

Lec-7. 正态分布、随机变量的函数分布

主讲教师：吴利苏 (wulisu@sdust.edu.cn)

主 页：wulisu.cn

目录

1. 正态分布

- 标准正态分布

2. 随机变量的函数分布

- 离散型随机变量的函数分布
- 连续型随机变量的函数分布

正态分布 (高斯分布 Gauss)

定义

若连续型随机变量 X 的概率密度为

$$f(x) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(x-\mu)^2}{2\sigma^2}}, \quad -\infty < x < +\infty,$$

其中 $\mu, \sigma (\sigma > 0)$ 为常数, 则称 X 服从参数 μ, σ 的**正态分布**(或**高斯分布**), 记为 $X \sim N(\mu, \sigma^2)$.

- 非负性: $f(x) \geq 0$.
- 规范性: $\int_{-\infty}^{+\infty} f(x) dx = 1$.

证明：令 $\frac{(x-\mu)}{\sigma} = t$, 则

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

记 $I = \int_{-\infty}^{+\infty} e^{-\frac{t^2}{2}} dt$, 则

$$I^2 = \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} e^{-\frac{t^2+u^2}{2}} dt du.$$

取极坐标变换, 令 $t = r \cos \theta$, $u = r \sin \theta$, 则

$$I^2 = \int_0^{2\pi} \int_0^{+\infty} r e^{-\frac{r^2}{2}} dr d\theta = 2\pi.$$

$I > 0$, 则 $I = \sqrt{2\pi}$, 故 $\int_{-\infty}^{+\infty} f(x) dx = 1$.

性质 (正态分布的性质)

(1) $f(x)$ 关于 $x = \mu$ 对称, 故 $\forall h > 0$,

$$P\{\mu - h < X \leq \mu\} = P\{\mu < X \leq \mu + h\}.$$

性质 (正态分布的性质)

(1) $f(x)$ 关于 $x = \mu$ 对称, 故 $\forall h > 0$,

$$P\{\mu - h < X \leq \mu\} = P\{\mu < X \leq \mu + h\}.$$

(2) 当 $x \leq \mu$ 时, $f(x)$ 严格单调增.

性质 (正态分布的性质)

(1) $f(x)$ 关于 $x = \mu$ 对称, 故 $\forall h > 0$,

$$P\{\mu - h < X \leq \mu\} = P\{\mu < X \leq \mu + h\}.$$

(2) 当 $x \leq \mu$ 时, $f(x)$ 严格单调增.

(3)

$$f_{\max} = f(\mu) = \frac{1}{\sqrt{2\pi}\sigma},$$

且 x 离 μ 越远, $f(x)$ 值越小, 落在 x 附近的概率越小.

性质 (正态分布的性质)

(4) 在 $x = \mu \pm \sigma$ 处有拐点.

性质 (正态分布的性质)

(4) 在 $x = \mu \pm \sigma$ 处有拐点.

(5) 以 x 轴为渐近线, 即 $\lim_{x \rightarrow \pm\infty} f(x) = 0$.

性质 (正态分布的性质)

(4) 在 $x = \mu \pm \sigma$ 处有拐点.

(5) 以 x 轴为渐近线, 即 $\lim_{x \rightarrow \pm\infty} f(x) = 0$.

(6) 当 σ 固定, 改变 μ 的大小时, $f(x)$ 的图像形状不变, 整体沿 x 轴平移. μ 为位置参数, 决定对称轴的位置.

性质 (正态分布的性质)

(4) 在 $x = \mu \pm \sigma$ 处有拐点.

(5) 以 x 轴为渐近线, 即 $\lim_{x \rightarrow \pm\infty} f(x) = 0$.

(6) 当 σ 固定, 改变 μ 的大小时, $f(x)$ 的图像形状不变, 整体沿 x 轴平移. μ 为位置参数, 决定对称轴的位置.

(7) 当 μ 固定, 改变 σ 的大小时, $f(x)$ 的图像形状变, σ 越小, 图像越高越瘦; σ 越大, 图像越胖. σ 为尺度参数, 决定曲线分散程度.

正态分布的用途

- 自然界和人类社会中很多现象可以看成正态分布. 比如, 人的身高, 体重, 医学检验指标, 测量误差.
- 正态分布是最常见的一种分布. 一个变量如果受到大量微小的, 独立的随机因素的影响, 则一般是正态随机变量.
- 二项分布、泊松分布的极限分布是正态分布.
(第五章)

正态分布的概率计算

$X \sim N(\mu, \sigma^2)$, 则 X 的分布函数为

$$\begin{aligned} P\{X \leq x\} &= F(x) \\ &= \frac{1}{\sqrt{2\pi}\sigma} \int_{-\infty}^x e^{-\frac{(t-\mu)^2}{2\sigma^2}} dt \end{aligned}$$

正态分布的概率计算

$X \sim N(\mu, \sigma^2)$, 则 X 的分布函数为

$$P\{X \leq x\} = F(x)$$

$$= \frac{1}{\sqrt{2\pi}\sigma} \int_{-\infty}^x e^{-\frac{(t-\mu)^2}{2\sigma^2}} dt$$

$$=?$$

(无初等原函数)

正态分布的概率计算

$X \sim N(\mu, \sigma^2)$, 则 X 的分布函数为

$$\begin{aligned} P\{X \leq x\} &= F(x) \\ &= \frac{1}{\sqrt{2\pi}\sigma} \int_{-\infty}^x e^{-\frac{(t-\mu)^2}{2\sigma^2}} dt \quad (\text{无初等原函数}) \\ &=? \end{aligned}$$

- 用 Matlab, excel, R 语言等;
- 询问 Deepseek, Chatgpt;
- 数值积分;
- 转为标准正态分布, 查标准正态分布表.

标准正态分布

- 若 $Z \sim N(0, 1)$, 称 Z 服从标准正态分布.

标准正态分布

- 若 $Z \sim N(0, 1)$, 称 Z 服从标准正态分布.
- Z 的概率密度函数

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

标准正态分布

- 若 $Z \sim N(0, 1)$, 称 Z 服从标准正态分布.
- Z 的概率密度函数

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

- 分布函数

$$\Phi(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^x e^{-\frac{t^2}{2}} dt.$$

标准正态分布

- 若 $Z \sim N(0, 1)$, 称 Z 服从标准正态分布.
- Z 的概率密度函数

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

- 分布函数

$$\Phi(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^x e^{-\frac{t^2}{2}} dt.$$

- $\Phi(-x) + \Phi(x) = 1.$

性质 (标准化)

若 $X \sim N(\mu, \sigma^2)$, 则 $Z = \frac{X - \mu}{\sigma} \sim N(0, 1)$.

性质 (标准化)

若 $X \sim N(\mu, \sigma^2)$, 则 $Z = \frac{X-\mu}{\sigma} \sim N(0, 1)$.

证: $Z = \frac{X-\mu}{\sigma}$ 的分布函数为

$$\begin{aligned} P\{Z \leq x\} &= P\left\{\frac{X-\mu}{\sigma} \leq x\right\} = P\{X \leq \mu + \sigma x\} \\ &= \frac{1}{\sqrt{2\pi}\sigma} \int_{-\infty}^{\mu+\sigma x} e^{-\frac{(t-\mu)^2}{2\sigma^2}} dt \\ &\stackrel{\text{令 } u=\frac{t-\mu}{\sigma}}{=} \frac{1}{\sqrt{2\pi}} \int_{-\infty}^x e^{-\frac{u^2}{2}} du \\ &= \Phi(x). \end{aligned}$$

故 $Z = \frac{X-\mu}{\sigma} \sim N(0, 1)$.

□

推论

若 $X \sim N(\mu, \sigma^2)$, 则

$$\begin{aligned} F(x) &= P\{X \leq x\} = P\left\{\frac{X - \mu}{\sigma} \leq \frac{x - \mu}{\sigma}\right\} \\ &= \Phi\left(\frac{x - \mu}{\sigma}\right). \end{aligned}$$

推论

若 $X \sim N(\mu, \sigma^2)$, 则

$$\begin{aligned} F(x) &= P\{X \leq x\} = P\left\{\frac{X - \mu}{\sigma} \leq \frac{x - \mu}{\sigma}\right\} \\ &= \Phi\left(\frac{x - \mu}{\sigma}\right). \end{aligned}$$

对 $\forall (x_1, x_2]$, 有

$$\begin{aligned} P\{x_1 < X \leq x_2\} &= P\left\{\frac{x_1 - \mu}{\sigma} < \frac{X - \mu}{\sigma} \leq \frac{x_2 - \mu}{\sigma}\right\} \\ &= \Phi\left(\frac{x_2 - \mu}{\sigma}\right) - \Phi\left(\frac{x_1 - \mu}{\sigma}\right). \end{aligned}$$

证:

$$\begin{aligned}P\{x_1 < X \leq x_2\} &= \int_{x_1}^{x_2} \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(x-\mu)^2}{2\sigma^2}} dx \\&= \int_{\frac{x_1-\mu}{\sigma}}^{\frac{x_2-\mu}{\sigma}} \frac{1}{\sqrt{2\pi}} e^{-\frac{t^2}{2}} dt \\&= \int_{-\infty}^{\frac{x_2-\mu}{\sigma}} \frac{1}{\sqrt{2\pi}} e^{-\frac{t^2}{2}} dt - \int_{-\infty}^{\frac{x_1-\mu}{\sigma}} \frac{1}{\sqrt{2\pi}} e^{-\frac{t^2}{2}} dt \\&= \Phi\left(\frac{x_2-\mu}{\sigma}\right) - \Phi\left(\frac{x_1-\mu}{\sigma}\right). \quad \square\end{aligned}$$

例

用天平称一实际重量为 μ 的物体, 天平得读数为随机变量 X , 若 $X \sim N(\mu, \sigma^2)$, 求读数与 μ 的偏差在 3σ 范围内的概率.

例

用天平称一实际重量为 μ 的物体, 天平得读数为随机变量 X , 若 $X \sim N(\mu, \sigma^2)$, 求读数与 μ 的偏差在 3σ 范围内的概率.

解:

$$\begin{aligned} P\{|X - \mu| < 3\sigma\} &= P\{-3\sigma < X - \mu < 3\sigma\} \\ &= P\left\{-3 < \frac{X - \mu}{\sigma} < 3\right\} \\ &= \Phi(3) - \Phi(-3) \\ &= 2\Phi(3) - 1 \\ &\approx 0.9974. \end{aligned}$$

在 $(\mu - 3\sigma, \mu + 3\sigma)$ 内几乎是肯定的概率.

σ 法则

$X \sim N(\mu, \sigma^2)$, 则

- $P\{|X - \mu| < \sigma\} = \Phi(1) - \Phi(-1) = 68.26\%$;
- $P\{|X - \mu| < 2\sigma\} = \Phi(2) - \Phi(-2) = 95.44\%$;
- $P\{|X - \mu| < 3\sigma\} = \Phi(3) - \Phi(-3) = 99.74\%$.

例

将一温度调节器放置在储存着某种液体的容器内, 调节器调整在 $d^{\circ}\text{C}$, 液体的温度 X (以 $^{\circ}\text{C}$ 计) 是一个随机变量, 且 $X \sim N(d, 0.5^2)$.

- (1) 若 $d = 90^{\circ}\text{C}$, 求 X 小于 89°C 的概率.
- (2) 若要求保持液体的温度至少为 80°C 的概率不低于 0.99, 问 d 至少为多少?

解: (1)

$$\begin{aligned}P\{X < 89\} &= P\left\{\frac{X - 90}{0.5} < \frac{89 - 90}{0.5}\right\} \\&= \Phi(-2) = 1 - \Phi(2) \\&\approx 0.0228.\end{aligned}$$

(2)

$$\begin{aligned}P\{X \geq 80\} &= P\left\{\frac{X - d}{0.5} \geq \frac{80 - d}{0.5}\right\} \\&= 1 - P\left\{\frac{X - d}{0.5} < \frac{80 - d}{0.5}\right\} \\&= 1 - \Phi\left(\frac{80 - d}{0.5}\right) \\&= \Phi\left(\frac{d - 80}{0.5}\right) \geq 0.99 = \Phi(2.327).\end{aligned}$$

故 $d \geq 81.1635$.

随机变量的函数的分布

- 离散型随机变量的函数的分布：
- 连续型随机变量的函数的分布.

随机变量的函数的分布

- 已知 X 的分布, g 为连续函数. 如何求 $Y = g(X)$ 的分布?

随机变量的函数的分布

- 已知 X 的分布, g 为连续函数. 如何求 $Y = g(X)$ 的分布?
- 例如: 把圆半径的测量值记为随机变量 X , 若 $X \sim N(\mu, \sigma^2)$, 问圆的面积 $Y = \pi X^2$ 的分布是?

随机变量的函数的分布

- 已知 X 的分布, g 为连续函数. 如何求 $Y = g(X)$ 的分布?
- 例如: 把圆半径的测量值记为随机变量 X , 若 $X \sim N(\mu, \sigma^2)$, 问圆的面积 $Y = \pi X^2$ 的分布是?
- 类似, 在第三章第五节中学习两个随机变量的函数分布.

例

设 X 的分布律为

X	-1	0	1
P	0.1	0.6	0.3

$Y = X^2$, 求 Y 的分布律.

解: X 的取值为 $-1, 0, 1$, 则 $Y = X^2$ 可能的取值为 $0, 1$.

$$\text{又 } \{Y = 0\} = \{X = 0\},$$

$$\{Y = 1\} = \{X = 1\} \cup \{X = -1\}, \text{ 所以}$$

$$P\{Y = 0\} = 0.6, P\{Y = 1\} = 0.4,$$

X	0	1
P	0.6	0.4



离散型随机变量函数的分布函数

设离散型随机变量 X 的分布律为：

X	x_1	x_2	\dots	x_k	\dots
P_k	p_1	p_2	\dots	p_k	\dots

则 $Y = g(X)$

Y	$g(x_1)$	$g(x_2)$	\dots	$g(x_k)$	\dots
P_k	p_1	p_2	\dots	p_k	\dots

若 $g(x_k)$ 给出 Y 的所有可能取值, 再利用等价事件来给出概率分布函数 $P\{Y = y_j\} = P\{X \in D\}$.

例

X	-1	1	2
P	$\frac{1}{6}$	$\frac{2}{6}$	$\frac{3}{6}$

求 $Y = X^2 - 5$ 的分布律.

例

X	-1	1	2
P	$\frac{1}{6}$	$\frac{2}{6}$	$\frac{3}{6}$

求 $Y = X^2 - 5$ 的分布律.

解:

Y	-4	-1
P	$\frac{1}{2}$	$\frac{1}{2}$



例

设 X 的概率密度 $f_X(x) = \begin{cases} \frac{x}{8} & 0 < x < 4; \\ 0 & \text{其他.} \end{cases}$ 求
 $Y = 2X + 8$ 的概率密度.

例

设 X 的概率密度 $f_X(x) = \begin{cases} \frac{x}{8} & 0 < x < 4; \\ 0 & \text{其他.} \end{cases}$ 求

$Y = 2X + 8$ 的概率密度.

思路：先求 Y 的分布函数, 再求导得概率密度.

$$\frac{dF(g(y))}{dy} = \left[\int_{-\infty}^{g(y)} f(x) dx \right]' = f(g(y)) \cdot g'(y).$$

解: 先求 $Y = 2X + 8$ 的分布函数 $F_Y(y)$,

$$\begin{aligned} F_Y(y) &= P\{Y \leq y\} = P\{2X + 8 \leq y\} \\ &= P\left\{X \leq \frac{y-8}{2}\right\} = \int_{-\infty}^{\frac{y-8}{2}} f_X(x) dx \end{aligned}$$

由分布函数 F_Y 求 f_Y .

$$\begin{aligned} f_Y(y) = F'_Y(y) &= \left[\int_{-\infty}^{\frac{y-8}{2}} f_X(x) dx \right]' = f_X\left(\frac{y-8}{2}\right) \cdot \left(\frac{y-8}{2}\right)' \\ &= \begin{cases} \frac{1}{8} \left(\frac{y-8}{2}\right) \cdot \left(\frac{y-8}{2}\right)' & 0 < \frac{y-8}{2} < 4; \\ 0 & \text{其他.} \end{cases} \\ &= \begin{cases} \frac{y-8}{32} & 8 < y < 16; \\ 0 & \text{其他.} \end{cases} \end{aligned}$$

□

例

设 X 的概率密度为

$$f_X(x) = \begin{cases} 0 & x < 0; \\ 2x^3 e^{-x^2} & x \geq 0. \end{cases}$$

求 $Y = X^2$ 和 $Y = 2X + 3$ 的概率密度.

解: ($Y = X^2$)

显然 $y \geq 0$ 时,

$$F_Y(y) = P\{Y \leq y\} = P\{X^2 \leq y\}$$

$$= P\{-\sqrt{y} \leq X \leq \sqrt{y}\} = P\{0 \leq X \leq \sqrt{y}\} = \int_0^{\sqrt{y}} f_X(x) dx.$$

$$f_Y(y) = F'_Y(y) = f_X(\sqrt{y})(\sqrt{y})' = \begin{cases} 0 & y < 0; \\ ye^{-y} & y \geq 0. \end{cases}$$

$$(Y = 2X + 3)$$

$$\begin{aligned} F_Y(y) &= P\{Y \leq y\} = P\{2X + 3 \leq y\} \\ &= P\{X \leq \frac{y-3}{2}\} = \int_{-\infty}^{\frac{y-3}{2}} f_X(x) dx \end{aligned}$$

$$\begin{aligned} f_Y(y) &= F'_Y(y) = f_X\left(\frac{y-3}{2}\right) \cdot \left(\frac{y-3}{2}\right)' \\ &= \begin{cases} 0 & y < 3; \\ \left(\frac{y-3}{2}\right)^3 e^{-\left(\frac{y-3}{2}\right)^2} & y \geq 3. \end{cases} \quad \square \end{aligned}$$