13 ЧИСЛОВЫЕ ХАРАКТЕРИСТИКИ СЛУЧАЙНЫХ ВЕЛИЧИН

13.1 Характеристики положения

Закон распределения случайной величины является ее исчерпывающей характеристикой. Однако на практике часто не требуется такого полного описания и достаточно знать только некоторые числовые характеристики случайной величины, такие, например, как показатели положения и разброса случайной величины.

Среди числовых характеристик положения случайной величины наиболее важной является *математическое ожидание*, которое характеризует центр рассеивания случайной величины.

Математическим ожиданием называется сумма произведений всех возможных значений случайной величины на вероятности этих значений. С точки зрения механики её можно интерпретировать, как координату центра тяжести вероятностных масс. Если дискретная случайная величина *X* может принимать значения $x_1,...,x_n$ с вероятностями $p_1,...,p_n$, то математическое ожидание *X*

$$M(X)=\frac{x_1p_1+x_1p_1+\ldots+x_np_n}{p_1+p_1+\ldots p_n}=\frac{\sum\limits_{i=1}^{n}x_ip_i}{\sum\limits_{i=1}^{n}p_i}.$$

Учитывая, что $\sum\limits_{i=1}^{n}p_i=1,$ получим, что

$$M(X)=\sum_{i=1}^{n}x_ip_i. \qquad (13.1)$$

Если случайная величина непрерывна, то отмеченная сумма превращается в интеграл

$$M(X) = \int_{-\infty}^{\infty} x f(x) dx, \quad (13.2)$$

где $f(x)$ плотность распределения случайной величины *X.* В литературе по теории вероятностей для математического ожидания применяется так же и более компактное обозначение $\overline{X}$, где черта над обозначением случайной величины обозначает математическое ожидание этой величины. Применяется так же обозначение m_x. То есть данном случае $M(X) = \overline{X} = m_x$.

Рассмотрим среднее арифметическое наблюдённых значений величины X:

$$M^*(x) = \frac{x_1 m_1 + x_1 m_1 + \ldots + x_1 m_1}{N} = \sum_{i=1}^{n} x_i \frac{m_i}{N}.$$

Но $\frac{m_i}{N}$ есть не что иное, как частота (или статистическая вероятность) события X. При увеличении числа опытов $\frac{m_i}{N} \to p_i$. Т. е. среднее арифметическое стремится к математическому ожиданию.

Для непрерывных величин:

$$M(x) = \int_{-\infty}^{+\infty} x f(x) dx.$$

<u>*Модой*</u> случайной величины называется её наиболее вероятное значение.

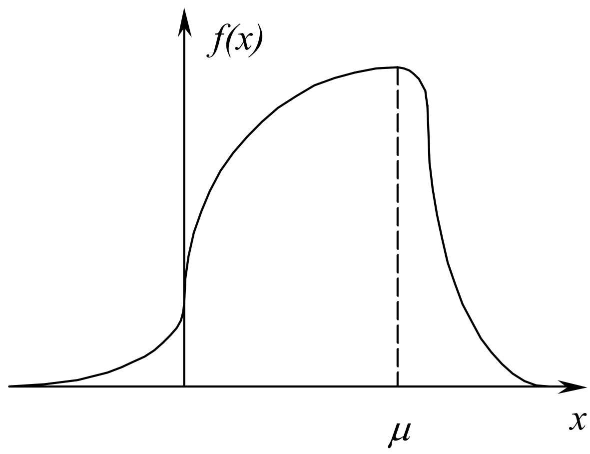

Рисунок. 13.1 – Иллюстрация к определению моды случайной величины

Для симметричного распределения мода и математическое ожидание совпадают.

Медианой случайной величины X называется такое её значение μ_e, для которого

$$P(X < \mu_e) = P(X > \mu_e) = 1/2,$$

т.е. одинаково вероятно, окажется ли случайная величина меньше или больше μ_e.

Геометрически это абсцисса точки, в которой площадь под кривой распределения делится пополам.

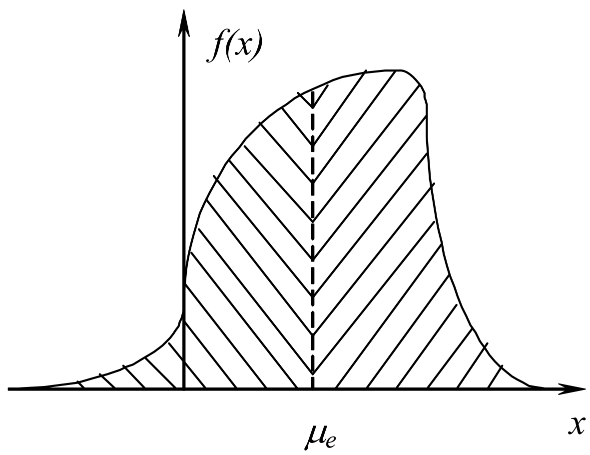

Рисунок 13.2 – Иллюстрация к определению медианы

Так же как в механике для описания распределения масс используются статические моменты и моменты инерции, в теории вероятности используются начальные и центральные моменты.

13.2 Моменты, дисперсия, квадратичное отклонение

Кроме рассмотренных выше характеристик положения используются и другие числовые характеристики случайной величины, характеризующие степень разброса, форму распределения и др. Для этого используют моменты распределения.

Начальным моментом s-го порядка называется сумма вида:

$$\alpha_s(X)=\sum_{i=1}^{n} x_i^s p_i .$$

для дискретной случайной величины и интеграл:

$$\alpha_s(X)=\int_{-\infty}^{+\infty} x^s f(x)dx. \qquad (13.3)$$

для непрерывной случайной величины.

Нетрудно убедиться, что математическое ожидание – это первый начальный момент, а математическое ожидание s-той степени X – это начальный момент порядка s случайной величины X. То есть

$$\alpha_s(X)=M\left(X^s\right)=\overline{X^s}\ . \qquad (13.4)$$

Центрированной случайной величиной, соответствующей величине X, называется отклонение случайной величины X от её математического ожидания:

$$X^\circ = X - m_x . \qquad (13.5)$$

Математическое ожидание центрированной случайной величины равно нулю.

$$M\left(X^\circ\right)=M\left(X-m_x\right)=\sum_{i=1}^{n}\left(x_i-m_x\right)p_i=\sum_{i=1}^{n} x_i p_i - m_x\sum_{i=1}^{n} p_i = m_x - m_x = 0.$$

Моменты центрированной случайной величины называются центральными моментами. Они аналогичны моментам относительно центра тяжести в механике.

Центральным моментом порядка s случайной величины X называется математическое ожидание s-й степени соответствующей центрированной случайной величины:

$$\mu_s(X)=M\left(X^{\circ s}\right)=M\left(\left(X-m_x\right)^s\right).$$

Для дискретной случайной величины s-й центральный момент:

$$\mu_s=\sum_{i=1}^{n}\left(x_i-m_x\right)^s p_i ,$$

для непрерывной:

$$\mu_s = \int_{-\infty}^{+\infty} (x - m_x)^s f(x) dx. \quad (13.6)$$

Рассмотрим второй центральный момент.

$$\mu_2 = M\left(\overset{o}{X}{}^2\right) = \sum_{i=1}^{n} (x_i - m_x)^2 p_i =$$
$$= \sum_{i=1}^{n} x_i^2 p_i - 2m_x \sum_{i=1}^{n} x_i p_i + m_x^2 \sum_{i=1}^{n} p_i =$$
$$= \alpha_2 - 2m_x^2 + m_x^2 = \alpha_2 - m_x^2.$$

Аналогично третий центральный момент.

$$\mu_3 = M\left(\overset{o}{X}{}^3\right) = \sum_{i=1}^{n} (x_i - m_x)^3 p_i =$$
$$= \sum_{i=1}^{n} x_i^3 p_i - 3m_x \sum_{i=1}^{n} x_i^2 p_i + 3m_x^2 \sum_{i=1}^{n} x_i p_i - m_x^3 \sum_{i=1}^{n} p_i =$$
$$= \alpha_3 - 3\alpha_2 m_x + 2m_x^3.$$

Вообще говоря, моменты могут рассматриваться относительно произвольной точки a:

$$\gamma_s = M\left((X - a)^s\right).$$

Но центральные моменты имеют преимущество: первый центральный момент равен нулю, а второй центральный момент минимален. Докажем это.

$$\gamma_2 = \sum_{i=1}^{n} (x_i - a)^2 p_i = \sum_{i=1}^{n} (x_i - m_x + m_x - a)^2 p_i =$$
$$= \sum_{i=1}^{n} (x_i - m_x)^2 p_i + 2(m_x - a) \sum_{i=1}^{n} (x_i - m_x) p_i + (m_x - a)^2 =$$
$$= \mu_2 + (m_x - a)^2.$$

Минимум будет при $a = m_x$.

Второй центральный момент называется *дисперсией* случайной величины. Обозначение:

$$D(X) = \mu_2 = M\left(X^{\circ 2}\right)$$

Т.е. *дисперсией* случайной величины X называется математическое ожидание квадрата соответствующей центрированной величины.

Другие формы записи:

$$D(X) = M\left(\left(X - m_x\right)^2\right);$$
$$D(X) = \sum_{i=1}^{n} \left(x_i - m_x\right)^2 p_i; \qquad (13.7)$$
$$D(X) = \int_{-\infty}^{+\infty} \left(x - m_x\right)^2 f(x) dx.$$

Дисперсия есть характеристика рассеивания случайной величины около её математического ожидания. В механике ей аналогичен момент инерции относительно центра тяжести. Формула для расчёта дисперсии:

$$D_x = \alpha_2 - m_x^2.$$

Дисперсия имеет размерность квадрата случайной величины. Чтобы иметь дело с величиной такой же размерности, что и случайная величина, введено *среднее квадратичное отклонение (стандарт)*.

Обозначение:

$$\sigma(X) = \sqrt{D(X)}.$$

Для положительных случайных величин в качестве меры разброса используют *коэффициент вариации* ν как отношение квадратичного отклонения к математическому ожиданию, то есть

$$\nu = \frac{\sigma}{m_x}. \qquad (13.8)$$

Третий центральный момент служит для характеристики асимметрии распределения. Он имеет размерность куба случайной величины. Чтобы получить безразмерную характеристику, его делят на куб среднего квадратичного отклонения. Получают *коэффициент асимметрии* или просто *асимметрию*:

$$Sk = \frac{\mu_3}{\sigma_3}. \qquad (13.9)$$

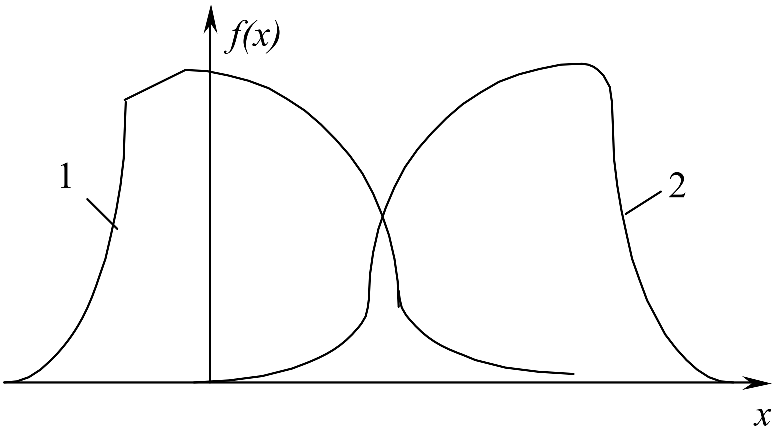

Рисунок 13.3 – Иллюстрация к определению ассиметрии. Распределение 1 имеет положиельную ассиметрию, а распределение 2 – отрицательную

Четвёртый центральный момент служит для характеристики “крутости”, то есть островершинности или плосковершинности распределения. Величина называется <u>*эксцессом*</u>. Запись:

$$Ex = \frac{\mu_4}{\sigma_4} - 3. \tag{13.10}$$

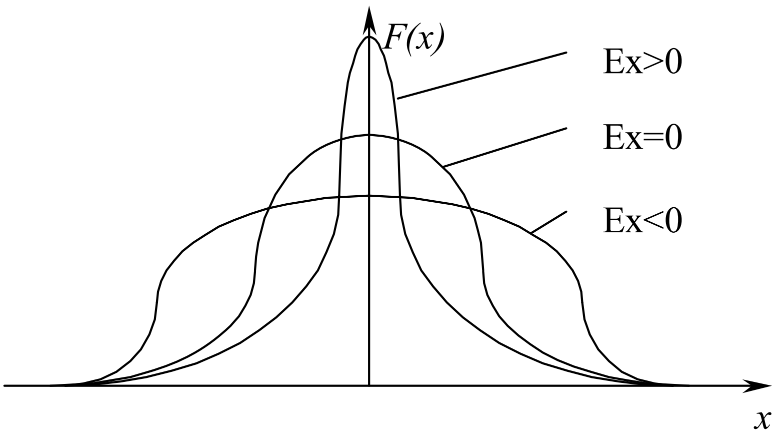

Рисунок 13.4 – Иллюстрация к определению эксцесса

Кроме начальных и центральных моментов иногда применяются абсолютные моменты (начальные и центральные), определяемые формулами

$$\beta_s = M\left(|X|^S\right); \quad \nu_s = M\left(\left|\overset{o}{X}\right|^S\right).$$

Первый абсолютный центральный момент называется ещё <u>*средним*</u>

<u>*арифметическим отклонением*</u>.

$$\nu_s = M\left(\left|\overset{o}{X}\right|\right) = M\left(\left|X - m_x\right|\right).$$

Квантиль порядка γ называют такое значение случайной величины *X,* обозначаемое X_γ, что вероятность того, что $X<X\gamma$ равна γ. Аналитически это можно записать, что X_γ является решением уравнения

$$F(X_\gamma) = \gamma \text{ или } \int_{-\infty}^{X_\gamma} f(x)dx = \gamma . \qquad (13.11)$$

Первое уравнение иллюстрируется рисунком 13.5

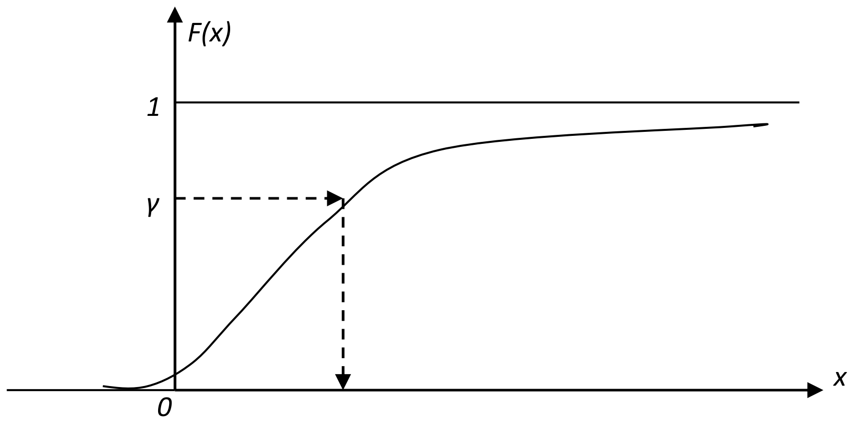

Рисунок 13.5 – Иллюстрация к первому уравнению (13.11)

Квантиль порядка $\gamma = 0.5$ совпадает с медианой. То есть $X_{0.5} = \mu_e$.

13.3 Контрольные вопросы

1. Что такое математическое ожидание случайной величины?
2. Что такое мода случайной величины?
3. Что такое медиана случайной величины?
4. Что такое центрированная случайная величина?
5. Что такое начальный моментом s-го порядка?
6. Что такое центральный момент порядка s?
7. Что такое дисперсия?

8. Что такое среднее квадратичное отклонение?

9. Что такое коэффициент вариации?

10. Что такое коэффициент ассиметрии?

11. Что такое эксцесс?

12. Что такое квантиль порядка γ?