

Электронный архив УГЛТУ

Министерство образования и науки Российской Федерации
ФГБОУ ВПО «Уральский государственный лесотехнический
университет»

Институт экономики и управления

Кафедра Информационных технологий и моделирования

Г.Л. Нохрина

СТАТИСТИКА

Методические указания по выполнению лабораторно-
практического цикла работ для студентов направления 09.03.03
всех форм обучения

ЕКАТЕРИНБУРГ 2015 г.

Оглавление

Лабораторная работа №1. Статистическая сводка и группировка	3
<i>Основные понятия</i>	<i>3</i>
<i>Сравнимость статистических группировок.....</i>	<i>5</i>
<i>Задачи к лабораторной работе №1</i>	<i>6</i>
Лабораторная работа №2. Формы выражения статистических показателей	7
<i>Относительные показатели.....</i>	<i>8</i>
<i>Средние величины</i>	<i>9</i>
<i>Задачи к лабораторной работе №2</i>	<i>9</i>
Лабораторная работа №3. Показатели вариации и анализ частотных распределений.....	12
<i>Правило сложения дисперсий.....</i>	<i>13</i>
<i>Характеристики вариационного ряда</i>	<i>14</i>
<i>Показатели асимметрии и эксцесса</i>	<i>16</i>
<i>Критерии согласия.....</i>	<i>17</i>
<i>Задачи к лабораторной работе №3</i>	<i>17</i>
Лабораторная работа №4. Статистическое изучение взаимосвязи социально - экономических явлений.....	19
<i>Причинность, регрессия, корреляция.....</i>	<i>19</i>
<i>Парная регрессия.....</i>	<i>20</i>
<i>Оценка существенности корреляции</i>	<i>20</i>
<i>Коэффициенты ассоциации и контингенции</i>	<i>21</i>
<i>Задачи к лабораторной работе №4</i>	<i>23</i>
Лабораторная работа №5. Статистическое изучение динамики социально–экономических явлений.....	24
<i>Показатели изменения уровней ряда динамики.....</i>	<i>24</i>
<i>Задачи к лабораторной работе №5</i>	<i>27</i>
Лабораторная работа №6. Экономические индексы.....	30
<i>Индивидуальные индексы.....</i>	<i>31</i>
<i>Сводные индексы</i>	<i>32</i>
<i>Системы индексов</i>	<i>34</i>
<i>Задачи к лабораторной работе №6</i>	<i>36</i>
Рекомендуемая литература.....	37

Лабораторная работа №1. Статистическая сводка и группировка

Основные понятия

Сводка – это комплекс последовательных операций по обобщению конкретных единичных фактов, образующих совокупность, для выявления типичных черт и закономерностей, присущих изучаемому явлению в целом. Проведение сводки выполняют по следующим этапам:

- 1) выбор группировочного признака;
- 2) определение порядка формирования групп;
- 3) разработка системы статистических показателей для характеристики групп и объекта в целом;
- 4) разработка макетов статистических таблиц.

Группировкой называется расчленение единиц изучаемой совокупности на однородные группы по существенным для них признакам. Различают типологические, структурные, аналитические группировки. Группировка, в которой группы расчленяются по одному признаку, называется простой, по двум и более признакам - сложной.

Признаки можно разделить на факторные и результативные. **Факторными** называются признаки, оказывающие влияние на изменение результативных. **Результативными** называются признаки, изменяющиеся под влиянием факторных.

Группировочным признаком называется признак, по которому производится разбиение единиц совокупности на отдельные группы. В основание группировки могут быть положены как количественные, так и атрибутивные признаки. Число групп зависит от задачи исследования и вида показателя, положенного в основание группировки, численности совокупности, степени вариации признака. Чем больше групп, тем точнее будет воспроизведен характер исследуемого объекта. Однако слишком большое число групп затрудняет выявление закономерностей. Определение числа групп можно осуществить и математическим путем с использованием формулы Стерджесса: $n = 1 + 3,322Lg(N)$, где n – число групп, N – число единиц совокупности.

После определения числа групп определяют интервалы группировки.

Интервал – это значение варьируемого признака, лежащее 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** – наибольшее значение признака в интервале. **Величина интервала** представляет собой разность между верхней и нижней границами интервала. Интервалы группировки в зависимости от их величины бывают равные и неравные. Неравные интервалы, в свою очередь, делятся на прогрессивно возрастающие, прогрессивно убывающие, произвольные и специализированные.

Если вариация признака проявляется в сравнительно узких границах и распределение носит равномерный характер, то строят группировку с равными интервалами. Величина равного интервала определяется по формуле

$$h = \frac{R}{n} = \frac{X_{\max} - X_{\min}}{n}$$

Если максимальное и минимальное значения сильно отличаются от смежных с ними вариантов в упорядоченном ряду значений группировочного признака, то для определения величины интервала используют значения, несколько больше, чем минимум, и несколько меньше, чем максимум. Существуют правила записи числа шага интервала. Если величина интервала представляет собой величину, которая имеет один знак до запятой то, в этом случае полученные значения округляют до десятых (например, 0,88 до 0,9; 1,585 до 1,6). Если рассчитанная величина интервала имеет две значащие цифры до запятой, то это значение округляют до целого (например, 15,985 до 16). В случае, когда рассчитанная величина интервала представляет собой трехзначное, четырехзначное и так далее число, то величину округляют до ближайшего числа, кратного 100 или 50 (например, 557 до 600).

Если размах вариации признака совокупности велик и значения признака варьируются неравномерно, то необходимо использовать группировку с неравными интервалами. Величина интервалов, изменяющихся в арифметической прогрессии, определяется следующим образом: $h_{i+1} = h_i + a$, а в геометрической прогрессии: $h_{i+1} = h_i q$, где a - константа, имеющая знак «+» или «-», q - константа, $q > 1$ или $q < 1$.

Интервалы 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.

От группировки следует отличать классификацию. Отличительной чертой классификации является то, что в ее основу кладется атрибутивный признак.

Ряды распределения – это упорядоченное распределение единиц совокупности на группы по определенному признаку. В зависимости от

признака, положенного в основу ряда распределения, различают атрибутивные и вариационные ряды распределения.

Атрибутивными называют ряды распределения, построенные по качественным признакам, т.е. признакам, не имеющим числового выражения.

Вариационными называют ряды распределения, построенные по количественному признаку. 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 или каждой группы вариационного ряда.

Сравнимость статистических группировок

Группировки, построенные за один и тот же период времени, но для разных регионов или, наоборот, для одного региона, но за два разных периода времени, могут оказаться несопоставимыми из-за различного числа выделенных групп или неодинаковости границ интервалов. **Вторичная группировка** – это операция по образованию новых групп на основе ранее выполненной группировки. Наиболее простым способом является изменение (чаще укрупнение) первоначальных интервалов. Второй способ называется долевой группировкой и состоит в образовании новых групп на основе закрепления за каждой группой определенной доли единиц совокупности. Например, пусть имеется следующее распределение работников фирмы по уровню дохода.

№ группы	Группы работников по уровню дохода, руб.	Число работников, чел.
1	до 400	16
2	400-1000	20
3	1000-1800	44
Итого		80

Необходимо произвести перегруппировку данных, образовав новые группы с интервалами до 500, 500-1000, свыше 1000 руб.

В первую группу войдет полностью первая группа и часть второй группы. Чтобы образовать группу до 500 необходимо от интервала второй группы взять 100 руб. Величина интервала этой группы составляет 600 руб., значит необходимо взять от нее $1/6$. Аналогичную часть во вновь образуемую группу надо взять и от численности работников, т.е. $20/6=3$ человека. Тогда в первой группе будет $16+3=19$ человек. Во вторую вновь образуемую группу войдут работники прежней второй группы за вычетом отнесенных к первой, т.е. $20-3=17$ человек. В новую третью группу войдут все работники третьей группы. Таким образом, имеем новую группировку.

№ группы	Группы работников по уровню дохода,	Число работников,
----------	-------------------------------------	-------------------

	руб.	чел.
1	до 500	19
2	500-1000	17
3	1000 и более	44
Итого		80

Задачи к лабораторной работе №1

Задача 1.1

Пользуясь формулой Стерджесса, определить интервал группировки сотрудников фирмы по уровню доходов, если общая численность сотрудников составляет 20 человек, а минимальный и максимальный доходы равны 500 и 3000 руб.

Задача 1.2

Известны данные об объеме импорта РФ с отдельными странами Европы в 1997 г. (в фактически действовавших ценах , млн дол США). Используя эти данные, построить интервальный вариационный ряд распределения стран Европы по объему импорта в РФ, выделив 4 группы стран с равными открытыми интервалами.

79 84 76 11 61

14 23 09 596 46

45 50 002 611 39

96 45 00 11 627

Задача 1.3

Известны данные по основным показателям деятельности крупнейших банков одной из областей России (данные условные). Построить группировку коммерческих банков по величине собственного капитала, выделив не более 5 групп с равными интервалами. Рассчитать по каждой группе сумму активов, собственный капитал, балансовую прибыль. Результаты представить в табличной форме, сформулировать выводы.

№ п/п	Сумма активов, ден.ед.	Собственный капитал, ден.ед.	Привлеченные ресурсы, ден.ед.	Балансовая прибыль, ден.ед.
1	645,60	12,00	27,10	8,10
2	636,90	70,40	56,30	9,50
3	629,00	41,00	95,70	38,40
4	619,60	120,80	44,80	38,40

№ п/п	Сумма активов, ден.ед.	Собственный капитал, ден.ед.	Привлеченные ресурсы, ден.ед.	Балансовая прибыль, ден.ед.
5	616,40	49,40	108,70	13,40
6	614,40	50,30	108,10	30,10
7	608,60	70,00	76,10	37,80
8	601,10	52,40	26,30	41,10
9	600,20	42,00	46,00	9,30
10	600,00	27,30	24,40	39,30
11	592,90	72,00	65,50	8,60
12	591,70	22,40	76,00	40,50
13	585,50	39,30	106,90	45,30
14	578,60	70,00	89,50	8,40
15	577,50	22,90	84,00	12,80
16	553,70	119,30	89,40	44,70
17	543,60	49,60	93,80	8,80
18	542,00	88,60	26,70	32,20
19	517,00	43,70	108,10	20,30
20	516,70	90,50	25,20	12,20

Задача 1.4

По данным задачи 1.3 построить группировку коммерческих банков по двум признакам: величине балансовой прибыли и сумме активов. По каждой группе и подгруппе определить число банков, величину балансовой прибыли и другие показатели.

Лабораторная работа №2. Формы выражения статистических показателей

Статистический показатель представляет собой количественную характеристику социально-экономических явлений и процессов. Все показатели делятся на абсолютные, относительные и средние. **Абсолютные показатели** отражают физические размеры, массу, площадь, объем, протяженность, временные характеристики. Эти показатели являются именованными числами и выражаются в натуральных, стоимостных или трудовых единицах измерения.

Относительный показатель представляет собой результат деления одного абсолютного показателя на другой и выражает соотношение между количественными характеристиками. При расчетах показатель, находящийся в числителе, называется текущим, или сравниваемым. Показатель в знаменателе называется основанием, или базой. Относительный показатель может выражаться в коэффициентах, процентах, промилле.

Относительные показатели

Относительный показатель динамики (ОПД) показывает, во сколько раз текущий уровень превышает предшествующий (базисный):

$$ОПД = \frac{\text{Текущий уровень}}{\text{Предшествующий или базисный уровень}}$$

Различают ОПД с постоянной и переменной базой сравнения. Если сравнение осуществляется с одним и тем же базисным уровнем, получают ОПД с постоянной базой. При расчетах ОПД с переменной базой (цепных) основание ОПД последовательно меняется.

Относительный показатель плана (ОПП) и реализации плана (ОПРП):

$$ОПП = \frac{\text{Уровень, планируемый на } (i + 1) - \text{й период}}{\text{Уровень, достигнутый в } i - \text{м периоде}};$$

$$ОПРП = \frac{\text{Уровень, достигнутый в } (i + 1) - \text{й период}}{\text{Уровень, планируемый на } (i + 1) - \text{й период}}.$$

Первый показатель характеризует напряженность плана. Второй показатель отражает фактический объем производства в процентах по сравнению с плановым уровнем.

Между показателями существует связь: $ОПД = ОПП \cdot ОПРП$.

Относительный показатель структуры (ОПС) представляет собой соотношение структурных частей изучаемого объекта и их целого:

$$ОПС = \frac{\text{Показатель, характеризующий часть совокупности}}{\text{Показатель по всей совокупности в целом}}.$$

Относительный показатель координации (ОПК) - это отношение одной части совокупности к другой части этой же совокупности:

$$ОПК = \frac{\text{Показатель, характеризующий } i - \text{ю часть совокупности}}{\text{Показатель, характеризующий часть совокупности, выбранную в качестве базы сравнения}}$$

Относительный показатель интенсивности (ОПИ) характеризует степень распространения изучаемого явления и представляет собой отношение исследуемого показателя к размеру присущей ему среды:

$$ОПИ = \frac{\text{Показатель, характеризующий явление } A}{\text{Показатель, характеризующий среду распространения явления } A}$$

Относительный показатель сравнения (ОПС) представляет собой соотношение одного и того же абсолютного показателя, характеризующего раз-

ные объекты: $ОПС = \frac{\text{Показатель, характеризующий объект } A}{\text{Показатель, характеризующий объект } B}$.

Средние величины

Наиболее распространенной формой статистических показателей является средняя величина. Показатели в форме средней величины дают обобщающую характеристику однотипных явлений по одному из варьирующих признаков. Средняя величина отражает то общее, что присуще всем единицам совокупности. Выделяют среднюю агрегатную, среднюю арифметическую, среднюю гармоническую величины. Все средние, за исключением средней агрегатной, могут рассчитываться в двух вариантах – как взвешенные или невзвешенные величины:

$$\text{средняя агрегатная величина: } \bar{X} = \frac{\sum W_i}{\sum f_i};$$

$$\text{средняя арифметическая взвешенная величина: } \bar{X} = \frac{\sum x_i f_i}{\sum f_i};$$

$$\text{средняя арифметическая невзвешенная величина: } \bar{X} = \frac{\sum x_i}{n};$$

$$\text{средняя гармоническая взвешенная величина: } \bar{X} = \frac{\sum w_i}{\sum \frac{w_i}{x_i}};$$

$$\text{средняя гармоническая невзвешенная величина: } \bar{X} = \frac{n}{\sum \frac{1}{x_i}}, \text{ где}$$

$$W_i = x_i f_i;$$

x_i – i – й вариант осредняемого признака;

f_i – вес i – го варианта;

n – объем совокупности.

Задачи к лабораторной работе №2

Задача 2.1

Имеются данные о производствах бумаги в РФ. Вычислить относительные показатели динамики с переменной и постоянной базой сравнения.

Годы	1992	1993	1994	1995
Произведено бумаги, тыс т	3603	2882	2215	2771

Задача 2.2

Добыча нефти и угля в РФ во 2-м квартале 1996 г характеризуется следующими данными.

Топливо	Объем добычи,		
	апрель	май	июнь
Нефть	23,80	25,00	24,20
Уголь	23,20	20,20	18,70

Теплота сгорания нефти 45 МДж/кг, угля - 26,8 МДж/кг. Сделать пересчет в условное топливо (29,3 МДж/кг) и провести анализ изменения совокупной добычи этих ресурсов.

Задача 2.3

Объем продаж в АО "ЛОМО" в 1995 г в сопоставимых ценах вырос по сравнению с предшествующим годом на 5% и составил 146 млрд руб.

Определить объем продаж в 1994 г.

Задача 2.4

Торговая фирма планировала в 1997 г. по сравнению с 1996 г. увеличить оборот на 14,5%. Выполнение установленного плана составило 102,7%. Определить относительный показатель динамики оборота.

Задача 2.5

Объем продаж компании Samsung в странах СНГ в первом полугодии 1996 г составил 250 млн долл. В целом же за год компания планировала реализовать товаров на 600 млн долл. Вычислить относительный показатель плана на второе полугодие.

Задача 2.6

Предприятие планировало увеличить выпуск продукции в 1997 году по сравнению с 1996 г. на 18%. Фактический же объем продукции составил 112,3% от прошлогоднего уровня. Определить ОПРП.

Задача 2.7

Известна структура произведенных затрат металлургических комбинатов России. Вычислить относительные показатели координации.

Статья затрат	Удельный вес в общих затратах, %
Сырье и материалы	33
Топливо и энергия	13
Оплата труда	4
Амортизация	10
Прочие расходы	40

Задача 2.8

Численность врачей в РФ характеризуется следующими данными (на начало года, тыс чел).

Врачи	1981	1995
Всего	560,70	663,10
в том числе:		
Педиатров	63,90	75,40
Терапевтов	127,70	169,00

Провести анализ изменения обеспеченности населения врачами, если известно, что численность постоянного населения на начало 1981 г. составляла 139 млн чел, в том числе в возрасте до 14 лет - 30,1 млн чел, а на начало 1995 г. - соответственно 147,9 и 31,8 млн чел.

Задача 2.9

Производственные мощности металлургических комбинатов и уровень их использования в 1995 г. характеризуется следующими данными.

Комбинат	Мощность, млн т/год			Загрузка, %		
	Чугун	Сталь	Прокат	Чугун	Сталь	Прокат
Магнитогорский	10,50	18,50	12,00	41,30	63,40	53,40
Череповецкий	9,50	13,50	11,50	60,50	70,40	58,50
Новолипецкий	9,50	9,90	7,00	71,40	73,70	89,00
Нижнетагильский	7,00	8,00	4,50	64,20	70,60	82,90
Западно-Сибирский	6,00	6,90	4,30	69,30	75,40	82,50
Челябинский	4,00	7,00	4,00	36,40	44,90	43,70
Кузнецкий	3,70	4,80	3,50	74,20	67,00	76,70
Орско-Халиловский	3,40	4,60	3,40	62,40	64,70	61,40

Рассчитать среднюю отраслевую загрузку производственных мощностей по каждому виду продукции.

Задача 2.10

Имеются данные о стоимости коттеджей, предлагаемых к продаже на начало 1996 г. Рассчитать среднюю цену 1 м².

Цена 1 м ² , долл. США	Общая площадь, тыс. м ²
300-400	29,40
400-500	20,50
500-600	7,30
600-700	7,00
700-800	4,00

Задача 2.11

Имеются данные об успеваемости студентов вуза. Определить долю отличников в общей численности студентов.

№ факультета	Доля отличников в общей численности студентов факультета	Доля студентов в общей численности студентов вуза
1	0,12	0,2
2	0,06	0,43
3	0,17	0,08
4	0,09	0,29

Задача 2.12

Использование складских помещений города характеризуется след данными. Вычислить средний % загрузки.

Группы складских помещений по площади, тыс м ²	Число помещений	Общая занятая площадь, тыс м ²
0-5	3	5,2
5-10	21	108,0
10-15	17	163,6
15-20	9	101,2
20-25	5	65,3
25-30	3	40,6
30-35	4	55,4
35и выше	2	29,0

Лабораторная работа №3. Показатели вариации и анализ частотных распределений

Величина вариации признака в статистической совокупности характеризует ее однородность. Для измерения вариации используют следующие величины: размах вариации, среднее линейное отклонение, средний квадрат отклонений (дисперсия), среднее квадратическое отклонение, коэффициент вариации.

Размах вариации $R = X_{\max} - X_{\min}$, где X_{\max} - наибольшее значение варьирующего признака; X_{\min} - наименьшее значение признака.

Среднее линейное отклонение представляет собой среднюю величину из отклонений вариантов признака от их средней. Можно рассчитать

невзвешенное среднее линейное отклонение $\bar{d} = \frac{\sum |X_i - \bar{X}|}{n}$ и взвешенное

среднее линейное отклонение $\bar{d} = \frac{\sum |X_i - \bar{X}| f_i}{\sum f_i}$.

Дисперсия представляет собой средний квадрат отклонений индивидуальных значений признака от их средней величины. Дисперсия вычисляется по формулам простой и взвешенной:

$$\sigma^2 = \frac{\sum (X_i - \bar{X})^2}{n}; \quad \sigma^2 = \frac{\sum (X_i - \bar{X})^2 f_i}{\sum f_i}.$$

Среднее квадратическое отклонение (имеет размерность осредняемого признака) - корень второй степени из дисперсии:

$$\sigma = \sqrt{\frac{\sum (X_i - \bar{X})^2}{n}}; \quad \sigma = \sqrt{\frac{\sum (X_i - \bar{X})^2 f_i}{\sum f_i}}.$$

Относительные показатели вариации используются для сравнения колеблемости различных признаков одной и той же совокупности или же при сравнении колеблемости одного и того же признака в нескольких совокупностях. Базой для сравнения служит средняя арифметическая. Совокупность считается однородной, если коэффициент вариации не превышает 33%. Различают следующие относительные показатели вариации:

коэффициент осцилляции $V_R = \frac{R}{\bar{X}} 100\%$;

линейный коэффициент вариации: $V_{\bar{d}} = \frac{\bar{d}}{\bar{X}} 100\%$;

коэффициент вариации $V_{\sigma} = \frac{\sigma}{\bar{X}} 100\%$.

Правило сложения дисперсий

Если данные представлены в виде аналитической группировки, то можно вычислить дисперсию общую, межгрупповую, внутригрупповую. Общая дисперсия измеряет вариацию признака во всей совокупности под влиянием всех факторов, обуславливающих эту вариацию:

$$\sigma^2 = \frac{\sum (X_i - \bar{X})^2}{\sum f_i}.$$

Межгрупповая дисперсия характеризует

систематическую вариацию, т.е. различия в величине изучаемого признака, возникающие под влиянием признака – фактора, положенного в основание группировки. Межгрупповая дисперсия рассчитывается по формуле

$$\delta_x^2 = \frac{\sum (X_i - \bar{X})^2 n_i}{\sum n_i}, \text{ где } X_i \text{ и } n_i$$

- средние и численности по отдельным группам. Внутригрупповая дисперсия рассчитывается по формуле

$$\sigma_i^2 = \frac{\sum (\bar{X}_i - \bar{X})^2 n_i}{\sum n_i}$$

и отражает случайную вариацию, т.е. часть вариации, происходящую под влиянием неучтенных факторов и не зависящую от признака – фактора, положенного в основание группировки. Средняя из

$$\bar{\sigma}_i^2 = \frac{\sum \sigma_i^2 n_i}{\sum n_i}$$

внутригрупповых имеет вид:

Существует закон, связывающий три вида дисперсии: $\sigma^2 = \bar{\sigma}_i^2 + \delta_x^2$.

Согласно этому правилу общая дисперсия, возникающая под влиянием всех факторов, равна сумме дисперсий, возникающих под влиянием прочих факторов, и дисперсии, возникающей за счет группировочного признака.

На основании правила сложения дисперсий можно определить показатель тесноты связи между группировочным (факторным) и результативными признаками. Он называется эмпирическим корреляционным

отношением, рассчитывается по формуле $\eta = \sqrt{\frac{\delta_x^2}{\sigma^2}}$.

Характеристики вариационного ряда

Для характеристики структуры вариационных рядов применяют структурные средние показатели: мода, медиана, квартили, децили.

Мода – значение признака, наиболее часто встречающееся в исследуемой совокупности. Для интервальных рядов мода рассчитывается по формуле

$$M_o = X_{M_o} + i_{M_o} \frac{f_{M_o} - f_{M_o-1}}{(f_{M_o} - f_{M_o-1}) + (f_{M_o} - f_{M_o+1})},$$

X_{M_o} - нижняя граница значения интервала, содержащего моду;

i_{M_o} - величина модального интервала;

f_{M_o} - частота модального интервала;

f_{M_o-1} - частота интервала, предшествующего модальному;

$f_{M_{o+1}}$ - частота интервала, следующего за модальным.

Для дискретных рядов модой будет значение варианта с наибольшей частотой.

Медиана – это значение признака, приходящееся на середину ранжированной (упорядоченной) совокупности. Медиана интервального ряда рассчиты-

вается по формуле
$$M_e = X_{Me} + i_{Me} \frac{\frac{1}{2}(\sum f) - S_{Me-1}}{f_{Me}},$$

X_{Me} - нижняя граница значения интервала, содержащего медиану;

i_{Me} - величина медианного интервала;

$\sum f$ - сумма частот;

S_{Me-1} - сумма накопленных частот, предшествующих медианному интервалу;

f_{Me} - частота медианного интервала.

Вычисление медианы в дискретных рядах распределения имеет специфику. Если такой ряд имеет нечетное число членов, то медианой будет вариант, находящийся в середине ранжированного ряда. Если ранжированный ряд состоит из четного числа членов, то медианой будет среднее арифметическое из двух значений признака, расположенных в середине ряда.

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

Нижний квартиль Q_1 , отделяет $\frac{1}{4}$ часть совокупности с наименьшими значениями признака, а **верхний квартиль** Q_3 , отсекает $\frac{1}{4}$ часть с наибольшими значениями признака. Это означает, что 25% единиц совокупности будут меньше по величине Q_1 ; 25% единиц будут заключены между Q_1 и Q_2 ; 25% - между Q_2 и Q_3 , а остальные 25% превосходят Q_3 . **Средним квартилем** Q_2 является медиана. Для расчета квартилей по

вариационному ряду используют формулы:
$$Q_1 = X_{Q_1} + i \frac{\frac{1}{4}(\sum f) - S_{Q_1-1}}{f_{Q_1}};$$

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

где X_{Q_1} - нижняя граница значения интервала, содержащего нижний квартиль (интервал определяется по накопленной частоте, первой превышающей 25%);

X_{Q_3} - нижняя граница значения интервала, содержащего верхний квартиль (интервал определяется по накопленной частоте, первой превышающей 75%);

i - величина интервала, содержащего нижний квартиль;

S_{Q_1-1} - накопленная частота интервала, предшествующего интервалу, содержащему нижний квартиль;

S_{Q_3-1} - то же для верхнего квартиля;

f_{Q_1} - частота интервала, содержащего нижний квартиль

f_{Q_3} - то же для верхнего квартиля.

Показатели асимметрии и эксцесса

Степень асимметрии может быть определена с помощью **коэффициента асимметрии**: $A_s = \frac{\bar{X} - M_o}{\sigma}$, где \bar{X} - среднее арифметическое ряда распределения, M_o - мода, σ - среднее квадратическое отклонение. При симметричном (нормальном) распределении $\bar{X} = M_o$, следовательно, коэффициент асимметрии равен нулю. Если $A_s > 0$, то имеет место правосторонняя асимметрия, если $A_s < 0$, то - левосторонняя асимметрия. Коэффициент асимметрии может меняться от -3 до 3 . В практических расчетах в качестве показателя асимметрии применяется соотношение $A_s = \frac{\mu_3}{\sigma^3}$. Принято считать, что асимметрия выше $0,5$ (независимо от знака) считается значительной. Асимметрия меньше $0,25$ – незначительная.

Для симметричных распределений может быть рассчитан **показатель эксцесса**: $E_k = \frac{\mu_4}{\sigma^4} - 3$. При симметричном распределении $E_k = 0$. Если $E_k > 0$, распределение считается островершинным, в противном случае –

плосковершинным.

Критерии согласия

Имея дело с эмпирическим распределением, стремятся описать эмпирический ряд с помощью математической модели. Количественная характеристика соответствия может быть получена с помощью статистических показателей – критериев согласия.

Критерий согласия Пирсона

Величина критерия вычисляется по формуле
$$\chi^2 = \sum \frac{(f_{\text{Э}} - f_{\text{Т}})^2}{f_{\text{Т}}}$$
, где $f_{\text{Т}}$ и $f_{\text{Э}}$ - теоретические и эмпирические частоты. С помощью величины критерия по таблицам определяется плотность вероятности $P(\chi^2)$. Входами в таблицу являются значения χ^2 и число степеней свободы $\gamma = n - 1$. При $P(\chi^2) > 0,5$ считается, что эмпирическое и теоретические распределения близки. При $P(\chi^2) \in (0,2; 0,5)$ совпадение между ними считается удовлетворительным, в остальных случаях – недостаточным.

Критерий Романовского

Критерий также используется для проверки близости эмпирического и теоретического распределений и определяется следующим образом:

$$C = \frac{\chi^2 - \gamma}{\sqrt{2\gamma}}$$
, где χ^2 - критерий Пирсона, γ - число степеней свободы. При $C < 3$ различие несущественно.

Критерий Колмогорова

Вычисляется по формуле:
$$\lambda = \frac{D}{\sqrt{\sum f}}$$
, где D - максимальное значение разности между накопленными эмпирическими и теоретическими частотами; $\sum f$ - сумма эмпирических частот. Необходимым условием использования этого критерия является достаточно большое число наблюдений (не меньше ста).

Задачи к лабораторной работе №3

Задача 3.1

Распределение студентов одного из факультетов характеризуется следующими данными.

Возраст студентов	17	18	19	20	21	22	23	24
Число студентов	20	80	90	110	130	170	90	60

Вычислить: размах вариации, среднее линейное отклонение дисперсию, среднее квадратическое отклонение, относительные показатели вариации возраста.

Задача 3.2

Определить среднюю длину пробега автофургона фирмы и вычислить все показатели вариации, если известны:

Длина пробега за 1 рейс, км	30-50	50-70	70-90	90-110	110-130	130-150
Число рейсов	20	25	14	18	90	6

Задача 3.3

Имеются данные выборочного обследования студентов вуза. Вычислить абсолютные и относительные показатели вариации.

Затраты времени на дорогу, ч	до 0,5	0,5-1	1-1,5	1,5-2	свыше 2
Число студентов, % к итогу	7	18	32	37	6

Задача 3.4

По данным таблицы о распределении пряжи по крепости нити вычислить все виды дисперсий, коэффициент корреляции.

1 группа пряжи (менее крепкая)		2 группа пряжи (более крепкая)	
Крепость нити, г	Число проб	Крепость нити, г	Число проб
120-130	2	200-210	25
130-140	6	210-220	28
140-150	8	220-230	16
150-160	15	230-240	10
160-170	25	240-250	8
170-180	29	250-260	7
180-190	35	260-270	5
190-200	30		

Задача 3.5

Имеются данные о распределении семей сотрудников по числу детей. Вычислить все дисперсии. Проверить правильность вычислений.

Число детей в семье	Число семей сотрудников по отделам
----------------------------	---

	1	2	3
0	4	7	5
1	6	10	13
2	3	3	3
3	2	1	0

Задача 3.6

Имеются данные о распределении семей города по числу детей. Рассчитать коэффициенты асимметрии и эксцесса.

Группы семей по числу детей	0	1	2	3	4	5
Число семей, % к итогу	10	26	29	17	13	5

Задача 3.7

С целью исследования качества деталей на предприятии проверена партия из 100 деталей. Определить моду, медиану, квартили, децили.

Группы деталей по весу, г	40-50	50-60	60-70	70-80	80-90	90-100	100-110	110-120
Число деталей	2	4	12	18	21	24	11	8

Лабораторная работа №4. Статистическое изучение взаимосвязи социально - экономических явлений

Социально–экономические явления представляют собой результат одновременного воздействия большого числа причин. При изучении этих явлений необходимо выявлять главные, основные причины. В основе первого этапа статистического изучения лежит качественный анализ явления, связанный с анализом его природы. Второй этап – построение модели связи. Он базируется на методах статистики: группировки, средних величин таблиц. Третий этап – интерпретация результатов.

Причинность, регрессия, корреляция

В статистике различают функциональную связь, и стохастическую зависимость. Функциональной называют такую связь, при которой определенному значению факторного признака соответствует одно и только одно значение результативного признака. Если причинная зависимость проявляется не в каждом отдельном случае, а в общем, при большом числе наблюдений, то такая зависимость называется стохастической. Частным случаем стохастической

связи является корреляционная связь, при которой изменение среднего значения результативного признака обусловлено изменением факторных признаков. По направлению связи выделяют прямую и обратную. По степени тесноты связи различают слабую, умеренную, сильную связи. По аналитическому выражению выделяют связи линейные и нелинейные.

Парная регрессия

Парная регрессия характеризует связь между двумя признаками: результативным и факторным. Определить тип уравнения можно, исследуя зависимость графически. Оценка параметров уравнения регрессии осуществляется методом наименьших квадратов, при котором минимизируется сумма квадратов отклонений эмпирических значений результативного признака от теоретических, полученных по уравнению регрессии $S = \sum (Y - \hat{Y}_x)^2 \rightarrow \min$.

Модель регрессии может быть построена как по индивидуальным значениям признака, так и по сгруппированным данным. Для выявления связи между признаками по достаточно большому числу наблюдений используется корреляционная таблица. В этой таблице можно отобразить только парную связь, т.е. связь результативного признака с одним фактором, и на ее основе построить уравнение регрессии и определить показатели тесноты связи. Для составления корреляционной таблицы парной связи необходимо:

- 1) данные предварительно сгруппировать по обоим признакам;
- 2) построить таблицу, по строкам в которой отложить группы результативного, а по столбцам – группы факторного признака.

Если оба признака X и Y располагаются в возрастающем порядке, а частоты f_{xy} сосредоточены по диагонали сверху вниз направо, то можно судить о прямой связи между признаками, в противном случае – об обратной связи. О тесноте связи между признаками можно судить по кучности расположения частот вокруг диагонали (насколько заполнены клетки таблицы в стороне от нее). Если клетки заполнены большими числами, то связь слабая. Чем ближе частоты располагаются к одной из диагоналей, тем теснее связь. Если в расположении частот нет системности, то можно судить об отсутствии связи.

Оценка существенности корреляции

Теснота связи при линейной зависимости измеряется с помощью **линейного коэффициента корреляции**. На практике применяются различные модификации формул расчета данного коэффициента:

$$r_{xy} = \frac{\overline{xy} - \bar{x}\bar{y}}{\sigma_x\sigma_y} = \frac{n\sum xy - \sum x\sum y}{\sqrt{[n\sum x^2 - (\sum x)^2][n\sum y^2 - (\sum y)^2]}}$$

Степень тесноты связи через линейный коэффициент связи оценивается следующим образом (табл. 1).

Таблица 1

Величина коэффициента корреляции	Характер связи
До $ \pm 0,3 $	Практически отсутствует
$ \pm 0,3 - \pm 0,5 $	Слабая
$ \pm 0,5 - \pm 0,7 $	Умеренная
$ \pm 0,7 - \pm 1,0 $	Сильная

Линейный коэффициент корреляции изменяется в пределах от -1 до 1 . При этом интерпретацию выходных значений коэффициента корреляции можно представить в табл. 2.

Таблица 2

Значение линейного коэффициента корреляции	Характер связи	Интерпретация связи
$r = 0$	Отсутствует	
$0 < r < 1$	Прямая	С увеличением X увеличивается Y
$-1 < r < 0$	Обратная	С увеличением X уменьшается Y
$r = 1$	Функциональная	Каждому значению факторного признака строго соответствует одно значение результативного признака

Значимость линейного коэффициента корреляции проверяется на основе

t -критерия Стьюдента: $t_p = \sqrt{\frac{r^2}{1-r^2}} \cdot (n-2) = \frac{|r|}{\sqrt{1-r^2}} \cdot \sqrt{n-2}$. Если $t_p > t_{kp}$ (табличное), то линейный коэффициент корреляции статистически значим, что свидетельствует и о статистической существенности зависимости между факторами.

Коэффициенты ассоциации и контингенции

Для определения связи двух качественных признаков, каждый из которых состоит только из двух групп, применяются коэффициенты ассоциации и контингенции. Для их вычисления строится таблица (табл.3), которая показывает связь между двумя явлениями, каждое из которых может быть альтерна-

ТИВНЫМ.

Таблица 3

a	b	a+b
c	d	c+d
a+c	b+d	a+b+c+d

Коэффициенты вычисляются по формулам:

ассоциации: $K_a = \frac{ad - bc}{ad + bc}$; контингенции: $K_k = \frac{ad - bc}{\sqrt{(a+b)(b+d)(a+c)(c+d)}}$.

Коэффициент контингенции всегда меньше коэффициента ассоциации. Связь считается подтвержденной, если $K_a > 0,5$ или $K_k > 0,3$.

Если каждый из качественных признаков состоит более чем из двух групп, то для определения тесноты связи возможно применение коэффициентов взаимной сопряженности Пирсона и Чупрова, которые вычисляются по следующим формулам:

$$K_{\Pi} = \sqrt{\frac{\varphi^2}{1 + \varphi^2}}; \quad K_{\text{Ч}} = \sqrt{\frac{\varphi^2}{\sqrt{(K_1 - 1)(K_2 - 1)}}}; \quad \varphi^2 = \sum \frac{n_{xy}^2}{n_x n_y} - 1,$$

φ^2 - показатель взаимной сопряженности;

n_x, n_y - итоговые частоты соответствующих столбцов;

K_1, K_2 - число значений (групп) первого и второго признака.

Чем ближе величины коэффициентов к 1, тем теснее связь. Значения частот рассчитываются с помощью табл. 4.

Таблица 4

Признаки	I	II	III	Всего
I	n_{xy}	...	n_{xy}	n_x
II		n_x
III	n_{xy}	...	n_{xy}	n_x
Итого	n_y	n_y	n_y	n

Существует модификация коэффициента Пирсона через расчет χ^2 -критерия. Коэффициент взаимной сопряженности вычисляется по формуле:

$$K_{II} = \sqrt{\frac{\chi^2}{n + \chi^2}}, \text{ где } \chi^2 = n \left\{ \sum \frac{n_{xy}^2}{xy n_x n_y} - 1 \right\}. \text{ Другой модификацией коэффициента сопряженности Чупрова является } K_{Ч} = \sqrt{\frac{\chi^2}{n \cdot \sqrt{(K_1 - 1)(K_2 - 1)}}}, \text{ где } K_1, K_2$$

- соответственно число строк и столбцов в таблице, n - число наблюдений.

Задачи к лабораторной работе №4

Задача 4.1

Имеются следующие данные о стоимости основных фондов и среднесуточной переработке сырья.

Стоимость основных фондов, тыс. руб.	Среднесуточная переработка сырья, тыс. ц			
	3-5	5-7	7-9	9-11
300-400	2			
400-500	5	2		
500-600	2	4	6	
600-700		2	3	5
700-800			2	2

Определить вид корреляционной зависимости, найти параметры уравнения регрессии, определить тесноту связи. Проанализировать результаты.

Задача 4.2

Взаимосвязь между стоимостью активной части основных фондов и затратами на производство работ по 35 строительным фирмам представлена в таблице. Найти линию регрессии.

Затраты на производство строительно – монтажных работ, % к стоимости активной части основных фондов	Стоимость активной части основных фондов, тыс. руб.			
	50-100	100-150	150-200	200-250
1-5				
5-9			2	4
9-13		2	6	4
13-17		5	3	

17-21	2	2		
-------	---	---	--	--

Задача 4.3

В ходе проведенного обследования оценки уровня жизни работающих на предприятиях различной формы собственности было опрошено 100 респондентов. Рассчитать коэффициенты ассоциации и контингенции. Сформулировать выводы из анализа коэффициентов.

Форма собственности предприятия	Удовлетворенность уровнем жизни	
	Вполне удовлетворен	Не удовлетворен
Государственное	30	55
Частное	10	5
Итого	40	60

Лабораторная работа №5. Статистическое изучение динамики социально–экономических явлений

Для отображения изменений явлений во времени строят ряды динамики. **Ряды динамики** представляют собой ряды изменяющихся во времени значений статистического показателя, расположенных в хронологическом порядке. В зависимости от способа выражения уровней ряды динамики подразделяются на ряды абсолютных, относительных, средних величин. В зависимости от того, выражают уровни ряда состояние явления на определенные моменты времени или его величину за определенные интервалы времени, различают моментные и интервальные ряды. Ряды динамики могут быть с равноотстоящими (по времени) уровнями и неравноотстоящими (по времени) уровнями.

Показатели изменения уровней ряда динамики

Абсолютные и относительные показатели

Для выражения абсолютной скорости роста (снижения) уровня ряда динамики исчисляют статистический показатель – абсолютный прирост базовый ($\Delta_B = y_i - y_o$) и цепной ($\Delta_{Ц} = y_i - y_{i-1}$).

Интенсивность изменения уровней ряда динамики оценивается темпом роста: $T_p = \frac{y_i}{y_o} 100\%$ или $T_p = \frac{y_i}{y_{i-1}} 100\%$.

Для выражения изменения величины абсолютного прироста уровней ряда динамики в относительных величинах определяется темп прироста:

$$T_{np} = \frac{\Delta}{y_{i-1}} 100\%; T_{np} = \frac{\Delta}{y_0} 100\% \text{ или } T_{np} = T_p - 100\% .$$

Показатель абсолютного значения одного процента прироста определяется следующим образом: $|\%| = \frac{\Delta}{T_{np}} = 0,01 y_{i-1}$. Расчет показателя имеет смысл только на цепной основе.

Средние показатели

Методы расчета среднего уровня ряда динамики зависят от его вида и способа получения статистических данных. В интервальном ряду динамики с равноотстоящими уровнями во времени расчет среднего уровня ряда производится по формуле средней арифметической простой: $\bar{y} = \frac{\sum y}{n}$. Если интервальный ряд динамики имеет неравноотстоящие уровни, то средний уровень ряда вычисляется по формуле $\bar{y} = \frac{\sum yt}{\sum t}$, где t – число периодов времени, в течение которых уровень не изменяется.

Для моментного ряда с равноотстоящими уровнями средняя хронологическая рассчитывается по формуле $\bar{y} = \frac{\frac{y_1}{2} + y_2 + \dots + y_{n-1} + \frac{y_n}{2}}{n-1}$ где n – число уровней ряда.

Средняя хронологическая для разноотстоящих уровней моментного ряда динамики вычисляется по формуле

$$\bar{y} = \frac{(y_1 + y_2)t_1 + (y_2 + y_3)t_2 + (y_3 + y_4)t_3 + \dots + (y_{n-1} + y_n)t_{n-1}}{2\sum t_i} .$$

Определение среднего абсолютного прироста производится по цепным абсолютным приростам по формуле: $\bar{\Delta} = \frac{\sum \Delta_{Ц}}{m-1} = \frac{y_n - y_0}{n-1}$.

Среднегодовой темп прироста вычисляется по формуле средней геометрической: $\bar{T}_p = \sqrt[m]{K_1 K_2 K_3 \dots K_m} = \sqrt[n]{\frac{y_n}{y_0}}$.

Среднегодовой темп прироста $\bar{T}_{np} = \bar{T}_p - 100\%$.

Приемы обработки и анализа рядов динамики

При анализе рядов динамики в случаях, когда уровни ряда несопостави-

мы, возникает необходимость смыкания рядов, т.е. объединения двух и более рядов, характеризующих изменение явления, в один ряд. Существует несколько способов приведения рядов динамики к сопоставимому виду.

Первый способ состоит в следующем:

- 1) определить коэффициент соотношения уровней двух рядов;
- 2) умножить на этот коэффициент уровни первого ряда.

Другой способ смыкания заключается в том, что уровни года, в котором произошли изменения, как до изменений, так и после изменений принимаются за 100%, а остальные пересчитываются в процентах по отношению к этим уровням соответственно.

Компоненты ряда динамики

Ряд динамики может быть подвержен влиянию факторов эволюционного, осциллятивного характера и факторов разного воздействия. Влияние эволюционного характера – это изменения, определяющие общее направление развития. Такие изменения называются тенденцией развития, или трендом. Влияние осциллятивного характера – это циклические или сезонные колебания. Таким образом, можно выделить следующие компоненты ряда динамики: **тренд Т, циклическую К, сезонную S, случайные колебания Е.**

Выявление основной тенденции ряда динамики

Рассмотрим три метода выявления основной тенденции.

Метод укрупнения интервалов основан на укрупнении периодов времени.

Метод скользящей средней состоит в замене абсолютных данных средними арифметическими за определенные периоды. Для каждого конкретного ряда алгоритм расчета скользящей средней следующий:

- 1) определить интервал сглаживания, т.е. число входящих в него уровней m ($m < n$);
- 2) вычислить среднее значение уровней, образующих интервал сглаживания;
- 3) если m – четное число, то необходимо провести центрирование, т.е. найти среднее из средних для отнесения полученного уровня к определенной дате;
- 4) сдвинуть интервал сглаживания на одну точку вправо, провести вычисления среднего значения, снова произвести сдвиг и т.д.

Метод аналитического выравнивания: уровни ряда динамики выражаются в виде функции времени. Выбор функции производится на основе

анализа характера закономерностей динамики данного явления. Выравнивание может быть осуществлено по любому рациональному многочлену.

Выявление сезонных колебаний

Для выявления сезонных колебаний обычно анализируются месячные и квартальные уровни ряда динамики за год или несколько лет. При изучении сезонных колебаний используются индексы сезонности. Совокупность этих показателей отражает сезонную волну. Для вычисления индексов сезонности применяют различные методы. Для внутригодовой динамики используется метод постоянной средней:

- 1) для каждого года рассчитывается средний уровень;
- 2) уровень каждого месяца сопоставляется (в процентах) с полученным

значением с помощью индекса сезонности: $I_s = \frac{Y_i}{\bar{Y}} 100\%$.

Для исключения элемента случайности для выявления закономерности колебаний используют данные за несколько лет. Тогда для вычисления индекса сезонности используют следующий алгоритм:

- 1) для каждого месяца рассчитывается средняя величина уровня (\bar{Y}_i);
- 2) из полученных значений вычисляется среднемесячный уровень для всего ряда (\bar{Y});

- 3) вычисляется индекс сезонности: $I_s = \frac{\bar{Y}_i}{\bar{Y}} 100\%$.

Задачи к лабораторной работе №5

Задача 5.1

Имеются данные о численности населения и производстве мяса в России. Определить среднюю численность населения за каждый год; построить произвольный ряд динамики производства мяса на душу населения и средние уровни рядов динамики.

Показатели	1990	1991	1992	1993	1994	1995	1996
Численность населения на начало года, млн чел	147,4	148,5	148,7	148,7	148,4	148,3	148
Производство мяса в убойном виде, млн т	10,11	9,38	8,26	7,51	6,86	5,9	

Задача 5.2

Списочная численность работников фирмы в 1997 г. составила: на 1 января - 530 чел., на 1 марта - 570 чел., на 1 июня - 520 чел., на 1 сентября - 430 чел., на 1 января 1998 г - 550 чел. Вычислить среднегодовую численность сотрудников в 1997 г

Задача 5.3

Имеются данные по объединению о производстве промышленной продукции за 1992 - 1997 гг (в сопоставимых ценах). Для анализа динамики определить: средний уровень ряда динамики; темпы роста и прироста; для каждого года абсолютное значение 1% прироста. Результаты оформить в табличном виде.

Года	Стоимость продукции, ден ед	Года	Стоимость продукции, ден ед
1992	67,7	1995	77,9
1993	73,2	1996	81,9
1994	75,7	1997	84,4

Задача 5.4

Производство электроэнергии в регионе в 1990 - 1997 гг характеризуется следующими данными, млрд кВт/ч:

1990	1991	1992	1993	1994	1995	1996	1997
915	976	1038	1111	1150	1202	1239	1294

Для анализа ряда динамики:

- 1) определить показатели, характеризующие динамику производства электроэнергии по годам к базисному 1991 г.:
 - 1.1) темпы роста;
 - 1.2) темпы прироста;
 - 1.3) абсолютные приросты;
- 2) рассчитать для каждого года абсолютное значение 1% прироста.

Результаты изложить в табличной форме.

Задача 5.5

Розничный товарооборот во всех каналах реализации в области увеличился в 1996 г. по сравнению с 1995 г. на 20%, а в 1997 г. по сравнению с 1996 г. - еще на 10%. Определить розничный товарооборот в области в 1995, 1996, 1997 гг., если абсолютный прирост розничного товарооборота в 1996 г. по сравнению с 1995 г. составил 3600 млн руб.

Задача 5.6

Используя взаимосвязь показателей динамики, определить уровни ряда

динамики и недостающие в таблице базисные показатели в регионе за 1989 - 1997 гг.

Годы	Производство часов, млн шт	Базисные показатели		
		Абсолютный прирост, млн шт	Темп роста, %	Темп прироста, %
1989	55,10		100,00	
1990		2,80		
1991			110,30	
1992				14,90
1993				17,10
1994			121,10	
1995		13,50		
1996	90,00			
1997		14,00		25,40

Задача 5.7

До 1992 г. в состав производственного объединения входили 20 предприятий. В 1992 г. в него влились еще 4 предприятия. Провести смыкание ряда динамики, используя следующие данные.

	1989	1990	1991	1992	1993	1994	1995	1996	1997
Реализованная продукция 20 предприятий, млн руб.	448,7	462,8	465,8	491,6					
Реализованная продукция 24 предприятий, млн руб.				559,5	578,7	580,5	610	612,9	915,5

Задача 5.8

Провести выравнивание данных о производстве электроэнергии за 1990 - 1997 гг. по показательной кривой. Построить график. Предполагая, что выявленная тенденция сохранится, определить ожидаемый объем производства на ближайшие годы.

Года	Производство, кВт	Года	Производство, кВт
1990	915	1994	1150
1991	976	1995	1202
1992	1038	1996	1239

1993	1111	1997	1294
------	------	------	------

Задача 5.9

Имеются данные о розничном товарообороте во всех каналах реализации в регионе, млрд руб.

Месяцы	1995	1996	1997
Январь	7,4	7,8	8,3
Февраль	7,9	8,2	8,6
Март	8,7	9,2	9,7
Апрель	8,2	8,6	9,1
Май	7,9	8,3	8,8
Июнь	8,2	8,7	9,1
Июль	8,3	8,8	9,3
Август	8,8	9,3	9,9
Сентябрь	8,7	8,9	9,3
Октябрь	8,8	8,2	9,9
Ноябрь	8,3	8,8	9,8
Декабрь	9	9,5	9,3

Для изучения общей тенденции розничного товарооборота региона по месяцам за 1995 - 1997 гг провести :

- 1) преобразование исходных данных, путем укрупнения периодов времени:
 - 1.1) в квартальные уровни;
 - 1.2) в годовые уровни;
- 2) сглаживание квартальных уровней с помощью скользящей средней.

Построить графики.

Задача 5.10

Используя данные задачи 5.9 для анализа внутригодовой динамики розничного товарооборота региона по месяцам 1995 - 1997 гг. определить индексы сезонности с применением:

- 1) 2-х месячной скользящей средней;
- 2) аналитического выравнивания по прямой.

Изобразить сезонную волну графически.

Лабораторная работа №6. Экономические индексы

Экономический индекс – это относительная величина, которая характеризует изменение исследуемого явления во времени, пространстве или по сравнению с некоторым эталоном.

По степени охвата явления индексы бывают индивидуальные и сводные (общие). По характеру объема исследования общие индексы подразделяются на индексы количественных (объемных) и качественных показателей. В зависимости от формы построения различают индексы агрегатные и средние.

Если в качестве базы сравнения используется уровень за какой-либо предшествующий период, получают динамический индекс; если же базой является уровень того же явления по другой территории – территориальный индекс. По виду весов индексы бывают базисные и цепные. По составу явления выделяют индексы постоянного (фиксированного) и переменного составов.

Индивидуальные индексы

Индивидуальные индексы служат для характеристики изменения отдельных элементов сложного явления. В зависимости от экономического назначения индивидуальные индексы бывают физического объема продукции, себестоимости, цен, трудоемкости.

Индекс физического объема продукции рассчитывается по формуле $i_q = \frac{q_1}{q_0}$. Этот индекс показывает, во сколько раз возрос (уменьшился) выпуск в отчетном периоде по сравнению с базисным. В знаменателе может быть плановое, нормативное или эталонное значение, принятое за базу сравнения.

Индексы других показателей строятся аналогично. Индивидуальный индекс цен $i_p = \frac{p_1}{p_0}$ характеризует изменение цены одного определенного товара в текущем периоде по сравнению с базисным. Индивидуальный индекс себестоимости единицы продукции $i_z = \frac{z_1}{z_0}$ показывает изменение себестоимости.

Производительность труда может быть измерена количеством продукции, производимой в единицу времени (v), или затратами рабочего времени на производство единицы продукции (t). Поэтому можно построить индекс количества продукции, произведенной в единицу времени $i_v = \frac{v_1}{v_0} = \frac{q_1}{T_1} : \frac{q_0}{T_0}$ и ин-

декс производительности труда по трудовым затратам $i_t = \frac{t_0}{t_1}$. Для характери-

стики производительности труда часто используют индивидуальный индекс выработки продукции в стоимостном выражении на одного рабочего:

$$i_w = \frac{w_1}{w_0} = \frac{q_1 p}{T_1} : \frac{q_0 p}{T_0}, \text{ где } p\text{- сопоставимые цены.}$$

Индивидуальный индекс стоимости продукции отражает, во сколько раз изменилась стоимость какого-либо товара в текущем периоде по сравнению с

базисным, и определяется по формуле: $i_{pq} = \frac{p_1q_1}{p_0q_0}$.

Изменение численности рабочих показывает индивидуальный индекс численности рабочих $i_T = \frac{t_1q_1}{t_0q_0} = \frac{T_1}{T_0}$.

Сводные индексы

Агрегатные индексы

Сводный индекс – это сложный относительный показатель, который характеризует динамику явления, составные части которого непосредственно несоизмеримы. Агрегатный индекс характеризует среднее изменение социально–экономического явления. В агрегатной форме непосредственно сравниваются две суммы одноименных показателей. Индексируемой величиной называется признак, изменение которого изучается. Вес индекса – это величина, служащая для целей соизмерения индексируемых величин. Методика построения агрегатного индекса предусматривает решение трех вопросов:

- 1) какая величина будет индексируемой;
- 2) по какому составу разнородных элементов необходимо исчислить индекс;
- 3) что будет служить весом при расчете индекса.

Рассмотрим три индекса:

индекс стоимости продукции или товарооборота: $I_{pq} = \frac{\sum p_1q_1}{\sum p_0q_0}$;

индекс физического объема продукции: $I_q = \frac{\sum q_1p_0}{\sum q_0p_0}$;

сводный индекс цен: $I_p = \frac{\sum p_1q_1}{\sum p_0q_1}$, где q - количество продукции p цена.

Между индексами существует связь: $I_{pq} = I_p I_q$.

Числитель и знаменатель сводного индекса цен можно интерпретировать следующим образом: числитель представляет собой сумму денег, фактически уплаченную покупателем за товара в текущем периоде, а знаменатель показывает, какую сумму покупатели заплатили бы за те же товары, если бы цены не изменились. Разность числителя и знаменателя отражает величину экономии или перерасхода от изменения цен: $E = \sum p_1q_1 - \sum p_0q_1$.

Располагая данными о трудоемкости и объемах производства можно рассчитать: сводный индекс производительности труда (по трудоемкости):

$I_w = \frac{\sum t_0 q_1}{\sum t_1 q_1}$. Этот индекс связан с индексом затрат рабочего времени (труда) и

с индексом физического объема продукции, взвешенным по трудоемкости:

$$I_w I_T = I_q \text{ или } I_q = \sum \frac{t_0 q_1}{t_1 q_1} \sum \frac{T_1}{T_0} = \sum \frac{q_1 t_0}{q_0 t_0}.$$

При расчете сводного индекса производительности труда в стоимостном выражении (по выработке) количество продукции, произведенной за каждый период взвешивают по ценам, принятым за сопоставимые. Индекс рассчитывают по формуле $I_w = \frac{\sum q_1 p}{\sum T_1} : \frac{\sum q_0 p}{\sum T_0}$.

Средние индексы

Средний индекс – это индекс, вычисленный как средняя величина из индивидуальных индексов. При вычислении средних индексов используют две формы средних: арифметическую и гармоническую.

Средний арифметический индекс тождествен агрегатному индексу, если весами индивидуальных индексов будут слагаемые знаменателя агрегатного индекса.

Средний арифметический индекс физического объема продукции вычисляется по формуле $I_q = \frac{\sum i_q p_0 q_0}{\sum p_0 q_0}$. Весами в формуле является стоимость продукции базисного периода.

Средний арифметический индекс производительности труда определяется следующим образом: $I_t = \frac{\sum i_t t_1 q_1}{\sum t_1 q_1} = \frac{\sum i_t T_1}{\sum T_1}$. Весами являются общие затраты

времени на производство продукции в текущем периоде. При анализе производительности труда используется и другой средний арифметический индекс -

индекс Струмилина: $I_v = \frac{\sum \frac{q_1}{T_1} : \frac{q_0}{T_0}}{\sum T_1}$.

Средний гармонический индекс тождествен агрегатному индексу, если индивидуальные индексы будут взвешены с помощью слагаемых числителя

агрегатного индекса. Например, индекс себестоимости можно вычислить по

$$\text{формуле } I_z = \frac{\sum z_1 q_1}{\sum \frac{z_1 q_1}{i_z}}, \text{ а индекс цен: } I_p = \frac{\sum p_1 q_1}{\sum \frac{p_1 q_1}{i_p}}.$$

Системы индексов

Выбор базы и весов индексов

Индексы могут использоваться для анализа динамики явлений за ряд последовательных периодов. В этом случае для достижения сопоставимости они должны рассчитываться по единой схеме. Такая схема расчета индексов за несколько временных периодов называется системой индексов. В зависимости от информационной базы и целей исследования индексная система может строиться в четырех вариантах. Рассмотрим систему индексов на примере индекса цен, рассчитанного за «n» периодов.

1. Цепные индексы цен с переменными весами:

$$I_{p\ 1/0} = \frac{\sum p_1 q_1}{\sum p_0 q_1}; I_{p\ 2/1} = \frac{\sum p_2 q_2}{\sum p_1 q_2}; I_{p\ 3/2} = \frac{\sum p_3 q_3}{\sum p_2 q_3}; \dots I_{p\ n/n-1} = \frac{\sum p_n q_n}{\sum p_{n-1} q_{n-1}}.$$

2. Цепные индексы цен с постоянными весами:

$$I_{p\ 1/0} = \frac{\sum p_1 q_0}{\sum p_0 q_0}; I_{p\ 2/1} = \frac{\sum p_2 q_0}{\sum p_1 q_0}; I_{p\ 3/2} = \frac{\sum p_3 q_0}{\sum p_2 q_0}; \dots I_{p\ n/n-1} = \frac{\sum p_n q_0}{\sum p_{n-1} q_0}.$$

3. Базисные индексы цен с переменными весами:

$$I_{p\ 1/0} = \frac{\sum p_1 q_1}{\sum p_0 q_1}; I_{p\ 2/0} = \frac{\sum p_2 q_2}{\sum p_0 q_2}; I_{p\ 3/0} = \frac{\sum p_3 q_3}{\sum p_0 q_3}; \dots I_{p\ n/0} = \frac{\sum p_n q_n}{\sum p_0 q_n}.$$

4. Базисные индексы цен с постоянными весами:

$$I_{p\ 1/0} = \frac{\sum p_1 q_1}{\sum p_0 q_0}; I_{p\ 2/0} = \frac{\sum p_2 q_2}{\sum p_0 q_0}; I_{p\ 3/0} = \frac{\sum p_3 q_3}{\sum p_0 q_0}; \dots I_{p\ n/0} = \frac{\sum p_n q_n}{\sum p_0 q_0}.$$

Индексы структурных сдвигов

При изучении динамики качественных показателей приходится определять изменение средней величины индексируемого показателя. Это изменение обусловлено взаимодействием двух факторов – изменением значения индексируемого показателя у отдельных групп единиц и изменением структуры явления. Под изменением структуры явления понимается изменение доли отдельных групп единиц совокупности в общей их численности. Так как на изменение среднего значения показателя оказывают воздействие два фактора, возникает задача определить степень влияния каждого из факторов на общую

динамику средней. Эта задача решается с помощью индексного метода, т.е. путем построения системы индексов, в которую включены три индекса: переменного состава, постоянного состава и структурных сдвигов.

Индексом переменного состава называется индекс, выражающий соотношение средних уровней изучаемого явления, относящихся к разным периодам. Например, индекс цен переменного состава рассчитывается по формуле

$$I_{ПС}^p = \frac{\sum p_1 q_1}{\sum q_1} \cdot \frac{\sum p_0 q_0}{\sum q_0}.$$

Индекс постоянного (фиксированного) состава – это индекс, исчисленный с весами, зафиксированными на уровне одного какого-либо периода, и показывающий изменение только индексируемой величины. Индекс цен фиксированного состава имеет вид:

$$I_{ФС}^p = \frac{\sum p_1 q_1}{\sum p_0 q_1}.$$

Под **индексом структурных сдвигов** понимают индекс, характеризующий влияние изменения структуры изучаемого явления на динамику среднего уровня этого явления. Для данного примера индекс имеет вид:

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

Между данными индексами имеется следующая связь:

$$I_{ПС} = I_{ФС} I_{СТР}.$$

Индексы пространственно – территориального сопоставления

Территориальные индексы служат для сравнения показателей в пространстве, т.е. по предприятиям, районам. Веса территорий имеют равные основания использоваться при расчете индекса. Для избежания такой неопределенности в качестве весов принимаются объемы проданных товаров по вместе взятым регионам: $Q = q_a + q_b$. Территориальный индекс цен в данном случае

рассчитывается по формуле $I_{pb/a} = \frac{\sum p_b Q}{\sum p_a Q}$. Второй способ расчета террито-

риальных индексов учитывает соотношение весов сравниваемых территорий. При этом способе выполняются следующие шаги:

1. расчет средней цены каждого товара по двум территориям, вместе

взятым: $\bar{p}_i = \frac{\sum p_i q_i}{\sum q_i};$

2. расчет территориального индекса: $I_{pb/a} = \frac{\sum \bar{p}_b q_b}{\sum \bar{p}_a q_a}.$

Задачи к лабораторной работе №6

Задача 6.1

Имеются данные о ценах на уголь и объемах его производства в РФ во 2-ом квартале 1996 г. При условии 100% реализации угля в каждом месяце определить цепные и базисные индивидуальные индексы цен, физического объема реализации и товарооборота.

Месяц	Цена за 1 т, тыс руб	Произведено, млн т
Апрель	120	23,2
Май	121	20,2
Июнь	116	18,7

Задача 6.2

Имеются следующие данные о реализации молочных продуктов на городском рынке. Рассчитать сводные индексы цен, товарооборота, физического объема реализации.

Продукт	Товарооборот, тыс руб		Изменение цены в декабре по сравнению с ноябрем, %
	Ноябрь	Декабрь	
Молоко	9,7	6,3	2,1
Сметана	4,5	4	3,5
Творог	12,9	11,5	4,2

Задача 6.3

По промышленному предприятию имеются следующие данные

Изделие	Общие затраты на производство в 1997 г., млн руб.	Изменение себестоимости изделия в 1997 г. по сравнению с 1996, %
Электромясорубка	1234	6
Кух. комбайн	5877	8,4
Миксер	980	1,3

Определить изменение себестоимости продукции в 1997 г по сравнению с 1996 г. и обусловленный этим размер экономии или дополнительных затрат предприятия.

Задача 6.4

Имеются данные о реализации картофеля на рынках города. Рассчитать индекс цен переменного состава, индекс цен фиксированного состава, индекс структурных сдвигов.

Рынок	Январь		Февраль	
	Цена за 1 кг, руб	Продано, ц	Цена за 1 кг, руб	Продано, ц
1	2,2	24,5	2,4	21,9
2	2	18,7	2,1	18,8
3	1,9	32	1,9	37,4

Задача 6.5

Уровень рыночных цен на молочные продукты и объем их реализации в двух городах характеризуется следующими данными.

Продукт	Город А		Город В	
	Цена за кг, руб	Продано, т	Цена за 1 кг, руб	Продано, т
Молоко	4	76	4	68
Масло	22	45	24	39
Творог	20	60	23	55
Сыр	18	32	16	24

Рассчитать двумя способами территориальный индекс цен города А по отношению к городу В.

Задача 6.6

Известны данные по промышленному предприятию за два года.

Вид продукции	Произведено, тыс шт		Среднесписочное число рабочих, чел	
	1996	1997	1996	1997
1	18,5	19,3	46	51
2	24,2	23,9	43	45

Определить индекс физического объема продукции, индекс производительности труда, индекс затрат труда

Рекомендуемая литература

1. Теория статистики: Учебник/Под ред. Проф. Р.А.Шмойловой.- М.:Финансы и статистика, 2010.
2. Практикум по теории статистики: Учебное пособие./Под ред. Проф. Р.А.Шмойловой.-М.:Финансы и статистика, 2010.