

Министерство образования Республики Беларусь
БЕЛОРУССКИЙ НАЦИОНАЛЬНЫЙ ТЕХНИЧЕСКИЙ УНИВЕРСИТЕТ
Факультет технологий управления и гуманитаризации
Кафедра менеджмента

СТАТИСТИКА

Учебно - методическое пособие
для студентов специальности
1-26 02 02 Менеджмент (по направлениям),
дневной и заочной формы обучения

Электронный учебный материал

Минск 2020

УДК 311.22 (075.8)

ББК 65.05197

С 78

Составители: Г.Д. Веренич, Марцева С. В.

Рецензент: доцент кафедры «Экономика и право» БНТУ, кандидат экономических наук Л.И. Дроздович

Рекомендовано кафедрой менеджмента ФТУГ БНТУ

Протокол № 5 от 30.11. 2020 г.

Рекомендовано методической комиссией ФТУГ БНТУ

Протокол № 2 от 17.12. 2020 г.

Данное учебно-методическое пособие представляет собой краткий курс лекций по дисциплине «Статистика». Рассмотрены общие вопросы теории статистики: сущность и задачи статистики на современном этапе, вопросы статистической методологии и особенности ее применения к изучению социально-экономических явлений и процессов.

Предназначено для студентов вузов, изучающих дисциплину «Статистика».

Белорусский национальный технический университет
пр-т Независимости, 65, г. Минск, Республика Беларусь.

Кафедра менеджмента ФТУГ.

Тел. (017) 231-71-25

Е-mail <http://www.bntu.by>

СОДЕРЖАНИЕ

Тема 1. Предмет, метод и задачи статистики.....	6
1.1 Зарождение статистической науки.....	6
1.2. Понятие и предмет статистики.....	8
1.3 Статистика как общественная наука.....	9
1.4 Массовые явления в обществе.....	9
1.5 Основные категории статистики.....	10
1.6 Закон больших чисел и его значение в статистике.....	12
1.7 Методы статистики.....	13
1.8 Этапы статистического исследования.....	14
1.9 Разделы статистической науки.....	15
1.10 Задачи статистики на современном этапе социально-экономического развития.....	16
1.11 Организация государственной статистики в Республике Беларусь...	16
Контрольные вопросы.....	17
Тема 2. Статистическое наблюдение.....	18
2.1 Источник статистических сведений.....	18
2.2 Сбор массовой первичной информации - исходный этап статистического исследования.....	18
2.3 Формы статистического наблюдения: отчетность, специально организованное статистическое наблюдение, регистры.....	20
2.4 Виды статистического наблюдения.....	24
2.5 План статистического наблюдения.....	25
2.6 Программно-методологические формуляры и принципы их разработки.....	28
2.7 Инструкция и ее содержание, последовательность разработки программы статистического наблюдения.....	29
2.8 Ошибки статистического наблюдения.....	30
2.9 Методы контроля материалов наблюдения.....	31
Контрольные вопросы.....	31
Тема 3 Сводка и группировка статистических данных.....	33
3.1 Статистическая сводка, ее составляющие и виды.....	33
3.2 Этапы сводки.....	34
3.3 Группировка – основной метод статистики.....	35
3.4 Задачи, решаемые на основе группировок.....	36
3.5 Типологические, структурные и аналитические группировки.....	36
3.6 Группировочные признаки, их виды.....	38
3.7 Простые и комбинационные группировки.....	39
3.8 Интервалы группировок.....	41
3.9 Классификации, применяемые в статистике.....	43
3.9 Метод вторичной группировки.....	46
3.10 Статистические ряды распределения.....	46
3.11 Статистические таблицы. Их виды и принципы построения.....	47

3.12 Статистические графики.....	49
Контрольные вопросы.....	51
Тема 4 Статистические показатели.....	53
4.1 Статистический показатель как количественная характеристика социально-экономических явлений в единстве с их качественной определенностью.....	53
4.2 Классификация показателей.....	54
4.3 Измерительные шкалы.....	56
4.4 Условия сопоставимости показателей. Приведение показателей в сопоставимый вид.....	57
4.5 Абсолютные величины как исходная форма статистических показателей. Прямые и косвенные методы их измерения.....	58
4.6 Относительные величины, их виды. Выбор базы при исчисления относительных величин.....	60
Контрольные вопросы.....	63
Тема 5. Средние величины статистики.....	65
5.1 Средняя величина, ее сущность и определение как категории статистической науки. Основные научные положения теории средних величин.....	65
5.2 Требования, предъявляемые к средним величинам.....	68
5.3 Принципы применения средних величин.....	68
5.4 Требование качественной однородности совокупности при расчете средних величин.....	69
5.5 Виды средних величин. Степенные средние.....	69
5.6 Средняя арифметическая величина.....	71
5.7 Свойства средних величин.....	72
5.8 Расчет средней арифметической в рядах распределения.....	73
5.9 Средняя гармоническая.....	73
5.10 Средняя геометрическая.....	75
5.11 Средняя квадратическая величина.....	75
5.12 Средняя кубическая величина.....	76
5.13 Средняя хронологическая величина.....	76
5.14 Структурные средние.....	78
5.15 Показатели дифференциации и концентрации.....	79
Контрольные вопросы.....	80
Тема 6. Анализ распределения (вариация).....	82
6.1 Понятие вариации.....	82
6.2 Задачи изучения вариации.....	83
6.3 Основные показатели вариации.....	83
6.4 Абсолютные показатели вариации.....	83
6.5 Относительные показатели вариации.....	85
6.6 Свойства дисперсий.....	86
6.7 Виды дисперсий и правила их сложения.....	87
Контрольные вопросы.....	89

Тема 7. Выборочный метод в статистических исследованиях.....	90
7.1 Понятие о выборочном наблюдении и его значение.....	90
7.2 Основные способы формирования выборочной совокупности.....	92
7.3 Ошибка выборки.....	94
7.4 Определение необходимой численности выборки.....	96
7.5 Распространение характеристик выборки на генеральную совокупность.....	98
Контрольные вопросы.....	99
Тема 8. Статистическое изучение динамики социально-экономических явлений (Анализ динамики).....	100
8.1 Ряды динамики и правила их построения.....	100
8.2 Сопоставимость рядов динамики.....	103
8.3 Аналитические показатели ряда динамики.....	104
8.4 Методы анализа основной тенденции ряда динамики.....	106
8.5 Методы изучения сезонных колебаний.....	108
Контрольные вопросы.....	110
Тема 9. Индексный метод в статистических исследованиях.....	112
9.1 Понятие об индексах, их классификация.....	112
9.2 Общие индексы количественных показателей.....	115
9.3 Общие индексы качественных показателей.....	116
9.4 Двухфакторные системы взаимосвязанных индексов.....	118
9.5 Индексный метод анализа динамики среднего уровня.....	118
9.6 Цепные и базисные индексы.....	120
Контрольные вопросы.....	121
Тема 10. Статистическое изучение связи социально-экономических явлений.....	123
10.1 Виды и формы взаимосвязей, изучаемых в статистике.....	123
10.2 Задачи статистического изучения связи.....	128
10.3 Определение формы связи.....	128
10.4 Измерение тесноты связи между признаками.....	130
10.5 Выявление влияния отдельных факторов.....	134
Контрольные вопросы.....	135
Список использованных источников.....	136

Тема 1. Предмет, метод и задачи статистики

1.1. Зарождение статистической науки

Современные реалии диктуют нам свои условия. Огромное информационное пространство и взаимодействие общества в различных сферах деятельности усилило интерес к статистике и как к науке, и как к сфере деятельности в целом. Данные статистики помогают ориентироваться нам в определенных важных сферах жизни, оценивать происходящие события в экономической, политической сферах общества, да и в повседневных бытовых делах.

Понятие «статистика» очень многогранно, нельзя остановиться на какой-то одной трактовке. Сейчас существует более тысячи определений этого слова в зависимости от того, в широком или узком смысле требуется объяснение. Статистика как термин корни свои берет от латинского слова «статус» - определенное положение вещей.

Изначально она все-таки была практической деятельностью. Историю статистики можно начинать еще с древних государств, когда там считали людей, скот, землю, имущество. Первые статистические сведения содержатся в китайских документальных источниках еще во II веке до нашей эры.

За триста лет до нашей эры Аристотель описал древние города, это тоже предпосылки возникновения статистики. Монгольские ханы переписывали население для взятия дани, всеобщую перепись в 1061 году проводили в Англии. Таким образом, в процессе осуществления практической деятельности человечество заложило основы статистики. Стали появляться определенные правила сбора и обработки информации.

Статистикой как наукой изначально стали заниматься в Англии и Германии. Две школы – английская научная и немецкая описательная - создавали статистику как науку. История статистики в Англии началась с английской научной школы политических арифметиков. У них была цель изучить явления общественной жизни с помощью чисел и их характеристик. В основе английской школы арифметиков было два направления. Первое представляли Д. Граунт и Э. Галей. Называлось оно демографическое. Второе направление, которое разработал В. Петти, называлось статистико-экономическим. Важные сферы жизни характеризовали статистическими данными, а не описательными оценками. Свое развитие английское направление получило также в Голландии и Франции.

Немецкая школа все сведения старалась описать, цифры здесь не использовались. Немецкие ученые описывали определенные периоды в развитии государства, быт и нравы, финансы, климат и так далее. Основал описательную немецкую школу Г. Конрринг.

Уже в 19 веке появилось еще одно, третье направление в истории развития статистики. Адольф Кетле (1796 – 1874 гг.), ученый из Бельгии, основал его. По основам, которые он разработал стали проводить перепись

населения. По его инициативе проводили статистические конгрессы на международном уровне. Он является также основоположником учения о средних величинах.

Историю развития статистики нельзя изучить, не оценив вклад слудеющих зарубежных ученых. Флоренс Найтингейл - реформатор из Великобритании - доказывал, что видные политики, предприниматели в своей деятельности должны основываться на статистических данных и статистическом анализе. Англичанин Карл Пирсон помог науке своими «критериями Пирсона». Он разработал корреляционную статистику. Большой вклад в развитие статистической науки внес Рональд Фишер.

Слово статистика имеет латинское происхождение: от лат. *status* — «состояние», что означало политическое состояние государства. В науку термин «статистика» ввел немецкий ученый Готфрид Ахенваль в XVIII веке, предложив заменить название курса «Государствоведение», преподававшегося в университетах Германии, на «Статистику», положив тем самым начало развитию статистики как науки и учебной дисциплины.

Конец 12 - начало 13 века в России можно смело отнести к развитию предпосылок истории возникновения статистики как науки. Уже в это время проводили перепись земель и населения (наличие храмов, монастырей, крепостных сооружений, жилых строений из дерева и камня).

Особое внимание в истории развития статистики в России стоит уделить периоду реформ Петра I. Многочисленные изменения во всех сферах государственного устройства определили потребность в учете. Сенат был центром проведения работ по статистике. Сюда собирали отчеты со всех ведомств государства Российского. Уже в восемнадцатом веке подавались сведения о тех, кто родился и умер. И тогда же начали переписывать рабочих различных фабрик. Государственный порядок регистрации актов гражданского состояния в России был введен раньше, чем в западных странах.

Новый этап в истории развития статистики в России пришелся на девятнадцатый век. 1802 год принято считать официальным днем рождения статистики, именно тогда министерства, по Манифесту Александра I, стали сдавать письменные отчеты. В 1811 году при Министерстве внутренних дел создали Статистическое отделение. Руководителем назначили К.Ф. Германа.

В России, как и в зарубежных государствах, существовали определенные направления (школы) в изучении статистики. Описательную школу представлял В.Н. Татищев. Что он сделал для науки? Ученый предложил определенные правила проведения ревизий, единый учет населения. М.В. Ломоносов усовершенствовал систему, предложенную В.Н. Татищевым. В работах М.В. Ломоносова были уже и аналитические данные. Описательной статистикой занимались И.К. Кириллов, М.И. Чулков, С.Н. Плещев и другие ученые. На основе их трудов была создана экономическая теория России. К.Ф. Герман написал книгу «Всеобщая история статистики. Для обучающихся сей науке». Здесь статистика рассматривалась именно как

наука. Много работ этой сфере посвятил К.И. Арсеньев. Основы статистики в России создал Журавский Д.П. Он системно изложил все вопросы этой науки. Экономическую и судебную статистику дополнили своими трудами А.Н. Радищев, Н.П. Огарев, А.И. Герцен.

Особую роль в формировании статистики как науки сыграло земство. Именно во времена земства были созданы статистические бюро. Земские статисты – В.И. Орлов, А.П. Шликевич и другие. Русские математики П.П. Чебышев, А.М. Ляпунова тоже внесли свой вклад в науку статистику, с помощью их работ развитие получило математическое направление в ней. Особенно интенсивно развивалась советская статистика. Анализу подвергались промышленность, народное хозяйство, население, строительство, сельское хозяйство, государственный бюджет и остальные сферы деятельности.

Таким образом, история развития статистики показывает, что статистическая наука сложилась в результате теоретического обогащения накопленного человечеством передового опыта учетно-статистических работ, обусловленных прежде всего потребностями управления жизнью общества.

1.2. Понятие и предмет статистики

Теоретической основой статистики являются положения социально-экономической теории. Опираясь на знания положений экономической теории, статистика анализирует конкретные формы проявления категорий, оценивает размеры явлений, осуществляет разработку методов их изучения и анализа.

В настоящее время термин статистика употребляется в 4 значениях:

1. – отрасль практической деятельности (статистический учет) по сбору, обработке, анализу и публикации массовых цифровых данных о самых различных процессах и явлениях общественной жизни. Эту деятельность на профессиональном уровне выполняет Национальный статистический комитет РБ (Белстат).

2. – совокупность цифровых сведений, характеризующих состояние массовых явлений и процессов общественной жизни; статистические данные, представляемые в отчетности предприятий, организаций, отраслей экономики, а также публикации в сборниках, справочниках, периодической прессе, которые являются результатом статистической работы.

3. – комплекс учебных дисциплин, обладающих определенной спецификой и изучающих количественную сторону массовых явлений и процессов в неразрывной связи с их качественным содержанием - учебный предмет в высших и средних специальных учреждениях.

4. – статистические методы (в том числе методы математической статистики), применяемые для изучения социально-экономических явлений и процессов.

Статистика, как любая наука, требует определения предмета исследования. В связи с этим различают статистику, занимающуюся изучением социально-экономических явлений, которая относится к циклу общественных наук, и статистику, занимающуюся закономерностями явлений природы, которая относится к естественным наукам.

Настоящий курс посвящен статистике социально-экономических явлений.

Объектом изучения социально-экономической статистики (или просто статистики) является общество во всем многообразии его форм и проявлений, процессы, происходящие в обществе, общественные явления.

Количественные изменения общественных явлений и процессов в неразрывной связи с их качественным содержанием и изучает статистика как наука.

Предметом изучения статистики выступают размеры и количественные соотношения качественно определенных социально-экономических явлений, закономерности их связи и развития в конкретных условиях места и времени.

Статистика – комплекс учебных дисциплин, обеспечивающих овладение методологией статистического исследования массовых социально-экономических явлений и процессов с целью выявления закономерностей их развития в конкретных условиях места и времени.

1.3. Статистика как общественная наука

Статистика – общественная наука, которая занимается сбором информации различного характера, упорядочиванием, сопоставлением, анализом и интерпретацией (объяснением). Она обладает следующими отличительными особенностями:

- изучает количественную сторону общественных явлений (данная сторона явления представляет его величину, размер, объем и имеет числовое измерение);
- исследует качественную сторону массовых явлений.

Предоставленная сторона явления выражает его специфику, внутреннюю особенность, отличающую его от других явлений. Качественная и количественная стороны явления всегда существуют вместе, образуют одно единое целое.

1.4. Массовые явления в обществе

Массовые явления - явления, любой природы, в том числе в экономике, повторяющиеся в пространстве и времени и отражающие некоторую статистическую закономерность.

1.5. Основные категории статистики

Статистика оперирует определенными категориями, то есть понятиями, отражающими существенные, всеобщие свойства и основные отношения явлений действительности.

Объект конкретного статистического исследования называют статистической совокупностью.

Статистическая совокупность – множество единиц (объектов или явлений), объединенных общей, единой закономерностью, но отличающихся между собой рядом признаков в пределах общего качества (например, совокупность коммерческих банков страны, население Республики Беларусь, совокупность вузов г. Минска и т.д.).

Специфическим свойством статистической совокупности является массовость единиц, поскольку явление характеризуется массовым процессом и всем многообразием определяющих его причин и форм.

Неделимые первичные элементы, выражающие качественную однородность статистической совокупности, то есть являющиеся носителями признаков называются единицами совокупности, а общее их количество – объемом совокупности. (Так, единицами совокупности могут выступать акционерные общества, фермерские хозяйства, человек, семья, станок, изделие и т.д.)

Свойства статистической совокупности:

- 1) массовость единиц совокупности;
- 2) количественная определенность единиц совокупности;
- 3) качественная однородность единиц совокупности (сходство единиц (объектов, явлений) по каким-либо существенным признакам, но различающимся по каким-либо другим признакам).

Единицы статистической совокупности характеризуются одним или несколькими свойствами, именуемыми в статистике признаками.

Признак – объективная характеристика единицы статистической совокупности, характерная черта или свойство, которое может быть определено или измерено и рассматривается как случайная величина.

Значения каждого признака отдельной единицы совокупности могут быть различными.

Вариант – возможное значение, которое может принимать тот или иной признак. Например, стаж работы может быть равен одному году, двум годам и т.д.

Вариация – различия в значениях того или иного признака у отдельных единиц, входящих в данную статистическую совокупность. Она возникает в результате того, что индивидуальные значения признака складываются под совокупным влиянием разнообразных факторов (условий), которые по-разному сочетаются в каждом отдельном случае. Например, успеваемость отдельного студента определяется затратами времени на подготовку к занятиям, способностью к обучению и т.п.

Наличие вариации является основной предпосылкой статистического исследования.

Варьирующие признаки могут быть количественными и неколичественными (атрибутивными)

Количественные признаки (признаки, варианты которых имеют числовое выражение): возраст, стаж работы, оплата труда. Атрибутивные, или качественные, признаки (признаки, не имеющие числового выражения и представляющие собой смысловые понятия): профессия, социальная принадлежность.

Количественные признаки могут быть дискретными и непрерывными.

Случаи, когда варианты признака могут принимать только одно из двух противоположных значений, говорит об альтернативных признаках (да, нет) (пол человека, городское и сельское население, продукция может быть годой или бракованной).

Признаки подразделяются на существенные признаки (признаки, выражающие содержательную сторону явлений, например, пол человека, возраст, стаж работы и др.) и несущественные признаки (второстепенные признаки, например, длина волос у человека и др.). Статистическому изучению подлежат существенные признаки.

Признаки могут быть порядковыми – признаки, имеющие несколько ранжированных, т. е. упорядоченных по возрастанию или убыванию, вариантов (уровень образования, разряд рабочего, экзаменационные оценки и др.).

Признаки, характеризующие статистическую совокупность, взаимосвязаны между собой, поэтому следует различать факторные (признаки-факторы) и результативные признаки.

Факторные признаки (независимые признаки, оказывающие влияние на другие связанные с ними признаки) и результативные признаки (зависимые признаки, изменяющиеся под влиянием факторных признаков). Так, квалификация, стаж работы рабочего – факторные признаки; производительность труда – результативный признак.

Статистическая совокупность состоит из массы отдельных единиц, разрозненных фактов. Задача статистики – установить общие свойства единиц совокупности, изучить имеющиеся взаимосвязи и закономерности развития. Достигается это с помощью расчета статистических показателей и их анализа.

Статистический показатель – это количественно-качественная обобщающая характеристика какого-либо свойства группы единиц или совокупности в целом. Этим он отличается от отдельных значений, которые, как отмечалось, называются признаками. Например, средний размер вкладов граждан в банках страны – статистический показатель, а размер вклада конкретного человека – признак.

Поскольку отдельные свойства совокупности не изолированы, а связаны между собой, то статистические показатели, характеризующие эти свойства, не являются разрозненными, а образуют систему показателей.

Система статистических показателей – это совокупность взаимосвязанных показателей, объективно отражающая существующие между явлениями взаимосвязи, она охватывает все стороны жизни общества как на макроуровне (страна, регион), так и на микроуровне (отдельное предприятие, домохозяйство, страна).

Важнейшей категорией статистики является статистическая закономерность. Закономерностью вообще принято называть повторяемость, последовательность и порядок изменений в явлениях.

Закономерность статистическая – количественная закономерность, которая проявляется в наличии однородных явлений или признаков при статистическом наборе данных. Она проявляется не в индивидуальном явлении, а в массе однородных явлений, при обобщении данных статистической совокупности, т.е. в среднем. Статистическая закономерность базируется на основе закона больших чисел. Статистика позволяет выявить обнаружить и распознать характер изменений, происходящих в социально-экономической среде. Статистика существует лишь там, где среди массы явлений можно выявить какие-либо общие черты, из которых и выявляются закономерности. Каждое отдельное измерение из выборки представляется как особый единичный случай рассматриваемой закономерности.

1.6. Закон больших чисел и его значение в статистике

Массовый характер общественных законов и своеобразие их действий предопределяет необходимость исследования совокупных данных. Закономерности, выявленные для той или иной совокупности, обнаруживаются при массовом наблюдении благодаря действию закона больших чисел.

Закон больших чисел порожден особыми свойствами массовых явлений. Последние в силу своей индивидуальности, с одной стороны, отличаются друг от друга, а с другой – имеют нечто общее, обусловленное их принадлежностью к определенному классу, виду. Причем единичные явления в большей степени подвержены воздействию случайных факторов, нежели их совокупность.

Закон больших чисел в наиболее простой форме гласит, что количественные закономерности массовых явлений отчетливо проявляются лишь в достаточно большом их числе.

Закон больших чисел в широком смысле – это общий принцип, согласно которому совокупное действие большого числа случайных величин приводит, при некоторых сравнительно широких условиях, к результату, почти независящему от случая, т.е. при большом числе случайных величин

их средний результат перестает быть случайным и может быть предсказан с большой степенью точности.

Случайное событие - событие, которое при заданной совокупности условий может произойти, а может и не произойти, но для которого определена вероятность его осуществления.

Содержание закона больших чисел сводится к следующему: в массе индивидуальных явлений общая закономерность проявляется тем полнее и точнее, чем больше их охвачено наблюдением.

Иными словами, закон больших чисел выражает общий принцип, в силу которого в большом числе явлений при некоторых общих условиях почти устраняется влияние случайных факторов.

Закон больших чисел получил свое математическое доказательство в теории вероятностей, а также подтверждение в многочисленных экспериментальных проверках. Так, французский естествоиспытатель А. Бюффон поставил следующий опыт: подбросил монету 4040 раз, при этом орел выпал 2048 раз, а решка 1992 раза. Отсюда, частность выпадения орла составила $2048/4040 = 0,507$ и отклонилась от вероятности его выпадения в каждом отдельном случае, равной $1/2$, лишь на $0,007$ ($0,507 - 0,500$).

Это говорит о том, что в рассматриваемом опыте почти полностью проявилось влияние основных, постоянных причин, а случайные причины отклонили результаты только на весьма незначительную величину. Т.е. в результате взаимопогашения случайных отклонений средние, исчисленные для величин одного и того же вида, становятся типичными, отражающими действие постоянных и существенных факторов в данных условиях времени и места.

В силу закона больших чисел случайные отклонения и ошибки в измерении величин взаимопогашаются в массе явлений. Опять-таки в силу этого же свойства следует изучать основные закономерности в большой совокупности объектов, а не на отдельных объектах, на величину которых, кроме основной закономерности, действуют двоякого рода погрешности: индивидуальные особенности данного события (объекта) и неточности, связанные с измерением его величины.

При определенных условиях величину отдельного элемента в совокупности можно рассматривать как случайную величину, имея в виду, что она является не только автоматическим результатом какой-то общей закономерности, но в то же время и сама определена действием множества факторов, не зависящих от этой общей закономерности.

Поэтому в основе статистического исследования всегда лежит массовое наблюдение фактов

1.7. Методы статистики

Метод статистики (или статистическая методология) представляет собой совокупность приемов, правил и принципов статистического

исследования социально-экономических явлений, т.е. сбора сведений, обработки их, вычисления показателей и анализа (оценки) полученных данных.

Общей основой статистического метода познания является диалектический метод, согласно которому общественные явления и процессы рассматриваются в развитии взаимной связи и причинной обусловленности.

Статистика опирается на такие диалектические категории, как количество и качество, причинность и закономерность, индивидуальное и общее.

В процессе исследования статистика может использовать и другие общенаучные методы:

- Аналогия — перенесение свойств одного объекта на другой.
 - Гипотезы — научно обоснованные предположения о возможных причинных связях между явлениями.
- Статистические методы используются комплексно — это обусловлено сложностью процесса экономико-статистического исследования.

Специфические статистические методы:

- метод массового статистического наблюдения - сбор первичных данных по единицам совокупности;
- метод статистической сводки и статистической группировки заключается в классификации, обобщении полученных первичных данных;
- методы анализа обобщающих статистических показателей позволяют дать характеристику изучаемому явлению при помощи статистических величин: абсолютных, относительных и средних с целью установления взаимосвязей и закономерностей развития процессов.

1.8. Этапы статистического исследования

Статистическое исследование — это научно организованный по единой программе сбор, сводка и анализ данных (фактов) о социально-экономических, демографических и других явлениях и процессах общественной жизни в государстве с регистрацией их наиболее существенных признаков в учетной документации.

В целом статистическое исследование должно:

- иметь общественно-полезную цель и всеобщую (государственную) значимость;
- относиться к предмету статистики в конкретных условиях его места и времени;
- выражать статистический вид учета (а не бухгалтерский и не оперативный);
- проводиться по заранее разработанной программе с ее научно обоснованным методологическим и другим обеспечением;

- осуществлять сбор массовых данных (фактов), в которых отражается вся совокупность причинно-следственных и других факторов, разносторонне характеризующих явление;
- регистрироваться в виде учетных документов установленного образца;
- гарантировать отсутствие ошибок наблюдения или же сводить их к возможному минимуму;
- предусматривать определенные критерии качества и способы контроля собранных данных, обеспечивая их достоверность, полноту и содержательность;
- ориентироваться на экономически эффективную технологию сбора и обработки данных;
- быть надежной информационной базой для всех последующих этапов статистического исследования и всех пользователей статистической информацией.

Различают три этапа работы со статистическими данными:

- 1) сбор – это массовое научно-организованное наблюдение, посредством которого получают первичную информацию об отдельных фактах (единицах) изучаемого явления. Данный статистический учет большого числа или всех входящих в состав изучаемого явления единиц является информационной базой для статистических обобщений, для формулирования выводов об изучаемом явлении или процессе;
- 2) группировка и сводка. Под этими данными понимают распределение множества фактов (единиц) на однородные группы и подгруппы, итоговый подсчет по каждой группе и подгруппе и оформление полученных итогов в виде статистической таблицы;
- 3) обработка и анализ. Статистический анализ включает стадию статистического исследования. Он содержит в себе обработку статистических данных, которые были получены при сводке, интерпретацию полученных результатов с целью получения объективных выводов о состоянии изучаемого явления и о закономерностях его развития. В процессе статистического анализа исследуются структура, динамика и взаимосвязь общественных явлений и процессов.

1.9. Разделы статистической науки

В процессе исторического развития в составе статистики как единой науки выделились и получили известную самостоятельность следующие разделы:

1. Общая теория статистики, которая разрабатывает понятие категорий и методы измерения количественных закономерностей общественной жизни. Она разрабатывает понятийный аппарат и систему категорий статистики, рассматривает методы сбора, сводки, обобщения и анализа статистических данных, то есть общую методологию статистического исследования.

2. Экономическая статистика, изучающая количественные закономерности процессов воспроизводства на различных уровнях.

3. Социальная статистика, изучающая количественную сторону развития социальной инфраструктуры общества (статистика здравоохранения, образования, культуры, моральная, судебная и др.).

4. Отраслевые статистики (статистика промышленности, агропромышленного комплекса, транспорта, связи и т.д.), которая изучает отдельные области общественных явлений или отдельные отрасли народного хозяйства.

Все отрасли статистики, развивая и совершенствуя свою методологию, способствуют развитию статистической науки в целом.

1.10. Задачи статистики на современном этапе социально-экономического развития

1. Сбор и обработка исходной информации массовых социально-экономических явлений с учетом рыночных отношений.

2. Определение статистических показателей, применяемых в исследованиях и их совершенствование параллельно совершенствованию бухгалтерского учета.

3. Определение прогнозов развития экономики на основе обобщения факторов, проявления этих событий.

1.11. Организация государственной статистики в Республике Беларусь

В основе организации статистики в РБ лежит Закон РБ «О государственной статистике». Данным законом определяется порядок организации и ведения государственной статистики в РБ;

регулируются отношения, связанные с осуществлением государственной статистической деятельности;

определяются: основные принципы и задачи государственной статистики; система органов государственной статистики, их права и обязанности; представление, распространение и использование статистических данных; ответственность за нарушение законодательства РБ в области государственной статистики.

Систему органов государственной статистики образуют:

- Республиканский орган государственного управления в области государственной статистики – Национальный статистический комитет РБ, который проводит государственную политику, осуществляет руководство государственной статистикой в РБ и координирует деятельность в этой области;

- территориальные органы государственной статистики, созданные республиканским органом государственного управления в области государственной статистики в административно-территориальных единицах;
- уполномоченные государственные организации – это государственные органы, за исключением органов государственной статистики, осуществляющие ведение государственной статистики в отношении находящихся в их подчинении организаций, а также по вопросам, входящим в их компетенцию.

Контрольные вопросы

1. Развитие статистики в древности.
2. Развитие статистики в западных странах.
3. Развитие статистики в России.
4. Что означает термин «статистика».
5. Предмет изучения статистики.
6. Статистика как учебная дисциплина.
7. Статистика как общественная наука.
8. Массовые явления в обществе.
9. Основные категории статистики.
10. Свойства статистической совокупности.
11. Варьирующие признаки.
12. Статистический показатель.
13. Система статистических показателей.
14. Статистическая закономерность.
15. Закон больших чисел.
16. Методы статистики.
17. Назовите этапы статистического исследования.
18. Разделы статистической науки.
19. Назовите основные задачи статистики.
20. Организация государственной статистики в Республике Беларусь.

Тема 2. Статистическое наблюдение

2.1. Источник статистических сведений

Государственная статистика выполняет важную роль в механизме управления экономикой, ориентированной на реализацию интересов государства в области информации.

Информация в переводе с латинского языка означает «осмысление, доведение сведений о чем-либо».

Статистическая информация (статистические данные) – это:

- 1.) первичный статистический материал о социально-экономических явлениях, формирующийся в процессе статистического наблюдения, который затем подвергается систематизации, сводке, анализу и обобщению;
- 2.) цифровые сведения в форме числовых рядов разнообразных величин, которые позволяют выявить определенные закономерности развития изучаемого явления, объекта или процесса;
- 3.) показатели, рассчитанные по совокупности предприятий, банков и других организаций, по рынкам, географическим и административным территориям.

Основными свойствами статистической информации являются массовость и стабильность. Первое свойство связано с особенностями предмета статистики, второе – с неизменностью однажды собранной информации, ее способностью устаревать и необходимостью получения новой информации.

Состав статистической информации во многом определяется потребностями общества в существующих условиях. Если раньше важнейшей задачей государственной статистики было обеспечение руководящих органов оперативной информацией о положении в стране (она часто носила закрытый характер), то в настоящее время почти вся информация, направляемая руководящим органам, становится достоянием общественности. Основными потребителями статистической информации является правительство, коммерческие организации, международные организации, общественность.

Качество, достоверность информации определяют эффективность использования статистики на любом уровне и в любой сфере.

Основой сбора статистической информации выступает наблюдение.

2.2. Сбор массовой первичной информации - исходный этап статистического исследования

Можно выделить четыре метода получения массовой первичной информации как исходного этапа статистического исследования: наблюдение, эксперимент, фокусирование и опрос.

Наблюдение – один из наиболее простых и дешевых методов исследования, осуществляемого в реальных условиях - заключается в проведении непосредственного наблюдения за людьми и окружающей обстановкой в районе интересующего объекта. Например, с его помощью можно оценивать половозрастной состав посетителей магазинов, ресторанов, культурных заведений, частоту и ассортимент покупаемых товаров.

Эксперимент позволяет выявить реальную реакцию потенциальных потребителей или иных групп людей на определенные факторы или их изменения. Например, для определения реакции покупателей на товар-новинку можно разместить его на прилавке рядом с аналогичным товаром и выяснять, как на него реагирует покупатель. Можно использовать и такой прием - ставить за прилавков молодых или опытных продавцов, мужчин или женщин. Эксперимент может применяться и при изучении таких проблем маркетинга, как определение наилучшего способа обучения торгового персонала или уровня цен.

Фокусирование заключается в целенаправленном отборе специальных фокус-групп, обычно в количестве от 7 до 15 чел., и обсуждении в их кругу интересующей исследователей проблемы под руководством профессионального психолога. Такие фокус-группы могут состоять из реальных или потенциальных потребителей товара, представителей определенных социальных слоев, экспертов или авторитетных людей. Обычно метод фокусирования хорошо подходит для поисковых исследований оценки продукции, организации рекламной кампании.

Опрос - наиболее универсальный и эффективный метод проведения маркетинговых исследований, особенно когда речь идет о сборе первичной информации.

Статистическое наблюдение – планомерная регистрация признаков у элементов статистической совокупности.

Не всякое соби́рание сведений является статистическим наблюдением (например, наблюдение покупателя за изменением цен на городских рынках). Статистическим можно назвать лишь такое наблюдение, которое обеспечивает регистрацию устанавливаемых фактов в учетных документах для последующего обобщения. Примерами статистического наблюдения служат систематический учет затрат на производство (его результат – бесперебойное производство) и популярные опросы общественного мнения с целью выявления отношения людей к представляющим интерес вопросам или событиям.

Статистическое наблюдение может проводиться органами государственной статистики, научно-исследовательскими институтами, экономическими службами банков, организаций. Оно обязательно должно быть массовым, систематическим, проводиться на научной основе по ранее разработанному плану и программе.

Планомерность статистического наблюдения заключается в том, что оно готовится и проводится по разработанному плану, который входит в план

всего статистического исследования и включает вопросы методологии, организации, техники сбора информации, контроля ее достоверности и оформления итоговых результатов.

Массовый характер статистического наблюдения предполагает, что оно охватывает большое число случаев проявления исследуемого явления или процесса, достаточное для получения правдивых статистических данных.

Систематичность статистического наблюдения определяется тем, что оно должно проводиться или систематически или непрерывно, регулярно. Только такой подход позволяет изучить тенденции и закономерности социально-экономических процессов, характеризующихся количественными и качественными изменениями.

Данные отдельных единиц наблюдения (людей, предприятий) должны быть сопоставимы друг с другом, иначе невозможно их последующее обобщение. Сопоставимость данных обеспечивается единством сроков наблюдения (например, численность студентов университета определяется на начало учебного года), его программы, методов регистрации данных.

Итак, в результате статистического наблюдения должна быть получена только объективная, сопоставимая и достаточно полная информация, позволяющая на последующих этапах исследования обеспечить научно обоснованные выводы о характере и закономерностях развития изучаемого явления.

Таким образом, *статистическое наблюдение* — это массовое (оно охватывает большое число случаев проявления исследуемого явления для получения правдивых статистических данных) планомерное (проводится по разработанному плану, включающему вопросы методологии, организации сбора и контроля достоверности информации), систематическое (проводится систематически, либо непрерывно, либо регулярно), научно организованное (для повышения достоверности данных, которая зависит от программы наблюдения, содержания анкет, качества подготовки инструкций) наблюдение за явлениями и процессами социально-экономической жизни, которое заключается в сборе и регистрации отдельных признаков у каждой единицы совокупности.

2.3. *Формы статистического наблюдения: отчетность, специально организованное статистическое наблюдение, регистры*

На этапе подготовки обследования нужно выяснить, как часто оно будет проводиться, будут ли обследоваться все единицы совокупности или только часть их, как получать информацию об объекте (путем интервью по телефону, по почте, простым наблюдением и т. п.) - необходимо определить формы, способы и виды статистического наблюдения.

В отечественной статистике используются три организационные формы (типы) статистического наблюдения:

- отчетность (предприятий, организаций, учреждений и т. п.);
- специально организованное статистическое наблюдение (переписи, единовременные учеты, обследования сплошного и не сплошного характера);
- регистры.

Отчетность - это основная форма статистического наблюдения, с помощью которой статистические органы в определенные сроки получают от предприятий, учреждений и организаций необходимые данные в виде установленных в законном порядке отчетных документов, скрепляемых подписями лиц, ответственных за их представление и достоверность собираемых сведений. Отчетность - это предусмотренная действующим законодательством форма организации статистического наблюдения за деятельностью предприятий и организаций, по которым органы государственной статистики получают информацию в виде установленных отчетных документов (форм отчетности), утвержденных Национальным статистическим комитетом РБ и Министерством финансов РБ, подписанных лицами, ответственными за достоверность сведений. Решающими являются две формы: баланс и отчет о прибылях и убытках.

Отчетность как форма статистического наблюдения основана на первичном учете и является его обобщением. Первичный учет представляет собой регистрацию различных фактов, событий, производимую по мере их совершения, как правило, на особом документе, называемом первичным учетным документом. Методы и формы организации статистической отчетности дифференцируются применительно к различным типам предприятий и формам предпринимательства, а также связанным с индивидуальными видами деятельности.

Действующую статистическую отчетность делят на типовую и специализированную. Состав показателей в типовой отчетности является единым для предприятий всех отраслей народного хозяйства. В специализированной отчетности состав показателей изменяется в зависимости от особенностей отдельных отраслей экономики.

По срокам представления отчетность бывает ежедневная, недельная, двухнедельная, месячная, квартальная и годовая. Кроме годовой отчетности все перечисленные виды представляют собой текущую отчетность.

По способу представления сведений отчетность делится на телеграфную, телетайпную, почтовую.

Специально организованное статистическое наблюдение. Специально организованное наблюдение проводится с целью получения сведений, отсутствующих в отчетности, или для проверки ее данных. Специально организованное наблюдение представляет собой сбор сведений посредством переписей, единовременных учетов и обследований. Наиболее простым примером такого наблюдения является перепись. Белорусская практическая статистика проводит переписи населения, материальных ресурсов, многолетних насаждений, неустановленного оборудования, строек незавершенного строительства, оборудования и др.

Перепись - это специально организованное наблюдение, повторяющееся, как правило, через равные промежутки времени, с целью получения данных о численности, составе и состоянии объекта статистического наблюдения по ряду признаков.

Характерными особенностями переписи являются: одновременность проведения ее на всей территории, которая должна быть охвачена обследованием; единство программы наблюдения; регистрация всех единиц наблюдения по состоянию на один и тот же критический момент времени. Программа наблюдения, приемы и способы получения данных по возможности должны оставаться неизменными. Это позволяет обеспечить сопоставимость собираемой информации и получаемых в ходе разработки материалов переписи обобщающих показателей. Тогда можно не только определить численность и состав исследуемой совокупности, но и проанализировать количественное изменение в период между двумя обследованиями.

Из всех переписей наиболее известны переписи населения. Цель последних состоит в установлении численности и размещения населения по территории страны, характеристики его состава по полу, возрасту, занятиям и другим показателям. Первая всеобщая перепись населения России была проведена в 1897 г., а последняя в Республике Беларусь проведена - в 2019 г.

В период подготовки всеобщей переписи для уточнения и апробации программно-методических и организационных вопросов наблюдения проводят пробную перепись. Например, такая перепись была осуществлена в декабре 1986 г. Это обследование охватило не все, а только пять процентов населения страны. Запись сведений при переписи населения всегда проводится на основе его опроса (без требования предъявить какие-либо документы, подтверждающие правильность ответа).

Кроме переписей статистика проводит и другие специально организованные наблюдения, в частности бюджетные обследования, которые характеризуют структуру потребительских расходов и доходов семей.

Регистровое наблюдение - это форма непрерывного статистического наблюдения за долговременными процессами, имеющими фиксированное начало, стадию развития и фиксированный конец. Оно основано на ведении статистического регистра. Регистр представляет собой систему, постоянно следящую за состоянием единицы наблюдения и оценивающую силу воздействия различных факторов на изучаемые показатели. В регистре каждая единица наблюдения характеризуется совокупностью показателей. Одни из них остаются неизменными в течение всего времени наблюдения и регистрируются один раз; другие показатели, периодичность изменения которых неизвестна, обновляются по мере изменения; третьи - представляют собой динамические ряды показателей с заранее известным периодом обновления. Все показатели хранятся до полного завершения наблюдения за единицей обследуемой совокупности.

Регистр населения - поименованный и регулярно актуализируемый перечень жителей страны. Программа наблюдения ограничена общими признаками, такими, как пол, дата и место рождения, дата вступления в брак (эти данные остаются неизменными в течение всего периода наблюдения) и брачное состояние (переменный признак). Как правило, регистры хранят информацию только по тем переменным признакам, изменение значений которых документально оформлено.

Информация в регистр заносится на каждого родившегося и прибывшего из-за границы. Если человек умер или выехал на постоянное место жительства из страны, то сведения о нем изымаются из регистра. Регистры населения ведутся по отдельным регионам страны. При перемене места жительства сведения по единице населения передаются в регистр соответствующей территории. В связи с тем, что правила регистрации довольно сложны и ведение регистра требует больших затрат, эта форма наблюдения практикуется в государствах с небольшой численностью и высокой культурой населения (в основном это европейские страны).

Необходимо отметить, что регистр населения, как любой регистр, охватывающий наблюдением значительную совокупность единиц, содержит данные по ограниченному числу признаков. Поэтому ведение регистра предполагает проведение специально организованных обследований, в том числе и переписей населения.

Регистр предприятий включает в себя все виды экономической деятельности и содержит значения основных признаков по каждой единице наблюдаемого объекта за определенный период или момент времени. Регистры предприятий содержат данные о времени создания (регистрации предприятия), его название и адрес, телефон, об организационно-правовой форме, структуре, виде экономической деятельности, количестве занятых (этот показатель отражает размер предприятия) и др.

Единый государственный регистр предприятий и организаций всех форм собственности (ЕГРПО) дает возможность организовать сплошное наблюдение по ограниченному кругу статистических показателей предприятий, зарегистрированных на территории РБ, позволяет получить непрерывные ряды показателей в случае изменения территориальной, отраслевой и других структур совокупности.

В регистр заносятся данные по всем предприятиям, организациям, учреждениям и объединениям независимо от их формы собственности, включая предприятия с иностранными инвестициями, банковские учреждения, общественные объединения и другие юридические лица.

Пользователями регистра могут быть любые юридические или физические лица, заинтересованные в информации.

2.4. Виды статистического наблюдения

Статистические наблюдения подразделяются на виды по следующим признакам:

- по времени регистрации данных;
- по полноте охвата единиц совокупности;

Виды статистического наблюдения по времени регистрации:

Текущее (непрерывное) наблюдение - проводится для изучения текущих явлений и процессов. Регистрация фактов осуществляется по мере их свершения непрерывно. (Регистрация семейных браков и разводов).

Прерывное наблюдение — проводится по мере необходимости, при этом допускаются временные разрывы в регистрации данных:

Периодическое наблюдение — регистрация данных по мере надобности, проводится через сравнительно равные интервалы времени (перепись населения).

Единовременное наблюдение — осуществляется без соблюдения строгой периодичности его проведения от случая к случаю, по мере необходимости (перепись жилого фонда).

По полноте охвата единиц совокупности различают следующие виды статистического наблюдения:

Сплошное наблюдение — представляет собой сбор и получение информации обо всех без исключения единицах изучаемой совокупности. Характеризуется высокими материальными и трудовыми затратами, недостаточной оперативностью информации. Применяется при переписи населения, при сборе данных в форме отчетности, охватывающей крупные и средние предприятия разных форм собственности.

Несплошное наблюдение — основано на принципе случайного отбора единиц изучаемой совокупности, при этом в выборочной совокупности должны быть представлены все типы единиц, имеющих в совокупности. Имеет ряд преимуществ перед сплошным наблюдением: сокращение временных и денежных затрат.

Несплошное наблюдение подразделяется на:

- *Выборочное наблюдение* - основано на случайном отборе единиц, которые подвергаются наблюдению. Получило наибольшее признание и распространение в статистической практике.

- *Монографическое наблюдение* — заключается в обследовании отдельных единиц совокупности, характеризующихся редкими качественными свойствами, при этом осуществляется подробное описание отдельных единиц совокупности для их углубленного изучения, которое не может быть столь результативным при массовом наблюдении. Обычно, монографическое наблюдение проводится в целях выявления имеющихся или намечающихся тенденций развития для изучения и распространения передового опыта отдельных хозяйств или выявления недостатков в работе отдельных предприятий. Пример монографического наблюдения:

характеристика работы отдельных предприятий, для выявления недостатков в работе или тенденций развития.

- *Метод основного массива* — состоит в изучении самых существенных, наиболее крупных единиц совокупности, имеющих по основному признаку наибольший удельный вес в изучаемой совокупности. Часть совокупности, о которой заведомо известно, что она не играет большой роли в характеристике совокупности, исключается из наблюдения. Например, структуру грузооборота можно изучить, исследовав только крупнейшие транспортные узлы.

- *Метод моментных наблюдений* — заключается в проведении наблюдений через случайные или постоянные интервалы времени с отметками о состоянии исследуемого объекта в тот или иной момент времени.

Способы получения статистической информации:

Непосредственное статистическое наблюдение — наблюдение, при котором сами регистраторы путем непосредственного замера, взвешивания, подсчета устанавливают факт подлежащий регистрации.

Документальное наблюдение — основано на использовании различного рода документов учетного характера. Включает в себя *отчетный* способ наблюдения — при котором предприятия представляют статистические отчеты о своей деятельности в строго обязательном порядке.

Опрос - заключается в получении необходимой информации непосредственно от респондента.

Существуют следующие виды опроса:

Экспедиционный — регистраторы получают необходимую информацию от опрашиваемых лиц и сами фиксируют ее в формулярах.

Способ саморегистрации — формуляры заполняются самими респондентами, регистраторы только раздают бланки и объясняют правила их заполнения.

Корреспондентский — сведения в соответствующие органы сообщает штат добровольных корреспондентов.

Анкетный — сбор информации осуществляется в виде анкет, представляющих собой специальные вопросники, удобен в случаях, когда не требуется высокая точность результатов. Организация рассылает (раздает) определенному кругу потребителей анкеты. Заполнение и возврат анкет является делом добровольным.

Явочный — заключается в предоставлении сведений в соответствующие органы в явочном порядке

2.5. План статистического наблюдения

Получение полной и качественной информации, которую можно использовать в практической работе или научном исследовании, зависит от

правильной организации и постановки статистического наблюдения. В связи с этим практическому проведению статистического наблюдения предшествует работа по составлению плана, обеспечивающего сбор первичных данных.

Любое статистическое наблюдение проводится по заранее и детально разработанному плану, который состоит из программно – методологических и организационных вопросов.

К программно - методологическим вопросам относятся:

1. Установление цели наблюдения.
2. Определение объекта и единицы наблюдения.
3. Выбор элемента совокупности.
4. Разработка программы наблюдения.
5. Выбор вида и способа наблюдения.

Основной практической целью статистического наблюдения является получение достоверной информации для выявления закономерностей развития явлений и процессов.

Установление *цели наблюдения* определяется конкретными потребностями в статистических данных. Она должна быть сформулирована ясно, четко, с указанием конкретных задач, стоящим перед данным наблюдением. Например, целью переписи населения является установление численности постоянного и наличного населения в стране, областях, населенных пунктах, определение состава и структуры населения по различным признакам.

Задача наблюдения непосредственно вытекает из задач статистического исследования и предопределяет его программу и формы организации.

В соответствии с целью определяется объект наблюдения.

Объект наблюдения - совокупность социально - экономических явлений и процессов, которые подлежат исследованию. Следует четко определить его границы, существенные признаки и характерные черты. Так, при переписи населения необходимо установить, какое население подлежит регистрации: постоянное или наличное.

Для ограничения объекта наблюдения часто пользуются *цензом*, то есть рядом ограничительных признаков. Нечеткое определение объекта наблюдения, его границ приводит к искажению действительности.

Определяя объект наблюдения, необходимо точно указать единицу наблюдения.

Единица наблюдения — источник информации, первичная ячейка, от которой должны быть получены необходимые сведения. В отдельных случаях может быть две единицы наблюдения, например, при переписи населения учитываются признаки семьи и каждого ее члена.

Единица (элемент) совокупности - первичный элемент объекта наблюдения, который обладает признаками, подлежащими регистрации при

проведении наблюдения. Единица наблюдения и единица совокупности могут совпадать.

Важным вопросом плана является составление *программы*, т.е. перечня признаков и показателей, подлежащих регистрации. Другими словами, это перечень вопросов, на которые должны быть получены ответы по каждой единице наблюдения.

В программу необходимо включать только те вопросы, которые необходимы, и ответы на которые будут использованы при разработке материалов или в контрольных целях. Вопросы должны быть сформулированы четко, кратко и ясно.

Организационная часть плана наблюдения определяет органы, место, время и сроки наблюдения и подготовительные работы.

Планом наблюдения четко определяются *органы наблюдения*, их права и обязанности, взаимоотношения между ними.

В плане статистического наблюдения указывается время и место наблюдения.

Местом наблюдения считается место, где непосредственно осуществляется регистрация признаков отдельных единиц. Место наблюдения необходимо определять точно. Например, при переписи регистрация осуществляется на месте проживания, а не на месте работы.

Выбор времени предусматривает решение двух вопросов – установление *критического момента* (даты) или *интервала времени* и определение *срока* (периода) наблюдения.

Время наблюдения - это время, к которому относятся регистрируемые события. Оно представляет собой определенный момент или период времени. Момент времени, к которому приурочены регистрируемые события, называется *критическим*. Например, критическим моментом переписи населения является 24 часа ночи. Показатель численности работающих или запас материалов могут быть представлены на определенный момент времени (на начало или конец месяца, года), а данные о количестве произведенной продукции могут быть представлены только за определенный интервал времени (день, месяц, квартал, год).

Период, в течение которого регистрируются события, называется *сроком (периодом) наблюдения*. Это время от начала до окончания сбора сведений, то есть время, в течение которого производится заполнение статистических формуляров (бланков определенных форм учета и отчетности).

Подготовительные работы включают:

1. Расчет потребности в кадрах, их подбор и подготовку.
2. Составление списков всех обследуемых единиц.
3. Печать бланков, инструкций и другой документации.
4. Проведение разъяснительной работы и т.д.

2.6. Программно-методологические формуляры и принципы их разработки

Для проведения наблюдения в каждом конкретном случае разрабатывается инструментарий статистического наблюдения, который включает в себя формуляр и инструкцию. *Формуляр статистического наблюдения* - это специальный документ, в котором регистрируются ответы на вопросы программы наблюдения. Он представляет собой разграфленный лист бумаги, в котором содержится перечень вопросов программы, свободные места для записи ответов (с указанием шифров и кодов) на них. Формуляр наблюдения состоит из двух частей: титульной и адресной. Титульная часть формуляра наблюдения включает: наименование статистического наблюдения и органа, его проводящего, а также дату и наименование органа, утвердившего данный формуляр. Адресная часть формуляра содержит запись точного адреса единицы или совокупности единиц наблюдения, их соподчиненность, иногда - сроки и место рассылки заполненных формуляров.

В зависимости от конкретного содержания и специфических особенностей проводимого статистического наблюдения формуляр наблюдения может иметь различные формы выражения и наименования: бланк, переписной лист, форма отчетности, анкета и т.д. В статистике различают две системы статистического формуляра: индивидуальную (формуляр - карточка) и списочная (формуляр - список).

Индивидуальный формуляр – это формуляр, предназначенный для регистрации в нем ответов на вопросы программы наблюдения только об одной единице наблюдения.

Списочный формуляр – это формуляр, предназначенный для регистрации в нем ответов на вопросы программы наблюдения о нескольких единицах наблюдения. Примером списочной системы формуляров служат переписные листы переписи населения страны, каждый из которых предназначается для регистрации нескольких лиц и характеризующих их признаков.

Для обеспечения единообразного толкования вопросов программы наблюдения и облегчения их понимания в статистических формулярах может быть дан статистический подсказ, либо в указании методики или способа расчета того или иного показателя.

Статистический подсказ – это перечень возможных ответов на поставленный вопрос. Например, при обследовании строительных организаций города после вопроса о величине рентабельности в скобках необходимо указать методику расчета рентабельности как отношение прибыли к себестоимости работ.

Полным называется статистический подсказ, содержащий исчерпывающий перечень ответов на поставленный в статистическом

формуляре вопрос. Полный подсказ предполагает выбор заполняющим формуляр только одного из перечисленных ответов на вопрос.

Неполным называется статистический подсказ, содержащий некоторые (но не все) из возможных ответов на поставленный в статистическом формуляре вопрос.

2.7. Инструкция и ее содержание, последовательность разработки программы статистического наблюдения

С целью обеспечения более полного и правильного процесса организации и проведения наблюдения к программе статистического наблюдения составляется подробная инструкция.

Статистическая инструкция - это документ, разъясняющий вопросы программы статистического наблюдения, порядок заполнения статистического формуляра и частично планово-организационные вопросы. В инструкции отражаются цели и задачи наблюдения, сведения об объекте и единицах наблюдения, о времени и сроках проведения наблюдения, об оформлении результатов и сроках их представления в соответствующие организации. В инструкции, так же как и в формуляре наблюдения, но в более развернутом виде, могут быть представлены толкование того или иного вопроса программы, примерные ответы на вопросы, типичные случаи заполнения формуляров и т.д. Инструкция статистического наблюдения должна быть составлена кратко, просто и ясно.

Разработка программы статистического наблюдения включает в себя:

- 1) определение цели и задач статистического наблюдения;
- 2) составление перечня показателей (вопросов) для сбора данных;
- 3) описание методики расчета показателей.

Проведение мероприятий по разработке программ статистических наблюдений осуществляется в следующем порядке:

- 1) издается приказ об организации работ по разработке и утверждению статистических форм на следующий отчетный год;
- 2) утверждается график разработки и утверждения статистических форм;
- 3) вносятся и анализируются предложения по совершенствованию действующих или разработке новых статистических форм;
- 4) определяется перечень статистических форм на основе планируемых статистических наблюдений по отраслям статистики;
- 5) осуществляется разработка статистических форм и инструкций по проведению статистических наблюдений;
- 6) проводится экспертиза статистических форм на соответствие действующим классификациям, стандартам оформления;
- 7) статистические формы согласовываются с заинтересованными государственными органами.

2.8. Ошибки статистического наблюдения

Каждое конкретное измерение величины данных, осуществляемое в процессе наблюдения, дает, как правило, приближенное значение величины явления, в той или иной мере отличающееся от истинного значения этой величины. *Точностью статистического наблюдения* называется степень соответствия какого-либо показателя или признака, исчисленного по материалам наблюдения, действительной его величине. Расхождение между результатом наблюдения и истинным значением величины наблюдаемого явления называется *ошибкой наблюдения*.

В зависимости от характера, стадии и причин возникновения различают несколько типов ошибок наблюдения.

Виды ошибок:

1. *Ошибки наблюдения (регистрации)* – это ошибки, связанные с нарушением процедуры регистрации исследуемого признака.

1.1. *Преднамеренные ошибки* – ошибки, тенденциозно искажающие информацию: завышение или занижение тенденций.

1.2. *Непреднамеренные ошибки:*

1.2.1. *Случайные ошибки* – описки, ошибки, оговорки (нивелируются посредством закона больших чисел). (К ним относятся оговорки и описки опрашиваемого лица. Они могут быть направлены в сторону уменьшения или увеличения значения признака. На конечном результате они, как правило, не отражаются, так как взаимопогашаются при сводной обработке результатов наблюдения.)

1.2.2. *Систематические ошибки* – ошибки аккумуляции, ошибки округления количественных параметров.

(Систематические ошибки имеют одинаковую тенденцию либо к уменьшению, либо к увеличению значения показателя признака. Это связано с тем, что измерения, например, производятся неисправным измерительным прибором или ошибки являются следствием неясной формулировки вопроса программы наблюдения и др. Систематические ошибки представляют большую опасность, так как в значительной мере искажают результаты наблюдения).

2. *Ошибки выборки* – ошибки, связанные с неправильным отбором единиц исследования.

3. *Ошибки исчисления, регистрации* – ошибки, связанные с вводом статистических данных и расчетом показателей и коэффициентов.

К *ошибкам регистрации* относятся те неточности, которые возникают при записи данных в статистический формуляр (первичный документ, бланк, отчет, переписной лист) или при вводе данных в вычислительную технику, искажение данных при передаче через линии связи (телефон, электронную почту). Часто ошибки регистрации возникают из-за несоблюдения формы бланка, т. е. запись производится не в установленную строку или графу

документа. Случается и преднамеренное искажение значений отдельных показателей.

2.9. Методы контроля материалов наблюдения

Желательно проводить некоторые мероприятия, которые помогут предупредить, выявить и исправить ошибки наблюдения. К таким мероприятиям относятся:

- 1) подбор квалифицированных кадров и качественное обучение персонала, связанного с проведением наблюдения;
- 2) организация контрольных проверок правильности заполнения документов сплошным или выборочным методом;
- 3) арифметический и логический контроль полученных данных после завершения сбора материалов наблюдения.

Основными видами контроля достоверности данных являются синтаксический, логический и арифметический.

1. *Синтаксический контроль* означает проверку правильности структуры документа, наличие необходимых и обязательных реквизитов, полноту заполнения строк формуляров в соответствии с установленными правилами. Важность и необходимость синтаксического контроля объясняются применением для обработки данных вычислительной техники, сканеров, которые предъявляют жесткие требования к соблюдению правил заполнения формуляров.

2. *Логическим контролем* проверяются правильность записи кодов, соответствие их наименованиям и значениям показателей. Выполняется проверка необходимых взаимосвязей между показателями, сопоставляются ответы на различные вопросы и выявляются несовместимые сочетания. Для исправления ошибок, выявленных при логическом контроле, возвращаются к исходным документам и делают поправки.

3. При *арифметическом контроле* сравниваются полученные итоги с предварительно подсчитанными контрольными суммами по строкам и по графам. Довольно часто арифметический контроль основывается на зависимости одного показателя от двух или нескольких других (например, является произведением других показателей). Если арифметический контроль итоговых показателей обнаружит, что данная зависимость не соблюдается, это будет свидетельствовать о неточности данных.

Таким образом, контроль достоверности статистической информации осуществляется на всех этапах проведения статистического наблюдения – начиная со сбора первичной информации и до этапа получения итогов

Контрольные вопросы

1. Понятие статистической информации.
2. Свойства статистической информации.

3. Понятие наблюдения.
4. Понятие эксперимента.
5. Понятие опроса.
6. Что такое статистическое наблюдение.
7. Формы статистического наблюдения.
8. Понятие отчетности.
9. Специально организованное статистическое наблюдение.
10. Виды статистического наблюдения по времени регистрации.
11. Виды статистического наблюдения по полноте охвата единиц совокупности.
12. Понятие выборочного наблюдения.
13. Понятие монографического наблюдения.
14. Понятие метода основного массива.
15. Понятие метода моментных наблюдений.
16. Способы получения статистической информации.
17. Виды опроса.
18. Программно-методологические вопросы статистического наблюдения.
19. Программно-методологические формуляры.
20. Понятие статистической инструкции.
21. Ошибки статистического наблюдения.
22. Методы контроля материалов наблюдения.

Тема 3. Сводка и группировка статистических данных

3.1 Статистическая сводка, ее составляющие и виды

Статистическое наблюдение позволяет получить сведения по каждой единице исследуемого объекта. В связи с этим данная информация не является обобщающей и с ее помощью нельзя сделать выводы об изучаемом явлении в целом.

Вторым этапом статистического исследования является систематизация исходных данных, полученных в результате наблюдения, их обработка и получение обобщающих статистических показателей, характеризующих весь объект, то есть сводка.

Сводка - это научно организованная обработка материалов наблюдения, включающая в себя систематизацию, группировку данных, составление таблиц, подсчет групповых и общих итогов, расчет производных показателей. Результатом сводки служат обобщающие характеристики, отражающие в целом всю совокупность.

Сводка проводится на основе предварительного статистического анализа явления и процессов, происходящих в нем. При осуществлении сводки все статистические итоги должны отражать важнейшие характеристики исследуемого объекта. Классификация видов статистической сводки зависит от ряда признаков, таких как: глубина обработки; место проведения сводки; способ обработки исходных данных.

В зависимости от глубины обработки первичной информации, полученной в результате статистического наблюдения, сводка делится на:

- простую;
- сложную.

Простая сводка предусматривает подсчет общих итогов по всей совокупности единиц статистического наблюдения. При этом определяется общий объем изучаемого явления. Например, для получения общей численности студентов высших учебных заведения Республики Беларусь достаточно сложить данные о численности студентов всех высших учебных заведений. Преимуществами данного вида сводки являются: быстрота получения общих итоговых данных по основным показателям; возможность принятия оперативных управленческих решений.

Сложная сводка представляет собой комплекс процедур, которые включают группировку единиц совокупности, подсчет итогов характеристик единиц совокупности по каждой группе и по совокупности в целом, а также представление полученных результатов в виде статистических таблиц и графиков.

Выделение однородных групп является основным содержанием сложной сводки.

По признаку места проведения сводки сводка делится на два вида:

- централизованную сводку;

- децентрализованную сводку.

Централизованная сводка предусматривает концентрацию всей исходной статистической информации в одном органе, в котором он полностью обрабатывается по соответствующей программе с получением выходных статистических таблиц. Такая сводка используется, как правило, при проведении единовременных статистических обследований.

При *децентрализованной сводке* обобщение исходных данных проводится последовательными этапами снизу доверху по иерархической системе административно-территориального управления. При этом на каждом этапе производится соответствующая обработка информации. К такой сводке прибегают при обработке статистической отчетности. На районном уровне осуществляется обработка первичных форм статистической отчетности предприятий и организаций, расположенных на данной территории, затем подводятся итоги по региону, которые передаются на областной уровень, где также осуществляется сводка отчетных данных по области. На последнем этапе итоги по области передаются на республиканский уровень, где сводятся данные по всей стране.

В зависимости от способа обработки исходных данных бывает: автоматизированной и ручной.

Автоматизированная сводка предусматривает электронную обработку исходных данных с помощью ЭВМ на всех этапах проведения сводки.

При *ручной сводке* все итоги подсчитываются вручную. Такой вид сводки в настоящее время встречается довольно редко, поскольку широкое распространение получили персональные компьютеры и созданы на их базе автоматизированные рабочие места (АРМ) с использованием пакетов прикладных программ (ППП), обеспечивающих автоматизацию практически всех статистических разработок.

3.2. Этапы сводки

Научная разработка статистической сводки проводится в несколько этапов:

- формирование целей и задач проведения статистической сводки;
- выбор группировочного признака;
- определение порядка формирования групп;
- осуществление группировки с расчетом интервалов группировки;
- разработка системы статистических показателей, характеризующих выделенные группы и объект исследования в целом;
- формирование макетов статистических таблиц с целью представления расчетных данных.

План статистической сводки содержит указания о последовательности и сроках выполнения отдельных частей сводки, ее исполнителях и о порядке изложения и представления результатов.

Для проведения статистической сводки разрабатывается программа сводки, которая содержит вышеперечисленные вопросы, включая определение вида сводки.

Следует учитывать, что перед проведением статистической сводки составляется организационный план, в котором указывается: организация, которая будет проводить сводку; время ее проведения; порядок проведения сводки; перечень статистических показателей, который должен быть опубликован.

3.3. Группировка – основной метод статистики

В сводке статистического материала отдельные единицы статистической совокупности объединяются в группы при помощи метода группировок.

Статистической группировкой называется разделение единиц изучаемой совокупности по существенным для них признакам на качественно однородные группы или объединение изучаемых единиц в частные совокупности по существенным для них признакам, каждая из которых характеризуется системой статистических показателей. Например, группировка промышленных предприятий по формам собственности, группировка населения по размеру среднедушевого дохода, группировка коммерческих банков по сумме активов баланса и т.д.

Научно обоснованное распределение на группы дает возможность сделать правильные выводы об изучаемой совокупности и происходящих в ней процессах.

Группировку можно трактовать также как объединение единиц совокупности в частные группы по типовым признакам. Группировка позволяет дать более глубокую и разностороннюю характеристику изучаемого объекта.

Группировка особенно важна в условиях рыночной экономики. Например, при оценке денежных доходов населения: их можно рассматривать как денежные доходы домашних хозяйств (семей), представляющих собой сумму денежных средств, полученных членами домашних хозяйств.

Причем все денежные доходы можно группировать следующим образом: как доходы, полученные в виде заработной платы, дохода от индивидуальной трудовой и предпринимательской деятельности, пенсий, стипендий, пособий, компенсационных и других дополнительных выплат (включая стоимость благотворительной помощи), процентов, дивидендов, ренты и других доходов от собственности, средств от продажи продукции личного подсобного хозяйства и прочих денежных поступлений.

Большие достижения в области применения метода группировок имеет современная отечественная статистика. Введение группировочных таблиц, содержащих показатели международной системы национальных счетов

(СНС), превращает группировки в эффективный метод анализа и вскрытия резервов в экономике.

В зависимости от глубины группировки можно выделить два ее вида: простую и сложную.

Простая или *монотетическая* группировка обеспечивает построение групп на основе одного группировочного признака.

При проведении группировки по нескольким признакам, она называется *сложной* или *политетической*. Как правило, сложная группировка бывает комбинационной, при проведении которой группы, выделенные по одному признаку затем подразделяются на подгруппы по другому признаку. Альтернативой комбинационных группировок выступают *многомерные группировки*, которые обеспечивают группировку по множеству признаков. К таким группировкам прибегают при классификации данных.

Группировка является одним из самых сложных в методологическом плане этапов статистического исследования, поскольку от правильности выбора группировочного признака и формирования групп зависит весь результат статистического анализа. Она позволяет раскрыть объективное состояние какого-либо явления, выявить наиболее существенные черты и его свойства, получить данные о размерности каждой группы, их соотношений в общей совокупности, выявить связи между изучаемыми показателями и признаками, лежащими в основе группировки.

3.4. Задачи, решаемые на основе группировок

С помощью группировки статистических данных могут быть решены следующие задачи:

- выделение социально-экономических типов явлений;
- изучение структуры явления и структурных сдвигов, происходящих в нем;
- установление причинно-следственной связи и зависимости между отдельными признаками явления.

3.5. Типологические, структурные и аналитические группировки

В соответствии с перечисленными задачами используются различные виды группировок, среди которых выделяются: типологические, структурные и аналитические группировки.

Типологические группировки формируются путем разделения качественно разнородной совокупности на классы, социально-экономические типы, однородные группы единиц в соответствии с правилами научной группировки. Они предназначены для исследования социально-экономических процессов и позволяют проследить жизненный цикл развития явлений, т.е. их зарождение, развитие и отмирание.

Основное внимание в типологической группировке должно уделяться идентификации типов и выбору группировочного признака. Для построения типологической группировки необходимо воспользоваться количественными и качественными (атрибутивными) признаками.

Примерами типологической группировки могут служить группировки секторов экономики, хозяйствующих субъектов по формам собственности (группы предприятий государственной собственности, частной собственности и смешанной собственности). Примером типологической группировки является также группировка населения по возрасту.

Структурные группировки - это группировки, когда проводится деление однородной совокупности на группы, характеризующие ее структуру по одному или нескольким варьирующим признакам. Данный вид группировок используется для изучения строения рассматриваемой совокупности. Как правило, структурные группировки формируются на основе образования качественно однородных групп.

На основе структурных группировок могут быть исследованы: состав товарооборота по формам собственности; социальный состав населения, состав населения по полу, возрасту, месту проживания, уровню дохода; состав предприятий торговли формам торговли, численности занятых, стоимости основных фондов, объему торговых залов, группировка хозяйств по объему продукции, структура депозитов по сроку их привлечения и т.д. Кроме того, с помощью структурной группировки может быть исследован процесс концентрации производства, специализация предприятий или отраслей, диверсификация капитала, степень монополизации рынка и т.п.

Аналитическая группировка – это группировка, выявляющая взаимосвязи между изучаемыми явлениями и их признаками. В основе аналитической группировки лежит факторный признак, и каждая выделенная группа характеризуется средними значениями результативного признака.

Все признаки можно подразделить на два вида: факторные и результативные. При этом следует выделить следующие взаимосвязи:

- в качестве факторного признака выступает количественный признак, а в качестве результативного - качественный, например: производительность труда и квалификация работников;

- в качестве факторного признака выступает качественный признак, а в качестве результативного – количественный, например: зависимость средней месячной заработной платы работника от его квалификации;

- в качестве факторного и результативного признаков выступает качественный признак, например: число малых предприятий по отраслям экономики;

- в качестве факторного и результативного признаков выступает количественный признак, например: объем прибыли и инвестиций.

Взаимосвязь проявляется в том, что с возрастанием (убыванием) значения факторного признака возрастает (убывает) значение результативного признака. Например, уровень жизни населения зависит от

получаемых доходов: чем выше доходы, тем выше уровень жизни (с учетом инфляционных процессов). Поэтому, при осуществлении группировки населения по уровню получаемых доходов и исчисляя для каждой группы средний уровень потребления на душу населения, можно статистически отразить данную зависимость между рассматриваемыми факторами.

Следует отметить особенности аналитической группировки:

в основу группировки в качестве группировочного признака, как правило, берется факторный признак;

каждая сформированная группа характеризуется средними значениями результативного признака.

С помощью аналитических группировок можно исследовать многообразные связи и зависимости между варьирующими признаками. Обязательным условием использования аналитических группировок является наличие качественно однородной статистической совокупности. Никакие другие условия применения данного вида группировок не существует. Это является одним из основных преимуществ метода аналитических группировок перед другими методами анализа взаимосвязи, такими как корреляционный и регрессионный анализ.

По взаимосвязи между признаками различают иерархические (когда есть строгая зависимость значения второго признака от первого) (например, классификация отраслей промышленности по подотраслям; вузов по факультетам) и неиерархические (когда строгой зависимости значений второго признака от первого не существует). Таковы, например, группировки населения по полу, семейному положению, образованию; распределение населения на городское и сельское.

3.6. Группировочные признаки, их виды

Признаки, положенные в основу группировки, т.е. по которым проводится распределение единиц исследуемой совокупности по группам, называются *группировочными признаками* или основанием группировки. От правильности выбора группировочного признака зависят качество группировки и достоверность полученных результатов.

Выделить типичное можно не по любому признаку, а только по определенному, который должен изменяться в зависимости от условий места и времени. Для правильного выбора группировочных признаков необходимо предварительно выявить возможные типы, четко сформулировать познавательную задачу.

Признаки, по которым может быть проведена группировка, в свою очередь, классифицируются следующим образом:

- по форме выражения группировочные признаки делятся на: атрибутивные и количественные.

Атрибутивные признаки не имеют количественного выражения, т.е.: для характеристики индивида они обозначают пол человека, его профессию,

квалификацию; для торговых предприятий характеризуют их спецификацию, - продовольственные, промтоварные; для товарных рынков - вещевые, сельскохозяйственные и т.п.

Количественные признаки принимают различные цифровые значения у отдельных единиц исследуемой совокупности, например: величина прибыли предприятия, объем совокупного дохода, численность работающих и т.д.

В свою очередь количественные признаки подразделяются на: дискретные и непрерывные.

Дискретные, или прерывные признаки выражаются только целыми числами, например: число торговых предприятий, численность населения, количество комнат в жилом помещении и т.д.

Непрерывные признаки принимают как целые, так и дробные значения, например: объем инвестированного капитала, расходы населения на приобретение того и или иного вида товара, объем предложения какого-либо товара на рынке товаров и услуг в стоимостном выражении и т.д.

- в зависимости от характера колеблемости признаки делятся на: альтернативные и признаки, имеющие множество количественных значений.

Альтернативные признаки - это признаки, которыми одни единицы совокупности обладают, а другие - не обладают, например: федеральная или муниципальная собственность, промышленные или сельскохозяйственные предприятия, банковские или коммерческие кредиты и т.д.

Второй вид признаков принимает множество количественных значений, например: объем розничной торговли, объем платных услуг, объем издержек обращения и т.д.

- по роли, которую играют признаки во взаимосвязи исследуемых совокупностей, они делятся на: факторные и результативные.

Факторными называются признаки, под воздействием которых изменяются другие (результативные) признаки.

Результативными называют признаки, которые изменяются в результате изменения факторных признаков. Следует иметь в виду, что в зависимости от условий проведения группировки и целей исследования признаки могут меняться ролями.

3.7. Простые и комбинационные группировки

В зависимости от степени сложности изучаемого явления и от поставленных задач статистические группировки могут выполняться по одному или нескольким группировочным признакам. По способу построения группировки бывают *простые и комбинационные*.

Простой (одномерной) называется группировка, в которой однородные группы образованы только по одному признаку одновременно. Например, распределение населения по возрастным группам, а семей по уровню дохода и т.д.

Алгоритм простой группировки:

1) Совокупность упорядочивается по значению группировочного признака.

2) Определяется число групп (n).

3) Единицы с одинаковыми или близкими значениями признака объединяются в группы.

4) Подсчитываются итоги по группам (число единиц совокупности и значений обобщающих показателей).

5) Результаты группировки оформляются в таблице.

Группировка по двум или нескольким признакам называется *сложной*.

Сложные группировки подразделяются на:

комбинационные:

если группы, образованные по одному признаку, делятся на подгруппы по второму, а последние – на подгруппы по третьему и т.д. признакам, т.е. в основании группировки лежит несколько признаков, взятых в комбинации, то такая группировка называется *комбинационной*.

- *многомерные* – осуществляются не последовательно по отдельным признакам, а одновременно по комплексу признаков.

При этом могут быть использованы два основных подхода:

1) каждая единица совокупности, характеризующаяся набором (из m) признаков, рассматривается как точка в m -мерном пространстве. Множество точек (единиц совокупности) разделяется на однородные группы. Мерой близости точек (сходства единиц совокупности) могут служить различные критерии.

2) заключается в расчете обобщающего показателя по комплексу группировочных признаков и проведении простой группировки по этому обобщающему показателю.

Комбинационные группировки дают возможность изучить единицы совокупности одновременно по нескольким взаимосвязанным 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.

Использование в статистических исследованиях ЭВМ и статистической теории распознавания образов позволило разработать метод группировки совокупности единиц одновременно по множеству характеризующих признаков. Такие группировки получили название *многомерных*. *Многомерная группировка* основана на измерении сходства или различия между объектами: единицы, отнесенные к одной группе различаются между собой меньше, чем единицы, отнесенные к различным группам. Мерой сходства между объектами могут служить различные критерии. Группы (кластеры) формируются на основании близости объектов одновременно ко всему комплексу признаков, описывающих объект. Нахождение этих групп осуществляется методами кластерного анализа на ЭВМ.

3.8. Интервалы группировок

При составлении структурных группировок на основе варьирующих количественных признаков необходимо определить количество групп.

Важным этапом построения статистической группировки является также определение интервала.

Интервал – это количественные значения варьирующего признака, лежащие в определенных границах и отделяющие одну единицу (группу) от другой.

Величина интервала – это разница между максимальным и минимальным значениями признака в каждой группе.

Интервал определяет количественные границы групп, что для статистической практики имеет большое значение, особенно когда нужно образовать качественно однородные группы. Например, исследуется совокупность предприятий по выполнению коллективных договоров. Здесь нельзя объединять предприятия, которые не выполнили обязательства, и те, которые их перевыполнили. Показатель здесь – величина интервала.

Другим примером является невозможность образовывать группу 95-105%, поскольку это разные части совокупности. Следует образовать две группы: 95 – 100% и 101 – 105%. В этом случае границы, по которым различаются совокупности, абсолютно соблюдаются.

Каждый интервал имеет верхнюю и нижнюю границы или одну из них. *Нижней границей* интервала называется наименьшее значение признака в интервале. *Верхней границей интервала* называется наибольшее значение признака в интервале.

Если у интервала указана только одна граница (у первого – верхняя, у последнего – нижняя), то речь идет об *открытых интервалах*. Если у интервала имеются и нижняя, и верхняя границы, то речь идет о *закрытых интервалах*.

Например, группы страховых компаний по числу работающих в них сотрудников (чел.): до 50, 50-100, 100-150, 150 и более.

До 50 и 150 и более – открытые интервалы, остальные закрытые интервалы.

Закрытые интервалы подразделяются на равные и неравные.

Равные интервалы применяются в том случае, когда разность между максимальным и минимальным значениями в каждом из интервалов одинакова. *Неравные интервалы* — когда величина интервала от одного к другому изменяется: может либо расти, либо убывать, либо меняться другим образом. Как правило, неравные интервалы отражают структуру совокупности или характеризуют границы изменения каждого типа.

В свою очередь неравные интервалы подразделяются на прогрессивно возрастающие, прогрессивно убывающие, произвольные и специализированные.

Группировку с равными интервалами строят тогда, когда исследуются количественные различия в величине признака внутри групп одинакового качества, а также, если распределение носит более или менее равномерный характер.

Вопрос о числе групп и величине интервала следует решать с учетом множества обстоятельств, прежде всего исходя из целей исследования, значения изучаемого признака.

Количество групп и величина интервала связаны между собой: чем больше образовано групп, тем меньше интервал, и наоборот. Количество групп зависит от числа единиц исследуемого объекта и степени колеблемости группировочного признака. При небольшом объеме совокупности нельзя образовывать большое число групп, так как группы будут малочисленными.

При определении количества групп необходимо стремиться к тому, чтобы были учтены особенности изучаемого явления. Поэтому число групп должно быть оптимальным, в каждую группу должно входить достаточно большое число единиц совокупности, что отвечает требованию закона больших чисел. Однако в отдельных случаях интерес могут представлять и малочисленные группы, поэтому задача статистики – выделить эти факты и изучить их.

Ориентировочно определить оптимальное количество групп с равными интервалами можно по формуле американского ученого Стерджесса:

$$n = 1 + 3,322 \lg N,$$

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

Для определения числа групп при невозможности вычисления логарифма числа можно воспользоваться таблицей (для равномерного распределения и для групп с равными интервалами).

Получаем следующее соотношение:

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

Формула Стерджесса пригодна при условии, что распределение единиц совокупности по данному признаку приближается к нормальному и при этом применяются равные интервалы в группах. Чтобы получить группы, адекватные действительности, необходимо руководствоваться сущностью изучаемого явления.

Величина равного интервала определяется по следующей формуле:

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

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

В статистической практике чаще применяются неравные интервалы (постепенно возрастающие или постепенно убывающие). При этом исследуемая совокупность делится на группы примерно равного заполнения

с большим числом единиц. Неравные интервалы могут использоваться, например, в таких случаях:

а) при исследовании группировки с применением нескольких признаков, дающих возможность составить несколько подгрупп, где требуются уже и более длинные и более короткие интервалы;

б) при образовании крупных групп с новым качеством на базе мелких групп при условии сохранения их однородности, что приводит к увеличению интервалов.

Пример неравных интервалов: численность работающих промышленных предприятий можно разбить на следующие группы: до 100 человек, 100-200, 200-300, 300-500, 500-1000, 1000 и более человек. Это объясняется тем, что количественные изменения размера признака имеют неодинаковые значения в низших и высших по размеру признака группах: изменение количества работающих из 50-10 человек имеет существенное значение для малых предприятий, а для крупных не имеет.

В статистической практике используются также специализированные интервалы. Интервалы называют специализированными, если речь идет об установлении границ интервала в группах, схожих по типу и по признаку, но имеющие отношение, скажем, к разным отраслям производства.

3.9. Классификации, применяемые в статистике

Особым видом группировок являются классификации.

Классификацией называется процесс разделения множества объектов определенной совокупности на подмножества по их сходству или различию в соответствии с принятыми методами. Имеется также другое определение классификации. Под классификацией понимается систематизация, упорядоченное расположение объектов соответствующего множества по общности какого-либо одного или нескольких признаков, отражающих общие свойства, характеристики или параметры объектов данного множества, его отдельных подмножеств или группировок. Под классификацией понимается также систематизированное распределение явлений и объектов на определенные группы, классы, разряды на основании их сходства и различия.

Признаком классификации является свойство или характеристика объекта классифицируемого множества, по которому производится классификация. Количество значений признака определяет число образуемых классификационных группировок по нему.

Отличительными особенностями классификации являются:

- в основу классификации закладывается качественный признак;
- классификации являются устойчивыми и остаются неизменными в течение длительного периода времени;
- классификации постоянно актуализируются при появлении новых объектов;

- классификации определяют важнейшие признаки статистической группировки.

Объектами классификации служат различные технико-экономические и социальные элементы и их свойства. В качестве объектов классификации выступают: виды экономической деятельности; предприятия и организации всех форм собственности и организационно-правовых форм; промышленная, сельскохозяйственная и строительная продукция; органы государственной власти и управления; административно-территориальные единицы; основные фонды; гидроэнергетические ресурсы; полезные ископаемые; профессии рабочих и должности служащих; технико-экономические и социальные показатели и др.

Применение классификаторов позволяет однозначно идентифицировать и кодировать объекты и признаки их характеризующие, что обеспечивает интеграцию государственных информационных ресурсов. Такая интеграция позволяет решать задачи, связанные с состоянием и динамикой развития национальной экономики, развитием демографической и социальной статистики, созданием статистических регистров и реестров в системах налогообложения, лицензирования, квотирования, государственной регистрации и учета хозяйствующих субъектов, формирования каталогов выпускаемой продукции и др.

Классификатором называется систематизированный свод конкретного множества группировок или объектов, классифицируемых по соответствующим признакам, оформленный в виде нормативного документа-стандарта.

Однозначная идентификация объектов классифицируемого множества и признаков классификации позволяет объективно отразить ту или иную сторону существующих отношений между экономическими и общественными явлениями, обеспечить сопоставимость показателей по качественным и количественным признакам. Создание каждого классификатора предусматривает решение следующих вопросов:

- определение множества объектов классификации;
- разработку структуры построения классификатора;
- определение целей и задач, возлагаемых на проектируемый классификатор;
- выбор методов классификации и кодирования информации.

Кодирование - это процесс присвоения кодовых обозначений соответствующим классифицируемым группировкам, признакам или объектам классификации, обеспечивающий их однозначную идентификацию.

Кодовым обозначением называют систему условных обозначений элементов информационной совокупности в виде знака или группы знаков, выраженных цифрами, буквами или символами.

Главной задачей кодирования является обеспечение однозначности в обозначении объектов системы, а также достижение необходимой достоверности информации.

К основным функциям кодирования следует отнести:

- минимизацию объема информации при ее вводе, хранении и передаче;
- группировку и поиск показателей экономической информации по ключевым признакам;
- разработку аналитических таблиц по различным признакам;
- декодирование информации при формировании выходной информации.

Для осуществления процесса кодирования необходимо придерживаться следующих требований:

- 1) охват всех объектов, подлежащих кодированию;
- 2) присвоение однозначных обозначений каждому объекту кодирования;
- 3) обеспечение возможного расширения совокупности кодируемых объектов с сохранением правил их обозначений;
- 4) использование минимальной значности кода при максимальной его информативности;
- 5) обеспечение удобства восприятия и запоминания кодовых обозначений;
- 6) обеспечение возможности автоматического контроля кодов.

Выбор соответствующего метода классификации и кодирования информации, структуры классификатора зависит от характера и специфики объектов рассматриваемого множества, состава и специфики признаков классификации, назначения и практического использования проектируемого классификатора. Разработка системы классификации и кодирования информации является сложным многоэтапным процессом. К основным этапам относятся:

- установление перечня и количества объектов, подлежащих кодированию;
- систематизация объектов по определенным классификационным признакам;
- выбор системы кодирования информации;
- разработка кодовых обозначений и положений по их ведению и внесению в них изменений.

На первом этапе с целью установления перечня и количества объектов, подлежащих кодированию, определяются объекты и совокупности, которые необходимо кодировать. К ним относятся, например, номенклатура промышленной и сельскохозяйственной продукции, виды экономической деятельности, предприятия и организации различных форм собственности, предприниматели, осуществляющие коммерческую деятельность и т.д. Такая работа проводится на этапе изучения состава экономической информации. Завершающим результатом первого этапа разработки системы классификации и кодирования информации является перечень всех позиций по каждой номенклатуре, который в дальнейшем должен быть закодирован.

На втором этапе осуществляется систематизация составленных перечней и номенклатур по определенным экономическим признакам на основе выбранной системы классификации. Таким образом, осуществляется упорядоченное расположение классифицируемых элементов на основе установленных взаимосвязей между признаками.

3.10. Метод вторичной группировки

В том случае, когда проведенная группировка не позволяет решить поставленные задачи, прибегают к перегруппировке, т.е. ко вторичной группировке.

Процесс образования новых групп на основе ранее осуществленной группировки, называется *вторичной группировкой*.

Необходимость во вторичной группировке возникает в случаях:

- когда в результате первоначальной группировки нечетко проявился характер распределения изучаемой совокупности (в этом случае производят укрупнение или уменьшение интервалов);
- когда требуется сопоставить между собой данные, имеющие различное число выделенных групп или неодинаковые границы интервалов.

Получение новых групп на основе имеющихся возможно двумя способами перегруппировки: объединение первоначальных интервалов (путем их укрупнения) и долевой перегруппировкой (на основе закрепления за каждой группой определенной доли единиц совокупности).

3.11. Статистические ряды распределения

Среди простых группировок особо выделяют *ряды распределения* – это группировка, в которой для характеристики групп применяется только один показатель. После определения группировочного признака и границ групп строится ряд распределения.

Статистический ряд распределения - упорядоченное распределение единиц изучаемой совокупности на группы по определенному варьирующему признаку. Он характеризует состав (структуру) изучаемого явления, позволяет судить об однородности совокупности, закономерности распределения и границах варьирования единиц совокупности.

Ряды распределения, построенные по атрибутивным (качественным) признакам, называются *атрибутивными* (распределение населения по полу, занятости, национальности, профессии и т.д.).

Ряды распределения, построенные по количественному признаку, называются *вариационными* (распределение населения по возрасту, рабочих – по стажу работы, зарплате и т.д.).

Вариационные ряды распределения состоят из двух элементов: вариантов и частот.

Варианты – отдельные числовые значения количественного признака, которые он принимает в вариационном ряду распределения. Так, при группировке предприятий по результатам хозяйственной деятельности варианты положительные (прибыль) и отрицательные (убыток) числа.

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 называется объемом совокупности и определяет число элементов всей совокупности.

Частоты – это частоты, выраженные в относительных величинах: в долях единиц или в % к итогу. Сумма частот равна единице или 100%.

Вариационные ряды в зависимости от характера вариации подразделяются на *дискретные и интервальные*.

Дискретные вариационные ряды основаны на дискретных (прерывных) признаках, имеющих только целые значения (тарифный разряд рабочих, число детей в семье), на дискретных признаках, представленных в виде интервалов.

Интервальные вариационные ряды основаны на непрерывных признаках (имеющих любые значения, даже дробные).

При наличии достаточно большого количества вариантов значений признака первичный ряд является трудно обозримым и непосредственное рассмотрение его не дает представления о распределении единиц по значению признака в совокупности. Поэтому первым шагом в упорядочении первичного ряда является его *ранжирование*, т.е. расположение всех вариантов в возрастающем или убывающем порядке.

3.12. Статистические таблицы. Их виды и принципы построения

Результаты сводки и группировки излагаются в виде таблиц. Таблица – рациональная, наглядная и компактная, краткая и последовательная форма изложения полученных цифровых данных.

Как заметили И. Ильф и Е. Петров, «сама жизнь смотрит на нас со статистических таблиц».

С помощью таблицы статистический материал излагается наиболее рационально.

Статистическая таблица – форма рационального и наглядного изложения цифровых характеристик исследуемых явлений, являющаяся итогом сводки первоначальной информации.

Таблица характеризует совокупность по одному или нескольким признакам, взаимосвязанным логикой.

Основанием любой таблицы является сетка-скелет, в которой вертикальные столбцы называются графами, а горизонтальные — строками. Внешне таблицы представляют собой сетку из вертикальных и горизонтальных линий, в которой записываются числовые данные. Графы и строки образуют макет таблицы.

Любую статистическую таблицу можно представить в виде следующего макета

Сказуемое Подлежащее	Заголовки граф			
ИТОГО				

Любая таблица состоит из 3 составляющих:

1. заголовок, в котором указывается основная цель или содержание таблицы и часто время и место, к которому относятся излагаемые в ней данные;

2. *статистическое подлежащее* – это единицы статистической совокупности или их группы (то, о чем говорится и что характеризуется в таблице – объект изучения);

3. *статистическое сказуемое* – это цифровой материал, которым характеризуется подлежащее (то, что говорится о подлежащем, его характеристика с помощью системы показателей, то, какими признаками характеризуется подлежащее).

Подлежащее обычно располагается в левой части таблиц; сказуемое – в верхней части таблицы в виде названия граф.

В зависимости от расположения подлежащего и группировки в нем признаков таблицы бывают простые и сложные

В *простой таблице* дается простой перечень объектов.

В простой таблице подлежащее не делится на группы. Простые таблицы бывают перечневые, динамические, территориальные.

Простые таблицы дают только итоговую сводку и не дают возможности выявить социально-экономические типы изучаемых явлений, их структуру, а также взаимосвязи и взаимозависимости между характеризующими их признаками.

Эти задачи более полно могут быть решены с помощью сложных - групповых и, особенно, комбинационных таблиц.

Сложная таблица содержит группировку единиц совокупности одновременно по 2-м и более признакам.

Групповыми называются таблицы, в которых подлежащее, т.е. объект исследования, подразделяется на группы по какому-либо одному признаку.

Групповыми называются таблицы, имеющие в подлежащем группировку единиц совокупности по одному признаку.

Комбинационными называются такие таблицы, в которых подлежащее делится на группы не по одному, а по нескольким признакам, причем каждая группа, образованная по одному признаку, делится на подгруппы по другому признаку.

По характеру разработки показателей сказуемого различают:

- таблицы с простой разработкой показателей сказуемого, в которых имеет место параллельное расположение показателей сказуемого.

- таблицы со сложной разработкой показателей сказуемого, в которых имеет место комбинирование показателей сказуемого: внутри групп, образованных по одному признаку, выделяют подгруппы по другому признаку.

В зависимости от этапа статистического исследования таблицы делятся на:

- *разработочные* (вспомогательные), цель которых обобщить информацию по отдельным единицам совокупности для получения итоговых показателей.

- *сводные*, задача которых показать итоги по группам и всей совокупности в целом.

- *аналитические* таблицы, задача которых — расчет обобщающих характеристик и подготовка информационной базы для анализа и структуры и структурных сдвигов, динамики изучаемых явлений и взаимосвязей между показателями.

При построении таблиц следует соблюдать следующие правила:

- таблица должна иметь небольшие размеры, чтобы ее удобно было читать и анализировать, то есть быть компактной;

- название таблицы, заголовки подлежащего и сказуемого должны быть точными, краткими и ясными;

- в таблице должны быть точно обозначены единицы измерения, а также территория и период, к которым относятся приводимые данные;

- информация в столбцах и графах должна завершаться итоговой строкой; т.е. должны быть подсчитаны итоги;

- каждая клетка таблицы должна быть заполнена, при отсутствии данных следует ставить знак тире, а при отсутствии сведений — многоточие или «нет сведений», если клетка не подлежит заполнению, ставится знак (х);

- цифровой материал должен даваться с одинаковой степенью точности.

3.13. Статистические графики

Практически любой пакет прикладных программ, предназначенный для статической обработки данных на компьютере, содержит графические методы представления данных.

Графический метод – это продолжение и дополнение табличного метода. Если при чтении таблицы что-то остается незамеченным, то обнаруживается на графике. Статистические графики показывают общую картину изучаемого явления, дают его обобщенное представление. При графическом изображении статистических данных становится более выразительной сравнительная характеристика изучаемых показателей, отчетливее проявляется тенденция развития изучаемого явления, лучше видны основные взаимосвязи.

Графический метод – это метод условных изображений статистических данных при помощи геометрических фигур, линий, точек и разнообразных символических образов.

Применение графиков в статистике насчитывает более чем двухсотлетнюю историю. Основателем графического метода в статистике коммерческой деятельности считают английского экономиста У. Плейфейра (1731 – 1798 гг.). В его работе «Коммерческий и политический атлас» (1786 г.) впервые были применены способы графического изображения статистических данных (линейные, столбиковые и другие диаграммы).

Статистический график – чертеж, на котором статистические совокупности, характеризуемые определенными показателями описываются с помощью условных геометрических образов или знаков.

При построении графика необходимо соблюдать следующие требования: наглядность, выразительность, понятность.

Любой график состоит из графического образа и вспомогательных элементов.

В статистическом графике существуют следующие основные элементы: поле графика, графический образ, пространственные и масштабные ориентиры, экспликация графика.

Поле графика – часть плоскости, где расположены графические образы. Это листы бумаги, географические карты, план местности. Поле графика характеризуется его форматом. Размер поля графика зависит от его назначения. Стороны поля статистического графика обычно находятся в определенной пропорции. Предназначенные для научного изучения графики изготавливаются в формате чертежного листа, т. е. размером 814x1152 мм, небольшие графики – 203x288 мм. Для зрительного восприятия наиболее приемлем график, выполненный на поле прямоугольной формы с соотношением сторон от 1:1,3 до 1:1,5. Размер поля графика и пропорции его сторон определяются исполнителем.

Графический образ – это совокупность символических знаков (точек, линий, геометрических фигур), с помощью которых изображаются статистические данные.

Эти знаки образуют основу графического изображения. В графиках также возможно использование негеометрических фигур в виде силуэтов или рисунков предмета.

При построении графика важен правильный подбор графического образа, который должен доходчиво отображать изучаемые показатели и соответствовать основному предназначению графика.

Пространственные ориентиры определяют размещение графических образов на поле графика. Они задаются координатной сеткой или контурными линиями и делят поле графика на части, соответствующие значениям изучаемых показателей.

В статистических графиках чаще всего применяется система прямоугольных (декартовых) координат. Но есть и графики, построенные по принципу полярных координат (круговые графики).

Масштабные ориентиры графика придают графическим образам количественную значимость, которая передается с помощью системы масштабных шкал.

Масштаб графика – это мера перевода численной величины в графическую и наоборот. Чем длиннее отрезок линии, принятой за числовую единицу, тем крупнее масштаб.

Масштабной шкалой является линия, отдельные точки которой читаются как определенные числа. В масштабной шкале различают линию – носитель информации – опора шкалы.

Шкала графика может быть прямолинейной и криволинейной. Различаются также шкалы равномерные и неравномерные.

Шкала является равномерной, если равным графическим отрезкам соответствуют равные числовые величины. Равномерные арифметические шкалы используются при построении большинства статистических графиков.

Неравномерным шкалам соответствуют неравные числовые значения. Шкала, по которой отсчитываются уровни изучаемых показателей, начинается с 0.

Экспликация графика – это словесное пояснение его содержания, включает в себя заголовок графика, объяснения масштабных шкал, объяснения отдельных элементов графического образа.

Чтобы понять график, дается пояснение знаков, масштаб и приводится наименование графика. Заголовок графика – это четкое пояснение основного содержания изображаемых данных.

Контрольные вопросы

1. Что такое сводка.
2. Группировка сводок от глубины обработки первичной информации.
3. Группировка сводок по признаку места проведения сводки.
4. Перечислите этапы сводок.
5. Понятие статистической группировки.
6. Задачи, решаемые на основе группировок.
7. Типологические группировки.
8. Структурные группировки.
9. Аналитические группировки.
10. Понятие группировочных признаков.
11. Атрибутивные признаки.
12. Количественные признаки.
13. Дискретные признаки.
14. Непрерывные признаки.

- 15.Альтернативные признаки.
- 16.Факторные признаки.
- 17.Результативные признаки.
- 18.Алгоритм простой группировки.
- 19.Сложные группировки.
- 20.Понятие интервала.
- 21.Нижняя и верхняя границы интервала.
- 22.Открытые и закрытые интервалы.
- 23.Равные и неравные интервалы.
- 24.Понятие классификации.
- 25.Признак классификации.
- 26.Объект классификации.
- 27.Понятие классификатора.
- 28.Понятие кодирования.
- 29.Метод вторичной группировки.
- 30.Статистический ряд распределения.
- 31.Понятие варианты.
- 32.Понятие частоты.
- 33.Статистическая таблица.
- 34.Составляющие таблицы.
- 35.Сказуемое таблицы.
- 36.Подлежащее таблицы.
- 37.Простая таблица.
- 38.Сложная таблица.
- 39.Правила построения таблиц.
- 40.Понятие графического метода.
- 41.Назовите основоположника графического метода.
- 42.Понятие статистического графика.
- 43.Что такое поле графика.
- 44.Что такое графический образ.
- 45.Масштаб графика.
- 46.Экспликация графика.

Тема 4. Статистические показатели

4.1. Статистический показатель как количественная характеристика социально-экономических явлений в единстве с их качественной определенностью

В итоге сводки статистических данных получают обобщающие статистические показатели, в которых отображаются результаты познания количественной стороны массовых общественных явлений.

Статистический показатель – это обобщающая количественно-качественная характеристика некоторого свойства статистической совокупности или ее части. Этим он отличается от признака (т.е. свойства, присущего единице совокупности). Например, средняя ожидаемая продолжительность жизни родившегося в 2010 году в РБ поколения людей – статистический показатель. Продолжительность жизни конкретного человека – признак. Или средний возраст студента в группе на начало учебного года – статистический показатель и возраст конкретного студента данной группы – признак.

Величина – характеристика объекта или явления материального мира, общая в качественном отношении, но индивидуальная для каждого из них в количественном отношении.

Значение конкретной величины – это ее оценка, выражаемая произведением отвлеченного числа на принятую для данной величины единицу. Значение показателя является функцией пространства и времени.

Статистический показатель характеризует абсолютный, относительный и средний размер социально-экономического явления.

Абсолютный размер явления – это величина, взятая сама по себе безотносительно к размерам других явлений.

Относительный размер явления – это соотношение величины данного явления с величиной какого-либо другого явления или величиной того же самого явления, но взятого за другое время или по другой территории.

Средний размер явления – выражает типичные черты и дает обобщенную характеристику однотипных явлений по одному из варьирующих признаков.

Выделяют следующие 4 атрибута статистического показателя: качественную сторону, количественную сторону, пространственные границы, временные границы.

Качественная сторона (основание или содержание) отражает сущность изучаемого свойства статистической совокупности без указания места, времени и возможности определения числового значения. Определяется она понятиями, входящими в наименование показателя и связана с функцией, которую выполняет показатель.

Понятия, входящие в наименование показателя можно разделить на 2 группы: а) понятия чисто статистические и б) понятия, являющиеся

предметом изучения других областей знания. Например, в показателе: средняя ожидаемая продолжительность жизни родившегося в 2010 году в РБ поколения людей содержание определяется понятиями: средняя величина (статистическое понятие) и продолжительность жизни (понятие демографии).

Количественная сторона включает методологию расчета (формулу), число и единицу измерения.

Пространственные границы представляют территориальные, отраслевые и иные границы статистического показателя.

Временные границы - это интервал или момент времени.

Чтобы статистические показатели правильно отражали изучаемые явления, при их построении следует соблюдать следующие правила:

1) опираясь на положения экономической теории, а также на статистическую методологию и опыт статистических работ, следует стремиться к тому, чтобы показатели выражали сущность изучаемых явлений и давали им точную количественную оценку;

2) следует добиваться полноты информации как по охвату единиц изучаемого объекта, так и по комплексному отображению всех сторон протекаемого процесса;

3) следует обеспечивать сравнимость статистических показателей. Условия сопоставимости показателей состоит в том, что при сопоставлении показатели должны отличаться не более чем одним атрибутом. Нельзя, например, сопоставлять (сравнивать) показатель добычи угля в США в 2019 г. с выплавкой стали в РБ в 2020 г.;

4) повышать степень точности исходной статистической информации, на основе которой исчисляются показатели.

4.2. Классификация показателей

Система статистических показателей - это совокупность взаимосвязанных показателей, имеющая одноуровневую или многоуровневую структуру, направленная на решение конкретной статистической задачи.

В основу классификации статистических показателей положены *следующие признаки*:

По форме выражения – абсолютные, относительные и средние.

По признаку пространственной определенности статистические показатели подразделяются на общетерриториальные, региональные и местные.

По характеру их выражения: описательные – выражаются словесно; количественные – числами.

По способу измерения: первичные – можно сосчитать, измерить, учесть; вторичные – рассчитываются на основе первичных.

По отношению к характеризующему объекту:

прямые – присущи непосредственно объекту, который характеризуют;

косвенные – описывают объект, имеющий отношение к изучаемому объекту.

По характеру вариации: альтернативные – принимают одно из двух или нескольких значений (обладания или не обладания каким-нибудь свойством); дискретные – количественные, принимающие отдельные, иногда только целочисленные значения; непрерывные – непрерывно изменяющиеся в определенных границах (интервалах) числовые признаки или показатели.

По отношению ко времени: моментные – характеризуют объект в определенный момент времени (например, среднесписочная численность работников за месяц); интервальные – характеризуют результаты процесса за определенный интервал времени (например, численность персонала на 01.01.ГГГГ).

Различают также *плановые, отчетные и оценочные показатели.*

Плановые показатели представляют собой конкретные задания по развитию народного хозяйства в целом или отдельных организаций на определенные периоды времени.

Отчетные показатели характеризуют реально сложившийся уровень экономического и социального развития, фактически достигнутый за определенный период.

Оценочные показатели представляют собой результат прогнозного расчета развития того или иного явления.

По сущности изучаемых явлений выделяют объемные и качественные статистические показатели.

Объемные показатели характеризуют размеры явления или процесса, рассчитываются путем суммирования. Выделяют 2 вида объемных показателей: 1) показатели объема совокупности (например, общая численность студентов вузов), 2) показатели объема признака в совокупности (объем выпускаемой продукции за год). Объемные показатели выражаются абсолютными величинами.

Качественные показатели характеризует размер явления или процесса в расчете на количественную единицу (человека, единицу объема выпуска и т.п.). Они измеряют не общий объем явления или процесса, а их интенсивность, эффективность. Как правило, они являются средними или относительными величинами. Например, цена, себестоимость, трудоемкость единицы продукции, производительность труда, средняя зарплата рабочих и т.п.

По степени агрегирования явлений, охвату единиц совокупности статистические показатели подразделяют на единичные, частные и сводные.

Индивидуальные (единичные) показатели характеризуют отдельные единичные процессы (изменение элементов сложного явления), размеры признака у отдельных единиц совокупности. Например, размер заработной платы отдельного работника, численность работников предприятия.

Частные показатели характеризуют части совокупности. Например, численность населения определенной области.

Сводные или общие статистические показатели характеризуют совокупность целиком, выражают размеры, величину того или иного признака у всех единиц данной совокупности. Например, численность населения республики.

Различают также *синтетические и аналитические* показатели.

Синтетические показатели – содержат обобщающие данные об объекте, характеризуют экономический объект, экономическую систему в целом, в органическом единстве ее частей.

Аналитические показатели – подробно раскрывают и конкретизируют обобщенные сведения. Это расчетные величины, критерии, уровни, измерители.

4.3. Измерительные шкалы.

При построении статистических показателей используют различные измерительные шкалы. *Измерительная шкала* – система чисел или иных элементов, принятых для измерения или оценки тех или иных величин. В определении шкал участвуют понятия равенства, порядка, расстояния между пунктами (интервалы), начала отсчета и единицы измерения. В зависимости от наличия или отсутствия этих элементов возникают различные типы шкал:

1. *Номинальная шкала (шкала наименований)*. Число на номинальной шкале служит лишь для опознавания, играет роль ярлыка (метки). (Например, свойство «цвет глаз» может принимать следующие значения: серые, карие, зеленые, голубые и т.д., которым ставятся в соответствие следующие числа: 1,2,3,4 и т.д. Еще пример - классификатор отраслей экономики, категорий работников). К таким числам неприменимы обычные правила арифметики. Номинальная шкала обладает только свойствами симметричности и транзитивности. Симметричность означает, что отношения, существующие между градациями x_1 и x_2 , имеют место и между x_2 и x_1 . Транзитивность выражается в следующем: если $x_1=x_2$, и $x_2=x_3$, то $x_1=x_3$. Примером измерения в номинальной шкале является классификация отраслей экономики, категорий работников и т.п.

2. *Порядковая (ординальная или ранговая) шкала*. Это шкала, на которой числа могут быть упорядочены. Однако определить и интерпретировать расстояние между числами на этой шкале невозможно. Например, показатель «качество продукции», принимающий значения: высшая категория (соответствие лучшим отечественным и мировым достижениям), первая категория (соответствие современным требованиям стандартов), вторая категория (морально устаревшая продукция) измеряются в ординальной шкале. Оценки: отлично, хорошо, удовлетворительно, неудовлетворительно также измеряются в ординальной шкале. Шкала порядка допускает операции: «равенство-неравенство», «больше-меньше».

3. *Количественные (метрические) шкалы* подразделяются на интервальные и пропорциональные.

Интервальная шкала (порядковая шкала с интервалом). Эта шкала позволяет не только установить порядок, но и определить интервал между числами. Величина интервала устанавливается по косвенным признакам, или на основе субъективных оценок. Интервальная шкала допускает операции: «равенство-неравенство», «больше-меньше», «равенство-неравенство интервалов» и операцию вычитания, на основе которой устанавливается величина интервала. По интервальной шкале измеряется календарное время, температура.

Пропорциональная шкала (отношений). Представляет собой интервальную шкалу с естественным началом отсчета (абсолютным нулем). Пропорциональная шкала в отличие от предыдущих шкал позволяет выяснить во сколько раз один признак больше или меньше другого. По шкале отношений можно измерить рост, вес, цену и т.д.

Метрические шкалы позволяют выполнять различные арифметические операции: сложение, умножение, деление. Такие шкалы – основа всевозможных статистических операций.

4.4. Условия сопоставимости показателей. Приведение показателей в сопоставимый вид.

Главнейшим требованием статистики является требование обеспечения сопоставимости показателей, так как без сопоставимости нет сравнения, а значит, нет объективных выводов об изучаемом социально-экономическом явлении или процессе.

При этом должны быть учтены следующие требования:

1. единство объемных, стоимостных, качественных, структурных факторов.
2. единство промежутков или моментов времени, за которые были вычислены показатели.
3. сопоставимость исходных условий производства (технических, природных, климатических).
4. единство методики вычисления показателей.

Несопоставимость показателей может быть вызвана различными причинами: разным уровнем цен, объемов деятельности, структурными изменениями, неодинаковыми календарными периодами.

Сравнение несопоставимых показателей приводит к неправильным результатам и выводам. Поэтому прежде чем проводить сравнение показателей необходимо привести их в сопоставимый вид.

В статистике выработаны определённые правила, обеспечивающие сопоставимость показателей:

- *показатели должны обладать общим содержанием*: еще древние говорили, что абсурдно сравнивать “что длиннее - дерево или ночь” или “чего больше – ума или зерна”;

- *статистические показатели должны выражаться в одинаковых единицах измерения: расстояние – в километрах, вес – в килограммах, и т.д.*

Если используются стоимостные измерители, то для обеспечения сопоставимости должны применяться сопоставимые цены – цены базисного (прошлого), либо отчётного периода (данного);

Например: если сравниваются стоимостные объёмы продаж 1995 и 2006 г., то для обеспечения корректности сравнения необходимо физические (натуральные) объёмы продаж выразить либо в ценах 1995 г, либо в ценах 2006 г. Кроме этого, сопоставимость разных по содержанию главного компонента разновидностей продукта может быть обеспечена применением условно-натуральных измерителей;

- *сравниваемые показатели должны рассчитываться по единой методике;*

- *сравниваемые статистические показатели должны быть однородными по времени и территории – они должны определяться за одинаковые периоды времени, на одни и те же даты, по одной территории.*

В соответствии с перечисленными правилами для обеспечения сопоставимости статистических показателей на практике используются следующие статистические приёмы:

- для обеспечения общего содержания – *разделение разнородных совокупностей на однородные части, т.е. группировку;*

- для приведения к одинаковым единицам измерения – *использование единой системы мер и весов, условно-натуральных измерителей, сопоставимых цен или индексов при сравнении стоимостных показателей;*

- *пересчёт несопоставимых показателей по единой методике;*

- *приведение показателей к одинаковым периодам и моментам времени;*

- *приведение показателей к единой территории или кругу охватываемых единиц.*

- *замена несравнимых абсолютных показателей относительными или средними показателями: показателями структуры, координации и т.д.*

Все статистические показатели исчисляются на основе первичных данных, собранных и определённым образом обработанных в процессе статистического исследования.

4.5. Абсолютные величины как исходная форма статистических показателей. Прямые и косвенные методы их измерения.

Статистические показатели могут выражаться абсолютными, относительными, либо средними величинами.

Абсолютными величинами выражаются объёмные статистические показатели. Они являются именованными величинами, имеющими определенную размерность и единицы измерения.

Абсолютные – это суммарные показатели, характеризующие размеры (уровни, объемы) общественных явлений в конкретных условиях места и времени.

В зависимости от целей анализа применяются натуральные, условно-натуральные, стоимостные и трудовые единицы измерения.

Натуральные единицы измерения в большинстве своем соответствуют природным или потребительским свойствам предмета и выражаются в физических мерах веса, длины и т.д. Они могут быть простыми (тонны, штуки, литры, метры) и сложными, являющимися комбинацией нескольких равноименных величин (грузооборот на транспорте выражается в тонно-километрах, пассажиро – километрах, производство электроэнергии – в киловатт-часах).

Наряду с натуральными применяются также *условно-натуральные* единицы для соизмерения разнородных, но взаимозаменяемых по какому-либо свойству объектов, причем мера этого свойства и становится средством соизмерения. Условно-натуральные единицы измерения используются для сведения воедино нескольких разновидностей одной и той же потребительской стоимости. Одну из них принимают за эталон, а другие пересчитываются с помощью специальных коэффициентов в единицы меры этого эталона.

Например, различные виды органического топлива переводят в условное топливо с теплотой сгорания 29,3 мДж/кг (7000 ккал/кг), мыло разных сортов – в условное мыло с 40%-ным содержанием жирных кислот, консервы различного объема – в условные консервные банки объемом 353,4 см³, алкогольные напитки учитываются в дкл 100% спирта, для подсчета общего объема работы транспорта складывают тонно-километры перевезенных грузов и пассажиро-километры, произведенные пассажирским транспортом, условно приравнивая при этом перевозку одного пассажира к перевозке одной тонны груза; общее количество школьных тетрадей измеряется в у.шк.т. — условные школьные тетради размером 12 листов и т.д.

Перевод в условные единицы измерения осуществляется на основе коэффициентов, рассчитываемых как отношение значений, характеризующих взятое за основу свойство по разновидностям продукции, к эталонному значению.

Трудовые единицы измерения позволяют учитывать общие затраты труда на предприятиях, а также трудоемкость отдельных операций технологического процесса. К ним относятся человеко-дни и человеко-часы.

Стоимостные единицы измерения используются для выражения объема разнородной продукции в единой денежной форме – рублях, долларах и т.п. В стоимостных единицах измерения выражают выпуск продукции, доходы населения.

Абсолютные величины могут быть положительными и отрицательными. Например, результат деятельности предприятия (прибыль/убыток).

Для измерения абсолютных величин применяют прямой и косвенный методы измерений.

При *прямом методе измерения* искомая величина находится напрямую не прибегая к арифметическим расчетам:

- либо непосредственным наблюдением. (Пример: счет продукции в штуках, табельный учет численности работающих, хронометраж времени обработки, снятие показателей измерительных приборов).

- либо опросом. (Пример: переписи, оценка спроса на товары, социологическое изучение мотивов поведения, склонностей и т.д.).

При *прямых измерениях* искомая величина определяется непосредственно показаниями прибора или измерительной шкалы инструмента. К прямым измерениям относятся измерения длин линейками, штанген-инструментом, микрометрами, высотомерами, измерения углов - угломерами и др.

При *косвенных измерениях* искомая величина (размер или отклонение) определяется по результатам прямых измерений одной или нескольких величин, связанных с искомой величиной определенной функциональной зависимостью, т. е. после определения косвенных величин, влияющих на искомую, определяют искомую величину, используя математические методы вычислений или преобразований. Примером косвенных измерений могут служить измерения диаметра вала по длине его окружности с помощью рулетки или обкатного ролика, измерения на координатно-измерительных машинах (КИМ), и др.

Абсолютные величины не дают полного представления об изучаемом явлении, не показывают его структуры, развития во времени, соотношения между отдельными частями, с другими абсолютными величинами.

4.6. Относительные величины, их виды. Выбор базы при исчисления относительных величин.

Относительная величина статистики - это обобщающий показатель, который дает числовую меру соотношения двух сопоставимых статистических величин. Относительная величина в статистике представляют собой результат отношения (сравнения) двух абсолютных или относительных величин. В числителе дроби стоит величина, которую сравнивают, а в знаменателе - величина, с которой сравнивают (*база сравнения*) или *основание*. Относительными величинами в статистике выражаются качественные показатели.

Относительные величины, получаемые при сопоставлении абсолютных величин, могут быть названы относительными величинами первого порядка,

а полученные при сопоставлении относительных величин – величинами высших (второго, третьего и т.д.) порядков.

Для выражения результата сопоставления одноименных величин используется:

- *коэффициенты*, если база сравнения принимается за единицу;
- *проценты %*, если база сравнения принимается за сто (умножить на 100);

Проценты используются в тех случаях, когда сравниваемый абсолютный показатель превосходит базисный не более чем в 2-3 раза (или базисный превосходит сравниваемый не более чем в 100 раз, например 174% или 5%). Проценты свыше 200-300 обычно заменяются коэффициентом, так 470% - 4,7 раза.

- *промилле (‰)*, если база сравнения принимается за тысячу (умножить на 1000);

Если базисный показатель превышает сравниваемый более чем в 100 раз, но менее чем в 1000 удобно использовать промилле (тысячную долю). Широко применяется в статистике населения: показатели рождаемости, смертности, брачности и т.п.

- *продецимилле (‱)*, если база сравнения принимается за десять тысяч (умножить на 10000);

Так в расчете на 10000 человек определяется численность студентов ВУЗов, численность врачей и т.п.

- *просантимилле (‱)*, если база сравнения принимается за сто тысяч (умножить на 100000).

При сопоставлении разноименных величин результат выражается сочетаниями наименований сравниваемых величин: производительность труда руб./чел.

Все используемые на практике относительные статистические показатели можно подразделить на 7 видов:

1.) *Относительная величина планового задания (ОВПЗ)* – отношение планового задания текущего (отчетного) периода к фактическому значению показателя в предыдущем периоде.

2.) *Относительная величина выполнения планового задания, реализации плана (ОВВП)* – отношение фактического значения к плановому заданию по нему за один и тот же период.

3.) *Относительная величина динамики (ОВД)* характеризует изменение явления во времени и представляет собой соотношение значений одного и того же показателя за различные периоды времени.

Относительные величины динамики за ряд лет могут быть исчислены как по отношению к предыдущему периоду, так и к одному и тому же базисному году. В первом случае они называются *цепными*, во втором – *базисными*.

Относительные величины динамики, планового задания и выполнения планового задания находятся во взаимосвязи – произведение относительных

величин выполнения плана и планового задания равно относительной величине динамики: $ОВД = ОВПЗ \cdot ОВВП$

4.) *Относительная величина интенсивности и уровня экономического развития (ОВИ)* характеризует степень распространения явления в определенной среде или по отношению к другому показателю. В отличие от всех других относительных величин, это единственная именованная величина, поскольку она представляет собой результат сопоставления разноименных величин и характеризует итог числителя, приходящийся на 1, 10, 100, 1000 и т. д. единиц знаменателя.

Эта величина определяется сопоставлением двух разноименных, но связанных между собой абсолютных величин (например, фондоотдача, фондоемкость, плотность населения на один кв. км, число автомашин на сто семей и т.п.). Наиболее часто данный показатель может быть выражен в процентах, промилле, продецимилле.

Обычно относительный показатель интенсивности рассчитывается в тех случаях, когда абсолютная величина оказывается недостаточной для формулировки обоснованных выводов о масштабах явления, его размерах, насыщенности, плотности распространения. Так, например, для определения уровня обеспеченности населения легковыми автомобилями рассчитывается число автомашин, приходящихся на 100 семей, для определения плотности населения рассчитывается число людей, приходящихся на 1 кв. км.

Разновидностью относительных показателей интенсивности являются относительные показатели уровня экономического развития, характеризующие размеры производства или потребления различных видов продукции в расчете на душу населения и играющие важную роль в оценке развития экономики государства или региона.

5.) *Относительная величина структуры (ОВС_т)* характеризует состав изучаемой совокупности. Относительная величина структуры показывает какую долю (или удельный вес) во всей совокупности составляет отдельная ее часть (например, удельный вес женщин, мужчин, малых предприятий, частных предприятий).

Выражается относительный показатель структуры в долях единицы, то есть коэффициентах или в процентах. Рассчитанные величины, соответственно называемые долями или удельными весами, показывают, какой долей обладает или какой удельный вес имеет та или иная часть в общем итоге.

6.) *Относительная величина сравнения (ОВС_р)* характеризует изменение в пространстве. Это – отношение одного и того же показателя за один и тот же момент времени, но по различным территориям или объектам.

Научная ценность относительных показателей высока, но их нельзя рассматривать в отрыве от абсолютных показателей, соотношения которых они выражают, иначе они не смогут точно характеризовать изучаемые явления. Пользуясь в анализе относительными величинами, необходимо показать, какие абсолютные показатели за ними скрываются.

Контрольные вопросы

1. Понятие статистического показателя.
2. Понятие величины.
3. Абсолютный размер явления.
4. Относительный размер явления.
5. Средний размер явления.
6. Качественная сторона статистического показателя.
7. Количественная сторона статистического показателя.
8. Пространственные границы статистического показателя.
9. Временные границы статистического показателя.
10. Система статистических показателей.
11. Классификация статистических показателей по форме выражения.
12. Классификация статистических показателей по признаку пространственной определенности.
13. Классификация статистических показателей по характеру выражения.
14. Классификация статистических показателей по способу измерения.
15. Классификация статистических показателей по отношению к характеризруемому объекту.
16. Классификация статистических показателей по характеру вариации.
17. Классификация статистических показателей по отношению ко времени.
18. Классификация статистических показателей по сущности изучаемых явлений.
19. Классификация статистических показателей по степени агрегирования явлений, охвату единиц совокупности.
20. Понятие измерительной шкалы.
21. Номинальная шкала.
22. Порядковая шкала.
23. Количественные шкалы.
24. Условия сопоставимости показателей.
25. Понятие абсолютной величины.
26. Натуральные единицы измерения.
27. Трудовые единицы измерения.
28. Стоимостные единицы измерения.
29. Прямой метод измерения.
30. Косвенные измерения.
31. Понятие относительной величины.
32. Выбор базы при исчислении относительных величин.
33. Относительная величина планового задания.

34. Относительная величина выполнения планового задания, реализации плана.
35. Относительная величина динамики.
36. Относительная величина интенсивности и уровня экономического развития.
37. Относительная величина структуры.
38. Относительная величина сравнения.

Тема 5. Средние величины статистики

5.1. Средняя величина, ее сущность и определение как категории статистической науки. Основные научные положения теории средних величин

Признаки единиц статистических совокупностей различны по своему значению, например, заработная плата одной профессии какого-либо предприятия не одинакова за один и тот же период времени, различны цены на рынке на одинаковую продукцию, урожайность сельскохозяйственных культур в хозяйствах района и т.д. Поэтому, чтобы определить значение признака, характерное для всей изучаемой совокупности единиц, рассчитывают средние величины.

Средней величиной в статистике называется обобщающий показатель, характеризующий типичный уровень явления в конкретных условиях места и времени, отражающий величину варьирующего признака в расчете на единицу качественно однородной совокупности.

Совокупность, изучаемая по количественному признаку, состоит из индивидуальных значений, на них оказывают влияние, как общие причины, так и индивидуальные условия. В среднем значении отклонения, характерные для индивидуальных значений, погашаются. Средняя, являясь функцией множества индивидуальных значений, представляет одним значением всю совокупность и отражает то общее, что присуще всем ее единицам.

Средняя, рассчитываемая для совокупностей, состоящих из качественно однородных единиц, называется типической средней. Например, можно рассчитать среднюю заработную плату работника той или иной профессиональной группы (слесаря, врача, библиотекаря). Разумеется, уровни месячной заработной платы слесарей в силу различия их квалификации, стажа работы, отработанного за месяц времени и многих других факторов отличаются друг от друга и от уровня средней заработной платы. Однако в среднем уровне отражены основные факторы, которые влияют на уровень заработной платы, и взаимно погашаются различия, которые возникают вследствие индивидуальных особенностей работника.

Средняя заработная плата отражает типичный уровень оплаты труда для данного вида работников. Получению типической средней величины должен предшествовать анализ того, насколько данная совокупность качественно однородна. Если совокупность состоит из отдельных частей, следует разбить ее на типические группы (средняя температура по больнице).

Средние величины, используемые в качестве характеристик для неоднородных совокупностей, называются системными средними. Например, средняя величина валового внутреннего продукта (ВВП) на душу населения, средняя величина потребления различных групп товаров на человека и

другие подобные величины, представляющие обобщающие характеристики государства как единой экономической системы.

Средняя должна вычисляться для совокупностей, состоящих из достаточно большого числа единиц. Соблюдение этого условия необходимо для того, чтобы вошел в силу закон больших чисел, в результате действия которого случайные отклонения индивидуальных величин от общей тенденции взаимно погашаются.

В анализе изучаемых явлений роль средних величин огромна. Английский экономист В. Петти (1623—1687 гг.) широко использовал средние величины. В. Петти хотел использовать средние величины в качестве меры стоимости расходов на среднее дневное пропитание одного работника. Устойчивость средней величины – это отражение закономерности изучаемых процессов. Он считал, что информацию можно преобразовать, даже если нет достаточного объема исходных данных.

Применял средние и относительные величины английский ученый Г. Кинг (1648—1712) при анализе данных о населении Англии.

Теоретические разработки бельгийского статистика А. Кетле (1796—1874 гг.) основаны на противоречивости природы социальных явлений – высокоустойчивых в массе, но сугубо индивидуальных.

Согласно А. Кетле постоянные причины действуют одинаково на каждое изучаемое явление и делают эти явления похожими друг на друга, создают общие для всех них закономерности.

Следствием учения А. Кетле явилось выделение средних величин в качестве основного приема статистического анализа. Он говорил, что статистические средние величины представляют собой не категорию объективной действительности.

А. Кетле выразил взгляды на среднюю величину в своей теории среднего человека. Средний человек – это человек, обладающий всеми качествами в среднем размере (средняя смертность или рождаемость, средний рост и вес, средняя быстрота бега, средняя склонность к браку и самоубийству, к добрым делам и т. д.). Для А. Кетле средний человек – это идеал человека. Несостоятельность теории среднего человека А. Кетле была доказана в русской статистической литературе в конце XIX—XX вв.

Известный русский статистик Ю. Э. Янсон (1835—1893 гг.) писал, что А. Кетле предполагает существование в природе типа среднего человека как чего-то данного, от которого жизнь отклонила средних людей данного общества и данного времени, а это приводит его к совершенно механическому взгляду и на законы движения социальной жизни: движение – это постепенное возрастание средних свойств человека, постепенное восстановление типа; следовательно, такое нивелирование всех проявлений жизни социального тела, за которым всякое поступательное движение прекращается.

Сущность данной теории нашла свое дальнейшее развитие в работах ряда теоретиков статистики как теория истинных величин. У А. Кетле были

последователи – немецкий экономист и статистик В. Лексис (1837—1914 гг.), перенесший теорию истинных величин на экономические явления общественной жизни. Его теория известна под названием теория устойчивости. Другая разновидность идеалистической теории средних величин основана на философии

Ее основатель – английский статистик А. Боули (1869– 1957гг.) – один из самых видных теоретиков новейшего времени в области теории средних величин. Его концепция средних величин изложена в книге «Элементы статистики».

А. Боули рассматривает средние величины лишь с количественной стороны, тем самым отрывает количество от качества. Определяя значение средних величин (или «их функцию»), А. Боули выдвигает махистский принцип мышления. А. Боули писал, что функция средних величин должна выражать сложную группу с помощью немногих простых чисел. Статистические данные должны быть упрощены, сгруппированы и приведены к средним Эти взгляды: разделяли Р. Фишер (1890—1968 гг.), Дж. Юл (1871 – 1951 гг.), Фредерик С. Миллс (1892 г) и др.

В 30—е гг. XX в. и последующие годы средняя величина рассматривается как социально значимая характеристика, информативность которой зависит от однородности данных.

Виднейшие представители итальянской школы Р. Бенини (1862—1956 гг.) и К. Джини (1884—1965 гг.), считая статистику отраслью логики, расширили область применения статистической индукции, но познавательные принципы логики и статистики они связывали с природой изучаемых явлений, следуя традициям социологической трактовки статистики.

В работах К. Маркса и В. И. Ленина средним величинам отводится особая роль.

К. Маркс утверждал, что в средней величине погашаются индивидуальные отклонения от общего уровня и средний уровень становится обобщающей характеристикой массового явления. Такой характеристикой массового явления средняя величина становится лишь при условии, если взято значительное число единиц и эти единицы качественно однородны. Маркс писал, чтобы находимая средняя величина была средней «...многих различных индивидуальных величин одного и того же вида».

Средняя величина приобретает особую значимость в условиях рыночной экономики. Она помогает определить необходимое и общее, тенденцию закономерности экономического развития непосредственно через единичное и случайное.

Средние величины являются обобщающими показателями, в которых находят выражение действие общих условий, закономерность изучаемого явления.

Статистические средние величины рассчитываются на основе массовых данных статистически правильно организованного массового наблюдения.

Если статистическая средняя рассчитывается по массовым данным для качественно однородной совокупности (массовых явлений), то она будет объективной.

Средняя величина абстрактна, так как характеризует значение абстрактной единицы.

От разнообразия признака у отдельных объектов абстрагируется средняя. Абстракция – степень научного исследования. В средней величине осуществляется диалектическое единство отдельного и общего.

Средние величины должны применяться исходя из диалектического понимания категорий индивидуального и общего, единичного и массового.

Средняя отображает что-то общее, которое складывается в определенном единичном объекте.

5.2. Требования, предъявляемые к средним величинам

К расчету средней величины предъявляются следующие основные требования:

(1) Среднюю величину нужно рассчитывать так, чтобы она погасила то, что мешает выявлению характерных черт и закономерностей в развитии явления, а не затушевывала развитие.

(2) Средняя величина может быть вычислена только для однородной совокупности. Для получения однородной совокупности необходима группировка данных, поэтому расчет средней должен сочетаться с методом группировок.

(3) Средняя только тогда будет верной обобщающей характеристикой совокупности по варьирующему признаку, когда при замене всех вариантов средней общий объем варьирующего признака остается неизменным.

5.3. Принципы применения средних величин

Общие принципы применения средних величин:

1. Средняя должна определяться для совокупностей, состоящих из качественно однородных единиц.
2. Средняя должна исчисляться для совокупности, состоящей из достаточно большого числа единиц.
3. Средняя должна рассчитываться для совокупности, единицы которой находятся в нормальном, естественном состоянии.
4. Средняя должна вычисляться с учетом экономического содержания исследуемого показателя.

5.4. Требование качественной однородности совокупности при расчете средних величин

Статистическая совокупность—это совокупность социально-экономических объектов или явлений общественной жизни, объединенных некоей качественной основой, общей связью, но отличающихся друг от друга отдельными признаками. Таковы, например, совокупность домохозяйств, совокупность семей, совокупность предприятий, фирм, объединений и т. п.. Совокупности могут быть однородными и разнородными.

Совокупность называется *однородной*, если один или несколько изучаемых существенных признаков ее объектов являются общими для всех единиц. Совокупность оказывается однородной именно с точки зрения этих признаков.

Совокупность, в которую входят явления разного типа, считается *разнородной*. Совокупность может быть однородна в одном отношении и разнородна в другом. В каждом отдельном случае однородность совокупности устанавливается путем проведения качественного анализа, выяснения содержания изучаемого общественного явления.

Статистическая совокупность состоит из отдельных единиц (в статистике внешней торговли - отдельных партий товаров), имеющих свои свойства, особенности.

Единица совокупности — это первичный элемент статистической совокупности, являющийся носителем признака.

Требование качественной однородности совокупности при расчете средних отражает типичный уровень признака в исследуемой совокупности в расчете на единицу совокупности в конкретных условиях места и времени. Типичность средней непосредственно связано с однородностью изучаемой совокупности. В случае неоднородности совокупности ее необходимо разбить на качественно однородные группы и рассчитать среднюю по каждой из однородных групп. От того, в каком виде представлены данные, зависит каким именно будет исходное соотношение средней величины.

5.5. Виды средних величин. Степенные средние

В зависимости от содержания поставленной задачи, от характера и взаимосвязи признаков, единицы изучаемых явлений средние величины делятся на следующие группы:

- I. Степенные средние величины
- II. Структурные средние величины
- III. Хронологические средние величины

К *степенным средним* относятся такие наиболее известные и часто применяемые виды, как средняя геометрическая, средняя арифметическая и

средняя квадратическая, средняя гармоническая и средняя кубическая величины.

Все степенные средние объединены общей формулой и отличаются только значением m .

$$\bar{x} = \sqrt[m]{\frac{\sum_{i=1}^n x_i^m}{n}}$$

где \bar{x} - среднее значение исследуемого признака; m - показатель степени средней, x_i - текущее значение (варианта) осредняемого признака, n - число признаков, количество x .

В зависимости от значения показателя степени m различают следующие виды степенных средних:

- $m = -1$ – средняя гармоническая,
- $m = 0$ – среднее геометрическая,
- $m = 1$ – средняя арифметическая,
- $m = 2$ – средняя квадратическая,
- $m = 3$ - средняя кубическая,

При использовании одних и тех же исходных данных, чем больше показатель степени m в вышеуказанной формуле, тем больше значение средней величины:

Свойство степенных средних возрастать с повышением показателя степени определяющей функции в статистике называется *правилом мажорантности средних*.

Степенные средние в зависимости от представления исходных данных могут быть простыми и взвешенными.

Степенная средняя взвешенная величина определяется по сгруппированным данным по формуле:

$$\bar{x} = \sqrt[m]{\frac{\sum x^m f}{\sum f}},$$

- где f – частота повторения варианта,
- x_i – текущее значение осредняемого признака,
- m – показатель степени средней.

Характер имеющихся данных определяет существование только одного истинного среднего значения показателя.

Исходя из вышеизложенного, отразим все виды степенных средних в таблице 1.

Значение k	Наименование средней	Формула средней	
		простая	взвешенная
-1	Гармоническая	$\bar{x} = \frac{n}{\sum \frac{1}{x}}$	$\bar{x} = \frac{\sum f}{\sum \frac{1}{x} \cdot f}; \bar{x} = \frac{\sum w}{\sum \frac{1}{x} \cdot w}$
0	Геометрическая	$\bar{x} = \sqrt[n]{x_1 \cdot x_2 \cdot \dots \cdot x_n} = \sqrt[n]{\prod x}$	$\bar{x} = \sqrt[n]{x_1^{f_1} \cdot x_2^{f_2} \cdot \dots}$
1	Арифметическая	$\bar{x} = \frac{\sum x}{n}$	$\bar{x} = \frac{\sum x \cdot f}{\sum f}; \bar{x} = \frac{\sum x \cdot w}{\sum w}$
2	Квадратическая	$\bar{x} = \sqrt{\frac{\sum x^2}{n}}$	$\bar{x} = \sqrt{\frac{\sum x^2 \cdot f}{\sum f}}$

Вид средней выбирается в каждом отдельном случае путем конкретного анализа изучаемой совокупности, он определяется материальным содержанием изучаемого явления, а также принципами суммирования и взвешивания.

5.6. Средняя арифметическая величина

Наиболее распространенным видом средних является *средняя арифметическая*. Она применяется в тех случаях, когда объем варьирующего признака для всей совокупности является суммой значений признаков отдельных ее единиц. Для общественных явлений характерна аддитивность (суммарность) объемов варьирующего признака, этим определяется область применения средней арифметической и объясняется ее распространенность как обобщающего показателя. Так, например, общий фонд заработной платы – это сумма заработных плат всех работников, валовой сбор урожая – сумма произведенной продукции со всей посевной площади.

Чтобы исчислить среднюю арифметическую, нужно сумму всех значений признаков разделить на их число.

Средняя арифметическая применяется в форме простой средней и взвешенной средней. Исходной, определяющей, формой служит простая средняя.

Типичность средней арифметической величины непосредственным образом связана также с *однородностью статистической совокупности*. Она только тогда будет отражать типичный уровень признака, когда рассчитана по однородной совокупности. В противном случае метод средних используется в сочетании с методом группировок: общие средние заменяются или дополняются групповыми средними, рассчитанными по однородным группам.

Средняя арифметическая простая равна простой сумме отдельных значений осредняемого признака, деленной на общее число этих значений (она применяется в тех случаях, когда имеются *несгруппированные индивидуальные значения признака*).

Средняя арифметическая простая рассчитывается по *несгруппированным данным* на основании формулы

$$\bar{X} = \frac{\sum x}{n},$$

где x - значение (варианта) осредняемого признака, n – число значений (вариант).

Средняя из вариантов, которые повторяются различное число раз или, как говорят, имеют различный вес, называется *взвешенной*. В качестве весов выступают численности единиц в разных группах совокупности (в группу объединяют одинаковые варианты).

Средняя арифметическая взвешенная – средняя сгруппированных величин x_1, x_2, \dots, x_n , - вычисляется по формуле:

$$\bar{X} = \frac{\sum x f}{\sum f}$$

где f_1, f_2, \dots, f_n – веса (частоты повторения одинаковых признаков);

$\sum x f$ – сумма произведений величины признаков на их частоты;

$\sum f$ – общая численность единиц совокупности.

Вычисление средней арифметической часто сопряжено с большими затратами времени и труда. Однако в ряде случаев процедуру расчета средней можно упростить и облегчить, если воспользоваться ее свойствами.

5.7. Свойства средних величин

Приведем (без доказательства) некоторые основные свойства средней арифметической величины.

Свойство 1. Сумма отклонений индивидуальных значений признака от его среднего значения равна нулю:

$$\sum (x_i - \bar{x}) = 0 \qquad \sum (x_i - \bar{x}) f = 0$$

Свойство 2. Если все индивидуальные значения признака (т.е. все варианты) уменьшить или увеличить в B раз, то среднее значение нового признака соответственно уменьшится или увеличится в B раз. В следствие этого свойства индивидуальные значения признака можно сократить в B раз, произвести расчет средней и результат умножить на B .

$$\bar{X} = \frac{\sum x f}{\sum f} \qquad \bar{X}' = \frac{\sum x' f}{\sum f}$$

Свойство 3. Если к каждому индивидуальному значению признака прибавить число A или из каждого значения вычесть постоянное число A , то средняя величина возрастет или уменьшится на это же число. Это свойство полезно использовать при расчете средней величины из многозначных и слабаварьирующих значений признака.

$$\bar{x}' = \frac{\sum (x_i \pm A)}{N} = \bar{x} \pm A$$

$$(\sum (x \pm A)f / \sum f = \bar{x} \pm A$$

Свойство 4. Если веса всех осредняемых вариантов уменьшить или увеличить в A раз, то средняя арифметическая не изменится.

$$\bar{x}' = \frac{\sum x_i (f_i \cdot A)}{\sum f_i \cdot A} = \bar{x}$$

Свойство 5. Если во всех интервалах частоты равны друг другу, то средняя арифметическая взвешенная равна простой средней арифметической:

$$\bar{x}_{\text{ар}} = \frac{\sum \bar{x}_i f}{\sum f} = \frac{f \cdot \sum \bar{x}_i}{f \cdot k} = \frac{\sum \bar{x}_i}{k}$$

где k - количество групп вариационного ряда.

Свойство 6. Сумма произведений вариант на частоты равна произведению средней величины на сумму частот:

$$\sum x_i f_i = \bar{x} \sum f_i$$

5.8. Расчет средней арифметической в рядах распределения

Если значения осредняемого признака заданы в виде интервалов («от – до»), т.е. интервальных рядов распределения, то при расчете средней арифметической величины в качестве значений признаков в группах принимают середины этих интервалов, в результате чего образуется *дискретный ряд*.

От интервального ряда перейдем к дискретному путем замены интервальных значений их средними значениями (простая средняя между верхней и нижней границами каждого интервала). При этом величины открытых интервалов (первый и последний) условно приравниваются к интервалам, примыкающим к ним (второй и предпоследний).

При таком исчислении средней допускается некоторая неточность, поскольку делается предположение о равномерности распределения единиц признака внутри группы. Однако ошибка будет тем меньше, чем уже интервал и чем больше единиц в интервале.

После того как найдены середины интервалов, вычисления делают так же, как и в дискретном ряду, – варианты умножают на частоты (веса) и сумму произведений делят на сумму частот (весов):

5.9. Средняя гармоническая

Когда статистическая информация не содержит частот f по отдельным вариантам x совокупности, а представлена как их произведение $x \cdot f$,

применяется формула *средней гармонической взвешенной*. Это величина обратная средней арифметической.

Чтобы исчислить среднюю гармоническую взвешенную, обозначим $x \cdot f = w$, откуда $f = w/x$. Теперь преобразуем формулу средней арифметической таким образом, чтобы по имеющимся данным x и w можно было исчислить среднюю. В формулу средней арифметической взвешенной величины вместо xf подставим w , вместо f – отношение w/x и получим формулу *средней гармонической взвешенной*:

$$\bar{x}_{\text{гар}} = \frac{\sum w}{\sum \frac{w}{x}} = \frac{w_1 + w_2 + \dots + w_n}{\frac{w_1}{x_1} + \frac{w_2}{x_2} + \dots + \frac{w_n}{x_n}}$$

Из формулы видно, что средняя гармоническая – средняя взвешенная из варьирующих обратных значений признака. Она является преобразованной формой арифметической средней и тождественна ей. Вместо гармонической всегда можно рассчитать среднюю арифметическую, но для этого сначала нужно определить веса отдельных значений признака, скрытые в весах средней гармонической.

Таким образом, средняя гармоническая применяется тогда, когда неизвестны действительные веса f , а известно $w = x \cdot f$, то есть в тех случаях, когда средняя предназначается для расчета сумм слагаемых, обратно пропорциональных величине данного признака, когда суммированию подлежат не сами варианты, а обратные им величины.

Исчисление средней гармонической взвешенной величины освобождает от необходимости предварительного расчета весов, поскольку эта операция заложена в саму формулу.

В тех случаях, когда вес каждого варианта равен единице (индивидуальные значения обратного признака встречаются по одному разу), применяется средняя гармоническая простая, которая исчисляется по формуле:

$$\bar{x}_{\text{гар}} = \frac{1+1+\dots+1}{\frac{1}{x_1} + \frac{1}{x_2} + \dots + \frac{1}{x_n}} = \frac{n}{\sum \frac{1}{x}}$$

где x – отдельные варианты обратного признака, встречающиеся по одному разу;

n – число вариантов.

Если по двум частям совокупности (численности n_1 и n_2) даны средние гармонические, то общую среднюю гармоническую по всей совокупности можно представить как взвешенную гармоническую среднюю из групповых средних:

$$\bar{x}_{\text{гар}} = \frac{n_1 + n_2}{\frac{n_1}{\bar{x}_{1\text{гар}}} + \frac{n_2}{\bar{x}_{2\text{гар}}}}$$

Средняя гармоническая взвешенная:

$$\bar{x} = \frac{\sum w}{\sum \frac{1}{x} w}, \text{ где } w - \text{ объем совокупности.}$$

5.10. Средняя геометрическая

Средняя геометрическая применяется в тех случаях, когда индивидуальные значения признака представляют собой, как правило, относительные величины динамики, построенные в виде цепных величин, как отношение к предыдущему уровню каждого уровня в ряду динамики, то есть характеризует средний коэффициент роста.

Средняя геометрическая равна корню степени n из произведения коэффициентов роста, характеризующих отношение величины каждого последующего периода к величине предыдущего.

Средняя геометрическая применяется в тех случаях, когда необходимо определить средний коэффициент (темп) роста при анализе динамики:

$\bar{K} = \sqrt[n]{K_1 * K_2 * K_3 * \dots * K_n} = \sqrt[n]{\text{ПК}}$, где K – коэффициент роста, n – количество K .

Для определения средней геометрической взвешенной применяется формула:

$$\bar{x} = \frac{\sum f_i \sqrt[f_i]{(x_1)^{f_1} (x_2)^{f_2} \dots (x_n)^{f_n}}}{\sum f_i} = \frac{\sum f_i \sqrt[f_i]{\Pi(x_i)^{f_i}}}{\sum f_i}$$

5.11. Средняя квадратическая величина

В ряде случаев в экономической практике возникает потребность расчета среднего размера признака, выраженного в квадратных или кубических единицах измерения.

Если осреднению подлежат величины, выраженные в виде квадратных функций, применяется *средняя квадратическая величина*. Например, с помощью средней квадратической можно определить средние диаметры труб, колес, средние стороны n квадратных участков и т. д.

Средняя квадратическая простая определяется путем извлечения квадратного корня из частного от деления суммы квадратов отдельных значений признака на их число.

$$\bar{x} = \sqrt{\frac{x_1^2 + x_2^2 + x_3^2 + \dots + x_n^2}{n}} = \sqrt{\frac{\sum x^2}{n}}$$

Средняя квадратическая взвешенная равна:

$$\bar{x} = \sqrt{\frac{\sum x^2 f}{\sum f}}$$

5.12. Средняя кубическая величина

При определении средней длины стороны n кубов определяется средняя кубическая.

Формулы для расчета *средней кубической* аналогичны средней квадратической:

- Средняя кубическая простая

$$\bar{x}_{\text{куб}} = \sqrt[3]{\frac{\sum x^3}{n}};$$

- Средняя кубическая взвешенная

$$\bar{x}_{\text{куб}} = \sqrt[3]{\frac{\sum x^3 f}{f}},$$

Средние квадратическая и кубическая величины имеют ограниченное применение в практике статистики. Широко пользуется статистика средней квадратической величины, но не из самих вариантов x , и из их отклонений от средней ($x - \bar{x}$) при расчете показателей вариации.

Средняя может быть вычислена не для всех, а для какой-либо части единиц совокупности. Примером такой средней может быть средняя прогрессивная величина как одна из частных средних, вычисляемая не для всех, а только для «лучших» (например, для показателей выше или ниже средних индивидуальных).

5.13. Средняя хронологическая величина

Средняя хронологическая — это средний уровень ряда динамики, т. е. средняя, исчисленная по совокупности значений показателя в разные моменты или периоды времени.

В зависимости от вида ряда динамики применяются различные способы ее расчета, а именно расчет средней хронологической интервального ряда и средней хронологической моментного ряда.

Ряды динамики состоят из числовых значений двух показателей: моментов или периодов времени t , к которым относятся приводимые данные, и соответствующих им статистических данных y , которые называются уровнями динамического ряда. В зависимости от того, к моментам или периодам времени привязываются статистические данные, различают два вида рядов динамики: моментные и интервальные.

Когда уровни ряда динамики характеризуют размеры общественных явлений за определенные интервалы (периоды) времени (за сутки, месяц, квартал, год и т. п.), то такие ряды называются интервальными (или

периодическими). В отличие от моментного ряда динамики уровни интервального ряда динамики могут быть суммированы. Например, сложив данные выпуска станков за четыре квартала, можно получить показатель их выпуска за год.

Для интервальных рядов с равноотстоящими уровнями средняя хронологическая имеет вид средней арифметической простой:

$$\bar{Y} = \sum y/n$$

где y – уровни интервального ряда;

n - количество равных периодов времени.

В интервальных рядах с неравноотстоящими уровнями средняя хронологическая имеет вид средней арифметической взвешенной:

$$\bar{Y} = \sum yt/\sum t,$$

где t – периоды времени, отделяющие один уровень ряда от другого

Если уровни ряда динамики выражают состояние явления на определенные моменты времени или даты, то такие ряды называют моментными рядами динамики. Особенность моментного ряда динамики в том, что некоторые его уровни содержат элементы повторного счета, т. е. каждый последующий уровень полностью или частично включает в себя предыдущий уровень. Поэтому суммирование уровней моментного динамического ряда не имеет смысла, но разность уровней имеет определенное значение.

В моментном ряду динамики с равноотстоящими уровнями средняя хронологическая имеет вид:

При равных промежутках времени между датами, на которые имеются данные, и равномерном изменении размера показателя между датами (то есть, в моментном ряду динамики с равноотстоящими уровнями) средняя хронологическая моментного ряда обычно исчисляется по формуле:

$$\bar{y} = \frac{\frac{1}{2}y_1 + y_2 + \dots + y_{n-1} + \frac{1}{2}y_n}{n-1},$$

где y — уровень ряда; n — число всех членов ряда; \bar{y} — средний уровень.

Характерной особенностью данного моментного ряда динамики с равными интервалами является то, что показатели о стоимости имущества предприятия представлены строго на определенную дату, притом — интервал составляет 1 месяц:

Средний уровень моментного ряда динамики с неравноотстоящими уровнями характеризует средняя хронологическая взвешенная, которая исчисляется по формуле:

$$\bar{Y} = (\sum (y_i + y_{i+1})/2) * t_i / \sum t_i$$

где y_i и y_{i+1} — значение уровня моментного ряда динамики и уровня, следующего за ним;

t_i — промежуток времени между датами.

5.14. Структурные средние

Особый вид средних величин – структурные средние – применяется для изучения внутреннего строения рядов распределения значений признака, а также для оценки средней величины (степенного типа), если по имеющимся статистическим данным ее расчет не может быть выполнен (например, если бы в рассмотренном примере отсутствовали данные и об объеме производства, и о сумме затрат по группам предприятий).

В качестве структурных средних чаще всего используют показатели *моды* – наиболее часто повторяющегося значения признака – и *медианы* – величины признака, которая делит упорядоченную последовательность его значений на две равные по численности части. В итоге у одной половины единиц совокупности значение признака не превышает медианного уровня, а у другой – не меньше его.

Мода (M_o) – чаще всего встречающийся вариант. Модой называется значение признака, которое соответствует максимальной точке теоретической кривой распределений.

Мода представляет наиболее часто встречающееся или типичное значение.

Мода применяется в коммерческой практике для изучения покупательского спроса и регистрации цен.

В дискретном ряду мода – это варианта с наибольшей частотой. В интервальном вариационном ряду модой считают центральный вариант интервала, который имеет наибольшую частоту (частность).

Для интервального ряда распределения мода определяется по формуле:

$$M_o = x_o + h \cdot \frac{f_m - f_{m-1}}{(f_m - f_{m-1}) + (f_m - f_{m+1})}$$

где x_o – нижняя граница модального интервала;

h – величина модального интервала;

f_m – частота модального интервала;

f_{m-1} – частота интервала, предшествующего модальному;

f_{m+1} – частота интервала, следующего за модальным.

Мода зависит от величины групп, от точного положения границ групп.

Мода – число, которое в действительности встречается чаще всего (является величиной определенной), в практике имеет самое широкое применение (наиболее часто встречающийся тип покупателя).

Медиана (M_e) – это величина, расположенная в середине упорядоченного ряда, которая делит статистическую совокупность на две равные части: одна часть имеет значения варьирующего признака меньшие, чем средний вариант, а другая – большие. Для определения медианы необходимо построить ранжированный ряд, т.е. ряд в порядке возрастания или убывания индивидуальных значений признака.

Медиана – это элемент, который больше или равен и одновременно меньше или равен половине остальных элементов ряда распределения.

Свойство медианы заключается в том, что сумма абсолютных отклонений значений признака от медианы меньше, чем от любой другой величины.

Применение медианы позволяет получить более точные результаты, чем при использовании других форм средних.

В ранжированном ряду медиана находится следующим образом. Во-первых, рассматривают номер медианной варианты: $n_{ме}=(n+1)/2$, где $n_{ме}$ – номер медианной варианты, n – общее число вариант в ряду.

Если общее число вариант нечетное, то медиана рассчитывается по вышеуказанной формуле.

В дискретном упорядоченном ряду с нечётным числом членов медианой будет варианта, расположенная в центре ряда.

Если дискретный упорядоченный ряд состоит из чётного числа членов, то медианой будет средняя арифметическая из двух смежных вариант, стоящих в центре ряда.

Если же сумма накопленных частот против одной из вариант равна точно половине суммы частот ряда, то медиана определяется как средняя арифметическая этой варианты и последующей.

Порядок нахождения медианы в интервальном вариационном ряду следующий: располагаем индивидуальные значения признака по ранжиру; определяем для данного ранжированного ряда накопленные частоты; по данным о накопленных частотах находим медианный интервал по формуле:

$$Me = X_{Me} + h_{Me} \frac{\sum f - S_{Me-1}}{f_{Me}}$$

где X_{Me} – нижняя граница медианного интервала;

h_{Me} – величина медианного интервала;

$\sum f$ – сумма частот ряда;

f_{Me} – частота медианного интервала;

S_{m-1} – сумма накопленных частот в интервалах предшествующих медианному.

5.15. Показатели дифференциации и концентрации

При изучении вариационного ряда для отражения его структуры, помимо моды и медианы, также применяются другие характеристики.

Для выражения особенностей формы распределения применяются ранговые характеристики.

Ранговые характеристики – это варианты, занимающие в вариационном ряду определенное место.

Их общее название – *квантили* – это порядковые характеристики признака, которые делят все единицы ряда на равные численности. Квантили

предназначены для более глубокого изучения структуры ряда распределения. Квантили, по сути, есть варианты, занимающие определенное место (десятое, двадцать пятое, пятидесятое и т.д.) в ранжированной совокупности, в которой в качестве весов используются частоты. К числу квантилей, наиболее часто используемых в статистическом анализе, относят:

медиану (Me),

квартили (Q) - значения признака, делящие упорядоченную совокупность на четыре равные части;

квинтили (quintile) – значения признака, делящие упорядоченную совокупность на пять равных частей;

децили (D) - значения признака, делящие упорядоченную совокупность на десять равных частей. Первый дециль делит совокупность в соотношении 1/10 к 9/10 второй дециль делит в соотношении 2/10 к 8/10 и т.д.;

и перцентили (P) - значения признака, делящие упорядоченную совокупность на сто равных частей.

Использование в анализе вариационных рядов распределения рассмотренных выше характеристик позволяет более глубоко и детально охарактеризовать изучаемую совокупность.

Контрольные вопросы

1. Дайте определение средней величины.
2. Какова роль средних величин в регулировании действия случайных причин и определении среднего уровня явления.
3. История развития средних величин.
4. Какие виды средних величин применяются в статистике?
5. Требования, предъявляемые к средним величинам.
6. Принципы применения средних величин.
7. Требование качественной однородности совокупности при расчете средних величин.
8. Правило мажорантности средних величин.
9. Как исчисляется средняя арифметическая простая и в каких случаях она применяется?
10. Как исчисляется средняя арифметическая взвешенная и в каких случаях она применяется?
11. Как исчисляется средняя арифметическая из вариационного ряда?
12. Каковы основные свойства средней арифметической?
13. Как исчисляется средняя гармоническая простая?
14. Как исчисляется средняя гармоническая взвешенная?
15. Как исчисляется средняя геометрическая и где она применяется?
16. Как исчисляется средняя квадратическая величина?
17. Как исчисляется средняя кубическая величина?

18. Понятие средней хронологической величины.
19. Понятие структурных средних.
20. Понятие моды.
21. Понятие медианы.
22. Ранговые характеристики.
23. Понятие квантиля.
24. Понятие квартиля.
25. Понятие квинтиля.
26. Понятие дециля.
27. Понятие перцентиля.

Тема 6. Анализ распределения (вариация)

6.1. Понятие вариации

Конкретные условия, в которых находится каждый из изучаемых объектов, а также особенности их собственного развития (социальные, экономические и пр.) влияют на формирование изучаемого признака у каждой единицы совокупности и, соответственно, выражаются определенными числовыми уровнями статистических показателей.

Средняя величина – это величина абстрактная, которая обобщает характеристику признака изучаемой совокупности, однако не дает представления о том, как отдельные значения изучаемого признака группируются вокруг средней: сосредоточены ли они вблизи или значительно отклоняются от неё. В некоторых случаях отдельные значения признака близко примыкают к средней арифметической и мало от нее отличаются – значит средняя хорошо представляет всю совокупность. В других случаях отдельные значения совокупности далеко отстают от средней и средняя плохо представляет всю совокупность.

Поэтому в социально-экономическом анализе важно знать не только среднее (или срединное) значение признака, но и насколько равномерно распределены эти значения относительно среднего значения.

Совокупности могут иметь одинаковые значения средней величины, но отличаться колеблемостью индивидуальных значений. Вот почему ограничиваться вычислением одной средней в ряде случаев недостаточно.

Это вызывает необходимость измерять вариацию признака в совокупности. Для этой цели в статистике применяют ряд обобщенных показателей.

Вариация признака – это различие индивидуальных значений признака внутри изучаемой совокупности в один и тот же период или момент времени.

Несовпадение уровней одного и того же показателя у разных объектов имеет объективный характер, индивидуальные значения складываются под совокупным влиянием разнообразных факторов (условий), которые по-разному сочетаются в каждом отдельном случае.

Термин «*вариация*» произошел от латинского *variatio* – «изменение, колеблемость, различие». Однако не всякие различия принято называть вариацией. Под вариацией в статистике понимают такие количественные изменения величины исследуемого признака в пределах однородной совокупности, которые обусловлены перекрещивающимся влиянием действия различных факторов. Различают вариацию признака: случайную и систематическую.

Анализ систематической вариации позволяет оценить степень зависимости изменений в изучаемом признаке от определяющих ее факторов. Например, изучая силу и характер вариации в выделяемой совокупности, можно оценить, насколько однородной является данная совокупность в

количественном, а иногда и качественном отношении, а следовательно, насколько характерной является исчисленная средняя величина. Степень близости данных отдельных единиц x_i к средней измеряется рядом абсолютных, средних и относительных показателей.

Возникает вариация в силу того, что отдельные значения признака статистической совокупности формируются под воздействием разнообразных факторов. Значение изучения вариации состоит в том, что по колеблемости признаков можно судить о качественной однородности совокупности.

6.2. Задачи изучения вариации

Наличие вариации в признаках изучаемых явлений ставит перед статистикой *задачи* ее исследования:

- 1) определение меры вариации и ее измерение,
- 2) нахождение соответствующих измерителей и показателей, характеризующих размеры вариации,
- 3) выявление сущности показателей вариации и методов вычисления факторов, ее определяющих.

По степени вариации можно судить о многих сторонах процесса развития изучаемых явлений, в частности об однородности совокупности, устойчивости индивидуальных значений признака, типичности средней, о взаимосвязи между признаками одного и того же явления и признаками разных явлений.

6.3. Основные показатели вариации

По степени вариации признака можно судить о процессах развития изучаемых явлений, о типичности средних величин.

Для определения меры вариации признака в статистике используются абсолютные и относительные *показатели вариации*.

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

К относительным показателям вариации относятся: коэффициент осцилляции, относительное линейное отклонение, коэффициент вариации.

6.4. Абсолютные показатели вариации

Размах вариации (R) является самым простым, элементарным из абсолютных показателей вариации и представляет собой разность между максимальным и минимальным значениями признака:

$$R = X_{\max} - X_{\min} ,$$

где X_{\max} и X_{\min} – максимальное и минимальное значение признака в статистической совокупности.

Величина размаха вариации зависит только от крайних значений и не учитывает всех изменений признака в пределах изучаемой совокупности. Поэтому при изучении вариации нельзя ограничиваться расчетом только этого показателя. Для анализа вариации необходимы показатели, дающие обобщенную характеристику всех колебаний варьирующего признака.

Имея ввиду, что алгебраическая сумма индивидуальных значений признаков при расчете отклонений от средней арифметической равна 0, то можно определить:

Среднее линейное отклонение – простейший обобщающий показатель. Он представляет собой среднюю величину абсолютных отклонений индивидуальных значений признака от их средней арифметической величины.

Среднее линейное отклонение для несгруппированных данных определяется по формуле

$$\bar{d} = \frac{\sum d_i}{n} = \frac{\sum |X_i - \bar{X}|}{n} ,$$

где $d_i = |X_i - \bar{X}|$ - индивидуальное линейное отклонение.

Среднее линейное отклонение для сгруппированных данных рассчитывается так:

$$\bar{d} = \frac{\sum d_i \times f_i}{\sum f_i} = \frac{\sum |X_i - \bar{X}| \times f_i}{\sum f_i}$$

Показатель среднего линейного отклонения нашел широкое применение на практике. С его помощью анализируют состав работающих, ритмичность производства, равномерность поставок материалов; разрабатывают системы материального стимулирования. Но этот показатель усложняет расчёты вероятностного типа, затрудняет применение методов математической статистики.

Среднее линейное отклонение не всегда улавливает степень вариации значений признака. Поэтому в статистике применяется более чувствительный обобщающий показатель – *дисперсия*, которая представляет собой средний квадрат отклонений индивидуальных значений признака от их средней арифметической величины.

Дисперсия для несгруппированных данных вычисляется по формуле

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

Дисперсия для сгруппированных данных рассчитывается так:

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

Среднее квадратическое отклонение является обобщающей характеристикой размеров вариации признака совокупности. Это - мера вариации, показатель надёжности средней. Чем меньше значение среднего квадратического отклонения, тем лучше средняя величина представляет собой рассматриваемую совокупность.

Среднее квадратическое отклонение представляет собой корень квадратный из дисперсии.

Среднее квадратическое отклонение рассчитывается по формуле:

$$\sigma = \sqrt{\sigma^2}$$

То есть, для несгруппированных данных:

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

Или, для сгруппированных данных:

$$\sigma = \sqrt{\frac{\sum (x - \bar{x})^2 f}{\sum f}}.$$

Среднее квадратическое отклонение также, как и среднее линейное отклонение, показывает, на сколько в среднем отличаются индивидуальные значения признака от их среднего значения. Однако по величине среднеквадратическое отклонение во всех случаях превышает среднее линейное, так как более чутко реагирует на вариацию.

6.5. Относительные показатели вариации

Размах вариации, среднее линейное и квадратическое отклонения выражаются именованными числами, т. е. имеют единицу измерения (такую же, как и значения признака). Поэтому их нельзя непосредственно использовать для сравнения степени вариации по одному и тому же признаку в двух группах с разным уровнем средних, а также для сравнения вариации двух различных признаков в одной группе. В этих случаях применяются *относительные* показатели вариации. Базой для сравнения должна служить средняя арифметическая. Эти показатели вычисляются как отношение размаха вариации, среднего линейного отклонения или среднего квадратического отклонения к средней арифметической (реже к медиане). Чаще всего они выражаются в процентах и определяют не только сравнительную оценку вариации, но и дают характеристику однородности совокупности.

Коэффициент осцилляции.

Отражает относительную меру колеблемости крайних значений признака вокруг средней и рассчитывается по формуле:

$$V_R = \frac{R}{\bar{x}} \cdot 100\% .$$

Относительное линейное отклонение (линейный коэффициент вариации).

Отражает долю усреднённого значения абсолютных отклонений от средней величины и рассчитывается по формуле:

$$V_{\bar{d}} = \frac{\bar{d}}{\bar{x}} \cdot 100\% .$$

Коэффициент вариации.

Это относительное квадратическое отклонение от средней величины, рассчитывается по формуле:

$$V_{\sigma} = \frac{\sigma}{\bar{x}} \cdot 100\% .$$

Коэффициент вариации позволяет не только получить обобщающую характеристику вариации признака в совокупности, но и дает возможность сделать выводы об однородности совокупности. Совокупность считается однородной, если коэффициент вариации не превышает 33%. Средние величины, рассчитанные по однородной совокупности.

Величина рассчитанных нами коэффициентов свидетельствует о том, что колеблемость индивидуальных значений выработки во второй бригаде высокая. Первую совокупность можно считать однородной, а её среднюю – надёжной. Вторую совокупность следует считать неоднородной, а её среднюю – ненадёжной.

6.6. Свойства дисперсий

Свойство 1. Дисперсия постоянной величины равна нулю.

Свойство 2. Уменьшение всех значений признака на одну и ту же величину А не меняет величины дисперсии: $\sigma_{(x-A)}^2 = \sigma_x^2$

$$D = \sigma^2$$

Значит, средний квадрат отклонений можно вычислить не по заданным значениям признака, а по отклонениям их от какого-то постоянного числа.

Свойство 3. Уменьшение всех значений признака в k раз уменьшает дисперсию в k² раз, а среднее квадратическое отклонение - в k раз:

$$\sigma_{\frac{x}{k}}^2 = \frac{\sigma_x^2}{k^2}$$

Значит, все значения признака можно разделить на какое-то постоянное число (скажем, на величину интервала ряда), исчислить среднее квадратическое отклонение, а затем умножить его на постоянное число:

$$\sigma_x = \sigma_{\frac{x}{k}} \times k$$

Свойство 4. Если исчислить средний квадрат отклонений от любой величины A , в той или иной степени отличающейся от средней арифметической (\bar{x}), то он всегда будет больше среднего квадрата отклонений, исчисленного от средней арифметической: $\sigma_A^2 > \sigma_x^2$

Средний квадрат отклонений при этом будет больше на вполне определенную величину - на квадрат разности средней и этой условно взятой величины, т. е. на $(\bar{x} - A)^2$:

$$\sigma_A^2 = \sigma_x^2 + (\bar{x} - A)^2 \quad \text{или} \quad \sigma_x^2 = \frac{\sum (x_i - A)^2 \times f_i}{\sum f_i} - (\bar{x} - A)^2$$

Значит, дисперсия от средней всегда меньше дисперсий, исчисленных от любых других величин, т. е. она имеет свойство минимальности.

В случае когда A приравнивается к нулю и, следовательно, не вычисляются отклонения, формула принимает такой вид:

$$\sigma_x^2 = \overline{x^2} - (\bar{x})^2 \quad \text{или} \quad \sigma_x^2 = \frac{\sum x_i^2 \times f_i}{\sum f_i} - \left(\frac{\sum x_i \times f_i}{\sum f_i} \right)^2$$

Значит, средний квадрат отклонений равен среднему квадрату значений признака минус квадрат среднего значения признака.

Между среднелинейным отклонением и среднеарифметической существует следующее примерное соотношение $\sigma = 1,25\bar{d}$, при условии, если фактическое распределение близко к нормальному.

По свойству мажорантности средних величин среднеквадратическое отклонение всегда больше среднелинейного отклонения.

Абсолютными показателями вариации нельзя пользоваться в двух случаях, так как они именованные:

- для сравнения степени вариации 2 различных признаков в одной и той же совокупности.
- для сравнения вариации по одному и тому же признаку, но в 2 различных группах с разным уровнем средних.

6.7. Виды дисперсий и правила их сложения

В статистическом исследовании очень часто бывает необходимо не только изучить вариации признака по всей совокупности, но и проследить количественные изменения признака по однородным группам совокупности, а также и между группами. Следовательно, помимо общей средней для всей совокупности необходимо просчитывать и частные средние величины по отдельным группам.

Различают три вида дисперсий:

- общая;
- средняя внутригрупповая;
- межгрупповая.

Общая дисперсия ($D = \sigma^2$) характеризует вариацию признака всей совокупности под влиянием всех тех факторов, которые обусловили данную вариацию. Эта величина определяется по формуле

$$D_x = \frac{\sum (x_i - \bar{x})^2 f_i}{\sum f_i}$$

где \bar{x} - общая средняя арифметическая всей исследуемой совокупности.

Средняя *внутригрупповая дисперсия* ($\bar{\sigma}_i^2$) свидетельствует о случайной вариации, которая может возникнуть под влиянием каких-либо неучтенных факторов и которая не зависит от признака-фактора, положенного в основу группировки. Данная дисперсия рассчитывается следующим образом: сначала рассчитываются дисперсии по отдельным группам (σ_i^2):

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

а затем - рассчитывается средняя внутригрупповая дисперсия :

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

где n_i - число единиц в группе

Межгрупповая дисперсия (δ_i^2) (дисперсия групповых средних) характеризует систематическую вариацию, т.е. различия в величине исследуемого признака, возникающие под влиянием признака-фактора, который положен в основу группировки. Эта дисперсия рассчитывается по формуле:

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

где \bar{X}_i - средняя величина по отдельной группе.

n_i - численность по группе

Все три вида дисперсии связаны между собой: общая дисперсия равна сумме средней внутригрупповой дисперсии и межгрупповой дисперсии:

$$\sigma_0^2 = \bar{\sigma}_i^2 + \delta^2 \quad \text{или} \quad D_x = \bar{D}_i + \delta_x^2$$

Данное соотношение отражает закон, который называют *правилом сложения дисперсий*. Согласно этому закону (правилу), общая дисперсия, которая возникает под влиянием всех факторов, равна сумме дисперсий, которые появляются как под влиянием признака-фактора, положенного в основу группировки, так и под влиянием других факторов. Благодаря правилу сложения дисперсий можно определить, какая часть общей дисперсии находится под влиянием признака-фактора, положенного в основу группировки.

Правило сложения дисперсий широко применяется при исчислении показателей тесноты связи, в дисперсионном анализе, при оценке точности типической выборки и в ряде других случаев.

Контрольные вопросы

1. Что представляет себя вариация признака?
2. Задачи изучения вариации.
3. Абсолютные и относительные показатели вариации.
4. Что такое размах вариации, и по какой формуле он исчисляется?
5. Что представляет собой среднее линейное отклонение, его формулы для сгруппированных и не сгруппированных данных?
6. Какой показатель вариации называется дисперсией, и по каким формулам она рассчитывается?
7. Что называется средним квадратическим отклонением? По каким формулам оно рассчитывается?
8. Что представляет собой коэффициент осцилляции?
9. Что представляет собой относительное линейное отклонение (линейный коэффициент вариации)?
10. Что представляет собой коэффициент вариации?
11. Каковы основные свойства дисперсии?
12. Виды дисперсии.
13. Что характеризует межгрупповая дисперсия, ее формула?
14. Как определяется средняя внутригрупповая дисперсия?
15. Что собой представляет правило сложения дисперсий, в чем его практическое значение?

Тема 7. Выборочный метод в статистических исследованиях

7.1. Понятие о выборочном наблюдении и его значение

В зависимости от полноты охвата единиц совокупности статистика различает *сплошное* и *несплошное* наблюдение.

Сплошным называется такое наблюдение, при котором обследованию подлежат все без исключения единицы изучаемой статистической совокупности. Например, перепись населения, оценка успеваемости студентов в учебные сессии. Такое наблюдение связано с большими трудовыми и материальными затратами. Изучение не всех единиц совокупности, а лишь некоторой части, по которой следует судить о свойствах всей совокупности в целом, можно осуществить не сплошным методом.

Несплошным называется такое наблюдение, при котором обследованию подвергаются не все единицы изучаемой совокупности, а только их часть, на основе которой можно получить обобщающую характеристику всей совокупности. Несплошное наблюдение подразделяется на обследование основного массива, монографическое наблюдение, анкетное наблюдение, выборочное наблюдение.

Обследование основного массива предполагает отбор и изучение наиболее крупных единиц совокупности, преобладающих в общей массе по изучаемому признаку (*например, наблюдение за работой городских продовольственных рынков можно организовать, обследовав наиболее крупные из них – Комаровский, Ждановичский*).

Монографическое наблюдение предполагает тщательное обследование отдельных единиц изучаемой совокупности, обычно представителей каких-либо новых типов явлений (*например, детальное изучение работы организации специалистами Госконтроля*).

Анкетное наблюдение – широкому кругу лиц рассылаются анкеты, содержащие вопросы программы наблюдения. Недостатки: возвращаются не все анкеты, их заполняют наиболее активные и заинтересованные лица, в результате чего может быть утрачен элемент случайности, правильность ответов проконтролировать невозможно.

Из всех видов несплошного наблюдения на практике наибольшее распространение получило *выборочное наблюдение*, под которым понимается такое наблюдение, при котором обобщающие показатели изучаемой совокупности устанавливаются по некоторой ее части, включающей единицы, отобранные случайным образом, а результаты распространяются на всю исходную совокупность. Наблюдение организуется таким образом, что эта часть отобранных единиц в уменьшенном масштабе репрезентативна (представляет) всю совокупность.

Изучаемая статистическая совокупность, из которой производится отбор части единиц, называется *генеральной совокупностью*, а все ее обобщающие показатели - *генеральными*.

Отобранная из генеральной совокупности часть единиц, подвергающихся обследованию, называется *выборочной совокупностью* или *выборкой*, и все ее обобщающие показатели - *выборочными*.

Выборочный метод применяется в тех случаях, когда проведение сплошного наблюдения *невозможно* или *экономически нецелесообразно*. В частности, проверка качества отдельных видов продукции может быть связана с ее уничтожением (*например, оценка крепости ниток на разрыв; дегустация продуктов питания; контроль качества фарфора, электроламп, свечек, различных сплавов, проверка консервов на доброкачественность и т. п.*). Отдельные статистические совокупности настолько велики, что было бы физически невозможно собрать данные в отношении каждой из единиц (*например, изучение цен на рынках, изучение бюджетов семей*). Выборочное наблюдение используют также для *проверки результатов сплошного наблюдения*.

Выборочный метод позволяет при минимальной численности обследованных единиц наблюдение произвести в наиболее короткие сроки с минимальными затратами труда и средств. Повышается оперативность информации, уменьшаются ошибки регистрации вследствие лучшего кадрового обеспечения наблюдения, так как к его проведению привлекаются наиболее квалифицированные кадры.

Преимущество выборочного наблюдения по сравнению со сплошным можно реализовать, если оно организовано и проведено в соответствии с научными принципами *теории выборочного метода*. Это обеспечение *случайности* (равной возможности попадания в выборку) отбора единиц и *достаточности* их числа. Соблюдение этих принципов позволяет получить объективную гарантию репрезентативности полученной выборочной совокупности. Понятие репрезентативности отобранной совокупности это ее представительство только в отношении тех признаков, которые изучаются и оказывают существенное влияние на формирование сводных обобщающих характеристик.

Основная задача выборочного наблюдения в экономике состоит в том, чтобы на основе характеристик выборочной совокупности (средней и доли) получить достоверные суждения о показателях средней и доли в генеральной совокупности.

Поскольку изучаемая статистическая совокупность состоит из единиц с варьирующими признаками, то состав выборочной совокупности может в той или иной мере отличаться от состава генеральной совокупности. Возникающее вследствие этого расхождение между характеристиками выборки и генеральной совокупности составляет *ошибку выборки*. Способы определения ошибки выборки и распространения характеристик выборки на

генеральную совокупность и составляют основное содержание статистической методологии выборочного метода.

7.2. Основные способы формирования выборочной совокупности

При выборочном наблюдении, как правило, обследованию подвергается сравнительно небольшая часть совокупности – 5–10%, реже 15–25%. Качество результатов выборочного наблюдения зависит от того, насколько *репрезентативна* (т. е. представительна) выборка. Для обеспечения репрезентативности выборки необходимо соблюдение принципа случайности отбора единиц. Принцип случайности предполагает, что на включение объекта в выборку не может повлиять какой-либо иной фактор, кроме случая.

Существуют различные *способы формирования* выборочной совокупности. В практике выборочных наблюдений наибольшее распространение получили собственно-случайная, механическая, типическая, серийная, комбинированная выборки. При этом основной вид выборки – *собственно-случайная*, все другие являются ее развитием или видоизменением.

Собственно-случайная выборка заключается в отборе единиц из генеральной совокупности наугад без каких-либо элементов системности. Отбор проводят методом *жеребьевки*. Всем элементам генеральной совокупности присваивается порядковый номер и для каждого элемента готовится жребий (пронумерованный шар или фишка). Жребии перемешиваются в специальном ящике, из которого затем отбираются наугад.

Собственно-случайный отбор может быть как *повторным*, так и *бесповторным*.

При *повторной выборке* общая численность единиц генеральной совокупности в процессе выборки остается неизменной. Ту или иную единицу, попавшую в выборку, после регистрации снова возвращают в генеральную совокупность, и она сохраняет равную возможность со всеми прочими единицами при повторном отборе единиц попасть в выборку («отбор по схеме возвращенного шара»). Повторная выборка в социально-экономической жизни встречается редко. Обычно выборку организуют по схеме бесповторной выборки.

Для проведения *бесповторного отбора* в процессе жеребьевки выпавшие жребии обратно в совокупность не возвращаются и в дальнейшем отборе не участвуют. Таким образом, при данной выборке численность единиц генеральной совокупности сокращается в процессе исследования.

Механическая выборку можно рассматривать как разновидность собственно-случайной бесповторной выборки, применяется в случаях, когда генеральная совокупность каким-либо образом упорядочена, т. е. имеется определенная последовательность в расположении единиц (*например,*

табельные номера работников, избирателей, телефонные номера респондентов, номера домов и квартир и т. п.). Поэтому для определения средней ошибки механической выборки используют формулы собственно-случайной бесповторной выборки. Для проведения механической выборки:

- устанавливается пропорция отбора, которая определяется соотношением объемов выборочной и генеральной совокупностей. Так, если из совокупности в 5000 единиц предполагается получить 2%-ную выборку, т. е. отобрать 100 единиц, то пропорция отбора составит 1:50 ($1/5000:100$). Это значит, что отбирается одна из каждых 50 единиц;

- в соответствии с установленной пропорцией отбора генеральная совокупность механически разбивается на равновеликие группы (в нашем примере по 50 единиц);

- из каждой группы в выборку отбирается одна единица. При этом устанавливается начало отсчета, т. е. номер единицы, которая должна быть обследована первой, и шаг отсчета, т. е. расстояние между отбираемыми единицами. За начало отсчета чаще всего принимают единицу, лежащую в середине первого интервала для исключения возможности возникновения систематической ошибки выборки. Шаг отсчета равен ширине интервала, на который разбивается совокупность.

Типическая выборка используется в тех случаях, когда все единицы генеральной совокупности можно разбить на несколько типических групп (т. е. осуществить типологическую группировку). Например, при обследованиях населения такими группами могут быть социальные, возрастные или образовательные группы; при обследовании организаций – виды экономической деятельности. Типический отбор предполагает выборку единиц из каждой типической группы собственно-случайным или механическим способом пропорционально объему групп либо внутригрупповой вариации признака.

Типическая выборка обычно применяется при изучении сложных статистических совокупностей. Например, при выборочном обследовании семейных бюджетов работников в отдельных отраслях экономики, производительности труда рабочих предприятия, представленных отдельными группами по квалификации.

Типическая выборка дает более точные результаты по сравнению с другими способами отбора единиц в выборочную совокупность. Типизация генеральной совокупности обеспечивает репрезентативность такой выборки, представительство в ней каждой типологической группы, что позволяет влияние межгрупповой дисперсии на среднюю ошибку выборки. При определении средней ошибки типической выборки в качестве показателя вариации выступает средняя из внутригрупповых дисперсий.

Серийный отбор удобен в тех случаях, когда единицы совокупности объединены в равновеликие группы или серии с тем чтобы в таких группах подвергать наблюдению все без исключения единицы. В качестве серий могут рассматриваться упаковки с готовой продукцией, партии товара,

студенческие группы, бригады и другие объединения. Сущность серийной выборки заключается в собственно-случайном или механическом отборе серий, внутри каждой из которых производится сплошное обследование единиц. Применение серийной выборки обусловлено тем, что многие товары для их транспортировки, хранения и продажи упаковываются в пачки, ящики и т.п. Поэтому при контроле качества упакованного товара рациональнее проверить несколько упаковок (серий), чем из всех упаковок отбирать необходимое количество товара.

Поскольку внутри групп (серий) обследуются все без исключения единицы, средняя ошибка выборки зависит только от межгрупповой (межсерийной) дисперсии.

Комбинированный отбор представляет собой различные сочетания уже рассмотренных видов выборки. Так, можно комбинировать типическую и серийную выборки, когда серии выбираются в установленном порядке из нескольких типических групп. Возможно также комбинирование серийного и собственно-случайного отбора, при котором отдельные единицы отбираются внутри серии в собственно-случайном порядке.

7.3. Ошибка выборки

Важнейшим требованием, предъявляемым к статистическому исследованию, является достоверность статистических данных. Расхождение между величиной какого-либо показателя, определенной по данным статистического исследования, и действительным значением изучаемого показателя называется *ошибкой наблюдения*. В зависимости от причин возникновения различают ошибки регистрации и ошибки репрезентативности.

Ошибки регистрации могут иметь место *при сплошном и несплошном наблюдении*. Они возникают вследствие неправильного установления фактов в процессе наблюдения или неправильной их записи и подразделяются на случайные и систематические.

Случайные ошибки регистрации представляют собой результат действия различных случайных факторов (*например, перепутаны строки или графы статистического формуляра*). Они имеют разную направленность: могут как повышать, так и понижать значения обобщающих статистических показателей. При достаточно большом объеме обследуемой совокупности в результате действия закона больших чисел эти ошибки взаимопогашаются. *Систематические ошибки регистрации* имеют одинаковую тенденцию либо к увеличению, либо к уменьшению значений признака по каждой единице наблюдения, поэтому они не компенсируют друг друга, а величина обобщающего статистического показателя по совокупности в целом будет включать в себя накопленную ошибку. *Например, снимаются показания с неправильно установленного прибора.*

Систематические ошибки регистрации могут быть преднамеренными и непреднамеренными. *Преднамеренные систематические ошибки регистрации* – сознательное тенденциозное искажение фактов как опрашиваемыми, так и регистраторами. *Непреднамеренные систематические ошибки регистрации* вызываются различными причинами, например небрежностью.

В отличие от ошибок регистрации ошибки репрезентативности характерны только для *несплошного наблюдения*. *Ошибкой репрезентативности* называется отклонение значения показателя по обследованной совокупности от его величины по исходной совокупности. *Случайные ошибки* возникают вследствие того, что отобранная совокупность неполно воспроизводит изучаемую совокупность. *Систематические ошибки репрезентативности* появляются вследствие нарушения принципов случайного отбора единиц из исходной совокупности.

Ошибок регистрации и систематических ошибок репрезентативности можно избежать при правильной организации и проведении наблюдения. Случайных ошибок репрезентативности избежать невозможно, поскольку они возникают в силу того, что выборочная совокупность не полностью воспроизводит генеральную. Однако *среднюю величину случайной ошибки можно рассчитать*, пользуясь методами теории вероятностей.

Применяя выборочной метод, в статистике используют два вида обобщающих показателей: 1) средняя величина количественного признака; 2) относительная величина альтернативного признака (доля единиц совокупности, обладающих изучаемым признаком).

В соответствии с двумя видами обобщающих показателей существует два вида формул средней ошибки выборки, которые мы приведем применительно к собственно случайной и механической выборке:

1) для расчета *средней ошибки средней величины количественного признака* в выборке:

при повторном отборе

$$\mu_x = \sqrt{\frac{\sigma^2}{n}};$$

для бесповторном отборе

$$\mu_x = \sqrt{\frac{\sigma^2}{n} \left(1 - \frac{n}{N}\right)},$$

где σ^2 – генеральная дисперсия; n – объем выборочной совокупности (число обследованных единиц); N – объем генеральной совокупности (число составляющих ее единиц).

2) для расчета *средней ошибки доли единиц совокупности, обладающих изучаемым признаком*:

при повторном отборе

$$\mu_w = \sqrt{\frac{w(1-w)}{n}};$$

при бесповторном отборе

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

где w – доля единиц, обладающих изучаемым признаком в выборочной совокупности (выборочная доля).

Исходя из формул, можно утверждать, что средняя величина случайной ошибки репрезентативности зависит от принятого способа формирования выборочной совокупности, объема выборки, степени колеблемости изучаемого признака в генеральной совокупности. Сопоставление формул повторного и бесповторного отбора свидетельствует о том, что применение последнего приводит к уменьшению ошибки выборки.

Для решения практических задач расчета средней ошибки выборки недостаточно. Так, из генеральной совокупности может быть получено несколько выборок и фактическая ошибка каждой конкретной выборки может быть больше или меньше средней ошибки выборки. Поэтому помимо средней, рассчитывается *предельная ошибка выборки*, величина которой зависит от того, с какой вероятностью должна гарантироваться ошибка выборки. Уровень доверительной вероятности определяется при помощи специального коэффициента t , называемого *коэффициентом доверия*. В социально-экономических исследованиях наиболее часто употребляются следующие уровни доверительной вероятности и значения t :

$$P = 0,683 \Rightarrow t = 1;$$

$$P = 0,954 \Rightarrow t = 2;$$

$$P = 0,997 \Rightarrow t = 3.$$

Расчет предельной ошибки выборки производится по формулам

$$\Delta_x = t \cdot \mu_x;$$

$$\Delta_w = t \cdot \mu_w.$$

Величина генеральной средней или доли представляется в виде пределов следующим образом:

$$\bar{x} - \Delta_x \leq \bar{x} \leq \bar{x} + \Delta_x;$$

$$w - \Delta_w \leq p \leq w + \Delta_w.$$

$\bar{\bar{o}}$ – средняя величина признака в генеральной совокупности (генеральная средняя); \bar{o} – средняя величина признака в выборочной совокупности (выборочная средняя); p – доля единиц, обладающих изучаемым признаком в генеральной совокупности (генеральная доля).

Таким образом, по результатам выборочного наблюдения с определенной степенью достоверности можно утверждать, что средняя величина признака или доля единиц, обладающих изучаемым признаком в генеральной совокупности, не выйдет за установленные пределы.

7.4. Определение необходимой численности выборки

На стадии организации выборочного наблюдения решается вопрос о том, каков должен быть объем выборочной совокупности, для того, чтобы была обеспечена требуемая точность результатов наблюдений. Уменьшение ошибки выборки, а, следовательно, увеличение точности определения параметров генеральной совокупности всегда связано с увеличением объема выборки.

Увеличивая численность выборки, можно довести ее ошибку до сколь угодно малых размеров. Однако из формул средней ошибки выборки следует, что уменьшение ошибки в k раз требует увеличения объема выборки в k^2 раз. Увеличение объема исследований, в свою очередь, вызывает дополнительные затраты труда и средств, снижает оперативность информации. Поэтому вопрос об оптимальной численности выборки имеет важное практическое значение.

Определение необходимой численности выборки основывается на формуле ее предельной ошибки. При случайном повторном отборе объем выборки получается в результате преобразования соответствующей формулы:

$$\Delta_x = t \mu_x = t \sqrt{\frac{\sigma^2}{n}}; \quad n = \frac{t^2 \sigma^2}{\Delta_x^2}.$$

Таким же образом выводятся формулы для расчета объема выборки при других способах отбора. Расчетную величину с целью получения запаса точности округляют в большую сторону. Для упрощения расчетов при определении объема бесповторной выборки может использоваться формула для повторной выборки, что также дает запас точности. Ниже представлены формулы для определения необходимой численности выборки применительно к собственно случайной и механической выборке:

1) при определении *средней величины количественного признака* в выборке:

при повторном отборе

$$n = \frac{t^2 \sigma^2}{\Delta_x^2};$$

для бесповторном отборе

$$n = \frac{t^2 \sigma^2 N}{\Delta_x^2 N + t^2 \sigma^2};$$

2) при определении *доли единиц совокупности*, обладающих изучаемым признаком:

при повторном отборе

$$n = \frac{t^2 w(1-w)}{\Delta_w^2};$$

при бесповторном отборе

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

Дисперсия признака в генеральной совокупности зачастую бывает неизвестна. Поэтому используют следующие приблизительные способы ее определения:

- 1) используются данные предыдущих обследований;
- 2) проводятся несколько пробных обследований и выбирается наибольшее значение дисперсии;
- 3) если распределение признака в генеральной совокупности подчиняется нормальному закону, $\delta = (x_{\max} - x_{\min}) / 6$.
- 4) при изучении альтернативного признака берется максимально возможная величина дисперсии, равная 0,25 при $w = 0,5$.

Если целью выборочного наблюдения является изучение различных признаков с неодинаковой колеблемостью, то при определении необходимого объема выборки следует ориентироваться на тот признак,

который при наибольшей колеблемости обладает наименьшей величиной допустимой ошибки.

Иногда на практике задается величина *относительной предельной ошибки*, выраженной в процентах к средней:

$$\Delta_{\%} = \frac{\Delta_{\bar{v}}}{\bar{v}} \cdot 100, \%$$

В этом случае формулы для расчета необходимого объема выборки также получаются в результате преобразования соответствующих формул предельной ошибки выборки:

$$\Delta_{\bar{x}} = \frac{\Delta_{\%} \cdot \bar{x}}{100} \Rightarrow n = \frac{t^2 \sigma^2}{\Delta_{\%}^2 \cdot (\bar{v})^2} \cdot 100^2 = \frac{t^2 v^2}{\Delta_{\%}^2}, \text{ поскольку } v = \frac{\sigma}{\bar{x}} \cdot 100.$$

7.5. Распространение характеристик выборки на генеральную совокупность

Заключительным этапом выборочного наблюдения является распространение его результатов на генеральную совокупность.

Выборочные средние и относительные величины распространяют на генеральную совокупность с учетом предела их возможной ошибки.

Существуют два основных способа распространения выборочных данных:

- 1) способ прямого пересчета;
- 2) способ поправочных коэффициентов.

Способ прямого пересчета применяется тогда, когда с помощью средних, полученных по данным выборочного наблюдения определяются объемные показатели. *Например, по данным выборочного наблюдения известно, что в микрорайоне число детей школьного возраста в семье в среднем составляет 1,2 чел. при ошибке выборки 0,1 чел. Общее количество семей в микрорайоне – 1000. Следовательно, для обучения всех детей школьного возраста, проживающих в микрорайоне, необходимо от 1100 до 1300 школьных мест.*

Способ поправочных коэффициентов применяется для уточнения и проверки данных сплошного наблюдения. На основании данных выборочного наблюдения рассчитываются поправочные коэффициенты, по которым производится корректировка итогов сплошного наблюдения. *Например, в городе зарегистрировано 2000 торговых палаток, в том числе в одном из его районов – 400. По результатам контрольных обходов установлено, что в этом районе находится 420 палаток. Коэффициент недосчета торговых палаток составляет 1,05 (420/400). Отсюда общее количество торговых палаток в городе составляет 2100 (2000*1,05).*

Контрольные вопросы

1. Какое наблюдение называется сплошным.
2. Какое наблюдение называется несплошным.
3. Обследование основного массива.
4. Монографическое наблюдение.
5. Анкетное наблюдение.
6. Какое наблюдение называется выборочным?
7. В чем преимущество выборочного наблюдения перед сплошным?
8. Какая совокупность называется генеральной?
9. Какая совокупность называется выборочной?
10. Принципы выборочного метода.
11. Как производится собственно-случайный отбор?
12. В чем различие повторной и бесповторной выборки?
13. Как производится механический отбор?
14. Как производится типический отбор?
15. Как производится серийный отбор?
16. В чем заключается комбинированный отбор?
17. Что такое ошибка регистрации?
18. Случайные ошибки регистрации.
19. Систематические ошибки регистрации.
20. Ошибка репрезентативности.
21. Что представляет собой средняя ошибка выборки?
22. По каким расчетным формулам находят средние ошибки выборки (для средней и доли) при повторном и бесповторном отборах?
23. Порядок определения необходимой численности выборки.
24. Распространение характеристик выборки на генеральную совокупность.

Тема 8. Статистическое изучение динамики социально-экономических явлений (Анализ динамики)

8.1. Ряды динамики и правила их построения

Одной из важнейших задач статистики является изучение изменений анализируемых показателей во времени, т.е. их динамики. Эта задача решается при помощи анализа рядов динамики (или временных рядов).

Рядами динамики являются статистические данные, отображающие развитие различных общественных явлений во времени.

Ряд динамики (или динамический ряд) представляет собой ряд расположенных в хронологической последовательности числовых значений статистического показателя, характеризующих изменение общественных явлений во времени.

В каждом ряду динамики имеются два основных элемента: показатели времени t и соответствующие им уровни изучаемого явления (уровни ряда) y .

Время – это моменты или периоды, к которым относятся уровни. В качестве показателей времени выступают определенные даты (моменты) времени либо отдельные периоды (годы, кварталы, месяцы, сутки).

Уровни ряда – это показатели, числовые значения которых составляют динамический ряд. Уровни ряда динамики отображают количественную оценку (меру) развития изучаемого явления во времени, могут выражаться абсолютными, относительными и средними величинами.

Ряды динамики, как правило, представляют в виде таблицы (см. таблицу 1) или графически (см. рисунок 1), причем по оси абсцисс строится шкала времени t , а по оси ординат – шкала уровней ряда y .

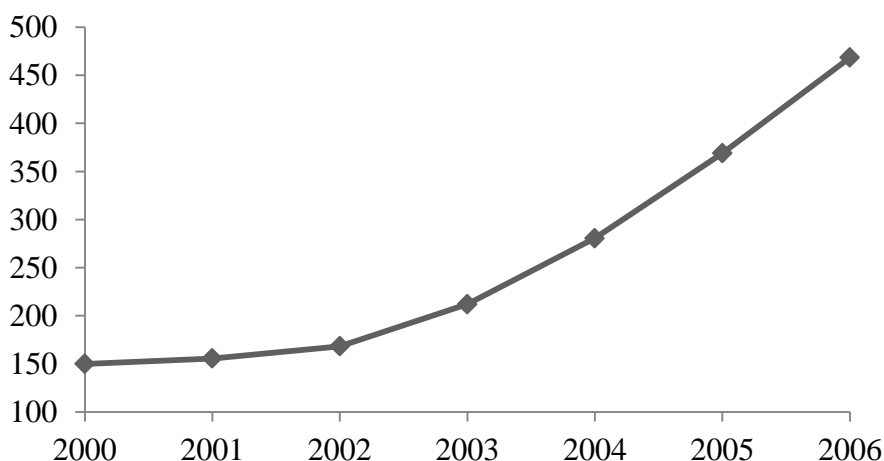


Рисунок 1 – Внешнеторговый оборот страны за 2000 – 2006 гг.

Таблица 1 – Внешнеторговый оборот страны за период 2000-2006 гг.

Год	2000	2001	2002	2003	2004	2005	2006
Млрд. долл. США	149,9	155,6	168,3	212,0	280,6	368,9	468,4

Графическое изображение позволяет наглядно представить развитие явления во времени и способствует проведению анализа уровней. Наиболее распространенным видом графического изображения для аналитических целей является линейная диаграмма, которая строится в прямоугольной системе координат.

Наряду с линейной диаграммой для графического ряда динамики в целях популяризации широко используются столбиковая диаграмма, секторная диаграмма и другие виды диаграмм (фигурные, квадратные, полосовые и т.п.)

Построение и анализ рядов динамики позволяют выявить и измерить закономерности развития общественных явлений во времени. Эти закономерности не проявляются четко на каждом конкретном уровне, а лишь в тенденции, в достаточно длительной динамике. На основную закономерность динамики накладываются другие, прежде всего, случайные, иногда сезонные влияния. Выявление основной тенденции в изменении уровней, именуемой *трендом*, является одной из главных задач анализа рядов динамики.

Ряды динамики состоят из 3 элементов: заголовков, моменты или промежутки времени, числовые характеристики моментов или промежутков.

Для сопоставимости данных ряда динамики необходимо соблюдать следующие условия:

- одинаковая полнота охвата явлений;
- показатели динамического ряда должны быть четко ограничены в отношении охвата территории;
- методология учета и расчета показателей должны быть одинакова;
- продолжительность периода, к которому относится уровень ряда, должна быть одинакова.

Все показатели ряда должны быть выражены в одинаковом измерении

В зависимости от характера показателя, использованного при построении рядов динамики, различают:

- а) ряд абсолютных величин (показатели объемов);
- б) ряд относительных величин;
- в) ряд средних величин.

В зависимости от характера изучаемого явления ряды абсолютных величин подразделяются на моментные и интервальные ряды динамики.

Моментный ряд динамики отображает состояние изучаемого явления на определенные даты (моменты) времени.

В моментных рядах уровни выражают состояние явления на критический момент времени – начало месяца, квартала, года и т.д.

Например, численность населения, численность работающих и т.д. Особенностью моментного ряда динамики является то, что в его уровни могут входить одни и те же единицы изучаемой совокупности.

В таких рядах каждый последующий уровень полностью или частично содержит значение предыдущего уровня, поэтому суммировать уровни нельзя, так как это приводит к повторному счету.

Интервальный (периодический) ряд динамики отображает итоги развития изучаемых явлений за отдельные периоды (интервалы времени).

В интервальных рядах уровни отражают состояние явления за определенный период времени – сутки, месяц, год и т.д. Это ряды показателей объема производства, объема продаж по месяцам года, количества отработанных человеко-дней и т.д.

Примером интервального ряда динамики могут служить данные о производстве продукции за отчетный год.

Особенностью интервального ряда динамики является то, что каждый его уровень складывается из данных за более короткие интервалы времени.

Значения уровней интервального ряда, в отличие от уровней моментного ряда, не содержатся в предыдущих или последующих показателях, их можно просуммировать, что позволяет получать ряды динамики более укрупненных периодов.

Например, суммируя объем производства за три месяца получают его объем за квартал. Суммирование уровней за последовательные интервалы времени позволяет получать ряды динамики более крупных периодов.

Интервальный ряд, где последовательные уровни могут суммироваться, можно представить как ряд с нарастающими итогами. При построении таких рядов производится последовательное суммирование смежных уровней. Этим достигается суммарное обобщение результата развития изучаемого явления с начала отчетного периода (месяца, квартала, года и т.д.).

Статистическое отображение развития изучаемого явления во времени с учетом предшествующих периодов может быть представлено интервальным рядом с нарастающими итогами. При составлении такого ряда производится последовательное суммирование смежных уровней интервального ряда динамики.

Обобщенной характеристикой ряда динамики служит, прежде всего, средний уровень ряда \bar{y} . Для разных видов рядов динамики он рассчитывается неодинаково. Ряды динамики бывают равномерные, с равностоящими уровнями (с равными интервалами времени между уровнями, то есть данные представлены через равные, следующие друг за другом интервалы времени), для которых средний уровень определяется по простой формуле средней величины, и неравномерные, с неравностоящими уровнями (с неравными интервалами, когда в рядах динамики прерывающиеся или неравномерные интервалы времени), для которых используются формулы средних взвешенных (по интервалам времени)

величин. В интервальном ряду динамики (в котором время задано в виде промежутков времени, к которым относятся уровни) \bar{y} определяется по формуле средней арифметической, а в моментном ряду (в котором время задано в виде конкретных моментов времени или дат, к которым относятся уровни) – по формуле средней хронологической. В таблице 2 приводятся виды рядов динамики и соответствующие формулы для расчета их среднего уровня \bar{y} .

Таблица 2 – Виды рядов динамики и формулы их среднего уровня

<i>Вид ряда динамики</i>	<i>Название средней величины</i>	<i>Формула средней величины</i>
Равномерный интервальный	Арифметическая простая	$\bar{y} = \frac{\sum y}{n}$
Равномерный моментный	Хронологическая простая	$\bar{y} = \frac{\frac{y_1 + y_2 + y_3 + \dots + y_{n-1} + y_n}{2}}{n-1} = \frac{\frac{y_1 + y_n}{2} + \sum_{i=2}^{n-1} y_i}{n-1}$
Неравномерный интервальный	Арифметическая взвешенная	$\bar{y} = \frac{\sum_{i=1}^n y_i t_i}{\sum_{i=1}^n t_i}$
Неравномерный моментный	Хронологическая взвешенная	$\bar{y} = \frac{\sum_{i=1}^{n-1} (y_i + y_{i+1}) t_i}{2 \sum_{i=1}^{n-1} t_i}$

8.2. Сопоставимость рядов динамики

Необходимым условием для получения правильных выводов при анализе рядов динамики является сопоставимость уровней ряда между собой. Статистические данные должны быть сопоставимы:

- 1) по территории (соблюдение одних и тех же территориальных границ для всех уровней ряда);
- 2) по кругу охватываемых объектов (сравнение совокупностей с равным числом элементов);
- 3) по времени регистрации (равенство периодов времени, за которые приводятся данные (для интервального ряда) или представление показателей на одну и ту же дату (для моментного ряда));
- 4) по методологии расчета (использование единой методологии расчета всех уровней ряда);
- 5) по единицам измерения (использование одних и тех же единиц измерения для всех уровней ряда);

б) по ценам (использование постоянных или сопоставимых цен).

Возможными являются и другие причины несопоставимости. В ряде случаев несопоставимые данные могут быть приведены к сопоставимому виду путем дополнительных расчетов. В частности, в статистике применяется прием, известный как *смыкание рядов динамики*. Этот прием позволяет преодолеть несопоставимость данных, возникающую вследствие изменения во времени территориальных границ, круга охватываемых объектов или методологии расчета показателей, и получить единый ряд условно сопоставимых данных за весь период времени.

Проблема приведения к сопоставимому виду возникает и при параллельном анализе развития во времени экономических показателей отдельных объектов. В таких случаях ряды динамики приводятся к общему основанию, т. е. к одному и тому же периоду или моменту времени, уровень которого принимается за базу сравнения, а все остальные уровни выражаются в виде коэффициентов или в процентах по отношению к нему.

8.3. Аналитические показатели ряда динамики

При изучении динамики общественных явлений возникает проблема описания интенсивности изменения и расчета средних показателей динамики.

При изучении динамики явлений или процессов для описания интенсивности происходящих изменений используются показатели, получаемые в результате сравнения уровней ряда динамики.

К аналитическим показателям ряда динамики относятся: *абсолютный прирост, темп (коэффициент) роста, темп (коэффициент) прироста, абсолютное значение одного процента прироста*.

Система средних показателей включает *средний уровень ряда, средний абсолютный прирост, средний темп роста, средний темп прироста*.

При расчете аналитических показателей сравниваемый уровень ряда динамики принято называть *отчетным*, а уровень, с которым производится сравнение, – *базой сравнения, базисным*. Показатели анализа динамики могут вычисляться с постоянной и переменной базой сравнения. Для расчета показателей динамики с постоянной базой, каждый уровень сравнивается с одним и тем же базисным уровнем. В качестве базисного выбирается начальный уровень в ряду динамики, либо уровень, с которого начинается какой-то новый этап развития явления. Исчисляемые при этом показатели называются базисными. Для расчета показателей динамики с переменной базой каждый последующий уровень ряда сравнивается с предыдущим.

Вычисленные таким образом показатели называются *цепными*.

Аналитические показатели ряда динамики представлены в следующей таблице. При этом использованы следующие условные обозначения:

y_i – отчетный уровень i -го периода;

y_{i-1} – уровень предшествующего периода;

y_0 – уровень базисного периода.

Важнейшим статистическим показателем анализа динамики является абсолютный прирост (сокращение), то есть абсолютное изменение, характеризующее увеличение или уменьшение уровня ряда за определенный промежуток времени. *Абсолютный прирост (A)* – абсолютное увеличение (уменьшение) уровня ряда за определенный промежуток времени. Он равен разности двух сравниваемых уровней и выражает абсолютную скорость роста. Абсолютный прирост с переменной базой называют *скоростью роста*.

Цепные и базисные абсолютные приросты связаны между собой: сумма последовательных цепных абсолютных приростов равна базисному (общему) приросту за весь рассматриваемый промежуток времени.

Для оценки интенсивности, то есть относительного изменения уровня динамического ряда за какой-либо период времени исчисляют *темпы роста (снижения)*. Интенсивность изменения уровня оценивается отношением отчетного уровня к базисному. Показатель интенсивности изменения уровня ряда, выраженный в долях единицы, называется *коэффициентом роста*, а в процентах – *темпом роста*. Эти показатели интенсивности изменения отличаются только единицами измерения.

Коэффициент роста показывает, во сколько раз данный сравниваемый уровень ряда больше базисного уровня с которым производится сравнение (если этот коэффициент больше единицы) или какую часть базисного уровня составляет уровень текущего периода за некоторый промежуток времени (если он меньше единицы). Темп роста всегда представляет собой положительное число. Если темп роста равен 100%, то значение уровня не изменилось, если больше 100%, то значение уровня повысилось, а если меньше 100% - понизилось.

Между цепными и базисными коэффициентами роста существует взаимосвязь. Произведение цепных коэффициентов роста равно общему (за весь период) базисному коэффициенту роста, а частное от деления i -го базисного коэффициента роста на предыдущий равно i -му цепному коэффициенту роста.

Относительную оценку скорости измерения уровня ряда в единицу времени дают показатели темпа прироста (сокращения).

Коэффициент прироста дает относительную оценка изменения уровня ряда в единицу времени. Темп прироста – коэффициент прироста, выраженный в %-тах, показывает, на сколько %-тов сравниваемый уровень больше уровня, с которым производится сравнение.

Темп прироста характеризует абсолютный прирост в относительных величинах. Определенный в процентах темп прироста показывает, на сколько процентов изменился сравниваемый уровень по отношению к уровню, принятому за базу сравнения. Темп прироста рассчитывается как отношение абсолютного прироста к уровню, принятому за базу сравнения.

Темп прироста может быть положительным, отрицательным или равным нулю.

Сравнение абсолютного прироста и темпа прироста за одни и те же периоды времени показывает, что в реальных экономических процессах замедление темпов прироста не всегда сопровождается уменьшением абсолютных приростов. Поэтому на практике часто проводят сопоставление этих показателей. Для этого рассчитывают *абсолютное значение одного процента прироста*. Оно представляет собой одну сотую часть базисного уровня и в то же время – отношение абсолютного прироста к соответствующему темпу прироста, %.

Абсолютное значение 1%-та прироста – «цена» каждого 1%-та прироста (сотая часть предыдущего или базисного уровня). Оно показывает, какое абсолютное значение скрывается за относительным показателем – одним процентом прироста.

Для более глубокого понимания характера явления необходимо показатели динамики анализировать комплексно, совместно.

Базисные показатели динамики характеризуют окончательный результат всех изменений в уровнях ряда от периода, к которому относится базисный уровень, до данного (*i*-го) периода.

Цепные показатели динамики характеризуют интенсивность изменения уровня от периода к периоду (или от даты к дате) в пределах изучаемого промежутка времени (рисунок 2).

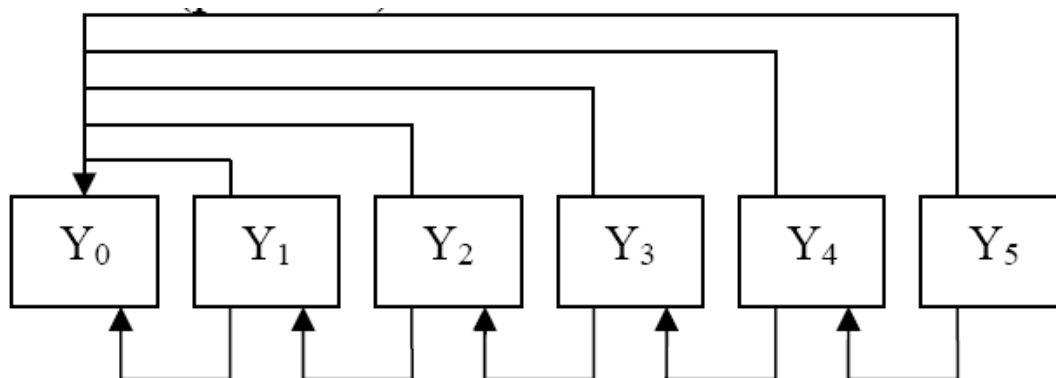


Рисунок 2 - Построение цепных и базисных аналитических показателей динамики

8.4. Методы анализа основной тенденции развития в рядах динамики

Одной из важнейших задач статистики является определение в рядах динамики общей тенденции развития явления.

Основной тенденцией ряда динамики (трендом) называется устойчивое изменение уровня явления во времени, обусловленное влиянием постоянно действующих факторов и свободное от случайных колебаний.

В случаях, когда уровни динамического ряда непрерывно растут или непрерывно снижаются, основная тенденция ряда является очевидной.

Однако чаще уровни динамических рядов претерпевают разнонаправленные изменения (то растут, то убывают), и общая тенденция неясна.

Для выявления основной тенденции такие ряды динамики подвергаются обработке методами усреднения по левой и правой колонке, укрупнения интервалов, скользящей средней или аналитического выравнивания ряда.

Метод усреднения по левой и правой половине: Разделяют ряд динамики на две части, находят для каждой из них среднее арифметическое значение и проводят через полученные точки линию тренда на графике.

Если рассматривать уровни экономических показателей за короткие промежутки времени, то в силу влияния различных факторов, действующих в разных направлениях, в рядах динамики наблюдаются снижение и повышение этих уровней. Суть данного метода заключается в том, что колебание абсолютных уровней динамики ряда, скрывающие ему тенденции, устраняются и закономерность проступает более отчетливо.

Укрупнение интервалов основано на увеличении периодов времени, к которым относятся уровни ряда динамики. Одновременно уменьшается количество интервалов. Недостатком этого метода является то, что необходимо всегда обеспечить равенство укрупненных интервалов.

По своей сути метод *скользящей средней* похож на метод укрупнения интервалов, но в данном случае фактические уровни заменяются средними уровнями, рассчитанными для последовательно подвижных (скользящих) укрупненных интервалов, охватывающих m уровней ряда. Например, если принять $m=3$, то сначала рассчитывается средняя величина из первых трех уровней, затем находится средняя величина из 2-го, 3-го и 4-го уровней, потом из 3-го, 4-го и 5-го и т.д., т.е. каждый раз в сумме трех уровней появляется новый уровень, а два остаются прежними, что и обуславливает взаимопогашение случайных колебаний в средних уровнях. Рассчитанные из m членов скользящие средние относятся к середине (центру) каждого рассматриваемого интервала. Таким образом, средняя «скользит» по ряду динамики, передвигаясь на один уровень. Недостатком метода является потеря информации, связанная с укорачиванием ряда.

Сглаживание методом скользящей средней можно проводить по любому числу членов m , но удобнее, если m – нечетное число, так как в этом случае скользящая средняя сразу относится к конкретной временной точке – середине (центру) интервала. Если же m – четное, то скользящая средняя относится к промежутку между временными точками: например, при сглаживании по четырем членам ($m=4$) средняя из первых четырех уровней будет находиться между второй и третьей временной точкой, следующая – между третьей и четвертой и т.д. Тогда, чтобы сглаженные уровни относились непосредственно к конкретным временным точкам, из каждой пары смежных промежуточных значений скользящих средних находят среднюю арифметическую, которую относят к временной точке,

находящейся между смежными. Такой прием двойного расчета сглаженных уровней называется *центрированием*.

Недостатком метода скользящей средней является то, что сглаженный ряд укорачивается по сравнению с фактическим с двух концов: при нечетном m на $(m-1)/2$, а при четном m – на $m/2$ с каждого конца. Применяя этот метод, надо помнить, что он сглаживает (устраняет) лишь случайные колебания. Если же, например, ряд содержит сезонную волну, она сохранится и после сглаживания методом скользящей средней. Кроме того, этот метод сглаживания, как и метод укрупнения интервалов не позволяет выразить общую тенденцию изменения уровней в виде математической модели.

Наиболее совершенным методом обработки рядов динамики в целях *устранения случайных колебаний и выявления тренда является выравнивание уровней ряда по аналитическим формулам (или аналитическое выравнивание)*. Суть аналитического выравнивания заключается в замене эмпирических (фактических, исходных) уровней y_i теоретическими \hat{y}_i , которые рассчитаны по определенному уравнению, принятому за математическую модель тренда, где теоретические уровни рассматриваются как функция времени: $\hat{y}_i = f(t)$.

Аналитическое выравнивание часто производится:

1. по прямой линии;
2. по параболе любого порядка;
3. по показательной кривой.

Выявление и характеристика трендов и моделей взаимосвязи создают базу для прогнозирования, т. е. для определения ориентировочных размеров явлений в будущем. Для этого используют метод экстраполяции, предполагающий подстановку в уравнение тренда значений t за пределами исследуемого ряда. Поскольку же в действительности тенденция развития не остается неизменной, то данные, получаемые путем экстраполяции ряда, рассматриваются как предварительные вероятностные оценки. Для составления окончательного прогноза, как правило, привлекается дополнительная информация, не содержащаяся в динамическом ряду.

8.5. Методы изучения сезонных колебаний

При сравнении месячных и квартальных данных многих социально-экономических явлений часто обнаруживаются периодические колебания, возникающие под влиянием смены времени года. Они являются результатом влияния природно-климатических условий, действия общеэкономических, а также других многочисленных и разнообразных факторов, которые часто являются регулируемы.

В статистике периодические колебания, которые имеют определенный и постоянный период, равный годовому промежутку, носят название сезонных колебаний или сезонной волны, а динамический ряд в этом случае называется сезонным рядом динамики.

Сезонные колебания наблюдаются во многих отраслях экономики (сельское хозяйство, строительство, транспорт, промышленность строительных материалов, энергетика и др.). В ряде случаев они отрицательно влияют на результаты производственной деятельности.

Поэтому встает вопрос о регулировании сезонных изменений. В основе этого регулирования должно лежать исследование сезонных колебаний.

К сезонным явлениям относятся вообще все явления, которые обнаруживают в своем развитии определенную закономерность внутри годичной динамики, более или менее устойчиво сохраняющуюся в течение ряда лет.

Важнейшими задачами при исследовании сезонности являются:

1. определение наличия сезонности, численное выражение проявления сезонных колебаний и выявление проявления сезонных колебаний и выявление их силы и характера в различных фазах годичного цикла;
2. характеристика факторов, вызывающих сезонные колебания;
3. оценка последствий, к которым приводит наличие сезонных колебаний;
4. математическое моделирование сезонности.

В статистике существует ряд методов изучения и измерения сезонных колебаний. Самый простой из них предполагает расчет индексов сезонности и построение сезонной волны. Для того чтобы выявить устойчивую сезонную волну, на которой не отражались бы случайные условия одного года, индексы сезонных колебаний вычисляются по данным за несколько лет (не менее трех).

Если ряд динамики не содержит ярко выраженной тенденции в развитии, то индексы сезонности вычисляются непосредственно по эмпирическим данным без их предварительного выравнивания. При этом используется следующий алгоритм:

- рассчитывается средний уровень каждого месяца по данным за несколько лет;
- вычисляется среднемесячный уровень для ряда динамики в целом;
- определяются индексы сезонности как процентные отношения средних для каждого месяца к общему среднемесячному уровню ряда.

В случае, когда динамический ряд проявляют тенденцию к росту или снижению, фактические (эмпирические) уровни сопоставляются с теоретическими (полученными на основе аналитического выравнивания по тренду) и рассчитываются индексы сезонности

Самыми простыми вариантами определения сезонных колебаний, не требующие предварительной обработки эмпирического ряда, является: способ А) и способ Б).

Измерение сезонных колебаний при использовании этих методов сводится к следующим расчетам:

Способ А) – способ абсолютных разниц:

- из имеющихся за ряд лет помесечных данных исчисляются среднеарифметические для каждого месяца;
- определяется общая средняя арифметическая для всего обрабатываемого ряда динамики;
- определяется отклонения (разности) средней по отдельным месяцам от общей средней месячной. Эти разности и характеризуют сезонные колебания являясь звеньями сезонной волны ($\bar{X}_1 - \bar{X}_{общ}$; $\bar{X}_2 - \bar{X}_{общ} \dots$)

Способ Б) – способ относительных разниц: Производятся все вычисления по аналогии со способом А, но только вычисленные абсолютные отклонения выражаются в % к общей средней месячной за весь период.

$$\left(\frac{\bar{X}_1 - \bar{X}_{общ}}{\bar{X}_{общ}} * 100\% \dots \right)$$

Индекс сезонности, показывает во сколько раз фактический уровень ряда в момент или интервал времени больше среднего уровня, либо уровня вычисляемого по уравнению тенденции.

Контрольные вопросы

1. Для чего нужно изучать динамику явлений?
2. Дайте определение ряда динамики.
3. Определите уровни ряда.
4. Из каких элементов состоит ряд динамики, и каков их смысл?
5. Какие динамические ряды называются моментными и почему их уровни нельзя суммировать?
6. Какие динамические ряды называются интервальными?
7. Условия сопоставимости данных ряда динамики.
8. Определите равномерный и неравномерный ряды динамики.
9. Сопоставимость рядов динамики.
10. Что характеризует показатели абсолютного прироста и как они исчисляются?
11. Что представляет собой темп роста? Как он исчисляется?
12. Какая существует взаимосвязь между последовательными цепными коэффициентами роста и базисным коэффициентом роста за соответствующий период?
13. Что показывает абсолютное значение одного процента прироста и как оно исчисляется?
14. Какими наиболее распространенными статистическими методами осуществляется изучение тренда в рядах динамики?
15. В чем сущность метода укрупнения интервалов и для чего он применяется?

16. Как производится сглаживание рядов динамики способом скользящей средней?
17. Что представляют собой сезонные колебания, в чем практическое значение их изучения?
18. Что представляет собой индекс сезонности?

Тема 9. Индексный метод в статистических исследованиях

9.1. Понятие об индексах, их классификация

К важнейшим обобщающим показателям относят индексы.

«Индекс» (index) в переводе с латинского буквально означает указатель или показатель. Обычно этот термин используется для обобщающей характеристики изменений.

Под *индексом* в статистике понимают относительный показатель, характеризующий изменение величины какого-либо явления (простого или сложного, состоящего из соизмеримых или несоизмеримых элементов) во времени, пространстве или по сравнению с любым эталоном (нормативом, планом, прогнозом и т. п.). Соответственно говорят об *индексах динамики (изменение во времени)*, *территориальных индексах (изменение в пространстве)* либо об *индексах выполнения планов, заданий, прогнозов (при сопоставлении с уровнем, например, договорных обязательств)* и т. п.

Индексы являются незаменимым инструментом исследования в тех случаях, когда необходимо сравнить во времени или в пространстве две совокупности, элементы которых являются несоизмеримыми величинами.

В индексе различают величину сравнения (числитель индексного отношения) и знаменатель индексного отношения (базу сравнения).

Основным элементом индексного отношения является *индексируемая величина*. Под *индексируемой величиной* понимают значение признака статистической совокупности, изменение которой является объектом изучения.

Поскольку объекты изучения индексов весьма разнообразны, то они широко применяются в экономической практике.

С помощью индексов решаются три главные задачи:

- 1) оценка изменения сложного явления;
- 2) определение влияния отдельных факторов на изменение сложного явления;
- 3) сравнение величины какого-то явления с величиной прошлого периода, величиной по другой территории, а также с нормативами, планами, прогнозами.

Индексы классифицируют по трем признакам:

- 1) по содержанию индексируемых величин;
- 2) по степени охвата элементов совокупности;
- 3) по методам расчета общих индексов.

По содержанию индексируемых величин, характеру изучаемых явлений индексы разделяются на индексы количественных показателей и индексы качественных показателей.

Индексы количественных (объемных) показателей – индексы физического объема промышленной и сельскохозяйственной продукции, физического объема розничного товарооборота, численности, национального дохода и др. Индексируемые показатели этих индексов являются *объемными*,

поскольку они характеризуют *общий, суммарный размер (объем) того или иного явления* и выражаются абсолютными величинами. При расчете таких индексов количества оцениваются в одинаковых, *сопоставимых ценах*.

Индексы качественных показателей – индексы курса валют, цен, себестоимости, производительности труда, средней заработной платы, урожайности и др. Индексируемые показатели этих индексов измеряют интенсивность (эффективность) явления или процесса, характеризуя *уровень явления в расчете на количественно измеримую единицу совокупности*: цена за единицу продукции, выработка в единицу времени (или на одного работника), себестоимость единицы продукции, заработная плата одного работника, выработка в единицу времени (или на одного работника), урожайность с одного гектара и т. д. Такие индексы носят расчетный, вторичный характер. Качественные показатели измеряют не общий объем, а *интенсивность, эффективность* явления или процесса. Как правило, они являются либо *средними*, либо *относительными* величинами. Расчет таких индексов производится на базе одинаковых, *неизменных количеств продукции*. Индексы качественных показателей включают также индексы сложных явлений (например, товарооборота, затрат на производство и т. д.).

По степени охвата единиц совокупности индексы делятся на индивидуальные и общие (агрегатные).

В дальнейшем изложении индексного метода будут использоваться следующие общепринятые обозначения:

- 1) i - индивидуальный индекс;
- 2) I - сводный индекс.

Индивидуальные индексы служат для изменения отдельных элементов сложного явления (например, изменение объема выпуска телевизоров определенной марки, рост или падение цен на акции в каком-либо акционерном обществе и т. д.).

Индивидуальный индекс будет снабжаться подстрочным знаком индексируемого показателя: i_q – индивидуальный индекс объема произведенной продукции отдельного вида или количества (объема) проданного товара данного вида, i_p – индивидуальный индекс цен; I_p – общий индекс цен, I_z – общий индекс себестоимости и т. д.

Для характеристики индексов введем следующие условные обозначения, принятые в практике применения индексного метода:

- q – количество (объем) какого-либо продукта в натуральном выражении;
- p – цена единицы продукции, товара;
- z – себестоимость единицы продукции;
- t – затраты времени на производство единицы продукции (трудоемкость);
- w – выработка продукции в натуральном выражении на одного работника или в единицу времени;
- pq – общая стоимость произведенной продукции данного вида или проданных товаров данного вида (товарооборот, выручка);
- zq – затраты на производство всей продукции (издержки производства).

Для того чтобы различать, к какому периоду или объекту относятся индексируемые величины, справа внизу за соответствующим символом указываются подстрочные знаки. Например, в индексах динамики сравниваемого (отчетного) периода используется подстрочный знак 1, а для периодов, с которыми производится сравнение, – 0. Если изменение явления изучается за ряд периодов, то каждый из них обозначается соответственно подстрочными знаками 0, 1, 2, 3 и т. д.

Индивидуальные индексы относятся к одному элементу (явлению) и не требуют суммирования данных. Они представляют собой относительные величины динамики, выполнения обязательств, сравнения. Выбор базы сравнения определяется целью исследования.

Индивидуальные индексы служат для характеристики изменения отдельных элементов сложного явления (*например, изменение объема выпуска продукции одного вида*). Они представляют собой относительные величины динамики, выполнения обязательств, сравнения. Индивидуальные индексы рассчитываются как соотношения двух значений индексируемой величины и выражают в коэффициентах или процентах:

$$i_p = \frac{p_1}{p_0}.$$

Индивидуальный индекс является простейшим показателем, используемым в индексном анализе, который характеризует изменение во времени экономических величин, относящихся к одному объекту:

1. индекс цен: $i_p = p_1/p_0$
2. индекс себестоимости: $i_z = z_1/z_0$
3. индекс производительности труда (выработка): $i_w = w_1/w_0$
4. индекс производительности труда (трудоемкость): $i_t = t_0/t_1$

Изменение физической массы проданного товара в натуральном выражении измеряется индивидуальным индексом физического объема реализации:

$$i_q = \frac{q_1}{q_0}.$$

Разность между числителем и знаменателем формулы представляет собой *абсолютное изменение* выручки, показывает на сколько в денежных единицах (например, рублях) изменилась выручка в отчетном периоде по сравнению с базисным:

$$\Delta Q = Q_1 - Q_0.$$

Аналогично определять индивидуальные индексы можно для любого интересующего показателя (производительности, заработной платы, себестоимости и т. д.).

В частности, поскольку сумма выручки определяется ценой товара p (от англ. «*price*») и количеством (физическим объемом, или объемом продаж в натуральном выражении) q (от англ. «*quantity*») т.е. $Q = qp$, можно

определить соответствующие индивидуальные индексы – цены и количества:

$$i_p = p_1/p_0, \quad i_q = q_1/q_0.$$

Очевидно, что произведение индивидуальных индексов цены и количества дает индивидуальный индекс выручки:

$$i_Q = i_q i_p.$$

Общий (сводный) индекс отражает изменение всех элементов сложного явления. При этом под *сложным явлением* понимают такую статистическую совокупность, отдельные элементы которой непосредственно не подлежат суммированию (физический объем разнородной продукции, цены на разные группы продуктов и т. д.). Общий индекс обозначается буквой *I* и сопровождается подстрочным знаком индексируемого показателя (*например, I_p – общий индекс цен, I_z – общий индекс себестоимости*).

Методика расчета общих индексов различна в зависимости от характера индексируемых показателей, наличия исходных данных и целей исследования.

Общие индексы могут быть построены двумя способами: как *агрегатные* и как *средние из индивидуальных (средние арифметические или средние гармонические)*. Последние, в свою очередь, делятся на *средние арифметические* и *средние гармонические*.

Агрегатный индекс является основной и наиболее распространенной формой индекса, его числитель и знаменатель представляют набор - «агрегат» (от латинского *aggregatus* – складываемый, суммируемый) непосредственно несоизмеримых и не поддающихся суммированию элементов – сумму произведений двух величин, одна из которых меняется (*индексируется*), а другая остается неизменной в числителе и знаменателе (*вес индекса*). Вес индекса служит для целей соизмерения индексируемых величин.

При этом агрегатные индексы качественных показателей могут быть рассчитаны как *индексы переменного состава* (на базе изменяющихся структур явлений) или *индексы постоянного (фиксированного) состава* (на базе неизменной структуры).

9.2. Общие индексы количественных показателей

Типичным индексом количественных показателей является *индекс физического объема*. Сложность при построении этого индекса заключается в том, что объемы разных видов продукции в натуральном выражении несоизмеримы (*например, столы и стулья*) и суммироваться не могут.

Единство различных видов продукции состоит в том, что они являются продуктами труда, имеют стоимость и ее денежный соизмеритель – цену (*p*). Каждый продукт характеризуется себестоимостью (*z*) и трудоемкостью (*t*). Эти качественные показатели и могут быть использованы как мера или коэффициент соизмерения разнородных продуктов.

Умножая физический объем продукции (q) на цену, себестоимость либо трудоемкость единицы продукции, получают сравнимые показатели, которые можно суммировать (qp , qz , $qt=T$). Показатели-сомножители, связанные с индексируемыми величинами, принято называть *весами индексов* или *соизмерителями*, а умножение на них – взвешиванием. Соизмеритель, как правило, указывается после индексируемой величины, его значение фиксируется на определенной уровне.

При построении индексов количественных показателей соизмерители, являющиеся качественными показателями, фиксируются на уровне базисного периода.

Построим агрегатный индекс физического объема продукции с учетом рассмотренных принципов, используя в качестве соизмерителя цену продукции в базисном периоде:

$$I_q = \frac{\sum q_1 p_0}{\sum q_0 p_0},$$

где $\sum q_1 p_0$ – условная стоимость продукции отчетного периода, исчисленная в ценах базисного периода;

$\sum q_0 p_0$ – фактическая стоимость продукции, произведенной в базисном периоде.

Индекс физического объема показывает, во сколько раз изменился физический объем продукции в отчетном периоде по сравнению с базисным периодом. Разность числителя и знаменателя индекса дает абсолютный прирост стоимости продукции, обусловленный изменением ее физического объема:

$$\Delta_{pq}^q = \sum q_1 p_0 - \sum q_0 p_0.$$

При построении агрегатного индекса физического объема может быть использована себестоимость или трудоемкость единицы продукции:

$$I_q = \frac{\sum q_1 z_0}{\sum q_0 z_0}; \quad I_q = \frac{\sum q_1 t_0}{\sum q_0 t_0}.$$

Разность числителя и знаменателя индекса себестоимости (трудоемкости) дает абсолютный прирост затрат на производство продукции (затрат труда), обусловленный изменением ее физического объема:

$$\Delta_{zq}^q = \sum q_1 z_0 - \sum q_0 z_0; \quad \Delta_{tq}^q = \sum q_1 t_0 - \sum q_0 t_0.$$

9.3. Общие индексы качественных показателей

В экономике каждый количественный показатель связан с тем или иным качественным показателем и, наоборот, каждый качественный показатель связан с каким-либо количественным показателем. *К примеру, с объемом продукции связаны такие качественные показатели, как цена, себестоимость и трудоемкость единицы продукции.*

Рассмотрим принципы построения агрегатных индексов качественных показателей на примере индекса цены. Индексируемой величиной в этом индексе выступает качественный показатель, а соизмерителем – связанный с ним количественный.

При построении качественных индексов соизмерители, как правило, фиксируются на уровне отчетного периода.

Построим агрегатный индекс цен с учетом рассмотренных принципов, используя в качестве соизмерителя физический объем продукции в отчетном периоде. Этот индекс цен известен в статистике как *индекс Пааше* и записывается следующим образом:

$$I_p = \frac{\sum p_1 q_1}{\sum p_0 q_1},$$

где $\sum p_1 q_1$ – стоимость продукции отчетного периода; $\sum p_0 q_1$ – стоимость продукции отчетного периода в ценах базисного периода.

Общий индекс цен характеризует, во сколько раз возрос в среднем уровень цен на массу продукции, произведенной в отчетном периоде. Разность числителя и знаменателя индекса цен представляет абсолютное изменение стоимости продукции, обусловленное изменением цен на нее:

$$\Delta_{pq}^p = \sum p_1 q_1 - \sum p_0 q_1.$$

По аналогии с индексом цен может быть записан агрегатный индекс себестоимости или трудоемкости:

$$I_z = \frac{\sum z_1 q_1}{\sum z_0 q_1}; \quad I_t = \frac{\sum t_1 q_1}{\sum t_0 q_1}.$$

Разность числителя и знаменателя индекса себестоимости представляет абсолютное изменение затрат, обусловленное изменением себестоимости:

$$\Delta_{zq}^z = \sum z_1 q_1 - \sum z_0 q_1$$

Отметим, что в статистике (*например, в отечественной статистике для исчисления индекса потребительских цен*) помимо индекса Пааше применяется *агрегатный индекс цен Ласпейреса с весами базисного периода*:

$$I_p^l = \frac{\sum p_1 q_0}{\sum p_0 q_0}.$$

Он показывает, на сколько бы изменились цены в отчетном периоде по сравнению с базисным по той продукции, которая была произведена в базисном периоде. Однако индекс Ласпейреса несет в себе тенденцию к завышению инфляции, тогда как индекс Пааше, наоборот, имеет тенденцию к ее занижению. Поэтому в ряде случаев используется индекс Фишера, определяемый как среднее геометрическое из индексов цен Пааше и Ласпейреса:

$$I_p^{\phi} = \sqrt{\frac{\sum p_1 q_1 \cdot \sum p_1 q_0}{\sum p_0 q_1 \cdot \sum p_0 q_0}}.$$

9.4. Двухфакторные системы взаимосвязанных индексов

Связь между экономическими показателями находит отражение во взаимосвязи их индексов, т. е. если $z = ux$, то $I_z = I_y I_x$. Индексы экономических показателей, связанных между собой, образуют *индексные системы*, которые дают возможность проведения факторного анализа.

Двухфакторные системы индексов предполагают наличие факторных индексов качественного и количественного показателей. Так, индексы цен и физического объема являются факторными по отношению к индексу стоимости продукции:

$$I_{pq} = I_p I_q \quad \text{или} \quad \frac{\sum p_1 q_1}{\sum p_0 q_0} = \frac{\sum p_1 q_1}{\sum p_0 q_1} \cdot \frac{\sum p_0 q_1}{\sum p_0 q_0}.$$

Разность числителя и знаменателя каждого из индексов позволяет определить абсолютное изменение стоимости показателя, в том числе обусловленное влиянием факторов (цен и физического объема продукции):

$$\begin{aligned} \Delta_{pq} &= \sum p_1 q_1 - \sum p_0 q_0; \\ \Delta_{pq}^p &= \sum p_1 q_1 - \sum p_0 q_1, \quad \Delta_{pq}^q = \sum q_1 p_0 - \sum q_0 p_0, \\ \Delta_{pq} &= \Delta_{pq}^p + \Delta_{pq}^q. \end{aligned}$$

По аналогии с индексом стоимости продукции можно представить индекс затрат на производство (I_{zq}) через взаимосвязь индексов себестоимости единицы продукции (I_z) и физического объема продукции (I_q):

$$I_{zq} = I_z I_q, \quad \text{или} \quad \frac{\sum z_1 q_1}{\sum z_0 q_0} = \frac{\sum z_1 q_1}{\sum z_0 q_1} \cdot \frac{\sum q_1 z_0}{\sum q_0 z_0}.$$

Разность числителя и знаменателя каждого из индексов позволяет определить абсолютное изменение затрат на производство продукции, в том числе обусловленное влиянием факторов (себестоимости единицы продукции и физического объема продукции):

$$\begin{aligned} \Delta_{zq} &= \sum z_1 q_1 - \sum z_0 q_0; \\ \Delta_{zq}^z &= \sum z_1 q_1 - \sum z_0 q_1, \quad \Delta_{zq}^q = \sum q_1 z_0 - \sum q_0 z_0, \\ \Delta_{zq} &= \Delta_{zq}^z + \Delta_{zq}^q. \end{aligned}$$

9.5. Индексный метод анализа динамики среднего уровня

Экономические явления часто характеризуются с помощью средних величин. Качественные показатели зачастую выражаются в виде средних величин: средняя цена единицы продукции (\bar{p}), средняя себестоимость

единицы продукции (\bar{z}), средняя заработная плата одного работника (\bar{z}), средняя выработка одного работника (\bar{w}), средняя трудоемкость единицы продукции (\bar{t}) и т. п. Для изучения их динамики в статистике применяются индексы средних показателей, в частности, индексы переменного состава, постоянного состава, структурных сдвигов. Рассмотрим их использование для изучения динамики средней трудоемкости единицы продукции.

Индексом переменного состава называется индекс, выражающий соотношение средних уровней изучаемого явления, относящихся к разным периодам времени.

Его экономический смысл: отражает изменение не только индексной величины, но и структуры совокупности, т.е. весов.

Индекс трудоемкости переменного состава определяется по формуле:

$$I_t = \frac{\bar{t}_1}{\bar{t}_0} = \frac{\sum t_1 q_1}{\sum q_1} \cdot \frac{\sum t_0 q_0}{\sum q_0},$$

где t_0 и t_1 – трудоемкость единицы продукции в базисном и отчетном периоде; q_0 и q_1 – количество единиц продукции в базисном и отчетном периоде.

Этот индекс характеризует изменение средней трудоемкости в целом за счет двух факторов: 1) изменения осредняемых значений индексируемой величины (трудоемкости); 2) влияния структурных сдвигов, т. е. изменение удельных весов единиц совокупности с различными значениями индексируемого показателя (трудоемкости). Поэтому индекс переменного состава можно разложить на два индекса-сомножителя, каждый из которых отражает влияние только одного из факторов, определяющих средний уровень.

Индекс постоянного состава (или фиксированного состава) — это индекс исчисленный с весами, зафиксированными на уровне одного какого-либо периода и показывающий изменение только индексируемой величины (т.е. изменение структуры не показывает).

Индекс-сомножитель, отражающий влияние первого фактора, называется *индексом трудоемкости постоянного состава*:

$$I_t = \frac{\bar{t}_1}{\bar{t}_0} = \frac{\sum t_1 q_1}{\sum q_1} \cdot \frac{\sum t_0 q_1}{\sum q_1}.$$

Второй индекс-сомножитель называется *индексом структурных сдвигов*.

Индекс структурных сдвигов характеризует влияние изменения структуры явления на динамику фиксированного уровня этого явления:

$$I_{стр} = \frac{\bar{t}'_0}{\bar{t}_0} = \frac{\sum t_0 q_1}{\sum q_1} \cdot \frac{\sum t_0 q_0}{\sum q_0}.$$

Между индексами существует взаимосвязь:

$$I_t = I_t \cdot I_{стр}.$$

9.6. Цепные и базисные индексы

Часто в ходе экономического анализа изменение индексируемых величин изучают не за два, а за ряд последовательных периодов. Следовательно, возникает необходимость построения индексов за ряд этих последовательных периодов, которые образуют индексные системы. Такие системы характеризуют изменения, происходящие в изучаемом явлении в течение исследуемого периода времени. В зависимости от базы сравнения индексы бывают базисные и цепные.

Базисные индексы – это такие, у которых в качестве базисного значения (показателя в знаменателе) берется какой-либо фиксированный период. Последовательность базисных индексов показывает динамику показателя относительно этого периода (обычно к началу месяца, началу года). Если взять кратность периода по месяцам, то это будет февраль к январю, март к январю, апрель к январю и т.д.

Цепные индексы – в качестве базисного значения (в знаменателе) выступает предыдущий период (не фиксируется, а изменяется в зависимости от анализируемого периода). При помесечных данных это будет: февраль к январю, март к февралю, апрель к марту и т.д.

Оба вида индексов взаимосвязаны между собой и, имея базисные, можно легко перейти к цепным индексам, и наоборот. Также легко из цепных и базисных индексов получить агрегированный общий индекс. В то же время агрегированный индекс без дополнительной информации невозможно разложить на индексы более коротких промежутков времени. Можно только рассчитать средний индекс. Здесь аналогия со средней величиной: из исходных данных можно посчитать общую сумму, но из общей суммы можно посчитать только среднее значение, но не каждое в отдельности. Оба вида индексных рядов отражают одну и ту же динамику, только немного под разным углом. В зависимости от цели выбирают тот и/или иной вид индексов.

Между цепными и базисными индексами существует взаимосвязь, которая для индивидуальных индексов проявляется всегда. А при агрегатных индексах лишь при определенных условиях.

Произведение последовательных цепных индивидуальных индексов дает базисный индекс последнего периода. Отношение базисного индекса отчетного периода к базисному индексу предшествующего периода дает цепной индекс отчетного периода. Это правило позволяет применять так называемый цепной метод, то есть находить неизвестный ряд базисных индексов по известным цепным индексам и наоборот.

В базисных агрегатных индексах все отчетные данные сопоставляются только с базисными (закрепленными) данными, а цепных – с предыдущими (в данном случае - смежными) данными.

Цепные и базисные индексы физического объема продукции представлены в таблице 3.

Таблица 3 - Базисные и цепные индексы физического объема продукции

Индекс физического объема продукции	
Базисные	Цепные
$J_1 = \frac{\sum q_1 p_0}{\sum q_0 p_0}$	$J_1 = \frac{\sum q_1 p_0}{\sum q_0 p_0}$
$J_2 = \frac{\sum q_2 p_0}{\sum q_0 p_0}$	$J_2 = \frac{\sum q_2 p_0}{\sum q_1 p_0}$
$J_3 = \frac{\sum q_3 p_0}{\sum q_0 p_0}$	$J_3 = \frac{\sum q_3 p_0}{\sum q_2 p_0}$
.....
$J_n = \frac{\sum q_n p_0}{\sum q_0 p_0}$	$J_n = \frac{\sum q_n p_0}{\sum q_{n-1} p_0}$
Индекс цен	
Базисные	Цепные
$J_1 = \frac{\sum q_1 p_1}{\sum q_1 p_0}$	$J_1 = \frac{\sum q_1 p_1}{\sum q_1 p_0}$
$J_2 = \frac{\sum q_2 p_2}{\sum q_2 p_0}$	$J_2 = \frac{\sum q_2 p_2}{\sum q_2 p_1}$
$J_3 = \frac{\sum q_3 p_3}{\sum q_3 p_0}$	$J_3 = \frac{\sum q_3 p_3}{\sum q_3 p_2}$
.....
$J_n = \frac{\sum q_n p_n}{\sum q_n p_0}$	$J_n = \frac{\sum q_n p_n}{\sum q_n p_{n-1}}$

Контрольные вопросы

1. Что называется индексом в статистике?
2. Какие задачи решаются при помощи индексов?
3. Что характеризуют индивидуальные индексы?
4. В чем сущность общих индексов?
5. Какие индексы относятся к индексам количественных показателей?
6. Какие индексы относятся к индексам качественных показателей?
7. Как исчисляются агрегатные индексы цен (Пааше и Лайспейреса)?
8. Что называется индексом переменного состава и как он исчисляется?

9. Какой индекс называется индексом постоянного состава, как он исчисляется?
10. Что характеризует индекс структурных сдвигов и как он исчисляется?
11. Какая взаимосвязь существует между индексами переменного, постоянного состава и структурных сдвигов?
12. Что представляют собой индексы с постоянными и переменными весами?

Тема 10. Статистическое изучение связи социально-экономических явлений

10.1. Виды и формы взаимосвязей, изучаемых в статистике

Исследуя состояние и развитие как природных, так и общественных явлений, необходимо изучать взаимосвязи наблюдаемых процессов и явлений. Современная наука об обществе объясняет суть явлений через изучение их взаимосвязи. При этом полнота описания, так или иначе, определяется количественными характеристиками причинно-следственных связей между ними. Например, объем валютных торгов зависит от спроса на валюту, который в свою очередь определяется состоянием экономики, активностью внешнеэкономической деятельности субъектов, политикой государства и др., объем продукции предприятия связан с численностью и квалификацией работников, стоимостью основных фондов, технологией производства и т.д.

Оценка наиболее существенных факторов, а также изучение воздействия одних факторов на другие является одной из основных задач статистики. Знание характера и силы связей позволяет управлять социально-экономическими процессами и предсказывать их развитие.

Среди многих форм связей важнейшей является *причинная*, определяющая все другие формы. Сущность причинности состоит в порождении одного явления другим. Вместе с тем, причина сама по себе еще не определяет следствия, она зависит также от условий, в которых протекает действие причины. Для возникновения следствия нужны все определяющие факторы – причины и условия. Необходимая обусловленность явлений множеством факторов называется *детерминизмом*. Объектами исследования при статистическом измерении связей служат, как правило, детерминированность следствия факторами (причиной и условиями).

Связи между признаками и явлениями ввиду их большого разнообразия, классифицируются по ряду оснований.

Признаки по их сущности и значению для изучения взаимосвязей делятся 2 класса:

1. Признаки, обуславливающие изменения других, связанных с ними признаков (характеризующие причины), называются *факторными*, или просто факторами.

2. Признаки, изменяющиеся под действием факторных признаков (характеризующие следствие), называются *результативными*.

Выявление связей между признаками основывается на результатах качественного теоретического анализа. Задача статистики – количественная оценка закономерности связей, *математическая определенность* позволяет использовать результаты экономических разработок для практических целей.

Вместе с тем, качественный анализ должен не только предшествовать статистическому, но и являться подтверждением справедливости его результатов.

Связи между явлениями и их признаками классифицируют *по степени тесноты связи, направлению и аналитическому выравниванию*.

Формы проявления факторных связей весьма разнообразны.

В качестве двух самых общих их видов выделяют *функциональную* или *жестко детерминированную* и *статистическую* или *стохастически детерминированную* связи.

Функциональная связь (жестко детерминированная) – это вид причинной зависимости, при которой определенному значению факторного признака строго соответствует одно или несколько значений результативного признака.

Связь признака y с признаком x называется функциональной, если каждому возможному значению независимого признака x соответствует одно или несколько строго определенных значений зависимого признака y . определение функциональной связи может быть легко обобщено для случая многих признаков $x_1, x_2, x_3, \dots, x_n$.

Характерной особенностью функциональных связей является то, что в каждом отдельном случае известен полный перечень факторов, определяющих значение зависимого (результативного) признака, а также точный механизм их влияния, выраженный определенным уравнением.

Функциональную связь можно представить уравнением:

$$y_i = f(x_i),$$

где y_i – результативный признак ($i = 1, \dots, n$);

$f(x_i)$ – известная функция связи результативного и факторного признаков;

x_i – факторный признак.

Чаще всего функциональные связи проявляется в физике, химии, математике и других точных науках. Имеют место функциональные связи и в социально-экономических процессах, но довольно редко (они отражают взаимосвязь отдельных сторон сложных явлений общественной жизни). В экономике примером может служить прямо пропорциональная зависимость между производительностью труда и увеличением производства продукции, а также связь между оплатой труда y и количеством изготовленных деталей x при простой сдельной оплате труда. Так, если расценка за одну деталь составляет 3 руб., то связь между признаками однозначно выразится простым линейным уравнением: $y = 3x$. Для каждого допустимого значения x можно указать определенное значение y . Например, $x = 5$, то соответственно $y = 15$.

Характерной особенностью функциональной связи является то, что она проявляется с одинаковой силой у каждой единицы изучаемой совокупности. Поэтому знание функциональной зависимости позволяет абсолютно точно прогнозировать события, например, наступление солнечных или лунных затмений.

В реальной общественной жизни, ввиду неполноты информации жестко детерминированной системы, может возникнуть неопределенность, из-за которой эта система по своей природе должна рассматриваться как *вероятностная*, при этом связь между признаками становится *стохастической*.

Если причинная зависимость проявляется не в каждом отдельном случае, а в общем, среднем, при большом числе наблюдений, то такая зависимость называется *стохастической* (неполной, статистической).

Термин «стохастический» происходит от греческого «stochos» - мишень или бычий глаз. Стреляя в мишень, даже хороший стрелок редко попадает в центр – «яблочко», выстрелы ложатся в некоторой, близкой к нему области. В этом смысле стохастическая связь означает приблизительный характер значений признака. Обычно стохастическая связь между двумя случайными величинами имеет место в случае наличия общих случайных факторов, воздействующих на каждую из них.

Например, зависимость цены товара от его качества. В отдельных случаях соотношение спроса и предложения может привести к тому, что товар худшего качества будет продан по более высокой цене, а аналогичный товар лучшего качества может иметь более низкую цену при достаточно большом объеме продаж. Однако же в среднем между ценой и качеством товара существует прямая зависимость – чем выше качество товара, тем выше его цена и, соответственно, чем ниже качество товара, тем ниже будет и цена на него.

Стохастическая связь – это связь между величинами, при которой одна из них, случайная величина y , реагирует на изменение другой величины x или других величин $x_1, x_2, x_3, \dots, x_n$ (случайных или неслучайных) изменением закона распределения. Это обуславливается тем, что зависимая переменная (результативный признак), кроме рассматриваемых независимых, подвержена влиянию ряда неучтенных или неконтролируемых (случайных) факторов, а также некоторых неизбежных ошибок измерения переменных. Поскольку значения зависимой переменной подвержены случайному разбросу, они не могут быть предсказаны с достаточной точностью, а только указаны с определенной *вероятностью*.

Модель стохастической связи может быть представлена в общем виде *уравнением*:

$$y_i = f(x_i) + e_i,$$

где y_i – расчетное значение результативного признака;

$f(x_i)$ – часть результативного признака, сформировавшаяся под воздействием учтенных известных факторных признаков (одного или множества), находящихся в стохастической связи с признаком;

e_i – часть результативного признака, возникшая вследствие действия неконтролируемых или неучтенных факторов, а также измерения признаков неизбежно сопровождающегося некоторыми случайными ошибками.

Появление стохастических связей подвержено действию закона больших чисел: лишь в достаточно большом числе единиц индивидуальные особенности сгладятся, случайности взаимопогасятся и зависимость, если она имеет существенную силу, проявится достаточно отчетливо.

При изучении массовых общественно-экономических явлений между факторными признаками проявляется корреляционная связь, т.е. такая связь, при которой на величину результативного признака одновременно или последовательно оказывают влияние, помимо факторного, множество других признаков, действующих в различных направлениях.

Корреляционная связь (от английского слова correlation – соотношение, соответствие) – частный случай стохастической связи, когда заданным значениям зависимой переменной соответствует некоторый ряд вероятных значений независимой переменной. Это обуславливается тем, что изменение результативного признака происходит под влиянием рассматриваемого факторного признака не всецело, а лишь частично, так как возможно влияние ряда неучтенных или неконтролируемых (случайных) факторов. Объяснением тому является сложность взаимосвязей между анализируемыми факторами, поэтому связь между признаками проявляется лишь в среднем, в массе случаев, т.е. с изменением факторного признака (x) закономерным образом изменяется среднее значение результативного признака (y), в то время как в каждом отдельном случае результативный признак может принимать множество различных значений.

Например, в сельском хозяйстве это может быть связь между урожайностью и количеством внесенных удобрений. Очевидно, что последние участвуют в формировании урожая, но для каждого поля одно и то же количество внесенных удобрений вызовет разный прирост урожайности, так как во взаимодействии находится еще целый ряд факторов (погода, состояние почвы, количество осадков и др.), которые и формируют конечный результат. Однако в среднем такая связь наблюдается, т.е. увеличение массы внесенных удобрений ведет к росту урожайности.

По направлению связи бывают *прямыми*, когда зависимая переменная растет с увеличением факторного признака, и *обратными*, при которых рост последнего сопровождается уменьшением функции. Такие связи также можно назвать, соответственно, положительными и отрицательными.

Относительно своей аналитической формы связи бывают *линейными* (*прямолинейными*) и *нелинейными* (*криволинейными*).

В первом случае между признаками, в среднем, проявляются линейные соотношения (с возрастанием факторного признака происходит непрерывное возрастание (или убывание) значений результативного признака). Математически такая связь представляется уравнением прямой, а графически – прямой линией. Отсюда ее короткое название – *линейная связь*.

Нелинейная взаимосвязь выражается нелинейной функцией, а переменные связаны между собой в среднем нелинейно. Здесь с возрастанием значения факторного признака возрастание (или убывание)

результативного признака происходит неравномерно или же направление его изменения меняется на обратное. Геометрически также связи представляются кривыми линиями (гиперболой, параболой и т.д.).

Существует еще одна достаточно важная характеристика связей с точки зрения взаимодействующих факторов. Если характеризуется связь двух признаков, то ее принято называть *парной (однофакторной)* (так как рассматривается пара признаков. Например, корреляционная связь между прибылью и производительностью труда).

Если изучаются более чем две переменные – *множественной (многофакторной)* (имеется в виду, что все факторы действуют комплексно, т.е. одновременно и во взаимосвязи, например, корреляционная связь между производительностью труда и уровнем организации труда, автоматизации производства, квалификации рабочих, производственным стажем, простоями и другими факторными признаками). С помощью множественной корреляции можно охватить весь комплекс факторных признаков и объективно отразить существующие множественные связи.

Указанные выше классификационные признаки наиболее часто встречаются в статистическом анализе, но кроме перечисленных, различают также *непосредственные, косвенные и ложные* связи. Собственно, суть каждой очевидна из её названия. В первом случае факторы взаимодействуют между собой непосредственно. Для косвенной связи характерно участие какой-то третьей переменной, которая опосредует связь между изучаемыми признаками. Ложная связь – это связь, установленная формально и, как правило, подтвержденная только количественными оценками. Она не имеет под собой качественной основы или же бессмысленна.

Здесь можно привести множество примеров. Известен классический пример, приведенный ведущим статистиком начала XX в. А.А. Чупровым. Если в качестве признака x взять число пожарных команд в городе, а за признак y – сумму убытков в городе от пожаров, то между признаками x и y в совокупности городов России возникнет существенная прямая корреляция. В среднем, чем больше пожарных в городе, тем больше убытков от пожаров. В чем же дело? Данную корреляцию нельзя интерпретировать как связь причины и следствия, оба признака – следствия общей причины – размера города. В крупных городах больше пожарных частей, но больше и пожаров, и убытков от них за год, чем в мелких городах.

По силе различаются *слабые* и *сильные* связи. Эта формальная характеристика выражается конкретными величинами и интерпретируется в соответствии с общепринятыми критериями силы связи для конкретных показателей.

Связи между явлениями и их признаками классифицируются по степени тесноты. По степени тесноты связи различают следующие количественные критерии оценки тесноты связи, представленные в таблице 4.

Таблица 4 – Количественные критерии оценки тесноты связи

Величина показателя связи	Характер связи
до $\pm 0,3$	практически отсутствует
$\pm 0,3 - \pm 0,5$	слабая
$\pm 0,5 - \pm 0,7$	умеренная
$\pm 0,7 - \pm 1,0$	сильная

10.2. Задачи статистического изучения связи

В наиболее общем виде задача статистики в области изучения взаимосвязей состоит в количественной оценке их наличия и направления, а также характеристике силы и формы влияния одних факторов на другие.

Изучение корреляционных связей сводится в основном к решению следующих задач:

- выявление наличия (или отсутствия) корреляционной связи между изучаемыми признаками;
- измерение степени тесноты связи между признаками;
- нахождение аналитического выражения связи, отражающей зависимость между результативным и факторным признаками;
- экономическая интерпретация и практическое использование полученного результата.

В начальной стадии анализа статистических зависимостей применяются простейшие методы наличия связи, ее направления и характера, выявляется форма воздействия одних признаков на другие.

10.3. Определение формы связи

Для выявления наличия и характера корреляционной связи между двумя признаками в статистике используется ряд методов:

1. Метод группировок

При составлении структурных группировок на основе варьирующих количественных признаков необходимо определить количество групп и интервалы группировки.

Интервал - количественное значение, отделяющее одну единицу (группу) от другой, т. е. интервал очерчивает количественные границы групп.

Как правило, величина интервала представляет собой разность между максимальным и минимальным значениями признака в каждой группе.

Вопрос о числе групп и величине интервала следует решать с учетом множества обстоятельств, прежде всего исходя из целей исследования, значения изучаемого признака и т.д.

Количество групп и величина интервала связаны между собой: чем больше образовано групп, тем меньше интервал, и наоборот. Количество групп зависит от числа единиц исследуемого объекта и степени колеблемости группировочного признака. При небольшом объеме совокупности нельзя образовывать большое число групп, так как группы будут малочисленными.

При определении количества групп необходимо стремиться к тому, чтобы были учтены особенности изучаемого явления. Поэтому число групп должно быть оптимальным, в каждую группу должно входить достаточно большое число единиц совокупности, что отвечает требованию закона больших чисел. Однако в отдельных случаях представляют интерес и малочисленные группы: новое, передовое, пока оно не станет массовым, проявляется в незначительном числе фактов; поэтому задача статистики - выделить эти факты, изучить их.

2. *Метод сравнения параллельных рядов.*

Суть данного метода заключается в том, что полученные данные в результате сводки и обработки располагаются в ранжированный ряд по факторному признаку, а результативный записывается соответственно своего значениям. Полученные ряды сравниваются.

Единицы наблюдения необходимо расположить по возрастанию значений факторного признака x и затем сравнить с ним (визуально) поведение результативного признака y .

3. *Графический метод* – это графическое изображение корреляционной зависимости. Для этого, имея n взаимосвязанных пар значений x и y и пользуясь прямоугольной системой координат, каждую такую пару изображают в виде точки на плоскости с координатами x и y .

Беспорядочное расположение говорит об отсутствии связи. Наоборот, чем сильнее связь, тем теснее точки группируются вокруг определенной линии.

Совокупность полученных точек представляет собой *корреляционное поле*, а соединяя последовательно нанесенные точки отрезками, получают ломаную линию, именуемую *эмпирической линией регрессии*.

Визуально анализируя график, можно предположить характер зависимости между признаками x и y .

4. *Корреляционный метод.*

Группировка исходной информации оформляется в виде корреляционной решетки.

5. *Балансовый метод.* Как правило, он позволяет охарактеризовать зависимость между источниками формирования ресурсов (средств) и их использование. Статистический баланс представляет собой систему показателей, которая состоит из 2-х сумм абсолютных величин, связанных между собой законом равенства. Эту сумму можно представить следующим равенством: остаток на начало + приход = расход + остаток на конец.

Балансовый метод дает возможность не только анализировать показатели во взаимосвязи, но и осуществлять взаимный контроль данных и рассчитывать недостающие показатели. Пример, продано в розницу = остаток на начало + приход – продано оптом – остаток на конец.

Всякий баланс состоит из 2 балансирующихся частей — приходной и расходной. В приходной части учитываются все ресурсы (запасы на начало года, производство, импорт и др.), а в расходной части — все виды использования (на производственные нужды, непроизводства, потребление, экспорт, создание резервов и др.) и запасы на конец года. В практике советской статистики разрабатываются различные виды балансов, они могут быть разделены на 4 группы. В 1-ю группу входят балансы, характеризующие ресурсы отдельных продуктов, их распределение и использование в народном хозяйстве. Составляются, как правило, в натуральном выражении. К ним относятся балансы, составляемые по отдельным видам чёрных и цветных металлов, топлива, химикатов, строительных и лесных материалов, по отдельным видам машин и оборудования. Они могут также составляться в сводном виде по группе продуктов (например, топливно-энергетический баланс). Ко 2-й группе относятся различные виды балансов труда — баланс трудовых ресурсов, баланс наличия и использования рабочей силы. К 3-й группе относятся балансы, характеризующие образование финансовых (денежных) ресурсов и их использование на различные цели в народном хозяйстве и в отдельных отраслях — сводный баланс финансовых ресурсов, балансы доходов и расходов отдельных отраслей, государственный бюджет, кредитный план Госбанка, баланс денежных доходов и расходов населения, кассовый план Госбанка и т. д. К 4-й группе относится баланс народного хозяйства, представляющий собой систему наиболее общих синтетических балансов, характеризующих в целом весь процесс общественного воспроизводства: сводный материальный баланс общественного продукта и национального дохода, баланс производства, распределения и перераспределения общественного продукта и национального дохода (финансовый баланс), баланс основных фондов, баланс труда и другие синтетические балансы. В балансе народного хозяйства балансовый метод находит своё наивысшее развитие.

Большое внимание уделяется разработке межотраслевого баланса производства и распределения общественного продукта и его использованию для анализа социалистического воспроизводства и планирования народного хозяйства. Этот баланс является частью баланса народного хозяйства и призван дать подробную характеристику межотраслевых связей в народном хозяйстве с выделением большого числа отраслей и продуктов.

10.4. Измерение тесноты связи между признаками

Для производства корреляционно-регрессионного анализа, для решения задач по измерению тесноты связи между варьирующими признаками и оценки факторов, оказывающие наибольшее влияние на результативный показатель, проводят следующие этапы:

- 1) предварительный анализ
- 2) сбор информации и ее первичная обработка
- 3) построение модели – уравнения регрессии.

Для оценки тесноты связи применяется ряд показателей и приёмов:

К простейшим показателям степени тесноты связи относится *коэффициент корреляции знаков* (коэффициент Фехнера). Этот показатель основан на оценке степени согласованности направлений отклонений индивидуальных значений факторного и результативного признаков от соответствующих средних. Для его расчета вычисляют среднее значение результативного и факторного признаков:

$$\bar{x} = \frac{\sum x}{n}; \quad \bar{y} = \frac{\sum y}{n},$$

где n – количество значений признаков.

Затем определяют знаки отклонений для всех взаимосвязанных пар признаков.

Во внимание принимаются не величины отклонений ($x_i - \bar{x}$) и ($y_i - \bar{y}$), а их знаки («+» или «-»). Определив знаки отклонений от средней величины в каждом ряду, рассматривают все пары знаков и подсчитывают число их совпадений (C) и несовпадений (H).

Коэффициент Фехнера определяется следующим образом:

$$K_{\Phi} = \frac{\sum C - \sum H}{\sum C + \sum H},$$

где C – число совпадений знаков отклонений индивидуальных значений от средней (согласованная вариация);

H – число несовпадений знаков отклонений индивидуальных значений от средней (несогласованная вариация).

Коэффициент Фехнера может принимать значения в пределах от -1 до +1. Положительное значение данного коэффициента позволяет сделать вывод о возможном наличии прямой связи, отрицательное – о возможном наличии обратной связи. Так как величина этого показателя не зависит от величины отклонений факторного и результативного признаков от соответствующих средних, то говорить о степени тесноты корреляционной связи и ее существенности на основании только коэффициента Фехнера нельзя.

Коэффициенты корреляции рангов – это менее точные, но более простые по расчету непараметрические показатели для измерения тесноты связи между двумя коррелируемыми признаками. К ним относятся

коэффициенты Спирмэна (ρ) и Кендэла (τ), основанные на корреляции не самих значений коррелируемых признаков, а их *рангов* – порядковых номеров, присваиваемых каждому индивидуальному значению x и y (отдельно) в ранжированном ряду. Оба признака необходимо ранжировать (нумеровать) в одном и том же порядке: от меньших значений к большим и наоборот. Если встречается несколько значений x (или y), то каждому из них присваивается ранг, равный частному от деления суммы рангов (мест в ряду), приходящихся на эти значения, на число равных значений. Ранги признаков x и y обозначают символами R_x и R_y (иногда N_x и N_y). Суждение о связи между изменениями значений x и y основано на сравнении поведения рангов по двум признакам параллельно. Если у каждой пары x и y ранги совпадают, это характеризует максимально тесную связь. Если же наблюдается полная противоположность рангов, т.е. в одном ряду ранги возрастают от 1 до n , а в другом – убывают от n до 1, это максимально возможная обратная связь. Подходы для оценки тесноты связи у Спирмэна и Кендэла несколько различаются. Для расчета *коэффициента Спирмэна* значения признаков x и y нумеруют (отдельно) в порядке возрастания от 1 до n , т.е. им присваивают определенный ранг (R_x и R_y) – порядковый номер в ранжированном ряду. Затем для каждой пары рангов находят их разность (обозначается как $d = R_x - R_y$), и квадраты этой разности суммируют.

$$\rho = 1 - \frac{6 \sum d^2}{n(n^2 - 1)} = 1 - \frac{6 \sum d^2}{n^3 - n},$$

где d – разность рангов x и y ;

n – число наблюдаемых пар значений x и y .

Коэффициент ρ может принимать значения от 0 до ± 1 . Следует иметь в виду, что, поскольку коэффициент Спирмэна учитывает разность только рангов, а не самих значений x и y , он менее точен по сравнению с линейным коэффициентом. Поэтому его крайние значения (1 или 0) нельзя безоговорочно расценивать как свидетельство функциональной связи или полного отсутствия зависимости между x и y . Во всех других случаях, т.е. когда ρ не принимает крайних значений, он довольно близок к r .

Формула применима строго теоретически только тогда, когда отдельные значения x (и y), а следовательно, и их ранги не повторяются. Для случая повторяющихся (связанных) рангов есть другая, более сложная формула, скорректированная на число повторяющихся рангов.

Для оценки степени тесноты связи между несколькими признаками, если используется коэффициент корреляции рангов, рассчитывается *коэффициент конкордации* (*множественный коэффициент ранговой корреляции*):

$$W = \frac{12S}{m^2(n^3 - n)}, \text{ где}$$

S — сумма квадратов отклонений суммы m рангов от их средней величины;

m — число ранжируемых признаков;

n — число ранжируемых единиц (число наблюдений).

Более совершенным показателем степени тесноты связи является *линейный коэффициент корреляции*. При расчете этого показателя учитываются не только знаки отклонений индивидуальных значений признаков от средней, но и сама величина таких отклонений. Формула для расчета линейного коэффициента корреляции (r) выглядит следующим образом:

$$r = \frac{\sum(x - \bar{x})(y - \bar{y})}{\sqrt{\sum(x - \bar{x})^2 \sum(y - \bar{y})^2}}.$$

При этом, если линейный коэффициент корреляции лежит в пределах $0 < r < 0,5$, то это говорит о прямой слабой связи между показательными. Если он лежит в пределах $> 0,5$, то это говорит о прямой довольно тесной связи. Если $-1 < r < 0$, - это свидетельствует об обратной связи. Если числовое значение данного коэффициента не подходит ни под один из вышеуказанных интервалов, то это говорит об отсутствии связей.

При наличии соотношения между вариацией качественных признаков говорят об их ассоциации, взаимосвязанности. Для определения тесноты связи двух качественных признаков, каждый из которых состоит только из двух групп, применяются коэффициенты ассоциации и контингенции. Для их вычисления строится таблица, которая показывает связь между двумя явлениями, каждое из которых должно быть альтернативным, т.е. состоящим из двух качественно отличных друг от друга значений признака (например, изделие качественное или бракованное).

К. Пирсон предложил показатель, названный *коэффициентом ассоциации*, который вычисляется по формуле:

$$A = \frac{ad - bc}{ad + bc}$$

Свойства коэффициента ассоциации такие же, как и у коэффициента корреляции.

Другой метод измерения связи предложен английскими статистиками Эдни Дж. Юлом (1871-1951) и Морисом Дж. Кендэлом (1907). Числитель этого коэффициента, называемого *коэффициентом контингенции*, совпадает с числителем коэффициента ассоциации Пирсона:

$$\hat{A} = \frac{ad - bc}{\sqrt{(a + b)(c + d)(a + c)(b + d)}} \quad (44)$$

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

a, b, c, d - значения признаков в клетках матрицы сопряженности альтернативных признаков.

Различия:

- В отличие от коэффициента ассоциации он учитывает двустороннюю взаимосвязь альтернативных признаков, т.е. производит измерение более чутко и корректно.

- Для одной и той совокупности численных показателей значение коэффициента сопряженности меньше, чем значение коэффициента ассоциации.

10.5. Выявление влияния отдельных факторов

Изучение влияния отдельных факторов на результативный показатель основывается на уравнении регрессии. При этом опираются на не сгруппированные данные.

Расчет имеет следующие этапы:

1. установить наличие связи между признаками.
2. нанести исходные данные на график и проследить зависимость.

При этом линейная зависимость может быть выражена уравнением:

$\bar{y}_x = ax + b$, где x – значение факторного признака, a – коэффициент регрессии, характеризующий изменение \bar{y} при изменении x на 1, b – значение \bar{y} при $x = 0$, т.е. \bar{y} за счет влияния других факторов кроме x .

При расчетах коэффициентов регрессии используют, как правило, *метод наименьших квадратов*. Он заключается в решении системы уравнений.

Система уравнений:

$$\begin{aligned} a \sum x + n \cdot b &= \sum y \\ a \sum x^2 + b \sum x &= \sum xy \end{aligned}$$

$$a = \frac{\sum Y - b \times \sum X}{n}$$

$$b = \frac{(\sum YX \times n - \sum Y \times \sum X)}{(n \times \sum X^2 - \sum X \times \sum X)}$$

Определить тип уравнения регрессии можно, исследуя зависимость графически на основе группировки. Так, если результативный и факторный признаки возрастают примерно одинаково, то это свидетельствует о том, что связь между ними линейная; если же один признак увеличивается, а другой неравномерно уменьшается — связь гиперболическая. Если с увеличением значений фактора результативный признак сначала растет, а потом снижается, то связь параболическая.

Линейная связь выражается уравнением $y = a + bx$.

Гиперболическая $\tilde{y} = a + \frac{b}{x}$.

Параболическая $\tilde{y} = a + bx + cx^2$.

Показательная $\tilde{y} = ab^x$.

Контрольные вопросы

1. В чем состоит отличие между функциональной и стохастической связью?
2. Что собой представляет корреляционная связь?
3. Перечислите основные методы обнаружения взаимосвязей между явлениями.
4. Как построить корреляционную таблицу?
5. Как рассчитывается коэффициент Фехнера?
6. Как рассчитывается коэффициент ранговой корреляции Спирмена?

Список использованных источников

1. Годин А. М. Статистика: учебник. – Москва: Дашков и К^о, 2016. – 451 с.
2. Гореева Н. М. Статистика в схемах и таблицах. Москва: Эксмо, 2017. – 414 с.
3. Гусаров В. М. Статистика: Учебное пособие для вузов.- М.: ЮНИТИ-ДАНА, 2003. – 463 с.
4. Долгова В. Н., Медведева Т. Ю. Статистика: Учебник и практикум для СПО. – М.: Юрайт, 2019. – 246 с.
5. Долгова В. Н., Медведева Т. Ю. Теория статистики: Учебник и практикум для академического бакалавриата. – М.: Юрайт, 2019. – 234 с.
6. Едроновва В. Н. Общая теория статистики. – М.: ЮРИСЬ, 2017. – 511 с.
7. Елисеева И. И. Статистика: учебник для бакалавров. Москва: Юрайт: ИД Брайт, 2016. – 565 с.
8. Зинченко А. П. Статистика: учебник. Москва: КолосС, 2016. – 566 с.
- Карпенко Л. И. Теория статистики: учебное пособие. – Минск: БГЭУ, 2013. – 591 с.
9. Лобан И. И. Статистика. Общая теория статистики. – Минск: БГСХА, 2019
10. Малых Н. и. Статистика в 2-х томах. Том 2. Социально-экономическая статистика: Учебник и практикум для академического бакалавриата. М.: Юрайт, 2017. – 474 с.
11. Нарбут В. В., Салин В. Н., Шпаковская Е. П. экономическая статистика: Учебник. М.: КноРус, 2019. – 300 с.
12. Ниворожкина Л. И. Статистика: учебник для бакалавров. – Москва: Дашков и Ко: Наука-Спектр, 2015. – 415 с.
13. Статистика для бакалавров: учебное пособие. / Очкин О. А. – М.: Феникс, 2015
14. Статистика: рабочая тетрадь для практических занятий студентов./ Касаева Т. В. – Витебск: ВГТУ, 2020.
15. Статистика: учебник / [И. И. Елисеева и др.]. – Москва: Проспект, 2015. – 443 с.
16. Статистика: учебно-практическое пособие / [М. Г. Назаров и др.] – Москва: КноРус, 2018. – 479 с.

17. Статистика: учебное пособие для высших учебных заведений по экономическим специальностям / В.М. Гусаров, Е. И. Кузнецова. – Москва: ЮНИТИ-ДАНА, 2016. – 479 с.
18. Статистика: теория и практика в Excel: учебное / В. С. Лялин, И.Г. Зверева, Н. Г. Никифорова. – Москва: Финансы и статистика: Инфра-М, 2016. – 446 с.
19. Статистика: учебное пособие. / Под ред. Н. В. Агабекова и др. – Мн: БГЭУ, 2020. – 303 с.
20. Статистика: учебное пособие. / Лацкевич Н. В. Минск: Вышэйшая школа, 2015.