

Министерство образования Республики Беларусь
Учреждение образования
«Брестский государственный технический университет»
Кафедра высшей математики

Практикум по общей теории статистики

для студентов экономических специальностей

ЧАСТЬ I

Брест 2010

УДК 311 (076.5)

Настоящее пособие содержит теоретические положения, образцы решений примеров и практические задания по общей теории статистики. Материалы данного пособия могут быть использованы на занятиях со студентами экономических специальностей всех форм обучения.

Издаётся в 2-х частях. Часть 1.

Составители: С.Т. Гусева, доцент,
М.Г. Журавель, ассистент,
Л.С. Золотухина, ст. преподаватель,
Г.В. Шамовская, ассистент,
Е.В. Кузьмина, ассистент

Рецензент: доцент кафедры информатики и прикладной математики УО «Брестский государственный университет им. А.С. Пушкина», к.ф.-м.н. Мирская Е.И.

Глава 1. СТАТИСТИЧЕСКАЯ СВОДКА И ГРУППИРОВКА. СТАТИСТИЧЕСКИЕ ТАБЛИЦЫ. ГРАФИЧЕСКОЕ ИЗОБРАЖЕНИЕ СТАТИСТИЧЕСКИХ ДАННЫХ

Статистическая сводка – систематизация единичных фактов, позволяющая перейти к обобщающим показателям, относящихся ко всей изучаемой совокупности и ее частям и осуществлять анализ и прогнозирование изучаемых явлений и процессов.

Группировка – это процесс образования однородных групп на основе разбиения статистической совокупности на части или объединение изучаемых единиц в частные совокупности по существенным для них признакам. Группировка выполняется по одному признаку (простая группировка) и по нескольким признакам (комбинированная группировка). Группировочные признаки делят на атрибутивные (профессия рабочих, социальная группа населения) и количественные (стаж работы, размер дохода). 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.

В соответствии с задачами группировки различают следующие ее виды: типологическая, структурная, аналитическая.

Типологическая группировка – это расчленение разнородной совокупности на отдельные качественно однородные группы и выявление на этой основе экономических типов явлений.

Структурная группировка предназначена для изучения состава однородной совокупности по какому-либо варьирующему признаку.

Группировка, выявляющая взаимосвязи между изучаемыми явлениями и их признаками, называется **аналитической группировкой**. Особенностью аналитической группировки является то, что единицы группируются по факторному признаку; каждая выделенная группа характеризуется средними значениями результативного признака. **Факторными** называются признаки, оказывающие влияние на изменение результативных. **Результативными** называются признаки, изменяющиеся под влиянием факторных.

Построение группировки начинается с определения состава группировочных признаков. После определения основания группировки, следует решить вопрос о количестве групп. 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. Если группировка проводится по **количественному** признаку, то тогда необходимо обратить особое внимание на число единиц исследуемого объекта и степень колеблемости группировочного признака.

Определение числа групп можно осуществить и математическим путем с использованием формулы Стерджесса: $n = 1 + 3,322 \cdot \lg N$, где n – число групп; N – число единиц совокупности.

Когда определено число групп, то следует определить интервалы группировки.

Интервал – это значение варьирующего признака, лежащее в определенных границах. **Нижней границей** интервала называется наименьшее значение признака в интервале, а **верхней границей** – наибольшее значение признака в интервале.

Если вариация признака проявляется в сравнительно узких границах и распределение носит равномерный характер, то строят группировку с **равными интервалами**. Величина равного интервала определяется по следующей формуле: $h = \frac{R}{n} = \frac{x_{\max} - x_{\min}}{n}$, где

x_{\max} и x_{\min} – максимальное и минимальное значения признака в совокупности; n – число групп, размах вариации $R = x_{\max} - x_{\min}$.

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: $h_{i+1} = h_i + a$; в геометрической прогрессии: $h_{i+1} = h_i \cdot q$, где a и q – константы.

Закрытыми называются интервалы, у которых имеются верхняя и нижняя границы. У **открытых** интервалов указана только одна граница: верхняя – у первого, нижняя – у последнего.

При определении границ интервалов статистических группировок необходимо исходить из того, что изменение количественного признака приводит к появлению нового качества. В этом случае граница интервала должна устанавливаться там, где происходит переход от одного качества в другое. Это достигается путем использования группировок со **специализированными интервалами**.

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

В статистической таблице выделяются два элемента: **подлежащее** (помещается в первой вертикальной или горизонтальной графе) – перечень единиц или групп, на которые подразделена вся масса единиц наблюдения; **сказуемое** – цифры, при помощи которых характеризуются выделенные в подлежащем единицы или группы. Над таблицей помещается заголовок, отражающий в сжатой форме ее основное содержание, время, место, к которому относятся изложенные в таблице данные.

После определения группировочного признака и границ групп строится ряд распределения.

Статистический ряд распределения – это упорядоченное распределение единиц совокупности на группы по определенному варьирующему признаку. В зависимости от признака, положенного в основу образования ряда распределения, различают **атрибутивные** (построенные по качественному признаку) и **вариационные** (построенные по количественному признаку) ряды распределения.

Любой вариационный ряд состоит из двух элементов: вариант и частот. **Вариантами** считаются отдельные значения признака, которые он принимает в вариационном ряду. **Частоты** – это численности отдельных вариант или каждой группы вариационного ряда. Распределение единиц совокупности по дискретному признаку характеризует дискретный вариационный ряд, по непрерывному признаку – интервальный вариационный ряд.

Графические методы в статистике являются способом наглядного изображения результатов статистической сводки и обработки массового материала.

Для анализа рядов распределения строят полигон, гистограмму, огиву и кумуляту распределения.

Для построения **полигона** в прямоугольной системе координат по оси абсцисс в одинаковом масштабе откладываются ранжированные значения варьирующего признака, а по оси ординат наносится шкала для выражения величины частот. Полученные на пересечении абсцисс и ординат точки соединяют прямыми линиями и получают ломанную линию, называемую полигоном частот.

При построении **гистограммы** на оси абсцисс откладываются величины интервалов, а частоты изображаются прямоугольниками, построенными на соответствующих интервалах. Высота столбиков должна быть пропорциональна частотам. Если середины верхних сторон прямоугольников соединить прямыми, то гистограмма может быть преобразована в полигон распределения.

При построении **кумуляты** интервального вариационного ряда по оси абсцисс откладываются варианты ряда, а по оси ординат – накопленные частоты, которые наносят на поле графика в виде перпендикуляров к оси абсцисс в верхних границах интервалов. Затем эти перпендикуляры соединяют и получают ломанную линию, т.е. кумуляту. Если при графическом изображении вариационного ряда в виде кумуляты оси поменять местами, то получим **огиву**.

Для графического изображения взаимосвязи между явлениями на оси абсцисс необходимо поместить значения признака-фактора, а на оси ординат – значения признака-результата.

Разновидностью линейной диаграммы является **радиальная диаграмма**, которая применяется для изображения рядов динамики при наличии в них сезонных колебаний.

1.1. Решение типовых задач

Задача 1.1. Произведем анализ 30 самых надежных среди малых и средних коммерческих банков одного из регионов (на 01.01.08 г.), применяя метод группировок (табл. 1.1).

Таблица 1.1 – Основные показатели деятельности коммерческих банков одного из регионов на 01.01.08 г. (цифры условные)

№ банка	Капитал	Работающие активы	Уставный капитал	№ банка	Капитал	Работающие активы	Уставный капитал
1	20710	11706	2351	16	55848	54435	7401
2	19942	19850	17469	17	10344	21430	4266
3	9273	2556	2626	18	16651	41119	5121
4	59256	43587	2100	19	15762	29771	9998
5	24654	29007	23100	20	6753	10857	2973
6	47719	98468	18684	21	22421	53445	3415
7	24236	25595	5265	22	13614	22625	4778
8	7782	6154	2227	23	9870	11744	5029
9	38290	79794	6799	24	24019	27333	6110
10	10276	10099	3484	25	22969	70229	5961
11	35662	30005	13594	26	75076	124204	17218
12	20702	21165	8973	27	56200	90367	20454
13	8153	16683	2245	28	60653	101714	10700
14	10215	9115	9063	29	14813	18245	2950
15	23459	31717	3572	30	41514	127732	12092

В качестве группировочного признака возьмем уставный капитал. Образует четыре группы банков с равными интервалами. Величину интервала определим по формуле

$$h = \frac{R}{n} = \frac{x_{\max} - x_{\min}}{n} = \frac{23100 - 2100}{4} = 5250 \text{ (тыс. руб.)}$$

Обозначим

- 2100-7350 1-я группа;
- 7350-12600 2-я группа;
- 12600-17850 3-я группа;
- 17850-23100 4-я группа.

После того как определен группировочный признак – уставный капитал, задано число групп – 4 и образованы сами группы, необходимо отобрать показатели, которые характеризуют группы, и определить их величины по каждой группе. Показатели, характеризующие банки, разносятся по указанным группам, и подсчитываются итоги по группам. Результаты группировки заносятся в таблицу, и определяются общие итоги по совокупности единиц наблюдения по каждому показателю (табл. 1.2).

Таблица 1.2 – Группировка малых и средних коммерческих банков одного из регионов по величине уставного капитала на 01.01.08 г.

№ группы	Группы банков по величине уставного капитала, тыс. руб.	Число банков, ед.	Работающие активы, тыс. руб.	Капитал, тыс. руб.	Уставный капитал, тыс. руб.
1	2100 - 7350	18	504898	342889	71272
2	7350 - 12600	6	343932	204694	58227
3	12600 - 17850	3	174059	130680	48281
4	17850 - 23100	3	217842	128573	62238
Итого		30	1240731	806836	240018

Структурная группировка коммерческих банков на основе данных таблицы 1.2 представлена в таблице 1.3.

Таблица 1.3 – Группировка малых и средних коммерческих банков одного из регионов по величине уставного капитала на 01.01.08 г.

№ группы	Группы банков по величине уставного капитала, тыс. руб.	Число банков, %	Работающие активы, % к итогу	Капитал, % к итогу	Уставный капитал, % к итогу
1	2100 - 7350	60	40,7	42,5	29,7
2	7350 - 12600	20	27,7	25,4	24,3
3	12600 - 17850	10	14,0	16,2	20,1
4	17850 - 23100	10	17,6	15,9	25,9
Итого		100	100,0	100,0	100,0

Из таблицы 1.3 видно, что в основном преобладают малые банки – 60%, на долю которых приходится 42,5% всего капитала. Более конкретный анализ взаимосвязи показателей можно сделать на основе аналитической группировки (таблица 1.4).

Таблица 1.4 – Группировка малых и средних коммерческих банков одного из регионов по величине уставного капитала на 01.01.08 г.

№ группы	Группы банков по величине уставного капитала, тыс. руб.	Число банков, ед.	Капитал, тыс. руб.		Работающие активы, тыс. руб.	
			Всего	в среднем на один банк	Всего	в среднем на один банк
1	2100 - 7350	18	342889	19049	504898	28050
2	7350 - 12600	6	204694	34116	343932	57322
3	12600 - 17850	3	130680	43560	174059	58020
4	17850 - 23100	3	128573	42858	217842	72614
Итого		30	806836	-	1240731	-
В среднем на один банк			-	26895	-	41358

Величины капитала и работающих активов прямо взаимосвязаны, и чем крупнее банк, тем эффективнее управление работающими активами.

Мы рассмотрели примеры группировок по одному признаку. Однако в ряде случаев для решения поставленных задач такая группировка является недостаточной. В этих случаях переходят к группировке исследуемой совокупности по двум и более существенным признакам во взаимосвязи (сложной группировке).

Произведем группировку данных коммерческих банков по двум признакам: величине капитала и работающих активов. Каждую группу и подгруппу охарактеризуем следующими показателями: число коммерческих банков, капитал, работающие активы (табл. 1.5).

Таблица 1.5 – Группировка коммерческих банков одного из регионов по величине капитала и работающих активов на 01.01.08 г.

№ гр.	Группы банков по велич. капитала, тыс. руб.	В том числе подгруппы по велич. работающих активов, тыс. руб.	Число банков, ед.	Капитал, тыс. руб.	Работающие активы, тыс. руб.
1	6753 - 29527	2556 - 65144	20	313649	420196
		65144 - 127732	1	22969	70229
Итого по группе			21	336618	490425

Продолжение табл. 1.5

2	29527 - 52301	2556 - 65144	1	35662	30005
		65144 - 127732	3	127523	305994
Итого по группе			4	163185	335999
3	52301 - 81828	2556 - 65144	2	115104	98022
		65144 - 127732	3	191929	316285
Итого по группе			5	307033	414307
4	Итого по подгруппам	2556 - 65144	23	464415	548223
		65144 - 127732	7	342421	692508
Всего			30	806836	1240731

Задача 1.2. Объем инвестиций в основной капитал характеризуется в России следующими данными (в фактически действовавших ценах, млрд. руб.): 1998 г. – 402,4; 1999 г. – 565,6; в том числе в отрасли: а) производящие товары – 1998 г. – 163,8; 1999 г. – 269,4; б) оказывающие рыночные и нерыночные услуги – 1998 г. – 238,6; 1999 г. – 296,2. Представить приведенные данные в виде статистической таблицы. Сформулировать выводы, охарактеризовав произошедшие изменения в объеме и составе инвестиций.

Решение

Таблица 1.6 – Инвестиции в России в основной капитал (в фактически действовавших ценах, млрд. руб.)

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

Изменение объема инвестиций в 1999 г. По сравнению с 1998 г. Определяется следующим образом, %:

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

$\frac{269,4}{163,8} \cdot 100 = 164,5$; инвестиции в отрасли, оказывающей рыночные и нерыночные услуги:

$\frac{296,2}{238,6} \cdot 100 = 124,1$. Структура инвестиций определялась так, %:

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

$$1999г.: \frac{264,4}{565,6} \cdot 100 = 47,6; \quad \frac{296,2}{565,6} \cdot 100 = 52,4.$$

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

Задача 1.3. Имеются следующие данные по заработной плате водителей за сентябрь.

Таблица 1.7

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

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

Решение. Для решения задачи необходимо произвести группировку водителей по двум признакам-факторам: сначала – на группы по квалификации, затем внутри каждой группы на подгруппы по проценту выполнения сменного задания. По проценту выполнения сменного задания принимаются две подгруппы: водители, выполняющие норму от 100% до 110%; водителей, выполняющие норму на 110% и выше.

Результаты группировки представлены во вспомогательной таблице 1.8.

На основе вспомогательной таблицы по каждой подгруппе определяют численность признака (общую сумму заработной платы), результаты оформляют в виде комбинационной таблицы 1.9.

Таблица 1.8

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

Таблица 1.9

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

Вывод: из данных табл. 1.9 следует, что с ростом квалификации водителей и процента выполнения сменного задания увеличивается заработная плата. Так, заработная плата водителей I класса, выполняющих норму выработки на 110% и выше, на 14,8% превышает заработную плату водителей II класса, выполняющих норму от 100-110%.

1.2. Задания для самостоятельной работы

1.2.1. Выполнение установленных норм выработки рабочими бригады строителей-монтажников характеризуется следующими данными за июнь:

№	Фамилия, И., О.	Выполнение нормы, %	№	Фамилия, И., О.	Выполнение нормы, %
1	Слицын М.И.	101,3	13	Щербakov А.В.	110,3
2	Кузьмин К.С.	124,6	14	Башлыков Б.Р.	102,6
3	Федоренко А.П.	103,5	15	Тенисон В.Ф.	100,1
4	Симонович В.В.	130,1	16	Устинов Г.Д.	153,0
5	Потапенко Н.М.	113,9	17	Ильин М.С.	109,9
6	Глебов Л.Н.	89,6	18	Дронов Б.Г.	117,6
7	Фролов И.И.	104,7	19	Киселев Ф.А.	102,4
8	Баринov Б.Я.	120,0	20	Вельдин М.С.	105,6
9	Кузнецов А.И.	119,5	21	Сомов Е.Н.	99,7
10	Петров С.С.	108,6	22	Федин З.П.	107,4
11	Синицын Н.К.	97,8	23	Мохов Р.И.	106,1
12	Королев А.М.	107,2	24	Стокос К.М.	112,2

Постройте ряд распределения рабочих-строителей по степени выполнения ими установленных норм выработки, приняв следующие интервалы: 1) до 100%; 2) от 100 до 110%; 3) от 110 до 120%; 4) от 120% и выше. Решение представьте в форме статистической таблицы.

1.2.2. Известны следующие данные по основным показателям деятельности крупнейших банков одной из областей России (данные условные, в тыс. руб.):

№ п/п	Сумма активов	Собственный капитал	Привлеченные ресурсы	Балансовая прибыль	Объем вложений в государственные ценные бумаги	Судная задолженность
1	645,6	12,0	27,1	8,1	3,5	30,8
2	636,9	70,4	56,3	9,5	12,6	25,7
3	629,0	41,0	95,7	38,4	13,3	26,4
4	619,6	120,8	44,8	38,4	4,4	25,3
5	616,4	49,4	108,7	13,4	15,0	20,9
6	614,4	50,3	108,1	30,1	19,1	47,3
7	608,6	70,0	76,1	37,8	19,2	43,7
8	601,1	52,4	26,3	41,1	3,7	29,1
9	600,2	42,0	46,0	9,3	5,2	56,1
10	600,0	27,3	24,4	39,3	13,1	24,9
11	592,9	72,0	65,5	8,6	16,7	39,6
12	591,7	22,4	76,0	40,5	7,5	59,6
13	585,5	39,3	106,9	45,3	6,7	44,9
14	578,6	70,0	89,5	8,4	11,2	32,2
15	577,5	22,9	84,0	12,8	19,3	45,1
16	553,7	119,3	89,4	44,7	15,4	24,5
17	543,6	49,6	93,8	8,8	5,7	31,1
18	542,0	88,6	26,7	32,2	7,8	37,1
19	517,0	43,7	108,1	20,3	8,3	23,1
20	516,7	90,5	25,2	12,2	9,7	15,8

Постройте группировку коммерческих банков по величине собственного капитала, выделив не более пяти групп с равными интервалами. Рассчитайте по каждой группе сумму активов, собственный капитал, привлеченные ресурсы, балансовую прибыль. Результаты группировки представьте в табличной форме и сформулируйте выводы.

1.2.3. Постройте структурную группировку банков по величине балансовой прибыли, выделив четыре группы банков с открытыми интервалами для характеристики структуры совокупности коммерческих банков, перечисленных в задаче 1.3. 2.

1.2.4. Постройте аналитическую группировку коммерческих банков, перечисленных в задаче 1.3.2, по величине балансовой прибыли, выделив четыре-пять групп. Рассчитайте по каждой группе два-три показателя, взаимосвязанных с балансовой прибылью. Результаты группировки изложите в табличной форме и сделайте выводы о взаимосвязи показателей.

1.2.5. Используя данные задачи 1.3.2, постройте группировку коммерческих банков в целях выявления взаимосвязи между показателями привлеченных ресурсов, объемом вложений в государственные ценные бумаги и ссудной задолженностью от результатов деятельности банков (показатель, выражающий результаты деятельности банков, определите самостоятельно).

1.2.6. По данным задачи 1.3.2 построите все возможные структурные и аналитические группировки коммерческих банков.

1.2.7. По данным задачи 1.3.2 построите группировку коммерческих банков по двум признакам: по величине балансовой прибыли и сумме активов. По каждой группе и подгруппе определите число банков, величину балансовой прибыли и сумму активов и другие два-три показателя, взаимосвязанные с группировочными. Результаты группировки оформите в виде таблицы и сформулируйте выводы.

1.3. Задания для индивидуальной работы

1.3.1. В результате 10-процентного обследования рабочих Минского тракторного завода по состоянию на 01.10.2008 г. получены следующие данные:

Номер п/п	Квалификационный разряд рабочего	Производственный стаж, лет	Зарплата за сентябрь, тыс. руб.	Процент выполнения нормы сменного задания, %
1	2	3	4	5
Цех № 1				
1	4	5	720	146
2	4	7	745	141
3	2	1	280	100
4	3	4	400	112
5	4	6	965	150
6	2	0	289	96
7	4	6	669	121
8	5	8	840	149
9	4	5	465	151
10	3	4	520	132
11	4	8	700	148
12	4	6	647	134
13	6	10	1020	153
14	6	9	983	151
15	3	4	510	118
16	2	2	316	102
17	3	2	439	109
18	4	1	742	146
19	4	6	752	131
20	3	4	465	108
21	3	4	438	106
22	4	3	673	107
23	4	7	650	116
24	5	12	781	136
25	5	10	800	126

26	4	8	750	103
27	3	4	469	100
28	5	7	843	98
29	6	18	900	101
30	5	12	863	105
31	4	6	720	111
32	4	9	657	110
33	3	4	537	108
34	5	12	890	102
35	6	20	1120	110
36	4	7	780	101
37	5	14	780	100
38	3	6	436	95
39	5	8	840	103
40	3	6	585	113

Lex № 2

1	2	3	4	5
1	3	2	540	104
2	4	9	723	108
3	5	13	780	112
4	6	15	985	110
5	4	6	680	111
6	4	8	732	103
7	5	9	841	106
8	3	4	485	106
9	3	2	468	101
10	4	2	675	102
11	4	8	463	94
12	6	14	958	100
13	5	10	852	100
14	6	18	1230	107
15	4	7	783	122
16	5	6	879	121
17	3	1	518	130
18	4	4	548	109
19	4	8	610	112
20	3	6	489	120
21	5	9	778	108
22	5	11	859	105
23	6	16	1100	128
24	3	7	469	107
25	2	3	286	98
26	2	4	300	104
27	4	3	670	96
28	5	8	865	114
29	6	14	980	109
30	4	7	825	98
31	5	12	1000	100
32	4	9	810	96
33	6	15	1310	151
34	5	16	1030	149
35	3	4	385	107
36	3	8	450	95
37	5	9	842	103
38	4	8	680	106
39	3	5	459	96
40	5	8	830	101
41	2	0	250	103
42	3	3	530	125
43	3	1	500	131

По данным задачи выявить зависимость заработной платы от:

1) производственного стажа (нечетные варианты), процента выполнения нормы выработки (четные варианты);

2) уровня квалификации и процента выполнения нормы выработки (нечетные варианты), производственного стажа и уровня квалификации (четные варианты).

Интервалы группировки разработать самостоятельно. На основе выполненной группировки построить комбинационную таблицу. Сформулировать вывод.

3) построить графическое изображение рядов распределения рабочих цехов и завода по признакам из п. 1).

№ вар-та	№ г/п рабочих цеха	№ г/п рабочих цеха	№ вар-та	№ г/п рабочих цеха	№ г/п рабочих цеха
1	1-10	1-13	16	2-12	1-14
2	3-13	3-13	17	4-18	4-16
3	5-17	1-10	18	14-24	14-28
4	7-17	3-16	19	16-28	16-30
5	9-19	9-29	20	18-30	18-29
6	10-20	10-25	21	20-31	20-33
7	12-25	12-20	22	23-35	23-37
8	14-27	14-23	23	25-33	25-38
9	16-26	16-29	24	27-38	27-36
10	18-30	18-28	25	29-40	29-42
11	20-34	20-30	26	21-33	21-38
12	22-35	22-30	27	27-35	27-40
13	24-34	24-38	28	29-38	29-41
14	26-36	26-36	29	28-39	28-43
15	25-35	25-35	30	30-40	30-43

Глава 2. СТАТИСТИЧЕСКИЕ ВЕЛИЧИНЫ

2.1. Понятие абсолютной и относительной величины

Статистический показатель представляет собой количественную характеристику социально-экономических процессов в условиях качественной определенности.

Все статистические показатели классифицируются на абсолютные, относительные и средние.

Абсолютным показателем является такой, который отражает либо суммарное число единиц, либо суммарное свойство объекта. Абсолютные показатели выражаются именованными величинами в натуральных единицах измерения (тоннах, штуках, часах, амперах и т.д.), условных единицах (нормо-сменах, киломерах пряжи и т.п.) или в стоимостных единицах (рублях, долларах, евро). Они характеризуют сумму значений первичных признаков объекта.

Относительный показатель представляет собой результат деления одного абсолютного показания на другой и выражает соотношение между количественными характеристиками социально-экономических процессов и явлений. Относительные показатели – величины производные, определяемые в форме коэффициентов, процентов, промилле и т.д.

Все используемые на практике относительные показатели можно подразделить на следующие виды: показатели динамики, плана, выполнения плана, структуры, координации, интенсивности, сравнения.

I. Относительный показатель динамики (ОПД) характеризует изменение уровня развития какого-либо явления во времени, т.е.

$$ОПД = \frac{\text{текущий уровень}}{\text{предшествующий уровень (базисный уровень)}}$$

Задача 1. Имеются данные о производстве бумаги (табл. 2.1.1):

Таблица 2.1.1

Год	Производство бумаги, тыс.т.	Расчетные графы	
		Переменная база сравнения	Постоянная база сравнения
2003	1600	-	-
2004	1282	$\frac{1282}{1600} \cdot 100\% = 80,1\%$	$\frac{1282}{1600} \cdot 100\% = 80,1\%$
2005	1015	$\frac{1015}{1282} \cdot 100\% = 79,2\%$	$\frac{1015}{1600} \cdot 100\% = 63,1\%$
2006	1171	$\frac{1171}{1015} \cdot 100\% = 115,4\%$	$\frac{1171}{1600} \cdot 100\% = 73,2\%$

Вычислите относительные показатели динамики с переменной базой сравнения. Проверьте их взаимосвязь.

Решение. Вычисление ОПД приведено в расчетных графах таблицы 2.1.1. Относительные показатели динамики с переменной и постоянной базой сравнения взаимосвязаны между собой: произведение всех относительных показателей с переменной базой равно относительному показателю с постоянной базой за исследуемый период.

В нашем случае: $0,801 \cdot 0,792 \cdot 1,154 = 0,732$ или 73,2%.

II. Относительный показатель планового задания (ОПП) рассчитывается как отношение уровня, запланированного на предстоящий период, к уровню, фактически сложившемуся в предшествующем периоде

$$ОПП = \frac{\text{уровень, планируемый на } (i+1)\text{-ый период}}{\text{уровень, достигнутый в } i\text{-ом периоде}}$$

Относительный показатель выполнения плана (ОПВП) рассчитывается как отношение фактически достигнутого в данном периоде уровня к запланированному

$$ОПВП = \frac{\text{уровень, достигнутый в } (i+1)\text{-ом периоде}}{\text{уровень, планированный на } (i+1)\text{-ый период}}$$

Между относительными показателями плана, выполнением плана и динамики существует следующая взаимосвязь: $ОПП \cdot ОПВП = ОПД$.

Задача 2а. Объем продаж компании Samsung в странах СНГ в первом полугодии 1996г. составил 250 млн. долларов. В целом же за год компания планировала реализовать товаров на 600 млн. долларов. Вычислить относительный показатель плана на второе полугодие.

Решение. Относительный показатель плана на второе полугодие есть отношение планируемой величины на второе полугодие к фактически достигнутой величине в первом полугодии: $\frac{600 - 250}{250} \cdot 100\% = 140\%$.

Задача 2б. Предприятие планировало увеличить выпуск продукции в 1997 г. по сравнению с 1996 г. на 18%. Фактически же объем продукции составил 112,3% от прошлогогоднего уровня. Определить относительный показатель выполнения плана.

Решение. Относительный показатель выполнения плана в 1997 г. есть отношение фактического выпуска продукции к планируемому в 1997 г.: $\frac{1,123}{1+0,18} \cdot 100\% = \frac{1,123}{1,18} \cdot 100\% = 95,2\%$, т.е. плановое задание было невыполнено на 4,8%.

III. Относительный показатель структуры (ОПС) характеризует доли, удельные веса составных элементов в общем итоге.

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

Задача 3. Имеются данные о внешнеторговом обороте России со странами дальнего зарубежья и СНГ

Таблица 2.1.2

	IV квартал 1995 г.		I квартал 1996 г.	
	млн. долларов, y_i	$d_i, \%$	млн. долларов, y_i	$d_i, \%$
Экспорт	22761	55,5	20972	60,0
Импорт	18274	44,5	13954	40,0
Итого	41035	100,0	34926	100,0

Вычислить относительные показатели структуры.

Решение. Расчет ОПС представлен в табл. 2.1.2 и произведен по формуле:

$$d_i (ОПС) = \frac{y_i}{\sum y_i} \cdot 100\%; \quad d_i = \frac{22761}{41035} \cdot 100\% = 55,5\% \text{ и т.д.}$$

Т.о., в 1995 г. доля экспорта составляла 55,5%, доля импорта – 44,5%, в 1996 г. – 60% и 40% соответственно.

IV. Относительный показатель координации (ОПК) – характеризует отношение частей данной совокупности к одной из них, принятой за базу сравнения. ОПК показывает, во сколько раз одна часть совокупности больше другой, либо сколько единиц одной части приходится на 1, 10, 100, 1000, ... единиц другой части.

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

В качестве базы сравнения выбирается та часть, которая имеет наибольший удельный вес или является приоритетной с экономической, социальной или какой-либо другой точки зрения.

Задача 4. Производство автомобилей в РФ в январе-мае 1996г. характеризуется следующими данными (тыс. шт.)

Таблица 2.1.3

	январь	февраль	март	апрель	май
Всего	65,0	83,2	79,3	89,9	76,6
В том числе:					
грузовые	11,0	11,5	12,0	11,0	9,3
легковые	54,0	71,5	67,3	78,9	67,3

Вычислить относительные показатели координации.

Решение. В январе – $\frac{54,0}{11,0} = 4,9$, т.е. на один произведенный грузовой автомобиль приходится 4,9 легковых или на 10 грузовых приходится 49 легковых.

В феврале $-\frac{71,7}{11,5} = 6,2$, т.е. на 10 произведенных грузовиков приходится 62 легковых автомобиля.

В марте $-\frac{67,3}{12,0} = 5,6$, т.е. на 10 произведенных грузовиков приходится 56 легковых автомобилей и т. д.

В апреле $-\frac{78,9}{11,0} = 7,2$. В мае $-\frac{67,3}{9,3} = 7,2$.

V. Относительный показатель сравнения (ОПС) представляет собой соотношение одного и того же абсолютного показателя, характеризующего разные объекты (предприятия, фирмы, районы, области, страны и т.д.).

$$ОПС = \frac{\text{показатель, характеризующий объект } A}{\text{показатель, характеризующий объект } B}$$

Задача 5. В апреле 1996 г. прожиточный минимум в РФ для трудоспособного населения составил 419,0 тыс. руб. в месяц на человека, для пенсионеров – 262,5 тыс. руб., для детей – 376,1 тыс. руб. Сделайте выводы о соотношении этих величин, используя ОПС.

Решение. Прожиточный минимум пенсионеров в 0,626 раз меньше, чем трудоспособного населения $\left(\frac{262,5}{419,0} = 0,626\right)$, или на 37,4% меньше, а детей – в 0,898 раз меньше, чем трудоспособного населения $\left(\frac{376,1}{419,0} = 0,898\right)$ или на 10,2%.

VI. Относительный показатель интенсивности (ОПИ) характеризует степень распределения или развития данного явления в той или иной среде

$$ОПИ = \frac{\text{показатель, характеризующий явление } A}{\text{показатель, характеризующий среду распределения явления } A}$$

ОПИ используется при изучении демографических процессов (показатели рождаемости, смертности). Относительными величинами интенсивности выступают показатели выработки продукции в единицу рабочего времени, затрат на единицу продукции, трудоемкости, эффективности использования производственных фондов.

Задача 6. Необходимо рассчитать и сравнить трудоемкость продукции в организациях по производству и переработке картофеля, если известно, что объем производства и переработки картофельного сырья составил 1000т; затраты труда соответственно 20 тыс. и 3 тыс. чел.- час.

Решение. Трудоемкость производства картофеля:

$$t_1 = \frac{y_1}{x} = \frac{\text{затраты труда}}{\text{объем производства}} = \frac{20000}{1000} = 20 \left(\frac{\text{чел.} - \text{час}}{\text{т}} \right).$$

Трудоемкость переработки картофеля:

$$t_2 = \frac{y_2}{x} = \frac{3000}{1000} = 3 \left(\frac{\text{чел.} - \text{час}}{\text{т}} \right).$$

Т.о., трудоемкость производства картофеля выше, чем трудоемкость его переработки.

2.2. Средняя величина. Виды средних величин.

Средняя величина – это обобщающий показатель, характеризующий типичный уровень явления в конкретных условиях места и времени.

Средние величины делятся на два больших класса:

- 1) Степенные средние $\bar{x} = \sqrt[m]{\frac{\sum x_i^m \cdot f_i}{\sum f_i}}$ - общее выражение.
- 2) Структурные средние (мода, медиана, квартили).

Виды степенных средних

Вид степенной средней	Показатель степени (m)	Формула расчета	
		простая	взвешенная
Гармоническая	-1	$\bar{x} = \frac{n}{\sum \frac{1}{x}}$	$\bar{x} = \frac{\sum w}{\sum \frac{w}{x}}$, где $w = xf$
Геометрическая	$\rightarrow 0$	$\bar{x} = \sqrt[m]{\prod x} = \sqrt[m]{x_1 \cdot x_2 \cdot \dots \cdot x_n}$	$\bar{x} = \sqrt[m]{\prod x^f} = \sqrt[m]{x_1^{f_1} \cdot x_2^{f_2} \cdot \dots \cdot x_n^{f_n}}$
Арифметическая	1	$\bar{x} = \frac{\sum x}{n}$	$\bar{x} = \frac{\sum xf}{\sum f}$
Квадратическая	2	$\bar{x} = \sqrt{\frac{\sum x^2}{n}}$	$\bar{x} = \sqrt{\frac{\sum x^2 f}{\sum f}}$
Кубическая	3	$\bar{x} = \sqrt[3]{\frac{\sum x^3}{n}}$	$\bar{x} = \sqrt[3]{\frac{\sum x^3 f}{\sum f}}$

Если рассчитаны все виды средних для одних и тех же исходных данных, то значения их окажутся неодинаковыми. С увеличением показателя степени m увеличивается и соответствующая средняя величина: $\bar{x}_{\text{гарм.}} \leq \bar{x}_{\text{геом.}} \leq \bar{x}_{\text{арифм.}} \leq \bar{x}_{\text{квдр.}} \leq \bar{x}_{\text{куб.}}$

2.2.1. Решение типовых задач

Задача 1. Имеются следующие данные о работе малых предприятий за текущий период.

Таблица 2.2.1

Предприятие	Фактический объем реализации, тыс. руб.	Средний объем реализации на одного работника, тыс. руб.	Прибыль к объему реализации, %	Процент совместителей в общей численности работников
	x	y	z	d
1	19000	3800	19	66
2	16000	4000	20	70
3	20000	5000	26	60
4	19200	3200	20	75

Определить по малым предприятиям района средние значения:

- 1) реализованной продукции на одно предприятие;
- 2) производительности труда;
- 3) рентабельности продукции;
- 4) доли совместителей в общей численности работников. Указать вид рассчитанных средних величин и сделать выводы.

Решение. Введем обозначения:

X – фактический объем реализации;

Y – средний объем реализации на одного работника;

Z – прибыль к объему реализации;

d – процент совместителей в общей численности работников.

Выбор вида средней величины необходимо начинать с построения логической формулы, исходя из качественного содержания содержания усредняемого показателя.

Все расчеты будем оформлять в таблице:

Таблица 2.2.2

№ предприятия	Фактический объем реализации, тыс. руб., (x_i)	Средний объем реализации на одного работника, тыс. руб., (y_i)	Кол-во работающих, чел., (x_i) y_i	Прибыль к объему реализации, % (z_i)	Прибыль, тыс.руб., ($\frac{x_i z_i}{100}$)	Совместители, чел., ($\frac{x_i \cdot d_i}{y_i \cdot 100}$)	Процент совместителей в общей численности работников, (d_i)
1	19000	3800	5	19	19000 · 0,19 = 3610	3,3	66
2	16000	4000	4	20	16000 · 0,20 = 3200	2,8	70
3	20000	5000	4	26	20000 · 0,26 = 5200	2,4	60
4	19200	3200	6	20	19200 · 0,20 = 3840	4,5	75
Σ	74200		19		15850	13	

$$1) \text{ Средний объем реализованной продукции} = \frac{\text{Совокупный объем реализованной продукции всеми предприятиями (тыс.руб.)}}{\text{Количество предприятий}}$$

Получаем, таким образом, среднюю арифметическую простую

$$x = \frac{\sum x}{n} = \frac{74200}{4} = 18550 \text{ (тыс. руб.)}$$

$$2) \text{ Средняя производительность труда} = \frac{\text{Совокупный объем реализованной продукции (тыс.руб.)}}{\text{Общее количество работающих}}$$

Количество работающих найдем, зная что:

$$\text{Количество работающих} = \frac{\text{Объем реализованной продукции (тыс.руб.)}}{\text{Средний объем реализованной продукции на одного работника (тыс.руб.)}}$$

Получаем среднюю производительность труда по четырем предприятиям как среднюю гармоническую взвешенную:

$$y = \frac{\sum x}{\sum \frac{x}{y}} = \frac{74200}{19} = 3905 \text{ (тыс. руб.)}, \text{ т.е. по четырем предприятиям объем реализации на одного работника предприятия составляет в среднем } 3905 \text{ тыс. руб.}$$

$$3) \text{ Средняя рентабельность продукции} = \frac{\text{Прибыль (тыс.руб.)}}{\text{Объем реализованной продукции (тыс.руб.)}}$$

Среднюю рентабельность продукции в процентах найдем как среднюю арифметическую взвешенную:

$$\bar{z} = \frac{\sum xz}{\sum x} \cdot 100\% = \frac{15850}{74200} \cdot 100 = 21,4\%, \text{ т.е. в среднем прибыль составляет } 21,4\% \text{ к объему реализации.}$$

$$4) \text{ Доля совместителей в общей численности работников} = \frac{\text{Общая численность совместителей}}{\text{Общая численность работников}}$$

$$\bar{d} = \frac{\sum x \cdot d_i}{\sum x} \cdot 100\% = \frac{\sum d_i \cdot x_i}{\sum x_i} \text{ -- средняя арифметическая взвешенная.}$$

$$\bar{d} = \frac{13}{19} \cdot 100\% = 68,4\% \text{ -- средний процент совместителей в общей численности работающих.}$$

Т.е. по четырем предприятиям совместители составляют в среднем 68,4% от общей численности работающих.

Задача 2. Рассчитать средний диаметр сосновых бревен по следующим данным:

Таблица 2.2.3

Диаметр бревен, см, $a_{i-1} - a_i$	Число бревен, шт, f_i	Середина интервала, см, x_i	x_i^2	$x_i^2 f_i$
25-35	10	30	900	9000
35-45	20	40	1600	32000
45-55	30	50	2500	75000
55-65	10	60	3600	36000
-	70	-	-	152000

Решение

Все промежуточные расчеты оформлены в таблице 2.2.3.

Средневзвешенный диаметр сосновых бревен

$$\bar{x} = \sqrt{\frac{\sum x_i^2 f_i}{\sum f_i}} = \sqrt{\frac{152000}{70}} = 46,6 \text{ (см)}$$

Главная сфера применения средней квадратической величины -- нахождение среднего квадратического отклонения.

Задача 3. Во сколько раз в среднем возросло производство сахарной свеклы в сельскохозяйственной организации за четыре года, если известно, что целные коэффициенты роста по годам составляли соответственно 1, 0,9, 1,3, 1,5 раза.

Решение. Так как целные коэффициенты взаимозависимы и связаны знаком произведения, то $x = \sqrt[4]{x_1 \cdot x_2 \cdot x_3 \cdot x_4} = \sqrt[4]{1 \cdot 0,9 \cdot 1,3 \cdot 1,5} = \sqrt[4]{1,755} = 1,151$.

Производство сахарной свеклы за каждый год в среднем возросло в 1,151 раза.

Задача 4. Динамика валового производства картофеля за 20-летний период приведена в таблице 2.2.4.

Таблица 2.2.4

Темп роста производства картофеля, %		Число лет в каждом периоде
Интервал	Середина интервала, x_i	
90 - 100	95	3
100 - 110	105	6
110 - 120	115	6
120 - 130	125	5
Итого	—	20

Рассчитать среднегодовой темп роста валового производства картофеля в районе.

Решение: Воспользуемся формулой средней геометрической взвешенной

$$\bar{x} = \sqrt[20]{\prod (x)^f} = \sqrt[20]{0,95^3 \cdot 1,05^6 \cdot 1,15^6 \cdot 1,25^5} = 1,002 \text{ (100,2\%)}$$

Таким образом, за 20-летний период производство картофеля развивалось со среднегодовым темпом роста 100,2%.

2.3. Структурные средние

Структурные средние применяются для получения внутреннего строения рядов распределений признака.

В качестве структурных средних чаще всего используются показатели моды, медианы и квартилей.

Мода – значение признака, наиболее часто встречающегося в исследуемой совокупности.

Медиана – значение признака, приходящегося на середине упорядоченной совокупности.

Квартили – значения признака, делящие ранжированную совокупность на четыре равновеликие части.

Мода для интервального ряда с равными интервалами рассчитывается по формуле:

$$M_o = x_{M_o} + h_{M_o} \frac{f_{M_o} - f_{M_{o-1}}}{(f_{M_o} - f_{M_{o-1}}) + (f_{M_o} - f_{M_{o+1}})},$$

где x_{M_o} – начало модального интервала, h_{M_o} – длина модального интервала,

f_{M_o} – частота модального интервала, $f_{M_{o-1}}$ – частота домодального интервала,

$f_{M_{o+1}}$ – частота послемодального интервала.

Медиана интервального распределения:

$$M_e = x_{M_e} + h_{M_e} \frac{\frac{\sum f_i}{2} - S_{M_e-1}}{f_{M_e}},$$

где x_{M_e} – начало медианного интервала,

h_{M_e} – длина медианного интервала,

S_{M_e-1} – накопленная частота домедианного интервала,

f_{M_e} – частота медианного интервала.

Квартили по интервальному ряду:

$$Q_1 = x_{Q_1} + h_{Q_1} \frac{\frac{\sum f_i}{4} - S_{Q_1-1}}{f_{Q_1}},$$

где x_{Q_1} – начало первого квартильного интервала,

h_{Q_1} – длина первого квартильного интервала,

S_{Q_1-1} – накопленная частота до первого квартильного интервала,

f_{Q_1} – частота первого квартильного интервала;

$Q_2 = Me$;

$$Q_3 = x_{Q_3} + h_{Q_3} \frac{\frac{3}{4} \sum f_i - S_{Q_3-1}}{f_{Q_3}},$$

где x_{Q_3} – начало третьего квартильного интервала, h_{Q_3} – его длина,

S_{Q_3-1} – накопленная частота до третьего квартильного интервала,

f_{Q_3} – частота третьего квартильного интервала.

2.3.1. Решение типовых задач

Задача 1. Известно распределение вкладчиков районного отделения сберегательного банка по размеру вкладов (табл. 4.2.5):

Таблица 4.2.5

Группы вкладчиков по размеру вклада, тыс. руб	Число вкладчиков, тыс. чел	Расчетная графа
		Накопленная частота
$x_{i-1} - x_i$	f_i	S_i
До 500	168	168
500 – 1000	29	197
1000 – 1500	13	210
1500 – 2000	3	213
2000 – 2500	1	214
2500 и более	1,109	214,109
Итого	214,109	

Расчитайте структурные средние – моду, медиану и квартили.

1) модальный интервал определяем по наибольшей частоте $f_i = 168$.

$$M_0 = 0 + 500 \frac{168 - 0}{(168 - 0) + (168 - 29)} = \frac{500 \cdot 168}{307} = 273,6 \text{ (тыс. руб.)},$$

следовательно, наиболее часто встречающийся размер вклада 273,6 тыс. руб.;

2) медианный интервал определим с помощью числа $\frac{214,109 + 1}{2} = 107,55$, следовательно, это первый интервал (0,500).

$$\text{Тогда, } M_e = 0 + 500 \frac{107,05 - 0}{168} = 318,6 \text{ (тыс. руб.)}.$$

Таким образом, половина вкладчиков имеют вклады менее 318,6 тыс. руб., остальные – более 318,6 тыс. руб.;

3) первый квартильный интервал тот же (0,500), т.к. $\frac{\sum f_i}{4} = 53,5 \in (0; 500)$, т.о.

$$Q_1 = 0 + 500 \frac{53,5 - 0}{168} = 159,3 \text{ (тыс. руб.)};$$

4) третий квартильный интервал определяется числом $\frac{3}{4} \sum f_i = 160,6 \in (0; 500)$,

$$Q_3 = 0 + 500 \frac{160,6 - 0}{168} = 477,9 \text{ (тыс. руб.)}. \text{ Таким образом, 25\% вкладчиков име}$$

ют вклады до 159,3 тыс. руб., 25% вкладчиков имеют вклады больше 477,9 тыс. руб. и 50% вкладчиков имеют вклады от 159,3 тыс. руб. до 477,9 тыс. руб.

2.3.2. Задания для самостоятельной работы

1. Крестьянские хозяйства подразделяются по размерам земельных угодий следующим образом:

Размеры земельных угодий, га	Число хозяйств, ед.	Размеры земельных угодий, га	Число хозяйств, ед.
До 4	34	7 – 8	1837
4 – 5	52	8 – 9	663
5 – 6	435	9 – 10	753
6 – 7	841	10 и более	133

Расчитайте: 1) средние размеры земельных угодий;

2) структурные средние – моду, медиану, квартили.

2. В трех партиях продукции, представленных на контроль качества, было обнаружено:
- а) первая партия – 1000 изделий, из них 920 годных и 80 бракованных;
 - б) вторая партия – 800 изделий, из них 730 годных, 70 бракованных единиц продукции;
 - в) третья партия – 900 изделий, из них 840 годных и 60 бракованных.

Определить в целом по трем партиям средний процент годной продукции и средний процент брака.

3. В результате проверки двух партий сыра перед отправкой его потребителям установлено, что в первой партии сыра высшего сорта было 3832 кг, что составляет 70,1% общего веса сыра этой партии; во второй партии сыра высшего сорта было 8520 кг, что составляет 75,4% общего веса сыра этой партии. Определите процент сыра высшего сорта в среднем по первой и второй партиям.

4. Четыре швеи – надомницы заняты пошивом головных уборов одной модели. Первая швея тратит на изготовление одного головного убора 30 мин, вторая – 40, третья – 50 и четвертая – 60 мин. Необходимо определить средние затраты времени на пошив 1-ого головного убора при условии, что каждая надомница работает по 10 часов в день.

5. Приведены данные об урожайности картофеля в личных подсобных хозяйствах населения:

Интервалы по урожайности, т/га	Число хозяйств, ед.
10 – 15	10
15 – 20	30
20 – 25	50
25 – 30	80
30 – 35	20
35 – 40	10

Рассчитайте структурные средние – моду, медиану, квартили.

6. Результаты торгов на российских биржах 1 июля 2007 г. характеризуются следующими данными:

Биржа	Курс \$, руб.	Объем продажи в млн. дол.
Московская	25,15	8,79
Петербургская	25,38	7,84
Самарская	25,26	1,88
Уральская	25,29	6,21
Азиатская	25,00	2,79
Ростовская	25,37	0,55
Нижегородская	25,25	0,03
Сибирская	25,08	2,48

Рассчитать средний курс доллара.

7. По трем районам города имеются следующие данные:

Район	Число отделений сбербанка	Среднее число вкладов в отделении	Средний размер вклада, тыс. руб.
1	4	1376	275
2	9	1359	293
3	5	1315	268

Определите средний размер вклада в сбербанке в целом по городу.

2.3.3. Задания для индивидуальной работы

Варианты 1,2,3

Имеются следующие данные о работе ТЭЦ области за месяц (см. табл.). Определить по совокупности ТЭЦ среднее значение: 1) количества электроэнергии, выработанной одной ТЭЦ; 2) процента выполнения плана по объему выработки одной ТЭЦ; 3) себестоимости 1 кВт.ч электроэнергии; 4) расхода условного топлива на 1кВт.ч электроэнергии.

Укажите виды рассчитанных средних величин. Сделайте выводы.

Вариант	ТЭЦ	Количество выработанной энергии, млн. кВт.ч	Процент выполнения плана по объему электроэнергии, %	Себестоимость 1кВт.ч, руб.	Расход условного топлива на 1кВт.ч, электроэнергии, грамм
1	1	5000	101,0	145,0	458
	2	1200	102,4	144,0	403
	3	3800	99,5	144,5	423
	4	4000	98,0	145,8	444
2	1	4800	101,8	140,5	460
	2	1300	102,1	142,0	420
	3	4100	100,5	143,0	425
	4	4800	99,3	144,0	450
3	1	4900	102,3	138,0	435
	2	1500	100,2	140,0	450
	3	3900	99,5	139,5	456
	4	4100	101,3	138,6	428

Варианты 4,5,6

Имеются следующие данные о работе коммерческих киосков за отчетный период. Определить по совокупности коммерческих киосков среднее значение:

- 1) товарооборота на один киоск;
- 2) товарооборота на одного работающего;
- 3) рентабельности;
- 4) заработной платы.

Укажите виды рассчитанных средних величин. Сделайте выводы.

Вариант	Киоск	Товарооборот		Рентабельность, %	Средняя заработная плата на одного работника, тыс. руб.
		Всего, млн. руб.	В расчете на одного работника, млн. руб.		
4	1	21	7,0	20	550
	2	15	5,5	22	490
	3	24	7,6	24	520
	4	20	4,8	21	495
5	1	31	8,0	21	601
	2	25	6,5	22	600
	3	34	7,0	23	590
	4	32	5,2	24	570
6	1	30	7,2	20	520
	2	32	7,0	19	580
	3	28	7,1	21	600
	4	29	6,8	22	610

Варианты 7,8,9

Имеются следующие данные по вузам города. Определить по всем вузам города среднее значение: 1) доли студентов заочной формы обучения в общей численности студентов города; 2) численности студентов, приходящихся на одного преподавателя; 3) доли выпускников очной формы обучения с отличием.

Укажите виды рассчитанных средних величин.

Вариант	Вуз	Численность студ. всех форм обучения, чел.	Доля студ. заочной формы обучения в общей численности студентов, %	Численность студ. всех форм обучения, приходящихся на одного преподавателя, чел.	Выпуск молодых спец-тов по заочной форме обучения, чел.	Доля выпускников очной формы обуч., получивших дипломы с отличием, %
7	1	1500	40	14	280	13
	2	3140	35	13	760	12
	3	2050	20	10	400	20
	4	1100	25	15	250	10

8	1	1600	30	15	250	15
	2	3200	32	18	700	20
	3	2100	25	20	350	13
	4	1200	20	17	200	14
9	1	1800	35	13	280	20
	2	3000	30	12	650	35
	3	2000	25	10	400	21
	4	1500	40	8	320	10

Вариант 10,11,12

Работа автокомбината за месяц характеризуется следующими данными. Определить по автокомбинату в целом:

- 1) среднюю себестоимость одного ткм;
- 2) среднее число машин в автоколонне;
- 3) средний месячный грузооборот автомашины.

Указать вид рассчитанных данных и сделать выводы.

Вариант	Автоколонна	Общие затраты на перевозку грузов, руб.	Средний месячный грузооборот автомашины, ткм.	Себестоимость одного ткм., руб
10	1	20286	4600	0,63
	2	17820	5400	0,98
	3	47628	4400	0,81
	4	35800	5000	0,88
11	1	22300	4000	0,65
	2	20527	4800	0,79
	3	45320	5000	0,58
	4	38430	4600	0,82
12	1	25720	4200	0,98
	2	28355	4600	0,85
	3	43435	5100	0,78
	4	37527	4000	0,69

Вариант 13,14,15

По результатам обследования сельхозпредприятий области получены следующие данные. Определить средний валовой надой молока, средний надой молока на одну корову и среднюю жирность молока.

Указать вид рассчитанных средних и сделать выводы.

Вариант	Группы с/х предприятий по среднему годовому надояю молока от одной коровы, кг	Число с/х предприятий	Среднегодовое поголовье коров на одно с/х предприятие	Процент жира в молоке, %
13	До 2000	4	417	3,0
	2000 – 2200	9	350	3,3
	2200 – 2400	15	483	3,8
	2400 и более	8	389	2,9
14	До 1500	5	415	3,1
	1500 – 1700	8	380	3,2
	1700 – 1900	18	450	3,3
	1900 и более	9	375	3,0
15	До 1800	6	325	2,9
	1800 – 2000	7	415	3,0
	2000 – 2200	15	460	3,2
	2200 и более	6	380	3,1

Вариант 16,17,18

Вариант	Группы с/х предприятий по годовому валовому надоям молока, т	Число с/х предприятий	Среднегодовой надой молока от одной коровы, на с/х предприятия, кг.	Затраты кормовых единиц в расчете на 1 корову на 1 с/х предприятие, кг
16	До 500	38	1500	3,5
	500 – 1000	40	1550	3,8
	1000 – 1500	45	1608	4,0
	1500 и более	50	1610	3,7
17	До 1000	35	1600	3,4
	1000 – 1500	45	1750	3,6
	1500 – 2000	43	1800	3,8
	2000 и более	48	1830	4,2
18	До 800	50	2000	3,0
	800 – 1100	52	2100	3,2
	1100 – 1400	48	1850	3,3
	1400 и более	49	1980	2,9

Определите по совокупности всех с/х предприятий среднегодовые значения:

- 1) валового надоя молока;
 - 2) надоя молока от одной коровы;
 - 3) затрат кормовых единиц в расчете на одну корову;
- Укажите вид рассчитанных единиц и сделайте вывод.

Вариант 19,20,21

Вариант	Предприятие	Общие затраты на производство, млн. руб.	Затраты на 1 руб. произведенной прод., руб.	Процент бракованной прод., %	Процент прод., идущей на экспорт, %
19	1	2323,4	175	1,3	18
	2	8215,9	171	0,9	20
	3	4420,6	173	1,2	19
	4	3525,3	178	1,0	21
20	1	2525,5	250	1,2	17
	2	8152,9	251	1,1	18
	3	4240,6	257	0,8	22
	4	3253,2	260	0,9	21
21	1	2540,1	198	0,7	25
	2	7835,2	189	0,9	23
	3	4084,4	185	1,0	22
	4	3528,5	186	1,1	20

Определить по совокупности предприятий средние значения:

- 1) затрат на 1 руб. произведенной продукции;
- 2) процент бракованной продукции;
- 3) процент продукции, идущей на экспорт.

Укажите виды рассчитанных средних величин и сделайте выводы.

Вариант 22,23,24

Вариант	Предприятие	Объем произведенной продукции, млн. руб.	Производительность труда, тыс. руб.	Доля пенсионеров в общей численности работников	Средняя заработная плата на 1 работника, тыс. руб.
22	1	22,85	15,1	0,20	450
	2	20,73	17,2	0,18	565
	3	21,68	16,3	0,22	487
	4	19,98	18,2	0,19	523
23	1	30,78	16,2	0,21	570
	2	32,12	17,3	0,22	602
	3	34,00	18,1	0,23	510
	4	36,15	20,2	0,20	475

24	1	65,3	17,0	0,25	443
	2	78,1	18,1	0,24	488
	3	68,3	21,3	0,23	535
	4	70,9	19,8	0,22	560

Определить по совокупности предприятий средние значения:

- 1) производительности труда;
- 2) доли пенсионеров в общей численности работников;
- 3) заработной платы на одного работника.

Укажите виды рассчитанных средних величин и сделайте выводы.

Варианты 25,26,27

Вариант	Подгруппы культур	Посевная площадь, га	Урожайность, т/га	Внесено органических удобрений в расчете на 1га, т
25	Озимые	800	4	1,9
	Ярвые	800	3	1,8
	Зернобобовые	400	2	2,0
	Картофель	500	12	61,8
26	Озимые	750	4	1,9
	Сахарная свекла	500	28	57,2
	Овощи	400	12	26,1
	Картофель	600	20	60,5
27	Зерновые	780	2	2,0
	Зернобобовые	450	3	1,9
	Столовая свекла	200	25	50,1
	Картофель	650	13	61,2

Определить по совокупности с/х культур средние значения:

- 1) валового сбора;
- 2) урожайности;
- 3) посевных площадей;
- 4) органических удобрений на 1 га.

Укажите вид рассчитанных средних и сделайте выводы.

Варианты 28,29,30

Вариант	Предприятие	Фондотдача	Объем произведенной продукции, млн. руб.	Доля производственных мощностей
28	1	17	37,1	0,45
	2	15	42,3	0,52
	3	16	38,0	0,38
	4	18	40,5	0,49
29	1	10	48,2	0,36
	2	12	53,4	0,41
	3	13	48,4	0,39
	4	14	39,8	0,4
30	1	11	56,1	0,48
	2	13	53,2	0,39
	3	18	50,8	0,47
	4	19	52,7	0,30

По совокупности предприятий определить средние значения:

- 1) объема произведенной продукции;
- 2) фондотдачи;
- 3) доли произведенных мощностей;
- 4) среднегодовую стоимость основных фондов.

Укажите вид рассчитанных средних и сделайте выводы.

Глава 3. ПОКАЗАТЕЛИ ВАРИАЦИИ ПРИЗНАКА

Колблемость, изменяемость, многообразие величины признака у единиц совокупности называется вариацией.

Для характеристики размера вариации признака используются абсолютные и относительные показатели.

К абсолютным показателям вариации относятся: размах вариации; среднее линейное отклонение; среднее квадратическое отклонение; дисперсия.

Размах вариации: $R = x_{\max} - x_{\min}$, где x_{\max} и x_{\min} – соответственно максимальное и минимальное значения признака.

Среднее линейное отклонение (\bar{d}):

а) для несгруппированных данных $\bar{d} = \frac{\sum |x - \bar{x}|}{n}$,

б) для сгруппированных данных $\bar{d} = \frac{\sum |x - \bar{x}| \cdot f}{\sum f}$.

Среднее квадратическое отклонение (σ) и дисперсия (σ^2):

а) для несгруппированных данных $\sigma = \sqrt{\frac{\sum (x - \bar{x})^2}{n}}$, $\sigma^2 = \frac{\sum (x - \bar{x})^2}{n}$;

б) для сгруппированных данных $\sigma = \sqrt{\frac{\sum (x - \bar{x})^2 \cdot f}{\sum f}}$, $\sigma^2 = \frac{\sum (x - \bar{x})^2 \cdot f}{\sum f}$.

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

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

При сравнении колеблемости различных признаков в одной и той же совокупности или же при сравнении колеблемости одного и того же признака в нескольких совокупностях с различной величиной средней арифметической используются относительные показатели вариации:

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

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

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

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

3.1. Решение типовых задач

Задача 3.1.1

По дискретному ряду распределения рабочих цеха по квалификации вычислить показатели вариации.

Тарифный разряд X	Число рабочих f
2	4
3	5
4	9
5	4
6	2
Итого:	24

Решение. Для расчета показателей вариации удобно воспользоваться таблицей 1. Расчет показателей вариации.

Таблица 1

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

Найдем среднее значение признака x:

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

Среднее линейное отклонение \bar{d} :

$$\bar{d} = \frac{\sum |x - \bar{x}| \cdot f}{\sum f} = \frac{\sum |d| \cdot f}{\sum f} = \frac{22,2}{24} = 0,9 \text{ (разряда)}$$

Среднее квадратическое отклонение σ :

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

Коэффициент вариации v:

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

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

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

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

3.2. Задания для самостоятельной работы

Задача 3.2.1.

Распределение сотрудников коммерческого банка по стажу работы не менее 5 лет характеризуется следующими данными:

Стаж, лет	5	6	7	8	9	10
Число сотрудников	43	32	25	13	10	7

Рассчитайте размах вариации, среднее линейное отклонение, среднее квадратическое отклонение и коэффициент вариации. Сделайте выводы.

Ответ: R = 5; $\bar{d} = 1,3$ года; $\sigma^2 = 2,22$; $\sigma = 1,5$ года; $v_c = 23,1\%$.

Задача 3.2.2.

Имеются следующие данные о товарообороте продовольственных магазинов города:

Группы магазинов по товарообороту, млн.руб	Число магазинов	Группы магазинов по товарообороту, млн.руб	Число магазинов
40-50	2	90-100	20
50-60	4	100-110	22
60-70	7	110-120	11
70-80	10	120-130	6
80-90	15	130-140	3

Определить размах вариации, среднее линейное отклонение, дисперсию, среднее квадратическое отклонение и коэффициент вариации. Проанализируйте полученные результаты.

Ответ: $R = 100$ млн.руб.; $\bar{d} = 13,5$ млн.руб.; $\sigma^2 = 399,36$; $\sigma = 19,98$ млн.руб.; $v_{\sigma} = 21,2\%$.

3.3. Задания для индивидуальной работы

Задача 3.3.1.

Для распределения рабочих Минского тракторного завода по квалификации (тарифному разряду) вычислить абсолютные и относительные показатели вариации (см. глава I, задача 1.3.1).

Задача 3.3.2.

Для распределения рабочих Минского тракторного завода по стажу рассчитайте размах вариации, среднее линейное отклонение, среднее квадратическое отклонение и коэффициент вариации (см. глава I, задача 1.3.1). Сделайте выводы.

Глава 4. СЛОЖЕНИЕ ДИСПЕРСИЙ ИЗУЧАЕМОГО ПРИЗНАКА

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

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

Величина **общей дисперсии** (σ_0^2) характеризует вариацию признака под влиянием всех факторов, формирующих уровень признака из единиц данной совокупности, и определяется по формуле:

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

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

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

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

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

n_i – число единиц совокупности в отдельной группе.

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

Средняя внутригрупповая дисперсия определяется по формуле:

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

где σ_i^2 – дисперсия по отдельной группе; $\sigma_i^2 = \frac{\sum (x - \bar{x}_i)^2 f}{\sum f}$.

Указанные дисперсии взаимосвязаны между собой следующим равенством: величина общей дисперсии равна сумме межгрупповой дисперсии и средней внутригрупповой дисперсии: $\sigma_o^2 = \delta^2 + \bar{\sigma}^2$.

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

Зная любые два вида дисперсий, можно определить или проверить правильность расчета третьего вида.

Правило сложения дисперсий применяется при исчислении показателей тесноты связи, в дисперсионном анализе и ряде других случаев. В статистическом анализе широко используется показатель, представляющий собой долю межгрупповой дисперсии в общей дисперсии. Он носит название эмпирического коэффициента детерминации (η^2):

$$\eta^2 = \frac{\delta^2}{\sigma^2}.$$

Этот коэффициент показывает долю (удельный вес) общей вариации изучаемого признака, обусловленную вариацией группировочного признака.

Эмпирическое корреляционное отношение (η): $\eta = \sqrt{\frac{\delta^2}{\sigma^2}}$. Оно характеризует влияние признака, положенного в основание группировки, на вариацию результативного признака.

Эмпирическое корреляционное отношение изменяется в пределах от 0 до 1. Если $\eta = 0$, то группировочный признак не оказывает влияние на результативный. Если $\eta = 1$, то результативный признак изменяется только в зависимости от признака, положенного в основание группировки, а влияние прочих факторов равно нулю. Промежуточные значения оцениваются в зависимости от их близости к предельным значениям.

Выше было рассмотрено правило сложения дисперсий для количественных признаков. Но наряду с вариацией количественных признаков может наблюдаться и вариация качественных признаков. Такое изучение вариации достигается, как и для количественных признаков, посредством вычисления и анализа следующих видов дисперсий.

Внутригрупповая дисперсия доли определяется по формуле:

$$\sigma_{p_i}^2 = p_i(1 - p_i)$$

средняя из внутригрупповых дисперсий рассчитывается так:

$$\bar{\sigma}_{p_i}^2 = \overline{p_i(1 - p_i)} = \frac{\sum p_i(1 - p_i)n_i}{\sum n_i}$$

Формула межгрупповой дисперсии имеет вид:

$$\delta_{p_1}^2 = \frac{\sum (p_i - \bar{p})^2 \cdot n_i}{\sum n_i},$$

где n_i – численность единиц в отдельных группах;

\bar{p} – доля изучаемого признака по всей совокупности, которая определяется по формуле:

$$\bar{p} = \frac{\sum p_i n_i}{\sum n_i}.$$

Общая дисперсия определяется по формуле:

$$\delta_p^2 = \bar{p}(1 - \bar{p}).$$

Три вида дисперсии связаны между собой следующим образом:

$$\sigma_p^2 = \sigma_{p_1}^2 + \delta_{p_1}^2$$

Это соотношение дисперсий называется теоремой сложения дисперсии доли признака.

4.1. Решение типовых задач

Задача 4.1.1

Имеются данные о производительности труда рабочих одного из цехов НПО «Циклон» (табл. 1). Требуется: проверить правило сложения дисперсий, вычислить эмпирическое корреляционное отношение.

Таблица 1 – Производительность труда двух групп рабочих одного из цехов НП «Циклон»

Производительность труда рабочих									
прошедших техническое обучение, деталей за смену					не прошедших техническое обучение, деталей за смену				
84	93	95	101	102	62	68	82	88	105

Рассчитаем общую и групповые средние и дисперсии:

$$\bar{x}_1 = \frac{\sum x_i}{n_1} = \frac{475}{5} = 95 \text{ (дет.)}; \quad \bar{x}_2 = \frac{\sum x_i}{n_2} = \frac{405}{5} = 81 \text{ (дет.)};$$

$$\bar{x} = \frac{475 + 405}{10} = 83 \text{ (дет.)};$$

$$\sigma_1^2 = \frac{\sum (x - \bar{x})^2}{n_1} = \frac{(84 - 95)^2 + (93 - 95)^2 + \dots + (102 - 95)^2}{5} = 42,$$

$$\sigma_2^2 = \frac{\sum (x - \bar{x})^2}{n_2} = \frac{(62 - 81)^2 + (68 - 81)^2 + \dots + (105 - 81)^2}{5} = 231,2,$$

$$\sigma^2 = \frac{(84 - 88)^2 + (93 - 88)^2 + \dots + (105 - 88)^2}{10} = 185,6.$$

Исходные данные для вычисления средней из внутригрупповых и межгрупповой дисперсии представлены в таблице 2.

Таблица 2 – Расчет σ_1^2 и δ^2 по двум группам рабочих

Группы рабочих	Численность рабочих, чел.	Средняя, дет./смен	Дисперсия
Прошедшая техн. обучение	5	95	42,0
Не прошедшая техн. обучение	5	81	231,2
Все рабочие	10	88	185,6

Рассчитаем следующие показатели.

Средняя из внутригрупповых дисперсий:

$$\bar{\sigma}_i^2 = \frac{42 \cdot 5 + 231,2 \cdot 5}{10} = 136,6;$$

межгрупповая дисперсия:

$$\delta^2 = \frac{(95 - 88)^2 \cdot 5 + (81 - 88)^2 \cdot 5}{10} = \frac{245 + 245}{10} = 49,0;$$

общая дисперсия $\sigma^2 = 136,6 + 49 = 185,6$

Таким образом, эмпирическое корреляционное отношение:

$$\eta = \sqrt{\frac{49,0}{185,6}} = 0,514, \text{ т.е. фактор технического обучения, объясняет в данном при-}$$

мере 51,4% вариации производительности труда рабочих, а неучтенные факторы 48,6%.

Задача 4.1.2.

Имеются следующие данные о результатах обследования рабочих предприятия по размеру заработной платы:

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

Общая дисперсия заработной платы в обследованной совокупности рабочих составила 450.

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

Решение. Средняя внутригрупповая дисперсия характеризует случайную вариацию под влиянием неучтенных факторов:

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

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

Межгрупповая дисперсия определяется по правилу сложения дисперсий:

$$\delta^2 = \sigma^2 - \bar{\sigma}^2 = 450 - 413,5 = 36,5$$

Отсюда соотношение дисперсий (эмпирический коэффициент детерминации):

$$\eta^2 = \frac{\delta^2}{\sigma^2} = \frac{36,5}{450} = 0,08 \text{ или } 8\% \text{ и эмпирическое корреляционное отношение}$$

$$\eta = \sqrt{\eta^2} = \sqrt{0,08} = 0,28.$$

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

Задача 4.1.3. Имеются следующие данные о времени простоя автомобиля под разгрузкой:

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

Проверить закон сложения дисперсий.

Решение

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

$$\sigma_o^2 = \frac{\sum(x - \bar{x}_o)^2 \cdot f}{\sum f}$$

Для расчета общей дисперсии составим дискретный ряд распределения, промежуточные расчеты поместим таблицу 3.

Таблица 3 – Вспомогательная таблица для расчета общей дисперсии

Время простоя под разгрузкой, мин, x	Число выполненных разгрузок, f	x · f	x - \bar{x}_o	(x - \bar{x}_o) ²	(x - \bar{x}_o) ² · f
8	3	24	-4	16	48
10	2	20	-2	4	8
12	2	24	0	0	0
15	1	15	3	9	9
18	1	18	6	36	36
19	1	19	7	49	49
Итого:	10	120	—	—	150

$$\bar{x}_o = \frac{\sum x_i f_i}{\sum f_i} = \frac{120}{10} = 12 \text{ мин.} - \text{среднее время простоя;}$$

$$\sigma_o^2 = \frac{\sum(x - \bar{x}_o)^2 f}{\sum f} = \frac{150}{10} = 15 - \text{общая дисперсия.}$$

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

Различия в величине изучаемого признака прежде всего возникают под влиянием числа грузчиков, принимающих участие в процессе разгрузки. В связи с этим в совокупности выделяются две однородные группы по числу грузчиков: в первую группу включаются наблюдения при числе грузчиков 3; во вторую группу попадают наблюдения при числе грузчиков 4. Для каждой из выделенных групп определяются внутригрупповая дисперсия, возникающая под влиянием неучтенных факторов. Для их расчета используем вспомогательные таблицы 4 и 5.

Таблица 4 – Расчет внутригрупповой дисперсии по первой группе (число грузчиков, участвующих в группе – 3).

Время простоя под разгрузкой, мин, x	Число выполненных разгрузок, f	x · f	x - \bar{x}_1	(x - \bar{x}_1) ² f _i
12	1	12	-4	16
15	1	15	-1	1
18	1	18	2	4
19	1	19	3	9
Итого:	4	64	—	30

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

Таблица 5 – Расчет внутригрупповой дисперсии по второй группе (число грузчиков, участвующих в группе – 4)

Время простоя под разгрузкой, мин, x	Число выполненных разгрузок, f	x · f	x – \bar{x}_2	(x – \bar{x}_2) ² · f
8	3	24	-1,33	5,31
10	2	20	0,67	0,90
12	1	12	2,57	7,13
Итого:	6	56	—	—

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

Средняя из внутригрупповых дисперсий

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

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

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

Общая дисперсия (σ_o^2) равна сумме средней внутригрупповой дисперсии и межгрупповой дисперсии: $\sigma_o^2 = 4,3 + 10,7 = 15$, что и соответствует полученной ранее величине.

Задача 4.1.4

Имеются следующие данные об удельном весе основных рабочих в трех цехах фирмы:

Цех	Удельный вес основных рабочих, в % (p_i)	Численность всех рабочих, чел. (n_i)
1	80	100
2	75	200
3	90	150
Итого:	—	450

Проверить теорему сложения дисперсии доли признака.

Решение. Определим долю основных рабочих в целом по фирме:

$$\bar{p} = \frac{\sum p_i n_i}{\sum n_i} = \frac{0,80 \cdot 100 + 0,75 \cdot 200 + 0,90 \cdot 150}{450} = \frac{365}{450} = 0,81$$

1. Общая дисперсия доли основных рабочих по всей фирме в целом будет равна:

$$\sigma_p^2 = \bar{p}(1 - \bar{p}) = 0,81 \cdot (1 - 0,81) = 0,154$$

2. Внутригрупповые дисперсии:

$$\sigma_{p_i}^2 = p_i(1 - p_i)$$

$$\sigma_{p_1}^2 = 0,8 \cdot 0,2 = 0,16; \sigma_{p_2}^2 = 0,75 \cdot 0,25 = 0,19; \sigma_{p_3}^2 = 0,9 \cdot 0,1 = 0,09.$$

3. Средняя из внутригрупповых дисперсий будет равна:

$$\bar{\sigma}_{p_i}^2 = \frac{\sum p_i(1 - p_i)n_i}{\sum n_i} = \frac{0,16 \cdot 100 + 0,19 \cdot 200 + 0,09 \cdot 150}{450} = \frac{67,5}{450} = 0,15$$

4. Межгрупповая дисперсия:

$$\delta_p^2 = \frac{\sum (p_i - \bar{p})^2 n_i}{\sum n_i} = \frac{(0,8 - 0,81)^2 \cdot 100 + (0,75 - 0,81)^2 \cdot 200 + (0,09 - 0,81)^2 \cdot 150}{450}$$

$$= \frac{0,01 + 0,72 + 1,215}{450} = \frac{1,945}{450} = 0,004$$

Проверка вычислений показывает: $0,154 = 0,15 + 0,004$.

4.2. Задания для самостоятельной работы

Задача 4.2.1. Совместные предприятия одной из отраслей сгруппированы по стоимости реализованной продукции и услуг за год. Результаты представлены в следующей таблице:

Стоимость реализованной продукции и услуг, млрд.руб.	Число предприятий	Средняя стоимость реализованной продукции и услуг по группе, млрд.руб	Групповая дисперсия
3,5-6,5	9	5,59	6,13
6,5-9,5	10	7,06	6,51
9,5 и выше	11	12,20	72,16

Определите:

- Среднюю из групповых дисперсий, межгрупповую дисперсию, общую дисперсию (по правилу сложения дисперсий);
- Эмпирическое корреляционное отношение.

Дайте экономическую интерпретацию полученных результатов.

Ответ: $\sigma_1^2 = 30,47$; $\delta^2 = 8,25$; $\sigma^2 = 38,72$; $\eta = 0,46$.

Задача 4.2.2. Зарботная плата 10 рабочих бригады характеризуется следующими данными:

Профессия	Число рабочих	Месячная зарботная плата каждого рабочего за март, ден. ед.
Токари	4	1252; 1548; 1600; 1400
Слесари	6	1450; 1380; 1260; 1700; 1250; 1372

Проверить правило сложения дисперсий и указать, велико ли влияние профессии на различие в уровне зарботной платы.

Ответ: $\sigma^2 = 21509,8$; $\delta^2 = 553,0$; $\sigma^2 = 20956,8$.

4.3. Задания для индивидуальной работы

По данным задачи 1.3.1 главы 1 требуется:

- определить общую дисперсию зарботной платы рабочих;
- произвести группировку рабочих завода по уровню квалификации; для каждой выделенной группы вычислить внутригрупповую дисперсию по уровню месячной зарботной платы;
- определить среднюю внутригрупповую дисперсию по уровню месячной зарботной платы;
- проверить правило сложения дисперсий.

Глава 5. ВЫБОРОЧНОЕ НАБЛЮДЕНИЕ

При проведении **выборочного наблюдения**, как и всякого несплошного наблюдения, обследуются не все единицы изучаемого объекта или **генеральной совокупности**, а лишь некоторая, так или иначе отобранная часть этих единиц. Однако наблюдение организовано таким образом, что эта часть отобранных единиц в уменьшенном масштабе репрезентирует всю совокупность. Часть единиц генеральной совокупности, подлежащей непосредственному наблюдению, называют **выборочной совокупностью**.

Ошибка выборочного наблюдения – это разность между величиной параметра в генеральной совокупности и его величиной, вычисленной по результатам выборочного наблюдения. Расчет ошибок позволяет решить одну из главных проблем организации выборочного наблюдения – оценить репрезентативность (представительность) выборочной совокупности. Различают среднюю и предельную ошибки выборки. Эти два вида ошибок связаны следующим соотношением: $\Delta = t\mu$, где Δ – предельная ошибка выборки; μ – средняя ошибка выборки; t – коэффициент доверия, определяемый в зависимости от уровня вероятности.

Величина средней ошибки выборки рассчитывается дифференцированно в зависимости от способа отбора и процедуры выборки. Так, при случайном повторном отборе

средняя ошибка определяется по формуле $\mu = \frac{\sigma}{\sqrt{n}}$, при бесповторном: $\mu = \sqrt{\frac{\sigma^2}{n} \left(1 - \frac{n}{N}\right)}$,

где σ^2 – выборочная (или генеральная) дисперсия; σ – выборочное (или генеральное) среднее квадратическое отклонение; n – объем выборочной совокупности; N – объем генеральной совокупности.

Расчет средней и предельной ошибок выборки позволяет определить возможные пределы, в которых будут находиться характеристики генеральной совокупности. Например, для выборочной средней такие пределы устанавливаются на основе следующих соотношений: $\bar{x} - \Delta_{\bar{x}} \leq \bar{X} \leq \bar{x} + \Delta_{\bar{x}}$, где \bar{X} и \bar{x} – генеральная и выборочная средние соответственно; $\Delta_{\bar{x}}$ – предельная ошибка выборочной средней.

5.1. Решение типовых задач

Задача 5.1. При проверке веса импортируемого груза на таможне методом случайной повторной выборки было отобрано 200 изделий. В результате был установлен средний вес изделия 30 г при среднем квадратическом отклонении 4 г. С вероятностью 0,997 определите пределы, в которых находится средний вес изделий в генеральной совокупности.

Решение. Рассчитаем сначала предельную ошибку выборки. Так, при доверительной вероятности $\gamma = 0,997$ из равенства $\Phi(t) = \frac{\gamma}{2} = \frac{0,997}{2} = 0,4985$ по таблице значений

функции $\Phi(t)$ найдем значение $t = 3$. Тогда получим предельную ошибку: $\Delta_{\bar{x}} = t \cdot \frac{\sigma_{\bar{x}}}{\sqrt{n}} =$

$$= 3 \cdot \frac{4}{\sqrt{200}} \approx 0,84.$$

Определим пределы генеральной средней:

$$30 - 0,84 \leq \bar{X} \leq 30 + 0,84 \text{ или } 29,16 \leq \bar{X} \leq 30,84.$$

Следовательно, с вероятностью 0,997 можно утверждать, что средний вес изделий в генеральной совокупности находится в пределах от 29,16 до 30,84 г.

Задача 5.2. В городе проживает 250 тыс. семей. Для определения среднего числа детей в семье была организована 2%-ная случайная бесповторная выборка семей. По ее результатам было получено следующее распределение семей по числу детей:

Число детей в семье	0	1	2	3	4	5
Количество семей	1000	2000	12000	400	200	200

С вероятностью 0,954 найдите пределы, в которых будет находиться среднее число детей в генеральной совокупности.

Решение. Вначале на основе имеющегося распределения семей определим выборочные среднюю и дисперсию:

Число детей в семье x_i	Количество семей f_i	$x_i \cdot f_i$	$x_i - \bar{x}$	$(x_i - \bar{x})^2$	$(x_i - \bar{x})^2 f_i$
0	1000	0	-1,5	2,25	2250
1	2000	2000	-0,5	0,25	500
2	1200	2400	0,5	0,25	300
3	400	1200	1,5	2,25	900
4	200	800	2,5	6,25	1250
5	200	1000	3,5	12,25	2450
Итого	5000	7400	-	-	7650

$$\bar{x} = \frac{7400}{5000} \approx 1,5, \quad \sigma_x^2 = \frac{7650}{5000} = 1,53.$$

Вычислим теперь предельную ошибку выборки (с учетом того, что $\gamma=0,954 \Rightarrow t=2$):

$$\Delta_{\bar{x}} = t \cdot \sqrt{\frac{\sigma_x^2}{n} \left(1 - \frac{n}{N}\right)} = 2 \cdot \sqrt{\frac{1,53}{5000} \left(1 - \frac{5000}{250000}\right)} \approx 0,035.$$

Следовательно, пределы генеральной средней: $\bar{x} = \bar{x} \pm \Delta_{\bar{x}} = 1,53 \pm 0,035$.

Таким образом, с вероятностью 0,954 можно утверждать, что среднее число детей в семьях города практически не отличается от 1,5, т.е. в среднем на каждые две семьи приходятся три ребенка.

Наряду с определением ошибок выборки и пределов для генеральной средней эти же показатели могут быть определены для доли признака. В этом случае особенности расчета связаны с определением дисперсии доли, которая вычисляется так:

$\sigma_w^2 = w(1-w)$, где $w = \frac{m}{n}$ – доля единиц, обладающих данным признаком в выборочной совокупности, определяемая как отношение количества соответствующих единиц к объему выборки.

Тогда, например, при собственно-случайном повторном и бесповторном отборах для определения предельной ошибки выборки используется следующие формулы соответственно:

$$\Delta_w = t \cdot \sqrt{\frac{\sigma_w^2}{n}} = t \cdot \sqrt{\frac{w(1-w)}{n}} \quad \text{и} \quad \Delta_w = t \cdot \sqrt{\frac{\sigma_w^2}{n} \left(1 - \frac{n}{N}\right)} = t \cdot \sqrt{\frac{w(1-w)}{n} \left(1 - \frac{n}{N}\right)}.$$

Пределы доли признака в генеральной совокупности p выглядят следующим образом: $w - \Delta_w \leq p \leq w + \Delta_w$.

Задача 5.3. С целью определения средней фактической продолжительности рабочего дня в государственном учреждении с численностью служащих 480 человек в июне 1996 г. была проведена 25%-ная механическая выборка. По результатам наблюдения оказалось, что у 10% обследованных потери времени достигали более 45 мин. в день. С вероятностью 0,683 установите пределы, в которых находится генеральная доля служащих с потерями рабочего времени более 45 мин в день.

Решение. Определим объем выборочной совокупности: $n = 480 \cdot 0,25 = 120$ человек. Выборочная доля w равна по условию 10%. Учитывая, что показатели точности механической и собственно-случайной бесповторной выборки определяются одинаково, а также то, что при $\gamma = 0,683 \Rightarrow t = 1$, вычислим предельную ошибку выборочной доли:

$$\Delta_w = t \cdot \sqrt{\frac{w(1-w)}{n} \left(1 - \frac{n}{N}\right)} = 1 \cdot \sqrt{\frac{0,1(1-0,1)}{120} \left(1 - \frac{120}{480}\right)} = 0,0237 \approx 0,024 \text{ или } 2,4\%.$$

Пределы доли признака в генеральной совокупности:

$$10 - 2,4 \leq p \leq 10 + 2,4 \text{ или } 7,6 \leq p \leq 12,4.$$

Таким образом, с вероятностью 0,683 можно утверждать, что доля работников учреждения с потерями рабочего времени более 45 мин. в день находится в пределах от 7,6 до 12,4%.

Ошибки и пределы генеральных характеристик при других способах формирования выборочной совокупности определяются на основе соответствующих формул, отражающих особенности этих видов выборки.

При применении выборочного наблюдения одной из основных задач является определение объема выборки, необходимого для получения требуемой точности результатов с заданной вероятностью. Объем выборки рассчитывается на стадии проектирования выборочного обследования. Для собственно-случайной и механической выборки необходимый объем выборки при повторном и бесповторном отборах находится соответственно по формулам:

$$n = \frac{t^2 \sigma^2}{\Delta^2} \text{ и } n = \frac{t^2 \sigma^2 N}{\Delta^2 N + t^2 \sigma^2}.$$

Задача 5.4. В 100 туристических агентствах города предполагается провести обследование среднемесячного количества реализованных путевок методом механического отбора. Какова должна быть численность выборки, чтобы с вероятностью 0,683 ошибка не превышала 3 путевок, если по данным пробного обследования дисперсия составляет 225?

Решение. Рассчитаем необходимый объем выборки:

$$n = \frac{1^2 \cdot 225 \cdot 100}{3^2 \cdot 100 + 1^2 \cdot 225} = \frac{22500}{1125} = 20 \text{ агентств.}$$

5.2. Задания для самостоятельной работы

5.2.1. Что произойдет с величиной предельной ошибки выборки, если вероятность, гарантирующую результат: а) увеличить с 0,954 до 0,997; б) уменьшить с 0,954 до 0,683; в) увеличить с 0,683 до 0,954; г) уменьшить с 0,997 до 0,954; д) увеличить с 0,683 до 0,997?

5.2.2. Определите, как изменится средняя ошибка случайной выборки, если необходимую численность выборочной совокупности: а) уменьшить в 2,5 раза; на 40%; б) увеличить в 1,5 раза; на 20%. Как нужно применить необходимую численность выборки, чтобы средняя ошибка уменьшилась в 2 раза; на 50%; на 30%?

5.2.3. Какой должна быть необходимая численность выборки при механическом отборе, чтобы установить генеральную долю с ошибкой не более 2%, если дисперсия доли неизвестна, а отбор производится из совокупности, включающей: а) 1000 единиц; б) 10 000 единиц; в) 100 000 единиц? Вероятность, гарантирующая результаты выборочного наблюдения, равна 0,954.

5.2.4. Каким должен быть объем случайной бесповторной выборки из генеральной совокупности численностью 10000 единиц при среднем квадратическом отклонении не более 20, предельной ошибке, не превышающей 5%, и вероятности 0,997?

5.2.5. С целью определения трудоемкости изготовления деталей на предприятии произведен хронометраж работы 50 рабочих, отобранных в случайном порядке. По данным обследований получили $\bar{x} = 10$ мин, при $\sigma = 1$ мин. Определите:

а) как изменится ошибка выборки, если объем выборочной совокупности увеличить в 1,5 раза?

б) как скажется на ошибке выборки увеличение дисперсии в 2 раза?

в) как изменится ошибка выборки, если с увеличением дисперсии в 1,44 раза объем выборочной совокупности увеличить в 2,56 раза?

г) как изменится ошибка выборки, если численность генеральной совокупности будет в 3 раза больше?

5.2.6. С целью определения средних затрат времени при поездках на работу населением города планируется выборочное наблюдение на основе случайного повторного отбора. Сколько людей должно быть обследовано, чтобы с вероятностью 0,954 ошибка выборочной средней не превышала 1 мин при среднем квадратическом отклонении 15 мин?

5.2.7. Из партии в 1 млн. шт. мелкокалиберных патронов путем случайного отбора взято для определения дальности боя 1000 шт. Результаты испытаний представлены в следующей таблице:

Дальность боя, м	25	30	35	40	45	50	Итого
Число патронов, шт.	120	180	280	170	140	110	1000

С вероятностью 0,954 определите среднюю дальность боя по выборке, ошибку выборки и возможные пределы средней дальности боя для всей партии патронов.

5.2.8. В порядке механической выборки обследован возраст 100 студентов вуза из общего числа 2000 человек. Результаты обработки материалов наблюдения приведены в таблице:

Возраст, лет	17	18	19	20	21	22	23
Число студентов, чел.	11	13	18	23	17	10	8

Установите: а) средний возраст студентов вуза по выборке; б) величину ошибки при определении возраста студентов на основе выборки; в) вероятные пределы колебания возраста для всех студентов при вероятности 0,997.

5.2.9. Для изучения заработной платы работников завода строительных материалов из общей численности 5400 человек были отобраны в порядке собственно случайной бесповторной выборки 600 человек, и зафиксированы следующие результаты:

Месячная зарплата, руб.	Число рабочих
от 15000 до 25000	20
от 25000 до 35000	79
от 35000 до 45000	130
от 45000 до 55000	150
от 55000 до 65000	110
от 65000 до 75000	100
от 75000 до 85000	11

Определите:

1. Среднюю выборочную заработную плату;
2. Среднюю ошибку выборочной средней;
3. Предельную ошибку выборочной средней с вероятностью 0,96 и укажите интервал, в котором находится генеральная средняя;
4. Долю рабочих в выборке с месячной зарплатой до 30000 руб.; свыше 72000 руб. и укажите интервалы, в которых находятся эти генеральные доли.

5.2.10. В процессе технического контроля из партии готовой продукции методом случайного бесповторного отбора было проверено 70 изделий, из которых 4 оказались бракованными. Можно ли с вероятностью 0,954 утверждать, что доля бракованных изделий во всей партии не превышает 7%, если процент отбора равен 10?

5.2.11. С целью определения среднего размера вклада в отделениях Сбербанка города предполагается провести механическую выборку лицевых счетов из общего числа 67800. По данным предыдущего обследования установлено среднее квадратическое отклонение размера вклада, равное 140 тыс. руб. С вероятностью 0,997 определите необходимый объем выборочной совокупности при условии, что ошибка выборки не превышает 10 тыс. руб.

5.2.12. Определите, сколько персональных компьютеров следует подвергнуть обследованию в порядке случайной бесповторной выборки, чтобы с вероятностью 0,954 предельная ошибка (в процентах к среднему сроку службы компьютера) не превышала 3%. Коэффициент вариации среднего срока службы компьютеров по данным предыдущих обследований составляет 15%, а вся партия состоит из 1250 компьютеров.

5.2.13. При планировании выборочного обследования занятости мужского населения сельских районов республики имеются следующие данные:

Район	Численность мужчин в трудоспособ. возрасте, тыс. чел.	Удельный вес занятых мужчин, % (оценка)
1	3,5	75
2	5,6	80
3	1,7	70
4	2,8	85

С вероятностью 0,954 определите необходимый объем типической пропорциональной выборки для установления границ генеральной доли: а) при повторном отборе; б) при бесповторном отборе в районах, чтобы ошибка выборки не превышала 5%.

5.2.14. На АО "Машиноаппарат" 2400 станков, в том числе токарных – 960, фрезерных – 720, шлифовальных – 480, прочих – 240. С целью исследования производительности станков планируется организовать типическую пропорциональную выборку станков с механическим отбором внутри групп. По результатам аналогичного обследования на другом подобном предприятии среднее квадратическое отклонение составило 60. Сколько станков необходимо отобрать из каждой группы, чтобы ошибка выборки не превышала 20 единиц при вероятности 0,997?

5.2.15. В процессе подготовки выборочного обследования качества импортируемых кондитерских изделий была проведена пробная проверка 8 ящиков для сбора данных о вариации их веса. Результаты проверки представлены в следующей таблице:

№ ящика	1	2	3	4	5	6	7	8
Средний вес коробки в ящике, г	540	520	550	500	510	530	560	520

Сколько ящиков с кондитерскими изделиями необходимо отобрать для проверки качества в порядке бесповторного отбора, чтобы с вероятностью 0,997 ошибка выборки не превышала 20 г, если генеральная совокупность включает 1000 равных по величине серий?

5.2.16. На склад коммерческой организации поступило 480 коробок с микрокалькуляторами, упакованными по 24 калькулятора в каждой коробке. Планируется проверить соответствие калькуляторов международным стандартам. Определите необходимый объем выборки, если результат требуется гарантировать с вероятностью 0,954 и ошибкой не более 5%, а межгрупповая дисперсия равна 51.

5.3. Задания для индивидуальной работы

1.1. В порядке механической выборки было подвергнуто испытанию на разрыв 100 нитей из партии. В результате обследования установлена средняя крепость пряжи 320 г при среднем квадратическом отклонении 22 г. С вероятностью 0,954 определите пределы, в которых находится средняя крепость пряжи.

1.2. Для определения среднего размера вклада определенной категории вкладчиков в банках города, где число вкладчиков 5000, необходимо провести выборку лицевых счетов методом механического отбора. Предварительно установлено, что среднее квадратическое отклонение размера вкладов составляет 120 у.е. Определите необходимую численность выборки при условии, что с вероятностью 0,954 ошибка выборки не превысит 10 у.е.

1.3. В области 10 тыс. семей. Из них 6 тыс. рабочих, 3 тыс. колхозников, 1 тыс. служащих. С целью определения доли семей, имеющих более трех детей, предполагается провести типическую выборку с пропорциональным отбором. Отбор внутри типов механический. Какое количество семей необходимо отобрать, чтобы с вероятностью 0,954 ошибка выборки не превышала 5%? Дисперсия типической выборки равна 2100.

2.1. Для определения зольности угля месторождения в порядке случайной выборки взято 400 проб. В результате исследования установлена средняя зольность угля в выработке 16% при среднем квадратическом отклонении 4%. С вероятностью 0,997 определите пределы, в которых находится средняя зольность угля.

2.2. На ткацкой фабрике работает 6 тыс. ткачих. Для установления норм выработки предполагается провести случайный бесповторный отбор ткачих. Предварительным обследованием установлено, что среднее квадратическое отклонение дневной выработки составляет 25 м. Определите необходимую численность выборки при условии, что с вероятностью 0,954 ошибка выборки не превысит 5 м.

2.3. На заводе 2000 рабочих. Из них 1500 мужчин и 500 женщин. С целью определения доли рабочих, которые проработали на заводе более 5 лет, предполагается провести типическую выборку рабочих с пропорциональным отбором. Отбор внутри групп механический. Какое число рабочих необходимо отобрать, чтобы с вероятностью 0,954 ошибка выборки не превышала 3%, если на основе обследования известно, что дисперсия типической выборки равна 2100?

3.1. В порядке случайной повторной выборки из партии было взято 100 проб продукта А. В результате исследования установлена средняя влажность продукта А в выборке 9% при среднем квадратическом отклонении 1,5%. С вероятностью 0,954 определите пределы, в которых находится средняя влажность продукта А в партии.

3.2. На заводе, где работает 10 тыс. рабочих, необходимо установить их средний стаж работы методом механической выборки. Предварительным обследованием установлено, что среднее квадратическое отклонение стажа работы равно 5 годам. Определите необходимую численность выборки при условии, что с вероятностью 0,997 ошибка выборки не превысит 1 года.

3.3. В городе 5 тыс. семей. Из них 1 тыс. составляют одинокие и 4 тыс. – семейные. С целью определения доли семей, имеющих отдельные квартиры, предполагается провести типическую выборку семей с пропорциональным отбором внутри типических групп. Отбор внутри групп механический. Какое количество семей необходимо отобрать, чтобы с вероятностью 0,954 ошибка выборки не превышала 5%? На основании предыдущих обследований известно, что дисперсия типической выборки равна 100.

4.1. В порядке случайной выборки обследовано 900 деревьев, при этом установлен средний диаметр одного дерева 235 мм и среднее квадратическое отклонение равно 27 мм. С вероятностью 0,954 определите границы, в которых будет находиться средний диаметр деревьев в генеральной совокупности.

4.2. В городе А с целью определения средней продолжительности поездки населения на работу предполагается провести выборочное обследование методом случайного повторного отбора. Какова должна быть численность выборки, чтобы с вероятностью 0,997 ошибка выборочной средней не превышала 5 мин при среднем квадратическом отклонении 20 мин?

4.3. В городе 15 тыс. семей. Из них 10 тыс. – семьи рабочих, 4 тыс. – служащих, 1 тыс. – колхозников. Для определения среднего размера семьи в городе предполагается провести типическую выборку семей с пропорциональным отбором внутри типических групп. Отбор внутри типов механический. Какое количество семей необходимо отобрать, чтобы с вероятностью 0,954 ошибка выборки не превышала 0,5 человека, если на основе предыдущих обследований известно, что дисперсия типической выборки равна 2,25?

5.1. На заводе с числом рабочих 1000 человек было проведено 2%-ное выборочное обследование возраста рабочих методом случайного бесповторного отбора. В результате обследования получены следующие данные:

Возраст рабочих, лет	До 30	30-40	40-50	50-60	Свыше 60
Число рабочих	8	22	19	6	4

С вероятностью 0,997 определите пределы, в которых находится средний возраст рабочих завода.

5.2. С целью определения среднего диаметра деревьев необходимо провести выборочное обследование методом случайного повторного отбора. Какова должна быть численность выборки, чтобы с вероятностью 0,954 ошибка выборочной средней не превышала 15 см при среднем квадратическом отклонении 25 мм?

5.3. На машиностроительном заводе 1600 станков четырех типов. Из них 1-го типа – 320, 2-го типа – 420, 3-го типа – 640 и 4-го типа – 160. Для изучения производительности станков предполагается провести типическую выборку станков с пропорциональным отбором. Отбор внутри типов механический. Какое количество станков необходимо отобрать, чтобы с вероятностью 0,954 ошибка выборки не превышала 10 ед. изделий? На основе предыдущих обследований известно, что дисперсия типической выборки равна 4900.

6.1. Для изучения производительности труда токарей на машиностроительном заводе было проведено 10-% выборочное обследование 100 рабочих методом случайного бесповторного отбора. В результате обследования получены данные о часовой выработке рабочих:

Часовая выработка	18-20	20-22	22-24	24-26	26-28	28-30
Число рабочих	2	8	24	50	12	4

С вероятностью 0,997 определите пределы, в которых находится среднее время обработки одной детали токарными завода.

6.2. С целью определения качества пряжи на прядильной фабрике предполагается провести выборочное обследование пряжи методом случайного повторного отбора. Какова должна быть численность выборки, чтобы с вероятностью 0,954 ошибка выборочной средней не превышала 4 г при среднем квадратическом отклонении 20 г?

6.3. В механическом цехе завода 1000 рабочих. Из них 800 квалифицированных и 200 неквалифицированных. С целью изучения производительности труда предполагается провести типическую выборку рабочих с пропорциональным отбором. Отбор внутри групп механический. Какое число рабочих необходимо отобрать, чтобы с вероятностью 0,997 ошибка выборки не превышала 6 ед. изделий при среднем квадратическом отклонении 25?

7.1. На машиностроительном заводе с числом рабочих 5000 человек было проведено 4%-ное выборочное обследование квалификации рабочих методом случайного бесповторного отбора. В результате обследования получены следующие данные:

Тарифный разряд	1	2	3	4	5	6
Число рабочих	10	30	40	70	30	50

С вероятностью 0,997 определите пределы, в которых находится средний тарифный разряд рабочих завода.

7.2. На заводе предполагается провести выборочное обследование средней часовой выработки рабочих методом случайного повторного отбора. Какова должна быть численность выборки, чтобы с вероятностью 0,954 ошибка выборки не превышала 5 шт., если на основе предыдущих обследований известно, что дисперсия равна 225?

7.3. В области 10 тыс. молочных коров. Из них в районе А – 5 тыс., В – 3 тыс., С – 2 тыс. С целью определения средней удойности коров предполагается провести типическую выборку с пропорциональным отбором внутри типических групп. Отбор внутри групп механический. Какое количество коров нужно отобрать, чтобы с вероятностью 0,954 ошибка выборки не превышала 5 л, если на основе предыдущих обследований известно, что дисперсия типической выборки равна 1600?

8.1. В районе 2000 семей. С целью определения среднего размера семьи района было проведено 3%-ное выборочное обследование семей методом случайного бесповторного отбора. В результате получены следующие данные:

Размер семьи, чел.	0	1	2	3	4	5	6	7	8
Число семей	4	8	14	16	8	4	3	2	1

С вероятностью 0,997 определите пределы, в которых находится средний размер семьи в районе.

8.2. На заводе с числом рабочих 15 тыс. чел. В порядке механической выборки предполагается определить долю рабочих со стажем работы 20 лет и более. Какова должна быть численность выборки, чтобы с вероятностью 0,954 ошибка выборки не превышала 0,03, если на основе предыдущих обследований известно, что дисперсия равна 0,2?

8.3. С целью определения доли расходов на питание населения города А методом типической выборки был произведен 5%-ный отбор семей. Внутри типов производилась механическая выборка. Результаты представлены в таблице:

Выборка	Численность выборки	Доля расходов на питание, %
Одинокие	30	33
Семейные	70	45

С вероятностью 0,954 определите предел, в котором находится доля расходов на питание семей города А.

9.1. При обследовании 500 образцов изделий, отобранных из партии готовой продукции в случайном порядке, 40 оказались нестандартными. С вероятностью 0,954 определите пределы, в которых находится доля нестандартной продукции, выпускаемой заводом.

9.2. В городе А проживает 20 тыс. семей. В порядке случайной бесповторной выборки предполагается установить долю семей с доходом на душу 500 у.е. и менее. Какова должна быть численность выборки, чтобы с вероятностью 0,997 ошибка выборки не превышала 0,03, если на основе предыдущих обследований известно, что дисперсия равна 0,16.

9.3. Для определения доли рабочих завода, не выполняющих норму выработки, была произведена 10%-ная типическая выборка рабочих с отбором числа рабочих пропорционально численности типических групп. Внутри типических групп применялся метод случайного бесповторного отбора. Результаты представлены в таблице:

Цех	Число рабочих в выборке	Доля рабочих, не выполняющих норму выработки, %
1	120	5
2	80	2

С вероятностью 0,954 определите пределы, в которых находится доля рабочих завода, не выполняющих норму выработки.

10.1. Научно-исследовательским институтом для изучения общественного мнения населения области о проведении определенных мероприятий в порядке случайного повторного отбора было опрошено 600 человек. Из числа опрошенных 360 человек одобрили мероприятия. С вероятностью 0,997 определите пределы, в которых находится доля лиц, одобрявших мероприятия.

10.2. В городе N с числом семей 10 тыс. предполагается методом случайного бесповторного отбора определить долю семей с детьми ясельного возраста. Какова должна быть численность выборки, чтобы с вероятностью 0,954 ошибка выборки не превышала 0,04, если на основе предыдущих обследований известно, что дисперсия равна 0,24?

10.3. С целью определения доли брака во всей партии изготовленных деталей была проведена 10%-ная типическая выборка с отбором единиц пропорционально численности типических групп. Внутри типических групп применялся метод механического отбора. Результаты представлены в таблице:

Тип станка	Выработка 1 станка, шт.	% брака по данным выборки
1	1500	2,0
2	2000	3,0
3	4000	1,5
4	5000	1,0
5	2500	1,8

С вероятностью 0,997 определите пределы, в которых находится доля брака во всей партии деталей, изготовленных на всех станках.

Литература

- Ефимова, М.Р. Общая теория статистики: учебник / М.Р. Ефимова, Е.В. Петрова, В.Н. Румянцев. – М.: ИНФРА-М, 1998. – 416 с.
- Статистика: Курс лекций / Л.П. Харченко, В.Г. Долженкова, В.Г. Ионин [и др.]. – Новосибирск: Изд-во НГАЭиУ, М.:ИНФРА-М, 1997. – 310 с.
- Теория статистики: учебник / Под ред. Р.А.Шмойловой. – М.: Финансы и статистика, 1998. – 576 с.
- Практикум по теории статистики: учеб. пособие / Под ред. проф. Р.А. Шмойловой. – М.: Финансы и статистика, 1999. – 416 с.
- Громька, Г.Л. Общая теория статистики: Практикум. – М.:ИНФРА-М, 1999. – 139 с.
- Ефимова, М.Р. Практикум по общей теории статистики: учеб. пособие. – 2-е изд., перераб. и доп. / М.Р. Ефимова, О.И. Ганченко, Е.В. Петрова. – М. Финансы и статистика, 2004. – 336 с.

ПРИЛОЖЕНИЯ

Приложение 1

$$\text{Значение функции } \varphi(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}$$

x	Сотые доли x									
	0	1	2	3	4	5	6	7	8	9
0,0	0,3989	3989	3989	3988	3986	3984	3982	3980	3977	3973
0,1	3970	3965	3961	3956	3951	3945	3939	3932	3925	3918
0,2	3910	3902	3894	3885	3876	3867	3857	3847	3836	3825
0,3	3814	3802	3790	3778	3765	3752	3739	3726	3712	3697
0,4	3683	3668	3653	3637	3721	3605	3588	3572	3555	3538
0,5	3521	3503	3485	3467	3448	3429	3411	3391	3372	3352
0,6	3332	3312	3292	3271	3251	3230	3209	3187	3166	3144
0,7	3123	3101	3079	3056	3034	3011	2989	2966	2943	2920
0,8	2897	2874	2850	2827	2803	2780	2756	2732	2709	2685
0,9	2661	2637	2613	2589	2565	2541	2516	2492	2468	2444
1,0	2420	2396	2371	2347	2323	2299	2275	2251	2227	2203
1,1	2179	2155	2131	2107	2083	2059	2036	2012	1989	1965
1,2	1942	1919	1895	1872	1849	1827	1804	1781	1759	1736
1,3	1714	1692	1669	1647	1626	1604	1582	1561	1540	1518
1,4	1497	1476	1456	1435	1415	1394	1374	1354	1334	1315
1,5	1295	1276	1257	1238	1219	1200	1181	1163	1145	1127
1,6	1109	1092	1074	1057	1040	1023	1006	0989	0973	0957
1,7	0941	0925	0909	0893	0878	0863	0848	0833	0818	0804
1,8	0790	0775	0761	0748	0734	0721	0707	0694	0681	0669
1,9	0656	0644	0632	0620	0608	0596	0584	0573	0562	0551
x	Десятые доли x									
	0	1	2	3	4	5	6	7	8	9
2,	0540	0440	0355	0283	0224	0175	0136	0104	0079	0060
3,	0044	0033	0024	0017	0012	0009	0006	0004	0030	0020
4,	0001									

$$\text{Значение функции } \Phi_0(x) = \frac{1}{\sqrt{2\pi}} \int_0^x e^{-\frac{t^2}{2}} dt$$

x	Сотые доли x									
	0	1	2	3	4	5	6	7	8	9
0,0	0,0000	0040	0080	0112	0160	0199	0239	0279	0319	0359
0,1	0398	0438	0478	0517	0557	0596	0636	0675	0714	0754
0,2	0793	0832	0871	0910	0948	0987	1026	1064	1103	1141
0,3	1179	1217	1255	1293	1331	1368	1406	1443	1480	1517
0,4	1554	1591	1628	1664	1700	1736	1772	1808	1844	1879
0,5	1915	1950	1985	2019	2054	2088	2123	2157	2190	2224
0,6	2258	2291	2324	2357	2389	2422	2454	2486	2518	2549
0,7	2580	2612	2642	2673	2704	2734	2764	2794	2823	2852
0,8	2881	2910	2939	2967	2996	3023	3051	3079	3106	3133
0,9	3159	3186	3212	3238	3264	3289	3315	3340	3365	3389
1,0	3413	3438	3461	3485	3508	3531	3553	3577	3599	3621
1,1	3643	3665	3686	3708	3729	3749	3770	3790	3810	3830
1,2	3849	3869	3888	3907	3925	3944	3962	3980	3997	4015
1,3	4032	4049	4066	4082	4099	4115	4131	4147	4162	4177
1,4	4192	4207	4222	4236	4251	4265	4279	4292	4306	4319
1,5	4332	4345	4357	4370	4382	4394	4406	4418	4430	4441
1,6	4452	4463	4474	4485	4495	4505	4515	4525	4535	4545
1,7	4554	4564	4573	4582	4591	4599	4608	4616	4625	4633
1,8	4641	4649	4656	4664	4671	4678	4686	4693	4700	4706
1,9	4713	4719	4726	4732	4738	4744	4750	4756	4762	4767
x	Десятые доли x									
	0	1	2	3	4	5	6	7	8	9
2,	4773	4821	4861	4893	4918	4938	4953	4965	4974	4961
3,	4987	4990	4993	4995	4997	4998	4998	4999	4999	5000 ¹

Квантили t -распределения Стьюдента (k - число степеней свободы)

k	Уровень значимости α (двусторонняя критическая область)					
	0,10	0,05	0,02	0,01	0,002	0,001
1	6,31	12,7	31,82	63,7	318,3	637,0
2	2,92	4,30	6,97	9,92	22,33	31,6
3	2,35	3,18	4,54	5,84	10,22	12,9
4	2,13	2,78	3,75	4,60	7,17	8,61
5	2,01	2,57	3,37	4,03	5,89	6,86
6	1,94	2,45	3,14	3,71	5,21	5,96
7	1,89	2,36	3,00	3,50	4,79	5,40
8	1,86	2,31	2,90	3,36	4,50	5,04
9	1,83	2,26	2,82	3,25	4,30	3,78
10	1,81	2,23	2,76	3,17	4,14	4,59
11	1,80	2,20	2,72	3,11	4,03	4,44
12	1,78	2,18	2,68	3,05	3,93	4,32
13	1,77	2,16	2,65	3,01	3,85	4,22
14	1,76	2,14	2,62	2,98	3,79	4,14
15	1,75	2,13	2,60	2,95	3,73	4,07
16	1,75	2,12	2,58	2,92	3,69	4,01
17	1,74	2,11	2,57	2,90	3,65	3,96
18	1,73	2,10	2,55	2,88	3,61	3,92
19	1,73	2,09	2,54	2,86	3,58	3,88
20	1,73	2,09	2,53	2,85	3,55	3,85
21	1,72	2,08	2,52	2,83	3,53	3,82
22	1,72	2,07	2,51	2,82	3,51	3,79
23	1,71	2,07	2,50	2,81	3,49	3,77
24	1,71	2,06	2,49	2,80	3,47	3,74
25	1,71	2,06	2,49	2,79	3,45	3,72
26	1,71	2,05	2,48	2,78	3,44	3,71
27	1,71	2,05	2,47	2,77	3,42	3,69
28	1,70	2,05	2,46	2,76	3,40	3,66
29	1,70	2,05	2,46	2,76	3,40	3,66
30	1,70	2,04	2,46	2,75	3,39	3,65
40	1,68	2,02	2,42	2,70	3,31	3,55
60	1,67	2,00	2,39	2,66	3,23	3,46
120	1,66	1,98	2,36	2,62	3,17	3,37
∞	1,64	1,96	2,33	2,58	2,99	3,29

Содержание

Глава 1. Статистическая сводка и группировка. Статистические таблицы	
Графическое изображение статистических данных	3
1.1 Решение типовых задач	5
1.2 Задания для самостоятельной работы	9
1.3 Задания для индивидуальной работы	10
Глава 2. Статистические величины	12
2.1 Понятие абсолютной и относительной величины	12
2.2 Средняя величина. Виды средних величин	15
2.2.1 Решение типовых задач	16
2.3 Структурные средние	19
2.3.1 Решение типовых задач	20
2.3.2 Задания для самостоятельной работы	20
2.3.3 Задания для индивидуальной работы	21
Глава 3. Показатели вариации признака	26
3.1 Решение типовых задач	26
3.2 Задания для самостоятельной работы	27
3.3 Задания для индивидуальной работы	28
Глава 4. Сложение дисперсий изучаемого признака	28
4.1 Решение типовых задач	30
4.2 Задания для самостоятельной работы	34
4.3 Задания для индивидуальной работы	34
Глава 5. Выборочное наблюдение	35
5.1 Решение типовых задач	35
5.2 Задания для самостоятельной работы	37
5.3 Задания для индивидуальной работы	40
Литература	43
Приложения	44

УЧЕБНОЕ ИЗДАНИЕ

Составители: Гусева Светлана Тадеушевна,
Журавель Мария Григорьевна,
Золотухина Лада Станиславовна,
Шамовская Галина Владимировна,
Кузьмина Елена Викторовна

Практикум
по общей теории статистики
для студентов экономических специальностей

ЧАСТЬ I

Ответственный за выпуск: Гусева С.Т.

Редактор: Строкач Т.В.

Компьютерная вёрстка: Боровикова Е.А.

Корректор: Никитчик Е.В.

Подписано к печати 23.09.2010 г. Формат 60*84 ¹/₁₆. Гарнитура Arial Narrow.
Бумага «Снегурочка». Усл. п. л. 2,79. Уч. изд. 3,0. Заказ № 938. Тираж 100 экз.

Отпечатано на ризографе учреждения образования
«Брестский государственный технический университет».
224017, Брест, ул. Московская, 267.