15 РАСПРЕДЕЛЕНИЯ СЛУЧАЙНЫХ ВЕЛИЧИН-2

15.1 Нормальное распределение

Это наиболее часто встречающийся на практике закон распределения. Он характеризуется плотностью вероятности вида:

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-m)^2}{2\sigma^2}}.$$

Кривая распределения по нормальному закону имеет симметричный холмообразный вид (см. рисунок 15.1)

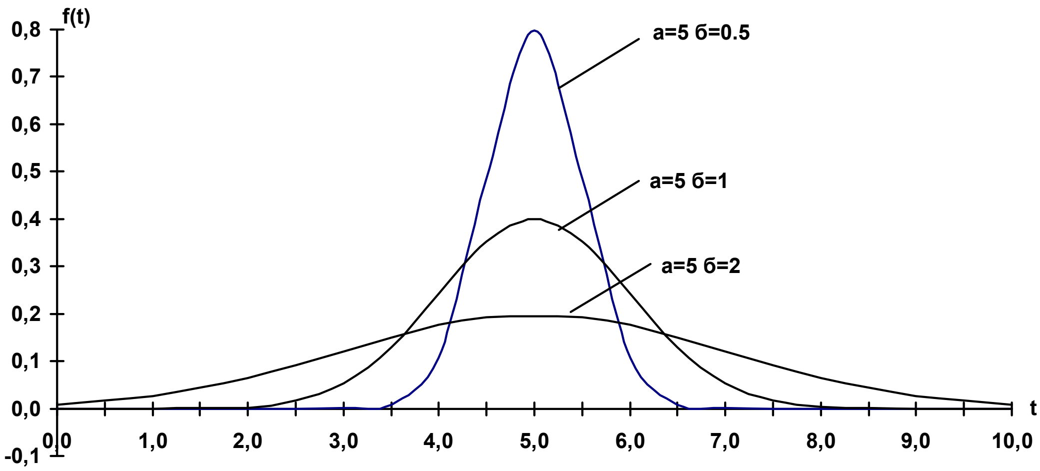


Рисунок 15.1 – Графики плотности нормального распределения при различных значениях квадратичного отклонения σ

Максимальная ордината кривой, равная $\frac{1}{\sigma\sqrt{2\pi}}$, соответствует точке $x = m$; по мере удаления от точки m плотность распределения падает и стремится к оси абсцисс.

Докажем, что m – есть математическое ожидание, а σ – есть среднее квадратическое отклонение. Для этого вычислим основные числовые характеристики случайной величины *X*.

$$M(X)=\int_{-\infty}^{+\infty} xf(x)dx=\frac{1}{\sigma\sqrt{2\pi}}\int_{-\infty}^{+\infty} xe^{-\frac{(x-m)^2}{2\sigma^2}}dx.$$

Применим замену переменной

$$\frac{x-m}{\sigma\sqrt{2}}=t.$$

$$M(X)=\frac{1}{\sqrt{\pi}}\int_{-\infty}^{+\infty}\left(\sigma\sqrt{2t}+m\right)e^{-t^2}dt=$$

$$=\frac{\sigma\sqrt{2}}{\sqrt{\pi}}\int_{-\infty}^{+\infty} te^{-t^2}dt+\frac{m}{\sqrt{\pi}}\int_{-\infty}^{+\infty} e^{-t^2}dt.$$

Первый интеграл равен нулю. Второй представляет собой интеграл Эйлера-Пуассона:

$$\int_{-\infty}^{+\infty} e^{-t^2}dt=2\int_{0}^{\infty} e^{-t^2}dt=\sqrt{\pi}.$$

Следовательно, $M(X)=m$. На практике параметр m часто называют центром рассеивания.

Вычислим дисперсию X:

$$D(X)=\frac{1}{\sigma\sqrt{2\pi}}\int_{-\infty}^{+\infty}(x-m)^2 e^{-\frac{(x-m)^2}{2\sigma^2}}dx.$$

Та же замена переменной:

$$\frac{x-m}{\sigma\sqrt{2}}=t.$$

Интегрирование по частям дает

$$D(X)=\frac{2\sigma^2}{\sqrt{\pi}}\int_{-\infty}^{+\infty} t^2 e^{-t^2}dt.$$

$$\mathrm{D(X)}=\frac{\sigma^2}{\sqrt{\pi}}\int_{-\infty}^{+\infty} t\cdot 2te^{-t^2}dt=$$

$$=\frac{\sigma^2}{\sqrt{\pi}}\left\{-te^{-t^2}\Big|_{-\infty}^{+\infty}+\int_{-\infty}^{+\infty} e^{-t^2}dt\right\}.$$

Первое слагаемое равно нулю, второе $\sqrt{\pi}$, откуда

$$D(X)=\sigma^2.$$

Геометрический смысл: m – центр симметрии кривой плотности распределения; σ – характеризует степень рассеивания случайной величины и одновременно расплывчатость кривой, поскольку площадь, ограниченная кривой плотности всегда равна единице. Размерность m и σ совпадает с размерностью случайной величины X. Выведем общие формулы для центральных моментов любого порядка.

$$\mu_S = \int_{-\infty}^{+\infty} (x-m)^S f(x)dx =$$

$$= \frac{1}{\sigma\sqrt{2\pi}} \int_{-\infty}^{+\infty} (x-m)^S e^{-\frac{(x-m)^2}{2\sigma^2}} dx.$$

Делаем замену переменной

$$\frac{x-m}{\sigma\sqrt{2}} = t.$$

$$\mu_S = \frac{\left(\sigma\sqrt{2}\right)^S}{\sqrt{\pi}} \int_{-\infty}^{+\infty} t^S e^{-t^2} dt.$$

Интегрируем по частям:

$$\mu_S = \frac{\left(\sigma\sqrt{2}\right)^S}{\sqrt{\pi}} \int_{-\infty}^{+\infty} t^{S-1} t e^{-t^2} dt =$$

$$= \frac{\left(\sigma\sqrt{2}\right)^S}{\sqrt{\pi}} \left\{ -\frac{1}{2} e^{-t^2} t^{S-1} \Big|_{-\infty}^{+\infty} + \frac{S-1}{2} \int_{-\infty}^{+\infty} t^{S-2} e^{-t^2} dt \right\}.$$

Первый член в скобках равен нулю. Получаем:

$$\mu_S = \frac{(S-1)\left(\sigma\sqrt{2}\right)^S}{2\sqrt{\pi}} \int_{-\infty}^{+\infty} t^{S-2} e^{-t^2} dt.$$

Но момент степени *S-2*:

$$\mu_{S-2} = \frac{\left(\sigma\sqrt{2}\right)^{S-2}}{\sqrt{\pi}} \int_{-\infty}^{+\infty} t^{S-2} e^{-t^2} dt.$$

Следовательно

$$\mu_S = (S-1)\sigma^2 \mu_{S-2}$$

Т. е. можно выражать чётные моменты через моменты на 2 порядка ниже. Нечетные моменты в силу симметрии распределения равны нулю. Т. е. для чётных моментов имеем:

$$\mu_2 = \sigma^2;$$
$$\mu_4 = 3\sigma^4;$$
$$\mu_6 = 15\sigma^6;$$
$$\dots$$

Общая формула для момента порядка S при чётном S:

$$\mu_S = (S-1)!!\sigma^S,$$

где под $(S-1)!!$ понимается произведение всех нечётных чисел от 1 до S-1.

Асимметрия:

$$S_k = \frac{\mu_3}{\sigma^3} = 0.$$

Эксцесс:

$$Ex = \frac{\mu_4}{\sigma^4} - 3 = \frac{3\sigma^4}{\sigma^4} - 3 = 0.$$

Т.е. эксцесс характеризует крутость конкретного закона распределения по отношению к нормальному.

Вычислим вероятность попадания случайной величины X, подчинённой нормальному закону с параметрами m, σ на участок от α до β.

$$P(\alpha < X < \beta) = F(\beta) - F(\alpha),$$

где $F(x)$ – функция распределения величины X.

$$F(x) = \int_{-\infty}^{x} f(x)dx = \frac{1}{\sigma\sqrt{2\pi}} \int_{-\infty}^{x} e^{-\frac{(x-m)^2}{2\sigma^2}} dx.$$

Замена переменной

$$\frac{x-m}{\sigma} = t.$$

$$F(x)=\frac{1}{\sqrt{2\pi}}\int\limits_{-\infty}^{\frac{x-m}{\sigma}} e^{-\frac{t^2}{2}}dx.$$

Этот интеграл сложный, но существуют специальные таблицы для функций:

$$\Phi(X)=\frac{2}{\sqrt{\pi}}\int\limits_{-\infty}^{x} e^{-t^2}dx, \quad \Phi^*(X)=\frac{1}{\sqrt{2\pi}}\int\limits_{-\infty}^{x} e^{-\frac{t^2}{2}}dx.$$

Φ^* есть нормальная функция распределения. Её таблицы приведены в приложениях учебников и задачников.

$$F(x)=\Phi^*\left(\frac{x-m}{\sigma}\right).$$

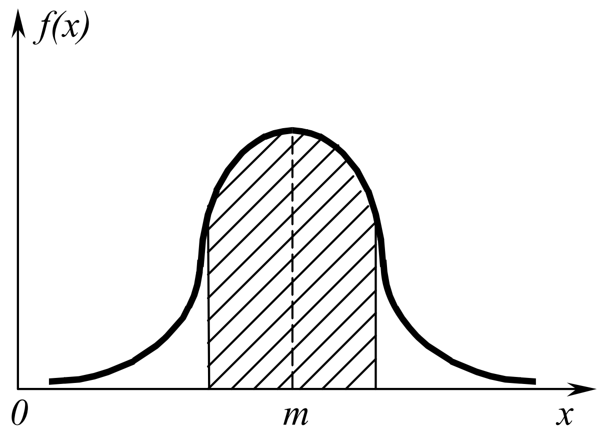
$$P(\alpha < X < \beta)=\Phi^*\left(\frac{\beta-m}{\sigma}\right)-\Phi^*\left(\frac{\alpha-m}{\sigma}\right).$$

Свойства функции Φ^*:

$$\Phi^*(-\infty)=0.$$
$$\Phi^*(+\infty)=1.$$
$$\Phi^*(-x)=1-\Phi^*(x).$$

Учитывая последнее свойство, рассмотрим вероятность попадания на участок, симметричный, относительно математического ожидания.

$$P(m-l < X < m+l)=\Phi^*\left(\frac{l}{\sigma}\right)-\Phi^*\left(-\frac{l}{\sigma}\right)=2\Phi^*\left(\frac{l}{\sigma}\right)-1.$$



Решим следующую задачу. Отложим от математического ожидания четыре отрезка длиной σ и вычислим вероятность попадания случайной величины X в каждый из них.

$$P(m < X < m+\sigma) = \Phi^*(1) - \Phi^*(0) \approx 0{,}341;$$
$$P(m+\sigma < X < m+2\sigma) = \Phi^*(2) - \Phi^*(1) \approx 0{,}136;$$
$$P(m+2\sigma < X < m+3\sigma) = \Phi^*(3) - \Phi^*(2) \approx 0{,}012;$$
$$P(m+3\sigma < X < m+4\sigma) = \Phi^*(4) - \Phi^*(3) \approx 0{,}001.$$

Вероятностью попадания в четвёртый участок уже практически можно пренебречь. Сумма же вероятностей для первых трёх равна 0,5 с точностью до 0,01 (1%). Т. е. можно сказать, что в интервале $m \pm 3\sigma$ укладывается практически всё рассеивание. Такой способ оценки диапазона возможных значений называется *правилом трёх сигм*. Это правило позволяет грубо оценить величину σ. Берут максимально возможное отклонение и делят его на три.

Часто (особенно в артиллерийской практике) для характеристики рассеяния кроме среднего квадратичного отклонения используют *вероятное (срединное) отклонение*, обозначается E или B.

Вероятным (срединным) отклонением случайной величины X, распределённой по нормальному закону, называется половина длины участка, симметричного относительно центра рассеивания, вероятность попадания в который равна 0,5.

Т. е. вероятность попадания в интервал $m \pm E$ равна 0,5.

$$P(|X-m| < E) = \frac{1}{2}; \quad P(|X-m| > E) = \frac{1}{2}.$$

Выразим E через σ:

$$P(|X-m| < E) = 2\Phi^*\left(\frac{E}{\sigma}\right) - 1 = \frac{1}{2};$$
$$2\Phi^*\left(\frac{E}{\sigma}\right) = \frac{3}{4} = 0{,}75;$$
$$\frac{E}{\sigma} = 0{,}674;$$
$$E = 0{,}674\sigma.$$

15.2 Показательное распределение

Плотность и функция *показательного распределения* положительной случайной величины T выражаются формулами:

$$f(t) = \frac{1}{a} e^{-\frac{t}{a}}, \quad F(t) = e^{-\frac{t}{a}}, \quad t \geq 0, \tag{15.1}$$

соответственно, a – параметр распределения. График плотности представлен на рисунке 15.2. В литературе это распределение называют также *экспоненциальным*.

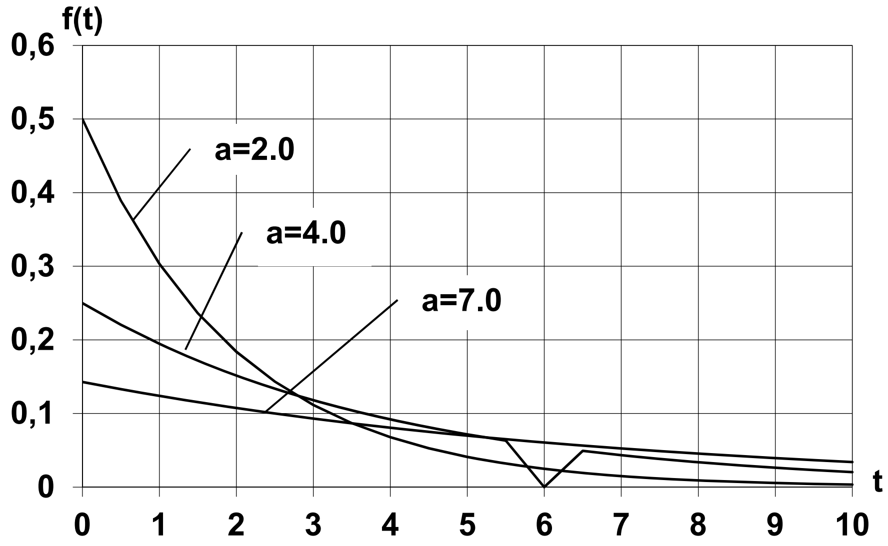


Рисунок 15.2 – График плотности показательного распределения при различных значениях параметра a

Математическое ожидание

$$m = M(T) = \int_{o}^{\infty} \frac{1}{a} \exp(-\frac{t}{a}) t dt = a .$$

Дисперсия

$$D(T) = \int_{0}^{\infty} \frac{1}{a} \exp(-\frac{t}{a}) t^2 dt - a^2 = a^2 .$$

Квадратичное отклонение

$$\sigma_T = \sqrt{D(T)} = a\,,$$

то есть для показательного распределения математическое ожидание и квадратичное отклонение совпадают.

Этот закон широко используется в теории надежности благодаря свойству "отсутствия памяти" (марковскому свойству), которое значительно облегчает выкладки и упрощает расчетные формулы. Суть свойства в том, что вероятность безотказной работы объекта в заданном интервале не зависит от времени предшествующей работы.

Показательный закон является предельным для вероятности безотказной работы сложных систем, если система состоит из элементов, каждый из которых отказывает и восстанавливается независимо, но при отказе хотя бы одного элемента простаивает вся система. Такая ситуация на практике весьма распространена. Она имеет место, например, для сложных станков, автоматических линий и др.

15.3 Логарифмически нормальное распределение

Логарифмически нормальное распределение для положительной случайной величины T имеет плотность

$$f(t) = \frac{1}{\sqrt{2\pi}dt}\exp\left[-\frac{(\ln t - \ln a)^2}{2d^2}\right], \quad t \ge 0\,, \qquad (15.2)$$

где a и d – параметры распределения.

Математическое ожидание

$$M(T) = \int_0^\infty \frac{1}{\sqrt{2\pi}dt}\exp\left[-\frac{(\ln t - \ln a)^2}{2d^2}\right]tdt = a \cdot \exp(\frac{d^2}{2}) \quad .$$

Дисперсия

$$D(T) = \int_0^\infty \frac{1}{\sqrt{2\pi}dt}\exp\left[-\frac{(\ln t - \ln a)^2}{2d^2}\right]t^2dt - (M(T))^2 = a^2 \cdot \exp(d^2) \cdot (\exp(d^2) - 1).$$

Квадратичное отклонение

$$\sigma_T = \sqrt{D(T)} = M(T) \cdot \sqrt{\exp(d^2) - 1}\,.$$

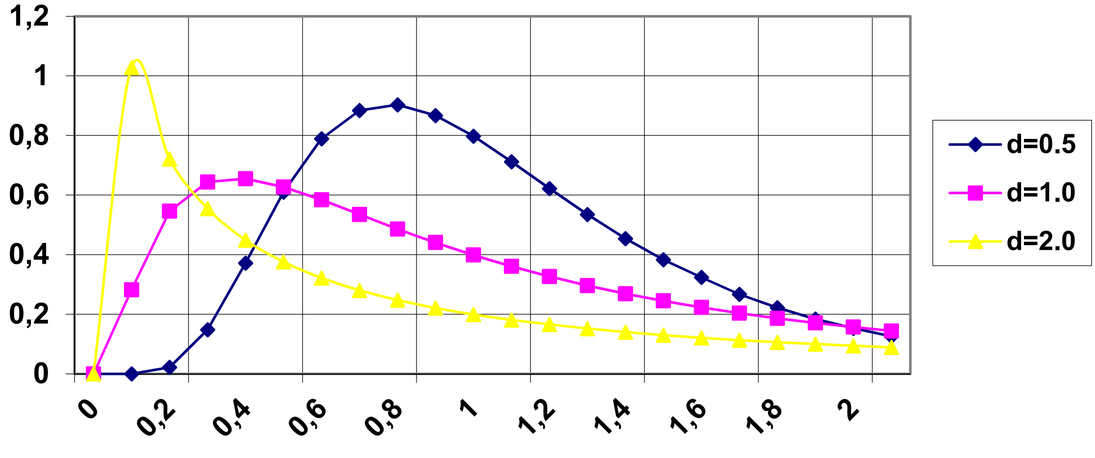


Рисунок 15.3 – График плотности логарифмически нормального распределения при значениях параметров: a=1.0, d=(0.5, 1.0, 2.0)

Коэффициент вариации

$$\nu_T = \frac{\sqrt{D(T)}}{M(T)} = \sqrt{\exp(d^2) - 1}\,.$$

Ассиметрия

$$S_k = \frac{\mu_3}{\sigma_T^3} = \nu_T \cdot (\exp(d^2) + 2)\,.$$

Мода

$$\mu_o = a\exp(-d^2)\,.$$

Функция распределения $F(t)$ выражается через рассмотренную ранее функцию нормированного нормального распределения $\Phi^*(x)$ следующим образом:

$$F(t) = \int_0^{\infty} \frac{1}{\sqrt{2\pi} td} \exp\left(-\frac{(\ln t - \ln a)^2}{2d^2}\right) dt = \int_{-\infty}^{x} \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{x^2}{2}\right) dx = \Phi^*\left(\frac{\ln t - \ln a}{\delta}\right),$$

где

$$x = \frac{\ln t - \ln a}{d}\,.$$

Здесь используется тот факт, что логарифм случайной величины *T*, распределенной по закону (15.2) имеет нормальное распределение с математическим ожиданием *a* и дисперсией d^2, по этому это распределение и называется логарифмически нормальным.

Если *X* имеет нормированное нормальное распределение то

$$T = a \cdot \exp(d \cdot X)$$

будет иметь логарифмически нормальное распределение. После логарифмирования получаем, что

$$X = \frac{\ln T - \ln a}{d}.$$

Медиана определяется из следующего уравнения:

$$F(\mu_e) = \Phi^*(\frac{\ln \mu_e - \ln a}{d}) = \frac{1}{2},$$

решение которого $\mu_e = a$.

Таким образом, параметр распределения *a* численно равен медиане распределения.

15.4 Контрольные вопросы

1. Как определяется плотность вероятности нормального закона распределения?

2. Как выглядит кривая распределения по нормальному закону?

3. Что такое Φ^*?

4. Какие основные свойства функции Φ^*?

5. В чем заключается правило трех сигм?

6. Что такое вероятное (срединное) отклонение случайной величины X?

7. Как выглядит плотность и функция показательного распределения положительной случайной величины *T*?

8. Где широко применяется показательное распределение?

9. Как определяется плотность вероятности логарифмически нормально-

го закона распределения?