

МИНИСТЕРСТВО НАУКИ И ВЫСШЕГО ОБРАЗОВАНИЯ РОССИЙСКОЙ ФЕДЕРАЦИИ
ФЕДЕРАЛЬНОЕ ГОСУДАРСТВЕННОЕ АВТОНОМНОЕ ОБРАЗОВАТЕЛЬНОЕ
УЧРЕЖДЕНИЕ ВЫСШЕГО ОБРАЗОВАНИЯ
«СЕВЕРО-КАВКАЗСКИЙ ФЕДЕРАЛЬНЫЙ УНИВЕРСИТЕТ»

МЕТОДИЧЕСКИЕ УКАЗАНИЯ

по выполнению лабораторных работ
по дисциплине «Эконометрика»
для студентов специальности 38.03.01 Экономика
направленность (профиль) «Экономика организаций»

Ставрополь, 2026

Методические указания по дисциплине «Эконометрика» предназначены для выполнения лабораторных заданий по изучаемой дисциплине.

Выполнение заданий направлено на формирование у студентов знаний о статистических и эконометрических методах составления и оценки моделей, поиска и анализа взаимосвязи между изучаемыми признаками, расчетах показателей и коэффициентов, решения основных аналитических задач, а также формирования навыков работы с данными, их сбор, классификация, группировка и анализ, в том числе применением информационных технологий.

Содержание

Цели и задачи изучения дисциплины	4
Компетенции обучающегося, формируемые в результате освоения дисциплины	4
<i>Тема 1 «Предмет и метод эконометрики»</i>	
Лабораторная работа 1. Предмет и метод эконометрики: типы данных и типы зависимостей	6
Лабораторная работа 2. Предмет и метод эконометрики: примеры баз данных и методы работы с ними	8
<i>Тема 2. «Базовые понятия эконометрики»</i>	
Лабораторная работа 3. Базовые понятия эконометрики: основные статистические пакеты, применяемые для оценки эконометрических моделей	10
Лабораторная работа 4. Базовые понятия эконометрики: подготовка данных к эконометрическим исследованиям	12
<i>Тема 3. «Элементы математической статистики»</i>	
Лабораторная работа 5. Элементы математической статистики: основные предпосылки эконометрического моделирования	14
Лабораторная работа 6. Элементы математической статистики: проверка статистических гипотез	17
<i>Тема 4. «Нормальное распределение и связанные с ним распределения»</i>	
Лабораторная работа 7. Нормальное распределение и связанные с ним распределения: распределение хи-квадрат, Стьюдента, Фишера	19
Лабораторная работа 8. Нормальное распределение и связанные с ним распределения: методы визуализации статистической информации.	21
<i>Тема 5. «Линейная регрессия с одной объясняющей переменной»</i>	
Лабораторная работа 9. Линейная регрессия с одной объясняющей переменной: метод наименьших квадратов для нахождения оценок коэффициентов парной регрессии	23
Лабораторная работа 10. Линейная регрессия с одной объясняющей переменной: линейная и нелинейная зависимости	24
<i>Тема 6. «Линейная регрессия с несколькими объясняющими переменными»</i>	
Лабораторная работа 11. Линейная регрессия с несколькими объясняющими переменными: МНК-оценки коэффициентов множественной регрессии	26
Лабораторная работа 12. Линейная регрессия с несколькими объясняющими переменными: показатели качества подгонки множественной регрессии	28
<i>Тема 7 «Выбор функциональной формы модели»</i>	
Лабораторная работа 13. Выбор функциональной формы модели: линейная модель	30
Лабораторная работа 14. Выбор функциональной формы модели: модели с постоянным темпом роста и эластичностью	32
<i>Тема 8. «Проблема пропущенных и избыточных факторов».</i>	
Лабораторная работа 15. Проблема пропущенных и избыточных факторов: мультиколлинеарность данных	34
Лабораторная работа 16. Проблема пропущенных и избыточных факторов: гетероскедастичность	36
<i>Тема 9. «Временные ряды и прогнозирование»</i>	
Лабораторная работа 17. Временные ряды и прогнозирование: типы моделей	37
Лабораторная работа 18. Временные ряды и прогнозирование: прогнозирование во временных рядах	39
Список рекомендуемой литературы	41

Цели и задачи изучения дисциплины

Целью освоения дисциплины является формирование части общепрофессиональных (ОПК-2, ОПК-5) компетенций будущего экономиста по специальности 38.03.01 Экономика направленность (профиль) «Экономика организаций».

Основными задачами изучения дисциплины «Эконометрика» являются:

- изучение принципов описания социальных, экономических, финансовых объектов, процессов и явлений языком математических моделей со случайной компонентой;
- изучение основных этапов эконометрического моделирования;
- формирование навыков подготовки статистической информации, предназначенной для построения эконометрических моделей и работы в основных статистических пакетах;
- формирование умений проверять статистические гипотезы и строить точечные и интервальные прогнозы;
- формирование навыков диагностики моделей, оценки регрессионных моделей;
- развитие навыков интерпретации основных результатов оценки моделей.

Компетенции обучающегося, формируемые в результате освоения дисциплины

В результате освоения дисциплины обучающийся должен:

ОПК-2: способен осуществлять сбор, обработку и статистический анализ данных, необходимых для решения поставленных задач

ОПК-5: способен использовать современные информационные технологии и программные средства при решении профессиональных задач

Тема 1. «Предмет и метод эконометрики»

Лабораторная работа 1. Предмет и метод эконометрики: типы данных и типы зависимостей

Цель работы: изучить основные понятия, сущность и методы эконометрики

Теоретическая часть:

Становление эконометрики как научной дисциплины представляет значительный интерес с точки зрения как определения объектов исследования, так и формирования набора методов. Сам термин «эконометрика» сформировался из двух частей: «эконо-» – от «экономика» и «-метрика» – от «измерение». Поэтому статистический анализ экономических данных называется эконометрикой, что буквально означает «наука об экономических измерениях». Существуют различные определения эконометрики. Их можно сформулировать на основе следующих задач, стоящих перед дисциплиной:

- унифицирование теоретико-количественных и эмпирико-количественных подходов к решению экономических проблем;
- развитие методов конструктивного точного анализа этих проблем, подобных существующим в естественных науках.

Задание:

Построить график зависимости Y от X . В таблице 1 представлены пространственные данные зависимости Y – розничный товарооборот от X – количество продавцов, при условии, что в магазине имеется очередь.

Таблица 1 – Исходные данные

№ магазина	X (чел)	Y (тыс. руб.)
1	1	4
2	3	6
3	2	7
4	4	10
5	5	?

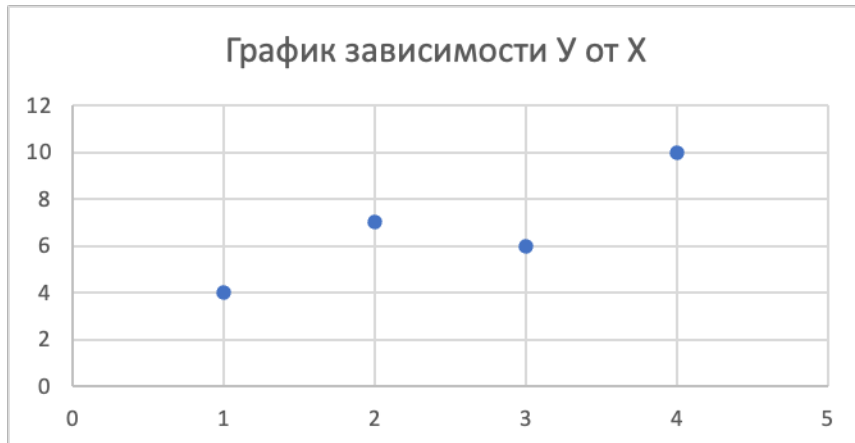
Y — зависимая переменная, которая изучается в соответствии с миссией, видением, стратегическими и оперативными целями и задачами деятельности предприятия, а также с учетом актуальности и важности проведения исследования.

Примечание. Одним из критериев правильно выбранной зависимой переменной является материальное вознаграждение руководством данного исследования. Различия в розничном товарообороте зависят от количества продавцов и неучтенных факторов. К неучтенным факторам можно отнести: наличие сертификатов соответствия на товар и систему качества, безопасность товара, насыщенность товаров по ассортименту, эффективность использования торгового оборудования, квалификацию продавцов, культуру обслуживания, стиль руководства

и эффект взаимодействия продавцов между собой, мотивацию, рабочую среду, изменение спроса в зависимости от места расположения магазина.

Ход решения задания:

1. Скопируйте табл. 1.1 и вставьте в лист Excel
2. В листе Excel выделите значения X и Y, мастер диаграмм, точечная, готово.



Визуально определим коэффициенты линейной модели $Y = a_0 + a_1X + e$ и точечный прогноз при ожидаемом значении количества продавцов, равного 5 чел., где Y — фактические значения зависимой величины; a_0 , a_1 — коэффициенты линейной модели; X — фактор; e — остатки модели, $e = Y - (a_0 + a_1X) = Y - Y_p$; $Y_p = a_0 + a_1X$ — расчетные значения Y . Возьмем линию из набора инструментов “рисование” и проведем ее по рис. 1.1 так, чтобы линия проходила: через точку $\bar{X}$ и $\bar{Y}$; линия прошла как можно ближе к исходным данным y .

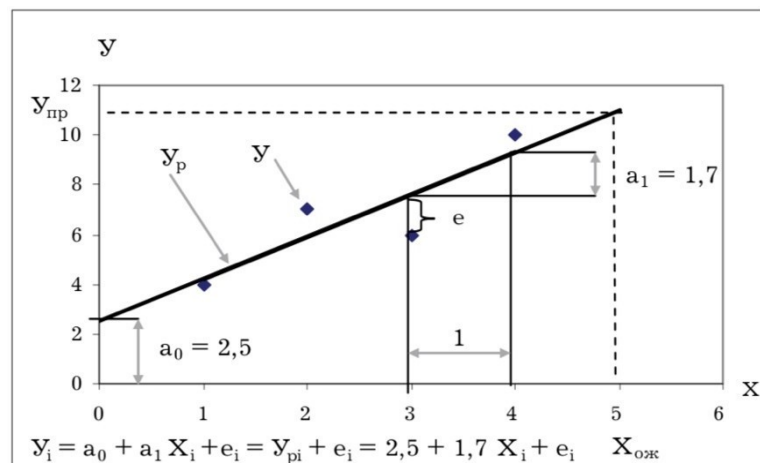


Рисунок 1 – Построении тенденции зависимости Y от X

a_0 равно расчетному значению Y в точке $X=0$,

a_1 равно приросту расчетного значения при изменении X на единицу.

Прогнозное значение $Y_{пр}$ равно расчетному значению Y при $X=5$. Таким образом, для выборочной совокупности магазинов зависимость розничного товарооборота от количества продавцов можно представить в виде однофакторного линейного уравнения регрессии:

$$Y_i = 2,5 + 1,7X_i + e_i$$

Средняя производительность труда продавцов составляет $a_1 = 1,7$ тыс. руб. Получается, что если в магазине не будет продавцов, то товарооборот составит $a_0 = 2,5$ тыс. руб. Этот вывод является абсурдным и объясняется тем, что:

-во-первых, расчетное значение Y при $X = 0$, равное a_0 , является прогнозным значением “назад” и не входит в нашу задачу;

-во-вторых, для исходных данных, изменяющихся от одного продавца и больше, модель хорошо отражает имеющуюся зависимость Y от X .

-в-третьих, для получения прогноза Y при $X = 0$ необходима другая модель, учитывающая особенность объекта моделирования, в частности необходимо вести ограничения $a_0 = 0$. При проведении расчетов в Excel с помощью функции “Линейн” имеется возможность ввести ограничение на $a_0 = 0$. Если в новом магазине ожидается 5 продавцов, то среднее прогнозное значение товарооборота будет равно 11 тыс. руб. Однако коэффициенты a_0 , a_1 , определенные по выборке, равной пяти магазинам, будут отличаться от истинных параметров a_0 , a_1 , которые можно определить по всей генеральной совокупности.

Содержание отчета о проделанной работе:

Отчет по лабораторной работе представляется в виде документа Word. В состав документа входят: название работы, цель работы, копии экрана, иллюстрирующие выполнение лабораторной работы, выводы по работе.

Лабораторная работа 2. Предмет и метод эконометрики:

примеры баз данных и методы работы с ними

Цель: изучить основные понятия, сущность и методы эконометрики

Теоретическая часть:

Объектом эконометрики выступают экономические процессы и явления, происходящие в экономической системе общества. Предметом эконометрики является количественная оценка взаимосвязи между случайными событиями, признаками, факторами, переменными экономических объектов, проверка теоретических моделей реальных экономических процессов для получения прогнозов, имитации и анализа экономической деятельности. С этой целью исследуются массовые экономические явления и процессы. Эконометрика через математические и статистические методы анализирует экономические закономерности, доказанные экономической теорией. Целью эконометрики является оценка точечных и интервальных прогнозов деятельности генеральной совокупности объектов экономической системы на основании расчетов по данным выборочной совокупности при реализации случайного процесса. К основным задачам эконометрики можно отнести следующие:

- построение эконометрических моделей, т. е. представление экономических моделей в математической форме, удобной для проведения эмпирического анализа. Данную проблему принято называть проблемой спецификации.

- оценка параметров построенной модели, делающих выбранную модель наиболее адекватной реальным данным. Это так называемый этап параметризации.

- проверка качества найденных параметров модели и самой модели в целом. Иногда этот этап анализа называют этапом верификации.

- использование построенных моделей для объяснения поведения исследуемых экономических показателей, прогнозирования и предсказания, а также для осмысленного проведения экономической политики.

Задание: Торговое предприятие имеет сеть, состоящую из 12 магазинов, информация о деятельности которых представлена в таблице 2.

Таблица 2 – Исходные данные

№	Годовой товарооборот, млн. руб.	Торговая площадь, тыс. кв.м.	Среднее число посетителей в день, тыс. чел.
1	19	0,25	8
2	38	0,31	10
3	41	0,55	9
4	41	0,48	11
5	56	0,78	9
6	69	0,98	7
7	75	0,94	12
8	89	1,21	10
9	91	1,29	9
10	91	1,12	14
11	99	1,29	12
12	109	1,49	13

Требуется построить диаграммы рассеяния годового товарооборота (Y) в зависимости от торговой площади (X1) и среднего числа посетителей в день (X2). Определить форму связи между результирующим показателем (Y) и каждым из факторов (X1 и X2). С помощью инструмента Excel «Вставка-диаграммы-точечная» строим диаграммы рассеяния в соответствии с требуемым заданием.

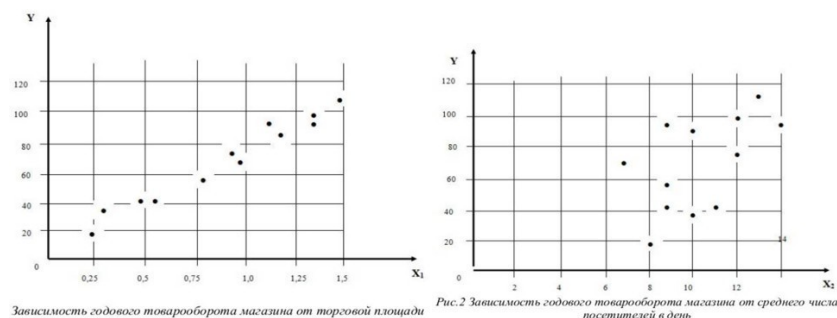


Рисунок 2 – Диаграмма рассеяния

Из диаграмм рассеяния можно сделать вывод, что с ростом торговой площади магазина и среднего числа посетителей растет его годовой товарооборот. При этом форма связи – линейная.

Задание:

Используя рисунок, визуальнo приблизительно определите индекс детерминации.

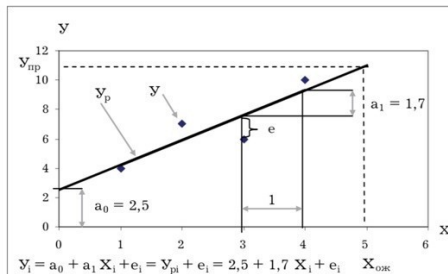


Рисунок 3 – Графическое изображение тренда

Ход решения задания:

Индекс детерминации показывает, насколько изменение результативного признака y зависит от исследуемого фактора x и какова доля влияния других факторов. Считается, что если $R^2 > 0,5$, то модель пригодна для практического применения, так как более половины общей вариации результативного признака объясняется воздействием факторного признака и рассчитывается по формуле:

$$R^2 = \frac{\sum (y_x - \bar{y})^2}{\sum (y - \bar{y})^2}$$

Для рисунка изменения Y при $X = 1$, $X = 2$ и изменения Y при $X = 3$, $X = 4$ совпадают с выбранной тенденцией. Следовательно, из трех изменений Y в двух случаях изменения совпадают с выбранной тенденцией, тогда индекс детерминации определяется по формуле $R^2 = \frac{2}{3}$

Вывод: 66 % исходных данных имеют линейную тенденцию.

Содержание отчета о проделанной работе

Отчет по лабораторной работе представляется в виде документа Word. В состав документа входят: название работы, цель работы, копии экрана, иллюстрирующие выполнение лабораторной работы, выводы по работе.

Тема 2. «Базовые понятия эконометрики»

Лабораторная работа 3: Базовые понятия эконометрики: основные статистические пакеты, применяемые для оценки эконометрических моделей

Цель: изучение основных понятий и составляющих эконометрического моделирования

Теоретическая часть:

В любой эконометрической модели зависимая переменная разбивается на две части: объясненную и случайную. Задача эконометрического моделирования может быть

сформулирована следующим образом: на основании экспериментальных данных определить объясненную часть и, рассматривая случайную составляющую как случайную величину, получить оценки параметров ее распределения. Эконометрическая модель имеет следующий вид: $Y = f(X) + e$, где: Y — наблюдаемое значение зависимой переменной; $f(X)$ — объясненная часть, зависящая от значений объясняющих переменных; e — случайная составляющая.

Задание:

Исследуется зависимость между среднемесячными доходами X на семью (в тыс. у.е.) и расходами Y на покупку кондитерских изделий (в у.е.) Данные опроса представлены в таблице 1:

Таблица 3 – Исходные данные

X	4,8	3,8	5,4	4,2	3,4	4,6	3,4	4,8	5,0	3,8	5,2	4,0	3,8	4,6	4,4
Y	75	68	78	71	64	73	66	75	75	65	77	69	67	72	70

Построить корреляционное поле и сделать предварительный вывод о форме зависимости случайных величин? Корреляционное поле – это график, на котором изображены x и y в виде точек с соответствующими абсциссами x и ординатами y . Анализ рисунка позволяет сделать вывод о наличии сильной линейной статистической связи между среднемесячными доходами на семью и расходами на покупку кондитерских изделий. График показывает что связь является прямой, т.е. с ростом переменной x наблюдается увеличение y .

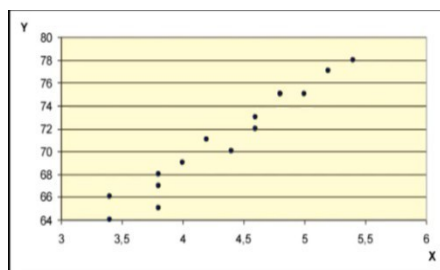
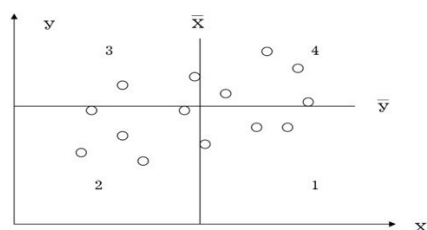


Рисунок 4 – Диаграмма рассеяния

Задание: определите приблизительно индекс детерминации при условии что зависимость Y от X представлена как облако значений на рисунке



Ход решения задания:

С помощью средних значений X и Y разделим все облако значений на четыре квадрата. Определим, какая доля данных находится во втором и четвертом квадратах. Тогда индекс детерминации будет равен $9/14=0,64$.

Содержание отчета о проделанной работе

Отчет по лабораторной работе представляется в виде документа Word. В состав документа входят: название работы, цель работы, копии экрана, иллюстрирующие выполнение лабораторной работы, выводы по работе.

Лабораторная работа 4: Базовые понятия эконометрики: подготовка данных к эконометрическим исследованиям

Цель: изучение основных понятий и составляющих эконометрического моделирования

Теоретическая часть:

Выделяют шесть основных этапов эконометрического моделирования:

1. Постановочный. На этом этапе формируется цель исследования и набор участвующих в модели экономических переменных.
2. Априорный. На данном этапе проводится анализ сущности изучаемого объекта, формирование и формализация априорной (известной до начала моделирования) информации.
3. Параметризация. На этом этапе осуществляется моделирование, то есть выбор общего вида функции $f(X)$ в эконометрической модели, выявление входящих в нее связей, то есть происходит непосредственно моделирование.
4. Информационный. На этом этапе осуществляется сбор и анализ необходимой статистической информации — наблюдаемых значений экономических переменных: $(x_{i1}, x_{i2}, \dots, x_{ip}, y_{i1}, y_{i2}, \dots, y_{iq})$, где $i=1, 2, \dots, n$. Статистическая информация может быть получена в условиях активного или пассивного эксперимента.
5. Идентификация модели. На этом этапе осуществляется статистический анализ модели и оценка ее параметров.
6. Верификация модели. На данном этапе проводится проверка адекватности модели.
7. Интерпретация результатов моделирования.

Задание: используя базу Росстата загрузите в пакет Excel выбранные данные.

Ход решения задания: Чтобы загрузить можно выбрать функцию Файл---Открыть в главном окне программы Excel, а затем выбрать нужный файл. В этом пакете для построения описательных статистик необходимо воспользоваться встроенными функциями (СЧЕТ, МИН, МАКС, ДИСП, СРЗНАЧ и т.д.). В пакете Excel на вкладке Вставка в разделе Диаграммы есть класс Гистограмма, однако он визуализирует вовсе не распределение переменной, а отдельные ее значения в виде столбцов, откладывая по оси абсцисс номер наблюдения. Чтобы построить

настоящую гистограмму, необходимо воспользоваться надстройкой «Пакет анализа». Нажмите на кнопку «Анализ данных» на вкладке Данные и выберите пункт Гистограмма. В появившемся окне в поле «Входной интервал» укажите столбец необходимой переменной. Если вы выделили столбец вместе с первой строкой, содержащей имя переменной, то поставьте галочку напротив «Метки». По умолчанию результаты разбиения сохраняются на отдельном листе. Выделите два полученных столбца и только теперь выберите пункт Вставка — Диаграммы — Гистограмма.

Задание: определите основные виды тенденций, которые имеются во временных рядах и остатках экономических показателей, дайте их характеристику:

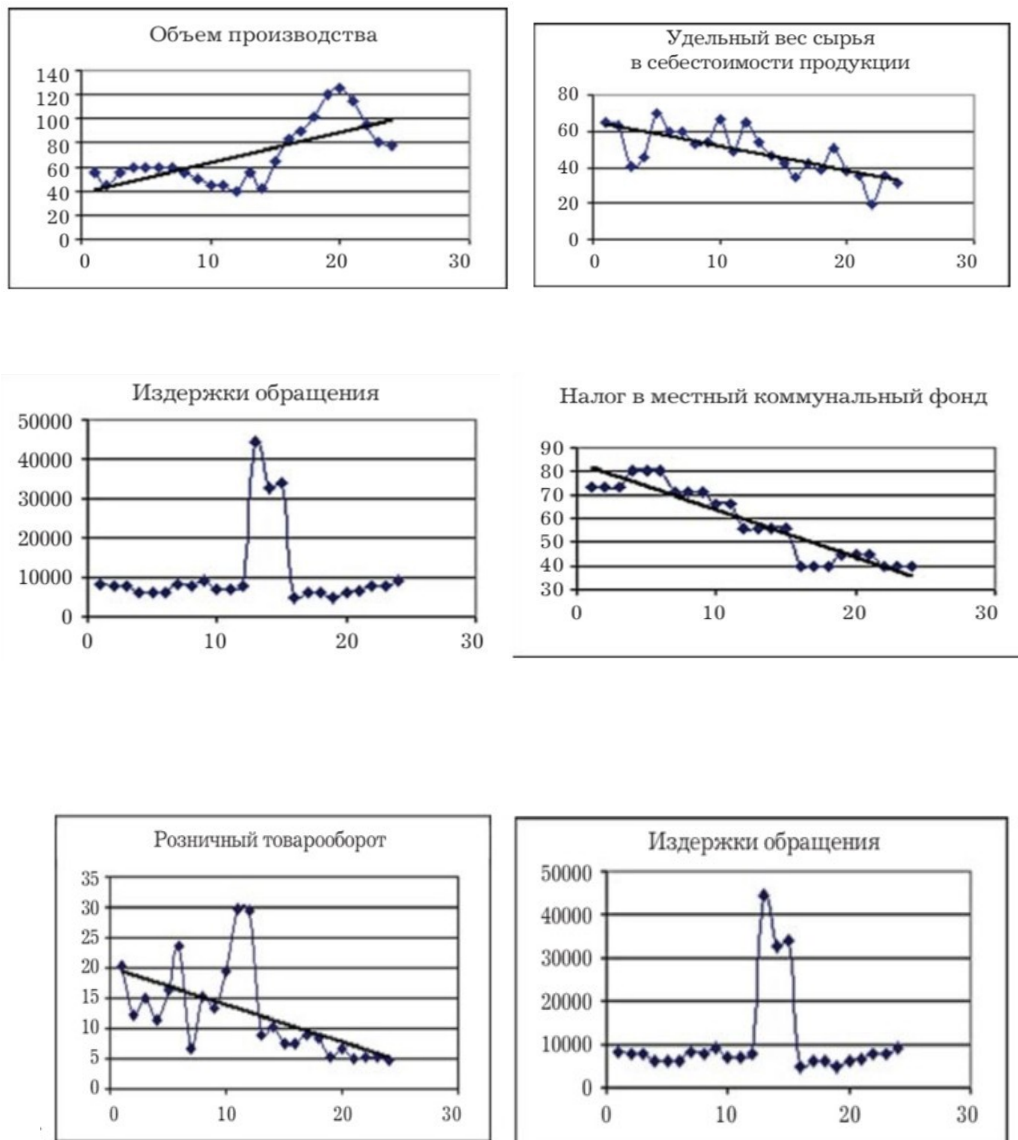


Рисунок 5 – Виды тенденций

Ход решения задания:

- 1) положительная линейная тенденция
- 2) отрицательная линейная тенденция
- 3) остатки содержат выбросы
- 4) остатки однородны гомоскедастичны

Содержание отчета о проделанной работе

Отчет по лабораторной работе представляется в виде документа Word. В состав документа входят: название работы, цель работы, копии экрана, иллюстрирующие выполнение лабораторной работы
-выводы по работе.

Тема 3. «Элементы математической статистики»

Лабораторная работа 5. Элементы математической статистики: основные предпосылки эконометрического моделирования

Цель: изучение методов и инструментов математической статистики, применяемых в эконометрике

Теоретическая часть:

Методы оценивания и проверки гипотез опираются на вероятностные модели происхождения данных. Эти модели делятся на параметрические и непараметрические. В параметрических моделях предполагается, что характеристику изучаемых объектов описывают посредством распределений, зависящих от (одного или нескольких) числовых параметров. Непараметрические модели связаны со спецификацией параметрического семейства для распределения изучаемых характеристик. В математической статистике оценивают параметры и функции от них, представляющие важные характеристики распределений (например, математическое ожидание, стандартное отклонение, квантили и др), плотности и функции распределения и другое. Выборочная совокупность. Пусть из генеральной совокупности извлекается выборка объема n . Выборочной средней называется арифметическое наблюдаемых значений случайной величины в выборке, т.е. $\bar{x} = \frac{1}{n} \sum x_i$ Выборочной дисперсией (вариацией) называется среднее арифметическое квадратов отклонений наблюдаемых значений случайной величины в выборке от их среднего значения, т.е. $ar(x) = \frac{1}{n} \sum (x_i - \bar{x})^2$ Значения $var(x)$ являются числовыми характеристиками выборочной совокупности. Для разных выборок, взятых из одной и той же генеральной совокупности, выборочные средние и дисперсии будут различны, то есть выборочные характеристики являются случайными величинами.

II Генеральная совокупность. Генеральной средней называется математическое ожидание случайной величины X , т.е. $\mu_x = M(X)$. Она характеризует среднее значение случайной величины в генеральной совокупности. Генеральной дисперсией называется математическое ожидание квадрата отклонений случайной величины X относительно ее средней, то есть:

Задание: вычислить выборочные характеристики по данным наблюдения:

Таблица 4 – Исходные данные

№	1	2	3	4	5
X	2	6	10	14	18

Исходные данные и расчетные показатели представлены в виде двух расчетных таблиц (два способа расчета $\bar{x}$, $\text{var}(x)$):

Таблица 5 – Расчетные данные

№	x	$(x - \bar{x})^2$
1	2	64
2	6	16
3	10	0
4	14	16
5	18	64
Итого	50	160
Среднее	10	32
	$\bar{x}$	$\text{var}(x)$
$\text{var}(x) = 160/5 = 32$		

№	x	x^2
1	2	4
2	6	36
3	10	100
4	14	196
5	18	324
Итого	50	660
Среднее	10	132
	$\bar{x}$	$\overline{x^2}$
$\text{var}(x) = \overline{x^2} - (\bar{x})^2 = 132 - 100 = 32$		

Ход решения задания:

Для вычисления выборочной средней и выборочной дисперсии в Excel имеются функции:

$\bar{x} = \text{СРЗНАЧ}(\text{массивx})$, $\text{var}(x) = \text{ДИСПР}(\text{массивx})$

Задание:

Определите параметры линейной функции применяя метод аналитического выравнивания.

Таблица 6 – Исходные данные

Период	Прибыль, млн. руб.
Январь	20
Февраль	24
Март	17
Апрель	18
Май	21
Июнь	25
Июль	24
Август	26
Сентябрь	22
Октябрь	24
Ноябрь	27
Декабрь	30

Итого	
-------	--

Ход решения задания:

Метод аналитического выравнивания предполагает определение функции изменения показателя от единственного фактора – времени. Подбор функции осуществляется на основе качественного анализа, который включает оценку возможности экономической интерпретации выбранной функции. Наиболее простой ситуацией является постоянный абсолютный прирост показателя. Тогда выбирается линейная функция. Следующим вариантом является постоянство абсолютных разностей, что можно выразить следующей формулой: $(Y_i - Y_{i-1}) - (Y_{i-1} - Y_{i-2}) = \text{const}$. В данном случае подойдет параболическая функция. Может встретиться постоянный цепной темп роста, что выражается через экспоненциальную функцию $Y = a e^{a + a^t}$. В качестве функции может быть гипербола или степенная функция $Y = a t^a$. Циклическое развитие экономики может быть охарактеризовано полиномом третьего порядка: $Y = a + a t + a t^2 + a t^3$. Подобрать функцию можно с помощью графика. При определении параметров функции обычно используют три метода: избранных точек, наименьших расстояний и наименьших квадратов. Для решения нашего задания остановимся на определении параметров линейного уравнения $Y = a + bx$ способом наименьших квадратов. Для упрощения расчетов в рядах динамики показатели времени t преобразуются таким образом, чтобы их сумма была равна нулю. Если ряд динамики содержит нечетное количество временных отрезков, например пять, то значения t будут выглядеть так: -2, -1, 0, +1, +2. Если ряд динамики состоит из четного количества временных отрезков, например шести, то значения t будут представлены как: -3, -2, -1, +1, +2, +3. Используя такие преобразования, можно получить параметры линейного уравнения. Находим параметры линейного уравнения: $=278/12=23,17$.

В результате линейное уравнение, показывающее зависимость размера прибыли Y от фактора времени t , будет таким: $Y = 23,17 + 0,68t$. Параметр a показывает средний размер прибыли. Показатель b характеризует силу влияния временного фактора на прибыль и означает изменение прибыли при увеличении отрезка времени на единицу. Изменяя временной параметр t в полученном уравнении, заполним колонку с теоретическими значениями прибыли Y_t выровненного динамического ряда. Проверим правильность вычислений, сравнив суммы годовой прибыли по отчетным данным Y и теоретическим значениям Y_t . В нашем примере они совпадают, что означает правильность проведенных расчетов.

После определения параметров функции следующим обязательным этапом работы является оценка надежности полученного уравнения с помощью критерия Фишера. Представим алгоритм такой оценки:

1. Определяем общую дисперсию анализируемого показателя Y по формуле:
2. Находим остаточную дисперсию, пользуясь формулой:
3. Получаем фактическую дисперсию вычитая из общей дисперсии остаточную дисперсию
4. Находим фактический критерий Фишера.

В этой формуле k – количество оцениваемых параметров (переменных величин), а n – число временных периодов статистического наблюдения. $K=1$ в нашем примере.

5. Определяем табличное значение критерия Фишера, используя уровень значимости $P=0,05$ (таблица). Уровень значимости показывает вероятность не принять гипотезу при условии, что она верна.

6. Сравниваем фактический критерий Фишера с его табличным значением. Если фактический критерий превышает табличное значение, то делается вывод о надежности полученной функциональной зависимости и что она адекватно отражает динамический тренд анализируемого показателя.

7. Завершим рассмотрение нашего примера аналитического выравнивания показателей, сделав оценку надежности полученного уравнения. Общая дисперсия равна 12,97. Остаточная дисперсия составляет 5,93. Определяем фактическую дисперсию, она равна 7,04. Находим фактический критерий Фишера, учитывая, что в нашем примере число степеней свободы большей дисперсии $f_1=k+1=2$, а число степеней свободы меньшей дисперсии $f_2=n-k-1=10$. В итоге получаем значение фактического критерия Фишера 5,93. Далее находим табличное значение Фишера на пересечении строки f_2 и колонки f_1 (таблица). Оно равно 4,10. Табличное значение меньше фактического значения Фишера. Следовательно, полученное уравнение надежно и адекватно отражает динамический тренд изменения прибыли предприятия.

Содержание отчета о проделанной работе

Отчет по лабораторной работе представляется в виде документа Word. В состав документа входят: название работы, цель работы, копии экрана, иллюстрирующие выполнение лабораторной работы, выводы по работе.

Лабораторная работа № 6. Элементы математической статистики: проверка статистических гипотез

Цель: изучение методов и инструментов математической статистики, применяемых в эконометрике

Теоретическая часть:

В эконометрических исследованиях чаще всего используются нормальное распределение, распределение Стьюдента и распределение Фишера. На основе этих распределений построены статистические критерии, служащие для проверки гипотез. Нормальное распределение случайной величины X характеризуется лишь двумя параметрами: средним значением и дисперсией. Это обозначается как $N(\mu, \sigma^2)$. График плотности нормального распределения $f(x)$ имеет колоколообразный вид, симметричный относительно центра распределения:

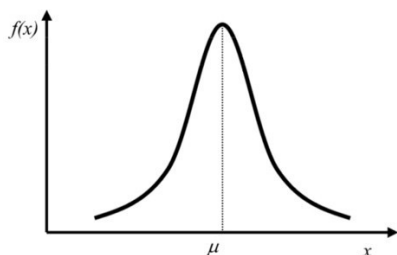


Рисунок 6 – График плотности нормального распределения

Максимум этой функции находится в точке , а разброс относительно этой точки определяется параметром . Чем меньше значение , тем более острый и высокий максимум $f(x)$. Если от точки среднего значения (или точки максимума плотности нормального распределения) отложить вправо и влево три величины стандартного отклонения, то площадь под графиком плотности распределения, подсчитанная по этому промежутку, будет равна 99,73 % всей площади под графиком (правило трех сигм). Другими словами, 99,73 % всех независимых наблюдений из нормального распределения лежат в радиусе трех стандартных отклонений от среднего значения.

Задание: по исходным данным вычислить ковариацию и коэффициент корреляции между переменными x , y . Представим исходные данные (x, y) и расчетные показатели в виде расчетной таблицы.

Таблица 7 – Исходные данные

№	x	y		xy	
1	2	1	4	2	1
2	6	2	36	12	4
3	10	4	100	40	16
4	14	11	196	154	121
5	18	12	324	216	144
Итого	50	30	660	424	286
Среднее	10	6	132	84,8	57,2

$$var(x) = \overline{x^2} - (\bar{x})^2 = 132 - 100 = 32$$

$$var(y) = \overline{y^2} - (\bar{y})^2 = 57,2 - 36 = 21,2$$

$$(x, y) = \overline{xy} - \bar{x}\bar{y} = 84,8 - 60 = 24,8$$

$$r = \frac{cov(x, y)}{\sqrt{var(x)var(y)}} = \frac{24,8}{\sqrt{21,2 * 32}} = 0,952$$

Ход решения задания:

Для вычисления коэффициента корреляции в Excel имеется функция: $r=\text{коррел}(\text{массив_x};\text{массив_y})$.

Содержание отчета о проделанной работе:

Отчет по лабораторной работе представляется в виде документа Word. В состав документа входят: название работы, цель работы, копии экрана, иллюстрирующие выполнение лабораторной работы, выводы по работе.

Тема 4. «Нормальное распределение и связанные с ним распределения»

Лабораторная работа № 7. Нормальное распределение и связанные с ним распределения: распределение хи-квадрат, Стьюдента, Фишера

Цель: знать основные сведения о нормальном и связанных с ним распределениях, уметь проверять несмещенность, эффективность, состоятельность статистических оценок.

Теоретическая часть:

Статистической гипотезой H называется предположение относительно параметра или вида распределения случайной величины.

Нулевой гипотезой H_0 называют выдвигаемую гипотезу. Обычно считают, что H_0 – гипотеза об отсутствии различий.

Конкурирующей гипотезой H_1 называют гипотезу, которая противоречит нулевой. Таким образом, H_1 – гипотеза о значимости различий.

Статистическим критерием называется случайная величина, которая служит для проверки нулевой гипотезы. В качестве статистического критерия выбирается такая случайная величина, точное или приближенное значение которой известно.

Если в качестве критерия проверки нулевой гипотезы используется случайная величина, подчиненная распределению Стьюдента, то ее обозначают через t (t -статистика), а подчиненная распределению Фишера – через F (F -статистика).

Критерий хи-квадрат Пирсона - это наиболее простой критерий проверки значимости связи между двумя категоризованными переменными. Критерий Пирсона основывается на том, что в таблице *ожидаемые* частоты при гипотезе "между переменными нет зависимости" можно вычислить непосредственно.

Задание: построить регрессионные зависимости: расходов на питание и личных доходов; расходов на питание и времени по следующим данным (усл. ед) и оценить качество подгонки:

Таблица 8 – Исходные данные

Год	2017	2018	2019	2020	2021
X	2	6	10	14	18
Y	1	2	4	11	12

Пусть истинная модель описывается выражением $y=a+bx+e$. Расчетные показатели

представлены в виде расчетной таблицы:

Таблица 9 – Расчетные данные

x^2	xy	$\hat{y}$	$(y - \bar{y})^2$	$(\hat{y} - \bar{y})^2$	$(y - \hat{y})^2$
4	2	-0,2	25	38,44	1,44
36	12	2,9	16	9,61	0,81
100	40	6	4	0	4
196	154	9,1	25	9,61	3,61
324	216	12,2	36	38,44	0,04
660	424	30	106	96,1	9,9
132	84,8	6	21,2	19,22	1,98
$\overline{x^2}$	$\overline{xy}$	$\hat{\bar{y}}$	$var(y)$	$var(\hat{y})$	$var(e)$

Ход решения задания:

Окончательно получаем:

$$var(x) = \overline{x^2} - (\bar{x})^2 = 132 - 100 = 32$$

$$cov(x, y) = \overline{xy} - \bar{x}\bar{y} = 84,8 - 10 * 6 = 24,8$$

$$b = \frac{cov(x, y)}{var(x)} = \frac{24,8}{32} = 0,775$$

$$a = \bar{y} - b\bar{x} = 6 - 0,775 * 10 = -1,75$$

Следовательно, $y = -1,75 + 0,775$ - линия регрессии. Коэффициент $b = 0,775$ показывает, что при увеличении дохода на 1 усл. ед. расходы на питание увеличиваются в среднем на 0,775 усл. ед. Качество подгонки оцениваем коэффициентом детерминации:

$$R^2 = \frac{\text{var}(\hat{y})}{\text{var}(y)} = \frac{19,22}{21,2} = 0,907$$

Значимость коэффициента детерминации проверяем по F-тесту.

Расчетное значение критерия (F_p):

$$F_p = \frac{R^2(n-2)}{1-R^2}$$

Зададим $\alpha = 0,05$. Число степени свободы $n-2 = 12 - 2 = 10$. Критическое значение критерия определяем с помощью функции $F_{кр} = F_{\text{РАСПОБР}}(0,05; 1; 10) = 10,13$. Поскольку $F_p = 29,12 > F_{кр} = 10,13$, то H_0 - отклоняется, т. е. коэффициент детерминации 0,907 статистически значим. Полученные значения представим в виде таблицы:

Таблица 10 – Полученные значения

Дисперсионный анализ					
	<i>df</i>	<i>SS</i>	<i>MS</i>	<i>F</i>	Значимость <i>F</i>
Регрессия	1	96,1	96,1	29,1212	0,01247
Остаток	3	9,9	3,3		
Итого	4	106			

Содержание отчета о проделанной работе:

Отчет по лабораторной работе представляется в виде документа Word. В состав документа входят: название работы, цель работы, копии экрана, иллюстрирующие выполнение лабораторной работы, выводы по работе.

Лабораторная работа № 8. Нормальное распределение и связанные с ним распределения: методы визуализации статистической информации

Цель: владеть навыками работы со статистическими таблицами для распределений; владеть методами визуализации имеющейся статистической информации

Теоретическая часть: Правила проверки нулевой гипотезы

В данном параграфе будут перечислены основные методы визуализации имеющейся статистической информации, а именно: описательные статистики, частотные распределения, гистограмма, диаграмма рассеяния, Q-Q-и P-P-графики, корреляционные таблицы, таблицы сопряженности. Визуализация крайне эффективна при первичном анализе данных, так как позволяет изучить закономерности в данных и получить представление об их

1. Описательные статистики (descriptive statistics). Изучение описательных статистик является весьма эффективным методом получения информации о данных. Чаще всего описательные статистики представляются в виде таблиц, в которых содержится следующая информация: количество наблюдений, максимум, минимум, выборочное среднее и др.

Гистограмма (histogram) - один из самых популярных способов визуализации данных. Гистограмма - это эмпирическая оценка плотности случайной величины по выборке. Графически гистограмма представляет собой набор прямоугольников. Высота каждого столбца

(прямоугольника) определяется количеством значений выборки в выбранном диапазоне, а количество столбцов задает число выбранных диапазонов. Гистограмма дает некоторое представление о распределении случайной величины, является ли оно симметричным, похоже ли на нормальное и т.д. Определяется с помощью функции Excel:

Диаграмма рассеяния (scatter plot) - метод визуализации данных, при котором на декартовой плоскости отображаются значения переменных в виде точек. Данный метод бывает крайне полезен при попарном сопоставлении переменных: благодаря таким графикам можно определить тип взаимосвязи переменных.

Графики Q-Q (quantile - quantile) и P-P (probability — probability).

Квантиль распределения - это значение, которого случайная величина не превышает с заданной вероятностью. Данный тип графиков часто используется при анализе распределений. Q- Q-график - это метод сравнения квантилей двух распределений для одного и того же набора данных.

Задание: покажем, что в модели регрессии $\hat{Y} = a$ оценка МНК для a есть $\bar{y}$. Выборочная регрессия для заданной модели есть $\hat{y}_i = a$. Наблюдаемые значения зависимой переменной связаны с расчетными значениями уравнением. Оценку a найдем из минимизации величины.

Ход решения задания:

$$y_i = \hat{y}_i + e_i = a + e_i$$

$$Q = \sum e_i^2 = \sum (y_i - a)^2 = \sum y_i^2 - 2a \sum y_i + na^2$$

Получаем

$$Q' = -2 \sum y_i + 2an = 0$$

$$a = \frac{\sum y_i}{n} = \bar{y}$$

Выборочная регрессия

Содержание отчета о проделанной работе:

Отчет по лабораторной работе представляется в виде документа Word. В состав документа входят: название работы, цель работы, копии экрана, иллюстрирующие выполнение лабораторной работы, выводы по работе.

Тема 5. «Линейная регрессия с одной объясняющей переменной»

Лабораторная работа № 9. Линейная регрессия с одной переменной: метод наименьших квадратов для нахождения оценок коэффициентов парной регрессии

Цель: изучить сущность и назначение моделей линейной регрессии в экономических исследованиях, научиться применять пакет Excel для работы с регрессионными моделями

Теоретическая часть:

В модели парной линейной регрессии зависимость между переменными в генеральной совокупности представляется в виде

$$Y = \alpha + \beta X + \varepsilon$$

где X - неслучайная величина;

Y , ε - случайные величины.

Величина Y называется объясняемой (зависимой) переменной, или результативным признаком, а X - объясняющей (независимой) переменной, или факторным признаком. Постоянные α, β - параметры уравнения.

Наличие случайного члена ε (ошибки регрессии) связано с воздействием на зависимую переменную других неучтенных в уравнении факторов.

На основе выборочного наблюдения оценивается выборочное уравнение регрессии (линия регрессии):

$$\hat{y} = a + bx,$$

где a, b - оценки параметров α, β .

Коэффициент регрессии показывает, на сколько единиц изменится результирующий признак при изменении факторного признака на единицу. Зависимость между факторным и результирующим признаком можно записать в виде уравнения регрессии.

Знак перед коэффициентом b характеризует нам тип связи. Если коэффициент b будет положительным, то можно говорить о том, что зависимость имеет прямой характер. Если же значение коэффициента будет отрицательным, то это будет обратная связь. Само значение коэффициента b будет говорить о силе влияния факторного признака x на результирующий y . Чем больше значения этого коэффициента, тем сильнее влияние факторного признака. А вот коэффициент a трактуется несколько в другом контексте. Его значение вбирает в себя силу влияния всех неучтенных факторных признаков. Поскольку рассматривается влияние на результирующий только одного фактора, это не исключает влияния некоторых других факторов. Их влияние в данном уравнении не рассматриваем. Но, тем не менее, это влияние существует, и оно накапливается в значении a . Если a будет существенно больше коэффициента b , то можно сказать, что для анализа выбран не самый влияющий факторный признак и надо искать другие показатели, которые более существенно влияют на результирующий признак.

При изучении множественных взаимосвязей могут встречаться факторы, которые действуют разнонаправленно. Влияние одного фактора может быть описано прямой зависимостью, влияние другого фактора - обратной зависимостью. Для расчета неизвестных значений коэффициентов a, b которые описывают силу влияния факторных признаков, используется система линейных уравнений.

$$y = \beta_0 + \beta_1 x_1 + \dots + \beta_k x_k + \varepsilon,$$

Обобщением линейной регрессионной модели с одной объясняющей переменной является линейная регрессионная модель с k -объясняющими переменными (модель множественной регрессии):

Построение экономической модели включает выбор объясняющих переменных. Свойства

оценок коэффициентов регрессии в значительной мере зависят от правильной спецификации модели. Предположим, что переменная y зависит от двух переменных x_1 и x_2 в соответствии с соотношением

$$y = \alpha + \beta_1 x_1 + \beta_2 x_2 + \varepsilon.$$

Однако считается, что модель выглядит как:

$$y = \alpha + \beta_1 x_1 + \varepsilon,$$

и оценивается регрессия

$$\hat{y} = a + b_1 x_1.$$

В этом случае оценка b_1 является смещенной. Смещенность оценки b_1 связана с тем, что если не учесть x_2 в регрессии, то переменная x_1 будет играть двойную роль: отражать свое прямое влияние и заменять переменную x_2 в описании ее влияния. Коэффициент R^2 для данной регрессии отражает общую объясняющую способность переменной x_1 в обеих ролях и является завышенной оценкой.

Задание: по данным бюджетного обследования семи случайно выбранных семей изучалась зависимость накопления y от дохода x_1 и стоимости имущества x_2 . Исходные данные приведены в таблице:

Таблица 11 – Исходные данные

y	3	7	5	4	2	7	6
X_1	40	55	45	30	30	60	50
X_2	60	40	40	15	90	30	30

Ход решения задания:

Используя пакет анализа Excel (регрессия), проанализировать зависимость накопления дохода и имущества. Результаты расчетов с использованием инструмента Регрессия выводятся под общим названием Вывод итогов в виде таблиц.

Таблица 12 – Итоговые данные

Регрессионная статистика	
Множественный R	0,970
R-квадрат	0,942
Нормированный R-квадрат	0,913
Стандартная ошибка	0,577
Наблюдения	7

Дисперсионный анализ					
	Df	SS	MS	F	Значимость F
Регрессия	2	21,527	10,763	32,355	0,0034
Остаток	4	1,331	0,333		
Итого	6	22,857			

	Коэффициенты	Стандартная ошибка	t-статистика	P-значение	Нижние 95%	Верхние 95%
Y-пересечение	0,493	1,229	0,402	0,708	-2,917	3,904
X ₁	0,128	0,022	5,866	0,004	0,067	0,188
X ₂	-0,030	0,010	-2,860	0,046	-0,058	-0,001

Коэффициенты регрессии $a=0,49$, $b_1=0,128$, $b_2=-0,030$. Запишем уравнение регрессии в виде: $Y=0,49+0,128x_1-0,030x_2$. Качество модели оценивается коэффициентом детерминации, его величина для нашей модели составила 0,942, что означает факторами дохода и имущества можно объяснить 94,2 % вариации накопления. Установим статистическую значимость коэффициента детерминации. Поскольку значимость $F=0,0034$, что меньше значимости на уровне 5 %, то коэффициент детерминации значим при 5%-ом уровне, то есть модель в целом адекватна исходным данным. Коэффициент b_1 равен 0,128, что больше 0 – показывает, что при увеличении дохода на 1 усл. ед. накопление увеличивается на 0,128 усл. ед. Коэффициент b_2 равен -0,030, что меньше 0 – показывает, что при увеличении имущества на 1 усл. ед. накопление уменьшается на 0,030 усл. ед. Результаты оценивания регрессии совместимы не только с полученными значениями коэффициентов регрессии, но и с некоторым их множеством (доверительным интервалом). С вероятностью 95 % доверительный интервал коэффициента b_1 есть (0,067 ... 0,188), а для b_2 есть (-0,058...-0,001).

Содержание отчета о проделанной работе:

Отчет по лабораторной работе представляется в виде документа Word. В состав документа входят: название работы, цель работы, копии экрана, иллюстрирующие выполнение лабораторной работы, выводы по работе.

Лабораторная работа № 10. Линейная регрессия с одной объясняющей переменной: линейная и нелинейная зависимости

Цель: изучить сущность и назначение моделей линейной регрессии в экономических исследованиях, научиться применять пакет Excel для работы с регрессионными моделями

Теоретическая часть: При изучении множественных взаимосвязей могут встречаться факторы, которые действуют разнонаправленно. Влияние одного фактора может быть описано прямой зависимостью, влияние другого фактора - обратной зависимостью. Для расчета неизвестных значений коэффициентов a , b которые описывают силу влияния факторных признаков, используется система линейных уравнений. Из данной системы уравнений можно рассчитать значение необходимых коэффициентов и получить уравнение множественной корреляции.

$$\sum (y_i - y_x)^2 \rightarrow \min$$

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

$$a = \bar{y} - b \bar{x} = \frac{\sum y_i \sum x_i^2 - \sum x_i y_i \sum x_i}{n \sum x_i^2 - (\sum x_i)^2}$$

$$b = \frac{\overline{xy} - \bar{x}\bar{y}}{\overline{x^2} - (\bar{x})^2} = \frac{n \sum x_i y_i - \sum x_i \sum y_i}{n \sum x_i^2 - (\sum x_i)^2}$$

Для оценки совместного влияния нескольких факторных признаков используется показатель, который называется коэффициент множественной корреляции или совокупный коэффициент корреляции.

$$R_{xyz} = \sqrt{\frac{r_{yx}^2 + r_{yz}^2 - 2(r_{yx} * r_{yz} * r_{xz})}{1 - r_{xz}^2}}$$

Задание: по четырем предприятиям региона изучается зависимость выработки продукции на одного работника у (тыс. руб.) от ввода в действие новых основных фондов x2 (% от стоимости фондов на конец года) и от удельного веса рабочих высокой квалификации в общей численности рабочих x1 (%). Определите модель множественной регрессии использованием метода наименьших квадратов. Сравните полученные расчетным путем результаты с результатами Excel.

Таблица 13 – Исходные данные

Номер предприятия	1	2	3	4
X1	1	2	3	5
X2	0	1	3	4
У	6	11	19	28

Ход решения задания:

Используя программу Excel составим таблицу промежуточных расчетов.

Таблица 14 – Промежуточные расчеты

№ компании	x1	x2	y	$\sum x_1 y$	x_1^2	$\sum x_2 y$	x_2^2	$x_1 x_2$
1	1	0	6	6	1	0	0	0
2	2	1	11	22	4	11	1	2
3	3	3	19	57	9	57	9	9
4	5	4	28	140	25	112	16	20
Итого	11	8	64	225	39	180	26	31

Из полученных промежуточных значения составим систему уравнений для последующего их решения.

$$\begin{cases} 64 = 4 \cdot a + 11 \cdot b_1 + 8 \cdot b_2 \\ 225 = 11 \cdot a + 39 \cdot b_1 + 31 \cdot b_2 \\ 180 = 8 \cdot a + 31 \cdot b_1 + 26 \cdot b_2 \end{cases}$$

Решим систему линейных уравнений (СЛУ) методом Крамера (составим матрицу и найдем определители). Получим уравнение следующего вида: $y = 2,4 + 3,4x_1 + 2,2x_2$. Сравним полученные значения параметров уравнения с расчетами Excel.

ВЫВОД ИТОГОВ					
Регрессионная статистика					
Множественный коэффициент	0,99972326				
R-квадрат	0,9994466				
Нормированный коэффициент	0,99833979				
Стандартная ошибка	0,39223227				
Наблюдения	4				
Дисперсионный анализ					
	<i>df</i>	<i>SS</i>	<i>MS</i>	<i>F</i>	<i>значимость F</i>
Регрессия	2	277,846154	138,923077	903	0,02352453
Остаток	1	0,15384615	0,15384615		
Итого	3	278			
Коэффициенты стандартной ошибки статистики: Р-Значение Нижние 95% Верхние 95% Нижние 95,0% Верхние 95,0%					
Y-пересечение	2,38461538	0,56000845	4,25817748	0,14684429	-4,7309667 9,50019744
x1	3,38461538	0,48650426	6,95701085	0,09088514	-2,7970073 9,56623806
x2	2,15384615	0,45508306	4,73286383	0,1325608	-3,6285324 7,93622469

Рисунок 7 – Итоговые расчеты

Оформите отчет о проделанной работе.

Содержание отчета о проделанной работе:

Отчет по лабораторной работе представляется в виде документа Word. В состав документа входят: название работы, цель работы, копии экрана, иллюстрирующие выполнение лабораторной работы, выводы по работе.

Тема 6: «Линейная регрессия с несколькими объясняющими переменными»

Лабораторная работа № 11. Линейная регрессия с несколькими объясняющими переменными: МНК – оценки коэффициентов множественной регрессии

Цель: формулу для нахождения МНК-оценок коэффициентов множественной регрессии, владеть навыками оценки множественной регрессии в основных статистических пакетах.

Теоретическая часть:

Обобщением линейной регрессионной модели с одной объясняющей переменной является линейная регрессионная модель с k-объясняющими переменными (модель множественной регрессии):

$$y = \beta_0 + \beta_1 x_1 + \dots + \beta_k x_k + \varepsilon,$$

где $\beta_0, \beta_1, \dots$ – параметры модели

x_1, x_2 – объясняющие переменные

ε – случайный член.

Случайный член ε удовлетворяет тем же предпосылкам (условия Гаусса-Маркова), что и в модели с парной регрессией, но на объясняющие переменные наложено условие: случайные члены в любом наблюдении должны быть статистически независимы от объясняющих переменных.

При условии Гаусса-Маркова модель называется классической нормальной линейной

регрессионной моделью.

Задание: имеются данные зависимости выпуска продукции обрабатывающей промышленности на душу населения y от валового внутреннего продукта x на душу населения в том же году для 17 стран.

Таблица 15 – Исходные данные

№	y	x	№	y	x
1	18	3	10	100	24
2	27		11	63	25
3	18		12	130	26
4	45		13	135	27
5	55		14	60	28
6	68		15	70	35
7	51		16	80	37
8	84		17	180	44
9	85				

Постройте регрессионную модель выпуска продукции от валового внутреннего продукта. Проверить наличие гетероскедастичности, используя тест Спирмена. Исходные и расчетные данные представим в таблице используя программу Excel.

y	x	Ранг	e	e	Ранг	D	D ²
18	3	1	-3,60	3,60	2	-1	1
27	6	2	-3,35	3,35	1	1	1
18	7	3	-15,27	15,27	9	-6	36
45	9	4	5,89	5,89	4	0	0
55	13	5	4,22	4,22	3	2	4
68	15	6	11,38	11,38	7	-1	1
51	18	7	-14,38	14,38	8	-1	1
84	21	8	9,87	9,87	6	2	4
85	22	9	7,95	7,95	5	4	16
100	24	10	17,11	17,11	10	0	0
63	25	11	-22,81	22,81	11	0	0
130	26	12	41,28	41,28	15	-3	9
135	27	13	43,36	43,36	16	-3	9
60	28	14	-34,56	34,56	12	2	4
70	35	15	-44,99	44,99	17	-2	4
80	37	16	-40,83	40,83	14	2	4
180	44	17	38,74	38,74	13	4	16
						Итого	110

Ход решения задания:

Пусть модель описывает выражение $y=a+bx+e$. Оцененное уравнением регрессии методом наименьших квадратов: $y=12,84+2,92x$, $a=12,84$, $b=2,92$. Из графика остатков видно, что с увеличением переменной x размах колебаний остатков e тоже возрастает, поэтому есть предположение о зависимости ошибки регрессии от независимой переменной (гетероскедастичность). Коэффициент ранговой корреляции вычисляется по формуле и равен 0,865.

$$r = 1 - \frac{6 \sum D_i^2}{n(n^2-1)}$$

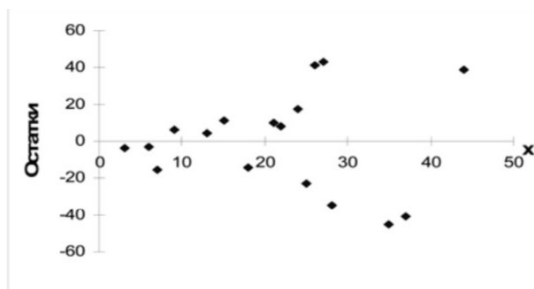


Рисунок 8 – Колебания остатков

Для проверки нулевой гипотезы используем t-тест, установим $\alpha=0,05$, число степеней свободы $\nu=n-2=15$. Критическую точку находим с помощью функции Excel: СТЬЮДРАСПОБР (0,05; 15)=2,13. Поскольку $|t_{\text{пл}}|=6,68$ больше $t_{\text{кр}}$ следовательно гипотеза H_0 отклоняется, таким образом имеет место гетероскедастичность.

Содержание отчета о проделанной работе:

Отчет по лабораторной работе представляется в виде документа Word. В состав документа входят: название работы, цель работы, копии экрана, иллюстрирующие выполнение лабораторной работы, выводы по работе.

Лабораторная работа № 12. Линейная регрессия с несколькими объясняющими переменными: показатели качества подгонки множественной регрессии

Цель: уметь проверять гипотезы о значимости коэффициентов множественной регрессии.

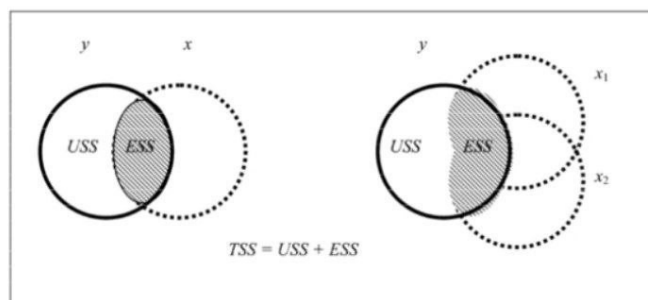
Теоретическая часть:

Коэффициент детерминации есть доля объясненной части разброса зависимой переменной, т.е.

$$R^2 = ESS / TSS = 1 - USS / TSS.$$

Он является мерой объясняющего качества уравнения регрессии по сравнению с горизонтальной линией $\hat{y} = \bar{y}$

Поскольку коэффициент R^2 измеряет долю дисперсии, совместно объясненной независимыми переменными, то, казалось бы, можно определить отдельный вклад каждой независимой переменной и таким образом получить меру ее относительной важности. Однако если объясняющие переменные, например x_1 и x_2 , сильно коррелируют между собой, то они объясняют одну и ту же часть разброса переменной y , и в этом случае трудно оценить вклад каждой из переменных в объяснение поведения y . Геометрический смысл коэффициента детерминации при использовании одной и двух объясняющих переменных показан на рис. 1.



Задание: по данным бюджетного обследования семи случайно выбранных семей изучалась зависимость накопления y от дохода x_1 стоимости имущества x_2 и расходов на питание x_3 . Исходные данные приведены в таблице (усл. ед.)

Таблица 16 – Исходные данные

y	3	7	5	4	2	7	6
x_1	40	55	45	30	30	60	50
x_2	60	40	40	15	90	30	30
x_3	10	15	12	8	10	20	15

Ход решения задания:

Используя Пакет анализа Excel (Корреляция), проанализируйте целесообразность включения в модель каждого фактора. Предположим, что имеется регрессионная модель

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \varepsilon$$

и необходимо оценить корреляцию между зависимой переменной y и объясняющей переменной x_1 после исключения влияния объясняющей переменной x_2 . С этой целью производим следующую процедуру:

- 1) оцениваем регрессию $y = a_1 + a_2 x_2$
- 2) оцениваем регрессию $x_1 = y_1 + y_2 x_2$
- 3) удаляем влияние x_2 , взяв остатки $e_y = y -$
- 4) определяем (выборочный) коэффициент частной корреляции между y и x_1 как (выборочный) коэффициент корреляции между e_y и e_{x_1} .

$$r(y, x_1 | x_2) = r(e_y, e_{x_1}).$$

Для коэффициента частной корреляции $r(y, x_1, x_2)$ используется также обозначение $r_{yx_1x_2}$.

Можно показать что справедлива следующая формула, связывающая коэффициенты частной и обычной корреляции:

$$r_{yx_1x_2} = \frac{r(y, x_1) - r(y, x_2)r(x_1, x_2)}{\sqrt{[1 - r^2(y, x_2)][1 - r^2(x_1, x_2)]}}.$$

Значения $r_{yx_1x_2}$ лежат в интервале $[-1, 1]$, а равенство нулю означает отсутствие прямого (линейного) влияния переменной x_1 на y .

Описанная выше процедура обобщается на случай, когда исключается влияние не одной, а нескольких переменных. Порядок частного коэффициента корреляции определяется количеством факторов, влияние которых исключается.

Содержание отчета о проделанной работе:

Отчет по лабораторной работе представляется в виде документа Word. В состав документа входят: название работы, цель работы, копии экрана, иллюстрирующие выполнение лабораторной работы, выводы по работе.

Тема 7: «Выбор функциональной формы модели»

Лабораторная работа № 13. Выбор функциональной формы модели: линейная модель.

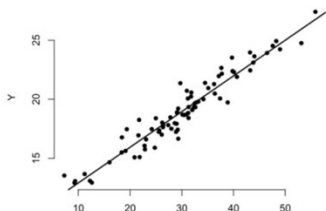
Цель: знать интерпретацию коэффициентов в линейной модели, уметь проводить выбор между линейной, линейной в логарифмах и полупологарифмической модели

Теоретическая часть:

Одним из условий теоремы Гаусса – Маркова является правильная спецификация модели

$$y = \beta_0 + \beta_1 x_1 + \dots + \beta_k x_k + \varepsilon$$

В модели зависимость Y от каждой объясняющей переменной является линейной. Как исследователь может предположить, что имеет место именно такая форма зависимости? Обычно первым шагом является изучение диаграмм рассеяния Y по каждой из объясняющих переменных. Если расположение точек похоже на вытянутое облако, то можно предположить, что имеет место линейная зависимость Y от соответствующей переменной.



Задание: проведем подбор ARIMA-модели по следующим данным, приведенным ниже и сделаем прогноз на предстоящий период.

Ход решения задания:

Исходный ряд y_t не является стационарным, поскольку имеет возрастающий тренд. Динамика имеет сезонный характер. Используя пакет Statistica можно перебрать различные варианты моделей ARIMA (p, d, q) при не очень больших значениях параметров.

Основными критериями отбора являются:

- остаток (остаточная дисперсия) должен быть по возможности минимизирован;
- все оцениваемые параметры (включая свободный параметр модели) должны быть высоко значимыми;

-прогноз должен учитывать сезонный характер данных.

Таблица 17 – Исходные данные

Дата	t	y_t	$z_t = \Delta y_t$	z_{t-1}	e_t
31.12.05	1	182,2			
06.01.06	2	182,3	0,1		
13.01.06	3	184,6	2,3	0,1	0,2693
20.01.06	4	185,2	0,6	2,3	-2,0422
27.01.06	5	188,2	3,0	0,6	0,8303
03.02.06	6	188,5	0,3	3,0	-2,5368
10.02.06	7	194,2	5,7	0,3	3,6137
17.02.06	8	195,4	1,2	5,7	-2,3872
24.02.06	9	195,6	0,2	1,2	-2,1365
03.03.06	10	197,9	2,3	0,2	0,2415
10.03.06	11	201,7	3,8	2,3	1,1578
17.03.06	12	204,1	2,4	3,8	-0,6591
24.03.06	13	204,4	0,3	2,4	-2,3700
31.03.06	14	205,9	1,5	0,3	-0,5863
07.04.06	15	208,1	2,2	1,5	-0,2199
14.04.06	16	212,0	3,9	2,2	1,2856
21.04.06	17	217,1	5,1	3,9	2,0131
28.04.06	18	225,7	8,6	5,1	5,1795
05.05.06	19	231,1	5,4	8,6	1,0067
12.05.06	20	236,1	5,0	5,4	1,4961
19.05.06	21	236,7	0,6	5,0	-2,7927
26.05.06	22	243,3	6,6	0,6	4,4303
02.06.06	23	247,0	3,7	6,6	-0,1374
09.06.06	24	247,9	0,9	3,7	-2,1313
16.06.06	25	246,0	-1,9	0,9	-4,1531
23.06.06	26	247,2	1,2	-1,9	-0,2748
30.06.06	27	250,6	3,4	1,2	1,0635
07.07.06	28	253,2	2,6	3,4	-0,3480

Применяя указанные критерии, найдем следующую идентификацию искомой модели ARIMA (0,1,0)(1,0,0) (без свободного параметра). Данная модель содержит первую разность для учета линейного тренда в динамике авиаперевозок y_t и коэффициент сезонной авторегрессии. В таблице приведены результаты оценивания параметров модели.

Исход.:Авиаперевозки					
Преобразования: D(1)					
Модель(0,1,0)(1,0,0) Сезонный лаг: 12 Остаток= 714,24					
Параметр	Парам.	Асимпт. ст.ошиб.	Асимпт. t(70)	p	Нижняя 95% дов. Верхняя 95% дов.
Ps(1)	0,935968	0,066979	13,97397	0,000000	0,802382 1,069555

Рисунок 10 - Результаты оценивания параметров модели

Приведенные в таблице параметры модели высокосignимы. На рисунках приведена автокорреляционная и частная автокорреляционная функция остатков модели.



Рисунок 11 – Автокорреляционная и частная автокорреляционная функция

Корреляции значимо отличаются от 0, следовательно выбранная модель является адекватной. Выполним прогнозирование в полученной модели с использованием программы. Результаты прогноза приведены в следующей таблице.

Прогнозы; Модель:(0,1,0)(1,0,0)
Исход.:Авиаперевозки
Начало исходных: 1 Конец исходн.:7:

Прогноз	Нижний 90%	Верхний 90%	Ст. ошиб.
439,5	392,6	486,5	28,2
392,8	326,4	459,2	39,9
448,0	366,6	529,3	48,8
453,6	359,6	547,5	56,4
484,4	379,4	589,5	63,0
534,9	419,8	650,0	69,0
601,3	477,0	725,6	74,6
635,9	503,0	768,8	79,7
546,1	405,2	687,1	84,6
489,1	340,5	637,7	89,1
438,6	282,8	594,4	93,5
463,9	301,1	626,6	97,6

Рисунок 12 – Результаты прогноза

Результаты прогноза предусматривают доверительный интервал. Это означает, что с вероятностью 90 % прогнозируемое значение попадает в заданный интервал.

Содержание отчета о проделанной работе:

Отчет по лабораторной работе представляется в виде документа Word. В состав документа входят: название работы, цель работы , копии экрана, иллюстрирующие выполнение лабораторной работы , выводы по работе.

Лабораторная работа № 14. Выбор функциональной формы модели: модели с постоянным темпом роста и эластичностью

Цель: владеть навыками создания новых переменных (в том числе логарифмов) в основных статистических пакетах

Теоретическая часть:

Если коэффициент В значим, то при прочих равных (*ceteris paribus*) условиях, т.е. фиксированных остальных факторах, увеличение X на одну единицу приводит к изменению Y на B, единиц (в сторону увеличения при положительной оценке и в сторону уменьшения при отрицательной)

Однако не всегда при анализе данных исследователь приходит к выводу, что имеет место линейная зависимость Y от X,..., X_n. Но в некоторых случаях монотонное преобразование Y и (или) каких-либо X, позволяет перейти к линейной зависимости. Примером является кривая

Филлипса $Y = B_0 + B_1X$, где Y - инфляция; X - уровень безработицы.

Наиболее распространенной на практике является трансформация Бокса - Кокса, частным случаем которой является логарифмическое преобразование зависимых и (или) независимых переменных. При этом интерпретация оценок коэффициентов меняется. Как же выбрать между моделями если каждая из них является адекватной? Поскольку в моделях одинаковые зависимые переменные, мы можем сравнить их по R^2 (лучше та модель, где этот показатель больше). Однако при разных зависимых переменных (Y и $\ln Y$) этот способ не годится.

Для сравнения линейной и линейной в логарифмах моделей можно и использовать тесты, предложенные А. Бера и М. МакАлером, а также Дж. МакКинноном, Х. Уайтом и Р. Дэвидсоном. Тест Бера и МакАлера предназначен для выбора между полулогарифмической ($\ln Y = B_0 + B_1X_1 + \dots + B_kX_k + \epsilon$) и линейной ($Y = B_0 + B_1X_1 + \dots + B_kX_k + \epsilon$)

Задание: исходя из корреляционной матрицы, полученной в Excel

	y	x_1	x_2	x_3
y	1			
x_1	0,907	1		
x_2	-0,664	-0,380	1	
x_3	0,830	0,935	-0,283	1

Рисунок 11 – Корреляционная матрица

Ход решения задания:

Из рассмотрения корреляционной матрицы делаем вывод, что факторы x_1 , x_3 дублируют друг друга $r(x_1, x_3) = 0,935$. В анализ целесообразно включить фактор x_1 а не x_3 , так как его связь с результативным признаком y сильнее, чем у фактора x_3 $r(y, x_1) = 0,907 > r(y, x_3) = 0,830$. Оценив регрессию y на x_1 , x_3 получим:

$$\hat{y} = 0,49 + 0,127x_1 - 0,029x_3, \quad R^2 = 0,941; \quad \bar{R}^2 = 0,912,$$

(0,021) (0,010)

где коэффициенты при факторных переменных значим. Заметим, что первоначальное значение коэффициента детерминации скорректировалось с первоначального значения 0,885, при отбраковке дублирующего фактора x_3 увеличилось до 0,912.

Содержание отчета о проделанной работе:

Отчет по лабораторной работе представляется в виде документа Word. В состав документа входят: название работы, цель работы, копии экрана, иллюстрирующие выполнение лабораторной работы, выводы по работе.

Тема 8. «Проблемы пропущенных и избыточных факторов»

Лабораторная работа № 15. Проблемы пропущенных и избыточных факторов: мультиколлинеарность данных

Цель: знать последствия включения в модель несущественных переменных, последствия не включения в модель существенных переменных

Теоретическая часть: При формулировке теоремы Гаусса - Маркова, эта теорема справедлива, если модель правильно специфицирована, т.е. имеет правильную функциональную форму (ее выбор обсуждался в предыдущей главе), все необходимые факторы в модель включены, а «лишних» факторов нет. Но как определить, какие факторы являются необходимыми, а какие - «лишними»? Каковы последствия неправильного выбора факторов?

Для того чтобы понять последствия не включения в модель необходимых факторов, рассмотрим следующий пример.

Предположим, что данные Y_i, X_{1i}, X_2 , $i = 1, \dots, n$, были порождены с помощью процесса $Y = B_0 + B_1 X_1 + B_2 X_2 + e$.

Однако исследователь об этом не подозревает и оценивает модель, в которой переменная X , пропущена:

$$Y = B_0 + B_1 X_1 + e.$$

Тогда МНК-оценка коэффициента B , будет иметь вид

$$B_1 = \text{COV}(X_1, Y) / \text{var}(X_1).$$

Мультиколлинеарность — это коррелированность двух или несколько объясняющих переменных в уравнении регрессии. При наличии мультиколлинеарности МНК-оценки формально существуют, но обладают рядом недостатков:

- небольшое изменение исходных данных приводит к существенному изменению оценок коэффициентов регрессии;

- 0 оценки коэффициентов регрессии имеют большие стандартные ошибки, малую значимость, в то время как модель в целом является значимой (высокое значение R^2).

Если при оценке уравнения регрессии несколько факторов оказались незначимы, то нужно выяснить, нет ли среди них сильно коррелированных между собой.

Для отбора факторов в модель регрессии и оценки их мультиколлинеарности анализируют корреляционную матрицу.

Общий вид корреляционной матрицы, составленной из переменных Y, X_1, X_2, X_3 приведен в следующей таблице.

Задание: Исследуется зависимость между средним курсом доллара и ценой на алюминий. Постройте корреляционное поле, выделите тренд и сделайте предварительный вывод форме

зависимости случайных величин.

Ход решения задания:

1. Зайдите на оф. сайт ЦБ РФ и экспортируйте информацию по динамике курса доллара за исследуемый период (оф. сайт: cbr.ru)
2. На сайте Лондонская биржа металл (оф. сайт lme.com) найдите стоимость металла
3. Создайте расчетный файл в Excel где консолидируйте найденную информацию
4. Найдите среднемесячное значение курса доллара по каждому анализируемому месяцу
5. Постройте корреляционное поле
6. Постройте тренд
7. Проведите анализ
8. При необходимости рассчитайте дополнительно необходимые коэффициенты

Таблица 17 – Исходные данные

-	Курс доллара						
	Янв	Фев	Мар	Апр	Май	Июн	Июл
1	76,2527			74,3823	73,5870	72,3723	73,1388
2	76,1854	74,4373	75,7023	74,939	73,4580	72,1777	73,1904
3	75,04	73,4747	75,8287	74,9578	73,4737	72,1694	73,6088
4	75,6354	73,7532	75,7576	74,768	73,3963	72,3260	73,8471
5	74,8569	73,9833	76,1741	75,0893	73,5266	72,6671	74,098
6	74,3615	73,9717	76,1535	76,4217	73,5803	73,1661	73,7663
7	73,3694	73,7755	75,3585	76,8198	73,6007	73,1987	73,6945
8	73,355	73,7669	74,6085	76,0155	73,6778	72,2216	74,4947
9	73,7243	73,2895	74,139	76,2491	73,6992	72,5048	74,491
10	73,9735	73,3092	73,6582	75,5535	73,8537	72,0323	74,3463
11	73,5453	73,9378	73,1019	76,9808	73,9968	71,8318	74,1656
12	73,7961	73,7579	72,9619	75,6826	74,3566	71,6797	74,2197
13	73,5264	73,8526	73,2317	77,2535	74,0400	72,1974	74,1236
14	74,2663	74,1192	73,5081	77,5104	74,1567	72,0829	74,0589
15	74,5157	74,2602	73,4996	77,1657	74,1373	72,8256	74,6336
16	73,8757	75,1107	74,0393	77,1011	74,5770	72,9294	74,4675
17	76,2527	75,7293	74,264	77,773	74,8617	73,2721	75,1952
18	76,1854	76,0801	74,4275	76,3802	75,2567	73,2636	74,058
19	75,04	75,9051	73,7864	76,6052	74,8451	73,4979	73,266
20	75,6354	75,5053	73,5187	76,0734	73,5870	73,2411	73,354
21	74,8569	74,4373	74,5755	75,8073	73,4580	73,2965	73,6175
22	74,3615	73,4747	74,0448	75,6373	73,4737	72,3723	72,9086
23	73,3694	73,7532	75,7023	74,3823	73,3963	72,1777	72,7234
24	73,355	73,9833	75,8287	74,939	73,5266	72,1694	73,1388
25	73,7243	73,9717	75,7576	74,9578	73,5803	72,3260	73,1904
26	73,9735	73,7755	76,1741	74,768	73,6007	72,6671	73,6088
27	73,5453	73,7669	76,1535	75,0893	73,6778	73,1661	73,8471
28	73,7961		75,3585	76,4217	73,6992	73,1987	74,098
29	73,5264		74,6085	76,8198	73,8537	72,2216	73,7663
30	74,2663		74,139	76,0155		72,5048	73,6945
31			73,6582				

Цена аллюм, \$ за тонну	январь	фев	мар	апр	май	июн	июл
	1987	2 202,5	2 212,5	2 445	2 404,5	2 523	2 556,5

Содержание отчета о проделанной работе:

Отчет по лабораторной работе представляется в виде документа Word. В состав документа входят: название работы, цель работы, копии экрана, иллюстрирующие выполнение лабораторной работы, выводы по работе.

Лабораторная работа № 16. Проблемы пропущенных и избыточных факторов: гетероскедастичность

Цель: владеть навыками применения методов пошагового включения и исключения переменных в основных статистических пакетах

Теоретическая часть:

Одной из предпосылок регрессионного анализа является предположение о постоянстве дисперсии случайного члена для всех наблюдений (гомоскедастичность). Это значит, что для каждого значения объясняющей переменной случайные члены имеют одинаковые дисперсии. Если это условие не соблюдается, то имеет место гетероскедастичность.

При отсутствии гетероскедастичности коэффициенты регрессии имеют наименьшую дисперсию среди всех несмещенных оценок, являющихся линейными функциями от наблюдений y .

Если наблюдается гетероскедастичность, то МНК-оценки будут неэффективными (они не будут иметь наименьшую дисперсию по сравнению с другими оценками данного параметра).

Оценки стандартных ошибок коэффициентов регрессии вычисляются в предположении, что распределение случайного члена гомоскедастично; если это не так, то они неверны (занижены), а, следовательно, t -статистика завышена. Это может привести к статистически значимым коэффициентам регрессии, тогда как в действительности это неверно.

Проблема гетероскедастичности характерна для пространственных данных, полученных от неоднородных объектов. Например, если исследуется зависимость прибыли предприятий от размера основного фонда, то можно ожидать, что для больших предприятий колебание прибыли будет выше, чем для малых.

Задание:

Используя исходные данные по товарным остаткам предприятия:

- 1) найдите показатели рядов динамики
- 2) используя метод скользящей средней и аналитического выравнивания, графически изобразите тренды изменения товарных остатков

Таблица 18 – Исходные данные

Показатель	янв	фев	мар	апр	май	июн	июл	авг	сен	окт	ноя	дек
Стоимость остатков, тыс. руб.	325	384	417	425	447	487	498	484	507	504	516	528

Ход решения задания:

- 1) Используя программу Excel рассчитаем показатели динамики: темп роста, темп

прироста, абсолютный прирост стоимости товарных остатков, найдем среднее значение.

Показатель	янв	фев	мар	апр	май	июн	июл	авг	сен	окт	нояб	дек	среднее
Стоимость товарных запасов, тыс. руб.	325	384	417	425	447	487	498	484	507	504	516	528	460,17
абсолютный прирост, тыс. руб.	0	59	92	100	122	162	173	159	182	179	191	203	135,17
темпл роста	1,00	1,18	1,28	1,31	1,38	1,50	1,53	1,49	1,56	1,55	1,59	1,62	1,40
прироста	0,00	18%	28%	31%	38%	50%	53%	49%	56%	55%	59%	62%	2%

Рисунок 12 – Результаты расчета показателей динамики

2) Для того чтобы выровнять ряд и найти тренд изменения воспользуемся методом скользящей средней и выровняем временной ряд.

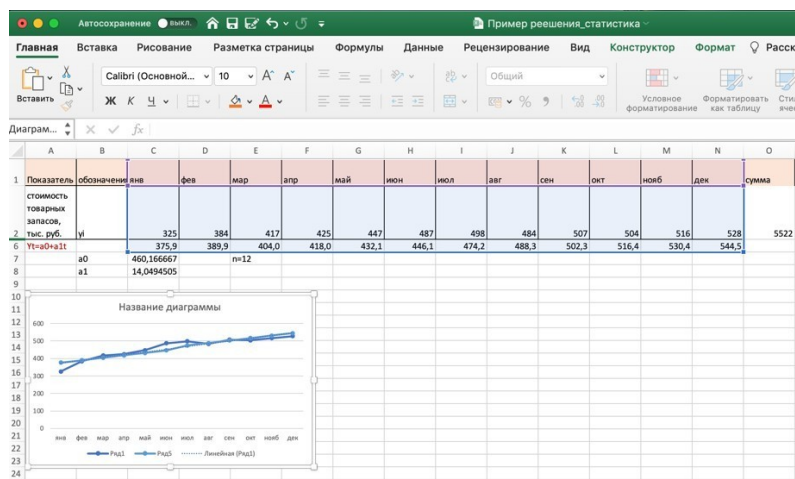


Рисунок 13 – Результаты выравнивания временного ряда

3) Для того чтобы построить прогноз временного ряда воспользуемся функцией Excel «добавить тренд». Так как в наших данных имеет место прямая линейная тенденция, которая описывается линейным уравнением, используем данное уравнение для прогноза на 2 периода вперед.

4) Сделайте выводы по проведенным расчетам

Содержание отчета о проделанной работе:

Отчет по лабораторной работе представляется в виде документа Word. В состав документа входят: название работы, цель работы, копии экрана, иллюстрирующие выполнение лабораторной работы, выводы по работе.

Тема 9. «Временные ряды и прогнозирование»

Лабораторная работа № 17. Временные ряды и прогнозирование: типы моделей

Цель: знать особенности анализа временных рядов; знать основные модели анализа временных рядов

Теоретическая часть:

Любой процесс, будь то динамика физической величины или поведение экономического индикатора, представляет собой изменение значения некоторой переменной во времени. Одни переменные ведут себя достаточно предсказуемо (среднюю температуру июля в Москве можно предсказать с высокой точностью), иные случайно колеблются внутри некоторого коридора (уровень безработицы), а некоторые процессы ведут себя безо всякой видимой закономерности, причем в любой момент минимальное изменение условий может привести к катастрофическим изменениям процесса.

Итак, временной ряд - это ряд некоторых числовых величин, измеренных в последовательные моменты времени. Существует принципиальное различие между анализом пространственных выборок (crossection) и временных рядов (time series): наблюдения в последних строго упорядочены

во времени, в то время как порядок наблюдений в перекрестных выборках можно менять произвольно. Переставлять наблюдения во временных рядах некорректно, так как из-за этого меняется закономерность развития ряда.

По той же причине во временных рядах нельзя удалять отдельные наблюдения: если в одномоментной кросс-секционной базе данных какие-то точки являются выбросами или закодированы с ошибкой, то тогда их удаление может улучшить качество подгонки оцениваемой модели, во временных рядах приходится работать со всем временным диапазоном без возможности удалить ухудшающие модель наблюдения (исключениями являются случаи, когда нестандартные наблюдения находятся в начале или конце временного периода).

Случайный процесс - процесс, реализациями которого в каждый момент времени являются случайные величины.

Задание: имеются поквартальные данные (усл. ед.) об объеме потребления электроэнергии у в некотором районе за 4 года.

Таблица 19 – Исходные данные

Квартал	Год			
	1	2	3	4
I	60	72	80	90
II	44	48	56	66
III	50	50	64	70
IV	90	100	110	108

Постройте аддитивную модель временного ряда. Дайте прогнозное потребление электроэнергии в течение четырех кварталов ближайшего следующего года?

Ход решения задания:

В качестве зависимой переменной при анализе временного ряда выступают фактические уровни ряда y_t , а в качестве независимой переменной - время. По графику ряда можно установить наличие приблизительно линейного тренда и сезонных колебаний.

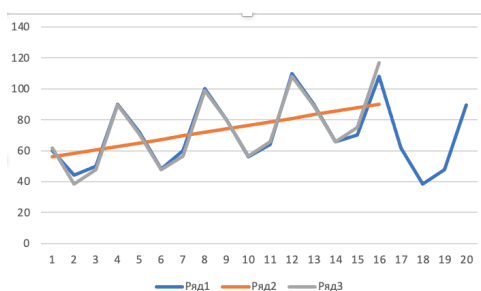


Рисунок 14 – Графическое изображение тенденций временного ряда

Для исключения влияния сезонной компоненты произведём выравнивание исходного ряда методом скользящей средней за 4 квартал.

	№ квартала						
год	1	2	3	4			
2018			-12,5	25,5			
2019	5,75	-20,75	-11	27			
2020	5,5	-20,25	-14,75	28,75			
2021	6,75	-17,75					
Сумма	18	-58,75	-38,25	81,25			
Средн оценка сез комп	6	-19,583333	-12,75	27,0833333	0,75	Корр	0,1875
S скоррект.	5,8125	-19,770833	-12,9375	26,8958333	0		

Рисунок 15 – Промежуточные результаты расчетов сезонной компоненты

В строке рассчитаны средние сезонной вариации по годам за каждый квартал и их сумма равная 0,75. В строке S скорректированное рассчитаны значения сезонных компонент как разность между сезонной вариацией и корректирующим коэффициентом $0,75/4$. Расчет трендовой компоненты произведем в таблице.

t	Потребление эл. Энергии y(t)	St	Yt-St	T	T+S	Et=(Yt-(S+T	Et^2	(Yt-Yсред)^2
1	60	5,8125	54,1875	55,9264706	61,7389706	-1,7389706	3,02401871	169
2	44	-19,770833	63,7708333	58,2029412	38,4321078	5,56789216	31,0014231	841
3	50	-12,9375	62,9375	60,4794118	47,5419118	2,45808824	6,04219777	529
4	90	26,8958333	63,1041667	62,7558824	89,6517157	0,34828431	0,12130196	289
5	72	5,8125	66,1875	65,0323529	70,8448529	1,15514706	1,33436473	1
6	48	-19,770833	67,7708333	67,3088235	47,5379902	0,4620098	0,21345306	625
7	60	-12,9375	72,9375	69,5852941	56,6477941	3,35220588	11,2372843	169
8	100	26,8958333	73,1041667	71,8617647	98,757598	1,24240196	1,54356263	729
9	80	5,8125	74,1875	74,1382353	79,9507353	0,04926471	0,00242701	49
10	56	-19,770833	75,7708333	76,4147059	56,6438725	-0,6438725	0,41457186	289
11	64	-12,9375	76,9375	78,6911765	65,7536765	-1,7536765	3,07538116	81
12	110	26,8958333	83,1041667	80,9676471	107,86348	2,13651961	4,56471603	1369
13	90	5,8125	84,1875	83,2441176	89,0566176	0,94338235	0,88997026	289
14	66	-19,770833	85,7708333	85,5205882	65,7497549	0,2502451	0,06262261	49
15	70	-12,9375	82,9375	87,7970588	74,8595588	-4,8595588	23,615312	9
16	108	26,8958333	81,1041667	90,0735294	116,969363	-8,9693627	80,4494681	1225

Рисунок 16 – Результаты расчета трендовой компоненты

С помощью пакета «анализ данных – регрессия» найдем значения параметров, учитывая трендовую и сезонные составляющие, которые приведены в таблице построим трендовую модель и дадим прогноз на 17,18,19,20 квартал.

Содержание отчета о проделанной работе:

Отчет по лабораторной работе представляется в виде документа Word. В состав документа входят: название работы, цель работы, копии экрана, иллюстрирующие выполнение лабораторной работы, выводы по работе.

Лабораторная работа № 18. Временные ряды и прогнозирование:

прогнозирование во временных рядах

Цель: владеть навыками проведения тестов на стационарность, знать методы анализа временных рядов

Теоретическая часть:

Все процессы в окружающем нас мире можно разделить на четыре типа:

1) непрерывное время, непрерывные значения случайной

Простой пример - координаты движения самолета;

2) непрерывное время, дискретные значения случайной величины. Пример - количество заявок на покупку акций определенной компании;

3) дискретное время, непрерывные значения случайной величины. В эту категорию попадает большинство исследуемых экономических процессов: статистические органы публикуют информацию через определенные промежутки времени, а показатели могут принимать значения, определенные с произвольной точностью. При дискретном времени наблюдения могут быть разделены как одинаковыми, так и различными по продолжительности

интервалами. Пример такого ряда с дискретным временем - значения годовой инфляции страны за несколько лет;

4) дискретное время, дискретные значения случайной величины. В эту категорию попадают показатели, связанные с неделимыми объектами, изменение которых происходит только в определенные моменты. Примером может послужить численность группы студентов в течение всего периода наблюдения (частота дискретизации - 1 день, минимальное изменение - 1 человек).

В некоторых случаях количество значений дискретной случайной величины столь велико, что можно полагать его непрерывным: у предприятия, выпускающего авторучки, отгрузка исчисляется миллионами, а ошибка измерения и несущественность для предприятия одной авторучки позволяет считать выпуск ручек непрерывной величиной.

Задание: производственные компании распределяют прибыль (Π), оставшуюся после уплаты налогов: одну часть на выплату доходов акционерам в форме дивидендов (D), другую – на финансирование инвестиций. Когда прибыль растет, дивиденды тоже увеличиваются, но, как правило, не в той пропорции. Причиной этого в основном является осторожность руководства компании.

Ход решения задания:

Предположим, что у фирмы имеется целевая долгосрочная доля выплат μ и что желаемый объем дивидендов D_t^* соотносится с текущей прибылью Π_t как $D_t^* = \mu \Pi_t + e_t$. Однако реальный объем дивидендов подвержен процессу частичной корректировки:

$$D_t - D_{t-1} = \mu(D_t^* - D_{t-1})$$

Используя данные о деятельности производственных компаний за ряд предыдущих лет:

Таблица 20 – Данные о деятельности производственных предприятий

t	D	Π
1	100	400
2	300	600
3	450	700
4	550	800
5	700	100
6	800	1100
7	900	1300
8	1000	1400
9	1100	1500
10	1200	1700

Можно построить уравнение регрессии

$$D_t = 68 + 0,293\Pi_t + 0,582D_{t-1}, R^2 = 0,999, \\ (28,9)(0,071) \quad (0,081)$$

где все коэффициенты значимы. Из соотношения $1 - \mu = 0,582$ определяется корректирующий коэффициент $\mu = 0,418$ а из соотношения $\mu\alpha = 0,293$ – оценка доли выплат

$\mu=0,7$.

Содержание отчета о проделанной работе:

Отчет по лабораторной работе представляется в виде документа Word. В состав документа входят: название работы, цель работы, копии экрана, иллюстрирующие выполнение лабораторной работы, выводы по работе.

Список рекомендуемой литературы

Основная литература:

1. Наумов, И. В. Эконометрика. Экономическое моделирование социально-экономических процессов в территориальных системах : учебное пособие / И. В. Наумов, Н. Л. Никулина. - Эконометрика. Экономическое моделирование социально-экономических процессов в территориальных системах, Весь срок охраны авторского права. - Электрон. дан. (1 файл). - Москва : Ай Пи Ар Медиа, 2021. - 127 с. - электронный. - Книга находится в премиум-версии ЭБС IPR BOOKS. - ISBN 978-5-4497-1408-4, экземпляров неограничено
2. Орлов, А. И. Эконометрика : учебное пособие / А. И. Орлов. - Эконометрика, 2021-12-05. - Электрон. дан. (1 файл). - Москва, Саратов : Интернет-Университет Информационных Технологий (ИНТУИТ), Ай Пи Ар Медиа, 2020. - 676 с. - электронный. - Книга находится в премиум-версии ЭБС IPR BOOKS. - ISBN 978-5-4497-0362-0, экземпляров неограничено
3. Яковлева, А. В. Эконометрика Электронный ресурс : Учебное пособие / А. В. Яковлева. - Эконометрика, 2020-02-05. - Саратов : Научная книга, 2019. - 223 с. - Книга находится в премиум-версии ЭБС IPR BOOKS. - ISBN 978-5-9758-1820-1, экземпляров неограничено

Дополнительная литература:

1. Воскобойников, Ю. Е. Эконометрика в Excel. Модели временных рядов Электронный ресурс / Воскобойников Ю. Е. : учебное пособие. - 2-е изд., стер. - Санкт-Петербург : Лань, 2020. - 152 с. - ISBN 978-5-8114-4863-0, экземпляров неограничено
2. Ершова, Н. А. Современная эконометрика Электронный ресурс : Учебное пособие / Н. А. Ершова, С. Н. Павлов. - Современная эконометрика, 2021-09-11. - Москва : Российский государственный университет правосудия, 2018. - 52 с. - Книга находится в премиум-версии ЭБС IPR BOOKS. - ISBN 978-5-93916-650-8, экземпляров неограничено
3. Кузнецова, О. А. Эконометрика (продвинутый уровень) Электронный ресурс / Кузнецова О. А. : учебно-методическое пособие. - Тольятти : ТГУ, 2020. - 125 с. - ISBN 978-5-8259-1525-8, экземпляров неограничено

МИНИСТЕРСТВО НАУКИ И ВЫСШЕГО ОБРАЗОВАНИЯ РОССИЙСКОЙ ФЕДЕРАЦИИ
ФЕДЕРАЛЬНОЕ ГОСУДАРСТВЕННОЕ АВТОНОМНОЕ ОБРАЗОВАТЕЛЬНОЕ УЧРЕЖДЕНИЕ
ВЫСШЕГО ОБРАЗОВАНИЯ
«СЕВЕРО-КАВКАЗСКИЙ ФЕДЕРАЛЬНЫЙ УНИВЕРСИТЕТ»

МЕТОДИЧЕСКИЕ УКАЗАНИЯ

по организации и проведению самостоятельной работы
по дисциплине «ЭКОНОМЕТРИКА»
для студентов специальности 38.03.01 Экономика
направленность (профиль) «Экономика организаций»

Ставрополь, 2026

Методические указания по дисциплине «Эконометрика» содержат задания для студентов, необходимые для организации самостоятельной работы.

Проработка предложенных заданий позволит студентам приобрести необходимые знания в области изучаемой дисциплины.

Предназначены для студентов специальности 38.03.01 Экономика, направленность, профиль «Экономика организаций».

Содержание

Введение.	4
1. Общая характеристика самостоятельной работы	4
2. План – график выполнения задания	4
3. Методические рекомендации по изучению теоретического материала	5
4. Методические рекомендации по подготовке доклада	6
5. Методические рекомендации по подготовке к тестированию	14
6. Методические рекомендации по подготовке к экзамену	19
Список рекомендуемой литературы	25

ВВЕДЕНИЕ

Данные методические указания по самостоятельной работе адресованы студентам очной и заочной форм обучения специальности 38.03.01 Экономика.

В целях успешного освоения дисциплины студентам необходимо:

- изучить теоретический материал, предусмотренный практическим курсом;
- изучить теоретический материал, предназначенный для самостоятельного изучения;
- закрепить теоретический материал, выполнив практические задания;
- выполнить предусмотренные курсом лабораторные работы.
- проводить сбор, обработку и систематизацию данных для их дальнейшего анализа и применения в эконометрическом моделировании;
- иметь представление о построении и оценке эконометрических моделей;
- применять результаты работы с эконометрическими моделями, интерпретировать полученные результаты;

Дисциплина изучается студентами в течение одного семестра.

1. ОБЩАЯ ХАРАКТЕРИСТИКА САМОСТОЯТЕЛЬНОЙ РАБОТЫ

Целью учебной дисциплины является освоение основ и практических навыков работы с эконометрическими моделями.

В ходе преподавания дисциплины ставятся следующие задачи:

- изучение принципов описания социальных, экономических, финансовых объектов, процессов и явлений языком математических моделей со случайной компонентой;
- изучение основных этапов эконометрического моделирования;
- формирование навыков подготовки статистической информации, предназначенной для построения эконометрических моделей и работы в основных статистических пакетах;
- формирование умений проверять статистические гипотезы и строить точечные и интервальные прогнозы;
- формирование навыков диагностики моделей, оценки регрессионных моделей;
- развитие навыков интерпретации основных результатов оценки моделей.

Самостоятельная работа по дисциплине «Эконометрика» выполняется с целью получения и закрепления знаний, приобретенных при изучении теоретического материала.

2. ПЛАН – ГРАФИК ВЫПОЛНЕНИЯ ЗАДАНИЯ

Виды и содержание самостоятельной работы студента; формы контроля и сроки сдачи содержатся в рабочей программе дисциплины.

3. МЕТОДИЧЕСКИЕ РЕКОМЕНДАЦИИ ПО ИЗУЧЕНИЮ ТЕОРЕТИЧЕСКОГО МАТЕРИАЛА

Изучение любого раздела или темы следует начинать с ознакомления с вопросами плана изучения темы. Теоретический материал представляет собой конспект лекций, содержащий необходимый набор утверждений и формул (без детальных подробностей), но с подробным обоснованием их использования при решении конкретных экономических задач. При изучении материала необходимо помимо лекционных материалов использовать рекомендуемую основную и дополнительную литературу для лучшего усвоения материала.

Осваивать теорию следует в соответствии с той последовательностью, которая представлена в плане лекции. Методика работы с литературой предусматривает ведение записи прочитанного в виде плана - конспекта, опорного конспекта. Это позволит сделать знания системными, зафиксировать и закрепить их в памяти.

Для успешного освоения дисциплины, необходимо самостоятельно детально изучить представленные темы по рекомендуемым источникам информации:

№ п/п	Виды самостоятельной работы	Рекомендуемые источники информации (№ источника)			
		Основная	Дополнительная	Методическая	Интернет-ресурсы
1	Выполнение лабораторных работ	1 2 3	1 2 3	1	1 2 3
2	Подготовка к лекции	1 2 3	1 2 3	1	1 2 3
3	Самостоятельное изучение литературы	1 2 3	1 2 3	1	1 2 3
4	Подготовка к зачету	1 2 3	1 2 3	1	1 2 3

4. МЕТОДИЧЕСКИЕ РЕКОМЕНДАЦИИ ПО ПОДГОТОВКЕ К ТЕСТИРОВАНИЮ

Успешное выполнение тестовых заданий является необходимым условием итоговой положительной оценки в соответствии с рейтинговой системой обучения. Выполнение тестовых заданий предоставляет студентам возможность самостоятельно контролировать уровень своих знаний, обнаруживать пробелы в знаниях и принимать меры по их ликвидации. Форма изложения тестовых заданий позволяет закрепить и восстановить в памяти пройденный материал. Тестовые задания охватывают основные вопросы по дисциплине «Эконометрика».

У студента есть возможность выбора правильного ответа или нескольких правильных ответов из числа предложенных вариантов. Для выполнения тестовых заданий студенты должны изучить лекционный материал по теме, соответствующие разделы учебников, учебных пособий и других источников.

Контрольный тест выполняется студентами самостоятельно во время семинарских занятий.

Тестовые задания

1. Эконометрика:

- это наука, которая использует методы математической статистики для описания теоретических моделей реальных хозяйственных экономических процессов
- это наука, которая использует методы статистики для описания теоретических моделей реальных хозяйственных экономических процессов

2. Объектом эконометрики являются:

- экономические процессы, происходящие в экономической системе общества
- финансовые процессы, происходящие в денежно-кредитной системе

3. Целью эконометрики является:

- а) оценка точечных и интервальных прогнозов деятельности генеральной совокупности объектов экономической системы на основании расчетов по данным выборочной совокупности или реализации случайного процесса;
- б) расчет точечных прогнозов деятельности генеральной совокупности объектов экономической системы по данным генеральной совокупности

4. Методами эконометрики являются:

- а) метод наибольших квадратов
- б) индексный метод
- в) метод наименьших квадратов

5. Задачами эконометрики являются:

- а) расчет финансово-экономических показателей деятельности экономического субъекта
- б) нахождение ошибок выборочной совокупности
- в) разработка эконометрических методов, повышающих точность прогноза

6. Построение графика в среде Excel производится в такой последовательности:

- а) ввести данные X и Y, выделить данные Y и X, мастер диаграмм, точечная, готово;
- б) ввести данные X и Y, выделить данные Y и X, мастер диаграмм, гистограмма, готово

7. Модель это:

- а) условный образ объекта исследования
- б) математическое уравнение объекта исследования
- в) точное копирование объекта исследования

8. Абстрактные модели подразделяются на:

- а) графические, математические, словесно-описательные
- б) математические
- в) логические

9. Этапы эконометрического моделирования:

- а) постановочный, априорный, информационный, спецификация, модели, идентификация модели, определение качества модели, верификация модели, выводы и предложения;
- б) постановочный, априорный, информационный, спецификация модели, идентификация модели, выводы и предложения;
- в) верификация модели, информационный, спецификация модели, априорный, идентификация модели, выводы и предложения;

10. Сущность метода экстраполяции ряда динамики заключается в:

- а) определении неизвестного уровня внутри ряда динамики
- б) нахождении неизвестного уровня за пределами ряда динамики
- в) нахождение среднего темпа роста явления за период
- г) выявлении тенденции развития явления в динамики

11. Нестационарный ряд динамики характеризует...

- а) нелинейную функцию изменений процесса во времени
- б) сезонные колебания показателей изучаемого процесса
- в) отсутствие тренда в изменениях показателя
- г) спонтанную колеблемость признака в течение исследуемого периода

12. Спецификация модели – это:

- а) определение вида математической функции, которая описывает влияние объясняемых переменных на зависимую переменную;
- б) определение факторов, влияющих на зависимую переменную;
- в) определение случайной величины, присутствующей в эконометрической модели

13. Регрессионная модель является линейной относительно переменных, если выполняются условия:

- а) переменные находятся в первой степени
- б) модель должна быть аддитивной
- в) модель должна быть мультипликативной

14. Регрессионная модель является нелинейной относительно связи между факторами, если выполняется условие: связь между факторами является мультипликативной:

- а) нет
- б) да

15. Метод наименьших квадратов применим к линейным относительно коэффициентов аддитивным регрессионным уравнениям следующего вида

:

- а) нет
- б) да

16. Общий вид множественного регрессионного уравнения для выборочных данных можно представить следующим математическим выражением

- а) нет
- б) да

17. Мультиколлинеарность – это свойство переменных объектов изучения, которое заключается в том, что:

- а) существует хотя бы одна пара переменных, между которыми имеется существенная связь, определяемая по коэффициенту корреляции;
- б) не существует хотя бы одной пары переменных между которыми имеется существенная связь, определяемая по коэффициенту корреляции

18. Известно два направления устранения мультиколлинеарности: а) методы, которые используют преобразования данных;

- б) методы, которые используют разные приемы построения моделей
- в) методы экономического анализа
- г) экспертные методы

19. Методами часто используемыми для построения моделей при наличии мультиколлинеарных факторов являются шаговый метод построения модели, метод корреляционных плеяд

- а) нет
- б) да

20. Регрессионная модель с переменной структурой - это такая модель, которая учитывает неоднородность количественной зависимой переменной с помощью качественного фактора:

- а) да
- б) нет

21. Различают два типа неоднородностей зависимой переменной:

- а) положительную и отрицательную
- б) прямую и обратную
- в) функциональную и структурную

22. Структурной неоднородностью называют такую тенденцию зависимой переменной:

- а) которая имеет четко выраженные выбросы для временных рядов или расслоения для пространственных данных;
- б) которая имеет прямую зависимость с результирующим признаком
- в) которая показывает удельный вес переменной в генеральной совокупности

23. Связь считается функциональной, если определенному значению факторного признака соответствует :

- а) два значения результирующего
- б) несколько значений результирующего признака
- в) нулевое значение результирующего признака

г) строго определенное значение результативного признака

24. Связь между функциональным и факторным признаком называется обратной если:

- а) с увеличением результативного признака факторный уменьшается
- б) с увеличением результативного признака факторный увеличивается
- в) коэффициент корреляции отрицательный
- г) с увеличением факторного признака результативный уменьшается
- д) коэффициент корреляции положительный
- е) с увеличением результативного признака факторный уменьшается

25. Для оценки параметров уравнения регрессии можно применить метод: а) проб и ошибок

- б) дифференциального и интегрального исчисления
- в) конечных разностей
- г) наименьших квадратов

26. Коэффициент детерминации может принимать значения:

- а) от 0 до +1
- б) любые положительные
- в) от -1 до +1
- г) от -1 до 0
- д) любые меньше 0

27. Средняя из внутригрупповых дисперсий:

- а) отражает вариацию признака, возникающего под влиянием случайных факторов
- б) может быть вычислена как разность общей и межгрупповой дисперсии
- в) характеризует вариацию признака, зависящую от всех условий в данной совокупности
- г) характеризует вариацию, возникающую под факторным признаком

28. Гетероскедастичность означает неоднородность. Гетероскедастичность остатков означает различие дисперсий остатков при всех фиксированных значениях объясняемой переменной

- а) да
- б) нет

29. Целями анализа временных рядов являются:

- выявление структуры временного ряда для более глубокого понимания динамических экономических процессов
- получение точечного и интервального прогноза динамического экономического показателя

- а) да
- б) нет

30. Во временных рядах не должно быть пропусков. Если имеется пропуск, то он восстанавливается средним значением, обычно из ближайших к нему чисел.

- а) да
- б) нет

Критерии оценивания компетенций

Оценка «отлично» выставляется студенту, если количество правильных ответов на тестовые задания превышает 90 %.

Оценка «хорошо» выставляется студенту, если количество правильных ответов на тестовые задания превышает 70 %).

Оценка «удовлетворительно» выставляется студенту, если количество правильных ответов на тестовые задания превышает 50 %).

Оценка «неудовлетворительно» выставляется студенту, если количество правильных ответов на тестовые задания составляет менее 50 %, либо ответы заимствованы.

Оформление ответов на тесты

Ответы на тесты оформляются на студентом на отдельном листе самостоятельно. В правом углу проставляется ФИО и группа, далее следует номер теста и выбранный вариант ответа.

6. МЕТОДИЧЕСКИЕ РЕКОМЕНДАЦИИ ПО ПОДГОТОВКЕ К ЗАЧЕТУ

Успеваемость студентов по дисциплине оценивается в ходе текущего контроля и промежуточной аттестации.

Текущий контроль

Текущее контрольное мероприятие считается сданным, если студент получил за него не менее 60 % от установленного для этого контроля максимального балла.

Промежуточная аттестация

Промежуточная аттестация в форме **зачета с оценкой**. Компетенции (знания, умения и навыки) студентов оцениваются по пятибалльной системе: «отлично», «хорошо», «удовлетворительно» и «неудовлетворительно». В ходе собеседования студенту предлагается теоретический вопрос и практическая задача. Для подготовки к ответу отводится 45 минут. При подготовке к ответу студенту предоставляется право пользования методическими материалами. При проверке практического задания, оцениваются: последовательность и рациональность выполнения, полнота ответа, правильность сделанных выводов.

Вопросы для собеседования:

1. Эконометрика как научная дисциплина
2. Цели и отличительные черты эконометрики
3. Типы эконометрических моделей
4. Методы оценки и верификации эконометрических моделей
5. Основные статистические пакеты, применяемые для оценки эконометрических моделей
6. Суть эконометрического моделирования
7. Основные этапы эконометрического моделирования
8. Проблемы, возникающие при построении эконометрических моделей
9. Основные виды эконометрических моделей
10. Классификация видов эконометрических переменных
11. Общая и нормальная линейная модели парной регрессии
12. Операции суммирования в математической статистики
13. Общие понятия о случайных величинах
14. Понятие о выборочной совокупности
15. Понятие о генеральной совокупности
16. Характеристика нормального распределения случайной величины
17. Статистические оценки и свойства
18. Понятие несмещенности оценок
19. Понятие эффективности оценок
20. Понятие состоятельности оценок
21. Понятие ковариации и корреляции
22. Суть проверки статистических гипотез
23. Правило проверки нулевой гипотезы (t-статистика)
24. Правило проверки нулевой гипотезы (F-статистика, P- значение)
25. Модели временных рядов: общая характеристика
26. Суть метода экспоненциального сглаживания
27. Простое экспоненциальное сглаживание
28. Подходы к определению средней ошибки при использовании метода экспоненциального сглаживания
29. Среднеквадратическая ошибка в экспоненциальном сглаживании и правила ее определения
30. Сезонная и несезонная модели с трендом и без тренда
31. Аддитивная и мультипликативная модели временных рядов
32. Анализ распределенных лагов
33. Суть методов экстраполяции и интерполяции

34. Понятие тренда, виды
35. Методы выравнивая временных рядов
36. Характеристика методов: укрупнения интервалов, скользящей средней
37. Суть и применение метода аналитического выравнивания
38. Классификация регрессионных моделей
39. Понятия о модели парной регрессии
40. Суть метода наименьших квадратов
41. Нахождение параметров линейной модели с применением метода наименьших квадратов
42. Нахождение параметров нелинейной модели с применением метода наименьших квадратов
43. Анализ вариации зависимой переменной
44. F-тест на качество оценивания моделей парной регрессии
45. Понятие гомоскедастичности и гетероскедастичности
46. Обнаружение гетероскедастичности: общая характеристика применяемых тестов
47. Тест ранговой корреляции Спирмена
48. Тест Голфельда-Квандта
49. Предпосылки регрессионного анализа (четыре условия)
50. Теорема Гаусса-Маркова
51. Расчет стандартной ошибки коэффициентов регрессии
52. Суть обобщенного метода наименьших квадратов
53. Обнаружение автокорреляции
54. Тест Дарбина-Уотсона
55. Автокорреляция с лаговой зависимой переменной
56. Модель множественной регрессии: общие понятия
57. Классическая нормальная линейная регрессионная модель
58. Анализ вариации зависимой переменной (модель множественной регрессии)
59. Понятие мультиколлинеарности
60. Характеристика коэффициента частной корреляции
61. Спецификация модели множественной регрессии
62. Прогнозирование в регрессионных моделях
63. Авторегрессионные процессы: понятие стационарности
64. Модели авторегрессии

Практические задания

Задание 1

Построить график и модель зависимости Y от X . В таблице представлены пространственные данные зависимости Y – розничный товарооборот от X – количество продавцов, при условии, что в магазине имеется очередь.

№ магазина	X (чел)	Y (тыс. руб.)
1	1	4
2	3	6
3	2	7
4	4	10
5	5	?

Задание 2

Торговое предприятие имеет сеть, состоящую из 12 магазинов, информация о деятельности которых представлена в таблице.

№	Годовой товарооборот, млн. руб.	Торговая площадь, тыс. кв.м.	Среднее число посетителей в день, тыс. чел.
---	---------------------------------	------------------------------	---

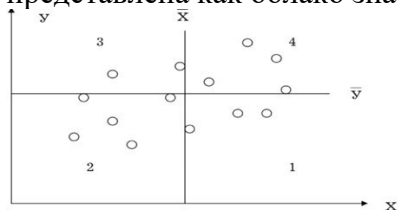
1	19	0,25	8
2	38	0,31	10
3	41	0,55	9
4	41	0,48	11
5	56	0,78	9
6	69	0,98	7
7	75	0,94	12
8	89	1,21	10
9	91	1,29	9
10	91	1,12	14
11	99	1,29	12
12	109	1,49	13

Требуется построить диаграммы рассеяния годового товарооборота (Y) в зависимости от торговой площади (X1) и среднего числа посетителей в день (X2).

Определить форму связи между результирующим показателем (Y) и каждым из факторов (X1 и X2).

Задание 3

Определите приблизительно индекс детерминации при условии что зависимость Y от X представлена как облако значений на рисунке.



Задание 4

Имеются данные о площади и стоимости элитного жилья в курортном городе, полученные на основании мониторинга рынка жилья в данном регионе. Позволяют ли полученные данные говорить о существовании связи? Определите ее характер и сделайте проверку гетероскедастичности используя тест ранговой корреляции Спирмена?

п/п №	Площадь, кв.м	Стоимость, млн. руб.
1	40	145
2	48	127
3	30	130
4	25	120
5	41	140
6	43	135
7	50	145
8	60	150
9	55	140
10	34	130
11	65	155
12	30	125

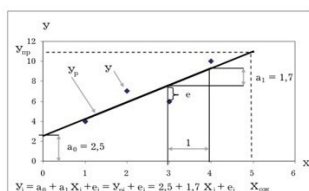
Задание 5

Построена модель зависимости накопления (Y) от дохода (X)

Требуется дать экономическую интерпретацию коэффициента () уравнения регрессии.

Задание 6

Используя рисунок, визуальным образом приблизительно определите индекс детерминации.



Задание 7

Имеются поквартальные данные (усл. ед.) об объеме потребления электроэнергии у в некотором районе за 4 года.

Квартал	Год			
	1	2	3	4
I	60	72	80	90
II	44	48	56	66
III	50	50	64	70
IV	90	100	110	108

Постройте аддитивную модель временного ряда. Дайте прогнозное потребление электроэнергии в течение четырех кварталов ближайшего следующего года?

Задание 8

Исследуется зависимость между среднемесячными доходами X на семью (в тыс. у.е.) и расходами Y на покупку кондитерских изделий (в у.е.) Данные опроса представлены в таблице:

X	4,8	3,8	5,4	4,2	3,4	4,6	3,4	4,8	5,0	3,8	5,2	4,0	3,8	4,6	4,4
Y	75	68	78	71	64	73	66	75	75	65	77	69	67	72	70

Построить корреляционное поле и сделать предварительный вывод о форме зависимости случайных величин?

Задание 9

Используя фактические данные приведенные в таблице изучите виды тенденций временных рядов условной области используя графический метод, Excel:

t	Y _t	t	Y _t
1	908	13	-5963
2	4219	14	804
3	-8914	15	-1002
4	4770	16	-6950
5	-6674	17	-4596
6	-7798	18	-9671
7	3882	19	-2793
8	3409	20	-6734
9	7294	21	-2104
10	-9623	22	7857
11	-1032	23	3003
12	-6734	24	-8964

Задание 10

Применяя метод экспоненциального сглаживания (Excel) постройте прогноз временного ряда, найдите ошибку прогноза?

Объем инвестиций	Период	Млрд., руб.
2020	январь	1000
	февраль	1200
	март	1260
	апрель	1197
	май	1790
	июнь	1969

	июль	1260
	август	1323
	сентябрь	1250
	октябрь	1790
	ноябрь	2148
	декабрь	2578
2021	январь	1050
	февраль	1260
	март	1323
	апрель	1257
	май	1880
	июнь	2068
	июль	1323
	август	1390
	сентябрь	1313
	октябрь	1880
	ноябрь	2256
	декабрь	2707

Задание 11

Расположите основные шаги эконометрического исследования в правильном порядке:

- 1) утверждение экономической теории;
- 2) поиск данных для эмпирической проверки эконометрической модели;
- 3) предложение соответствующей статистической (или эконометрической) модели;
- 4) предложение соответствующей математической модели;
- 5) верификация модели;
- 6) оценка параметров эконометрической модели с помощью выбранного метода оценивания;
- 7) выбор другой модели или способа оценивания при признании модели неподходящей;
- 8) использование результатов оценки при признании модели подходящей.

Задание 12

По данным, приведенным в таблице, установите наличие гетероскедастичности, используя тест Голдфелда-Квандта.

№ п/п	y	x
1	18	3
2	27	6
3	19	7
4	45	9
5	55	13
6	68	15
7	51	18
8	84	21
9	85	22
10	100	24
11	63	25
12	130	26
13	135	27
14	60	28
15	70	35
16	80	37
17	180	44

Задание 13

Используя исходные данные по товарным остаткам предприятия найдите показатели рядов динамики:

Показатель	янв	фев	мар	апр	май	июн	июл	авг	сен	окт	ноя	дек
Стоимость товара	325	384	417	425	447	487	498	484	507	504	516	528

Задание 14

Имеются данные о площади и стоимости элитного жилья в курортном городе, полученные на основании мониторинга рынка жилья в данном регионе. Позволяют ли полученные данные говорить о существовании связи? Уровень значимости равен 0,05.

п/п №	Площадь, кв.м	Стоимость, млн. руб.
1	40	145
2	48	127
3	30	130
4	25	120
5	41	140
6	43	135
7	50	145
8	60	150
9	55	140
10	34	130
11	65	155
12	30	125

Задание 15

В таблице приведены данные за 8 месяцев по расходам на покупку недвижимости и личным доходам граждан (средним) города К.

-постройте регрессионную зависимость

-рассчитайте стандартную ошибку, значимость и доверительный интервал коэффициентов модели.

Месяц	Расходы на покупку недвижимости, млн. руб.	Личные доходы
Янв	6,00	200,00
Фев	7,00	280,00
Март	9,00	310,00
Апр	15,00	380,00
Май	21,00	400,00
Июн	19,00	420,00
Июл	22,00	440,00
Авг	24,00	430,00

Задание 16

Построено двухфакторное уравнение годового товарооборота (Y) в зависимости от торговой площади магазина (x1) и среднего числа посетителей в день (x2): $y = -10,8 + 62x_1 + 2,3x_2$. Требуется дать экономическую интерпретацию коэффициентов уравнения регрессии.

Задание 17

Имеются данные затрат на устранение брака в сборочном цехе, вызванные ошибками в чертежах, составленных конструкторским отделом завода

t	Y _t	t	Y _t
1	12	6	15
2	15	7	12
3	16	8	16
4	12	9	14
5	13	10	15

где t – время (дни), Y – расходы на устранение брака (тыс. руб.). Необходимо определить основные характеристики временного ряда: критерий восходящих и нисходящих серий, дисперсию значений, критерий Дарбина-Уотсона.

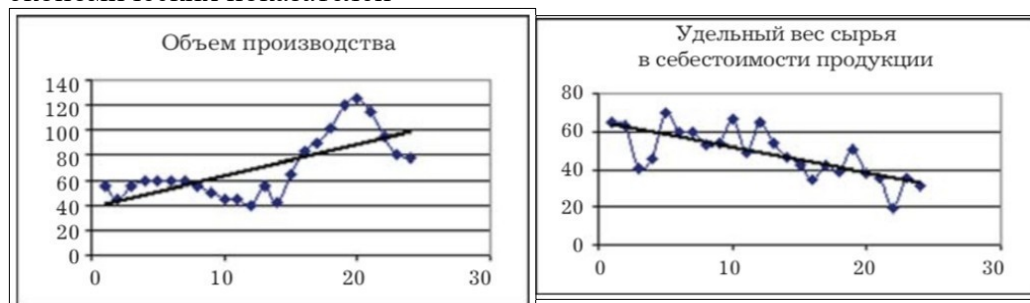
Задание 18

Имеются данные об объеме предложения товара (y) его цены (x₁) и заработной сотрудников (x₂) за 10 месяцев. Выявите на уровне значимости 0,05 наличие автокорреляции остатков в модели регрессии используя статистику Дарбина - Уотсона.

y	x ₁	x ₂
20	10	12
35	15	10
30	20	9
45	25	9
60	40	8
70	37	8
75	43	6
90	35	4
105	40	4
110	55	5

Задание 19

Изучите основные виды тенденций, которые имеются во временных рядах и остатках экономических показателей



Задание 20

Имеется временной ряд Y_t розничного товарооборота (тыс. руб.) киоска розничной торговли в течение недели:

t	Y _t
1	5
2	8
3	6
4	9
5	5
6	10
7	8

Где t - время, Y_t – розничный товарооборот. Необходимо определить автокорреляцию первого порядка для временного ряда Y_t .

Задание 21

Необходимо определить коэффициент и достоверность автокорреляции остатков. База данных представлена в таблице

t	Y_t	Y_{pt}	e_t	e_{t-1}		()
1	5	6	-1		1	
2	8	6,4286	1,5714	-1	2,4694	6,6122
3	6	6,8571	-0,8571	1,5714	0,7347	5,898
4	9	7,2857	1,7143	-0,8571	2,9388	6,6122
5	5	7,7143	-2,7143	1,7143	7,3673	19,612
6	10	8,1429	1,8571	-2,7143	3,449	20,898
7	8	8,5714	-0,5714	1,8571	0,3265	5,898
8				-0,5714		
Сумма					18,286	65,531

где t – время, Y_t – розничный товарооборот

Задание 22

Имеются данные (усл. ед) зависимости выпуска продукции обрабатывающей промышленности на душу населения (y) от валового внутреннего продукта (x) на душу населения в том же году для 17 стран.

№ п/п	y	x
1	18	3
2	27	6
3	19	7
4	45	9
5	55	13
6	68	15
7	51	18
8	84	21
9	85	22
10	100	24
11	63	25
12	130	26
13	135	27
14	60	28
15	70	35
16	80	37
17	180	44

-постройте регрессионную модель выпуска продукции от валового внутреннего продукта, оцените качество модели, определите доверительный интервал параметров уравнения регрессии?

Задание 23

По выборке объема $n=5$ получен выборочный коэффициент корреляции $r=0,952$. На 5-% уровне установите статистическую значимость выборочного коэффициента корреляции.

Задание 24

Имеется база данных зависимости величины розничного товарооборота Y магазина от известных затрат на рекламу X .

i	X_i	Y_i
1	12	35
2	13	40
3	11	36

4	14	48
5	16	?

Найдите выборочные коэффициенты линейной парной регрессии визуальным способом и решением системы нормальных уравнений, определите прогнозное значение Упр при ожидаемом расходе на рекламу равном $X = 16$ тыс. рублей.

Задание 25

Используя метод скользящей средней и аналитического выравнивания (Excel) графически изобразите тренды изменения товарных остатков:

Показатель	янв	фев	мар	апр	май	июн	июл	авг	сен	окт	ноя	дек
Стоимость товара	325	384	417	425	447	487	498	484	507	504	516	528

Задание 26

По исходным данным вычислите ковариацию и коэффициент корреляции между переменными x , y .

№	x
1	8
2	12
3	18
4	20
5	24
Итого	
Среднее	

Задание 27

Для некоторой страны имеются данные (усл. ед.) о совокупном доходе используя модель формирования доходов Y , объеме потребления C и инвестициях I , полученные за 10 лет.

Постройте функцию потребления

C_t	190	198	200	180	200	210	220	210	205
I_t	10	20	30	20	10	20	30	20	15
Y_t	200	218	230	200	210	230	250	230	220

Задание 28

Постройте модель множественной линейной регрессии (с учетом проверки на мультиколлинеарность), дайте оценку?

п/п	Y	X_1	X_2	X_3	X_4	X_5
1	126	4	3,5	14	17	100
2	137	5	4,8	13,8	17,3	98,2
3	148	6	3,8	14,2	16,8	101,1
4	191	8	8,6	14,5	16,2	103,5
5	274	12	8,1	14,5	16	104,1
6	370	13	9,6	15	18	107
7	432	14	14,5	17,1	20,2	107,4
8	445	16	18,6	12	15,8	108,5
9	367	16	19,7	14,8	18,2	108,3
10	367	16	10,6	15,9	16,8	109,2

Задание 29

Составьте модель множественной регрессии зависимости производительности труда одного рабочего от уровня модернизации оборудования и от удельного веса квалифицированности рабочих в общей численности персонала, для расчетов используйте метод наименьших квадратов.

Производительность, млн. руб.	Уровень модернизации оборудования, %	Удельный вес квалифицированности рабочих, %
6	1	0
11	2	1
19	3	3
28	5	4

Задание 30

По данным бюджетного обследования семи случайно выбранных семей изучалась зависимость накопления (y) от дохода (x_1), стоимости имущества (x_2) и расходов на питание (x_3). Исходные данные (усл. ед.)

y	x_1	x_2	x_3
3	40	60	10
7	55	40	15
5	45	40	12
4	30	15	8
2	30	90	10
7	60	30	20
6	50	30	15

Используя Пакет анализа Excel проанализируйте целесообразность включения в модель каждого фактора?

СПИСОК РЕКОМЕНДУЕМОЙ ЛИТЕРАТУРЫ

Основная литература

1. Наумов, И. В. Эконометрика. Экономическое моделирование социально-экономических процессов в территориальных системах : учебное пособие / И. В. Наумов, Н. Л. Никулина. - Эконометрика. Экономическое моделирование социально-экономических процессов в территориальных системах, Весь срок охраны авторского права. - Электрон. дан. (1 файл). - Москва : Ай Пи Ар Медиа, 2021. - 127 с. - электронный. - Книга находится в премиум-версии ЭБС IPR BOOKS. - ISBN 978-5-4497-1408-4, экземпляров неограничено
2. Орлов, А. И. Эконометрика : учебное пособие / А. И. Орлов. - Эконометрика, 2021-12-05. - Электрон. дан. (1 файл). - Москва, Саратов : Интернет-Университет Информационных Технологий (ИНТУИТ), Ай Пи Ар Медиа, 2020. - 676 с. - электронный. - Книга находится в премиум-версии ЭБС IPR BOOKS. - ISBN 978-5-4497-0362-0, экземпляров неограничено
3. Яковлева, А. В. Эконометрика Электронный ресурс : Учебное пособие / А. В. Яковлева. - Эконометрика, 2020-02-05. - Саратов : Научная книга, 2019. - 223 с. - Книга находится в премиум-версии ЭБС IPR BOOKS. - ISBN 978-5-9758-1820-1, экземпляров неограничено

Дополнительная литература

1. Воскобойников, Ю. Е. Эконометрика в Excel. Модели временных рядов Электронный ресурс / Воскобойников Ю. Е. : учебное пособие. - 2-е изд., стер. - Санкт-Петербург : Лань, 2020. - 152 с. - ISBN 978-5-8114-4863-0, экземпляров неограничено
2. Ершова, Н. А. Современная эконометрика Электронный ресурс : Учебное пособие / Н. А. Ершова, С. Н. Павлов. - Современная эконометрика, 2021-09-11. - Москва : Российский государственный университет правосудия, 2018. - 52 с. - Книга находится в премиум-версии ЭБС IPR BOOKS. - ISBN 978-5-93916-650-8, экземпляров неограничено
3. Кузнецова, О. А. Эконометрика (продвинутый уровень) Электронный ресурс / Кузнецова О. А. : учебно-методическое пособие. - Тольятти : ТГУ, 2020. - 125 с. - ISBN 978-5-8259-1525-8, экземпляров неограничено

Интернет-ресурсы

1. Информационная справочная система ГАРАНТ.РУ // Режим доступа: <http://www.garant.ru/>
2. Информационная справочная система КонсультантПлюс. // Режим доступа: <http://www.consultant.ru>
3. Профессиональная база данных «Всероссийская система данных о компаниях и бизнесе «За честный бизнес» // Режим доступа: <https://zachestnyibiznes.ru>
4. Профессиональная база данных Росстата // Режим доступа: Росстат — Базы данных (<https://rosstat.gov.ru>)
5. Профессиональная база данных Интерфакс «СПАРК» // Режим доступа: <https://spark-interfax.ru/>