

МИНИСТЕРСТВО НАУКИ И ВЫСШЕГО ОБРАЗОВАНИЯ РОССИЙСКОЙ ФЕДЕРАЦИИ

ФЕДЕРАЛЬНОЕ ГОСУДАРСТВЕННОЕ АВТОНОМНОЕ
ОБРАЗОВАТЕЛЬНОЕ УЧРЕЖДЕНИЕ ВЫСШЕГО ОБРАЗОВАНИЯ
«САМАРСКИЙ НАЦИОНАЛЬНЫЙ ИССЛЕДОВАТЕЛЬСКИЙ
УНИВЕРСИТЕТ ИМЕНИ АКАДЕМИКА С.П. КОРОЛЕВА»
(САМАРСКИЙ УНИВЕРСИТЕТ)

О.А. КУЗНЕЦОВА, О.Н. МАЗУРМОВИЧ

РАБОЧАЯ ТЕТРАДЬ ПО ДИСЦИПЛИНЕ ЭКОНОМЕТРИКА

Рекомендовано редакционно-издательским советом федерального государственного автономного образовательного учреждения высшего образования «Самарский национальный исследовательский университет имени академика С.П. Королева» в качестве учебного пособия для обучающихся по основной образовательной программе высшего образования по направлению подготовки 38.03.01 Экономика

САМАРА
Издательство Самарского университета
2022

УДК 330.115(075)

ББК 65в6я7

К891

Рецензенты: д-р экон. наук, проф. Д. Ю. И в а н о в,

д-р экон. наук, проф. В. М. Р а м з а е в

Кузнецова, Ольга Александровна

К891 **Рабочая тетрадь по дисциплине эконометрика:** учебное пособие / *О.А. Кузнецова, О.Н. Мазурмович.* – Самара: Издательство Самарского университета, 2022. – 60 с.

ISBN 978-5-7883-1837-0

Учебное пособие составлено применительно к учебному плану по названным направлениям. В пособии учтены требования государственного образовательного стандарта высшего профессионального образования по вышеуказанным направлениям и стандарта организации СТО Самарского университета 02068410-003–2016. Описаны методы решения эконометрических задач, приведены примеры.

Учебное пособие предназначено для очной формы обучения.

Подготовлено на кафедре математических методов в экономике.

УДК 330.115(075)

ББК 65в6я7

ISBN 978-5-7883-1837-0

© Самарский университет, 2022

ОГЛАВЛЕНИЕ

ТЕМА 1. ПАРНАЯ ЛИНЕЙНАЯ РЕГРЕССИЯ.....	5
Задача 1.....	5
Задание 1.1.	7
Задание 1.2.	8
Задание 1.3.	10
Задание 1.4.	11
Задание 1.5.	12
Задание 1.6.	12
Задание 1.7.	13
Задание 1.8.	14
Задание 1.9.	15
Задание 1.10.	16
Задание 1.11.	16
Задание 1.12.	18
Задание 1.13.	19
Задание 1.14.	20
Задание 1.15.	21
Задание 1.16.	21
Задание 1.17.	23
Задание 1.18.	24
Задание 1.19.	27
Задание 1.20.	28
ТЕМА 2. МНОЖЕСТВЕННАЯ ЛИНЕЙНАЯ РЕГРЕССИЯ	30
Задача 2.....	30
Задание 2.1.	31
Задание 2.2.	35
Задание 2.3.	36
Задание 2.4.	37
Задание 2.5.	37

ТЕМА 3. ЯВЛЕНИЕ МУЛЬТИКОЛЛИНЕАРНОСТИ	39
Задача 3.....	39
ТЕМА 4. ГЕТЕРОСКЕДАСТИЧНОСТЬ И АВТОКОРРЕЛЯЦИЯ ОСТАТКОВ.....	43
Задача 4.....	43
Задание 4.1.	43
Задание 4.2.	44
Задание 4.3.	44
Задание 4.4.	44
ТЕМА 5. ВРЕМЕННЫЕ РЯДЫ.....	49
Задача 5.....	49
Задание 5.1.	49
Задание 5.2.	49
Задание 5.3.	49
Задание 5.4.	49
Задание 5.5.	49
ТЕМА 6. СИСТЕМЫ ЭКОНОМЕТРИЧЕСКИХ УРАВНЕНИЙ..	54
Задача 6.....	54
Задание 6.1.	54
Задание 6.2.	54
Задание 6.3.	54
Задание 6.4.	54
Задание 6.5.	54
СПИСОК ЛИТЕРАТУРЫ.....	57
ПРИЛОЖЕНИЕ 1. Таблица значений F – критерия ФИШЕРА...	58
ПРИЛОЖЕНИЕ 2. Таблица критических значений t-критерия Стьюдента	59

ТЕМА 1. ПАРНАЯ ЛИНЕЙНАЯ РЕГРЕССИЯ

Задача 1.

Имеются следующие данные о зависимости заработной платы по Центральному федеральному округу от курса доллара (табл. 1): x – курс доллара, руб; y – заработная плата, руб.

Таблица 1. Исходные данные к задаче 1

год	x	y		год	x	y
2000	28,16	2173		2011	29,39	28449,4
2001	29,17	3265,9		2012	31,09	32186,2
2002	31,36	4432,6		2013	31,82	36212,5
2003	30,67	5872,8		2014	38,33	39945
2004	28,81	7276,3		2015	61,15	41961
2005	28,30	9621,8		2016	66,96	45943
2006	27,19	12117,4		2017	58,25	48593
2007	25,58	15876,7		2018	62,78	54689
2008	24,85	20665,7		2019	69,47	60770
2009	31,85	22404,6		2020	71,8	65319
2010	30,34	25376,9		2021	76,09	52143

Построить эмпирическое уравнение парной линейной регрессии. Оценить его качество и сделать экономические выводы и прогнозы.

Для этого необходимо выполнить следующие задания:

Вычислить выборочную ковариацию. Сделать вывод о направлении взаимосвязи между изучаемыми признаками.

Построить поле корреляции. Сделать вывод о форме взаимосвязи между признаками X и Y .

Найти выборочный коэффициент корреляции r_{xy} , вычисляя непосредственно по формуле и с помощью функции КОРРЕЛ.

Проверить значимость коэффициента корреляции при уровне значимости 0,05.

Найти коэффициент детерминации R^2 и сформулировать его экономический смысл.

Рассчитать оценки параметров регрессии по формулам и дать экономическую интерпретацию параметров уравнения регрессии.

Добавить на поле корреляции линию тренда, уравнение регрессии и коэффициент детерминации R^2 .

Вычислить остатки регрессии, суммы квадратов отклонений ESS, RSS, TSS и проверить выполнение равенства $TSS=ESS+RSS$.

Округлить оценки параметров и рассчитать остатки регрессии и сумму их квадратов для новых значений оценок параметров. Попытайся так изменить эти значения, чтобы ESS получилась меньше той, что была вычислена в задании 8. Сделать вывод.

Рассчитать стандартные ошибки остатков (S_e) и коэффициентов регрессии (S_{b_0} , S_{b_1}).

Проверить значимость оценок параметров регрессии b_0 и b_1 при уровне значимости $\alpha = 0,05$. Сделать выводы и записать экономический смысл параметров регрессии.

Проверить существенность отличия b_1 от некоторого числа β .

Рассчитать R^2 двумя способами (через ESS и RSS), сравнить с ранее найденным значением коэффициента детерминации и сделать вывод.

Проверить значимость модели в целом с помощью F-теста, сделать вывод и записать экономическую интерпретацию коэффициента детерминации.

Проверить эквивалентность критериев, то есть сравнить $t_{набл}(r)$, $t_{набл}(b_1)$ и $\sqrt{F_{набл}}$. Сделать вывод.

Использовать функцию ЛИНЕЙН для решения задачи 1 и проанализировать полученные результаты.

Задав уровень достоверности $\gamma = 0,95$, построить интервальные оценки для параметров b_0 и b_1 и сделать выводы.

Применить процедуру РЕГРЕССИЯ пакета анализа для решения задачи 1. Проанализировать результаты.

С надежностью $\gamma = 0,95$ построить по модели точечный и интервальный прогнозы заработной платы по Центральному Федеральному округу, курс доллара $x_p = 2$. Сформулировать выводы.

Составить итоговый отчет по задаче 1.

Выполнение заданий

Эмпирическое уравнение парной линейной регрессии имеет вид:

$$\hat{y} = \hat{b}_0 + \hat{b}_1 * x,$$

где \hat{b}_0 и \hat{b}_1 – оценки параметров теоретического уравнения регрессии.

Задание 1.1. Вычислить выборочную ковариацию. Сделать вывод о направлении взаимосвязи между изучаемыми признаками.

Выборочную ковариацию можно вычислить непосредственно по определению:

$$cov(x, y) = \overline{(x - \bar{x}) \cdot (y - \bar{y})} = \frac{1}{n} \sum_{i=1}^n (x_i - \bar{x}) \cdot (y_i - \bar{y}).$$

Для вычисления по этой формуле надо составить расчетную таблицу, первые три столбца которой скопировать из таблицы с исходными данными. Затем добавить столбец для промежуточных вычислений и две строки (*Сумма, Среднее*). Заполнить столбец и строки, выполнив соответствующие расчеты. Значение ковариации будет находиться в добавленном столбце и строке *Среднее*.

Внимание! Справа от расчетной таблицы следует оставить свободные ячейки, так как таблица будет продолжена во время выполнения следующих заданий.

Добавив в расчетную таблицу столбец XY для вычисления произведений $x_i y_i$ и заполнив его, можно найти выборочную ковариацию по альтернативной формуле:

$$\widehat{\text{cov}}(x, y) = \overline{xy} - \bar{x} * \bar{y} = \underline{\hspace{2cm}}$$

Для нахождения $\widehat{\text{cov}}(x, y)$ с помощью функции КОВАР следует:

Установить табличный курсор с свободную ячейку. Нажать на панели инструментов кнопку Вставка функции (f_x).

В появившемся диалоговом окне Мастер функций выбрать категорию Статистические и функцию КОВАР. Нажать кнопку ОК.

Появляется диалоговое окно. В рабочем поле Массив 1 ввести значения переменной X, в рабочее поле Массив 2 ввести значения переменной Y. Нажать кнопку ОК.

В выбранной ячейке появится значение выборочного коэффициента ковариации:

$$\widehat{\text{cov}}(x, y) = \underline{\hspace{2cm}}.$$

Вывод:

Задание 1.2. Построить поле корреляции. Сделать вывод о форме взаимосвязи между признаками X и Y.

Для построения поля корреляции необходимо выполнить следующие действия:

Выделить массив ячеек с исходными числовыми данными (в примере-столбцы X и Y без заголовков). При этом столбец X должен находиться слева от столбца Y.

Щелчком указателя мыши выбрать вкладку Вставка на панели управления. В открывшемся меню щелчком указателя мыши сначала выбрать тип диаграммы (в примере- точечная); затем выбрать подтип диаграммы (в примере – левая, верхняя). Появится диаграмма.

В окне Макет диаграмм (Экспресс-макет) щелчком указателя мыши выбрать первый макет. Заменить название диаграммы на Поле корреляции, название осей – на соответствующие названия признаков (в примере – Курс доллара и Заработная плата) (рис.1).

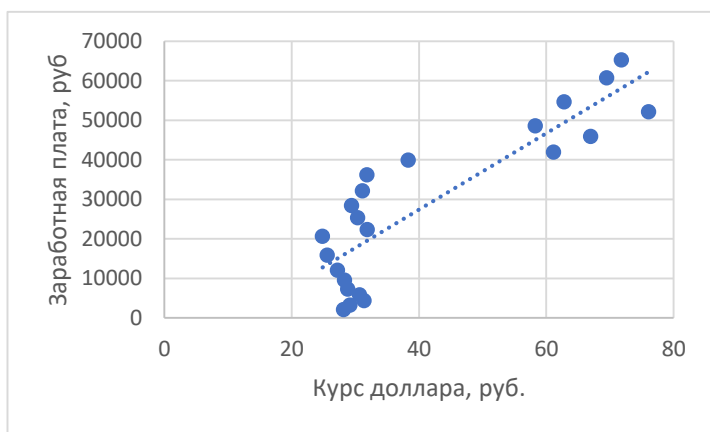


Рис. 1. Диаграмма корреляции заработной платы от курса доллара

Вывод:

Задание 1.3. Найти выборочный коэффициент корреляции r_{xy} , вычисляя непосредственно по формуле и с помощью функции КОРРЕЛ.

Для вычисления выборочного коэффициента корреляции по формуле $r_{xy} = \frac{\overline{x * y - \bar{x} * \bar{y}}}{\sigma_x * \sigma_y} = \frac{\text{COV}(x, y)}{\sigma_x * \sigma_y}$ надо в расчетную таблицу добавить два столбца (X^2, Y^2) и заполнить эти столбцы и строки, выполнив соответствующие расчеты.

Далее нужно найти выборочные дисперсии и средние квадратические отклонения по формулам:

$$D_x = \overline{x^2} - \bar{x}^2 = \underline{\hspace{2cm}};$$

$$\sigma_x = \sqrt{D_x} = \underline{\hspace{2cm}};$$

$$D_y = \overline{y^2} - \bar{y}^2 = \underline{\hspace{2cm}};$$

$$\sigma_y = \sqrt{D_y} = \underline{\hspace{2cm}}.$$

Выборочные дисперсии и средние квадратические отклонения можно найти с помощью функций ДИСПР и СТАНДОТКЛОНП (категория – Статистические) соответственно.

Подставив найденные значения в формулу для r_{xy} , получим

$$r_{xy} = \underline{\hspace{2cm}}.$$

Для нахождения r_{xy} с помощью функции КОРРЕЛ следует:

Установить табличный курсор в свободную ячейку. Нажать на панели инструментов кнопку Вставка функций (f_x).

В появившемся диалоговом окне Мастер функций выбрать категорию Статистические и функцию КОРРЕЛ. Нажать кнопку ОК.

Появится диалоговое окно. В рабочее поле Массив 1 ввести значения переменной X, в рабочее поле Массив 2 ввести значение переменной Y. Нажать кнопку ОК.

В выбранной ячейке появится значение выборочного коэффициента корреляции: $r_{xy} = \underline{\hspace{2cm}}$.

Задание 1.4. Проверить значимость коэффициента корреляции при уровне значимости 0,05.

Для проверки значимости выборочного коэффициента корреляции следует проверить нулевую гипотезу

$$H_0 : r(X, Y) = 0 \text{ (} r_{xy} \text{ не значим).}$$

$$H_1 : r(X, Y) \neq 0 \text{ (} r_{xy} \text{ значим).}$$

Критическая область – $\underline{\hspace{2cm}}$.

Гипотеза H_0 проверяется с помощью критерия

$$t = \frac{r_{xy} * \sqrt{n-2}}{\sqrt{1-r_{xy}^2}}.$$

Подставляя в эту формулу значение r_{xy} и $n = \underline{\hspace{1cm}}$, найти $t_{набл}(r) = \underline{\hspace{1cm}}$.

Критическую точку $t_{кр.05}(\alpha, \kappa)$, где $\alpha = 0,05$, $\kappa = n - 2 = \underline{\hspace{1cm}}$, найти с помощью функции СТЬЮДРАСПОБР следующим образом:

Установить табличный курсор в свободную ячейку. Нажать на панели инструментов кнопку Вставка функций (f_x).

В появившемся диалоговом окне Мастер функций выбрать категорию Статистические и функцию СТЬЮДРАСПОБР. Нажать кнопку ОК.

Появляется диалоговое окно СТЬЮДРАСПОБР. В рабочее поле Вероятность ввести с клавиатуры значение уровня значимости α (в примере -0,05), в рабочее поле Степени свободы ввести число степеней свободы k (в примере – 30). Нажать кнопку ОК.

В ячейке появится значение $t_{кр.об}(0,05;30) = \underline{\hspace{2cm}}$.

Далее сравнить $|t_{набл}(r)|$ и $t_{кр.об}(0,05;30)$ и сделать вывод:

Задание 1.5. Найти коэффициент детерминации R^2 и сформулировать его экономический смысл.

Необходимо найти коэффициент детерминации

$$R^2 = r_{xy}^2 = \underline{\hspace{2cm}}.$$

Учитывая, что $R^2 * 100\%$ показывает, на сколько процентов в среднем вариация объясняемой переменной Y обусловлена вариацией объясняющей переменной X , следует устно сформулировать экономический смысл коэффициента детерминации.

Задание 1.6. Рассчитать оценки параметров регрессии по формулам и дать экономическую интерпретацию параметров уравнения регрессии.

Необходимо найти оценки уравнения регрессии по формулам:

$$\widehat{b}_1 = \frac{\overline{x * y} - \bar{x} * \bar{y}}{D_x} = r_{xy} * \frac{\sigma_y}{\sigma_x} = \underline{\hspace{2cm}};$$

$$\widehat{b}_0 = \bar{y} - \widehat{b}_1 * \bar{x} = \underline{\hspace{2cm}}.$$

Тогда эмпирическое уравнение регрессии запишется в виде:

Принимая во внимание то, что оценка параметра \widehat{b}_1 показывает, как в среднем изменится объясняемая переменная Y , если объясняющая переменная X увеличится на единицу своего измерения, следует устно сформулировать экономический смысл оценки параметра \widehat{b}_1 .

Учитывая, что оценка параметра \widehat{b}_0 показывает, какое в среднем значение примет объясняемая переменная Y , если объясняющая переменная X примет значение, равное нулю, необходимо устно сформулировать экономический смысл оценки параметра \widehat{b}_0 .

Замечание. Оценка константы \widehat{b}_0 уравнения регрессии не всегда допускает экономическую интерпретацию. Она имеет экономический смысл только тогда, когда факторные признаки могут принимать нулевые значения.

Задание 1.7. Добавьте на поле корреляции линию тренда, уравнение регрессии и коэффициент детерминации R^2 .

Для добавления на поле корреляции линии тренда, уравнения регрессии и коэффициента детерминации R^2 необходимо выполнить следующие действия:

Установить указатель мыши на одну из точек поля корреляции и щелкнуть правой кнопкой мыши. В появившемся контекстном меню выбрать пункт Добавить линию тренда.

В появившемся диалоговом окне выбрать тип линии тренда – Линейная и установить флажки в поля показывать уравнение на диаграмме и поместить на диаграмму величину достоверности аппроксимации (R^2). Нажать кнопку Закрыть.

В результате получим на диаграмме линию тренда, уравнение регрессии _____ и $R^2 =$ _____ (рис.2).

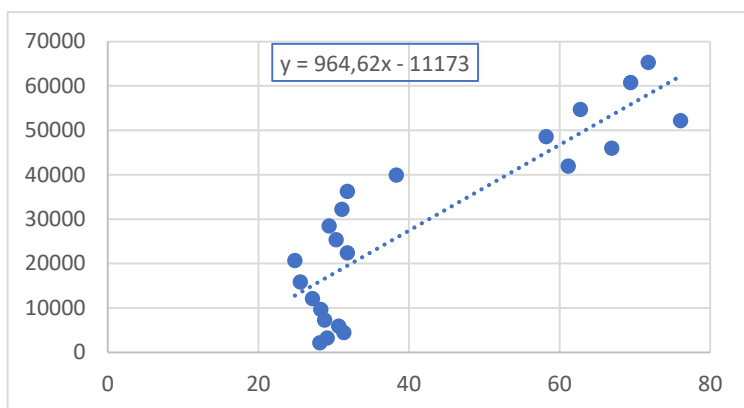


Рис. 2. График линии тренда

Задание 1.8. Вычислить остатки регрессии, суммы квадратов отклонений ESS, RSS, TSS и проверить выполнение равенства $TSS=ESS+RSS$.

Необходимо добавить в расчетную таблицу столбцы для расчетных значений \hat{Y} (\hat{y}), остатков (e) и их квадратов (e^2); заполнить эти столбцы, учитывая, что $\hat{y}_i = \hat{b}_0 + \hat{b}_1 * x_i$, $e_i = y_i - \hat{y}_i$; записать, чему равна сумма остатков и сумма их квадратов:

$$\sum_{i=1}^n e_i = \text{_____}; \quad ESS = \sum_{i=1}^n e_i^2 = \text{_____}.$$

Для вычисления RSS и TSS в расчетную таблицу надо добавить столбцы для вычисления значений $(\hat{y}_i - \bar{y})^2$ и $(y_i - \bar{y})^2$, суммы которых и дадут, соответственно, $RSS = \underline{\hspace{2cm}}$ и $TSS = \underline{\hspace{2cm}}$.

Принимая во внимание, что $ESS + RSS = \underline{\hspace{2cm}}$, можно сделать вывод:

_____.

Задание 1.9. Округлить оценки параметров и рассчитать остатки регрессии и сумму их квадратов для новых значений оценок параметров. Попытаться так изменить эти значения, чтобы ESS получилась меньше той, что была вычислена в задании 8. Сделать вывод.

Необходимо округлить оценки параметров до одной значащей цифры:

$$\hat{b}_0 \approx \underline{\hspace{1cm}}; \hat{b}_1 \approx \underline{\hspace{1cm}}.$$

Далее следует добавить в расчетную таблицу три столбца и сделать расчеты, аналогичные тем, что сделаны в задании 8 при вычислении ESS, но с округленными значениями оценок параметров. Затем необходимо записать в таблицу, расположенную ниже, значения \hat{b}_0, \hat{b}_1 и ESS. Нужно попытаться изменить значения оценок параметров так, чтобы ESS получилась меньше, чем в задании 8 (попробовать несколько вариантов). Результаты требуется занести в таблицу 2:

Таблица 2. **Расчетная таблица для задания 1.9**

\hat{b}_0	\hat{b}_1	ESS

Вывод:

Задание 1.10. Рассчитать стандартные ошибки остатков (S_e) и коэффициентов регрессии ($S_{\hat{b}_0}, S_{\hat{b}_1}$).

Вычислим стандартные ошибки остатков и коэффициентов регрессии по формулам:

$$S_e^2 = \frac{\sum_{i=1}^n e_i^2}{n-2} = \frac{ESS}{n-2} = \underline{\hspace{2cm}};$$

$$S_e = \sqrt{S_e^2} = \underline{\hspace{2cm}}.$$

$$S_{\hat{b}_1} = \frac{S_e}{\sqrt{n} * \sigma_x} = \underline{\hspace{2cm}};$$

$$S_{\hat{b}_0} = S_{\hat{b}_1} * \sqrt{x^2} = \underline{\hspace{2cm}}.$$

Задание 1.11. Проверить значимость оценок параметров регрессии \hat{b}_0 и \hat{b}_1 при условии значимости $\alpha = 0,05$. Сделать выводы и записать эконометрический смысл параметров регрессии.

В задании 6 была найдена оценка параметра $\hat{b}_0 = \underline{\hspace{2cm}}$.

Для проверки значимости этой оценки следует проверить гипотезу:

$H_0 : b_0 = 0$ (оценка \hat{b}_0 не значима).

$H_1 : b_0 \neq 0$ (оценка \hat{b}_0 значима).

Критическая область – _____.

Гипотеза H_0 проверяется с помощью критерия $t = \frac{\hat{b}_0}{S_{\hat{b}_0}}$. Под-

ставляя в эту формулу найденные ранее значения \hat{b}_0 и $S_{\hat{b}_0}$, найти

$$t_{набл}(b_0) = \underline{\hspace{10cm}}.$$

Критическую точку $t_{кр.обс}(\alpha, \kappa)$, где $\alpha = 0,05$, $\kappa = n - 2 = \underline{\hspace{2cm}}$, найти с помощью функции СТЬЮДРАСПОБР аналогично тому, как это сделано в задании 4: $t_{кр.обс}(0,05; \underline{\hspace{2cm}}) = \underline{\hspace{2cm}}$.

Далее сравнить $|t_{набл}(b_0)|$ и $t_{кр.обс}(0,05; 30)$ и сделать вывод:

Замечание. Даже если оценка константы уравнения регрессии \hat{b}_0 оказалась незначимой, то ее все равно оставляют в модели, так как модель с константой всегда лучше, чем модель без константы.

Записать эконометрический смысл оценки параметра \hat{b}_0 (если это возможно):

Аналогично проверим гипотезу о значимости оценки параметра $\hat{b}_1 = \underline{\hspace{2cm}}$, найденной в задании 6. Выдвигаем гипотезы:

H_0 : _____;

H_1 : _____.

Критическая область – _____.

Гипотеза H_0 проверяется с помощью критерия $t = \frac{\widehat{b}_1}{S_{\widehat{b}_1}}$. Под-

ставляя в эту формулу найденные ранее значения \widehat{b}_1 и $S_{\widehat{b}_1}$, найдем

$t_{набл}(b_1) =$ _____.

Далее сравнить $|t_{набл}(b_1)|$ и $t_{кр.дв}(0,05;30)$ и сделать вывод:

Записать экономический смысл оценки параметра \widehat{b}_1 :

Задание 1.12. Проверить существенность отличия b_1 от некоторого числа β .

Для проверки существенности отличия b_1 от некоторого числа, например, от числа $\beta = 50$ следует проверить гипотезу:

H_0 : _____;

H_1 : _____.

Критическая область – _____.

Гипотеза H_0 проверяется с помощью критерия $t = \frac{\hat{b}_1 - \beta}{S_{\hat{b}_1}}$.

Подставляя в эту формулу $\beta = 50$ и найденные ранее значения \hat{b}_1 и $S_{\hat{b}_1}$, найдем $t_{набл} =$ _____. Сравнить его модуль с найденным выше

$$t_{кр.дв}(0,05;30) = \underline{\hspace{2cm}}.$$

Вывод:

Замечание 1. Можно проверить существенность отличия b_1 от $\beta = 50$ при уровне значимости α с помощью интервальной оценки этого параметра, построенной с достоверностью $\gamma = 1 - \alpha$. Если $\beta = 50$ принадлежит полученному интервалу, то это отличие несущественное, а если не принадлежит, то отличие значимое.

Замечание 2. Гипотеза такого вида проверяется только в случае практической необходимости, когда есть основания предполагать, что параметр b_1 равен некоторому числу β .

Задание 1.13. Рассчитать R^2 двумя способами (через ESS и через RSS), сравнить с ранее найденным значением коэффициента детерминации и сделать вывод.

Вычислить R^2 двумя способами:

$$R^2 = \frac{RSS}{TSS} = \underline{\hspace{2cm}}; \quad R^2 = 1 - \frac{RSS}{TSS} = \underline{\hspace{2cm}}.$$

Сравнить полученные числа между собой, со значением R^2 , вычисленным в задании 5, и значением R^2 , показанным на поле корреляции (рис. 2).

Вывод:

Задание 1.14. Проверить значимость модели в целом с помощью F-теста, сделать вывод и записать экономическую интерпретацию коэффициента детерминации ($\alpha = 0,05$).

В заданиях 5 и 13 найдено значение $R^2 =$ _____.

Для проверки гипотезы о значимости R^2 , а следовательно, о значимости модели в целом выдвигают нулевую гипотезу:

H_0 : _____

H_1 : _____

Критическая область – _____.

Подставляя найденные значения RSS , ESS и R^2 в формулы $F = \frac{RSS * (n - 2)}{ESS} = \frac{R^2 * (n - 2)}{1 - R^2}$, найти наблюдаемое значение критерия $F_{набл} =$ _____.

Затем найти с помощью функции FРАСПОБР критическую точку $F_{кр}(\alpha; k_1; k_2)$, учитывая, что $\alpha = 0,05$, $k_1 = m =$ _____,

$k_2 = n - m - 1 =$ _____. Функция ФРАСПОБР используется аналогично функции СТЬЮДРАСПОБР, только необходимо ввести в соответствующие рабочие поля не одну, а две степени свободы.

$$F_{кр}(0,05; \text{____}; \text{____}) = \text{_____}.$$

Далее сравнить $F_{набл}$ и $F_{кр}(0,05; 1; 30)$ и сделать вывод:

Экономический смысл коэффициента детерминации:

Задание 1.15. Проверить эквивалентность критериев, то есть сравнить $|t_{набл}(r)|, |t_{набл}(b_1)|, u\sqrt{F_{набл}}$. Сделать вывод.

Найти и сравнить значения:

$$|t_{набл}(r)| = \text{_____}; |t_{набл}(b_1)| = \text{_____}; \sqrt{F_{набл}} = \text{_____}.$$

Вывод:

Задание 1.16. Использовать функцию ЛИНЕЙН для решения задачи 1 и проанализировать полученные результаты.

Функция ЛИНЕЙН имеет параметры;

ЛИНЕЙН (*известные_значения_у*; *известные_значения_x*; *константа*; *статистика*). Здесь:

известные_значения_у – множество наблюдаемых значений y ;

известные_значения_x – множество наблюдаемых значений x .

Причем, если массив *известные_значения_у* имеет один столбец, то каждый столбец массива *известные_значения_x* интерпретируется как отдельная переменная;

Константа – логическое значение, которое указывает, требуется ли, чтобы константа b_0 была равна нулю. При этом, если константа имеет значение 1 (ИСТИНА) или по умолчанию, то b_0 вычисляется обычным образом, а если константа имеет значение 0 (ЛОЖЬ), то b_0 полагается равным нулю;

Статистика – логическое значение, которое указывает, требуется ли вычислить дополнительную статистику по регрессии. Если введено значение 1 (ИСТИНА), то дополнительные параметры вычисляются, если введено значение 0 (ЛОЖЬ), то – нет.

Для использования функций ЛИНЕЙН надо выполнить следующие действия:

Выделить блок ячеек под массив результатов – 5 строк и 2 столбца (по числу определяемых параметров). Нажать на панели инструментов кнопку Вставка функций (f_x).

В появившемся диалоговом окне Мастер функций выбрать категорию Статистические и функцию ЛИНЕЙН. Нажать кнопку ОК.

Появляется диалоговое окно ЛИНЕЙН. В рабочее поле Изв_знач_у ввести наблюдаемые значения y ; в рабочее поле Изв_знач_x ввести наблюдаемые значения x ; в рабочее поле Константа ввести с клавиатуры 1; в рабочее поле Статистика ввести с клавиатуры 1. Нажать на кнопку ОК.

Нажать клавишу F2. Затем нажать сочетание клавиш CTRL+SHIFT+ENTER.

Внимание! Не следует пытаться сделать это одномоментно. Надо левой рукой нажать CTRL+SHIFT и, удерживая эти клавиши, нажать клавишу ENTER правой рукой.

В результате в выделенном блоке ячеек получим числовые данные.

Анализируя полученные результаты, приходим к выводу, что эта таблица содержит значения следующих величин:

Таблица 3. Результаты выполнения задания 1.16

Задание 1.17. Задав уровень достоверности $\gamma = 0,95$, построить интервальные оценки для параметров b_0 и b_1 и сделать выводы.

Используя формулу

$$\widehat{b}_0 - S_{\widehat{b}_0} * t_{кр.дв}(\alpha; \kappa) < b_0 < \widehat{b}_0 + S_{\widehat{b}_0} * t_{кр.дв}(\alpha; \kappa)$$

С учетом того, что при уровне достоверности $\gamma = 0,95(95\%)$ $\alpha = 1 - \gamma = \underline{\hspace{2cm}}$ и $\kappa = n - 2 = \underline{\hspace{2cm}}$, построить интервальную оценку параметра b_0 :

$$\underline{\hspace{4cm}} < b_0 < \underline{\hspace{4cm}}$$

Принимая во внимание экономическую интерпретацию параметра b_0 , сделать вывод:

Аналогично, используя формулу

$$\widehat{b}_1 - S_{\widehat{b}_1} * t_{кр.об}(\alpha; k) < b_1 < \widehat{b}_1 + S_{\widehat{b}_1} * t_{кр.об}(\alpha; k),$$

Построить интервальную оценку для параметра b_1 :

$$\underline{\hspace{2cm}} < b_1 < \underline{\hspace{2cm}}$$

и записать экономический смысл:

Замечание. Интервальные оценки параметров можно использовать для проверки значимости соответствующего параметра. Если интервальная оценка параметра содержит ноль, то параметр не значим, в противном случае параметр значим.

Задание 1.18. Применить процедуру РЕГРЕССИЯ пакета анализа для решения задачи 1. Проанализировать результаты.

Табличный процессор MS Excel позволяет быстро делать необходимые статистические расчеты с помощью надстройки «Пакета анализа».

Внимание! При работе с пакетом анализа в Excel на листе не должно быть активных диаграмм.

Для реализации процедуры РЕГРЕССИЯ необходимо:

На панели управления выбрать вкладку Данные. В открывшемся окне щелчком указателя мыши выбрать Анализ данных и указать курсором не строке Регрессия.

Внимание! Если в окне Данные нет пункта Анализ данных, то следует нажать кнопку настройки панели быстрого доступа и в открывшемся списке выбрать пункт Другие команды. В левом списке открывшегося окна надо щелкнуть по строке Надстройки, и в нижней части появившегося справа окна нажать на кнопку Перейти... В открывшемся окне следует установить флажок в поле Пакет анализа. Нажать на кнопку ОК и перейти к выполнению пункта 1.

В появившемся диалоговом окне в рабочее поле Входной интервал Y ввести ссылку на диапазон анализируемых значений зависимой переменной, содержащих один столбец данных.

В рабочее поле Входной интервал X ввести ссылку на диапазон значений независимых переменных. Этот диапазон может содержать несколько столбцов.

Если столбцы анализируемых данных вводились с шапкой, то установить флажок в поле Метки. Если нужно вывести остатки, графики остатков или график подбора, то установить флажки в соответствующие поля.

Указать выходной диапазон. Для этого поставить флажок перед словами Выходной интервал. Затем установить курсор в поле Выходной интервал и щелкнуть левой кнопкой мыши по пустой ячейке, справа и снизу от которой все ячейки свободны.

Если необходимо получить остатки регрессии и построить графики, то надо установить флажки в следующих полях: Остатки, График остатков, График подбора. Нажать кнопку ОК.

Результаты реализации процедуры РЕГРЕССИЯ

В выходном диапазоне появятся результаты. Переписать значения в таблицы, которые располагаются ниже:

Таблица 4. Результаты расчёта показателей для регрессионного анализа

<i>Регрессионная статистика</i>	
Множественный R	
R-квадрат	
Нормированный R-квадрат	
Стандартная ошибка	
Наблюдения	

<i>Дисперсионный анализ</i>					
	<i>df</i>	<i>SS</i>	<i>MS</i>	<i>F</i>	<i>Значимость F</i>
Регрессия					*
Остаток					
Итого					

	Коэф.	Ст. ошиб.	t-стат.	P-знач.	Нижн. 95%	Верх. 95%
Y-перес.				**		
X				**		

*

**

Замечание. Здесь в последней таблице не показаны еще два столбца, аналогичные двум последним столбцам.

В таблице Вывод остатка указаны следующие величины:

В столбце Наблюдение _____;

В столбце Предсказанное Y _____;

В столбце Остатки _____.

Анализ графика подбора: _____

Анализ графика остатков: _____

Задание 1.19. С надежностью $\gamma = 0,95$ построить по модели точечный и интервальные прогнозы цены кондиционера, потребляемая мощность которого равна 2 кВт ($x_p=2$). Сформулировать выводы.

Построить точечный и интервальные прогнозы для прогнозируемого значения потребляемой мощности ($x_p=2$ кВт) по формулам:

$$\hat{y}_p = \hat{b}_0 + \hat{b}_1 * x_p = \text{_____} \text{ (точечный прогноз),}$$

$\hat{y}_p - S_{\square p} * t_{кр.дв}(\alpha; k) < y_p < \hat{y}_p + S_{\square p} * t_{кр.дв}(\alpha; k)$ (интервальный прогноз),

Где $\alpha = 1 - \gamma = \text{_____}$, $k = n - 2 = \text{_____}$,

$$S_{\square p} = S_e \sqrt{1 + \frac{1}{n} + \frac{(x_p - \bar{x})^2}{n * \sigma_x^2}} = \text{_____} \text{ – стандартная}$$

ошибка прогноза.

Получим _____ $< y_p <$ _____.

Вывод:

Задание 1.20. Составить итоговый отчет по задаче 1.

В задаче 1 изучалась зависимость между признаками:

(указать источник информации). При этом были найдены следующие значения:

Таблица 5. Результаты выполнения задания 1.20

ESS	RSS	TSS	S_e	$S_{\hat{b}_0}$	$S_{\hat{b}_1}$

Показатель	Значение	Наблюдаемое значение критерия	Критическая точка	Значимость
r_{xy}				
\hat{b}_0				
\hat{b}_1				
R^2				

В результате получили:

Эмпирическое уравнение регрессии

Интервальные оценки параметров при $\gamma = 0,95$

_____ $< b_0 <$ _____

_____ $< b_1 <$ _____

Точечный и интервальный ($\gamma = 0,95$) прогнозы значения резуль-
тативного признака для $x_p = 2кВт$

$$\hat{y}_p = \text{_____}; \text{_____} < y_p < \text{_____}.$$

Экономические выводы:

ТЕМА 2. МНОЖЕСТВЕННАЯ ЛИНЕЙНАЯ РЕГРЕССИЯ

Задача 2.

Имеются данные по прибыли (Y , тыс. у.е.), среднему числу посетителей (X_1 , чел/день) и площади (X_2 , m^2) двенадцати интернет-кафе (табл. 4):

Таблица 6. Исходные данные для задачи 4

№	Y	X_1	X_2
1	3,9	5	59,7
2	17,1	15	71,8
3	16,1	16	82,4
4	3,1	9	37
5	6,3	12	40,2
6	17,8	14	110,4
7	11,3	22	60
8	3,4	10	24
9	4,7	10	49,1
10	19,7	19	99,4
11	18,2	21	67,1
12	22,3	23	118,4

Построить эмпирическое уравнение множественной линейной регрессии, оценить его качество и сделать экономические выводы и прогнозы (уровень значимости взять равным 0,02).

Для этого выполнить следующие задания:

1. Найти МНК-оценки коэффициентов уравнения регрессии:
 - а) с помощью матричных преобразований;
 - б) применив функцию ЛИНЕЙН;

в) используя процедуру РЕГРЕССИЯ пакета анализа.

Проанализировать результаты.

2. Проверить статистическую значимость всех параметров модели, применяя различные способы ($\alpha = 0,02$). Дать экономическую интерпретацию оценок параметров уравнения регрессии.

3. Найти коэффициент детерминации и проверить качество модели в целом при уровне значимости 0,02. Сформулировать экономический смысл коэффициента детерминации.

4. С надежностью 0,98 найти интервальные оценки истинных коэффициентов регрессии. Сделать экономические выводы.

5. С надежностью 0,98 найти интервальную оценку для индивидуального прогнозного значения прибыли интернет-кафе, площадь которого равна $x_{p2} = 70 \text{ м}^2$, если среднее число посетителей в день будет равно $x_{p1} = 15$ человек. Записать экономический смысл точечного и интервального прогноза.

Выполнение заданий

Эмпирическое уравнение множественной линейной регрессии имеет вид $\hat{y} = \hat{b}_0 + \hat{b}_1 * x_1 + \hat{b}_2 * x_2$, где \hat{b}_0, \hat{b}_1 и \hat{b}_2 – оценки параметров теоретического уравнения регрессии $y = b_0 + b_1 * x_1 + b_2 * x_2 + \varepsilon$.

Задание 2.1. Найти МНК-оценки коэффициентов уравнения регрессии.

Вычислим МНК – оценки коэффициентов уравнения регрессии:

А. Использовать матричные преобразования по формуле

$$\vec{\hat{b}}_{\text{МНК}} = (X^T X)^{-1} X^T Y, \text{ где}$$

$$X = \begin{bmatrix} 1 & 5 & 59,7 \\ 1 & 15 & 71,8 \\ 1 & 16 & 82,4 \\ 1 & 9 & 37 \\ 1 & 12 & 40,2 \\ 1 & 14 & 110,4 \\ 1 & 22 & 60 \\ 1 & 10 & 24 \\ 1 & 10 & 49,1 \\ 1 & 19 & 99,4 \\ 1 & 21 & 67,1 \\ 1 & 23 & 118,4 \end{bmatrix}; \quad Y = \begin{bmatrix} 3,9 \\ 17,1 \\ 16,1 \\ 3,1 \\ 6,3 \\ 17,8 \\ 11,3 \\ 3,4 \\ 4,7 \\ 19,7 \\ 18,2 \\ 22,3 \end{bmatrix}; \quad \vec{\hat{b}}_{МНК} = \begin{bmatrix} \hat{b}_0 \\ \hat{b}_1 \\ \hat{b}_2 \end{bmatrix}.$$

X^T – матрица, транспонированная к матрице X ;

$(X^T X)^{-1}$ – матрица, обратная к матрице $(X^T X)$.

Транспонированную матрицу X^T можно найти с помощью функции ТРАНСП следующим образом:

1. Выделить блок ячеек под транспонированную матрицу – 3 строк, 12 столбцов. Нажать на панели инструментов кнопку Вставка функции (f_x).

2. В появившемся диалоговом окне Мастер функций выбрать категорию Ссылки и массивы и функцию ТРАНСП. Нажать кнопку ОК.

3. В рабочее поле Массив ввести матрицу X . Нажать кнопку ОК.

4. Нажать клавишу F2. Затем нажать сочетание клавиш CTRL+SHIFT+ENTER.

5. В результате в выделенном блоке ячеек получим X^T .

Для нахождения произведения матриц используют функцию МУМНОЖ:

1. Выделить блок ячеек под результат – 3 строки, 3 столбца. Нажать на панели инструментов кнопку Вставка функций (f_x).

2. В появившемся диалоговом окне Мастер функций выбрать категорию Математические и функцию МУМНОЖ. Нажать кнопку ОК.

3. В рабочее поле Массив1 ввести матрицу X^T ; в рабочее поле Массив2 – матрицу X. Нажать кнопку ОК.

4. Нажать клавишу F2. Затем нажать сочетание клавиш CTRL+SHIFT+ENTER.

5. В результате в выделенном блоке ячеек получим произведение матриц ($X^T X$).

Для нахождения обратной матрицы надо аналогично воспользоваться функцией МОБР (категория – Математические). Затем дважды применить функцию МУМНОЖ, сначала для вычисления произведения ($X^T Y$), и далее для нахождения вектора эмпирических коэффициентов регрессии. Получим

$$\vec{b}_{МНК} = (\text{_____}; \text{_____}; \text{_____})^T.$$

Б. Применить функцию ЛИНЕЙН аналогично тому, как это сделано при выполнении задания 16 задачи 1, с той лишь разницей, что блок ячеек под результат будет состоять из 3 столбцов и 5 строк; и в рабочее поле Изв_знач_x следует ввести наблюдаемые значения X_1 и X_2 (оба столбца вместе).

Получим следующие результаты:

Таблица 7. Результаты выполнения задания 2.1

В. Использовать процедуру РЕГРЕССИЯ пакета анализа аналогично тому, как это сделано при выполнении задания 1.19, толь-

ко при заполнении рабочего поля Входной интервал X следует ввести не один столбец, а все столбцы, содержащие значения факторных признаков.

Таблица 8. Результаты применения процедуры РЕГРЕССИЯ

<i>Регрессионная статистика</i>	
Множественный R	
R- квадрат	
Нормированный R- квадрат	
Стандартная ошибка	
Наблюдения	

<i>Дисперсионный анализ</i>					
	<i>df</i>	<i>SS</i>	<i>MS</i>	<i>F</i>	<i>Значимость F</i>
Регрессия					
Остаток					
Итого					

	Коэф.	Ст.ошиб	t-стат.	P-знач.	Нижн. 95%	Верх. 95%
Y-перес.						
X ₁						
X ₂						

Выводы:

Задание 2.2. Проверить статистическую значимость всех параметров модели, применяя различные способы ($\alpha = 0,02$). Дать экономическую интерпретацию оценок параметров уравнения регрессии.

Для проверки значимости параметров выдвигаются гипотезы:

$H_0 : b_j = 0$ (оценка \widehat{b}_j статистически не значима);

$H_1 : b_j \neq 0$ (оценка \widehat{b}_j статистически значима).

Наблюдаемое значение t-критерия для каждого из параметров вычисляют по формуле

$$t = \widehat{b}_j / S_{\widehat{b}_j}, j = \overline{0,2}.$$

$$t_{набл}(b_0) = \underline{\hspace{2cm}}; t_{набл}(b_1) = \underline{\hspace{2cm}};$$

$$t_{набл}(b_2) = \underline{\hspace{2cm}}.$$

Критическую точку находят, используя функцию СТЬЮДРАС-ПОБР, учитывая, что число степеней свободы

$$k = n - m - 1 = 12 - 2 - 1 = 9.$$

$$t_{кр.дв}(0,02;9) = \underline{\hspace{2cm}}.$$

Вывод:

Экономическая интерпретация оценок параметров уравнения регрессии:

Проверить значимость эмпирических коэффициентов регрессии, используя столбец Р-значение и интервальные оценки, полученные при выполнении процедуры РЕГРЕССИЯ.

Задание 2.3. Найти коэффициент детерминации и проверить качество модели в целом при уровне значимости 0,02. Сформулировать экономический смысл коэффициента детерминации.

Коэффициент детерминации построенной модели

$$R^2 = \underline{\hspace{2cm}} .$$

Для проверки качества модели в целом выдвигают гипотезы:

$$H_0 : R^2 = 0 \text{ (модель в целом не значима);}$$

$$H_1 : R^2 > 0 \text{ (модель в целом значима).}$$

Наблюдаемое значение критерия $F_{набл} = \underline{\hspace{2cm}} .$

Критическую точку находят с помощью функции ФРАСПОБР, учитывая, что уровень значимости равен 0,02 и число степеней свободы $k_1 = m = \underline{\hspace{2cm}} ; k_2 = n - m - 1 = \underline{\hspace{2cm}} .$

$$F_{кр} (\underline{\hspace{2cm}} ; \underline{\hspace{2cm}} ; \underline{\hspace{2cm}}) = \underline{\hspace{2cm}} .$$

Вывод:

Экономический смысл коэффициента детерминации:

Проверить значимость модели в целом с помощью столбца Значимость F процедуры РЕГРЕССИЯ.

Задание 2.4. С надежностью 0,98 найти интервальные оценки истинных коэффициентов регрессии. Сделать экономические выводы.

Интервальные оценки истинных коэффициентов регрессии получены с надежностью 0,98 при проведении процедуры РЕГРЕССИЯ:

$$\text{_____} < b_0 < \text{_____};$$

$$\text{_____} < b_1 < \text{_____};$$

$$\text{_____} < b_2 < \text{_____}.$$

Экономические выводы:

Задание 2.5. С надежностью 0,98 найти интервальную оценку для индивидуального прогноза значения прибыли интернет-кафе, площадь которого равна $x_{p2} = 70\text{м}^2$, если среднее число посетителей в день будет равно $x_{p1} = 15\text{чел}$. Записать экономический смысл точечного и интервального прогноза.

Интервальная оценка для индивидуального прогноза значения y_p имеет вид:

$$\hat{y}_p - t_{кр.06} (1 - \gamma; \kappa) * S_{y_p - \hat{y}_p} < y_p < \hat{y}_p + t_{кр.06} (1 - \gamma; \kappa) * S_{y_p - \hat{y}_p},$$

где $\hat{y}_p = \hat{b}_0 + \hat{b}_1 x_{p1} + \hat{b}_2 x_{p2} = \underline{\hspace{2cm}}$ – точечная оценка индивидуального прогнозного значения y ;

$$k = n - m - 1 = \underline{\hspace{2cm}};$$

$$t_{кр.дв}(1 - \gamma; k) = t_{кр.дв}(\underline{\hspace{1cm}}; \underline{\hspace{1cm}}) = \underline{\hspace{2cm}};$$

$$S_{y_p - \hat{y}_p} = S_e \sqrt{1 + \vec{x}_p^T (X^T X)^{-1} \vec{x}_p} = \underline{\hspace{2cm}};$$

$$\vec{x}_p^T = (1; x_{p1}; x_{p2}) = (1; 15; 70).$$

Величину $\vec{x}_p^T (X^T X)^{-1} \vec{x}_p = \underline{\hspace{2cm}}$ находят, используя найденную при выполнении задания 1 обратную матрицу $(X^T X)^{-1}$ и функцию МУМНОЖ.

В результате вычислений получим:

$$\underline{\hspace{2cm}} < y_p < \underline{\hspace{2cm}}$$

Экономический смысл точечного и интервального прогноза:

ТЕМА 3. ЯВЛЕНИЕ МУЛЬТИКОЛЛИНЕАРНОСТИ

Задача 3.

В результате маркетингового исследования получены статистические данные зависимости цены ноутбуков (Y , руб.) от следующих технических характеристик:

X_1 – диагональ, дюйм;

X_2 – частота процессора, ГГц;

X_3 – размер оперативной памяти, Гб;

X_4 – объем жесткого диска, Гб;

X_5 – время непрерывной работы, ч.

Таблица 5. Исходные данные для задачи 3

№	Y	X_1	X_2	X_3	X_4	X_5
1	32789	15,6	2,6	8	1000	6
2	15999	15,5	1,65	4	500	5
3	15999	13	1,7	2	500	8
4	16999	15,6	2,2	2	320	3
5	27199	17,3	2,3	8	750	4,5
6	25299	14	2,5	4	640	4,3
7	31649	14	1,7	4	500	7
8	10699	10,1	1,6	1	320	5
9	10399	10,1	1,6	1	320	5
10	11099	10,1	1	2	320	4
11	15999	15,6	2,2	4	500	3
12	27199	15,6	2,4	4	750	4,5
13	23999	15,6	2,3	8	1000	4
14	24299	13,3	2,3	4	320	5
15	11599	11,6	1	2	320	7
16	27899	15,6	2,5	4	750	4
17	12599	15,6	1,6	2	320	4
18	16599	15,6	1,3	4	320	5
19	40699	14,4	1,7	6	500	7,5

20	11149	10,1	1,6	2	320	8
21	17999	15,6	1,5	4	500	4
22	18599	11,6	1,65	2	320	4
23	14999	15,6	1,6	4	500	6
24	17299	11,2	1,65	2	320	5
25	19799	14	2,1	4	500	6
26	19599	15,6	1,4	4	320	5
27	20399	15,6	2,7	6	750	6
28	22599	13,3	1,5	4	500	6,5
29	25999	15,6	2,3	4	500	5
30	30499	15,6	2,3	4	750	6

Построить линейную регрессионную модель и оценить ее качество. При наличии признаков мультиколлинеарности провести анализ матрицы парных коэффициентов корреляции и построить прогнозную модель с учётом мультиколлинеарности.

Порядок выполнения работы:

1. Построить матрицу коэффициентов парной корреляции, используя MS Excel «Данные – Анализ данных – Корреляция».

На рис. 3. Показано окно для ввода данных в котором нужно ввести соответствующую информацию.

«Входной интервал» предполагает выделение всех данных исходной таблицы вместе с заголовками. Важно!!! поставить значок «Метки в первой строке», тогда в итоговой таблице будут отражены названия переменных.

Выходной интервал – координаты любой свободной ячейки.

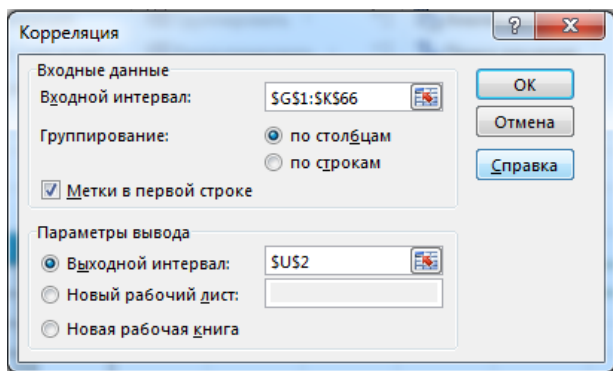


Рис. 3. Ввод данных для построения матрицы парных коэффициентов корреляции

В результате будет получена матрица коэффициентов парной корреляции (рис. 4).

	y_1	x_1	x_2	...	x_n
y_1	1				
x_1	0,03	1			
x_2	0,7	-0,4	1		
...	
x_n	-0,18	0,08	0,03	-0,4	1

Рис. 4. Матрица коэффициентов парной корреляции

2. Выбрать переменные, входящие в модель.

1) Первый столбец показывает коэффициенты парной корреляции между переменными x и y . Важна сильная связь. Переменные, соответствующие значениям, выделенным в первом столбце, включаем в модель. Для этого в первом столбце отмечаем все значения больше или равные 0,7 по модулю. Указываем подходящие переменные (x_2 , ... и т.д.).

2) Остальные столбцы показывают коэффициенты парной корреляции между переменными x . Сильная связь (больше или равно 0,5 по

модулю) означает наличие мультиколлинеарности. Переменные, соответствующие этим значениям, исключаем из модели. Переменные, соответствующие этим значениям, нельзя включать в одну модель и нужно составить несколько моделей.

Например, на первом этапе выявили сильную корреляционную связь между y и переменными $r_{y1x2}=0,8$, $r_{y1x3}=-0,7$, $r_{y1x5}=0,9$.. Далее определили $r_{x2x3}=-0,6$, $r_{x2x5}=0,1$, $r_{x3x5}=-0,2$. Соответственно, мы можем построить модели $y_1=f(x_2, x_5)$ и $y_1=f(x_3, x_5)$, т.е. $y_1=b_0+b_2x_2+b_5x_5+e$ и $y_1=b_0+b_3x_3+b_5x_5+e$. Потом находим коэффициенты регрессий и проверяем качество полученных моделей. Выбираем для прогнозирования наилучшую, при условии, что коэффициент детерминации модели $> 0,8$.

Указание. Решить задачу с помощью инструментов табличного процессора MS Excel. Файл с решением и необходимыми пояснениями показать преподавателю. Записать итоговую модель и сделать экономические выводы:

ТЕМА 4. ГЕТЕРОСКЕДАСТИЧНОСТЬ И АВТОКОРРЕЛЯЦИЯ ОСТАТКОВ

Задача 4.

В результате проведенного исследования была выявлена зависимость цены ультрабука (Y , тыс.руб.) от массы (X_1 , кг) и общего индекса производительности Windows (X_2). Исходные данные приведены в таблице 6.

Таблица 6. Исходные данные для задачи 4

№	Y	X_1	X_2	№	Y	X_1	X_2
1	61,8	1,30	9,0	16	35,5	1,30	6,0
2	60,0	1,25	8,5	17	33,8	1,35	6,0
3	58,3	1,20	8,6	18	32,0	1,40	6,0
4	56,5	1,15	8,4	19	30,3	1,45	6,0
5	54,8	1,10	7,0	20	30,5	1,50	6,0
6	53,0	1,05	8,0	21	26,8	1,55	5,0
7	51,3	1,00	7,00	22	25,0	1,60	4,8
8	49,5	0,95	7,0	23	23,3	1,65	4,6
9	47,8	0,90	7,0	24	21,5	1,70	4,4
10	46,0	1,00	6,0	25	19,8	1,75	4,2
11	44,3	1,05	8,0	26	18,0	1,80	4,0
12	42,5	1,10	5,0	27	17,5	1,85	3,8
13	40,8	1,15	6,0	28	17,0	1,90	3,6
14	39,0	1,20	6,4	29	16,3	1,50	3,4
15	37,3	1,25	6,2	30	16,0	0,90	3,2

Задание 4.1. Построить по МНК линейную регрессионную модель и оценить ее качество.

Задание 4.2. Провести графический анализ остатков.

Задание 4.3. Для каждой независимой переменной выполнить проверку наличия гетероскедастичности остатков с помощью теста ранговой корреляции Спирмена.

Задание 4.4. В случае выявления гетероскедастичности применить ОМНК. Сделать выводы.

Краткие результаты решения и выводы:

Тест Спирмена:

При использовании теста Спирмена предполагается, что дисперсия отклонения будет либо увеличиваться или уменьшаться, или уменьшаться с увеличением значений X . Поэтому для регрессии, построенной по методу МНК, абсолютные величины отклонение e и значения x объясняющей переменной X будут коррелированы. Значения x и e ранжируются, а затем определяется коэффициент ранговой корреляции.

Построить уравнение регрессии, любым известным способом найти его коэффициенты:

$$y = 0.16x + 7.04 + e$$

Проверка наличия гетероскедастичности

1) Методом графического анализа остатков.

В этом случае по оси абсцисс откладываются значения объясняющей переменной X , а по оси ординат либо отклонения e_i , либо их квадраты e_i^2 .

Если имеется определенная связь между отклонениями, то гетероскедастичность имеет место. Отсутствие зависимости скорее всего будет свидетельствовать об отсутствии гетероскедастичности (рис. 5).

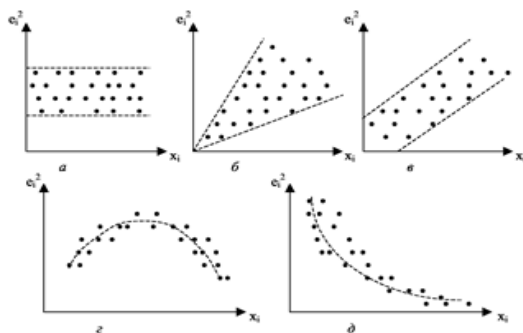


Рис. 5. Графический анализ остатков

2) При помощи теста ранговой корреляции Спирмена.

Присвоим ранги признаку e_i и фактору X . Найдем сумму разности квадратов d^2 (табл.7).

Таблица 7. Вспомогательная таблица для теста Спирмена

X	26	27	27	30	36	39	39	39	40	42	43	44	45	46	45,5	48	50	52	56	59
e_i	3,5	0,1	3,4	1,5	0,9	3,2	1	4,6	3	0,3	1,2	2,4	7,4	3,4	0,38	1,4	3,4	4	0,5	1
ранг X_i d_x	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
ранг e_i , d_y	26	1	23	15	8	21	7	29	19	2	10	17	37	22	4	13	24	27	5	11
$(d_x - d_y)^2$	625	1	400	121	9	225	0	441	100	64	1	25	576	56	110	9	49	81	196	81
X	61	62	63	65	70	71	74	75	76	77	79	82	82	83	83	86	86	87	88	89
e_i	5,7	0,6	6,6	6,6	11	10	4	3	7,2	0,9	0,3	1,4	9	2,7	3,44	1,5	2,3	7	6,4	6
ранг X_i d_x	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36	37	38	39	40
ранг e_i , d_y	30	6	32	34	40	39	28	20	36	9	3	12	38	18	25	14	16	35	33	31
$(d_x - d_y)^2$	81	256	81	100	225	169	1	64	49	441	784	400	25	256	100	484	441	9	36	81

Сумма $(d_x - d_y)^2$ равна 7253,5.

По формуле вычислим коэффициент ранговой корреляции Спирмена.

$$r = 1 - 6 \frac{\sqrt{d^2}}{n^3 - n}$$

Если среди значений признаков x и y встречается несколько одинаковых, образуются связанные ранги, т.е. одинаковые средние номера; например, вместо одинаковых по порядку третьего и четвертого значений признака будут два ранга по 3,5. В таком случае коэффициент Спирмена вычисляется как:

$$r = 1 - 6 \frac{\sqrt{d^2} - A - B}{\sqrt{(n^3 - n - 12A)(n^3 - n - 12B)}}, \text{ где}$$

$$A = \frac{1}{12} \sum (A_j^3 - A_j), \quad B = \frac{1}{12} \sum (B_k^3 - B_k)$$

j – номера связок по порядку для признака x ;

A_j – число одинаковых рангов в j -й связке по x ;

k – номера связок по порядку для признака y ;

B_k – число одинаковых рангов в k -й связке по y .

$$A = 6/12 = 0.5$$

$$B = 0/12 = 0$$

$$r = 1 - 6 \frac{7253.5 - 0.5 - 0}{\sqrt{(40^3 - 40 - 12 \times 0.5)(40^3 - 40 - 12 \times 0)}} = 0.32$$

Связь между признаком e_i и фактором X слабая и прямая
Оценка коэффициента ранговой корреляции Спирмена.
Значимость коэффициента ранговой корреляции Спирмена

$$T_n = r_{xy} \frac{\sqrt{n-2}}{\sqrt{1-r_{xy}^2}} = 0.32 \frac{\sqrt{38}}{\sqrt{1-0.32^2}} = 2.08$$

По таблице Стьюдента находим $t_{\text{табл}}$:

$$t_{\text{табл}}(n-m-1; \alpha/2) = (38; 0.05/2) = 2.021$$

Поскольку $T_{\text{набл}} > t_{\text{табл}}$, то отклоняем гипотезу о равенстве 0 коэффициента ранговой корреляции. Другими словами, коэффициент ранговой корреляции статистически – значим.

Проверим гипотезу H_0 : гетероскедастичность отсутствует.

Поскольку $2.021 < 2.08$, то гипотеза об отсутствии гетероскедастичности отвергается.

Краткие результаты решения и выводы:

ТЕМА 5. ВРЕМЕННЫЕ РЯДЫ

Задача 5. Составить аддитивную модель временного ряда.

Задание 5.1. Провести сглаживание данных методом скользящей средней (интервал сглаживания = 3).

Задание 5.2. По уравнению линии тренда $= 0,3044*t+47,979$ найти значения линии тренда и вписать в таблицу.

Задание 5.3. Определить значения оценки сезонной компоненты по формуле:

Оценка сезонной компоненты = Скользящая средняя – Тренд.

Задание 5.4. Рассчитать значения сезонной компоненты S (составить шаблон для S), пользуясь вспомогательной таблицей 3.

Задание 5.5. Оценить качество построенной модели.

Краткие результаты решения и выводы:

Таблица 8. Вспомогательная таблица для построения модели временного ряда

Месяц	Объём продаж	Скольльзящая средняя	Тренд	Сезон. компонента	S	T+S	E	E ²	(Y-Y _{ср}) ²
1	40,98								
2	37,086								
3	42,522								
4	48,99								
5	50,79								
6	57,882								
7	62,814								
8	64,506								
9	59,796								
10	49,182								
11	41,106								
12	42,18								
13	43,632								
14	40,668								
15	46,932								
16	50,244								
17	54,432								

Продолжение табл. 8

Месяц	Объём продаж	Скользящая средняя	Тренд	Сезон. компонента	S	T+S	E	E ²	(Y-Y _{ср}) ²
18	61,506								
19	66,198								
20	65,31								
21	62,016								
22	54,672								
23	46,128								
24	45,63								
25	50,118								
26	46,992								
27	52,992								
28	59,706								
29	63,846								
30	67,536								
31	68,562								
32	68,364								
33	64,008								
34	56,394								

Окончание табл. 8

Месяц	Объём продаж	Скользкая средняя	Тренд	Сезон. компонента	S	T+S	E	E ²	(Y-Y _{ср}) ²
35	46,668								
36	47,616								
37	49,134								
38	44,682								
39	50,922								
40	59,202								
41	61,53								
42	73,71								
43	69,84								
44	69,48								
45	74,52								
46	57,84								
47	48,582								
48	55,698								

Таблица 9. **Вспомогательная таблица для шаблона сезонной компоненты временного ряда**

	1 год	2 год	3 год	4 год	среднее	скор. средн.
1						
2						
3						
4						
5						
6						
7						
8						
9						
10						
11						
12						
сумма						

ТЕМА 6. СИСТЕМЫ ЭКОНОМЕТРИЧЕСКИХ УРАВНЕНИЙ

Задача 6. Провести идентификацию системы уравнений.

Пример:

Исходные данные:

$$y_1 = a_{11}x_1 + a_{12}x_2$$

$$y_2 = b_{21}y_1 + a_{22}x_2$$

Задание 6.1. Определить количество predetermined переменных в каждом уравнении.

Задание 6.2. Определить количество факторных переменных в системе.

Задание 6.3. Определить количество факторных переменных в каждом уравнении.

Задание 6.4. Определить уровень идентифицируемости каждого уравнения системы.

Задание 6.5. Определить уровень идентифицируемости всей системы уравнений.

Таблица 10. Системы структурных уравнений для идентификации

Вариант 1.	Вариант 2.	Вариант 3.	Вариант 4.	Вариант 5.
$y_1 = b_{11}y_3 + a_{11}x_1 + a_{13}x_3$	$y_1 = b_{12}y_2 + a_{11}x_1 + a_{13}x_3$	$y_1 = b_{12}y_2 + a_{12}y_2 + a_{13}x_3$	$y_1 = a_{11}x_1 + a_{12}x_2 + a_{13}x_3$	$y_1 = b_{12}y_2 + a_{11}x_1 + a_{13}x_3$
$y_2 = b_{21}y_1 + b_{23}y_3 + a_{22}x_2$	$y_2 = b_{23}y_3 + b_{23}y_3 + a_{22}x_2$	$y_2 = b_{23}y_3 + b_{23}y_3 + a_{22}x_2$	$y_2 = b_{21}y_1 + b_{23}y_3 + a_{22}x_2$	$y_2 = b_{23}y_3 + b_{23}y_3 + a_{22}x_2$
$y_3 = b_{32}y_2 + a_{31}x_1 + a_{33}x_3$	$y_3 = b_{31}y_1 + a_{31}x_1 + a_{33}x_3$	$y_3 = b_{31}y_1 + a_{31}x_1 + a_{33}x_3$	$y_3 = b_{32}y_2 + a_{31}x_1 + a_{33}x_3$	$y_3 = b_{31}y_1 + a_{31}x_1 + a_{33}x_3$
Вариант 6.	Вариант 7.	Вариант 8.	Вариант 9.	Вариант 10.
$y_1 = b_{12}y_2 + a_{11}x_1 + a_{13}x_3$	$y_1 = b_{12}y_2 + a_{12}x_2 + a_{13}x_3$	$y_1 = b_{12}y_2 + b_{13}y_3 + a_{13}x_3$	$y_1 = b_{12}y_2 + a_{11}x_1 + a_{13}x_3$	$y_1 = b_{12}y_2 + a_{12}x_2 + a_{13}x_3$
$y_2 = b_{21}y_1 + b_{23}y_3 + a_{22}x_2$	$y_2 = b_{23}y_3 + a_{22}x_2 + a_{23}x_3$	$y_2 = b_{23}y_3 + b_{23}y_3 + a_{22}x_2$	$y_2 = b_{23}y_3 + a_{22}x_2 + a_{22}x_2$	$y_2 = b_{23}y_3 + a_{22}x_2 + a_{23}x_3$
$y_3 = b_{32}y_2 + a_{31}x_1 + a_{33}x_3$	$y_3 = b_{31}y_1 + a_{32}x_2 + a_{33}x_3$	$y_3 = b_{31}y_1 + a_{31}x_1 + a_{33}x_3$	$y_3 = b_{31}y_1 + a_{31}x_1 + a_{33}x_3$	$y_3 = b_{31}y_1 + a_{32}x_2 + a_{33}x_3$

<p>Вариант 11.</p> $y_1 = a_{11}x_1 + a_{13}x_3$ $y_2 = b_{21}y_1 + b_{23}y_3 + a_{22}x_2$ $y_3 = b_{32}y_2 + a_{31}x_1 + a_{33}x_3$	<p>Вариант 12.</p> $y_1 = b_{12}y_2 + a_{11}x_1 + a_{13}x_3$ $y_2 = b_{23}y_3 + a_{22}x_2$ $y_3 = b_{31}y_1 + a_{31}x_1 + a_{33}x_3$	<p>Вариант 13.</p> $y_1 = b_{12}y_2 + a_{12}y_2 + a_{13}x_3$ $y_2 = b_{23}y_3 + b_{23}y_3 + a_{22}x_2$ $y_3 = a_{31}x_1 + a_{33}x_3$	<p>Вариант 14.</p> $y_1 = b_{12}y_2 + b_{13}y_3 + a_{13}x_3$ $y_2 = a_{11}x_1 + a_{22}x_2 + b_{23}y_3$ $y_3 = b_{31}y_1 + a_{31}x_1 + a_{33}x_3$	<p>Вариант 15.</p> $y_1 = b_{12}y_2 + a_{11}x_1 + a_{13}x_3$ $y_2 = b_{23}y_3 + a_{22}x_2$ $y_3 = b_{31}y_1 + a_{31}x_1 + a_{33}x_3$
<p>Вариант 16.</p> $y_1 = a_{11}x_1 + a_{13}x_3$ $y_2 = b_{21}y_1 + a_{22}x_2$ $y_3 = b_{32}y_2 + a_{31}x_1 + a_{33}x_3$	<p>Вариант 17.</p> $y_1 = b_{12}y_2 + a_{11}x_1 + a_{13}x_3$ $y_2 = b_{23}y_3 + a_{22}x_2$ $y_3 = a_{31}x_1 + a_{33}x_3$	<p>Вариант 18.</p> $y_1 = b_{12}y_2 + a_{12}y_2 + a_{13}x_3$ $y_2 = b_{23}y_3 + a_{22}x_2$ $y_3 = b_{31}y_1 + a_{31}x_1 + a_{33}x_3$	<p>Вариант 19.</p> $y_1 = b_{12}y_2 + a_{11}x_1$ $y_2 = b_{21}y_1 + a_{22}x_2$	<p>Вариант 20.</p> $y_1 = b_{12}y_2 + a_{11}x_1 + a_{13}x_3$ $y_2 = b_{23}y_3 + a_{22}x_2$ $y_3 = a_{31}x_1 + a_{33}x_3$

СПИСОК ЛИТЕРАТУРЫ

1. Айвазян С.А. Эконометрика. Краткий курс. – Москва: Маркет Дс, 2010. – 104 с.
2. Айвазян С.А., Фантаццини Д. Эконометрика-2. Продвинутый курс с приложениями в финансах: учебник. – Москва: Магистр, 2014, – 944 с.
3. Бородич С.А. Эконометрика: практикум. – Москва: ИНФРА-М, 2014. – 329 с.
4. Буравлёв А. Эконометрика: учебное пособие. – Москва: Бином, 2012. – 166 с.
5. Герасимов А.Н., Гладилин А.В. Эконометрика. Теория и практика. – Москва: Кнорус, 2011.
6. Гладилин А.В., Герасимов А.Н., Громов Е.И. Эконометрика. – Москва: Феникс, 2011. – 304 с.
7. Костромин А.В. Эконометрика.– Москва: КноРус, 2015. – 232 с.
8. Котенко А.П., Кузнецова О.А. Эконометрика. Множественная регрессия: методические указания к лабораторным работам по направлениям «Экономика», «Менеджмент», «Бизнес-информатика» / сост. А.П. Котенко, О.А. Кузнецова. – Самара, 2016. – 32 с.
9. Кузнецова О.А., Мазурмович О.Н. Эконометрика: практикум. – Самара: Издательство Самарского университета, 2019.
10. Новиков А.И. Эконометрика. – Москва: ИНФРА-М, 2014. – 272 с.

ПРИЛОЖЕНИЕ 1

Таблица значений F-критерия ФИШЕРА

k 2	α	k1(число степеней свободы)											
		1	2	3	4	5	6	7	8	9	10	11	12
1	0,1	39,9	49,5	53,6	55,8	57,2	58,2	58,9	59,4	59,9	60,2	60,5	60,7
	0,05	161	200	216	225	230	234	237	239	241	242	243	244
2	0,10	8,53	9,00	9,16	9,24	9,29	9,33	9,35	9,37	9,38	9,39	9,40	9,41
	0,05	18,5	19,0	19,2	19,2	19,3	19,3	19,4	19,4	19,4	19,4	19,4	19,4
	0,01	98,5	99,2	99,2	99,2	99,3	99,3	99,4	99,4	99,4	99,4	99,4	99,4
3	0,10	5,54	5,46	5,39	5,34	5,31	5,28	5,27	5,25	5,24	5,23	5,22	5,22
	0,05	10,01	9,55	9,28	9,12	9,01	8,94	8,89	8,85	8,81	8,79	8,76	8,74
	0,01	34,1	30,3	29,5	28,7	28,2	27,9	27,7	27,5	27,3	27,2	27,1	27,1
4	0,10	4,54	4,32	4,19	4,11	4,05	4,01	3,98	3,95	3,94	3,92	3,91	3,90
	0,05	7,71	6,94	6,59	6,39	6,26	6,16	6,09	6,04	6,00	5,96	5,94	5,91
	0,01	21,2	18,0	16,7	16,0	15,5	15,2	15,0	14,8	14,7	14,5	14,4	14,4
5	0,10	4,06	3,78	3,62	3,52	3,45	3,40	3,37	3,34	3,32	3,30	3,28	3,24
	0,05	6,61	5,79	5,41	5,19	5,05	4,95	4,88	4,82	4,77	4,74	4,71	4,68
	0,01	16,3	13,3	12,1	11,4	11,0	10,7	10,5	10,3	10,2	10,1	9,96	9,89
6	0,10	3,78	3,46	3,29	3,18	3,11	3,05	3,01	2,98	2,96	2,94	2,92	2,90
	0,05	5,99	5,14	4,76	4,53	4,39	4,28	4,21	4,15	4,10	4,06	4,03	4,00
	0,01	13,7	10,9	9,78	9,15	8,75	8,47	8,26	8,10	7,98	7,87	7,79	7,72
7	0,10	3,59	3,26	3,07	2,96	2,88	2,83	2,78	2,75	2,72	2,70	2,68	2,67
	0,05	5,59	4,74	4,35	4,12	3,97	3,87	3,79	3,73	3,68	3,64	3,60	3,57
	0,01	12,2	9,55	8,45	7,85	7,46	7,19	6,99	6,84	6,72	6,62	6,54	6,47
8	0,10	3,46	3,11	2,92	2,81	2,73	2,67	2,62	2,59	2,56	2,54	2,52	2,50
	0,05	5,32	4,46	4,07	3,84	3,69	3,58	3,50	3,44	3,39	3,35	3,31	3,28
	0,01	11,3	8,65	7,59	7,01	6,63	6,37	6,18	6,03	5,91	5,81	5,73	5,67
9	0,10	3,36	3,01	2,81	2,69	2,61	2,55	2,51	2,47	2,44	2,42	2,40	2,38
	0,05	5,12	4,26	3,86	3,63	3,48	3,37	3,29	3,23	3,18	3,14	3,10	3,07
	0,01	10,5	8,02	6,99	6,42	6,06	5,80	5,61	5,47	5,35	5,26	5,18	5,11

ПРИЛОЖЕНИЕ 2

Таблица критических значений t-критерия Стьюдента
 k – число степеней свободы, α – уровень значимости

k	α			k	α		
	0,100	0,050	0,010		0,100	0,050	0,010
1,000	6,314	12,706	63,657	21,000	1,721	2,080	2,831
2,000	2,920	4,303	9,925	22,000	1,717	2,074	2,819
3,000	2,353	3,182	5,841	23,000	1,714	2,069	2,807
4,000	2,132	2,776	4,604	24,000	1,711	2,064	2,797
5,000	2,015	2,571	4,032	25,000	1,708	2,060	2,787
6,000	1,943	2,447	3,707	26,000	1,706	2,056	2,779
7,000	1,895	2,365	3,499	27,000	1,703	2,052	2,771
8,000	1,860	2,306	3,355	28,000	1,701	2,048	2,763
9,000	1,833	2,262	3,250	29,000	1,699	2,045	2,756
10,000	1,812	2,228	3,169	30,000	1,697	2,042	2,750
11,000	1,796	2,201	3,106	31,000	1,696	2,040	2,744
12,000	1,782	2,179	3,055	32,000	1,694	2,037	2,738
13,000	1,771	2,160	3,012	33,000	1,692	2,035	2,733
14,000	1,761	2,145	2,977	34,000	1,691	2,032	2,728
15,000	1,753	2,131	2,947	35,000	1,690	2,030	2,724
16,000	1,746	2,120	2,921	36,000	1,688	2,028	2,719
17,000	1,740	2,110	2,898	37,000	1,687	2,026	2,715
18,000	1,734	2,101	2,878	38,000	1,686	2,024	2,712
19,000	1,729	2,093	2,861	39,000	1,685	2,023	2,708
20,000	1,725	2,086	2,845	40,000	1,684	2,021	2,704

Учебное издание

***Кузнецова Ольга Александровна,
Мазурмович Ольга Николаевна***

РАБОЧАЯ ТЕТРАДЬ ПО ДИСЦИПЛИНЕ ЭКОНОМЕТРИКА

Учебное пособие

Редакционно-издательская обработка И. П. В е д м и д с к о й

Подписано в печать 12.12.2022. Формат 60x84 1/16.

Бумага офсетная. Печ. л. 3,75.

Тираж 120 экз. (1-й з-д 1-25). Заказ . Арт. – 32(Р2УП)/2022.

ФЕДЕРАЛЬНОЕ ГОСУДАРСТВЕННОЕ АВТОНОМНОЕ
ОБРАЗОВАТЕЛЬНОЕ УЧРЕЖДЕНИЕ ВЫСШЕГО ОБРАЗОВАНИЯ
«САМАРСКИЙ НАЦИОНАЛЬНЫЙ ИССЛЕДОВАТЕЛЬСКИЙ
УНИВЕРСИТЕТ ИМЕНИ АКАДЕМИКА С.П. КОРОЛЕВА»
(САМАРСКИЙ УНИВЕРСИТЕТ)

443086, САМАРА, МОСКОВСКОЕ ШОССЕ, 34.

Издательство Самарского университета.
443086, Самара, Московское шоссе, 34.