М.Р. Ефимова
О.И. Ганченко
Е.В. Петрова

ПРАКТИКУМ

по общей теории статистики

М.Р. Ефимова, О.И. Ганченко, Е.В. Петрова

ПРАКТИКУМ

по общей теории статистики

Издание второе, переработанное и дополненное

Рекомендовано
Советом Учебно-методического объединения вузов
России по образованию в области менеджмента
в качестве учебного пособия по специальностям
"Менеджмент организации", "Государственное
и муниципальное управление", "Маркетинг",
"Управление персоналом"

Москва
"Финансы и статистика"
2004

УДК 311(076.5)
ББК 60.6я73
Е91

РЕЦЕНЗЕНТЫ:

Кафедра статистики
Российского государственного торгово-экономического университета;
Р. А. Шмойлова,
кандидат экономических наук, профессор
кафедры теории статистики и прогнозирования
Московского государственного университета
экономики, статистики и информатики

Ефимова М. Р., Ганченко О. И., Петрова Е. В.
Е91 Практикум по общей теории статистики: Учеб. пособие. – 2-е изд., перераб. и доп. – М.: Финансы и статистика, 2004. – 336 с.: ил.

ISBN 5-279-02555-0

Содержит краткий обзор основных понятий общей теории статистики. Рассматриваются группировка статистических данных и их роль в анализе информации, абсолютные, относительные и средние величины, статистические распределения, выборочное наблюдение, корреляционная связь, ряды динамики, индексы и их использование в экономико-статистических исследованиях. В отличие от первого издания (1999 г.) значительно переработана глава, излагающая абсолютные, относительные и средние величины; она дополнена материалом по их графическому изображению. Представлены типовые примеры с решениями и задачи с ответами. В приложениях даны математико-статистические таблицы.

Для преподавателей, аспирантов, студентов экономических вузов, экономистов, менеджеров, руководителей предприятий.

Е $\frac{0702000000 - 150}{010 (01) - 2004}$ 287 – 2004

УДК 311(076.5)
ББК 60.6я73

ISBN 5-279-02555-0

© Ефимова М. Р., Ганченко О. И., Петрова Е. В., 2002

Предисловие

В современном обществе статистика стала одним из важнейших инструментов управления национальной экономикой. Развитие рыночных отношений в стране поставило перед статистикой новую задачу – реформирование общеметодологических и организационных основ статистической теории и практики.

Улучшение хозяйственного руководства неразрывно связано с возрастанием роли статистики и повышением научного уровня статистических исследований.

Главной задачей статистики является исчисление и анализ статистических показателей, благодаря чему органы управления получают всестороннюю характеристику управляемых объектов, будь то вся национальная экономика или отдельные ее отрасли, предприятия и их подразделения.

Общая теория статистики разрабатывает общие принципы и методы статистического исследования общественных явлений, наиболее общие категории (показатели) статистики. Она является учебной дисциплиной, формирующей необходимые профессиональные знания у экономистов, менеджеров, руководителей предприятий.

Данный практикум составлен в соответствии с учебной программой курса и предназначен для студентов экономических специальностей высших учебных заведений, а также для практических работников, занимающихся вопросами планирования, учета и анализа производственно-хозяйственной деятельности. Практикумом могут воспользоваться лица, самостоятельно изучающие статистику.

Практикум состоит из семи глав, соответствующих темам курса общей теории статистики. Каждая глава построена так, что сначала излагается теория по изучаемому материалу, приводятся сжатые теоретические указания о методах расчета и анализа показателей. Затем рассматриваются типовые примеры с решениями и приводятся задачи для самостоятельной работы (большинство задач имеет ответы), которые позволяют обучаемому наиболее полно и глубоко изучить материал.

В приложениях учебного пособия представлены математико-статистические таблицы.

Второе издание практикума значительно переработано. В отличие от первого издания (1999 г.) расширена глава, изучающая абсолютные, относительные и средние величины. Она дополнена новым материалом по их графическому изображению.

При составлении задач были использованы данные статистических сборников и отчеты предприятий. Отчетные данные предприятий методически обработаны и не могут служить справочным материалом.

ГЛАВА 1

Группировка статистических данных и ее роль в анализе информации

Одним из основных наиболее распространенных методов обработки и анализа первичной статистической информации является группировка.

Под **группировкой** понимают расчленение единиц статистической совокупности на группы, однородные в каком-либо существенном отношении, и характеристику таких групп системой показателей в целях выделения типов явлений, изучения структуры и взаимосвязей. Следовательно, **с помощью группировок решаются три задачи:**

• разделение всей совокупности на качественно однородные группы – выделение социально-экономических типов. Эти группировки называются **типологическими** (например, группировки хозяйствующих субъектов по формам собственности, населения по общественным группам и др.);

• характеристика структуры явления и структурных сдвигов. Эти группировки называются **структурными** (например, определение значения каждого вида транспорта в транспортном балансе страны, изучение состава населения по полу, возрасту и другим признакам и т. д.);

• изучение взаимосвязей между отдельными признаками изучаемого явления. Такие группировки называются **аналитическими** (например, группировка предприятий определенной отрасли экономики по уровню производительности труда для выявления ее влияния на себестоимость продукции).

Разграничение трех видов группировки является в известной мере условным. Во многих случаях одна и та же группировка дает возможность решать все три задачи.

Группировка является аналитико-синтетическим процессом. Выделенные при группировке однородные части, отличающиеся друг от друга качеством или условиями своего развития, детально изучаются. После этого решается синтетическая задача – отражение процесса в целом, т. е. характеристика соотношения между выделенными группами.

Признак, на основе которого производится подразделение единиц наблюдения на группы, называется **группировочным признаком или основанием группировки.** Группировка может выполняться по одному признаку (*простая группировка*) и по нескольким признакам (*комбинированная группировка*).

Выбор группировочных признаков всегда должен быть основан на анализе качественной природы изучаемого явления. Всесторонний теоретико-экономический анализ сущности и закономерностей развития явления должен быть направлен на то, чтобы в соответствии с целью исследования положить в основание группировки существенные признаки.

Группировочные признаки могут быть атрибутивными и количественными. **Атрибутивные признаки** регистрируются в виде текстовой записи (например, профессия рабочих, социальная группа населения). **Количественные признаки** имеют цифровое выражение (стаж работы, размер дохода).

При **группировке по атрибутивному признаку** число групп определяется количеством соответствующих наименований, если число этих наименований не очень велико. Если признак имеет большое количество разновидностей, то при группировке ряд наименований объединяют в одну группу. Для обоснованного объединения их в группы разрабатываются классификации. В отличие от группировок при классификации группировочные признаки установлены заранее на длительный период для решения многих задач, в то время как группировки выполняются для целей конкретного исследования. Примерами могут служить классификации отраслей экономики, автотранспортных предприятий по целевому назначению (грузовые, автобусные, таксомоторные и др.).

При **группировке по количественному признаку** число групп определяется в зависимости от характера изменения признака и задач исследования. Если количественный признак меняется *прерывно* (*дискретно*), т. е. может принимать только некоторые – чаще целые значения (например, тарифный разряд рабочих), то число групп должно соответствовать количеству значений признака.

При *непрерывном* изменении признак принимает любые значения (например, стаж работы или возраст рабочих), поэтому группы ограничиваются значениями признака в интервале «от – до». **Интервалом** называется разница между максимальным и минимальным значениями признака в каждой группе. На практике используются три вида интервалов: равные, неравные (постепенно увеличивающиеся) и специализированные.

Равные интервалы используются, если нужно охарактеризовать количественные различия в величине признака внутри групп одинакового качества (например, при группировке рабочих определенной профессии по проценту выполнения норм выработки).

Величина равного интервала исчисляется по формуле

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

где $x_{\max}, x_{\min}$ – соответственно наибольшее и наименьшее значения признака в изучаемой совокупности;
m – принятое число групп.

Для расчета величины интервала по этой формуле необходимо заранее установить число групп (при числе наблюдений более 200 используют 10 – 15 групп).

Возможен и другой способ определения величины интервала, не требующий предварительного установления числа групп. В этом случае используется **формула Стерджесса:**

$$i = \frac{x_{\max} - x_{\min}}{1{,}000 + 3{,}322 \lg n},$$

где n – число наблюдений.

Выполняя расчет величины интервала по этой формуле, следует знаменатель предварительно округлить до целого бóльшего числа, поскольку количество групп не может быть дробным.

Величину интервала обычно округляют до целого (всегда большего) числа, исключение составляют случаи, когда изучаются малейшие колебания признака.

Неравные интервалы (постепенно увеличивающиеся) часто применяются в аналитических группировках. В этом случае интервалы выбираются так, чтобы число единиц в образованных группах было достаточно велико (т. е. чтобы группы были приблизительно одинаково заполнены).

Специализированные интервалы используются в типологических группировках; границы устанавливаются там, где намечается переход от одного качества к другому. Наметить точки перехода можно только на основе теоретического анализа, используя для выделения типов не отдельные, изолированные признаки, а совокупность признаков, характеризующих различные стороны изучаемого явления.

Интервалы группировки могут быть закрытыми и открытыми. **Закрытые интервалы** – это обычные интервалы, имеющие как нижние (т. е. «от»), так и верхние (т. е. «до») границы. **Открытые интервалы** – это интервалы, имеющие какую-либо одну границу – верхнюю или нижнюю. Они применяются тогда, когда признак изменяется неравномерно в широких пределах, причем большие (или малые) значения признака встречаются нечасто.

Иногда имеющуюся группировку необходимо несколько изменить: объединить ранее выделенные относительно мелкие группы в небольшое число более крупных, типичных групп или изменить границы прежних групп, с тем чтобы сделать группировку сопоставимой с другими. Такая переработка результатов первичной группировки называется **перегруппировкой** или **вторичной группировкой.**

Следующей за группировкой ступенью систематизации и обобщения материалов статистического наблюдения является статистическая сводка. Под **статистической сводкой** в узком смысле слова понимается подсчет числа единиц в подгруппах и группах, выделенных при группировке, и подведение итогов по количественным признакам.

Результаты группировки и сводки материалов оформляются в виде **статистических таблиц.**

В статистической таблице выделяются два элемента:

• **подлежащее** (обычно помещается в первой вертикальной или в горизонтальной графе) – перечень единиц или групп, на которые подразделена вся масса единиц наблюдения.

• **сказуемое** – цифры, при помощи которых характеризуются выделенные в подлежащем единицы или группы.

Над таблицей помещается заголовок, отражающий в сжатой форме ее основное содержание, время и место, к которым относятся изложенные в таблице данные.

Данные статистических таблиц используются для целей оперативного руководства, научного анализа, позволяющего вскрыть

взаимосвязи и имеющиеся резервы. Различие целей сказывается на характере подлежащего.

В зависимости от характера подлежащего различают три вида таблиц: простые, групповые, комбинационные.

В подлежащем **простых таблиц** дается перечень единиц или групп, составляющих объект изучения (предприятия, районы и др.), однако части подлежащего не являются группами одинакового качества. В сказуемом этих таблиц основное значение имеют абсолютные величины, выражающие объемы изучаемых общественных явлений. Простые таблицы дают справочный материал; они, как правило, отражают наличие и распределение ресурсов в стране и регионах.

Для целей научного анализа используются групповые и комбинационные таблицы.

Групповой таблицей называется таблица, подлежащее которой образовано в результате группировки единиц по одному какому-то признаку. Если в сказуемом групповой таблицы только одна графа, характеризующая численность группы (частота), то такая таблица называется **рядом распределения.**

В комбинационной таблице подлежащее образовано в результате группировки единиц совокупности по двум и более признакам. В этом случае все единицы распределяются на группы сначала по одному признаку, а затем внутри каждой из выделенных групп – на подгруппы по другому признаку.

В сказуемом групповых и комбинационных таблиц на основе абсолютных величин исчисляют средние и относительные величины, позволяющие раскрыть особенности и закономерности развития изучаемого явления.

1.1. Решение типовых задач

1.1. Объем инвестиций в основной капитал характеризуется в России следующими данными (в фактически действовавших ценах, млрд руб.): 1998 г. – 402,4; 1999 г. – 565,6; в том числе в отрасли: а) производящие товары – 1998 г. – 163,8; 1999 г. – 269,4; б) оказывающие рыночные и нерыночные услуги – 1998 г. – 238,6; 1999 г. – 296,2.

Представить приведенные данные в виде статистической таблицы.

Сформулировать выводы, охарактеризовав произошедшие изменения в объеме и составе инвестиций.

Решение

В табл. 1.1 представлены данные об инвестициях в России в основной капитал.

Таблица 1.1

Инвестиции в России в основной капитал (в фактически действовавших ценах, млрд руб.)

Показатели	1998 г.	1999 г.	Изменение объема инвестиций в 1999 г. по сравнению с 1998 г., %	Структура инвестиций, %	
				1998 г.	1999 г.
Инвестиции в основной капитал	402,4	565,6	140,6	100,0	100,0
в том числе в отрасли:					
производящие товары	163,8	269,4	164,5	40,7	47,6
оказывающие рыночные и нерыночные услуги	238,6	296,2	124,1	59,3	52,4

Изменение объема инвестиций в 1999 г. по сравнению с 1998 г. определялось следующим образом, %:

все инвестиции: $\frac{565,6}{402,4} \cdot 100 = 140,6;$

инвестиции в отрасли, производящие товары:

$\frac{269,4}{163,8} \cdot 100 = 164,5;$

инвестиции в отрасли, оказывающие рыночные и нерыночные услуги: $\frac{269,2}{238,6} \cdot 100 = 124,1.$

Структура инвестиций определялась так, %:

$$1998 \text{ г.}: \frac{163,8}{402,4} \cdot 100 = 40,7; \quad \frac{238,6}{402,4} \cdot 100 = 59,3;$$

$$1999 \text{ г.}: \frac{269,4}{565,6} \cdot 100 = 47,6; \quad \frac{296,2}{565,6} \cdot 100 = 52,4.$$

Как видно из данных построенной групповой таблицы, в 1999 г. произошли существенные изменения объема инвестиций: общий объем инвестиций увеличился на 40,6%, инвестиции в отрасли, производящие товары, возросли на 64,5%, инвестиции в отрасли, оказывающие услуги, выросли на 24,1%. Различный темп роста инвестиций по отраслям повлек изменения в структуре инвестиций. Так, удельный вес инвестиций в отрасли, производящие товары, в 1998 г. составил 40,7%, а в 1999 г. он увеличился до 47,6%, соответственно произошло снижение удельного веса инвестиций в отрасли, оказывающие услуги.

1.2. Имеются следующие данные по заработной плате водителей за сентябрь (табл. 1.2).

Таблица 1.2

Табельный номер водителя	Класс водителя	Процент выполнения сменных заданий	Заработная плата за месяц, руб.
1	I	110,2	4100,3
2	II	102,0	3600,8
3	II	111,0	3970,7
4	I	107,9	4050,2
5	II	106,4	3740,5
6	I	109,0	3985,4
7	I	115,0	4300,8
8	II	112,2	4015,7
9	I	105,0	3790,2
10	II	107,4	3700,7
11	I	112,5	4280,2
12	I	108,6	4170,1

Требуется для выявления зависимости заработной платы водителей от уровня квалификации и процента выполнения сменных заданий произвести аналитическую группировку. Интервалы группировки водителей по проценту выполнения норм выработки разработать самостоятельно. На основе выполненной группировки построить комбинационную таблицу.

Сформулировать вывод.

Решение

Для решения задачи необходимо произвести группировку водителей по двум признакам-факторам: сначала – на группы по квалификации, затем внутри каждой группы – на подгруппы по проценту выполнения сменного задания.

По проценту выполнения сменного задания принимаются две подгруппы: водители, выполняющие норму от 100 до 110%; водители, выполняющие норму на 110% и выше.

Результаты группировки представлены во вспомогательной табл. 1.3.

Таблица 1.3

Вспомогательная таблица

Группы водителей по уровню квалификации	Водители II класса		Водители I класса	
Подгруппы водителей по проценту выполнения сменного задания	100–110	110 и выше	100–110	110 и выше
Табельный номер водителя	2; 5; 10	3; 8	4; 6; 9; 12	1; 7; 11
Заработная плата за месяц, руб.	3600,8 3740,5 3700,7	3970,7 4015,7	4050,2 3985,4 3790,2 4170,1	4100,3 4300,8 4280,2

На основе вспомогательной таблицы по каждой подгруппе определяют численность и итог признака (общую сумму заработной платы), результаты оформляют в виде комбинационной таблицы (табл. 1.4).

Из данных табл. 1.4 следует, что с ростом квалификации водителей и процента выполнения сменного задания увеличивается заработная плата. Так, заработная плата водителей I класса, выполняющих норму выработки на 110% и выше, на 14,8% превышает заработную плату водителей II класса, выполняющих норму от 100 до 110%.

Таблица 1.4

Зависимость заработной платы водителей от квалификации и процента выполнения задания

Группы водителей по уровню квалификации	Подгруппы водителей по проценту выполнения сменного задания	Число водителей	Общая сумма заработной платы, руб.	Средняя заработная плата одного водителя, руб.	Изменение средней заработной платы по сравнению с низшей подгруппой, %
II класс	100–110 110 и выше	3 2	11042,0 7986,4	3680,7 3993,2	100,0 108,5
Итого по группе		5	19028,4	3805,7	–
I класс	100–110 110 и выше	4 3	15995,9 12681,3	3999,0 4227,1	108,6 114,8
Итого по группе		7	28677,2	4096,7	–
Всего		12	47705,6	3975,5	–

1.2. Задачи для самостоятельной работы

1.3. Выпуск продукции по предприятию следующий (млн руб.): 1999 г. – 123,0; 2000 г. – 187,5; 2001 г. – 210,0. Из общего объема продукции было предназначено на экспорт (млн руб.): 1999 г. – 50,8; 2000 г. – 92,7; 2001 г. – 122,8.

Представить приведенные данные в виде статистической таблицы; указать тип таблицы.

1.4. Перевозка грузов автотранспортным предприятием характеризуется следующими данными (тыс. т): 1999 г. – 2238,9; 2000 г. – 2175,8; 2001 г. – 2485,5, в том числе по договорной клиентуре – соответственно 1308,0; 1025,5; 1390,7.

Представить приведенные данные в виде статистической таблицы.

1.5. Имеются данные о численности и составе населения России (на начало года, млн чел.).

Все население: 1997 г. – 147,1; 1998 г. – 146,7; 1999 г. – 146,3; 2000 г. – 145,6, в том числе мужчины составили: 1997 г. – 69,9; 1998 г. – 68,8; 1999 г. – 68,6; 2000 г. – 68,2.

Построить статистическую таблицу, характеризующую динамику численности и состава населения России.

1.6. *Построить* макеты статистических таблиц, характеризующих за период 1999 – 2001 гг. динамику следующих показателей:

а) объем выпуска продукции (млн руб.) предприятиями добывающей и обрабатывающей промышленности России;

б) выработку электроэнергии (тыс. кВт · ч) электростанциями различных типов;

в) объем перевозок (тыс. т) и объем выполненной транспортной работы (грузооборот, млн ткм) по предприятиям региона различной организационно-правовой формы (государственные, арендные, акционерные).

Разработать для макета каждой таблицы подлежащее и сказуемое.

Определить, к какому виду таблиц относится построенный макет.

1.7. Имеются следующие данные о численности и составе населения России (на начало года, млн чел.).

Все население: 1997 г. – 147,1; 1998 г. – 146,7; 1999 г. – 146,3; 2000 г. – 145,6; в том числе городское население составило: 1997 г. – 107,3; 1998 г. – 107,1; 1999 г. – 106,8; 2000 г. – 106,1.

Построить статистическую таблицу, характеризующую динамику численности и состава населения России.

1.8. *Построить* макет статистической таблицы, характеризующий изменение численности работников предприятия по категориям (рабочие, служащие) и их средней заработной платы по кварталам отчетного года.

К какому виду таблиц может быть отнесен построенный макет?

1.9. Имеются следующие данные о численности постоянного населения России (на начало года, млн чел.).

Все постоянное население: 1998 г. – 146,7; 1999 г. – 146,3; 2000 г. – 145,6. Из общей численности постоянного населения численность населения моложе трудоспособного возраста составила: 1998 г. – 31,3; 1999 г. – 30,3; 2000 г. – 29,1. Численность трудоспособного населения следующая: 1998 г. – 84,8; 1999 г. – 85,6; 2000 г. – 86,3. К остальному населению относится население старше трудоспособного возраста.

Построить статистическую таблицу, характеризующую динамику численности постоянного населения России и его возрастной состав.

1.10. Имеются данные о заработной плате за месяц рабочих бригады (табл. 1.5).

Таблица 1.5

Табельный номер рабочего	1	2	3	4	5	6	7	8
Процент выполнения норм выработки	110,8	102,0	111,0	107,8	106,4	109,0	100,0	105,0
Заработная плата за месяц, руб.	3910	3600	4100	4800	3850	3980	3400	3700

Требуется для выявления зависимости заработной платы рабочих от процента выполнения норм выработки произвести аналитическую группировку рабочих бригады по проценту выполнения норм выработки, выделив три группы: а) рабочие, выполняющие норму до 105,0%; б) рабочие, выполняющие норму от 105 до 110%; в) рабочие, выполняющие норму на 110% и более.

На основе выполненной группировки построить групповую таблицу.

Сформулировать вывод.

1.11. По группе грузовых автотранспортных предприятий города имеется следующая информация за отчетный год (табл. 1.6).

Таблица 1.6

№ предприятия	Грузооборот, млн ткм	Сумма затрат на перевозки, тыс. руб.	№ предприятия	Грузооборот, млн ткм	Сумма затрат на перевозки, тыс. руб.
1	62	29140	9	47	26790
2	40	22040	10	24	14160
3	38	21660	11	18	11700
4	25	14625	12	58	27750
5	15	9900	13	44	22000
6	30	17100	14	23	13317
7	52	25272	15	32	17280
8	27	30800	16	20	12000

Требуется:

1) произвести группировку грузовых автотранспортных предприятий по размеру грузооборота, выделив следующие группы: до 20 млн ткм; 20 – 40; 40 млн ткм и более;

2) по каждой группе определить: число предприятий, общий объем грузооборота, общую сумму затрат на перевозки, среднюю величину затрат на 10 ткм;

3) представить решение в форме статистической таблицы.

Сформулировать вывод.

1.12. По годовым отчетам промышленных предприятий района получена следующая информация (табл. 1.7).

Таблица 1.7

№ предприятия	Объем продукции, млн руб.	Среднегодовая стоимость основных фондов, млн руб.	Среднесписочное число работников, чел.	№ предприятия	Объем продукции, млн руб.	Среднегодовая стоимость основных фондов, млн руб.	Среднесписочное число работников, чел.
1	402	7,2	700	11	1756	21,0	1425
2	792	11,6	1100	12	1014	14,0	1208
3	1116	15,6	1285	13	1440	19,0	1400
4	435	7,6	705	14	720	11,0	900
5	1281	16,0	1300	15	1086	14,8	1300
6	1756	22,0	1450	16	1809	23,0	1480
7	510	8,4	800	17	1125	15,6	1295
8	1392	18,8	1380	18	648	10,0	895
9	540	9,2	825	19	1716	19,8	1440
10	924	13,2	1210	20	881	12,4	1180

Требуется:

1) выполнить группировку промышленных предприятий по стоимости основных фондов, положив в основание группировки стоимость основных фондов: до 10 млн руб.; 10 – 15 млн руб.; 15 – 20 млн руб.; 20 млн руб. и выше;

2) определить по каждой группе число предприятий, объем продукции, среднесписочное число работников, объем продукции в расчете на 1 тыс. руб. стоимости основных фондов;

3) оформить результаты в виде статистической таблицы.

Сформулировать вывод.

1.13. По отдельным бригадам строительной организации имеются следующие данные за август (табл. 1.8).

Таблица 1.8

Показатель	№ бригады							
	1	2	3	4	5	6	7	8
Объем работ, тыс. руб.	819	1296	1340	1008	1468	1772	720	1904
Численность рабочих, чел.	16	24	25	21	27	32	15	34

Требуется:

1) для выявления зависимости объема работ от числа рабочих, занятых в строительных бригадах, произвести группировку бригад по численности рабочих, выделив три группы с равными интервалами;

2) на основе выполненной группировки построить групповую таблицу.

Сформулировать вывод.

1.14. Объем инвестиций, поступивших в Россию от иностранных инвесторов в 1999 г., следующий (млн долл. США): всего инвестиций 9560, в том числе инвестиций в промышленность – 4876, строительство – 97, транспорт – 521, связь –386, торговлю и общественное питание – 1622, общую коммерческую деятельность по обеспечению функционирования рынка – 190, финансы, кредит, страхование, пенсионное обеспечение – 114, прочие отрасли – 1754.

Построить статистическую таблицу, характеризующую структуру иностранных инвестиций, поступивших в Россию за 1999 г.

1.15. По 12 партиям деталей, обрабатываемых рабочими производственного участка, имеются следующие данные (табл. 1.9).

Таблица 1.9

№ партии	Число операций, выполняемых при обработке одной детали	Число деталей в партии	Время на обработку всей партии, ч.
1	2	12	3,86
2	3	16	1,97
3	3	4	1,83
4	4	12	8,10
5	5	20	4,40
6	5	8	4,70
7	6	12	5,90
8	8	4	5,38
9	11	4	3,80
10	12	4	4,40
11	11	3	3,75
12	9	1	1,45

Требуется:

1) выполнить аналитическую группировку, с тем чтобы выявить, наблюдается ли в условиях работы данного участка связь между количеством операций по обработке одной детали и временем ее обработки;

2) результаты группировки представить в форме групповой таблицы.

Сформулировать вывод.

1.16. Имеются следующие данные по группе промышленных предприятий за отчетный год (табл. 1.10).

Требуется:

1) выполнить группировку предприятий по объему продукции, приняв следующие интервалы: а) до 600 млн руб.; б) от 600 до 1200 млн руб.; в) 1200 млн руб. и более;

2) по каждой группе и в целом по всем предприятиям определить: число предприятий, объем продукции, среднесписочное число работников, среднюю выработку продукции на одного работника;

3) представить результаты группировки в виде статистической таблицы.

Сформулировать вывод.

1.17. По данным задачи 1.16 произвести группировку предприятий по стоимости основных фондов, приняв следующие интервалы: стоимость основных фондов: а) до 12,0 млн руб.; б) от 12,0 до 18,0 млн руб.; в) от 18,0 млн руб. и выше.

Таблица 1.10

№ предприятия	Объем продукции, млн руб.	Среднегодовая стоимость основных фондов, млн руб.	Среднесписочное число работников, чел.	Прибыль, млн руб.
1	591	10,0	900	27
2	1776	22,8	1500	272
3	1395	18,4	1412	194
4	888	12,6	1200	88
5	1752	22,0	1485	292
6	1440	19,0	1420	220
7	1734	21,6	1390	276
8	612	9,4	817	60
9	1398	19,4	1375	224
10	876	13,6	1200	100
11	1269	17,6	1365	110
12	576	8,8	850	61
13	1080	14,0	1290	128
14	624	10,2	900	67

По каждой группе и в целом по всем предприятиям *определить*: число предприятий, среднегодовую стоимость основных фондов, объем продукции, сумму прибыли, а также объем продукции в расчете на 1 млн руб. стоимости основных фондов и размер прибыли в расчете на 1 млн руб. стоимости основных фондов. Результаты группировки оформить в виде статистической таблицы.

Сформулировать вывод.

1.18. По данным задачи 1.16 произвести группировку предприятий по численности работников, приняв следующие интервалы: а) до 1000 человек; б) от 1000 до 1300 человек; в) 1300 человек и более.

По каждой группе и в целом по всем предприятиям *определить*: число предприятий, объем продукции, среднесписочное число работников, среднегодовую стоимость основных фондов, а также размер среднегодовой стоимости основных фондов в расчете на одного работника и среднюю выработку продукции на одного работника. Результаты группировки представить в виде статистической таблицы.

Сформулировать вывод.

1.19. По промышленным предприятиям города имеются следующие данные за отчетный год (табл. 1.11).

Таблица 1.11

№ предприятия	Объем продукции, млн руб.	Фонд заработной платы, млн руб.	№ предприятия	Объем продукции, млн руб.	Фонд заработной платы, млн руб.
1	124,8	19,8	9	110,0	17,7
2	256,0	38,4	10	256,3	40,9
3	190,7	31,3	11	187,5	30,7
4	185,0	31,4	12	140,8	23,2
5	403,2	56,4	13	167,3	27,0
6	115,0	19,6	14	208,2	32,2
7	106,5	17,2	15	135,4	21,9
8	350,0	49,7	16	370,2	51,8

Требуется:

1) сгруппировать предприятия по объему выработанной продукции, выделив три группы (интервалы группировки разработать самостоятельно);

2) определить по каждой группе число предприятий, объем продукции, фонд заработной платы, размер заработной платы (тыс. руб.) на 1 млн руб. объема продукции;

3) оформить решение в виде статистической таблицы.

Сформулировать вывод.

1.20. По группе промышленных предприятий, выпускающих одинаковые виды продукции, имеются следующие данные за отчетный год (табл. 1.12).

Для выявления зависимости производительности труда работников, представляющей объем продукции на одного списочного работника, от фондовооруженности (фондовооруженность – стоимость основных производственных фондов, приходящаяся на одного работника) произвести аналитическую группировку предприятий по показателю фондовооруженности труда, выделив три группы предприятий. Интервалы группировки разработать самостоятельно. На основе группировки построить групповую таблицу.

Сформулировать вывод.

Таблица 1.12

Показатель	№ предприятия								
	1	2	3	4	5	6	7	8	9
Общий объем продукции, млн руб.	249,6	391,6	734,4	512,0	369,4	806,4	224,6	213,8	696,0
Среднесписочная численность работников, чел.	520	680	1020	970	855	1200	585	594	1000
Фондовооруженность работников, тыс. руб./чел.	30,8	36,0	41,5	35,5	27,6	40,2	25,5	24,0	40,8

1.21. Заработная плата рабочих бригады за сентябрь следующая (табл. 1.13).

Таблица 1.13

Табельный номер	Тарифный разряд	Процент выполнения норм выработки	Заработная плата за месяц, руб.
1	6	110,2	4820,3
2	5	102,0	3520,0
3	5	111,0	2782,4
4	6	107,9	3800,0
5	5	106,4	3615,2
6	6	109,0	4790,3
7	6	115,0	4830,5
8	5	112,2	3995,0
9	6	105,0	3612,3
10	5	107,4	3570,1
11	6	112,5	4827,4
12	6	108,0	3788,4

На основе приведенных данных произвести аналитическую группировку и построить комбинационную таблицу, характеризующую зависимость размера заработной платы рабочих от уровня их квалификации и процента выполнения норм выработки. По проценту выполнения норм выработки принять две подгруппы: а) рабочие, выполняющие норму выработки до 110,0%; б) рабочие, выполняющие норму на 110% и выше.

Сформулировать вывод.

1.22. Произвести перегруппировку данных об уровне выполнения норм выработки рабочими двух цехов с целью получения сопоставимых показателей и их анализа (табл. 1.14).

Таблица 1.14

Цех № 1		Цех № 2	
группы рабочих по проценту выполнения норм выработки	число рабочих, % к итогу	группы рабочих по проценту выполнения норм выработки	число рабочих, % к итогу
До 90	2,0	До 100	9,0
90 – 100	8,0	100 – 120	40,0
100 – 110	40,0	120 – 150	25,0
110 – 120	25,0	150 – 180	15,0
120 – 150	20,0	180 – 200	7,0
150 и выше	5,0	200 и выше	4,0
Итого	100,0	Итого	100,0

1.23. По цехам предприятия имеются следующие данные за сентябрь (табл. 1.15).

Таблица 1.15

№ цеха	Среднесписочное число работников			Общий фонд заработной платы, тыс. руб.			Среднемесячная заработная плата работника, руб.		
	план	отчет	процент выполнения плана	план	отчет	процент выполнения плана	план	отчет	процент выполнения плана
1	1200		105,0			108,0		3800,0	
2		1078	98,0					4100,0	105,0
3			106,0						
Итого по предприятию	4100				17412,2		4000,0		

Вычислить и *проставить* в таблицу все недостающие данные (точность расчета: численность работников – до целого числа, остальные показатели – с точностью до 0,1).

1.24. *Построить* макет статистической таблицы, отражающей расход различных видов топлива (твердое, жидкое, газообразное) в тыс. т на производство электрической и тепловой энергии в регионе за отчетный год.

1.25. Объем инвестиций, поступивших в Россию от иностранных инвесторов в 1999 г., следующий (млн. долл. США): всего инвестиций 9560, в том числе инвестиций в промышленность – 4876, строительство – 97, транспорт – 521, связь – 386, торговля и общественное питание – 1622, общая коммерческая деятельность по обеспечению функционирования рынка – 190, финансы, кредит, страхование, пенсионное обеспечение – 114, прочие отрасли – 1754.

Построить статистическую таблицу, характеризующую структуру иностранных инвестиций, поступивших в Россию за 1999 г.

ГЛАВА 2

Абсолютные, относительные, средние величины и их графические изображения

Абсолютные величины

Абсолютные величины характеризуют численность совокупности и объем (размер) изучаемого социально-экономического явления в определенных границах времени и места. Они являются всегда именованными числами, т. е. имеют какую-либо единицу измерения. Единицы измерения могут быть натуральные, условно-натуральные, стоимостные (денежные) и трудовые. Выбор единицы измерения зависит от сущности изучаемого явления и конкретных задач исследования.

Абсолютные величины подразделяются на две группы:

• абсолютные величины, характеризующие объем явления на определенную дату (например, стоимость основного капитала предприятия на 1 января);

• абсолютные величины, характеризующие объем явления за определенный период времени – результат процесса (например, выпуск продукции предприятием за месяц или за год).

Абсолютные величины первой группы имеют особенность: если они характеризуют объем явления на определенную дату по нескольким единицам (например, стоимость основного капитала по предприятиям фирмы), то их можно суммировать и получить общий объем явления. Если данные характеризуют объем явления по одной единице на несколько моментов (например, стоимость основного капитала на начало каждого квартала), то эти абсолютные величины суммировать нельзя.

Абсолютные величины второй группы можно суммировать за одинаковые периоды по нескольким единицам, а также по одной единице за несколько периодов, получая итог за более длительный период (например, можно складывать объем продукции предприятия в целом по месяцам или объем продукции по предприятиям, получая итог в целом по фирме).

Абсолютные величины могут быть получены путем суммирования данных статистического наблюдения или расчетным путем. Например, численность населения страны определяется по результатам сводки данных единовременного наблюдения. При определении стоимостных показателей объема продукции абсолютные величины получают расчетным путем.

Относительные величины

Относительные величины исчисляются при выполнении третьего этапа статистического исследования. **Относительная величина** представляет собой результат сопоставления двух статистических показателей, дает цифровую меру их соотношения. Она получается путем деления сравниваемого показателя на другой показатель, принимаемый за базу сравнения.

Относительные величины делятся на две группы:

• относительные величины, полученные в результате соотношения одноименных статистических показателей;

• относительные величины, представляющие результат сопоставления разноименных статистических показателей.

К относительным величинам первой группы относятся: относительные величины динамики, относительные величины планового задания и выполнения плана, относительные величины структуры, координации и наглядности.

Результат сопоставления одноименных показателей представляет собой краткое отношение (коэффициент), показывающее, во сколько раз сравниваемая величина больше (или меньше) базисной. Результат может быть выражен в процентах, показывая, сколько процентов сравниваемая величина составляет от базы.

Относительные величины динамики характеризуют изменение явления во времени. Они показывают, во сколько раз увеличился (или уменьшился) объем явления за определенный период времени, их называют коэффициентами роста. Коэффициенты роста можно исчислять в процентах, для этого отношения умножают на 100. Их называют темпами роста. Коэффициенты роста и темпы роста можно определять с переменной или постоянной базой.

Темпы роста с переменной базой получают при сравнении уровня явления каждого периода с уровнем предшествующего периода. Темпы роста с постоянной базой сравнения получают путем сопоставления уровня явления в каждом отдельном периоде с уровнем одного периода, принятого за базу. Выбор базы сравнения нередко имеет существенное значение. Так, в ряде случаев в качестве базы сравнения принимаются годы, являющиеся исторически обусловленной границей отдельных периодов времени.

y_1; y_2; y_3; y_4 – уровни явления за одинаковые последовательные периоды (например, выпуск продукции по кварталам года).

Темпы роста в процентах с переменной базой (цепные темпы роста):

$$T_{p_1} = \frac{y_2}{y_1} \cdot 100; \quad T_{p_2} = \frac{y_3}{y_2} \cdot 100; \quad T_{p_3} = \frac{y_4}{y_3} \cdot 100.$$

Темпы роста с постоянной базой (базисные темпы роста):

$$T'_{p_1} = \frac{y_1}{y_k} \cdot 100; \quad T'_{p_2} = \frac{y_2}{y_k} \cdot 100; \quad T'_{p_3} = \frac{y_3}{y_k} \cdot 100; \quad T'_{p_4} = \frac{y_4}{y_k},$$

где y_k – постоянная база сравнения.

Относительная величина планового задания – отношение величины показателя по плану ($y_{пл}$) к его фактической величине в предшествующем периоде (y_0), т. е. $y_{пл} : y_0$.

Относительная величина выполнения плана – отношение фактической (отчетной) величины показателя (y_1) к запланированной на тот же период его величине ($y_{пл}$), т. е. $y_1 : y_{пл}$.

Относительная величина динамики – отношение фактической (отчетной) величины показателя (y_1) к фактической величине предшествующего периода (y_0):

$$y_1 : y_0.$$

Относительные величины планового задания, выполнения плана и динамики связаны между собой.

Так,

$$\frac{y_1}{y_0} = \frac{y_{пл}}{y_0} \cdot \frac{y_1}{y_{пл}}$$

или

$$\frac{y_{пл}}{y_0} = \frac{y_1}{y_0} : \frac{y_1}{y_{пл}}; \quad \frac{y_1}{y_{пл}} = \frac{y_1}{y_0} : \frac{y_{пл}}{y_0}.$$

В ряде случаев расчет относительной величины выполнения плана может производиться по методу нарастающего итога. Так, оценка выполнения квартального плана по объему продукции выполняется по данным, взятым нарастающим итогом с начала квартала.

Относительные величины структуры характеризуют долю отдельных частей в общем объеме совокупности и выражаются в долях единицы или в процентах. Они исчисляются по сгруппированным данным:

$$\text{Относительная величина структуры, \%} = \frac{\text{Число единиц (или объем признака) по группе}}{\text{Общее число единиц (или объем признака) по всей совокупности}} \cdot 100.$$

Каждую относительную величину структуры называют *удельным весом.*

Относительные величины координации отражают отношение численности двух частей единого целого, т. е. показывают, сколько единиц одной группы приходится в среднем на одну, на десять или на сто единиц другой группы изучаемой совокупности (например, сколько служащих приходится на 100 рабочих).

Относительные величины наглядности отражают результаты сопоставления одноименных показателей, относящихся к одному и тому же периоду (или моменту) времени, но к разным объектам или территориям (например, сравнивается годовая производительность труда по двум предприятиям).

Вторая группа относительных величин, представляющая собой результат сопоставления разноименных статистических показателей, носит название **относительных величин интенсивности.**

Они являются именованными числами и показывают итог числителя, приходящийся на одну, на десять, на сто единиц знаменателя.

В эту группу относительных величин включаются показатели производства продукции на душу населения; показатели потребления продуктов питания и непродовольственных товаров на душу населения; показатели, отражающие обеспеченность населения материальными и культурными благами; показатели, характеризующие техническую оснащенность производства, рациональность расходования ресурсов:

$$\text{Показатель производства продукции на душу населения} = \frac{\text{Выпуск определенного вида продукции в натуральном выражении за год}}{\text{Среднегодовая численность населения}}.$$

$$\text{Обеспеченность населения материальными и культурными благами} = \frac{\text{Наличие определенных благ на начало (или конец) года}}{\text{Численность населения на начало (или конец) года}}.$$

Средние величины

Средней величиной называется обобщающий показатель, характеризующий типичный уровень варьирующего количественного признака на единицу совокупности в определенных условиях места и времени.

Объективность и типичность статистической средней обеспечивается лишь при определенных условиях. Первое условие – средняя должна вычисляться для качественно однородной совокупности. Для получения однородной совокупности необходима группировка данных, поэтому расчет средней должен сочетаться с методом группировок. Второе условие – для исчисления средних должны быть использованы массовые данные. В средней величине, исчисленной на основе данных о большом числе единиц (массовых данных), колебания в величине признака, вызванные случайными причинами, погашаются и проявляется общее свойство (типичный размер признака) для всей совокупности.

Средняя величина всегда именованная, она имеет ту же размерность, что и признак у отдельных единиц совокупности.

При использовании средних в практической работе и научных исследованиях необходимо иметь в виду, что за средним показателем скрываются особенности различных частей изучаемой совокупности, поэтому общие средние для однородной совокупности должны дополняться групповыми средними, характеризующими части совокупности.

В экономических исследованиях и плановых расчетах применяются две категории средних:

- степенные средние;
- структурные средние.

К категории **степенных средних** относятся: средняя арифметическая, средняя гармоническая, средняя квадратическая, средняя геометрическая. Величины, для которых исчисляется средняя, обозначаются буквой x_i. Средняя обозначается через $\bar{x}$. Такой

способ обозначения указывает на происхождение средней из конкретных величин. Черта вверху символизирует процесс осреднения индивидуальных значений. **Частота** – повторяемость индивидуальных значений признака – обозначается буквой f; $\Sigma f = n$.

Формулы средних величин могут быть получены на основе степенной средней, для которой определяющей функцией является уравнение

$$\sum_{i=1}^{n} x_i^k f_i = \sum_{i=1}^{n} \bar{x}^k f_i,$$

откуда

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

В дальнейшем при написании формул средних подстрочные значки i, n использоваться не будут, но подразумевается, что суммируются все произведения $x_i^k f_i$.

В зависимости от степени k получаются различные виды средних величин, их формулы представлены в табл. 2.1.

Таблица 2.1

Формулы различных видов степенных средних величин

Значение, k	Наименование средней	Формула средней: простая	Формула средней: взвешенная
–1	Гармоническая	$\bar{x} = \frac{n}{\sum \frac{1}{x}}$	$\bar{x} = \frac{\sum f}{\sum \frac{1}{x} \cdot f}$; $\bar{x} = \frac{\sum w}{\sum \frac{1}{x} \cdot w}$
0	Геометрическая	$\bar{x} = \sqrt[n]{x_1 \cdot x_2 \cdot x_n} = \sqrt[n]{\Pi_x}$	$\bar{x} = \sqrt[\Sigma f]{x_1^{f_1} \cdot x_2^{f_2} \ldots}$
1	Арифметическая	$\bar{x} = \frac{\sum x}{n}$	$\bar{x} = \frac{\sum x \cdot f}{\sum f}$; $\bar{x} = \frac{\sum x \cdot w}{\sum w}$
2	Квадратическая	$\bar{x} = \sqrt{\frac{\sum x^2}{n}}$	$\bar{x} = \sqrt{\frac{\sum x^2 \cdot f}{\sum f}}$

Как видно из данных табл. 2.1, взвешенные средние учитывают, что отдельные варианты значений признака имеют различную численность, поэтому каждый вариант «взвешивают» по своей частоте, т. е. умножают на нее. Частоты f при этом называются **статистическими весами** или **просто весами средней.** Однако необходимо учитывать, что статистический вес — понятие более широкое, чем частота. В качестве веса могут применяться какие-либо другие величины (в табл. 2.1 они обозначены буквой w). Например, при расчете средней продолжительности рабочего дня по предприятию единственно правильным будет взвешивание по количеству отработанных человеко-дней. Частоты отдельных вариантов могут быть выражены не только абсолютными величинами, но и относительными — **частостями.**

Величины степенных средних, рассчитанных на основе одних и тех же индивидуальных значений признака при различных значениях степени (k), не одинаковы. Чем выше степень k средней, тем больше величина самой средней:

$$\bar{x}_{\text{гарм}} < \bar{x}_{\text{геом}} < \bar{x}_{\text{арифм}} < \bar{x}_{\text{квадр}}.$$

Вопрос о выборе средней решается в каждом отдельном случае, исходя из задачи исследования, материального содержания изучаемого явления и наличия исходной информации. Он состоит из нескольких этапов:

1) устанавливается определяющий показатель, т. е. обобщающий показатель совокупности, от которого зависит величина средней;

2) определяется математическое выражение для определяющего показателя;

3) производится замена индивидуальных значений средними величинами;

4) решение уравнения средней.

Основополагающее правило при этом заключается в том, что величины, представляющие собой числитель и знаменатель средней, должны иметь определенный логический смысл.

Средняя арифметическая и средняя гармоническая наиболее распространенные виды средней, получившие широкое применение в плановых расчетах, при расчете общей средней из средних групповых, а также при выявлении взаимосвязи между признаками с помощью группировок. Выбор средней арифметической и средней гармонической определяется характером имеющейся в распоряжении исследователя информации.

Средняя квадратическая применяется для расчета среднего квадратического отклонения (σ), являющегося показателем вариации признаков, а также в технике (например, при сооружении трубопроводов).

Средняя геометрическая (простая) используется при вычислении среднего коэффициента роста (темпа) в рядах динамики, если промежутки времени, к которым относятся коэффициенты роста, одинаковы. Если средние коэффициенты роста относятся к периодам различной продолжительности, то общий средний коэффициент роста за весь период определяется по формуле средней геометрической взвешенной (f_i — продолжительность периода, к которому относится средний коэффициент роста).

Структурные средние — мода и медиана — в отличие от степенных средних, которые в значительной степени являются абстрактной характеристикой совокупности, выступают как конкретные величины, совпадающие с вполне определенными вариантами совокупности. Это делает их незаменимыми при решении ряда практических задач.

Модой называется значение признака, которое наиболее часто встречается в совокупности (в статистическом ряду).

Медианой называется значение признака, которое лежит в середине ранжированного ряда и делит этот ряд на две равные по численности части.

Ранжированный ряд — ряд, расположенный в порядке возрастания или убывания значений признака.

Для определения медианы сначала определяют ее место в ряду, используя формулу

$$N_{\text{Me}} = \frac{n+1}{2},$$

где n — число членов ряда.

Если ряд состоит из четного числа членов, то за медиану условно принимают среднюю арифметическую из двух срединных значений.

Применяется мода при экспертных оценках, при определении наиболее ходовых размеров обуви, одежды, что учитывается при планировании их производства. Медиана используется при статистическом контроле качества продукции и технологического процесса на промышленных предприятиях; при изучении распределения семей по величине дохода и др. Мода и медиана имеют преимущество перед средней арифметической для ряда распределения с открытыми интервалами.

Графические изображения статистических данных

Графические изображения статистических данных облегчают их обобщение и анализ. Графики применяются для характеристики развития явления во времени, в пространстве, отображения структуры явления и структурных сдвигов, при контроле за выполнением плана, изучении взаимосвязи между явлениями.

По способу построения графики делятся на диаграммы, картограммы и картодиаграммы.

Диаграмма – изображение статистических данных при помощи геометрических фигур, линий, точек.

Картограмма – это географическая (контурная) карта, которая графически характеризует пространственное распределение какого-либо статистического показателя путем различной окраски, штриховки и т. д. (например, плотность населения в различных регионах).

Картодиаграмма – это совмещение картограммы с диаграммой, т. е. в отдельных районах условными знаками наносят абсолютные значения статистических показателей.

Самым распространенным видом графиков являются диаграммы, которые делятся на линейные, столбиковые, структурные, фигурные, знаки Варзара и др.

Линейные диаграммы – наиболее простой способ наглядного изображения статистических данных, когда изучаемое явление представляется в виде отрезков ломаной линии, называемой статистической кривой. Они применяются для характеристики и сравнения развития различных явлений во времени, пространстве, а также для отображения взаимосвязи между явлениями.

Для построения линейной диаграммы используется прямоугольная система координат. На оси абсцисс (по горизонтальной шкале) откладываются равные отрезки, представляющие собой периоды времени, на ось ординат наносят масштаб для отображения уровня явления. Соединение точек, построенных на координатной системе, дает ломаную линию, представляющую собой закономерность развития явления.

Рекомендуется строить координатную сетку с учетом соотношения масштабов по осям координат примерно 1 : 1,5 (правило «золотого сечения»), т. е. с учетом соотношения масштабов по сторонам занятого графиком пространства по вертикали и горизонтали.

Преимуществом линейных графиков является то, что на одном графике имеется возможность отображения закономерности нескольких явлений. Разновидностью линейных диаграмм являются контрольно-плановые графики, обеспечивающие оперативный контроль за ходом выполнения задания как за отдельные промежутки (дни, пятидневки), так и нарастающим итогом с начала периода.

Для сравнения различных величин между собой и для изображения динамики могут быть использованы **столбиковые (ленточные) диаграммы**. Для их построения также используется система прямоугольных координат. Основания столбиков одинакового размера, представляющие собой периоды времени (годы, месяцы, дни), размещаются на оси абсцисс, а вершины столбиков соответствуют величине изучаемого показателя. Столбиковые диаграммы называют ленточными, если столбики расположены горизонтально в виде лент.

Структурные диаграммы применяются для изображения структуры явления и характеристики структурных сдвигов. При построении таких графиков состав совокупности выражается относительными величинами структуры, исчисленными в процентах. Они могут быть двух видов: столбиковые и круговые. Общая высота столбика и площадь круга отображают целое и принимаются соответственно за 100%.

При построении круговой диаграммы необходимо проценты перевести в градусы, учитывая, что каждый процент равен 3,6° (360 : 100).

Знаки Варзара (по имени статистика В. Е. Варзара) являются разновидностью столбиковых диаграмм. Они позволяют отобразить на графике сложное явление, представляющее собой произведение двух показателей. Например, объем продукции – произведение производительности труда и численности работников. Если в прямоугольнике одну сторону взять пропорционально уровню производительности труда, а другую – пропорционально численности работников, то площадь прямоугольника будет пропорциональна объему продукции.

2.1. Решение типовых задач

2.1. Расход топлива на производственные нужды предприятия характеризуется в отчетном периоде следующими данными (табл. 2.2).

Таблица 2.2

Вид топлива	Единица измерения	Расход	
		по плану	фактически
Мазут топочный	т	500	520
Уголь	»	320	300
Газ природный	тыс. $м^3$	650	690

Средние калорийные эквиваленты (коэффициенты) перевода в условное топливо составили: мазут – 1,37 т; уголь – 0,9 т; газ – 1,2 тыс. $м^3$.

Определить:

1) общее потребление условного топлива по плану и фактически;

2) процент выполнения плана по общему расходу топлива;

3) удельные веса фактически израсходованного топлива по видам (расчет с точностью до 0,1%).

Решение

1. Для определения общего потребления топлива используется условно-натуральный метод; расходы по плану и фактически исчисляются в единицах условного топлива (усл. ед.):

$$y_{пл} = 500 \cdot 1{,}37 + 320 \cdot 0{,}9 + 650 \cdot 1{,}2 = 1753;$$

$$y_1 = 520 \cdot 1{,}37 + 300 \cdot 0{,}9 + 690 \cdot 1{,}2 = 1810{,}4.$$

2. Процент выполнения плана по общему расходу топлива:

$$\frac{y_1}{y_{пл}} \cdot 100 = \frac{1810{,}4}{1753} \cdot 100 = 103{,}27\%.$$

Следовательно, фактический расход топлива превышает плановый на 3,27%.

3. Для определения удельного веса израсходованного топлива по видам (структуры расхода топлива) используется вспомогательная табл. 2.3.

$$(712{,}4 : 1810{,}4) \cdot 100 = 39{,}4\% \text{ и т. д.}$$

Таблица 2.3

Фактический расход топлива

Вид топлива	Израсходовано условных единиц	Удельный вес в общем объеме расхода, %
Мазут топочный	712,4	39,4
Уголь	270,0	14,9
Газ природный	828,0	45,7
Итого	1810,4	100,0

2.2. По региону имеются следующие данные о вводе в эксплуатацию жилой площади (табл. 2.4).

Таблица 2.4

Вид жилых домов	Введено в эксплуатацию, тыс. $м^2$	
	прошлый год	отчетный год
Кирпичные многоквартирные	4400	4200
Панельные многоквартирные	2800	2100
Коттеджи	800	2100

Определить:

1) динамику ввода в эксплуатацию жилой площади по каждому виду жилых домов и в целом по региону;

2) структуру введенной в эксплуатацию жилой площади в прошлом и отчетном годах (расчет с точностью до 0,1%).

3) структуру введенной в эксплуатацию площади представить на графике.

Сформулировать вывод.

Решение

1. Показатели динамики (темпы роста) следующие, %:

кирпичные дома – $T_р = \frac{y_1}{y_0} = \frac{4200}{4400} = 0{,}955$, или 95,5;

панельные дома – $T_р = \frac{2100}{2800} = 0{,}750$, или 75,0;

коттеджи – $T_р = \frac{2100}{800} = 2{,}625$, или 262,5.

В целом по региону – $T_р = \frac{4200 + 2100 + 2100}{4400 + 2800 + 800} = \frac{8400}{8000} = 1{,}05$,

или 105,0%.

Следовательно, ввод в эксплуатацию жилой площади в кирпичных домах уменьшился на 4,5% (95,5 – 100); в панельных домах снизился на 25,0% (75,0 – 100); по коттеджам увеличился на 162,5% (262,5 – 100), в целом по региону ввод жилой площади возрос на 5,0% (105 – 100).

2. Структура введенной в эксплуатацию жилой площади по региону представлена в табл. 2.5 и на рис. 2.1.

Как видно из данных табл. 2.5, существенно увеличился удельный вес вводимой жилой площади по коттеджам.

Таблица 2.5

Структура введенной в эксплуатацию жилой площади по региону

Вид жилых домов	Прошлый год		Отчетный год	
	введено, тыс. м^2	в % к итогу	введено, тыс. м^2	в % к итогу
Кирпичные многоквартирные	4400	55,0	4200	50,0
Панельные многоквартирные	2800	35,0	2100	25,0
Коттеджи	200	10,0	2100	25,0
Итого	8000	100,0	8400	100,0

Рис. 2.1. Структура введенной в эксплуатацию жилой площади по видам домов по региону

2.3. Имеются следующие данные о стоимости основного капитала по фирме (табл. 2.6).

Таблица 2.6

№ предприятия, входящего в фирму	Стоимость основного капитала, млн руб.		
	на 1 января 2000 г.	на 1 января 2001 г.	на 1 января 2002 г.
1	22 150	24 855	26 970
2	7 380	9 100	12 550
3	13 970	16 700	20 800

Определить показатели динамики стоимости основного капитала фирмы.

Решение

Стоимость основного капитала фирмы определяется путем суммирования данных по предприятиям (млн руб.):

на 1 января 2000 г. – y_1 = 22 150 + 7 380 + 13 970 = 43 500;

на 1 января 2001 г. – y_2 = 24 855 + 9 100 + 16 700 = 50 655;

на 1 января 2002 г. – y_3 = 26 970 + 12 550 + 20 800 = 60 320.

Для характеристики динамики определяются темпы роста:

а) с переменной базой:

$$T_{p_1} = \frac{y_2}{y_1} \cdot 100 = \frac{50\,655}{43\,500} \cdot 100 = 116{,}4\%;$$

$$T_{p_2} = \frac{y_3}{y_2} \cdot 100 = \frac{60\,320}{50\,655} \cdot 100 = 119{,}1\%.$$

Следовательно, на 1 января 2001 г. стоимость основного капитала увеличилась по сравнению с началом предыдущего года на 16,4% (116,4 – 100), а на 1 января 2002 г. увеличение составило по сравнению с данными на 1 января 2001 г. 19,1% (119,1 – 100);

б) с постоянной базой (за постоянную базу принимаются данные о стоимости основного капитала на 1 января 2000 г.):

$$T'_{p_1} = \frac{y_2}{y_1} \cdot 100 = \frac{50\,655}{43\,500} \cdot 100 = 116{,}4\%;$$

$$T'_{p_2} = \frac{y_3}{y_1} \cdot 100 = \frac{60\,320}{43\,500} \cdot 100 = 138{,}7\%.$$

Таким образом, на 1 января 2001 г. стоимость основного капитала увеличилась по сравнению с данными на 1 января 2000 г. на 16,4%, а на 1 января 2002 г. – на 38,7%.

2.4. По промышленному предприятию за отчетный год имеются следующие данные о выпуске продукции (табл. 2.7).

Таблица 2.7

Наименование продукции	План на I квартал, тыс. т	Фактический выпуск, тыс. т			Отпускная цена за 1 т, руб.
		январь	февраль	март	
Сталь арматурная	335	110	115	108	1700
Прокат листовой	255	75	90	100	2080

Определить процент выполнения квартального плана по выпуску каждого вида продукции и в целом по выпуску всей продукции.

Решение

Фактический выпуск каждого вида продукции за I квартал следующий, тыс. т.:

сталь арматурная – $y_1 = 110 + 115 + 108 = 333$;

прокат листовой – $y_1 = 75 + 90 + 110 = 265$.

Процент выполнения квартального плана по выпуску каждого вида продукции:

сталь арматурная:

$$\frac{y_1}{y_{пл}} \cdot 100 = \frac{333}{335} \cdot 100 = 99{,}4\%,$$

т. е. фактический выпуск ниже плана на 0,6% (99,4 – 100);

прокат листовой:

$$\frac{y_1}{y_{пл}} \cdot 100 = \frac{265}{255} \cdot 100 = 103{,}9\%,$$

т. е. план перевыполнен на 3,9% (103,9 – 100).

Для расчета процента выполнения плана по выпуску всей продукции необходимо определить общий итог продукции по плану и фактический в денежном выражении, руб.:

$$\sum y_{пл} = 335\,000 \cdot 1\,700 + 255\,000 \cdot 2\,080 = 1\,099\,900\,000;$$

$$\sum y_1 = 333\,000 \cdot 1\,700 + 265\,000 \cdot 2\,080 = 1\,117\,300\,000.$$

Процент выполнения плана по выпуску всей продукции –

$$\frac{1\,117\,300\,000}{1\,099\,900\,000} \cdot 100 = 101{,}6\%.$$

Следовательно, план выпуска всей продукции перевыполнен на 1,6%.

2.5. По фирме имеются следующие данные о выпуске продукции за год (табл. 2.8).

Определить процент выполнения плана выпуска продукции в целом по фирме.

Таблица 2.8

№ предприятия, входящего в фирму	Фактический выпуск продукции, млн руб.	Процент выполнения плана
1	29,4	105,0
2	42,6	100,0
3	24,0	96,0

Решение

Для расчета процента выполнения плана выпуска продукции по фирме определяется плановый выпуск:

$$y_{пл} = \frac{29{,}4 \cdot 100{,}0}{105{,}0} + \frac{42{,}6 \cdot 100{,}0}{100{,}0} + \frac{24{,}0 \cdot 100{,}0}{96{,}0} =$$

$$= 28{,}0 + 42{,}6 + 25{,}0 = 95{,}6 \text{ млн руб.}$$

Фактический выпуск продукции по фирме

$$y_1 = 29{,}4 + 42{,}6 + 24{,}0 = 96{,}0 \text{ млн руб.}$$

Процент выполнения плана по фирме

$$\frac{96{,}0}{95{,}6} \cdot 100 = 100{,}4\%.$$

2.6. В прошлом году объем грузооборота по грузовому автотранспортному предприятию составил 210,0 млн ткм. Планом текущего года было предусмотрено довести объем грузооборота до 220,5 млн ткм; фактический объем грузооборота в текущем году составил 229,32 млн ткм.

Определить:

1) относительную величину планового задания по росту грузооборота;

2) относительную величину динамики грузооборота;

3) относительную величину выполнения плана по грузообороту.

Решение

Введем условные обозначения:

y_0 = 210 млн ткм – фактический грузооборот в прошлом году;

$y_{пл}$ = 220,5 млн ткм – плановый грузооборот в текущем году;

y_1 = 229,32 млн ткм – фактический грузооборот в текущем году.

Относительная величина планового задания

$$\frac{y_{пл}}{y_0} = \frac{220{,}5}{210{,}0} = 1{,}05, \text{ или } 105{,}0\%.$$

Следовательно, по плану предусмотрено увеличение грузооборота на 5,0% (105,0 – 100).

Относительная величина динамики

$$\frac{y_1}{y_0} = \frac{229{,}32}{210{,}0} = 1{,}092, \text{ или } 109{,}2\%.$$

Следовательно, фактически по сравнению с прошлым годом грузооборот повысился на 9,2%.

Относительная величина выполнения плана

$$\frac{y_1}{y_{пл}} = \frac{229{,}32}{220{,}5} = 1{,}04, \text{ или } 104{,}0\%,$$

т. е. план по размеру грузооборота превышен на 4,0%.

2.7. Планом предусмотрено увеличение годовой производительности труда работников против прошлого года на 4,0%. Фактически против прошлого года производительность труда увеличилась на 6,2%.

Определить процент выполнения плана по уровню производительности труда.

Решение

Введем условные обозначения:

y_0 – уровень производительности труда в прошлом году;

$y_{пл}$ – плановый уровень производительности труда в отчетном году;

y_1 – фактический уровень производительности труда в отчетном году.

Относительная величина планового задания

$$\frac{y_{пл}}{y_0} = 1{,}04.$$

Относительная величина динамики

$$\frac{y_1}{y_0} = 1{,}062.$$

Относительная величина выполнения плана

$$\frac{y_1}{y_{пл}} = \frac{y_1}{y_0} : \frac{y_{пл}}{y_1} = 1{,}062 : 1{,}04 = 1{,}021, \quad \text{или } 102{,}1\%.$$

2.8. По плану объем продукции в отчетном году должен возрасти против прошлого года на 2,5%. План выпуска продукции перевыполнен на 3,0%.

Определить фактический выпуск продукции в отчетном году, если известно, что объем продукции в прошлом году составил 25 300 тыс. руб.

Решение

Введем условные обозначения:
y_0 – выпуск продукции в прошлом году;
$y_{пл}$ – выпуск продукции по плану в отчетном году;
y_1 – фактический выпуск продукции в отчетном году.
По условию задачи y_0 = 25 300 тыс. руб.
Относительная величина планового задания

$$\frac{y_{пл}}{y_0} = 1{,}025,$$

отсюда

$$y_{пл} = y_0 \cdot 1{,}025 = 25\,300 \cdot 1{,}025 = 25932{,}5 \text{ тыс. руб.}$$

Относительная величина выполнения плана

$$\frac{y_1}{y_{пл}} = 1{,}03,$$

отсюда $y_1 = y_{пл} \cdot 1{,}03 = 25932{,}5 \cdot 1{,}03 = 26710{,}5$ тыс. руб.

2.9. По автотранспортному предприятию за два года имеются данные о численности рабочих (табл. 2.9).

Таблица 2.9

Показатель	Прошлый год	Отчетный год
Среднесписочная численность рабочих	1092	1251
В том числе:		
водители	780	900
ремонтно-вспомогательные рабочие	312	351

Охарактеризовать изменения в соотношениях численности водителей и ремонтно-вспомогательных рабочих с помощью относительных величин координации.

Решение

Относительные величины координации следующие:

прошлый год – $\frac{312}{780} \cdot 100 = 40$ чел.,

т. е. на 100 водителей приходилось 40 человек ремонтно-вспомогательных рабочих;

отчетный год – $\frac{351}{900} \cdot 100 = 39$ чел.

Следовательно, в текущем (отчетном) году численность ремонтно-вспомогательных рабочих на 100 водителей уменьшилась на одного человека (39 – 40), т. е. на 2,5% [(1 : 40) · 100].

2.10. По двум промышленным предприятиям за отчетный год имеются следующие данные (табл. 2.10).

Таблица 2.10

№ предприятия	Выпуск продукции, млн руб.	Среднесписочная численность рабочих, чел.
1	360,0	1200
2	693,0	1980

Определить различие (в %) в уровне годовой производительности труда работников двух предприятий.

Решение

Определяется уровень годовой производительности труда работников каждого предприятия:

$$\text{предприятие 1: } y_1 = \frac{360\,000}{1\,200} = 300 \text{ тыс. руб.;}$$

$$\text{предприятие 2: } y_2 = \frac{693\,000}{1\,980} = 350 \text{ тыс. руб.}$$

Для ответа определяется относительная величина наглядности:

$$\frac{y_2}{y_1} = \frac{350}{300} = 1{,}167, \ 116{,}7\%.$$

Следовательно, на предприятии № 2 уровень производительности труда на 16,7% выше, чем на предприятии № 1.

2.11. Потребление электроэнергии в регионе характеризуется следующими данными: 2000 г. – 43,1 млрд кВт · ч; 2001 г. – 49,8 млрд кВт · ч. Численность населения региона составила (млн чел.): на 1 января 2000 г. – 8,8; 1 января 2001 г. – 9,0; 1 января 2002 г. – 9,3.

Определить, на сколько процентов изменилось потребление электроэнергии на душу населения.

Решение

Определяется потребление электроэнергии на душу населения по формуле:

$$y = \frac{\text{Количество потребляемой электроэнергии за год}}{\text{Среднегодовая численность населения}}.$$

$$2000 \text{ г.} - y_1 \frac{43\,100\,000\,000}{(8\,800\,000 + 9\,000\,000):2} = 4842{,}7 \text{ кВт·ч;}$$

$$2001 \text{ г.} - y_2 = \frac{49\,800\,000\,000}{(9\,000\,000 + 9\,300\,000):2} = 5442{,}6 \text{ кВт·ч.}$$

Относительная величина динамики

$$\frac{y_2}{y_1} = \frac{5442{,}6}{4842{,}7} = 1{,}124.$$

Следовательно, потребление энергии на душу населения увеличилось на 12,4% (1,124 · 100 – 100).

2.12. Имеются следующие данные о заработной плате рабочих участка (табл. 2.11).

Таблица 2.11

Профессия	Количество рабочих	Заработная плата каждого рабочего за сентябрь, руб.
Токари	5	4700; 4208; 1917; 3620; 4400
Фрезеровщики	2	3810; 4550
Слесари	3	5210; 3380; 1870

Вычислить среднюю месячную заработную плату рабочих участка.

Решение

Процесс выбора средней таков:

- определяющий показатель – общая сумма начисленной заработной платы;
- математическое выражение определяющего показателя $-\sum x$;
- замена индивидуальных значений средними $-\sum x = n \cdot \bar{x}$;
- решение уравнения.

$$\bar{x} = \frac{\sum x}{n} =$$

$$= \frac{4700 + 4208 + 1917 + 3620 + 4400 + 3810 + 4550 + 5210 + 3380 + 1870}{10} =$$

$$= \frac{37\,665}{10} = 3766{,}5 \text{ руб.}$$

Следовательно, использовалась формула простой средней арифметической.

2.13. Распределение рабочих участка по стажу работы следующее (табл. 2.12).

Таблица 2.12

Стаж работы (лет), x	До 5 лет	5 – 10	10 – 15	15 и более
Количество рабочих, f	2	6	15	7

Определить средний стаж работы рабочих участка.

Решение

Определяющий показатель – общий стаж работы всех рабочих $-\Sigma xf$.

Средний стаж работы –

$$\bar{x} = \frac{\Sigma x'f}{\Sigma f} = \frac{2{,}5 \cdot 2 + 7{,}5 \cdot 6 + 12{,}5 \cdot 15 + 17{,}5 \cdot 7}{2 + 6 + 15 + 7} = 12{,}0 \text{ лет.}$$

Для каждого интервала предварительно вычислялось среднее значение признака как полусумма нижнего и верхнего значений интервала. Величина открытых интервалов приравнивается к величине примыкающих к ним соседних интервалов:

$$x_1' = \frac{0+5}{2} = 2{,}5; \quad x_2' = \frac{5+10}{2} = 7{,}5;$$

$$x_3' = \frac{10+15}{2} = 12{,}5; \quad x_4' = \frac{15+20}{2} = 17{,}5.$$

Для решения задачи использовалась формула средней арифметической взвешенной.

2.14. За два месяца по цехам завода имеются следующие данные (табл. 2.13).

Определить, за какой месяц и на сколько процентов была выше средняя месячная заработная плата работников предприятия.

Таблица 2.13

№ цеха	Сентябрь		Октябрь	
	численность работников	средняя месячная заработная плата, руб.	средняя месячная заработная плата, руб.	фонд заработной платы руб.
1	140	3 560	3 600	486 000
2	200	3 600	3 580	751 800
3	260	3 330	3 340	835 000

Решение

Введем условные обозначения для сентября:

f – численность работников по каждому цеху;

x – средняя месячная заработная плата работников каждого цеха.

Определяющий показатель – общий фонд заработной платы – Σxf.

Средняя месячная заработная плата работников предприятия за сентябрь составила:

$$\bar{x}_c = \frac{\Sigma x \cdot f}{\Sigma f} = \frac{3560 \cdot 140 + 3600 \cdot 200 + 3330 \cdot 260}{140 + 200 + 260} = \frac{2\ 084\ 200}{600} = 3473{,}7 \text{ руб.}$$

Условные обозначения для октября следующие:

w – фонд заработной платы по каждому цеху;

x – средняя месячная заработная плата работников каждого цеха.

Определяющий показатель $-\Sigma w$.

Средняя месячная заработная плата работников предприятия за октябрь равна:

$$\bar{x}_0 = \frac{\Sigma W}{\Sigma \frac{W}{x}} = \frac{486\ 000 + 751\ 800 + 835\ 000}{\frac{486\ 000}{3\ 600} + \frac{751\ 800}{3\ 580} + \frac{835\ 000}{3\ 340}} = \frac{2\ 072\ 800}{135 + 210 + 250} = 3483{,}7 \text{руб.}$$

где $\frac{W}{x}$ – численность работников каждого цеха в октябре.

Средняя заработная плата в октябре исчислена по формуле средней взвешенной гармонической.

Динамика средней месячной заработной платы работников предприятия:

$$\frac{\bar{x}_0}{\bar{x}_c}=\frac{3483{,}7}{3473{,}7}=1{,}003,\ \text{или}\ 100{,}3\%.$$

Следовательно, средняя месячная заработная плата работников предприятия в октябре повысилась на 0,3% по сравнению с сентябрем.

2.15. Имеются следующие данные об экспорте продукции металлургического комбината (табл. 2.14).

Таблица 2.14

Вид продукции	Удельный вес продукции на экспорт, %	Стоимость продукции на экспорт, тыс. руб.
Сталь арматурная	40,0	32 100
Прокат листовой	32,0	42 500

Определить средний удельный вес продукции на экспорт.

Решение

$$\text{Удельный вес продукции на экспорт, \%}=\frac{\text{Стоимость продукции на экспорт}}{\text{Стоимость всей продукции}}\cdot 100.$$

$\frac{W}{x}\cdot 100$ – стоимость всей продукции,

где w – стоимость продукции на экспорт;

x – удельный вес продукции на экспорт.

Средний удельный вес продукции на экспорт:

$$\bar{x}=\frac{\sum W}{\sum\frac{W}{x}\cdot 100}\cdot 100=\frac{32\,100+42\,500}{\frac{32\,100}{40{,}0}\cdot 100+\frac{42\,500}{32{,}0}\cdot 100}\cdot 100=$$

$$=\frac{32\,100+42\,500}{80\,250+132\,812{,}5}\cdot 100=\frac{74\,600}{213\,062{,}5}\cdot 100=35{,}0\%.$$

2.16. Распределение промышленных предприятий региона по показателю затрат на 1 тыс. руб. продукции в сентябре следующее (табл. 2.15).

Таблица 2.15

Затраты на 1 тыс. руб. продукции, руб.	Число предприятий	Общая стоимость продукции, тыс. руб.
600 – 650	2	19 800
650 – 700	8	66 000
700 – 750	4	32 000
750 – 800	3	21 450

Определить:

1) средний размер затрат на 1 тыс. руб. продукции по предприятиям региона;

2) средний объем продукции на одно предприятие.

Решение

Введем условные обозначения:

x – размер затрат на 1 тыс. руб. продукции по каждой группе предприятий;

f – число предприятий в каждой группе;

w – общая стоимость продукции по каждой группе предприятий.

Определяющий показатель – общая сумма затрат на выпуск продукции по всем предприятиям региона – $\Sigma x\cdot w$; $\Sigma xw=\bar{x}\cdot\Sigma w$.

Средний размер затрат на 1 тыс. руб. продукции по предприятиям региона – $\bar{x}=\frac{\Sigma xw}{\Sigma w}$;

$$\bar{x}=\frac{625\cdot 19\,800+675\cdot 66\,000+725\cdot 32\,000+775\cdot 21\,450}{19\,800+66\,000+32\,000+21\,450}=$$

$$= \frac{96\,748\,750}{139\,250} = 694{,}8 \text{ руб.}$$

Средний размер объема продукции на одно предприятие:

$$\bar{x} = \frac{\Sigma w}{\Sigma f} = \frac{139\,250}{17} = 8191{,}2 \text{ тыс. руб.}$$

2.17. Подача жидкого топлива для технологического процесса осуществляется в цехе тремя трубопроводами с диаметрами 2, 5 и 6 см. При капитальном ремонте здания цеха эти трубопроводы будут заменены на три новых одинакового диаметра при сохранении их общей пропускной способности.

Определить средний диаметр трубы (диаметр новой трубы).

Решение

Пропускная способность труб – определяющий показатель. r – радиус труб;

$$r_1 = 1;\ r_2 = 2{,}5;\ r_3 = 3{,}0.$$

$$\pi r_1^{\,2} + \pi r_2^2 + \pi r_3^2 = 3\pi \bar{r}^{\,2}.$$

$$\bar{r} = \sqrt{\frac{r_1^2 + r_2^2 + r_3^2}{3}} = \sqrt{\frac{1^2 + 2{,}5^2 + 3^2}{3}} = \sqrt{\frac{1 + 6{,}25 + 9}{3}} = \sqrt{\frac{16{,}25}{3}} = 1{,}73 \text{ см.}$$

$$\text{Д} = 2\,\bar{r} = 2 \cdot 1{,}73 = 3{,}46 \text{ см.}$$

2.18. Проведена малая выборка из партии электрических лампочек для определения продолжительности их службы. Результаты представлены в табл. 2.16.

Таблица 2.16

№ лампочки	1	2	3	4	5	6	7	8	9
Срок горения, ч.	1450	1400	1370	1430	1400	1380	1270	1420	1400

Определить моду и медиану.

Решение

Для определения моды и медианы производится ранжирование данных.

Ранжированный ряд: 1270; 1370; 1380; 1400; 1400; 1400; 1420; 1430; 1450.

Мода – Мо = 1400 ч (1400 – значение признака, встречающееся три раза).

Место медианы – $N_{\text{Me}} = \frac{n+1}{2} = \frac{9+1}{2} = 5.$

Ме = 1400 ч (1400 – значение признака, находящееся на 5-м месте в ранжированном ряду).

2.19. Данные о перевозке грузов по автотранспортному предприятию представлены в табл. 2.17.

Таблица 2.17

	Январь	Февраль	Март	Апрель
Перевезено грузов, тыс. т	35	40	42	50

Требуется:

1) определить среднемесячный темп роста объема грузовых перевозок;

2) построить график, отражающий динамику объема грузовых перевозок.

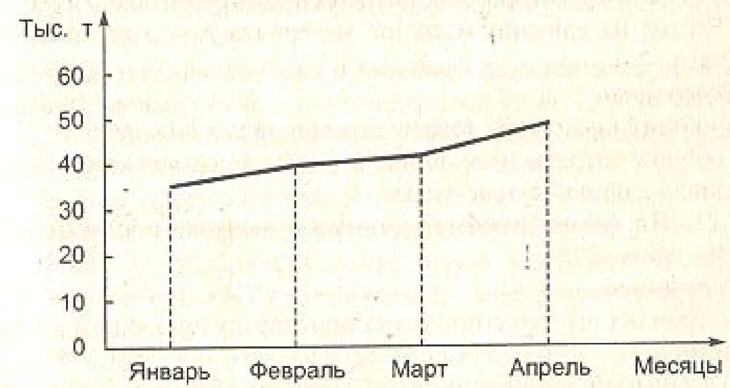

Рис. 2.2. Динамика перевозки грузов

1. Для характеристики развития грузовых перевозок во времени целесообразно использовать линейную диаграмму; она представлена на рис. 2.2.

2. Чтобы определить средний месячный темп роста перевозок, сначала определяется средний месячный коэффициент роста по формуле средней геометрической:

$$\bar{x} = \sqrt[3]{x_1 \cdot x_2 \cdot x_3},$$

где x_1, x_2, x_3 – коэффициент роста с переменной базой.

$$x_1 = \frac{40,0}{35,0} = 1,143; \quad x_2 = \frac{42,5}{40,0} = 1,063; \quad x_3 = \frac{50,0}{42,5} = 1,176;$$

$$\bar{x} = \sqrt[3]{1,143 \cdot 1,063 \cdot 1,176} = \sqrt[3]{1,429} = 1,126.$$

Средний месячный темп роста перевозок

$$\bar{T}_р = \bar{x} \cdot 100 = 1,126 \cdot 100 = 112,6\%.$$

2.2. Задачи для самостоятельной работы

2.20. В январе отчетного года фирма выпустила 2000 шт. изделий. На изготовление изделия расходуется два вида материала: А и Б, цена за 1 т которых составляет соответственно 6,0 и 9,0 тыс. руб. Расход на единицу изделия: материала А – 3 кг, материала Б – 2 кг.

Определить:

1) общий расход каждого вида материала в январе;

2) общие затраты на материалы в январе и долю каждого вида материала в общей сумме затрат.

2.21. По фирме имеются данные о выпуске продукции за I квартал (табл. 2.18).

Определить:

1) процент выполнения плана по выпуску продукции в целом по фирме;

2) удельный вес предприятий в общем объеме фактического выпуска продукции (расчет с точностью до 0,1%).

Таблица 2.18

№ предприятия фирмы	Выпуск продукции по плану, млн руб.	Процент выполнения плана по выпуску продукции
1	10,0	103,5
2	24,0	98,0
3	42,5	106,0

2.22. По металлургическому комбинату имеются следующие данные о выпуске продукции (табл. 2.19).

Таблица 2.19

Наименование продукции	Стоимость продукции в фиксированных ценах, млн руб.		Процент выполнения плана по выпуску продукции
	по плану	фактически	
Сталь арматурная	440	452	
Прокат листовой	500		97,0
Гнутые профили стальные		208	104,0

Требуется:

1) проставить в таблице недостающие данные;

2) определить процент выполнения плана выпуска продукции в целом по комбинату;

3) структуру фактического выпуска продукции представить в виде диаграммы.

2.23. В прошлом году себестоимость производства грузового автомобиля КамАЗ-55111 составила 70,0 тыс. руб. По плану отчетного года предусматривалось снизить себестоимость на 1400 руб., фактическая себестоимость составила 68,2 тыс. руб.

Определить относительные величины планового задания по снижению себестоимости и динамики себестоимости производства автомобиля.

2.24. Планом предусмотрено увеличение объема продукции предприятия против прошлого года на 2,1%. Фактически прирост продукции против прошлого года составил 4,8%.

Определить процент выполнения плана по выпуску продукции.

2.25. По плану отчетного года уровень годовой производительности труда работников должен возрасти против прошлого года на 3,0%. План по уровню производительности труда перевыполнен на 2,0%.

Определить фактический уровень производительности труда, если известно, что в прошлом году уровень годовой производительности труда составил 680 тыс. руб.

2.26. Предприятие перевыполнило план реализации продукции в отчетном году на 3,8%. Увеличение реализации продукции в отчетном году по сравнению с прошлым составило 5,6%.

Определить, каково было плановое задание по росту объема реализации продукции.

2.27. За отчетный квартал потребление топлива на производственные нужды по предприятию следующее: уголь – 1200 т, газ – 380 тыс. м^3, нефть – 210 т.

Определить, какую долю в общем объеме потребленного топлива занимает уголь, если коэффициенты пересчета в условное топливо следующие: уголь – 0,9 т; газ – 1,2 тыс. м^3; нефть – 1,3 т.

2.28. Планом предусмотрено снижение затрат на 1 руб. продукции на 4,0%; фактически по сравнению с прошлым годом затраты возросли на 1,8%.

Определить, на сколько процентов фактические затраты на 1 руб. продукции отличаются от плановых.

2.29. По отделению дороги планом предусмотрено увеличение объема отправок груза на 10,0%. Фактически объем отправок против прошлого года повысился на 12,2%.

Определить, на сколько процентов перевыполнен план по объему отправок груза.

2.30. Данные о жилищном фонде и численности населения России представлены в табл. 2.20.

Таблица 2.20

Показатель	1998 г.	1999 г.	2000 г.
Весь жилой фонд на начало года, млн м^2	2715	2745	2770
Численность населения на начало года, млн чел.	146,7	146,3	145,6

Требуется:

1) охарактеризовать изменение обеспеченности населения жилой площадью;

2) перечислить, какие виды относительных величин использовались.

2.31. По предприятиям фирмы имеются следующие данные (табл. 2.21).

Таблица 2.21

№ предприятия, входящего в фирму	Фактический объем реализованной продукции в 1999 г., млн руб.	Плановое задание по росту реализованной продукции в 2000 г., %	Фактический объем реализованной продукции в 2000 г., млн руб.
1	30,0	104,0	32,6
2	48,5	106,0	52,7
3	60,0	102,5	63,0

Определить в целом по фирме:

1) размер планового задания по росту объема реализованной продукции в 2000 г.;

2) процент выполнения плана по объему реализованной продукции в 2000 г.;

3) показатель динамики реализованной продукции.

2.32. Данные о численности экономически активного населения и безработных в России представлены в табл. 2.22.

Таблица 2.22

Показатели	1997 г.	1998 г.	1999 г.
Экономически активное население – всего	68079	66736	69701
в том числе:			
мужчины	35925	35273	36767
женщины	32154	31463	32934
Безработные – всего	6732	8058	9070
в том числе:			
мужчины	3662	4371	4757
женщины	3070	3687	4313

Требуется:

1) определить удельный вес численности безработных в общей численности экономически активного населения и динамику этого показателя для каждой группы населения;

2) с помощью относительных величин наглядности дать сравнительную оценку уровня безработицы среди мужчин и женщин.

2.33. Имеются следующие данные о квалификации рабочих двух бригад (табл. 2.23).

Таблица 2.23

№ бригады	Число рабочих	Уровень квалификации каждого рабочего бригады (тарифный разряд)
1	12	4; 3; 2; 4; 5; 6; 4; 3; 4; 3; 5; 4
2	10	3; 5; 6; 5; 4; 3; 2; 3; 3; 4

Определить средний уровень квалификации рабочих каждой бригады.

2.34. Продажа грузовых автомобилей КамАЗ-55111 на товарной бирже города характеризуется следующими данными (табл. 2.24).

Таблица 2.24

Дата торга	Реализовано автомобилей, шт.	Средняя цена одного автомобиля, тыс. руб.	Дата торга	Общая сумма выручки от реализации автомобилей, тыс. руб.	Средняя цена одного автомобиля, тыс. руб.
4.01	18	120,5	3.02	1830	122,0
17.01	25	118,7	9.02	2651	120,5
28.01	24	116,0	20.02	4165	119,0
			26.02	1232	123,2

Определить, на сколько процентов изменилась средняя цена одного грузового автомобиля в феврале по сравнению с январем.

2.35. Распределение автомобилей автотранспортного предприятия по величине суточного пробега за 25 сентября следующее (табл. 2.25).

Таблица 2.25

Суточный пробег автомобиля, км	До 160	160 – 180	180 – 200	200 и более
Число автомобилей	12	36	28	25

Определить средний суточный пробег одного автомобиля.

2.36. По двум цехам имеются следующие данные о распределении рабочих по уровню месячной заработной платы за апрель (табл. 2.26).

Таблица 2.26

Месячная заработная плата, руб.	Число рабочих		Месячная заработная плата, руб.	Число рабочих	
	цех № 1	цех № 2		цех № 1	цех № 2
4000 – 4200	32	17	4600 – 4800	70	110
4200 – 4400	36	40	4800 – 5000	32	83
4400 – 4600	150	220			

Определить, в каком цехе и на сколько процентов была выше средняя заработная плата рабочих.

2.37. По двум предприятиям фирмы имеются следующие данные о затратах на производство продукции (табл. 2.27).

Таблица 2.27

№ предприятия, входящего в фирму	Прошлый год		Отчетный год	
	доля затрат на оплату труда в общих затратах на производство, %	общие затраты на производство, млн руб.	затраты на оплату труда, млн руб.	доля затрат на оплату труда в общих затратах на производство, %
1	18,0	200	40,7	18,5
2	19,5	180	38,0	20,2

Определить изменение (в %) доли затрат на оплату труда в общих затратах на производство в целом по фирме в отчетном году по сравнению с прошлым годом.

2.38. Автобус на междугородной линии протяженностью 625 км прошел путь в прямом направлении со скоростью 68 км/ч, в обратном направлении – со скоростью 52 км/ч.

Определить среднюю скорость сообщения за оборотный рейс.

2.39. Имеются следующие данные по предприятиям фирмы (табл. 2.28).

Таблица 2.28

№ предприятия, входящего в фирму	I квартал		II квартал	
	выпуск продукции, тыс. руб.	средняя выработка на одного рабочего в день, руб.	отработано рабочими, человеко-дней	средняя выработка[1] на одного рабочего в день, руб.
1	59390,13	1540,6	79 200	1600,4
2	34246,10	1421,0	50 400	1500,0
3	72000,00	1600,0	90 300	1621,0

[1]Средняя выработка на одного рабочего в день определяется путем деления общей стоимости продукции на количество отработанных человеко-дней.

Определить:

1) среднюю выработку на одного рабочего в день в целом по фирме в I и II кварталах;

2) на сколько процентов изменилась средняя выработка на одного рабочего в день во II квартале по сравнению с I кварталом;

3) среднюю выработку на одного рабочего в день по фирме за первое полугодие.

2.40. Имеются данные о себестоимости транспортной работы по автотранспортным предприятиям объединения (табл. 2.29).

Определить, на сколько процентов изменилась средняя себестоимость 10 ткм по объединению в сентябре по сравнению с августом.

2.41. По металлургическому заводу имеются следующие данные об экспорте продукции (табл. 2.30).

Определить средний удельный вес продукции на экспорт.

2.42. Распределение промышленных предприятий региона по показателям затрат на 1 тыс. руб. продукции за два месяца следующее (табл. 2.31).

Таблица 2.29

№ предприятия, входящего в объединение	Август		Сентябрь	
	транспортная работа, тыс. ткм	себестоимость 10 ткм, руб.	общая сумма затрат на транспортную работу, тыс. руб.	себестоимость 10 ткм, руб.
1	20 800	5,12	10784,7	5,21
2	8 500	5,40	4609,6	5,36
3	30 000	4,97	14526,2	4,81

Таблица 2.30

Вид продукции	Стоимость всей реализованной продукции, тыс. руб.	Удельный вес продукции на экспорт, %
Чугун	68 200	35,5
Прокат листовой	75 100	22,8

Таблица 2.31

Затраты на 1 тыс. руб. продукции, руб.	Март		Апрель	
	число предприятий	общая стоимость продукции, тыс. руб.	число предприятий	средний объем продукции на одно предприятие, тыс. руб.
600 – 650	5	48 500	16	9 800
650 – 700	18	178 200	20	10 200
700 – 750	7	52 500	2	7 650

Определить:

1) изменение (в %) среднего размера затрат на 1 тыс. руб. продукции по предприятиям региона;

2) средний объем продукции на одно предприятие региона в марте и в апреле.

2.43. *Определить* средний удельный вес (в %) бракованной продукции за I квартал по данным табл. 2.32.

Таблица 2.32

Показатель	Январь	Февраль	Март
Выпуск годной продукции, млн руб.	80	96	100
Удельный вес бракованной продукции, %	5,0	3,2	3,8

2.44. Цехом произведены бракованные детали в трех партиях: в первой партии – 90 шт., что составило 3,0% общего числа деталей; во второй партии – 140 шт., или 2,8%; в третьей партии – 160 шт., или 2,0%.

Определить средний процент бракованных деталей.

2.45. По предприятию имеются следующие данные за два месяца (табл. 2.33).

Таблица 2.33

Категории работников	Апрель		Декабрь	
	численность работников	фонд заработной платы, руб.	средняя месячная заработная плата, руб.	фонд заработной платы, руб.
Рабочие	1 400	3 710 000	3800	5 358 000
Служащие	300	540 000	2780	750 600

Определить изменение (в %) среднего уровня месячной заработной платы рабочих и служащих, а также средней заработной платы всех работников предприятия в декабре по сравнению с апрелем.

2.46. *Определить* среднюю долю брака за год по следующим данным (табл. 2.34).

2.47. Предприятие реализует на рынке четыре вида изделий (табл. 2.35).

Таблица 2.34

Наименование продукции	Доля брака, %	Стоимость всей продукции, тыс. руб.
А	1,5	900
Б	2,0	1200
В	0,8	4000
Г	3,0	2700

Таблица 2.35

Наименование изделия	Физический объем продаж, тыс. шт.	Цена производителя за 1 шт., долл. США
А	24,7	110
Б	14,3	125
С	17,2	150
Д	4,3	250

Определить долю каждого вида изделий в выручке предприятия от их реализации.

2.48. По предприятиям объединения имеются следующие данные об удельных расходах топлива на производство теплоэнергии и общем расходе условного топлива за отчетный год (табл. 2.36).

Таблица 2.36

№ предприятия, входящего в объединение	Удельный расход условного топлива, кг/Гкал	Общий расход условного топлива, тыс. т
1	172,3	13,1
2	189,4	33,4
3	187,3	8,4

Определить средний по объединению удельный расход топлива на производство теплоэнергии.

2.49. Выпуск стального проката по сортам характеризуется следующими данными (табл. 2.37).

Таблица 2.37

Сорт стального проката	Отпускная цена за 1 т, тыс. руб.	Выпуск, т	
		по плану	фактически
I	1,3	21 000	24 050
II	0,97	7 260	6 800

Определить:

1) удельный вес продукции каждого сорта по плану и фактически;

2) среднюю плановую и фактическую цены за 1 т проката.

2.50. Отдел маркетинга фирмы организует реализацию производимого товара А по двум каналам распределения следующим образом (табл. 2.38).

Таблица 2.38

Показатель	Каналы распределения	
	1	2
Физический объем продаж, тыс. шт.	24,7	11,3
Цена производителя за единицу, у. е.	110,0	112,0
Себестоимость единицы, у. е.	48,0	48,0

Определить:

1) долю каждого канала распределения в общем объеме продаж;

2) долю каждого канала распределения в формировании прибыли фирмы.

2.51. Основные показатели здравоохранения в России на начало года представлены в табл. 2.39.

Таблица 2.39

Показатель	1998 г.	1999 г.	2000 г.
Численность врачей – всего, тыс. чел.	673	680	686
Число больничных коек, тыс. шт.	1761	1717	1668
Мощность амбулаторно-поликлинических учреждений, число посещений в смену, тыс. чел.	3475	3483	3490
Численность населения, млн чел.	146,7	146,3	145,6

Оценить обеспеченность населения медицинским обслуживанием, используя относительные величины интенсивности и динамики.

Сформулировать выводы.

2.52. Продажа грузовых автомобилей в России характеризуется следующими данными, тыс. шт. (табл. 2.40).

Таблица 2.40

Продажа	1997 г.	1998 г.	1999 г.
Промышленными предприятиями	148,3	151,0	178,0
Организациями оптовой торговли	3,6	12,1	21,9

Сопоставить среднегодовые темпы роста продаж грузовых автомобилей промышленными предприятиями и организациями оптовой торговли.

2.53. Данные о выпуске продукции по двум металлургическим предприятиям представлены в табл. 2.41.

Таблица 2.41

Вид продукции	Выплавлено, тыс. т	
	предприятие 1	предприятие 2
Чугун передельный	250	280
Чугун литейный	110	130
Чугун зеркальный	90	70

Определить, на сколько процентов отличается выпуск чугуна в условных тоннах на предприятии 2 по сравнению с предприятием 1, используя следующие коэффициенты пересчета: передельный чугун – 1,0; литейный чугун – 1,15; зеркальный чугун – 1,5.

2.54. По грузовому автотранспортному предприятию имеются следующие данные (табл. 2.42).

Таблица 2.42

№ автоколонны	Объем грузовых перевозок, тыс. т		
	2000 г.	2001 г.	
		план	отчет
1	520	570	576
2	380	390	385
3	540	600	620

Определить по каждой автоколонне и в целом по предприятию относительные величины:

1) планового задания;

2) выполнения плана за 2001 г.;

3) динамики объема перевозок в 2001 г. по сравнению с 2000 г.

2.55. Данные о среднегодовой численности работников, занятых в экономике России, представлены в табл. 2.43.

Таблица 2.43

Показатель	1997 г.	1998 г.	1999 г.
Всего занято в экономике, тыс. чел.	64639	63642	64500
в том числе по формам собственности:			
государственная и муниципальная	25897	24256	23865
собственность общественных организаций	411	445	450
частная	25752	27464	29250
смешанная российская (без иностранного участия)	11840	10473	9675
смешанная российская с иностранным участием	739	1004	1260

Определить показатели структуры численности работников, занятых в экономике, и представить их графически.

Сформулировать выводы.

2.56. Портфель состоит из акций трех компаний. Их доходность равна соответственно 15, 18 и 20 %, а доля в портфеле – 20, 50 и 30%.

Какова средняя доходность портфеля? Как следует увеличить доходность портфеля по сравнению с существующей?

Ответы к задачам для самостоятельной работы

2.20. Общий расход А – 6000 кг; Б – 4000 кг; общие затраты: 72 тыс. руб.; 0,5; 0,5. **2.21.** 1) 103,2; 2) 13,1; 29,8; 57,1. **2.22.** 1) 102,7%, 485 млн руб.; 200 млн руб.; 2) 100,4%. **2.23.** 0,98; 0,974. **2.24.** 102,6%. **2.25.** 714,4 тыс. руб. **2.26.** Увеличить на 1,7%. **2.27.** 59,7%. **2.28.** Больше на 6,04%. **2.29.** 2,0%. **2.30.** На одного жителя, кв. м: 18,5; 18,8; 19,0. Темпы роста с постоянной базой, %:

10,1,6; 102,7. **2.31.** 1) 104,05%; 2) 102,91%; 3) 107,07%. **2.32.** 1) Удельный вес безработных, %: всего – 9,9; 12,0; 13,0; мужчины – 10,2; 12,4; 12,9; женщины – 9,5; 11,7; 13,1. Темпы роста с постоянной базой, %: все население – 121,2; 131,3; мужчины – 121,6; 126,4; женщины – 123,2; 137,9. Относительные величины наглядности (мужчины по сравнению с женщинами): 1,07; 1,06; 0,98. **2.33.** 3,9 разряда; 3,8 разряда. **2.34.** Увеличение на 1,9%. **2.35.** 183,1 км. **2.36.** Во втором цехе на 1,43%. **2.37.** Увеличение на 3,1%. **2.38.** 59,0 км/ч. **2.39.** 1) 1538,6; 1585,8; 2) увеличение на 3,07%; 3) 1570,3. **2.40.** Снижение на 1,0%. **2.41.** 28,8%. **2.42.** 1) Снижение на 2,9%; 2) 9306,7 тыс. руб.; 9897,4 тыс. руб. **2.43.** 3,9%. **2.44.** 2,44%. **2.45.** 143,4%; 154,4%; 145,4%. **2.46.** 1,7%. **2.47.** 33,3%; 21,9%; 31,6%; 13,2%. **2.48.** 184,7 кг/Гкал. **2.49.** 1) 74,3 и 25,7%; 78,0 и 22,0%; 2) 1,215 тыс. руб.; 1,227 тыс. руб. **2.50.** 1) 68,6%; 31,4%; 2) 67,9%; 32,1%. **2.51.** Число врачей на 10000 населения: 45,9; 46,5; 47,1; изменение по сравнению с 1998 г., %: 101,3; 102,6; число больничных коек на 10000 населения: 120; 117,4; 114,6; изменение по сравнению с 1998 г., %: 97,8; 95,5; мощность лечебных учреждений (число посещений в смену) на 10000 населения: 236,7; 238,1; 239,7: изменение по сравнению с 1998 г., %: 100,6; 101,3. **2.52.** $\overline{T}_{p_1} = 109{,}6\%$; $\overline{T}_{p_2} = 246{,}6\%$. **2.53.** $y_1 = 511{,}5$ тыс. т.; $y_2 = 534{,}5$; увеличение на 4,5%. **2.54.** 1) плановое задание: +9,6%; +2,7%; +11,1%; в целом +8,3%; 2) выполнение плана: 101,1%; 98,7%; 103,3%; в целом 101,3%; 3) динамика, %: 110,8; 101,3; 114,8; в целом 109,8%. **2.55.** 1997 г, %: 40,1; 0,6; 39,9; 18,3; 1,1; 1998 г., %: 38,1; 0,7; 43,2; 16,4; 1,6; 1999 г., %: 37,0; 0,7; 45,3; 15,0; 2,0. **2.56.** 18,0%.

ГЛАВА 3

Статистические распределения и их основные характеристики

Ряды распределения и приемы их построения

Ряд распределения – это групповая таблица, имеющая две графы: группы по выделенному признаку (графа вариант) и численность групп (графа частот).

Ряды распределения делятся на вариационные и атрибутивные.

Вариационный ряд – групповая таблица, построенная по количественному признаку, в сказуемом которой показывается число единиц в каждой группе. В атрибутивных рядах представлены группировка по атрибутивным (качественным) признакам (например, деление рабочих предприятия по полу, профессиям и т. д.) и численность каждой группы.

Главное предназначение рядов распределения – изучение вариации признаков.

Различия индивидуальных значений признака у единиц совокупности называются **вариацией признака**. Она возникает в результате того, что индивидуальные значения складываются под совместным влиянием разнообразных условий (факторов), по-разному сочетающихся в каждом отдельном случае.

Вариация наблюдается и в пределах однородной, выделенной по тому или другому группировочному признаку, группы. Вариация, которая не зависит от факторов, положенных в основу выделения групп, называется **случайной вариацией.**

Изучение вариации в пределах однородной группы предполагает использование следующих приемов: построение ряда распределения, его графическое изображение, исчисление основных характеристик распределения.

Форма построения вариационного ряда зависит от характера изменения изучаемого признака, он может быть построен в форме дискретного ряда или в форме интервального ряда.

По характеру вариации значений признака различают:

- признаки с прерывным изменением (дискретные);
- признаки с непрерывным изменением (непрерывные).

Признаки с прерывным изменением могут принимать лишь конечное число определенных значений (например, тарифный разряд рабочих, число детей в семье, число станков, обслуживаемых одним рабочим). Признаки с непрерывным изменением могут принимать в определенных границах любые значения (например, стаж работы, пробег автомобиля, размер дохода и т. д.).

Для признака, имеющего прерывное изменение и принимающего небольшое количество значений, применяется построение **дискретного ряда.** В первой графе ряда указываются конкретные значения каждого индивидуального значения признака, во второй графе – численность единиц с определенным значением признака.

Для признака, имеющего непрерывное изменение, строится интервальный вариационный ряд, состоящий, так же как и дискретный ряд, из двух граф (варианты и частоты). При его построении в первой графе отдельные значения признака указываются в интервалах «от – до», во второй графе – число единиц, входящих в интервал. Интервалы образуются, как правило, равные и закрытые.

Величина интервала определяется по формуле

$$i = R / m,$$

где R – размах колебания (варьирования) признака;
$R = x_{max} - x_{min}$; x_{max}, x_{min} соответственно максимальное и минимальное значения признака в совокупности;
m – число групп.

Число групп приближенно определяется по формуле Стерджесса:

$$m = 1 + 3{,}322 \lg n,$$

где n – общее число единиц совокупности.

Полученную по этой формуле величину округляют до целого бо́льшего числа, поскольку количество групп не может быть дробным числом.

При небольшом объеме информации (численности единиц в совокупности) число групп может быть установлено исследователем без использования формулы Стерджесса.

Величину интервала обычно округляют до целого (всегда бо́льшего) числа, исключение составляют лишь случаи, когда изучаются малейшие колебания признака (например, при группировке деталей по величине размера отклонений от номинала, измеряемого в долях миллиметра).

Нижнюю границу первого интервала принимают равной минимальному значению признака (чаще всего его предварительно округляют до целого меньшего числа); верхняя граница первого интервала соответствует значению $(x_{min} + i)$. Для последующих групп границы определяются аналогично, т. е. последовательно прибавляется величина интервала. Если единица обладает значением признака, равным величине верхней границы интервала, то ее следует относить к следующей группе.

Примером интервального вариационного ряда служит табл. 3.1.

Таблица 3.1

Выполнение норм выработки рабочими цеха

Группы рабочих по выполнению норм выработки, %	Число рабочих
80 – 90	2
90 – 100	22
100 – 110	48
110 – 120	16
120 – 130	2
Итого	90

В каждой выделенной группе различают нижнюю и верхнюю границы интервала. Так, в последней группе рабочих по выполнению норм выработки (табл. 3.1) нижняя граница – 120%, верхняя – 130%.

При построении атрибутивных рядов число групп соответствует числу разновидностей признака.

Ряд распределения, состоящий из двух граф (варианты и частоты), иногда дополняется другими графами, необходимыми для вычисления отдельных статистических показателей или для более

отчетливого выражения характера вариации изучаемого признака. Достаточно часто в ряд вводится графа, в которой подсчитываются накопленные частоты (S). **Накопленные частоты** показывают, сколько единиц совокупности имеют значение признака не больше, чем данное значение, и исчисляются путем последовательного прибавления к частоте первого интервала частот последующих интервалов.

Частоты ряда (f) могут быть заменены **частостями** (w), которые представляют собой частоты, выраженные в относительных числах (долях или процентах) и рассчитанные путем деления частоты каждого интервала на их общую сумму, т. е.

$$w_1 = \frac{f_1}{\Sigma f}; \quad w_2 = \frac{f_2}{\Sigma f} \text{ и т. д.}$$

Замена частот частостями позволяет сопоставлять вариационные ряды с различным числом наблюдений. В табл. 3.2 по данным табл. 3.1 исчислены частости и накопленные частоты. Частости в долях исчислялись так:

$$\frac{2}{90} = 0{,}022; \quad \frac{22}{90} = 0{,}245 \text{ и т. д.}$$

Таблица 3.2

Выполнение норм выработки – рабочими цеха

Группы рабочих по выполнению норм выработки, (%), x	Число рабочих, f	Частости, w		Накопленная частота, S
		в долях	в %	
80 – 90	2	0,022	2,2	2
90 – 100	22	0,245	24,5	24
100 – 110	48	0,533	53,3	72
110 – 120	16	0,178	17,8	88
120 – 130	2	0,022	2,2	90
Итого	90	1,000	100,0	

Частости в процентах:

$$0{,}022 \cdot 100 = 2{,}2\%; \; 0{,}245 \cdot 100 = 24{,}5\% \text{ и т. д.}$$

Накопленные частоты:

$$2 + 22 = 24;\ 24 + 48 = 72;\ 72 + 16 = 88;\ 88 + 2 = 90.$$

Если вариационный ряд дан с неравными интервалами, то для правильного представления о характере распределения необходимо произвести расчет абсолютной или относительной плотности распределения.

Абсолютная плотность распределения (p) представляет собой величину частоты, приходящейся на единицу размера интервала отдельной группы ряда: $p = f / i$.

Относительная плотность распределения (p') – частное от деления частости (w) отдельной группы на размер ее интервала: $p' = w/i$.

Эти показатели используются для преобразования интервалов, что бывает необходимо при сравнительной оценке двух группировок. Для перегруппированных данных (с равными интервалами) частоты (или частости) для каждой вновь выделенной группы определяются по формуле:

$$f = \sum p_i \cdot i_i \text{ (или } \omega = \sum p'_i \cdot i_i),$$

где p_i – абсолютная плотность распределения i-й группы первоначальной группировки;

i_i – часть величины интервала новой группировки, приходящаяся на i-ю группу первоначальной группировки.

Первым этапом изучения вариационного ряда является его **графическое изображение.** Дискретный вариационный ряд изображается в виде так называемого **полигона, или многоугольника, распределения частот,** являющегося разновидностью статистических ломаных. Для изображения интервального ряда применяются **полигон распределения частот** и **гистограмма частот**.

Строятся графики в прямоугольной системе координат. При построении полигона частот на оси абсцисс в одинаковом масштабе откладываются направо в порядке возрастания значения признака (для дискретного характера) или центральные значения интервалов (для интервальных рядов); по оси ординат наносится шкала для выражения величин частот. Из точек на оси абсцисс, соответствующих величине признака, восстанавливаются перпендикуляры высотой, соответствующей частоте; вершины перпендикуляров соединяются отрезками прямой. Крайние точки полученной ломаной соединяются с лежащими на оси абсцисс следующими (меньшими и большими) возможными, но фактически не наблюдающимися значениями признака, частота которых, очевидно, равна 0. Замкнутая с осью абсцисс ломаная линия представляет полигон распределения частот.

Для построения гистограммы по оси абсцисс откладывают величины интервалов, а частоты изображаются прямоугольниками, построенными на интервалах с высотой в масштабе оси ординат.

В случае неравенства интервалов гистограмма строится не по частотам или частостям, а по плотности распределения.

В ряде случаев для изображения вариационных рядов используется **кумулятивная кривая (кумулята)**, она особенно удобна для сравнения вариационных рядов. Накопленные частоты наносятся на чертеж в виде ординат; соединяя вершины отдельных ординат прямыми, получают ломаную линию, которая, начиная с нуля, непрерывно поднимается над осью абсцисс до тех пор пока не достигнет высоты, соответствующей общей сумме частот.

Если поменять местами оси координат в кумуляте, то получаем новый вид графического изображения – **огиву.**

При изучении процессов концентрации (концентрации производства, концентрации капитала и др.) используется графическое изображение вариационного ряда в виде кривой Лоренца. Для ее построения абсолютные показатели числа единиц в группах и размер изучаемого признака выражаются в относительных показателях (в долях или процентах к итогу) и исчисляются их накопленные значения. При построении графика на горизонтальной линии наносится шкала для ряда накопленных частостей, а на вертикальной линии – шкала для накопленных относительных величин размера изучаемого признака. Далее наносятся точки в соответствии с накопленными значениями двух рядов. Соединив все точки прямыми линиями, получают кривую, характеризующую степень неравномерности распределения. Линия, соединяющая нижний левый угол графика с верхним правым (диагональ четырехугольника), является линией равномерного распределения. Чем больше кривая отличается от диагонали, тем больше неравномерность.

На основе графика можно рассчитать коэффициент концентрации (индекс Джини):

$$K_к = S_0 \cdot S_1,$$

где S_0 – площадь между линией равномерного и фактического распределения;

S_1 – площадь треугольника, образуемого линией равномерного распределения и горизонтальной линией графика (соответствует половине площади четырехугольника).

Величина индекса изменяется в пределах от 0 до 1; для равномерного распределения она равна 0; чем больше степень концентрации, тем больше величина индекса.

При построении графических изображений вариационного ряда большое значение имеет соотношение масштабов по оси абсцисс (x) и оси ординат (f). В этом случае следует руководствоваться так называемым «правилом золотого сечения», в соответствии с которым высота графика должна быть примерно в 1,5 раза меньше его основания.

Для анализа вариационных рядов используются три группы показателей:

- показатели центра распределения;
- показатели степени вариации;
- показатели формы распределения.

Показатели центра распределения

Для характеристики среднего значения признака в вариационном ряду применяются: средняя арифметическая, медиана, мода.

Средняя арифметическая для дискретного ряда распределения исчисляется по формуле

$$\bar{x} = \frac{\Sigma xf}{\Sigma f},$$

где x – варианты значений признака;

f – частота повторения данного варианта.

Средняя арифметическая для интервального ряда распределения:

$$\bar{x} = \frac{\Sigma x'f}{\Sigma f},$$

где x' – середина соответствующего интервала значения признака; вычисляется как средняя из значений границ интервала.

Медиана (Ме) соответствует варианту, стоящему в середине ранжированного ряда. Положение медианы определяется ее номером:

$$N_{\text{Me}} = \frac{n+1}{2},$$

где n – число единиц в совокупности.

По накопленным частотам определяют ее численное значение в дискретном вариационном ряду.

Если совокупность содержит четное число значений варьирующего признака ($n = 2к$; $к = n/2$), то в этом случае за медиану условно принимают значение:

$$\text{Me} = \frac{1}{2}(x_k + x_{k+1}),$$

так как в ряду нет члена, который делил бы совокупность на две равные по объему группы.

В интервальном ряду распределения сначала указывают интервал, в котором находится медиана.

Медианным является первый интервал, в котором сумма накопленных частот превысит половину общего числа наблюдений.

Численное значение медианы определяется по формуле

$$\text{Me} = x_{\text{Me}} + i\frac{\frac{(n+1)}{2} - S_{(\text{Me}-1)}}{f_{\text{Me}}},$$

где x_{Me} – нижняя граница медианного интервала;

i – величина интервала;

$S_{(\text{Me}-1)}$ – накопленная частота интервала, предшествующего медианному;

f_{Me} – частота медианного интервала.

Мода (Мо) – наиболее часто встречающееся значение признака. В дискретном ряду – это варианта с наибольшей частотой. В интервальном ряду сначала определяется модальный интервал, т. е. тот интервал, который имеет наибольшую частоту.

Конкретное значение моды определяется по формуле:

$$\text{Mo} = x_{\text{Mo}} + i\frac{f_{\text{Mo}} - f_{(\text{Mo}-1)}}{[f_{\text{Mo}} - f_{(\text{Mo}-1)}] + [f_{\text{Mo}} - f_{(\text{Mo}+1)}]},$$

где x_{Mo} – нижняя граница модального интервала;

f_{Mo} – частота модального интервала;

$f_{(\text{Mo}-1)}$ – частота интервала, предшествующего модальному;

$f_{(\text{Mo}+1)}$ – частота интервала, следующего за модальным.

Моду и медиану можно определить на основе графического изображения ряда. Медиана определяется по кумуляте. Для ее определения высоту наибольшей ординаты, которая соответствует общей численности, делят пополам. Через полученную точку проводят прямую, параллельную оси абсцисс, до пересечения ее с кумулятой. Абсцисса точки пересечения является медианной величиной.

Мода определяется по гистограмме распределения. Для этого правую вершину модального прямоугольника соединяют с правым верхним углом предыдущего прямоугольника, а левую вершину модального прямоугольника – с левым верхним углом последующего прямоугольника. Абсцисса точки пересечения этих прямых и будет модой распределения.

Показатели вариации (колеблемости) признака

Для характеристики размера вариации признака используются абсолютные и относительные показатели. К абсолютным показателям вариации относятся:

- размах колебаний;
- среднее линейное отклонение;
- среднее квадратическое отклонение;
- дисперсия;
- квартильное отклонение.

Размах колебаний (размах вариации)

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

где $x_{\max}$, $x_{\min}$ – соответственно максимальное и минимальное значения признака.

Величина показателя зависит от величины только двух крайних вариант и не учитывает степени колеблемости основной массы членов ряда.

Среднее линейное отклонение ($\bar{d}$) и среднее квадратическое отклонение (σ) показывают, на сколько в среднем отличаются индивидуальные значения признака от среднего его значения.

Среднее линейное отклонение определяется по формулам:

а) для несгруппированных данных (первичного ряда)

$$\bar{d} = \frac{\sum|x - \bar{x}|}{n};$$

б) для вариационного ряда

$$\bar{d} = \frac{\sum|x - \bar{x}|f}{\sum f}.$$

Среднее квадратическое отклонение (σ) **и дисперсия** (σ^2) определяются так:

а) для несгруппированных данных

$$\sigma = \sqrt{\frac{\sum(x - \bar{x})^2}{n}}; \quad \sigma^2 = \frac{\sum(x - \bar{x})^2}{n};$$

б) для вариационного ряда

$$\sigma = \sqrt{\frac{\sum(x - \bar{x})^2 f}{\sum f}}; \quad \sigma^2 = \frac{\sum(x - \bar{x})^2 f}{\sum f}.$$

Формула для расчета дисперсии может быть преобразована:

$$\sigma^2 = \frac{\sum_{i=1}^{n}(x_i - \bar{x})^2}{n} = \frac{\sum_{i=1}^{n}[x_i^2 - 2x_i\bar{x} + (\bar{x})^2]}{n} =$$

$$\frac{\sum_{i=1}^{n}x_i^2 - 2\bar{x}\sum_{i=1}^{n}x_i + n(\bar{x})^2}{n} = \frac{\sum_{i=1}^{n}x_i^2}{n} - (\bar{x})^2,$$

т. е. дисперсия равна средней из квадратов индивидуальных значений признака минус квадрат средней величины. Следовательно,

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

Среднее квадратическое отклонение по своей величине всегда превышает значение среднего линейного отклонения в соответствии со **свойством мажорантности средних.**

Квартильное отклонение (d_k) применяется вместо размаха вариации, чтобы избежать недостатков, связанных с использованием крайних значений:

$$d_k = \frac{Q_3 - Q_1}{2},$$

где Q_3 и Q_1 – соответственно третья и первая квартили распределения.

Квартиль – значения признака, которые делят ранжированный ряд на четыре равные по численности части. Таких величин будет три: первая квартиль (Q_1), вторая квартиль (Q_2), третья квартиль (Q_3). Вторая квартиль является медианой. Вычисление квартилей аналогично вычислению медианы.

Сначала определяют положение или место квартили:

$$N_{Q_1} = \frac{n+1}{4}; \quad N_{Q_2} = \frac{n+1}{4} \cdot 2 = \frac{n+1}{2}; \quad N_{Q_3} = \frac{n+1}{4} \cdot 3.$$

Затем по накопленным частотам в дискретном ряду определяют численное значение.

В интервальном ряду распределения сначала указывают интервал, в котором лежит квартиль, затем определяют ее численное значение по формуле

$$Q = x_Q + i\frac{N_Q - S_{(Q-1)}}{f_Q},$$

где x_Q – нижняя граница интервала, в котором находится квартиль;

$S_{(Q-1)}$ – накопленная частота интервала, предшествующего тому, в котором находится квартиль;

f_Q – частота интервала, в котором находится квартиль.

При сравнении колеблемости различных признаков в одной и той же совокупности или же при сравнении колеблемости одного и того же признака в нескольких совокупностях с различной величиной средней арифметической используются **относительные показатели вариации.** Они вычисляются как отношение абсолютных показателей вариации к средней арифметической (или медиане) и чаще всего выражаются в процентах.

Формулы расчета относительных показателей вариации следующие:

коэффициент осцилляции $K_R = \frac{R}{\bar{x}} \cdot 100\%;$

относительное линейное отклонение $K_{\bar{d}} = \frac{\bar{d}}{\bar{x}} \cdot 100\%;$

коэффициент вариации $v = \frac{\sigma}{\bar{x}} \cdot 100\%;$

относительный показатель квартильной вариации

$$K_{d_k} = \frac{d_k}{\text{Me}} \cdot 100\%, \quad \text{или} \quad K_Q = \frac{Q_3 - Q_1}{2Q_2} \cdot 100\%.$$

Наиболее часто применяется **коэффициент вариации.** Его применяют не только для сравнительной оценки вариации, но и для характеристики однородности совокупности. Совокупность считается однородной, если коэффициент вариации не превышает 33% (для распределений, близких к нормальному).

Сложение дисперсий изучаемого признака

Изучая дисперсию интересующего нас признака в пределах исследуемой совокупности и опираясь на общую среднюю в расчетах, нельзя оценить влияние отдельных факторов, определяющих колеблемость индивидуальных значений (вариант) признака. Это можно сделать при помощи метода группировок, когда единицы изучаемой совокупности подразделяются на однородные группы по признаку-фактору. При этом кроме общей средней для всей совокупности исчисляются средние по отдельным группам (групповые или частные средние) и *три показателя дисперсии:*

- общая дисперсия;
- межгрупповая дисперсия;
- средняя внутригрупповая дисперсия.

Величина общей дисперсии (σ_o^2) характеризует вариацию признака под влиянием всех факторов, формирующих уровень признака у единиц данной совокупности, и определяется по формуле

$$\sigma_0^2 = \frac{\sum(x-\bar{x}_0)^2 f}{\sum f},$$

где $\bar{x}_0$ – общая средняя арифметическая для всей изучаемой совокупности.

Межгрупповая дисперсия (дисперсия групповых средних δ^2) отражает систематическую вариацию, т. е. те различия в величине изучаемого признака, которые возникают под влиянием фактора, положенного в основу группировки. Межгрупповая дисперсия определяется по формуле

$$\delta^2 = \frac{\sum(\bar{x}_i-\bar{x}_0)^2 n_i}{\sum n_i},$$

где $\bar{x}_i$ – средняя по отдельной группе;

n_i – число единиц в определенной группе.

Средняя внутригрупповая дисперсия ($\overline{\sigma^2}$) характеризует случайную вариацию, возникающую под влиянием других, неучтенных факторов, и не зависит от условия (признака-фактора), положенного в основу группировки.

Средняя внутригрупповая дисперсия определяется по формуле

$$\overline{\sigma^2} = \frac{\sum\sigma_i^2 n_i}{\sum n_i} \quad \text{или} \quad \overline{\sigma^2} = \frac{\sum\sum(x-\bar{x}_i)^2 f}{\sum f},$$

где σ_i^2 – дисперсия по отдельной группе;

$$\sigma_i^2 = \frac{\sum(x-\bar{x}_i)^2 f}{\sum f}.$$

Указанные дисперсии взаимосвязаны между собой следующим равенством: величина общей дисперсии равна сумме межгрупповой дисперсии и средней внутригрупповой дисперсии:

$$\sigma^2_o = \delta^2 + \overline{\sigma^2}.$$

Это тождество отражает закон (правило) сложения дисперсий. Опираясь на это правило, можно определить, какая часть (доля) общей дисперсии складывается под влиянием признака-фактора, положенного в основу группировки.

Вариации альтернативного признака

Альтернативный признак – качественный признак, имеющий две взаимоисключающие разновидности (например, работники предприятия подразделяются на мужчин и женщин; продукция – на годную и бракованную и т. д.).

Альтернативный признак принимает всего два значения:

1 – наличие признака;

0 – отсутствие признака.

$$p + q = 1,$$

где p – доли единиц, обладающих признаком;

q – доли единиц, не обладающих признаком.

Среднее значение альтернативного признака

$$\bar{x} = \frac{(1\cdot p)+(0\cdot q)}{p+q} = p.$$

Дисперсия альтернативного признака

$$\sigma^2 = \frac{(1-p)^2\cdot p + (0-p)^2\cdot q}{p+q} = p\cdot q.$$

Предельное значение вариации альтернативного признака равно 0,25; оно получается при $p = q = 0{,}5$.

Показатели формы распределения

Для получения приблизительного представления о форме распределения строят графики распределения (полигон и гистограмму). В практике статистических исследований приходится встречаться с самыми различными распределениями. Однородные совокупности характеризуются, как правило, одновершинными распределениями. Многовершинность свидетельствует о неоднородности изучаемой совокупности. В этом случае необходима перегруппировка данных с целью выделения более однородных групп.

Обобщающие характеристики (показатели) центра распределения и степени вариации на дают представления о форме распределения, так как не вскрывают характера изменения частот.

Для выражения особенностей формы распределения применяются ранговые характеристики, показатели дифференциации, асимметрии и эксцесса, кривые распределения.

Ранговые характеристики – варианты, занимающие в ранжированном вариационном ряду определенное место. К их числу относятся квартили (Q), децили (D), перцентили (P).

Расчет квартилей и их практическое использование даны при рассмотрении показателей вариации.

Децили – значения признака, которые делят ранжированный ряд на десять равных по численности частей.

Перцентили – значения признака, делящие ранжированный ряд на 100 равных частей.

Расчет децилей и перцентилей выполняется аналогично исчислению квартилей.

Так, при расчете децилей сначала определяют место девяти децилей:

$$N_{D_1} = \frac{n+1}{10};\ N_{D_2} = \frac{2(n+1)}{10};\ldots\ N_{D_9} = \frac{9(n+1)}{10},$$

где n – общее число единиц в совокупности.

В дискретном ряду по накопленным частотам определяют численные значения децилей.

В интервальном ряду сначала определяют интервал, в котором лежит дециль. Ее численное значение определяют по формуле:

$$D = X_D + i\frac{N_D - S_{D-1}}{f_D},$$

где X_D – нижняя граница интервала, в котором находится дециль;
i – величина интервала;
N_D – место децили;
S_{D-1} – накопленная частота интервала, предшествующего тому, в котором находится дециль;
f_D – частота интервала, в котором находится дециль.

Анализ вариационного ряда дополняется расчетом показателя дифференциации.

По ряду распределения определяется коэффициент децильной дифференциации по формуле:

$$K_D = \frac{D_9}{D_1},$$

где D_9 – девятая дециль;
D_1 – первая дециль.

Он показывает, во сколько раз наименьший уровень признака из 10% единиц, имеющих наибольший уровень признака, больше наибольшего уровня признака, из 10% единиц совокупности, имеющих наименьший уровень признака.

По первичным данным исчисляется коэффициент фондовой дифференциации по формуле:

$$K_Ф = \frac{\overline{X}_{наиб}}{\overline{X}_{наим}},$$

где $\overline{X}_{наиб}$ – средний уровень признака из 10% наибольших значений признака;
$\overline{X}_{наим}$ – средний уровень признака из 10% наименьших значений признака.

Для сравнительного анализа степени асимметрии нескольких распределений рассчитывается относительный показатель асимметрии (A_s):

$$A_s = \frac{\overline{x} - Mo}{\sigma}.$$

Величина показателя асимметрии A_s может быть положительной и отрицательной. Положительная величина показателя асимметрии указывает на наличие правосторонней асимметрии. Отрицательный знак показателя асимметрии говорит о наличии левосторонней асимметрии. Чем больше абсолютная величина коэффициента, тем больше степень скошенности. Принято считать, что если коэффициент асимметрии меньше 0,25, то асимметрия незначительная, если свыше 0,5, то асимметрия значительная.

Другой показатель асимметрии, предложенный шведским математиком Линдбергом, исчисляется по формуле

$$A_s = П - 50,$$

где П – процент тех значений признака, которые превышают величину средней арифметической;
50 – процент вариант, превосходящих среднюю арифметическую ряда нормального распределения.

6–1622

Наиболее распространенным является показатель асимметрии, исчисляемый по формуле

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

где μ_3 – центральный момент третьего порядка;

$$\mu_3 = \frac{\sum (x-\bar{x})^3 \cdot f}{\sum f}.$$

Этот показатель асимметрии не только определяет степень асимметрии, но и указывает на наличие или отсутствие асимметрии в распределении признака в генеральной совокупности. Оценка степени существенности этого показателя дается с помощью средней квадратической ошибки, рассчитываемой по формуле

$$\sigma_{A_s} = \sqrt{\frac{6 \cdot (n-1)}{(n+1) \cdot (n+3)}},$$

где n – число наблюдений.

Если $\frac{|A_s|}{\sigma_{A_s}} > 3$, асимметрия существенна и распределение признака в генеральной совокупности не является симметричным.

Если $\frac{|A_s|}{\sigma_{A_s}} < 3$, асимметрия несущественна, ее наличие объясняется влиянием случайных обстоятельств.

Для симметричных распределений рассчитывается показатель эксцесса (островершинности):

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

где μ_4 – центральный момент четвертого порядка.

$$\mu_4 = \frac{\Sigma (x-\bar{x})^4 \cdot f}{\sum f}.$$

Эксцесс может быть положительным и отрицательным. У высоковершинных распределений показатель эксцесса имеет положительный знак (+), а у низковершинных – отрицательный знак (–). Предельным значением отрицательного эксцесса является значение $E_x = -2$; величина положительного эксцесса является величиной бесконечной. В нормальном распределении $\frac{\mu_4}{\sigma^4} = 3$. Следовательно, для нормального закона $E_x = 0$.

Средняя квадратическая ошибка эксцесса исчисляется по формуле

$$\sigma_{E_x} = \sqrt{\frac{24 \cdot n \cdot (n-2) \cdot (n-3)}{(n-1)^2 \cdot (n+3) \cdot (n+5)}},$$

где n – число наблюдений.

Для приближенного определения величины эксцесса может быть использована **формула Линдберга:**

$$E_x = \Pi - 38{,}29,$$

где П – процент количества вариант, лежащих в интервале, равном половине среднего квадратического отклонения (в ту и другую сторону от величины средней);

38, 29 – процент количества вариант, лежащих в интервале, равном половине среднего квадратического отклонения, в общем количестве вариант ряда нормального распределения.

Кривые распределения

Наиболее надежный путь выявления закономерностей распределения – увеличение количества наблюдений. По мере увеличения количества наблюдений (в пределах той же однородной совокупности) при одновременном уменьшении величины интервала закономерность, характерная для данного распределения, будет выступать все более и более ясно, а представляющая полигон частот ломаная линия будет приближаться к некоторой плавной линии и в пределе должна превратиться в кривую линию.

Кривая линия, которая отражает закономерность изменения частот в чистом, исключающем влияние случайных факторов виде, называется **кривой распределения.**

В настоящее время изучено значительное число различных форм распределений. В практике статистических исследований часто используется *распределение Пуассона, Максвелла*, особенно *нормальное распределение*. Распределения, близкие к нормальному распределению, были обнаружены при изучении самых различных явлений как в природе, так и в развитии общества.

В статистической практике большой интерес представляет решение вопроса о том, в какой мере можно считать полученное в результате статистического наблюдения распределение признака в исследуемой совокупности, соответствующее нормальному распределению.

Для решения этого вопроса следует рассчитать теоретические частоты нормального распределения, т. е. те частоты, которые были бы, если бы данное распределение в точности следовало закону нормального распределения. Для расчета теоретических частот применяется следующая формула:

$$f_t' = \frac{n \cdot i}{\sigma} \cdot \frac{1}{\sqrt{2\pi}} \cdot e^{-\frac{t^2}{2}},$$

где t – нормированное отклонение.

$$t = \frac{x - \bar{x}}{\sigma}.$$

Величина $\frac{1}{\sqrt{2\pi}} \cdot e^{-\frac{t^2}{2}}$ определяется по специальной таблице (см. приложение 1).

Следовательно, в зависимости от величины t для каждого интервала эмпирического ряда определяются теоретические частоты.

Для проверки близости теоретического и эмпирического распределений используются специальные показатели, называемые **критериями согласия.** Наиболее распространенным является критерий согласия **К. Пирсона** χ^2 («хи – квадрат»), исчисляемый по формуле

$$\chi^2 = \sum \frac{(f - f')^2}{f'},$$

где f – эмпирические частоты (частости) в интервале;
f' – теоретические частоты (частости) в интервале.

Полученное значение критерия ($\chi^2_{\text{расч}}$) сравнивается с табличным значением ($\chi^2_{\text{табл}}$). Последнее определяется по специальной таблице (см. приложение 2) в зависимости от принятой вероятности (P) и числа степеней свободы k (для нормального распределения k равно числу групп в ряду распределения минус 3).

Если $\chi^2_{\text{расч}} \le \chi^2_{\text{табл}}$, то гипотеза о близости эмпирического распределения к нормальному не отвергается.

При расчете критерия Пирсона необходимо соблюдать условия: число наблюдений должно быть достаточно велико ($n \ge 50$); если теоретические частоты в некоторых интервалах меньше 5, то интервалы объединяют так, чтобы частоты были больше 5.

Используя величину χ^2, **В. И. Романовский** предложил оценивать близость эмпирического распределения кривой нормального распределения по отношению

$$\frac{\chi^2 - (m-3)}{\sqrt{2 \cdot (m-3)}},$$

где m – число групп;
$m-3$ – число степеней свободы при исчислении частот нормального распределения.

Если $\frac{\chi^2 - (m-3)}{\sqrt{2 \cdot (m-3)}} < 3$, то можно принять гипотезу о нормальном характере эмпирического распределения.

Распространенным критерием согласия является критерий **А. Н. Колмогорова:**

$$\lambda = \frac{D}{\sqrt{n}},$$

где D – максимальное значение разности между накопленными эмпирическими и теоретическими частотами;
n – сумма эмпирических частот.

По таблице значений вероятностей λ-критерия находят соответствующую вероятность (P). Если найденной величине λ соответствует значительная по величине вероятность (P), то расхождения между эмпирическим и теоретическим распределениями несущественны.

Практическое и научное значение имеет распределение **Пуассона.** Оно характерно для редко встречающихся явлений, поэтому его называют «законом редких явлений» (или «законом малых чисел»).

Закон Пуассона применяется для совокупностей, достаточно больших по объему ($n \geq 100$) и имеющих достаточно малую долю единиц, обладающих данным признаком ($p \leq 0,1$), например для распределения партий готовой продукции по числу забракованных изделий, печатных страниц по числу опечаток, станков по числу отказов, ткацких станков по числу обрывов нити и т. д.

Теоретические частоты распределения Пуассона определяются формулой

$$f'_m = n \cdot \frac{\lambda^m \cdot e^{-\lambda}}{m!},$$

где n — общее число независимых испытаний;
λ — среднее число появления редкого события в n одинаковых независимых испытаниях;
m — частота данного события ($m = 0, 1, 2$ ...);
e — основание натуральных логарифмов, e = 2,71828.

Величина $e^{-\lambda}$ определяется по специальной таблице (приложение 8); $m!$ — произведение $1 \cdot 2 \cdot 3 ... \cdot m$; $0!$ — считается равным единице.

Степень расхождения теоретических и эмпирических частот оценивается с помощью критериев согласия.

3.1. Решение типовых задач

3.1. По приведенным ниже данным о квалификации рабочих цеха *требуется*:

1) построить дискретный ряд распределения;

2) дать графическое изображение ряда;

3) вычислить показатели центра распределения, показатели вариации и формы распределения.

Тарифные разряды 24 рабочих цеха: 4; 3; 6; 4; 4; 2; 3; 5; 4; 4; 5; 2; 3; 4; 4; 5; 2; 3; 6; 5; 4; 2; 4; 3.

Решение

1. Дискретный ряд распределения имеет вид (табл. 3.3).

Таблица 3.3

Распределение рабочих цеха по квалификации

Тарифный разряд, x	Число рабочих, f	Накопленная частота, S
2	4	4
3	5	9
4	9	18
5	4	22
6	2	24
Итого	24	–

2. На рис. 3.1 представлено графическое изображение построенного дискретного вариационного ряда в виде полигона частот.

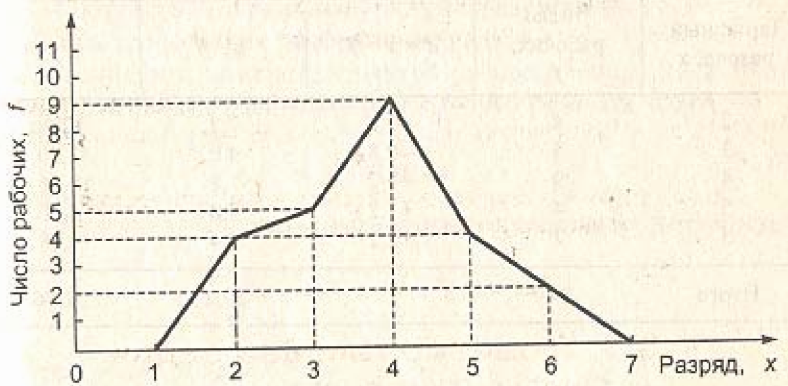

Рис. 3.1. Полигон распределения рабочих цеха по квалификации

Полигон частот замыкается, для этого крайние вершины соединяются с точками на оси абсцисс, отстоящими на одно деление в принятом масштабе (в данном случае $x = 1$ и $x = 7$).

3. К показателям центра распределения относятся: средняя арифметическая, мода и медиана.

Средняя арифметическая –

$$\bar{x} = \frac{\sum x \cdot f}{\sum f} = \frac{2 \cdot 4 + 3 \cdot 5 + 1 \cdot 9 + 5 \cdot 4 + 6 \cdot 2}{4 + 5 + 9 + 4 + 2} = \frac{91}{24} = 3{,}8 \text{ разряда.}$$

Мо = 4-му разряду (4-й разряд встречается 9 раз, т. е. это наибольшая частота).

$$N_{\text{Me}} = \frac{24+1}{2} = 12{,}5;$$

Ме = 4-му разряду (так как номера 12 и 13 соответствуют 4-му разряду).

К показателям вариации относятся: среднее линейное отклонение ($\bar{d}$), среднее квадратическое отклонение (σ), коэффициент вариации (V). Для расчета показателей ряда распределения удобно использовать вспомогательную табл. 3.4.

Таблица 3.4

Расчет показателей вариации

Тарифный разряд, x	Число рабочих, f	$d = x - \bar{x}$	$\lvert d \rvert \cdot f$	$d^2 \cdot f$
2	4	–1,8	7,2	12,96
3	5	–0,8	4,0	3,20
4	9	+0,2	1,8	0,36
5	4	+1,2	4,8	5,76
6	2	+2,2	4,4	9,68
Итого	24		22,2	31,96

$$\bar{d} = \frac{\sum \lvert x - \bar{x} \rvert f}{\sum f} = \frac{\sum \lvert d \rvert \cdot f}{\sum f} = \frac{22{,}2}{24} = 0{,}9 \text{ разряда;}$$

$$\sigma = \sqrt{\frac{\sum (x - \bar{x})^2 f}{\sum f}} = \sqrt{\frac{\sum d^2 f}{\sum f}} = \sqrt{\frac{31{,}96}{24}} = 1{,}15 \text{ разряда;}$$

$$V = \frac{\sigma}{\bar{x}} \cdot 100 = \frac{1{,}15}{3{,}8} \cdot 100 = 30{,}3\%.$$

Следовательно, индивидуальные значения отличаются в среднем от средней арифметической на 1,15 разряда, или на 30,3%.

Среднее квадратическое отклонение превышает среднее линейное отклонение в соответствии со свойствами мажорантности средних.

Значение коэффициента вариации (30,3%) свидетельствует о том, что совокупность достаточно однородна.

Как видно на рис. 3.1, распределение рабочих по тарифному разряду несимметрично, поэтому определяется показатель асимметрии:

$$A_s = \frac{\bar{x} - \text{Mo}}{\sigma} = \frac{3{,}8 - 4}{1{,}15} = -0{,}17.$$

Следовательно, асимметрия левосторонняя, незначительная.

3.2. Имеются следующие данные о возрастном составе рабочих цеха (лет): 18; 38; 28; 29; 26; 38; 34; 22; 28; 30; 22; 23; 35; 33; 27; 24; 30; 32; 28; 25; 29; 26; 31; 24; 29; 27; 32; 25; 29; 29.

Для анализа распределения рабочих цеха по возрасту *требуется*:

1) построить интервальный ряд распределения;

2) дать графическое изображение ряда;

3) исчислить показатели центра распределения, показатели вариации и формы распределения. Сформулировать вывод.

Решение

1. Величина интервала группировки определяется по формуле

$$i = \frac{R}{m} = \frac{38-18}{7} = \frac{20}{7} = 2{,}85 \approx 3 \text{ года,}$$

где m принимаем равным 7.

Интервальный ряд представлен в табл. 3.5.

2. Графически интервальный вариационный ряд может быть представлен в виде гистограммы, полигона, кумуляты.

Гистограмма строится в прямоугольной системе координат. По оси абсцисс откладывают интервалы значений вариационного признака, причем число интервалов целесообразно увеличить на два (по одному в начале и в конце имеющегося ряда) для удобства преобразования гистограммы в полигон частот. На отрезках (интервалах) строятся прямоугольники, высота которых соответствует частоте.

Таблица 3.5

Интервальный ряд распределения

Группы рабочих по возрасту (лет), x	Число рабочих, f	Накопленная частота, S
18 – 21	1	1
21 – 24	3	4
24 – 27	6	10
27 – 30	10	20
30 – 33	5	25
33 – 36	3	28
36 – 39	2	30
Итого	30	

Для преобразования гистограммы в полигон частот середины верхних сторон прямоугольников соединяют отрезками прямой, и две крайние точки прямоугольников замыкаются по оси абсцисс на середине интервалов, в которых частоты равны нулю.

На рис. 3.2 представлено графическое изображение построенного интервального вариационного ряда в виде гистограммы и полигона частот.

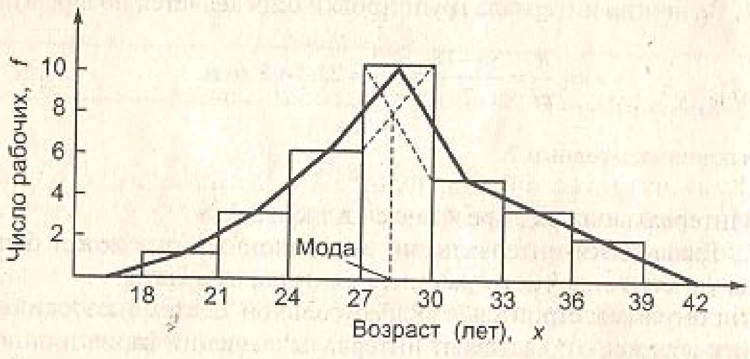

Рис. 3.2. Гистограмма и полигон распределения рабочих цеха по возрасту

Как видно из графика, треугольники, относящиеся к площади гистограммы и к площади полигона, попарно равны между собой, и, следовательно, площадь гистограммы и площадь полигона данного вариационного ряда также совпадают.

На основе построенной гистограммы графически можно определить значение моды. Для этого правую вершину модального прямоугольника соединяют прямой с правым верхним углом предыдущего прямоугольника, а левую вершину модального прямоугольника соединяют с левым верхним углом последующего прямоугольника. Абсцисса точки пересечения этих прямых и будет модой распределения. Мо = 28,3 года. На рис. 3.2 эти прямые линии, соединяющие вершины прямоугольников, и перпендикуляр из точки их пересечения показаны пунктирной линией.

На рис. 3.3 представлена кумулятивная кривая (кумулята).

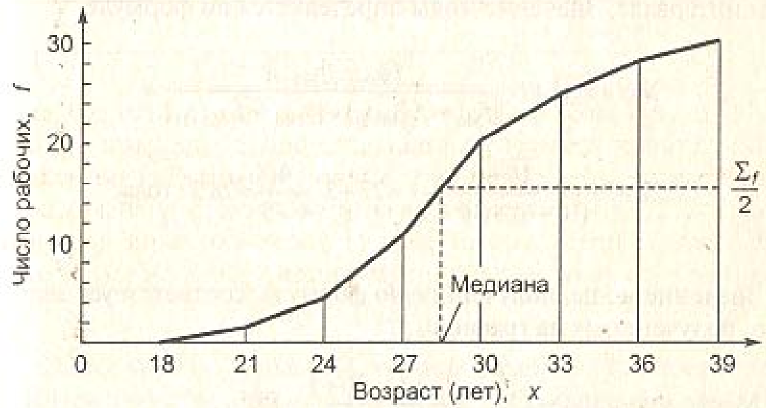

Рис. 3.3. Кумулята распределения рабочих цеха по возрасту

Кумулята может быть использована для графического определения медианы. Для этого последнюю ординату кумуляты делят пополам. Через полученную точку проводят прямую, параллельную оси x, до пересечения ее с кумулятой. Из точки пересечения опускается перпендикуляр до оси абсцисс. Абсцисса точки пересечения является медианой. Линии, определяющие медиану, на рис. 3.3 показаны пунктирными линиями. Ме = 28,6 года.

3. Расчет показателей центра распределения:

$$\bar{x} = \frac{\sum x'f}{\sum f} = \frac{19{,}5 \cdot 1 + 22{,}5 \cdot 3 + 25{,}5 \cdot 6 + 28{,}5 \cdot 10 + 31{,}5 \cdot 5 + 34{,}5 \cdot 3 + 37{,}5 \cdot 2}{30} =$$

$$= \frac{861}{30} = 28{,}7 \text{ года,}$$

где x' – среднее значение признака в интервале (центр интервала).

Для интервального вариационного ряда порядок расчета структурных средних следующий: сначала находят интервал, содержащий моду или медиану, а затем рассчитывают соответствующие значения названных показателей.

Модальным в данном распределении является интервал 27 – 30 лет, так как наибольшее число рабочих ($f = 10$) находится в этом интервале. Значение моды определяется по формуле

$$\text{Mo} = x_{\text{Mo}} + i \cdot \frac{f_{\text{Mo}} - f_{(\text{Mo}-1)}}{[f_{\text{Mo}} - f_{(\text{Mo}-1)}] + [f_{\text{Mo}} - f_{(\text{Mo}+1)}]} =$$

$$= 27 + 3 \cdot \frac{10 - 6}{(10 - 6) + (10 - 5)} = 27 + 3 \cdot \frac{4}{4 + 5} = 28{,}3 \text{ года.}$$

Значение моды, полученное по формуле, соответствует значению, полученному на графике.

Место медианы – $N_{\text{Me}} = \frac{n+1}{2} = \frac{30+1}{2} = 15{,}5.$

Медианным является также интервал 27 – 30 лет, так как в этом интервале находятся номера 15 и 16 ряда.

$$\text{Me} = x_{\text{Me}} + i \cdot \frac{\frac{n+1}{2} - S_{(\text{Me}-1)}}{f_{\text{Me}}} = 27 + 3 \cdot \frac{\frac{30+1}{2} - 10}{10} = 28{,}6 \text{ года.}$$

Для расчета показателей вариации составляется вспомогательная таблица (табл. 3.6).

Таблица 3.6

Вспомогательная таблица для расчета показателей вариации

Группы рабочих по возрасту, лет	Центр интервала, (лет), (x')	f	$x' \cdot f$	$d = x' - \bar{x}$	$\lvert d \rvert \cdot f$	d^2	$d^2 \cdot f$
18 – 21	19,5	1	19,5	–9,2	9,2	84,64	84,64
21 – 24	22,5	3	67,5	–6,2	18,6	38,44	115,32
24 – 27	25,5	6	153,0	–3,2	19,2	10,24	61,44
27 – 30	28,5	10	285,0	–0,2	20,0	0,04	0,40
30 – 33	31,5	5	157,5	2,8	14,0	7,84	39,20
33 – 36	34,5	3	103,5	5,8	17,4	33,64	100,92
36 – 39	37,5	2	75,0	8,8	17,6	77,44	154,88
Итого	–	30	861,0	–	116,0	–	556,80

$$\bar{d} = \frac{\sum \lvert x' - \bar{x} \rvert \cdot f}{\sum f} = \frac{\sum \lvert d \rvert \cdot f}{\sum f} = \frac{116{,}0}{30} = 3{,}87 \text{ года}$$

$$\sigma = \sqrt{\frac{\sum (x' - \bar{x})^2 f}{\sum f}} = \sqrt{\frac{\sum d^2 f}{\sum f}} = \sqrt{\frac{556{,}8}{30}} = 4{,}31 \text{ года;}$$

$$V = \frac{\sigma}{\bar{x}} \cdot 100 = \frac{4{,}31}{28{,}7} \cdot 100 = 15{,}0\%.$$

Следовательно, вариация возраста у рабочих данного цеха не является значительной, что подтверждает достаточную однородность совокупности.

Как видно на рис. 3.2, распределение рабочих по возрасту несимметрично, поэтому определяется показатель асимметрии:

$$A_s = \frac{\bar{x} - \text{Mo}}{\sigma} = \frac{28{,}7 - 28{,}3}{4{,}31} = +0{,}09.$$

Следовательно, асимметрия правосторонняя, незначительная.

При правосторонней асимметрии между показателями центра распределения существует соотношение

$$\text{Mo} < \text{Me} < \bar{x}.$$

Для данного распределения это соотношение выполняется, т. е. 28, 3 < 28, 6 < 28, 7. При левосторонней асимметрии (A_S со знаком минус) соотношение между показателями центра распределения будет иметь вид:

$$\text{Mo} > \text{Me} > \bar{x}.$$

Для имеющегося распределения, учитывая незначительную асимметрию, целесообразно определить показатель эксцесса (островершинности):

$$E_x = \frac{\mu_4}{\sigma^4} - 3;$$

$$\mu_4 = \frac{\sum (x' - \bar{x})^4 \cdot f}{\sum f},$$

где μ_4 — центральный момент четвертого порядка.

$$\mu_4 = \frac{(19{,}5-28{,}7)^4 \cdot 1 + (22{,}5-28{,}7)^4 \cdot 3 + (25{,}5-28{,}7)^4 \cdot 6 + (28{,}5-28{,}7)^4 \cdot 10 +}{}$$

$$\frac{+(31{,}5-28{,}7)^4 \cdot 5 + (34{,}5-28{,}7)^4 \cdot 3 + (37{,}5-28{,}7)^4 \cdot 2}{30} = 930{,}7;$$

$$E_x = \frac{930{,}7}{4{,}31^4} - 3 = 2{,}7 - 3 = -0{,}3.$$

Отрицательное значение эксцесса свидетельствует о плосковершинности распределения.

3.3. Распределение рабочих предприятий по размеру заработной платы за август следующее (табл. 3.7).

Таблица 3.7

Месячная заработная плата, руб.	500–1000	1000–1500	1500–2500	2500–4000	4000 и более	Итого
Число рабочих	44	113	245	537	61	1000

Требуется дать графическое изображение ряда в виде гистограммы.

Решение

Представлен ряд с неравными интервалами, поэтому следует построить гистограмму плотностей распределения, так как плотность дает представление о заполненности каждого интервала (табл. 3.8).

Таблица 3.8

Месячная заработная плата, руб., x	500–1000	1000–1500	1500–2500	2500–4000	4000–5500
Абсолютная плотность распределения, p	0,088	0,226	0,245	0,358	0,041

Абсолютная плотность распределения определялась так:

$$p_1 = \frac{44}{500} = 0{,}088; \quad p_2 = \frac{113}{500} = 0{,}226 \text{ и т. д.}$$

На рис. 3.4 представлена гистограмма.

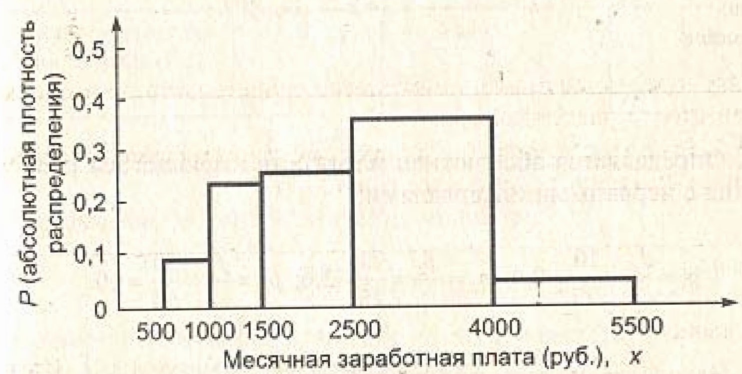

Рис. 3.4. Гистограмма распределения рабочих предприятия по размеру месячной заработной платы

3.4. Распределение грузовых автомобилей автотранспортного предприятия по величине суточного пробега за 10 мая следующее (табл. 3.9).

Таблица 3.9

Суточный пробег (км), x	100–125	125–150	150–200	Итого
Число автомобилей, f	10	90	50	150

Требуется построить вариационный ряд с равными интервалами в 20 км.

Решение

Построение нового ряда с равными интервалами осуществляется в следующей последовательности.

1. Записывается макет таблицы нового ряда (табл. 3.10).

Таблица 3.10

Суточный пробег (км), x	100–120	120–140	140–160	160–180	180–200	Итого
Число автомобилей, f						

2. Определяется абсолютная плотность имеющегося распределения с неравными интервалами:

$$p_1 = \frac{f_1}{i_1} = \frac{10}{25} = 0{,}4;\ p_2 = \frac{f_2}{i_2} = \frac{90}{25} = 3{,}6;\ p_3 = \frac{f_3}{i_3} = \frac{50}{50} = 1{,}0.$$

3. Находятся частоты для каждой группы нового ряда по формуле:

$$f = \Sigma p_i \cdot i_i;$$

$$f_1 = 0{,}4 \cdot 20 = 8;$$
$$f_2 = 0{,}4 \cdot 5 + 3{,}6 \cdot 15 = 56;$$
$$f_3 = 3{,}6 \cdot 10 + 1{,}0 \cdot 10 = 46;$$
$$f_4 = 1{,}0 \cdot 20 = 20;$$
$$f_5 = 1{,}0 \cdot 20 = 20.$$

4. Оформляется новый ряд распределения с равными интервалами.

Распределение грузовых автомобилей автотранспортного предприятия по величине суточного пробега за 10 мая представлено в табл. 3.11.

Таблица 3.11

Суточный пробег (км), x	100–120	120–140	140–160	160–180	180–200	Итого
Число автомобилей, f	8	56	46	20	20	150

3.5. Дневная производительность труда рабочих бригады, выполняющих одинаковую операцию по обработке детали № 408, следующая (табл. 3.12).

Таблица 3.12

Дневная производительность труда (шт.), x	18	19	20	21	22	23	Итого
Число рабочих, f	1	3	6	5	3	2	20

Определить численное значение медианы.

Решение

Совокупность содержит четное число значений признака, поэтому медиана определяется по формуле:

$$\text{Me} = \frac{1}{2}...(x_k + x_{k+1});$$

$k = n : 2 = 10.$

Следовательно

$$\mathrm{Me} = \frac{1}{2}(x_{10} + x_{11}).$$

Для определения численных значений x_{10} и x_{11} исчисляются накопленные частоты (табл. 3.13).

Таблица 3.13

Дневная производительность труда (шт.), x	18	19	20	21	22	23
Число рабочих, f	1	4	10	15	18	20

$$x_{10} = 20 \text{ шт.}; x_{11} = 21.$$

Отсюда,

$$\mathrm{Me} = \frac{1}{2}(20 + 21) = 20{,}5 \text{ шт.}$$

3.6. Имеются следующие данные о времени простоя автомобиля под разгрузкой (табл. 3.14).

Таблица 3.14

№ пункта разгрузки	1	2	3	4	5	6	7	8	9	10
Число грузчиков	3	4	4	3	3	4	4	4	3	4
Время простоя, мин.	12	10	8	15	19	12	8	10	18	8

Проверить закон сложения дисперсий.

Решение

В данной задаче варьирующим признаком является время простоя автомобиля под разгрузкой. Общая дисперсия времени простоя под разгрузкой определяется по формуле

$$\sigma_0^2 = \frac{\sum (x - \overline{x}_0)^2 f}{\sum f}.$$

Для расчета общей дисперсии составляется дискретный ряд распределения, промежуточные расчеты помещены в табл. 3.15.

Таблица 3.15

Вспомогательная таблица для расчета общей дисперсии

Время простоя под разгрузкой (мин. т), x	Число выполненных разгрузок, f	$x \cdot f$	$x - \overline{x}_0$	$(x - \overline{x}_0)^2$	$(x - \overline{x}_0)^2 \cdot f$
8	3	24	−4	16	48
10	2	20	−2	4	8
12	2	24	0	0	0
15	1	15	3	9	9
18	1	18	6	36	36
19	1	19	7	49	49
Итого	10	120	–	–	150

$\overline{x}_0 = \frac{120}{10} = 12{,}0$ мин. – среднее время простоя;

$\sigma_0^2 = \frac{150}{10} = 15$ – общая дисперсия.

Величина этой дисперсии характеризует вариацию времени простоя под разгрузкой под влиянием всех условий.

Различия в величине изучаемого признака прежде всего возникают под влиянием числа грузчиков, принимающих участие в процессе разгрузки. В связи с этим в совокупности выделяются две однородные группы по числу грузчиков: в первую группу включаются наблюдения при числе грузчиков 3; во вторую группу попадают наблюдения при числе грузчиков 4. Для каждой из выделенных групп определяется внутригрупповая дисперсия, возникающая под влиянием неучтенных факторов. Для их расчета использованы вспомогательные табл. 3.16 и 3.17.

Таблица 3.16

Расчет внутригрупповой дисперсии по первой группе (число грузчиков, участвующих в разгрузке, – 3)

Время простоя под разгрузкой (мин.), x	Число выполненных разгрузок, f	$x \cdot f$	$x-\overline{x}_1$	$(x-\overline{x}_1)^2 \cdot f$
12	1	12	–4	16
15	1	15	–1	1
18	1	18	2	4
19	1	19	3	9
Итого	4	64	–	30

Таблица 3.17

Расчет внутригрупповой дисперсии по второй группе (число грузчиков, участвующих в разгрузке, – 4)

Время простоя под разгрузкой (мин.), x	Число выполненных разгрузок, f	$x \cdot f$	$x-\overline{x}_2$	$(x-\overline{x}_2)^2 \cdot f$
8	3	24	–1,33	5,31
10	2	20	0,67	0,90
12	1	12	2,67	7,13
Итого	6	56	–	13,37

$$\overline{x}_1 = \frac{64}{4} = 16 \text{ мин.} \quad \sigma_1^2 = \frac{30}{4} = 7{,}5.$$

$$\overline{x}_2 = \frac{56}{6} = 9{,}33 \text{ мин.}; \quad \sigma_2^2 = \frac{13{,}37}{6} = 2{,}23.$$

Средняя из внутригрупповых дисперсий

$$\overline{\sigma^2} = \frac{\sum \sigma_i^2 \cdot n_i}{\sum n_i} = \frac{7{,}5 \cdot 4 + 2{,}23 \cdot 6}{10} = 4{,}3.$$

Межгрупповая дисперсия, отражающая различия в величине признака под влиянием фактора, положенного в основу группировки, определяется по формуле

$$\delta^2 = \frac{\sum(\overline{x}_i - \overline{x}_0)^2 \cdot n_i}{\sum n_i} = \frac{(16-12)^2 \cdot 4 + (9{,}33-12)^2 \cdot 6}{10} = 10{,}7.$$

Общая дисперсия (σ^2_o) равна сумме средней внутригрупповой дисперсии и межгрупповой дисперсии:

$$\sigma^2_o = 4{,}3 + 10{,}7 = 15{,}0,$$

что и соответствует полученной ранее величине.

3.7. Имеются следующие данные о результатах обследования рабочих предприятия по размеру месячной заработной платы (табл. 3.18).

Таблица 3.18

Группы рабочих по возрасту, лет	Число рабочих	Дисперсия заработной платы
До 20	100	300
20 – 30	120	400
30 и старше	150	500

Общая дисперсия заработной платы в обследованной совокупности рабочих составила 450.

Определить, в какой степени вариация заработной платы рабочих предприятия зависит от возраста.

Решение

Средняя внутригрупповая дисперсия характеризует случайную вариацию под влиянием неучтенных факторов:

$$\overline{\sigma^2} = \frac{\sum \sigma_i^2 n_i}{\sum n_i} = \frac{300 \cdot 100 + 400 \cdot 120 + 500 \cdot 150}{100 + 120 + 150} = 413{,}5.$$

Межгрупповая дисперсия отражает систематическую вариацию под влиянием фактора, положенного в основу группировки (возраста рабочих).

Межгрупповая дисперсия определяется по правилу сложения дисперсий:

$$\delta^2 = \sigma^2_о - \overline{\sigma^2} = 450 - 413,5 = 36,5.$$

Отсюда соотношение дисперсий

$$\delta^2 : \sigma^2_о = 36,5 : 450 = 0,08, \text{ или } 8,0\%.$$

Полученный результат показывает, что возраст на варьирование заработной платы рабочих предприятия не оказывает существенного влияния.

3.8. Удельный вес основных рабочих в трех цехах предприятия составил: 80, 75 и 90% общей численности рабочих.

Определить дисперсию и среднее квадратическое отклонение доли основных рабочих по предприятию в целом, если численность всех рабочих трех цехов составила соответственно 100, 200 и 150 человек.

Решение

Рабочие предприятия подразделяются на две группы: основные и ремонтно-вспомогательные рабочие.

Общая численность основных рабочих по предприятию в целом составит:

$$\sum n_0 = 0,8 \cdot 100 + 0,75 \cdot 200 + 0,9 \cdot 150 = 365 \text{ человек.}$$

Доля основных рабочих по предприятию

$$p = \frac{\sum n_0}{\sum n} = \frac{365}{100 + 200 + 150} = \frac{365}{450} = 0,811.$$

Дисперсия альтернативного признака

$$\sigma^2 = p \cdot q,$$

где p – доля единиц, обладающих данным признаком (доля основных рабочих);

q – доля единиц, не обладающих данным признаком (доля ремонтно-вспомогательных рабочих).

Поскольку $p + q = 1$, следовательно, $q = 1 - p$ и формула дисперсии имеет вид:

$$\sigma^2 = p \cdot (1 - p) = 0,811 \cdot (1 - 0,811) = 0,1533.$$

Среднее квадратическое отклонение доли основных рабочих по предприятию в целом:

$$\sigma = \sqrt{pq} = \sqrt{0,1533} = 0,3915.$$

3.9. Дисперсия признака равна 600. Объем совокупности равен 10. Сумма квадратов индивидуальных значений признака равна 6250.

Найти среднюю величину.

Решение

Для нахождения средней величины воспользуемся формулой

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

где $\overline{x^2}$ – средняя арифметическая из квадратов индивидуальных значений признака;

$(\overline{x})^2$ – квадрат среднего значения признака.

Тогда

$$\overline{x^2} = \frac{\sum x^2}{n} = \frac{6250}{10} = 625.$$

Средняя величина признака

$$\overline{x} = \sqrt{\overline{x^2} - \sigma^2} = \sqrt{625 - 600} = 5.$$

3.10. Распределение работников предприятия по размеру месячной заработной платы следующее (табл. 3.19).

Определить коэффициент децильной дифференциации.

Сформулировать вывод.

Таблица 3.19

Месячная заработная плата, тыс. руб.	До 1,0	1,0–2,0	2,0–3,0	3,0–4,0	4,0–5,0	5,0 и более	Итого
Число работников	25	89	145	215	65	20	559

Решение

Коэффициент децильной дифференциации

$$K_D = \frac{D_9}{D_1},$$

где D_9 – девятая дециль;
D_1 – первая дециль.

Определяется место децилей:

$$N_{D_1} = \frac{n+1}{10} = \frac{559+1}{10} = 56; \; N_{D_9} = \frac{9(n+1)}{10} = \frac{9(559+1)}{10} = 504.$$

Для расчета численных значений децилей определяются интервалы, в которых они находятся, для чего исчисляются накопленные частоты в табл. 3.20.

Таблица 3.20

Месячная заработная плата, тыс. руб.	До 1,0	1,0–2,0	2,0–3,0	3,0–4,0	4,0–5,0	5,0 и более
Накопленные частоты	25	114	259	474	539	559

Первая дециль находится в интервале 1,0–2,0, девятая – в интервале 4,0–5,0.

Численные значения децилей следующие:

$$D_1 = 1{,}0 + 1{,}0\frac{56-25}{89} = 1{,}3483 \text{ тыс. руб. или } 1348{,}3 \text{ руб.};$$

$$D_9 = 4{,}0 + 1{,}0\frac{504-474}{65} = 4{,}4615 \text{ тыс. руб. или } 4461{,}5 \text{ руб.};$$

$$K_D = \frac{4461.5}{1348.3} = 3.31.$$

Следовательно, наименьший размер месячной заработной платы 10% наиболее обеспеченных работников в 3,31 раза выше наивысшего размера месячной заработной платы 10% наименее обеспеченных работников.

3.11. По цеху имеются следующие данные о распределении рабочих по стажу работы (табл. 3.21).

Таблица 3.21

Группы рабочих по стажу работы (лет), x	0–2	2–4	4–6	6–8	8–10	10–12	12–14	Итого
Число рабочих, f	6	8	12	24	17	8	5	80

На основе приведенных данных *проверить* соответствие эмпирического распределения закону нормального распределения, используя критерий согласия К. Пирсона.

Решение

Для ответа на поставленный вопрос прежде всего исчисляются теоретические частоты нормального распределения по формуле

$$f' = \frac{n \cdot i}{\sigma} \cdot \frac{1}{\sqrt{2\pi}} \cdot e^{-\frac{t^2}{2}},$$

где t – нормированное отклонение.

$$t = \frac{x - \bar{x}}{\sigma}.$$

$$\bar{x} = \frac{\sum x'f}{\sum f} = \frac{1 \cdot 6 + 3 \cdot 8 + 5 \cdot 12 + 7 \cdot 24 + 9 \cdot 17 + 11 \cdot 8 + 13 \cdot 5}{80} = 7{,}0 \text{ года.}$$

В формуле x' – центральное значение интервала.

$$\sigma=\sqrt{\frac{\sum(x-\bar{x})^2 f}{\sum f}}=$$

$$=\sqrt{\frac{(1-7)^2 6+(3-7)^2 8+(5-7)^2 12+(7-7)^2 24+(9-7)^2 17+(11-7)^2 8+(13-7)^2 5}{80}}=$$

$$=3{,}1\ \text{года}.$$

$$\frac{n\cdot i}{\sigma}=\frac{80\cdot 2}{3{,}1}=51{,}6.$$

Величина $\frac{1}{\sqrt{2\pi}}\cdot e^{-\frac{t^2}{2}}$ определяется по таблице (см. приложение 1).

Расчет теоретических частот выполнен в табл. 3.22.

Таблица 3.22

Вспомогательная таблица для расчета теоретических частот нормального распределения

Группы рабочих по стажу работы (лет), x	Число рабочих, f	Центр интервала (лет), x'	$t=\frac{x'-\bar{x}}{\sigma}$	$\frac{1}{\sqrt{2\pi}}e^{-\frac{t^2}{2}}$	Теоретические частоты, f'	Округленные теоретические частоты, f'	$f-f'$	$\frac{(f-f')^2}{f'}$
0–2	6	1	–1,935	0,063	3,251	3,3	2,7	2,21
2–4	8	3	–1,290	0,174	8,978	9,0	–1,0	0,11
4–6	12	5	–0,645	0,325	16,770	16,8	–4,8	1,37
6–8	24	7	0	0,399	20,588	20,6	3,4	0,56
8–10	17	9	0,645	0,325	16,770	16,8	0,2	0,00
10–12	8	11	1,290	0,174	8,978	9,0	–1,0	0,11
12–14	5	13	1,935	0,063	3,251	3,3	1,7	0,88
Итого	80					78,8[1]		5,24

[1] $\sum f'$ оказалась равной 78,8, а не 80, что объясняется округлением цифр при расчете.

На рис. 3.5 представлено распределение рабочих по стажу работы.

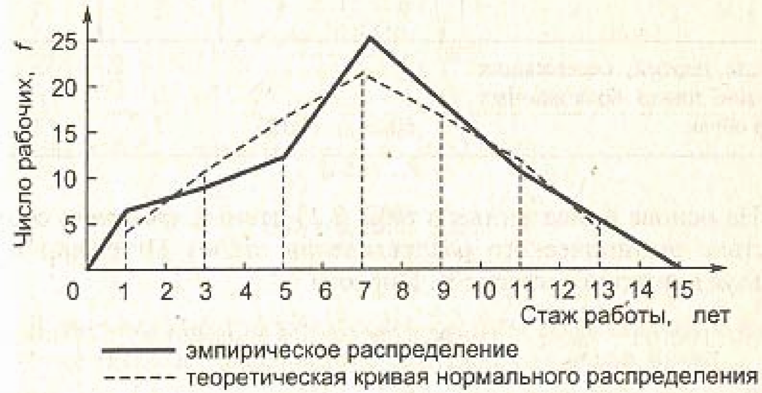

Рис. 3.5. Распределение рабочих по стажу работы

Эмпирическое распределение замыкается в точке x = 15 лет (это середина интервала 14 – 16, где частота равна 0) и в точке x = 0. Теоретическая кривая нормального распределения с осью абсцисс не замыкается.

Сопоставление на графике эмпирического распределения с теоретической кривой нормального распределения свидетельствует о достаточно хорошем согласовании распределений.

Степень расхождения теоретических и эмпирических частот оценивается с помощью критерия «хи-квадрат» (χ^2) К. Пирсона:

$$\chi^2=\sum\frac{(f-f')^2}{f'}=5{,}24\ \text{(см. табл. 3.22)}.$$

Полученное значение критерия ($\chi^2_{расч}$) сравнивается с табличным значением $\chi^2_{табл}$, которое определяется по таблице (см. приложение 2).

При вероятности P = 0,95 и числе степеней свободы k = 4 (7 – 3) $\chi^2_{табл}$ = 9,5. $\chi^2_{расч} < \chi^2_{табл}$ (5,24 < 9,5), следовательно, гипотеза о близости эмпирического распределения к нормальному не отвергается.

3.12. По предприятию имеются следующие данные о проверке 100 партий мужской обуви, передаваемых в торговую сеть (табл. 3.23).

Таблица 3.23

Число бракованных пар обуви	0	1	2	3
Число партий, содержащих данное число бракованных пар обуви	60	32	7	1

На основе приведенных в табл. 3.23 данных *проверить* соответствие эмпирического распределения закону Пуассона, используя критерий согласия К. Пирсона.

Решение

Для ответа на поставленный вопрос определяются теоретические частоты распределения Пуассона по формуле

$$f' = n \cdot \frac{\lambda^m \cdot e^{-\lambda}}{m!}$$

Нахождение теоретических частот производится в следующей последовательности.

1. Определяется величина λ, равная среднему числу появления события (среднее число бракованных пар обуви в одной партии):

$$\lambda = \frac{\sum m \cdot f}{\sum f},$$

где m – число бракованных пар обуви;

f – число партий, содержащих данное число бракованных пар обуви.

$$\lambda = \frac{0 \cdot 60 + 1 \cdot 32 + 2 \cdot 7 + 3 \cdot 1}{100} = 0{,}44.$$

2. По приложению 8 определяется значение $e^{-\lambda}$:

$$e^{-\lambda} = 2{,}71828^{-0{,}44} = 0{,}6440.$$

3. Для каждого значения m по приведенной выше формуле определяется теоретическая частота:

при $m = 0$ $f' = 100 \cdot \frac{0{,}44^0 \cdot 0{,}6440}{0!} = 64{,}4;$

при $m = 1$ $f' = 100 \cdot \frac{0{,}44^1 \cdot 0{,}6440}{1} = 28{,}3;$

при $m = 2$ $f' = 100 \cdot \frac{0{,}44^2 \cdot 0{,}6440}{1 \cdot 2} = 6{,}2;$

при $m = 3$ $f' = 100 \cdot \frac{0{,}44^3 \cdot 0{,}6440}{1 \cdot 2 \cdot 3} = 0{,}9.$

Эмпирические и теоретические частоты представлены в табл. 3.24. Все теоретические частоты округлены до целых чисел.

Таблица 3.24

Распределение партий обуви по числу бракованных пар

Число бракованных пар обуви	Число партий		$f - f'$	$\frac{(f-f')^2}{f'}$
	эмпирические частоты, f	теоретические частоты, f'		
0	60	65	5	0,38
1	32	28	4	0,57
2	7	6	1	0,17
3	1	1	0	–
Итого	100	100	–	1,11

На рис. 3.6 представлено распределение партий обуви по числу бракованных пар.

Графическое сопоставление обоих распределений говорит о достаточной близости между эмпирическим и теоретическим распределениями.

Степень расхождения теоретических и эмпирических частот оцениваем с помощью критерия «хи-квадрат» (χ^2) К. Пирсона:

$$\chi^2_{расч} = \sum \frac{(f - f')^2}{f'} = 1{,}11 \text{ (см. табл. 3.24).}$$

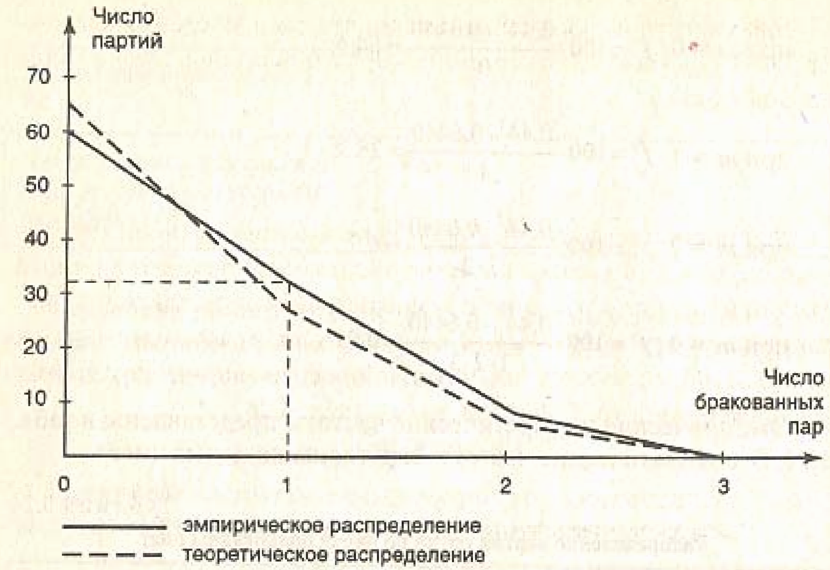

Рис. 3.6. Распределение партий обуви по числу бракованных пар

Полученное значение критерия ($\chi^2_{расч}$) сравнивается с табличным значением $\chi^2_{табл}$, которое определяется по таблице (см. приложение 2). При вероятности $P = 0{,}95$ и числе степеней свободы $k = 2$ $(4 - 2)$ $\chi^2_{табл} = 3{,}8$.

$$\chi^2_{расч} < \chi^2_{табл} \ (1{,}11 < 3{,}8).$$

Следовательно, гипотеза близости эмпирического распределения к распределению Пуассона не отвергается.

3.13. Распределение коммерческих банков региона на 1 января 2002 г. по размеру капитала следующее (табл. 3.25).

Таблица 3.25

Группы банков по размеру активов, млн руб.	200–250	250–300	300–350	350–400	400 и более	Итого
Число банков	3	8	12	5	3	31
Общая сумма активов, млн руб.	675	2200	3900	1875	1275	9925

Для характеристики неравномерности распределения капитала банков региона дать графическое изображение ряда в виде кривой Лоренца.

Решение

Для построения кривой Лоренца необходимо выполнить следующие дополнительные расчеты: абсолютные показатели числа единиц в группах и размерах признака (общая сумма активов) выражаются в относительных величинах (в долях или процентах к итогу) и определяются их накопленные значения. Результаты этих преобразований представлены в табл. 3.26.

Таблица 3.26

Распределение коммерческих банков по размеру капитала

Группы банков по размеру активов, млн руб.	Число банков			Общая сумма активов		
	частота, f	частость (%), $(f{:}\Sigma f)\cdot 100$	накопленная частость, %	млн руб.	в % к итогу	накопленный % к итогу
1	2	3	4	5	6	7
200–250	3	9,7	9,7	675	6,8	6,8
250–300	8	25,8	35,5	2200	22,2	29,0
300–350	12	38,7	74,2	3900	39,3	68,3
350–400	5	16,1	90,3	1875	18,9	87,2
400 и более	3	9,7	100,0	1275	12,8	100,0
Итого	31	100,0	–	9925	100,0	–

Кривая Лоренца представлена на рис. 3.7. На график нанесены накопленные значения двух рядов (табл. 3.26 гр. 4 и 7). При соединении всех точек получена кривая, отражающая процесс концентрации капитала.

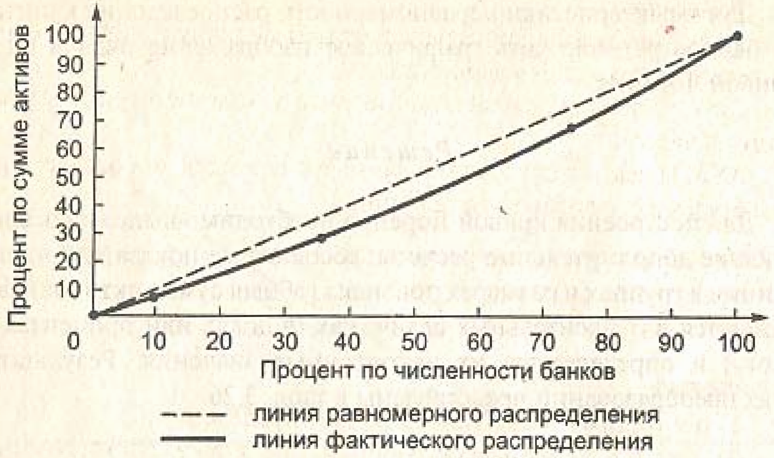

Рис. 3.7. Кривая Лоренца

3.2. Задачи для самостоятельной работы

3.14. Имеются следующие данные о размере семьи работников цеха (число человек в семье):

3	4	5	2	3	6	4	2	5	3	4	2	7	3	3	6
2	3	8	5	6	7	3	4	5	4	3	3	4			

Требуется:

1) составить дискретный вариационный ряд;

2) определить показатели центра распределения, показатели вариации;

3) дать графическое изображение ряда в виде полигона распределения.

Сформулировать краткие выводы.

3.15. Хронометраж операций пайки радиаторов на ремонтном предприятии дал следующие результаты (табл. 3.27).

Таблица 3.27

Время пайки, мин.	20–30	30–40	40–50	50–60	60–70	Итого
Количество радиаторов	2	5	10	17	1	35

Вычислить:

а) среднее время пайки радиатора; б) медиану и моду; в) относительный показатель вариации.

Дать графическое изображение ряда в виде гистограммы и полигона частот.

3.16. Имеются следующие данные о возрастном составе группы студентов вечернего отделения:

18	38	28	29	26	38	34	22	28	30
22	23	35	33	27	24	30	32	28	25
29	26	31	24	29	27	32	25	29	20

Требуется:

1) построить интервальный ряд распределения;

2) дать его графическое изображение в виде гистограммы и кумуляты;

3) определить численное значение моды и медианы, используя графическое изображение.

3.17. По предприятию получены данные о расстоянии перевозки партий груза в междугородном сообщении (км):

560	1060	420	1410	1500	400	3800	700	1780	450
449	285	1850	2200	800	1200	1540	1150	180	452
452	2500	300	400	900	1800	452	1850	1225	220
1800	300	920	1400	1400	480	850	200	400	1440
420	1700	1615	3500	300	320	600	965	450	245

Для анализа работы предприятия *требуется*:

1) построить интервальный ряд распределения партий груза по дальности перевозки, определив величину интервала по формуле Стерджесса;

2) дать графическое изображение ряда;

3) исчислить показатели центра распределения и показатели вариации.

Сформулировать вывод.

3.18. Имеются следующие данные о распределении продовольственных магазинов региона по размеру товарооборота за месяц (табл. 3.28).

Требуется вычислить средний месячный размер товарооборота магазинов региона, дисперсию и коэффициент вариации.

Таблица 3.28

Группы магазинов по товарообороту, млн руб.	40–50	50–60	60–70	70–80	80–90	90–100	100–110	110–120	120–130	130–140
Число магазинов	2	4	7	10	15	20	22	11	6	3

3.19. По автотранспортному предприятию, осуществляющему перевозку грузов автомобилями КамАЗ-5320 грузоподъемностью 16 т, имеются следующие данные о весе партий груза (т):

8	11	14	6	10	13	12	16	15	16
16	10	16	13	14	16	16	4	16	14
5	13	11	2	16	8	16	7	14	16

Требуется:

1) построить интервальный ряд распределения партий груза по весу;

2) вычислить для построенного ряда показатели центра распределения и вариации.

Сформулировать вывод об использовании автомобилей КамАЗ-5320.

3.20. Выходной контроль качества поступающих комплектующих изделий дал следующие результаты (табл. 3.29).

Таблица 3.29

№ партии изделий	1	2	3	4	5
Процент брака	2	5	12	1	3

Вычислить дисперсию доли брака по каждой поступившей партии.

3.21. Распределение рабочих двух участков по стажу работы следующее (табл. 3.30).

Определить, на каком участке состав рабочих по стажу работы более однороден.

Таблица 3.30

Стаж работы, лет	Число рабочих	
	Участок № 1	Участок № 2
0 – 5	2	7
5 – 10	15	25
10 – 15	20	12
15 – 20	3	8

3.22. Заработная плата 10 рабочих бригады характеризуется следующими данными (табл. 3.31).

Таблица 3.31

Профессия	Число рабочих	Месячная заработная плата каждого рабочего за март, руб.					
Токари	4	3252;	3548;	3600;	3400		
Слесари	6	3450;	3380;	3260;	3700;	3250;	3372

Проверить правило сложения дисперсий и указать, велико ли влияние профессии на различие в уровне заработной платы.

3.23. Имеются следующие данные о часовой интенсивности движения автомобилей на автомагистрали (авт/ч):

140	99	80	140	218	340	92	152	120	130
50	110	130	96	48	36	60	30	86	102
90	210	220	261	282	312	68	80	131	190

Требуется:

1) построить интервальный ряд распределения;

2) вычислить: среднее линейное отклонение, дисперсию, среднее квадратическое отклонение и коэффициент вариации.

3.24. Имеются следующие данные о заработной плате рабочих автотранспортного предприятия за январь (табл. 3.32).

Определить общую дисперсию заработной платы рабочих предприятия, а также сопоставить однородность двух групп рабочих по уровню месячной заработной платы.

Таблица 3.32

Группы рабочих	Число рабочих	Средняя месячная заработная плата, руб.	Внутригрупповая дисперсия заработной платы
Водители	840	4100,0	6600
Ремонтно-вспомогательные рабочие	125	3397,5	2800

3.25. Имеются следующие данные о количестве заявок на автомобили технической помощи по дням:

11	2	5	14	7	2	8	10	2	6
10	8	3	13	11	8	8	2	9	8
5	14	4	10	12	6	8	2	8	7
9	2	8	4	6	13	5	3	12	2
2	7	9	8	5	8	6	10	11	5

Требуется:

1) построить: а) интервальный ряд распределения суточного количества заявок на автомобили технической помощи, определив величину интервала группировки по формуле Стерджесса; б) теоретическую кривую нормального распределения;

2) проверить соответствие эмпирического и теоретического распределений по критерию К. Пирсона.

Сформулировать вывод.

3.26. Средняя величина в совокупности равна 15, среднее квадратическое отклонение равно 10.

Чему равен средний квадрат индивидуальных значений этого признака?

3.27. Средняя величина признака в совокупности равна 13, а средний квадрат индивидуальных значений этого признака равен 174.

Определить коэффициент вариации.

3.28. Дисперсия признака равна 360 000, коэффициент вариации равен 50%.

Чему равна средняя величина признака?

3.29. Дисперсия признака равна 25, средний квадрат индивидуальных значений равен 250.

Чему равна средняя?

3.30. Средняя величина признака равна 2600 единицам, а коэффициент вариации равен 30%.

Определить дисперсию признака.

3.31. Общая дисперсия равна 8,4. Средняя величина признака для всей совокупности равна 13. Средние по группам равны соответственно 10, 15 и 12. Численность единиц в каждой группе составляет 32, 53 и 45.

Определить среднюю внутригрупповую дисперсию.

3.32. На двух предприятиях фирмы выпускается одинаковый вид изделий. На первом предприятии изготовили 12 тыс. изделий; на втором – 10 тыс. Средняя себестоимость изделий на первом предприятии – 100 тыс. руб., на втором – 110 тыс. руб. Дисперсия себестоимости на первом предприятии – 20 тыс. руб., на втором – 2,5 тыс. руб.

Вычислите дисперсию себестоимости изделий в целом по фирме.

3.33. Имеются следующие данные о размере заработной платы рабочих цеха за апрель (табл. 3.33).

Таблица 3.33

Профессия	Число рабочих	Средняя заработная плата, руб.	Внутригрупповая дисперсия заработной платы
Токари	50	4650	6500
Фрезеровщики	25	4800	5025
Слесари	40	4500	4910

Требуется:

1) определить общую дисперсию заработной платы рабочих цеха;

2) оценить однородность совокупности рабочих цеха по уровню месячной заработной платы;

3) определить, на сколько процентов дисперсия в размере заработной платы обусловлена различием в профессии рабочих и влиянием прочих причин.

Сформулировать вывод.

3.34. Имеются следующие данные 10%-ного случайного бесповторного выборочного обследования рабочих механического цеха (табл. 3.34).

Таблица 3.34

Табельный номер рабочего	Возраст, лет	Заработная плата за сентябрь, руб.	Стаж работы, лет	Тарифный разряд
2	25	4480	7	3
17	24	2360	7	2
28	43	4510	25	4
35	41	4670	23	5
44	37	3880	18	5
47	42	4965	24	5
102	29	2744	11	5
112	36	4030	16	5
123	56	5150	34	6
135	29	3740	11	5
138	18	2215	1	2
140	37	3582	20	4
147	25	2500	8	3
149	30	3630	12	4
150	26	3520	9	3

Требуется:

1) определить дисперсию заработной платы рабочих;

2) произвести группировку рабочих по стажу работы, выделив три группы; для каждой выделенной группы исчислить внутригрупповую дисперсию по уровню месячной заработной платы;

3) определить среднюю внутригрупповую дисперсию по уровню месячной заработной платы и ее долю в общей дисперсии.

Сформулировать вывод.

3.35. По данным задачи 3.34 *требуется*:

1) определить общую дисперсию заработной платы рабочих;

2) произвести группировку рабочих по уровню квалификации; для каждой выделенной группы исчислить внутригрупповую дисперсию по уровню месячной заработной платы;

3) определить среднюю внутригрупповую дисперсию по уровню месячной заработной платы;

4) проверить правило сложения дисперсий.

3.36. По данным задачи 3.34 требуется определить, по какому признаку более однородна группа рабочих цеха – по стажу работы или по уровню квалификации.

3.37. По данным задачи 3.34 требуется определить, по какому признаку более однородна группа рабочих цеха – по стажу работы или по возрасту.

3.38. По совокупности, состоящей из 100 единиц, известны: 1) средняя арифметическая – 47,0; 2) сумма квадратов индивидуальных значений признака – 231 592.

Определить, достаточно ли однородна изучаемая совокупность.

3.39. По группе промышленных предприятий имеются следующие данные (табл. 3.35).

Таблица 3.35

Группы предприятий по стоимости основного капитала, млн руб.	Число предприятий	Средний объем продукции в группе, млн руб.	Внутригрупповая дисперсия объема продукции
40 – 50	15	290	190,7
50 – 60	8	410	115,8
60 – 70	2	520	84,0

Определить общую дисперсию объема продукции.

3.40. Выпуск продукции по предприятию следующий, млн руб. (табл. 3.36).

Таблица 3.36

Показатель	Квартал			
	I	II	III	IV
Выпуск продукции – всего	150	120	160	180
В том числе продукция на экспорт	85	60	128	130

Вычислить за год дисперсию удельного веса продукции на экспорт.

3.41. Имеются следующие данные о составе рабочих машиностроительного предприятия (табл. 3.37).

Таблица 3.37

Показатель	Цех № 1	Цех № 2	Цех № 3	Итого по предприятию
Число рабочих	200	300	500	1000
Из них:				
в возрасте до 30 лет	50	110	170	330
с заработной платой до 4500 руб.	30	70	50	150
с общим стажем работы 15 и более лет	60	70	50	180
имеющие 5-й и 6-й разряды	90	60	50	200

С целью сравнительного анализа по каждому из указанных признаков для каждого цеха и предприятия в целом *определить*:

1) долю рабочих, обладающих данным признаком;

2) дисперсию доли рабочих, обладающих данным признаком;

3) среднее квадратическое отклонение.

Указать наибольшую и наименьшую дисперсию по каждому признаку.

3.42. Обработка детали № 37 производится в цехе на токарном полуавтомате. На 25 января получены следующие данные о размере обработанных деталей (в отклонениях от номинала) (табл. 3.38).

Таблица 3.38

Отклонения от номинала (в сотых долях мм)	0 – 2	2 – 4	4 – 6	6 – 8	8 – 10	10 – 12	12 – 14
Число деталей	6	15	18	36	30	9	6

Для характеристики состояния технологического процесса требуется проверить соответствие эмпирического распределения закону нормального распределения, используя критерий согласия К. Пирсона.

3.43. Имеются следующие данные о величине межремонтного пробега автомобилей ЗИЛ-133 (табл. 3.39).

Таблица 3.39

Величина межремонтного пробега, тыс. км	80 – 100	100 – 120	120 – 140	140 – 160	160 – 180
Число автомобилей	10	60	100	26	14

Требуется:

1) дать графическое изображение ряда в виде гистограммы и кумуляты;

2) определить численное значение моды и медианы, используя графические изображения;

3) определить показатель асимметрии.

Сформулировать вывод.

3.44. При проверке партии электроламп из 1000 шт. 30 шт. оказались бракованными.

Определить дисперсию и среднее квадратическое отклонение.

3.45. Средняя арифметическая по совокупности из 20 единиц равна 2,0. Сумма квадратов значений равна 120.

Чему равна величина среднего квадратического отклонения?

3.46. Какому значению x_i соответствует нормированное отклонение 1,50, если для нормально распределенной величины средняя арифметическая и среднее квадратическое отклонения равны соответственно 1 и 0,2.

3.47. Распределение доходности акций характеризуется левосторонней асимметрией. Модальное значение равно 14%, а соответствующая ей частость – 0,345. Будет ли частость, соответствующая средней арифметической: а) больше 0,345; б) меньше 0,345; в) равна 0,345; г) все ответы неверны.

3.48. Распределение рабочих предприятия по размеру месячного дохода следующее (табл. 3.40).

Определить коэффициент децильной дифференциации. Сформулировать вывод.

Таблица 3.40

Месячный доход, руб.	4200–4500	4500–4800	4800–5100	5100–5400	5400 и более	Итого
Число рабочих	75	218	300	201	25	819

3.49. Распределение парка трамвайных вагонов по длительности нахождения в эксплуатации в двух городах следующее (табл. 3.41).

Таблица 3.41

Город А		Город Б	
группы вагонов по длительности нахождения в эксплуатации, лет	число вагонов, % к итогу	группы вагонов по длительности нахождения в эксплуатации, лет	число вагонов, % к итогу
До 5	8	До 3	12
5–12	28	3–7	22
12–20	64	7–12	35
		12–20	24
		20 и более	7
Итого	100		100

Произвести перегруппировку данных с целью изучения сопоставимых рядов для двух городов, приняв равную величину интервала в 5 лет.

3.50. Распределение промышленных предприятий города по численности работников следующее (табл. 3.42).

Таблица 3.42

Группы предприятий по численности работников, чел.	До 50	50–100	100–200	200–400	400–800	800–1200	1200 и более
Число предприятий	140	80	35	60	45	12	10

Оценить уровень неравномерности распределения работников, используя для этого построение кривой Лоренца.

Сформулировать вывод.

Ответы к задачам для самостоятельной работы

3.14. $\overline{x}_0$ =4,1 чел.; Мо = 3 чел.; Ме = 4 чел.; σ = 1,6 чел.; v = 39,0%. **3.15.** $\overline{x}$ = 47,9 мин.; Мо = 53,0 мин.; Ме = 50,6 мин.; v = 19,7%. **3.18.** $\overline{x}$ = 94,2 млн руб.; σ^2 = 399,36; v = 21,2%. **3.20.** 0,0196; 0,0475; 0,1056; 0,0099; 0,0291. **3.21.** На участке № 1. **3.22.** σ^2_0 = 21509,8; δ^2 = 553,0; $\overline{\sigma^2}$ = 20956,8. **3.24.** $\overline{x}_0$ = 4009; $\overline{\sigma^2}$ = 6107,8; δ^2 = 55645,2; σ^2_0 = 61753; $V_в$ = 1,98%; $V_р$ = 1,56%. **3.26.** $\overline{x^2}$ = 325. **3.27.** v = 17,20%. **3.28.** $\overline{x}$ = 1200. **3.29.** $\overline{x}$ = 15. **3.30.** σ^2 = 608400. **3.31.** $\overline{\sigma^2}$ = 4,2. **3.33.** 1) $\overline{x}_0$ = 4630,4; δ^2 = 12334,6; $\overline{\sigma^2}$ = 5626,3; σ^2_0 = 17960,9; 2) v = (134,0 : 4630,4) · 100 = 2,89% – однородна; 3) 68,7%; 31,3%. **3.38.** v = 22,0% – однородна. **3.39.** $\overline{\sigma^2}$ = 98,2; δ^2 = 5613,8; σ^2_0 = 5712,0. **3.40.** p = 0,661; q = 0,339; σ^2 = 0,224. **3.44.** p = 0,03; q = 0,97; σ^2 = 0,0291; σ = 0,171. **3.45.** σ =1,41. **3.46.** x_i = 1,3. **3.48.** D_1 = 4509,6; D_9 = 5316,4; K_D = 1,179. **3.49.** А: 8; 20; 32; 40; Б: 23; 32; 23; 15; 7.

ГЛАВА 4

Выборочное наблюдение

Понятие о выборочном наблюдении

Наиболее совершенным и научно обоснованным способом несплошного наблюдения является выборочное наблюдение, получившее в настоящее время широкое применение в работе органов государственной статистики, научно-исследовательских лабораторий, институтов, предприятий. Его использование позволяет лучше организовать наблюдение, обеспечивает быстроту проведения, экономию труда и средств на получение и обработку информации.

Выборочное наблюдение при строгом соблюдении условий случайности и достаточно большой численности отобранных единиц репрезентативно (представительно); по результатам изучения определенной части единиц с достаточной для практики степенью точности можно судить о всей совокупности. Однако вычисленные по материалам выборочного наблюдения статистические показатели не будут точно совпадать с соответствующими характеристиками для всей совокупности (генеральной совокупности). Величина этих отклонений называется **ошибкой наблюдения**, которая складывается из ошибок двоякого рода: ошибки регистрации (точности) и ошибки репрезентативности.

Ошибки регистрации свойственны любому наблюдению (сплошному и несплошному). Они вызываются несовершенством измерительных приборов, недостаточной квалификацией наблюдателя, неточностью подсчетов и т. п. Однако при выборочном наблюдении они значительно меньше, так как в этом случае используются более квалифицированные и подготовленные кадры.

Ошибки репрезентативности свойственны только несплошным наблюдениям. Они характеризуют размер расхождений между величинами показателя, полученного в выборочной и генеральной совокупности в условиях одинаковой точности единичных наблюдений. Ошибки репрезентативности могут быть систематическими и случайными. Систематические ошибки возникают при нарушении установленных правил отбора единиц. Случайные ошибки репрезентативности обязаны своим возник-

новением недостаточно равномерным представлением в выборочной совокупности различных категорий единиц генеральной совокупности.

Величина случайной ошибки определяет надежность данных выборочного наблюдения, их пригодность для суждения о генеральной совокупности. При помощи формул теории вероятностей можно рассчитать возможную максимальную случайную ошибку – вероятный (стохастический) предел ошибки.

Максимально возможная ошибка – это такая величина отклонения выборочной средней (доли) от генеральной, вероятность превышения которой вследствие случайных причин в условиях данной выборки очень мала.

Величина случайной ошибки репрезентативности зависит от:

- степени колеблемости изучаемого признака в генеральной совокупности;
- способа формирования выборочной совокупности;
- объема выборки.

По **степени охвата единиц** исследуемой совокупности различают большие и малые выборки.

По **способу формирования** выборочной совокупности различают следующие виды выборочного наблюдения: простая случайная (собственно случайная) выборка, расслоенная (типическая или районированная), серийная, механическая, комбинированная, ступенчатая, многофазная.

Совокупность единиц, из которых производится отбор, принято называть **генеральной совокупностью.** Совокупность отобранных единиц из генеральной совокупности называется **выборочной совокупностью.**

N – объем генеральной совокупности (число входящих в нее единиц);

n – объем выборочной совокупности (число единиц, попавших в выборку);

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

$\tilde{x}$ – выборочная средняя (среднее значение признака в выборочной совокупности);

p – генеральная доля (доля единиц, обладающих данным признаком в генеральной совокупности);

w – выборочная доля (доля единиц, обладающих данным признаком в выборочной совокупности);

σ^2 – генеральная дисперсия (дисперсия признака в генеральной совокупности);

S^2 – выборочная дисперсия (дисперсия признака в выборочной совокупности);

σ – среднее квадратическое отклонение признака в генеральной совокупности;

S – среднее квадратическое отклонение признака в выборочной совокупности.

Простая случайная выборка

При простой случайной выборке отбор единиц в выборочную совокупность производится непосредственно из всей массы единиц генеральной совокупности в форме случайного отбора, при котором каждой единице генеральной совокупности обеспечивается одинаковая вероятность (возможность) быть выбранной. Единица отбора совпадает с единицей наблюдения. Случайный отбор осуществляется путем применения жеребьевки (лотереи) или путем использования таблиц случайных чисел.

Случайный отбор может быть проведен в двух формах: в форме возвратной (повторной) выборки и в форме безвозвратной (бесповторной) выборки. При повторном отборе вероятность попадания каждой единицы генеральной совокупности остается постоянной, так как после отбора какой-то единицы она снова возвращается в генеральную совокупность и может быть выбранной. При бесповторном отборе выбранная единица не возвращается в генеральную совокупность и вероятность попадания отдельных единиц в выборку все время изменяется (для оставшихся единиц она возрастает).

Применение простой случайной повторной выборки на практике весьма ограниченно; обычно используется бесповторная выборка.

Теорема П. Л. Чебышева утверждает принципиальную возможность определения генеральной средней по данным случайной повторной выборки. Теорема Чебышева дополняется **теоремой А. М. Ляпунова,** которая позволяет рассчитать максимальную ошибку выборочной средней при данном достаточно большом числе независимых наблюдений. Согласно этой теореме при достаточно большом числе независимых наблюдений в генеральной

совокупности с конечной средней и ограниченной дисперсией вероятность того, что расхождение между выборочной и генеральной средней $(\tilde{x} - \bar{x})$ не превзойдет по абсолютной величине некоторую величину $t\mu$, равна **интегралу Лапласа.** Это можно записать так:

$$P(|\tilde{x} - \bar{x}| \le t\mu) = \Phi(t);$$

$$\Phi(t) = \frac{1}{\sqrt{2\pi}} \int_{-t}^{+t} e^{-\frac{t^2}{2}} dt,$$

где $\Phi(t)$ – интеграл Лапласа (нормированная функция Лапласа).

Величина $t\mu$, обозначаемая Δ, называется **предельной ошибкой выборки.** Следовательно,

$$\Delta_{\tilde{x}} = t\mu_{\tilde{x}};\ \Delta_p = t\mu_p,$$

где $\Delta_{\tilde{x}}$ – предельная (максимально возможная) ошибка средней;
Δ_p – предельная (максимально возможная) ошибка доли;
μ – величина средней квадратической стандартной ошибки;
t – коэффициент кратности средней ошибки выборки, зависящий от вероятности, с которой гарантируется величина предельной ошибки.

В зависимости от принятой вероятности P определяется значение коэффициента кратности (t) по удвоенной нормированной функции Лапласа (см. приложение 3).

Величина средней ошибки в условиях большой выборки ($n > 30$) рассчитывается по известным из теории вероятностей формулам:

а) при случайной повторной выборке:

$$\mu_{\tilde{x}} = \sqrt{\frac{\sigma^2}{n}};\quad \mu_p = \sqrt{\frac{p\cdot(1-p)}{n}};$$

б) при случайной бесповторной выборке:

$$\mu_{\tilde{x}} = \sqrt{\frac{\sigma^2}{n}(1-\frac{n}{N})};$$

$$\mu_p = \sqrt{\frac{p\cdot(1-p)}{n}\cdot\left(1-\frac{n}{N}\right)}.$$

При расчете ошибок возникает существенное затруднение: величины σ и p по генеральной совокупности неизвестны. Эти величины в условиях большой выборки заменяют величинами S (выборочная дисперсия) и w (выборочная доля), рассчитанными по выборочным данным. В табл. 4.1 приведены формулы расчета ошибок простой случайной выборки.

Таблица 4.1

Формулы ошибок простой случайной выборки

	Способ отбора единиц	
	повторный	бесповторный
Средняя ошибка μ: для средней	$\mu_{\bar{x}} = \sqrt{\frac{S^2}{n}}$	$\mu_x = \sqrt{\frac{S^2}{n}\left(1-\frac{n}{N}\right)}$
для доли	$\mu_p = \sqrt{\frac{w(1-w)}{n}}$	$\mu_p = \sqrt{\frac{w(1-w)}{n}\left(1-\frac{n}{N}\right)}$
Предельная ошибка Δ: для средней	$\Delta_{\bar{x}} = t\sqrt{\frac{S^2}{n}}$	$\Delta_{\bar{x}} = t\sqrt{\frac{S^2}{n}\left(1-\frac{n}{N}\right)}$
для доли	$\Delta_p = t\sqrt{\frac{w(1-w)}{n}}$	$\Delta_p = t\sqrt{\frac{w(1-w)}{n}\left(1-\frac{n}{N}\right)}$

Формулы предельной ошибки позволяют решать задачи трех видов:

1. Определение пределов генеральных характеристик с заданной степенью надежности (доверительной вероятностью) на основе показателей, полученных по данным выборки.

Доверительные интервалы для генеральной средней –

$$\bar{x} = \tilde{x} \pm \Delta_{\bar{x}};$$
$$\tilde{x} - \Delta_{\bar{x}} \le \bar{x} \le \tilde{x} + \Delta_{\bar{x}}.$$

Доверительные интервалы для генеральной доли –

$$p = w \pm \Delta_p;$$

$$w - \Delta_p \le p \le w + \Delta_p.$$

2. Определение доверительной вероятности того, что генеральная характеристика может отличаться от выборочной не более чем на определенную заданную величину.

Доверительная вероятность является функцией от t, определяемой по формуле

$$t = \frac{\Delta_{\bar{x}}}{\mu_{\bar{x}}}.$$

По величине t определяется доверительная вероятность (приложение 3).

3. Определение необходимого объема выборки, который с практической вероятностью обеспечивает заданную точность выборки.

Для расчета объема выборки необходимо иметь следующие данные:

а) размер доверительной вероятности (P);

б) коэффициент t, зависящий от принятой вероятности (определяется по приложению 3);

в) величину σ^2 (или pq) в генеральной совокупности; они заменяются величинами, полученными в предшествующих обследованиях или при пробных выборках;

Таблица 4.2

Формулы для определения численности простой случайной выборки

	Способ отбора единиц	
	повторный	бесповторный
Численность выборки (n): для средней	$n = \frac{t^2 S^2}{\Delta_{\bar{x}}^2}$	$n = \frac{t^2 N S^2}{\Delta_{\bar{x}}^2 N + t^2 S^2}$
для доли[1]	$n = \frac{t^2 w(1-w)}{\Delta_p^2}$	$n = \frac{t^2 N w(1-w)}{\Delta_p^2 N + t^2 w(1-w)}$

[1] В случаях, когда частость w даже приблизительно неизвестна, в расчет вводят максимальную величину дисперсии доли, равную 0,25 (если $w = 0{,}5$, то $w(1-w) = 0{,}25$).

г) величину максимально допустимой ошибки ($\Delta_{\bar{x}}$ или Δ_p);
д) объем генеральной совокупности (N).

Необходимый объем выборки определяется на основе допустимой величины ошибки: $\Delta_{\bar{x}} = t\mu_{\bar{x}}$ или $\Delta_p = t\mu_p$.

В табл. 4.2 приведены формулы для расчета численности простой случайной выборки.

Расслоенная (типическая или районированная) выборка

В составе генеральной совокупности с различным уровнем изучаемого признака желательно обеспечить более равномерное представительство в выборочной совокупности различных типов. Эта цель достигается при применении расслоенной (типической или стратифицированной) выборки. Эту выборку применяют также в целях более равномерного представления в выборке различных районов, и в этом случае ее называют районированной выборкой.

При типической выборке неоднородная генеральная совокупность подразделяется на более однородные в отношении изучаемых признаков группы (типы, районы). По каждой группе определяются ее объем (N_i) и число подлежащих наблюдению единиц (n_i). Отбор обследуемых единиц производится в каждой группе при помощи одного из способов случайного отбора – повторного или бесповторного.

Общее число единиц выборочной совокупности распределяется между группами пропорционально численности групп в составе генеральной совокупности. Такой отбор называется **пропорциональным.**

N – общая численность единиц в генеральной совокупности

$$N = N_1 + N_2 + ... + N_k,$$

где $N_1, N_2, ..., N_k$ – численность отдельных групп генеральной совокупности;
n – общий объем выборочной совокупности.

Объем выборки для каждой группы –

$$n_i = n \cdot \frac{N_i}{N},$$

где $\frac{N_i}{N}$ – удельный вес данной (i-й) группы в генеральной совокупности;

$$n = n_1 + n_2 + ... + n_k.$$

Кроме пропорционального размещения по группам численности единиц выборочной совокупности применяется так называемое оптимальное размещение, при котором число наблюдений в группе определяется по формуле

$$n_i = \frac{N_i \sigma_i}{\sum_1^k N_i \sigma_i} \cdot n.$$

Формулы для расчета ошибок типической выборки приведены в табл. 4.3.

В табл. 4.3 приняты следующие *условные обозначения:*

$\overline{S^2}$ – средняя групповая выборочная дисперсия средней;

$$\overline{S^2} = \frac{\sum S_i^2 n_i}{\sum n_i};$$

S_i^2 – внутригрупповая дисперсия данной (i-й) группы в выборочной совокупности;

$\overline{w(1-w)}$ – средняя групповая выборочная дисперсия доли;

$$\overline{w(1-w)} = \frac{\sum w_i(1-w_i)n_i}{\sum n_i}.$$

Как видно из приведенных формул, величина стандартной ошибки типической выборки зависит только от точности определения групповых средних, т. е. от величины внутригрупповых дисперсий. Согласно правилу сложения дисперсий общая дисперсия слагается из межгрупповой дисперсии и средней из внутригрупповых дисперсий. Отсюда следует, что ошибка типической случайной выборки меньше, чем ошибка простой случайной выборки.

Предельная (максимально возможная) ошибка типической выборки:

$$\Delta_{\bar{x}} = t\mu_{\bar{x}}; \Delta_p = t\mu_p.$$

Таблица 4.3

Формулы ошибок типической выборки

	Способ отбора единиц	
	повторный	бесповторный
Средняя ошибка (μ): для средней: а) при пропорциональном размещении единиц	$\mu_{\tilde{x}} = \sqrt{\frac{\overline{S^2}}{n}}$	$\mu_{\tilde{x}} = \sqrt{\frac{\overline{S^2}}{n}(1-\frac{n}{N})}$
б) при оптимальном размещении единиц	$\mu_{\tilde{x}} = \frac{1}{N}\sqrt{\sum\frac{S_i^2 N_i^2}{n_i}}$	$\mu_{\tilde{x}} = \frac{1}{N}\sqrt{\sum\frac{S_i^2 N_i^2}{n_i}(1-\frac{n_i}{N_i})}$
для доли: а) при пропорциональном размещении единиц	$\mu_p = \sqrt{\frac{\overline{w(1-w)}}{n}}$	$\mu_p = \sqrt{\frac{\overline{w(1-w)}}{n}(1-\frac{n}{N})}$
б) при оптимальном размещении единиц	$\mu_p = \frac{1}{N}\sqrt{\sum\frac{w_i(1-w_i)N_i^2}{n_i}}$	$\mu_p = \frac{1}{N}\sqrt{\sum\frac{w_i(1-w_i)N_i^2}{n_i}(1-\frac{n_i}{N_i})}$

Необходимый объем выборки определяется на основе формулы и величины допустимой ошибки.

Серийная выборка

Сущность серийной выборки заключается в том, что вместо случайного отбора единиц совокупности осуществляется отбор групп (серий, гнезд). Внутри отобранных серий производится сплошное наблюдение. Серии (гнезда) состоят из единиц, связанных между собой или территориально, или организационно, или, наконец, во времени. Отбор серий может производиться в

порядке повторного и бесповторного отбора. Серии могут быть равновеликими и неравновеликими. На практике чаще применяется серийный отбор с равными сериями.

Стандартная ошибка при равновеликих сериях определяется по формулам, представленным в табл. 4.4.

Таблица 4.4

	Способ отбора серий	
	повторный	бесповторный
Средняя ошибка (μ): для средней	$\mu_{\tilde{x}} = \sqrt{\frac{\delta^2_{\tilde{x}}}{m}}$	$\mu_{\tilde{x}} = \sqrt{\frac{\delta^2_{\tilde{x}}}{m}(1-\frac{m}{M})}$
для доли	$\mu_p = \sqrt{\frac{\delta^2_w}{m}}$	$\mu_p = \sqrt{\frac{\delta^2_w}{m}(1-\frac{m}{M})}$

В табл. 4.4 приняты следующие *условные обозначения:*

$\delta^2_{\tilde{x}}$ – межгрупповая выборочная дисперсия средней;

$$\delta^2_{\tilde{x}} = \frac{\sum(\tilde{x}_i - \tilde{x}_0)^2}{m},$$

где $\tilde{x}_i$ – средний уровень признака в серии;
$\tilde{x}_0$ – средний уровень признака для всей выборочной совокупности;
m – число равных серий в выборочной совокупности;
М – число равных серий в генеральной совокупности;
δ^2_w – межгрупповая выборочная дисперсия доли;

$$\delta^2_w = \frac{\sum(w_i - w)^2}{m},$$

где w_i – доля единиц, обладающих данным признаком в серии;
w – доля единиц, обладающих данным признаком во всей выборочной совокупности.

Ошибка серийной выборки больше, чем при любом другом способе отбора. Тем не менее серийный отбор широко применяется на практике, что объясняется его организационными преимуществами.

Механическая выборка

Механическая выборка заключается в отборе единиц из генеральной совокупности через равные промежутки из определенного расположения их в генеральной совокупности (по алфавиту, в пространстве, последовательности появления во времени).

При организации механического отбора возникают две задачи:

- определение «шага отчета» (расстояния между отбираемыми единицами);
- выбор единицы, с которой надо начинать отчет.

«Шаг отчета» определяется путем деления численности генеральной совокупности на численность выборочной совокупности: $\frac{N}{n}$.

Выбор начала отчета рекомендуется производить путем случайного отбора из единиц первого интервала — первого «шага отчета». Механический отбор может осуществляться в самом процессе наблюдения, и его удобно применять тогда, когда выборочно наблюдается масса постепенно возникающих перед наблюдателем единиц (например, производят проверку каждой 10-й, 20-й и т. д. детали, обработанной на станке).

Если в генеральной совокупности единицы располагаются случайным образом по отношению к изучаемому признаку, то механический отбор можно рассматривать как разновидность случайного бесповторного отбора; поэтому для оценки ошибки механической выборки применяются формулы случайной бесповторной выборки:

$$\mu_{\bar{x}} = \sqrt{\frac{S^2}{n}\left(1-\frac{n}{N}\right)}; \qquad \Delta_{\bar{x}} = t \cdot \mu_{\bar{x}};$$

$$\mu_p = \sqrt{\frac{w(1-w)}{n}\left(1-\frac{n}{N}\right)}; \qquad \Delta_p = t \cdot \mu_p.$$

Комбинированная выборка

Комбинированная выборка предполагает использование нескольких способов выборки. Можно комбинировать, например, серийную выборку и случайную. В этом случае, разбив генеральную совокупность на серии (группы) и отобрав нужное число серий, производят случайную выборку единиц в серии. Такая комбинированная выборка может быть повторной (для групп и единиц) и бесповторной.

Средняя ошибка комбинированной выборки определяется по формулам:

при повторном отборе — $\mu_{\bar{x}} = \sqrt{\frac{S^2}{n} + \frac{\delta^2}{m}}$;

при бесповторном отборе — $\mu_{\bar{x}} = \sqrt{\frac{S^2}{n}\left(1-\frac{n}{N}\right) + \frac{\delta^2}{m}\left(1-\frac{m}{M}\right)}$.

Многоступенчатая выборка

Многоступенчатая выборка предполагает извлечение из генеральной совокупности сначала укрупненных групп единиц, затем групп, меньших по объему, и так до тех пор, пока не будут отобраны те группы (серии) или отдельные единицы, которые будут подвергнуты наблюдению. Выборка может быть двухступенчатой, когда генеральная совокупность разбивается на группы и производится отбор групп, а затем внутри групп — отбор единиц наблюдения. На обеих ступенях отбор может вестись в случайном порядке.

В отличие от типического отбора, где отбор производится из всех без исключения групп, при многоступенчатом отборе производится отбор самих групп, и, следовательно, не все они попадают в выборку.

Число ступеней отбора может быть и более трех. Если число ступеней отбора больше двух, то средняя ошибка выборки определяется по формуле

$$\mu = \sqrt{\mu_1^2 + \frac{\mu_2^2}{n_1} + \frac{\mu_3^2}{n_1 n_2} + \ldots},$$

где μ_1, μ_2, μ_3 — средние ошибки выборки на отдельных ступенях отбора;

n_1, n_2 — численность выборок на соответствующих ступенях.

Многофазная выборка

При многофазной выборке выборочные совокупности образуются так, что одни сведения собираются от всех единиц отбора, затем отбираются еще некоторые единицы, которые и обследуются по более широкой программе. Расчет ошибки многофазной выборки производится для каждой фазы в отдельности.

Малые выборки

Выборки, при которых наблюдением охватывается небольшое число единиц ($n < 30$), принято называть малыми выборками. Они обычно применяются в том случае, когда невозможно или нецелесообразно использовать большую выборку (исследование качества продукции, если это связано с ее разрушением, в частности на прочность, на продолжительность срока службы и т. д.).

Предельная ошибка малой выборки определяется по формуле

$$\Delta_{\text{М.В}} = t \cdot \mu_{\text{М.В}}.$$

Средняя ошибка малой выборки:

$$\mu_{\text{М.В}} = \sqrt{\frac{S^2}{n}},$$

где S^2 – дисперсия малой выборки.

$$S^2 = \frac{\sum (x - \tilde{x})^2}{n-1},$$

где $\tilde{x}$ – среднее значение признака по выборке;
$n - 1$ – число степеней свободы ($n - 1 = k$);
t – коэффициент доверия малой выборки, зависящий не только от заданной доверительной вероятности, но и от численности единиц выборки.

Вероятность того, что генеральная средняя находится в определенных границах, определяется по формуле

$$P\,[\tilde{x} - t\mu_{\text{М.В}} \leq \overline{x} \leq \tilde{x} + t\mu_{\text{М.В}}] = 2S_{(t)} - 1,$$

где $S_{(t)}$ – значение функции Стьюдента.

Для расчета коэффициента доверия t определяют значение функции $S_{(t)}$ по формуле

$$S_{(t)} = (P + 1) : 2.$$

Затем по таблице распределения Стьюдента (см. приложение 4) в зависимости от значения функции $S_{(t)}$ и числа степеней $k(n - 1)$ определяют значение t.

Функция $S_{(t)}$ используется также для определения вероятностей того, что фактическое нормированное отклонение

$\left(t_{\text{ф}} = \frac{\tilde{x} - \overline{x}}{\mu_{\text{М.В}}}\right)$ не превзойдет или превзойдет табличное значение.

Вероятность того, что фактическое отношение ($t_{\text{ф}}$) не превзойдет по абсолютной величине табличное значение (t), определяется по формуле

$$P\,[t_{\text{ф}} < |t|] = 2S_{(t)} - 1.$$

Вероятность того, что фактическое отношение превзойдет по абсолютной величине табличное значение

$$P\,[t_{\text{ф}} > |t|] = 2[1 - S_{(t)}].$$

Метод моментных наблюдений

Метод моментных наблюдений применяется для получения структуры затрат рабочего времени, характеристики использования оборудования. Сущность метода состоит в периодической фиксации состояния наблюдаемых единиц в заранее установленные или случайно выбранные моменты времени. При этом заранее составляется перечень всех возможных состояний процесса или видов затрат времени. По окончании наблюдения исследователь подсчитывает долю отметок о каждом состоянии или виде затрат времени в общем числе наблюдений; при этом считается, что доля времени, затраченного на данный вид работы, может быть оценена с помощью доли моментов, когда выполнялась эта работа, в общем числе наблюдений.

Средняя ошибка доли определяется по формуле простой случайной выборки:

$$\mu_p = \sqrt{\frac{w(1-w)}{n}},$$

где w — доля отметок о данном состоянии процесса;

n — количество моментов наблюдения.

Предельная ошибка $\Delta_p = t\mu_p$.

Для определения численности моментов наблюдения применяется следующая формула:

$$n = \frac{t^2(1-w)}{\Delta_p^2}.$$

Доверительная вероятность и величина допустимой ошибки устанавливаются исследователем; величина w, как правило, неизвестна, поэтому обычно ориентируются на наибольшую дисперсию, когда $w = 0,5$ [$w (1 - w) = 0,25$].

После определения числа наблюдений устанавливается необходимое число обходов как частное от деления количества наблюдений на число рабочих мест, подлежащих обследованию.

Следующим этапом является составление графика проведения наблюдения. Наблюдение за состоянием процесса осуществляется через определенные промежутки времени. При моментном наблюдении со случайным отбором моменты отбираются при помощи таблиц случайных чисел. Второй способ целесообразен в тех случаях, когда наблюдение для объекта должно быть неожиданным.

Проверка гипотезы о существенности расхождения средних (долей)

К расчетам ошибок случайной выборки прибегают в тех случаях, когда необходимо сравнить между собой средние величины данного признака по двум совокупностям, т. е. определить, существенно ли расхождение между двумя выборочными средними или несущественно; следовательно, превосходит или не превосходит максимальной величины случайного расхождения, которое можно ожидать с определенной вероятностью. Другими словами, можно ли считать, что генеральные средние в двух подгруппах

одинаковы, и эти подгруппы можно объединить в одну группу и характеризовать последнюю общей средней?

Для ответа на вопрос определяют среднюю (стандартную) случайную ошибку разности двух выборочных средних (μ_δ); для двух независимых выборок она определяется по формуле

$$\mu_\delta = \sqrt{\frac{S_1^2}{n_1} + \frac{S_2^2}{n_2}} = \sqrt{\mu_1^2 + \mu_2^2},$$

где S^2_1 и S^2_2 — выборочные дисперсии соответственно в первой и во второй выборках;

μ^2_1 и μ^2_2 — стандартная ошибка средней соответственно в первой и во второй выборках.

$$S_1^2 = \frac{\sum(x-\tilde{x}_1)^2 \cdot f}{n_1 - 1};\quad S_2^2 = \frac{\sum(x-\bar{x})^2 \cdot f}{n_2 - 1};\quad \frac{|\tilde{x}_1 - \tilde{x}_2|}{\mu_\delta} = t_{\text{расч}}.$$

По таблице Лапласа определяют по полученному значению t соответствующую вероятность (P) (см. приложение 3). Если вероятность значительна, то нулевая гипотеза, т. е. предположение об отсутствии существенного различия, не опровергается.

Получить ответ на выдвинутую нулевую гипотезу можно иначе. Зная величину ошибки разности выборочных двух средних, можно с заданной вероятностью указать предел возможных расхождений двух выборочных средних. Если $t_{\text{расч}} > t_{\text{табл}}$ при определенной заданной доверительной вероятности, то это свидетельствует о том, что нулевая гипотеза не подтверждается.

При малом объеме выборок используется распределение Стьюдента (см. приложение 4).

Проверка гипотезы о существенности расхождения двух выборочных долей осуществляется аналогично.

Ошибка разности двух долей при справедливости нулевой гипотезы исчисляется по формуле

$$\mu_\delta = \sqrt{\frac{w_1(1-w_1)}{n_1} + \frac{w_2(1-w_2)}{n_2}}.$$

Расчетное значение t-критерия: $t_{\text{расч}} = \dfrac{|w_1 - w_2|}{\mu_\delta}$.

При заданной доверительной вероятности определяется $t_{табл}$. Если $t_{расч} > t_{табл}$, то это говорит о том, что нулевая гипотеза не подтверждается.

Элементы дисперсионного анализа

Дисперсионный анализ является одним из методов изучения влияния одного или нескольких факторных признаков на результативный признак. В зависимости от количества факторов дисперсионный анализ подразделяется на однофакторный и многофакторный. Ниже рассмотрено применение дисперсионного анализа для случая однофакторного комплекса.

В основе дисперсионного анализа лежит *расчленение общей вариации изучаемого признака по источникам ее происхождения на два вида вариации:*

- систематическую вариацию, которая обусловлена изменением признака-фактора;
- остаточную (случайную) вариацию, обусловленную действием прочих, случайных, не связанных с данным фактором обстоятельств.

Для разграничения этих вариаций всю совокупность наблюдавшихся единиц разбивают на группы (классы) по факторному признаку и исчисляют средние результативного признака по группам.

Групповые средние

$$\tilde{x}_i = \frac{\sum x_i}{n_i};$$

общая средняя

$$\tilde{x}_0 = \frac{\sum x_i}{n},$$

где x_i – индивидуальные значения признака в группе;
n_i – число единиц, входящих в группу;
n – общее число наблюдений.

Если сравнение групповых средних показывает определенное различие в их уровне, то необходимо установить, является ли это различие существенным и вызвано ли оно влиянием признака-фактора.

Для ответа на поставленный вопрос определяют два показателя дисперсии:

1) показатель $S^2{}_1$, характеризующий колеблемость групповых средних вокруг общей средней (межгрупповая дисперсия);

2) показатель $S^2{}_2$, отражающий остаточную, внутригрупповую дисперсию. Полученные показатели сравнивают, получая фактическое дисперсионное отношение:

$$F_{расч} = \frac{S_1^2}{S_2^2}.$$

При дисперсионном анализе межгрупповую $S^2{}_1$ и внутригрупповую $S^2{}_2$ дисперсии определяют путем деления суммы квадратов отклонений на соответствующее число степеней свободы:

$$S^2{}_1 = \frac{\text{Сумма квадратов отклонений групповых средних от общей средней}}{\text{Число степеней вариации между группами}};$$

$$S_1^2 = \frac{\sum(\tilde{x}_i - \tilde{x}_0)^2 \cdot n_i}{K_1},$$

где n_i – число единиц в группе;
$K_1 = m - 1$;
m – число групповых средних (число выделенных групп по признаку-фактору);

$$S^2{}_2 = \frac{\text{Сумма квадратов отклонений индивидуальных значений от групповых средних}}{\text{Число степеней вариации внутри групп}};$$

$$S_2^2 = \frac{\sum\sum(x - \tilde{x}_i)^2}{K_2},$$

где $K_2 = n - m$.

По **таблице F-распределения Р. Фишера** (см. приложение 5) при определенном уровне значимости (или доверительной вероятности)[1] и числе степеней свободы (K_1 и K_2) определяется табличное дисперсионное отношение ($F_{табл}$).

[1] Доверительная вероятность $P = 1 - \alpha$, где α – уровень значимости.

Если $F_{расч} > F_{табл}$, то следует считать, что гипотеза о влиянии признака-фактора не опровергается.

4.1. Решение типовых задач

4.1. Из партии электроламп взята 20%-ная случайная бесповторная выборка для определения среднего веса спирали.

Результаты выборки следующие (табл. 4.5).

Таблица 4.5

Вес, мг	38 – 40	40 – 42	42 – 44	44 – 46
Число спиралей	15	30	45	10

Определить с вероятностью 0,95 доверительные пределы, в которых лежит средний вес спирали, для всей партии электроламп.

Решение

Доверительные интервалы для генеральной средней с вероятностью P:

$$\tilde{x} - \Delta_{\bar{x}} \le \bar{x} \le \tilde{x} + \Delta_{\bar{x}},$$

где $\tilde{x}$ – средний уровень признака по выборке:

$$\tilde{x} = \frac{\sum x'f}{\sum f} = \frac{39 \cdot 15 + 41 \cdot 30 + 43 \cdot 45 + 45 \cdot 10}{15 + 30 + 45 + 10} = \frac{4200}{100} = 42{,}0 \text{ мг}$$

$$\Delta_{\bar{x}} = t\mu_{\bar{x}} = t\sqrt{\frac{S^2}{n} \cdot \left(1 - \frac{n}{N}\right)}; \quad N = \frac{100}{0{,}2} = 500.$$

При вероятности $P = 0{,}95$ $t = 1{,}96$ (по таблице приложения 3).

$$S^2 = \frac{\sum (x' - \tilde{x})^2 \cdot f}{\sum f} =$$

$$= \frac{(39-42)^2 \cdot 15 + (41-42)^2 \cdot 30 + (43-42)^2 \cdot 45 + (45-42)^2 \cdot 10}{100} =$$

$$= \frac{300}{100} = 3{,}0.$$

$$\Delta_{\bar{x}} = 1{,}96\sqrt{\frac{3{,}0}{100} \cdot \left(1 - \frac{100}{500}\right)} = 0{,}3 \text{ мг}.$$

Доверительные интервалы для генеральной средней с вероятностью $P = 0{,}95$:

$$42{,}0 - 0{,}3 \le \bar{x} \le 42{,}0 + 0{,}3; \; 41{,}7 \text{ мг} \le \bar{x} \le 42{,}3 \text{ мг}.$$

4.2. На заводе электроламп из партии продукции в количестве 16000 шт. ламп взято на выборку 1600 шт. (случайный, бесповторный отбор), из которых 40 шт. оказались бракованными.

Определить с вероятностью 0,997 пределы, в которых будет находиться процент брака для всей партии продукции.

Решение

Определяется доля бракованной продукции по выборке:

$$w = \frac{40}{1600} = 0{,}025, \text{ или } 2{,}5\%.$$

При вероятности $P = 0{,}997$ $t = 3{,}0$ (по таблице приложения 3). Размер предельной ошибки

$$\Delta_p = t\mu_p = t\sqrt{\frac{w(1-w)}{n} \cdot \left(1 - \frac{n}{N}\right)} = 3{,}0\sqrt{\frac{0{,}025 \cdot (1 - 0{,}025)}{16000} \cdot \left(1 - \frac{1600}{16000}\right)} =$$

$$= 3{,}0 \cdot 0{,}0037 = 0{,}011, \text{ или } 1{,}1\%.$$

Доверительные интервалы для генеральной доли с вероятностью $P = 0{,}997$

$$w - \Delta_p \le p \le w + \Delta_p; \; 2{,}5 - 1{,}1 \le p \le 2{,}5 + 1{,}1;$$

$$1{,}4\% \le p \le 3{,}6\%.$$

4.3. По городской телефонной сети в порядке случайной выборки (механический) отбор произвели 100 наблюдений и установили среднюю продолжительность одного телефонного разговора 5 мин. при среднем квадратическом отклонении 2 мин.

Какова вероятность того, что ошибка репрезентативности при определении средней продолжительности телефонного разговора не превысит 18 с?

Решение

По условию задачи известны:
объем выборки – $n = 100$;
выборочная средняя – $\tilde{x} = 5$ мин.;
выборочное среднее квадратическое отклонение – $S = 2$ мин.;
предельная ошибка выборки – $\Delta_{\bar{x}} = 18$ сек. = 0,3 мин.

$$\Delta_{\bar{x}} = t \cdot \mu_{\bar{x}}; \quad \mu_{\bar{x}} \sqrt{\frac{S^2}{n}} = \sqrt{\frac{2^2}{100}} = 0{,}2 \text{ мин.}$$

$$t = \frac{\Delta_{\bar{x}}}{\mu_{\bar{x}}} = \frac{0{,}3}{0{,}2} = 1{,}5.$$

Затем по таблице (см. приложение 3) на основе значения t определяется вероятность того, что ошибка не превысит заданной величины.

При $t = 1{,}5$ вероятность $P = 0{,}866$.

4.4. На основе выборочного обследования в отделении связи города предполагается определить долю писем частных лиц в общем объеме отправляемой корреспонденции. Никаких предварительных данных об удельном весе этих писем в общей массе отправляемой корреспонденции не имеется.

Определить численность выборки, если результаты выборки дать с точностью до 1% и гарантировать это с вероятностью 0,95.

Решение

По условию задачи известны:
размер допустимой (предельной) ошибки – $\Delta_p = 1\%$, или 0,01;
принятая вероятность – $P = 0{,}95$;
при $P = 0{,}95$ $t = 1{,}96$ (см. приложение 3).
Необходимая численность выборки:

$$n = \frac{t^2 \cdot w(1-w)}{\Delta_p^2}.$$

Так как значение w не дано, то следует ориентироваться на наибольшую дисперсию, которой соответствует значение $w = 0{,}5$.

$$n = \frac{1{,}96^2 \cdot 0{,}5 \cdot 0{,}5}{0{,}01^2} = 9604.$$

Таким образом, чтобы с заданной точностью определить долю частных писем в общем объеме отправляемой корреспонденции, необходимо в порядке случайной выборки отобрать 9604 письма.

4.5. На предприятии в порядке случайной бесповторной выборки было опрошено 100 рабочих из 1000 и получены следующие данные об их доходе за октябрь (табл. 4.6).

Таблица 4.6

Месячный доход, руб.	1200 – 2000	2000 – 2800	2800 – 3600	3600 – 4400
Число рабочих	12	60	20	8

Определить:

1) среднемесячный размер дохода у работников данного предприятия, гарантируя результат с вероятностью 0,997;

2) долю рабочих предприятия, имеющих месячный доход 2800 руб. и выше, гарантируя результат с вероятностью 0,954;

3) необходимую численность выборки при определении среднего месячного дохода работников предприятия, чтобы с вероятностью 0,954 предельная ошибка выборки не превышала 100 руб.;

4) необходимую численность выборки при определении доли рабочих с размером месячного дохода 2800 руб. и выше, чтобы с вероятностью 0,954 предельная ошибка не превышала 4%.

Решение

1. Доверительный интервал среднего размера месячного дохода работников предприятия: $\tilde{x} - \Delta_{\bar{x}} \le \bar{x} \le \tilde{x} + \Delta_{\bar{x}}$.

Средний месячный доход по выборке

$$\tilde{x} = \frac{\sum x'f}{\sum f}; \quad x = \frac{1600 \cdot 12 + 2400 \cdot 60 + 3200 \cdot 20 + 4000 \cdot 8}{12 + 60 + 20 + 8} = 2592{,}0 \text{ руб.}$$

10–1622

Предельная ошибка выборки $\Delta_{\bar{x}} = t\mu_{\bar{x}}$.

При вероятности $P = 0{,}997$ $t = 3{,}0$ (см. приложение 3)

$$\Delta_{\bar{x}} = t\mu_{\bar{x}} = t\sqrt{\frac{S^2}{n}(1-\frac{n}{N})}.$$

$$S^2 = \frac{\sum(x'-\bar{x})^2 f}{\sum f} =$$

$$= \frac{(1600-2592)^2 12 + (2400-2592)^2 60 + (3200-2592)^2 20 + (4000-2592)^2 8}{12+60+20+8} =$$

$$=372736.$$

$$\Delta_{\bar{x}} = 3\sqrt{\frac{372736}{100}(1-\frac{100}{1000})} = 173{,}76 \text{ руб.}$$

$2592{,}0 - 173{,}76 \le \bar{x} \le 2592 + 173{,}76$; $2418{,}24$ руб. $\le \bar{x} \le 2765{,}76$ руб.

2. w – доля рабочих, имеющих размер месячного дохода 2800 руб. и выше:

$$w = \frac{20+8}{100} = 0{,}28.$$

Предельная ошибка доли $\Delta_p = t \cdot \mu_p = t\sqrt{\frac{w(1-w)}{n}(1-\frac{n}{N})}$.

При вероятности $P = 0{,}954$ $t = 2$.

$$\Delta_p = 2\sqrt{\frac{0{,}28(1-0{,}28)}{100}(1-\frac{100}{1000})} = 0{,}085.$$

Доверительные интервалы для генеральной доли:

$$w - \Delta_p \le p \le w + \Delta_p;$$

$$0{,}28 - 0{,}085 \le p \le 0{,}28 + 0{,}085;$$

$$0{,}195 \le p \le 0{,}365.$$

3. Необходимая численность выборки для определения среднего месячного дохода определяется по формуле

$$n = \frac{t^2 N S^2}{\Delta_{\bar{x}}^2 N + t^2 S^2}.$$

По условию задачи известны:

при вероятности $P = 0{,}954$ $t = 2$ (см. приложение 3);

$\Delta_{\bar{x}} = 100$ руб.; $S^2 = 372\,736$ (по данным предыдущей выборки).

$$n = \frac{2^2 \cdot 1000 \cdot 372736}{10^2 \cdot 1000 + 2^2 \cdot 372736} = 142 \text{ чел.}$$

4. Необходимая численность выборки для определения доли рабочих, имеющих доход 2800 руб. и выше, определяется по формуле

$$n = \frac{t^2 N w(1-w)}{\Delta_p^2 N + t^2 w(1-w)}.$$

По условию задачи известны:

$\Delta_p = 4\%$, или $0{,}04$; при вероятности $P = 0{,}954$ $t = 2$;

$w = 0{,}28$ (по данным предыдущей выборки).

$$n = \frac{2^2 \cdot 1000 \cdot 0{,}28(1-0{,}28)}{0{,}04^2 \cdot 1000 + 2^2 \cdot 0{,}28(1-0{,}28)} = 480 \text{ чел.}$$

4.6. Операция шлифования при обработке детали № 312 производится в цехе на трех станках. Для определения процента брака для всей партии продукции, выработанной за день, проведена расслоенная (типическая) 10%-ная выборка. Отбор деталей из выработки каждого станка – случайный бесповторный; объем выборки пропорционален размеру выпуска. На первом станке было обработано 1700 деталей, на втором – 2000, на третьем – 1800. Число забракованных деталей в выборке: по первому станку – 2, по второму – 3, по третьему – 3.

Определить:

1) доверительные интервалы, в которых с вероятностью 0,95 заключен процент брака для всей партии продукции;

2) вероятность того, что процент брака для всей партии продукции отличается от полученного по выборке не более чем на 0,6%.

Решение

1. Общий объем генеральной совокупности

$$N = N_1 + N_2 + N_3 = 1700 + 2000 + 1800 = 5500 \text{ деталей.}$$

Численность выборки

$$n = \frac{5500 \cdot 10}{100} = 550.$$

Численность выборки по станкам

$$n_1 = 170 \text{ деталей; } n_2 = 200 \text{ деталей; } n_3 = 180 \text{ деталей.}$$

$$(n_i = n \cdot \frac{N_i}{N}).$$

Доверительные интервалы процента брака для всей партии продукции:

$$w - \Delta_p \le p \le w + \Delta_p,$$

где w — процент брака для всей выборочной совокупности;

$$w = \frac{\text{Общее количество забракованных деталей}}{\text{Общая численность выборочной совокупности}} \cdot 100.$$

Δ_p — предельная ошибка выборки.

$$w = \frac{2+3+3}{550} \cdot 100 = 1{,}45\%.$$

$$\Delta_p = t\mu_p; \quad \mu_p = \sqrt{\frac{\overline{w(1-w)}}{n}(1-\frac{n}{N})},$$

где $\overline{w(1-w)}$ — среднегрупповая выборочная дисперсия доли.

$$\overline{w(1-w)} = \frac{\sum w_i(1-w_i)n_i}{\sum n_i} =$$

$$= \frac{0{,}0118(1-0{,}0118)\cdot 170 + 0{,}015(1-0{,}015)\cdot 200 + 0{,}0167(1-0{,}0167)\cdot 180}{170+200+180} =$$

$$= 0{,}0144,$$

где w_i — доля брака для каждого станка.

$$w_1 = \frac{2}{170} = 0{,}0118; \quad w_2 = \frac{3}{200} = 0{,}015; \quad w_3 = \frac{3}{180} = 0{,}0167.$$

$$\mu_p = \sqrt{\frac{0{,}0144}{550}(1-\frac{500}{5500})} = 0{,}0048, \text{ или } 0{,}48\%.$$

При вероятности $P = 0{,}95$ $t = 1{,}96$ (см. приложение 3).

$$\Delta_p = 1{,}96 \cdot 0{,}48 = 0{,}94\%.$$

$$w - \Delta_p \le p \le w + \Delta_p; \ 1{,}45 - 0{,}94 \le p \le 1{,}45 + 0{,}94; \ 0{,}51\% \le p \le 2{,}39\%.$$

2. Для решения второго задания известна допустимая ошибка $\Delta_p = 0{,}6\%$, или 0,006.

$$\Delta_p = t\mu_p; \text{ отсюда } t = \frac{\Delta_p}{\mu_p} = \frac{0{,}006}{0{,}0048} = 1{,}25.$$

$\mu_p = 0{,}0048$ (см. решение первого задания).

Величина $t = 1{,}25$; ей соответствует вероятность $P = 0{,}7887$ (см. приложение 3).

4.7. При контрольной проверке качества апельсинов проведена 10%-ная серийная выборка. Из партии, содержащей 50 ящиков апельсинов (вес ящиков одинаков), методом механического отбора взято 5 ящиков. В результате сплошного обследования находящихся в ящике апельсинов получили данные об удельном весе бракованных апельсинов. Результаты следующие (табл. 4.7).

Таблица 4.7

№ ящика, попавшего в выборку	1	2	3	4	5
Удельный вес бракованной продукции, %	1,2	1,8	2,0	1,0	1,5

Требуется с вероятностью 0,95 установить доверительные интервалы удельного веса бракованной продукции для всей партии апельсинов.

Решение

Для установления доверительного интервала, в котором для всей партии поставки находится доля бракованной продукции, используется формула $p = w \pm \Delta_p$.

$$\Delta_p = t\mu_p = t\sqrt{\frac{\delta_w^2}{m}(1-\frac{m}{M})},$$

где δ_w^2 — межсерийная (межгрупповая) выборочная дисперсия доли;
m — число ящиков, попавших в выборку;
M — общее число ящиков.

При вероятности $P = 0,95$ $t = 1,96$ (см. приложение 3).

$$w = \frac{1,2+1,8+2,0+1,0+1,5}{5} = 1,5\%, \text{ или } 0,015$$

(при расчете использована простая арифметическая, так как вес ящиков одинаков).

$$\delta_w^2 = \frac{\sum(w_i-w)^2}{m} = \frac{(0,012-0,015)^2+(0,018-0,015)^2+}{5}$$

$$\frac{+(0,020-0,015)^2+(0,010-0,015)^2+(0,015-0,015)^2}{5} = 0,0000136.$$

$$\Delta_P = 1,96\sqrt{\frac{0,0000136}{5}(1-\frac{5}{50})} = 0,003, \text{ или } 0,3\%;$$

$$P = 1,5 \pm 0,3; \quad 1,2\% \le P \le 1,8\%.$$

4.8. Из партии электроламп произведена малая выборка (отбор случайный, бесповторный) для определения продолжительности службы ламп. Результаты выборки следующие (табл. 4.8).

Таблица 4.8

№ лампы	1	2	3	4	5	6	7	8	9	10
Срок горения, ч.	1450	1370	1250	1400	1360	1420	1400	1320	1300	1430

На основе приведенных данных *требуется*:

1) определить доверительные интервалы, в которых заключена средняя продолжительность службы ламп для всей партии, гарантируя результат с вероятностью 0,99;

2) определить вероятность того, что средний срок службы ламп для всей партии отличается от полученного по выборке не более чем на 40 ч.

Решение

1. Доверительные интервалы для генеральной средней:

$$\tilde{x} - \Delta_{\text{М.В}} \le \bar{x} \le \tilde{x} + \Delta_{\text{М.В}};$$

$$\Delta_{\text{М.В}} = t \cdot \mu_{\text{М.В}}; \; \mu_{\text{М.В}} = \frac{S}{\sqrt{n}}; \; S = \sqrt{\frac{\sum(x-\tilde{x})^2}{n-1}}.$$

Для расчета S использована вспомогательная табл. 4.9.

$$\tilde{x} = \frac{13\ 700}{10} = 1370 \text{ ч}; \; S = \sqrt{\frac{36\ 200}{10-1}} = 63,42 \text{ ч};$$

$$\mu_{\text{М.В}} = \frac{63,42}{\sqrt{10}} = 20,06.$$

Для определения t сначала исчисляется $S_{(t)}$:

$$P = 2S_{(t)} - 1;\ 2S_{(t)} - 1 = 0{,}99;\ 2S_{(t)} = 1{,}99;\ S_{(t)} = 0{,}995.$$

Таблица 4.9

№ лампы	Срок горения (ч), x	$x - \tilde{x}$	$(x - \tilde{x})^2$
1	1450	80	6400
2	1370	0	0
3	1250	−120	14400
4	1400	30	900
5	1360	−10	100
6	1420	50	2500
7	1400	30	900
8	1320	−50	2500
9	1300	−70	4900
10	1430	60	3600
Итого	13 700	0	36 200

По таблице (см. приложение 4) при $k = 9$ $(10 - 1)$ и $S_{(t)} = 0{,}995$ $t = 3{,}2$.

Тогда $\bar{x} = \tilde{x} \pm t \cdot \mu_{\text{М.В}} = 1370 \pm 3{,}2 \cdot 20{,}06$.

Следовательно, с вероятностью 0,99 можно утверждать, что генеральная средняя колеблется в пределах от 1305,8 до 1434,2 ч.

2. Для определения вероятности отклонения генеральной средней от выборочной средней не более чем на 40 ч имеются следующие данные:

$$\Delta_{\text{М.В}} = \tilde{x} - \bar{x} = 40 \text{ ч};\quad \mu_{\text{М.В}} = 20{,}06 \text{ (см. п. 1).}$$

Отсюда

$$t_{\Phi} = \frac{\tilde{x} - \bar{x}}{\mu_{\text{М.В}}} = \frac{40}{20{,}06} = 1{,}994 \approx 2{,}0.$$

Пользуясь таблицей распределения Стьюдента (см. приложение 4), определяется значение функции $S_{(t)}$ при $t = 2{,}0$ и $k = 10 - 1 = 9$; $S_{(t)} = 0{,}962$.

Вероятность того, что отклонение генеральной средней от выборочной средней не превзойдет 40 ч:

$$P\,[t_{\Phi} < |t|] = 2S_{(t)} - 1 = 2 \cdot 0{,}962 - 1 = 0{,}924.$$

Вероятность того, что ошибка будет превышать 40 ч:

$$P\,[t_{\Phi} > |t|] = 2[1 - S_{(t)}] = 2\,[1 - 0{,}962] = 0{,}076.$$

Следовательно, вероятность отклонения генеральной средней от выборочной средней по абсолютной величине, превышающей 40 ч, очень мала.

4.9. Обработка детали № 427 производится в цехе на двух однотипных станках. При выборочном наблюдении (механический отбор единиц) были зарегистрированы следующие затраты на обработку одной детали (табл. 4.10).

Таблица 4.10

Затраты времени на одну деталь, мин.	Число деталей	
	станок № 1	станок № 2
1,5 – 2,5	7	–
2,5 – 3,5	10	12
3,5 – 4,5	15	17
4,5 – 5,5	8	11
Итого	40	40

Определить на основе приведенных данных, существенно ли расхождение в затратах времени на обработку одной детали для этих двух станков, гарантируя результат с вероятностью 0,95.

Решение

Для ответа на поставленный вопрос определяется отношение

$$\frac{|\tilde{x}_1 - \tilde{x}_2|}{\mu_\delta} = t_{\text{расч}}$$

где μ_δ – стандартная ошибка разности двух выборочных средних;

$$\mu_\delta = \sqrt{\frac{S_1^2}{n_1} + \frac{S_1^2}{n_2}},$$

где S^2_1 и S^2_2 – выборочные дисперсии для первой и второй выборки.

$$S_1^2 = \frac{\sum (x - \tilde{x}_1)^2 f}{n_1 - 1}; \quad S_2^2 = \frac{\sum (x - \tilde{x}_2)^2 f}{n_2 - 1}.$$

Вспомогательные расчеты для определения средних и дисперсий выполнены в табл. 4.11.

Таблица 4.11

Вспомогательная таблица для расчета $\tilde{x}$ и S^2

Затраты времени на 1 деталь, мин.	Станок № 1				
	число деталей, f	x'	$x'f$	$x' - \tilde{x}_2$	$(x' - \tilde{x}_2)^2 f$
1,5–2,5	7	2	14	−1,6	17,92
2,5–3,5	10	3	30	−0,6	3,6
3,5–4,5	15	4	60	+,04	2,4
4,5–5,5	8	5	40	1,4	15,68
Итого	40		144		39,6

Продолжение

Затраты времени на 1 деталь, мин.	Станок № 2				
	число деталей, f	x'	$x'f$	$x' - \tilde{x}_1$	$(x' - \tilde{x}_1)^2 f$
1,5–2,5	–	–	–	–	–
2,5–3,5	12	3	36	0,98	11,52
3,5–4,5	17	4	68	0,02	0,007
4,5–5,5	11	5	55	1,02	5,20
Итого	40		159		16,73

$$\tilde{x}_1 = \frac{\sum x'f}{\sum f} = \frac{144}{40} = 3{,}6 \text{ мин.}; \quad S_1^2 = \frac{39{,}6}{40-1} = 1{,}015;$$

$$\tilde{x}_2 = \frac{\sum x'f}{\sum f} = \frac{159}{40} = 3{,}98 \text{ мин.}; \quad S_2^2 = \frac{16{,}73}{40-1} = 0{,}429;$$

$$\mu_\delta = \sqrt{\frac{1{,}015}{40} + \frac{0{,}429}{40}} = 0{,}19; \quad t_{\text{расч}} = \frac{|3{,}6 - 3{,}98|}{0{,}19} = 2{,}0.$$

При заданной доверительной вероятности $P = 0{,}95$ $t_{\text{табл}} = 1{,}96$ (по таблице Лапласа, приложение 3).

$t_{\text{расч}} > t_{\text{табл}}$ – это свидетельствует о том, что нулевая гипотеза не подтверждается, т. е. расхождение между средними затратами времени на обработку одной детали на двух станках существенно и не может быть объяснено случайностями выборки.

4.10. Известны результаты выборочного обследования пробега автомобильных шин нового типа в различных условиях эксплуатации (табл. 4.12).

Таблица 4.12

№, п/п	Условия эксплуатации	Пробег шин, тыс. км	№ п/п	Условия эксплуатации	Пробег шин, тыс. км
1	Загородные	54,2	13	Загородные	56,6
2	Городские	70,5	14	Смешанные	60,5
3	Смешанные	58,9	15	Городские	70,3
4	Городские	71,8	16	Загородные	55,0
5	Смешанные	59,1	17	Смешанные	58,4
6	Городские	69,8	18	Городские	69,1
7	Загородные	58,8	19	Городские	72,0
8	Городские	58,9	20	Смешанные	59,0
9	Городские	68,7	21	Загородные	56,4
10	Смешанные	60,1	22	Городские	58,7
11	Городские	72,1	23	Смешанные	61,8
12	Смешанные	62,2	24	Городские	66,2

Установить, существует ли зависимость между условиями эксплуатации и величиной пробега шин, гарантируя результат с вероятностью 0,95.

Решение

Для выявления влияния факторного признака (условий эксплуатации) на результативный признак (величину пробега шин) используется дисперсионный анализ.

Всю совокупность единиц подразделяем по факторному признаку на группы (результаты группировки представлены в табл. 4.13).

Таблица 4.13

Пробег шин в различных условиях эксплуатации

Условия эксплуатации	Пробег шин, тыс. км
Городские	70,5; 71,8; 69,8; 58,9; 68,7; 72,1; 70,3; 69,1; 72,0; 58,7; 66,2.
Смешанные	58,9; 59,1; 60,1; 62,2; 60,5; 58,4; 59,0; 61,8.
Загородные	54,2; 58,8; 56,6; 55,0; 56,4.

Для каждой группы определяется средний пробег шин:

городские условия эксплуатации $\tilde{x}_1 = \frac{\sum x}{n} = \frac{748{,}1}{11} = 68{,}0$ тыс. км;

смешанные условия эксплуатации $\tilde{x}_2 = \frac{\sum x}{n} = \frac{480}{8} = 60{,}0$ тыс. км;

загородные условия эксплуатации $\tilde{x}_3 = \frac{\sum x}{n} = \frac{281}{5} = 56{,}2$ тыс. км.

Общая средняя $\tilde{x}_0 = \frac{\sum x}{n} = \frac{748{,}1 + 480 + 281}{11 + 8 + 5} = \frac{1509{,}1}{24} = 62{,}9$ тыс. км.

Полученные средние величины пробега шин для разных условий эксплуатации отличаются друг от друга. Для того чтобы установить, является ли это различие существенным и вызвано различными условиями эксплуатации, определяется фактическое дисперсионное отношение:

$$F_{\text{расч}} = \frac{S_1^2}{S_2^2},$$

где S_1^2 – межгрупповая дисперсия;

S_2^2 – внутригрупповая (остаточная) дисперсия.

$$S_1^2 = \frac{\sum(\tilde{x}_i - \tilde{x}_0)^2 n_i}{k_1} = \frac{(68{,}0 - 62{,}9)^2 11 + (60 - 62{,}9)^2 8 + (56{,}2 - 62{,}9)^2 \cdot 5}{2} =$$

$$= \frac{597{,}84}{2} = 298{,}92.$$

$$k_1 = 3 - 1 = 2.$$

$$S_2^2 = \frac{\sum\sum(x - \tilde{x}_i)^2}{k_2}.$$

Вспомогательные расчеты для определения S_2^2 выполнены в табл. 4.14; $k_2 = 24 - 3 = 21$.

Таблица 4.14

Вспомогательная таблица для расчета S_2^2

№ п/п	Условия эксплуатации городские			Условия эксплуатации смешанные			Условия эксплуатации загородные		
	x	$x - \tilde{x}_1$	$(x-\tilde{x}_1)^2$	x	$x - \tilde{x}_2$	$(x-\tilde{x}_2)^2$	x	$x-\tilde{x}_3$	$(x-\tilde{x}_3)^2$
1	70,5	2,5	6,25	58,9	–1,1	1,21	54,2	–2,0	4,00
2	71,8	3,8	14,44	59,1	–0,9	0,81	58,8	2,6	6,76
3	69,8	1,8	3,24	60,1	0,1	0,01	56,6	0,4	0,16
4	58,9	–0,1	82,81	62,2	2,2	4,84	55,0	–1,2	1,44
5	68,7	0,7	0,49	60,5	0,5	0,25	56,4	0,2	0,04
6	72,1	4,1	16,81	58,4	–1,6	2,56			
7	70,3	2,3	5,29	59,0	–1,0	1,00			
8	69,1	1,1	1,21	61,8	1,8	3,24			
9	72,0	4,0	16,00						
10	58,7	–9,3	86,49						
11	66,2	–1,8	3,24						
Итого			236,27			13,92			12,40

$$S_2^2 = \frac{236{,}27 + 13{,}92 + 12{,}40}{21} = \frac{262{,}59}{21} = 12{,}50;$$

$$F_{\text{расч}} = \frac{298{,}92}{12{,}50} = 23{,}91.$$

При вероятности $P = 0,95$ (уровень значимости 0,05) и числе степеней свободы (k_1 и k_2) по таблице F-распределений (см. приложение 5) $F_{табл} = 3,47$.

$F_{расч} > F_{табл}$ (23,91 > 3,47), следовательно, условия эксплуатации оказывают существенное влияние на величину пробега шин.

4.11. Для выявления текущих потерь рабочего времени на производственном участке цеха был использован метод моментных наблюдений. За смену было произведено 64 наблюдения и получено 6 отметок о простое рабочих.

Определить с вероятностью 0,954 доверительные интервалы текущих потерь рабочего времени на производственном участке цеха.

Решение

Показатель w – доля отметок о простое (доля потерь рабочего времени по выборке);

$$w = \frac{6}{64} = 0,094, \text{ или } 9,4\%.$$

Средняя ошибка доли $\mu_p = \sqrt{\frac{w(1-w)}{n}} = \sqrt{\frac{0,094 \cdot 0,906}{64}} = 0,0365.$

Предельная ошибка

$\Delta_p = t\mu_p = 2 \cdot 0,0365 = 0,073$, или 7,3% (при $P = 0,954$ $t = 2$).

Доверительные интервалы потерь рабочего времени:

$$9,4\% \pm 7,3\%;\ 2,1\% - 16,7\%.$$

4.12. В цехе проектируется проведение моментных наблюдений для выявления текущих простоев производственного оборудования.

Требуется для организации моментных наблюдений определить необходимое число наблюдений и число обходов, если в цехе имеется 20 единиц предназначенного к работе оборудования. Никаких предварительных данных о доле простоев в сменном фонде не имеется. Ошибка наблюдения не должна превышать 5% и гарантирована с вероятностью 0,954.

Решение

Необходимая численность моментов наблюдения определяется по формуле

$$n = \frac{t^2 \cdot w(1-w)}{\Delta_p^2}.$$

По условию задачи:

$t = 2$ (так как вероятность $P = 0,954$);

w – доля простоев по условию не дана, поэтому принимается наибольшая дисперсия, когда $w = 0,5$.

$$\Delta_p = 5\%, \text{ или } 0,05,$$

$$n = \frac{2^2 \cdot 0,5 \cdot 0,5}{0,05^2} = 400.$$

Число обходов (т. е. число записей о каждой единице оборудования) определяется путем деления числа наблюдений на число единиц оборудования:

$$m = \frac{400}{20} = 20.$$

4.2. Задачи для самостоятельной работы

4.13. Из общего количества рабочих предприятия была проведена 30%-ная случайная бесповторная выборка с целью определения затрат времени на проезд к месту работы. Результаты выборки следующие (табл. 4.15).

Таблица 4.15

Затраты времени на проезд к месту работы, мин.	До 30	30 – 40	40 – 50	50 – 60	60 – 70
Число рабочих	70	80	200	55	45

Определить:

1) средние затраты времени на проезд к месту работы у рабочих данного предприятия, гарантируя результат с вероятностью 0,997;

2) долю рабочих предприятия, у которых затраты времени на проезд к месту работы составляют 60 мин. и более, гарантируя результат с вероятностью 0,954.

4.14. Выходной контроль качества поступающих на предприятие комплектующих изделий, осуществляемый в порядке механической выборки, дал следующие результаты (табл. 4.16).

Таблица 4.16

Отклонение размера изделия от принятого по ГОСТу, %	Число изделий
От −2,0 до −3,0	5
» −1,0 » −2,0	15
» 0,0 » −1,0	20
» 1,0 » 0,0	80
» 2,0 » 1,0	50
» 3,0 » 2,0	20
» 4,0 » 3,0	5
» 5,0 » 4,0	5

Определить:

1) пределы значений среднего отклонения размера изделий от стандарта по ГОСТу с вероятностью 0,997;

2) пределы доли изделий с отрицательным отклонением в общей совокупности изделий с вероятностью 0,954.

4.15. Произведен 10%-ный пропорциональный типический отбор рабочих со сдельной и повременной системами оплаты труда для изучения показателей выполнения сменного задания. Отбор единиц в каждой группе бесповторный. Выборка дала следующее распределение численности рабочих по проценту выполнения норм выработки (табл. 4.17).

Определить:

1) доверительные интервалы, в которых с вероятностью 0,954 заключен средний процент выполнения сменного задания для всех рабочих предприятия;

2) возможные пределы доли рабочих, выполняющих сменное задание не менее чем на 120% (с вероятностью 0,954);

Таблица 4.17

Группы рабочих по оплате труда	Группы рабочих по проценту выполнения сменного задания				Итого рабочих
	до 100	100 – 120	120 – 140	140 и выше	
Рабочие-сдельщики	20	150	80	30	280
Рабочие-повременщики	40	100	60	20	220
Итого	60	250	140	50	500

3) необходимую численность выборки при определении доли рабочих, выполняющих сменное задание не менее чем на 120%, чтобы с вероятностью 0,954 предельная ошибка выборки не превышала 3%.

4.16. В АО «Прогресс» работает 3000 человек. Методом случайной бесповторной выборки обследовано 1000 человек, из которых 820 выполняли и перевыполняли дневную норму выработки.

Определить:

1) долю рабочих, не выполняющих норму выработки, по данным выборочного обследования;

2) долю всех рабочих акционерного общества, не выполняющих норму (с вероятностью 0,954).

4.17. Из партии изготовленных изделий общим объемом 2000 единиц проверено посредством механической выборки 30% изделий, из которых бракованными оказались 12 изделий.

Определить:

1) долю бракованных изделий по данным выборки;

2) пределы, в которых находится процент бракованных изделий, для всей партии (с вероятностью 0,954).

4.18. По данным выборочного обследования 10 000 пассажиров пригородного сообщения средняя дальность поездки пассажира составила 35,5 км, а среднее квадратическое отклонение – 16,0 км.

Определить:

1) пределы средней дальности поездки пассажиров с вероятностью 0,954;

11-1622

2) как изменится предельная ошибка выборки, если вероятность будет принята равной 0,997?

4.19. Хронометраж работы станочника дал следующие результаты (табл. 4.18).

Таблица 4.18

Затраты времени на изготовление одной детали, с	50–60	60 – 70	70 – 80	80 – 90	90 – 100	Итого
Количество деталей, шт.	3	27	35	29	6	100

Определить:

1) средние затраты времени на обработку одной детали по данным наблюдения;

2) предельную ошибку выборки с вероятностью 0,954, учитывая, что речь идет о массовом производстве, т. е. выборка производится из генеральной совокупности бесконечно большого объема.

4.20. В механическом цехе завода в порядке малой выборки изучались фотографии рабочего дня 10 рабочих. Время непроизводительной работы и перерывов, зависящих от рабочего и по организационно-техническим причинам, для обследованных рабочих составило: 52, 48, 60, 46, 62, 54, 51, 49, 55, 53 мин.

Определить:

1) доверительные пределы, в которых находится среднее время непроизводительной работы и перерывов для всех рабочих цеха, гарантируя результат с вероятностью 0,99;

2) вероятность того, что среднее время непроизводительной работы и перерывов всех рабочих цеха отличалось от полученного по выборке не более чем на 3 мин.

4.21. Из 200 ящиков по 100 деталей в каждом, поступивших на склад готовой продукции, в порядке случайной бесповторной серийной выборки отобрано 5 ящиков, все детали которых проверены на вес. Результаты проверки следующие (табл. 4.19).

Определить:

1) возможные пределы среднего веса детали для всей партии, поступившей на склад (с вероятностью 0,954);

Таблица 4.19

	№ ящика				
	1	2	3	4	5
Средний вес 1 детали, г	50	49	53	53	55

2) объем случайной бесповторной серийной выборки, чтобы с вероятностью 0,683 предельная ошибка выборки при определении среднего веса одной детали для всей партии не превышала 0,7 г.

4.22. На предприятии с числом установленных металлорежущих станков 120 единиц необходимо на основе выборочного обследования определить долю станков возрастом свыше 10 лет. Никаких предварительных данных об удельном весе этого оборудования в общей численности установленного оборудования нет.

Определить, каков должен быть объем выборки с механическим отбором, чтобы при вероятности 0,954 предельная ошибка выборки не превышала 5%.

4.23. Объем выборки: 1) увеличился в 2 раза; 2) уменьшился в 2 раза.

Определить, как изменится ошибка простой случайной повторной выборки.

4.24. На основе 5%-ной бесповторной выборки получены следующие данные о пробеге автомобильных шин, эксплуатируемых в городских условиях (табл. 4.20).

Таблица 4.20

Пробег шин, тыс. км	40 – 42	42 – 44	44 – 46	46 – 48	48 – 50	50 – 52
Число шин	4	8	22	26	40	20

Определить доверительные интервалы среднего пробега шин в городских условиях, гарантируя результат с вероятностью 0,954.

4.25. По 25 рабочим механического цеха собраны данные о прохождении этими рабочими технического обучения и проценте выполнения норм выработки. Результаты обследования следующие (табл. 4.21).

Таблица 4.21

Группы рабочих	Число рабочих	Процент выполнения норм выработки каждым рабочим
Не прошедшие техническое обучение	11	98,0; 102,0; 108,0; 103,2; 97,5; 100,0; 104,0; 100,8; 107,2; 105,4; 99,2
Прошедшие техническое обучение	14	112,8; 118,4; 106,8; 103,1; 108,9; 111,4; 100,8; 114,1; 110,8; 112,0; 107,9; 106,9; 118,7; 110,2

Установить, используя метод дисперсионного анализа, существует ли зависимость между процентом выполнения норм выработки и повышением квалификации, гарантируя результат с вероятностью 0,95.

4.26. Для определения средней из совокупности произведена типическая выборка. Совокупность разделена на три однородные группы численностью 3000, 5000 и 10 000 единиц соответственно. Отбор 5%-ный. Результаты, полученные по данным выборки, следующие (табл. 4.22).

Таблица 4.22

Группы	Выборочная средняя	Выборочная дисперсия
1	12	9
2	15	16
3	18	25

Гарантийную вероятность принять равной 0,997.

Определить доверительные интервалы средней.

4.27. Методом собственно случайной бесповторной выборки было обследовано 150 студентов дневного отделения одного из высших учебных заведений. Доля студентов, совмещающих работу и учебу, составила, по данным выборки, 30%.

Определить вероятность того, что ошибка доли студентов дневного отделения этого учебного заведения, работающих в течение учебного года, не превысит 5%; 10%.

4.28. Сколько фирм необходимо проверить налоговой инспекции района, чтобы ошибка доли фирм, несвоевременно уплачивающих налоги, не превысила 5%? По данным предыдущей проверки, доля таких фирм составила 32%. Доверительную вероятность принять равной 0,954 (0,997).

4.29. Общая численность служащих предприятия составляет 324 человека.

Рассчитайте численность механической выборки для определения доли служащих, прошедших повышение квалификации по использованию вычислительной техники, чтобы с вероятностью 0,954 ошибка репрезентативности не превышала 10%.

4.30. Из 220 отобранных изделий 5% не соответствуют ГОСТу.

Определить среднюю ошибку повторной выборки и границы, в которых находится доля продукции, соответствующая ГОСТу, для всей партии с вероятностью 0,997.

4.31. В сберегательных банках города методом случайной повторной выборки было отобрано 1600 счетов вкладчиков. Средний размер остатков вклада по этим счетам составил 3,2 тыс. руб. при коэффициенте вариации 30%.

Какова вероятность того, что ошибка репрезентативности при определении среднего размера остатков вклада не превысит 0,05 тыс. руб.?

4.32. Для определения средней продолжительности телефонного разговора и доли разговоров, продолжительность которых превышает 5 мин., предполагается провести выборочное наблюдение методом случайной выборки. По данным аналогичных обследований, среднее квадратическое отклонение продолжительности разговора составило 3,5 мин., а доля телефонных разговоров, продолжительность которых превышает 5 мин., составила 0,4.

Сколько телефонных разговоров необходимо обследовать для того, чтобы с вероятностью 0,954 (0,997) найти среднюю продолжительность телефонного разговора, с ошибкой, не превышающей 30 с, а также долю телефонных разговоров, продолжительность которых превышает 5 мин., с ошибкой, не превышающей 5%?

4.33. На автотранспортном предприятии известны следующие результаты выборочного обследования пробега автомобильных шин одного типоразмера в городских условиях при работе водителей различной квалификации (табл. 4.23).

Таблица 4.23

Пробег автомобильных шин, тыс. км	Число шин	
	при работе водителей I класса	при работе водителей II класса
50 – 52	2	10
52 – 54	6	26
54 – 56	18	10
56 – 58	10	8
58 – 60	4	6
Итого	40	60

Определить на основе приведенных данных существенно ли расхождение среднего пробега автомобильных шин для двух групп водителей, гарантируя результат с вероятностью 0,954.

4.34. Обработка детали № 318 производится в цехе на трех станках, имеющих различную производительность. Для определения доли бракованных деталей для всей партии продукции организована типическая выборка. Методом бесповторного отбора от каждого станка взято 10% деталей из числа обработанных за день и получены следующие результаты (табл. 4.24).

Таблица 4.24

№ станка	№ 1	№ 2	№ 3
Число проверенных деталей, шт.	200	120	250
В том числе брак	4	3	6

Определить:

1) предельную ошибку выборки и доверительные интервалы, в которых с вероятностью 0,997 будет находиться процент брака для всей партии деталей, обработанных за день;

2) вероятность того, что процент брака для всей партии деталей будет отличаться от полученного по выборке не более, чем на 0,7%.

4.35. *Построить* 95%-ный доверительный интервал для оценки генерального среднего размера детали по данным 12 деталей, произведенных на токарном автомате, если отклонения размеров этих деталей от середины поля допуска оказались следующими (табл. 4.25).

Таблица 4.25

№ детали	1	2	3	4	5	6	7	8	9	10	11	12
Отклонение размера в МК	–1	+2	–2	+4	–3	+2	+6	–1	0	+4	+2	–1

4.36. Для характеристики использования рабочего времени в механическом цехе проектируется повторное проведение моментного наблюдения. Проведение предыдущего наблюдения дало следующие результаты: 420 отметок состояния «работа» и 60 – состояния «простой».

Определить необходимое число моментных наблюдений и обходов рабочих мест с вероятностью 0,954 (0,997), приняв точность результатов в пределах 1% (2%). Число рабочих мест в цехе – 60.

4.37. При проверке автомобильных шин на сопротивление разрыву была проведена малая выборка и получены следующие результаты (табл. 4.26).

Таблица 4.26

№ шины	1	2	3	4	5	6	7	8
Сопротивление разрыву, кг/см	164	180	176	168	156	186	190	170

Определить доверительные интервалы, в которых заключен средний уровень сопротивления материала разрыву, гарантируя результат с вероятностью 0,99.

4.38. Компания, сдающая автомобили в аренду, решила оценить размеры простоев автомобилей в ремонте в течение года. Выборка по 10 автомобилям показала, что в прошлом году количество дней, в течение которых автомобили находились на ремонте, составило: 15; 11; 19; 24; 6; 18; 20; 15; 18; 9.

Определить с вероятностью 95% доверительный интервал для среднего числа дней в году, когда автомобили не используются в связи с ремонтом, полагая, что распределение времени простоя автомобиля в ремонте подчиняется нормальному закону. Не производя дополнительных расчетов, указать, будет ли доверительный интервал шире или уже, если нужно будет его рассчитать с вероятностью 90%.

4.39. По результатам выборки имеются следующие данные: средняя равна 8, среднее квадратическое отклонение 2,6, а объем выборки – 32 единицы.

Какому уровню доверительной вероятности соответствует доверительный интервал средней $7{,}195 < \bar{x} < 8{,}805$?

4.40. Для изучения важности сторон маркетинговой деятельности была проведена простая случайная выборка, в процессе которой были изучены мнения 50 руководителей маркетинговых служб предприятий пищевой промышленности. Из 11 вопросов 16% участников выборочного обследования наиболее важным посчитали ценовую политику.

Определить с вероятностью 99% доверительный интервал доли руководителей маркетинговых служб в генеральной совокупности предприятий пищевой промышленности, оценивающих ценовую политику, как наиболее важную сторону маркетинга.

Как изменится величина доверительного интервала, если будет обследовано не 50, а 70 руководителей?

4.41. По данным выборочного обследования, средняя арифметическая величина равна 100. При уровне доверительной вероятности 90% верхняя граница доверительного интервала генеральной средней составила 112.

Какой величине равна нижняя граница доверительного интервала?

4.42. Из партии в 4000 электрических лампочек было отобрано по схеме собственно случайной бесповторной выборки 200 лампочек. Средняя продолжительность горения лампочек в выборке оказалась равной 1250 ч.

Какова вероятность того, что средний срок службы лампочек во всей партии заключен в пределах от 1220 до 1280 ч? Среднее квадратическое отклонение по предшествующим исследованиям равно 150 ч.

4.43. Для определения процента нестандартных деталей в партии по схеме повторной выборки было отобрано 500 деталей, среди которых оказалось 12 нестандартных.

Какова вероятность того, что доля нестандартных деталей во всей партии отличается от доли деталей, полученной по выборке не более чем на 0,02 (по абсолютной величине).

4.44. Из партии в 8000 деталей было подвергнуто контролю 12,5% деталей. Среди них оказалось 4% нестандартных.

Определить вероятность того, что доля нестандартных деталей во всей партии отличается от выборочной доли не более чем на 1,5%, если выборочная совокупность образована по схеме:

а) повторной выборки;

б) бесповторной выборки.

4.45. Выборочное обследование дальности поездок населения в пригородных электропоездах трех дорог, организованное по схеме 10%-ной типической бесповторной выборки, дало следующие результаты (табл. 4.27).

Таблица 4.27

Дальность поездки, км	Число пассажиров			Итого
	дорога 1	дорога 2	дорога 3	
0–10	5	5	–	10
10–20	15	20	10	45
20–30	20	40	20	80
30–40	40	25	35	100
40–50	30	10	70	110
50–60	10	–	45	55
Итого	120	100	180	400

Определить:

1) доверительные интервалы, в которых с вероятностью 0,997 заключена средняя дальность поездки пассажира по каждой дороге и в целом по пригородному сообщению;

2) вероятность того, что средняя дальность поездок по трем дорогам вместе отличается от полученной по выборке не более чем на 0,8 км.

4.46. Из генеральной совокупности численностью в 400 единиц планируется 10%-ная выборка с механическим отбором единиц.

Определить:

1) объем выборки;

2) интервал отбора;

3) на сколько частей делится генеральная совокупность.

4.47. Предельная ошибка доли признака при случайной повторной выборке равна 8%.

Определить, как следует изменить объем выборки, если величина ошибки должна быть уменьшена до 5%.

4.48. Из партии изготовленных изделий в 1800 шт. проверено посредством механической выборки 25% изделий, из которых бракованными оказались 18.

Определить:

1) долю бракованных изделий по данным выборки;

2) пределы, в которых находится процент бракованных изделий в партии с вероятностью 0,954.

Ответы к задачам для самостоятельной работы

4.13. 1) 43,3 ± 1,35; 2) 0,1 ± 0,024. **4.14.** 1) 0,8 ± 0,221; 2) 0,2 ± 0,057, или 20,0% ± 5,7%. **4.15.** 1) 117,2 ± 1,38; 2) 0,38 ± 0,041; 3) 866. **4.16.** 1) 0,18; 2) 0,18 ± 0,02. **4.17.** 1) 0,02, или 2,0%; 2) 2,0 ± 0,96. **4.18.** 1) 35,5 ± 0,32; 2) увеличение в 1,5 раза. **4.19.** 1) 75,8; 2) 1,9. **4.20.** 1) 53 ± 5,1 ($S = 5{,}05$; $\mu_{М.В} = 1{,}6$; $t = 3{,}2$); 2) 0,91 $t_ф = 1{,}9$; $S_{(t)} = 0{,}955$. **4.21.** 1) 52 ± 1,94; 2) 10 ящиков. **4.22.** 93. **4.23.** 1) уменьшение в 1,41 раза; 2) увеличение в 1,41 раза. **4.24.** 47,5 ± 0,46. **4.25.** $\tilde{x}_1 = 102{,}3$; $\tilde{x}_2 = 110{,}2$; $\tilde{x}_0 = 106{,}7$; $S_1^2 = 384{,}46$; $S_2^2 = 20{,}14$; $F_{расч} = 19{,}09$; $F_{табл} = 4{,}28$; факторный признак оказывает влияние. **4.26.** 16,167 ± 0,434. **4.27.** $\mu_p = 0{,}0374$; $P = 0{,}8198$; $P = 0{,}9924$. **4.28.** 348; 783. **4.29.** 76. **4.30.** 0,0147; 0,95 ± 0,0441. **4.31.** 0,9625. **4.32.** 196 (441); 384 (864). **4.33.** $\tilde{x}_1 = 55{,}4$; $\tilde{x}_2 = 54{,}1$; $S^1{}_2 = 3{,}94$; $S^2{}_2 = 5{,}88$; $\mu_\delta = 0{,}443$; $t_{расч} = 2{,}93$; $t_{табл} = 2{,}0$; расхождение существенно. **4.34.** 1) $w = 2{,}28\%$; $\Delta_p = 3 \cdot 0{,}006 = 0{,}018$, или 1,8%; 2,28 ± 1,8; 2) $t = 1{,}17$; $P = 0{,}758$. **4.38.** $\tilde{x} = 15{,}5$; $\mu = 1{,}83$; 15,5 ± 4,0; у́же. **4.39.** $\Delta = 0{,}805$; $\mu = 0{,}46$; $P = 0{,}9199$. **4.40.** $\mu = 0{,}052$; $t = 2{,}58$; 16,0 ± 13,4; $\mu = 0{,}044$; 16,0 ± 11,4. **4.41.** 88. **4.42.** $t = 2{,}90$; $P = 0{,}9963$. **4.43.** $t = 2{,}94$; $P = 0{,}9967$. **4.44.** а) $t = 2{,}42$; $P = 0{,}9845$; б) $t = 2{,}58$; $P = 0{,}9901$. **4.45.** 1) $\tilde{x}_1 = 33{,}75$; $S^2{}_1 = 160{,}9$; $\Delta_1 = 3{,}3$; 33,75 ± 3,3; $\tilde{x}_2 = 26{,}5$; $S^2{}_2 = 102{,}7$; $\Delta_2 = 2{,}88$; 26,5 ± 2,88; $\tilde{x}_3 = 41{,}7$; $S^2{}_3 = 127{,}8$; $\Delta_3 = 2{,}40$; 41,7 ± 2,4; $\tilde{x}_0 = 35{,}5$; $S^2{}_0 = 131{,}5$; $\Delta_0 = 1{,}63$; 35,5 ± 1,63; 2) $t = 1{,}47$; $P = 0{,}8584$. **4.46.** 1) 40 ед.; 2) 10; 3) 40. **4.47.** Увеличить в 2,56 раза. **4.48.** 0,04; 0,04 ± 0,016.

ГЛАВА 5

Корреляционная связь и ее статистическое изучение

Корреляционная связь – связь, проявляющаяся не в каждом отдельном случае, а в массе случаев в средних величинах в форме тенденции.

Статистическое исследование ставит своей конечной целью получение модели зависимости для ее практического использования. Решение этой задачи осуществляется в следующей последовательности.

1. Логический анализ сущности изучаемого явления и причинно-следственных связей.

В результате устанавливаются результативный показатель (y), факторы его изменения, характеризуемые показателями ($x_1, x_2, x_3, ..., x_n$). Связь двух признаков (y и x) называется **парной корреляцией.** Влияние нескольких факторов на результативный признак называется **множественной корреляцией.**

По общему направлению связи могут быть прямые и обратные. При прямых связях с увеличением признака x увеличивается и признак y, при обратных – с увеличением признака x признак y уменьшается.

2. Сбор первичной информации и проверка ее на однородность и нормальность распределения.

Для оценки однородности совокупности используется коэффициент вариации по факторным признакам

$$V_{x_i} = \frac{\sigma_{x_i}}{\bar{x}_i} \cdot 100.$$

Совокупность считается однородной, если коэффициент вариации не превышает 33%. Проверка нормальности распределения исследуемых факторных признаков ($x_1, x_2, x_3, ..., x_n$) проводится с помощью правила «трех сигм». Результаты проверки на нормальность распределения следует представлять в табличной форме (табл. 5.1).

Таблица 5.1

Интервалы значений признака-фактора	Число единиц, входящих в интервал	Удельный вес единиц, входящих в интервал, в общем их числе, %	Удельный вес единиц, входящих в интервал, при нормальном распределении, %
1	2	3	4
$(\overline{x}_i - \sigma_{x_i}) - (\overline{x}_i + \sigma_{x_i})$			68,3
$(\overline{x}_i - 2\sigma_{x_i}) - (\overline{x}_i + 2\sigma_{x_i})$			95,4
$(\overline{x}_i - 3\sigma_{x_i}) - (\overline{x}_i + 3\sigma_{x_i})$			99,7

Сопоставление данных граф 3 и 4 позволяет судить о наличии или об отсутствии нормальности распределения.

На практике часто встречаются случаи отклонения от этих двух предпосылок. Однако это не означает, что следует отказаться от применения корреляционного анализа.

3. Исключение из массива первичной информации всех резко выделяющихся (аномальных) единиц по уровню признаков-факторов.

Исключаются все единицы, у которых уровень признака-фактора не попадает в интервал

$$\overline{x}_i \pm 3\sigma_{x_i} \; (\overline{x}_i - 3\sigma_{x_i} \le x_i \le \overline{x}_i + 3\sigma_{x_i}),$$

и формируется новый массив для последующего анализа.

4. Установление факта наличия и направления корреляционной зависимости между результативным (y) и факторным (x) признаками.

Для установления наличия корреляционной связи используется ряд специфических методов: параллельное сопоставление рядов результативного и факторного признака, графическое изображение фактических данных с помощью поля корреляции, построения корреляционной таблицы.

Основным методом выявления наличия корреляционной связи является метод аналитической группировки и определения групповых средних. Он заключается в том, что все единицы совокупности разбиваются на группы по величине признака-фактора и для каждой группы определяется средняя величина результативного признака. На основе данных аналитической группиров-

ки строится график эмпирической линии связи (линии регрессии), вид которой не только позволяет судить о возможном наличии связи, но и дает некоторое представление о форме корреляционной связи. Если эмпирическая линия связи по своему виду приближается к прямой линии, то можно предположить наличие прямолинейной корреляционной связи; если эмпирическая линия приближается к какой-либо кривой, то это связано с наличием криволинейной связи.

5. После установления факта наличия связи и ее формы измеряется степень тесноты связи и проводится оценка ее существенности.

Для определения степени тесноты парной линейной зависимости служит **линейный коэффициент корреляции** (r); при любой форме зависимости (линейной и криволинейной) – **эмпирическое корреляционное отношение** (η).

Для расчета линейного коэффициента корреляции по несгруппированным данным могут быть использованы следующие формулы:

$$r = \frac{\sum(x-\overline{x})(y-\overline{y})}{n\sigma_x\sigma_y}; \quad r = \frac{\sum dxdy}{\sqrt{\sum dx^2 \sum dy^2}};$$

$$r = \frac{\sum xy - \frac{\sum x \cdot \sum y}{n}}{\sqrt{\left(\sum x^2 - \frac{(\sum x)^2}{n}\right)\cdot\left(\sum y^2 - \frac{(\sum y)^2}{n}\right)}},$$

где $(x-\overline{x})$ – отклонения вариантов значений признака-фактора от их средней величины; $(x-\overline{x}) = dx$;

$(y-\overline{y})$ – отклонения вариантов значений результативного признака от их средней величины; $(y-\overline{y}) = dy$;

n – число единиц в совокупности;

σ_x, σ_y – среднее квадратическое отклонение соответственно признака-фактора и результативного признака.

Линейный коэффициент корреляции может принимать значения в пределах от –1 до +1. Чем ближе он по абсолютной величине к 1, тем теснее связь. Знак при нем указывает направление связи: знак «+» соответствует прямой зависимости, знак «–» – обратной.

Если коэффициент корреляции равен нулю, то связи между признаками нет; если он равен единице (с любым знаком), то между признаками существует функциональная связь.

Оценка существенности линейного коэффициента корреляции при большом объеме выборки (свыше 50) проводится с использованием отношения коэффициента корреляции (r) к его средней квадратической ошибке (σ_r):

$$t_{расч.} = \frac{|r|}{\sigma_r},$$

где $\sigma_r = \frac{1-r^2}{\sqrt{n-1}}$.

Если это отношение окажется больше значения t-критерия Стьюдента, определяемого по приложению 6 при числе степеней свободы $k = n - 2$ и с вероятностью $(1 - \alpha)$, то следует говорить о существенности коэффициента корреляции (α – уровень значимости 0,01 или 0,05).

При недостаточно большом объеме выборки величину средней квадратической ошибки коэффициента корреляции определяют по формуле:

$$\sigma_r = \frac{\sqrt{1-r^2}}{\sqrt{n-2}}.$$

В этом случае $t_{расч} = \frac{r\sqrt{n-2}}{\sqrt{1-r^2}}$.

Полученная величина $t_{расч}$ сравнивается с табличным значением t-критерия Стьюдента.

В тех случаях, когда r получен по данным малой выборки, для проверки его существенности целесообразно использовать метод преобразованной корреляции, предложенный Р. Фишером.

Средняя квадратическая ошибка Z-распределения зависит только от объема выборки и определяется по формуле:

$$\sigma_z = \frac{1}{\sqrt{n-3}}.$$

По таблице соотношений между y и Z (приложение 9) находят значение Z, соответствующее рассчитанному коэффициенту корреляции.

Если соотношение Z к средней квадратической ошибке ($Z : \sigma_Z$) окажется больше табличного значения критерия Стьюдента при определенном уровне значимости, то можно говорить о наличии связи между признаками в генеральной совокупности.

Корреляционное отношение определяется по формулам:

$$\eta = \sqrt{\frac{\delta^2}{\sigma_y^2}}; \quad \eta = \sqrt{1 - \frac{\overline{\sigma^2}}{\sigma_y^2}},$$

где δ^2 – межгрупповая дисперсия результативного признака, вызванная влиянием признака-фактора;

σ_y^2 – общая дисперсия результативного признака;

$\overline{\sigma^2}$ – средняя внутригрупповая дисперсия результативного признака.

$$\delta^2 = \frac{\sum(\overline{y}_i - \overline{y}_0)^2 n_i}{\sum n_i};$$

$$\sigma_y^2 = \frac{\sum(y - \overline{y}_0)^2}{n};$$

$$\overline{\sigma^2} = \frac{\sum \sigma_i^2 n_i}{\sum n_i},$$

где $\overline{y}_i$ – среднее значение результативного признака в соответствующих группах, выделенных по величине признака-фактора;

$\overline{y}_0$ – общая средняя для всей совокупности;

n_i – число единиц в соответствующих группах;

σ_i^2 – внутригрупповая дисперсия.

Вычисление корреляционного отношения требует достаточно большого объема информации, которая должна быть представлена в форме групповой таблицы или в форме корреляционной таблицы, т. е. обязательным условием является группировка данных по признаку-фактору (изменяется от 0 до 1).

При недостаточном количестве данных в выделенных группах к рассчитанной величине корреляционного отношения вносится поправка:

$$\eta^2_{скор} = 1 - (1 - \eta^2)(\frac{n-1}{n-m}),$$

где m – число выделенных групп.

Корреляционное отношение в квадрате (η^2) называют коэффициентом детерминации (причинности), он отражает долю факторной дисперсии в общей дисперсии.

В практике могут быть использованы и другие показатели для определения степени тесноты связи.

Элементарной характеристикой степени тесноты связи является **коэффициент Фехнера:**

$$K_Ф = \frac{n_a - n_b}{n_a + n_b},$$

где n_a – количество совпадений знаков отклонений индивидуальных величин факторного признака x и результативного признака y от их средней арифметической величины (например, «плюс» и «плюс», «минус» и «минус», «отсутствие отклонения» и «отсутствие отклонения»);

n_b – количество несовпадений знаков отклонений индивидуальных значений изучаемых признаков от значения их средней арифметической.

Коэффициент Фехнера целесообразно использовать для установления факта наличия связи при небольшом объеме исходной информации. Он изменяется в пределах $-1{,}0 \le K_Ф \le +1{,}0$.

Для определения тесноты связи как между количественными, так и между качественными признаками, при условии, что значения этих признаков могут быть проранжированы по степени убывания или возрастания, используется **коэффициент корреляции рангов Спирмэна:**

$$\rho = 1 - \frac{6\sum d_i^2}{n(n^2-1)},$$

где d_i – разность между величинами рангов признака-фактора и результативного признака;

n – число показателей (рангов) изучаемого ряда.

Он варьирует в пределах от $-1{,}0$ до $+1{,}0$.

Ранговый коэффициент обычно исчисляется на основе небольшого объема исходной информации, поэтому необходимо выполнить проверку его существенности. В приложении 7 приводится таблица предельных значений коэффициента корреляции рангов Спирмэна при условии верности нулевой гипотезы об отсутствии корреляционной связи при заданном уровне значимости и определенном объеме выборочных данных.

Если полученное значение ρ превышает критическую величину при данном уровне значимости, то нулевая гипотеза может быть отвергнута, т. е. величина ρ не является результатом случайных совпадений рангов.

Для исследования степени тесноты связи между качественными признаками, каждый из которых представлен в виде альтернативных признаков, может быть использован **коэффициент ассоциации Д. Юла или коэффициент контингенции К. Пирсона.**

Расчетная таблица в этом случае состоит из четырех ячеек (таблица «четырех полей»), статистическое сказуемое которой схематически может быть представлено в следующем виде (табл. 5.2).

Таблица 5.2

Признаки	А (да)	$\overline{А}$ (нет)	Итого
В (да)	a	b	$a+b$
$\overline{В}$ (нет)	c	d	$c+d$
Итого	$a+c$	$b+d$	n

В расчетной таблице:

a, b, c, d – частоты взаимного сочетания (комбинации) двух альтернативных признаков – А – $\overline{А}$ и В – $\overline{В}$;

n – общая сумма частот.

Коэффициент ассоциации исчисляется по формуле

$$K_A = \frac{ad - bc}{ad + bc}.$$

Коэффициент контингенции:

$$K_A = \frac{ad - bc}{\sqrt{(a+b)(b+d)(a+c)(c+d)}},$$

где a, b, c, d – числа в четырехклеточной таблице.

Коэффициент контингенции также изменяется от –1 до +1, но всегда его величина для тех же данных меньше коэффициента ассоциации.

Для оценки тесноты связи между альтернативными признаками, принимающими любое число вариантов значений, применяется **коэффициент взаимной сопряженности К. Пирсона и коэффициент взаимной сопряженности А. А .Чупрова.**

Первичная статистическая информация для исследования этой связи располагается в форме таблицы (табл. 5.3).

Таблица 5.3

Признаки	A	B	C	Итого
D	f_{11}	f_{12}	f_{13}	A_{1_i}
E	f_{21}	f_{22}	f_{23}	A_{2_i}
F	f_{31}	f_{32}	f_{33}	A_{3_i}
Итого	A_{1_j}	A_{2_j}	A_{3_j}	n

В табл. 5.3:

f_{ij} – частоты взаимного сочетания двух атрибутивных признаков;

n – число пар наблюдений.

Коэффициент взаимной сопряженности К. Пирсона определяется по формуле

$$C = \sqrt{\frac{\varphi^2}{1+\varphi^2}},$$

где φ^2 – показатель средней квадратической сопряженности.

Показатель φ^2 определяется как сумма отношений квадратов частот каждой клетки таблицы к произведению итоговых частот соответствующего столбца и строки за минусом единицы.

$$\varphi^2 = \left(\sum \frac{f_{ij}^2}{A_i \; A_j}\right) - 1,$$

где f_{ij} – частоты каждой клетки;

i – номер строки;

A_i – итоговые частоты по строкам;

A_j – итоговые частоты по графам.

Коэффициент взаимной сопряженности А. А. Чупрова исчисляется по формуле:

$$K = \sqrt{\frac{\varphi^2}{\sqrt{(K_1 - 1)(K_2 - 1)}}},$$

где φ^2 – имеет одинаковое значение с показателем ψ^2 Пирсона и является показателем взаимной сопряженности;

K_1 – число групп по столбцам таблицы;

K_2 – число групп по строкам таблицы.

Коэффициент взаимной сопряженности Чупрова (K) является более гибким, поскольку он учитывает число образуемых по каждому признаку групп (K_1 и K_2), поэтому результат является более точным по сравнению с коэффициентом взаимной сопряженности по формуле Пирсона.

Коэффициент взаимной сопряженности изменяется от 0 до 1.

6. После установления достаточной степени тесноты связи выполняется построение модели связи (уравнения регрессии).

Тип модели выбирается на основе сочетания теоретического анализа и исследования эмпирических данных посредством построения эмпирической линии регрессии. Чаще всего используются следующие типы функций:

а) линейная – $\hat{y}_x = a + bx$;

б) гиперболическая – $\hat{y}_x = a + b\frac{1}{x}$;

в) параболическая – $\hat{y}_x = a + bx + cx^2$;

г) показательная – $\hat{y}_x = ab^x$.

Для определения численных значений параметров уравнения связи (линии регрессии) используется метод наименьших квадратов и решается система нормальных уравнений.

Для определения параметров a и b уравнения прямолинейной корреляционной связи система нормальных уравнений (для несгруппированных данных) следующая:

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

Параметры a и b можно определить по следующим формулам:

$$a = \bar{y} - b\bar{x};\quad b = \frac{\sum xy - n\bar{x}\,\bar{y}}{\sum x^2 - n(\bar{x})^2}\ \text{ или }\ b = \frac{\sum d_x d_y}{\sum d_x^2}.$$

Для проверки возможности использования линейной функции определяется разность $(\eta^2 - r^2)$; если она менее 0,1, то считается возможным применение линейной функции. Для решения этой же задачи можно использовать величину ω^2, определяемую по формуле

$$\omega^2 = \frac{\eta^2 - r^2}{m-2} : \frac{1-\eta^2}{n-m},$$

где m – число групп, на которое разделен диапазон значений факторного признака.

Если ω^2 окажется меньше табличного значения F-критерия, то нулевая гипотеза о возможности использования в качестве уравнения регрессии линейной функции не опровергается. Значение F-критерия определяется по таблице в зависимости от уровня значимости $\alpha = 0{,}05$ (вероятность $P = 0{,}95$) и числа степеней свободы числителя ($k_1 = m - 2$) и знаменателя ($k_2 = n - m$) (см. приложение 5).

Для определения параметров гиперболической функции система нормальных уравнений следующая:

$$\begin{cases} \sum y = an + b\sum \frac{1}{x} \\ \sum y\frac{1}{x} = a\sum\frac{1}{x} + b\sum(\frac{1}{x})^2. \end{cases}$$

Для определения параметров параболы второго порядка система нормальных уравнений такова:

$$\begin{cases} \sum y = na + b\sum x + c\sum x^2 \\ \sum yx = a\sum x + b\sum x^2 + c\sum x^3 \\ \sum yx^2 = a\sum x^2 + b\sum x^3 + c\sum x^4. \end{cases}$$

В качестве меры достоверности уравнения корреляционной зависимости используется процентное отношение средней квадратической ошибки уравнения (S_e) к среднему уровню результативного признака ($\bar{y}$):

$$\frac{S_e}{\bar{y}}\cdot 100,\quad S_e = \sqrt{\frac{\sum(y - \hat{y}_x)^2}{n - l}},$$

где y – фактические значения результативного признака;

$\hat{y}_x$ – значения результативного признака, рассчитанные по уравнению регрессии;

l – число параметров в уравнении регрессии.

Если это отношение не превышает 10 – 15%, то следует считать, что уравнение регрессии достаточно хорошо отображает изучаемую взаимосвязь.

Полученное уравнение регрессии используется для экстраполяции. Однако ее можно применять лишь тогда, когда существенно не изменились условия формирования уровней признаков.

Для результативного признака определяются доверительные границы, в пределах которых с заданной доверительной вероятностью будет находиться теоретическое значение y. Доверительные границы результативного признака y при значении факторного признака x_0 определяются следующим образом:

$$\hat{y}_{x_0} - t_\alpha \frac{S_e}{\sqrt{n}} \sqrt{1 + \frac{(x_0 - \bar{x})^2}{\sigma_x^2}} \le y \le \hat{y}_{x_0} + t_\alpha \frac{S_e}{\sqrt{n}} \sqrt{1 + \frac{(x_0 - \bar{x})^2}{\sigma_x^2}},$$

где t_α – определяется в соответствии с уровнем значимости, по t-распределению Стьюдента с $(n-l)$ степенями свободы.

7. Изучение множественной корреляционной зависимости начинается с анализа матрицы парных коэффициентов корреляции, что позволяет произвести отбор факторов, включаемых в модель множественной зависимости.

Матрица имеет следующий вид (табл. 5.4).

Таблица 5.4

Признак	y_0	x_1	x_2	...	x_k
y_0	1	r_{01}	r_{02}	...	r_{0k}
x_1	r_{01}	1	r_{21}	...	r_{k1}
x_2	r_{02}	r_{12}	1	...	r_{k2}
...	...	...	...	...	...
x_k	r_{0k}	r_{1k}	r_{2k}	1	1

Анализ первой строки матрицы позволяет выявить факторы, у которых степень тесноты связи с результативным показателем значительна, а поэтому они могут быть включены в модель. Однако при построении многофакторных моделей должно соблюдаться требование возможно меньшей коррелированности включенных в модель признаков-факторов (отсутствие мультиколлинеарности). В качестве критерия мультиколлинеарности может быть принято соблюдение следующих неравенств:

$$r_{x_j y} > r_{x_j x_k};\ r_{x_k y} > r_{x_j x_k}.$$

Если приведенные неравенства (или хотя бы одно из них) не выполняются, то исключается тот фактор x_j или x_k, связь которого с результативным признаком y будет менее тесной.

8. Отобранные факторы включаются в модель множественной зависимости. При этом следует учитывать, что число факторов, включаемых в модель, должно быть в 5 – 6 раз меньше, чем число единиц, входящих в совокупность.

Линейное уравнение множественной зависимости имеет следующий вид:

$$\hat{y}_{x_1 x_2 \dots x_k} = a + b_1 x_1 + b_2 x_2 + \dots + b_k x_k.$$

Параметры уравнения определяются из системы нормальных уравнений, отвечающей требованиям способа наименьших квадратов.

Если зависимость выражена уравнением

$$\hat{y}_{x_1 x_2} = a + b_1 x_1 + b_2 x_2,$$

то система нормальных уравнений следующая:

$$\begin{cases} \sum y = an + b_1 \sum x_1 + b_2 \sum x_2 \\ \sum y x_1 = a \sum x_1 + b_1 \sum x_1^2 + b_2 \sum x_2 x_1 \\ \sum y x_2 = a \sum x_2 + b_1 \sum x_1 x_2 + b_2 \sum x_2^2. \end{cases}$$

Мерой достоверности уравнения является процентное отношение средней квадратической ошибки уравнения к среднему уровню результативного показателя, так же как в случае парной корреляции.

9. Для измерения степени тесноты связи между изменениями величины результативного признака (y) и изменениями значений факторных признаков определяется коэффициент множественной (совокупной) корреляции (R).

Для случая зависимости результативного признака от двух факторных признаков формула совокупного коэффициента корреляции имеет вид:

$$R_{yx_1 x_2} = \sqrt{\frac{r_{yx_1}^2 + r_{yx_2}^2 - 2 r_{yx_1} \cdot r_{yx_2} \cdot r_{x_1 x_2}}{1 - r_{x_1 x_2}^2}}.$$

Если число факторов-признаков более двух, то совокупный коэффициент корреляции определяется следующим образом:

$$R^2_{yx_1x_2...x_k} = 1 - \frac{\Delta}{\Delta^*},$$

где Δ – матрица парных коэффициентов корреляции;

Δ^* – соответствует матрице парных коэффициентов корреляции (Δ) без верхней строки и первого столбца.

Величина R^2 называется коэффициентом детерминации; она показывает, в какой мере вариация результативного признака обусловлена влиянием признаков-факторов, включенных в уравнение множественной зависимости.

Величина совокупного коэффициента корреляции изменяется в пределах от 0 до 1 и численно не может быть меньше, чем любой из образующих его парных коэффициентов корреляции. Чем ближе он к единице, тем меньше роль неучтенных в модели факторов и тем более оснований считать, что параметры регрессионной модели отражают степень эффективности включенных в нее факторов.

Для оценки существенности (значимости) совокупного коэффициента корреляции используется критерий F-Фишера.

Для этого определяется F-расчетное по следующей формуле:

$$F_{расч.} = \frac{\sigma^2_{\hat{y}}}{\sigma^2_E} \cdot \frac{n-l}{l-1},$$

где $\sigma^2_{\hat{y}}$ – факторная дисперсия результативного признака, обусловленная вариацией признаков-факторов;

$$\sigma^2_{\hat{y}} = \frac{\sum(\hat{y} - \bar{y})^2}{n},$$

где $\hat{y}$ – значения результативного признака, рассчитанные по уравнению регрессии;

σ^2_E – остаточная дисперсия: $\sigma^2_E = \sigma^2_y - \sigma^2_{\hat{y}}$;

σ^2_y – общая дисперсия результативного признака;

n – число данных;

l – число параметров уравнения.

По таблице F-распределения (см. приложение 5) следует отыскать табличное значение $F_{табл}$ при числе степеней свободы $k_1 = l - 1$, $k_2 = n - 1$ и уровне значимости $\alpha = 0,05$ ($P = 1 - 0,05$).

Если $F_{расч} < F_{табл}$, то с вероятностью 0,95 можно утверждать, что связь между результативным и факторными признаками существенна.

Кроме совокупного коэффициента корреляции познавательное значение имеют частные коэффициенты корреляции, позволяющие установить степень тесноты связи между результативным признаком y и каждым из факторных признаков при исключении искажающего влияния других факторных признаков. Следовательно, коэффициенты частной корреляции отражают степень «чистого» влияния факторного признака на результативный признак. Для их расчета могут быть использованы парные коэффициенты корреляции.

Для случая зависимости результативного признака y от двух признаков-факторов (x_1 и x_2) определяются два коэффициента частной корреляции:

- частный коэффициент корреляции между результативным признаком y и фактором x_1 при элиминировании фактора x_2:

$$r_{yx_1(x_2)} = \frac{r_{yx_1} - r_{yx_2} \cdot r_{x_1x_2}}{\sqrt{(1 - r^2_{yx_2})(1 - r^2_{x_1x_2})}};$$

- частный коэффициент корреляции между результативным признаком y и фактором x_2 при элиминировании фактора x_1:

$$r_{yx_2(x_1)} = \frac{r_{yx_2} - r_{yx_1} \cdot r_{x_1x_2}}{\sqrt{(1 - r^2_{yx_1})(1 - r^2_{x_1x_2})}}.$$

Для общего случая частные коэффициенты корреляции определяются по формуле

$$r_{yx_k}(x_1, x_2, ..., x_{k-1}) = \sqrt{\frac{R^2_k - R^2_{k-1}}{1 - R^2_k}},$$

где R_k^2 – коэффициент детерминации результативного признака y с комплексом факторных признаков $x_1, x_2, ..., x_{k-1}, x_k$;

R^2_{k-1} – коэффициент детерминации результативного признака с комплексом признаков $x_1, x_2, ..., x_{k-1}$;

$r_{yx_k(x_1, x_2, x_{k-1})}$ – частный коэффициент корреляции результативного признака y с факторным признаком x_k при исключении влияния факторных признаков $x_1, x_2, ..., x_{k-1}$.

Величина частного коэффициента корреляции лежит в пределах от 0 до 1, а знак определяется знаком соответствующих параметров регрессии.

Рассчитывая величины частных коэффициентов корреляции, следует иметь в виду, что каждый из них по своей абсолютной величине не может быть больше величины коэффициента множественной (совокупной) корреляции $R_{yx_1, x_2, ..., x_k}$.

10. Для сравнения роли различных факторов в формировании моделируемого показателя определяется коэффициент эластичности ($Э_j$) или β-коэффициент (β_j).

Частный коэффициент эластичности показывает, на сколько процентов в среднем изменяется результативный признак y с изменением признака-фактора x на 1%, и определяется по формуле

$$Э_j = b_j \cdot \frac{\overline{x}_j}{\overline{y}},$$

где b_j – коэффициент регрессии при j-м факторе.

β-коэффициент показывает, на какую часть среднего квадратического отклонения изменится результативный показатель при изменении соответствующего фактора x на величину его среднего квадратического отклонения; его формула имеет вид:

$$\beta_j = b_j \cdot \frac{\sigma_{x_j}}{\sigma_y}.$$

5.1. Решение типовых задач

5.1. Имеются экспериментальные данные исследования влияния времени вулканизации на сопротивление резины разрыву (табл. 5.5).

Провести на основе приведенных данных исследование взаимосвязи сопротивления резины разрыву и времени ее вулканизации; аналитическое выражение связи проверить на достоверность.

Таблица 5.5

№ анализа	Время вулканизации (мин.), x	Сопротивление разрыву, кг/см²	№ анализа	Время вулканизации (мин.), x	Сопротивление разрыву, кг/см²
1	35	162	8	33	160
2	40	174	9	36	167
3	30	155	10	31	153
4	42	172	11	36	163
5	37	173	12	43	173
6	38	166	13	39	168
7	34	162	14	44	176

Решение

Результативный признак – сопротивление разрыву (y); факторный признак – время вулканизации (x).

1. Первичная информация проверяется на однородность по признаку-фактору с помощью коэффициента вариации.

$$V_x = \frac{\sigma_x}{\overline{x}} \cdot 100; \quad \overline{x} = \frac{\sum x}{n} = \frac{518}{14} = 37{,}0 \text{ мин.}$$

Для расчета σ_x использована вспомогательная таблица (табл. 5.6).

$$\sigma_x = \sqrt{\frac{\sum (x - \overline{x})^2}{n}} = \frac{240}{14} = 4{,}1 \text{ мин.}; \quad V_x = \frac{4{,}1}{37{,}0} \cdot 100 = 11{,}08\%.$$

Следовательно, совокупность можно считать однородной.

2. Проверка первичной информации на нормальность распределения с помощью правила «трех сигм» (табл. 5.7).

Интервалы для значений признака-фактора: ($\overline{x} \pm \sigma$); ($\overline{x} \pm 2\sigma$); ($\overline{x} \pm 3\sigma$), т. е. (37,0 – 4,1) – (37,0 + 4,1); (37,0 – 8,2) – – (37,0 + 8,2); (37,0 – 12,3) – (37,0 + 12,3).

Первичная информация по признаку-фактору не подчиняется закону нормального распределения, однако это не является основанием для отказа использования корреляционно-регрессионного анализа.

Таблица 5.6

Вспомогательная таблица для расчета σ_x

№ анализа	Время вулканизации (мин.), x	$x-\bar{x}$	$(x-\bar{x})^2$	№ анализа	Время вулканизации (мин.), x	$x-\bar{x}$	$(x-\bar{x})^2$
1	35	−2	4	8	33	−4	16
2	40	+3	9	9	36	−1	1
3	30	−7	49	10	31	−6	36
4	42	+5	25	11	36	−1	1
5	37	0	0	12	43	+6	36
6	38	+1	1	13	39	+2	4
7	34	−3	9	14	44	−7	49
Итого	–	–	–	–	–	0	240

Таблица 5.7

Интервалы значений признака x, мин.	Число единиц, входящих в интервал	Удельный вес числа единиц, входящих в интервал, в общем их числе, %	Удельный вес числа единиц, входящих в интервал, при нормальном распределении, %
32,9–41,1	9	64,3	68,3
28,8–45,2	14	100,0	95,4
24,7–49,3	14	100,0	99,7

3. Исключение из первичной информации резко выделяющихся единиц, которые по признаку-фактору не попадают в интервал $\bar{x} - 3\sigma \le x_i \le \bar{x} + 3\sigma$, т. е., по имеющимся данным:

$$37{,}0 - 3 \cdot 4{,}1 \le x_i \le 37{,}0 + 3 \cdot 4{,}1;\ 24{,}7 \le x_i \le 49{,}3.$$

Резко выделяющихся единиц в первичной информации нет.

4. Для установления факта наличия связи производится аналитическая группировка по признаку-фактору. Группировка выполняется при равных интервалах и числе групп 4. Величина интервала определяется по формуле

$$i = \frac{x_{\max} - x_{\min}}{m} = \frac{44-30}{4} = 3{,}5 \approx 4{,}0 \text{ мин.},$$

где m – число групп.

Величина интервала принимается равной 4,0 мин. Ниже построена групповая таблица (табл. 5.8).

Таблица 5.8

Зависимость сопротивления резины разрыву от времени вулканизации

Время вулканизации, мин.	Число анализов, n_i	Σy_i	Средняя величина сопротивления разрыву (кг/см^2), $\bar{y}_i$
30–34	3	468	156,0
34–38	5	827	165,4
38–42	3	508	169,3
42–46	3	521	173,7
Итого	14	2324	–

Для заполнения групповой таблицы была использована вспомогательная таблица (табл. 5.9).

Таблица 5.9

Время вулканизации, мин.	30–34	35–38	38–42	42–46
№ анализа	3; 8; 10	1; 5; 7; 9; 11	2; 6; 13	4; 12; 14
Сопротивление разрыву, кг/см^2	155; 160; 153	162; 173; 162; 167; 163	174; 166; 168	172; 173; 176

Как видно из данных групповой таблицы (см. табл. 5.8), с увеличением времени вулканизации возрастает величина сопротивления резины разрыву. На рис. 5.1 представлен график связи.

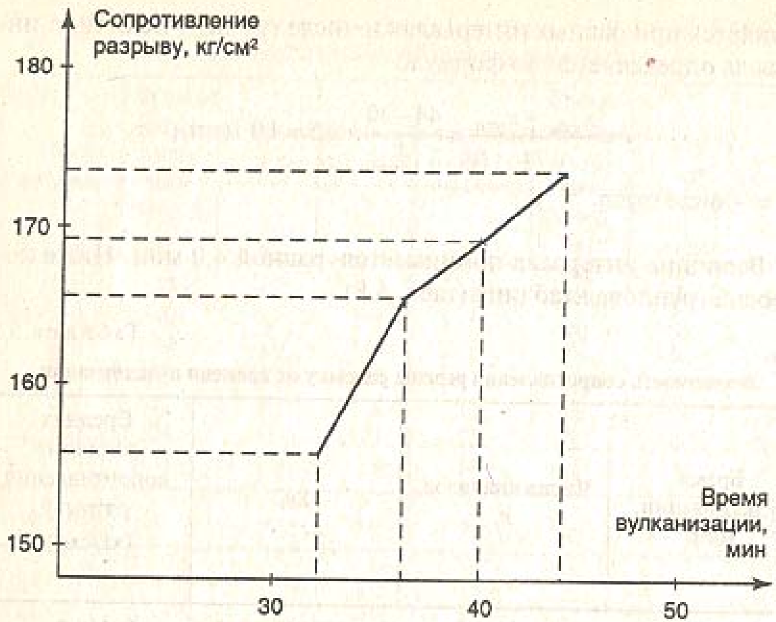

Рис. 5.1. Зависимость сопротивления резины разрыву от времени вулканизации

Эмпирическая линия связи приближается к прямой линии. Следовательно, можно считать наличие прямолинейной корреляции.

5. Для измерения степени тесноты связи используется линейный коэффициент корреляции:

$$r=\frac{\sum xy-\frac{\sum x\cdot\sum y}{n}}{\sqrt{[\sum x^2-\frac{(\sum x)^2}{n}][\sum y^2-\frac{(\sum y)^2}{n}]}}=$$

$$=\frac{86\,359-\frac{518\cdot 2324}{14}}{\sqrt{[19\,406-\frac{(518)^2}{14}][386\,454-\frac{(2324)^2}{14}]}}=+0{,}925.$$

Для расчета r использована вспомогательная таблица (табл. 5.10).

Таблица 5.10

Вспомогательная таблица для расчета линейного коэффициента корреляции и уравнения связи

№ анализа	Время вулканизации (мин.), x	Сопротивление разрыву (кг/см²), y	x^2	y^2	xy	$\hat{y}$	$y-\hat{y}$	$(y-\hat{y})^2$
1	35	162	1 225	26 244	5 670	163,0	−1,0	1,00
2	40	174	1 600	30 276	6 960	170,5	+3,5	12,25
3	30	155	900	24 025	4 650	155,5	−0,5	0,25
4	42	172	1 764	29 584	7 224	173,5	−1,5	2,25
5	37	173	1 369	29 929	6 401	166,0	+7,0	49,00
6	38	166	1 444	27 556	6 308	167,5	−1,5	2,25
7	34	162	1 156	26 244	5 508	161,5	+0,5	0,25
8	33	160	1 089	25 600	5 280	160,0	0,0	0,00
9	36	167	1 296	27 889	6 012	164,5	+2,5	6,25
10	31	153	961	23 409	4 743	157,0	−4,0	16,00
11	36	163	1 296	26 569	5 868	164,5	−1,5	2,25
12	43	173	1 849	29 929	7 439	175,0	−2,0	4,00
13	39	168	1 521	28 224	6 552	169,0	−1,0	1,00
14	44	176	1 936	30 976	7 744	176,5	−0,5	0,25
Итого	518	2 324	19 406	386 454	86 359	–	–	97,00

Значение линейного коэффициента корреляции ($r = +0{,}925$) свидетельствует о наличии прямой и очень тесной связи.

Средняя квадратическая ошибка коэффициента корреляции

$$\sigma_r=\frac{\sqrt{1-r^2}}{\sqrt{n-2}}=\frac{\sqrt{1-0{,}925^2}}{\sqrt{14-2}}=\frac{0{,}38}{3{,}46}=0{,}11;\quad \frac{|r|}{\sigma_r}=\frac{0{,}925}{0{,}11}=8{,}41.$$

По таблице (см. приложение 6) определяется t-критерий Стьюдента при $P = 0{,}95$ и $k = 14 - 2$; $t_{табл} = 2{,}179$.

$$\frac{|r|}{\sigma_r} > t_{\text{табл.}}\text{-критерия } (8{,}41 > 2{,}179).$$

Следовательно, можно утверждать существенность коэффициента корреляции.

6. Определяется модель связи. График линии средних показывает наличие линейной связи, поэтому используется функция $\hat{y} = a + bx$.

$$b = \frac{\sum xy - n\bar{x}\cdot\bar{y}}{\sum x^2 - n(\bar{x})^2} = \frac{86\,359 - 14\cdot 37\cdot 166}{19\,406 - 14\cdot(37)^2} = \frac{371}{240} = 1{,}5;$$

$$\left(\bar{x} = \frac{518}{14} = 37;\quad \bar{y} = \frac{2324}{14} = 166\right);$$

$$a = \bar{y} - b\bar{x} = 166 - 1{,}5\cdot 37 = 110{,}5.$$

Модель связи следующая: $\hat{y} = 110{,}5 + 1{,}5x$.

Для возможности использования линейной функции определяется величина $\omega^2 = \dfrac{\eta^2 - r^2}{m-2} : \dfrac{1-\eta^2}{n-m}$, которая сравнивается с F-критерием.

Для расчета ω^2 исчисляется корреляционное отношение

$$\eta = \sqrt{\frac{\delta^2}{\sigma_y^2}};\quad \delta^2 = \frac{\sum(\bar{y}_i - \bar{y}_0)^2\cdot n_i}{\sum n_i}.$$

Для расчета используются данные табл. 5.8.

$$\delta^2 = \frac{(156{,}0-166)^2\cdot 3 + (165{,}4-166)^2\cdot 5 + (169{,}3-166)^2\cdot 3 + {}}{}$$

$$\frac{+(173{,}7-166)^2\cdot 3}{14} = 36{,}6;$$

$$\sigma_y^2 = \frac{\sum d_y^2}{n} = \frac{\sum(y-\bar{y})^2}{n} =$$

$$= \frac{(162-166)^2 + (174-166)^2 + \ldots + (176-166)^2}{14} = \frac{670}{14} = 47{,}8;$$

$$\eta = \sqrt{\frac{36{,}6}{47{,}8}} = 0{,}88.$$

Следовательно, корреляционное отношение показывает наличие достаточно тесной связи.

$$\omega^2 = \frac{0{,}88^2 - 0{,}93^2}{4-2} : \frac{1-0{,}88^2}{14-4} = -1{,}95,$$

При вероятности $P = 0{,}95$ ($\alpha = 0{,}05$) $k_1 = m - 2 = 4 - 2 = 2$ и $k_2 = n - m = 14 - 4 = 10$; $F_{\text{табл}} = 4{,}10$ (см. приложение 5). Так как ω^2 меньше $F_{\text{табл}}$, то возможность использования линейной функции не опровергается.

Средняя квадратическая ошибка уравнения

$$S_e = \sqrt{\frac{\sum(y-\hat{y})^2}{n-l}} = \sqrt{\frac{97}{14-2}} = 2{,}84.$$

(см. табл. 5.10). В формуле $\hat{y}$ – значения результативного признака, рассчитанные по уравнению связи.

Так, для анализа № 1 $\hat{y} = 110{,}5 + 1{,}5\cdot 35 = 163$; для остальных анализов расчет выполняется аналогично. Результаты расчета записаны в табл. 5.10.

$$\frac{S_e}{\bar{y}}\cdot 100 = \frac{2{,}84}{166}\cdot 100 = 1{,}7\%.$$

Полученное отношение значительно меньше 15%, поэтому уравнение достаточно хорошо отображает взаимосвязь двух признаков и может быть использовано в практической работе.

5.2. По группе предприятий за отчетный год имеются следующие данные (табл. 5.11).

Таблица 5.11

№ предприятия	Годовая производительность труда работника, тыс. руб.	Вооруженность труда основным капиталом, тыс. руб./чел.	Удельный вес оборудования в стоимости основного капитала	Текучесть кадров, %	Интегральный показатель использования рабочего времени
1	360	15,2	0,39	9,1	0,96
2	298	12,8	0,29	10,1	0,80
3	328	13,8	0,34	5,0	0,84
4	330	14,0	0,36	7,0	0,86
5	366	16,3	0,47	9,0	0,98
6	316	12,6	0,28	4,0	0,83
7	334	13,2	0,32	12,0	0,87
8	300	12,9	0,29	6,5	0,84
9	314	13,1	0,33	8,0	0,81
10	320	12,5	0,28	7,0	0,85
11	362	15,7	0,40	8,5	0,97
12	332	13,5	0,34	5,0	0,83

На основании приведенных данных *требуется*:

1) составить уравнение множественной зависимости производительности труда, обосновав систему факторов, включенных в модель;

2) определить совокупный коэффициент корреляции и частные коэффициенты корреляции;

3) сопоставить роль различных факторов в формировании моделируемого показателя.

Решение

Результативный показатель – годовая производительность труда работника y.

Признаки-факторы:

вооруженность труда основным капиталом – x_1;

удельный вес оборудования в стоимости основного капитала – x_2;

текучесть кадров – x_3;

интегральный показатель использования рабочего времени – x_4.

Для определения возможности включения факторов в модель строится матрица парных коэффициентов корреляции (с использованием ЭВМ). Результаты расчета дали следующую матрицу (табл. 5.12).

Таблица 5.12

	y	x_1	x_2	x_3	x_4
y	1	0,911	0,903	0,249	0,979
x_1	0,911	1	0,977	0,264	0,812
x_2	0,903	0,977	1	–0,009	0,710
x_3	0,249	0,264	–0,009	1	0,930
x_4	0,979	0,812	0,710	0,930	1

Цифры первой строки матрицы парных коэффициентов корреляции показывают, что фактор x_3 (текучесть кадров) не следует включать в модель, так как связь результативного показателя с ним слабая ($r_{x_3y} = 0,249$). С остальными факторами связь тесная, и, если нет мультиколлинеарности, они могут быть включены в модель.

Сначала проверяется возможность включения в модель факторов x_1 и x_2. В качестве критерия принимается соблюдение следующих неравенств:

$$r_{x_1y} > r_{x_1x_2};$$
$$r_{x_2y} > r_{x_1x_2}.$$

Фактически эти неравенства не соблюдаются, так как

$$0,911 < 0,977;$$
$$0,903 < 0,977.$$

Следовательно, в модель должен быть включен фактор x_1, так как связь результативного показателя с ним более тесная ($r_{x_1y} = 0,911$).

Далее проверяется возможность включения в модель факторов x_1 и x_4 на основе следующих неравенств:

$$r_{x_1y} > r_{x_1x_4};$$
$$r_{x_4y} > r_{x_1x_4}.$$

Фактически эти неравенства соблюдаются, так как

$$0{,}911 > 0{,}812;$$

$$0{,}979 > 0{,}812.$$

Таким образом, в модель множественной зависимости могут быть включены два фактора: x_1 и x_4. Линейное уравнение имеет следующий вид:

$$\hat{y} = a + b_1 x_1 + b_4 x_4.$$

Система нормальных уравнений для нахождения параметров a, b_1 и b_4 следующая:

$$\begin{cases} \sum y = na + b_1 \sum x_1 + b_4 \sum x_4; \\ \sum y x_1 = a \sum x_1 + b_1 \sum x_1^2 + b_4 \sum x_4 \cdot x_1; \\ \sum y x_4 = a \sum x_4 + b_1 \sum x_1 \cdot x_4 + b_4 \sum x_4^2; \end{cases}$$

$$\begin{cases} 3960 = 12a + b_1 \cdot 165{,}6 + b_4 \cdot 10{,}44; \\ 54937{,}8 = a \cdot 165{,}6 + b_1 \cdot 2303 + b_4 \cdot 144{,}85; \\ 3459{,}9 = a \cdot 10{,}44 + b_1 \cdot 144{,}85 + b_4 \cdot 9{,}127. \end{cases}$$

Решение системы уравнений дает следующие значения параметров:

$$a = 48{,}27;\ b_1 = 4{,}43;\ b_4 = 253{,}63.$$

Модель зависимости производительности труда от факторов имеет следующий вид:

$$\hat{y} = 48{,}27 + 4{,}43 x_1 + 253{,}63 x_4.$$

Значения результативного признака, рассчитанные по уравнению связи, представлены в табл. 5.13. По данным этой же таблицы исчисляется средняя квадратическая ошибка уравнения:

$$S_e = \sqrt{\frac{\sum (y - \hat{y})^2}{n - l}} = \sqrt{\frac{587}{12 - 3}} = 8{,}08.$$

$$\frac{S_e}{\bar{y}} \cdot 100 = \frac{8{,}08}{330} \cdot 100 = 2{,}45\%.$$

Таблица 5.13

Вспомогательная таблица для расчета ошибки модели

№ предприятия	Годовая производительность труда работника (тыс. руб.), y	Теоретический уровень годовой производительности труда работника по уравнению связи (тыс. руб.), $\hat{y}$	$y - \hat{y}$	$(y - \hat{y})^2$
1	360	359,1	+0,9	0,81
2	298	307,8	−9,8	96,04
3	328	332,4	−4,4	19,36
4	330	328,4	+1,6	2,56
5	366	369,0	−3,0	9,00
6	316	314,6	+1,4	1,96
7	334	324,4	+9,6	92,16
8	300	318,9	−18,5	342,25
9	314	311,7	+2,3	5,29
10	320	319,2	+0,8	0,64
11	362	363,8	−1,8	3,24
12	332	328,3	+3,7	13,69
Итого				587,00

Следовательно, уравнение хорошо отображает взаимосвязь производительности труда и двух ее факторов.

Определяется совокупный коэффициент корреляции по следующей формуле:

$$R = \sqrt{\frac{r_{yx_1}^2 + r_{yx_4}^2 - 2 r_{yx_1} \cdot r_{yx_4} \cdot r_{x_1 x_4}}{1 - r_{x_1 x_4}^2}};$$

$$R \sqrt{\frac{0{,}911^2 + 0{,}979^2 - 2 \cdot 0{,}911 \cdot 0{,}979 \cdot 0{,}812}{1 - 0{,}812^2}} = \sqrt{0{,}9979} = 0{,}99$$

(парные коэффициенты корреляции взяты из матрицы).

Близость совокупного коэффициента корреляции к единице означает: роль не учтенных в модели факторов ничтожна, и есть основания считать, что параметры регрессионной модели отражают степень эффективности включенных в нее факторов.

Частные коэффициенты корреляции следующие:

а) частный коэффициент корреляции между результативным признаком y и фактором x_1 при элиминировании фактора x_4:

$$r_{yx_1(x_4)} = \frac{r_{yx_1} - r_{yx_4} \cdot r_{x_1x_4}}{\sqrt{(1-r^2_{yx_4})(1-r^2_{x_1x_4})}} = \frac{0{,}911 - 0{,}979 \cdot 0{,}812}{\sqrt{(1-0{,}979^2)(1-0{,}812^2)}} = 0{,}967;$$

б) частный коэффициент корреляции между результативным признаком y и фактором x_4 при элиминировании фактора x_1:

$$r_{yx_4(x_1)} = \frac{r_{yx_4} - r_{yx_1} \cdot r_{x_1x_4}}{\sqrt{(1-r^2_{yx_1})(1-r^2_{x_1x_4})}} = \frac{0{,}979 - 0{,}911 \cdot 0{,}812}{\sqrt{(1-0{,}911^2)(1-0{,}812^2)}} = 0{,}987.$$

Сопоставление полученных частных коэффициентов корреляции с вычисленными ранее парными коэффициентами корреляции подтверждает наличие тесной связи между результативным и факторными признаками.

Для сравнения роли отдельных факторов в формировании показателя производительности труда определяются коэффициенты эластичности:

а) для фактора $x_1 - Э_1 = b_1 \cdot \frac{\bar{x}_1}{\bar{y}} = 4{,}43 \cdot \frac{13{,}8}{330} = 0{,}185;$

$$\bar{x}_1 = \frac{165{,}6}{12} = 13{,}8; \quad \bar{y} = \frac{3960}{12} = 330;$$

б) для фактора $x_4 - Э_4 = b_4 \cdot \frac{\bar{x}_4}{\bar{y}} = 253{,}63 \cdot \frac{0{,}87}{330} = 0{,}669;$

$$\bar{x}_4 = \frac{10{,}44}{12} = 0{,}87.$$

Следовательно, при увеличении вооруженности труда основным капиталом на 1% производительность труда возрастает лишь на 0,185%. Увеличение показателя использования рабочего времени на 1% повлечет рост производительности труда на 0,669%.

5.3. По группе однородных предприятий имеются данные об объеме выпущенной продукции и уровне механизации трудоемких и тяжелых работ (табл. 5.14).

Таблица 5.14

№ предприятия	Уровень механизации трудоемких и тяжелых работ, %	Объем продукции, млн руб.
1	22	117
2	85	186
3	67	86
4	36	112
5	21	52
6	40	132
7	39	141
8	39	158
9	31	120
10	62	197
11	36	106
12	50	189

Требуется оценить степень тесноты связи между показателями механизации трудоемких и тяжелых работ и объемом продукции при помощи коэффициента Фехнера.

Решение

Для расчета коэффициента Фехнера составляется вспомогательная таблица (табл. 5.15).

$$\bar{x} = \frac{\sum x}{n} = \frac{528}{12} = 44{,}0\%; \quad \bar{y} = \frac{\sum y}{n} = \frac{1596}{12} = 133{,}0 \text{ млн руб.}$$

Коэффициент Фехнера определяется по формуле

$$К_Ф = \frac{n_a - n_b}{n_a + n_b} = \frac{9-3}{9+3} = \frac{6}{12} = 0{,}5.$$

Таблица 5.15

Уровень механизации трудоемких и тяжелых работ (%), x	$x-\overline{x}$	Объем продукции (млн руб.), y	$y-\overline{y}$
22	−22	117	−16
85	41	186	53
67	23	86	−47
36	−8	112	−21
21	−23	52	−81
40	−4	132	−1
39	−5	141	8
39	−5	158	25
31	−13	120	−13
62	18	197	64
36	−8	106	−27
50	6	189	56

Полученное значение коэффициента свидетельствует о наличии связи между уровнем механизации работ и объемом продукции.

5.4. По группе акционерных коммерческих банков региона имеются следующие данные (табл. 5.16).

Исчислить коэффициент корреляции рангов для оценки тесноты связи между суммой прибыли банка и размером его активов.

Таблица 5.16

№ банка	Активы банка, млн руб.	Прибыль, млн руб.
1	866	39,6
2	328	17,8
3	207	12,7
4	185	14,9
5	109	4,0
6	104	15,5
7	327	6,4
8	113	10,1
9	91	3,4
10	849	13,4

Решение

Для расчета коэффициента корреляции рангов предварительно выполняется ранжирование банков по уровню каждого признака (табл. 5.17).

Таблица 5.17

№ банка	Активы банка (млн руб.), x	Ранг по x	№ банка	Прибыль банка (млн руб.), y	Ранг по y
9	191	1	9	3,4	1
6	104	2	5	4,0	2
5	109	3	7	6,4	3
8	113	4	8	10,1	4
4	185	5	3	12,7	5
3	207	6	10	13,4	6
7	327	7	4	14,9	7
2	328	8	6	15,5	8
10	849	9	2	17,8	9
1	866	10	1	39,6	10

Дальнейшие расчеты даны в табл. 5.18.

$$\rho = 1-\frac{6\sum d_i^2}{n(n^2-1)} = 1-\frac{6\cdot 68}{10(100-1)} = +0{,}588.$$

По таблице (приложение 7) определяется при объеме выборки 10 единиц ($n = 10$) и уровне значимости 5% ($\alpha = 0{,}05$) критическая величина для рангового коэффициента корреляции. Она составляет ± 0,6364. Поэтому общий вывод по результату анализа: есть необходимость увеличивать объем выборки.

5.5. В результате обследования работников предприятия получены следующие данные (чел.) (табл. 5.19).

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

Таблица 5.18

Вспомогательная таблица для расчета коэффициента корреляции рангов

№ п/п	Активы (млн руб.), x	Прибыль (млн руб.), y	Ранги		d_i (ранг x – ранг y)	d_i^2
			x	y		
1	866	39,6	10	10	0	0
2	328	17,8	8	9	–1	1
3	207	12,7	6	5	1	1
4	185	14,9	5	7	–2	4
5	109	4,0	3	2	1	1
6	104	15,5	2	8	–6	36
7	327	6,4	7	3	4	16
8	113	10,1	4	4	0	0
9	91	3,4	1	1	0	0
10	849	13,4	9	6	3	9
Итого					0	68

Таблица 5.19

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

Решение

Коэффициент ассоциации –

$$K_A = \frac{ad - bc}{ad + bc} = \frac{300 \cdot 250 - 50 \cdot 200}{300 \cdot 250 + 50 \cdot 200} = \frac{65\ 000}{85\ 000} = +0{,}765.$$

Коэффициент контингенции –

$$K_k = \frac{ad - bc}{\sqrt{(a+b)\cdot(b+d)\cdot(a+c)\cdot(c+d)}} = \frac{300 \cdot 250 - 50 \cdot 200}{\sqrt{350 \cdot 300 \cdot 500 \cdot 450}} =$$

$$= \frac{65\,000}{153\,707} = +0{,}423.$$

Полученные коэффициенты подтверждают наличие существенной связи между исследуемыми признаками. Однако коэффициент контингенции всегда бывает меньше коэффициента ассоциации и дает более корректную оценку тесноты связи.

5.6. Для изучения влияния условий производства на взаимоотношения в коллективе было проведено выборочное обследование 250 рабочих, ответы которых распределились следующим образом (табл. 5.20).

Таблица 5.20

Условия производства	Взаимоотношения в коллективе			
	хорошие	удовлетворительные	неудовлетворительные	итого
Соответствуют требованиям	30	20	10	60
Не полностью соответствуют	25	50	15	90
Не соответствуют	10	40	50	100
Итого	65	110	75	250

Требуется охарактеризовать связь между исследуемыми показателями с помощью коэффициента взаимной сопряженности К. Пирсона и А. А. Чупрова.

Сформулировать вывод.

Решение

Коэффициент взаимной сопряженности К. Пирсона определяется по формуле

$$C=\sqrt{\frac{\varphi^2}{1+\varphi^2}};$$

$$\varphi^2=(\frac{30^2}{60\cdot65}+\frac{25^2}{90\cdot65}+\frac{10^2}{100\cdot65}+\frac{20^2}{60\cdot110}+\frac{50^2}{90\cdot110}+\frac{40^2}{100\cdot110}+$$

$$+\frac{10^2}{60\cdot75}+\frac{15^2}{90\cdot75}+\frac{50^2}{100\cdot75})-1=1{,}2003-1=0{,}2003;$$

$$C=\sqrt{\frac{0{,}2003}{1+0{,}2003}}=\sqrt{0{,}1669}=0{,}408.$$

Коэффициент взаимной сопряженности А. А. Чупрова:

$$K=\sqrt{\frac{\psi^2}{\sqrt{(K_1-1)\cdot(K_2-1)}}}=\sqrt{\frac{0{,}2003}{\sqrt{(3-1)\cdot(3-1)}}}=\sqrt{\frac{0{,}2003}{2}}=0{,}316.$$

Полученное значение коэффициента взаимной сопряженности К. Пирсона свидетельствует, что связь между условиями производства и взаимоотношениями в коллективе весьма заметна. Коэффициент А. А. Чупрова также не опровергает наличие установленной связи.

5.2. Задачи для самостоятельной работы

5.7. В мартеновском цехе завода произведены испытания для определения зависимости производительности печи от содержания углерода в металле. Результаты следующие (табл. 5.21).

На основе приведенных данных *требуется*:

1) проверить первичную информацию по признаку-фактору на однородность;

2) установить факт наличия связи с помощью аналитической группировки;

Таблица 5.21

№ анализа	Процент углерода в металле	Производительность печи, т/ч	№ анализа	Процент углерода в металле	Производительность печи, т/ч
1	0,95	16,3	7	0,82	16,7
2	0,98	16,0	8	1,12	15,8
3	0,65	17,3	9	0,92	16,4
4	0,94	16,5	10	1,12	15,7
5	0,99	16,0	11	1,00	16,0
6	0,78	17,0	12	1,13	15,9

3) с помощью линейного коэффициента корреляции измерить степень тесноты связи; оценить существенность полученного значения коэффициента корреляции с помощью t-критерия Стьюдента при вероятности 0,95;

4) определить модель линейной зависимости, оценить ее достоверность.

5.8. Имеются следующие данные о производительности труда рабочих, выполняющих одинаковую операцию по обработке детали № 408 (табл. 5.22).

Таблица 5.22

Группы рабочих по стажу работы	Число рабочих	Дневная производительность труда, шт.	Дисперсия производительности труда в группе
До 5 лет	6	40	5,0
5 – 10 лет	8	45	2,0
10 лет и более	2	60	1,0

Определить степень тесноты связи между уровнем производительности труда рабочих и стажем их работы.

5.9. Для выявления зависимости производительности труда рабочих, выполняющих в цехе одинаковую операцию по обработке детали № 312, от стажа их работы был найден линейный коэффициент корреляции, равный 0,80.

Кроме того, известны такие данные:

1) средний стаж работы рабочих – $\bar{x} = 5$ лет;

2) среднее квадратическое отклонение по стажу – $\sigma_x = 2$ года;

3) среднее квадратическое отклонение по производительности труда – $\sigma_y = 4{,}4$ шт. (число обработанных деталей);

4) коэффициент вариации по производительности труда – $v_y = 40{,}0\%$.

Найти аналитическое уравнение связи, характеризующее зависимость производительности труда рабочих от стажа их работы.

5.10. Для определения степени влияния стоимости основного капитала на выпуск продукции по 20 предприятиям рассчитаны следующие показатели:

а) линейный коэффициент корреляции, равный 0,8;

б) эмпирическое корреляционное отношение, равное 0,84.

Возможно ли в качестве уравнения связи использовать функцию вида $\hat{y} = a + bx$?

5.11. Для оценки степени тесноты связи между уровнем выработки рабочих и стажем их непрерывной работы была рассчитана величина корреляционного отношения, оказавшаяся равной 0,9 (объем выборки был равен 100).

Определить величину средней внутригрупповой дисперсии, если известно, что общая дисперсия выработки рабочих составляет 6,6.

5.12. В табл. 5.23 представлены следующие данные.

Таблица 5.23

Группы рабочих по стажу работы	Число рабочих в группе	Средняя месячная заработная плата, руб.	Дисперсия месячной заработной платы в группе
До 5 лет	75	3600	14400
5 лет и более	425	4500	15625

Определить степень тесноты связи между стажем работы и размером заработной платы рабочих.

5.13. По 20 однородным предприятиям была получена модель, отражающая зависимость выпуска продукции (y) за месяц от размера основного капитала (x) : $\hat{y} = 12{,}0 + 0{,}5x$.

Кроме того, по этой совокупности предприятий известны следующие данные:

а) средняя стоимость основного капитала на одно предприятие – $\bar{x} = 12{,}0$ млн руб.;

б) средний размер выпуска продукции на одно предприятие – $\bar{y} = 18{,}0$ млн руб.;

в) среднее квадратическое отклонение по стоимости основного капитала – $\sigma_x = 3{,}5$ млн руб.;

г) среднее квадратическое отклонение по размеру выпуска продукции – $\sigma_y = 2{,}0$ млн руб.

Определить степень тесноты связи между размером выпуска продукции и стоимостью основного капитала, учитывая форму связи и используя для этого необходимые данные, из числа приведенных выше.

5.14. В результате обследования студентов экономического факультета института получены следующие данные (табл. 5.24).

Таблица 5.24

Успеваемость	Количество студентов		
	посещающих спортивные секции	не посещающих спортивные секции	итого
Удовлетворительная	220	60	280
Неудовлетворительная	10	30	40

Определить коэффициент контингенции между успеваемостью и посещаемостью спортивных секций.

5.15. По результатам социологического обследования получены следующие данные (табл. 5.25).

Таблица 5.25

Удовлетворенность работой	Мужчины, чел.	Женщины, чел.	Итого
Удовлетворены своей работой	270	80	350
Не удовлетворены своей работой	30	120	150
Итого	300	200	500

Определить коэффициенты ассоциации и контингенции между удовлетворенностью работой и полом.

5.16. Имеются следующие данные по 20 предприятиям городского хозяйства об объеме продукции (услуг) за месяц и уровне механизации труда (табл. 5.26).

Таблица 5.26

№ предприятия	Объем продукции (услуг) за месяц, млн руб.	Уровень механизации, %	№ предприятия	Объем продукции (услуг) за месяц, млн руб.	Уровень механизации, %
1	90	95	11	65	70
2	77	64	12	95	90
3	80	77	13	90	85
4	90	93	14	91	90
5	91	64	15	100	99
6	100	98	16	110	100
7	101	99	17	109	98
8	105	100	18	107	89
9	110	100	19	89	95
10	99	96	20	98	99

Требуется по приведенным данным для выявления наличия связи между объемом продукции и уровнем механизации труда:

1) построить аналитическую таблицу и дать графическое изображение линии связи;

2) измерить тесноту связи между признаками с помощью коэффициента корреляции рангов; проверить его достоверность.

5.17. *Составить* линейное уравнение регрессии, если известно, что $a = 2{,}8$; линейный коэффициент корреляции $r = 0{,}9$; дисперсии признаков x и y соответственно равны 25 и 36.

5.18. По группе однородных предприятий для построения многофакторной модели, отражающей зависимость уровня годовой производительности труда работников, получена следующая матрица парных коэффициентов корреляции (табл. 5.27).

Таблица 5.27

	y	x_1	x_2	x_3	x_4
y	1	0,91	0,90	0,85	0,89
x_1	0,91	1	0,75	0,67	0,70
x_2	0,90	0,75	1	0,52	0,64
x_3	0,85	0,67	0,52	1	0,97
x_4	0,89	0,70	0,64	0,97	1

В табл. 5.27:

y – годовая производительность труда работников;

x_1 – вооруженность труда основными средствами;

x_2 – удельный вес производственного оборудования в общей стоимости основных средств;

x_3 – энерговооруженность труда;

x_4 – коэффициент загрузки оборудования.

Требуется на основе анализа матрицы парных коэффициентов корреляции указать факторы, которые следует включить в многофакторную модель производительности труда.

5.19. По предприятиям имеются следующие данные о емкости электросталеплавильных печей (т) и расходе электроэнергии на 1 т выплавленной стали (кВт · ч/т) (табл. 5.28).

Таблица 5.28

№ п/п	Емкость печи, т	Расход электроэнергии на 1 т стали, кВт · ч/т	№ п/п	Емкость печи, т	Расход электроэнергии на 1 т стали, кВт · ч/т
1	1,0	924	1	10,0	664
2	1,5	909	2	1,5	850
3	1,0	1010	3	3,0	731
4	10,0	541	4	3,5	719
5	10,0	681	5	1,1	793
6	5,0	657	6	0,5	968
7	2,0	888	7	3,5	696
8	1,5	835	8	2,0	892
9	3,5	602	9	3,5	790
10	2,0	890	10	1,0	900

По приведенным данным *требуется*:

1) проверить первичную информацию на однородность и нормальность распределения;

2) построить аналитическую таблицу для выявления зависимости расхода электроэнергии от емкости печи;

3) дать графическое изображение связи;

4) измерить степень тесноты связи с помощью корреляционного отношения;

5) рассчитать параметры линейного уравнения связи и его среднюю квадратическую ошибку.

5.20. По 100 однородным предприятиям было получено уравнение, характеризующее зависимость себестоимости продукции (y) от уровня производительности труда работников (x):

$$\hat{y} = 15{,}0 + \frac{400{,}0}{x}.$$

Кроме того, по этой же совокупности предприятий известны следующие данные:

$$\text{а)}\ \sum_{1}^{100}(y-\bar{y})^2 = 900,$$

где $\bar{y}$ – средняя себестоимость продукции по всем предприятиям;

$$\text{б)}\ \sum_{1}^{100}(x-\bar{x})\cdot(y-\bar{y}) = -250,$$

где $\bar{x}$ – средний уровень производительности труда по всем предприятиям;

$$\text{в)}\ \frac{\sum \sigma_i^2 \cdot n_i}{\sum n_i} = 2{,}7,$$

где σ_i^2 – дисперсия себестоимости по группам предприятий, выделенным по уровню производительности труда;

n_i – число предприятий в каждой группе.

Определить степень тесноты связи между себестоимостью продукции и уровнем производительности труда, учитывая форму связи и используя для этого необходимые данные из числа приведенных выше.

Сформулировать вывод.

5.21. Распределение грузовых автотранспортных предприятий города по формам собственности и уровню рентабельности следующее (табл. 5.29).

Таблица 5.29

Группы предприятий по формам собственности	Число предприятий с уровнем рентабельности		
	ниже среднего	средним	выше среднего
Государственная и муниципальная	15	35	20
Частная	5	42	30
Смешанная (без иностранного участия)	10	20	15

Требуется для оценки влияния формы собственности на уровень рентабельности определить коэффициенты взаимной сопряженности К. Пирсона и А. А. Чупрова.

Сформулировать вывод.

5.22. Ниже приведены экспериментальные данные исследования качества ковкого чугуна от его химического состава (табл. 5.30).

На основе приведенных данных *требуется*:

1) проверить первичную информацию по признаку-фактору на однородность и нормальность распределения;

2) исключить из первичной информации резко выделяющийся анализ, в котором признак-фактор не попадает в интервал $\bar{x} \pm 3\sigma_x$;

3) построить аналитическую таблицу для установления факта наличия связи;

4) по данным аналитической группировки построить график эмпирической линии связи;

5) измерить степень тесноты связи при помощи линейного коэффициента корреляции, оценив его существенность с помощью t-критерия Стьюдента при вероятности 0,954;

Таблица 5.30

№ анализа	Содержание углерода, %	Относительное удлинение, %	№ анализа	Содержание углерода, %	Относительное удлинение, %
1	2,58	13,5	26	2,44	12,3
2	2,30	10,8	27	2,43	11,9
3	2,55	12,9	28	2,70	15,0
4	2,69	14,5	29	2,31	10,5
5	2,40	11,8	30	2,25	10,4
6	2,57	13,2	31	2,47	12,5
7	2,20	9,6	32	2,50	12,6
8	2,40	11,8	33	2,42	11,9
9	2,53	12,9	34	2,48	12,2
10	2,44	12,0	35	2,61	13,7
11	2,32	11,0	36	2,43	11,9
12	2,64	14,2	37	2,30	10,6
13	2,36	11,3	38	2,55	13,1
14	2,41	11,9	39	2,69	14,9
15	2,64	14,0	40	2,35	10,5
16	2,45	12,3	41	2,53	12,9
17	2,37	11,5	42	2,48	12,4
18	2,63	10,9	43	2,38	11,4
19	2,35	10,9	44	2,47	12,7
20	2,50	12,6	45	2,58	13,0
21	2,42	11,7	46	2,24	10,2
22	2,57	13,0	47	2,43	11,9
23	2,46	12,6	48	2,46	12,3
24	2,42	11,3	49	2,51	12,6
25	2,52	12,8	50	2,58	12,9

б) определить модель линейной зависимости, оценив ее достоверность.

5.23. Имеются следующие данные о колеблемости пробега автобусов одной модели до капитального ремонта (табл. 5.31).

Определить долю вариации под влиянием условий эксплуатации в общей вариации пробега до капитального ремонта.

Сформулировать вывод.

5.24. Как повысить точность оценки по уравнению регрессии:

а) увеличить объем исходной информации, используемой для расчета параметров уравнения регрессии;

Таблица 5.31

Группы автобусов по условиям эксплуатации	Число автобусов	Средний пробег в группе, тыс. км	Внутригрупповая дисперсия пробега
Городские	80	135,7	1225
Загородные	120	114,2	784

б) использовать более высокий уровень доверительной вероятности (например, $P = 0{,}997$ вместо $P = 0{,}95$);

в) уменьшить уровень доверительной вероятности (например, вероятность 0,954 вместо вероятности $P = 0{,}997$);

г) все утверждения неверны.

Выберите правильный вариант ответа.

5.25. При производстве керамических изделий была выявлена зависимость уровня брака от влажности используемой массы. Линейный коэффициент корреляции составил 0,69, корреляционное отношение – 0,78, общее число наблюдений – 50. При расчете корреляционного отношения были выделены 4 группы, на которые был разделен диапазон факторного признака.

Определить, возможно ли применение линейного уравнения регрессии, если использовать показатель ω^2 при вероятности 0,95.

Ответы к задачам для самостоятельной работы

5.7. 3) $r = -0{,}967$; $\sigma_r = 0{,}020$; $\dfrac{|r|}{\sigma_r} = 48{,}35$; $t_{\text{табл}} = 2{,}228$; 4) $\hat{y} = 20{,}1 - 4{,}0x$. **5.8.** $\bar{x}_0 = 45$; $\overline{\sigma^2} = 3$; $\delta^2 = 37{,}5$; $\eta = 0{,}96$. **5.9.** $\hat{y} = 2{,}2 + 1{,}76x$. **5.10.** $(\eta^2 - r^2) = 0{,}0656$; возможно использовать указанную функцию. **5.11.** $\overline{\sigma^2} = 1{,}254$. **5.12.** $\bar{x}_0 = 4365$; $\overline{\sigma^2} = 15441{,}0$; $\delta^2 = 103275$; $\eta = 0{,}932$. **5.13.** $r = 0{,}875$. **5.14.** 0,394; **5.15.** $K_A = 0{,}86$; $K_K = 0{,}53$. **5.16.** $\rho = 0{,}762$. **5.17.** $\hat{y} = 2{,}8 + 1{,}08x$. **5.20.** $\eta = 0{,}837$. **5.21.** 0,206; 0,149. **5.23.** 0,1036 или 10,36%. **5.25.** Нет.

ГЛАВА 6
Ряды динамики

Социально-экономические явления общественной жизни находятся в непрерывном развитии. Их изменение во времени статистика изучает при помощи построения и анализа рядов динамики.

Ряд динамики – числовые значения статистического показателя, представленные во временной последовательности. Он состоит из двух граф: в первой указываются периоды (или даты), во второй – показатели, характеризующие изучаемый объект за эти периоды (или на эти даты).

Показатели второй графы носят название *уровней ряда:* первый показатель называется *начальным* уровнем, последний – *конечным.* Уровни ряда могут быть выражены абсолютными, средними или относительными величинами. Ряды динамики относительных и средних величин строятся на основе рядов абсолютных величин. Для наглядного представления ряда динамики широко используются графические изображения, чаще всего линейные диаграммы.

Ряды динамики могут быть двух видов: интервальные и моментные.

В **интервальном ряду** приводятся данные, характеризующие величину показателя за определенные периоды (сутки, месяц, квартал, год и т. д.). Особенностью интервальных рядов из абсолютных величин является то, что их уровни можно суммировать, получая новые численные значения *объема явления*, относящиеся к более длительным периодам.

В **моментном ряду** динамики приводятся данные, характеризующие размеры явления на определенные моменты (даты) времени. Уровни моментных динамических рядов суммировать нельзя; сумма не имеет смысла, так как каждый последующий уровень полностью или частично включает в себя предыдущий уровень. Однако разность уровней имеет смысл, характеризуя увеличение или уменьшение уровня ряда между датами учета.

Важнейшим условием правильного формирования рядов динамики является **сопоставимость уровней, образующих ряд.** Основным требованием сопоставимости уровней является одинаковая

методология их исчисления для всех периодов или дат. При этом все уровни должны быть даны не только в одинаковых, но и в равноценных единицах измерения. Условием сопоставимости данных является также одинаковая полнота охвата различных частей явления, представленного рядом динамики. Уровни показателей в интервальных динамических рядах должны относиться к периодам с одинаковой продолжительностью. Для моментных рядов должна соблюдаться неизменность даты учета (например, наличие материалов на складе предприятия на первое число каждого месяца или квартала).

Вопрос о том, следует ли считать условием сопоставимости данных динамического ряда одинаковость границ территории, к которой относятся данные, решается по-разному. Если ставится задача изучения изменения явления в связи с изменением территории, то в этом случае сопоставляются данные, относящиеся к различной территории. Если же ставится задача изучения темпов развития явления, то сравниваемые показатели должны относиться к неизменной территории.

Следовательно, прежде чем анализировать ряд динамики, надо, исходя из цели исследования, обеспечить сопоставимость уровней ряда дополнительными расчетами, т. е. произвести так называемое **смыкание рядов динамики.**

Специальным условием сопоставимости абсолютных величин интервального динамического ряда является равенство периодов, за которые приводятся данные; если это условие нарушено, то ряд подвергают дополнительной обработке – рассчитывают величины явления в среднем на единицу времени.

Например, объем капитальных вложений (инвестиций) по фирме характеризуется следующими данными (табл. 6.1).

Таблица 6.1

	Период			
	1987–1992 гг.	1993–1996 гг.	1997–1999 гг.	2000 г.
Объем капитальных вложений (в сопоставимых ценах), млн руб.	840,7	420,8	540,3	200,5

Приведенный ряд дает неправильное представление о динамике капитальных вложений, так как показатели относятся к периодам с различной продолжительностью. Чтобы выявить изменение объема капитальных вложений во времени, следует определить величину капитальных вложений на одну и ту же единицу каждого периода – один год.

Объем капитальных вложений за один год составляет (млн руб.):

1987 – 1992 гг. – 140,1 (840,7 : 6);
1993 – 1996 гг. – 105,2 (420,8 : 4);
1997 – 1999 гг. – 180,1 (540,3 : 3);
2000 г. – 200,5.

Как видно из этих данных, объем капитальных вложений по фирме снижался до 1997 г., и лишь начиная с 1997 г. наметилось некоторое его повышение.

Если несопоставимость в рядах динамики вызвана административно-территориальными изменениями, то для изучения развития явления необходимо построить ряд сопоставимых уровней в новых территориальных границах.

Например, имеются данные об объеме транспортной работы (грузообороте) автотранспортных предприятий, обслуживающих регион грузовыми перевозками (млн ткм) (табл. 6.2).

Таблица 6.2

	1998 г.	1999 г.	2000 г.	2001 г.
В старых границах региона	215	238	250	–
В новых границах региона	–	–	300	380

Для приведения этой информации к сопоставимому виду определяется коэффициент пересчета (коэффициент соотношения двух уровней):

$$K_{\text{П}} = \frac{\text{Уровень явления в новых границах}}{\text{Уровень явления в старых границах}};$$

$$K_{\text{П}} = \frac{300}{250} = 1{,}2.$$

Умножая на этот коэффициент уровни объема грузооборота 1998 и 1999 гг., можно построить ряд динамики сопоставимых уровней в новых территориальных границах региона (табл. 6.3).

Таблица 6.3

	1998 г.	1999 г.	2000 г.	2001 г.
Грузооборот, млн ткм	258	285,6	300	380

При изучении рядов динамики перед статистикой стоят следующие задачи: охарактеризовать интенсивность развития явления от периода к периоду (от даты к дате), а также среднюю интенсивность развития за исследуемый период, выявить основную тенденцию в развитии явления, осуществить прогноз развития на будущее, а также изучить сезонные колебания.

Показатели ряда динамики

Для изучения интенсивности изменения уровней ряда во времени исчисляются следующие показатели динамики: абсолютные приросты, коэффициенты роста, темпы роста, темпы прироста, абсолютные значения одного процента прироста.

Перечисленные показатели динамики можно исчислять с переменной или постоянной базой. Если производится сравнение каждого уровня с предыдущим уровнем, то получаются показатели динамики с переменной базой (цепные показатели динамики). Если каждый уровень сравнивается с начальным уровнем или каким-то другим, принятым за базу сравнения, то получаются показатели динамики с постоянной базой (базисные показатели динамики). База сравнения должна выбираться обоснованно, в зависимости от экономических особенностей явления и задач исследования.

Методы расчета показателей динамики представлены в табл. 6.4; они одинаковы для моментных и для интервальных рядов.

При расчете показателей приняты следующие *условные обозначения*:

y_i – уровень любого периода (кроме первого), называемый уровнем текущего периода;

y_{i-1} – уровень периода, предшествующего текущему;

y_k – уровень, принятый за постоянную базу сравнения (часто начальный уровень).

Таблица 6.4

Показатели динамики

Показатель	Метод расчета	
	с переменной базой (цепные)	с постоянной базой (базисные)
1. Абсолютный прирост (Δ)	$\Delta = y_i - y_{i-1}$	$\Delta' = y_i - y_к$
2. Коэффициент роста ($K_р$)	$K_р = \dfrac{y_i}{y_{i-1}}$	$K'_р = \dfrac{y_i}{y_к}$
3. Темп роста ($Т_р$), %	$Т_р = K_р \cdot 100;$	$Т'_р = K'_р \cdot 100$
4. Темп прироста ($Т_п$), %	$Т_п = (K_р - 1) \cdot 100;$ $Т_п = Т_р - 100;$ $Т_п = \dfrac{\Delta}{y_{i-1}} \cdot 100.$	$Т'_п = (K'_р - 1) \cdot 100;$ $Т'_п = Т'_р - 100;$ $Т'_п = \dfrac{\Delta'}{y_к} \cdot 100.$
5. Абсолютное значение 1% прироста (А)	$A = \dfrac{\Delta}{T_п};$ $A = \dfrac{y_{i-1}}{100}$	$A' = \dfrac{\Delta'}{Т'_п};$ $A' = \dfrac{y_к}{100}$

Абсолютный прирост показывает на сколько в абсолютном выражении уровень текущего периода больше (меньше) базисного.

Коэффициент роста показывает, во сколько раз уровень текущего периода больше (или меньше) базисного.

Темп роста – это коэффициент роста, выраженный в процентах; он показывает, сколько процентов уровень текущего периода составляет по отношению к уровню базисного периода.

Темп прироста показывает, на сколько процентов уровень текущего периода больше (или меньше) уровня базисного периода.

Абсолютное значение 1% прироста показывает, какая абсолютная величина скрывается за относительным показателем – одним процентом прироста.

Между базисными и цепными абсолютными приростами существует взаимосвязь: сумма цепных абсолютных приростов равна базисному абсолютному приросту последнего периода ряда динамики.

Например, имеются данные об уровне явления за четыре периода: y_1; y_2; y_3; y_4.

Цепные абсолютные приросты:

$$\Delta_1 = y_2 - y_1;\ \Delta_2 = y_3 - y_2;\ \Delta_3 = y_4 - y_3;$$

$$\Delta_1 + \Delta_2 + \Delta_3 = (y_2 - y_1) + (y_3 - y_2) + (y_4 - y_3) = y_4 - y_1 = \Delta^1{}_3.$$

Взаимосвязь между базисными и цепными коэффициентами роста такова: произведение последовательных цепных коэффициентов роста равно базисному коэффициенту роста, а частное от деления последующего базисного коэффициента роста на предыдущий равно соответствующему цепному коэффициенту роста.

Цепные коэффициенты роста:

$$K_1 = \frac{y_2}{y_1};\ K_2 = \frac{y_3}{y_2};\ K_3 = \frac{y_4}{y_3};$$

$$K_1 \cdot K_2 \cdot K_3 = \frac{y_2}{y_1} \cdot \frac{y_3}{y_2} \cdot \frac{y_4}{y_3} = K'_3;$$

$$K'_3 : K'_2 = \frac{y_4}{y_1} : \frac{y_3}{y_1} = \frac{y_4}{y_3} = K_3,$$

где K'_2, K'_3 – базисные коэффициенты роста.

Для характеристики динамики явлений в ряде случаев используются **пункты роста** (%), когда сравнение производится с отдаленным периодом. Пункты роста представляют собой разность темпов прироста с постоянной базой двух смежных периодов. Пункты роста можно складывать, в результате получают темп прироста соответствующего периода по сравнению с базисным (табл. 6.5).

Таблица 6.5

Показатель	1998 г.	1999 г.	2000 г.	2001 г.
Уровень ряда	200	220	245	254
Темп роста с постоянной базой, %	–	110	122,5	127
Темп прироста с постоянной базой, %	–	10	22,5	27,0
Пункты роста, %	–	10	12,5	4,5

Для характеристики интенсивности развития за длительный период рассчитываются средние показатели динамики; метод их расчета представлен в табл. 6.6.

Таблица 6.6

Средние показатели динамики

Показатель	Метод расчета
1. Средний уровень ряда ($\bar{y}$): а) для интервального ряда	$\bar{y} = \frac{\sum y}{n}$
б) для моментного ряда с равными интервалами	$\bar{y} = \frac{\frac{1}{2}\cdot y_1 + y_2 + ... + y_{n-1} + \frac{1}{2}\cdot y_n}{n-1}$
в) для моментного ряда с неравными интервалами	$\bar{y} = \frac{\sum y \cdot t}{\sum t}$
2. Средний абсолютный прирост ($\bar{\Delta}$)	$\bar{\Delta} = \frac{\sum \Delta}{n-1}$ или $\bar{\Delta} = \frac{y_n - y_1}{n-1}$
3. Средний коэффициент роста ($\bar{K}_р$)	$\bar{K}_р = \sqrt[n-1]{K_{р_1} \cdot K_{р_2}, ..., K_{n-1}}$, или $\bar{K}_р = \sqrt[n-1]{\Pi K_р}$; $\bar{K}_р = \sqrt[n-1]{\frac{y_n}{y_1}}$;
4. Средний темп роста ($\bar{Т}_р$), %	$\bar{Т}_р = \bar{K}_р \cdot 100$
5. Средний темп прироста ($\bar{Т}_п$), %	$\bar{Т}_п = \bar{Т}_р - 100$ или $\bar{Т}_п = (\bar{K}_р - 1)\cdot 100$
6. Средняя величина абсолютного значения 1% прироста ($\bar{А}$)	$\bar{А} = \frac{\bar{\Delta}}{\bar{Т}_п}$

Средние показатели динамики исчисляются одинаковым методом для интервальных и моментных рядов, исключение составляет лишь расчет среднего уровня ряда.

При написании формул приняты следующие условные обозначения:

$y_1, y_2, ..., y_n$ – все уровни последовательных периодов (дат);

n – число уровней ряда;

t – продолжительность периода, в течение которого уровень не изменялся.

При статистическом анализе и сопоставлении стохастически взаимосвязанных рядов динамики, характеризующих различные социально-экономические явления, рассчитывают коэффициент опережения. Он показывает, во сколько раз один ряд динамики растет быстрее другого, и определяется сопоставлением коэффициентов роста двух рядов. Коэффициенты опережения можно также определить путем сопоставления темпов прироста:

$$K_{оп} = \frac{K_р(>)}{K_р(<)};\quad K_{оп} = \frac{Т_п(>)}{Т_п(<)},$$

где $K_р(>)$ – больший коэффициент роста;

$K_р(<)$ – меньший коэффициент роста;

$Т_п(>)$ – больший темп прироста;

$Т_п(<)$ – меньший темп прироста.

Выявление и характеристика основной тенденции развития

Выявление общей тенденции изменения динамического ряда обеспечивается при помощи особых приемов. Наиболее простым способом является укрупнение интервалов и определение итога уровня для этих интервалов или исчисление средних для каждого укрупненного интервала. При этом используют либо переменную среднюю, либо скользящую среднюю. Исчисление итогов за укрупненный период возможно только по интервальным рядам абсолютных величин. Во всех других случаях следует исчислять среднюю величину уровня в укрупненном интервале.

При использовании переменной средней укрупнение интервала обычно начинают с наименьшего возможного, т. е. с интервала, объединяющего два периода. Если в этом случае тенденция развития четко не проявляется, переходят к следующему возможному интервалу, объединяющему три периода. Недостатком этого способа является то, что из поля зрения исследователя выпадает процесс изменения внутри укрупненного интервала, что вызвано сокращением числа уровней изучаемого ряда. Однако преимуществом данного способа является сохранение экономической природы явления.

Расчет переменной средней осуществляется по формулам простой средней арифметической. Например, если укрупненный интервал образован объединением трех периодов, средние для укрупненных интервалов определяются следующим образом:

$$\bar{y}_1 = \frac{y_1 + y_2 + y_3}{3}; \quad \bar{y}_2 = \frac{y_4 + y_5 + y_6}{3} \text{ и т. д.},$$

где $y_1, y_2, ..., y_6$ – уровни исходного ряда динамики.

Скользящая средняя – подвижная динамическая средняя, которая исчисляется по ряду при последовательном передвижении на один интервал, т. е. сначала вычисляют средний уровень из определенного числа первых по порядку уровней ряда, затем – средний уровень из такого же числа членов, начиная со второго. Если в ряду динамики имеются периодические колебания, то период скользящей средней должен совпадать с периодом колебания или быть кратным ему. Если в ряду периодических колебаний нет, то период скользящей подбирают, начиная с наименьшего (т. е. с двух уровней), если в этом случае тенденция не проявляется, то период укрупняют. Период скользящей может быть четным и нечетным, практически удобнее использовать нечетный период, так как в этом случае скользящая средняя будет отнесена к середине периода скольжения.

Скользящие средние с продолжительностью периода, равной 3, следующие:

$$\bar{y}_1 = \frac{y_1 + y_2 + y_3}{3}; \quad \bar{y}_2 = \frac{y_2 + y_3 + y_4}{3}; \quad \bar{y}_3 = \frac{y_3 + y_4 + y_5}{3} \text{ и т. д.}$$

Полученные средние записываются к соответствующему срединному интервалу (второму, третьему, четвертому и т. д.).

Если период скользящей четный, то выполняют центрирование данных, т. е. определение средней из найденных средних, что необходимо для определения срединного периода. Например, если исчисляется скользящая с продолжительностью периода, равной 2, то расчет производится следующим образом:

$$\bar{y}_1 = \frac{y_1 + y_2}{2}; \quad \bar{y}_2 = \frac{y_2 + y_3}{2}; \quad \bar{y}_3 = \frac{y_3 + y_4}{2} \text{ и т. д.}$$

Тогда центрированные средние равны:

$$\bar{y}_1^1 = \frac{\bar{y}_1 + \bar{y}_2}{2}; \quad \bar{y}_2^1 = \frac{\bar{y}_2 + \bar{y}_3}{2} \text{ и т. д.}$$

Первая центрированная средняя будет отнесена ко второму периоду, вторая – к третьему и т. д.

Сглаженный ряд «укорачивается» по сравнению с фактическим на $\frac{m-1}{2}$ члена с одного и другого конца, где m – количество уровней, входящих в интервал.

Обеспечиваемое при применении способа скользящей средней погашение колебаний величин индивидуальных уровней ряда динамики называется сглаживанием динамического ряда.

Рассмотренные приемы выявления общей тенденции изменения динамического ряда не позволяют получить описание плавной линии развития (тренда) данного ряда. Для этой цели используется **аналитическое выравнивание**, сущность которого заключается в нахождении уравнения, выражающего закономерность изменения явления как функцию времени $\hat{y}_t = f(t)$.

Вид уравнения определяется характером динамики развития конкретного явления.

Логический анализ при выборе вида уравнения может быть основан на рассчитанных показателях динамики, а именно:

- если относительно стабильны абсолютные приросты (первые разности уровней приблизительно равны), сглаживание может быть выполнено по прямой;
- если абсолютные приросты равномерно увеличиваются (вторые разности уровней приблизительно равны), можно принять параболу второго порядка;
- при ускоренно возрастающих (замедляющихся) абсолютных приростах принимают параболу третьего порядка;
- при относительно стабильных темпах роста принимают показательную функцию.

На практике выбор формы кривой может быть основан на анализе графического изображения уровней динамического ряда (линейной диаграммы); при этом целесообразнее воспользоваться графическим изображением сглаженных уровней, в которых случайные колебания погашены. Если условия формирования уровней ряда изменяются, то расчет параметров уравнения не следует вести по данным за весь рассматриваемый период. В этом случае было бы целесообразно разбить ряд динамики на ряд периодов, основываясь на оценке устойчивости показателей динамики.

В табл. 6.7 приводятся различные виды трендовых моделей, наиболее часто используемых для аналитического выравнивания.

Таблица 6.7

Виды трендовых моделей

№ п/п	Наименование функции	Вид функции	Система нормальных уравнений для нахождения параметров уравнения
1	Линейная	$\hat{y}_t = a_0 + a_1 \cdot t$	$\sum y = a_0 n + a_1 \sum t$ $\sum yt = a_0 \sum t + a_1 \sum t^2$
2	Парабола второго порядка	$\hat{y}_t = a_0 + a_1 \cdot t + a_2 t^2$	$\sum y = a_0 n + a_1 \sum t + a_2 \sum t^2$ $\sum yt = a_0 \sum t + a_1 \sum t^2 + a_2 \sum t^3$ $\sum yt^2 = a_0 \sum t^2 + a_1 \sum t^3 + a_2 \sum t^4$
3	Парабола третьего порядка	$\hat{y}_t = a_0 \cdot a_1 t + a_2 t^2 + a_3 t^3$	$\sum y = a_0 n + a_1 \sum t + a_2 \sum t^2 + a_3 \sum t^3$ $\sum yt = a_0 \sum t + a_1 \sum t^2 + a_2 \sum t^3 + a_3 \sum t^4$ $\sum yt^2 = a_0 \sum t^2 + a_1 \sum t^3 + a_2 \sum t^4 + a_3 \sum t^5$ $\sum yt^3 = a_0 \sum t^3 + a_1 \sum t^4 + a_2 \sum t^5 + a_3 \sum t^6$
4	Показательная	$\hat{y}_t = a_0 \cdot a_1^t$	$\sum \lg y = n \cdot \lg a_0 + \lg a_1 \sum t$ $\sum \lg y \cdot t = \lg a_0 \sum t + \lg a_1 \sum t^2$
5	Гиперболическая	$\hat{y}_t = a_0 + a_1 \cdot \frac{1}{t}$	$\sum y = a_0 n + a_1 \sum \frac{1}{t}$ $\sum y \frac{1}{t} = a_0 \sum \frac{1}{t} + a_1 \sum \frac{1}{t^2}$

Вычислительный процесс нахождения параметров уравнения при сохранении полной идентичности конечных результатов может быть значительно упрощен, если ввести обозначения дат (периодов) времени с помощью натуральных чисел (t), с тем, чтобы $\sum t = 0$. Если $\sum t = 0$, то $\sum t^3 = 0$; $\sum t^5 = 0$.

Так, если количество уровней в ряду динамики нечетное, то временные даты (t) обозначаются следующим образом (табл. 6.8).

Таблица 6.8

Временные даты (периоды)	Январь	Февраль	Март	Апрель	Май
Уровни ряда динамики	y_1	y_2	y_3	y_4	y_5
Обозначения временных дат, t	−2	−1	0	+1	+2

Если же количество уровней в ряду динамики четное, то обозначения временных дат (t) принимают следующий вид (табл. 6.9).

Таблица 6.9

Временные даты (периоды)	Январь	Февраль	Март	Апрель	Май	Июнь
Уровни ряда динамики	y_1	y_2	y_3	y_4	y_5	y_6
Обозначения временных дат, (t)	−5	−3	−1	+1	+3	+5

Тогда система нормальных уравнений при выравнивании по прямой примет вид:

$$\begin{cases} \sum y = a_0 n \\ \sum yt = a_1 \sum t^2. \end{cases}$$

Откуда

$$a_0 = \frac{\sum y}{n}; \; a_1 = \frac{\sum yt}{\sum t^2}.$$

Система нормальных уравнений при выравнивании по параболе второго порядка будет следующей:

$$\sum y = a_0 n + a_2 \sum t^2;$$

$$\sum yt = a_1 \sum t^2;$$
$$\sum yt^2 = a_0 \sum t^2 + a_3 \sum t^4.$$

Расчет сумм слагаемых целесообразно вести в таблице. Например, при выравнивании по показательной функции вид таблицы следующий (табл. 6.10).

Таблица 6.10

Дата	Уровни ряда динамики	Обозначения временных дат	$\lg y$	t^2	$\lg y \cdot t$	$\hat{y}_t$
1	2	3	4	5	6	7
. . .						
	$\sum y =$	$\sum t = 0$	$\sum \lg y =$	$\sum t^2 =$	$\sum \lg y \cdot t =$	$\sum \hat{y}_t =$

По полученной модели для каждого периода (каждой даты) определяются теоретические уровни тренда ($\hat{y}_t$) и стандартная ошибка аппроксимации (среднее квадратическое отклонение тренда) по формуле

$$S_{\hat{y}} = \sqrt{\frac{\sum(\hat{y} - y_t)^2}{n - l}},$$

где y и $\hat{y}_t$ – соответственно фактические и расчетные значения уровней динамического ряда;
n – число уровней ряда;
l – число параметров в уравнении тренда.

Аналитическое сглаживание позволяет не только определить общую тенденцию изменения явления на рассматриваемом отрезке времени, но и выполнять расчеты для таких периодов, в отношении которых нет исходных данных.

Нахождение по имеющимся данным за определенный период времени некоторых недостающих значений признака внутри этого периода называется **интерполяцией.** Нахождение значений признака за пределами анализируемого периода называется **экстраполяцией.**

Применение экстраполяции для прогнозирования должно основываться на предположении, что найденная закономерность развития внутри динамического ряда сохраняется и вне этого ряда. Это означает, что основные факторы, сформировавшие выявленную закономерность изменения уровней ряда во времени, сохранятся и в будущем.

При составлении прогнозов уровней социально-экономических явлений обычно оперируют не точечной, а интервальной оценкой, рассчитывая так называемые доверительные интервалы прогноза. Границы интервалов определяются по формуле

$$\hat{y}_t \pm t_\alpha \cdot S_{\hat{y}},$$

где $\hat{y}_t$ – точечный прогноз, рассчитанный по модели;
t_α – коэффициент доверия по распределению Стьюдента при уровне значимости α.

Статистическое изучение сезонных колебаний

Сезонные колебания (сезонная неравномерность) – это сравнительно устойчивые внутригодичные колебания, т. е. когда из года в год в одни месяцы уровень явления повышается, а в другие – снижается. Они обусловливаются специфическими условиями, влиянием многочисленных факторов, в том числе и природно-климатических.

Перед статистикой стоит задача выявить колебания и их измерить. Наличие сезонных колебаний выявляют с помощью графического метода. В этом случае применяют линейные диаграммы, на которые наносят данные об объеме явления по месяцам не менее чем за три года.

Целесообразно для выявления сезонных колебаний использовать среднесуточные уровни за каждый месяц, что позволяет исключить влияние различной продолжительности месяцев. Эти уровни исчисляются путем деления общего объема явления за месяц на число календарных дней в месяце.

Измеряются сезонные колебания (сезонная волна) при помощи особых показателей, которые называются **индексами сезонности.** Их расчет выполняют двумя методами в зависимости от характера динамики.

Если годовой уровень явления из года в год остается относительно неизменным, то индексы сезонности исчисляются по формуле

$$i_c = \frac{\overline{y}_i}{\overline{y}_0} \cdot 100,$$

где $\overline{y}_i$ – средняя из фактических уровней одноименных месяцев;

$\overline{y}_0$ – общая средняя за исследуемый период.

Индексы сезонности исчисляются в три этапа:

1. Рассчитываются средние уровни для каждого месяца по данным за все годы исследуемого периода ($\overline{y}_i$), что позволяет избавиться от случайных колебаний месячных уровней по годам.

2. Определяется общая средняя ($\overline{y}_0$) за весь исследуемый период. При расчете сезонных колебаний по абсолютным данным об объеме явления за каждый месяц $\overline{y}_0$ исчисляется путем деления общего объема явления за весь исследуемый период (сумма исходных данных) на число месяцев в исследуемом периоде (так, при периоде 3 года – 36 месяцев). При расчете сезонных колебаний на основе среднесуточных уровней $\overline{y}_0$ определяется как средняя взвешенная арифметическая из среднесуточных объемов по месяцам исследуемого периода; в качестве веса используется число календарных дней каждого месяца:

$$\overline{y}_0 = (\Sigma \overline{y}_i \cdot t_k) : 365,$$

где t_k – число календарных дней каждого месяца.

3. Исчисляются индексы сезонности по приведенной формуле.

Если уровни сезонного явления имеют тенденции к развитию (от года к году повышаются или снижаются), то индексы сезонности исчисляются по формуле

$$i_c = \frac{\overline{y}_i}{\overline{\overline{y}}'_i} \cdot 100,$$

где $\overline{y}_i$ – средняя из фактических уровней одноименных месяцев;

$\overline{\overline{y}}'_i$ – средняя из сглаженных (выравненных) уровней одноименных месяцев.

Расчет индексов сезонности осуществляется в следующей последовательности.

1. Определяются средние уровни для каждого месяца исследуемого периода ($\overline{y}_i$).

2. Для выявления общей тенденции ряда производится аналитическое выравнивание или сглаживание 12-месячной скользящей средней, условно центрированной на 7-й месяц.

3. Определяются для каждого месяца средние из выравненных или сглаженных (центрированных) скользящих средних $\overline{\overline{y}}'_i$.

4. Исчисляются индексы сезонности для каждого месяца.

Для сопоставления величины сезонных колебаний по нескольким предприятиям или периодам может быть использовано среднее квадратическое отклонение, исчисляемое по формуле

$$\sigma = \sqrt{\frac{(i_c - 100)^2}{n}},$$

где i_c – индекс сезонности для каждого месяца;

n – число месяцев (12).

Чем меньше среднее квадратическое отклонение, тем меньше величина сезонных колебаний.

Для анализа и прогнозирования внутригодичных колебаний может быть построена модель сезонных колебаний с помощью **гармоник ряда Фурье:**

$$\hat{y}_t = a_0 + \sum_{k=1}^{m} (a \cos kt + b \sin kt),$$

где k – номер гармоники, определяющий степень точности (адекватности) модели (обычно k берется в пределах от 1 до 4).

При $k = 1$ $\hat{y}_t = a_0 + a_1 \cos t + b_1 \sin t$.

При $k = 2$ $\hat{y}_t = a_0 + a_1 \cos t + b_1 \sin t + a_2 \cos 2t + b_2 \sin 2t$.

Параметры уравнения ($\hat{y}_t$) определяются методом наименьших квадратов:

$$a_0 = \frac{\sum y}{n};\; a_k = \frac{2}{n}\sum y \cdot \cos kt;\; b_k = \frac{2}{n}\sum y \cdot \sin kt.$$

При анализе внутригодовой динамики $n = 12$ – по числу месяцев в году. Представляя месячные периоды как части окружности, ряд внутригодовой динамики имеет следующий вид (табл. 6.11).

Таблица 6.11

Периоды, t	0	$\frac{1}{6}\pi$	$\frac{1}{3}\pi$	$\frac{1}{2}\pi$	$\frac{2}{3}\pi$	$\frac{5}{6}\pi$	π	$\frac{7}{6}\pi$	$\frac{4}{3}\pi$	$\frac{3}{2}\pi$	$\frac{5}{3}\pi$	$\frac{11}{6}\pi$
Уровни, y	y_1	y_2	y_3	y_4	y_5	y_6	y_7	y_8	y_9	y_{10}	y_{11}	y_{12}

Для расчета гармоники используется вспомогательная таблица. Для первой гармоники она имеет следующий вид (табл. 6.12).

Таблица 6.12

Месяц	Условное обозначение месяца, t	Уровни ряда, y	$\cos t$	$\sin t$	$y \cdot \cos t$	$y \cdot \sin t$	$\hat{y}_t$
1	0	y_1					
2	$1/6\,\pi$	y_2					
3	$1/3\,\pi$	y_3					
.							
.							
и т. д.							
		$\sum y$			$\sum y \cdot \cos t$;	$\sum y \cdot \sin t$;	$\sum \hat{y}_t$

6.1. Решение типовых задач

6.1. Имеются следующие данные о выпуске легковых автомобилей в России (табл. 6.13).

Определить показатели динамики выпуска легковых автомобилей от года к году и средние за весь анализируемый период.

Таблица 6.13

	1996 г.	1997 г.	1998 г.	1999 г.
Произведено легковых автомобилей, тыс. шт.	868	986	840	956

Решение

Расчет показателей динамики от года к году представлен в табл. 6.14.

Таблица 6.14

Расчет показателей динамики от года к году

Показатель		Год 1996	1997	1998	1999
Абсолютный прирост Δ, тыс. шт.	с переменной базой	–	$\Delta_1 = 986 - 868 = 118$	$\Delta_2 = 840 - 986 = -146$	$\Delta_3 = 956 - 840 = 116$
	с постоянной базой	–	$\Delta'_1 = 986 - 868 = 118$	$\Delta'_2 = 840 - 868 = -28$	$\Delta'_3 = 956 - 868 = 88$
Коэффициент роста ($K_р$)	с переменной базой	–	$K_{р_1} = \frac{986}{868} = 1{,}136$	$K_{р_2} = \frac{840}{986} = 0{,}852$	$K_{р_3} = \frac{956}{840} = 1{,}138$
	с постоянной базой	–	$K'_{р_1} = \frac{986}{868} = 1{,}136$	$K'_{р_2} = \frac{840}{868} = 0{,}968$	$K'_{р_3} = \frac{956}{868} = 1{,}101$
Темп роста $Т_р$, %	с переменной базой	–	$Т_{р_1} = 1{,}136 \cdot 100 = 113{,}6$	$Т_{р_2} = 0{,}852 \cdot 100 = 85{,}2$	$Т_{р_3} = 1{,}138 \cdot 100 = 113{,}8$
	с постоянной базой	–	$Т'_{р_1} = 1{,}136 \cdot 100 = 113{,}6$	$Т'_{р_2} = 0{,}968 \cdot 100 = 96{,}8$	$Т'_{р_3} = 1{,}101 \cdot 100 = 110{,}1$
Темп прироста $Т_п$, %	с переменной базой	–	$Т_{п_1} = 113{,}6 - 100 = 13{,}6$	$Т_{п_2} = 85{,}2 - 100 = -14{,}8$	$Т_{п_3} = 113{,}8 - 100 = 13{,}8$
	с постоянной базой	–	$Т'_{п_1} = 113{,}6 - 100 = 13{,}6$	$Т'_{п_2} = 96{,}8 - 100 = -3{,}2$	$Т'_{п_3} = 110{,}1 - 100 = 10{,}1$
Абсолютное значение 1% прироста А, тыс. шт.	с переменной базой	–	$А_1 = \frac{118}{13{,}6} = 8{,}68$	$А_2 = \frac{-14{,}6}{-14{,}8} = 9{,}86$	$А_3 = \frac{116}{13{,}8} = 8{,}41$
	с постоянной базой	–	$А'_1 = \frac{868}{100} = 8{,}68$	$А'_2 = \frac{868}{100} = 8{,}68$	$А'_3 = \frac{868}{100} = 8{,}68$

Средний уровень интервального ряда динамики:

$$\bar{y} = \frac{\sum y}{n} = \frac{868 + 986 + 840 + 956}{4} = \frac{3650}{4} = 912{,}5 \text{ тыс. шт.}$$

Средний абсолютный прирост:

$$\bar{\Delta} = \frac{\sum \Delta}{n-1} = \frac{118 + (-146) + 116}{4-1} = \frac{88}{3} = 29{,}3 \text{ тыс. шт.}$$

или

$$\bar{\Delta} = \frac{y_n - y_1}{n-1} = \frac{956 - 868}{4-1} = 29{,}3 \text{ тыс. шт.}$$

Средний коэффициент роста:

$$\bar{K}_\text{р} = \sqrt[n-1]{K_{\text{р}_1} \cdot K_{\text{р}_2}, \ldots, K_{\text{р}_{n-1}}} = \sqrt[3]{1{,}136 \cdot 0{,}852 \cdot 1{,}138} = \sqrt[3]{1{,}101} = 1{,}032$$

или

$$\bar{K}_\text{р} = \sqrt[n-1]{\frac{y_n}{y_1}} = \sqrt[3]{\frac{956}{868}} = \sqrt[3]{1{,}101} = 1{,}032.$$

Средний темп роста:

$$\bar{\text{Т}}_\text{р} = \bar{K}_\text{р} \cdot 100 = 1{,}032 \cdot 100 = 103{,}2\%.$$

Средний темп прироста:

$$\bar{\text{Т}}_\text{П} = (\bar{K}_\text{р} - 1) \cdot 100 = (1{,}032 - 1) \cdot 100 = 3{,}2\%$$

или

$$\bar{\text{Т}}_\text{П} = \bar{\text{Т}}_\text{р} - 100 = 103{,}2 - 100 = 3{,}2\%.$$

Средняя величина абсолютного значения 1% прироста:

$$\bar{A} = \frac{\bar{\Delta}}{\bar{\text{Т}}_\text{П}} = \frac{29{,}3}{3{,}2} = 9{,}2 \text{ тыс. шт.}$$

6.2. Имеются следующие данные о стоимости имущества предприятия (млн руб.) (табл. 6.15).

Таблица 6.15

Год	Отчетные данные			
	1.01	1.04	1.07	1.10
1998	62	65	70	68
1999	68	70	75	78
2000	80	84	88	90
2001	95	–	–	–

Определить абсолютное и относительное изменение среднегодовой стоимости имущества предприятия в 2000 г. по сравнению с 1998 и 1999 гг.

Решение

Поскольку промежутки времени между датами равны, средний уровень моментного ряда динамики исчисляется по формуле

$$\bar{y} = \frac{\frac{1}{2} \cdot y_1 + y_2 + \ldots + y_{n-1} + \frac{1}{2} y_n}{n-1},$$

где y_1 и y_n – уровни соответственно на начало и на конец периода, за который исчисляется средний уровень;

n – число уровней ряда.

$$\bar{y}_{1998} = \frac{\frac{1}{2} \cdot 62 + 65 + 70 + 68 + \frac{1}{2} \cdot 68}{4} = \frac{268}{4} = 67 \text{ млн руб.}$$

$$\bar{y}_{1999} = \frac{\frac{1}{2} \cdot 68 + 70 + 75 + 78 + \frac{1}{2} \cdot 80}{4} = \frac{297}{4} = 74{,}25 \text{ млн руб.}$$

$$\bar{y}_{2000} = \frac{\frac{1}{2} \cdot 80 + 84 + 88 + 90 + \frac{1}{2} \cdot 95}{4} = \frac{349{,}5}{4} = 87{,}357 \text{ млн руб.}$$

В 2000 г. среднегодовая стоимость имущества предприятия возросла по сравнению с 1998 г. на 20,375 млн руб. ($\Delta = 87{,}375 - 67$), или на 30,4% ($\text{К}_\text{р} = 87{,}375 : 67 = 1{,}304$), и по сравнению с 1999 г. – на 13,125 млн руб., или на 17,7%.

6.3. На 1 января 2001 г. остаток составлял: по вкладу № 1 500 руб., по вкладу № 2 – 700 руб. В течение I квартала имели место следующие изменения величины остатков вкладов (руб.) (табл. 6.16).

Таблица 6.16

№ вклада	Дата изменения размера вклада, руб.						
	05.01	17.01	02.02	21.02	13.03	20.03	28.03
1	+150	–200	×	+500	×	×	+100
2	×	×	+300	+150	–550	–200	+400

Определить, на сколько рублей и процентов различаются средние остатки по вкладам за I квартал.

Решение

Для определения среднего уровня моментного ряда при наличии исчерпывающих данных об его изменении используется формула

$$\bar{y} = \frac{\sum y \cdot t}{\sum t},$$

где y – уровни, сохранившиеся без изменения в течение времени t.

Результаты расчета сведены в табл. 6.17 и 6.18.

Таблица 6.17

Вклад № 1

Период	Число дней в периоде, t	Размер вклада (руб.), y	$y \cdot t$
01.01 – 05.01	4	500	2 000
05.01 – 17.01	12	650	7 800
17.01 – 21.02	35	450	15 750
21.02 – 28.03	35	950	33 250
28.03 – 1.04	4	1 050	4 200
Итого	90	–	63 000

Таблица 6.18

Вклад № 2

Периоды	Число дней в периоде, t	Размер вклада (руб.), y	$y \cdot t$
01.01 – 02.02	32	700	22 400
02.02 – 21.02	19	1 000	19 000
21.02 – 13.03	20	1 150	23 000
13.03 – 20.03	7	600	4 200
20.03 – 28.03	8	400	3 200
28.03 – 01.04	4	800	3 200
Итого	90	–	75 000

Средние остатки по вкладам составляют, руб.:

$$\text{вклад № 1: } \bar{y}_1 = \frac{63\,000}{90} = 700;$$

$$\text{вклад № 2: } \bar{y}_2 = \frac{75\,000}{90} = 833{,}3.$$

Следовательно, средний остаток вклада № 2 больше на 133,3 руб., или на 19%.

6.4. Количество дорожно-транспортных происшествий (ДТП), совершенных водителями в регионе, увеличилось в 1995 г. по сравнению с 1990 г. на 2 тыс., или на 4%; в 1997 г. по сравнению с 1995 г. их число возросло на 30%, а в 2000 г. по сравнению с 1997 г. – на 2%.

Определите количество ДТП в 1990, в 1995, в 1997 и в 2000 гг.

Решение

Уровень ДТП в 1990 г. определяется по формуле

$$y_{1990} = A_{1995} \cdot 100,$$

где A_{1995} – абсолютная величина 1% прироста для 1995 г.;

$$A_{1995} = \frac{\Delta}{T_П} = \frac{2,0}{4} = 0,5 \text{ тыс. ед.};$$

$$y_{1990} = 0,5 \,.\, 100 = 50 \text{ тыс. ед.}$$

Далее, недостающие уровни 1995, 1997 и 2000 гг. определим, зная темпы роста для соответствующего периода, тыс. ед.:

$$y_{1995} = y_{1990} \cdot K_р = 50 \cdot 1,04 = 52,0;$$
$$y_{1997} = y_{1995} \cdot K_р = 52 \cdot 1,3 = 67,6;$$
$$y_{2000} = y_{1997} \cdot K_р = 67,6 \cdot 1,02 = 68,952.$$

6.5. Имеются следующие данные об объеме пассажирооборота по автобусным предприятиям города (табл. 6.19).

Таблица 6.19

Год	Пассажирооборот, млрд пасс.-км	Цепные показатели динамики			
		абсолютный прирост, млрд пасс.-км	коэффициент роста	темпы прироста, %	абсолютное значение 1% прироста, млрд пасс.-км
1	2	3	4	5	6
1995	127,0	–	–	–	–
1996			1,102		
1997				7,1	
1998	164,60				
1999					
2000				9,9	1,75

Вычислить и *проставить* в таблицу уровни ряда динамики и недостающие показатели динамики.

Решение

Решение задачи целесообразно начать с определения отсутствующих в таблице уровней ряда динамики, используя для этого данные об уровне предыдущего года и об одном из известных показателей динамики.

Уровень 1996 г. можно найти, используя уровень 1995 г. и коэффициент роста для 1996 г.:

$$y_{1996} = y_{1995} \cdot K_р = 127,0 \,.\, 1,102 = 139,95 \text{ млрд пасс.-км.}$$

Уровень 1997 г. определяется так:

$$y_{1997} = y_{1996} \cdot K_р = 139,95 \,.\, 1,071 = 149,87 \text{ млрд пасс.-км.}$$

Для 1997 г. известен темп прироста:

$$T_п = K_р \cdot 100 - 100,$$

отсюда

$$K_р = \frac{T_п + 100}{100} = \frac{7,1 + 100}{100} = 1,071.$$

Для определения уровня 1999 г. исходим из того, что в 2000 г. каждый процент прироста составлял 1,75 млрд пасс.-км. Следовательно, базисный уровень, т. е. уровень 1999 г. (y_{1999}), принимаемый за 100%, составил 175 млрд пасс.-км (т. е. в 100 раз больше абсолютного значения 1% прироста).

Уровень 2000 г.:

$$y_{2000} = y_{1999} \cdot K_р = 175 \cdot 1,099 = 192,33 \text{ млрд пасс.-км.}$$

$$K_р = \frac{T_п + 100}{100} = \frac{9,9 + 100}{100} = 1,099.$$

Далее выполняется расчет всех недостающих показателей динамики. Ниже представлена полностью заполненная табл. 6.20.

6.6. Численность населения региона возросла за период с 01.01.1998 по 01.01.2000 г. на 4,2%, при этом удельный вес мужского населения за этот период увеличился с 42,1 до 44,3%.

Определить показатели динамики численности мужского и женского населения региона.

Решение

При решении данной задачи исходим из следующих соотношений:

Численность женского населения = Общая численность населения · Удельный вес женщин

Численность мужского населения = Общая численность населения · Удельный вес мужчин.

Таблица 6.20

Год	Пассажирооборот, млрд пасс.-км	Цепные показатели динамики			
		абсолютный прирост, млрд пасс.-км	коэффициент роста	темпы прироста, %	абсолютное значение 1% прироста, млрд пасс.-км
1	2	3	4	5	6
1995	127,00	–	–	–	–
1996	139,95	12,95	1,102	10,2	1,270
1997	149,87	9,92	1,071	7,1	1,399
1998	164,60	14,73	1,098	9,8	1,498
1999	175,00	10,4	1,064	6,4	1,646
2000	192,33	17,3	1,099	9,9	1,750

Изменение численности населения характеризуется коэффициентами роста $K_р$. По условию задачи $K_р$ общей численности населения за рассматриваемый период составил:

$$K_р^0 = \frac{4{,}2+100}{100} = 1{,}042.$$

Коэффициент роста удельного веса мужского населения

$$K_р^{d_м} = \frac{44{,}3}{42{,}1} = 1{,}052.$$

Аналогично находим $K_р$ удельного веса женского населения:

$$K_р^{d_ж} = \frac{100-44{,}3}{100-42{,}1} = \frac{55{,}7}{57{,}9} = 0{,}962.$$

На основе приведенных соотношений получим:

$$K_р^{м} = K_р^0 \cdot K_р^{dм} = 1{,}042 \cdot 1{,}052 = 1{,}096;$$
$$K_р^{ж} = K_р^0 \cdot K_р^{dж} = 1{,}042 \cdot 0{,}962 = 1{,}002,$$

где $K_р^{м}$ и $K_р^{ж}$ – соответственно коэффициенты роста численности мужчин и женщин.

Следовательно, численность мужского населения региона возросла за указанный период на 9,6%, а женского – на 0,2%.

6.7. Среднегодовые темпы роста продукции фермерского хозяйства за период 1996–2000 гг. в земледелии составили 102,6%, а в животноводстве – 105,3%. Величина продукции в 2000 г. (в условных единицах) составила: в земледелии – 7820, в животноводстве – 8590.

Определите среднегодовой темп роста продукции в целом за период 1996–2000 гг.

Решение

Среднегодовой темп роста всей продукции фермерского хозяйства за период 1996–2000 гг.:

$$\overline{K}_р = \sqrt[4]{\frac{y_{2000}^{зем} + y_{2000}^{жив}}{y_{1996}^{зем} + y_{1996}^{жив}}}.$$

По условию задачи известны уровни 2000 г. Начальные уровни 1996 г. определим отдельно по земледелию и животноводству исходя из их среднегодовых темпов роста:

$$\overline{K}_р^{зем} = \sqrt[4]{\frac{y_{2000}^{зем}}{y_{1996}^{зем}}};$$

$$y_{1996}^{зем} = \frac{y_{2000}^{зем}}{(\overline{K}_р^{зем})^4} = \frac{7820}{1{,}026^4} = \frac{7820}{1{,}108} = 7057{,}8 \quad усл.\ ед.;$$

$$y_{1996}^{жив} = \frac{y_{2000}^{жив}}{(\overline{K}_р^{жив})^4} = \frac{8590}{1{,}053^4} = \frac{8590}{1{,}229} = 6989{,}4 \quad усл.\ ед.$$

Отсюда

$$\bar{K}_р = \sqrt[4]{\frac{7820+8590}{7057{,}8+6989{,}4}} = \sqrt[4]{\frac{16410{,}0}{14047{,}2}} = \sqrt[4]{1{,}1682} = 1{,}0396.$$

Среднегодовой темп роста всей продукции фермерского хозяйства за период 1996–2000 гг. составил 103,96%.

6.8. Численность специалистов с высшим и специальным средним образованием (человек) двух регионов представлена в табл. 6.21.

Таблица 6.21

Дата	I регион	II регион
1 января 2000 г.	1850	1720
1 апреля 2000 г.	1866	1810
1 декабря 2000 г.	1910	1860
1 января 2001 г.	1960	1900

Требуется:

1) сопоставить среднегодовую численность специалистов по двум регионам;

2) определить, в каком регионе и на сколько средняя численность специалистов больше (в абсолютном и относительном выражении).

Решение

Для определения среднего уровня моментного ряда динамики с неравными интервалами между отдельными датами, по состоянию на которые дается размер изучаемого явления, используется формула средней арифметической взвешенной:

$$\bar{y} = \frac{\sum \bar{y}_i t}{\sum t},$$

где $\bar{y}_i$ – средние уровни за промежуток времени между двумя соседними датами;

t – продолжительность соответствующих промежутков времени.

Уровни $\bar{y}_i$ рассчитываются по формуле простой средней арифметической.

Тогда средняя численность специалистов по I региону составит, чел.:

за I квартал 2000 г. – $\bar{y}_1 = (1850 + 1866) : 2 = 1858$;

за апрель – ноябрь 2000 г. – $\bar{y}_2 = (1866 + 1910) : 2 = 1888$;

за декабрь 2000 г. – $\bar{y}_3 = (1910 + 1960) : 2 = 1935$.

Средний уровень ряда за 2000 г. по I региону:

$$\bar{y} = \frac{1858 \cdot 3 + 1888 \cdot 8 + 1935 \cdot 1}{12} = 1884 \text{ чел.}$$

Аналогично рассчитывается средний уровень ряда за 2000 г. для II региона; он равен $\bar{y} = 1821$ чел.

В I регионе численность специалистов с высшим и средним образованием больше на 63 чел. (1884 – 1821), или на 3,46% [(63 : 1821) · 100].

6.9. *Рассчитать* интервальный прогноз объема перевозок на 2001 г. с вероятностью 0,99 на основе следующих отчетных данных по грузовому автотранспортному предприятию (табл. 6.22).

Таблица 6.22

Показатель	Год							
	1993	1994	1995	1996	1997	1998	1999	2000
Перевезено груза, тыс. т	360	381	401	422	443	463	485	505

Решение

Для определения формы тренда и расчета его параметров составляется вспомогательная табл. 6.23.

Первые разности (гр. 3) приблизительно равны между собой, что позволяет в виде модели принять уравнение прямой:

$$\hat{y}_t = a_0 + a_1 t.$$

Таблица 6.23

Год	Объем перевозок (тыс. т), y	Первые разности	t	t^2	yt	Теоретический уровень, $\hat{y}_t$	$(y-\hat{y}_t)^2$
1	2	3	4	5	6	7	8
1993	360	–	−7	49	−2520	359,7	0,09
1994	381	21	−5	25	−1905	380,5	0,25
1995	401	20	−3	9	−1203	401,3	0,09
1996	422	21	−1	1	−422	422,1	0,01
1997	443	21	+1	1	443	442,9	0,01
1998	463	20	+3	9	1389	463,7	0,49
1999	485	22	+5	25	2425	484,5	0,25
2000	505	20	+7	49	3535	505,3	0,09
Итого	$\sum y = 3460$		0	168	1742	$\sum \hat{y}_t = 3460$	1,28

Для нахождения a_0 и a_1 используется система нормальных уравнений:

$$\begin{cases} \sum y = n \cdot a_0 + a_1 \sum t \\ \sum yt = a_0 \sum t + a_1 \sum t^2. \end{cases}$$

Для упрощения системы уравнений показатели времени t обозначаются так, чтобы $t = 0$; тогда система принимает вид:

$$\sum y = n \cdot a_0;$$
$$\sum yt = a_1 \sum t_2,$$

откуда

$$a_0 = \frac{\sum y}{n} = \frac{3460}{8} = 432{,}5 \text{ тыс. т}; \quad a_1 = \frac{\sum yt}{\sum t^2} = \frac{1742}{168} = 10{,}4 \text{ тыс. т}$$

Модель тренда:

$$\hat{y}_t = 432{,}5 + 10{,}4t.$$

Точечный прогноз для 2001 г.:

$$\hat{y}_t = 432{,}5 + 10{,}4 \cdot 9 = 526{,}1 \text{ тыс. т.}$$

Интервальный прогноз объема перевозок для 2001 г.:

$$\hat{y}_t \pm t\alpha \cdot S_{\hat{y}}.$$

$$S_{\hat{y}} = \sqrt{\frac{\sum (y-\hat{y}_t)^2}{n-l}} = \sqrt{\frac{1{,}28}{8-2}} = 0{,}462 \text{ тыс. т}$$

(вспомогательные расчеты выполнены в таблице).

Теоретические уровни ($\hat{y}_t$) исчисляются так, тыс. т:

$$1993 \text{ г.} - \hat{y}_t = 432{,}5 + 10{,}4 \cdot (-7) = 359{,}7;$$
$$1994 \text{ г.} - \hat{y}_t = 432{,}5 + 10{,}4 \cdot (-5) = 380{,}5 \text{ и т. д.}$$

Интервальный прогноз объема перевозок для 2001 г.: $\hat{y}_{\text{прогн}} = 526{,}1 \pm 3{,}4 \cdot 0{,}462$. $t_\alpha = 3{,}4$ (при вероятности $P = 0{,}99$, $S_{(t)} = 0{,}995$, $k = n - 1 = 7$, по таблице распределения Стьюдента, приложение 4).

$$524{,}53 \text{ тыс. т} \le \hat{y}_{\text{прогн}} \le 527{,}67 \text{ тыс. т.}$$

6.10. По группе таксомоторных предприятий города имеются следующие данные (табл. 6.24).

Таблица 6.24

Показатель	Год							
	1993	1994	1995	1996	1997	1998	1999	2000
Выполнено оплаченных км, млн	2,3	2,8	3,6	4,8	6,3	8,1	10,3	12,9

Требуется на основе приведенных данных составить интервальный прогноз для 2001 и 2002 гг., гарантируя результат с вероятностью 0,954.

Решение

Для определения формы тренда и расчета его параметров составляется вспомогательная табл. 6.25.

Таблица 6.25

Год	Выполнено оплаченных км (млн), y	Первые разности	Вторые разности	t	t^2	t^4	yt	yt^2
1993	2,3	–	–	–7	49	2401	–16,1	112,7
1994	2,8	0,5	–	–5	25	625	–14,0	70,0
1995	3,6	0,8	0,3	–3	9	81	–10,8	32,4
1996	4,8	1,2	0,4	–1	1	1	–4,8	4,8
1997	6,3	1,5	0,3	+1	1	1	+6,3	6,3
1998	8,1	1,8	0,3	+3	9	81	+24,3	72,9
1999	10,3	2,2	0,4	+5	25	625	+51,5	257,5
2000	12,9	2,6	0,4	+7	49	2401	+90,3	632,1
Итого	51,1			0	168	6216	126,7	1188,7

В данном случае выравнивание проводится по параболе второго порядка, так как вторые разности приблизительно равны.

Парабола второго порядка имеет вид

$$\hat{y}_t = a_0 + a_1 \cdot t + a_2 \cdot t^2.$$

Система нормальных уравнений для нахождения параметров уравнения:

$$\begin{cases} \sum y = a_0 \cdot n + a_1 \sum t + a_2 \sum t^2 \\ \sum yt = a_0 \sum t + a_1 \sum t^2 + a_2 \sum t^3; \\ \sum yt^2 = a_0 \sum t^2 + a_1 \sum t^3 + a_2 \sum t^4 \end{cases}$$

$$\sum t = 0; \sum t^3 = 0;$$

$$\sum y = a_0 \cdot n + a_2 \sum t^2; \qquad 51{,}1 = a_0 \cdot 8 + a_2 \cdot 168;$$
$$\sum yt = a_1 \sum t^2; \qquad 126{,}7 = a_1 \cdot 168;$$
$$\sum yt^2 = a_0 \sum t^2 + a_2 \sum t^4; \qquad 1188{,}7 = a_0 \cdot 168 + a_2 \cdot 6216.$$

Решение уравнений дает следующие параметры, млн км:

$$a_0 = 5{,}6;\ a_1 = 0{,}75;\ a_2 = 0{,}04.$$

Модель тренда:

$$\hat{y}_t = 5{,}6 + 0{,}75t + 0{,}04t^2.$$

Точечный прогноз, млн км:

для 2001 г. – $\hat{y}_t = 5{,}6 + 0{,}75 \cdot 9 + 0{,}04 \cdot 9^2 = 15{,}6$;

для 2002 г. – $\hat{y}_t = 5{,}6 + 0{,}75 \cdot 11 + 0{,}04 \cdot 11^2 = 18{,}7$.

Интервальный прогноз:

$$\hat{y}_t \pm t_\alpha \cdot S_{\hat{y}}.$$

Среднее квадратическое отклонение тренда

$$S_{\hat{y}} = \sqrt{\frac{\sum (y - \hat{y}_t)^2}{n - l}} = \sqrt{\frac{0{,}05}{8-3}} = 0{,}10 \quad \text{млн км}.$$

Теоретический уровень исчисляется так, млн км:

1993 г. – $\hat{y}_t = 5{,}6 + 0{,}75 \cdot (-7) + 0{,}04 \cdot (-7)^2 = 2{,}3$;
1994 г. – $\hat{y}_t = 5{,}6 + 0{,}75 \cdot (-5) + 0{,}04 \cdot (-5)^2 = 2{,}8$.
$t_\alpha = 2{,}4$ (при $P = 0{,}954$, $S_{(i)} = 0{,}977$, $k = n - 1 = 7$ по приложению 4).

Составим вспомогательную табл. 6.26.
Интервальный прогноз:

для 2001 г. – $15{,}6 \pm 2{,}4 \cdot 0{,}10$;
$15{,}6 - 0{,}24 \le \hat{y}_{\text{прогн}} \le 15{,}6 + 0{,}24$;
15,36 млн км $\le \hat{y}_{\text{прогн}} \le$ 15,84 млн км;

для 2002 г. – $18{,}7 \pm 2{,}4 \cdot 0{,}10$;
18,46 млн км $\le \hat{y}_{\text{прогн}} \le$ 18,94 млн км.

Вспомогательная таблица для расчета $S_{\hat{y}}$

Год	Выполнено оплаченных км (млн), y	Теоретический уровень, $\hat{y}_t$	$y - \hat{y}_t$	$(y - \hat{y}_t)^2$
1993	2,3	2,3	0	0
1994	2,8	2,8	0	0
1995	3,6	3,7	–0,1	0,01
1996	4,8	4,9	–0,1	0,01
1997	6,3	6,4	–0,1	0,01
1998	8,1	8,2	–0,1	0,01
1999	10,3	10,3	0	0
2000	12,9	12,8	0,1	0,01
Итого				0,05

6.11. По грузовому автотранспортному предприятию имеются следующие данные об объеме перевозок (табл. 6.27).

Таблица 6.27

Месяц	Среднесуточный объем перевозок, тыс. т		
	1999 г.	2000 г.	2001 г.
Январь	10,2	10,7	10,3
Февраль	10,4	10,4	10,6
Март	10,6	10,8	10,9
Апрель	11,0	11,1	11,3
Май	11,3	11,2	11,2
Июнь	11,5	11,0	11,7
Июль	11,6	11,3	11,8
Август	12,0	11,7	12,4
Сентябрь	11,2	11,6	11,7
Октябрь	10,9	10,7	11,2
Ноябрь	10,2	10,4	10,8
Декабрь	10,0	10,3	10,5

На основе приведенных данных *требуется:*

1) выявить наличие сезонной неравномерности;

2) определить величину сезонной волны, используя индексы сезонности.

Решение

Для выявления наличия сезонной неравномерности используется графический метод (рис. 6.1).

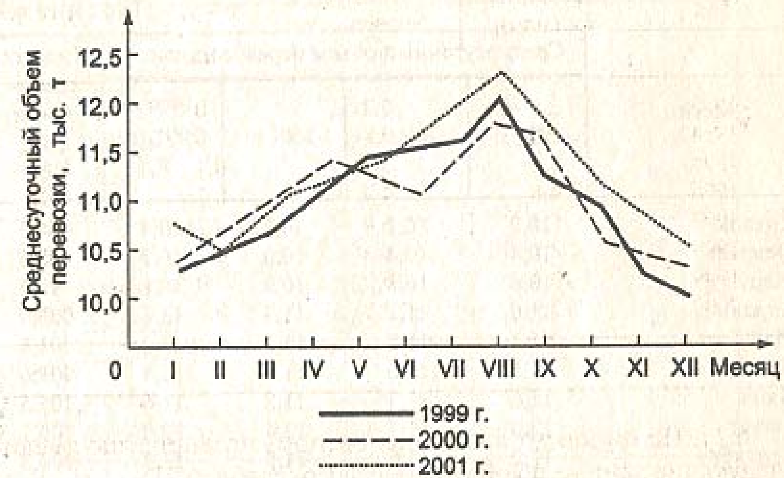

Рис. 6.1. Динамика перевозки грузов

На рис. 6.1 видно, что максимальные и минимальные объемы перевозок практически приходятся на одинаковые месяцы.

Поскольку объем перевозок от года к году существенно не меняется, индекс сезонности определяется по формуле

$$i_c = \frac{\overline{y}_i}{\overline{y}_0} \cdot 100.$$

Расчет индексов сезонности выполнен во вспомогательной табл. 6.28.

Расчет индексов осуществляется так:

1. Определяются средние суточные уровни для каждого месяца:

$$\overline{y}_1 = \frac{10{,}2 + 10{,}7 + 10{,}3}{3} = 10{,}4;$$

$$\overline{y}_2 = \frac{10{,}4 + 10{,}4 + 10{,}6}{3} = 10{,}5 \text{ и т.д.}$$

2. Определяется общая средняя за весь период:

$$\bar{y}_0 = (10{,}4 \cdot 31 + 10{,}5 \cdot 28 + 10{,}8 \cdot 31 + 11{,}1 \cdot 30 + 11{,}2 \cdot 31 + 11{,}4 \cdot 30 + 11{,}6 \cdot 31 + 12{,}0 \cdot 31 + 11{,}5 \cdot 30 + 10{,}9 \cdot 31 + 10{,}5 \cdot 30 + 10{,}3 \cdot 31) : 365 = 11{,}0 \text{ тыс. т.}$$

Таблица 6.28

Месяц	Среднесуточный объем перевозок, тыс. т				Индексы сезонности (%), i_0
	1999 г.	2000 г.	2001 г.	1999 – 2001 гг., $\bar{y}_i$	
Январь	10,2	10,7	10,3	10,4	94,5
Февраль	10,4	10,4	10,6	10,5	95,5
Март	10,6	10,8	10,9	10,8	98,2
Апрель	11,0	11,1	11,3	11,1	100,9
Май	11,3	11,2	11,2	11,2	101,8
Июнь	11,5	11,0	11,7	11,4	103,6
Июль	11,6	11,3	11,8	11,6	105,5
Август	12,0	11,7	12,4	12,0	109,1
Сентябрь	11,2	11,6	11,7	11,5	104,5
Октябрь	10,9	10,7	11,2	10,9	99,0
Ноябрь	10,2	10,4	10,8	10,5	95,5
Декабрь	10,0	10,3	10,5	10,3	93,6

3. Исчисляются индексы сезонности:

для января – $i_c = \dfrac{10{,}4}{11{,}0} \cdot 100 = 94{,}5;$

для февраля – $i_c = \dfrac{10{,}5}{11{,}0} \cdot 100 = 95{,}5$ и т.д.

Индексы сезонности показывают, что среднесуточный объем перевозок в январе меньше среднесуточного за весь период на 5,5% (94,5 – 100), а в августе превышает его на 9,1% (109,1 – 100).

6.12. По станциям технического обслуживания легковых автомобилей города имеются следующие данные (табл. 6.29).

Требуется на основе приведенных данных выявить наличие сезонной неравномерности и рассчитать величину сезонной волны.

Таблица 6.29

Месяц	Число поступивших заявок, тыс.		
	1999 г.	2000 г.	2001 г.
Январь	10,3	13,6	14,0
Февраль	11,1	14,3	14,7
Март	11,5	14,4	15,1
Апрель	12,0	14,6	15,6
Май	12,6	15,6	16,0
Июнь	16,0	17,1	17,4
Июль	15,9	16,9	18,2
Август	16,2	17,0	18,4
Сентябрь	16,4	16,5	17,8
Октябрь	15,2	16,0	17,5
Ноябрь	15,0	14,9	17,0
Декабрь	12,8	13,8	16,5

Решение

Анализ исходной информации позволяет сделать вывод о наличии сезонной неравномерности при росте годового объема технических обслуживаний. Так как уровни явления имеют тенденцию к развитию, то индексы сезонности исчисляются по формуле

$$i_c = \frac{\bar{y}_i}{\bar{\bar{y}}'_i} \cdot 100,$$

где $\bar{y}_i$ – среднее число заявок для одноименных месяцев;

$\bar{\bar{y}}'_i$ – среднее число из сглаженных (выравненных) уровней одноименных месяцев.

Результаты расчета индексов сезонности представлены в следующей таблице (табл. 6.30).

Расчет индексов сезонности проводится в следующей последовательности.

1. Определяются средние из фактических уровней одноименных месяцев ($\bar{y}_i$, гр. 2).

Таблица 6.30

Месяц	Среднее число заявок за 1999 – 2001 гг. (тыс.), $\overline{y}_i$	Скользящая 12-месячная средняя, центрированная на 7-м месяце				Индексы сезонности (%), i_c
		1999 г., $\overline{y}'_i$	2000 г., $\overline{y}'_i$	2001 г., $\overline{y}'_i$	1999 – 2001 гг., $\overline{\overline{y}}'_i$	
1	2	3	4	5	6	7
Январь	12,6	–	15,1	15,7	15,4	81,8
Февраль	13,4	–	15,2	15,8	15,5	86,5
Март	13,7	–	15,3	15,9	15,6	87,8
Апрель	14,1	–	15,3	16,0	15,7	90,4
Май	14,7	–	15,3	16,2	15,8	93,1
Июнь	16,8	–	15,4	16,4	15,9	105,7
Июль	17,0	13,9	15,4	–	14,7	115,6
Август	17,2	14,2	15,5	–	14,9	115,4
Сентябрь	16,9	14,4	15,5	–	15,0	112,7
Октябрь	16,2	14,6	15,6	–	15,1	107,3
Ноябрь	15,6	14,9	15,6	–	15,3	101,9
Декабрь	14,4	15,1	15,6	–	15,4	93,5

Например,

$$\text{для января} - \ \overline{y}_1 = \frac{10{,}3+13{,}6+14{,}0}{3} = 12{,}6;$$

$$\text{для февраля} - \ \overline{y}_2 = \frac{11{,}1+14{,}1+14{,}7}{3} = 13{,}4 \text{ и т.д.}$$

2. Для выявления общей тенденции ряда производится сглаживание с помощью 12-месячной скользящей средней y'_i:

$$y'_1 = (10{,}3 + 11{,}1 + 11{,}5 + 12{,}0 + 12{,}6 + 16{,}0 + 15{,}9 + 16{,}2 + 16{,}4 + 15{,}2 + 15{,}0 + 12{,}8) : 12 = 13{,}75;$$

$$y'_2 = (11{,}1 + 11{,}5 + 12{,}0 + 12{,}6 + 16{,}0 + 15{,}9 + 16{,}2 + 16{,}4 + 15{,}2 + 15{,}0 + 12{,}8 + 13{,}6) : 12 = 14{,}03;$$

$$y'_3 = (11{,}5 + 12{,}0 + 12{,}6 + 16{,}0 + 15{,}9 + 16{,}2 + 16{,}4 + 15{,}2 + 15{,}0 + 12{,}8 + 13{,}6 + 14{,}3) : 12 = 14{,}28.$$

Всего таких средних будет 25.

3. Для нахождения срединного периода, к которому может быть отнесена скользящая средняя, выполняется центрирование, т. е. определение средней из найденных скользящих средних ($\overline{y}'_i$):

$$\overline{y}'_1 = \frac{y'_1 + y'_2}{2} = \frac{13{,}75+14{,}03}{3} = 13{,}9;$$

$$\overline{y}'_2 = \frac{y'_2 + y'_3}{2} = \frac{14{,}03+14{,}28}{2} = 14{,}2 \text{ и т.д.}$$

Первая средняя ($\overline{y}'_1$) может быть отнесена к июлю 1999 г., так как данный месяц будет срединный; вторая средняя ($\overline{y}'_1$) – к августу 1999 г. и т. д.

4. Определяются средние из сглаженных (центрированных) скользящих для одноименных месяцев ($\overline{\overline{y}}'_1$ гр. 6):

$$\text{январь} - \ \overline{\overline{y}}'_1 = \frac{15{,}1+15{,}7}{2} = 15{,}4;$$

$$\text{февраль} - \ \overline{\overline{y}}'_2 = \frac{15{,}2+15{,}8}{2} = 15{,}5 \text{ и т. д.}$$

5. Исчисляются индексы сезонности для каждого месяца, %:

$$\text{январь} - \ i_c = \frac{12{,}6}{15{,}4} \cdot 100 = 81{,}8;$$

$$\text{февраль} - \ i_c = \frac{13{,}4}{15{,}5} \cdot 100 = 86{,}5.$$

6.13. По грузовому автотранспортному предприятию имеются следующие данные об объеме перевозок грузов за отчетный год (табл. 6.31).

Построить модель сезонных колебаний в объеме перевозок, используя первую гармонику ряда Фурье.

Таблица 6.31

Месяц	Среднесуточный объем перевозок, тыс. т	Месяц	Среднесуточный объем перевозок, тыс. т
Январь	10,3	Июль	11,8
Февраль	10,6	Август	12,4
Март	10,9	Сентябрь	11,7
Апрель	11,3	Октябрь	11,2
Май	11,2	Ноябрь	10,8
Июнь	11,7	Декабрь	10,5

Решение

Первая гармоника имеет вид:

$$\hat{y}_t = a_0 + a_1 \cdot \cos t + b_1 \cdot \sin t.$$

Параметры уравнения определяются по формулам

$$a_0 = \frac{\sum y}{n}; \quad a_1 = \frac{2}{12}\sum y \cdot \cos t; \quad b_1 = \frac{2}{12}\sum y \cdot \sin t.$$

Для расчета параметров уравнения использована следующая вспомогательная табл. 6.32.

Параметры искомого уравнения следующие:

$$a_0 = \frac{134{,}4}{12} = 11{,}2; \quad a_1 = \frac{2 \cdot (-4{,}7)}{12} = -0{,}78; \quad b_1 = \frac{2 \cdot (-0{,}5)}{12} = -0{,}08.$$

Модель сезонной волны объема перевозок примет вид:

$$\hat{y}_t = 11{,}2 - 0{,}78 \cdot \cos t - 0{,}08 \cdot \sin t.$$

На основе полученного уравнения определяются расчетные (теоретические) уровни для каждого месяца, тыс. т:

для января: $\hat{y}_t = 11{,}2 - 0{,}78 \cdot 1{,}0 - 0{,}08 \cdot 0 = 10{,}4$;

для февраля: $\hat{y}_t = 11{,}2 - 0{,}78 \cdot 0{,}866 - 0{,}08 \cdot 0{,}5 = 10{,}5$ и т. д.

Таблица 6.32

Месяц	Условное обозначение месяца, t	Среднесуточный объем перевозок, тыс. т	$\cos t$	$\sin t$	$y \cdot \cos t$	$y \cdot \sin t$	$\hat{y}_t$
Январь	0	10,3	1,0	0,0	10,3	0	10,4
Февраль	$1/6\pi$	10,6	0,866	0,5	9,2	5,3	10,5
Март	$1/3\pi$	10,9	0,5	0,866	5,5	9,4	10,7
Апрель	$1/2\pi$	11,3	0,0	1,0	0	11,3	11,1
Май	$2/3\pi$	11,2	–0,5	0,866	–5,6	9,7	11,5
Июнь	$5/6\pi$	11,7	–0,866	0,5	–10,1	5,9	11,8
Июль	π	11,8	–1,0	0,0	–11,8	0	12,0
Август	$7/6\pi$	12,4	–0,866	–0,5	–10,7	–6,2	11,9
Сентябрь	$4/3\pi$	11,7	–0,5	–0,866	–5,9	–10,1	11,7
Октябрь	$3/2\pi$	11,2	0,0	–1,0	0	–11,2	11,3
Ноябрь	$5/3\pi$	10,8	0,5	–0,866	5,4	–9,4	10,9
Декабрь	$11/6\pi$	10,5	0,866	–0,5	9,1	–5,2	10,5
Итого		134,4	–	–	–4,7	–0,5	134,3

6.2. Задачи для самостоятельной работы

6.14. Укажите, к какому виду относятся ряды, характеризующие размеры (объемы) следующих социально-экономических явлений:

а) численность населения (по данным переписей населения);

б) протяженность автомобильных дорог с усовершенствованным покрытием (по состоянию на конец каждого года);

в) объем реализованной продукции по кварталам;

г) жилищный фонд (общая площадь на конец года);

д) удельный вес объема перевезенного железнодорожным транспортом груза в общем объеме перевозок по годам;

е) средний размер дохода населения по годам;

ж) удельный вес городского и сельского населения региона;

з) среднемесячная (списочная) численность работников предприятия;

и) численность студентов (на конец учебного года);

к) объем инвестиций, вложенных в различные отрасли экономики;

л) количество дорожно-транспортных происшествий в регионе;

м) численность врачей на 1000 жителей района;

н) коэффициент текучести кадров на предприятии по месяцам;

о) число вкладов населения в учреждениях Сберегательного банка России;

п) удельный вес затрат на услуги связи в общем объеме затрат предприятий и организаций, отдельных отраслей экономики;

р) удельный вес иностранных инвестиций в предприятия и организации транспорта и связи;

с) число приватизированных предприятий (объектов) транспорта.

6.15. Производство основных товаров длительного пользования для населения России характеризуется следующими данными (тыс. шт.) (табл. 6.33).

Таблица 6.33

Наименование товара	1995 г.	1996 г.	1997 г.	1998 г.	1999 г.
Телевизоры	1005	313	327	329	278
В том числе цветного изображения	370	102	252	293	260
Холодильники и морозильники	1789	1064	1186	1043	1168
Легковые автомобили	896	868	986	840	956
Фотоаппараты	296	217	143	60,1	81,2

Определите показатели динамики (цепные, базисные) производства каждого вида товара длительного пользования. Сопоставьте приведенные ряды динамики, используя среднегодовые показатели динамики.

Сформулировать выводы.

6.16. Численность населения РФ на начало года характеризуется следующими данными (табл. 6.34).

Таблица 6.34

Год	1995	1996	1997	1998	1999	2000
Численность населения, млн чел.	147,9	147,6	147,1	146,7	146,3	145,6

Требуется, используя данные о численности населения и производстве товаров длительного пользования (задача 6.15):

а) построить ряды динамики выпуска каждого вида товаров на душу населения;

б) проанализировать динамику полученных показателей, исчислив коэффициенты опережения среднегодовых темпов прироста;

в) изобразить графически динамику выпуска каждого вида товаров на душу населения.

6.17. Число вкладов населения в учреждениях Сберегательного банка России по региону на начало года представлено в табл. 6.35.

Таблица 6.35

Год	1997	1998	1999	2000
Число вкладов, млн	141,0	203,7	210,9	234,2

Определить ежегодные абсолютные приросты, коэффициенты роста и темпы прироста числа вкладов с постоянной и переменной базой.

6.18. Известны темпы роста инвестиций по двум регионам (в % к 1998 г.) (табл. 6.36).

Таблица 6.36

Регион	1999 г.	2000 г.	2001 г.
А	120	254	308
Б	108	190	240

Определить темпы роста инвестиций за каждый год по сравнению с предшествующим годом и среднегодовые темпы роста инвестиций для каждого региона.

Сформулируйте вывод.

6.19. Имеются следующие данные о мощности электростанций региона (на конец года, млн кВт), табл. 6.37.

Таблица 6.37

Год	Мощность электростанций (на конец года), млн кВт	Цепные показатели динамики			
		абсолютный прирост, млн кВт	коэффициент роста	темп прироста, %	абсолютное значение 1% прироста, млн кВт
1995	22,3	1,3			
1996					
1997				2,12	0,24
1998			1,041		
1999			1,071		
2000				1,85	

Требуется исчислить отсутствующие в таблице сведения за 1995 – 2000 гг., а также определить, в каком периоде (в 1995 – 1997 гг. или в 1998 – 2000 гг.) были более высокие абсолютный и относительный приросты мощности электростанций региона.

6.20. Имеются следующие данные о приеме студентов в высшие учебные заведения России, тыс. чел. (табл. 6.38).

Требуется:

1) исчислить отсутствующие в таблице сведения о приеме студентов за 1996 – 2000 гг.;

2) проанализировать динамику изучаемого явления, опираясь на рассчитанные показатели динамики.

Таблица 6.38

Год	Принято студентов, тыс. чел.	Цепные показатели динамики			
		абсолютный прирост, тыс. чел.	темп роста, %	темп прироста, %	абсолютное значение 1% прироста, тыс. чел.
1996	2791	146			
1997			106,2		
1998				9,5	
1999					
2000		475			35,98

6.21. В табл. 6.39 представлены данные о пассажирообороте автобусного транспорта региона.

Таблица 6.39

Год	Пассажирооборот, млрд пасс.-км	Цепные показатели динамики			
		абсолютный прирост, млрд пасс.-км	коэффициент роста	темп прироста, %	абсолютное значение 1% прироста, млрд пасс.-км
1996	360,2	–	–	–	–
1997		14,5			
1998			1,037		
1999					
2000		10,8			4,018

Определить недостающие уровни и цепные показатели динамики.

6.22. В табл. 6.40 представлены данные о перевозке грузов речным пароходством.

Таблица 6.40

Год	Объем перевозок грузов, млн т	Цепные показатели динамики		
		абсолютный прирост, млн т	темп роста, %	темп прироста, %
1995	520,6	–	–	–
1996			105,4	
1997		–9,0		
1998				5,8
1999		26,4		
2000			101,7	

Определить недостающие уровни и цепные показатели динамики.

6.23. По группе предприятий имеются следующие данные (табл. 6.41).

Таблица 6.41

№ предприятия	Удельный вес в общем объеме продукции в прошлом году, %	Прирост объема производства продукции против прошлого года, %
1	30,5	7,3
2	24,3	10,5
3	45,2	18,4

Определить удельные веса предприятий в общем объеме продукции в отчетном году.

6.24. Выработка изделия на предприятии в 1997 г. составила 4 тыс. шт., а в 2000 г. 4,6 тыс. шт.

Определить методом интерполяции выработку изделия в 1998 и 1999 гг.

6.25. Приведены следующие данные (табл. 6.42).

Таблица 6.42

Год	1997	1998	1999	2000
Уровень ряда	10			14

Определить неизвестные уровни, предполагая их линейное изменение.

6.26. Среднее расстояние перевозки грузов в международном сообщении по годам характеризуется следующими данными (табл. 6.43).

Таблица 6.43

Год	1996	1997	1998	1999	2000
Среднее расстояние перевозки, км	512	255	223	210	185

Произвести аналитическое выравнивание с последующей экстраполяцией для 2001 г.

6.27. Удельный вес городского населения региона увеличился с 1 января 1995 г. по 1 января 2000 г. с 36,2 до 42,8%.

Определить показатели динамики численности городского и сельского населения региона, если общая численность населения данного региона за этот период возросла на 8,4%.

6.28. Поступление денежных средств от реализации продукции, работ и услуг за отчетный год по предприятию следующее (табл. 6.44).

Таблица 6.44

Месяц	Фактически поступило на расчетный счет, млн руб.	Месяц	Фактически поступило на расчетный счет, млн руб.
Январь	15,2	Июль	16,1
Февраль	14,8	Август	17,3
Март	14,5	Сентябрь	16,9
Апрель	16,0	Октябрь	15,8
Май	16,7	Ноябрь	17,5
Июнь	15,4	Декабрь	18,0

Требуется:

1) определить начальный, конечный и средний уровень ряда динамики;

2) построить ряд динамики с нарастающими итогами по кварталам года;

3) определить среднемесячный уровень поступления денежных средств за каждый квартал.

6.29. Динамика объема реализации услуг коммунальных предприятий города в процентах к 1996 г. составила: 1997 г. – 108,0; 1998 г. – 110,5; 1999 г. – 125,0; 2000 г. – 153,2.

Определить:

1) коэффициенты роста для 1999 и 2000 гг. по сравнению с 1998 г.;

2) среднегодовой темп прироста за период 1996 – 2000 гг.

6.30. На 1 октября в списке предприятия «Акрос» числилось 25 человек; с 10 октября были приняты на работу 6 человек, а с 12 октября были уволены по собственному желанию 4 человека. С 25 октября на предприятие были приняты 6 человек.

На предприятии «Восход» на 1 октября числилось 32 человека; с 15 октября были приняты на работу 5 человек, а с 28 октября уволилось 6 человек.

Требуется:

1) определить, на каком предприятии и насколько среднесписочная численность в октябре была больше (в абсолютном выражении и в процентах);

2) изобразить динамику численности работников каждого предприятия с помощью линейной диаграммы.

6.31. Жилищный фонд России характеризуется следующими данными (табл. 6.45).

Таблица 6.45

	На 1 января 1997 г.	На 1 января 1998 г.	На 1 января 1999 г.	На 1 января 2000 г.
Весь жилищный фонд, млн кв. м	2580	2715	2745	2770

Определить, на сколько кв. м и процентов в среднем ежегодно увеличивался жилищный фонд России.

6.32. Имеются следующие данные о числе пассажиров, перевезенных метрополитеном по городам России (табл. 6.46).

Таблица 6.46

	1996 г.	1997 г.	1998 г.
Перевезено пассажиров, млн чел.	4173	4128	4146

Определить среднегодовой темп роста и темп прироста перевозок пассажиров метрополитеном.

6.33. Погрузка вагонов по отделению дороги характеризуется следующими данными за апрель отчетного года (табл. 6.47).

Таблица 6.47

Числа месяца	Погружено вагонов	Числа месяца	Погружено вагонов	Числа месяца	Погружено вагонов
1	218	11	210	21	203
2	190	12	184	22	195
3	105	13	200	23	214
4	185	14	163	24	177
5	200	15	112	25	209
6	170	16	174	26	197
7	175	17	103	27	169
8	98	18	170	28	181
9	208	19	188	29	170
10	164	20	152	30	210

Требуется:

1) для погашения колебаний и выявления основной тенденции роста числа погруженных вагонов произвести сглаживание динамического ряда с помощью трехчленной переменной и скользящей средней;

2) объяснить полученные результаты.

6.34. Стоимость основных средств на предприятии за отчетный год составила (млн руб.):

на 1 января – 4,8;
на 1 апреля – 4,0;
на 1 июля – 5,0;
на 1 октября – 6,0;
на 1 января (следующего года) – 5,2.

Определить среднегодовую стоимость основных средств предприятия и величину 1% прироста за год.

6.35. Остаток средств на расчетном счете предприятия составил на 1 января 2000 г. 180 тыс. руб.; 15 января поступило на расчетный счет 900 тыс. руб.; 22 января списано с расчетного счета 530 тыс. руб.; 27 января поступило на расчетный счет 380 тыс. руб. С 28 января до конца месяца остаток средств на расчетном счете не изменился.

Определить среднесуточный остаток средств на расчетном счете предприятия в январе.

6.36. По ТЭЦ имеются следующие данные об отпуске теплоэнергии (тыс. Гкал) за 3 года по месяцам (табл. 6.48).

Таблица 6.48

Месяц	1998 г.	1999 г.	2000 г.	Месяц	1998 г.	1999 г.	2000 г.
Январь	715	735	816	Июль	111	116	122
Февраль	608	612	697	Август	115	124	126
Март	502	540	638	Сентябрь	168	204	177
Апрель	401	405	442	Октябрь	340	450	428
Май	153	185	206	Ноябрь	443	545	467
Июнь	181	126	137	Декабрь	668	631	666

Требуется на основе приведенных данных выявить наличие сезонной неравномерности в отпуске теплоэнергии и измерить ее степень.

6.37. По таксомоторному предприятию имеются следующие данные о величине платного пробега за 3 года, тыс. км (табл. 6.49).

Требуется на основе приведенных данных выявить наличие сезонной неравномерности в таксомоторных перевозках и измерить ее степень.

Таблица 6.49

Месяц	1998 г.	1999 г.	2000 г.	Месяц	1998 г.	1999 г.	2000 г.
Январь	60,0	70,0	102,0	Июль	80,3	90,3	116,4
Февраль	62,0	77,4	105,0	Август	86,5	94,6	120,0
Март	66,4	78,2	107,0	Сентябрь	79,0	94,0	118,7
Апрель	70,0	80,0	110,5	Октябрь	76,4	92,0	115,0
Май	78,4	88,4	113,7	Ноябрь	75,0	90,0	107,5
Июнь	80,0	89,5	115,0	Декабрь	70,4	85,5	98,2

6.38. Имеются следующие данные о ежесуточной добыче угля по шахте за первую декаду (табл. 6.50).

Таблица 6.50

День	1	2	3	4	5	6	7	8	9	10
Добыча угля, т	800	790	804	808	805	810	800	817	820	832

Произвести сглаживание ряда методом трехчленной переменной и трехчленной скользящей средней. Дать график первичного и сглаженного рядов.

6.39. Производство цемента в регионе характеризуется следующими данными (табл. 6.51).

Таблица 6.51

Год	1991	1992	1993	1994	1995	1996	1997	1998	1999	2000
Производство цемента, млн т	64	72	80	84	86	90	95	100	104	109

Требуется:

1) провести аналитическое выравнивание по прямой и использовать полученное уравнение для экстраполяции уровней 2001 и 2002 гг.;

2) построить график первичного и выравненного рядов.

6.40. Имеются следующие данные о росте производительности труда в отрасли (к 1994 г.) (табл. 6.52).

Таблица 6.52

Год	1995	1996	1997	1998	1999	2000
Коэффициент роста	1,29	1,37	1,48	1,58	1,64	1,77

Определить:

1) на сколько процентов возросла производительность труда в 2000 г. по сравнению с 1995 и 1999 гг.;

2) среднегодовой темп роста производительности труда за период 1995 – 2000 гг.

6.41. Имеются следующие данные об удельных расходах условного топлива на производство теплоэнергии (кг/Гкал) на ТЭЦ по годам (табл. 6.53).

Таблица 6.53

Год	1992	1993	1994	1995	1996	1997	1998	1999	2000
Удельный расход условного топлива, кг/Гкал	167,6	165,8	167,4	168,0	167,5	167,2	166,5	166,5	166,4

Требуется:

1) произвести сглаживание ряда методом трехлетней скользящей средней;

2) произвести аналитическое выравнивание ряда по прямой;

3) методом экстраполяции определить уровни 2001 и 2002 гг.;

4) начертить график первичного и выравненного рядов.

6.42. Продажа детских велосипедов в магазинах региона характеризуется следующими данными (табл. 6.54).

Построить модель уровня продажи детских велосипедов, используя первую гармонику ряда Фурье.

Таблица 6.54

Месяц	Продано, тыс. шт.	Месяц	Продано, тыс. шт.
Январь	30,8	Июль	190,1
Февраль	61,2	Август	138,9
Март	120,7	Сентябрь	82,3
Апрель	240,6	Октябрь	50,4
Май	280,1	Ноябрь	36,7
Июнь	255,3	Декабрь	35,3

6.43. Добыча и производство газа в регионе характеризуются следующими данными (табл. 6.55).

Таблица 6.55

	1994 г.	1995 г.	1996 г.	1997 г.	1998 г.	1999 г.	2000 г.
Добыча и производство газа, млн $м^3$	157	169	181	198	212	221	236

Требуется:

1) произвести аналитическое выравнивание по показательной кривой;

2) определить интервальный прогноз для 2001 г., гарантируя результат с вероятностью 0,95.

6.44. Золотовалютные резервы РФ в 2000 г. представлены в табл. 6.56.

Таблица 6.56

Дата	Млн долл. США	Дата	Млн долл. США
1 января	12456	1 июля	20996
1 февраля	12948	1 августа	23302
1 марта	13657	1 сентября	23731
1 апреля	15532	1 октября	24997
1 мая	17091	1 ноября	25880
1 июня	19570	1 декабря	27667
		1 января 2001 г.	27951

Определить:

1) средний уровень золотовалютных резервов РФ за 2000 г.;

2) пункты роста золотовалютных резервов в 2000 г.

Сформулировать вывод.

6.45. Котировка драгоценных металлов, установленная ЦБ РФ представлена в табл. 6.57.

Таблица 6.57

Дата котировки на декабрь 2000 г.	Покупка, руб./г		
	золото	серебро	платина
25	232,77	3,76	478,70
26	233,60	3,78	480,40
27	235,05	3,79	486,06
28	234,97	3,82	493,23
29	233,30	3,79	486,85

Сопоставить средний темп прироста покупных цен на драгоценные металлы в последнюю пятидневку 2000 г.

Сформулировать вывод.

6.46. В табл. 6.58 представлены данные о квартальном изменении цен на реализуемый продукт.

Таблица 6.58

Кварталы	I	II	III	IV
Изменение цен (в % к предыдущему кварталу)	+3,0	+1,5	+2,0	+3,5

Определить, на сколько процентов выросли цены за год.

6.47. Производство стали в России характеризуется следующими данными (табл. 6.59).

Требуется:

1) определить объем производства кислородно-конвертерной стали и электростали;

2) сопоставить три ряда динамики, используя среднегодовые показатели динамики.

6.48. Внешняя торговля России со странами дальнего зарубежья (млрд руб.) характеризуется следующими данными (табл. 6.60).

Таблица 6.59

Показатель	1995 г.	1996 г.	1997 г.	1998 г.	1999 г.
Произведено стали – всего, млн т	51,6	49,3	48,5	43,7	51,5
в том числе (% от общего объема производства):					
кислородно-конвертерная	45,2	51,5	55,1	59,5	58,8
электросталь	12,8	12,6	12,8	12,8	13,2

Таблица 6.60

Показатель	1998 г.	1999 г.	2000 г.
Экспорт	59,1	63,5	90,8
Импорт	43,7	29,2	31,4

Сопоставить средние годовые показатели динамики экспорта и импорта.

Сформулировать вывод.

Ответы к задачам для самостоятельной работы

6.19. 1994 г.: y = 21,0; 1995 г.: $K_р$ = 1,0619, $Т_п$ = 6,19%, А = 0,21 млн кВт; 1996 г.: y = 24 млн кВт, Δ = 1,7 млн кВт; $K_р$ = 1,0762; $Т_п$ = 7,62%; А = 0,22 млн кВт; 1997 г.: y = 24,51 млн кВт, Δ = 0,51 млн кВт, $K_р$ = 1,0212; 1998 г.: y =25,51 млн кВт, Δ = 1 млн кВт, $Т_п$ = 4,1%, А = 0,24 млн кВт; 1999 г.: y = 27,32 млн кВт, Δ = 1,81 млн кВт, $Т_п$ = 7,1%, А = 0,25 млн кВт; 2000 г.: y = 27,83 млн кВт, Δ = 0,51 млн кВт, $K_р$ = 1,0185, А = 0,28 млн кВт; 1995–1997 гг.: $\overline{\Delta}$ = (24,5 – 22,3) : 2 = 1,1 млн кВт, $K_р = \sqrt{24,51 : 22,3} =$ =1,048; 1998 – 2000 гг.: $\overline{\Delta}$ = 1,16 млн кВт, $\overline{K}_р$ = 1,044. **6.27.** Численность городского населения увеличилась на 28,2%, сельского населения снизилась на 2,8%. **6.29.** $K_р$ = 1,131; $K_р$ = 1,386; $\overline{K}_р$ = 1,1125; $\overline{Т}_п$ = 11,25%. **6.31.** $\overline{\Delta}$ = 63,33 млн кв. м, $\overline{Т}_п$ = 2,4%. **6.32.** $\overline{K}_р$ = 0,997, $\overline{Т}_р$ = 99,7%; $\overline{Т}_п$ = –0,3%. **6.34.** $\overline{y}$ = 5,0 млн руб., $\overline{Т}_п$ = 2,02%, $\overline{\Delta}$ = 0,1 млн руб., $\overline{А}$ = 49,5 тыс. руб. **6.35.** $\overline{y}$ = 563,871 тыс. руб. **6.40.** Увеличение на 37,21%; увеличение на 7,93%; $Т_р = \sqrt[5]{1,77 : 1,29} \cdot 100 = 106,55\%$. **6.45.** Золото – 0,5%; серебро – 0,19%; платина – 0,42%. **6.46.** Увеличение на 10,37%. **6.48.** Экспорт: $\overline{\Delta}$ = 15,85; $\overline{K}_р$ = 1,24; $\overline{Т}_п$ = 24,0%; $\overline{А}$ = 0,66 млрд руб.; импорт: $\overline{\Delta}$ = –6,15; $\overline{K}_р$ = 0,847; $\overline{Т}_п$ = –15,3%; $\overline{А}$ = –0,40 млрд руб.

ГЛАВА 7

Индексы и их использование в экономико-статистических исследованиях

Индекс – относительная величина, характеризующая изменение уровней сложных социально-экономических показателей во времени, в пространстве или по сравнению с планом. Сложный показатель состоит из непосредственно несоизмеримых (несуммируемых) элементов. Например, предприятие выпускает несколько видов продукции, но получить общий итог объема продукции путем суммирования количества различных ее видов в натуральном выражении нельзя.

Индексные показатели вычисляются на высшей ступени статистического обобщения и опираются на результаты сводки и обработки данных статистического наблюдения. С их помощью решаются следующие *основные задачи*:

- характеристика общего изменения сложного экономического показателя и отдельных его элементов;
- измерение влияния факторов на общую динамику сложного показателя, включая характеристику влияния изменения структуры явления.

Индекс является результатом сравнения двух одноименных показателей, поэтому при их вычислении различают сравниваемый уровень (числитель индексного отношения), называемый **текущим** или **отчетным**, и уровень, с которым производится сравнение (знаменатель индексного отношения), называемый **базисным.** Выбор базы определяется целью исследования.

При изучении динамики за базисную величину принимают размер показателя в каком-либо периоде, предшествующем отчетному. При этом возможны два способа расчета индексов – цепной и базисный. **Цепные индексы** получают путем сопоставления текущих уровней с предшествующим. Следовательно, база сравнения непрерывно меняется. **Базисные индексы** получают путем сопоставления с уровнем какого-либо одного периода, принятого за базу сравнения.

При территориальных сравнениях за базу принимают данные другой территории.

При использовании индексов как показателей выполнения плана за базу сравнения принимаются плановые показатели.

В зависимости от содержания и характера изучаемых социально-экономических показателей различают индексы количественных (объемных) показателей и индексы качественных показателей.

К индексам количественных (объемных) показателей относятся индексы физического объема производства продукции, физического объема потребления продукции (производственного и личного) и индексы других показателей, размеры которых характеризуются абсолютными величинами.

К индексам качественных показателей относятся индексы цен, себестоимости, индексы средней заработной платы, производительности труда. Качественный показатель характеризует уровень изучаемого результативного показателя в расчете на количественную единицу и определяется путем деления результативного показателя на количественный показатель, на единицу которого он определяется. Например, средняя заработная плата определяется путем деления фонда заработной платы на численность работников; производительность труда определяется путем деления общего объема выработанной продукции на численность работников.

По степени охвата элементов совокупности различают индивидуальные и сводные (общие) индексы. **Индивидуальные индексы** характеризуют изменение одного элемента совокупности. **Сводные индексы** характеризуют изменение сложного явления в целом. Если индексы охватывают не все элементы сложного явления, а лишь некоторую часть, то их принято называть групповыми. В зависимости от способа исчисления общих (сводных) индексов различаются агрегатные индексы и средние взвешенные индексы.

Для удобства применения индексного метода, составления формул индексов и их использования в статистико-экономическом анализе в теории статистики разработана определенная символика и применяются соответствующие условные обозначения.

Каждая индексируемая величина имеет свое символическое обозначение:

q – количество продукции одного вида в натуральном выражении;

p – цена за единицу продукции;

z – себестоимость единицы продукции;

t – затраты труда (рабочего времени) на единицу продукции.

Индексы по отдельным элементам изучаемого сложного экономического явления (т. е. индивидуальные индексы) обозначаются символом i, у которого проставляется символ соответствующей индексируемой величины. Например:

i_q – индивидуальный индекс объема (количества) отдельного вида продукции;

i_p – индивидуальный индекс цен на отдельный вид продукции (товара);

i_z – индивидуальный индекс себестоимости единицы отдельного вида продукции;

i_{qp} – индекс стоимости отдельного вида продукции;

i_{qz} – индекс денежных затрат на выпуск одного вида продукции;

i_{qt} – индекс затрат труда на выпуск (производство) одного вида продукции.

Общий (сводный) индекс изучаемого сложного экономического явления обозначается символом I, у которого отражается символ индексируемой величины. Например:

I_q – общий индекс физического объема продукции;

I_p – общий индекс цен;

I_z – общий индекс себестоимости;

I_{qp} – общий индекс стоимости всех видов продукции;

I_{qz} – общий индекс затрат на производство всех видов продукции;

I_{qt} – общий индекс затрат труда на выпуск всех видов продукции.

Для отражения базисных периодов времени применяются специальные обозначения, которые пишутся внизу символа используемых при написании индекса величин. Базисный период, с данными которого производится сравнение, обозначается нулевым значением, первый отчетный период – единицей и т. д. Кроме того, обозначения сравниваемого и базисного периодов можно проставлять внизу символа индекса (например, $I_{q_{1/0}}$).

Индексы количественных показателей

Индивидуальный индекс физического объема выпуска продукции характеризует изменение выпуска (реализации или потребления) одного вида продукции и определяется по формуле

$$i_{q_{1/0}} = q_1 : q_0,$$

где q_1 и q_0 – количество продукции данного вида в натуральном выражении соответственно в текущем и базисном периодах.

Индивидуальный индекс затрат на выпуск продукции показывает изменение затрат на производство одного вида продукции и имеет следующий вид:

$$i_{qz_{1/0}} = q_1 z_1 : q_0 z_0,$$

где z_1 и z_0 – себестоимость единицы продукции данного вида соответственно в текущем и базисном периодах;

$q_1 z_1$ и $q_0 z_0$ – сумма затрат на выпуск продукции данного вида соответственно в текущем и базисном периодах.

Индивидуальный индекс стоимости продукции:

$$i_{qp_{1/0}} = q_1 p_1 : q_0 p_0,$$

где p_1 и p_0 – цена единицы продукции данного вида соответственно в текущем и базисном периодах;

$q_1 p_1$ и $q_0 p_0$ – стоимость продукции данного вида соответственно в текущем и базисном периодах.

Агрегатный индекс физического объема продукции $Iq_{1/0}$ характеризует изменение выпуска всей совокупности продукции и исчисляется по формуле

$$I_{q_{1/0}} = \frac{\sum q_1 p_0}{\sum q_0 p_0},$$

где q_1 и q_0 – количество выработанных единиц отдельных видов продукции соответственно в отчетном и базисном периодах;

p_0 – цена единицы отдельного вида продукции в базисном периоде.

Такой вариант построения агрегатного индекса был предложен Э. Ласпейресом в 1864 г.

В агрегатном индексе физического объема продукции индексируемой величиной является количество продукции (q); цена (p) служит коэффициентом соизмерения (соизмерителем).

$$\Delta^q_{\sum qp} = \sum q_1 p_0 - \sum q_0 p_0,$$

где $\Delta^q_{\sum qp}$ – абсолютное изменение общей стоимости продукции за счет изменения выпуска продукции.

При вычислении индекса физического объема продукции возможны разные решения – в зависимости от выбора коэффициента соизмерения. В качестве коэффициента соизмерения можно также использовать цены отчетного периода (p_1) или сопоставимые (фиксированные – P_c). Тогда формулы агрегатного индекса имеют следующий вид:

$$I_{q_{1/0}} = \frac{\sum q_1 p_1}{\sum q_0 p_1}; \quad I_{q_{1/0}} = \frac{\sum q_1 p_c}{\sum q_0 p_c}.$$

Агрегатный индекс с соизмерителями отчетного периода был предложен в 1874 г. Г. Пааше.

Сопоставимые (фиксированные или неизменные) цены применяются в тех случаях, когда изучается динамика объемов явлений за несколько последовательных периодов времени. Периодически эти цены пересматриваются или заменяются в соответствии с изменением особенностей ценообразования.

Кроме того, в качестве соизмерителей может быть использована себестоимость единицы продукции, а также затраты рабочего времени на единицу продукции. В этом случае агрегатный индекс физического объема определяется по формулам

$$I_{q_{1/0}} = \frac{\sum q_1 z_0}{\sum q_0 z_0}; \quad I_{q_{1/0}} = \frac{\sum q_1 t_0}{\sum q_0 t_0},$$

где z_0 – себестоимость единицы продукции каждого вида в базисном периоде;

t_0 – затраты рабочего времени на производство единицы продукции каждого вида в базисном периоде.

Средние взвешенные индексы физического объема продукции применяются в том случае, если известны индивидуальные индексы объема по отдельным видам продукции и стоимость отдельных видов продукции (или затраты на отдельные виды продукции) в базисном или отчетном периоде.

Средний взвешенный арифметический индекс физического объема продукции

$$I_{q_{1/0}} = \frac{\sum i_q q_0 p_0}{\sum q_0 p_0},$$

где i_q – индивидуальный индекс по каждому виду продукции;
$q_0 p_0$ – стоимость продукции каждого вида в базисном периоде.

Средний взвешенный гармонический индекс физического объема продукции:

$$I_{q_{1/0}} = \frac{\sum q_1 p_1}{\sum \frac{1}{i_q} q_1 p_1},$$

где $q_1 p_1$ – стоимость продукции каждого вида в текущем периоде.

Агрегатный индекс затрат на выпуск всей продукции имеет следующий вид:

$$I_{qz_{1/0}} = \frac{\sum q_1 z_1}{\sum q_0 z_0},$$

где $q_1 z_1$ и $q_0 z_0$ – затраты на выпуск продукции каждого вида соответственно в отчетном и базисном периодах.

$$\Delta^{q,z}{}_{\sum q,z} = \sum q_1 z_1 - \sum q_0 z_0,$$

где $\Delta^{q,z}{}_{\sum q,z}$ – абсолютное изменение общей суммы затрат на выпуск продукции за счет изменения количества выработанной продукции и ее себестоимости.

Агрегатный индекс стоимости продукции (товарооборота):

$$I_{qp_{1/0}} = \frac{\sum q_1 p_1}{\sum q_0 p_0},$$

$$\Delta^{q,p}{}_{\sum q,p} = \sum q_1 p_1 - \sum q_0 p_0,$$

где $\Delta^{q,p}{}_{\sum q,p}$ – абсолютное изменение общей стоимости продукции за счет изменения количества продукции и цен.

Агрегатный территориальный индекс физического объема производства (реализации) продукции имеет вид:

$$I_{q_{В/Г}} = \frac{\sum q_В \cdot \overline{p}}{\sum q_Г \cdot \overline{p}},$$

где $q_В$ и $q_Г$ – количество выпущенной (реализованной) продукции каждого вида в натуральном выражении соответственно на территории В и Г;
$\overline{p}$ – средняя цена каждого вида продукции по сравниваемым территориям, определяемая как средняя взвешенная арифметическая.

Индексы качественных показателей

Индивидуальные индексы цен, себестоимости, затрат рабочего времени на единицу продукции характеризуют изменение цен, себестоимости, затрат рабочего времени по каждому виду продукции:

$$i_{p_{1/0}} = p_1 : p_0;\ i_z = z_1 : z_0;\ i_t = t_1 : t_0,$$

где p_1 и p_0 – цена за единицу продукции каждого вида соответственно в текущем и базисном периодах;
z_1 и z_0 – себестоимость единицы продукции каждого вида соответственно в текущем и базисном периодах;
t_1 и t_0 – затраты рабочего времени на единицу продукции каждого вида соответственно в текущем и базисном периодах.

Агрегатный индекс цен характеризует среднее изменение цен по совокупности различных видов продукции и исчисляется по формуле

$$I_{p_{1/0}} = \frac{\sum p_1 q_1}{\sum p_0 q_1}$$

(этот вариант индекса был предложен Г. Пааше).

Индексируемой величиной является цена (p), количество продукции (q) носит название веса.

$$\Delta^{p}_{\sum qp} = \sum p_1 q_1 - \sum p_0 q_1,$$

где $\Delta^{p}_{\sum qp}$ – абсолютное изменение всей стоимости продукции за счет изменения цен.

Для характеристики среднего изменения цен на потребительские товары (потребительскую корзину) агрегатный индекс цен целесообразно определять по формуле

$$I_{p_{1/0}} = \frac{\sum p_1 q_0}{\sum p_0 q_0},$$

(вариант индекса был предложен Э. Ласпейресом).

На основе этого индекса целесообразно определять индекс покупательной способности рубля.

$$I_{\text{покуп. спос.}} = 1 : I_p.$$

Средние взвешенные индексы цен применяются в том случае, если известны индивидуальные индексы цен по отдельным видам продукции, а также стоимость отдельных видов продукции.

Средний взвешенный арифметический индекс цен:

$$I_{p_{1/0}} = \frac{\sum i_p p_0 q_0}{\sum p_0 q_0},$$

где i_p – индивидуальный индекс цен по каждому виду продукции;
$p_0 q_0$ – стоимость продукции каждого вида в базисном периоде.

Средний взвешенный гармонический индекс цен:

$$I_{p_{1/0}} = \frac{\sum p_1 q_1}{\sum \frac{1}{i_p} p_1 q_1},$$

где $p_1 q_1$ – стоимость продукции каждого вида в текущем периоде.

Агрегатный территориальный индекс цен имеет вид:

$$I_{p_{\text{В/Г}}} = \frac{\sum p_{\text{В}} \cdot q_{\text{В}}}{\sum p_{\text{Г}} \cdot q_{\text{В}}},$$

где $p_{\text{В}}, p_{\text{Г}}$ – цена за единицу продукции каждого вида соответственно на территории В и Г;
$q_{\text{В}}$ – количество выработанной (реализованной) продукции каждого вида в натуральном выражении по территории В.

В качестве фиксированного показателя (веса) в данном индексе принят объем продукции территории В.

При построении данного индекса в качестве веса может быть принят также объем продукции той территории, с которой производится сравнение (Г) или суммарный объем продукции двух территорий:

$$I_{p_{\text{В/Г}}} = \frac{\sum p_{\text{В}} \cdot q_{\text{Г}}}{\sum p_{\text{Г}} \cdot q_{\text{Г}}}; \quad I_{p_{\text{В/Г}}} = \frac{\sum p_{\text{В}} \cdot q_{(\text{В+Г})}}{\sum p_{\text{Г}} \cdot q_{(\text{В+Г})}}.$$

При различных приемах «взвешивания» получаются различные числовые значения территориального индекса цен. В практике расчетов предпочтение отдается первому варианту.

Агрегатные индексы себестоимости и затрат рабочего времени на единицу продукции исчисляются по такому же принципу, как и агрегатные индексы цен. Их формулы следующие:

$$I_{z_{1/0}} = \frac{\sum z_1 \cdot q_1}{\sum z_0 \cdot q_1}; \quad I_{t_{1/0}} = \frac{\sum t_1 \cdot q_1}{\sum t_0 \cdot q_1}.$$

Цепные и базисные индексы

Цепные индексы получают путем сопоставления показателя любого периода с показателем предшествующего периода.

Цепные индивидуальные индексы физического объема продукции следующие:

$$i_{q_{1/0}} = \frac{q_1}{q_0}; \quad i_{q_{2/1}} = \frac{q_2}{q_1}; \quad i_{q_{3/2}} = \frac{q_3}{q_2} \text{ и т. д.},$$

где q_1, q_2, q_3 – количество продукции одного вида в последовательных периодах времени.

Цепные агрегатные индексы физического объема продукции следующие:

$$I_{q_{1/0}} = \frac{\sum q_1 p_0}{\sum q_0 p_0}; \quad I_{q_{2/1}} = \frac{\sum q_2 p_1}{\sum q_1 p_1}; \quad I_{q_{3/2}} = \frac{\sum q_3 p_2}{\sum q_2 p_2} \text{ и т. д.}$$

Базисные индексы получают сравнением показателя любого периода с показателем какого-нибудь одного периода, принятого за базу сравнения.

Базисные индивидуальные индексы физического объема продукции следующие:

$$i_{q_{1/0}} = \frac{q_1}{q_0}; \quad i_{q_{2/0}} = \frac{q_2}{q_0}; \quad i_{q_{3/0}} = \frac{q_3}{q_0} \text{ и т. д.}$$

Произведение цепных индивидуальных индексов равно последнему базисному индексу:

$$i_{q_{1/0}} \cdot i_{q_{2/1}} = i_{q_{2/0}}; \; i_{q_{1/0}} \cdot i_{q_{2/1}} \cdot i_{q_{3/2}} = i_{q_{3/0}}.$$

Базисные агрегатные индексы физического объема продукции следующие:

$$I_{q_{1/0}} = \frac{\sum q_1 p_0}{\sum q_0 p_0}; \quad I_{q_{2/0}} = \frac{\sum q_2 p_0}{\sum q_0 p_0}; \quad I_{q_{3/0}} = \frac{\sum q_3 p_0}{\sum q_0 p_0}.$$

Базисный агрегатный индекс физического объема продукции может быть получен как произведение цепных агрегатных индексов при постоянных соизмерителях:

$$I_{q_{2/0}} = I_{q_{1/0}} \cdot I_{q_{2/1}} = \frac{\sum q_1 p_0}{\sum q_0 p_0} \cdot \frac{\sum q_2 p_0}{\sum q_1 p_0} = \frac{\sum q_2 p_0}{\sum q_0 p_0}$$

или

$$I_{q_{2/0}} = I_{q_{1/0}} \cdot I_{q_{2/1}} = \frac{\sum q_1 p_C}{\sum q_0 p_C} \cdot \frac{\sum q_2 p_C}{\sum q_1 p_C} = \frac{\sum q_2 p_C}{\sum q_0 p_C}.$$

Цепные индивидуальные индексы цен:

$$i_{p_{1/0}} = \frac{p_1}{p_0}; \quad i_{p_{2/1}} = \frac{p_2}{p_1}; \quad i_{p_{3/2}} = \frac{p_3}{p_2} \text{ и т. д.}$$

Базисные индивидуальные индексы цен:

$$i_{p_{1/0}} = \frac{p_1}{p_0}; \quad i_{p_{2/0}} = \frac{p_2}{p_0}; \quad i_{p_{3/0}} = \frac{p_3}{p_0} \text{ и т. д.}$$

Произведение цепных индивидуальных индексов цен равно последнему базисному индексу:

$$i_{p_{3/0}} = i_{p_{1/0}} \cdot i_{p_{2/1}} \cdot i_{p_{3/2}}.$$

Цепные агрегатные индексы цен:

$$I_{p_{1/0}} = \frac{\sum p_1 q_1}{\sum p_0 q_1}; \quad I_{p_{2/1}} = \frac{\sum p_2 q_2}{\sum p_1 q_2}; \quad I_{p_{3/2}} = \frac{\sum p_3 q_3}{\sum p_2 q_3} \text{ и т. д.}$$

Базисные агрегатные индексы цен:

$$I_{p_{1/0}} = \frac{\sum p_1 q_1}{\sum p_0 q_1}; \quad I_{p_{2/0}} = \frac{\sum p_2 q_2}{\sum p_0 q_2}; \quad I_{p_{3/0}} = \frac{\sum p_3 q_3}{\sum p_0 q_3} \text{ и т. д.}$$

Агрегатные индексы качественных показателей, рассчитанные по формуле Пааше, всегда являются индексами с меняющимися весами, поэтому цепной метод исчисления базисных индексов не применяется.

Если же воспользоваться формулой Ласпейреса для расчета агрегатных индексов цен при условии постоянных весов q для всех периодов, то базисные индексы могут быть определены на основе цепных.

Изучение динамики качественных показателей по нескольким единицам (предприятиям, территориям, странам)

Анализ динамики уровней качественных показателей по нескольким единицам означает анализ динамики уровней средних величин различных экономических показателей (средней себестоимости, средней цены, средней заработной платы и т. д.). Этот анализ выполняется с помощью системы взаимосвязанных индексов: индекса переменного состава, индекса фиксированного состава и индекса влияния структурных сдвигов.

Построение этой системы индексов показано на примере анализа себестоимости *одного вида продукции* А, выпускаемой несколькими предприятиями фирмы.

Изменение себестоимости продукта А по фирме (по группе предприятий) определяется следующим индексом:

$$I_{\bar{z}_{1/0}} = \frac{\bar{z}_1}{\bar{z}_0},$$

где $\bar{z}_1$ и $\bar{z}_0$ – средняя себестоимость единицы продукции по группе предприятий соответственно в отчетном и базисном периодах.

Средняя себестоимость единицы продукции в базисном и отчетном периодах исчисляется по формулам средней арифметической взвешенной:

$$\bar{z}_0 = \frac{\sum z_0 q_0}{\sum q_0}; \quad \bar{z}_1 = \frac{\sum z_1 q_1}{\sum q_1},$$

где z_0 и z_1 – себестоимость единицы продукции каждого предприятия соответственно в базисном и отчетном периодах;

q_0 и q_1 – выпуск продукции в натуральном выражении каждым предприятием соответственно в базисном и отчетном периодах.

Следовательно,

$$I_{\bar{z}_{1/0}} = \frac{\sum z_1 q_1}{\sum q_1} : \frac{\sum z_0 q_0}{\sum q_0}.$$

Этот индекс носит название *индекса переменного состава.* Это объясняется тем, что при исчислении средней себестоимости единицы продукции в отчетном периоде весами служило количество продукции отчетного периода. При определении средней себестоимости единицы продукции базисного периода весами было количество продукции базисного периода, т. е. исчислялись средние с меняющимися (переменными) весами.

Величины $\frac{q_1}{\sum q_1}$ и $\frac{q_0}{\sum q_0}$ отражают распределение продукции по предприятиям, поэтому формула индекса себестоимости переменного состава может быть записана так:

$$I_{\bar{z}_{1/0}} = \frac{\sum z_1 d_1}{\sum z_0 d_0},$$

где d_1 и d_0 – удельный вес каждого предприятия в общем объеме выпуска продукта А соответственно в базисном и отчетном периодах.

$$\Delta_{\bar{z}} = \bar{z}_1 - \bar{z}_0 = \sum z_1 d_1 - \sum z_0 d_0,$$

где $\Delta_{\bar{z}}$ – абсолютное изменение средней себестоимости по группе предприятий.

Величина индекса переменного состава зависит от изменения уровня себестоимости по предприятиям и изменения в распределении физического объема продукции между предприятиями.

Чтобы устранить влияние изменений в структуре весов на показатель изменения уровня себестоимости, рассчитывается отношение средних с одними и теми же весами, т. е. исчисляется *индекс себестоимости фиксированного состава.* Для этого среднюю себестоимость определяют при структуре фактического объема продукции в текущем периоде.

Формула индекса себестоимости фиксированного состава записывается так:

$$I_{\bar{z}}^{z} = \frac{\sum z_1 q_1}{\sum q_1} : \frac{\sum z_0 q_1}{\sum q_1} = \frac{\sum z_1 q_1}{\sum z_0 q_1} = \frac{\sum z_1 d_1}{\sum z_0 d_1}.$$

Полученный индекс себестоимости фиксированного состава отражает изменение уровня средней себестоимости в связи с изменениями значений себестоимости по отдельным предприятиям:

$$\Delta_{\bar{z}}^{z} = \sum z_1 d_1 - \sum z_0 d_1,$$

где $\Delta_{\bar{z}}^{z}$ – абсолютное изменение средней себестоимости по группе предприятий за счет изменения уровня себестоимости по предприятиям.

Индекс влияния структурных сдвигов в объеме продукции определяется по формуле

$$I_d = \frac{\sum z_0 d_1}{\sum z_0 d_0}; \quad \Delta^d_{\bar{z}} = \sum z_0 d_1 - \sum z_0 d_0,$$

где $\Delta^d_{\bar{z}}$ – абсолютное изменение средней себестоимости по группе предприятий за счет структурных сдвигов в объеме выпуска продукции.

Поскольку изменение средней себестоимости в целом по группе предприятий определяется изменением двух факторов, то

$$I_{\bar{z}} = I'_{\bar{z}} \cdot I_d; \; \Delta_{\bar{z}} = \Delta^z_{\bar{z}} + \Delta^d_{\bar{z}} = \bar{z}_1 - \bar{z}_0.$$

Использование индексного метода в анализе взаимосвязи экономических явлений

Индексный метод используется при изучении роли отдельных факторов в динамике какого-либо сложного явления, позволяя определить размер абсолютного и относительного изменения сложного явления за счет каждого фактора в отдельности.

Роль отдельных факторов изменения результативного показателя оценивается путем построения системы взаимосвязанных индексов. В основе приема аналитических индексных расчетов лежит принцип элиминирования изменений величины всех факторов, кроме изучаемого. Предпосылкой такого анализа является возможность представления результативного экономического показателя произведением двух или более определяющих его величину показателей (факторов) или суммой таких произведений.

Предположим, что сложный результативный показатель $A = a \cdot b$, где a и b – показатели-факторы.

Изменение сложного явления может быть представлено индексом:

$$I_A = \frac{A_1}{A_0} = \frac{a_1 \cdot b_1}{a_0 \cdot b_0} = I_a \cdot I_b.$$

Абсолютное изменение явления А под влиянием всех факторов представляет собой разность между числителем и знаменателем индекса:

$$\Delta_A = A_1 - A_0 = a_1 b_1 - a_0 b_0.$$

Для выявления влияния каждого фактора в отдельности индекс сложного показателя разлагают на частные (факторные) индексы, характеризующие роль каждого фактора.

Применяются *два метода разложения общего индекса на частные:*

- метод обособленного изучения факторов;
- метод последовательно-цепной.

Сущность метода обособленного изучения факторов заключается в том, что при выявлении влияния отдельного фактора сложный показатель берется в том виде, какой бы он имел, если бы изменился лишь один данный фактор, а все прочие остались неизменными на уровне базисного периода. Роль каждого фактора определяется по следующим формулам:

$$\text{фактор } a\text{:} \quad I_a = \frac{a_1 b_0}{a_0 b_0};$$

$$\text{фактор } b\text{:} \quad I_b = \frac{a_0 b_1}{a_0 b_0}.$$

Абсолютное изменение результативного показателя за счет каждого фактора получается как разность между числителем и знаменателем индекса:

$$\text{фактор } a\text{: } \Delta^a_A = a_1 b_0 - a_0 b_0;$$

$$\text{фактор } b\text{: } \Delta^b_A = a_0 b_1 - a_0 b_0.$$

Однако необходимо иметь в виду, что факторные индексы при данном методе не разлагают полностью величины абсолютного изменения результативного показателя. Получается некоторый неразложенный остаток, который следует рассматривать как результат совместного действия факторов, т. е.

$$\Delta_A \neq \Delta^a_A + \Delta^b_A.$$

При последовательно-цепном методе используется система взаимосвязанных индексов, требующая правильного расположения факторов при написании модели результативного показателя (например, $А = a \cdot b \cdot c$). На первом месте в модели следует ставить качественный фактор. Увеличение цепи факторов на один фактор (например, $a \cdot b$) каждый раз должно давать показатель, имеющий реальный экономический смысл. При выявлении влияния факторов определяются факторные индексы. При определении влияния первого фактора все остальные факторы сохраняются в числителе и знаменателе на уровне отчетного периода. При построении второго факторного индекса первый фактор сохраняется на уровне базисного периода, третий и все последующие – на уровне отчетного периода; при построении третьего факторного индекса первый и второй факторы сохраняются на уровне базисного периода, четвертый и все последующие – на уровне отчетного периода и т. д.

Предположим, что $А = a \cdot b \cdot c$, при этом обеспечена правильность расположения факторов:

$$I_A = \frac{A_1}{A_0} = \frac{a_1 \cdot b_1 \cdot c_1}{a_0 \cdot b_0 \cdot c_0} = I_a \cdot I_b \cdot I_c.$$

Частные индексы следующие:

для фактора a: $I_a = \dfrac{a_1 \cdot b_1 \cdot c_1}{a_0 \cdot b_1 \cdot c_1}$;

для фактора b: $I_b = \dfrac{a_0 \cdot b_1 \cdot c_1}{a_0 \cdot b_0 \cdot c_1}$;

для фактора c: $I_c = \dfrac{a_0 \cdot b_0 \cdot c_1}{a_0 \cdot b_0 \cdot c_0}$.

Абсолютное изменение результативного показателя за счет каждого фактора:

$$\Delta^a{}_А = (a_1 - a_0) \cdot b_1 \cdot c_1;$$
$$\Delta^b{}_А = a_0 (b_1 - b_0) \cdot c_1; \ \Delta^c{}_А = a_0 \cdot b_0 (c_1 - c_0);$$
$$\Delta_А = А_1 - А_0 = \Delta^a{}_А + \Delta^b{}_А + \Delta^c{}_А.$$

Абсолютное изменение сложного экономического показателя за счет каждого фактора может быть определено и в том случае, если этот показатель представляет собой сумму произведений, определяющих его величину показателей. К числу таких показателей относятся: общая стоимость всей выработанной (или реализованной) продукции, общая сумма затрат на ее производство, общая сумма затрат труда на производство всей продукции.

Агрегатный индекс общей стоимости продукции (I_{qp}) равен произведению агрегатного индекса физического объема продукции (I_q) и агрегатного индекса цен (I_p), если коэффициенты соизмерения в индексе физического объема взяты из базисного периода, а веса в индексе цен – из текущего периода. В общем виде это равенство должно быть записано так:

$$I_{qp_{1/0}} = I_{q_{1/0}} \cdot I_{p_{1/0}} = \frac{\sum q_1 p_0}{\sum q_0 p_0} \cdot \frac{\sum p_1 q_1}{\sum p_0 q_1} = \frac{\sum q_1 p_1}{\sum q_0 p_0}.$$

Общее абсолютное изменение общей стоимости продукции за счет двух факторов составляет:

$$\Delta^{qp}{}_{\sum qp} = \sum q_1 p_1 - \sum q_0 p_0.$$

Абсолютное изменение общей стоимости продукции за счет отдельных факторов:

а) изменения физического объема продукции

$$\Delta^q{}_{\sum qp} = \sum q_1 p_0 - \sum q_0 p_0.$$

б) среднего изменения цен на продукцию

$$\Delta^p{}_{\sum qp} = \sum p_1 q_1 - \sum p_0 q_1.$$

Общее абсолютное изменение результативного показателя составит алгебраическую сумму абсолютных изменений за счет отдельных факторов, т. е.

$$\Delta^{qp}_{\sum qp} = \Delta^q_{\sum qp} + \Delta^p_{\sum qp} = \sum q_1 p_1 - \sum q_0 p_0.$$

Доля каждого фактора в общем абсолютном изменении результативного показателя определяется следующим образом:

а) физического объема продукции – $\frac{\Delta^q_{\sum qp}}{\Delta^{qp}_{\sum qp}}$;

б) среднего изменения цен на продукцию – $\frac{\Delta^p_{\sum qp}}{\Delta^{qp}_{\sum qp}}$.

Агрегатный индекс общих затрат (I_{qz}) представляет собой произведение агрегатного индекса физического объема продукции (I_q) и агрегатного индекса себестоимости продукции (I_z) при условии, что коэффициенты соизмерения в индексе объема и веса в индексе себестоимости взяты из разных периодов:

$$I_{qz_{1/0}} = I_{q_{1/0}} \cdot I_{z_{1/0}} = \frac{\sum q_1 z_0}{\sum q_0 z_0} \cdot \frac{\sum z_1 q_1}{\sum z_0 q_1} = \frac{\sum q_1 z_1}{\sum q_0 z_0}.$$

Абсолютное изменение общих затрат на выпуск продукции за счет отдельных факторов:

а) изменения физического объема продукции

$$\Delta^q_{\sum qz} = \sum q_1 z_0 - \sum q_0 z_0;$$

б) среднего изменения себестоимости единицы продукции

$$\Delta^z_{\sum qz} = \sum z_1 q_1 - \sum z_0 q_1.$$

Общее абсолютное изменение общих затрат составит:

$$\Delta^{qz}_{\sum qz} = \Delta^q_{\sum qz} + \Delta^z_{\sum qz} = \sum z_1 q_1 - \sum z_0 q_0.$$

Агрегатный индекс общих затрат рабочего времени на выпуск продукции (J_{qt}) равен произведению агрегатного индекса физического объема продукции (J_q) и агрегатного индекса затрат рабочего времени на единицу продукции (J_t), при условии, если коэффициенты в первом индексе взяты из базисного периода, а веса во втором индексе – из текущего периода.

$$J_{qt} = J_q \cdot J_t = \frac{\sum q_1 t_0}{\sum q_0 t_0} \cdot \frac{\sum t_1 q_1}{\sum t_0 q_1} = \frac{\sum q_1 t_1}{\sum q_0 t_0}.$$

Выявления влияния отдельных факторов на абсолютное изменение общего объема затрат рабочего времени выполняется аналогично двум предыдущим результативным показателям (общей стоимости продукции и общих затрат на производство продукции).

7.1. Решение типовых задач

7.1. Выпуск продукции по заводу почвообрабатывающих машин за два квартала следующий (табл. 7.1).

Таблица 7.1

Вид продукции	Выпуск, шт.		Отпускная цена за шт., тыс. руб.	
	I кв., q_0	II кв., q_1	I кв., p_0	II кв., p_1
Плуги навесные	2500	2610	4,8	5,4
Плуги прицепные	3000	2950	7,1	7,6
Культиваторы навесные	3600	3700	5,0	5,7

Определить:

1) изменение (в %) выпуска каждого вида продукции, а также изменение выпуска продукции в целом по предприятию;

2) изменение цен (в %) по каждому виду продукции и среднее изменение цен по всему ассортименту продукции;

3) абсолютное изменение общей стоимости продукции, выделив из общей суммы изменение за счет изменения количества продукции и за счет изменения цен.

Решение

1. Для характеристики изменения выпуска каждого вида продукции исчисляются индивидуальные индексы физического объема продукции:

$$i_{q_{1/0}} = \frac{q_1}{q_0}$$

(условные обозначения представлены в табл. 7.1).

Плуги навесные $i_{q_{1/0}} = \frac{2610}{2500} = 1{,}044$, или 104,4%, т. е. выпуск увеличился на 4,4% (104,4 – 100).

Плуги прицепные $i_{q_{1/0}} = \frac{2950}{3000} = 0{,}983$, или 98,3%, следовательно, выпуск снизился на 1,7% (98,3 – 100).

Культиваторы навесные $i_{q_{1/0}} = \frac{3700}{3600} = 1{,}028$, или 102,8%, т. е. увеличение выпуска на 2,8%.

Для характеристики изменения выпуска продукции в целом по предприятию исчисляется агрегатный индекс физического объема продукции:

$$I_{q_{1/0}} = \frac{\sum q_1 p_0}{\sum q_0 p_0} = \frac{2610 \cdot 4{,}8 + 2950 \cdot 7{,}1 + 3700 \cdot 5{,}0}{2500 \cdot 4{,}8 + 3000 \cdot 7{,}1 + 3600 \cdot 5{,}0} = \frac{51\,973}{51\,300} = 1{,}0131,$$

или 101,31%, т. е. в целом по предприятию выпуск продукции увеличился на 1,31%, в результате стоимость продукции увеличилась на 673 тыс. руб. (51 973 – 51 300).

2. Для характеристики изменения цен по каждому виду продукции используются индивидуальные индексы цен:

$$i_{p_{1/0}} = \frac{p_1}{p_0}.$$

Плуги навесные $i_{p_{1/0}} = \frac{5{,}4}{4{,}8} = 1{,}125$, или 112,5%, следовательно, цена повысилась на 12,5% (112,5 – 100).

Плуги прицепные $i_{p_{1/0}} = \frac{7{,}6}{7{,}1} = 1{,}070$, или 107,0%, т. е. цена возросла на 7,0%.

Культиваторы навесные $i_{p_{1/0}} = \frac{5{,}7}{5{,}0} = 1{,}14$, или 114,0%, т. е. увеличение цены на 14,0%.

Среднее изменение цен по всему ассортименту продукции определяется по формуле агрегатного индекса цен:

$$I_{p_{1/0}} = \frac{\sum p_1 q_1}{\sum p_0 q_1} = \frac{5{,}4 \cdot 2\,610 + 7{,}6 \cdot 2\,950 + 5{,}7 \cdot 3\,700}{4{,}8 \cdot 2\,610 + 7{,}1 \cdot 2\,950 + 5{,}0 \cdot 3\,700} = \frac{57\,604}{51\,973} = 1{,}1083,$$

или 110,83%.

Таким образом, цены на продукцию предприятия повышены в среднем на 10,83%, за счет чего стоимость продукции повысилась на 5631 тыс. руб. (57 604 – 51 973).

3. Абсолютное изменение стоимости продукции определяется по формуле

$$\Delta^{qp}_{\Sigma qp} = \Sigma q_1 p_1 - \Sigma q_0 p_0 = 57\,604 - 51\,300 = 6304 \text{ тыс. руб.},$$

где $\Delta^{q}_{\Sigma qp}$ – абсолютное изменение стоимости продукции за счет изменения выпуска продукции; $\Delta^{q}_{\Sigma qp} = 673$ тыс. руб. (см. решение в п. 1);

$\Delta^{p}_{\Sigma qp}$ – абсолютное изменение стоимости продукции за счет изменения цен; $\Delta^{p}_{\Sigma qp} = 5631$ тыс. руб. (см. решение в п. 2);

$\Delta^{qp}_{\Sigma qp} = \Delta^{q}_{\Sigma qp} + \Delta^{p}_{\Sigma qp} = 673 + 5631 = 6304$, что и соответствует ранее полученной цифре.

Доля каждого фактора в общем абсолютном размере изменения результативного показателя следующая:

а) физического объема продукции – $\frac{\Delta^{q}_{\Sigma qp}}{\Delta^{qp}_{\Sigma qp}} = \frac{673}{6304} = 0{,}1068$, или 10,68%.

б) среднего изменения цен – $\frac{\Delta^{p}_{\Sigma qp}}{\Delta^{qp}_{\Sigma qp}} = \frac{5631}{6304} = 0{,}8932$, или 89,32%.

7.2. По предприятию имеются следующие данные о реализации продукции (табл. 7.2).

Таблица 7.2

Вид продукции	Единицы измерения	Реализовано		Общая стоимость реализованной продукции, тыс. руб.	
		сентябрь, q_0	октябрь, q_1	сентябрь, q_0p_0	октябрь, q_1p_1
Цемент М-400 портланд	т	18 200	19 500	8 918	9 594
Кирпич красный М-100	тыс. шт.	3 400	4 000	2 958	3 520

Определить:

1) среднее изменение цен на реализованную продукцию и абсолютное изменение стоимости реализованной продукции за счет изменения цен;

2) общее изменение физического объема реализованной продукции предприятия и абсолютное изменение стоимости реализованной продукции за счет изменения ее физического объема.

Решение

1. Среднее изменение цен на реализованную продукцию определяется по формуле агрегатного индекса цен:

$$I_{p_{1/0}} = \frac{\sum p_1 q_1}{\sum p_0 q_1},$$

где $\sum p_1q_1$ – стоимость реализованной продукции за октябрь;

$$\sum p_1q_1 = 9\,594 + 3\,520 = 13\,114 \text{ тыс. руб.}$$

(условные обозначения проставлены в табл. 7.2).

Для расчета $\sum p_0q_1$ необходимо по первичной информации исчислить цены за единицу продукции в сентябре:

цемент $p_0 = \frac{p_0 q_0}{q_0} = \frac{8\,918\,000}{18\,200} = 490$ руб.

кирпич $p_0 = \frac{p_0 q_0}{q_0} = \frac{2\,958\,000}{3\,400} = 870$ руб.

$$\sum p_0q_1 = 490 \cdot 19\,500 + 870 \cdot 4\,000 = 9\,555\,000 + 3\,480\,000 = \\ = 13\,035\,000 \text{ руб.} = 13\,035 \text{ тыс. руб.}$$

$$I_{p_{1/0}} = \frac{13\,114}{13\,035} = 1{,}0060, \text{ или } 100{,}60\%.$$

Следовательно, цены на продукцию предприятия повышены в среднем на 0,60% (100,60 – 100).

Изменение стоимости реализованной продукции за счет изменения цен на продукцию:

$$\Delta^p_{\Sigma qp} = \sum p_1q_1 - \sum p_0q_1 = 13\,114 - 13\,035 = 79 \text{ тыс. руб.},$$

т. е. увеличение на 79 тыс. руб.

2. Для характеристики изменения физического объема реализованной продукции предприятия исчисляется агрегатный индекс физического объема продукции:

$$I_{q_{1/0}} = \frac{\sum q_1 p_0}{\sum q_0 p_0} = \frac{13\,035}{8\,918 + 2\,958} = \frac{13\,035}{11\,876} = 1{,}098, \text{ или } 109{,}8\%,$$

т. е. в целом по предприятию физический объем реализованной продукции увеличился на 9,8% (109,8 – 100).

Изменение стоимости реализованной продукции за счет изменения объема реализованной продукции:

$$\Delta^q_{\Sigma qp} = \sum q_1p_0 - \sum q_0p_0 = 13\,035 - 11\,876 = 1\,159 \text{ тыс. руб.},$$

т. е. имеется увеличение на 1159 тыс. руб.

Проверка. Общий прирост стоимости реализованной продукции составил, тыс. руб.:

$$\Delta^{qp}_{\Sigma qp} = \sum q_1p_1 - \sum q_0p_0 = 13\,114 - 11\,876 = 1\,238;$$

$$\Delta^{qp}_{\Sigma qp} = \Delta^p_{\Sigma qp} + \Delta^q_{\Sigma qp} = 79 + 1\,159 = 1\,238,$$

что и соответствует ранее полученному результату.

7.3. Объем товарной продукции авторемонтного завода (в действующих ценах) составил, тыс. руб.: апрель – 12 000, май – 14 400, июнь – 24 000.

Отпускные цены на продукцию завода снижены в среднем в мае по сравнению с апрелем на 0,6%, а в июне повышены на 5,0% по сравнению с маем.

Определить изменение физического объема продукции.

Решение

Условие задачи представлено в табличной форме (табл. 7.3).

Таблица 7.3

Месяц	Стоимость товарной продукции, тыс. руб.	Цепные агрегатные индексы цен
Апрель	$\sum q_0p_0 = 12\ 000$	
Май	$\sum q_1p_1 = 14\ 400$	$I_{p_{1/0}} = 0{,}994\left(\frac{100-0{,}6}{100}\right)$
Июнь	$\sum q_2p_2 = 24\ 000$	$I_{p_{2/1}} = 1{,}05\left(\frac{100+5{,}0}{100}\right)$

На основе имеющихся данных следует сначала определить цепные индексы общей стоимости продукции, учитывая, что по условию задачи известны цепные агрегатные индексы цен.

$$I_{qp_{1/0}} = \frac{\sum q_1 p_1}{\sum q_0 p_0} = \frac{14\ 400}{12\ 000} = 1{,}20;$$

$$I_{qp_{2/1}} = \frac{\sum q_2 p_2}{\sum q_1 p_1} = \frac{24\ 000}{14\ 000} = 1{,}71.$$

Для определения агрегатных индексов физического объема продукции используется взаимосвязь трех индексов: агрегатный индекс общей стоимости продукции равен произведению агрегатного индекса физического объема продукции и агрегатного индекса цен:

$$I_{qp_{1/0}} = I_{q_{1/0}} \cdot I_{p_{1/0}}; \quad I_{qp_{2/1}} = I_{q_{2/1}} \cdot I_{p_{2/1}}.$$

Отсюда

$$I_{q_{1/0}} = I_{qp_{1/0}} : I_{p_{1/0}} = 1{,}2 : 0{,}994 = 1{,}207, \text{ или } 120{,}7\%.$$

Следовательно, в мае по сравнению с апрелем физический объем продукции предприятия увеличился на 20,7% (120,7 – 100).

$$I_{q_{2/1}} = I_{qp_{2/1}} : I_{p_{2/1}} = 1{,}71 : 1{,}05 = 1{,}629, \text{ или } 162{,}9\%,$$

т. е. в июне по сравнению с маем физический объем продукции возрос на 62,9%.

7.4. По металлургическому комбинату имеются следующие данные о выпуске продукции (табл. 7.4).

Таблица 7.4

Вид продукции	I квартал		II квартал		III квартал	
	выпуск (т), q_0	отпускная цена за 1 т (руб.), p_0	выпуск (т), q_1	отпускная цена за 1 т (руб.), p_1	выпуск (т), q_2	отпускная цена за 1 т (руб.), p_2
Прокат листовой	5 000	1 900	5 100	1 900	5 400	2 090
Сталь арматурная	4 500	1 650	4 500	1 680	4 700	1 700
Швеллер	800	1900	1 000	1 910	1 100	1 940

Определить агрегатные цепные и базисные индексы физического объема продукции, цен и общей стоимости продукции. Показать взаимосвязь вычисленных индексов.

Сформулировать вывод.

Решение

1. Агрегатные индексы физического объема характеризуют изменение выпуска в целом по предприятию.

Цепные агрегатные индексы физического объема продукции (условные обозначения проставлены в табл. 7.4):

$$I_{q_{1/0}} = \frac{\sum q_1 p_0}{\sum q_0 p_0} =$$

$$= \frac{5100 \cdot 1900 + 4500 \cdot 1650 + 1000 \cdot 1900}{5000 \cdot 1900 + 4500 \cdot 1650 + 800 \cdot 1900} = \frac{19\,015\,000}{18\,445\,000} = 1{,}031.$$

Следовательно, во II квартале по сравнению с I кварталом физический объем продукции по предприятию увеличился на 3,1%.

$$I_{q_{2/1}} = \frac{\sum q_2 p_1}{\sum q_1 p_1} = \frac{5400 \cdot 1900 + 4700 \cdot 1680 + 1100 \cdot 1910}{5100 \cdot 1900 + 4500 \cdot 1680 + 1000 \cdot 1910} =$$

$$= \frac{20\,257\,000}{19\,160\,000} = 1{,}057, \text{ или } 105{,}7\%.$$

В III квартале по сравнению со II кварталом физический объем продукции возрос на 5,7%.

Базисные агрегатные индексы физического объема продукции:

$$I_{q_{1/0}} = \frac{\sum q_1 p_0}{\sum q_0 p_0} = 1{,}031$$ (соответствует первому цепному индексу);

$$I_{q_{2/0}} = \frac{\sum q_2 p_0}{\sum q_0 p_0} = \frac{5400 \cdot 1900 + 4700 \cdot 1650 + 1100 \cdot 1900}{5000 \cdot 1900 + 4500 \cdot 1650 + 800 \cdot 1900} =$$

$$= \frac{20\,105\,000}{18\,445\,000} = 1{,}090, \text{ или } 109{,}0\%.$$

В III квартале физический объем продукции предприятия возрос по сравнению с I кварталом на 9%.

2. Агрегатные индексы цен характеризуют среднее изменение цен по всему ассортименту продукции.

Цепные агрегатные индексы цен:

$$I_{p_{1/0}} = \frac{\sum p_1 q_1}{\sum p_0 q_1} = \frac{19\,160\,000}{19\,015\,000} = 1{,}0076, \text{ или } 100{,}76\%.$$

Следовательно, цены на продукцию возросли во II квартале против I квартала в среднем на 0,76%:

$$I_{p_{2/1}} = \frac{\sum p_2 q_2}{\sum p_1 q_2} = \frac{5400 \cdot 2090 + 4700 \cdot 1700 + 1100 \cdot 1940}{20\,257\,000} =$$

$$= \frac{21\,410\,000}{20\,257\,000} = 1{,}057, \text{ или } 105{,}7\%.$$

Цены на продукцию предприятия увеличились в среднем в III квартале по сравнению со II кварталом на 5,7%.

Базисные агрегатные индексы цен:

$$I_{p_{1/0}} = \frac{\sum p_1 q_1}{\sum p_0 q_1} = 1{,}0076$$ (соответствует первому цепному индексу);

$$I_{p_{2/0}} = \frac{\sum p_2 q_2}{\sum p_0 q_2} = \frac{21\,410\,000}{20\,105\,000} = 1{,}065, \text{ или } 106{,}5\%.$$

В III квартале по сравнению с I кварталом цены выросли в среднем на 6,5%.

3. Агрегатные индексы общей стоимости продукции отражают ее изменение по всему составу продукции.

Цепные агрегатные индексы общей стоимости продукции:

$$I_{qp_{1/0}} = \frac{\sum p_1 q_1}{\sum p_0 q_0} = \frac{19\,160\,000}{18\,445\,000} = 1{,}039, \text{ или } 103{,}9\%,$$

т. е. увеличение стоимости на 3,9%.

$$I_{qp_{2/1}} = \frac{\sum p_2 q_2}{\sum p_1 q_1} = \frac{21\,410\,000}{19\,160\,000} = 1{,}117, \text{ или } 111{,}7\%.$$

Базисные агрегатные индексы общей стоимости продукции:

$$I_{qp_{1/0}} = \frac{\sum p_1 q_1}{\sum p_0 q_0} = 1{,}039;$$

$$I_{qp_{2/0}} = \frac{\sum p_2 q_2}{\sum p_1 q_1} = \frac{21\,410\,000}{18\,445\,000} = 1{,}161, \text{ или } 116{,}1\%,$$

т. е. увеличение на 16,1%.

Таким образом, общая стоимость продукции предприятия во II квартале увеличилась по сравнению с I кварталом на 3,9%, а в III квартале по сравнению с I кварталом – на 16,1%.

4. Взаимосвязь вычисленных индексов.

Изменение общей стоимости продукции происходит за счет двух факторов: изменения объема продукции и изменения цен. Агрегатный индекс общей стоимости продукции равен произведению агрегатного индекса физического объема продукции на агрегатный индекс цен, если коэффициенты соизмерения в первом индексе взяты из базисного периода, а веса во втором индексе – из текущего периода.

$$I_{qp_{1/0}} = I_{q_{1/0}} \cdot I_{p_{1/0}} = 1{,}031 \cdot 1{,}076 = 1{,}039,$$

что и соответствует полученной ранее цифре.

Абсолютное изменение общей стоимости продукции:

$$\Delta^{qp}{}_{\Sigma qp} = \Sigma p_1 q_1 - \Sigma p_0 q_0 = \Delta^{q}{}_{\Sigma qp} + \Delta^{p}{}_{\Sigma qp},$$

где $\Delta^{q}{}_{\Sigma qp}$ – абсолютное изменение общей стоимости продукции за счет изменения выпуска продукции;

$\Delta^{p}{}_{\Sigma qp}$ – абсолютное изменение общей стоимости продукции за счет изменения цен на продукцию.

$$\Delta^{q}{}_{\Sigma qp} = \Sigma q_1 p_0 - \Sigma q_0 p_0 = 19\,015\,000 - 18\,445\,000 = 570\,000 \text{ руб.};$$

$$\Delta^{p}{}_{\Sigma qp} = \Sigma p_1 q_1 - \Sigma p_0 q_1 = 19\,160\,000 - 19\,015\,000 = 145\,000 \text{ руб.};$$

$$\Delta^{qp}{}_{\Sigma qp} = 19\,160\,000 - 18\,445\,000 = 715\,000 \text{ руб.};$$

$$\Delta^{qp}{}_{\Sigma qp} = 570\,000 + 145\,000 = 715\,000 \text{ руб.};$$

$$I_{qp_{2/1}} = I_{q_{2/1}} \cdot I_{p_{2/1}} = 1{,}057 \cdot 1{,}057 = 1{,}117,$$

что соответствует ранее полученной величине.

$$\Delta^{qp}{}_{\Sigma qp_{2/1}} = \Sigma p_2 q_2 - \Sigma p_1 q_1 = 21\,410\,000 - 19\,160\,000 = 2\,250\,000 \text{ руб.};$$

$$\Delta^{q}{}_{\Sigma qp} = \Sigma q_2 p_1 - \Sigma q_1 p_1 = 20\,257\,000 - 19\,160\,000 = 1\,097\,000 \text{ руб.};$$

$$\Delta^{p}{}_{\Sigma qp} = \Sigma p_2 q_2 - \Sigma p_1 q_2 = 21\,410\,000 - 20\,257\,000 = 1\,153\,000 \text{ руб.};$$

$$\Delta^{q}{}_{\Sigma qp} + \Delta^{p}{}_{\Sigma qp} = 1\,097\,000 + 1\,153\,000 = 2\,250\,000 \text{ руб.}$$

7.5. Имеются следующие данные о выпуске кирпича (красный кирпич М-100) тремя предприятиями фирмы (табл. 7.5).

Таблица 7.5

№ предприятия	Выпуск, тыс. шт.		Себестоимость 1000 шт., руб.	
	сентябрь, q_0	октябрь, q_1	сентябрь, z_0	октябрь, z_1
1	3000	3500	610	608
2	6000	7700	590	580
3	3000	2800	630	628

Определить:

1) изменение себестоимости по каждому предприятию фирмы;

2) изменение себестоимости по фирме, выяснив, за счет действия каких факторов это произошло.

Решение

1. Изменение себестоимости по каждому предприятию определяется с помощью индивидуального индекса себестоимости:

$i_{z_{1/0}} = \frac{z_1}{z_0}$ (условные обозначения записаны в табл. 7.5).

Предприятие № 1 – $i_{z_{1/0}} = \frac{608}{610} = 0{,}996$, или 99,6%,

т. е. себестоимость продукции снизилась на 0,4% (100 – 99,6).

Предприятие № 2 – $i_{z_{1/0}} = \frac{580}{590} = 0{,}983$, или 98,3%,

т. е. снижение себестоимости на 1,7%.

Предприятие № 3 – $i_{z_{1/0}} = \frac{628}{630} = 0{,}997$, или 99,7%,

т. е. себестоимость снизилась на 0,3%.

2. Изменение себестоимости продукта по фирме определяется с помощью индекса переменного состава:

$$I_{\bar{z}_{1/0}} = \frac{\bar{z}_1}{\bar{z}_0},$$

где $\bar{z}_1$ и $\bar{z}_0$ – средняя себестоимость продукта по фирме соответственно в текущем и базисном периодах.

$$\bar{z}_0 = \frac{610 \cdot 3\ 000 + 590 \cdot 6\ 000 + 630 \cdot 3\ 000}{3\ 000 + 6\ 000 + 3\ 000} = \frac{7\ 260\ 000}{12\ 000} = 605{,}0 \text{ руб.};$$

$$\bar{z}_1 = \frac{608 \cdot 3\ 500 + 580 \cdot 7\ 700 + 628 \cdot 2\ 800}{3\ 500 + 7\ 700 + 2\ 800} = \frac{8\ 352\ 400}{14\ 000} = 596{,}6 \text{ руб.}$$

$$I_{\bar{z}_{1/0}} = \frac{596{,}6}{605{,}0} = 0{,}986, \text{ или } 98{,}6\%.$$

Следовательно, в целом по фирме себестоимость кирпича снизилась на 1,4% (98,6 – 100).

Абсолютное изменение (снижение) себестоимости по фирме составило:

$$\Delta_{\bar{z}} = \bar{z}_1 - \bar{z}_0 = 596{,}6 - 605{,}0 = -8{,}4 \text{ руб.}$$

Себестоимость продукта по фирме изменилась за счет двух факторов:

а) изменения уровня себестоимости по каждому предприятию;

б) изменения удельного веса предприятий в общем объеме выпуска продукции (структурный фактор).

Для выявления влияния первого фактора (изменения уровня себестоимости по предприятиям) определяется индекс себестоимости фиксированного (постоянного) состава:

$$I'_{\bar{z}_{1/0}} = \frac{\sum z_1 d_1}{\sum z_0 d_1},$$

где d_1 – удельный вес предприятия в общем объеме выпуска продукции в текущем периоде.

Предприятие № 1 – $d_1 = \frac{3\,500}{3\,500 + 7\,700 + 2\,800} = \frac{3\,500}{14\,000} = 0{,}25.$

Предприятие № 2 – $d_1 = \frac{7\,700}{14\,000} = 0{,}55.$

Предприятие № 3 – $d_1 = \frac{2\,800}{14\,000} = 0{,}20.$

$$I'_{\bar{z}_{1/0}} = \frac{608 \cdot 0{,}25 + 580 \cdot 0{,}55 + 628 \cdot 0{,}20}{610 \cdot 0{,}25 + 590 \cdot 0{,}55 + 630 \cdot 0{,}20} = \frac{596{,}6}{603{,}0} = 0{,}989, \ 98{,}9\%.$$

Следовательно, за счет изменения себестоимости продукции по каждому предприятию себестоимость по фирме снизилась на 1,1%, что в абсолютном размере составило:

$$\Delta^{z}_{\bar{z}} = \sum z_1 d_1 - \sum z_0 d_1 = 596{,}6 - 603{,}0 = -6{,}4 \text{ руб.}$$

Для выявления влияния второго фактора (изменения в распределении продукции по предприятиям) определяется индекс влияния структурных сдвигов в объеме продукции:

$$I_{d_{1/0}} = \frac{\sum z_0 d_1}{\sum z_0 d_0},$$

где d_0 – удельный вес предприятия в общем объеме выпуска продукции в базисном периоде.

Предприятие № 1 – $d_0 = \frac{3\ 000}{12\,000} = 0{,}25.$

Предприятие № 2 – $d_0 = \frac{6\,000}{12\,000} = 0{,}5.$

Предприятие № 3 – $d_0 = \frac{3\,000}{12\,000} = 0{,}25.$

$$I_{d_{1/0}} = \frac{610 \cdot 0{,}25 + 590 \cdot 0{,}55 + 630 \cdot 0{,}20}{610 \cdot 0{,}25 + 590 \cdot 0{,}50 + 630 \cdot 0{,}25} = \frac{603{,}0}{605{,}0} = 0{,}997, \text{ или } 99{,}7\%.$$

За счет изменений в распределении продукции по предприятиям себестоимость продукта по фирме снизилась на 0,3%, что в абсолютном размере составило:

$$\Delta^{d}_{\bar{z}} = \sum z_0 d_1 - \sum z_0 d_0 = 603{,}0 - 605{,}0 = -2{,}0 \text{ руб.}$$

Совместное влияние двух факторов обеспечило следующее изменение себестоимости по фирме:

$$I_{\bar{z}_{1/0}} = I'_{\bar{z}_{1/0}} \cdot I_{d_{1/0}} = 0{,}989 \cdot 0{,}997 = 0{,}986,$$

что соответствует ранее полученной величине.

Абсолютное изменение себестоимости по фирме за счет двух факторов:

$$\Delta_{\bar{z}} = \Delta^{z}_{\bar{z}} + \Delta^{d}_{\bar{z}} = (-6{,}4) + (-2{,}0) = -8{,}4 \text{ руб.}$$

7.6. Количество произведенной продукции в отчетном периоде по сравнению с базисным увеличилось на 8,0%, а общая стоимость продукции уменьшилась на 5,0%.

Определить, как изменились в среднем отпускные цены на продукцию.

Решение

Для определения индекса цен используется взаимосвязь между тремя индексами:

$$I_{qp_{1/0}} = I_{q_{1/0}} \cdot I_{p_{1/0}}.$$

Отсюда

$$I_{p_{1/0}} = I_{qp_{1/0}} : I_{q_{1/0}} = 1{,}05 : 1{,}08 = 0{,}972, \text{ или } 97{,}2\%.$$

Следовательно, отпускные цены на продукцию снижены в среднем на 2,8%.

7.7. По заводу имеются следующие данные о выпуске продукции (табл. 7.6).

Определить, на сколько процентов увеличился выпуск продукции по предприятию.

Таблица 7.6

Вид продукции	Выпуск продукции в I квартале, тыс. руб.	Увеличение (+) или уменьшение (–) выпуска продукции во II квартале по сравнению с I кварталом, %
Рельсы трамвайные	22 300	+3,0
Чугун литейный	15 800	–2,0
Железо листовое	10 500	+1,5

Решение

Для определения изменения физического объема продукции в целом по предприятию используется формула среднего взвешенного арифметического индекса, так как по условию задачи известны индивидуальные индексы физического объема:

$$I_{q_{1/0}} = \frac{\sum i_q q_0 p_0}{\sum q_0 p_0}.$$

Индивидуальные индексы по видам продукции следующие:

рельсы трамвайные – $i_q = \frac{100+3}{100} = 1{,}03;$

чугун литейный – $i_q = \frac{100-2}{100} = 0{,}98;$

железо листовое – $i_q = \frac{100+1{,}5}{100} = 1{,}015.$

$$I_{q_{1/0}} = \frac{1{,}03 \cdot 22\,300 + 0{,}98 \cdot 15\,800 + 1{,}015 \cdot 10\,500}{22\,300 + 15\,800 + 10\,500} = \frac{49110{,}5}{48600{,}0} = 1{,}0105,$$

или 101,05%.

Следовательно, физический объем продукции в целом по предприятию увеличился на 1,05%.

7.8. По предприятию за два месяца имеются следующие данные о выпуске продукции и затратах на нее (табл. 7.7).

Таблица 7.7

Показатель	Январь	Февраль
Прокат листовой, тыс. т	45	50
Общая сумма затрат на выпуск, тыс. руб.	65 835	73 500

Определить абсолютное изменение общей суммы затрат предприятия за счет изменения выпуска продукции и ее себестоимости.

Решение

Для выявления влияния двух факторов на изменение общей суммы затрат используется последовательно-цепной индексный метод.

Себестоимость продукции (z) – фактор качественный; выпуск продукции (q) – фактор объемный.

Изменение общей суммы затрат (А) можно представить следующим индексом:

$$I_A = \frac{A_1}{A_0} = \frac{z_1 q_1}{z_0 q_0}.$$

По условию задачи $z_0 = \frac{65\,835\,000}{45\,000} = 1463$ руб.

$$z_1 = \frac{73\,500\,000}{50\,000} = 1470 \text{ руб}$$

Влияние факторов следующее:

а) себестоимости: $I_z = \frac{z_1 q_1}{z_0 q_1} = \frac{1470 \cdot 50\,000}{1463 \cdot 50\,000} = 1{,}005.$

Следовательно, повышение себестоимости продукции на 0,5% обеспечило увеличение общей суммы затрат на 350 тыс. руб. $[\Delta^z{}_A = (z_1 - z_0) \cdot q_1 = (1\,470 - 1\,463) \cdot 50\,000]$;

б) количества продукции: $I_q = \frac{z_0 q_1}{z_0 q_0} = \frac{1\,463 \cdot 50\,000}{1\,463 \cdot 45\,000} = 1{,}111.$

Рост количества выпущенной продукции на 11,1% повлек за собой увеличение суммы затрат на 7 315 тыс. руб. $[\Delta^q{}_A = (q_1 - q_0) \cdot z_0 = (50\,000 - 45\,000)\ 1463]$.

За счет двух факторов общая сумма затрат увеличилась на 7665 тыс. руб. (350 + 7315), что и соответствует цифре фактического изменения (73 500 – 65 835).

7.9. Производительность труда в отчетном году по сравнению с прошлым годом возросла на 12% и составила 336 тыс. руб. на одного работающего. За этот же период численность работающих сократилась на 20 человек и составила 380 человек.

Определить индексы численности работающих, физического объема продукции и абсолютный прирост физического объема продукции за счет роста производительности труда и изменения численности работающих.

Решение

Индекс численности работающих определяется по формуле

$$I_N = \frac{N_1}{N_0} = \frac{380}{380 + 20} = 0{,}95.$$

Используя взаимосвязь индексов физического объема продукции (I_q), производительности труда ($I_{ПТ}$) и численности (I_N), находим:

$$I_q = I_{ПТ} \cdot I_N.$$

По условию задачи $I_{ПТ} = 1{,}12$. Тогда $I_q = 1{,}12 \cdot 0{,}95 = 1{,}064$.

Прирост физического объема продукции находим как разность числителя и знаменателя соответствующего индекса:

$$I_q = \frac{ПТ_1 \cdot N_1}{ПТ_0 \cdot N_0}; \quad \Delta_q^{ПТ,\,N} = ПТ_1 \cdot N_1 - ПТ_0 \cdot N_0.$$

По условию $ПТ_1 = 336$ тыс. руб.;

$$I_{ПТ} = \frac{ПТ_1}{ПТ_0};\ ПТ_0 = \frac{ПТ_1}{I_{ПТ}};$$

Тогда общий прирост продукции составит:

$$\Delta_q^{ПТ,\ N} = ПТ_1 \cdot N_1 - \frac{ПТ_1}{I_{ПТ}} \cdot N_0 =$$

$$= 336 \cdot 380 - \frac{336}{1{,}12} \cdot 400 = 127\ 680 - 120\ 000 = 7680 \text{ тыс. руб.}$$

Прирост объема продукции за счет:

а) роста производительности труда:

$$\Delta_q^{ПТ} = (ПТ_1 \cdot ПТ_0) \cdot N_1 = (336 - \frac{336}{1.12}) \cdot 380 = 13\ 680 \text{ тыс. руб.};$$

б) уменьшения численности работающих:

$$\Delta^N{}_q = (N_1 - N_0) \cdot ПТ_0 =$$

$$= (380 - 400) \cdot 300 = -6\ 000 \text{ тыс. руб.}$$

Общий прирост физического объема продукции равен сумме приростов, полученных за счет каждого фактора:

$$\Delta^{ПТ,\ N}{}_q = \Delta^{ПТ}{}_q + \Delta^N{}_q =$$

$$= 13\ 680 - 6000 = 7680 \text{ тыс. руб.}$$

7.10. В отчетном году было реализовано товара А на 300 млн руб., товара Б – на 5 млрд руб., товара В – на 412 млн руб., товара Г – на 143 млн руб.

Исчислить общий индекс цен на все товары, если известно, что цены на товар А были снижены на 4%, на товар Б остались без изменения, а на товары В и Г повысились на 3 и 10% соответственно.

Решение

Для характеристики изменения цен на все виды товара вместе исчисляется общий индекс цен по формуле средневзвешенного

гармонического индекса, поскольку известны индивидуальные индексы цен по каждому виду товара:

$$I_{p_{1/0}} = \frac{\sum p_1 q_1}{\sum \frac{p_1 q_1}{i_p}} = \frac{300 + 5\,000 + 412 + 143}{\frac{300}{0{,}96} + \frac{5\,000}{1{,}0} + \frac{412}{1{,}03} + \frac{143}{1{,}1}} = \frac{5855}{5842{,}5} = 1{,}002,$$

т. е. по всему ассортименту товаров цены в отчетном году повысились в среднем на 0,2%.

7.11. По предприятию за два года имеются следующие данные (табл. 7.8).

Таблица 7.8

Показатель	Прошлый год	Отчетный год
Объем произведенной продукции в сопоставимых ценах, млн руб.	1176	1530
Среднесписочная численность персонала основной деятельности	1200	1250
Доля рабочих в общей среднесписочной численности персонала основной деятельности, %	70,0	72,0

Определить абсолютный прирост объема произведенной продукции за счет отдельных факторов, используя последовательно-цепной индексный метод выявления влияния факторов.

Решение

На объем произведенной продукции (Q) влияют два фактора: годовая производительность труда персонала основной деятельности (ПТ) и численность персонала основной деятельности (N).

Годовая производительность труда персонала основной деятельности

$$ПТ = \frac{Q}{N};\ Q = ПТ \cdot N.$$

Изменение годовой производительности персонала основной деятельности объясняется изменением годовой производительности труда рабочих (ПТр) и их доли в общей численности (a).

Годовая производительность персонала основной деятельности

$$ПТ = ПТр \cdot a.$$

Таким образом, на объем произведенной продукции влияют три фактора:

уровень годовой производительности труда рабочих (ПТр);

доля рабочих в общей численности персонала (a);

численность персонала основной деятельности (N).

Изменение объема произведенной продукции определяется следующим индексом:

$$J_Q = \frac{Q_1}{Q_0} = \frac{ПТр_1 a_1 N_1}{ПТр_0 a_0 N_0}.$$

В данном индексе требование правильности расположения факторов учтено: увеличение цепи сомножителей на каждое новое звено приводит к произведению, имеющему конкретное экономическое содержание.

Необходимые для расчета показатели следующие:

$$ПТр_0 = \frac{Q_1}{Np_0},$$

где N_{p_0} – численность рабочих в прошлом году.

$$N_{p_0} = 1200 \cdot 0,7 = 840 \text{ чел.};$$

$$ПТр_0 = \frac{1176000}{840} = 1400 \text{ тыс. руб.};$$

$$ПТр_1 = \frac{Q_1}{Np_1},$$

где N_{p_1} – численность рабочих в текущем году.

$$N_{p_1} = 1250 \cdot 0,72 = 900 \text{ чел.};$$

$$ПТр_1 = \frac{1530000}{900} = 1700 \text{ тыс. руб.}$$

$$J_Q = \frac{1700 \cdot 0,72 \cdot 1250}{1400 \cdot 0,70 \cdot 1200} = \frac{1530000}{1176000} = 1,301, \text{ или } 130,1\%.$$

Следовательно, объем продукции увеличился на 30,1%, что в абсолютном размере составило 354 млн руб. (354 000 тыс. руб.).

Для выявления влияния факторов исчисляются факторные индексы:

а) изменения годовой производительности труда рабочих –

$$J_{ПТр} = \frac{ПТр_1 \cdot a_1 \cdot N_1}{ПТр_0 \cdot a_1 \cdot N_1} = \frac{1700 \cdot 0,72 \cdot 1250}{1400 \cdot 0,72 \cdot 1250} = \frac{1530000}{1260000} = 1,214, \text{ или } 121,4\%.$$

Следовательно, годовая производительность труда рабочих возросла на 21,4%; объем продукции увеличился на 270 млн руб. (1530000 – 1260000 = 270000 тыс. руб.);

б) изменения доли рабочих в общей численности персонала основной деятельности –

$$J_a = \frac{ПТр_0 \cdot a_1 \cdot N_1}{ПТр_0 \cdot a_0 \cdot N_1} = \frac{1400 \cdot 0,72 \cdot 1250}{1400 \cdot 0,70 \cdot 1200} = \frac{1260000}{1225000} = 1,029, \text{ или } 102,9\%.$$

Удельный вес рабочих в общей численности персонала увеличился на 2,9%, за счет этого объем продукции возрос на 35 млн руб. (1260000 – 1225000 = 35000 тыс. руб.);

в) изменения численности персонала основной деятельности –

$$J_N = \frac{ПТр_0 \cdot a_0 \cdot N_1}{ПТр_0 \cdot a_0 \cdot N_0} = \frac{1400 \cdot 0,70 \cdot 1250}{1400 \cdot 0,70 \cdot 1200} = \frac{1225000}{1176000} = 1,042, \text{ или } 104,2\%.$$

Численность всего персонала возросла на 4,2%, поэтому объем продукции увеличился на 49 млн руб. (1225000 – 1176000 = = 49000 тыс. руб.).

Произведение факторных индексов соответствует общему относительному изменению объема продукции:

$$J_Q = J_{ПТр} \cdot J_a \cdot J_N = 1,214 \cdot 1,029 \cdot 1,42 = 1,301,$$

что соответствует ранее полученной величине.

Алгебраическая сумма абсолютных изменений объема продукции за счет отдельных факторов соответствует фактическому изменению.

$$\Delta_Q = 270 + 35 + 49 = 354 \text{ млн руб.}$$

7.2. Задачи для самостоятельной работы

7.12. По предприятию имеются следующие данные о выпуске продукции (табл. 7.9).

Таблица 7.9

Вид продукции	Выпуск, тыс. т			Сопоставимая цена за 1 т, руб.
	прошлый год	отчетный год		
		план	фактически	
Сталь арматурная	20	25	24	1600
Прокат листовой	18	20	22	1950
Швеллер	4	5	6	1900

Определить в целом по предприятию:

1) запланированный процент увеличения физического объема продукции для отчетного года;

2) фактический процент изменения физического объема продукции по сравнению с прошлым годом;

3) процент выполнения плана выпуска продукции.

7.13. По заводу нефтепродуктов имеются следующие данные (табл. 7.10).

Определить:

1) процент выполнения плана по выпуску продукции каждого вида;

2) процент выполнения плана выпуска продукции в целом по предприятию.

Таблица 7.10

Вид продукции	Выпуск, тыс. т		Общие затраты на производство продукции по плану, тыс. руб.
	план	фактически	
Бензин А-76 неэтилированный	40	42	46 800
Бензин А-92/93 неэтилированный	38	38	51 870
Авиакерасин ТС-1	70	77	51 170
Дизельное топливо	110	99	82 500

7.14. Количество произведенной продукции в натуральном выражении уменьшилось на 2,5%, а отпускные цены на продукцию увеличились на 5,2%.

Определить, на сколько процентов изменилась стоимость продукции в отчетном году по сравнению с прошлым годом.

7.15. По товарной бирже имеются следующие данные о реализации грузовых автомобилей (табл. 7.11).

Таблица 7.11

Марка автомобиля	Процент снижения (–), повышения (+) оптовых цен в отчетном периоде по сравнению с базисным	Стоимость реализованной продукции в отчетном периоде, тыс. руб.
МАЗ-5551	–2,0	7 360
КамАЗ-55111	+3,8	15 200
КамАЗ-53212	–0,6	9 000

Определить среднее изменение цен на грузовые автомобили.

7.16. Стоимость продукции в ценах соответствующих лет составила: в 1999 г. – 25 млн руб., в 2000 г. – 32,5 млн руб. Индекс цен в 2000 г. по сравнению с 1999 г. составил 115%. Производительность труда на одного работающего возросла за этот период со 120 до 144 тыс. руб.

Определить индексы физического объема продукции, производительности труда и численности работающих.

7.17. Имеются следующие данные о выработке и себестоимости кирпича по двум однородным предприятиям (табл. 7.12).

Таблица 7.12

№ предприятия	Базисный период		Отчетный период	
	произведено продукции, тыс. шт.	себестоимость 1000 шт., руб.	произведено продукции, тыс. шт.	себестоимость 1000 шт., руб.
1	4000	560	4500	540
2	5500	520	6000	520

Определить:

1) изменение себестоимости производства 1000 шт. кирпича по каждому предприятию;

2) абсолютную экономию (перерасход) за счет изменения себестоимости, полученную каждым предприятием;

3) среднюю себестоимость производства 1000 шт. кирпича в отчетном и базисном периодах в целом по группе предприятий;

4) изменение средней себестоимости по группе предприятий в отчетном периоде по сравнению с базисным;

5) влияние изменения себестоимости производства 1000 шт. кирпича на отдельных предприятиях на изменение средней себестоимости;

6) влияние на изменение средней себестоимости структурных сдвигов в составе продукции;

7) абсолютное изменение затрат на производство, в том числе за счет увеличения продукции, а также за счет снижения себестоимости производства 1000 шт. кирпича: а) по каждому предприятию; б) в целом по двум предприятиям.

7.18. В прошлом году металлургический завод выпустил чугуна на 5000 тыс. руб., стали – на 3 500, проката – на 2 100 тыс. руб. На отчетный год предусмотрено увеличение производства чугуна на 12,0%, стали – на 7,5, проката – на 3,2%.

Определить, на сколько процентов должно увеличиться производство продукции в целом по предприятию.

7.19. В табл. 7.13 представлены данные о себестоимости продукции машиностроительного завода.

Таблица 7.13

Вид продукции	Произведено продукции, ед.			Себестоимость единицы продукции, тыс. руб.		
	апрель	май	июнь	апрель	май	июнь
Погрузчики	22	24	25	26	24	22
Электрокары	16	18	19	38	40	48
Конвейеры	18	19	20	53	52	50

Исчислить цепные агрегатные индексы себестоимости.

7.20. В табл. 7.14 представлены данные по предприятию, осуществляющему капитальный ремонт автомобилей КамАЗ.

Таблица 7.14

Вид продукции	Число ремонтов в прошлом году, ед.	Изменение числа ремонтов в текущем году по сравнению с прошлым годом, %	Сопоставимая цена за единицу продукции, тыс. руб.
Капитальный ремонт автомобилей	2500	105	30,0
Капитальный ремонт двигателей	1200	107	8,5
Капитальный ремонт агрегатов	850	98	0,4

Определить общий индекс объема продукции.

7.21. По машиностроительному предприятию объем выпущенной продукции во II квартале увеличился по сравнению с I кварталом на 10%, в III квартале по сравнению со II кварталом он снизился на 1,2%, а в IV квартале по сравнению с III кварталом объем выпущенной продукции увеличился на 12,5%.

Определить, как изменился объем выпущенной продукции на предприятии в IV квартале по сравнению с I кварталом.

7.22. Общие затраты на производство продукции завода составили: в 1998 г. – 7 800 тыс. руб.; 1999 г. – 8 500; 2000 г. – 9 000 тыс. руб. Себестоимость продукции в 1999 г. снизилась в среднем по сравнению с 1998 г. на 3,0%, а в 2000 г. по сравнению с 1999 г. – на 1,2%.

Определить изменение физического объема продукции завода за эти годы.

7.23. Табл. 7.15 содержит следующие данные по нефтеперерабатывающему заводу.

Таблица 7.15

Вид продукции	Январь		Февраль		Март	
	количество, тыс. т	оптовая цена за 1 т, руб.	количество, тыс. т	оптовая цена за 1 т, руб.	количество, тыс. т	оптовая цена за 1 т, руб.
Бензин А-76 (неэтилированный)	40	1800	42	1800	45	1820
Авиакеросин ТС-1	60	1120	65	1150	70	1150
Дизельное топливо	120	1160	125	1180	130	1080

Определить базисные агрегатные индексы физического объема продукции и агрегатные индексы стоимости продукции. Объяснить, чем вызвано расхождение этих индексов.

7.24. По деревообрабатывающему комбинату имеются следующие данные о выпуске продукции и ее себестоимости (табл. 7.16).

Таблица 7.16

Вид продукции	Себестоимость 1 м^3, руб.			Произведено в отчетном году, тыс. м^3
	в предыдущем году	по плану текущего года	фактически в текущем году	
Доска обрезная	440	480	475	850
Доска-вагонка	750	825	820	490

Определить среднее изменение себестоимости единицы продукции в целом по комбинату и соответствующие суммы экономии (перерасхода):

1) фактически по сравнению с планом;

2) фактически по сравнению с прошлым годом.

7.25. Имеются следующие данные о выработке продукции, нормах затрат сырья и ценах на сырье (табл. 7.17).

Таблица 7.17

Изделие	Выработано продукции в отчетном периоде, шт.	Расход сырья на единицу продукции по норме, кг		Цена 1 кг сырья, руб.	
		базисный период	отчетный период	базисный период	отчетный период
1	2000	10	8	5,1	5,0
2	400	15	13	10,2	10,5

Определить:

1) индивидуальные индексы норм и цен;

2) агрегатные индексы цен, норм и затрат на сырье;

3) размер уменьшения (увеличения) затрат на сырье в отчетном периоде по сравнению с базисным на производство всей продукции.

7.26. По предприятию за два месяца имеются следующие данные о выпуске арматурной стали и затратах на нее (табл. 7.18).

Таблица 7.18

Показатель	Январь	Февраль
Выпуск арматурной стали, тыс. т	120	140
Общая сумма затрат на выпуск, тыс. руб.	132 000	147 000

Определить абсолютное изменение общей суммы затрат предприятия за счет:

1) изменения выпуска продукции;

2) изменения ее себестоимости, используя последовательно-цепной индексный метод разложения по факторам.

7.27. Товарооборот предприятия увеличился в отчетном году по сравнению с прошлым годом в 1,2 раза при снижении цен за этот же период в среднем на 5%.

Как изменился объем реализованной товарной массы в отчетном году?

7.28. Работа предприятия характеризуется следующими данными (табл. 7.19).

Таблица 7.19

Вид продукции	Выпуск по плану, тыс. шт.	Процент выполнения плана по количеству	Затраты времени на единицу продукции, чел.-ч	
			план	отчет
А	140	120,0	15	14
Б	200	108,0	35	33

Определить в целом по предприятию:

1) процент изменения общих затрат рабочего времени по сравнению с планом;

2) среднее изменение затрат рабочего времени на единицу продукции;

3) процент выполнения плана по выпуску продукции.

7.29. По автотранспортному предприятию объем выполненной транспортной работы во II квартале отчетного года увеличился по сравнению с I кварталом на 15,2%; в III квартале по сравнению со II кварталом увеличение составило 11,2%, а в IV квартале по сравнению с III кварталом произошло снижение на 2,5%.

Определить, как изменился (в %) объем транспортной работы в IV квартале по сравнению с I и II кварталами.

7.30. По двум ТЭЦ за два месяца имеются следующие данные о себестоимости выработанной электроэнергии (табл. 7.20).

Таблица 7.20

№ ТЭЦ	Сентябрь		Октябрь	
	выработано электроэнергии, млн кВт · ч	себестоимость 1 кВт · ч, руб.	выработано электроэнергии, млн кВт · ч	себестоимость 1 кВт · ч, руб.
1	4000	0,08	5000	0,10
2	600	0,09	700	0,12

Определить в целом по двум ТЭЦ:

1) изменение средней себестоимости выработанной электроэнергии в процентах и в абсолютном размере;

2) абсолютное изменение средней себестоимости за счет отдельных факторов: а) изменения себестоимости на отдельных ТЭЦ; б) структурных сдвигов в общем объеме выработанной электроэнергии.

7.31. Выпуск продукции по зоне технического обслуживания грузовых автомобилей характеризуется следующими данными (табл. 7.21).

Таблица 7.21

Марка автомобиля	Выполнено технических обслуживаний № 2, ед.			Трудоемкость одного обслуживания, чел.-ч		
	январь	февраль	март	январь	февраль	март
ЗИЛ-130В1	35	36	38	13,6	13,6	12,8
КамАЗ-55111	12	18	22	20,3	20,0	20,1

Определить в целом по предприятию цепные и базисные индексы физического объема продукции, затрат рабочего времени на единицу продукции и общих затрат рабочего времени. Проверить взаимосвязь вычисленных индексов.

Сформулировать вывод.

7.32. Реализация грузовых автомобилей на рынках двух регионов в отчетном периоде следующая (табл. 7.22).

Таблица 7.22

Марка автомобиля	Регион А		Регион Б	
	реализовано, шт.	модальная цена за 1 шт., тыс. руб.	реализовано, шт.	модальная цена за 1 шт., тыс. руб.
МАЗ-5551	600	92,0	620	90,0
КамАЗ-55111	800	127,3	790	128,0
КамАЗ-53212	1100	120,0	1150	121,5

Определить индивидуальные и общие территориальные индексы физического объема реализации продукции и цен.

7.33. Имеются следующие данные о производительности труда и структуре численности работающих по двум предприятиям (табл. 7.23).

Таблица 7.23

Предприятие	Выработка продукции в сопоставимых ценах на одного работающего, тыс. руб.		Удельный вес по численности работающих, %	
	Базисный период	Отчетный период	Базисный период	Отчетный период
1	12,0	12,6	35	30
2	15,0	16,5	65	70

Известно также, что общая численность работающих на обоих предприятиях составляет 1200 человек в отчетном периоде.

Определить абсолютное изменение объема продукции по предприятиям за счет изменения средней выработки на одного работающего.

7.34. Имеются следующие данные о реализации товаров на рынках двух городов за квартал (табл. 7.24).

Таблица 7.24

Наименование товара	Единица измерения	Город В		Город Г	
		реализовано	цена за единицу, руб.	реализовано	цена за единицу, руб.
А	кг	200	20	250	18
Б	шт.	120	400	180	360

Рассчитайте территориальные индексы цен по всей совокупности товаров при различных приемах взвешивания.

7.35. Машиностроительный завод производит поставку трех видов продукции на экспорт (табл. 7.25).

Таблица 7.25

Вид продукции	Объем поставок, тыс. шт.			Внешнеторговая цена за единицу продукции, долл. США		
	I квартал	II квартал	III квартал	I квартал	II квартал	III квартал
А	200	220	260	100	102	104
Б	410	390	400	340	325	330
В	100	120	130	408	410	420

Определить изменение (в абсолютном размере) общей суммы выручки от поставок продукции на экспорт во II квартале по сравнению с I кварталом и в III квартале по сравнению со II кварталом, в том числе за счет изменения физического объема поставок и изменения внешнеторговых цен.

7.36. Продукт А производится на трех предприятиях региона (табл. 7.26).

Таблица 7.26

№ предприятия	Себестоимость за единицу продукта, долл. США		Физический объем выпуска, тыс. шт.	
	базисный период	отчетный период	базисный период	отчетный период
1	75	72	8	14
2	56	57	18	17
3	60	62	7	5

Определить:

1) изменение средней себестоимости продукта А в процентах и в абсолютном размере;

2) абсолютное изменение средней себестоимости за счет отдельных факторов: изменения себестоимости по отдельным предприятиям, изменения их структурных сдвигов в общем объеме выпуска продукции.

7.37. Общие затраты на производство продукции увеличились во II квартале по сравнению с I кварталом на 3,8%, объем продукции предприятия снизился на 2,0%.

Определить, на сколько процентов изменилась в среднем себестоимость продукции.

7.38. По грузовому автотранспортному предприятию имеются следующие данные (табл. 7.27).

Таблица 7.27

Показатель	Январь	Февраль
Грузооборот, млн ткм	200	250
Себестоимость 10 ткм, руб.	4,8	4,5

Определить абсолютное изменение общей суммы затрат предприятия на транспортную работу, в том числе за счет изменения объема работы и изменения себестоимости работы.

7.39. Общая стоимость продукции завода составила, млн руб.: в 1998 г. – 120; 1999 г. – 135; 2000 г. – 145,8. Физический объем продукции в 1999 г. по сравнению с 1998 г. снизился на 2,5%, а в 2000 г. по сравнению с 1999 г. повысился на 4,0%.

Определить среднее изменение отпускных цен на продукцию завода за эти годы.

7.40. Металлургический завод выпускает сталь арматурную и листовой прокат. Выпуск стали в 2000 г. по сравнению с 1999 г. возрос на 8,0%, а проката – на 16,0%.

Определить общий индекс физического объема продукции, если известно, что в общем выпуске продукции 1999 г. удельный вес стали составил 70%, а проката – 30%.

7.41. По предприятию за два месяца имеются следующие данные (табл. 7.28).

Таблица 7.28

Показатель	Март	Апрель
Выпуск арматурной стали, тыс. т	120	140
Общая сумма затрат на выпуск, тыс. руб.	132000	147000

Определить абсолютное изменение общей суммы затрат предприятия за счет изменения выпуска продукции.

7.42. Работа предприятия характеризуется следующими данными (табл. 7.29).

Определить в целом по предприятию:

1) процент выполнения плана по выпуску продукции;

2) процент изменения общих затрат рабочего времени по сравнению с планом.

Таблица 7.29

Вид продукции	Выпуск по плану, шт.	Процент выполнения плана по выпуску продукции	Затраты рабочего времени на единицу продукции, чел.-ч	
			план	отчет
А	140	120,0	15	14
Б	200	108,0	35	33

Ответы к задачам для самостоятельной работы

7.12. 1) увеличение на 18,5%; 2) увеличение на 24,1%; 3) 104,7%. **7.13.** 1) 105,0; 100,0; 110,0; 90,9; 2) 99,7%. **7.14.** Увеличение на 2,6%. **7.15.** Повышение на 1,13%. **7.16.** 1,13; 1,20; 0,94. **7.18.** Увеличение на 8,8%. **7.19.** 0,987; 1,026. **7.20.** 1,052. **7.21.** Увеличение на 22,27%. **7.24.** 1) снижение на 0,82%, экономия 6 700 тыс. руб; 2) повышение на 8,64%, перерасход 64 050 тыс. руб. **7.26.** За счет: 1) снижения себестоимости уменьшение на 7 000 тыс. руб.; 2) роста выпуска продукции увеличение на 22 000 тыс. руб. **7.27.** Увеличение на 26,3%. **7.28.** 1) 104,18%; 2) 0,9405, или снижение на 5,95%; 3) 110,77%. **7.29.** Увеличение на 24,9%; увеличение на 8,4%. **7.30.** 1) $\overline{x}_0 = 0{,}0813$ руб.; $\overline{x}_1 = 0{,}1025$ руб.; увеличение на 26,1%, или на 0,0212 руб.; 2) а) увеличение на 0,0213 руб.; б) снижение на 0,001 руб. **7.32.** $i_{q_{А/Б}}$: 0,968; 1,013; 0,957; $i_{p_{АБ}}$: :1,022; 0,995 : 0,988; $\overline{P}_0$: 91,0; 127,6; 120,8; $I_{q_{А/Б}} = 0{,}978$; $I_{p_{А/Б}} = 0{,}997$. **7.33.** Предприятие № 1 – 216 тыс. руб.; предприятие № 2 – 1 260 тыс. руб. **7.35.** II кв.: $\Delta_0 = -1810$; $\Delta^q = 3360$; $\Delta^p = -5170$; III кв.: $\Delta_0 = 15250$; $\Delta^q = 11430$; $\Delta^p = 3820$. **7.36.** $\overline{z}_0 = 61{,}5$; $\overline{z}_1 = 63{,}5$; $\Delta_{\overline{z}} = 2{,}0$; 103,3%; $\Delta^z_{\overline{z}} = -0{,}5$; $\Delta^q_{\overline{z}} = 2{,}5$. **7.37.** Увеличение на 5,9%. **7.38.** $\Delta_{qz} = 16{,}5$ млн руб.; 24,0 млн руб.; (–7,5) млн руб. **7.39.** $J_{p_{1/0}} = 1{,}154$; $J_{p_{2/1}} = 1{,}038$; **7.40.** 1,104; **7.41.** 22000 тыс. руб. **7.42.** 1) 110,76%; 2) увеличение на 4,18%.

ПРИЛОЖЕНИЯ

Приложение 1

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

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	0355	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
2,7	0104	0101	0099	0096	0093	0091	0088	0086	0084	0081
2,8	0079	0077	0075	0073	0071	0069	0067	0065	0063	0061
2,9	0060	0058	0056	0055	0053	0051	0050	0048	0047	0046
3,0	0044	0043	0042	0040	0039	0038	0037	0036	0035	0034
4,0	0001	0001	0001	0000	0000	0000	0000	0000	0000	0000

* Значения ординат увеличины в 10 000 раз.

Приложение 2

Значение верхнего q предела $\chi^2 q$ в зависимости от вероятности $P(\chi^2 > \chi^2 q)$ и числа степеней свободы χ^2-распределения

Число степеней свободы	Вероятность $P(\chi^2 > \chi^2 q)$							
	0,99	0,98	0,95	0,90	0,80	0,70	0,50	0,30
1	1,64	2,7	3,8	5,4	6,6	7,9	9,5	10,8
2	3,22	4,6	6,0	7,8	9,2	11,6	12,3	13,8
3	4,64	6,3	7,8	9,8	11,3	12,8	14,8	16,3
4	6,0	7,8	9,5	11,7	13,3	13,9	16,9	18,5
5	7,3	9,2	11,1	13,4	15,1	16,3	18,9	20,5
6	8,6	10,6	12,6	15,0	16,8	18,6	20,7	22,5
7	9,8	12,0	14,1	16,6	18,5	20,3	22,6	24,3
8	11,0	13,4	15,5	18,2	20,1	21,9	24,3	26,1
9	12,2	14,7	16,9	19,7	21,7	23,6	26,1	27,9
10	13,4	16,0	18,3	21,2	23,2	25,2	27,7	29,6
11	14,6	17,3	19,7	22,6	24,7	26,8	29,4	31,3
12	15,8	18,5	21,0	24,1	26,2	28,3	31,0	32,9
13	17,0	19,8	22,4	25,5	27,7	29,8	32,5	34,5
14	18,2	21,1	23,7	26,9	29,1	31,0	34,0	36,1
15	19,3	22,3	25,0	28,3	30,6	32,5	35,5	37,7
16	20,5	23,5	26,3	29,6	32,0	34,0	37,0	39,2
17	21,6	24,8	27,6	31,0	33,4	35,5	38,5	40,8
18	22,8	26,0	28,9	32,3	34,8	37,0	40,0	42,3
19	23,9	27,2	30,1	33,7	36,2	38,9	41,5	43,8
20	25,0	28,4	31,4	35,0	37,6	40,0	43,0	45,3
21	26,2	29,6	32,7	36,3	38,9	41,5	44,5	46,8
22	27,3	30,8	33,9	37,7	40,3	42,5	46,0	48,3
23	28,4	32,0	35,2	39,0	41,6	44,0	47,5	49,7
24	29,6	33,2	36,4	40,3	43,0	45,5	48,5	51,2
25	30,7	34,4	37,7	41,6	44,3	47,0	50,0	52,6
26	31,8	35,6	38,9	42,9	45,6	48,0	51,5	54,1
27	32,9	36,7	40,1	44,1	47,0	49,5	53,0	55,5
28	34,0	37,9	41,3	45,4	48,3	51,0	54,5	56,9
29	35,1	39,1	42,6	46,7	49,9	52,5	56,0	58,3
30	36,3	40,3	43,8	48,0	50,9	54,0	57,5	59,7

Приложение 3

Удвоенная нормированная функция Лапласа

$$\Phi_{(t)} = \frac{2}{\sqrt{2\pi}} \int_0^t e^{-\frac{t^2}{2}} \cdot dt = \frac{1}{\sqrt{2\pi}} \int_{-t}^{+t} e^{-\frac{t^2}{2}} \cdot dt$$

t	$\Phi(t)$	t	$\Phi(t)$	t	$\Phi(t)$	t	$\Phi(t)$
0,00	0,00000	0,30	0,23582	0,60	0,45149	0,90	0,63188
01	00798	31	24344	61	45814	91	63718
02	01596	32	25103	62	46474	92	64243
03	02393	33	25860	63	47131	93	64763
04	03191	34	26614	64	47783	94	65278
05	03988	35	27366	65	48431	95	65789
06	04784	36	28115	66	49075	96	66294
07	05581	37	28862	67	49714	97	66795
08	06376	38	29605	68	50350	98	67291
09	07171	39	30346	69	50981	99	67783
0,10	0,07966	0,40	0,31084	0,70	0,51607	1,00	0,68269
11	08759	41	31819	71	52230	01	68750
12	09552	42	32552	72	52848	02	69227
13	10348	43	33280	73	53461	03	69699
14	11134	44	34006	74	54070	04	70166
15	11924	45	34729	75	54675	05	70628
16	12712	46	35448	76	55275	06	71086
17	13499	47	36164	77	55870	07	71538
18	14285	48	36877	78	56461	08	71986
19	15069	49	37587	79	57047	09	72429
0,20	0,15852	0,50	0,38292	0,80	0,57629	1,10	0,72867
21	16633	51	38995	81	58206	11	73300
22	17413	52	39694	82	58778	12	73729
23	18191	53	40389	83	59346	13	74152
24	18967	54	41080	84	59909	14	74571
25	19741	55	41768	85	60468	15	74986
26	20514	56	42452	86	61021	16	75395
27	21284	57	43132	87	61570	17	75800
28	22052	58	43809	88	62114	18	76200
29	22818	59	44481	89	62953	19	76595

Продолжение

t	$\Phi(t)$	t	$\Phi(t)$	t	$\Phi(t)$	t	$\Phi(t)$
1,20	0,76986	1,55	0,87886	1,90	0,94257	2,25	0,97555
21	77372	56	88124	91	94387	26	97618
22	77754	57	88358	92	94514	27	97679
23	78130	58	88589	93	94639	28	97739
24	78502	59	88817	94	94792	29	97798
25	78870	1,60	0,89040	95	94882	2,30	0,97855
26	79233	61	89260	96	95000	31	97911
27	79592	62	89477	97	95116	32	97966
28	79945	63	89690	98	95230	33	98019
29	80295	64	89899	99	95341	34	98072
1,30	0,80640	65	90106	2,00	0,95450	35	98123
31	80980	66	90309	01	95557	36	98172
32	81316	67	90508	02	95662	37	98221
33	81948	68	90704	03	95764	38	98269
34	81975	69	90897	04	95865	39	98315
35	82298	1,70	0,91087	05	95964	2,40	0,98360
36	82617	71	91273	06	96060	41	98405
37	82931	72	91457	07	96155	42	98448
38	83241	73	91637	08	96247	43	98490
39	83547	74	91814	09	96338	44	98531
1,40	0,83849	75	91988	2,10	0,96427	45	98571
41	84146	76	92159	11	96514	46	98611
42	84439	77	92327	12	96599	47	98649
43	84728	78	92492	13	96683	48	98686
44	85013	79	92655	14	96765	49	98723
45	85294	1,80	0,92814	15	96844	2,50	0,98758
46	85571	81	92970	16	96923	51	98793
47	85844	82	93124	17	96999	52	98826
48	86113	83	93275	18	97074	53	98859
49	86328	84	93423	19	97148	54	98891
1,50	0,86639	85	93569	2,20	0,97219	55	98923
51	86696	86	93711	21	97289	56	98953
52	87140	87	93852	22	97358	57	98983
53	87398	88	93989	23	97425	58	99012
54	87644	89	94124	24	97491	59	99040

Продолжение

t	$\Phi(t)$	t	$\Phi(t)$	t	$\Phi(t)$	t	$\Phi(t)$
2,60	0,99068	2,95	0,99682	3,30	0,99903	3,65	0,99974
61	99095	96	99692	31	99907	66	99975
62	99121	97	99702	32	99910	67	99976
63	99146	98	99712	33	99913	68	99977
64	99171	99	99721	34	99916	69	99978
65	99195	3,00	0,99730	35	99919	3,70	0,99978
66	99219	01	99739	36	99922	71	99979
67	99241	02	99747	37	99925	72	99980
68	99283	03	99755	38	99928	73	99981
69	99285	04	99763	39	99930	74	99982
2,70	0,99307	05	99771	3,40	0,99933	75	99982
71	99327	06	99779	41	99935	76	99983
72	99347	07	99786	42	99937	77	99984
73	99367	08	99793	43	99940	78	99984
74	99386	09	99800	44	99942	79	99985
75	99404	3,10	0,99806	45	99944	3,80	0,99986
76	99422	11	99813	46	99946	81	99986
77	99439	12	99819	47	99948	82	99987
78	99456	13	99825	48	99950	83	99987
79	99473	14	99831	49	99952	84	99988
2,80	0,99489	15	99837	3,50	0,99953	85	99988
81	99505	16	99842	51	99955	86	99989
82	99520	17	99848	52	99957	87	99989
83	99535	18	99853	53	99958	88	99990
84	99549	19	99858	54	99960	89	99990
85	99563	3,20	0,99863	55	99961	3,90	0,9990
86	99576	21	99867	56	99963	91	99991
87	99590	22	99872	57	99964	92	99992
88	99502	23	99876	58	99966	93	99992
89	99615	24	99880	59	99967	94	99992
2,90	0,99627	25	99855	3,60	0,99968	95	99992
91	99639	26	99889	61	99969	96	99993
92	99650	27	99892	62	99971	97	99993
93	99661	28	99896	63	99972	98	99993
94	99672	29	99800	64	99973	99	99993

Значения функции $S(t)$ для распределения Стьюдента в зависимости от t и числа k степеней свободы

t \ k	1	2	3	4	5	6	7	8	9	10
0,0	0,500	0,500	0,500	0,500	0,500	0,500	0,500	0,500	0,500	0,500
0,1	532	535	537	537	538	538	538	539	539	539
0,2	563	570	573	574	575	576	576	577	577	577
0,3	593	604	608	610	612	613	614	614	614	615
0,4	621	636	642	645	647	648	650	650	651	651
0,5	648	667	674	678	681	683	684	685	686	686
0,6	672	695	705	710	713	715	716	717	718	719
0,7	694	722	733	739	742	745	747	748	749	750
0,8	715	746	759	766	770	773	775	777	778	779
0,9	733	768	783	790	795	799	801	803	804	805
1,0	0,750	789	804	813	818	822	825	827	828	830
1,1	765	807	824	834	839	843	846	848	850	851
1,2	779	824	842	852	858	862	865	868	870	871
1,3	791	838	858	868	875	879	883	885	887	889
1,4	803	852	872	883	890	894	898	900	902	904
1,5	813	864	885	896	903	908	911	914	916	918
1,6	822	875	896	908	915	920	923	926	928	930
1,7	831	884	906	918	925	930	934	936	938	940
1,8	839	893	915	927	934	939	943	945	947	949
1,9	846	901	923	935	942	947	950	953	955	957
2,0	0,852	908	930	942	949	954	957	960	962	963
2,2	864	921	942	954	960	965	968	970	972	974
2,4	874	931	952	963	969	973	976	978	980	981
2,6	883	938	960	970	976	980	982	984	986	987
2,8	891	946	966	976	981	984	987	988	990	991
3,0	898	952	971	980	985	988	990	992	992	993
3,2	904	957	975	984	988	991	992	994	995	995
3,4	909	962	979	986	990	993	994	995	996	997
3,6	914	965	982	989	992	994	996	996	997	998
3,8	918	969	984	990	994	996	997	997	998	998
4,0	922	971	986	992	995	996	997	998	998	999
4,2	926	974	988	993	996	997	998	998	999	999
4,4	929	976	989	994	996	998	998	999	999	999
4,6	932	978	990	995	997	998	999	999	999	1,000
4,8	935	980	991	996	998	998	999	999	1,000	
5,0	937	981	992	996	998	999	999	1,000		
5,2	940	982	993	997	998	999	999			

Продолжение

t \ k	1	2	3	4	5	6	7	8	9	10
5,4	942	984	994	997	998	999	1,000			
5,6	944	985	994	998	999	999				
5,8	946	986	995	998	999	999				
6	947	987	995	998	999	1,000				

t \ k	11	12	13	14	15	16	17	18	19	20
0,0	0,500	0,500	0,500	0,500	0,500	0,500	0,500	0,500	0,500	0,50000
0,1	539	539	539	539	539	539	539	539	539	53983
0,2	577	578	578	578	578	578	578	578	578	57926
0,3	615	615	616	616	616	616	616	616	616	61791
0,4	652	652	652	652	653	653	653	653	653	65542
0,5	686	687	687	688	688	688	688	688	689	69146
0,6	720	720	721	721	721	722	722	722	722	72575
0,7	751	751	752	752	753	753	753	754	754	75804
0,8	780	780	781	782	782	782	783	783	783	78814
0,9	806	807	808	808	809	809	810	810	810	81594
1,0	831	832	832	833	833	834	834	835	835	84134
1,1	853	854	854	855	856	856	857	857	858	86433
1,2	872	873	874	875	876	876	877	877	878	88493
1,3	890	891	892	893	893	894	894	895	895	90320
1,4	906	907	908	908	909	910	910	911	911	91924
1,5	919	920	921	922	923	924	924	924	925	93319
1,6	931	932	933	934	935	935	936	936	937	94520
1,7	941	943	944	945	945	946	946	947	947	95543
1,8	950	952	952	953	954	955	955	956	956	96407
1,9	958	959	960	961	962	962	963	963	964	97128
2,0	965	967	967	967	968	969	969	970	970	97725
2,2	975	976	977	977	978	979	979	979	980	98610
2,4	982	983	984	985	985	986	986	986	987	99180
2,6	988	988	989	990	990	990	991	991	991	99534
2,8	0,991	0,992	0,992	0,993	0,993	0,994	0,994	0,994	0,994	0,99744
3,0	994	994	995	995	996	996	996	996	996	99865
3,2	996	996	996	997	997	997	997	998	998	99931
3,4	997	997	998	998	998	998	998	998	998	99966
3,6	998	998	998	999	999	999	999	999	999	99984
3,8	998	999	999	999	999	999	999	999	999	99993
4,0	999	999	999	999	999	1,000	1,000	1,000	1,000	99997
4,2	999	999	1,000	1,000	1,000					99999
4,4	1,000	1,000								99999

Таблица значения *F* для доверительной вероятности *P* = (1 – 0,05) = 0,95

K_2 \ K_1	1	2	3	4	5	6	8	12	24
1	161,45	199,50	215,72	224,57	230,17	233,97	238,89	243,91	249,04
2	18,54	19,00	19,16	19,25	19,30	19,33	19,37	19,41	19,45
3	10,18	9,55	9,28	9,12	9,01	8,94	8,84	8,74	8,64
4	7,71	6,94	6,59	6,39	6,26	6,16	6,04	5,91	5,77
5	6,61	5,79	5,41	5,19	5,05	4,95	4,82	4,68	4,53
6	5,99	5,14	4,76	4,53	4,39	4,28	4,15	4,00	3,84
7	6,59	4,74	4,35	4,12	3,97	3,87	3,73	3,57	3,41
8	5,32	4,46	4,07	3,84	3,69	3,58	3,44	3,28	3,12
9	5,12	4,26	3,86	3,63	3,48	3,37	3,23	3,07	2,90
10	4,96	4,10	3,71	3,48	3,33	3,22	3,07	2,91	2,74
11	4,84	3,98	3,59	3,36	3,20	3,09	2,95	2,79	2,61
12	4,75	3,88	3,49	3,26	3,11	3,00	2,85	2,69	2,50
13	4,67	3,80	3,41	3,18	3,02	2,92	2,77	2,60	2,42
14	4,60	3,74	3,34	3,11	2,96	2,85	2,70	2,53	2,35
15	4,54	3,68	3,29	3,06	2,90	2,79	2,64	2,48	2,29
16	4,49	3,63	3,24	3,01	2,85	2,74	2,59	2,42	2,24
17	4,45	3,59	3,20	2,96	2,81	2,70	2,55	2,38	2,19
18	4,41	3,55	3,16	2,93	2,77	2,66	2,51	2,34	2,15
19	4,38	3,52	3,13	2,90	2,74	2,63	2,48	2,31	2,11
20	4,35	3,49	3,10	2,87	2,71	2,60	2,45	2,28	2,08
21	4,32	3,47	3,07	2,84	2,68	2,57	2,42	2,25	2,05
22	4,30	3,44	3,05	2,82	2,66	2,55	2,40	2,23	2,03
23	4,28	3,42	3,03	2,80	2,64	2,53	2,38	2,20	2,00
24	4,26	3,40	3,01	2,78	2,62	2,51	2,36	2,18	1,98
25	4,24	3,38	2,99	2,76	2,60	2,49	2,34	2,16	1,96
26	4,22	3,37	2,98	2,74	2,59	2,47	2,32	2,15	1,95
27	4,21	3,35	2,96	2,73	2,57	2,46	2,30	2,13	1,93
28	4,20	3,34	2,95	2,71	2,56	2,44	2,29	2,12	1,91
29	4,18	3,33	2,93	2,70	2,54	2,43	2,28	2,10	1,90
30	4,17	3,32	2,92	2,69	2,53	2,42	2,27	2,09	1,89
35	4,12	3,26	2,87	2,64	2,48	2,37	2,22	2,04	1,83
40	4,08	3,23	2,84	2,61	2,45	2,34	2,18	2,00	1,79
45	4,06	3,21	2,81	2,58	2,42	2,31	2,15	1,97	1,76
50	4,03	3,18	2,79	2,56	2,40	2,29	2,13	1,95	1,74

Приложение 6

Значения α-процентных пределов $t_{\alpha,k}$ в зависимости от k степеней свободы и заданного уровня значимости α для распределения Стьюдента

k \ α	10,0	5,0	2,5	2,0	1,0	0,5	0,3	0,2	0,1
1	6,314	12,706	25,452	31,821	63,657	127,3	212,2	318,3	636,6
2	2,920	4,303	6,205	6,965	9,925	14,089	18,216	22,327	31,600
3	2,353	3,182	4,177	4,541	5,841	7,453	8,891	10,214	12,922
4	2,132	2,776	3,495	3,747	4,604	5,597	6,435	7,173	8,610
5	2,015	2,571	3,163	3,365	4,032	4,773	5,376	5,893	6,869
6	1,943	2,447	2,969	3,143	3,707	4,317	4,800	5,208	5,959
7	1,895	2,365	2,841	2,998	3,499	4,029	4,442	4,785	5,408
8	1,860	2,306	2,752	2,696	3,355	3,833	4,199	4,501	5,041
9	1,833	2,262	2,685	2,821	3,250	3,690	4,024	4,297	4,781
10	1,812	2,228	2,634	2,764	3,169	3,581	3,892	4,144	4,587
12	1,782	2,179	2,560	2,681	3,055	3,428	3,706	3,930	4,318
14	1,761	2,145	2,510	2,624	2,977	3,326	3,583	3,787	4,140
16	1,746	2,120	2,473	2,583	2,921	3,252	3,494	3,686	4,015
18	1,734	2,101	2,445	2,552	2,878	3,193	8,428	3,610	3,922
20	1,725	2,086	2,423	2,528	2,845	3,153	3,376	3,552	3,849
22	1,717	2,074	2,405	2,508	2,819	3,119	3,335	3,505	3,792
24	1,711	2,064	2,391	2,492	2,797	3,092	3,302	3,467	3,745
26	1,706	2,056	2,379	2,479	2,779	3,067	3,274	3,435	3,704
28	1,701	2,048	2,369	2,467	2,763	3,047	3,250	3,408	3,674
30	1,697	2,042	2,360	2,457	2,750	3,030	3,230	3,386	3,646
∞	1,645	1,960	2,241	2,326	2,576	2,807	2,968	3,090	3,291

Приложение 7

Значения коэффициента корреляции рангов Спирмэна для двухсторонних пределов уровня значимости α

n \ α	0,20	0,10	0,05	0,02	0,01	0,002
4	0,8000	0,8000				
5	0,7000	0,8000	0,9000	0,9000		
6	0,6000	0,7714	0,8286	0,8857	0,9429	
7	0,5357	0,6786	0,7450	0,8571	0,8929	0,9643
8	0,5000	0,6190	0,7143	0,8095	0,8571	0,9286
9	0,4667	0,5833	0,6833	0,7667	0,8167	0,9000
10	0,4424	0,5515	0,6364	0,7333	0,7818	0,8667
11	0,4182	0,5273	0,6091	0,7000	0,7455	0,8364
12	0,3986	0,4965	0,5804	0,6713	0,7273	0,8182
13	0,3791	0,4780	0,5549	0,6429	0,6978	0,7912
14	0,3626	0,4593	0,5341	0,6220	0,6747	0,7670
15	0,3500	0,4429	0,5179	0,6000	0,6536	0,7464
16	0,3382	0,4265	0,5000	0,5824	0,6324	0,7265
17	0,3260	0,4118	0,4853	0,5637	0,6152	0,7083
18	0,3148	0,3994	0,4716	0,5480	0,5975	0,6904
19	0,3070	0,3895	0,4579	0,5333	0,5825	0,6737
20	0,2977	0,3789	0,4451	0,5203	0,5684	0,6586
21	0,2909	0,3688	0,4351	0,5078	0,5545	0,6455
22	0,2829	0,3597	0,4241	0,4963	0,5426	0,6318
23	0,2767	0,3518	0,4150	0,4852	0,5306	0,6186
24	0,2704	0,3435	0,4061	0,4748	0,5200	0,6070
25	0,2646	0,3362	0,3977	0,4654	0,5100	0,5962
26	0,2588	0,3299	0,3894	0,4564	0,5002	0,5856
27	0,2540	0,3236	0,3822	0,4481	0,4915	0,5757
28	0,2490	0,3175	0,3749	0,4401	0,4828	0,5660
29	0,2443	0,3113	0,3685	0,4320	0,4744	0,5567
30	0,2400	0,3059	0,3620	0,4251	0,4665	0,5479

Приложение 8

Таблица значений $e^{-\lambda}$

λ	$e^{-\lambda}$	λ	$e^{-\lambda}$	λ	$e^{-\lambda}$	λ	$e^{-\lambda}$
1	2	1	2	1	2	1	2
0,00	1,0000	0,24	0,7866	0,48	0,6188	0,72	0,4868
01	0,9900	25	0,7788	49	0,6126	73	0,4819
02	0,9802	26	0,7711	0,50	0,6065	74	0,4771
03	0,9704	27	0,7634	51	0,6005	75	0,4724
04	0,9608	28	0,7558	52	0,5945	76	0,4677
05	0,9512	29	0,7483	53	0,5886	77	0,4630
06	0,9418	0,30	0,7408	54	0,5827	78	0,4584
07	0,9324	31	0,7334	55	0,5769	$^1/_4\pi$= =0,7854	0,4559
08	0,9231	32	0,7261	56	0,5712	79	0,4538
09	0,9139	33	0,7189	57	0,5655	0,80	0,4493
0,10	0,9048	34	0,7118	58	0,5599	81	0,4449
11	0,8958	35	0,7047	59	0,5543	82	0,4404
12	0,8869	36	0,6977	0,60	0,5488	83	0,4360
13	0,8781	37	0,6907	61	0,5434	84	0,4317
14	0,8694	38	0,6839	62	0,5379	85	0,4274
15	0,8607	39	0,6771	63	0,5326	86	0,4232
16	0,8251	0,40	0,6703	64	0,5273	87	0,4190
17	0,8437	41	0,6637	65	0,5220	88	0,4148
18	0,8353	42	0,6570	66	0,5169	89	0,4107
19	0,8270	43	0,6505	67	0,5117	0,90	0,4066
0,20	0,8187	44	0,6440	68	0,5066	91	0,4025
21	0,8106	45	0,6376	69	0,5016	92	0,3985
22	0,8025	46	0,6313	0,70	0,4966	93	0,3946
23	0,7945	47	0,6250	71	0,4916	94	0,3906

Продолжение

λ	$e^{-\lambda}$	λ	$e^{-\lambda}$	λ	$e^{-\lambda}$	λ	$e^{-\lambda}$
1	2	1	2	1	2	1	2
0,95	0,3867	1,21	0,2982	1,47	0,2299	1,72	0,1791
96	0,3829	22	0,2952	48	0,2276	73	0,1773
97	0,3791	23	0,2923	49	0,2254	74	0,1755
98	0,3753	24	0,2894	1,50	0,2231	75	0,1738
99	0,3716	25	0,2865	51	0,2209	76	0,1720
1,00	0,3679	26	0,2837	52	0,2187	77	0,1703
01	0,3642	27	0,2808	53	0,2165	78	0,1683
02	0,3606	28	0,2780	54	0,2144	79	0,1670
03	0,3570	29	0,2753	55	0,2122	1,80	0,1653
04	0,3535	1,30	0,2725	56	0,2101	81	0,1637
05	0,3499	31	0,2698	57	0,2080	82	0,1620
06	0,3465	32	0,2671	$^1/_2\pi$= =1,57008	0,207	83	0,1604
07	0,3430	33	0,2645	58	0,2060	84	0,1588
08	0,3396	34	0,2618	59	0,2039	85	0,1572
09	0,3362	35	0,2592	1,60	0,2019	86	0,1557
1,10	0,3329	36	0,2567	61	0,1999	87	0,1541
11	0,3296	37	0,2541	62	0,1979	88	0,1526
12	0,3263	38	0,2561	63	0,1959	89	0,1511
13	0,3230	39	0,2491	64	0,1940	1,90	0,1496
14	0,3198	1,40	0,2466	65	0,1920	91	0,1481
15	0,3166	41	0,2441	66	0,1901	92	0,1466
16	0,3135	42	0,2417	67	0,1882	93	0,1451
17	0,3104	43	0,2393	68	0,1864	94	0,1437
18	0,3073	44	0,2369	69	0,1845	95	0,1423
19	0,3042	45	0,2346	1,70	0,1827	96	01409
1,20	0,3012	46	0,2322	71	0,1809	97	0,1395
						98	0,1381
						99	0,1367
						2,00	0,1353

Приложение 9

Соотношение между r и z' для z' значений от 0 до 5*

z'	0,00	0,01	0,02	0,03	0,04	0,05	0,06	0,07	0,08	0,09
0,0	0,0000	0,0100	0,0200	0,0300	0,0400	0,0500	0,0599	0,0699	0,0798	0,0898
0,1	0,0997	0,1096	0,1194	0,1293	0,1391	0,1489	0,1587	0,1684	0,1781	0,1878
0,2	0,1974	0,2070	0,2165	0,2260	0,2355	0,2449	0,2543	0,2636	0,2729	0,2821
0,3	0,2913	0,3004	0,3095	0,3185	0,3275	0,3364	0,3452	0,3540	0,3627	0,3714
0,4	0,3800	0,3885	0,3969	0,4053	0,4136	0,4219	0,4301	0,4382	0,4462	0,4542
0,5	0,4621	0,4700	0,4777	0,4854	0,4930	0,5005	0,5080	0,5154	0,5227	0,5299
0,6	0,5370	0,5441	0,5511	0,5581	0,5649	0,5717	0,5784	0,5850	0,5915	0,5980
0,7	0,6044	0,6107	0,6169	0,6231	0,6291	0,6352	0,6411	0,6469	0,6527	0,6584
0,8	0,6640	0,6696	0,6751	0,6805	0,6858	0,6911	0,6963	0,7014	0,7064	0,7114
0,9	0,7163	0,7211	0,7259	0,7306	0,7352	0,7398	0,7443	0,7487	0,7531	0,7574
1,0	0,7616	0,7658	0,7699	0,7739	0,7779	0,7818	0,7857	0,7895	0,7932	0,7969
1,1	0,8005	0,8041	0,8076	0,8110	0,8144	0,8178	0,8210	0,8243	0,8275	0,8306
1,2	0,8337	0,8367	0,8397	0,8426	0,8455	0,8483	0,8511	0,8538	0,8565	0,8591
1,3	0,8617	0,8643	0,8668	0,8693	0,8717	0,8741	0,8764	0,8787	0,8810	0,8832
1,4	0,8854	0,8875	0,8896	0,8917	0,8937	0,8957	0,8977	0,8996	0,9015	0,9033
1,5	0,9052	0,9069	0,9087	0,9104	0,9121	0,9138	0,9154	0,9170	0,9186	0,9202
1,6	0,9217	0,9232	0,9246	0,9261	0,9275	0,9289	0,9302	0,9316	0,9329	0,9342
1,7	0,9354	0,9367	0,9379	0,9391	0,9402	0,9414	0,9425	0,9436	0,9447	0,9458
1,8	0,9468	0,9478	0,9498	0,9488	0,9508	0,9518	0,9527	0,9536	0,9545	0,9554
1,9	0,9562	0,9571	0,9579	0,9587	0,9595	0,9603	0,9611	0,9619	0,9626	0,9633
2,0	0,9640	0,9647	0,9654	0,9661	0,9668	0,9674	0,9680	0,9687	0,9693	0,9699
2,1	0,9705	0,9710	0,9716	0,9722	0,9727	0,9732	0,9738	0,9743	0,9748	0,9753
2,2	0,9757	0,9762	0,9767	0,9771	0,9776	0,9780	0,9785	0,9789	0,9793	0,9797
2,3	0,9801	0,9805	0,9809	0,9812	0,9816	0,9820	0,9823	0,9827	0,9830	0,9834
2,4	0,9837	0,9840	0,9843	0,9846	0,9849	0,9852	0,9855	0,9858	0,9861	0,9863
2,5	0,9866	0,9869	0,9871	0,9874	0,9876	0,9879	0,9881	0,9884	0,9886	0,9888
2,6	0,9890	0,9892	0,9895	0,9897	0,9899	0,9901	0,9903	0,9905	0,9906	0,9908
2,7	0,9910	0,9912	0,9914	0,9915	0,9917	0,9919	0,9920	0,9922	0,9923	0,9925
2,8	0,9926	0,9928	0,9929	0,9931	0,9932	0,9933	0,9935	0,9936	0,9937	0,9938
2,9	0,9940	0,9941	0,9942	0,9943	0,9944	0,9945	0,9946	0,9947	0,9949	0,9950
3,0	0,9951									
4,0	0,9993									
5,0	0,9999									

* Цифры таблицы являются значениями коэффициента корреляции r, соответствующими значениям z', указанным слева и сверху таблицы.

Рекомендуемая литература

1. Общая теория статистики: Учебник / Под ред. О. Э. Башиной, А. А. Спирина. – 5-е изд., доп. и перераб. – М.: Финансы и статистика, 2003.

2. *Ефимова М. Р. и др.* Общая теория статистики: Учебник / *М. Р. Ефимова, Е. В. Петрова, В. Н. Румянцев* и др. – М.: Инфра-М, 2004.

3. *Елисеева И. И., Юзбашев М. М.* Общая теория статистики: Учебник. – 5-е изд., перераб. и доп. – М.: Финансы и статистика, 2004.

4. Статистика: Учеб. пособие / Под ред. М. Р. Ефимовой. – М.: Инфра-М, 2000.

Оглавление

Учебное издание

Ефимова Марина Романовна
Ганченко Ольга Ивановна
Петрова Екатерина Валериановна

ПРАКТИКУМ
ПО ОБЩЕЙ ТЕОРИИ СТАТИСТИКИ

Заведующая редакцией *Л. А. Табакова*
Редактор *Е. В. Стадниченко*
Художественный редактор *Ю. И. Артюхов*
Технический редактор *В. Ю. Фотиева*
Корректоры *Н. Н. Зубенко, Н. П. Сперанская*
Компьютерная верстка *Е. А. Федорова*
Оформление художника *Н. М. Биксентеева*

ИБ № 4452

Подписано в печать 27.04.2004
Формат 60 × 88/16. Гарнитура «Таймс». Печать офсетная.
Усл. п. л. 20,58. Уч.-изд. л. 18,37. Тираж 3000 экз. Заказ 1622. «С» 150

Издательство «Финансы и статистика»
101000, Москва, ул. Покровка, 7
Телефон (095) 925-35-02, факс (095) 925-09-57
E-mail: mail@finstat.ru http:/www.finstat.ru

ГП Псковской области «Великолукская городская типография»
Комитета по средствам массовой информации
182100, Великие Луки, ул. Полиграфистов, 78/12
Тел./факс: (811-53) 3-62-95
E-mail: VTL@MART.RU