

Т.М. Сизова, Л.Г. Мишура
СТАТИСТИКА: ПРАКТИКУМ



МИНИСТЕРСТВО ОБРАЗОВАНИЯ И НАУКИ РОССИЙСКОЙ ФЕДЕРАЦИИ

УНИВЕРСИТЕТ ИТМО

Т.М. Сизова, Л.Г. Мишура
СТАТИСТИКА: ПРАКТИКУМ
Учебное пособие



Санкт-Петербург

2016

Т.М. Сизова, Л.Г. Мишура Статистика: практикум. - СПб: Университет ИТМО, 2016. - 61 с.

Учебное пособие предназначено для использования в учебном процессе студентами Факультета технологического менеджмента и инноваций Университета ИТМО, обучающихся по направлениям 38.03.01 «Экономика» и 38.03.02 «Менеджмент» и изучающих дисциплину «Статистика». Пособие содержит краткий обзор основных понятий теории распределения, динамических рядов, выборочного исследования, корреляционного анализа и индексов, методики выполнения работ по указанным темам, а также необходимые экономико-статистические таблицы.

Рекомендовано к печати Советом Гуманитарного факультета 20.01.2015, протокол № 2.



Университет ИТМО - ведущий вуз России в области информационных и фотонных технологий, один из немногих российских вузов, получивших в 2009 году статус национального исследовательского университета. С 2013 года Университет ИТМО - участник программы повышения конкурентоспособности российских университетов среди ведущих мировых научно-образовательных центров, известной как проект «5 в 100». Цель Университета ИТМО - становление исследовательского университета мирового уровня, предпринимательского по типу, ориентированного на интернационализацию всех направлений деятельности.

© Университет ИТМО, 2016

©Т.М. Сизова, Л.Г. Мишура, 2016

Содержание

Введение.....	4
1. Практическая работа №1 «Построение статистических таблиц».....	5
2. Практическая работа №2 «Расчет статистических показателей».....	9
3. Практическая работа №3 «Построение и исследование вариационного ряда».....	17
4. Практическая работа №4 «Исследование корреляционной связи».....	26
5. Практическая работа №5 «Исследование ряда динамики».....	33
6. Практическая работа №6 «Выборочное наблюдение».....	42
7. Практическая работа №7 «Индексы».....	47
Список литературы.....	55
Приложения.....	56

Введение

Практикум предназначен для бакалавров Университета ИТМО, обучающихся по направлениям подготовки 38.03.01 «экономика», 38.03.02 «менеджмент» и изучающих дисциплину «статистика».

Содержание и структура практикума соответствует требованиям Государственных образовательных стандартов высшего профессионального образования и действующих в университете учебных планов подготовки бакалавров.

Практикум дополняет учебное пособие «Статистика», написанное теми же авторами.

В практикум включены практические работы по ключевым темам первого раздела дисциплины - «Общей теории статистики». По каждой работе даются основные определения и необходимые для расчетов методики, излагается порядок выполнения, предложены вопросы и тестовые задания для контроля полученных знаний. Форма проведения занятий в виде выполнения практических работ базируется на изложении основных положений изучаемой дисциплины в виде лекций, опросах по итогам освоения каждой темы, выполнении практических работ на основе индивидуальных исходных данных и консультировании.

Пособие обеспечивает методическую поддержку практических занятий по курсу «Статистика» и ориентировано на создание необходимых условий для успешной самостоятельной работы студентов.

Построение статистических таблиц

Основные определения

Статистическая таблица представляет собой компактную и наглядную форму изложения статистических данных об исследуемых явлениях. Основу статистической таблицы составляет графленная сетка, вертикальные столбцы которой называются графами, а горизонтальные - строками. Статистическая таблица отличается от других табличных форм наличием итоговой строки и (или) итогового столбца.

Основными элементами статистической таблицы являются: заголовок, подлежащее и сказуемое. Заголовок таблицы отражает содержание таблицы, место и время, к которому относятся ее данные, единицы измерения, если они являются общими для приведенных данных. *Подлежащим* таблицы являются перечень единиц совокупности или группы, т. е. объект изучения, *сказуемым* - система показателей, характеризующих объект изучения, т. е. подлежащее. Как правило, *подлежащее таблицы слева в виде названий строк, а сказуемое - сверху в виде названий граф.* Подлежащее формирует боковые заголовки, а сказуемое - верхние. Статистическая информация располагается в таблице в *клетках таблицы*, на пересечении соответствующих строки и столбца.

Основу любой статистической таблицы составляет ее **макет** - форма с названием строк и граф, но незаполненная статистическими данными.

По содержанию подлежащего все статистические таблицы можно разделить на следующие группы:

- **Простые** таблицы, в подлежащем которых отсутствуют группировки по одному признаку, то есть дается перечень единиц, составляющих объект изучения. Они содержат обобщающие показатели, относящиеся к перечню единиц совокупности (перечневые таблицы), к перечню хронологических дат (хронологические таблицы) или к перечню территорий (территориальные таблицы). В простых таблицах обычно отражается справочный материал.
- **Групповые** таблицы, в подлежащем которых изучаемый объект разделен на группы по определенному признаку. При этом каждая группа может быть охарактеризована рядом показателей.
- **Комбинационные** таблицы, в подлежащем которых дана группировка единиц совокупности по двум и более признакам, взятым в комбинации. В комбинационных таблицах статистические единицы совокупности по наиболее важному для изучения признаку распределяются на группы, затем по следующему признаку внутри каждой из полученных групп - на подгруппы и т. д.

Сказуемое таблицы может быть *простым или сложным*.

Простое сказуемое предусматривает параллельное расположение показателей, а **сложное** - комбинированное.

При построении статистических таблиц необходимо соблюдать определенные **правила по их оформлению**:

1. Таблица должна быть компактной, легко обозримой. Ее не следует загружать излишними подробностями, затрудняющими анализ.

2. Заголовок таблицы должен ясно и кратко выражать ее содержание. Заголовки строк подлежащего и граф сказуемого также должны быть сформулированы точно и кратко.

3. В таблице желательно давать нумерацию граф. Это облегчает пользование таблицей, показывает способ расчета чисел в графах. Графы, содержащие подлежащее, обозначаются заглавными буквами алфавита; графы, содержащие сказуемое, нумеруются арабскими числами. Не допускается в заголовках подлежащего и сказуемого сокращение слов.

4. Если единицы измерения различны, то они указываются в названиях строк и граф.

5. Приводимые в подлежащем и сказуемом признаки должны располагаться в логическом порядке с учетом необходимости их совместного рассмотрения. Информация размещается от частного к общему, т. е. сначала показываются слагаемые, а в конце подводят итоги.

6. Если в таблице приводятся не все данные, а только наиболее значимые из них то сначала показывают итог, а затем выделяют наиболее важные части с помощью оборотов «в том числе», «из них».

7. Следует различать «Итого» и «Всего». «Итого» является итогом для определенной части совокупности, а «Всего» - итог для всей совокупности.

8. При оформлении таблицы применяются следующие обозначения:

- прочерк (-) - когда явление отсутствует;
- символ «х» - если явление не имеет осмысленного содержания;
- многоточие (...) - если отсутствуют сведения (или делается запись «нет сведений»);
- если сведения имеются, но числовое их значение меньше принятой в таблице точности, оно выражается дробным числом 0,0.

9. Округление чисел, приводимых в таблице, должно проводиться с одинаковой степенью точности.

10. Если одна величина превосходит другую многократно, то полученные относительные показатели лучше выражать не в процентах, а в количестве раз.

Соблюдение приведенных правил построения и оформления статистических таблиц делает их основным средством представления, обработки и обобщения статистической информации.

Анализ статистических таблиц

Анализ статистических таблиц включает в себя:

- *Структурный анализ;*
- *Содержательный анализ.*

Структурный анализ заключается в анализе строения таблицы и статистической характеристике представленных в ней данных:

- изучаемой совокупности и ее единиц;
- исследуемых признаков и их сочетаний, формирующих подлежащее и сказуемое таблицы;
- соотношение признаков, формирующих подлежащее и сказуемое;

Содержательный анализ заключается в изучении статистической информации, содержащейся в таблице:

- анализ содержания подлежащего (отдельных его групп по соответствующим признакам сказуемого)
- выявление существующих соотношений и пропорций между частями изучаемого явления по одному или нескольким статистическим признакам;
- сравнительный анализ по отдельным группам и по изучаемому явлению в целом.
- установление и формулировка закономерности.

Кроме этого, анализ таблицы предполагает проведение *логической и счетной проверки* данных таблицы. Логическая проверка заключается в выявлении соответствия числовых значений признака их содержанию. Счетная проверка сводится к проведению выборочного расчета значений признака по выделенным однородным группам и итоговых значений строк или граф.

Содержание работы

Цель работы:

Освоить *навыки построения и анализа статистических таблиц*, как одного из важных средств представления, обработки и обобщения статистической информации представления статистической информации.

Порядок выполнения:

- По полученным от преподавателя данным построить три вида статистических таблиц - простую, групповую и комбинационную ;
- Провести структурный анализ каждой из построенных таблиц
- Выполнить содержательный анализ данных, приведенных в построенных таблицах.
- Сделать выводы

Тесты:

1. Сказуемое статистической таблицы составляют:

- а) объект изучения;
- б) числовая характеристика изучаемой статистической совокупности;
- в) система показателей, характеризующих объект изучения.

2. Подлежащим статистической таблицы называется :

- а) система показателей, характеризующих объект изучения;
- б) статистические совокупности, характеризующиеся различными показателями;

в) сведения, расположенные в боковых заголовках таблицы.

3. Статистическая таблица представляет собой:

а) совокупность числовых данных об отдельных единицах изучаемой совокупности;

б) компактную и наглядную форму изложения статистических данных об исследуемых явлениях;

в) сведения, расположенные по строкам и графам.

4. Основными элементами статистической таблицы являются:

а) строки и графы;

б) заголовки, подлежащее и сказуемое;

в) числовые данные, содержащиеся в таблице.

Контрольные вопросы:

1. Дать определение статистической таблицы.
2. Чем отличается подлежащее таблицы от сказуемого?
3. В чем состоит различие между таблицами по характеру подлежащего?
4. Перечислить правила построения статистических таблиц.
5. Чем различаются таблицы с простой и сложной разработкой сказуемого?
6. Что такое комбинационная таблица и в каких случаях она применяется?
7. Что из себя представляет макет статистической таблицы?
8. В чем заключается анализ статистической таблицы?
9. Какова цель структурного анализа статистической таблицы?
10. Для чего проводится содержательный анализ статистических таблиц?

Расчет статистических показателей

Основные определения

Статистическим показателем называется обобщающая количественно-качественная характеристика статистического признака (определенного свойства статистической совокупности).

Статистические показатели могут выражаться в форме *абсолютных, относительных и средних величин*.

Характеристика абсолютных показателей

Абсолютные показатели характеризуют численность совокупности, либо объём изучаемого явления в конкретных границах пространства и времени, т. е. отражают уровень развития явления, его размер.

Абсолютный показатель можно получить одним из двух способов:

- *путём подсчёта единиц совокупности, обладающих конкретным значением признака;*
- *путём суммирования значения признака по всей статистической совокупности.*

Абсолютные показатели всегда являются **именованными числами**. В зависимости от социально-экономической сущности исследуемых явлений они выражаются в **натуральных, стоимостных и трудовых единицах измерения**.

Характеристика относительных показателей

Относительным статистическим показателем называется обобщающая характеристика, выраженная в виде числовой меры соотношения двух сопоставляемых абсолютных величин.

Такие показатели используются в различных целях: для выяснения структуры изучаемого явления, для сравнения его уровня развития с уровнем развития другого явления, для оценки происходящих в изучаемом явлении изменений и т. д.

Относительный статистический показатель получают путём деления одного абсолютного показателя на другой. Схема расчета относительного показателя (*ОП*) выглядит следующим образом:

$$ОП = \frac{\text{величина}_\text{сравнения}}{\text{база}_\text{сравнения}};$$

где *величина сравнения* и *база сравнения* - сравниваемые абсолютные показатели.

Таким образом, по способу получения относительные показатели всегда являются величинами производными по отношению к абсолютным, их можно получить только расчетным путем.

Относительные показатели выражаются в разных формах - *коэффициентов, процентов, промилле, продецимилле*. Если база сравнения принимается за единицу, то относительный показатель выражается в форме *коэффициента*. Если база сравнения принимается за 100 единиц, то относительный показатель выражается *в процентах*. Если база сравнения принимается за 1000 единиц, то относительный показатель выражается в *промилле* (десятая часть процента), если - за 10 000, то относительный показатель выражается в *продецимилле* (сотая часть процента).

По содержанию выражаемых количественных соотношений выделяют следующие виды относительных показателей:

- **Относительный показатель динамики** характеризует изменение изучаемого явления во времени и представляет собой соотношение показателей, характеризующих явление в текущем периоде и предшествующем (базисном) периоде.

$$ОПД = \frac{\text{показатель_текущего_периода}}{\text{показатель_предшествующего_ (базисного) _периода}}.$$

- **Относительный показатель плана (прогноза) и выполнения плана**
Относительный показатель плана (ОПП) и относительный показатель выполнения плана (ОПВП) используют все субъекты финансово-хозяйственной деятельности, осуществляющие текущее и стратегическое планирование. Они рассчитываются следующим образом:

$$ОПП = \frac{\text{Плановый_показатель_текущего_периода}}{\text{Фактический_показатель_предшествующего_периода}};$$

$$ОПВП = \frac{\text{Фактический_показатель_текущего_периода}}{\text{Плановый_показатель_текущего_периода}}.$$

Относительный показатель выполнения плана характеризует напряженность планового задания, а относительный показатель выполнения плана - степень его выполнения.

- **Относительные показатели структуры (ОПС)**

Относительные показатели структуры (ОПС) характеризуют доли (удельные веса) составных частей совокупности в общем ее объеме. Они показывают структуру совокупности, ее строение. Расчет относительных показателей структуры заключается в исчислении удельных весов отдельных частей во всей совокупности:

$$ОПС = \frac{\text{показатель_части_совокупности}}{\text{показатель_всей_совокупности}}.$$

ОПС обычно выражаются в форме коэффициентов или процентах; при этом сумма коэффициентов должна составлять 1, а сумма процентов - 100, так как удельные веса приведены к общему основанию.

- **Относительные показатели координации (ОПК)** характеризуют отношение частей данных статистической совокупности к одной из них, взятой за базу сравнения, и показывают, во сколько раз одна часть совокупности больше другой, или сколько единиц одной части совокупности приходится на 1,10,100 и т.д. единиц другой части. За базу сравнения выбирается часть, имеющая наибольший удельный вес или являющаяся приоритетной в совокупности.

$$ОПК_i = \frac{\text{показатель}_\text{части}}{\text{показатель}_\text{всей}_\text{совокупности}}.$$

- **Относительные показатели интенсивности и уровня экономического развития (ОПИ)** характеризуют степень распространения или уровень развития изучаемых явлений или процессов в определённой среде и образуются как результат сравнения разноименных, но определенным образом связанных между собой величин. Указанные показатели рассчитываются следующим образом:

$$ОПИ = \frac{\text{Показатель}_\text{характеризующий}_\text{явление}}{\text{Показатель}_\text{характеризующий}_\text{среду}_\text{распространения}_\text{явления}}.$$

ОПИ исчисляются в расчете на 100, 1000, 1000 и т.д. единиц изучаемой совокупности и используются в тех случаях, когда невозможно по значению абсолютного показателя оценить масштаб распространения явления.

- **Относительные показатели сравнения (ОПСр)** характеризуют сравнительные размеры одноименных абсолютных показателей, относящихся к различным объектам или территориям, но за одинаковый период времени. Их получают как частные от деления одноименных абсолютных показателей, характеризующих разные объекты, относящихся к одному и тому же периоду или моменту времени.

$$ОПСр = \frac{\text{Показатель}_\text{характеризующий}_\text{объект}_\text{А}}{\text{Показатель}_\text{характеризующий}_\text{объект}_\text{Б}}.$$

Характеристика средних величин

Средним называется обобщающий показатель статистической совокупности, характеризующий наиболее типичный уровень явления. Средний показатель выражает величину признака на единицу совокупности. **Особенности средних** показателей заключаются в том, что они, во-первых, отражают общее, присущее всем единицам совокупности; во-вторых, в них

взаимопогашаются отклонения значений признака, возникающие под воздействием случайных факторов.

В социально-экономическом анализе используются два класса средних величин:

- *степенные средние;*
- *структурные средние.*

К классу степенных средних *относятся несколько видов средних, построенных по одному общему принципу:*

$$\bar{x} = \sqrt[k]{\frac{\sum_{i=1}^n x_i^k}{n}},$$

где x_i - варианты,

$n=N$ - объем статистической совокупности,

k - показатель степени.

Показатель степени k может принимать любые значения, но на практике обычно используются следующие: *при $k = 1$ получают среднюю арифметическую; $k = -1$ - среднюю гармоническую; $k = 0$ - среднюю геометрическую; $k = 2$ - среднюю квадратическую.*

Степенные средние в зависимости от формы представления исходных данных могут быть **простыми и взвешенными**.

- **Средняя арифметическая** является наиболее распространенным видом степенных средних; **используется в случаях, когда объём усредняемого признака является аддитивной величиной, т.е. образуется как сумма его значений.**

Средняя арифметическая простая используется при работе с не сгруппированными данными и рассчитывается по формуле:

$$\bar{x} = \frac{\sum_{i=1}^N x_i}{N}.$$

Среднюю арифметическую взвешенную используют при работе с предварительно сгруппированными данными:

$$\bar{x} = \frac{\sum_{i=1}^m x_i \cdot n_i}{\sum_{i=1}^m n_i} ; \text{ или } \bar{x}_i = \sum_{i=1}^m x_i \cdot q_i ;$$

где $q_i = \frac{n_i}{\sum_{i=1}^m n_i}$ - частость, т. е. удельный вес статистических единиц,

обладающих определенным значением признака в общем объеме совокупности.

- **Средняя гармоническая** используется в тех случаях, когда **статистическая информация не содержит частот по отдельным значениям признака, а представлена произведением значения признака на частоту.**

Средняя гармоническая простая рассчитывается по формуле:

$$\bar{x} = \frac{N}{\sum_{i=1}^n \frac{1}{x_i}};$$

При работе со сгруппированными данными используется **средняя гармоническая взвешенная:**

$$\bar{x} = \frac{\sum_{i=1}^m w_i}{\sum_{i=1}^m \frac{w_i}{x_i}},$$

где w_i – статистический вес; $w_i = x_i \cdot n_i$.

- **Средняя геометрическая** применяется в тех случаях, когда **общий объем усредняемого признака является мультипликативной величиной, т.е. определяется не суммированием, а умножением индивидуальных значений признака.** Она применяется только в форме средней арифметической простой:

$$\bar{x}_{geom} = \sqrt[n]{x_1 \cdot x_2 \cdot \dots \cdot x_n} = \sqrt[n]{\prod_{i=1}^n x_i}.$$

- **Средняя квадратическая** применяется в тех случаях, когда **при замене индивидуальных значений признака на среднюю величину необходимо сохранить неизменной сумму квадратов исходных величин.** Средняя квадратическая рассчитывается, как простая:

$$\bar{x}_{kv} = \sqrt{\frac{\sum_{i=1}^n x_i^2}{N}};$$

как взвешенная:

$$\bar{x}_{kv} = \sqrt{\frac{\sum_{i=1}^m x_i^2 \cdot n_i}{\sum_{i=1}^m n_i}}.$$

Степенные средние связаны между собой правилом «мажорантности средних»:

чем выше показатель степени, тем больше количественное значение среднего показателя:

$$\bar{x}_{гарм} \leq \bar{x}_{geom} \leq \bar{x}_{арифм} \leq \bar{x}_{kv}.$$

В класс структурных средних входят мода и медиана - численные значения признаков у статистических единиц, занимающих определенное положение в ранжированном ряду статистических единиц (в вариационных рядах).

- **Медиана (Me)** - значение признака у статистической единицы, стоящей в середине ранжированного ряда и делящей совокупность на две равные по численности части.
- **Мода (Mo)** - наиболее часто встречаемое значение признак в совокупности. Мода широко используется в статистической практике при изучении покупательского спроса, регистрации цен и др.

Для дискретных вариационных рядов Mo и Me выбираются в соответствии с определениями: Mo - как значение признака с наибольшей частотой n_i ; положение Me при нечетном объеме совокупности определяется ее номером $n_{Me} = \frac{N+1}{2}$, где N - объем статистической совокупности. При четном объеме ряда Me равна средней из двух вариантов, находящихся в середине ряда.

Если статистический признак меняется непрерывно (интервальные вариационные ряды), то структурные средние рассчитываются.

Мода определяется следующим образом:

- По максимальному значению частоты определяется интервал, в котором находится значение моды. Он называется модальным.
- Внутри модального интервала значение моды вычисляется по формуле:

$$Mo = x_{Mo}'' + a_{Mo} \times \frac{n_{Mo} - n_{Mo-1}}{(n_{Mo} - n_{Mo-1}) + (n_{Mo} + n_{Mo+1})},$$

где x_{Mo}'' - нижняя граница модального интервала,

a_{Mo} - ширина модального интервала,

n_{Mo} , n_{Mo-1} , n_{Mo+1} - соответственно частоты модального, предмодального (предшествующего модальному) и постмодального (следующего за модальным) интервалов.

Для расчета медианы в интервальных рядах используется следующий подход:

- По накопленным частотам находится медианный интервал. Медианным называется интервал, содержащий центральную единицу.
- Внутри медианного интервала значение Me определяется по формуле:

$$Me = x_{Me}'' + a_{Me} \cdot \frac{\frac{N}{2} - N_{Me-1}}{n_{Me}},$$

где x_{Me}'' - нижняя граница медианного интервала,

a_{Me} - ширина медианного интервала,

N - объем статистической совокупности,

N_{Me-1} - накопленная частота предмедианного интервала,

n_{Me} - частота медианного интервала.

Содержание работы

Цель работы:

Приобретение навыков расчета абсолютных, относительных и средних величин.

Порядок выполнения работы:

- По полученным от преподавателя исходным данным выполнить расчеты абсолютной, относительной и средних величин (одной из степенных средних и одной из структурных средних).
- Выполнить анализ рассчитанных показателей:
 - объяснить выбор соответствующего исходным данным вида относительной и средней величины;
 - Объяснить содержание полученных показателей.

Тесты:

1. Показатель фондоотдачи, рассчитываемый как отношение стоимости произведенной продукции к стоимости основных производственных фондов, является:
 - а) абсолютным показателем;
 - б) интервальным показателем;
 - в) относительным показателем сравнения;
 - г) относительным показателем интенсивности развития;
 - д) относительным показателем координации.
2. Мода - это значение признака:
 - а) наиболее часто встречающегося в совокупности;
 - б) минимальное значение признака в совокупности;
 - в) максимальное значение признака в совокупности;
 - г) среднее значение признака.
3. Выберите формулу для расчета средней цены, если имеются данные о количестве проданных компьютеров в трех фирмах и ценах на них:
 - а) средней арифметической;
 - б) средней гармонической;
 - г) средней геометрической;
 - д) средней квадратической.
4. *Относительный показатель структуры характеризует:*
 - а) состав изучаемой совокупности;
 - б) интенсивность развития явления;
 - в) динамику развития явления;
 - г) степень выполнения планового задания;
 - д) степень распространения изучаемого явления в определенной среде.
5. Медиана находится в:
 - а) в середине ряда распределения;
 - б) в начале ряда распределения;

- в) в конце ряда распределения;
- г) в центре ранжированной совокупности .

Контрольные вопросы:

1. Что такое средняя величина?
2. Перечислить основные виды степенных средних. Каково их назначение?
3. Даете определения относительных показателей динамики и объясните, что она показывает?
4. Что называется относительными величинами, для каких целей они применяются?
5. Что такое средние и каковы их роль и значение?
6. Что характеризуют относительные показатели плана и выполнения плана?
7. Дайте характеристику относительных показателей интенсивности развития.
8. Для чего используются относительные показатели координации?
9. Что такое медиана, как она определяется в структурированных совокупностях?
10. Дать определение моды; как она рассчитывается для дискретных и интервальных рядов распределения?
11. Можно ли определить моду по не сгруппированным данным?
12. Назначение структурных средних.

Построение и исследование вариационного ряда

Основные определения

Вариационным рядом называют систематизированную последовательность статистических единиц, сгруппированных по количественному признаку.

Вариационный ряд состоит из двух элементов: вариант и частот.

Вариантами называются значения признака, которые он принимает в вариационном ряду; *частотой* - число статистических единиц, обладающих данным значением признака. Частота показывает, как часто то или иное значение признака встречается в статистической совокупности.

В зависимости от характера вариации статистического признака различают дискретные и интервальные ряды.

В дискретных вариационных рядах статистический признак принимает только целочисленное значение, в интервальных рядах - любое.

Построение интервального ряда

Как правило, интервальный ряд являются результатом выполнения **структурной группировки**. Изначально структурные группировки, а значит, и интервальные ряды распределения строятся как *равноинтервальные*. Построение *равноинтервального* ряда распределения осуществляется в такой же последовательности как структурных группировок:

1. Определяется количество групп, на которое разбивается статистическая совокупность. При этом учитывается только размер совокупности, поэтому расчет носит приближенный характер.

Ориентировочно число групп m можно определить по формуле Стерджесса:

$$m = 1 + 3,322 \cdot \lg N,$$

где N - объем статистической совокупности; или по соответствующей данной формуле таблице 3.1.

	15 - 24	25 - 44	45 - 89	90 - 179	180 - 339	340 и более
m	5	6	7	8	9	10

2. Рассчитываются параметры групп

К ним относятся: *ширина интервала*, *нижняя и верхняя граница интервала*.

При построении равноинтервальных рядов распределения **ширина интервала** определяется по формуле:

$$a_i = \frac{R}{m} = \frac{x_{\max} - x_{\min}}{m},$$

где a_i - ширина интервала

R - размах вариации;

$x_{\max}$, $x_{\min}$ - максимальное и минимальное значения признака в статистической совокупности, свободной от экстремальных значений.

Полученное значение ширины интервала округляется в бóльшую сторону последующему правилу:

- если эта величина имеет один знак до запятой, то округление осуществляется до десятых;
- если рассчитанная величина имеет две значимые цифры до запятой и несколько после запятой, то округление осуществляется до целого числа;
- если ширина интервала является трех и более значимым числом, то она округляется до ближайшего числа, кратного 50 или 100.

Границы интервалов определяются следующим образом:

Нижняя граница первого интервала x_1^H выбирается из условия $x_1^H \leq x_{\min}$, верхняя граница всех интервалов x_i^G , в том числе и первого, определяется по формуле:

$$x_i^G = x_i^H + a_i$$

При построении интервальных группировок обычно используются *смыкающиеся интервалы*, поэтому *нижняя граница второго и всех последующих интервалов принимаются равной верхней границе предыдущего интервала*, т.е.:

$$x_i^H = x_{i-1}^G.$$

При исследовании интервальных рядов распределения необходимо использовать характеристику, называемую *серединой интервала*. Она рассчитывается по формуле:

$$b_i = \frac{x_i^H + x_i^G}{2},$$

где b_i - середина i -ого интервала.

3. Распределение единиц совокупности по образованным группам

Такое распределение можно сделать:

- путем последовательного перебора всех единиц совокупности;
- на основе предварительного ранжирования единиц и расчленения совокупности в соответствии с установленной шириной интервала

4. **Подсчет числа единиц, попавших в каждую группу**, т.е. определение частоты n_i . При определении частот необходимо помнить, что

$$\sum_{i=1}^m n_i = N.$$

Определение частотных характеристик вариационного ряда

К частотным характеристикам вариационного ряда относятся: частота, частость, накопленная частота, накопленная частость, плотности распределения.

Частота n_i - количество статистических единиц статистической совокупности, обладающих определенным значением признака. Она является исходной нерасчетной характеристикой вариационного ряда, на основе которой рассчитываются остальные частотные характеристики. Частота определяется количеством статистических единиц статистической совокупности, попавших в определенную группу.

Частость q_i - удельный вес (доля) статистических единиц, обладающих данным значением признака:

$$q_i = \frac{n_i}{N}, \quad \text{при этом } \sum_{i=1}^m q_i = 1.$$

Накопленная частота N_{i*} - число статистических единиц, у которых значение признака не превышает данного, рассчитывается по формуле:

$$N_{i*} = \sum_{i=1}^{i*} n_i$$

Для последней m -ой группы построенного ряда распределения:

$$N_m = \sum_{i=1}^m n_i = N.$$

Накопленная частость Q_{i*} - удельный вес (доля) статистических единиц, у которых значение признака не превышает данного:

$$Q_{i*} = \sum_{i=1}^{i*} q_i;$$

Для последней m -ой группы:

$$Q_m = 1.$$

Графическое представление вариационных рядов

Вариационный ряд может быть представлен двумя графиками:

- **основным** - в виде полигона или гистограммы распределения;
- **дополнительными** - в виде кумуляты или огивы распределения.

Полигон распределения используется для изображения дискретных вариационных рядов, представляет собой ломаную линию, соединяющую точки с координатами $\{x_i; n_i(q_i)\}$.

Гистограмма распределения используется для изображения интервального вариационного ряда. **Гистограмма** представляет собой

геометрическую фигуру, состоящую из смыкающихся прямоугольников с основанием, равным ширине соответствующего интервала a_i , и высотой, равной соответствующей частоте n_i или частоте q_i .

Дополнительным графиком является график кумулятивного распределения частот (кумулята). С её помощью изображается ряд накопленных частот (частостей).

Графиком кумулятивного распределения называется ломаная линия, проходящая через точки с координатами $\{x_i; N_i^*(Q_i^*)\}$ в дискретных рядах; в интервальных рядах - с координатами $\{x_i^6; N_i^*(Q_i^*)\}$

Если при графическом представлении вариационного ряда в виде кумуляты, оси поменять местами, то получим график, называемый **огивой**.

При построении графиков рядов распределения важно соблюдать соотношение масштабов по оси абсцисс и оси ординат. Данное соотношение выбирается исходя из «правила золотого сечения», согласно которому высота графика должна быть примерно в два раза меньше его основания.

Исследование вариационного ряда

При исследовании вариационного ряда последовательно решаются следующие задачи:

- определяется положение центра распределения,
- оценивается степень однородности распределения,
- исследуется форма распределения,
- подбирается теоретический аналог.

Определение положения центра распределения

Положение центра распределения определяется с помощью следующих показателей: среднего значения признака, моды и медианы.

Среднее значение признака $\bar{x}$ рассчитывается по формуле средней арифметической взвешенной:

$$\bar{x} = \frac{\sum_{i=1}^m b_i \cdot n_i}{\sum_{i=1}^m n_i}$$

где b_i - середина соответствующего интервала.

Мода (Мо) - наиболее часто встречаемое значение признака.

В дискретных вариационных рядах Мода выбирается в соответствии с определением по максимальному значению частоты n_i .

В интервальных вариационных рядах Мода определяется следующим образом:

- по максимальному значению частоты n_i выбирается **модальный интервал** - интервал, содержащий статистическую единицу, носящую моду;
- в рамках модального интервала значение моды вычисляется по формуле:

$$Mo = x_{Mo}'' + a_{Mo} * \frac{n_{Mo} - n_{Mo-1}}{(n_{Mo} - n_{Mo-1}) + (n_{Mo} - n_{Mo+1})}$$

где x_{Mo}'' - нижняя граница модального интервала,

a_{Mo} - ширина модального интервала,

n_{Mo} , n_{Mo-1} , n_{Mo+1} - соответственно частоты модального, предмодального и послемодального интервалов.

Медиана (Me) - значение признака у статистической единицы, находящейся в центре распределения.

В дискретных вариационных рядах Медиана определяется в соответствии с определением по накопленным частотам.

В интервальных вариационных рядах Медиана определяется следующим образом:

- по накопленным частотам определяется **медианный интервал** - интервал, содержащий статистическую единицу, носящую медиану.

- в пределах медианного интервала медиана вычисляется по формуле:

$$Me = x_{Me}' + a_{Me} * \frac{N/2 - N_{Me-1}}{n_i}$$

где x_{Me}' - нижняя граница медианного интервала,

a_{Me} - ширина медианного интервала,

N_{Me-1} - накопленная частота предмедианного интервала.

Оценка однородности распределения

Для оценки степени однородности ряда используются показатели вариации: размах вариации R , дисперсия σ^2 , среднеквадратичное отклонение σ , коэффициент вариации V . Указанные показатели рассчитываются по следующим формулам:

– Размах вариации

$$R = x_{\max} - x_{\min}$$

– Дисперсия

$$\sigma^2 = \overline{x^2} - (\bar{x})^2;$$

$$\text{где } \overline{x^2} = \frac{\sum_{i=1}^m b_i^2 * n_i}{\sum_{i=1}^m n_i} - \text{средняя из квадратов значений признака.}$$

– Среднее квадратическое отклонение

$$\sigma = \sqrt{\sigma^2}$$

– Коэффициент вариации

$$V = \frac{\sigma}{\bar{x}} * 100\%.$$

Если $V > 33\%$, то ряд считается неоднородным, колеблемость признака признается высокой. В таких рядах среднее значение признака не является устойчивым и не может использоваться в качестве характеристики положения центра распределения.

Оценка формы распределения

При исследовании формы распределения оценивается степень асимметрии As и эксцесс Ex распределения.

Симметричным считается распределение, в котором частоты двух равностоящих от центра значений признака равны между собой.

Для оценки симметричности распределения могут использоваться следующие показатели:

- коэффициент асимметрии Пирсона
- коэффициент асимметрии, рассчитанный на основе центрального момента третьего порядка

Коэффициент асимметрии Пирсона применяется для оценки асимметрии в центральной части распределения:

$$As = \frac{\bar{x} - M_0}{\sigma}.$$

Степень асимметрии всего ряда распределения определяется через *коэффициент асимметрии, рассчитанный на основе центрального момента третьего порядка:*

$$As = \frac{\mu_3}{\sigma^3},$$

где μ_3 - центральный момент третьего порядка.

$$\mu_3 = \frac{\sum_{i=1}^m (x_i - \bar{x})^3 * n_i}{\sum_{i=1}^m n_i}.$$

При $As = 0$ распределение считается симметричным; при $As > 0$ наблюдается правосторонняя асимметрия; при $As < 0$ наблюдается левосторонняя асимметрия. Диапазон изменения коэффициента асимметрии : $-3 \leq As \leq +3$.

По численному значению показателей асимметрии определяется ее степень:

- при $|As| \leq 0,25$ имеет место незначительная асимметрия;
- при $0,25 < |As| < 0,5$ - умеренная асимметрия;
- при $|As| > 0,5$ - значительная.

Показатель эксцесса (плоско- или островершинности распределения) рассчитывается по формуле:

$$Ex = \frac{\mu_4}{\sigma^4} - 3,$$

где μ_4 - центральный момент четвертого порядка, рассчитываемый по формуле:

$$\mu_4 = \frac{\sum_{i=1}^m (x_i - \bar{x})^4 * n_i}{\sum_{i=1}^m n_i}$$

Если $Ex > 0$, то распределение относится к островершинным, если $Ex < 0$, то распределение относится к плосковершинным.

Содержание работы

Цель работы:

Построение и исследование вариационного ряда.

Порядок выполнения работы:

- Проверить полученные от преподавателя исходные данные на однородность; для этого исключить из них аномальные значения признака.
- По исходным данным построить вариационный ряд и рассчитать его частотные характеристики.
- Представить построенный ряд в табличной форме (форма № 3.1) и графически в виде полигона или гистограммы распределения и графика накопленных частот.
- Выполнить расчет показателей центра распределения: степенной средней, моды, медианы. Определить положение центра распределения.
- Рассчитать показатели, характеризующие вариацию признака: размах R , дисперсию σ^2 , среднее квадратическое отклонение σ , коэффициент вариации V . Оценить степень однородности распределения.
- Рассчитать показатели асимметрии - коэффициент асимметрии Пирсона As , коэффициент асимметрии на основе центрального момента третьего порядка As и эксцесс Ex . Оценить форму распределения.
- Дать полную характеристику построенного и исследованного ряда.

Построение и расчет частотных характеристик вариационного ряда распределения

№гр.	Группировочный признак		n_i	Частость q_i	Накопленн ая частота N_i , чел.	Накопленн ая частость Q_i	
	интервал						a_i
	x_i^H	x_i^6					

Тесты:

- В однородных рядах распределения положение центра распределение определяется:
 - модальным значением признака;
 - средним арифметическим значением признака;
 - медианным значением признака.
- Какой показатель позволяет сравнить степень однородности разных рядов распределения:
 - размах вариации;
 - коэффициент вариации;
 - дисперсия.
- Графически интервальный ряд распределения может быть представлен:
 - линейной диаграммой;
 - полигоном распределения;
 - гистограммой распределения.
- Для характеристики формы распределения применяются показатели:
 - мода и медиана;
 - коэффициент асимметрии Пирсона и показатель эксцесса;
 - коэффициент асимметрии на основе центрального момента 3-его порядка и медиана.
- Ряд распределения студентов университета по уровню успеваемости относится к:
 - атрибутивным;
 - интервальным;
 - дискретным

6. Частотой называется:

- а) количество статистических единиц, обладающих данным значением признака;
- б) доля статистических единиц, у которых значение признака не превышает данное;
- в) доля статистических единиц, обладающих данным значением признака.

7. К дискретным рядам распределения относится:

- а) ряд распределения населения района города по национальному составу;
- б) ряд распределения населения района города по уровню доходов;
- в) ряд распределения семей района города по количеству детей.

8. В равно интервальных рядах распределения количество однородных групп может определяться на основе:

- а) формулы Стерджесса;
- б) критерия Пирсона;
- в) эксцесса.

Контрольные вопросы:

1. Как строятся вариационные ряды распределения?
2. Чем отличаются вариационные ряды распределения от дискретных?
3. Какими частотными характеристиками обладает интервальный вариационный ряд распределения?
4. Что такое гистограмма?
5. Какие графики используются для описания дискретных рядов распределения?
6. Какие показатели используются для оценки однородности рядов распределения?
7. В каких случаях положение центра ряда распределения определяется через медиану?
8. Какой ряд распределения считается однородным?
9. Какие показатели используются для оценки формы распределения?
10. Что называется эксцессом распределения?
11. Какое распределение считается симметричным?
12. Какие показатели применяются для оценки симметричности распределения?

Исследование корреляционной связи

Основные определения

Корреляционная связь является важнейшим случаем проявления **статистической связи** между изучаемыми явлениями. **При статистической связи** строго определенному значению одного признака (факторного x_i^*) соответствует множество значений другого (результативного, y_{ij}).

Статистические связи имеют место в тех случаях, когда на результативный признак действует несколько факторных, но для описания связи используются только определяющие (учтенные) признаки.

Математически статистические связи описываются зависимостью:

$$Y_i = f(x_i) + e_i,$$

где $f(x_i)$ - влияние учтенных факторов;

e_i - влияние неучтенных факторов

Статистическую связь можно представить в виде набора локальных распределений результативного признака при каждом фиксированном значении факторного:

$$x_1 : y_{11}, y_{12}, \dots, y_{1p}$$

$$x_2 : y_{21}, y_{22}, \dots, y_{2p}$$

$$x_n : y_{n1}, y_{n2}, \dots, y_{np}$$

Каждое локальное распределение результативного признака можно описать на эмпирическом и теоретическом уровнях, рассчитать его характеристики, определить форму и положение центра в виде локальных средних $\tilde{y}_i$, $i = \overline{1, m}$.

Если при изменении значения факторного признака не меняется форма локальных распределений, но смещаются их центры, то такую связь относят к корреляционным.

Корреляционной называется связь между случайными значениями факторного признака и средними значениями результативного.

Порядок исследования корреляционной связи

Исследование корреляционной связи проводится в такой последовательности:

1. Выявление связи между признаками.
2. Описание связи в табличной, графической и аналитической форме.
3. Измерение тесноты связи.
4. Оценка достоверности связи.
5. Интерпретация модели, формулировка выводов.

Выявление связи

Наличие связи между признаками может быть выявлено на основе **построения и анализа корреляционного поля**, представляющего собой совокупность точек с координатами $\{x_i; y_i\}$. Если в расположении точек наблюдается определенная зависимость, то связь между признаками существует.

Описание связи

Выявленную связь представляют в табличной, графической и аналитической форме.

Табличное описание связи есть набор пар значений факторного и локального среднего значения результативного признака; при этом проводится ранжирование пар по факторному признаку :

x_i				
$\tilde{y}_i$				

Графически корреляционная связь описывается линией эмпирической регрессии - ломаной линией, соединяющей точки с координатами $\{x_i, \tilde{y}_i\}$.

Аналитически выявленная связь описывается корреляционной моделью. Для построения корреляционных моделей используется различные математические функции. Подбор функции осуществляется на основе анализа эмпирической линии регрессии.

Корреляционная модель считается построенной, если определены численные значения параметров в соответствующих аналитических выражениях функций. Параметры модели рассчитываются с использованием **метода наименьших квадратов**, в соответствии с которым наилучшие оценки параметров корреляционной модели можно получить при минимизации суммы квадратов отклонений эмпирических значений факторного признака от соответствующих теоретических (модельных) значений вычисленных по уравнению регрессии:

$$\sum_{i=1}^n (y_i - \hat{y}_i)^2 \rightarrow \min ;$$

где $\hat{y}_i$ - модельные значения результативного признака.

Для нахождения минимума указанного функционала необходимо найти частные производные по параметрам уравнения регрессии и приравнять их к 0. Решение полученной системы уравнений дает численное значение искомых параметров модели.

Наиболее часто для построения корреляционных моделей используются следующие функции:

а) **полином первой степени** $\hat{y} = a + bx$

Для нахождения численного значения параметров в соответствии с методом наименьших квадратов используется система нормальных уравнений

$$\begin{cases} n \cdot a + b \cdot \Sigma x = \Sigma y \\ a \cdot \Sigma x + b \cdot \Sigma x^2 = \Sigma xy \end{cases}$$

б) **полином второй степени** $\hat{y} = a + bx + cx^2$

Система нормальных уравнений для нахождения параметров имеет вид:

$$\begin{cases} n \cdot a + b \cdot \Sigma x + c \cdot \Sigma x^2 = \Sigma y \\ a \cdot \Sigma x + b \cdot \Sigma x^2 + c \cdot \Sigma x^3 = \Sigma x \cdot y \\ a \cdot \Sigma x^2 + b \cdot \Sigma x^3 + c \cdot \Sigma x^4 = \Sigma x^2 \cdot y \end{cases}$$

в) **гипербола** $\hat{y} = a + \frac{b}{x}$

Система нормальных уравнений для нахождения параметров имеет вид:

$$\begin{cases} n \cdot a + b \cdot \sum \frac{1}{x} = \Sigma y \\ a \cdot \sum \frac{1}{x} + b \cdot \sum \frac{1}{x^2} = \sum \frac{y}{x} \end{cases}$$

г) **степенная** $\hat{y} = a \cdot b^x$

Для расчета параметров по методу наименьших квадратов предварительно осуществляется линеаризация функции:

$$\ln \hat{y} = \ln a + x \cdot \ln b$$

$$\ln a = A$$

$$\ln b = B$$

В результате получаем линейную зависимость вида:

$$\hat{y} = A + Bx.$$

Построенная корреляционная модель представляется графически в виде линии теоретической регрессии - ломаной линии, соединяющей точки с координатами $(x_i, \hat{y}_i)$.

Измерение тесноты связи

Теснота связи показывает степень влияния вариации факторного признака на общую вариацию результативного.

Показателем тесноты связи для линейных моделей является **линейный коэффициент корреляции** r_{yx} , рассчитываемый по формуле:

$$r_{yx} = \frac{\Sigma xy - \frac{\Sigma x \cdot \Sigma y}{n}}{\sqrt{\left[\Sigma x^2 - \frac{(\Sigma x)^2}{n} \right] \cdot \left[\Sigma y^2 - \frac{(\Sigma y)^2}{n} \right]}}$$

В нелинейных моделях для измерения тесноты связи используют **эмпирическое корреляционное отношение** η , рассчитываемое по формуле:

$$\eta = \sqrt{\frac{\delta^2}{\sigma_y^2}},$$

где δ^2 - межгрупповая дисперсия

σ_y^2 - общая дисперсия результативного признака.

Межгрупповая дисперсия рассчитывается по формуле:

$$\delta^2 = \frac{\sum (\tilde{y}_i - \bar{y})^2 \cdot n_i}{\sum n_i},$$

Общая дисперсия результативного признака может быть рассчитана «сокращенным способом»:

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

По численному значению линейного коэффициента корреляции или эмпирического корреляционного отношения связь классифицируется по степени ее тесноты. Классификация проводится с использованием шкалы Чеддока, приведенной в таблице 4.1.

Таблица 4.1.

Шкала Чеддока

r, η	$0 \div 0,1$	$0,11 \div 0,3$	$0,31 \div 0,5$	$0,51 \div 0,70$	$0,71 \div 0,9$	$0,91 \div 0,99$	свыше 0,999
Характеристика связи	отсутствует	слабая	умеренная	заметная	заметная	сильная	Функциональная

Оценка достоверности связи

Оценка достоверности линейной связи проводится на основе проверки значимости линейного коэффициента корреляции. Значимость линейного коэффициента корреляции проверяется на основе **t - критерия Стьюдента**:

$$t_p = r_{yx} \cdot \sqrt{\frac{n-2}{1-r_{yx}^2}}$$

Если расчетное значение t - критерия больше критического значения $t_{кр}$, выбранного по таблицам распределения Стьюдента (t - распределения), то связь между результативным и факторным признаками признается достоверной, а построенная модель - надежной. Критическое значение t - критерия $t_{кр}$ выбирается в зависимости от числа степеней свободы t -распределения, определяемого как $v = n - m$ (m - число параметров уравнения регрессии), и уровня значимости $\alpha = 0.05$. Таблица распределения Стьюдента приведена в Приложении 3.

Оценка достоверности нелинейной связи проводится на основе проверки значимости эмпирического корреляционного отношения η .

Для оценки надежности построенной модели используется **критерий Фишера (F-критерий)**. Фактический уровень F -критерия сравнивается с табличным (критическим) уровнем.

Если $F_{факт} > F_{кр}$, корреляционная модель признается надежной. Фактический уровень критерия Фишера рассчитывается по формуле:

$$F_{факт} = \frac{n_T^2}{1 - \eta_T^2} \cdot \frac{n - m}{m - 1};$$

где η_T^2 - теоретический коэффициент детерминации,

m - количество параметров уравнения регрессии,

$$\eta_r^2 = 1 - \frac{\sigma_{y-\hat{y}}^2}{\sigma_y^2},$$

где $\sigma_{y-\hat{y}}^2 = \frac{\sum_{i=1}^n (y_i - \hat{y}_i)^2}{n}$ - остаточная дисперсия результативного признака

$\sigma_y^2 = \frac{\sum_{i=1}^n (y_i - \bar{y})^2}{n}$ - общая дисперсия результативного признака.

Критическое значение F -критерия выбирается по таблицам распределения

Фишера- Снедекора (F -распределения) при числе степеней свободы $V_1 = m-1$ и $V_2 = n-m$ и уровне значимости $\alpha = 0,05$.

Если $F_p > F_{кр}$, выбранного по таблицам распределения Фишера-Снедекора (F -распределения) при числе степеней свободы $V_1 = m-1$ и $V_2 = n - m$ и заданном уровне значимости, связь признается достоверной. Таблица распределения Фишера приведена в Приложении 4.

Интерпретация модели

Интерпретация заключается в статистической оценке модели и включенного в нее факторного признака, то есть в выяснении, как факторный признак влияет на результативный признак. В линейных моделях влияние факторного признака оценивается через параметр b - коэффициент регрессии: чем больше величина коэффициента регрессии, тем значительнее влияние факторного признака на результат. Особое значение имеет знак перед коэффициентом регрессии. “Плюс” - означает увеличение результативного признака при увеличении факторного, “минус” - уменьшение результативного при увеличении факторного.

Для расширения возможностей экономического анализа рассчитывается **коэффициент эластичности** Θ_x :

$$\Theta_x = K \cdot \frac{\bar{x}}{\bar{y}},$$

где K - коэффициент регрессии (в линейной модели $K = b$, в параболической - $K = c$, в гиперболической - $K = b$).

Коэффициент эластичности Θ_x показывает, на сколько процентов в среднем изменится значение результативного признака при изменении факторного на один процент.

Содержание работы

Цель работы:

Построение однофакторной корреляционной модели и ее интерпретация.

Порядок выполнения работы:

- По полученным исходным данным построить корреляционное поле и установить факт наличия, либо отсутствия связи между признаками.
- Описать выявленную связь в табличной и графической форме (в виде эмпирической линии регрессии). Эмпирическую линию регрессии представить на рис. 1
- Построить корреляционную модель, рассчитав ее параметры по методу наименьших квадратов. Корреляционную модель представить графически в виде теоретической линии регрессии на рис. 1.
- Рассчитать показатели тесноты связи, связь квалифицировать по шкале Чеддока.
- Оценить достоверность корреляционной связи (адекватность модели) по критерию Стьюдента или критерию Фишера.
- Дать интерпретацию модели, сделать общие выводы о характере связи.

Тесты

1. Функциональной считается связь между признаками, при которой:
 - а) определенному значению факторного признака соответствует строго определенное значение результативного признака;
 - б) при определенном значении факторного признака результативный признак может иметь множество значений;
 - в) при определенном значении результативного признака факторный признак может принимать несколько значений.
2. Корреляционной связью называется:
 - а) связь между средними значениями результативного и факторного признаков;
 - б) связь между средними значениями результативного признака и наблюдаемыми значениями факторного признака;
 - в) связь между наблюдаемыми значениями результативного признака и факторного признака.
3. Линейный коэффициент парной корреляции может принимать значения в диапазоне:
 - а) от -1 до +1;
 - б) от 0 до +1;
 - в) от -1 до 0.
4. Для измерения тесноты нелинейной связи используется:
 - а) коэффициент регрессии;
 - б) коэффициент линейной корреляции;
 - в) индекс корреляции.
6. Метод наименьших квадратов основан на минимизации:

а) $\sum_{i=1}^n (y_i^2 - \bar{y}_i^2);$

б) $\sum_{i=1}^n (\bar{y} - \bar{y}_i)^2;$

$$в) \sum_{i=1}^n (y_i - \hat{y}_i)^2.$$

7. Уравнение регрессии имеет вид $\hat{y} = 6.5 - 2.3x$. Коэффициент регрессии показывает, что при изменении факторного признака на одну единицу своего измерения значение результативного признака:

- а) увеличится на 2.3 единицы своего измерения;
- б) увеличится на $(6.5 - 2.3) = 4.2$ единицы;
- в) уменьшится на 2.3 единицы своего измерения.

8. Какой показатель определяет в среднем процент изменения результативного признака y при изменении факторного x на 1%:

- а) коэффициент регрессии;
- б) индекс корреляции;
- в) коэффициент эластичности.

Контрольные вопросы:

1. Чем отличается статистическая связь от функциональной связи? Приведите примеры функциональной и статистической связи.
2. Какая связь называется корреляционной?
3. Как рассчитываются параметры уравнений регрессии?
4. Чем отличается эмпирическая линия регрессии от теоретической линии регрессии?
5. Что понимается под теснотой корреляционной связи?
6. Какие показатели используются для оценки тесноты связи?
7. Каким образом можно выявить наличие статистической связи между признаками?
8. Что такое коэффициенты эластичности и зачем они рассчитываются?
9. Как интерпретируется коэффициент регрессии в линейной парной регрессии?
10. Как рассчитывается коэффициент линейной корреляции?
11. При каком значении коэффициента линейной корреляции связь между признаками будет считаться функциональной?
12. Каким образом оценивается надежность линейных и нелинейных моделей парной регрессии?

Исследование ряда динамики

Основные определения

Рядом динамики называется последовательность упорядоченных во времени значений одноименного статистического показателя.

Любой ряд динамики состоит из 2-х элементов: показателя времени t_i и соответствующего ему уровня развития явления $y_i, i = \overline{1, n}$. **Уровнем** называется численное значение статистического показателя в определенный момент (за определенный период) времени. В зависимости от формы представления показателя времени различают моментные и интервальные ряды динамики. В *моментных рядах* уровни выражают состояние явления на критический момент времени (начало месяца, квартала, года и т. д.); в *интервальных* - уровни отражают состояние явления за определенные интервалы времени (сутки, месяц, квартал и т. д.).

По форме представления уровней различают ряды абсолютных, относительных и средних величин.

Ряды динамики могут быть представлены в табличной форме и графической - в виде линейной диаграммы.

При исследовании рядов динамики решаются следующие задачи:

- определение системы характеристик динамического ряда,
- разложение динамического ряда на компоненты,
- выявление тренда,
- прогнозирование на основе экстраполяции тренда,
- выявление сезонности.

Система характеристик динамического ряда

Система характеристик динамического ряда включает в себя:

- индивидуальные характеристики,
- обобщающие характеристики.

Индивидуальные характеристики рядов динамики рассчитываются цепным и базисным методами. К ним относятся:

- **Абсолютное изменение уровней (прирост или снижение) уровней**

Абсолютное изменение уровней (прирост или снижение) показывает, на сколько данный уровень выше или ниже уровня, принятого за базу.

– *Цепное абсолютное изменение уровня:*

$$\Delta y_{цi} = y_i - y_{i-1}, \quad i = \overline{2, n}.$$

где y_i, y_{i-1} - соответственно данный и предыдущий уровень ряда
 n - количество уровней динамического ряда

– *Базисное абсолютное изменение уровня:*

$$\Delta y_{\delta i} = y_i - y_1, \quad i = \overline{2, n}.$$

Абсолютные базисные и цепные изменение уровня связаны между собой:

$$\sum_{i=2}^n \Delta y_{\psi i} = \Delta y_{\delta n}$$

○ **Коэффициент (темн) роста**

Коэффициентом роста называется относительная величина, выраженная в долях единицы (K_i), та же самая величина, выраженная в процентах называется темпом роста (T_i).

Цепные коэффициент $K_{\psi i}$ и темп роста $T_{\psi i}$ рассчитываются по формулам:

$$K_{\psi i} = \frac{y_i}{y_{i-1}}, \quad i = \overline{2, n}$$

$$T_{\psi i} = \frac{y_i}{y_{i-1}} \cdot 100, \quad i = \overline{2, n}$$

– *Базисные:*

$$K_{\delta i} = \frac{y_i}{y_1}, \quad i = \overline{2, n}$$

$$T_{\delta i} = \frac{y_i}{y_1} \cdot 100, \quad i = \overline{2, n}$$

Цепные и базисные коэффициенты роста связаны между собой:

$$K_{\delta n} = \prod_{i=2}^n K_{\psi i}$$

○ **Коэффициент (темн) прироста**

– *цепной*

$$K'_{\psi i} = \frac{\Delta y_{\psi i}}{y_1}, \quad i = \overline{2, n}$$

$$T'_{\psi i} = \frac{\Delta y_{\psi i}}{y_1} \cdot 100, \quad i = \overline{2, n}$$

– *базисный*

$$K'_{\delta i} = \frac{\Delta y_{\delta i}}{y_1}, \quad i = \overline{2, n}$$

$$T'_{\delta i} = \frac{\Delta y_{\delta i}}{y_1} \cdot 100, \quad i = \overline{2, n}$$

Коэффициенты прироста и роста связаны между собой зависимостью:

$$K'_i = K_i - 1.$$

○ ***Абсолютное значение 1% прироста***

Абсолютное значение одного процента прироста используется для оценки значения полученного темпа прироста. Он показывает, какое абсолютное значение соответствует одному проценту прироста. Показатель считается по цепным характеристикам:

$$A_i = \frac{\Delta y_{ti}}{T_{ti}} = \frac{y_i - y_{i-1}}{\frac{y_i - y_{i-1}}{y_{i-1}} \cdot 100} = \frac{y_{i-1}}{100}.$$

• **Обобщающие характеристики ряда динамики**

К ним относятся:

○ ***средний уровень ряда $\bar{y}$***

В интервальных рядах значение среднего уровня ряда рассчитывается по формуле средней арифметической простой:

$$\bar{y} = \frac{\sum_{i=1}^n y_i}{n};$$

в моментных рядах - по формуле средней хронологической

$$\bar{y} = \frac{1/2 y_1 + y_2 + \dots + y_{n-1} + 1/2 y_n}{n-1}.$$

○ ***среднее абсолютное изменение (прирост или снижение) уровней:***

$$\bar{\Delta y} = \frac{\sum \Delta y_{ti}}{n-1} = \frac{\Delta y_{\bar{6}n}}{n-1};$$

○ ***Средний коэффициент роста $\bar{K}$***

$$\bar{K} = \sqrt[n-1]{\prod_{i=2}^n K_{ti}} = \sqrt[n-1]{K_{\bar{6}n}};$$

○ ***Средний коэффициент прироста $\bar{K}'$***

$$\bar{K}' = \bar{K} - 1$$

○ ***Дисперсия уровней динамического ряда***

$$\sigma_y^2 = \frac{\sum_{i=1}^n (y_i - \bar{y})^2}{n}$$

○ ***Среднее квадратическое отклонение уровней***

$$\sigma_y = \sqrt{\sigma_y^2};$$

○ **Коэффициент вариации уровней ряда динамики**

$$V_y = \frac{\sigma_y}{y}.$$

Анализ рядов динамики

Анализ ряда динамики в данной работе сводится к выявлению и оценке основной долговременной тенденции развития (тренда).

Для выявления тренда применяются следующие методы:

- *метод укрупнения интервалов;*
- *метод скользящей средней;*
- *метод аналитического выравнивания.*

Метод укрупнения интервалов основан на укрупнении периодов времени, к которым относятся уровни ряда. Для каждого укрупненного интервала рассчитывается новое значение уровня как среднее арифметическое по этому интервалу.

Метод скользящих средних заключается в замене уровней исходного ряда теоретическими уровнями, рассчитанными по формуле скользящей средней. При этом, как и в предыдущем методе, происходит укрупнение интервалов. Число уровней, по которым укрупняется интервал, называется *диапазоном укрупнения* или *периодом скольжения* α . При расчете скользящих средних последовательно исключаются из принятого периода скольжения первые уровни и включаются последующие. В практических расчетах часто период скольжения выбирается равным $\alpha=3$. Скользящая средняя при $\alpha=3$ вычисляется по формуле:

$$\hat{y}_i = \frac{y_{i-1} + y_i + y_{i+1}}{3}, \quad i = \overline{2, n-1}$$

Метод аналитического выравнивания предполагает, что основная тенденция развития рассчитывается как временная функция $\hat{y} = f(t)$. Подбор функции производится на основе анализа закономерностей динамики данного явления. Подобранная функция используется для построения трендовой модели. Модель считается построенной, если рассчитаны её параметры. Расчет параметров проводится методом наименьших квадратов (МНК).

Наиболее часто в практических расчетах используются следующие функции:

- *при стабильных абсолютных цепных приростах $\Delta y_{\text{цч}} \approx \text{const}$ -*
полином первой степени $y = a + bt$.

Для нахождения статистических оценок параметров модели необходимо решить систему уравнений вида:

$$\begin{cases} n \cdot a + b \sum t = \sum y, \\ a \cdot \sum t + b \sum t^2 = \sum yt \end{cases};$$

где y - фактические (эмпирические) уровни ряда;

t - условные показатели времени.

- для рядов динамики, в которых имеет место смена направления развития - полином второй степени вида $\hat{y} = a + b \cdot t + c \cdot t^2$.

Для нахождения параметров которого необходимо найти решение систему уравнений следующего вида:

$$\begin{cases} n \cdot a + b \cdot \sum t + c \cdot \sum t^2 = \sum y, \\ a \cdot \sum t + b \cdot \sum t^2 + c \cdot \sum t^3 = \sum yt, \\ a \cdot \sum t^2 + b \cdot \sum t^3 + c \cdot \sum t^4 = \sum yt^2. \end{cases}$$

- при стабильных цепных коэффициентах роста $K_{ци} \approx \text{const}$ показательная функция $\hat{y} = a \cdot b^t$.

Для расчета параметров по методу наименьших квадратов предварительно осуществляется линеаризация функции:

$$\ln \hat{y} = \ln a + t \cdot \ln b$$

$$\ln a = A$$

$$\ln b = B$$

$$\hat{y} = A + Bt.$$

Для получения статистических оценок параметров временных функций необходимо решить проблему преобразования хронологических показателей времени в условные показатели. Для этого используется “метод отсчета от условного нуля”, в соответствии с которым начало координат по признаку времени переносится в середину ряда так, чтобы $\sum_{i=1}^n t_i = 0$. Построенная

трендовая модель **проверяется на надежность** (существенность) корреляционной связи. Для оценки надежности трендовой модели используется **критерий Фишера (F-критерий)**. Фактический уровень F -критерия сравнивается с табличным (критическим) уровнем.

Если $F_{\text{факт}} > F_{\text{кр}}$, трендовая модель признается надежной. Фактический уровень критерия Фишера рассчитывается по формуле:

$$F_{\text{факт}} = \frac{n_T^2}{1 - \eta_T^2} \cdot \frac{n - m}{m - 1};$$

где η_T^2 - теоретический коэффициент детерминации,
 m - количество параметров трендового уравнения,

$$\eta_T^2 = 1 - \frac{\sigma_{y-\hat{y}}^2}{\sigma_y^2},$$

где $\sigma_{y-\hat{y}}^2 = \frac{\sum_{i=1}^n (y_i - \hat{y}_i)^2}{n}$ - остаточная дисперсия уровней.

Критическое значение F -критерия выбирается по таблицам распределения Фишера- Снедекора (F -распределения) при числе степеней свободы $V_1 = m-1$ и $V_2=n-m$ и уровне значимости $\alpha = 0,05$ (Приложение 4)

Прогнозирование на основе экстраполяции тренда

Выявление и оценка основной тенденции развития дают основание для **прогнозирования** - *определения возможных значений уровней ряда динамики в будущем*. Прогнозирование предполагает, что **закономерность развития, действующая в прошлом (внутри ряда динамики), сохранится и в будущем, то есть прогноз основан на экстраполяции**.

Экстраполяция, в общем виде, может быть представлена зависимостью

$$\hat{y}_{n+l} = f(y_i, l, q),$$

где $\hat{y}_{n+l}$ - прогнозируемое значение уровня;

y_i - текущий уровень ряда динамики;

l - период упреждения прогноза;

q_j - параметры уравнения тренда.

В статистике рассчитываются два вида прогноза: точечный и интервальный. Для прогнозирования используются различные методы экстраполяции. Наиболее эффективным является экстраполяция на основе трендовой модели. На основе этого метода строятся два вида прогноза - *точечный и интервальный*.

Точечным прогнозом называется *единственное значение прогнозируемого показателя*. Точечный прогноз на основе трендовой модели определяется как:

$$\hat{y}_{n+l} = f(t_i) \quad i = n + l.$$

Точечные прогнозы используются для построения **интервальных прогнозов**, считающихся более качественными. В интервальных прогнозах с определенной вероятностью можно ожидать появления фактического значения прогнозируемого экономического показателя.

Интервальный прогноз предполагает определение величины его **доверительного интервала**:

$$\hat{y}_{n+l} \pm t_{\alpha} \cdot \sigma_{y-\hat{y}}$$

t_{α} - коэффициент доверия по распределению Стьюдента

$\sigma_{y-\hat{y}}$ - средняя квадратическая ошибка тренда

$$\sigma_{y-\hat{y}} = \sqrt{\frac{\sum_{i=1}^n (y_i - \hat{y}_i)^2}{n}}.$$

Содержание работы

Цель работы:

Расчет динамических характеристик ряда, выявление и описание тренда; прогнозирование на основе экстраполяции тренда.

Порядок выполнения работы:

- Анализировать ряд динамики представить графически на рис. 1
- По исходным данным построить систему динамических характеристик ряда. Для этого рассчитать следующие показатели:
 - а) абсолютные приросты: цепные $\Delta y_{ци}$ и базисные $\Delta y_{би}$,
 - б) коэффициенты роста: цепные $K_{ци}$ и базисные $K_{би}$,
 - в) коэффициенты прироста: цепные $K'_{ци}$ и базисные $K'_{би}$,
 - г) абсолютное значение 1% прироста,
 - д) средний уровень ряда $\bar{y}$
 - е) средний абсолютный прирост $\overline{\Delta y}$,
 - ж) средний коэффициент роста $\bar{K}$,
 - з) средний коэффициент прироста $\bar{K}'$,
 - и) дисперсию и среднее квадратичное отклонение уровней ряда σ_y^2 и σ_y ,
 - к) коэффициент вариации уровней ряда V_y ,

Полученные результаты представить в форме 5.1.

- Выявить основную тенденцию методом укрупнения интервалов. Исходные интервалы увеличить в 3 раза. Расчеты провести в форме 5.2. Выявленную тенденцию представить графически на рис. 1.
- Выявить основную тенденцию методом скользящей средней. Скользящую среднюю рассчитать при $\alpha=3$. Расчет провести в форме 2. Выявленную тенденцию представить графически на рис. 1.
- Выявить основную тенденцию методом аналитического выравнивания. Выявленную тенденцию представить графически на рис. 1.
- Оценить надежность трендовой модели по F -критерию при $\alpha=0,05$. $F_{кр.}$ определить по Приложению 1.
- Рассчитать точечные и интервальные прогнозы при периодах упреждения $l=1$ и $l=3$.
- Рассчитать относительную ошибку прогноза:

$$E_{np} = \frac{y_{иф} - \hat{y}_i}{\hat{y}_i} \cdot 100\%$$

- Сделать выводы о характере динамики изучаемого явления, качестве построенных прогнозов.

Форма 5.1

Индивидуальные динамические характеристики ряда

Хронологический период времени t_i	Уровни ряда y_i	Абсолютные приросты		Коэффициенты роста		Коэффициенты прироста		Абсолютное значение 1% прироста
		Цепные $\Delta y_{ци}$	Базисные	Цепные $K_{ци}$	Базисные $K_{би}$	Цепные $K'_{ци}$	Базисные $K'_{би}$	A_i
1								
2								
n								

Форма 5.2

Расчетная таблица для определения средних по методу укрупнения интервалов и по методу скользящих средних

Показатель времени t_i	Уровни ряда y_i	Метод укрупнения интервалов			Метод скользящей средней	
		новый показатель в времени t_i	сумма уровней за интервал	среднее значение	Скользящая сумма	Скользящая средняя
1					-	-
2		I				
3						
					-	-
		K				
Σ		-		-	-	-

Тесты:

1. Что характеризует ряд динамики?

- изменение явления во времени;
- распределение единиц совокупности по какому-либо признаку;
- распространение изучаемого явления в пространстве.

2. На расчетном счете предприятия остаток средств на 1 января (тыс. руб.): 2014г. - 4000, 2013г. - 4100. Отношение второй величины к первой, 1,025 называется:

- темп роста;

- б) коэффициент роста;
 - в) коэффициент прироста.
3. Количественное описание тренда (трендовую модель) можно получить на основе:
- а) метода укрупнения интервалов;
 - б) гармонического анализа;
 - в) метода аналитического выравнивания.
4. Статистические прогнозы делятся на:
- а) точные и приближенные;
 - б) точечные и интервальные;
 - в) моментные и интервальные.
5. Если фирма продала в сентябре 20 персональных компьютеров, а в октябре - 30, то абсолютное изменение объема продаж составит:
- а) $\Delta_q = q_1 - q_0 = 30 - 20 = +10$ штук;
 - б) $\Delta_q = q_0 - q_1 = 20 - 30 = -10$ штук;
 - в) $\Delta_q = q_0 - q_1 = 20 + 30 = 50$ штук.
6. Под экстраполяцией понимается:
- а) определение значений уровней прошлых периодов на основе тренда;
 - б) определение значений уровней будущих периодов на основе продолжения тренда;
 - в) построение трендовой модели;
 - г) предположение о том, что закономерность развития, действующая в прошлом, сохранится и в будущих периодах.
7. Оценка надежности трендовой модели определяется на основе:
- а) критерия согласия Пирсона (χ^2 -критерия);
 - б) критерия согласия Фишера (F-критерия);
 - в) критерия согласия Пуассона.

Контрольные вопросы

1. Чем отличаются интервальные ряды динамики от моментных?
2. Для решения каких задач предназначены ряды динамики?
3. Дайте определение тренда.
4. Дайте характеристику метода укрупнения интервалов.
5. В чем сущность выравнивания рядов динамики методом скользящей средней?
6. В чем сущность методов аналитического выравнивания?
7. Какова структура ряда динамики?
8. Чем отличается аддитивная модель разложения ряда динамики от мультипликативной?
9. Что представляют собой сезонные колебания, для чего необходимо их изучение?
10. Что такое экстраполяция ряда динамики?

11. Какие виды статистических прогнозов можно получить на основе экстраполяции тренда?
12. Как исчисляются показатели абсолютного изменения уровней и что они характеризуют?
13. Как исчисляется и что характеризует показатель «абсолютное значение одного процента прироста»?
14. Как рассчитывается и что характеризует средний коэффициент роста?
15. Чем отличаются ряды динамики от рядов распределения?

Выборочное наблюдение

Основные определения

Выборочным называется несплошное наблюдение, при проведении которого обследуется определенная часть единиц изучаемой статистической совокупности. При этом изучаемая совокупность называется *Генеральной*, а обследуемая ее часть - **выборочной совокупностью или выборкой**. Выборочные наблюдения проводятся с целью получения обобщающих характеристик генеральной совокупности на основе характеристик выборки.

Для обеспечения необходимой точности результатов **выборочная совокупность должна быть репрезентативной**, то есть полностью представлять все свойства и пропорции генеральной совокупности. Репрезентативность выборки обеспечивается применением случайных методов отбора при ее формировании.

К обобщающим характеристикам генеральной совокупности относятся:

$\bar{d}$ - **генеральная доля** (удельный вес статистических единиц генеральной совокупности, обладающих определенным значением признака);

$\bar{x}$ - **генеральная средняя** (среднее значение признака в генеральной совокупности);

σ^2 - **генеральная дисперсия** (средний квадрат отклонений индивидуальных значений признака в генеральной совокупности от генеральной средней);

σ - **генеральное среднее квадратическое отклонение**.

Перечисленные обобщающие характеристики рассчитываются на основе соответствующих характеристик выборки:

ω - **выборочная доля** (удельный вес статистических единиц, обладающих определенным значением признака в выборке);

$\tilde{x}$ - **выборочная средняя** (среднее значение признака в выборке);

s^2 - **выборочная дисперсия**;

s - **выборочное среднее квадратическое отклонение**.

Основные обобщающие характеристики генеральной совокупности получают в виде **статистических оценок - точечных и интервальных**. Более качественными из указанных статистических оценок являются интервальные оценки.

Расчет **интервальных оценок** соответствующих обобщающих характеристик генеральной совокупности производится по формулам:

$$\bar{d} = \omega \pm \Delta_{\omega},$$

$$\bar{x} = \tilde{x} \pm \Delta_{\tilde{x}}$$

где Δ_{ω} , $\Delta_{\tilde{x}}$ - предельные ошибки выборочной доли и выборочной средней соответственно.

В свою очередь, **предельные ошибки выборки** рассчитываются по формулам:

$$\Delta\omega = t\mu_{\omega}$$

$$\Delta\tilde{x} = t \cdot \mu_{\tilde{x}}$$

где t - коэффициент доверия Стьюдента;

$\mu_{\omega}, \mu_{\tilde{x}}$ - средние ошибки выборочной доли и выборочной средней соответственно.

Значение коэффициента доверия Стьюдента выбирается по специальным математическим таблицам. Наиболее часто употребляются следующие значения t и соответствующих уровней вероятностей:

t	1.0	1.5	2.0	2.5	3.0	3.5
$\varphi(t)$	0.683	0.866	0.954	0.988	0.997	0.999

Определение ошибок выборки

Величина **средней ошибки выборки** зависит от способа формирования выборки. При случайном повторном отборе средняя ошибка определяется по формуле:

$$\mu = \sqrt{\frac{s^2}{n}},$$

где n - объем выборочной совокупности.

При случайном бесповторном отборе средняя ошибка равна:

$$\mu = \sqrt{\frac{s^2}{n} \left(1 - \frac{n}{N}\right)},$$

где N - объем генеральной совокупности.

Обоснование необходимого объема выборки

При подготовке выборочного наблюдения одним из важнейших вопросов является **определение необходимого объема выборки**, так как это влияет на величину фактической ошибки.

Объем выборки при повторном отборе определяется по формуле:

$$n = \frac{t^2 s^2 N}{\Delta},$$

при бесповторном:

$$n = \frac{t^2 s^2 N}{\Delta^2 N + t^2 s^2}.$$

Одним из наиболее сложных вопросов при определении необходимого объема выборки является определение величины выборочной дисперсии, так как данные для ее расчета перед началом выборочного наблюдения бывают неизвестными. Поэтому для оценки вариации изучаемого признака используются данные пробного обследования.

По пробной выборке дисперсия рассчитывается по формуле:

$$s^2 = \frac{\sum_{i=1}^n (x_i - \tilde{x})^2}{n_{np}}$$

где $x_i, \tilde{x}$ - соответственно наблюдаемое значение признака и выборочная средняя по данным пробного обследования;

n_{np} - объем пробной выборки.

Выборочная средняя $\tilde{x}$ рассчитывается по формуле:

$$\tilde{x} = \frac{\sum_{i=1}^{n_{np}} x_i}{n_{np}}$$

Вычисленная таким способом *дисперсия пробной выборки используется в дальнейшем для обоснования необходимого объема выборки.*

Содержание работы:

Цель работы:

Получение статистических оценок обобщающих характеристик генеральной совокупности на основе выборочного наблюдения

Порядок выполнения работы:

- По исходным данным, полученным от преподавателя и представляющим собой генеральную совокупность, провести 10%-ую пробную выборку $n_{пр}$ методом случайного повторного или бесповторного отбора (выбирается студентом самостоятельно);
- По данным пробной выборки рассчитать ее дисперсию;
- Используя значение дисперсии пробной выборки и заданную величину предельной ошибки выборки, рассчитать необходимый объем выборки n при повторном или бесповторном отборе;
- Сформировать выборку необходимого объема, используя выбранный метод отбора;
- Обработать результаты сформированной выборки: рассчитать выборочную среднюю, выборочную дисперсию, фактические среднюю и предельную ошибки выборки;
- Рассчитать интервальную оценку параметра генеральной совокупности - генеральной средней с вероятностью $P=0.95$;
- Сделать выводы о соответствии исследуемой продукции предъявляемым к ней требованиям, сравнив интервальную оценку генеральной средней с накладываемыми ограничениями в виде допусков, ГОСТов или технических условий (ТУ);
- Определить долю бракованной продукции в исследуемой партии продукции. Для этого:

- по выборке определить выборочную долю как $w = \frac{m}{n}$, где m - количество бракованных изделий (несоответствующих допускам, ГОСТам или ТУ);
- рассчитать интервальную оценку генеральной доли.
- Сделать выводы.

Тесты:

1. Под ошибкой выборки понимается:

- а) расхождение между максимальными значениями характеристики выборки и генеральной совокупности;
- б) расхождение между минимальными значениями характеристики выборки и генеральной совокупности;
- в) объективно существующее расхождение между численными значениями наблюдаемого параметра в выборке и Генеральной совокупности.

2. На величину ошибки выборки не влияют следующие факторы:

- а) объем выборки;
- б) уровень колеблемости признака в генеральной совокупности;
- в) объем генеральной совокупности.

3. В городе проведено обследование санитарного состояния организаций сферы услуг. Были обследованы гостиницы с числом мест от 50, каждое пятое кафе и все рестораны. Выборочной была проверка:

- а) гостиниц
- б) кафе
- в) ресторанов.

4. Проведено выборочное обследование студентов двух институтов, в каждом объем выборки составил 300 чел. В первом вузе обучалось 5000 студентов, во втором - 3500. Основываясь на этих данных можно утверждать, что ошибка выборки:

- а) больше в первом вузе
- б) больше во втором
- в) одинакова в обоих вузах.

5. Чтобы провести 5%-ую механическую выборку предприятий приборостроения необходимо отобрать каждое:

- а) пятое предприятие
- б) второе предприятие
- в) двадцатое предприятие.

6. Выборкой называется:

- а) часть единиц генеральной совокупности, отобранных для ее изучения;
- б) совокупность случайно отобранных единиц генеральной совокупности с целью получения статистических оценок её параметров;
- в) совокупность закономерно отобранных единиц генеральной совокупности с целью получения статистических оценок её характеристик.

7. Наименьшую ошибку выборки обеспечивает:

- а) механическая выборка;
- б) типическая выборка;
- в) серийная.

Контрольные вопросы:

1. Даете определение выборочному наблюдению.
2. Что называется ошибкой выборки?
3. Какие ошибки называются ошибками репрезентативности?
4. Что является причиной возникновения ошибок репрезентативности?
5. Чем отличается бесповторная выборка от повторной? Какая из них точнее?
6. Перечислить преимущества выборочного метода.
7. Что характеризует средняя ошибка выборки?
8. Как связаны между собой предельная и средняя ошибки выборки?
9. Как формулируется задача выборочного исследования?
10. Как организуется механический отбор?
11. Статистические оценки каких параметров генеральной совокупности можно получить по данным выборки?
12. Какие требования предъявляются к статистическим оценкам?

Индексы

Основные определения

Индексами называются относительные величины, характеризующие изменение уровней простых или сложных социально-экономических явлений во времени, пространстве или по сравнению с планом.

Индексы относятся к обобщающим характеристикам изучаемых явлений, предназначенным для решения следующих задач:

- оценка изменений сложных явлений и отдельных их частей;
- выявление влияния отдельных факторов на общую динамику сложного явления.

Для анализа изменений простых единичных показателей (например, цен на определенный товар за несколько лет) используют **индивидуальные индексы**. Индивидуальные индексы представляют собой относительные величины (коэффициенты роста), показывающие, во сколько раз уровень простого единичного показателя в данный период больше или меньше по сравнению с какой-либо базой.

Для оценки изменений сложных явлений, состоящих из несоизмеримых элементов применяются **общие индексы**. В зависимости от формы представления исходных данных и применяемых методов расчета общие индексы делятся на *агрегатные и средние из индивидуальных*. Ведущей формой *общих индексов* являются **агрегатные**. В агрегатных индексах отдельные элементы индексируемой величины не являются сопоставимыми. Для оценки изменения таких величин необходимо *отдельные их элементы, составляющие содержание сложных явлений, привести к соизмеримому виду*; и только после этого сопоставлять полученные по всем элементам суммы за два периода.

В тех случаях, когда исходные данные не позволяют использовать агрегатную форму расчета, применяются средние из индивидуальных.

По содержанию и характеру изучаемых показателей различают индексы количественных и качественных показателей.

Для удобства применения индексов используется определенная символика и специальная терминология.

Каждая индексируемая величина имеет свое обозначение:

q - количество продукции одного вида в натуральном выражении,

p - цена единицы продукции.

Индивидуальные индексы обозначаются следующими символами:

i_q - индивидуальный индекс физического объема,

i_p - индивидуальный индекс цен,

Общие (сводные) индексы имеют обозначения:

I_q - общий индекс физического объема,

I_p - общий индекс цен,

При расчете индексов используются **два вида данных**:

- **данные базисного уровня** - уровня, с которым производится сравнение; для их обозначения к символу соответствующего показателя добавляется «0».
- **данные текущего уровня** - уровня, который сравнивается. Они обозначаются добавлением «1» к символу соответствующего показателя.

Расчет общих индексов

Для расчета общего индекса результирующего показателя в агрегатной форме, например, агрегатного индекса стоимости, необходимо перейти от количества несоизмеримых в физических единицах объемов реализации к **стоимостной оценке**.

Стоимость реализованной продукции определяется по формуле:

$$Q_i = \sum_{i=1}^n p_i \cdot q_i$$

где p_i - цена единицы i -ой продукции;

q_i - количество (физический объем) единиц i -ой продукции;

i - вид продукции, $i = \overline{1, n}$

n - число видов продукции.

В этом случае стоимость продукции определяется:
в отчетном периоде:

$$Q_1 = \sum_{i=1}^n p_1 \cdot q_i,$$

в базисном периоде:

$$Q_0 = \sum_{i=1}^n p_0 \cdot q_0.$$

Агрегатный индекс стоимости, показывающий, во сколько раз стоимость продукции отчетного периода отличаются от стоимости продукции базисного периода, имеет вид:

$$I_{pq} = \frac{Q_1}{Q_0} = \frac{\sum_{i=1}^n p_1 \cdot q_1}{\sum_{i=1}^n p_0 \cdot q_0}.$$

Основная функция общих (агрегатных, т. е. относящихся к наборам видов продукции) индексов - аналитическая. С их помощью выявляется мера влияния на рост затрат показателей двух видов: количественного и качественного. Соответственно и индексы делятся на две группы: **индексы количественного показателя и индекс качественного показателя**.

Агрегатный индекс количественного показателя характеризует изменение общей стоимости произведенной продукции под влиянием изменения количества (физических объемов) разнородной продукции. В этом индексе переменными (индексируемыми) являются физические объемы. Качественный показатель в индексе физического объема является показателем-соизмерителем.

Агрегатный индекс физического объема определяется по формуле Ласпейреса:

$$I_q = \frac{\sum q_1 \cdot p_0}{\sum q_0 \cdot p_0}.$$

Общий индекс количественного показателя может быть рассчитан и через индивидуальные индексы физического объема отдельных видов продукции

$$i_q = \frac{q_1}{q_0}.$$

Такой индекс называется **средним арифметическим индексом количественного показателя** и рассчитывается по формуле:

$$I_q = \frac{\sum i_q p_0 q_0}{\sum p_0 q_0}.$$

Вторым фактором, вызывающим изменение общей стоимости произведенной продукции, является качественный показатель. **Агрегатный индекс качественного показателя** показывает, во сколько раз изменилась стоимость продукции отчетного периода по сравнению с базисным периодом в связи с изменением цен на произведенную продукцию, и рассчитывается по формуле Пааше:

$$I_p = \frac{\sum p_1 \cdot q_1}{\sum p_0 \cdot q_1}.$$

Общий индекс качественного показателя может быть рассчитан и по индивидуальным индексам качественного показателя (индивидуальным индексам цен) как среднее взвешенное гармоническое значение этих индивидуальных индексов с весами, соответствующими стоимости отдельных видов продукции отчетного периода:

$$I_p = \frac{\sum p_1 \cdot q_1}{\sum \frac{1}{i_p} p_1 \cdot q_1}.$$

Построение индексной системы

Существует очевидная зависимость агрегатных индексов общей стоимости произведенной продукции, цен и физического объема:

$$I_{pq} = I_p \cdot I_q$$

Приведенная взаимосвязь экономических индексов относится к **двухфакторной индексной системе**, с помощью которой можно разложить изменение общей стоимости произведенной продукции как результирующего показателя на изменение, вызванное влиянием изменений физических объемов по отдельным видам продукции и цен на нее.

На основе построенной индексной системы можно разложить **абсолютное** изменение общей стоимости на сумму факторных изменений, вызванных динамикой количественного и качественного показателей. Общее абсолютное изменение стоимости равно:

$$\Delta_{pq} = Q_1 - Q_0 = \sum_i p_1 q_1 - \sum_i p_0 q_0.$$

Абсолютные изменения факторного вида также равны разности числителя и знаменателя соответствующих индексов.

Изменение стоимости в связи с изменением количественного показателя определяется как:

$$\Delta_q = \sum q_1 \cdot p_0 - \sum q_0 \cdot p_0,$$

в связи с динамикой качественного показателя:

$$\Delta_p = \sum p_1 \cdot q_1 - \sum p_0 \cdot q_1.$$

При этом сумма абсолютных изменений общей стоимости произведенной продукции по факторам равна общему изменению общей стоимости произведенной продукции в отчетном периоде по сравнению с базисным:

$$\Delta_{pq} = \Delta_p + \Delta_q.$$

Индексный анализ динамики среднего уровня.

Индексный метод в некоторых случаях используется и для изучения динамики соизмеримых между собой величин. При этом он позволяет выявлять факторы, влияющие на динамику среднего уровня качественного показателя для продукции одного и того же назначения, отличающейся только ценами за единицу физического объема. Для такой продукции в случае объединения ее физических объемов можно вводить *средние уровни соответствующих качественных признаков* (например, средняя цена цемента разных марок, выпускаемого предприятием).

Средний уровень качественного показателя в отчетном периоде может отличаться от среднего уровня в базисном периоде по двум причинам: динамикой качественного показателя на отдельные элементы затрат (например, с изменением цены тонны цемента каждой марки) и изменением соотношения (структуры) физических объемов.

Если анализируется однородная (соизмеримая) продукция одного и того же назначения, то для характеристики изменения среднего уровня качественного показателя рассчитываются следующие индексы:

– **индекс среднего уровня качественного показателя переменного состава;**

– *индекс среднего уровня качественного показателя постоянного (фиксированного) состава;*

– *индекс структурного сдвига.*

Индекс переменного состава характеризует общее изменение среднего уровня качественного показателя (цены) вследствие изменения качественного и количественного (структурного) факторов/ Он рассчитывается по формуле:

$$I_{p_{\text{перем.}}} = \frac{\bar{p}_1}{\bar{p}_0} = \frac{\sum p_{1i} q_{1i} / \sum q_{1i}}{\sum p_{0i} q_{0i} / \sum q_{0i}}.$$

Индекс постоянного состава характеризует динамику среднего уровня качественного показателя в связи с изменением качественного показателя при неизменности физических объемов. При этом количественный (неиндексируемый) признак фиксируется на отчетном уровне:

$$I_{p_{\text{(пост.сост.)}}} = \frac{\sum p_1 q_1 / \sum q_1}{\sum p_0 q_1 / \sum q_1} = \frac{\sum p_1 q_1}{\sum p_0 q_1}.$$

Индекс структурного сдвига характеризует влияние изменения структуры. Этот индекс предполагает учет качественного показателя (неиндексируемого в нем) на базисном уровне:

$$I_{\text{стр.сдв}} = \frac{\sum p_0 q_{1i} / \sum q_{1i}}{\sum p_0 q_{0i} / \sum q_{0i}}.$$

Между *индексами переменного состава, переменного состава и структурного сдвига* существует взаимосвязь следующего вида:

$$I_{\bar{p}_{\text{(перем.сост.)}}} = I_{\bar{p}_{\text{(пост.сост.)}}} \cdot I_{\text{стр.сдв.}}$$

Содержание работы:

Цель работы:

Изучение и расчет индексов на примере анализа объема произведенной продукции; построение индексных систем и проведение факторного анализа.

Порядок выполнения работы:

- ***Считать продукцию предприятия разнородной*** и выполнить следующие расчеты:
- Определить, как изменился общий объем произведенной продукции в отчетном периоде по сравнению с базисным периодом: во сколько раз, на сколько процентов, тыс. рублей; для этого рассчитать индекс общего объема произведенной продукции и абсолютное изменение объема произведенной продукции;
- Определить, как повлияло изменение физических объемов выпускаемой продукции на динамику общего объема произведенной продукции: во сколько раз, на сколько процентов, каково абсолютное изменение объема произведенной продукции в связи с ростом физических объемов продукции. Для этого рассчитать индекс физического объема и изменение общего объема произведенной продукции в связи с изменением количества продукции отдельных видов;
- Определить, какие изменения общего объема произведенной продукции (во сколько раз, на сколько процентов, тыс. рублей) вызвало изменение цены единицы отдельных видов продукции. Для этого рассчитать индекс качественного показателя (цены) и изменение общего объема произведенной продукции в связи с динамикой этого показателя. Расчеты проводить в форме 7.1;
- Определить, как изменились физические объемы и индивидуальные цены единицы продукции каждого вида. Для этого рассчитать индивидуальные индексы физического объема отдельных видов продукции и индивидуальные индексы цен. Расчеты проводить в форме 7.1;
- Оценить динамику физического объема и цен через индивидуальные индексы количественного и качественного показателей. Для расчета использовать среднюю форму общих индексов;
- Провести факторный анализ динамики общего объема произведенной продукции. Для этого построить индексную систему и определить абсолютное и относительное влияние качественного и количественного показателей (цены и физического объема).
- ***Считать продукцию предприятия однородной*** и выполнить на основании тех же исходных данных следующие задания:
- Определить изменение среднего значения качественного показателя, учитывающее влияние динамики качественного показателя и

количественного (структурного) показателей. Рассчитать индекс переменного состава;

- Определить влияние на динамику средней себестоимости единицы продукции следующих факторов:
 - а) изменение цены единицы отдельных видов продукции; рассчитать индекс среднего уровня качественного показателя постоянного состава;
 - б) структурных сдвигов в объеме выпускаемой продукции; рассчитать индекс структурного сдвига

Форма 7. 1.

Промежуточные данные для расчета индексов

i	$p_0 q_0$	$p_1 q_1$	$p_0 q_1 i$	i_q	i_p	$i_q p_0 q_0$	$\frac{1}{i_p} p_1 q_1$
1							
2							
3							
$\sum_{i=1}^3$				-	-		

Тесты:

1. В агрегатных индексах качественных показателей весами (показателями-соизмерителями) служат:

- а) объемные (количественные) показатели;
- б) средние уровни объемных показателей;
- в) качественные показатели.

2. Из нижеследующих формул выберите формулу для расчета агрегатного индекса физического объема продаж:

а) $I_p = \frac{\sum i_p p_0 q_0}{\sum p_0 q_0}$;

б) $I_q = \frac{\sum q_1 p_0}{\sum q_0 p_0}$;

в) $I_q = \frac{\sum i_q p_0 q_0}{\sum p_0 q_0}$.

3. Цены на реализованную АО продукцию в отчетном году по сравнению с базисным увеличены в среднем в 1,2 раза, а количество реализованной продукции стало меньше на 20%. Индекс товарооборота следует исчислить по формуле:

а) $I_{pq} = I_q \cdot I_p = 0,8 \cdot 1,2 = 0,96$,

б) $I_q = I_{pq} \div I_p = 0,8 \div 1,2 = 0,667$

в) $I_p = I_{pq} \div I_q = 1,2 \div 0,8 = 1,5$

4. Из нижеследующих формул выберите формулу для расчета агрегатного индекса цен:

а) $I_p = \frac{\sum i_p p_0 q_0}{\sum p_0 q_0}$;

б) $I_p = \frac{\sum p_1 q_0}{\sum p_0 q_0}$;

в) $I_p = \frac{\sum p_1 q_1}{\sum \frac{p_1 q_1}{i_p}}$

5. Индексами называются:

- а) относительные величины, характеризующие изменения сложных социально-экономических явлений;
- б) абсолютные величины, характеризующие изменения простых явлений во времени, в пространстве и по сравнению с планом;
- в) Средние величины, характеризующие уровень развития простых и сложных социально-экономических явлений

6. Индексные системы предназначены для:

- а) оценки изменения сложного социально-экономического явления во времени;
- б) для проведения факторного анализа;
- в) расчета неизвестных средних индексов.

7. Для оценки изменений сложных неоднородных явлений используются:

- а) агрегатные индексы;
- б) индивидуальные индексы;
- в) индексы переменного состава.

Контрольные вопросы:

1. Чем отличается индекс от относительных величин?
2. Какие формы индексов Вам известны?
3. Опишите структуру агрегатного индекса.
4. Какие задачи решаются с помощью индексов?
5. Что представляют собой индивидуальные индексы?
6. Как исчисляются агрегатные индексы цен (Пааше и Ласпейреса)?
7. Что представляет собой индексная система? С какой целью она строится?
8. Что представляют собой средние индексы? Каким образом они выводятся?
9. Что характеризует индекс структурных сдвигов? Как он рассчитывается?
10. Что представляет собой индекс постоянного состава? Для чего он предназначен?
11. Дайте характеристику индекса переменного состава.
12. В чем проявляется связь индексов переменного состава, постоянного состава и структурного сдвига?

Литература

1. Годин А.М. Статистика: учебник, М.: Дашков и К, 2014
2. Назаров М.Г., Минашкин В.Г., Мхитарян В.С. Статистика: учебник для ВУЗов. - Омега-Л, 2011.
3. Ефимова М.Р., Петрова Е.В., Румянцев В.Н. Общая теория статистики: учебник - ИНФРА-М, 2013
4. Ефимова М.Р. Практикум по общей теории статистики: учебное пособие - М. Финансы и статистика, 2014
5. Статистика: учебник для теория и практика в Excel - М. Финансы и статистика.2010
6. Минашкин В.Г. Статистика: учебник для бакалавров - М. Юрайт, 2013
7. Вуколов Э.А. Практикум по статистическим методам и исследованию операций с использованием пакетов STATISTICA и Excel - Форум-Инфра-М, 2014
8. Просветов Г.И. Статистика: задачи и решения. Учебно-методическое пособие - М. Альфа-Пресс, 2014
9. Сизова Т.М. Статистика: учебное пособие - СПб. НИУ ИТМО, 2013
10. Электронно-библиотечная система. Издательство «Лань» [Электронный ресурс] Годин А.М. Статистика: учебник, М.: Дашков и К, 2014
http://e.lanbook.com.academicnt.ru/books/element.php?pl1_id=56301
11. Электронно-библиотечная система. Издательство «Лань» [Электронный ресурс] Елисеева И.И. Практикум по общей теории статистики - М. Финансы и статистика, 2008.
http://e.lanbook.com.academicnt.ru/books/element.php?pl1_id=53865
12. Электронно-библиотечная система. Издательство «Лань» [Электронный ресурс] Лялин И.С. Статистика: теория и практика в Exel: учебное пособие - М. Финансы и статистика, 2010
http://e.lanbook.com.academicnt.ru/books/element.php?pl1_id=1048

Таблица значений функции $f(t) = \frac{1}{\sqrt{2\pi}} e^{-\frac{t^2}{2}}$

t	0	1	2	3	4	5	6	7	8	9
0,0	3989	3989	3989	3988	3986	3984	3982	3980	3977	3973
0,1	3970	3965	3961	3956	3951	3945	3939	3932	3925	3918
0,2	3910	3902	3894	3885	3876	3867	3857	3847	3836	3825
0,3	3814	3802	3790	3778	3765	3752	3739	3725	3712	3697
0,4	3683	3668	3653	3637	3621	3605	3589	3572	3555	3538
0,5	3521	3503	3485	3467	3448	3429	3410	3391	3372	3352
0,6	3332	3312	3292	3271	3251	3230	3209	3187	3166	3144
0,7	3123	3101	3079	3056	3034	3011	2989	2966	2943	2920
0,8	2897	2874	2850	2827	2803	2780	2756	2732	2709	2685
0,9	2661	2637	2613	2589	2565	2541	2516	2492	2468	2444
1,0	2420	2396	2371	2347	2323	2299	2275	2251	2227	2203
1,1	2179	2155	2131	2107	2083	2059	2036	2012	1989	1965
1,2	1942	1919	1895	1872	1849	1826	1804	1781	1758	1736
1,3	1714	1691	1669	1647	1626	1604	1582	1561	1539	1518
1,4	1497	1476	1456	1435	1415	1394	1374	1354	1334	1315
1,5	1295	1276	1257	1238	1219	1200	1182	1163	1145	1127
1,6	1109	1092	1074	1057	1040	1023	1006	0989	0973	0957
1,7	0940	0925	0909	0893	0878	0863	0848	0833	0818	0804
1,8	0790	0775	0761	0748	0734	0721	0707	0694	0681	0669
1,9	0656	0644	0632	0620	0608	0596	0584	0573	0562	0551
2,0	0540	0529	0519	0508	0498	0488	0478	0468	0459	0449
2,1	0440	0431	0422	0413	0404	0396	0387	0379	0371	0363
2,2	0855	0347	0339	0332	0325	0317	0310	0303	0297	0290
2,3	0283	0277	0270	0264	0258	0252	0246	0241	0235	0229
2,4	0224	0219	0213	0203	0203	0198	0194	0189	0184	0180
2,5	0175	0171	0167	0163	0158	0154	0151	0147	0143	0139
2,6	0136	0132	0129	0126	0122	0119	0116	0113	0110	0107

ЗНАЧЕНИЯ χ^2 - КРИТЕРИЯ ПИРСОНА

Число степеней свободы $K=(M_1-1) \cdot (M_2-1)$	Уровень значимости			Число степене й свободы	Уровень значимости		
	0,10	0,05	0,01		0,10	0,05	0,01
1	2	3	4	1	2	3	4
1	2,71	3,84	6,63	21	29,62	32,67	38,93
2	4,61	5,99	9,21	22	30,81	33,92	40,29
3	6,25	7,81	11,34	23	32,01	35,17	41,64
4	7,78	9,49	13,28	24	33,20	36,42	42,98
5	9,24	11,07	15,09	25	34,38	37,65	44,31
6	10,64	12,59	16,81	26	35,56	38,89	45,64
7	12,02	14,07	18,48	27	36,74	40,11	46,96
8	13,36	15,51	20,09	28	37,92	41,34	48,28
9	14,68	16,92	21,67	29	39,09	42,56	49,59
10	15,99	18,31	23,21	30	40,26	43,77	50,89
11	17,28	19,68	24,72	40	51,80	55,76	63,69
12	18,55	21,03	26,22	50	63,17	67,50	76,15
13	19,81	22,36	27,69	60	74,40	79,08	88,38
14	21,06	23,68	29,14	70	85,53	90,53	100,42
15	22,31	25,00	30,58	80	96,58	101,88	112,33
16	23,54	26,30	32,00	90	107,56	113,14	124,12
17	24,77	27,59	33,41	100	118,50	124,34	135,81
18	25,99	28,87	34,81				
19	27,20	30,14	36,19				
20	28,41	31,41	37,57				

ЗНАЧЕНИЯ t -КРИТЕРИЯ СТЬЮДЕНТА

Число СТЕПЕНЕЙ СВОБОДЫ N-1	УРОВЕНЬ ЗНАЧИМОСТИ			Число СТЕПЕНЕЙ СВОБОДЫ	УРОВЕНЬ ЗНАЧИМОСТИ		
	0,10	0,05	0,01		0,10	0,05	0,01
1	2	3	4	1	2	3	4
1	6,314	12,706	63,657	18	1,734	2,101	2,878
2	2,920	4,303	9,925	19	1,729	2,093	2,861
3	2,353	3,182	5,841	20	1,725	2,086	2,845
4	2,132	2,776	4,604	21	1,721	2,080	2,831
5	2,015	2,571	4,032	22	1,717	2,074	2,819
6	1,943	2,447	3,707	23	1,714	2,069	2,807
7	1,895	2,365	3,499	24	1,711	2,064	2,797
8	1,859	2,306	3,355	25	1,708	2,059	2,787
9	1,833	2,262	3,249	26	1,706	2,055	2,779
10	1,812	2,228	3,169	27	1,703	2,052	2,771
11	1,796	2,201	3,106	28	1,701	2,048	2,763
12	1,782	2,179	3,054	29	1,699	2,045	2,756
13	1,771	2,160	3,012	30	1,697	2,042	2,750
14	1,761	2,145	2,977	40	1,684	2,02	2,704
15	1,753	2,131	2,947	60	1,671	2,000	2,660
16	1,746	2,120	2,921	120	1,658	1,980	2,617
17	1,740	2,110	2,898				

Значения F-критерия Фишера при уровне значимости 0,05

$v_1 \backslash v_2$	1	2	3	$v_1 \backslash v_2$	1	2	3
1	161,00	200,00	216,00	18	4,41	3,55	3,16
2	18,51	19,00	19,16	19	4,38	3,52	3,13
3	10,13	9,55	9,28	20	4,35	3,49	3,10
4	7,71	6,94	6,59	21	4,32	3,47	3,07
5	6,61	5,79	5,41	22	4,30	3,44	3,05
6	5,99	5,14	4,76	23	4,28	3,42	3,03
7	5,59	4,74	4,35	24	4,26	3,40	3,01
8	5,32	4,46	4,07	25	4,24	3,88	2,99
9	5,12	4,26	3,86	26	4,22	3,37	2,98
10	4,96	4,10	3,71	27	4,21	3,35	2,96
11	4,84	3,98	3,59	28	4,20	3,34	2,95
12	4,75	3,88	3,49	29	4,18	3,33	2,93
13	4,67	3,80	3,41	30	4,17	3,32	2,92
14	4,60	3,74	3,34	40	4,08	3,23	2,84
15	4,54	3,68	3,29	50	4,03	3,18	2,79
16	4,49	3,63	3,24	60	4,00	3,15	2,76
17	4,45	3,59	3,20	100	3,94	3,09	2,70

$V_1=m-1$; $V_2=n-m$; n -число наблюдений; m -число признаков.

Распределение вероятности в малых выборках в зависимости от коэффициента доверия t и объема выборки n

$t \backslash N$	4	5	6	7	8	9	10	15	20	∞
0,5	348	356	362	366	368	370	372	376	378	383
1,0	608	626	636	644	650	654	656	666	670	683
1,5	770	792	806	816	832	828	832	846	850	865
2,0	860	884	908	908	914	920	924	936	940	954
2,5	933	946	955	959	963	966	968	975	978	988
3,0	942	960	970	970	980	938	984	992	992	997

При $n=\infty$ в таблице даны вероятности нормального распределения. Для определения вероятности соответствующие табличные значения следует разделить на 1000.

Миссия университета – генерация передовых знаний, внедрение инновационных разработок и подготовка элитных кадров, способных действовать в условиях быстро меняющегося мира и обеспечивать опережающее развитие науки, технологий и других областей для содействия решению актуальных задач.

КАФЕДРА ФИНАНСОВОГО МЕНЕДЖМЕНТА И АУДИТА

Кафедра финансового менеджмента и аудита (ФМиА) создана в 2015 году на базе трех кафедр: экономики и финансов, экономики и предпринимательской деятельности, финансового менеджмента. Заведующий кафедрой ФМиА – профессор, доктор экономических наук Сергеева Ирина Григорьевна. Кафедра является выпускающей кафедрой факультета технологического менеджмента и инноваций. Кафедра проводит обучение бакалавров по направлению 38.03.02 «Менеджмент», программы «Финансовый менеджмент», «Логистика», «Управление малым бизнесом». Кафедрой ФМиА осуществляется подготовка магистров по направлению 38.04.02 «Менеджмент», магистерские программы «Финансовый менеджмент», «Ресурсный менеджмент в инновационной деятельности», «Антикризисное управление и аудит»; по направлению 27.04.05 «Инноватика», магистерская программа «Экономика и управление инновационной деятельностью в областях науки»; по направлению 27.04.02 «Управление качеством», магистерская программа «Аудит качества и стандартизация».

Тамара Митрофановна Сизова, Людмила Геннадьевна Мишура

Статистика: практикум

Учебное пособие

В авторской редакции

Редакционно-издательский отдел Университета ИТМО

Зав. РИО

Н.Ф. Гусарова

Подписано к печати

Заказ №

Тираж 100 экз.

Отпечатано на ризографе

Редакционно-издательский отдел
Университета ИТМО
197101, Санкт-Петербург, Кронверкский пр., 49