Числовые характеристики вариационного ряда.

Средней арифметической вариационного ряда называется сумма произведений всех вариантов на соответствующие частоты, деленная на сумму частот: $\bar{x} = \frac{\sum_{i=1}^{m} x_i n_i}{n}$

где x_i – варианты дискретного вариационного ряда или середины интервалов интервального вариационного ряда; n_i – соответствующие им частоты; m – число неповторяющихся вариантов или число интервалов; $n = \sum_{i=1}^{m} n_i$

Очевидно, что $\bar{x} = \sum_{i=1}^{m} x_i w_i$

где $w_i = n_i/n$ – частости вариантов или интервалов.

Пример. Найти среднюю выработку рабочих:

i	Выработка в отчетном году в процентах к предыдущему x	Частота (количество рабочих) n_i	Частость (доля рабочих) $w_i = \frac{n_i}{n}$	Накопленная частота $n_i^{\text{íàê}}$	Накопленная частость $w_i^{\text{íàê}} = \frac{n_i^{\text{íàê}}}{n}$
1	94-100	3	0,03	3	0,03
2	100-106	7	0,07	10	0,1
3	106-112	11	0,11	21	0,21
4	112-118	20	0,2	41	0,41
5	118-124	28	0,28	69	0,69
6	124-130	19	0,19	88	0,88
7	130-136	10	0,1	98	0,98
8	136-142	2	0,02	100	1
$\sum$		100	1	-	-

$$\bar{x} = \frac{97 \cdot 3 + 103 \cdot 7 + ... + 133 \cdot 10 + 139 \cdot 2}{100} = 119\%.$$

Для несгруппированного ряда все частоты n_i=1 (i=1,2…,n), а $\bar{x} = \frac{\sum_{i=1}^{n} x_i}{n}$ есть "невзвешенная" средняя арифметическая.

Основные свойства средней арифметической.

Свойства средней арифметической аналогичны свойствам математического ожидания.

1. Средняя арифметическая постоянной равна самой постоянной.
2. Если все варианты увеличить (уменьшить) в одно и то же число раз, то средняя арифметическая увеличиться (уменьшится) во столько же раз: $\overline{kx} = k\bar{x}$

   или $\frac{\sum_{i=1}^{m} (kx_i) n_i}{n} = k \frac{\sum_{i=1}^{m} x_i n_i}{n}$

3. Если все варианты увеличить (уменьшить) на одно и то же число, то средняя арифметическая увеличиться (уменьшится) на это же число:

$$\overline{x + c} = \bar{x} + c$$

4.

   или $\frac{\sum_{i=1}^{m} (x_i + c) n_i}{n} = \frac{\sum_{i=1}^{m} x_i n_i}{n} + c$.

5. Средняя арифметическая отклонений вариантов от средней арифметической равна нулю:

   $\overline{x - \bar{x}} = 0$

   или $\sum_{i=1}^{m} (x_i - \bar{x}) n_i = 0$

6. Средняя арифметическая алгебраической суммынескольких признаков равна такой же сумме средних арифметических этих признаков:

$$\overline{x+y} = \bar{x} + \bar{y}$$

.

7. Если ряд состоит из нескольких групп, общая средняя равна средней арифметической групповых средних, причем весами являются объем групп:

$$\bar{x} = \frac{\sum_{i=1}^{l} \bar{x}_i n_i}{n}$$

где $\bar{x}$ – общая средняя, $\bar{x}_i$ – групповая средняя i-ой группы, объем которой равен n_i, l – число групп.

При решении практических задач могут применяться и иные формы средней, которые можно получить из средней степенной к-го порядка:

$$\overline{x_k} = \sqrt[k]{\frac{\sum_{i=1}^{m} x_i^k n_i}{n}},$$

где $x_i > 0$.

Легко убедиться в том, что при k=1 получаем формулу средней арифметической. При k=-1 получаем формулу средней гармонической :

$$\overline{x_{-1}} = \left(\frac{\sum_{i=1}^{m} x^{-1} n_i}{n} \right)^{-1} = \frac{n}{\sum_{i=1}^{m} \frac{n_i}{x_i}}.$$

При k=0 получаем формулу средней геометрической:

$$\overline{x_0} = \sqrt[n]{x_1^{n_1} \cdot x_2^{n_2} \dots x_m^{n_m}} = \sqrt[n]{\prod_{i=1}^{m} x_i^{n_i}}$$

При k=2 получаем формулу средней квадратической

$$\overline{x_2} = \sqrt{\frac{\sum_{i=1}^{m} x_i^2 n_i}{n}}$$

Можно показать, что с ростом порядка k степенная средняя возрастает, т.е.

$$\overline{x_{-1}} < \overline{x_0} < \overline{x_1} < \overline{x_2} < \dots$$

(свойство мажорантности средних)

Кроме рассмотренных средних величин, называемых аналитическими, в статистическом анализе применяют структурные, или порядковые, средние. Из них наиболее широко применяется медиана и мода.

Медианой $\tilde{M}e$ вариационного ряда называется значение признака, приходящееся на середину ранжированного ряда наблюдений.

Для дискретного вариационного ряда с нечетным числом членов медиана равна серединному варианту, а для ряда с четным числом членов – полусумме двух серединных вариантов.

Пример. Найти медиану распределения рабочих по тарифному разряду:

Тарифный разряд x_i	1	2	3	4	5	6	$\sum$
Частота (количество рабочих) n_i	2	3	6	8	22	9	50

n=50 – четное, следовательно средеинных вариантов два: x_{25}=5 и x_{26}=5.

Поэтому $\tilde{Me}$ =5%.

Для интервального вариационного ряда находится медианный интервал, на который приходится середина ряда, а значение медианы на этом интервале находят с помощью линейного интерполирования.

$$\tilde{Me} = a_{\tilde{Me}} + k\frac{\frac{n}{2} - m_{\tilde{Me}}^{íàê}}{m_{\tilde{Me}}}$$

где $a_{\tilde{Me}}$ – начало медианного интервала, для которого $n_i^{нак}$ >n/2; k – величина интервала (интервальная разность); $m_{\tilde{Me}}^{íàê}$ частота, накопленная к началу медианного интервала; $m_{\tilde{Må}}$ частота медианного интервала.

Так же медианный интервал можно приближенно с помощью кумуляты как значение признака, для которого $n_x^{íàê} = \frac{n}{2}$ или $w_x^{íàê} = \frac{1}{2}$.

Достоинство медианы как меры центральной тенденции заключается в том, что на нее не влияет изменение крайних членов вариационного ряда, если любой из них, меньший медианы, остается меньше ее, а любой, больший медианы, продолжает быть больше ее. Медиана предпочтительнее средней арифметической для ряда, у которого крайние варианты по сравнению с остальными оказались чрезмерно большими или малыми.

Модой $\tilde{Mo}$ вариационного ряда называется вариант, которому соответствует наибольшая частота.

Например для вариационного ряда

Тарифный разряд x_i	1	2	3	4	5	6	$\sum$
Частота (количество рабочих) n_i	2	3	6	8	22	9	50

мода $\tilde{Mo}=5,$ так как этому варианту соответствует наибольшая частота n_i=22. Для интервального ряда находится модальный интервал, имеющий наибольшую частоту, а значение моды на этом интервале определяют с помощью линейного интерполирования.

$$\tilde{Mo} = a_{\tilde{Mo}} + k\frac{m_{\tilde{Mo}} - m_{\tilde{Mo}-1}}{2m_{\tilde{Mo}} - m_{\tilde{Mo}-1} - m_{\tilde{Mo}+1}}$$

где $a_{\tilde{Mo}} -$ начало модального интервала, для которого n_i – макимально; k – величина интервала (интервальная разность); $m_{\tilde{Mo}}$ частота модального интервала; $m_{\tilde{Mo}-1}$ $m_{\tilde{Mo}+1}$ частота интервала, предшествующего модальному (следующему за модальным).

Однако проще моду можно найти графическим путем с помощью гистограммы.

Особенность моды как меры центральной тенденции заключается в том, что она не изменяется при изменениикрайних членов ряда, т.е. обладает определенной устойчивочтью к вариации признака.