1 + \ldots + a_{3m} \cdot x_m + \varepsilon_3; \\
&\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\
&y_n = b_{n1} \cdot y_1 + b_{n2} \cdot y_2 + \ldots + b_{n,n-1} \cdot y_{n-1} + a_{n1} \cdot x_1 + \ldots + a_{nm} \cdot x_m + \varepsilon_n,
\end{aligned}
\tag{4.3}
$$

когда каждая зависимая переменная y_i является функцией только предопределенных переменных x_i и зависимых переменных y_i, определяемых в предыдущих уравнениях системы.

В системах независимых и рекурсивных уравнений отсутствует взаимное влияние зависимых переменных, предпосылки регрессионного анализа не нарушаются и поэтому для нахождения параметров a_{ij} и b_{ij}, называемых *структурными коэффициентами*, можно применять обычный МНК.

В моделях (4.1), (4.2), (4.3) отсутствуют свободные члены в каждом уравнении системы, так как предполагается, что значения переменных предварительно центрированы (выражены в отклонениях от среднего уровня).

Следует отметить, что структурная форма модели может включать не только уравнения, содержащие параметры (константы, подлежащие определению) и называемые поведенческими уравнениями, но и тождества, т. е. уравнения, не содержащие параметров и определяющие фиксированные отношения между переменными, например $Y_t = C_t + I_t + G_t$.

Наличие взаимозависимости между эндогенными переменными в системе одновременных уравнений (4.1) приводит к нарушению предпосылки о независимости объясняющих переменных и случайных членов, в результате чего обычный метод наименьших квадратов будет давать несостоятельные и смещенные оценки параметров.

Если с помощью преобразований исключить зависимые переменные из правых частей уравнений (4.1), то полученная система уравнений называется *приведенной формой модели* (ПФМ)

$$
\begin{aligned}
&y_1 = \delta_{11} \cdot x_1 + \delta_{12} \cdot x_2 + \ldots + \delta_{1m} \cdot x_m; \\
&y_2 = \delta_{21} \cdot x_1 + \delta_{22} \cdot x_2 + \ldots + \delta_{2m} \cdot x_m; \\
&\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\
&y_n = \delta_{n1} \cdot x_1 + \delta_{n2} \cdot x_2 + \ldots + \delta_{nm} \cdot x_m.
\end{aligned}
\tag{4.4}
$$

Параметры δ_{ij} приведенной формой модели являются алгебраическими функциями от структурных коэффициентов a_{ij} и b_{ij} и называются *приведенными коэффициентами*.

Например, для конъюнктурной модели, определяемой соотношениями:

$$
\begin{aligned}
&C_t = a_1 + b_{11} \cdot Y_t + b_{12} \cdot C_{t-1} + u_1 && \text{(функция потребления);} \\
&I_t = a_2 + b_{21} \cdot r_t + b_{22} \cdot I_{t-1} + u_2 && \text{(функция инвестиций);} \\
&r_t = a_3 + b_{31} \cdot Y_t + b_{32} \cdot M_t + u_3 && \text{(функция денежного рынка);} \\
&Y_t = C_t + I_t + G_t && \text{(тождество дохода),}
\end{aligned}
\tag{4.5}
$$

где C – расходы на потребление, Y – ВВП, I – инвестиции, r – процентная ставка, M – денежная масса, G – государственные расходы, t и t–1 обозначают текущий и предыдущий периоды, u_1, u_2, u_3 – случайные ошибки, эндогенными переменными являются переменные C_t, I_t, r_t, Y_t, а экзогенными (предопределенными) переменными – $M_t, G_t, C_{t-1}, I_{t-1}$, поэтому приведенная форма модели будет иметь следующий вид:

$$
\begin{aligned}
C_t &= \delta_{11} \cdot M_t + \delta_{12} \cdot G_t + \delta_{13} \cdot C_{t-1} + \delta_{14} \cdot I_{t-1} + \nu_1, \\
I_t &= \delta_{21} \cdot M_t + \delta_{22} \cdot G_t + \delta_{23} \cdot C_{t-1} + \delta_{24} \cdot I_{t-1} + \nu_2, \\
r_t &= \delta_{31} \cdot M_t + \delta_{32} \cdot G_t + \delta_{33} \cdot C_{t-1} + \delta_{34} \cdot I_{t-1} + \nu_3, \\
Y_t &= \delta_{41} \cdot M_t + \delta_{42} \cdot G_t + \delta_{43} \cdot C_{t-1} + \delta_{44} \cdot I_{t-1} + \nu_4.
\end{aligned}
\tag{4.6}
$$

По своей структуре приведенная форма модели представляет собой систему независимых уравнений, поэтому ее параметры δ_{ij} можно оценивать с помощью обычного метода наименьших квадратов. Полученные численные значения параметров δ_{ij} позволяют вычислять модельные значения эндогенных переменных через предопределенные переменные. На этом процесс построения модели не заканчивается, так как для исследователя наибольший интерес представляют значения именно структурных коэффициентов a_{ij} и b_{ij}, характеризующих внутренние взаимосвязи в системе и допускающих экономическую интерпретацию и позволяющих осуществлять экономическое прогнозирование.

4.2. Оценка параметров структурной формы модели

Получение оценок параметров приведенной формы модели, как уже отмечалось, затруднений не представляет. Следующим этапом должно быть определение оценок параметров структурной формы модели по оценкам приведенной формы модели, что в ряде случаев можно осуществить с помощью обратного преобразования. Здесь возникает проблема идентификации, заключающаяся в том, что не всегда возможно по приведенным коэффициентам модели однозначно определить ее структурные коэффициенты, так как в общем случае структурная и приведенная формы модели содержат разное число параметров.

С позиции идентифицируемости можно выделить три вида структурных моделей:

– *идентифицируемые* системы, в которых число параметров структурной и приведенной форм модели совпадает, и структурные коэффициенты модели однозначно оцениваются через параметры приведенной формы модели;

– *неидентифицируемые* системы, в которых число структурных параметров превышает число приведенных, и структурные коэффициенты не могут быть получены из коэффициентов приведенной формы модели;

– *сверхидентифицируемые* системы с числом приведенных параметров, превышающих число структурных. В этом случае возможно неоднозначное определение значений структурных коэффициентов при полученных значениях приведенных коэффициентов.

Наиболее часто для оценки параметров системы одновременных уравнений применяются косвенный и двухшаговый методы наименьших квадратов (КМНК, ДМНК).

Косвенный МНК используется в случае идентифицируемой системы уравнений и заключается в получении уравнений структурной формы модели с помощью алгебраического преобразования уравнений приведенной формы.

Двухшаговый МНК является универсальным методом. Он основан на использовании, так называемых, «инструментальных» переменных, которые заменяют эндогенные переменные в правых частях уравнений структурной формы модели.

Инструментальными называются такие переменные $\hat{y}_i$, которые близки к исходным эндогенным переменным y и при этом не коррелируют с ошибками уравнений модели. В качестве таких переменных предлагается использовать переменные, определяемые уравнениями приведенной формы модели (4.4).

Алгоритм двухшагового МНК:

1) Исходная система уравнений (4.1) преобразуется в приведенную форму модели (4.4) и определяются численные значения параметров δ_{ij} для каждого ее уравнения в отдельности с помощью традиционного МНК;

2) Полученные уравнения приведенной формы используются для введения инструментальных переменных $\hat{y}_i$, соответствующих исходным переменным y_i

$$\hat{y}_1 = \delta_{11} \cdot x_1 + \delta_{12} \cdot x_2 + ... + \delta_{1m} \cdot x_m;$$
$$\hat{y}_2 = \delta_{21} \cdot x_1 + \delta_{22} \cdot x_2 + ... + \delta_{2m} \cdot x_m;$$
$$\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots$$
$$\hat{y}_n = \delta_{n1} \cdot x_1 + \delta_{n2} \cdot x_2 + ... + \delta_{nm} \cdot x_m.$$

Находятся расчетные значения инструментальных переменных $\hat{y}_i$ для каждого наблюдения;

3) С помощью обычного МНК определяются параметры каждого структурного уравнения в отдельности, используя в качестве значений факторов фактические значения предопределенных переменных x_i и полученные расчетные значения инструментальных переменных $\hat{y}_i$.

Контрольные вопросы

1. Какие виды систем уравнений применяются в эконометрике?
2. Что представляют собой структурная и приведенная форма модели?
3. Какие переменные называются эндогенными, экзогенными, предопределенными?
4. В чем заключается проблема идентификации модели?
5. Какие методы применяются для нахождения структурных коэффициентов модели для различных видов систем уравнений?
6. В каких случаях применяется косвенный МНК?
7. В каких случаях применяется двухшаговый МНК?
8. Опишите алгоритм применения двухшагового МНК.

Задачи

1. В системе уравнений

$$C_t = a + b_1 \cdot Y_t + b_2 \cdot Y_{t-1};$$
$$Y_t = C_t + I_t + G_t,$$

где a, b_1, b_2 – константы, выделить эндогенные и предопределенные переменные.

(Эндогенные: C_t, Y_t; предопределенные: I_t, G_t, Y_{t-1}).

2. Определить, к какому типу относится следующая система уравнений

$$C_t = a_1 + b_{11} \cdot Y_t;$$
$$I_t = a_2 + b_{21} \cdot Y_t;$$
$$Y_t = C_t + I_t + G_t,$$

где a_1, a_2, b_1, b_{21} – константы.

(Система одновременных уравнений).

3. Определить, к какому типу относится следующая система уравнений

$$C_t = a_1 + b_{11} \cdot Y_t;$$
$$I_t = a_2 + b_{21} \cdot Y_t,$$

где a_1, a_2, b_1, b_{21} – константы.

(Система независимых уравнений).

4. Для макроэкономической модели

$$C_t = b_{11} \cdot Y_t + \varepsilon_1;$$
$$I_t = b_{21} \cdot Y_t + \varepsilon_2;$$
$$Y_t = C_t + I_t + G_t,$$

где Y – валовой национальный доход; C – личное потребление; I – инвестиции; G – государственные расходы; t и t–1 обозначают текущий и предыдущий периоды; ε_1 и ε_2 – случайные ошибки, построить систему уравнений приведенной формы модели.

$$\begin{pmatrix} C_t = \delta_{11} \cdot G_t \\ I_t = \delta_{21} \cdot G_t \\ Y_t = \delta_{31} \cdot G_t \end{pmatrix}.$$

5. Проверить следующую систему уравнений на идентифицируемость

$$C_t = b_{11} \cdot Y_t + \varepsilon_1;$$
$$I_t = b_{21} \cdot Y_t + \varepsilon_2;$$
$$Y_t = C_t + I_t + G_t,$$

(Система сверхидентифицируема).

Лабораторная работа №4. Системы эконометрических уравнений: построение модели в виде системы взаимосвязанных эконометрических уравнений

Задание. По заданным исходным данным для заданной модели (в соответствии с вариантом):

1. Выделить эндогенные и экзогенные переменные.
2. Записать приведенную форму модели.
3. Определить коэффициенты приведенной формы модели.
4. Вычислить значения инструментальных переменных.
5. Определить коэффициенты структурной формы модели двухшаговым методом наименьших квадратов.
6. Проверить значимость полученных уравнений и их коэффициентов.

Указания к решению. Для нахождения приведенных уравнений (а также коэффициентов структурных уравнений при применении ДМНК) рекомендуется использовать функцию «Сервис.Анализ данных.Регрессия» табличного процессора MS Excel. (рис 3.2) либо программный пакет Matrixer 5.1.

Варианты заданий к лабораторным работам №4

Ниже приведены варианты систем уравнений, исходные данные к которым берутся из таблицы П4.3 приложения.

Вариант 1

$$C_t = a_1 + b_{11} \cdot Y_t + b_{12} \cdot Y_{t-1} + \varepsilon_{1t}, \quad \text{(функция потребления)}$$
$$I_t = a_2 + b_{21} \cdot Y_t + \varepsilon_{2t}, \quad \text{(функция инвестиций)}$$
$$Y_t = C_t + I_t + G_t. \quad \text{(тождество дохода)}$$

где C_t – потребление;
Y_t – ВВП;
I_t – валовые инвестиции;
G_t – государственные расходы;
t, t–1 – текущий и предыдущий периоды;
ε_1 и ε_2 – случайные ошибки.

Вариант 2

$$C_t = a_1 + b_{11} \cdot Y_t + b_{12} \cdot Y_{t-1} + \varepsilon_1;$$
$$I_t = a_2 + b_{21} \cdot Y_t + b_{22} \cdot Y_{t-1} + \varepsilon_2;$$
$$Y_t = C_t + I_t + G_t,$$

где C – расходы на потребление;
Y – ВВП;
I – инвестиции;
G – государственные расходы;
t, t–1 – текущий и предыдущий периоды;
ε_1 и ε_2 – случайные ошибки.

Вариант 3

$$C_t = a_1 + b_{11} \cdot Y_t + \varepsilon_1;$$
$$I_t = a_2 + b_{21} \cdot Y_t + b_{22} \cdot Y_{t-1} + \varepsilon_2;$$
$$Y_t = C_t + I_t + G_t,$$

где Y – ВВП; C – личное потребление; I – инвестиции; G – государственные расходы; t и t–1 обозначают текущий и предыдущий периоды; ε_1 и ε_2 – случайные ошибки.

Вариант 4

$$C_t = a_1 + b_{11} \cdot Y_t + b_{12} \cdot Y_{t-1} + \varepsilon_{1t},$$
$$I_t = a_2 + b_{21} \cdot Y_t + \varepsilon_{2t},$$
$$Y_t = C_t + I_t + G_t.$$

где C_t – потребление;
Y_t – валовой национальный доход;
I_t – валовые инвестиции;
G_t – государственные расходы;
t, t–1 – текущий и предыдущий периоды;
ε_1 и ε_2 – случайные ошибки.

Вариант 5

$$C_t = a_1 + b_{11} \cdot Y_t + b_{12} \cdot Y_{t-1} + \varepsilon_1;$$
$$I_t = a_2 + b_{21} \cdot Y_t + b_{22} \cdot Y_{t-1} + \varepsilon_2;$$
$$Y_t = C_t + I_t + G_t,$$

где C – расходы на потребление;
Y – валовой национальный доход;
I – инвестиции;
G – государственные расходы;
t, t–1 – текущий и предыдущий периоды;
ε_1 и ε_2 – случайные ошибки.

Вариант 6

$$C_t = a_1 + b_{11} \cdot Y_t + \varepsilon_1;$$
$$I_t = a_2 + b_{21} \cdot Y_t + b_{22} \cdot Y_{t-1} + \varepsilon_2;$$
$$Y_t = C_t + I_t + G_t,$$

где Y – валовой национальный доход;
C – личное потребление;
I – инвестиции;
G – государственные расходы;
t и t–1 обозначают текущий и предыдущий периоды;
ε_1 и ε_2 – случайные ошибки.

Вариант 7

Модель Менгеса

$$Y_t = a_1 + b_{11} \cdot Y_{t-1} + b_{12} \cdot I_t + \varepsilon_1;$$

$$I_t = a_2 + b_{21} \cdot Y_t + b_{22} \cdot Q_t + \varepsilon_2;$$

$$C_t = a_3 + b_{31} \cdot Y_t + b_{32} \cdot C_{t-1} + b_{33} \cdot P_t + \varepsilon_3;$$

$$Q_t = a_4 + b_{41} \cdot Q_{t-1} + b_{42} \cdot R_t + \varepsilon_4,$$

где Y – национальный доход;
C – расходы на личное потребление;
I – чистые инвестиции;
Q – валовая прибыль экономики;
P – индекс стоимости жизни;
R – объем продукции промышленности;
t, t–1 – текущий и предыдущий периоды;
ε_1 и ε_2 – случайные ошибки.

Вариант 8

$$C_t = a_1 + b_{11} \cdot R_t + b_{12} \cdot C_{t-1} + \varepsilon_1;$$

$$I_t = a_2 + b_{21} \cdot R_t + b_{22} \cdot I_{t-1} + \varepsilon_2;$$

$$R_t = C_t + I_t,$$

где C – расходы на потребление;
R – национальный доход;
I – инвестиции;
t, t–1 – текущий и предыдущий периоды;
ε_1 и ε_2 – случайные ошибки.

Вариант 9

$$C_t = a_1 + b_{11} \cdot Y_t + b_{13} \cdot T_t + \varepsilon_1;$$

$$I_t = a_2 + b_{21} \cdot Y_t + b_{24} \cdot K_{t-1} + \varepsilon_2;$$

$$Y_t = C_t + I_t,$$

где C – потребление;
I – инвестиции;
Y – национальный доход;
T – налоги;
K – запас капитала;
t, t–1 – текущий и предыдущий периоды;
ε_1 и ε_2 – случайные ошибки.

Вариант 10

$$C_t = a_1 + b_{11} \cdot Y_t + \varepsilon_1;$$

$$Y_t = C_t + I_t,$$

где C – потребление;
Y – ВВП;
I – валовые инвестиции;
t, t–1 – текущий и предыдущий периоды;
ε_1 – случайная ошибка.

Вариант 11

$$C_t = a_1 + b_{11} \cdot Y_t + b_{12} \cdot Y_{t-1} + \varepsilon_1;$$

$$Y_t = C_t + I_t + G_t,$$

где C – потребление;
Y – ВВП;
I – валовые инвестиции;
G – государственные расходы;
t, t–1 – текущий и предыдущий периоды;
ε_1 – случайная ошибка.

Вариант 12

$$C_t = a_1 + b_{11} \cdot D_t + \varepsilon_1;$$

$$I_t = a_2 + b_{22} \cdot Y_t + b_{23} \cdot Y_{t-1} + \varepsilon_2;$$

$$Y_t = D_t + T_t;$$

$$D_t = C_t + I_t + G_t,$$

где C – расходы на потребление;
Y – чистый национальный продукт;
D – чистый национальный доход;
I – инвестиции;
T –налоги;
G – государственные расходы;
t, t–1 – текущий и предыдущий периоды;
ε_1 и ε_2 – случайные ошибки.

Вариант 13

$$C_t = b_1 + b_2 \cdot S_t + b_3 \cdot P_t + \varepsilon_1;$$

$$S_t = a_1 + a_2 \cdot R_t + a_3 \cdot R_{t-1} + a_4 \cdot t + \varepsilon_2;$$

$$R_t = S_t + P_t,$$

где C_t – личное потребление;
S_t – зарплата;
P_t – прибыль;
R_t – национальный доход;
t, t–1 – текущий и предыдущий периоды;
ε_1 и ε_2 – случайные ошибки.

Вариант 14

$$C_t = b_1 + b_2 \cdot S_t + b_3 \cdot P_{t-1} + \varepsilon_1;$$

$$S_t = a_1 + a_2 \cdot R_t + a_3 \cdot R_{t-1} + a_4 \cdot t + \varepsilon_2;$$

$$R_t = S_t + P_t,$$

где C_t – личное потребление;
S_t – зарплата;
P_t – прибыль;
R_t – национальный доход;
t, t–1 – текущий и предыдущий периоды;

ε_1 и ε_2 – случайные ошибки.

Вариант 15

$$C_t = b_1 + b_2 \cdot S_t + b_3 \cdot P_t + \varepsilon_1;$$

$$S_t = a_1 + a_2 \cdot R_t + a_3 \cdot S_{t-1} + a_4 \cdot t + \varepsilon_2;$$

$$R_t = S_t + P_t,$$

где C_t – личное потребление;

S_t – зарплата;

P_t – прибыль;

R_t – национальный доход;

t, t–1 – текущий и предыдущий периоды;

ε_1 и ε_2 – случайные ошибки.

Вариант 16

$$C_t = b_1 + b_2 \cdot S_t + b_3 \cdot P_t + \varepsilon_1;$$

$$S_t = a_1 + a_2 \cdot R_t + a_3 \cdot R_{t-1} + \varepsilon_2;$$

$$R_t = S_t + P_t,$$

где C_t – личное потребление;

S_t – зарплата;

P_t – прибыль;

R_t – национальный доход;

t, t–1 – текущий и предыдущий периоды;

ε_1 и ε_2 – случайные ошибки.

Вариант 17

$$C_t = a_1 + b_{11} \cdot Y_t + b_{12} \cdot I_t + \varepsilon_1;$$

$$I_t = a_2 + b_{21} \cdot Y_{t-1} + \varepsilon_2;$$

$$T_t = a_3 + b_{31} \cdot Y_t + \varepsilon_3;$$

$$Y_t = C_t + I_t + G_t,$$

где C – совокупное потребление;

Y – совокупный доход;

I – инвестиции;

T – налоги;

G – государственные расходы в период t.

t, t–1 – текущий и предыдущий периоды;

ε_1 и ε_2 – случайные ошибки.

Вариант 18

$$C_t = a + b \cdot Y_t + \varepsilon;$$

$$Y_t = C_t + I_t + G_t,$$

где C – расходы на потребление;

Y – доход;

I – инвестиции;

G – государственные расходы;

t – текущий период.

ε_1 – случайная ошибка.

Вариант 19

$$C_t = a + b_1 \cdot Y_t + b_2 \cdot Y_{t-1} + \varepsilon;$$

$$Y_t = C_t + I_t + G_t,$$

где C – расходы на потребление;

Y – доход;

I – инвестиции;

G – государственные расходы;

t, t–1 – текущий и предыдущий периоды;

ε_1 – случайная ошибка.

Вариант 20

$$C_t = a_1 + b_{11} \cdot Y_t + b_{12} \cdot I_{t-1} + \varepsilon_1;$$

$$I_t = a_2 + b_{21} \cdot Y_{t-1} + \varepsilon_2;$$

$$T_t = a_3 + b_{31} \cdot Y_t + \varepsilon_3;$$

$$Y_t = C_t + I_t + G_t,$$

где C – совокупное потребление;

Y – совокупный доход;

I – инвестиции;

T – налоги;

G – государственные расходы в период t.

t, t–1 – текущий и предыдущий периоды;

ε_1 и ε_2 – случайные ошибки.

Пример выполнения лабораторной работы №4

Исходные данные:

– уровень значимости α = 0,05;

– система уравнений представляет собой модифицированную модель Кейнса

$$\begin{aligned} C_t &= a_1 + b_{11} \cdot Y_t + \varepsilon_1; \\ I_t &= a_2 + b_{21} \cdot Y_t + b_{22} \cdot Y_{t-1} + \varepsilon_2; \\ Y_t &= C_t + I_t + G_t, \end{aligned} \qquad (4.7)$$

где Y – валовой национальный доход; C – личное потребление; I – инвестиции; G – государственные расходы; t и t–1 обозначают текущий и предыдущий периоды; ε_1 и ε_2 – случайные ошибки.

Таблица 4.1

Данные наблюдений для макроэкономической модели Кейнса

Год наблюдения	C_t	I_t	Y_t	$Y_{t\text{-}1}$	G_t	Расчетные значения $\hat{Y}_t$
1	1016,6	267,0	1412,7	–	486,1	–
2	1435,9	376,0	1978,9	1412,7	652,7	2243,7
3	1776,1	408,8	2292,0	1978,9	839,0	2899,5
4	2003,8	407,1	2514,4	2292,0	842,1	3158,6
5	3265,7	670,4	4632,0	2514,4	1258,0	3771,6
6	4476,9	1165,2	7116,6	4632,0	1960,1	6230,0
7	5886,9	1504,7	8819,9	7116,6	2419,4	8736,4

Окончание табл. 4.1

1	2	3	4	5	6	7
8	7443,2	1762,4	10627,5	8819,9	3422,3	11168,2
9	9024,8	2186,4	12886,1	10627,5	3964,9	13207,8
10	11401,4	2865,0	16679,9	12886,1	4669,7	15784,2
11	14363,5	3611,1	21079,5	16679,9	6820,6	21114,7
12	17742,6	4580,5	26009,7	21079,5	8375,2	26321,7

1) Выделение эндогенных и предопределеных переменных.

Эндогенные переменные: Y_t, C_t, I_t

Предопределенные переменные Y_{t-1} и G_t.

2) Приведенная форма модели имеет вид;

$$\begin{aligned} C_t &= \lambda_1 + \delta_{11} \cdot Y_{t-1} + \delta_{12} \cdot G_t; \\ I_t &= \lambda_2 + \delta_{21} \cdot Y_{t-1} + \delta_{22} \cdot G_t; \\ Y_t &= \lambda_3 + \delta_{31} \cdot Y_{t-1} + \delta_{32} \cdot G_t. \end{aligned} \tag{4.8}$$

3) Определение коэффициентов приведенной формы модели.

Для построения определения параметров 1-го уравнения системы (4.8) используем функцию «Сервис.Анализ данных.Регрессия» табличного процессора MS Excel (рис 3.2).

Задав соответствующие диапазоны данных в окне определения параметров регрессии, получим следующие результаты (табл. 4.2).

Таблица 4.2

Результаты регрессионного анализа

Показатели	Коэффициенты уравнения регрессии	Стандартная ошибка определения коэффициентов	t-статистика	Вероятность ошибки α	Нижние 95%-пределы	Верхние 95%-пределы
Y-пересечение	377,52	179,041	2,109	0,068	–35,353	790,388
Переменная Y_{t-1}	0,582	0,195	2,987	0,017	0,133	1,031
Переменная G_t	0,633	0,497	1,272	0,239	–0,514	1,780

Из таблицы следует, что уравнение регрессии имеет вид

$$C_t = 377{,}52 + 0{,}582 \cdot Y_{t-1} + 0{,}633 \cdot G_t. \tag{4.9}$$

Аналогично получим значения коэффициентов следующих двух уравнений системы (4.8)

$$I_t = 19{,}26 + 0{,}154 \cdot Y_{t-1} + 0{,}155 \cdot G_t. \tag{4.10}$$

$$Y_t = 412{,}51 + 0{,}817 \cdot Y_{t-1} + 1{,}037 \cdot G_t. \tag{4.11}$$

4) Вычисление значений инструментальных переменных.

В правую часть уравнений системы входит только переменная Y_t, поэтому достаточно вычислить только значения инструментальной переменной $\hat{Y}_t$ по уравнению (4.11). Результаты расчетов $\hat{Y}_t$ показаны в последнем столбце таблицы 4.1.

5) Определение коэффициентов структурной формы модели.

Для построения определения параметров 1-го уравнения системы (4.8) используем функцию «Сервис.Анализ данных.Регрессия» табличного процессора MS Excel (рис 3.2).

Задав соответствующие диапазоны данных в окне определения параметров регрессии для 1-го уравнения системы (4,7), в котором переменная Y_t заменена на инструментальную переменную $\hat{Y}_t$,

$$C_t = a_1 + b_{11} \cdot \hat{Y}_t + \varepsilon_1 \qquad (4.12)$$

получим:

множественный коэффициент корреляции R = 0,9982,

коэффициент детерминации R^2 = 0,9965,

$F_{факт}$ = 2557,

уровень значимости уравнения регрессии $\alpha = 2{,}32 \cdot 10^{-12}$.

Таблица 4.3

Результаты регрессионного анализа

Показатели	Коэффициенты уравнения регрессии	Стандартная ошибка определения коэффициентов	t-статистика	Вероятность ошибки α	Нижние 95%-пределы	Верхние 95%-пределы
Y-пересечение	97,653	173,441	0,563	0,587	–294,697	490,003
Переменная $\hat{Y}_t$	0,678	0,013	50,570	0,000	0,648	0,709

Из таблицы 4.3 следует, что уравнение регрессии имеет вид

$$C_t = 97{,}653 + 0{,}678 \cdot Y_t. \qquad (4.13)$$

Аналогично для 2-го 1-го уравнения системы (4.8)

$$I_t = a_2 + b_{21} \cdot \hat{Y}_t + b_{22} \cdot Y_{t-1} + \varepsilon_2 \qquad (4.14)$$

получим:

множественный коэффициент корреляции R = 0,9959,

коэффициент детерминации R^2 = 0,9979,

$F_{факт}$ = 960,2,

уровень значимости уравнения регрессии $\alpha = 2{,}96 \cdot 10^{-10}$.

Таблица 4.4

Результаты регрессионного анализа

Показатели	Коэффициенты уравнения регрессии	Стандартная ошибка определения коэффициентов	t-статистика	Вероятность ошибки α	Нижние 95%-пределы	Верхние 95%-пределы
Y-пересечение	–42,481	76,234	–0,557	0,593	–218,277	133,315
Переменная $\hat{Y}_t$	0,150	0,135	1,107	0,300	–0,162	0,461
Переменная Y_{t-1}	0,032	0,165	0,191	0,853	–0,349	0,413

Из таблицы 4.4 следует, что уравнение регрессии имеет вид

$$I_t = -42{,}48 + 0{,}150 \cdot Y_t + 0{,}032 \cdot Y_{t-1}. \qquad (4.15)$$

6) Проверка значимости полученных уравнений и их коэффициентов.

Уравнение (4.13) значимо при $\alpha = 0{,}05$, так как его значимость $\alpha = 2{,}32 \cdot 10^{-12}$.

Из таблицы 4.3 следуют следующие значения уровней значимости значений параметров уравнения (4.13):

- параметр 97,653: $\alpha = 0,\ 587$;
- параметр 0,678: $\alpha = 2{,}3210^{-12}$.

Следовательно, при уровне значимости α = 0,05 параметр 97,653 – не значим, а параметр 0,678 – значим.

Уравнение (4.14) значимо при $\alpha = 0{,}05$, так как его значимость $\alpha = 2{,}96 \cdot 10^{-10}$.

Из таблицы 4.4 следуют следующие значения уровней значимости значений параметров уравнения (4.13):

- параметр -42,481: $\alpha = 0{,}593$;
- параметр 0,150: $\alpha = 0{,}300$.
- параметр 0,032: $\alpha = 0{,}853$.

Следовательно при уровне значимости $\alpha = 0{,}05$ все параметры не значимы.

Результаты

1) Эндогенные переменные: Y_t, C_t, I_t

Предопределенные переменные Y_{t-1} и G_t.

2) Приведенная форма модели имеет вид;

$$C_t = \lambda_1 + \delta_{11} \cdot Y_{t-1} + \delta_{12} \cdot G_t;$$

$$I_t = \lambda_2 + \delta_{21} \cdot Y_{t-1} + \delta_{22} \cdot G_t;$$

$$Y_t = \lambda_3 + \delta_{31} \cdot Y_{t-1} + \delta_{32} \cdot G_t.$$

3) Коэффициенты приведенной формы модели.

$$C_t = 377{,}5 + 0{,}582 \cdot Y_{t-1} + 0{,}632 \cdot G_t;$$

$$I_t = 19{,}3 + 0{,}154 \cdot Y_{t-1} 0{,}155 \cdot G_t;$$

$$Y_t = 412{,}5 + 0{,}817 \cdot Y_{t-1} + 1{,}037 \cdot G_t.$$

4) Значения инструментальных переменных.

Результаты расчетов инструментальной переменной $\hat{Y}_t$ показаны в последнем столбце таблицы 4.1.

5) Коэффициенты структурной формы модели.

$$C_t = 97{,}66 + 0{,}678 \cdot Y_t;$$

$$I_t = -42{,}47 + 0{,}150 \cdot Y_t + 0{,}031 \cdot Y_{t-1}.$$

Из уравнений (4.22) следует, что 67,8% прироста национального дохода идет на увеличение потребления. На увеличение инвестиций направляется соответственно 15% и 3,1% прироста национального дохода текущего и предыдущего года.

6) Проверка значимости полученных уравнений и их коэффициентов.

Первое уравнение системы

$$C_t = 97{,}653 + 0{,}678 \cdot Y_t$$

значимо при $\alpha = 0{,}05$.

При уровне значимости $\alpha = 0{,}05$ параметр 97,653 – не значим, а параметр 0,678 – значим.

Второе уравнение системы

$$I_t = -42{,}48 + 0{,}150 \cdot Y_t + 0{,}032 \cdot Y_{t-1}$$

значимо при $\alpha = 0{,}05$.

При уровне значимости $\alpha = 0{,}05$ все параметры не значимы.

5. Динамические эконометрические модели

5.1. Общая характеристика динамических моделей

Состояние экономического явления в данный момент или период времени часто зависит от его состояний либо состояний окружающей среды в предшествующие моменты или периоды времени. Данное обстоятельство является следствием наличия запаздывания в действии факторов либо инерционностью изучаемых процессов.

Модели, связывающие состояния экономических явлений в последовательные моменты (периоды) времени, принято называть *динамическими*. Такие модели позволяют изучать явления в динамике, в развитии.

Аналитическое представление динамических моделей включает значения переменных, относящиеся как к текущему, так и к предыдущим моментам (периодам) времени.

Эконометрические модели, включающие в качестве факторов значения факторных переменных в предыдущие моменты времени, называются *моделями с распределенным лагом*.

$$y_t = a + b_0 \cdot x_t + b_1 \cdot x_{t-1} + b_2 \cdot x_{t-2} + ... + b_p \cdot x_{t-p} + \varepsilon_t . \qquad (5.1)$$

Моделями этого типа описываются ситуации, когда влияние причины (независимых факторов) на следствие (зависимую переменную) проявляется с некоторым запаздыванием. Например, при изучении зависимости объемов выпуска продукции от величины инвестиций, выручки от расходов на рекламу и т. п.

Эконометрические модели, включающие в качестве факторов значения результативной переменной в предыдущие моменты времени. Эти модели называются *моделями авторегрессии*.

$$y_t = a + b_0 \cdot x_t + c_1 \cdot y_{t-1} + c_2 \cdot y_{t-2} + ... + c_q \cdot y_{t-q} + \varepsilon_t . \qquad (5.2)$$

Модели такого типа предполагают наличие определенной инерционности в изменении рассматриваемого явления, когда уровень изучаемого явления существенно зависит от его уровней, достигнутых в предыдущих периодах. Например, уровень спроса на товар либо уровень ВВП в данном периоде во многом определяется уровнями, достигнутыми в предшествующем периоде.

Применяются и различные комбинации упомянутых выше моделей.

Включение в эконометрическую модель лаговых переменных вызывает следующие проблемы.

Во-первых, наличие нескольких лаговых переменных y_{t-1}, y_{t-2}, ... либо x_{t-1}, x_{t-2}, ... , зачастую сильно коррелирующих между собой, ведет к потере качества модели вследствие ухудшения точности оценок ее параметров, снижению их эффективности и устойчивости к незначительным колебаниям исходной информации, ошибкам округления.

Во-вторых, как правило, существует сильная корреляционная зависимость между переменными y_{t-1}, y_{t-2}, ... и ошибкой ε_t, ведущая к появлению смещения в оценках параметров при использовании МНК.

В-третьих, временной ряд ошибки модели ε_t часто характеризуется наличием автокорреляционной связи, вследствие чего оценки параметров модели, полученные непосредственно на основе МНК, являются неэффективными.

Отметим, что важным этапом при построении моделей с распределенным лагом и моделей авторегрессии является выбор оптимальной величины лага и определение его структуры.

5.2. Интерпретация параметров динамических моделей

5.2.1. Интерпретация параметров моделей с распределенным лагом

Рассмотрим модель с распределенным лагом

$$y_t = a + b_0 \cdot x_t + b_1 \cdot x_{t-1} + b_2 \cdot x_{t-2} + ... + b_p \cdot x_{t-p} + \varepsilon_t. \qquad (5.3)$$

Из соотношения (5.3) следует, что изменение независимой переменной x в каком-либо периоде времени t влияет на значение переменной y в данном периоде и в течение p следующих периодов времени. В последующие периоды это влияние проявляться уже не будет. Таким образом, временной интервал влияния конечен и ограничен p+1 периодом.

Коэффициент регрессии b_0 при переменной x_t называют *краткосрочным мультипликатором*. Он характеризует среднее абсолютное изменение y_t при изменении x_t на одну единицу своего измерения в некотором периоде времени t, без учета воздействия лаговых значений фактора x.

Величины (b_0 + b_1), (b_0 + b_1 + b_2) и т. д. называются *промежуточными мультипликаторами*. Они характеризует изменение y_t в течение двух, трех и т. д. периодов после изменения x_t на одну единицу.

Величина

$$b = b_0 + b_1 +...+ b_l \qquad (5.4)$$

показывает максимальное суммарное изменение результирующей переменной y, которое будет достигнуто (по окончании текущего и p следующих периодов) под влиянием изменения фактора x на единицу в текущем периоде, и называется *долгосрочным мультипликатором*.

Например, для модели

$$y_{\text{t}} = 100 + 70x_t + 25x_{t-1} + 5x_{t-2}$$

краткосрочный мультипликатор равен 70, т. е. увеличение x_t на 1 единицу ведет в среднем к росту показателя y_t на 70 единиц в том же периоде. В течение двух периодов показатель y_t возрастет на 70 + 25 = 95 единиц, а долгосрочный мультипликатор равен

$$b= (b_0 + b_1 + b_2) = 70+25+5 =100,$$

и, следовательно, суммарное изменение показателя y_t составит 100 единиц.

5.2.2. Интерпретация параметров моделей авторегрессии

Рассмотрим модель авторегрессии первого порядка

$$y_t = a + b_0 \cdot x_t + c_1 \cdot y_{t-1} + \varepsilon_t. \qquad (5.5)$$

Коэффициент b_0, как и ранее, характеризует краткосрочное изменение y_t под воздействием изменения x_t на единицу в том же периоде. Изменение y_t на b_0 в данном периоде в силу соотношения (7.13) повлечет в следующем периоде изменение y_{t+1} на величину $b_0 \cdot c_1$. В периоде t + 2 изменение y_{t+2} составит $b_0 \cdot c_1^2$ и т. д. Долгосрочный мультипликатор в модели авторегрессии рассчитывается как бесконечная сумма

$$b = b_0 + b_0 c_1 + b_0 c_1^2 + b_0 c_1^3 + \ldots \quad (5.6)$$

Если выполняется условие $|c_1| < 1$, то сумма в правой части (5.6), т. е. величина долгосрочного мультипликатора, будет конечная

$$b = b_0 \cdot (1 + c_1 + c_1^2 + c_1^3 \ldots) = \frac{b_0}{1 - c_1}, \text{ где } |c_1| < 1. \quad (5.7)$$

В модельном примере

$$y_t = 200 + 50x_t + 0{,}6\, y_{t-1},$$

краткосрочный мультипликатор равен 50, следовательно, увеличение x_t на 1 единицу приводит к росту y_t в том же периоде в среднем на 50 единиц. Долгосрочное изменение y_t составит $b = 50 /(1–0{,}6) = 125$ единиц, т. е. изменение x_t на 1 единицу в каком-либо периоде приведет к изменению y_t в долгосрочной перспективе в среднем на 125 единиц.

5.3. Оценка параметров моделей авторегрессии

Рассмотрим модель авторегрессии первого порядка

$$y_t = a + b_0 \cdot x_t + c_1 \cdot y_{t-1} + \varepsilon_t. \quad (5.8)$$

Одна из основных проблем при построении моделей авторегрессии (при оценке параметров) связана с наличием корреляционной зависимости между переменной y_{t-1} и остатками ε_t в уравнении регрессии, что приводит при применении обычного МНК к получению смещенной оценки параметра при переменной y_{t-1}.

Для преодоления этой проблемы обычно используется метод *инструментальных переменных*, согласно которому переменная y_{t-1} из правой части модели заменяется на новую переменную $\hat{y}_{t-1}$, которая, во-первых, должна тесно коррелировать с y_{t-1}, и, во-вторых, не коррелировать с ошибкой модели ε_t.

В качестве такой переменной можно взять регрессию переменной y_{t-1} на переменную x_{t-1}, определяемую соотношением

$$\hat{y}_{t-1} = d_0 + d_1 \cdot x_{t-1}, \quad (5.9)$$

где константы d_1, d_2 являются коэффициентами уравнения регрессии

$$y_{t-1} = d_0 + d_1 \cdot x_{t-1} + u_t, \quad (5.10)$$

полученными с помощью обычного МНК.

В результате, для оценки параметров уравнения (5.8) используется уравнение

$$y_t = a + b_0 \cdot x_t + c_1 \cdot \hat{y}_{t-1} + \nu_t, \quad (5.11)$$

где значения переменной $\hat{y}_{t-1}$ рассчитаны по формуле (5.9).

Заметим, что функциональная связь между переменными $\hat{y}_{t-1}$ и x_{t-1} (5.9) приводит к появлению высокой корреляционной связи между переменными $\hat{y}_{t-1}$ и x_t. Для преодоления этой проблем в модель (5.8) и, соответственно, в модель (5.11) можно включить фактор времени в качестве независимой переменной. Модель при этом примет вид

$$y_t = a + b_0 \cdot x_t + c_1 \cdot \hat{y}_{t-1} + c_2 \cdot t + \nu_t. \quad (5.12)$$

Для проверки гипотезы об автокорреляции остатков в моделях авторегрессии (5.8) используется критерий h Дарбина. Фактическое значение критерия вычисляется по формуле

$$h = (1 - \frac{d}{2}) \cdot \sqrt{\frac{n}{1 - n \cdot V}}, \tag{5.13}$$

где d – фактическое значение критерия Дарбина–Уотсона для данной модели

$$d = \frac{\sum_{i=2}^{n} (\nu_t - \nu_{t-1})^2}{\sum_{i=1}^{n} {\nu_t}^2}, \tag{5.14}$$

n – число наблюдений в модели;

V – квадрат стандартной ошибки при лаговой результативной переменной.

В качестве критических значений критерия при уровне значимости α берутся значения $t_{\alpha/2}$ и $t_{1-\alpha/2}$ квантилей порядка α/2 и1–α/2 стандартизованного нормального распределения. Нулевая гипотезы об отсутствии автокорреляции не отвергается, если выполняется условие

$$t_{\alpha/2} < h < t_{1-\alpha/2}. \tag{5.15}$$

Заметим, что этот критерий применим, если $n \cdot V < 1$.

Контрольные вопросы

1. Какие эконометрические модели называются динамическими?
2. Какой вид имеют модели авторегрессии?
3. Какой вид имеют из себя модели с распределенным лагом?
4. Что является значениями лаговых переменных?
5. Как интерпретируются параметры модели с распределенным лагом?
6. Как интерпретируются параметры модели авторегрессии?
7. Как осуществляется оценка параметров модели авторегрессии?
8. Что используется в качестве инструментальной переменной при оценке параметров модели авторегрессии?

Задачи

1. Определить к какому классу относится следующая модель

$y_t = 100 + 70 \cdot x_t + 25 \cdot x_{t-1} + 5 \cdot x_{t-2}$.

(Модель с распределенным лагом второго порядка).

2. Определить к какому классу относится следующая модель

$y_t = 200 + 50 \cdot x_t + 0{,}6 \cdot y_{t-1}$.

(Модель авторегрессии первого порядка).

1. Определить краткосрочный и долгосрочный мультипликаторы для модели

$y_t = 100 + 70 \cdot x_t + 25 \cdot x_{t-1} + 5 \cdot x_{t-2}$. ($b_0 = 70$; $b = 100$).

2. Определить краткосрочный и долгосрочный мультипликаторы для модели

$y_t = 200 + 50 \cdot x_t + 0{,}6 \cdot y_{t-1}$. ($b_0 = 50$; $b = 125$).

3. Если величина критерия h Дарбина равна 0,6, то что можно сказать о наличии автокорреляции в остатках? (Автокорреляция отсутствует).

Лабораторная работа №5. Динамические эконометрические модели: построение модели авторегрессии и оценка ее качества

Задание. На основании данных табл. П1.3 для соответствующего варианта (табл. 5.1):

1. Построить уравнение авторегрессии.

$$y_t = a + b_0 \cdot x_t + c_1 \cdot y_{t-1} + \varepsilon_t .$$

2. Проверить значимость уравнения регрессии и отдельных коэффициентов.
3. Дать интерпретацию полученным значениям параметров уравнения.
4. Проверить наличие автокорреляции в остатках.

Таблица 5.1

Варианты выполнения лабораторных работ №5

Вариант	Номер графы табл. П1.3 для результативной переменной y	Номер графы табл. П1.3 для факторной переменной x	Уровень значимости
1	3	4	0,05
2	3	5	0,01
3	3	11	0,05
4	3	15	0,01
5	10	4	0,05
6	10	5	0,01
7	10	11	0,05
8	10	15	0,01
9	16	4	0,05
10	16	5	0,01
11	16	11	0,05
12	16	15	0,01
13	7	4	0,05
14	7	5	0,01
15	7	11	0,05
16	7	15	0,01
17	14	4	0,05
18	14	5	0,01
19	14	11	0,05
20	14	15	0,01
21	6	4	0,05
22	6	5	0,01
23	6	11	0,05
24	6	15	0,01
25	11	8	0,05

Пример выполнения лабораторной работы №5

Исходные данные:

– данные наблюдений даны в таблице 5.2;

– уровень значимости α = 0,05.

Таблица 5.2

Данные наблюдений

Год наблюдения	Y_t	X_t	Y_{t-1}	Расчетные значения $\hat{Y}_{t-1}$
1	1016,6	1412,7	–	–
2	1435,9	1978,9	1016,6	1060,6
3	1776,1	2292,0	1435,9	1443,8
4	2003,8	2514,4	1776,1	1655,8

Окончание табл. 5.2

1	2	3	4	5
5	3265,7	4632,0	2003,8	1806,3
6	4476,9	7116,6	3265,7	3239,7
7	5886,9	8819,9	4476,9	4921,5
8	7443,2	10627,5	5886,9	6074,5
9	9024,8	12886,1	7443,2	7298,0
10	11401,4	16679,9	9024,8	8826,9
11	14363,5	21079,5	11401,4	11394,9
12	17742,6	26009,7	14363,5	14372,9

1) Построение уравнения авторегрессии.

$$y_t = a + b_0 \cdot x_t + c_1 \cdot y_{t-1} + \varepsilon_t.$$

Для введения инструментальной переменной построим уравнение регрессии

$$\hat{y}_{t-1} = d_0 + d_1 \cdot x_{t-1} + u_t,$$

используя функцию «Сервис.Анализ данных.Регрессия» табличного процессора MS Excel (рис 3.2). Задавая соответствующие диапазоны данных в окне определения параметров регрессии, получим

Таблица 5.3

Результаты регрессионного анализа

Показатели	Коэффициенты уравнения регрессии	Стандартная ошибка определения коэффициентов	t-статистика	Вероятность ошибки α	Нижние 95%–пределы	Верхние 95%–пределы
Y-пересечение	104,31	97,732	1,067	0,314	–116,778	325,394
Переменная X_{t-1}	0,677	0,009	71,398	0,000	0,655	0,698

Уравнение, определяющее инструментальную переменную $\hat{y}_{t-1}$ имеет вид

$$\hat{y}_{t-1} = 104{,}31 + 0{,}677 \cdot x_{t-1}. \tag{5.16}$$

Расчетные значения инструментальной переменной $\hat{y}_{t-1}$ приведены в таблице 5.2.

Используя функцию «Сервис.Анализ данных.Регрессия» табличного процессора MS Excel получим

множественный коэффициент корреляции R = 0,9962,

коэффициент детерминации R^2 = 0,9993,

$F_{факт}$ = 5737,

уровень значимости уравнения регрессии $\alpha = 2{,}36 \cdot 10^{-13}$.

Таблица 5.4

Результаты регрессионного анализа

Показатели	Коэффициенты уравнения регрессии	Стандартная ошибка определения коэффициентов	t-статистика	Вероятность ошибки α	Нижние 95%–пределы	Верхние 95%–пределы
Y-пересечение	139,80	82,870	1,687	0,130	-51,296	330,903
Переменная X_t	0,496	0,078	6,348	0,00022	0,316	0,676
Переменная $\hat{y}_{t-1}$	0,329	0,141	2,328	0,048	0,003	0,655

Уравнение регрессии

$$y_t = 139{,}80 + 0{,}496 \cdot x_t + 0{,}329 \cdot y_{t-1}. \qquad (5.17)$$

2) Проверка значимости.

Уравнение (5.17) значимо при $\alpha = 0{,}05$, так как его значимость $\alpha = 2{,}36 \cdot 10^{-13}$.

Из таблицы 4.4 следуют следующие значения уровней значимости значений параметров уравнения (4.13):

– параметр 139,80: $\alpha = 0, 130$;

– параметр 0,496: $\alpha = 0{,}00022$;

– параметр 0,329: $\alpha = 0{,}048$.

Следовательно при уровне значимости $\alpha = 0{,}05$ параметр 139,80 – не значим, а параметры 0,496 и 0,329 – значимы.

3) Интерпретация значений параметров уравнения.

Краткосрочный мультипликатор $b_0 = 0{,}496$.

Долгосрочный мультипликатор $b = \dfrac{b_0}{1-c_1} = \dfrac{0{,}496}{1-0{,}329} = 0{,}739.$

Таким образом, увеличение x_t на 1 единицу приводит к росту y_t в том же периоде в среднем на 0,496 единиц. Долгосрочное изменение y_t составит 0,739 единиц, т. е. изменение x_t на 1 единицу в каком-либо периоде приведет к изменению y_t в долгосрочной перспективе в среднем на 0,739 единиц.

4) Проверим наличие автокорреляции в остатках для уравнения (5.17) с помощью критерия h Дарбина.

Результаты расчетов значений инструментальной переменной $\hat{y}_{t-1}$ по уравнению (5.16) и остатки показаны в таблице 5.5.

Вычислим величину d (5.14)

$$d = \frac{\sum_{i=2}^{n}(e_t - e_{t-1})^2}{\sum_{i=1}^{n} e_t^{\,2}} = \frac{363417}{202262{,}6} = 1{,}762.$$

Квадрат стандартной ошибки коэффициента при переменой y_{t-1} в (5.17)

$$V = 0{,}141^2 = 0{,}02.$$

Таблица 5.5

Результаты расчетов для проверки автокорреляции остатков

Год наблюдения	Порядковый номер Переменная t	Данные наблюдений Y_t	Расчетные значения $\hat{Y}_t$	Остатки $e_t = \hat{Y}_t - Y_t$	e_t^2	$e_t - e_{t-1}$	$(e_t - e_{t-1})^2$
1		1016,6	–	–	–	–	–
2	1	1435,9	1470,2	34,3	1173,6	–	–
3	2	1776,1	1751,6	-24,5	601,2	-58,8	3454,78
4	3	2003,8	1931,6	-72,2	5208,9	-47,7	2270,81
5	4	3265,7	3031,1	-234,6	55015,6	-162,4	26367,5
6	5	4476,9	4735,0	258,1	66608,5	492,6	242694

Окончание табл. 5.5

1	2	3	4	5	6	7	8
7	6	5886,9	6133,2	246,3	60677,1	-11,8	138,277
8	7	7443,2	7409,1	-34,1	1165,1	-280,5	78658,4
9	8	9024,8	8931,8	-93,0	8655,4	-58,9	3469,30
10	9	11401,4	11316,2	-85,2	7265,8	7,8	60,7634
11	10	14363,5	14343,1	-20,4	418,1	64,8	4197,92
12	11	17742,6	17768,0	25,4	646,8	45,9	2105,04
Сумма					206262,6		363417

Вычислим величину критерия h (5.13)

$$h = (1 - \frac{d}{2}) \cdot \sqrt{\frac{n}{1 - n \cdot V}} = (1 - \frac{1{,}762}{2}) \cdot \sqrt{\frac{12}{1 - 12 \cdot 0.02}} = 0{,}473.$$

Определим значения $t_{\alpha/2}$ и $t_{1-\alpha/2}$ квантилей порядка α/2 и1–α/2 стандартизованного нормального распределения при уровне значимости α = 0,05:

$t_{\alpha/2}$ = НОРМСТОБР(α/2) = НОРМСТОБР(0,05/2) = – 1,96;

$t_{1-\alpha/2}$ = НОРМСТОБР(1–α/2) = НОРМСТОБР(1–0,05/2) = 1,96.

Так как выполняется условие

$$t_{\alpha/2} = -1{,}96 < h = 0{,}473 < t_{1-\alpha/2} = 1{,}96,$$

то делаем вывод об отсутствии автокорреляции в остатках для уравнения (5.17).

Результаты

1) Уравнение авторегрессии

$$y_t = 139{,}80 + 0{,}496 \cdot x_t + 0{,}329 \cdot y_{t-1}.$$

2) Проверка значимости.

Уравнение (5.17) значимо при α = 0,05.

При уровне значимости α = 0,05 параметр 139,80 не значим, а параметры 0,496 и 0,329 – значимы.

3) Интерпретация значений параметров уравнения.

Краткосрочный мультипликатор b_0 = 0,496.

Долгосрочный мультипликатор b = 0,739.

4) Проверка наличия автокорреляции в остатках для уравнения (5.17).

Автокорреляции в остатках отсутствует.

6. Линейные модели стохастических процессов

6.1. Стационарные стохастические процессы

Уровни временного ряда $x_1, x_2, ..., x_n$ при наличии случайной составляющей могут рассматриваться как проявления случайных величин $X_1, X_2,..., X_n$, соответствующих моментам времени $t_1, t_2, ..., t_n$, т. е. как отдельная реализация дискретного стохастического процесса.

Стохастическим процессом называется случайная функция $X(t)$ вещественного аргумента t. Иными словами, если каждому значению аргумента t из некоторого множества $\check{T}$ действительных чисел поставлена в соответствие случайная величина $X_t = X(t)$, то совокупность случайных величин $\{X_t\}$ представляет собой стохастический процесс.

Если множество определения $\check{T}$ случайной функции $X(t)$ дискретно, т. е. $\check{T} = \{t_i\}$, то стохастический процесс называется *дискретным*.

Дискретный стохастический процесс представляет собой последовательность случайных величин X_t, соответствующих моментам времени $t_1, t_2, ..., t_T, ...$.

Стохастический процесс называется *стационарным процессом в широком (слабом) смысле*, если математическое ожидание μ_t и дисперсия σ^2_t не зависят от времени (одинаковы для всех X_t), а автоковариация $\gamma_{t_1 t_2}$ зависит только от величины лага $\tau = t_2 - t_1$, т. е.

$$\begin{aligned} &\mu_t = \mu = \text{const}; \\ &\sigma^2_t = \sigma^2 = \text{const}; \\ &\gamma_{t_1 t_2} = \text{cov}(X_{t_1}, X_{t_2}) = E[(X_{t_1} - \mu)(X_{t_1+\tau} - \mu)] = \gamma(\tau). \end{aligned} \tag{6.1}$$

«Белым шумом» называется последовательность независимых, одинаково распределенных случайных величин a_t. Из определения «белого шума» следует, что

$$\mu_t = \text{const} = \mu; \quad D_t = \sigma^2_t = \text{const} = \sigma^2; \quad \gamma_{t_1 t_2} = 0, \text{ если } t_1 \neq t_2 . \tag{6.2}$$

«Белый шум» является стационарным стохастическим процессом и играет важную роль при моделировании остатков стохастического процесса в уравнениях регрессии.

Зависимость автоковариации $\gamma_\tau = \gamma(\tau)$ от длины лага τ называется *автоковариационной функцией*. При $\tau = 0$ ее значение равно дисперсии, т. е. $\gamma_0 = \gamma(\tau) = \sigma^2$.

Отношение автоковариации $\gamma_\tau = \gamma(\tau)$ к дисперсии $\sigma^2 = \gamma_0$ называется *автокорреляционной функций* стационарного стохастического процесса:

$$\rho_\tau = \frac{\gamma_\tau}{\gamma_0}, \tag{6.3}$$

причем $-1 \le \rho_\tau \le 1$.

Стационарному стохастическому процессу X_t соответствует стационарный временной ряд $x_1, x_2, ..., x_n$.

Признаками стационарности временного ряда являются отсутствие тенденции и периодической составляющей, а также систематических изменений размаха колебаний и систематически изменяющихся взаимозависимостей между элементами временного ряда.

Для распознавания стационарности временных рядов могут использоваться следующие подходы:

– визуальный анализ графического представления временного ряда на наличие тенденции и периодической составляющей, на постоянство дисперсии и т. п.;

– анализ временного ряда на наличие автокорреляции;

– тесты на присутствие детерминистического тренда;

– тесты на постоянство статистических характеристик;

– тесты на наличие стохастического тренда, например, тесты на единичный корень.

6.2. Линейные модели стационарных временных рядов. Процессы ARMA

6.2.1. Модели авторегрессии (AR)

Авторегрессионным процессом порядка p (обозначается AR(p)) называется стохастический процесс X_t, определяемый соотношением

$$X_t = \alpha_0 + \alpha_1 X_{t-1} + \alpha_2 X_{t-2} + \ldots + \alpha_p X_{t-p} + \varepsilon_t, \qquad (6.4)$$

где ε_t – процесс типа «белый шум» с $\mu_\varepsilon = 0$. Свободный член α_0 часто приравнивается нулю (т. е. рассматриваются центрированные процессы, средний уровень которых равен нулю).

Авторегрессионная модель временного ряда основана на предположении, что поведение какого-либо экономического явления в будущем определяется только его текущим и предыдущими состояниями.

AR-процесс является *стационарным* тогда и только тогда, когда комплексные решения (корни) его характеристического уравнения

$$1 - \alpha_1 z - \alpha_2 z^2 - \ldots - \alpha_p z^p = 0 \qquad (6.5)$$

лежат вне единичного круга, т. е. $|z| > 1$ (z – комплексное число).

Процессы, у которых $|z| = 1$, называются *процессами единичного корня* и являются *нестационарными*.

Для процесса AR(1)

$$X_t = \alpha_0 + \alpha_1 X_{t-1} + \varepsilon_t$$

характеристическое уравнение имеет вид

$$1 - \alpha_1 z = 0.$$

Неравенство $|z| > 1$ выполняется, если $|\alpha_1| < 1$. Следовательно, соотношение $|\alpha_1| < 1$ есть условие стационарности процесса AR(1).

6.2.2. Модели скользящего среднего (MA)

В моделях скользящего среднего порядка среднее текущее значение стационарного стохастического процесса представляется в виде линейной комбинации текущего и прошедших значений ошибки ε_t, ε_{t-1}, …, ε_{t-p}, обладающей свойствами «белого шума».

Процессом скользящего среднего порядка q (обозначается MA(q)) называется стохастический процесс X_t, определяемый соотношением

$$X_t = \varepsilon_t - \beta_1\varepsilon_{t-1} - \beta_2\varepsilon_{t-2} - ... - \beta_q\varepsilon_{t-q}, \qquad (6.6)$$

где ε_t – процесс типа «белый шум» с $\mu_\varepsilon = 0$, $\sigma^2_\varepsilon = \sigma^2$.

Процесс MA(q) обладает следующими свойствами:

$$E[X_t] = 0; \qquad D[X_t] = \sigma^2 \sum_{i=0}^{q} \beta_i^2. \qquad (6.7)$$

Согласно (6.7), среднее значение, дисперсия и ковариация не зависят от времени, поэтому процесс MA стационарен в широком смысле.

6.2.3. Модели авторегрессии-скользящего среднего (ARMA)

Комбинация процессов авторегрессии и скользящего среднего порядков p и q соответственно называется **авторегрессионным процессом скользящего среднего** (ARMA(p,q))

$$X_t = \alpha_0 + \alpha_1X_{t-1} + \alpha_2X_{t-2} + ... + \alpha_pX_{t-p} + \varepsilon_t - \beta_1\varepsilon_{t-1} - \beta_2\varepsilon_{t-2} - ... - \beta_q\varepsilon_{t-q}. \qquad (6.8)$$

Использование ARMA-процессов позволяет строить более компактные модели реальных временных рядов по сравнению со схожими по поведению AR- или MA-процессами.

6.3. Автокорреляционные функции

6.3.1. Автокорреляционная функция

Автокорреляционная функция (ACF) процесса X_t, определяющая зависимость коэффициентов автокорреляции ρ_τ от величины лага τ, определяется с помощью соотношения (см. (6.3))

$$\rho_\tau = \rho(\tau) = \frac{\gamma_\tau}{\gamma_0} = \frac{1}{\gamma_0} E[(X_t - \mu)(X_{t+\tau} - \mu)]. \qquad (6.9)$$

График ρ_τ называется *коррелограммой*.

Для идентификации модели стационарного временного ряда, т. е. для определения типа и порядка процесса, могут быть использованы следующие свойства автокорреляционной функции:

а) Для процесса AR(p) коррелограмма представляет собой смесь экспоненциальной кривой и синусоиды.

б) Для процесса MA(q) только первые q автокорреляционных коэффициентов значимо отличны от нуля.

В качестве примера рассмотрим автокорреляционные функции процессов AR(1) и MA(1).

Для процесса AR(1) без свободного члена и с $\alpha_1 < 1$

$$X_t = \alpha_1X_{t-1} + \varepsilon_t \qquad (6.10)$$

автокорреляционная функция определяется соотношениями $\rho_1 = \alpha_1$, и $\rho_k = \alpha_1^k$ (рис. 6.1, а, б).

Для процесса MA(1)

$$X_t = \varepsilon_t - \beta_1 \cdot \varepsilon_{t-1} \qquad (6.11)$$

автокорреляционная функция определяется соотношениями $\rho_1 = \dfrac{-\beta_1}{(1+\beta_1^2)}$, $\rho_2 = 0$, $\rho_3 = 0, \ldots$ (рис. 6.2, а, б).

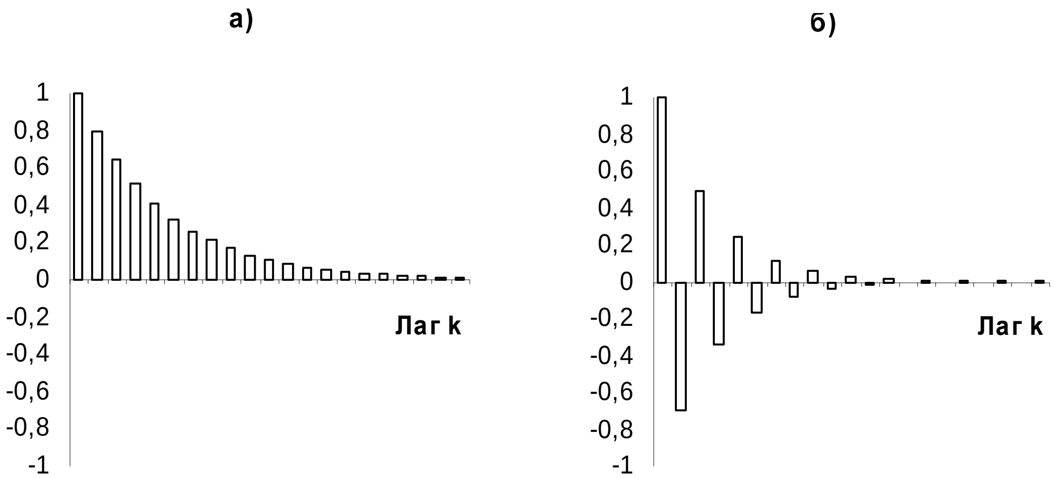


Рис. 6.1. Кореллограмма процесса AR(1)
а) $\alpha_1 > 0$; б) $\alpha_1 < 0$

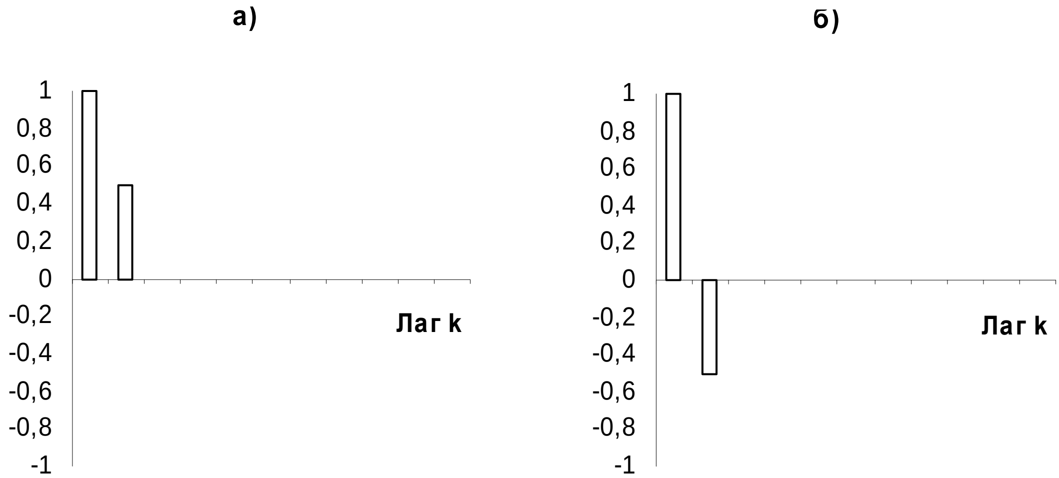


Рис. 6.2. Кореллограмма процесса MA(1)
а) $\beta_1 < 0$; б) $\beta_1 > 0$

6.3.2. Частная автокорреляционная функция

Важную информацию о структуре модели стационарного стохастического процесса можно получить, используя частную автокорреляционную функцию.

Рассмотрим аппроксимацию AR(k) стационарного стохастического процесса X_t

$$X^{(k)}_t = \alpha_{0k} + \alpha_{1k} X^{(k)}_{t-1} + \alpha_{2k} X^{(k)}_{t-2} + \ldots + \alpha_{kk} X^{(k)}_{t-k}. \qquad (6.12)$$

Коэффициент α_{kk} называется *коэффициентом частной автокорреляции* X_t для величины лага k.

Ряд $p_{\mathrm{part}}(k) = \alpha_{kk}$ с различными k называется *частной автокореляционной функцией* (PACF).

Для процесса AR(p) значения частной автокореляционной функции $\rho_{\mathrm{part}}(\tau)$ равны нулю для величины лага $\tau > p$.

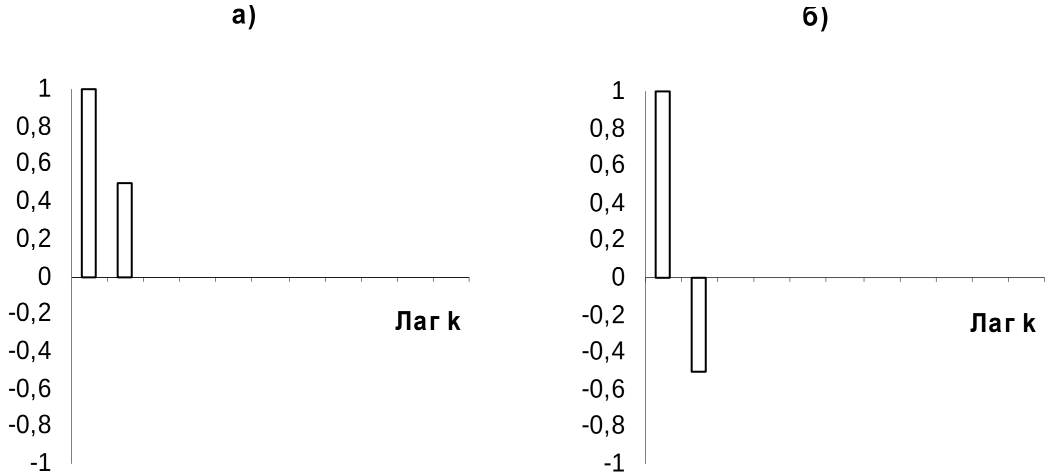


Рис. 6.3. Частная автокорреляционная функция процесса AR(1)
а) $\alpha_1 > 0$; б) $\alpha_1 < 0$

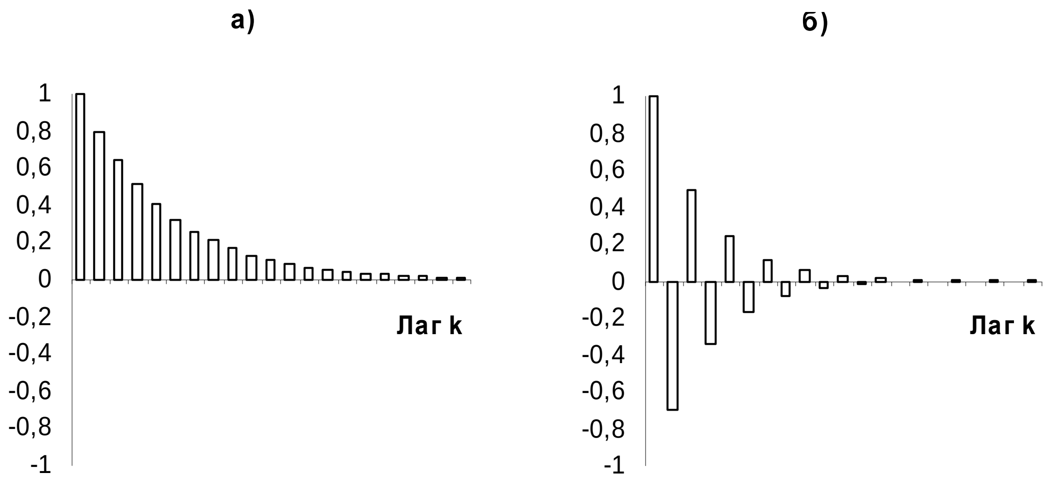


Рис. 6.4. Частная автокорреляционная функция процесса MA(1)
а) $\beta_1 < 0$; б) $\beta_1 > 0$

Для процессов MA(q) значения частной автокореляционной функции экспоненциально убывают с величиной лага q.

В качестве значения частной автокореляционной функции $\rho_{\text{part}}(k)$ при заданной величине лага k может быть использована оценка коэффициента $\acute{\alpha}_{kk}$ модели AR(k) (6.12), полученная с помощью МНК-оценивания.

6.4. Прогнозирование ARMA-процессов

Формулы прогнозирования для процессов ARMA(p,q) получаются с помощью подстановки в правую часть уравнения модели (6.8) известных значений переменной и ошибки и прогнозных значений переменной, полученных на предыдущих шагах прогнозирования. Покажем это на примере модели ARMA(3,3)

$$Y_t = \alpha_0 + \alpha_1 Y_{t-1} + \alpha_2 Y_{t-2} + \alpha_p Y_{t-3} + \varepsilon_t - \beta_1 \varepsilon_{t-1} - \beta_2 \varepsilon_{t-2} - \beta_3 \varepsilon_{t-3.} \qquad (6.13)$$

При расчете прогнозного значения $\hat{Y}_{\text{T}}(h)$ в правую часть (6.13) вместо Y_{T+i} ($i > 0$) следует подставлять вычисленное ранее прогнозное значение $\hat{Y}_{\text{T}}(i)$. Тогда точечный прогноз будет определяться соотношениями:

$$\hat{Y}_T(1) = \alpha_0 + \alpha_1 Y_T + \alpha_2 Y_{T-1} + \alpha_3 Y_{T-2} - \beta_1 \cdot \varepsilon_T - \beta_2 \varepsilon_{T-1} - \beta_3 \varepsilon_{T-2},$$
$$\hat{Y}_T(2) = \alpha_0 + \alpha_1 \hat{Y}_T(1) + \alpha_2 Y_T + \alpha_3 Y_{T-1} - \beta_2 \varepsilon_T - \beta_3 \varepsilon_{T-1}, \qquad (6.14)$$
$$\hat{Y}_T(3) = \alpha_0 + \alpha_1 \hat{Y}_T(2) + \alpha_2 \hat{Y}_T(1) + \alpha_3 Y_T - \beta_3 \varepsilon_T,$$
$$\hat{Y}_T(h) = \alpha_0 + \alpha_1 \hat{Y}_T(h-1) + \alpha_2 \hat{Y}_T(h-2) + \alpha_3 \hat{Y}_T(h-3) \quad \text{при } h > 3.$$

6.5. Нестационарные интегрируемые процессы

6.5.1. Нестационарные стохастические процессы. Нестационарные временные ряды

Признаком нестационарного стохастического процесса является нарушение одного из условий стационарности (6.1). Конкретная реализация нестационарного стохастического процесса представляет собой нестационарный временной ряд. Признаками нестационарности временного ряда могут служить наличие тенденции, систематических изменений дисперсии, периодической составляющей, систематически изменяющихся взаимозависимостей между элементами временного ряда.

Заметим, что, как правило, значения, характеризующие изменение экономических показателей во времени, образуют нестационарные временные ряды.

Рассмотрим авторегрессионный процесс первого порядка, определяемый моделью

$$Y_t = \alpha_0 + \alpha_1 \cdot Y_{t-1} + \varepsilon_t, \qquad (6.15)$$

где ε_t – процесс типа «белый шум» с $\mu_\varepsilon = 0$. При $|\alpha_1| < 1$ случайный процесс Y_t будет стационарным. Процесс, определяемый соотношением (6.21) при $\alpha_1 = 1$

$$Y_t = Y_{t-1} + \varepsilon_t \qquad (6.16)$$

является нестационарным и называется «случайным блужданием». Такие нестационарные процессы называют *процессами единичного корня*.

Среднее процесса Y_t постоянно $E(Y_t) = E(Y_{t-1}) + E(\varepsilon_t) = \mu = \text{const}$, а дисперсия $\text{var}(Y_t) = t\sigma^2$ неограниченно возрастает с течением времени. Первые разности Y_t являются «белым шумом» ε_t и стационарны:

$$\Delta Y_t = Y_t - Y_{t-1} = \varepsilon_t.$$

Как показывает практика, рассматриваемые в эконометрических исследованиях нестационарные временные ряды чаще всего относятся именно к этому типу и проблема выявления нестационарности временного ряда сводится к проверке $\alpha_1 = 1$ в модели (6.15). Соответствующие тесты называются «тестами единичного корня».

6.5.2. Тесты Дики-Фуллера

Тест Дики-Фуллера (Dickey-Fuller test, DF-тест) основан на оценке параметра $\lambda = \alpha_1 - 1$ уравнения

$$\Delta Y_t = \lambda \cdot Y_{t-1} + \varepsilon_t, \qquad (6.17)$$

эквивалентного уравнению авторегрессии (6.15). Его называют также *тестом на единичный корень*.

Нулевая H_0 и ей альтернативная H_1 гипотезы определяются соотношениями: H_0: $\lambda = 0$; H_1: $\lambda < 0$.

Если значение *t*-статистики Стьюдента для параметра λ меньше нижнего порогового значения *DF*-статистики, то нулевую гипотезу $\lambda = 0$ (о наличии единичного корня $\alpha_1 = 1$) следует отклонить и принять альтернативную о стационарности процесса Y_t.

Таблицы теста Дики-Фуллера (DF-теста) рассчитаны для уровней значимости в 1, 5, 10%. Указанные в таблице значения *DF-теста* – отрицательные.

DF-тест применим также для тестирования на единичный корень *случайных процессов со смещением* и *со смещением и линейным детерминистическим трендом*, определяемых уравнениями:

$$\Delta Y_t = \alpha_0 + \alpha_1 \cdot Y_{t-1} + \varepsilon_t, \tag{6.18}$$

$$\Delta Y_t = \alpha_0 + \alpha_1 \cdot Y_{t-1} + \alpha_2 \cdot t + \varepsilon_t, \tag{6.19}$$

где α_0 – константа, называемая смещением. При этом используются соответствующие таблицы критических значений DF-теста.

Отметим, что на практике трудно различить ситуации, когда следует применять DF-тест, а когда – DF-тест со смещением.

6.5.3. Метод разностей и интегрируемость

Для практики большой интерес представляют, так называемые, интегрируемые нестационарные процессы. Это процессы, для которых с помощью последовательного применения операции взятия последовательных разностей из нестационарных временных рядов можно получить стационарные ряды.

Последовательные разности стохастического процесса определяются соотношениями:

$\Delta Y_t = Y_t - Y_{t-1}$ – первые последовательные разности

$\Delta^2 Y_t = \Delta Y_t - \Delta Y_{t-1}$ – вторые последовательные разности и т. д.

Если первые разности нестационарного ряда Y_t стационарны, то ряд Y_t называется интегрируемым первого порядка. Стационарный временной ряд называется интегрируемым нулевого порядка.

Если первые разности нестационарного ряда нестационарны, а вторые разности стационарны, то ряд Y_t называется интегрируемым второго порядка. Если первый стационарный ряд получается после *k*-кратного взятия разностей, то ряд Y_t называется интегрируемым *k*-го порядка.

6.6. Модели ARIMA

6.6.1. Определение и идентификация модели

Рассмотрим интегрируемый порядка *d* нестационарный процесс X_t. Если при этом процесс $Y_t = \Delta^d X_t$, составленный из первых разностей *d*-порядка исходного процесса, является процессом ARMA(p,q), т. е.

$$Y_t = \alpha_0 + \alpha_1 Y_{t-1} + \alpha_2 Y_{t-2} + \ldots + \alpha_p Y_{t-p} + \varepsilon_t - \beta_1 \varepsilon_{t-1} - \beta_2 \varepsilon_{t-2} - \ldots - \beta_q \varepsilon_{t-q}, \tag{6.20}$$

тогда X_t называется процессом ARIMA(p,d,q). На практике свободный член α_0 часто опускается (приравнивается к нулю).

Можно считать, что большинство эмпирических временных рядов является реализациями процессов ARIMA.

Основная проблема в анализе временных рядов заключается в определении порядка модели ARIMA(p,d,q).

Необходимо оценить три основных параметра: d – порядок интегрируемости, порядок p компоненты AR и порядок q компоненты MA. Для экономических временных рядов параметр d обычно равен 1, возможны также значения 0 или 2. При определении параметров p и q используются характеристики автокорреляционной функции (ACF) и частной автокорреляционной функции (PACF). При этом предпочтение отдается моделям с наименьшим числом параметров.

При выборе наилучшей модели из нескольких вариантов более предпочтительной при прочих равных условиях считается модель:

1) с меньшим числом параметров;

2) с меньшим значением R^2;

3) с меньшим значением суммы квадратов остатков $\sum(\hat{y}_t - y_t)^2$;

4) с меньшим значением информационного критерия Акаике AIC.

6.6.2. Прогнозирование ARIMA-процессов

Для прогнозирования ARIMA-процессов X_t могут быть применены два подхода:

1) Получение прогнозных значений $\hat{Y}_T(h)$ ARMA-процесса $Y_t = \Delta^d X_t$ по методике прогнозирования ARMA-процессов (см. разд. 6.4) с последующим последовательным вычислением прогнозных значений $\Delta^{d-1}\hat{X}_T(h)$, $\Delta^{d-2}\hat{X}_T(h)$ и т. д., пока не будут получены $\hat{X}_T(h)$.

2) Построение прогнозной формулы с помощью модификации уравнения (6.20) путем подстановки разностей $\Delta^d X_t$ вместо Y_t и последующего разрешения полученного уравнения относительно X_t. В результате, будет получена ARMA-модель нестационарного процесса, которая может быть преобразована в формулу для прогнозирования на h шагов вперед величин $\hat{X}_T(h)$ с началом отсчета в момент времени T по методике, описанной в разделе 6.4.

Рассмотрим ARIMA(0,1,0)–модель случайного блуждания $Y_t = \Delta X_t = \varepsilon_t$ или в преобразованном виде $X_t = X_{t-1} + \varepsilon_t$.

Формула экстраполяции имеет вид

$$X_{T+h} = X_{T+h-1} + \varepsilon_t, \tag{6.21}$$

а формула прогноза дается соотношением

$$\hat{X}_T(h) = X_T, \qquad \text{для } h \geq 1. \tag{6.22}$$

Дисперсия ошибки прогноза var(e_T(h))= $h \cdot \sigma^2_\varepsilon$. увеличивается с ростом h. Ширина доверительного интервала прогноза возрастает пропорционально $\sqrt{h}$.

Если X_t – случайное блуждание со сдвигом

$$X_t = X_{t-1} + \alpha_0 + \varepsilon_t, \tag{6.23}$$

тогда формула для прогнозирования имеет вид

$$\hat{X}_{T(+h)} = X_T + h\alpha_0, \tag{6.24}$$

что соответствует простому линейному тренду. Дисперсия ошибки прогноза такая же, как и в предыдущем случае с $\alpha_0 = 0$.

Рассмотрим ARIMA(1,1,1)–модель

$$\Delta X_t - \alpha_1 \Delta X_{t-1} = X_t - X_{t-1} - \alpha_1 \cdot (X_{t-1} - X_{t-2}) = \alpha_0 + \varepsilon_t - \beta_1 \cdot \varepsilon_{t-1},$$

которая после преобразования принимает вид

$$X_t = \alpha_0 + (1 + \alpha_1) X_{t-1} - \alpha_1 X_{t-2} + \varepsilon_t - \beta_1 \cdot \varepsilon_{t-1}. \quad (6.25)$$

Формулы для прогнозирования в момент $t = T + h$ определяются соотношениями

$$\hat{X}_T(1) = \alpha_0 + (1+\alpha_1) X_T - \alpha_1 X_{T-1} - \beta_1 \varepsilon_T,$$
$$\hat{X}_T(2) = \alpha_0 + (1+\alpha_1)\hat{X}_T(1) - \alpha_1 X_T, \quad (6.26)$$
$$\hat{X}_T(h) = \alpha_0 + (1+\alpha_1)\hat{X}_T(h-1) - \alpha_1 X_T(h-2) \text{ для } h \geq 3.$$

Контрольные вопросы

1. Дайте определение стохастического процесса.
2. Дайте определение стационарного стохастического процесса в слабом (широком) смысле.
3. Какой стохастический процесс называется «белый шумом»?
4. Какими параметрами характеризуется стационарный процесс?
5. Дайте определение автоковариационной функции.
6. Охарактеризуйте процессы AR.
7. В каких случаях процессы AR являются стационарными?
8. Охарактеризуйте процессы MA.
9. Охарактеризуйте процессы ARMA.
10. Опишите модель ARMA(3,2).
11. Как используется автокорреляционная функция для идентификации модели стационарного стохастического процесса?
12. Как используется частная автокорреляционная функция для идентификации модели стационарного стохастического процесса?
13. Как осуществляется прогнозирование ARMA-процессов?
14. Что может служить признаком нестационарности временного ряда?
15. Для чего применяются тесты Дики-Фуллера?
16. Охарактеризуйте процессы ARIMA.
17. Как осуществляется прогнозирование ARMA-процессов на примере процесса ARMA(2,2)?

Задачи

1. По представленному уравнению модели стохастического процесса определить ее тип.

$$Y_t = 10 + 0{,}7 \cdot Y_{t-1} - 0{,}08 \cdot Y_{t-2} + 0{,}02 \cdot Y_{t-3} + \varepsilon_t - 0{,}24 \cdot \varepsilon_{t-1}.$$

(ARMA(3,1)).

2. Является ли стационарным стохастический процесс $Y_t = 10 + 0{,}7 \cdot Y_{t-1} + \varepsilon_t$?

(Является).

3. Если значение t-статистики Стьюдента для параметра λ в уравнении

$$\Delta Y_t = \lambda \cdot Y_{t-1} + \varepsilon_t,$$

принимает значение $t_\lambda = -1{,}1$, то что можно сказать о стационарности стохастического процесса при уровне значимости 0,05? (Гипотеза о нестационарности не отвергается)

4. Написать формулу для вычисления прогнозного значения $\hat{Y}_T(2)$ переменной Y для модели

$$Y_t = 5 + 0{,}6Y_{t-1} + 0{,}04Y_{t-2} + \varepsilon_t - 0{,}3 \cdot \varepsilon_{t-1}.$$

($\hat{Y}_T(2) = 5 + 0{,}6\ \hat{Y}_T(1) + 0{,}04Y_T$).

Лабораторная работа №6. Моделирование стохастических процессов

Задание: На основании данных табл. П1.4 для соответствующего варианта

1) проверить ряд Y_t на нестационарность и интегрируемость не более, чем до 3-го порядка с помощью теста Дики-Фуллера.

2) построить модель ARIMA(p,d,q) временного ряда Y_t порядка не выше, чем p=2 и q=2. (Лучшая модель выбирается из возможных как имеющая наименьшее значение R^2, наименьшее значение суммы квадратов остатков $\Sigma\varepsilon^2 = \sum (\hat{y}_t - y_t)^2$, наименьшее значение информационного критерия Акаике AIC).

3) Расчитать прогнозные значения по полученной модели на 1 и 2 шага вперед.

Пример выполнения лабораторной работы №6

Исходные данные:

- данные наблюдений Y_t даны в таблице 6.1 (столбец Y_t);
- уровень значимости $\alpha = 0{,}05$.

Таблица 6.1

Данные наблюдений и результаты промежуточных расчетов

	Y_t	Y_{t-1}	ΔY_t	$(\Delta Y)_{t-1}$	$\Delta^2 Y_t$	$\hat{Y}_t$ ARMA(1,0)
1	2143	–	–	–	–	–
2	2174	2143	31	–	–	2143
3	2136	2174	–38	31	–69	2174
4	2312	2136	176	–38	214	2136
5	2262	2312	–50	176	–226	2313
6	2188	2262	–74	–50	–24	2263
7	2163	2188	–25	–74	49	2188
8	2172	2163	9	–25	34	2163
9	2135	2172	–37	9	–46	2172
10	2139	2135	4	–37	41	2135
11	2034	2139	–105	4	–109	2139
12	2077	2034	43	–105	148	2034
13	2059	2077	–18	43	–61	2077

Окончание таблицы 6.1

	Y_t	Y_{t-1}	ΔY_t	$(\Delta Y)_{t-1}$	$\Delta^2 Y_t$	$\hat{Y}_t$ AR-MA(1,0)
14	2003	2059	–56	–18	–38	2059
15	1894	2003	–109	–56	–53	2002
16	1879	1894	–15	–109	94	1893
17	1710	1879	–169	–15	–154	1878
18	1770	1710	60	–169	229	1708
19	1585	1770	–185	60	–245	1768
20	1735	1585	150	–185	335	1582
21	1894	1735	159	150	9	1733
22	1753	1894	–141	159	–300	1893
23	1640	1753	–113	–141	28	1751
24	1677	1640	37	–113	150	1638
25	1737	1677	60	37	23	1675
26	1742	1737	5	60	–55	1735
27	1673	1742	–69	5	–74	1740
28	1620	1673	–53	–69	16	1671
29	1493	1620	–127	–53	–74	1618
30	1529	1493	36	–127	163	1490
31	1549	1529	20	36	–16	1526
32	1646	1549	97	20	77	1546
33	1698	1646	52	97	–45	1644
34	1687	1698	–11	52	–63	1696
35	1721	1687	34	–11	45	1685
36	1762	1721	41	34	7	1719
37	1760	1762	–2	41	–43	1760
38	1660	1760	–100	–2	–98	1758
39	1629	1660	–31	–100	69	1658
40	1682	1629	53	–31	84	1627
41	1598	1682	–84	53	–137	1680
42	1553	1598	–45	–84	39	1595
43	1497	1553	–56	–45	–11	1550
44	1558	1497	61	–56	117	1494
45	1496	1558	–62	61	–123	1555
46	1489	1496	–7	–62	55	1493
47	1478	1489	–11	–7	–4	1486
48	1373	1478	–105	–11	–94	1475
49	1340	1373	–33	–105	72	1369
50	1409	1340	69	–33	102	1336
51	1516	1409	107	69	38	1405
52	1490	1516	–26	107	–133	1513
h = 1	Прогнозные значения					1487
h =2						1483

1) Проверка ряда Y_t на нестационарность. Построим график (t, Y_t) (рис.6.5)

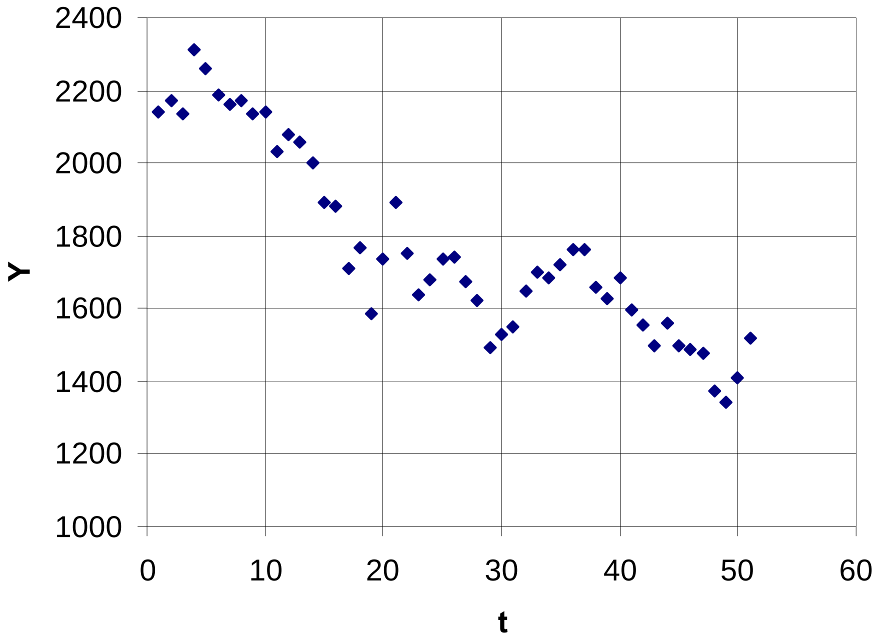


Рис. 6.5. График (t, Y_t)

На рис. 6.5 ясно проглядывается наличие нисходящей тенденции, что говорит о нестационарности временного ряда Y_t.

Проверим нестационарность временного ряда Y_t с помощью теста Дики-Фуллера. Необходимо оценить параметр λ уравнения

$$\Delta Y_t = \lambda \cdot Y_{t-1} + \varepsilon_t, \qquad (6.33)$$

где $\Delta Y_t = Y_t - Y_{t-1}$ являются первыми разностями ряда Y_t (табл. 6.1).

Будем использовать программный пакет Matrixer 5.1. Введем два вектора DY = ΔY_t и Y1 = Y_{t-1}, с предварительно рассчитанными значениями (табл. 6.1), наберем в командном окне программы формулу DY: Y1 и нажмем кнопку . Результат показан на рис. 6.6.

Из рис. 6.6 следует, что $\lambda = -0{,}0082$ и t-статистика равна $-1{,}327 > f_{крит} = -1{,}95$ (критическое значение при уровне значимости 0,05). Следовательно, гипотеза о нестационарности не отвергается.

Проверим на нестационарность ряд ΔY_t, предварительно рассчитав вторые последовательные разности $\Delta^2 Y_t = \Delta Y_t - \Delta Y_{t-1}$ (табл. 6.1). Построим график (t, ΔY_t) (рис. 6.7). Анализ рис. 6.7 показывает, что временной ряд ΔY_t больше похож на стационарный.

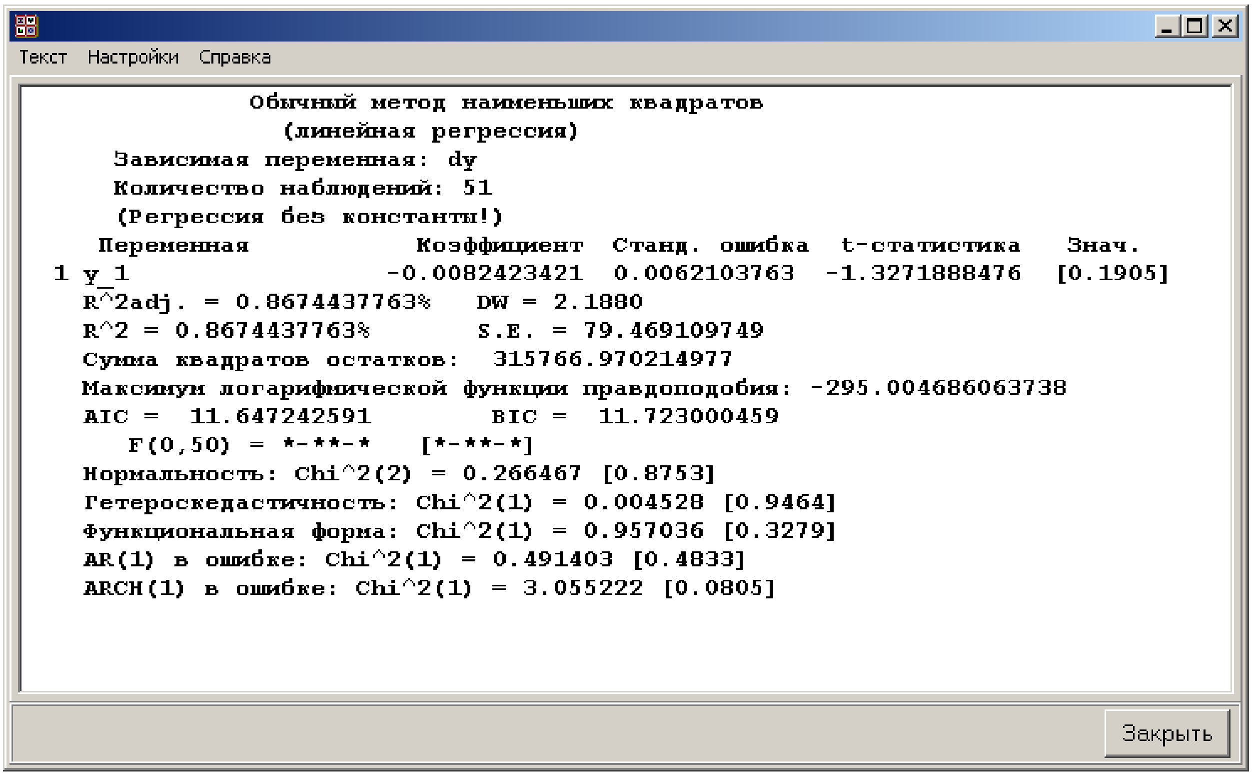


Рис. 6.6. Результаты оценки параметра λ

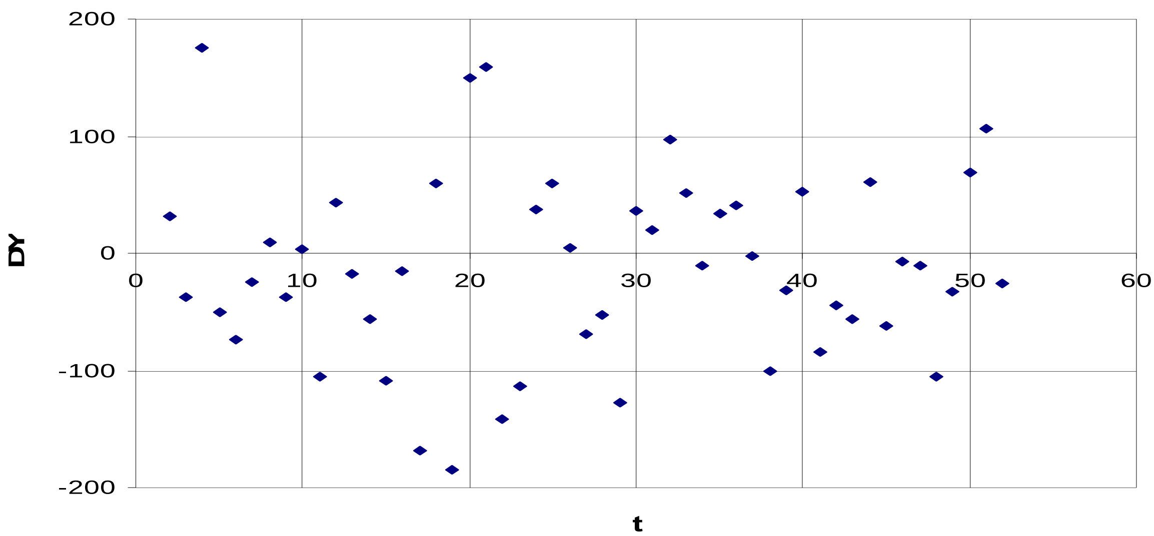


Рис. 6.7. График (t, Y_t)

Согласно тесту Дики-Фуллера, для оценки параметра λ_1 в уравнении $\Delta^2 Y_t = \lambda_1 \cdot \Delta Y_{t-1} + \varepsilon_{t1}$, введем два вектора D2Y = $\Delta^2 Y_t$ и DY1 = ΔY_{t-1}, с предварительно рассчитанными значениями (табл. 6.1), наберем в командном окне программы формулу DY: Y1 и нажмем кнопку ▶. Результат показан на рис. 6.8.

```
              Обычный метод наименьших квадратов
                  (линейная регрессия)
   Зависимая переменная: d2y
   Количество наблюдений: 50
   (Регрессия без константы!)
  Переменная              Коэффициент  Станд. ошибка  t-статистика   Знач.
1 Dy_1                   -1.068154438  0.1424624058  -7.4977986814  [0.0000]
  R^2adj. = 53.425227757%    DW = 2.0015
  R^2 = 53.425227757%        S.E. = 81.367736972
  Сумма квадратов остатков:  324414.722379412
  Максимум логарифмической функции правдоподобия: -290.390801114713
  AIC =  11.695632045        BIC =  11.772112965
```

Рис. 6.8. Результаты оценки параметра λ_1

Из рис. следует, что $\lambda_1 = -1{,}068 > 0$ и t-статистика равна $-7{,}498 < f_{крит} = -1{,}95$ (критическое значение при уровне значимости 0,05). Следовательно, гипотеза о нестационарности отвергается.

Таким образом, временной ряд Y_t является интегрируемым первого порядка нестационарным временным рядом.

2) Построение модели нестационарного ряда Y. Построим автокорреляционную и частную автокорреляционную функции. Для этого введем вектор Y со значениями ряда Y_t, наберем в в командном окне программы формулу «acf! Y &pacf» и нажмем кнопку . Результат показан на рис. 6.9.

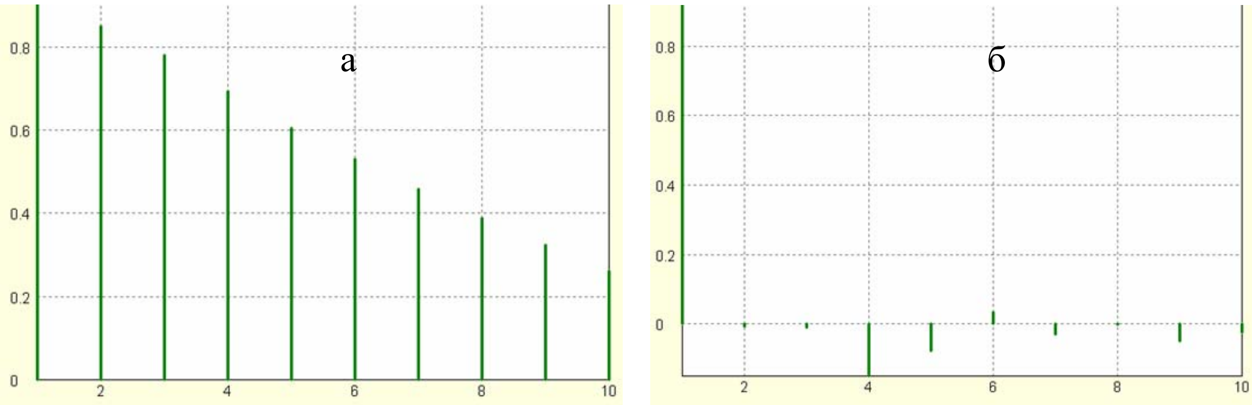


Рис. 6.9. Автокорреляционная (а) и частная автокорреляционная (б) функции

Так как, согласно рис. 6.9, значимо отличается от нуля только коэффициент частной автокорреляции α_{11} для величины лага 1, то делаем вывод о том, что ряд Y_t является реализацией стохастического процесса ARMA(1,0).

$$Y_t = \alpha_0 + \alpha_1 \cdot Y_{t-1}, \quad \text{с } |\alpha_1| > 1.$$

Для построения модели ARMA(1,0) наберем в командном окне программы формулу «boxjen! (1,0) Y » и нажмем кнопку .

В появившемся окне выбора метода нахождения решения (рис. 6.10) следует выбрать метод (рис. 6.11) и нажать кнопку .

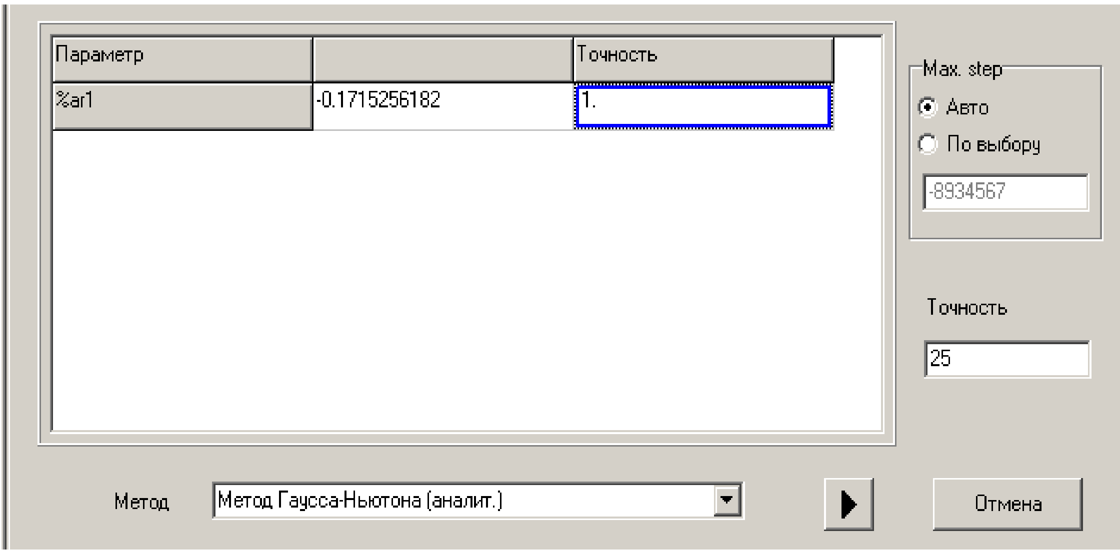


Рис. 6.10. Окно выбора метода нахождения решения

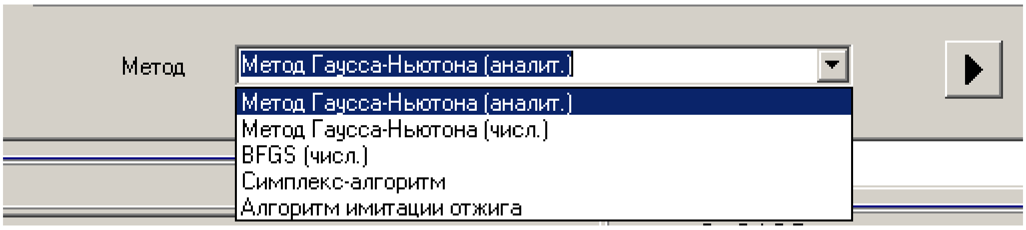


Рис. 6.11. Выбор метода оптимизации

Построение ARMA осуществляется с помощью нелинейного метода наименьших квадратов. Соответствующие процедуры носят итерационный характер и могут продолжаться бесконечно долго, если не достигаются условия окончания процесса по точности получаемого решения. В этом случае следует задавать конечное число шагов итерационного процесса, либо принудительно останавливать процесс решения с помощью кнопки «Результаты» в окне процесса решения рис. 6.12.

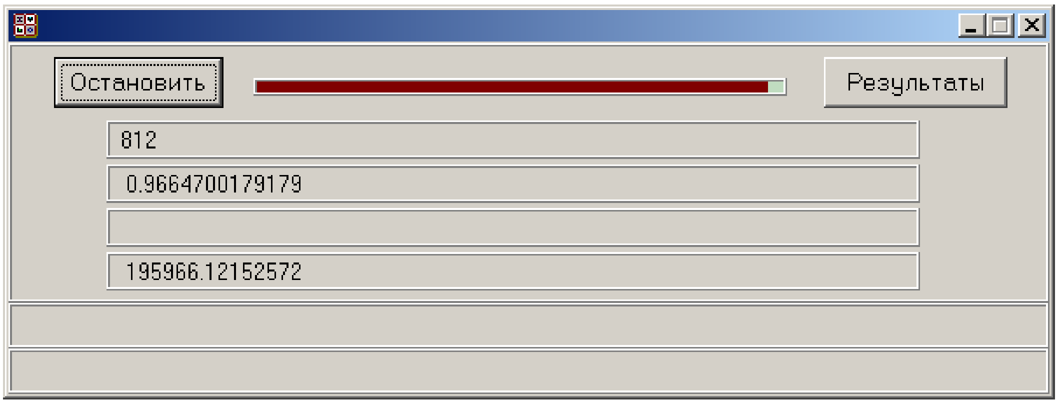


Рис. 6.12. Окно процесса решения с помощью нелинейного МНК

Результаты расчета модели ARMA(1,0) показаны на рис. 6.13.

```
              Модель Бокса-Дженкинса (ARIMA)
    Метод оценивания: Условный метод наименьших квадратов
    Зависимая переменная: Y, уровни
    Количество наблюдений: 52
   Переменная            Коэффициент  Станд. ошибка  t-статистика   Знач.
 1 Константа            -10.552765202  53.578999132  -0.196957117   [0.8447]
 2 %ar1                  1.0049179762  0.0249744387   40.237860264  [0.0000]
  R^2 = 90.612708887%        S.E. = 80.826358398
  Сумма квадратов остатков:  326645.010595287
  Максимум логарифмической функции правдоподобия: -301.164827951337
  AIC =  11.698647229         BIC =  11.811218982
     F(1,50) = 482.635   [0.0000]
```

Рис. 6.13. Результаты расчета модели ARMA(1,0)

Значимым является только коэффициент при Y_{t-1} (графа «Знач.»).

Для проверки оптимальности модели ARMA(1,0) рассчитаем параметры моделей ARMA(2,0) и ARMA(2,1). Набирая в командном окне программы формулы «boxjen! (2,0) Y» и «boxjen! (2,1) Y» получим (рис. 6.14 и 6.15):

```
              Модель Бокса-Дженкинса (ARIMA)
    Метод оценивания: Условный метод наименьших квадратов
    Зависимая переменная: Y, уровни
    Количество наблюдений: 52
   Переменная            Коэффициент  Станд. ошибка  t-статистика   Знач.
 1 Константа            -16.634326494  54.759147499  -0.3037725614  [0.7626]
 2 %ar1                  0.9303305851  0.1425663274   6.5255983088  [0.0000]
 3 %ar2                  0.0774075847  0.1453339343   0.5326187935  [0.5967]
  R^2 = 90.666317489%        S.E. = 81.41348458
  Сумма квадратов остатков:  324779.618097815
  Максимум логарифмической функции правдоподобия: -301.015922505134
  AIC =  11.731381635         BIC =  11.881477305
     F(2,49) = 237.9902 [0.0000]
```

Рис. 6.14. Результаты расчета модели ARMA(2,0)

```
              Модель Бокса-Дженкинса (ARIMA)
    Метод оценивания: Условный метод наименьших квадратов
    Зависимая переменная: Y, уровни
    Количество наблюдений: 52
   Переменная            Коэффициент  Станд. ошибка  t-статистика   Знач.
 1 Константа            -14.172754911  23.44177883   -0.6045938328  [0.5483]
 2 %ar1                  1.5478339656  0.6689820619   2.3137152006  [0.0250]
 3 %ar2                 -0.5412734023  0.6746066931  -0.8023540351  [0.4263]
 4 %ma1                 -0.6625943829  0.6020610004  -1.1005436035  [0.2766]
  R^2 = 90.794084396%        S.E. = 81.692229215
  Сумма квадратов остатков:  320333.775081054
  Максимум логарифмической функции правдоподобия: -300.657555279547
  AIC =  11.756059818         BIC =  11.943679407
     F(3,48) = 157.8013 [0.0000]
```

Рис. 6.15. Результаты расчета модели ARMA(2,1)

Значимым в обоих моделях является только коэффициент при Y_{t-1}.

Сведем результаты расчетов параметров моделей ARMA(1,0), ARMA(2,0) и ARMA(2,1) в таблицу 6.2.

Таблица 6.2

Результаты расчетов параметров моделей ARMA(1,0), ARMA(2,0) и ARMA(2,1)

Пара-метр	Модели ARMA					
	(1,0)		(2,0)		(2,1)	
	Значение	Значи-мость	Значение	Значи-мость	Значение	Значи-мость
α_0	-10,553	0,844	-16,63	0,763	-14,17	0,548
α_1	1,0049	0,000	0,930	0,000	1,548	0,025
α_2			0,077	0,597	-0,541	0,426
β_1					-0,663	0,277
R^2	0,906		0,907		0,908	
$\Sigma\varepsilon^2$	80,82		81,41		81,69	
AIC	11,70		11,73		11,76	

Значимыми во всех моделях является только коэффициент при Y_{t-1}.

Значения критериев R^2, $\Sigma\varepsilon^2$, AIC отличаются незначительно, причем для модели ARMA(1,0) значения критериев $\Sigma\varepsilon^2$, AIC минимальны, что говорит об оптимальности модели ARMA(1,0).

Результаты расчетов по модели ARMA(1,0) показаны в таблице 6.1 (последняя графа) и на рис. 6.16.

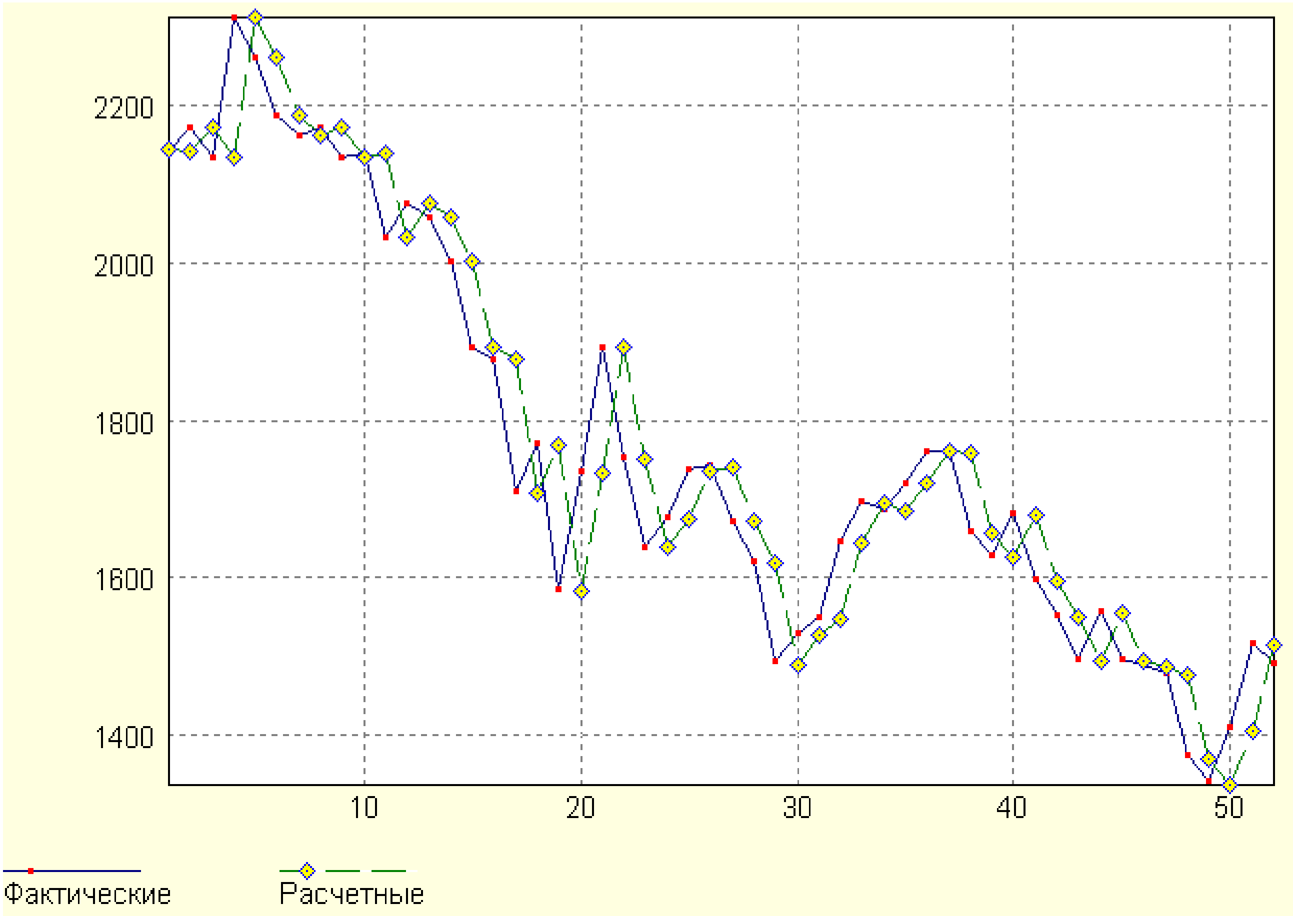


Рис. 6.16. Результаты расчетов по модели ARMA(1,0)

3) Расчет прогнозных значений по модели $Y_t = -10{,}553 + 1{,}0049 \cdot Y_{t-1} + \varepsilon_t$.

$$\hat{Y}_T(1) = -10{,}553 + 1{,}0049 \cdot Y_T = -10{,}553 + 1{,}0049 \cdot 1490 = 1487,$$

$$\hat{Y}_T(2) = -10{,}553 + 1{,}0049 \cdot \hat{Y}_T(1) = -10{,}553 + 1{,}0049 \cdot 1487 = 1483.$$

Результаты

1) Временной ряд Y_t является интегрируемым первого порядка нестационарным временным рядом.

2) Модель ARMA(1,0)

$$Y_t = -10{,}553 + 1{,}0049 \cdot Y_{t-1} + \varepsilon_t.$$

3) Прогнозные значения:

$$\hat{Y}_T(1) = -10{,}553 + 1{,}0049 \cdot Y_T = -10{,}553 + 1{,}0049 \cdot 1490 = 1487,$$

$$\hat{Y}_T(2) = -10{,}553 + 1{,}0049 \cdot \hat{Y}_T(1) = -10{,}553 + 1{,}0049 \cdot 1487 = 1483.$$

ПРИЛОЖЕНИЯ

Приложение 1. Таблицы исходных данных

Таблица П1.1

Исходные данные к лабораторной работе №1

	1	2	3	4	5	6	7	8
1	3	14	11	12	6	9	32	45
2	6	24	18	14	5	16	29	55
3	6	35	18	11	11	12	22	64
4	9	53	23	17	11	16	25	63
5	12	67	32	13	18	29	19	67
6	13	61	33	11	14	31	18	47
7	13	55	34	15	6	28	19	50
8	14	47	38	20	7	34	22	62
9	15	38	35	19	21	33	19	63
10	19	66	45	19	36	30	20	54
11	24	79	62	25	24	32	20	70
12	26	79	66	26	62	33	18	63
13	26	78	61	19	33	54	18	86
14	26	95	66	21	17	42	18	78
15	27	99	65	12	45	45	18	78
16	28	79	73	15	51	52	19	80
17	30	80	74	26	28	49	18	80
18	31	91	69	24	67	55	18	74
19	31	88	70	16	39	52	17	95
20	33	77	86	17	53	42	19	80
21	34	74	79	21	48	40	19	79
22	35	107	82	18	57	64	19	80
23	66	172	162	16	327	112	17	117
24	36	80	79	19	66	43	18	100
25	36	110	91	20	51	43	17	81
26	37	88	97	25	40	48	19	76
27	67	166	145	14	388	90	18	139
28	38	109	98	26	100	58	18	95
29	39	93	98	29	84	67	18	106
30	40	120	85	18	86	65	17	80
31	40	92	101	29	54	77	17	82
32	40	110	95	22	120	74	17	102
33	41	93	101	24	61	50	17	83
34	44	127	104	26	79	50	18	102
35	47	109	117	18	149	48	18	99
36	48	126	102	17	101	43	18	97
37	48	133	116	16	129	85	18	93
38	49	132	123	18	157	70	18	103
39	52	136	133	25	148	46	17	99
40	54	114	131	24	125	50	18	101

Таблица П1.2

Исходные данные к лабораторным работам №2, 3

Факторные переменные												
1	2	3	4	5	6	7	8	9	10	11	12	13
0,16	0,11	2,40	0,16	14,99	0,80	0,57	12,01	0,81	74,96	3,35	2,39	50,46
0,80	0,19	5,44	0,24	14,72	4,01	0,96	27,20	1,20	73,61	16,83	4,02	114,26
0,94	0,20	5,87	0,32	4,55	4,72	1,00	29,33	1,58	22,77	19,80	4,19	123,17
0,26	0,15	9,65	0,48	11,57	1,29	0,75	48,27	2,39	57,86	5,44	3,13	202,72
0,27	0,05	8,11	0,13	1,64	1,36	0,23	40,53	0,66	8,18	5,73	0,99	170,22
0,47	0,05	8,23	0,10	16,75	2,35	0,25	41,17	0,50	83,73	9,87	1,04	172,90
0,34	0,16	3,89	0,22	17,56	1,71	0,80	19,47	1,12	87,81	7,17	3,35	81,79
0,31	0,06	7,97	0,13	18,92	1,55	0,28	39,87	0,66	94,60	6,53	1,18	167,45
0,65	0,12	0,97	0,11	6,11	3,27	0,61	4,83	0,53	30,57	13,74	2,55	20,29
0,06	0,08	9,05	0,30	14,92	0,31	0,40	45,24	1,48	74,62	1,28	1,66	190,01
0,14	0,05	9,20	0,18	19,72	0,70	0,25	46,02	0,90	98,58	2,95	1,05	193,29
0,10	0,09	5,00	0,14	19,90	0,52	0,46	25,01	0,68	99,52	2,20	1,94	105,02
0,81	0,15	5,69	0,24	9,36	4,04	0,75	28,47	1,18	46,81	16,95	3,14	119,58
0,19	0,11	7,32	0,36	3,61	0,93	0,53	36,59	1,81	18,07	3,89	2,22	153,66
0,39	0,05	2,99	0,17	1,02	1,97	0,26	14,93	0,85	5,08	8,26	1,10	62,71
0,91	0,11	3,05	0,19	3,11	4,55	0,55	15,27	0,94	15,53	19,12	2,32	64,15
0,64	0,02	4,17	0,48	1,89	3,20	0,11	20,87	2,39	9,45	13,44	0,46	87,66
0,87	0,01	9,54	0,03	9,53	4,33	0,06	47,72	0,16	47,64	18,17	0,24	200,41
0,33	0,14	6,88	0,15	0,01	1,64	0,69	34,41	0,76	0,03	6,90	2,92	144,51
0,92	0,16	8,51	0,04	11,08	4,62	0,78	42,57	0,20	55,39	19,39	3,26	178,80
0,49	0,11	0,94	0,09	2,98	2,44	0,53	4,69	0,45	14,92	10,24	2,24	19,70
0,17	0,01	7,51	0,10	1,88	0,83	0,03	37,56	0,52	9,39	3,48	0,14	157,75
0,47	0,07	0,81	0,17	7,65	2,33	0,35	4,06	0,84	38,25	9,80	1,45	17,06
0,79	0,15	5,16	0,44	0,02	3,97	0,75	25,80	2,18	0,08	16,67	3,17	108,35
0,22	0,08	6,21	0,33	4,25	1,08	0,38	31,06	1,66	21,25	4,53	1,60	130,45
0,39	0,04	9,38	0,47	0,60	1,95	0,21	46,88	2,34	3,01	8,20	0,86	196,90
0,57	0,04	4,28	0,10	2,30	2,84	0,22	21,38	0,52	11,51	11,94	0,93	89,80
0,24	0,05	3,42	0,30	10,11	1,20	0,24	17,10	1,52	50,53	5,03	1,00	71,83
0,08	0,20	3,90	0,06	0,10	0,39	1,00	19,52	0,28	0,48	1,64	4,19	81,98
0,53	0,08	4,38	0,11	17,98	2,66	0,40	21,90	0,56	89,90	11,18	1,69	91,97
0,24	0,11	5,30	0,28	1,34	1,19	0,55	26,51	1,41	6,70	5,01	2,29	111,35
0,12	0,13	1,63	0,39	6,40	0,58	0,65	8,15	1,95	32,01	2,45	2,74	34,25
0,76	0,09	5,71	0,47	1,86	3,78	0,44	28,53	2,37	9,32	15,87	1,85	119,84
0,61	0,01	7,65	0,45	3,49	3,07	0,03	38,25	2,24	17,47	12,90	0,12	160,64
0,85	0,13	0,82	0,41	15,02	4,23	0,63	4,10	2,05	75,10	17,78	2,63	17,24
0,39	0,20	4,50	0,38	10,15	1,95	0,99	22,50	1,89	50,76	8,17	4,18	94,51
0,23	0,17	1,17	0,09	14,31	1,17	0,85	5,87	0,47	71,55	4,90	3,58	24,66
0,77	0,10	5,71	0,28	6,39	3,85	0,51	28,53	1,42	31,96	16,15	2,15	119,82
0,29	0,15	8,93	0,48	11,19	1,46	0,76	44,64	2,38	55,95	6,13	3,20	187,48
0,99	0,15	1,63	0,12	0,30	4,97	0,76	8,17	0,59	1,51	20,89	3,19	34,32
0,83	0,17	8,58	0,08	17,06	4,17	0,86	42,91	0,42	85,30	17,50	3,61	180,22
0,44	0,13	4,19	0,46	1,50	2,19	0,65	20,96	2,28	7,50	9,20	2,73	88,02
0,03	0,16	6,48	0,34	12,22	0,14	0,81	32,40	1,69	61,08	0,60	3,38	136,08
0,07	0,11	2,37	0,34	5,36	0,33	0,57	11,85	1,71	26,79	1,41	2,40	49,77
0,75	0,16	2,80	0,10	3,24	3,74	0,80	14,00	0,50	16,22	15,69	3,36	58,78

Окончание таблицы П1.2

Переменная y					
14	15	16	17	18	19
21,1	15,0	-0,6	20,7	23,0	12,7
20,6	10,3	43,0	20,5	24,4	45,7
20,7	4,5	49,2	20,6	25,7	51,0
22,1	-0,5	51,7	21,0	28,5	101,7
20,2	-4,5	27,6	19,9	20,3	49,5
19,8	2,9	41,7	19,6	19,4	62,4
21,3	14,2	26,2	20,8	24,8	34,8
20,2	4,6	23,7	20,0	20,4	41,4
20,0	13,6	24,2	19,9	20,6	5,9
21,4	0,9	36,8	20,6	23,4	80,1
20,7	2,4	42,4	20,2	21,1	83,9
20,9	11,8	32,0	20,6	22,3	51,6
20,3	6,3	36,4	20,2	22,7	38,9
21,6	-0,7	17,4	20,7	24,8	42,5
20,1	5,6	33,0	19,8	20,2	30,8
19,6	7,7	45,1	19,6	20,2	29,7
20,3	3,0	31,6	19,8	18,8	38,6
18,5	-4,1	70,8	18,4	16,8	94,3
20,9	-1,0	26,3	20,6	23,0	44,8
19,5	1,6	71,4	19,5	19,4	88,9
20,1	11,8	23,8	20,0	20,7	6,2
20,0	-3,9	34,6	19,8	19,5	64,1
20,1	13,6	4,7	19,9	20,4	3,6
20,9	2,7	25,3	20,7	25,8	30,4
21,2	1,2	31,1	20,4	22,9	57,3
21,0	-7,6	46,4	20,1	21,0	83,4
19,5	3,5	19,9	19,4	18,9	23,3
20,8	9,2	27,5	20,1	21,2	36,4
21,6	6,2	1,0	21,5	24,0	18,0
19,9	11,8	19,6	19,8	20,1	24,5
21,2	2,2	30,2	20,6	23,9	50,3
22,0	12,5	1,4	20,9	26,7	14,7
20,6	1,3	40,0	20,3	22,6	49,6
20,2	-3,4	54,4	19,6	17,9	81,0
20,5	18,4	29,6	20,4	23,8	10,7
21,9	10,1	22,8	21,3	29,1	36,9
21,2	18,2	-2,8	21,0	23,4	4,0
20,1	3,8	50,4	19,9	21,5	57,1
22,1	0,8	49,5	21,0	28,5	93,8
19,6	9,9	37,4	19,6	20,2	12,4
20,0	4,8	44,1	19,9	20,9	47,8
21,5	5,0	28,8	20,8	26,2	40,9
22,2	6,4	18,1	21,3	27,9	48,8
21,8	10,2	20,0	20,8	25,5	27,2
20,1	9,2	14,5	20,0	21,2	11,5

Таблица П1.3

Исходные данные к лабораторным работам №4, 5

Текущий период t	Год	ВВП Y (млрд руб.)	Денежная масса *M* (млрд руб.)	Внутренние инвестиции I (млрд руб.)	Национальный доход Y (млрд руб.)	Расходы на личное потребление C* (млрд руб.)	Валовая прибыль экономики Q (млрд руб.)
1		3	4	5	6	7	8
1	1995	1428,5	98,7	267,0	1412,7	872	610,8
2	1996	2007,8	220,8	376,0	1978,9	1313	699,4
3	1997	2342,5	288,3	408,8	2292,0	1597	783,3
4	1998	2629,6	374,1	407,1	2514,4	1811	948,7
5	1999	4823,2	453,7	670,4	4632,0	2870	2131,5
6	2000	7305,6	714,6	1165,2	7116,6	3813	3119,9
7	2001	8943,6	1154,4	1504,7	8819,9	5014	3692,6
8	2002	10834,2	1612,6	1762,4	10627,5	6390	3919,7
9	2003	13285,2	2134,5	2186,4	12886,1	7710	4901,6
10	2004	17048,1	3212,6	2865,0	16679,9	9814	6330,4
11	2005	21620,1	4363,3	3611,1	21079,5	12391	7908,1
12	2006	26781,1	6044,7	4730	26078	15147	9659,3
13	2007	32987,4	8995,8	6627	32213	18644	11702
14	2008	41428,6	13493	8782	40222	27423	13696
15	2009	39100,7	15698	7930	37862	29191	12440

Окончание таблицы П1.3

Текущий период t	Индекс стоимости жизни P (%)	Объем продукции промышленности *R* (млрд руб.)	Государственные расходы G (млрд руб.)	Доля импорта в ВВП M	Реальный объем чистого экспорта X (млрд руб.)	Налоги T (млрд руб.)	Запас капитала K (млрд руб.)	Зарплата S (тыс.руб.)
1	9	10	11	12	13	14	15	16
1	195	1108,4	486,1	0,243	418,5	364,3	301,1	472,4
2	208	1468,8	652,7	0,206	523,5	473	401,6	790,2
3	229	1626,5	839,0	0,209	579,3	594,1	428,5	950,2
4	204	1706,6	842,1	0,235	821,0	564,6	424,7	1051,5
5	180	3359,9	1258,0	0,269	2085	1007,6	694,0	1522,6
6	181	5174,1	1960,1	0,240	3219	1707,6	1232,0	2223,4
7	205	6329,7	2419,4	0,242	3300	2345	1689,3	3240,4
8	220	7294,8	3422,3	0,244	3790	3137	1943,4	4360,3
9	244	8755,2	3964,9	0,237	4656	3735	2417,7	5498,3
10	270	11994	4669,7	0,221	5860	4942	3130,5	6739.5
11	269	15609	6820,6	0,216	7592	5949	3848,4	8555
12	298	19258	8375,2	0,212	9069	7162	4968	10634
13	328	24400	11377	0,210	9667	9469	6951	13593
14	325	24710	13992	0,207	13855	12008	9172	17290
15	327	22493	16048	0,168	9174	10430	8384	18637

Таблица П1.4

Исходные данные к лабораторной работе №6

	Варианты							
	1	2	3	4	5	6	7	8
1	72,2	105,0	4171	4134	2643	3199	2265	2394
2	68,8	106,1	4198	4083	2556	3224	2309	2371
3	69,0	108,2	4339	4091	2587	3289	2365	2394
4	67,4	111,3	4338	4100	2628	3286	2346	2399
5	70,1	111,7	4345	4168	2730	3292	2371	2364
6	70,8	111,0	4342	4092	2749	3279	2387	2363
7	72,4	110,8	4344	4104	2719	3283	2346	2391
8	71,5	111,4	4339	4140	2652	3278	2349	2391
9	72,8	110,6	4283	4216	2679	3249	2326	2403
10	76,7	110,8	4306	4163	2763	3198	2326	2348
11	77,1	108,6	4251	4137	2741	3185	2285	2316
12	77,0	109,9	4254	4121	2705	3220	2289	2241
13	77,3	108,7	4284	4015	2741	3257	2313	2300
14	79,2	109,8	4397	4000	2781	3271	2348	2268
15	80,7	112,9	4373	4085	2792	3289	2398	2236
16	81,9	115,6	4446	3961	2690	3327	2394	2198
17	79,8	116,3	4423	4011	2743	3341	2397	2241
18	77,3	117,1	4380	4094	2732	3363	2393	2196
19	77,8	118,6	4390	4099	2758	3381	2410	2143
20	80,1	121,1	4429	3960	2784	3409	2425	2174
21	81,9	124,2	4457	4007	2741	3394	2423	2136
22	80,4	122,9	4465	3844	2744	3403	2418	2312
23	78,4	123,8	4465	3991	2807	3386	2433	2262
24	79,3	121,8	4450	4089	2814	3400	2416	2188
25	80,9	122,9	4543	3939	2839	3370	2497	2163
26	82,3	124,3	4551	3889	2939	3346	2475	2172
27	83,9	123,8	4518	3865	2925	3321	2453	2135
28	86,3	123,1	4546	3930	2957	3361	2484	2139
29	92,1	123,9	4454	3943	2919	3345	2451	2034
30	91,9	124,4	4471	3711	2965	3328	2430	2077
31	98,0	124,4	4442	3828	2900	3257	2451	2059
32	96,4	125,5	4490	3816	2892	3244	2492	2003
33	93,3	122,6	4486	3627	2805	3288	2506	1894
34	93,6	123,2	4518	3580	2820	3320	2497	1879
35	93,5	126,9	4550	3471	2800	3321	2544	1710
36	97,1	128,4	4545	3307	2866	3355	2532	1770
37	101,7	132,3	4569	3261	2869	3375	2501	1585
38	100,1	136,9	4527	3040	2850	3389	2495	1735
39	101,8	140,1	4457	3008	2790	3380	2453	1894
40	101,1	140,9	4464	3326	2787	3446	2456	1753
41	98,7	144,2	4395	3294	2867	3457	2448	1640
42	101,3	143,8	4443	3020	2805	3444	2493	1677
43	103,8	141,8	4469	3153	2844	3474	2486	1737
44	105,7	143,2	4510	3114	2898	3457	2518	1742
45	112,9	145,3	4514	3278	2996	3414	2516	1673
46	117,3	144,6	4465	3190	2995	3441	2502	1620
47	114,0	144,9	4435	3020	2952	3375	2470	1493
48	110,8	146,8	4453	3070	2984	3341	2510	1529
49	115,8	146,9	4527	2964	2988	3321	2529	1549
50	112,0	142,9	4374	2888	2983	3276	2483	1646

Продолжение табл. П1.4

	Варианты							
	9	10	11	12	13	14	15	16
1	721	987	12,60	10,370	3,805	4,880	3,119	2,590
2	683	1019	12,67	10,000	3,750	4,985	3,120	2,550
3	690	1051	12,83	10,100	3,700	5,360	3,250	2,560
4	686	1094	12,80	10,330	4,020	5,300	3,365	2,570
5	734	1103	12,65	9,710	4,150	5,270	3,274	2,400
6	747	1072	12,78	9,300	4,340	5,240	3,320	2,300
7	749	1070	12,88	9,580	4,170	5,290	3,340	2,300
8	740	1089	12,84	9,560	4,000	5,260	3,280	2,400
9	761	1067	13,15	9,750	4,250	5,230	3,340	2,320
10	811	1073	13,43	9,700	4,330	5,100	3,300	2,350
11	817	1053	12,94	9,900	4,370	5,045	3,120	2,450
12	815	1074	13,18	9,810	4,350	5,080	3,160	2,320
13	807	1064	13,30	9,230	4,250	5,070	3,160	2,150
14	807	1093	13,49	8,740	4,270	5,280	3,220	2,300
15	820	1136	13,32	9,200	4,331	5,325	3,170	2,195
16	835	1160	13,58	8,510	4,250	5,630	3,274	1,960
17	796	1164	13,45	8,100	3,890	5,695	3,210	1,920
18	775	1197	13,43	7,920	3,950	5,570	3,210	1,770
19	785	1194	13,42	7,830	4,180	5,699	3,150	1,770
20	821	1225	13,23	8,000	4,400	5,920	3,160	1,730
21	831	1247	13,35	7,400	4,400	6,050	3,195	1,475
22	803	1231	13,20	6,600	4,250	5,920	3,230	1,600
23	784	1246	13,26	7,000	4,340	6,000	3,215	1,750
24	815	1211	13,22	7,650	4,580	6,080	3,298	1,820
25	833	1249	13,25	8,500	4,550	6,080	3,380	1,640
26	856	1254	13,26	8,200	4,700	5,920	3,425	1,720
27	872	1242	13,40	8,130	4,930	5,930	3,400	1,840
28	898	1225	13,83	8,530	5,380	6,000	3,410	1,810
29	942	1249	14,90	8,400	5,330	6,040	3,430	1,670
30	938	1261	14,74	8,200	5,480	5,960	3,540	1,450
31	980	1255	14,64	7,250	5,730	5,930	3,470	1,720
32	948	1267	15,02	7,900	5,210	5,690	3,520	1,610
33	933	1225	14,97	7,620	5,240	5,750	3,540	1,580
34	936	1227	15,42	7,230	5,050	5,850	3,630	1,360
35	940	1271	15,40	6,320	5,500	5,950	3,670	1,320
36	971	1291	15,29	5,650	5,600	6,000	3,650	1,350
37	1024	1335	15,45	4,750	5,670	6,130	3,680	1,200
38	1001	1374	15,11	4,900	5,565	6,300	3,630	1,230
39	1014	1428	15,35	5,000	5,500	6,550	3,620	1,270
40	1018	1400	15,40	5,100	5,400	6,710	3,600	1,250
41	991	1441	14,90	5,120	5,515	6,880	3,610	1,250
42	1031	1434	15,19	4,950	5,450	6,740	3,580	1,020
43	1054	1409	15,16	4,180	5,580	6,600	3,590	1,060
44	1089	1433	15,10	4,250	6,020	6,750	3,550	1,050
45	1167	1446	14,93	4,100	6,195	6,700	3,510	1,025
46	1181	1447	14,81	4,350	6,200	6,510	3,475	0,960
47	1125	1448	14,28	4,150	5,780	6,680	3,465	0,920
48	1109	1461	14,60	3,970	5,910	6,800	3,480	0,760
49	1150	1474	14,47	3,540	5,700	6,630	3,345	0,700
50	1097	1421	14,43	3,150	5,745	6,150	3,360	0,650

Окончание табл. П1.4

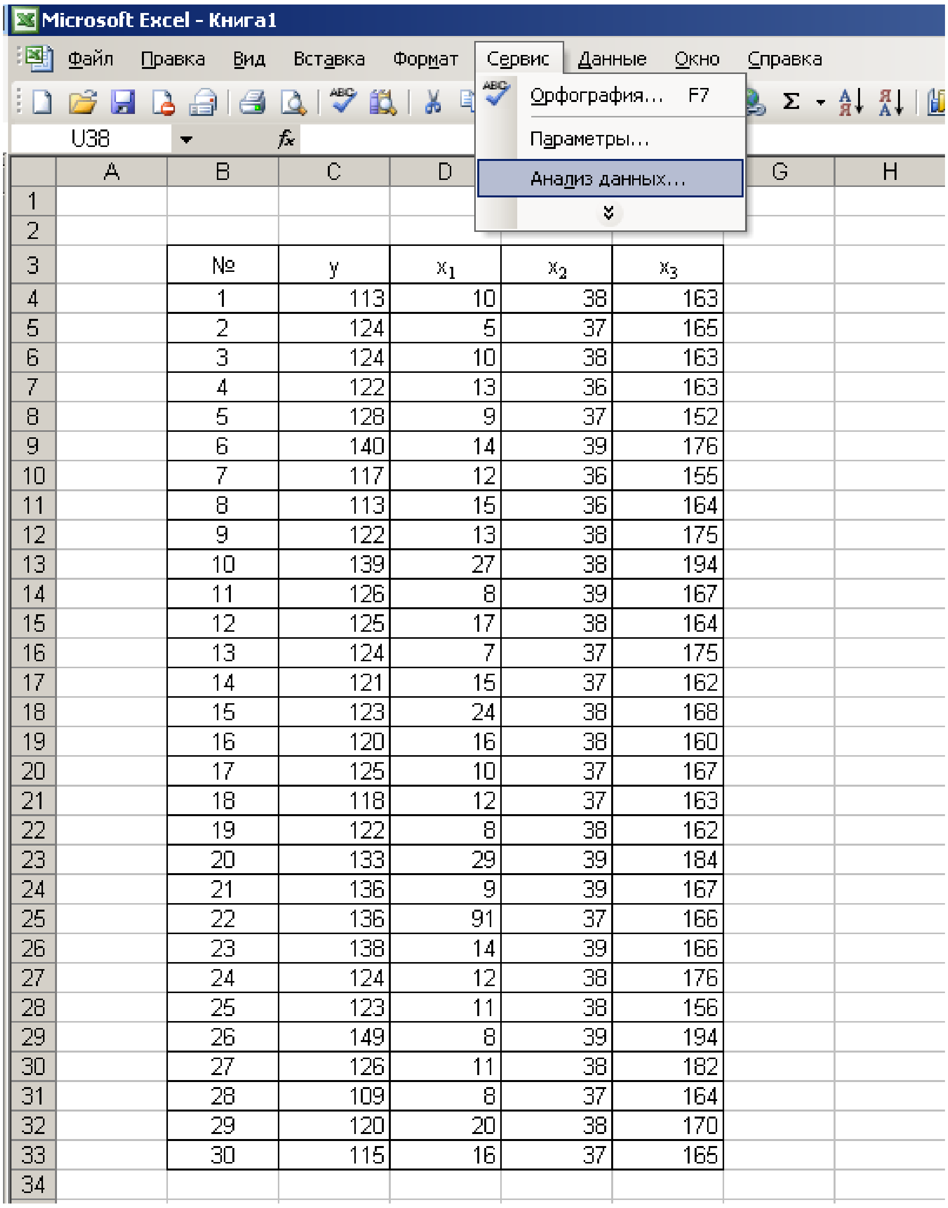
	Варианты								
	17	18	19	20	21	22	23	24	25
1	0,610	1,417	9,030	8,350	4,440	5,614	9,230	0,635	4,400
2	0,613	1,488	8,880	8,380	4,380	5,773	9,130	0,623	4,570
3	0,620	1,445	9,100	8,050	4,285	5,880	9,120	0,610	4,425
4	0,645	1,560	9,150	7,570	4,650	6,360	9,100	0,540	4,250
5	0,690	1,540	9,060	7,970	4,920	6,410	9,070	0,500	4,420
6	0,727	1,565	9,100	8,000	5,300	6,505	8,880	0,555	4,600
7	0,709	1,523	9,340	8,360	5,161	6,423	8,594	0,470	4,630
8	0,673	1,530	9,240	8,580	4,839	6,373	8,280	0,455	4,900
9	0,750	1,497	9,470	8,730	5,290	6,218	8,050	0,500	5,150
10	0,795	1,510	9,640	8,630	5,400	6,420	7,770	0,500	5,400
11	0,810	1,530	9,370	8,190	5,430	6,140	7,890	0,498	5,350
12	0,820	1,556	9,650	7,610	5,380	6,310	7,750	0,460	5,580
13	0,845	1,720	9,850	8,180	5,150	6,390	7,530	0,448	5,730
14	0,850	1,750	9,850	8,000	5,177		7,918	0,430	5,210
15	0,885	1,795	9,840	7,400	5,373	6,544	8,190	0,427	5,240
16	0,865	1,880	9,940	7,620	5,360	6,700	7,879	0,412	5,250
17	0,778	1,890	9,820	7,250	4,999	6,815	7,400	0,414	5,540
18	0,783	1,840	9,950	7,150	5,070	6,850	7,380	0,445	5,680
19	0,800	1,880	9,870	6,330	5,325	6,868	7,200	0,465	5,680
20	0,835	1,940	9,740	5,500	5,370	7,050	7,347	0,490	5,570
21	0,830	2,057	9,870	6,000	5,100	7,430	7,450	0,549	5,500
22	0,768	2,055	9,730	6,650	4,810	7,550	7,514	0,553	5,400
23	0,788	1,900	9,690	7,500	4,971	7,450	7,480	0,530	5,570
24	0,838	1,885	9,720	7,450	5,370	7,450	7,640	0,505	5,610
25	0,874	1,850	9,650	8,000	5,450	7,545	7,800	0,425	5,780
26	0,885	1,980	9,810	7,800	5,500	7,580	7,650	0,414	6,190
27	0,885	1,980	9,760	7,400	5,400	7,605	7,550	0,400	6,280
28	0,930	1,930	9,840	7,300	5,750	7,433	7,770	0,395	6,200
29	0,975	1,940	10,700	6,350	5,680	7,300	7,600	0,415	5,800
30	1,000	2,030	10,990	6,750	5,620	7,610	7,605	0,404	6,050
31	1,090	2,030	11,170	6,380	5,970	7,610	7,700	0,397	5,750
32	0,975	2,013	11,790	5,800	5,483	7,610	7,800	0,380	5,745
33	1,020	1,940	11,650	5,200	5,500	7,390	7,800	0,402	5,850
34	0,990	1,940	11,620	4,200	5,280	7,347	7,940	0,435	5,970
35	1,080	1,992	11,250	3,500	5,750	7,680	8,000	0,445	5,700
36	1,120	2,050	11,280	4,300	5,700	7,842	8,000	0,439	5,710
37	1,241	2,120	11,340	4,000	5,950	7,950	8,110	0,465	5,610
38	1,225	2,260	11,200	3,800	5,892	8,130	8,069	0,495	5,200
39	1,280	2,353	11,470	4,200	5,880	8,500	8,066	0,490	5,180
40	1,245	2,417	11,850	3,800	5,860	8,520	8,150	0,500	5,000
41	1,300	2,400	11,450	3,500	6,009	8,430	8,080	0,495	4,910
42	1,305	2,400	11,900	3,850	6,100	8,370	8,040	0,525	4,850
43	1,400	2,335	11,910	3,750	6,550	8,490	8,130	0,630	4,850
44	1,550	2,260	12,100	4,000	6,999	8,370	7,950	0,650	4,970
45	1,660	2,410	12,060	3,500	7,000	8,550	7,920	0,710	4,905
46	1,765	2,375	11,900	3,600	7,240	8,380	7,740	0,744	5,260
47	1,565	2,420	11,450	3,300	6,570	8,340	7,700	0,795	5,220
48	1,700	2,447	11,900	3,000	6,750	8,620	7,610	0,675	5,140
49	1,620	2,430	11,850	2,790	6,317	8,453	7,590	0,760	5,035
50	1,550	2,355	11,860	3,250	6,500	8,228	7,650	0,725	4,680

Приложение 2. Использование возможностей MS Excel для проведения корреляционного и регрессионного анализа

Корреляционный анализ

Рассмотрим построение корреляционной матрицы (матрицы парных корреляций) $\{r_{x_i x_j}\}$по данным наблюдений за совместным изменением n переменных x_j (табл. 3.1).

Расположим исходные данные в ячейках с C4 по F33 (рис. П2.1) и вызовем функцию «Сервис.Анализ данных.Корреляция» табличного процессора MS Excel (рис. П2.1, П2.2).



№	y	x_1	x_2	x_3
1	113	10	38	163
2	124	5	37	165
3	124	10	38	163
4	122	13	36	163
5	128	9	37	152
6	140	14	39	176
7	117	12	36	155
8	113	15	36	164
9	122	13	38	175
10	139	27	38	194
11	126	8	39	167
12	125	17	38	164
13	124	7	37	175
14	121	15	37	162
15	123	24	38	168
16	120	16	38	160
17	125	10	37	167
18	118	12	37	163
19	122	8	38	162
20	133	29	39	184
21	136	9	39	167
22	136	91	37	166
23	138	14	39	166
24	124	12	38	176
25	123	11	38	156
26	149	8	39	194
27	126	11	38	182
28	109	8	37	164
29	120	20	38	170
30	115	16	37	165

Рис. П2.1. Вызов функции «Сервис.Анализ данных»

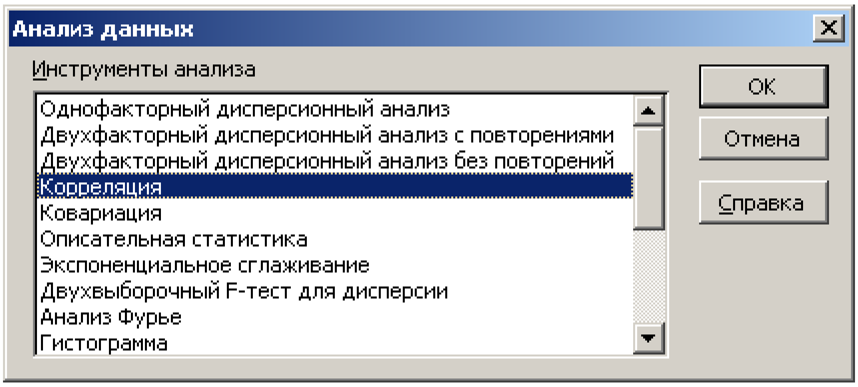


Рис. П2.2. Вызов функции «Сервис.Анализ данных.Корреляция»

В окне ввода параметров функции «Сервис.Анализ данных.Корреляция» (рис. П2.3) необходимо указать диапазон ячеек, содержащих исходные данные («Входной интервал»), и диапазон ячеек, в которых будет располагаться полученная корреляционная матрица («Выходной интервал»).

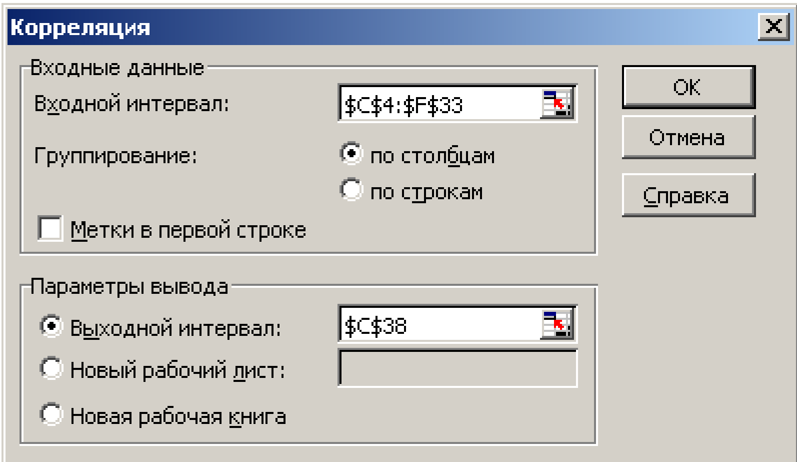


Рис. П2.3. Окно ввода параметров функции «Сервис.Анализ данных.Корреляция»

В области ячеек, начиная с указанной ячейки C38 получим искомую матрицу (рис. П2.4):

	Столбец 1	Столбец 2	Столбец 3	Столбец 4
Столбец 1	1			
Столбец 2	0,263	1		
Столбец 3	0,605	–0,071	1	
Столбец 4	0,599	0,091	0,471	1

Рис. П2.4. Корреляционная матрица

Регрессионный анализ

Рассмотрим построение уравнения линейной множественной регрессии по данным наблюдений за совместным изменением p+1 переменной y *и* x_j и ((y_i, $x_{j,i}$); j=1, 2, ..., p; i=1, 2, ..., n) (табл. 3.1).

Будем считать, что имеется три факторные переменные (p = 3) и число наблюдений равно 30. Расположим исходные данные в ячейках с C4 по F33 и вызовем функцию «Сервис.Анализ данных.Регрессия» табличного процессора MS Excel (рис.П2.1, П2.5), в результате чего на экране появится окно ввода параметров данной функции (рис. П2.6).

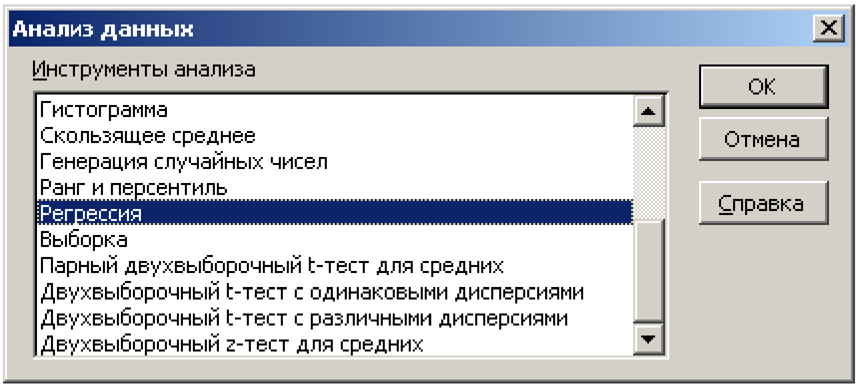


Рис. П2.5. Вызов функции «Сервис.Анализ данных. Регрессия»

В окне ввода параметров функции «Сервис.Анализ данных. Регрессия» (рис. П2.6) необходимо указать диапазон ячеек, содержащих исходные данные («Входной интервал по Y», «Входной интервал по X»), и место, где будут располагаться результаты: диапазон ячеек на данном рабочем листе, новый рабочий лист, новая рабочая книга («Выходной интервал»). Если требуется получить уравнение регрессии без свободного члена, то нужно установить флажок «Константа–ноль» (рис. П2.6);

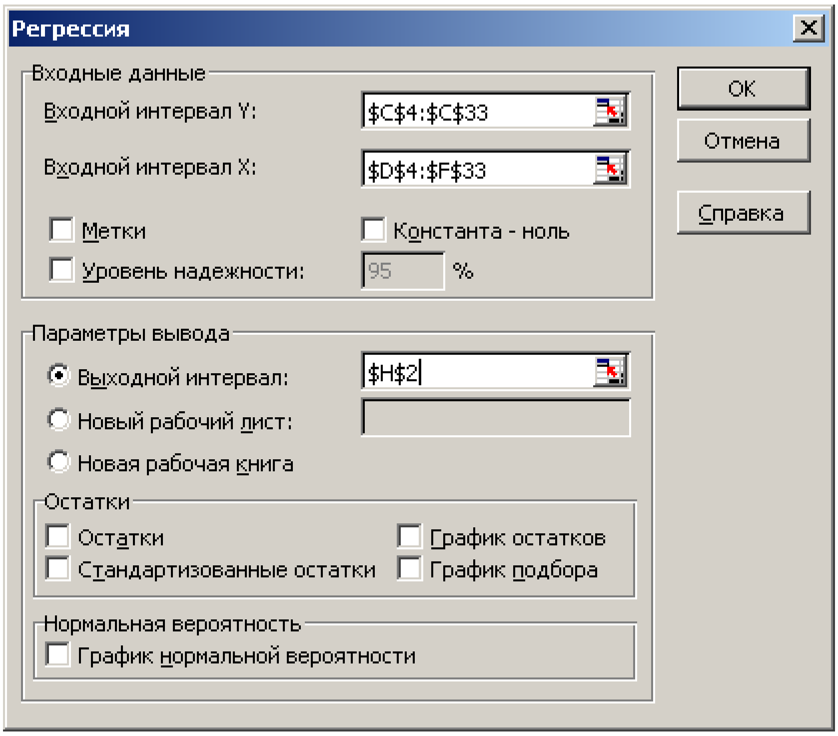
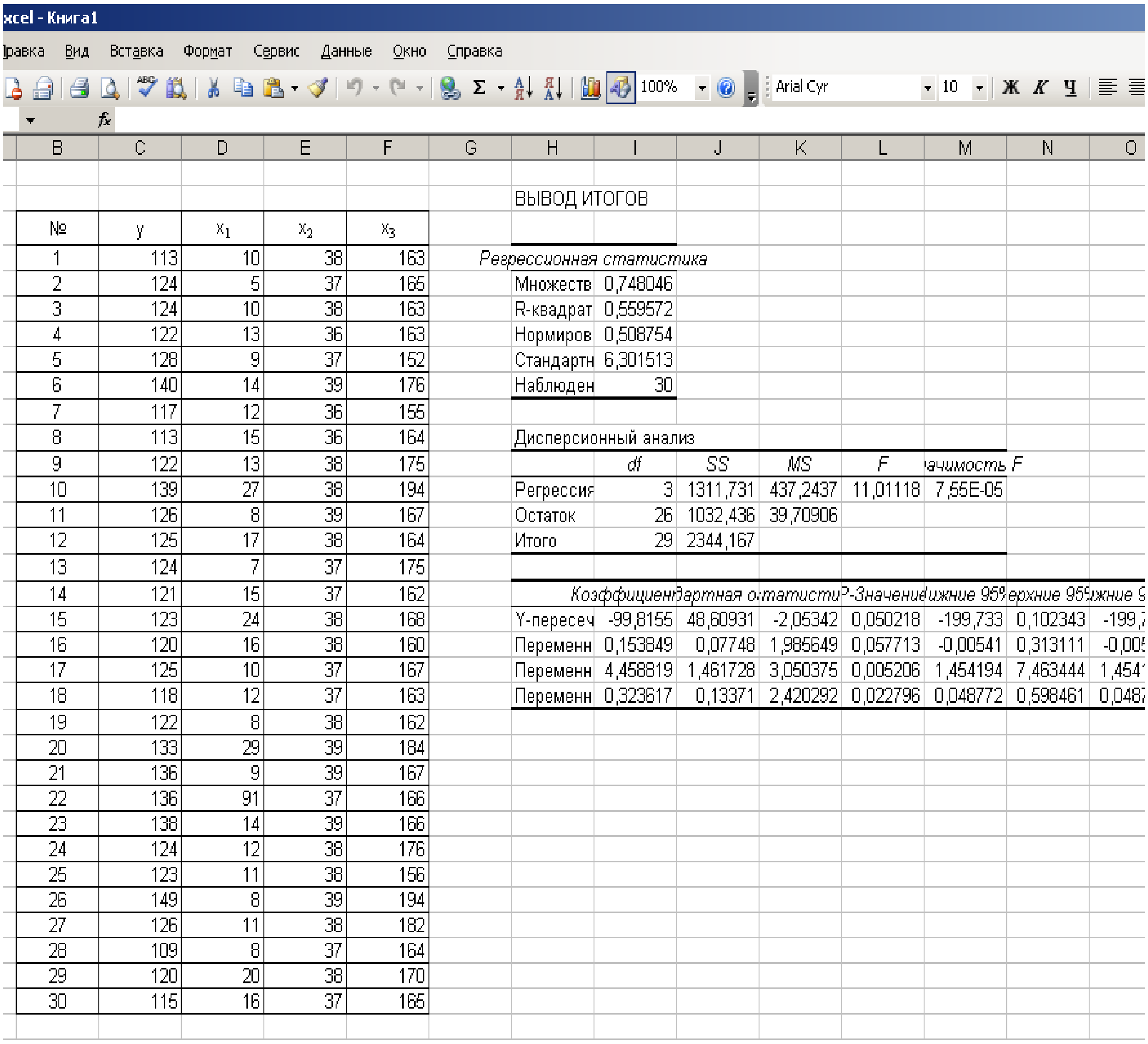


Рис. П2.6. Окно ввода параметров функции «Сервис.Анализ данных. Регрессия»

После выполнения функции «Сервис.Анализ данных. Регрессия» рабочий лист Excel примет вид (рис. П2.7).



№	y	x_1	x_2	x_3
1	113	10	38	163
2	124	5	37	165
3	124	10	38	163
4	122	13	36	163
5	128	9	37	152
6	140	14	39	176
7	117	12	36	155
8	113	15	36	164
9	122	13	38	175
10	139	27	38	194
11	126	8	39	167
12	125	17	38	164
13	124	7	37	175
14	121	15	37	162
15	123	24	38	168
16	120	16	38	160
17	125	10	37	167
18	118	12	37	163
19	122	8	38	162
20	133	29	39	184
21	136	9	39	167
22	136	91	37	166
23	138	14	39	166
24	124	12	38	176
25	123	11	38	156
26	149	8	39	194
27	126	11	38	182
28	109	8	37	164
29	120	20	38	170
30	115	16	37	165

ВЫВОД ИТОГОВ

Регрессионная статистика	
Множеств	0,748046
R-квадрат	0,559572
Нормиров	0,508754
Стандартн	6,301513
Наблюден	30

Дисперсионный анализ

	df	SS	MS	F	ачимость F
Регрессия	3	1311,731	437,2437	11,01118	7,55E-05
Остаток	26	1032,436	39,70906		
Итого	29	2344,167			

	Коэффициент	дартная о	статисти	P-Значени	ижние 95%	ерхние 95%	жние 9
Y-пересеч	-99,8155	48,60931	-2,05342	0,050218	-199,733	0,102343	-199,7
Переменн	0,153849	0,07748	1,985649	0,057713	-0,00541	0,313111	-0,005
Переменн	4,458819	1,461728	3,050375	0,005206	1,454194	7,463444	1,454
Переменн	0,323617	0,13371	2,420292	0,022796	0,048772	0,598461	0,048

Рис. П2.7. Результаты регрессионного анализа – вызова функции «Сервис.Анализ данных. Регрессия»

Результаты регрессионного анализа представлены в виде трех таблиц

Таблица П2.1

Результаты корреляционного анализа («Регрессионная статистика»)

Множественный R	0,748	*Множественный коэффициент корреляции R*
R-квадрат	0,560	*Коэффициент детерминации* R^2
Нормированный R-квадрат	0,509	*Модифицированный коэффициент детерминации* $\overline{R}^2$ (У3.28)
Стандартная ошибка	6,302	*Стандартная ошибка определения R*
Наблюдения	30	*Число наблюдений*

Таблица П2.2

Результаты дисперсионного анализа

Пояснения	Число степеней свободы *df*	Сумма квадратов отклонений *SS*	Дисперсия на 1 степень свободы *MS*	Статистика Фишера *F*	Уровень значимости *Значимость F*
Регрессия	3	1311,7	437,2	11,011	7,55E-05
Остаток	26	1032,4	39,7		
Итого	29	2344,2			

Столбец «Сумма квадратов отклонений» содержит следующие суммы:

$$\text{Регрессия} = \frac{1}{n}\sum\left(\hat{y}_i - \bar{y}\right)^2; \quad \text{Остаток} = \frac{1}{n}\sum\left(\hat{y}_i - y_i\right)^2; \quad \text{Итого} = \frac{1}{n}\sum\left(y_i - \bar{y}\right)^2.$$

Столбцы «Статистика Фишера *F*» и «Уровень значимости» содержат фактическое значение критерия Фишера $F = 11{,}011$ и минимальный уровень значимости уравнения регрессии α_0, т. е. уравнение регрессии значимо при всех $\alpha > \alpha_0$.

Таблица П2.3

Результаты регрессионного анализа

Пояснения	Коэффициенты уравнения регрессии	Стандартная ошибка определения коэффициентов	t-статистика	Вероятность ошибки α	Нижние 95%-пределы	Верхние 95%-пределы
Показатели	Коэффициенты	Стандартная ошибка	t-статистика	P-значение	Нижние 95%	Верхние 95%
Y-пересечение	–99,816	48,6093	–2,0534	0,0502	–199,7334	0,1023
Переменная X 1	0,154	0,0775	1,9856	0,0577	–0,0054	0,3131
Переменная X 2	4,459	1,4617	3,0504	0,0052	1,4542	7,4634
Переменная X 3	0,324	0,1337	2,4203	0,0228	0,0488	0,5985

Искомые значения коэффициентов линейного уравнения регрессии (a, b_i) берутся из столбца «Коэффициенты» таблицы результатов регрессии (табл. 2.3), из которой следует, что уравнение регрессии имеет вид

$$\hat{y} = -99{,}816 + 0{,}154 \cdot x_1 + 4{,}459 \cdot x_2 + 0{,}324 \cdot x_3.$$

Столбец «Стандартная ошибка определения коэффициентов» содержит стандартные ошибки определения коэффициентов уравнения регрессии.

Столбец «t-статистика» содержит фактические значения критерия Стьюдента для соответствующего коэффициента.

Столбец «Вероятность ошибки» содержит минимальный уровень значимости коэффициента α_0.

Столбцы «Нижние 95%–пределы» и «Верхние 95%–пределы» содержат границы доверительных интервалов для значений коэффициентов. Разные знаки нижней и верхней границы доверительного интервала говорят о ненадежности полученного значения соответствующего коэффициента (свободный член и первый коэффициент в нашем примере).

Приложение 3. Использование возможностей пакета Matrixer 5.1 для проведения корреляционного и регрессионного анализа

Программа предназначена для анализа и обработки данных, проведения эконометрических и статистических расчетов. Основной тип объектов – матрицы (вектор рассматривается как одномерная матрица). Имеется возможность обмена данными с электронными таблицами Excel.

Главное окно программы имеет вид

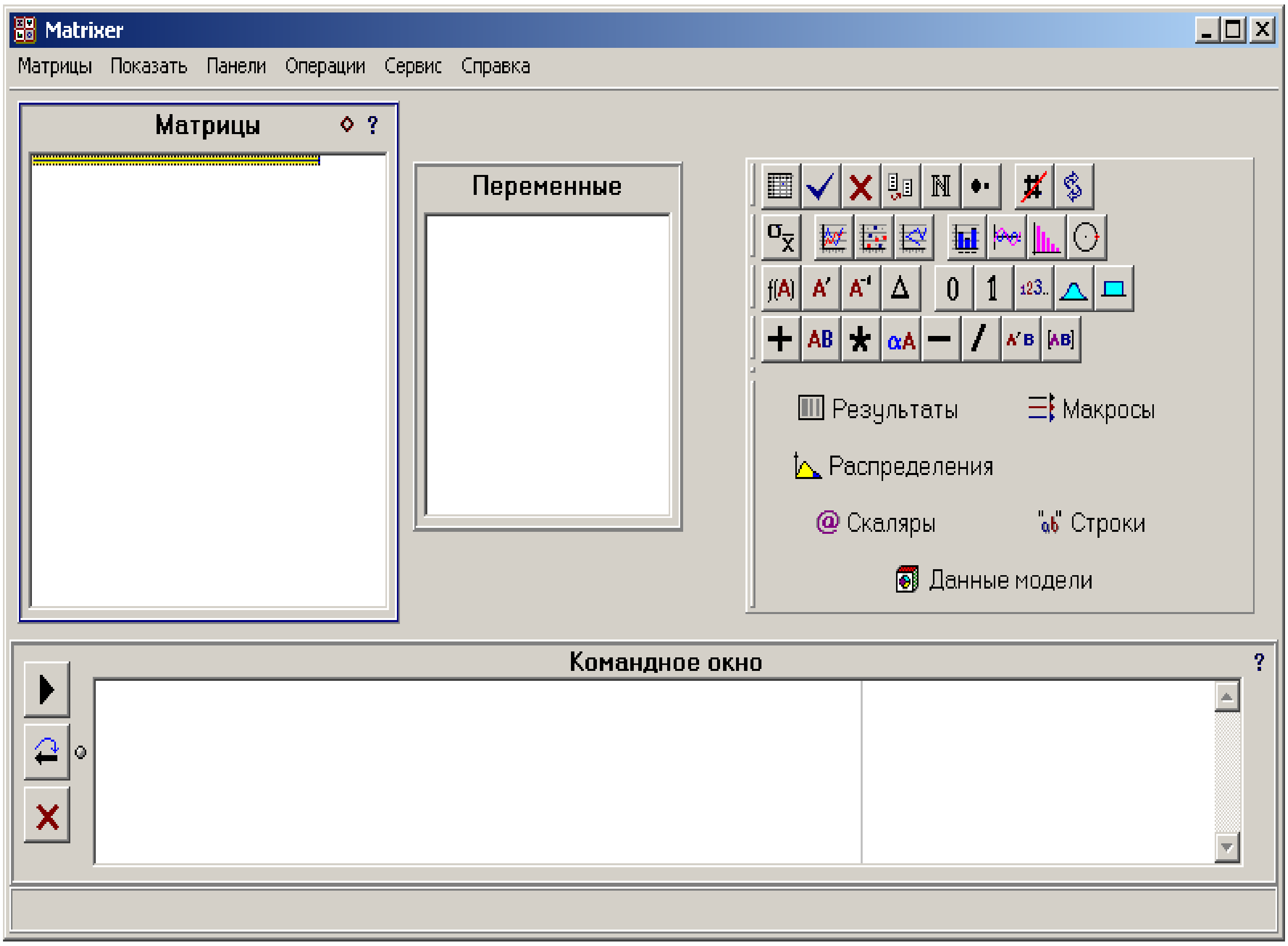


Рис. П3.1. Главное окно программы Matrixer 5.1

3.1. Корреляционный анализ

Рассмотрим построение корреляционной матрицы (матрицы парных корреляций) $\{r_{x_i x_j}\}$по данным наблюдений за совместным изменением n переменных x_j (табл. 3.1).

Создадим новый объект – матрицу X, столбцами которой являются столбцы данных наблюдений. Для этого вызовем пункт меню «Матрицы.Создать» либо нажмем кнопку .

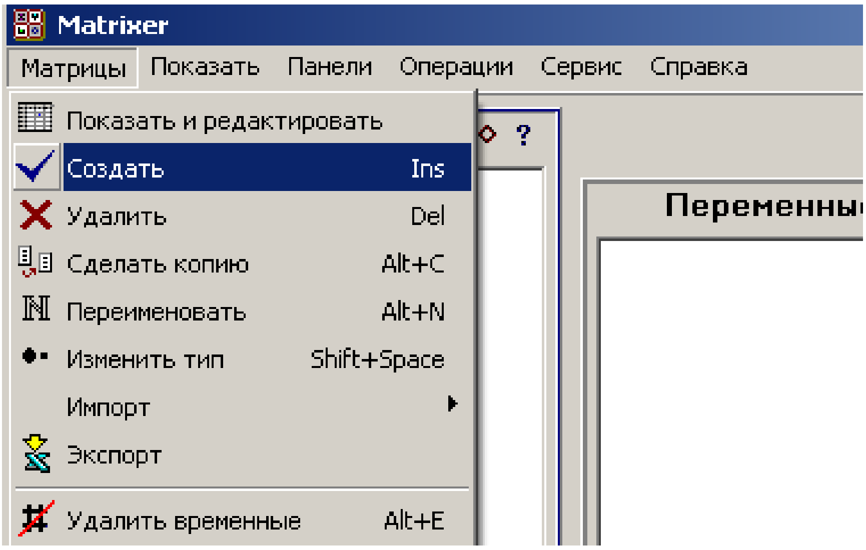


Рис. П3.2. Вызов пункта меню «Матрицы.Создать»

После задания имени (X) объекту на экране появится окно ввода данных (рис. П3.3)

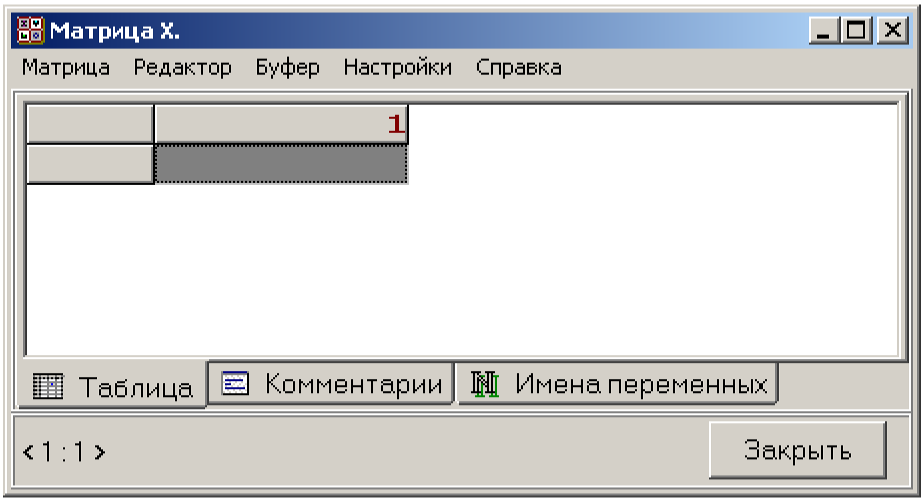
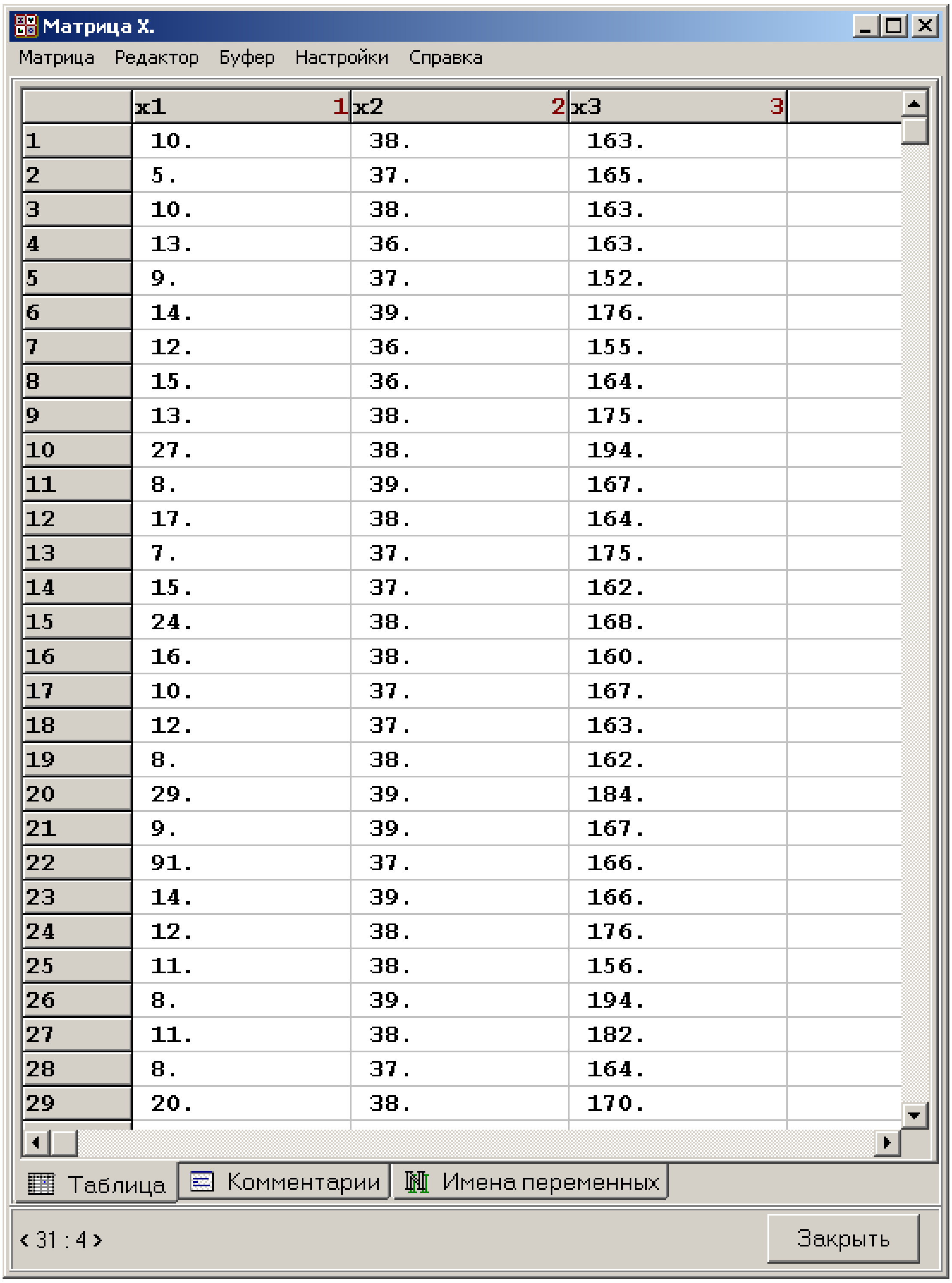


Рис. П3.4. Окно ввода данных

Ввод данных можно осуществлять с клавиатуры либо использовать копирование табличных данных через буфер обмена. После ввода данных окно данных матрицы X принимает вид



	x1 1	x2 2	x3 3
1	10.	38.	163.
2	5.	37.	165.
3	10.	38.	163.
4	13.	36.	163.
5	9.	37.	152.
6	14.	39.	176.
7	12.	36.	155.
8	15.	36.	164.
9	13.	38.	175.
10	27.	38.	194.
11	8.	39.	167.
12	17.	38.	164.
13	7.	37.	175.
14	15.	37.	162.
15	24.	38.	168.
16	16.	38.	160.
17	10.	37.	167.
18	12.	37.	163.
19	8.	38.	162.
20	29.	39.	184.
21	9.	39.	167.
22	91.	37.	166.
23	14.	39.	166.
24	12.	38.	176.
25	11.	38.	156.
26	8.	39.	194.
27	11.	38.	182.
28	8.	37.	164.
29	20.	38.	170.

Рис. П3.5. Окно данных матрицы X

Далее для получения корреляционной матрицы необходимо набрать формулу «corr! X» в Командном окне (рис. П3.6) и нажать кнопку .

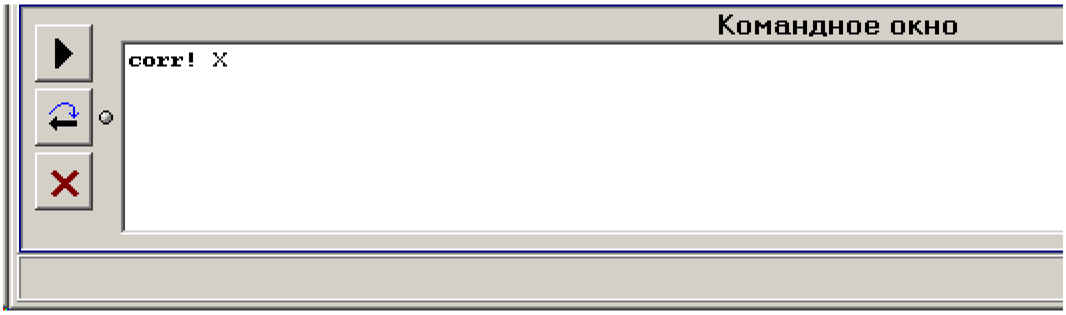


Рис. П3.6. Командное окно программы Matrixer 5.1

Результат будет иметь вид

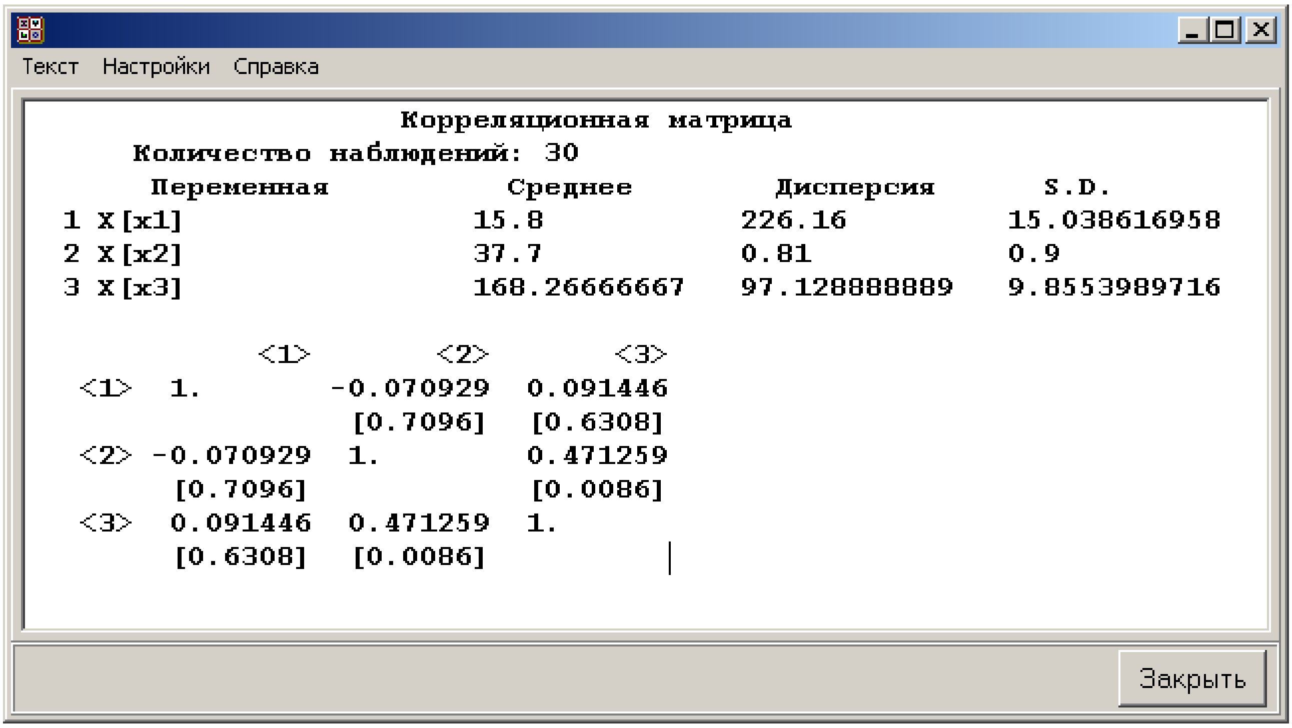


Рис. П3.7. Окно результатов регрессионного анализа

Числа в квадратных скобках означают стандартные ошибки определения соответствующего коэффициента корреляции, т. е. корреляционная матрица имеет вид

	Столбец 1	Столбец 2	Столбец 3
Столбец 1	1	–0,071	0,091
Столбец 2	–0,071	1	0,471
Столбец 3	0,091	0,471	1

Рис. П3.8. Корреляционная матрица

3.2. Регрессионный анализ

Рассмотрим построение уравнения линейной множественной регрессии по данным наблюдений за совместным изменением p+1 переменной y *и* x_j и $((y_i,\ x_{j,i})$; $j = 1, 2, ..., p$; $i = 1, 2, ..., n)$ (табл. 3.1).

Создадим новый вектор Y, как было показано в предыдущем пункте со значениями, показанными на рис. П2.1. Для построения уравнения линейной множественной регрессии необходимо набрать формулу «Y: 1 X» в Командном окне (рис. П3.9) и нажать кнопку ▶.. Единица в формуле означает наличие свободного члена в уравнении регрессии.

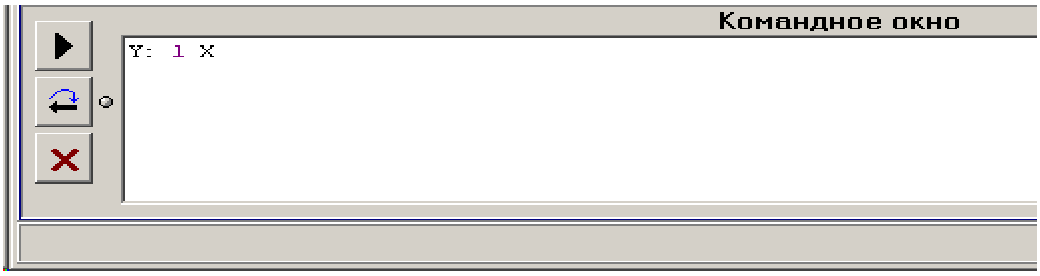
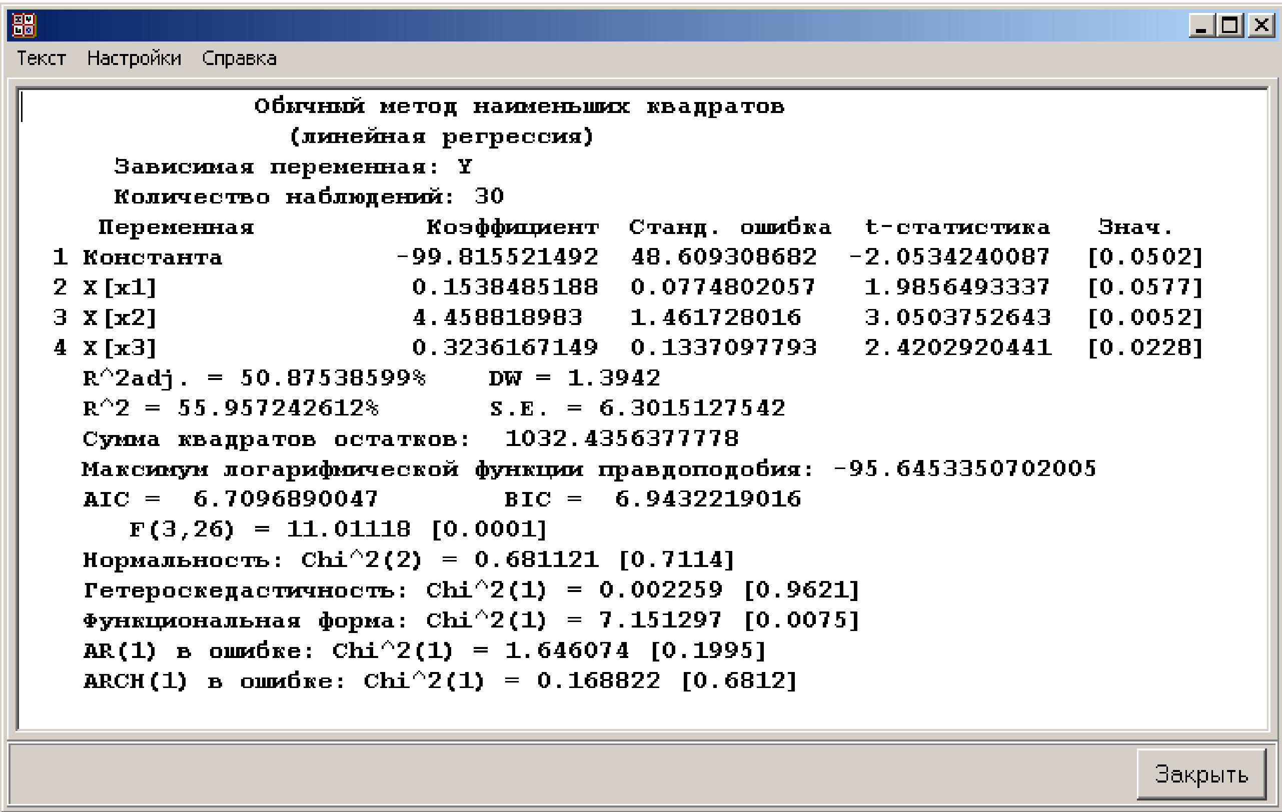


Рис. П3.9. Командное окно программы Matrixer 5.1

Результат будет иметь вид




```
                 Обычный метод наименьших квадратов
                      (линейная регрессия)
         Зависимая переменная: Y
         Количество наблюдений: 30
        Переменная                 Коэффициент  Станд. ошибка  t-статистика   Знач.
   1 Константа                  -99.815521492  48.609308682  -2.0534240087  [0.0502]
   2 X[x1]                        0.1538485188  0.0774802057   1.9856493337  [0.0577]
   3 X[x2]                        4.458818983   1.461728016    3.0503752643  [0.0052]
   4 X[x3]                        0.3236167149  0.1337097793   2.4202920441  [0.0228]
     R^2adj. = 50.87538599%     DW = 1.3942
     R^2 = 55.957242612%        S.E. = 6.3015127542
     Сумма квадратов остатков:  1032.4356377778
     Максимум логарифмической функции правдоподобия: -95.6453350702005
     AIC =  6.7096890047         BIC =  6.9432219016
        F(3,26) = 11.01118 [0.0001]
     Нормальность: Chi^2(2) = 0.681121 [0.7114]
     Гетероскедастичность: Chi^2(1) = 0.002259 [0.9621]
     Функциональная форма: Chi^2(1) = 7.151297 [0.0075]
     AR(1) в ошибке: Chi^2(1) = 1.646074 [0.1995]
     ARCH(1) в ошибке: Chi^2(1) = 0.168822 [0.6812]
```

Рис. П3.10. Окно результатов построения корреляционной матрицы

Искомые значения коэффициентов линейного уравнения регрессии (a, b_i) берутся из столбца «Коэффициенты» таблицы результатов регрессии, из которой следует, что уравнение регрессии имеет вид

$$\hat{y} = -99{,}816 + 0{,}154 \cdot x_1 + 4{,}459 \cdot x_2 + 0{,}324 \cdot x_3.$$

Столбец «Станд. ошибка» содержит стандартные ошибки определения коэффициентов уравнения регрессии.

Столбец «t-статистика» содержит фактические значения критерия Стьюдента для соответствующего коэффициента.

Столбец «Знач.» содержит минимальный уровень значимости коэффициента α_0.

Коэффициент детерминации $R^2 = 0{,}55957$.

Модифицированный коэффициент детерминации (У3.28) $\overline{R}^2 = 0{,}50875$.

Приложение 4. Статистические таблицы

4.1. Нормированная функция Лапласа

$$\Phi(e) = \frac{1}{\sqrt{2\pi}} \int_0^e e^{-\frac{t^2}{2}} \cdot dt = \frac{1}{\sqrt{2\pi}} \int_{-e}^0 e^{-\frac{t^2}{2}} \cdot dt$$

t	0	1	2	3	4	5	6	7	8	9
0	0	40	80	120	160	199	239	279	319	359
0,1	398	438	478	517	557	596	636	675	714	753
0,2	793	832	871	910	948	987	1026	1064	1103	1141
0,3	1179	1217	1255	1293	1331	1368	1406	1443	1480	1517
0,4	1554	1591	1628	1664	1700	1736	1772	1808	1844	1879
0,5	1915	1950	1985	2019	2054	2088	2123	2157	2190	2224
0,6	2257	2291	2324	2357	2389	2422	2454	2486	2517	2549
0,7	2580	2611	2642	2673	2704	2734	2764	2794	2823	2852
0,8	2881	2910	2939	2967	2995	3023	3051	3078	3106	3133
0,9	3159	3186	3212	3238	3264	3289	3315	3340	3365	3389
1	3413	3438	3461	3485	3508	3531	3554	3577	3599	3621
1,1	3643	3665	3686	3708	3729	3749	3770	3790	3810	3830
1,2	3849	3869	3888	3907	3925	3944	3962	3980	3997	4015
1,3	4032	4049	4066	4082	4099	4115	4131	4147	4162	4177
1,4	4192	4207	4222	4236	4251	4265	4279	4292	4306	4319
1,5	4332	4345	4357	4370	4382	4394	4406	4418	4429	4441
1,6	4452	4463	4474	4484	4495	4505	4515	4525	4535	4545
1,7	4554	4564	4573	4582	4591	4599	4608	4616	4625	4633
1,8	4641	4649	4656	4664	4671	4678	4686	4693	4699	4706
1,9	4713	4719	4726	4732	4738	4744	4750	4756	4761	4767
2	4772	4778	4783	4788	4793	4798	4803	4808	4812	4817
2,1	4821	4826	4830	4834	4838	4842	4846	4850	4854	4857
2,2	4861	4864	4868	4871	4875	4878	4881	4884	4887	4890
2,3	4893	4896	4898	4901	4904	4906	4909	4911	4913	4916
2,4	4918	4920	4922	4925	4927	4929	4931	4932	4934	4936
2,5	4938	4940	4941	4943	4945	4946	4948	4949	4951	4952
2,6	4953	4955	4956	4957	4959	4960	4961	4962	4963	4964
2,7	4965	4966	4967	4968	4969	4970	4971	4972	4973	4974
2,8	4974	4975	4976	4977	4977	4978	4979	4979	4980	4981
2,9	4981	4982	4982	4983	4984	4984	4985	4985	4986	4986
3	4987	4987	4987	4988	4988	4989	4989	4989	4990	4990
3,1	4990	4991	4991	4991	4992	4992	4992	4992	4993	4993
3,2	4993	4993	4994	4994	4994	4994	4994	4995	4995	4995
3,3	4995	4995	4995	4996	4996	4996	4996	4996	4996	4997
3,4	4997	4997	4997	4997	4997	4997	4997	4997	4997	4998
3,5	4998	4998	4998	4998	4998	4998	4998	4998	4998	4998
*Значения ординат увеличены в 10 000 раз										

4.2. Значения критических уровней $t_{\alpha,k}$ в зависимости от k степеней свободы и заданного уровня значимости α для распределения Стьюдента

k \ α	0,1	0,05	0,025	0,02	0,01	0,005	0,001
1	6,31	12,71	25,45	31,82	63,66	127,32	636,62
2	2,92	4,30	6,21	6,96	9,92	14,09	31,60
3	2,35	3,18	4,18	4,54	5,84	7,45	12,92
4	2,13	2,78	3,50	3,75	4,60	5,60	8,61
5	2,02	2,57	3,16	3,36	4,03	4,77	6,87
6	1,94	2,45	2,97	3,14	3,71	4,32	5,96
7	1,89	2,36	2,84	3,00	3,50	4,03	5,41
8	1,86	2,31	2,75	2,90	3,36	3,83	5,04
9	1,83	2,26	2,69	2,82	3,25	3,69	4,78
10	1,81	2,23	2,63	2,76	3,17	3,58	4,59
11	1,80	2,20	2,59	2,72	3,11	3,50	4,44
12	1,78	2,18	2,56	2,68	3,05	3,43	4,32
13	1,77	2,16	2,53	2,65	3,01	3,37	4,22
14	1,76	2,14	2,51	2,62	2,98	3,33	4,14
15	1,75	2,13	2,49	2,60	2,95	3,29	4,07
16	1,75	2,12	2,47	2,58	2,92	3,25	4,01
17	1,74	2,11	2,46	2,57	2,90	3,22	3,97
18	1,73	2,10	2,45	2,55	2,88	3,20	3,92
19	1,73	2,09	2,43	2,54	2,86	3,17	3,88
20	1,72	2,09	2,42	2,53	2,85	3,15	3,85
21	1,72	2,08	2,41	2,52	2,83	3,14	3,82
22	1,72	2,07	2,41	2,51	2,82	3,12	3,79
23	1,71	2,07	2,40	2,50	2,81	3,10	3,77
24	1,71	2,06	2,39	2,49	2,80	3,09	3,75
25	1,71	2,06	2,38	2,49	2,79	3,08	3,73
26	1,71	2,06	2,38	2,48	2,78	3,07	3,71
27	1,70	2,05	2,37	2,47	2,77	3,06	3,69
28	1,70	2,05	2,37	2,47	2,76	3,05	3,67
29	1,70	2,05	2,36	2,46	2,76	3,04	3,66
30	1,70	2,04	2,36	2,46	2,75	3,03	3,65
35	1,69	2,03	2,34	2,44	2,72	3,00	3,59
40	1,68	2,02	2,33	2,42	2,70	2,97	3,55
45	1,68	2,01	2,32	2,41	2,69	2,95	3,52
50	1,68	2,01	2,31	2,40	2,68	2,94	3,50
60	1,67	2,00	2,30	2,39	2,66	2,91	3,46
70	1,67	1,99	2,29	2,38	2,65	2,90	3,44
80	1,66	1,99	2,28	2,37	2,64	2,89	3,42
100	1,66	1,98	2,28	2,36	2,63	2,87	3,39
∞	1,64	1,96	2,24	2,33	2,58	2,81	3,29

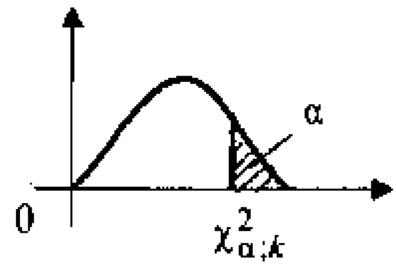
4.3. Значения *F*-критерия Фишера на уровне значимости $\alpha = 0{,}05$

k_2 \ k_1	**1**	**2**	**3**	**4**	**5**	**6**	**8**	**12**	**24**	**∞**
1	161,45	199,50	215,71	224,58	230,16	233,99	238,88	243,91	249,05	254,31
2	18,51	19,00	19,16	19,25	19,30	19,33	19,37	19,41	19,45	19,50
3	10,13	9,55	9,28	9,12	9,01	8,94	8,85	8,74	8,64	8,53
4	7,71	6,94	6,59	6,39	6,26	6,16	6,04	5,91	5,77	5,63
5	6,61	5,79	5,41	5,19	5,05	4,95	4,82	4,68	4,53	4,36
6	5,99	5,14	4,76	4,53	4,39	4,28	4,15	4,00	3,84	3,67
7	5,59	4,74	4,35	4,12	3,97	3,87	3,73	3,57	3,41	3,23
8	5,32	4,46	4,07	3,84	3,69	3,58	3,44	3,28	3,12	2,93
9	5,12	4,26	3,86	3,63	3,48	3,37	3,23	3,07	2,90	2,71
10	4,96	4,10	3,71	3,48	3,33	3,22	3,07	2,91	2,74	2,54
11	4,84	3,98	3,59	3,36	3,20	3,09	2,95	2,79	2,61	2,40
12	4,75	3,89	3,49	3,26	3,11	3,00	2,85	2,69	2,51	2,30
13	4,67	3,81	3,41	3,18	3,03	2,92	2,77	2,60	2,42	2,21
14	4,60	3,74	3,34	3,11	2,96	2,85	2,70	2,53	2,35	2,13
15	4,54	3,68	3,29	3,06	2,90	2,79	2,64	2,48	2,29	2,07
16	4,49	3,63	3,24	3,01	2,85	2,74	2,59	2,42	2,24	2,01
17	4,45	3,59	3,20	2,96	2,81	2,70	2,55	2,38	2,19	1,96
18	4,41	3,55	3,16	2,93	2,77	2,66	2,51	2,34	2,15	1,92
19	4,38	3,52	3,13	2,90	2,74	2,63	2,48	2,31	2,11	1,88
20	4,35	3,49	3,10	2,87	2,71	2,60	2,45	2,28	2,08	1,84
21	4,32	3,47	3,07	2,84	2,68	2,57	2,42	2,25	2,05	1,81
22	4,30	3,44	3,05	2,82	2,66	2,55	2,40	2,23	2,03	1,78
23	4,28	3,42	3,03	2,80	2,64	2,53	2,37	2,20	2,01	1,76
24	4,26	3,40	3,01	2,78	2,62	2,51	2,36	2,18	1,98	1,73
25	4,24	3,39	2,99	2,76	2,60	2,49	2,34	2,16	1,96	1,71
26	4,23	3,37	2,98	2,74	2,59	2,47	2,32	2,15	1,95	1,69
27	4,21	3,35	2,96	2,73	2,57	2,46	2,31	2,13	1,93	1,67
28	4,20	3,34	2,95	2,71	2,56	2,45	2,29	2,12	1,91	1,65
29	4,18	3,33	2,93	2,70	2,55	2,43	2,28	2,10	1,90	1,64
30	4,17	3,32	2,92	2,69	2,53	2,42	2,27	2,09	1,89	1,62
35	4,12	3,27	2,87	2,64	2,49	2,37	2,22	2,04	1,83	1,56
40	4,08	3,23	2,84	2,61	2,45	2,34	2,18	2,00	1,79	1,51
45	4,06	3,20	2,81	2,58	2,42	2,31	2,15	1,97	1,76	1,47
50	4,03	3,18	2,79	2,56	2,40	2,29	2,13	1,95	1,74	1,44
60	4,00	3,15	2,76	2,53	2,37	2,25	2,10	1,92	1,70	1,39
70	3,98	3,13	2,74	2,50	2,35	2,23	2,07	1,89	1,67	1,35
80	3,96	3,11	2,72	2,49	2,33	2,21	2,06	1,88	1,65	1,32
90	3,95	3,10	2,71	2,47	2,32	2,20	2,04	1,86	1,64	1,30
100	3,94	3,09	2,70	2,46	2,31	2,19	2,03	1,85	1,63	1,28
∞	3,84	3,00	2,60	2,37	2,21	2,10	1,94	1,75	1,52	1,00

4.4. Значения F-критерия Фишера на уровне значимости α = 0,01

k_2 \ k_1	1	2	3	4	5	6	8	12	24	∞
1	4052,18	4999,50	5403,35	5624,58	5763,65	5858,99	5981,07	6106,32	6234,63	6365,86
2	98,50	99,00	99,17	99,25	99,30	99,33	99,37	99,42	99,46	99,50
3	34,12	30,82	29,46	28,71	28,24	27,91	27,49	27,05	26,60	26,13
4	21,20	18,00	16,69	15,98	15,52	15,21	14,80	14,37	13,93	13,46
5	16,26	13,27	12,06	11,39	10,97	10,67	10,29	9,89	9,47	9,02
6	13,75	10,92	9,78	9,15	8,75	8,47	8,10	7,72	7,31	6,88
7	12,25	9,55	8,45	7,85	7,46	7,19	6,84	6,47	6,07	5,65
8	11,26	8,65	7,59	7,01	6,63	6,37	6,03	5,67	5,28	4,86
9	10,56	8,02	6,99	6,42	6,06	5,80	5,47	5,11	4,73	4,31
10	10,04	7,56	6,55	5,99	5,64	5,39	5,06	4,71	4,33	3,91
11	9,65	7,21	6,22	5,67	5,32	5,07	4,74	4,40	4,02	3,60
12	9,33	6,93	5,95	5,41	5,06	4,82	4,50	4,16	3,78	3,36
13	9,07	6,70	5,74	5,21	4,86	4,62	4,30	3,96	3,59	3,17
14	8,86	6,51	5,56	5,04	4,69	4,46	4,14	3,80	3,43	3,00
15	8,68	6,36	5,42	4,89	4,56	4,32	4,00	3,67	3,29	2,87
16	8,53	6,23	5,29	4,77	4,44	4,20	3,89	3,55	3,18	2,75
17	8,40	6,11	5,18	4,67	4,34	4,10	3,79	3,46	3,08	2,65
18	8,29	6,01	5,09	4,58	4,25	4,01	3,71	3,37	3,00	2,57
19	8,18	5,93	5,01	4,50	4,17	3,94	3,63	3,30	2,92	2,49
20	8,10	5,85	4,94	4,43	4,10	3,87	3,56	3,23	2,86	2,42
21	8,02	5,78	4,87	4,37	4,04	3,81	3,51	3,17	2,80	2,36
22	7,95	5,72	4,82	4,31	3,99	3,76	3,45	3,12	2,75	2,31
23	7,88	5,66	4,76	4,26	3,94	3,71	3,41	3,07	2,70	2,26
24	7,82	5,61	4,72	4,22	3,90	3,67	3,36	3,03	2,66	2,21
25	7,77	5,57	4,68	4,18	3,85	3,63	3,32	2,99	2,62	2,17
26	7,72	5,53	4,64	4,14	3,82	3,59	3,29	2,96	2,58	2,13
27	7,68	5,49	4,60	4,11	3,78	3,56	3,26	2,93	2,55	2,10
28	7,64	5,45	4,57	4,07	3,75	3,53	3,23	2,90	2,52	2,06
29	7,60	5,42	4,54	4,04	3,73	3,50	3,20	2,87	2,49	2,03
30	7,56	5,39	4,51	4,02	3,70	3,47	3,17	2,84	2,47	2,01
35	7,42	5,27	4,40	3,91	3,59	3,37	3,07	2,74	2,36	1,89
40	7,31	5,18	4,31	3,83	3,51	3,29	2,99	2,66	2,29	1,80
45	7,23	5,11	4,25	3,77	3,45	3,23	2,94	2,61	2,23	1,74
50	7,17	5,06	4,20	3,72	3,41	3,19	2,89	2,56	2,18	1,68
60	7,08	4,98	4,13	3,65	3,34	3,12	2,82	2,50	2,12	1,60
70	7,01	4,92	4,07	3,60	3,29	3,07	2,78	2,45	2,07	1,54
80	6,96	4,88	4,04	3,56	3,26	3,04	2,74	2,42	2,03	1,49
90	6,93	4,85	4,01	3,53	3,23	3,01	2,72	2,39	2,00	1,46
100	6,90	4,82	3,98	3,51	3,21	2,99	2,69	2,37	1,98	1,43
∞	6,63	4,61	3,78	3,32	3,02	2,80	2,51	2,18	1,79	1,00

4.5. Значения $\chi^2_{\alpha;k}$ критерия Пирсона



Число степеней свободы *к*	Вероятность α												
	0,99	0,98	0,95	0,90	0,80	0,70	0,50	0,30	0,20	0,10	0,05	0,02	0,01
1	0,00	0,00	0,00	0,02	0,06	0,15	0,45	1,07	1.64	2,71	3,84	5,41	6,64
2	0,02	0,04	0,10	0,21	0,45	0,71	1,39	2,41	3,22	4,60	5,99	7,82	9,21
3	0,11	0,18	0,35	0,58	1,00	1,42	2,37	3,66	4,64	6,25	7,82	9,84	11,3
4	0,30	0,43	0,71	1,06	1,65	2,20	3,36	4,88	5,99	7,78	9,49	11,7	13,3
5	0,55	0,75	1,14	1,61	2,34	3,00	4,35	6,06	7,29	9,24	11,1	13,4	15,1
6	0,87	1,13	1,63	2,20	3,07	3,83	5,35	7,23	8,56	10,6	12,6	15,0	16,8
7	1,24	1,56	2,17	2,83	3,82	4,67	6,35	8,38	9,80	12,0	14,1	16,6	18,5
8	1,65	2,03	2,73	3,49	4,59	5,53	7,34	9,52	11,0	13,4	15,5	18,2	20,1
9	2,09	2,53	3,32	4,17	5,38	6,39	8,34	10,7	12,2	14,7	16,9	19,7	21,7
10	2,56	3,06	3,94	4,86	6,18	7,27	9,34	11,8	13,4	16,0	18,3	21,2	23,2
11	3,05	3,61	4,58	5,58	6,99	8,15	10,3	12,9	14,6	17,3	19,7	22,6	24,7
12	3,57	4,18	5,23	6,30	7,81	9,03	11,3	14,0	15,8	18,5	21,0	24,1	26,2
13	4,11	4,76	5,89	7,04	8,63	9,93	12,3	15,1	17,0	19,8	22,4	25,5	27,7
14	4,66	5,37	6,57	7,79	9,47	10,8	13,3	16,2	18,1	21,1	23,7	26,9	29,1
15	5,23	5,98	7,26	8,55	10,3	П,7	14,3	17,3	19,3	22,3	25,0	28,3	30,6
16	5,81	6,61	7,96	9,31	11,1	12,6	15,3	18,4	20,5	23,5	26,3	29,6	32,0
17	6,41	7,26	8,67	10,1	12,0	13,5	16,3	19,5	21,6	24,8	27,6	31,0	33,4
18	7.02	7,91	9,39	10,9	12,9	14,4	17,3	20,6	22,8	26,0	28,9	32,3	34,8
19	7,63	8,57	10,1	11,6	13,7	15,3	18,3	21,7	23,9	27,2	30,1	33,7	36,2
20	8.26	9,24	10,8	12,4	14,6	16,3	19,3	22,8	25,0	28,4	31,4	35,0	37,6
21	8,90	9,92	11,6	13,2	15,4	17,2	20,3	23,9	26,2	29,6	32,7	36,3	38,9
22	9,54	10,6	12,3	14,0	16,3	18,1	21,3	24,9	27,3	30,8	33,9	37,7	40,3
23	10,2	11,3	13,1	14,8	17,2	19,0	22,3	26,0	28,4	32,0	35,2	39,0	41,6
24	10,9	12,0	13,8	15,7	18,1	19,9	23,3	27,1	29,6	33,2	36,4	40,3	43,0
25	11,5	12,7	14,6	16,5	18,9	20,9	24,3	28,2	30,7	34,4	37,7	41,7	44,3
26	12,2	13,4	15,4	17,3	19,8	21,8	25,3	29,2	31,8	35,6	38,9	42,9	45,6
27	12,9	14,1	16,1	18,1	20,7	22,7	26,3	30,3	32,9	36,7	40,1	44,1	47,0
28	13,6	14,8	16,9	18,9	21,6	23,6	27,3	31,4	34,0	37,9	41,3	45,4	48,3
29	14,3	15,6	17,7	19,8	22,5	24,6	28,3	32,5	35,1	39,1	42.6	46,7	49,6
30	14,9	16,3	18,5	20,6	23,4	25,5	29,3	33,5	36,2	40,3	43,8	48,0	50,9

4.6. Значения статистик Дарбина-Уотсона d_L d_U при уровне значимости α = 0,05

(n – число наблюдений; m – число объясняющих переменных)

n	m=1		m=2		m=3		m=4		m=5	
	d_L	d_U	d_L	d_U	d_L	d_U	d_L	d_U	d_L	d_U
	0,61	1,40	–	–	–	–				
7	0,70	1,36	0,47	1,90	–	–				
8	0,76	1,33	0,56	1,78	0,37	2,29				
9	0,82	1,32	0,63	1,70	0,46	2,13				
10	0,88	1,32	0,70	1,64	0,53	2,02				
11	0,93	1,32	0,66	1,60	0,60	1,93				
12	0,97	1,33	0,81	1,58	0,66	1,86				
13	1,01	1,34	0,86	1,56	0,72	1,82				
14	1,05	1,35	0,91	1,55	0,77	1,78				
15	1,08	1,36	0,95	1,54	0,82	1,75	0,69	1,97	0,56	2,21
16	1,10	1,37	0,98	1,54	0,86	1,73	0,74	1,93	0,62	2,15
17	1,13	1,38	1,02	1,54	0,90	1,71	0,78	1,90	0,67	2,10
18	1,16	1,39	1,05	1,53	0,93	1,69	0,82	1,87	0,71	2,06
19	1,18	1,40	1,08	1,53	0,97	1,68	0,86	1,85	0,75	2,02
20	1,20	1,41	1,10	1,54	1,00	1,68	0,90	1,83	0,79	1,99
21	1,22	1,42	1,13	1,54	1,03	1,67	0,93	1,81	0,83	1,96
22	1,24	1,43	1,15	1,54	1,05	1,66	0,96	1,80	0,86	1,94
23	1,26	1,44	1,17	1,54	1,08	1,66	0,99	1,79	0,90	1,92
24	1,27	1,45	1,19	1,55	1,10	1,66	1,01	1,78	0,93	1,90
25	1,29	1,45	1,21	1,55	1,12	1,66	1,04	1,77	0,95	1,89
26	1,30	1,46	1,22	1,55	1,14	1,65	1,06	1,76	0,98	1,88
27	1,32	1,47	1,24	1,56	1,16	1,65	1,08	1,76	1,01	1,86
28	1,33	1,48	1,26	1,56	1,18	1,65	1,10	1,75	1,03	1,85
29	1,34	1,48	1,27	1,56	1,20	1,65	1,12	1,74	1,05	1,84
30	1,35	1,49	1,28	1,57	1,21	1,65	1,14	1,74	1,07	1,83
31	1,36	1,50	1,30	1,57	1,23	1,65	1,16	1,74	1,09	1,83
32	1,37	1,50	1,31	1,57	1,24	1,65	1,18	1,73	1,11	1,82
33	1,38	1,51	1,32	1,58	1,26	1,65	1,19	1,73	1,13	1,18
34	1,39	1,51	1,33	1,58	1,27	1,65	1,21	1,73	1,15	1,81
35	140	1,52	1,34	1,58	1,28	1,65	1,22	1,73	1,16	1,80
36	1,41	1,52	1,35	1,59	1,29	1,65	1,24	1,73	1,18	1,80
37	1,42	1,53	1,36	1,59	1,31	1,66	1,25	1,72	1,19	1,80
38	1,43	1,54	1,37	1,59	1,32	1,66	1,26	1,72	1,21	1,79
39	1,43	1,54	1,38	1,60	1,33	1,66	1,27	1,72	1,22	1,79
40	1,44	1,54	1,39	1,60	1,34	1,66	1,28	1,72	1,23	1,78
100	1,65	1,69	1,63	1,72	1,61	1,74	1,59	1,76	1,57	1,78

4.7. Критические значения *f*-критерия для DF-, ADF- и PP-тестов

Уровень значимости	Тип функции		
	без параметра смещения и тренда	включая параметр смещения	включая параметр смещения и тренд
	T=25		
0,01	–2,66	–3,75	–4,38
0,025	-2,26	-3,33	-3,95
0,05	–1,95	–3,00	–3,60
	T=50		
0,01	–2,62	–3,58	–4,15
0,025	–2,25	–3,22	–3,80
0,05	–1,95	–2,93	–3,50
	T=100		
0,01	–2,60	–3,51	–4,04
0,025	–2,24	–3,17	–3,69
0,05	–1,95	–2,89	–3,45
	T= ∞		
0,01	–2,58	–3,43	–3,96
0,025	–2,23	–3,12	–3,66
0,05	–1,951	–2,86	–3,41

Приложение 5. Функции табличных процессоров MS Excel и OpenOffice.org Calc

5.1. Функции табличного процессора MS Excel

Значение	Обозначение	Функция
Критическое значение F-критерия Фишера при уровне значимости α = 0,05 и числе степеней свободы k_1 и k_2	$F_{крит}$	FРАСПОБР(α; k_1; k_2)
Критическое значение t-критерия Стьюдента при уровне значимости α = 0,05 и числе степеней свободы k	$t_{\text{крит}}$	СТЬЮДРАСПОБР(α; k)
Квантиль стандартного нормального распределения порядка 1–α/2	$t_{1-\alpha/2}$	НОРМСТОБР(1–α/2)
Среднее квадратическое отклонение	σ	СТАНДОТКЛОН()
Среднее значение	$\bar{x}$	СРЗНАЧ()
Преобразование Фишера	$z = \mathrm{Z}(r)$	ФИШЕР(r)
Обратное преобразование Фишера	$r = Z^{-1}(z)$	ФИШЕРОБР(z)
Вычисление коэффициента корреляции	r_{xy}	КОРРЕЛ()

5.2 Функции табличного процессора OpenOffice.org Calc

Значение	Обозначение	Функция
Критическое значение F-критерия Фишера при уровне значимости α = 0,05 и числе степеней свободы k_1 и k_2	$F_{крит}$	FINV(α; k_1; k_2)
Критическое значение t-критерия Стьюдента при уровне значимости α = 0,05 и числе степеней свободы k	$t_{\text{крит}}$	TINV(α; k)
Квантиль стандартного нормального распределения порядка 1–α/2	$t_{1-\alpha/2}$	NORMSINV(1–α/2)
Среднее квадратическое отклонение	σ	STDEV()
Среднее значение	$\overline{x}$	AVERAGE()
Преобразование Фишера	$z = \mathrm{Z}(r)$	FISHER(r)
Обратное преобразование Фишера	$r = Z^{-1}(z)$	FISHERINV(z)
Вычисление коэффициента корреляции	r_{xy}	correl()

Приложение 6. Пример отчета по лабораторной работе

Ульяновский государственный технический университет
Кафедра Информационные системы
Дисциплина Эконометрика

ОТЧЕТ
по лабораторной работе №1
«Парный корреляционный анализ: проверка наличия и степени тесноты линейной и нелинейной связи»
Вариант №1

Группа ИСЭд-31 Студент Шамирзаев Н.

Ульяновск 2010

Задание. На основании данных таблицы П1.1 для соответствующего варианта (табл. 1.3):

1. Вычислить линейный коэффициент парной корреляции r_{xy} и индекс корреляции R.
2. Проверить значимость коэффициента парной корреляции r_{xy} и индекса корреляции R при уровне значимости α = 0,05 для нечетного варианта и α = 0,01 для четного варианта.
3. Построить доверительный интервал для значимого линейного коэффициента парной корреляции r_{xy}.

Исходные данные:

– переменные x и y задаются графами 3 и 4 таблицы П1.1;

– зависимости: линейная и $\hat{y} = 0{,}09 \cdot x^{1.25}$;

– уровень значимости α = 0,05.

Решение

1) Вычисление σ_x, σ_y и r_{xy} (1.3), (1.4). Используя данные таблицы 1 получим

Таблица 1

Промежуточные результаты расчетов

Номер наблюдения	x	y	x^2	y^2	xy	$\hat{y}$	$(\hat{y}-y)^2$	$(y-\bar{y})^2$
1	113	39	12769	1521	4407	33,16	34,13	1,00
2	124	37	15376	1369	4588	37,24	0,06	9,00
3	124	36	15376	1296	4464	37,24	1,54	16,00
4	122	36	14884	1296	4392	36,49	0,24	16,00
5	128	26	16384	676	3328	38,75	162,52	196,00
6	140	43	19600	1849	6020	43,34	0,12	9,00
7	117	31	13689	961	3627	34,63	13,19	81,00
8	113	40	12769	1600	4520	33,16	46,81	0,00
9	122	48	14884	2304	5856	36,49	132,44	64,00
10	139	64	19321	4096	8896	42,95	442,90	576,00
11	126	39	15876	1521	4914	37,99	1,01	1,00
12	120	34	14400	1156	4080	35,75	3,05	36,00
13	125	39	15625	1521	4875	37,62	1,91	1,00
14	118	37	13924	1369	4366	35,00	3,99	9,00
15	122	35	14884	1225	4270	36,49	2,22	25,00
16	133	54	17689	2916	7182	40,65	178,23	196,00
17	136	36	18496	1296	4896	41,80	33,63	16,00

Окончание таблицы 1

Номер наблюдения	x	y	x^2	y^2	xy	$\hat{y}$	$(\hat{y}-y)^2$	$(y-\bar{y})^2$
18	136	35	18496	1225	4760	41,80	46,23	25,00
19	138	34	19044	1156	4692	42,57	73,42	36,00
20	124	48	15376	2304	5952	37,24	115,76	64,00
21	123	30	15129	900	3690	36,87	47,14	100,00
22	149	59	22201	3481	8791	46,85	147,58	361,00
Сумма	2792	880	356192	37038	112566	844,081	1488,136	1838
Среднее значение	126,91	40	16190,55	1683,545	5116,636	38,367	67,643	83,545

$$\sigma_x = \sqrt{=\overline{x^2} - \bar{x}^2} = \sqrt{(16190{,}55 - (126{,}91)^2} = 9{,}199,$$

$$\sigma_y = \sqrt{\overline{y^2} - \bar{y}^2} = \sqrt{(1683{,}545 - 40^2)} = 9{,}140,$$

$$r_{xy} = \frac{\overline{yx} - \bar{y}\cdot\bar{x}}{\sigma_x\sigma_y} = \frac{5116{,}636 - 126{,}91\cdot 40}{9{,}199\cdot 9{,}140} = 0{,}479.$$

Вычисление R (1.5):

$$R = \sqrt{1 - \frac{\sum_{i=1}^{n}(\hat{y}_i - y_i)^2}{\sum_{i=1}^{n}(y_i - \bar{y})^2}} = \sqrt{1 - \frac{1486{,}136}{1838}} = 0{,}436\,.$$

2) Проверка значимости r_{xy} (1.6).

$$t_r = \frac{r_{xy}}{\sqrt{\frac{1-r_{xy}^2}{n-2}}} = \frac{0{,}479}{\sqrt{\frac{1-0{,}479^2}{22-2}}} = 2{,}44.$$

$t_{крит} = t_{1-\alpha,n-2}$ = СТЬЮДРАСПОБР(0,05;20) =2,086.

Так как

$$t_r = 2{,}44 > t_{1-\alpha,n-2} = 2{,}086,$$

то делаем вывод о **статистической значимости** линейного коэффициента парной корреляции r_{xy}.

Проверка значимости индекса корреляции R (1.7). Значение F-критерия Фишера

$$F_r = \frac{R^2}{1-R^2}(n-2) = \frac{0{,}436^2}{1-0{,}436^2}(22-2) = 4{,}702$$

При α = 0,05 и степенях свободы $k_1 = 1$, $k_2 = n - 2 = 20 - 2 = 20$.

$F_{крит}$ = FРАСПОБР(0,05;1;20) = 4,35

Так как

$$F_r = 4{,}702 > F_{\text{крит}} = 4{,}35,$$

то формально с погрешностью 5% индекс корреляции следует считать значимым и следовательно с вероятность 95% нельзя отвергать наличие исследуемой зависимости, но учитывая близость значений $F_r = 4{,}702$ и $F_{\text{крит}} = 4{,}35$ зависимость следует считать практически отсутствующей, нельзя считать адекватной.

3) Построение доверительного интервала для линейного коэффициента корреляции r_{xy} (1.8) – (1.10).

Определим величину z (1.8) Z-преобразования Фишера

$$z = \frac{1}{2} \cdot \ln \frac{1 + r_{xy}}{1 - r_{xy}} = \frac{1}{2} \cdot \ln \frac{1 + 0{,}479}{1 - 0{,}479} = 0{,}522.$$

$t_{1-\alpha/2}$ = НОРМСТОБР(0,975) = 1,96.

Вычислим $t_{1-\alpha/2} \cdot \sqrt{\dfrac{1}{n-3}} = 1{,}96 \cdot \sqrt{\dfrac{1}{22-3}} = 1{,}96 \cdot 0{,}2294 = 0{,}450$.

Вычислим границы доверительного интервала (z^-, z^+) для величины z

$$z^- = z' - t_{1-\alpha/2} \cdot \sqrt{\frac{1}{n-3}} = 0{,}522 - 0{,}45 = 0{,}072,$$

$$z^+ = z' + t_{1-\alpha/2} \cdot \sqrt{\frac{1}{n-3}} = 0{,}522 + 0{,}45 = 0{,}972.$$

Определим граничные значения доверительного интервала (r^-, r^+) для r_{xy}.

$r^- = Z^{-1}(z^-) = Z^{-1}(0{,}072) = 0{,}072; \quad r^+ = Z^{-1}(z^+) = Z^{-1}(0{,}972) = 0{,}75.$

Искомый доверительный интервал для r_{xy} имеет вид (0,072; 0,75).

Результаты:

1) $r_{xy} = 0{,}479. \quad R = 0{,}436.$

2) Коэффициенты для r_{xy} и R – статистически значимы.

3) Доверительный интервал для r_{xy} – (0,072; 0,75).

Библиографический список

1. Айвазян, С. А. Прикладная статистика. Основы эконометрики : учебник для вузов. В 2 т. / С. А. Айвазян, В. С. Мхитарян. – 2-е изд., испр. – М. : ЮНИТИ, 2001. – Т. 1: Теория вероятностей и прикладная статистика. – 656 с.
2. Доугерти, К. Введение в эконометрику / К. Доугерти. – М. : ИНФРА-М, 1997. – 402 с.
3. Елисеева, И. И. Эконометрика : учебное пособие / И. И. Елисеева, С. В. Курышева, Д. М. Гордиенко и др. – М. : Финансы и статистика, 2001.
4. Кремер, Н. Ш. Эконометрика / Н. Ш. Кремер, Б. А. Путко. – М. : ЮНИТИ, 2005. – 311 с.
5. Практикум по эконометрике : учебное пособие / под ред. И. И. Елисеевой. – М. : Финансы и статистика, 2002. – 191 с.
6. Шанченко, Н. И. Лекции по эконометрике : учебное пособие для студентов высших учебных заведений, обучающихся по специальности «Прикладная информатика (в экономике)» / Н. И. Шанченко. – Ульяновск : УлГТУ, 2008. – 139 с.

Интернет-ресурсы

7. Рассылка «Эконометрика». − Режим доступа : http://subscribe.ru/archive/ science.humanity. econometrika.
8. Ресурсы по статистике и эконометрике. − Режим доступа : http://dist-economics.eu.spb.ru/ HTML/predmet/econometrics.htm.
9. Ресурсы по статистике и эконометрике. − Режим доступа : http://research.by/rus/links.
10. Ресурсы по статистике и эконометрике. − Режим доступа : http://www.ecsocman.edu.ru/ db/msg/163749.html.
11. Сайт фирмы Statsoft разработчика пакета STATISTICA. − Режим доступа : http://www.statsoft.ru.
12. Статистический Портал StatSoft. − Режим доступа : http://www. statistica.ru.
13. Эконометрика. Библиотека. Единое окно доступа к образовательным ресурсам. − Режим доступа : http://window.edu.ru/window/library?p_rubr= 2.2.76.4.8.
14. Эконометрическая программа Matrixer 5.1. А.Цыплаков (tsy@land4.nsu.ru). − Режим доступа : http://www.nsu.ru/ef/tsy/ecmr/mtx/index.htm.
15. Эконометрическая страничка НГУ. − Режим доступа : http://www.nsu.ru/ ef/tsy/ecmr/ index.htm.
16. Экономическая библиотека онлайн. − Режим доступа : http://www. elobook.com/ ekonometr/index.htm.
17. Электронные библиотеки России. Полнотекстовые pdf-учебники. − Режим доступа : http://www.gaudeamus.omskcity.com/PDF_library_economic_7.html.
18. Электронный учебник по статистике. М. : StatSoft, Inc, 2001. − Режим доступа : http://www. statsoft.ru/home/download/textbook/default.htm.

Учебное издание

ШАНЧЕНКО Николай Иванович

Эконометрика. Лабораторный практикум

Учебное пособие

Редактор М. В. Штаева
Компьютерная верстка Н. И. Шанченко

Подписано в печать 19.05.2011. Формат 60×84/16.
Усл. печ. л. 6,98. Тираж 80 экз. Заказ 560.

Ульяновский государственный технический университет
432027, г. Ульяновск, ул. Сев. Венец, 32.

Типография УлГТУ, 432027, Ульяновск, ул. Сев. Венец, 32.