



ДОНСКОЙ ГОСУДАРСТВЕННЫЙ ТЕХНИЧЕСКИЙ УНИВЕРСИТЕТ

УПРАВЛЕНИЕ ЦИФРОВЫХ ОБРАЗОВАТЕЛЬНЫХ ТЕХНОЛОГИЙ  
ПИ (филиал) ДГТУ в г. Таганроге

Кафедра «Гуманитарные и социально-экономические науки»

## **ПРАКТИКУМ**

по выполнению практических работ  
по дисциплине

## **«ЭКОНОМЕТРИКА»**

Автор  
Чернова Т.В.

Таганрог, 2025

## Аннотация

Практикум по выполнению практических работ по дисциплине «Эконометрика». ПИ (филиал) ДГТУ в г. Таганроге, 2025 г.

В практикуме кратко изложены теоретические вопросы, необходимые для успешного выполнения практических работ, рабочее задание для самостоятельной работы и тестовые задания для самопроверки.

Предназначено для обучающихся по направлению подготовки 38.03.01 Экономика.

## Авторы

д-р экон. наук, доцент кафедры «Гуманитарные и социально-экономические науки»

Чернова Т.В.



## Оглавление

Введение	4
1.Теоретическая часть	5
1.1 Предмет и основные понятия эконометрики	5
1.2 Введение в корреляционно-регрессионный анализ	9
1.3 Спецификация эконометрической модели	17
1.4 Множественная регрессия и корреляция	20
1.5 Моделирование одномерных временных рядов	22
1.6 Изучение сезонных колебаний	28
2.Практическая часть	34
2.1 Задачи	34
2.2 Индивидуальные задания	36
2.3 Тестовые задания для самоконтроля	44
3.Перечень использованных информационных ресурсов	47

## ВВЕДЕНИЕ

Эконометрика – одна из базовых дисциплин экономического образования во всем мире. Эконометрика представляет собой самостоятельную научную дисциплину, которая объединяет совокупность теоретических результатов, приемов, методов и моделей, предназначенных для того, чтобы на базе экономической теории, экономической статистики и экономических измерений, математико-статистического инструментария придавать конкретное количественное выражение общим (качественным) закономерностям, обусловленным экономической теорией.

Практикум содержит теоретическую часть, где представлены краткие сведения по каждой теме курса и формулы для расчета эконометрических показателей и построения графиков. В практической части представлены задачи и тестовые задания.

## I. ТЕОРЕТИЧЕСКАЯ ЧАСТЬ

### 1.1 ПРЕДМЕТ И ОСНОВНЫЕ ПОНЯТИЯ ЭКОНОМЕТРИКИ

Эконометрика представляет собой комбинацию двух слов: «экономика» и «метрика» (от греческого «метрон», «метрия», т.е. измерение. Измерение предполагает получение численного значения). Сам термин подчеркивает специфику, содержание эконометрики как науки: ***количественное выражение тех связей и соотношений, которые раскрыты и обоснованы экономической теорией.*** Наука сравнительно молодая, стала интенсивно развиваться с середины 20 века. Зарождение эконометрики является следствием междисциплинарного подхода к изучению экономики. Название этой науки дал норвежский ученый Р.Фриш. Эта наука возникла в результате взаимодействия и объединения в особый «сплав» трех компонент: экономической теории, статистики и математики (статистических и математических методов).

***Эконометрика даёт количественное выражение взаимосвязей экономических явлений и процессов, занимается определением наблюдаемых в экономической жизни конкретных количественных закономерностей, применяя для этой цели статистические и математические методы.***

Цель эконометрики – модельное описание конкретных количественных взаимосвязей, обусловленных закономерностями развития общественных явлений, изучаемых экономической теорией. Для построения модели необходимы:

- гипотеза, т.е. теоретическое описание процесса;
- данные, т.е. информационная база;
- методы обработки (модели, расчеты и т.д.).

Задачи и область применения эконометрических моделей:

1. прогнозирование экономических и социально-экономических показателей, характеризующих анализируемую систему;
2. имитация различных сценариев развития анализируемой системы.

В качестве системы может выступать:

- страна/страны – макроэкономический уровень;
- регионы, отрасли, корпорации – мезоуровень;
- индивиды, отдельные предприятия, фирмы, домохозяйства – микроуровень.

Для построения эконометрических моделей используются следующие виды переменных:

- экзогенные – переменные, значение которых определяется вне модели, т.е. заранее известны, заданы ( $x$ );
- лаговые – переменные, значения которых измерены в предыдущие моменты времени;

– предопределенные – переменные, объединяющие экзогенные и лаговые переменные;

– эндогенные – значение этих переменных определяется по модели под влиянием экзогенных и лаговых переменных ( $y$ ).

Для описания сущности эконометрической модели удобно разбить весь процесс моделирования на шесть основных этапов:

1-й этап (постановочный) – определение конечных целей моделирования, набора участвующих в модели факторов и показателей, их роли;

2-й этап (априорный) – предмодельный анализ экономической сущности изучаемого явления, формирование и формализация априорной информации, в частности, относящейся к природе и генезису исходных статистических данных и случайных остаточных составляющих;

3-й этап (параметризация) – собственно моделирование, т.е. выбор общего вида модели, в том числе состава и формы входящих в нее связей;

4-й этап (информационный) – сбор необходимой статистической информации, т.е. регистрация значений участвующих в модели факторов и показателей на различных временных или пространственных тактах функционирования изучаемого явления;

5-й этап (идентификация модели) – статистический анализ модели и в первую очередь статистическое оценивание неизвестных параметров модели;

6-й этап (верификация модели) – сопоставление реальных и модельных данных, проверка адекватности модели, оценка точности модельных данных.

#### Типы данных, используемых в эконометрических моделях:

– Пространственные данные – собираются/регистрируются по нескольким объектам по ряду показателей в фиксированный момент времени (данные поперечного среза)

– Временные ряды – данные продольного среза, для таких данных существует порядок следования значений. Данные регистрируются по одному объекту по ряду показателей через определенный интервал

– Панельные данные (пространственно-временные данные). Регистрируются по каждому из объектов по ряду показателей в течение некоторого периода времени, как правило, число объектов велико, а число периодов мало.

#### Основные понятия, виды и формы связей, различаемые в статистике

***Выявление и оценка влияния наиболее существенных факторов, а также изучение воздействия факторных признаков на резульативные показатели является одной из основных задач эконометрики.*** Разработка эконометрической модели, учитывающей закономерности развития изучаемого явления позволяет в дальнейшем управлять социально-экономическими процессами и предсказывать их развитие.

Формы проявления факторных связей весьма разнообразны. В качестве двух самых общих их видов выделяют функциональную или жестко детерминированную и статистическую или стохастически детерминированную связи.

Функциональная связь (жестко детерминированная) – это вид причинной зависимости, при которой определенному значению факторного признака строго соответствует одно или несколько значений результативного признака. Чаще всего функциональные связи проявляются в физике, химии и других точных науках. В экономике примером может служить прямо пропорциональная зависимость между производительностью труда и увеличением производства продукции.

Характерной особенностью функциональной связи является то, что она проявляется с одинаковой силой у каждой единицы изучаемой совокупности. Поэтому знание функциональной зависимости позволяет абсолютно точно прогнозировать события, например, наступление солнечных или лунных затмений.

Стохастическая связь (которую также называют неполной, или статистической) проявляется не в каждом отдельном случае, а в среднем, при массовых наблюдениях.

Термин «стохастический» происходит от греческого «stochos» – мишень или бычий глаз. Стреляя в мишень, даже хороший стрелок редко попадает в центр – «яблочко», выстрелы ложатся в некоторой, близкой к нему области. В этом смысле стохастическая связь означает приблизительный характер значений признака. Обычно стохастическая связь между двумя случайными величинами имеет место в случае наличия общих случайных факторов, воздействующих на каждую из них.

Например, зависимость цены товара от его качества. В отдельных случаях соотношение спроса и предложения может привести к тому, что товар худшего качества будет продан по более высокой цене, а аналогичный товар лучшего качества может иметь более низкую цену при достаточно большом объеме продаж. Однако же в среднем между ценой и качеством товара существует прямая зависимость – чем выше качество товара, тем выше его цена и, соответственно, чем ниже качество товара, тем ниже будет и цена на него.

Среди взаимосвязанных признаков одни могут рассматриваться как факторные, а другие как результативные признаки. Признак, от которого зависит другой, принято называть факторным, зависимый же признак называют результативным. Часто рассматривают зависимость признака не от одного, а от нескольких факторных признаков.

При изучении массовых общественно-экономических явлений между факторными признаками проявляется корреляционная связь, т.е. такая связь, при которой на величину результативного признака одновременно или последовательно

оказывают влияние, помимо факторного, множество других признаков, действующих в различных направлениях.

**Корреляционная связь (от английского слова *correlation* – соотношение, соответствие) – частный случай стохастической связи, которая выражается в том, что при определенном значении одной переменной величины (независимая переменная – аргумент ( $x$ )) другая переменная величина (зависимая переменная – функция ( $y$ )) принимает определенное значение.**

Корреляционную связь можно выявить только в виде общей тенденции при массовом сопоставлении фактов. При этом каждому значению факторного признака будет соответствовать не одно определенное значение результативного признака, а целая совокупность их.

Это обуславливается тем, что изменение результативного признака происходит под влиянием рассматриваемого факторного признака не всецело, а лишь частично, так как в каждом конкретном случае (факте) возможно влияние ряда неучтенных или неконтролируемых (случайных) факторов. Объяснением тому является сложность взаимосвязей между анализируемыми факторами, поэтому связь между признаками проявляется лишь в среднем, в массе случаев, т.е. с изменением факторного признака ( $x$ ) закономерным образом изменяется среднее значение результативного признака ( $y$ ), в то время как в каждом отдельном случае результативный признак может принимать множество различных значений.

Например, в сельском хозяйстве это может быть связь между урожайностью и количеством внесенных удобрений. Очевидно, что последние участвуют в формировании урожая, но для каждого поля одно и то же количество внесенных удобрений вызовет разный прирост урожайности, так как во взаимодействии находится еще целый ряд факторов (погода, состояние почвы, количество осадков и др.), которые и формируют конечный результат. Однако в среднем такая связь наблюдается, т.е. увеличение массы внесенных удобрений ведет к росту урожайности.

### **Виды и формы корреляционных связей**

**По направлению** связи бывают прямыми, когда зависимая переменная растет с увеличением факторного признака, и обратными, при которых рост последнего сопровождается уменьшением функции. Такие связи также можно назвать, соответственно, положительными и отрицательными.

**Относительно своей аналитической формы** связи бывают линейными и нелинейными. В первом случае между признаками, в среднем, проявляются линейные соотношения. Нелинейная взаимосвязь выражается нелинейной функцией, а переменные связаны между собой в среднем нелинейно.

Существует еще одна достаточно важная характеристика связей **с точки зрения взаимодействующих факторов.** Если характеризуется связь двух



признаков, то ее принято называть парной. Если изучаются более чем две переменные – множественной.

Указанные выше классификационные признаки наиболее часто встречаются в статистическом анализе, но кроме перечисленных различают *также непосредственные, косвенные и ложные связи*. Собственно, суть каждой очевидна из её названия. В первом случае факторы взаимодействуют между собой непосредственно. Для косвенной связи характерно участие какой-то третьей переменной, которая опосредует связь между изучаемыми признаками. Ложная связь – это связь, установленная формально и, как правило, подтвержденная только количественными оценками. Она не имеет под собой качественной основы или же бессмысленна.

**По силе различаются слабые и сильные связи.** Эта формальная характеристика выражается конкретными величинами и интерпретируется в соответствии с общепринятыми критериями силы связи для конкретных показателей.

Изучение корреляционных связей сводится в основном к решению следующих задач:

- выявление наличия (или отсутствия) корреляционной связи между изучаемыми признаками;
- измерение степени тесноты связи между признаками;
- нахождение аналитического выражения связи, отражающей зависимость между результативным и факторным признаками;
- экономическая интерпретация и практическое использование полученного результата.

В начальной стадии анализа статистических зависимостей применяются простейшие методы наличия связи, ее направления и характера, выявляется форма воздействия одних признаков на другие.

## 1.2 ВВЕДЕНИЕ В КОРРЕЛЯЦИОННО-РЕГРЕССИОННЫЙ АНАЛИЗ

Задача эконометрики в области изучения взаимосвязей состоит в построении модели, т.е. нужно найти аналитическое выражение корреляционной связи, отражающей зависимость между результативным и факторным признаками; дать экономическую интерпретацию результатов моделирования и возможностей практического использования полученного результата.

При количественной оценки связи, для характеристики силы и формы влияния одних признаков на другие применяются две группы методов, одна из которых включает методы корреляционного анализа, другая – регрессионный анализ. Ряд исследователей объединяет эти методы в корреляционно-регрессионный анализ, что имеет под собой некоторые основания: наличие целого ряда общих вычислительных процедур, взаимодополнения при интерпретации

результатов и др.

Задачи собственно **корреляционного анализа** сводятся к измерению тесноты связи между варьирующими признаками, определению неизвестных причинных связей и оценке факторов, оказывающих наибольшее влияние на результативный признак.

Задачи **регрессионного анализа** лежат в сфере установления формы зависимости, определения функции регрессии, использования уравнения для оценки неизвестных значений зависимой переменной.

Решение названных задач опирается на соответствующие приемы, алгоритмы, показатели, применение которых дает основание говорить о статистическом изучении взаимосвязей.

Статистические методы выявления наличия корреляционной связи между признаками

**1 – метод параллельного сопоставления рядов.** Сущность метода заключается в том, что значения факторного признака располагают в возрастающем порядке и затем прослеживают направление изменения величины результативного признака. В тех случаях, когда возрастание величины факторного признака влечет за собой возрастание и величины результативного признака, можно говорить о возможном наличии прямой корреляционной связи. Если же с увеличением факторного признака результативный имеет тенденцию к уменьшению, то можно предполагать обратную связь между признаками.

Допустим, есть данные о выпуске продукции на 6 однотипных предприятиях ( $x$ ) и потреблении на них электричества ( $y$ ). Сравним изменения двух величин и если их вариация согласована, то можно сделать вывод о наличии связи:

Выпуск продукции ( $x$ )	5	7	10	12	15	17
Потребление электричества ( $y$ )	17	22	26	24	30	42

Таблица наглядно демонстрирует, что с увеличением  $x$  возрастает и  $y$ , поэтому связь между ними можно считать прямой.

Однако при большом числе различных значений результативного признака, соответствующих одному и тому же значению признака-фактора, параллельные ряды плохо воспринимаются, особенно в больших совокупностях. В подобных случаях для установления факта наличия связи лучше воспользоваться методом группировок.

**2 – метод аналитических группировок.** При аналитической группировке исследуется связь между двумя и более признаками, из которых один рассматривается как результат, а другой как фактор (факторных признаков может быть несколько). Для установления связи между признаками данные группируются по

признаку-фактору, который располагается в подлежащем аналитической таблицы. Если результативный признак зависит от факторного, то в изменении факторного признака и среднего значения результативного будет обнаруживаться определенная закономерность.

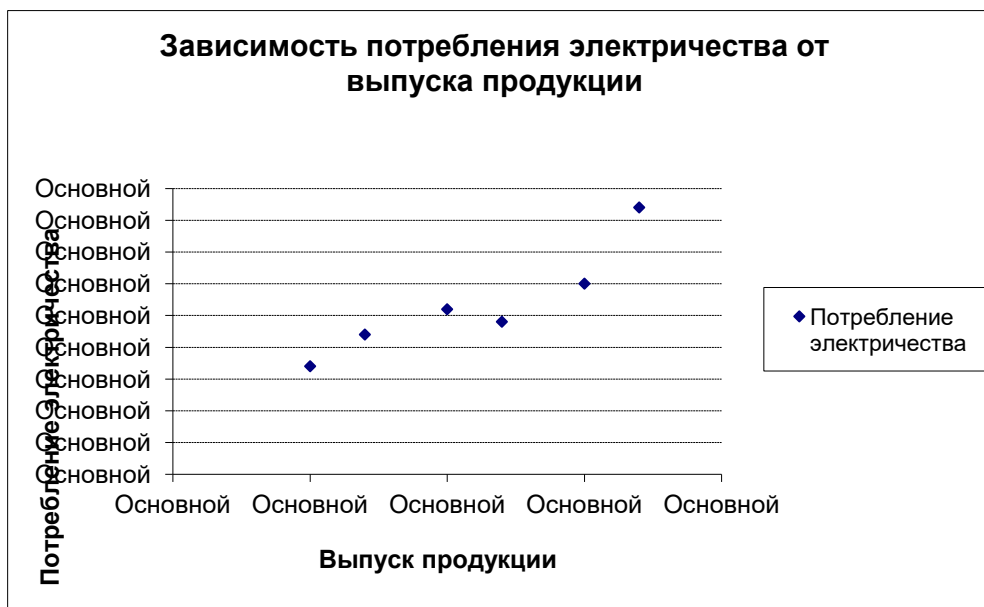
Таблица 1 – Характеристика зависимости прибыли малых предприятий РО от оборачиваемости оборотных средств по состоянию на конец отчетного года

Продолжительность оборота средств, в днях ( $x$ )	Число малых предприятий, ед.	Средняя прибыль, у.е. ( $y$ )
40 - 50	6	14,57
51 - 70	8	12,95
71 - 101	6	7,40
Итого	20	11,77

Например, связь между продолжительностью оборота оборотных средств и размером прибыли предприятия можно легко обнаружить, сгруппировав предприятия по длительности оборота оборотных средств и рассчитав среднюю прибыль по каждой группе предприятий. Оборачиваемость в днях – фактор, обозначаемый обычно  $x$ , а прибыль – результат –  $y$ . Таблица ясно демонстрирует присутствие обратной (отрицательной) связи между признаками.

Говоря об использовании метода группировок для выявления корреляционной зависимости, следует иметь в виду, что это только прием, при помощи которого можно иллюстрировать наличие или отсутствие связи.

**3 –графический метод** используется для наглядного изображения формы связи между изучаемыми признаками. Для этого в прямоугольных осях координат строят график, по оси ординат которого откладывают индивидуальные значения результативного признака –  $y$ , а по оси абсцисс – индивидуальные значения факторного признака –  $x$ . Точками показывается сочетание  $x$  и  $y$ . По расположению точек, их концентрации в определенном направлении можно судить о наличии связи и ее приблизительной форме. Полученная совокупность точек называется *полем корреляции*.



График, построенный по индивидуальным значениям признаков примера, приводимого для параллельного сопоставления рядов, подтверждает то, что связь носит приблизительно линейный характер. По существу, рассмотренные методы характеризуют взаимосвязь, когда выбраны факторный и результативный признаки и требуется сформулировать предположения о форме и направленности связи. В то же время количественная оценка тесноты связи требует дополнительных расчетов и более глубокого ее изучения, и анализа.

Измерение тесноты связи в случае наличия корреляционной зависимости осуществляется на основе применения специальных коэффициентов. Один из простейших показателей тесноты связи разработан австрийским психиатром Г.Фехнером:

**Ошибка! Объект не может быть создан из кодов полей редактирования.,** где  $n_a$  и  $n_b$  – соответственно количество совпадений и несовпадений знаков отклонений.

Показатель Фехнера изменяется от  $-1$  до  $1$ , при значении равном  $1$  он указывает на полную прямую связь, при значении  $-1$  на полную обратную связь, при  $i=0$  связь отсутствует. Промежуточные значения  $i$  характеризуют степень близости связи к функциональной. Например, по данным о выпуске продукции ( $x$ ) на 6 однотипных предприятиях и потреблении на них электроэнергии ( $y$ ), рассчитаем средние значения для  $x$  и  $y$ :

**Ошибка! Объект не может быть создан из кодов полей редактирования.;**

**Ошибка! Объект не может быть создан из кодов полей редактирования..**

Тогда расчет коэффициента Фехнера для оценки степени зависимости потребления электричества от объема выпуска продукции будет следующим

Выпуск продукции	5	7	10	12	15	17
Потребление электричества	17	22	26	24	30	42

Ошибка! Объект не может быть создан из кодов полей редактирования.	-6	-4	-1	1	4	6
Ошибка! Объект не может быть создан из кодов полей редактирования.	-9,83	-4,83	-0,83	-2,83	3,17	15,17

**Ошибка! Объект не может быть создан из кодов полей редактирования..**

Судя по полученному значению коэффициента, связь можно считать достаточно сильной. Недостаток показателя Фехнера состоит в том, что разные по абсолютной величине отклонения имеют одинаковый вес. Более совершенным показателем степени тесноты связи является линейный коэффициент корреляции.

Самый популярный измеритель тесноты связи между признаками – линейный коэффициент корреляции Пирсона (назван по имени английского статистика К. Пирсона, который ввел это коэффициент в научный анализ). Данный коэффициент характеризует тесноту и направление связи между двумя коррелируемыми признаками в случае наличия между ними линейной зависимости.

Что понять смысл (содержание) этого показателя, воспользуемся для расчета данными вышеприведенного примера. Запишем расчеты в виде среднего значения совместных отклонений признаков от их средних значений:

**Ошибка! Объект не может быть создан из кодов полей редактирования.Ошибка! Объект не может быть создан из кодов полей редактирования.**

Полученное значение называется **коэффициентом ковариации**. Это – мера совместной вариации признаков, т.е. можно сказать, что это мера соответствия вариации результативного признака вариации факторного. Недостатком коэффициента ковариации является то, что он не нормирован. Для преодоления этого недостатка можно полученное выражение разделить на среднее квадратическое отклонение по  $x$  и по  $y$  **Ошибка! Объект не может быть создан из кодов полей редактирования.:**

**Ошибка! Объект не может быть создан из кодов полей редактирования..**

Полученное значение есть **линейный коэффициент корреляции**, показатель интенсивности линейной связи, величина безразмерная, которая изменяется в интервале **Ошибка! Объект не может быть создан из кодов полей редактирования..** Существует эмпирическое правило (шкала Чеддока), согласно которому принято считать, что если  $r$  находится в пределах  $0,100 - 0,300$ , то связь слабая; при  $r = 0,300 \div 0,500$  – умеренная;  $0,500 - 0,700$  связь заметная; при  $r > 0,700$  - сильная, или тесная. Когда  $r = 1$  - связь функциональная. Если же  $r \approx 0$ , то это дает основание говорить об отсутствии линейной связи между  $x$  и  $y$ .

Путем ряда преобразований можно получить следующие аналитические выражения для коэффициента корреляции:

**Ошибка! Объект не может быть создан из кодов полей редактирования.;**

Ошибка! Объект не может быть создан из кодов полей редактирования..

Производя расчет по итоговым значениям исходных переменных, линейный коэффициент корреляции можно вычислить по формуле:

**Ошибка! Объект не может быть создан из кодов полей редактирования..**

Линейный коэффициент корреляции имеет большое значение при исследовании социально-экономических явлений и процессов, распределения которых близки к нормальному.

Непараметрические методы оценки связи. Методы корреляционного и дисперсионного анализа не универсальны: их можно применять, если все изучаемые признаки являются количественными. При использовании этих методов нельзя обойтись без вычисления основных параметров распределения (средних величин, дисперсий), поэтому они получили название параметрических методов. Между тем в статистической практике изучения общественных явлений приходится сталкиваться с задачами измерения связи между качественными признаками, к которым параметрические методы анализа в их обычном виде неприменимы. Статистической наукой разработаны методы, с помощью которых можно измерить связь между явлениями, не используя при этом количественные значения признака, а значит, и параметры распределения. Такие методы получили название непараметрических.

Например, измерение связи между альтернативными признаками, т.е. признаками, принимающими только два возможных значения (дихотомические). Исследования их корреляции основано на показателях, построенных на четырехклеточных таблицах, в которые сводятся значения признаков.

<i>A</i>	<i>B</i>
<i>C</i>	<i>D</i>

Допустим, требуется измерить связь между прививками от гриппа и пониженной заболеваемостью от гриппа в группе случайно выбранных студентов.

	Заболели	Не заболели	Итого
Привитые	30	20	50
Не привитые	15	5	20
Итого	45	25	70

Существует ли связь между заболеваемостью и прививками? Для измерения тесноты взаимосвязи признаков производится расчет коэффициента контингенции по формуле:

**Ошибка! Объект не может быть создан из кодов полей редактирования..**

Коэффициент контингенции принимает значения в интервале от  $-1$  до  $1$ , а интерпретация аналогична коэффициенту корреляции. Судя по значению коэффициента связь между заболеваемостью и прививками слабая отрицательная.

Другой метод измерения связи основан на расчете коэффициента ассоциации:

**Ошибка! Объект не может быть создан из кодов полей редактирования..**

Минус перед коэффициентом говорит об обратном направлении связи, т.е. чем больше прививок, тем меньше заболеваний. Коэффициент контингенции всегда бывает меньше коэффициента ассоциации и дает более корректную оценку тесноты связи.

Парная линейная регрессия. Для характеристики влияния изменений  $x$  на вариацию  $y$  служат методы регрессионного анализа. *Регрессия – это линия, показывающая вид зависимости средней результативного признака от факторного.* Термин, введенный английским статистиком Гальтоном при изучении наследственности. Уравнение регрессии является аналитическим уравнением, с помощью которого выражается связь между признаками. В случае парной линейной зависимости строится регрессионная модель

**Ошибка! Объект не может быть создан из кодов полей редактирования.,** где  $n$  - число наблюдений;  $a_0, a_1$  - неизвестные параметры уравнения;  $\varepsilon_i$  - ошибка случайной переменной  $Y$ .

Уравнение регрессии записывается как:

$$Y_{i\text{теор}} = a_0 + a_1 X_i,$$

где  $Y_{i\text{теор}}$  - рассчитанное выравненное значение результативного признака после подстановки в уравнение  $X$ .

Параметры  $a_0$  и  $a_1$  оцениваются с помощью процедур, наибольшее распространение из которых получил *метод наименьших квадратов.* Линия должна быть проведена так, чтобы сумма отклонений точек поля корреляции от соответствующих точек теоретической линии регрессии равнялась нулю, а сумма квадратов этих отклонений была бы минимальной величиной. Значит, наилучшие оценки  $a_0$  и  $a_1$ , получают, когда

**Ошибка! Объект не может быть создан из кодов полей редактирования.**

т.е. сумма квадратов отклонений эмпирических значений зависимой переменной от вычисленных по уравнению регрессии должна быть минимальной. Сумма квадратов отклонений является функцией параметров  $a_0$  и  $a_1$ . Ее минимизация осуществляется решением системы уравнений

**Ошибка! Объект не может быть создан из кодов полей редактирования.;**

**Ошибка! Объект не может быть создан из кодов полей редактирования.;**

**Ошибка! Объект не может быть создан из кодов полей редактирования..**

Важен смысл параметров:  $a_1$  – это коэффициент регрессии, характеризующий влияние, которое оказывает изменение  $X$  на  $Y$ . Он показывает, на сколько единиц, в среднем, изменится  $Y$  при изменении  $X$  на одну единицу. Если  $a$  больше 0, то наблюдается положительная связь. Если  $a$  имеет отрицательное значение, то увеличение  $X$  на единицу влечет за собой уменьшение  $Y$  в среднем на  $a_1$ . Параметр  $a_1$  обладает размерностью отношения  $Y$  к  $X$ .

Параметр  $a_0$  – это постоянная величина в уравнении регрессии. На наш взгляд, экономического смысла он не имеет, но в ряде случаев его интерпретируют как начальное значение  $Y$ .

Значение функции  $Y_{\text{теор}} = a_0 + a_1X$  называется расчетным значением и на графике образует *теоретическую линию регрессии*. Смысл теоретической регрессии в том, что это оценка среднего значения переменной  $Y$  для заданного значения  $X$ .

Например, рассмотрим построение однофакторного уравнения регрессии зависимости производительности труда  $Y$  от стажа работы  $X$  по представленным данным о бригаде из десяти рабочих, занятых производством электротехнических изделий. Для расчета коэффициентов уравнения регрессии воспользуемся расчетной таблицей:

Исходные данные			Расчетные значения			
Номер рабочего	Стаж работы, лет $X$	Выработка рабочего, шт. в смену $Y$	$X^2$	$Y^2$	$X Y$	$Y_{\text{теор}}$
1	4	7	16	49	28	6,4
2	5	7	25	49	35	7,0
3	3	6	9	36	18	5,8
4	1	4	1	16	4	4,6
5	10	9	100	81	90	10,0
6	2	5	4	25	10	5,2
7	6	8	36	64	48	7,6
8	9	10	81	100	90	9,4
9	7	8	49	64	56	8,2
10	8	9	64	81	72	8,8
Итого	55	73	385	565	451	73,0

Расчетные значения подставим в формулу и получим параметры для данного уравнения регрессии:

Ошибка! Объект не может быть создан из кодов полей редактирования.;

Ошибка! Объект не может быть создан из кодов полей редактирования..

Следовательно, регрессионная модель распределения выработки по стажу работы для данного примера имеет вид:  $Y_{\text{теор}} = 4 + 0.6X$ . Это уравнение



характеризует зависимость среднего уровня выработки рабочим бригады от стажа работы.

### **Средняя ошибка аппроксимации**

Фактические значения результативного признака  $y_i$  отличаются от теоретических, рассчитанных по уравнению регрессии  $y_{\text{теор}}$ . Величина отклонений по каждому наблюдению представляет собой ошибку аппроксимации. Их число соответствует объему совокупности ( $k$ ). В отдельных случаях ошибка аппроксимации может оказаться равной нулю. Для сравнения используются величины отклонений, выраженные в процентах к фактическим значениям.

Отклонения  $(y_i - y_{\text{теор}})$  можно рассматривать как абсолютную ошибку аппроксимации, а  $(y_i - y_{\text{теор}})/y_i \times 100\%$  – как относительную ошибку аппроксимации.

Чтобы иметь общее суждение о качестве модели из относительных отклонений по каждому наблюдению, определяют среднюю ошибку аппроксимации как среднюю арифметическую простую:

$$A = \frac{1}{k} \sum_{i=1}^k \left| \frac{y_i - \hat{y}_i}{y_i} \right| \cdot 100\%$$

Ошибка аппроксимации в пределах 5 – 7 % свидетельствует о хорошем подборе модели к исходным данным. *Допустимый* предел значений  $A$  – не более 8 – 10% (допускается 8– 15 %).

Возможно и иное определение средней ошибки аппроксимации:

$$A = \frac{100\%}{\bar{y}} \sqrt{\frac{\sum_{i=1}^k (y_i - \hat{y}_i)^2}{k}}$$

## **1.3 СПЕЦИФИКАЦИЯ ЭКОНОМЕТРИЧЕСКОЙ МОДЕЛИ**

В зависимости от количества факторов, включенных в уравнение регрессии, принято различать простую (парную) и множественную регрессии.

**Простая регрессия** представляет собой регрессию между двумя переменными –  $x$  и  $y$ , т.е. модель вида  $Y = f(x)$

где:  $Y$  – зависимая переменная (результативный признак);  $x$  – независимая, или объясняющая, переменная (признак-фактор).

Множественная регрессия соответственно представляет собой регрессию результативного признака с двумя и большим числом факторов, т.е. модель вида

$$Y = f(x_1, x_2, \dots, x_k).$$

Любое эконометрическое исследование начинается со спецификации модели, т.е. с формулировки вида модели, исходя из соответствующей теории связи между переменными. Нужно определить наличие связи между явлениями.

Прежде всего из всего круга факторов, влияющих на результативный признак, необходимо выделить наиболее существенно влияющие факторы. Парная регрессия достаточна, если имеется доминирующий фактор, который и используется в качестве объясняющей переменной.

Предположим, что выдвигается гипотеза о том, что величина спроса  $U$  на товар  $A$  находится в обратной зависимости от цены  $X$ , т.е.  $Y_x = a - bx$ ,

Уравнение простой регрессии характеризует связь между двумя переменными, которая проявляется как некоторая закономерность лишь в среднем в целом по совокупности наблюдений. Так, если зависимость спроса  $U$  от цены  $X$  характеризуется, например, уравнением  $U = 5000 - 2x$ , то это означает, что с ростом цены на 1 д.е. спрос в среднем уменьшается на 2 д.е. В уравнении регрессии корреляционная по сути связь признаков представляется в виде функциональной связи, выраженной соответствующей математической функцией.

Практически в каждом отдельном случае величина  $U$  складывается из двух слагаемых:  $Y_j = y_{xj} + \varepsilon_j$

где  $Y_j$  – фактическое значение результативного признака;  $y_{xj}$  – теоретическое значение результативного признака, найденное исходя из соответствующей математической функции связи  $U$  и  $X$ , т.е. из уравнения регрессии;  $\varepsilon_j$  – случайная величина, характеризующая отклонения реального значения результативного признака от теоретического, найденного по уравнению

**Случайная величина**  $\varepsilon_j$  называется также **возмущением**. Она включает влияние не учтенных в модели факторов, случайных ошибок и особенностей измерения. Её присутствие в модели порождено тремя источниками: спецификацией модели, выборочным характером исходных данных, особенностями измерения переменных. Приведенное ранее уравнение зависимости спроса  $U$  от цены  $X$  следует записывать как  $U = 5000 - 2x + \varepsilon$ .

Потому что всегда есть место для действия случайности. Обратная зависимость спроса от цены не обязательно характеризуется линейной функцией. От правильно выбранной спецификации модели зависит величина случайных ошибок: они тем меньше, чем в большей мере теоретические значения результативного признака  $U_x$ , подходят к фактическим данным  $y$ .

К ошибкам спецификации относятся:

- неправильный выбор той или иной математической функции для  $U_x$ ;
- недоучет в уравнении регрессии какого-либо существенного фактора, т.е. использование парной регрессии вместо множественной;
- ошибки выборки, которые имеют место в силу неоднородности данных в исходной статистической совокупности, что, как правило, бывает при изучении экономических процессов. Если совокупность неоднородна, то уравнение регрессии не имеет практического смысла. Для получения хорошего результата обычно

исключают из совокупности единицы с аномальными значениями исследуемых признаков. И в этом случае результаты регрессии представляют собой выборочные характеристики.

Использование временной информации также представляет собой выборку из всего множества хронологических дат. Изменив временной интервал, можно получить другие результаты регрессии.

Наибольшую опасность в практическом использовании методов регрессии представляют ошибки измерения. Если ошибки спецификации можно уменьшить, изменяя форму модели (вид математической формулы), а ошибки выборки – увеличивая объем исходных данных, то ошибки измерения практически сводят на нет все усилия, связанные с количественной оценкой связи между признаками.

Особенно велика роль ошибок измерения при исследовании на макроуровне. Так, в исследованиях спроса и потребления в качестве объясняющей переменной широко используется «доход на душу населения». Вместе с тем, статистическое измерение величины дохода сопряжено с рядом трудностей и не лишено возможных ошибок, например, в результате наличия скрытых доходов. Предполагая, что ошибки измерения сведены к минимуму, основное внимание в эконометрических исследованиях уделяется ошибкам спецификации модели.

В парной регрессии выбор вида математической функции  $Yx = f(x)$  может быть осуществлен тремя методами:

- 1) графическим;
- 2) аналитическим, т.е. исходя из теории изучаемой взаимосвязи;
- 3) экспериментальным.

При изучении зависимости между двумя признаками графический метод подбора вида уравнения регрессии достаточно нагляден. Он основан на поле корреляции. Основные типы кривых, используемые при количественной оценке связей, линейная, степенная, параболическая, гипербола, показательная, логистическая и др.

Значительный интерес представляет аналитический метод выбора типа уравнения регрессии. Он основан на изучении материальной природы связи исследуемых признаков.

Если уравнение регрессии проходит через все точки корреляционного поля, что возможно только при функциональной связи, когда все точки лежат на линии регрессии  $Yx = f(x)$ , то фактические значения результативного признака совпадают с теоретическими  $Y = Y_x$ , т.е. они полностью обусловлены влиянием фактора  $X$ . В этом случае остаточная дисперсия равна нулю ( $D_{ост} = 0$ ). В реальных условиях практических исследований, как правило, имеет место некоторое рассеяние точек относительно линии регрессии. Оно обусловлено влиянием прочих не учитываемых в уравнении регрессии факторов. Иными словами, имеют место отклонения

фактических данных от теоретических ( $Y - Y_x$ ). Величина этих отклонений и лежит в основе расчета остаточной дисперсии:

$$D_{ост} = 1/n \sum (y - y_x)^2$$

Чем меньше величина остаточной дисперсии, тем в меньшей мере наблюдается влияние прочих не учитываемых в уравнении регрессии факторов лучше уравнение регрессии подходит к исходным данным. При обработке статистических данных на компьютере перебираются разные математические функции в автоматическом режиме, и из них выбирается та, для которой остаточная дисперсия является наименьшей.

Если остаточная дисперсия оказывается примерно одинаковой для нескольких функций, то предпочтение лучше отдавать простым видам функций, поскольку они в большей степени поддаются интерпретации и требуют меньшего объема наблюдений. Результаты многих исследований подтверждают, что *число наблюдений должно в шесть – семь раз превышать число рассчитываемых параметров при переменной X. Это означает, что искать линейную регрессию, имея менее 7 наблюдений, вообще не имеет смысла. Если вид функции усложняется, то требуется увеличение объема наблюдений, ибо каждый параметр при X должен рассчитываться хотя бы по семи наблюдениям. Значит, если мы выбираем параболу второй степени, то требуется 14 наблюдений.*

#### 1.4 МНОЖЕСТВЕННАЯ РЕГРЕССИЯ И КОРРЕЛЯЦИЯ

Множественная регрессия широко используется в решении проблем спроса, доходности акций, при изучении издержек производства и т.п. В настоящее время множественная регрессия – один из наиболее распространенных методов в эконометрике.

Основная цель множественной регрессии – построить модель с большим числом факторов, определив при этом влияние каждого из них в отдельности, а также их совокупное воздействие на моделируемый показатель.

*Многофакторный корреляционный и регрессионный анализ позволяет оценить влияние на исследуемый результативный показатель каждого из включенных в модель (уравнение) факторов при фиксированном положении (на среднем уровне) остальных факторов, а также при любых возможных сочетаниях факторов с определенной степенью точности найти теоретическое значение этого показателя.*

Факторы, включаемые во множественную регрессию, должны отвечать следующим требованиям:

1. Они должны быть количественно измеримы. Если необходимо включить в модель качественный фактор, не имеющий количественного измерения, то ему нужно придать количественную определенность (например, в модели

урожайности качество почвы задается в виде баллов; в модели стоимости объектов недвижимости учитывается место нахождения недвижимости: районы могут быть проранжированы).

2. Факторы не должны коррелировать между собой (зарплата, производительность) и важным условием является отсутствие между факторами функциональной связи.

Если между факторами существует высокая корреляция, то нельзя определить их изолированное влияние на результирующий показатель и параметры уравнения регрессии оказываются неинтерпретируемыми. Включаемые во множественную регрессию факторы должны объяснить вариацию независимой переменной.

Таким образом, хотя теоретически регрессионная модель позволяет учесть любое число факторов, практически в этом нет необходимости. Отбор факторов производится на основе качественного теоретико-экономического анализа. Однако теоретический анализ часто не позволяет однозначно ответить на вопрос о количественной взаимосвязи рассматриваемых признаков и целесообразности включения фактора в модель. ***Поэтому отбор факторов обычно осуществляется в две стадии: на первой подбираются факторы исходя из сущности проблемы; на второй – на основе матрицы показателей корреляции определяют статистики для параметров регрессии.***

Наиболее широкое применение получили следующие методы построения уравнения множественной регрессии:

1. Метод исключения – отсев факторов из полного его набора.
2. Метод включения – дополнительное введение фактора.
3. Шаговый регрессионный анализ – исключение ранее введенного фактора.

*При отборе факторов также рекомендуется пользоваться следующим правилом: число включаемых факторов обычно в 6–7 раз меньше объема совокупности, по которой строится регрессия. Если это соотношение нарушено, то число степеней свободы остаточной дисперсии очень мало.*

Возможны разные виды уравнений множественной регрессии: линейные и нелинейные. Ввиду четкой интерпретации параметров наиболее широко используется линейная функция. В линейной множественной регрессии

$$y = a + a_1 x_1 + a_2 x_2 + \dots + a_m x_m$$

параметры при  $x$  называются коэффициентами «чистой» регрессии. Они характеризуют среднее изменение результата с изменением соответствующего фактора на единицу при неизменном значении других факторов, закрепленных на среднем уровне.

Классический подход к оцениванию параметров линейной модели множественной регрессии основан на методе наименьших квадратов (МНК). МНК позволяет получить такие оценки параметров, при которых сумма квадратов отклонений фактических значений результативного признака  $y$  от расчетных минимальна.

Для оценки взаимосвязи между признаком  $Y$  и несколькими факторными признаками  $X_i$  используется **коэффициент множественной корреляции** (совокупный коэффициент корреляции, коэффициент детерминации). Для случая двух факторных признаков  $x_1$  и  $x_2$  он имеет вид:

**Ошибка! Объект не может быть создан из кодов полей редактирования.**,  
где  $r$  – парные коэффициенты корреляции между признаками.

**Множественный коэффициент корреляции изменяется от 0 до 1 и численно не может быть меньше, чем любой из образующих его парных коэффициентов.** Приближение  $R$  к единице свидетельствует о сильной зависимости между признаками.

Если **Ошибка! Объект не может быть создан из кодов полей редактирования.**  $> 3$ , то с вероятностью близкой к 0,99 можно считать  $R$  значимым.

## 1.5 МОДЕЛИРОВАНИЕ ОДНОМЕРНЫХ ВРЕМЕННЫХ РЯДОВ

Помимо пространственных данных в эконометрических исследованиях используются и временные ряды. При построении эконометрической модели используются два типа временных данных:

- 1) данные, характеризующие совокупность различных объектов в определенный момент времени (например, товарооборот и прибыль семи магазинов, торгующих обувью, на 1 число отчетного месяца);
- 2) данные, характеризующие один объект за ряд последовательных моментов времени (товарооборот одного магазина за все месяцы отчетного года).

Модели, построенные по данным первого типа, называются *пространственными моделями*. Модели, построенные на основе второго типа данных, называются *моделями временных рядов*.

*Временной ряд (ряд динамики)* – это совокупность значений какого-либо показателя за несколько последовательных моментов или периодов времени. Каждый уровень временного ряда формируется под воздействием большого числа факторов, которые условно можно подразделить на три группы:

- 1) факторы, формирующие тенденцию ряда;
- 2) факторы, формирующие циклические колебания ряда;
- 3) случайные факторы.

В большинстве случаев фактический уровень временного ряда можно представить, как сумму или произведение трендовой, циклической и случайной

компонент. Модель, в которой временной ряд представлен как сумма перечисленных компонент, называется *аддитивной моделью* временного ряда. Модель, в которой временной ряд представлен как произведение перечисленных компонент, называется *мультипликативной моделью* временного ряда. Основная задача эконометрического исследования отдельного временного ряда – выявление и придание количественного выражения каждой из перечисленных выше компонент с тем, чтобы использовать полученную информацию для прогнозирования будущих значений ряда или при построении моделей взаимосвязи двух или более временных рядов.

Моделирование тенденции временного ряда. Распространенным способом моделирования тенденции временного ряда является построение аналитической функции, характеризующей зависимость уровней ряда от времени, или тренда. Этот способ называют *аналитическим выравниванием временного ряда*. Поскольку зависимость от времени может принимать разные формы, для ее формализации можно использовать различные виды функций. Для построения трендов чаще всего применяются следующие функции:

*линейный тренд*  $Y_{t_{теор}} = a_0 + a_1 t$ ,

– парабола второго порядка и полиномы различных степеней

$Y_t = a + a_1 t + a_2 t^2$ ;  $Y_t = a + a_1 t + a_2 t^2 + \dots + a_m t^m$

– равнобочная гиперболола  $Y_t = a + b/t$

– степенная  $Y_t = a t^b$

– показательная  $Y_t = a b^t$

Параметры каждого из перечисленных выше трендов можно определить обычным МНК, используя в качестве независимой переменной время  $t = 1, 2, \dots, n$ , а в качестве зависимой переменной – фактические уровни временного ряда  $y_t$ . Для нелинейных трендов предварительно проводят стандартную процедуру их линеаризации.

Существует несколько способов определения типа тенденции. К числу наиболее распространенных способов относятся качественный анализ изучаемого процесса, построение и визуальный анализ графика зависимости уровней ряда от времени. В этих же целях можно использовать и коэффициенты автокорреляции уровней ряда.

Тип тенденции можно определить путем сравнения коэффициентов автокорреляции первого порядка, рассчитанных по исходным и преобразованным уровням ряда. Если временной ряд имеет линейную тенденцию, то его соседние уровни  $y_t$  и  $y_{t-1}$  тесно коррелируют. В этом случае коэффициент автокорреляции первого порядка уровней исходного ряда должен быть высоким. Если временной ряд содержит нелинейную тенденцию, например, в форме экспоненты, то коэффициент автокорреляции первого порядка по логарифмам уровней исходного

ряда будет выше, чем соответствующий коэффициент, рассчитанный по уровням ряда. Чем сильнее выражена нелинейная тенденция в изучаемом временном ряде, тем в большей степени будут различаться значения указанных коэффициентов.

Выбор наилучшего уравнения в случае, когда ряд содержит нелинейную тенденцию, можно осуществить путем перебора основных форм тренда, расчета по каждому уравнению скорректированного коэффициента детерминации и средней ошибки аппроксимации. Этот метод легко реализуется при компьютерной обработке данных.

Аналитическое выравнивание. Основным содержанием данного метода является то, что основная тенденция развития изучаемого явления рассчитывается как функция времени. В итоге выравнивания временного ряда получают наиболее общий, суммарный, проявляющийся во времени результат действия всех влияющих факторов, а отклонение конкретных уровней ряда от уровней, соответствующих общей тенденции, объясняют действием факторов, проявляющихся случайно или периодически. В результате приходят к трендовой модели:  $Y_t = f(t) + \varepsilon_t$ , где  $f(t)$  – уровень, определяемый тенденцией развития;  $\varepsilon_t$  – случайное и циклическое отклонение от тенденции.

При этом определение теоретических, т.е. расчетных уровней временного ряда производится на основе так называемой *адекватной математической функции*, которая наилучшим образом отображает основную тенденцию изучаемого явления.

Целью аналитического выравнивания динамического ряда является определение аналитической или графической зависимости  $f(t)$ . На практике по имеющемуся временному ряду задают вид и находят параметры функции  $f(t)$ , а затем анализируют поведение отклонений от тенденции. Вид функции выбирают таким образом, чтобы она давала содержательное объяснение изучаемого процесса. На данном этапе очень важно произвести качественный анализ изучаемого явления. Чтобы решить проблему подбора математической функции, по которой рассчитываются теоретические уровни тренда, необходимо изучить природу исследуемого явления или процесса. От правильности решения этой проблемы зависит обоснованность выводов о закономерностях тренда.

Чаще всего при выравнивании используются следующие зависимости:

– линейная  $\hat{Y}_t = a_0 + a_1 t$ ;

– параболическая  $\hat{Y}_t = a_0 + a_1 t + a_2 t^2$ ;

– экспоненциальные  $\hat{Y}_t = \exp(a_0 + a_1 t)$  или  $\hat{Y}_t = \exp(a_0 + a_1 t + a_2 t^2)$ .

*Линейная зависимость* выбирается в тех случаях, когда в исходном временном ряду наблюдаются более или менее постоянные абсолютные цепные



приросты, не проявляющие тенденции ни к увеличению, ни к снижению.

*Параболическая зависимость* используется, если абсолютные цепные приросты сами по себе обнаруживают некоторую тенденцию развития, но абсолютные цепные приросты абсолютных цепных приростов (разности второго порядка) никакой тенденции развития не проявляют.

*Экспоненциальные зависимости* применяются, если в исходном временном ряду наблюдается либо более или менее постоянный относительный рост (устойчивость цепных темпов роста, темпов прироста, коэффициентов роста), либо, при отсутствии такого постоянства, - устойчивость в изменении показателей относительного роста (цепных темпов роста цепных же темпов роста, цепных коэффициентов роста цепных же коэффициентов или темпов роста и т. п.).

Наилучшим методом оценки параметров ( $a_0, a_1, a_2, \dots$ ) можно признать метод наименьших квадратов (МНК), который обеспечивает наименьшую сумму квадратов отклонений фактических уровней от выравненных:

$$\sum (\hat{Y}_t - y_i)^2 \rightarrow \min$$

В качестве примера произведем выравнивание данных о производстве зерна в России по уравнению прямой ( $\hat{Y}_t = a_0 + a_1 t$ ), используя два метода. Это метод наименьших квадратов и упрощенный метод выравнивания.

При использовании метода наименьших квадратов необходимо, чтобы сумма квадратов отклонений фактических данных от выравненных была наименьшей. Для нахождения параметров  $a_0$  и  $a_1$ , строится система нормальных уравнений, которая имеет вид:

**Ошибка! Объект не может быть создан из кодов полей редактирования.**

В рядах с равноотстоящими значениями можно выбрать упрощенный способ – способ условного отсчета времени ( $t$ ), так, чтобы  $\sum t=0$

Таблица исходных и расчетных данных

Годы	Производство зерна, млн. т. $Y$	Условные годы, $t$	$t^2$	$Y \cdot t$	$Y_t$
2014	10,0	- 4	16	- 40	9,3
2015	10,7	- 3	9	- 32,1	10,41
2016	12,0	- 2	4	- 24,0	11,52
2017	10,3	- 1	1	- 10,3	12,63
2018	12,9	0	0	0	13,74
2019	16,3	1	1	16,3	14,85
2020	15,6	2	4	31,2	15,96
2021	17,8	3	9	53,4	17,07
2022	18,0	4	16	72,0	18,18
Итого	123,6	0	60	66,5	123,66

Следовательно,

**Ошибка! Объект не может быть создан из кодов полей редактирования.**

млн. т.;

**Ошибка! Объект не может быть создан из кодов полей редактирования.**

млн. т.

Таким образом, уравнение прямой примет вид:  $Y_t = 13,74 + 1,11 t$ .

Подставив в это уравнение значение  $t$ , получим выровненные теоретические значения  $Y_t$ . Они показаны в последней колонке таблицы, причем общий объем производства зерна остался неизменным.

*Пояснения к таблице.* Первые две колонки – ряд динамики, подвергаемый выравниванию, дополняется колонкой, в которой показана система отсчета времени " $t$ ". Причем эта система выбирается таким образом, чтобы  $t = 0$ . Если число уровней ряда четное, то вместо нуля в центре мы поставили бы единицу с противоположными знаками у двух уровней, находящихся в середине ряда. Тогда разница между годами составляла бы две единицы времени и общий вид системы был бы таким (например, для ряда из 6 уровней):

2016	2017	2018	2019	2020	2021
- 5	- 3	- 1	+ 1	+ 3	+ 5

Параметры  $a_0$  и  $a_1$  можно исчислить иначе с помощью определителей, построенных на основе системы нормальных уравнений:

**Ошибка! Объект не может быть создан из кодов полей редактирования.**

**Ошибка! Объект не может быть создан из кодов полей редактирования.**

Приведенные формулы показывают, что для нахождения параметров  $a_0$  и  $a_1$  необходимо получить следующие значения: **Ошибка! Объект не может быть создан из кодов полей редактирования.**

Обозначив годы  $t$  порядковыми номерами, определим эти величины и представим их значения в таблице.

Расчет параметров  $a_0$  и  $a_1$  с помощью определителей

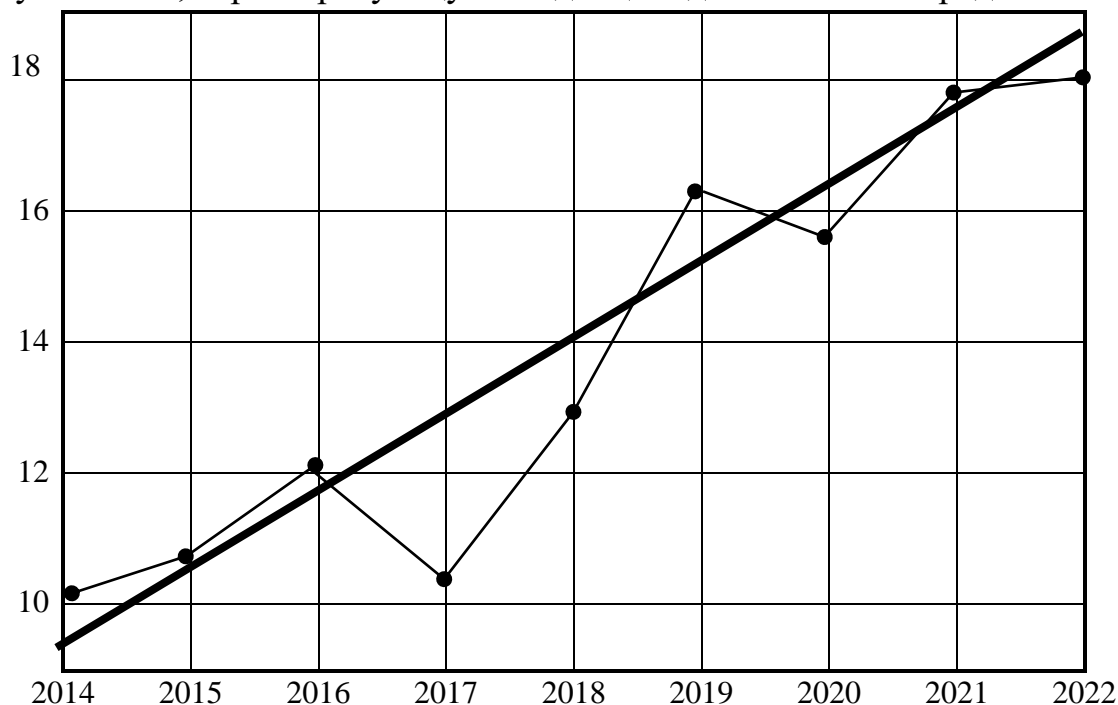
Годы	Производство зерна, млн т., $Y$	$T$	$t^2$	$Y \cdot t$	$Y_t$
2014	10,0	1	1	10	9,30
2015	10,7	2	4	21,4	10,41
2016	12,0	3	9	36,0	11,52
2017	10,3	4	16	41,2	12,63
2018	12,9	5	25	64,5	13,74
2019	16,3	6	36	97,8	14,85
2020	15,6	7	49	109,2	15,96
2021	17,8	8	64	142,4	17,07
2022	18,0	9	81	162	18,18
Итого	123,6	45	285	684,5	123,66

Далее определим параметры  $a_0$  и  $a_1$ :

**Ошибка! Объект не может быть создан из кодов полей редактирования.;**  
**Ошибка! Объект не может быть создан из кодов полей редактирования..**

Следовательно,  $Y = 8,19 + 1,66t$ .

После решения уравнения наносим на график фактические уровни и исчисленную прямую, характеризующую тенденцию динамического ряда.



При решении некоторых вопросов приходится определять неизвестные промежуточные значения динамического ряда. Эта задача решается способом интерполяции. **Интерполяция** – это нахождение по ряду данных значений функции промежуточных ее значений, что по сути есть приближенное отражение сложившейся закономерности внутри определенного отрезка времени. Однако **при прогнозировании** процессов и явлений необходимо по ряду данных значений функции найти другие ее значения, находящиеся вне этого ряда. Для этих целей используется экстраполяция. **Экстраполяция** заключается, по существу, в определении количественных характеристик для совокупностей и явлений, не подвергшихся наблюдению, путем распространения на них результатов, полученных из наблюдения над аналогичными совокупностями за прошедшее время, на будущее. Естественно, данные, получаемые путем экстраполяции ряда, следует рассматривать как вероятные оценки будущего, поскольку в действительности тенденция развития не остается неизменной.

## 1.6 ИЗУЧЕНИЕ СЕЗОННЫХ КОЛЕБАНИЙ

При сравнении квартальных и месячных данных многих социально-экономических явлений часто обнаруживаются периодические колебания, возникающие под влиянием смены времен года, обусловленные природно-климатическими, экономическими и социальными факторами. Такие специфические изменения проявляются в виде внутригодовых чередований подъемов и спадов выпуска продукции, колебаний производительности труда, объемами пассажироперевозок и т.п.

В широком понимании к сезонным относят все явления, которые обнаруживают в своем развитии отчетливо выраженную закономерность внутригодовых изменений, т.е. более или менее устойчиво повторяющиеся из года в год колебания уровней.

В статистике периодические колебания, которые имеют определенный и постоянный период, равный годовому промежутку свидетельствуют о наличии сезонных колебаний. Статистическое изучение сезонных колебаний имеет большое практическое значение, так как позволяет разрабатывать более квалифицированные и объективные планы по обеспечению договорных обязательств.

Существует ряд методов изучения и измерения сезонных колебаний. Самый простой метод заключается в построении специальных показателей, которые называются индексами сезонности.

Под индексами сезонности ( $I_s$ ) понимают процентное отношение фактических (эмпирических) внутригрупповых уровней к теоретическим (расчетным) уровням, выступающим в качестве базы сравнения.

$$I_s = \frac{\bar{y}_i}{\bar{y}} \cdot 100\%$$

Где **Ошибка! Объект не может быть создан из кодов полей редактирования.** - средняя величина уровня ряда  $i$ -го месяца за несколько лет; **Ошибка! Объект не может быть создан из кодов полей редактирования.** - среднемесячный уровень для всего временного ряда.

Для того, чтобы выявить устойчивую сезонную волну, на которой не отражались бы случайные условия одного года, индексы сезонности вычисляют по данным за несколько лет (не менее трех), распределенным по месяцам. Совокупность рассчитанных индексов отражает сезонную волну и позволяет определить, каким может быть влияние сезонных факторов на изменение уровней временного ряда.

Моделирование сезонных колебаний. Статистическое изучение сезонных колебаний имеет большое практическое значение, так как позволяет разрабатывать более квалифицированные и объективные планы по обеспечению договорных обязательств.

Простейший подход к моделированию сезонных колебаний – это расчет значений сезонной компоненты методом скользящей средней и построение аддитивной или мультипликативной модели временного ряда. Общий вид аддитивной модели следующий:

$$Y = T + S + E$$

Эта модель предполагает, что каждый уровень временного ряда может быть представлен как сумма трендовой (T), сезонной (S) и случайной (E) компонент. Общий вид мультипликативной модели выглядит так:

$$Y = T \times S \times E$$

Эта модель предполагает, что каждый уровень временного ряда может быть представлен как произведение трендовой (T), сезонной (S) и случайной (E) компонент.

Выбор одной из двух моделей осуществляется на основе анализа структуры сезонных колебаний. Если амплитуда колебаний приблизительно постоянна, строят аддитивную модель временного ряда, в которой значения сезонной компоненты предполагаются постоянными для различных циклов. Если амплитуда сезонных колебаний возрастает или уменьшается, строят мультипликативную модель временного ряда, которая ставит уровни ряда в зависимость от значений сезонной компоненты.

Построение аддитивной и мультипликативной моделей сводится к расчету значений T, S и E для каждого уровня временного ряда. Процесс построения модели включает в себя следующие шаги:

- 1) Выравнивание исходного ряда методом скользящей средней;
- 2) Расчет значений сезонной компоненты S;
- 3) Устранение сезонной компоненты из исходных уровней ряда и получение выровненных данных (T+E) в аддитивной или (T×E) в мультипликативной модели;
- 4) Аналитическое выравнивание уровней (T+ E) или (T×E) и расчет значений T с использованием полученного уравнения тренда;
- 5) Расчет полученных по модели значений (T + E) или (T × E);
- 6) Прогноз будущих значений уровней временного ряда на основе построенной модели.

Методику построения каждой из моделей рассмотрим на примерах. Обратимся к данным об объеме правонарушений на таможне за четыре года, представленным в таблице 1. (например, Ростовская область).

Таблица 1.

Год	Квартал	t	Количество возбужденных дел, $y_t$
2020	1	1	36

	2	2	37
	3	3	86
	4	4	101
2021	1	5	35
	2	6	47
	3	7	99
	4	8	102
2022	1	9	39
	2	10	35
	3	11	99
	4	12	90
2023	1	13	46
	2	14	45
	3	15	92
	4	16	95

Было показано, что данный временной ряд содержит сезонные колебания периодичностью 4, т.к. количество правонарушений в первый-второй кварталы ниже, чем в третий-четвертый. Рассчитаем компоненты аддитивной модели временного ряда.

**Шаг 1.** Проведем выравнивание исходных уровней ряда методом скользящей средней. Для этого:

1.1. Просуммируем уровни ряда последовательно за каждые четыре квартала со сдвигом на один момент времени и определим условные годовые объемы правонарушений

1.2. Разделив полученные суммы на 4, найдем скользящие средние (гр. 4 табл. 2). Полученные таким образом выровненные значения уже не содержат сезонной компоненты.

1.3. Приведем эти значения в соответствие с фактическими моментами времени, для чего найдем средние значения из двух последовательных скользящих средних – центрированные скользящие средние (гр. 5 табл. 2).

Таблица 2.

№ кварт	Количество правонарушений	Итого за 4 квартала	Средняя скользящая	Центрированная ср.ск.	Оценка сезонной комп.
<b>1</b>	<b>2</b>	<b>3</b>	<b>4</b>	<b>5</b>	<b>6</b>
1	<b>36</b>	-	-	-	-
2	<b>37</b>	<b>260</b>	65	-	-
3	<b>86</b>	<b>259</b>	64.7	64.8	21.2

4	<b>101</b>	<b>269</b>	67.2	65.9	35.1
5	<b>35</b>	<b>282</b>	70.5	68.8	– 33.8
6	<b>47</b>	<b>283</b>	70.7	70.6	– 23.6
7	<b>99</b>	<b>287</b>	71.7	71.2	27.8
8	<b>102</b>	<b>275</b>	68.7	70.2	31.8
9	<b>39</b>	<b>275</b>	68.7	68.7	– 29,7
10	<b>35</b>	<b>263</b>	65.7	67.2	– 32.2
11	<b>99</b>	<b>270</b>	67.5	66.6	32.4
12	<b>90</b>	<b>280</b>	70.0	68.7	21.3
13	<b>46</b>	<b>273</b>	68.2	69.1	– 23.1
14	<b>45</b>	<b>278</b>	69.5	68.8	– 23.8
15	<b>92</b>	-	-	-	-
16	<b>95</b>	-	-	-	-

**Шаг 2.** Найдем оценки сезонной компоненты как разность между фактическими уровнями ряда и центрированными скользящими средними (гр. 6 табл. 3). Используем эти оценки для расчета значений сезонной компоненты  $S_i$  (табл. 3). Для этого найдем средние за каждый квартал (по всем годам) оценки сезонной компоненты  $S_i$ .

Таблица 3

Показатели	год	№ квартала			
		1	2	3	4
	2020	-	-	21.2	35.1
	2021	– 33.8	– 23.6	27.8	31.8
	2022	– 29.7	– 32.2	32.4	21.3
	2023	– 23.1	– 23.8	–	–
Всего за $i$ -й квартал		– 86.6	– 79.6	81.4	88.2
Средняя оценка сезонной компоненты для $i$ -го квартала, $S_i$		– 28.8	– 26.5	27.1	29.4
Скорректированная сезонная компонента, $S_i$		–29.15	– 26.85	26.75	29.05

В моделях с сезонной компонентой обычно предполагается, что сезонные воздействия за период взаимопогашаются. В аддитивной модели это выражается в том, что сумма значений сезонной компоненты по всем кварталам должна быть равна нулю.

Для данной модели имеем:  $- 28.8 - 26.3 + 27.1 + 29.4 = 1.4$

Корректирующий коэффициент:  $k = 1,4 / 4 = 0,35$ .

Рассчитываем скорректированные значения сезонной компоненты

$(S_i = S_{i-k})$  и заносим полученные данные в таблицу 4.

Проверим равенство нулю суммы значений сезонной компоненты:

$$-29.15 - 26.85 + 26.75 + 29.05 = 0$$

**Шаг 3.** Исключим влияние сезонной компоненты, вычитая ее значение из каждого уровня исходного временного ряда. Получим величины  $T + E = Y - S$  (гр. 4 табл. 4) Эти значения рассчитываются за каждый момент времени и содержат только тенденцию и случайную компоненту.

**Шаг 4.** Определим компоненту  $T$  данной модели. Для этого проведем аналитическое выравнивание ряда  $(T+E)$  с помощью линейного тренда. Результаты аналитического выравнивания следующие:  $T = 67,2 + 0,092t$

Подставляя в это уравнение значения  $t = 1, 2, \dots, 16$ , найдем уровни  $T$  для каждого момента времени (гр. 5 табл. 4).

Таблица 4

t	$Y_t$	$S_i$	$Y_t - S_i$	T	T + S	$E = Y_t - (T+S)$	$E^2$
1	2	3	4	5	6	7	8
1	36	-29.15	65.15	67.29	38.14	- 2.14	4.57
2	37	- 26.85	63.85	67,38	40.53	- 3.53	12.46
3	86	26.75	59.25	67.47	94.22	- 8.22	67.56
4	101	29.05	71.95	67.56	96.61	4.39	19.27
5	35	-29.15	64.15	67.66	38.51	- 3.51	12.32
6	47	- 26.85	73.85	67.75	40,90	6.1	37.21
7	99	26.75	72.25	67.84	94.59	4.41	19.44
8	102	29.05	72.95	67.93	96.98	5.02	25.2
9	39	-29.15	68.15	68.02	38.87	0.13	0,02
10	35	- 26.85	61.85	68.12	41.27	- 6.27	39.31
11	99	26.75	72.25	68.21	94.96	4.04	16.32
12	90	29.05	60.95	68.30	97.35	- 7.35	54.02
13	46	-29.15	75.15	68.39	39.24	6.76	45.69
14	45	- 26.85	71.85	68.48	41.63	3.37	11.35
15	92	26.75	65.25	68.58	95.33	- 3.33	11.08
16	95	29.05	65.95	68.67	97.72	- 2.72	7.39

**Шаг 5.** Найдем значения уровней ряда, полученные по аддитивной модели. Для этого прибавим к уровням  $T$  значения сезонной компоненты для соответствующих кварталов (гр. 6 табл. 4). На одном графике отложим фактические значения уровней временного ряда и теоретические, полученные по аддитивной модели.



Для оценки качества построенной модели применим сумму квадратов полученных абсолютных ошибок.

$$R^2 = 1 - \frac{\sum E^2}{\sum (y_t - y)^2}$$

Следовательно, можно сказать, что аддитивная модель объясняет 97 % общей вариации уровней временного ряда количества правонарушений по кварталам за 4 года.

**Шаг 6.** Прогнозирование по аддитивной модели. Предположим, что по нашему примеру необходимо дать прогноз об общем объеме правонарушений на I и II кварталы 2026 года. Прогнозное значение  $F_t$  уровня временного ряда в аддитивной модели есть сумма трендовой и сезонной компонент. Для определения трендовой компоненты воспользуемся уравнением тренда

$$T = 67,2 + 0,092t$$

## II. ПРАКТИЧЕСКАЯ ЧАСТЬ

### 2.1 ЗАДАЧИ

**Задача 1.** Имеются следующие данные о работе за отчетный месяц 22 продовольственных магазинов одного из районов города: На основе приведенных данных определите наличие корреляционной связи, используя метод параллельного сопоставления рядов, аналитических группировок, постройте корреляционное поле, рассчитайте коэффициенты корреляции (коэффициент Фехнера и линейный коэффициент Пирсона) постройте линейное уравнение парной регрессии, рассчитайте параметры уравнения. Объясните полученные статистические характеристики.

#### Основные экономические показатели торговых организаций

№ магазина	Среднесписочная численность, чел.	Объем розничного товарооборота, млн руб.	Прибыль, млн руб.
1	8	1,5	1,8
2	19	2,9	3,8
3	7	1,3	1,7
4	43	7,1	9,2
5	29	3,3	4,6
6	38	3,5	4,2
7	25	2,9	4,3
8	9	2,0	2,8
9	37	5,8	8,1
10	10	1,9	2,7
11	30	3,1	4,3
12	21	2,3	3,2
13	12	1,7	2,4
14	11	1,6	2,4
15	34	4,0	5,2
16	23	2,6	3,9
17	27	3,0	3,8
18	40	5,8	8,0
19	36	4,9	6,7
20	17	2,1	2,8
21	6	1,0	1,3
22	42	5,3	7,2

**Задача 2.** В результате обследования работников предприятия получены следующие данные (чел.):

Образование	Удовлетворены работой	Не удовлетворены работой	Итого
Высшее и среднее	300	50	350
Незаконченное среднее	200	250	450
Итого	500	300	800

Требуется оценить тесноту взаимосвязи между уровнем образования и удовлетворенностью своей работой с помощью коэффициентов контингенции и ассоциации.

**Задача 3.** Проанализируйте связь между полом работника и характером труда в сезонных отраслях:

Численность занятых в отраслях			
	Сезонных	Не сезонных	Всего
Мужчины	187	265	452
Женщины	307	272	579
Всего	494	537	1031

**Задача 4.** Распределение предприятий пищевой промышленности Ростовской области по источникам формирования оборотных средств характеризуется следующими данными:

Источники средств	Малые предприятия	Средние предприятия	Итого
Банковский кредит	31	32	63
Собственные средства	38	15	53
Итого	69	47	116

Вычислите коэффициенты ассоциации и контингенции. Какие выводы можно сделать на основании этих коэффициентов?

**Задача 5.** Туристическая компания предлагает места в гостиницах приморского курорта. Менеджера компании интересует, насколько возрастает привлекательность гостиницы в зависимости от её удаленности от пляжа. С этой целью по 14 гостиницам города была выяснена среднегодовая наполняемость номеров и расстояние от пляжа:

Расстояние, км	0,1	0,1	0,2	0,3	0,4	0,4	0,5	0,6	0,7	0,7	0,8	0,8	0,9	0,9
Наполняемость, %	92	95	96	90	89	86	90	83	85	80	78	76	72	75

**Задание:**

1. Постройте поле корреляции результативного и факторного признаков.
2. Рассчитайте линейный коэффициент корреляции и поясните его смысл.
3. Определите коэффициент детерминации и дайте его интерпретацию.
4. Определите параметры уравнения парной линейной регрессии и дайте интерпретацию коэффициента регрессии  $a_1$ .
5. Рассчитайте ошибку аппроксимации и сделайте выводы.

## 2.2 ИНДИВИДУАЛЬНЫЕ ЗАДАНИЯ

### ИНДИВИДУАЛЬНОЕ ЗАДАНИЕ 1

На основе приведенных данных постройте эконометрические модели корреляции и регрессии

Регионы	Внесение минеральных удобрений на один га посевных площадей, т	Использование органических удобрений на один га посевных площадей, т	Урожайность овощей, центнеров с одного га
1.Брянская область	1,3	0,6	131
2.Владимирская область	1,9	0,7	170
3.Ивановская область	1,2	0,2	103
4.Калужская область	1,6	0,8	175
5.Костромская область	2,5	1,8	211
6.Московская область	0,9	1,3	176
7.Орловская область	1,1	0,7	175
8.Рязанская область	1,4	0,9	166
9.Смоленская область	2,0	1,3	187
10.Тверская область	2,6	1,5	236
11.Тульская область	1,7	0,6	167
12.Ярославская область	1,3	0,3	156
13.Ростовская область	1,8	1,1	209
14.Волгоградская область	0,6	0,5	100

- 1.Сформулируйте гипотезу исследования, определите факторные и результативный признаки.
- 2.Постройте два корреляционных поля на основе взаимосвязи признаков.
- 3.Рассчитайте линейные коэффициенты парной корреляции Пирсона между каждым их факторных признаков и результативным признаком. Определите коэффициент детерминации и дайте его интерпретацию. Определите, какой признак имеет наиболее тесную связь с результативным.
- 4.Постройте линейное уравнение регрессии на основе взаимосвязи доминирующего признака из числа факторных и результативного признака, рассчитайте параметры уравнения. Нанесите на корреляционное поле полученные результаты моделирования.
- 5.Рассчитайте ошибку аппроксимации. Сделайте прогнозный расчет на основе выявленной взаимосвязи. Объясните полученные статистические характеристики.
- 6.Рассчитайте коэффициент множественной корреляции. Сделайте выводы.

**ИНДИВИДУАЛЬНОЕ ЗАДАНИЕ 2**

На основе приведенных данных постройте эконометрические модели корреляции и регрессии показателей деятельности малых предприятий города

№ п/п.	Объем продукции, млн руб.	Среднегодовая стоимость основных фондов, млн руб.	Среднесписочная численность рабочих, чел.
1	9	10	11
2	17	32	15
3	19	38	16
4	12	13	14
5	22	20	18
6	39	59	42
7	34	42	39
8	12	19	11
9	38	44	37
10	21	26	25
11	28	36	26
12	16	29	15
13	10	14	12
14	25	18	20

1. Сформулируйте гипотезу исследования, определите факторные и результативный признаки.
2. Постройте два корреляционных поля на основе взаимосвязи признаков.
3. Рассчитайте линейные коэффициенты парной корреляции Пирсона между каждым их факторных признаков и результативным признаком. Определите коэффициент детерминации и дайте его интерпретацию. Определите, какой признак имеет наиболее тесную связь с результативным.
4. Постройте линейное уравнение регрессии на основе взаимосвязи доминирующего признака из числа факторных и результативного признака, рассчитайте параметры уравнения. Нанесите на корреляционные поле полученные результаты моделирования.
5. Рассчитайте ошибку аппроксимации. Сделайте прогнозный расчет на основе выявленной взаимосвязи. Объясните полученные статистические характеристики.
6. Рассчитайте коэффициент множественной корреляции. Сделайте выводы.

**ИНДИВИДУАЛЬНОЕ ЗАДАНИЕ 3**

На основе приведенных данных постройте эконометрические модели корреляции и регрессии по основным показателям деятельности крупных банков одной из областей России (данные условные, млрд руб.):

№ п/п	Сумма активов	Собственный капитал	Балансовая прибыль
1	5,4	2,5	15
2	3,6	1,6	9
3	5,2	2,8	20
4	6,0	4,4	28
5	9,0	7,1	36
6	9,5	7,6	40
7	6,7	7,5	39
8	9,6	5,1	27
9	7,1	3,5	9
10	9,8	8,0	38
11	7,3	6,5	34
12	7,0	6,3	35
13	7,2	4,9	27
14	6,8	5,0	26

1. Сформулируйте гипотезу исследования, определите факторные и результативный признаки.
2. Постройте два корреляционных поля на основе взаимосвязи признаков.
3. Рассчитайте линейные коэффициенты парной корреляции Пирсона между каждым их факторных признаков и результативным признаком. Определите коэффициент детерминации и дайте его интерпретацию. Определите, какой признак имеет наиболее тесную связь с результативным.
4. Постройте линейное уравнение регрессии на основе взаимосвязи доминирующего признака из числа факторных и результативного признака, рассчитайте параметры уравнения. Нанесите на корреляционные поле полученные результаты моделирования.
5. Рассчитайте ошибку аппроксимации. Сделайте прогнозный расчет на основе выявленной взаимосвязи. Объясните полученные статистические характеристики.
6. Рассчитайте коэффициент множественной корреляции. Сделайте выводы.

**ИНДИВИДУАЛЬНОЕ ЗАДАНИЕ 4**

На основе приведенных данных постройте эконометрические модели парной корреляции и регрессии по основным показателям деятельности бригады рабочих 1. Сформулируйте гипотезу исследования, определите факторные и результативный признаки.

№ п/п	Стаж работы, лет	Часовая выработка рабочего, тыс. руб.	Среднемесячная заработная плата рабочего, тыс. руб.
1	1,0	0,5	22,0
2	1,5	1,2	31,0
3	3,0	1,5	36,0
4	6,5	1,9	38,0
5	9,0	2,0	44,0
6	4,4	1,6	37,0
7	6,0	1,8	39,0
8	2,5	1,4	35,0
9	2,7	1,4	38,0
10	6,0	1,6	36,0
11	3,2	1,8	40,0
12	4,0	1,7	32,0
13	10,0	2,5	45,0
14	7,0	1,8	40,0

2. Постройте два корреляционных поля на основе взаимосвязи признаков.

3. Рассчитайте линейные коэффициенты парной корреляции Пирсона между каждым их факторных признаков и результативным признаком. Определите коэффициент детерминации и дайте его интерпретацию. Определите, какой признак имеет наиболее тесную связь с результативным.

4. Постройте линейное уравнение регрессии на основе взаимосвязи доминирующего признака из числа факторных и результативного признака, рассчитайте параметры уравнения. Нанесите на корреляционные поле полученные результаты моделирования.

5. Рассчитайте ошибку аппроксимации. Сделайте прогнозный расчет на основе выявленной взаимосвязи. Объясните полученные статистические характеристики.

6. Рассчитайте коэффициент множественной корреляции. Сделайте выводы.

**ИНДИВИДУАЛЬНОЕ ЗАДАНИЕ 5**

На основе приведенных данных постройте эконометрические модели по анализу работы туристического комплекса

Месяцы	1	2	3	4	5	6	7	8	9	10	11	12
Средняя наполняемость, номеров, %	92	88	79	76	79	86	90	95	100	90	78	76
Цена за номер, тыс. руб.	1,5	1,2	1,1	1,0	0,9	1,2	1,5	2,0	2,0	2,0	1,5	1,2

1. Постройте графики рядов динамики.
2. Рассчитайте коэффициент автокорреляции уровней ряда первого порядка. Сделайте выводы.
3. Проведите аналитическое выравнивание рядов на основе МНК. Рассчитайте ошибку аппроксимации. Нанесите на график полученные значения параметров каждой модели.
4. Определите величину отклонения каждого фактического уровня ряда динамики от соответствующего ему выравненного значения;
5. Определите линейный коэффициент корреляции полученных отклонений. Сделайте выводы.
6. Сделайте прогнозный расчет по выравненным рядам на май следующего года без учета сезонной компоненты.

**ИНДИВИДУАЛЬНОЕ ЗАДАНИЕ 6**

На основе приведенных данных постройте эконометрические модели по анализу работы туристического комплекса

Расстояние от лыжной трассы, км	0,2	0,3	0,4	0,5	0,6	0,7	0,8	0,8	0,9	0,9	1,0	1,3
Средняя наполняемость, номеров, %	100	98	95	90	92	90	86	89	80	87	85	80
Цена за номер, тыс. руб.	3,5	2,8	2,5	2,3	2,9	1,8	1,5	2,0	1,80	2,0	1,5	1,0

1. Сформулируйте гипотезу исследования, определите факторные и результативный признаки.
2. Постройте два корреляционных поля на основе взаимосвязи признаков.
3. Рассчитайте линейные коэффициенты парной корреляции Пирсона между каждым их факторных признаков и результативным признаком. Определите коэффициент детерминации и дайте его интерпретацию. Определите, какой признак имеет наиболее тесную связь с результативным.



4. Постройте линейное уравнение регрессии на основе взаимосвязи доминирующего признака из числа факторных и результативного признака, рассчитайте параметры уравнения. Нанесите на корреляционное поле полученные результаты моделирования.
5. Рассчитайте ошибку аппроксимации. Сделайте прогнозный расчет на основе выявленной взаимосвязи. Объясните полученные статистические характеристики.
6. Рассчитайте коэффициент множественной корреляции. Сделайте выводы.

### ИНДИВИДУАЛЬНОЕ ЗАДАНИЕ 7

По данным станций технического обслуживания легковых автомобилей города требуется провести эконометрическое исследование и выявить наличие сезонной компоненты в рядах динамики

Месяц	Число поступивших заявок, тыс.			
	2021 г.	2022 г.	2023 г.	2024 г.
Январь	10,3	13,6	14,0	15,0
Февраль	11,1	14,3	14,7	15,7
Март	11,5	14,4	15,1	15,8
Апрель	12,0	14,6	15,6	16,6
Май	12,6	15,6	16,0	16,7
Июнь	16,0	17,1	17,3	17,4
Июль	15,9	16,9	18,2	18,1
Август	16,2	17,0	18,3	18,4
Сентябрь	16,4	16,5	17,8	17,9
Октябрь	15,2	16,0	17,5	18,5
Ноябрь	15,0	14,9	17,0	18,0
Декабрь	12,8	13,8	16,5	16,9

1. Постройте график рядов динамики, используя укрупнение интервалов до квартальных данных.
2. Произведите выравнивание исходных квартальных данных методом скользящей средней.
3. Найдите оценку сезонной компоненты и исключите её влияние в модели
4. Определите трендовую компоненту в данной модели.
5. Найдите значения уровней ряда, полученные по аддитивной модели. Рассчитайте ошибку аппроксимации.
6. Сделайте прогноз на два квартала вперед и обоснуйте свои выводы.

**ИНДИВИДУАЛЬНОЕ ЗАДАНИЕ 8**

На основе приведенных данных постройте эконометрические модели по анализу взаимосвязи экономических показателей торговой сети города

**Основные экономические показатели торговых организаций**

№ магазина	Объем розничного товарооборота, млн руб.	Издержки обращения, млн руб.	Прибыль, млн руб.
1	1,5	0,9	1,8
2	2,9	1,5	3,8
3	1,3	0,8	1,7
4	7,1	4,4	9,2
5	3,3	2,0	4,6
6	3,5	2,1	4,2
7	2,9	1,7	4,3
8	2,0	1,3	2,8
9	5,8	3,6	8,1
10	1,9	1,1	2,7
11	3,1	1,9	4,3
12	2,3	1,4	3,2
13	1,7	1,0	2,4
14	1,6	0,9	2,4

1. Сформулируйте гипотезу исследования, определите факторные и результативный признаки.
2. Постройте два корреляционных поля на основе взаимосвязи признаков.
3. Рассчитайте линейные коэффициенты парной корреляции Пирсона между каждым их факторных признаков и результативным признаком. Определите коэффициент детерминации и дайте его интерпретацию. Определите, какой признак имеет наиболее тесную связь с результативным.
4. Постройте линейное уравнение регрессии на основе взаимосвязи доминирующего признака из числа факторных и результативного признака, рассчитайте параметры уравнения. Нанесите на корреляционное поле полученные результаты моделирования.
5. Рассчитайте ошибку аппроксимации. Сделайте прогнозный расчет на основе выявленной взаимосвязи. Объясните полученные статистические характеристики.
6. Рассчитайте коэффициент множественной корреляции. Сделайте выводы.

**Контрольная работа для текущего контроля**

1. Дайте определение эконометрики и укажите, для чего могут применяться эконометрические модели.
2. Для построения эконометрических моделей используются экзогенные и эндогенные переменные. В чём различие этих переменных. Поясните на примере.
3. На основе данных об объемах продажи продукции птицефабрики определить основную тенденцию развития, используя метод скользящей средней. Данные представлены в тысячах руб. Полученную тенденцию изобразите графически.

Год	Месяц											
	1	2	3	4	5	6	7	8	9	10	11	12
2023	35	29	28	37	32	38	47	47	49	43	38	37
2024	33	34	37	38	40	47	49	52	51	54	55	56

### 2.3 ТЕСТОВЫЕ ЗАДАНИЯ ДЛЯ САМОПРОВЕРКИ

**1. Дайте определение эконометрики и укажите, для чего могут применяться эконометрические модели.**

**2. Для построения эконометрических моделей используются экзогенные и эндогенные переменные. В чём различие этих переменных. Поясните на примере.**

**3. Типы данных, используемых в эконометрических моделях:**

а) географические данные; б) панельные данные; в) гипотетические данные.

**4. В зависимости от причинно-следственных связей статистические признаки разделяют на следующие виды ...**

а) результативные; б) факторные; в) зависимые; г) независимые.

**5. Если определённому значению факторного признака соответствует строго определённое значение результативного признака, то это:**

а) альтернативная связь; б) функциональная связь; в) корреляционная связь.

**6. Виды корреляционных связей между признаками по направлению различают на:**

а) линейные; б) прямые; в) нелинейные; г) обратные.

**7. В теории статистики по аналитическому выражению выделяют следующие виды связей между признаками**

а) линейные; б) нелинейные; в) прямые; г) обратные.

**8. Наиболее наглядным видом выбора уравнения парной регрессии является:**

а) аналитический; б) графический; в) экспериментальный (табличный).

**9. Рассчитывать параметры парной линейной регрессии можно, если у нас есть:**

а) не менее 5 наблюдений; б) не менее 7 наблюдений; в) не менее 10 наблюдений.

**10. Прямолинейная связь между факторами исследуется с помощью уравнения регрессии ...**

**А.** Ошибка! Объект не может быть создан из кодов полей редактирования.;

**Б.** Ошибка! Объект не может быть создан из кодов полей редактирования.;

**В.** Ошибка! Объект не может быть создан из кодов полей редактирования.;

**Г.** Ошибка! Объект не может быть создан из кодов полей редактирования.

**11. Линейный коэффициент корреляции может принимать значения...**

- а) любые меньше нуля; б) от  $-1$  до  $1$ ; в) от  $-1$  до  $0$ ; г) любые положительные; д) от  $0$  до  $1$ .

**12. Параметр  $a_1$  ( $a_1 = -1,04$ ) линейного уравнения регрессии  $y_x = 36,5 - 1,04x$  показывает, что:**

- а) с увеличением признака «х» на 1 признак «у» уменьшается на 36,5; б) связь между признаками «х» и «у» прямая; в) связь между признаками «х» и «у» обратная; г) с увеличением признака «х» на 1 признак «у» уменьшается на 1,04

**13. Укажите соответствие между значениями коэффициента корреляции и смысловой интерпретацией связи**

Значение а)  $-0,5$ ; б)  $1,9$ ; в)  $-0,1$ ; г)  $0,95$

**Смысловая интерпретация:**

1. Связь обратная, практически отсутствует; 2. Связь прямая сильная 3. Связь обратная умеренная. 4 Показатель не имеет смысла

**14. Эконометрическая модель парной корреляции должна отражать:**

- а) существенные свойства объекта; б) математические уравнения; в) причинно-следственные связи между признаками; г) реальную действительность.

**15. Укажите соответствие между классификационными признаками и видами корреляционной связи**

Классификационные признаки	Виды корреляционной связи
1. теснота связи	а) парная, множественная
2. направление связи	б) линейная, нелинейная
3. аналитическое выражение связи	в) прямая, обратная
4. число взаимосвязанных статистических признаков	г) практически отсутствует, слабая, умеренная, сильная

**16. Суть метода наименьших квадратов состоит в:**

- а) минимизации суммы остаточных величин; б) минимизации дисперсии резуль- тативного признака; в) минимизации суммы квадратов остаточных величин.

**17. Коэффициент линейного парного уравнения регрессии:**

- а) показывает среднее изменение результата с изменением фактора на одну единицу; б) оценивает статистическую значимость уравнения регрессии; в) показывает, как тесно взаимосвязаны анализируемые признаки.

**18. Множественный коэффициент корреляции может принимать значения...**

- а) от  $0$  до  $1$ ; б) любые меньше  $0$ ; в) любые положительные; г) от  $-1$  до  $0$ ; д) от  $-1$  до  $1$ .

- 19. Если число отобранных факторов в регрессионной модели равно 3, то какой минимальный объем совокупности должен быть взят для анализа?**
- 20. Произведите отбор факторов для построения множественной регрессии на условном примере с учетом требований отбора. Например, какие факторы влияют на товарооборот в магазине.**
- 21. Криволинейная связь между факторами исследуется с помощью ...**  
**А. Ошибка! Объект не может быть создан из кодов полей редактирования.; Б. Ошибка! Объект не может быть создан из кодов полей редактирования.; В. Ошибка! Объект не может быть создан из кодов полей редактирования.; Г. Ошибка! Объект не может быть создан из кодов полей редактирования.**
- 22. Дайте определение временного ряда (ряда динамики) и укажите, какие группы факторов влияют на его формирование.**
- 23. Раскройте понятие автокорреляции уровней временного ряда**
- 24. Раскройте понятие – экстраполяция и интерполяция значений временных рядов**
- 25. Моделирование сезонных колебаний – понятие и сущность метода скользящей средней**
- 26. Специфика статистической оценки взаимосвязи двух временных рядов.**

### **3. ПЕРЕЧЕНЬ ИСПОЛЬЗОВАННЫХ ИНФОРМАЦИОННЫХ РЕСУРСОВ**

1. Демидова О.А., Малахов Д.И. Эконометрика. Учебник. – М.: Издательство Юрайт, 2018. – 334 с.
2. Елисеева И.И., Статистика: учебник для прикладного бакалавриата. – 3-е изд., перераб. и доп. – М.: Издательство Юрайт, 2017. – 361 с.
3. Ниворожкина Л.И. Эконометрика: теория и практика / Л.И. Ниворожкина, С.В. Арженовский, Е.П. Кокина. – Москва: ИЦ РИОР, 2018. – 207 с.