



ОБЩАЯ ТЕОРИЯ СТАТИСТИКИ

Валерия Андреевна Колычева, кафедра
статистики, учета и аудита



**Случай – только
«мера нашего незнания».**

А.В. Васильев.

Теория вероятностей. Лекции.

Казань, 1886

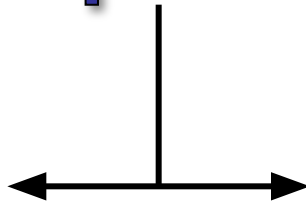
Описательная статистика

Стадии статистического исследования:

- 1. Статистическое наблюдение*
- 2. Сводка и группировка результатов наблюдения*
- 3. Анализ полученных обобщающих показаний*

Общая теория статистики

*Экономическая
статистика*



*Социальная
статистика*

- 1. Промышленности*
- 2. Сельского хозяйства*
- 3. Транспорта*
- 4. Связи*

- 1. Правовая*
- 2. Демографическая*
- 3. Политики*
- 4. Здравоохранения*
- 5. Науки*
- 6. Культуры*

1) Теория статистического наблюдения

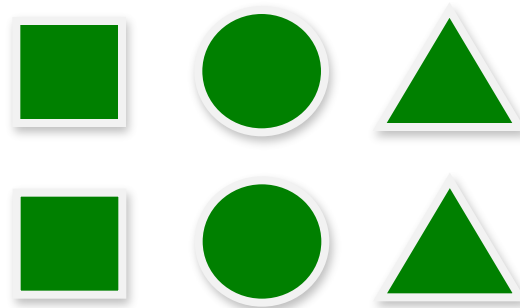
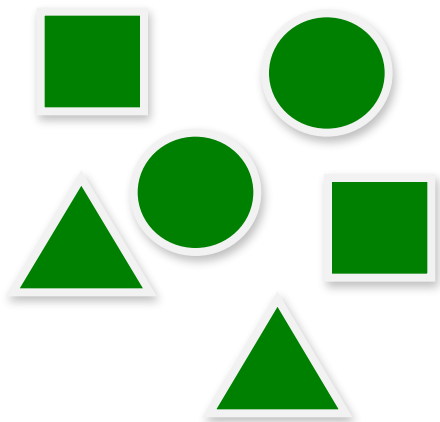
Основные формы наблюдения:

- 1. Отчетность** (*первичный учет*)
- 2. Специально организованное наблюдение**
(*перепись*)
- 3. Регистры** (*ЕГРПО – единый государственный регистр предприятий и организаций всех форм собственности*)

2) Статистическая сводка и группировка

Виды статистических группировок:

1. Типологическая

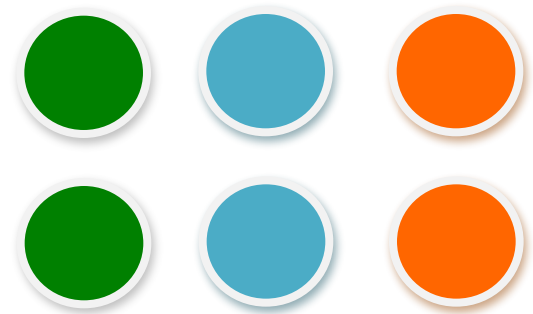
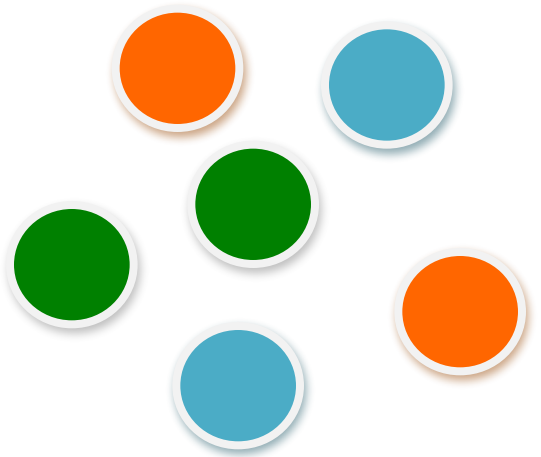


Разнородная совокупность

Однородные группы

Виды статистических группировок:

2. Структурная

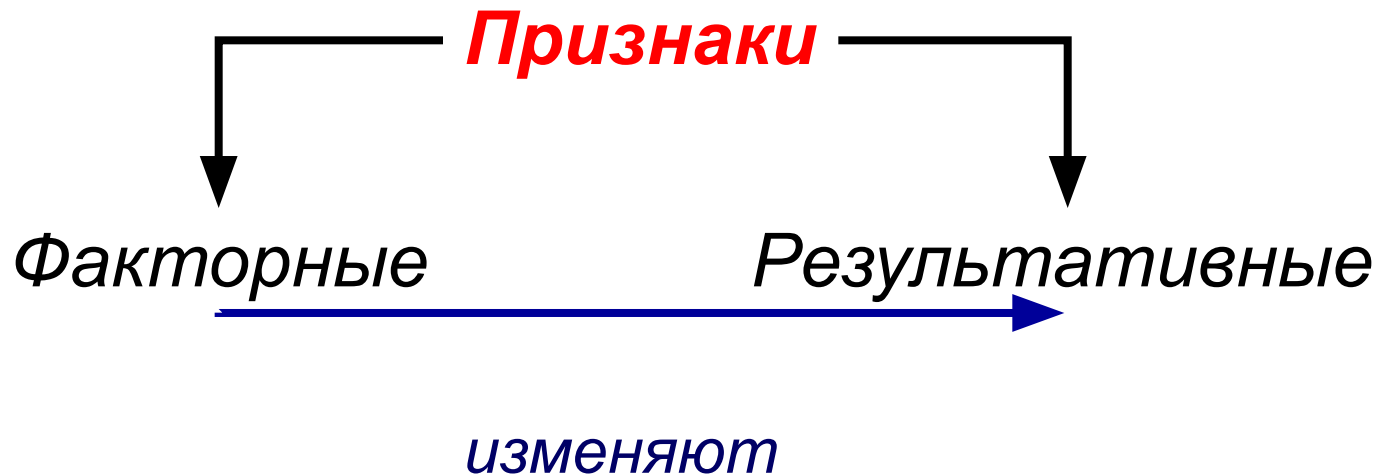


Однородная совокупность

Однородные группы

Виды статистических группировок:

3. Аналитическая



Принципы построения статистических группировок:

- 1 Выбор **группировочного признака**, по которому производится объединение единиц исследуемой совокупности в группы

Признаки:

- ❖ **количественные** (имеют числовое выражение)
- ❖ **атрибутивные** (отражают состояние единиц совокупности)

Принципы построения статистических группировок:

② Определить количество групп

✓ *Математический путь*

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

*Формула
Стерджесса*

n – число групп

N – число единиц совокупности

Принципы построения статистических группировок:

- ③ Определить интервалы группировки, т. е. значения варьирующего признака, лежащего в определенных границах

✓ Величина равного интервала:

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

x_{\max} ,
 x_{\min}

– max и min значения признака в совокупности

n – число групп

Правила округления

✓1 знак до запятой – округление до десятых:

$$0,88 \approx 0,9$$

$$1,585 \approx 1,6$$

$$4,71 \approx 4,7$$

✓2 знака до запятой – округление до целого числа:

$$15,985 \approx 16$$

✓3^х значное и т.д. – округление до ближайшего целого числа:

$$557 \approx 600$$

Принципы построения статистических группировок:

- ③ ✓ Величина неравного интервала:

Арифметическая прогрессия

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

a – константа, имеющая для прогрессивно
возрастающих интервалов знак $+$, а для прогрессивно
убывающих знак $-$.

Принципы построения статистических группировок:

- ③ ✓ Величина неравного интервала:

Геометрическая прогрессия

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

q – константа (для прогрессивно возрастающих интервалов $q > 1$, в другом случае – $q < 1$).

Пример:

Требуется произвести группировку с равными интервалами предприятий по стоимости основных фондов;

При этом:

- max значение признака=2040 млн.руб,*
- min значение=290 млн.руб.*
- совокупность включает 80 единиц.*

Решение:

- Согласно формуле Стерджесса совокупность должна быть разбита на 7 групп

- Размах вариации:

$$R=2040-290=1750 \text{ млн.руб}$$

- Величина интервала:

$$h=1750:7=250 \text{ млн.руб}$$

Группа	Интервалы	
	1-ый вариант (закрытые)	2-ой вариант (открытые)
1	<i>от 290 до 540</i>	<i>до 540</i>
2	<i>от 540 до 790</i>	<i>540-790</i>
3	<i>от 790 до 1040</i>	<i>790-1040</i>
4	<i>от 1040 до 1290</i>	<i>1040-1290</i>
5	<i>от 1290 до 1540</i>	<i>1290-1540</i>
6	<i>от 1540 до 1790</i>	<i>1540-1790</i>
7	<i>от 1790 до 2040</i>	<i>1790 и более (более 1790)</i>

Пример:

Необходимо построить группировку предприятий отрасли по показателю выручки от реализации продукции, который варьируется от 500 млн.руб до 4000 млн.руб.

Строить группировку с равными интервалами нецелесообразно, т.к. распределение числа предприятий по величине выручки является неравномерным, поэтому следует построить группировку с неравными интервалами.

Группа	Интервал
1	500-800
2	800-1300
3	1300-2000
4	2000-2900
5	2900-4000

*Величина каждого последующего интервала
больше предыдущего на 200 млн.руб, т.е.
увеличивается в арифметической
прогрессии*

Пример:

Пусть:

Совокупность надо разделить на группы по численности занятых.

- *max значение признака=2040 чел.,*
- *min значение=290 чел.*
- *совокупность включает 80 единиц.*

В основании группировки лежит дискретный (прерывный) признак.

Группа	Интервал	
	1-ый вариант	2-ой вариант
1	<i>290-540</i>	<i>До 540</i>
2	<i>541-790</i>	<i>541-790</i>
3	<i>791-1040</i>	<i>791-1040</i>
4	<i>1041-1290</i>	<i>1041-1290</i>
5	<i>1291-1540</i>	<i>1291-1540</i>
6	<i>1541-1790</i>	<i>1541-1790</i>
7	<i>1791-2040</i>	<i>1791 и более</i>

Ряд распределения

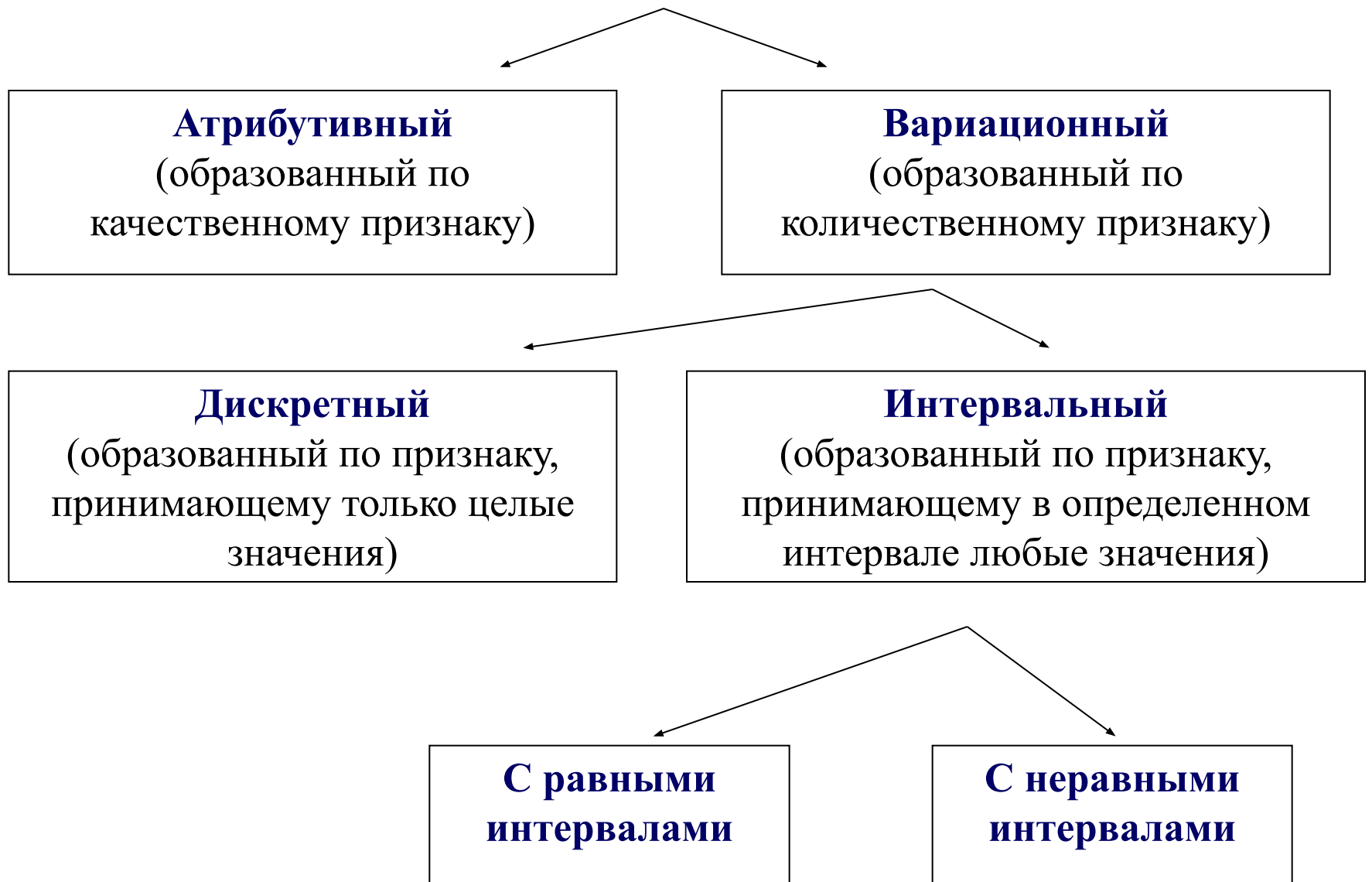
*Разновидностью структурной группировки является **ряд распределения** (правила построения аналогичны правилам построениям группировок)*

Состоит из 2-х элементов:

- **Вариантов** (групп по выделенному признаку)
- **Частот** (численности групп)

<i>Вариант, x_i</i>	<i>Частота, f_i</i>
x_1	f_1
x_2	f_2
x_3	f_3
\dots	\dots
x_n	f_n
<i>Итого</i>	$\sum_i f_i$ (<i>или N</i>)

Ряд распределения



Плотность распределения

Чтобы частоты можно было сравнивать, исчисляют **плотность распределения**.

Абсолютная плотность распределения –
это частота, приходящаяся на единицу длины
интервала

$$\frac{f_i}{h_i}$$

Относительная плотность распределения –
частота, приходящаяся на единицу длины
интервала

$$\frac{w_i}{h_i}$$

№ п/п	Группы кредитных организаций по уставному капиталу, млрд. руб.	Ширина интервала, млрд.руб.	Число кредитных организаций (частота)	Плотность распределения
		1	2	3 (гр.2 : гр.1)
1	до 3*	2	150	75
2	3 - 10	7	254	36,3
3	10 - 30	20	316	15,8
4	30 - 60	30	256	8,5
5	60 - 150	90	144	1,6
6	150 - 300	150	90	0,6
7	300 - 600	300	112	0,37
	Итого	-	1322	-

***** *Неравные интервалы не позволяют сравнивать частоты в разных интервалах*

Число случаев ниже или выше определенного уровня показателей: **ряд накопленных частот**
(кумулятивный ряд)

№ п/п	Группы семей по числу детей, x	число семей		Накоплен ные частоты, S
		тыс. (частоты), f	% к итогу (частости), w	
1	0	6	5,9	6
2	1	28	27,5	34
3	2	22	21,6	56
4	3	20	19,6	76
5	4	13	12,7	89
6	5	8	7,8	96
7	6 и более	5	4,9	102
Итого		102	100	

Результат:

- *Меньше 1 ребенка имеют 6 тыс. семей*
- *Меньше 2 детей имеют $6+28=34$ тыс. семей*
- *102 тыс. семей = объему совокупности*

Аналогично определяются **накопленные частоты**
(показывают долю членов совокупности, у которых интересующий нас признак меньше данного значения)

Графическое изображение рядов распределения

Распределение работников строительной фирмы по уровню дохода

Группы работников по уровню дохода, тыс. руб.	Число работников, чел.	Удельный вес, % к итогу
до 20	60	52,2
20 – 25	30	26,1
25 – 30	15	13,0
30 и более	10	8,7
Итого		

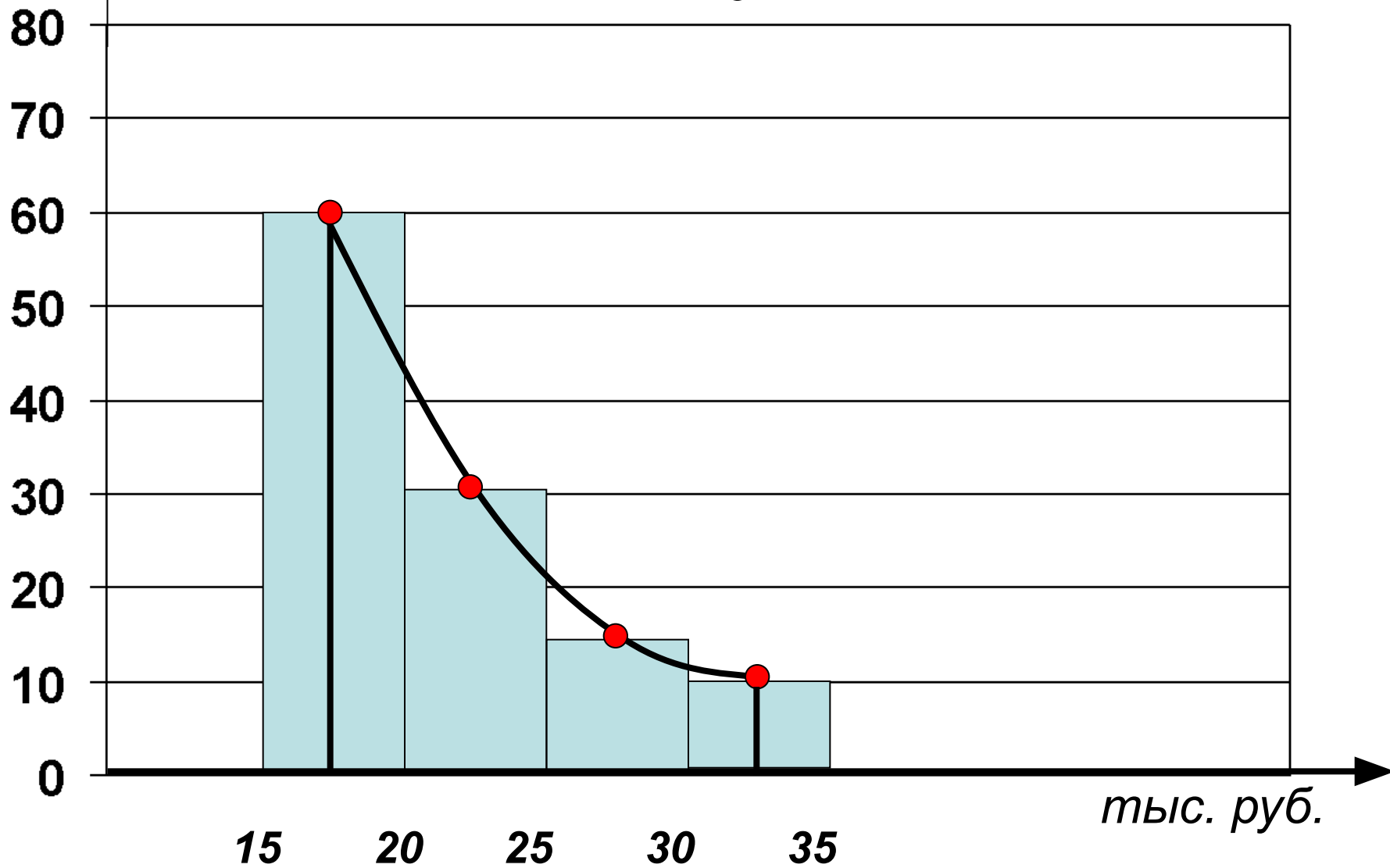
Число рабочих,
чел.



- гистограмма



- полигон



Кумулятивная кривая



кумулята

- *распределение «не меньше чем»*
- *ось абсцисс – варианты ряда*
- *ось ординат – накопленные частоты*

огива

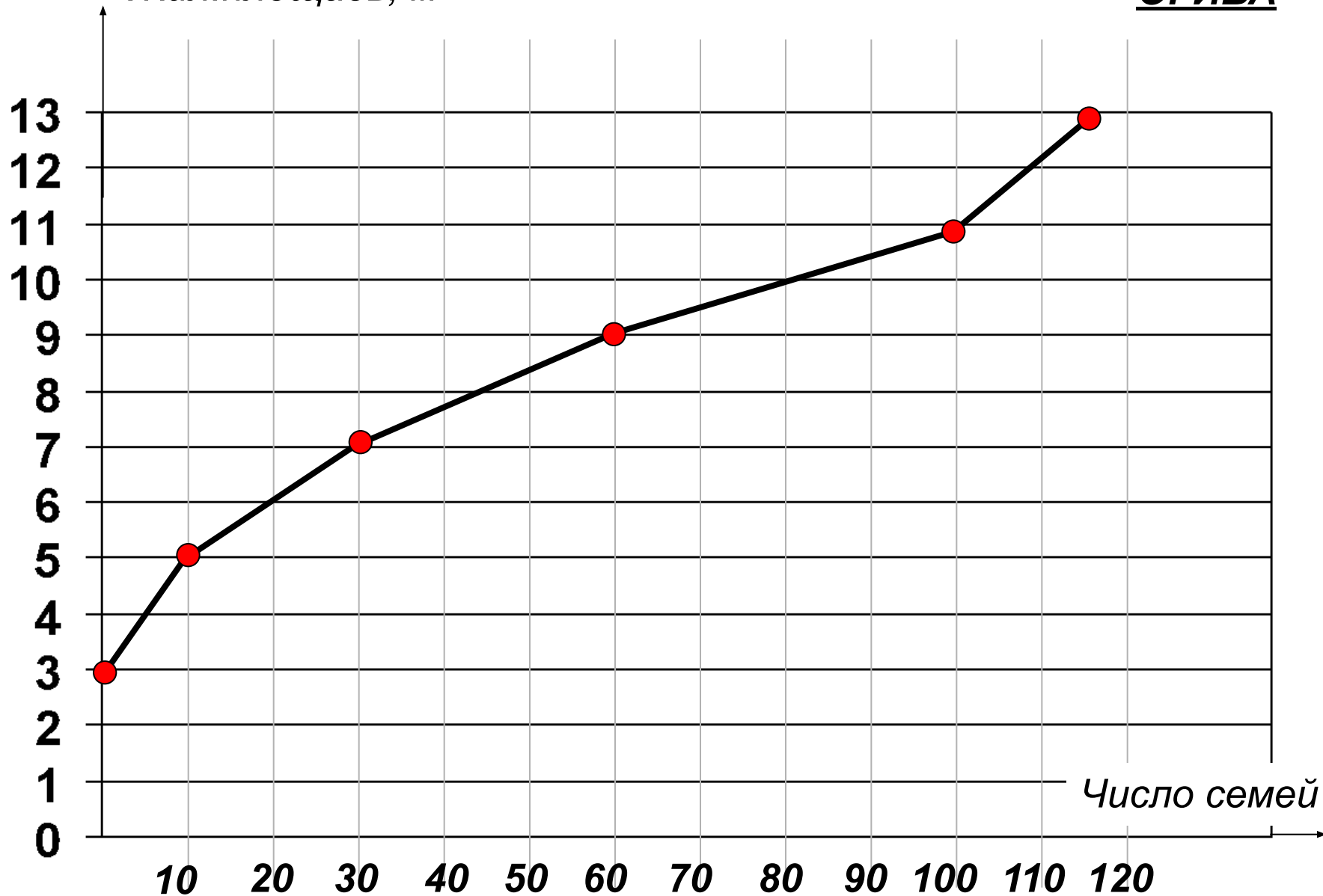
- *распределение «больше чем»*
- *абсциссы – накопленные частоты*
- *ординаты – варианты ряда*

*Распределение семей по размеру жилищной площади,
приходящейся на 1 человека*

Группы семей по размеру жилищной площади, приходящейся на 1 человека, м² <i>x</i>	Число семей с данным размером жил. площади <i>f</i>	Накопленное число семей <i>S</i>
3-5	10	10
5-7	20	30
7-9	30	60
9-11	40	100
11-13	15	115
Итого	115	

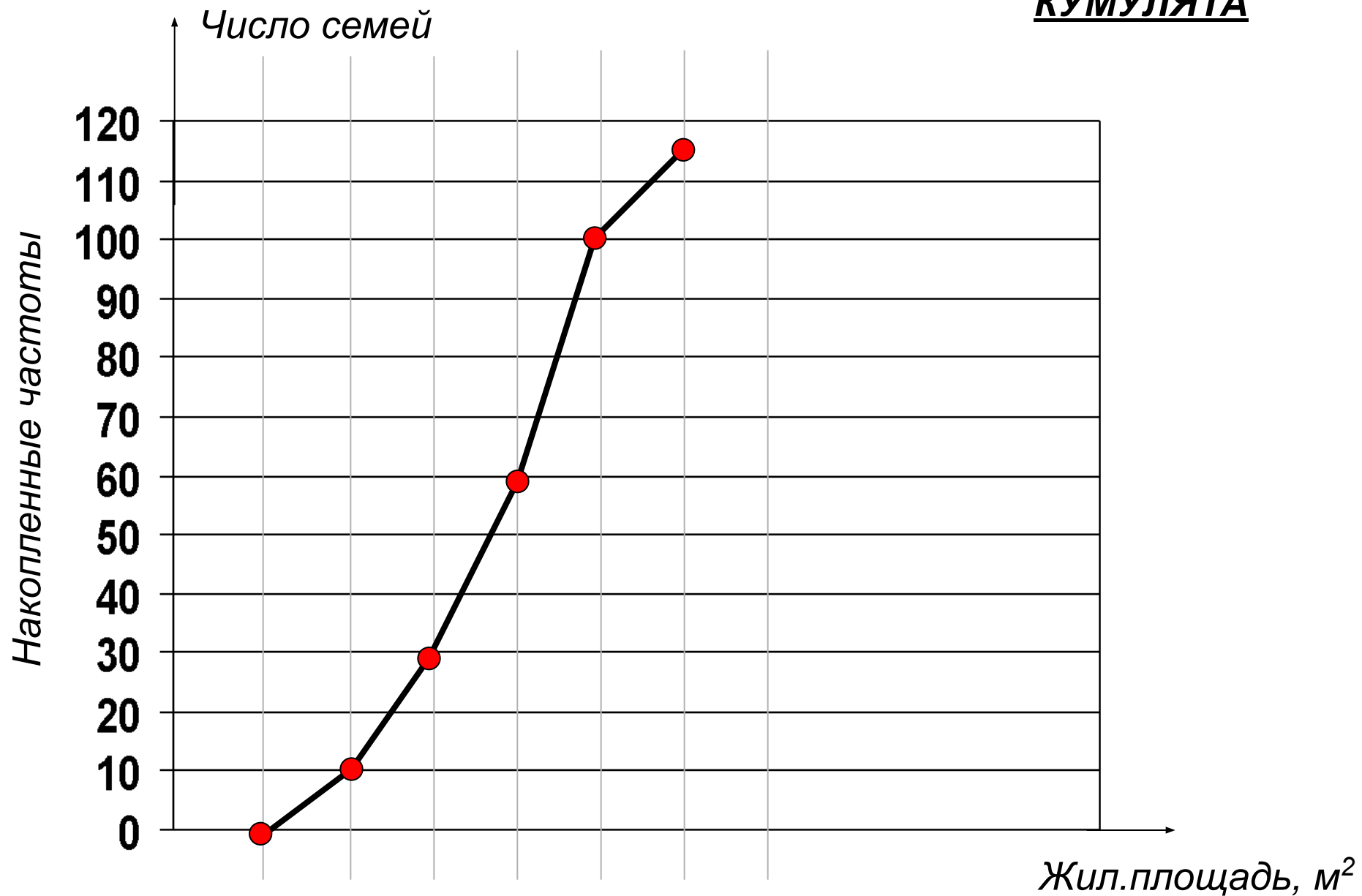
Жил.площадь, м²

ОГИВА



Накопленные частоты

КУМУЛЯТА



Метод вторичной группировки

Пример:

(Способ объединения первоначальных интервалов)

Группировка кредитов коммерческих банков по сроку выдачи в ноябре:

№	Группы кредитов по сроку выдачи, мес.	Число заключенных договоров, % от их общего кол-ва	Сумма выданных кредитов, % от общей суммы
1	1-3	87,05	66,87
2	3-6	10,43	24,86
3	6-12	1,80	8,17
4	Более 12	0,72	0,10
Итого		100,000	100,000

Пример:

(Способ объединения первоначальных интервалов)

Группировка кредитов коммерческих банков по сроку выдачи в декабре

№	Группы кредитов по сроку выдачи, мес.	Число заключенных договоров, % от их общего кол-ва	Сумма выданных кредитов, % от общей суммы
1	Краткосрочные (1-6)	86,54	97,91
2	Среднесрочные (6-12)	1,92	1,70
3	Долгосрочные (более 12)	11,54	0,39
Итого		100,000	100,000

Пример:

(Способ объединения первоначальных интервалов)

Группировка кредитов коммерческих банков по сроку выдачи в ноябре - декабре

№	Группы кредитов по сроку выдачи, мес.	Число заключенных договоров, % от их общего кол-ва		Сумма выданных кредитов, % от общей суммы	
		ноябрь	декабрь	ноябрь	декабрь
1	Краткосрочные(1-6)	97,48	86,54	91,73	97,91
2	Среднесрочные(6-12)	1,80	1,92	8,17	1,70
3	Долгосрочные (более 12)	0,72	11,54	0,10	0,39
Итого		100,000	100,000	100,000	100,000

Пример: *(Способ долевой перегруппировки)*

№	Группы 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 бол.	9
	Итого	200

Необходимо провести перегруппировку данных, образовав новые группы с интервалами до 500, 500-1000, 1000-2000, 2000-3000, 3000 и выше

1) В 1-ую новую группу войдет полностью 1 группа и часть 2-ой

- Необходимо от интервала 2-ой группы взять 100 тыс. руб.
- Величина интервала 2-ой группы составляет 600 тыс. руб.
- Необходимо взять от 2-ой группы $1/6$ ($100:600$)



2) 2-ую новую группу образуют контракты 2-ой группы (за вычетом отнесенных к 1-й)

3) 3-ю новую группу составляют все контракты 3-ей группы и часть 4-ой

- **Ширина интервала 4-ой группы = 1200 тыс. руб.**
- **Необходимо взять от 4-ой группы $1/6$ или 200 тыс. руб.**



4) 4-ую новую группу образуют контракты 4-ой старой за вычетом, отнесенных к 3-ей

5) 5-ую – все контракты 5-ой и 6-ой групп

Результаты перегруппировки:

№	Группы контрактов по величине прибыли, тыс. руб.	Число контрактов, ед.
1	До 500	$16+20*1/6=16+3=19$
2	500-1000	$20-3=17$
3	1000-2000	$44+74*1/6=44+12=56$
4	2000-3000	$74-12=62$
5	3000 и более	$37+9=46$
Итого		200