

Показатели вариации

В социально-экономическом анализе важно знать не только среднее (или срединное) значение признака, но и насколько равномерно распределены эти значения относительно среднего значения, а так же знать количественную меру степени этой неравномерности.

Вариация (variatio, латинск. - различие, изменение, колеблемость) -

- количественное различие значений одного и того же признака у отдельных единиц совокупности
- позволяет установить зависимость между изменением, которое происходит в исследуемом признаке, и теми факторами, которые вызывают данное изменение
- по степени вариации можно судить о границах вариации признака, однородности совокупности по данному признаку, типичности средней, взаимосвязи факторов, определяющих вариацию

Пример

Пусть заданы два ряда.

Ряд I: 1, 2, 3, 4, 5, 6, 6, 7, 8, 9, 10, 11

Ряд II: 4, 5, 5, 5, 6, 6, 6, 6, 7, 7, 7, 8

Рассчитаем для этих рядов среднюю арифметическую, моду и медиану.

Ряд I. $\bar{x} = 6$, $Me = 6$, $Mo = 6$, $n = 12$.

Ряд II. $\bar{x} = 6$, $Me = 6$, $Mo = 6$, $n = 12$.

Закономерность распределения признака в совокупности описывается:

- частотными показателями;
- показателями распределения – структурные средние;
- показателями степени вариации;
- показателями формы распределения.

Частотные показатели вариации

- абсолютная численность i -той группы – частота f_i

$$\sum_1^m f_i = n$$

- относительная частота – частость d_i

$$\sum_1^m d_i = 1 \quad d_i = \frac{f_i}{\sum_{i=1}^k f_i}$$

- кумулятивная (накопленная) частота S_i (частость S_d , кумулята) характеризует объем совокупности со значениями вариантов, не превышающих X_i .

$$S_1=f_1, S_2=f_1+f_2, S_3=f_1+f_2+f_3;$$

- плотность частоты (частости) представляет собой частоту, приходящуюся на единицу интервала,

$$q_i=f_i/h_i \text{ или } q_i=d_i/h_i$$

где h_i – величина i -того интервала.

Mo и Me

- В симметричных распределениях средняя арифметическая, мода и медиана совпадают

$$\bar{X} = M_e = M_0$$

- Если это равенство нарушается — распределение асимметрично.

- Простейшим показателем асимметрии является разность , $\bar{x} - M_0$

которая в случае правосторонней асимметрии положительна, а при левосторонней — отрицательна.

порядковый номер M_e

- при нечетном числе единиц к сумме всех частот прибавляется единица и все делится на два.
- при четном числе единиц медиана = значению признака у единицы совокупности, порядковый номер который определяется по общей сумме частот, деленной на два.
- В дискретном ряду распределения медиана находится непосредственно по накопленной частоте, соответствующей номеру медианы

$$N_{M_e} = \frac{n + 1}{2};$$

Изменение вариации признака в совокупности осуществляется с помощью **абсолютных и относительных показателей**

Абсолютные показатели вариации

включают:

- размах вариации
- среднее линейное отклонение
- дисперсию
- среднее квадратическое отклонение

Размах вариации -

- это разность между максимальным и минимальным значениями признака

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

- показывает пределы, в которых изменяется величина признака в изучаемой совокупности
- Размах вариации в первом ряду равен 10.
Размах вариации во втором ряду равен 4.

средние показатели вариации, основанные на учете отклонений от центра ряда распределения - средней арифметической

- За отклонение от средней принимается разность

$$(x_i - \bar{x})$$

- Для превращения в нуль суммы отклонений вариантов признака от средней (нулевое свойство средней) приходится либо не учитывать знаки отклонения, то есть брать эту сумму по модулю

$$\sum |x_i - \bar{x}|$$

$$\sum |x_i - \bar{x}|^2$$

- либо возводить значения отклонений в квадрат

Среднее линейное отклонение

• это средняя арифметическая из абсолютных отклонений отдельных значений признака от средней; **показывает, насколько в среднем колеблется величина признака у единиц исследуемой совокупности.**

• **Среднее линейное отклонение простое:**

$$\bar{d} = \frac{\sum |x_i - \bar{x}|}{n}$$

• **Среднее линейное отклонение взвешенное:**

$$\bar{d} = \frac{\sum |x_i - \bar{x}| * f}{\sum f}$$

Среднее квадратическое отклонение показывает во сколько раз в среднем колеблется величина признака совокупности; является мерой надежности средней величины: чем оно меньше, тем точнее средняя арифметическая.

- **Средняя квадратическая простая**

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

- **Средняя квадратическая взвешенная**

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

Опыт работы у пяти претендентов на работе составляет: 2,3,4,7 и 9 лет.

Х_{ср.}=5 лет

$$\bar{d} = \frac{|2-5| + |3-5| + |4-5| + |7-5| + |9-5|}{5} = \frac{12}{5} = 2,4 \text{ (года).}$$

В нашем примере

$$\sigma = \sqrt{\frac{(2-5)^2 + (3-5)^2 + (4-5)^2 + (7-5)^2 + (9-5)^2}{5}} = \sqrt{6,80} = 2,607 \sim$$

2,61 (года).

Дисперсия - средний квадрат отклонений индивидуальных значений признака от его средней величины (1- дисперсия взвешенная, 2- дисперсия простая)

$$\sigma_{вз}^2 = \frac{\sum (\bar{x}_i - \bar{x})^2 f}{\sum f} ; \quad \sigma_{пр}^2 = \frac{\sum (\bar{x}_i - \bar{x})^2}{n}$$

$$\sigma^2 = \overline{x^2} - (\bar{x})^2$$

средний квадрат отклонений равен средней из квадратов значений признака минус квадрат средней

- Если из генеральной совокупности сделать несколько выборок и каждый раз при этом определять среднее значение признака, то возникает задача **оценки колеблемости средних**. Оценить дисперсию *среднего значения* можно и на основе всего одного выборочного наблюдения по формуле

$$\sigma^2 (X) = \sigma^2 / n$$

- где n - объем выборки; σ^2 - дисперсия признака, рассчитанная по данным выборки.

Для оценки достоверности результатов выборочного наблюдения применяется показатель средней ошибки выборки

- Величина

$$\mu = \sqrt{\sigma^2(X)} = \sqrt{\sigma^2 / n}$$

носит название *средней ошибки выборки* и является характеристикой отклонения выборочного среднего значения признака X от его истинной средней величины

Свойства дисперсии:

1. Дисперсия постоянной величины равна 0.
2. Если все значения вариантов признака X уменьшить на постоянную величину A , то дисперсия не изменится.
3. Если все значения вариантов X уменьшить в K раз, то дисперсия уменьшится в K^2 раз.
4. При малом числе наблюдений (< 30):

$$\sigma^2 = \frac{n}{n-1} \left(\overline{X^2} - (\overline{X})^2 \right)$$

Сравнение вариации нескольких совокупностей по одному и тому же признаку, а тем более по различным признакам с помощью абсолютных показателей не представляется возможным. В этих случаях для сравнительной оценки степени различия строят относительные показатели вариации. Они вычисляются как отношения абсолютных показателей вариации к средней

Относительные показатели вариации -

<p><i>Коэффициент осцилляции</i> - отношение размаха вариации к средней величине признака</p>	$K_0 = \frac{R}{\bar{x}} \cdot 100\%$
<p><i>Линейный коэффициент вариации</i> - отношение среднего линейного отклонения к средней величине признака</p>	$K_L = \frac{\bar{L}}{\bar{x}} \cdot 100\%$
<p><i>Коэффициент вариации</i> - отношение среднего квадратического отклонения к средней величине признака</p>	$V = \frac{\sigma}{\bar{x}} \cdot 100\%$

• Коэффициент осцилляции – это отношение размаха вариации к средней, в процентах. Отражает относительную колеблемость крайних значений признака вокруг средней.

• Линейный коэффициент вариации характеризует долю усредненного значения абсолютного отклонения от средней величины.

Коэффициент вариации позволяет судить об однородности совокупности:

- $< 17\%$ – абсолютно однородная;
- $17\text{--}33\%$ – достаточно однородная;
- $35\text{--}40\%$ – недостаточно однородная;
- $40\text{--}60\%$ – это говорит о большой колеблемости совокупности.

Рассчитать коэффициент вариации на основе среднего квадратического отклонения для следующего примера:

- Расход сырья на единицу продукции составил (кг): по одной технологии $X1_{ср.}=10$ при $\sigma 1=4$; по другой- $X2_{ср.}=6$ при $\sigma 2=3$.
- Какая вариация расхода сырья интенсивнее?

$$(V_1 = \frac{4}{10} = 0,40 \text{ или } 40,0\%, V_2 = \frac{3}{6} = 0,50 \text{ или } 50,0\%).$$

Пример 1. $X_{\text{ср}}=4$, $X_{\text{ср.вз}}=3.98$

Значение (X)	Частота (f)	Кумулята (S)	$X \cdot f$	$ X - \bar{X} $	$5 \cdot 2$	$5 \cdot 5$	$7 \cdot 2$
1	2	3	4	5	6	7	8
2	5						
3	13						
4	16						
5	10						
6	6						
ИТОГО	50						

Пример 1

Значение (X)	Частота (f)	Кумулята (S)	$X * f$	$ X - \bar{X} $	5*2	5*5	7*2
1	2	3	4	5	6	7	8
2	5	5					
3	13	18					
4	16	34					
5	10	44					
6	6	50					
ИТОГО	50						

Пример 1. $X_{\text{ср}}=4$, $X_{\text{ср.вз}}=3.98$

Значение (X)	Частота (f)	Кумулянта (S)	$X \cdot f$	$ X - \bar{X} $	$5 \cdot 2$	$5 \cdot 5$	$7 \cdot 2$
1	2	3	4	5	6	7	8
2	5	5	10				
3	13	18	39				
4	16	34	64				
5	10	44	50				
6	6	50	36				
ИТОГО	50		199				

Значение (X)	Частота (f)	Кумулята (S)	$X \cdot f$	$ X - \bar{X} $	$5 \cdot 2$	$5 \cdot 5$	$7 \cdot 2$
1	2	3	4	5	6	7	8
2	5	5	10	1,98			
3	13	18	39	0,98			
4	16	34	64	0,02			
5	10	44	50	1,02			
6	6	50	36	2,02			
ИТОГО	50		199				

Значение (X)	Частота (f)	Кумулята (S)	$X \cdot f$	$ X - \bar{X} $	$5 \cdot 2$	$5 \cdot 5$	$7 \cdot 2$
1	2	3	4	5	6	7	8
2	5	5	10	1,98	9,9		
3	13	18	39	0,98	12,7 4		
4	16	34	64	0,02	0,32		
5	10	44	50	1,02	10,2		
6	6	50	36	2,02	12,1 2		
ИТОГО	50		199		45,2 8		

Значение (X)	Частота (f)	Кумулята (S)	X*f	$ X - \bar{X} $	5*2	5*5	7*2
1	2	3	4	5	6	7	8
2	5	5	10	1,98	9,9	3,92	
3	13	18	39	0,98	12,74	0,96	
4	16	34	64	0,02	0,32	0,00	
5	10	44	50	1,02	10,2	1,04	
6	6	50	36	2,02	12,12	4,08	
Итого	50		199		45,28	10,0 0	

Значение (X)	Частота (f)	Кумулята (S)	$X \cdot f$	$ X - \bar{X} $	5*2	5*5	7*2
1	2	3	4	5	6	7	8
2	5	5	10	1,98	9,9	3,92	19,60
3	13	18	39	0,98	12,74	0,96	12,49
4	16	34	64	0,02	0,32	0,00	0,01
5	10	44	50	1,02	10,2	1,04	10,40
6	6	50	36	2,02	12,12	4,08	24,48
ИТОГО	50		199		45,28	10,00	66,98

Показатели вариации (пример 1)

max	6
min	2
n	50
среднее	4
средневзвешенное	3,98
Мода	4
Номер медианы	25,5
Медиана	4
Размах вариации	4
Среднее линейное отклонение	0,91
Дисперсия	1,34
Среднее квадратическое отклонение	1,16
Коэффициент осциляции	101%
Линейная вариация	23%

Пример 2

Значение (X)	Частота (f)	Кумулята (S)	Средина интервала	$4*f$	$!X-X_{ср}$!	$6*2$	$6*6$	$8*2$
1	2	3	4	5	6	7	8	9
5-7	9							
7-9	16							
9-11	11							
11-13	8							
13-15	6							
	50							

Пример 2

Значение (X)	Частота (f)	Кумулята (S)	Средина интервала	$4 \cdot f$	$!X - X_{\text{ср}}!$	$6 \cdot 2$	$6 \cdot 6$	$8 \cdot 2$
1	2	3	4	5	6	7	8	9
5-7	9	9						
7-9	16	25						
9-11	11	36						
11-13	8	44						
13-15	6	50						
	50							

Пример 2

Значение (X)	Частота (f)	Кумулята (S)	Средина интервала	$4*f$	$!X-X_{ср}!$	$6*2$	$6*6$	$8*2$
1	2	3	4	5	6	7	8	9
5-7	9	9	6					
7-9	16	25	8					
9-11	11	36	10					
11-13	8	44	12					
13-15	6	50	14					
	50							

Пример 2

Значение (X)	Частота (f)	Кумулята (S)	Средина интервала	$4 \cdot f$	$!X - X_{\text{ср}}!$	$6 \cdot 2$	$6 \cdot 6$	$8 \cdot 2$
1	2	3	4	5	6	7	8	9
5-7	9	9	6	54				
7-9	16	25	8	128				
9-11	11	36	10	110				
11-13	8	44	12	96				
13-15	6	50	14	84				
	50			472				

Пример 2

Значение (X)	Частота (f)	Кумулята (S)	Средина интервала	$4 \cdot f$	$!X - X_{\text{ср}}!$	$6 \cdot 2$	$6 \cdot 6$	$8 \cdot 2$
1	2	3	4	5	6	7	8	9
5-7	9	9	6	54	3,44			
7-9	16	25	8	128	1,44			
9-11	11	36	10	110	0,56			
11-13	8	44	12	96	2,56			
13-15	6	50	14	84	4,56			
	50			472				

Пример 2

Значение (X)	Частота (f)	Кумулята (S)	Средина интервала	$4 \cdot f$	$!X - X_{\text{ср}}!$	$6 \cdot 2$	$6 \cdot 6$	$8 \cdot 2$
1	2	3	4	5	6	7	8	9
5-7	9	9	6	54	3,44	30,96		
7-9	16	25	8	128	1,44	23,04		
9-11	11	36	10	110	0,56	6,16		
11-13	8	44	12	96	2,56	20,48		
13-15	6	50	14	84	4,56	27,36		
	50			472		108		

Пример 2

Значение (X)	Частота (f)	Кумулята (S)	Средина интервала	$4 \cdot f$	$!X - X_{\text{ср}}!$	$6 \cdot 2$	$6 \cdot 6$	$8 \cdot 2$
1	2	3	4	5	6	7	8	9
5-7	9	9	6	54	3,44	30,96	11,83	
7-9	16	25	8	128	1,44	23,04	2,07	
9-11	11	36	10	110	0,56	6,16	0,31	
11-13	8	44	12	96	2,56	20,48	6,55	
13-15	6	50	14	84	4,56	27,36	20,79	
	50			472		108	41,57	

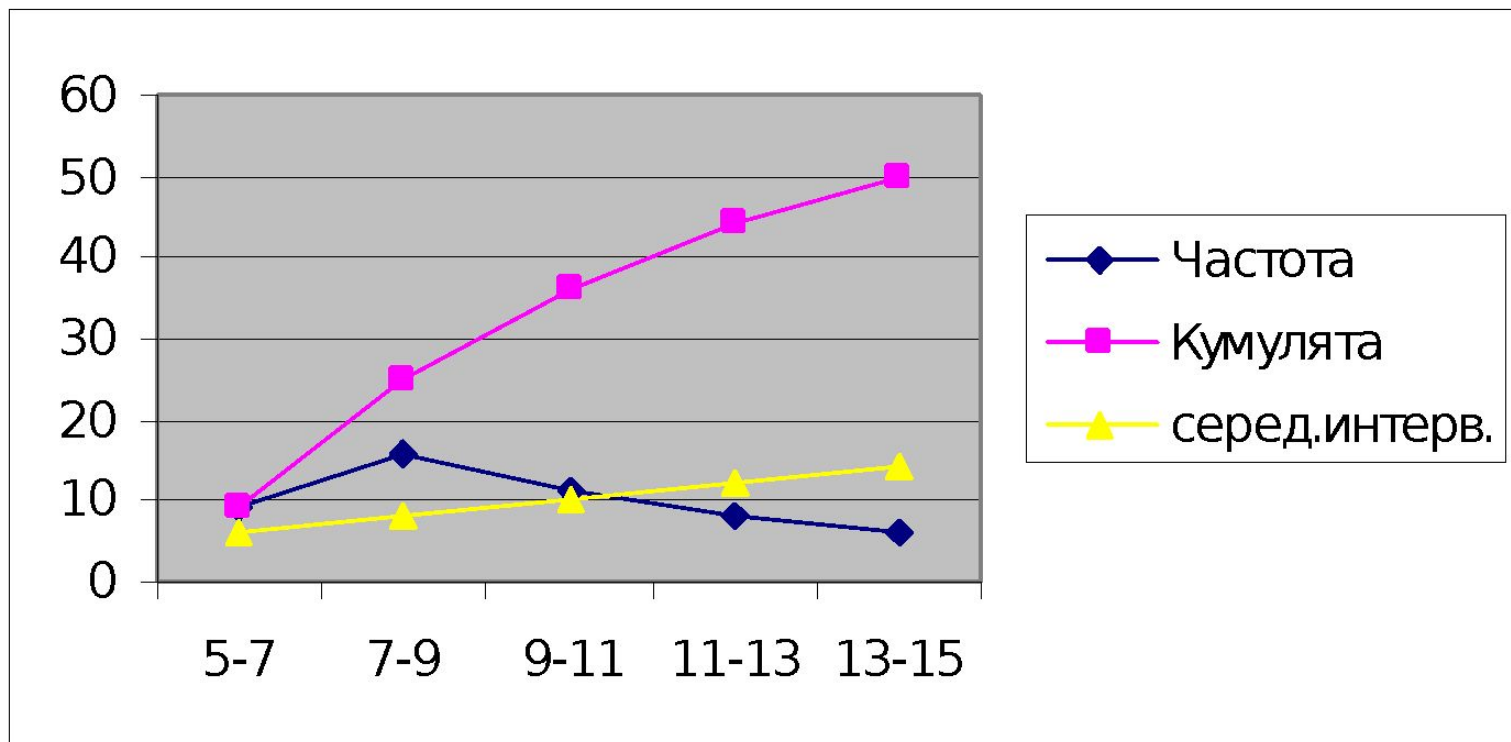
Пример 2

Значение (X)	Частота (f)	Кумулята (S)	Средина интервала	$4*f$	$!X-X_{cp}!$	$6*2$	$6*6$	$8*2$
1	2	3	4	5	6	7	8	9
5-7	9	9	6	54	3,44	30,96	11,83	106,50
7-9	16	25	8	128	1,44	23,04	2,07	33,18
9-11	11	36	10	110	0,56	6,16	0,31	3,45
11-13	8	44	12	96	2,56	20,48	6,55	52,43
13-15	6	50	14	84	4,56	27,36	20,79	124,76
	50			472		108	41,57	320,32

Показатели вариации (пример 2)

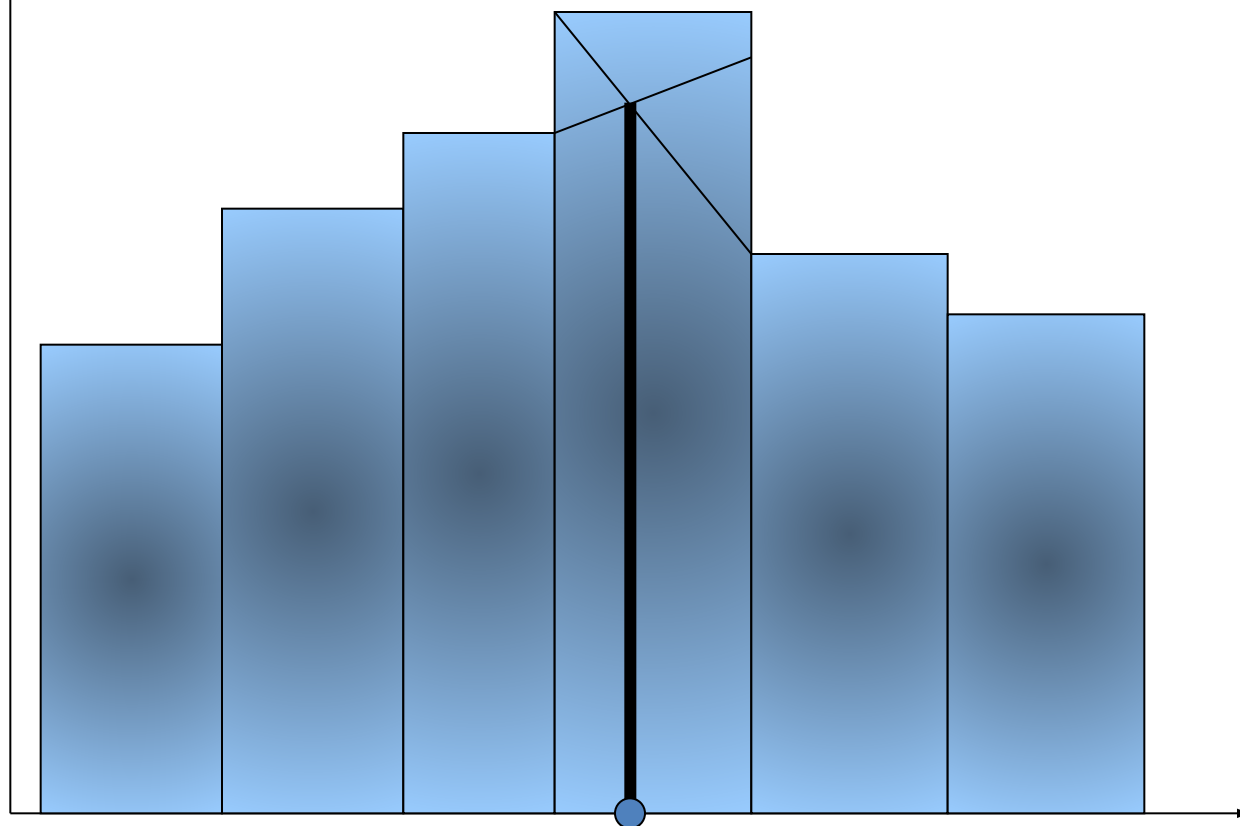
max	15
min	5
n	50
среднее	9,44
средневзвешенное	9,44
Мода	8,17
Номер медианы	25,50
Медиана	9,00
Размах вариации	9,90
Среднее линейное отклонение	2,16
Дисперсия	6,41
Среднее квадратическое отклонение	2,53
Коэффициент осциляции	1,05
Линейная вариация	0,23
Показатель колеблемости	0,27
Средняя ошибка выборки	0,36

Графики



Графическое определение моды

Частота (f)



Гистограмма

Признак (X)

Графическое определение моды

