

Кашина И. А., Кашин В. К., Нечаев Д. Ю., Чекмарев Ю. В.

Автоматизация процессов обработки информации в статистике

**Допущено
учебно-методическим
объединением**

Учебное пособие

ФМК
ИЗДАТЕЛЬСТВО

И .А. Кашина, В. К. Кашин,
Д. Ю. Нечаев, Ю. В. Чекмарев

АВТОМАТИЗАЦИЯ ПРОЦЕССОВ ОБРАБОТКИ ИНФОРМАЦИИ В СТАТИСТИКЕ

2-е издание, электронное

*Допущено учебно-методическим объединением
по образованию в области коммерции и маркетинга
в качестве учебного пособия для студентов
высших учебных заведений, обучающихся
по специальностям
080301 (351300) – «Коммерция (Торговое дело)»,
080111 (061500) – «Маркетинг»
и 032401 (350700) – «Реклама»*



Москва, 2023

УДК 681.142.2

ББК 32.97

К31

Кашина, И. А.

К31 Автоматизация процессов обработки информации в статистике : учебное пособие / И. А. Кашина, В. К. Кашин, Д. Ю. Нечаев, Ю. В. Чекмарев. — 2-е изд., эл. — 1 файл pdf : 200 с. — Москва : ДМК Пресс, 2023. — Систем. требования: Adobe Reader XI либо Adobe Digital Editions 4.5 ; экран 10". — Текст : электронный.

ISBN 978-5-89818-471-1

Управление предприятием требует решения задач, связанных с различными производственными вопросами, такими как: повышение эффективности управления персоналом, сбыт готовой продукции, оптимизация поставок, планирование и прогнозирование деятельности предприятия в целом и его подразделений. Определение базовых понятий статистического анализа, статистические методы и инструментарий решения таких задач изложен в данном пособии.

Это пособие предназначено главным образом для специалистов, решающих практические задачи, связанные с обработкой и анализом статистической информации с использованием пакета Microsoft Excel.

УДК 681.142.2

ББК 32.97

Электронное издание на основе печатного издания: Автоматизация процессов обработки информации в статистике : учебное пособие / И. А. Кашина, В. К. Кашин, Д. Ю. Нечаев, Ю. В. Чекмарев. — Москва : ДМК Пресс, 2012. — 200 с. — ISBN 978-5-94074-499-3. — Текст : непосредственный.

Все права защищены. Любая часть этой книги не может быть воспроизведена в какой бы то ни было форме и какими бы то ни было средствами без письменного разрешения владельцев авторских прав.

Материал, изложенный в данной книге, многократно проверен. Но поскольку вероятность технических ошибок все равно существует, издательство не может гарантировать абсолютную точность и правильность приводимых сведений. В связи с этим издательство не несет ответственности за возможные ошибки, связанные с использованием книги.

В соответствии со ст. 1299 и 1301 ГК РФ при устранении ограничений, установленных техническими средствами защиты авторских прав, правообладатель вправе требовать от нарушителя возмещения убытков или выплаты компенсации.

ISBN 978-5-89818-471-1

© Кашина И. А., Кашин В. К.,
Нечаев Д. Ю., Чекмарев Ю. В., 2009
© Оформление, ДМК Пресс, 2009

Содержание

Введение	6
1. СТАТИСТИЧЕСКИЕ НАБЛЮДЕНИЯ	7
1.1. Определение статистического наблюдения	7
2. Статистическая сводка и группировка	13
2.1 Задачи сводки и ее содержание	13
2.2 Виды статистических группировок	14
2.3. Принципы построения статистических группировок и классификаций.....	15
2.4. Сравнимость статистических группировок. Вторичная группировка	23
2.5. Статистическая таблица и ее элементы.....	24
2.6. Статистические графики и их классификация	25
2.7. Построение диаграмм в MS Excel	28
2.8. Построение рядов распределения в MS Excel.....	29
Тесты и задачи для самостоятельной работы	36
3. Абсолютные и относительные статистические показатели	44
3.1. Классификация статистических показателей	44
3.2. Абсолютные показатели.....	47
3.3. Относительные показатели	49
Тесты и задачи для самостоятельной работы	54
4. Средние показатели и показатели вариации	64
4.1. Сущность средних показателей	64
4.2. Средняя арифметическая	65
4.3. Другие виды средних	69
4.4. Структурные средние	72

4.5. Показатели вариации	73
4.6. Описательная статистика в Microsoft Excel.....	81
4.7. Установка и работа с инструментом Ms Excel «Пакета анализа».....	86
Тесты и задачи для самостоятельной работы	90
5. Статистическое изучение динамики социально-экономических явлений	106
5.1. Понятие и классификации рядов динамики	106
5.2. Сопоставимость уровней и смыкание рядов динамики	108
5.3. Показатели изменения уровней ряда динамики	111
5.4. Методы анализа основной тенденции (тренда) в рядах динамики.....	116
Тесты и задачи для самостоятельной работы	122
6. Индексы	132
6.1. Общие понятия об индексах	132
6.2. Агрегатные индексы	135
6.3. Сводные индексы в средней арифметической и средней гармонической формах.....	138
Тесты и задачи для самостоятельной работы	141
7. Выборочное наблюдение	155
7.1. Выборочное наблюдение как важнейший источник статистической информации	155
7.2. Основные способы формирования выборочной совокупности	159
Тесты и задачи для самостоятельной работы	166
8. Статистическое изучение взаимосвязи социально-экономических явлений	173
8.1. Причинность, регрессия, корреляция.....	173
8.2. Парная регрессия на основе метода наименьших квадратов и метода группировок	177

8.3. Собственно-корреляционные параметрические методы изучения связи	179
8.4. Методы изучения связи качественных признаков.....	182
8.5. Технология решения задач корреляционного и регрессионного анализа с помощью надстроек MS Excel	183
Тесты и задачи для самостоятельной работы	189
Толковый словарь	197

Введение

Управление предприятием требует решения задач, связанных с различными производственными вопросами, такими как: повышение эффективности управления персоналом, сбыт готовой продукции, оптимизация поставок, планирование и прогнозирование деятельности предприятия в целом и его подразделений. Определение базовых понятий статистического анализа, статистические методы и инструментарий решения таких задач изложен в данном пособии.

Microsoft Excel – это приложение, которое входит в состав Microsoft Office и на сегодняшний день установлено практически на каждом компьютере. Обладая мощными вычислительными средствами, Microsoft Excel позволяет решать достаточно сложные и разнообразные задачи. В книге рассмотрен набор инструментов, который входит в пакет анализа данных для решения статистических задач различного рода.

Описание каждого метода включает изложение теоретического материала и сопровождается большим количеством практических примеров, решением задач, выполнением тестовых заданий, которые приведены в конце каждой главы данного пособия. Для каждого примера приведена последовательность действий, соблюдение которой позволяет правильно использовать выбранный метод с учетом поставленной задачи и особенностей обрабатываемых данных. Такой подход позволит читателю понять, каким образом задачи статистического анализа и создания прогнозов решаются в среде Microsoft Excel, как правильно интерпретировать данные и как на основании полученных результатов анализа делать обоснованные выводы.

Это пособие предназначено главным образом для специалистов, решающих практические задачи, связанные с обработкой и анализом статистической информации с использованием пакета Microsoft Excel версии 2007.



СТАТИСТИЧЕСКИЕ НАБЛЮДЕНИЯ

1.1. Определение статистического наблюдения

Статистическое наблюдение представляет собой планомерный, научно организованный и, как правило, систематический сбор данных о явлениях и процессах общественной жизни путем регистрации, заранее намеченных существенных признаков с целью получения в дальнейшем характеристик этих явлений и процессов.

Например, при переписи населения специальные работники, привлеченные к ее проведению, по поручению органов статистики записывают для каждого жителя страны сведения о его поле, возрасте, семейном положении, образовании и др., а затем на основе этих сведений статистические органы определяют численность населения, его возрастную структуру, размещение по территории страны и многие другие показатели (табл. 1.1–1.2).

Основные формы, виды и способы статистического наблюдения

Статистическое наблюдение можно классифицировать по различным признакам. Один из вариантов такой классификации представлен на рис. 1.1.

Отчетностью называют такую организационную форму статистического наблюдения, при которой сведения поступают в статистические органы от предприятий, учреждений и организаций в виде обязательных отчетов об их деятельности.

Таблица 1.1.
Численность населения Российской Федерации
в 1897–2002 г.г.

Годы	Все население, тыс.чел	В том числе	
		городское	сельское
1897	67473	9894	57579
1826	92735	16455	76280
1939	108377	36296	72081
1959	117240	61143	56097
1970	129941	80631	49310
1979	137410	94942	42468
1989	147022	107959	39063
2002	145164	106427	38737

Таблица 1.2
Распределение населения РФ по возрасту и полу по данным
переписи населения 2002 г.

Тысяч человек				Доля (%) в общей численности населения		
Все население	Мужчины и женщины	Мужчины	Женщины	Мужчины и женщины	Мужчины	Женщины
	1451644	67604	77560	100	100	100
в том числе в возрасте, лет:						
0-4	6399	3276	3123	4,4	4,8	4,0
5-9	6940	3548	3392	4,8	5,2	4,4
10-14	10407	5313	5094	7,2	7,9	6,6
15-19	12800	6504	6296	8,8	9,6	8,1
20-24	11466	5783	5683	7,9	8,6	7,3
25-29	10613	5314	5299	7,3	7,9	6,8
30-34	9836	4914	4922	6,8	7,3	6,4
35-39	10216	5025	5191	7,0	7,4	6,7
40-44	12546	6084	6462	8,6	9,0	8,3
45-49	11606	5494	6112	8,0	8,1	7,9
50-54	10071	4642	5429	6,9	6,9	7,0

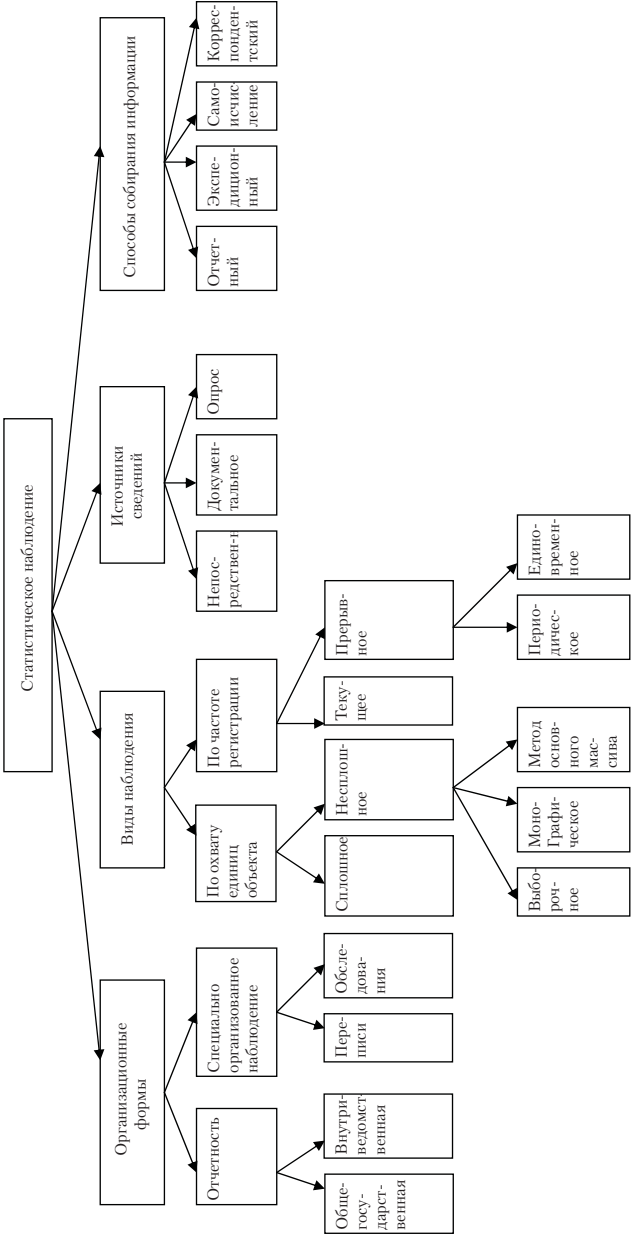


Рис. 1.1. Классификация форм, видов и способов статистических наблюдений

55-59	5347	2366	2981	3,7	3,5	3,8
60-64	7983	3251	4732	5,5	4,8	6,1
65-69	6344	2444	3900	4,4	3,6	5,0
70 и старше	12590	3646	8944	8,7	5,4	11,6

Специально организованное статистическое наблюдение представляет собой наблюдение, организуемое с какой-либо особой целью для получения данных, которые в силу тех или иных причин не собираются посредством отчетности, или для проверки, уточнения данных отчетности.

Примерами такой формы статистического наблюдения могут служить уже упоминавшиеся переписи населения, а также бюджетные обследования домашних хозяйств, опросы общественного мнения и т. п.

Сплошным называется такое наблюдение, при котором обследованию подвергаются все без исключения единицы изучаемой совокупности (объекта наблюдения).

Примером такого наблюдения являются переписи, при которых по основной программе обследованию подлежит все без исключения население страны.

Несплошное - это такое наблюдение, при котором обследованию подвергаются не все единицы совокупности, а только часть их.

В статистической практике применяется несколько видов несплошного наблюдения. Главными из них являются выборочное наблюдение, монографическое обследование и метод основного массива.

Выборочным называют наблюдение, основанное на принципе случайного отбора тех единиц изучаемой совокупности, которые должны быть подвергнуты наблюдению.

Выборочное наблюдение при правильной его организации и проведении дает достаточно достоверные данные для характеристики изучаемой совокупности в целом. Во многих случаях им вполне можно заменить сплошной учет. В условиях рыночной экономики сфера применения выборочного наблюдения постоянно расширяется.

Монографическое обследование представляет собой детальное, глубокое изучение и описание отдельных, характерных в каком-либо отношении единиц совокупности.

Монографическое исследование проводится с целью выявления имеющихся или намечающихся тенденций в развитии явления, для выявления имеющихся резервов, изучения опыта отдельных субъектов рыночной экономики и т.п.

Метод основного массива заключается в том, что обследованию подвергаются наиболее крупные единицы, которые вместе взятые

имеют преобладающий удельный вес в совокупности по основному для данного исследования признаку (признакам).

Единицы совокупности, обладающие незначительной величиной изучаемого признака, обследованию не подвергаются. Например, цены на продовольственных рынках могут регистрироваться лишь в крупных городах, где проживает большая часть населения России.

По частоте (срокам регистрации) наблюдение может быть непрерывным (текущим) и прерывным. Последнее, в свою очередь, подразделяется на периодическое и единовременное.

Текущим называют такое наблюдение, которое ведется непрерывно, и регистрация фактов производится по мере их свершения.

Пример такого наблюдения - регистрация актов гражданского состояния: рождений, смертей, браков, разводов.

Периодическое - это наблюдение, которое повторяется через определенные, равные промежутки времени.

Таковым является, в частности, ежеквартальное представление финансовых отчетов в налоговые службы.

Единовременным называется такое наблюдение, которое проводится по мере необходимости, время от времени, без соблюдения строгой периодичности или вообще проводится один раз и больше не повторяется.

Примером такого рода наблюдения может служить учет товарных остатков и денежной наличности на момент денежной реформы.

По источнику сведений различают непосредственное наблюдение, документальное наблюдение и опрос.

Непосредственным называют такое наблюдение, при котором сами регистраторы путем непосредственного замера, взвешивания или подсчета устанавливают факт и на этом основании производят запись в формуляре наблюдения.

Таковы, например, инвентаризация имущества, снятие остатков товаров в магазине и др.

Документальное наблюдение предполагает запись ответов на вопросы формуляра на основании соответствующих документов.

Примером такого наблюдения является сбор данных об успеваемости студентов вуза на основе зачетно - экзаменационных ведомостей.

Опрос - это наблюдение, при котором ответы на вопросы формуляра наблюдения записываются со слов опрашиваемого.

Так проводятся, в частности, переписи населения. Основанием записи сведений о возрасте, семейном положении, образовании и т.д. служат ответы опрашиваемого.

В статистике применяются следующие способы собирания сведений: отчетный, экспедиционный, самоисчисление, анкетный, корреспондентский.

Сущность **отчетного** способа заключается, как уже отмечалось, в представлении предприятиями, учреждениями и организациями статистических отчетов о своей деятельности в строго обязательном порядке.

Экспедиционный способ наблюдения заключается в том, что специально привлеченные и обученные работники посещают каждую единицу наблюдения и сами заполняют формуляр наблюдения.

Этим способом собираются сведения при переписях населения.

При способе **самоисчисления (саморегистрации)** формуляры заполняют сами опрашиваемые.

Обязанность специально привлеченных для получения информации сотрудников состоит в раздаче формуляров опрашиваемым, инструктаже их, сборе заполненных формуляров и проверке правильности их заполнения. Способ саморегистрации применяется органами статистики, например, для изучения так называемой маятниковой миграции - передвижения населения от места жительства до места работы и обратно.

Анкетный способ – это сбор статистических данных с помощью специальных вопросников, рассылаемых определенному кругу лиц или публикуемых в периодической печати.

В современных условиях данный способ собирания сведений применяется очень широко, особенно в различных социологических исследованиях.

Сущность **корреспондентского** способа наблюдения заключается в том, что статистические органы договариваются с определенными лицами, которые берут на себя обязательство вести наблюдение за какими-либо явлениями, процессами и в установленные сроки сообщать результаты наблюдений статистическим органам.

Таким образом, проводятся, в частности, экспертные оценки по конкретным вопросам социально - экономического развития страны.

При статистическом наблюдении необходимо, прежде всего, определить его объект и единицу.

Объектом статистического наблюдения называется та совокупность, о которой должны быть собраны нужные сведения.

Объектом наблюдения может быть, например, совокупность жителей страны, промышленных предприятий, коммерческих банков, высших учебных заведений и т.п.

Единицей наблюдения называют тот составной элемент объекта наблюдения, который является носителем признаков, подлежащих регистрации.

Единица наблюдения, как и объект в целом, обладают множеством различных признаков.



СТАТИСТИЧЕСКАЯ СВОДКА И ГРУППИРОВКА

Важнейшим этапом исследования социально-экономических явлений и процессов является систематизация первичных данных и получение на этой основе сводной характеристики всего объекта при помощи обобщающих показателей, что достигается путем сводки и группировки первичного статистического материала.

2.1. Задачи сводки и ее содержание

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

Статистическая сводка представляет собой вторую стадию статистического исследования. Она состоит в том, что первичные материалы сводятся вместе, образуя статистические совокупности, которые характеризуются итоговыми обобщающими показателями в виде абсолютных, относительных и средних величин.

Сводка в виде простого подытоживания отдельных фактов не дает возможности проводить глубокий статистический анализ. Для того чтобы выявить все закономерности, содержащиеся в первичном статистическом материале, при сводке этого материала необходимо применять статистические группировки. Обычно в процессе сводки статистический материал упорядочивается, систематизируется, делится на группы по существенным признакам, и на этой основе образуются статистические совокупности. Таким образом, статистической группировкой называется разделение единиц совокупности на

группы, по существенным варьирующим признакам.

По глубине и точности обработки материала различают сводку простую и сложную.

Простая сводка – это операция подсчета общих итогов по совокупности единиц наблюдения.

Сложная сводка – это комплекс операций, включающих группировку единиц наблюдения, подсчет итогов по каждой группе и по всему объекту и представление результатов в виде статистических таблиц.

Проведение сводки необходимо включает следующие этапы:

- выбор группировочного признака;
- определение порядка формирования групп;
- разработка системы статистических показателей для характеристики групп и объекта в целом;
- разработка макетов статистических таблиц для представления результатов сводки.

По форме обработки материала сводка бывает:

- централизованная, когда весь первичный материал поступает в одну организацию, подвергается в ней обработке от начала до конца;
- децентрализованная, когда отчеты предприятий сводятся статистическими органами субъектов РФ, а полученные итоги поступают в Госкомстат РФ и там определяются итоговые показатели в целом по народному хозяйству страны.

По технике выполнения статистическая сводка бывает механизированная (с использованием электронно-вычислительной техники) и ручная.

2.2. Виды статистических группировок

Группировкой называется расчленение единиц изучаемой совокупности по определенным существенным для них признакам.

Группировочным признаком называется признак, по которому проводится разбиение единиц совокупности на отдельные группы. От правильного выбора группировочного признака зависят выводы статистического исследования. В качестве основания группировки необходимо использовать существенные, теоретически обоснованные признаки.

В основание группировки могут быть положены как количественные, так и качественные (атрибутивные) признаки.

Атрибутивные признаки не имеют количественного варьирования. Это, например, род занятий, образование, пол, форма собствен-

ности предприятия, отрасль производства и т.д.

Количественные признаки – это такие признаки, у которых варьирование проявляется в изменении количественного значения признака у отдельных единиц совокупности, т.е. имеют числовое выражение (объем торгов, возраст человека, доход семьи и т. д.),

Группировки делятся на три основных вида: **типологические, структурные и аналитические.**

Типологическими называются группировки, в которых вся совокупность делится на качественно однородные совокупности, т.е. выделяются социально-экономические типы. Например, совокупность работников какого-либо предприятия можно разделить на группы по уровню образования или по размеру заработной платы и др.

Структурные группировки – это группировки, которые применяются для изучения состава совокупности (структуры) по тем или иным признакам. С помощью таких группировок можно изучать, например, состав рабочих по стажу работы, профессиям, возрасту.

Аналитические группировки – это такие группировки, которые применяются для изучения взаимосвязанного изменения варьирующих признаков в пределах той или иной совокупности.

Группировки могут быть простыми и комбинационными.

Простая группировка имеет место, когда единицы совокупности группируются по одному признаку.

Комбинационной называется группировка, в которой расчленение совокупности на группы производится по двум и более признакам, взятым в сочетании (комбинации).

Сначала группы формируются по одному признаку, затем группы делятся на подгруппы по другому признаку, а эти в свою очередь делятся по третьему и так далее. Таким образом, комбинационные группировки дают возможность изучить единицы совокупности одновременно по нескольким признакам.

При построении комбинационной группировки возникает вопрос о последовательности разбиения единиц объекта по признакам. Как правило, рекомендуется сначала производить группировку по атрибутивным признакам, значения которых имеют ярко выраженные качественные различия.

2.3. Принципы построения статистических группировок и классификаций

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

Число групп зависит от задач исследования и вида показателя, положенного в основание группировки, объема совокупности, степени вариации признака. Например, группировка предприятий по формам собственности учитывает муниципальную, федеральную и собственность субъектов федерации.

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

При небольшом объеме совокупности не следует образовывать большого количества групп, так как группы будут включать недостаточное число единиц объекта. Поэтому показатели, рассчитанные для таких групп, не будут представительными и не позволят получить адекватную характеристику исследуемого явления.

Часто группировка по количественному признаку имеет задачу отразить распределение единиц совокупности по этому признаку. В этом случае количество групп зависит, в первую очередь, от степени колеблемости группировочного признака: чем больше его колеблемость, тем больше можно образовать групп. Чем больше групп, тем точнее будет воспроизведен характер исследуемого объекта. Однако, слишком большое число групп затрудняет выявление закономерностей при исследовании социально-экономических явлений и процессов. Поэтому в каждом конкретном случае при определении числа групп следует исходить не только из степени колеблемости признака, но и из особенностей объекта и цели исследования.

Определение числа групп можно осуществить и математическим путем с использованием формулы Стерджесса:

$$n = 1 + 3,322 \times \lg N, \quad (2.1)$$

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

Значения n , подсчитанные по этой формуле, представлены в следующей таблице:

N	15-24	25-44	45-89	90-179	180-359	360-719	720-1439
n	5	6	7	8	9	10	11

Согласно этой формуле выбор числа групп зависит от объема совокупности.

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

Когда определено число групп, то следует определить интервалы группировки.

Интервал – это значения варьирующего признака, лежащие в определенных границах. Каждый интервал имеет свою величину, верхнюю и нижнюю границы или хотя бы одну из них.

Нижней границей интервала называется наименьшее значение признака в интервале, а **верхней границей** – наибольшее значение признака в интервале. Величина интервала представляет собой разность между верхней и нижней границами.

Интервалы группировки в зависимости от их величины бывают: равные и неравные. Последние делятся на прогрессивно возрастающие, прогрессивно убывающие, произвольные и специализированные.

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

Величина равного интервала определяется по следующей формуле:

$$i = \frac{X_{\max} - X_{\min}}{n}, \quad (2.1)$$

где $X_{\max}$, $X_{\min}$ – максимальное и минимальное значения признака в совокупности; n – число групп.

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

Полученную по формуле (2.2) величину округляют и она будет являться шагом интервала.

Существуют следующие правила определения шага интервала.

Если величина интервала, рассчитанная по формуле (2.2) представляет собой величину, которая имеет один знак до запятой (например: 0,88; 1,585; 4,8), то в этом случае полученные значения целесообразно округлить до десятых и их использовать в качестве шага интервала. В приведенном выше примере это будут соответственно значения: 0,9; 1,6; 4,7.

Если рассчитанная величина интервала имеет две значащие цифры до запятой и несколько после запятой (например 15,985), то это значение необходимо округлить до целого числа (до 16).

В случае, когда рассчитанная величина интервала представляет собой трехзначное, четырехзначное и так далее число, то эту величину следует округлить до ближайшего числа, кратного 100 или 50. Например, 557 следует округлить до 550 или до 600.

Если размах вариации признака в совокупности велик и значения признака варьируют неравномерно, то надо использовать группировку с неравными интервалами.

Неравные интервалы могут быть прогрессивно возрастающие или убывающие в арифметической или геометрической прогрессии. Величина интервалов, изменяющихся в арифметической и геометрической прогрессии определяются следующим образом:

$$h_{i+1} = h_i + a, \quad (2.3)$$

а в геометрической прогрессии:

$$h_{i+1} = h_i \times q,$$

где a – константа: для прогрессивно возрастающих интервалов имеет знак «+», и знак «-» – при прогрессивно – убывающих;

q – константа: больше «1» – для прогрессивно-возрастающих и меньше «1» – в другом случае.

Применение неравных интервалов обусловлено тем, что в первых группах небольшая разница в показателях имеет большое значение, а в последних группах эта разница не существенна.

Например, при построении группировки предприятий отрасли по показателю численности промышленно-производственного персонала, который варьирует от 200 человек до 2000 человек, нецелесообразно рассматривать равные интервалы, т. к. учитываются как малые, так и крупнейшие предприятия отрасли. Поэтому следует образовывать неравные интервалы: 200–500, 500–1100, 1100–2000, т. е. величина каждого последующего интервала больше предыдущего на 300 человек и увеличивается в арифметической прогрессии.

Интервалы группировок могут быть закрытыми и открытыми.

Закрытыми называются интервалы, у которых имеются верхняя и нижняя границы.

Открытые – это те интервалы, у которых указана только одна граница: верхняя – у первого, нижняя – у последнего. Например, группы предприятий по числу работающих в них сотрудников (чел.): до 200, 200–300, 300–400, 400 и более.

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

Если основанием группировки служит непрерывный признак (например, группы строительных фирм по объему работ (млн. руб.): 1200–1400, 1400–1600, 1600–1800, 1800–2000), то одно и то же значение признака выступает и верхней и нижней границами двух смежных интервалов. В данном случае объем работ 1400 млн. руб. составляет верхнюю границу первого интервала и нижнюю границу второго, 1600 млн. руб. – соответственно второго и третьего и т. д., то

есть верхняя граница i -го интервала равна нижней границе $(i+1)$ -го интервала.

При таком обозначении границ может возникнуть вопрос, в какую группу включать единицы объекта, значения признака у которых совпадают с границами интервалов. Например, во вторую или третью группу должна войти строительная фирма с объемом работ 1600 млн. рублей. Если верхняя граница формируется по принципу «исключительно», то фирма должна быть отнесена к третьей группе, в противном случае – ко второй. Для того, чтобы правильно отнести к той или иной группе единицу объекта, значение признака которой совпадает с границами интервалов, можно использовать открытые интервалы (по нашему примеру группы строительных фирм по объему работ преобразуются в следующие: до 1400, 1400–1600, 1600–1800, 1800 и более). В данном случае, вопрос отнесения отдельных единиц совокупности, значения которых являются граничными, к той или иной группе решается на основе анализа последнего открытого интервала. Возможны два случая обозначения последнего открытого интервала: 1). 1800 млн. руб. и более; 2). более 1800 млн. руб. В первом случае, строительные фирмы с объемом работ 1600 млн. руб. попадут в третью группу; во втором случае – во вторую группу.

Если в основании группировки лежит дискретный признак, то нижняя граница i -го интервала равна верхней границе i -го интервала увеличенной на 1. Например группы предприятий по числу занятого персонала (чел) будут иметь вид: 100-150; 151-200; 201-250.

Пример:

По нижеследующим данным произвести группировку 40 шоколадных батончиков по их массе (в граммах).

52,8	51,0	50,0	51,5	53,1	51,0	49,5	54,5
50,5	49,5	49,3	49,5	50,0	52,0	51,0	51,2
49,5	50,0	49,7	53,7	52,0	50,2	50,1	50,5
50,5	50,2	49,5	51,9	51,0	49,5	51,1	52,5
51,0	51,0	52,0	51,2	49,8	50,4	48,7	51,6

Результаты представить в виде таблицы и проанализировать их. Построить гистограмму распределения.

Решение

Преобразуем вышеуказанный неупорядоченный ряд в ранжированный

48,7	49,3	49,5	49,5	49,5	49,5	49,5	49,5
49,7	49,8	50,0	50,0	50,0	50,1	50,2	50,2
50,4	50,5	50,5	50,5	51,0	51,0	51,0	51,0
51,0	51,0	51,1	51,2	51,2	51,5	51,6	51,9
52,0	52,0	52,2	52,5	52,8	53,1	53,7	54,5