

МИНИСТЕРСТВО НАУКИ И ВЫСШЕГО ОБРАЗОВАНИЯ
РОССИЙСКОЙ ФЕДЕРАЦИИ

Федеральное государственное бюджетное образовательное учреждение
высшего образования

«Воронежский государственный технический университет»

Строительно-политехнический колледж

ОСНОВЫ СТАТИСТИЧЕСКОГО АНАЛИЗА

МЕТОДИЧЕСКИЕ УКАЗАНИЯ
к выполнению лабораторных работ
для студентов специальности 27.02.07 «Управление качеством продукции,
процессов и услуг (по отраслям)» всех форм обучения

Воронеж 2021

УДК 519.23(07)
ББК 60.6я723

Составитель канд. техн. наук С. И. Сергеева

Основы статистического анализа: методические указания к выполнению лабораторных работ для студентов специальности 27.02.07 «Управление качеством продукции, процессов и услуг (по отраслям)» / ФГБОУ ВО «Воронежский государственный технический университет»; сост.: С. И. Сергеева. - Воронеж: Изд-во ВГТУ, 2021. - 43 с.

Содержат комплекс заданий и методических рекомендаций к выполнению лабораторных работ. Выполнение предусмотренных заданий позволит студентам закрепить теоретические знания и приобрести необходимые практические навыки.

Предназначены для студентов специальности 27.02.07 «Управление качеством продукции, процессов и услуг (по отраслям)».

Методические указания подготовлены в электронном виде и содержатся в файле ЛР _ ОА. pdf.

Ил. 5. Табл. 12. Библиогр.: 5 назв.

УДК 519.23(07)
ББК 60.6я723

Рецензент - И. В. Поцбнева, канд. техн. наук, доц. кафедры систем управления и информационных технологий в строительстве ВГТУ

*Издается по решению редакционно-издательского совета
Воронежского государственного технического университета*

ВВЕДЕНИЕ

Статистика как общественная наука изучает количественную сторону различных массовых экономических и социальных явлений и процессов общественной жизни с учётом их качественной стороны. С помощью системы количественных показателей статистика отражает процесс и результаты функционирования рыночной экономики, анализирует тенденции и закономерности развития общества.

Данный курс лабораторных работ включает и общую теорию статистики, разрабатывающую методологию получения, обработки и анализа статистических данных, и экономическую статистику, уделяющую основное внимание количественному описанию экономического процесса, состоянию и развитию экономики в целом, и социальную статистику, дающую количественную характеристику уровня жизни населения, состояния и развития социальной сферы.

Важную роль в освоении курса статистики играют практические занятия, на которых рассматриваются актуальные вопросы, связанные с переходом экономики на международные стандарты учёта и статистики.

Основная цель практических занятий – применение методов статистического анализа в конкретных областях исследования социально-экономических процессов, происходящих в обществе.

Применяя современную методологию исчисления статистических показателей, учитывающую изменившуюся хозяйственную и законодательную практику, необходимо изучать взаимосвязи между важнейшими статистическими показателями и на примерах решения задач показывать возможные направления их анализа.

ЛАБОРАТОРНАЯ РАБОТА № 1 СВОДКА И ГРУППИРОВКА СТАТИСТИЧЕСКОЙ ИНФОРМАЦИИ

Цель работы. Изучить основные принципы проведения сводки и группировки статистических данных и освоить их практическое применение с использованием стандартных функций MicrosoftExcel.

Краткие теоретические сведения

Сводка — это комплекс последовательных операций по обобщению конкретных единичных факторов, образующих совокупность, для выявления типичных черт и закономерностей, присущих изучаемому явлению в целом.

Группировкой называется разбиение совокупности на группы, однородные по какому-либо признаку.

Группировка является важнейшим статистическим методом обобщения статистических данных, основой для правильного исчисления статистических показателей.

С помощью метода группировок решаются следующие задачи:

- выделение социально-экономических типов явлений;
- изучение структуры явления и структурных сдвигов, происходящих в нем;
- выявление связи и зависимости между явлениями.

Построение группировки начинается с определения состава группировочных признаков.

Группировочным признаком называется признак, по которому проводится разбиение единиц совокупности на отдельные группы. От правильного выбора группировочного признака зависят выводы статистического исследования.

После того как определено основание группировки, следует решить вопрос о количестве групп, на которые надо разбить исследуемую совокупность.

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

Если группировка строится по атрибутивному признаку, то групп, как правило, будет столько, сколько имеется градаций, видов состояний у этого признака (например, группировка работников по образованию).

Если группировка проводится по количественному признаку, то необходимо обратить особое внимание на число единиц исследуемого объекта и степень колеблемости группировочного признака.

Определение числа групп можно осуществить математическим путем, используя формулу Стерджесса:

$$n = 1 + 3,322 \cdot \lg N \quad (1.1)$$

где n - число групп; N - число единиц совокупности.

Недостаток формулы Стерджесса состоит в том, что ее применение дает хорошие результаты, если совокупность состоит из большого числа единиц, распределение единиц по признаку, положенному в основание группировки, близко к нормальному и при этом в группах применяются равные интервалы.

Чтобы получить группы, адекватные действительности, необходимо руководствоваться сущностью изучаемого явления.

Когда определено число групп, следует определить интервалы группировки.

Интервал — это значения варьирующего признака, лежащие в определенных границах. Каждый интервал имеет свою величину, верхнюю и нижнюю границы или хотя бы одну из них. Нижней границей интервала называется наименьшее значение признака в интервале, а верхней границей является наибольшее значение признака в нем. Величина интервала представляет собой разность между верхней и нижней границами интервала.

Интервалы группировки в зависимости от их величины бывают равные и неравные.

Если вариация признака проявляется в сравнительно узких границах и распределение носит равномерный характер, то строят группировку с равными интервалами.

Величина равного интервала вычисляется по формуле

$$i = (X_{\max} - X_{\min}) / n, \quad (1.2)$$

где X_{\max} и X_{\min} - максимальное и минимальное значения признака в совокупности.

Прежде чем определять максимальное и минимальное значения, из совокупности рекомендуется исключить аномальные наблюдения.

Если максимальное и минимальное значения сильно отличаются от смежных с ними значений вариантов в упорядоченном ряду значений группировочного признака, то для определения величины интервала следует использовать не максимальное и минимальное значения, а значения, несколько превышающие минимум и несколько меньше, чем максимум.

Полученное значение величины интервала необходимо округлить, как правило, в большую сторону.

Если размах вариации признака совокупности велик и значения признака варьируются неравномерно, то следует использовать группировку с неравными интервалами. Использование неравных интервалов характерно для большинства социально-экономических явлений, особенно при анализе макроэкономических показателей. |

Неравные интервалы могут быть прогрессивно возрастающими или прогрессивно убывающими в арифметической или геометрической прогрессии, а также произвольными и специализированными.

Специализированные интервалы используются для выделения из совокупности одних и тех же типов по одному и тому же признаку для явлений, находящихся в различных условиях.

Интервалы группировок могут быть закрытыми и открытыми. Закрытыми называются интервалы, у которых имеются верхняя и нижняя границы. У открытых интервалов указана только одна граница: верхняя - у первого, нижняя - у последнего.

При группировке единиц совокупности по количественному признаку границы интервалов могут быть обозначены по-разному, в зависимости от того, непрерывный это признак или прерывный.

Если основанием группировки служит непрерывный признак, то одно и то же значение признака выступает и верхней, и нижней границами двух смежных интервалов. Таким образом, нижняя граница $(i + 1)$ -го интервала равна верхней границе i -го интервала.

Если в основании группировки лежит прерывный признак, то нижняя граница $(i + 1)$ -го интервала равна верхней границе i -го интервала, увеличенной на 1.

После определения границ групп строится ряд распределения.

Статистический ряд распределения - это упорядоченное распределение единиц совокупности на группы по определенному варьирующему признаку.

В зависимости от признака, взятого за основу группировки, различают атрибутивные и вариационные ряды распределения.

Атрибутивные ряды распределения построены по качественным признакам (распределение по полу, национальности, профессии и т.д.).

Вариационные ряды распределения построены по количественному признаку.

Ряд распределения принято оформлять в виде таблиц.

Любой вариационный ряд состоит из двух элементов: вариантов и частот.

Вариантами считаются отдельные значения признака, которые он принимает в вариационном ряду, т.е. конкретное значение варьирующего значения признака.

Частоты - это численности отдельных вариантов или каждой группы вариационного ряда, т.е. это числа, показывающие как часто встречаются те или иные варианты в ряду распределения. Сумма всех частот определяет численность всей совокупности, ее объем.

Порядок выполнения работы

1. Выбор группировочного признака.
2. Определение количества групп.
3. Расчет величины интервала группировки.
4. Установление границ групп.
5. Построение ряда распределения.
6. Выбор показателей для характеристики групп.
7. Расчет величины показателей по каждой группе.
8. Сведение данных группировки в таблицу.

Пример выполнения работы

Задание. Произвести группировку предприятий региона по стоимости основных средств и определить средний объем выпуска продукции в каждой группе. Исходные данные для расчета приведены в табл. 1.1.

Таблица 1.1

Характеристика предприятий региона по стоимости основных средств и объему выпуска продукции

Номер предприятия	Стоимость основных средств, млн руб.	Объем выпуска продукции, млн руб.	Номер предприятия	Стоимость основных средств, млн руб.	Объем выпуска продукции, млн руб.
1	5999	5349	16	6413	7079
2	6925	6882	17	9387	6339
3	6925	7046	18	3949	1544
4	10097	7248	19	10826	11431
5	8097	5256	20	6695	4105
6	11116	14090	21	6633	13366
7	5880	3525	22	6472	6340
8	7355	5431	23	6183	3624
9	9566	7680	24	8107	4917
10	7884	8226	25	9369	9040
11	7038	4081	26	11817	5359
12	4950	10473	27	4894	6266
13	4550	6097	28	8488	17093
14	3427	5307	29	7560	11641
15	6062	7400	30	6429	12328

В соответствии с заданием группировочным признаком является стоимость основных средств предприятия.

Количество групп $n = 1 + 3,322 * \lg(30) = 6$. Следовательно, совокупность необходимо разбить на 6 групп.

Теперь определяется минимальное и максимальное значения признака в совокупности, для этого используются стандартные функции *MSExcel* МАКС, МИН (категория «Статистические»). Минимальное значение составляет 3427 млн руб., максимальное -118/7 млн руб.

Рассчитывается величина равного интервала группировки:

$$i = (11817 - 3427) / 6 = 1400 \text{ млн руб.}$$

Таким образом, величина интервала составляет 1400 млн руб. Определяются границы групп (табл. 1.2).

Таблица 1.2

Расчет границ групп и частот

Номер группы	Нижняя граница	Верхняя граница	Количество предприятий
1	3427	3427+1400=4827	3
2	4827	4827+1400=6227	6
3	6227	6227+1400=7627	10
4	7627	7627+1400=9027	4
5	9027	9027+1400=10427	4
6	10427	10427+1400=11827	3

Рассчитывается количество предприятий (частота), вошедших в каждую группу с использованием стандартной функции *MicrosoftExcel* ЧАСТОТА (категория «Статистические»).

Функция ЧАСТОТА должна задаваться в качестве формулы массива. Для этого необходимо сделать следующее:

1. Выделить диапазон ячеек, равный количеству интервалов, начиная с ячейки, содержащей формулу.

2. Нажать клавишу F2.

3. Нажать одновременно клавиши CTRL + SHIFT + ENTER

Если формула не будет введена как формула массива, единственное значение будет равно 1.

Результаты расчетов сводятся в табл. 1.3. В таблице представлен вариационный ряд распределения промышленных предприятий региона по стоимости основных средств.

Таблица 1.3

Группировка предприятий региона по стоимости основных средств

Группы предприятий по стоимости основных средств	Количество предприятий	Средняя стоимость основных средств, млн руб.	Средний объем производства, млн руб.
3427-4827	3	4127	4316
4827-6227	6	5527	6106
6227-7627	10	6927	7830
7627-9027	4	8327	8873
9027-10427	4	9727	7577
10427-11827	3	11127	10293
Итого	30	7347	7485

Вариантами ряда распределения являются отдельные значения стоимости основных средств промышленных предприятий региона, частотами - количество предприятий, вошедших в каждую группу.

По каждой группе предприятий определяется средняя стоимость основных средств предприятия и средний объем выпуска продукции.

Выводы

В ходе проведения сводки и группировки статистической информации разбили совокупность, состоящую из 30 промышленных предприятий, на 6 групп.

В качестве группировочного признака взяли показатель стоимости основных средств предприятий. Данный признак является количественным. Получилась простая группировка.

По каждой группе предприятий рассчитали показатели - среднюю стоимость основных средств предприятия и средний объем выпуска продукции.

ЛАБОРАТОРНАЯ РАБОТА № 2 СРЕДНИЕ ВЕЛИЧИНЫ. ПОКАЗАТЕЛИ ВАРИАЦИИ

Цель работы. Изучить основные принципы расчета средних величин и показателей вариации и по несгруппированным и сгруппированным данным и освоить их практическое применение с использованием стандартных функций *Microsoft Excel*.

Краткие теоретические сведения

Изучая зарегистрированные в процессе статистического наблюдения величины того или иного признака у отдельных единиц совокупности, можно обнаружить различия между ними. Поэтому, чтобы определить значение признака, характерное для всей изучаемой совокупности единиц, используют средние величины.

Средняя величина представляет собой обобщенную количественную характеристику признака в статистической совокупности в конкретных условиях места и времени. Показатель в форме средней величины выражает типичные черты и дает обобщенную характеристику однотипных явлений по одному из варьирующих признаков.

Простая средняя рассчитывается по не сгруппированным данным и имеет следующий общий вид:

$$\bar{X} = \frac{\sum x_i^z}{n} \quad (2.1)$$

где x_i - величина осредняемого признака у каждой единицы совокупности; n - объем совокупности; z - показатель степени средней.

Взвешенная средняя вычисляется по сгруппированным данным и имеет общий вид:

$$\bar{X} = \sqrt{\frac{\sum x_i^2 \cdot f_i}{\sum f_i}}, \quad (2.2)$$

где f_i - частота или повторяемость индивидуальных значений признака в совокупности.

Выбор вида средней определяется экономическим содержанием показателя и исходных данных. В каждом конкретном случае применяется одна из степенных средних.

Наиболее распространенным видом средних величин является средняя арифметическая. Она применяется в тех случаях, когда объем варьирующего признака для всей совокупности выражается суммой значений признаков отдельных ее единиц.

Средняя арифметическая простая рассчитывается по формуле:

$$\bar{X} = \frac{\sum x_i}{n}, \quad (2.3)$$

Средняя арифметическая взвешенная по формуле:

$$\bar{X} = \frac{\sum x_i \cdot f_i}{\sum f_i} \quad (2.4)$$

Кроме степенных средних в статистической практике применяются структурные (или распределительные) средние. Они используются для изучения внутреннего строения и структуры рядов распределения признака. К ним относятся мода и медиана.

Модой называется наиболее часто встречающийся вариант или то значение признака, которое соответствует максимальной точке теоретической кривой распределения. В дискретном ряду мода - это вариант с наибольшей частотой.

В интервальном вариационном ряду модой приближенно считают центральный вариант так называемого модального интервала.

Модальный интервал - это интервал, который имеет наибольшую частоту.

Конкретное значение моды для интервального ряда распределения с равными интервалами определяется формулой

$$M_o = X_{mo} + i_{mo} \frac{f_{mo} - f_{mo-1}}{(f_{mo} - f_{mo-1}) + (f_{mo} - f_{mo+1})} \quad (2.5)$$

где x_{mo} - нижняя граница модального интервала; i_{mo} - величина модального интервала; f_{mo} - частота, соответствующая модальному интервалу;

f_{mo-1} - частота интервала, предшествующая модальному интервалу; f_{mo+1} - частота интервала, следующего за модальным.

Медиана - это величина, которая делит численность упорядоченного вариационного ряда на две равные части: одна часть имеет значение варьирующего признака меньше, чем средний вариант, другая - больше, чем средний вариант.

В дискретном ряду с четным числом индивидуальных величин медианой будет средняя арифметическая из двух смежных вариантов, а с нечетным числом

медианой будет варианта, расположенная в центре ряда.

Для интервального ряда медиана определяется по формуле:

$$M_{Me} = X_{Me} + i_{Me} \frac{\sum f / 2 - S_{Me-1}}{f_{Me}}, \quad (2.6)$$

где X_{Me} - нижняя граница медианного интервала; i_{Me} - величина медианного интервала; S_{Me-1} - сумма частот, накопленная до медианного интервала; f_{Me} - частота, соответствующая медианному интервалу.

Медианный интервал - это первый интервал, в котором накопленная частота составляет половину или больше половины общей суммы частот.

Различие индивидуальных значения признака внутри изучаемой совокупности в статистике называется *вариацией признака*. Она возникает в результате того, что его индивидуальные значения признака складываются под совокупным влиянием разнообразных факторов, которые по-разному сочетаются в каждом отдельном случае.

Колеблемость отдельных значений характеризуют показатели вариации. Статистические показатели, определяющие вариацию, делятся на две группы: абсолютные и относительные. К абсолютным относятся размах вариации, среднее линейное отклонение, дисперсия и среднее квадратическое отклонение. Вторая группа показателей вычисляется как отношение абсолютных показателей вариации к средней арифметической. Относительными показателями вариации являются коэффициенты осцилляции, вариации, относительное линейное отклонение.

Самым простым абсолютным показателем является размах вариации. *Размах вариации* показывает, насколько велико различие между единицами совокупности, имеющими наименьшее и наибольшее значение признака:

$$R = X_{\max} - X_{\min}, \quad (2.7)$$

где X_{\max} и X_{\min} - соответственно, наибольшее и наименьшее значения признака в совокупности.

Среднее линейное отклонение вычисляется как средняя арифметическая (простая или взвешенная в зависимости от исходных данных) из абсолютных значений отклонений вариант и среднего значения признака в совокупности по следующим формулам:

$$\bar{d} = \frac{\sum |x_i - \bar{x}|}{n}, \quad (2.8)$$

$$\bar{d} = \frac{\sum |x_i - \bar{x}| f_i}{\sum f_i}, \quad (2.9)$$

Среднее линейное отклонение дает обобщенную характеристику степени колеблемости признака в совокупности.

Общепринятыми мерами вариации являются дисперсия и среднее квадратическое отклонение. Данные показатели нашли широкое применение в статистических исследованиях, а также в других отраслях знаний.

Дисперсия представляет собой средний квадрат отклонений индивидуальных значений признака от их средней величины и в зависимости от исходных данных вычисляется по формулам простой и взвешенно

$$\sigma^2 = \frac{\sum(x_i - \bar{x})^2}{n}, \quad (2.10)$$

$$\sigma^2 = \frac{\sum(x_i' - \bar{x})^2 \cdot f_i}{\sum f_i} \quad (2.11)$$

Среднее квадратическое отклонение равно корню квадратному из дисперсии:

$$\sigma = \sqrt{\sigma^2} \quad (2.12)$$

Среднее квадратическое отклонение - это обобщающая характеристика размеров вариации признака в совокупности. Оно выражается в тех же единица измерения, что и признак (в метрах, тоннах, рублях, процентах и т.д.)

Относительные показатели вариации выражаются в процентах и определяют не только сравнительную оценку вариации, но и дают характеристику однородности совокупности.

Коэффициент осцилляции вычисляется как отношение размаха вариации к средней арифметической:

$$K_o = \frac{R}{\bar{x}} \cdot 100\% \quad (2.13)$$

Относительное линейное отклонение вычисляется как отношение среднего линейного отношения к средней арифметической:

$$K_{\bar{d}} = \frac{\bar{d}}{\bar{x}} \cdot 100\% \quad (2.14)$$

Наиболее часто в практических расчетах применяется показатель относительной вариации - коэффициент вариации.

Коэффициент вариации вычисляется как отношение среднего квадратического отклонения к средней арифметической:

$$K_{\sigma} = \frac{\sigma}{\bar{x}} \cdot 100\% \quad (2.15)$$

Для распределений, близких к нормальному, совокупность считается однородной, если коэффициент вариации не превышает 33 %

Средняя величина только тогда отражает типичный уровень признака, когда она рассчитана по качественно однородной совокупности.

При расчете средних величин и показателей вариации по не сгруппированным данным используют стандартные функции *Microsoft Excel* (табл. 2.1).

Таблица 2.1

Стандартные функции Microsoft Excel' используемые при расчете показателей по не сгруппированным данным

Показатель	Используемая функция
Средняя арифметическая простая	СРЗНАЧ
Мода дискретного ряда	МОДА
Медиана дискретного ряда	МЕДИАНА
Минимальное значение	МИН
Максимальное значение	МАКС
Среднее линейное отклонение	СРОТКЛ

Порядок выполнения работы

1. Определение средних величин и показателей вариации по не сгруппированным данным.
2. Определение средних величин и показателей вариации по сгруппированным данным.
3. Формулировка выводов об однородности исследуемой совокупности по изучаемому признаку.
4. Формулировка выводов о типичности среднего значения признака.

Пример выполнения работы

Задание. Рассчитать среднее значение, структурные средние, показатели вариации, используя данные о стоимости основных средств предприятий региона. Определить однородность совокупности предприятий региона по стоимости основных средств. Оценить типичность рассчитанного среднего значения.

При расчете показателей по сгруппированным данным необходимо использовать формулы для вычисления взвешенных величин. Для удобства расчета используется табл. 2.2.

Таблица 2.2

Вспомогательная таблица для расчета средних показателей и показателей вариации

Группы предприятий по стоимости основных средств x_i , млн руб.	Количество предприятий f_i	Середина интервала x'_i , млн руб.	$x'_i \cdot f_i$	$ x'_i - x_{cp} \cdot f_i$	$(x'_i - \bar{x})^2 \cdot f_i$	Накопленные частоты S_i
3427-4827	3	4127	12381		31105200	3
4827-6227	6	5527	33162	10920	19874400	9
6227-7627	10	6927	69270	4200	1764000	19
7627-9027	4	8327	33308	3920	3841600	23
9027-10427	4	9727	38908	9520	22657600	27
10427-11827	3	11127	33381	11340	42865200	30
Итого	30	-	220410	49560	122108000	-

Среднее значение признака $\bar{x} = 220410/30 = 7347$ млн руб.

Модальным интервалом является интервал (6227-7627), имеющий максимальную частоту (10ед.)

Мода $M_o = 6227 + (7627 - 6227) \cdot \frac{10 - 6}{(10 - 6) + (10 - 4)} \approx 6787$ млн руб.

Медианным интервалом является интервал 6227-7627, в котором накопленная частота $S = 19$, что больше половины общей суммы частот.

Медиана $M_e = 6227 + (7627 - 6227) \cdot \frac{30/2 - 9}{10} = 7067$ млн. руб.

Размах вариации $R = 11827 - 3427 = 8400$ млн. руб.

Среднее линейное отклонение $\bar{d} = 49560/30 = 1652$ млн. руб.

Дисперсия $\sigma^2 = 122108000/30 = 4070267$

Среднеквадратическое отклонение $\sigma = \sqrt{4070267} = 2017$ млн. руб.

Коэффициент осцилляции $K_o = 8400 \cdot 100/7347 = 114\%$.

Относительное линейное отклонение $K_{\bar{d}} = 1652 \cdot 100/7347 = 22\%$.

Коэффициент вариации $v = 1652 \cdot 100/7347 = 27\%$.

Результаты расчета по не сгруппированным данным с использованием стандартных функций *Microsoft Excel* составляют: среднее значение - 7302 млн. руб., мода - 6925 млн. руб., медиана - 6914 млн. руб., размах вариации - 8390 млн. руб., среднее линейное отклонение - 1649 млн. руб., дисперсия - 4382880, среднее квадратическое отклонение - 2094 млн. руб., коэффициент осцилляции - 115%, линейный коэффициент вариации - 23 %, коэффициент вариации - 29 %.

Выводы

Приведенные расчеты показали, что коэффициент вариации не превышает 33 %, следовательно, совокупность можно считать однородной. Типичность среднего значения в данной совокупности - удовлетворительная.

ЛАБОРАТОРНАЯ РАБОТА № 3 АНАЛИЗ ВАРИАЦИОННЫХ РЯДОВ

Цель работы. Изучить основные принципы определения общего характера распределения и освоить их практическое применение.

Краткие теоретические сведения

Выяснение общего характера распределения предполагает оценку степени его однородности, а также вычисление показателей асимметрии и эксцесса.

Статистический ряд распределения — это упорядоченное распределение единиц совокупности на группы по определенному варьирующему признаку.

В зависимости от признака, взятого за основу группировки, различают атрибутивные и вариационные ряды распределения.

Атрибутивные ряды распределения построены по качественным признакам (распределение по полу, национальности, профессии и т.д.).

Вариационные ряды распределения построены по количественному признаку.

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

Дискретный вариационный ряд характеризует распределение единиц совокупности по дискретному признаку.

Интервальный вариационный ряд используется в случае непрерывной вариации признака, а также если дискретная вариация проявляется в широких пределах (т.е. число вариантов дискретного признака достаточно велико).

В вариационных рядах существует определенная связь между изменением частот и изменением величины варьирующего признака: с увеличением варьирующего признака величина частот вначале возрастает до определенной величины, а затем уменьшается. Такого рода изменения называются *закономерностями распределения*.

Одна из важных целей статистического изучения вариационных рядов состоит в том, чтобы выявить закономерность распределения и определить ее характер.

Под *кривой распределения* понимается графическое изображение в виде непрерывной линии изменения частот в вариационном ряду, функционально связанного с изменением вариант.

В практике статистических исследований приходится встречаться с самыми различными распределениями. Различают следующие разновидности кривых распределения: одновершинные кривые (симметричные, умеренно асимметричные и крайне асимметричные) и многовершинные кривые. Для однородных совокупностей, как правило, характерны одновершинные распределения.

Определение формы кривой является важной задачей, так как статистический материал в обычных для него условиях дает по определенному признаку характерную, типичную для него кривую распределения.

Выяснение общего характера распределения предполагает оценку степени его однородности, а также вычисления показателей асимметрии и эксцесса.

При сравнительном изучении асимметрии нескольких распределений с разными единицами измерения вычисляется относительный показатель асимметрии.

Наиболее точным и распространенным является показатель асимметрии, основанный на определении центрального момента третьего порядка:

$$A_s = \frac{\mu_3}{\sigma_3} \quad (3.1)$$

где μ_3 – центральный момент третьего порядка.

Центральный момент третьего порядка рассчитывается по формуле:

$$\mu_3 = \frac{\sum(x-\bar{x})^3 \cdot f}{\sum f} \quad (3.2)$$

При симметричном распределении коэффициент асимметрии равен нулю. При $A_s < 0$ асимметрия левосторонняя. При $A_s > 0$ асимметрия правосторонняя.

Оценка степени существования показателя асимметрии дается с помощью среднеквадратической ошибки:

$$\sigma_{A_s} = \sqrt{\frac{6(n-1)}{(n+1)(n+3)}} \quad (3.3)$$

где n – объем наблюдений.

Если $\frac{|E_x|}{\sigma_{E_x}} < 3$, то асимметрия существенная и распределение $\frac{|A_s|}{\sigma_{A_s}} > 3$

признака в совокупности не симметрично. Если $\frac{A_s}{\sigma_{A_s}} < 3$, то асимметрия несущественна и ее наличие может быть объяснено влиянием различных случайных факторов.

Кроме того, в симметричных распределениях рассчитывается показатель эксцесса (островершинной), использующий центральный момент четвертого порядка:

$$E_x = \frac{\mu_4}{\sigma^4} - 3, \quad (3.4)$$

где μ_4 – центральный момент четвертого порядка.

Центральный момент четвертого порядка вычисляется по формуле:

$$\mu_4 = \frac{\sum(x-\bar{x})^4 \cdot f}{\sum f} \quad (3.5)$$

Эксцесс представляет собой выпад вершины эмпирического распределения вверх или вниз от вершины кривой нормального распределения, где $\mu_4 / \sigma^4 = 3$

Если $E_x=0$, то распределение является нормальным. Если $E_x<0$ (отрицательный эксцесс), то распределение является плосковершинным (низковершинным).

Низковершинность означает отрицательный эксцесс и характеризует большую разбросанность членов ряда.

Если $E_x >0$ (положительный эксцесс), то распределение является островершинным (высоковершинным).

Высоковершинность означает положительный эксцесс и характеризует скопление частот (т.е. членов вариационного ряда) в середине.

Оценка степени существенности этого показателя дается с помощью среднеквадратической ошибки эксцесса:

$$\sigma_{E_x} = \sqrt{\frac{24n(n-2)(n-3)}{(n-1)^2(n+3)(n+5)}} \quad (3.6)$$

Если $\frac{|E_x|}{\sigma_{E_x}} < 3$, то эксцесс существенный.

Оценка существенности показателей асимметрии и эксцесса позволяет сделать вывод о том, можно ли отнести данное эмпирическое распределение к типу кривых нормального распределения.

Для того чтобы эмпирическое распределение можно было отнести к нормальному типу, необходимо, чтобы и асимметрия и эксцесс были признаны несущественными.

Порядок выполнения работы

1. Построение графического изображения эмпирического распределения.
2. Анализ вариационного ряда распределения.
3. Формулирование вывода о возможности отнесения данного эмпирического распределения к типу кривых нормального распределения.

Пример выполнения работы

Задание. На основании анализа вариационного ряда распределения сделайте вывод о возможности отнесения данного эмпирического распределения к типу кривых нормального распределения.

Для удобства вычисления показателей асимметрии и эксцесса используется табл. 3.1.

Таблица 3.1

Вспомогательная таблица для расчета показателей асимметрии и эксцесса

Группы предприятий по стоимости основных средств x_i , млн. руб.	Количество предприятий f_i	Середина интервала x_i , млн. руб.	$(x_i' - x_{cp})^3 * f_i$	$(x_i' - x_{cp})^4 * f_i$
3427-4827	3	4127	-100158744000	1000121643731520
4827-6227	6	5527	-36171408000	408783873291846
6227-7627	10	6927	-740880000	47077239334410
7627-9027	4	8327	3764768000	113592964
9027-10427	4	9727	53925088000	12403483772164
10427-11827	3	11127	162030456000	165905890703523
Итого	30	-	82649280000	1634292244426430

Среднее значение признака составляет 7347 млн. руб., мода- 6787 млн. руб., медиана – 7067 млн. руб., среднеквадратическое отклонение – 2017 млн. руб.

Центральный момент третьего порядка

$$\mu_3 = 82649280000 / 30 = 2754976000.$$

Показатель асимметрии: $A_s = 2754976000 / 2017^3 = 0,3$.

Положительный показатель асимметрии ($A_s > 0$) свидетельствует о наличии правосторонней асимметрии.

Среднеквадратическая ошибка асимметрии составляет^

$$\sigma_{A_s} = \sqrt{\frac{6(30-1)}{(30+1)(30+3)}} = 1,6$$

Отношение $\frac{A_s}{\sigma_{A_s}} < 3$, следовательно, асимметрия незначительна и ее наличие может быть объяснено влиянием различных случайных обстоятельств.

В симметричных распределениях рассчитывается показатель эксцесса (островершинности).

Центральный момент четвертого порядка $\mu_4 = 1634292244426430 / 30 = 54476408147548$.

Показатель эксцесса (островершинной) $E_x = 54476408147548 / 2017^4 - 3 = 0,3$.

Положительный эксцесс ($E_x > 0$) говорит об островершинности распределения. Островершинность означает положительный эксцесс и характеризует скопление членов вариационного ряда в середине.

Среднеквадратическая ошибка эксцесса составляет^

$$\sigma_{E_x} = \sqrt{\frac{24 \cdot 30(30-2)(30-3)}{(30-1)^2(30+3)(30+5)}} = 0,56$$

Отношение $\frac{|E_x|}{\sigma_{E_x}} < 3$, следовательно, эксцесс незначительный.

Так как и асимметрия и эксцесс несущественны, то данное эмпирическое распределение можно отнести к типу кривых нормального распределения.

Выводы

Вариационный ряд имеет правостороннюю асимметрию, которая признана несущественной и ее наличие может быть объяснено влиянием различных случайных обстоятельств.

Положительный эксцесс указывает на качественную однородность исследуемой совокупности.

Оценка существенности показателей асимметрии и эксцесса позволяет сделать вывод о том, что данное эмпирическое распределение можно отнести к типу кривых нормального распределения.

Положительный эксцесс указывает на качественную однородность исследуемой совокупности.

Оценка существенности показателей асимметрии и эксцесса позволяет сделать вывод о том, что данное эмпирическое распределение можно отнести к типу кривых нормального распределения.

ЛАБОРАТОРНАЯ РАБОТА № 4 ВИДЫ ДИСПЕРСИЙ

Цель работы. Изучить основные принципы расчета различных видов дисперсий и освоить их практическое применение для анализа взаимосвязей.

Краткие теоретические сведения

Если статистическая совокупность разбита на группы по какому-либо признаку, то для оценки влияния различных факторов, определяющих колеблемость индивидуальных значений признака, можно воспользоваться разложением дисперсии на составляющие: на межгрупповую и внутригрупповую дисперсии.

Общая дисперсия измеряет вариацию признака по всей совокупности под влиянием всех факторов, обусловивших эту вариацию.

Она определяется по формулам как простая или взвешенная дисперсия:

$$\sigma_{\text{общ}}^2 = \frac{\sum(x_i - \bar{x}_{\text{общ}})^2}{n}, \quad (4.1)$$

$$\sigma_{\text{общ}}^2 = \frac{\sum(x_i - \bar{x}_{\text{общ}})f}{\sum f}, \quad (4.2)$$

где $\bar{x}_{\text{общ}}$ - общая средняя для всей изучаемой совокупности.

Межгрупповая дисперсия характеризует систематическую вариацию, т.е. вариацию изучаемого признака под воздействием факторного признака, положенного в основу группировки.

Она характеризует колеблемость групповых средних около общей средней:

$$\delta^2 = \frac{\sum(\bar{x}_{\text{гр}} - \bar{x}_{\text{общ}})^2}{n}, \quad (4.3)$$

$$\delta^2 = \frac{\sum(\bar{x}_{гр} - \bar{x}_{общ})^2 f}{\sum f}, \quad (4.4)$$

где f_i – численность отдельных групп; $\bar{x}_{гр}$ – средняя в группах.

Внутригрупповая (частная) дисперсия характеризует случайную вариацию в каждой отдельной группе, т.е. часть вариации, возникающую под влиянием других, неучтенных факторов, и не зависящую от фактора, положенного в основу группировки.

Внутригрупповые дисперсии рассчитываются по каждой группе:

$$\sigma_{гр}^2 = \frac{\sum(x_i - \bar{x}_{гр})^2}{n}, \quad (4.5)$$

$$\sigma_{гр}^2 = \frac{\sum(x_i - \bar{x}_{гр})^2 f}{\sum f}. \quad (4.6)$$

Средняя изнутригрупповых дисперсий определяется на основании внутригрупповых дисперсий по каждой группе:

$$\overline{\sigma_{гр}^2} = \frac{\sum \sigma_{гр}^2 \cdot f}{\sum f}. \quad (4.7)$$

Общая дисперсия равна сумме величин межгрупповой дисперсии и средней из внутригрупповой дисперсии:

$$\sigma_{общ}^2 = \delta^2 + \overline{\sigma_{гр}^2}. \quad (4.8)$$

Это правило имеет большую практическую значимость, так как позволяет выявить зависимость результатов от определяющих факторов.

Очевидно, что чем больше доля межгрупповой дисперсии в общей дисперсии, тем сильнее влияние группировочного признака на изучаемый признак.

Поэтому в статистическом анализе широко используется такой показатель, как эмпирический коэффициент детерминации:

$$\eta^2 = \frac{\delta^2}{\sigma_{общ}^2}. \quad (4.9)$$

Выраженный в процентах, он показывает, какая доля всей вариации результативного признака обусловлена факторным признаком, положенным в основу группировки.

Эмпирический коэффициент детерминации изменяется в пределах: $0 \leq \eta^2 \leq 1$.

Эмпирическое корреляционное отношение показывает тесноту связи между группировочным и результативным признаками:

$$\eta = \sqrt{\eta^2}. \quad (4.10)$$

Эмпирическое корреляционное отношение также изменяется в пределах $0 \leq \eta^2 \leq 1$.

Для качественной оценки тесноты связи на основе показателя эмпирического корреляционного отношения можно воспользоваться соотношениями Чэддока (табл. 4.1).

Таблица 4.1

		Соотношения Чэддока					
$ \eta $	0	0,1-0,3	0,3-0,5	0,5-0,7	0,7-0,9	0,9-0,99	1
Сила связи	Отсутствует	Слабая	Умеренная	Заметная	Тесная	Весьма тесная	Функциональная

Порядок выполнения работы

1. Определение средних значений выработки по каждой группе и в целом по совокупности.
2. Определение групповых дисперсий по каждой группе и средней из групповых дисперсий.
3. Определение межгрупповой дисперсии.
4. Определение общей дисперсии.
5. Проверка правильности расчетов с использованием закона сложения дисперсий.
6. Определение эмпирического коэффициента детерминации.
7. Определение эмпирического корреляционного отношения.
8. Формулировка вывода о наличии связи между группировочным (факторным) и результативным признаком.

Пример выполнения работы

Задание. Используя данные табл. 8, 9, сделайте вывод о наличии связи между прохождением рабочими повышения квалификации и выработкой.

Таблица 8.1

Выработка рабочих, не прошедших повышение квалификации

Табельный номер рабочего	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
Выработка рабочих, шт. /см	50	42	46	46	50	47	45	47	46	48	49	40	50	50	49	45	50	47	45	49
Табельный номер рабочего	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36	37	38	39	40
Выработка рабочих, шт. /см	44	41	41	41	43	44	46	40	50	45	42	45	46	45	43	50	45	40	45	42
Табельный номер рабочего	41	42	43	44	45	46	47	48	49	50	51	52	53	54	55	56	57	58	59	60

Табельный номер рабочего	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
Выработка рабочих, шт./см	52	45	45	48	55	50	54	60	61	54	57	61	57	49	50	45	60	47	46	48
Табельный номер рабочего	61	62	63	64	65	66	67	68	69	70	71	72	73	74	75	76	77	78	79	80
Выработка рабочих, шт./см	61	60	53	48	55	51	58	49	56	49	55	50	49	54	46	52	52	60	51	49
Табельный номер рабочего	81	82	83	84	85	86	87	88	89	90	91	92	93	94	95	96	97	98	99	100
Выработка рабочих, шт./см	59	47	45	56	53	59	55	49	50	58	60	56	49	50	58	59	49	53	60	61

Средняя выработка рабочих, не прошедших повышение квалификации, составляет 45,48 шт./см.

Средняя выработка рабочих, прошедших повышение квалификации, составляет 53,13 шт./см.

Средняя выработка рабочих цеха составляет 50,07 шт./см.

Определим дисперсии с помощью стандартных функций *Microsoft Excel*.
Общая дисперсия выработки всех рабочих равна 33,03.

Дисперсия выработки рабочих, не прошедших повышение квалификации, составляет 10,00.

Дисперсия выработки рабочих, прошедших повышение квалификации, составляет 24,92.

Средняя из внутригрупповых дисперсий[^]

$$\overline{\sigma_1^2} = \frac{10,00 \cdot 40 + 24,92 \cdot 60}{100} = 18,95 .$$

Межгрупповая дисперсия

$$\delta^2 = \frac{(45,48 - 50,07)^2 \cdot 40 + (53,13 - 50,07)^2 \cdot 60}{100} = 14,08 .$$

Используя закон сложения дисперсий, проверим правильность расчетов:

$$\sigma^2 = 14,08 + 18,95 = 33,03.$$

Эмпирический коэффициент детерминации:

$$\eta^2 = 14,08 / 33,03 = 0,43.$$

Рассчитаем эмпирическое корреляционное отношение:

$$\eta = \sqrt{0,43} = 0,65.$$

Выводы

Средняя выработка рабочих, не прошедших повышение квалификации, составляет 45,48 шт./см, прошедших повышение квалификации – 53,13 шт./см. Средняя выработка рабочих цеха равна 50,07 шт./см.

Доля вариации выработки рабочего, которая обусловлена прохождением рабочими повышения квалификации, равна 43 %.

Эмпирическое корреляционное отношение равно 0,65. Следовательно, связь между выработкой и прохождением рабочими повышения квалификации заметная.

ЛАБОРАТОРНАЯ РАБОТА № 5 ВЫБОРОЧНОЕ НАБЛЮДЕНИЕ

Цель работы. Изучить основные принципы проведения выборочного исследования и освоить практическое применение результатов выборочного наблюдения для оценки величины параметров генеральной совокупности.

Краткие теоретические сведения

Под выборочным наблюдением понимается такое не сплошное наблюдение, при котором статистическому обследованию (наблюдению) подвергаются единицы изучаемой совокупности, отобранные случайным способом.

Выборочное наблюдение ставит перед собой задачу по обследуемой части дать характеристику всей совокупности единиц при условии соблюдения всех правил и принципов проведения статистического наблюдения и научно-организованной работы по отбору единиц.

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

Совокупность отобранных для обследования единиц в статистике принято называть *выборочной* и все ее обобщающие показатели – *выборочными*.

При *собственно-случайной выборке* отбор из генеральной совокупности осуществляется случайным образом, наугад, без каких-либо элементов системности.

Собственно-случайный отбор может быть повторный и бесповторный.

Разность между величиной параметра в генеральной совокупности и его величиной, вычисленной по результатам выборочного наблюдения, называется *ошибкой выборочного наблюдения*.

Величина ошибки выборки зависит от способа отбора единиц совокупности в выборку, от объема выборки и от вариации изучаемого признака в генеральной совокупности.

Для определения возможных границ генеральных характеристик рассчитываются средняя и предельная ошибки выборки.

Средняя ошибка выборочной средней при повторном отборе вычисляется по формуле:

$$\mu_{\bar{x}} = \sqrt{\frac{\sigma_{\bar{x}}^2}{n}}, \quad (5.1)$$

где $\sigma_{\bar{x}}^2$ – дисперсия изучаемого показателя в генеральной совокупности (при большом числе наблюдений выборочная дисперсия равна генеральной);
 n – численность выборки,

Средняя ошибка выборочной средней при бесповторном отборе равна:

$$\mu_{\bar{x}} = \sqrt{\frac{\sigma_{\bar{x}}^2}{n} \left(1 - \frac{n}{N}\right)}, \quad (5.2)$$

где N – численность генеральной совокупности.

Средняя ошибка выборочной доли (альтернативного признака) при повторном отборе и бесповторном отборе рассчитывается по формулам:

$$\mu_w = \sqrt{\frac{w(1-w)}{n}}, \quad (5.3)$$

$$\mu_w = \sqrt{\frac{w(1-w)}{n} \left(1 - \frac{n}{N}\right)}. \quad (5.4)$$

где w – выборочная доля единиц, обладающих изучаемым признаком;
 $w(1-w)$ – дисперсия доли (альтернативного признака).

Выборочная доля единиц, обладающих изучаемым признаком, определяется по формуле:

$$w = m / n, \quad (5.5)$$

где m – число единиц, обладающих изучаемым (альтернативным) признаком.

Предельная ошибка выборки дает возможность выяснить, в каких пределах находится величина генеральной средней.

Предельная ошибка выборки рассчитывается по формуле:

$$\Delta = t\mu, \quad (5.6)$$

где t — коэффициент доверия, показатель, определяющий размер ошибки в зависимости от того, с какой вероятностью $\Phi(t)$ она находится.

Наиболее часто употребляемые уровни доверительной вероятности и соответствующие значения t для выборок достаточно большого объема ($n \geq 30$) приведены в табл. 5.1.

Таблица 5.1

Наиболее часто употребляемые уровни доверительной вероятности и соответствующие значения t

t	1,00	1,50	1,96	2,00	2,50	2,58	3,00	3,50
$\Phi(t)$	0,683	0,866	0,95	0,954	0,988	0,990	0,997	0,999

Предельная ошибка выборки позволяет определить предельные значения характеристик генеральной совокупности и их доверительные интервалы:

$$\tilde{x} - \Delta_{\tilde{x}} \leq \bar{x} \leq \tilde{x} + \Delta_{\tilde{x}}, \quad (5.7)$$

$$w - \Delta_w \leq p \leq w + \Delta_w. \quad (5.8)$$

Это означает, что с заданной вероятностью можно утверждать, что значение генеральной средней (доли) следует ожидать в указанных пределах.

При подготовке выборочного наблюдения с заранее заданным значением допустимой ошибки выборки очень важно правильно определить численность выборочной совокупности, которая с определенной вероятностью обеспечит заданную точность результатов наблюдения.

Объем выборки при определении среднего размера признака при повторном и бесповторном отборе вычисляется следующим образом:

$$n = \frac{t^2 \cdot \sigma_{\tilde{x}}^2}{\Delta_{\tilde{x}}^2}, \quad (5.9)$$

$$n = \frac{t^2 \cdot \sigma_{\tilde{x}}^2 \cdot N}{\Delta_{\tilde{x}}^2 \cdot N + t^2 \cdot \sigma_{\tilde{x}}^2}. \quad (5.10)$$

Объем выборки при определении доли альтернативного признака при повторном и бесповторном отборе вычисляется следующим образом:

$$n = \frac{t^2 \cdot w(1-w)}{\Delta_w^2}, \quad (5.11)$$

$$n = \frac{t^2 \cdot w(1-w) \cdot N}{\Delta_w^2 \cdot N + t^2 \cdot w(1-w)}. \quad (5.12)$$

Механический отбор единиц из генеральной совокупности осуществляется в установленной пропорции через равные интервалы.

При расчете средней ошибки механической выборки используется формула средней ошибки при собственно-случайном бесповторном отборе.

Для определения средней ошибки механической выборки используется формула средней ошибки при собственно-случайном бесповторном отборе.

Порядок выполнения работы

1. Проведение отбора единиц в выборку с помощью процедуры *Microsoft Excel* СЕРВИС / АНАЛИЗ ДАННЫХ / ВЫБОРКА.

2. Определение основных выборочных характеристик (выборочной средней и выборочной доли изучаемого признака). Расчет количества единиц, обладающих альтернативным признаком, производится с помощью процедуры *Microsoft Excel* СЧЕТЕСЛИ.

3. Определение средней ошибки выборки.

4. Определение предельной ошибки выборки с заданной вероятностью.

5. Определение пределов, в которых с заданной вероятностью будут находиться генеральные характеристики.

6. Определение необходимой численности выборки, при которой средняя ошибка выборки снизится в X раз.

Пример выполнения работы

Задание. В табл. 5.2 приведены примеры случайного повторного 10 %-го отбора изделий из партии готовой продукции в выборку. Оцените средний вес изделия и долю стандартных изделий во всей партии с вероятностью 0,866. Определите оптимальный объем выборки при условии, что предельная ошибка выборки сократится в 2 раза.

Таблица 5.2

Вес изделия, г, по результатам случайного повторного 10 %-го отбора изделий в выборку

48,0	47,0	48,0	51,0	50,0	51,0	47,0	51,0	47,0	47,0
50,0	51,0	51,0	50,0	48,0	49,0	48,0	49,0	51,0	48,0
51,0	48,0	47,0	49,0	48,0	51,0	50,0	47,0	47,0	49,0
48,0	48,0	51,0	51,0	51,0	51,0	51,0	50,0	47,0	48,0

Средний вес изделий в выборке составил 49,13 г.

Средняя ошибка выборки $\mu_{\bar{x}} = \sqrt{2,46/40} = 0,25$ г.

Предельная ошибка выборки с вероятностью 0,866 составляет

$$\Delta_{\bar{x}} = 1,5 \cdot 0,25 = 0,375 \text{ г.}$$

Таким образом, с вероятностью 0,866 можно утверждать, что вес изделия во всей партии находится в границах

$$49,13 - 0,375 \leq \bar{x} \leq 49,13 + 0,375 ,$$

$$48,755 \leq \bar{x} \leq 49,505 .$$

Для того чтобы сократить предельную ошибку выборки в 2 раза, необходимо, чтобы численность выборки составила

$$n = \frac{1,5^2 \cdot 2,46}{(0,375/2)^2} = 158 \text{ изделий.}$$

В выборке лишь 5 изделий имеют стандартный вес, равный 50 г.

Доля изделий, соответствующих требованию стандарта, составляет

$$w = 5/40 = 0,125.$$

Дисперсия доли изделий, соответствующих требованию стандарта, составляет $w(1-w) = 0,125(1-0,125) = 0,109$.

$$\text{Средняя ошибка доли } \mu_w = \sqrt{\frac{0,109}{40}} = 0,052.$$

Предельная ошибка доли равна $\Delta_w = 1,5 \cdot 0,052 = 0,078$.

Таким образом, с вероятностью 0,866 можно утверждать, что доля стандартных изделий во всей партии находится в границах

$$0,125 - 0,078 \leq p \leq 0,125 + 0,078,$$

$$0,047 \leq p \leq 0,203.$$

Для того чтобы сократить предельную ошибку выборки в 2 раза, необходимо, чтобы численность выборки составила

$$n = \frac{1,5^2 \cdot 0,109}{(0,078/2)^2} = 162 \text{ изделия.}$$

Выводы

В результате 10 %-го выборочного обследования партии изделий (используя случайную схему повторного отбора) в выборку отобрано 40 изделий. Средний вес изделия в выборке составил 49,13 г.

С вероятностью 0,866 можно утверждать, что вес одного изделия во всей партии составляет от 48,755 до 49,505

Для того чтобы сократить среднюю ошибку выборки в 2 раза, необходимо, чтобы численность выборки составила 158 изделий.

В выборке лишь 5 изделий имеют стандартный вес, равный 50 г.

Доля изделий, соответствующих требованию стандарта, составляет 12,5 %.

С вероятностью 0,866 можно утверждать, что доля стандартных изделий во всей партии составляет 4,7–20,3 %.

Для того чтобы сократить среднюю ошибку в выборки в 2 раза, необходимо, чтобы численность выборки составила 162 изделия.

ЛАБОРАТОРНАЯ РАБОТА № 6 АНАЛИЗ РЯДОВ ДИНАМИКИ

Цель работы. Изучить основные принципы расчета показателей динамики и освоить их практическое применение для анализа рядов динамики и выявления тенденций развития явлений и процессов.

Краткие теоретические сведения

Процесс развития, движения социально-экономических явлений во времени в статике принято называть динамикой.

Рядами динамики называются последовательно расположенные в хронологическом порядке статистические данные, отражающие развитие изучаемого явления во времени.

В каждом ряду динамики имеются два основных элемента: *показатель времени* и соответствующие им *уровни* развития изучаемого явления у.

Существуют различные виды рядов динамики.

В зависимости от способа выражения уровней ряды динамики подразделяются на ряды абсолютных, относительных и средних величин.

В зависимости от того, как выражают уровни ряда состояние явления на определенные моменты времени или его величину за определенные интервалы времени, различают, соответственно, моментные и интервальные ряды динамики.

Моментные ряды динамики отражают состояние изучаемых явлений на определенный момент времени (дату). *Интервальные ряды динамики* отображают итоги развития изучаемых явлений за отдельные периоды (интервалы).

В зависимости от расстояния между уровнями ряды динамики подразделяются на ряды с равноотстоящими и неравноотстоящими уровнями во времени.

В основе расчета показателей рядов динамики лежит сравнение его уровней.

Абсолютный прирост является важнейшим из показателей динамики. Он характеризует увеличение (уменьшение) уровня ряда за определенный промежуток времени.

Абсолютный прирост базисный выражается формулой:

$$\Delta y_{\sigma i} = y_i - y_1, \quad (6.1)$$

где y_i – уровень сравниваемого периода; y_1 – уровень, базисного периода.

Абсолютный прирост цепной выражается формулой:

$$\Delta y_{\pi i} = y_i - y_{i-1}, \quad (6.2)$$

где y_{i-1} - уровень предшествующего периода.

Между базисными и цепными абсолютными приростами имеется связь: сумма цепных абсолютных приростов равна базисному абсолютному приросту последнего периода ряда динамики:

$$\sum \Delta y_{\pi i} = \Delta y_{\sigma n} = y_n - y_1. \quad (6.3)$$

Темп роста характеризует интенсивность изменения уровня ряда. Он может выражаться в виде коэффициента или процента:

$$Kp_{\sigma i} = y_i : y_1, \quad (6.3)$$

$$Kp_{\pi i} = y_i : y_{i-1}, \quad (6.4)$$

$$Tr = Kp \cdot 100. \quad (6.5)$$

Между базисными и цепными коэффициентами роста имеется взаимосвязь: произведение последовательных цепных коэффициентов роста равно базисному коэффициенту роста последнего периода ряда динамики:

$$Kp_{\pi 2} \cdot Kp_{\pi 3} \cdot Kp_{\pi 4} = Kp_{\sigma 4} = y_4 / y_1. \quad (6.6)$$

Частное от деления последующего базисного коэффициента роста на предыдущий равно соответствующему цепному коэффициенту роста:

$$Kp_{\sigma 2} / Kp_{\sigma 1} = Kp_{\pi 2}. \quad (6.7)$$

Темп прироста показывает, на сколько процентов сравниваемый уровень больше (меньше) уровня, принятого за базу сравнения:

$$Tp_{\sigma i} = \Delta y_{\sigma i} \cdot 100 / y_1, \quad (6.8)$$

$$Tp_{\pi i} = \Delta y_{\pi i} \cdot 100 / y_{i-1}, \quad (6.9)$$

Между темпами роста и прироста имеется взаимосвязь:

$$Тп_i (\%) = Тр_i (\%) - 100, \quad (6.10)$$

$$Тп_i = Кр_i - 1. \quad (6.11)$$

Абсолютное значение одного процента прироста определяется как частное абсолютного прироста к темпу прироста:

$$A_i = \Delta y_{цi} / Тп_{цi} = (y_i - y_{i-1}) / Тп_{цi} = y_{i-1} / 100 = 0,01 y_{i-1}. \quad (6.12)$$

Темп наращивания говорит о затухании или усилении темпов роста:

$$ТН_i = \frac{\Delta y_{цi}}{y_1} = \frac{y_i - y_{i-1}}{y_1} = \frac{y_i}{y_1} \cdot \frac{y_{i-1}}{y_1} = Кр_{\bar{y}_i} - Кр_{\bar{y}_{i-1}}. \quad (6.13)$$

Для получения обобщающих показателей динамики рассчитываются средние показатели.

Средний уровень ряда динамики характеризует типическую величину абсолютных уровней.

В зависимости от вида ряда динамики средний уровень определяется по-разному.

В интервальных рядах динамики с равностоящими уровнями его рассчитывают по формуле:

$$\bar{y} = \frac{\sum y_i}{n} = \frac{y_1 + y_2 + y_3 \dots + y_n}{n}, \quad (6.14)$$

где n – число уровней ряда;

в интервальных рядах динамики с неравностоящими уровнями – по формуле:

$$\bar{y} = \frac{\sum y_i t_i}{\sum t_i} = \frac{y_1 t_1 + y_2 t_2 + y_3 t_3 \dots + y_n t_n}{t_1 + t_2 + t_3 \dots + t_n}, \quad (6.15)$$

где y_i – уровни ряда динамики, сохранившиеся без изменения в течение промежутка времени t_i ; t_i – интервал времени между смежными уровнями.

В моментных рядах динамики с равноотстоящими уровнями его вычисляют по формуле

$$\bar{y} = \frac{0,5 y_1 + y_2 t_2 + y_3 t_3 \dots + 0,5 y_n}{n-1}, \quad (6.16)$$

где n – число уровней ряда; y_i – уровни ряда динамики;

в моментных рядах динамики с неравностоящими уровнями – по формуле:

$$\bar{y} = \frac{\sum (y_i - y_{i+1}) \cdot t_i}{2 \cdot \sum_{i=1}^{n-1} t_i} = \frac{(y_1 + y_2) \cdot t_1 + (y_2 + y_3) \cdot t_2 + \dots + (y_{n-1} + y_n) \cdot t_{n-1}}{2 \cdot (t_1 + t_2 + t_3 \dots + t_{n-1})}. \quad (6.17)$$

Средний абсолютный прирост является обобщающим показателем скорости изменения явления во времени. Этот показатель дает возможность установить, насколько в среднем за единицу времени должен увеличиться уровень ряда (в абсолютном выражении), чтобы, отправляясь от начального уровня за данное число периодов (например, лет), достигнуть конечного уровня:

$$\overline{\Delta y} = \frac{\sum \Delta y_{ii}}{n-1} = \frac{y_n - y_1}{n-1}. \quad (6.18)$$

Сводной обобщающей характеристикой интенсивности изменения уровней ряда динамики служит средний темп роста, показывающий, во сколько раз в среднем за единицу времени изменился уровень динамического ряда. Необходимость исчисления среднего темпа роста возникает вследствие того, что темпы роста из года в год колеблются. Кроме того, средний темп роста часто следует определить в тех случаях, когда имеются данные об уровне в начале какого-либо периода и в конце его, а промежуточные данные отсутствуют.

Средний коэффициент роста вычисляется по формуле

$$\overline{K_p} = \sqrt[n-1]{K_{p_{ц2}} \cdot K_{p_{ц3}} \cdot \dots \cdot K_{p_{цn}}} = \sqrt[n-1]{\prod K_{p_{ц}}} = \sqrt[n-1]{K_{p_{6n}}} = \sqrt[n-1]{\frac{y_n}{y_1}}. \quad (6.19)$$

Средний темп прироста можно определить на основе взаимосвязи между коэффициентами (темпами) роста и прироста:

$$\overline{K_n} = \overline{K_p} - 1, \quad (6.20)$$

$$\overline{T_n} = \overline{T_p} - 100. \quad (6.21)$$

Одна из задач анализа рядов динамики – установить закономерность изменения уровней изучаемого показателя во времени, т.е. определить основную тенденцию развития явления (тренд).

Основная тенденция развития (тренд) это достаточно плавное и устойчивое изменение уровня явления во времени, более или менее свободное от случайных колебаний.

Для определения тренда используют специальные методы выравнивания (сглаживания) рядов динамики. При выравнивании отклонения, обусловленные случайными причинами, взаимопогашаются (сглаживаются), в результате четко обнаруживается действие основных факторов изменения уровней (общая тенденция).

Основную тенденцию можно представить либо графически, либо аналитически – в виде уравнения (модели) тренда.

Методы выравнивания рядов динамики таковы:

1. Метод укрупнения интервалов;
2. Метод усреднения по левой и правой половине;
3. Метод простой скользящей средней;
4. Аналитическое выравнивание.

Наиболее эффективным способом выявления тренда является аналитическое выравнивание, при этом уровни ряда динамики выражаются в виде функции: $\bar{y}_t = f(t)$.

Аналитическое выравнивание проводится на основе адекватной математической формулы – полиномов разной степени, экспонент и других функций. Выбор функции производится на основе анализа характера закономерности динамики данного явления.

Для оценки надежности линии тренда применяется величина R^2 :

$$R^2 = 1 - \frac{\sum (y_i - \hat{y}_i)^2}{(\sum y_i^2) - \frac{(\sum y_i)^2}{n}}, \quad (6.22)$$

где y – исходный уровень ряда динамики.

Наиболее надежной является та функция, для которой значение R^2 равно или близко к 1.

Порядок выполнения работы

1. Определение показателей ряда динамики.
2. Определение средних показателей ряда динамики.
3. Проведение аналитического выравнивания ряда динамики с использованием процедуры *MicrosoftExcel* МАСТЕР ДИАГРАММ / ДОБАВЛЕНИЕ ЛИНИИ ТРЕНДА.
4. Проверка надежности моделей тренда и выбор наиболее адекватной модели основной тенденции развития исследуемого явления.
5. Прогнозирование развития изучаемого явления в будущем.
6. Формулировка вывода о наличии основной тенденции развития исследуемого явления.

Пример выполнения работы

Задание. Используя данные табл. 6.1, рассчитайте показатели динамики. Определите наличие основной тенденции развития ряда динамики. Сделайте прогноз на 3 периода вперед.

Исследуемый ряд динамики является интервальным рядом с равноотстоящими уровнями.

Рассчитанные показатели динамики сведены в табл. 6.1.

Таблица 6.1

Расчет показателей динамики

Период	Выпуск продукции, тыс. т	Абсолютный прирост, тыс. т		Темп роста, %		Темп прироста, %		Абсолютное значение 1 прироста, тыс. т	Темп наращивания, %
		ба-зис-ный	цеп-ной	ба-зис-ный	цеп-ной	ба-зис-ный	цеп-ной		
1	21	—	—	—	—	—	—	—	—
2	22	1	1	105	105	5	5	0,21	0,048
3	24	3	2	114	109	14	9	0,22	0,095
4	25	4	1	119	104	19	4	0,24	0,048
5	27	6	2	129	108	29	8	0,25	0,095
6	28	7	1	133	104	33	4	0,27	0,048
7	31	10	3	148	111	48	11	0,28	0,143
8	33	12	2	157	106	57	6	0,31	0,095
9	33	12	0	157	100	57	0	0,33	0,0000,143
10	36	15	3	171	109	71	9	0,33	0,095
11	38	17	2	181	106	81	6	0,36	0,048
12	39	18	1	186	103	86	3	0,38	0,095
13	41	20	2	195	105	95	5	0,39	0,095
14	43	22	2	205	105	105	5	0,41	0,095
15	45	24	2	214	105	114	5	0,43	

Средний выпуск продукции, рассчитанный с помощью стандартной функции *Microsoft Excel*, составляет 32,4 тыс. т., средний абсолютный прирост равен:

$$\bar{\Delta y} = \frac{45-21}{15-1} = 1,7 \text{ тыс. т.}$$

Средний темп роста выпуска продукции $\bar{K}_p = \sqrt[15-1]{45/21} = 1,05$, средний темп прироста $\bar{T}_n = 105 - 100 = 5\%$.

Используя процедуру *Microsoft Excel*/МАСТЕР ДИАГРАММ/ДОБАВЛЕНИЕ ЛИНИИ ТРЕНДА, построим линии тренда нескольких моделей (рис. 6.1–6.5).

Наиболее надежной моделью основной тенденции развития исследуемого ряда динамики является полиномиальный (2-ой степени) тренд: $y = 0,0136x^2 + 1,5185x + 19,13$, так как значение параметра $R^2 = 0,9961$ является максимально приближенным к 1.

Прогнозное значение объема выпуска продукции в 16-м периоде составит $0,0136 \cdot 16^2 + 1,5185 \cdot 16 + 19,13 = 46,9$ тыс. т.

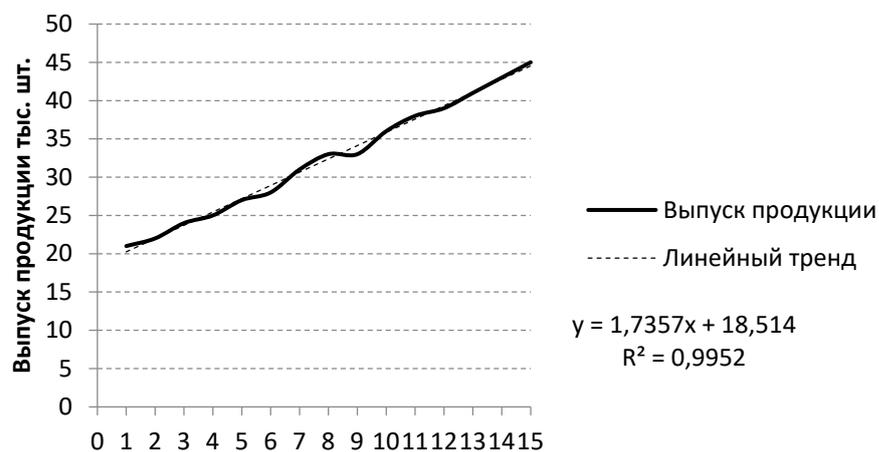


Рис. 6.1. Динамика выпуска продукции и линейный тренд, полученный методом аналитического выравнивания

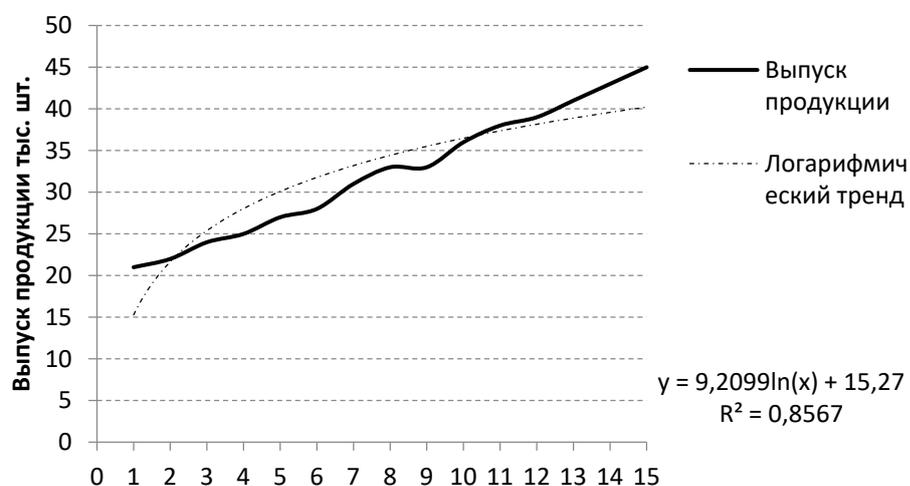


Рис. 6.2. Динамика выпуска продукции и логарифмический тренд, полученный методом аналитического выравнивания

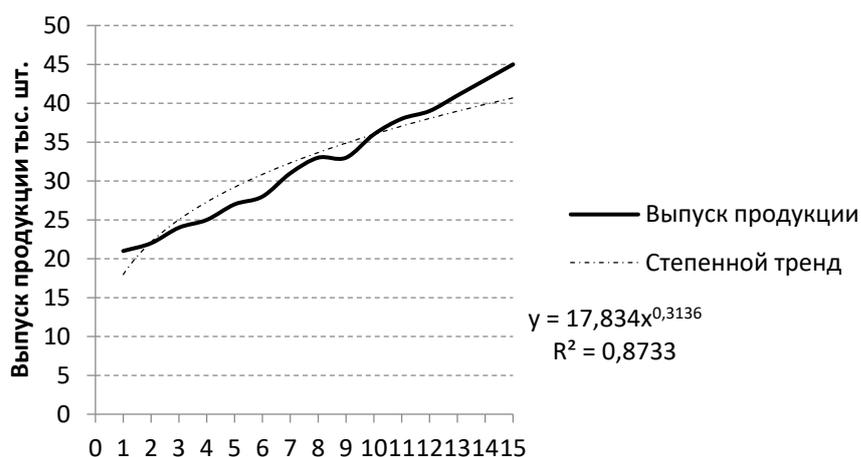


Рис. 6.3. Динамика выпуска продукции и степенной тренд, полученный методом аналитического выравнивания

Прогнозное значение объема выпуска продукции в 17-м периоде составит $0,0136 \cdot 17^2 + 1,5185 \cdot 17 + 19,13 = 48,9$ тыс. т.

Прогнозное значение объема выпуска продукции в 18-м периоде составит $0,0136 \cdot 18^2 + 1,5185 \cdot 18 + 19,13 = 50,9$ тыс. т.

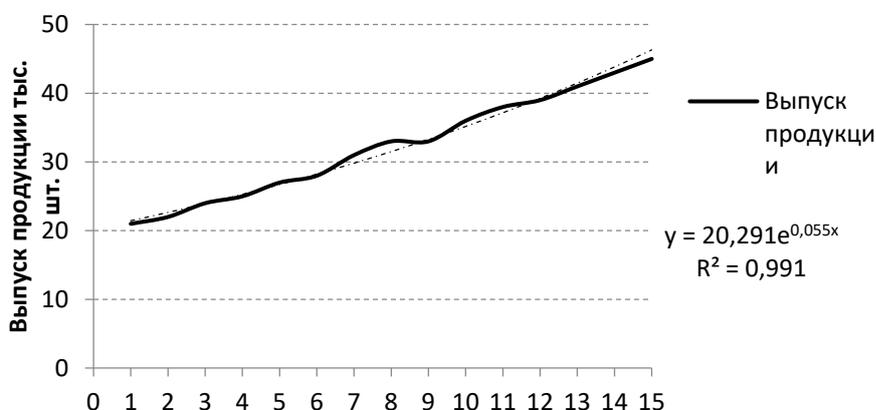


Рис. 6.4. Динамика выпуска продукции и экспоненциальный тренд, полученный методом аналитического выравнивания

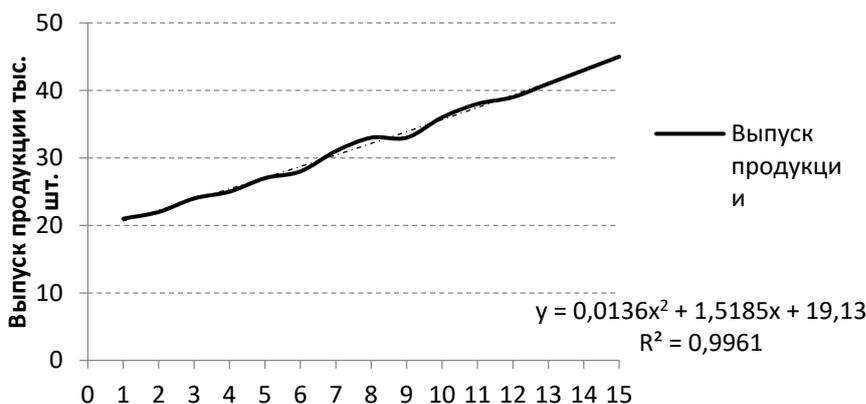


Рис. 6.5. Динамика выпуска продукции и полиномиальный тренд, полученный методом аналитического выравнивания

Выводы

Средний выпуск продукции в исследуемом периоде составил 32,4 тыс. т., средний абсолютный прирост — 1,7 тыс. т., средний темп роста — 105 %. Наблюдается постоянный рост объемов производства продукции. В процессе анализа ряда динамики выявлена основная тенденция развития выпуска по полиному 2-й степени ($y = 0,0136x^2 + 1,5185x + 19,13$).

Прогнозное значение объема выпуска продукции в 16-м периоде составит 46,9 тыс. т, в 17-м 48,9 тыс. т, в 18-м периоде — 50,9 тыс. т.

ЛАБОРАТОРНАЯ РАБОТА № 7

ИНДЕКСЫ КОЛИЧЕСТВЕННЫХ И КАЧЕСТВЕННЫХ ПОКАЗАТЕЛЕЙ

Цель работы. Изучить теоретические основы индексов и освоить их практическое применение с использованием стандартных функций *Microsoft Excel*.

Краткие теоретические сведения

Индексами называют сравнительные относительные величины, которые характеризуют изменение сложных социально-экономических показателей (показатели, состоящие из несуммируемых элементов) во времени, в пространстве, по сравнению с планом.

Индексные показатели позволяют осуществить анализ результатов деятельности предприятий и организаций, выпускающих самую разнообразную продукцию или занимающихся различными видами деятельности. С помощью индексов можно проследить роль отдельных факторов при формировании важнейших экономических показателей, выявить основные резервы производства. Индексы широко используются в сопоставлении международных экономических показателей при определении уровня жизни, деловой активности, ценовой политики и т.д.

Существует два подхода в интерпретации возможностей индексных показателей: **обобщающий (синтетический) и аналитический**, которые в свою очередь определяются разными задачами.

Суть обобщающего подхода - в трактовке индекса как показателя среднего изменения уровня исследуемого явления. В этом случае основной задачей, решаемой с помощью индексных показателей, будет характеристика общего изменения многофакторного экономического показателя.

Аналитический подход рассматривает индекс как показатель изменения уровня результативной величины, на которую оказывает влияние величина, изучаемая с помощью индекса. Отсюда и иная задача, которая решается с помощью индексных показателей: выделить влияние одного из факторов в изменении многофакторного показателя.

От содержания изучаемых показателей, методологии расчета первичных показателей, целей и задач исследования зависят и способы построения индексов.

По степени охвата элементов явления индексы делят на *индивидуальные и общие* (сводные).

Индивидуальные индексы (i) - это индексы, которые характеризуют изменение только одного элемента совокупности.

Общий (сводный) индекс (I) характеризует изменение по всей совокупности элементов сложного явления. Если индексы охватывают только часть явления, то их называют *групповыми*.

В зависимости от способа изучения общие индексы могут быть построены или как *агрегатные* (от лат. aggrega - присоединяю) индексы, или как *средние взвешенные* индексы (средние из индивидуальных).

Способ построения агрегатных индексов заключается в том, что при помощи так называемых соизмерителей можно выразить итоговые величины

сложной совокупности в отчетном и базисном периодах, а затем первую сопоставить со второй.

В статистике имеют большое значение индексы **переменного и фиксированного состава**, которые используются при анализе динамики средних показателей.

Индексом *переменного состава* называют отношение двух средних уровней.

Индекс *фиксированного состава* есть средний из индивидуальных индексов. Он рассчитывается как отношение двух стандартизованных средних, где влияние изменения структурного фактора устранено, поэтому данный индекс называют еще индексом постоянного состава.

В зависимости от характера и содержания индексируемых величин различают индексы количественных (объемных) показателей и индексы качественных показателей.

К индексам **количественных** (объемных) показателей относятся такие индексы, как индексы физического объема производства продукции, затрат на выпуск продукции, стоимости продукции, а также индексы показателей, размеры которых определяются абсолютными величинами. Используются различные виды индексов количественных показателей.

Индекс физического объема продукции (ФОП) отражает изменение выпуска продукции. Индивидуальный индекс ФОП отражает изменение выпуска продукции одного вида и определяется по формуле:

$$i_{q_{1/0}} = \frac{q_1}{q_0} \quad (7.1)$$

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

Агрегатный индекс ФОП (предложен Э. Ласпейресом) отражает изменение выпуска всей совокупности продукции, где индексируемой величиной является количество продукции q , а соизмерителем - цена p :

$$I_{q_{1/0}} = \frac{\sum q_1 p_0}{\sum q_0 p_0} \quad (7.2)$$

где q_1 и q_0 - количество выработанных единиц отдельных видов продукции соответственно в отчетном и базисном периодах; p_0 - цена единицы продукции (отдельного вида) в базисном периоде.

При вычислении индекса ФОП в качестве соизмерителей может выступать также себестоимость продукции или трудоемкость.

Средние взвешенные индексы ФОП используются в том случае, если известны индивидуальные индексы объема по отдельным видам продукции и стоимость отдельных видов продукции (или затраты) в базисном или отчетном периоде.

Средний взвешенный арифметический индекс ФОП определяется по формуле:

$$I_{q_{1/0}} = \frac{\sum i_q q_0 p_0}{\sum q_0 p_0} \quad (7.3)$$

где i_q - индивидуальный индекс по каждому виду продукции; $q_0 p_0$ - стоимость продукции каждого вида в базисном периоде.

Средний взвешенный гармонический индекс ФОП:

$$I_{q_{1/0}} = \frac{\sum q_1 p_1}{\sum \frac{1}{i_q} q_1 p_1} \quad (7.4)$$

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

Аналогично рассчитывается индекс затрат на выпуск продукции (ЗВП), который отражает изменение затрат на производство и может быть как индивидуальным, так и агрегатным.

Индивидуальный индекс ЗВП отражает изменение затрат на производство одного вида и определяется по формуле:

$$I_{qz_{1/0}} = \frac{q_1 z_1}{q_0 z_0} \quad (7.5)$$

где z_1 и z_0 - себестоимость единицы продукции искомого вида в текущем и базисном периодах; $q_1 z_1$ и $q_0 z_0$ - суммы затрат на выпуск продукции искомого вида в текущем и базисном периодах.

Агрегатный индекс ЗВП характеризует изменение общей суммы затрат на выпуск продукции за счет изменения количества выработанной продукции и ее себестоимости и определяется по формуле:

$$I_{qz_{1/0}} = \frac{\sum q_1 z_1}{\sum q_0 z_0} \quad (7.6)$$

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

Рассмотрим построение индекса стоимости продукции (СП), который может определяться и как индивидуальный, и как агрегатный.

Индивидуальный индекс СП характеризует изменение стоимости продукции данного вида и имеет вид:

$$i_{qp} = \frac{q_1 p_1}{q_0 p_0} \quad (7.7)$$

где p_1 и p_0 - цена единицы продукции данного вида в текущем и базисном периодах; $q_1 p_1$ и $q_0 p_0$ - стоимость продукции данного вида в текущем и базисном периодах.

Агрегатный индекс СП (товарооборота) характеризует изменение общей стоимости продукции за счет изменения количества продукции и цен и определяется по формуле:

$$I_{qp_{1/0}} = \frac{\sum q_1 p_1}{\sum q_0 p_0} \quad (7.8)$$

Качественные показатели определяют уровень исследуемого итогового показателя и определяются путем соотношения итогового показателя и определенного количественного показателя (например, средняя заработная плата определяется путем соотношения фонда заработной платы и количества работников). К индексам качественных показателей относятся *индексы цен, себестоимости, средней заработной платы, производительности труда*.

Самым распространенным индексом в этой группе является **индекс цен**.

Индивидуальный индекс цен характеризует изменение цен по одному виду продукции и определяется по формуле:

$$i_{p_{1/0}} = \frac{p_1}{p_0} \quad (7.9)$$

где p_1 и p_0 - цена за единицу продукции в текущем и базисном периодах.

Соответственно определяются индексы себестоимости и затрат рабочего времени по каждому виду продукции.

Агрегатный индекс цен определяет среднее изменение цены p по совокупности определенных видов продукции q .

Для характеристики среднего изменения цен на потребительские товары используют *индекс цен, предложенный Э. Ласпейресом*:

$$I_{\text{ил}} = \frac{\sum q_0 p_1}{\sum q_0 p_0} \quad (7.10)$$

где q_0 - потребительская корзина (базовый период); p_0 и p_1 - соответственно цены базисного и отчетного периодов.

Если количество набора продуктов принимается на уровне отчетного периода (q_1), то в этом случае *индекс цен именуется индексом Пааше*:

$$I_{\text{иПааше}} = \frac{\sum q_1 p_1}{\sum q_1 p_0} \quad (7.11)$$

Если известны индивидуальные индексы цен по отдельным видам продукции и стоимость отдельных видов продукции, то применяются средние взвешенные индексы цен (средний взвешенный арифметический и средний взвешенный гармонический индексы цен).

Формула *среднего взвешенного арифметического индекса цен*:

$$I_{\text{сзА}} = \frac{\sum i p_0 q_0}{\sum p_0 q_0} \quad (7.12)$$

где i - индивидуальный индекс по каждому виду продукции; p_0 q_0 - стоимость продукции каждого вида в базисном периоде.

Формула *среднего взвешенного гармонического индекса цен*

$$I_{\text{свг}} = \frac{\sum p_1 q_1}{\sum \frac{1}{p_1} p_1 q_1} \quad (7.13)$$

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

В статистической практике очень широко используется *агрегатный территориальный индекс цен*, который может быть рассчитан по следующей формуле:

$$I_{\text{рвв}} = \frac{\sum p_A q_A}{\sum p_B q_A} \quad (7.14)$$

где $p_A p_B$ - цена за единицу продукции каждого вида соответственно на территории А и В; q_A - количество выработанной или реализованной продукции каждого вида по территории А (в натуральном выражении).

Из формулы видно, что в данном индексе в качестве фиксированного показателя (веса) принят объем продукции территории А. При расчете данного индекса в качестве веса можно принять также объем продукции территории В или суммарный объем продукции двух территорий.

Пример выполнения работы

Задание. Динамика средних цен и объема продажи в крупных супермаркетах города характеризуется следующими данными, представленными в табл. 7.1 ниже.

Таблица 7.1

Данные для расчета

Продукция	Продано продукции, тыс.кг		Средняя цена за 1 кг, тыс.руб.	
	Базисный период	Отчетный период	Базисный период	Отчетный период
Супермаркет №1				
помидоры	4.0	4.2	6.4	7.6
огурцы	2.5	2.4	7.2	8.4
Супермаркет №2				
помидоры	10.0	12.0	7.6	7.0

Вычислите

1. Для супермаркета № 1 (по двум видам продукции вместе):

1.1. Общий индекс товарооборота

1.2. Общий индекс цен

1.3. Общий индекс физического объема товарооборота

Определите в отчетном периоде абсолютный прирост товарооборота и разложите по факторам (за счет изменения цен и объемов продаж овощей) .

2. Для супермаркетов вместе по помидорам:

2.1. Индекс цен переменного состава

2.2. Индекс цен постоянного (фиксированного) состава

2.3. Индекс влияния изменения структуры объема продаж помидор на динамику средней цены

Сформулируйте выводы.

Решение

1.1. Общий индекс товарооборота можно рассчитать по формуле:

$$i_{pq} = \frac{\sum p_1 q_1}{\sum q_0 p_0}$$

где p -цена, q -количество проданной продукции

$$i_{pq} = \frac{\sum p_1 q_1}{\sum q_0 p_0} = \frac{7,6 \cdot 4,2 + 8,4 \cdot 2,4}{6,4 \cdot 4 + 7,2 \cdot 2,5} = 1,194 = 119,4\%$$

1.2 Общий индекс цен вычисляем по формуле:

$$i_p = \frac{\sum p_1 q_1}{\sum p_0 q_1} = \frac{7,6 \cdot 4,2 + 8,4 \cdot 2,4}{6,4 \cdot 4,2 + 7,2 \cdot 2,4} = 1,179 = 117,9\%$$

1.3 Общий индекс цен физического товарооборота можно рассчитать по формуле:

$$i_q = \frac{\sum p_0 q_1}{\sum p_0 q_0} = \frac{6,4 \cdot 4,2 + 7,2 \cdot 2,4}{6,4 \cdot 4 + 7,2 \cdot 2,5} = 1,013 = 101,3\%$$

Эти индексы связаны между собой формулой:

$$i_{pq} = i_p \cdot i_q = 1,179 \cdot 1,013 = 1,194$$

Таким образом, товарооборот увеличился на 19,4%, в том числе за счет увеличения физического объема товарооборота на 1,3 %.

Абсолютный прирост товарооборота:

$$\Delta pq = \sum p_1 q_1 - \sum p_0 q_0 = (7,6 \cdot 4,2 + 8,4 \cdot 2,4) - (6,4 \cdot 4 + 7,2 \cdot 2,5) = 8,48 \text{ млн. р.}$$

В том числе за счет изменения цены:

$$\Delta pq_{(p)} = \sum p_1 q_1 - \sum p_0 q_1 = (7,6 \cdot 4,2 + 8,4 \cdot 2,4) - (6,4 \cdot 4,2 + 7,2 \cdot 2,4) = 7,92 \text{ млн. р.}$$

В том числе за счет изменения продажи товаров:

$$\Delta pq_{(q)} = \sum p_0 q_1 - \sum p_0 q_0 = (6,4 \cdot 4,2 + 7,2 \cdot 2,4) - (6,4 \cdot 4 + 7,2 \cdot 2,5) = 0,56 \text{ млн. р.}$$

Абсолютные приросты связаны между собой формулами:

$$\Delta pq = \Delta pq_{(p)} + \Delta pq_{(q)} = 0,56 + 7,92 = 8,48 \text{ тыс. р.}$$

Таким образом, товарооборот увеличился на 8,48 млн.руб., в том числе за счет увеличения цен на 7,92 млн.р. за счет увеличения физического объема товарооборота на 0,56 млн.р.

2.1. Вычислим для 2-х супермаркетов по помидорам индекс цен переменного состава

$$i_{\bar{p}} = \frac{\sum p_1 q_1}{\sum q_1} + \frac{\sum p_0 q_0}{\sum q_0} = \frac{7,6 \cdot 4,2 + 7,0 \cdot 12,0}{4,2 + 12,0} + \frac{6,4 \cdot 4 + 7,6 \cdot 10,0}{4,0 + 10,0} = 0,986 = 98,6\%$$

Вычислим индекс цен постоянного состава:

$$i_p = \frac{\sum p_1 q_1}{\sum q_1} + \frac{\sum p_0 q_1}{\sum q_1} = \frac{7,6 \cdot 4,2 + 7,0 \cdot 12,0}{4,2 + 12,0} + \frac{6,4 \cdot 4,2 + 7,6 \cdot 12,0}{4,2 + 12,0} = 0,981 = 98,1\%$$

2.2 Вычислим индекс влияния структуры объема продаж помидор на динамику средней цены:

$$i_{cmp} = \frac{\sum p_0 q_1}{\sum q_1} + \frac{\sum p_0 q_0}{\sum q_0} = \frac{7,6 \cdot 4,2 + 7,0 \cdot 12,0}{4,2 + 12,0} + \frac{6,4 \cdot 4,0 + 7,6 \cdot 10,0}{4,0 + 10,0} = 1,004 = 100,4\%$$

Разница между индексами переменного и постоянного состава заключается в том, что индекс переменного состава равен соотношению средних уровней цены, а постоянного характеризует изменение средней цены за счет изменения только цен на каждом рынке.

Таким образом, средняя цена в супермаркетах уменьшилась на 1,4%. Если бы в обоих магазинах структура рынка была одна и та же, средняя цена уменьшилась на 1,9%. Увеличение доли более дорогого супермаркета в структуре продаж увеличило среднюю цену на 0,4%

2.3 Определим общее абсолютное изменение цены помидор:

$$\Delta \bar{p} = \frac{\sum p_1 q_1}{\sum q_1} - \frac{\sum p_0 q_0}{\sum q_0} = \frac{7,6 \cdot 4,2 + 7,0 \cdot 12,0}{4,2 + 12,0} - \frac{6,4 \cdot 4 + 7,6 \cdot 10,0}{4,0 + 10,0} = -0,11 \text{ тыс. р.}$$

Общее абсолютное изменение цены из-за непосредственного изменения уровней цен на помидоры:

$$\Delta \bar{p}_{(p)} = \frac{\sum p_1 q_1}{\sum q_1} - \frac{\sum p_0 q_1}{\sum q_1} = \frac{7,6 \cdot 4,2 + 7,0 \cdot 12,0}{4,2 + 12,0} - \frac{6,4 \cdot 4,2 + 7,6 \cdot 12,0}{4,2 + 12,0} = -0,14 \text{ тыс. р.}$$

Общее абсолютное изменение цены за счет изменения структуры продажи помидор:

$$\Delta \bar{p}_{(cmp)} = \frac{\sum p_0 q_1}{\sum q_1} - \frac{\sum p_0 q_0}{\sum q_0} = \frac{6,4 \cdot 4,2 + 7,6 \cdot 12,0}{4,2 + 12,0} - \frac{6,4 \cdot 4,2 + 7,6 \cdot 10,0}{4,0 + 10,0} = 0,03 \text{ тыс. р.}$$

Таким образом, средняя цена на помидоры снизилась на 0,11 тыс.р., в том числе за счет непосредственного изменения уровней цен на 0,14 тыс.р. Увеличение доли рынка с более дорогими помидорами увеличило результативный показатель на 0,03 тыс.р.

БИБЛИОГРАФИЧЕСКИЙ СПИСОК

1. Шорохова, И. С. Статистические методы анализа: учебное пособие для СПО / И. С. Шорохова, Н. В. Кисляк, О. С. Мариев. — 2-е изд. — Саратов, Екатеринбург: Профобразование, Уральский федеральный университет, 2019. — 298 с. — ISBN 978-5-4488-0520-2, 978-5-7996-2853-6. — Текст: электронный // Электронно-библиотечная система IPR BOOKS: [сайт]. — URL: <http://www.iprbookshop.ru/87873.html>
2. Выгодчикова, И. Ю. Методы процентных вычислений: учебное пособие для СПО / И. Ю. Выгодчикова. — Саратов: Профобразование, Ай Пи Ар Медиа, 2019. — 60 с. — ISBN 978-5-4488-0254-6, 978-5-4497-0012-4. — Текст : электронный // Электронно-библиотечная система IPR BOOKS: [сайт]. — URL: <http://www.iprbookshop.ru/81307.html>
3. Бородачѳв, С. М. Статистические методы в управлении качеством: учебное пособие для СПО / С. М. Бородачѳв; под редакцией О. И. Никонова. — 2-е изд. — Саратов, Екатеринбург : Профобразование, Уральский федеральный университет, 2019. — 86 с. — ISBN 978-5-4488-0411-3, 978-5-7996-2810-9. — Текст: электронный // Электронно-библиотечная система IPR BOOKS: [сайт]. — URL: <http://www.iprbookshop.ru/87874.html>
4. Черткова Е. А., Статистика. Автоматизация обработки информации: Учебное пособие Для СПО / Черткова Е. А. - 2-е изд. ; испр. и доп. - Москва: Издательство Юрайт, 2019. - 195. - (Профессиональное образование). - ISBN 978-5-9916-9342-4: 409.00. URL: <https://www.biblio-online.ru/bcode/437290>

ОГЛАВЛЕНИЕ

ВВЕДЕНИЕ	3
ЛАБОРАТОРНАЯ РАБОТА № 1	
ОСНОВНЫЕ ПОЛОЖЕНИЯ МЕТРОЛОГИИ.....	3
ЛАБОРАТОРНАЯ РАБОТА № 2	
СРЕДНИЕ ВЕЛИЧИНЫ. ПОКАЗАТЕЛИ ВАРИАЦИИ.....	8
ЛАБОРАТОРНАЯ РАБОТА № 3	
АНАЛИЗ ВАРИАЦИОННЫХ РЯДОВ.....	14
ЛАБОРАТОРНАЯ РАБОТА № 4	
ВИДЫ ДИСПЕРСИЙ.....	18
ЛАБОРАТОРНАЯ РАБОТА № 5	
ВЫБОРОЧНОЕ НАБЛЮДЕНИЕ.....	22
ЛАБОРАТОРНАЯ РАБОТА № 6	
АНАЛИЗ РЯДОВ ДИНАМИКИ.....	26
ЛАБОРАТОРНАЯ РАБОТА № 7	
ИНДЕКСЫ КОЛИЧЕСТВЕННЫХ И КАЧЕСТВЕННЫХ ПОКАЗАТЕЛЕЙ.....	34
БИБЛИОГРАФИЧЕСКИЙ СПИСОК.....	41

ОСНОВЫ СТАТИСТИЧЕСКОГО АНАЛИЗА

МЕТОДИЧЕСКИЕ УКАЗАНИЯ
к выполнению лабораторных работ
для студентов специальности 27.02.07 «Управление качеством продукции,
процессов и услуг (по отраслям)» всех форм обучения

Составитель:
Сергеева Светлана Ивановна

Издается в авторской редакции

Подписано к изданию 17.11. 2021.

Уч.-изд. л. 2,3.

ФГБОУ ВО «Воронежский государственный технический университет»

394026 Воронеж, Московский проспект 14