

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

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 направлениям подготовки 38.03.01 Экономика, 38.03.02 Менеджмент, 38.03.05 Бизнес-информатика

САМАРА
Издательство Самарского университета
2019

УДК 330.115 (075)

ББК 65вбя7

К891

Рецензенты: д-р экон. наук, доц. Д.Ю. Иванов,
доц. А.П. Котенко

Кузнецова, Ольга Александровна

К891 Эконометрика: практикум / О.А. Кузнецова, О.Н. Мазурмович. –
Самара: Изд-во Самарского университета, 2019. – 72 с.

ISBN 978-5-7883-1413-6

Практикум содержит сведения методического и теоретического характера по эконометрике.

Методика решения задач иллюстрируется на конкретном примере. Приводится база исходных данных и необходимые таблицы справочного характера.

Предназначается для обучающихся по основной образовательной программе высшего образования по направлениям подготовки Экономика», «Менеджмент», «Бизнес-информатика».

Разработано на кафедре «Математических методы в экономике». Соответствует государственному образовательному стандарту.

УДК 330.115 (075)

ББК 65вбя7

ISBN 978-5-7883-1413-6

© Самарский университет, 2019

ОГЛАВЛЕНИЕ

ВВЕДЕНИЕ	5
I Парная регрессия	6
Основные понятия.....	6
Методы наименьших квадратов и максимального правдоподобия.....	8
Линейная и нелинейная регрессия.....	9
Коэффициент и индекс корреляции.....	10
Коэффициент и индекс детерминации.....	11
F-критерий Фишера.....	12
Расчёт F-критерия Фишера парной регрессии.....	12
Использование t-критерия Стьюдента.....	13
Задания по теме «Парная регрессия».....	15
Задачи по вариантам.....	18
Парная линейная регрессия.....	20
Свойства логарифмов.....	23
Оценка качества модели.....	25
Задания по вариантам.....	30
II Множественная регрессия	35
Общий и частный F-критерии Фишера.....	36
Индекс множественной корреляции.....	38
Задания по теме «Множественная регрессия».....	38
III Временные ряды	45
Стационарность и нестационарность временного ряда.....	46
Автоковариация и автокорреляция.....	47
Метод скользящего среднего.....	48
Задания по теме «Временные ряды».....	48
Системы эконометрических уравнений	52
Достаточное условие идентификации.....	54

СПИСОК ЛИТЕРАТУРЫ	61
ПРИЛОЖЕНИЕ 1 Таблица значений F – критерия ФИШЕРА....	63
ПРИЛОЖЕНИЕ 2. Таблица критических значений t-критерия Стьюдента	65
ПРИЛОЖЕНИЕ 3. Исходные данные к лабораторной работе по теме «множественная регрессия»	67

ВВЕДЕНИЕ

В настоящем практикуме в краткой форме представлены все необходимые теоретические и справочные материалы, необходимые для приобретения студентами опыта построения эконометрических моделей, принятия решений о спецификации и идентификации модели, выбора метода оценки параметров модели, интерпретации результатов, получения прогнозных оценок по курсу эконометрики (специальности: «Менеджмент», «Экономика», «Бизнес-информатика»).

Для подготовки к зачёту или экзамену (в зависимости от специальности) следует самостоятельно проработать учебную литературу (рекомендуемые источники приведены в приложении), руководствуясь списком вопросов для изучения, который также представлен в приложении.

Методически оправдано предварительный разбор практических задач с последующей более глубокой и расширенной проработкой теории, выходящей за рамки данного практикума. Такая организация изучения эконометрики позволяет сосредоточиться сначала на практической стороне предмета, а затем углубить понимание теоретической базы. При этом сведений представленных в данном практикуме вполне достаточно, чтобы самостоятельно справиться с решением задач.

І ПАРНАЯ РЕГРЕССИЯ

Основные понятия

Парная регрессия – это уравнение $y = \tilde{f}(x) + \varepsilon$ связи переменных y и x , где y – **зависимая переменная (результативный признак)**, x – **независимая объясняющая переменная (признак-фактор, регрессор)**, ε – **случайный член регрессии (остаток, невязка)** в **аддитивной** форме записи, учитывающий случайные (стохастические, недетерминированные) воздействия.

Возможна **мультипликативная форма** записи регрессии $y = \tilde{f}(x) \cdot \varepsilon$, а также другие математические способы учёта стохастических влияний.

Идентификация уравнения регрессии сводится к оценке её параметров. Для **линейных по неизвестным параметрам** регрессий применим метод наименьших квадратов (МНК), позволяющий найти оценки, минимизирующие сумму квадратов отклонений **фактических (наблюдаемых, экспериментальных)** значений результативного признака y от **теоретических (гипотетических, регрессионных, сглаженных)** значений

$$\tilde{y} : \sum (y_x - \tilde{y}_x)^2 \rightarrow \min .$$

Для линейных, а также нелинейных, **приводимых к линейным по идентифицируемым параметрам**, регрессионных моделей реша-

ется следующая, линейная по искомым точечным оценкам идентифицируемых параметров, система алгебраических уравнений в нормальной форме:

$$\begin{cases} an + b \sum x = \sum y, \\ a \sum x + b \sum x^2 = \sum xy. \end{cases}$$

Её решения имеют вид $a = \bar{y} - b\bar{x}$, $b = \frac{\text{cov}(x, y)}{\sigma_x^2} = \frac{\overline{xy} - \bar{x} \cdot \bar{y}}{\overline{x^2} - \bar{x}^2}$.

Интерпретация коэффициента регрессии b – изменение регрессора x на 1 единицу влечёт изменение исследуемого фактора y на b единиц. Т.о. коэффициент b – **экономически интерпретируемый** параметр (это иногда позволяет, к примеру, проверить его экспериментально).

Математическая интерпретация свободного члена регрессии a – значение y при нулевом значении регрессора $x=0$. Однако, в большинстве экономических задач регрессор не может принять нулевое или даже близкое к 0 значение. Поэтому свободный член линейной регрессии a относят к **неинтерпретируемым параметрам**, которые почти никогда не удаётся проверить экспериментально.

При интерпретации параметров линейной регрессии отметим 3 факта.

1. Параметры a и b являются лишь выборочными (точечными) оценками параметров линейной регрессии $\tilde{y} = \alpha + \beta x$ генеральной совокупности, поэтому возможно сколь угодно большое их отклонение от «истинных» значений.

2. Уравнение регрессии (не только линейное) отражает лишь **общую тенденцию** выборки, когда каждое отдельное исследование подвержено воздействию сколь угодно больших случайностей.

3. Верность интерпретации зависит от правильности *спецификации* уравнения. «Наилучшая» **МНК-прямая** всегда существует, но не всегда является достаточно хорошей.

**Методы наименьших квадратов
и максимального правдоподобия**

Пусть имеется n совместных наблюдений дискретных случайных величин X и Y . Представим их n точками (x_i, y_i) плоскости \mathfrak{R}^2 . Спроектируем каждую точку параллельно оси Oy на прямую искомой линейной регрессии $\hat{y} = ax + b$. Получим **регрессионные (теоретические, сглаженные, гипотетические)** значения $\hat{y}_i = ax_i + b$.

Подберём теперь параметры a и b искомой регрессии из условия

$$S(a, b) := \sum_{i=1}^n (y_i - \hat{y}_i)^2 = \sum_{i=1}^n (y_i - ax_i - b)^2 \rightarrow \min .$$

Для этого воспользуемся **необходимым условием экстремума** функции двух переменных:
$$\begin{cases} \partial S / \partial a = 0, \\ \partial S / \partial b = 0. \end{cases}$$

Получим
$$\begin{cases} \partial S / \partial a = 2 \sum_{i=1}^n (y_i - ax_i - b)(-x_i) = 0, \\ \partial S / \partial b = 2 \sum_{i=1}^n (y_i - ax_i - b)(-1) = 0. \end{cases}$$

Приведём эту систему 2-х линейных неоднородных уравнений с 2-мя неизвестными к **нормальному виду**, когда неизвестные находятся в левой, а известные слагаемые в правой части уравнений:

$$\begin{cases} a \sum_{i=1}^n x_i^2 + b \sum_{i=1}^n x_i = \sum_{i=1}^n x_i y_i, \\ a \sum_{i=1}^n x_i + b \cdot n = \sum_{i=1}^n y_i. \end{cases}$$

Решить её можно известными методами исключения неизвестных, обратной матрицы или по формулам Крамера. К примеру, по формулам Крамера решение примет вид $a = \Delta_a / \Delta$, $b = \Delta_b / \Delta$, где **определитель системы**

$$\Delta := \begin{vmatrix} \sum_{i=1}^n x_i^2 & \sum_{i=1}^n x_i \\ \sum_{i=1}^n x_i & n \end{vmatrix} = \begin{vmatrix} \sum_{i=1}^n x_i^2 & \sum_{i=1}^n x_i \\ \sum_{i=1}^n x_i & n \end{vmatrix} \neq 0,$$

$$\Delta_a := \begin{vmatrix} \sum_{i=1}^n x_i y_i & \sum_{i=1}^n x_i \\ \sum_{i=1}^n y_i & n \end{vmatrix}, \quad \Delta_b := \begin{vmatrix} \sum_{i=1}^n x_i^2 & \sum_{i=1}^n x_i y_i \\ \sum_{i=1}^n x_i & \sum_{i=1}^n y_i \end{vmatrix}.$$

Из первого уравнения СЛАУ (1) получим $a = \text{Cov}(X, Y) / \text{Var}(X)$. Тогда из второго уравнения $b = \bar{y} - a\bar{x}$. Последнее равенство означает, что МНК-прямая всегда проходит через **центр** (\bar{x}, \bar{y}) диаграммы рассеяния.

Линейная и нелинейная регрессия

Различают **линейные** и **нелинейные** регрессии.

Линейная регрессия: $y = a + bx + \varepsilon$.

Нелинейные регрессии делятся на регрессии, нелинейные относительно регрессоров, и регрессии, нелинейные по оцениваемым параметрам.

Регрессии, *нелинейные по регрессорам*:

полиномы $y = a + b_1 x + b_2 x^2 + \dots + b_k x^k + \varepsilon$;

(данная зависимость **линеаризуется** сведением к множественной линейной регрессии путём введения **новых регрессоров** $X_1 := x$, $X_2 := x^2$, ..., $X_k := x^k$);

равносторонняя гипербола $y = a + b/x + \varepsilon$;

(данная зависимость линеаризуется заменой $X := 1/x$);

и т.п.

Регрессии, *нелинейные по оцениваемым параметрам*:

показательная зависимость $y = ab^x \cdot \varepsilon$;

(данная зависимость линеаризуется заменами

$Y := \ln y$, $A := \ln a$, $B := \ln b$, $E := \ln \varepsilon$);

степенная зависимость $y = ax^b \cdot \varepsilon$;

(данная зависимость линеаризуется заменами

$Y := \ln y$, $X := \ln x$, $A := \ln a$, $E := \ln \varepsilon$);

экспоненциальная $y = e^{a+bx} \cdot \varepsilon$;

(данная зависимость линеаризуется заменами

$Y := \ln y$, $E := \ln \varepsilon$);

и т.п.

Регрессии, которые можно свести к линейным, называют **нелинейными внутренне линейными**. К ним применим МНК.

Регрессии, которые не сводятся к линейным, называют **нелинейными внутренне нелинейными**. К ним МНК не применим, т.к. система уравнений для идентификации **МНК-параметров** является нелинейной и требует специальных приёмов нахождения точного или приближённого решения.

Коэффициент и индекс корреляции

Коэффициент парной линейной корреляции

$$r_{xy} := b \frac{\sigma_x}{\sigma_y} = \frac{\text{cov}(x, y)}{\sigma_x \sigma_y} = \frac{\overline{xy} - \bar{x} \cdot \bar{y}}{\sqrt{\overline{x^2} - \bar{x}^2} \cdot \sqrt{\overline{y^2} - \bar{y}^2}}, \quad -1 \leq r_{xy} \leq 1,$$

оценивает тесноту *линейной* связи случайных величин, а **индекс корреляции**

$$\rho_{xy} := \sqrt{1 - \frac{\sigma_{\text{ост}}^2}{\sigma_y^2}} = \sqrt{1 - \frac{\sum (y_x - \tilde{y}_x)^2}{\sum (y_x - \bar{y})^2}}, \quad 0 \leq \rho_{xy} \leq 1,$$

оценивает тесноту *нелинейной* связи.

Коэффициент и индекс детерминации

Оценку **адекватности** построенной модели экспериментальным наблюдениям даёт коэффициент (индекс) детерминации, а также средняя ошибка аппроксимации.

Долю дисперсии, объясняемую регрессией, в общей дисперсии результативного признака у характеризует **коэффициент (индекс) детерминации**

$$R^2 := \frac{\text{Var}(\tilde{y})}{\text{Var}(y)} = \frac{\sum (\tilde{y}_x - \bar{y})^2}{\sum (y_x - \bar{y})^2},$$

или

$$R^2 := 1 - \frac{\text{Var}(\varepsilon)}{\text{Var}(y)} = 1 - \frac{\sum (y_x - \tilde{y}_x)^2}{\sum (y_x - \bar{y})^2}.$$

Коэффициент (индекс) *детерминации* равен квадрату коэффициента (индекса) *линейной парной корреляции*.

Средняя ошибка аппроксимации

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

оценивает **среднее отклонение** расчётных значений от фактических. Допустимый предел значений \bar{A} – не более 8-10%.

F-критерий Фишера

Расчёт F-критерия Фишера парной регрессии

Оценка качества уравнения регрессии состоит в проверке *нулевой гипотезы* H_0 о **статистической незначимости** уравнения регрессии и показателя тесноты связи. Для этого *фактическое* значение статистики $F_{\text{факт}}$ сравнивается с *критическим (табличным)* значением $F_{\text{табл}}$ критерия Фишера.

Фактическое значение статистики Фишера равно отношению удельных (рассчитанных на одну степень свободы) факторной и остаточной дисперсий:

$$F_{\text{факт}} := \frac{\sum (\tilde{y}_x - \bar{y})^2 / m}{\sum (y_x - \tilde{y}_x)^2 / (n - m - 1)} = \frac{r_{xy}^2}{1 - r_{xy}^2} (n - 2),$$

где n – объём выборки, m – число параметров при регрессорах.

Табличное значение $F_{\text{табл}}$ – это максимально возможное значение F -статистики при действии случайных факторов при заданных уровне значимости α и числе степеней свободы.

F -статистика Фишера имеет 2 числа **степеней свободы**:

– **верхнее число** = числу объясняющих переменных (в парной регрессии =1);

– **нижнее число** = числу наблюдений в выборке минус число оцениваемых параметров (в парной регрессии = $n-2$).

$$F_{\text{факт}} := \frac{ESS/p}{RSS/q},$$

где p – верхнее число степеней свободы, q – нижнее число степеней свободы.

$$\text{Для парной регрессии } F_{\text{факт}} := \frac{ESS}{RSS} \cdot (n - 2).$$

Уровень значимости α – вероятность *отвергнуть верную* гипотезу. Стандартное значение $\alpha=0,05$; часто табулируются значения 0,01 и 0,1.

При $F_{\text{табл}} < F_{\text{факт}}$ нулевая гипотеза H_0 о случайном характере оцениваемых характеристик отклоняется и принимается их **статистическая значимость и надёжность**. В противном случае $F_{\text{табл}} > F_{\text{факт}}$ гипотеза H_0 принимается, т.е. признаётся **статистическая незначимость и ненадёжность** найденного уравнения регрессии.

Если $1 \leq F < +\infty$, следует рассмотреть F^{-1} .

Использование t -критерия Стьюдента

Для оценки **статистической значимости** коэффициентов регрессии и корреляции рассчитывается t -критерий Стьюдента и доверительные интервалы каждого показателя. *Нулевая гипотеза H_0* говорит о **случайной природе показателей**, т.е. незначимом их отклонении от 0. С табличным значением сравнивается отношение значения оценки и её случайной ошибки

$$t_a := a/m_a, \quad t_b := b/m_b, \quad t_r := r/m_r,$$

где **случайные ошибки** точечных оценок параметров определяются формулами

$$m_a := \sqrt{\frac{\sum (y_x - \tilde{y}_x)^2}{n-2} \cdot \frac{\sum x^2}{n \sum (x - \bar{x})^2}} = \sqrt{S_{\text{ост}}^2 \frac{\sum x^2}{n^2 \sigma_x^2}} = \frac{S_{\text{ост}}}{n \sigma_x} \sqrt{\sum x^2},$$

$$m_b := \sqrt{\frac{\sum (y_x - \tilde{y}_x)^2}{(n-2)\sum (x - \bar{x})^2}} = \sqrt{\frac{S_{\text{ост}}^2}{\sum (x - \bar{x})^2}} = \frac{S_{\text{ост}}}{\sigma_x \sqrt{n}},$$

$$m_{r_{xy}} := \sqrt{\frac{1 - r_{xy}^2}{n - 2}}.$$

Сравнивая *фактическое* $t_{\text{факт}}$ и *критическое* (табличное) $t_{\text{табл}}$ значения, принимаем или отвергаем гипотезу H_0 :

- при $t_{\text{табл}} < t_{\text{факт}}$ – отвергаем (т.е. a , b и r_{xy} под действием систематически действующего фактора x отличаются от 0 **не случайно**);
- при $t_{\text{табл}} > t_{\text{факт}}$ – принимаем (т.е. a , b и r_{xy} формируются **случайно**).

Статистики Фишера и Стьюдента связаны соотношением $t_b^2 = t_r^2 = \sqrt{F}$.

Предельные ошибки для расчёта доверительных интервалов найдём по формулам $\Delta_a := t_{\text{табл}} m_a$, $\Delta_b := t_{\text{табл}} m_b$.

Для проверки **гипотезы о линейной зависимости** между СВ X и Y справедлива следующая процедура:

1. Вычислить t -статистику $t := r \cdot \sqrt{\frac{n-2}{1-r^2}}$.

2. По заданному уровню значимости α найти критическое значение t -статистики с $(n-2)$ степенями свободы.

3. Если величина t превосходит критическое значение в положительную или отрицательную сторону, то нулевая гипотеза $H_0: r=0$ отклоняется, т.е. принимается *альтернативная* гипотеза H_1 о **наличии линейной положительной или отрицательной зависимости** соответственно.

Замечание: тест справедлив лишь для проверки нулевой гипотезы об отсутствии линейной зависимости; для проверки какой-либо другой гипотезы (например, о равенстве коэффициента корреляции заданному ненулевому числу) требуется иная, более сложная процедура.

Для *парного* регрессионного анализа t -критерий для нулевой гипотезы $H_0: r_{xy}=0$, F -критерий для гипотезы $H_0: R^2=0$ и t -критерий для $H_0: \beta=0$ эквивалентны.

Задания по теме «Парная регрессия»

Экономический смысл параметров a, b, α, β

Задание 1.1 Методика решения задач для функций вида $y = a + b \cdot x + \varepsilon$. Необходимо определить прогнозное значение y' .

x – текущее значение x , x' – прогнозное значение x , y – текущее значение y ,

y' – прогнозное значение y

Пример:

Исходные данные: $y = a + b \cdot x + \varepsilon$, $x = 2$, $y = 100$, $b = 3$, $x' = 4$.

Найти y' .

Решение:

Параметр b в парной линейной регрессии показывает, на сколько единиц изменится результат при изменении фактора на одну единицу.

Необходимые формулы

$\Delta x = x - x$ – на сколько единиц изменился фактор

$\Delta y = \Delta x \cdot b$ – на сколько единиц изменился результат

$y' = y + \Delta y$ новое значение результата.

Подстановка данных в формулы:

$\Delta x = 3 - 2 = 1$

$\Delta y = 1 \cdot 3 = 3$

$$y' = 100 + 3 = 103$$

Ответ: $y' = 103$.

Задачи по вариантам

1.1. $x = 2, y = 100, b = 3, x' = 1, y' = ?$

1.2. $x = 5, y = 50, b = -2, x' = 4, y' = ?$

1.3. $x = 5, y = 50, b = -2, x' = 6, y' = ?$

1.4. $x = 5, y = 70, b = 2, x' = 7, y' = ?$

1.5. $x = 1, y = 70, b = 2, x' = 1, y' = ?$

1.6. $x = 5, y = 60, b = -2, x' = 7, y' = ?$

1.7. $x = 5, y = 60, b = -2, x' = 2, y' = ?$

1.8. $x = 2, y = 100, b = 3, x' = 1, y' = ?$

1.9. $x = 5, y = 50, b = -2, x' = 4, y' = ?$

1.10. $x = 5, y = 50, b = -2, x' = 6, y' = ?$

1.11. $x = 5, y = 70, b = 2, x' = 7, y' = ?$

1.12. $x = 1, y = 70, b = 2, x' = 1, y' = ?$

1.13. $x = 5, y = 60, b = -2, x' = 7, y' = ?$

1.14. $x = 5, y = 60, b = -2, x' = 2, y' = ?$

1.15. $x = 2, y = 100, b = 3, x' = 1, y' = ?$

1.16. $x = 5, y = 50, b = -2, x' = 4, y' = ?$

1.17. $x = 5, y = 50, b = -2, x' = 6, y' = ?$

1.18. $x = 5, y = 70, b = 2, x' = 7, y' = ?$

1.19. $x = 1, y = 70, b = 2, x' = 1, y' = ?$

1.20. $x = 5, y = 60, b = -2, x' = 7, y' = ?$

1.21. $x = 5, y = 60, b = -2, x' = 2, y' = ?$

Задание 1.2. Методика решения задач для функции вида $y = a * x^b * \varepsilon$. Необходимо определить прогнозное значение y' .

x – текущее значение x , x' – прогнозное значение x , y – текущее значение y ,

y' – прогнозное значение y

Пример:

Исходные данные: $y = a * x^b * \varepsilon$, $x = 2$, $y = 100$, $b = 3$,
 $x' = 3$, $y' = ?$

Решение:

Параметр b в парной степенной регрессии показывает, на сколько процентов изменится результат при изменении фактора на один процент.

Необходимые формулы

$\Delta x = \frac{x' - x}{x} 100\%$ – на сколько % изменился фактор

$\Delta y = \Delta x * b(\%)$ – на сколько % изменился результат

$y' = (1 + \frac{\Delta y\%}{100\%})$ – новое значение результата.

Подстановка данных в формулы:

$$\Delta x = \frac{3 - 2}{2} 100\% = 50\%$$

$$\Delta y = 50\% * 3 = 150\%$$

$$y = 100 \left(1 + \frac{150\%}{100\%} \right) = 250 \text{шт}$$

Ответ: $y' = 250$.

Задачи по вариантам

- 2.1 $x = 2, y = 100, b = 3, x' = 1, y' = ?$
2.2 $x = 2, y = 50, b = 3, x' = 3, y' = ?$
2.3 $x = 2, y = 50, b = 3, x' = 1, y' = ?$
2.4 $x = 2, y = 100, b = -3, x' = 3, y' = ?$
2.5 $x = 2, y = 100, b = -3, x' = 1, y' = ?$
2.6 $x = 5, y = 50, b = -2, x' = 7, y' = ?$
2.7 $x = 5, y = 50, b = -2, x' = 3, y' = ?$
2.8 $x = 100, y = 20, b = 4, x' = 101, y' = ?$
2.9 $x = 100, y = 20, b = 4, x' = 99, y' = ?$
2.10 $x = 100, y = 20, b = -4, x' = 102, y' = ?$
2.11 $x = 2, y = 100, b = 3, x' = 1, y' = ?$
2.12 $x = 2, y = 50, b = 3, x' = 3, y' = ?$
2.13 $x = 2, y = 50, b = 3, x' = 1, y' = ?$
2.14 $x = 2, y = 100, b = -3, x' = 3, y' = ?$
2.15 $x = 2, y = 100, b = -3, x' = 1, y' = ?$
2.16 $x = 5, y = 50, b = -2, x' = 7, y' = ?$
2.17 $x = 5, y = 50, b = -2, x' = 3, y' = ?$
2.18 $x = 100, y = 20, b = 4, x' = 101, y' = ?$
2.19 $x = 100, y = 20, b = 4, x' = 99, y' = ?$
2.20 $x = 100, y = 20, b = -4, x' = 102, y' = ?$

Задание 1.3. Определить параметры линейной регрессии:

- а) с помощью системы нормальных уравнений;
б) по расчетной формуле;

в) в матричной форме.

г) записать уравнение регрессии с найденными параметрами.

Исходные данные к работе: В табл.1 приведена статистика распределения расходов на потребление продуктов питания и средней заработной платы по годам.

Таблица 1 – Статистика распределения расходов на потребление продуктов питания и средней заработной платы

№	y	x
1	33,66	120,00
2	28,56	112,30
3	20,40	107,25
4	27,54	107,25
5	48,96	127,25
6	21,42	112,20
7	22,44	113,85
8	43,86	122,10
9	47,94	122,10
10	72,42	132,10
11	79,56	127,05
12	55,08	127,05
13	72,42	128,70
14	77,52	132,00
15	81,60	133,65
16	56,10	128,60
17	77,52	140,25
18	51,00	130,25
19	66,30	130,25
20	64,26	131,90
21	66,30	136,90

Данные для индивидуальных заданий рассчитываются по формуле $y = y + 2 \cdot N$, где « N » обозначен номер варианта работы, соответствующий номеру студента в списке группы.

Парная линейная регрессия

$$y = a + b \cdot x + \varepsilon.$$

Задание: определить коэффициенты парной линейной регрессии методами определителей и наименьших квадратов, оценить качество полученной модели.

1. Составить систему нормальных уравнений и найти параметры регрессии методом определителей

Согласно методу наименьших квадратов

$$\sum (y - y_{расч})^2 \rightarrow \min.$$

Система нормальных уравнений для определения параметров a и b линейной регрессии выглядит следующим образом:

$$\begin{cases} na + b \sum x = \sum y \\ a \sum x + b \sum x^2 = \sum xy \end{cases}$$

где n – количество наблюдений.

Для подстановки числовых параметров в систему уравнений необходимо заполнить вспомогательную таблицу (табл. 2).

Из системы получаем матрицу

$$\begin{bmatrix} n & \sum x & \sum y \\ \sum x & \sum x^2 & \sum xy \end{bmatrix}$$

Таблица 2 – Вспомогательная таблица для подстановки числовых параметров в систему уравнений

№	y	x	x*y	x ²	y ²	\hat{y}	$(y - \hat{y})^2$	$(x - \bar{x})^2$
1	33,66	120,00	4039,20	14400,00	1133,00	43,39	94,74	24,06
2	28,56	112,30	3207,29	12611,29	815,67	28,17	0,15	158,88
3	20,40	107,25	2187,90	11502,56	416,16	18,19	4,88	311,69
4	27,54	107,25	2953,67	11502,56	758,45	18,19	87,42	311,69
5	48,96	127,25	6230,16	16192,56	2397,08	27,72	76,82	5,50
6	21,42	112,20	2403,32	12588,84	458,82	27,97	42,96	161,41
7	22,44	113,85	2554,79	12961,82	503,55	31,24	77,38	122,21
8	43,86	122,10	5355,31	14908,41	1923,70	47,54	13,57	7,87
9	47,94	122,10	5853,47	14908,41	2298,24	47,54	0,16	7,87
10	72,42	132,10	9566,68	17450,41	5244,66	67,31	26,10	51,77
11	79,56	127,05	10108,10	16141,70	6329,79	57,33	494,21	4,60
12	55,08	127,05	6997,91	16141,70	3033,81	57,33	5,06	4,60
13	72,42	128,70	9320,45	16563,69	5244,66	60,59	139,93	14,40
14	77,52	132,00	10232,64	17424,00	6009,35	67,11	108,29	50,34
15	81,60	133,65	10905,84	17862,32	6658,56	70,38	125,99	76,48
16	56,10	128,60	7214,46	16537,96	3147,21	60,39	18,43	13,65
17	77,52	140,25	10872,18	19670,06	6009,35	83,42	34,83	235,48
18	51,00	130,25	6642,75	16965,06	2601,00	63,65	160,14	28,57
19	66,30	130,25	8635,58	16965,06	4395,69	63,65	7,00	28,57
20	64,26	131,90	8475,89	17397,61	4129,35	66,92	7,06	48,93
21	66,30	136,90	9076,47	18741,61	4395,69	76,80	110,25	143,89
Σ	1114,86	2623,00	142834,07	329437,66	67903,79	1114,86	1635,36	1812,46
Среднее	53,09	124,90	6801,62	15687,51	3233,51	53,09	77,87	86,31

И считаем определители

$$\Delta = n \cdot \sum x^2 - \sum x \cdot \sum x$$

$$\Delta a = \sum y \cdot \sum x^2 - \sum x \cdot \sum xy$$

$$\Delta b = n \cdot \sum xy - \sum y \cdot \sum x$$

Δ – главный определитель матрицы

Δa – определитель матрицы a

Δb – определитель матрицы b

$$a = \frac{\Delta a}{\Delta}$$

$$b = \frac{\Delta b}{\Delta}$$

2. Найти параметры регрессии методом наименьших квадратов

$$b = \frac{\bar{x} \cdot \bar{y} - \bar{x} \cdot \bar{y}}{x^2 - \bar{x}^2}$$

$$a = \bar{y} - b \cdot \bar{x}$$

3. Записать уравнение регрессии с найденными параметрам.

Задание 1.4. Привести нелинейную регрессию к линейному виду.

Применять метод наименьших квадратов для нахождения коэффициентов регрессии можно только для линейной модели. В случае, когда мы имеем дело с нелинейными моделями типа $y = a * b^x * \varepsilon$,

$y = a + \frac{b}{x} + \varepsilon$, необходимо привести их к линейному виду. Для этого можно использовать способы логарифмирования и замены. Логарифмирование предполагает знание свойств логарифмов.

Свойства логарифмов

1. $a^{\log_a b} = b$ – основное логарифмическое тождество.

2. $\log_a a = 1, a > 0, a \neq 1$

3. $\log_a 1 = 0, a > 0, a \neq 1$

Логарифм единицы по любому положительному, отличному от 1, основанию равен нулю. Это возможно потому, что из любого действительного числа можно получить 1 только возведя его в нулевую степень.

4. $\log_a (bc) = \log_a b + \log_a c$ -логарифм произведения. Логарифм произведения равен сумме логарифмов сомножителей.

5. $\log_a \frac{b}{c} = \log_a b - \log_a c$ – логарифм частного. Логарифм частного (дроби) равен разности логарифмов сомножителей.

6. $\log_a b^p = p * \log_a b$ – логарифм степени. Логарифм степени равен произведению показателя степени на логарифм ее основания.

$$7. \log_{a^k} b = \frac{1}{k} * \log_a b$$

$$8. \log_a b = \frac{1}{\log_b a}$$

$$9. \log_a b = \frac{\log_c b}{\log_c a} \text{ – переход к новому основанию.}$$

Пример:

Исходные данные: $y = a * x^b * \varepsilon$

Решение:

$$\ln y = \ln(a * x^b * \varepsilon)$$

$$\ln y = \ln a + \ln x^b + \ln \varepsilon$$

$$\ln y = \ln a + b * \ln x + \ln \varepsilon$$

Ответ: $\ln y = \ln a + b * \ln x + \ln \varepsilon$

Задания по вариантам

1. $y = a + \frac{b}{x} + \varepsilon$

2. $y = a + bx + cx + \varepsilon$

3. $y = a + bx + cx^2 + dx^2 + \varepsilon$

4. $y = a * x^b * \varepsilon$

5. $y = a * b^x * \varepsilon$

6. $y = e^{a+bx} * \varepsilon$

7. $y = a + \frac{b}{x} + \varepsilon$

8. $y = a + bx + cx + \varepsilon$

9. $y = a + bx + cx^2 + dx^2 + \varepsilon$

10. $y = a * x^b * \varepsilon$

11. $y = a * b^x * \varepsilon$

12. $y = e^{a+bx} * \varepsilon$

13. $y = a + \frac{b}{x} + \varepsilon$

14. $y = a + bx + cx + \varepsilon$

15. $y = a + bx + cx^2 + dx^2 + \varepsilon$

16. $y = a * x^b * \varepsilon$

17. $y = a * b^x * \varepsilon$

$$18. y = e^{a+bx} * \varepsilon$$

$$19. y = a + \frac{b}{x} + \varepsilon$$

$$20. y = a + bx + cx + \varepsilon$$

Оценка качества модели

Качество построенной модели определяется с помощью показателей корреляции, детерминации, критериев Фишера и Стьюдента.

Задание 1.5. Методика решения задач для функции вида $y = a + b * x + \varepsilon$ определить **линейный коэффициент корреляции**.

Пример:

Исходные данные:

$$b = 36,84$$

$$\sigma_x = 1,25$$

$$\sigma_y = 46,29$$

Найти: **линейный коэффициент корреляции**.

Решение:

При использовании линейной регрессии используется линейный коэффициент корреляции r_{xy} . Линейный коэффициент корреляции находится в определенных пределах

$$(-1) \leq r_{xy} \leq (+1).$$

При этом чем ближе r_{xy} к нулю, тем слабее корреляция, чем ближе r_{xy} к (-1) или к (+1), тем сильнее корреляция, т.е. зависимость x и y близка к линейной

$$r_{xy} = b \frac{\sigma_x}{\sigma_y}$$

где σ_x – среднее квадратическое отклонение фактора,
 σ_y – среднее квадратическое отклонение результата.

$$\sigma_x^2 = \overline{x^2} - \bar{x}^2$$

$$\sigma_y^2 = \overline{y^2} - \bar{y}^2,$$

$$r_{xy} = 36,84 * \frac{1,25}{46,29} = b \frac{\sigma_x}{\sigma_y} = 0,75$$

Ответ: линейный коэффициент корреляции равен 0,75. Он показывает сильную прямую взаимосвязь между фактором и результатом.

Задание 1.6. Методика решения задач для функции вида $y = a + b * x + \varepsilon$ определить коэффициент детерминации.

Пример:

Исходные данные: значение коэффициента корреляции взять из предыдущего задания.

Решение: линейный коэффициент корреляции равен 0,75.

Коэффициент (индекс) детерминации характеризует долю дисперсии, объясняемую регрессией, в общей дисперсии результативного признака y .

$$r^2 = 0,75 * 0,75 = 0,5625$$

Ответ: коэффициент детерминации равен 0,5625. Это означает, что изменение результата на 56% зависит от изменения x .

Задания по вариантам

Значение коэффициента корреляции взять из задачи 5 своего варианта.

Задание 1.7. Методика решения задач для функции вида $y = a + b * x + \varepsilon$ определить значение критерия Фишера, при $\alpha = 0,05$

Пример:

Исходные данные: $TSS = 15000$, $ESS = 14735$, $RSS = 265$,
 $n = 7$.

Решение:

$$F = \frac{\sum (\hat{y} - \bar{y})^2}{\sum (y - \hat{y})^2} * (n - 2) \Rightarrow \frac{ESS}{RSS} (n - 2)$$

$$ESS = \sum (\hat{y} - \bar{y})^2$$

$$RSS = \sum (y - \hat{y})^2$$

$$TSS = ESS - RSS$$

$$F_{расч} > F_{табл}$$

По формуле находим расчётное значение критерия ФИШЕРА

$$F_{расч} = \frac{14735}{265} * (7 - 2) = 278$$

Учитывая число степеней свободы $k_1 = m - 1$ $k_2 = n - m$

$$k_1 = 2 - 1 = 1$$

$$k_2 = 7 - 2 = 5$$

По таблице значений F – критерия ФИШЕРА на пересечении столбца 1 и строки 5 находим значение критерия ФИШЕРА.

$$F_{T(0.05)} = 6.61$$

Сравниваем расчётное и табличное значение критерия ФИШЕРА.
 $278 > 6.61$

Вывод: так как расчётное значение критерия ФИШЕРА больше табличного, то уравнение регрессии значимо.

Ответ: $F_{расч} = 278$ $F_{T(0.05)} = 6.61$ так как расчётное значение критерия ФИШЕРА больше табличного, то уравнение регрессии значимо.

Задания по вариантам

1. TSS =100, ESS = 99, RSS = 1, n= 15, m =2.
2. TSS =100, ESS = 60, RSS = 0, n= 15, m =2.
3. TSS =100, ESS = 70, RSS = 30, n= 15, m =2.
4. TSS =100, ESS = 80, RSS = 20, n= 15, m =2.
5. TSS =100, ESS = 85, RSS = 15, n= 15, m =2.
6. TSS =100, ESS = 99, RSS = 1, n= 21, m =2.
7. TSS =100, ESS = 60, RSS = 0, n= 21, m =2.
8. TSS =100, ESS = 70, RSS = 30, n= 21, m =2.
9. TSS =100, ESS = 80, RSS = 20, n= 21, m =2.
10. TSS =100, ESS = 85, RSS = 15, n= 21, m =2.
11. TSS =100, ESS = 99, RSS = 1, n= 23, m =2.
12. TSS =100, ESS = 60, RSS = 0, n= 23, m =2.
13. TSS =100, ESS = 70, RSS = 30, n= 23, m =2.
14. TSS =100, ESS = 80, RSS = 20, n= 23, m =2.
15. TSS =100, ESS = 85, RSS = 15, n= 23, m =2.
16. TSS =100, ESS = 99, RSS = 1, n= 21, m = 3.
17. TSS =100, ESS = 60, RSS = 0, n= 21, m = 3.
18. TSS =100, ESS = 70, RSS = 30, n= 21, m = 3.
19. TSS =100, ESS = 80, RSS = 20, n= 21, m = 3.
20. TSS =100, ESS = 85, RSS = 15, n= 21, m = 3.

Задание 1.8. Для функции вида $y = a + b * x + \varepsilon$ определить значение критерия Стьюдента, при $\alpha = 0,05$

а) Найти стандартизированную ошибки коэффициентов регрессии;

б) найти значение критерия Стьюдента для всех переменных;

в) сделать вывод о значимости параметров регрессии.

Пример:

Исходные данные: $a = -5,79$, $b = 36,84$, $r = 0,75$, $RSS = 265$,
 $n = 7$, $\sum (x - \bar{x})^2 = 10,857$, $\sum x^2 = 80$.

Решение:

а) стандартизированные ошибки коэффициентов регрессии а и b.

$$m_a = \sqrt{\frac{RSS}{n-2} * \frac{\sum x^2}{n * \sum (x - \bar{x})^2}} = \sqrt{\frac{265}{7-2} * \frac{80}{7 * 10,857}} =$$

$$m_a = 7,45$$

$$m_b = \sqrt{\frac{RSS / (n-2)}{\sum (x - \bar{x})^2}} = \sqrt{\frac{265 / (7-2)}{10,857}} = 2,21$$

б) значение критерия Стьюдента считается для каждого коэффициента регрессии.

$$t_a = \frac{a}{m_a}$$

$t_a = -5,79/7,45 = -0,78$ значение критерия берём по модулю, следо-

вательно, $t_a = 0,78$ $t_b = \frac{b}{m_b} = \frac{36,84}{2,21} = 16,67$

в) сделать вывод о значимости параметров регрессии.

Учитывая число степеней свободы $k=n-m$ и $\alpha = 0,05$

$$k=7-2=5$$

По таблице значений критерия Стьюдента на пересечении столбца $\alpha = 0,05$ и строки 5 находим значение критерия Стьюдента.

$$t_{\text{табл}} = 2,57$$

Сравниваем расчётное и табличное значение критерия Стьюдента.

$0,78 < 2,57$ так как расчётное значение критерия Стьюдента меньше табличного, то принимаем теорию H_0 о том, что параметр a незначим.

$16,65 > 2,57$ так как расчётное значение критерия Стьюдента больше табличного, то отвергаем теорию H_0 о том, что параметр b значим.

Ответ: $m_a = 7,45, m_b = 2,21, t_a = 0,78, t_b = 16,65$, так как расчётное значение критерия Стьюдента меньше табличного, то принимаем теорию H_0 о том, что параметр a незначим. так как расчётное значение критерия Стьюдента больше табличного, то отвергаем теорию H_0 о том, что параметр b значим.

Задания по вариантам

1. $a = 20, b = 5, r = 0,8, RSS = 100, n = 5, \Sigma(x-x)^2 = 3, \Sigma x^2 = 45.$
2. $a = 20, b = 5, r = 0,5, RSS = 100, n = 5, \Sigma(x-x)^2 = 3, \Sigma x^2 = 45.$
3. $a = 20, b = 5, r = 0,7, RSS = 100, n = 5, \Sigma(x-x)^2 = 3, \Sigma x^2 = 45.$
4. $a = 20, b = 5, r = 0,3, RSS = 100, n = 5, \Sigma(x-x)^2 = 3, \Sigma x^2 = 45.$
5. $a = 20, b = 15, r = 0,8, RSS = 100, n = 5, \Sigma(x-x)^2 = 3, \Sigma x^2 = 45.$
6. $a = 20, b = 15, r = 0,5, RSS = 100, n = 5, \Sigma(x-x)^2 = 5, \Sigma x^2 = 49.$

7. $a = 20, b = 15, r = 0,8, RSS = 20, n = 5, \Sigma(x-x)^2 = 5, \Sigma x^2 = 49.$
8. $a = 20, b = 15, r = 0,6, RSS = 100, n = 5, \Sigma(x-x)^2 = 5, \Sigma x^2 = 49.$
9. $a = 20, b = 15, r = 0,9, RSS = 100, n = 5, \Sigma(x-x)^2 = 5, \Sigma x^2 = 49.$
10. $a = 200, b = 15, r = 0,8, RSS = 50, n = 5, \Sigma(x-x)^2 = 5, \Sigma x^2 = 16.$
11. $a = 200, b = 5, r = 0,8, RSS = 50, n = 5, \Sigma(x-x)^2 = 5, \Sigma x^2 = 16.$
12. $a = 100, b = 15, r = 0,6, RSS = 100, n = 5, \Sigma(x-x)^2 = 5, \Sigma x^2 = 16.$
13. $a = 200, b = 15, r = 0,8, RSS = 70, n = 5, \Sigma(x-x)^2 = 5, \Sigma x^2 = 16.$
14. $a = 200, b = 1, r = 0,8, RSS = 100, n = 5, \Sigma(x-x)^2 = 5, \Sigma x^2 = 16.$
15. $a = 200, b = 15, r = 0,8, RSS = 100, n = 5, \Sigma(x-x)^2 = 5, \Sigma x^2 = 49.$
16. $a = 100, b = 5, r = 0,9, RSS = 100, n = 5, \Sigma(x-x)^2 = 3, \Sigma x^2 = 45.$
17. $a = 200, b = 50, r = 0,95, RSS = 100, n = 5, \Sigma(x-x)^2 = 3, \Sigma x^2 = 45.$
18. $a = 20, b = 50, r = 0,7, RSS = 100, n = 5, \Sigma(x-x)^2 = 3, \Sigma x^2 = 45.$
19. $a = 20, b = 5, r = 0,3, RSS = 100, n = 5, \Sigma(x-x)^2 = 3, \Sigma x^2 = 16.$
20. $a = 20, b = 15, r = 0,7, RSS = 100, n = 5, \Sigma(x-x)^2 = 5, \Sigma x^2 = 45.$

Задание 1.9. Методика решения задач для функции вида $y = a + b * x + \varepsilon$ рассчитать доверительные интервалы для параметров.

Пример:

Исходные данные: данные для выполнения данного задания взять из предыдущего задания, для этого примера $a = -5,79; b = 36,84, m_a = 7,45, m_b = 2,21, t_a = 0,78, t_b = 16,65.$

Решение:

$$\underline{a - t_a * m_a < a < a + t_a * m_a}$$

$$\underline{b - t_b * m_b < b < b + t_b * m_b}$$

$$-5,79 - 0,78 * 7,45 < a < -5,79 + 0,78 * 7,45$$

$$36,84 - 16,65 * 2,21 < b < 36,84 + 16,65 * 2,21$$

Ответ: доверительный интервал параметра а (-11,6; 0,021), доверительный интервал параметра b (0,04; 73,64)

Задания по вариантам

данные взять из предыдущего задания своего варианта

Задание 1.10. Методика решения задач для функции вида $y = a + b * x + \varepsilon$ определить интервальный прогноз на основе линейного уравнения регрессии

Пример:

Исходные данные: $\bar{X} = 3,3143$, $\sum (X_i - \bar{X})^2 = 10,857$, $n = 7$

$$RSS = 265$$

Решение:

$$\hat{y}_x - m_{\hat{y}_x} \leq y^U \leq \hat{y}_x + m_{\hat{y}_x}$$

$$m_{\hat{y}_x} = \sqrt{\frac{RSS}{(n-2)} \left(\frac{1}{n} + \frac{(X_k - \bar{X})^2}{\sum (X - \bar{X})^2} \right)}$$

При $x_k = \bar{x}$

$$m_{\hat{y}_x} = \sqrt{\frac{265}{(7-2)} + \left(\frac{1}{7} + \frac{(3,3143 - 3,3143)^2}{10,857} \right)} = 2,75$$

При $x_k = 4$

$$m_{\hat{y}_x} = \sqrt{\frac{265}{(7-2)} + \left(\frac{1}{7} + \frac{(4 - 3,3143)^2}{10,857} \right)} = 3,34$$

$$\hat{y}_{x_k} \pm t_a * m_{\hat{y}_x}$$

$$\hat{y}_{x_k} \pm 2,57 * 3,34$$

$$y_p = -5,79 + 36,84 * 4 = 141,57$$

$$\underline{132,99 < \hat{y}_{x_k} < 150,15}$$

Ответ: интервальный прогноз при $x=4$ равен $132,99 < \hat{y}_{x_k} < 150,15$

Задания по вариантам

1. $\bar{X}=3, \Sigma(X - \bar{X})^2 = 10, n=5, RSS = 200$
2. $\bar{X}=7, \Sigma(X - \bar{X})^2 = 3, n=5, RSS = 100$
3. $\bar{X}=15, \Sigma(X - \bar{X})^2 = 7, n=5, RSS = 200$
4. $\bar{X}=25, \Sigma(X - \bar{X})^2 = 10, n=5, RSS = 50$
5. $\bar{X}=30, \Sigma(X - \bar{X})^2 = 12, n=5, RSS = 70$
6. $\bar{X}=3, \Sigma(X - \bar{X})^2 = 10, n=3, RSS = 200$
7. $\bar{X}=7, \Sigma(X - \bar{X})^2 = 3, n=3, RSS = 100$
8. $\bar{X}=15, \Sigma(X - \bar{X})^2 = 7, n=3, RSS = 200$
9. $\bar{X}=25, \Sigma(X - \bar{X})^2 = 10, n=3, RSS = 50$
10. $\bar{X}=30, \Sigma(X - \bar{X})^2 = 12, n=15, RSS = 70$
11. $\bar{X}=3, \Sigma(X - \bar{X})^2 = 10, n=15, RSS = 200$
12. $\bar{X}=7, \Sigma(X - \bar{X})^2 = 3, n=15, RSS = 100$
13. $\bar{X}=15, \Sigma(X - \bar{X})^2 = 7, n=15, RSS = 200$
14. $\bar{X}=25, \Sigma(X - \bar{X})^2 = 10, n=15, RSS = 50$
15. $\bar{X}=30, \Sigma(X - \bar{X})^2 = 12, n=15, RSS = 70$
16. $\bar{X}=3, \Sigma(X - \bar{X})^2 = 10, n=25, RSS = 200$

17. $\bar{X} = 7, \Sigma(X - \bar{X})^2 = 3, n = 25, \text{RSS} = 100$
18. $\bar{X} = 15, \Sigma(X - \bar{X})^2 = 7, n = 25, \text{RSS} = 200$
19. $\bar{X} = 25, \Sigma(X - \bar{X})^2 = 10, n = 25, \text{RSS} = 50$
20. $\bar{X} = 30, \Sigma(X - \bar{X})^2 = 12, n = 25, \text{RSS} = 70$

II МНОЖЕСТВЕННАЯ РЕГРЕССИЯ

Множественная регрессия $y=f(x_1, x_2, \dots, x_k)$ – уравнение связи с несколькими регрессорами.

Например,

линейная модель $y = a + b_1x_1 + b_2x_2 + \dots + b_kx_k + \varepsilon$; (Плоскость регрессии является k -мерной гиперплоскостью в $(k+1)$ -мерном пространстве.)

линейная модель без свободного члена

$$y = b_1x_1 + b_2x_2 + \dots + b_kx_k + \varepsilon;$$

степенная $y = ax_1^{b_1} x_2^{b_2} \dots x_k^{b_k} \cdot \varepsilon$;

экспонента $y = e^{a+b_1x_1+b_2x_2+\dots+b_kx_k} \cdot \varepsilon$;

гипербола $y = (a + b_1x_1 + b_2x_2 + \dots + b_kx_k + \varepsilon)^{-1}$ и т.п.

Для *идентификации* неизвестных параметров множественной регрессии, *линейно зависящей от неизвестных параметров* вновь применим МНК. **Нормальная форма** системы линейных уравнений примет вид:

$$\begin{cases} an + b_1 \sum x_1 + b_2 \sum x_2 + \dots + b_k \sum x_k = \sum y, \\ a \sum x_1 + b_1 \sum x_1^2 + b_2 \sum x_1x_2 + \dots + b_k \sum x_kx_1 = \sum x_1y, \\ \dots \\ a \sum x_k + b_1 \sum x_1x_k + b_2 \sum x_2x_k + \dots + b_k \sum x_k^2 = \sum x_ky. \end{cases}$$

Её решения по формулам Крамера: $a = \Delta_a / \Delta$, $b_1 = \Delta_{b_1} / \Delta$, ...,

$$b_k = \Delta_{b_k} / \Delta,$$

где определитель системы

$$\Delta := \begin{vmatrix} n & \sum x_1 & \sum x_2 & \cdots & \sum x_k \\ \sum x_1 & \sum x_1^2 & \sum x_2 x_1 & \cdots & \sum x_k x_1 \\ \sum x_2 & \sum x_1 x_2 & \sum x_2^2 & \cdots & \sum x_k x_2 \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ \sum x_k & \sum x_1 x_k & \sum x_2 x_k & \cdots & \sum x_k^2 \end{vmatrix} \neq 0.$$

Для случая 2-х регрессоров x_1 и x_2 можно преобразовать МНК-параметры к виду

$$b_1 = \frac{\text{Cov}(x_1, y)\text{Var}(x_2) - \text{Cov}(x_2, y)\text{Cov}(x_1, x_2)}{\text{Var}(x_1)\text{Var}(x_2) - \text{Cov}^2(x_1, x_2)},$$

$$b_2 = \frac{\text{Cov}(x_2, y)\text{Var}(x_1) - \text{Cov}(x_1, y)\text{Cov}(x_1, x_2)}{\text{Var}(x_1)\text{Var}(x_2) - \text{Cov}^2(x_1, x_2)},$$

$$a = \bar{y} - b_1 \bar{x}_1 - b_2 \bar{x}_2.$$

Для большего числа регрессоров формулы для коэффициентов значительно усложняются, но формула свободного члена регрессии остаётся практически той же: $a = \bar{y} - b_1 \bar{x}_1 - b_2 \bar{x}_2 - \dots - b_k \bar{x}_k$. Как и в случае парной регрессии, она гарантирует, что найденная **МНК-плоскость** линейной выборочной регрессии всегда проходит через **центр диаграммы рассеяния** $(\bar{x}_1, \bar{x}_2, \dots, \bar{x}_k, \bar{y}) \in \mathcal{R}^{k+1}$.

Общий и частный F-критерии Фишера

Значимость уравнения множественной регрессии **в целом** оценивается **общим F-критерием** Фишера

$$F := \frac{R^2}{1-R^2} \cdot \frac{n-m-1}{m}.$$

Частный F -критерий позволяет оценить статистическую значимость вклада в уравнение регрессии фактора x_i .

$$F_{x_i} := \frac{R_{yx_1 \dots x_{i-1} x_{i+1} \dots x_k}^2 - R_{yx_1 \dots x_{i-1} x_{i+1} \dots x_k}^2}{1 - R_{yx_1 \dots x_{i-1} x_{i+1} \dots x_k}^2} \cdot \frac{n-m-1}{1}$$

Стандартизованное уравнение регрессии

Для ранжирования значимости регрессоров x_i перейдём к уравнению множественной регрессии в **стандартизованном масштабе**

$$t_y = \beta_1 t_{x_1} + \beta_2 t_{x_2} + \dots + \beta_k t_{x_k},$$

$$t_y := \frac{y_x - \bar{y}}{\sigma_y}, \quad t_{x_i} := \frac{x_i - \bar{x}_i}{\sigma_{x_i}}$$

где t_y, t_{x_i} – **стандартизованные переменные**, β_i – **стандартизованные коэффициенты** линейной регрессии.

Применение МНК для идентификации стандартизованных коэффициентов даёт следующую систему линейных уравнений в нормальной форме:

$$\begin{cases} \beta_1 + \beta_2 r_{x_2 x_1} + \beta_3 r_{x_3 x_1} + \dots + \beta_k r_{x_k x_1} = r_{yx_1}, \\ \beta_1 r_{x_1 x_2} + \beta_2 + \beta_3 r_{x_3 x_2} + \dots + \beta_k r_{x_k x_2} = r_{yx_2}, \\ \dots \\ \beta_1 r_{x_1 x_k} + \beta_2 r_{x_2 x_k} + \beta_3 r_{x_3 x_k} + \dots + \beta_k = r_{yx_k}. \end{cases}$$

Коэффициенты b_i линейной множественной регрессии в исходных (**натуральных**) показателях связаны со стандартизованными коэффици-

циентами β_i соотношением $b_i = \beta_i \sigma_y / \sigma_{x_i}$. Свободный член a уравнения в натуральных показателях определяется как $a = \bar{y} - b_1 \bar{x}_1 - b_2 \bar{x}_2 - \dots - b_k \bar{x}_k$.

Индекс множественной корреляции

Тесноту совместного влияния регрессоров на результат при *нелинейной* регрессии оценивает **индекс множественной корреляции**

$$R_{yx_1x_2\dots x_k} := \sqrt{1 - \sigma_{\text{ост}}^2 / \sigma_y^2}, \quad 0 \leq R_{yx_1x_2\dots x_k} \leq 1.$$

При этом справедлива **нижняя оценка** $R_{yx_1x_2\dots x_k} \geq \max_{1 \leq i \leq k} r_{yx_i}$, т.е. индекс множественной корреляции мажорирует коэффициенты парной линейной корреляции между регрессорами x_i и результирующим признаком y .

Через стандартизованные коэффициенты и парные коэффициенты линейной корреляции его можно выразить в виде

$$R_{yx_1x_2\dots x_k} = \sqrt{\sum \beta_i r_{yx_i}}.$$

Задания по теме «Множественная регрессия»

Задание 2.1. Найти частные уравнения множественной регрессии.

Пример:

Исходные данные: $\hat{y} = 0,5 + 0,35x_1 + 0,77x_2$, x_1 ср = 5, x_2 ср = 7.

Решение:

$$y = 0,5 + 0,35 * x_1 + 0,77 * 7 = 5,89 + 0,35 * x_1$$

$$y = 0,5 + 0,35 * 5 + 0,77 * x_2 = 2,25 + 0,77 * x_2$$

Ответ: $y = 5,89 + 0,35 * x_1$, $y = 2,25 + 0,77 * x_2$

Задания по вариантам

1. $y = 20 + 5x_1 + 3x_2$, x_1 ср = 5, x_2 ср = 7.
2. $y = 12 + 7x_1 + 44x_2$, x_1 ср = 5, x_2 ср = 7.
3. $y = 20 + 51x_1 + 13x_2$, x_1 ср = 5, x_2 ср = 7.
4. $y = 20 + 15x_1 + 31x_2$, x_1 ср = 5, x_2 ср = 7.
5. $y = 20 + 0,5x_1 + 3x_2$, x_1 ср = 5, x_2 ср = 7.
6. $y = 20 + 5x_1 + 3x_2$, x_1 ср = 7, x_2 ср = 5.
7. $y = 12 + 7x_1 + 44x_2$, x_1 ср = 7, x_2 ср = 5.
8. $y = 20 + 51x_1 + 13x_2$, x_1 ср = 7, x_2 ср = 5.
9. $y = 20 + 15x_1 + 31x_2$, x_1 ср = 7, x_2 ср = 5.
10. $y = 20 + 0,5x_1 + 3x_2$, x_1 ср = 7, x_2 ср = 5.
11. $y = 20 + 5x_1 + 3x_2$, x_1 ср = 15, x_2 ср = 7.
12. $y = 12 + 7x_1 + 44x_2$, x_1 ср = 15, x_2 ср = 7.
13. $y = 20 + 51x_1 + 13x_2$, x_1 ср = 15, x_2 ср = 7.
14. $y = 20 + 15x_1 + 31x_2$, x_1 ср = 15, x_2 ср = 7.
15. $y = 20 + 0,5x_1 + 3x_2$, x_1 ср = 15, x_2 ср = 7.
16. $y = 20 + 5x_1 + 3x_2$, x_1 ср = 5, x_2 ср = 27.
17. $y = 12 + 7x_1 + 44x_2$, x_1 ср = 5, x_2 ср = 27.
18. $y = 20 + 51x_1 + 13x_2$, x_1 ср = 5, x_2 ср = 27.
19. $y = 20 + 15x_1 + 31x_2$, x_1 ср = 5, x_2 ср = 27.
20. $y = 20 + 0,5x_1 + 3x_2$, x_1 ср = 5, x_2 ср = 27.

Задание 2.2. Найти уравнение регрессии в стандартизированном масштабе

Пример:

Исходные данные: $r_{yx_1} = 0,7$, $r_{yx_2} = 0,9$, $r_{x_1x_2} = 0,2$, $\sigma_{x_1} = 47$, $\sigma_{x_2} = 300$, $\sigma_y = 15$, y ср = 15, x_1 ср = 7, x_2 ср = 3.

Решение:

Для вычисления коэффициентов множественной регрессии применим метод стандартизации переменных и построим искомое уравнение в стандартизованном масштабе

$$t_y = \beta_1 * t_{x1} + \beta_2 * t_{x2}$$

Расчёт β -коэффициентов выполняется по формулам:

$$\beta_1 = \frac{r_{yx1} - r_{yx2} \cdot r_{x1x2}}{1 - r_{x1x2}^2}$$

$$\beta_2 = \frac{r_{yx2} - r_{yx1} \cdot r_{x1x2}}{1 - r_{x1x2}^2}$$

$$\beta_1 = (0,7 - 0,9 * 0,2) / (1 - 0,2^2) = 0,54$$

$$\beta_2 = (0,9 - 0,7 * 0,2) / (1 - 0,2^2) = 0,79$$

Для построения уравнения в естественной форме рассчитаем b_1 и b_2 , используя формулы перехода;

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

$$b_1 = 0,54 * (47/15) = 1,69$$

$$b_2 = 0,79 * (300/15) = 15,8$$

Значение a определим из соотношения

$$a = 15 - 7 * 1,69 - 3 * 15,8 = -44,23$$

Ответ: $-44,23 + 1,69x_1 + 15,8x_2$

$$a = \bar{y} - b_1 \bar{x}_1 - b_2 \bar{x}_2$$

Задания по вариантам

1. $r_{yx1} = 0,9$, $r_{yx2} = 0,7$, $r_{x1x2} = 0,2$, $\sigma_{x1} = 47$, $\sigma_{x2} = 300$, $\sigma_y = 15$,
 $y_{cp} = 15$, $x_{1cp} = 7$, $x_{2cp} = 3$.

2. $r_{yx1} = 0,7$, $r_{yx2} = 0,9$, $r_{x1x2} = 0,2$, $\sigma_{x1} = 17$, $\sigma_{x2} = 100$, $\sigma_y = 15$,
 $y_{cp} = 5$, $x1_{cp} =$, $x2_{cp} = .7$
3. $r_{yx1} = 0,75$, $r_{yx2} = 0,9$, $r_{x1x2} = 0,2$, $\sigma_{x1} = 47$, $\sigma_{x2} = 30$, $\sigma_y = 15$,
 $y_{cp} = 15$, $x1_{cp} = 7$, $x2_{cp} = 3$.
4. $r_{yx1} = 0,7$, $r_{yx2} = 0,95$, $r_{x1x2} = 0,2$, $\sigma_{x1} = 47$, $\sigma_{x2} = 100$, $\sigma_y = 15$,
 $y_{cp} = 15$, $x1_{cp} = 7$, $x2_{cp} = 3$.
5. $r_{yx1} = 0,7$, $r_{yx2} = 0,9$, $r_{x1x2} = 0,2$, $\sigma_{x1} = 17$, $\sigma_{x2} = 300$, $\sigma_y = 15$,
 $y_{cp} = 15$, $x1_{cp} = 7$, $x2_{cp} = 3$.
6. $r_{yx1} = 0,9$, $r_{yx2} = 0,7$, $r_{x1x2} = 0,1$, $\sigma_{x1} = 47$, $\sigma_{x2} = 300$, $\sigma_y = 15$,
 $y_{cp} = 15$, $x1_{cp} = 7$, $x2_{cp} = 3$.
7. $r_{yx1} = 0,7$, $r_{yx2} = 0,9$, $r_{x1x2} = 0,5$, $\sigma_{x1} = 17$, $\sigma_{x2} = 100$, $\sigma_y = 15$,
 $y_{cp} = 5$, $x1_{cp} =$, $x2_{cp} = .7$
8. $r_{yx1} = 0,75$, $r_{yx2} = 0,9$, $r_{x1x2} = 0,3$, $\sigma_{x1} = 47$, $\sigma_{x2} = 30$, $\sigma_y = 15$,
 $y_{cp} = 15$, $x1_{cp} = 7$, $x2_{cp} = 3$.
9. $r_{yx1} = 0,7$, $r_{yx2} = 0,95$, $r_{x1x2} = 0,4$, $\sigma_{x1} = 47$, $\sigma_{x2} = 100$, $\sigma_y = 15$,
 $y_{cp} = 15$, $x1_{cp} = 7$, $x2_{cp} = 3$.
10. $r_{yx1} = 0,75$, $r_{yx2} = 0,9$, $r_{x1x2} = 0,2$, $\sigma_{x1} = 17$, $\sigma_{x2} = 300$, $\sigma_y = 15$,
 $y_{cp} = 15$, $x1_{cp} = 7$, $x2_{cp} = 3$.
11. $r_{yx1} = 0,9$, $r_{yx2} = 0,7$, $r_{x1x2} = 0,2$, $\sigma_{x1} = 47$, $\sigma_{x2} = 300$, $\sigma_y = 15$,
 $y_{cp} = 15$, $x1_{cp} = 7$, $x2_{cp} = 3$.
12. $r_{yx1} = 0,7$, $r_{yx2} = 0,9$, $r_{x1x2} = 0,2$, $\sigma_{x1} = 17$, $\sigma_{x2} = 100$, $\sigma_y = 15$,
 $y_{cp} = 5$, $x1_{cp} =$, $x2_{cp} = .7$
13. $r_{yx1} = 0,75$, $r_{yx2} = 0,9$, $r_{x1x2} = 0,2$, $\sigma_{x1} = 47$, $\sigma_{x2} = 30$, $\sigma_y = 15$,
 $y_{cp} = 15$, $x1_{cp} = 7$, $x2_{cp} = 3$.
14. $r_{yx1} = 0,7$, $r_{yx2} = 0,95$, $r_{x1x2} = 0,2$, $\sigma_{x1} = 47$, $\sigma_{x2} = 100$, $\sigma_y = 15$,
 $y_{cp} = 15$, $x1_{cp} = 7$, $x2_{cp} = 3$.
15. $r_{yx1} = 0,7$, $r_{yx2} = 0,9$, $r_{x1x2} = 0,2$, $\sigma_{x1} = 17$, $\sigma_{x2} = 300$, $\sigma_y = 15$,
 $y_{cp} = 15$, $x1_{cp} = 7$, $x2_{cp} = 3$.

16. $r_{yx1} = 0,9$, $r_{yx2} = 0,7$, $r_{x1x2} = 0,3$, $\sigma_{x1} = 47$, $\sigma_{x2} = 300$, $\sigma_y = 15$,
 $y_{cp} = 15$, $x1_{cp} = 7$, $x2_{cp} = 3$.
17. $r_{yx1} = 0,7$, $r_{yx2} = 0,9$, $r_{x1x2} = 0,4$, $\sigma_{x1} = 17$, $\sigma_{x2} = 100$, $\sigma_y = 15$,
 $y_{cp} = 5$, $x1_{cp} =$, $x2_{cp} = .7$
18. $r_{yx1} = 0,75$, $r_{yx2} = 0,9$, $r_{x1x2} = 0,5$, $\sigma_{x1} = 47$, $\sigma_{x2} = 30$, $\sigma_y = 15$,
 $y_{cp} = 15$, $x1_{cp} = 7$, $x2_{cp} = 3$.
19. $r_{yx1} = 0,7$, $r_{yx2} = 0,6$, $r_{x1x2} = 0,2$, $\sigma_{x1} = 47$, $\sigma_{x2} = 100$, $\sigma_y = 15$,
 $y_{cp} = 15$, $x1_{cp} = 7$, $x2_{cp} = 3$.
20. $r_{yx1} = 0,7$, $r_{yx2} = 0,9$, $r_{x1x2} = 0,3$, $\sigma_{x1} = 17$, $\sigma_{x2} = 300$, $\sigma_y = 15$,
 $y_{cp} = 15$, $x1_{cp} = 7$, $x2_{cp} = 3$.

Задание 2.3. Найти индекс множественной корреляции

Пример:

Исходные данные: $TSS = 15000$, $ESS = 14735$, $RSS = 265$, $n = 7$.

Решение:

$$R_{yx_1x_2 \dots x_n} = \sqrt{1 - \frac{RSS}{TSS}}$$

$$R_{yx_1x_2 \dots x_n} = \sqrt{1 - \frac{265}{15000}} = 0,99$$

Ответ: $R = 0,99$.

Задания по вариантам

1. $TSS = 10000$, $ESS = 9000$, $RSS = 1000$, $n=5$
2. $TSS = 10000$, $ESS = 9500$, $RSS = 500$, $n=5$
3. $TSS = 10000$, $ESS = 9800$, $RSS = 200$, $n=5$
4. $TSS = 9000$, $ESS = 8000$, $RSS = 1000$, $n=5$
5. $TSS = 9000$, $ESS = 8500$, $RSS = 500$, $n=5$

6. TSS = 10000, ESS = 9000, RSS = 1000, n=7
7. TSS = 10000, ESS = 9500, RSS = 500, n=7
8. TSS = 10000, ESS = 9800, RSS = 200, n=7
9. TSS = 9000, ESS = 8000, RSS = 1000, n=7
10. TSS = 9000, ESS = 8500, RSS = 500, n=7
11. TSS = 10000, ESS = 9000, RSS = 1000, n=3
12. TSS = 10000, ESS = 9500, RSS = 500, n=3
13. TSS = 10000, ESS = 9800, RSS = 200, n=3
14. TSS = 9000, ESS = 8000, RSS = 1000, n=3
15. TSS = 9000, ESS = 8500, RSS = 500, n=3
16. TSS = 1000, ESS = 900, RSS = 100, n=3
17. TSS = 1000, ESS = 950, RSS = 50, n=3
18. TSS = 1000, ESS = 980, RSS = 20, n=3
19. TSS = 900, ESS = 800, RSS = 100, n=3
20. TSS = 900, ESS = 850, RSS = 50, n=3

Задание 2.4. Найти средний показатель эластичности:

Пример:

Исходные данные: $y_{cp} = 15$, $x_{cp} = 7$, $b = 2$.

Решение:

$$\bar{\varepsilon} = b * \frac{\bar{x}}{\bar{y}}$$

$$\varepsilon_{cp} = 2 * 7 / 15 = 0,93$$

Ответ: $\varepsilon_{cp} = 0,93$

Задания по вариантам

1. $y_{cp} = 15, x_{cp} = 7, b = 2.$
2. $y_{cp} = 5, x_{cp} = 7, b = 2.$
3. $y_{cp} = 12, x_{cp} = 7, b = 2.$
4. $y_{cp} = 10, x_{cp} = 7, b = 2.$
5. $y_{cp} = 7, x_{cp} = 7, b = 2.$
6. $y_{cp} = 15, x_{cp} = 17, b = 2.$
7. $y_{cp} = 5, x_{cp} = 71, b = 2.$
8. $y_{cp} = 12, x_{cp} = 27, b = 2.$
9. $y_{cp} = 10, x_{cp} = 5, b = 2.$
10. $y_{cp} = 7, x_{cp} = 3, b = 2.$
11. $y_{cp} = 15, x_{cp} = 12, b = 2.$
12. $y_{cp} = 5, x_{cp} = 7, b = 2.$
13. $y_{cp} = 12, x_{cp} = 7, b = 23.$
14. $y_{cp} = 10, x_{cp} = 7, b = 32.$
15. $y_{cp} = 7, x_{cp} = 7, b = 25.$
16. $y_{cp} = 15, x_{cp} = 7, b = 12.$
17. $y_{cp} = 5, x_{cp} = 7, b = 20.$
18. $y_{cp} = 12, x_{cp} = 17, b = 12.$
19. $y_{cp} = 10, x_{cp} = 7, b = 22.$
20. $y_{cp} = 7, x_{cp} = 27, b = 3.$

III Временные ряды

Временной ряд — это совокупность наблюдений какого-либо показателя $x(t_1), x(t_2), \dots, x(t_N)$ за несколько последовательных **моментов** или **периодов** времени. Включает как динамические, так и статические последовательности уровней какого — либо показателя.

Динамические ряды — ряды уровней, в которых содержится тенденция изменения.

Интервальным вариационным рядом называют упорядоченную совокупность интервалов варьирования значений случайной величины с соответствующими частотами или относительными частотами попаданий в каждый из них значений величины.

Для построения интервального ряда необходимо:

- 1) определить **величину** частичных интервалов;
- 2) определить **ширину** интервалов;
- 3) установить для каждого интервала его **верхнюю** и **нижнюю границы**;
- 4) сгруппировать результаты наблюдения.

Динамический моментный ряд отражает значения показателей на определенном момент времени (дату времени).

Основные компоненты ряда:

- тренд;
- сезонность;
- случайная компонента.

Тренд — это долговременная тенденция изменения исследуемого временного ряда. Тренды могут быть описаны различными уравнениями — линейными, логарифмическими, степенными и так далее.

Сезонность — периодически колебания, наблюдаемые на временных рядах.

Каждый его **уровень** формируется из **трендовой** (T), **циклической** (S) и случайной (E) компонент.

Аддитивная модель: $Y = T + S + E$; **мультипликативная** модель: $Y = T \cdot S \cdot E$.

Стационарность и нестационарность временного ряда

Интуитивное представление – ряд имеет постоянное среднее значение и постоянную дисперсию. Однако необходимо учесть и менее очевидную внутреннюю связь между наблюдениями временного ряда в разные моменты времени. Поэтому необходимо добавить требование **постоянства автокорреляционной функции по времени**.

Ряд $x(t)$ называется **строго стационарным (сильно стационарным, стационарным в узком смысле)**, если совместное распределение вероятностей m наблюдений $x(t_1), x(t_2), \dots, x(t_m)$ такое же, как и совместное распределение наблюдений $x(t_1 + \tau), x(t_2 + \tau), \dots, x(t_m + \tau)$ при любых $m, t_1, t_2, \dots, t_m, \tau$.

Таким образом, свойства строго стационарного временного ряда не меняются при изменении **начала отсчёта времени**. В частности, при $m = 1$ закон распределения $x(t)$ не зависит от t , а значит, не зависят от t мат. ожидание $Mx(t) = a = \text{Const}$ и дисперсия $Dx(t) = M(x(t) - a)^2 = \sigma^2 = \text{Const}$.

Выборочными оценками этих моментов являются **выборочное среднее**

$$\hat{a} := \frac{1}{N} \sum_{t=1}^N x(t)$$

и выборочная дисперсия

$$\hat{\sigma}^2 := \frac{1}{N} \sum_{t=1}^N (x(t) - \hat{a})^2 = \frac{1}{N} \sum_{t=1}^N x^2(t) - \hat{a}^2$$

или **исправленная выборочная дисперсия**

$$s^2 := \frac{1}{N-1} \sum_{t=1}^N (x(t) - \hat{a})^2 = \frac{N}{N-1} \hat{\sigma}^2.$$

Проверка *строгой* стационарности на практике невозможна. Поэтому вводится ослабленное с точки зрения математической теории, но экономически обоснованное понятие *слабой* стационарности.

Ряд $x(t)$ называют **слабо стационарным** (стационарным в широком смысле), если не зависят от времени его среднее значение и дисперсия.

Не удовлетворяющие этим определениям ряды называют **нестационарными**.

Из строгой стационарности очевидно следует слабая; обратное в общем случае не верно.

Автоковариация и автокорреляция

Из предположения о строгой стационарности временного ряда $x(t)$ при $m = 2$ следует совпадение совместных двумерных распределений пар СВ $(x(t_1), x(t_2))$ и $(x(\tau), x(t_2 - t_1 + \tau))$. Они зависят лишь от разности $t_2 - t_1$, но не от начала отсчёта τ .

Тогда ковариация СВ $x(t)$ и $x(t \pm \tau)$ зависит только от сдвига по времени τ , но не от t .

Соответственно, **автоковариационная функция** $\gamma(\tau) := \text{cov}(x(t), x(t+\tau))$ будет зависеть только от сдвига τ и будет чётна: $\gamma(-\tau) = \gamma(\tau)$.

Метод скользящего среднего

Его идея заключается в замене исходного временного ряда $x(1), x(2), \dots, x(n)$ с дисперсией σ^2 **сглаженным рядом** из средних взвешенных соседних $2m+1$ значений

$$\hat{f}(t) := \sum_{k=-m}^m w_k x(t+k), \quad t \in \overline{m+1, n-m}, \quad (1)$$

с весовыми коэффициентами

$$w_k \in [0,1]: \sum_{k=-m}^m w_k = 1$$

и меньшей дисперсией $\sigma^2 / (2m + 1)$.

При запуске t от $m + 1$ до $n - m$ «маска» для расчёта (1) скользит по оси времени так, что при каждом следующем пересчёте происходит замена только одного слагаемого $x(t - m)$ слагаемым $x(t + m + 1)$. Поэтому этот метод назван **методом скользящего среднего**.

Задания по теме «Временные ряды»

Задание 3.1. Провести сглаживание ряда методом скользящей средней.

Пример:

Месяц	1	2	3	4	5	6	7	8	9	10	11	12
Объём продаж	71	68	67	68	68	64	62	64	58	62	51	53
Скользящая средняя	69	68	68	67	65	63	61	61	57	55	53	50

Объём продаж – первоначальный «несглаженный» ряд данных.

Для сглаживания выбираем окно равное 3, т.е. берём первые 3 значения и находим среднее. Результат записываем в первую ячейку ряда (можно начать со второй или третьей ячеек).

$$\text{Скольльзящая средняя } 1 = (71+68+67)/3=69$$

Потом берём 2,3 и 4 значения и находим среднее. Результат записываем в следующую свободную ячейку ряда.

$$\text{Скольльзящая средняя } 2 = (68+67+68)/3=68$$

Задания по вариантам

Месяц	1	2	3	4	5	6	7	8	9	10	11	12
Вариант 1	41	37	43	49	51	58	63	65	60	49	41	42
Вариант 2	44	41	47	50	54	62	66	65	62	55	46	46
Вариант 3	50	47	53	60	64	68	69	68	64	56	47	48
Вариант 4	49	45	51	59	62	74	70	69	75	58	49	56
Вариант 5	41	37	43	49	51	58	63	65	60	49	41	42
Вариант 6	44	41	47	50	54	62	66	65	62	55	46	46
Вариант 7	50	47	53	60	64	68	69	68	64	56	47	48
Вариант 8	49	45	51	59	62	74	70	69	75	58	49	56
Вариант 9	41	37	43	49	51	58	63	65	60	49	41	42
Вариант 10	44	41	47	50	54	62	66	65	62	55	46	46
Вариант 11	50	47	53	60	64	68	69	68	64	56	47	48
Вариант 12	49	45	51	59	62	74	70	69	75	58	49	56
Вариант 13	41	37	43	49	51	58	63	65	60	49	41	42
Вариант 14	44	41	47	50	54	62	66	65	62	55	46	46
Вариант 15	50	47	53	60	64	68	69	68	64	56	47	48
Вариант 16	49	45	51	59	62	74	70	69	75	58	49	56
Вариант 17	41	37	43	49	51	58	63	65	60	49	41	42
Вариант 18	44	41	47	50	54	62	66	65	62	55	46	46

Вариант 19	50	47	53	60	64	68	69	68	64	56	47	48
Вариант 20	49	45	51	59	62	74	70	69	75	58	49	56

Задание 3.2. Известно уравнение тренда и значения сезонной компоненты, определить значение Y по аддитивной модели в периоде t .

Пример:

Известно уравнение тренда $T = 2 + 5 * t$, $t = 15$. Значения сезонной компоненты $S_3 = 7$, $S_1 = 3$.

Аддитивная модель выглядит $Y = T + S$

Значение сезонной компоненты $T = 2 + 5 * 15 = 32$.

Значения сезонной компоненты одинаковы для любого года.

Так как $t = 15$, значит нам необходимо значение сезонной компоненты в третьем месяце.

S_3 – это значение сезонной компоненты в третьем месяце каждого года.

Таким образом,

$$Y = 32 + 7 = 39.$$

Ответ: $Y = 39$.

Задачи по вариантам

Сделать прогноз по аддитивной модели для заданного периода t .

Уравнение тренда $T = 12 + 3 * t$.

№	S1	S2	S3	S4	S5	S6	S7	S8	S9	S10	S11	S12
Значение сезонной компоненты	1	3	5	3	1	-1	-3	-5	-3	-1	1	3

№ варианта	1	2	3	4	5	6	7	8	9	10
t	7	13	25	15	44	45	3	21	73	18
№ варианта	11	12	13	14	15	16	17	18	19	20
t	4	17	35	27	32	19	11	23	52	26

СИСТЕМЫ ЭКОНОМЕТРИЧЕСКИХ УРАВНЕНИЙ

Переменные, входящие в систему уравнений подразделяют на *экзогенные, эндогенные* и лаговые (эндогенные переменные, влияние которых характеризуется некоторым запаздыванием, временным лагом τ).

Экзогенные и лаговые переменные называют *предопределенными*, т. е. определенными заранее.

Классификация переменных на эндогенные и экзогенные зависит от принятой теоретической концепции модели. Экономические показатели могут выступать в одних моделях как эндогенные, а в других как экзогенные переменные. Внеэкономические переменные (например, климатические условия, социальное положение, пол, возраст) входят в систему только как экзогенные переменные. В качестве экзогенных переменных могут рассматриваться значения эндогенных переменных за предшествующий период времени (*лаговые переменные*).

Структурная форма модели описывает реальное экономическое явление или процесс. При структурной форме в системе одни и те же переменные одновременно рассматриваются как зависимые в одних уравнениях и как независимые в других.

Классификация систем эконометрических уравнений:

1) **система независимых уравнений** – каждая зависимая переменная (y) рассматривается как функция одного набора регрессоров (x):

$$\begin{cases} y_1 = a_{11}x_1 + a_{12}x_2 + \dots + a_{1m}x_m + \varepsilon_1, \\ y_2 = a_{21}x_1 + a_{22}x_2 + \dots + a_{2m}x_m + \varepsilon_2, \\ \dots \\ y_n = a_{n1}x_1 + a_{n2}x_2 + \dots + a_{nm}x_m + \varepsilon_n. \end{cases}$$

Для её решения используется МНК.

2) **система рекурсивных уравнений** – зависимая переменная (y) одного уравнения является регрессором в следующем уравнении:

$$\begin{cases} y_1 = a_{11}x_1 + a_{12}x_2 + \dots + a_{1m}x_m + \varepsilon_1, \\ y_2 = b_{21}y_1 + a_{21}x_1 + a_{22}x_2 + \dots + a_{2m}x_m + \varepsilon_2, \\ y_3 = b_{31}y_1 + b_{32}y_2 + a_{31}x_1 + a_{32}x_2 + \dots + a_{3m}x_m + \varepsilon_3, \\ \dots \\ y_n = b_{n1}y_1 + b_{n2}y_2 + \dots + b_{n,n-1}y_{n-1} + a_{n1}x_1 + a_{n2}x_2 + \dots + a_{nm}x_m + \varepsilon_n. \end{cases}$$

Для её решения используется МНК, который применяется к уравнениям системы по очереди, начиная с первого.

3) **система взаимосвязанных (одномоментных) уравнений** – зависимые переменные (y) в одних уравнениях входят в левую часть, а в других – в правую:

$$\begin{cases} y_1 = b_{12}y_2 + b_{13}y_3 + \dots + b_{1n}y_n + a_{11}x_1 + a_{12}x_2 + \dots + a_{1m}x_m + \varepsilon_1, \\ y_2 = b_{21}y_1 + b_{23}y_3 + \dots + b_{2n}y_n + a_{21}x_1 + a_{22}x_2 + \dots + a_{2m}x_m + \varepsilon_2, \\ \dots \\ y_n = b_{n1}y_1 + b_{n2}y_2 + \dots + b_{n,n-1}y_{n-1} + a_{n1}x_1 + a_{n2}x_2 + \dots + a_{nm}x_m + \varepsilon_n. \end{cases}$$

Её называют ещё **структурной формой модели**, а её коэффициенты – **структурными**. МНК для её решения не применим, т.к. он даёт несостоятельные оценки идентифицируемых параметров.

Эндогенные переменные – взаимозависимые переменные (y), определённые внутри модели.

Экзогенные переменные – независимые переменные (x), определённые вне модели.

Предопределённые переменные – экзогенные и лаговые (за предыдущие моменты времени) эндогенные переменные системы.

Система линейных зависимостей всех эндогенных переменных от всех predetermined называется **приведённой формой модели**:

$$\begin{cases} \tilde{y}_1 = \delta_{11}x_1 + \delta_{12}x_2 + \dots + \delta_{1m}x_m, \\ \tilde{y}_2 = \delta_{21}x_1 + \delta_{22}x_2 + \dots + \delta_{2m}x_m, \\ \dots \\ \tilde{y}_n = \delta_{n1}x_1 + \delta_{n2}x_2 + \dots + \delta_{nm}x_m. \end{cases}$$

Её коэффициенты называются **приведёнными**.

Идентифицируемость и идентификация уравнений системы

Идентифицируемость системы уравнений – возможность определения коэффициентов системы уравнений.

Идентификация системы уравнений – процесс проверки идентифицируемости каждого уравнения системы.

Задача идентификации системы уравнений сводится к корректной и однозначной оценке ее коэффициентов.

Счётное правило (необходимое условие идентифицируемости):

$M-m=k-1$ – уравнение идентифицируемо;

$M-m < k-1$ – уравнение неидентифицируемо;

$M-m > k-1$ – уравнение сверхидентифицируемо;

M – количество predetermined переменных в системе; m – количество predetermined переменных в уравнении; k – количество y в уравнении.

Достаточное условие идентификации

Если определитель матрицы, составленной из коэффициентов при переменных, отсутствующих в рассматриваемом структурном уравнении, не равен нулю и ранг этой матрицы не меньше числа эндогенных

переменных системы без единицы, то это уравнение точно (просто) идентифицируемо.

Модель считается идентифицируемой, если каждое уравнение системы идентифицируемо.

Если хотя бы одно из уравнений системы неидентифицируемо, то и вся модель считается неидентифицируемой.

Сверхидентифицируемая модель содержит хотя бы одно сверхидентифицируемое уравнение.

Оценка неидентифицированного уравнения осуществляется с помощью *косвенного метода наименьших квадратов (КМНК)*.

Алгоритм КМНК включает 3 шага: 1) составление приведенной формы модели и выражение каждого коэффициента приведенной формы через структурные параметры; 2) применение обычного МНК к каждому уравнению приведенной формы и получение численных оценок приведенных параметров; 3) определение оценок параметров структурной формы по оценкам приведенных коэффициентов, используя соотношения, найденные на шаге 1.

Оценка сверхидентифицированного уравнения осуществляется при помощи *двухшагового метода наименьших квадратов*.

Алгоритм двухшагового МНК включает следующие шаги: 1) составление приведенной формы модели; 2) применение обычного МНК к каждому уравнению приведенной формы и получение численных оценок приведенных параметров; 3) определение расчетных значений эндогенных переменных, которые фигурируют в качестве факторов в структурной форме модели; 4) определение структурных параметров каждого уравнения в отдельности обычным МНК, используя в качестве факторов входящие в это уравнение predetermined переменные и расчетные значения эндогенных переменных, полученные на шаге 1.

Задания по теме «Системы эконометрических уравнений»

Задание 3.3. Провести идентификацию системы уравнений.

Пример:

Исходные данные:

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

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

Решение:

1) количество предопределённых переменных в системе:

предопределёнными являются x_1 и x_2 , следовательно

$$M_1 = M_2 = M_3 = 2$$

2) количество предопределённых переменных в уравнении:

$$m_1 = 2, m_2 = 1$$

3) количество y в уравнении:

$$k_1 = 1, k_2 = 2.$$

Для первого уравнения $M_1 - m_1$ и $k_1 - 1$, $2 - 2 = 1 - 1$ следовательно, уравнение идентифицируемо и решается МНК.

Для второго уравнения $M_2 - m_2 = k_2 - 1$, $2 - 1 = 2 - 1$ следовательно, уравнение идентифицируемо и решается МНК.

Задачи по вариантам

Таблица 3– Системы структурных уравнений для идентификации

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

<p>Вариант 7.</p> $y_1 = a_{11}x_1 + a_{12}x_2 + a_{13}x_3$ $y_2 = b_{21}y_1 + b_{23}y_3 + a_{22}x_2$ $y_3 = b_{32}y_2 + a_{31}x_1 + a_{33}x_3$	<p>Вариант 8.</p> $y_1 = b_{12}y_2 + a_{11}x_1 + a_{13}x_3$ $y_2 = b_{23}y_3 + a_{22}x_2 + a_{23}x_3$ $y_3 = b_{31}y_1 + a_{31}x_1 + a_{33}x_3$	<p>Вариант 9.</p> $y_1 = b_{12}y_2 + a_{11}x_1 + a_{13}x_3$ $y_2 = b_{23}y_3 + a_{11}x_1 + a_{22}x_2$ $y_3 = a_{31}x_1 + a_{32}x_2 + a_{33}x_3$
<p>Вариант 10.</p> $y_1 = a_{11}x_1 + a_{12}x_2 + a_{13}x_3$ $y_2 = b_{21}y_1 + a_{22}x_2 + a_{23}x_3$ $y_3 = b_{32}y_2 + a_{31}x_1 + a_{33}x_3$	<p>Вариант 11.</p> $y_1 = b_{12}y_2 + a_{11}x_1 + a_{13}x_3$ $y_2 = b_{23}y_3 + a_{22}x_2 + a_{23}x_3$ $y_3 = a_{31}x_1 + a_{32}x_2 + a_{33}x_3$	<p>Вариант 12.</p> $y_1 = b_{12}y_2 + a_{11}x_1 + a_{13}x_3$ $y_2 = b_{23}y_3 + a_{22}x_2 + a_{23}x_3$ $y_3 = b_{31}y_1 + a_{31}x_1 + a_{33}x_3$
<p>Вариант 13.</p> $y_1 = a_{11}x_1 + a_{12}x_2 + a_{13}x_3$ $y_2 = b_{21}y_1 + a_{22}x_2 + a_{23}x_3$ $y_3 = b_{32}y_2 + a_{31}x_1 + a_{33}x_3$	<p>Вариант 14.</p> $y_1 = b_{12}y_2 + a_{11}x_1 + a_{13}x_3$ $y_2 = b_{23}y_3 + a_{22}x_2 + a_{22}x_2$ $y_3 = b_{31}y_1 + a_{31}x_1 + a_{33}x_3$	<p>Вариант 15.</p> $y_1 = b_{12}y_2 + a_{11}x_1 + a_{13}x_3$ $y_2 = a_{11}x_1 + a_{22}x_2 + b_{23}y_3$ $y_3 = b_{31}y_1 + a_{31}x_1 + a_{33}x_3$
<p>Вариант 16.</p> $y_1 = b_{12}y_2 + a_{11}x_1 + a_{13}x_3$ $y_2 = b_{21}y_1 + a_{22}x_2 + a_{23}x_3$ $y_3 = b_{31}y_1 + a_{31}x_1 + a_{33}x_3$	<p>Вариант 17.</p> $y_1 = b_{12}y_2 + a_{11}x_1 + a_{13}x_3$ $y_2 = b_{23}y_3 + a_{21}x_1 + a_{22}x_2$ $y_3 = b_{31}y_1 + a_{31}x_1 + a_{33}x_3$	<p>Вариант 18.</p> $y_1 = b_{12}y_2 + a_{12}x_2 + a_{13}x_3$ $y_2 = b_{23}y_3 + a_{22}x_2 + a_{23}x_3$ $y_3 = b_{31}y_1 + a_{32}x_2 + a_{33}x_3$
<p>Вариант 19.</p> $y_1 = b_{12}y_2 + a_{11}x_1 + a_{13}x_3$ $y_2 = b_{23}y_3 + a_{22}x_2 + a_{23}x_3$ $y_3 = b_{31}y_1 + a_{31}x_1 + a_{33}x_3$	<p>Вариант 20.</p> $y_1 = b_{12}y_2 + a_{11}x_1 + a_{13}x_3$ $y_2 = b_{23}y_3 + a_{22}x_2 + a_{23}x_3$ $y_3 = a_{31}x_1 + a_{32}x_2 + a_{33}x_3$	

Задание 3.4. Найти структурные коэффициенты системы уравнений, исходя из приведённой формы модели.

Пример:

Исходные данные:

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

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

коэффициенты системы уравнений в приведённой форме:

2; 4;

3; -6;

Решение:

Так как система уравнений в структурной форме выглядит как

$$\begin{aligned}y_1 &= a_{11}x_1 + a_{12}x_2 \\ y_2 &= b_{21}y_1 + a_{22}x_2 ,\end{aligned}$$

то в приведённой форме она будет выглядеть следующим образом:

$$\begin{aligned}y_1 &= \delta_{11}x_1 + \delta_{11}x_2 \\ y_2 &= \delta_{21}x_1 + \delta_{22}x_2.\end{aligned}$$

Подставляем коэффициенты в приведённую форму модели

$$y_1 = 2x_1 + 4x_2$$

$$y_2 = 3x_1 - 6x_2.$$

Найдём коэффициенты первого уравнения системы:

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

$$y_1 = 2x_1 + 4x_2$$

Так как структурная и приведённая форма уравнения полностью совпадают (одни и те же переменные в левой и правой частях уравнений), просто подставляем коэффициенты из приведённой формы уравнения в структурную.

$$y_1 = 2x_1 + 4x_2$$

Найдём коэффициенты второго уравнения системы:

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

$$y_2 = 3x_1 - 6x_2.$$

Оцениваем структурную форму уравнения – по сравнению в приведённой формой в уравнении не хватает переменной x_1 , но есть лишняя y_1 , следовательно, нужно выразить x_1 через y_1 (обе переменные есть в первом уравнении модели в приведённой форме).

$$y_1 = 2x_1 + 4x_2$$

$$x_1 = y_1 / 2 - 4x_2 / 2$$

$$x_1 = y_1 - 2x_2$$

Теперь подставим x_1 во второе уравнение модели в приведённой форме

$$y_2 = 3x_1 - 6x_2$$

$$y_2 = 3(y_1 - 2x_2) - 6x_2$$

$$y_2 = 3y_1 - 6x_2 - 6x_2$$

$$y_2 = 3y_1 - 12x_2$$

Полученное уравнение выглядит так же как уравнение в структурной форме, но с числовыми значениями коэффициентов. Задача решена.

Задачи по вариантам

Система уравнений в структурной форме для всех вариантов одинакова:

$$y_1 = b_{12}y_2 + a_{12}x_2 + a_{13}x_3$$

$$y_2 = b_{23}y_3 + a_{22}x_2 + a_{23}x_3$$

$$y_3 = b_{31}y_1 + a_{32}x_2 + a_{33}x_3$$

Таблица – Исходные данные по вариантам для приведённой формы модели

Вариант 1	Вариант 2	Вариант 11	Вариант 12
2; 4; 10	2; 6; -5	2; 6; -5	-3; 4; 10
3; -6; 2	-3; 4; 2	3; 4; 2	2; 5; -2
-5; 8; 5	5; 8; -10	5; -8; 10	6; 8; 5

Вариант 3 -2; 4; 10 3; 6; 2 5; -8; 5	Вариант 4 2; 8; 10 3; -4; 2 -5; 6; 5	Вариант 13 2; 4; 10 3; -6; 2 -5; 8; 5	Вариант 14 2; 6; -5 -3; 4; 2 5; 8; -10
Вариант 5 2; 4; -10 3; 6; -2 5; 8; 5	Вариант 6 3; 4; 10 2; -5; 2 -6; 8; 5	Вариант 15 -2; 4; 10 3; 6; 2 5; -8; 5	Вариант 16 2; 8; 10 3; -4; 2 -5; 6; 5
Вариант 7 2; 6; 5 3; 4; 2 5; 8; 10	Вариант 8 3; 4; 10 2; 5; 2 6; 8; 5	Вариант 17 2; 4; -10 3; 6; -2 5; 8; 5	Вариант 18 3; 4; 10 2; -5; 2 -6; 8; 5
Вариант 9 2; 6; -5 -3; 4; 2 5; 8; 10	Вариант 10 3; 4; -10 2; 5; 2 -6; 8; 5	Вариант 19 2; 6; 5 3; 4; 2 5; 8; 10	Вариант 20 3; 4; 10 2; 5; 2 6; 8; 5

СПИСОК ЛИТЕРАТУРЫ

1. Айвазян, С.А. Эконометрика: краткий курс / С.А. Айвазян. – М: Маркет Дс, 2010. – 104 с.
2. Айвазян, С.А. Эконометрика-2. Продвинутый курс с приложениями в финансах: учебник / Магистр, 2015. – 944с.
3. Бородич, С.А. Эконометрика: практикум. М: ИНФРА-М, 2014. – 329 с.
4. Буравлёв, А. Эконометрика: учеб. пособие / М: Бином, 2012. – 166 с.
5. Герасимов, А.Н., Гладилин А.В. Эконометрика. Теория и практика. – Кнорус, 2011, CD
6. Гладилин, А.В. Эконометрика / А.В. Гладилин, А.Н. Герасимов, Е.И. Громов. – М: Феникс, 2011. – 304 с.
7. Горлач, Б.А. Теория вероятностей и математическая статистика: учеб. пособие. – СПб.: Издательство «Лань», 2013. – 320 с.: ил.
8. Кремер, Н.Ш., Путко Б.А.. Эконометрика / Н.Ш. Кремер, Б.А. Путко; под ред. Н.Ш. Кремера. – М.: ЮНИТИ, 2010. – 328 с.
9. Клентак, Л.С. Элементы теории вероятностей и математической статистики: учеб. пособие / Л.С. Клентак. – Самара: Изд-во Самар. гос. аэрокосм. ун-та, 2013. – 156 с.
10. Костромин, А.В. Эконометрика / А.В. Костромин. – Издательство КноРус, 2015. – 232 с.
11. Новиков, А.И. Эконометрика, М: ИНФРА-М, 2014. – 272с.
12. Озерная, С. А. Эконометрика: метод. указания к лабораторному практикуму по специальностям «Бизнес-информатика», «Менеджмент», «Финансы и кредит» / Самар. гос. аэрокосм. ун-т; Сост. С. А. Озерная. – Самара, 2013. – 76 с.

13. Соколов, Г.А. Эконометрика. Теоретические основы. учеб. пособие. – М: ИНФРА-М, 2012. – 216 с.
14. Эконометрика / И. И. Елисева [и др.] ; под редакцией И. И. Елисейевой. — М.: Юрайт, 2019. — 449 с.
15. Эконометрика: учебник для бакалавров. В.С. Тимофеев, А.В. Фаддеенков, В.Ю. Щеколдин, – М: Юрайт, 2013. – 328 с.
16. Эконометрика: учебник для магистров. Елисева И.И. – М: Юрайт, 2014. – 449 с.

ПРИЛОЖЕНИЕ 1

Таблица значений F – критерия ФИШЕРА

k2	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	3,18	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
10	0,10	3,29	2,92	2,73	2,61	2,52	2,46	2,41	2,38	2,35	2,32	2,30	2,28
	0,05	4,96	4,10	3,71	3,48	3,33	3,22	3,14	3,07	3,02	2,98	2,94	2,91
	0,01	10,0	7,56	6,55	5,99	5,54	5,39	5,20	5,06	4,94	4,85	4,77	4,71
11	0,10	3,23	2,86	2,66	2,54	2,45	2,39	2,34	2,30	2,27	2,25	2,23	2,21
	0,05	4,84	3,98	3,59	3,36	3,20	3,09	3,01	2,95	2,90	2,85	2,82	2,79
	0,01	9,65	7,21	6,22	5,67	5,32	5,07	4,89	4,74	4,63	4,54	4,46	4,40
12	0,05	4,75	3,89	3,49	3,26	3,11	3,00	2,91	2,85	2,80	2,75	12	0,05
13	0,05	4,67	3,81	3,41	3,18	3,03	2,92	2,83	2,77	2,71	2,67	13	0,05
14	0,05	4,60	3,74	3,34	3,11	2,96	2,85	2,76	2,70	2,65	2,60	14	0,05
15	0,05	4,54	3,68	3,29	3,06	2,90	2,79	2,71	2,64	2,59	2,54	15	0,05
16	0,05	4,49	3,63	3,24	3,01	2,85	2,74	2,66	2,59	2,54	2,49	16	0,05
17	0,05	4,45	3,59	3,20	2,96	2,81	2,70	2,61	2,55	2,49	2,45	17	0,05
18	0,05	4,41	3,55	3,16	2,93	2,77	2,66	2,58	2,51	2,46	2,41	18	0,05
19	0,05	4,38	3,52	3,13	2,90	2,74	2,63	2,54	2,48	2,42	2,38	19	0,05
20	0,05	4,35	3,49	3,10	2,87	2,71	2,60	2,51	2,45	2,39	2,35	20	0,05

ПРИЛОЖЕНИЕ 2

Таблица критических значений t-критерия Стьюдента

k- число степеней свободы, α – уровень значимости

$\alpha \backslash$ k	0,4	0,25	0,10	0,05	0,025	0,01	0,005	0,001	0,0005
1	0,325	1,000	3,078	6,341	12,706	31,821	63,657	31,831	6,366
2	0,289	0,816	1,886	2,920	4,303	6,965	9,925	22,327	31,6
3	0,277	0,765	1,638	2,353	3,182	4,541	5,841	10,214	12,94
4	0,271	0,741	1,533	2,132	2,776	3,747	4,604	7,173	8,610
5	0,267	0,727	1,476	2,015	2,571	3,365	4,032	5,893	6,859
6	0,265	0,718	1,440	1,943	2,447	3,143	3,707	5,208	5,959
7	0,263	0,711	1,415	1,895	2,365	2,998	3,499	4,785	5,405
8	0,262	0,706	1,397	1,860	2,306	2,896	3,355	4,501	5,041
9	0,261	0,703	1,383	1,833	2,262	2,821	3,250	4,297	4,781
10	0,260	0,700	1,372	1,812	2,228	2,764	3,169	4,144	4,587
11	0,259	0,697	1,363	1,796	2,201	2,718	3,106	4,025	4,437
12	0,259	0,695	1,356	1,782	2,179	2,681	3,055	3,930	4,318
13	0,258	0,694	1,350	1,771	2,160	2,650	3,012	3,852	4,221
14	0,258	0,692	1,345	1,761	2,145	2,624	2,977	3,787	4,140
15	0,258	0,691	1,341	1,753	2,131	2,602	2,947	3,733	4,073

16	0,258	0,690	1,337	1,746	2,120	2,583	2,921	3,686	4,015
17	0,257	0,689	1,333	1,740	2,110	2,567	2,898	3,646	3,965
18	0,257	0,688	1,330	1,734	2,101	2,552	2,878	3,610	3,922
19	0,257	0,688	1,328	1,729	2,093	2,539	2,861	3,579	3,883
20	0,257	0,687	1,325	1,725	2,086	2,528	2,845	3,552	3,850

ПРИЛОЖЕНИЕ 3

Исходные данные к лабораторной работе по теме «Множественная регрессия»

№	y1	y2	y3	y4	y5	y6	x1	x2	x3	x4	x5	x6	x7	x8
1	71	31	60	203	8,5	12	0,55	30,1	250	5	11	8,60	17	14,8
2	68	30	72	213	8	16	0,55	29,8	250	5	10,5	8,60	22	12,8
3	67	27	65	224	7,3	18	0,55	28,4	250	5	9,2	8,60	29,7	7,3
4	68	30	67	243	8,1	16	0,55	27,1	250	5	9,5	8,60	34,1	14
5	68	29	68	256	7,8	18	0,55	25,3	250	5	9,4	8,60	38	9,5
6	64	26	92	314	6,8	22	0,55	18,3	250	5	9	8,60	58,7	6,7
7	62	25	102	356	6	37	0,55	18	250	5	8,5	8,60	64,2	6,2
8	64	30	77	450	7,5	19	0,55	17,7	250	5	8,5	8,60	64,4	8,5
9	58	21	120	412	5,5	65	0,55	21,7	250	5	8,5	8,60	68	7,5
10	62	24	132	349	5	34	0,55	17,3	250	5	8,4	8,60	71	5,6
11	51	20	142	375	5	35	0,55	16,6	250	5	6,5	8,60	82	5,2
12	53	22	150	366	5	41	0,55	16,6	230	5	7,1	8,60	82	5,2
13	49	17	132	394	6,5	30	0,55	16,6	230	6	6,1	8,60	82	5,2
14	55	18	132	432	5	65	0,55	19	250	5	7	8,60	85,1	3,1
15	53	18	110	459	5,2	62	0,55	19	230	6	7	8,60	85,1	3,1
16	57	19	127	432	4,5	65	0,55	19	230	5	8	8,60	85,1	3,1
17	59	20	136	421,9	4	70	0,55	19	220	5	8,5	8,60	85,1	3,1
18	45	17	150	498	5,7	46	0,55	15,2	250	5	2,5	8,60	170,2	3
19	50	18	150	451	5,4	78	0,55	15,2	230	5	3,5	8,60	170,2	3
20	43	16	150	448	3,5	150	0,55	13	250	5	2	8,60	270	2,1
21	47	18	150	439	3	150	0,55	13	230	5	3	8,60	270	2,1
22	48	18	150	445	3	150	0,55	13	220	5	3,5	8,60	270	2,1
23	54	21	150	400	7,2	70	0,55	18,3	250	5	7,3	8,60	55,5	6,4

24	58	23	150	476	6,9	71	0,55	20,4	250	5	7,5	8,60	54,1	7,9
25	55	23	150	483	6,8	65	0,55	16,5	250	5	7	8,60	67	6,2
26	54	18	150	371	6,5	65	0,55	16,9	250	5	7,5	8,60	81,5	7,7
27	53	17	150	381	6	68	0,55	16,6	250	5	7	8,60	90,1	7,9
28	52	17	150	466	4,7	63	0,55	15,9	250	5	6,5	8,60	99,5	3,5
29	69	25	67	149	6,9	14	1,85	27,1	250	5	7,2	8,9	34	14
30	68	19	65	210	6	18	1,85	19,9	250	5	8,5	8,9	41	13,1
31	63	18	113	245	5	98	1,85	21,2	250	5	6,5	8,9	58	11,7
32	63	18	113	245	5	38	1,85	18,3	250	5	6,5	8,9	83,64	4,3
33	68	19	133	235	2,5	47	1,85	18,3	230	5	4,1	8,9	83,64	4,3
34	54	17	132	263	5,3	27	1,85	18,3	230	6	3,5	8,9	83,64	4,3
35	63	20	143	271	4,9	44	1,85	15,9	250	5	6,3	8,9	96	11,1
36	57	21	128	294,9	5	43	1,85	16,3	250	5	5,5	8,9	101	6,9
37	56	17	134	306	4,5	41	1,85	15,8	250	5	5,6	8,9	103	6,6
38	66	18	150	300	4,3	49	1,85	15,8	230	5	4,5	8,9	103	6,6
39	68	19	150	257	4,3	66	1,85	15,8	220	5	4,2	8,9	103	6,6
40	65	18	134	309	4,1	51	1,85	15,7	250	5	6	8,9	107	6,1
41	64	17	143	306	4,7	49	1,85	15,6	230	5	6,5	8,9	115	7,2
42	53	16	150	326	3,8	44	1,85	13,9	250	5	5	8,9	119	10,4
43	55	17	150	329	3,4	50	1,85	13,9	230	6	5,1	8,9	119	10,4
44	51	18	150	296	3,7	40	1,85	13,7	250	5	5	8,9	120	10,3
45	56	19	150	295,3	3,6	49	1,85	13,7	220	5	3,5	8,9	120	10,3
46	51	18	150	308	3,7	31	1,85	13,7	220	6	4,4	8,9	120	10,3
47	49	17	96	318,5	3,8	31	1,85	13,2	220	6	5	8,9	130,5	8,2
48	57	18	150	302,3	3	66	1,85	13,2	220	5	5,5	8,9	130,5	8,2
49	53	16	150	319	4	49	1,85	13,2	250	5	4,5	8,9	130,5	8,2
50	52	17	150	340	3,5	150	1,85	12,6	230	5	4,1	8,9	146,3	6,5
51	50	16	150	331	3,9	100	1,85	12,8	250	5	4	8,9	148,2	6,9
52	49	17	150	342	3,6	84	1,85	10,2	250	5	3,3	8,9	171	5,7

53	49	15	150	312	3,5	86	1,85	10,2	250	5	3,5	8,9	200	9,3
54	59	19	150	186	4,8	106	1,85	9,8	230	4	4	8,9	230	5,7
55	45	12	150	338	3,3	65	1,85	7,9	250	5	1,5	8,9	302	5,6
56	61	20	142	306	4,3	43	1,85	14,8	250	5	6	8,9	98,4	4,1
57	64	22	132	300	5	39	1,85	15,4	250	5	5,5	8,9	113	9,4

Учебное издание

*Кузнецова Ольга Александровна,
Мазурмович Ольга Николаевна*

ЭКОНОМЕТРИКА

Практикум

В авторской редакции

Компьютерная верстка Л.Р. Д м и т р и е н к о

Подписано в печать 15.07.2019. Формат 60x84 1/16.

Бумага офсетная. Печ. л. 4,5.

Тираж 120 экз (1-й з-д 1-30). Заказ .

ФЕДЕРАЛЬНОЕ ГОСУДАРСТВЕННОЕ АВТОНОМНОЕ
ОБРАЗОВАТЕЛЬНОЕ УЧРЕЖДЕНИЕ ВЫСШЕГО ОБРАЗОВАНИЯ
«САМАРСКИЙ НАЦИОНАЛЬНЫЙ ИССЛЕДОВАТЕЛЬСКИЙ
УНИВЕРСИТЕТ ИМЕНИ АКАДЕМИКА С.П. КОРОЛЕВА»
(САМАРСКИЙ УНИВЕРСИТЕТ)
443086 Самара, Московское шоссе, 34.

Изд-во Самарского университета.
443086 Самара, Московское шоссе, 34.

