СТАТИСТИЧЕСКОЕ ИЗУЧЕНИЕ ВАРИАЦИИ

Абсолютные показатели математической оценки размера вариации

Средние величины и показатели вариации широко применяются для характеристики статистических совокупностей по варьирующим признакам.

Средняя величина является обобщенной характеристикой для однородной совокупности. Она определяет общие условия в отношении изучаемого признака. Но для всесторонней характеристики вариационного ряда необходимо установить степень колеблемости отдельных значений признака. *Вариация* представляет собой изменение значений этого признака или его колеблемость за определенный период или на момент времени. Показатели, характеризующие колеблемость признаков, получили общее название *показатели вариации*. Для определения абсолютной меры колеблемости признаков или величины вариации применяются абсолютные средние размеры вариации: размах вариации, среднее линейное отклонение, дисперсия признака или средний квадрат отклонения, среднее квадратическое отклонение.

Размах вариации - это абсолютная разность между максимальным x_{max} и минимальным x_{min} значениями признака в изучаемой совокупности:

$$R = x_{\max} - x_{\min} .$$

Он измеряет только крайние отклонения вариант в ряду.

Для измерения отклонения каждой варианты от средней величины в ряду распределения или в группировке применяется *среднее линейное отклонение.*

Все вышеперечисленные абсолютные показатели вариации исчисляются путем нахождения отклонения каждой единицы изучаемой совокупности от средней величины. Упрощенно эти расчеты можно произвести, применив способ отсчета от условного нуля, или способ моментов.

Вернемся к среднему линейному отклонению, которое определяется из отклонений индивидуальных значений признака от средней величины без учета знака этих отклонений. Среднее линейное отклонение определяется по следующим формулам:

Для ранжированного ряда: $\bar{d} = \dfrac{\sum|x-\bar{x}|}{n}$ (простое);

Для интервального ряда: $\bar{d} = \dfrac{\sum|x-\bar{x}|\cdot f}{\sum f}$ (взвешенное);

Этот показатель не устанавливает степень рассеивания признаков.

Наибольшее применение в практике статистических работ и в математической статистике находит показатель дисперсия признака, или средний квадрат отклонений, или квадрат среднего квадратического отклонения. Он определяется как средняя из отклонений индивидуальных значений признака от средней величины, возведенных в квадрат. Обозначается греческой буквой сигма (σ2). Дисперсия σ^2 определяется по формулам:

для ранжированного ряда: $\sigma^2 = \dfrac{\sum(x-\bar{x})^2}{n}$ (простая);

для интервального ряда $\sigma^2 = \dfrac{\sum(x-\bar{x})^2\cdot f}{\sum f}$ (взвешенная).

Корень квадратный из дисперсии σ^2 представляет собой *среднее квадратическое отклонение:*

$$\sigma = \sqrt{\sigma^2} .$$

Среднее квадратическое отклонение σ *определяется по формулам :*

для ранжированного ряда: $\sigma = \sqrt{\dfrac{\sum(x-\bar{x})^2}{n}}$ (простое);

для интервального ряда: $\sigma = \sqrt{\dfrac{\sum(x-\bar{x})^2 \cdot f}{\sum f}}$ (взвешенное).

Среднее квадратическое отклонение по величине всегда больше среднего линейного отклонения. Соотношение $\sigma : \bar{d}$ для нормального закона распределения должно равняться примерно 1:2. Если оно больше, то это свидетельствует о наличии в совокупностях резких выделяющихся отклонений не однородных с основной массой элементов, нарушающих развитие основной тенденции или закономерности совокупности. Поэтому в исследованиях необходимо применять тот или иной показатель (σ или d). Среднее квадратическое отклонение является мерой надежности средней величины: чем оно меньше, тем точнее средняя арифметическая отражает собой всю изучаемую совокупность.

Относительные показатели интенсивности вариации

Для оценки интенсивности вариации и сравнения ее в разных совокупностях и различных признаков применяются относительные показатели вариации. Они вычисляются как отношения абсолютных показателей силы вариации к средней арифметической величине признака.

Существуют три относительных показателя, выраженных в процентах:

относительный размах вариации

$$K_p = \frac{R}{\bar{x}} \cdot 100$$

отражает относительную меру колеблемости крайних значений признака вокруг средней;

относительное линейное отклонение

$$K_d = \frac{\bar{d}}{\bar{x}} \cdot 100$$

отражает долю усредненного значения абсолютных отклонений от средней величины;

коэффициент вариации как относительное квадратическое отклонение от средней величины

$$V = \frac{\sigma}{\bar{x}} \cdot 100$$

Наиболее распространенным показателем колеблемости признаков является коэффициент вариации, потому что среднее квадратическое отклонение дает общую или обобщенную характеристику колеблемости всех вариантов совокупности.

Статистическое изучение вариации в прерывных, ранжированных, непрерывных и интервальных рядах

Все абсолютные показатели колеблемости признаков выражаются в единицах измерения признака x, в основном в именованных числах.

В дискретном прерывном или ранжированном ряду абсолютные показатели колеблемости признаков или размеры вариации определяются по следующим простым формулам:

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

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

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

В интервальном ряду или по сгруппированным данным абсолютные показатели колеблемости признаков или показатели размера вариации определяются как отклонения середины интервалов от средней арифметической величины, взвешенные (умноженные) по соответствующей частоте f по следующим взвешенным формулам:

$$\bar{d} = \frac{\sum|x-\bar{x}|\cdot f}{\sum f};$$

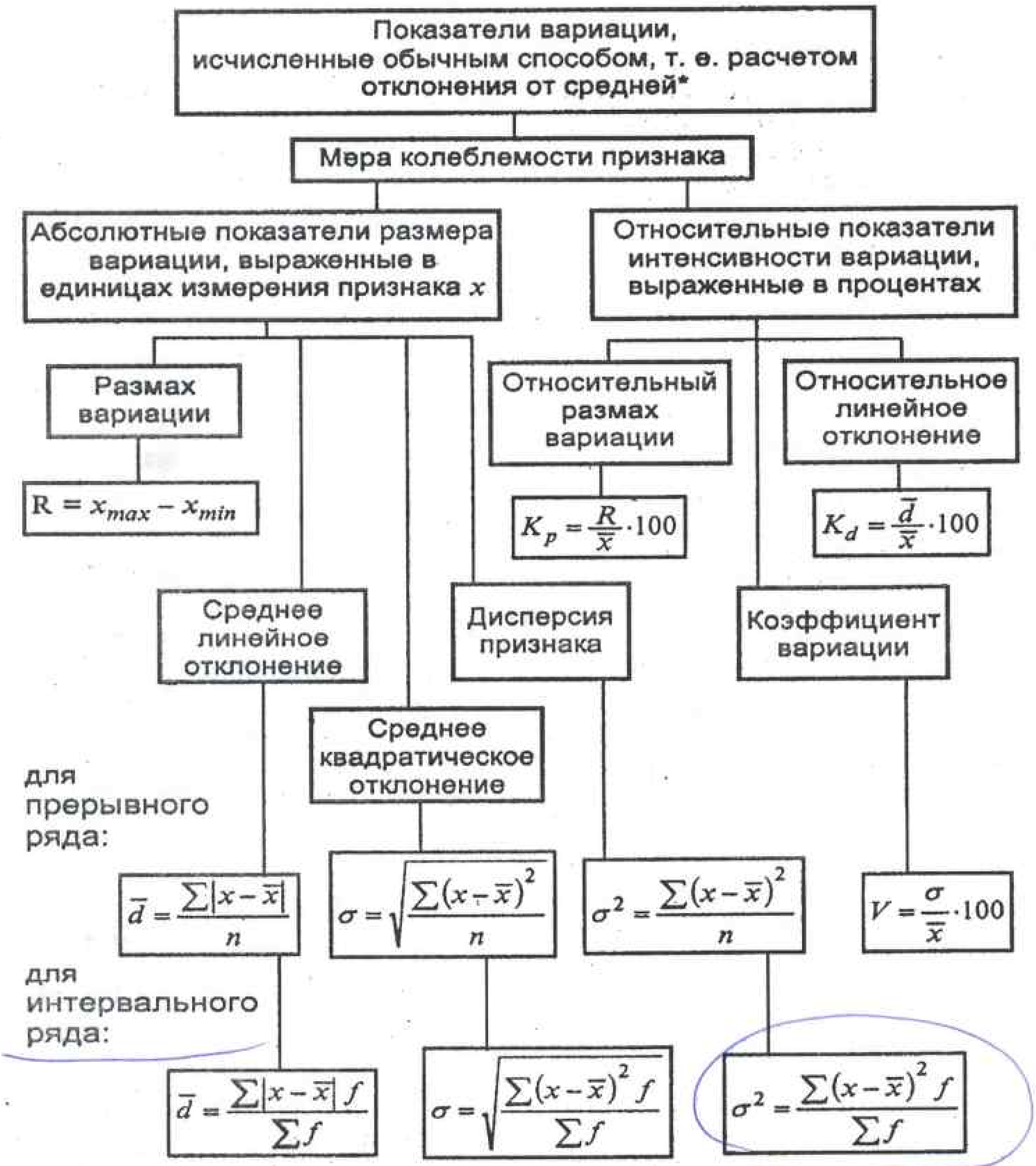
$$\sigma^2 = \frac{\sum(x-\bar{x})^2\cdot f}{\sum f};$$

$$\sigma = \sqrt{\frac{\sum(x-\bar{x})^2\cdot f}{\sum f}}.$$

Для лучшего запоминания, более ясного представления и использования в экономико-статистическом анализе абсолютные и относительные показатели вариации представлены на схеме 6.1.

Схема 6.1

Статистическое изучение вариации



* *Примечание*: существует способ, упрощающий расчеты: способ отсчета от условного нуля, или способ моментов.

Методика определения среднего квадратического отклонения или дисперсии признака

Пример 1. Для изучения естественной убыли произведено 5%-ное выборочное обследование партии хранящихся на базе товаров. В результате лабораторного анализа установлено следующее распределение образцов (табл. 1).

Таблица 1

Процент естественной убыли	Количество образцов f
до 4	6
4-6	14
6-8	22
8-10	48
от 10 и выше	10
Итого	**100**

Исчислить средний процент естественной убыли в выборке и среднее квадратическое отклонение или дисперсию.

Решение

Для определения середины интервала в каждой группе, т. е. среднего значения, необходимо от интервального перейти к дискретному ряду (прерывному). Величина интервала равна 2 (6-4, 8-6 и т.д.), среднее значение для второй и третьей групп составит 5 и 7, т. е. $\frac{x_{min} + x_{max}}{2}$;

вторая группа $\frac{2+6}{2} = 5$;

третья группа $\frac{6+8}{2} = 7$ и т.д.

Но первая и последняя группы имеют варианты с открытыми интервалами, у которых верхняя или нижняя границы точно не определены и сама граница остается как бы открытой. В этом случае поступают следующим образом. Величину интервала (2) делят на два и из верхнего значения отнимают полученную величину (1); к нижнему значению прибавляют эту величину (1) и получают значения вариантов: 4-1=3 и 10+1 = 11.

Расчеты для средней величины и дисперсии признака произведем в табл. 2.

Таблица 2. Расчетная таблица.

Процент естественной убыли x	Количество образцов f	Середина интервала x	xf	$x-\bar{x}$	$(x-\bar{x})^2$	$(x-\bar{x})^2 f$
До 4	6	3	18	-5	25	150
4-6	14	5	70	-3	9	126
6-8	22	7	154	-1	1	22
8-10	48	9	432	+1	1	48
от 10 и выше	10	11	110	+3	9	90
Итого	**100**		**784**			**436**

Средняя величина определяется по формуле средней арифметической взвешенной:

$$\bar{x}_{арифмвзв.} = \frac{\sum x \cdot f}{\sum f}$$

а дисперсия признака - по формуле взвешенной:

$$\sigma_x^2 = \frac{\sum (x-\bar{x})^2 \cdot f}{\sum f}.$$

Расчет σ_x^2 произведем обычным способом, т. е. отсчетом от средней арифметической величины, в следующем порядке:

1) найдем отклонения каждой варианты от средней величины - $(x-\bar{x})$;
2) полученные отклонения возведем в квадрат - $(x-\bar{x})^2$;
3) квадраты отклонения умножим на соответствующие частоты - $(x-\bar{x})^2 f$
4) полученные результаты суммируем и поделим на сумму частот:

$$\sigma_x^2 = \frac{\sum(x-\bar{x})^2 \cdot f}{\sum f}$$

Средняя величина в выборочной совокупности $\tilde{x}_{ариф}$ обозначается сверху волнистой чертой. Средний процент естественной убыли в выборочной совокупности

$$\tilde{x} = \frac{\sum x \cdot f}{\sum f} = \frac{784}{100} = 7{,}84\% \approx 8\%$$

Дисперсия $\sigma^2 = \frac{436}{100} = 4{,}36\%$.

Решение примера 1 будет продолжено в следующей теме.

В некоторых случаях необходимо определить коэффициент равномерности. Он исчисляется на основании коэффициента вариации, который показывает отношение среднего квадратического отклонения и средней величины, выраженное в процентах: $V = \frac{\sigma \cdot 100}{\bar{x}}$.

Отсюда коэффициент равномерности равен: $P = 100 - V$.