



РОЗЕНЦВАЙГ А.К.
ИСАВНИН А.Г.

Статистика. Сводка и группировка данных статистического наблюдения

УЧЕБНО-МЕТОДИЧЕСКОЕ
ПОСОБИЕ

УДК 519.252

ББК 22.172

Печатается по решению учебно-методической комиссии экономического отделения Набережночелнинского института (филиала) федерального государственного автономного образовательного учреждения высшего образования «Казанский (Приволжский) федеральный университет», от «05» сентября 2023г. (протокол №5)

Рецензенты:

Доктор экономических наук, профессор А.Н. Макаров

Кандидат экономических наук, доцент М.Р. Хамидуллин

Розенцвайг А.К, Исавнин А.Г. Статистика. Сводка и группировка данных статистического наблюдения: учебно-методическое пособие / А.К Розенцвайг, А.Г. Исавнин – Набережные Челны: Изд-во Набережночелнинского института КФУ, 2024. – 27 с.

Учебно-методическое пособие содержит последовательное изложение базовых понятий теории сводки и группировки данных статистического наблюдения. Подробно изложены: этапы статистического исследования, сводка статистических данных, метод группировки, принципы построения статистических группировок, а также виды статистических группировок.

Учебно-методическое пособие предназначено для использования в учебном процессе студентами технических и экономических направлений дневной и заочной форм обучения.

© Розенцвайг А.К., Исавнин А.Г., 2024

© НЧИ КФУ, 2024

© Кафедра бизнес-информатики и математических методов в экономике, 2024 г.

1. ОСНОВНЫЕ ЭТАПЫ СТАТИСТИЧЕСКОГО ИССЛЕДОВАНИЯ.

Количественная характеристика социально-экономических процессов, учитывающая особенности их организации и места в системе общественного производства, невозможна без применения статистических методов. Использование любых способов и приемов статистического анализа предполагает наличие исчерпывающей и достоверной информации об изучаемых объектах. Исследование массовых социально-общественных явлений включает этапы:

- 1) сбор исходных статистических данных и их первичная обработка;
- 2) сведение и группировка результатов наблюдения в определенные совокупности;
- 3) обобщение и анализ сводных материалов второго этапа.

На первом этапе статистического исследования собираются первичные данные, на которые опирается статистическая информация, которая является фундаментом всего предстоящего анализа. Чтобы здание было прочным, добротной и качественной должна быть его основа. Если при сборе этих данных допущены методологические ошибки или материал оказался недоброкачественным, то это напрямую повлияет на правильность и достоверность результатов анализа и выводов, а также на принимаемые на их основе практические решения. Поэтому статистическое наблюдение является наиболее ответственным этапом и от начальной до завершающей стадии должно быть тщательно продуманным, правильно организованным и четко исполненным.

Статистическое наблюдение дает исходный материал для второго этапа исследования, началом которого служит сводка. Если при статистическом наблюдении получают сведения о каждой единице изучаемой совокупности,

которые характеризуют их с разных сторон, то данные сводки характеризуют всю статистическую совокупность в целом и выделяют ее характерные части. На этой стадии совокупность делится по признакам различия и объединяется по признакам сходства изученных единиц, затем подсчитываются суммарные показатели по отдельным группам и в целом по совокупности. С помощью метода группировок изучаемые социально-общественные явления делятся на важнейшие типы, характерные группы и подгруппы по существенным признакам. С помощью группировок выделяют и ограничивают качественно однородные в отношении изучаемых признаков совокупности, что является основой для проведения третьего этапа, состоящего в определении и применении обобщающих показателей.

На заключительном 3-ем этапе анализа, где рассчитывают относительные и средние величины, дают сводную оценку вариации признаков, характеризуют динамику явлений, рассчитывают показатели, характеризующие тесноту и форму статистических взаимосвязей признаков. С целью наиболее яркого и наглядного изложения цифрового статистического материала, его представляют в виде статистических таблиц и статистических графиков.

2. СТАТИСТИЧЕСКАЯ СВОДКА И ГРУППИРОВКА.

2.1. Сводка статистических данных.

Сводка представляет собой второй этап любого статистического исследования.

Первичный материал, собранный в процессе статистического наблюдения, нуждается в определенной обработке и сведении многочисленных разрозненных данных воедино.

Сводкой в статистике называется научно организованная (по заранее разработанной программе) обработка материалов наблюдения, которая включает в себя кроме обязательного контроля собранных данных

систематизацию, группировку, составление таблиц, получение итоговых и производных показателей (относительных и средних величин).

Целью сводки является получение обобщающих статистических показателей, отражающих сущность социально-экономических явлений в целом и наличие определенных статистических тенденций в виде сводных материалов.

Статистическая сводка выполняется по программе, которая должна разрабатываться еще до сбора данных, практически одновременно с составлением плана и программы статистического наблюдения. Программа сводки включает определение:

- 1) групп и подгрупп признаков;
- 2) системы показателей (для групп);
- 3) видов сводных таблиц.

Все эти вопросы решают не механически, а с учетом цели исследования и особенностей изучаемой статистической совокупности. Выделение тех или иных групп должно быть обоснованным экономической природой изучаемых социально-общественных явлений. Кроме итоговых и групповых показателей сводка создает основу для последующего анализа и выявления различного рода закономерностей. По технике или способу выполнения сводка может быть ручной либо механизированной. *Ручная сводка* применяется в основном для небольших массивов данных. Начинается она с шифровки статистических формуляров (карточек), которые затем определенным образом группируются, подсчитывается их число и другие показатели. При *механизированной сводке* большие объемы исходных данных сразу заносятся на машиночитаемые носители информации и полностью обрабатываются на компьютере.

2.2. Метод группировки.

Одним из основных и наиболее распространенных методов обработки и анализа первичных статистических данных является группировка. Это один из

самых сложных в методологическом плане этапов статистического исследования.

Позволяет представить в «сжатом» и обозримом виде большой объем первичных данных, собранных в ходе статистического наблюдения, и на этой основе судить о наличии статистических закономерностей, которые могут представлять характерные особенности изучаемого социально-общественного явления или процесса.

Среди бесконечного разнообразия явлений мы, как правило, улавливаем наличие некоторого конечного числа групп или типов. Например, лицо каждого человека неповторимо, и все-таки можно классифицировать лица по типам (скуластое, продолговатое, круглое и т.д.); предприятия образуют группы по формам собственности, характеру производимой продукции, размерам (крупные, средние, мелкие); государства делятся на группы по уровню экономического развития и т.д.

Группировка – это распределение множества единиц изучаемой совокупности на однородные по какому-либо признаку группы.

Группировка лежит в основе всей дальнейшей работы с собранной информацией. На основе группировки рассчитываются сводные показатели по группам, появляется возможность их сравнения, анализа причин различий между группами, изучения взаимосвязей между признаками.

От группировок следует отличать классификацию. **Классификацией** называется систематизированное распределение явлений и объектов на определенные группы, классы, разряды на основании их сходства и различия. В основу классификации кладется качественный признак. Нередки случаи, когда качественный признак имеет большое число разновидностей, и перечислить их все не представляется целесообразным. Например, профессии рабочих, номенклатура выпускаемой продукции, виды основных фондов и т. п. В таких случаях разрабатывают *классификацию разновидностей*, т. е. сходные по своим особенностям разновидности объединяются в группы (классы). Под классификацией обычно понимается более устойчивое разграничение единиц

наблюдения, чем при группировке. Классификации стандартны, они устанавливаются органами государственной и международной статистики. Классификации остаются неизменными в течение длительного времени. Однако, если появляются новые группы единиц, их классы, разряды, то в классификации вносятся соответствующие изменения и дополнения.

Классификация, определяя важнейшие признаки группировки единиц совокупности, является основой группировок. В классификации точно определены всевозможные группы и имеются подробные указатели, которые помогают отнести любую единицу объекта в ту или иную группу в каждом конкретном случае.

2.3 Принципы построения статистических группировок.

Метод группировки основывается на двух категориях – группировочном признаке и интервале группирования.

Группировочным называется признак, по которому проводится разбивка единиц совокупности на отдельные группы. Его часто называют основанием группировки. От правильного выбора группировочного признака зависят выводы статистического исследования.

В качестве основания группировки следует использовать существенные признаки. В каждом конкретном исследовании включение признака в состав группировочных должно быть теоретически обосновано. Только на базе теоретического анализа экономических законов развития исследуемого явления можно правильно определить состав признаков.

В основание группировки могут быть положены как количественные, так и качественные (атрибутивные) признаки. Первые имеют числовое выражение (объем торгов, курс доллара в рублях, возраст человека, денежный доход семьи и т. д.), а вторые отражают состояние единицы совокупности (пол человека, его национальность, семейное положение, отраслевая принадлежность предприятия, его форма собственности и организационно-правовая форма и т. д.).

Следует иметь в виду, что в ряде случаев классификация, которая представляется чисто качественной, в конечном итоге оказывается основанной на количественном признаке. Такова, например, классификация промышленных предприятий по отраслям. Поскольку одно и то же предприятие выпускает продукцию разных видов, статистика решает этот вопрос по количественному преобладанию того или иного вида продукции.

После определения основания группировки, следует решить вопрос о количестве групп, на которые нужно разбить исследуемую совокупность.

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

Если группировка строится по *атрибутивному признаку*, то групп будет столько, сколько видов состояний у этого признака. Например, группировка предприятий по формам собственности учитывает муниципальную, федеральную и собственность субъектов Федерации.

Если группировка проводится по *количественному признаку*, то нужно обратить внимание на число единиц исследуемого объекта и степень колеблемости группировочного признака. При небольшом объеме совокупности и малой степени колеблемости признака не следует образовывать большое число групп, т.к. группы будут включать недостаточное число единиц объекта. Поэтому показатели, рассчитанные для таких групп, не будут представительными и не позволят получить адекватную характеристику исследуемого явления.

Чем больше степень колеблемости группировочного признака, 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.

После определения числа групп определяют интервалы группировки.

Интервал группирования – устанавливает количественные границы групп. Как правило, он представляет собой промежуток между минимальными и максимальными значениями признака в группе. Каждый интервал имеет свою величину, верхнюю и нижнюю границы или хотя бы одну из них. *Нижняя граница* интервала – наименьшее значение признака в интервале, *верхняя граница* – наибольшее значение признака в интервале. *Величина интервала* – это разность между верхней и нижней границами интервала.

Закрытыми называются интервалы, у которых обозначены обе границы.

При группировке по количественному признаку границы интервалов могут быть обозначены по-разному. Если основанием группировки служит непрерывный признак, то одно и то же значение признака выступает и верхней, и нижней границами у двух смежных интервалов. Например, группы предприятий по численности работников, человек:

100-200 200-600 600-1000 1000-2000.

При таком обозначении границ может возникнуть вопрос, в какую группу должно войти предприятие с численностью работников 200 человек. Если нижняя граница интервала формируется по принципу «включительно», а верхняя – по принципу «исключительно», то предприятие должно быть отнесено ко второй группе, в противном случае – к первой.

Для того, чтобы правильно отнести к той или иной группе единицу объекта, у которой значение признака совпадает с границами интервалов, можно использовать открытые интервалы.

Открытые – это те интервалы, у которых указана только одна граница: верхняя у первого, нижняя – у последнего:

до 200 200-600 600-1000 1000 и более.

В данном случае предприятие с численностью работников 200 человек относится ко второй группе, 600 – к третьей и т. д. (так как верхняя граница первой группы построена по принципу «исключительно»).

При выполнении расчетов величины открытых интервалов условно приравниваются к величинам интервалов, примыкающих к ним.

Если в основании группировки лежит дискретный признак, то нижняя граница i – го интервала равна верхней границе $(i - 1)$ – го интервала, увеличенной на 1. Например, группировка населения по возрасту, лет:

0 - 4, 5 - 9, 10 - 14, 15 - 19, 20 - 24, 25 - 29 и т. д.

Закрытые интервалы группировки в зависимости от их величины бывают **равные** и **неравные**. Величина *равного интервала* определяется по формуле:

$$h = \frac{R}{k} = \frac{x_{\max} - x_{\min}}{k},$$

где x_{\max} и x_{\min} – максимальное и минимальное значения признака в совокупности, k – число групп, R – размах вариации.

Равные интервалы применяются, если вариация признака проявляется в сравнительно узких границах и распределение носит равномерный характер.

Неравные интервалы применяются в статистике, когда значения признака варьируют неравномерно и в значительных размерах, что характерно для большинства социально-экономических явлений, особенно при анализе макроэкономических показателей. Неравные интервалы могут быть прогрессивно возрастающие или убывающие в арифметической или геометрической прогрессии, а также могут определяться как равнонаполненные.

Величина интервалов, изменяющихся в *арифметической прогрессии*, определяется следующим образом:

$$h_{i+1} = h_i + a,$$

где a – константа ($a > 0$ при прогрессивно возрастающих интервалах и $a < 0$ при прогрессивно убывающих);

в *геометрической прогрессии*: $h_{i+1} = h_i \cdot q,$

где q - константа ($q > 1$ при прогрессивно возрастающих интервалах и $q < 1$ при прогрессивно убывающих). На практике чаще используется значение $q = 10$.

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

Пример. Построить группировку предприятий отрасли по показателю выручки от реализации продукции, который варьирует от 500 млн. руб. до 4 000 млн. руб.

Решение. В данном примере строить группировку с равными интервалами нецелесообразно, потому что, как правило, совокупность предприятий любой отрасли промышленности, торговли включает большое число малых предприятий, имеющих небольшую выручку. С ростом выручки от реализации продукции значительно уменьшается число предприятий. Таким образом, распределение числа предприятий по величине выручки является *неравномерным*. Поэтому следует построить *группировку с неравными интервалами*. Один из вариантов такой группировки

| № группы | Интервал | Величина каждого следующего интервала больше предыдущего на 200 млн. руб., т.е. увеличивается в арифметической прогрессии. |
|----------|---------------|--|
| 1. | 500 - 800 | |
| 2. | 800 – 1 300 | |
| 3. | 1 300 – 2 000 | |
| 4. | 2 000 – 2 900 | |
| 5. | 2 900 – 4 000 | |

При определении границ интервалов статистических группировок исходят из того, что изменение количественного признака приводит к появлению нового качества. В этом случае граница интервала устанавливается там, где происходит переход от одного качества к другому. Например, группировка предприятий по числу занятых показывает, что в промышленности и строительстве предприятия со среднесписочной численностью работающих 75 – 100 человек относятся к группе малых

предприятий, а в отраслях непроеизводственной сферы и в розничной торговле – к крупным.

В случае равнонаполненных интервалов вся совокупность разделяется на группы равного объема с числом единиц: $m = \frac{n}{k}$, где n – численность единиц совокупности, k – число групп, m – численность единиц в группе.

Данные ранжируются (упорядочиваются), отсчитывается число единиц m , составляющих первую группу, затем – вторую и т. д. Границы интервалов будут соответствовать фактическим значениям признака в каждой группе.

Бывает, что число групп заранее неизвестно и определяется опытным путем на основе перебора вариантов группировки, выявления такого варианта, который наилучшим образом позволяет увидеть различия между группами.

Пример. Распределите районы Новосибирской области по величине розничного товарооборота за 2005 г. Число групп задайте равным 5.

| № | Название района | Тыс. руб. |
|-----|-----------------|-----------|
| 1. | Баганский | 31 331 |
| 2. | Батабинский | 56 440 |
| 3. | Болотнинский | 99 212 |
| 4. | Венгеровский | 34 088 |
| 5. | Доволенский | 43 520 |
| 6. | Здвинский | 38 196 |
| 7. | Искитимский | 208 492 |
| 8. | Красноозерский | 51 387 |
| 9. | Купонский | 65 680 |
| 10. | Каргатский | 82 972 |
| 11. | Колыванский | 45 561 |
| 12. | Коченевский | 137 445 |
| 13. | Кочковский | 28 970 |
| 14. | Карасукский | 104 518 |
| 15. | Кыштовский | 21 253 |

| № | Название района | Тыс. руб. |
|-----|-----------------|-----------|
| 16. | Куйбышевский | 36 775 |
| 17. | Маслянинский | 47 248 |
| 18. | Мошковский | 92 955 |
| 19. | Новосибирский | 178 291 |
| 20. | Ордынский | 68 865 |
| 21. | Сузунский | 60 674 |
| 22. | Северный | 9 767 |
| 23. | Татарский | 23 944 |
| 24. | Тогучинский | 127 725 |
| 25. | Убинский | 24 559 |
| 26. | Усть-Тарский | 21 946 |
| 27. | Чановский | 44 876 |
| 28. | Черепановский | 117 021 |
| 29. | Чулымский | 36 637 |
| 30. | Чистоозерный | 33 775 |

Решение. Так как число групп заранее определено, остается решить вопрос о величине интервалов.

Возможно несколько вариантов.

а). Равные интервалы. Найдем величину интервала:

$$h = \frac{x_{\max} - x_{\min}}{k} = \frac{208492 - 9767}{5} = 39745.$$

Результаты представим в табличном виде.

Группировка районов по объему товарооборота (вариант а).

| Группы районов по объему товарооборота | Число районов в группе | Район | Средний размер товарооборота по группе, тыс.руб. |
|--|------------------------|---|--|
| 9767 – 49512 | 16 | 1, 4, 5, 6, 11, 13, 15, 16, 17, 22, 23, | 32652,8 |

| Группы районов по объему товарооборота | Число районов в группе | Район | Средний размер товарооборота по группе, тыс.руб. |
|--|------------------------|---------------------|--|
| | | 25, 26, 27, 29, 30 | |
| 49512 – 89257 | 6 | 2, 8, 9, 10, 20, 21 | 64336,3 |
| 89257 – 129002 | 5 | 3, 14, 18, 24, 28 | 108286,2 |
| 129002– 168747 | 1 | 12 | 137445,0 |
| 168 747- 208492 | 2 | 7, 19 | 193391,5 |
| Итого | 30 | | 65804,1 |

б). *Равнонаполненные интервалы.* Найдем численность единиц в группе:

$$m = \frac{n}{k} = \frac{30}{5} = 6.$$

Затем упорядочим районы по величине товарооборота.

Группировка районов по объему товарооборота (вариант б).

| Группы районов по объему товарооборота | Группы | Число районов в группе | Район | Средний размер товарооборота по группе, тыс.руб. |
|--|-------------|------------------------|------------------------|--|
| 9767 – 28970 | До 30 | 6 | 13, 15, 22, 23, 25, 26 | 21739,8 |
| 31331 - 38196 | 30-40 | 6 | 1, 4, 6, 16, 29, 30 | 35133,7 |
| 43520 – 56440 | 40-60 | 6 | 2, 5, 8, 11, 17, 27 | 48172,0 |
| 60674 - 99212 | 60-100 | 6 | 3, 9, 10, 18, 20, 21 | 78393,0 |
| 104518 - 208492 | 100 и более | 6 | 7, 12, 14, 19, 24,28 | 145582,0 |
| Итого | - | 30 | - | 65804,1 |

в). Величина интервала меняется в *арифметической прогрессии*. Тогда, например:

$$h_k = k \cdot h,$$

здесь h – величина первого интервала, h_k – величина k -го интервала, k - номер интервала.

$$h = \frac{x_{\max} - x_{\min}}{\sum k} = \frac{208492 - 9767}{1 + 2 + 3 + 4 + 5} = \frac{198725}{15} = 13248$$

$$h_1 = 1 \cdot 13248 = 13248$$

$$h_2 = 2 \cdot 13248 = 26496$$

.....

$$h_5 = 5 \cdot 13248 = 66240$$

Расчеты представим в таблице.

Группировка районов по объему товарооборота (вариант в).

| Группы районов по объему товарооборота | Число районов в группе | Район | Средний размер товарооборота по группе, тыс.руб. |
|--|------------------------|--|--|
| До 23016 | 3 | 15, 22, 26 | 17655,3 |
| 23016 – 49512 | 13 | 1, 4, 5, 6, 11, 13, 16, 17, 23, 25, 27, 29, 30 | 36113,8 |
| 49512 – 89256 | 6 | 2, 8, 9, 10, 20, 21 | 64336,3 |
| 89256 - 142248 | 6 | 3, 12, 14, 18, 24, 28 | 113146,0 |
| 142248 и более | 2 | 7,19 | 193391,5 |
| Итого | 30 | | 65804,1 |

Рассмотрение вариантов может быть продолжено. Как из этих вариантов выбрать лучший? В качестве критерия рекомендуется использовать межгрупповую дисперсию. Группировка, в которой межгрупповая дисперсия максимальная, считается лучшей. По данным таблиц рассчитаем:

$$S_a^2 = 2143673000$$

$$S_o^2 = 1943244100$$

$$S_6^2 = 2147738200.$$

Соответственно, лучший из трех рассмотренных вариантов – вариант **в**, при котором величина интервалов изменяется в арифметической прогрессии.

2.4 Виды статистических группировок:

Статистические группировки и классификации преследуют три основные цели:

- а) выделение качественно однородных совокупностей;
- б) изучение структуры совокупности;
- в) исследование существующих статистических зависимостей.

Каждой из этих целей соответствует особый вид группировки: типологическая, структурная, аналитическая (факторная).

Типологическая группировка решает задачу выявления и характеристики социально-экономических типов (частных подсовокупностей).

Структурная группировка дает возможность описать составные части (структуру) совокупности, а также проанализировать структурные сдвиги, состоящие в изменении величин.

Аналитическая (факторная) группировка позволяет оценивать статистические связи между взаимодействующими признаками изучаемой совокупности.

В зависимости от числа положенных в их основание признаков различают *простые* и *многомерные* группировки.

Группировка, выполненная по одному признаку, называется *простой*.

Многомерная группировка производится по двум и более признакам. Частным случаем многомерной группировки является *комбинационная* группировка, базирующаяся на двух и более признаках, взятых во взаимосвязи, в комбинации.

По отношениям между группировочными признаками выделяют *иерархические* и *неиерархические* группировки.

Иерархические группировки выполняются по двум и более признакам, когда значения второго признака определяются областью значений первого (например, классификация отраслей промышленности по подотраслям).

Неиерархические группировки строятся, когда нет строгой зависимости между значениями признаков.

Среди простых группировок особо выделяют *ряды распределения*.

Ряд распределения – это группировка, в которой для характеристики групп (упорядоченно расположенных в соответствии со значением изучаемого признака) применяется один показатель – абсолютная частота или численность группы. Другими словами, это ряд чисел, показывающий, как распределяются единицы некоторой совокупности по отношению к изучаемому признаку.

Ряды, построенные по качественному признаку, называются **атрибутивными** рядами распределения.

Ряды распределения, построенные по количественному признаку, называются **вариационными** рядами. Любой вариационный ряд состоит из двух элементов: вариантов и частот. *Вариантами* считаются отдельные значения признака, которые он принимает в вариационном ряду, т. е. конкретное значение варьирующего признака. *Частоты* – это численности отдельных вариантов или каждой группы вариационного ряда, т.е. это числа, показывающие, как часто встречаются те или иные варианты в ряду распределения. Сумма всех частот определяет численность всей совокупности, ее объем. *Частотами* называются частоты, выраженные в долях единицы или в процентах к итогу. Соответственно сумма всех частостей равна 1 или 100%.

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

Как известно, вариация количественных признаков может быть дискретной (прерывной) или непрерывной.

В случае дискретной вариации величина количественного признака принимает только целые значения. Следовательно, **дискретный**

вариационный ряд характеризует распределение единиц совокупности по дискретному признаку.

В случае непрерывной вариации величина признака у единиц совокупности может принимать в определенных пределах любые значения, отличающиеся друг от друга на сколь угодно малую величину. Построение **интервальных вариационных рядов** целесообразно, прежде всего, при непрерывной вариации признака, а также, если дискретная вариация проявляется в широких пределах, т. е. число вариантов дискретного признака достаточно велико.

По очередности обработки данных группировки бывают первичные (составленные на основе измеренных или собранных данных) и вторичные, являющиеся результатом перегруппировки ранее уже сгруппированного материала.

Относительно временного критерия группировки бывают *статические*, характеризующие совокупность на определенный момент времени или за определенный период, и *динамические*. Последние – это группировки, показывающие переходы единиц из одних групп в другие (а также вход и выход из совокупности). Количества таких переходов, рисующие внутреннюю динамику совокупности, удобно располагать в «шахматную» таблицу, которую называют матрицей перехода. Такую матрицу также часто называют миграционной или матрицей мобильности.

При проведении группировки решают следующие вопросы:

- 1) выделить группировочный признак;
- 2) определить число групп и величины интервалов;
- 3) при наличии нескольких группировочных признаков установить порядок, в котором они группируются между собой;
- 4) установить показатели, которыми должны характеризоваться группы (их называют сказуемым группировки).

Рассмотрим принципы построения различных видов группировок.

1. Типологическая группировка – это разделение исследуемой качественно разнородной совокупности на классы, социально-экономические типы. Примером типологической группировки является группировка промышленных предприятий по формам собственности.

Группировка промышленных предприятий одного из регионов России по формам собственности в 2004 г.:

| № п/п | Группы предприятий по формам собственности | Число предприятий | |
|-------|--|-------------------|--------------|
| | | Всего единиц | % к итогу |
| 1 | Федеральная собственность | 26326 | 93,6 |
| 2 | Муниципальная собственность | 89 | 0,3 |
| 3 | Частная собственность | 1366 | 4,9 |
| 4 | Смешанная собственность | 331 | 1,2 |
| | Всего: | 28112 | 100,0 |

Согласно данным таблицы, подавляющее большинство предприятий находилось в федеральной собственности, менее 5% составляли предприятия с частной формой собственности и только 1,5% - с муниципальной и смешанной собственностью.

Типологическая группировка строится с целью выделения из общей совокупности типов, различия между которыми очевидны и устойчивы и могут быть описаны одним или несколькими признаками. Однако на практике это бывает редко. Принадлежность группируемых объектов к общей совокупности приводит обычно к появлению у них некоторых общих особенностей, маскирующих различия между типами. Кроме того, недостаточно четкое обозначение отдельных типов, множественность признаков описания объекта и ряд других обстоятельств еще более усложняют группировку. Поэтому качественное проведение типологической группировки совокупности является весьма сложной задачей.

2. Структурная группировка характеризует структуру совокупности по какому-либо одному признаку. С помощью таких группировок могут изучаться:

состав населения по полу, возрасту, месту проживания и т.д. Пример такой группировки:

**Группировка населения России по размеру среднедушевого дохода в апреле
2004 г.:**

| № п/п | Группировка населения по размеру среднедушевого дохода, тыс. руб. в месяц | Численность населения | |
|----------|--|------------------------|--------------|
| | | Всего, млн. человек | % к итогу |
| 1 | До 40 | 2,4 | 1,6 |
| 2 | 40-80 | 23,4 | 15,8 |
| 3 | 80-120 | 34,8 | 23,5 |
| 4 | 120-160 | 29,4 | 19,8 |
| 5 | 160-200 | 20,7 | 13,9 |
| 6 | 200-240 | 13,5 | 9,1 |
| 7 | 240-280 | 8,7 | 5,9 |
| 8 | 280 и более | 15,5 | 10,4 |
| | Всего: | 148,4 | 100,0 |

Если для типологической группировки чаще используются открытые и неравные интервалы, то для структурной группировки более характерны закрытые равные интервалы.

Структурная группировка применяется для характеристики структуры совокупности и анализа структурных сдвигов.

Структура – это строение, форма организации совокупности, состоящей из отдельных элементов, единиц или групп единиц совокупности. Структура характеризуется с помощью показателей удельного веса $\left\{ \frac{f_i}{\sum f_i} \right\}$, доли, частости отдельных элементов в общей численности совокупности.

При проведении структурной группировки решаются следующие вопросы:

1. Выбор группировочного признака. В данном случае в качестве такового может выступать как существенный, так и несущественный признак.
2. При определении числа групп и величины интервала учитывают следующие факторы:
 - а) число групп связано с уровнем колеблемости группировочного признака. Чем значительнее вариация признака, тем больше при прочих равных условиях должно быть групп;
 - б) число групп должно отражать реальную структуру изучаемой совокупности;
 - в) не допускается выделение пустых групп. Если они все же появляются, то при проведении структурных группировок используют неравные интервалы.
3. Определение системы показателей для характеристики групп. Обязателен показатель численности групп. Он может быть представлен либо абсолютной частотой (количеством единиц в каждой группе), либо частотностью (удельным весом каждой группы в составе совокупности).

Для анализа изменения структуры совокупности во времени используют различные обобщающие показатели структурных сдвигов. Наиболее часто рассчитывают *линейный коэффициент абсолютных структурных сдвигов*:

$$S_d = \frac{\sum |d_1 - d_0|}{n},$$

где d_1 , d_0 – удельные веса (в %) отдельных элементов совокупности в рассматриваемом и предыдущем периодах; n – число выделяемых элементов совокупности.

Чтобы избежать взаимопогашения разных по знаку изменений долей, вместо модулей иногда применяют квадраты отклонений и получают *квадратический коэффициент абсолютных структурных сдвигов*:

$$S_\sigma = \sqrt{\frac{\sum (d_1 - d_0)^2}{n}}.$$

Линейный и квадратический коэффициенты абсолютных структурных сдвигов показывают, на сколько процентных пунктов в среднем отклоняются друг от друга удельные веса отдельных групп сравниваемых структур. При отсутствии структурных сдвигов оба эти показателя равны 0. Верхней границы изменения коэффициенты не имеют: чем больше изменения структуры, тем выше значения коэффициентов. Использование квадратического коэффициента предпочтительнее, т. к. он более чутко реагирует на сильные колебания структуры.

К наиболее простым сводным показателям структурных сдвигов относится и *индекс различий*, который в отличие от предыдущих коэффициентов имеет не только нижнюю, но и верхнюю границу изменения:

$$I_{разл} = \frac{1}{2} \sum |d_1 - d_0|.$$

Очевидно, что максимальная сумма модулей изменения долей может быть равна 2. Это возможно в гипотетической ситуации, когда в одном периоде вся совокупность сосредоточена в одной группе, а в следующем периоде – в другой. Поэтому теоретически индекс различий может иметь верхнюю границу, равную 1, однако в действительности он всегда меньше 1. Если изменений в структуре не происходило, индекс различий будет равен 0. Таким образом, $0 < I_{разл} < 1$. Чем ближе значение индекса различий к 1, тем более значительны изменения структуры.

3. Аналитическая группировка предназначена для установления тесноты связи между взаимодействующими признаками – факторным и результативным

Факторными называются признаки, под воздействием которых изменяются другие признаки, которые и образуют группу результативных признаков. Взаимосвязь проявляется в том, что с возрастанием значения факторного признака возрастает или убывает среднее значение результативного признака.

Например, производительность труда зависит от технического уровня предприятия: чем он выше, тем при прочих равных условиях выше производительность труда занятых на предприятии. Поэтому, группируя промышленные предприятия, производящие одну и ту же продукцию, по техническому уровню производства и исчисляя для каждой группы среднюю выработку товарной продукции на одного работающего, можно статистически отразить эту зависимость между факторами.

Принципиальными вопросами построения факторной группировки также являются:

- а) выбор группировочного признака;
- б) определение числа групп и величины интервала;
- в) выбор системы показателей для характеристики групп.

Чаще всего в качестве группировочного принимают факторный признак, выделенный на основе предварительного анализа. Интервалы в аналитической группировке берутся преимущественно равные либо равнонаполненные (группы с приблизительно одинаковой частотой). Величина интервала рассчитывается так же, как при построении структурной группировки. Среди показателей групп обязательным является среднее значение результативного показателя по каждой группе.

Группировка коммерческих банков России по сумме активов баланса

(данные условные):

| № п/п | Группа банков по сумме активов баланса, млн. руб. | Кол-во банков, единиц | В среднем на один банк | |
|----------|---|-----------------------------|-----------------------------------|--------------------------------------|
| | | | Численность занятых человек | Балансовая прибыль, млрд. руб. |
| 1 | до 20 000 | 19 | 184 | 22,5 |
| 2 | 20 000-30 000 | 8 | 313 | 31,6 |
| 3 | 30 000-40 000 | 7 | 374 | 36,0 |
| 4 | 40 000-50 000 | 9 | 468 | 69,2 |

| № п/п | Группа банков по сумме активов баланса, млн. руб. | Кол-во банков, единиц | В среднем на один банк | |
|----------|---|-----------------------------|-----------------------------------|--------------------------------------|
| | | | Численность занятых человек | Балансовая прибыль, млрд. руб. |
| 5 | 50 000 и более | 7 | 516 | 205,6 |
| | Всего: | 50 | 1855 | 364,9 |

Данные таблицы характеризуют зависимость между суммой активов банка и численностью занятых, а также суммой балансовой прибыли. Чем больше сумма активов, тем больше прибыль банка и численность его сотрудников.

СПИСОК ИСПОЛЬЗОВАННЫХ ИСТОЧНИКОВ

1. Теория статистики: учебник/ Р.А.Шмойлова, В.Г.Минашкин, Н.А.Садовникова, Е.Б.Шувалова; под ред. Р.А.Шмойловой.- 5-е изд., перераб. и доп. – М.: Финансы и статистика, 2008. – 656 с.
2. Статистика. Практикум [Текст] : учебное пособие для бакалавров / И. И. Елисеева [и др.] ; С.-Петерб. гос. ун-т экономики и финансов ; под ред. И. И. Елисеевой .— Москва : Юрайт, 2013 .— 514 с .
3. Ефимова М.Р. Практикум по общей теории статистики: учеб. пособие / М.Р. Ефимова, О.И. Ганченко, Е.В. Петрова.- 3-е изд., перераб. и доп.- М.: ИНФРА-М, 2008.- 368 с.
4. Елисеева И.И. Общая теория статистики: учебник/ И.И. Елисеева, М.М. Юзбашев; под ред. И.И. Елисеевой.- 5-е изд., перераб. и доп.- М.: Финансы и статистики, 2008.- 656 с.
5. Елисеева И.И. Практикум по общей теории статистики: учеб. пособие / И.И. Елисеева, Н.А. Флуд, М.М. Юзбашев; под ред. И.И. Елисеевой.- 2008. -512 с.
6. Бычков А. Г. Сборник задач по теории вероятностей, математической статистике и методам оптимизации. [Текст] : учебное пособие / А. Г. Бычков .— Москва : Форум, 2008 .— 224 с : ил.
7. Годин А.М. Статистика: Учебник. – 2-е изд., перераб.-М.: Издательско-торговая корпорация «Дашков и К^о», 2003.-472 с.

ОГЛАВЛЕНИЕ

| | |
|--|-----------|
| 1. ОСНОВНЫЕ ЭТАПЫ СТАТИСТИЧЕСКОГО ИССЛЕДОВАНИЯ... | 3 |
| 2. СТАТИСТИЧЕСКАЯ СВОДКА И ГРУППИРОВКА..... | 4 |
| 2.1. Сводка статистических данных..... | 4 |
| 2.2. Метод группировки..... | 5 |
| 2.3 Принципы построения статистических группировок..... | 7 |
| 2.4 Виды статистических группировок:..... | 16 |
| СПИСОК ИСПОЛЬЗОВАННЫХ ИСТОЧНИКОВ | 25 |

Розенцвайг А.К, Исавнин А.Г

СТАТИСТИКА
Сводка и группировка данных статистического
наблюдения

Учебно-методическое пособие

Подписано в печать 23.01.2024.
Формат 60x84/16. Печать ризографическая.
Бумага офсетная. Гарнитура «Times New Roman».
Усл.п.л. 1,75 Уч.-изд. л. 1,68

Тираж 100 экз. Заказ № 1241

Отпечатано в Издательско-полиграфическом центре
Набережночелнинского института
Казанского (Приволжского) федерального университета

423810, г. Набережные Челны, Новый город, пр.Мира, 68/19
тел./факс (8552) 39-65-99 e-mail: ic-nchi-kpfu@mail.ru