

Документ подписан простой электронной подписью

Информация о владельце:

ФИО: Макаренко Елена Николаевна

Должность: Ректор

Дата подписания: 2021.11.25 10:02

Уникальный программный ключ:

c098bc0c1041cb3a4cf92af171db715d99a6ae00adc8e27b55cbe1e2dbd7c98



МИНИСТЕРСТВО ОБРАЗОВАНИЯ И НАУКИ
РОССИЙСКОЙ ФЕДЕРАЦИИ

РОСТОВСКИЙ ГОСУДАРСТВЕННЫЙ
ЭКОНОМИЧЕСКИЙ УНИВЕРСИТЕТ "РИНХ"

ЭКОНОМЕТРИКА

Методические указания и задания к контрольной работе

Ростов-на-Дону

2004

УДК 330.43 (075.8)
Н 60

Ниворожкина Л.И., Арженовский С.В., Федосова О.Н. Эконометрика: Методические указания и задания к контрольной работе/Рост. гос. экон. ун-т. – Ростов н/Д, 2004. – 37 с.

Методические указания составлены на основе государственного стандарта высшего профессионального образования и включают программу курса "Эконометрика", список рекомендуемой литературы, задания для самостоятельной работы студентов, методические указания к решению типовых задач.

Для студентов и аспирантов всех форм обучения.

Рецензенты: к.э.н., доц. Князевская Н.В., к.э.н., доц. Кокина Е.П.

© Ростовский государственный экономический университет "РИНХ", 2004

© Ниворожкина Л.И., Арженовский С.В., Федосова О.Н., 2004

Введение

Целью изучения "Эконометрики" является знакомство с эконометрическими моделями и методами, выработка навыков их применения для анализа социально-экономических явлений и процессов.

Студенты при изучении данной дисциплины должны уметь использовать методы эконометрики для прикладных целей. В частности, студенты должны уметь обосновывать закономерности изучаемого экономического объекта; определять основные показатели, характеризующие объект; устанавливать взаимосвязи между этими показателями; формировать статистическую информацию о процессе; специфицировать систему совместных уравнений (одно уравнение); идентифицировать взаимосвязи между показателями в системе уравнений (одном уравнении); оценивать качество расчетов по модели; выполнять практические расчеты по модели и делать экономико-математический анализ результатов.

Учебная программа дисциплины "Эконометрика"

Тема 1. Предмет и задачи курса.

Определение эконометрики. Эконометрика и экономическая теория. Эконометрика и статистика. Эконометрика и экономико-математические методы. Области применения эконометрических моделей. Методологические вопросы построения эконометрических моделей: обзор используемых методов.

Литература.

1. Айвазян С.А., Мхитарян В.С. Прикладная статистика и основы эконометрики. Учебник для вузов. – М.: ЮНИТИ, 1998.
2. Айвазян С.А., Мхитарян В.С., Балалова Е.И. Эконометрика: этапы развития и причина популярности//Вопросы статистики, №2, 2001.
3. Магнус Я.Р., Катышев П.К., Пересецкий А.А. Эконометрика. Начальный курс. – М.: Дело, 2000.
4. Эконометрика. Учебное пособие/Под ред. И.И. Елисеевой. – М.: Финансы и статистика, 2001.

Тема 2. Парная регрессия и корреляция.

Понятие о функциональной, статистической и корреляционной связях. Основные задачи прикладного корреляционно-регрессионного анализа.

Уравнение регрессии, его смысл и назначение. Выбор типа математической функции при построении уравнения регрессии.

Парная регрессия. Метод наименьших квадратов и условия его применения для определения параметров уравнения парной регрессии.

Нелинейные модели регрессии и их линеаризация.

Оценка степени тесноты связи между количественными переменными. Коэффициент ковариации. Показатели корреляции: линейный коэффициент корреляции, индекс корреляции, теоретическое корреляционное отношение. Коэффициент детерминации.

Стандартная ошибка уравнения регрессии.

Оценка статистической значимости показателей корреляции, параметров уравнения регрессии, уравнения регрессии в целом: t - критерий Стьюдента, F - критерий Фишера.

Литература.

1. Айвазян С.А., Мхитарян В.С. Прикладная статистика и основы эконометрики. Учебник для вузов. – М.: ЮНИТИ, 1998.
2. Джонстон Дж. Эконометрические методы. – М.: Статистика, 1980.
3. Доугерти К. Введение в эконометрику. – М.: ИНФРА-М, 1997.
4. Магнус Я.Р., Катышев П.К., Пересецкий А.А. Эконометрика. Начальный курс. – М.: Дело, 2000.
5. Эконометрика. Учебное пособие/Под ред. И.И. Елисеевой. – М.: Финансы и статистика, 2001.

Тема 3. Множественная регрессия и корреляция.

Понятие о множественной регрессии. Классическая линейная модель множественной регрессии (КЛММР). Определение параметров уравнения множественной регрессии методом наименьших квадратов.

Стандартизованные коэффициенты регрессии, их интерпретация. Парные и частные коэффициенты корреляции. Множественный коэффициент корреляции и множественный коэффициент детерминации. Оценка надежности показателей корреляции.

Оценка качества модели множественной регрессии: F – критерий Фишера, t - критерий Стьюдента.

Мультиколлинеарность. Методы устранения мультиколлинеарности.

Литература.

1. Айвазян С.А., Мхитарян В.С. Прикладная статистика и основы эконометрики. Учебник для вузов. – М.: ЮНИТИ, 1998.
2. Джонстон Дж. Эконометрические методы. – М.: Статистика, 1980.
3. Доугерти К. Введение в эконометрику. – М.: ИНФРА-М, 1997.
4. Магнус Я.Р., Катышев П.К., Пересецкий А.А. Эконометрика. Начальный курс. – М.: Дело, 2000.
5. Эконометрика. Учебное пособие/Под ред. И.И. Елисеевой. – М.: Финансы и статистика, 2001.

Тема 4. Спецификация переменных в уравнениях регрессии.

Эконометрические модели: общая характеристика, различия статистического и эконометрического подхода к моделированию.

Спецификация переменных в уравнениях регрессии. Ошибки спецификации.

Обобщенная линейная модель множественной регрессии. Обобщенный метод наименьших квадратов.

Проблема гетероскедастичности. Автокорреляция. Анализ линейной модели множественной регрессии при гетероскедастичности и автокорреляции.

Фиктивные переменные: общий случай. Множественные совокупности фиктивных переменных. Фиктивные переменные для коэффициентов наклона. Тест Чоу.

Моделирование: влияние отсутствия переменной, которая должна быть включена; влияние включения в модель переменной, которая не должна быть включена. Замещающие переменные.

Литература.

1. Айвазян С.А., Мхитарян В.С. Прикладная статистика и основы эконометрики. Учебник для вузов. – М.: ЮНИТИ, 1998.
2. Доугерти К. Введение в эконометрику. – М.: ИНФРА-М, 1997.
3. Кремер Н., Путко Б. Эконометрика. – М.: ЮНИТИ-ДАНА, 2002.
4. Магнус Я.Р., Катышев П.К., Пересецкий А.А. Эконометрика. Начальный курс. – М.: Дело, 2000.

Тема 5. Временные ряды в эконометрических исследованиях.

Специфика временных рядов как источника данных в эконометрическом моделировании.

Аналитическое выравнивание временных рядов. Оценка параметров уравнения тренда.

Автокорреляция в остатках, ее измерение и интерпретация. Критерий Дарбина-Уотсона в оценке качества трендового уравнения регрессии.

Анализ временных рядов при наличии периодических колебаний: аддитивная и мультипликативная модели.

Особенности изучения взаимосвязанных временных рядов. Автокорреляция рядов динамики и методы ее устранения.

Метод последовательных разностей. Интерпретация параметров уравнения регрессии, построенного по первым и вторым разностям.

Метод отклонения уровней ряда от основной тенденции. Метод включения фактора времени.

Литература.

1. Айвазян С.А., Мхитарян В.С. Прикладная статистика и основы эконометрики. Учебник для вузов. – М.: ЮНИТИ, 1998.
2. Магнус Я.Р., Катышев П.К., Пересецкий А.А. Эконометрика. Начальный курс. – М.: Дело, 2000.
3. Эконометрика. Учебное пособие/Под ред. И.И. Елисеевой. – М.: Финансы и статистика, 2001.

Тема 6. Системы эконометрических уравнений.

Виды систем эконометрических уравнений. Независимые системы. Рекурсивные системы. Системы одновременных (совместных) уравнений.

Структурная и приведенная формы эконометрической модели. Проблемы идентификации.

Косвенный и двухшаговый метод наименьших квадратов, общая схема алгоритма расчетов.

Применение эконометрических моделей. Модель Кейнса (статистическая и динамическая формы). Модель Клейна.

Литература.

1. Айвазян С.А., Мхитарян В.С. Прикладная статистика и основы эконометрики. Учебник для вузов. – М.: ЮНИТИ, 1998.
2. Арженовский С.В. Системы одновременных уравнений. Текст лекций. – Ростов-на-Дону: РГЭУ, 2002.
3. Джонстон Дж. Эконометрические методы. – М.: Статистика, 1980.
4. Магнус Я.Р., Катышев П.К., Пересецкий А.А. Эконометрика. Начальный курс. – М.: Дело, 2000.
5. Эконометрика. Учебное пособие/Под ред. И.И. Елисеевой. – М.: Финансы и статистика, 2001.

Общие указания по выполнению контрольной работы

Номер варианта задания к контрольной работе соответствует последней цифре номера зачетной книжки. Если последняя цифра номера зачетной книжки 0, то следует выполнить 10-й вариант.

Каждый вариант контрольной работы содержит 4 задачи по основным разделам дисциплины. Порядковый номер задачи по каждой теме соответствует номеру варианта. Исходные данные к задачам для самостоятельной работы приведены в пунктах 1.3, 2.3, 3.3 и 4.3. Задание каждой задачи соответствует заданию типового примера.

Результаты расчетов всех величин необходимо приводить с точностью до 0,001.

Все расчеты могут быть выполнены с использованием пакетов прикладных программ на персональном компьютере. В этом случае следует обязательно указывать название и версию использованного программного обеспечения, и также соответствующие распечатки необходимо привести в тексте работы или оформить в качестве приложения.

Все расчеты должны сопровождаться комментариями и интерпретацией полученных результатов.

Указания к выполнению контрольной работы содержат все необходимые формулы, а также содержат примеры решения типовых задач.

Задача 1

"Парная регрессия и корреляция"*1.1. Краткие сведения из теории.*

Построение поля корреляции производится по исходным данным о парах значений факторного и результативного признаков с соблюдением масштаба. На основе поля корреляции делаются выводы о возможной функциональной форме связи между факторным и результативным признаками.

Оценки $\hat{\alpha}$ и $\hat{\beta}$ параметров α и β в уравнении парной линейной регрессии $Y_i = \alpha + \beta X_i + u_i$ производится методом наименьших квадратов (МНК) из условия минимума суммы квадратов отклонений Y_i от полученных по модели \hat{Y}_i . Расчетные формулы:

$$\hat{\beta} = \frac{n \sum_{i=1}^n X_i Y_i - \left(\sum_{i=1}^n X_i \right) \left(\sum_{i=1}^n Y_i \right)}{n \sum_{i=1}^n X_i^2 - \left(\sum_{i=1}^n X_i \right)^2}, \quad (1)$$

$$\hat{\alpha} = \bar{Y} - \hat{\beta} \bar{X}, \quad (2)$$

где n – количество наблюдений в выборке, $i = 1, \dots, n$, \bar{X} и \bar{Y} – средние арифметические соответственно $\bar{X} = \frac{1}{n} \sum_{i=1}^n X_i$ и $\bar{Y} = \frac{1}{n} \sum_{i=1}^n Y_i$.

Коэффициент регрессии линейной функции $\hat{\beta}$ есть абсолютный показатель силы связи, характеризующий среднее абсолютное изменение результата при изменении факторного признака на единицу своего измерения.

Линейный коэффициент корреляции характеризует тесноту линейной связи между изучаемыми признаками. Его можно определить по следующей формуле:

$$r = \frac{\sum_{i=1}^n (X_i - \bar{X})(Y_i - \bar{Y})}{n \sigma_X \sigma_Y}, \quad (3)$$

где σ_X и σ_Y среднеквадратические отклонения для переменных X и Y соответ-

$$\text{ственно } \sigma_X = \sqrt{\frac{\sum_{i=1}^n (X_i - \bar{X})^2}{n}} \text{ и } \sigma_Y = \sqrt{\frac{\sum_{i=1}^n (Y_i - \bar{Y})^2}{n}}.$$

Значения линейного коэффициента корреляции принадлежит промежутку $[-1; 1]$. Чем ближе его абсолютное значение к 1, тем теснее линейная связь между признаками. Положительная величина свидетельствует о прямой связи между изучаемыми признаками, отрицательная – о наличии обратной связи между признаками.

Квадрат коэффициента корреляции называется коэффициентом детерминации, который показывает долю вариации результативного признака, объяс-

ненную вариацией факторного признака. Чаще всего, давая интерпретацию коэффициента детерминации, его выражают в процентах.

Для проверки гипотезы о статистической значимости коэффициента регрессии формулируется нулевая гипотеза о том, что коэффициент регрессии статистически незначим: $H_0: \beta=0$ (линейной зависимости нет) при конкурирующей гипотезе $H_1: \beta \neq 0$ (линейная зависимость есть).

Фактическое значение критерия для проверки указанной гипотезы имеет t -распределение Стьюдента и рассчитывается по формуле:

$$t = \frac{\hat{\beta}}{\mu_{\beta}}, \quad (4)$$

где μ_{β} - стандартная ошибка коэффициента регрессии, рассчитываемая как

$$\mu_{\beta} = \sqrt{\frac{\sigma_u^2}{\sum_{i=1}^n (X_i - \bar{X})^2}}, \quad (5)$$

здесь σ_u^2 - оценка дисперсии случайных остатков регрессии $\sigma_u^2 = \frac{\sum_{i=1}^n e_i^2}{n-2}$ или с учетом формулы для ошибок регрессии $e_i = Y_i - \hat{Y}_i$ получим

$$\sigma_u^2 = \frac{\sum_{i=1}^n (Y_i - \hat{Y}_i)^2}{n-2}. \quad (6)$$

Для определения табличного значения $t_{\varepsilon, n-2}$ пользуются таблицами распределения Стьюдента для заданного уровня значимости ε , принимая во внимание, что число степеней свободы для t -статистики равно $(n-2)$.

Далее сравнивают полученное фактическое значение с табличным. Если фактическое значение используемого критерия превышает табличное $t > t_{\varepsilon, n-2}$, то нулевая гипотеза отклоняется, и с вероятностью $(1-\varepsilon)$ принимается альтернативная гипотеза о статистической значимости коэффициента регрессии. Если фактическое значение t -критерия меньше табличного $t < t_{\varepsilon, n-2}$, то оснований отклонять нулевую гипотезу нет.

Оценка статистической значимости построенной модели регрессии в целом производится с помощью F -критерия Фишера. В случае парной регрессии нулевая гипотеза формулируется как $H_0: \beta=0$ при конкурирующей гипотезе $H_1: \beta \neq 0$. Фактическое значение F -критерия может быть определено по формуле

$$F = \frac{R^2}{1-R^2}(n-2). \quad (7)$$

Для определения табличного значения $F_{\varepsilon; 1; n-2}$ пользуются таблицами распределения Фишера для заданного уровня значимости ε , принимая во внимание, что в случае парной регрессии число степеней свободы большей дисперсии равно 1, а число степеней свободы меньшей дисперсии равно $(n-2)$.

Если $F > F_{\varepsilon; 1; n-2}$, то нулевая гипотеза отклоняется, и с вероятностью $(1-\varepsilon)$ принимается альтернативная гипотеза о статистической значимости коэффициента регрессии. Если $F < F_{\varepsilon; 1; n-2}$, то оснований отклонять нулевую гипотезу нет.

Отметим, что между значениями t и F , рассчитанных по формулам (4) и (7) соответственно, существует взаимосвязь $t = \sqrt{F}$.

Для расчета точечного прогноза \hat{Y}^* необходимо подставить в уравнение регрессии заданное значение факторного признака X^* , т.е. $\hat{Y}^* = \hat{\alpha} + \hat{\beta}X^*$.

Для расчета интервального прогноза построим доверительный интервал для значения \hat{Y}^* , лежащего на линии регрессии

$$\hat{Y}^* \pm t_{\varepsilon, (n-2)} \sqrt{\sigma_u^2 (1+h^*)}, \quad (8)$$

$$\text{где } h^* = \frac{1}{n} + \frac{(X^* - \bar{X})^2}{\sum_{i=1}^n (X_i - \bar{X})^2}.$$

Полученный интервал будет характеризовать значения результативного признака при заданном значении факторного признака X^* для отдельной наблюдаемой единицы.

1.2. Пример решения типовой задачи.

Имеются данные о ежемесячном количестве посетителей и выручке крупных супермаркетов г. Ростова-на-Дону, приведенные ниже в таблице.

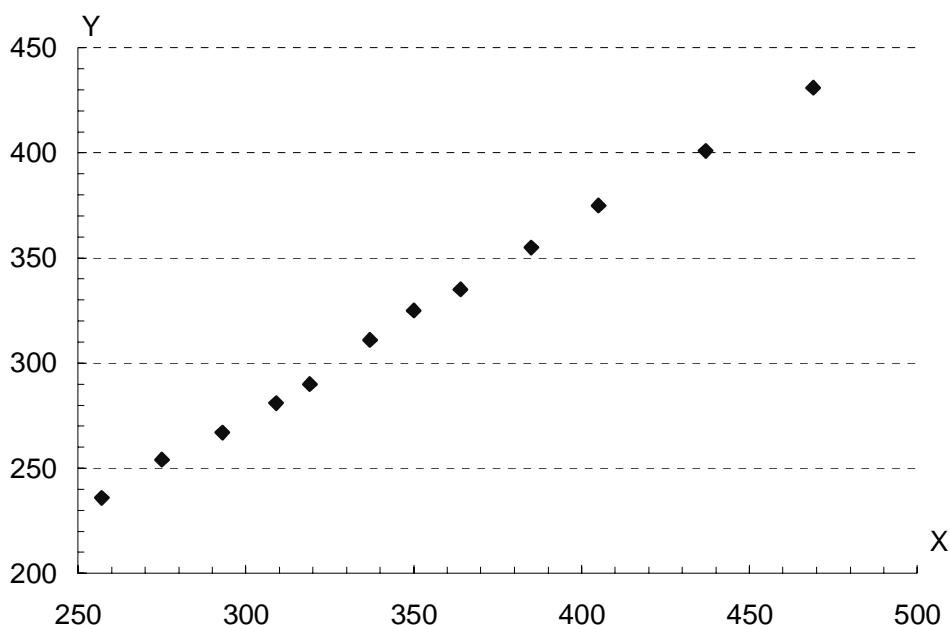
Задание.

- 1). Постройте поле корреляции результативного и факторного признаков.
- 2). Определите параметры уравнения парной линейной регрессии и дайте интерпретацию коэффициента регрессии β .
- 3). Рассчитайте линейный коэффициент корреляции и поясните его смысл. Определите коэффициент детерминации и дайте его интерпретацию.
- 4). С вероятностью 0,95 оцените статистическую значимость коэффициента регрессии β и уравнения регрессии в целом. Сделайте выводы.
- 5). Рассчитайте прогнозное значение \hat{Y}^* для заданного $X^*=500$ и постройте 95% доверительный интервал для прогноза.

Супермаркет	Выручка за месяц, тыс. у.е., Y	Число посетителей за месяц, тыс. чел., X
Пять золотых	236	257
Солнечный круг - 1	254	275
Арагат	267	293
Солнечный круг - 2	281	309
Театральный	290	319
Пчелка	311	337
Пятый элемент	325	350
Северный	335	364
Вавилон - 2	355	385
Западный	375	405
Вавилон - 1	401	437
Рамстор	431	469

Решение.

1). Для условия задачи поле корреляции выглядит следующим образом:



Между выручкой (Y) и количеством посетителей (X) визуально определяется прямая линейная зависимость.

2). Определим параметры уравнения парной линейной регрессии. Вычисления удобно организовать в таблицу. При этом сначала рассчитываются средние значения \bar{X} и \bar{Y} по данным столбцов 2 и 3. Затем в столбцах 4 и 5 рассчитываются $X_i - \bar{X}$, $Y_i - \bar{Y}$, $i = 1, \dots, n$, и в столбце 8 их произведение.

№ п/п	X	Y	$X_i - \bar{X}$	$Y_i - \bar{Y}$	$(X_i - \bar{X})^2$	$(Y_i - \bar{Y})^2$	$(X_i - \bar{X}) * (Y_i - \bar{Y})$	\hat{Y}	e
1	2	3	4	5	6	7	8	9	10
1	257	236	-93	-85,75	8649	7353,06	7974,75	235,48	0,52
2	275	254	-75	-67,75	5625	4590,06	5081,25	252,18	1,82
3	293	267	-57	-54,75	3249	2997,56	3120,75	268,88	-1,88
4	309	281	-41	-40,75	1681	1660,56	1670,75	283,72	-2,72
5	319	290	-31	-31,75	961	1008,06	984,25	292,99	-2,99
6	337	311	-13	-10,75	169	115,56	139,75	309,69	1,31
7	350	325	0	3,25	0	10,56	0	321,75	3,25
8	364	335	14	13,25	196	175,56	185,5	334,74	0,26
9	385	355	35	33,25	1225	1105,56	1163,75	354,22	0,78
10	405	375	55	53,25	3025	2835,56	2928,75	372,77	2,23
11	437	401	87	79,25	7569	6280,56	6894,75	402,45	-1,45
12	469	431	119	109,25	14161	11935,56	13000,75	432,13	-1,13
Сумма	4200	3861	0	0,00	46510	40068,25	43145	3861	0,00
Среднее	350,0	321,75							

По формуле (1) получим: $\hat{\beta} = 43145/46510 = 0,93$. По формуле (2) получим: $\hat{\alpha} = 321,75 - 0,93 \cdot 350 = -2,91$.

Оцененное уравнение регрессии запишется в виде $\hat{Y} = -2,91 + 0,93X$.

Интерпретация коэффициента регрессии. С увеличением числа посетителей на 1 тыс. человек выручка увеличится на 0,93 тыс.у.е.

3). Расчет линейного коэффициента корреляции проведем по формуле (3). С учетом вычислений в столбцах 6, 7 и 8 таблицы, получим:

$$r=43145/\left(\sqrt{46510 \cdot 40068,25}\right)=0,99.$$

Т.е. связь между изучаемыми переменными прямая (коэффициент корреляции положителен) линейная.

Определим коэффициент детерминации $R^2=0,99^2=0,998$. Т.е. 99,8% вариации выручки объясняется вариацией числа посетителей.

4). Оценим статистическую значимость коэффициента регрессии β .

Рассчитаем дисперсию ошибки регрессии по формуле (6) с учетом столбца 10 таблицы: $\sigma_u^2=44,76/(12-2)=4,476$.

Рассчитаем стандартную ошибку коэффициента регрессии по формуле (5): $\mu_\beta=\sqrt{4,476/46510}=0,0098$.

Тогда по формуле (4) фактическое значение t статистики составит

$$t=0,93/0,0098=94,9.$$

По таблице находим для уровня значимости по условию $1-0,95=0,05$ и числа степеней свободы 10: $t_{0,05;10}=2,228$. Поскольку $t_{0,05;10}<t$, то коэффициент регрессии β значим, т.е. наличие статистической связи между выручкой и числом посетителей супермаркетов статистически подтверждается.

Для проверки значимости уравнения регрессии в целом воспользуемся формулой (7):

$$F=10 \cdot 0,998/(1-0,998)=8935,4.$$

Поскольку табличное значение F распределения Фишера $F_{0,05;1;10}=4,96$ меньше расчетного, то гипотеза о статистической незначимости коэффициента регрессии должна быть отвергнута.

5). Рассчитаем прогнозное значение \hat{Y}^* для X^* . Полученный интервал будет характеризовать значения результативного признака (выручки) при заданном значении факторного признака (числа посетителей) для отдельной наблюдаемой единицы. Построим точечный прогноз:

$$\hat{Y}^*=-2,91+0,93 \cdot 500=462,1.$$

Построим 95% доверительный интервал для прогноза по формуле (8). Определим сначала

$$h^*=\frac{1}{n}+\frac{\left(X^*-\bar{X}\right)^2}{\sum_{i=1}^n\left(X_i-\bar{X}\right)^2}=\frac{1}{12}+\frac{\left(500-350\right)^2}{46510}=0,567.$$

И, следовательно, $462,1 \pm 2,228 \sqrt{4,476 \cdot (1+0,567)}=462,1 \pm 5,9$ или

$$456,2 \leq \hat{Y}^* \leq 468.$$

Таким образом, с вероятностью 0,95 можно утверждать, что прогнозное значение выручки отдельного супермаркета, который за месяц посетят 500 тыс. чел. будет находиться в интервале от 456,2 до 468 тыс.у.е.

1.3. Задачи для самостоятельной работы.

1. Имеются данные о количестве копий (тыс. шт.), сделанных копирувальными машинами различных марок в издательских центрах города и стоимости технического обслуживания копирувальных машин (тыс. у. е.):

Количество копий	16	19	24	26	28	29	33	39	40	41	44	45
Стоимость техобслуживания	1,4	1,6	1,7	1,75	1,85	2,4	2,7	2,8	2,8	2,7	2,9	3,0

К заданию 5) $X^* = 21$.

2. Имеются данные по 12 группам населения о среднегодовом доходе и уровне потребления мяса жителями штата Канзас (США):

Среднегодовой доход в среднем по группе, тыс. дол.	41,5	29,6	31,8	69,8	100,5	93,3	82,1	77,4	55,7	38,9	45,2	60,2
Годовое потребление мяса на душу населения в среднем по группе, кг.	41,2	35,3	40,7	55,1	80,1	65,9	64,2	70,5	61,1	51,7	59,4	65,8

К заданию 5) $X^* = 51,4$.

3. По однородным предприятиям имеются данные о количестве рабочих с профессиональной подготовкой и количестве бракованной продукции:

№ предприятия	1	2	3	4	5	6	7	8	9	10
Количество рабочих с проф. подготовкой, %	39,4	40,1	48,6	54,7	58,4	70,5	85,1	65,3	57,3	50,6
Количество бракованной продукции, %	17,1	18,3	11,2	9,3	10,8	5,9	2,8	6,7	8,4	9,5

К заданию 5) $X^* = 50,9$.

4. Периодически в средствах массовой информации обсуждаются высокие должностные оклады президентов благотворительных организаций. Дана информация о десяти крупнейших филиалах общества United Way.

Город	Должностной оклад президента, тыс. дол.	Собранная сумма пожертвований в расчете на душу населения, дол.
Атланта	154	16
Чикаго	189	22
Кливленд	226	31
Денвер	177	30
Детройт	134	14
Хьюстон	195	25
Канзас-Сити	161	21
Лос-Анжелес	142	19
Миннеаполис	159	23
Сиэтл	203	29

К заданию 5) $X^* = 26$.

5. Компания «Вест», состоящая из 12 региональных представительств, продает кухонные принадлежности, рассылая каталоги по почте. Данные, иллюстрирующие количество рассылок (тыс. адресов) и объем выручки региональных представительств компании (млн. у. е.):

Количество адресов рассылки, тыс. адресов	65	55	67	41	59	78	105	110	125	91	47	93
Выручка, млн. у. е.	28	19	20	17,5	20	24,5	29,5	31	35	30	18,5	27

К заданию 5) $X^* = 100$.

6. Данные о тираже бесплатной рекламной газеты «Реклама для вас», распространяемой в различных регионах РФ и стоимости размещения в ней рекламы стандартного размера (1/4 газетной полосы):

Регион	Тираж, тыс.экз.	Стоимость рекламы, тыс.у.е.
Ростовская область	350	1,6
Курская область	125	1,2
Воронежская область	400	2
Московская область	875	2,6
Ставропольский край	450	2
Хабаровский край	200	1,3
Вологодская область	100	0,8
Волгоградская область	300	1,7
Рязанская область	175	1,3
Красноярский край	225	1,5
Иркутская область	85	0,9
Томская область	325	1,6
Тюменская область	400	2,1
Краснодарский край	500	2,2
Ленинградская область	800	2,7

К заданию 5) $X^*=150$.

7. Администрация страховой компании приняла решение о введении нового вида услуг – страхования на случай пожара. С целью определения тарифов по выборке анализируется зависимость стоимости ущерба, нанесенного пожаром, от расстояния до ближайшей пожарной станции:

Общая сумма ущерба, млн. руб.	25,0	38,9	68,1	75,4	91,4	55,3	40,7	79,3	88,8	19,1
Расстояние до ближайшей пожарной станции, км.	4,5	3,8	5,1	4,8	10,1	8,2	6,1	9,2	3,1	2,1

К заданию 5) $X^*=5,5$.

8. Компанию по прокату автомобилей интересует зависимость между пробегом автомобилей и стоимостью ежемесячного обслуживания:

Пробег, тыс. км	8	11	31	15	27	30	13	17	21	23	35	9
Стоимость обслуживания, у.е.	15	17	26	20	24	25	18	21	24	20	30	12

К заданию 5) $X^*=25$.

9. Исследуется зависимость между урожайностью зерновых и количеством внесенных удобрений. Данные по 11 фермерским хозяйствам области:

Урожайность зерновых ц/га	29	36	41	48	46	31	25	22	38	32	26
Внесено удобрений на 1 га посева, кг	12	18	25	19	23	17	11	19	14	18	11

К заданию 5) $X^*=22$.

10. При исследовании годового дохода и сбережений населения получены следующие данные:

Доход, тыс. у.е.	59	83	55	47	39	97	125	150	74	71
Сбережения, тыс. у.е.	8	13	9	3	4	15	16	20	15	17

К заданию 5) $X^*=100$.

Задача 2

"Множественная регрессия и корреляция"*2.1. Краткие сведения из теории.*

Пусть имеется n наблюдений над переменными $Y_i, X_{1i}, X_{2i}, i=1, \dots, n$. Парные коэффициенты корреляции рассчитываются по формулам:

$$r_{YX_j} = \frac{\sum_{i=1}^n (Y_i - \bar{Y})(X_{ji} - \bar{X}_j)}{\sqrt{\sum_{i=1}^n (Y_i - \bar{Y})^2} \sqrt{\sum_{i=1}^n (X_{ji} - \bar{X}_j)^2}}, j \in \{1, 2\}, \quad (9)$$

$$r_{X_1X_2} = \frac{\sum_{i=1}^n (X_{1i} - \bar{X}_1)(X_{2i} - \bar{X}_2)}{\sqrt{\sum_{i=1}^n (X_{1i} - \bar{X}_1)^2} \sqrt{\sum_{i=1}^n (X_{2i} - \bar{X}_2)^2}} \dots \quad (10)$$

Частные коэффициенты корреляции между двумя переменными при фиксированном воздействии другой переменной рассчитываются через значения парных коэффициентов корреляции. Коэффициент частной корреляции между Y и X_1 , когда X_2 является константой:

$$r_{YX_1 \cdot X_2} = \frac{r_{YX_1} - r_{YX_2} r_{X_1X_2}}{\sqrt{(1 - r_{YX_2}^2)(1 - r_{X_1X_2}^2)}}. \quad (11)$$

Коэффициент частной корреляции между Y и X_2 , когда X_1 является константой:

$$r_{YX_2 \cdot X_1} = \frac{r_{YX_2} - r_{YX_1} r_{X_1X_2}}{\sqrt{(1 - r_{YX_1}^2)(1 - r_{X_1X_2}^2)}}. \quad (12)$$

Коэффициент частной корреляции между X_1 и X_2 , когда Y является константой:

$$r_{X_1X_2 \cdot Y} = \frac{r_{X_1X_2} - r_{YX_1} r_{YX_2}}{\sqrt{(1 - r_{YX_1}^2)(1 - r_{YX_2}^2)}}. \quad (13)$$

Классическая линейная модель множественной регрессии записывается в виде:

$$Y_i = \beta_0 + \beta_1 X_{1i} + \beta_2 X_{2i} + u_i, i = 1, \dots, n. \quad (14)$$

где u – случайная величина ошибки.

Применение МНК к (14) позволяет получить формулы для оценок коэффициентов:

$$\hat{\beta}_1 = \frac{r_{YX_1} - r_{YX_2} r_{X_1X_2}}{1 - r_{X_1X_2}^2} \frac{\sigma_Y}{\sigma_{X_1}}, \hat{\beta}_2 = \frac{r_{YX_2} - r_{YX_1} r_{X_1X_2}}{1 - r_{X_1X_2}^2} \frac{\sigma_Y}{\sigma_{X_2}}, \hat{\beta}_0 = \bar{Y} - \hat{\beta}_1 \bar{X}_1 - \hat{\beta}_2 \bar{X}_2, \quad (15)$$

где σ_Y, σ_{X_1} и σ_{X_2} среднеквадратические отклонения для переменных Y, X_1 и X_2

соответственно $\sigma_Y = \sqrt{\frac{\sum_{i=1}^n (Y_i - \bar{Y})^2}{n}}, \quad \sigma_{X_1} = \sqrt{\frac{\sum_{i=1}^n (X_{1i} - \bar{X}_1)^2}{n}}$ и

$\sigma_{X_2} = \sqrt{\frac{\sum_{i=1}^n (X_{2i} - \bar{X}_2)^2}{n}}$, а \bar{Y} , \bar{X}_1 и \bar{X}_2 средние арифметические для переменных Y , X_1 и X_2 соответственно.

Параметры регрессии $\hat{\beta}_1$ и $\hat{\beta}_2$ это показатели, характеризующие абсолютное (в натуральных единицах измерения) изменение резульативного признака при изменении соответствующего факторного признака на единицу своего измерения при фиксированном влиянии другого фактора.

Значение коэффициента детерминации (квадрат коэффициента множественной корреляции) получается по формуле:

$$R^2 = \frac{r_{YX_1}^2 + r_{YX_2}^2 - 2r_{YX_1}r_{YX_2}r_{X_1X_2}}{1 - r_{X_1X_2}^2}. \quad (16)$$

Скорректированный на число степеней свободы коэффициент детерминации для модели (14) равен:

$$R_{adj}^2 = 1 - \frac{n-1}{n-3}(1 - R^2). \quad (17)$$

Для проверки статистической значимости построенной модели регрессии в целом используется F -критерий Фишера. В случае регрессии (14) нулевая гипотеза формулируется как $H_0: \hat{\beta}_1 = \hat{\beta}_2 = 0$ при конкурирующей гипотезе $H_1: \hat{\beta}_1 \neq 0, \hat{\beta}_2 \neq 0$. Фактическое значение F -критерия может быть определено по формуле

$$F = \frac{R^2}{1 - R^2} \frac{n-3}{2}. \quad (18)$$

Для определения табличного значения $F_{\varepsilon; 2; n-3}$ пользуются таблицами распределения Фишера для заданного уровня значимости ε , принимая во внимание, что число степеней свободы большей дисперсии равно 2, а число степеней свободы меньшей дисперсии равно $(n-3)$.

Если $F > F_{\varepsilon; 2; n-3}$, то нулевая гипотеза отклоняется, и с вероятностью $(1-\varepsilon)$ принимается альтернативная гипотеза о статистической значимости регрессии. Если $F < F_{\varepsilon; 2; n-3}$, то оснований отклонять нулевую гипотезу нет.

Расчет относительных показателей силы связи в уравнении множественной регрессии, частных коэффициентов эластичности, производится по формуле:

$$\mathcal{E}_j = \hat{\beta}_j \cdot \frac{\bar{X}_j}{\bar{Y}}, j=1, 2. \quad (19)$$

Частные коэффициенты эластичности \mathcal{E}_j показывают, на сколько процентов от значения своей средней \bar{Y} изменяется результат при изменении фактора X_j на 1% от своей средней \bar{X}_j и при фиксированном воздействии на Y других факторов, включенных в уравнение регрессии.

Доверительные интервалы для коэффициентов уравнения регрессии строятся так:

$$\beta_j = \hat{\beta}_j \pm t_{\varepsilon; n-3} \mu_{\hat{\beta}_j}, j=1, 2. \quad (20)$$

В (20) $t_{\varepsilon; n-3}$ – табличное значение t распределения Стьюдента для уровня значимости ε и $n-3$ степеней свободы, а $\mu_{\hat{\beta}_j}$ – стандартная ошибка коэффициента регрессии $\hat{\beta}_j$, вычисляемая по формуле

$$\mu_{\hat{\beta}_j} = \frac{\sigma_Y \sqrt{(1-R^2)}}{\sqrt{1-r_{x_1x_2}^2} \sigma_{X_j}} : \sqrt{n-3}. \quad (21)$$

Если в доверительный интервал попадает ноль, то делается вывод о незначимости коэффициента регрессии, для которого построен этот интервал, иначе коэффициент регрессии значим.

2.2. Пример решения типовой задачи.

Компания, производящая моющие средства, предприняла рекламную акцию в магазинах с демонстрацией антисептических свойств нового моющего средства. В этот же период компания использовала обычную теле- и радиорекламу. Через некоторое время компания решила проанализировать сравнительную эффективность различных видов рекламы по данным об объеме продаж моющего средства (Y), расходах на теле и радио рекламу (X_1), расходах на демонстрацию товара в магазинах (X_2). Расходы приведены в условных денежных единицах.

№ п.п.	Y	X_1	X_2
1	72	12	5
2	76	11	7
3	78	15	6
4	70	10	5
5	68	11	3
6	80	16	7
7	82	14	3
8	65	8	4
9	62	8	3
10	90	18	5

Задание.

1). Определите парные и частные коэффициенты корреляции. Сделайте выводы.

2). Постройте линейное уравнение множественной регрессии и поясните смысл его параметров. Рассчитайте скорректированный коэффициент детерминации.

3). Проверьте значимость уравнения регрессии на 95% уровне.

4). Рассчитайте коэффициенты эластичности. Дайте их интерпретацию.

5). Постройте 95% доверительные интервалы для коэффициентов регрессии. Проверьте значимость каждого из коэффициентов.

Решение.

1). Рассчитаем средние значения переменных $\bar{Y}=74,3$, $\bar{X}_1=12,3$ и $\bar{X}_2=4,8$. Определим парные коэффициенты корреляции по формулам (9) и (10).

$$r_{YX_1} = \frac{\sum_{i=1}^n (Y_i - \bar{Y})(X_{1i} - \bar{X}_1)}{\sqrt{\sum_{i=1}^n (Y_i - \bar{Y})^2} \sqrt{\sum_{i=1}^n (X_{1i} - \bar{X}_1)^2}} = \frac{243,1}{\sqrt{656,1} \sqrt{102,1}} = 0,94.$$

$$r_{YX_2} = \frac{\sum_{i=1}^n (Y_i - \bar{Y})(X_{2i} - \bar{X}_2)}{\sqrt{\sum_{i=1}^n (Y_i - \bar{Y})^2} \sqrt{\sum_{i=1}^n (X_{2i} - \bar{X}_2)^2}} = \frac{49,6}{\sqrt{656,1} \sqrt{21,6}} = 0,42.$$

$$r_{X_1X_2} = \frac{\sum_{i=1}^n (X_{1i} - \bar{X}_1)(X_{2i} - \bar{X}_2)}{\sqrt{\sum_{i=1}^n (X_{1i} - \bar{X}_1)^2} \sqrt{\sum_{i=1}^n (X_{2i} - \bar{X}_2)^2}} = \frac{19,6}{\sqrt{102,1} \sqrt{21,6}} = 0,42.$$

Значение парного коэффициента корреляции свидетельствует о сильной линейной связи между переменными Y и X_1 . Связь между X_2 и Y существенно слабее. Кроме того, теснота связи между X_1 и X_2 небольшая. Таким образом, можно сделать предварительное заключение, что расходы на демонстрацию моющего средства в магазинах, существенно не влияют на рост объема продаж нового моющего средства.

Расчет частных коэффициентов корреляции по формулам (11)-(13) дает соответственно:

$$r_{YX_1 \cdot X_2} = \frac{r_{YX_1} - r_{YX_2} r_{X_1X_2}}{\sqrt{(1 - r_{YX_2}^2)(1 - r_{X_1X_2}^2)}} = \frac{0,94 - 0,42^2}{1 - 0,42^2} = 0,93.$$

$$r_{YX_2 \cdot X_1} = \frac{r_{YX_2} - r_{YX_1} r_{X_1X_2}}{\sqrt{(1 - r_{YX_1}^2)(1 - r_{X_1X_2}^2)}} = \frac{0,42 - 0,94 \cdot 0,42}{\sqrt{(1 - 0,94^2)(1 - 0,42^2)}} = 0,08.$$

$$r_{X_1X_2 \cdot Y} = \frac{r_{X_1X_2} - r_{YX_1} r_{YX_2}}{\sqrt{(1 - r_{YX_1}^2)(1 - r_{YX_2}^2)}} = \frac{0,42 - 0,94 \cdot 0,42}{\sqrt{(1 - 0,94^2)(1 - 0,42^2)}} = 0,08.$$

Коэффициенты частной корреляции дают более точную характеристику тесноты зависимости двух признаков, чем коэффициенты парной корреляции, так как "очищают" парную зависимость от взаимодействия данной пары переменных с другими переменными, представленными в модели. Наиболее тесно связаны Y и X_1 . Другие взаимосвязи существенно слабее. При сравнении коэффициентов парной и частной корреляции видно, что из-за влияния межфакторной зависимости между X_1 и X_2 происходит завышение оценки тесноты связи между переменными.

2). Построим линейное уравнение множественной регрессии $Y_i = \beta_0 + \beta_1 X_{1i} + \beta_2 X_{2i} + u_i, i = 1, \dots, 10$.

Вычислим сначала $\sigma_Y = \sqrt{\frac{\sum_{i=1}^n (Y_i - \bar{Y})^2}{n}} = \sqrt{\frac{656,1}{10}} = 8,1$,

$$\sigma_{X_1} = \sqrt{\frac{\sum_{i=1}^n (X_{1i} - \bar{X}_1)^2}{n}} = \sqrt{\frac{102,1}{10}} = 3,2, \quad \sigma_{X_2} = \sqrt{\frac{\sum_{i=1}^n (X_{2i} - \bar{X}_2)^2}{n}} = \sqrt{\frac{21,6}{10}} = 1,5.$$

Затем, применяя формулы (15) получим оценки уравнения регрессии:

$$\hat{\beta}_1 = \frac{0,94 - 0,42 \cdot 0,42}{1 - 0,42^2} \cdot \frac{8,1}{3,2} = 2,35. \quad \hat{\beta}_2 = \frac{0,42 - 0,94 \cdot 0,42}{1 - 0,42^2} \cdot \frac{8,1}{1,5} = 0,16.$$

$$\hat{\beta}_0 = 74,3 - 2,35 \cdot 12,3 - 0,16 \cdot 4,8 = 44,62.$$

Уравнение регрессии имеет вид:

$$\hat{Y} = 44,62 + 2,35X_1 + 0,16X_2.$$

Интерпретация коэффициентов регрессии. Константа оценивает агрегированное влияние прочих (кроме учтенных в модели X_1 и X_2) факторов на результат Y и означает, что объем продаж при отсутствии рекламы составил бы 44,62 у.е. Коэффициенты $\hat{\beta}_1$ и $\hat{\beta}_2$ указывают, что с увеличением X_1 и X_2 на единицу их значений объем продаж нового моющего средства увеличивается, соответственно, на 2,35 и 0,16 условных денежных единиц.

Рассчитаем скорректированный коэффициент детерминации по формуле

$$(17) R_{adj}^2 = 1 - \frac{10-1}{10-3} \left(1 - \frac{0,94^2 + 0,42^2 - 2 \cdot 0,94 \cdot 0,42 \cdot 0,42}{1 - 0,42^2} \right) = 0,85.$$

Т.о. 85% вариации зависимой переменной объясняется вариацией независимых переменных.

3). Проверим значимость уравнения регрессии на 95% уровне. Рассчитаем фактическое значение F статистики по формуле (18):

$$F = \frac{R^2}{1-R^2} \frac{n-3}{2} = 0,88 / (1-0,88) \cdot (10-3) / 2 = 26,4.$$

Для определения табличного значения $F_{0,05; 2; 7}$ воспользуемся таблицами распределения Фишера для заданного уровня значимости 0,05, принимая во внимание, что число степеней свободы большей дисперсии равно 2, а число степеней свободы меньшей дисперсии равно 7, $F_{0,05; 2; 7} = 4,74$.

Так как $26,4 = F > F_{0,05; 2; 7} = 4,74$, то нулевая гипотеза отклоняется, и с вероятностью 0,95 можно говорить о статистической значимости уравнения регрессии.

4). Рассчитаем коэффициенты эластичности. Воспользуемся формулой (19). Получим

$$\varepsilon_1 = \hat{\beta}_1 \cdot \frac{\bar{X}_1}{\bar{Y}} = 2,35 \cdot 12,3 / 74,3 = 0,39. \quad \varepsilon_2 = \hat{\beta}_2 \cdot \frac{\bar{X}_2}{\bar{Y}} = 0,16 \cdot 4,8 / 74,3 = 0,01.$$

Значения коэффициентов эластичности позволяют сделать вывод о большем влиянии на объем продаж рекламы на радио и телевидении, нежели рекламы путем демонстрации товара в магазинах. В частности, при изменении фактора X_1 на 1% от своего среднего значения и при фиксированном воздействии на Y другого фактора, включенного в уравнение регрессии, объем продаж изменится на 0,39% от среднего значения.

5). Построим 95% доверительные интервалы для коэффициентов регрессии.

По формуле (21) вычислим стандартные ошибки коэффициентов

$$\mu_{\hat{\beta}_1} = \frac{\sigma_Y \sqrt{(1-R^2)}}{\sqrt{1-r_{x_1x_2}^2} \sigma_{X_1}} : \sqrt{10-3} = \frac{8,1\sqrt{1-0,88}}{\sqrt{1-0,42^2} \cdot 3,2} : \sqrt{7} = 0,36,$$

$$\text{аналогично } \mu_{\hat{\beta}_2} = \frac{8,1\sqrt{1-0,88}}{\sqrt{1-0,42^2} \cdot 1,5} : \sqrt{7} = 0,78.$$

Тогда по (20) с учетом $t_{0,05; 7}=2,36$ получим:

$$\beta_1 = \hat{\beta}_1 \pm t_{\epsilon} (10-3) \mu_{\hat{\beta}_1} = 2,35 \pm 2,36 \cdot 0,36 = 2,35 \pm 0,85 \text{ или } 1,5 \leq \beta_1 \leq 3,2.$$

$$\beta_2 = \hat{\beta}_2 \pm t_{\epsilon} (10-3) \mu_{\hat{\beta}_2} = 0,16 \pm 2,36 \cdot 0,78 = 0,16 \pm 1,84 \text{ или } -1,7 \leq \beta_2 \leq 2,0.$$

Проверка значимости каждого из коэффициентов: поскольку доверительный интервал для β_1 не включает ноль, то коэффициент β_1 значим на 95% уровне, коэффициент β_2 не значим, поскольку доверительный интервал для β_2 включает в себя ноль.

2.3. Задачи для самостоятельной работы.

1. Имеются данные по ряду стран за 199* год:

Страна	Продолжительность жизни при рождении, лет	Численность населения, обслуживаемого одним врачом, чел.	Валовой внутренний продукт на душу населения, тыс. долл.
Россия	65,9	230	9,5
Австрия	77,1	298	24,1
Белоруссия	67,9	230	7,1
Великобритания	77,2	590	21,5
Германия	77,4	295	24,5
Норвегия	78,7	350	29,5
Финляндия	77,2	340	22,4
Франция	78,2	370	25,1
Турция	69,3	820	7,4
Узбекистан	70,3	326	3,5
США	76,9	395	34,3
Япония	81,5	556	31,1

2. Приведены данные о тарифах на размещение одной страницы цветной рекламы в ведущих американских журналах (тыс. долл.), численности планируемой аудитории (млн. чел.), проценте мужчин-читателей.

Издание	Тариф, тыс. долл.	Численность планируемой аудитории, млн. чел.	Процент мужчин-читателей, %
Business Week	115,1	5,9	71,1
Cosmopolitan	97,1	17	15,2
Elle	53,6	4,1	8,5

Издание	Тариф, тыс. долл.	Численность планируемой аудитории, млн. чел.	Процент мужчин-читателей, %
Fortune	61,5	4,6	69,1
Forbes	55,3	5,2	70,3
Life	68,9	16,8	49,7
People	130	41,3	33,1
Reader's Digest	197	56,4	40,3
Newsweek	145,1	24,7	55
National Geographic	167	36,5	59,6
Seventeen	77,5	6,3	8,5
The New Yorker	63,1	4,3	44,3
Time	158	29,9	53,9
TV Guide	135	51,9	40,1
Vogue	65,8	10,1	11,3

3. Представлены сведения о биржевой стоимости одной акции (условных денежных единиц), величине активов (млн.у.е.) и численности служащих (тыс. человек) ряда промышленных компаний.

Компания	1	2	3	4	5	6	7	8	9	10	11	12	13	14
Биржевая стоимость акции, у.е.	10	15	12	41	18	19	25	39	29	33	38	19	24	28
Активы компании, млн.у.е.	120	140	130	250	160	155	175	214	180	200	260	170	184	210
Численность служащих, тыс. чел.	8	11	9	22	12	15	17	18	16	18	20	12	15	17

4. Представлены данные о величине ежемесячных доходов (тыс.у.е.), затратах на питание (тыс.у.е.) и численности членов семьи (человек).

Семья	1	2	3	4	5	6	7	8	9	10	11	12
Расходы на питание, тыс.у.е.	2,4	4,8	3,9	5,9	7,5	3,5	2,1	5,4	3,3	4,2	3,4	1,6
Доход семьи, тыс.у.е.	7,8	12,5	12,9	14,6	22,7	10,5	5,4	18,8	9,6	14,6	9,1	5,2
Количество членов семьи	1	3	2	3	5	2	1	4	3	3	2	1

5. Изучается зависимость потребления электроэнергии (тыс. кВт ч) от производства продукции (тыс. ед.), и уровня механизации труда. Для этого по 20 производственным компаниям, выпускающим одноименную продукцию были получены следующие данные:

№ п/п	Потребление электроэнергии (тыс. кВт ч)	Производство продукции (тыс. ед.)	Уровень механизации труда, %
1	120	12	19
2	130	16	25
3	185	19	28
4	195	22	30
5	200	23	35
6	260	26	47
7	400	38	58
8	450	41	60
9	470	45	65
10	370	39	53
11	230	25	45
12	410	43	67
13	340	33	50
14	230	20	41
15	110	12	33

№ п/п	Потребление электроэнергии (тыс. кВт ч)	Производство продукции (тыс. ед.)	Уровень механизации труда, %
16	100	10	28
17	240	28	44
18	500	60	80
19	550	70	87
20	510	58	85

6. Изучается влияние стоимости основных и оборотных средств на величину валового дохода торговых предприятий г. Ростова-на-Дону. Для этого по 15 торговым предприятиям были получены следующие данные в млн. руб.:

Предприятие	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
Валовой доход за год	29	47	83	46	52	39	71	21	33	68	95	57	43	92	34
Среднегодовая стоимость оборотных средств	9	14	26	14	17	12	23	8	10	21	30	18	13	29	11
Среднегодовая стоимость основных фондов	19	34	60	34	36	29	51	14	21	47	67	40	30	64	24

7. Для исследования зависимости между стоимостью мужских рубашек (у.е.) и составом тканей, использовавшихся при их изготовлении, в магазине мужской одежды было отобрано 15 образцов.

Образец	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
Содержание натуральных волокон, %	70	65	30	40	35	45	50	95	85	90	85	80	65	75	50
Содержание полиэстера, %	25	25	50	40	60	43	40	2	7	5	10	10	27	15	31
Стоимость рубашки, у.е.	30	21	12	16	10	17	19	47	37	42	37	35	28	35	19

8. Для изучения зависимости между производительностью труда, уровнем механизации работ и количеством рабочих, имеющих специальную подготовку, представлены следующие данные.

Предприятие	1	2	3	4	5	6	7	8	9	10	11	12	13	14
Кол-во рабочих с проф. подготовкой, %	38	46	73	92	81	62	55	71	45	56	77	88	65	47
Коэффициент механизации работ, %	46	59	87	98	92	70	65	82	5	60	88	95	75	55
Производительность труда, шт.	25	32	48	60	53	41	38	47	29	36	50	56	43	30

9. Медицинская компания провела обследования людей, имеющих лишний вес. В ходе обследования изучалась зависимость между величиной лишних килограммов (Y), возрастом пациентов (X_1) и среднесуточной калорийностью (X_2) питания. В таблице приведены результаты обследования за один год.

Y	15	17	19	22	35	8	23	11	6	19	17	9	16	23	30
X_1	26	33	39	48	55	25	40	31	22	45	41	23	39	60	58
X_2	2,7	2,9	3,6	4,0	4,1	2,4	3,5	3,0	2,2	3,5	2,9	2,3	3,0	3,6	4,3

10. Изучается зависимость между стоимостью номера, уровнем сервиса и удаленностью от моря в отелях на курортах Турции.

Название отеля	Классность отеля (количество звезд)	Удаленность от моря, метров	Стоимость одноместного номера, у.е.
Туана	2	800	35
Фортуна	3	700	40
Корингия	4	800	60
Мираж	4	400	80

Название отеля	Классность отеля (количество звезд)	Удаленность от моря, метров	Стоимость одноместного номера, у.е.
Амос	5	200	90
Посейдон	2	500	45
Мунамар	4	150	95
Атлантика	3	300	70
Викинги	3	500	55
Венеция	2	400	45
Олимпус	5	300	85
Лимра	4	600	75
Коллекция	2	900	30
Браво	2	300	40
Гавайи	3	200	70

Задача 3

"Временные ряды в эконометрических исследованиях"

3.1. Краткие сведения из теории.

Совокупность наблюдений $y(t_1), y(t_2), \dots, y(t_n)$ анализируемой случайной величины $Y(t)$, произведенных в последовательные моменты времени t_1, t_2, \dots, t_n называется временным рядом, причем n - число наблюдений. Значения элементов временного ряда формируются под воздействием ряда факторов, среди которых выделяют:

- долговременные, формирующие в длительной перспективе тенденцию анализируемого признака. Эта тенденция описывается с помощью некоторой функции, называемой трендом (T);
- сезонные, формирующие периодически повторяемые в определенное время года колебания анализируемого признака (S);
- циклические, формирующие изменения признака в результате воздействия циклов экономической, демографической или астрофизической природы (C);
- случайные, не поддающиеся учету и регистрации, как результат воздействия случайных, внешних факторов (E).

Предметом анализа временного ряда является выделение и изучение вышеуказанных факторов, в дальнейшем именуемых компонентами временного ряда, как правило, в рамках одной из моделей ряда: аддитивной ($Y=T+C+S+E$) или мультипликативной ($Y=T \cdot C \cdot S \cdot E$). Некоторые составляющие могут отсутствовать в тех или иных рядах. При решении задачи 3 предполагается, что речь идет об аддитивной модели временного ряда с отсутствующей циклической компонентой.

Процесс построения аддитивной модели включает следующие этапы:

1. Расчет значений сезонной компоненты S . Простейший путь оценки сезонности для ряда $y_1, y_2, \dots, y_t, \dots, y_n$ с периодом сезонности τ ($\tau=12$ для ежемесячных данных, $\tau=4$ для ежеквартальных данных, и $h=n/\tau$) заключается в вычислении разности между средним по всем одноименным месяцам (кварталам) и средним по всем данным:

$$\frac{1}{h} \sum_{j=0}^{h-1} y_{k+\tau \cdot j} - \frac{1}{n} \sum_{i=1}^n y_i, \quad k = 1, \dots, \tau. \quad (22)$$

Подробно процесс выделения сезонной компоненты представлен при решении типовой задачи в п. 3.2.

2. Устранение сезонной компоненты из исходных уровней ряда и получение z_i суммы трендовой и случайной компонент ($z_i = T_i + E_i$) согласно аддитивной модели $y_i = T_i + S_i + E_i$ откуда $T_i + E_i = z_i = y_i - S_i$.

3. Аналитическое выравнивание уровней (z_i), то есть расчет значений T_i с использованием уравнения тренда. Аналитическое выравнивание осуществляется по математической модели тренда. Выбор модели тренда может быть осуществлен несколькими способами. Предполагая, что тренд имеет вид полинома, анализируют цепные абсолютные приросты (первые разности уровней ряда) $\Delta_i = z_i - z_{i-1}$ и абсолютные ускорения уровней ряда (вторые разности ряда) $\Delta_{\Delta} = \Delta_i - \Delta_{i-1}$. Если примерно одинаковы Δ_i , то ряд имеет линейный тренд $\hat{T}_i = a + b \cdot t$. При вычислениях удобно моменты времени t пронумеровать так, чтобы получилось $\sum_{i=1}^n t_i = 0$. Тогда параметры линейного тренда могут быть найдены по следующим формулам

$$b = \frac{\sum_{i=1}^n z_i t_i}{\sum_{i=1}^n t_i^2}; \quad a = \bar{z}. \quad (23)$$

Если же примерно постоянны Δ_{Δ} , то для описания тенденции временного ряда следует выбрать параболу второго порядка $\hat{T}_i = a_0 + a_1 \cdot t + a_2 \cdot t^2$. Параметры такого тренда могут быть найдены по следующим формулам

$$a_0 = \frac{\sum_{i=1}^n t_i^4 \sum_{i=1}^n z_i - \sum_{i=1}^n t_i^2 \sum_{i=1}^n t_i^2 z_i}{n \sum_{i=1}^n t_i^4 - \left(\sum_{i=1}^n t_i^2 \right)^2}, \quad (24)$$

$$a_1 = \frac{\sum_{i=1}^n t_i z_i}{\sum_{i=1}^n t_i^2}, \quad (25)$$

$$a_2 = \frac{n \sum_{i=1}^n t_i^2 z_i - \sum_{i=1}^n t_i^2 \sum_{i=1}^n z_i}{n \sum_{i=1}^n t_i^4 - \left(\sum_{i=1}^n t_i^2 \right)^2}, \quad (26)$$

Методика нумерации моментов времени, так, чтобы $\sum_{i=1}^n t_i = 0$, различна для рядов имеющих четное и нечетное число наблюдений. Так, если число наблюдений нечетное, то нумерация проводится так:

Год	1997	1998	1999	2000	2001	2002	2003
t	-3	-2	-1	0	1	2	3

Если же число наблюдений четное, то нумерация соответственно:

Год	1997	1998	1999	2000	2001	2002
t	-2,5	-1,5	-0,5	0,5	1,5	2,5

После выделения трендовой компоненты \hat{T} , случайная компонента получается как разность $\hat{E}_i = z_i - \hat{T}_i$.

4. Расчет абсолютных и/или относительных ошибок. Часто рассчитывают: среднюю абсолютную процентную ошибку (Mean Absolute Percentage Error):

$$MAPE = \frac{1}{n} \sum_{i=1}^n \frac{|\hat{E}_i|}{y_i} \cdot 100\% . \quad (27)$$

Если модель подогнана с высокой точностью $MAPE < 10\%$, хорошей – $10\% < MAPE < 20\%$, удовлетворительной – $20\% < MAPE < 50\%$, неудовлетворительной – $MAPE > 50\%$.

Целесообразно пропускать значения ряда, для которых $y_i = 0$.

Средняя процентная ошибка (Mean Percentage Error) и средняя ошибка (Mean Error). Средняя процентная ошибка не определена при нулевых данных и не должна превышать 5% для хорошо подогнанной модели:

$$MPE = \frac{1}{n} \sum_{i=1}^n \frac{E_i}{y_i} 100\% . \quad (28)$$

Средняя ошибка:

$$ME = \frac{1}{n} \sum_{i=1}^n E_i . \quad (29)$$

3.2. Пример решения типовой задачи.

Задание.

1). Постройте аддитивную модель временного ряда, последовательно выделив сезонную, трендовую и случайную компоненты.

2). Используйте полученную модель для краткосрочного прогнозирования объема товарооборота в 3 квартале 2003 года.

3). Проверьте качество модели.

Имеются сведения о квартальном товарообороте торговой компании "Дельта" за 20 последних кварталов.

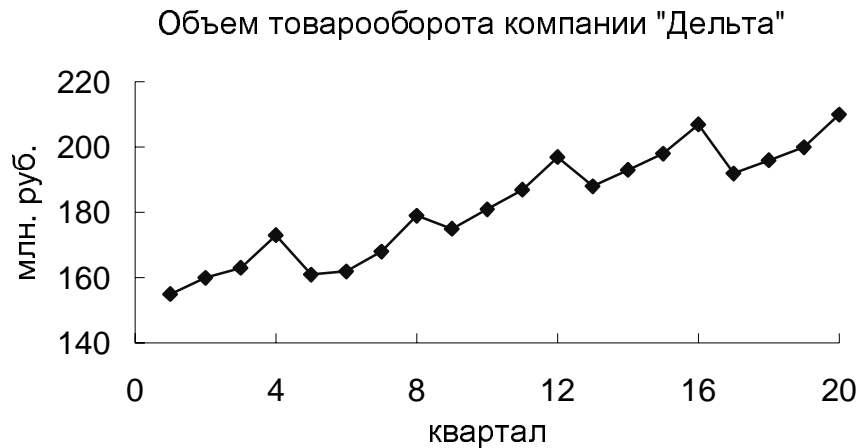
Год / Квартал	Товарообо- рот	Год / Квартал	Товарообо- рот	Год / Квартал	Товарообо- рот	Год / Квартал	Товарообо- рот
1998/1	155	1999/2	162	2000/3	187	2001/4	207
1998/2	160	1999/3	168	2000/4	197	2002/1	192
1998/3	163	1999/4	179	2001/1	188	2002/2	196
1998/4	173	2000/1	175	2001/2	193	2002/3	200
1999/1	161	2000/2	181	2001/3	198	2002/4	210

Решение.

1). Построим график ряда динамики.

На графике отчетливо видно, что товарооборот изменяется под воздействием сезонных колебаний. Заметен рост в 4 квартале каждого года (соответственно 4, 8, 12, 16 и 20 кварталы), а затем снижение товарооборота в первом квартале каждого года (соответственно 5, 9, 13 и 17 кварталы).

Построение аддитивной модели начнем с выделения сезонной компоненты временного ряда.



В нашем случае $\tau=4$ и $h=n/\tau=20/4=5$. Применим формулу (22). Для этого просуммируем уровни ряда по 1-му кварталу по всем пяти годам $155+161+175+188+192=871$ и найдем среднее значение $871/5=174,2$, и аналогично для остальных кварталов. Получим расчетную таблицу. Среднее значение по всем 20-ти наблюдениям равно 182,25. Вычитая из средних значений по кварталам 182,25 (например, для первого квартала $174,2-182,25=-8,05$), получим последнюю строку расчетной таблицы, в которой и содержатся значения сезонной компоненты S_t .

Квартал/Год	I	II	III	IV
1998	155	160	163	173
1999	161	162	168	179
2000	175	181	187	197
2001	188	193	198	207
2002	192	196	200	210
Среднее по одноименным кварталам	174,2	178,4	183,2	193,2
S_t	-8,05	-3,85	0,95	10,95

Устраним сезонную компоненту из исходных уровней ряда и получим $z_i = T_i + E_i = y_i - S_i$ в столбце 4 расчетной таблицы, которая дана ниже.

Далее рассчитаем значения $\Delta_i = z_i - z_{i-1}$ представленные в столбце 5 расчетной таблицы. Поскольку первые разности являются примерно одинаковыми (см. столбец 5), считаем, что ряд z имеет линейный тренд. Рассчитаем значения тренда T . Модель тренда имеет вид $\hat{T}_i = a + b \cdot t$. Расчет параметров уравнения проведем по формуле (23). Необходимые предварительные расчеты приведены в таблице в столбцах 6-8: столбец 7 получается путем возведения в квадрат значений столбца 6, столбец 8 равен произведению столбца 4 на столбец 6.

i	y_i	S_i	$z_i = y_i - S_i$	Δ_i	t	t^2	zt	T	E_i	E_i/y_i	$ E_i/y_i $
1	2	3	4	5	6	7	8	9	10	11	12
1	155	-8,05	163,05	-	-9,5	90,25	-1548,98	158,785	4,265	0,027516	0,027516
2	160	-3,85	163,85	0,8	-8,5	72,25	-1392,73	161,255	2,595	0,016219	0,016219
3	163	0,95	162,05	-1,8	-7,5	56,25	-1215,38	163,725	-1,675	-0,01028	0,010276
4	173	10,95	162,05	0	-6,5	42,25	-1053,33	166,195	-4,145	-0,02396	0,02396
5	161	-8,05	169,05	7	-5,5	30,25	-929,78	168,665	0,385	0,002391	0,002391
6	162	-3,85	165,85	-3,2	-4,5	20,25	-746,33	171,135	-5,285	-0,03262	0,032623
7	168	0,95	167,05	1,2	-3,5	12,25	-584,68	173,605	-6,555	-0,03902	0,039018
8	179	10,95	168,05	1	-2,5	6,25	-420,13	176,075	-8,025	-0,04483	0,044832
9	175	-8,05	183,05	15	-1,5	2,25	-274,58	178,545	4,505	0,025743	0,025743
10	181	-3,85	184,85	1,8	-0,5	0,25	-92,43	181,015	3,835	0,021188	0,021188
11	187	0,95	186,05	1,2	0,5	0,25	93,03	183,485	2,565	0,013717	0,013717
12	197	10,95	186,05	0	1,5	2,25	279,08	185,955	0,095	0,000482	0,000482
13	188	-8,05	196,05	10	2,5	6,25	490,13	188,425	7,625	0,040559	0,040559
14	193	-3,85	196,85	0,8	3,5	12,25	688,98	190,895	5,955	0,030855	0,030855
15	198	0,95	197,05	0,2	4,5	20,25	886,73	193,365	3,685	0,018611	0,018611
16	207	10,95	196,05	-1	5,5	30,25	1078,28	195,835	0,215	0,001039	0,001039
17	192	-8,05	200,05	4	6,5	42,25	1300,33	198,305	1,745	0,009089	0,009089
18	196	-3,85	199,85	-0,2	7,5	56,25	1498,88	200,775	-0,925	-0,00472	0,004719
19	200	0,95	199,05	-0,8	8,5	72,25	1691,93	203,245	-4,195	-0,02098	0,020975
20	210	10,95	199,05	0	9,5	90,25	1890,98	205,715	-6,665	-0,03174	0,031738
Сумма			3649	36	0	665	1640	3654	10	-0,00073	0,415549

Параметры уравнения линейного тренда:

$$b = \frac{1640}{665} = 2,47, \quad a = \frac{3649}{20} = 182,25.$$

Таким образом, уравнение тренда имеет вид:

$$T = 182,25 + 2,47t.$$

Подставляя в уравнение тренда последовательно соответствующие значения t , получим значения тренда для каждого уровня временного ряда (столбец 9 расчетной таблицы), например, для $t = -9,5$ получим

$$T(-9,5) = 182,25 + 2,47 \cdot (-9,5) = 158,785.$$

После выделения тренда остаток E получается как разность между z и T (разность значений в столбцах 4 и 9) и представлен в столбце 10 расчетной таблицы.

Заметим в целях самопроверки, что значения в столбце 2 для y_i должны получаться как сумма значений в столбцах 3, 9 и 10 согласно принятой аддитивной модели.

2). Полученное уравнение временного ряда может быть использовано для краткосрочного прогнозирования. Так, если необходимо спрогнозировать значение товарооборота для III квартала 2003 года, то определим $t=12,5$ (для IV квартала 2002 года $t=9,5$, I квартала 2003 года 10,5, II квартала 2003 года 11,5). Подставим значение t в уравнение тренда:

$$T(12,5)=182,25+2,47 \cdot (12,5)=213,55.$$

С учетом того, что сезонная компонента равна для III квартала 0,95, получим окончательно $213,55+0,95=214,5$.

Таким образом, в третьем квартале 2003 года прогнозируется выручка 214,5 млн. руб.

3). Проверим качество полученной модели.

Рассчитаем среднюю процентную ошибку по формуле (28). Расчет суммы приведен в столбце 11 таблицы выше. Таким образом:

$$MPE=(-0,00073) \cdot 100/20=-0,00367\%,$$

что существенно меньше 5%.

Рассчитаем среднюю абсолютную процентную ошибку по формуле (27). Расчет суммы приведен в столбце 12 таблицы выше. Таким образом:

$$MAPE=(0,415549) \cdot 100/20=2,08\%.$$

Поскольку $MAPE < 10\%$, то модель подогнана с высокой точностью.

Средняя ошибка, рассчитанная по формуле (29) с учетом суммы в столбце 10 таблицы, суть $ME=10/20=0,5$.

3.3. Задачи для самостоятельной работы.

1. Имеются данные о расстоянии, пройденном самолетами Великобритании, с янв. 1963 г. по дек. 1964 г., млн. миль.

Год/ Месяц	I	II	III	IV	V	VI	VII	VIII	IX	X	XI	XII
1963	5,327	4,678	5,584	6,762	7,062	8,144	8,566	9,268	8,463	6,694	5,348	6,080
1964	5,769	5,275	6,319	6,871	7,569	8,748	9,530	9,382	8,733	7,609	6,185	6,825

К заданию 2) – прогноз на февраль 1965г.

2. Динамика импорта КНР характеризуется поквартальными данными за 1999-2002 гг., млрд. \$

Год/ Квартал	1999				2000				2001				2002
	I	II	III	IV	I	II	III	IV	I	II	III	IV	I
Значение импорта	19,3	24,0	28,4	36,6	18,7	22,3	30,2	38,1	25,3	33,6	34,7	41,7	29,8

К заданию 2) – прогноз на 4 квартал 2002 г.

3. Динамика добычи газа в РФ характеризуется по месяцам 2002 – 2003 гг., млрд. м³:

Год/Месяц	I	II	III	IV	V	VI	VII	VIII	IX	X	XI	XII
2002	70,8	67,2	68,3	55,7	59,9	56,3	57,0	55,2	57,2	65,5	65,6	69,1
2003	71,4	64,5	67,2	62,7	58,0	52,3	51,9	51,6	53,7	61,6	67,8	69,9

К заданию 2) – прогноз на февраль 2003 г.

4. Индекс объема выпуска промышленной продукции в РФ с 2002 по 2003 гг. характеризуется следующими данными:

Год/Месяц	I	II	III	IV	V	VI	VII	VIII	IX	X	XI	XII
2002	125,01	126,41	137,67	125,95	119,12	123,49	112,3	108,83	112,09	117,61	121,89	125,6
2003	146,0	141,52	156,23	143,82	139,01	143,49	142,91	148,31	147,92	154,72	145,97	151,35

К заданию 2) – прогноз на 2 квартал 2003 г.

5. Динамика потребления мороженого в Бельгии на одного человека, в пинтах с марта 2002 г. по февраль 2004 г., характеризуется следующими данными:

Год/Месяц	III	IV	V	VI	VII	VIII	IX	X	XI	XII	I	II
2002/03	0,495	0,483	0,492	0,534	0,515	0,453	0,436	0,397	0,378	0,365	0,395	0,398
2003/04	0,438	0,427	0,490	0,480	0,579	0,552	0,497	0,451	0,428	0,416	0,393	0,435

К заданию 2) – прогноз на 4 квартал 2004 г.

6. Доля сбережений населения в РФ с 1991 по 1992 гг. имеет следующую динамику:

Год/Месяц	I	II	III	IV	V	VI	VII	VIII	IX	X	XI	XII
1991	11,28	20,93	13,65	17,78	12,99	22,57	69,23	24,89	23,59	22,86	22,46	29,65
1992	15,23	15,14	13,11	17,81	12,97	19,01	25,15	19,49	15,73	16,31	17,38	21,08

К заданию 2) – прогноз на апрель 1993г.

7. Уровень безработицы в %, на конец месяца по методологии МОТ, с 2001 по 2002 гг. имеют следующую динамику:

Год/Месяц	I	II	III	IV	V	VI	VII	VIII	IX	X	XI	XII
2001	13,38	13,22	12,58	11,95	11,37	11,23	11,1	10,97	10,96	10,95	10,94	11,06
2002	11,19	11,31	10,79	10,29	9,78	9,76	9,75	9,74	9,83	9,93	10,03	9,87

К заданию 2) – прогноз на март 2003 г.

8. Ряд динамики ВВП РФ по кварталам за 2000-2003 гг. имеет следующий вид (трлн. у.е.):

Год/Квартал	2000				2001				2002				2003
	I	II	III	IV	I	II	III	IV	I	II	III	IV	I
ВВП	295,7	350,4	376,1	432,9	461,4	561,2	650,9	700,2	654,3	716,6	777,8	820,3	746,8

К заданию 2) – прогноз на 4 квартал 2003 г.

9. Индекс цен топливной промышленности в РФ характеризуется следующими данными:

Год/Месяц	I	II	III	IV	V	VI	VII	VIII	IX	X	XI	XII
2001	29,32	30,26	30,97	31,85	35,74	38,12	41,33	45,26	52,16	59,46	64,77	65,12
2002	69,45	73,59	76,11	78,02	79,22	81,43	84,97	88,15	94,15	96,86	98,72	100,00

К заданию 2) – прогноз на 3 квартал 2003 г.

10. Экспорт машин и оборудования в РФ имеет следующую динамику, млрд. долл.

Год/ Месяц	I	II	III	IV	V	VI	VII	VIII	IX	X	XI	XII
2002	0,82	1,06	1,28	1,11	1,06	1,15	1,03	1,42	1,17	1,22	1,67	2,19
2003	0,93	1,09	1,38	1,62	1,16	1,69	1,01	1,69	1,62	1,27	1,88	1,43

К заданию 2) – прогноз на 2 квартал 2004 г.

Задача 4

"Системы эконометрических уравнений"

4.1. Краткие сведения из теории.

Пусть имеется выборка, содержащая значения переменных $Y_{1t}, \dots, Y_{nt}, X_{jt}, j=1, \dots, m, t=1, \dots, N$, где t – номер наблюдения, N – количество наблюдений.

Рассмотрим систему уравнений вида:

$$\begin{cases} \gamma_{11}Y_{1t} + \gamma_{12}Y_{2t} + \dots + \gamma_{1n}Y_{nt} + \beta_{11}X_{1t} + \beta_{12}X_{2t} + \dots + \beta_{1m}X_{mt} = \varepsilon_{1t}, \\ \gamma_{21}Y_{1t} + \gamma_{22}Y_{2t} + \dots + \gamma_{2n}Y_{nt} + \beta_{21}X_{1t} + \beta_{22}X_{2t} + \dots + \beta_{2m}X_{mt} = \varepsilon_{2t}, \\ \dots \\ \gamma_{n1}Y_{1t} + \gamma_{n2}Y_{2t} + \dots + \gamma_{nn}Y_{nt} + \beta_{n1}X_{1t} + \beta_{n2}X_{2t} + \dots + \beta_{nm}X_{mt} = \varepsilon_{nt}. \end{cases} \quad (30)$$

В (30) переменные $Y_{it}, i=1, \dots, n, t=1, \dots, N$, являются эндогенными, т.е. их значения определяются внутри системы, а переменные $X_{jt}, j=1, \dots, m, t=1, \dots, N$ являются предопределенными, т.е. включают как экзогенные, т.е. внешние по отношению к системе переменные, так и лагированные значения эндогенных переменных. Разделение переменных на экзогенные и эндогенные должно осуществляться вне модели, исходя из экономической сущности изучаемой системы. В (30) $\varepsilon_{it}, i=1, \dots, n, t=1, \dots, N$ суть величины случайных ошибок уравнений модели.

Система (30) может быть записана в матричном виде. Обозначим:

$$\mathbf{Y}_t = \begin{bmatrix} Y_{1t} \\ Y_{2t} \\ \dots \\ Y_{nt} \end{bmatrix}, \quad \mathbf{X}_t = \begin{bmatrix} X_{1t} \\ X_{2t} \\ \dots \\ X_{mt} \end{bmatrix}, \quad \boldsymbol{\varepsilon}_t = \begin{bmatrix} \varepsilon_{1t} \\ \varepsilon_{2t} \\ \dots \\ \varepsilon_{nt} \end{bmatrix},$$

$$\boldsymbol{\Gamma} = \begin{bmatrix} \gamma_{11} & \gamma_{12} & \dots & \gamma_{1n} \\ \gamma_{21} & \gamma_{22} & \dots & \gamma_{2n} \\ \dots & \dots & \dots & \dots \\ \gamma_{n1} & \gamma_{n2} & \dots & \gamma_{nn} \end{bmatrix}, \quad \mathbf{B} = \begin{bmatrix} \beta_{11} & \beta_{12} & \dots & \beta_{1m} \\ \beta_{21} & \beta_{22} & \dots & \beta_{2m} \\ \dots & \dots & \dots & \dots \\ \beta_{n1} & \beta_{n2} & \dots & \beta_{nm} \end{bmatrix}.$$

Тогда (30) может быть записана в следующем виде:

$$\boldsymbol{\Gamma}\mathbf{Y}_t + \mathbf{B}\mathbf{X}_t = \boldsymbol{\varepsilon}_t. \quad (31)$$

Система одновременных уравнений, записанная в форме (30) или в матричном виде (31) называется структурной формой эконометрической модели. Структурная форма модели лучше отражает экономические взаимосвязи между переменными, позволяет увидеть влияние изменений экзогенной переменной на значения эндогенной переменной. Коэффициенты при переменных в (31), т.е. элементы матриц Γ и \mathbf{B} называются структурными коэффициентами модели.

Система (31) может быть записана в виде зависимости эндогенных переменных от predetermined

$$\mathbf{Y}_t = \mathbf{\Pi X}_t + \mathbf{v}_t, \quad (32)$$

где $\mathbf{\Pi}' = \begin{bmatrix} \pi_{11} & \pi_{12} & \dots & \pi_{1n} \\ \pi_{21} & \pi_{22} & \dots & \pi_{2n} \\ \dots & \dots & \dots & \dots \\ \pi_{m1} & \pi_{m2} & \dots & \pi_{mn} \end{bmatrix}$, $\mathbf{v}_t = \begin{bmatrix} v_1 \\ v_2 \\ \dots \\ v_n \end{bmatrix}$, здесь $\mathbf{\Pi} = -\Gamma^{-1}\mathbf{B}$, $\mathbf{v}_t = \Gamma^{-1}\boldsymbol{\varepsilon}_t$.

Система (32) называется приведенной формой модели (30), (31). Элементы матрицы $\mathbf{\Pi}$ называются коэффициентами приведенной формы модели.

Как видно из (32) приведенная форма модели представляет собой систему линейных зависимостей эндогенных переменных от экзогенных. Коэффициенты приведенной формы модели представляют собой нелинейные функции коэффициентов структурной формы модели. Приведенная форма модели хотя и позволяет получить значения эндогенных переменных по значениям экзогенных переменных, аналитически уступает структурной форме модели, так как в ней отсутствуют оценки взаимосвязи между эндогенными переменными. Кроме того, приведенная форма модели не отражает, как правило, содержательной экономической сущности исследуемого объекта, явления, процесса.

4.2. Пример решения типовой задачи.

Рассмотрим вариант макроэкономической модели. Пусть C_t – потребление, Y_t – валовой внутренний продукт (ВВП), I_t – инвестиции, r_t – процентная ставка, G_t – государственные расходы. Структурная форма модели имеет вид:

$$\begin{cases} C_t = \alpha_0 + \alpha_1 Y_t + \alpha_2 C_{t-1} + \varepsilon_{t1}, \\ I_t = \beta_0 + \beta_1 r_t + \beta_2 (Y_t - Y_{t-1}) + \varepsilon_{t2}, \\ Y_t = C_t + I_t + G_t. \end{cases} \quad (33)$$

Задание.

1). Определить эндогенные, экзогенные, лаговые и predetermined переменные модели.

2). Записать приведенную форму модели.

Решение.

1). Эндогенными переменными являются C_t , Y_t , I_t . Экзогенные переменные - r_t и G_t . Лаговые переменные C_{t-1} и Y_{t-1} . Предопределенные переменные r_t , G_t , C_{t-1} и Y_{t-1} .

2). Приведенную форму модели (33) можно получить путем алгебраических преобразований. Подставим выражение для Y_t из третьего уравнения системы в первое и второе:

$$\begin{aligned} C_t &= \alpha_0 + \alpha_1 (C_t + I_t + G_t) + \alpha_2 C_{t-1} + \varepsilon_{t1}, \\ I_t &= \beta_0 + \beta_1 r_t + \beta_2 (C_t + I_t + G_t - Y_{t-1}) + \varepsilon_{t2}. \end{aligned} \quad (34)$$

Выражая I_t из второго из уравнений (34) и подставляя формулу для него в первое уравнение (34), получим:

$$C_t = \frac{1}{1 - \alpha_1} \left(\alpha_0 + \alpha_1 \frac{1}{1 - \beta_2} (\beta_0 + \beta_1 r_t + \beta_2 C_t + \beta_2 G_t - \beta_2 Y_{t-1} + \varepsilon_{t2}) + \alpha_1 G_t + \alpha_2 C_{t-1} + \varepsilon_{t1} \right),$$

исключая C_t из последнего уравнения, имеем:

$$\begin{aligned} C_t &= \frac{1}{1 - \alpha_1 - \beta_2} (\alpha_0 (1 - \beta_2) + \alpha_1 \beta_0 + \alpha_1 \beta_1 r_t + \alpha_1 G_t + \alpha_1 (1 - \beta_2) C_{t-1} - \beta_2 \alpha_1 Y_{t-1}) + \\ &\quad + \frac{1}{1 - \alpha_1 - \beta_2} ((1 - \beta_2) \varepsilon_{t1} + \alpha_1 \varepsilon_{t2}). \end{aligned}$$

Выражая C_t из первого из уравнений (34) и подставляя формулу для него во второе уравнение (34), получим, после исключения I_t :

$$\begin{aligned} I_t &= \frac{1}{1 - \alpha_1 - \beta_2} (\alpha_0 \beta_2 + \beta_0 (1 - \alpha_1) + \beta_1 (1 - \alpha_1) r_t + \beta_2 G_t + \alpha_2 \beta_2 C_{t-1} - \beta_2 (1 - \alpha_1) Y_{t-1}) + \\ &\quad + \frac{1}{1 - \alpha_1 - \beta_2} (\beta_2 \varepsilon_{t1} + (1 - \alpha_1) \varepsilon_{t2}). \end{aligned}$$

Подставляя полученные выражения для C_t и I_t в последнее равенство в (33) и приводя подобные, получим:

$$Y_t = \frac{1}{1 - \alpha_1 - \beta_2} (\alpha_0 + \beta_0 + \beta_1 r_t + G_t + \alpha_2 C_{t-1} - \beta_2 Y_{t-1} + \varepsilon_{t1} + \varepsilon_{t2}).$$

Таким образом, приведенная форма модели (33) имеет вид:

$$\left\{ \begin{aligned} C_t &= \frac{1}{1 - \alpha_1 - \beta_2} (\alpha_0 (1 - \beta_2) + \alpha_1 \beta_0 + \alpha_1 \beta_1 r_t + \alpha_1 G_t + \alpha_1 (1 - \beta_2) C_{t-1} - \beta_2 \alpha_1 Y_{t-1}) + \\ &\quad + \frac{1}{1 - \alpha_1 - \beta_2} ((1 - \beta_2) \varepsilon_{t1} + \alpha_1 \varepsilon_{t2}), \\ I_t &= \frac{1}{1 - \alpha_1 - \beta_2} (\alpha_0 \beta_2 + \beta_0 (1 - \alpha_1) + \beta_1 (1 - \alpha_1) r_t + \beta_2 G_t + \alpha_2 \beta_2 C_{t-1} - \beta_2 (1 - \alpha_1) Y_{t-1}) + \\ &\quad + \frac{1}{1 - \alpha_1 - \beta_2} (\beta_2 \varepsilon_{t1} + (1 - \alpha_1) \varepsilon_{t2}), \\ Y_t &= \frac{1}{1 - \alpha_1 - \beta_2} (\alpha_0 + \beta_0 + \beta_1 r_t + G_t + \alpha_2 C_{t-1} - \beta_2 Y_{t-1} + \varepsilon_{t1} + \varepsilon_{t2}). \end{aligned} \right.$$

4.3. Задачи для самостоятельной работы.

1. Конъюнктурная модель имеет вид:

$$\begin{cases} C_t = a_1 + b_1 Y_t + c_1 C_{t-1} + \varepsilon_1, \\ I_t = a_2 + b_2 r_t + c_2 I_{t-1} + \varepsilon_2, \\ r_t = a_3 + b_3 Y_t + c_3 M_t + \varepsilon_3, \\ Y_t = C_t + I_t + G_t. \end{cases}$$

где C – расходы на потребление, Y – ВВП, I – инвестиции, r – процентная ставка, M – денежная масса, G – государственные расходы, t – текущий период, $t-1$ – предыдущий период.

2. Модель мультипликатора-акселератора:

$$\begin{cases} C_t = a_1 + b_1 R_t + c_1 C_{t-1} + \varepsilon_1, \\ I_t = a_2 + b_2 (R_t - R_{t-1}) + \varepsilon_2, \\ R_t = C_t + I_t. \end{cases}$$

где C – расходы на потребление, R – доход, I – инвестиции, t – текущий период, $t-1$ – предыдущий период.

3. Модель денежного и товарного рынков:

$$\begin{cases} R_t = a_1 + b_1 Y_t + c_1 M_t + \varepsilon_1, \\ Y_t = a_2 + b_2 R_t + c_2 I_t + d_2 G_t + \varepsilon_2, \\ I_t = a_3 + b_3 R_t + \varepsilon_3. \end{cases}$$

где Y – реальный ВВП, I – внутренние инвестиции, R – процентная ставка, M – денежная масса, G – реальные государственные расходы, t – текущий период, $t-1$ – предыдущий период.

4. Макроэкономическая модель:

$$\begin{cases} C_t = a_1 + b_1 D_t + \varepsilon_1, \\ I_t = a_2 + b_2 Y_t + c_2 Y_{t-1} + \varepsilon_2, \\ Y_t = D_t + T_t, \\ D_t = C_t + I_t + G_t. \end{cases}$$

где C – расходы на потребление, Y – чистый национальный продукт, I – инвестиции, D – чистый национальный доход, T – косвенные налоги, G – государственные расходы, t – текущий период, $t-1$ – предыдущий период.

5. Модель:

$$\begin{cases} C_t = a_1 + b_1 S_t + c_1 P_t + \varepsilon_1, \\ S_t = a_2 + b_2 R_t + c_2 R_{t-1} + d_2 t + \varepsilon_2, \\ R_t = S_t + P_t. \end{cases}$$

где C – личное потребление, S – зарплата, P – прибыль, R – общий доход, t – текущий период, $t-1$ – предыдущий период.

6. Макроэкономическая модель (упрощенная версия модели Клейна):

$$\begin{cases} C_t = a_1 + b_1 Y_t + c_1 T + \varepsilon_1, \\ I_t = a_2 + b_2 Y_t + c_2 K_{t-1} + \varepsilon_2, \\ Y_t = C_t + I_t. \end{cases}$$

где C – потребление, Y – доход, I – инвестиции, T – налоги, K – запас капитала, t – текущий период, $t-1$ – предыдущий период.

7. Модель Менгеса:

$$\begin{cases} C_t = a_1 + b_1 Y_t + c_1 C_{t-1} + d_1 P_t + \varepsilon_1, \\ I_t = a_2 + b_2 Y_t + c_2 Q_t + \varepsilon_2, \\ Y_t = a_3 + b_3 Y_{t-1} + c_3 I_t + \varepsilon_3, \\ Q_t = a_4 + b_4 Q_{t-1} + c_4 R_t + \varepsilon_4. \end{cases}$$

где C – личное потребление, Y – национальный доход, I – чистые инвестиции, Q – валовая прибыль экономики, P – индекс стоимости жизни, R – объем продукции промышленности, t – текущий период, $t-1$ – предыдущий период.

8. Модель денежного рынка:

$$\begin{cases} R_t = a_1 + b_1 M_t + c_1 Y_t + \varepsilon_1, \\ Y_t = a_2 + b_2 R_t + c_2 I_t + \varepsilon_2, \end{cases}$$

где Y – ВВП, I – внутренние инвестиции, R – процентная ставка, M – денежная масса, t – текущий период, $t-1$ – предыдущий период.

9. Модифицированная модель Кейнса:

$$\begin{cases} C_t = a_1 + b_1 Y_t + \varepsilon_1, \\ I_t = a_2 + b_2 Y_t + c_2 Y_{t-1} + \varepsilon_2, \\ Y_t = C_t + I_t + G_t. \end{cases}$$

где C – расходы на потребление, Y – доход, I – инвестиции, G – государственные расходы, t – текущий период, $t-1$ – предыдущий период.

10. Одна из версий модели Кейнса:

$$\begin{cases} C_t = a_1 + b_1 Y_t + c_1 Y_{t-1} + \varepsilon_1, \\ I_t = a_2 + b_2 Y_t + \varepsilon_2, \\ Y_t = C_t + I_t + G_t. \end{cases}$$

где C – расходы на потребление, Y – доход, I – инвестиции, G – государственные расходы, t – текущий период, $t-1$ – предыдущий период.

Рекомендуемая литература

1. Айвазян С.А., Мхитарян В.С. Прикладная статистика и основы эконометрики. – М.: ЮНИТИ, 1998.
2. Арженовский С.В. Системы одновременных уравнений. Текст лекций. – Ростов-на-Дону: РГЭУ, 2002.

3. Арженовский С.В., Федосова О.Н. Эконометрика. – Ростов н/Д.: РГЭУ, 2002.
4. Джонстон Дж. Эконометрические методы. – М.: Статистика, 1980.
5. Доугерти К. Введение в эконометрику. – М.: ИНФРА-М, 2001.
6. Дрейпер Н., Смит Г. Прикладной регрессионный анализ. – М.: Финансы и статистика, 1986.
7. Катышев П.К., Пересецкий А.А. Сборник задач к начальному курсу эконометрики. – М.: Дело, 1999.
8. Кейн Э. Экономическая статистика и эконометрика. Введение в количественный анализ. Пер. с англ. – М.: Статистика, 1977.
9. Кремер Н., Путко Б. Эконометрика. – М.: ЮНИТИ-ДАНА, 2002.
10. Магнус Я.Р., Катышев П.К., Пересецкий А.А. Эконометрика. Начальный курс. – М.: Дело, 2000.
11. Ниворожкина Л.И., Морозова З.А., Житников И.В., Герасимова И.А. Основы статистики с элементами теории вероятностей. – Ростов-на-Дону: Феникс, 1999.
12. Ниворожкина Л.И., Князевский В.С., Житников И.В., Федосова О.Н. Эконометрика: методические рекомендации по изучению курса. – Ростов-на-Дону: РГЭУ, 2001.
13. Практикум по эконометрике/Под ред. И.И.Елисеевой. – М.: Финансы и статистика, 2001.
14. Фестер Э., Ренц Б. Методы корреляционного и регрессионного анализа: руководство для экономистов. – М.: Финансы и статистика, 1983.
15. Эконометрика/Под ред. И.И. Елисеевой. – М.: Финансы и статистика, 2001.

Статистические таблицы

Двухсторонние квантили t - распределения Стьюдента

m	ε						
	0,10	0,05	0,025	0,020	0,010	0,005	0,001
1	6,314	12,706	25,452	31,821	63,657	127,3	636,6
2	2,920	4,303	6,205	6,965	9,925	14,089	31,598
3	2,353	3,182	4,177	4,541	5,841	7,453	12,941
4	2,132	2,776	3,495	3,747	4,604	5,597	8,610
5	2,015	2,571	3,163	3,365	4,032	4,773	6,859
6	1,943	2,447	2,969	3,143	3,707	4,317	5,959
7	1,895	2,365	2,841	2,998	3,499	4,029	5,405
8	1,860	2,306	2,752	2,896	3,355	3,833	5,041
9	1,833	2,262	2,685	2,821	3,250	3,690	4,781
10	1,812	2,228	2,634	2,764	3,169	3,581	4,587
12	1,782	2,179	2,560	2,681	3,055	3,428	4,318
14	1,761	2,145	2,510	2,624	2,977	3,326	4,140
16	1,746	2,120	2,473	2,583	2,921	3,252	4,015
18	1,734	2,101	2,445	2,552	2,878	3,193	3,922
20	1,725	2,086	2,423	2,528	2,845	3,153	3,849
22	1,717	2,074	2,405	2,508	2,819	3,119	3,792
24	1,711	2,064	2,391	2,492	2,797	3,092	3,745
26	1,706	2,056	2,379	2,479	2,779	3,067	3,707
28	1,701	2,048	2,369	2,467	2,763	3,047	3,674
30	1,697	2,042	2,360	2,457	2,750	3,030	3,646
∞	1,645	1,960	2,241	2,326	2,576	2,807	3,291

m - число степеней свободы

95% квантили распределения Фишера $F(n_1, n_2)$

n_2	n_1																		
	1	2	3	4	5	6	7	8	9	10	12	15	20	24	30	40	60	120	∞
1	161	200	216	225	230	234	237	239	241	242	244	246	248	249	250	251	252	253	254
2	18,5	19,0	19,2	19,2	19,3	19,3	19,4	19,4	19,4	19,4	19,4	19,4	19,4	19,5	19,5	19,5	19,5	19,5	19,5
3	10,1	9,55	9,28	9,12	9,01	8,94	8,89	8,85	8,81	8,79	8,74	8,70	8,66	8,64	8,62	8,59	8,57	8,55	8,53
4	7,71	6,94	6,59	6,39	6,26	6,16	6,09	6,04	6,00	5,96	5,91	5,86	5,80	5,77	5,75	5,72	5,69	5,66	5,63
5	6,61	5,79	5,41	5,19	5,05	4,95	4,88	4,82	4,77	4,74	4,68	4,62	4,56	4,53	4,50	4,46	4,43	4,40	4,37
6	5,99	5,14	4,76	4,53	4,39	4,28	4,21	4,15	4,10	4,06	4,00	3,94	3,87	3,84	3,81	3,77	3,74	3,70	3,67
7	5,59	4,74	4,35	4,12	3,97	3,87	3,79	3,73	3,68	3,64	3,57	3,51	3,44	3,41	3,38	3,34	3,30	3,27	3,23
8	5,32	4,46	4,07	3,84	3,69	3,58	3,50	3,44	3,39	3,35	3,28	3,22	3,15	3,12	3,08	3,04	3,01	2,97	2,93
9	5,12	4,26	3,86	3,63	3,48	3,37	3,29	3,23	3,18	3,14	3,07	3,01	2,94	2,90	2,86	2,83	2,79	2,75	2,71
10	4,96	4,10	3,71	3,48	3,33	3,22	3,14	3,07	3,02	2,98	2,91	2,85	2,77	2,74	2,70	2,66	2,62	2,58	2,54
11	4,84	3,98	3,59	3,36	3,20	3,09	3,01	2,95	2,90	2,85	2,79	2,72	2,65	2,61	2,57	2,53	2,49	2,45	2,40
12	4,75	3,89	3,49	3,26	3,11	3,00	2,91	2,85	2,80	2,75	2,69	2,62	2,54	2,51	2,47	2,43	2,38	2,34	2,30
13	4,67	3,81	3,41	3,18	3,03	2,92	2,83	2,77	2,71	2,67	2,60	2,53	2,46	2,42	2,38	2,34	2,30	2,25	2,21
14	4,60	3,74	3,34	3,11	2,96	2,85	2,76	2,70	2,65	2,60	2,53	2,46	2,39	2,35	3,31	2,27	2,22	2,18	2,13
15	4,54	3,68	3,29	3,06	2,90	2,79	2,71	2,64	2,59	2,54	2,48	2,40	2,33	2,29	2,25	2,20	2,16	2,11	2,07
16	4,49	3,53	3,24	3,01	2,85	2,74	2,66	2,59	2,54	2,49	2,42	2,35	2,28	2,24	2,19	2,15	2,11	2,06	2,01
18	4,41	3,55	3,16	2,93	2,77	2,66	2,58	2,51	2,46	2,41	2,34	2,27	2,19	2,15	2,11	2,06	2,02	1,97	1,92
20	4,35	3,49	3,10	2,87	2,71	2,60	2,51	2,45	2,39	2,35	2,28	2,20	2,12	2,08	2,04	1,99	1,95	1,90	1,84
22	4,30	3,44	3,05	2,82	2,66	2,55	2,46	2,40	2,34	2,30	2,23	2,15	2,07	2,03	1,98	1,94	1,89	1,84	1,78
24	4,26	3,40	3,01	2,78	2,62	2,51	2,42	2,36	2,30	2,25	2,18	2,11	2,03	1,98	1,94	1,89	1,84	1,79	1,73
25	4,24	3,39	2,99	2,76	2,60	2,49	2,40	2,34	2,28	2,24	2,16	2,09	2,01	1,96	1,92	1,87	1,82	1,77	1,71
30	4,17	3,32	2,92	2,69	2,53	2,42	2,33	2,27	2,21	2,16	2,09	2,01	1,93	1,89	1,84	1,79	1,74	1,68	1,62
40	4,08	3,23	2,84	2,61	2,45	2,34	2,25	2,18	2,12	2,08	2,00	1,92	1,84	1,79	1,74	1,69	1,64	1,58	1,51
60	4,00	3,15	2,76	2,53	2,37	2,25	2,17	2,10	2,04	1,99	1,92	1,84	1,75	1,70	1,65	1,59	1,53	1,47	1,39
120	3,92	3,07	2,68	2,45	2,29	2,18	2,09	2,02	1,96	1,91	1,83	1,75	1,66	1,61	1,55	1,50	1,43	1,35	1,25
∞	3,84	3,00	2,60	2,37	2,21	2,10	2,01	1,94	1,88	1,83	1,75	1,67	1,57	1,52	1,46	1,39	1,32	1,22	1,00

n_1 – число степеней свободы числителя, n_2 – число степеней свободы знаменателя

Эконометрика
Методические указания и задания к контрольной работе

Ответственная за выпуск
Начальник РИО РГЭУ
Редактирование, корректура и оригинал-макет С.В.Арженовский

В.Е. Смейле

Изд. №189/5968	Подписано к печати 08.08.03	Бумага офсетная
Печать офсетная	Формат 60/84/16	Объем 2,5 уч.- изд.л.
Заказ № 183	Тираж 100 экз.	"С" 189
