Министерство образования и науки Российской Федерации

Орский гуманитарно-технологический институт (филиал)
Федерального государственного бюджетного образовательного учреждения
высшего профессионального образования
«Оренбургский государственный университет»

В. В. Свечникова

СТАТИСТИКА

Учебное пособие

ИЗДАТЕЛЬСТВО ОРСКОГО ГУМАНИТАРНО-ТЕХНОЛОГИЧЕСКОГО ИНСТИТУТА

Орск 2012

УДК 311
ББК 60.6
С24

Научный редактор

Ермакова Ж. А., *доктор экономических наук, профессор,
директор НИИ региональной экономики
ФГБОУ ВПО «Оренбургский государственный университет»*

Рецензенты:

Чмышенко Е. Г., *доктор экономических наук, профессор,
заведующий кафедрой экономики и управления на предприятии
ФГБОУ ВПО «Оренбургский государственный университет»;*

Корабейников И. Н., *кандидат экономических наук,
заведующий отделом региональной конкурентоспособности
и инвестиционного развития НИИ региональной экономики
ФГБОУ ВПО «Оренбургский государственный университет»*

С24 Свечникова, В. В. Статистика : учебное пособие / В. В. Свечникова. – Орск : Издательство Орского гуманитарно-технологического института (филиала) ОГУ, 2012. – 91 с. – ISBN 978-5-8424-0622-7.

В учебном пособии рассмотрены основные темы дисциплины «Статистика», посвященные вопросам сбора, обработки и анализа статистических данных. По всем темам приведены типовые задачи и их решение.

Рекомендуется студентам, обучающимся по направлению 080100 – Экономика, а также будет полезно всем интересующимся вопросами статистики.

ISBN 978-5-8424-0622-7

СОДЕРЖАНИЕ

ВВЕДЕНИЕ

Неотъемлемой частью экономического образования являются статистическая грамотность, умение пользоваться статистическими методами исследования, позволяющими обнаружить различные статистические закономерности.

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

В данном пособии рассмотрены основные темы дисциплины «Статистика», приведено решение типовых задач, построенных на условных данных. В приложении приведены основные математико-статистические таблицы, использование которых предусмотрено в ряде задач.

ГЛАВА 1. ГРУППИРОВКА СТАТИСТИЧЕСКИХ ДАННЫХ

1.1. Сущность группировки и ее виды

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

Группировка – расчленение единиц изучаемой совокупности на однородные группы по определенным, существенным для них признакам.

Признак, на основе которого производится подразделение единиц наблюдения на группы, называется ***группировочным признаком,*** или ***основанием группировки.***

Группировки классифицируются следующим образом:

I. В соответствии с задачами группировки выделяются следующие виды группировок:

1) **Типологическая** – расчленение разнородной совокупности на отдельные качественно однородные группы и выявление на этой основе экономических типов явлений.

Пример: Группировка хозяйствующих субъектов по формам собственности – группы предприятий государственной собственности, федеральной, муниципальной и частной собственности.

2) **Структурная** – группировка, предназначенная для изучения состава однородной совокупности по какому-либо варьирующему признаку.

Пример: Группировка населения по размеру среднедушевого дохода, группировка коммерческих банков по величине уставного капитала и др.

3) **Аналитическая** – группировка, выявляющая взаимосвязи между изучаемыми явлениями и их признаками.

Пример: Группировка рабочих по квалификации (тарифному разряду) с указанием их средней месячной заработной платы.

II. По количеству группировочных признаков группировки бывают:

1) **простая** – группировка, в которой группы образованы по одному признаку;

2) **сложная** – группировка, в которой расчленение совокупности на группы производится по двум и более признакам.

III. В зависимости от вида группировочного признака группировки бывают:

1) **группировки по атрибутивному признаку** – группы образуются по признаку, отражающему состояние единицы совокупности (пол человека, семейное положение, организационно-правовая форма предприятия и др.);

2) **группировки по количественному признаку** – группы образуются по признаку, имеющему числовое выражение (возраст человека, доход семьи и др.).

1.2. Выполнение группировки по количественному и атрибутивному признакам

После того как определено основание группировки, решается вопрос о количестве групп, на которые разбивается исследуемая совокупность.

Число групп зависит от задач исследования, вида показателя, положенного в основание группировки, степени вариации признака и численности совокупности.

Если группировка строится по ***атрибутивному признаку***, то групп будет столько, сколько видов состояний у данного признака.

При группировке по ***количественному признаку*** число групп определяется в зависимости от характера изменения признака и задач исследования:

1) если количественный признак меняется прерывно (дискретно), то число групп должно соответствовать количеству значений признака;

2) при непрерывном изменении признака число групп определяется по формуле ***Стерджесса*** (1.2.1):

$$n = 1 + 3{,}322 \cdot \lg N, \qquad (1.2.1)$$

где N – число единиц совокупности.

После установления числа групп необходимо определить интервалы группировки, которые могут быть равные и неравные.

Равные интервалы используются, если нужно охарактеризовать количественные различия в величине признака внутри групп одинакового качества.

Величина равного интервала определяется по формуле (1.2.2):

$$h = \frac{X_{max} - X_{min}}{n}, \qquad (1.2.2)$$

где X_{max}, X_{min} – соответственно наибольшее и наименьшее значения признака в изучаемой совокупности;

n – число групп.

Неравные интервалы используются в случае значительного размаха вариации признака и неравномерного варьирования признака в совокупности. В интервальных рядах с неравными интервалами определяется плотность распределения, которая может быть:

1) ***абсолютная***, если рассчитана на основе частот (f_i) по формуле (1.2.3):

$$m_i^{абс} = \frac{f_i}{h_i}, \qquad (1.2.3)$$

где h_i – величина интервала;

2) ***относительная***, рассчитанная на основе частостей (w_i) по формуле (1.2.4):

$$m_i^{отн} = \frac{w_i}{h_i}. \qquad (1.2.4)$$

Плотность распределения показывает заполненность различных интервалов, то есть сколько единиц совокупности приходится на каждую единицу интервала.

Интервалы группировок могут быть закрытыми и открытыми.

Закрытый интервал – интервал, у которого имеются и верхняя, и нижняя границы.

Открытый интервал – интервал, у которого указана только одна граница (например, группы строительных фирм по объему работ (тыс. руб.): до 5 000, 5 000-10 000, свыше 10 000).

На практике возможно, что группировки, построенные за один и тот же период времени, но для разных регионов или, наоборот, для одного региона, но за два разных периода времени, окажутся несопоставимыми из-за различного числа выделенных групп или границ интервалов.

В этом случае необходимо перегруппировать уже сгруппированные данные или провести вторичную группировку.

Вторичная группировка – операция по образованию новых групп на основе ранее осуществленной группировки.

Наиболее распространенным способом проведения вторичной группировки является изменение первоначальных интервалов. Второй способ получил название долевой перегруппировки и состоит в образовании новых групп на основе закрепления за каждой группой определенной доли единиц совокупности.

1.3. Решение типовых задач

Задача 1. Известны данные о количестве членов семьи 50 работников цеха предприятия:

2	5	5	6	3	2	5	6	5	6
6	6	4	3	3	5	7	3	5	5
5	4	5	6	4	4	4	4	7	4
4	3	5	3	7	4	6	6	4	7
4	4	6	7	6	3	3	5	8	5

Произвести группировку работников цеха по количеству членов их семей.

Решение:

В соответствии с условием задачи в основу группировки положен количественный дискретный признак, следовательно, число групп соответствует числу значений признака.

Для проведения группировки рабочих по количеству членов их семей выписываются все значения признака в порядке возрастания, и затем подсчитывается число семей рабочих в каждой группе (табл. 1.3.1).

Таблица 1.3.1

Группировка семей рабочих по количеству членов семьи

Количество членов семьи	**Число семей работников**
2	2
3	8
4	12
5	12
6	10
7	5
8	1
Итого	50

Задача 2. Известны следующие данные об урожайности озимой пшеницы в 40 обследованных фермерских хозяйствах, ц/га:

22,8 27,0 20,4 27,1 18,2 16,3 22,0 24,3 24,8 33,0

27,3 23,1 21,1 22,6 14,0 29,5 22,9 28,5 15,1 19,5

28,1 25,1 26,7 28,4 29,6 19,9 27,0 25,3 23,9 21,5

23,7 18,0 31,0 19,8 26,0 23,5 20,2 25,1 25,8 33,8

Произвести группировку данных, выделив 6 групп хозяйств по величине урожайности с равными интервалами.

Решение:

В основу группировки положен количественный непрерывный признак, следовательно, для проведения группировки данных необходимо знать число групп и величину интервала. Так как число групп известно, определяем величину интервала по формуле (1.2.2):

$$h = \frac{33{,}8 - 14{,}0}{6} = 3{,}3.$$

Следовательно, первая группа фермерских хозяйств по величине урожайности находится в интервале 14,0-17,3, вторая – 17,3-20,6 и т. д.

При такой записи непрерывного признака, когда одна и та же величина встречается дважды (как верхняя граница одного интервала и как нижняя граница другого интервала), единица, обладающая этим значением, относится к той группе, где эта величина выступает в роли верхней границы.

Результат группировки фермерских хозяйств по величине урожайности представлен в таблице 1.3.2.

Таблица 1.3.2

Группировка фермерских хозяйств по величине урожайности пшеницы

Урожайность пшеницы, ц/га	**Число хозяйств**
14,0-17,3	3
17,3-20,6	7
20,6-23,9	10
23,9-27,2	11
27,2-30,5	6
30,5-33,8	3
Итого	40

Задача 3. Распределение акционеров двух районов области по размеру выплаты дивидендов на одну акцию представлено следующими данными:

Таблица 1.3.3

Первый район		**Второй район**	
Размер дивидендов, руб.	Число акционерных обществ, %	Размер дивидендов, руб.	Число акционерных обществ, %
100-400	18	100-600	10
400-800	12	600-1200	20

800-1200	40	1200-2000	40
1200-1600	25	2000-3000	30
1600-2000	5	-	-
Итого	100	Итого	100

Произвести перегруппировку данных первого района в соответствии с группировкой второго района.

***Решение*:**

Приведенные данные не позволяют сравнить распределение акций двух районов по размеру дивидендов на одну акцию, так как в этих районах имеется различное число групп акционеров и различны величины интервалов.

По условию задачи необходимо по первому району произвести вторичную группировку, образовав такое же число групп и с теми же интервалами, как во втором районе.

Для решения воспользуемся наиболее распространенным способом вторичной группировки – способом изменения первоначальных интервалов.

В первую новую группу [100-600] войдет полностью первая группа акционерных обществ [100-400] и часть второй группы. Чтобы образовать группу по размеру дивидендов от 100 до 600 руб., необходимо от интервала второй группы взять 200 руб. Величина интервала второй группы составляет 400 руб. Следовательно, необходимо взять от нее 1/2 (200 : 400) часть. Аналогичную часть во вновь образуемую новую группу надо взять и от числа акционерных обществ, то есть $12 \cdot 1/2 = 6$ %. Тогда число акционерных обществ в первой группе будет $18 + 6 = 24$ %.

Вторую новую группу [600-1200] образуют общества второй группы за вычетом отнесенных к первой ($12 - 6 = 6$ %) и все общества третьей группы. Следовательно, число акционерных обществ данной группы составит $6 + 40 = 46$ %.

Третью новую группу [1200-2000] составят акционерные общества четвертой [1200-1600] и пятой [1600-2000] прежних групп: $25 + 5 = 30$%.

Четвертую новую группу образуют лишь акционерные общества второго района, акционерных обществ с размером дивидендов на одну акцию от 2000 до 3000 руб. в первом районе нет.

Результаты вторичной группировки представлены в таблице 1.3.4.

Таблица 1.3.4

Вторичная группировка акционеров по размеру дивидендов на одну акцию

Группы акционеров по размеру дивидендов на одну акцию, руб.	**Удельный вес акционеров группы, %**		**Расчет числа акционеров для первого района**
	Первый район	**второй район**	
100-600	24	10	$18 + \frac{1}{2} \times 12$
600-1200	46	20	$\frac{1}{2} \times 12 + 40$
1200-2000	30	40	25 + 5
2000-3000	-	30	-
Итого	100	100	100

ГЛАВА 2. АБСОЛЮТНЫЕ, ОТНОСИТЕЛЬНЫЕ И СРЕДНИЕ ВЕЛИЧИНЫ

Все используемые в статистической практике показатели по форме выражения классифицируются на абсолютные, относительные и средние.

2.1. Абсолютные величины

Абсолютные показатели отражают физические размеры изучаемых статистикой процессов и явлений, а именно: их массу, площадь, временные характеристики и др.

Абсолютные показатели всегда являются именованными величинами, то есть имеют какую-либо единицу измерения. В зависимости от сущности исследуемого социально-экономического явления абсолютные статистические величины выражаются в натуральных, стоимостных или трудовых единицах измерения.

Натуральные единицы измерения применяются в тех случаях, когда единица измерения соответствует потребительским свойствам продукта. Например, производство тканей оценивается в квадратных метрах, автомобилей – в штуках и т. д.

Для определения объема выпуска однородной, но не одинаковой продукции, обладающей общностью основного потребительского свойства, используются ***условно-натуральные единицы.***

В этом случае одна из разновидностей продукции принимается в качестве единого измерителя, а другие приводятся к этому измерителю с помощью коэффициента пересчета, определяемого по формуле (2.1.1):

$$K_{\text{пер}_i} = \frac{\text{потребительское значение i-й продукции}}{\text{потребительское значение условной продукции}}. \quad (2.1.1)$$

Например, различные виды органического топлива переводятся в условное топливо с теплотой сгорания 7000 ккал/кг (29,3 мДж/кг), мыло разных сортов – в условное мыло с 40 % содержанием жирных кислот и т. д.

Тогда объем выпуска продукции в условно-натуральных единицах определяется по формуле (2.1.2):

$$Q_{\text{усл-нат}} = \sum_{i=1}^{n} g_i \times K_{\text{пер}_i}, \quad (2.1.2)$$

где g_i – количество i-й продукции в натуральном выражении.

Стоимостные показатели позволяют соизмерить в денежной форме величины, которые несоизмеримы в натуральной форме. Для получения общего объема продукции в денежном выражении количество единиц каждого вида продукции в натуральном выражении умножается на цену соответствующего вида, а затем полученные произведения суммируют по всем видам.

К ***трудовым единицам измерения***, позволяющим учитывать как общие затраты труда на предприятии, так и трудоемкость отдельных операций технологического процесса, относятся человеко-дни и человеко-часы.

Значение абсолютных показателей в статистике, бесспорно, велико. Однако в научном анализе для раскрытия явления выявления определенных закономерностей прибегают к сопоставлению абсолютных показателей друг с другом и определению на основе этих сопоставлений относительных и средних величин.

2.2. Относительные величины

Относительный показатель – это результат сопоставления двух абсолютных показателей.

При расчете относительного показателя величина, с которой производится сравнение, называется ***базой сравнения***, или ***основанием.***

Относительная величина может быть выражена в форме коэффициента, процента или промилле. Причем выбор формы выражения относительной величины определяется размерностью сравниваемых величин и стремлением придать относительной величине наибольшую выразительность.

Различают относительные величины динамики, планового задания, выполнения плана, структуры, сравнения, интенсивности и координации.

Относительная величина динамики представляет собой отношение уровня исследуемого процесса или явления за данный (отчетный) период времени ($у_1$) к уровню этого же процесса или явления в прошлом ($у_0$) (формула 2.2.1):

$$ОВ_{дин} = \frac{у_1}{у_0}. \qquad (2.2.1)$$

Относительная величина динамики характеризует изменение явления во времени и показывает, во сколько раз изменился уровень показателя по сравнению с каким-либо предшествующим периодом.

Относительная величина планового задания показывает, как изменится величина показателя по плану ($у_{пл}$) в сравнении с его уровнем в предшествующем периоде ($у_0$) (формула 2.2.2):

$$ОВ_{пл} = \frac{у_{пл}}{у_0}. \quad (2.2.2)$$

Относительная величина выполнения плана представляет собой отношение фактической величины показателя отчетного периода ($у_1$) к запланированной на тот же период его величине ($у_{пл}$) (формула 2.2.3):

$$ОВ_{вып.пл} = \frac{у_1}{у_{пл}}. \quad (2.2.3)$$

Относительные величины планового задания, выполнения плана и динамики связаны между собой соотношением, представленным формулой (2.2.4):

$$ОВ_{дин} = ОВ_{пл} \times ОВ_{вып.пл} = \frac{у_{пл}}{у_0} \times \frac{у_1}{у_{пл}} = \frac{у_1}{у_0}. \quad (2.2.4)$$

Относительная величина структуры характеризует долю отдельных частей в общем объеме совокупности и выражается в долях единицы или в процентах. Например, в составе промышленно-производственного персонала выделяются следующие категории работников: рабочие, руководители, специалисты, служащие и младший обслуживающий персонал (МОП). Анализ показателей удельного веса каждой категории в составе промышленно-производственного персонала позволяет сопоставлять состав работников по категориям на различных предприятиях отрасли, в различных отраслях и т.д.

Относительная величина координации устанавливает соотношение между частями одного целого. К таким величинам относятся, например, показатели, характеризующие соотношение между численностью городского и сельского населения, численностью мужчин и женщин, между величиной собственного и заемного капиталов.

Относительная величина наглядности (сравнения) отражает результат сопоставления одноименных показателей, относящихся к одному и тому же периоду (или моменту времени), но к разным объектам или территориям. Например, сравнение размера основных фондов пищевой промышленности двух регионов по состоянию на 1 января 2006 г. или сравнение годовой производительности труда по двум предприятиям.

Относительная величина интенсивности характеризует степень распространения изучаемого процесса или явления в присущей ему среде. Примерами такого рода показателей могут служить коэффициент рождаемости (число родившихся в расчете на 1000 человек населения), плотность населения (количество людей, приходящихся на 1 $км^2$ территории) и др.

Относительные величины интенсивности, в отличие от других видов относительных величин, являются именованными числами.

Одно из главных требований, которое предъявляется при исчислении относительных величин, заключается в необходимости обеспечения сопоставимости сравниваемой величины и величины, принятой за базу сравнения.

2.3. Виды средних величин и способы их вычисления

Средняя величина – это обобщающая характеристика изучаемого признака в исследуемой совокупности.

Объективность и типичность статистической средней обеспечивается при выполнении следующих условий:

1) средняя должна вычисляться для качественно однородной совокупности;

2) 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.

Существуют две категории средних величин:

1) ***степенные средние***;

2) ***структурные средние***.

Степенные средние представлены средней арифметической, средней гармонической, средней квадратической, средней геометрической и средней кубической.

Средняя величина признака в общем виде определяется по формуле (2.3.1):

$$\overline{X} = \sqrt[k]{\frac{\Sigma x_i^k f_i}{\Sigma f_i}}. \qquad (2.3.1)$$

В зависимости от степени k получаются различные виды средних величин, формулы которых представлены в таблице 2.3.1.

Таблица 2.3.1

Формулы различных видов степенных средних величин

k	Наименование средней	Формула средней	
		Простая	Взвешенная
- 1	гармоническая	$\overline{X}_{гарм} = \frac{n}{\Sigma \frac{1}{x_i}}$	$\overline{X}_{гарм} = \frac{\Sigma M_i}{\Sigma \frac{M_i}{x_i}}$
0	геометрическая	$\overline{X}_{геом} = \sqrt[n]{x_1 x_2 \ldots x_n}$	$\overline{X}_{геом} = \sqrt[\Sigma f]{x_1^{f_1} x_2^{f_2} \ldots x_n^{f_n}}$
1	арифметическая	$\overline{X}_{арифм} = \frac{\Sigma x_i}{n}$	$\overline{X}_{арифм} = \frac{\Sigma x_i f_i}{\Sigma f_i}$
2	квадратическая	$\overline{X}_{квадр} = \sqrt{\frac{\Sigma x_i^2}{n}}$	$\overline{X}_{квадр} = \sqrt{\frac{\Sigma x_i^2 f_i}{\Sigma f_i}}$
3	кубическая	$\overline{X}_{куб} = \sqrt[3]{\frac{\Sigma x_i^3}{n}}$	$\overline{X}_{куб} = \sqrt[3]{\frac{\Sigma x_i^3 f_i}{\Sigma f_i}}$

В таблице 2.3.1 использованы следующие обозначения:

x_i – индивидуальные значения признака совокупности;

f_i – частота повторения индивидуальных значений признака;

n – число единиц совокупности;

M_i – произведение индивидуальных значений признака на количество единиц, обладающих этим значением.

Средняя арифметическая простая применяется, когда каждое индивидуальное значение признака встречается только один или одинаковое число раз.

Если отдельные значения исследуемой совокупности встречаются не один, а много, причем неодинаковое число раз, то рассчитывают ***среднюю арифметическую взвешенную***.

В практике расчетов есть ситуации, когда данные о частоте (весах) признака отсутствуют, но известны варианты признака (х) и произведение значений этих вариантов на количество единиц, обладающих этим значением (x·f). В этих случаях среднее значение признака необходимо рассчитывать по формуле ***средней гармонической***.

Применение средней арифметической или средней гармонической определяется наличием исходных данных.

Средняя геометрическая используется для анализа динамики явлений и определения среднего коэффициента роста, а также для определения равноудаленной величины от максимального и минимального значений признака.

В ряде случаев в экономической практике возникает потребность расчета среднего размера признака, выраженного в квадратных или кубических единицах измерения. Тогда применяется ***средняя квадратическая*** (например, для определения среднего диаметра колес, труб, средней величины стороны n квадратных участков) и ***средняя кубическая*** (например, для определения средней длины стороны n кубов).

Степенные средние разных видов, исчисленные по одной и той же совокупности, имеют различные количественные значения, между которыми существует следующее соотношение (формула (2.3.2)):

$$X_{\text{гарм}} \le X_{\text{геом}} \le X_{\text{арифм}} \le X_{\text{квадр}} \le X_{\text{куб}} \qquad (2.3.2)$$

Это свойство степенных средних возрастать с повышением показателя степени определяющей функции (формула (2.3.1)) называется ***мажорантностью средних***.

Структурные средние представлены модой и медианой и применяются для изучения внутреннего строения и структуры рядов распределения значений признака (см. главу 3).

2.4. Решение типовых задач

Задача 1. Расход топлива на производственные нужды предприятия характеризуется в отчетном периоде следующими данными:

Таблица 2.4.1

Вид топлива	**Единица измерения**	**Расход топлива**	
		По плану	**Фактически**

Мазут топочный	т	500	520
Уголь	т	320	300
Газ природный	тыс. $м^3$	650	690

Коэффициенты перевода в условное топливо составили: мазут – 1,37 т; уголь – 0,9 т; газ – 1,2 тыс. $м^3$.

Определить общее потребление условного топлива по плану и фактически и процент выполнения плана по общему расходу топлива.

Решение:

Для определения общего потребления топлива используется условно-натуральный метод (формула (2.1.2)). Расходы по плану и фактически исчисляются в единицах условного топлива.

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

$$у_1 = 520 \cdot 1{,}37 + 300 \cdot 0{,}9 + 690 \cdot 1{,}2 = 1810{,}4$$

Процент выполнения плана по общему расходу топлива составит:

$$В_{пл} = \frac{у_1}{у_{пл}} \times 100 = \frac{1810{,}4}{1753} \times 100 = 103{,}27\%.$$

Задача 2. Планом предусматривалось повышение выпуска продукции на 5%, фактически произведено на 10,25% больше, чем в базисном периоде. Определить процент выполнения плана по выпуску продукции.

Решение:

Так как по плану предусматривалось увеличение выпуска продукции на 5%, следовательно, относительная величина планового задания составит:

$$ОВ_{пл} = \frac{у_{пл}}{у_0} = 1{,}05 \text{ или } 105{,}0\%.$$

По условию задачи фактически произведено на 10,25% больше, чем в базисном периоде, следовательно, относительная величина динамики составит:

$$ОВ_{дин} = \frac{у_1}{у_0} = 1{,}1025.$$

Относительная величина выполнения плана определяется из равенства (2.2.4):

$$ОВ_{дин} = ОВ_{пл} \times ОВ_{вып.пл.},$$

$$OB_{\text{вып.пл.}} = \frac{OB_{\text{дин}}}{OB_{\text{пл}}}.$$

Следовательно,

$$OB_{\text{вып.пл.}} = \frac{1{,}1025}{1{,}05} = 1{,}05.$$

Таким образом, степень выполнения плана составила 105% или план перевыполнен на 5% (105%-100%).

Задача 3. На начало 2006 г. в России численность мужчин составляла 118,7 млн чел., а численность женщин – 136,8 млн чел. Определить относительную величину координации.

Решение:

Относительная величина координации характеризует соотношение между частями одного целого.

Найдем соотношение между численностью мужчин и женщин:

$$OB_{\text{коорд}} = \frac{118{,}7}{136{,}8} \times 1000 = 867{,}7.$$

Следовательно, на 1000 женщин приходится 868 мужчин.

Задача 4. Определить среднюю месячную заработную плату работников предприятия за сентябрь и октябрь.

Таблица 2.4.2

№ цеха	Сентябрь		Октябрь	
	Численность работников	**Средняя месячная заработная плата, руб.**	**Средняя месячная заработная плата, руб.**	**Фонд заработной платы, тыс. руб.**
1	140	3 560	3 600	486
2	200	3 600	3 580	751,8
3	260	3 330	3 340	835

Решение:

Для определения средней месячной заработной платы работников предприятия за сентябрь по условию задачи известны:

1) частота признака (f_i) – количество работников;

2) индивидуальные значения признака (x_i) – средняя месячная заработная плата в подразделении.

Тогда для определения средней величины признака применим формулу ***средней арифметической взвешенной***:

$$\overline{X}_{арифм} = \frac{\Sigma x_i f_i}{\Sigma f_i} = \frac{3560 \times 140 + 3600 \times 200 + 3330 \times 260}{140 + 200 + 260} = \frac{2084200}{600} = 3473{,}7 \text{ руб.}$$

За октябрь по условию задачи известны следующие данные:

1) индивидуальные значения признака (x_i) – средняя месячная заработная плата в подразделении;

2) произведение значений признака на количество единиц, обладающих этим значением ($M_i = x_i \times f_i$), – фонд заработной платы работников предприятия.

Таким образом, применить формулу средней арифметической взвешенной невозможно, так как неизвестна частота признака.

В этом случае логические рассуждения остаются те же: для определения средней месячной заработной платы работников необходимо фонд заработной платы за месяц (ΣM_i) по предприятию разделить на общую численность работников предприятия ($\Sigma Ч_i$).

В свою очередь, численность работников в цехе определяется как отношение фонда заработной платы за месяц в данном цехе к средней заработной плате работников цеха:

$$Ч_i = \frac{M_i}{\overline{X}_i}.$$

Тогда расчет средней заработной платы производится по формуле средней гармонической взвешенной:

$$\overline{X}_{гарм} = \frac{\Sigma M_i}{\Sigma Ч_i} = \frac{\Sigma M_i}{\Sigma \frac{M_i}{\overline{X}_i}} = \frac{486000 + 751800 + 835000}{\frac{486000}{3600} + \frac{751800}{3580} + \frac{835000}{3340}} = 3483{,}7 \text{ руб.}$$

Задача 5. Подача жидкого топлива для технологического процесса осуществляется в цехе тремя трубопроводами с диаметрами 2 см, 5 см 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 новой трубы.

Решение:

Средний диаметр труб определяется по формуле:

$$\overline{Д} = 2\,\overline{r},$$

где $\overline{r}$ – средний радиус труб (r_1=1, r_2=2,5, r_3=3).

Так как необходимо определить средний размер признака, выраженного в квадратных единицах измерения, применим формулу средней квадратической простой.

$$\bar{r} = \sqrt{\frac{\Sigma r_i^2}{n}} = \sqrt{\frac{1^2 + 2{,}5^2 + 3^2}{3}} = 1{,}73 \text{ см.}$$

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

ГЛАВА 3. ВАРИАЦИОННЫЙ РЯД И ЕГО ОСНОВНЫЕ ХАРАКТЕРИСТИКИ

3.1. Понятие, виды и основные элементы вариационных рядов

В результате группировки единиц совокупности по величине какого-либо варьирующего признака получают ряды распределения. В зависимости от признака, положенного в основу ряда распределения, различают атрибутивные и вариационные ряды распределения.

Ряды распределения, построенные по качественным признакам, называются ***атрибутивными***.

Ряды распределения, построенные по количественному признаку, называются ***вариационными***. Любой вариационный ряд состоит из двух элементов: вариантов и частот. ***Вариантами*** называются отдельные значения признака, которые он принимает в вариационном ряду. ***Частотами*** называются численности отдельных вариантов, то есть числа, которые показывают, как часто встречаются те или иные варианты в ряду распределения. Частоты, выраженные в долях единицы или в процентах к итогу, называются ***частостями*** и определяются по формуле (3.1.1):

$$w_i = \frac{f_i}{\Sigma f_i}, \qquad (3.1.1)$$

где f_i – частота признака.

В зависимости от характера вариации признака различают дискретные (прерывные) и интервальные (непрерывные) ряды. Способы построения вариационных рядов для этих видов признаков рассматривались в главе 1.

Для анализа вариационных рядов используются три группы показателей:

1) показатели центра распределения;

2) показатели степени вариации;

3) показатели формы распределения.

3.2. Показатели центра распределения

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

- ***Средняя арифметическая***:

1) для дискретного ряда распределения определяется по формуле (3.2.1):

$$\overline{X} = \frac{\Sigma x_i f_i}{\Sigma f_i}, \tag{3.2.1}$$

где x_i – варианты значений признака;

f_i – частота повторения данного варианта.

2) для интервального ряда распределения исчисляется по формуле (3.2.2):

$$\overline{X} = \frac{\Sigma x_i^{'} f_i}{\Sigma f_i}, \tag{3.2.2}$$

где $x_i^{'}$ – середина соответствующего интервала значения признака.

- ***Мода*** (Мо) – значение признака, наиболее часто встречающееся в совокупности.

Порядок расчета моды определяется видом вариационного ряда:

1) для ***дискретных рядов*** модой является вариант, имеющий наибольшую частоту;

2) в ***интервальных вариационных рядах*** прежде всего определяют интервал, в котором находится мода (модальный интервал).

В вариационном ряду с равными интервалами модальный интервал определяется по наибольшей частоте, в рядах с неравными интервалами – по наибольшей плотности распределения.

После определения модального интервала мода определяется по формуле (3.2.3):

$$Mo = X_{мо} + h\frac{f_{мо} - f_{мо-1}}{(f_{мо} - f_{мо-1}) + (f_{мо} - f_{мо+1})}, \tag{3.2.3}$$

где $X_{мо}$ – нижняя граница модального интервала;

$f_{мо}, f_{мо-1}, f_{мо+1}$ – соответственно частота (плотность распределения) модального, предмодального и послемодального интервалов;

h – величина модального интервала.

• ***Медиана*** (Ме) – значение признака у средней единицы ранжированного ряда. Медиана делит ряд на две равные (по числу единиц) части – со значениями признака меньше медианы и со значениями признака больше медианы.

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

Порядок расчета медианы также определяется видом вариационного ряда.

1) ***Дискретный ряд***

– если ряд имеет нечетное число членов, то медианой будет вариант, находящийся в середине ранжированного ряда;

– если ряд состоит из четного числа членов, то медианой будет средняя арифметическая из двух значений признака, расположенных в середине ряда.

2) ***Интервальный ряд***

В интервальном ряду распределения сначала определяют интервал, в котором находится медиана (медианный интервал).

Медианный интервал – интервал, в котором сумма накопленных частот впервые превысит полусумму всех частот ряда.

Численное значение медианы определяется по формуле (3.2.4):

$$M_e = x_{ме} + h\frac{\frac{1}{2}\Sigma f - S_{ме-1}}{f_{ме}}, \qquad (3.2.4)$$

где $x_{ме}$ – нижняя граница медианного интервала;

h – величина медианного интервала;

$S_{ме-1}$ – накопленная частота интервала, предшествующая медианному;

$f_{ме}$ – частота медианного интервала.

Аналогично с нахождением медианы в вариационных рядах можно определить значение признака у любой по порядку единицы ранжированного ряда. Например, можно определить значение признака у единиц, делящих ряд на четыре равные части, десять или сто частей. Эти величины называются квартили, децили и перцентили.

Рассмотрим порядок расчета квартилей.

Квартили представляют собой значение признака, делящее ранжированную совокупность на четыре равные части. Различают следующие виды квартилей:

1) ***нижний (первый) квартиль*** (Q_1) – значение признака у единицы ранжированного ряда, делящее совокупность в соотношении 1/4 к 3/4 и определяемое по формуле (3.2.5):

$$Q_1 = X_{Q_1} + h\frac{\frac{1}{4}\sum f - S_{Q_1 - 1}}{f_{Q_1}}, \qquad (3.2.5)$$

2) ***верхний (третий) квартиль*** (Q_3) – значение признака у единицы ранжированного ряда, делящее совокупность в соотношении 3/4 к 1/4 и определяемое по формуле (3.2.6):

$$Q_3 = X_{Q_3} + h\frac{\frac{3}{4}\sum f - S_{Q_3 - 1}}{f_{Q_3}}, \qquad (3.2.6)$$

3) ***средний (второй) квартиль*** является медианой.

3.3. Показатели степени вариации признака

Для характеристики размера вариации признака используются абсолютные и относительные показатели.

1. Абсолютные показатели вариации:

• ***размах вариации*** – разность между максимальным (X_{max}) и минимальным (X_{min}) значениями признака в совокупности. Определяется по формуле (3.3.1):

$$R = X_{max} - X_{min}. \qquad (3.3.1)$$

Недостатком данного показателя является то, что его значение зависит от величины только двух крайних вариант и не учитывает степени колеблемости основной массы членов ряда;

• ***среднее линейное отклонение*** – представляет собой среднюю арифметическую абсолютных значений отклонений отдельных вариантов от их средней арифметической и определяется по формулам (3.3.2) и (3.3.3):

1) для несгруппированных данных:

$$\bar{d} = \frac{\sum\left|x_i - \bar{x}\right|}{n}, \qquad (3.3.2)$$

2) для сгруппированных данных:

$$\bar{d} = \frac{\Sigma\left|x_i - \bar{x}\right| f_i}{\Sigma f_i}, \qquad (3.3.3)$$

• ***дисперсия*** – средний квадрат отклонений вариантов от их средней величины и определяется по формулам (3.3.4) и (3.3.5):

1) для несгруппированных данных:

$$\sigma^2 = \frac{\Sigma\left(x_i - \bar{x}\right)^2}{n}, \qquad (3.3.4)$$

2) для сгруппированных данных:

$$\sigma^2 = \frac{\Sigma\left(x_i - \bar{x}\right)^2 f_i}{\Sigma f_i}. \qquad (3.3.5)$$

Преобразовав формулу (3.3.4), получим упрощенную формулу дисперсии (3.3.6):

$$\sigma^2 = \overline{x^2} - (\bar{x})^2, \qquad (3.3.6)$$

то есть дисперсия равна разности средней из квадратов вариантов и квадрата их средней.

Дисперсия обладает рядом свойств, которые позволяют упростить ее вычисление:

1) дисперсия постоянной величины равна нулю;

2) если все варианты значений признака уменьшить на одно и то же число, то дисперсия не уменьшится;

3) если все варианты значений признака уменьшить в одно и то же число раз (k раз), то дисперсия уменьшится в k^2 раз;

• ***среднее квадратическое отклонение*** – представляет собой корень квадратный из дисперсии и определяется по формулам (3.3.7) и (3.3.8):

1) для несгруппированных данных:

$$\sigma = \sqrt{\frac{\Sigma(x_i - \bar{x})^2}{n}}, \qquad (3.3.7)$$

2) для сгруппированных данных:

$$\sigma = \sqrt{\frac{\Sigma(x_i - \bar{x})^2 f_i}{\Sigma f_i}}. \qquad (3.3.8)$$

Размах вариации, среднее линейное и среднее квадратическое отклонение являются величинами именованными, то есть имеют те же единицы измерения, что и индивидуальные значения признака.

2. Относительные показатели вариации:

• ***коэффициент осцилляции*** определяется по формуле (3.3.9):

$$K_R = \frac{R}{\overline{\overline{X}}} \times 100, \qquad (3.3.9)$$

• ***относительное линейное отклонение*** исчисляется по формуле (3.3.10):

$$K_{\bar{d}} = \frac{\bar{d}}{\overline{\overline{X}}} \times 100, \qquad (3.3.10)$$

• ***коэффициент вариации***, определяемый по формуле (3.3.11):

$$V = \frac{\sigma}{\overline{\overline{X}}} \times 100. \qquad (3.3.11)$$

Коэффициент вариации используют не только для сравнительной оценки вариации единиц совокупности, но и как характеристику однородности совокупности. Совокупность считается однородной, если коэффициент вариации не превышает 33%.

Правило сложения дисперсий

Если данные представлены в виде аналитической группировки, то можно вычислить общую, межгрупповую и внутригрупповую дисперсии.

Общая дисперсия (σ^2) характеризует вариацию признака под влиянием всех факторов, формирующих уровень признака у единиц данной совокупности, и определяется по формуле (3.3.5) или (3.3.6).

Межгрупповая дисперсия (δ^2) отражает вариацию признака под влиянием одного фактора, положенного в основу группировки и определяется по формуле (3.3.12):

$$\delta^2 = \frac{\Sigma(\bar{x}_i - \bar{x})^2 n_i}{\Sigma n_i}, \qquad (3.3.12)$$

где $\overline{x_i}$ – средняя по отдельной группе;

$\bar{x}$ – общая средняя для всей совокупности;

n_i – число единиц в определенной группе.

Средняя из внутригрупповых дисперсий $(\overline{\sigma_i^2})$ характеризует случайную вариацию, возникающую под влиянием других факторов, кроме положенного в основу группировки, и рассчитывается по формуле (3.3.13):

$$\overline{\sigma_i^2} = \frac{\Sigma \sigma_i^2 n_i}{\Sigma n_i}, \qquad (3.3.13)$$

где σ_i^2 – дисперсия по отдельной группе.

Дисперсия по отдельной группе равна среднему квадрату отклонений отдельных значений признака *x* внутри группы от средней арифметической этой группы $\bar{x}_i$ и может быть исчислена как простая дисперсия или как взвешенная дисперсия по формулам (3.3.14) и (3.3.15) соответственно:

$$\sigma_i^2 = \frac{\Sigma(x_i - \bar{x}_i)^2}{n}, \qquad (3.3.14)$$

$$\sigma_i^2 = \frac{\Sigma(x_i - \bar{x}_i)^2 f_i}{\Sigma f_i}. \qquad (3.3.15)$$

Указанные дисперсии взаимосвязаны между собой равенством, выраженным формулой (3.3.16):

$$\sigma^2 = \delta^2 + \overline{\sigma_i^2}. \qquad (3.3.16)$$

Это тождество отражает правило сложения дисперсий: величина общей дисперсии равна сумме межгрупповой дисперсии и средней из внутригрупповых дисперсий.

Очевидно, чем больше доля межгрупповой дисперсии в общей дисперсии, тем сильнее влияние группировочного признака на изучаемый признак. Поэтому в статистическом анализе используется ***эмпирический коэффициент детерминации*** (η^2), характеризующий силу влияния группировочного признака на образование общей вариации и определяемый по формуле (3.3.17):

$$\eta^2 = \frac{\delta^2}{\sigma^2}. \qquad (3.3.17)$$

Дисперсия альтернативного признака

Наряду с измерением вариации количественных признаков в статистике измеряют вариацию и альтернативных (качественных) признаков.

Альтернативный признак – признак, которым одни единицы изучаемой совокупности обладают, а другие нет.

Рассчитаем среднее значение альтернативного признака и его дисперсию, для чего введем следующие обозначения:

1 – наличие признака у единиц совокупности;

0 – отсутствие признака;

p – доля единиц, обладающих данным признаком;

q – доля единиц, не обладающих данным признаком.

Тогда ряд распределения, построенный по качественному признаку, будет выглядеть следующим образом:

Таблица 3.3.1

Значение переменной (x_i)	Частота повторений (f_i)
1	p
0	q
Итого	1

Среднее значение альтернативного признака составит (формула (3.3.18)):

$$\bar{x} = \frac{\Sigma x_i f_i}{\Sigma f_i} = \frac{1 \times p + 0 \times q}{p + q} = p \qquad (3.3.18)$$

Дисперсия альтернативного признака определяется формулой (3.3.19):

$$\sigma^2 = \frac{(1 - p)^2 \times p + (0 - p)^2 \times q}{p + q} = p \times q \qquad (3.3.19)$$

Предельное значение вариации альтернативного признака равно 0,25. Оно получается при p = q = 0,5.

3.4. Показатели формы распределения

Обобщающие характеристики центра распределения и степени вариации не дают представления о форме распределения, так как не вскрывают характера изменения частот. Для выражения особенностей формы распределения применяются показатели асимметрии и эксцесса.

Асимметрия

Для сравнительного анализа степени асимметрии нескольких распределений рассчитывают относительный показатель асимметрии по формуле (3.4.1):

$$As = \frac{\bar{X} - Mo}{\sigma}. \qquad (3.4.1)$$

Величина асимметрии может быть:

1) положительная, что указывает на наличие правосторонней асимметрии;

2) отрицательная, что свидетельствует о наличии левосторонней асимметрии.

Наиболее точным и распространенным показателем асимметрии является показатель, основанный на определении центрального момента третьего порядка. Применение этого показателя дает возможность не только определить степень асимметрии, но и ответить на вопрос о наличии или отсутствии асимметрии в распределении признака в генеральной совокупности.

Показатель асимметрии определяется по формуле (3.4.2):

$$As = \frac{\mu_3}{\sigma^3}, \qquad (3.4.2)$$

где μ_3 – центральный момент третьего порядка.

Центральный момент третьего порядка определяется по формуле (3.4.3):

$$\mu_3 = \frac{\Sigma(x_i - \bar{x})^3 f_i}{\Sigma f_i}. \qquad (3.4.3)$$

Эксцесс

Эксцесс представляет собой выпад вершины эмпирического распределения вверх или вниз от вершины кривой нормального распределения и определяется по формуле (3.4.4):

$$Ex = \frac{\mu_4}{\sigma^4} - 3, \qquad (3.4.4)$$

где μ_4 – центральный момент четвертого порядка.

Центральный момент четвертого порядка определяется по формуле (3.4.5):

$$\mu_4 = \frac{\Sigma(x_i - \bar{x})^4 f_i}{\Sigma f_i}. \qquad (3.4.5)$$

Величина эксцесса может быть:

1) положительная, что свидетельствует об островершинном распределении;

2) отрицательная, что указывает на плосковершинное распределение.

Предельным значением отрицательного эксцесса является Ex=-2; величина положительного эксцесса является величиной бесконечной.

3.5. Решение типовых задач

Задача 1. Известны следующие данные о распределении продовольственных магазинов региона по размеру товарооборота за месяц (табл. 3.5.1).

Определить средний месячный размер товарооборота магазинов региона, моду и медиану.

Таблица 3.5.1

Распределение магазинов по размеру товарооборота

Группы магазинов по размеру товарооборота, млн руб. (x_i)	**Число магазинов (f_i)**
40-50	2
50-60	4
60-70	7
70-80	10
80-90	15
90-100	20
100-110	22
110-120	11
120-130	6
130-140	3
Итого	100

Решение:

Для определения требуемых показателей составим вспомогательную таблицу 3.5.2.

Таблица 3.5.2

Вспомогательная таблица для определения показателей центра распределения

Группы магазинов по размеру товарооборота, млн руб.	**Число магазинов (f_i)**	**Середина интервала (x_i')**	$x_i' f_i$	**Накопленные частоты (S)**
1	2	3	4	5
40-50	2	45	90	2
50-60	4	55	220	6
60-70	7	65	455	13
70-80	10	75	750	23
80-90	15	85	1275	38($S_{ме-1}$)
90-100	($f_{мо-1}$) 20 ($f_{ме}$)	95	1900	58
100-110	($f_{мо}$) 22	105	2310	80
1	2	3	4	5
110-120	($f_{мо+1}$) 11	115	1265	91
120-130	6	125	750	97
130-140	3	135	405	100
Итого	100	-	9 420	-

Данный ряд является интервальным рядом распределения, для которого расчет средней арифметической осуществляется по формуле (3.2.2):

$$\overline{X} = \frac{\Sigma x_i' f_i}{\Sigma f_i} = \frac{9420}{100} = 94{,}2 \text{ млн руб.}$$

Расчет моды начинается с определения модального интервала. В вариационном ряду с равными интервалами модальный интервал определяется по наибольшей частоте. Наибольшая частота ($f_{max} = 22$) соответствует интервалу [100-110], который и является модальным интервалом.

Используя формулу (3.2.3), определяем моду:

$$Mo = X_{мо} + h\frac{f_{мо} - f_{мо-1}}{(f_{мо} - f_{мо-1}) + (f_{мо} - f_{мо+1})} = 100 + 10\frac{22 - 20}{(22 - 20) + (22 - 11)} = 101{,}54.$$

Полученное значение моды свидетельствует о том, что в данном регионе чаще всего встречаются предприятия с размером товарооборота 101,54 млн руб.

Расчет медианы в вариационном ряду также начинается с определения медианного интервала.

Интервалом, в котором сумма накопленных частот впервые превысила полусумму всех частот ряда ($\frac{1}{2}\Sigma f_i = 50$), является [90-100].

Численное значение медианы определяется по формуле (3.2.4):

$$M_e = x_{ме} + h\frac{\frac{1}{2}\Sigma f - S_{ме-1}}{f_{ме}} = 90 + 10\frac{50 - 38}{20} = 96 \text{ млн руб.}$$

Полученный результат говорит о том, что размер товарооборота у 50 (половины) магазинов менее 96 млн руб., а у другой половины – более 96 млн руб.

Задача 2. Известны данные о сменной выработке рабочих бригады, представленных интервальным рядом распределения (табл. 3.5.3).

Таблица 3.5.3

Распределение рабочих по сменной выработке изделия

Группы рабочих по сменной выработке изделий, шт.	**Число рабочих**
170-190	10
190-210	20

210-230	50
230-250	20
Итого	100

Определить дисперсию, среднее квадратическое отклонение и показатель вариации.

Решение:

Для определения показателей вариации составим вспомогательную таблицу 3.5.4.

Таблица 3.5.4

Вспомогательная таблица для определения показателей вариации

Группы рабочих по сменной выработке изделий, шт.	**Число рабочих**	**Середина интервала (x_i')**	$x_i' f_i$	$(x_i' - \bar{x})$	$(x_i' - \bar{x})^2$	$(x_i' - \bar{x})^2 f_i$
170-190	10	180	1800	-36	1296	12 960
190-210	20	200	4000	-16	256	5 120
210-230	50	220	11000	4	16	800
230-250	20	240	4800	24	576	11 520
Итого	100	-	21600	-	-	30 400

Определим сменную выработку по формуле (3.2.2):

$$\bar{X} = \frac{\Sigma x_i' f_i}{\Sigma f_i} = \frac{21600}{100} = 216 \text{ шт.}$$

Рассчитаем дисперсию выработки по формуле (3.3.5):

$$\sigma^2 = \frac{\Sigma (x_i - \bar{x})^2 f_i}{\Sigma f_i} = \frac{30400}{100} = 304.$$

Найдем квадратическое отклонение, используя формулу (3.3.8):

$$\sigma = \sqrt{\frac{\Sigma (x_i - \bar{x})^2 f_i}{\Sigma f_i}} = \sqrt{304} = 17{,}44 \text{ шт.}$$

Для расчета коэффициента вариации используем формулу (3.3.11):

$$V = \frac{\sigma}{\bar{X}} \times 100 = \frac{17{,}44}{216} \times 100 = 8\%.$$

Таким образом, данная бригада рабочих достаточно однородна по выработке, поскольку вариация признака составляет лишь 8 %.

Задача 3. Заработная плата 10 рабочих бригады характеризуется следующими данными (табл. 3.5.5):

Таблица 3.5.5

Распределение рабочих бригады по профессиям

Профессия	Число рабочих	Месячная заработная плата каждого работника за месяц, руб.
Токарь	4	3 252; 3 548; 3 600; 3 400
Слесарь	6	3 450; 3 380; 3 260; 3 700; 3 250; 3 372

Проверить правило сложения дисперсий и указать, каково влияние профессии (в %) на различие в уровне заработной платы.

Решение:

1. Рассчитаем среднюю заработную плату бригады по формуле средней арифметической простой (табл. 2.3.1):

$$\overline{X}_{\text{арифм}} = \frac{3252 + 3548 + 3600 + 3400 + 3450 + 3380 + 3260 + 3700 + 3250 + 3372}{10} =$$

$$=2\ 421{,}2 \text{ руб.}$$

2. Определим среднюю заработную плату по профессиям (внутригрупповые средние):

$$\overline{X}^{\text{ток}}_{\text{арифм}} = \frac{3252 + 3548 + 3600 + 3400}{4} = 3450 \text{ руб.}$$

$$\overline{X}^{\text{слес}}_{\text{арифм}} = \frac{3450 + 3380 + 3260 + 3700 + 3250 + 3372}{6} = 3402 \text{ руб.}$$

3. Рассчитаем внутригрупповые дисперсии по формуле (3.3.14):

$$\sigma^2_{\text{ток}} = \frac{\Sigma\left(x_i - \overline{x}_i\right)^2}{n} = \frac{(3252 - 3450)^2 + (3548 - 3450)^2 + (3600 - 3450)^2 + (3400 - 3450)^2}{4}$$

$$= 18\ 452.$$

$$\sigma^2_{\text{слес}} = \frac{\Sigma\left(x_i - \overline{x}_i\right)^2}{n} = \frac{(3450 - 3402)^2 + (3380 - 3402)^2 + \ldots + (3372 - 3402)^2}{6} = 2\ 626{,}7.$$

Средняя из внутригрупповых дисперсий определяется по формуле (3.3.13):

$$\overline{\sigma^2_i} = \frac{\Sigma\sigma^2_i n_i}{\Sigma n_i} = \frac{18452 \times 4 + 22626{,}7 \times 6}{10} = 20956{,}8.$$

4. Используя формулу (3.3.12), рассчитаем межгрупповую дисперсию:

$$\delta^2 = \frac{\Sigma(\overline{x_i} - \overline{x})^2 n_i}{\Sigma n_i} = \frac{(3450 - 3421{,}2)^2 \times 4 + (3402 - 3421{,}2)^2 \times 6}{10} = 552{,}96.$$

5. Определим общую дисперсию:

– по правилу сложения дисперсий, используя формулу (3.3.16):

$$\sigma^2 = \delta^2 + \overline{\sigma_i^2} = 552{,}96 + 20956{,}8 = 21509{,}76;$$

– по формуле (3.3.4):

$$\sigma^2 = \frac{(3252 - 3421{,}2)^2 + (3548 - 3421{,}2)^2 + (3600 - 3421{,}2)^2 + \ldots + (3372 - 3421{,}2)^2}{10}.$$

$$= 21\,509{,}76$$

6. Влияние профессии на различие в уровне заработной платы определяется на основе эмпирического коэффициента детерминации по формуле (3.3.17):

$$\eta^2 = \frac{\delta^2}{\sigma^2} = \frac{552{,}96}{21509{,}76} = 0{,}0257.$$

Таким образом, вариация уровня заработной платы лишь на 2,6% обусловлена влиянием профессии и на 97,4% – влиянием других факторов (образование, стаж, квалификация и др.).

Задача 4. Дисперсия признака равна 360000, коэффициент вариации 50%. Чему равна средняя величина признака?

Решение:

Коэффициент вариации признака определяется по формуле (3.3.11):

$$V = \frac{\sigma}{\overline{X}} \times 100.$$

Зная, что $\sigma = \sqrt{\sigma^2}$, выразим из приведенной формулы среднее значение признака:

$$\overline{X} = \frac{\sqrt{\sigma^2}}{V} \times 100 = \frac{\sqrt{360000}}{50\%} \times 100\% = 1200 \text{ усл. ед.}$$

Задача 5. В проверенной партии готовых изделий из 400 штук восемь оказались бракованными. Определите дисперсию доли бракованных изделий.

Решение:

В данной задаче признаки «бракованное изделие» и «годное изделие» являются альтернативными.

Для определения дисперсии альтернативного признака необходимо определить доли единиц, обладающих данным признаком (наличие брака) или нет.

Определяем долю бракованных изделий:

$$p = \frac{8}{400} = 0{,}02.$$

Тогда доля годных изделий (q) на основе соотношения p + q = 1 составит:

$$q = 1 – 0{,}02 = 0{,}98.$$

Дисперсия доли бракованных изделий по формуле (3.3.19) составит:

$$\sigma = p \times q = 0{,}02 \times 0{,}98 = 0{,}0196.$$

ГЛАВА 4. ВЫБОРОЧНОЕ НАБЛЮДЕНИЕ

4.1. Понятие о выборочном наблюдении. Виды выборок

Выборочное наблюдение – один из наиболее применяемых видов несплошного наблюдения, при котором обследуются не все единицы изучаемой совокупности, а некоторая часть этих единиц. Причем наблюдение организовывается таким образом, чтобы эта часть отобранных единиц в уменьшенном масштабе репрезентировала (представляла) всю совокупность.

Совокупность, из которой производится отбор, называется ***генеральной***, и все ее обобщающие показатели – генеральными.

Совокупность отобранных единиц именуют ***выборочной*** совокупностью, и все ее обобщающие показатели – выборочными.

Имеется ряд причин, в силу которых выборочному наблюдению отдается предпочтение перед сплошным наблюдением:

1) экономия времени и средств в результате сокращения объема работы;

2) сведение к минимуму порчи или уничтожения исследуемых объектов (например, испытание электрических лампочек на продолжительность горения);

3) достижение большой точности результатов обследования благодаря сокращению ошибок, происходящих при регистрации.

Основная задача выборочного наблюдения состоит в определении характеристик генеральной совокупности по выборочным данным. Причем между характеристиками выборочной совокупности и искомыми характеристиками генеральной совокупности существует некоторое расхождение, которое называется ошибкой.

Выделяют два вида ошибок:

1) ***ошибки регистрации*** – свойственны как сплошному, так и несплошному наблюдению и вызываются недостаточной квалификацией наблюдателя, неточностью подсчетов и др.;

2) ***ошибки репрезентативности*** – свойственны только выборочному наблюдению и представляют собой величину возможных расхождений между показателями выборочной и генеральной совокупности.

Для каждого конкретного выборочного наблюдения значение ошибки репрезентативности может быть определено по соответствующим формулам, которые зависят от способа отбора и формирования выборочной совокупности, а также степени охвата единиц совокупности.

По способу отбора различают повторную и бесповторную выборки.

При ***повторной выборке*** общая численность единиц генеральной совокупности в процессе выборки остается неизменной. Ту или иную единицу, попавшую в выборку, после регистрации снова возвращают в генеральную совокупность, и она сохраняет равную возможность со всеми прочими единицами при повторном отборе единиц вновь попасть в выборку.

При ***бесповторной выборке*** единица совокупности, попавшая в выборку, в генеральную совокупность не возвращается и в дальнейшем в выборке не участвует.

По способу формирования различают простую случайную, механическую, типическую, серийную и комбинированную выборки.

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 быть отобранной. Случайность отбора можно обеспечить путем применения жеребьевки или использования таблицы случайных чисел.

Механическая выборка заключается в отборе единиц из генеральной совокупности через равные промежутки из определенного расположения их в генеральной совокупности (например, по алфавиту).

При ***типической*** выборке вся генеральная совокупность предварительно подразделяется на качественно-однородные по изучаемому признаку группы, а затем из этих групп производится случайный отбор n единиц.

При ***серийной*** выборке отбору подлежат не отдельные единицы совокупности, а целые их группы или серии, в состав которых входят единицы, связанные определенным образом: территориально, организационно, во времени и др. Отбор серий может производиться в порядке повторного и бесповторного отбора, а внутри отобранных серий обследуются все единицы, то есть проводится сплошное наблюдение.

Комбинированная выборка предполагает использование нескольких способов выборки.

По степени охвата единиц совокупности различают большие и малые ($n<30$) выборки.

Оценить репрезентативность выборки можно, рассчитав ошибки выборки.

4.2. Ошибки выборки

Выделяют среднюю и предельную ошибки выборки. ***Средняя ошибка*** представляет среднее квадратическое отклонение возможных значений выборочной средней от генеральной средней. Фактические расхождения между выборочной средней и генеральной представляют собой ***предельную*** ошибку выборки.

Использование той или иной формулы для определения ошибки выборки зависит от способа отбора единиц (повторный или бесповторный) и вида выборочной характеристики (средняя или доля). В таблице 4.2.1 приведены формулы расчета ошибок простой случайной выборки.

Таблица 4.2.1

Формулы ошибок простой случайной выборки

Вид ошибки	**Выборочная характеристика**	**Способ отбора единиц**	
		Повторный	**Бесповторный**
Средняя ошибка (μ)	средняя	$\mu_{\tilde{x}} = \sqrt{\frac{\sigma_в^2}{n}} = \frac{\sigma}{\sqrt{n}}$	$\mu_{\tilde{x}} = \sqrt{\frac{\sigma_в^2}{n}(1 - \frac{n}{N})}$
	доля	$\mu_p = \sqrt{\frac{w(1-w)}{n}}$,	$\mu_p = \sqrt{\frac{w(1-w)}{n}(1 - \frac{n}{N})}$
Предельная ошибка (Δ)	средняя	$\Delta_{\tilde{x}} = t\sqrt{\frac{\sigma_в^2}{n}}$	$\Delta_{\tilde{x}} = t\sqrt{\frac{\sigma_в^2}{n}(1 - \frac{n}{N})}$
	доля	$\Delta_p = t\sqrt{\frac{w(1-w)}{n}}$	$\Delta_p = t\sqrt{\frac{w(1-w)}{n}(1 - \frac{n}{N})}$

В таблице использованы следующие обозначения:

n – объем выборки;

N – объем генеральной совокупности;

$\sigma_в^2$ – выборочная дисперсия;

w – выборочная доля;

p – генеральная доля;

$\bar{x}$ – генеральная средняя;

$\tilde{x}$ – выборочная средняя;

t – коэффициент доверия, зависящий от вероятности, с которой гарантируется предельная ошибка, определяется по таблице приложения 1.

Рассмотренные выше формулы средней и предельной ошибок выборки разработаны применительно к простой случайной выборке и ее разновидности – механической выборке. Для других видов выборок эти формулы несколько видоизменяются, конкретизируются в соответствии с особенностями видов выборок.

Формулы предельной ошибки выборки позволяют решать задачи трех видов:

• ***Определение доверительных пределов генеральных характеристик*** с заданной степенью надежности на основе показателей, полученных по данным выборки.

При решении задач этого типа рассчитываются выборочная средняя ($\tilde{x}$) (или выборочная доля) и с заданной вероятностью (Р) предельная ошибка выборки (Δ), на основании которой определяются доверительные пределы:

1) для генеральной средней по формуле (4.2.1):

$$\tilde{x} - t \cdot \mu_{\tilde{x}} \le \bar{x} \le \tilde{x} + t \cdot \mu_{\tilde{x}}, \qquad (4.2.1)$$

2) для генеральной доли по формуле (4.2.2):

$$w - t \cdot \mu_p \le p \le w + t \cdot \mu_p. \qquad (4.2.2)$$

• ***Определение доверительной вероятности того, что расхождение*** между выборочными и генеральными характеристиками не превзойдет определенную заданную величину.

Доверительная вероятность является функцией от *t*, определяемой по формуле (4.2.3):

$$t = \frac{\Delta_{\tilde{x}}}{\mu_{\tilde{x}}}. \qquad (4.2.3)$$

По величине *t* определяется доверительная вероятность (приложение 1).

• ***Определение необходимого объема выборки***, который с определенной вероятностью обеспечит заданную точность выборочных показателей.

В таблице 4.2.2 приведены формулы для расчета численности простой случайной выборки, полученные в ходе преобразования формулы предельной ошибки выборки.

Таблица 4.2.2

Формулы для определения численности простой случайной выборки

Численность выборки	**Способ отбора единиц**	
	Повторный	**Бесповторный**
для средней	$n = \frac{t^2 \sigma_в^2}{\Delta_{\tilde{x}}^2}$	$n = \frac{t^2 \sigma_в^2 N}{N\Delta_{\tilde{x}}^2 + t^2 \sigma_в^2}$
для доли*	$n = \frac{t^2 w(1-w)}{\Delta_p^2}$	$n = \frac{t^2 N w(1-w)}{N\Delta_p^2 + t^2 w(1-w)}$

* В случаях, когда дисперсия альтернативного признака неизвестна, в расчет вводят максимальную величину дисперсии доли, равную 0,25 (если w = 0,5, то w·(1-w) = 0,25).

4.3. Малая выборка

Выборки, при которых наблюдением охватывается небольшое число единиц ($n<30$), принято называть ***малыми выборками***.

При малых выборках распределение выборочных средних, а следовательно, и ошибок выборки отличается от нормального. Поэтому для оценки результатов малой выборки используют несколько видоизмененные формулы.

Во-первых, средняя ошибка малой выборки определяется по формуле (4.3.1):

$$\mu_{м.в} = \sqrt{\frac{\sigma_в^2}{n-1}}. \tag{4.3.1}$$

Во-вторых, при недостаточном объеме выборки разность между выборочной и генеральной средней ($t = \frac{\tilde{x} - \bar{x}}{\mu_{м.в}}$) имеет распределение Стьюдента, а не нормальное распределение. Тогда вероятность того, что генеральная средняя находится в определенных пределах, устанавливается по формуле (4.3.2):

$$P(|\tilde{x} - \bar{x}| \leq t\mu_{м.в.}) = 2S(t) - 1. \tag{4.3.2}$$

Величина S(t) зависит от коэффициента доверия (t) и числа степеней свободы (k = n -1).

На практике пользуются готовыми таблицами распределений Стьюдента, которые позволяют определить величину S(t) двумя способами:

1) искомая вероятность находится по приложению 2 на пересечении строки, соответствующей значению коэффициента доверия (t) и столбца, соответствующего числу степеней свободы (k);

2) на основе приложения 3 значение коэффициента доверия t определяется по числу степеней свободы (k) и доверительной вероятности P(t) (0,9; 0,95; 0,99) или уровню значимости $\alpha = 1 - P(t)$ (0,1; 0,05; 0,01).

4.4. Решение типовых задач

Задача 1. Для определения скорости расчетов с кредиторами предприятий корпорации в коммерческом банке была проведена случайная выборка 100 платежных документов, по которым средний срок перечисления и получения денег оказался равным 22 дням ($\tilde{x} = 22$) со средним квадратическим отклонением 6 дней ($\sigma_в = 6$).

Определить с вероятностью 0,9545 предельную ошибку выборочной средней и доверительные пределы средней продолжительности расчетов предприятий данной корпорации.

Решение:

Предельная ошибка выборочной средней (Δ) определяется по формуле повторного отбора, так как численность генеральной совокупности (N) не известна.

По данным приложения 1 для вероятности P = 0,9545 определяем t = 2.

Предельная ошибка выборки составит:

$$\Delta_{\tilde{x}} = t \times \mu = t\sqrt{\frac{\sigma_в^2}{n}} = 2\sqrt{\frac{6^2}{100}} = 1{,}2 \text{ дня.}$$

Доверительные пределы генеральной средней определяем по формуле (4.2.1):

$$\tilde{x} - t \cdot \mu_{\tilde{x}} \le \bar{x} \le \tilde{x} + t \cdot \mu_{\tilde{x}},$$

$$22 - 1{,}2 \le \bar{x} \le 22 + 1{,}2$$

$$20{,}8 \le \bar{x} \le 23{,}2$$

Таким образом, с вероятностью 0,9545 можно утверждать, что средняя продолжительность расчетов предприятий данной корпорации колеблется в пределах от 20,8 до 23,2 дней.

Задача 2. По городской телефонной сети в порядке случайной выборки провели 100 наблюдений (n = 100) и установили среднюю продолжительность одного телефонного разговора 5 мин. ($\tilde{x} = 5$) при среднем квадратическом отклонении 2 мин. ($\sigma_в = 2$).

Какова вероятность того, что ошибка репрезентативности при определении средней продолжительности телефонных разговоров не превысит 18 с. ($\Delta = 18$)?

Решение:

Так как объем генеральной совокупности неизвестен, расчет ведется для повторного отбора.

Первоначально определяем среднюю ошибку выборки для средней величины:

$$\mu_{\tilde{x}} = \sqrt{\frac{\sigma_в^2}{n}} = \frac{\sigma}{\sqrt{n}} = \frac{2}{\sqrt{100}} = 0{,}2 \text{ мин.}$$

Далее определяется параметр t по формуле (4.2.3):

$$t = \frac{\Delta_{\tilde{x}}}{\mu_{\tilde{x}}} = \frac{0,3}{0,2} = 1,5.$$

Затем по таблице приложения 1 на основе значения t определяется искомая вероятность. При t = 1,5 вероятность P = 0,8664.

Задача 3. Для определения среднего возраста 1200 студентов факультета необходимо провести выборочное обследование методом случайного бесповторного отбора. Предварительно установлено, что среднее квадратическое отклонение возраста студентов составляет 10 лет.

Сколько студентов нужно обследовать, чтобы с вероятностью 0,9545 средняя ошибка выборки не превышала 3 года?

Решение:

Рассчитаем необходимую численность выборки по формуле бесповторного отбора, учитывая, что при P = 0,9545 t = 2:

$$n = \frac{t^2\sigma_в^2 N}{N\Delta_{\tilde{x}}^2 + t^2\sigma_в^2} = \frac{2^2 \times 10^2 \times 1200}{1200 \times 3^2 + 2^2 \times 10^2} \approx 43 \text{ чел.}$$

Таким образом, выборка численностью 43 чел. обеспечит заданную точность при бесповторном отборе.

Задача 4. Среди выборочно обследованных 1 000 семей из 50 000 семей региона по уровню душевого дохода малообеспеченных оказалось 300 семей.

Определить с вероятностью 0,9973 долю малообеспеченных семей во всем регионе.

Решение:

Определяем выборочную долю – долю малообеспеченных семей среди обследованных семей:

$$w = \frac{300}{1000} = 0,3.$$

По таблице приложения 1 для вероятности 0,9973 определяем t = 3. Предельную ошибку доли рассчитываем по формуле бесповторного отбора:

$$\Delta_p = t\sqrt{\frac{w(1-w)}{n}(1-\frac{n}{N})} = 3\sqrt{\frac{0,3(1-0,3)}{1000}(1-\frac{1000}{50000})} = 0,014.$$

Доверительные пределы генеральной доли определяем по формуле (4.2.2):

$$w - t \cdot \mu_p \le p \le w + t \cdot \mu_p$$

$$0,3 - 0,014 \le p \le 0,3 + 0,014$$

$$0,286 \le p \le 0,314.$$

Таким образом, почти достоверно, с вероятностью 0,997, можно утверждать, что доля малообеспеченных семей среди всех семей региона колеблется от 28,6% до 31,4%.

Задача 5. На участке леса выборочным методом обследовано 8 деревьев с целью определения объема деловой древесины в одном дереве. Выборочная средняя составила 1,8 $м^3$ при среднем квадратическом отклонении 0,28 $м^3$.

Определить с вероятностью 0,955 доверительные пределы для генеральной средней.

Решение: Так как объем выборочной совокупности меньше 30 единиц, в расчетах используются формулы для малой выборки.

Определяем среднюю ошибку выборки по формуле (4.3.1):

$$\mu_{м.в} = \sqrt{\frac{\sigma_в^2}{n-1}} = \frac{0{,}28}{\sqrt{8-1}} = 0{,}11 \text{ м}^3.$$

Из условия $P(|\tilde{x} - \bar{x}| \le t\mu_{м.в.}) = 2S(t) - 1$ определяем $S(t)$:

$$S(t) = \frac{0{,}955 + 1}{2} = 0{,}978.$$

По таблице распределения вероятностей Стьюдента (приложение 2) находим, что для $n = 8$ вероятности $P = 0{,}978$ соответствует $t = 2{,}3$.

Следовательно, предельная ошибка составит:

$$\Delta = t \times \mu_{м.в.} = 2{,}3 \times 0{,}11 = 0{,}253.$$

Отсюда доверительные пределы генеральной средней, рассчитанные по формуле (4.2.1), равны:

$$\tilde{x} - t \cdot \mu_{м.в.} \le \bar{x} \le \tilde{x} + t \cdot \mu_{м.в.},$$

$$1{,}8 - 0{,}253 \le \bar{x} \le 1{,}8 + 0{,}253$$

$$1{,}547 \le \bar{x} \le 2{,}053.$$

Следовательно, с вероятностью 0,955 можно утверждать, что в исследуемой совокупности в одном дереве содержится деловой древесины от 1,547 до 2,053 $м^3$.

ГЛАВА 5. СТАТИСТИЧЕСКОЕ ИЗУЧЕНИЕ КОРРЕЛЯЦИОННЫХ ВЗАИМОСВЯЗЕЙ

5.1. Виды зависимостей между признаками

Исследуя явления в различных областях, статистика неизбежно сталкивается с зависимостями как между количественными, так и между качественными показателями.

Основная задача статистики заключается в выявлении зависимости между признаками и их количественной характеристикой.

Признаки по их значению для изучения взаимосвязи делятся на две группы:

1) ***факторные*** – признаки, обуславливающие изменение других связанных с ними признаков;

2) ***результативные*** – признаки, являющиеся результатом влияния других признаков.

Связи между признаками и явлениями, ввиду их большого разнообразия, классифицируются по следующим основаниям:

1) По ***характеру связи*** различают функциональную и стохастическую.

Функциональная связь – связь, при которой определенному значению факторного признака соответствует одно и только одно значение результативного признака. Эти связи обычно встречаются в точных науках: математике, физике и др.

Стохастическая связь – связь, проявляющаяся не в каждом отдельном случае, а в общем, среднем при большом числе наблюдений. Например, урожайность определенной культуры зависит от количества выпавших осадков. Между этими признаками нет функциональной связи, то есть при одном и том же количестве выпавших осадков урожайность будет неодинакова, так как на нее кроме осадков влияет много других факторов (качество семян, своевременность уборки, уход за посевами и другие).

Частным случаем стохастической связи является ***корреляционная связь***, под которой понимают связь, проявляющуюся при большом числе наблюдений в виде зависимости между факторным признаком и средним значением результативного признака.

2) По ***направлению действия*** выделяют связь прямую и обратную.

При ***прямой связи*** значения факторного и результативного признаков изменяются в одном направлении, то есть с увеличением значений факторного признака увеличивается и значение результативного признака, и наоборот. Например, чем выше квалификация рабочего, тем выше уровень производительности труда – прямая связь.

При ***обратной зависимости*** значения факторного и результативного признаков изменяются в разных направлениях. Например, чем выше производительность труда, тем ниже себестоимость единицы продукции.

3) По ***аналитическому выражению*** связи могут быть прямолинейными и криволинейными.

Прямолинейная связь – связь, приближенно выраженная уравнением прямой линии.

Криволинейная связь – связь, выражаемая уравнением какой-либо кривой линии (параболы, гиперболы и т.д.).

4) По ***степени тесноты*** связи различают количественные критерии оценки тесноты связи (табл. 5.1.1).

Таблица 5.1.1

Количественные критерии оценки тесноты связи

Величина коэффициента корреляции	**Характер связи**
$0 - \lvert 0{,}3 \rvert$	практически отсутствует
$\lvert 0{,}3 \rvert - \lvert 0{,}5 \rvert$	слабая
$\lvert 0{,}5 \rvert - \lvert 0{,}7 \rvert$	умеренная
$\lvert 0{,}7 \rvert - \lvert 1{,}0 \rvert$	сильная

Показатели тесноты связи могут принимать значения в пределах от -1 до + 1. Чем ближе он по абсолютной величине к 1, тем теснее связь. Знак при количественном значении показателя тесноты связи указывает направление связи: знак «+» соответствует прямой связи, знак «–» – обратной.

Определение степени тесноты связи дает возможность охарактеризовать зависимость вариации результативного признака от вариации факторного признака.

Показатели тесноты связи между признаками делятся на две группы в зависимости от вида признака:

5.2. Показатели тесноты связи между количественными признаками

Коэффициент Фехнера основан на сравнении поведения отклонений индивидуальных значений факторного и результативного признака от своей средней величины. При этом во внимание принимаются не величины отклонений, а их знаки. Определив знаки отклонения от средней величины в каждом ряду, рассматривают все пары знаков и подсчитывают число их совпадений и несовпадений по формуле (5.2.1):

$$K_ф = \frac{\Sigma C - \Sigma H}{\Sigma C + \Sigma H}, \qquad (5.2.1)$$

где $\sum C$ – число совпадений знаков отклонений индивидуальных величин от средней;

$\sum H$ – число несовпадений знаков отклонений индивидуальных величин от средней.

Линейный коэффициент корреляции применяется в случае линейной зависимости между признаками и позволяет учесть не только знаки отклонений индивидуальных значений признака от средней, но и саму величину таких отклонений. Линейный коэффициент корреляции определяется по формуле (5.2.2):

$$r = \frac{\overline{xy} - \bar{x}\bar{y}}{\sigma_x \sigma_y}, \qquad (5.2.2)$$

где $\overline{xy}$ – среднее значение произведения значений признака *x* и *y*;

$\bar{x}$ – среднее значение признака *x*;

$\bar{y}$ – среднее значение признака *y*;

σ_x, σ_y – среднее квадратическое отклонение соответственно признака *x* и *y*.

Оценка существенности линейного коэффициента корреляции при большом объеме выборки (свыше 500) проводится по формуле (5.2.3):

$$t_{расч} = \frac{|r|}{\sigma_r}, \qquad (5.2.3)$$

где r – линейный коэффициент корреляции;

σ_r – средняя квадратическая ошибка линейного коэффициента корреляции.

Средняя квадратическая ошибка линейного коэффициента корреляции определяется по формуле (5.2.4):

$$\sigma_r = \frac{1 - r^2}{\sqrt{n - 1}}. \qquad (5.2.4)$$

Если это отношение окажется больше значения t – критерия Стьюдента, определяемого по приложению 4 при числе степеней свободы k = n – 2 и с вероятностью (1 – α), то следует говорить о существенности коэффициента корреляции (α – уровень значимости, принимающий значения 0,01 или 0,05).

При недостаточно большом объеме выборки величину средней квадратической ошибки коэффициента корреляции определяют по формуле (5.2.5):

$$\sigma_r = \frac{\sqrt{1 - r^2}}{\sqrt{n - 2}}. \qquad (5.2.5)$$

В тех случаях, когда линейный коэффициент корреляции 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 (5.2.6):

$$\sigma_z = \frac{1}{\sqrt{n - 3}}. \qquad (5.2.6)$$

По таблице соотношений между *y* и Z находят значение Z, соответствующее рассчитанному коэффициенту корреляции.

Если соотношение Z к средней квадратической ошибке окажется больше табличного значения критерия Стьюдента при определенном уровне значимости, то можно говорить о наличии связи между признаками в генеральной совокупности.

Коэффициенты корреляции рангов Спирмэна и Кендэла основаны на корреляции не самих значений коррелируемых величин, а их рангов.

Ранг – порядковый номер значений признака, расположенных в порядке возрастания или убывания их величин. Если отдельные значения признака имеют одинаковую количественную оценку, то ранг всех этих значений принимается равным средней арифметической от соответствующих номеров мест.

Коэффициент корреляции рангов Спирмена определяется по формуле (5.2.7):

$$\rho = 1 - \frac{6\,\Sigma d^2}{n(n^2 - 1)}, \qquad (5.2.7)$$

где d – разность рангов x и y;

n – число пар наблюдения.

Коэффициент корреляции рангов Кендэла определяется по формуле (5.2.8):

$$\tau = \frac{2S}{n(n - 1)}, \qquad (5.2.8)$$

где S – сумма положительных (P) и отрицательных баллов (Q).

Приведенные формулы коэффициентов корреляции рангов Спирмэна и Кендэла применяются для случаев, когда отдельные значения признака не повторяются.

5.3. Показатели тесноты связи между качественными признаками

Изучение зависимости между качественными признаками осуществляется при помощи таблицы взаимной сопряженности (ТВС).

В простейшем случае, если совокупность единиц по каждому из двух взаимосвязанных признаков подразделяется на две подгруппы (альтернативные признаки), зависимость между ними изучается при помощи четырехклеточной таблицы корреляции, или таблицы «четырех полей». В клетках таблицы на пересечении строк и граф проставляются числа (a, b, c, d), показывающие, сколько единиц совокупности встречается с сочетанием одного и другого признаков (табл. 5.3.1).

Таблица 5.3.1

Структура таблицы четырех полей

Признаки	**A**	$\overline{A}$	**Итого**
B	a	b	a + b
$\overline{B}$	c	d	c + d
Итого	a + c	b + d	a + b + c + d

Применительно к таблице «четырех полей» для измерения тесноты связи между признаками используют коэффициенты ассоциации или контингенции.

Коэффициент ассоциации определяется по формуле (5.3.1):

$$К_а = \frac{ad - bc}{ad + bc}. \quad (5.3.1)$$

Недостатком коэффициента ассоциации является то, что в случае отсутствия частоты в одной из клеток данный показатель всегда будет равен по модулю 1, что даст преувеличенную оценку степени тесноты связи между признаками.

Чтобы этого избежать, применяется другой показатель – ***коэффициент контингенции***, определяемый по формуле (5.3.2):

$$К_{конт} = \frac{ad - bc}{\sqrt{(a + b)(c + d)(a + c)(b + d)}}. \quad (5.3.2)$$

Между коэффициентами контингенции и ассоциации существует следующее соотношение, выраженное формулой (5.3.3):

$$К_{конт} < К_а. \quad (5.3.3)$$

Связь считается подтвержденной, если $|К_а|>0{,}5$ и $|К_{конт}|>0{,}3$.

Если по каждому из двух качественных признаков выделяется число групп больше двух, то для определения тесноты связи применяются коэффициенты взаимной сопряженности ***Пирсона*** и ***Чупрова***, определяемые соответственно по формулам (5.3.4) и (5.3.5):

$$К_П = \sqrt{\frac{\ ^2}{1 + \ ^2}}, \quad (5.3.4)$$

$$К_Ч = \sqrt{\frac{\ ^2}{\sqrt{(К_1 - 1)(К_2 - 1)}}}, \quad (5.3.5)$$

где φ^2 – показатель взаимной сопряженности;

$К_1$ – число групп по столбцам таблицы;

$К_2$ – число групп по строкам таблицы.

Показатель взаимной сопряженности определяется по формуле (5.3.6):

$$\varphi^2 = \Sigma \frac{f_{ij}^2}{A_i A_j} - 1, \quad (5.3.6)$$

где f_{ij} – частоты каждой клетки;

A_i – итоговые частоты по строкам;

A_j – итоговые частоты по графам.

Коэффициент взаимной сопряженности Чупрова является более гибким, так как учитывает число образуемых по каждому признаку групп ($К_1$ и $К_2$), поэтому результат является более точным по сравнению с коэффициентом взаимной сопряженности по формуле Пирсона.

После установления достаточной степени тесноты связи выполняется построение модели связи (уравнения регрессии). ***Уравнение регрессии*** позволяет определить, каким в среднем будет значение результативного признака при том или ином значении факторного признака, если остальные факторы, влияющие на результативный признак и не связанные с факторным, не учитывать.

Определить тип уравнения можно, исследуя зависимость графически. Однако существуют общие указания, позволяющие выявить уравнение связи, не прибегая к графическому изображению:

1) если результативный и факторный признаки возрастают одинаково, это говорит о наличии линейной связи между признаками;

2) обратная зависимость между факторным и результативным признаками свидетельствует о гиперболической связи между признаками;

3) если результативный признак увеличивается в арифметической прогрессии, а факторный значительно быстрее, то используется параболическая или степенная функции.

После выявления вида уравнения регрессии, по эмпирическим данным определяются параметры искомого уравнения с помощью способа наименьших квадратов. В зависимости от формы связи в каждом отдельном случае устанавливается своя система уравнений.

Так, в случае прямолинейной связи между x и y вида $\bar{y}_x = a_0 + a_1 x$ параметры уравнения (a_0 и a_1) определяются через следующую систему нормальных уравнений:

$$\begin{cases} na_0 + a_1 \sum x = \sum y \\ a_0 \sum x + a_1 \sum x^2 = \sum xy \end{cases}.$$

Вычислив по эмпирическим данным все записанные в уравнении суммы и подставив их в систему уравнений, находят параметры искомой прямой.

5.4. Решение типовых задач

Задача 1. По группе однородных предприятий имеются данные об объеме выпущенной продукции и уровне механизации трудоемких и тяжелых работ, представленные в таблице 5.4.1.

Оценить степень тесноты связи между уровнем механизации трудоемких работ и объемом произведенной продукции при помощи коэффициента Фехнера.

Таблица 5.4.1

Взаимосвязь объема выпущенной продукции и уровня механизации трудоемких работ

№ предприятия	Уровень механизации трудоемких и тяжелых работ, % (x)	Объем произведенной продукции, млн руб. (y)
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
Итого	528	1596

Решение:

Рассчитаем среднее значение признаков *x* и *y*:

$$\bar{x} = \frac{\sum_{i=1}^{n} x_i}{n} = \frac{528}{12} = 44{,}0\ \%;$$

$$\bar{y} = \frac{\sum_{i=1}^{n} y_i}{n} = \frac{1596}{12} = 133{,}0 \text{ млн руб.}$$

Для расчета коэффициента составляется вспомогательная таблица 5.4.2, в которой помимо исходных данных определяются знаки отклонений индивидуальных значений признаков от их средних величин.

Используя формулу (5.2.1), определяем коэффициент Фехнера:

$$K_{ф} = \frac{\Sigma C - \Sigma H}{\Sigma C + \Sigma H} = \frac{9 - 3}{9 + 3} = 0{,}5.$$

Полученное значение коэффициента свидетельствует о наличии связи между уровнем механизации работ и объемом продукции.

Таблица 5.4.2

Вспомогательная таблица для определения тесноты связи между признаками

Уровень механизации трудоемких и тяжелых работ, % (x)	Объем продукции, млн руб. (y)	Знаки отклонений	
		$x_i - \bar{x}$	$y_i - \bar{y}$
22	117	-	-
85	186	+	+
67	86	+	-
36	112	-	-
21	52	-	-
40	132	-	-
39	141	-	+
39	158	-	+
31	120	-	-
62	197	+	+
36	106	-	-
50	189	+	+

Задача 2. По группе однородных предприятий имеются данные об объеме выпущенной продукции и стоимости основных производственных фондов (ОПФ), представленные в таблице 5.4.3.

Оценить степень тесноты связи между стоимостью основных производственных фондов и объемом произведенной продукции, используя коэффициенты корреляции Кендэла и Спирмэна.

Таблица 5.4.3

Взаимосвязь объема произведенной продукции со стоимостью ОПФ

№ предприятия	Стоимость основных производственных фондов, млн руб. (x)	Объем произведенной продукции, млн руб. (y)
1	1,2	2,8
2	1,6	4,0
3	2,5	3,8
4	3,8	6,5
5	4,3	8,0
6	5,5	10,1
7	6,0	9,5
8	8,0	12,5
9	9,1	18,3

10	10,0	24,5

Решение:

Для расчета коэффициента корреляции рангов Спирмэна значения случайных величин *x* и *y* ранжируются, то есть нумеруются в порядке возрастания (столбцы 3 и 4 таблицы 5.4.4).

Таблица 5.4.4

Вспомогательная таблица для определения коэффициента корреляции рангов Спирмэна

Стоимость основных производственных фондов, млн руб. (x)	**Объем произведенной продукции, млн руб. (y)**	**Ранги**		**Разность рангов** $d = N_x - N_y$	d^2
		N_x	N_y		
1	2	3	4	5	6
1,2	2,8	1	1	0	0
1,6	4,0	2	3	-1	1
2,5	3,8	3	2	1	1
3,8	6,5	4	4	0	0
4,3	8,0	5	5	0	0
1	2	3	4	5	6
5,5	10,1	6	7	-1	1
6,0	9,5	7	6	1	1
8,0	12,5	8	8	0	0
9,1	18,3	9	9	0	0
10,0	24,5	10	10	0	0

Используя формулу (5.2.7), определяем коэффициент корреляции рангов Спирмэна:

$$\rho = 1 - \frac{6\,\Sigma d^2}{n(n^2-1)} = 1 \quad \frac{6 \times 4}{10(100-1)} = 0{,}98.$$

Полученное значение коэффициента корреляции рангов Спирмэна свидетельствует об очень тесной связи между стоимостью основных производственных фондов и объемом выпущенной продукции.

Для расчета коэффициента корреляции рангов Кендэла значения признаков *x* и *y* также ранжируются. Затем определяют меру соответствия последовательности рангов *y* последовательности рангов *x*. При этом для каждого ранга *y* определяют число следующих за ним значений рангов, превышающих его величину. Сумму чисел таких превышений считают со знаком плюс (+). Аналогично для каждого ранга *y* определяют число следующих за ним рангов, имеющих значение меньше его величины. Сумму чисел таких случаев считают со знаком минус (–).

Для расчета коэффициента корреляции рангов Кендэла составим вспомогательную таблицу 5.4.5, в которой в столбцах 5 и 6 представлен расчет положительных и отрицательных баллов.

Сумма положительных (P) и отрицательных (Q) баллов составит:

P = 9+7+7+6+5+3+3+2+1 = 43

Q = -1-1 = -2

Используя формулу (5.2.8), рассчитаем коэффициент корреляции рангов Кендэла:

$$\tau = \frac{2S}{n(n-1)} = \frac{2 \times (43-2)}{10(10-1)} = 0{,}91.$$

Полученное значение коэффициента корреляции рангов Кендэла также свидетельствует о значительной тесноте связи между признаками x и y.

Таблица 5.4.5

Вспомогательная таблица для определения коэффициента корреляции рангов Кендэла

Стоимость основных производственных фондов, млн руб. (x)	**Объем произведенной продукции, млн руб. (y)**	**Ранги**		**P**	**Q**
		N_x	N_y		
1,2	2,8	1	1	9	0
1,6	4,0	2	3	7	-1
2,5	3,8	3	2	7	0
3,8	6,5	4	4	6	0
4,3	8,0	5	5	5	0
5,5	10,1	6	7	3	-1
6,0	9,5	7	6	3	0
8,0	12,5	8	8	2	0
9,1	18,3	9	9	1	0
10,0	24,5	10	10	-	-

Задача 3. В результате обследования работников предприятия получены следующие данные, представленные в таблице 5.4.6

.

Таблица 5.4.6

Взаимосвязь уровня образования работников со степенью их удовлетворения своей работой

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

Определить тесноту связи между уровнем образования и удовлетворенностью своей работой с помощью коэффициентов ассоциации и контингенции.

Решение:

Используя формулы (5.3.1) и (5.3.2) рассчитаем коэффициенты ассоциации и контингенции:

$$К_а = \frac{ad - bc}{ad + bc} = \frac{300 \times 250 - 50 \times 200}{300 \times 250 + 50 \times 200} = \frac{65000}{85000} = 0{,}765,$$

$$К_{конт} = \frac{ad - bc}{\sqrt{(a+b)(c+d)(a+c)(b+d)}} = \frac{65000}{\sqrt{350 \times 450 \times 500 \times 300}} = 0{,}423.$$

Полученные коэффициенты подтверждают наличие существенной связи между исследуемыми признаками. Однако коэффициент контингенции всегда меньше коэффициента ассоциации и дает более корректную оценку тесноты связи.

Задача 4. По данным о сумме активов и кредитных вложениях коммерческих банков определить тесноту связи между данными признаками, рассчитав линейный коэффициент корреляции.

Исходные данные для расчета представлены в таблице 5.4.7.

Таблица 5.4.7

Взаимосвязь величины активов банка с объемом кредитных вложений

Сумма активов, млн руб.	Кредитные вложения, млн руб.
3176	2496
3066	1962
2941	783
1997	1319
1865	1142
1194	658
518	311

Решение:

Для расчета линейного коэффициента корреляции составим вспомогательную таблицу 5.4.8.

Таблица 5.4.8

Вспомогательная таблица для расчета линейного коэффициента корреляции

№ банка	Сумма активов, млн руб. (y)	Кредитные вложения, млн руб. (x)	x^2	y^2	xy
1	3176	2496	6 230 016	10 086 976	7 927 296
2	3066	1962	3 849 444	9 400 356	6 015 492
3	2941	783	613 089	8 649 481	2 302 803
4	1997	1319	1 739 761	3 988 009	2 634 043
5	1865	1142	1 304 164	3 478 225	2 129 830
6	1194	658	432 964	1 425 636	785 652
7	518	311	96 721	268 324	161 098
Всего	14 757	8 671	14 266 159	37 297 007	21 956 214

По формуле 5.2.2 рассчитаем линейный коэффициент корреляции:

$$r = \frac{\overline{xy} - \bar{x}\bar{y}}{\sigma_x \sigma_y}.$$

$$\overline{xy} = \frac{\Sigma xy}{n} = \frac{21956214}{7} = 3\ 136\ 602$$

$$\bar{x} = \frac{\Sigma x}{n} = \frac{8671}{7} = 1\ 238{,}7$$

$$\bar{y} = \frac{\Sigma y}{n} = \frac{14757}{7} = 2\ 108$$

$$\sigma_x = \overline{x^2} - (\bar{x})^2 = \frac{14266159}{7} - 1238{,}7^2 = 709{,}7$$

$$\sigma_y = \overline{y^2} - (\bar{y})^2 = \frac{37297007}{7} - 2108^2 = 949{,}6$$

$$r = \frac{3136602 - 1238{,}7 \times 2108}{709{,}7 \times 949{,}6} = 0{,}78$$

Значение линейного коэффициента корреляции свидетельствует о наличии достаточно тесной связи между признаками.

ГЛАВА 6. РЯДЫ ДИНАМИКИ

6.1. Понятие о рядах динамики. Сопоставимость уровней ряда

Социально-экономические явления общественной жизни находятся в непрерывном развитии. Их изменение во времени статистика изучает при помощи построения и анализа рядов динамики.

Ряд динамики – ряд расположенных в хронологической последовательности значений статистических показателей.

Ряд динамики состоит из двух элементов:

1) момент или период времени, к которому относятся приводимые статистические данные;

2) статистические показатели, которые характеризуют изучаемый объект на определенный момент или за указанный период времени.

Статистические показатели, характеризующие изучаемый объект, называются ***уровнями ряда***. Уровни ряда могут быть представлены абсолютными, относительными и средними величинами.

По времени, отраженному в динамических рядах, они разделяются на моментные и интервальные.

Моментный ряд динамики – ряд, уровни которого характеризуют состояние явления на определенные даты (моменты времени).

Интервальный ряд динамики – ряд, уровни которого характеризуют размер явления за конкретный период времени (год, квартал, месяц).

Условием правильного формирования рядов динамики является сопоставимость уровней, образующих ряд. Основными требованиями сопоставимости уровней являются:

1) ***Одинаковая методология исчисления уровней ряда***. Например, если до 2000 г. на предприятии производительность труда определялась на одного рабочего, а после 2000 г. – на одного работника промышленно-производственного персонала (ППП), то для динамического анализа уровни производительности, рассчитанные до 2000 г., необходимо пересчитать по новой методологии.

2) ***Одинаковая полнота охвата различных частей явления***. Например, при характеристике динамики численности студентов высших учебных заведений по годам нельзя в одни годы учитывать

только студентов дневного отделения, а в другие – численность студентов всех форм обучения.

3) ***Одинаковая продолжительность периодов, к которым относятся уровни***. Например, нельзя строить ряд, где одни уровни являются месячными показателями, а другие – квартальными или годовыми.

4) ***Сопоставимость единиц измерения***. С этим часто приходится сталкиваться при учете продукции в натуральном и стоимостном выражении.

Например, данные о производстве ткани могут быть выражены в погонных метрах и в квадратных; количество произведенного молока – в литрах и килограммах.

При проведении к сопоставимому виду продукции, измеренной в стоимостных показателях, трудность заключается в том, что, во-первых, с течением времени происходит непрерывное изменение цен, а во-вторых, существует несколько видов цен. Для характеристики изменения объема продукции должно быть устранено влияние изменения цен. Поэтому на практике количество продукции, произведенной в разные периоды, оценивают в ценах одного и того же базисного периода, которые называют фиксированными, или сопоставимыми.

В ряде случаев несопоставимость рядов динамики может быть устранена путем проведения смыкания рядов динамики.

Смыкание рядов динамики – объединение в один ряд двух или несколько рядов, уровни которых исчислены по разной методологии или в разных границах. При этом для осуществления смыкания необходимо, чтобы для одного из периодов имелись данные, исчисленные по разной методологии.

Например, известны данные о выпуске продукции, оцененной за разные периоды в различных ценах (табл. 6.1.1).

Показатели объема выпуска продукции за 2001-2003 гг. несопоставимы с показателями объема выпуска продукции за 2003-2005 гг., так как рассчитаны в разных ценах. Для приведения этой информации к сопоставимому виду определяется коэффициент соотношения двух уровней по формуле (6.1.1):

Таблица 6.1.1

Динамика объема выпуска продукции предприятия за 2001-2005 гг.

Объем выпуска продукции, млн руб.	**2001 г.**	**2002 г.**	**2003 г.**	**2004 г.**	**2005 г.**
В ценах 2001 г.	215	238	250	-	-

В ценах 2004 г.	-	-	300	380	415
Сопоставимый ряд	$215\times1{,}2=258$	$238\times1{,}2=285{,}6$	300	380	415

Показатели объема выпуска продукции за 2001-2003 гг. несопоставимы с показателями объема выпуска продукции за 2003-2005 гг., так как рассчитаны в разных ценах. Для приведения этой информации к сопоставимому виду определяется коэффициент соотношения двух уровней по формуле (6.1.1):

$$Кс = \frac{\text{уровень явления в новых ценах}}{\text{уровень явления в старых ценах}}. \qquad (6.1.1)$$

$$Кс = \frac{300}{250} = 1{,}2$$

Умножая на полученный коэффициент данные за 2001 и 2002 гг., приводим их в сопоставимый вид с последующими уровнями. Сомкнутый ряд динамики показан в нижней строке таблицы.

6.2. Показатели ряда динамики и методы их определения

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).

Методы расчета показателей динамики 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_0 – уровень, принятый за постоянную базу сравнения (часто начальный уровень).

Методы расчета показателей динамики представлены в таблице 6.2.1.

Таблица 6.2.1

Методика расчета показателей ряда динамики

Наименование показателя	Метод расчета	
	С переменной базой (цепные)	**С постоянной базой (базисные)**
Абсолютный прирост (Δ)	$\Delta_ц = y_i - y_{i-1}$	$\Delta_б = y_i - y_0$
Коэффициент роста (Кр)	$Кр = \dfrac{y_i}{y_{i-1}}$	$Кр = \dfrac{y_i}{y_0}$
Темп роста (Тр), %	$Тр = \dfrac{y_i}{y_{i-1}} \times 100$ $Тр = Кр \times 100$	$Тр = \dfrac{y_i}{y_0} \times 100$ $Тр = Кр \times 100$
Темп прироста (Тпр), %	$Т\ пр = (Кр - 1) \times 100$ $Тпр = \dfrac{y_i - y_{i-1}}{y_{i-1}} \times 100$ $Тпр = Тр - 100$	$Т\ пр = (Кр - 1) \times 100$ $Тпр = \dfrac{y_i - y_0}{y_0} \times 100$ $Тпр = Тр - 100$
Абсолютное значение 1% прироста (А)	$А = \dfrac{y_i - y_{i-1}}{Тпр} = \dfrac{\Delta_ц}{Тпр}$ $А = \dfrac{y_{i-1}}{100}$	$А = \dfrac{\Delta_б}{Тпр} = \dfrac{y_0}{100}$

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, показывающий, 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.

Обобщающей характеристикой динамики исследуемого явления являются средние показатели динамики: средний уровень ряда динамики, средний абсолютный прирост, средний темп роста и прироста.

Метод расчета средних показателей динамики представлен в таблице 6.2.2.

Таблица 6.2.2

Методика расчета средних показателей динамики

Наименование показателя	**Метод расчета**	
	С равноотстоящими уровнями во времени	**С неравноотстоящими уровнями во времени**
1. Средний уровень ряда: – для интервального ряда	$\bar{y} = \frac{\sum_{i=1}^{n} y_i}{n}$	$\bar{y} = \frac{\sum_{i=1}^{n} y_i t_i}{\sum_{i=1}^{n} t_i}$
– для моментного ряда	$\bar{y} = \frac{\frac{1}{2} y_1 + y_2 + \ldots + y_{n-1} + \frac{1}{2} y_n}{n-1}$	$\bar{y} = \frac{\sum_{i=1}^{n} (y_i + y_{i+1}) t_i}{2 \sum_{i=1}^{n} t_i}$
2. Средний абсолютный прирост	$\bar{\Delta} = \frac{y_n - y_0}{n-1}$	
3. Средний коэффициент роста	$\bar{K}_р = \sqrt[n]{K_1 K_2 \ldots K_n}$ $\bar{K}_р = \sqrt[n-1]{\frac{y_n}{y_0}}$	
4. Средний темп роста	$\bar{T}_р = \bar{K}_р \times 100$	
5. Средний темп прироста, %	$\bar{T}_{пр} = \bar{T}_р - 100$ $\bar{T}_{пр} = (\bar{K}_р - 1) \times 100$	

При написании формул приняты следующие условные обозначения:

$y_1, y_2, \ldots y_n$ – уровни ряда динамики;

n – число уровней ряда;

t – продолжительность периода, в течение которого уровень не изменялся.

6.3. Выявление основной тенденции развития в рядах динамики

При анализе рядов динамики важной задачей статистики является определение ***основной тенденции развития (тренда)***. С этой целью ряды динамики подвергаются обработке методами укрупнения интервалов, скользящей средней и аналитического выравнивания.

Метод укрупнения интервалов заключается в том, что первоначальный ряд динамики преобразуется в другой ряд, показатели которого относятся к большим по продолжительности периодам времени и более четко выражают закономерность изменения показателя.

Метод скользящей средней заключается в замене абсолютных данных средними арифметическими за определенные периоды. Расчет средних ведется способом скольжения, то есть постепенным исключением из принятого периода скольжения первого уровня и включением следующего.

Сглаживание можно проводить по любому числу членов, но чаще применяется нечетное число.

Метод аналитического выравнивания заключается в нахождении уравнения, наиболее адекватно отражающего тенденцию развития исследуемого показателя.

При выборе уравнения можно руководствоваться следующими указаниями:

1) если относительно стабильны абсолютные приросты (первые разности уровней приблизительно равны), сглаживание может быть выполнено по ***прямой***;

2) если абсолютные приросты равномерно увеличиваются (вторые разности уровней приблизительно равны), можно принять ***параболу второго порядка***;

3) при ускоренно возрастающих (замедляющихся) абсолютных приростах зависимости отражаются ***параболой третьего порядка***;

4) при относительно стабильных темпах роста принимают ***показательную функцию***.

После выбора вида уравнения рассчитываются параметры уравнения методом наименьших квадратов путем решения соответствующей системы нормальных уравнений.

Вычислительный процесс нахождения параметров уравнения может быть упрощен, если ввести обозначения периодов времени с помощью натуральных чисел (t), с тем чтобы $\Sigma t = 0$.

Если число уровней в ряду динамики нечетное, то временные даты (t) обозначаются следующим образом (табл. 6.3.1):

Таблица 6.3.1

Обозначение временных дат при нечетном количестве уровней ряда динамики

Временные 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 количество уровней в ряду динамики четное, то обозначения временных дат (t) принимают следующий вид (табл. 6.3.2).

Таблица 6.3.2

Обозначение временных дат при четном количестве уровней ряда динамики

Временные 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

В таблице 6.3.3 приведены различные виды трендовых моделей, наиболее часто используемых для аналитического выравнивания.

Таблица 6.3.3

Виды трендовых моделей

№ п/п	Наименование функции	Вид функции	Система нормальных уравнений для нахождения параметров уравнения
1	Линейная	$\hat{y}_t = a_0 + a_1 t$	$\begin{cases} \sum y = a_0 n + a_1 \sum t \\ \sum yt = a_0 \sum t + a_1 \sum t^2 \end{cases}$
2	Парабола второго порядка	$\hat{y}_t = a_0 + a_1 t + a_2 t^2$	$\begin{cases} \sum y = a_0 n + a_1 \sum t + a_2 \sum t^2 \\ \sum yt = a_0 \sum t + a_1 \sum t^2 + a_2 \sum t^3 \\ \sum yt^2 = a_0 \sum t^2 + a_1 \sum t^3 + a_2 \sum t^4 \end{cases}$
3	Парабола третьего порядка	$\hat{y}_t = a_0 + a_1 t + a_2 t^2 + a_3 t^3$	$\begin{cases} \sum y = a_0 n + a_1 \sum t + a_2 \sum t^2 + a_3 \sum t^3 \\ \sum yt = a_0 \sum t + a_1 \sum t^2 + a_2 \sum t^3 + a_3 \sum t_4 \\ \sum yt^2 = a_0 \sum t^2 + a_1 \sum t^3 + a_2 \sum t^4 + a_3 \sum t^5 \\ \sum yt^3 = a_0 \sum t^3 + a_1 \sum t^4 + a_2 \sum t^5 + a_3 \sum t^6 \end{cases}$
4	Показательная	$\hat{y}_t = a_0 a_1^t$	$\begin{cases} \sum \lg y = n \lg a_0 + \lg a_1 \sum t \\ \sum \lg y \cdot t = \lg a_0 \sum t + \lg a_1 \sum t^2 \end{cases}$
5	Гиперболическая	$\hat{y}_t = a_0 + a_1 \frac{1}{t}$	$\begin{cases} \sum y = a_0 n + a_1 \sum \frac{1}{t} \\ \sum y \frac{1}{t} = a_0 \sum \frac{1}{t} + a_1 \sum \frac{1}{t^2} \end{cases}$

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.

6.4. Статистическое изучение сезонных колебаний

Сезонные колебания – это сравнительно устойчивые внутригодичные колебания уровня явления. Основной задачей статистики в изучении сезонных колебаний является их выявление и измерение.

Наличие сезонных колебаний выявляют с помощью графического метода.

Для количественной характеристики сезонных колебаний используются индексы сезонности, расчет которых выполняется двумя методами в зависимости от характера динамики:

1) если годовой уровень явления остается относительно неизменным, то индексы сезонности определяются по формуле (6.4.1):

$$J_c = \frac{\bar{y}_i}{\bar{\bar{y}}_0} \times 100, \qquad (6.4.1)$$

где $\bar{y}_i$ – средняя из фактических уровней одноименных месяцев;

$\bar{y}_0$ – общая средняя за исследуемый период.

2) если уровни сезонного явления имеют тенденцию к развитию (повышаются или понижаются), то индексы сезонности рассчитываются по формуле (6.4.2):

$$J_c = \frac{\bar{y}_i}{\bar{y}'_i} \times 100, \qquad (6.4.2)$$

где $\bar{y}'_i$ – средняя из сглаженных (выровненных) уровней одноименных месяцев.

6.5. Решение типовых задач

Задача 1. Имеются следующие данные о выпуске легковых автомобилей в России (табл. 6.5.1).

Таблица 6.5.1

Динамика производства легковых автомобилей в регионе за 2002-2005 гг.

Год	2002	2003	2004	2005
Произведено автомобилей, тыс. шт.	92	88	103	98

Определить абсолютные, относительные и средние показатели ряда динамики.

Решение:

Руководствуясь методикой расчета показателей ряда динамики (таблица 6.2.1), рассчитаем абсолютные и относительные показатели динамики. Расчет сведен в таблицу 6.5.2.

Расчет средних показателей ряда динамики осуществляется в соответствии с методикой, приведенной в таблице 6.2.2:

1) средний уровень интервального ряда динамики:

$$\bar{y} = \frac{92 + 88 + 103 + 98}{4} = 95{,}25 \text{ тыс. шт.},$$

2) средний абсолютный прирост:

$$\bar{\Delta} = \frac{98 - 92}{4 - 1} = \frac{6}{3} = 2 \text{ тыс. шт.},$$

3) средний коэффициент роста:

$$\bar{К}_р = \sqrt[n-1]{\frac{y_n}{y_0}} = \sqrt[3]{\frac{98}{92}} = \sqrt[3]{1{,}065} = 1{,}021,$$

4) средний темп роста:

$$\bar{Т}_р = \bar{К}_р \times 100 = 1{,}021 \times 100 = 102{,}1\%,$$

5) средний темп прироста:

$$\bar{Т}_{пр} = \bar{Т}_р - 100 = 102{,}1 - 100 = 2{,}1\%.$$

Таблица 6.5.2

Абсолютные и относительные показатели динамики производства автомобилей

Показатель		**Год**			
		2002	**2003**	**2004**	**2005**
1. Абсолютный прирост (Δ), тыс. шт.	переменная база	-	$\Delta_ц^{03/02} = 88 - 92 = -4$	$\Delta_ц^{04/03} = 103 - 88 = 15$	$\Delta_ц^{05/04} = 98 - 103 = -5$
	постоянная база	-	$\Delta_б^{03/02} = 88 - 92 = -4$	$\Delta_б^{04/02} = 103 - 92 = 11$	$\Delta_б^{05/02} = 98 - 92 = 6$
2. Коэффициент роста, Кр	переменная база	-	$Кр_ц^{03/02} = \frac{88}{92} = 0{,}9565$	$Кр_ц^{04/03} = \frac{103}{88} = 1{,}1705$	$Кр_ц^{05/04} = \frac{98}{103} = 0{,}9515$
	постоянная база	-	$Кр_б^{03/02} = \frac{88}{92} = 0{,}9565$	$Кр_б^{04/02} = \frac{103}{92} = 1{,}1196$	$Кр_б^{05/02} = \frac{98}{92} = 1{,}0652$
3. Темп роста, (Тр), %	переменная база	-	$Тр_ц^{03/02} = 0{,}9565 \times 100 = 95{,}65$	$Тр_ц^{04/03} = 1{,}1705 \times 100 = 117{,}05$	$Тр_ц^{05/04} = 0{,}9515 \times 100 = 95{,}15$
	постоянная база	-	$Тр_б^{03/02} = 0{,}9565 \times 100 = 95{,}65$	$Тр_б^{04/02} = 1{,}1196 \times 100 = 111{,}96$	$Тр_б^{05/02} = 1{,}0652 \times 100 = 106{,}52$
4. Темп прироста (Тпр), %	переменная база	-	$Тпр_ц^{03/02} = 95{,}65 - 100 = -4{,}35$	$Тпр_ц^{04/03} = 117{,}05 - 100 = 17{,}05$	$Тпр_ц^{05/04} = 95{,}15 - 100 = -4{,}85$
	постоянная база	-	$Тпр_б^{03/02} = 95{,}65 \quad 100 = -4{,}5$	$Тпр_б^{04/02} = 111{,}96 - 100 = 11{,}96$	$Тпр_б^{05/02} = 106{,}52 - 100 = 6{,}52$
5. Абсолютное значение 1% рироста (А), тыс. шт.	переменная база	-	$А^{03/02} = \frac{92}{100} = 0{,}92$	$А^{04/03} = \frac{88}{100} = 0{,}88$	$А^{05/04} = \frac{103}{100} = 1{,}03$

Задача 2. Известно, что с 1-го по 15-е число месяца в организации работали 20 человек, с 16-го по 25-е – 27 человек, а с 26-го по 30-е – 30 человек.

Определить среднесписочную численность работников за месяц.

Решение:

По условию задачи имеем интервальный ряд динамики с неравными интервалами, поэтому средний уровень ряда необходимо определить по формуле средней арифметической взвешенной:

$$\bar{y} = \frac{\sum_{i=1}^{n} y_i t_i}{\sum_{i=1}^{n} t_i} = \frac{20 \times 15 + 27 \times 10 + 30 \times 5}{30} = 24 \text{ чел.}$$

Задача 3. Имеются следующие данные о запасах дизельного топлива в фермерском хозяйстве (табл. 6.5.3):

Таблица 6.5.3

Динамика запасов дизельного топлива в фермерском хозяйстве в 2005 г.

Период	**На 01.01. 2005 г.**	**На 01.03. 2005 г.**	**На 01.04. 2005 г.**	**На 01.08. 2005 г.**	**На 01.01. 2006 г.**
Запасы топлива, т.	40	60	100	10	30

Определить среднегодовой запас дизельного топлива в фермерском хозяйстве за 2005 г.

Решение:

По условию задачи имеем моментный ряд динамики с неравноотстоящими уровнями, поэтому средний уровень ряда определяется по формуле средней хронологической взвешенной:

$$\bar{y} = \frac{\sum_{i=1}^{n} (y_i + y_{i+1}) t_i}{2 \sum_{i=1}^{n} t_i} = \frac{(40 + 60)2 + (60 + 100)1 + (100 + 10)4 + (10 + 30)5}{2(2 + 1 + 4 + 5)} = 41{,}67 \text{ т.}$$

Задача 4. Объемы перевозки грузов по автотранспортному предприятию в динамике характеризуются следующими данными (табл. 6.5.4):

Таблица 6.5.4

Динамика объемов перевозки грузов по предприятию за 1999-2006 гг.

Показатель	**Годы**							
	1999	**2000**	**2001**	**2002**	**2003**	**2004**	**2005**	**2006**
Перевезено груза, тыс. т	360	381	401	422	443	463	485	505

Провести аналитическое выравнивание с последующей экстраполяцией до 2008 г.

Решение:

Для определения формы тренда и расчета его параметров составляется вспомогательная таблица 6.5.5.

Так как первые разности (графа 3 таблицы 6.5.5) приблизительно равны между собой, это позволяет в виде модели принять уравнение прямой $\hat{y}_t = a_0 + a_1 t$.

Для нахождения параметров a_0 и a_1 используется система нормальных уравнений (см. табл. 6.3.3).

$$\begin{cases} \sum y = a_0 n + a_1 \sum t \\ \sum yt = a_0 \sum t + a_1 \sum t^2 \end{cases}$$

Для упрощения системы уравнений показатели времени обозначены так, чтобы $\Sigma t = 0$. Тогда система примет следующий вид:

$$\begin{cases} \sum y = a_0 n \\ \sum yt = a_1 \sum t^2 \end{cases}$$

Откуда определяем параметры:

$$a_0 = \frac{\Sigma y}{n} = \frac{3460}{8} = 432{,}5 \text{ тыс. т}$$

$$a_1 = \frac{\Sigma yt}{\Sigma t^2} = \frac{1742}{168} = 10{,}4 \text{ тыс. т}$$

Соответственно, модель тренда будет:

$$\hat{y}_t = 432{,}5 + 10{,}4t.$$

Таблица 6.5.5

Вспомогательная таблица для определения основной тенденции развития ряда динамики

Год	Объем перевозок, тыс. т (у)	Первые разности	t	t^2	yt	Теоретический уровень, $\hat{y}_t$
1999	360	-	-7	49	-2520	359,7
2000	381	21	-5	25	-1905	380,5
2001	401	20	-3	9	-1203	401,3
2002	422	21	-1	1	-422	422,1
2003	443	21	1	1	443	442,9
2004	463	20	3	9	1389	463,7
2005	485	22	5	25	2425	484,5
2006	505	20	7	79	3535	505,3
Итого	3460	-	0	168	1742	3460

Подставив в это уравнение значение t (графа 4 таблицы 6.5.5), получим выравненные теоретические значения $\hat{y}_t$ (графа 7).

Для определения объема перевозок в 2008 г. подставляем в найденное уравнение тренда значение t, соответствующее 2008 г.

$\hat{y}_{2008} = 432{,}5 + 10{,}4 \times 11 = 546{,}9$ тыс. т.

Задача 5. Производство яиц в акционерном обществе за три года характеризуется данными таблицы 6.5.6.

Таблица 6.5.6

Динамика производства яиц на предприятии за 2004-2006 гг., тыс. шт.

Годы	Месяцы												Итого
	I	II	III	VI	V	VI	VII	VIII	XI	X	XI	XII	
2004	10,2	15,2	17,3	19,4	21,2	26,1	28,3	21,4	22,1	14,6	9,5	12,4	217,7
2005	9,7	16,1	14,8	22,7	25,4	28,2	25,8	23,3	20,7	15,2	8,6	12,9	223,4
2006	11,8	14,4	15,6	16,5	29,1	25,2	23,5	23,6	18,2	16,3	13,3	14,6	221,1

Определить индексы сезонности.

Решение:

Индекс сезонности определяется по формуле (6.4.1):

$$J_c = \frac{\bar{y}_i}{\bar{\bar{y}}_0} \times 100 .$$

Применяя формулу средней арифметической простой, определим средние месячные уровни за три года:

Январь: $\bar{y}_1 = \frac{10,2 + 9,7 + 11,8}{3} = 10,6$ тыс. шт.

Февраль: $\bar{y}_2 = \frac{15,2 + 16,1 + 14,4}{3} = 15,2$ тыс. шт.

Март: $\bar{y}_3 = \frac{17,3 + 14,8 + 15,6}{3} = 15,9$ тыс. шт. и т. д.

Далее определяем общую среднюю за исследуемый период:

$\bar{y}_0 = \frac{217,7 + 223,4 + 222,1}{12 + 12 + 12} = 18,42$ тыс. шт.

Расчет индексов сезонности представлен в таблице 6.5.7.

Анализ данных таблицы 6.5.7 позволяет сделать следующие выводы:

1) производство яиц характеризуется резко выраженной сезонностью;

2) яйценоскость по отдельным месяцам года отклоняется от среднемесячной на 42-44%;

3) наименьшей яйценоскостью характеризуется ноябрь (57 %), наибольшей – июнь (143,9 %).

Таблица 6.5.7

Вспомогательная таблица для определения сезонной неравномерности производства яиц на предприятии за 2004-2006 гг.

Месяц	**Производство яиц, тыс. шт.**				**Индексы сезонности (J_c), %**
	2004 г.	**2005 г.**	**2006 г.**	**Среднемесячная за три года ($\bar{y}_i$)**	
I	10,2	9,7	11,8	10,6	57,5
II	15,2	16,1	14,4	15,2	82,5
III	17,3	14,8	15,6	15,9	86,3
VI	19,4	22,7	16,5	19,5	105,9
V	21,2	25,4	29,1	25,2	136,8
VI	26,1	28,2	25,2	26,5	143,8
VII	28,3	25,8	23,5	25,9	140,6
VIII	21,4	23,3	23,6	22,8	123,8
XI	22,1	20,7	18,2	20,3	110,2
X	14,6	15,2	16,3	15,4	83,5
XI	9,5	8,6	13,3	10,5	57,0
XII	12,4	12,9	14,6	13,3	72,1
Итого	217,7	223,4	222,1	221,1	1200,0
В среднем	18,14	18,61	18,51	18,42	100

Для наглядного представления сезонной волны исчисленные индексы сезонности представлены графически (рис. 6.5.1).

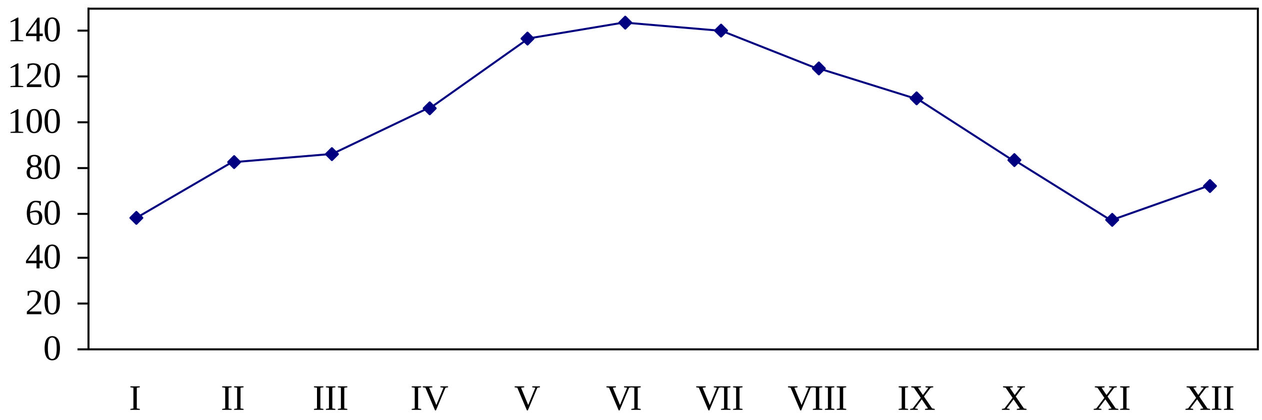

Рис. 6.5.1 Сезонная волна яйценоскости (изменение индексов сезонности в течение года)

ГЛАВА 7. ИНДЕКСЫ

7.1. Понятие об индексах, их классификация

Индекс – относительная величина, характеризующая соотношение значений определенного показателя во времени, пространстве или с планом.

С помощью индексов решаются следующие основные задачи:

1) характеристика общего изменения сложного экономического показателя и отдельных его элементов;

2) измерение влияния факторов на общую динамику сложного показателя, включая характеристику влияния изменения структуры явления.

Индексы классифицируются по следующим признакам:

1) ***По содержанию изучаемых объектов*** различают индексы качественных и количественных показателей.

К индексам количественных показателей относятся индексы физического объема производства продукции, индекс национального дохода и др. Все индексируемые показатели этих индексов являются объемными, так как характеризуют ***общий, суммарный размер*** того или иного явления.

К индексам качественных явлений относятся индексы цен, индексы средней заработной платы, себестоимости, производительности труда и др.

Индексируемые показатели этих индексов характеризуют уровень явления в расчете на ту или иную единицу совокупности: цена за единицу продукции, себестоимость единицы продукции, выработка в единицу времени (или на одного работника), заработная плата одного работника, урожайность с одного гектара и т. д. Качественные показатели измеряют не общий объем, а интенсивность, эффективность явления или процесса.

2) ***По степени охвата*** элементов совокупности различают индивидуальные и сводные (общие) индексы.

Индивидуальные индексы характеризуют изменение одного элемента совокупности (например, изменение объема выпуска телевизоров определенной марки).

Сводные индексы характеризуют изменение сложного явления в целом. Под сложным явлением понимают такую статистическую совокупность, отдельные элементы которой непосредственно не подлежат суммированию (физический объем продукции, включающий разноименные товары, цены на разные группы продуктов).

Если индексы охватывают не все элементы сложного явления, а лишь часть, то их называют ***групповыми*** индексами*,* ***или субиндексами***. Например, общий индекс характеризует динамику объема промышленной продукции, а к субиндексам в этом случае могут быть отнесены индексы продукции по отдельным отраслям промышленности.

3) ***По способу расчета*** индексы делятся на базисные и цепные.

Базисные индексы получают сопоставлением текущих уровней с уровнем периода, принятого за базу сравнения.

Цепные индексы получают сопоставлением текущих уровней с предшествующим.

4) В зависимости ***от методологии расчета*** различают агрегатные и средние из индивидуальных индексов.

Агрегатный индекс – общий индекс, полученный путем сопоставления итогов, выражающих величину сложного показателя в отчетном и базисном периодах при помощи соизмерителей.

Суть ***среднего из индивидуальных*** индекса состоит в том, что по отдельным видам продукции рассчитываются индивидуальные индексы, а затем из них рассчитывается средний индекс.

Индексный метод имеет свою символику. Обычно используются следующие обозначения индексируемых величин:

g – количество продукции в натуральном выражении;

p – цена единицы продукции;

z – себестоимость единицы продукции;

t – затраты времени на производство продукции (трудоемкость);

w – выработка продукции в стоимостном выражении на одного работника или в единицу времени.

Каждый уровень, который характеризует отчетный период, обозначается в индексной системе через 1, а уровень, характеризующий базисный период, – через 0.

7.2. Индексы количественных показателей

1. ***Индивидуальный индекс физического объема выпуска продукции*** (i_q) характеризует изменение выпуска одного вида продукции и определяется по формуле (7.2.1):

$$i_q = \frac{q_1}{q_0}, \qquad (7.2.1)$$

где q_1 и q_0 – количество продукции данного вида в натуральном выражении соответственно в текущем и базисном периодах.

2. ***Агрегатный индекс физического объема продукции*** характеризует изменение выпуска всей продукции и определяется по формуле (7.2.2):

$$J_q = \frac{\Sigma q_1 p}{\Sigma q_0 p}, \tag{7.2.2}$$

где q_1 и q_0 – количество произведенных единиц отдельных видов продукции соответственно в отчетном и базисном периодах;

p – сопоставимая цена единицы отдельного вида продукции (соизмеритель).

Если за коэффициент соизмерения принять цены базисного периода (p_0), то индекс физического объема определяют по методу ***Ласпейреса*** (формула (7.2.3)):

$$J_q^{Л} = \frac{\Sigma q_1 p_0}{\Sigma q_0 p_0}. \tag{7.2.3}$$

Если в качестве соизмерителей использовать цены отчетного периода (p_1), то индекс физического объема определяют по методу ***Пааше*** (формула (7.2.4)):

$$J_q^{П} = \frac{\Sigma q_1 p_1}{\Sigma q_0 p_1}. \tag{7.2.4}$$

3. ***Агрегатный индекс затрат на выпуск всей продукции*** характеризует изменение общей суммы затрат на выпуск продукции и определяется по формуле (7.2.5):

$$J_{qz} = \frac{\Sigma q_1 z_1}{\Sigma q_0 z_0}, \tag{7.2.5}$$

где $q_1 z_1$ и $q_0 z_0$ – затраты на выпуск продукции каждого вида соответственно в отчетном и базисном периодах.

4. ***Агрегатный индекс стоимости продукции*** (товарооборота) определяется по формуле (7.2.6):

$$J_{qp} = \frac{\Sigma q_1 p_1}{\Sigma q_0 p_0}. \tag{7.2.6}$$

7.3. Индексы качественных показателей

1. ***Индивидуальные индексы цен, себестоимости, затрат рабочего времени на единицу продукции*** характеризуют изменение цен, себестоимости, затрат рабочего времени по каждому виду продукции и определяются по формулам (7.3.1)-(7.3.3):

$$i_p = p_1 : p_0, \tag{7.3.1}$$

$$i_z = z_1 : z_0, \tag{7.3.2}$$

$$i_t = t_1 : t_0, \tag{7.3.3}$$

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.

2. ***Агрегатный индекс цен*** – в общем виде определяется по формуле (7.3.4):

$$J_p = \frac{\Sigma q p_1}{\Sigma q p_0}. \tag{7.3.4}$$

Индексируемой величиной является цена (p), количество продукции (q) – соизмерителем.

Если за коэффициент соизмерения принять количество продукции каждого вида в базисном периоде (q_0), то агрегатный индекс цен определяют по методу ***Ласпейреса*** (формула (7.3.5)):

$$J_p^{Л} = \frac{\Sigma q_0 p_1}{\Sigma q_0 p_0}. \tag{7.3.5}$$

Если за коэффициент соизмерения принять количество продукции каждого вида в отчетном периоде (q_1), то агрегатный индекс цен определяют по методу ***Пааше*** (формула (7.3.6)):

$$J_p^{П} = \frac{\Sigma q_1 p_1}{\Sigma q_1 p_0}. \tag{7.3.6}$$

3. ***Агрегатные индексы себестоимости и затрат рабочего времени на единицу продукции*** определяются по формулам (7.3.7)-(7.3.8):

$$J_z = \frac{\Sigma q_1 z_1}{\Sigma q_1 z_0}, \tag{7.3.7}$$

$$J_t = \frac{\Sigma t_1 q_1}{\Sigma t_0 q_1}. \qquad (7.3.8)$$

4. ***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 определяется по формуле (7.3.9):

$$J_p = \frac{\Sigma i_p q_0 p_0}{\Sigma p_0 q_0}, \qquad (7.3.9)$$

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 определяется по формуле (7.3.10):

$$J_p = \frac{\Sigma q_1 p_1}{\Sigma \frac{1}{i_p} q_1 p_1}, \qquad (7.3.10)$$

где $q_1 p_1$ – стоимость продукции каждого вида в текущем периоде.

7.4. Индексы средних величин

Индексы средних величин используются, когда необходимо определить изменение средней величины индексируемого показателя для какой-либо однородной совокупности. Например, по совокупности предприятий, выпускающих одинаковую продукцию, но с разным уровнем себестоимости, можно показать изменение средней себестоимости; при реализации одной и той же продукции на разных рынках – изменение средней цены.

К индексам средних величин относятся:

1. ***Индекс переменного состава*** отражает динамику среднего показателя за счет изменения индексируемой величины (X) и за счет изменения весов (f) и определяется по формуле (7.4.1):

$$J^{ПС} = \frac{\bar{x}_1}{\bar{x}_0} = \frac{\Sigma x_1 f_1}{\Sigma f_1} : \frac{\Sigma x_0 f_0}{\Sigma f_0}. \qquad (7.4.1)$$

2. ***Индекс фиксированного состава*** отражает динамику среднего показателя лишь за счет изменения индексируемой величины x при фиксировании весов на уровне, как правило, отчетного периода (формула (7.4.2)):

$$J^{ФС} = \frac{\Sigma x_1 f_1}{\Sigma f_1} : \frac{\Sigma x_0 f_1}{\Sigma f_1} = \frac{\Sigma x_1 f_1}{\Sigma x_0 f_1}. \quad (7.4.2)$$

3. ***Индекс структурных сдвигов*** характеризует динамику среднего показателя за счет изменения весов при фиксировании индексируемой величины на уровне базисного периода (формула (7.4.3)):

$$J^{СТР} = \frac{\Sigma x_0 f_1}{\Sigma f_1} : \frac{\Sigma x_0 f_0}{\Sigma f_0}. \quad (7.4.3)$$

Между данными индексами существует взаимосвязь, представленная формулой (7.4.4):

$$J^{ПС} = J^{ФС} \times J^{СТР}. \quad (7.4.4)$$

Используя символику индексного метода, можно записать индексы переменного, фиксированного состава и структурных сдвигов для конкретных экономических явлений. Например, в индексе себестоимости индексируемой величиной будет себестоимость единицы продукции (z), а весами (частотой) – количество продукции.

Тогда индекс себестоимости переменного состава будет определяться по формуле (7.4.5):

$$J^{ПС} = \frac{\Sigma z_1 q_1}{\Sigma q_1} : \frac{\Sigma z_0 q_0}{\Sigma q_0}. \quad (7.4.5)$$

Индекс себестоимости фиксированного (постоянного) состава рассчитывается по формуле (7.4.6):

$$J^{ФС} = \frac{\Sigma z_1 q_1}{\Sigma q_1} : \frac{\Sigma z_0 q_1}{\Sigma q_1} = \frac{\Sigma z_1 q_1}{\Sigma z_0 q_1}. \quad (7.4.6)$$

Индекс себестоимости структурных сдвигов определяется по формуле (7.4.7):

$$J^{СТР} = \frac{\Sigma z_0 q_1}{\Sigma q_1} : \frac{\Sigma z_0 q_0}{\Sigma q_0}. \quad (7.4.7)$$

Основные формулы для расчета общих индексов приведены в таблице 7.4.1.

Таблица 7.4.1

Основные формулы для расчета индексов

Наименование индекса		Формула расчета индексов		
		Индивидуальный индекс	**Агрегатный индекс**	**Средний индекс**
Индекс физического объема продукции	с базисными весами	$i_q = \frac{q_1}{q_0}$	$J_q = \frac{\Sigma q_1 p_0}{\Sigma q_0 p_0}$	$J_q = \frac{\Sigma i_q q_0 p_0}{\Sigma q_0 p_0}$
	с отчетными весами		$J_q = \frac{\Sigma q_1 p_1}{\Sigma q_0 p_1}$	$J_q = \frac{\Sigma q_1 p_1}{\Sigma \frac{1}{i_q} q_1 p_1}$
Индекс цен	с базисными весами	$i_p = \frac{p_1}{p_0}$	$J_p = \frac{\Sigma p_1 q_0}{\Sigma p_0 q_0}$	$J_p = \frac{\Sigma i_p q_0 p_0}{\Sigma q_0 p_0}$
	с отчетными весами		$J_p = \frac{\Sigma p_1 q_1}{\Sigma p_0 q_1}$	$J_p = \frac{\Sigma q_1 p_1}{\Sigma \frac{1}{i_p} q_1 p_1}$
Индекс себестоимости продукции		$i_z = \frac{z_1}{z_0}$	$J_z = \frac{\Sigma z_1 q_1}{\Sigma z_0 q_1}$	$J_z = \frac{\Sigma z_1 q_1}{\Sigma \frac{1}{i_z} z_1 q_1}$
Индексы производительности труда		$i_w = \frac{t_1}{t_0}$	$J_w = \frac{\Sigma t_0 q_1}{\Sigma t_1 q_1}$	$J_w = \frac{\Sigma t_1 q_1}{\Sigma i_w t_1 q_1}$

7.5. Использование индексного метода в анализе взаимосвязи экономических явлений

Индексный метод не только характеризует динамику сложного явления, но и позволяет проанализировать влияние на него отдельных факторов.

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).

При этом в статистике руководствуются следующим правилом: ***если результативный показатель можно представить как произведение количественного и качественного факторов, то, определяя влияние количественного фактора на изменение результативного показателя, качественный фактор фиксируется на уровне базисного периода; если же определяется влияние качественного показателя, то количественный фактор фиксируется на уровне отчетного периода.***

Применяются два метода разложения общего индекса на частные:

1) метод обособленного изучения факторов;

2) метод взаимосвязанного изучения факторов (последовательно-цепной).

Поскольку в действительности явления взаимосвязаны, то основным методом следует считать последовательно-цепной, требующий правильного расположения факторов при построении модели результативного показателя (например, $A = a \times b$): на первом месте в модели ставится качественный фактор (a), затем – количественный (b).

Связь между экономическими показателями находит отражение и во взаимосвязи характеризующих их индексов, отраженной формулой (7.5.1):

$$J_{ab} = J_a \times J_b. \tag{7.5.1}$$

Общее изменение результирующего показателя выражается формулой (7.5.2):

$$J_{ab} = \frac{\Sigma a_1 b_1}{\Sigma a_0 b_0}. \tag{7.5.2}$$

Данный индекс показывает, как изменится результирующий показатель за счет изменения качественного и количественного факторов.

Изменение результирующего показателя за счет изменения качественного фактора определяется по формуле (7.5.3) на основе вышеизложенного правила:

$$J_a = \frac{\Sigma a_1 b_1}{\Sigma a_0 b_1}. \tag{7.5.3}$$

Изменение результирующего показателя за счет изменения количественного фактора определяется по формуле (7.5.4):

$$J_b = \frac{\Sigma a_0 b_1}{\Sigma a_0 b_0}. \tag{7.5.4}$$

Индексные системы могут применяться и для определения в абсолютном выражении изменения сложного явления за счет влияния отдельных факторов. Расчеты, связанные с определением в абсолютном выражении изменения результативного показателя за счет отдельных факторов, называют ***разложением абсолютного прироста (сокращения) по факторам***.

Так, рассмотренная выше индексная система двухфакторной связи может быть представлена в абсолютных величинах следующим образом (формула (7.5.5)):

$$a_1b_1 - a_0b_0 = (a_1 - a_0)b_1 + (b_1 - b_0)a_0. \qquad (7.5.5)$$

Рассмотрим построение взаимосвязанных индексов на примере индексов физического объема, цен и индекса стоимости продукции.

Произведение индекса цен на индекс физического объема продукции дает индекс стоимости продукции, то есть образует систему из трех индексов (формула (7.5.6)):

$$J_{pq} = J_p \times J_q. \qquad (7.5.6)$$

Изменение стоимости продукции за счет изменения количества продукции и цены единицы продукции определяется по формуле (7.5.2):

$$J_{pq} = \frac{\Sigma p_1q_1}{\Sigma p_0q_0}.$$

Разность между числителем и знаменателем этого индекса отражает абсолютный прирост (снижение) стоимости реализованной продукции в текущем периоде по сравнению с базисным:

$$\Delta pq = \Sigma p_1q_1 - \Sigma p_0q_0.$$

Изменение общей стоимости продукции за счет отдельных факторов:

1) изменение физического объема продукции:

– в индексной системе

$$J_q = \frac{\Sigma q_1p_0}{\Sigma q_0p_0},$$

– в абсолютном выражении

$$\Delta^q pq = \Sigma q_1p_0 - \Sigma q_0p_0.$$

2) изменение цен на продукцию:

– в индексной системе

$$J_p = \frac{\Sigma q_1p_1}{\Sigma q_1p_0},$$

– в абсолютном выражении

$$\Delta^p pq = \Sigma q_1 p_1 - \Sigma q_1 p_0 .$$

Общее абсолютное изменение результативного показателя составит алгебраическую сумму абсолютных изменений за счет отдельных факторов:

$$\Delta pq = \Delta^q pq + \Delta^p pq = \Sigma p_1 q_1 - \Sigma p_0 q_0 .$$

7.6. Решение типовых задач

Задача 1. Известны следующие данные о продаже и ценах на продукты на одном из рынков города (табл. 7.6.1):

Таблица 7.6.1

Динамика объемов производства и стоимости продукции на рынке города

Продукты	**Продано, тыс. ед.**		**Цена единицы продукции, руб.**	
	базисный период	**отчетный период**	**базисный период**	**отчетный период**
А	50	60	3	2,5
В	40	50	2	1,5
С	1,5	2	20	18

Определить:

1) агрегатный индекс физического объема продаж по методу Ласпейреса;

2) агрегатный индекс цен на продукты по методу Пааше;

3) абсолютную экономию (перерасход) денежных средств населения от изменения цен.

Решение:

Агрегатный индекс физического объема продаж по методу Ласпейреса определяем по формуле (7.3):

$$J_q^{Л} = \frac{\Sigma q_1 p_0}{\Sigma q_0 p_0} = \frac{60 \times 3 + 50 \times 2 + 2 \times 20}{50 \times 3 + 40 \times 2 + 1{,}5 \times 20} = \frac{320}{260} = 1{,}23 .$$

Агрегатный индекс цен на продукты по методу Пааше определяется по формуле (7.3.6):

$$J_p^{П} = \frac{\Sigma q_1 p_1}{\Sigma q_1 p_0} = \frac{60 \times 2{,}5 + 50 \times 1{,}5 + 2 \times 18}{60 \times 3 + 50 \times 2 + 2 \times 20} = \frac{261}{320} = 0{,}816 .$$

Абсолютная экономия (перерасход) денежных средств населения от изменения цен определяется как разность между числителем и знаменателем агрегатного индекса цен, рассчитанного по методу Пааше:

$$Э^{p}_{абс} = \Sigma q_1 p_1 - \Sigma q_1 p_0 = 261 - 320 = -59 \text{ тыс. руб.}$$

Задача 2. Известны данные о выпуске продукции по предприятию строительных пластмасс (табл. 7.6.2).

Таблица 7.6.2

Динамика выпуска продукции на предприятии

Вид продукции	**Выпуск продукции в I квартале, млн руб.**	**Изменение объема производства во II квартале в натуральном выражении, %**
Пленка	30	+ 10
Пеноплен	25	- 10
Линолеум	40	- 25

Определить индекс физического объема продукции.

Решение:

Из условия следует, что индивидуальные индексы по видам продукции имеют следующие значения:

$$i_{пленка} = 1{,}1; \; i_{пеноплен} = 0{,}9; \; i_{линолиум} = 0{,}75.$$

Используя формулу среднего арифметического объема продукции, получим:

$$J_q = \frac{\Sigma i_q q_0 p_0}{\Sigma q_0 p_0} = \frac{1{,}1 \times 30 + 0{,}9 \times 25 + 0{,}75 \times 40}{30 + 25 + 40} = 0{,}9.$$

Следовательно, объем производства в натуральном выражении во втором квартале по сравнению с первым уменьшился на 10%.

Задача 3. Известны данные о заработной плате работников организаций по трем отраслям экономики региона (табл. 7.6.3):

Таблица 7.6.3

Динамика среднемесячной заработной платы работников организаций ряда отраслей экономики региона

Отрасль экономики	Заработная плата, руб.		Число работников, чел.	
	Январь	Июль	Январь	Июль
Здравоохранение	6 000	7 000	2 400	1 600
Образование	5 500	6 200	2 100	2 000
Культура и искусство	5 100	5 900	1 500	1 400

Определить индекс заработной платы переменного состава; постоянного (фиксированного) состава и структурных сдвигов.

Решение:

Для определения индекса заработной платы переменного состава сначала определим среднюю заработную плату в январе и июле месяцах.

Обозначим заработную плату через З, а число работников – Ч.

Заработная плата в январе составит:

$$\bar{З}_0 = \frac{\Sigma З_0 Ч_0}{\Sigma Ч_0} = \frac{6000 \times 2400 + 5500 \times 2100 + 5100 \times 1500}{2400 + 2100 + 1500} = 5600 \text{ руб.}$$

Заработная плата в июле равна:

$$\bar{З}_1 = \frac{\Sigma З_1 Ч_1}{\Sigma Ч_1} = \frac{7000 \times 1600 + 6200 \times 2000 + 5900 \times 1400}{1600 + 2000 + 1400} = 6372 \text{ руб.}$$

Индекс заработной платы переменного состава определим, используя формулу (7.17):

$$J^{ПС} = \frac{\bar{З}_1}{\bar{З}_0} = \frac{\Sigma З_1 Ч_1}{\Sigma Ч_1} : \frac{\Sigma З_0 Ч_0}{\Sigma Ч_0} = \frac{6372}{5600} = 1{,}138.$$

Следовательно, средняя заработная плата работников по трем отраслям экономики в июле по сравнению с январем выросла на 13,8%.

Абсолютный прирост средней заработной платы составил:

$$\Delta З = 6372 - 5600 = 772 \text{ руб.}$$

Данное значение показывает, как изменилась средняя заработная плата по отраслям под влиянием двух факторов: изменения уровня заработной платы в каждой отрасли экономики и изменения структуры численности работников.

Определим индекс заработной платы постоянного состава, руководствуясь формулой (7.18):

$$J^{ФС} = \frac{\Sigma З_1 Ч_1}{\Sigma Ч_1} : \frac{\Sigma З_0 Ч_1}{\Sigma Ч_1} = 6372 : \frac{6000 \times 1600 + 5500 \times 2000 + 5100 \times 1400}{1600 + 2000 + 1400} =$$

$$= \frac{6372}{5548} = 1{,}149.$$

Следовательно, средняя заработная плата работников по данным отраслям экономики в июле по сравнению с январем выросла на 14,9% в результате изменения только одного фактора – самой заработной платы по каждой отрасли.

Абсолютный прирост средней заработной платы за счет изменения заработной платы по каждой отрасли составил:

$$\Delta З^{З} = 6372 – 5548 = 824 \text{ руб.}$$

Определим влияние изменения структуры численности работников на динамику средней заработной платы на основе индекса структурных сдвигов (7.19):

$$J^{СТР} = \frac{\Sigma З_0 Ч_1}{\Sigma Ч_1} : \frac{\Sigma З_0 Ч_0}{\Sigma Ч_0} = \frac{5548}{5600} = 0{,}9907.$$

Следовательно, увеличение доли работников с меньшей заработной платой в общей их численности привело к снижению средней заработной платы по трем отраслям вместе на 0,03 %.

Абсолютное снижение средней заработной платы за счет изменения структуры численности работников составит:

$$\Delta З^{Ч} = 5548 – 5600 = 52 \text{ руб.}$$

Отрицательный эффект структурных сдвигов объясняется тем, что в июле по сравнению с январем в большей мере сократилась доля работников с наиболее высоким уровнем заработной платы.

Проверим взаимосвязь рассчитанных индексов и абсолютных отклонений:

$$J^{ПС} = J^{ФС} \times J^{СТР} = 1{,}149 \times 0{,}9907 = 1{,}138$$

$$\Delta З = \Delta З^{З} + \Delta З^{Ч} = 824 - 52 = 772 \text{ руб.}$$

БИБЛИОГРАФИЧЕСКИЙ СПИСОК

1. Гусаров, В. М. Статистика : учебное пособие для вузов / В. М. Гусаров. – М. : ЮНИТИ – ДАНА, 2003. – 463 с.

2. Ефимова, М. Р. Практикум по общей теории статистики : учебное пособие / М. Р. Ефимова, О. И. Ганченко, Е. В. Петрова. – М. : Финансы и статистика, 2003. – 336 с.

3. Практикум по теории статистики : учебное пособие / под ред. проф. Р. А. Шмойловой. – М. : Финансы и статистика, 2003. – 416 с.

4. Статистика : учебник / под ред. И. И. Елисеевой. – М. : Проспект, 2010. – 448 с. – (Рек. М-вом образов. РФ).

5. Теория статистики : учебник / под ред. проф. Г. Л. Громыко. – М. : ИНФРА-М, 2002. – 560 с.

6. Харченко, Н. М. Статистика : учебник / Н. М. Харченко. – М. : Издательско-торговая корпорация «Дашков и К», 2006. – 368 с.

ПРИЛОЖЕНИЯ

Приложение 1

Значения функции Лапласа при разных значениях t Ф(t) = $\frac{1}{\sqrt{2\pi}} \int_{t}^{+t} e^{-\frac{t^2}{2}} dt$

Таблица 1

t	Сотые доли									
	0	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**
0,0	0000	0080	0160	0239	0319	0399	0478	0558	0638	0717
0,1	0797	0876	0955	1034	1114	1192	1271	1350	1428	1507
0,2	1585	1663	1741	1819	1897	1974	2051	2028	2205	2282
0,3	2358	2434	2510	2586	2661	2737	2812	2886	2961	3035
0,4	3108	3182	3255	3328	3401	3473	3545	3616	3688	3759
0,5	3829	3899	3969	4039	4108	4177	4245	4313	4381	4448
0,6	4515	4581	4647	4713	4778	4843	4907	4971	5035	5098
0,7	5161	5223	5285	5346	5407	5467	5527	5587	5646	5705
0,8	5763	5821	5878	5935	5991	6047	6102	6157	6211	6265
0,9	6319	6372	6424	6476	6528	6579	6629	6679	6727	6778
1,0	6827	6875	6923	7415	7017	7063	7109	7154	7199	7243
1,1	7287	7330	7373	7415	7457	7499	7540	7580	7620	7660
1,2	7699	7737	7775	7813	7850	7887	7923	7959	7994	8029
1,3	8064	8098	8132	8165	8198	8230	8262	8293	8324	8355
1,4	8385	8415	8444	8473	8501	8529	8557	8584	8611	8638
1,5	8664	8690	8715	8740	8764	8789	8812	8836	8859	8882
1,6	8904	8926	8948	8969	8990	9011	9031	9051	9070	9089
1,7	9109	9127	9146	9164	9182	9199	9216	9233	9249	9265
1,8	9281	9297	9312	9327	9342	9357	9371	9385	9399	9412
1,9	9425	9439	9451	9464	9476	9488	9500	9512	9523	9534
2,0	9545	9556	9566	9576	9586	9596	9606	9615	9627	9634
2,1	9643	9651	9660	9668	9676	9684	9692	9700	9707	9715
2,2	9722	9729	9736	9743	9749	9755	9762	9768	9774	9780
2,3	9785	9791	9797	9802	9807	9812	9817	9822	9827	9832
2,4	9836	9840	9845	9849	9853	9857	9861	9865	9869	9872
2,5	9876	9879	9883	9886	9889	9892	9895	9898	9901	9904
2,6	9907	9909	9912	9915	9917	9920	9924	9926	9927	9929
2,7	9931	9933	9935	9937	9939	9940	9942	9944	9946	9947
2,8	9949	9950	9952	9953	9955	9956	9958	9959	9960	9961
2,9	9963	9964	9965	9966	9967	9968	9969	9970	9971	9972
3,0	99730	99739	99747	99755	99763	99771	99779	99786	99793	99800
3,1	99807	99813	99819	99825	99831	99837	99842	99847	99853	99858
3,2	99863	99867	99872	99876	99880	99884	99889	99892	99896	99900
3,3	99903	99907	99910	99913	99916	99919	99922	99925	99928	99930
3,4	99933	99935	99937	99940	99942	99944	99946	99948	99950	99952
3,5	99953	99955	99957	99958	99960	99961	99963	99964	99966	99967
4,0	99994	-	-	-	-	-	-	-	-	-
5,0	99999	-	-	-	-	-	-	-	-	-

Приложение 2

S (t) в распределении Стьюдента

Таблица 1

t \ v	1	2	3	4	5	6-7	8-10	11-15	16-25	25-30	∞
1	*2*	*3*	*4*	*5*	*6*	*7*	*8*	*9*	*10*	*11*	*12*
0,0	500	500	500	500	500	500	500	500	500	500	500
0,1	532	535	537	537	538	538	539	539	539	539	539,827
0,2	563	570	573	574	576	575	578	578	578	578	579,259
0,3	593	606	608	610	612	613	615	616	616	616	617,911
0,4	621	636	642	645	647	649	651	652	653	654	655,421
0,5	648	667	674	678	681	683	685	687	689	689	691,462
0,6	672	695	705	710	713	715	718	721	722	724	725,746
0,7	694	723	733	739	742	746	749	752	754	756	758,036
0,8	715	746	759	766	770	774	778	781	783	785	788,144
0,9	733	768	783	790	795	800	804	808	811	813	815,939
1,0	750	789	804	813	818	823	828	832	835	838	841,344
1,1	765	807	824	834	839	844	850	854	858	860	864,333
1,2	779	824	842	852	858	864	870	874	878	881	884,930
1,3	791	838	858	868	875	881	887	892	896	899	903,199
1,4	803	852	872	883	890	896	902	907	912	915	919,243
1,5	813	864	885	896	903	909	916	921	925	928	933,192
1,6	822	875	896	908	915	921	928	933	937	940	945,200
1,7	831	884	906	918	925	932	938	943	948	951	955,434
1,8	839	893	915	927	934	941	947	952	956	959	964,069
1,9	846	901	923	935	942	948	955	960	964	967	971,283
2,0	852	908	930	942	949	955	962	967	970	973	977,249
2,1	858	915	937	948	955	961	967	972	976	978	982,135
2,2	864	921	942	954	960	966	972	977	980	982	986,096
2,3	870	926	948	958	965	971	977	981	984	986	989,275
2,4	874	931	952	963	969	975	980	984	987	989	991,802
2,5	879	935	956	966	973	978	983	987	989	991	993,79
2,6	883	939	960	970	976	981	986	989	991	993	995,338
2,7	887	943	963	973	979	983	988	991	993	995	996,533
2,8	891	946	966	976	981	985	990	993	995	996	997,444
2,9	894	949	969	978	983	987	991	994	996	997	998,134
3,0	898	952	971	980	985	989	993	995	997	997	998,650
3,1	901	955	973	982	987	990	994	996	997	998	999,032
3,2	904	957	975	984	988	991	995	997	998	998	999,312
3,3	906	960	977	985	989	992	995	997	998	999	999,516
3,4	909	962	979	986	990	993	996	998	998	999	999,663
3,5	911	964	980	988	991	993	997	998	999	-	999,767
3,6	914	965	982	989	992	994	997	998	-	-	999,840

Окончание приложения 2

Окончание таблицы 1

1	*2*	*3*	*4*	*5*	*6*	*7*	*8*	*9*	*10*	*11*	*12*
3,7	916	967	983	990	993	995	998	999	-	-	999,892
3,8	918	969	984	990	994	996	998	999	-	-	999,927
3,9	920	970	985	991	994	996	998	999	-	-	999,951
4,0	922	971	986	992	995	997	998	-	-	-	999,968
4,1	924	973	987	993	995	997	999	-	-	-	999,979
4,2	926	974	988	993	996	998	999	-	-	-	999,986
4,3	927	975	988	994	996	998	999	-	-	-	999,991
4,4	929	976	989	994	996	998	-	-	-	-	999,994
4,5	930	977	989	995	997	998	-	-	-	-	999,996
4,6	932	978	990	995	997	998	-	-	-	-	999,997
4,7	933	979	991	995	997	999	-	-	-	-	999,998
4,8	935	980	991	996	998	999	-	-	-	-	999,999
4,9	936	980	992	996	998	999	-	-	-	-	999,999
5,0	937	981	992	996	998	999	-	-	-	-	999,999
5,1	938	982	993	996	998	-	-	-	-	-	999,999
5,2	940	982	993	997	998	-	-	-	-	-	999,999
5,3	941	983	993	997	998	-	-	-	-	-	999,999
5,4	942	984	994	997	998	-	-	-	-	-	-
5,5	943	984	994	997	999	-	-	-	-	-	-
5,6	943	984	994	997	999	-	-	-	-	-	-
5,7	945	985	995	998	999	-	-	-	-	-	-
5,8	946	986	995	998	999	-	-	-	-	-	-
5,9	947	986	995	998	999	-	-	-	-	-	-
6,0	947	987	995	998	-	-	-		-	-	-

Приложение 3

Значение t-критерия Стьюдента при уровне значимости 0,1, 0,05 и 0,01

Таблица 1

Число степеней свободы	α		
	0,1	**0,05**	**0,01**
1	6,3138	12,706	63,657
2	2,9200	4,3027	9,9248
3	2,3534	3,1825	5,8409
4	2,1318	2,7764	4,6041
5	2,0150	2,5706	4,0321
6	1,9432	2,4469	3,7074
7	1,8946	2,3646	3,4995
8	1,8595	2,3060	3,3554
9	1,8331	2,2622	3,2498
10	1,8125	2,2281	3,1693
11	1,7959	2,2010	3,1058
12	1,7823	2,1788	3,0545
13	1,7709	2,1604	3,0123
14	1,7613	2,1448	2,9768
15	1,7530	2,1315	2,9467
16	1,7459	2,1199	2,9208
17	1,7396	2,1098	2,8982
18	1,7341	2,1009	2,8784
19	1,7291	2,0930	2,8609
20	1,7247	2,0860	2,8453
21	1,7207	2,0796	2,8314
22	1,7171	2,0739	2,8188
23	1,7139	2,0687	2,8073
24	1,7109	2,0639	2,7969
25	1,7081	2,0595	2,7874
26	1,7056	2,0555	2,7787
27	1,7033	2,0518	2,7707
28	1,7011	2,0484	2,7633
29	1,6991	2,0452	2,7564
30	1,6973	2,0423	2,7500
40	1,6839	2,0211	2,7045
60	1,6707	2,0003	2,6603
120	1,6577	1,9799	2,6174
∞	1,6449	1,9600	2,5758

Приложение 4

Значения процентных пределов $t_{\alpha,k}$ в зависимости от k степеней свободы и заданного уровня значимости α для распределения Стьюдента

Таблица 1

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

Учебное издание

Виктория Владимировна Свечникова

СТАТИСТИКА

Учебное пособие

Ведущий редактор
Е. В. Кондаева

Старший корректор
Е. А. Феонова

Ведущий инженер
Г. А. Чумак

Подписано в печать 27.09.2012 г.
Формат 60x84 1/16. Усл. печ. 5,7.
Тираж 50 экз. Заказ ________

Издательство Орского гуманитарно-технологического института (филиала) Федерального государственного бюджетного образовательного учреждения высшего профессионального образования «Оренбургский государственный университет»

462403, г. Орск Оренбургской обл., пр. Мира, 15 А