Теория вероятностей и математическая статистика—2

Семинар Борзых Д.А.

Винер Даниил @danya_vin

17 января 2025 г.

Определение. Случайная величина X имеет нормальное распределение с параметрами $\mu \in \mathbb{R}$ и σ^2, если она имеет плотность вида

$$f_X(x) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}, \quad x \in \mathbb{R}$$

Обозначение: $X \sim \mathcal{N}(\mu, \sigma^2)$

Примечание. Если $X \sim \mathcal{N}(\mu, \sigma^2)$, то $\mathbb{E}[X] = \mu$ и $\mathbb{D}[X] = \sigma^2$

Определение. Если $X \sim \mathcal{N}(0, 1)$, то говорят, что X имеет *стандартное нормальное распределение*

Определение. Функцию распределения стандартной нормальной величины X обозначаем как:

$$\Phi(x) = \int\limits_{-\infty}^{x} \frac{1}{\sqrt{2\pi}} e^{-\frac{t^2}{2}} \mathrm{d}t$$

Теорема. $\forall x \in \mathbb{R}: \ \Phi(-x) = 1 - \Phi(x)$

Доказательство.

$$\begin{aligned}
\Phi(-x) &= \int\limits_{-\infty}^{-x} \frac{1}{\sqrt{2\pi}} e^{-\frac{t^2}{2}} \mathrm{d}t \\
&\text{пусть } -t = u|_{+\infty}^{x} \text{ и } \mathrm{d}t = -\mathrm{d}u \\
&= -\int\limits_{+\infty}^{x} \frac{1}{\sqrt{2\pi}} e^{-\frac{u^2}{2}} \mathrm{d}u \\
&= \int\limits_{x}^{+\infty} \frac{1}{\sqrt{2\pi}} e^{-\frac{u^2}{2}} \mathrm{d}u \\
&= 1 - \int\limits_{-\infty}^{x} \frac{1}{\sqrt{2\pi}} e^{-\frac{u^2}{2}} \mathrm{d}u \\
&= 1 - \Phi(x)
\end{aligned}$$

Пример 1

Пусть $X \sim \mathcal{N}(0, 1)$

а) Найдите $\mathbb{P}(-1 < X < 1)$

$$\begin{aligned}
\mathbb{P}(-1 < X < 1) &= \mathbb{P}(X < 1) - \mathbb{P}(X \leqslant -1) \\
&= \mathbb{P}(X \leqslant 1) - \mathbb{P}(X \leqslant -1) \\
&= \Phi(1) - \Phi(-1) \\
&= 2\Phi(1) - 1
\end{aligned}$$

Открываем таблицу нормального распределения, где в левом столбце целые и десятые доли числа, а сотые — в верхней строке. Получаем, что $\Phi(1) = 0.8413$ и $2\Phi(1) - 1 = 0.6826$

b) Найдите $\mathbb{P}(X \in [0;1])$

$$\begin{aligned}\mathbb{P}(X \in [0;1]) &= \mathbb{P}(X \leqslant 1) - \mathbb{P}(X \leqslant 0)\\ &= \Phi(1) - \Phi(0)\\ &= 0.8413 - 0.5\\ &= 0.3413\end{aligned}$$

c) Найдите $\mathbb{P}(X > 2)$

$$\begin{aligned}\mathbb{P}(X > 2) &= 1 - \mathbb{P}(X \leqslant 2)\\ &= 1 - \Phi(2)\\ &= 1 - 0.9772\\ &= 0.0228\end{aligned}$$

d) Найдите $\mathbb{P}(-2 < -X + 1 \leqslant 0)$

$$\begin{aligned}\mathbb{P}(-2 < -X + 1 \leqslant 0) &= \mathbb{P}(-3 < -X \leqslant -1)\\ &= \mathbb{P}(3 > X \leqslant 1)\\ &= \Phi(3) - \Phi(1)\\ &= 0.1574\end{aligned}$$

Теорема. Если $X \sim \mathcal{N}(\mu, \sigma^2)$, то $y = \dfrac{\overbrace{x - \mathbb{E}[X]}^{\text{центрирование}}}{\underbrace{\sqrt{\mathbb{D}[X]}}_{\text{нормирование}}} = \dfrac{x-\mu}{\sigma} \sim \mathcal{N}(0,1)$

Примечание. Стандартизация — применение центрирования и нормирования

Доказательство. $F_Y(x) = \mathbb{P}(Y \leqslant x) = \mathbb{P}\left(\frac{x-\mu}{\sigma} \leqslant x\right) = \mathbb{P}(X \leqslant \mu + \sigma x) = F_X(\mu + \sigma x)$

$f_Y(x) = f_X(\mu + \sigma x) \cdot \sigma$

$f_X(x) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}$

$f_Y(x) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(\mu+\sigma x-\mu)^2}{2\sigma^2}} \cdot \sigma = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}} \Longrightarrow Y \sim \mathcal{N}(0,1)$

Пример 2

Вычислить $\mathbb{P}(1 < X < 4)$. Сделаем двумя способами

1. Через MatLab: $\mathbb{P}(1 < X < 4) = F_X(4) - F_X(1) = \text{normcdf}(4, \mu, \sigma) - \text{normcdf}(1, \mu, \sigma) = 0.4332$
2. По таблице: $\mathbb{P}(1 < X < 4) = \mathbb{P}\left(\frac{1-1}{\sqrt{4}} < \frac{X-1}{\sqrt{4}} < \frac{4-1}{\sqrt{4}}\right) = F_X(1.5) - F_X(0) = \Phi(1.5) - \Phi(0) = 0.4332$

Теорема. Пусть $X \sim \mathcal{N}(0,1)$, тогда

- $\mathbb{E}[X] = 0$
- $\mathbb{E}\left[X^2\right] = 1$
- $\mathbb{E}\left[X^3\right] = 0$
- $\mathbb{E}\left[X^4\right] = 3$

Теорема. Если случайные величины $X \sim \mathcal{N}(\mu_X, \sigma_X^2),\ Y \sim \mathcal{N}(\mu_Y, \sigma_Y^2)$ независимые, то

$$aX + bY + c \sim \mathcal{N}(a\mu_X + b\mu_Y + c, a^2\sigma_Y^2 + b^2\sigma_Y^2)$$

Доказательство.

$$\begin{aligned}\mathbb{E}[aX + bY + c] &= a\mathbb{E}[X] + b\mathbb{E}[Y] + c\\ &= a\mu_X + b\mu_Y + c\end{aligned}$$

$$\begin{aligned}\mathbb{D}[aX + bY + c] &= a^2\mathbb{D}[X] + b^2\mathbb{D}[Y] + 2ab\text{cov}(aX, bY)\\ &= a^2\sigma_Y^2 + b^2\sigma_Y^2\end{aligned}$$

Пример 3

Даны $X \sim \mathcal{N}(\prime, \infty)$ и $Y \sim \mathcal{N}(3, 7)$ — независимые. Вычислим $\mathbb{P}\left(1 < 3X + Y < 5\right)$ двумя способами

- $\mathbb{P}\left(1 < 3X + Y < 5\right) = F_{3X+Y}(5) - F_{3X+Y}(1) = \text{normcdf}(5, 3, \sqrt{16}) - \text{normcdf}(1, 3, \sqrt{16}) = 0.3829$
  Тогда, $3X + Y \sim \mathcal{N}(3 \cdot 0 + 3, 9 \cdot 1 + 1 \cdot 7) = \mathcal{N}(3, 16)$
- Пусть $W = 3X + Y$, тогда

$$\begin{aligned} \mathbb{P}\left(1 < 3X + Y < 5\right) &= \mathbb{P}\left(\frac{1-3}{\sqrt{16}} \leqslant \frac{W-3}{\sqrt{16}} < \frac{5-3}{\sqrt{16}}\right) \\ &= \mathbb{P}\left(-\frac{1}{2} \leqslant \frac{W-3}{\sqrt{16}} < \frac{1}{2}\right) \\ &= \Phi(0.5) - \Phi(-0.5) \\ &= 2\Phi(0.5) - 1 \\ &= 0.383 \end{aligned}$$

Определение. Квантилью уровня $u \in (0; 1)$ для случайной величины X с плотностью распределения f_X называется наименьшее число $q \in \mathbb{R}$, такое что

$$\int\limits_{-\infty}^{q} f_X(x) \mathrm{d}x = u$$

Примечание. В MatLab $q =$norminv(u)

Пример 4

Нужно найти $q_{0.3}$. По таблице распределения его невозможно найти, однако, можно найти $q_{0.7}$, тогда $q_{0.3} = -q_{0.7}$

Для нахождения $q_{0.7}$ найдем в таблице распределения самое близкое к 0.7 число, это будет 0.6985. Оно соответствует $q_{0.7} = 0.52$