Свойства дисперсии

Расчет дисперсии можно упростить.
В случае равных интервалов в вариационном ряду распределения используется способ отсчета от условного нуля(метод моментов).

Математические свойства дисперсии:

1. Дисперсия постоянной величины равна нулю.
2. Уменьшение всех значений признака на одно и тоже число **А** не меняет величину дисперсии.
3. Уменьшение всех значений признака в k раз уменьшает дисперсию в k^2 раз, а среднее квадратическое отклонение в k раз.

Формы распределения

Различают эмпирические и теоретические кривые распределения.

***Эмпирическая кривая распределения* - это фактическая** кривая распределения, полученная по данным наблюдения, в которой отражаются как общие, так и случайные условия, определяющие распределение.

Теоретическая кривая распределения - это кривая, выражающая **функциональную связь** между изменением варьирующего признака и изменением частот и характеризующая определенный тип распределения.

Кривые распределения бывают симметричными и асимметричными.
В зависимости от того, какая ветвь кривой вытянута - правая или левая, различают правостороннюю или левостороннюю асимметрию.
Кривые распределения могут быть одно-, двух- и многовершинными.

Для однородных совокупностей, как правило, характерны одновершинные распределения.

Многовершинность свидетельствует о неоднородности изучаемой совокупности. Появление двух и более вершин делает необходимой перегруппировку данных с целью выделения более однородных групп.

Средние арифметические разных степеней отклонения индивидуальных признаков от определённой исходной величины называются **моментами распределения**.

При исчислении средней в качестве весов могут быть использованы частоты или частости (тогда эти моменты называются **эмпирическими**), а могут вероятности (тогда моменты называются **теоретическими**).

Формула момента k-го порядка:

$$M_k = \frac{\sum_{i=1}^{n}(x_i - A)^k \cdot f_i}{\sum_{i=1}^{n} f_i}$$

x – варианты
k – показатель степени
f – частоты
A – выбранная исходная величина

1. При А = 0 получаем **системy начальных моментов**.

2. При А равном не нулю, а не которой величине x_0 получаем **систему условных моментов**.

3. При А равной средней арифметической величине получаем систему **центральных моментов**.

Центральный момент первого порядка всегда равен нулю в соответствии с нулевым свойством средней арифметической

$$\mu_1 = 0$$

Центральный момент второго порядка представляет собой дисперсию и служит основной мерой колеблемости признака

$$\mu_2 = \sigma^2$$

Центральный момент третьего порядка служит **мерой асимметрии** распределения. Если распределение симметрично, то

$$\mu_3 = 0$$

Центральный момент четвёртого порядка применяется для вычисления показателя **эксцесса** (остро- или плосковершинного распределения).

Отношение центрального момента k-го порядка к k-ой степени среднего квадратического отклонения называется **нормированным моментом**:

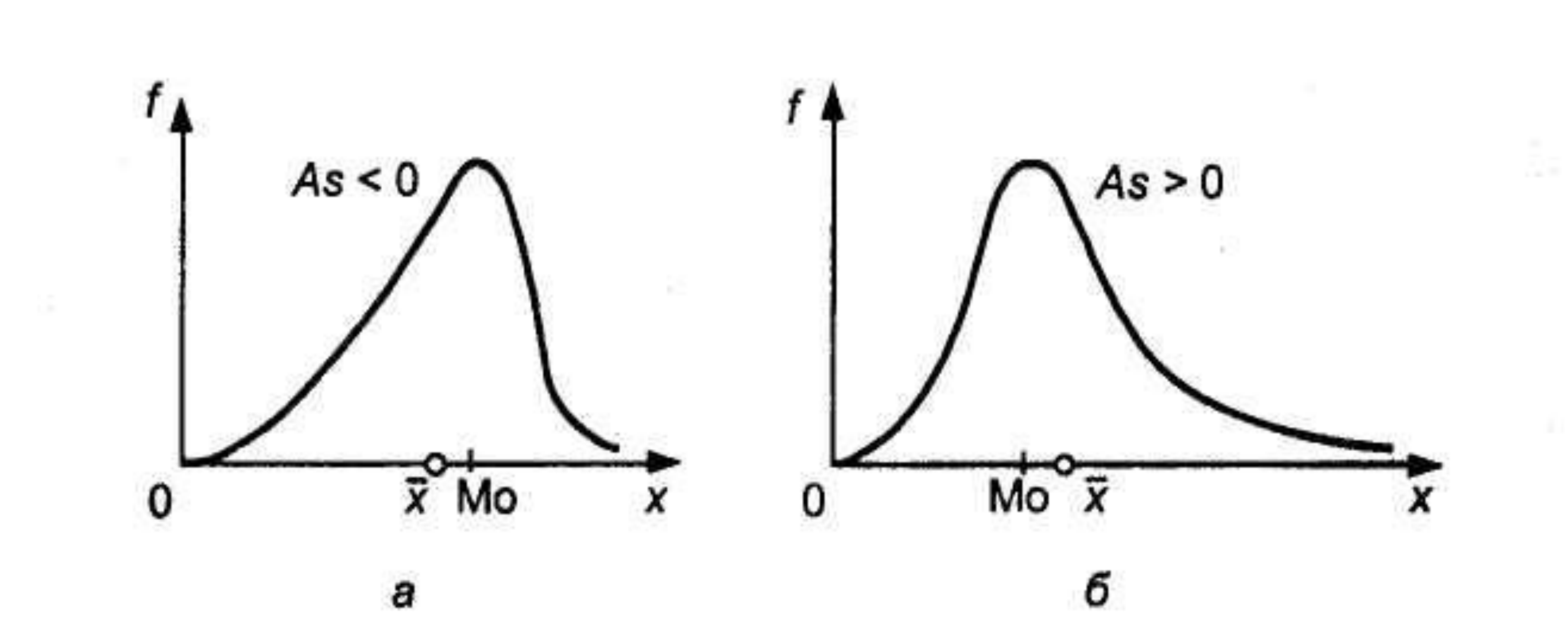
$$r_k = \frac{\mu_k}{\sigma^k}$$

Нормированный момент

- первого порядка равен 0;
- второго порядка равен 1;
- третьего порядка используется для характеристики асимметрии;
- четвертого порядка используется для характеристики эксцессов.

Для характеристики асимметрии более широко применяется нормированный момент третьего порядка:

$$As = \frac{\mu_3}{\sigma^3} = \frac{\sum_{i=1}^{n} \left(x_i - \overline{x}\right)^3 \cdot f_i}{\sum_{i=1}^{n} f_i \cdot \sigma^3}$$



Асимметрия распределения:
а — левосторонняя;
б — правосторонняя.

As=0 если ряд распределения симметричен, т. е.

$$\bar{x} = Mo$$

As>0 если скошенность ряда правосторонняя, т.е.

$$\bar{x} > Mo$$

As<0 если скошенность ряда левосторонняя, т.е.

$$\bar{x} < Mo$$

Если As <0,5 (независимо от знака) то асимметрия считается незначительной.

Если As>0,5 то асимметрия считается значительной.

Расчет асимметрии и эксцесса распределения.

Группы студентов по возрасту **x**, лет	Число студентов f_i	$(x_i - \bar{x})$	$(x_i - \bar{x})^3 f_i$	$(x_i - \bar{x})^4 f_i$
17	10	-3,7	-506,53	1874,161
18	70	-2,7	-1377,81	3720,087
19	80	-1,7	-393,04	668,168
20	100	-0,7	-34,3	24,01
21	120	0,3	3,24	0,972
22	160	1,3	351,52	456,976
23	90	2,3	1095,03	2518,569
Итого	630		-861,89	9262,943

$$As = \frac{\mu_3}{\sigma^3} = \frac{\sum_{i=1}^{n}(x_i - \bar{x})^3 \cdot f_i}{\sum_{i=1}^{n} f_i \cdot \sigma^3}$$

$$As = \frac{-861,89}{630 * 1,63^3} = -0,32$$

$$As = -0{,}32 < 0$$

Вывод: скошенность распределения левосторонняя, а асимметрия небольшая (незначительная).

Оценка степени существенности этого показателя дается с помощью средней квадратической ошибки, рассчитываемой по формуле

$$\sigma_{As} = \sqrt{\frac{6 \cdot (n-1)}{(n+1) \cdot (n+3)}}$$

где n — число наблюдений.

Если $\dfrac{|As|}{\sigma_{As}} > 3$

то **асимметрия существенна** и распределение признака в генеральной совокупности не является симметричным.

Если $\dfrac{|As|}{\sigma_{As}} < 3$

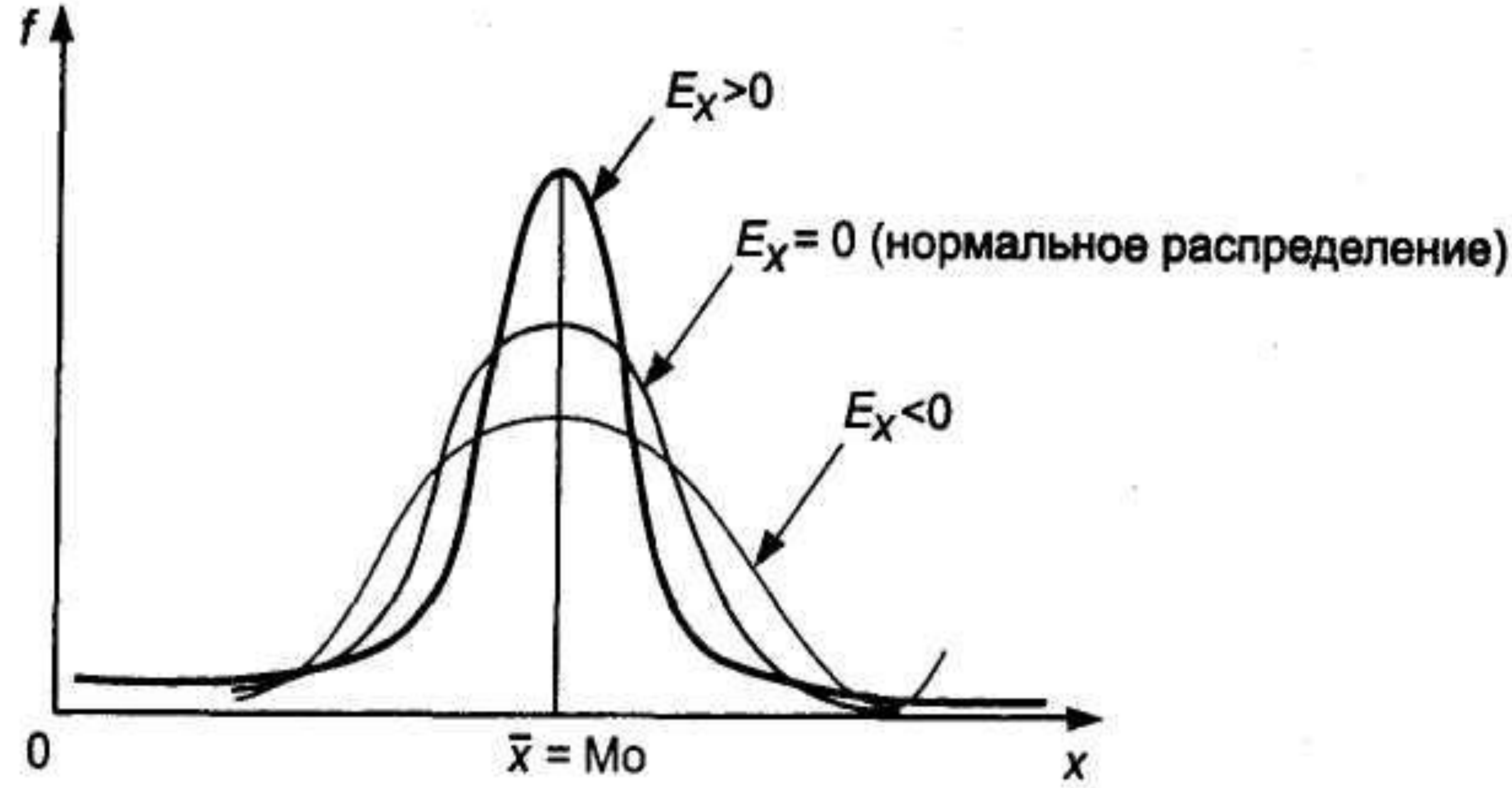
то **асимметрия несущественна**, ее наличие объясняется влиянием случайных обстоятельств.

$$\sigma_{As} = \sqrt{\frac{6\cdot(7-1)}{(7+1)\cdot(7+3)}} = 0{,}67 \qquad \frac{|-0{,}32|}{0{,}67} = 0{,}48$$

Т.е., асимметрия несущественна.

Под **эксцессом** понимается **степень островершинности** (крутизны) распределения, при этом в качестве эталона берется нормальное распределение. Характеристикой эксцесса является нормированный момент четвертого порядка:

$$Ex = \frac{\mu_4}{\sigma^4} - 3 = \frac{\sum_{i=1}^{n}(x_i - \bar{x})^4 \cdot f_i}{\sum_{i=1`}^{n} f_i \cdot \sigma^4} - 3$$



Для вариационного ряда нормального распределения Ex=0.
Для более островершинных распределений, чем нормальное, Ex > 0,
для более плосковершинных Ex < 0.

$$Ex = \frac{9262,94}{630 * 1,63^4} - 3 = -0,92$$

Вывод: распределение плосковершинное

Предельным значением отрицательного эксцесса является значение Ex= - 2; величина положительного эксцесса является величиной бесконечной.

В нормальном распределении

$$\frac{\mu_4}{\sigma^4} = 3$$

Средняя квадратическая ошибка эксцесса исчисляется по формуле

$$\sigma_{Ex} = \sqrt{\frac{24 \cdot n \cdot (n-2) \cdot (n-3)}{(n-1)^2 \cdot (n+3) \cdot (n+5)}}$$

где *n* — число наблюдений.

Если $\frac{|Ex|}{\sigma_{Ex}} \leq 2$, то распределение можно считать нормальным.

$$\sigma_{Ex} = \sqrt{\frac{24 \cdot 7 \cdot (7-2) \cdot (7-3)}{(7-1)^2 \cdot (7+3) \cdot (7+5)}} = 0{,}88$$

$$\frac{Ex}{\sigma_{Ex}} = \frac{|-0{,}92|}{0{,}88} = 1{,}05 < 2$$

т.е. распределение можно считать нормальным.