

### Задача с решением по правовой статистике

ЗАДАНИЕ.

*Дан ряд распределения территорий по количеству преступлений, совершаемых за 1 год:*

Интервалы значений	Число регионов	
	В абсолютных единицах	В относительных единицах, %
198-1198	34	
1198-2198	12	
2198-3198	5	
3198-4198	8	
4198-5198	13	
Всего		

*Необходимо:*

- 1. Заполнить недостающие данные в таблице.*
- 2. Определить величину интервала, максимальное и минимальное значение признака.*
- 3. Построить гистограмму и полигон распределения.*
- 4. Рассчитать статистические характеристики ряда распределения: моду, медиану, среднее значение признака, вид распределения, среднее линейное отклонение, дисперсию, среднее квадратическое отклонение, размах вариации, коэффициент вариации.*

РЕШЕНИЕ.

1. Заполним недостающие данные в таблице:

Интервалы значений	Число регионов	
	В абсолютных единицах	В относительных единицах, %
198-1198	34	47,22
1198-2198	12	16,67
2198-3198	5	6,94
3198-4198	8	11,11
4198-5198	13	18,06

Всего	72	100,00
-------	----	--------

1. Величина интервала – разница между максимальным и минимальным значением в интервале:

$$h = 1198 - 198 = 1000.$$

3. Гистограмма распределения:



4. Полигон распределения (по серединам интервалов):



### 5. Вычисляем отдельные показатели.

Сначала определим ряд сумм:

$x_i$  – середина интервала;

$f_i$  – частота интервала.

$x_i$	$f_i$	$x_i \cdot f_i$	$ x_i - x_{sr}  \cdot f_i$	$(x_i - x_{sr})^2 \cdot f_i$
698	34	23732	46277,74	62989094,7
1698	12	20376	4333,32	1564805,19
2698	5	13490	3194,45	2040902,16
3698	8	29584	13111,12	21487683,5
4698	13	61074	34305,57	90528625,6
	72	148256	101222,2	178611111

Мода:

$$M_o = x_{M_o} + h \frac{f_{M_o} - f_{M_o-1}}{(f_{M_o} - f_{M_o-1}) + (f_{M_o} - f_{M_o+1})} = 198 + 1000 \frac{34 - 0}{(34 - 0) + (34 - 12)} = 805.14$$

$x_{M_o}$  – нижняя граница модального интервала;

$h$  – интервал группирования, у нас равен 2;

$f_{M_o}, f_{M_o-1}, f_{M_o+1}$  – частоты модального, передмодального и послемодального интервалов.

Мода определяет величину наиболее вероятного значения количества преступлений в регионах.

Медиана:

$$M_e = x_{M_e} + i \frac{0,5 \sum f - S_{M_e-1}}{f_{M_e}} = 1198 + 1000 \frac{0,5 * 72 - 34}{12} = 1364.67$$

$x_{M_e}$  – нижняя граница медианного интервала;

$S_{M_e-1}$  – суммарная частота передмедианного интервала;

$f_{M_e}$  – частота медианного интервала.

Значение медианы в данном случае характеризует середину распределения регионов по количеству преступлений.

Размах вариации – разница между максимальным и минимальным значением вариант:

$$R = x_{\max} - x_{\min} = 5198 - 198 = 5000.$$

Остальные показатели вариации рассчитаем по формулам:

$$\bar{x} = \frac{\sum x_i f_i}{\sum f_i} = 148256/72 = 2059,11.$$

$$\text{среднее линейное отклонение} - d = \frac{\sum |x_i - \bar{x}| f_i}{\sum f_i} = 101222,2/72 = 1405,86;$$

$$\text{дисперсия} - \sigma^2 = \frac{\sum (x_i - \bar{x})^2 f_i}{\sum f_i} = 178611111/72 = 2480709,875;$$

среднее квадратическое отклонение – корень с дисперсии – 1575,03;

$$\text{коэффициент вариации} - V = \frac{\sigma}{\bar{x}} * 100 = 1575,03/2059,11 * 100 = 76,49\%,$$

Вывод: распределение регионов по количеству преступлений – неоднородное, а полученное значение среднего – ненадежное, поскольку коэффициент вариации превышает 33%.