

Министерство науки и высшего образования Российской Федерации

Федеральное государственное бюджетное образовательное учреждение
высшего образования
«Оренбургский государственный университет»

Т.В. Лебедева, Н.С. Еремеева

ЭКОНОМЕТРИКА. РАСЧЕТЫ В MICROSOFT EXCEL

Практикум

Рекомендовано ученым советом федерального государственного бюджетного образовательного учреждения высшего образования «Оренбургский государственный университет» для обучающихся по образовательным программам высшего образования по направлению подготовки 38.05.01 Экономическая безопасность

Оренбург
2021

УДК 330.4(075.8)

ББК 65в631я73

Л 33

Рецензент – кандидат экономических наук, заместитель министра экономического развития, инвестиций, туризма и внешних связей Оренбургской области Здорова Е.С.

Лебедева Т.В.

Л 33 Эконометрика. Расчеты в Microsoft Excel : практикум / Т.В. Лебедева, Н.С. Еремеева ; Оренбургский гос. ун-т. – Оренбург : ОГУ, 2021. – 196 с.
ISBN 978-5-7410-2533-8

В практикуме подробно, вплоть до пошаговых инструкций, описаны способы выполнения лабораторных работ по дисциплине «Эконометрика» с помощью Microsoft Excel – самой известной и наиболее часто используемой программы для осуществления подобных расчетов.

Наличие в практикуме контрольных вопросов и тестов позволяет проверить усвоение основных определений, категорий и методов эконометрики. В приложениях приводится информационная база для выполнения лабораторных работ.

Практикум предназначен для обучающихся всех форм обучения направления подготовки 38.05.01 Экономическая безопасность.

УДК 330.4(075.8)

ББК 65в631я73

ISBN 978-5-7410-2533-8

© Лебедева Т.В.,
Еремеева Н.С., 2021
© ОГУ, 2021

Содержание

Введение.....	5
Общие указания.....	7
1 Лабораторная работа № 1. Классическая модель линейной регрессии: построение парной линейной регрессии	8
Контрольные вопросы к лабораторной работе № 1	29
Тесты для самоконтроля к лабораторной работе № 1	29
2 Лабораторная работа № 2. Классическая модель линейной регрессии: построение множественной линейной регрессии.....	33
Контрольные вопросы к лабораторной работе № 2	52
Тесты для самоконтроля к лабораторной работе № 2.....	53
3 Лабораторная работа № 3. Регрессионные модели с переменной структурой.....	56
Контрольные вопросы к лабораторной работе № 3	68
Тесты для самоконтроля к лабораторной работе № 3	69
4 Лабораторная работа № 4. Нарушения допущений классической модели линейной регрессии: гетероскедастичность.....	72
Контрольные вопросы к лабораторной работе № 4	86
Тесты для самоконтроля к лабораторной работе № 4.....	86
5 Лабораторная работа № 5. Нарушения допущений классической модели линейной регрессии: автокорреляция	90
Контрольные вопросы к лабораторной работе № 5	99
Тесты для самоконтроля к лабораторной работе № 5.....	100
6 Лабораторная работа № 6 Нелинейная регрессия и способы линеаризации моделей нелинейных по переменным.....	103
Контрольные вопросы к лабораторной работе № 6	112
Тесты для самоконтроля к лабораторной работе № 6.....	112
7 Лабораторная работа № 7 Нелинейная регрессия и способы линеаризации моделей нелинейных по параметрам	116

Контрольные вопросы к лабораторной работе № 7	128
Тесты для самоконтроля к лабораторной работе № 7.....	129
8 Лабораторная работа № 8 Моделирование временных рядов: исследование тенденции	132
Контрольные вопросы к лабораторной работе № 8	147
Тесты для самоконтроля к лабораторной работе № 8.....	147
9 Лабораторная работа № 9. Моделирование временных рядов:	
исследование сезонности	151
Контрольные вопросы к лабораторной работе № 9	172
Тесты для самоконтроля к лабораторной работе № 9.....	172
Список использованных источников	175
Приложение А Исходные данные для выполнения лабораторных работ	177
Приложение Б Статистико-математические таблицы.....	192

Введение

В современных программах подготовки экономистов курс эконометрики занял одно из ключевых мест, поскольку сегодня деятельность в любой области экономики требует от специалиста применения современных методов оценки, анализа и интерпретации экономических данных.

Сегодня эконометрические методы применяются в качестве стандартных в различных отраслях прикладной экономики, изучающей все, начиная от расходов домашних хозяйств и предпринимательских инвестиций и заканчивая организацией производств, рынков труда и эффектами государственной политики. Особенность эконометрических моделей, обусловленная наличием в них случайных составляющих, требует использования основ теории вероятностей и статистики.

Наиболее важная задача эконометрики состоит в том, чтобы количественно измерить связи между различными экономическими процессами и явлениями на основе имеющихся данных при помощи статистических методов, а также соответствующим образом интерпретировать и использовать полученные результаты. Следовательно, эконометрика – это взаимодействие экономической теории, наблюдаемых данных и статистических методов. Именно взаимодействие этих трех составляющих делает эконометрику интересной, многообещающей и, не самой легкой в понимании.

Именно для овладения обучающимися необходимыми основами знаний по эконометрике разработан данный лабораторный практикум.

Практикум составлен в соответствии с учебной программой курса и предназначен для обучающихся направления подготовки 38.05.01 Экономическая безопасность. Особенности преподавания эконометрики для различных направлений определяются уровнем базовой математической

подготовки обучающихся, степенью адаптации теоретических основ и лабораторного практикума к наиболее актуальным для экономических специальностей задачам исследования, а также объемом выделяемых для данной дисциплины учебных часов. Практикум рассчитан на лиц, имеющих подготовку по статистике, теории вероятностей, информатике, хотя авторы стремились сделать изложение максимально доступным.

Задания к лабораторным работам составлены по основным темам курса: «Классическая модель множественной линейной регрессии», «Регрессионные модели с переменной структурой», «Нарушения допущений классической модели линейной регрессии», «Нелинейная регрессия», «Моделирование временных рядов». Для самоконтроля, а также подготовке к защите лабораторных работ, в практикуме предложены контрольные вопросы и тесты.

Применение *Microsoft Excel*, при выполнении лабораторных работ, дает возможность обучающимся повысить темп решения задач, и сосредоточить их внимание на понимании экономического смысла исчисленных показателей, на объяснении полученных результатов и формулировании выводов. Кроме того, табличный редактор *MS Excel* обладает довольно широкими возможностями решения задач эконометрики, при этом он способствует более глубокому усвоению теории, так как в отличие от ППП *Statistica*, *Eviews*, *СЭМП* и других, требует самостоятельных корректировок и этапов построения моделей, основанных на положениях теории.

Общие указания

В лабораторных работах рассмотрены основные вопросы касающиеся теории изучаемой темы, приводятся решения типовых заданий и варианты индивидуальных заданий для самостоятельной работы. Обучающимся должен быть представлен отчет по каждой выполненной лабораторной работе. Содержание отчета выполняется в соответствии с [общими требованиями и правилами оформления студенческих работ](#) (Требования СТО 02069024.101-2015. Работы студенческие). Защита отчетов проводится по вопросам, размещенным в конце каждой лабораторной работы.

Задания для лабораторных работ распределяются между обучающимися по вариантам в зависимости от начальной буквы фамилии (таблица 1), либо выдаются преподавателем.

Таблица 1 – Распределение заданий по вариантам

Первая буква фамилии	А, Я	Б, Ю	В, Э	Г, Щ	Д, Ш	Е, Ё	Ж, З	И, Ч	К, Ц	Л, Х	М, Ф	Н, У	О, Т	П, С	Р
Номер варианта	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15

В пособии представлены типовые задания для выполнения лабораторных работ по дисциплине. Обучающему представляется возможность выбора выполнения заданий либо по предложенному массиву данных (приложение А), либо сформировать самостоятельно массив данных обосновав систему показателей.

При выполнении лабораторных работ необходимо в отчетах очень подробно отображать этапы решения по каждому заданию, с представлением исходных данных.

1 Лабораторная работа № 1. Классическая модель линейной регрессии: построение парной линейной регрессии

Цель выполнения лабораторной работы: научиться оценивать параметры эконометрических моделей, представленных линейными уравнениями парной регрессии, оценивать их качество, разрабатывать вариантный прогноз и интерпретировать полученные результаты.

Задания:

Для эндогенной переменной Y – интегральный показатель экономической безопасности организации и экзогенной переменной X – объем реализованной продукции (млн. рублей) выполните следующие задания:

1 Постройте поле корреляции и сформулируйте гипотезу о форме связи между Y и X .

2 Рассчитайте оценки параметров уравнения парной линейной регрессии.

3 Оцените тесноту связи между Y и X с помощью выборочного коэффициента корреляции. Проверьте значимость коэффициента корреляции ($\alpha = 0,05$).

4 Рассчитайте выборочный коэффициент детерминации. Сделайте экономический вывод.

5 Проверьте значимость оценки коэффициента регрессии с помощью критерия Стьюдента при уровне значимости $\alpha = 0,05$.

6 Постройте доверительный интервал для коэффициента регрессии. Дайте экономическую интерпретацию.

7 Составьте таблицу дисперсионного анализа.

8 Оцените с помощью F-критерия Фишера – Снедекора значимость уравнения линейной регрессии ($\alpha = 0,05$).

9 Рассчитайте интегральный показатель экономической безопасности организации, если объем реализованной продукции снизится на 10 % относительно последнего периода. Постройте доверительный интервал для прогнозного значения эндогенной переменной. Сделайте экономический вывод.

10 Определить среднюю ошибку аппроксимации.

11 На поле корреляции постройте линию регрессии.

Примечание: Y – интегральный показатель экономической безопасности организации:

- ниже 0,3 характеризует высокий уровень безопасности предприятия;
- от 0,31 до 0,6 – достаточная безопасность;
- от 0,61 до 0,9 – умеренная безопасность;
- от 0,91 до 1,2 – допустимая безопасность;
- от 1,21 и выше – кризисное состояние предприятия.

Реализация типовых задач

1 Постройте поле корреляции и сформулируйте гипотезу о форме связи между Y и X

Для построения «Поля корреляции» воспользуемся табличным редактором Excel. Выделяем столбец со значениями эндогенной переменной Y , затем во вкладке «Вставка» выбираем «Диagramмы» – «Точечная» (рисунок 1.1).

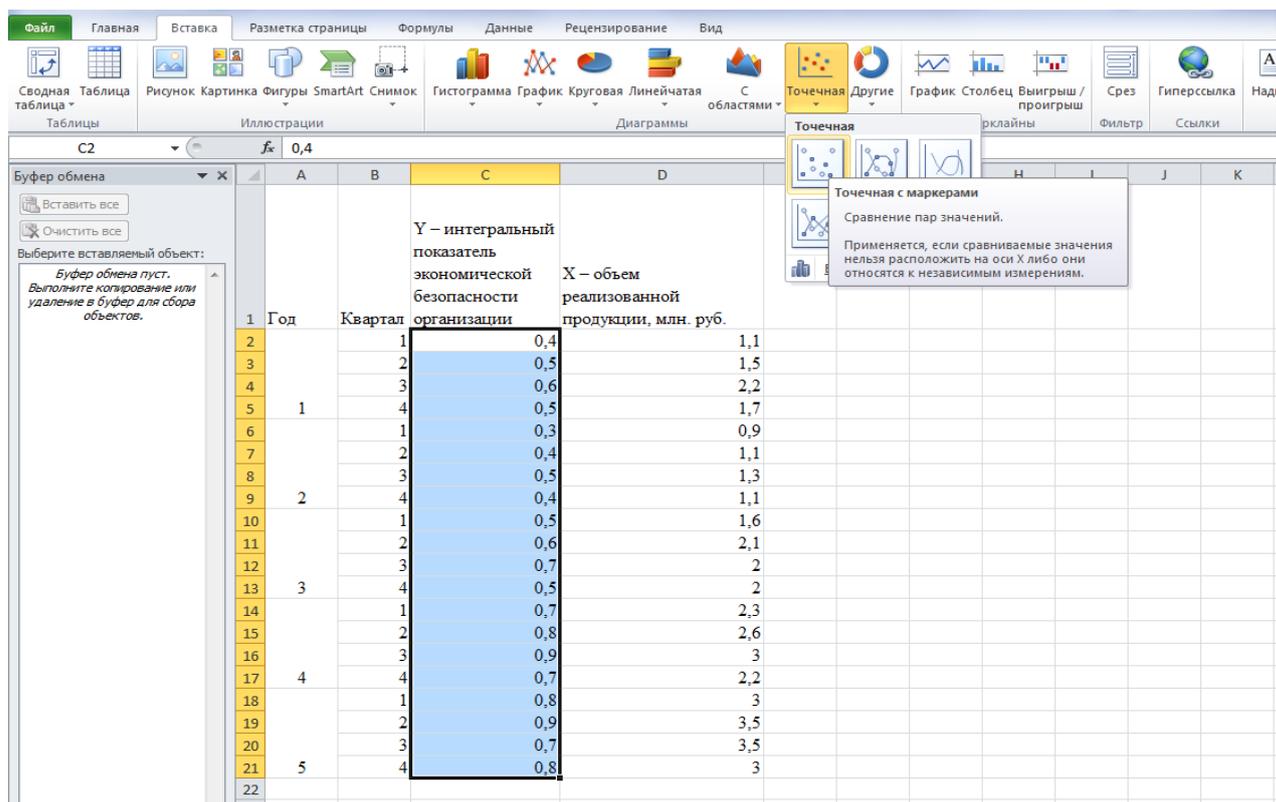


Рисунок 1.1 – Выбор вида диаграммы в MS Excel

В результате получится график, представленный на рисунке 1.2. При нем по *оси абсцисс* отмечены *номера наблюдений*, нам необходимо отметить по этой оси *значения экзогенных переменных*. Для этого во вкладке «Работа с диаграммами» заходим в «Конструктор» – «Выбрать данные» (также можно открыть это окно нажав правой клавишей «мыши» по полю графика), появится диалоговое окно «Выбор источника данных» (рисунок 1.2).

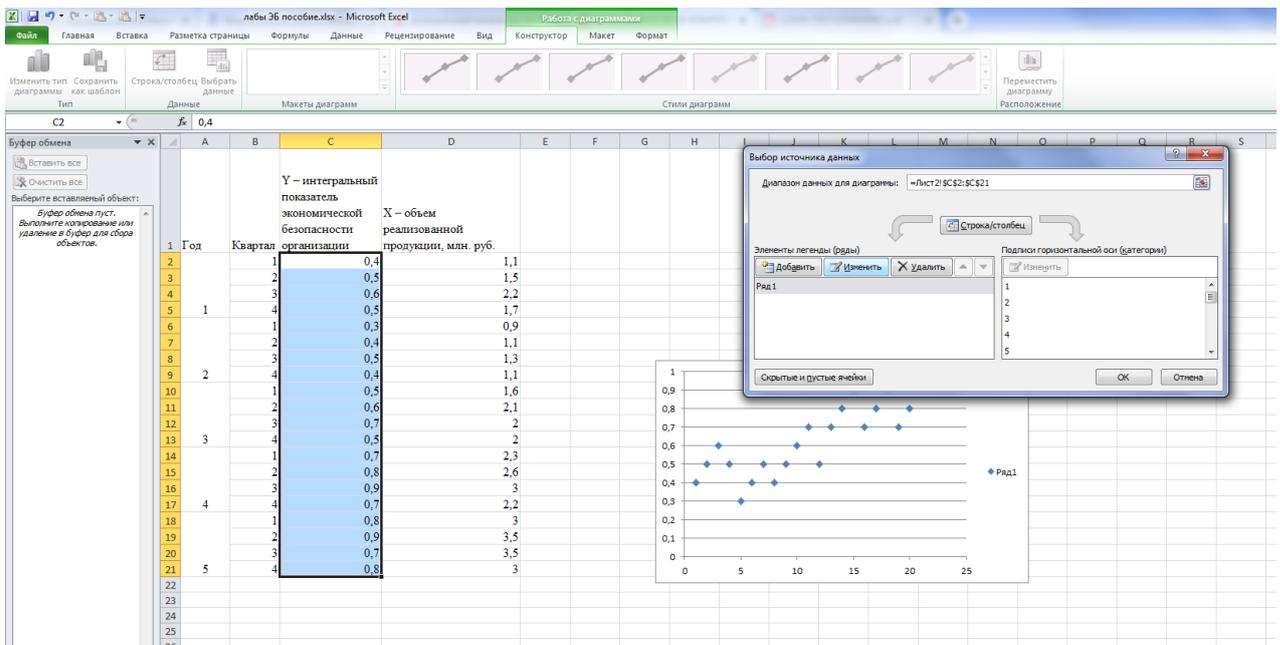


Рисунок 1.2 – Работа с полем «Выбор источника данных»

В диалоговом окне «Выбор источника данных» выбираем «Изменить» (рисунок 1.2), появится диалоговое окно «Изменение ряда». В строке «Значения X» выделяем столбец со значениями экзогенной переменной (рисунок 1.3). В результате преобразований график примет вид, как показано на рисунке 1.3.

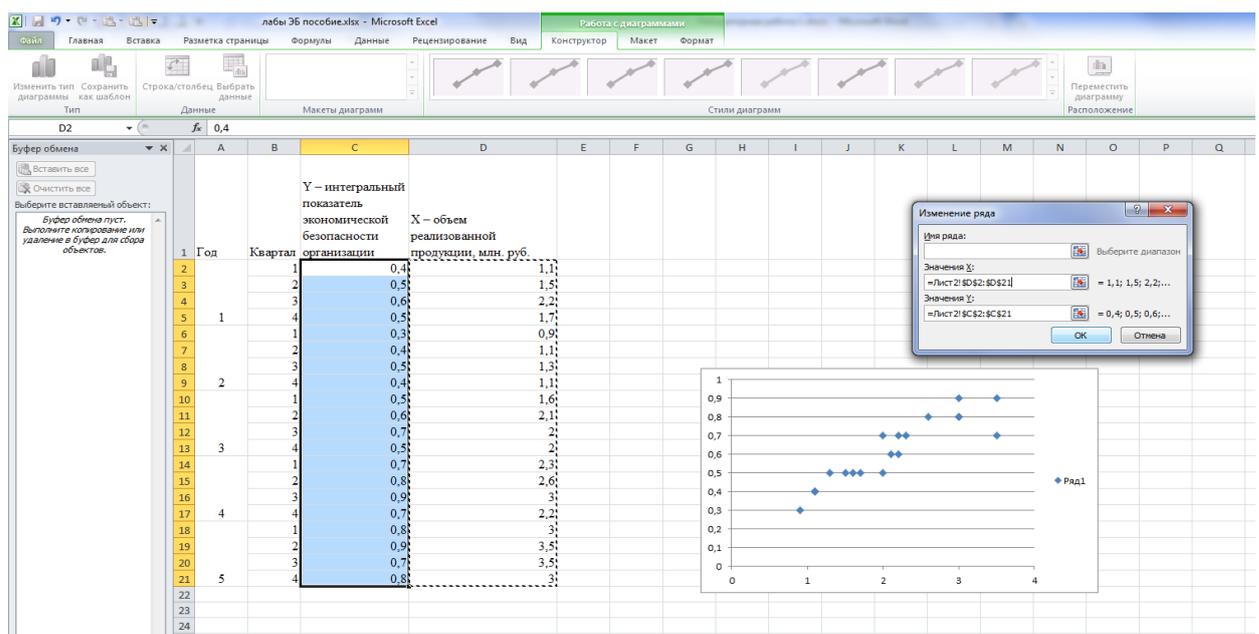


Рисунок 1.3 – Работа с полем «Изменение ряда»

Для подписи названия осей графика необходимо во вкладке «Работа с диаграммами» зайти в «Макет» – «Названия осей». Для подписи оси абсцисс выбираем «Название основной горизонтальной оси» (рисунок 1.4).

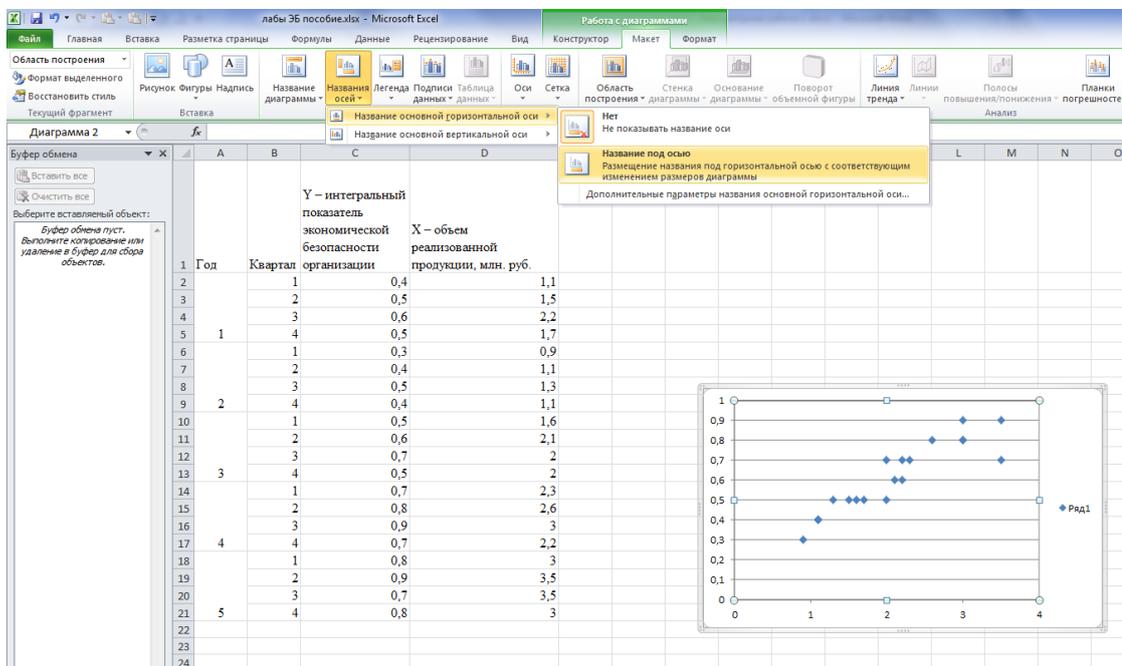


Рисунок 1.4 – Подпись горизонтальной оси

Для подписи оси ординат выбираем «Название основной вертикальной оси» (рисунок 1.5).

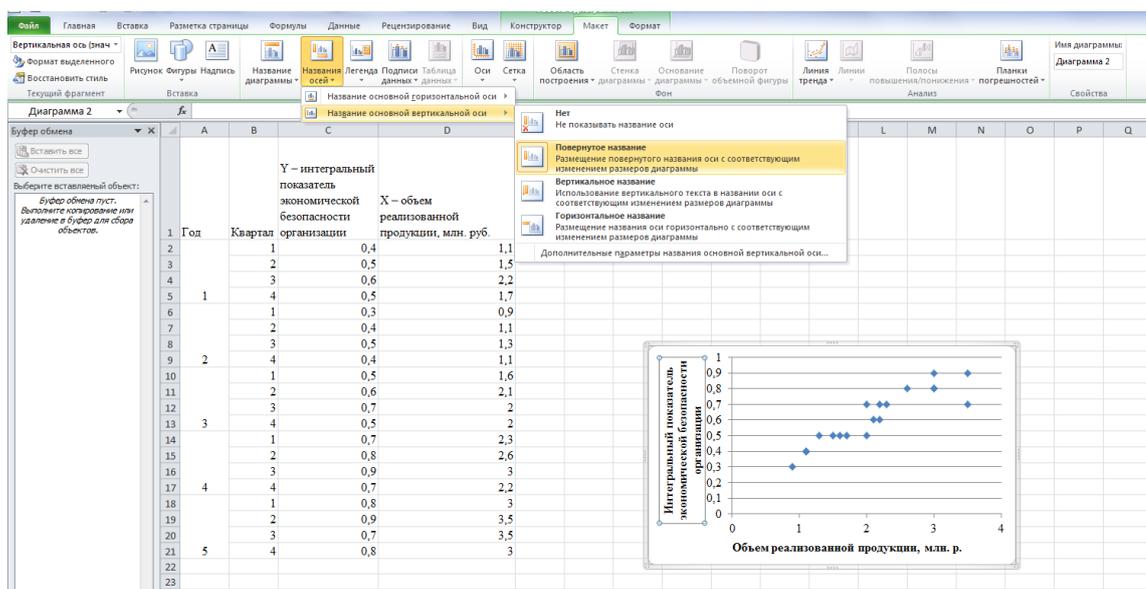


Рисунок 1.5 – Подпись вертикальной оси

На рисунке 1.6 представлены результаты построения «Поля корреляции», отражающего зависимость между эндогенной и экзогенной переменными.

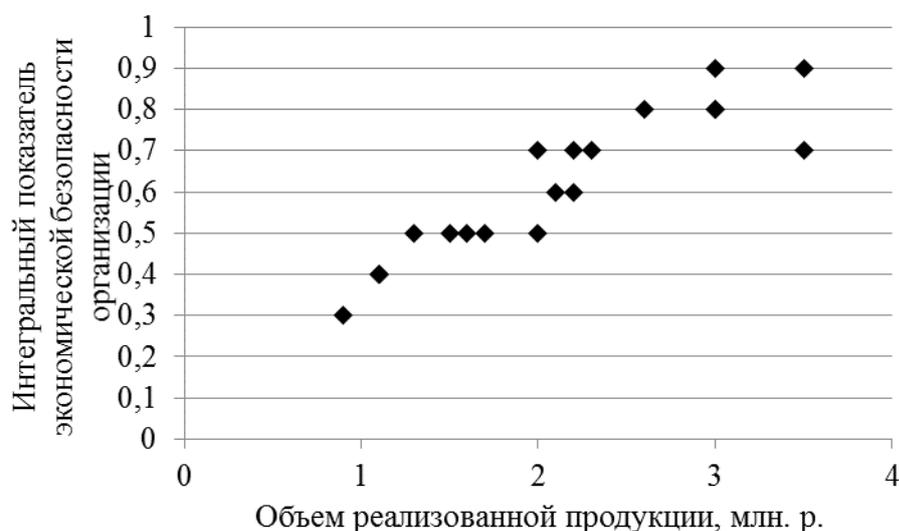


Рисунок 1.6 – Поле корреляции

По расположению точек на поле корреляции можно предположить наличие прямой линейной связи между интегральным показателем экономической безопасности организации и объемом реализованной продукции.

2 Рассчитайте оценки параметров уравнения парной линейной регрессии

Построим уравнение парной линейной регрессии с помощью MS Excel.

Для построения модели можно воспользоваться инструментом анализа данных **Регрессия**. В главном меню выберите **Данные / Анализ данных / Регрессия** (рисунок 1.7). Щелкните по кнопке **ОК**;

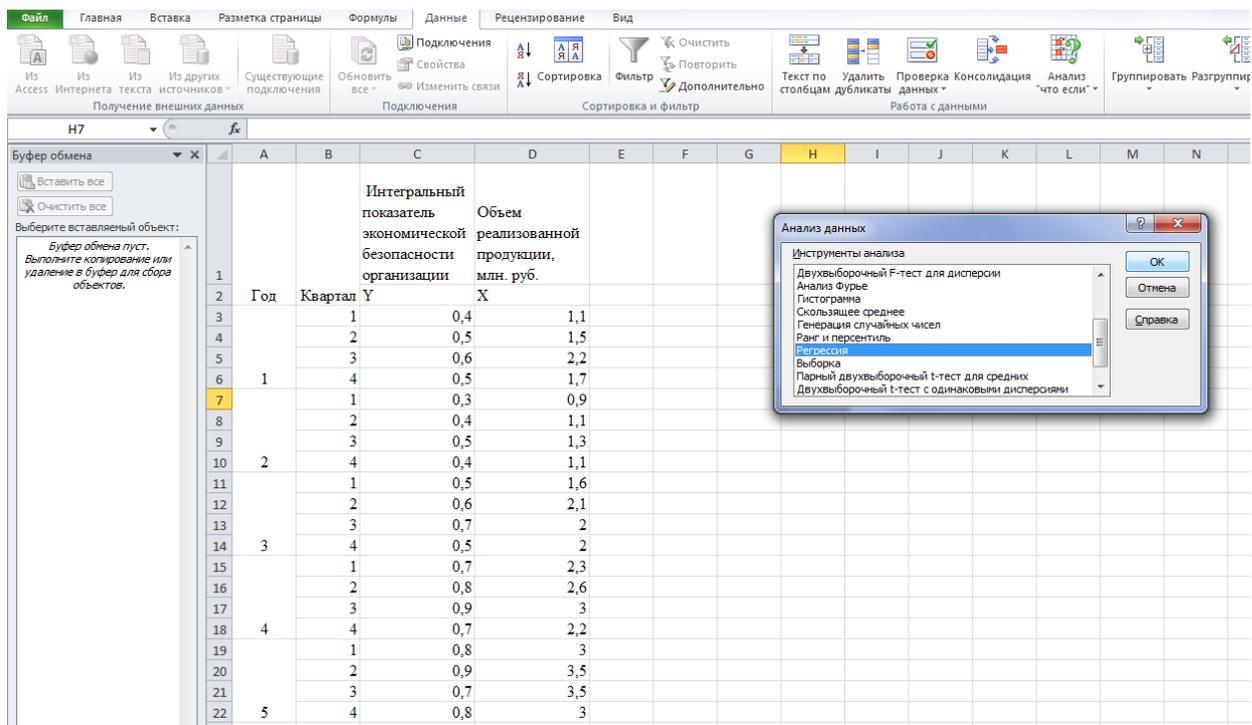


Рисунок 1.7 – Работа в диалоговом окне «Анализ данных»

Заполните диалоговое окно ввода данных и параметров ввода (рисунок 1.8):

Входной интервал Y – диапазон, содержащий данные эндогенной переменной;

Входной интервал X – диапазон, содержащий данные экзогенной переменной;

Метки – флажок, который указывает, содержит ли первая строка названия столбцов или нет;

Константа – ноль – флажок, указывающий на наличие или отсутствие свободного члена в уравнении;

Выходной интервал – достаточно указать левую верхнюю ячейку будущего диапазона;

Новый рабочий лист - можно задать произвольное имя нового листа.

ВАЖНО: удобно выделять показатели вместе с их названиями, но при этом необходимо ставить галочку в поле «Метки».

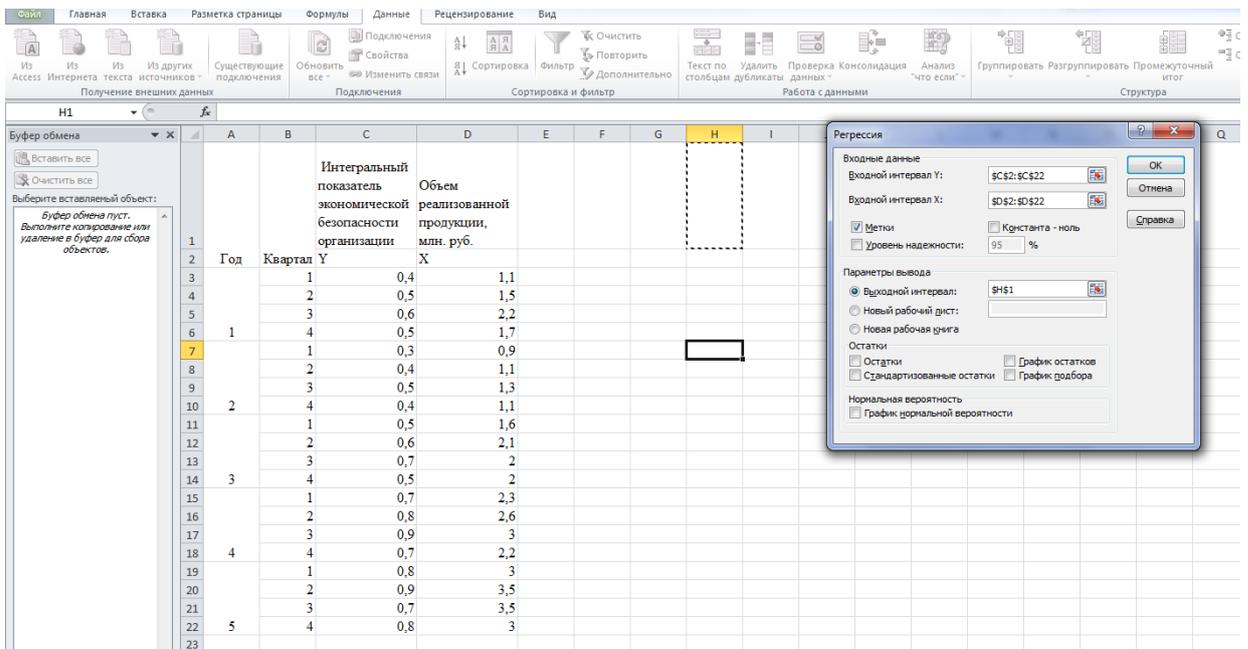


Рисунок 1.8 – Работа в диалоговом окне «Регрессия»

Результаты регрессионного анализа представлены на рисунке 1.9.

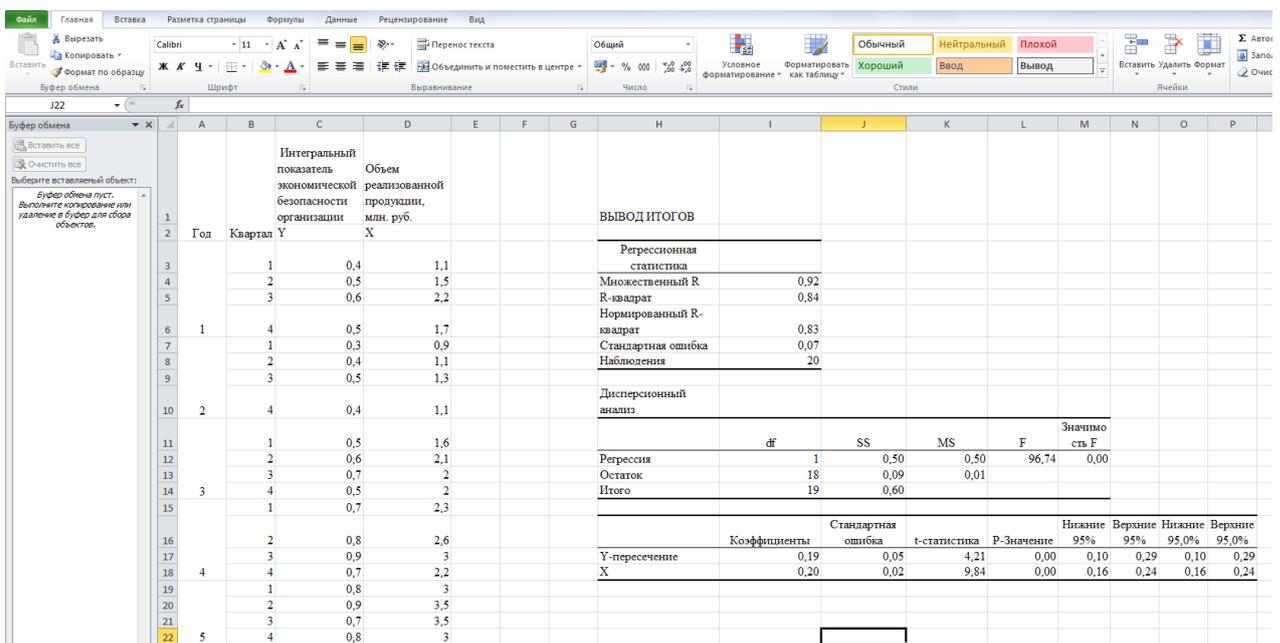


Рисунок 1.9 – Результат применения инструмента «Регрессия»

Составим уравнение парной регрессии:

$$\tilde{y}_i = 0,19 + 0,20x_i.$$

Коэффициент регрессии показывает среднее изменение эндогенной переменной с изменением на 1 единицу своего измерения экзогенной переменной.

Таким образом, коэффициент регрессии при x показывает, что с увеличением объема реализованной продукции на 1 млн. рублей, интегральный показатель экономической безопасности организации увеличится на 0,20.

Параметр b_0 экономической интерпретации не имеет.

3 Оцените тесноту связи между Y и X с помощью выборочного коэффициента корреляции. Проверьте значимость коэффициента корреляции ($\alpha = 0,05$)

Выборочный коэффициент корреляции можно рассчитать, используя инструмент анализа данных **Корреляция**. В главном меню последовательно выберите **Данные/ Анализ данных/ Корреляция** (рисунок 1.10). Щелкните по кнопке **ОК**.

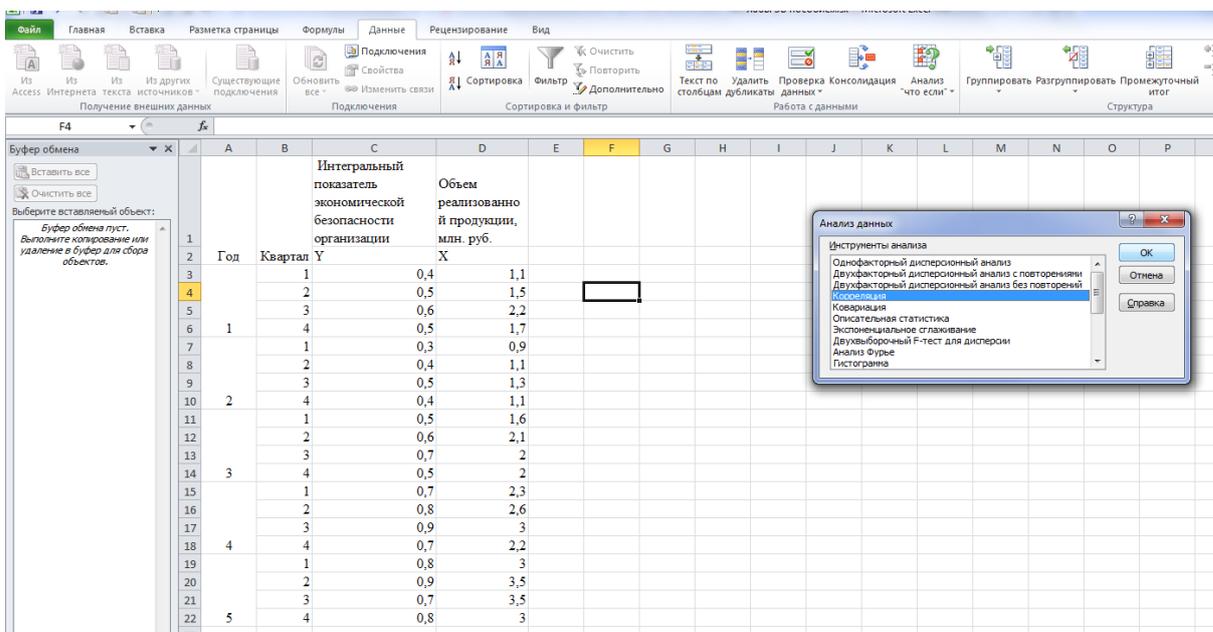


Рисунок 1.10 – Работа в диалоговом окне «Анализ данных»

Заполните диалоговое окно ввода данных и параметров вывода (рисунок 1.11).

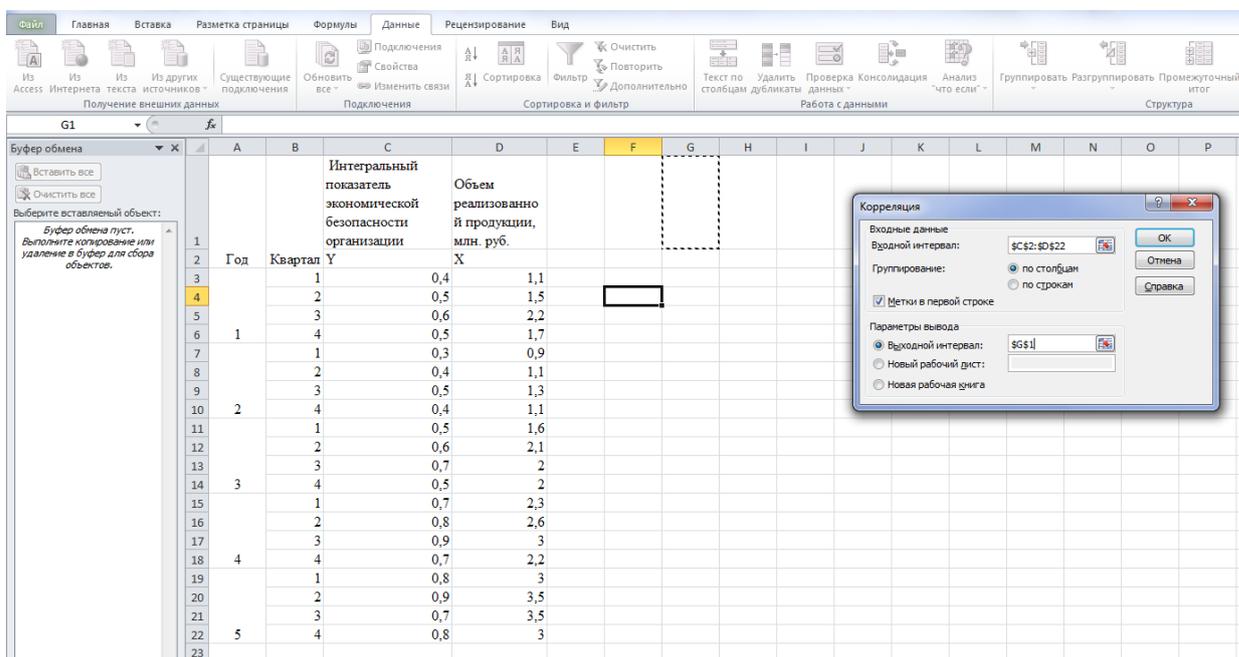


Рисунок 1.11 – Диалоговое окно ввода параметров инструмента «Корреляция»

Результаты вычислений – матрица коэффициентов парной корреляции – представлены на рисунке 1.12.

	Y	X
Y	1	
X	0,918215	1

Рисунок 1.12 – Матрица коэффициентов парной корреляции

Матрица симметрична относительно главной диагонали, поэтому приводится только одна из половин. В первом столбце стоят значения парных коэффициентов корреляции, отражающие зависимость между Y и фактором X.

Значение выборочного коэффициента корреляции составило $r_{xy}=0,92$. Проверим его статистическую значимость.

Значимость парных коэффициентов корреляции проверяют по t -критерию Стьюдента. Выдвигается гипотеза о равенстве нулю генерального коэффициента корреляции:

$$H_0 : \rho_{yx_i} = 0 ;$$

$$H_1 : \rho_{yx_i} \neq 0 .$$

Затем задаются параметры: уровень значимости α и число степеней свободы $k = n - 2$. Используя эти параметры по таблице критических точек распределения Стьюдента, находят $t_{кр}$, а по имеющимся данным вычисляют наблюдаемое значение критерия:

$$t_{набл} = \frac{r\sqrt{n-2}}{\sqrt{1-r^2}} ,$$

где r – парный коэффициент корреляции, рассчитанный по отобранным для исследования данным.

Парный коэффициент корреляции считается значимым (гипотеза о равенстве коэффициента нулю отвергается) с доверительной вероятностью $1 - \alpha$, если $t_{набл}$ по модулю будет больше, чем $t_{кр}$.

Вычислим наблюдаемые значения t -критерия Стьюдента:

$$t_{набл} = \frac{0,92\sqrt{20-2}}{\sqrt{1-0,92^2}} = 9,84 .$$

Так как наблюдаемые значения t -критерия превышают его критическое значение на 5 % уровне значимости с 18 степенями свободы $t_{крит}(0,05;18)=2,1$ (таблица Б.1), можно сделать вывод о значимости коэффициента корреляции между эндогенной и экзогенной переменными переменной, т.е. выполняется альтернативная гипотеза.

Коэффициент корреляции близок к 1, значит связь между признаками тесная, а положительный знак указывает на прямую зависимость между интегральным показателем экономической безопасности организации и объемом реализованной продукции, что подтверждается экономической теорией.

4 Рассчитайте выборочный коэффициент детерминации. Сделайте экономический вывод

Для расчета выборочного коэффициента детерминации возведем коэффициент корреляции в квадрат:

$$R_{yx}^2 = r_{yx}^2 = 0,92^2 = 0,84.$$

Таким образом, в среднем 84 % вариации интегрального показателя экономической безопасности организации объясняется вариацией объема реализованной продукции, а 16 % зависит от вариации неучтенных в модели факторов.

5 Проверьте значимость оценки коэффициента регрессии с помощью критерия Стьюдента при уровне значимости $\alpha = 0,05$

Анализ статистической значимости коэффициента регрессии b_j осуществляется по схеме статистической проверки гипотез. Проверяют гипотезу $H_0 : \beta_j = 0$ (j -я независимая переменная не влияет на результат) при альтернативной гипотезе $H_1 : \beta_j \neq 0$ (j -я независимая переменная влияет на результат).

Для проверки гипотезы используется t -статистика:

$$t = \frac{b_j}{m_{b_j}},$$

где $m_{b_j} = \sqrt{\frac{\sum_{i=1}^n (y_i - \tilde{y}_i)^2}{(n-p-1)\sum_{i=1}^n (x_i - \bar{x})^2}}$ – стандартная ошибка коэффициента регрессии

которая при справедливости H_0 имеет распределение Стьюдента с числом степеней свободы $n-p-1$.

При требуемом уровне значимости α наблюдаемое значение t -статистики сравнивается с критической точкой $t_{кр} = t_{\alpha, n-p-1}$ распределения Стьюдента.

Если $|t_{факт}| > t_{крит}$, то коэффициент считается статистически значимым, т. е. гипотеза $H_0 : \beta_j = 0$ отклоняется.

В противном случае ($|t_{факт}| < t_{крит}$) коэффициент считается статистически незначимым (статистически близким к нулю). Это означает, что фактор x_j линейно не связан с зависимой переменной y . Поэтому после установления того факта, что коэффициент b_j статистически незначим, рекомендуется исключить из уравнения регрессии переменную x_j .

Проверим значимость оценки параметра регрессии β_j с помощью критерия Стьюдента.

Выдвигаем две гипотезы:

$$H_0 : \beta_j = 0;$$

$$H_1 : \beta_j \neq 0.$$

Значение стандартной ошибки параметра b с учетом округления составило $m_b=0,020$ (рисунок 1.9).

Оно показывает, какое значение данной характеристики сформировались под влиянием случайных факторов. Эти значения используются для расчета t -критерия Стьюдента (рисунок 1.9): $t_b=9,835$.

Так как наблюдаемые значения t -критерия превышают его критическое значение на 5 % уровне значимости с 18 степенями свободы $t_{крит}(0,05;18)=2,1$ (таблица Б.1), можно сделать вывод о существенности параметра, который

формируется под воздействием неслучайных причин. Здесь коэффициент регрессии является статистически значимыми, т.е. выполняется альтернативная гипотеза.

На это же указывает показатель вероятности случайных значений параметров регрессии: если α меньше принятого нами уровня (обычно 0,1; 0,05 или 0,01), делают вывод о неслучайной природе данного значения параметра, т.е. о том, что он статистически значим и надежен. В противном случае принимается гипотеза о случайной природе значения коэффициентов уравнения.

6 Постройте доверительный интервал для коэффициента регрессии. Дайте экономическую интерпретацию

Рассчитаем доверительный интервал для коэффициента регрессии, для чего определим предельную ошибку для параметра b .

$$\Delta_b = t_{\text{крит}} \cdot m_b = 2,1 \cdot 0,02 = 0,04.$$

Доверительные интервалы: $\gamma_b = b \pm \Delta_b = 0,2 \pm 0,04$, т.е. $0,158 \leq b \leq 0,243$.

Анализ верхней и нижней границ доверительного интервала приводит к выводу о том, что с вероятностью 95 % коэффициент регрессии, находясь в указанных границах, не принимает нулевых значение, т.е. не является статистически незначимым и существенно отличен от нуля. Значения доверительного интервала позволяют сделать вывод, что при увеличении объем реализованной продукции на 1 млн. рублей, интегральный показатель экономической безопасности организации с вероятностью 95 % увеличится в среднем от 0,158 до 0,243.

7 Составьте таблицу дисперсионного анализа

Результаты дисперсионного анализа приведены в таблице 1.1 (значения для заполнения таблицы на рисунке 1.9).

Таблица 1.1 – Таблица дисперсионного анализа

Вариация результата	Число степеней свободы	Сумма квадратов отклонений	Дисперсия на одну степень свободы	F – критерий (набл.)
Факторная	1	0,504	0,504	96,736
Остаточная	18	0,094	0,005	
Общая	19	0,598		

8 Оцените с помощью F-критерия Фишера – Снедекора значимость уравнения линейной регрессии ($\alpha = 0,05$)

Наиболее часто в практических расчетах для оценки качества всего уравнения, в целом, применяется коэффициент детерминации R^2 .

Оценка значимости уравнения регрессии осуществляется путем проверки гипотезы о статистической значимости коэффициента детерминации:

$$H_0 : R^2 = 0 ,$$

$$H_1 : R^2 > 0 .$$

Для проверки гипотезы используется следующая F– статистика:

$$F = \frac{R^2}{1 - R^2} \cdot \frac{n - p - 1}{p}$$

или

$$F = \frac{\sum(\tilde{y}_i - \bar{y})^2}{p} \cdot \frac{\sum(y - \tilde{y}_i)^2}{n - p - 1} .$$

Проверка данной гипотезы равносильна проверке гипотезы о статистической незначимости уравнения регрессии:

$$H_0 : b = 0$$

(коэффициент линейной регрессии, за исключением свободного члена, равен нулю). Ведь если коэффициенты равны нулю для генеральной совокупности, то уравнение регрессии должно иметь вид $\tilde{y}_i = \bar{y}$, а коэффициент детерминации R^2 и F- статистика Фишера также равны нулю. При этом их оценки для случайной выборки, конечно, отличаются от нуля.

Величина F при выполнении предпосылок МНК и при справедливости имеет распределение Фишера. При заданном уровне значимости α по таблицам критических точек распределения Фишера находится критическое значение $F_{кр} = F_{\alpha;p;n-p-1}$. Если $F_{набл} > F_{кр}$, то основную гипотезу отвергают и принимают альтернативную гипотезу о статистической значимости уравнения регрессии. Если $F_{набл} < F_{кр}$, то основную гипотезу о незначимости уравнения регрессии не отвергают (уравнение модели признается незначимым).

Оценим значимость уравнения линейной регрессии, построенного в задании 3.

Для проверки значимости уравнения выдвигаем две гипотезы:

$$H_0 : b = 0 \text{ и } R_{yx}^2 = 0;$$

$$H_1 : b \neq 0 \text{ и } R_{yx}^2 \neq 0.$$

Оценим значимость уравнения линейной регрессии, построенного в задании 2.

По данным таблицы дисперсионного анализа (таблица 1.1), $F_{набл} = 96,736$, критическое значение $F_{крит}(0,05;1;18) = 4,41$ (приложение Б), т.е. выполняется неравенство факт $F_{крит} < F_{набл}$, т.е. именно принимаем гипотезу H_1 .

Следовательно, полученное значение не случайно, оно сформировалось под влиянием существенных экзогенных переменных, т.е. подтверждается статистическая значимость всего уравнения, показателя тесноты связи, уравнение можно использовать для прогнозирования. Кроме того, вероятность случайно получить такое значение F-критерия составляет 0,000, что не превышает допустимый уровень значимости 5 %; об этом свидетельствует величина «Значимость F» (рисунок 1.9).

9 *Рассчитайте интегральный показатель экономической безопасности организации, если объем реализованной продукции снизится на 10 % относительно последнего периода. Постройте доверительный интервал для прогнозного значения эндогенной переменной. Сделайте экономический вывод*

Определим прогнозное значение экзогенной переменной.

Значение экзогенной переменной для последнего периода (4 квартал 5 года) составляет равно 3 млн. р.

Тогда, согласно заданию, $x_p = 3 - 0,1 \cdot 3 = 2,7$ млн. р.

Подставляем полученное значение в уравнение регрессии:

$$\tilde{y}_p = 0,19 + 0,20 \cdot 2,7 = 0,73.$$

Чтобы получить интервальный прогноз, найдем стандартную ошибку предсказываемого значения интегрального показателя экономической безопасности организации m_p .

$$m_p = S \sqrt{1 + \frac{1}{n} + \frac{(x^p - \bar{x})^2}{\sum (x - \bar{x})^2}},$$

где $S = \sqrt{\frac{\sum (y - \tilde{y}_x)^2}{n-2}}$ – стандартная ошибка регрессии;

\tilde{y}_x – теоретические (расчетные) значения эндогенной переменной полученные по уравнению регрессии (рисунок 1.13).

	A	B	C	D	E	F
			Y – интегральный показатель экономической безопасности организации	X – объем реализованной продукции, млн. руб.	\tilde{y}_x	
1	Год	Квартал				
2		1	0,4	1,1	=0,19+0,2*D2	
3		2	0,5	1,5		
4		3	0,6	2,2		
5	1	4	0,5	1,7		
6		1	0,3	0,9		
7		2	0,4	1,1		
8		3	0,5	1,3		
9	2	4	0,4	1,1		
10		1	0,5	1,6		
11		2	0,6	2,1		
12		3	0,7	2		

Рисунок 1.13 – Расчет теоретических значений эндогенной переменной по уравнению регрессии

Расчета числителя формулы стандартной ошибки регрессии выполним в MS Excel (рисунок 1.14).

	A	B	C	D	E	F	G
			Y – интегральный показатель экономической безопасности организации	X – объем реализованной продукции, млн. руб.	\tilde{y}_x	$(y - \tilde{y}_x)^2$	
1	Год	Квартал					
2		1	0,4	1,1	0,41	=(C2-E2)^2	
3		2	0,5	1,5	0,49		
4		3	0,6	2,2	0,63		
5	1	4	0,5	1,7	0,53		
6		1	0,3	0,9	0,37		
7		2	0,4	1,1	0,41		
8		3	0,5	1,3	0,45		

Рисунок 1.14 – Расчет стандартной ошибки регрессии

Таблица с промежуточными расчетами представлена на рисунке 1.15.

	A	B	C	D	E	F	G	H
			Y – интегральный показатель экономической безопасности организации	X – объем реализованной продукции, млн. руб.	\bar{y}_x	$(y - \bar{y}_x)^2$	$(x - \bar{x})^2$	
1	Год	Квартал						
2		1	0,4	1,1	0,41	0,0001	0,9702	
3		2	0,5	1,5	0,49	0,0001	0,3422	
4		3	0,6	2,2	0,63	0,0009	0,0132	
5	1	4	0,5	1,7	0,53	0,0009	0,1482	
6		1	0,3	0,9	0,37	0,0049	1,4042	
7		2	0,4	1,1	0,41	0,0001	0,9702	
8		3	0,5	1,3	0,45	0,0025	0,6162	
9	2	4	0,4	1,1	0,41	0,0001	0,9702	
10		1	0,5	1,6	0,51	0,0001	0,2352	
11		2	0,6	2,1	0,61	0,0001	0,0002	
12		3	0,7	2	0,59	0,0121	0,0072	
13	3	4	0,5	2	0,59	0,0081	0,0072	
14		1	0,7	2,3	0,65	0,0025	0,0462	
15		2	0,8	2,6	0,71	0,0081	0,2652	
16		3	0,9	3	0,79	0,0121	0,8372	
17	4	4	0,7	2,2	0,63	0,0049	0,0132	
18		1	0,8	3	0,79	0,0001	0,8372	
19		2	0,9	3,5	0,89	0,0001	2,0022	
20		3	0,7	3,5	0,89	0,0361	2,0022	
21	5	4	0,8	3	0,79	0,0001	0,8372	
22	Итого		12,2	41,7	12,14	0,094	12,526	
23	В среднем		0,61	2,085				

Рисунок 1.15 – Промежуточные расчеты для оценки стандартной ошибки предсказываемого значения интегрального показателя экономической безопасности организации

Подставим полученные значения в формулы:

$$S = \sqrt{\frac{0,094}{18}} = 0,072,$$

$$m_p = 0,072 \sqrt{1 + 20 + \frac{(2,7 - 2,085)^2}{12,526}} = 0,076.$$

Предельная ошибка прогнозируемого расхода составит:

$$\Delta_{y^p} = t_{крит} \cdot m_p = 2,01 \cdot 0,076 = 0,152.$$

Доверительный интервал прогнозируемого расхода составит:

$$y_p = 0,73 \pm 0,152,$$

т.е. при объеме реализованной продукции (млн. рублей), равной 2,7 млн. рублей, интегральный показатель экономической безопасности организации составит не меньше, чем

$$y_{p \min} = 0,73 - 0,152 = 0,578$$

и не больше, чем

$$y_{p \max} = 0,73 + 0,152 = 0,882.$$

10 Определить среднюю ошибку аппроксимации

На практике, как правило имеет место некоторое рассеивание точек корреляционного поля относительно теоретической линии регрессии, т.е. отклонения эмпирических данных от теоретических. Величина этих отклонений и лежит в основе расчета показателя качества уравнения.

В качестве такого показателя выступает средняя ошибка аппроксимации, которая рассчитывается по формуле:

$$\bar{A} = \frac{1}{n} \sum \left| \frac{y - \tilde{y}_i}{y} \right| \cdot 100\%.$$

Если $|\bar{A}| < 10\%$, это свидетельствует о высокой точности модели, при $10\% \leq |\bar{A}| \leq 20\%$ – точность хорошая, при $20\% \leq |\bar{A}| \leq 50\%$ – удовлетворительная.

Необходимые данные для расчета средней ошибки аппроксимации представлены на рисунке 1.16.

	A	B	C	D	E	F	G
	Год	Квартал	Y – интегральный показатель экономической безопасности организации	X – объем реализованной продукции, млн. руб.	\tilde{y}_x	$\left \frac{y - \tilde{y}_x}{y} \right $	
2		1	0,4	1,1	0,41	0,025	
3		2	0,5	1,5	0,49	0,02	
4		3	0,6	2,2	0,63	0,05	
5	1	4	0,5	1,7	0,53	0,06	
6		1	0,3	0,9	0,37	0,2333	
7		2	0,4	1,1	0,41	0,025	
8		3	0,5	1,3	0,45	0,1	
9	2	4	0,4	1,1	0,41	0,025	
10		1	0,5	1,6	0,51	0,02	
11		2	0,6	2,1	0,61	0,0167	
12		3	0,7	2	0,59	0,1571	
13	3	4	0,5	2	0,59	0,18	
14		1	0,7	2,3	0,65	0,0714	
15		2	0,8	2,6	0,71	0,1125	
16		3	0,9	3	0,79	0,1222	
17	4	4	0,7	2,2	0,63	0,1	
18		1	0,8	3	0,79	0,0125	
19		2	0,9	3,5	0,89	0,0111	
20		3	0,7	3,5	0,89	0,2714	
21	5	4	0,8	3	0,79	0,0125	
22	Итого		12,2	41,7	12,14	1,6258	
23	В среднем					8,1292	
24							
25							

Рисунок 1.16 – Данные для расчета средней ошибки аппроксимации

$$\bar{A} = \frac{1}{n} \cdot \sum \left| \frac{y - y_{\text{теор}}}{y} \right| \cdot 100 \% = \frac{1}{20} \cdot 1,626 \cdot 100 \% = 8,13 \%$$

Таким образом, фактические значения эндогенной переменной отличаются от теоретических значений на 8,13 %. Следовательно, построенная модель имеет высокую точность.

11 На поле корреляции постройте линию регрессии

Графиком является прямая, которую можно построить по данным столбцов «Y – интегральный показатель экономической безопасности организации» и « \hat{y}_x ».

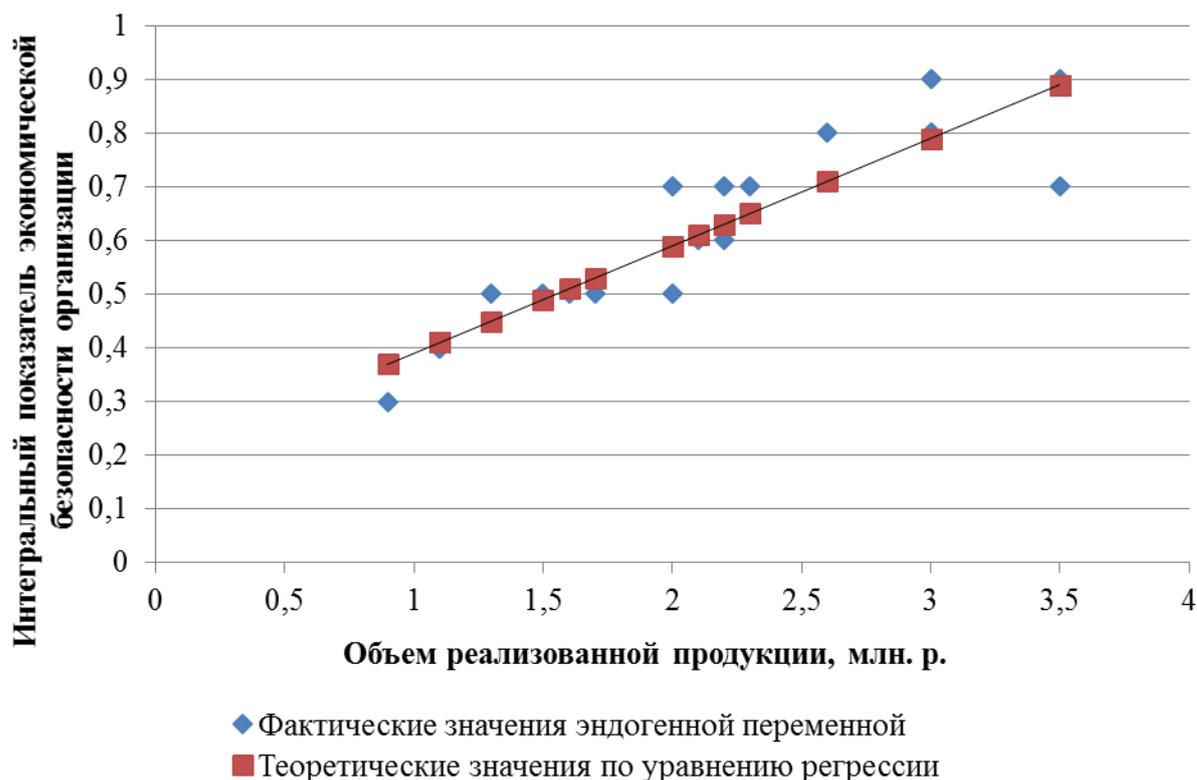


Рисунок 1.17 – График уравнения линейной регрессии

Контрольные вопросы к лабораторной работе № 1

- 1 Дайте определение парной (простой) модели линейной регрессии.
- 2 В чем суть метода наименьших квадратов?
- 3 Как проводится проверка статистической значимости уравнения регрессии и коэффициентов регрессии.
- 4 Как оценивается точность уравнения регрессии?
- 5 Как строится доверительный интервал для функции регрессии?

Тесты для самоконтроля к лабораторной работе № 1

- 1 Анализ тесноты и направления связей двух признаков осуществляется на основе:
 - А) парного коэффициента корреляции;
 - Б) среднего квадратического отклонения;
 - В) коэффициента детерминации;
 - Г) множественного коэффициента корреляции.

2 Известно, что между величинами X и Y существует отрицательная связь. В каких пределах находится парный коэффициент корреляции?

- А) от -1 до 0 ;
- Б) от $-\infty$ до $+\infty$;
- В) от 0 до 1 ;
- Г) от -1 до 1 .

3 С помощью какого метода можно найти оценки параметров уравнения линейной регрессии:

- А) методом наименьшего квадрата;
- Б) индексным методом;
- В) методом сумм;
- Г) методом аналитического выравнивания.

4 В линейном уравнении $\hat{y} = \beta_0 + \beta_1 x$ коэффициент регрессии показывает:

- А) тесноту связи;
- Б) долю дисперсии « Y », зависимую от « X »;
- В) на сколько в среднем изменится « Y » при изменении « X » на одну единицу;
- Г) ошибку коэффициента корреляции.

5 По 16 наблюдениям построено парное линейное уравнение регрессии. Для проверки значимости коэффициента регрессии вычислено $t_{\text{набл}}=2.5$.

- А) коэффициент незначим при $\alpha=0,05$;
- Б) коэффициент значим при $\alpha=0,05$;
- В) коэффициент значим при $\alpha=0,01$;
- Г) коэффициент незначим при $\alpha=0,01$.

6 Суть метода наименьших квадратов заключается в том, что:

А) оценка определяется из условия минимизации суммы квадратов отклонений выборочных данных от определяемой оценки;

Б) оценка определяется из условия минимизации суммы отклонений выборочных данных от определяемой оценки;

В) оценка определяется из условия максимизации суммы квадратов отклонений выборочных данных от определяемой оценки;

Г) оценка определяется из условия минимизации суммы квадратов отклонений выборочной средней от выборочной дисперсии.

7 Коэффициент регрессии в уравнении $\hat{y} = 9,2 + 1,5 \cdot x$, характеризующем связь между объемом реализованной продукции (млн. р.) и прибылью предприятий автомобильной промышленности за год (млн. р.) означает, что при увеличении объема реализованной продукции на 1 млн. р. прибыль увеличивается на:

А) 0,5 %;

Б) 1,5 %;

В) 500 тыс. р.;

Г) 1,5 млн. р.

8 Какое уравнение соответствует следующей системе нормальных уравнений, применяемой для определения параметров уравнения (по методу наименьших квадратов)

$$\begin{cases} \sum y_i = n \cdot a + b \cdot \sum x_i \\ \sum y_i \cdot x_i = a \cdot \sum x_i + b \cdot \sum x_i^2 \end{cases}$$

А) $\tilde{y}_i = a_0 + a_1 \cdot \frac{1}{x_i}$;

Б) $\tilde{y}_i = a_0 + a_1 \cdot x_i + a_2 \cdot x_i^2$;

В) $\tilde{y}_i = a_0 + a_1 \cdot x_i^2$;

Г) $\tilde{y}_i = a_0 + a_1 \cdot x_i$.

9 Сущность метода наименьших квадратов состоит в:

А) $\sum_{i=1}^n |y_i - \tilde{y}_i| \rightarrow \min ;$

Б) $\sum_{i=1}^n (y_i - \tilde{y}_i)^2 \rightarrow \max ;$

В) $\sum_{i=1}^n (y_i - \tilde{y}_i)^2 \rightarrow \min ;$

Г) $\sum_{i=1}^n (y_i - \tilde{y}_i) \rightarrow \min .$

10 Эконометрическая модель – это модель:

А) гипотетического экономического объекта;

Б) теоретического экономического объекта;

В) конкретно-существующего экономического объекта, построенная на гипотетических данных;

Г) конкретно-существующего экономического объекта, построенная на статистических данных.

2 Лабораторная работа № 2. Классическая модель линейной регрессии: построение множественной линейной регрессии

Цель выполнения лабораторной работы: научиться строить статистически значимое уравнение множественной линейной регрессии, уравнение регрессии в стандартизованном масштабе, разрабатывать факторный прогноз и интерпретировать полученные результаты.

Задания:

Для эндогенной переменной Y – интегральный показатель экономической безопасности организации и экзогенных переменных $X_1 – X_5$ выполните следующие задания:

1 Рассчитайте параметры линейного уравнения множественной регрессии с полным перечнем факторов.

2 Рассчитайте матрицу парных коэффициентов корреляции и отберите информативные факторы в модели. Проверьте значимость парных коэффициентов корреляции. Укажите коллинеарные факторы.

3 Постройте модель в естественной форме только с информативными факторами. Оцените качество построенного уравнения регрессии.

4 Оцените с помощью F-критерия Фишера-Снедекора значимость уравнения линейной регрессии и показателя тесноты связи.

5 Оцените статистическую значимость коэффициентов регрессии с помощью t- критерия Стьюдента.

6 Оцените качество уравнения через среднюю ошибку аппроксимации.

7 Постройте модель в стандартизованном масштабе и проинтерпретируйте ее параметры.

8 Рассчитайте прогнозное значение результата, если прогнозное значение факторов составляют 80 % от их максимальных значений.

9 Рассчитайте ошибки и доверительный интервал прогноза для уровня значимости $\alpha = 0,05$.

10 По полученным результатам сделайте экономический вывод.

Примечание: Y – интегральный показатель экономической безопасности организации;

X_1 – Объем реализованной продукции, млн. рублей;

X_2 –оборотный капитал, млн. рублей;

X_3 – Основной капитал, млн. рублей;

X_4 – Производительность труда, рублей на чел.;

X_5 – Объем привлеченных инвестиций, млн. рублей.

Реализация типовых задач

1 Рассчитайте параметры линейного уравнения множественной регрессии с полным перечнем факторов по данным о деятельности крупнейших компаний в текущем году (условные данные)

Построим уравнение множественной линейной регрессии.

Для этого проведем регрессионный анализ данных факторов с помощью ППП MS Excel.

Для построения модели можно воспользоваться инструментом анализа данных **Регрессия**. Порядок действий, следующий:

а) в главном меню выберите **Данные / Анализ данных / Регрессия**. Щелкните по кнопке **ОК**;

б) заполните диалоговое окно ввода данных и параметров ввода (рисунок 2.1):

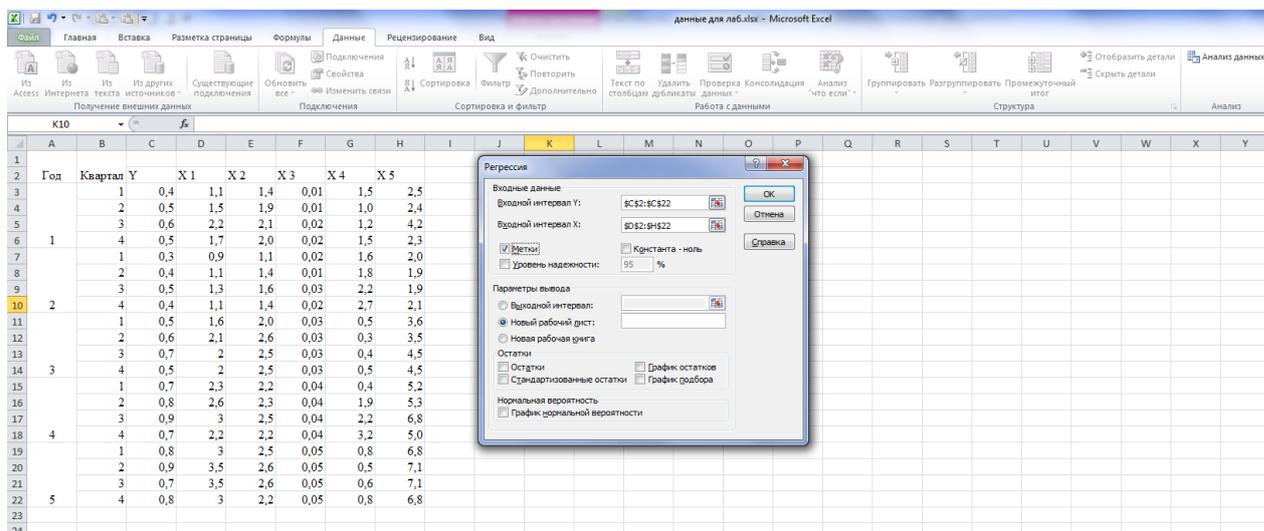


Рисунок 2.1 – Диалоговое окно ввода параметров инструмента «Регрессия»

Входной интервал Y – диапазон, содержащий данные результивного признака;

Входной интервал X – диапазон, содержащий данные всех пяти факторов;

Метки – флажок, который указывает, содержит ли первая строка названия столбцов или нет;

Константа – ноль – флажок, указывающий на наличие или отсутствие свободного члена в уравнении;

Выходной интервал – достаточно указать левую верхнюю ячейку будущего диапазона;

Новый рабочий лист - можно задать произвольное имя нового листа.

Результаты регрессионного анализа представлены на рисунке 2.2.

Регрессионная статистика									
Множественный R		0,942071954							
R-квадрат		0,887499567							
Нормированный R-квадрат		0,847320841							
Стандартная ошибка		0,069320817							
Наблюдения		20							
Дисперсионный анализ									
		df	SS	MS	F	Значимость F			
Регрессия		5	0,530724741	0,106144948	22,08879307	3,47656E-06			
Остаток		14	0,067275259	0,004805376					
Итого		19	0,598						
	Коэффициенты	Стандартная ошибка	t-статистика	P-значение	Нижние 95%	Верхние 95%	Нижние 95,0%	Верхние 95,0%	
Y-пересечение	0,003307416	0,109425131	0,030225378	0,976313971	-0,231386148	0,23800098	-0,231386148	0,23800098	
X 1	0,078381508	0,085531339	0,916406883	0,374969807	-0,10506497	0,261827986	-0,10506497	0,261827986	
X 2	0,115574523	0,071016937	1,627421972	0,125936225	-0,036741657	0,267890703	-0,036741657	0,267890703	
X 3	-0,709321494	2,757035776	-0,257276855	0,800707871	-6,622575124	5,203932136	-6,622575124	5,203932136	
X 4	0,043393858	0,022451464	1,932785254	0,073749844	-0,004759743	0,091547459	-0,004759743	0,091547459	
X 5	0,039680751	0,035638399	1,113426878	0,284277188	-0,036756012	0,116117514	-0,036756012	0,116117514	

Рисунок 2.2 – Результат применения инструмента «Регрессия»

Составим уравнение множественной регрессии:

$$\tilde{y}_i = 0,003 + 0,078x_1 + 0,116x_2 + 0,709x_3 + 0,043x_4 + 0,04x_5.$$

Коэффициенты регрессии показывают среднее изменение результативного признака с изменением на 1 единицу своего измерения данного фактора при условии постоянства всех остальных.

Таким образом, коэффициент регрессии при x_1 показывает, что при увеличении объема реализованной продукции на 1 млн. рублей, интегральный показатель экономической безопасности увеличится на 0,078, при фиксированном значении остальных факторов.

Аналогичным образом делаются выводы по остальным коэффициентам регрессии.

Параметр b_0 экономической интерпретации не имеет.

2 Рассчитайте матрицу парных коэффициентов корреляции и отберите информативные факторы в модели. Проверьте значимость парных коэффициентов корреляции. Укажите коллинеарные факторы

Матрицу парных коэффициентов корреляции переменных можно рассчитать, используя инструмент анализа данных **Корреляция**. Для этого:

1) в главном меню последовательно выберите пункты **Данные / Анализ данных / Корреляция**. Щелкните по кнопке **ОК**;

2) заполнит диалоговое окно ввода данных и параметров вывода (рисунок 2.3);

3) результаты вычислений – матрица коэффициентов парной корреляции – представлены на рисунке 2.4.

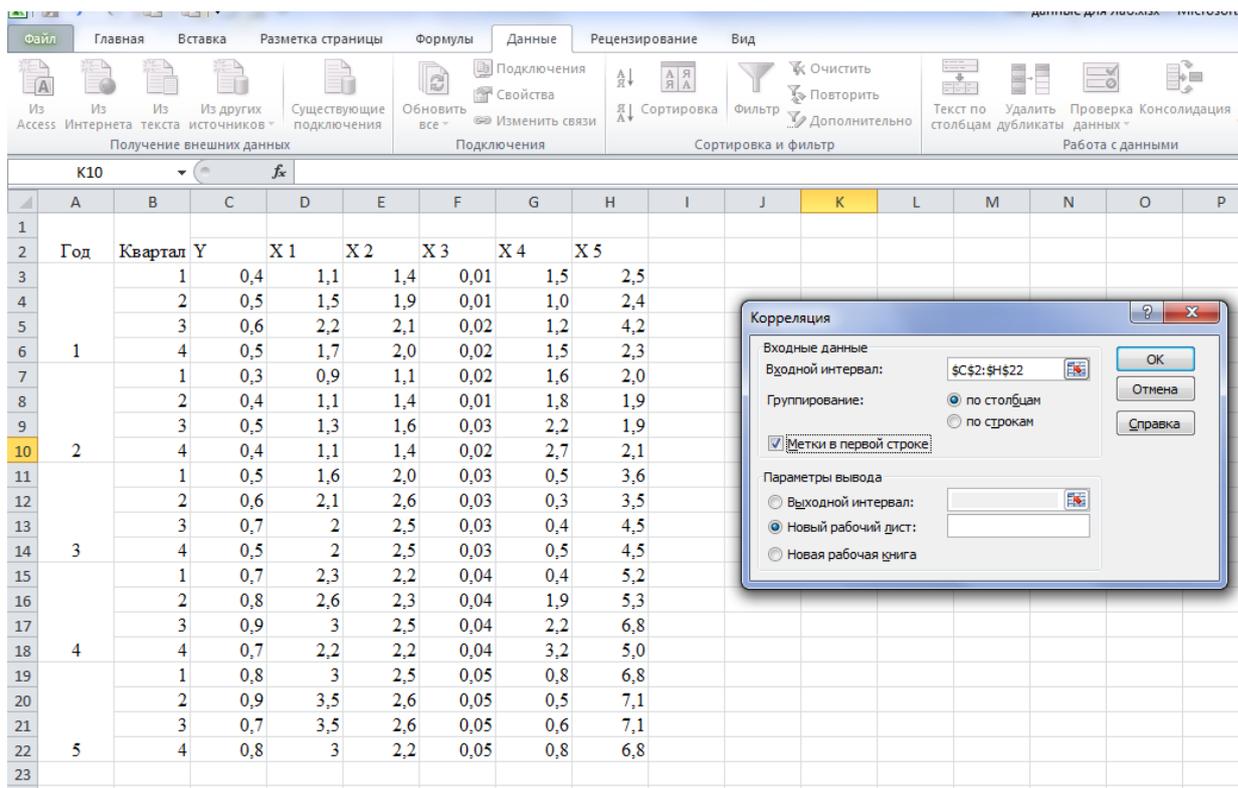


Рисунок 2.3 – Диалоговое окно ввода параметров инструмента «Корреляция»

	Y	X 1	X 2	X 3	X 4	X 5
Y	1					
X 1	0,918215	1				
X 2	0,813413	0,846627	1			
X 3	0,826695	0,86741	0,73281	1		
X 4	-0,18443	-0,33205	-0,48868	-0,19361	1	
X 5	0,908832	0,964957	0,798673	0,893305	-0,31165	1

Рисунок 2.4 – Матрица коэффициентов парной корреляции

Проверим значимость парных коэффициентов корреляции между эндогенными и экзогенными переменными.

Выдвигаем две гипотезы:

$$H_0 : \rho_{yx_i} = 0 ;$$

$$H_1 : \rho_{yx_i} \neq 0 .$$

Вычислим наблюдаемые значения t-критерия Стьюдента по формуле:

$$t_{набл} = \frac{r}{\sqrt{1-r^2}} \sqrt{n-2} .$$

$$t_{r_{yx1}} = \frac{0,92}{\sqrt{1-0,92^2}} \sqrt{18} = 24,83 ,$$

$$t_{r_{yx2}} = \frac{0,81}{\sqrt{1-0,81^2}} \sqrt{18} = 10,2 ,$$

$$t_{r_{yx3}} = \frac{0,83}{\sqrt{1-0,83^2}} \sqrt{18} = 11,08 ,$$

$$t_{r_{yx4}} = \frac{-0,18}{\sqrt{1-0,18^2}} \sqrt{18} = -0,81 ,$$

$$t_{r_{yx5}} = \frac{0,91}{\sqrt{1-0,91^2}} \sqrt{18} = 22,16 .$$

Так как наблюдаемые значения t-критерия превышают его критическое значение на 5 % уровне значимости с 18 степенями свободы $t_{крит(0,05;18)} = 2,1$

(таблице Б.1), можно сделать вывод о значимости коэффициентов корреляции между эндогенной переменной y с экзогенными переменными x_1, x_2, x_3, x_5 , т.е. выполняется альтернативная гипотеза. Коэффициент корреляции r_{yx_4} является статистически незначимым, т.к. $|t_{r_{yx_4}}| = 0,85 < t_{\text{кр}} = 2,1$, т.е. выполняется гипотеза H_0 . Таким образом, экзогенная переменная x_4 из уравнения множественной регрессии, построенной в задании 1, исключается.

Из матрицы парных коэффициентов корреляции можно заметить, что факторы x_1 и x_2 , x_1 и x_3 , x_1 и x_5 , x_3 и x_2 , x_3 и x_5 мультиколлинеарны, т.к. коэффициенты корреляции между ними превышают 0,75. Таким образом, можно сказать, что они дублируют друг друга, т.е. факторы x_1 и x_3 тоже исключаются из уравнения множественной регрессии.

В нашем примере получаем, что информативными факторами являются: x_2 и x_5 .

Построим новое уравнение множественной регрессии с информативными факторами.

3 Постройте модель в естественной форме только с информативными факторами. Оцените качество построенного уравнения регрессии

Построим уравнение множественной линейной регрессии с информативными факторами следующего вида:

$$\tilde{y}_i = a + b_2x_2 + b_5x_5.$$

Параметры вычисляем аналогично заданию 1 (рисунок 2.5).

	A	B	C	D	E	F	G	H	I
1	Вывод итогов								
2									
3	<i>Регрессионная статистика</i>								
4	Множественный R	0,881199186							
5	R-квадрат	0,776512005							
6	Нормированный R-квадрат	0,7502193							
7	Стандартная ошибка	0,088665207							
8	Наблюдения	20							
9									
10	Дисперсионный анализ								
11		df	SS	MS	F	Значимость F			
12	Регрессия	2	0,464354179	0,23217709	29,5333629	2,94212E-06			
13	Остаток	17	0,133645821	0,007861519					
14	Итого	19	0,598						
15									
16		Коэффициенты	Стандартная ошибка	t-статистика	P-Значение	Нижние 95%	Верхние 95%	Нижние 95,0%	Верхние 95,0%
17	Y-пересечение	0,054836109	0,095800019	0,572401862	0,574542404	-0,147284263	0,256956482	-0,147284263	0,256956482
18	X 2	0,16855401	0,063342444	2,660996305	0,016459967	0,034913134	0,302194885	0,034913134	0,302194885
19	X 3	6,645015549	2,247974672	2,956001076	0,008845818	1,902203568	11,38782753	1,902203568	11,38782753
20									

Рисунок 2.5 – Результат применения инструмента Регрессия

Получаем уравнение следующего вида: $\tilde{y}_i = 0,055 + 0,169x_2 + 0,645x_3$.

Выводы по коэффициентам регрессии делаются аналогично заданию 1.

Коэффициент множественной корреляции $R_{yx_2x_3} = 0,88$ свидетельствует о тесной связи между эндогенной и экзогенными переменными, вошедшими в уравнение регрессии. Коэффициент множественной детерминации $R^2_{yx_2x_3} = 0,78$, показывает, что 78 % вариации интегрального показателя экономической безопасности организации (y) обусловлено вариацией оборотного капитала (x_2) и основного капитала (x_3), а влияние неучтенных в уравнении факторов составляет 22 %.

Скорректированный коэффициент множественной детерминации $\bar{R}^2_{yx_2x_3} = 0,75$ определяет тесноту связи с учетом степеней свободы общей и остаточной дисперсий. Он дает такую оценку тесноты связи, которая не зависит от числа экзогенных переменных в модели и потому может сравниваться по разным моделям с разным числом экзогенных переменных.

4 Оцените с помощью F-критерия Фишера-Снедекора значимость уравнения линейной регрессии и показателя тесноты связи

Для проверки значимости уравнения выдвигаем две гипотезы:

$$H_0 : b_2 = b_3 = 0 \text{ и } R^2_{yx_2x_3} = 0;$$

$$H_1 : b_2 \neq b_3 \neq 0 \text{ и } R_{yx_2x_3}^2 \neq 0.$$

По данным таблицы дисперсионного анализа, представленной на рисунке 5, $F_{\text{факт}} = 29,53$, критическое значение $F_{\text{крит}(0,05;2;17)} = 3,59$ (приложение Б), т.е. выполняется неравенство $F_{\text{факт}} > F_{\text{крит}}$, а именно принимаем гипотезу H_1 . Следовательно, полученное значение не случайно, оно сформировалось под влиянием существенных экзогенных переменных, т.е. подтверждается статистическая значимость всего уравнения, показателя тесноты связи $R_{yx_2x_3}$ и уравнение можно использовать для прогнозирования. Кроме того, вероятность случайно получить такое значение F-критерия составляет 0,0000, что не превышает допустимый уровень значимости 5 %; об этом свидетельствует величина «Значимость F» (рисунок 2.5).

5 Оцените статистическую значимость коэффициентов регрессии с помощью t- критерия Стьюдента

Поскольку уравнение регрессии определяется на основе выборочных данных, то коэффициенты этого уравнения являются точечными оценками (случайными величинами), значения которых изменяются от выборки к выборке. А значит, необходима проверка значимости этих коэффициентов.

Для оценки статистической значимости коэффициентов регрессии рассчитывают доверительные интервалы и t – критерии.

Анализ статистической значимости коэффициента регрессии b_j осуществляется по схеме статистической проверки гипотез. Проверяют гипотезу $H_0 : \beta_j = 0$ (j -я независимая переменная не влияет на результат) при альтернативной гипотезе $H_1 : \beta_j \neq 0$ (j -я независимая переменная влияет на результат).

Для проверки гипотезы используется t -статистика:

$$t = \frac{b_j}{m_{b_j}},$$

где $m_{b_j} = \sqrt{\frac{\sum_{i=1}^n (y_i - \tilde{y}_i)^2}{(n-p-1) \sum_{i=1}^n (x_i - \bar{x})^2}}$ – стандартная ошибка коэффициента

регрессии, которая при справедливости H_0 имеет распределение Стьюдента с числом степеней свободы $n - p - 1$.

При требуемом уровне значимости α наблюдаемое значение t -статистики сравнивается с критической точкой $t_{кр} = t_{\alpha, n-p-1}$ распределения Стьюдента.

Если $|t_{факт}| > t_{крит}$, то коэффициент считается статистически значимым, т. е. гипотеза $H_0 : \beta_j = 0$ отклоняется.

В противном случае ($|t_{факт}| < t_{крит}$) коэффициент считается статистически незначимым (статистически близким к нулю). Это означает, что фактор x_j линейно не связан с зависимой переменной y . Поэтому после установления того факта, что коэффициент b_j статистически незначим, рекомендуется исключить из уравнения регрессии переменную x_j .

Проверим значимость оценки параметра регрессии β_j с помощью критерия Стьюдента.

Выдвигаем две гипотезы:

$$H_0 : \beta_j = 0;$$

$$H_1 : \beta_j \neq 0.$$

Значения стандартных ошибок параметров b_2, b_3 с учетом округления равны (рисунок 2.5):

$$m_{b_2} = 0,06, \quad m_{b_3} = 2,25.$$

Они показывают, какое значение данной характеристики сформировались под влиянием случайных факторов. Эти значения используются для расчета t-критерия Стьюдента (рисунок 5):

$$t_{b_2} = 2,66, \quad t_{b_3} = 2,96.$$

Так как наблюдаемые значения t-критерия превышают его критическое значение на 5 % уровне значимости с 20 степенями свободы $t_{крит(0,05;18)} = 2,1$ (приложение Б), можно сделать вывод о существенности параметра, который формируется под воздействием неслучайных причин. Здесь коэффициенты регрессии являются статистически значимыми, т.е. выполняется альтернативная гипотеза.

На это же указывает показатель вероятности случайных значений параметров регрессии: если α меньше принятого нами уровня (обычно 0,1; 0,05 или 0,01), делают вывод о неслучайной природе данного значения параметра, т.е. о том, что он статистически значим и надежен. В противном случае принимается гипотеза о случайной природе значения коэффициентов уравнения.

6 Оцените качество уравнения через среднюю ошибку аппроксимации

Необходимые данные для расчета средней ошибки аппроксимации представлены в таблице 2.1.

$$\bar{A} = \frac{1}{n} \sum \left| \frac{y - \tilde{y}_i}{y} \right| \cdot 100\% = \frac{1}{20} \cdot 2,26 \cdot 100 = 11,3\%.$$

Таблица 2.1 – Данные для расчета средней ошибки аппроксимации

Год	Квартал	y	\tilde{y}_i	$\left \frac{y - \tilde{y}_i}{y} \right $
1	1	0,4	0,36	0,11
	2	0,5	0,44	0,12
	3	0,6	0,54	0,10
	4	0,5	0,52	0,05
2	1	0,3	0,37	0,24
	2	0,4	0,36	0,11
	3	0,5	0,52	0,05
	4	0,4	0,42	0,06
3	1	0,5	0,58	0,15
	2	0,6	0,69	0,15
	3	0,7	0,69	0,01
	4	0,5	0,70	0,41
4	1	0,7	0,67	0,05
	2	0,8	0,70	0,13
	3	0,9	0,75	0,17
	4	0,7	0,71	0,01
5	1	0,8	0,78	0,03
	2	0,9	0,81	0,10
	3	0,7	0,82	0,17
	4	0,8	0,77	0,04
Итого		12,2	12,2	2,26

Таким образом, фактические значения эндогенной переменной отличаются от теоретических значений на 11,3 %. Следовательно, построенная модель имеет хорошую точность.

7 Постройте модель в стандартизованном масштабе и интерпретируйте ее параметры

К определению параметров множественной регрессии возможен иной подход, когда на основе матрицы парных коэффициентов корреляции строится уравнение регрессии в стандартизованном масштабе:

$$t_y = \beta_1 \cdot t_{x_1} + \beta_2 \cdot t_{x_2} + \dots + \beta_p \cdot t_{x_p} + \varepsilon,$$

где $t_y, t_{x_1}, \dots, t_{x_p}$ – стандартизованные переменные: $t_y = \frac{y - \bar{y}}{\sigma_y}$, $t_{x_i} = \frac{x_i - \bar{x}_i}{\sigma_{x_i}}$, для

которых среднее значение равно нулю: $\bar{t}_y = \bar{t}_{x_i} = 0$, а среднее квадратическое отклонение равно единице: $\sigma_{t_y} = \sigma_{t_{x_i}} = 1$;

β – стандартизованные коэффициенты регрессии.

Применив МНК к уравнению множественной регрессии в стандартизованном масштабе, после соответствующих преобразований получим систему нормальных уравнений вида

$$\begin{cases} r_{yx_1} = \beta_1 + \beta_2 \cdot r_{x_2x_1} + \beta_3 \cdot r_{x_3x_1} + \dots + \beta_p \cdot r_{x_px_1}, \\ r_{yx_2} = \beta_1 \cdot r_{x_2x_1} + \beta_2 + \beta_3 \cdot r_{x_3x_2} + \dots + \beta_p \cdot r_{x_px_2}, \\ \dots \\ r_{yx_p} = \beta_1 \cdot r_{x_px_1} + \beta_2 \cdot r_{x_px_2} + \beta_3 \cdot r_{x_px_3} + \dots + \beta_p. \end{cases}$$

Решая ее методом определителей, найдем параметры – стандартизованные коэффициенты регрессии (β -коэффициенты).

Стандартизованные коэффициенты регрессии показывают, на сколько сигм изменится в среднем результат, если соответствующий фактор x_i изменится на одну сигму при неизменном среднем уровне других факторов. В силу того, что все переменные заданы как центрированные и нормированные, стандартизованные коэффициенты регрессии β_i сравнимы между собой.

Сравнивая, их друг с другом, можно ранжировать факторы по силе их воздействия на результат. В этом основное достоинство стандартизованных коэффициентов регрессии в отличие от коэффициентов «чистой» регрессии, которые несравнимы между собой.

В парной зависимости стандартизованный коэффициент регрессии есть не что иное, как линейный коэффициент корреляции r_{xy} . Подобно тому, как в парной зависимости коэффициенты регрессии и корреляции связаны между собой, так и во множественной регрессии коэффициенты «чистой» регрессии b_i связаны со стандартизованными коэффициентами регрессии β_i , а именно

$$b_i = \beta_i \frac{\sigma_y}{\sigma_{x_i}}.$$

Это позволяет от уравнения регрессии в стандартизованном масштабе

$$\tilde{t}_y = \beta_1 \cdot t_{x_1} + \beta_2 \cdot t_{x_2} + \dots + \beta_p \cdot t_{x_p}$$

переходить к уравнению регрессии в натуральном масштабе переменных

$$\tilde{y}_i = b_0 + b_1 \cdot x_1 + b_2 \cdot x_2 + \dots + b_p \cdot x_p.$$

Параметр b_0 определяется как

$$b_0 = \bar{y} - b_1 \cdot \bar{x}_1 - b_2 \cdot \bar{x}_2 - \dots - b_p \cdot \bar{x}_p.$$

Содержание стандартизованных коэффициентов регрессии позволяет использовать их при отсеке факторов – из модели исключаются факторы с наименьшим значением β_i .

В нашем примере для информативных экзогенных переменных уравнение в стандартизованном масштабе примет вид:

$$\tilde{t}_y = \beta_2 t_{x_2} + \beta_3 t_{x_3}.$$

Расчет β – коэффициентов выполним по формулам:

$$\beta_2 = \frac{r_{yx_2} - r_{yx_3} r_{x_2x_3}}{1 - r_{x_2x_3}^2}, \quad \beta_3 = \frac{r_{yx_3} - r_{yx_2} r_{x_2x_3}}{1 - r_{x_2x_3}^2}.$$

Парные коэффициенты корреляции берутся из матрицы (рисунок 4):

$$\beta_2 = \frac{r_{yx_2} - r_{yx_3} r_{x_2x_3}}{1 - r_{x_2x_3}^2} = \frac{0,81 - 0,83 \cdot 0,73}{1 - 0,73^2} = 0,45,$$

$$\beta_3 = \frac{r_{yx_3} - r_{yx_2} r_{x_2x_3}}{1 - r_{x_2x_3}^2} = \frac{0,83 - 0,81 \cdot 0,73}{1 - 0,73^2} = 0,5.$$

Получим уравнение $\tilde{t}_y = 0,45 t_{x_2} + 0,5 t_{x_3}$.

Стандартизованные коэффициенты регрессии показывают, на сколько сигм изменится в среднем результативный признак, если соответствующий фактор изменится на 1 сигму при неизменном среднем уровне других факторов.

В нашем случае, при увеличении оборотного капитала на 1 сигму показатель экономической безопасности увеличится на 0,45 сигм, при условии, что основной капитал останется на прежнем уровне.

Аналогично вывод для β_3 .

8 Рассчитайте прогнозное значение результата, если прогнозное значение факторов составляют 80 % от их максимальных значений

Пусть требуется оценить прогнозное значение эндогенной переменной для заданных значений экзогенных переменных. Тогда точечный прогноз будет рассчитываться по линейному уравнению регрессии следующим образом:

$$\tilde{y}_{x_{\text{прог}}} = b_0 + b_1 \cdot x_1^{\text{прогн}} + b_2 \cdot x_2^{\text{прогн}} + \dots + b_p \cdot x_p^{\text{прогн}}.$$

Рассчитаем ожидаемое прогнозное значение чистого дохода как точечный прогноз путем подстановки в уравнение регрессии прогнозные значения экзогенных переменных.

Найдем максимальные значения для экзогенных переменных x_2 и x_3 с помощью инструмента анализа данных Описательная статистика (рисунок 2.6).

Результаты вычисления соответствующих показателей для каждого признака представлены на рисунке 2.7.

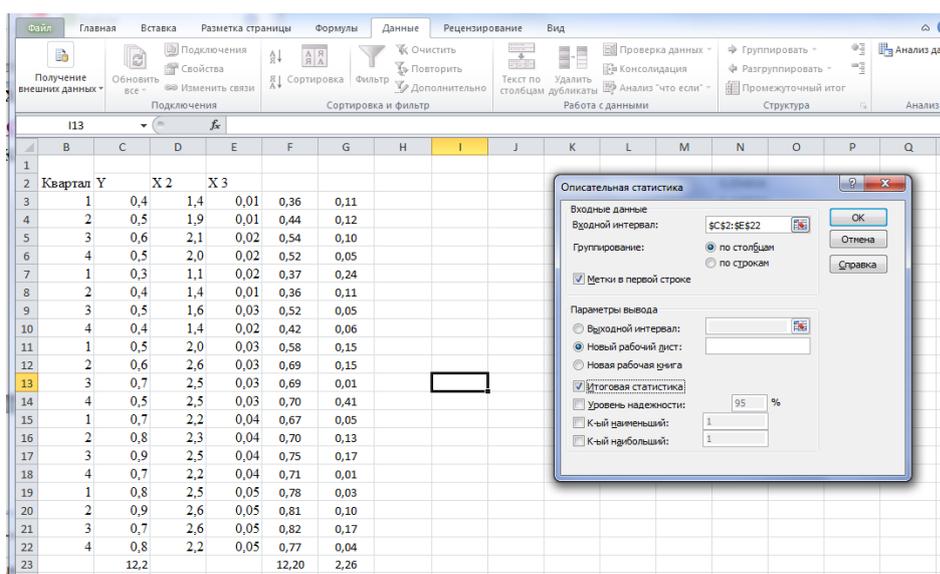


Рисунок 2.6 – Применения инструмента «Описательная статистика»

	Y	X 2	X 3	
3	Среднее	0,61	Среднее	2,08
4	Стандартная ошибка	0,03967	Стандартная ошибка	0,105531
5	Медиана	0,6	Медиана	2,2
6	Мода	0,5	Мода	2,5
7	Стандартное отклонение	0,177408	Стандартное отклонение	0,47195
8	Дисперсия выборки	0,031474	Дисперсия выборки	0,222737
9	Эксцесс	-1,05104	Эксцесс	-0,67827
10	Асимметричность	0,084205	Асимметричность	-0,70535
11	Интервал	0,6	Интервал	1,5
12	Минимум	0,3	Минимум	1,1
13	Максимум	0,9	Максимум	2,6
14	Сумма	12,2	Сумма	41,6
15	Счет	20	Счет	20

Рисунок 2.7 – Результат применения инструмента «Описательная статистика»

Следовательно, максимальные значения для экзогенных переменных составят: $x_2^{\max} = 2,6$ и $x_3^{\max} = 0,05$ (рисунок 2.7).

Найдем прогнозные значения экзогенных переменных:

для экзогенной переменной x_2 : $x_2^{\text{прогн}} = 0,8 \cdot 2,6 = 2,08$;

для экзогенной переменной x_3 : $x_3^{\text{прогн}} = 0,8 \cdot 0,05 = 0,04$;

Подставим прогнозные значения экзогенных переменных в уравнение $\tilde{y}_i = 0,055 + 0,169x_2 + 0,645x_3$. В результате получим:

$$\tilde{y}_i = 0,055 + 0,169 \cdot 2,08 + 0,645 \cdot 0,04 = 0,432$$

Таким образом, при прогнозных значениях оборотного капитала 2,08 млн. р. и основного капитала 0,04 млн. р. показатель экономической безопасности составит 0,432 млрд. долл.

9 Рассчитайте ошибки и доверительный интервал прогноза для уровня значимости $\alpha = 0,05$

Доверительный интервал прогноза имеет следующий вид:

$$\tilde{y}_{\text{прог}} - t_{\alpha} m_{\tilde{y}_i} \leq y^* \leq \tilde{y}_{\text{прог}} + t_{\alpha} m_{\tilde{y}_i},$$

где $m_{\tilde{y}_i}$ – средняя ошибка прогнозируемого значения $m_{\tilde{y}_i} = S \sqrt{X_0^T (X^T X)^{-1} X_0}$;

X_0 – вектор-столбец прогнозных значений факторов;

S – стандартная ошибка $S = \sqrt{\frac{(1-R^2)\sigma_y^2 n}{n-p-1}}$.

Поэтапно определим доверительный интервал для прогнозного значения эндогенной переменной:

1) составим вектор-столбец $X_0 = \begin{pmatrix} 1 \\ 2,08 \\ 0,04 \end{pmatrix}$;

2) найдем транспонируемый вектор-столбец $X_0^T = (1 \ 2,08 \ 0,04)$;

3) из описательной статистики (рисунок 2.7) находим дисперсию эндогенной переменной: $\sigma_y^2 = 0,03$, а из результатов регрессионного анализа (рисунок 5) определяем коэффициент множественной детерминации $R^2 = 0,78$;

4) найдем стандартную ошибку $S = \sqrt{\frac{(1-0,78^2) \cdot 0,03 \cdot 20}{20-2-1}} = 0,014$;

5) составим матрицу X , размер которой 20×3 , состоящую из 20 наблюдаемых значений независимых переменных x_2 и x_3 , а также единичного столбца:

$$X = \begin{pmatrix} 1 & 1,4 & 0,01 \\ 1 & 1,9 & 0,01 \\ 1 & 2,1 & 0,02 \\ 1 & 2,0 & 0,02 \\ 1 & 1,1 & 0,02 \\ 1 & 1,4 & 0,01 \\ 1 & 1,6 & 0,03 \\ 1 & 1,4 & 0,02 \\ 1 & 2,0 & 0,03 \\ 1 & 2,6 & 0,03 \\ 1 & 2,5 & 0,03 \\ 1 & 2,5 & 0,03 \\ 1 & 2,2 & 0,04 \\ 1 & 2,3 & 0,04 \\ 1 & 2,5 & 0,04 \\ 1 & 2,2 & 0,04 \\ 1 & 2,5 & 0,05 \\ 1 & 2,6 & 0,05 \\ 1 & 2,6 & 0,05 \\ 1 & 2,2 & 0,05 \end{pmatrix}$$

б) найдем произведение $(X^T \cdot X)$

$$X = \begin{pmatrix} 20 & 41,6 & 0,616 \\ 41,6 & 90,76 & 1,368 \\ 0,616 & 1,368 & 0,022 \end{pmatrix}$$

7) найдем $(X^T \cdot X)^{-1}$

$$(X^T \cdot X)^{-1} = \begin{pmatrix} 1,167 & -0,653 & 7,819 \\ -0,653 & 0,510 & -13,273 \\ 7,819 & -13,273 & 642,801 \end{pmatrix}$$

- 8) найдем выражение $X_0^T (X^T \cdot X)^{-1} \cdot X_0 = 0,105$;
- 9) вычислим среднюю ошибку прогнозируемого значения

$$m_{\bar{y}_i} = 0,014 \cdot \sqrt{0,105} = 0,0045;$$

10) по таблицам распределения Стьюдента находим табличное значение t_α при уровне значимости 0,05 и числе степеней свободы 22:

$$t_{\text{крит}(0,05;18)} = 2,1 \text{ (приложение Б);}$$

11) составляем доверительный интервал:

$$0,432 - 2,1 \cdot 0,0045 \leq y^* \leq 0,432 + 2,1 \cdot 0,0045$$

$$0,423 \leq y^* \leq 0,442$$

Значит, с вероятностью 95 % можно сказать, что показатель экономической безопасности будет варьировать от 0,423 до 0,442 при оборотном капитале в 2,08 млн. р. и основном капитале 0,04 млн. р.

10 По полученным результатам сделайте экономический вывод

Делается общий вывод по проделанной лабораторной работе.

Контрольные вопросы к лабораторной работе № 2

- 1 Дайте определение классической модели линейной регрессии.
- 2 Перечислите допущения лежащие в основе метода наименьших квадратов.
- 3 Какие свойства оценок МНК?
- 4 Уравнение регрессии в стандартизованном масштабе.

5 Проверка статистической значимости уравнения множественной регрессии и коэффициентов регрессии.

6 Частная регрессия и корреляция.

7 Определение доверительного интервала для уравнения множественной регрессии.

Тесты для самоконтроля к лабораторной работе № 2

1 Какой показатель используется для определения части вариации, обусловленной изменением величины изучаемого фактора?

- А) коэффициент вариации;
- Б) коэффициент корреляции;
- В) коэффициент детерминации;
- Г) коэффициент эластичности.

2 По 20 наблюдениям построено уравнение регрессии: $\hat{y} = b_0 + b_1x_1 + b_2x_2$. Для проверки значимости уравнения вычислено значение статистики: 4.2. Выводы:

- А) уравнение значимо при $\alpha=0,05$;
- Б) уравнение незначимо при $\alpha=0,05$;
- В) уравнение незначимо при $\alpha=0,01$.
- Г) уравнение значимо при $\alpha=0,01$.

3 В множественном линейном уравнении регрессии строятся доверительные интервалы для коэффициентов регрессии с помощью распределения:

- А) Нормального;
- Б) Стьюдента;
- В) Пирсона;
- Г) Фишера-Снедекора.

4 Множественный коэффициент корреляции равен 0.9. Какой процент дисперсии результативного признака объясняется влиянием всех факторных признаков?

- А) 90 %;
- Б) 81 %;
- В) 95 %;
- Г) 45 %.

5 В каких пределах меняется частный коэффициент корреляции вычисленный по рекуррентным формулам?

- А) от $-\infty$ до $+\infty$;
- Б) от 0 до 1;
- В) от 0 до $+\infty$;
- Г) от -1 до $+1$.

6 При добавлении в уравнение регрессии еще одного объясняющего фактора множественный коэффициент корреляции:

- А) уменьшится;
- Б) возрастет;
- В) сохранит свое значение;
- Г) зависит от направления связи между «Y» и «X».

7 Оценка значимости параметров уравнения регрессии осуществляется на основе:

- А) t - критерия Стьюдента;
- Б) F - критерия Фишера – Снедекора;
- В) средней квадратической ошибки;
- Г) средней ошибки аппроксимации.

8 Приведенная формула $\beta_j = a_j \cdot \frac{\sigma_{xj}}{\sigma_y}$ необходима для расчета:

- А) параметра уравнения;
- Б) коэффициента корреляции;
- В) стандартизованным коэффициентом регрессии;
- Г) коэффициента эластичности.

9 В чем принципиальное отличие скорректированного коэффициента детерминации от обычного коэффициента детерминации

- А) учитывает число переменных в уравнении регрессии;
- Б) позволяет оценить значимость модели;
- В) позволяет оценить качество и точность модели;
- Г) учитывает дисперсию остатков;

10 Допустим в результате исследования было получено 3 статистически значимых регрессионных уравнения, различающихся количеством и составом независимых переменных, на основе какого показателя производится отбор наилучшей?

- А) множественного коэффициента корреляции;
- Б) коэффициента взаимной сопряженности К. Пирсона;
- В) множественного коэффициента детерминации;
- Г) скорректированного значения множественного коэффициента детерминации.

3 Лабораторная работа № 3. Регрессионные модели с переменной структурой

Цель выполнения лабораторной работы: научиться использовать в качестве экзогенных переменных качественные признаки и применять тест Г. Чоу для обнаружения наличия структурного сдвига.

Задания:

По данным лабораторной работы № 2:

- 1) оцените линейную регрессию, включив в модель информативные факторы и фиктивную переменную;
- 2) проверти данные на наличие структурного сдвига при помощи теста Чоу.

Реализация типовых заданий

1 Оцените линейную регрессию, включив в модель информативные факторы и фиктивную переменную

При изучении взаимосвязей между социально-экономическими показателями зачастую возникает необходимость наряду с количественными признаками рассматривать также качественные признаки, имеющих две или более градаций: пол, образование, сезон, национальность и т.п. Для отражения влияния качественных признаков на эндогенную переменную применяются так называемые фиктивные переменные.

Фиктивные (искусственные) переменные (dummy variables) – это переменные с дискретным множеством значений, которые количественным образом описывают качественные признаки. В эконометрических моделях обычно используют фиктивные переменные бинарного типа (0 – 1):

$$D = \begin{cases} 1 - \text{наличие признака} \\ 0 - \text{отсутствие признака} \end{cases}.$$

Если качественный признак имеет более двух значений, вводится несколько бинарных переменных по правилу: при k альтернативных значениях у качественной переменной, используется $k-1$ фиктивная переменная.

В регрессионных моделях используются фиктивные переменные двух типов: переменные сдвига и переменные наклона.

Если изменение значения фиктивной переменной в модели регрессии приводит к изменению значения эндогенной переменной на некую среднюю величину, не зависящую от значений экзогенных переменных, такую переменную называют фиктивной переменной сдвига.

Спецификация множественной регрессионной модели с фиктивными переменными сдвига имеет вид:

$$y = \beta_0 + \beta_1 \cdot x_1 + \beta_2 \cdot x_2 + \dots + \beta_p \cdot x_p + \delta_{11} \cdot D_{11} + \delta_{12} \cdot D_{12} + \dots + \delta_{21} \cdot D_{21} + \delta_{22} \cdot D_{22} + \dots + \delta_{j1} \cdot D_{j1} + \delta_{j2} \cdot D_{j2} + \dots + u,$$

где β_0 – свободный член,

$\beta_1, \beta_2, \dots, \beta_p$ – коэффициенты при экзогенных переменных x_1, x_2, \dots, x_p ;

$\delta_{11}, \delta_{12}, \dots$ – коэффициенты при первой фиктивной переменной D_{11}, D_{12}, \dots ;

$\delta_{21}, \delta_{22}, \dots$ – коэффициенты при второй фиктивной переменной D_{21}, D_{22}, \dots ;

$\delta_{j1}, \delta_{j2}, \dots$ – коэффициенты при j -ой фиктивной переменной D_{j1}, D_{j2}, \dots ;

u – случайная ошибка (отклонение).

Коэффициент при фиктивной переменной сдвига (δ) показывает среднее изменение эндогенной переменной при переходе из одной категории в другую при неизменных значениях остальных переменных.

Значение фиктивной переменной $D=0$ называется базовым, или сравнительным, оно определяется целями исследования или принимается произвольно. При замене базового значения переменной суть модели не меняется, а меняется знак параметра δ на противоположный.

Если включение фиктивной переменной в модель регрессии позволяет изменить угол наклона линии регрессии, отражающей зависимость эндогенной переменной от экзогенных переменных, то такую переменную называют фиктивной переменной наклона.

Спецификация множественной регрессионной модели с фиктивной переменной наклона имеет вид:

$$\begin{aligned}
 y = & \beta_0 + \beta_1 \cdot x_1 + \beta_2 \cdot x_2 + \dots + \beta_p \cdot x_p + \\
 & + \delta_{111} \cdot x_1 \cdot D_{11} + \delta_{112} \cdot x_1 \cdot D_{12} + \dots + \delta_{1j1} \cdot x_1 \cdot D_{j1} + \delta_{1j2} \cdot x_1 \cdot D_{j2} + \dots + \\
 & + \delta_{211} \cdot x_2 \cdot D_{11} + \delta_{212} \cdot x_2 \cdot D_{12} + \dots + \delta_{2j1} \cdot x_2 \cdot D_{j1} + \delta_{2j2} \cdot x_2 \cdot D_{j2} + \dots + \\
 & + \dots + \\
 & + \delta_{p11} \cdot x_p \cdot D_{11} + \delta_{p12} \cdot x_p \cdot D_{12} + \dots + \delta_{pj1} \cdot x_p \cdot D_{j1} + \delta_{pj2} \cdot x_p \cdot D_{j2} + \dots + u.
 \end{aligned}$$

Коэффициент при фиктивной переменной наклона δ показывает среднее изменение силы влияния экзогенной переменной при переходе от одного значения фиктивной переменной к другому.

По исходным данным из лабораторной работы № 2, построим уравнение регрессии с качественной переменной «пол руководителя компании». Для этого введем фиктивную переменную сдвига:

$$D = \begin{cases} 1 - \text{руководитель компании мужчина} \\ 0 - \text{в другом случае} \end{cases} .$$

Исходные данные по эндогенной и экзогенным переменным, а также значения фиктивной переменной представлены в таблице 3.1.

Таблица 3.1 – Исходные данные для построения модели с фиктивной переменной

Год	Квартал	Интегральный показатель экономической безопасности организации	Оборотный капитал (млн. рублей)	Основной капитал (млн. рублей)	Пол руководителя компании	Фиктивная переменная
		Y	X2	X3		D
1	1	0,4	1,4	0,01	муж	1
	2	0,5	1,9	0,01	муж	1
	3	0,6	2,1	0,02	муж	1
	4	0,5	2,0	0,02	жен	0
2	1	0,3	1,1	0,02	жен	0
	2	0,4	1,4	0,01	жен	0
	3	0,5	1,6	0,03	муж	1
	4	0,4	1,4	0,02	муж	1
3	1	0,5	2,0	0,03	жен	0
	2	0,6	2,6	0,03	муж	1
	3	0,7	2,5	0,03	муж	1
	4	0,5	2,5	0,03	муж	1
4	1	0,7	2,2	0,04	муж	1
	2	0,8	2,3	0,04	муж	1
	3	0,9	2,5	0,04	жен	0
	4	0,7	2,2	0,04	жен	0
5	1	0,8	2,5	0,05	жен	0
	2	0,9	2,6	0,05	жен	0
	3	0,7	2,6	0,05	муж	1
	4	0,8	2,2	0,05	жен	0

Предварительно проведем корреляционный анализ, построив матрицу парных коэффициентов корреляции между рассматриваемыми переменными, используя для этого инструмент анализа данных **Корреляция** (также как в задании 2 лабораторной работы № 2).

Результаты корреляционного анализа представлены на рисунке 2.1.

	Y	X 2	X 3	D
Y	1,000			
X 2	0,813	1,000		
X 3	0,827	0,733	1,000	
D	-0,180	0,048	-0,217	1,000

Рисунок 3.1 – Матрица парных коэффициентов корреляции по объединенной выборке

По матрице коэффициентов корреляции видно, что фиктивная переменная не коллинеарна с отобранными в лабораторной работе № 2 экзогенными переменными x_2 и x_3 (определено так же, как во втором задании лабораторной работы № 2). Следовательно, можно построить модель множественной регрессии, используя в качестве независимых переменных D, x_2 , и x_3 вида:

$$y = \beta_0 + \beta_2 \cdot x_2 + \beta_3 \cdot x_3 + \delta \cdot D + u .$$

Для оценки параметров используем инструмент анализа данных **Регрессия** (также как в первом задании лабораторной работы № 2). Результаты регрессионного анализа представлены на рисунке 3.2.

Регрессионная статистика		Дисперсионный анализ						
		df	SS	MS	F	начисность F		
Множественный	0,887	3	0,470	0,157	19,676	0,000013		
R-квадрат	0,787	16	0,128	0,008				
Нормированный	0,747	19	0,598					
Стандартная	0,089							
Наблюдения	20							
	Коэффициенты	Стандартная ошибка	t-статистика	P-Значение	Нижние 95%	Верхние 95%	Нижние 95,0%	Верхние 95,0%
Y-пересечение	0,062	0,097	0,640	0,531	-0,143	0,267	-0,143	0,267
X 2	0,187	0,067	2,784	0,013	0,045	0,329	0,045	0,329
X 3	5,854	2,437	2,402	0,029	0,688	11,020	0,688	11,020
D	-0,038	0,043	-0,876	0,394	-0,130	0,054	-0,130	0,054

Рисунок 3.2 – Вывод итогов регрессионного анализа

Уравнение примет вид: $\tilde{y} = 0,062 + 0,187 \cdot x_2 + 5,854 \cdot x_3 - 0,038 \cdot D$.

Уравнение регрессии статистически значимо по F – критерию Фишера на 5 % уровне значимости (определено так же, как в четвертом задании лабораторной работы № 2).

Коэффициент регрессии, стоящий при фиктивной переменной показывает, что при одном и том же объеме оборотного капитала и основного капитала, у предприятий руководителями, которых являются мужчины, интегральный показатель экономической безопасности организации ниже в среднем на 0,038, чем у остальных компаний.

Проверка статистической значимости параметров при фиктивных переменных осуществляется на основе t-критерия Стьюдента, также как для коэффициентов регрессии. Как видно по данным рисунка 3.2, коэффициент при D статистически незначим (определено как в пятом задании лабораторной работы № 2). Следовательно, наблюдаются несущественные различия интегрального показателя экономической безопасности организации между компаниями с руководителями разного пола.

Подставляя значение фиктивной переменной $D = 0$, получим уравнение зависимости интегрального показателя экономической безопасности организации от объема оборотного и основного капитала, при условии, что руководитель компании женщина, $D = 1$ - то же уравнение с руководителем компании мужчиной:

$$D = 0 \quad \tilde{y} = 0,062 + 0,187 \cdot x_2 + 5,854 \cdot x_3,$$

$$D = 1 \quad \tilde{y} = 0,024 + 0,187 \cdot x_2 + 5,854 \cdot x_3.$$

Как видим, при разных значениях фиктивной переменной получаем разные уравнения регрессии. Целесообразность применения двух уравнений регрессии вместо одного можно оценить, не прибегая к вводу фиктивных переменных. Для этого используется тест Чоу, к рассмотрению которого мы и переходим.

2 Проверти данные на наличие структурного сдвига при помощи теста Чоу

При эконометрических исследованиях нередко предполагается, что совокупность неоднородна с точки зрения числовых характеристик воздействия экзогенных переменных на эндогенные, при этом однородность может быть достигнута при разбиении совокупности по определенному критерию на две части. Проверку наличия (или отсутствия) в выборочных данных структурных изменений можно выполнить при помощи теста Чоу. Выдвигаются гипотезы:

H_0 : в выборочных данных отсутствует структурный сдвиг;

H_1 : в выборочных данных присутствует структурный сдвиг (целесообразно строить уравнение регрессии с соответствующей фиктивной переменной или для каждой качественно однородной совокупности строить свое уравнение регрессии).

Алгоритм теста Чоу включает следующие этапы.

1 Оцениваются параметры трех уравнений регрессии: первое уравнение строится для всей совокупности наблюдений, второе и третье – для соответствующих выделенных подмножеств совокупности наблюдений.

2 Для каждого из этих уравнений находят остаточную сумму квадратов:

$$S_{ост} = \sum (y - \tilde{y}_i)^2 .$$

Обозначим остаточную сумму квадратов для общего уравнения регрессии через S_0 , для уравнений регрессии по подмножествам наблюдений – через S_1 и S_2 .

3 Проверяется существенность различия дисперсий с помощью F -критерия. Наблюдаемое значение критерия находят по формуле:

$$F_{набл} = \frac{S_0 - S_1 - S_2}{S_1 + S_2} \cdot \frac{n - 2p - 2}{p + 1},$$

где n – число наблюдений по всей совокупности,

p – число экзогенных переменных в уравнении (должно быть одинаково во всех трех уравнениях).

Если $F_{набл} > F_{крит(\alpha; p+1; n-2p-2)}$, то гипотеза H_0 отклоняется.

Используя критерий Г. Чоу, проверим анализируемую совокупность компаний на наличие/отсутствие структурного сдвига.

По 11 наблюдениям с руководителями - мужчинами, построим уравнение регрессии от факторов x_2 и x_3 . Исходные данные представлены в таблице 3.2.

Таблица 3.2 – Данные по первому подмножеству (руководитель компании – мужчина)

Год	Квартал	Интегральный показатель экономической безопасности организации	Оборотный капитал (млн. рублей)	Основной капитал (млн. рублей)	Пол руководителя компании
		Y	X ₂	X ₃	D
1	1	0,4	1,4	0,01	муж
	2	0,5	1,9	0,01	муж
	3	0,6	2,1	0,02	муж
2	3	0,5	1,6	0,03	муж
	4	0,4	1,4	0,02	муж
3	2	0,6	2,6	0,03	муж
	3	0,7	2,5	0,03	муж
	4	0,5	2,5	0,03	муж
4	1	0,7	2,2	0,04	муж
	2	0,8	2,3	0,04	муж
5	3	0,7	2,6	0,05	муж

Для оценки параметров используем инструмент анализа данных Регрессия. Результаты регрессионного анализа представлены на рисунке 3.3.

Регрессионная статистика								
Множественный R	0,82							
R-квадрат	0,67							
Нормированный R-квадрат	0,59							
Стандартная ошибка	0,08							
Наблюдения	11							
Дисперсионный анализ								
	df	SS	MS	F	значимость F			
Регрессия	2	0,119	0,059	8,278	0,011			
Остаток	8	0,057	0,007					
Итого	10	0,176						
	Коэффициенты	Стандартная ошибка	t-статистика	P-Значение	Нижние 95%	Верхние 95%	Нижние 95,0%	Верхние 95,0%
Y-пересечение	0,19	0,13	1,48	0,18	-0,10	0,48	-0,10	0,48
X 2	0,12	0,08	1,53	0,16	-0,06	0,29	-0,06	0,29
X 3	5,23	2,84	1,84	0,10	-1,33	11,78	-1,33	11,78

Рисунок 3.3 – Вывод итогов регрессионного анализа по первому подмножеству

$$\text{Уравнение примет вид: } \tilde{y}_i = 0,19 + 0,12x_2 + 5,23x_3 .$$

Расчетные значения эндогенных переменных представлены в таблице 3.3.

Таблица 3.3 – Промежуточные расчеты для критерия Чоу по первому подмножеству

Год	Квартал	Y	X ₂	X ₃	\tilde{y}_i	$(y - \tilde{y}_i)^2$
1	1	0,4	1,4	0,01	0,405	0,00
	2	0,5	1,9	0,01	0,463	0,00
	3	0,6	2,1	0,02	0,539	0,00
2	3	0,5	1,6	0,03	0,533	0,00
	4	0,4	1,4	0,02	0,457	0,00
3	2	0,6	2,6	0,03	0,650	0,00
	3	0,7	2,5	0,03	0,638	0,00
	4	0,5	2,5	0,03	0,638	0,02
4	1	0,7	2,2	0,04	0,655	0,00
	2	0,8	2,3	0,04	0,667	0,02
5	3	0,7	2,6	0,05	0,755	0,00
Итого		6,4	23,1	0,31	6,400	0,06

Построим модель регрессии по 9 наблюдениям с руководителями женщинами (исходные данные представлены в таблице 3.4).

Таблица 3.4 – Данные по второму подмножеству (руководитель компании – женщина)

Год	Квартал	Интегральный показатель экономической безопасности организации	оборотный капитал (млн. рублей)	основной капитал (млн. рублей)	Пол руководителя компании
		Y	X ₂	X ₃	
1	4	0,5	2,0	0,02	жен
2	1	0,3	1,1	0,02	жен
3	4	0,5	2,0	0,02	жен
4	1	0,5	2,0	0,03	жен
5	3	0,9	2,5	0,04	жен
	4	0,7	2,2	0,04	жен
6	1	0,8	2,5	0,05	жен
	2	0,9	2,6	0,05	жен
	4	0,8	2,2	0,05	жен

Результаты регрессионного анализа представлены на рисунке 3.4.

Регрессионная статистика								
Множественный R	0,82							
R-квадрат	0,67							
Нормированный R-квадрат	0,59							
Стандартная ошибка	0,08							
Наблюдения	11							
Дисперсионный анализ								
	df	SS	MS	F	Значимость F			
Регрессия	2	0,12	0,06	8,28	0,011			
Остаток	8	0,06	0,01					
Итого	10	0,18						
	Коэффициенты	Стандартная ошибка	t-статистика	P-Значение	Нижние 95%	Верхние 95%	Нижние 95,0%	Верхние 95,0%
Y-пересечение	0,188	0,127	1,479	0,177	-0,105	0,481	-0,105	0,481
X 2	0,117	0,077	1,529	0,165	-0,060	0,295	-0,060	0,295
X 3	5,227	2,843	1,839	0,103	-1,328	11,781	-1,328	11,781

Рисунок 3.4 – Вывод итогов регрессионного анализа по второму подмножеству

Уравнение регрессии примет вид: $\tilde{y}_i = 0,188 + 0,117x_2 + 5,227x_3$.

Теоретические значения эндогенных переменных по уравнению представлены в таблицы 3.5.

Таблица 3.5 – Промежуточные расчеты для критерия Чоу по второму подмножеству

Год	Квартал	Y	X ₂	X ₃	\tilde{y}_i	$(y - \tilde{y}_i)^2$
1	4	0,5	2,0	0,02	0,55	0,30
2	1	0,3	1,1	0,02	0,30	0,09
	2	0,4	1,4	0,01	0,33	0,11
3	1	0,5	2,0	0,03	0,60	0,37
4	3	0,9	2,5	0,04	0,80	0,64
	4	0,7	2,2	0,04	0,71	0,51
5	1	0,8	2,5	0,05	0,85	0,73
	2	0,9	2,6	0,05	0,88	0,77
	4	0,8	2,2	0,05	0,77	0,59
Итого		5,8	18,5	0,31	5,80	4,11

По всем 20 наблюдениям (таблица 3.6) уравнение регрессии было оценено в лабораторной работе № 2.

Таблица 3.6 – Данные по всей совокупности

Год	Квартал	Y	X ₂	X ₃	\tilde{y}_i	$(y - \tilde{y}_i)^2$
1	1	0,4	1,4	0,01	0,357	0,128
	2	0,5	1,9	0,01	0,442	0,195
	3	0,6	2,1	0,02	0,542	0,293
	4	0,5	2,0	0,02	0,525	0,275
2	1	0,3	1,1	0,02	0,373	0,139
	2	0,4	1,4	0,01	0,357	0,128
	3	0,5	1,6	0,03	0,524	0,274
	4	0,4	1,4	0,02	0,424	0,180
3	1	0,5	2,0	0,03	0,577	0,333
	2	0,6	2,6	0,03	0,692	0,479

Продолжение таблицы 3.6

Год	Квартал	Y	X2	X3	\tilde{y}_i	$(y - \tilde{y}_i)^2$
	3	0,7	2,5	0,03	0,690	0,476
	4	0,5	2,5	0,03	0,704	0,496
4	1	0,7	2,2	0,04	0,668	0,446
	2	0,8	2,3	0,04	0,699	0,488
	3	0,9	2,5	0,04	0,747	0,558
	4	0,7	2,2	0,04	0,710	0,505
5	1	0,8	2,5	0,05	0,775	0,601
	2	0,9	2,6	0,05	0,806	0,650
	3	0,7	2,6	0,05	0,821	0,673
	4	0,8	2,2	0,05	0,767	0,589
Итого		12,2	41,6	0,62	12,200	7,906

Теоретические значения эндогенных переменных по данной модели представлены в графе 6 таблицы 3.6.

Рассчитываем F-критерий по формуле:

$$F = \frac{7,906 - 0,06 - 4,11}{0,06 + 4,11} \cdot \frac{20 - 2 \cdot 2 - 2}{2 + 1} = 4,18,$$

где $S_0 = 7,906$ – сумма квадратов отклонений фактических значений результативного признака от теоретических для объединенной выборки (таблица 3.6, итог графы 7);

$S_1 = 0,06$ – сумма квадратов отклонений фактических значений результативного признака от расчетных для первой подвыборки (таблица 3.3, итог графы 7);

$S_2 = 4,11$ – сумма квадратов отклонений фактических значений результативного признака от расчетных для второй подвыборки (таблица 3.5, итог графы 7).

Критическое значение критерия Фишера $F_{крит}(0,05;3;14)$ составило 3,34 (приложение Б). Так как расчетное значение критерия больше табличного, то гипотеза H_0 отклоняется, в выборочных данных присутствует структурный сдвиг, т.е. целесообразно строить уравнение регрессии с фиктивной переменной «пол» или отдельные уравнения регрессии для наблюдений, с руководителями разного пола.

Тест Чоу вполне достаточен для определения структурных изменений в выборочных данных. Однако более информативным является использование фиктивных переменных, так как их анализ позволяет не только установить наличие или отсутствие структурных изменений, но и оценить степень влияния как фиктивных, так и экзогенных переменных на эндогенную переменную, оценить качество регрессии в целом.

Контрольные вопросы к лабораторной работе № 3

- 1 Можно ли учесть в уравнении регрессии неколичественные факторы?
- 2 Дайте определение фиктивной переменной.
- 3 Существуют ли ограничения на количество вводимых переменных в модель регрессии?
- 4 При каких условиях строится уравнение множественной регрессии с фиктивными переменными.
- 5 Каков общий вид регрессионной модели с фиктивной переменной сдвига?
- 6 Каков общий вид регрессионной модели с фиктивной переменной наклона?
- 7 Как интерпретируются коэффициенты регрессии при фиктивной переменной сдвига и наклона.
- 8 Какова область применения теста Чоу?
- 9 Опишите методику теста Чоу.

Тесты для самоконтроля к лабораторной работе № 3

1 Если качественная переменная имеет k альтернативных значений, то при моделировании используются:

- А) $(k-1)$ фиктивная переменная;
- Б) k фиктивных переменных;
- В) $2*k$ фиктивных переменных;
- Г) $(k+1)$ фиктивная переменная.

2 Критерий Чоу основывается на применении:

- А) F - статистики;
- Б) t - статистики;
- В) χ^2 - статистики;
- Г) критерии Дарбина – Уотсона.

3 Фиктивные переменные могут принимать значения:

- А) 1 и 0;
- Б) 2;
- В) минус 1 и 1;
- Г) любые значения.

4 Тест Чоу основан на сравнении:

- А) дисперсий;
- Б) коэффициентов детерминации;
- В) математических ожиданий;
- Г) средних.

5 При замене базового значения фиктивной переменной меняется:

- А) знак параметра при фиктивной переменной;
- Б) значение параметра при фиктивной переменной;
- В) спецификация модели;

Г) значение коэффициента детерминации.

6 Какое влияние оказывает на линию регрессии фиктивная переменная сдвига:

- А) смещение линии регрессии относительно оси абсцисс;
- Б) параллельное смещение линии регрессии;
- В) не оказывает влияния;
- Г) изменение наклона линии регрессии.

7 Сезонные колебания учитываются включением в спецификацию регрессионной модели:

- А) одной фиктивной переменной сдвига;
- Б) одной фиктивной переменной наклона;
- В) нескольких фиктивных переменных наклона;
- Г) нескольких фиктивных переменных сдвига.

8 Если параметры, стоящие при фиктивных переменных оказываются статистически незначимыми, то можно сделать вывод:

- А) что фиктивная переменная не оказывает существенного влияния на результирующий признак;
- Б) что модель статистически незначима;
- В) что модель статистически значима;
- Г) что фиктивная переменная оказывает существенного влияния на результирующий признак.

9 Регрессионная модель с переменной структурой это:

- А) регрессия, построенная для каждой качественно отличной группы единиц совокупности;
- Б) регрессия по панельным данным;
- В) регрессия по временным рядам;

Г) общая регрессионная модель, построенная для совокупности в целом, учитывающей неоднородность данных (вводятся фиктивные переменные).

10 При наличии взаимодействия факторов x_j и фиктивных переменных z целесообразно построение модели:

А) $\hat{y} = a + bx + cz$;

Б) $\hat{y} = a + bx + cz + d(xz)$;

В) $\hat{y} = a + bx + c(xz)$;

Г) $\hat{y} = a + bx$.

4 Лабораторная работа № 4. Нарушения допущений классической модели линейной регрессии: гетероскедастичность

Цель выполнения лабораторной работы: научиться проверять регрессионные остатки на гомоскедастичность, смягчать последствия нарушения допущений классической модели линейной регрессии.

Задания:

По данным лабораторной работы № 2 выполните следующие задания:

1 Проверьте регрессионные остатки на гетероскедастичность с помощью 2 любых тестов:

- теста Голдфелда-Квандта,
- теста ранговой корреляции Спирмена,
- теста Уайта (White test).

2 Если будет обнаружена гетероскедастичность остатков, примените для исходных данных ОМНК, предполагая, что $\sigma^2(u_i) = \sigma^2 x_i^2$.

Реализация типовых заданий

1 Проверьте регрессионные остатки на гетероскедастичность

Использование МНК для нахождения параметров уравнения регрессии предполагает выполнение предпосылок. На практике проверка их выполнения осуществляется уже после того, как были найдены параметры уравнения регрессии. В данной лабораторной работе регрессионные остатки проверим на гетероскедастичность.

Постоянство дисперсии регрессионных остатков называют гомоскедастичностью остатков. Если эта дисперсия не постоянна, то такое явление называют гетероскедастичностью остатков.

Проверка выполнения требования гомоскедастичности остатков может быть проведена визуально на основе графика остатков или с помощью специальных критериев.

К тестам, позволяющим выявить наличие гетероскедастичности остатков, относятся тест Голдфелда-Квандта, ранговой корреляции Спирмена, Уайта, Парка, Глейзера и т.д.

Тест Голдфелда-Квандта применяется, если регрессионные остатки предполагаются нормально распределенными случайными величинами. Тогда выдвигаются гипотезы:

$H_0: \sigma_i = \sigma(u_i)$ – гомоскедастичность;

$H_1: \sigma_i \neq \sigma(u_i)$ – гетероскедастичность.

Порядок проведения теста следующий:

1 Все n наблюдений упорядочиваются по величине x_2 и x_3 , выбрав для этого на вкладке **Главная/Сортировка и фильтр/Автоматически расширить выделенный диапазон**, представив данные как показано на рисунке 4.1

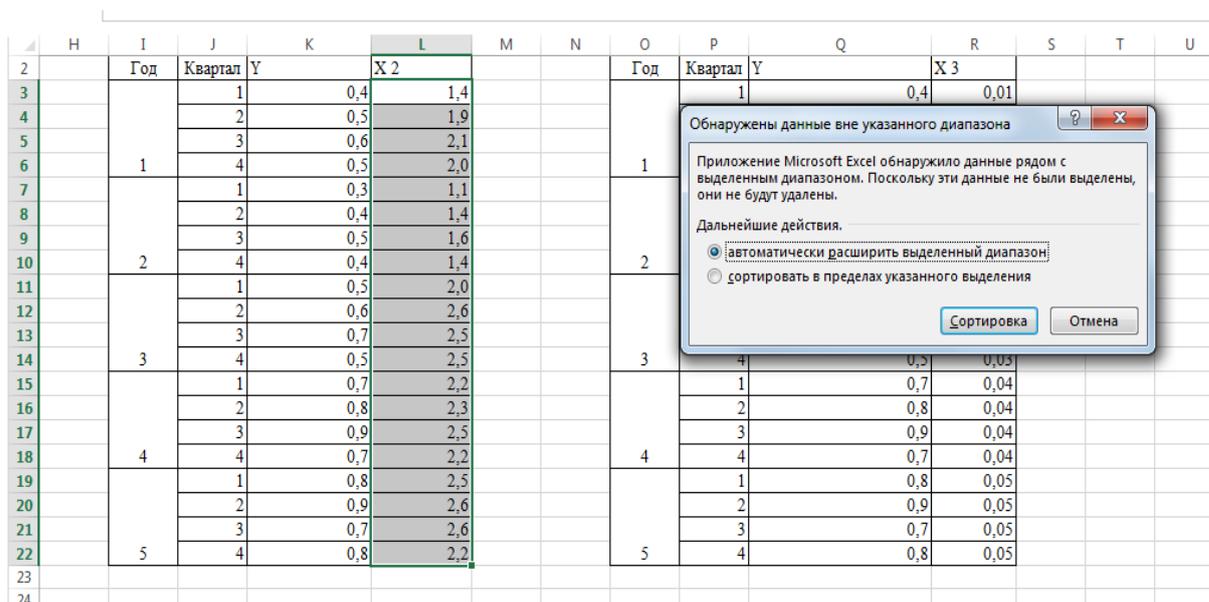


Рисунок 4.1 – Сортировка данных в табличном редакторе Excel

Аналогично упорядочиваются данные по переменной X_3 . Результаты представлены в таблице 4.1.

Таблица 4.1 – Упорядоченные значения по экзогенным переменным

Год	Квартал	Интегральный показатель экономической безопасности организации	Оборотный капитал (млн. рублей)	Год	Квартал	Интегральный показатель экономической безопасности организации	Основной капитал (млн. рублей)
		Y	X_2			Y	X_3
2	1	0,3	1,1	1	1	0,4	0,01
1	1	0,4	1,4	1	2	0,5	0,01
2	2	0,4	1,4	2	2	0,4	0,01
2	4	0,4	1,4	1	3	0,6	0,02
2	3	0,5	1,6	1	4	0,5	0,02
1	2	0,5	1,9	2	1	0,3	0,02
1	4	0,5	2,0	2	4	0,4	0,02
3	1	0,5	2,0	2	3	0,5	0,03
1	3	0,6	2,1	3	1	0,5	0,03
4	1	0,7	2,2	3	2	0,6	0,03
4	4	0,7	2,2	3	3	0,7	0,03
5	4	0,8	2,2	3	4	0,5	0,03
4	2	0,8	2,3	4	1	0,7	0,04
3	3	0,7	2,5	4	2	0,8	0,04
3	4	0,5	2,5	4	3	0,9	0,04
4	3	0,9	2,5	4	4	0,7	0,04
5	1	0,8	2,5	5	1	0,8	0,05
3	2	0,6	2,6	5	2	0,9	0,05
5	2	0,9	2,6	5	3	0,7	0,05
5	3	0,7	2,6	5	4	0,8	0,05

2 Упорядоченную совокупность делят на три группы, причем первая и третья должны быть равного объема с числом единиц большим, чем число параметров в уравнении регрессии. Пусть в первую и третью группы отобрано по k единиц.

Для рассматриваемого примера возьмем $k = 8$.

3 Оцениваются отдельные регрессии для первой подвыборки (8 первых наблюдений) и для третьей подвыборки (8 последних наблюдений). Если предположение о пропорциональности дисперсий

отклонений значениям X верно, то дисперсия регрессии по первой подвыборке (сумма квадратов отклонений $S_1 = \sum_{i=1}^k u_i^2$) будет существенно меньше дисперсии регрессии по третьей подвыборке (суммы квадратов отклонений $S_3 = \sum_{i=n-k+1}^n u_i^2$).

4 По каждой части находим уравнение регрессии (рисунок 4.2):

$$y'_1 = 0,11 + 0,21x_2 \quad S_1 = \sum u_i^2 = 0,006,$$

$$y'_3 = 1,24 - 0,20x_2 \quad S_3 = \sum u_i^2 = 0,136.$$

Год	Квартал	Y (рублей)	X2	Порядок подвыборки										
				Порядок подвыборки										
				вывод итогов										
				Y	X2									
2	1	0,3	1,1	Регрессионная статистика										
1	1	0,4	1,4	Множествен	0,92									
2	2	0,4	1,4	К-квадрат	0,83									
2	4	0,4	1,4	Нормирован	0,82									
2	3	0,5	1,6	Стандартна	0,05									
1	2	0,5	1,9	Наблюдени	8									
1	4	0,5	2,0	Дисперсионный анализ										
3	1	0,5	2,0		df	SS	MS	F	значимость F					
1	3	0,6	2,1	Регрессия	1	0,033	0,033	33,211	0,001					
4	1	0,7	2,2	Остаток	6	0,008	0,001							
4	4	0,7	2,2	Итого	7	0,039								
5	4	0,8	2,2											
				Третья подвыборка										
				Y	X2	Коэффициент								
4	2	0,8	2,3	Упорядочен	0,11	0,06	1,88	0,11	-0,05	0,25	-0,05	0,25		
3	3	0,7	2,5	X 2	0,21	0,04	5,76	0,00	0,12	0,29	0,12	0,29		
3	4	0,5	2,5	Третья подвыборка										
4	3	0,9	2,5	вывод итогов										
5	1	0,8	2,5											
3	2	0,6	2,6	Регрессионная статистика										
5	2	0,9	2,6	Множествен	0,141									
5	3	0,7	2,6	К-квадрат	0,020									
				Нормирован	-0,144									
				Стандартна	0,151									
				Наблюдени	8									
				Дисперсионный анализ										
					df	SS	MS	F	значимость F					
				Регрессия	1	0,005	0,005	0,121	0,739					
				Остаток	6	0,136	0,023							
				Итого	7	0,139								
				Коэффициент										
				Упорядочен	1,24	1,44	0,86	0,42	-2,29	4,77	-2,29	4,77		
				X 2	-0,20	0,37	-0,35	0,74	-1,60	1,20	-1,60	1,20		

Рисунок 4.2 – Вывод итогов для подвыборок для фактора x_2

5 Для сравнения соответствующих дисперсий строится следующая F-статистика:

$$\text{если } S_3 > S_1 \text{ то } F_{\text{набл}} = \frac{S_3}{S_1},$$

$$\text{если } S_1 > S_3 \text{ то } F_{\text{набл}} = \frac{S_1}{S_3},$$

(в числителе всегда большая сумма квадратов отклонений).

Если предполагается обратная зависимость дисперсии от значения экзогенной переменной, то числитель и знаменатель меняются местами. Для рассматриваемого примера наблюдаемое значение F – критерия составит:

$$F_{\text{набл}} = \frac{0,136}{0,006} = 22,94.$$

При сделанных предположениях относительно случайных отклонений построенная F -статистика имеет распределение Фишера с числами степеней свободы $v_1=v_2=k-2$, следовательно, критическое значение составит $F_{\text{крит}}(0,05;6;6) = 4,28$.

6 Если $F_{\text{набл}} > F_{\text{кр}}$, то гипотеза об отсутствии гетероскедастичности отклоняется.

По проведенным расчетам мы получили, что $F_{\text{набл}} > F_{\text{крит}}$ следовательно в ряду остатков обнаружена гетероскедастичность.

Аналогично проводится анализ для фактора x_3 .

Тест ранговой корреляции Спирмена основывается на предположении о зависимости величины дисперсии регрессионных остатков от значений экзогенных переменных. Выдвигаются гипотезы:

$$H_0: \rho_{x,u} = 0 \text{ (гомоскедастичность);}$$

$$H_1: \rho_{x,u} \neq 0 \text{ (гетероскедастичность),}$$

где $\rho_{x,u}$ – коэффициент корреляции рангов Спирмена в генеральной совокупности.

Для проведения теста значения x_i и абсолютные величины u_i ранжируются (упорядочиваются по величинам).

Предварительно найдем ряд остатков.

Для нахождения остатков u_i можно воспользоваться инструментом анализа данных **Регрессия**. Порядок действий следующий:

а) в главном меню выберите **Сервис / Анализ данных / Регрессия** или **Данные / Анализ данных / Регрессия**. Щелкните по кнопке **ОК**;

б) заполните диалоговое окно ввода данных и параметров ввода как показано на рисунке 4.3:

входной интервал Y – диапазон, содержащий данные эндогенной переменной;

входной интервал X – диапазон, содержащий данные экзогенных переменных;

метки – флажок, который указывает, содержит ли первая строка названия столбцов или нет;

выходной интервал – достаточно указать левую верхнюю ячейку будущего диапазона;

новый рабочий лист - можно задать произвольное имя нового листа;

остаток - флажок, указывает вывод остатков u_i и теоретические значения результирующего признака.

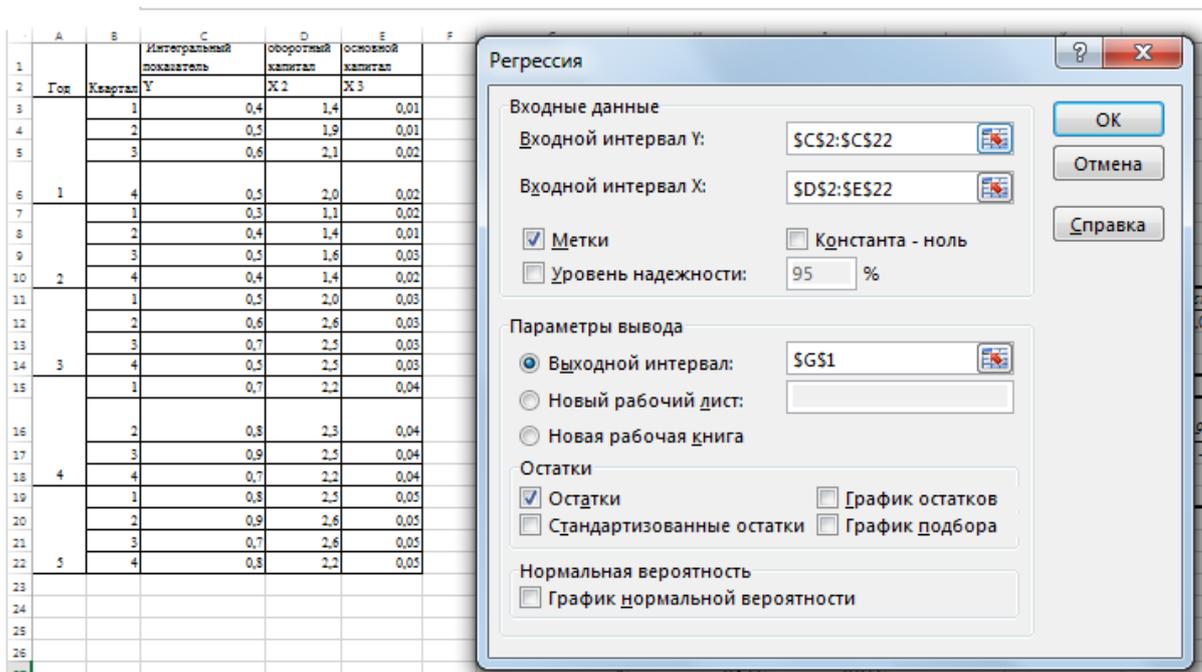


Рисунок 4.3 – Регрессия с остатками

Результаты регрессионного анализа, а также вспомогательные характеристики представлены на рисунке 4.4.

ВЫВОДИТОГОВ						
<i>Регрессионная статистика</i>						
Множественный R		0,89				
R-квадрат		0,80				
Нормированный R-квадрат		0,77				
Стандартная ошибка		0,08				
Наблюдения		20				
<i>Дисперсионный анализ</i>						
	<i>df</i>	<i>SS</i>	<i>MS</i>	<i>F</i>	<i>Значимость F</i>	
Регрессия	2	0,476	0,238	33,022	0,00000	
Остаток	17	0,122	0,007			
Итого	19	0,598				
	<i>Коэффициенты</i>	<i>Стандартная ошибка</i>	<i>t-статистика</i>	<i>P-Значение</i>	<i>Нижние 95%</i>	<i>Верхние 95%</i>
Y-пересечение	0,059	0,091	0,640	0,531	-0,134	0,252
X 2	0,164	0,059	2,765	0,013	0,039	0,289
X 3	6,790	2,038	3,331	0,004	2,490	11,090
ВЫВОД ОСТАТКА						
<i>Наблюдение</i>	<i>Предсказанное Y</i>	<i>Остатки</i>				
1	0,356	0,044				
2	0,438	0,062				
3	0,539	0,061				
4	0,522	-0,022				
5	0,375	-0,075				
6	0,356	0,044				
7	0,525	-0,025				
8	0,424	-0,024				
9	0,590	-0,090				

Рисунок 4.4 – Вывод остатков

Расставляем ранги для экзогенной переменной и модуля остатков, используя для этого встроенную функцию табличного редактора Excel – **Функции / Статистические/Ранг. Ср** (рисунок 4.5).

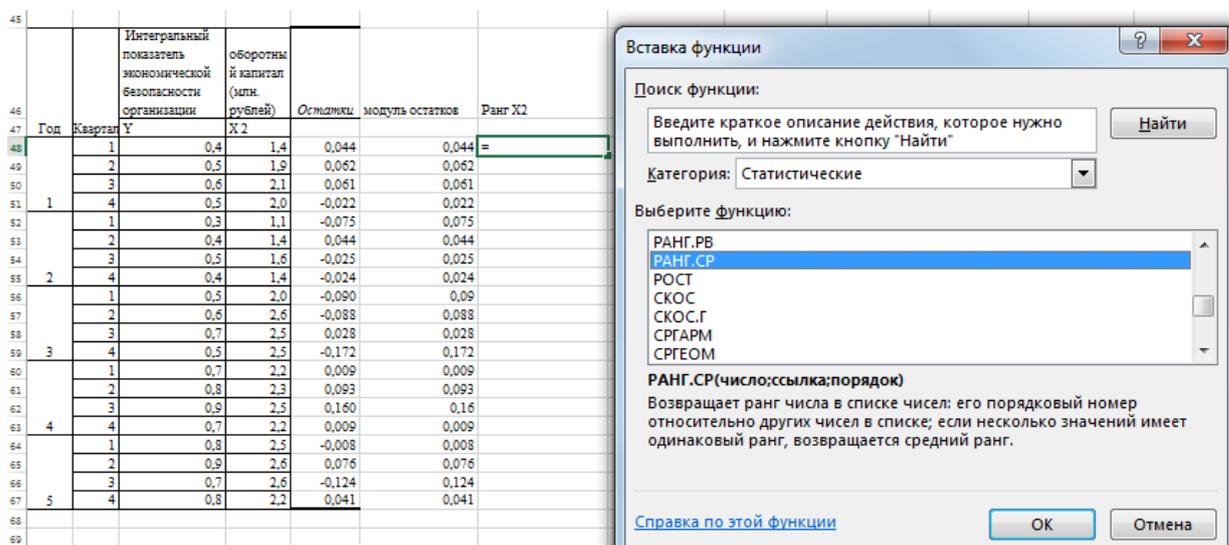


Рисунок 4.5 – Выбор функции «Ранг» в табличном редакторе Excel

В выбранной функции необходимо задать аргументы, как показано на рисунке 4.6.

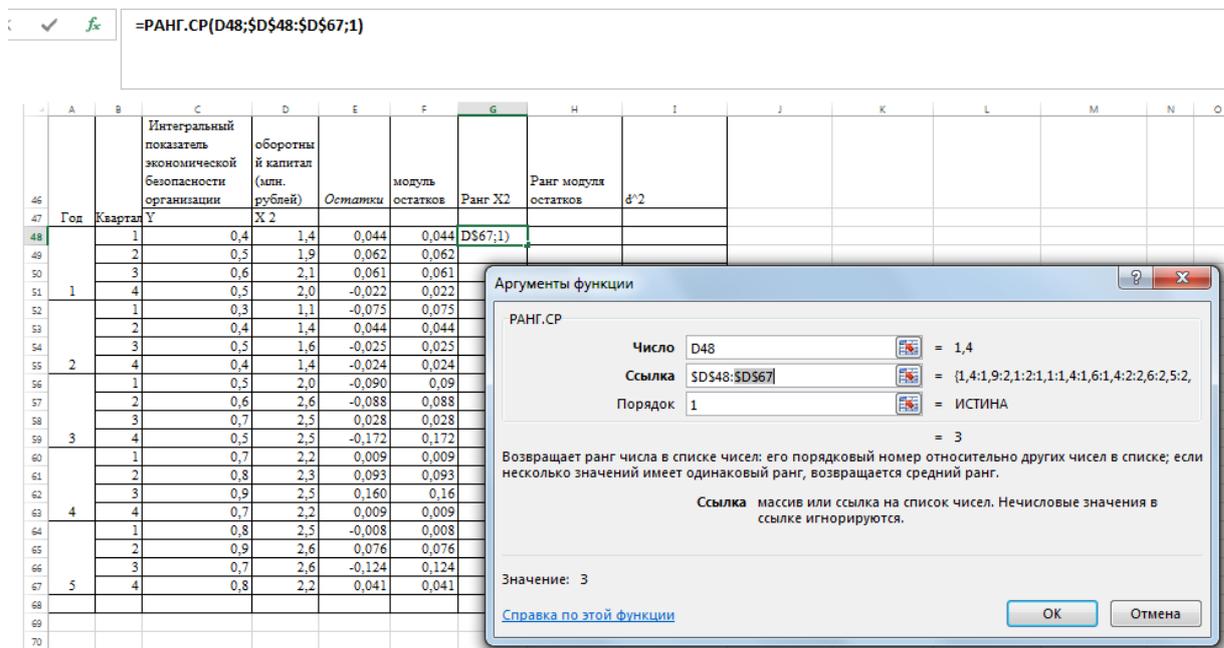


Рисунок 4.6 – Выбор аргументов функции «Ранг» в табличном редакторе Excel

Результаты ранжирования, а также вспомогательные расчеты для расчета коэффициента ранговой корреляции Спирмена представлены на рисунке 4.7.

45	А	В	С	U	Е	Р	С	Н	І
46	Год	Квартал	Интегральный показатель экономической безопасности организации	оборотный капитал (млн. рублей)	Остатки	модуль остатков	Ранг X2	Ранг модуля остатков	d^2
47	Год	Квартал	Y	X 2					
48	1	1	0,4	1,4	0,044	0,044	3	9,5	42,25
49		2	0,5	1,9	0,062	0,062	6	12	36
50		3	0,6	2,1	0,061	0,061	9	11	4
51		4	0,5	2,0	-0,022	0,022	7,5	4	12,25
52	2	1	0,3	1,1	-0,075	0,075	1	13	144
53		2	0,4	1,4	0,044	0,044	3	9,5	42,25
54		3	0,5	1,6	-0,025	0,025	5	6	1
55	3	4	0,4	1,4	-0,024	0,024	3	5	4
56		1	0,5	2,0	-0,090	0,09	7,5	16	72,25
57		2	0,6	2,6	-0,088	0,088	19	15	16
58		3	0,7	2,5	0,028	0,028	15,5	7	72,25
59	4	4	0,5	2,5	-0,172	0,172	15,5	20	20,25
60		1	0,7	2,2	0,009	0,009	11	2,5	72,25
61		2	0,8	2,3	0,093	0,093	13	17	16
62	5	3	0,9	2,5	0,160	0,16	15,5	19	12,25
63		4	0,7	2,2	0,009	0,009	11	2,5	72,25
64		1	0,8	2,5	-0,008	0,008	15,5	1	210,25
65		2	0,9	2,6	0,076	0,076	19	14	25
66	5	3	0,7	2,6	-0,124	0,124	19	18	1
67		4	0,8	2,2	0,041	0,041	11	8	9
68									884,5

Рисунок 4.7 – Расчетная таблица для проведения теста Спирмена

Затем определяется коэффициент ранговой корреляции Спирмена:

$$r_{x,u} = 1 - 6 \cdot \frac{\sum d_i^2}{n(n^2 - 1)},$$

где d_i – разность между рангами x_i и u_i , $i = 1, 2, \dots, n$;

n – число наблюдений.

$$\text{Тогда } r_{x,u} = 1 - 6 \cdot \frac{884,5}{20(20^2 - 1)} = 0,33.$$

Если коэффициент корреляции $\rho_{x,u}$ для генеральной совокупности равен нулю, то статистика

$$t = \frac{r_{x,u}}{\sqrt{1 - r_{x,u}^2}} \cdot \sqrt{n - 2}$$

имеет распределение Стьюдента с числом степеней свободы $v=n-2$. Следовательно, если наблюдаемое значение t -статистики превышает табличное, то необходимо отклонить гипотезу о равенстве нулю коэффициента корреляции $\rho_{x,u}$, а следовательно, и об отсутствии гетероскедастичности.

В нашем примере статистика Стьюдента равна: $t = \frac{0,33 \cdot \sqrt{20-2}}{\sqrt{1-0,33^2}} = 1,51$.

Табличное значение статистики Стьюдента составит $t_{(0,05; 18)}=2,1$.

Таким образом, мы получили, что расчетное значение меньше табличного, следовательно, гипотеза об отсутствии гетероскедастичности принимается на уровне значимости 5 %.

Аналогично проводится анализ для фактора x_3 .

В основе применения **теста Уайта** лежит предположение, что дисперсия регрессионных остатков представляет собой квадратичную функцию от одного или нескольких факторов. Выдвигаются гипотезы:

$H_0: \alpha_i = 0$ – параметры уравнения вспомогательной регрессии статистически незначимы (гомоскедастичность);

$H_1: \alpha_i \neq 0$ – параметры уравнения вспомогательной регрессии статистически значимы (гетероскедастичность).

Алгоритм теста включает следующие этапы:

1) допустим, исходная модель имеет вид:

$$y_i = \beta_0 + \beta_2 x_{2i} + \beta_3 x_{3i} + u_i .$$

В результате оценки данной модели получаем регрессионные остатки u_i (регрессионные остатки нами были получены ранее (см. тест ранговой корреляции Спирмена));

2) оцениваем вспомогательную регрессию вида:

$$u^2 = \alpha_0 + \alpha_1 x_2 + \alpha_2 x_3 + \alpha_3 x_2^2 + \alpha_4 x_3^2 + \alpha_5 x_2 x_3 + v,$$

где v – нормально распределенная ошибка.

Используем для этого инструмент **Регрессия в Excel** (рисунок 4.8).

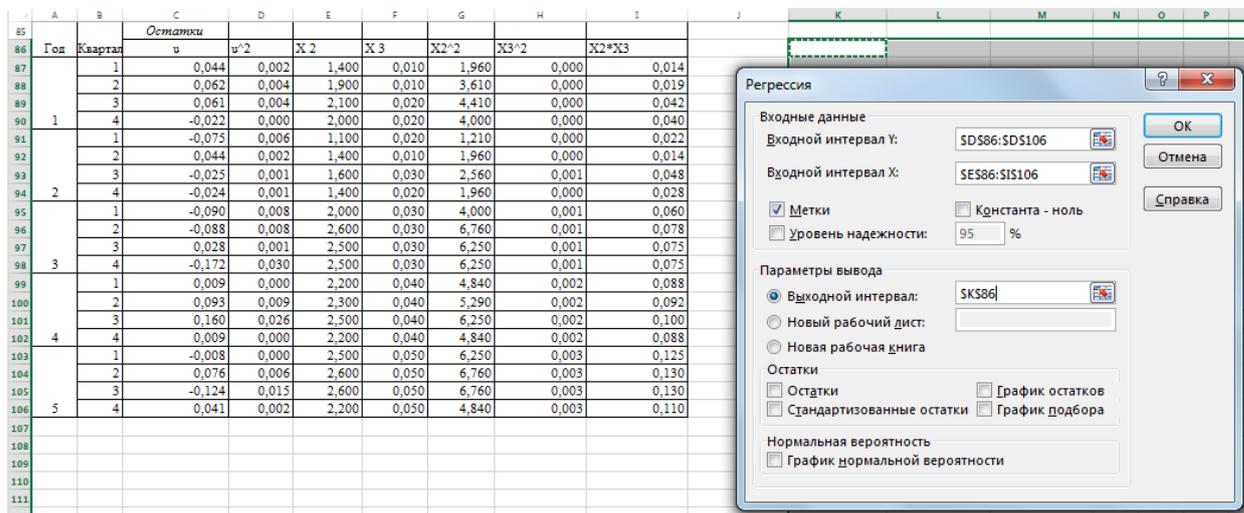


Рисунок 4.8 – Диалоговое окно ввода параметров инструмента Регрессия

Результаты оценки параметров вспомогательной регрессии представлены на рисунке 4.9.

Вывод итогов									
Регрессионная статистика									
Множественный	0,587								
R-квадрат	0,345								
Нормированный	0,111								
Стандартная ош	0,008								
Наблюдения	20								
Дисперсионный анализ									
	df	SS	MS	F	Значимость F				
Регрессия	5	0,0005	9,1383E-05	1,4730048	0,260076121				
Остаток	14	0,0009	6,2038E-05						
Итого	19	0,0013							
Коэффициенты стандартная ош-статистика		P-Значение	Нижние 95%	Верхние 95%	Нижние 95,0%	Верхние 95,0%			
Y-пересечение	0,033	0,035	0,937	0,364	-0,043	0,109	-0,043	0,109	
X 2	-0,036	0,044	-0,827	0,422	-0,130	0,058	-0,130	0,058	
X 3	-0,190	1,237	-0,154	0,880	-2,842	2,463	-2,842	2,463	
X2^2	0,008	0,017	0,465	0,649	-0,028	0,043	-0,028	0,043	
X3^2	-20,869	23,834	-0,876	0,396	-71,987	30,249	-71,987	30,249	
X2*X3	0,622	1,100	0,566	0,580	-1,737	2,981	-1,737	2,981	

Рисунок 4.9 – Вывод итогов вспомогательной регрессии теста Уайта

Тогда вспомогательная регрессия принимает вид:

$$\hat{u}^2 = 0,033 - 0,036x_2 - 0,190x_3 + 0,008x_2^2 - 20,869x_3^2 + 0,622x_2x_3.$$

3) проверяется гипотеза о статистической значимости уравнения вспомогательной регрессии. Исследуется статистика

$$n \cdot R^2 \sim \chi_m^2,$$

где R^2 – множественный коэффициент детерминации для вспомогательной регрессии;

n – количество наблюдений;

m – количество оцениваемых параметров во вспомогательной регрессии за исключением свободного члена.

Если $n \cdot R^2 > \chi_m^2$, то нулевая гипотеза о гомоскедастичности остатков отвергается, то есть делается вывод о присутствии *гетероскедастичности*.

По данным рисунка 4.9 находим: $n \cdot R^2 = 20 \cdot 0,345 = 6,89$. Критическое значение рассматриваемой статистики на 5 % уровне значимости при 5 степенях свободы составляет $\chi_5^2 = 11,07$. Следовательно, $n \cdot R^2 < \chi_m^2$, т.е. нулевая гипотеза о гомоскедастичности остатков принимается.

2 Если будет обнаружена гетероскедастичность остатков, примените для исходных данных обобщенный метод наименьших квадратов (ОМНК), предполагая, что $\sigma^2(e_i) = \sigma^2 x_i^2$

Для устранения гетероскедастичности необходимо применить к исходным данным обобщенный метод наименьших квадратов в предположении, что $\sigma^2(u_i) = \sigma^2 x_i^2$. Проверка по тесту Голдфелда-Квандта показала, что имеет место гетероскедастичность остатков, причем их дисперсия зависит от фактора x_2 . Тогда для устранения гетероскедастичности необходимо найти параметры уравнения:

$$\frac{y_i}{x_2} = a_0 \frac{1}{x_2} + a_1 + a_2 \frac{x_3}{x_2} + \frac{u}{x_2}.$$

После применения МНК к этому уравнению (рисунок 4.10) получим:

$$\frac{y_i}{x_2} = 0,218 \frac{1}{x_2} + 0,002 + 5,022 \frac{x_3}{x_2}.$$

	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U					
112	Год	Квартал	Y/X2	1/X2	X3/X2		ВЫВОД ИТОГОВ													
113		1	0,286	0,714	0,007															
114		2	0,263	0,526	0,005		Регрессионная статистика													
115		3	0,286	0,476	0,010		Множественный R									0,573				
116	1	4	0,250	0,500	0,010		R-квадрат									0,328				
117		1	0,273	0,909	0,018		Нормированный									0,249				
118		2	0,286	0,714	0,007		Стандартная ошибка									0,038				
119		3	0,313	0,625	0,019		Наблюдения									20				
120	2	4	0,286	0,714	0,014															
121		1	0,250	0,500	0,015		Дисперсионный анализ													
122		2	0,231	0,385	0,012			df	SS	MS	F	Значимость F								
123		3	0,280	0,400	0,012		Регрессия	2	0,012	0,006	4,144	0,034								
124	3	4	0,200	0,400	0,012		Остаток	17	0,024	0,001										
125		1	0,318	0,455	0,018		Итого	19	0,036											
126		2	0,348	0,435	0,017															
127		3	0,360	0,400	0,016		Коэффициенты, стандартная ошибка статистика									P-Значение	Нижние 95%	Верхние 95%	нижние 95,0%	Верхние 95,0%
128	4	4	0,318	0,455	0,018		Y-пересечение	0,218	0,046	4,781	0,000	0,122	0,314	0,122	0,314					
129		1	0,320	0,400	0,020		1/X2	0,002	0,061	0,031	0,976	-0,126	0,130	-0,126	0,130					
130		2	0,346	0,385	0,019		X3/X2	5,022	1,787	2,811	0,012	1,253	8,791	1,253	8,791					
131		3	0,269	0,385	0,019															
132	5	4	0,364	0,455	0,023		ВЫВОД ОСТАТКА													
133							Прогнозы													
134							Наблюдение	Y/X2	Остатки											
135							1	0,255	0,030											
136							2	0,245	0,018											
137							3	0,267	0,019											
138							4	0,269	-0,019											
139							5	0,311	-0,038											
139							6	0,255	0,030											

Рисунок 4.10 – Вывод итога ОМНК

Проверка регрессионных остатков этого уравнения по тесту Уайта дает следующий результат.

Оцениваем вспомогательную регрессию вида:

$$u^2 = \alpha_0 + \alpha_1 \frac{1}{x_2} + \alpha_2 \frac{x_3}{x_2} + \alpha_3 \left(\frac{1}{x_2} \right)^2 + \alpha_4 \left(\frac{x_3}{x_2} \right)^2 + \alpha_5 \frac{1}{x_2} \cdot \frac{x_3}{x_2} + \nu,$$

где ν – нормально распределенная ошибка.

Используем для этого инструмент **Регрессия в Excel** (рисунок 4.11).

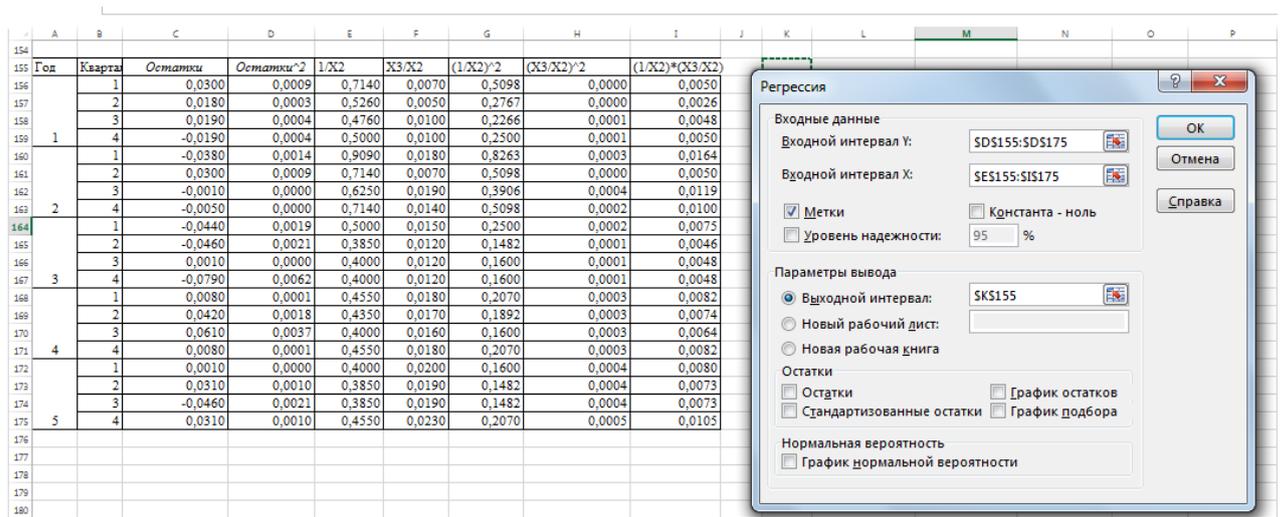


Рисунок 4.11 – Диалоговое окно ввода параметров инструмента «Регрессия»

Результаты оценки параметров вспомогательной регрессии представлены на рисунке 4.12.

Вывод итогов								
Регрессионная статистика								
Множественный R	0,52							
R-квадрат	0,27							
Нормированный R-кв:	0,01							
Стандартная ошибка	0,00							
Наблюдения	20							
Дисперсионный анализ								
	df	SS	MS	F	Значимость F			
Регрессия	5	0,00001	0,00000	1,03652	0,43427			
Остаток	14	0,00003	0,00000					
Итого	19	0,00005						
	Коэффициенты	стандартная ошибка	t-статистика	P-значения	Нижние 95%	Верхние 95%	Нижние 95,0%	Верхние 95,0%
Y-пересечение	0,011	0,009	1,142	0,273	-0,009	0,031	-0,009	0,031
1/X2	-0,037	0,023	-1,623	0,127	-0,086	0,012	-0,086	0,012
X3/X2	0,322	0,704	0,457	0,655	-1,189	1,833	-1,189	1,833
(1/X2)^2	0,030	0,019	1,566	0,140	-0,011	0,070	-0,011	0,070
(X3/X2)^2	-11,889	16,915	-0,703	0,494	-48,168	24,390	-48,168	24,390
(1/X2)*(X3/X2)	-0,131	0,669	-0,196	0,847	-1,567	1,304	-1,567	1,304

Рисунок 4.12 – Вывод итогов вспомогательной регрессии теста Уайта

Вспомогательная регрессия принимает вид:

$$\hat{u}^2 = 0,011 - 0,037 \frac{1}{x_2} + 0,322 \frac{x_3}{x_2} + 0,030 \left(\frac{1}{x_2} \right)^2 - 11,889 \left(\frac{x_3}{x_2} \right)^2 - 0,131 \left(\frac{1}{x_2} \cdot \frac{x_3}{x_2} \right).$$

По данным рисунка 4.12 находим: $n \cdot R^2 = 20 \cdot 0,27 = 5,4$. Критическое значение рассматриваемой статистики на 5 % уровне значимости при 5 степенях свободы составляет $\chi_5^2 = 11,07$. Следовательно, $n \cdot R^2 < \chi_m^2$, т.е. нулевая гипотеза о гомоскедастичности остатков принимается.

Контрольные вопросы к лабораторной работе № 4

- 1 Что такое гетероскедастичность регрессионных остатков?
- 2 Каковы причины возникновения гетероскедастичности регрессионных остатков?
- 3 Как проверить наличие гомо- или гетероскедастичности остатков?
- 4 Какие методы используются для устранения и смягчения гетероскедастичности регрессионных остатков?
- 5 Каковы последствия гетероскедастичности остатков?
- 6 В чем сущность обобщенного МНК?

Тесты для самоконтроля к лабораторной работе № 4

- 1 Способ корректировки гетероскедастичности:
 - А) метод взвешенных наименьших квадратов;
 - Б) доступный обобщенный метод наименьших квадратов;
 - В) косвенный метод наименьших квадратов;
 - Г) двухшаговый метод наименьших квадратов;
 - Д) авторегрессионные модели.

- 2 Последствия гетероскедастичности:
 - А) оценка дисперсии возмущений – смещена;
 - Б) оценка автоковариационной матрицы оценок параметров – смещена;
 - В) оценка дисперсии возмущений – несмещена;
 - Г) оценка параметров модели смещена;
 - Д) интервальные оценки параметров и значений эндогенной переменной адекватны.

3 Причины гетероскедастичности:

- А) исследование неоднородных объектов;
- Б) характер наблюдений
- В) ошибки спецификации;
- Г) избыточный объем совокупности;
- Д) ошибки измерений.

4 Для регрессионной модели с гетероскедастичностью автоковариационная матрица вектора возмущений:

- А) симметричная положительно определенная, с произвольными элементами;
- Б) диагональная, с различными элементами;
- В) диагональная, с равными элементами;
- Г) симметричная положительно определенная, с равными диагональными элементами.

5 Гетероскедастичность ошибок в регрессионных моделях означает, что они имеют:

- А) одинаковую дисперсию для всех наблюдений;
- Б) увеличивающую (уменьшающуюся) дисперсию для всех наблюдений;
- В) одинаковое математическое ожидание для всех наблюдений;
- Г) увеличивающее (уменьшающееся) математическое ожидание для всех наблюдений.

6 Какое из следующих утверждений не верно в случае гетероскедастичности остатков?

- А) Выводы по t и F - статистикам являются ненадежными;
- Б) Гетероскедастичность проявляется через низкое значение статистики Дарбина-Уотсона;

- В) При гетероскедастичности оценки остаются эффективными;
- Г) Оценки являются смещенными;
- Д) Гетероскедастичность проявляется через высокое значение статистики Дарбина-Уотсона.

7 Каким методом можно воспользоваться для устранения гетероскедастичности?

- А) Обобщенным методом наименьших квадратов;
- Б) Взвешенным методом наименьших квадратов;
- В) Методом максимального правдоподобия;
- Г) Двухшаговым методом наименьших квадратов.

8 Какие методы можно применить для обнаружения гетероскедастичности?

- А) тест Голфелда-Квандта;
- Б) тест ранговой корреляции Спирмена;
- В) тест Дарбина- Уотсона;
- Г) тест Чоу;
- Д) метод рядов.

9 На чем основан тест ранговой корреляции Спирмена?

- А) На использовании t – статистики;
- Б) На использовании F – статистики;
- В) На использовании χ^2 ;
- Г) На графическом анализе остатков.

10 Какое из следующих утверждений верно в случае гетероскедастичности остатков?

- А) Выводы по t и F - статистикам являются ненадежными;

Б) Гетероскедастичность проявляется через низкое значение статистики Дарбина-Уотсона;

В) При гетероскедастичности оценки остаются эффективными;

Г) Оценки параметров уравнения регрессии являются смещенными.

5 Лабораторная работа № 5. Нарушения допущений классической модели линейной регрессии: автокорреляция

Цель выполнения лабораторной работы: научиться проверять регрессионные остатки на независимость, а также смягчать последствия нарушения допущений классической модели линейной регрессии.

Задания:

По данным лабораторной работы № 2 выполните следующие задания:

1 Проверить остатки на наличие автокорреляции первого порядка, используя критерий Дарбина – Уотсона и Q - статистику Льюинга – Бокса.

2 Если гипотеза об отсутствии автокорреляции первого порядка не будет отвергнута, то применить ОМНК для оценивания параметров уравнения регрессии.

Реализация типовых заданий

1 Проверить остатки на наличие автокорреляции первого порядка, используя критерий Дарбина – Уотсона и Q - статистику Льюинга – Бокса

Важной предпосылкой построения качественной регрессионной модели по МНК является независимость значений регрессионных остатков (u_i) от значений остатков во всех других наблюдениях. Отсутствие зависимости гарантирует отсутствие коррелированности между регрессионными остатками в любых номерах наблюдений ($\text{cov}(u_i, u_j) = 0$ при $i \neq j$) и, в частности между соседними остатками ($\text{cov}(u_i, u_{i-1}) = 0$ при $i = 2, 3, \dots, n$).

Автокорреляция (последовательная корреляция) определяется как корреляция между показателями, упорядоченными во времени (временные ряды) или в пространстве (перекрестные данные). Автокорреляция регрессионных остатков обычно встречается в регрессионном анализе при использовании данных временных рядов. При использовании перекрестных данных наличие автокорреляции (пространственной корреляции) крайне редко.

Для проверки на наличие (отсутствие) автокорреляции первого порядка широкое применение получил тест **Дарбина-Уотсона**. Тест основан на вычислении статистики DW :

$$DW = \frac{\sum_{i=2}^n (u_i - u_{i-1})^2}{\sum_{i=1}^n u_i^2},$$

где u_i – регрессионные остатки.

Выдвигается гипотеза H_0 об отсутствии автокорреляции остатков. При сравнении расчетного значения статистики ($DW < 2$) с критическими d_l и d_u на заданном уровне значимости α , с числами степеней свободы p (число экзогенных переменных в уравнении регрессии) и n (объем выборки) возможны следующие варианты.

1 Если $DW < d_l$, то гипотеза H_0 отвергается.

2 Если $DW > d_u$, то гипотеза H_0 не отвергается.

3 Если $d_l < DW < d_u$, то нельзя сделать определенного вывода по имеющимся исходным данным (зона неопределенности).

При $DW > 2$, то с табличными значениями сравнивается величина $(4 - DW)$.

Как видно по формуле критерия, в расчетах участвуют только регрессионные остатки. Их значения были найдены в лабораторной работе

№4 с использованием инструмента анализа данных **Регрессия** (см. лаб. раб. №4, тест ранговой корреляции Спирмена).

В результате проведенных расчетов получено значение критерия Дарбина-Уотсона $DW=1,87$ (рисунок 5.1). Критические значения критерия Дарбина-Уотсона, при $p=2$ и $n=20$ составили $d_l=1,10$ и $d_u=1,54$ (таблица Б.3).

Так как выполняется неравенство $DW > d_u$, то мы не можем отвергнуть гипотезу H_0 – в ряду остатков отсутствует автокорреляция первого порядка.

	A	B	C	D	E	F	G	H	I	J	K
1	Расчетные значения для критерия Дарбина-Уотсона										
2	Остатки	ui	ui-1	(ui-ui-1)^2	остатки^2						
3	0,044	0,062	0,044	0,0003	0,002	DW	=D23/E23	=	1,87		
4	0,062	0,061	0,062	0,0000	0,004	dl	1,1				
5	0,061	-0,022	0,061	0,0069	0,004	du	1,54				
6	-0,022	-0,075	-0,022	0,0028	0,000						
7	-0,075	0,044	-0,075	0,0142	0,006						
8	0,044	-0,025	0,044	0,0048	0,002						
9	-0,025	-0,024	-0,025	0,0000	0,001						
10	-0,024	-0,090	-0,024	0,0044	0,001						
11	-0,090	-0,088	-0,090	0,0000	0,008						
12	-0,088	0,028	-0,088	0,0135	0,008						
13	0,028	-0,172	0,028	0,0400	0,001						
14	-0,172	0,009	-0,172	0,0328	0,030						
15	0,009	0,093	0,009	0,0071	0,000						
16	0,093	0,160	0,093	0,0045	0,009						
17	0,160	0,009	0,160	0,0228	0,026						
18	0,009	-0,008	0,009	0,0003	0,000						
19	-0,008	0,076	-0,008	0,0071	0,000						
20	0,076	-0,124	0,076	0,0400	0,006						
21	-0,124	0,041	-0,124	0,0272	0,015						
22	0,041				0,002						
23	Итого	-	-	0,2284	0,12227						

Рисунок 5.1 – Расчет критерия Дарбина-Уотсона

На практике могут существовать автокорреляции в остатках не только первого, но и более высокого порядков. В таких случаях критерий Дарбина-Уотсона не применим.

Рассмотрим гипотезу об отсутствии автокорреляции до порядка ρ :

$$H_0 : r(1) = r(2) = \dots = r(\rho) = 0,$$

где $r(1), r(2), \dots, r(\rho)$ $r(\rho)$ – коэффициенты автокорреляции первого, второго и т.д. порядков.

Для проверки этой гипотезы используется Q-тест Льюинга – Бокса, дающий корректные выводы даже для выборок малого объема.

Реализация данного теста предполагает использование Q- статистики, значение которой определяется по формуле:

$$Q = n(n+2) \cdot \sum_{\tau=1}^{\rho} \frac{r^2(\tau)}{n-\tau},$$

где $r(\tau)$ – выборочные значения автокорреляционной функции;

τ – величина лага;

ρ – максимальная величина лага;

n – число наблюдений.

Q - статистика имеет χ^2 – распределение с ρ степенями свободы. Если Q – статистика меньше табличного χ^2 , то гипотеза H_0 об отсутствии автокорреляции не выше порядка ρ не отвергается.

Рассчитаем для нашего примера Q – статистику. Для этого необходимо определить коэффициенты автокорреляции. Максимальная величина лага не должна превышать $\frac{1}{4}$ числа наблюдений, т.е. в рассматриваемом примере $\rho \leq 5$ (т.к. $n=20$, то $20/4=5$). Следовательно, нужно определить коэффициенты автокорреляции до пятого порядка. Для этого используем функцию Excel **Сервис – Анализ данных – Корреляция** или **Данные – Анализ данных – Корреляция** (аналогично второму заданию лабораторной работы №2).

Предварительно необходимо сформировать матрицу из регрессионных остатков, как на рисунке 5.2.

Столбец u_i состоит из значений столбца «Остатки» взятых со второго по последнее наблюдение; столбец u_{i-1} состоит из значений столбца

«Остатки» взятых с первого по предпоследнего наблюдения, т.е. лаг (расхождение, шаг между наблюдениями) равен 1. Используя данные этих столбцов мы рассчитаем коэффициент автокорреляции первого порядка (r_1).

Для расчета коэффициента автокорреляции второго порядка (r_2), столбец u_i состоит из третьего по последнее наблюдение столбца «Остатки»; столбец u_{i-2} состоит из значений столбца «Остатки» взятых с первого до $n-2$ наблюдения, т.е. лаг (расхождение, шаг между наблюдениями) равен 2. Аналогично формируем столбцы для расчета коэффициента автокорреляции третьего и следующих порядков ($r_3 - r_5$).

Ввод данных для расчета коэффициента автокорреляции первого порядка (r_1) с помощью функции «Корреляция» представлен на рисунке 5.2.

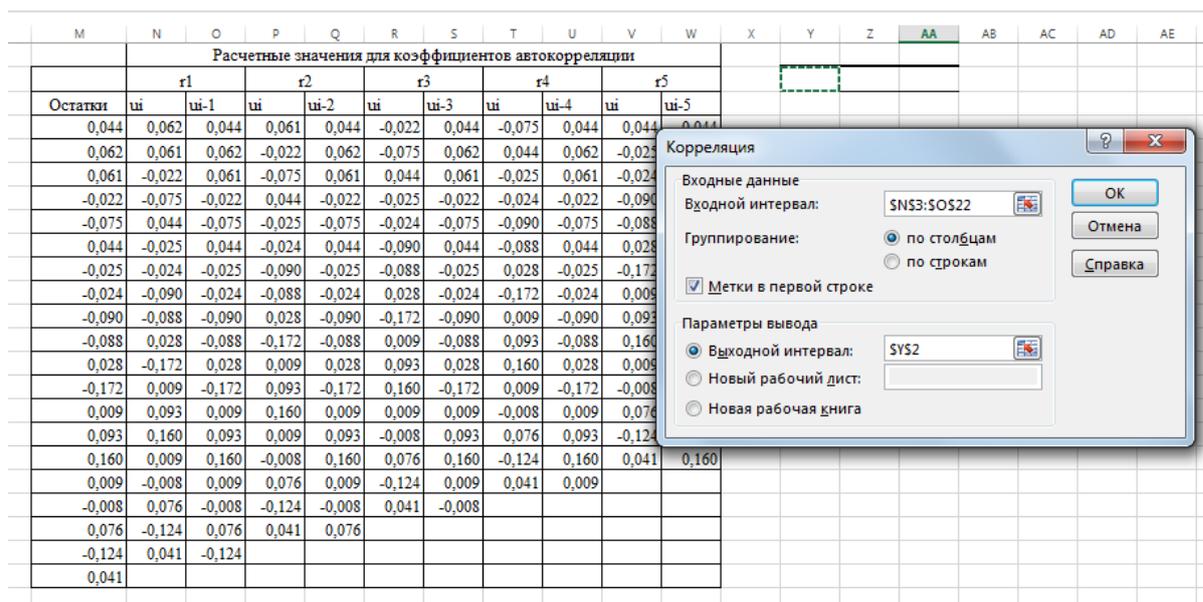


Рисунок 5.2 – Расчет коэффициента автокорреляции первого порядка с помощью функции «Корреляция»

Ввод данных для расчета коэффициента автокорреляции второго порядка (r_2) с помощью функции «Корреляция» представлен на рисунке 5.3.

Аналогично рассчитываются коэффициенты автокорреляции более высоких порядков.

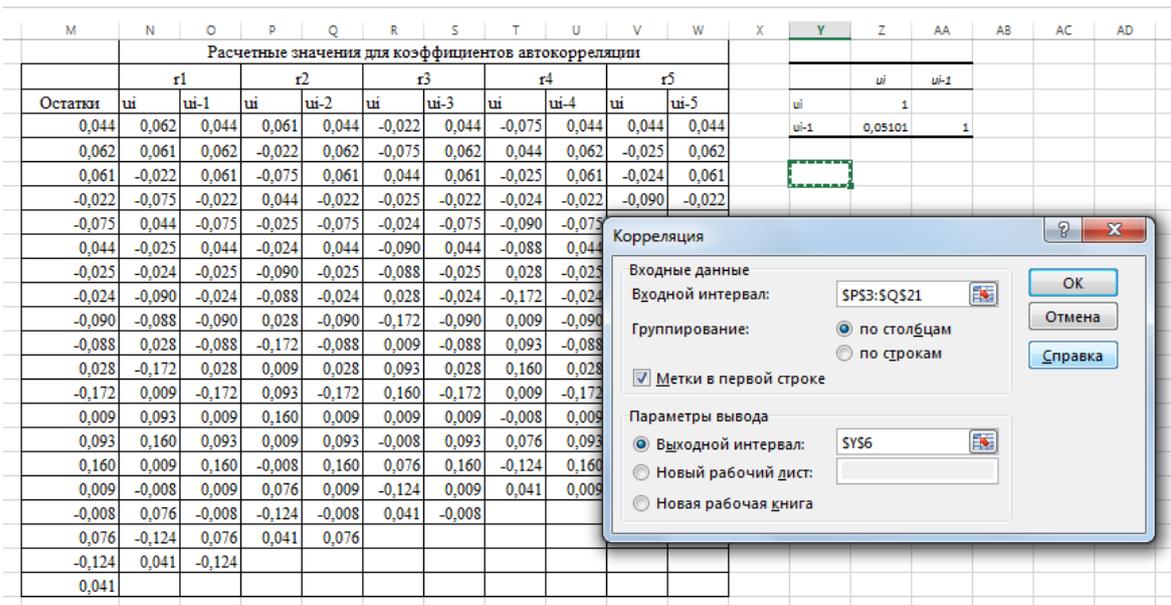


Рисунок 5.3 – Расчет коэффициента автокорреляции второго порядка с помощью функции «Корреляция»

Результаты расчетов всех коэффициентов автокорреляции с помощью функции «Корреляция» а также Q-статистики Льюинга – Бокса представлены на рисунке 5.4.

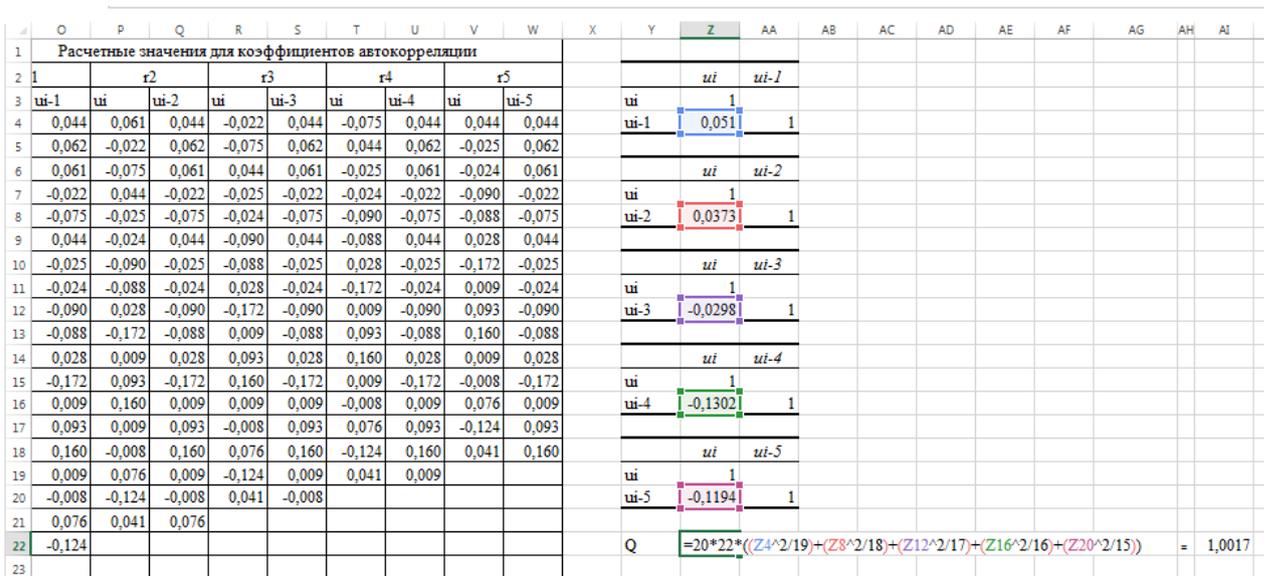


Рисунок 5.4 – Расчет Q-статистики Льюинга – Бокса

Подставив полученное значение в формулу, получим:

$$Q = 20(20 + 2) \cdot \left(\frac{(0,051)^2}{20-1} + \frac{(0,0373)^2}{20-2} + \frac{(-0,0298)^2}{20-3} + \frac{-(0,1302)^2}{20-4} + \frac{(-0,1194)^2}{20-5} \right) = 1,0017.$$

Критическое значение $\chi_{0,05;5}^2 = 11,07$ (таблица Б.4).

Фактическое значение статистики меньше критического, следовательно, гипотеза принимается, т.е. в ряду остатков отсутствует автокорреляция не выше пятого порядка.

По всем проведенным тестам можно сделать вывод об отсутствии автокорреляции в регрессионных остатках.

2 Если гипотеза об отсутствии автокорреляции первого порядка не будет отвергнута, то применить ОМНК для оценивания параметров уравнения регрессии

При обнаружении автокорреляции в регрессионных остатках необходимо в первую очередь проанализировать правильность спецификации модели. Если после ряда возможных усовершенствований регрессии (уточнения состава объясняющих переменных либо изменения формы зависимости) автокорреляция по-прежнему имеет место, то возможно, это связано с внутренними свойствами ряда регрессионных остатков. В этом случае возможны преобразования, устраняющие автокорреляцию. Среди них выделяется авторегрессионная схема первого порядка AR(1), которая может быть обобщена в AR(k), k=2,3,... Для применения указанных схем необходимо оценить коэффициент корреляции между регрессионными остатками. Это может быть сделано различными методами: на основе статистики Дарбина-Уотсона, Кохрана-Оркатта, Хилдрета-Лу и др.

В линейной регрессионной модели либо в моделях, сводящихся к линейной, наиболее целесообразным и простым преобразованием является авторегрессионная схема первого порядка AR(1). Согласно данному подходу,

выполняется преобразование эндогенных и экзогенных переменных, а также свободного члена следующим образом:

$$\begin{aligned}
 y_i^* &= y_i - \rho \cdot y_{i-1}; \\
 x_{1i}^* &= x_{1i} - \rho \cdot x_{1i-1}; \\
 x_{2i}^* &= x_{2i} - \rho \cdot x_{2i-1}; \\
 &\dots \\
 x_{pi}^* &= x_{pi} - \rho \cdot x_{pi-1}; \\
 b_0^* &= b_0(1 - \rho).
 \end{aligned}$$

Тогда, уравнение множественной регрессии примет вид:

$$y_i^* = b_0^* + b_1 x_{i1}^* + b_2 x_{i2}^* + \dots + b_p x_{ip}^*.$$

Однако способ вычисления y_i^* и x_i^* приводит к потере первого наблюдения если мы не обладаем предшествующим ему наблюдением). Число степеней свободы уменьшится на единицу, что при больших выборках не так существенно, но при малых выборках может привести к потере эффективности. Эта проблема обычно преодолевается с помощью поправки Прайса-Винстена:

$$\begin{aligned}
 y_1^* &= y_1 \cdot \sqrt{1 - \rho^2}; \\
 x_{11}^* &= x_{11} \cdot \sqrt{1 - \rho^2}; \\
 x_{21}^* &= x_{21} \cdot \sqrt{1 - \rho^2}; \\
 &\dots \\
 x_{p1}^* &= x_{p1} \cdot \sqrt{1 - \rho^2}.
 \end{aligned}$$

На практике значение коэффициента ρ обычно неизвестно и его необходимо оценивать. Существует несколько методов оценивания. Рассмотрим некоторые из них.

При большом числе наблюдений в качестве оценки коэффициента ρ может быть взят коэффициент автокорреляции первого порядка $r(1)$ найденный на основе статистики [Дарбина-Уотсона](#):

$$\rho = r(1) \approx 1 - \frac{DW}{2}.$$

Другим возможным методом оценивания ρ является итеративный процесс, называемый **методом Кохрана-Оркатта**. Он включает следующие этапы:

- 1 Оценивается по МНК регрессия

$$\tilde{y}_i = b_0 + b_1 x_{i1} + b_2 x_{i2} + \dots + b_p x_{ip}$$

и для нее определяются оценки регрессионных остатков $u_i = y_i - \tilde{y}_i$.

- 2 С использованием схем AR(1) оценивается регрессионная зависимость

$$u_i = \tilde{\rho} \cdot u_{i-1} + v_i,$$

где $\tilde{\rho}$ - оценка коэффициента ρ .

- 3 На основе данной оценки строится уравнение:

$$y_i - \tilde{\rho} \cdot y_{i-1} = \alpha \cdot (1 - \tilde{\rho}) + \beta_1 \cdot (x_{i1} - \tilde{\rho} \cdot x_{i-1,1}) + \beta_2 \cdot (x_{i2} - \tilde{\rho} \cdot x_{i-1,2}) + \dots + \beta_p \cdot (x_{ip} - \tilde{\rho} \cdot x_{i-1,p}) + (u_i - \tilde{\rho} \cdot u_{i-1}),$$

с помощью которого оцениваются коэффициенты α и β (в этом случае значение $\tilde{\rho}$ известно).

4 Значения $b_0 = \alpha(1 - \tilde{\rho})$, $b_1 = \beta_1$, $b_2 = \beta_2, \dots$, $b_p = \beta_p$ подставляют в исходное уравнение регрессии. Вновь вычисляются оценки u_i отклонений и процесс возвращается ко второму этапу.

Чередование этапов осуществляется до тех пор, пока не будет достигнута требуемая точность, т.е. пока разность между предыдущей и последующей оценками $\tilde{\rho}$ не станет меньше любого наперед заданного числа.

В пакетах прикладных программ широко используется метод Хилдрета-Лу. По данному методу регрессия вида:

$$y_i - \rho \cdot y_{i-1} = \beta_0 \cdot (1 - \rho) + \beta_1 \cdot (x_{1i} - \rho \cdot x_{1i-1}) + \beta_2 \cdot (x_{2i} - \rho \cdot x_{2i-1}) + \dots + \beta_p \cdot (x_{pi} - \rho \cdot x_{pi-1}) + (u_i - \rho \cdot u_{i-1})$$

оценивается для каждого возможного значения ρ из отрезка $[-1; 1]$ с любым шагом (например, 0,001; 0,01 и т.д.). Величина $\tilde{\rho}$, дающая наименьшую стандартную ошибку регрессии, принимается в качестве оценки коэффициента ρ . И значения b_0^*, b_1, \dots, b_p оцениваются из уравнения регрессии

$$y_i^* = b_0^* + b_1 x_{i1}^* + b_2 x_{i2}^* + \dots + b_p x_{ip}^*$$

именно с данным значением $\tilde{\rho}$.

Контрольные вопросы к лабораторной работе № 5

- 1 Что понимается под автокорреляцией регрессионных остатков?
- 2 Как оценивается отсутствие автокорреляции остатков при построении регрессионной модели?
- 3 Какие методы используются для устранения и смягчения автокорреляции в регрессионных остатках?

- 4 Каковы последствия автокорреляции остатков?
- 5 В чем сущность обобщенного МНК?

Тесты для самоконтроля к лабораторной работе № 5

- 1 Для чего применяется критерий Дарбина - Уотсона:
 - А) обнаружения автокорреляции в остатках;
 - Б) обнаружения циклической составляющей;
 - В) обнаружения гетероскедастичности в остатках;
 - Г) для проверки подчинения случайного компонента нормальному закону распределения.

- 2 Каким методом можно воспользоваться для устранения автокорреляции?
 - А) Обобщенным методом наименьших квадратов;
 - Б) Взвешенным методом наименьших квадратов;
 - В) Методом максимального правдоподобия;
 - Г) Двухшаговым методом наименьших квадратов.

- 3 С помощью каких методов нельзя устранить автокорреляцию остатков?
 - А) Обобщенным методом наименьших квадратов;
 - Б) Взвешенным методом наименьших квадратов;
 - В) Методом максимального правдоподобия;
 - Г) Двухшаговым методом наименьших квадратов;
 - Д) Косвенным методом наименьших квадратов.

- 4 Как называется нарушение допущения о независимости остатков?
 - А) Мультиколлинеарность;
 - Б) Автокорреляция;
 - В) Гетероскедастичность;

Г) Гомоскедастичность.

5 Автокорреляцией в статистике называется:

А) зависимость вариации значений одного показателя от вариации значений другого;

Б) зависимость между цепными уровнями;

В) отклонения от тенденции;

Г) зависимость последующего уровня динамического ряда от предыдущего.

6 Для регрессионной модели с автокорреляцией автоковариационная матрица вектора возмущений:

А) симметричная положительно определенная, с произвольными элементами;

Б) диагональная, с различными элементами;

В) диагональная, с равными элементами;

Г) симметричная положительно определенная, с равными диагональными элементами.

7 Причины автокорреляции:

А) исследование неоднородных объектов;

Б) характер наблюдений;

В) ошибки параметризации;

Г) избыточный объем совокупности;

Д) ошибки измерений.

8 Последствия автокорреляции:

А) оценка дисперсии возмущений – несмещена;

Б) оценка автоковариационной матрицы оценок параметров – смещена;

В) оценка автоковариационной матрицы оценок параметров – несмещена;

Г) оценка параметров модели смещена;

Д) интервальные оценки параметров и значений эндогенной переменной неадекватна.

9 Способ корректировки автокорреляции:

А) метод взвешенных наименьших квадратов;

Б) доступный обобщенный метод наименьших квадратов;

В) косвенный метод наименьших квадратов

Г) авторегрессионные модели.

10 Какой из перечисленных методов не может быть применен для обнаружения автокорреляции?

А) метод рядов;

Б) критерий Льюинга-Бокса;

В) критерий Дарбина-Уотсона;

Г) тест ранговой корреляции Спирмена;

Д) тест Уайта.

6 Лабораторная работа № 6 Нелинейная регрессия и способы линеаризации моделей нелинейных по переменным

Цель выполнения лабораторной работы: научить строить статистически значимые модели некоторых видов нелинейной регрессии и давать содержательный анализ полученным результатам.

Задания:

По исходным данным лабораторной работы № 2 для эндогенной переменной и экзогенными переменными, не вошедшими в линейное уравнение множественной регрессии выполните следующие задания:

1 Рассчитайте параметры уравнения гиперболической (обратной) регрессии вида:

$$\tilde{y}_x = b_0 + \frac{b_1}{x_1} + \frac{b_2}{x_2} + \dots + \frac{b_p}{x_p}.$$

2 Оцените тесноту связи между эндогенной переменной и экзогенными переменными, вошедшими в уравнение регрессии.

3 Рассчитайте средний коэффициент эластичности.

4 Оцените качество уравнения через среднюю ошибку аппроксимации.

5 Оцените с помощью F-критерия Фишера-Снедекора значимость уравнения гиперболической (обратной) регрессии.

6 Проверьте регрессионные остатки на гетероскедастичность с помощью любых тестов, представленных в лабораторной работе № 4. Если будет обнаружена гетероскедастичность остатков, примените для исходных данных ОМНК, предполагая, что $\sigma^2(u_i) = \sigma^2 x_i^2$.

7 Проверить остатки на наличие автокорреляции первого порядка, используя тесты из лабораторной работы № 5. Если гипотеза об отсутствии автокорреляции первого порядка не будет отвергнута, то применить ОМНК для оценивания параметров уравнения регрессии.

Реализация типовых заданий

1 Рассчитайте параметры уравнения гиперболической (обратной) регрессии вида:

$$\tilde{y}_x = b_0 + \frac{b_1}{x_1} + \frac{b_2}{x_2} + \dots + \frac{b_p}{x_p}.$$

В линейное уравнение регрессии, полученное в лабораторной работе № 2, не вошли три экзогенные переменные x_1 , x_4 и x_5 .

Построим матрицу парных коэффициентов корреляции и отберем информативные факторы в модели (рисунок 6.1).

	Y	X1	X4	X5
Y	1			
X1	0,918215	1		
X4	-0,18443	-0,33205	1	
X5	0,908832	0,964957	-0,31165	1

Рисунок 6.1 – Матрица коэффициентов парной корреляции

По матрице парных коэффициентов корреляции (рисунок 6.1) видно, что переменные x_1 и x_5 коллинеарны, т.к. коэффициент корреляции между ними превышает 0,75. Таким образом, фактор x_5 исключается из уравнения регрессии.

Поэтому в качестве экзогенных переменных в нелинейную регрессионную модель войдут x_1 , x_4 .

Следовательно, уравнение примет вид: $\tilde{y}_i = b_0 + \frac{b_1}{x_1} + \frac{b_4}{x_4}$.

Оно относится к первому классу нелинейных регрессионных моделей «нелинейные относительно включенных в анализ объясняющих переменных, но линейные по оцениваемым параметрам».

Для оценивания неизвестных параметров моделей первого класса применяют метод «замена переменных», суть которого состоит в замене нелинейных объясняющих переменных новыми линейными переменными и сведении нелинейной регрессии к линейной.

Тогда, для рассматриваемой модели необходимо провести замену переменных:

$$z_1 = \frac{1}{x_1}, z_4 = \frac{1}{x_4}.$$

Получим уравнение множественной линейной регрессии:

$$\tilde{y}_i = b_0 + b_1 \cdot z_1 + b_4 \cdot z_4.$$

Для нахождения параметров данного уравнения воспользуемся инструментом анализа данных **Регрессия** (см. лабораторную работу № 2). Уравнение построим используя данные таблицы, представленной на рисунке 6.2.

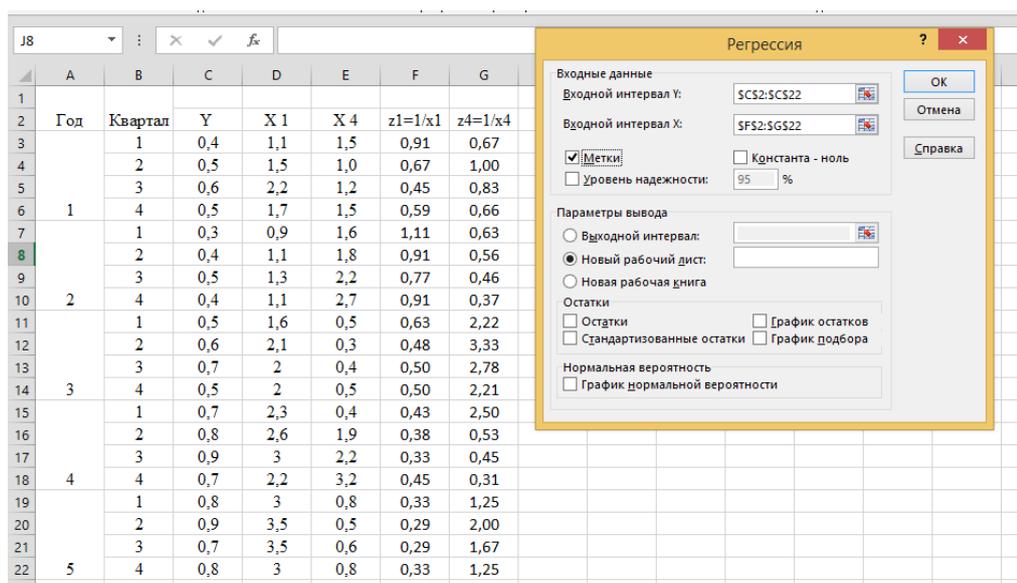


Рисунок 6.2 – Исходные данные для построения гиперболической модели

Результаты регрессионного анализа представлены на рисунке 6.3.

	A	B	C	D	E	F	G	H	I
1	Вывод итогов								
2									
3	Регрессионная статистика								
4	Множественный R	0,920264016							
5	R-квадрат	0,84688586							
6	Нормированный R-квадрат	0,828872431							
7	Стандартная ошибка	0,073389475							
8	Наблюдения	20							
9									
10	Дисперсионный анализ								
11		df	SS	MS	F	Значимость F			
12	Регрессия	2	0,506437744	0,253218872	47,01414117	1,18203E-07			
13	Остаток	17	0,091562256	0,005386015					
14	Итого	19	0,598						
15									
16		Коэффициенты	Стандартная ошибка	t-статистика	P-Значение	Нижние 95%	Верхние 95%	Нижние 95,0%	Верхние 95,0%
17	Y-пересечение	1,047063727	0,059211423	17,68347508	2,21482E-12	0,922138544	1,17198891	0,922138544	1,17198891
18	z1=1/x1	-0,707899669	0,07497847	-9,441372562	3,57133E-08	-0,866090413	-0,549708925	-0,866090413	-0,549708925
19	z4=1/x4	-0,02990625	0,019785576	-1,511517797	0,149022238	-0,071650167	0,011837666	-0,071650167	0,011837666
20									

Рисунок 6.3 – Результат применения инструмента «Регрессия»

Получено уравнение множественной линейной регрессии:

$$\tilde{y}_i = 1,047 - 0,708 \cdot z1 - 0,03 \cdot z4.$$

Оценивая параметры данного уравнения, замечаем, что статистически значимым является параметр при X1, (значимость параметров уравнения регрессии определяем, как в пятом задании лабораторной работы № 2). Следовательно, целесообразно строить уравнение гиперболической регрессии только с данным фактором.

В результате получаем равнение следующего вида: $\tilde{y}_i = 0,985 - 0,665 \cdot z1$ (рисунок 6.4).

	A	B	C	D	E	F	G	H	I
1	Вывод итогов								
2									
3	Регрессионная статистика								
4	Множественный R	0,909015041							
5	R-квадрат	0,826308346							
6	Нормированный R-квадрат	0,816658809							
7	Стандартная ошибка	0,075963299							
8	Наблюдения	20							
9									
10	Дисперсионный анализ								
11		df	SS	MS	F	Значимость F			
12	Регрессия	1	0,494132391	0,494132391	85,63192209	2,90571E-08			
13	Остаток	18	0,103867609	0,005770423					
14	Итого	19	0,598						
15									
16		Коэффициенты	Стандартная ошибка	t-статистика	P-Значение	Нижние 95%	Верхние 95%	Нижние 95,0%	Верхние 95,0%
17	Y-пересечение	0,9846132	0,043901459	22,42780112	1,3213E-14	0,892379656	1,076846744	0,892379656	1,076846744
18	z1=1/x1	-0,665173452	0,071881488	-9,253751785	2,90571E-08	-0,816190855	-0,514156048	-0,816190855	-0,514156048
19									

Рисунок 6.4 – Результат применения инструмента «Регрессия»

Следовательно, получим уравнение регрессии: $\tilde{y}_i = 0,985 - \frac{0,665}{x_i}$.

2 Оцените тесноту связи между эндогенной переменной и экзогенными переменными, вошедшими в уравнение регрессии

Подставляя в данное уравнение фактические значения x_1 , получаем теоретические значения эндогенной переменной \tilde{y}_i (рисунок 6.5 графа 7).

	A	B	C	D	E	F	G	H	I
1	Квартал	Y	X 1	X 4	z1=1/x1	z4=1/x4	\tilde{y}_i	$(y - \tilde{y}_i)^2$	$(y - \bar{y})^2$
2	1	2	3	4	5	6	7	8	9
3	1	0,4	1,1	1,5	0,91	0,67	0,4	0,0004	0,0961
4	2	0,5	1,5	1,0	0,67	1,00	0,5	0,0017	0,0441
5	3	0,6	2,2	1,2	0,45	0,83	0,7	0,0068	0,0121
6	4	0,5	1,7	1,5	0,59	0,66	0,6	0,0087	0,0441
7	1	0,3	0,9	1,6	1,11	0,63	0,2	0,0030	0,1681
8	2	0,4	1,1	1,8	0,91	0,56	0,4	0,0004	0,0961
9	3	0,5	1,3	2,2	0,77	0,46	0,5	0,0007	0,0441
10	4	0,4	1,1	2,7	0,91	0,37	0,4	0,0004	0,0961
11	1	0,5	1,6	0,5	0,63	2,22	0,6	0,0047	0,0441
12	2	0,6	2,1	0,3	0,48	3,33	0,7	0,0046	0,0121
13	3	0,7	2	0,4	0,50	2,78	0,7	0,0023	0,0001
14	4	0,5	2	0,5	0,50	2,21	0,7	0,0231	0,0441
15	1	0,7	2,3	0,4	0,43	2,50	0,7	0,0000	0,0001
16	2	0,8	2,6	1,9	0,38	0,53	0,7	0,0051	0,0081
17	3	0,9	3	2,2	0,33	0,45	0,8	0,0188	0,0361
18	4	0,7	2,2	3,2	0,45	0,31	0,7	0,0003	0,0001
19	1	0,8	3	0,8	0,33	1,25	0,8	0,0014	0,0081
20	2	0,9	3,5	0,5	0,29	2,00	0,8	0,0111	0,0361
21	3	0,7	3,5	0,6	0,29	1,67	0,8	0,0089	0,0001
22	4	0,8	3	0,8	0,33	1,25	0,8	0,0014	0,0081
23	Сумма	14,2	44,7				19,2	8,1039	9,7980
24	Средняя	0,71	2,235						
25									

Рисунок 6.5 – Исходные данные для нахождения индекса корреляции

По ним рассчитаем показатель тесноты связи – **индекс корреляции**:

$$\rho_{xy} = \sqrt{1 - \frac{\sum (y - \tilde{y}_i)^2}{\sum (y - \bar{y})^2}}.$$

Величина данного показателя находится в границах: $0 < \rho_{xy} < 1$. Чем ближе значение корреляционного отношения к единице, тем теснее связь рассматриваемых признаков, тем более надежно найденное уравнение регрессии. При $|\rho_{xy}| < 0,30$ связь считается слабой, при $0,30 < |\rho_{xy}| < 0,70$ - средней, при $|\rho_{xy}| > 0,70$ - сильной (тесной). Поскольку в расчете индекса корреляции используется соотношение факторной и общей суммы квадратов отклонений, то ρ_{xy}^2 имеет тот же смысл, что и коэффициент детерминации (индекс детерминации).

По данным рисунка 6.6 индекс корреляции составит:
 $\rho_{xy} = \sqrt{1 - \frac{8,1039}{9,7980}} = 0,42$. Следовательно, связь между признаками средней силы, тогда индекс детерминации составит $\rho_{xy}^2 = 0,17$, т.е. 17 % вариации интегрального показателя экономической безопасности организации объясняется вариацией объема реализованной продукции.

3 Рассчитайте средний коэффициент эластичности

Средние коэффициенты эластичности показывают, на сколько процентов от значения своей средней (\bar{y}) изменяется эндогенная переменная при изменении экзогенной переменной (x_i) на 1 % от своей средней (\bar{x}_i) при фиксированном воздействии на y всех прочих факторов, включенных в уравнение регрессии. Формула для расчета коэффициента эластичности имеет вид:

$$\varepsilon = f'(x) \cdot \frac{x}{y},$$

где $f'(x)$ – первая производная, характеризующая соотношение приростов эндогенной и экзогенной переменных для соответствующей формы связи.

Коэффициент эластичности для оцененного уравнения гиперболической регрессии составит:

$$\bar{\varepsilon} = \frac{-b}{a \cdot \bar{x} + b} = \frac{0,665}{0,985 \cdot 2,235 - 0,665} = 0,43 \%$$

Он показывает, что при увеличении на 1 % от своей средней объема реализованной продукции (x_1), интегральный показатель экономической безопасности организации (y) увеличивается на 0,43 % от своей средней.

4 Оцените качество уравнения через среднюю ошибку аппроксимации

По данным рисунка 6.6, графы 10 средняя ошибка аппроксимации для гиперболического уравнения регрессии будет равна:

$$\bar{A} = \frac{1}{20} \cdot 12,0342 \cdot 100 = 60,17 \%$$

Полученное уравнение имеет удовлетворительную точность.

	A	B	C	D	E	F	G	H	I	J
							\tilde{y}_i	$(y - \tilde{y}_i)^2$	$(y - \bar{y})^2$	$\frac{y - \tilde{y}_i}{y}$
1	Квартал	Y	X 1	X 4	z1=1/x1	z4=1/x4				
2	1	2	3	4	5	6	7	8	9	10
3	1	0,4	1,1	1,5	0,91	0,67	0,4	0,0004	0,0961	0,050225
4	2	0,5	1,5	1,0	0,67	1,00	0,5	0,0017	0,0441	0,082328
5	3	0,6	2,2	1,2	0,45	0,83	0,7	0,0068	0,0121	0,137103
6	4	0,5	1,7	1,5	0,59	0,66	0,6	0,0087	0,0441	0,186669
7	1	0,3	0,9	1,6	1,11	0,63	0,2	0,0030	0,1681	0,181561
8	2	0,4	1,1	1,8	0,91	0,56	0,4	0,0004	0,0961	0,050225
9	3	0,5	1,3	2,2	0,77	0,46	0,5	0,0007	0,0441	0,054117
10	4	0,4	1,1	2,7	0,91	0,37	0,4	0,0004	0,0961	0,050225
11	1	0,5	1,6	0,5	0,63	2,22	0,6	0,0047	0,0441	0,13776
12	2	0,6	2,1	0,3	0,48	3,33	0,7	0,0046	0,0121	0,113107
13	3	0,7	2	0,4	0,50	2,78	0,7	0,0023	0,0001	0,068534
14	4	0,5	2	0,5	0,50	2,21	0,7	0,0231	0,0441	0,304053
15	1	0,7	2,3	0,4	0,43	2,50	0,7	0,0000	0,0001	0,006561
16	2	0,8	2,6	1,9	0,38	0,53	0,7	0,0051	0,0081	0,089028
17	3	0,9	3	2,2	0,33	0,45	0,8	0,0188	0,0361	0,152346
18	4	0,7	2,2	3,2	0,45	0,31	0,7	0,0003	0,0001	0,025341
19	1	0,8	3	0,8	0,33	1,25	0,8	0,0014	0,0081	0,046389
20	2	0,9	3,5	0,5	0,29	2,00	0,8	0,0111	0,0361	0,117152
21	3	0,7	3,5	0,6	0,29	1,67	0,8	0,0089	0,0001	0,135091
22	4	0,8	3	0,8	0,33	1,25	0,8	0,0014	0,0081	0,046389
23	Сумма	14,2	44,7				19,2	8,1039	9,7980	12,0342
24	Средняя	0,71	2,235							0,60171
25										

Рисунок 6.6 – Исходные данные для нахождения ошибки аппроксимации

5 Оцените с помощью F-критерия Фишера-Снедекора значимость уравнения гиперболической (обратной) регрессии

Значимость уравнения гиперболической регрессии оценивается с помощью F-критерий Фишера, который рассчитывается по формуле

$$F = \frac{\rho_{xy}^2}{1 - \rho_{xy}^2} \cdot \frac{n - m}{m - 1}$$

Наблюдаемое значение F-критерия Фишера, по данным рисунка 6.4, составило $F_{набл} = 85,63$. Критическое значение F-критерия Фишера

$F_{крит(0,05;1;18)} = 4,41$. Так как $F_{набл} > F_{крит}$ уравнение статистически значимо.

6 Проверьте регрессионные остатки на гетероскедастичность с помощью любых тестов, представленных в лабораторной работе № 4. Если будет обнаружена гетероскедастичность остатков, примените для исходных данных ОМНК, предполагая, что $\sigma^2(u_i) = \sigma^2 x_i^2$.

Например, для проверки регрессионных остатков на гетероскедастичность воспользуемся [тестом Уайта](#) (см. лабораторную работу № 4).

Вспомогательная регрессия имеет вид:

$$\hat{u}^2 = 0,014 - 0,022z_1 + 0,009z_1^2 .$$

По результатам регрессионного анализа вспомогательной регрессии находим: $n \cdot R^2 = 20 \cdot 0,16 = 3,11$.

Критическое значение рассматриваемой статистики на 5 % уровне значимости при 2 степенях свободы составляет $\chi_2^2 = 5,99$. Следовательно, $n \cdot R^2 < \chi_m^2$, т.е. нулевая гипотеза о гомоскедастичности остатков принимается.

7 Проверить остатки на наличие автокорреляции первого порядка, используя тесты из лабораторной работы № 5. Если гипотеза об отсутствии автокорреляции первого порядка не будет отвергнута, то применить ОМНК для оценивания параметров уравнения регрессии

Например, наличие/отсутствие автокорреляции в регрессионных остатках проверим с помощью [Q-теста Льюинга-Бокса](#) (см. лабораторную работу № 5).

Подставив полученное значение в формулу, получим:

$$Q = 20(20 + 2) \cdot \left(\frac{(0,0287)^2}{20-1} + \frac{(0,0008)^2}{20-2} + \frac{(0,0051)^2}{20-3} + \frac{-(0,1892)^2}{20-4} + \frac{-(0,2701)^2}{20-5} \right) = 3,14.$$

Фактическое значение статистики меньше критического, следовательно, гипотеза принимается, т.е. в ряду остатков отсутствует автокорреляция не выше пятого порядка.

Контрольные вопросы к лабораторной работе № 6

- 1 Что понимается под спецификацией модели?
- 2 Назовите классы нелинейных моделей.
- 3 Запишите все виды моделей, нелинейных относительно включаемых переменных.
- 4 Как применяется МНК к моделям, нелинейным относительно включаемых переменных?
- 5 Какой нелинейной функцией может быть заменена парабола второй степени, если не наблюдается смена направленности связи признаков?
- 6 Как определить коэффициент эластичности для моделей, нелинейных относительно включаемых переменных?
- 7 Назовите показатели корреляции, используемых для нелинейных моделей относительно включаемых переменных.

Тесты для самоконтроля к лабораторной работе № 6

1 Какое уравнение соответствует следующей системе нормальных уравнений, применяемой для определения параметров уравнения (по методу наименьших квадратов):

$$\begin{cases} \sum y = n \cdot a_0 + a_1 \cdot \sum \frac{1}{x} \\ \sum \frac{y}{x} = a_0 \cdot \sum \frac{1}{x} + a_1 \cdot \sum \frac{1}{x^2} \end{cases}$$

А) $\tilde{y}_i = a_0 + a_1 \cdot \frac{1}{x_i}$;

Б) $\tilde{y}_i = a_0 + a_1 \cdot x_i + a_2 \cdot x_i^2$;

В) $\tilde{y}_i = a_0 + a_1 \cdot x_i$;

Г) $\tilde{y}_i = a_0 x^{a_1}$.

2 Уравнение $\hat{y}_i = a + b \cdot x + c \cdot x^2$ называется:

А) линейным;

Б) параболическим;

В) гиперболическим;

Г) экспоненциальным.

3 Уравнение $\hat{y} = a + \frac{b}{x}$ называется:

А) линейным трендом;

Б) параболическим трендом;

В) гиперболическим трендом;

Г) экспоненциальным трендом.

4 Коэффициент эластичности определяется по формуле

$\varepsilon = \frac{(b + 2 \cdot c \cdot x) \cdot x}{a + b \cdot x + c \cdot x^2}$ для модели регрессии в форме:

А) линейной функции;

Б) параболы;

В) гиперболы;

Г) показательной кривой;

Д) степенной.

5 В уравнении регрессии в форме гиперболы $\hat{y} = a + \frac{b}{x}$ если величина

$b > 0$, то:

А) при увеличении факторного признака x значения результативного признака y замедленно уменьшаются, и при $x \rightarrow \infty$ средняя величина y будет равна a ;

Б) то значение результативного признака y возрастает с замедленным ростом при увеличении факторного признака x , и при $x \rightarrow \infty$ $\bar{y} = a$;

В) при уменьшении факторного признака x значения результативного признака y замедленно увеличиваются, и при $x \rightarrow 0$ средняя величина y будет равна a ;

Г) то значение результативного признака y уменьшается с ускорением при увеличении факторного признака x , и при $x \rightarrow 0$ $\bar{y} = a$.

6 К какому классу нелинейных регрессий относится функция вида $\hat{y} = a + bx + cx^2$:

А) регрессии, нелинейные относительно включенных в анализ переменных, но линейных по оцениваемым параметрам;

Б) нелинейные регрессии по оцениваемым параметрам;

В) кривые насыщения;

Г) S-образные кривые.

7 К какому классу нелинейных регрессий относится функция вида $\hat{y} = a + \frac{b}{x}$:

А) регрессии, нелинейные относительно включенных в анализ переменных, но линейных по оцениваемым параметрам;

Б) нелинейные регрессии по оцениваемым параметрам;

В) кривые насыщения;

Г) S-образные кривые.

8 К какому классу нелинейных регрессий относится парабола:

А) регрессии, нелинейные относительно включенных в анализ переменных, но линейных по оцениваемым параметрам;

Б) нелинейные регрессии по оцениваемым параметрам;

В) кривые насыщения;

Г) S-образные кривые.

9 Коэффициент эластичности показывает:

А) на сколько процентов изменяется функция с изменением аргумента на одну единицу своего измерения;

Б) на сколько процентов изменяется функция с изменением аргумента на 1%;

В) на сколько единиц своего измерения изменяется функция с изменением аргумента на 1%;

Г) на сколько единиц своего измерения изменяется функция с изменением аргумента на единицу своего измерения.

10 Величина индекса корреляции, равная 1,587, свидетельствует:

А) о слабой их зависимости;

Б) о связи средней силы;

В) о сильной взаимосвязи;

Г) об ошибках в вычислениях.

7 Лабораторная работа № 7 Нелинейная регрессия и способы линеаризации моделей нелинейных по параметрам

Цель выполнения лабораторной работы: научить строить статистически значимые модели некоторых видов нелинейной регрессии и давать содержательный анализ полученным результатам.

Задания:

По исходным данным лабораторной работы № 2 для эндогенной переменной и экзогенными переменными, не вошедшими в линейное уравнение множественной регрессии выполните следующие задания:

- 1 Рассчитайте параметры уравнения степенной регрессии вида:

$$\tilde{y}_x = b_0 \cdot x_1^{b_1} \cdot x_2^{b_2} \cdot \dots \cdot x_p^{b_p} .$$

- 2 Рассчитайте параметры уравнения показательной регрессии вида:

$$\tilde{y}_x = b_0 \cdot b_1^{x_1} \cdot b_2^{x_2} \cdot \dots \cdot b_p^{x_p} .$$

- 3 Для каждого уравнения оцените тесноту связи между эндогенной переменной и экзогенными переменными, вошедшими в уравнения регрессии.

- 4 Рассчитайте средний коэффициент эластичности для каждой модели.

- 5 Оцените качество уравнений через среднюю ошибку аппроксимации.

- 6 Оцените с помощью F-критерия Фишера-Снедекора значимость уравнений.

7 Проверьте регрессионные остатки на гетероскедастичность с помощью любых тестов, представленных в лабораторной работе № 4. Если будет обнаружена гетероскедастичность остатков, примените для исходных данных ОМНК, предполагая, что $\sigma^2(u_i) = \sigma^2 x_i^2$.

8 Проверить остатки на наличие автокорреляции первого порядка, используя тесты из лабораторной работы № 5. Если гипотеза об отсутствии автокорреляции первого порядка не будет отвергнута, то применить ОМНК для оценивания параметров уравнения регрессии.

9 По значениям характеристик, рассчитанных в 3 – 8 заданиях выберете лучшее уравнение регрессии и оцените прогнозное значение эндогенной переменной, если прогнозные значения факторов составляют 80 % от их максимальных значений

Реализация типовых заданий

1 Рассчитайте параметры уравнения степенной регрессии вида:

$$\tilde{y}_x = b_0 \cdot x_1^{b_1} \cdot x_2^{b_2} \cdot \dots \cdot x_p^{b_p}.$$

В линейное уравнение регрессии, полученное в лабораторной работе № 2, не вошли три экзогенные переменные x_1 , x_4 и x_5 . Поэтому уравнение степенной модели, для рассматриваемых данных, примет вид:

$$\tilde{y}_i = b_0 \cdot x_1^{b_1} \cdot x_4^{b_4} \cdot x_5^{b_5}.$$

Данный вид регрессии относится ко второму классу нелинейных регрессий «нелинейные как относительно включенных в анализ объясняющих переменных, так и по оцениваемым параметрам».

Для оценивания параметров необходимо провести процедуру линеаризации переменных путем логарифмирования обеих частей уравнения:

$$\ln y = \ln b_0 + b_1 \ln x_1 + b_4 \ln x_4 + b_5 \ln x_5.$$

Введем новые переменные:

$$Y = \ln y, X1 = \ln x_1, X4 = \ln x_4, X5 = \ln x_5, C = \ln b_0.$$

Тогда уравнение примет вид множественной линейной регрессии:

$$Y = C + b_1 X1 + b_4 X4 + b_5 X5.$$

Для нахождения параметров данного уравнения воспользуемся инструментом анализа данных **Регрессия**. Для расчетов параметров используем данные рисунка 7.1.

	A	B	C	D	E	F	G	H	I	J
1										
2	Год	Квартал	y	x 1	x 4	x5	Y	X1	X4	X5
3		1	0,4	1,1	1,5	2,5	-0,91629	0,09531	0,405465	0,907839
4		2	0,5	1,5	1,0	2,4	-0,69315	0,405465	0	0,875469
5		3	0,6	2,2	1,2	4,2	-0,51083	0,788457	0,182322	1,435085
6	1	4	0,5	1,7	1,5	2,3	-0,69315	0,530628	0,412845	0,832909
7		1	0,3	0,9	1,6	2,0	-1,20397	-0,10536	0,470004	0,707168
8		2	0,4	1,1	1,8	1,9	-0,91629	0,09531	0,587787	0,641854
9		3	0,5	1,3	2,2	1,9	-0,69315	0,262364	0,770108	0,641854
10	2	4	0,4	1,1	2,7	2,1	-0,91629	0,09531	1,000632	0,741937
11		1	0,5	1,6	0,5	3,6	-0,69315	0,470004	-0,79851	1,282533
12		2	0,6	2,1	0,3	3,5	-0,51083	0,741937	-1,20397	1,252763
13		3	0,7	2	0,4	4,5	-0,35667	0,693147	-1,02165	1,505676
14	3	4	0,5	2	0,5	4,5	-0,69315	0,693147	-0,79113	1,504077
15		1	0,7	2,3	0,4	5,2	-0,35667	0,832909	-0,91629	1,645438
16		2	0,8	2,6	1,9	5,3	-0,22314	0,955511	0,641854	1,667707
17		3	0,9	3	2,2	6,8	-0,10536	1,098612	0,788457	1,911141
18	4	4	0,7	2,2	3,2	5,0	-0,35667	0,788457	1,163151	1,600986
19		1	0,8	3	0,8	6,8	-0,22314	1,098612	-0,22314	1,916923
20		2	0,9	3,5	0,5	7,1	-0,10536	1,252763	-0,69315	1,960095
21		3	0,7	3,5	0,6	7,1	-0,35667	1,252763	-0,51083	1,960095
22	5	4	0,8	3	0,8	6,8	-0,22314	1,098612	-0,22314	1,916923

Рисунок 7.1 – Исходные данные для степенной модели

Результаты регрессионного анализа представлены на рисунке 7.2.

	A	B	C	D	E	F	G	H	I
1	Вывод итогов								
2									
3	Регрессионная статистика								
4	Множественный R	0,949591339							
5	R-квадрат	0,901723711							
6	Нормированный R-квадрат	0,883296906							
7	Стандартная ошибка	0,105087311							
8	Наблюдения	20							
9									
10	Дисперсионный анализ								
11		<i>df</i>	<i>SS</i>	<i>MS</i>	<i>F</i>	<i>Значимость F</i>			
12	Регрессия	3	1,621232414	0,540410805	48,93543657	2,77512E-08			
13	Остаток	16	0,176693486	0,011043343					
14	Итого	19	1,7979259						
15									
16		<i>Коэффициенты</i>	<i>Стандартная ошибка</i>	<i>t-статистика</i>	<i>P-Значение</i>	<i>Нижние 95%</i>	<i>Верхние 95%</i>	<i>Нижние 95,0%</i>	<i>Верхние 95,0%</i>
17	Y-пересечение	-1,034095965	0,108085569	-9,567382388	5,06427E-08	-1,263227134	-0,804964795	-1,263227134	-0,804964795
18	X1	0,70386911	0,181358593	3,881090492	0,001325546	0,319406068	1,088332151	0,319406068	1,088332151
19	X4	0,044657045	0,035529941	1,256884856	0,226834948	-0,030663066	0,119977156	-0,030663066	0,119977156
20	X5	0,025322318	0,155189355	0,163170456	0,8724274	-0,303664418	0,354309054	-0,303664418	0,354309054

Рисунок 7.2 – Результат применения инструмента Регрессия

Полученное уравнение множественной линейной регрессии будет иметь следующий вид:

$$Y = -1,034 + 0,704X_1 + 0,045X_4 + 0,025X_5 .$$

Оценивая параметры данного уравнения, замечаем, что статистически значимым является параметр при X1, следовательно, целесообразно строить уравнение степенной регрессии только с данным фактором. В результате получаем уравнение следующего вида: $Y = -0,999 + 0,702X_1$ (рисунок 7.3).

	A	B	C	D	E	F	G	H	I
1	Вывод итогов								
2									
3	Регрессионная статистика								
4	Множественный R	0,944465784							
5	R-квадрат	0,892015616							
6	Нормированный R-квадрат	0,886016484							
7	Стандартная ошибка	0,103855648							
8	Наблюдения	20							
9									
10	Дисперсионный анализ								
11		<i>df</i>	<i>SS</i>	<i>MS</i>	<i>F</i>	<i>Значимость F</i>			
12	Регрессия	1	1,60377798	1,60377798	148,6907695	3,89712E-10			
13	Остаток	18	0,19414792	0,010785996					
14	Итого	19	1,7979259						
15									
16		<i>Коэффициенты</i>	<i>Стандартная ошибка</i>	<i>t-статистика</i>	<i>P-Значение</i>	<i>Нижние 95%</i>	<i>Верхние 95%</i>	<i>Нижние 95,0%</i>	<i>Верхние 95,0%</i>
17	Y-пересечение	-0,998522706	0,044380476	-22,49914346	1,25022E-14	-1,091762627	-0,905282785	-1,091762627	-0,905282785
18	X1	0,70171925	0,057546827	12,19388246	3,89712E-10	0,580817852	0,822620648	0,580817852	0,822620648

Рисунок 7.3 – Результат применения инструмента Регрессия

Потенцируя параметр $C = \ln b_0$ уравнения, получим $b_0=0,368$ (используем **Мастер функций / Категория «Математические» / EXP(ln b₀)**). Следовательно, уравнение примет вид: $\tilde{y}_i = 0,368 \cdot x_i^{0,702}$.

2 Рассчитайте параметры уравнения показательной регрессии вида:

$$\tilde{y}_x = b_0 \cdot b_1^{x_1} \cdot b_2^{x_2} \cdot \dots \cdot b_p^{x_p}.$$

Данный вид регрессии также относится ко второму классу нелинейных регрессий «нелинейные как относительно включенных в анализ объясняющих переменных, так и по оцениваемым параметрам».

Построению уравнения показательной кривой $\tilde{y}_i = b_0 \cdot b_1^{x_1} \cdot b_4^{x_4} \cdot b_5^{x_5}$ предшествует процедура линеаризации переменных при логарифмировании обеих частей уравнения:

$$\ln y = \ln b_0 + x_1 \cdot \ln b_1 + x_4 \cdot \ln b_4 + x_5 \cdot \ln b_5; Y = C + B_1 \cdot x_1 + B_4 \cdot x_4 + B_5 \cdot x_5,$$

где $Y = \ln y$, $C = \ln b_0$, $B_1 = \ln b_1$, $B_4 = \ln b_4$, $B_5 = \ln b_5$.

Значения параметров уравнения регрессии определим аналогично степенной модели. Для их расчета используем данные таблицы (рисунок 7.4):

	A	B	C	D	E	F	G
1							
2	Год	Квартал	y	x 1	x 4	x5	Y
3		1	0,4	1,1	1,5	2,5	-0,91629
4		2	0,5	1,5	1,0	2,4	-0,69315
5		3	0,6	2,2	1,2	4,2	-0,51083
6	1	4	0,5	1,7	1,5	2,3	-0,69315
7		1	0,3	0,9	1,6	2,0	-1,20397
8		2	0,4	1,1	1,8	1,9	-0,91629
9		3	0,5	1,3	2,2	1,9	-0,69315
10	2	4	0,4	1,1	2,7	2,1	-0,91629
11		1	0,5	1,6	0,5	3,6	-0,69315
12		2	0,6	2,1	0,3	3,5	-0,51083
13		3	0,7	2	0,4	4,5	-0,35667
14	3	4	0,5	2	0,5	4,5	-0,69315
15		1	0,7	2,3	0,4	5,2	-0,35667
16		2	0,8	2,6	1,9	5,3	-0,22314
17		3	0,9	3	2,2	6,8	-0,10536
18	4	4	0,7	2,2	3,2	5,0	-0,35667
19		1	0,8	3	0,8	6,8	-0,22314
20		2	0,9	3,5	0,5	7,1	-0,10536
21		3	0,7	3,5	0,6	7,1	-0,35667
22	5	4	0,8	3	0,8	6,8	-0,22314

Рисунок 7.4 – Исходные данные для построения показательной модели

Результаты регрессионного анализа представлены на рисунке 7.5.

	A	B	C	D	E	F	G	H	I
1	Вывод итогов								
2									
3	<i>Регрессионная статистика</i>								
4	Множественный R	0,914974							
5	R-квадрат	0,837178							
6	Нормированный R-квадрат	0,806649							
7	Стандартная ошибка	0,135264							
8	Наблюдения	20							
9									
10	<i>Дисперсионный анализ</i>								
11		<i>df</i>	<i>SS</i>	<i>MS</i>	<i>F</i>	<i>Значимость F</i>			
12	Регрессия	3	1,505183716	0,501727905	27,42224015	1,52506E-06			
13	Остаток	16	0,292742184	0,018296387					
14	Итого	19	1,7979259						
15									
16		<i>Коэффициент</i>	<i>Стандартная ошибка</i>	<i>t-статистика</i>	<i>P-Значение</i>	<i>Нижние 95%</i>	<i>Верхние 95%</i>	<i>Нижние 95,0%</i>	<i>Верхние 95,0%</i>
17	Y-пересечение	-1,31405	0,11956876	-10,98992677	7,27418E-09	-1,567526371	-1,060577473	-1,567526371	-1,060577473
18	x 1	0,311866	0,146811166	2,124266577	0,049583802	0,000640284	0,623091822	0,000640284	0,623091822
19	x 4	0,033689	0,038614264	0,872461597	0,39586119	-0,04816912	0,115548044	-0,04816912	0,115548044
20	x5	0,019569	0,061607767	0,317632369	0,754871473	-0,111034011	0,150171253	-0,111034011	0,150171253

Рисунок 7.5 – Результат применения инструмента Регрессия для показательной регрессии

Получено уравнение множественной линейной регрессии:

$$Y = -1,314 + 0,312 \cdot x_1 + 0,034 \cdot x_4 + 0,02 \cdot x_5.$$

Оценивая параметры данного уравнения, замечаем, что статистически значимым является параметр при X_1 , следовательно, целесообразно строить уравнение показательной регрессии только с данным фактором. В результате получаем уравнение следующего вида: $Y = 1,256 + 0,345x_1$ (рисунок 7.6).

	A	B	C	D	E	F	G	H	I
1	Вывод итогов								
2									
3	<i>Регрессионная статистика</i>								
4	Множественный R	0,910051685							
5	R-квадрат	0,82819407							
6	Нормированный R-квадрат	0,818649296							
7	Стандартная ошибка	0,130999222							
8	Наблюдения	20							
9									
10	<i>Дисперсионный анализ</i>								
11		<i>df</i>	<i>SS</i>	<i>MS</i>	<i>F</i>	<i>Значимость F</i>			
12	Регрессия	1	1,489031568	1,489031568	86,76937532	2,63116E-08			
13	Остаток	18	0,308894332	0,017160796					
14	Итого	19	1,7979259						
15									
16		<i>Коэффициенты</i>	<i>Стандартная ошибка</i>	<i>t-статистика</i>	<i>P-Значение</i>	<i>Нижние 95%</i>	<i>Верхние 95%</i>	<i>Нижние 95,0%</i>	<i>Верхние 95,0%</i>
17	Y-пересечение	-1,256240926	0,08254718	-15,21845958	1,01177E-11	-1,429666116	-1,082815735	-1,429666116	-1,082815735
18	x 1	0,344789806	0,03701444	9,315008069	2,63116E-08	0,267025353	0,422554258	0,267025353	0,422554258

Рисунок 7.6 – Результат применения инструмента Регрессия для показательной регрессии

Произведем потенцирование полученного уравнения и запишем его в обычном виде: $\tilde{y}_i = 0,285 \cdot 1,417^{x_i}$.

3 Для каждого уравнения оцените тесноту связи между эндогенной переменной и экзогенными переменными, вошедшими в уравнения регрессии.

Оценка тесноты связи между экзогенной и эндогенной переменной, рассчитывается аналогично лабораторной работы № 6, задание 2.

Подставляя в данные уравнения фактические значения x_1 , получаем теоретические значения эндогенной переменной \tilde{y}_i , и по ним рассчитаем

показатель тесноты связи – **индекс корреляции** $\rho_{xy} = \sqrt{1 - \frac{\sum (y - \tilde{y}_i)^2}{\sum (y - \bar{y})^2}}$.

4 Рассчитайте средний коэффициент эластичности для каждой модели.

Средние коэффициенты эластичности показывают, на сколько процентов от значения своей средней (\bar{y}) изменяется эндогенная переменная при изменении экзогенной переменной (x_i) на 1 % от своей средней (\bar{x}_i) при фиксированном воздействии на y всех прочих факторов, включенных в уравнение регрессии. Формула для расчета коэффициента эластичности имеет вид:

$$\Theta = f'(x) \cdot \frac{x}{y},$$

где $f'(x)$ – первая производная, характеризующая соотношение приростов эндогенной и экзогенной переменных для соответствующей формы связи.

Коэффициент эластичности для оцененного уравнения степенной регрессии составит:

$$\bar{\Theta} = b = 0,702.$$

Он показывает, что при увеличении на 1 % от своей средней объема реализованной продукции (x_1), интегральный показатель экономической безопасности организации (y) увеличивается на 0,702 % от своей средней.

Коэффициент эластичности для оцененного уравнения показательной регрессии составит:

$$\bar{\Theta} = \bar{x} \cdot \ln b = 2,085 \cdot 0,3448 = 0,72\%.$$

Аналогично делаем вывод по коэффициенту эластичности.

5 Оцените качество уравнений через среднюю ошибку аппроксимации.

Ошибка аппроксимации оценивается аналогично лабораторной работе № 6, задание 4.

6 Оцените с помощью *F*-критерия Фишера-Снедекора значимость уравнений.

Оценка значимости уравнение проводится аналогично согласно лабораторной работы №6, задание 5.

7 Проверьте регрессионные остатки на гетероскедастичность с помощью любых тестов, представленных в лабораторной работе № 4. Если будет обнаружена гетероскедастичность остатков, примените для исходных данных ОМНК, предполагая, что $\sigma^2(u_i) = \sigma^2 x_i^2$.

Задание выполняется по лабораторной работе №6, задание 6.

8 Проверить остатки на наличие автокорреляции первого порядка, используя тесты из лабораторной работы № 5. Если гипотеза об отсутствии автокорреляции первого порядка не будет отвергнута, то применить ОМНК для оценивания параметров уравнения регрессии.

Задание выполняется по лабораторной работе №6, задание 7.

Таблица 7.1 – Основные характеристики качества регрессионных моделей

Название модели	Оценка уравнения регрессии	ρ_{xy}^2	\bar{A} , %	Результаты проверки гипотезы о		
				значимости уравнения регрессии	гомоскедастичности регрессионных остатков	автокорреляции регрессионных остатков
Гиперболическая	$\tilde{y}_i = 0,985 - \frac{0,665}{x_i}$	0,42	60,17	статистически значимо	гомоскедастичны	не автокоррелированы
Степенная	$\tilde{y}_i = 0,368 \cdot x_i^{0,702}$	0,86	7,5	статистически значимо	гомоскедастичны	не автокоррелированы
Показательная	$\tilde{y}_i = 0,285 \cdot 1,417^{x_i}$	0,78	9,26	статистически значимо	гомоскедастичны	не автокоррелированы

9 Выберите из построенных уравнений регрессий, по рассчитанным значениям характеристик, лучшее уравнение. Оцените прогнозное значение эндогенной переменной, самостоятельно спрогнозировав прогнозные значения факторов.

Результаты оценки полученных нелинейных уравнений регрессии на значимость, точность и адекватность представлены в таблице 7.1.

Как видно по данным таблицам 7.1 лучшими характеристиками обладает уравнение степенной регрессии. Рассчитаем по нему интервальный прогноз для эндогенной переменной, если прогнозное значение экзогенной переменной составит $X_1^{прогн} = 2,506$.

При нелинейных соотношениях рассматриваемых признаков, приводимых к линейному виду, возможно интервальное оценивание прогнозируемых значений эндогенной переменной.

Для показательной кривой

$$\tilde{y}_i = b_0 \cdot b_1^{x_1} \cdot b_2^{x_2} \cdot \dots \cdot b_p^{x_p}$$

сначала находится точечный прогноз по новому преобразованному уравнению

$$Y = C + B_1 \cdot x_1 + B_2 \cdot x_2 + \dots + B_p \cdot x_p,$$

затем строится доверительный интервал

$$\tilde{Y}_{прог} - t_\alpha m_{\tilde{Y}_x} \leq Y^* \leq \tilde{Y}_{прог} + t_\alpha m_{\tilde{Y}_x},$$

где $m_{\tilde{Y}_x}$ – средняя ошибка прогнозируемого значения: $m_{\tilde{Y}_x} = S \sqrt{X_0^T (X^T X)^{-1} X_0}$;

X_0 – вектор-столбец прогнозных значений факторов;

$$S - \text{стандартная ошибка: } S = \sqrt{\frac{(1 - R^2) \sigma_Y^2 n}{n - p - 1}}.$$

В случае простой (парной) зависимости средняя ошибка прогнозируемого значения рассчитывается по формуле:

$$m_{\tilde{Y}_x} = \sqrt{\frac{\sum (Y_i - \tilde{Y}_i)^2}{n - 2}} \cdot \sqrt{1 + \frac{1}{n} + \frac{(x^{\text{прогн}} - \bar{x})^2}{\sum (x_i - \bar{x})^2}}.$$

Далее с помощью обратного преобразования определяются доверительные интервалы для эндогенной переменной в исходном соотношении (используем **Мастер функций / Категория «Математические» / EXP(Y)**).

В **степенной функции** $\tilde{y}_i = b_0 \cdot x_1^{b_1} \cdot x_2^{b_2} \cdot \dots \cdot x_p^{b_p}$ доверительный интервал для прогнозного значения эндогенной переменной строиться так же, как в показательной функции. Отличие состоит лишь в том, что при расчете средней ошибки прогнозируемого значения эндогенной переменной ($m_{\tilde{Y}_x}$) используются не исходные данные экзогенных переменных, а их логарифмы.

Подставим прогнозное значение фактора X_1 в уравнение $Y = -0,999 + 0,702X_1$. В результате получим:

$$\tilde{Y}_{\text{прог}} = -0,999 + 0,702 \cdot 2,506 = 0,8.$$

Вычислим среднюю ошибку прогнозируемого значения

$$m_{\tilde{Y}_x} = \sqrt{\frac{0,19415}{20 - 2}} \cdot \sqrt{1 + \frac{1}{20} + \frac{(2,506 - 0,657)^2}{3,257}} = 0,132,$$

по таблицам распределения Стьюдента находим табличное значение t_α при уровне значимости 0,05 и числе степеней свободы 18: $t_{крит}(0,05;18) = 1,73$.

Составляем доверительный интервал:

$$0,8 - 1,73 \cdot 0,132 \leq Y_{прог} \leq 0,8 + 1,73 \cdot 0,132$$
$$0,6 \leq Y_{прог} \leq 1$$

Тогда доверительный интервал для эндогенной переменной составит:

$$1,82 \leq y_{прог} \leq 2,72$$

Значит, с вероятностью 95 % можно сказать, что прогнозируется кризисное состояние предприятия.

Для **гиперболической кривой** $\tilde{y}_i = b_0 + \frac{b_1}{x_1} + \frac{b_2}{x_2} + \dots + \frac{b_p}{x_p}$ доверительный интервал прогнозного значения эндогенной переменной строиться так же, как в линейной функции, заменяя при расчетах значения экзогенных переменных x_i на $1/x_i$.

Контрольные вопросы к лабораторной работе № 7

1 Запишите все виды моделей, нелинейных относительно оцениваемых параметров.

2 Как применяется МНК к моделям, нелинейным относительно оцениваемых параметров?

3 Каков смысл коэффициентов регрессии в логарифмических регрессионных моделях?

4 Как определить коэффициент эластичности для моделей, нелинейных относительно оцениваемых параметров?

5 Назовите показатели корреляции, используемых для нелинейных моделей относительно оцениваемых параметров.

Тесты для самоконтроля к лабораторной работе № 7

1 Отметьте правильную форму показательной функции:

А) $\hat{y} = a + \frac{b}{x}$

Б) $\hat{y} = a \cdot b^x$

В) $\hat{y} = a \cdot x^b$

Г) $\hat{y} = a + bx$

2 Величина индекса корреляции, равная -1,00, свидетельствует:

А) о слабой их зависимости;

Б) о связи средней силы;

В) о сильной взаимосвязи;

Г) об ошибках в вычислениях.

3 К какому классу нелинейных регрессий относится показательная кривая:

А) регрессии, нелинейные относительно включенных в анализ переменных, но линейных по оцениваемым параметрам;

Б) нелинейные регрессии по оцениваемым параметрам;

В) кривые насыщения;

Г) S-образные кривые.

4 К какому классу нелинейных регрессий относится функция вида $\hat{y} = e^{a+bx}$:

А) регрессии, нелинейные относительно включенных в анализ переменных, но линейных по оцениваемым параметрам;

Б) нелинейные регрессии по оцениваемым параметрам;

В) кривые насыщения;

Г) S-образные кривые.

5 Коэффициент эластичности определяется по формуле $\varepsilon = b$ для модели регрессии в форме:

- А) линейной функции;
- Б) параболы;
- В) гиперболы;
- Г) показательной кривой;
- Д) степенной.

6 Уравнение $\hat{y}_i = a \cdot k^{t_i}$ называется:

- А) линейным;
- Б) параболическим;
- В) гиперболическим;
- Г) экспоненциальным.

7 Какой из тестов используют для выбора формы модели:

- А) тест Рамсея;
- Б) тест Дарбина-Уотсона;
- В) тест Бреуша-Годфри;
- Г) тест Бокса-Кокса.

8 Экономический смысл параметра a_2 уравнения регрессии $y = a_0 \cdot K^{a_1} \cdot L^{a_2}$ в модели производственной функции:

- А) эластичность выпуска по труду;
- Б) эластичность выпуска по капиталу;
- В) средняя производительность труда;
- Г) предельная производительность труда.

9 Построено уравнение регрессии: $y = a_0 \cdot K^{a_1} \cdot L^{a_2}$. Для проверки значимости уравнения используется распределение:

- А) Нормальное;
- Б) Стьюдента;
- В) Пирсона;
- Г) Фишера-Снедекора.

10 Коэффициент эластичности показывает:

А) на сколько процентов изменяется функция с изменением аргумента на одну единицу своего измерения;

Б) на сколько процентов изменяется функция с изменением аргумента на 1%;

В) на сколько единиц своего измерения изменяется функция с изменением аргумента на 1%;

Г) на сколько единиц изменяется функция при изменении аргумента на одну единицу.

8 Лабораторная работа № 8 Моделирование временных рядов: исследование тенденции

Цель выполнения лабораторной работы: научить тестировать временные ряды на стационарность; строить статистически значимые модели тренда; проводить экстраполяцию тенденции динамики и давать содержательный анализ полученным результатам.

Задания:

По исходным данным лабораторной работы № 1 для эндогенной переменной Y – интегральный показатель экономической безопасности организации выполните следующие задания:

1 Проверить утверждение об отсутствии тенденции во временном ряду, используя критерии серий, основанный на медиане выборке.

2 Оцените параметры следующих трендов:

- а) линейного;
- б) полиномиального второй степени;
- в) экспоненциального;
- г) степенного;
- д) логарифмического.

3 Оцените точность моделей трендов с помощью ошибки аппроксимации.

4 По модели имеющей самую высокую точность рассчитайте точечный и интервальный прогноз на следующий год в поквартальном разрезе.

Реализация типовых заданий

1 Проверти утверждение об отсутствии тенденции во временном ряду, используя критерий серий, основанный на медиане выборке

Критерий серий, основанный на медиане выборки, включает следующие шаги:

1 Из исходного временного ряда интегрального показателя экономической безопасности организации (y_t), образуем ранжированный ряд (y'_t) (столбец 4, таблица 8.1).

Таблица 8.1 – Формирование серий

Год	Квартал	y_t	y'_t	δ_i	Номер серии
1	2	3	4	5	6
1	1	0,4	0,3	-	1
	2	0,5	0,4	-	
	3	0,6	0,4		
	4	0,5	0,4	-	
2	1	0,3	0,5	-	
	2	0,4	0,5	-	
	3	0,5	0,5	-	
	4	0,4	0,5	-	
3	1	0,5	0,5	-	
	2	0,6	0,6		
	3	0,7	0,6	+	2
	4	0,5	0,7	-	3
4	1	0,7	0,7	+	4
	2	0,8	0,7	+	
	3	0,9	0,7	+	
	4	0,7	0,8	+	
5	1	0,8	0,8	+	
	2	0,9	0,8	+	
	3	0,7	0,9	+	
	4	0,8	0,9	+	

2 Определяем медиану (Me) ранжированного временного ряда. Так как значение длины ряда четное, то $Me = \frac{y'_m + y'_{m+1}}{2} = \frac{0,6 + 0,6}{2} = 0,6$

3 Образует последовательность δ_i из плюсов и минусов по правилу:

$$\delta_i = \begin{cases} +, \text{ если } y_t > M_e, t = 1, 2, \dots, n \\ -, \text{ если } y_t < M_e, t = 1, 2, \dots, n \end{cases}$$

Если значение уровня исходного ряда y_t равно медиане, то это значение пропускается (столбец 5 таблицы 1).

4 Подсчитывается $\nu(n)$ – число серий в совокупности δ_i , где под серией понимается последовательность подряд идущих плюсов и минусов. Один плюс или один минус тоже будут считаться серией. Определяется $\tau_{\max}(n)$ – протяженность самой длинной серии.

Получаем: $\nu(n)=4$ (столбец 6 таблицы 1), $\tau_{\max}(n)=8$ (столбец 5 таблицы 1).

5 Чтобы не была отвергнута гипотеза о случайности исходного ряда, должны выполняться следующие неравенства:

$$\nu(n) > \left[\frac{1}{2} (n + 1 - 1,96\sqrt{n-1}) \right]$$

$$\tau_{\max}(n) < [1,43 \ln(n+1)]$$

где n – длина временного ряда.

Рассчитаем правые части неравенств:

$$\frac{1}{2} (20 + 1 - 1,96\sqrt{20-1}) = 6,23.$$

$$1,43 \cdot \ln(20+1) = 4,35$$

Так как в правой части неравенства стоят квадратные скобки, означающие целую часть числа, то сравнения будем проводить с целыми числами – соответственно с 6 и 4.

Получим:

$$4 > 6 \text{ и } 8 < 4.$$

Оба неравенства нарушаются, следовательно, гипотеза отвергается, временной ряд интегрального показателя экономической безопасности организации содержит тенденцию развития.

2 Оцените параметры следующих трендов:

- а) линейного;
- б) полиномиального второй степени;
- в) экспоненциального;
- г) степенного;
- д) логарифмического.

Для оценки параметров трендов используем встроенную функцию MS Excel. Строим график со значениями эндогенной переменной Y , открываем диалоговое окно нажав правой клавишей «мыши» по полю графика (рисунок 1).

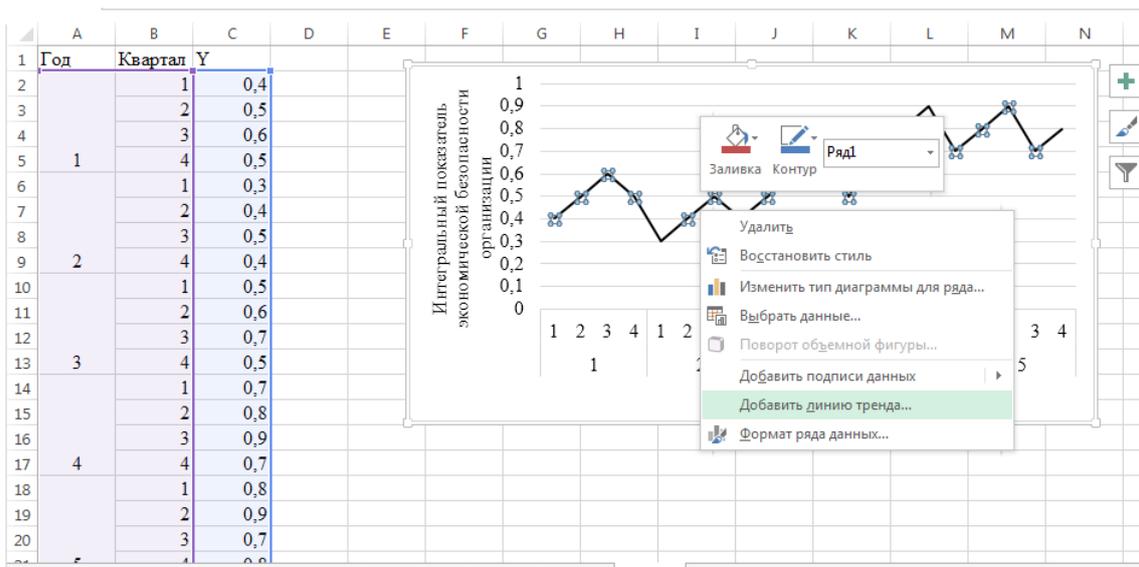


Рисунок 8.1 – Добавление линии тренда на график в MS Excel

Для построения линейного тренда в появившемся окне «Формат линии тренда» выбираем «Линейная», в разделе «Название аппроксимирующей (сглаженной) кривой» выбираем «Другое» и вводим название «Линейный тренд»; ниже отмечаем «Показывать уравнение на диаграмме» и «Пометитить на диаграмму величину достоверности аппроксимации» (рисунок 8.2).

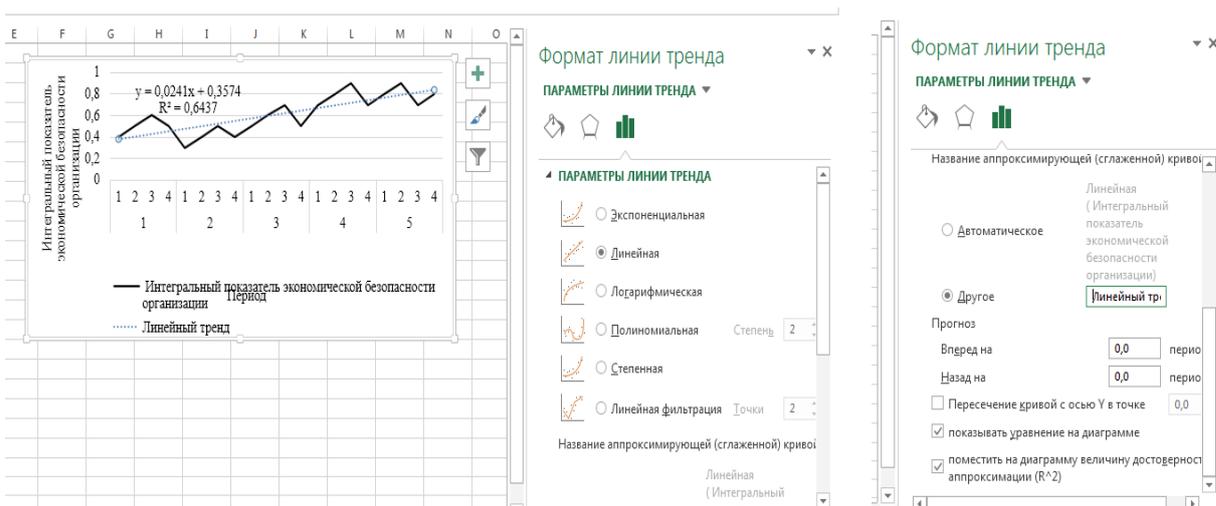


Рисунок 8.2 – Работа в диалоговом окне «Формат линии тренда» при построении линейного тренда

В результате должен получиться график, представленный на рисунке 8.3.

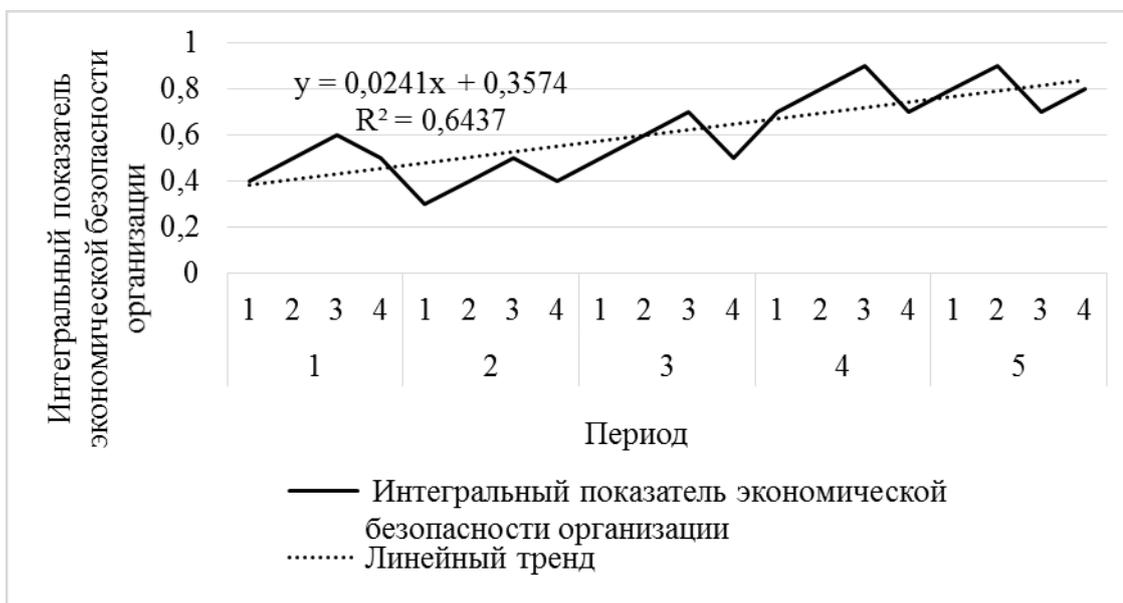


Рисунок 8.3 – Результаты аналитического выравнивания по линейному тренду

Аналогично строим полиномиальный тренд 2 степени (рисунок 8.4).

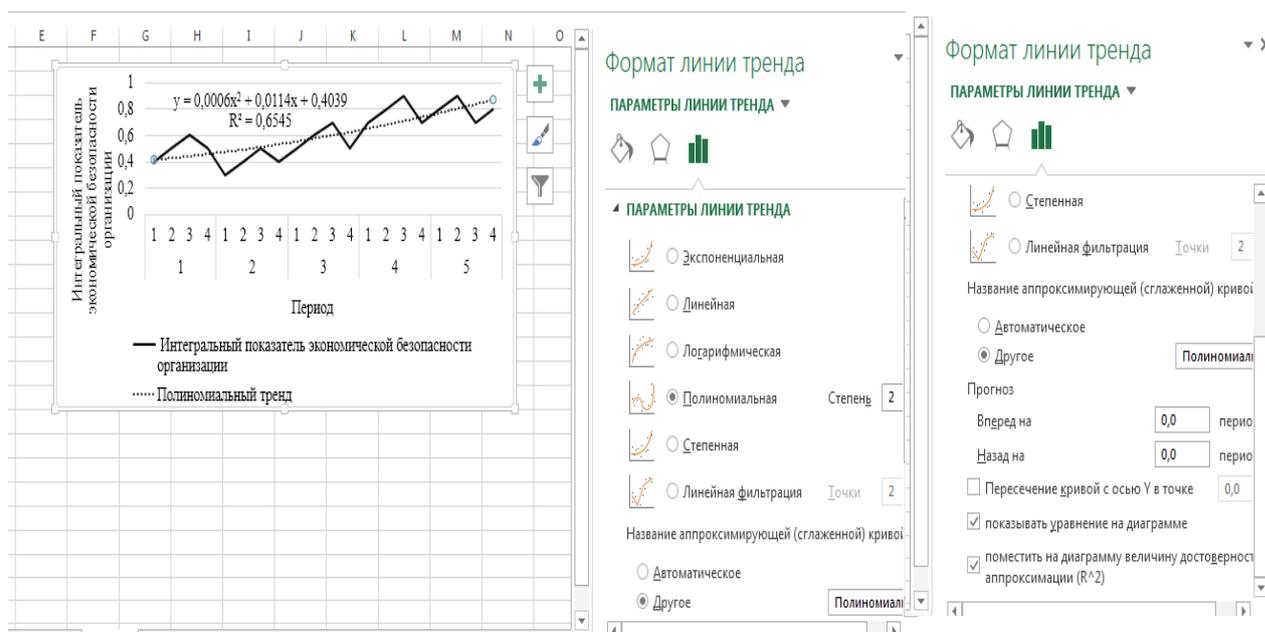


Рисунок 8.4 – Работа в диалоговом окне «Формат линии тренда» при построении полиномиального тренда

В результате должен получиться график, представленный на рисунке 8.5.

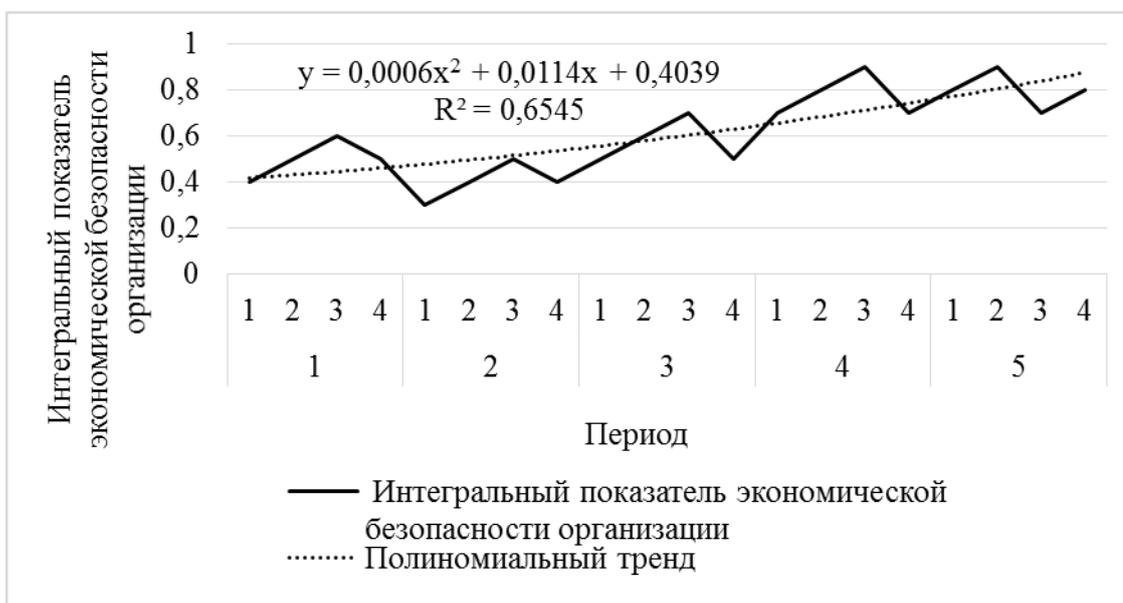


Рисунок 8.5 – Результаты аналитического выравнивания по полиномиальному тренду

На рисунках 8.6 – 8.8 представлены результаты аналитического выравнивания по экспоненциальному, степенному и логарифмическому трендам.

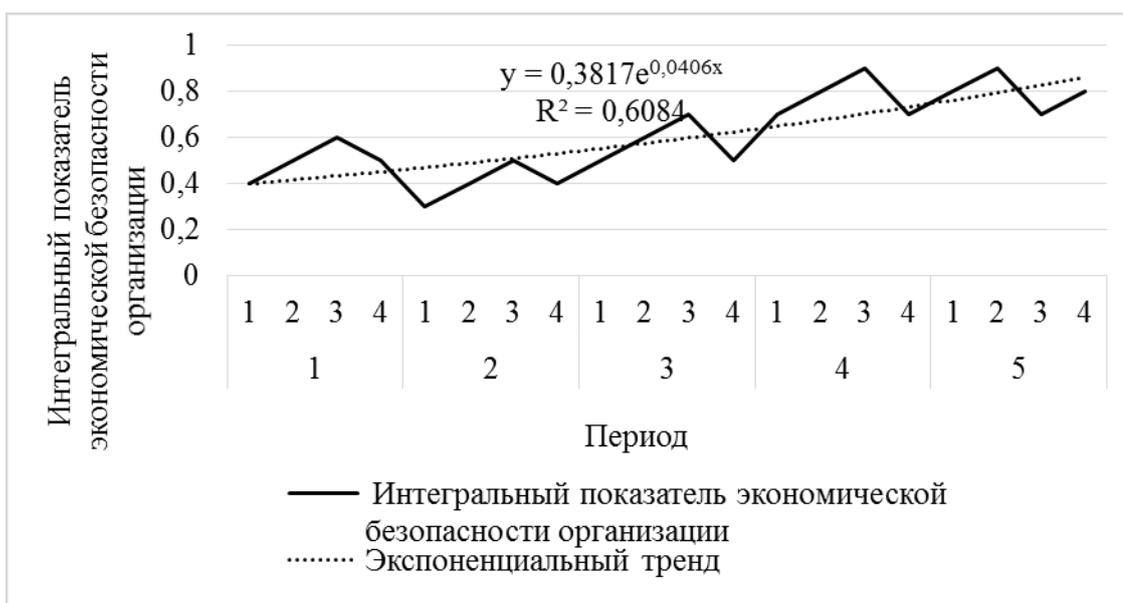


Рисунок 8.6 – Результаты аналитического выравнивания по экспоненциальному тренду

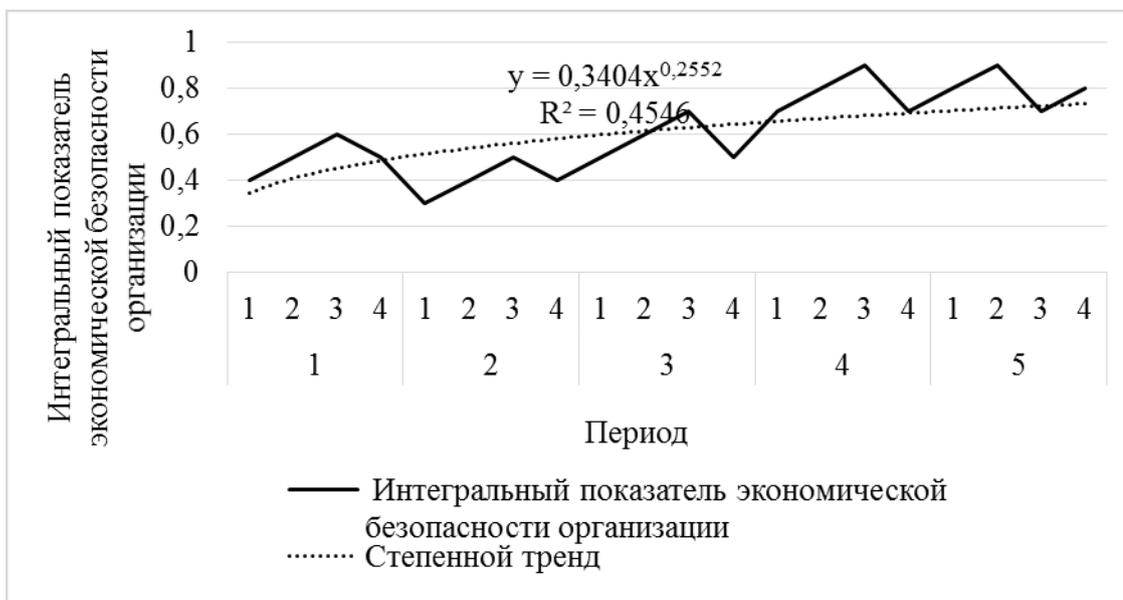


Рисунок 8.7 – Результаты аналитического выравнивания по степенному тренду

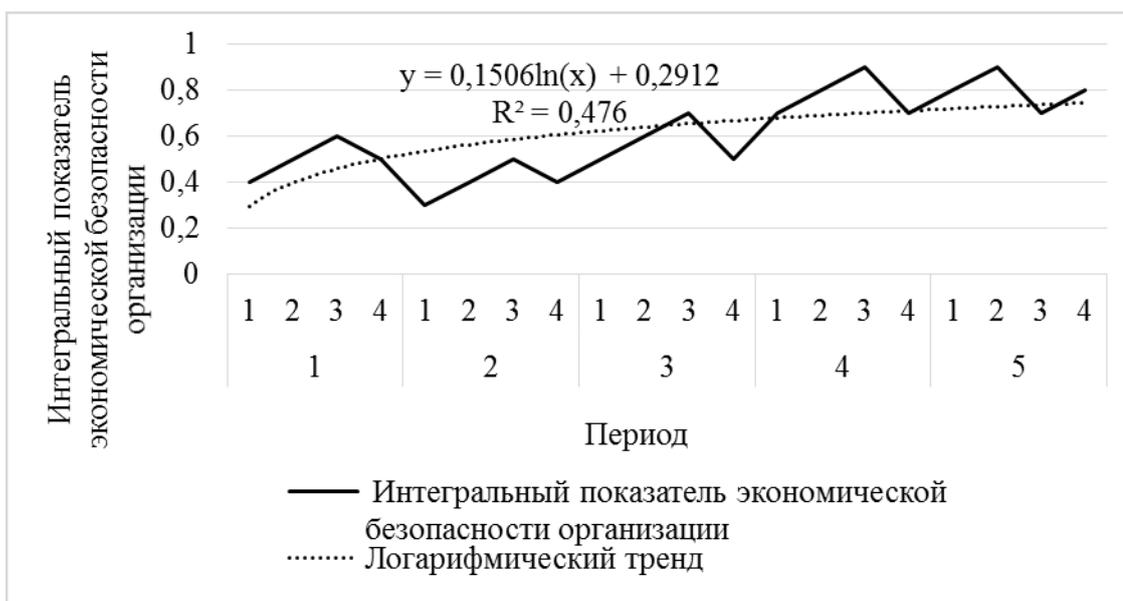


Рисунок 8.8 – Результаты аналитического выравнивания по логарифмическому тренду

Так как в уравнение трендов входит фактор времени t , то необходимо заменить X в уравнениях выведенных на графиках в MS Excel автоматически, на t ; в уравнение экспоненциального тренда заменим e на число 2,72.

Оценки уравнений трендов, соответствующие их теоретической форме, представлены в таблице 8.2.

Таблица 8.2 – Результаты аналитического выравнивания эндогенной переменной Y – интегральный показатель экономической безопасности организации

Вид тренда	Оценка уравнения тренда
линейный	$\tilde{y} = 0,36 + 0,02 \cdot t$
полиномиальный второй степени	$\tilde{y} = 0,40 + 0,01 \cdot t + 0,0006 \cdot t^2$
экспоненциальный	$\tilde{y} = 0,38 \cdot 2,72^{0,04t}$
степенной	$\tilde{y} = 0,34 \cdot t^{0,26}$
логарифмический	$\tilde{y} = 0,29 + 0,15 \cdot \ln(t)$

Примечание: в уравнение экспоненциального тренда число $e=2,72$.

Для расчета теоретических (выровненных) значений эндогенной переменной по полученным уравнениям трендов, необходимо подставить в уравнение значение t . При расчете параметров уравнений трендов в MS Excel фактор времени t нумеруется от начала ряда и принимает значения от 1 до n с шагом 1. Результаты расчета по всем уравнениям трендов представлены на рисунке 8.9.

	A	B	C	D	E	F	G	H	I
1					Теоретические значения по уравнению тренда				
2	Год	Квартал	Y	t	линейного	логарифми- ческого	полиномаль- ного	экспонен- циального	степенного
3		1	0,4	1	$= 0,02 * D3 + 0,36$		0,42	0,40	0,34
4		2	0,5	2	0,41	0,40	0,43	0,41	0,41
5		3	0,6	3	0,43	0,46	0,44	0,43	0,45
6	1	4	0,5	4	0,45	0,50	0,46	0,45	0,48
7		1	0,3	5	0,48	0,53	0,48	0,47	0,51
8		2	0,4	6	0,50	0,56	0,49	0,49	0,54
9		3	0,5	7	0,53	0,58	0,51	0,51	0,56
10	2	4	0,4	8	0,55	0,60	0,53	0,53	0,58
11		1	0,5	9	0,57	0,62	0,56	0,55	0,60
12		2	0,6	10	0,60	0,64	0,58	0,57	0,61
13		3	0,7	11	0,62	0,65	0,60	0,60	0,63
14	3	4	0,5	12	0,65	0,67	0,63	0,62	0,64
15		1	0,7	13	0,67	0,68	0,65	0,65	0,66
16		2	0,8	14	0,69	0,69	0,68	0,67	0,67
17		3	0,9	15	0,72	0,70	0,71	0,70	0,68
18	4	4	0,7	16	0,74	0,71	0,74	0,73	0,69
19		1	0,8	17	0,77	0,72	0,77	0,76	0,70
20		2	0,9	18	0,79	0,73	0,80	0,79	0,71
21		3	0,7	19	0,82	0,73	0,84	0,83	0,72
22	5	4	0,8	20	0,84	0,74	0,87	0,86	0,73
23	Итого		12,2		12,21	12,20	12,19	12,02	11,91

Рисунок 8.9 – Расчет теоретических (выравненных) значений эндогенной переменной по уравнениям трендов

3 Оцените точность моделей трендов с помощью ошибки аппроксимации.

Ошибка аппроксимации использовалась для оценки точности моделей в лабораторных работах 1 – 7.

На рисунке 8.10 приведены расчеты средней ошибки аппроксимации для оцененных уравнений трендов.

Год	Квартал	Y	t	Теоретические значения по уравнению тренда					Ошибка аппроксимации тренда			
				линейного	логарифмического	полиномиального	экспоненциального	степенного	линейного	логарифмического	полиномиального	экспоненциального
3	1	0,41	1	0,38	0,29	0,42	0,40	0,34	=ABS(C3-E3)/C3*100	3,97	0,62	14,90
4	2	0,5	2	0,41	0,40	0,43	0,41	0,41	ABS(число)	20,88	14,18	17,20
5	3	0,6	3	0,43	0,46	0,44	0,43	0,45	28,38	23,89	26,08	28,14
6	4	0,5	4	0,45	0,50	0,46	0,45	0,48	9,24	0,00	8,18	10,19
7	1	0,3	5	0,48	0,53	0,48	0,47	0,51	59,30	77,86	58,63	55,89
8	2	0,4	6	0,50	0,56	0,49	0,49	0,54	25,50	40,26	23,48	21,77
9	3	0,5	7	0,53	0,58	0,51	0,51	0,56	5,22	16,85	2,62	1,45
10	4	0,4	8	0,55	0,60	0,53	0,53	0,58	37,55	51,09	33,38	32,07
11	1	0,5	9	0,57	0,62	0,56	0,55	0,60	14,86	24,42	11,02	10,04
12	2	0,6	10	0,60	0,64	0,58	0,57	0,61	0,27	6,33	3,68	4,50
13	3	0,7	11	0,62	0,65	0,60	0,60	0,63	11,07	6,81	14,01	14,75
14	4	0,5	12	0,65	0,67	0,63	0,62	0,64	29,32	33,09	25,42	24,30
15	1	0,7	13	0,67	0,68	0,65	0,65	0,66	4,19	3,22	6,64	7,53
16	2	0,8	14	0,69	0,69	0,68	0,67	0,67	13,15	13,92	14,86	15,74
17	3	0,9	15	0,72	0,70	0,71	0,70	0,68	20,12	22,33	21,12	21,99
18	4	0,7	16	0,74	0,71	0,74	0,73	0,69	6,14	1,25	5,70	4,45
19	1	0,8	17	0,77	0,72	0,77	0,76	0,70	4,11	10,26	3,61	4,81
20	2	0,9	18	0,79	0,73	0,80	0,79	0,71	12,09	19,28	10,72	11,88
21	3	0,7	19	0,82	0,73	0,84	0,83	0,72	16,47	4,95	19,59	17,99
22	4	0,8	20	0,84	0,74	0,87	0,86	0,73	4,92	7,21	8,99	7,52
23	Итого	12,2		12,21	12,20	12,19	12,02	11,91	325,42	411,10	315,89	312,84
24	В среднем								16,27	20,55	15,79	15,64
25												18,87

Рисунок 8.10 – Расчет средней ошибки аппроксимации по уравнениям трендов

Таблица 8.3 – Оценка точности результатов аналитического выравнивания эндогенной переменной Y – интегральный показатель экономической безопасности организации

Вид тренда	Оценка уравнения тренда	Ошибка аппроксимации, %
линейный	$\tilde{y} = 0,36 + 0,02 \cdot t$	16,27
полиномиальный второй степени	$\tilde{y} = 0,40 + 0,01 \cdot t + 0,0006 \cdot t^2$	15,79
экспоненциальный	$\tilde{y} = 0,38 \cdot 2,72^{0,04t}$	15,64
степенной	$\tilde{y} = 0,34 \cdot t^{0,26}$	18,87
логарифмический	$\tilde{y} = 0,29 + 0,15 \cdot \ln(t)$	20,55

Как видно по данным таблицы 8.3, все модели имеют хорошую точность. Наименьшая ошибка аппроксимации у экспоненциального тренда. Поэтому в 4 задании по уравнению экспоненциального тренда будет разработан прогноз.

4 По модели имеющей самую высокую точность рассчитайте точечный и интервальный прогноз на следующий год

Прогноз на основании трендовых моделей (кривых роста) содержит два элемента: точечный и интервальный прогнозы.

Точечный прогноз – это прогноз, которым называется единственное значение прогнозируемого показателя. Это значение определяется подстановкой в уравнение выбранной кривой роста величины времени t , соответствующей периоду упреждения.

Расчет точечного прогноза на следующий, 6 год, в поквартальной динамике по уравнению экспоненциального тренда представлен на рисунке 8.11.

	A	B	C	D	E	F	G	H
1	Год	Квартал	Y	t	\tilde{y}			
2	1	1	0,4	1	0,40			
3		2	0,5	2	0,41			
4		3	0,6	3	0,43			
5		4	0,5	4	0,45			
6	2	1	0,3	5	0,47			
7		2	0,4	6	0,49			
8		3	0,5	7	0,51			
9		4	0,4	8	0,53			
10	3	1	0,5	9	0,55			
11		2	0,6	10	0,57			
12		3	0,7	11	0,60			
13		4	0,5	12	0,62			
14	4	1	0,7	13	0,65			
15		2	0,8	14	0,67			
16		3	0,9	15	0,70			
17		4	0,7	16	0,73			
18	5	1	0,8	17	0,76			
19		2	0,9	18	0,79			
20		3	0,7	19	0,83			
21		4	0,8	20	0,86			
22	6*	1*	-	21	$= 0,38 * 2,72^{(0,04 * D22)}$			
23		2*	-	22	0,93			
24		3*	-	23	0,97			
25		4*	-	24	1,01			

Рисунок 8.11 – Расчет точечного прогноза по уравнению экспоненциального тренда

Интервальный прогноз на базе трендовых моделей осуществляется путем расчета доверительного интервала такого интервала, в котором с определенной вероятностью можно ожидать появления фактического значения прогнозируемого экономического показателя.

Интервальный прогноз производится в соответствии с формулой:

$$\tilde{y}_{np} \pm t_{\alpha} \cdot S_{\tilde{y}},$$

где $S_{\tilde{y}}$ – среднее квадратическое отклонение уровней временного ряда от тренда;

t_{α} – критическое значение критерия Стьюдента при уровне значимости α с числом степеней свободы $(n - m)$.

$$S = \sqrt{\frac{\sum (y_i - \tilde{y}_i)^2}{n - m}},$$

где y_i и \tilde{y}_i – соответственно фактические и теоретические значения уровней временного ряда;

n – число уровней ряда;

m – число параметров в уравнении тренда.

Расчет числителя среднего квадратического отклонения уровней временного ряда от экспоненциального тренда представлен на рисунке 8.12.

	A	B	C	D	E	F
1	Год	Квартал	Y	t	\tilde{y}	$(y_i - \tilde{y}_i)^2$
2		1	0,4	1	0,40	$=(C2-E2)^2$
3		2	0,5	2	0,41	0,0074
4		3	0,6	3	0,43	0,0285
5	1	4	0,5	4	0,45	0,0026
6		1	0,3	5	0,47	0,0281
7		2	0,4	6	0,49	0,0076
8		3	0,5	7	0,51	0,0001
9	2	4	0,4	8	0,53	0,0165
10		1	0,5	9	0,55	0,0025
11		2	0,6	10	0,57	0,0007
12		3	0,7	11	0,60	0,0107
13	3	4	0,5	12	0,62	0,0148
14		1	0,7	13	0,65	0,0028
15		2	0,8	14	0,67	0,0158
16		3	0,9	15	0,70	0,0392
17	4	4	0,7	16	0,73	0,0010
18		1	0,8	17	0,76	0,0015
19		2	0,9	18	0,79	0,0114
20		3	0,7	19	0,83	0,0159
21	5	4	0,8	20	0,86	0,0036
22		1*	-	21	0,88	-
23		2*	-	22	0,92	-
24		3*	-	23	0,95	-
25	6*	4*	-	24	0,99	-
26	Итого					0,2106

Рисунок 8.12 – Расчет среднего квадратического отклонения уровней временного ряда от экспоненциального тренда

Подставим в формулу полученное значение суммы квадратов отклонений фактических значений от теоретических и найдем значение среднего квадратического отклонения:

$$S = \sqrt{\frac{0,2106}{20-2}} = \sqrt{0,012} = 0,108.$$

Критическое значение t – критерия Стьюдента на 5 % уровне значимости при 18 степенях свободы составляет $t(0,05;18) = 2,1$.

L	M	N	O	P
			Доверительный интервал	
Год	Квартал	точечный прогноз	нижняя граница	верхняя граница
6*	1*	0,88	$=N3-0,108*2,1$	
	2*	0,92	0,69	1,14
	3*	0,95	0,73	1,18
	4*	0,99	0,77	1,22

L	M	N	O	P	Q
			Доверительный интервал		
Год	Квартал	точечный прогноз	нижняя граница	верхняя граница	
6*	1*	0,88	0,65	$=N3-0,108*2,1$	
	2*	0,92	0,69	1,14	
	3*	0,95	0,73	1,18	
	4*	0,99	0,77	1,22	

Рисунок 8.13 – Расчет доверительных границ прогноза по экспоненциальному тренду

Прогноз по экспоненциальному тренду эндогенной переменной Y – интегральный показатель экономической безопасности организации на следующий год в поквартальной динамике представлен в таблице 8.3.

Таблица 8.3 – Результаты экстраполяции тенденции эндогенной переменной Y – интегральный показатель экономической безопасности организации по экспоненциальному тренду

Год	Квартал	Точечный прогноз	Доверительный интервал	
			нижняя граница	верхняя граница
6*	1*	0,88	0,65	1,11
	2*	0,92	0,69	1,14
	3*	0,95	0,73	1,18
	4*	0,99	0,77	1,22

При условии сохранения тенденции, наблюдавшейся в 1 – 5 годы в динамике интегрального показателя экономической безопасности организации, в периоде упреждения (6 году) будет наблюдаться рост показателя: в 4 квартале 6 года он может достигнуть 0,99 и с вероятностью 95 % будет колебаться от 0,77 до 1,22.

Контрольные вопросы к лабораторной работе № 8

- 1 Охарактеризуйте основные типы кривых роста, наиболее часто используемые на практике при построении трендовых моделей.
- 2 Какие методы проверки ряда на стационарность вы знаете?
- 3 Охарактеризуйте алгоритм одного из методов проверки ряда на стационарность.
- 4 Как оцениваются параметры основных моделей трендов?
- 5 Что понимается под экстраполяцией?
- 6 Как строятся точечный и интервальный прогнозы по трендовым моделям?

Тесты для самоконтроля к лабораторной работе № 8

- 1 Отметьте методы, используемые для проверки наличия тенденции во временном ряду:
 - А) критерий Дарбина-Уотсона;
 - Б) критерий Голдфелда-Квандта;
 - В) Критерий Спирмена;
 - Г) метод проверки разностей средних уровней;
 - Д) метод Фостера-Стюарта;
 - Е) критерий серий, основанный на медиане выборки;
 - Ж) критерий восходящих и нисходящих серий.

- 2 Получена кривая роста прибыли компании $y = 4,5 + 0,5t$. Это означает,

что:

- А) среднегодовой прирост прибыли компании составит 0,5 тыс. р.;
- Б) среднегодовой прирост прибыли компании составит $2 \cdot 0,5$ тыс. р.;
- В) прибыль компании возрастала со среднегодовым ускорением 0,5 тыс. р.
- Г) среднегодовой прирост прибыли компании составит 4,5 тыс. р.

3 Методом аналитического выравнивания по прямой выявлена тенденция ряда динамики: $y_t = 770,71 + 46,71t$.

Годы	Объем выручки предприятия (y), тыс. р.	t
2011	800	1
2012	857	2
2013	915	3
2014	976	4
2015	1038	5
2016	1052	6
2017	1065	7

Определите теоретическое значение показателя объема выручки в 2019 году:

- А) 1150;
- Б) 1200;
- В) 1038;
- Г) 1191.

4 Получена экспоненциальная кривая естественного прироста населения $y = 109 \cdot 0,5^t$. Это означает, что:

- А) среднегодовой темп роста естественного прироста составил 5 %;
- Б) среднегодовой темп роста естественного прироста составил 109 %;
- В) среднегодовой темп роста естественного прироста составил 0,5 %;
- Г) среднегодовой темп роста естественного прироста составил 50 %.

5 К какому классу кривых роста относится экспоненциальная кривая:

- А) функции, используемые для описания процессов с монотонным характером развития и отсутствием пределов роста;

- Б) кривые насыщения;
- В) кривые насыщения с отсутствием пределов роста;
- Г) кривые насыщения, имеющие точку перегиба.

6 Чему равно абсолютное ускорение по данным таблицы:

Годы	\hat{y} , тыс.р.	Δ_y , тыс.р.	T_{p_n} ,%	Δ'
2010	9,2			
2011	9,3	0,1	101,09	
2012	9,4	0,1	101,08	0
2013	9,5	0,1	101,06	0
2014	9,6	0,1	101,05	0
2015	9,7	0,1	101,04	0
2016	9,8	0,1	101,03	0
2017	9,9	0,1	101,02	0

- А) 1;
- Б) 0;
- В) 1,64;
- Г) 0,24.

7 Какой вид тренда характерен временному ряду прибыли организации, представленному в таблице:

Годы	\hat{y} , тыс.р.	Δ_y , тыс.р.	T_{p_n} ,%	Δ'
2010	9,37			
2011	9,44	0,07	100,75	
2012	9,5	0,06	100,64	-0,01
2013	9,55	0,05	100,53	-0,01
2014	9,59	0,04	100,42	-0,01
2015	9,62	0,03	100,31	-0,01
2016	9,64	0,02	100,21	-0,01
2017	9,65	0,01	100,10	-0,01

- А) параболический;
- Б) линейный;
- В) логистическая кривая;
- Г) экспоненциальный.

8 Временной ряд описывается параболическим трендом $\hat{y}_i = 8,83 + 0,12t_i - 0,004t_i^2$. Чему равно ускорение:

- А) -0,004;
- Б) минус 0,008;
- В) минус 0,002;
- Г) 0,008.

9 Временной ряд описывается параболическим трендом $\hat{y}_i = 8,83 + 0,12t_i - 0,004t_i^2$. Следовательно, имеем:

- А) восходящую ветвь с замедляющимся ростом уровней;
- Б) нисходящую ветвь с замедляющимся ростом уровней;
- В) нисходящую ветвь с ускоренным ростом уровней;
- Г) восходящую ветвь с замедляющимся снижением уровней.

10 Если при проверке надежности параметра тренда фактическое значение критерия Стьюдента больше табличного, это означает что:

- А) гипотеза о равенстве параметра тренда нулю отклоняется;
- Б) гипотеза о равенстве параметра тренда нулю принимается;
- В) вид уравнения тренда выбран неудачно;
- Г) нет достаточных оснований для принятия решения.

9 Лабораторная работа № 9. Моделирование временных рядов: исследование сезонности

Цель выполнения лабораторной работы: ознакомится с методами, позволяющими выявить наличие периодической составляющей во временном ряду, а также научиться оценивать уровень сезонности, осуществлять фильтрацию периодических составляющих временного ряда и их моделирование. Разрабатывать прогноз с учетом тенденции и сезонности.

Задания:

По исходным данным лабораторной работы № 1 для эндогенной переменной Y – интегральный показатель экономической безопасности организации выполните следующие задания:

1 На основе графического анализа проведите исследование компонентного состава временного ряда.

2 Оцените параметры следующих моделей:

а) тренд-сезонной аддитивной;

б) тренд-сезонной мультипликативной;

в) регрессии с включением фактора времени и фиктивных переменных.

3 Оцените точность полученных моделей с помощью ошибки аппроксимации.

4 По модели имеющей самую высокую точность рассчитайте точечный и интервальный прогноз на следующий год в поквартальном разрезе.

Реализация типовых заданий

1 На основе графического анализа проведите исследование компонентного состава временного ряда

Графический анализ исходного временного ряда (рисунок 9.1) свидетельствует о наличии трендовой компоненты, характер которой близок к линейному развитию: имеется устойчивая, ярко выраженная тенденция роста интегрального показателя экономической безопасности организации.

Также отчетливо видны сезонные колебания (период которых равен 1 году). Наиболее существенные «всплески» в динамике показателя просматриваются в 3 квартале.

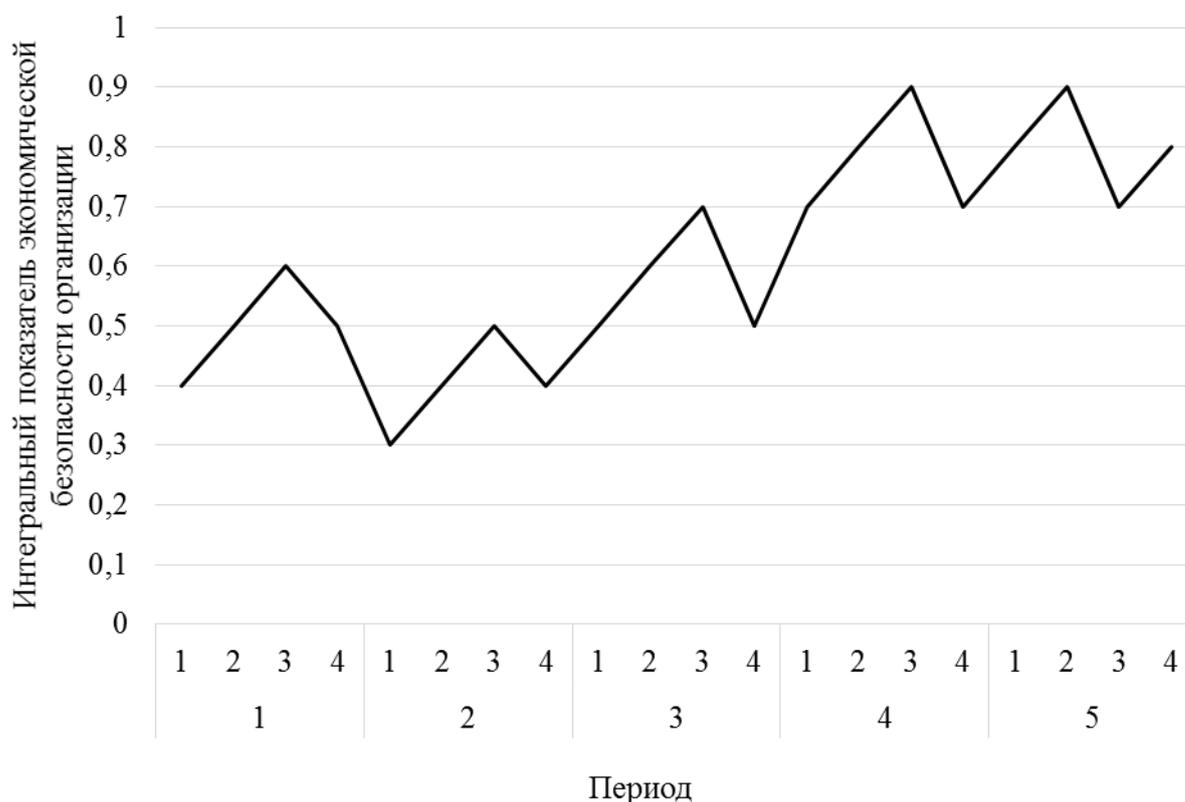


Рисунок 9.1 – Квартальная динамика интегрального показателя экономической безопасности организации

2 Оцените параметры следующих моделей:

а) тренд-сезонной аддитивной;

б) тренд-сезонной мультипликативной;

в) регрессии с включением фактора времени и фиктивных переменных.

Алгоритм построения тренд – сезонной аддитивной модели включает следующие шаги:

1 Проведем сглаживание временного ряда с помощью центрированной скользящей средней, по формуле (период скользящего равен 1 году, т.е. для нашего примера он равен 4):

$$\tilde{y}_i = \frac{1/2 \cdot y_{i-2} + y_{i-1} + y_i + y_{i+1} + 1/2 \cdot y_{i+2}}{4}.$$

Полученная скользящая средняя относится к середине базы скользящего, тогда:

$$\tilde{y}_3 = \frac{1/2 \cdot 0,4 + 0,5 + 0,6 + 0,5 + 1/2 \cdot 0,3}{4} = 0,5;$$

$$\tilde{y}_4 = \frac{1/2 \cdot 0,5 + 0,6 + 0,5 + 0,3 + 1/2 \cdot 0,4}{4} = 0,5;$$

...

$$\tilde{y}_{18} = \frac{1/2 \cdot 0,7 + 0,8 + 0,9 + 0,7 + 1/2 \cdot 0,8}{4} = 0,8.$$

Результаты сглаживания по центрированной скользящей средней представлены на рисунке 9.2.

	A	B	C	D	E	F	G
2	Год	Квартал	Y	\tilde{y}_i			
3	1	1	0,4				
4		2	0,5				
5		3	0,6	$= (1/2 * C3 + C4 + C5 + C6 + 1/2 * C7) / 4$			
6		4	0,5	0,5			
7	2	1	0,3	0,4			
8		2	0,4	0,4			
9		3	0,5	0,4			
10		4	0,4	0,5			
11	3	1	0,5	0,5			
12		2	0,6	0,6			
13		3	0,7	0,6			
14		4	0,5	0,7			
15	4	1	0,7	0,7			
16		2	0,8	0,8			
17		3	0,9	0,8			
18		4	0,7	0,8			
19	5	1	0,8	0,8			
20		2	0,9	0,8			
21		3	0,7				
22		4	0,8				

Рисунок 9.2 – Расчет скользящей средней для интегрального показателя экономической безопасности организации

2 Рассчитаем абсолютные показатели сезонности по формуле:

$$S_i = y_i - \tilde{y},$$

где \tilde{y} – выровненные скользящие средние.

Результаты расчетов представлены на рисунке 9.3.

	A	B	C	D	E	F
2	Год	Квартал	Y	\bar{y}_i	S_i	
3		1	0,4			
4		2	0,5			
5		3	0,6	0,5	=C5-D5	
6	1	4	0,5	0,5	0,0	
7		1	0,3	0,4	-0,1	
8		2	0,4	0,4	0,0	
9		3	0,5	0,4	0,1	
10	2	4	0,4	0,5	-0,1	
11		1	0,5	0,5	0,0	
12		2	0,6	0,6	0,0	
13		3	0,7	0,6	0,1	
14	3	4	0,5	0,7	-0,2	
15		1	0,7	0,7	0,0	
16		2	0,8	0,8	0,1	
17		3	0,9	0,8	0,1	
18	4	4	0,7	0,8	-0,1	
19		1	0,8	0,8	0,0	
20		2	0,9	0,8	0,1	
21		3	0,7			
22	5	4	0,8			

Рисунок 9.3 – Расчет абсолютных показателей сезонности для интегрального показателя экономической безопасности организации

3 Рассчитываем средние показатели сезонности для одноименных кварталов (месяцев):

$$\bar{S}_j = \frac{1}{n} \sum S_i .$$

Для 1 квартала средний показатель сезонности составит:

$$\bar{S}_1 = \frac{-0,1+0,0+0,0+0,0}{4} = -0,041 .$$

Аналогично рассчитываем средние показатели сезонности для других кварталов. Результаты расчетов представлены на рисунке 9.4.

	A	B	C	D	E	F	G	H	I	J
2	Год	Квартал	Y	\tilde{y}_i	S_i		Квартал	S_j		
3	1	1		0,4			1	$=(E7+E11+E15+E19)/4$		
4		2		0,5			2	0,047		
5		3		0,6	0,5	0,1	3	0,072		
6		4		0,5	0,5	0,0	4	-0,084		
7	2	1		0,3	0,4	-0,1	Итого	-0,006		
8		2		0,4	0,4	0,0				
9		3		0,5	0,4	0,1				
10		4		0,4	0,5	-0,1				
11	3	1		0,5	0,5	0,0				
12		2		0,6	0,6	0,0				
13		3		0,7	0,6	0,1				
14	4	4		0,5	0,7	-0,2				
15		1		0,7	0,7	0,0				
16		2		0,8	0,8	0,1				
17	5	3		0,9	0,8	0,1				
18		4		0,7	0,8	-0,1				
19		1		0,8	0,8	0,0				
20	5	2		0,9	0,8	0,1				
21		3		0,7						
22		4		0,8						

Рисунок 9.4 – Расчет средних показателей сезонности для интегрального показателя экономической безопасности организации

4 Так как $\sum \bar{S}_j \neq 0$, проводим корректировку сезонной компоненты по формуле:

$$\hat{S}_j = \bar{S}_j - \frac{1}{m} \sum_{i=1}^m \bar{S}_j .$$

Для 1 квартала скорректированный показатель сезонности составит:

$$\hat{S}_1 = -0,041 - \frac{-0,006}{4} = -0,039 .$$

Аналогично рассчитываем скорректированные показатели сезонности для других кварталов. Результаты расчетов представлены на рисунке 9.5.

Квартал	\bar{S}_j	\hat{S}_j
1	-0,041	$=H3 - \$H\$7/4$
2	0,047	0,048
3	0,072	0,073
4	-0,084	-0,083
Итого	-0,006	0,000

Рисунок 9.5 – Расчет скорректированных показателей сезонности для интегрального показателя экономической безопасности организации

5 Проводим десеонализацию временного ряда: из исходных уровней вычитаем скорректированную сезонную компоненту:

$$y_i - \hat{S}_i.$$

Результаты расчетов представлены на рисунке 9.6.

	A	B	C	D	E	F	G	H
2	Год	Квартал	Y	\tilde{y}_i	S_i	S_j	$y_i - S_i$	
3		1	0,4			-0,039	$=C3-F3$	
4		2	0,5			0,048	0,452	
5		3	0,6	0,5	0,1	0,073	0,527	
6	1	4	0,5	0,5	0,0	-0,083	0,583	
7		1	0,3	0,4	-0,1	-0,039	0,339	
8		2	0,4	0,4	0,0	0,048	0,352	
9		3	0,5	0,4	0,1	0,073	0,427	
10	2	4	0,4	0,5	-0,1	-0,083	0,483	
11		1	0,5	0,5	0,0	-0,039	0,539	
12		2	0,6	0,6	0,0	0,048	0,552	
13		3	0,7	0,6	0,1	0,073	0,627	
14	3	4	0,5	0,7	-0,2	-0,083	0,583	
15		1	0,7	0,7	0,0	-0,039	0,739	
16		2	0,8	0,8	0,1	0,048	0,752	
17		3	0,9	0,8	0,1	0,073	0,827	
18	4	4	0,7	0,8	-0,1	-0,083	0,783	
19		1	0,8	0,8	0,0	-0,039	0,839	
20		2	0,9	0,8	0,1	0,048	0,852	
21		3	0,7			0,073	0,627	
22	5	4	0,8			-0,083	0,883	
23								

Рисунок 9.6 – Расчет десеонализованного временного ряда интегрального показателя экономической безопасности организации

6 По десеоналированному временного ряда проводим аналитическое выравнивание.

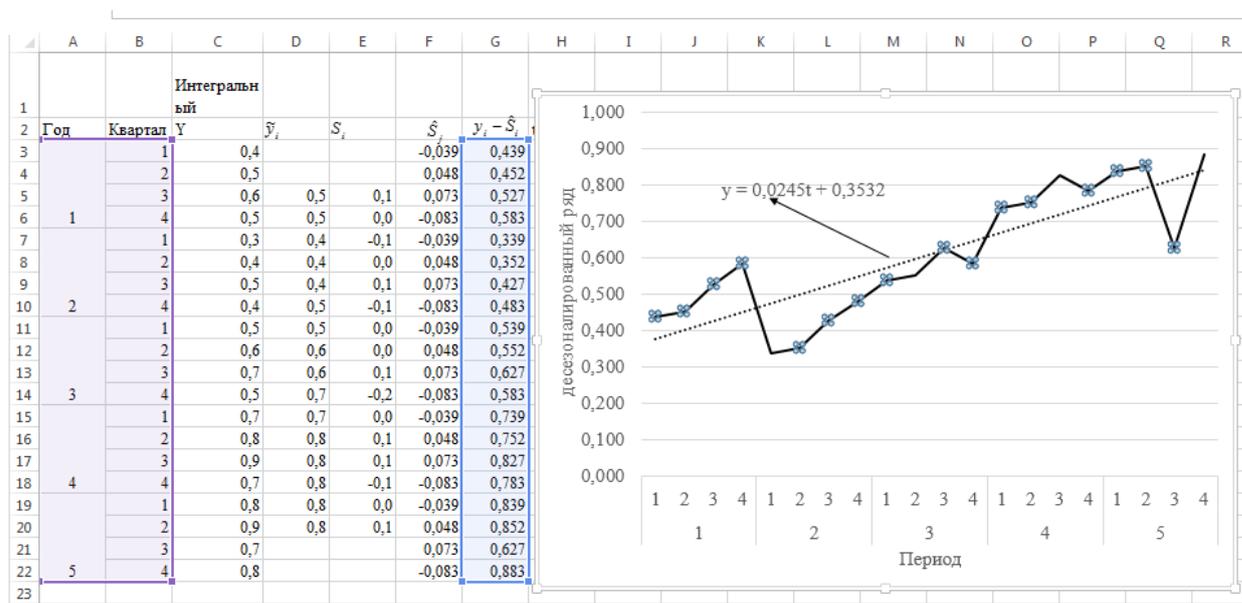


Рисунок 9.7 – Аналитическое выравнивание десеоналированного временного ряда интегрального показателя экономической безопасности организации

Уравнение тренда имеет вид: $\hat{y}_i = 0,35 + 0,03 t$

7 Рассчитываем тренд с учетом сезонности.

Предварительно рассчитаем теоретические значения по линейному тренду (рисунок 9.8).

	A	B	C	D	E	F	G	H	I	J
2	Год	Квартал	Y	\tilde{y}_i	S_i	S_j	$y_i - S_i$	t	\hat{y}_i	
3		1		0,4		-0,039	0,439	1	$= 0,03 * 1 + 0,35$	
4		2		0,5		0,048	0,452	2	0,4	
5		3		0,6	0,5	0,073	0,527	3	0,4	
6	1	4		0,5	0,5	-0,083	0,583	4	0,5	
7		1		0,3	0,4	-0,039	0,339	5	0,5	
8		2		0,4	0,4	0,048	0,352	6	0,5	
9		3		0,5	0,4	0,073	0,427	7	0,5	
10	2	4		0,4	0,5	-0,083	0,483	8	0,5	
11		1		0,5	0,5	0,0	-0,039	0,539	9	0,6
12		2		0,6	0,6	0,0	0,048	0,552	10	0,6
13		3		0,7	0,6	0,1	0,073	0,627	11	0,6
14	3	4		0,5	0,7	-0,2	-0,083	0,583	12	0,6
15		1		0,7	0,7	0,0	-0,039	0,739	13	0,7
16		2		0,8	0,8	0,1	0,048	0,752	14	0,7
17		3		0,9	0,8	0,1	0,073	0,827	15	0,7
18	4	4		0,7	0,8	-0,1	-0,083	0,783	16	0,7
19		1		0,8	0,8	0,0	-0,039	0,839	17	0,8
20		2		0,9	0,8	0,1	0,048	0,852	18	0,8
21		3		0,7			0,073	0,627	19	0,8
22	5	4		0,8			-0,083	0,883	20	0,8

Рисунок 9.8 – Расчет теоретических значений десезонализованного временного ряда интегрального показателя экономической безопасности организации по линейному тренду

По формуле $y_s = \hat{y}_i + \hat{S}_i$ рассчитываем теоретические значения тренд-сезонной аддитивной модели (рисунок 9.9).

	A	B	C	D	E	F	G	H	I	J	K
2	Год	Квартал	Y	\tilde{y}_i	S_i	\hat{S}_i	$y_i - S_i$	t	\hat{y}_i	y_s	
3		1		0,4		-0,039	0,439	1	0,4	$= F3 + I3$	
4		2		0,5		0,048	0,452	2	0,4	0,5	
5		3		0,6	0,5	0,073	0,527	3	0,4	0,5	
6	1	4		0,5	0,5	-0,083	0,583	4	0,5	0,4	
7		1		0,3	0,4	-0,039	0,339	5	0,5	0,4	
8		2		0,4	0,4	0,048	0,352	6	0,5	0,5	
9		3		0,5	0,4	0,073	0,427	7	0,5	0,6	
10	2	4		0,4	0,5	-0,083	0,483	8	0,5	0,5	
11		1		0,5	0,5	0,0	-0,039	0,539	9	0,6	0,5
12		2		0,6	0,6	0,0	0,048	0,552	10	0,6	0,6
13		3		0,7	0,6	0,1	0,073	0,627	11	0,6	0,7
14	3	4		0,5	0,7	-0,2	-0,083	0,583	12	0,6	0,6
15		1		0,7	0,7	0,0	-0,039	0,739	13	0,7	0,6
16		2		0,8	0,8	0,1	0,048	0,752	14	0,7	0,7
17		3		0,9	0,8	0,1	0,073	0,827	15	0,7	0,8
18	4	4		0,7	0,8	-0,1	-0,083	0,783	16	0,7	0,7
19		1		0,8	0,8	0,0	-0,039	0,839	17	0,8	0,7
20		2		0,9	0,8	0,1	0,048	0,852	18	0,8	0,8
21		3		0,7			0,073	0,627	19	0,8	0,9
22	5	4		0,8			-0,083	0,883	20	0,8	0,8

Рисунок 9.9 – Расчет тренд-сезонной аддитивной модели временного ряда интегрального показателя экономической безопасности организации

Прогнозирование по тренд – сезонной мультипликативной модели

1 Проведем сглаживание временного ряда с помощью центрированной скользящей средней, аналогично как при построении тренд-сезонной аддитивной модели

2 Рассчитаем коэффициенты сезонности по формуле

$$K_s = \frac{y_i}{\tilde{y}_i}.$$

Результаты расчетов скользящей средней и коэффициента сезонности представлены на рисунке 9.10.

	A	B	C	D	E
2	Год	Квартал	Y	\tilde{y}_i	K_s
3		1		0,4	
4		2		0,5	
5		3		0,6	0,5 = C5/D5
6	1	4		0,5	0,5
7		1		0,3	0,4
8		2		0,4	0,4
9		3		0,5	0,4
10	2	4		0,4	0,5
11		1		0,5	0,5
12		2		0,6	0,6
13		3		0,7	0,6
14	3	4		0,5	0,7
15		1		0,7	0,7
16		2		0,8	0,8
17		3		0,9	0,8
18	4	4		0,7	0,8
19		1		0,8	0,8
20		2		0,9	0,8
21		3		0,7	
22	5	4		0,8	

Рисунок 9.10 – Расчет коэффициентов сезонности для интегрального показателя экономической безопасности организации

3 Определяем средние показатели сезонности для одноименных кварталов (месяцев):

$$\bar{K}_j = \frac{1}{n} \sum K_{Si}.$$

т.е. для 1 квартала средний коэффициент сезонности составит:

$$\bar{K}_1 = \frac{0,7 + 1,0 + 1,0 + 1,0}{4} = 0,91.$$

Аналогично рассчитывают для других кварталов. Результаты расчета средних коэффициентов сезонности представлены на рисунке 9.11.

	A	B	C	D	E	F	G	H	I	J
2	Год	Квартал	Y	\bar{y}_i	K_s		Квартал	\bar{K}_j		
3		1		0,4			1	$= (E7+E11+E15+E19)/4$		
4		2		0,5			2	1,061		
5		3		0,6	0,5	1,2	3	0,871		
6	1	4		0,5	0,5	1,1	4	0,618		
7		1		0,3	0,4	0,7	Итого	3,461		
8		2		0,4	0,4	1,0				
9		3		0,5	0,4	1,2				
10	2	4		0,4	0,5	0,8				
11		1		0,5	0,5	1,0				
12		2		0,6	0,6	1,1				
13		3		0,7	0,6	1,2				
14	3	4		0,5	0,7	0,8				
15		1		0,7	0,7	1,0				
16		2		0,8	0,8	1,1				
17		3		0,9	0,8	1,1				
18	4	4		0,7	0,8	0,9				
19		1		0,8	0,8	1,0				
20		2		0,9	0,8	1,1				
21		3		0,7						
22	5	4		0,8						

Рисунок 9.11 – Расчет средних коэффициентов сезонности для интегрального показателя экономической безопасности организации

4 Так как сумма средних коэффициентов сезонности не равна 4, проведем их корректировку по формуле:

$$\hat{K}_j = \bar{K}_j \cdot \frac{4}{\sum \bar{K}_j}$$

Так скорректированный коэффициент сезонности для 1 квартала составит:

$$\hat{K}_1 = 0,91 \cdot \frac{4}{3,461} = 1,051 \text{ и т.д.}$$

Результаты расчетов средних и скорректированных показателей сезонности представлены на рисунке 9.12.

G	H	I	J
Квартал	\bar{K}_j	\hat{K}_j	
1	0,910	$=H3*4/H\$7$	
2	1,061	1,227	
3	0,871	1,007	
4	0,618	0,715	
Итого	3,461	4,000	

Рисунок 9.12 – Расчет скорректированных коэффициентов сезонности для интегрального показателя экономической безопасности организации

5 Определяем десеоналированный ряд объема производства:

$$\frac{y_i}{\hat{K}_j}$$

По десеоналированному временному ряду проводим аналитическое выравнивание по линейному тренду.

Уравнение тренда имеет вид: $\hat{y}_t = 0,3382 + 0,0278 \cdot t$

Рассчитываем тренд с учетом сезонности:

$$y_s = \hat{y}_t \cdot \hat{K}_j$$

Результаты расчетов представлены на рисунке 9.13.

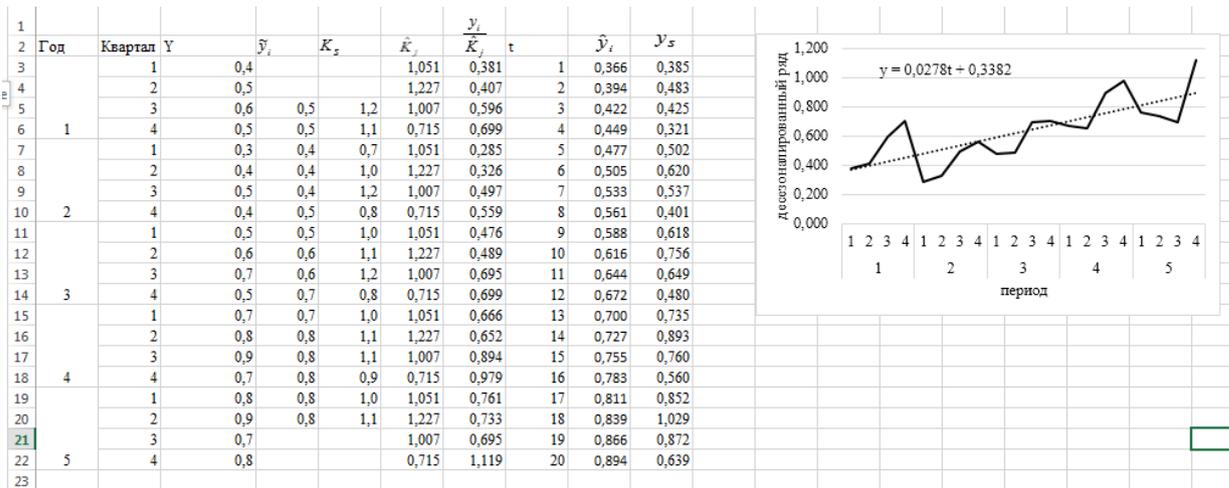


Рисунок 9.13 – Расчет тренд-сезонной мультипликативной модели временного ряда интегрального показателя экономической безопасности организации

Прогнозирование по модели регрессии с включением фактора времени и фиктивных переменных

Спрогнозируем интегральный показатель экономической безопасности организации с помощью модели регрессии с включением фактора времени и фиктивных переменных. Модель для квартальной динамики имеет вид:

$$y_t = a + bt + c_1x_1 + c_2x_2 + c_3x_3 + \varepsilon_t,$$

где $x_1 = \begin{cases} 1 & \text{для 1 квартала} \\ 0 & \text{во всех остальных случаях} \end{cases}$

$$x_2 = \begin{cases} 1 & \text{для 2 квартала} \\ 0 & \text{во всех остальных случаях} \end{cases}$$

$$x_3 = \begin{cases} 1 & \text{для 3 квартала} \\ 0 & \text{во всех остальных случаях} \end{cases}$$

Занесем значение фиктивных переменных и фактора времени в таблицу (рисунок 9.14).

	A	B	C	D	E	F	G
2	Год	Квартал	Y	t	X ₁	X ₂	X ₃
3	1	1	0,4	1	1	0	0
4		2	0,5	2	0	1	0
5		3	0,6	3	0	0	1
6		4	0,5	4	0	0	0
7	2	1	0,3	5	1	0	0
8		2	0,4	6	0	1	0
9		3	0,5	7	0	0	1
10	4	0,4	8	8	0	0	0
11	3	1	0,5	9	1	0	0
12		2	0,6	10	0	1	0
13		3	0,7	11	0	0	1
14		4	0,5	12	0	0	0
15	4	1	0,7	13	1	0	0
16		2	0,8	14	0	1	0
17		3	0,9	15	0	0	1
18		4	0,7	16	0	0	0
19	5	1	0,8	17	1	0	0
20		2	0,9	18	0	1	0
21		3	0,7	19	0	0	1
22		4	0,8	20	0	0	0

Рисунок 9.14 – Исходные данные для расчета параметров уравнения регрессии с фиктивными переменными во временном ряду интегрального показателя экономической безопасности организации

Оценим параметры уравнения традиционным МНК с помощью табличного редактора Excel (рисунок 9.15).

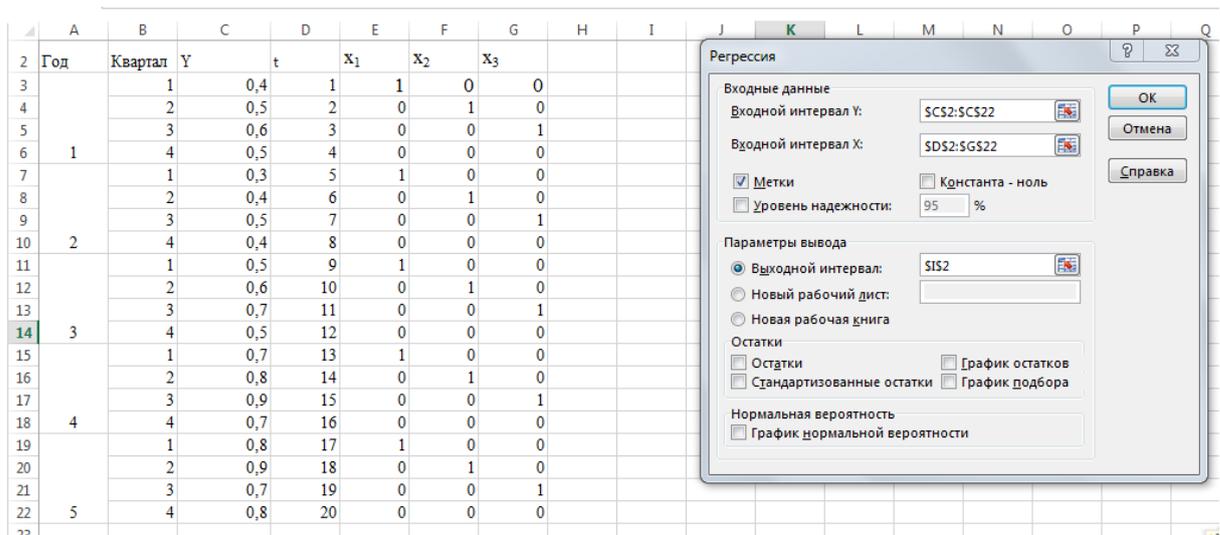


Рисунок 9.15 – Построение регрессионной модели с фиктивными переменными

Результаты регрессионного анализа представлены на рисунке 9.16.

	I	J	K	L	M	N	O	P	Q
Регрессионная статистика									
Множественный R		0,86							
R-квадрат		0,73							
Нормированный R-квадрат		0,66							
Стандартная ошибка		0,10							
Наблюдения		20							
Дисперсионный анализ									
		<i>df</i>	<i>SS</i>	<i>MS</i>	<i>F</i>	<i>Значимость F</i>			
Регрессия		4	0,44	0,11	10,29	0,00			
Остаток		15	0,16	0,01					
Итого		19	0,60						
		<i>Коэффициенты</i>	<i>Стандартная ошибка</i>	<i>t-статистика</i>	<i>P-значение</i>	<i>Верхние 95%</i>	<i>Нижние 95,0%</i>	<i>Верхние 95,0%</i>	<i>Нижние 95,0%</i>
Y-пересечение		0,29	0,07	4,27	0,00	0,14	0,43	0,14	0,43
t		0,02	0,00	5,98	0,00	0,02	0,03	0,02	0,03
x1		0,03	0,07	0,50	0,63	-0,11	0,17	-0,11	0,17
x2		0,11	0,07	1,65	0,12	-0,03	0,25	-0,03	0,25
x3		0,12	0,07	1,90	0,08	-0,02	0,26	-0,02	0,26

Рисунок 9.16 – Результаты оценивания регрессионной модели с фиктивными переменными

Уравнение регрессии примет вид:

$$\hat{y}_t = 0,29 + 0,02 \cdot t + 0,103 \cdot x_1 + 0,11 \cdot x_2 + 0,12 \cdot x_3.$$

Параметры c_1, c_2, c_3 характеризуют отклонения уровней временного ряда от уровней, учитывающих сезонные воздействия в 4 квартале. Величина параметра $b = 0,02$ говорит о том, что в среднем за квартал происходит рост интегрального показателя экономической безопасности организации на 0,02.

На рисунке 9.17 представлены результаты расчета теоретических значений интегрального показателя экономической безопасности организации по уравнению регрессии, а также график фактических и теоретических значений эндогенной переменной.

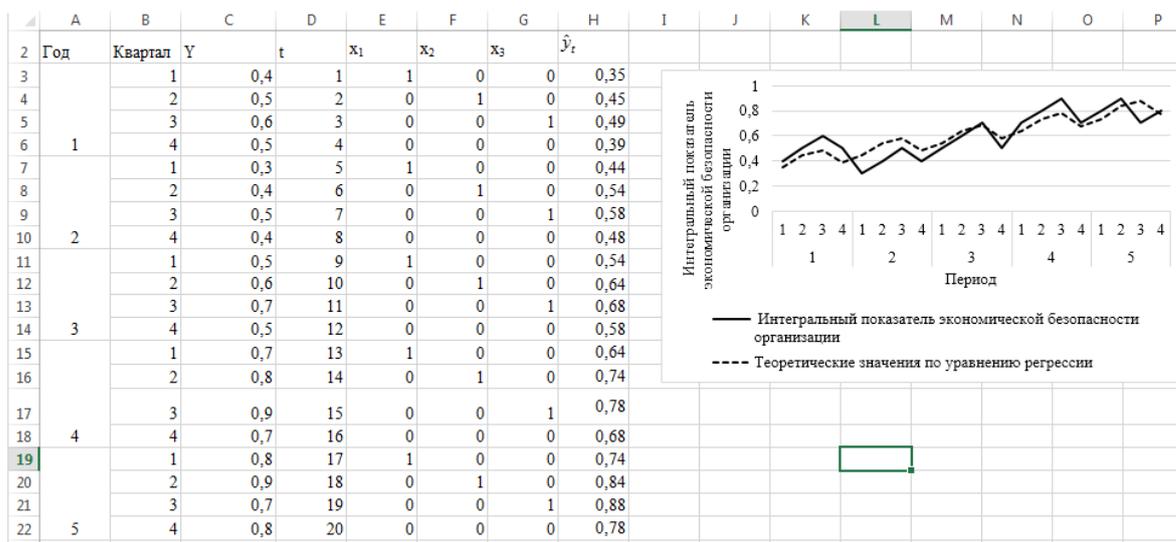


Рисунок 9.17 – Моделирование сезонных колебаний интегрального показателя экономической безопасности организации с помощью фиктивных переменных

3 Оцените точность полученных моделей с помощью ошибки аппроксимации

Ошибка аппроксимации использовалась для оценки точности моделей в лабораторных работах 1 – 8.

На рисунке 9.18 приведены расчеты средней ошибки аппроксимации для оцененных моделей.

№	A	B	C	D			E			F			G			H			I		
				Теоретические значения по модели			Ошибка аппроксимации модели														
1			Интегральный показатель экономической безопасности организации	аддитивной тренд-сезонной	мультипликативной тренд-сезонной	регрессии	аддитивной тренд-сезонной	мультипликативной тренд-сезонной	регрессии												
2	Год	Квартал																			
3		1	0,4	0,3	0,4	0,4	25,00	3,75	12,50												
4		2	0,5	0,5	0,5	0,5	0,00	3,40	10,00												
5		3	0,6	0,5	0,4	0,5	16,67	29,17	18,33												
6	1	4	0,5	0,4	0,3	0,4	20,00	35,80	22,00												
7		1	0,3	0,4	0,5	0,4	33,33	67,33	46,67												
8		2	0,4	0,5	0,6	0,5	25,00	55,00	35,00												
9		3	0,5	0,6	0,5	0,6	20,00	7,40	16,00												
10	2	4	0,4	0,5	0,4	0,5	25,00	0,25	20,00												
11		1	0,5	0,5	0,6	0,5	0,00	23,60	8,00												
12		2	0,6	0,6	0,8	0,6	0,00	26,00	6,67												
13		3	0,7	0,7	0,6	0,7	0,00	7,29	2,86												
14	3	4	0,5	0,6	0,5	0,6	20,00	4,00	16,00												
15		1	0,7	0,6	0,7	0,6	14,29	5,00	8,57												
16		2	0,8	0,7	0,9	0,7	12,50	11,63	7,50												
17		3	0,9	0,8	0,8	0,8	11,11	15,56	13,33												
18	4	4	0,7	0,7	0,6	0,7	0,00	20,00	2,86												
19		1	0,8	0,7	0,9	0,7	12,50	6,50	7,50												
20		2	0,9	0,8	1,0	0,8	11,11	14,33	6,67												
21		3	0,7	0,9	0,9	0,9	28,57	24,57	25,71												
22	5	4	0,8	0,8	0,6	0,8	0,00	20,13	2,50												
23	итого						275,08	380,70	288,67												
24	в среднем						13,75	19,03	14,43												

Рисунок 9.18 – Расчет средней ошибки аппроксимации по тренд-сезонным моделям и уравнению регрессии

Как видно на рисунке 9.18, все модели имеют хорошую точность. Наименьшая ошибка аппроксимации у аддитивной тренд-сезонной модели. Поэтому в 4 задании по этой модели будет разработан прогноз.

4 По модели имеющей самую высокую точность рассчитайте точечный и интервальный прогноз на следующий год.

Точечный прогноз по аддитивной тренд-сезонной модели определяется по формуле:

$$y_{Sp} = \hat{y}_p + \hat{S}_p,$$

где \hat{y}_p – точечный прогноз по тренду для десеоналированного ряда;

\hat{S}_p – скорректированный показатель сезонности для прогнозируемого периода (квартала, месяца);

по мультипликативной тренд-сезонной модели определяется по формуле:

$$y_{Sp} = \hat{y}_p \cdot \hat{K}_p,$$

где \hat{y}_p – точечный прогноз по тренду для десеоналированного ряда;

\hat{K}_p – скорректированный коэффициент сезонности для прогнозируемого периода (квартала, месяца);

по уравнению регрессии с фактором времени и фиктивными переменными:

$$y_p = a + bt_p + c_1x_{1p} + c_2x_{2p} + c_3x_{3p},$$

где t_p – значение фактора времени для прогнозируемого периода (квартала, месяца);

x_{1p}, x_{2p}, x_{3p} – значения фиктивных переменных для прогнозируемого периода (квартала, месяца).

Расчет точечного прогноза на следующий, 6 год, в поквартальной динамике по тренд–сезонной аддитивной модели представлен на рисунке 9.19.

	A	B	C	D	E	F	G	H	I	J
2	Год	Квартал	Y	\hat{y}_i	S_i	\hat{S}_i	$y_i - S_i$	t	\hat{y}_i	y_s
3		1	0,4			-0,039	0,439	1	0,4	0,3
4		2	0,5			0,048	0,452	2	0,4	0,5
5		3	0,6	0,5	0,1	0,073	0,527	3	0,4	0,5
6	1	4	0,5	0,5	0,0	-0,083	0,583	4	0,5	0,4
7		1	0,3	0,4	-0,1	-0,039	0,339	5	0,5	0,4
8		2	0,4	0,4	0,0	0,048	0,352	6	0,5	0,5
9		3	0,5	0,4	0,1	0,073	0,427	7	0,5	0,6
10	2	4	0,4	0,5	-0,1	-0,083	0,483	8	0,5	0,5
11		1	0,5	0,5	0,0	-0,039	0,539	9	0,6	0,5
12		2	0,6	0,6	0,0	0,048	0,552	10	0,6	0,6
13		3	0,7	0,6	0,1	0,073	0,627	11	0,6	0,7
14	3	4	0,5	0,7	-0,2	-0,083	0,583	12	0,6	0,6
15		1	0,7	0,7	0,0	-0,039	0,739	13	0,7	0,6
16		2	0,8	0,8	0,1	0,048	0,752	14	0,7	0,7
17		3	0,9	0,8	0,1	0,073	0,827	15	0,7	0,8
18	4	4	0,7	0,8	-0,1	-0,083	0,783	16	0,7	0,7
19		1	0,8	0,8	0,0	-0,039	0,839	17	0,8	0,7
20		2	0,9	0,8	0,1	0,048	0,852	18	0,8	0,8
21		3	0,7			0,073	0,627	19	0,8	0,9
22	5	4	0,8			-0,083	0,883	20	0,8	0,8
23		1*				-0,039		21	1,0	=F23+I23
24		2*				0,048		22	0,9	0,9
25		3*				0,073		23	0,9	1,0
26	6*	4*				-0,083		24	0,9	0,9

Рисунок 9.19 – Расчет точечного прогноза тренд-сезонной аддитивной модели

Интервальный прогноз производится в соответствии с формулой:

$$y_p \pm t_\alpha \cdot S_{\hat{y}}$$

где $S_{\hat{y}}$ – среднее квадратическое отклонение уровней временного ряда от теоретических уровней, рассчитанных по тренд-сезонной аддитивной модели (мультипликативной модели, модели регрессии с включением фактора времени и фиктивных переменных);

t_α – критическое значение критерия Стьюдента при уровне значимости α с числом степеней свободы $(n - m)$.

$$S = \sqrt{\frac{\sum (y_i - \tilde{y}_i)^2}{n - m}},$$

где y_i и \tilde{y}_i – соответственно фактические и теоретические значения уровней временного ряда;

n – число уровней ряда;

m – число параметров в модели.

Расчет числителя среднего квадратического отклонения уровней временного ряда от аддитивной тренд-сезонной модели представлен на рисунке 9.20.

№	A	B	C	D	E	F	G	H	I	J	K
2	Год	Квартал	Y	\tilde{y}_i	S_i	\hat{S}_i	$y_i - S_i$	t	Y_i	Y_s	$(y_i - \tilde{y}_i)^2$
3		1	0,4			-0,039	0,439	1	0,4	0,3	0,003
4		2	0,5			0,048	0,452	2	0,4	0,5	0,002
5		3	0,6	0,5	0,1	0,073	0,527	3	0,4	0,5	0,010
6	1	4	0,5	0,5	0,0	-0,083	0,583	4	0,5	0,4	0,017
7		1	0,3	0,4	-0,1	-0,039	0,339	5	0,5	0,4	0,019
8		2	0,4	0,4	0,0	0,048	0,352	6	0,5	0,5	0,022
9		3	0,5	0,4	0,1	0,073	0,427	7	0,5	0,6	0,010
10	2	4	0,4	0,5	-0,1	-0,083	0,483	8	0,5	0,5	0,004
11		1	0,5	0,5	0,0	-0,039	0,539	9	0,6	0,5	0,001
12		2	0,6	0,6	0,0	0,048	0,552	10	0,6	0,6	0,002
13		3	0,7	0,6	0,1	0,073	0,627	11	0,6	0,7	0,000
14	3	4	0,5	0,7	-0,2	-0,083	0,583	12	0,6	0,6	0,004
15		1	0,7	0,7	0,0	-0,039	0,739	13	0,7	0,6	0,005
16		2	0,8	0,8	0,1	0,048	0,752	14	0,7	0,7	0,003
17		3	0,9	0,8	0,1	0,073	0,827	15	0,7	0,8	0,011
18	4	4	0,7	0,8	-0,1	-0,083	0,783	16	0,7	0,7	0,001
19		1	0,8	0,8	0,0	-0,039	0,839	17	0,8	0,7	0,005
20		2	0,9	0,8	0,1	0,048	0,852	18	0,8	0,8	0,003
21		3	0,7			0,073	0,627	19	0,8	0,9	0,037
22	5	4	0,8			-0,083	0,883	20	0,8	0,8	0,002
23		1*				-0,039		21	1,0	0,9	-
24		2*				0,048		22	0,9	0,9	-
25		3*				0,073		23	0,9	1,0	-
26	6*	4*				-0,083		24	0,9	0,9	-
27	итого										0,162

Рисунок 9.20 – Расчет среднего квадратического отклонения уровней временного ряда от экспоненциального тренда

Подставим в формулу полученное значение суммы квадратов отклонений фактических значений от теоретических и найдем значение среднего квадратического отклонения:

$$S = \sqrt{\frac{0,162}{20-2}} = \sqrt{0,009} = 0,095.$$

Критическое значение t – критерия Стьюдента на 5 % уровне значимости при 18 степенях свободы составляет $t(0,05;18) = 2,1$.

Подставив полученные значения в формулу, найдем интервальный прогноз по аддитивной тренд-сезонной модели интегрального показателя экономической безопасности организации на следующий год в поквартальной динамике (таблица 9.1).

Таблица 9.1 – Результаты прогнозирования по аддитивной тренд-сезонной модели интегрального показателя экономической безопасности организации

Год	Квартал	Точечный прогноз	Доверительный интервал	
			нижняя граница	верхняя граница
6*	1*	0,94	0,74	1,14
	2*	0,94	0,74	1,14
	3*	0,99	0,79	1,19
	4*	0,86	0,66	1,06

При условии сохранения выявленных тенденций и закономерностей, наблюдавшихся в 1 – 5 годы в динамике интегрального показателя экономической безопасности организации, в периоде упреждения (6 году) будет наблюдаться рост показателя в 3 квартале 6 года он может достигнуть 0,99 и с вероятностью 95 % будет колебаться от 0,79 до 1,19; а в 4 квартале 6

года он снизится до 0,86 и с вероятностью 95 % будет колебаться от 0,66 до 1,06.

Контрольные вопросы к лабораторной работе № 9

- 1 Что понимается под сезонными колебаниями?
- 2 Назовите основные этапы построения аддитивной модели сезонности.
- 3 Назовите основные этапы построения мультипликативной модели сезонности.
- 4 Как проводится моделирование сезонных колебаний с помощью фиктивных переменных?
- 5 Как рассчитывается прогноз по аддитивной и мультипликативной тренд-сезонным моделям?

Тесты для самоконтроля к лабораторной работе № 9

1 Для расчета теоретических значений по тренд - сезонной аддитивной модели используется формула:

А) $y_s = \hat{y}_t + \bar{S}$;

Б) $y_s = \hat{y}_t \cdot \hat{S}_j$;

В) $y_s = \frac{\hat{S}_j}{\hat{y}}$;

Г) $y_s = \hat{y}_t + \hat{S}_j$.

2 Чему должна быть равна сумма средних абсолютных показателей сезонности?

А) 0;

Б) 1;

В) 4;

Г) нет ограничений.

3 Для расчета теоретических значения по тренд - сезонной мультипликативной модели используется формула:

А) $y_s = \hat{y}_t \cdot \hat{S}_j$;

Б) $y_s = \hat{y}_t / \hat{S}_j$;

В) $y_s = \hat{y}_t \cdot \hat{K}_j$;

Г) $y_s = \hat{y}_t \cdot \bar{K}$.

4 Прогнозная аддитивная модель динамического ряда с сезонными колебаниями при отсутствии в нем тенденции имеет вид:

А) $y_p = \bar{y}_p \pm S$;

Б) $y_p = y_n \pm \bar{S}$;

В) $y_p = y_n / \bar{S}$;

Г) $y_p = \bar{y}_p \cdot S$.

5 Что характеризует параметр b в модели регрессии с включением фактора времени и фиктивных переменных ($y_t = a + bt + c_1x_1 + \dots + c_jx_j + \dots + c_{k-1}x_{k-1} + \varepsilon_t$)?

А) отклонение уровней временного ряда от уровней, учитывающих сезонные воздействия;

Б) сумму начального уровня ряда и сезонной компоненты в 4 квартале;

В) абсолютное изменение уровней ряда под воздействием тенденции;

Г) относительное изменение уровней ряда под воздействием тенденции.

6 Прогнозная мультипликативная модель динамического ряда с сезонными колебаниями при отсутствии в нем тенденции имеет вид:

А) $y_p = \bar{y}_p \cdot K_S$;

Б) $y_p = y_n \cdot K_S$;

В) $y_p = y_n \pm \bar{S}$;

Г) $y_p = \bar{y}_p \pm K_s$.

7 Чему должна быть равна сумма средних коэффициентов сезонности при поквартальном наблюдении?

А) 0;

Б) 1;

В) 4;

Г) нет ограничений.

8 В мультипликативной модели сезонность выражается:

А) в виде абсолютной величины;

Б) в виде среднего уровня ряда;

В) в виде среднего абсолютного прироста;

Г) в виде относительной величины.

9 Периодические колебания, возникающие под влиянием смены времени года называются:

А) циклическими;

Б) сезонными;

В) тенденцией;

Г) случайными.

10 Аддитивная стационарная модель при наличии сезонных колебаний имеет вид:

А) $y_i = \hat{y}_i \cdot K_s \cdot E$;

Б) $y_i = \bar{y} + S + \varepsilon$;

В) $y_i = \hat{y}_i + S + \varepsilon$;

Г) $y_i = \bar{y} \cdot \frac{\bar{y}_s}{\bar{y}} \cdot \frac{y_i}{\bar{y}_s}$.

Список использованных источников

- 1 Айвазян, С.А. Прикладная статистика. Основы эконометрики : учеб. для вузов : в 2 т. / С.А. Айвазян, В.С. Мхитарян. – Москва : ЮНИТИ-ДАНА, 2001. – 2 т. – ISBN 5-238-00304-8.
- 2 Афанасьев, В.Н. Анализ временных рядов и прогнозирование : учебник / В.Н. Афанасьев, М.М. Юзбашев. – 2-е изд., перераб. и доп. – Москва : Финансы и статистика; ИНФРА-М, 2010. – 320 с. – ISBN 978-5-279-03400-0 (Финансы и статистика); ISBN 978-5-16-004248-0 (ИНФРА-М).
- 3 Афанасьев, В.Н. Эконометрика: учеб. для вузов / В.Н. Афанасьев, М.М. Юзбашев, Т.И. Гуляева; под ред. В.Н. Афанасьева. – Москва : Финансы и статистика, 2005. – 256 с. – ISBN 5-279-02738-3.
- 4 Бабешко, Л.О. Основы эконометрического моделирования : учеб. пособие / Л.О. Бабешко. – Изд. 2-е, испр. – Москва : КомКнига, 2006. – 432 с. – ISBN 978 – 5- 484-00757-8.
- 5 Еремеева, Н.С. Эконометрика : учеб. пособие для вузов / Н.С. Еремеева, Т.В. Лебедева – Оренбург : ОАО «ИПК «Южный Урал», 2010. – 296 с. – ISBN 978 -5-94162-074-6.
- 6 Кремер, Н.Ш. Эконометрика : учебник для вузов / Н.Ш. Кремер, Б.А. Путко. – Москва : ЮНИТИ-ДАНА, 2003. – 311 с. – ISBN5-238-00333-1.
- 7 Официальный сайт Высшей школы экономики. – Режим доступа: <http://www.hse.ru> .
- 8 Официальный сайт Федеральной службы государственной статистики. – Режим доступа : <http://www.gks.ru>.
- 9 Практикум по эконометрике : учеб. пособие для экон. вузов / под ред. И. И. Елисеевой. – Москва : Финансы и статистика, 2008. – 344 с. – ISBN 978-5-279-02785-9.

10 Тихомиров, Н.П. Эконометрика : учебник / Н.П. Тихомиров, Е.Ю. Дорохина – Москва: Изд-во «Экзамен», 2003. – 512 с. – ISBN 5-94692-438-9.

11 Эконометрика для бакалавров : учебник для студентов, обучающихся по программам высшего профессионального образования по направлению подготовки 080100 Экономика / под ред. В. Н. Афанасьева; Оренбургский гос. ун-т.- 2-е изд., перераб. и доп. – Оренбург : Университет, 2014. – 422 с. – ISBN 978-5-4417-0345-1.

12 Эконометрика : учебник / под ред. д-ра экон. наук, проф. В.С. Мхитаряна. – Москва : Проспект, 2009. – 384 с. – ISBN 978-5-392-00188-0.

13 Эконометрика : учебник / под ред. И.И. Елисеевой. – Москва : Проспект, 2008. – 288 с. – ISBN 978-5-392-00186-6.

14 Эконометрика : учебник / И.И. Елисеева [и др.]. – 2-е изд., перераб. и доп. – Москва : Финансы и статистика, 2007. – 576 с. – ISBN 5-279-02786-3.

Приложение А

(обязательное)

Исходные данные для выполнения лабораторных работ

Таблица А.1 – Вариант 1

Год	Квартал	y	x_1	x_2	x_3	x_4	x_5	Пол руководителя компании
1	1	0,5	4,6	3,8	40,7	19,2	40,0	муж
	2	0,9	7,8	2,7	17,4	18,8	61,7	муж
	3	0,4	3,4	3,7	5,9	17,2	21,0	муж
	4	0,9	5,2	1,0	13,0	16,8	47,2	муж
2	1	1,4	6,2	4,4	26,0	15,6	103,0	муж
	2	0,7	1,2	1,2	19,5	4,8	93,6	жен
	3	2,2	10,2	19,8	178,2	15,3	344,0	муж
	4	0,9	5,3	4,0	23,3	15,1	82,6	муж
3	1	3,7	9,0	12,1	215,0	14,6	742,0	жен
	2	0,2	3,9	0,3	2,6	13,6	1,1	муж
	3	0,7	6,0	1,6	8,8	13,6	23,8	жен
	4	1,0	7,2	3,8	35,2	13,3	39,7	муж
4	1	1,0	7,4	2,6	17,4	4,5	58,8	муж
	2	0,8	5,0	2,5	12,7	11,4	209,0	жен
	3	0,2	6,1	2,4	25,4	11,0	102,0	жен
	4	0,4	0,6	0,6	8,8	10,4	30,5	жен

Y – интегральный показатель экономической безопасности организации;

X_1 – объем реализованной продукции (млн. рублей);

X_2 – оборотный капитал (млн. рублей);

X_3 – основной капитал (млн. рублей);

X_4 – производительность труда (рублей на чел.);

X_5 – объем привлеченных инвестиций (млн. рублей).

Таблица А.2 – Вариант 2

Год	Квартал	y	x_1	x_2	x_3	x_4	x_5	Пол руководителя компании
1	1	0,4	38,1	29,3	16,4	1,1	5,2	муж
	2	0,6	20,2	28,9	24,8	1,9	3,7	муж
	3	0,4	11,3	27,3	9,1	0,8	5,0	муж
	4	0,6	16,8	26,9	19,2	1,2	1,2	муж
2	1	0,7	26,8	25,7	40,7	1,5	6,0	муж
	2	0,5	21,8	14,9	37,1	0,2	1,5	жен
	3	1,0	143,0	25,4	133,4	2,5	27,4	муж
	4	0,5	24,7	25,2	32,8	1,2	5,5	муж
3	1	1,4	172,2	24,7	286,4	2,2	16,7	муж
	2	0,3	8,8	23,7	1,5	0,9	0,3	жен
	3	0,5	13,6	23,7	10,2	1,4	2,1	жен
	4	0,6	33,0	23,4	16,3	1,7	5,2	муж
4	1	0,6	20,2	14,6	23,7	1,8	3,6	жен
	2	0,5	16,6	21,5	81,4	1,2	3,4	жен
	3	0,3	26,3	21,1	40,3	1,4	3,3	жен
	4	0,4	13,6	20,5	12,8	0,1	0,8	жен
5	1	0,6	33,8	18,9	54,5	1,0	3,5	муж
	2	0,4	19,2	18,2	36,8	1,6	1,8	жен
	3	0,5	24,5	13,8	53,7	2,0	4,1	жен
	4	0,6	19,5	17,7	22,7	2,0	3,2	муж
6	1	0,1	28,2	17,6	50,3	1,3	0,3	муж
	2	0,5	20,3	17,6	27,1	1,7	2,3	жен
	3	0,6	20,2	17,5	25,1	1,9	3,1	муж
	4	0,4	11,0	16,3	8,8	1,2	0,4	жен

Y – интегральный показатель экономической безопасности организации;

X_1 – объем реализованной продукции (млн. рублей);

X_2 – оборотный капитал (млн. рублей);

X_3 – основной капитал (млн. рублей);

X_4 – производительность труда (рублей на чел.);

X_5 – объем привлеченных инвестиций (млн. рублей).

Таблица А.3 – Вариант 3

Год	Квартал	y	x_1	x_2	x_3	x_4	x_5	Пол руководителя компании
1	1	1,5	12,5	68,0	34,0	43,3	2,3	муж
	2	1,6	18,8	49,3	16,1	42,9	3,9	жен
	3	1,5	7,0	66,6	7,2	41,3	1,7	муж
	4	1,6	14,6	17,3	12,7	40,9	2,6	муж
2	1	1,6	30,7	78,5	22,7	39,7	3,1	муж
	2	1,5	28,0	20,9	17,7	28,9	0,6	жен
	3	1,6	100,6	356,4	139,8	39,4	5,1	муж
	4	1,6	24,8	72,4	20,6	39,2	2,6	муж
3	1	1,7	216,1	218,2	168,1	38,7	4,5	муж
	2	1,5	1,2	5,0	4,7	37,7	1,9	жен
	3	1,5	7,8	28,8	9,5	37,7	3,0	жен
	4	1,6	12,4	68,0	29,8	37,4	3,6	жен
4	1	1,6	17,9	47,5	16,1	28,6	3,7	муж
	2	1,5	61,5	45,4	12,5	35,5	2,5	жен
	3	1,5	30,5	43,9	22,2	35,1	3,1	жен
	4	1,5	9,7	11,5	9,5	34,5	0,3	жен
5	1	1,6	41,2	46,8	29,7	32,9	2,2	муж
	2	1,5	27,8	24,8	15,1	32,2	3,5	жен
	3	1,5	40,6	54,0	20,4	27,8	4,1	жен
	4	1,6	17,2	42,8	15,4	31,7	4,3	муж

Y – интегральный показатель экономической безопасности организации;

X_1 – объем реализованной продукции (млн. рублей);

X_2 – оборотный капитал (млн. рублей);

X_3 – основной капитал (млн. рублей);

X_4 – производительность труда (рублей на чел.);

X_5 – объем привлеченных инвестиций (млн. рублей).

Таблица А.4 – Вариант 4

Год	Квартал	y	x_1	x_2	x_3	x_4	x_5	Пол руководителя компании
1	1	2,0	3	14,1	12,9	20,3	21,3	муж
	2	1,7	0,8	19,1	4,8	19,5	12,4	муж
	3	2,0	1,7	4,9	10	19,3	17,9	муж
	4	2,1	2,2	22,5	21,2	18,7	27,9	муж
2	1	1,9	0,3	6	19,3	13,3	22,9	жен
	2	2,5	4,2	102,1	69,4	18,5	145	муж
	3	1,9	1,8	20,7	17,1	18,4	25,8	муж
	4	3,1	3,6	62,5	149	18,2	173,3	муж
3	1	1,7	1,1	1,4	0,8	17,7	9,9	жен
	2	1,9	2,1	8,2	5,4	17,7	14,7	жен
	3	2,0	2,7	19,5	8,5	17,5	35	жен
	4	2,0	2,8	13,6	12,4	13,1	21,3	муж
4	1	1,9	1,6	13	42,4	16,6	17,7	жен
	2	1,7	2,2	12,6	21	16,4	27,4	жен
	3	1,8	0,6	3,3	6,7	16,1	14,7	жен
	4	2,0	1,4	13,4	28,4	15,3	34,9	муж

Y – интегральный показатель экономической безопасности организации;

X_1 – объем реализованной продукции (млн. рублей);

X_2 – оборотный капитал (млн. рублей);

X_3 – основной капитал (млн. рублей);

X_4 – производительность труда (рублей на чел.);

X_5 – объем привлеченных инвестиций (млн. рублей).

Таблица А.5 – Вариант 5

Год	Квартал	y	x_1	x_2	x_3	x_4	x_5	Пол руководителя компании
1	1	1,7	13,7	40,5	64,7	13,4	3,9	муж
	2	0,7	18,5	38,9	24	4,5	1,7	муж
	3	1,7	4,8	38,5	50,2	10	2,6	муж
	4	2,6	21,8	37,3	106	20	3,1	муж
2	1	1,3	5,8	26,5	96,6	15	0,6	жен
	2	4,1	99	37	347	137,1	5,1	муж
	3	1,6	20,1	36,8	85,6	17,9	2,7	муж
	4	6,9	60,6	36,3	745	165,4	4,5	муж
3	1	0,4	1,4	35,3	4,1	2	2	жен
	2	1,3	8	35,3	26,8	6,8	3	жен
	3	1,9	18,9	35	42,7	27,1	3,6	муж
	4	1,9	13,2	26,2	61,8	13,4	3,7	жен
4	1	1,4	12,6	33,1	212	9,8	2,5	жен
	2	0,4	12,2	32,7	105	19,5	3,1	жен
	3	0,8	3,2	32,1	33,5	6,8	0,3	жен
	4	1,8	13	30,5	142	27	2,3	муж
5	1	0,9	6,9	29,8	96	12,4	3,5	жен
	2	1,1	15	25,4	140	17,7	4,1	жен
	3	1,9	11,9	29,3	59,3	12,7	4,3	муж
	4	0,9	1,6	29,2	131	21,4	2,9	муж

Y – интегральный показатель экономической безопасности организации;

X_1 – объем реализованной продукции (млн. рублей);

X_2 – оборотный капитал (млн. рублей);

X_3 – основной капитал (млн. рублей);

X_4 – производительность труда (рублей на чел.);

X_5 – объем привлеченных инвестиций (млн. рублей).

Таблица А.6 – Вариант 6

Год	Квартал	y	x_1	x_2	x_3	x_4	x_5	Пол руководителя компании
1	1	0,99	18,2	27,4	3,4	6,8	9,2	муж
	2	0,84	9,3	37	1,2	6,5	3,4	муж
	3	0,99	14,8	9,6	2,1	6,4	7,2	муж
	4	1,11	24,8	43,6	2,6	6,2	15,1	муж
2	1	0,93	19,8	11,6	0,1	4,4	13,8	жен
	2	1,33	141,9	198	4,6	6,2	49,6	муж
	3	0,97	22,7	40,2	2,1	6,1	12,2	муж
	4	1,73	170,2	121,2	4	6,1	106,4	муж
3	1	0,80	6,8	2,8	1,4	5,9	0,6	жен
	2	0,93	11,6	16	2,5	5,9	3,8	жен
	3	1,01	31,9	37,8	3	5,8	6,1	муж
	4	1,01	18,2	26,4	3,2	4,4	8,8	жен
4	1	0,94	14,6	25,2	2	5,5	30,3	жен
	2	0,80	24,3	24,4	2,5	5,5	15	жен
	3	0,86	11,6	6,4	0,2	5,4	4,8	жен
	4	1,00	31,8	26	1,7	5,1	20,3	муж
5	1	0,87	17,2	13,8	2,9	5	13,7	жен
	2	0,90	22,5	30	3,6	4,2	20	жен
	3	1,01	17,5	23,8	3,7	4,9	8,5	муж
	4	0,61	26,2	3,2	2,4	4,9	18,7	муж
6	1	0,93	18,3	17,2	3	4,9	10,1	муж
	2	1,03	18,2	23	3,5	4,9	9,3	муж
	3	0,83	9	3,8	2,1	4,7	3,3	жен
	4	0,84	20,3	11,6	3,5	4,5	11,5	жен

Y – интегральный показатель экономической безопасности организации;

X_1 – объем реализованной продукции (млн. рублей);

X_2 – оборотный капитал (млн. рублей);

X_3 – основной капитал (млн. рублей);

X_4 – производительность труда (рублей на чел.);

X_5 – объем привлеченных инвестиций (млн. рублей).

Таблица А.7 – Вариант 7

Год	Квартал	y	x_1	x_2	x_3	x_4	x_5	Пол руководителя компании
1	1	0,6	5,5	4,6	48,8	23,0	48,0	муж
	2	1,1	9,4	3,2	20,9	22,6	74,0	муж
	3	0,4	4,1	4,4	7,1	20,6	25,2	муж
	4	1,1	6,2	1,2	15,6	20,2	56,6	муж
2	1	1,7	7,4	5,3	31,2	18,7	123,6	муж
	2	0,8	1,4	1,4	23,4	5,8	112,3	жен
	3	2,6	12,2	23,8	213,8	18,4	412,8	муж
	4	1,0	6,4	4,8	28,0	18,1	99,1	муж
3	1	4,4	10,8	14,5	258,0	17,5	890,4	жен
	2	0,3	4,7	0,4	3,1	16,3	1,3	муж
	3	0,8	7,2	1,9	10,6	16,3	28,6	жен
	4	1,2	8,6	4,6	42,2	16,0	47,6	муж
4	1	1,2	8,9	3,1	20,9	5,4	70,6	муж
	2	0,9	6,0	3,0	15,2	13,7	250,8	жен
	3	0,3	7,3	2,9	30,5	13,2	122,4	жен
	4	0,5	0,7	0,7	10,6	12,5	36,6	жен

Y – интегральный показатель экономической безопасности организации;

X_1 – объем реализованной продукции (млн. рублей);

X_2 – оборотный капитал (млн. рублей);

X_3 – основной капитал (млн. рублей);

X_4 – производительность труда (рублей на чел.);

X_5 – объем привлеченных инвестиций (млн. рублей).

Таблица А.8 – Вариант 8

Год	Квартал	y	x_1	x_2	x_3	x_4	x_5	Пол руководителя компании
1	1	0,3	31,8	24,4	13,7	0,9	4,3	муж
	2	0,5	16,8	24,1	20,7	1,6	3,1	муж
	3	0,3	9,4	22,8	7,6	0,7	4,2	муж
	4	0,5	14,0	22,4	16,0	1,0	1,0	муж
2	1	0,6	22,3	21,4	33,9	1,3	5,0	муж
	2	0,4	18,2	12,4	30,9	0,2	1,3	жен
	3	0,8	119,2	21,2	111,2	2,1	22,8	муж
	4	0,4	20,6	21,0	27,3	1,0	4,6	муж
3	1	1,2	143,5	20,6	238,7	1,8	13,9	муж
	2	0,3	7,3	19,8	1,3	0,8	0,3	жен
	3	0,4	11,3	19,8	8,5	1,2	1,8	жен
	4	0,5	27,5	19,5	13,6	1,4	4,3	муж
4	1	0,5	16,8	12,2	19,8	1,5	3,0	жен
	2	0,4	13,8	17,9	67,8	1,0	2,8	жен
	3	0,3	21,9	17,6	33,6	1,2	2,8	жен
	4	0,3	11,3	17,1	10,7	0,1	0,7	жен
5	1	0,5	28,2	15,8	45,4	0,8	2,9	муж
	2	0,3	16,0	15,2	30,7	1,3	1,5	жен
	3	0,4	20,4	11,5	44,8	1,7	3,4	жен
	4	0,5	16,3	14,8	18,9	1,7	2,7	муж
6	1	0,1	23,5	14,7	41,9	1,1	0,3	муж
	2	0,4	16,9	14,7	22,6	1,4	1,9	жен
	3	0,5	16,8	14,6	20,9	1,6	2,6	муж
	4	0,3	9,2	13,6	7,3	1,0	0,3	жен

Y – интегральный показатель экономической безопасности организации;

X_1 – объем реализованной продукции (млн. рублей);

X_2 – оборотный капитал (млн. рублей);

X_3 – основной капитал (млн. рублей);

X_4 – производительность труда (рублей на чел.);

X_5 – объем привлеченных инвестиций (млн. рублей).

Таблица А.9 – Вариант 9

Год	Квартал	y	x_1	x_2	x_3	x_4	x_5	Пол руководителя компании
1	1	1,9	1,8	9,7	4,9	6,2	0,3	муж
	2	1,8	2,7	7,0	2,3	6,1	0,6	жен
	3	1,7	1,0	9,5	1,0	5,9	0,2	муж
	4	1,8	2,1	2,5	1,8	5,8	0,4	муж
2	1	1,8	4,4	11,2	3,2	5,7	0,4	муж
	2	1,7	4,0	3,0	2,5	4,1	0,1	жен
	3	1,8	14,4	50,9	20,0	5,6	0,7	муж
	4	1,7	3,5	10,3	2,9	5,6	0,4	муж
3	1	1,9	30,9	31,2	24,0	5,5	0,6	муж
	2	1,7	0,2	0,7	0,7	5,4	0,3	жен
	3	1,7	1,1	4,1	1,4	5,4	0,4	жен
	4	1,8	1,8	9,7	4,3	5,3	0,5	жен
4	1	1,8	2,6	6,8	2,3	4,1	0,5	муж
	2	1,7	8,8	6,5	1,8	5,1	0,4	жен
	3	1,7	4,4	6,3	3,2	5,0	0,4	жен
	4	1,7	1,4	1,6	1,4	4,9	0,0	жен
5	1	1,8	5,9	6,7	4,2	4,7	0,3	муж
	2	1,7	4,0	3,5	2,2	4,6	0,5	жен
	3	1,7	5,8	7,7	2,9	4,0	0,6	жен
	4	1,8	2,5	6,1	2,2	4,5	0,6	муж

Y – интегральный показатель экономической безопасности организации;

X_1 – объем реализованной продукции (млн. рублей);

X_2 – оборотный капитал (млн. рублей);

X_3 – основной капитал (млн. рублей);

X_4 – производительность труда (рублей на чел.);

X_5 – объем привлеченных инвестиций (млн. рублей).

Таблица А.10 – Вариант 10

Год	Квартал	y	x_1	x_2	x_3	x_4	x_5	Пол руководителя компании
1	1	0,8	1,2	5,6	5,2	8,1	8,5	муж
	2	0,7	0,3	7,6	1,9	7,8	5,0	муж
	3	0,8	0,7	2,0	4,0	7,7	7,2	муж
	4	0,9	0,9	9,0	8,5	7,5	11,2	муж
2	1	0,7	0,1	2,4	7,7	5,3	9,2	жен
	2	1,0	1,7	40,8	27,8	7,4	58,0	муж
	3	0,8	0,7	8,3	6,8	7,4	10,3	муж
	4	1,2	1,4	25,0	59,6	7,3	69,3	муж
3	1	0,7	0,4	0,6	0,3	7,1	4,0	жен
	2	0,7	0,8	3,3	2,2	7,1	5,9	жен
	3	0,8	1,1	7,8	3,4	7,0	14,0	жен
	4	0,8	1,1	5,4	5,0	5,2	8,5	муж
4	1	0,8	0,6	5,2	17,0	6,6	7,1	жен
	2	0,7	0,9	5,0	8,4	6,6	11,0	жен
	3	0,7	0,2	1,3	2,7	6,4	5,9	жен
	4	0,8	0,6	5,4	11,4	6,1	14,0	муж

Y – интегральный показатель экономической безопасности организации;

X_1 – объем реализованной продукции (млн. рублей);

X_2 – оборотный капитал (млн. рублей);

X_3 – основной капитал (млн. рублей);

X_4 – производительность труда (рублей на чел.);

X_5 – объем привлеченных инвестиций (млн. рублей).

Таблица А.11 – Вариант 11

Год	Квартал	y	x_1	x_2	x_3	x_4	x_5	Пол руководителя компании
1	1	1,6	13,0	38,5	61,5	12,7	3,7	муж
	2	0,7	17,6	37,0	22,8	4,3	1,6	муж
	3	1,6	4,6	36,6	47,7	9,5	2,5	муж
	4	2,5	20,7	35,4	100,7	19,0	2,9	муж
2	1	1,2	5,5	25,2	91,8	14,3	0,6	жен
	2	3,9	94,1	35,2	329,7	130,2	4,8	муж
	3	1,5	19,1	35,0	81,3	17,0	2,6	муж
	4	2,6	57,6	34,5	707,8	157,1	4,3	муж
3	1	0,4	1,3	33,5	3,9	1,9	1,9	жен
	2	1,2	7,6	33,5	25,5	6,5	2,9	жен
	3	1,8	18,0	33,3	40,6	25,7	3,4	муж
	4	1,8	12,5	24,9	58,7	12,7	3,5	жен
4	1	1,3	12,0	31,4	201,4	9,3	2,4	жен
	2	0,4	11,6	31,1	99,8	18,5	2,9	жен
	3	0,8	3,0	30,5	31,8	6,5	0,3	жен
	4	1,7	12,4	29,0	134,9	25,7	2,2	муж
5	1	0,9	6,6	28,3	91,2	11,8	3,3	жен
	2	1,0	14,3	24,1	133,0	16,8	3,9	жен
	3	1,8	11,3	27,8	56,3	12,1	4,1	муж
	4	0,9	1,5	27,7	124,5	20,3	2,8	муж

Y – интегральный показатель экономической безопасности организации;

X_1 – объем реализованной продукции (млн. рублей);

X_2 – оборотный капитал (млн. рублей);

X_3 – основной капитал (млн. рублей);

X_4 – производительность труда (рублей на чел.);

X_5 – объем привлеченных инвестиций (млн. рублей).

Таблица А.12 – Вариант 12

Год	Квартал	y	x_1	x_2	x_3	x_4	x_5	Пол руководителя компании
1	1	0,12	2,28	3,43	0,43	0,85	1,15	муж
	2	0,11	1,16	4,63	0,15	0,81	0,43	муж
	3	0,12	1,85	1,20	0,26	0,80	0,90	муж
	4	0,14	3,10	5,45	0,33	0,78	1,89	муж
2	1	0,12	2,48	1,45	0,01	0,55	1,73	жен
	2	0,17	17,74	24,75	0,58	0,78	6,20	муж
	3	0,12	2,84	5,03	0,26	0,76	1,53	муж
	4	0,22	21,28	15,15	0,50	0,76	13,30	муж
3	1	0,10	0,85	0,35	0,18	0,74	0,08	жен
	2	0,12	1,45	2,00	0,31	0,74	0,48	жен
	3	0,13	3,99	4,73	0,38	0,73	0,76	муж
	4	0,13	2,28	3,30	0,40	0,55	1,10	жен
4	1	0,12	1,83	3,15	0,25	0,69	3,79	жен
	2	0,10	3,04	3,05	0,31	0,69	1,88	жен
	3	0,11	1,45	0,80	0,03	0,68	0,60	жен
	4	0,13	3,98	3,25	0,21	0,64	2,54	муж
5	1	0,11	2,15	1,73	0,36	0,63	1,71	жен
	2	0,11	2,81	3,75	0,45	0,53	2,50	жен
	3	0,13	2,19	2,98	0,46	0,61	1,06	муж
	4	0,08	3,28	0,40	0,30	0,61	2,34	муж
6	1	0,12	2,29	2,15	0,38	0,61	1,26	муж
	2	0,13	2,28	2,88	0,44	0,61	1,16	муж
	3	0,10	1,13	0,48	0,26	0,59	0,41	жен
	4	0,11	2,54	1,45	0,44	0,56	1,44	жен

Y – интегральный показатель экономической безопасности организации;

X_1 – объем реализованной продукции (млн. рублей);

X_2 – оборотный капитал (млн. рублей);

X_3 – основной капитал (млн. рублей);

X_4 – производительность труда (рублей на чел.);

X_5 – объем привлеченных инвестиций (млн. рублей).

Таблица А.13 – Вариант 13

Год	Квартал	y	x_1	x_2	x_3	x_4	x_5	Пол руководителя компании
1	1	0,41	3,9	3,2	34,2	16,1	33,6	муж
	2	0,76	6,6	2,3	14,6	15,8	51,9	муж
	3	0,31	2,9	3,1	5,0	14,5	17,7	муж
	4	0,76	4,4	0,8	10,9	14,1	39,7	муж
2	1	1,16	5,2	3,7	21,9	13,1	86,6	муж
	2	0,59	1,0	1,0	16,4	4,0	78,7	жен
	3	1,84	8,6	16,7	149,9	12,9	289,3	муж
	4	0,71	4,5	3,4	19,6	12,7	69,5	муж
3	1	3,10	7,6	10,2	180,8	12,3	624,1	жен
	2	0,18	3,3	0,3	2,2	11,4	0,9	муж
	3	0,59	5,0	1,3	7,4	11,4	20,0	жен
	4	0,86	6,1	3,2	29,6	11,2	33,4	муж
4	1	0,86	6,2	2,2	14,6	3,8	49,5	муж
	2	0,63	4,2	2,1	10,7	9,6	175,8	жен
	3	0,18	5,1	2,0	21,4	9,3	85,8	жен
	4	0,36	0,5	0,5	7,4	8,7	25,7	жен

Y – интегральный показатель экономической безопасности организации;

X_1 – объем реализованной продукции (млн. рублей);

X_2 – оборотный капитал (млн. рублей);

X_3 – основной капитал (млн. рублей);

X_4 – производительность труда (рублей на чел.);

X_5 – объем привлеченных инвестиций (млн. рублей).

Таблица А.14 – Вариант 14

Год	Квартал	y	x_1	x_2	x_3	x_4	x_5	Пол руководителя компании
1	1	0,35	32,2	24,7	13,8	0,9	4,4	муж
	2	0,46	17,0	24,4	20,9	1,6	3,1	муж
	3	0,32	9,5	23,0	7,7	0,7	4,2	муж
	4	0,46	14,2	22,7	16,2	1,0	1,0	муж
2	1	0,59	22,6	21,7	34,3	1,3	5,1	муж
	2	0,41	18,4	12,6	31,3	0,2	1,3	жен
	3	0,80	120,7	21,4	112,6	2,1	23,1	муж
	4	0,45	20,8	21,3	27,7	1,0	4,6	муж
3	1	1,20	145,3	20,8	241,7	1,9	14,1	муж
	2	0,28	7,4	20,0	1,3	0,8	0,3	жен
	3	0,41	11,5	20,0	8,6	1,2	1,8	жен
	4	0,49	27,8	19,7	13,8	1,4	4,4	муж
4	1	0,49	17,0	12,3	20,0	1,5	3,0	жен
	2	0,42	14,0	18,1	68,7	1,0	2,9	жен
	3	0,28	22,2	17,8	34,0	1,2	2,8	жен
	4	0,34	11,5	17,3	10,8	0,1	0,7	жен
5	1	0,48	28,5	15,9	46,0	0,8	3,0	муж
	2	0,35	16,2	15,4	31,1	1,4	1,5	жен
	3	0,38	20,7	11,6	45,3	1,7	3,5	жен
	4	0,49	16,5	14,9	19,2	1,7	2,7	муж
6	1	0,10	23,8	14,9	42,4	1,1	0,3	муж
	2	0,41	17,1	14,9	22,9	1,4	1,9	жен
	3	0,51	17,0	14,8	21,2	1,6	2,6	муж
	4	0,31	9,3	13,8	7,4	1,0	0,3	жен

Y – интегральный показатель экономической безопасности организации;

X_1 – объем реализованной продукции (млн. рублей);

X_2 – оборотный капитал (млн. рублей);

X_3 – основной капитал (млн. рублей);

X_4 – производительность труда (рублей на чел.);

X_5 – объем привлеченных инвестиций (млн. рублей).

Таблица А.15 – Вариант 15

Год	Квартал	y	x_1	x_2	x_3	x_4	x_5	Пол руководителя компании
1	1	0,86	7,0	38,1	19,0	24,3	1,3	муж
	2	0,87	10,5	27,6	9,0	24,0	2,2	жен
	3	0,85	3,9	37,3	4,0	23,1	1,0	муж
	4	0,87	8,2	9,7	7,1	22,9	1,5	муж
2	1	0,89	17,2	44,0	12,7	22,2	1,7	муж
	2	0,86	15,7	11,7	9,9	16,2	0,3	жен
	3	0,92	56,4	199,7	78,3	22,1	2,9	муж
	4	0,87	13,9	40,6	11,5	22,0	1,5	муж
3	1	0,97	121,1	122,2	94,2	21,7	2,5	муж
	2	0,85	0,7	2,8	2,6	21,1	1,1	жен
	3	0,86	4,4	16,1	5,3	21,1	1,7	жен
	4	0,88	6,9	38,1	16,7	21,0	2,0	жен
4	1	0,88	10,0	26,6	9,0	16,0	2,1	муж
	2	0,87	34,5	25,4	7,0	19,9	1,4	жен
	3	0,85	17,1	24,6	12,4	19,7	1,7	жен
	4	0,86	5,4	6,4	5,3	19,3	0,2	жен
5	1	0,87	23,1	26,2	16,6	18,4	1,2	муж
	2	0,86	15,6	13,9	8,5	18,0	2,0	жен
	3	0,86	22,7	30,3	11,4	15,6	2,3	жен
	4	0,88	9,6	24,0	8,6	17,8	2,4	муж

Y – интегральный показатель экономической безопасности организации;

X_1 – объем реализованной продукции (млн. рублей);

X_2 – оборотный капитал (млн. рублей);

X_3 – основной капитал (млн. рублей);

X_4 – производительность труда (рублей на чел.);

X_5 – объем привлеченных инвестиций (млн. рублей).

Приложение Б

(обязательное)

Статистико-математические таблицы

Таблица Б.1 – Критические значения t-критерия Стьюдента при уровне значимости 0,10; 0,05; 0,01 (двухсторонний)

Число степеней свободы	α			Число степеней свободы	α		
	0,10	0,05	0,01		0,10	0,05	0,01
1	6,3138	12,706	63,657	18	1,7341	2,1009	2,8784
2	2,9200	4,3027	9,9248	19	1,7291	2,0930	2,8609
3	2,3534	3,1825	5,8409	20	1,7247	2,0860	2,8453
4	2,1318	2,7764	4,6041	21	1,7207	2,0796	2,8314
5	2,0150	2,5706	4,0321	22	1,7171	2,0739	2,8188
6	1,9432	2,4469	3,7074	23	1,7139	2,0687	2,8073
7	1,8946	2,3646	3,495	24	1,7109	2,0639	2,7969
8	1,8595	2,3060	3,3554	25	1,7081	2,0595	2,7874
9	1,8331	2,2622	3,2498	26	1,7056	2,0555	2,7787
10	1,8125	2,2281	3,1693	27	1,7033	2,0518	2,7707
11	1,7959	2,2010	3,1058	28	1,7011	2,0484	2,7633
12	1,7823	2,1788	3,0545	29	1,6991	2,0452	2,7564
13	1,7709	2,1604	3,0123	30	1,6973	2,0423	2,7500
14	1,7613	2,1448	2,9768	40	1,6839	2,0211	2,7045
15	1,7530	2,1315	2,9467	60	1,6707	2,0003	2,6603
16	1,7459	2,1199	2,9208	120	1,6577	1,9799	2,6174
17	1,7396	2,1098	2,8982	∞	1,6449	1,9600	2,5758

Таблица Б.2 – Критические значения F-критерия Фишера ($\alpha = 0,05$)

$k_2=n-m$	$k_1=m-1$									
	1	2	3	4	5	6	8	12	24	∞
1	161,45	199,5	215,72	224,57	230,17	233,97	238,89	243,91	249,04	254,32
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	8,84	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	3,88	3,49	3,26	3,11	3,00	2,85	2,69	2,50	2,30
13	4,67	3,80	3,41	3,18	3,02	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

Продолжение таблицы Б.2

$k_2=n-m$	$k_1=m-1$									
	1	2	3	4	5	6	8	12	24	∞
23	4,28	3,42	3,03	2,80	2,64	2,53	2,38	2,20	2,00	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	3,38	2,99	2,76	2,60	2,49	2,34	2,16	1,96	1,71
26	4,22	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	2,30	2,13	1,93	1,67
28	4,20	3,34	2,95	2,71	2,56	2,44	2,29	2,12	1,91	1,65
29	4,18	3,33	2,93	2,70	2,54	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	3,26	2,87	2,64	2,48	2,37	2,22	2,04	1,83	1,57
40	4,08	3,23	2,84	2,61	2,45	2,34	2,18	2,00	1,79	1,51
45	4,06	3,21	2,81	2,58	2,42	2,31	2,15	1,97	1,76	1,48
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	2,52	2,37	2,25	2,10	1,92	1,70	1,39
70	3,89	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	1,31
90	3,95	3,10	2,71	2,47	2,32	2,20	2,04	1,86	1,64	1,28
100	3,94	3,09	2,70	2,46	2,30	2,19	2,03	1,85	1,63	1,26
125	3,92	3,07	2,68	2,44	2,29	2,17	2,01	1,83	1,60	1,21
150	3,90	3,06	2,66	2,43	2,27	2,16	2,00	1,82	1,59	1,18
200	3,89	3,04	2,65	2,42	2,26	2,14	1,98	1,80	1,57	1,14
300	3,87	3,03	2,64	2,41	2,25	2,13	1,97	1,79	1,55	1,10
400	3,86	3,02	2,63	2,40	2,24	2,12	1,96	1,78	1,54	1,07
500	3,86	3,01	2,62	2,39	2,23	2,11	1,96	1,77	1,54	1,06
1000	3,85	3,00	2,61	2,38	2,22	2,10	1,95	1,76	1,53	1,03
∞	3,84	2,99	2,60	2,37	2,21	2,09	1,94	1,75	1,52	1,00

Таблица Б.3 - Значения критерия Дарбина-Уотсона для положительной автокорреляции при 5%-ном уровне значимости

n	$p=1$		$p=2$		$p=3$		$p=4$		$p=5$	
	d_L	d_U								
15	1,08	1,36	0,95	1,54	0,82	1,75	0,69	1,97	0,56	2,21
16	1,1	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	0,9	1,71	0,78	1,9	0,67	2,1
18	1,16	1,39	1,05	1,53	0,93	1,69	0,82	1,87	0,71	2,06
19	1,18	1,4	1,08	1,53	0,97	1,68	0,86	1,85	0,75	2,02
20	1,2	1,41	1,10	1,54	1	1,68	0,9	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	1,8	0,86	1,94
23	1,26	1,44	1,17	1,54	1,08	1,66	0,99	1,79	0,9	1,92
24	1,27	1,45	1,19	1,55	1,1	1,66	1,01	1,78	0,93	1,9
25	1,29	1,45	1,21	1,55	1,12	1,66	1,04	1,77	0,95	1,89
26	1,3	1,46	1,22	1,55	1,14	1,65	1,06	1,76	0,98	1,89
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	1,1	1,75	1,03	1,85
29	1,34	1,48	1,27	1,56	1,2	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	1,3	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	1,63	1,19	1,73	1,13	1,81
34	1,39	1,51	1,33	1,58	1,27	1,65	1,21	1,73	1,15	1,81
35	1,40	1,52	1,34	1,58	1,28	1,65	1,22	1,73	1,16	1,8
36	1,41	1,52	1,35	1,59	1,29	1,65	1,24	1,73	1,18	1,8
37	1,42	1,53	1,36	1,59	1,31	1,66	1,25	1,72	1,19	1,8
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	1,29	1,72	1,23	1,79
45	1,48	1,57	1,43	1,62	1,38	1,67	1,34	1,72	1,29	1,78
50	1,5	1,59	1,46	1,63	1,42	1,67	1,38	1,72	1,34	1,77
55	1,53	1,6	1,49	1,64	1,45	1,68	1,41	1,72	1,38	1,77
60	1,55	1,62	1,51	1,65	1,48	1,69	1,44	1,73	1,41	1,77
65	1,57	1,63	1,54	1,66	1,5	1,70	1,47	1,73	1,44	1,77
70	1,58	1,64	1,55	1,67	1,52	1,70	1,49	1,74	1,46	1,77
75	1,60	1,65	1,57	1,68	1,54	1,71	1,51	1,74	1,49	1,77
80	1,61	1,66	1,59	1,69	1,56	1,72	1,53	1,74	1,51	1,77
85	1,62	1,67	1,6	1,70	1,57	1,72	1,55	1,75	1,52	1,77
90	1,63	1,68	1,61	1,70	1,59	1,73	1,57	1,75	1,54	1,78
95	1,64	1,69	1,62	1,71	1,6	1,73	1,58	1,75	1,56	1,78
100	1,65	1,69	1,63	1,72	1,61	1,74	1,59	1,76	1,57	1,78

Таблица Б.4 – Квантили распределения $\chi^2(\nu)$

ν	$1-\alpha$									
	0,005	0,010	0,025	0,050	0,100	0,900	0,950	0,975	0,990	0,995
1	0,000039	0,00016	0,00098	0,0039	0,0158	2,71	3,84	5,02	6,63	7,88
2	0,0100	0,0201	0,0506	0,1026	0,2107	4,61	5,99	7,38	9,21	10,60
3	0,0717	0,115	0,216	0,352	0,584	6,25	7,81	9,35	11,34	12,84
4	0,207	0,297	0,484	0,711	1,064	7,78	9,49	11,14	13,28	14,86
5	0,412	0,554	0,831	1,15	1,61	9,24	11,07	12,83	15,09	16,75
6	0,676	0,872	1,24	1,64	2,20	10,64	12,59	14,45	16,81	18,55
7	0,989	1,24	1,69	2,17	2,83	12,02	14,07	16,01	18,48	20,28
8	1,34	1,65	2,18	2,73	3,49	13,36	15,51	17,53	20,09	21,96
9	1,73	2,09	2,70	3,33	4,17	14,68	16,92	19,02	21,67	23,59
10	2,16	2,56	3,25	3,94	4,87	15,99	18,31	20,48	23,21	25,19
11	0,60	3,05	3,82	4,57	5,58	17,28	19,68	21,92	24,73	26,76
12	3,07	3,57	4,40	5,23	6,30	18,55	21,03	23,34	26,22	28,30
13	3,57	4,11	5,01	5,89	7,04	19,81	22,36	24,74	27,69	29,82
14	4,07	4,66	5,63	6,57	7,79	21,06	23,68	26,12	29,14	31,32
15	4,60	5,23	6,26	7,26	8,55	22,31	25,00	27,49	30,58	32,80
16	5,14	5,81	6,91	7,96	9,31	23,54	26,30	28,85	32,00	34,27
18	6,26	7,01	8,23	9,39	10,86	25,99	28,87	31,53	34,81	37,16
20	7,43	8,26	9,59	10,85	12,44	28,41	31,41	34,17	37,57	40,00
24	9,89	10,86	12,40	13,85	15,66	33,20	36,42	39,36	42,98	45,56
30	13,79	14,95	16,79	18,49	20,60	40,26	43,77	46,98	50,89	63,67
40	20,71	22,16	24,43	26,51	29,05	51,81	55,76	59,34	63,69	66,77
60	35,53	37,48	40,48	43,19	46,46	74,40	79,08	83,30	88,38	91,95
80	51,17	53,54	57,15	60,39	64,28	96,58	101,88	106,6	112,3	116,3
100	67,33	70,06	74,22	77,93	82,36	118,50	124,34	129,6	135,8	140,2
120	83,85	86,92	91,58	95,70	100,62	140,2	146,57	152,2	159,0	163,6