

**МИНИСТЕРСТВО ОБРАЗОВАНИЯ И НАУКИ  
РОССИЙСКОЙ ФЕДЕРАЦИИ**  
Государственное образовательное учреждение  
высшего профессионального образования  
**ВСЕРОССИЙСКИЙ ЗАОЧНЫЙ  
ФИНАНСОВО-ЭКОНОМИЧЕСКИЙ ИНСТИТУТ**

**ЭКОНОМЕТРИКА**

Компьютерный практикум  
для студентов третьего курса,  
обучающихся по специальностям  
080105.65 «Финансы и кредит»,  
080109.65 «Бухгалтерский учет, анализ и аудит»

**Финансово-кредитный факультет**  
**Кафедра экономико-математических методов и моделей**

**Москва 2011**

ББК 65.61

**Компьютерный практикум разработали:**

кандидат экономических наук, профессор **И.В. Орлова**,  
кандидат физико-математических наук, доцент **Е.С. Филонова**,  
кандидат экономических наук, доцент **А.В. Агеев**

Компьютерный практикум обсужден на заседании кафедры  
экономико-математических методов и моделей  
Зав. кафедрой доктор физико-математических наук,  
профессор **В.В. Угрозов**

Учебно-методическое издание одобрено  
на заседании Научно-методического совета ВЗФЭИ

Проректор, председатель НМС, профессор **А.В. Rogozёнков**

**Эконометрика.** Компьютерный практикум для студентов третьего  
курса, обучающихся по специальностям 080105.65 «Финансы и кредит»,  
080109.65 «Бухгалтерский учет, анализ и аудит». – М.: ВЗФЭИ, 2011.

ББК 65.61

© Всероссийский заочный  
финансово-экономический  
институт (ВЗФЭИ), 2011

## **Введение**

Основной целью изучения дисциплины «Эконометрика» является формирование у студентов теоретических знаний по методологии и методике построения эконометрических моделей и навыков их применения для анализа состояния и оценки закономерностей развития экономических и социальных систем.

В соответствии с учебным планом по дисциплине «Эконометрика» каждый студент должен выполнить две зачетные работы – контрольную и лабораторную.

В данном практикуме представлены методические рекомендации по вопросам, которые часто возникают у студентов при подготовке и оформлении контрольной и лабораторной работ по дисциплине «Эконометрика». В частности, на конкретных примерах рассмотрены некоторые особенности проведения корреляционно-регрессионного анализа экономических данных.

Рекомендации компьютерного практикума (особенно по оформлению лабораторной работы) не являются строго обязательными. Последующая оценка преподавателем результатов практической работы студента основывается на проверке его теоретических знаний и практических навыков по дисциплине. В то же время рекомендуемые приемы выполнения и оформления работ позволяют студенту достичь некоторого необходимого (начального!) уровня квалификации в области эконометрических исследований, а также в использовании математических и компьютерных средств анализа данных.

Основой подготовки студента к выполнению контрольной и лабораторной работ по эконометрике служит учебное пособие [4]. В нем рассматриваются основные этапы построения и анализа различных регрессий, а также даются описания всех расчетов и статистических тестов, выполнение которых необходимо при решении эконометрических задач.

## **1. Общие рекомендации по выполнению и оформлению зачетных работ**

### **1.1. Требования к оформлению контрольной работы**

Номер варианта контрольной и лабораторной работы, как правило, соответствует порядковому номеру фамилии студента в журнале посещаемости занятий, если преподавателем не установлен другой порядок выбора варианта.

*Контрольную работу* все студенты вечерних и дневных групп выполняют дома по приведенным в данном пособии вариантам и направляют в институт для проверки в сроки, указанные индивидуальным графиком студента. Однако эти сроки являются крайними. Чтобы работа была своевременно проверена, а при необходимости доработана и сдана повторно, ее надлежит сдать на проверку раньше указанного срока.

Студентам дневных групп рекомендуется во время установочной (зимней экзаменационной) сессии выполнить большую часть домашней контрольной работы, чтобы получить своевременную консультацию по вопросам, возникшим в процессе ее выполнения. В течение двух недель после окончания сессии контрольная работа должна быть завершена и представлена на проверку.

Решение задач контрольной работы должно сопровождаться необходимыми расчетами и комментариями, то есть все основные моменты процесса решения задачи должны быть раскрыты и обоснованы соответствующими теоретическими положениями.

Для решения задач рекомендуется использовать средства MS Excel (надстройка **Анализ данных**), а также встроенные статистические и математические функции.

Титульный лист контрольной работы должен содержать все необходимые реквизиты: названия института, факультета и специ-

альности, наименование учебной дисциплины с указанием вида отчетности (контрольная или лабораторная работа), номер курса, группы и номер зачетной книжки, ф.и.о. студента и преподавателя.

Работа без указания номера группы и номера зачетной книжки проверке не подлежит. При отсутствии ф.и.о. преподавателя установленные сроки проверки работы могут быть изменены.

К собеседованию допускаются студенты, выполнившие правильно и в полном объеме все задания контрольной работы.

*Контрольная работа не засчитывается, если ее вариант не совпадает с номером варианта, указанным преподавателем, или она выполнена по вариантам прошлых лет.*

*Цель лабораторной работы по эконометрике – закрепить у студентов теоретические основы и практические навыки проведения эконометрических исследований. Лабораторная работа выполняется студентами в компьютерном зале с использованием табличного процессора Microsoft Excel или SPSS<sup>1</sup> под руководством преподавателя.*

При решении задач средствами Microsoft Excel могут использоваться разнообразные с точки зрения содержательности, наглядности, удобства и дизайна подходы к оформлению таблиц и результатов решения.

## **1.2. Требования к оформлению отчета по лабораторной работе**

Оформление отчета по лабораторной работе осуществляется студентом самостоятельно во время или после занятий в установленные преподавателем сроки. Файл (книга Excel) с префиксом «лр» сохраняется на сетевом диске в папке с номером группы. Структура имени файла: **лрИвановИИ**.

Полный отчет по лабораторной работе должен содержать:

- 1) титульный лист (с указанием всех необходимых реквизитов);
- 2) постановку экономической задачи (задач);
- 3) описание компьютерной информационной технологии, с помощью которой получено решение;
- 4) содержательный отчет по каждому пункту задания.

---

<sup>1</sup> Используется при наличии данного программного продукта.

Обязательно соблюдение нумерации пунктов задания. Отчет должен содержать команды MS Excel (необходимо выделить их полужирным шрифтом), применяемые для получения результатов каждого пункта. Каждый пункт исследования необходимо завершать содержательной экономической интерпретацией полученных результатов;

5) при построении графиков и гистограмм подписываются оси (в том числе указываются единицы измерения) и числовые метки. Графики, помимо этого, должны содержать титульные подписи, чтобы было понятно, что на них изображено. При наличии выбросов или нескольких групп, явно видимых на графике, проводится корректировка данных с целью исключения тех наблюдений, которые выбиваются из основной совокупности (и из-за которых на графике бывает неудачный масштаб), или рассматриваются группы наблюдений по отдельности.

Не следует приводить лишние результаты! Некоторые второстепенные таблицы и графики можно опускать (главное, чтобы они содержались в рабочем файле!).

К зачету допускаются студенты, выполнившие все пункты задания и оформившие результаты работы в соответствии с установленными требованиями.

Зачет по лабораторной работе каждый студент сдает персонально преподавателю, ведущему занятия в данной группе.

Для получения зачета студент должен:

- знать теоретические основы тематики лабораторной работы в объеме содержания материалов учебного пособия [4] и лекций;
- уметь ответить на вопросы преподавателя по содержанию выполненной лабораторной работы.

Собеседование (зачет) проходят все студенты отдельно по контрольной и лабораторной работам. По результатам собеседования выставляется «зачет» или «незачет».

*Студенты, не получившие зачета хотя бы по одной из двух зачетных работ, к экзамену не допускаются.*

## **2. Рекомендации по выполнению и оформлению расчетов в Microsoft Excel**

Контрольная и лабораторная работы выполняются с использованием средств MS Excel. В большинстве компьютерных классов установлено программное обеспечение MS Excel 2003. При выполнении контрольных работ дома студенты, как правило, используют более поздние версии MS Excel. Ниже приводятся некоторые отличия в версиях MS Excel, знание которых облегчит студенту выполнение работы.

Основное отличие MS Excel 2010 от MS Excel 2003 – это полностью переработанный пользовательский интерфейс.

Компания Microsoft постаралась сделать так, чтобы на каждом этапе работы над документом все необходимые опции были под рукой, а на экране в то же время не было ничего лишнего. Новый интерфейс является динамическим, то есть привычные панели инструментов заменены на наборы тематических команд, которые появляются на экране в ответ на те или иные действия пользователя.

Панели инструментов и строку меню заменили «ленты» (ribbons), то есть новые наборы команд. Переключение между ними может происходить не только автоматически – основные наборы команд представлены на экране в виде вкладок, поэтому для доступа к ним достаточно щелкнуть по заголовку мышкой.

Главное отличие заключается в том, что команды в них сгруппированы по темам, благодаря чему легко можно найти то, что требуется в данный момент.

Если необходимо получить доступ к дополнительным инструментам, которые в версии MS Excel 2003 вызывались через пункты меню, то нужно щелкнуть по небольшому значку, который расположен в нижней части практически каждой группы.

Функциональные возможности версий MS Excel 2003 и MS Excel 2010 практически не отличаются.

### **►► Названия листов в книге Excel**

Каждый лист книги Excel должен иметь содержательное название (например, «Мультиколлинеарность» или «Гетероскедастичность»), соответствующее выполняемому пункту задания. Это необходимо для удобной навигации по книге при просмотре. Достаточно

щелкнуть правой кнопкой мыши на ярлычке листа, в раскрывшемся контекстном меню выбрать пункт «Переименовать», ввести новое имя и нажать клавишу **Enter**.

Название листа (чуть более развернутое) нужно ввести в ячейку A1 рабочего листа и отформатировать как заголовок полужирным шрифтом более крупного размера (например, Приложение 3. Пошаговый метод).

#### »» Обозначения и описание используемых величин

Вычисление любой величины в некоторой ячейке должно быть документировано. Это означает, что рядом с каждой вычисляемой величиной необходимо ввести ее обозначение (как правило, в соседней ячейке слева) и содержательное описание или текст расчетной формулы (как правило, в соседней ячейке справа):

tkr=	2,02439415	=СТЫЮДРАСПОБР( <i>Alpha</i> ;dfk)
------	------------	-----------------------------------

Для того чтобы в некоторой ячейке Excel отобразить формулу, но не выполнять расчет по ней, а оставить ее в виде текста, нужно перед знаком равенства ввести одиночный апостроф «'».

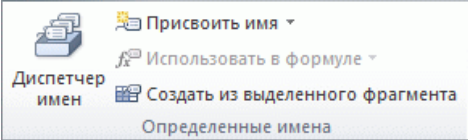
Все обозначения величин должны быть уникальны – никакие две разные величины не должны иметь одинаковое обозначение.

#### »» Именованные ячейки и диапазоны Microsoft Excel

При выполнении расчетов в MS Excel удобно использовать именованные диапазоны и ячейки. При этом каждая используемая в расчетах или анализе ячейка (или блок ячеек) получает помимо формального адреса (буква колонки и номер строки) еще и содержательное имя.

Для того чтобы присвоить ячейке или блоку ячеек некоторое имя, нужно выделить эту ячейку или блок ячеек и последовательно выбрать в меню следующие пункты:



Excel 2003	Excel 2010
<p>1. В меню <b>Вставка</b> укажите на пункт <b>Имя</b> и выберите команду <b>Присвоить</b>.</p> <p>2. Введите имя в диалоговом окне <b>Присвоить имя</b>.</p> <p>3. Чтобы закончить и вернуться на лист, нажмите кнопку <b>ОК</b>.</p>	<p>1. На вкладке <b>Формулы</b> в группе <b>Определенные имена</b> выберите команду <b>Присвоить имя</b>.</p>  <p>2. В диалоговом окне <b>Создание имени</b> введите имя, которое нужно использовать в качестве ссылки, в поле <b>Имя</b>.</p> <p>3. Чтобы закончить и вернуться на лист, нажмите кнопку <b>ОК</b>.</p>

### » Ссылки (или имена) как аргументы функций и формул

Нежелательным является использование конкретных чисел в качестве аргументов функций и формул MS Excel. Этого нужно по возможности избегать.

Например, при вычислении критического значения  $t$ -статистики Стьюдента для уровня значимости<sup>1</sup>  $\alpha = 0,05$  и числа степеней свободы 38 используется запись:

$$=\text{СТЮДРАСПОБР}(0.05;38).$$

Данная формула не является универсальной, так как при смене аргументов необходимо будет каждый раз ее редактировать. Правильнее было бы вместо этого использовать три ячейки следующего содержания:

Alpha	0,05
Dfk	38
Tkr	=СТЮДРАСПОБР( <i>Alpha</i> ;dfk)

Это позволит при необходимости быстро менять аргументы функций, не редактируя саму формулу.

<sup>1</sup> *Уровнем значимости* в статистике называется вероятность совершить ошибку первого рода, то есть отвергнуть верную гипотезу.

### » Вставка функций


Кроме ввода формул для выполнения базовых математических операций, таких как сложение, вычитание, умножение и деление, в MS Excel можно использовать большую библиотеку встроенных функций для выполнения других операций. Для вставки функции необходимо:

Excel 2003	Excel 2010
<ol style="list-style-type: none"> <li>1. В меню <b>Вставка</b> выберите команду <b>Функции....</b></li> <li>2. В раскрывшемся диалоговом окне следуйте инструкциям <b>Мастера функций</b>.</li> </ol>	<ol style="list-style-type: none"> <li>1. На вкладке <b>Формулы</b> в группе <b>Библиотека функций</b> нажмите кнопку <b>Вставить функцию</b>.</li> <li>2. В раскрывшемся диалоговом окне следуйте инструкциям <b>Мастера функций</b>.</li> </ol>

### » Использование надстройки «Пакет анализа»

При проведении сложного статистического или инженерного анализа можно упростить процесс и сэкономить время, используя надстройку **Пакет анализа**.

Загрузка надстройки **Анализ данных**<sup>1</sup>

1. Откройте вкладку **Файл** и выберите пункт **Параметры** (в Microsoft Office 2007 щелкните значок **Кнопка Microsoft Office** , а затем **Параметры Excel**).

2. Выберите команду **Надстройки**, а затем в поле **Управление** выберите пункт **Надстройки Excel**.

3. Нажмите кнопку **Перейти**.

4. В окне **Доступные надстройки** установите флажок **Пакет анализа** и нажмите кнопку **ОК**.

5. После загрузки надстройки **Пакет анализа** в группе **Анализ** на вкладке **Данные** становится доступна команда **Пакет анализа**.

Для анализа данных с помощью этого пакета следует указать входные данные и выбрать параметры. Расчет будет выполнен с помощью подходящей статистической или инженерной макрофункции, а результат – помещен в выходной диапазон. Некоторые инст-

<sup>1</sup> В более ранних версиях в MS Office для запуска **Анализа данных** необходимо выбрать команду **Сервис** → **Надстройки** и активизировать надстройку **Поиск решения**.

рументы позволяют представить результаты анализа в графическом виде.

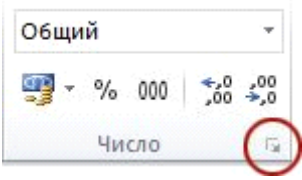
Для выполнения зачетных заданий по эконометрике из **Пакета анализа** необходимо использовать инструменты **Корреляция** и **Регрессия**.

### »» Форматы чисел в MS Excel

Применяя различные числовые форматы, можно изменить формат числа, не изменяя само число. Числовой формат не влияет на фактическое значение ячейки, используемое в MS Excel для осуществления расчетов.



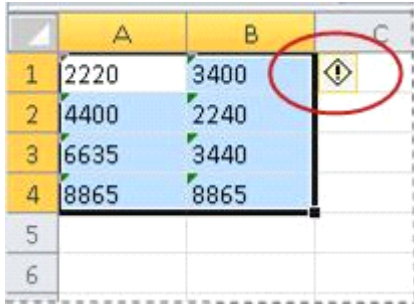
С помощью различных числовых форматов можно выводить числовые данные как текст, число, проценты, даты, валюты и т.д. Например, экспоненциальный формат используется для отображения чисел в экспоненциальном представлении и замены части числа на  $E + n$ , где  $E$  (экспонента) равно предыдущему числу, умноженному на 10 в степени  $n$ . Например, в экспоненциальном формате, где количество знаков после запятой равно двум, число 12345678901 отобразится как  $1,23E + 10$ , то есть как 1,23, умноженное на 10 в 10-й степени. Можно указать используемое количество десятичных знаков. Величина  $8,069E - 10$  в экспоненциальной записи – это то же, что 0,0000000008 в обычной записи.

Для смены формата числа необходимо выполнить следующие действия:

Excel 2003	Excel 2010
<ol style="list-style-type: none"> <li>1. В меню <b>Формат</b> выберите команду <b>Ячейки...</b> (или просто нажмите клавиши Ctrl + 1).</li> <li>2. В списке <b>Числовые форматы</b> выберите нужный формат и, если это необходимо, настройте его параметры.</li> <li>3. Чтобы закончить и вернуться на лист, нажмите кнопку <b>ОК</b>.</li> </ol>	<ol style="list-style-type: none"> <li>1. На вкладке <b>Главная</b> в группе <b>Число</b> нажмите кнопку вызова диалогового окна рядом с надписью <b>Число</b> (или просто нажмите клавиши Ctrl + 1).</li> </ol>  <ol style="list-style-type: none"> <li>2. В списке <b>Числовые форматы</b> выберите нужный формат и, если это необходимо, настройте его параметры.</li> <li>3. Чтобы закончить и вернуться на лист, нажмите кнопку <b>ОК</b>.</li> </ol>

Если после изменения числового формата в ячейке Microsoft Excel отображаются символы #####, то вероятно, что ширина ячейки недостаточна для отображения данных. Чтобы увеличить ширину ячейки, дважды щелкните правую границу столбца, содержащего ячейки с ошибкой #####. Размер столбца автоматически изменится таким образом, чтобы отобразить число. Кроме того, можно перетащить правую границу столбца, увеличив его ширину.

Чаще всего числовые данные отображаются правильно независимо от того, вводятся ли они в таблицу вручную или импортируются из базы данных или другого внешнего источника. Однако иногда MS Excel применяет к данным неправильный числовой формат, из-за чего приходится изменять некоторые настройки. Например, при копировании данных из MS Word к числовым данным может быть применен текстовый формат, что впоследствии приводит к проблемам при вычислениях или нарушению порядка сортировки. Индикатором данной ошибки служит маленький зеленый треугольник в левом верхнем углу ячейки. Для ее исправления:

Excel 2003	Excel 2010
<ol style="list-style-type: none"> <li>1. В меню <b>Сервис</b> выберите команду <b>Параметры</b> и откройте вкладку <b>Контроль ошибок</b>.</li> <li>2. Убедитесь, что установлены флажки <b>Включить фоновую проверку ошибок</b> и <b>Число сохранено как текст</b>.</li> <li>3. Выделите ячейки с зеленым индикатором ошибки в верхнем левом углу .</li> <li>4. Нажмите на появившуюся рядом с ячейкой кнопку  и выберите команду <b>Преобразовать в число</b>.</li> </ol>	<ol style="list-style-type: none"> <li>1. Выделите любую ячейку или диапазон смежных ячеек с индикатором ошибки в верхнем левом углу.</li> <li>2. Нажмите появившуюся рядом с выделенной ячейкой или диапазоном ячеек кнопку ошибки.</li> </ol>  <ol style="list-style-type: none"> <li>3. Выберите в меню пункт <b>Преобразовать в число</b>.</li> </ol>

При выполнении вычислений ошибка также может возникнуть, если при вводе чисел использовать разделитель целой и дробной части, отличный от принятого в системе. Как правило, для разделения целой и дробной частей используются «.» или «,». Определить, какой разделитель используется, можно щелкнув по ярлыку **Язык и региональные стандарты** в окне **Панель управления Windows**.

### » Копирование листов Microsoft Excel

Лист с исходными данными не следует изменять ни при каких обстоятельствах. Для выполнения различных расчетов или тестов нужно сначала скопировать данные, а затем уже на листе-копии произвести необходимые изменения.

Наиболее удобным способом является копирование целого листа. Для этого:

Excel 2003	Excel 2010
<ol style="list-style-type: none"> <li>1. В меню <b>Правка</b> выберите команду <b>Переместить/скопировать лист....</b></li> <li>2. В диалоговом окне <b>Переместить или скопировать</b> в списке <b>Перед листом</b> укажите, куда необходимо скопировать лист, установите флажок <b>Создать копию</b>.</li> <li>3. Чтобы закончить и вернуться на лист, нажмите кнопку <b>ОК</b>.</li> </ol>	<ol style="list-style-type: none"> <li>1. На вкладке <b>Главная</b> в группе <b>Ячейки</b> нажмите кнопку <b>Формат</b> и в разделе <b>Упорядочить листы</b> выберите пункт <b>Переместить или скопировать лист</b>. <div data-bbox="890 1151 1075 1326" data-label="Image"> </div> </li> <li>2. В диалоговом окне <b>Переместить или скопировать</b> в списке <b>Перед листом</b> укажите, куда необходимо скопировать лист, установите флажок <b>Создать копию</b>.</li> <li>3. Чтобы закончить и вернуться на лист, нажмите кнопку <b>ОК</b>.</li> </ol>

Можно, конечно, не создавать копию целого листа, а скопировать сами данные на новый лист, однако при этом придется заново изменять ширину столбцов, если она была специально настроена на листе с исходными данными, для того чтобы нагляднее отобразить их.

### »» Создание парных диаграмм рассеяния

Важнейшим элементом эконометрического исследования является графический анализ исходных данных. В случае множественной регрессии, то есть когда у нас несколько показателей, необходимо построить парные диаграммы зависимости объясняемой переменной  $Y$  от каждой из объясняющих переменных  $X$  – диаграммы рассеяния.

В эконометрическом анализе диаграммы, используемые при построении регрессионной модели, имеют тип «Точечная».

Каждая диаграмма должна иметь содержательный заголовок, пусть и такой простой, как «Зависимость объема продаж транспортного средства от расходов на бензин и индекса потребительских расходов». Оси диаграммы тоже должны быть названы сокращенными именами соответствующих переменных (обязательно включая единицы измерения). А вот легенды на парной диаграмме рассеяния может и не быть, если на ней не добавлена линия тренда.

Для создания диаграммы рассеяния нужно выделить два столбца данных со значениями показателей, включая их названия (метки) в первой строке матрицы данных, и выполнить следующие действия:

Excel 2003	Excel 2010
<ol style="list-style-type: none"> <li>1. В меню <b>Вставка</b> выберите команду <b>Диаграмма...</b></li> <li>2. Следуйте инструкциям в раскрывшемся диалоговом окне <b>Мастер диаграмм</b>.</li> </ol>	<ol style="list-style-type: none"> <li>1. На вкладке <b>Вставка</b> в группе <b>Диаграммы</b> выберите тип и вид диаграммы, диаграмма добавится на лист.</li> <li>2. Щелкните в любом месте внедренной диаграммы, чтобы активизировать ее. Откроется панель <b>Работа с диаграммами</b> с дополнительными вкладками <b>Конструктор</b>, <b>Макет</b> и <b>Формат</b>.</li> <li>3. Выберите необходимые команды для оформления диаграммы.</li> </ol>

Нужно помнить, что для того чтобы MS Excel правильно определил переменные, объясняемая переменная  $Y$  должна быть расположена в правом из двух выделенных столбцов, а объясняющая переменная  $X$  – в левом.

Для добавления данных в уже построенную диаграмму необходимо предварительно внести их в исходную таблицу, а затем выполнить следующие действия:

Excel 2003	Excel 2010
<ol style="list-style-type: none"><li>1. Щелкните правой кнопкой мыши на области диаграммы.</li><li>2. В раскрывшемся контекстном меню выберите команду <b>Исходные данные</b>.</li><li>3. В окне <b>Исходные данные</b> перейдите на вкладку <b>Ряд</b> и укажите новые диапазоны данных для значений Y и X.</li><li>4. Чтобы закончить и вернуться на лист, нажмите кнопку <b>ОК</b>.</li></ol>	<ol style="list-style-type: none"><li>1. Щелкните правой кнопкой мыши на области диаграммы.</li><li>2. В раскрывшемся контекстном меню выберите команду <b>Выбрать данные</b>.</li><li>3. В окне <b>Выбор источника данных</b> в поле <b>Элементы легенды (ряды)</b> выделите нужный ряд и нажмите кнопку <b>Изменить</b>.</li><li>4. В окне <b>Изменение ряда</b> укажите новые диапазоны данных для значений Y и X.</li><li>5. Чтобы закончить и вернуться на лист, последовательно нажимайте кнопку <b>ОК</b>.</li></ol>

### 3. Справочные материалы для выполнения расчетов

#### Формулы, используемые при корреляционном анализе

Формула для вычислений	Функция или инструмент Анализа данных в Excel	Результат вычислений/ Примечания
Среднее значение $\bar{x} = \frac{1}{n} \sum_{i=1}^n x_i$	СРЗНАЧ(число1; число2;...)	Возвращает среднее значение (среднее арифметическое) аргументов
Дисперсия $S_x^2 = \frac{1}{n-1} \sum (x_i - \bar{x})^2$	ДИСП(число1; число2;...)	Оценивает дисперсию по выборке
Стандартное отклонение $S_x = \sqrt{S_x^2}$	СТАНДОТКЛОН(число1;число2;...)	Оценивает стандартное отклонение по выборке. Стандартное отклонение – это мера того, насколько широко разбросаны точки данных относительно их среднего
Сумма квадратов отклонений $\sum_{i=1}^n (x_i - \bar{x})^2$	КВАДРОТКЛ(число1;число2;...)	Возвращает сумму квадратов отклонений точек данных от их среднего
Коэффициент корреляции $r_{x,y} = \frac{\sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_{i=1}^n (x_i - \bar{x})^2 \cdot \sum_{i=1}^n (y_i - \bar{y})^2}}$	КОРРЕЛ(массив1;массив2)	Возвращает коэффициент корреляции между интервалами ячеек массив1 и массив2
t-критерий Стьюдента для проверки значимости коэффициента корреляции $t_{\text{набл}} = \sqrt{\frac{r_{y,x}^2}{1-r_{y,x}^2}} (n-2)$	СТЮДРАСПОБР(вероятность; степени_свободы)	Вычисленное по этой формуле значение $t_{\text{набл}}$ сравнивается с критическим значением t-критерия, которое берется из таблицы значений t-распределения Стьюдента с учетом заданного уровня значимости и числа степеней свободы ( $n-2$ ) или определяется с помощью функции СТЮДРАСПОБР( )



<p><b>Матрица коэффициентов парной корреляции</b></p> $R = \begin{pmatrix} 1 & r_{yx_1} & r_{yx_2} & \dots & r_{yx_m} \\ r_{yx_1} & 1 & r_{x_1x_2} & \dots & r_{x_1x_m} \\ r_{yx_2} & r_{x_1x_2} & 1 & \dots & r_{x_2x_m} \\ \dots & \dots & \dots & \dots & \dots \\ r_{yx_m} & r_{x_1x_m} & r_{x_2x_m} & \dots & 1 \end{pmatrix}$	<p><b>Обращение к средствам анализа данных.</b> Для вычисления матрицы коэффициентов парной корреляции <math>R</math> следует воспользоваться инструментом <b>Корреляция</b> из пакета <b>Анализ данных</b></p>	<p>Инструмент <b>Корреляция</b> применяется, если имеется более двух переменных измерений для каждого объекта. В результате выдается таблица – корреляционная матрица, показывающая значение функции КОРРЕЛ( ) для каждой возможной пары переменных измерений. Любое значение коэффициента корреляции должно находиться в диапазоне от <math>-1</math> до <math>+1</math> включительно</p>
---	---	--

### Формулы, используемые при регрессионном анализе

Формула для вычислений	Функция или инструмент Анализа данных в Excel	Результат вычислений / Примечания
<p><b>Оценка параметров модели парной и множественной линейной регрессии</b></p> $A = (X'X)^{-1} X'Y$	<p>Для вычисления параметров уравнения регрессии следует воспользоваться инструментом <b>Регрессия</b> из пакета <b>Анализ данных</b></p>	<p>Возвращает подробную информацию о параметрах модели, качестве модели, расчетных значениях и остатках в виде четырех таблиц: <i>Регрессионная статистика</i>, <i>Дисперсионный анализ</i>, <i>Коэффициенты</i>, <i>Вывод остатка</i>. Также могут быть получены <i>график подбора</i> и <i>график остатков</i></p>

## Оценка качества модели регрессии

<p><b>F-критерий Фишера</b> для проверки значимости модели регрессии</p> $F = \frac{R^2 / k}{(1 - R^2) / (n - k - 1)}$	<p><b>=ФРАСПОБР(вероятность; степени_свободы1; степени_свободы2)</b></p> <p><b>вероятность</b> – это вероятность, связанная с F-распределением</p> <p><b>степени_свободы 1</b> – это числитель степеней свободы (<math>v_1 = k</math>)</p> <p><b>степени_свободы 2</b> – это знаменатель степеней свободы (<math>v_2 = (n - k - 1)</math>), где <math>k</math> – количество факторов, включенных в модель)</p>	<p>Возвращает обратное значение для F-распределения вероятностей. ФРАСПОБР() можно использовать, чтобы определить критические значения F-распределения. Чтобы определить критическое значение <math>F</math>, нужно использовать уровень значимости <math>\alpha</math> как аргумент <b>вероятность</b> для ФРАСПОБР().</p>
<p><b>Коэффициент детерминации</b></p> $R^2 = \frac{\sum_{i=1}^n (\hat{y}_i - \bar{y})^2}{\sum_{i=1}^n (y_i - \bar{y})^2} = 1 - \frac{\sum_{i=1}^n \varepsilon_i^2}{\sum_{i=1}^n (y_i - \bar{y})^2}$	<p><b>Коэффициент детерминации</b> показывает долю вариации результирующего признака, находящегося под воздействием изучаемых факторов, то есть определяет, какая доля вариации признака <math>Y</math> учтена в модели и обусловлена влиянием на него факторов. Чем ближе <math>R^2</math> к 1, тем выше качество модели</p>	
<p><b>Коэффициент множественной корреляции</b> (индекс корреляции) <math>R</math></p> $R = \sqrt{1 - \frac{\sum_{i=1}^n \varepsilon_i^2}{\sum_{i=1}^n (y_i - \bar{y})^2}} = \sqrt{\frac{\sum_{i=1}^n (\hat{y}_i - \bar{y})^2}{\sum_{i=1}^n (y_i - \bar{y})^2}}$	<p>Данный коэффициент является универсальным, так как он отражает тесноту связи и точность модели, а также может использоваться при любой форме связи переменных. Чем ближе <math>R</math> к 1, тем выше качество модели</p>	
<p><b>t-критерий Стьюдента</b> для оценки значимости параметров модели линейной регрессии:</p> $t_{aj} = \hat{a}_j / \sigma_{aj}$	<p>Вычисленное значение <math>t_{aj}</math> сравнивается с критическим значением <math>t</math>-критерия, которое берется из таблицы значений <math>t</math>-распределения Стьюдента с учетом заданного уровня значимости и числа степеней свободы (<math>n - k - 1</math>). В Excel критическое значение <math>t</math>-критерия можно получить с помощью функции <b>СТЮДРАСПОБР(вероятность; степени_свободы)</b></p> <p><b>вероятность</b> – вероятность, соответствующая двустороннему распределению Стьюдента</p> <p><b>степени_свободы</b> – число степеней свободы, характеризующее распределение</p>	

<b>Средняя относительная ошибка аппроксимации</b>		Средняя относительная ошибка аппроксимации – оценка точности модели
$E_{отн} = \frac{1}{n} \sum_{i=1}^n \frac{ e_i }{y_i} \times 100\%$		
<b>Оценка влияния отдельных факторов на зависимую переменную на основе модели</b>		
<b>Коэффициенты эластичности</b>	$\varepsilon_j = a_j \cdot \frac{\bar{x}_j}{\bar{y}}$	Коэффициент эластичности показывает, на сколько процентов изменится значение исследуемой величины при изменении соответствующего фактора на 1%
<b>Бета-коэффициенты</b>	$\beta_j = \hat{a}_j \cdot \frac{S_{x_j}}{S_y}$	Бета-коэффициент показывает, на какую часть своего СКО изменится значение исследуемой переменной при изменении соответствующего фактора на 1 СКО
<b>Дельта-коэффициенты</b>	$\Delta_j = r_{y,x_j} \cdot \beta_j / R^2$	Дельта-коэффициент показывает среднюю долю влияния соответствующего фактора в совокупном влиянии всех факторов, включенных в модель
<b>Построение интервальных прогнозов по модели регрессии</b>		
$U(X_{прогн}) = \sigma_e \cdot t_{\alpha} \cdot \sqrt{1 + X_{прогн}^T \cdot (X^T \cdot X)^{-1} \cdot X_{прогн}}$ – ошибка прогнозирования, которая позволяет определить доверительный интервал прогноза,		
где $\sigma_e = \sqrt{\frac{\sum e_i^2}{n - k - 1}}$ – стандартная ошибка модели		

### Регрессионная статистика в отчете Excel

Наименование в отчете Excel	Принятое наименование	Формула
Множественный $R$	Коэффициент множественной корреляции, индекс корреляции	$R = \sqrt{R^2}$
$R$ -квадрат	Коэффициент детерминации, $R^2$	$R^2 = \frac{\sum_{i=1}^n (\hat{y}_i - \bar{y})^2}{\sum_{i=1}^n (y_i - \bar{y})^2} = 1 - \frac{\sum_{i=1}^n e_i^2}{\sum_{i=1}^n (y_i - \bar{y})^2}$
Нормированный $R$ -квадрат	Скорректированный $R^2$	$\bar{R}^2 = 1 - \left(1 - R^2\right) \frac{n - 1}{n - k - 1}$
Стандартная ошибка	Среднеквадратическое отклонение от модели	$\sigma_e = \sqrt{\sum e_i^2 / (n - k - 1)} = \sqrt{ESS / (n - k - 1)}$

### Дисперсионный анализ в отчете Excel

Наименование в отчете Excel	Df – число степеней свободы	SS – сумма квадратов	MS – дисперсия на одну степень свободы	F-критерий Фишера
Регрессия	$k$	$RSS = \sum (\hat{y}_i - \bar{y})^2$	$\sum (\hat{y}_i - \bar{y})^2 / k =$ $= RSS / k$	$F = \frac{R^2 / k}{(1 - R^2) / (n - k - 1)}$
Остаток	$n - k - 1$	$ESS = \sum e_i^2$	$\sum e_i^2 / (n - k - 1) =$ $= ESS / (n - k - 1)$	
Итого	$n - 1$	$TSS = \sum (y_i - \bar{y})^2$		

Названия некоторых функций в Excel 2010 были изменены по сравнению с более ранними версиями.

Чтобы повысить точность работы функций MS Excel, обеспечить их согласованность и привести имена функций в соответствие с их назначением, корпорация Microsoft изменила, переименовала и добавила несколько функций в библиотеку MS Excel 2010.

Для обеспечения обратной совместимости переименованные функции доступны также и по их старым именам.

Название функции в Excel более ранних версий	Название функции в Excel 2010	Примечания
ДИСП(число1,[число2],...]	ДИСП.В(число1,[число2],...]	Оценивает дисперсию по выборке
СТЮДРАС-ПОБР(вероятность; степени_свободы)	СТЪДЕНТ.ОБР.2Х(вероятность, степени_свободы)	Возвращает двустороннее обратное $t$ -распределение Стьюдента
ФРАСПОБР(вероятность; степени_свободы1; степени_свободы2)	F.ОБР.ПХ(вероятность, степени_свободы1, степени_свободы2)	Возвращает значение, обратное (правостороннему) $F$ -распределению вероятностей
ХИ2ОБР(вероятность, степени_свободы)	ХИ2.ОБР.ПХ(вероятность, степени_свободы)	Возвращает обратное значение односторонней вероятности распределения хи-квадрат

#### 4. Комплексный пример исследования экономических данных с использованием корреляционно-регрессионного анализа

На основе статистических данных за 16 месяцев, представленных в табл. 4.1, проведите корреляционно-регрессионный анализ с целью прогнозирования объема реализации продукции фирмы на два месяца вперед.

Таблица 4.1. Исходные данные

Y	X <sub>1</sub>	X <sub>2</sub>	X <sub>3</sub>	X <sub>4</sub>	X <sub>5</sub>
Объем реализации	Время	Затраты на рекламу	Цена товара	Средняя цена товара у конкурентов	Индекс потребительских расходов
126	1	4,0	15,0	17,0	100,0
137	2	4,8	14,8	17,3	98,4
148	3	3,8	15,2	16,8	101,2
191	4	8,7	15,5	16,2	103,5
274	5	8,2	15,5	16,0	104,1
370	6	9,7	16,0	18,0	107,0
432	7	14,7	18,1	20,2	107,4
445	8	18,7	13,0	15,8	108,5
367	9	19,8	15,8	18,2	108,3
367	10	10,6	16,9	16,8	109,2
321	11	8,6	16,3	17,0	110,1
307	12	6,5	16,1	18,3	110,7
331	13	12,6	15,4	16,4	110,3
345	14	6,5	15,7	16,2	111,8
364	15	5,8	16,0	17,7	112,3
384	16	5,7	15,1	16,2	112,9

**[?] 1.** Осуществите двумя способами выбор факторных признаков для построения регрессионной модели:

а) на основе анализа матрицы коэффициентов парной корреляции с проверкой гипотезы о независимости объясняющих перемен-

ных (тест на выявление мультиколлинеарности Фаррара–Глоубера);

б) с помощью пошагового отбора методом исключения.

2. Оцените параметры модели. Дайте экономическую интерпретацию коэффициентов регрессии.

3. Для оценки качества модели определите:

а) коэффициент детерминации;

б) коэффициент множественной корреляции;

в) среднюю относительную ошибку аппроксимации.

4. Проведите оценку значимости уравнения регрессии и его коэффициентов.

5. По диаграммам остатков определите ту объясняющую переменную, от которой может зависеть дисперсия случайных возмущений. Проверьте выполнение условия гомоскедастичности остатков по тесту Голдфелда–Кванда.

6. Оцените по модели влияние факторов на зависимую переменную.

7. Постройте точечный и интервальный прогнозы результирующего показателя на два месяца вперед ( $\alpha = 0,1$ ).

### **Решение.**

Содержательная интерпретация конечной цели задачи – прогнозирования объема продаж:

*прогноз объема продаж* – это предсказание будущего спроса, выраженное в денежных единицах или единицах продаваемого товара; в более узком смысле – это процесс определения объема реализации товара или группы товаров на несколько ближайших периодов времени.

## **4.1. Выбор факторных признаков для построения регрессионной модели**

### ***Корреляционный анализ данных***

Объем реализации – это зависимая переменная  $Y$  (тыс. руб.).

В качестве независимых, объясняющих переменных выбраны:

$X_1$  – время, дни;

$X_2$  – затраты на рекламу, тыс. руб.;

$X_3$  – цена товара, руб.;

$X_4$  – средняя цена товара у конкурентов, руб.;

$X_5$  – индекс потребительских расходов, %.

В этом примере количество наблюдений  $n = 16$ , количество объясняющих переменных  $m = 5$ .

Для проведения корреляционного анализа используем инструмент **Корреляция** (надстройка **Анализ данных Excel**).

В результате будет получена матрица коэффициентов парной корреляции (табл. 4.2).

Таблица 4.2. *Результат корреляционного анализа*

	Объем реализации	Время	Затраты на рекламу	Цена товара	Средняя цена товара у конкурентов	Индекс потребительских расходов
Объем реализации	1					
Время	0,678	1				
Затраты на рекламу	0,646	0,106	1			
Цена товара	0,233	0,174	–0,003	1		
Средняя цена товара у конкурентов	0,226	–0,051	0,204	0,698	1	
Индекс потребительских расходов	0,816	0,960	0,273	0,235	0,03	1

Анализ матрицы коэффициентов парной корреляции начнем с анализа первого столбца матрицы, в котором расположены коэффициенты корреляции, отражающие тесноту связи зависимой переменной **Объем реализации** с включенными в анализ факторами. Анализ показывает, что зависимая переменная, то есть объем реализации, имеет тесную связь с индексом потребительских расходов ( $r_{x_5} = 0,816$ ), с затратами на рекламу ( $r_{x_2} = 0,646$ ) и временем ( $r_{x_1} = 0,678$ ). Факторы  $X_3$  и  $X_4$  имеют слабую связь с зависимой переменной и их не рекомендуется включать в модель регрессии.

Затем перейдем к анализу остальных столбцов матрицы с целью выявления коллинеарности. Факторы  $X_1$  и  $X_5$  тесно связаны между собой ( $= 0,960$ ), что свидетельствует о наличии коллинеарности. Из этих двух переменных оставим  $X_5$  – индекс потребительских расходов, так как  $r_{x_1y} = 0,678 < r_{x_5y} = 0,816$ .

Таким образом, на основе анализа только корреляционной матрицы остаются два фактора – **Затраты на рекламу** и **Индекс потребительских расходов** ( $n = 16, k = 2$ ).

Одним из условий классической регрессионной модели является предположение о *независимости* объясняющих переменных.

В нашем примере из двух тесно связанных друг с другом факторов  $X_1$  и  $X_5$  ( $= 0,960$ ) один,  $X_1$ , был исключен.

Для выявления мультиколлинеарности оставшихся факторов выполняем *тест Фаррара–Глоубера* по факторам  $X_2, X_3, X_4, X_5$ .

#### **1. Проверка наличия мультиколлинеарности всего массива переменных**

1. Построим матрицу межфакторных корреляций  $R_1$  (табл. 4.3) и найдем ее определитель  $\det[R_1] = 0,373$  с помощью функции МОПРЕД.

Таблица 4.3. Матрица  $R_1$

	$X_2$	$X_3$	$X_4$	$X_5$
$X_2$	1	–0,003	0,204	0,273
$X_3$	–0,003	1	0,698	0,235
$X_4$	0,204	0,698	1	0,031
$X_5$	0,273	0,235	0,031	1



2. Вычислим наблюдаемое значение статистики Фаррара–Глоубера по следующей формуле:

$$FG_{\text{набл}} = - \left[ n - 1 - \frac{1}{6}(2k + 5) \right] \ln(\det[R_1]) =$$

$$= - [15 - 13/6] \cdot \ln(0,373) = 12,66,$$

где  $n = 16$  – количество наблюдений;  
 $k = 4$  – количество факторов.

Фактическое значение этого критерия  $FG_{\text{набл}}$  сравниваем с табличным значением  $\chi^2$  при  $\frac{1}{2}k(k-1) = 6$  степенях свободы и уровне значимости  $\alpha = 0,05$ . Табличное значение  $\chi^2$  можно найти с помощью функции ХИ2.ОБР.ПХ<sup>1</sup> (рис. 4.1).

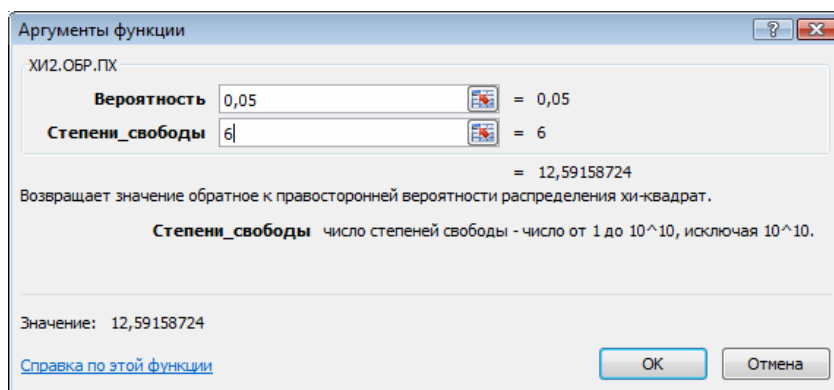


Рис. 4.1. Получение табличного значения  $\chi^2$

Так как  $FG_{\text{набл}} > FG_{\text{крит}}$  ( $12,66 > 12,59$ ), то в массиве объясняющих переменных существует мультиколлинеарность.

<sup>1</sup> В более ранних версиях Excel – ХИ2ОБР.

2. Проверка наличия мультиколлинеарности каждой переменной с другими переменными

1. Вычислим обратную матрицу

$$C \Rightarrow R_1^{-1} =$$

	$X_2$	$X_3$	$X_4$	$X_5$
$X_2$	1,252	0,544	-0,621	-0,451
$X_3$	0,544	2,376	-1,749	-0,654
$X_4$	-0,621	-1,749	2,331	0,510
$X_5$	-0,451	-0,654	0,510	1,262

2. Вычислим  $F$ -критерии  $F_j = (c_{jj} - 1) \frac{n-k-1}{k}$ , где  $c_{jj}$  – диагональные элементы матрицы  $C$ :

$F_2$	$F_3$	$F_4$	$F_5$
0,692	3,784	3,660	0,719

3. Фактические значения  $F$ -критериев сравниваем с табличным значением  $F_{\text{табл}} = 3,357$  при  $v_1 = 4$  и  $v_2 = (n - k - 1) = 11$  степенях свободы и уровне значимости  $\alpha = 0,05$ , где  $k$  – количество факторов.

4. Так как  $F_3 > F_{\text{табл}}$  и  $F_4 > F_{\text{табл}}$ , то независимые переменные  $X_3$  и  $X_4$  мультиколлинеарны с другими.

3. Проверка наличия мультиколлинеарности каждой пары переменных

1. Вычислим частные коэффициенты корреляции по формуле

$$r_{ij}(\ ) = \frac{-c_{ij}}{\sqrt{c_{ii} \cdot c_{jj}}}, \text{ где } c_{ij} - \text{элементы матрицы } C:$$

$$r_{2,3(4,5)} = \frac{-(0,544)}{\sqrt{1,252 \cdot 2,376}} = -0,315;$$

$$r_{2,4(3,5)} = \frac{-(-0,621)}{\sqrt{1,252 \cdot 2,331}} = 0,363;$$

$$r_{2,5(3,4)} = \frac{-(-0,451)}{\sqrt{1,252 \cdot 1,262}} = 0,359;$$

$$r_{3,4(2,5)} = \frac{-(-1,749)}{\sqrt{2,376 \cdot 2,331}} = 0,743;$$

$$r_{3,5(2,4)} = \frac{-(-0,654)}{\sqrt{2,376 \cdot 1,262}} = 0,378;$$

$$r_{4,5(2,3)} = \frac{-(0,510)}{\sqrt{2,331 \cdot 1,262}} = -0,297.$$

2. Вычислим  $t$ -критерии по формуле  $t_{ij} = \frac{r_{ij( )} \sqrt{n-k-1}}{\sqrt{1-r_{ij( )}^2}}$  :

$$\begin{aligned} t_{2,3} &= -1,102; \\ t_{2,4} &= 1,293; \\ t_{2,5} &= 1,275; \\ t_{3,4} &= 3,682; \\ t_{3,5} &= 1,353; \\ t_{4,5} &= -1,032. \end{aligned}$$

Фактические значения  $t$ -критериев сравниваются с табличным значением при степенях свободы  $(n-k-1)=11$  и уровне значимости  $\alpha = 0,05$ :  $t_{\text{табл}} = 2,201$ . Так как  $|t_{3,4}| > t_{\text{табл}}$  и  $r_{3,4(2,5)} = 0,743 \rightarrow 1$ , то между независимыми переменными  $X_3$  и  $X_4$  существует мультиколлинеарность.

Для того чтобы избавиться от мультиколлинеарности, можно исключить одну из переменных мультиколлинеарной пары  $X_3, X_4$ . Удалить следует переменную  $X_3$ , так как у нее больше значение  $F$ -критерия. Следовательно, она больше влияет на общую мультиколлинеарность факторов.

Результаты проведенного теста не опровергают выводы, сделанные ранее только на основе корреляционной матрицы.

Целесообразность включения фактора  $X_4$  рассмотрим с помощью *теста на выбор «длинной» и «короткой» регрессии*. Этот тест используется для отбора наиболее существенных объясняющих переменных. Иногда переход от большего числа исходных показателей анализируемой системы к меньшему числу наиболее информативных факторов может быть объяснен дублированием информации из-за сильно взаимосвязанных факторов. Стремление к построению более простой модели приводит к идее уменьшения размерности модели без потери ее качества. Для этого используют тест проверки «длинной» и «короткой» регрессий.

Рассмотрим две модели регрессии:

$$\begin{aligned} y_i &= \beta_0 + \beta_1 x_{i1} + \dots + \beta_k x_{ik} + \varepsilon_i \text{ (длинную),} \\ y_i &= \beta_0 + \beta_1 x_{i1} + \dots + \beta_k x_{ik-q} + \varepsilon_i \text{ (короткую).} \end{aligned}$$

Предположим, что модель не зависит от последних  $q$  объясняющих переменных и их можно исключить из модели. Это соответствует гипотезе

$$H_0: \beta_{k-q+1} = \beta_{k-q+2} = \dots = \beta_k = 0,$$

то есть последние  $q$  коэффициентов  $\beta_i$  равны нулю.

Алгоритм проверки следующий:

1. Построим по МНК «длинную» регрессию по всем факторам  $X_1, \dots, X_k$  и найдем для нее сумму квадратов остатков  $ESS_{\text{длин}}$ .

2. Построим по МНК «короткую» регрессию по первым  $(k - q)$  факторам  $X_1, \dots, X_{k-q}$  и найдем для нее сумму квадратов остатков  $ESS_{\text{кор}}$ .

3. Вычислим  $F$ -статистику:

$$F_{\text{набл}} = \frac{(ESS_{\text{кор}} - ESS_{\text{длин}}) / q}{ESS_{\text{длин}} / (n - k - 1)}$$

4. Если  $F_{\text{набл}} > F_{\text{табл}}(\alpha, v_1 = q, v_2 = n - k - 1)$ , то гипотеза отвергается (выбираем «длинную» регрессию), в противном случае – «короткую» регрессию.

На основании данных примера сравним две модели – «длинную» (с факторами  $X_2, X_4, X_5$ ) и «короткую» (с факторами  $X_2, X_5$ ).

1. Построим «длинную» регрессию по всем факторам  $X_2, X_4, X_5$  и найдем для нее сумму квадратов остатков  $ESS_{\text{длин}}$ .

Дисперсионный анализ					
	<i>df</i>	<i>SS</i>	<i>MS</i>	<i>F</i>	Значимость <i>F</i>
Регрессия	3	138 429,778	46 143,259	27,292	1,20724E-05
Остаток	12	20 288,659	1 690,722		
Итого	15	158 718,438			

	Коэффициенты	Стандартная ошибка	t-статистика	P-значение	Нижние 95%	Верхние 95%
Y – пересечение	–1654,763	306,264	–5,403	0,000	–2322,054	–987,472
$X_2$	9,052	2,295	3,945	0,002	4,052	14,051
$X_5$	15,825	2,447	6,468	0,000	10,494	21,156
$X_4$	10,539	9,521	1,107	0,290	–10,206	31,284

2. Построим «короткую» регрессию по первым факторам  $X_2, X_5$  и найдем для нее сумму квадратов остатков  $ESS_{\text{кор}}$ .

Дисперсионный анализ					
	<i>df</i>	<i>SS</i>	<i>MS</i>	<i>F</i>	Значимость <i>F</i>
Регрессия	2	136 358,334	68 179,167	39,639	2,93428E-06
Остаток	13	22 360,104	1 720,008		
Итого	15	158 718,438			

	Коэффициенты	Стандартная ошибка	t-статистика	P-значение	Нижние 95%	Верхние 95%
Y – пересечение	–1471,314	259,766	–5,664	0,000	–2032,505	–910,124
$X_2$	9,568	2,266	4,223	0,001	4,673	14,464
$X_5$	15,753	2,467	6,386	0,000	10,424	21,082

3. Вычислим  $F$ -статистику:

$$F_{\text{набл}} = \frac{(ESS_{\text{кор}} - ESS_{\text{длин}})/q}{ESS_{\text{длин}}/(n - k - 1)} = \frac{(22460,104 - 20288,659)/1}{20288,659/(16 - 3 - 1)} = 1,225$$

$$F_{\text{табл}} = 4,747.$$

4. Так как  $F_{\text{набл}} < F_{\text{табл}}$  ( $1,125 < 4,747$ ), выбираем «короткую» регрессию

$$\hat{y}_i = -1471,31 + 9,57x_2 + 15,75x_5.$$

**Выбор факторных признаков для построения регрессионной модели методом исключения**

Для проведения регрессионного анализа используем инструмент **Регрессия** (надстройка **Анализ данных** в Excel).

На первом шаге строится модель регрессии по всем факторам:

$$\hat{y}_i = -3017,40 - 13,42x_1 + 6,67x_2 - 6,48x_3 + 12,24x_4 + 30,48x_5.$$

$$(10,38) \quad (3,01) \quad (15,78) \quad (14,41) \quad (11,52)$$

В скобках указаны значения стандартных ошибок коэффициентов регрессии.

Фрагмент протокола регрессионного анализа приведен в табл. 4.4.

*Таблица 4.4. Модель регрессии по пяти факторам*

	Кэф- фициен- ты	Стан- дартная ошибка	t-ста- тистика	P-зна- чение	Нижние 95%	Верхние 95%
У – пересече- ние	-3017,40	1094,49	-2,76	0,02	-5456,06	-578,73
Время – $X_1$	-13,42	10,38	-1,29	0,23	-36,54	9,71
Затраты на рекламу – $X_2$	6,67	3,01	2,22	0,05	-0,03	13,38
Цена товара – $X_3$	-6,48	15,78	-0,41	0,69	-41,63	28,68
Средняя цена товара у конку- рентов – $X_4$	12,24	14,41	0,85	0,42	-19,87	44,34
Индекс потре- бительских расходов – $X_5$	30,48	11,52	2,64	0,02	4,80	56,15

В данном случае коэффициенты уравнения регрессии при  $X_1$ ,  $X_3$ ,  $X_4$  незначимы при 5%-ном уровне значимости. После построения уравнения регрессии и оценки значимости всех коэффициентов регрессии из модели исключают тот фактор, коэффициент при котором незначим и имеет наименьший по абсолютной величине коэффициент  $t$ , а именно  $X_3$ .

После этого получают новое уравнение множественной регрессии

$$\hat{y}_i = -2914,33 - 12,57x_1 + 7,13x_2 + 7,93x_4 + 29,15x_5$$

(9,78)   (2,69)   (9,49)   (10,64)

и снова производят оценку значимости всех оставшихся коэффициентов регрессии (табл. 4.5).

*Таблица 4.5. Модель регрессии по четырем факторам*

	Кэф- фици- енты	Стан- дартная ошибка	t-ста- тистика	P-зна- чение	Нижние 95%	Верхние 95%
<b>Y – пересечение</b>	-2914,33	1024,23	-2,85	0,02	-5168,65	-66,00
<b>Время – <math>X_1</math></b>	-12,57	9,78	-1,29	0,23	-34,09	8,95
<b>Затраты на рек- ламу – <math>X_2</math></b>	7,13	2,69	2,65	0,02	1,20	13,05
<b>Средняя цена товара у конку- рентов – <math>X_4</math></b>	7,93	9,49	0,84	0,42	-12,96	28,82
<b>Индекс потреби- тельских расхо- дов – <math>X_5</math></b>	29,15	10,64	2,74	0,02	5,74	52,56

Так как среди них есть незначимые ( $X_1$  и  $X_4$ ), то исключают фактор с наименьшим значением  $t$ -критерия –  $X_4$ . В табл. 4.6 представлены результаты, полученные после исключения фактора  $X_4$ . На следующем шаге исключаем незначимый фактор  $X_1$ .

Таблица 4.6. Модель регрессии по трем факторам

	Коэф- фициен- ты	Стан- дартная ошибка	t-ста- тистика	P-зна- чение	Нижние 95%	Верхние 95%
Y – пересече- ние	–2957,61	1009,97	–2,93	0,01	–5158,15	–2957,61
Время – X <sub>1</sub>	–14,32	9,43	–1,52	0,15	–34,86	–14,32
Затраты на рекламу – X <sub>2</sub>	7,23	2,65	2,72	0,02	1,45	7,23
Индекс потре- бительских расходов – X <sub>5</sub>	30,95	10,28	3,01	0,01	8,54	30,95

Процесс исключения факторов останавливается на том шаге, при котором все регрессионные коэффициенты значимы (табл. 4.7).

Таблица 4.7. Модель регрессии со значимыми факторами

	Коэф- фициен- ты	Стан- дартная ошибка	t-ста- тистика	P-зна- чение	Нижние 95%	Верхние 95%
Y – пересече- ние	–1471,31	259,77	–5,66	0,00	–2032,50	–910,12
Затраты на рекламу – X <sub>2</sub>	9,57	2,27	4,22	0,00	4,67	14,46
Индекс потре- бительских расходов – X <sub>5</sub>	15,75	2,47	6,39	0,00	10,42	21,08
Y – пересече- ние	–1471,31	259,77	–5,66	0,00	–2032,50	–910,12

Получено уравнение регрессии, все коэффициенты которого значимы не только при 5%-ном уровне значимости, но и при 1%-ном уровне значимости:

$$\hat{y}_i = -1471,31 + 9,57x_2 + 15,75x_5.$$

(2,27)      (2,47)



## 4.2. Оценка параметров модели.

### Экономическая интерпретация коэффициентов регрессии

В результате применения различных подходов к выбору факторов пришли к выводу о необходимости включения в модель двух факторов – **Затраты на рекламу** и **Индекс потребительских расходов**.

Выполняя матричные вычисления по формуле  $A = (X'X)^{-1} X'Y$ , естественно, получим такое же уравнение регрессии, как и при использовании инструмента **Регрессия** в **Анализе данных** (рис. 4.2). Уравнение зависимости объема реализации от затрат на рекламу и индекса потребительских расходов можно записать в следующем виде:

$$\hat{y}_i = -1471,31 + 9,57x_2 + 15,75x_5$$

	A	B	C	D	E	F	G
1	ВЫВОД ИТОГОВ						
2							
3	<b>Регрессионная статистика</b>						
4	Множественный R	0,927					
5	R-квадрат	0,859					
6	Нормированный R-квадрат	0,837					
7	Стандартная ошибка	41,473					
8	Наблюдения	16					
9							
10	<b>Дисперсионный анализ</b>						
11		df	SS	MS	F	Значимость F	
12	Регрессия	2	136358,334	68179,167	39,639	2,93428E-06	
13	Остаток	13	22360,104	1720,008			
14	Итого	15	158718,438				
15							
16		<b>Коэффициенты</b>					
17	Y-пересечение	-1471,314	259,766	-5,664	0,000	-2032,505	-910,124
18	X2	9,568	2,266	4,223	0,001	4,673	14,464
19	X5	15,753	2,467	6,386	0,000	10,424	21,082
20							
21							
22							
23	ВЫВОД ОСТАТКА						
24							
25	Наблюдение	Предсказанное Y	Остатки				
26	1	142,247	-16,247				
27	2	124,697	12,303				
28	3	159,237	-11,237				
29	4	242,353	-51,353				
30	5	247,021	26,979				
31	6	307,057	62,943				
32	7	361,200	70,800				

Рис. 4.2. Результаты работы с инструментом Регрессия

Коэффициент регрессии  $\alpha_j$  показывает, на какую величину в среднем изменится результативный признак  $Y$ , если переменную  $x_j$  увеличить на единицу измерения, то есть  $\alpha_j$  является нормативным коэффициентом.

В нашей задаче величина, равная 9,57 (коэффициент при  $x_2$ ), показывает, что при увеличении затрат на рекламу на 1000 руб. объем реализации увеличится на 9,57 тыс. руб., а если на 1% увеличится индекс потребительских расходов, то объем реализации увеличится на 15,75 тыс. руб.

Расчетные значения  $Y$  определяются путем последовательной подстановки в эту модель значений факторов, взятых для каждого наблюдения, или из последней таблицы регрессионного анализа **Вывод остатка** (столбец **Предсказанное  $Y$** ).

### 4.3. Оценка качества модели регрессии

Для оценки качества модели множественной регрессии вычисляют коэффициент детерминации  $R^2$  и коэффициент множественной корреляции (индекс корреляции)  $R$ . Чем ближе к 1 значение этих характеристик, тем выше качество модели.

Значение коэффициентов детерминации и множественной корреляции можно найти в таблице **Регрессионная статистика** (см. рис. 2) или вычислить по формулам:

а) коэффициент детерминации:

$$R^2 = 1 - \frac{\sum e_i^2}{\sum (y_i - \bar{y})^2} = 1 - \frac{22360,104}{158718,438} = 0,859.$$

Коэффициент детерминации показывает долю вариации результативного признака под воздействием изучаемых факторов. Следовательно, около 86% вариации зависимой переменной учтено в модели и обусловлено влиянием факторов, включенных в модель;

б) коэффициент множественной корреляции:

$$R = \sqrt{R^2} = 0,927.$$

Коэффициент множественной корреляции показывает высокую тесноту связи зависимой переменной  $Y$  с двумя включенными в модель объясняющими факторами.

Точность модели оценим с помощью средней ошибки аппроксимации:

$$E_{\text{отн}} = \frac{1}{n} \sum_{i=1}^n \frac{|e_i|}{y_i} \times 100\% = 10,65\%.$$

Модель неточная. Фактические значения объема реализации отличаются от расчетных в среднем на 10,65%.

#### 4.4. Оценка значимости уравнения регрессии и его коэффициентов

Проверку значимости уравнения регрессии произведем на основе  $F$ -критерия Фишера:

$$F = \frac{R^2/k}{(1-R^2)/(n-k-1)} = \frac{0,859/2}{(1-0,859)/(16-2-1)} = 39,6.$$

Значение  $F$ -критерия Фишера можно найти в таблице **Дисперсионный анализ** протокола Excel (см. рис. 4.2).

Табличное значение  $F$ -критерия при доверительной вероятности  $\alpha = 0,95$  и числе степеней свободы, равном  $\nu_1 = k = 2$  и  $\nu_2 = n - k - 1 = 16 - 2 - 1 = 13$ , составляет 3,81.

Поскольку  $F_{\text{расч}} > F_{\text{табл}}$ , уравнение регрессии следует признать значимым, то есть его можно использовать для анализа и прогнозирования.

Оценку значимости коэффициентов полученной модели, используя результаты отчета Excel, можно осуществить тремя способами.

Коэффициент уравнения регрессии признается значимым в том случае, если:

1) наблюдаемое значение  $t$ -статистики Стьюдента для этого коэффициента больше, чем критическое (табличное) значение статистики Стьюдента (для заданного уровня значимости, например  $\alpha = 0,05$ , и числа степеней свободы  $df = n - k - 1$ , где  $n$  – число наблюдений, а  $k$  – число факторов в модели);

2)  $P$ -значение  $t$ -статистики Стьюдента для этого коэффициента меньше, чем уровень значимости, например,  $\alpha = 0,05$ ;

3) доверительный интервал для этого коэффициента, вычисленный с некоторой доверительной вероятностью (например, 95%), не содержит ноль внутри себя, то есть нижняя 95% и верхняя 95% границы доверительного интервала имеют одинаковые знаки.

Значимость коэффициентов  $\hat{a}_1$  и  $\hat{a}_2$  проверим по второму и третьему способам, используя данные рис. 4.2:

$$P\text{-значение}(\hat{a}_1) = 0,00 < 0,01 < 0,05.$$

$$P\text{-значение}(\hat{a}_2) = 0,00 < 0,01 < 0,05.$$

Следовательно, коэффициенты  $\hat{a}_1$  и  $\hat{a}_2$  значимы при 1%-ном уровне, а тем более при 5%-ном уровне значимости.

Нижние и верхние 95% границы доверительного интервала имеют одинаковые знаки (см. рис. 4.2), следовательно, коэффициенты  $\hat{a}_1$  и  $\hat{a}_2$  значимы.

#### **4.5. Определение объясняющей переменной, от которой может зависеть дисперсия случайных возмущений.**

##### **Проверка выполнения условия гомоскедастичности остатков по тесту Гольдфелда–Квандта**

При проверке предпосылки МНК о гомоскедастичности остатков в модели множественной регрессии следует вначале определить, по отношению к какому из факторов дисперсия остатков более всего нарушена. Это можно сделать в результате визуального исследования графиков остатков, построенных по каждому из факторов, включенных в модель (рис 4.3). Та из объясняющих переменных, от которой больше зависит дисперсия случайных возмущений, и будет упорядочена по возрастанию фактических значений при проверке теста Гольдфелда–Квандта.

Для двухфакторной модели нашего примера графики остатков относительно каждого из двух факторов имеют вид, представленный на рис. 4.3 (эти графики легко получить в отчете, который формируется в результате использования инструмента **Регрессия** в пакете **Анализ данных**).

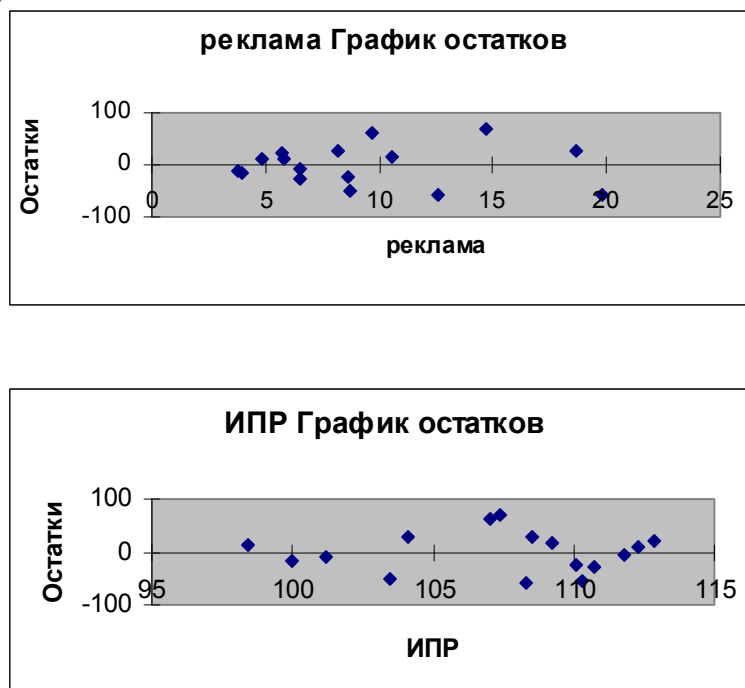


Рис. 4.3. Графики остатков по каждому из факторов двухфакторной модели

Из графиков, представленных на рис. 4.3, видно, что дисперсия остатков более всего нарушена по отношению к фактору **Затраты на рекламу**.

Проверим наличие гомоскедастичности в остатках двухфакторной модели на основе *теста Гольдфелда–Квандта*.

1. Упорядочим переменные  $Y$  и  $X_3$  по возрастанию фактора  $X_2$  (в Excel для этого можно использовать команду **Данные – Сортировка по возрастанию  $X_2$** ):

## Исходные данные

$Y$	$X_2$	$X_5$
Объем реализации	Затраты на рекламу	Индекс потребительских расходов
126	4,0	100,0
137	4,8	98,4
148	3,8	101,2
191	8,7	103,5
274	8,2	104,1
370	9,7	107,0
432	14,7	107,4
445	18,7	108,5
367	19,8	108,3
367	10,6	109,2
321	8,6	110,1
307	6,5	110,7
331	12,6	110,3
345	6,5	111,8
364	5,8	112,3
384	5,7	112,9

Данные, отсортированные по возрастанию  $X_2$ 

$Y$	$X_2$	$X_5$
148	3,8	101,2
126	4,0	100,0
137	4,8	98,4
384	5,7	112,9
364	5,8	112,3
307	6,5	110,7
345	6,5	111,8
274	8,2	104,1
321	8,6	110,1
191	8,7	103,5
370	9,7	107,0
367	10,6	109,2
331	12,6	110,3
432	14,7	107,4
445	18,7	108,5
367	19,8	108,3

2. Уберем из середины упорядоченной совокупности  $C = 1/4 \cdot n = 1/4 \cdot 16 = 4$  значения. В результате получим две совокупности соответственно с малыми и большими значениями  $X_2$ .

3. Для каждой совокупности выполним расчеты:

Уравнения	Y	X <sub>2</sub>	X <sub>5</sub>	Y <sub>p</sub>	e	e <sup>2</sup>
	148	3,8	101,2	157,9192	-9,91918	98,39019
Y = -1588,77 +	126	4,0	100,0	138,2998	-12,29980	151,28460
+ 4,458X <sub>1</sub> +	137	4,8	98,4	114,5179	22,48206	505,44280
+ 17,09X <sub>2</sub>	384	5,7	112,9	366,3700	17,62997	310,81580
	364	5,8	112,3	356,5603	7,439672	55,34873
	307	6,5	110,7	332,3327	-25,33270	641,74750
<b>Сумма</b>						<b>1 763,03000</b>
	370	9,7	107,0	390,6914	-20,69140	428,13250
Y = 2333,286 +	367	10,6	109,2	354,0009	12,99911	168,97680
+ 4,64X <sub>1</sub> -	331	12,6	110,3	342,8479	-11,84790	140,37320
- 18,576X <sub>2</sub>	432	14,7	107,4	406,4619	25,53808	652,19360
	445	18,7	108,5	404,5893	40,41071	1 633,02600
	367	19,8	108,3	413,4086	-46,40860	2 153,76000
<b>Сумма</b>						<b>5 176,46200</b>

Результаты данной таблицы получены с помощью инструмента **Регрессия** поочередно к каждой из полученных совокупностей.

4. Найдем отношение полученных остаточных сумм квадратов (в числителе должна быть большая сумма):

$$F = 5176,462/1763,03 = 2,936117.$$

5. Вывод о наличии гомоскедастичности остатков делаем с помощью  $F$ -критерия Фишера с уровнем значимости  $\alpha = 0,05$  и двумя одинаковыми степенями свободы  $k_1 = k_2 = \frac{n - C - 2 \cdot p}{2} = \frac{16 - 4 - 2 \cdot 3}{2} = 3$ ,

где  $p$  – число параметров уравнения регрессии:

$$F_{\text{табл}}(0,05; 3; 3) = 9,28.$$

Так как  $F_{\text{табл}} > R$ , то подтверждается гомоскедастичность в остатках двухфакторной регрессии.

#### 4.6. Оценка влияния факторов, включенных в модель, на объем реализации

Учитывая, что коэффициент регрессии невозможно использовать для непосредственной оценки влияния факторов на зависимую переменную из-за различия единиц измерения и разной колеблемости факторов, используем коэффициенты эластичности и бета-коэффициенты:

$$\begin{aligned}\Theta_j &= a_j \times \bar{x}_j / \bar{y} \\ \Theta_2 &= 9,568 \times 9,294 / 306,813 = 0,2898; \\ \Theta_5 &= 15,7529 \times 107,231 / 306,813 = 5,506.\end{aligned}$$

Коэффициент эластичности показывает, на сколько процентов изменяется зависимая переменная при изменении фактора на один процент:

$$\begin{aligned}\beta_j &= a_j \times S_{xj} / S_y \\ \beta_2 &= 9,568 \times 4,913 / 102,865 = 0,457; \\ \beta_5 &= 15,7529 \times 4,5128 / 102,865 = 0,691.\end{aligned}$$

Бета-коэффициент с математической точки зрения показывает, на какую часть величины среднеквадратического отклонения меняется среднее значение зависимой переменной с изменением независимой переменной на одно среднеквадратическое отклонение при фиксированных на постоянном уровне значениях остальных независимых переменных. Это означает, что при увеличении затрат на рекламу на 4,91 тыс. руб. объем реализации увеличится на 47 тыс. руб. ( $0,457 \times 102,865$ ).

Среднеквадратическое отклонение затрат на рекламу, равное 4,91, можно вычислить с помощью функции СТАНДОТКЛОН.

Долю влияния фактора в суммарном влиянии всех факторов можно оценить по величине дельта-коэффициентов  $\Delta_j$ :

$$\begin{aligned}\Delta_j &= r_{y,x_j} \cdot \beta_j / R^2 \\ \Delta_2 &= 0,646 \cdot 0,457 / 0,859 = 0,344; \\ \Delta_5 &= 0,816 \cdot 0,691 / 0,859 = 0,656.\end{aligned}$$

**Вывод:** на объем реализации более сильное влияние оказывает фактор **Индекс потребительских расходов**.



#### 4.7. Прогнозирование объема реализации на два месяца вперед

Прогнозируемое значение переменной получается при подстановке в уравнение регрессии ожидаемых значений объясняющих факторов  $X$ .

В нашей задаче необходимо построить прогноз объема реализации на два месяца вперед ( $Y_{17}$ ,  $Y_{18}$ ).

Сначала найдем прогнозные значения факторов  $X_2$  (затраты на рекламу) и  $X_5$  (индекс потребительских расходов).

Так как исходные данные представлены временными рядами, то для получения прогнозных значений факторов  $X_{2,17}$ ,  $X_{5,17}$  и  $X_{2,18}$ ,  $X_{5,18}$  воспользуемся инструментом **Мастер диаграмм** Excel для построения трендовых моделей затрат на рекламу и индекса потребительских расходов.

Для фактора  $X_2$  (затраты на рекламу) выбрана модель:

$$X_2 = 12,83 - 11,616t + 4,319t^2 - 0,552t^3 + 0,0292t^4 - 0,0006t^5,$$

по которой получен прогноз на два месяца вперед:  $X_{2,17} = 5,75$ ,  $X_{2,18} = 4,85$ .

График модели временного ряда **Затраты на рекламу** приведен на рис. 4.

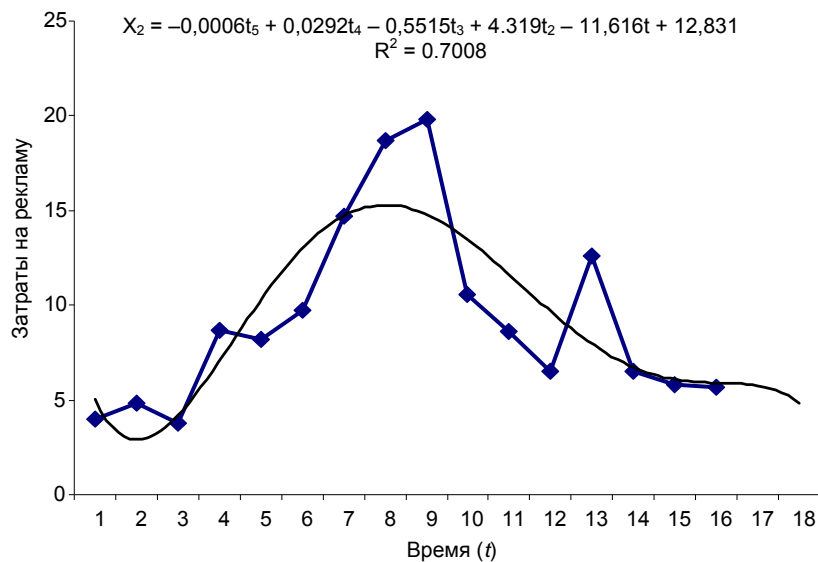


Рис. 4.4. Прогноз показателя «Затраты на рекламу»  
с помощью Мастера диаграмм

Для временного ряда **Индекс потребительских расходов** в качестве аппроксимирующей функции выбран полином второй степени (парабола), по которой построен прогноз на два шага вперед. На рис. 5 приведен результат построения тренда для временного ряда **Индекс потребительских расходов**. Прогноз этого фактора на два месяца получен по модели

$$X_5 = -0,0488t^2 + 1,739t + 97,008;$$

$$X_{5,17} = 112,468, X_{5,18} = 112,488.$$

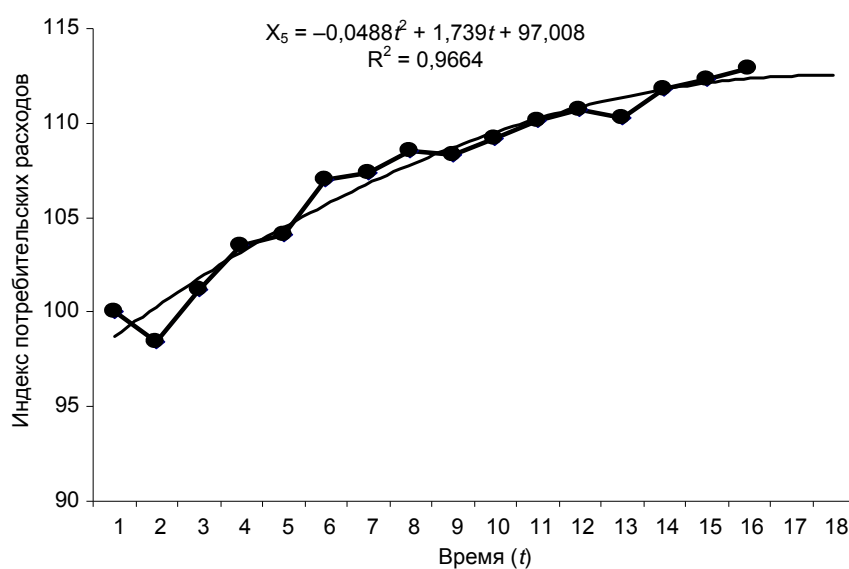


Рис. 5. Прогноз показателя «Индекс потребительских расходов» с помощью Мастера диаграмм

Для получения прогнозных оценок зависимой переменной по модели

$$Y = -1471,438 + 9,568X_2 + 15,753X_5$$

подставим в нее найденные прогнозные значения факторов  $X_2$  и  $X_5$ :

$$Y_{t=17} = -1471,438 + 9,568 \times 5,75 + 15,753 \times 112,468 = 355,398;$$

$$Y_{t=18} = -1471,438 + 9,568 \times 4,85 + 15,753 \times 112,488 = 347,102.$$

Доверительный интервал прогноза будет иметь следующие границы:

- а) верхняя граница прогноза:  $Y_{\text{прогн}(n+1)} + U(l)$ ,
- б) нижняя граница прогноза:  $Y_{\text{прогн}(n+1)} - U(l)$ ,

где  $U = \sigma_e \cdot t_{\text{кр}} \cdot \sqrt{1 + X_{\text{пр}}^T (X^T X)^{-1} X_{\text{пр}}}$ .

Стандартная ошибка ( $\sigma_e = 41,473$ ) получена из таблицы *Регрессионная статистика* (см. рис. 2). Значение  $t_{\text{кр}}$  ( $t_{\text{кр}} = 1,77$ ) получено с помощью функции СТЬЮДРАСПРОБР(0.1;13) для выбранной вероятности 90% с числом степеней свободы, равным 13.

Для первого шага прогноза:

$$l = 1;$$

$$X_{\text{пр}}^T = (1; 5,75; 112,468);$$

$$(X^T X)^{-1} = \begin{pmatrix} 39,2314 & 0,06752 & -0,3711 \\ 0,06752 & 0,00299 & -0,00088 \\ -0,3711 & -0,00088 & 0,00354 \end{pmatrix}$$

$$u(1) = 81,45.$$

Для второго шага прогноза:

$$l = 2;$$

$$X_{\text{пр}}^T = (1; 4,85; 112,488);$$

$$u(2) = 82,47.$$

Результаты прогнозных оценок модели регрессии представим в табл. 4.8.

**Таблица 4.8. Прогнозные оценки объема реализации**

Таблица прогнозов ( $p = 90\%$ )			
Упреждение	Прогноз	Нижняя граница	Верхняя граница
1	355,398	273,94	436,85
2	344,179	261,71	426,65

С вероятностью 90% объем реализации в первом прогнозируемом месяце составит от 273,94 млн руб. до 436,85 млн руб., а во втором – попадет в интервал от 261,71 млн руб. до 426,65 млн руб.

## 5. Задания для выполнения контрольной работы

На основании данных, приведенных в табл. 5.1:

**1.** Постройте диаграммы рассеяния, представляющие собой зависимости  $Y$  от каждого из факторов  $X$ . Сделайте выводы о характере взаимосвязи переменных.

**2.** Осуществите двумя способами выбор факторных признаков для построения регрессионной модели:

а) на основе анализа матрицы коэффициентов парной корреляции, включая проверку гипотезы о независимости объясняющих переменных (тест на выявление мультиколлинеарности Фаррара–Глоубера);

б) с помощью пошагового отбора методом исключения.

**3.** Постройте уравнение множественной регрессии в линейной форме с выбранными факторами. Дайте экономическую интерпретацию коэффициентов модели регрессии.

**4.** Дайте сравнительную оценку силы связи факторов с результатом с помощью коэффициентов эластичности,  $\beta$ - и  $\Delta$ -коэффициентов.

**5.** Рассчитайте параметры линейной парной регрессии для наиболее подходящего фактора  $X_j$ .

**6.** Оцените качество построенной модели с помощью коэффициента детерминации, средней относительной ошибки аппроксимации и  $F$ -критерия Фишера.

**7.** Проверьте выполнение условия гомоскедастичности.

**8.** Используя результаты регрессионного анализа, ранжируйте компании по степени эффективности.

**9.** Осуществите прогнозирование среднего значения показателя  $Y$  при уровне значимости  $\alpha = 0,1$ , если прогнозное значение фактора  $X_j$  составит 80% от его максимального значения. Представьте на графике фактические данные  $Y$ , результаты моделирования, прогнозные оценки и границы доверительного интервала.

**10.** Составьте уравнения нелинейной регрессии:

а) гиперболической;

б) степенной;

в) показательной.

**11.** Приведите графики построенных уравнений регрессии.

**12.** Для нелинейных моделей найдите коэффициенты детерминации и средние относительные ошибки аппроксимации. Сравните модели по этим характеристикам и сделайте вывод о лучшей модели.

Варианты заданий для контрольной работы представлены в табл. 5.2.

Таблица 5.1. Добыча сырой нефти и природного газа,  
предоставление услуг в этих областях (данные за 2009 г.)

ОАО по добыче сырой нефти и природного газа; предоставление услуг в этих областях	№ п/п	Прибыль (убыток)	Долгосрочные обязательства	Краткосрочные обязательства	Оборотные активы	Основные средства	Дебиторская задолженность (краткосрочная)	Запасы готовой продукции и товаров для переработки
		Y	X <sub>1</sub>	X <sub>2</sub>	X <sub>3</sub>	X <sub>4</sub>	X <sub>5</sub>	X <sub>6</sub>
Аганнефтегазгеология, открытое акционерное общество, многопрофильная компания	1	1 440 075	61 749	1 007 355	4 920 199	5 165 712	3 490 541	31 365
Азнамаевский горизонт, открытое акционерное общество	2	5 146	17 532	5 8110	50 798	19 595	23 014	0
Акмай, открытое акционерное общество	3	13 612	20 268	51 271	18 903	81 072	8 678	84
Аксоль, открытое акционерное общество, производственно-коммерческая фирма	4	964	211	5 827	13 398	8 446	4 821	0
Акционерная нефтяная компания «Башнефть», открытое акционерное общество	5	19 513 178	52 034 182	2 411 352	63 269 757	47 002 385	23 780 450	1 696 853
АПРОСА-Газ, открытое акционерное общество	6	28 973	602 229	74 839	367 880	1 545 052	204 181	19 474
Арктическая газовая компания, открытое акционерное общество	7	-780 599	311 268	15 737 048	3 933 712	740 437	1 456 438	176

ОАО по добыче сырья нефти и природного газа; предоставление услуг в этих областях	№ п/п	Прибыль (убыток)	Долгосрочные обязательства	Краткосрочные обязательства	Оборотные активы	Основные средства	Дебиторская задолженность (краткосрочная)	Запасы готовой продукции и товаров для переработки
			X <sub>1</sub>	X <sub>2</sub>	X <sub>3</sub>	X <sub>4</sub>	X <sub>5</sub>	X <sub>6</sub>
		Y						
Барьеганнефтегаз, открытое акционерное общество	8	2 598 165	464 651	4 381 403	5 910 831	11 925 177	5 566 412	127 937
Белкамнефть, открытое акционерное общество	9	628 091	2 14 411	3 728 587	5 325 806	2 580 485	4 285 041	73 823
Белорусское управление по повышению нефтеотдачи пластов и капитальному ремонту скважин, открытое акционерное общество	10	29 204	12 039	7 38 811	705 877	269 908	624 393	130
Битран, открытое акционерное общество	11	1 945 560	9 670	7 16 648	2 964 277	229 855	2 918 345	39 667
Богородскнефть, открытое акционерное общество	12	366 170	287 992	239 076	624 661	349 643	484 537	5 733
Братскгаз, открытое акционерное общество	13	-20 493	1 105 293	8 855	46 728	934 881	9 865	3 319
Булгарнефть, открытое акционерное общество	14	381 558	27 265	265 569	582 581	697 664	196 045	5 763
Варьеганнефть, открытое акционерное общество	15	1 225 908	431 231	1 525 379	3 463 511	2 231 651	1 095 263	430 844
Верхнеюганнефтегаз, открытое акционерное общество	16	3 293 989	37 315 847	8 556 455	5 891 049	23 170 344	2 477 424	38 133

ОАО по добыче сырья нефти и природного газа; предоставление услуг в этих областях	№ п/п	Прибыль (убыток)	Долгосрочные обязательства	Краткосрочные обязательства	Оборотные активы	Основные средства	Дебиторская задолженность (краткосрочная)	Запасы готовой продукции и товаров для перепродажи
		Y	X <sub>1</sub>	X <sub>2</sub>	X <sub>3</sub>	X <sub>4</sub>	X <sub>5</sub>	X <sub>6</sub>
Восточная транснациональная компания, открытое акционерное общество	17	4 16 616	2 122 138	258 120	299 286	3 509 537	48 174	28 393
Восточно-Сибирская нефтегазовая компания, открытое акционерное общество	18	-564 258	1 395 080	7 958 766	801 276	1 290 245	286 058	236 642
Геолого-разведочный исследовательский центр, открытое акционерное общество	19	221 194	13 429	105 123	257 633	607 249	72 854	4548
Грознефтегаз, открытое акционерное общество	20	701 035	75 554	497 028	1 566 040	4 616 250	1 304 084	8 773
Губкинский газоперерабатывающий комплекс, открытое акционерное общество	21	62 200	22 195	1 659 245	528 912	991 114	294 575	0
Дагнефтегаз, открытое акционерное общество	22	123 440	12 350	84 026	167 297	438 262	44 889	24 866
Елабуганефть, открытое акционерное общество	23	55 528	14 686	137 348	52 042	75 442	24 275	3 949
Иделюйл, открытое акционерное общество	24	422 070	52 443	662 299	188 662	1 269 731	140 535	8 212

ОАО по добыче сырья нефти и природного газа; предоставление услуг в этих областях	№ п/п	Прибыль (убыток)	Долгосрочные обязательства	Краткосрочные обязательства	Оборотные активы	Основные средства	Дебиторская задолженность (краткосрочная)	Запасы готовой продукции и товаров для переработки
		Y	X <sub>1</sub>	X <sub>2</sub>	X <sub>3</sub>	X <sub>4</sub>	X <sub>5</sub>	X <sub>6</sub>
Избешанефть, открытое акционерное общество	25	-468	239 255	29 880	130 350	10 870	114 444	940
Инвестиционная нефтяная компания, открытое акционерное общество	26	225 452	1 292	87 112	585 017	227 132	272 147	0
Инга, открытое акционерное общество	27	-61 237	924 951	299 733	344 398	110 970	76 561	11 218
Кабалнефтепромпром, открытое акционерное общество	28	-540	0	46 139	36 641	21 278	25 017	127
Калининграднефть, открытое акционерное общество	29	40 588	1 638	22 683	215 106	139 209	18 072	7 569
Камчатгазпром, открытое акционерное общество	30	53 182	54 758	1 909 328	998 875	113 113	496 994	0
Кировское нефтегазодобывающее управление, открытое акционерное общество	31	-210	8	16 191	1 702	12 685	602	46
Когалымнефтепрогресс, открытое акционерное общество	32	63 058	235 731	563 481	807 686	873 886	474 612	0
Комнедра, открытое акционерное общество	33	1 197 196	2 232 742	1 083 829	1 567 998	2 307 478	1 040 387	25 862



ОАО по добыче сырой нефти и природного газа; предоставление услуг в этих областях	№ п/п	Прибыль (убыток)	Долгосрочные обязательства	Краткосрочные обязательства	Оборотные активы	Основные средства	Дебиторская задолженность (краткосрочная)	Запасы готовой продукции и товаров для перепродажи
			X <sub>1</sub>	X <sub>2</sub>	X <sub>3</sub>	X <sub>4</sub>	X <sub>5</sub>	X <sub>6</sub>
Кондурчанефть, открытое акционерное общество	34	221 177	4 682	40 664	128 256	331 954	55 155	1 260
Корпорация «Юганефть», открытое акционерное общество	35	1 548 768	84 262	413 994	7 720 298	1 138 707	7 613 662	14 716
Краснодарское опытно-экспериментальное управление по повышению нефтеотдачи пластов и капитальному ремонту скважин, открытое акционерное общество	36	-33 030	106	52 575	14 412	16 705	5 038	0
Ленинградсланец, открытое акционерное общество	37	-34 929	103 567	1 769 300	921 832	393 717	61 353	833 099
Меллянефть, открытое акционерное общество	38	115 847	275 386	432 312	233 340	517 290	122 062	6 824
МНКТ, общество с ограниченной ответственностью	39	35 198	20 624	169 155	361 672	484 228	168 314	3 227
Мохтикнефть, открытое акционерное общество	40	788 567	33 879	647 914	458 233	402 613	317 153	14 021
Научно-производственное объединение «Спецэлектро-механика», открытое акционерное общество	41	309 053	99 670	211 624	619 452	18 776	212 882	1 909

ОАО по добыче сырья нефти и природного газа; предоставление услуг в этих областях	№ п/п	Прибыль (убыток)	Долгосрочные обязательства	Краткосрочные обязательства	Оборотные активы	Основные средства	Дебиторская задолженность (краткосрочная)	Запасы готовой продукции и товаров для переработки
			X <sub>1</sub>	X <sub>2</sub>	X <sub>3</sub>	X <sub>4</sub>	X <sub>5</sub>	X <sub>6</sub>
		Y						
Научно-производственное предприятие «Бурсервис», открытое акционерное общество	42	8 552	257	99 815	119 434	12 381	63 550	2 558
НГДУ «Пензанефть», открытое акционерное общество	43	173 079	6120	114 223	257 140	176 126	147 549	16 197
Негуснефть, открытое акционерное общество	44	1 227 017	33 757	1 930 517	4 215 454	2 063 285	171 162	63 810
Немецкая нефтяная компания, открытое акционерное общество	45	701 728	381 050	335 238	324 968	59 353	237 083	3 886
Нефтебурсервис, открытое акционерное общество	46	17 927	53 260	101 834	81 960	84 818	73 343	963
Нефтегазовая компания «Славнефть», открытое акционерное общество	47	2 557 698	4 537 040	21 786 237	35 232 071	3 841 845	33 477 251	26 578
Нефтерафедка, открытое акционерное общество	48	0	194 091	64 889	76 430	33 112	15 161	7
Нефть, открытое акционерное общество	49	5 406	1 185	27 941	21 132	38 560	7 540	6 465

ОАО по добыче сырой нефти и природного газа; предоставление услуг в этих областях	№ п/п	Прибыль (убыток)	Долгосрочные обязательства	Краткосрочные обязательства	Оборотные активы	Основные средства	Дебиторская задолженность (краткосрочная)	Запасы готовой продукции и товаров для перепродажи
		У	X <sub>1</sub>	X <sub>2</sub>	X <sub>3</sub>	X <sub>4</sub>	X <sub>5</sub>	X <sub>6</sub>
Нефтьинвест, открытое акционерное общество	50	40 997	101 706	39 653	79 930	178 604	58 762	1 035
Нефтяная акционерная компания «АКИ-ОТЪР», открытое акционерное общество	51	1 580 624	9 285 230	1 476 613	1 553 508	6 546 853	259 519	13 516
Нефтяная компания «Магма», открытое акционерное общество	52	9 990 896	1 645 470	5 066 776	26 312 477	2 329 554	7 271 400	391 744
Нефтяная компания «Мангазея», открытое акционерное общество	53	6 649	82 229	1 486 511	972 138	78 526	444 251	24 001
Нефтяная компания «Нефтьтиса», открытое акционерное общество	54	22 868	3	76 455	132 783	9 067	28 536	0

Источники: <http://www.fira.ru/>

Таблица 5.2. *Варианты заданий для контрольной работы*

Вариант	Номер строки	Факторы	Вариант	Номер строки	Факторы
1	1-50	$X_1, X_2, X_3$	31	3-52	$X_1, X_3, X_5$
2	1-50	$X_1, X_2, X_4$	32	3-52	$X_1, X_5, X_4$
3	1-50	$X_1, X_2, X_6$	33	3-52	$X_1, X_5, X_6$
4	1-50	$X_2, X_3, X_4$	34	3-52	$X_3, X_5, X_4$
5	1-50	$X_2, X_3, X_6$	35	3-52	$X_3, X_5, X_6$
6	1-50	$X_2, X_4, X_6$	36	3-52	$X_4, X_5, X_6$
7	1-50	$X_1, X_3, X_5$	37	4-53	$X_1, X_2, X_3$
8	1-50	$X_1, X_5, X_4$	38	4-53	$X_1, X_2, X_4$
9	1-50	$X_1, X_5, X_6$	39	4-53	$X_1, X_2, X_6$
10	1-50	$X_3, X_5, X_4$	40	4-53	$X_2, X_3, X_4$
11	1-50	$X_3, X_5, X_6$	41	4-53	$X_2, X_3, X_6$
12	1-50	$X_4, X_5, X_6$	42	4-53	$X_2, X_4, X_6$
13	2-51	$X_1, X_2, X_3$	43	4-53	$X_1, X_3, X_5$
14	2-51	$X_1, X_2, X_4$	44	4-53	$X_1, X_5, X_4$
15	2-51	$X_1, X_2, X_6$	45	4-53	$X_1, X_5, X_6$
16	2-51	$X_2, X_3, X_4$	46	4-53	$X_3, X_5, X_4$
17	2-51	$X_2, X_3, X_6$	47	4-53	$X_3, X_5, X_6$
18	2-51	$X_2, X_4, X_6$	48	4-53	$X_4, X_5, X_6$
19	2-51	$X_1, X_3, X_5$	49	5-54	$X_1, X_2, X_3$
20	2-51	$X_1, X_5, X_4$	50	5-54	$X_1, X_2, X_4$
21	2-51	$X_1, X_5, X_6$	51	5-54	$X_1, X_2, X_6$
22	2-51	$X_3, X_5, X_4$	52	5-54	$X_2, X_3, X_4$
23	2-51	$X_3, X_5, X_6$	53	5-54	$X_2, X_3, X_6$
24	2-51	$X_4, X_5, X_6$	54	5-54	$X_2, X_4, X_6$
25	3-52	$X_1, X_2, X_3$	55	5-54	$X_1, X_3, X_5$
26	3-52	$X_1, X_2, X_4$	56	5-54	$X_1, X_5, X_4$
27	3-52	$X_1, X_2, X_6$	57	5-54	$X_1, X_5, X_6$
28	3-52	$X_2, X_3, X_4$	58	5-54	$X_3, X_5, X_4$
29	3-52	$X_2, X_3, X_6$	59	5-54	$X_3, X_5, X_6$
30	3-52	$X_2, X_4, X_6$	60	5-54	$X_4, X_5, X_6$

## 6. Задания для выполнения лабораторной работы

### Вариант 1

По данным, полученным от 40 промышленных предприятий одного из регионов, изучается зависимость объема выпуска продукции  $Y$  (млн руб.) от четырех факторов:

$X_1$  – численность промышленно-производственного персонала, чел.;

$X_2$  – среднегодовая стоимость основных фондов, млн руб.;

$X_3$  – электровооруженность 1 чел.-ч., кВт·ч;

$X_4$  – прибыль от реализации продукции, млн руб.

№	$Y$	$X_1$	$X_2$	$X_3$	$X_4$
1	62 240	2 890	30 434	13,9	1 258
2	88 569	4 409	162 229	55,3	16 304
3	3 118	210	2 370	5,7	250
4	186 256	5 436	802 725	87,2	8 306
5	56 262	1 559	48 075	56,2	17 663
6	19 216	940	18 894	27,6	2 861
7	16 567	1 197	19 621	31,1	734
8	203 456	8 212	336 472	60,5	42 392
9	13 425	459	9 843	33,5	4 740
10	31 163	1 405	44 217	35,1	9 469
11	30 109	1 575	34 142	26,5	5 206
12	14 781	964	6 412	4,4	-1 437
13	41 138	1 866	39 208	24,9	9 948
14	69 202	4 419	80 694	13,2	-9 135
15	10 201	802	6 714	14,9	857
16	75 282	2 600	28 148	2,4	12 729
17	47 064	1 110	11 861	5,8	8 887
18	57 342	1 147	63 273	50,4	15 503
19	32 900	864	16 144	4,9	7 960
20	18 135	763	14 758	25,9	2 522
21	29 589	1 003	27 642	43,5	4 412
22	22 604	1 680	23 968	3,1	3 304
23	1 878	77	250	0,6	172
24	49 378	2 505	85 105	43,1	6 264
25	6 896	1 556	12 612	0,7	1 745
26	3 190	442	478	0,2	116
27	5 384	305	3 667	6,6	1 225
28	17 668	875	16 250	4,8	4 652
29	24 119	1 142	28 266	48,7	5 278
30	16 649	710	8 228	17,8	5 431

№	Y	X <sub>1</sub>	X <sub>2</sub>	X <sub>3</sub>	X <sub>4</sub>
31	6 925	162	1 481	3,0	2 213
32	5 394	331	3 797	20,8	1 454
33	4 330	217	2 950	23,9	1 764
34	1 087 322	46 142	972 349	27,5	163 420
35	284 154	10 469	272 622	33,2	27 506
36	363 204	16 063	267 652	23,1	72 782
37	162 216	6 139	128 731	35,1	35 218
38	97 070	4 560	108 549	32,5	10 035
39	55 410	2 797	60 209	37,2	1 021
40	32 654	1 922	60 669	33,8	–5 192

- ?** 1. Постройте матрицу парных коэффициентов корреляции. Выполните тест Фаррара–Глоубера на мультиколлинеарность. Обоснуйте отбор факторов для регрессионного анализа.
2. Постройте уравнение регрессии. Дайте экономическую интерпретацию коэффициентов регрессии.
3. Приведите график остатков. Проверьте условие гомоскедастичности остатков.
4. Рассчитайте прогнозное значение результата, если прогнозные значения факторов составляют 80% от их максимальных значений.
5. Рассчитайте ошибки и доверительный интервал прогноза для уровня значимости 5% или 10% ( $\alpha_1 = 0,05$ ,  $\alpha_2 = 0,1$ ).

### Вариант 2

По данным, представленным в таблице, исследуется зависимость между величиной накладных расходов Y (млн руб.) 40 строительных организаций и следующими тремя основными факторами:

- X<sub>1</sub> – объем выполненных работ, млн руб.;
- X<sub>2</sub> – численность рабочих, чел.;
- X<sub>3</sub> – фонд заработной платы, млн руб.

№	Y	X <sub>1</sub>	X <sub>2</sub>	X <sub>3</sub>
1	5,7	26,9	1 276	12,250
2	5,0	24,5	975	10,627
3	4,5	18,4	869	6,865
4	4,0	18,1	757	6,964
5	4,4	18,1	740	7,622
6	3,5	17,9	699	6,291
7	3,5	15,7	840	7,980
8	3,8	14,2	744	6,770
9	5,1	13,3	725	7,105
10	3,4	15,0	670	5,762
11	4,1	14,7	622	6,096
12	4,1	13,3	566	6,056
13	3,1	14,6	518	4,921
14	2,8	11,7	510	4,131
15	2,1	10,6	452	4,384
16	2,5	10,0	447	4,157
17	2,0	9,0	497	4,324
18	2,4	9,5	428	4,023
19	2,3	7,0	381	3,315
20	2,4	9,1	385	3,619
21	2,5	6,8	412	3,461
22	2,2	5,5	293	2,139
23	1,6	5,1	284	2,244
24	3,4	12,2	514	3,958
25	2,7	11,0	407	3,337
26	3,2	9,3	577	3,676
27	2,9	5,9	265	2,120
28	4,8	25,9	977	10,649
29	3,7	23,5	724	6,806
30	4,4	19,8	983	9,240
31	3,7	18,8	828	8,860
32	4,8	19,1	766	7,354
33	3,7	18,8	615	5,289
34	3,6	17,4	583	5,830
35	4,0	14,1	591	6,265
36	3,8	13,8	593	5,396
37	3,7	13,7	611	5,194
38	4,1	13,8	562	4,608
39	2,4	13,9	488	5,856
40	2,5	10,6	740	7,326

- [?]** 1. Постройте матрицу парных коэффициентов корреляции. Выполните тест Фаррара–Глоубера на мультиколлинеарность. Обоснуйте отбор факторов для регрессионного анализа.
2. Постройте уравнение множественной регрессии.
3. Оцените качество уравнения регрессии с помощью коэффициента детерминации  $R^2$ , индекса корреляции. Оцените точность модели.
4. Оцените статистическую значимость уравнения регрессии, используя  $F$ -критерий Фишера ( $\alpha = 0,05$ ), и статистическую значимость параметров регрессии, используя критерий Стьюдента.
5. Дайте сравнительную оценку силы связи факторов с результатом с помощью коэффициентов эластичности,  $\beta$ - и  $\Delta$ -коэффициентов.
6. Рассчитайте прогнозное значение результата, если прогнозные значения факторов составляют 70% от их максимальных значений.

### Вариант 3

По предприятию легкой промышленности оценивается эффективность использования активов. Для этой цели анализируется зависимость квартальной выручки от продаж ( $Y$ ) от среднеквартальной стоимости внеоборотных и оборотных активов (переменные  $X_1$  и  $X_2$  соответственно). Имеются данные за три года, млн руб.:

Год	Квартал	$Y$	$X_1$	$X_2$
2006	1	412	110	112
	2	456	145	132
	3	434	152	96
	4	476	155	144
2007	1	482	170	192
	2	467	185	164
	3	478	200	150
	4	489	215	192
2008	1	503	230	176
	2	492	245	138
	3	522	260	150
	4	564	245	212



**[?]** 1. Постройте матрицу парных коэффициентов корреляции между исследуемыми переменными. Проверьте значимость коэффициентов корреляции. Проверьте факторы на коллинеарность.

2. Постройте уравнение линейной регрессии выручки от продаж в зависимости от среднеквартальной стоимости внеоборотных и оборотных активов.

3. Проверьте статистическую значимость уравнения и его коэффициентов на уровне значимости  $\alpha = 0,05$ . Сделайте вывод о существенном либо несущественном влиянии изменения стоимости активов на изменение выручки от продаж.

4. Оцените точность модели.

5. Дайте экономическую интерпретацию уравнения регрессии и оцените степень влияния каждой из групп активов на выручку от продаж (вычислите коэффициенты эластичности,  $\beta$ - и  $\Delta$ -коэффициенты).

6. Постройте прогноз квартальной выручки от продаж на два следующих квартала.

#### Вариант 4

По машиностроительному предприятию оценивается эффективность использования материальных и трудовых ресурсов. Для этого анализируется зависимость годового объема выпущенной продукции  $Y$  (млн руб.) от среднегодовой стоимости основных средств  $X_1$  (млн руб.) и среднегодовой численности работников предприятия  $X_2$  (чел.). Имеются данные за десять лет:

Год	$Y$	$X_1$	$X_2$
1999	405,3	41,8	1305,2
2000	428,1	66,3	1330,1
2001	423,9	69,6	1295,3
2002	433,2	76,8	1302,9
2003	456,5	89,4	1334,1
2004	464,7	95,3	1320,7
2005	542,1	92,9	1303,5
2006	599,9	95,1	1456,9
2007	599,2	122,5	1478,2
2008	576,5	135,9	1390,3

- [?]** 1. Постройте матрицу парных коэффициентов корреляции между исследуемыми переменными. Проверьте факторы на коллинеарность.
2. Постройте уравнение линейной регрессии объема выпущенной продукции без включения фактора времени и с включением фактора времени для учета тренда.
3. Проверьте статистическую значимость уравнения и его коэффициентов на уровне значимости  $\alpha = 0,05$ . Сделайте выводы о существенном либо несущественном влиянии факторных переменных  $X_1$  и  $X_2$  на изменение объема выпущенной продукции и целесообразности включения фактора времени в регрессионную модель.
4. Оцените точность модели.
5. Дайте экономическую интерпретацию уравнения регрессии.
6. Постройте прогноз годового объема выпущенной продукции на следующий год.

### Вариант 5

Исследуется влияние некоторых показателей социально-экономического развития субъектов Центрального федерального округа России на региональный коэффициент смертности. В таблице приведены официальные статистические данные по субъектам Центрального федерального округа за 2005 г.<sup>1</sup>, где:

$Y$  – коэффициент смертности, то есть число умерших за год на 1000 человек населения в 2006 г., ‰;

$X_1$  – индекс (темпы роста) промышленного производства, % к 2004 г.;

$X_2$  – индекс производства сельскохозяйственной продукции, % к 2004 г. (для г. Москвы условно принято за 100 %);

$X_3$  – численность работников малых предприятий, ‰ (чел. на 1000 чел. населения);

$X_4$  – среднемесячная номинальная начисленная заработная плата по региону, тыс. руб.;

$X_5$  – численность населения на 1 января 2005 г., тыс. чел.

<sup>1</sup> Российская газета. – 2006. – 24 марта. – № 60.

Область	Y	X <sub>1</sub>	X <sub>2</sub>	X <sub>3</sub>	X <sub>4</sub>	X <sub>5</sub>
Белгородская	16,0	108,8	115,8	35,4	6,86	1 512
Брянская	19,8	116,0	95,7	25,0	5,24	1 346
Владимирская	20,3	100,2	113,3	43,1	6,07	1 487
Воронежская	18,8	109,6	102,1	53,3	5,60	2 334
Ивановская	22,0	107,6	96,8	36,5	5,37	1 115
Калужская	19,2	105,0	94,7	58,4	6,98	1 022
Костромская	21,0	108,4	100,3	30,1	5,84	717
Курская	19,7	104,0	101,1	29,8	5,65	1199
Липецкая	17,9	102,5	108,2	33,6	7,19	1 190
Московская	17,5	129,6	101,2	61,5	9,51	6 630
Орловская	18,5	110,3	101,7	28,4	5,46	842
Рязанская	20,3	106,2	100,9	49,4	6,22	1 195
Смоленская	21,5	104,3	92,3	26,3	6,30	1 019

**[?] 1.** Постройте матрицу парных коэффициентов линейной корреляции и выявите коллинеарные факторы.

2. Постройте линейную регрессионную модель коэффициента смертности, обосновав отбор факторов. Если из-за коллинеарности факторов невозможно построить уравнение регрессии с полным перечнем факторов, то построьте несколько моделей.

3. Оцените качество построенных моделей.

4. Дайте экономическую интерпретацию параметров лучшего уравнения регрессии и оцените вклад каждого из факторов в вариацию коэффициента смертности с помощью  $\Delta$ -коэффициентов.

5. Постройте три однофакторные нелинейные регрессионные модели зависимой переменной с наиболее подходящим фактором: степенную, гиперболическую и показательную. Сравните качество моделей. Выберите лучшую модель.

Примечание. При проверке статистических гипотез уровень значимости  $\alpha$  примите равным 0,05.

### Вариант 6

По тринадцати супермаркетам исследуется зависимость квартального торгового оборота от размера торговых площадей, района расположения (центральный или периферийный) и формы собственности (муниципальный или частный). Имеются следующие данные:

№ магазина	Торговый оборот, млн руб.	Торговые площади, м <sup>2</sup>	Район расположения	Форма собственности
1	59	2 500	периферийный	муниципальная
2	85	2 172	периферийный	частная
3	127	2 928	центральный	муниципальная
4	178	3 943	центральный	муниципальная
5	156	2 819	центральный	частная
6	122	4 902	периферийный	муниципальная
7	89	4 236	центральный	муниципальная
8	159	5 486	периферийный	муниципальная
9	256	7 186	центральный	частная
10	156	4 501	центральный	частная
11	149	3 495	центральный	муниципальная
12	122	4 562	периферийный	частная
13	178	2 706	центральный	частная

1. Проанализируйте тесноту и направление связи между переменными, отберите факторы для регрессионного анализа.
2. Постройте линейную регрессионную модель торгового оборота магазина, не содержащую коллинеарных факторов. Оцените параметры модели.
3. Существенна ли разница в торговом обороте магазинов:
  - а) расположенных в центральном и периферийных районах города;
  - б) частных и муниципальных?
4. Соответствуют ли остатки регрессии нормальному закону распределения?
5. Выполняется ли условие гомоскедастичности?

6. Спрогнозируйте значение торгового оборота муниципального магазина с торговой площадью 4000 м<sup>2</sup>, расположенного в центральном районе города.

### Вариант 7

По хладокомбинату изучается зависимость месячного объема реализации мороженого от средней цены выпускаемой продукции, затрат на рекламу, среднемесячной температуры воздуха и месячного темпа инфляции. Имеются данные за двенадцать месяцев:

Месяц	Объем реализации, тыс. руб.	Цена, руб.	Затраты на рекламу, тыс. руб.	Температура воздуха, °С	Темп инфляции, %
1	185	8,3	6	2	0,3
2	162	8,3	7	4	0,4
3	182	8,9	5	7	0,3
4	195	10,6	5	10	0,2
5	226	10,7	7	13	0,7
6	279	10,8	22	18	0,9
7	312	12,2	12	22	0,9
8	286	14,2	17	24	0,4
9	212	14,5	22	17	0,1
10	178	13,7	26	13	0,1
11	182	13,3	8	8	0,5
12	173	12,1	4	5	0,9

**?** 1. Проанализируйте тесноту и направление связи между переменными, отберите факторы для регрессионного анализа.

2. Постройте линейную регрессионную модель объема реализации мороженого, не содержащую коллинеарных факторов. Оцените параметры модели. Являются ли уравнение регрессии и его коэффициенты статистически значимыми?

3. Присутствует ли в остатках регрессии автокорреляция первого порядка?

4. Приемлема ли точность регрессионной модели?

5. Изменение какого фактора сильнее всего влияет на изменение объема реализации мороженого? Оцените вклад каждого из факторов в вариацию объема прибыли с помощью  $\Delta$ -коэффициентов.

6. Спрогнозируйте величину объема реализации мороженого на следующий месяц в зависимости от прогнозных значений влияющих факторов.

### Вариант 8

По 14 страховым компаниям имеются данные, характеризующие зависимость объема чистой годовой прибыли от годовых объемов собственных средств, страховых резервов, страховых премий и страховых выплат, тыс. руб.:

№ компании	Годовая прибыль	Собственные средства	Страховые резервы	Страховые премии	Страховые выплаты
1	92	3 444	9 563	11 456	1 659
2	42	2 658	6 354	5 249	2 625
3	186	9 723	10 245	12 968	4 489
4	48	4 526	6 398	7 589	6 896
5	38	5 369	5 692	7 256	5 698
6	74	2 248	6 359	4 963	4 321
7	48	5 671	6 892	7 259	6 692
8	82	4 312	7 256	6 935	756
9	45	2 226	8 256	2 693	5 532
10	46	3 654	5 982	6 324	3 235
11	65	2 635	6 359	7 853	5 325
12	29	2 463	7 532	8 253	6 862
13	34	3 265	5 632	7 564	6 325
14	66	7 546	7 625	9 638	4 569

**?** 1. Постройте линейную регрессионную модель объема чистой годовой прибыли страховой компании, не содержащую коллинеарных факторов. Оцените параметры модели.

2. Являются ли уравнение регрессии и его коэффициенты статистически значимыми?

3. Имеют ли остатки регрессии одинаковую дисперсию?
4. Приемлема ли точность регрессионной модели?
5. Дайте экономическую интерпретацию коэффициентов уравнения регрессии.
6. Изменение какого фактора сильнее всего влияет на изменение объема годовой прибыли?
7. Используя результаты регрессионного анализа, ранжируйте компании по степени эффективности деятельности.

### Вариант 9

Строится модель цены автомобиля на вторичном рынке в зависимости от пробега, срока эксплуатации и объема двигателя. Имеются данные по 15 автомобилям одной и той же модели:

№ автомобиля	Цена автомобиля, долл.	Пробег, тыс. км	Срок эксплуатации, лет	Объем двигателя, л
1	12 500	130	12	2,3
2	13 700	120	10	1,9
3	9 200	300	15	1,8
4	11 400	180	13	2,1
5	15 800	150	14	2,6
6	12 300	80	8	1,7
7	16 300	170	10	2,4
8	10 200	210	11	1,9
9	11 000	250	7	1,9
10	12 700	150	9	1,7
11	15 000	90	4	2,2
12	10 500	230	13	2,4
13	17 200	120	8	2,3
14	16 000	110	9	2,5
15	17 100	120	6	2,6

- [?]** 1. Постройте матрицу парных коэффициентов линейной корреляции. Выполните тест Фаррара–Глоубера на мультиколлинеарность.

2. Постройте линейную регрессионную модель цены автомобиля, обосновав отбор факторов. Оцените параметры модели.
3. Оцените качество построенной модели.
4. Упорядочите факторы по степени их влияния на изменение цены автомобиля.
5. Спрогнозируйте цену автомобиля с пробегом 150 тыс. км, сроком эксплуатации 10 лет и объемом двигателя 2 л.

### Вариант 10

По 14 страховым компаниям исследуется зависимость месячной прибыли от численности страховых агентов, затрат на рекламу и расположения офиса компании (центральный или периферийный район города):

№ компании	Прибыль, тыс. руб.	Численность страховых агентов, чел.	Затраты на рекламу, тыс. руб.	Район расположения
1	726	14	75	периферийный
2	550	8	36	центральный
3	429	4	55	периферийный
4	439	4	45	периферийный
5	646	10	79	периферийный
6	507	10	53	периферийный
7	834	13	69	центральный
8	579	9	47	периферийный
9	701	16	45	центральный
10	532	14	49	периферийный
11	281	7	53	периферийный
12	349	5	45	периферийный
13	625	10	68	периферийный
14	533	11	38	центральный

- ?** 1. Постройте линейную регрессионную модель прибыли страховой компании методом пошагового исключения факторов.
2. Оцените качество построенной модели.
3. Существенна ли разница в прибыли компаний, офисы которых расположены в центральном и периферийных районах города?
4. Изменение какого фактора сильнее всего влияет на изменение прибыли? Оцените вклад каждого из факторов в вариацию объема прибыли с помощью  $\Delta$ -коэффициентов.



5. Используя результаты регрессионного анализа, ранжируйте компании по степени эффективности деятельности.

6. Спрогнозируйте месячную прибыль страховой компании, если прогнозные значения факторов равны своим средним значениям, а офис расположен: а) в центре города; б) на окраине.

### Вариант 11

По хлебобулочному предприятию исследуется зависимость месячного объема реализованной продукции от затрат в предыдущем месяце на теле-, радио-, газетную и наружную рекламу. Имеются данные за двенадцать месяцев:

Месяц	Объем реализованной продукции, тыс. руб.	Затраты на рекламу, тыс. руб.			
		телерекламу	радиорекламу	газетную рекламу	наружную рекламу
1	14 050	240	42	42	34
2	16 310	263	47	44	36
3	15 632	241	55	45	35
4	15 126	276	47	42	32
5	13 972	236	49	47	25
6	15 753	272	44	45	39
7	16 661	276	57	55	45
8	15 584	260	46	47	36
9	15 326	280	40	35	34
10	14 077	248	38	38	29
11	15 528	289	49	45	25
12	15 755	258	56	52	26

**?** 1. Постройте линейную регрессионную модель объема реализованной продукции, не содержащую коллинеарных факторов. Оцените параметры модели.

2. Какая доля вариации объема реализованной продукции объясняется вариацией факторов, включенных в модель регрессии?

3. Присутствует ли в остатках регрессии автокорреляция первого порядка?

4. Можно ли считать остатки случайными?

5. Приемлема ли точность регрессионной модели?

6. Изменение какого фактора сильнее всего влияет на изменение объема реализованной продукции? Оцените вклад каждого из фак-

торов в вариацию объема реализованной продукции с помощью  $\Delta$ -коэффициентов.

7. Спрогнозируйте величину объема реализованной продукции на следующие два месяца.

### Вариант 12

Исследуется взаимосвязь курса доллара США с курсами евро, японской иены и английского фунта стерлингов. Имеются данные об официальных курсах валют, установленных Центральным банком Российской Федерации, за двенадцать дней:

День	Доллар США, руб./долл.	Евро, руб./евро	Японская иена, руб./100 иен	Английский фунт стерлингов, руб./фунт
1	28,12	36,13	26,97	52,63
2	28,18	35,97	26,80	52,32
3	28,13	35,97	26,77	52,26
4	28,08	36,00	26,63	52,28
5	28,06	36,13	26,53	52,43
6	28,03	36,28	26,70	52,58
7	28,02	36,34	26,67	52,90
8	28,00	36,47	26,63	52,99
9	27,99	36,54	26,60	52,81
10	27,93	36,50	26,50	52,89
11	27,95	36,52	26,55	52,62
12	27,97	36,54	26,52	52,67

**?** 1. Постройте матрицу парных коэффициентов линейной корреляции. Выполните тест Фаррара–Глоубера на мультиколлинеарность.

2. Постройте линейную регрессионную модель курса доллара США, обосновав отбор факторов. Оцените параметры модели.

3. Оцените качество построенной модели.

4. Изменение курсов каких валют существенно влияет на изменение курса доллара США? Изменение какого фактора сильнее все-

го влияет на изменение курса доллара США? Оцените вклад каждого из факторов в вариацию курса доллара США с помощью  $\Delta$ -коэффициентов.

5. Присутствует ли в остатках регрессии автокорреляция первого порядка?

6. Можно ли считать остатки случайными?

7. Спрогнозируйте курс доллара на следующие два дня.

### Вариант 13

По 20 предприятиям региона изучается зависимость выработки продукции на одного работника  $Y$  (тыс. руб.) от ввода в действие новых основных фондов  $X_1$  (% от стоимости фондов на конец года) и удельного веса рабочих высокой квалификации в общей численности рабочих  $X_2$  (%):

№ предприятия	$Y$	$X_1$	$X_2$	№ предприятия	$Y$	$X_1$	$X_2$
1	16	3,6	9	11	19	6,3	21
2	16	3,5	12	12	21	6,4	22
3	16	3,9	14	13	21	7,4	24
4	17	4,1	17	14	22	7,5	25
5	17	3,9	18	15	22	7,9	28
6	17	4,5	19	16	23	8,2	30
7	18	5,3	19	17	23	8,0	30
8	18	5,3	19	18	23	8,6	31
9	19	5,6	20	19	24	9,5	33
10	20	6,8	21	20	24	9,0	36

**?** 1. Найдите коэффициенты парной, частной и множественной корреляции, проанализируйте их.

2. Постройте линейную модель множественной регрессии. Запишите стандартизованное уравнение множественной регрессии. На основе  $\beta$ -коэффициентов регрессии и средних коэффициентов эластичности ранжируйте факторы по степени их влияния на результат.

3. Предполагая прогнозные значения переменных равными  $x_1 = 7$  и  $x_2 = 35$ , найдите с вероятностью 0,90 доверительный интервал для прогнозного значения выработки продукции на одного работника.

4. Составьте уравнение линейной парной регрессии, оставив лишь один значимый фактор.

5. Постройте три однофакторные нелинейные регрессионные модели зависимой переменной с наиболее подходящим фактором: степенную, гиперболическую и показательную. Приведите графики всех моделей.

6. Сравните качество моделей. Выберите лучшую модель.

### Вариант 14

По 11 металлообрабатывающим цехам машиностроительного предприятия изучается зависимость фактических затрат на 1 рубль валовой продукции от среднего уровня производительности труда (отношение объема продукции в денежном выражении к затратам труда на ее изготовление) и средней энергоотдачи (отношение объема продукции в денежном выражении к затратам электроэнергии на ее изготовление). Имеются данные за последний квартал:

№ цеха	Затраты на 1 рубль валовой продукции, руб.	Уровень производительности труда, руб./чел.-ч	Энергоотдача, руб./кВт-ч
1	0,38	675	42
2	0,53	375	30
3	0,49	421	18
4	0,35	428	72
5	0,23	721	75
6	0,52	420	32
7	0,44	284	44
8	0,34	522	42
9	0,42	431	40
10	0,48	422	55
11	0,53	223	52

**?** 1. Проверьте, связаны ли между собой показатели значимыми парными линейными зависимостями.

2. Постройте все возможные линейные регрессионные модели затрат, оцените параметры моделей и выберите одну из них в качестве лучшей.

3. Можно ли использовать лучшую модель для анализа и прогнозирования затрат?

4. Приемлема ли точность лучшей модели?

5. Рассчитайте затраты на 1 рубль валовой продукции, если прогнозные значения факторов на 25% превышают свои средние значения.

6. Постройте три однофакторные нелинейные регрессионные модели зависимой переменной с наиболее подходящим фактором: степенную, гиперболическую и показательную. Приведите графики всех моделей.

7. Сравните качество моделей. Выберите лучшую модель.

*Примечание.* Там, где это необходимо, примите уровень значимости  $\alpha$  равным 0,05.

### Вариант 15

Исследуется зависимость цены системного блока компьютера от тактовой частоты процессора, размера оперативной памяти и наличия DVD-накопителя. Имеются данные по 13 компьютерам:

№ компьютера	Цена системного блока, руб.	Тактовая частота процессора, МГц	Оперативная память, Мбайт	DVD-накопитель
1	12 500	2 000	256	отсутствует
2	13 700	2 800	256	имеется
3	16 250	2 700	512	отсутствует
4	13 580	2 800	256	отсутствует
5	19 840	3 200	512	имеется
6	16 570	2 400	512	отсутствует
7	12 560	2 700	128	отсутствует
8	18 260	3 200	512	имеется
9	14 590	2 700	256	отсутствует
10	17 250	2 400	512	имеется
11	14 890	2 700	256	отсутствует
12	11 560	1 800	128	отсутствует
13	15 870	2 700	512	отсутствует

- [?]** 1. Постройте линейную регрессионную модель цены системного блока компьютера, не содержащую коллинеарных факторов. Оцените параметры модели. Если имеется возможность построить несколько моделей, то выберите одну из них в качестве лучшей.
2. Дайте экономическую интерпретацию коэффициентов уравнения регрессии.
3. Существенно ли влияет на цену системного блока:
- а) тактовая частота процессора;
  - б) размер оперативной памяти;
  - в) наличие или отсутствие DVD-накопителя?
- Приведите количественные соотношения.
4. Имеют ли остатки регрессии одинаковую дисперсию?
5. Приемлема ли точность регрессионной модели?
6. Рассчитайте стоимость системного блока, если тактовая частота процессора составляет 3000 МГц, оперативная память – 256 Мбайт, а DVD-накопитель: а) имеется; б) отсутствует.

### Вариант 16

Исследуется зависимость цены квартиры от размера ее общей площади, типа дома (кирпичный или панельный) и этажа, на котором расположена квартира (средний или крайний). Имеются данные по 16 квартирам в домах, расположенных в одном и том же районе города:

№ квартиры	Цена квартиры, долл.	Общая площадь, м <sup>2</sup>	Тип дома	Этаж
1	38 500	72	панельный	крайний
2	45 000	83	кирпичный	крайний
3	42 800	79	кирпичный	крайний
4	34 200	65	панельный	крайний
5	46 700	85	кирпичный	средний
6	48 500	70	кирпичный	крайний
7	52 300	104	кирпичный	крайний
8	44 600	72	панельный	средний
9	42 300	65	кирпичный	крайний
10	48 100	69	кирпичный	средний
11	37 400	55	кирпичный	крайний
12	35 200	54	панельный	крайний
13	49 000	72	кирпичный	средний
14	47 600	70	кирпичный	средний
15	56 000	98	кирпичный	средний
16	38 500	69	панельный	крайний

**[?]** 1. Постройте линейную регрессионную модель цены квартиры, не содержащую коллинеарных факторов на уровне значимости  $\alpha = 0,05$ . Оцените параметры модели. Если имеется возможность построить несколько моделей, то выберите одну из них в качестве лучшей.

2. Значимо ли уравнение регрессии и его коэффициенты на уровне значимости  $\alpha = 0,01$ ?

3. Какая доля вариации цены квартиры объясняется вариацией факторов, включенных в модель?

4. Приемлема ли точность модели?

5. Выполняется ли условие гомоскедастичности остатков?

6. Что в большей степени влияет на цену квартиры – тип дома или этаж, на котором она расположена? Оцените вклад каждого из факторов в вариацию цены квартиры с помощью  $\Delta$ -коэффициентов.

7. Спрогнозируйте среднюю цену квартиры общей площадью 80 м<sup>2</sup>, расположенной в панельном доме на одном из крайних этажей.

### Вариант 17

Исследуется влияние стажа работы, уровня образования и пола менеджера по продажам на размер дохода от реализации товаров, принесенного торговой фирме за последний год. Имеются сведения по 10 менеджерам:

Менеджер	Доход, млн руб.	Стаж, лет	Образова- ние	Пол
Иванова	286	7	высшее	женский
Петров	143	6	среднее	мужской
Кузнецов	187	3	высшее	мужской
Светлова	110	4	среднее	женский
Сидоренко	253	7	высшее	женский
Калинин	352	8	высшее	мужской
Крымова	154	3	высшее	женский
Жуков	308	5	высшее	мужской
Баранова	187	8	среднее	женский
Семенов	242	8	высшее	мужской

- [?]** 1. Постройте линейную регрессионную модель дохода с полным набором факторов. Оцените параметры модели.
2. Пригодно ли уравнение регрессии для анализа и прогнозирования?
3. Существенна ли разница в размере дохода, принесенного фирме менеджерами с высшим и средним образованием?
4. Существенна ли разница в размере дохода, принесенного фирме менеджерами-мужчинами и менеджерами-женщинами?
5. Постройте линейную регрессионную модель только со статистически значимыми факторами. Оцените параметры модели. Дайте экономическую интерпретацию коэффициентов уравнения регрессии.
6. Оцените точность построенной модели.
7. Спрогнозируйте средний доход менеджера с высшим образованием и стажем работы 7 лет.

*Примечание.* Там, где это необходимо, примите уровень значимости  $\alpha$  равным 0,05.

### Вариант 18

По данным, представленным в таблице, исследуется зависимость между величиной накладных расходов 40 строительных организаций  $Y$  (млн руб.) и следующими тремя основными факторами:

- $X_1$  – объем выполненных работ, млн руб.;  
 $X_2$  – численность рабочих, чел.;  
 $X_3$  – фонд заработной платы, млн. руб.



№	Y	X <sub>1</sub>	X <sub>2</sub>	X <sub>3</sub>
1	3,5	11,9	980	5,754
2	4,0	12,1	675	5,820
3	3,1	11,2	1 020	4,267
4	2,7	10,8	509	4,581
5	3,6	11,7	499	5,190
6	2,7	11,8	483	4,830
7	2,9	9,8	502	4,518
8	1,6	2,8	275	0,840
9	1,3	5,9	250	2,150
10	2,5	8,7	359	2,482
11	2,1	7,6	363	3,231
12	2,4	7,3	373	2,060
13	2,0	7,9	387	3,212
14	2,5	8,9	595	3,634
15	1,8	5,4	253	2,125
16	2,8	10,2	965	3,008
17	4,0	25,1	861	9,213
18	3,9	22,7	1 320	8,990
19	4,7	20,3	993	6,265
20	4,8	19,9	607	7,347
21	4,3	18,2	760	7,524
22	3,5	17,3	738	6,642
23	3,0	16,5	634	5,833
24	3,6	17,0	683	12,059
25	3,3	17,1	424	7,051
26	2,9	16,2	593	6,404
27	3,1	17,3	406	5,575
28	2,8	16,3	807	5,019
29	3,5	12,9	629	10,485
30	4,6	13,8	1 060	5,820
31	3,5	10,1	588	5,116
32	2,9	10,9	625	5,510
33	2,7	11,4	500	5,200
34	2,8	11,3	450	4,455
35	3,0	8,7	510	4,488
36	2,9	10,0	232	4,968
37	2,4	5,2	419	4,022
38	1,6	7,4	159	1,570
39	1,2	2,2	162	1,142
40	1,5	2,6	101	0,429

**?** 1. Осуществите двумя способами выбор факторных признаков для построения регрессионной модели:

а) на основе анализа матрицы коэффициентов парной корреляции, включая проверку гипотезы о независимости объясняющих переменных (тест Фаррара–Глоубера на выявление мультиколлинеарности);

- б) с помощью пошагового отбора методом исключения.
2. Дайте экономическую интерпретацию коэффициентов уравнения регрессии.
3. Дайте сравнительную оценку силы связи факторов с результатом с помощью коэффициентов эластичности,  $\beta$ - и  $\Delta$ -коэффициентов.
4. Проведите тестирование ошибок уравнения регрессии на гетероскедастичность, используя тест Гольдфельда–Квандта.
5. Рассчитайте прогнозное значение результата, если прогнозные значения факторов составляют 75% от их максимальных значений.

### Вариант 19

Исследуется влияние некоторых показателей социально-экономического развития субъектов Центрального федерального округа России на региональный индекс потребительских цен. В таблице приведены официальные статистические данные по субъектам Центрального федерального округа Российской Федерации за 2005 г.<sup>1</sup>

Область	Y	X <sub>1</sub>	X <sub>2</sub>	X <sub>3</sub>	X <sub>4</sub>	X <sub>5</sub>
Белгородская	112,6	110,5	106,1	126,2	6,86	1 512
Брянская	111,9	109,2	108,3	126,8	5,24	1 346
Владимирская	110,9	107,7	104,7	126,9	6,07	1 487
Воронежская	109,7	108,8	105,9	120,9	5,60	2 334
Ивановская	108,7	106,0	107,2	116,6	5,37	1 115
Калужская	111,2	109,1	108,7	120,9	6,98	1 022
Костромская	109,2	106,6	105,1	123,4	5,84	717
Курская	109,4	107,7	106,1	116,9	5,65	1 199
Липецкая	110,9	109,0	106,3	125,4	7,19	1 190
Московская	111,3	107,5	105,6	126,8	9,51	6 630
Орловская	109,5	107,5	106,1	122,4	5,46	842
Рязанская	110,1	109,7	106,1	118,4	6,22	1 195
Смоленская	111,9	110,3	107,5	122,4	6,30	1 019
Тамбовская	109,8	109,8	106,7	115,0	5,08	1 145
Тверская	110,9	108,7	105,2	125,6	6,64	1 425
Тульская	110,8	108,0	106,7	126,1	6,34	1 622
Ярославская	112,3	109,2	106,6	130,6	7,39	1 339
г. Москва	110,4	109,5	105,8	118,1	13,74	10 407

<sup>1</sup> Российская газета. – 2006. – 24 марта. – № 60.

В таблице приняты следующие обозначения:

1) переменные  $Y, X_1, X_2, X_3$  представляют собой индексы потребительских цен в декабре 2005 г., выраженные в процентах к декабрю 2004 г.:

$Y$  – общий индекс потребительских цен;

$X_1$  – индекс потребительских цен на продовольственные товары;

$X_2$  – индекс потребительских цен на непродовольственные товары;

$X_3$  – индекс потребительских цен на платные услуги населению;

2) переменные  $X_4$  и  $X_5$ :

$X_4$  – среднемесячная номинальная начисленная заработная плата в 2005 г., тыс. руб.;

$X_5$  – численность населения региона на 1 января 2005 г., тыс. чел.

Переменная  $Y$  рассматривается как результирующая, переменные  $X_1, X_2, X_3, X_4, X_5$  – как факторные.

**[?] 1.** Постройте матрицу парных коэффициентов линейной корреляции и выявите наличие либо отсутствие коллинеарности факторов.

2. Постройте линейную регрессионную модель общего индекса потребительских цен, не содержащую коллинеарных факторов. Проверьте статистическую значимость уравнения регрессии и его параметров. Сделайте выводы о существенности либо несущественности влияния каждого из факторов на общий индекс потребительских цен.

3. Постройте линейную регрессионную модель общего индекса потребительских цен, включающую в себя только те факторы, которые оказывают существенное влияние на результирующую переменную. Оцените качество модели.

4. Дайте экономическую интерпретацию параметров уравнения регрессии и оцените вклад каждого из факторов в общий индекс потребительских цен с помощью  $\Delta$ -коэффициентов.

5. Для анализа влияния численности населения региона на средний уровень доходов его жителей постройте линейную модель регрессионной зависимости среднемесячной заработной платы  $X_4$  от численности населения  $X_5$ . Проверьте статистическую значимость уравнения регрессии и оцените среднюю эластичность изменения заработной платы в зависимости от числа жителей.

6. Используя результаты регрессионного анализа, выявите наиболее привлекательные регионы.

### Вариант 20

Исследуется влияние объема промышленного производства и размера инвестиций в основной капитал на региональный коэффициент смертности. В таблице приведены официальные статистические данные по субъектам Центрального федерального округа за 2005 и 2006 гг.<sup>1</sup>:

Область	Y	X <sub>1</sub>	X <sub>2</sub>	X <sub>3</sub>
Белгородская	15,3	135,2	109,5	125,2
Брянская	18,6	86,5	111,4	112,2
Владимирская	19,4	107,4	105,3	105,1
Воронежская	18,1	108,1	105,1	112,4
Ивановская	19,9	104,6	112,0	106,2
Калужская	17,7	103,6	106,4	107,8
Костромская	18,9	114,8	110,8	73,5
Курская	19,0	92,3	107,4	100,8
Липецкая	17,4	100,2	110,4	130,5
Московская	17,1	90,5	118,0	106,6
Орловская	17,9	100,2	108,9	112,1
Рязанская	19,2	89,7	110,2	99,8
Смоленская	20,8	114,7	106,3	88,6
Тамбовская	18,3	116,3	108,1	118,9
Тверская	21,8	66,3	111,3	84,5
Тульская	20,9	101,9	107,9	100,0
Ярославская	18,3	125,7	105,6	76,9
г. Москва	12,2	106,3	118,5	109,4

В таблице приняты следующие обозначения:

Y – коэффициент смертности, то есть число умерших за год на 1000 человек населения в 2006 г., ‰;

X<sub>1</sub> – индекс (темп роста) инвестиций в основной капитал в 2005 г., % к 2004 г.;

X<sub>2</sub> – индекс промышленного производства в 2006 г., % к 2005 г.;

X<sub>3</sub> – индекс инвестиций в основной капитал в 2006 г., % к 2005 г.

<sup>1</sup> Российская газета. – 2006. – 24 марта. – № 60; Российская газета. – 2007. – 14 марта. – № 51.

**[?]** 1. Для выявления линейных связей в исходных данных постройте матрицу парных коэффициентов корреляции. Проверьте статистическую значимость коэффициентов корреляции и сделайте выводы о наличии либо отсутствии устойчивых зависимостей между исследуемыми показателями.

2. Постройте линейную модель регрессионной зависимости коэффициента смертности в 2006 г. от индексов инвестиций в основной капитал в 2005 г., промышленного производства в 2006 г. и инвестиций в основной капитал в 2006 г. Проверьте статистическую значимость полученного уравнения регрессии и его параметров. Сделайте выводы о существенности либо несущественности влияния индекса инвестиций в основной капитал и индекса промышленного производства на коэффициент смертности.

3. Дайте экономическую интерпретацию параметров уравнения регрессии и оцените вклад каждого из факторов в вариацию коэффициента смертности с помощью  $\Delta$ -коэффициентов.

4. На основе анализа остатков регрессии ранжируйте регионы по эффективности снижения коэффициента смертности под влиянием роста инвестиций в основной капитал и промышленного производства. Выявите наиболее «передовые» и наиболее «отсталые» субъекты.

*Примечание.* При проверке статистических гипотез уровень значимости  $\alpha$  примите равным 0,05.

### Вариант 21

По данным, представленным в таблице ( $n = 25$ ), изучается зависимость объема выпуска продукции  $Y$  (млн руб.) от следующих факторов (переменных):

$X_1$  – численность промышленно-производственного персонала, чел.;

$X_2$  – среднегодовая стоимость основных фондов, млн руб.;

$X_3$  – износ основных фондов, %;

$X_4$  – электровооруженность, кВт·ч;

$X_5$  – техническая вооруженность одного рабочего, млн руб.;

$X_6$  – выработка товарной продукции на одного работающего, руб.

№ наблюде- ния	Y	X <sub>1</sub>	X <sub>2</sub>	X <sub>3</sub>	X <sub>4</sub>	X <sub>5</sub>	X <sub>6</sub>
1	32 900	864	16 144	39,5	4,9	3,2	36 354
2	203 456	8 212	336 472	46,4	60,5	20,4	23 486
3	41 138	1 866	39 208	43,7	24,9	9,5	20 866
4	57 342	1 147	63 273	35,7	50,4	34,7	47 318
5	27 294	1 514	31 271	41,8	5,1	17,9	17 230
6	94 552	4 970	86 129	49,8	35,9	12,1	19 025
7	28 507	1 561	48 461	44,1	48,1	18,9	18 262
8	97 788	4 197	138 657	48,1	69,5	12,2	23 360
9	101 734	6 696	127 570	47,6	31,9	8,1	15 223
10	175 322	5 237	208 900	58,6	139,4	29,7	32 920
11	2 894	547	6 922	70,4	16,9	5,3	5 291
12	16 649	710	8 228	37,5	17,8	5,6	23 125
13	19 216	940	18 894	62,0	27,6	12,3	20 848
14	23 684	3 528	27 486	34,4	13,9	3,2	6 713
15	1 237 132	52 412	1 974 472	35,4	37,3	19,0	22 581
16	88 569	4 409	162 229	40,8	55,3	19,3	20 522
17	162 216	6 139	128 731	48,1	35,1	12,4	26 396
18	10 201	802	6 714	43,4	14,9	3,1	13 064
19	3 190	442	478	43,2	0,2	0,6	6 847
20	55 410	2 797	60 209	57,1	37,2	13,1	20 335
21	332 448	10 280	540 780	51,5	74,45	21,5	32 339
22	97 070	4 560	108 549	53,6	32,5	13,2	20 675
23	98 010	3 801	169 995	60,4	75,9	27,2	26 756
24	1 087 322	46 142	972 349	50,0	27,5	10,8	23 176
25	55 004	2 535	163 695	25,5	65,5	19,9	21 698

- [?]** 1. Постройте матрицу парных коэффициентов корреляции. Проверьте наличие мультиколлинеарности. Обоснуйте отбор факторов в модель.
2. Постройте уравнение множественной регрессии в линейной форме, обоснуйте выбор факторов.
3. Оцените статистическую значимость уравнения регрессии и его параметров с помощью критерия Фишера и критерия Стьюдента.

4. Постройте уравнение регрессии со статистически значимыми факторами. Оцените качество уравнения регрессии с помощью коэффициента детерминации  $R^2$ . Оцените точность построенной модели.

5. Проверьте выполнение предпосылок МНК, в том числе проведите тестирование ошибок уравнения регрессии со значимыми факторами на гетероскедастичность, используя тест Гольдфелда–Квандта.

6. Оцените прогноз объема выпуска продукции, если прогнозные значения факторов составляют 75% от их максимальных значений.

### Вариант 22

В таблице представлены данные об источниках финансирования аграрного сектора и результатах деятельности 24 районов Орловской области в 2007 г. (по данным Орловского областного Департамента аграрной политики), тыс. руб. на 100 га сельскохозяйственных угодий.

Район	Средства федерального бюджета	Средства областного бюджета	Кредиты
Болховский	76 803,35372	33 267,280	144,431600
Знаменский	11 192,96598	4 548,019	0,779589
Хотынецкий	15 367,32481	9 507,580	192,901700
Шаблыкинский	11 751,83347	5 033,002	10,870350
Сосковский	23 029,84660	9 304,449	46,749450
Урицкий	16 200,32172	4 781,344	68,130280
Дмитровский	8 809,14503	4 794,554	41,022560
Троснянский	340 383,88920	14 925,670	140,650100
Кромской	16 273,40470	8 604,795	361,595200
Орловский	52 128,02153	24 846,490	1 343,826000
Мценский	138 044,11110	18 055,920	1 141,875000
Корсаковский	34 224,60674	14 835,770	316,5608000
Новосильский	51 778,32526	15 801,740	0
Залегощенский	26 295,58714	6 113,528	78,251280
Свердловский	103 652,14960	19 527,240	368,824500
Глазуновский	11 516,87061	4 989,014	1 068,922000
Малоархангельский	31 647,06381	14 460,130	105,448200
Покровский	32 090,93950	9 732,127	196,579800
Верховский	26 087,04346	30 753,110	77,558730
Новодеревеньковский	25 759,89303	21 186,150	66,292520
Краснозоренский	33 370,16999	17 810,100	35,603340
Ливенский	73 574,94683	26 367,330	564,035300
Колпнянский	82 399,20846	37 606,950	209,350500
Должанский	84 133,46450	30 418,570	269,796200

Район	Собственный капитал	Привлеченные средства	Выручка от продаж
Болховский	851,85090	34,54439992	384,98207130
Знаменский	121,21900	22,82423813	129,00070870
Хотынецкий	188,29870	133,88347160	411,48022530
Шаблыкинский	205,05110	52,86614972	229,46952680
Сосковский	231,95030	107,43608470	291,46822500
Урицкий	272,71570	60,10326681	204,41068410
Дмитровский	221,99540	61,18031584	130,32501210
Троснянский	114,39140	209,81378150	499,31871270
Кромской	196,89750	400,53086180	535,00641820
Орловский	984,75480	282,18499830	1165,04944500
Мценский	761,39020	291,81858010	1562,75787200
Корсаковский	2152,80900	2168,13874000	524,52369320
Новосильский	753,10990	115,71082390	476,84437260
Залогощенский	585,44530	470,15932470	801,49831950
Свердловский	1033,33300	377,86411100	937,10230890
Глазуновский	-7,09466	395,04228900	341,71765340
Малоархангельский	478,21950	213,13253740	65,45663899
Покровский	807,20930	377,89721430	647,17160320
Верховский	503,06000	153,58321350	778,85889950
Новодеревеньковский	339,99770	477,03868320	665,18136280
Краснозоренский	220,33930	350,42409860	375,82463610
Ливенский	1223,96800	219,16902000	1329,72492400
Колпнянский	1123,11700	208,58115110	747,28510130
Должанский	1672,28900	49,79274512	700,69912700

**[?] 1.** Методом исключения отберите факторы, наиболее подходящие для исследования результативного признака «Выручка от продаж». Постройте соответствующее регрессионное уравнение, выявите его недостатки. Сделайте выводы экономического характера.

**2.** Оцените степень влияния факторов на результат (коэффициенты эластичности,  $\beta$ - и  $\Delta$ -коэффициенты).



3. Приемлема ли точность модели?
4. Выполняется ли условие гомоскедастичности остатков?
5. Постройте 90%-ные доверительные интервалы для результативного признака. Определите, в каких районах выручка занижена (завышена) по сравнению с полученными интервалами.

### Вариант 23

В таблице представлены данные по трем источникам финансирования аграрного сектора 24 районов Орловской области в 2008 г. (по данным Орловского областного Департамента аграрной политики), руб.

Район	Средства федерального бюджета	Средства областного бюджета	Кредиты
	$X_1$	$X_2$	$Y$
Болховский	3 654 893	4 614 092	7 890
Верховский	22 918 887	8 837 266	27 319
Глазуновский	9 798 328	6 661 913	151 300
Дмитровский	12 056 843	6 250 182	12 045
Должанский	704 478 409	25 547 013	48 097
Залегощенский	32 470 495	5 286 063	3 400
Знаменский	7 999 957	5 229 540	0
Колпнянский	14 559 803	5 583 613	8 696
Корсаковский	12 036 677	2 351 210	13 000
Краснозороенский	13 107 929	4 148 032	42 000
Кромской	15 413 019	3 905 414	205 644
Ливенский	112 567 704	28 015 838	317 378
Малоархангельский	16 916 111	3 360 425	16 000
Мценский	199 847 559	27 759 036	581 324
Новодеревеньковский	16 125 153	2 661 459	3 000
Новосильский	3 575 760	985 800	0
Орловский	71 883 509	32 223 701	1 454 130
Покровский	22 469 778	7 849 744	500
Свердловский	69 730 395	18 562 686	116 247
Сосковский	11 153 935	6 967 994	0
Троснянский	15 712 151	5 036 214	9 675
Урицкий	14 698 546	2 917 100	16 800
Хотынецкий	12 207 639	10 007 704	2 000
Шаблыкинский	9 318 512	5 328 866	1 199

- [?] 1.** Проведите корреляционный анализ данных для определения переменных, наиболее подходящих для дальнейшего регрессионного анализа.
2. Осуществите регрессионный анализ выбранных переменных.
3. Сравните качественные характеристики построенных моделей. Определите вариант наиболее целесообразной регрессии. Сделайте выводы экономического характера.
4. Оцените точность лучшей модели.
5. По лучшей модели оцените прогноз результативного признака с вероятностью 90%, если прогнозное значение фактора (факторов) увеличится на 30% от своего среднего.
6. Результаты аппроксимации и прогнозирования представьте на графике.

### Вариант 24

Обозначение	Наименование показателя	Единица измерения (возможные значения)
Y	цена квартиры	тыс. долл.
X <sub>1</sub>	город области	1 – Подольск
		0 – Люберцы
X <sub>2</sub>	число комнат в квартире	
X <sub>3</sub>	общая площадь квартиры	м <sup>2</sup>
X <sub>4</sub>	площадь кухни	м <sup>2</sup>

№	Y	X <sub>1</sub>	X <sub>2</sub>	X <sub>3</sub>	X <sub>4</sub>
1	115,00	0	4	70,4	7,0
2	85,00	1	3	82,8	10,0
3	69,00	1	2	64,5	10,0
4	57,00	1	2	55,1	9,0
5	184,60	0	3	83,9	9,0
6	56,00	1	1	32,2	7,0
7	85,00	0	3	65,0	8,3
8	265,00	0	4	169,5	16,5
9	60,65	1	2	74,0	12,1
10	130,00	0	4	87,0	6,0
11	46,00	1	1	44,0	10,0
12	115,00	0	3	60,0	7,0
13	70,96	0	2	65,7	12,5

№	Y	X <sub>1</sub>	X <sub>2</sub>	X <sub>3</sub>	X <sub>4</sub>
14	39,5)	1	1	42,0	11,0
15	78,9)	0	1	49,3	13,6
16	60,00	1	2	64,5	12,0
17	100,00	1	4	93,8	9,0
18	51,00	1	2	64,0	12,0
19	157,00	0	4	98,0	11,0
20	123,50	1	4	107,5	12,3
21	55,20	0	1	48,0	12,0
22	95,50	1	3	80,)	12,5
23	57,60	0	2	63,9	11,4
24	64,50	1	2	58,1	10,6
25	92,00	1	4	83,0	6,5
26	100,00	1	3	73,4	7,0
27	81,00	0	2	45,5	6,3
28	65,00	1	1	32,0	6,6
29	110,00	0	3	65,2	9,6
30	42,10	1	1	40,3	10,8
31	135,00	0	2	72,0	10,0
32	39,60	1	1	36,0	8,6
33	57,00	1	2	61,6	10,0
34	80,00	0	1	35,5	8,5
35	61,00	1	2	58,1	10,6
36	69,60	1	3	83,0	12,0
37	250,00	1	4	152,0	13,3
38	64,50	1	2	64,5	8,6
39	125,00	0	2	54,0	9,0
40	152,30	0	3	89,0	13,0

**?** 1. Рассчитайте матрицу парных коэффициентов корреляции. Оцените статистическую значимость коэффициентов корреляции.

2. Выполните тест Фаррара–Глоубера на мультиколлинеарность. Обоснуйте отбор факторов в модель регрессии.

3. Постройте модель регрессии с выбранными факторами и дайте экономическую интерпретацию ее коэффициентов.

4. Оцените качество модели, используя коэффициент детерминации, среднюю ошибку аппроксимации и  $F$ -критерий Фишера. Проверьте значимость коэффициентов регрессии.

5. Дайте для выбранной модели оценку влияния значимых факторов на результат с помощью коэффициентов эластичности,  $\beta$ - и  $\Delta$ -коэффициентов.

6. Постройте модель регрессии с наиболее влиятельным фактором. Сравните ее качественные характеристики с характеристиками модели из п. 3.

### Вариант 25

Обозначение	Наименование показателя	Единица измерения (возможные значения)
$Y$	цена квартиры	тыс. долл.
$X_1$	город области	1 – Подольск
		0 – Люберцы
$X_2$	число комнат в квартире	
$X_3$	жилая площадь квартиры	$m^2$
$X_4$	площадь кухни	$m^2$

№	$Y$	$X_1$	$X_2$	$X_3$	$X_4$
1	38,0	1	1	19,0	9,5
2	62,2	1	2	36,0	10,0
3	125,0	0	3	41,0	8,0
4	61,1	1	2	34,8	10,6
5	67,0	0	1	18,7	6,0
6	93,0	0	2	27,7	11,3
7	118,0	1	3	59,0	13,0
8	132,0	0	3	44,0	11,0
9	92,5	0	3	56,0	12,0
10	105,0	1	4	47,0	12,0
11	42,0	1	1	18,0	8,0
12	125,0	1	3	44,0	9,0
13	170,0	0	4	56,0	8,5
14	38,0	0	1	16,0	7,0
15	130,5	0	4	66,0	9,8
16	85,0	0	2	34,0	12,0
17	98,0	0	4	43,0	7,0
18	128,0	0	4	59,2	13,0
19	85,0	0	3	50,0	13,0
20	160,0	1	3	42,0	10,0
21	60,0	0	1	20,0	13,0
22	41,0	1	1	14,0	10,0
23	90,0	1	4	47,0	12,0
24	83,0	0	4	49,5	7,0
25	45,0	0	1	18,9	5,8
26	39,0	0	1	18,0	6,5
27	86,9	0	3	58,7	14,0
28	40,0	0	1	22,0	12,0

№	Y	X <sub>1</sub>	X <sub>2</sub>	X <sub>3</sub>	X <sub>4</sub>
29	80,0	0	2	40,0	10,0
30	227,0	0	4	91,0	20,5
31	235,0	0	4	90,0	18,0
32	40,0	1	1	15,0	11,0
33	67,0	1	1	18,5	12,0
34	123,0	1	4	55,0	7,5
35	100,0	0	3	37,0	7,5
36	105,0	1	3	48,0	12,0
37	70,3	1	2	34,8	10,6
38	82,0	1	3	48,0	10,0
39	280,0	1	4	85,0	21,0
40	200,0	1	4	60,0	10,0

**?** 1. Используя пошаговую множественную регрессию (метод исключения или метод включения), постройте модель формирования цены квартиры за счет значимых факторов. Дайте экономическую интерпретацию коэффициентов модели регрессии.

2. Оцените качество модели, используя коэффициент детерминации, среднюю ошибку аппроксимации и *F*-критерий Фишера.

3. По диаграммам остатков определите ту объясняющую переменную, от которой может зависеть дисперсия случайных возмущений. Проверьте выполнение условия гомоскедастичности модели по тесту Гольдфельда–Квандта.

4. Осуществите прогнозирование среднего значения показателя *Y* при уровне значимости  $\alpha = 0,1$ , если прогнозное значение фактора *X* составляет 80% от его максимального значения. Представьте графически фактические и модельные значения, точки прогноза.

### Вариант 26

В таблице представлены финансовые показатели деятельности предприятий отрасли «Производство автомобилей, прицепов и полуприцепов» за 2009 г.<sup>1</sup>

Для вашего варианта используйте данные строк 1–40.

<sup>1</sup> <http://www.fira.ru/>

№ п/п	Предприятие отрасли	Прибыль (убыток) от про- даж	Долго- срочные обяза- тельства	Запасы готовой продук- ции и товаров для пе- репрода- жи	Кратко- срочные обяза- тельства	Оборот- ные ак- тивы	Основ- ные средства	Дебитор- ская за- должен- ность
		Y	X <sub>1</sub>	X <sub>2</sub>	X <sub>3</sub>	X <sub>4</sub>	X <sub>5</sub>	X <sub>6</sub>
1	136 центральная база производ- ства и ремонта вооружения и средств радиационной, химиче- ской и биологической защиты, открытое акционерное общество	-9,48	0,89	1,44	60,82	73,17	61,56	17,42
2	Абаканвагонмаш, открытое ак- ционерное общество	-29,11	1,75	3,97	104,41	160,94	413,73	69,91
3	Автоагрегат, открытое акционер- ное общество	6,15	1,06	21,78	62,55	213,70	107,58	70,74
4	АВТОВАЗ, открытое акционерное общество	-9 482,00	25 867,00	8 961,00	83 953,00	38 997,00	51 899,00	5 857,00
5	АВТОВАЗАГРЕГАТ, открытое акционерное общество	19,16	271,93	60,65	2 290,04	1 731,87	807,30	1 147,49
6	АВТОГИДРОПОДЪЕМНИК, от- крытое акционерное общество	14,30	5,54	2,92	57,43	72,43	38,04	37,37
7	АВТОДЕТАЛЬ-СЕРВИС, открытое акционерное общество	-12,27	573,91	135,32	1 117,02	1 152,17	1 272,18	571,19
8	Автодизель (Ярославский Мотор- ный Завод), открытое акционер- ное общество	-220,51	745,71	1,68	6 945,37	3 857,65	2 756,8	2 586,83
9	Автомобильный завод «УРАЛ», открытое акционерное общество	116,49	95,00	21,65	2 698,84	1 331,41	1 052,31	497,91

№ п/п	Предприятие отрасли	Прибыль (убыток) от про- даж	Долго- срочные обяза- тельства	Запасы готовой продук- ции и товаров для пе- репрода- жи	Кратко- срочные обяза- тельства	Оборот- ные ак- тивы	Основ- ные средства	Дебитор- ская за- должен- ность
		Y	X <sub>1</sub>	X <sub>2</sub>	X <sub>3</sub>	X <sub>4</sub>	X <sub>5</sub>	X <sub>6</sub>
10	Автополимер, закрытое акцио- нерное общество	-6,24	9,00	1,11	114,68	101,41	6,29	73,87
11	АВТОПРИЦЕП-КАМАЗ, открытое акционерное общество	27,10	0	14,09	37,31	291,14	130,84	50,03
12	Арзамасский завод коммунально- го машиностроения, открытое акционерное общество	162,78	12,56	121,09	1 133,60	1 017,28	179,69	537,09
13	Баксанский завод «Автозап- часть», общество с ограниченной ответственностью	-11,27	0	22,75	152,21	0	36,13	89,51
14	Башкирский троллейбусный за- вод, открытое акционерное обще- ство	37,78	1,13	0,93	66,71	103,30	24,96	35,59
15	Бесланский завод «Автоспецебо- рудование», открытое акционер- ное общество	-0,30	0	1,67	34,77	22,31	6,40	2,31
16	Бугурусланский завод «Радиа- тор», открытое акционерное об- щество	-29,01	95,39	8,57	376,75	280,78	156,95	193,40
17	Бузулукский механический завод, открытое акционерное общество	7,72	82,90	41,28	215,88	261,54	22,20	149,07
18	Варгашинский завод противопопо- жарного и специального оборудо- вания, открытое акционерное общество	51,90	31,00	10,53	54,98	132,97	143,00	30,12

№ п/п	Предприятие отрасли	Прибыль (убыток) от про- даж	Долго- срочные обяза- тельства	Запасы готовой продук- ции и товаров для пе- репрода- жи	Кратко- срочные обяза- тельства	Оборот- ные ак- тивы	Основ- ные средства	Дебитор- ская за- должен- ность
		Y	X <sub>1</sub>	X <sub>2</sub>	X <sub>3</sub>	X <sub>4</sub>	X <sub>5</sub>	X <sub>6</sub>
19	ГАЗ, открытое акционерное об- щество	-1 263,74	2 025,04	4,11	16 260,50	5 881,73	5 403,58	2 783,68
20	ГЕОМАШ, Щигровское открытое акционерное общество	69,76	26,22	83,69	146,08	294,60	226,52	14,65
21	Грабовский автомобильный за- вод, открытое акционерное обще- ство	39,97	42,00	44,12	336,72	409,01	104,56	199,69
22	Дмитровградский автоагрегат- ный завод, открытое акционерное общество	-48,54	16,15	39,86	1 943,93	1 264,44	907,88	523,31
23	Завод Автотехнологий, открытое акционерное общество	2,80	2,34	8,00	19,96	17,89	13,95	6,13
24	Завод «ВОЛНА», открытое акцио- нерное общество	-0,45	0	5,00	40,76	52,33	11,76	31,59
25	Завод имени И.А. Пихачева, от- крытое акционерное московское общество	-1 323,82	819,77	65,46	14 808,23	3 206,03	1 796,14	1 813,48
26	Завод коммунальной энергетики, открытое акционерное общество	1,85	14,16	0,59	34,38	42,92	6,26	2,90
27	Завод «Старт», открытое акцио- нерное общество	9,54	3,61	24,30	54,44	164,73	120,68	43,66
28	Завод технологического оборудо- вания «Каменя», открытое акцио- нерное общество	6,48	0,01	0,97	52,30	62,29	21,17	41,37



№ п/п	Предприятие отрасли	Прибыль (убыток) от продаж	Долгосрочные обязательства	Запасы готовой продукции и товаров для перепродажи	Краткосрочные обязательства	Оборотные активы	Основные средства	Дебиторская задолженность
		У	X <sub>1</sub>	X <sub>2</sub>	X <sub>3</sub>	X <sub>4</sub>	X <sub>5</sub>	X <sub>6</sub>
29	Заволжский завод гусеничных тягачей, открытое акционерное общество	-5,17	5,32	23,96	60,84	333,97	98,58	270,21
30	Заволжский моторный завод, открытое акционерное общество	-86,31	2 511,64	214,73	1 941,50	5 367,62	1 739,15	4 090,04
31	Ижевский автомобильный завод, открытое акционерное общество	-892,34	733,65	260,43	14 117,10	9 016,58	3 254,13	7 385,73
32	ИНПРОКОМ-КИРИШИ, открытое акционерное общество	0,01	0	6,31	7,90	13,46	0	6,41
33	Ишимский машиностроительный завод, открытое акционерное общество	-37,80	19,34	6,88	74,88	57,34	34,73	29,24
34	КАМАЗ, открытое акционерное общество	-671,99	19 107,65	520,41	19 422,45	33 861,44	24 745,77	15 519,55
35	КАМАЗ-Дизель, открытое акционерное общество	-616,10	1 320,71	102,23	2 318,64	2 817,81	1 168,15	1 905,08
36	КАМАЗинструментспецмаш, открытое акционерное общество	-113,66	331,79	93,20	473,51	687,73	162,68	297,71
37	Камский прессово-рамный завод, открытое акционерное общество	-336,34	457,52	65,29	1 497,33	1 348,34	333,49	545,62

№ п/п	Предприятие отрасли	Прибыль (убыток) от про- даж	Долго- срочные обяза- тельства	Запасы готовой продук- ции и товаров для пе- репрода- жи	Кратко- срочные обяза- тельства	Оборот- ные ак- тивы	Основ- ные средства	Дебитор- ская за- должен- ность
		Y	X <sub>1</sub>	X <sub>2</sub>	X <sub>3</sub>	X <sub>4</sub>	X <sub>5</sub>	X <sub>6</sub>
38	Канашский автоагрегатный завод, открытое акционерное общество	10,00	14,11	2,02	315,73	247,36	116,79	164,64
39	Клинцовский автокрановый завод, открытое акционерное общество	5,19	2,34	78,43	166,06	696,85	241,90	93,02
40	Козельский механический завод, открытое акционерное общество	20,97	1,54	14,29	105,37	113,39	21,75	65,96
41	Комбинат автомобильных фургон- ов, открытое акционерное обще- ство	308,00	32,00	52,25	192,10	328,53	217,41	131,41
42	Костромской завод «Моторде- таль», открытое акционерное общество	-328,76	18,04	4,61	1 530,36	512,65	764,89	209,16
43	Красногорский комбинат авто- фургонов, открытое акционерное общество	239,36	0,32	11,07	329,86	722,68	205,82	162,10
44	Кубжик, закрытое акционерное общество	-0,50	0	0,14	0,69	5,41	1,74	1,70

- [?]** 1. Проведите корреляционный анализ данных с целью определения переменных, наиболее подходящих для дальнейшего регрессионного анализа.
2. Осуществите парный регрессионный анализ с каждой из выбранных переменных.
3. Сравните качественные характеристики построенных моделей. Определите вариант наиболее целесообразной регрессии. Сделайте выводы экономического характера.
4. По лучшей модели оцените прогноз результативного признака с вероятностью 90%, если прогнозное значение фактора увеличится на 30% от своего среднего значения.
5. Обоснуйте целесообразность или нецелесообразность использования множественной регрессии для анализа переменной  $Y$ .

### Вариант 27

В таблице (см. вариант 26) представлены финансовые показатели деятельности предприятий отрасли «Производство автомобилей, прицепов и полуприцепов» за 2009 г.

Для вашего варианта используйте числовые данные строк 2–41.

- [?]** 1. Методом исключения отберите факторы, наиболее подходящие для исследования результативного признака «Прибыль (убыток) от продаж». Постройте соответствующее регрессионное уравнение, выявите его недостатки. Сделайте выводы экономического характера.
2. Постройте 90%-ные доверительные интервалы для результативного признака. Определите, на каких предприятиях прибыль (убыток) занижена (завышена) по сравнению с полученными интервалами? Продемонстрируйте результаты выполнения этого пункта задания на графике.
3. Оцените степень влияния факторов на результат.
4. Постройте уравнение парной регрессии с наиболее влиятельным фактором и сравните его качественные характеристики с уравнением п. 1.

### Вариант 28

В таблице (см. вариант 26) представлены финансовые показатели деятельности предприятий отрасли «Производство автомобилей, прицепов и полуприцепов» за 2009 г.

Для вашего варианта используйте числовые данные строк 3–42.

**[?]** 1. Методом пошагового включения постройте уравнение регрессии с наиболее подходящим набором факторов. Дайте экономическую интерпретацию коэффициентов регрессии.

2. Какова роль факторов, не учтенных в модели, в вариации прибыли (убытка) от продаж?

3. Постройте матрицу коэффициентов парной корреляции по всем имеющимся переменным. Отберите факторы для регрессионного анализа. При наличии коллинеарности факторов воспользуйтесь коэффициентами частной корреляции.

4. Постройте уравнение регрессии с факторами, отобранными в п. 3.

5. Сравните модели п. 1 и 4. Какая из моделей является более предпочтительной и почему?

### Вариант 29

В таблице (см. вариант 26) представлены финансовые показатели деятельности предприятий отрасли «Производство автомобилей, прицепов и полуприцепов» за 2009 г.

Для вашего варианта используйте числовые данные строк 4–43.

**[?]** 1. Определите парные и частные коэффициенты корреляции прибыли (убытка) от продаж с каждым из факторов. Прокомментируйте различие парных и частных коэффициентов. Выберите факторы (фактор), наиболее подходящие для построения регрессионного уравнения.

2. Постройте регрессионное уравнение с выбранными факторами (фактором), определите недостатки его качества. Сделайте выводы экономического характера.

3. Постройте регрессионное уравнение методом пошагового отбора, сравните его с уравнением п. 2.

4. На основе уравнения п. 3 определите с вероятностью 85% прогнозные оценки прибыли (убытка) от продаж при увеличении значений факторов (фактора) на 20% от их (его) среднего значения.

5. Результаты моделирования и прогнозирования представьте графически.

### Вариант 30

В таблице (см. вариант 26) представлены финансовые показатели деятельности предприятий отрасли «Производство автомобилей, прицепов и полуприцепов» за 2009 г.

Для вашего варианта используйте данные строк 5–44.

**[?]** 1. По результатам корреляционного анализа имеющейся совокупности выберите ведущий фактор, наиболее подходящий для анализа исследуемой переменной  $Y$ .

2. Постройте парную линейную регрессию с выбранным фактором. Поясните экономический смысл ее параметров.

3. Рассчитайте эластичность переменной  $Y$  по фактору и сделайте вывод об эластичной или неэластичной зависимости.

4. Постройте степенную модель парной регрессии.

5. Оцените качество построенных линейной и степенной моделей:

а) проверьте ряды остатков на наличие гетероскедастичности;

б) оцените долю случайной вариации переменной  $Y$ , учтенную в модели и обусловленную случайными колебаниями включенного в нее фактора;

в) на основе средней относительной ошибки аппроксимации определите средний процент отклонений фактических значений  $Y$  от рассчитанных по модели;

г) проверьте статистическую значимость уравнений регрессии ( $P = 95\%$ ).

6. Сравните качественные характеристики моделей, сделайте вывод о лучшей регрессии.

7. Методом пошагового отбора (последовательным исключением или включением факторов) обоснуйте целесообразность или нецелесообразность использования модели множественной регрессии для анализа поведения переменной  $Y$  в зависимости от значений факторов, представленных в таблице варианта.

### Литература

1. **Орлова И.В., Половников В.А.** Экономико-математические методы и модели: Компьютерное моделирование: учебное пособие. – М.: Вузовский учебник, 2007, 2011.
2. Эконометрика: учебник / под ред. И.И. Елисеевой. – 2-е изд., перераб. и доп. – М.: Финансы и статистика, 2005–2008.
3. Многомерный статистический анализ в экономических задачах: Компьютерное моделирование в SPSS / под ред. И.В. Орловой. – М.: Вузовский учебник, 2008, 2011.
4. **Кремер Н.Ш., Путко Б.А.** Эконометрика: учебник для вузов / под ред. проф. Н.Ш. Кремера. – М.: ЮНИТИ-ДАНА, 2003–2008.
5. **Дайитбегов Д.М.** Компьютерные технологии анализа данных в эконометрике. – 2-е изд., испр. и доп. – М.: Вузовский учебник, 2010.

## Содержание

Введение .....	3
1. Общие рекомендации по выполнению и оформлению зачетных работ .....	4
1.1. Требования к оформлению контрольной работы .....	4
1.2. Требования к оформлению отчета по лабораторной работе .....	5
2. Рекомендации по выполнению и оформлению расчетов в Microsoft Excel .....	7
3. Справочные материалы для выполнения расчетов .....	16
4. Комплексный пример исследования экономических данных с использованием корреляционно-регрессионного анализа .....	21
4.1. Выбор факторных признаков для построения регрессионной модели .....	22
4.2. Оценка параметров модели. Экономическая интерпретация коэффициентов регрессии .....	33
4.3. Оценка качества модели регрессии .....	34
4.4. Оценка значимости уравнения регрессии и его коэффициентов .....	35
4.5. Определение объясняющей переменной, от которой может зависеть дисперсия случайных возмущений. Проверка выполнения условия гомоскедастичности остатков по тесту Гольдфельда–Квандта .....	36
4.6. Оценка влияния факторов, включенных в модель, на объем реализации .....	40
4.7. Прогнозирование объема реализации на два месяца вперед .....	41
5. Задания для выполнения контрольной работы .....	44
6. Задания для выполнения лабораторной работы .....	53
Литература .....	94

**Эконометрика.** Компьютерный практикум для студентов третьего курса, обучающихся по специальностям 080105.65 «Финансы и кредит», 080109.65 «Бухгалтерский учет, анализ и аудит». – М.: ВЗФЭИ, 2011.

Редактор Т.А. Балашова  
Корректор О.Э. Стрекачёва  
Компьютерная верстка О.В. Бельнской

ЛР ИД № 00009 от 25.08.99 г.

Подписано в печать 07.09.11. Формат 60×90<sup>1</sup>/<sub>16</sub>.  
Бумага офсетная. Гарнитура Times. Усл.-печ. л. 6,0.  
Изд. № 1/124-11.  
Тираж 200 экз. Заказ № 2282.

Редакционно-издательский отдел  
Всероссийского заочного  
финансово-экономического института (ВЗФЭИ)  
Олеко Дундича, 23, Москва, Г-96, ГСП-5, 123995