

СВОДКА И ГРУППИРОВКА ДАННЫХ СТАТИСТИЧЕСКОГО НАБЛЮДЕНИЯ

Понятия сводки и группировки статистических данных

Собранный в процессе статистического наблюдения материал нуждается в определенной обработке, сведении разрозненных данных воедино.

Научно организованная обработка материалов наблюдения (по заранее разработанной программе), включающая в себя кроме обязательного контроля собранных данных систематизацию, группировку материалов, составление таблиц, получение итогов и производных показателей (средних, относительных величин), называется в статистике **сводкой**.

Сводка

Сводка представляет собой второй этап статистического исследования. **Целью** сводки **является получение** на основе сведенных материалов **обобщающих статистических показателей, отражающих сущность социально-экономических явлений и определенные статистические закономерности.**

Сводка

Сводка представляет собой второй этап статистического исследования. **Целью** сводки **является получение** на основе сведенных материалов **обобщающих статистических показателей, отражающих сущность социально-экономических явлений и определенные статистические закономерности.**

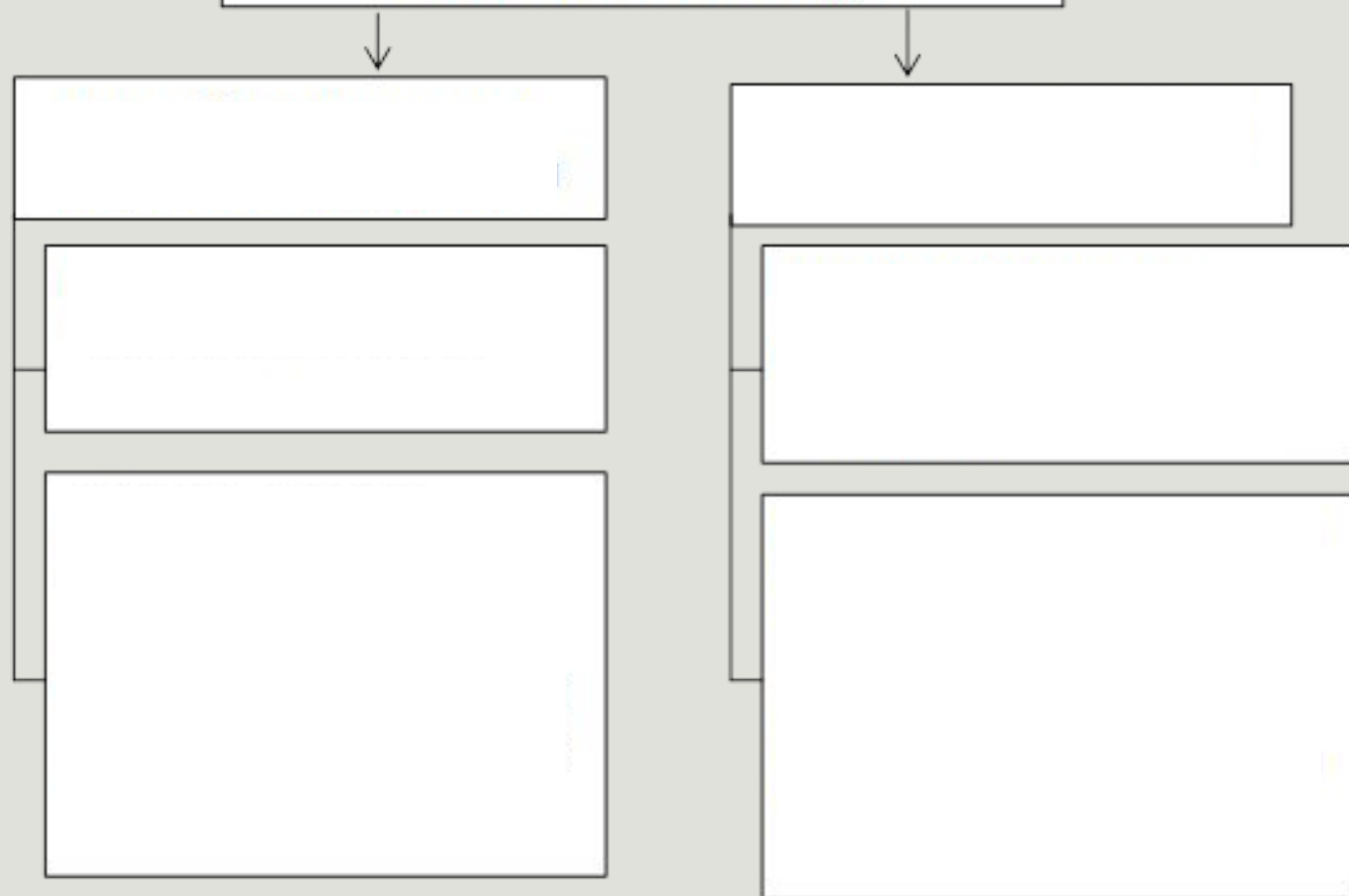
Сводка - это комплекс последовательных операций по обобщению конкретных единичных фактов, образующих совокупность, для выявления типичных черт и закономерностей, присущих изучаемому явлению в целом.

Сводка

Статистическая сводка осуществляется по программе, которая должна разрабатываться еще до сбора статистических данных, практически одновременно с составлением плана и программы статистического наблюдения.

Программа сводки включает определение групп и подгрупп; системы показателей; видов таблиц.

Статистическая сводка



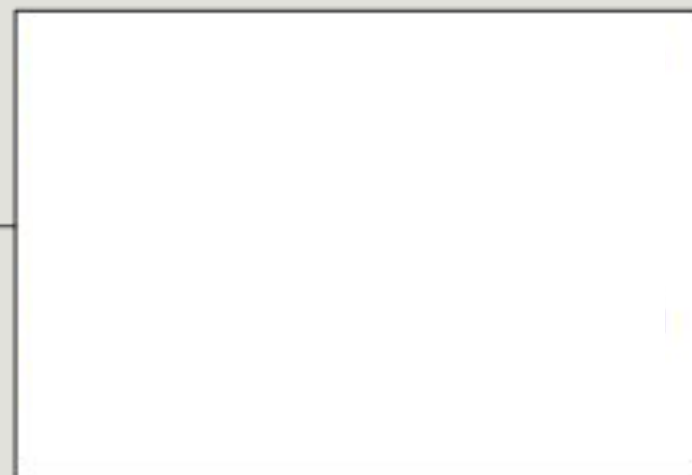
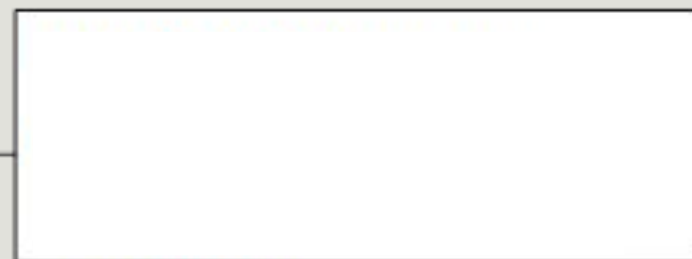
Статистическая сводка



по глубине и точности
обработки информации



по форме обработки
информации



Статистическая сводка

```
graph TD; A[Статистическая сводка] --> B[по глубине и точности обработки информации]; A --> C[по форме обработки информации]; B --> D[простая – подсчет общих итогов по совокупности]; B --> E[ ]; C --> F[ ]; C --> G[ ]
```

по глубине и точности
обработки информации

простая – подсчет общих
итогов по совокупности

по форме обработки
информации

Статистическая сводка

```
graph TD; A[Статистическая сводка] --> B[по глубине и точности обработки информации]; A --> C[по форме обработки информации]; B --> D[простая – подсчет общих итогов по совокупности]; B --> E[сложная – комплекс действий по разделению совокупности на группы, составлению системы показателей, подсчету числа единиц и итогов по группам, оформлению таблиц]; C --> F[ ]; C --> G[ ]
```

по глубине и точности
обработки информации

простая – подсчет общих
итогов по совокупности

сложная – комплекс
действий по разделению
совокупности на группы,
составлению системы
показателей, подсчету числа
единиц и итогов по группам,
оформлению таблиц

по форме обработки
информации

Статистическая сводка

```
graph TD; A[Статистическая сводка] --> B[по глубине и точности обработки информации]; A --> C[по форме обработки информации]; B --> D[простая – подсчет общих итогов по совокупности]; B --> E[сложная – комплекс действий по разделению совокупности на группы, составлению системы показателей, подсчету числа единиц и итогов по группам, оформлению таблиц]; C --> F[централизованная – все данные поступают и обрабатываются в одной организации]; C --> G[ ];
```

**по глубине и точности
обработки информации**

простая – подсчет общих
итогов по совокупности

сложная – комплекс
действий по разделению
совокупности на группы,
составлению системы
показателей, подсчету числа
единиц и итогов по группам,
оформлению таблиц

**по форме обработки
информации**

централизованная – все
данные поступают и
обрабатываются в одной
организации

Статистическая сводка

```
graph TD; A[Статистическая сводка] --> B[по глубине и точности обработки информации]; A --> C[по форме обработки информации]; B --> D[простая – подсчет общих итогов по совокупности]; B --> E[сложная – комплекс действий по разделению совокупности на группы, составлению системы показателей, подсчету числа единиц и итогов по группам, оформлению таблиц]; C --> F[централизованная – все данные поступают и обрабатываются в одной организации]; C --> G[децентрализованная – отчеты предприятий поступают и сводятся в статорганах субъектов РФ, затем итоги поступают в Федеральную службу и сводятся там в целом по стране];
```

по глубине и точности обработки информации

простая – подсчет общих итогов по совокупности

сложная – комплекс действий по разделению совокупности на группы, составлению системы показателей, подсчету числа единиц и итогов по группам, оформлению таблиц

по форме обработки информации

централизованная – все данные поступают и обрабатываются в одной организации

децентрализованная – отчеты предприятий поступают и сводятся в статорганах субъектов РФ, затем итоги поступают в Федеральную службу и сводятся там в целом по стране

Группировка

Группировка – это разделение единиц изучаемой совокупности на однородные группы по определенным, существенным для них признакам.

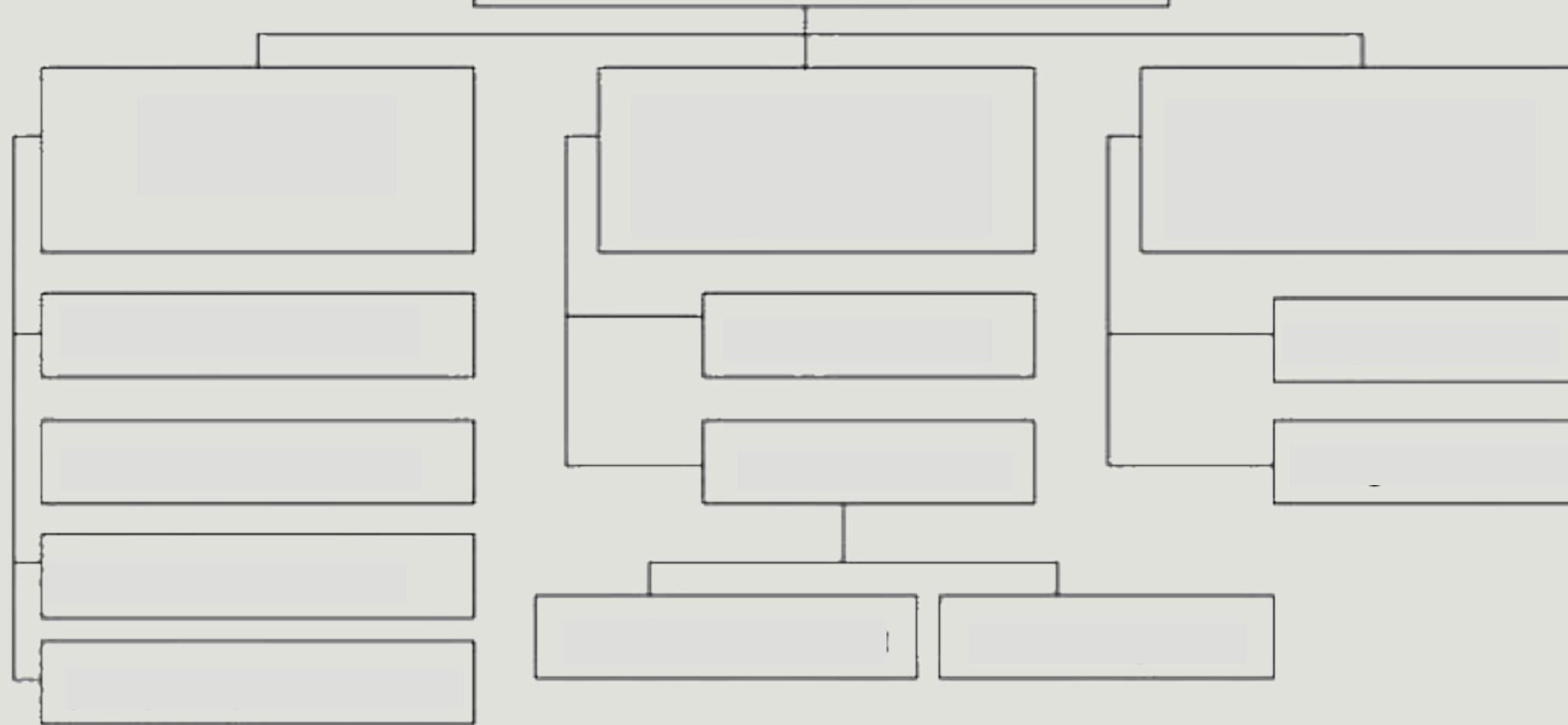
Группировка

Группировка – это разделение единиц изучаемой совокупности на однородные группы по определенным, существенным для них признакам.

Группировка в статистическом анализе выполняет следующие определенные **функции**:

- **выделение социально-экономических типов явлений;**
- **изучение структуры и структурных сдвигов, происходящих в социально-экономических явлениях;**
- **анализ взаимосвязей между явлениями.**

Виды группировок

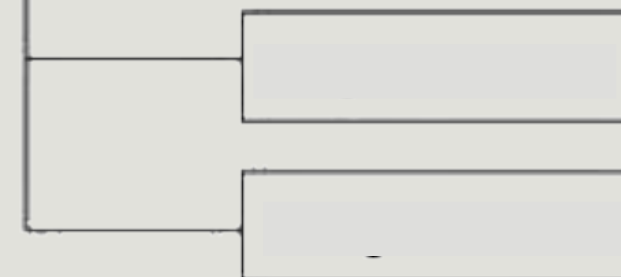
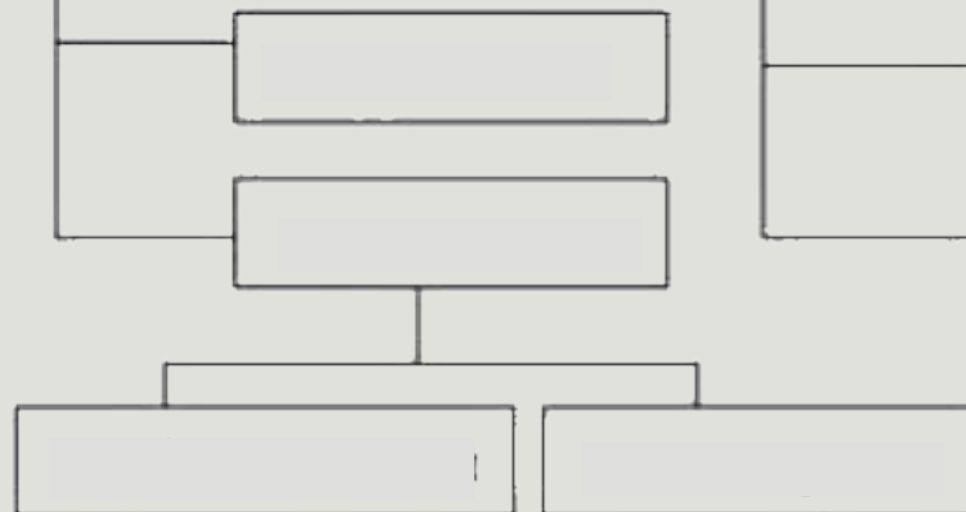
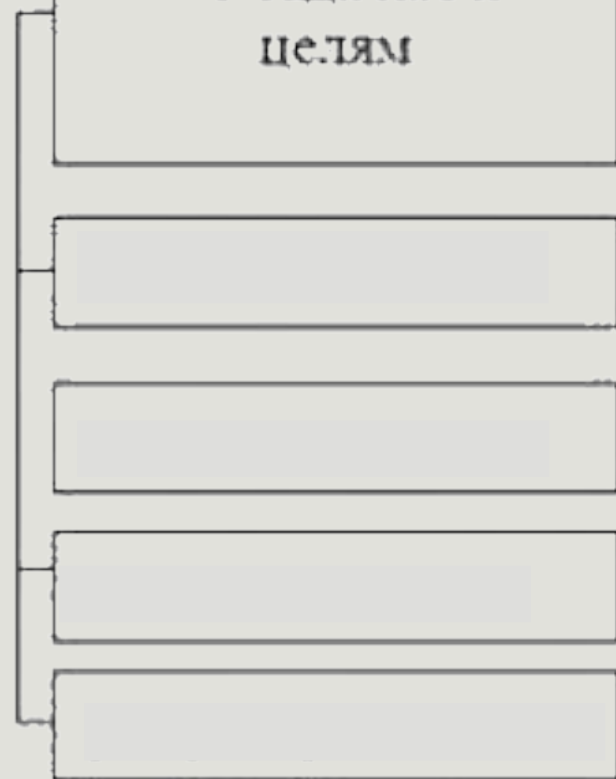


Виды группировок

по задачам и
целям

по числу
группировочных
признаков

по
упорядоченности
исходных данных



Виды группировок

по задачам и
целям

типологическая

структурная

аналитическая

ряд распределения

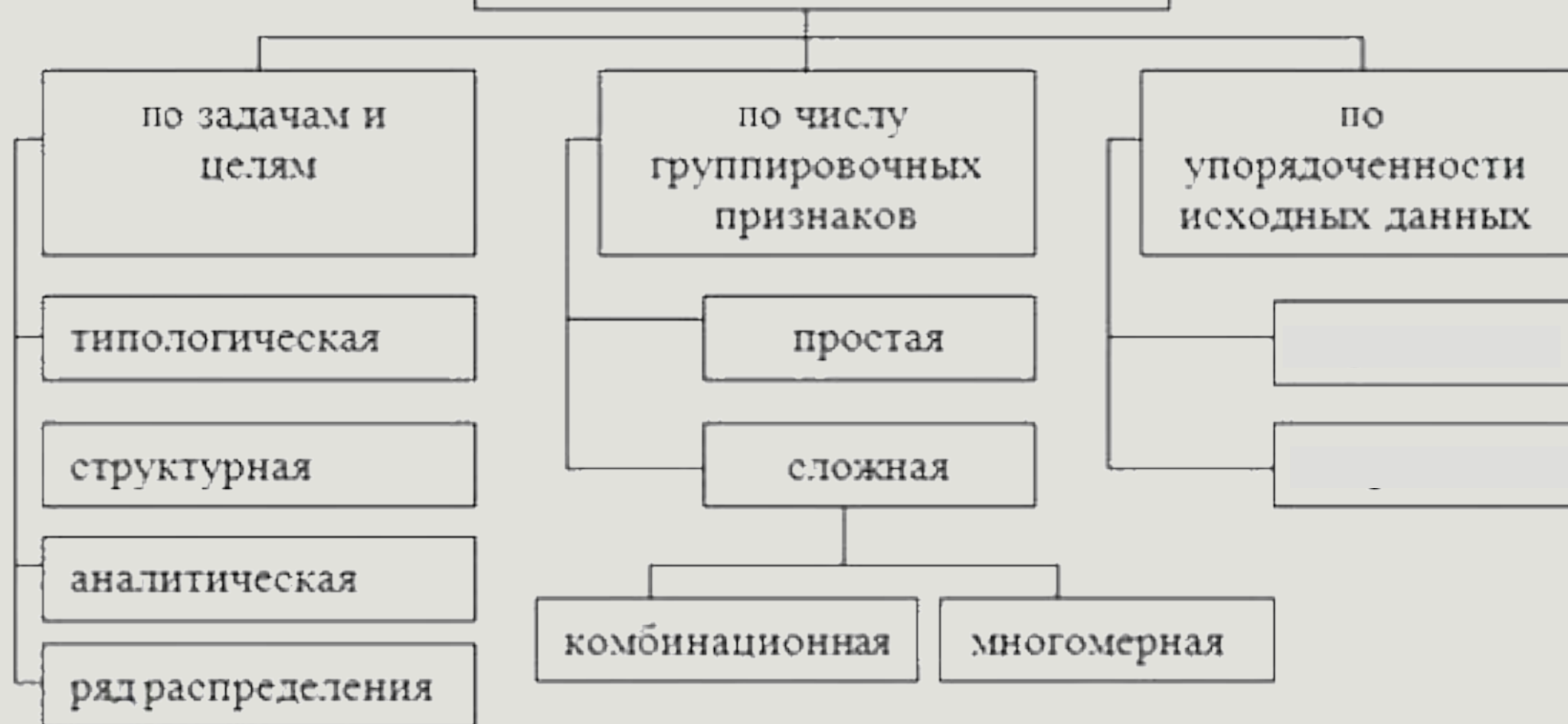
по числу
группировочных
признаков

по
упорядоченности
исходных данных

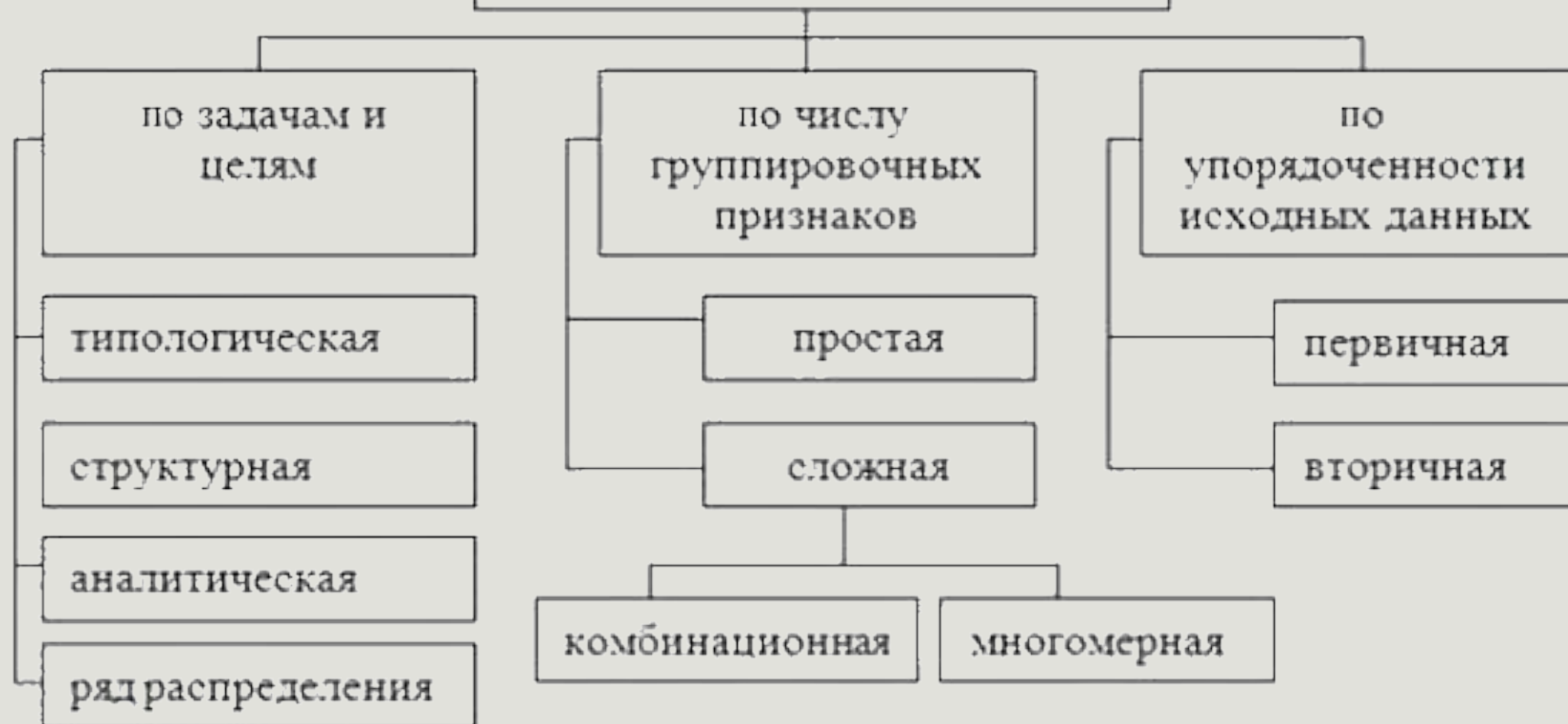
Виды группировок



Виды группировок



Виды группировок



Типологическая группировка

это разделение качественно неоднородной совокупности на отдельные качественно однородные группы и выявление на этой основе экономических типов явлений.

Таким образом, основная задача такой группировки — это идентификация типов социально-экономических явлений, поэтому важное значение при ее построении должно уделяться выбору группировочного признака.

Аналитическая группировка

это исследование взаимосвязей варьирующих признаков в пределах однородной совокупности.

Аналитическая группировка

это исследование взаимосвязей варьирующих признаков в пределах однородной совокупности.

При ее построении можно установить взаимосвязи между двумя признаками и более.

При этом один признак будет результативным, а другой (другие) - факторным.

Структурная группировка

это выявление закономерностей распределения единиц однородной совокупности по варьирующим значениям исследуемого признака.

Она позволяет изучить структуру совокупности и происходящих в ней сдвигов.

Надобность в таких группировках возникает потому, что однородность однокачественных явлений, элементов, входящих в статистическую совокупность, отнюдь не означает их тождественности.

Структурные группировки отличаются от типологических не столько по внешнему виду, сколько по целям, т. е. отличаются по уровню качественных различий между группами.

Ряд распределения

Разновидностью структурной группировки является ряд распределения.

Ряд распределения - это упорядоченное распределение единиц совокупности на группы по определенному признаку.

Правила построения ряда распределения аналогичны правилам построения группировки.

Ряд распределений

Рядом распределения в статистике называется упорядоченное распределение единиц совокупности на группы по какому-либо одному признаку: по качественному или количественному.

Если ряд построен по качественному признаку, то он называется атрибутивным, а если по количественному признаку, то вариационный.

Ряд распределения

В зависимости от признака, положенного в основу распределения, различают атрибутивные и вариационные ряды распределения.

Атрибутивными называют распределения построенные по качественным признакам, т.е. признаками не имеющих числового выражения.

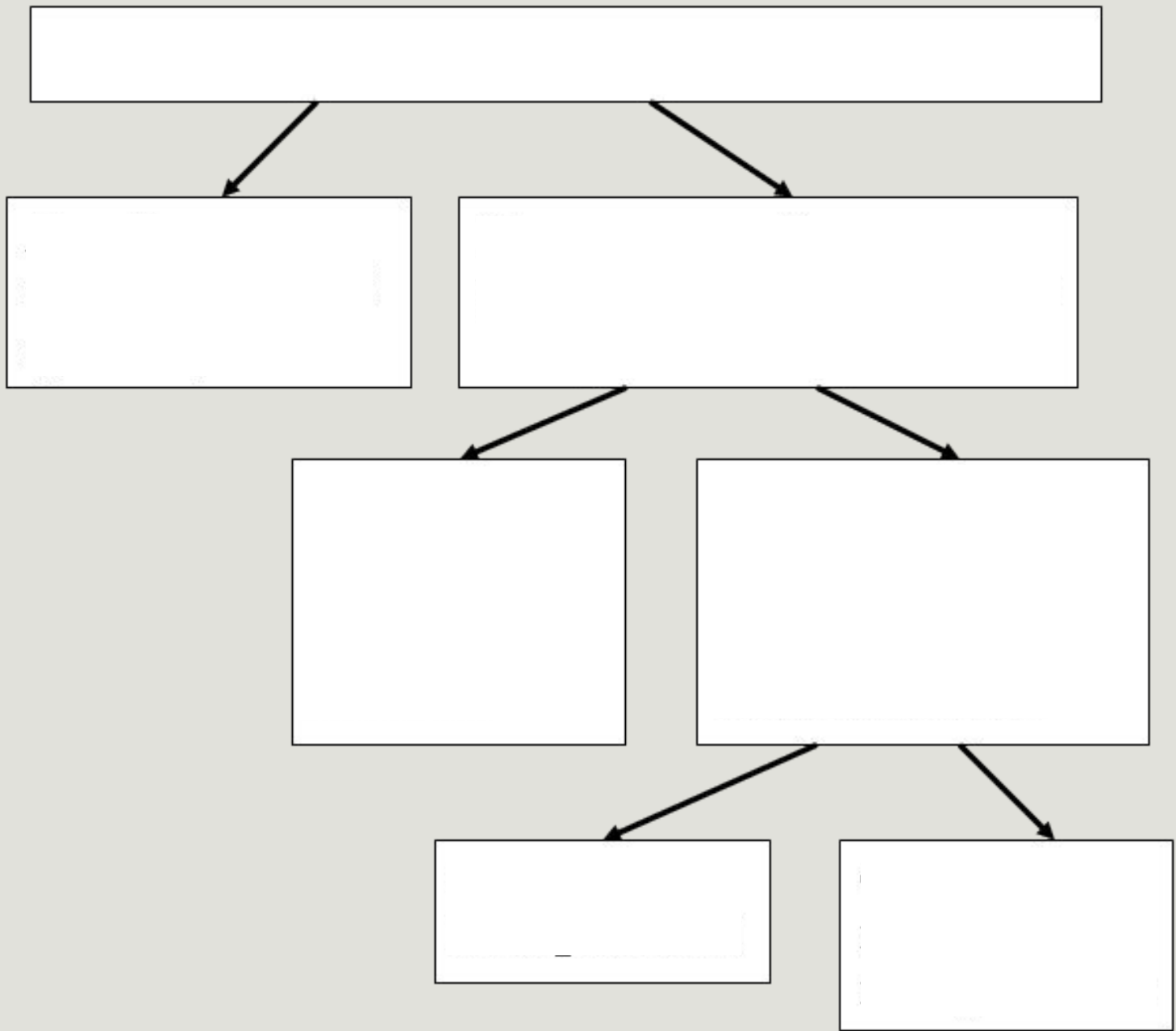
Ряд распределений

Вариационный ряд характеризуется двумя элементами: вариантой (X) и частотой (f).

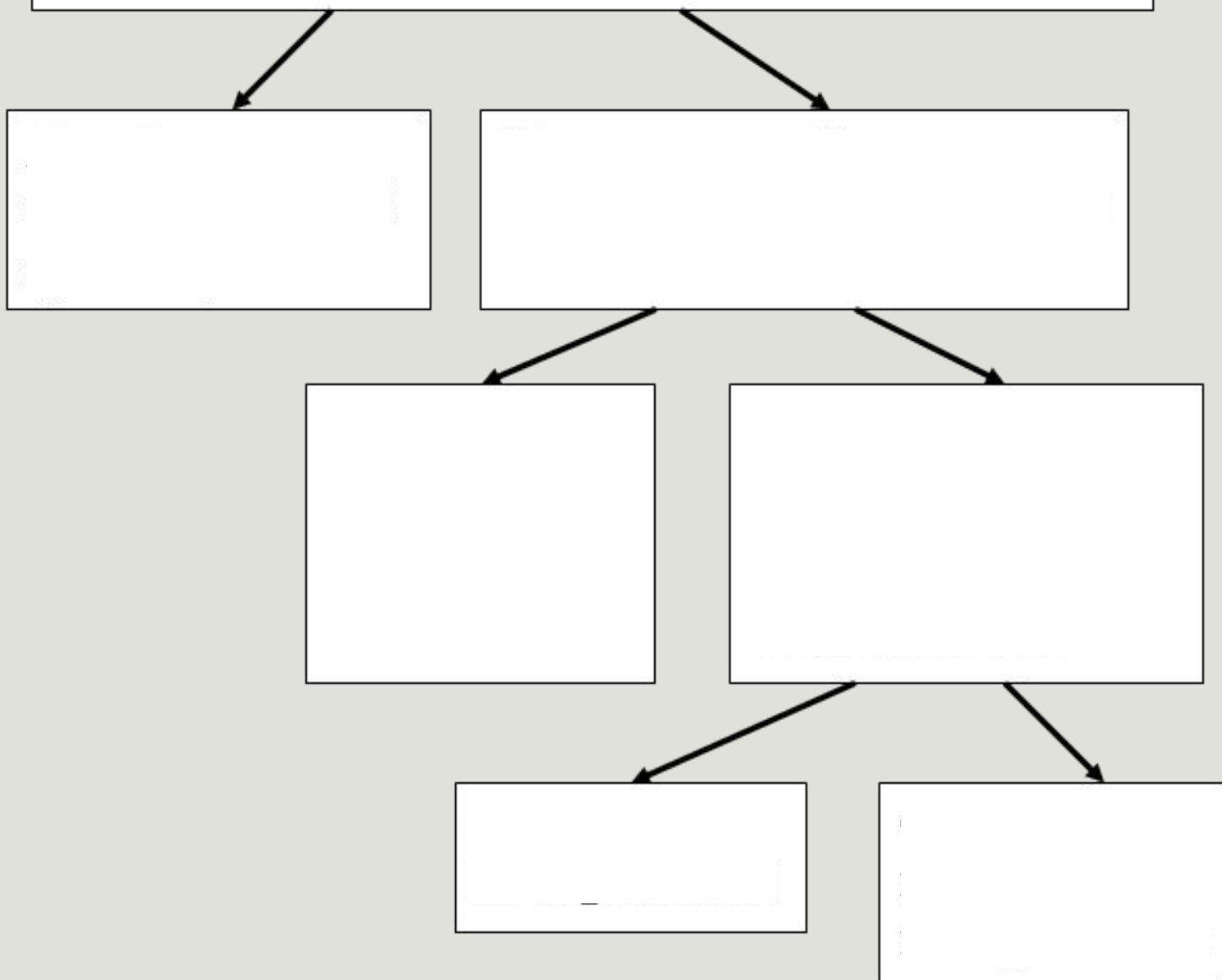
Варианта – это отдельное значение признака отдельной единицы или группы совокупности.

Число, показывающее, сколько раз встречается то или иное значение признака, называется частотой. Если частота выражена относительным числом, то она называется частотостью.

Вариационный ряд может быть интервальным, когда определены границы «от» и «до», а может быть дискретным, когда изучаемый признак характеризуется определенным числом.



Классификация рядов распределения

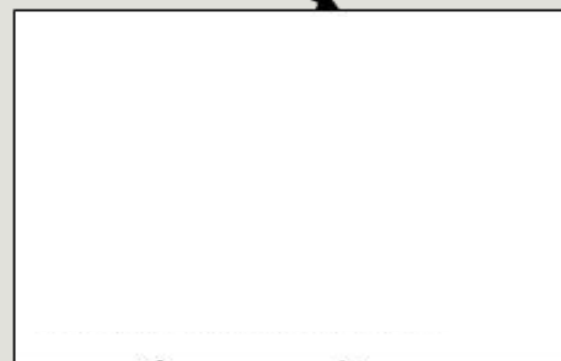
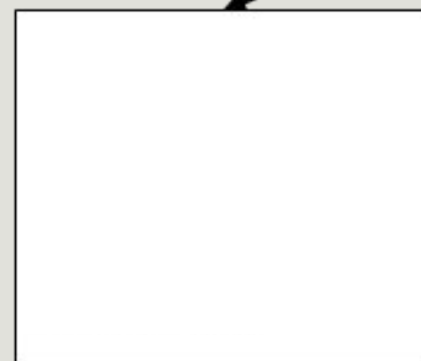


Классификация рядов распределения

```
graph TD; A[Классификация рядов распределения] --> B[Атрибутивные<br/>(атрибутивный признак)]; A --> C[Вариационный<br/>(количественный признак)]; C --> D[ ]; C --> E[ ]; D --> F[ ]; E --> G[ ]
```

Атрибутивные
(атрибутивный
признак)

Вариационный
(количественный признак)



Классификация рядов распределения

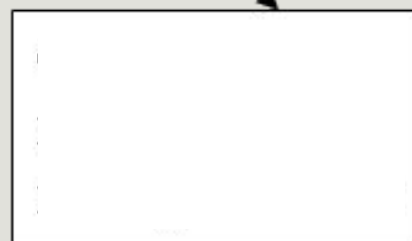
```
graph TD; A[Классификация рядов распределения] --> B[Атрибутивные  
(атрибутивный признак)]; A --> C[Вариационный  
(количественный признак)]; C --> D[Дискретный  
(дискретный, изменяющийся в небольших пределах)]; C --> E[Интервальный  
(непрерывный; дискретный, изменяющийся в больших пределах)]; E --> F[ ]; E --> G[ ]
```

Атрибутивные
(*атрибутивный признак*)

Вариационный
(*количественный признак*)

Дискретный
(*дискретный, изменяющийся в небольших пределах*)

Интервальный
(*непрерывный; дискретный, изменяющийся в больших пределах*)



Классификация рядов распределения

```
graph TD; A[Классификация рядов распределения] --> B[Атрибутивные  
(атрибутивный признак)]; A --> C[Вариационный  
(количественный признак)]; C --> D[Дискретный  
(дискретный, изменяющийся в небольших пределах)]; C --> E[Интервальный  
(непрерывный; дискретный, изменяющийся в больших пределах)]; E --> F[С равными интервалами]; E --> G[С неравными интервалами];
```

Атрибутивные
(*атрибутивный признак*)

Вариационный
(*количественный признак*)

Дискретный
(*дискретный, изменяющийся в небольших пределах*)

Интервальный
(*непрерывный; дискретный, изменяющийся в больших пределах*)

С равными интервалами

С неравными интервалами

Пример: распределение студентов в группе по полу

Женщины	18	60
Мужчины	12	40
Всего	30	100

Пример: распределение студентов в группе по полу

Группа студентов по полу	Число студентов	Удельный вес в общей численности студентов, %
Женщины	18	60
Мужчины	12	40
Всего	30	100

Пример: распределение студентов в группе по полу

Группа студентов по полу	Число студентов	Удельный вес в общей численности студентов, %
Женщины	18	60
Мужчины	12	40
Всего	30	100

Пример: распределение студентов в группе по полу

Группа студентов по полу	Число студентов	Удельный вес в общей численности студентов, %
Женщины	18	60
Мужчины	12	40
Всего	30	100

Пример: распределение студентов в группе по полу

Группа студентов по полу	Число студентов	Удельный вес в общей численности студентов, %
Женщины	18	60
Мужчины	12	40
Всего	30	100

Вариационный ряд

Вариационным рядом называют ряд распределения, построенный по количественному признаку.

Любой вариационный ряд состоит из двух элементов: вариантов и частот.

Вариационный ряд

Вариантами называются отдельные значения признака, которые он принимает в вариационном ряду, т. е. конкретное значение варьирующего признака.

Частотами называются численности отдельных вариантов или каждой группы вариационного ряда, т. е. это числа, которые показывают, как часто встречаются те или иные варианты в ряду распределения.

Сумма всех частот определяет численность всей совокупности, ее объем.

Вариационный ряд

Частостями называются частоты, выраженные в долях единицы или в процентах к итогу.

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

Вариационный ряд



Дискретный вариационный ряд

Дискретный вариационный ряд характеризует распределение единиц совокупности по дискретному признаку, принимающему только целые значения



Пример. Распределение семей по числу детей

Число детей в семье, чел.	Число семей, ед.	Удельный вес, % к итогу

Пример. Распределение семей по числу детей

Число детей в семье, чел.	Число семей, ед.	Удельный вес, % к итогу
1		
2		
3		
Итого		

Пример. Распределение семей по числу детей

Число детей в семье, чел.	Число семей, ед.	Удельный вес, % к итогу
1	600	
2	300	
3	100	
Итого	1000	

Пример. Распределение семей по числу детей

Число детей в семье, чел.	Число семей, ед.	Удельный вес, % к итогу
1	600	60
2	300	30
3	100	10
Итого	1000	100

Интервальные вариационные ряды

Построение интервальных вариационных рядов целесообразно прежде всего при непрерывной вариации признака, а также если дискретная вариация проявляется в широких пределах, т. е. число вариантов прерывного признака достаточно велико.

Пример. Распределение работников фирмы по уровню дохода

Группы работников по уровню доходов, тыс.руб.	Число работников, чел.	Удельный вес, % к итогу

Пример. Распределение работников фирмы по уровню дохода

Группы работников по уровню доходов, тыс.руб.	Число работников, чел.	Удельный вес, % к итогу
До 5,0		
5,0 – 7,5		
7,5 – 10,0		
10,0 и более		

Пример. Распределение работников фирмы по уровню дохода

Группы работников по уровню доходов, тыс.руб.	Число работников, чел.	Удельный вес, % к итогу
До 5,0	60	
5,0 – 7,5	30	
7,5 – 10,0	15	
10,0 и более	10	

Пример. Распределение работников фирмы по уровню дохода

Группы работников по уровню доходов, тыс.руб.	Число работников, чел.	Удельный вес, % к итогу
До 5,0	60	52,2
5,0 – 7,5	30	26,1
7,5 – 10,0	15	13,0
10,0 и более	10	8,7

Пример. Распределение работников фирмы по уровню дохода

Группы работников по уровню доходов, тыс.руб.	Число работников, чел.	Удельный вес, % к итогу
До 5,0	60	52,2
5,0 – 7,5	30	26,1
7,5 – 10,0	15	13,0
10,0 и более	10	8,7
Итого	115	100

Ранжированный вариационный ряд распределения		Дискретный вариационный ряд распределения	
Количество полученных книг	Число студентов, получивших такое количество книг	Количество полученных книг	Доля студентов в общей совокупности
2	7	2	$7/40 = 0,175$
3	9	3	$9/40 = 0,225$
4	9	4	$9/40 = 0,225$
5	5	5	$5/40 = 0,125$
6	6	6	$6/40 = 0,150$
7	3	7	$3/40 = 0,075$
10	1	10	$1/40 = 0,025$
Итого:	40	Итого:	1

Пример построения вариационного ряда

Пример. Имеются данные о тарифных разрядах 60 рабочих одного их цехов завода.

2	4
4	3
3	4
6	5
5	2

Распределить рабочих по тарифному разряду, построить вариационный ряд.

Пример построения вариационного ряда

Пример. Имеются данные о тарифных разрядах 60 рабочих одного их цехов завода.

2	4	5	6
4	3	3	4
3	4	3	5
6	5	4	4
5	2	6	2

Распределить рабочих по тарифному разряду, построить вариационный ряд.

Пример построения вариационного ряда

Пример. Имеются данные о тарифных разрядах 60 рабочих одного их цехов завода.

2	4	5	6	5	2
4	3	3	4	4	4
3	4	3	5	4	3
6	5	4	4	4	2
5	2	6	2	3	3

Распределить рабочих по тарифному разряду, построить вариационный ряд.

Пример построения вариационного ряда

Пример. Имеются данные о тарифных разрядах 60 рабочих одного их цехов завода.

2	4	5	6	5	2	3	4
4	3	3	4	4	4	4	5
3	4	3	5	4	3	5	3
6	5	4	4	4	2	3	4
5	2	6	2	3	3	4	5

Распределить рабочих по тарифному разряду, построить вариационный ряд.

Пример построения вариационного ряда

Пример. Имеются данные о тарифных разрядах 60 рабочих одного их цехов завода.

2	4	5	6	5	2	3	4	1	4
4	3	3	4	4	4	4	5	5	3
3	4	3	5	4	3	5	3	3	2
6	5	4	4	4	2	3	4	4	6
5	2	6	2	3	3	4	5	4	4

Распределить рабочих по тарифному разряду, построить вариационный ряд.

Пример построения вариационного ряда

Пример. Имеются данные о тарифных разрядах 60 рабочих одного их цехов завода.

2	4	5	6	5	2	3	4	1	4	3	3
4	3	3	4	4	4	4	5	5	3	4	1
3	4	3	5	4	3	5	3	3	2	3	4
6	5	4	4	4	2	3	4	4	6	5	1
5	2	6	2	3	3	4	5	4	4	6	4

Распределить рабочих по тарифному разряду, построить вариационный ряд.

Выпишем все значения признака в порядке возрастания и посчитаем число рабочих в каждой группе

Распределение рабочих по разряду

Разряд рабочих (X)	Число рабочих	
	человек (f)	в % к итогу (частность)
1		
2		
3		
4		
5		
6		
Итого		

Выпишем все значения признака в порядке возрастания и посчитаем число рабочих в каждой группе

Распределение рабочих по разряду

Разряд рабочих (X)	Число рабочих	
	человек (f)	в % к итогу (частность)
1	3	
2	6	
3	15	
4	20	
5	10	
6	6	
Итого	60	

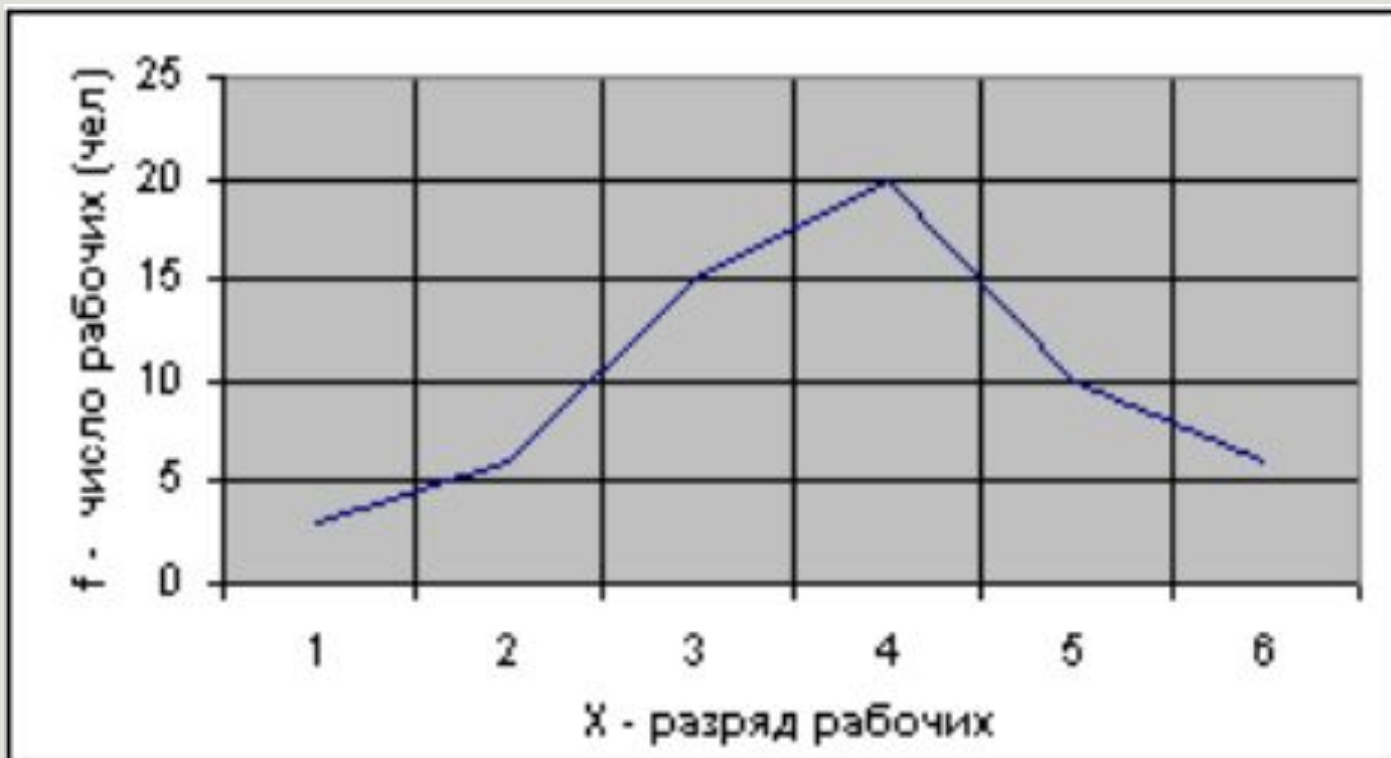
Выпишем все значения признака в порядке возрастания и посчитаем число рабочих в каждой группе

Распределение рабочих по разряду

Разряд рабочих (X)	Число рабочих	
	человек (f)	в % к итогу (частность)
1	3	5,0%
2	6	10,0%
3	15	25,0%
4	20	33,3%
5	10	16,7%
6	6	10,0%
Итого	60	100,0%

Мы получили вариационный дискретный ряд, в котором изучаемый признак (разряд рабочего) представлен определенным числом.

На основании данного ряда распределения построили поверхность распределения.



Построение интервального ряда с равными интервалами

Известны данные о стоимости основного капитала 50 фирм в млн руб. Требуется показать распределение фирм по стоимости основного капитала.

10,4	18,6
28,2	25,2
18,1	16,8
14,2	20,8
12,9	12,6

Построение интервального ряда с равными интервалами

Известны данные о стоимости основного капитала 50 фирм в млн руб. Требуется показать распределение фирм по стоимости основного капитала.

10,4	18,6	10,3	26,0
28,2	25,2	18,4	17,5
18,1	16,8	38,5	37,7
14,2	20,8	13,5	42,4
12,9	12,6	16,8	19,7

Построение интервального ряда с равными интервалами

Известны данные о стоимости основного капитала 50 фирм в млн руб. Требуется показать распределение фирм по стоимости основного капитала.

10,4	18,6	10,3	26,0	45,0	18,2
28,2	25,2	18,4	17,5	41,8	14,6
18,1	16,8	38,5	37,7	17,9	29,0
14,2	20,8	13,5	42,4	15,5	17,9
12,9	12,6	16,8	19,7	18,3	36,8

Построение интервального ряда с равными интервалами

Известны данные о стоимости основного капитала 50 фирм в млн руб. Требуется показать распределение фирм по стоимости основного капитала.

10,4	18,6	10,3	26,0	45,0	18,2	17,3	19,2
28,2	25,2	18,4	17,5	41,8	14,6	10,0	37,8
18,1	16,8	38,5	37,7	17,9	29,0	10,1	28,0
14,2	20,8	13,5	42,4	15,5	17,9	19,2	10,8
12,9	12,6	16,8	19,7	18,3	36,8	15,0	37,0

Построение интервального ряда с равными интервалами

Известны данные о стоимости основного капитала 50 фирм в млн руб. Требуется показать распределение фирм по стоимости основного капитала.

10,4	18,6	10,3	26,0	45,0	18,2	17,3	19,2	25,8	18,7
28,2	25,2	18,4	17,5	41,8	14,6	10,0	37,8	10,5	16,0
18,1	16,8	38,5	37,7	17,9	29,0	10,1	28,0	12,0	14,0
14,2	20,8	13,5	42,4	15,5	17,9	19,2	10,8	12,1	12,4
12,9	12,6	16,8	19,7	18,3	36,8	15,0	37,0	13,0	19,5

Построение интервального ряда с равными интервалами

Чтобы показать распределение фирм по стоимости основного капитала, сначала решим вопрос о количестве групп, которые хотим выделить. Предположим, решили выделить 5 групп предприятий. Затем определим величину интервала в группе. Для этого воспользуемся формулой

$$i = \frac{X_{\max} - X_{\min}}{n} .$$

Построение интервального ряда с равными интервалами

Чтобы показать распределение фирм по стоимости основного капитала, сначала решим вопрос о количестве групп, которые хотим выделить. Предположим, решили выделить 5 групп предприятий. Затем определим величину интервала в группе. Для этого воспользуемся формулой

$$i = \frac{X_{\max} - X_{\min}}{n} .$$

$$i = \frac{45 - 10}{5} = 7 .$$

Построение интервального ряда с равными интервалами

Чтобы показать распределение фирм по стоимости основного капитала, сначала решим вопрос о количестве групп, которые хотим выделить. Предположим, решили выделить 5 групп предприятий. Затем определим величину интервала в группе. Для этого воспользуемся формулой

$$i = \frac{X_{\max} - X_{\min}}{n} .$$

$$i = \frac{45 - 10}{5} = 7 .$$

Путем прибавления величины интервала к минимальному значению признака, получим группы фирм по стоимости основного капитала.

Единица, обладающая двойным значением, относится к той группе, где она выступает в роли верхней границы (т.е. значение признака 17 пойдет в первую группу, 24 – во вторую и т.д.).

Подсчитаем число заводов в каждой группе.

Построение интервального ряда с равными интервалами

Распределение фирм по стоимости основного капитала (млн руб.)

Стоимость основного капитала в млн руб. (X)	Число фирм (частота) (f)	Накопленные частоты (кумулятивные)
10–17	22	22
17–24	18	40
24–31	12	52
31–38	8	60
38–45	5	65

Построение интервального ряда с равными интервалами

Распределение фирм по стоимости основного капитала (млн руб.)

Стоимость основного капитала в млн руб. (X)	Число фирм (частота) (f)	Накопленные частоты (кумулятивные)
10–17	22	22
17–24	14	36

Построение интервального ряда с равными интервалами

Распределение фирм по стоимости основного капитала (млн руб.)

Стоимость основного капитала в млн руб. (X)	Число фирм (частота) (f)	Накопленные частоты (кумулятивные)
10–17	22	22
17–24	14	36
24–31	16	42

Построение интервального ряда с равными интервалами

Распределение фирм по стоимости основного капитала (млн руб.)

Стоимость основного капитала в млн руб. (X)	Число фирм (частота) (f)	Накопленные частоты (кумулятивные)
10–17	22	22
17–24	14	36
24–31	16	42
31–36	4	46

Построение интервального ряда с равными интервалами

Распределение фирм по стоимости основного капитала (млн руб.)

Стоимость основного капитала в млн руб. (X)	Число фирм (частота) (f)	Накопленные частоты (кумулятивные)
10–17	22	22
17–24	14	36
24–31	16	42
31–36	4	46
38–45	4	50

Согласно данному распределению получили вариационный интервальный ряд, из которого следует, что 36 фирм имеют основной капитал стоимостью от 10 до 24 млн руб. и т.д.

Признаки группировки

Факторными называются признаки, оказывающие влияние на изменение результативных.

Результативными называются признаки, изменяющиеся под влиянием факторных.

Основные этапы построения аналитической группировки

- обоснование и выбор факторного и результативного признаков;
- группировка единиц совокупности по факторному признаку;
- подсчет числа единиц в каждой из образованных групп, а также определение объема варьирующих признаков в пределах созданных групп;
- исчисление средних размеров результативного показателя (признака) по каждой из образованных групп;
- оформление результатов группировки в таблице;
- сопоставление изменения значений факторного и результативного признаков, определяющее характер связи между ними, т. е. выявление взаимосвязи между признаками, когда с возрастанием значения факторного признака систематически возрастает или убывает значение признака результативного.

Принципы построения статистических группировок и классификаций

Построение группировки начинается с определения состава группировочных признаков.

Принципы построения статистических группировок и классификаций.

Построение группировки начинается с определения состава группировочных признаков.

Выбор группировочного признака, т. е. признака, по которому производится объединение единиц исследуемой совокупности в группы, — один из самых существенных и сложных вопросов теории группировки и статистического исследования.

Принципы построения статистических группировок и классификаций.

Построение группировки начинается с определения состава группировочных признаков.

Выбор группировочного признака, т. е. признака, по которому производится объединение единиц исследуемой совокупности в группы, — один из самых существенных и сложных вопросов теории группировки и статистического исследования.

От правильного выбора группировочного признака зависят выводы статистического исследования.

Принципы построения статистических группировок и классификаций

Построение группировки начинается с определения состава группировочных признаков.

Выбор группировочного признака, т. е. признака, по которому производится объединение единиц исследуемой совокупности в группы, — один из самых существенных и сложных вопросов теории группировки и статистического исследования.

От правильного выбора группировочного признака зависят выводы статистического исследования.

В качестве основания группировки необходимо использовать существенные обоснованные признаки.

Группировочный признак

Это признак, по которому проводится разбиение единиц совокупности на отдельные группы.

В основание группировки могут быть положены как количественные, так и атрибутивные признаки. Первые имеют числовое выражение (объем торгов, возраст человека, доход семьи и т. д.), а вторые отражают состояние единицы совокупности (пол человека, семейное положение, отраслевую принадлежность предприятия, его форму собственности и т. д.).

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

Группировочный признак

Если группировка строится по атрибутивному признаку, то число групп, как правило, будет столько, сколько имеется градаций, видов состояний у этого признака.

Например, группировка предприятий по формам собственности учитывает муниципальную, федеральную и собственность субъектов Федерации.

Группировочный признак



Группировочный признак

Если группировка проводится по количественному признаку, то число групп зависит от числа единиц исследуемого объекта и степени колеблемости группировочного признака, в каждом отдельном случае его необходимо об



Обоснование по формуле Стерджесса:

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

где n — число групп;

N — число единиц совокупности.

Согласно формуле выбор числа групп зависит от объема совокупности.

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

Обоснование по среднеквадратичному отклонению

Если величина интервала равна $0,5a$, то совокупность разбивается на 12 групп,

Если величина интервала равна $2/3a$ и a , то совокупность делится соответственно на 9 и 6 групп.

Однако при определении групп данными методами существует большая вероятность получения «пустых», или малочисленных, групп.

Когда определено число групп, то следует определить интервалы группировки.

Интервал

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

Нижней границей интервала называется наименьшее значение признака в интервале, а верхней границей - наибольшее значение признака в интервале.

Величина интервала представляет собой разность между верхней и нижней границами интервала.

Интервалы группировки

Интервалы группировки в зависимости от их величины бывают равные и неравные. Если вариация признака проявляется в сравнительно узких границах и распределение носит равномерный характер, то строят группировку с равными интервалами. Величина равного интервала определяется по следующей формуле:

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

где $x_{\max}$ и $x_{\min}$ — максимальное и минимальное значения признака в совокупности;
 n — число групп.

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

Правила записи числа шага интервала

Если величина интервала представляет собой величину, которая имеет один знак до запятой, то в этом случае полученные значения целесообразно округлить до десятых и их использовать в качестве шага интервала.

В приведенном выше примере это будут соответственно значения; 0,9; 1,6; 4,7.

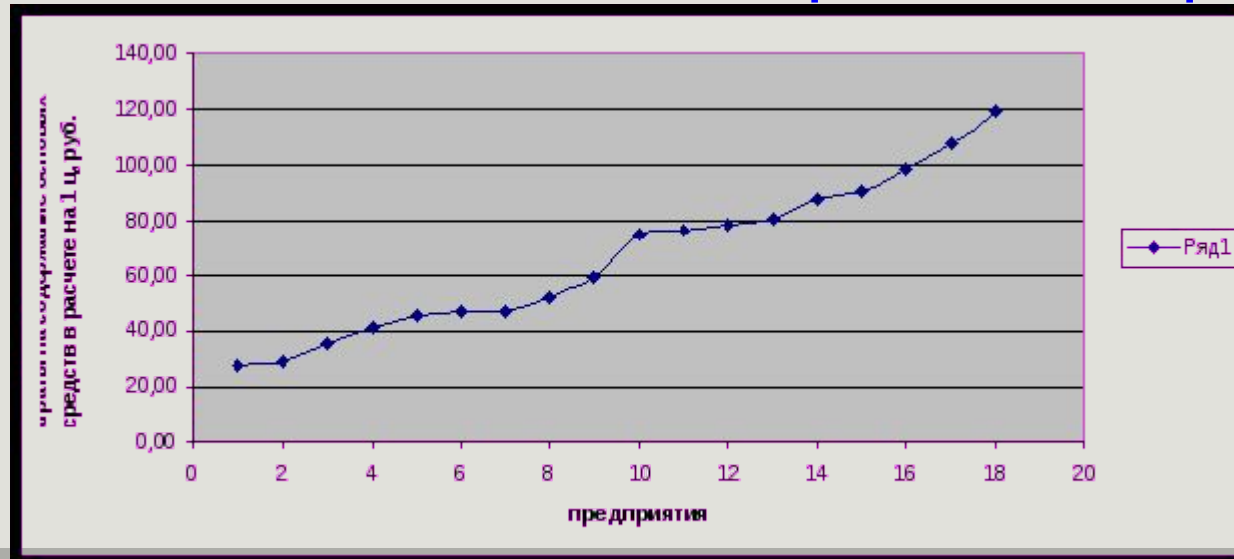
Если рассчитанная величина интервала имеет две значащие цифры до запятой и несколько знаков после запятой, то это значение необходимо округлить до целого числа.

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

Например, 557 следует округлить до 600.

Группировка с неравными интервалами

Если размах вариации признака велик и его значения варьируются неравномерно, то необходимо использовать группировку с неравными интервалами. Неравные интервалы могут быть прогрессивно-возрастающими или прогрессивно-убывающими в арифметической или геометрической прогрессии.



Величина интервалов, изменяющихся в прогрессии

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

а в геометрической прогрессии:

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

где a — константа, имеющая для прогрессивно-возрастающих интервалов знак «+», а для прогрессивно-убывающих интервалов знак «-»;
 q — константа (для прогрессивно-убывающих интервалов $q > 1$; в другом случае — $q < 1$).

Применение неравных интервалов

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

Например, при построении группировки промышленных предприятий строительного комплекса по показателю численности работающих, который варьирует от 400 до 2800 чел., целесообразно рассматривать неравные интервалы.

Поэтому следует образовывать неравные интервалы: 400 - 800; 800 - 1600; 1600 - 2800, т. е. величина каждого последующего интервала больше предыдущего на 400 чел. и увеличивается в арифметической прогрессии.

Интервалы группировок

Интервалы группировок могут быть закрытыми и открытыми.

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

У открытых интервалов указана только одна граница: верхняя - у первого, нижняя — у последнего.

Например, группы коммерческих банков по уровню дохода работающих в них сотрудников (тыс. руб.): до 10; 10 — 20; 20 — 30; 30-40; 40 и более.

Группировка единиц совокупности по количественному признаку

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

Если основанием группировки служит непрерывный признак (например, группы строительных фирм по объему работ (млн. руб.): 12 - 14, 14 - 16, 16 - 18, 18 - 20), то одно и то же значение признака выступает и верхней, и нижней границами двух смежных интервалов.

В данном случае объем работ 14 млн руб. составляет верхнюю границу первого и нижнюю границу второго интервалов; 16 млн руб. - соответственно второго и третьего и т. д., т. е. верхняя граница i -го интервала равна нижней границе ($Ж$) интервала.

Группировка единиц совокупности по количественному признаку

При таком обозначении границ может возникнуть вопрос, в какую группу включать единицы объекта, значения признака у которых совпадают с границами интервалов.

Например, во вторую или третью группу должна войти строительная фирма с объемом работ 16 млн руб. Если нижняя граница формируется по принципу «включительно», а верхняя - по принципу «исключительно», то фирма должна быть отнесена к третьей группе, в противном случае - ко второй.

Для того чтобы правильно отнести к той или иной группе единицу объекта, значение признака у которой совпадает с границами интервалов, можно использовать открытые интервалы (по нашему примеру группы строительных фирм по объему работ преобразуются в следующие: до 14, 14 — 16, 16 — 18, 18 и более).

Группировка единиц совокупности по количественному признаку

В данном случае вопрос отнесения отдельных единиц совокупности, значения которых являются граничными, к той или иной группе решается на основе анализа последнего открытого интервала.

Возможны два случая обозначения открытого последнего интервала:

- 1) 18 млн руб. и более;
- 2) более 18 млн руб.

В первом случае строительные фирмы с объемом работ 16 млн руб. попадут в третью группу; во втором случае — во вторую группу.

Группировка единиц совокупности по прерывному признаку

Если в основании группировки лежит прерывный признак, то нижняя граница i -го интервала равна верхней границе $(i-1)$ интервала, увеличенной на 1.

Например, группы строительных фирм по числу занятого персонала будут иметь вид (чел.): 100 — 150, 151-200, 201-300.

При определении границ интервалов статистических группировок необходимо исходить из того, что изменение количественного признака приводит к появлению нового качества.

В этом случае граница интервала должна устанавливаться там, где происходит переход одного качества в другое.

Это достигается путем использования группировок со специализированными интервалами.

Специализированные интервалы

Специализированными называются интервалы, которые применяются для выделения из совокупности одних и тех же типов по одному и тому же признаку для явлений, находящихся в различных условиях.

Например, группировка по отраслям экономики. При изучении социально-экономических явлений на макроуровне часто применяют группировки, интервалы которых не будут ни прогрессивно-возрастающими, ни прогрессивно-убывающими.

Такие интервалы называются произвольными и, как правило, используются при группировке предприятий, например, по уровню рентабельности

Произвольные интервалы

При изучении социально-экономических явлений на макроуровне часто применяют группировки, интервалы которых не будут ни прогрессивно-возрастающими, ни прогрессивно-убывающими.

Такие интервалы называются произвольными и, как правило, используются при группировке предприятий, например, по уровню рентабельности.

Анализ рядов распределения

Анализ рядов распределения наглядно можно проводить на основе их графического изображения.

Для этой цели строят графики

- ☐ полигон,
- ☐ гистограмму,
- ☐ огиву ,
- ☐ кумуляту распределения.

Полигон

Полигон используется при изображении дискретных вариационных рядов.

Для его построения в прямоугольной системе координат по оси абсцисс в одинаковом масштабе откладываются ранжированные значения варьирующего признака, а по оси ординат наносится шкала для выражения величины частот.

Полученные на пересечении абсцисс и ординат точки соединяют прямыми линиями и получают ломаную линию, называемую **полигоном частот**. Для замыкания полигона крайние точки (слева и справа на ломаной линии) соединяют с точками на оси абсцисс и получают многоугольник

Пример : Пусть дискретный вариационный ряд задан таблицей:

α_i							
k_i							
n_i							

Пример : Пусть дискретный вариационный ряд задан таблицей:

α_i	0	1	2	3	4	5	7
k_i							
n_i							

Пример : Пусть дискретный вариационный ряд задан таблицей:

α_i	0	1	2	3	4	5	7
k_i	8	17	16	10	6	2	1
n_i							

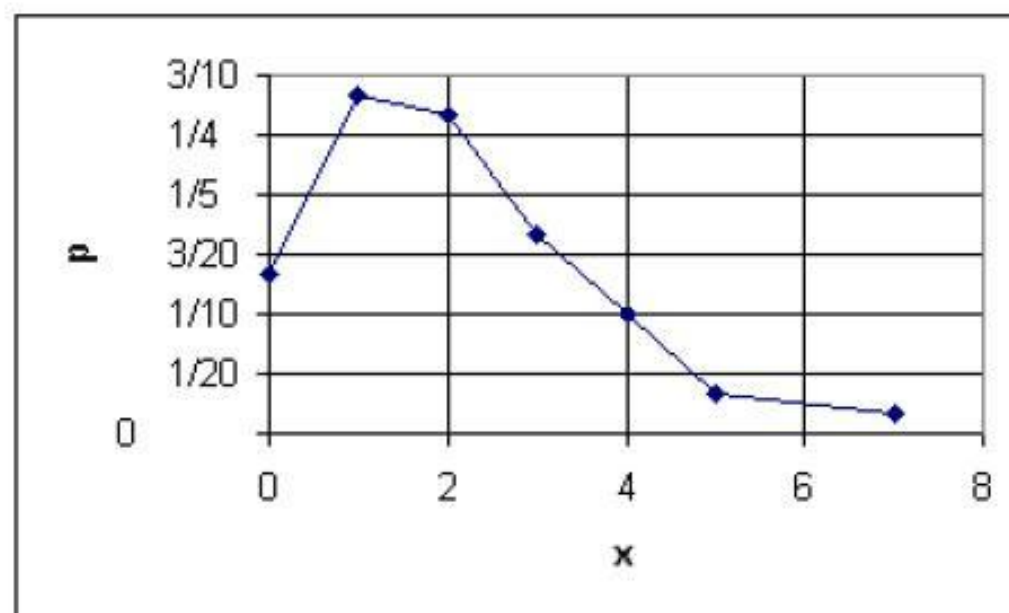
Пример : Пусть дискретный вариационный ряд задан таблицей:

α_i	0	1	2	3	4	5	7
k_i	8	17	16	10	6	2	1
n_i	8/60	17/60	16/60	10/60	6/60	2/60	1/60

Пример : Пусть дискретный вариационный ряд задан таблицей:

α_i	0	1	2	3	4	5	7
k_i	8	17	16	10	6	2	1
n_i	8/60	17/60	16/60	10/60	6/60	2/60	1/60

На рисунке представлен полигон его относительных частот



Гистограмма

Гистограмма применяется для изображения интервального вариационного ряда.

При построении гистограммы на оси абсцисс откладываются величины интервалов, а частоты изображаются прямоугольниками, построенными на соответствующих интервалах.

Высота столбиков должна быть пропорциональна частотам. В результате мы получим на графике гистограмму, где ряд распределения изображен в виде смежных друг с другом столбиков.

Если середины верхних сторон прямоугольников соединить прямыми, то гистограмма может быть преобразована в полигон

Кумулятивная кривая

Для графического изображения вариационных рядов может использоваться также кумулятивная кривая.

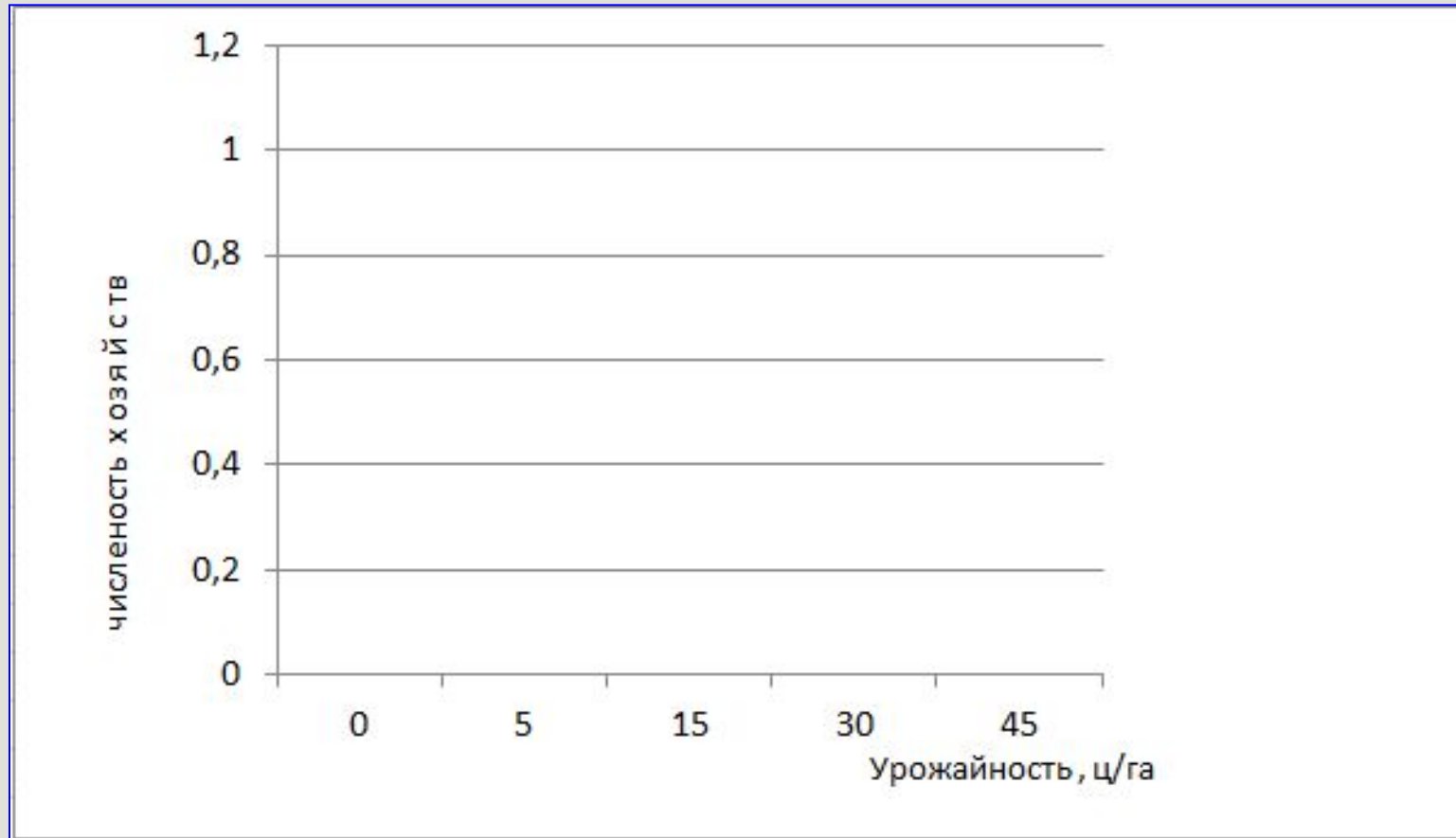
При построении кумулятивной кривой по интервальному вариационному ряду на оси абсцисс откладываются варианты ряда, а на оси ординат — накопленные частоты, которые наносят на поле графика в виде перпендикуляров к оси абсцисс в верхних границах интервалов.

Затем эти перпендикуляры соединяют и получают ломаную линию, т. е. кумулятивную кривую.

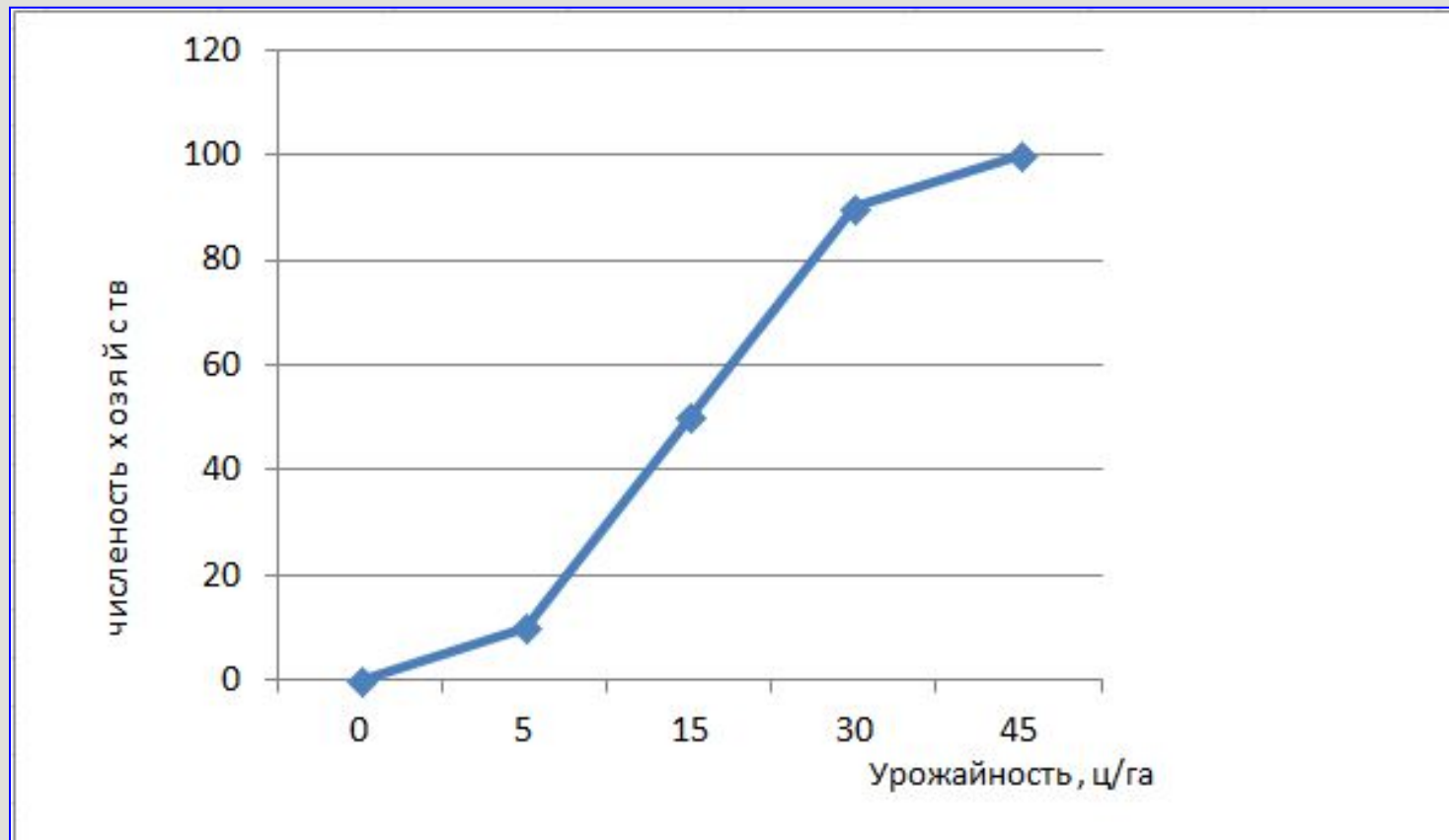
Кумулятивная кривая

Можно построить кумулятивное распределение «не меньше чем», а можно «больше чем». В первом случае график кумулятивного распределения называется кумулятой, во втором - огивой.

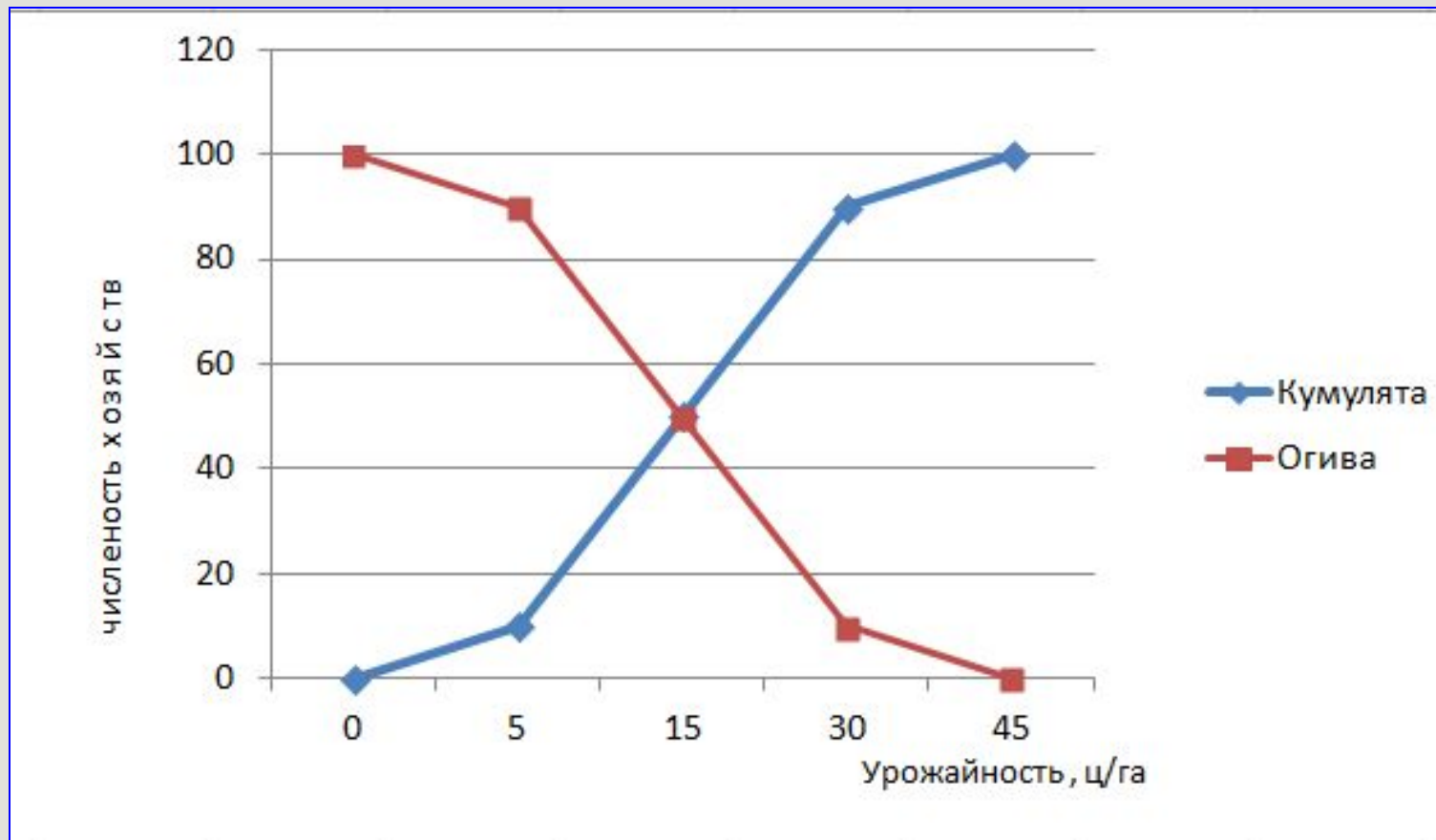
Кумулятивная кривая



Кумулятивная кривая



Кумулятивная кривая



Вторичная группировка

Группировки, построенные за один и тот же период времени, но для разных регионов или, наоборот, для одного региона, но за два разных периода времени, **могут оказаться несопоставимыми из-за различного числа выделенных групп или неодинаковости границ интервалов.**

В таком случае необходима перегруппировка данных с помощью вторичной группировки.

Вторичная группировка

Вторичная группировка - операция по образованию новых групп на основе ранее осуществленной группировки.

Применяют два способа образования новых групп.

Вторичная группировка

Вторичная группировка - операция по образованию новых групп на основе ранее осуществленной группировки.

Применяют два способа образования новых групп.

Первым, наиболее простым и распространенным способом является изменение (чаще укрупнение) первоначальных интервалов.

Вторичная группировка

Вторичная группировка - операция по образованию новых групп на основе ранее осуществленной группировки.

Применяют два способа образования новых групп.

Первым, наиболее простым и распространенным способом является изменение (чаще укрупнение) первоначальных интервалов.

Второй способ получил название долевого перегруппировки и состоит в образовании новых групп на основе закрепления за каждой группой определенной доли единиц совокупности.

Пример

Необходимо провести перегруппировку данных, образовав новые группы с интервалами до 500, 500 - 1000, 1000 - 2000, 2000 — 3000, свыше 3000 руб. поданным о распределении контрактов строительной фирмы по величине прибыли

Номер группы	Число контрактов по величине прибыли, тыс.руб.	Число контрактов, ед.
1		
2		
3		
4		
5		
6		

Номер группы	Число контрактов по величине прибыли, тыс.руб.	Число контрактов, ед.
1	до 400	
2	400 – 1000	
3	1000 – 1800	
4	1800 – 3000	
5	3000- 4000	
6	4000 и более	
	Итог о	

Номер группы	Число контрактов по величине прибыли, тыс.руб.	Число контрактов, ед.
1	до 400	16
2	400 – 1000	20
3	1000 – 1800	44
4	1800 – 3000	74
5	3000- 4000	37
6	4000 и более	9
	Итог о	200

Первая группа

В первую новую группу войдет полностью 1-я группа контрактов и часть 2-й группы.

Чтобы образовать группу до 500 тыс. руб., необходимо от интервала 2-й группы взять 100 тыс. руб.

Величина интервала этой группы составит 600 тыс. руб.

Следовательно, необходимо взять от нее $1/6$ ($100 : 600$).

Аналогичную же часть во вновь образуемую новую группу надо взять и от числа контрактов, т. е. $20 \cdot 1/6 = 3$ контракта.

Тогда в 1-й группе будет контрактов $16 + 3 = 19$ контрактов.

Вторая группа

Вторую новую группу образуют контракты 2-й группы за вычетом отнесенных к 1-й,

т. е, $20 - 3 = 17$ ед.

Третья группа

Во вновь образованную третью группу войдут все контракты 3-й группы и часть контрактов 4-й.

Для определения этой части от интервала 1800 -3000 (ширина интервала равна 1200 тыс. руб.) нужно добавить к предыдущему 200 тыс. руб. (чтобы верхняя граница интервала была равна 2000 руб.).

Следовательно, необходимо взять часть интервала, равную $200 : 1200$, т. е. $1/6$. В этой группе 74 контракта, значит, надо взять 12 ед.

В третью новую группу войдет: $44 + 12 = 56$ контрактов

Четвертая группа

Во вновь образованную четвертую группу войдет:

74 - 12 =; 62 контракта, оставшихся от прежней 4-й группы

Пятая группа

Пятую, вновь образованную группу составят контракты 5-й и 6-й прежних групп: $37 + 9 = 46$ контрактов.

Номер группы	Число контрактов по величине прибыли, тыс.руб.	Число контрактов, ед.
1		
2		
3		
4		
5		

Номер группы	Число контрактов по величине прибыли, тыс.руб.	Число контрактов, ед.
1	до 500	
2	500 – 1000	
3	1000 – 2000	
4	2000 – 3000	
5	3000 и более	
	Итог о	

Номер группы	Число контрактов по величине прибыли, тыс.руб.	Число контрактов, ед.
1	до 500	19
2	500 – 1000	17
3	1000 – 2000	56
4	2000 – 3000	62
5	3000 и более	46
	Итог о	200

Номер группы	Число контрактов по величине прибыли, тыс.руб.	Число контрактов, ед.
1	до 400	16
2	400 – 1000	20
3	1000 – 1800	44
4	1800 – 3000	74
5	3000- 4000	37
6	4000 и более	9
	Итог о	200

Номер группы	Число контрактов по величине прибыли, тыс.руб.	Число контрактов, ед.
1	до 500	19
2	500 – 1000	17
3	1000 – 2000	56
4	2000 – 3000	62
5	3000 и более	46
	Итог о	200