

Lec-14. 方差

主讲教师：吴利苏 (wulisu@sdust.edu.cn)

主页：wulisu.cn

本次课内容

1. 方差的定义

2. 方差的计算

3. 方差的性质

4. 切比雪夫不等式

引理

例

有两批灯泡，寿命分布如下：

X	$X < 950$	$950 \leq X \leq 1050$	$X > 1050$
n	0.005	0.99	0.005

Y	$Y = 700$	$700 < Y < 1300$	$Y = 1300$
n	0.5	0	0.5

假设平均寿命都是 $E(X) = E(Y) = 1000h$ ，
如何判定这两批灯泡的质量好坏。

- 随机变量 X 的均值： $E(X)$
- X 对于均值的离差： $X - E(X)$
- X 对于均值的平均离差： $E(X - E(X)) = 0$

- 随机变量 X 的均值： $E(X)$
- X 对于均值的离差： $X - E(X)$
- X 对于均值的平均离差： $E(X - E(X)) = 0$
- 反映随机变量波动性可以用：

$$E[X - E(X)]^2$$

\Rightarrow 方差.

方差

定义

设 X 是一个随机变量, 若 $E\{[X - E(X)]^2\}$ 存在, 则称它为 X 的 **方差**, 记为 $D(X)$, 或 $Var(X)$. 即

$$D(X) = Var(X) = E\{[X - E(X)]^2\}$$

将 $\sqrt{D(X)}$ 记为 $\sigma(X)$, 称为 **标准差** 或 **均方差**.

方差

定义

设 X 是一个随机变量, 若 $E\{[X - E(X)]^2\}$ 存在, 则称它为 X 的**方差**, 记为 $D(X)$, 或 $Var(X)$. 即

$$D(X) = Var(X) = E\{[X - E(X)]^2\}$$

将 $\sqrt{D(X)}$ 记为 $\sigma(X)$, 称为**标准差**或**均方差**.

$D(X)$ 或 $\sigma(X)$ 体现 X 取值的波动性, 是衡量 X 取值分散程度的数字特征. 若 $D(X)$ 较小, 则 X 取值比较集中; 反之, 若 $D(X)$ 越大, 则说明 X 取值较分散.

方差的计算

- 令 $g(X) = [X - E(X)]^2$, 则

$$D(X) = E(g(X)).$$

方差是 X 的函数 $g(X)$ 的数学期望.

方差的计算

- 离散型: $P\{X = x_k\} = p_k,$

$$D(X) = \sum_{k=1}^{\infty} (x_k - E(X))^2 p_k.$$

- 连续型:

$$D(X) = \int_{-\infty}^{+\infty} (x - E(X))^2 f(x) dx.$$

常见离散型随机变量的方差

- 两点分布 $X \sim 0-1(p)$

$$E(X) = p, D(X) = p(1 - p).$$

- 二项分布 $X \sim b(n, p)$

$$E(X) = np, D(X) = np(1 - p).$$

- 泊松分布 $X \sim \pi(\lambda)$

$$E(X) = \lambda, D(X) = \lambda.$$

- 几何分布 $X \sim Geom(p)$

$$E(X) = \frac{1}{p}, D(X) = \frac{p - 1}{p^2}.$$

常见连续型随机变量的方差

- 均匀分布 $X \sim U(a, b)$

$$E(X) = \frac{a+b}{2}, D(X) = \frac{(b-a)^2}{12}.$$

- 指数分布 $X \sim E(\theta)$

$$E(X) = \theta, D(X) = \theta^2.$$

- 正态分布 $X \sim N(\mu, \sigma^2)$

$$E(X) = \mu, D(X) = \sigma^2.$$

性质

$$D(X) = E(X^2) - [E(X)]^2.$$

性质

$$D(X) = E(X^2) - [E(X)]^2.$$

证明:

$$\begin{aligned} D(X) &= E\{[X - E(X)]^2\} \\ &= E\{X^2 - 2XE(X) + (E(X))^2\} \\ &= E(X^2) - 2E(X)E(X) + (E(X))^2 \\ &= E(X^2) - [E(X)]^2. \end{aligned}$$

□

例

设 $E(X) = \mu$, $D(X) = \sigma^2 \neq 0$. 记

$$X^* = \frac{X - \mu}{\sigma},$$

$$\text{则 } E(X^*) = \frac{1}{\sigma} E(X - \mu) = 0$$

$$\begin{aligned} D(X^*) &= E(X^{*2}) - [E(X^*)]^2 \\ &= E\left[\left(\frac{X - \mu}{\sigma}\right)^2\right] = \frac{1}{\sigma^2} E((X - \mu)^2) = 1. \end{aligned}$$

X^* 为 X 的标准化变量.

方差的性质

性质

1. 设 C 是常数, 则有 $D(C) = 0$.
2. 设 X 是随机变量, C 是常数, 则

$$D(CX) = C^2 D(X), D(X + C) = D(X)$$

特别地, $D(X) = D(-X)$.

方差的性质

性质

3. 设 X, Y 是两个随机变量, 则有

$$D(X+Y) = D(X) + D(Y) + 2E\{[X-E(X)][Y-E(Y)]\}.$$

特别地, 若 X 与 Y 独立, 则

$$D(X+Y) = D(X) + D(Y).$$

证明:

$$\begin{aligned}D(X + Y) &= E\{[(X + Y) - E(X + Y)]^2\} \\&= E\{[(X - E(X)) + (Y - E(Y))]^2\} \\&= D(X) + D(Y) + 2E\{[X - E(X)][Y - E(Y)]\}\end{aligned}$$

□

证明:

$$\begin{aligned}D(X + Y) &= E\{[(X + Y) - E(X + Y)]^2\} \\&= E\{[(X - E(X)) + (Y - E(Y))]^2\} \\&= D(X) + D(Y) + 2E\{[X - E(X)][Y - E(Y)]\}\end{aligned}$$

□

- 由性质 1-3, 若 X, Y 相互独立, 则

$$D(aX + bY + C) = a^2 D(X) + b^2 D(Y).$$

- 进一步, 若 X_1, \dots, X_n 相互独立, 则

$$D(c_0 + \sum_i^n c_i X_i) = \sum_i^n c_i^2 D(X_i).$$

方差的性质

性质

4. $D(X) = 0 \Leftrightarrow P\{X = E(X)\} = 1.$

方差的性质

性质

4. $D(X) = 0 \Leftrightarrow P\{X = E(X)\} = 1.$

证明: (充分性) $P\{X = E(X)\} = 1$, 则有
 $P\{X^2 = (E(X))^2\} = 1$, 则

$$D(X) = E(X^2) - (E(X))^2 = 0.$$

(必要性) 通过切比雪夫不等式证明.

□

例

设 X 的概率密度

$$f(x) = \begin{cases} 1+x & -1 \leq x < 0; \\ 1-x & 0 \leq x < 1; \\ 0 & \text{其他.} \end{cases}$$

求 $D(X)$.

例

设 X 的概率密度

$$f(x) = \begin{cases} 1+x & -1 \leq x < 0; \\ 1-x & 0 \leq x < 1; \\ 0 & \text{其他.} \end{cases}$$

求 $D(X)$.

$$\text{解: } E(X) = \int_{-1}^0 x(1+x)dx + \int_0^1 x(1-x)dx = 0$$

$$E(X^2) = \int_{-1}^0 x^2(1+x)dx + \int_0^1 x^3(1-x)dx = \frac{1}{6}$$

$$D(X) = E(X^2) - (E(X))^2 = \frac{1}{6}.$$

□

例

设活塞的直径 (以 cm 计) $X \sim N(22.40, 0.03^2)$,
气缸的直径为 $Y \sim N(22.50, 0.04^2)$. X, Y 相互独立, 任取一只活塞, 求活塞能装入气缸的概率.

例

设活塞的直径 (以 cm 计) $X \sim N(22.40, 0.03^2)$, 气缸的直径为 $Y \sim N(22.50, 0.04^2)$. X, Y 相互独立, 任取一只活塞, 求活塞能装入气缸的概率.

解: $X - Y \sim N(-0.10, 0.05^2)$

$$\begin{aligned} P\{X < Y\} &= P\{X - Y < 0\} \\ &= P\left\{\frac{X - Y - (-0.1)}{0.05} < \frac{0 - (-0.1)}{0.05}\right\} \\ &= \Phi(2) = 0.9772. \end{aligned}$$

例

$$f(x) = \begin{cases} \cos x & 0 \leq x \leq \frac{\pi}{2}; \\ 0 & \text{其他.} \end{cases}$$

求 $Y = X^2$ 的方差 $D(Y)$.

例

$$f(x) = \begin{cases} \cos x & 0 \leq x \leq \frac{\pi}{2}; \\ 0 & \text{其他.} \end{cases}$$

求 $Y = X^2$ 的方差 $D(Y)$.

$$\text{解: } E(X^2) = \int_{-\infty}^{+\infty} x^2 f(x) dx = \int_0^{\frac{\pi}{2}} x^2 \cos x dx = \frac{\pi^2}{4} - 2$$

$$E(X^4) = \int_{-\infty}^{+\infty} x^4 f(x) dx = \int_0^{\frac{\pi}{2}} x^4 \cos x dx = \frac{\pi^4}{16} - 3\pi^2 + 24$$

$$D(X^2) = E(X^4) - (E(X^2))^2 = 20 - 2\pi^2.$$

□

例

设

X	-2	0	1	3
P	$\frac{1}{3}$	$\frac{1}{2}$	$\frac{1}{12}$	$\frac{1}{12}$

求 $D(2X^3 + 5)$.

例

设

X	-2	0	1	3
P	$\frac{1}{3}$	$\frac{1}{2}$	$\frac{1}{12}$	$\frac{1}{12}$

求 $D(2X^3 + 5)$.

解: $E(X^6) = \frac{493}{6}$, $E(X^3)^2 = \frac{1}{9}$

$$D(2X^3 + 5) = 4D(X^3) = 4(E(X^6) - (E(X^3))^2) = \frac{2954}{9}. \quad \square$$

切比雪夫不等式 (Chebyshev)

定理

设 X 具有 $E(X) = \mu$, $D(X) = \sigma^2$, 则对任意正数 ϵ , 有

$$P\{|X - \mu| \geq \epsilon\} \leq \frac{\sigma^2}{\epsilon^2}$$

切比雪夫不等式 (Chebyshev)

定理

设 X 具有 $E(X) = \mu$, $D(X) = \sigma^2$, 则对任意正数 ϵ , 有

$$P\{|X - \mu| \geq \epsilon\} \leq \frac{\sigma^2}{\epsilon^2}$$

意义：分布未知, $E(X)$ 和 $D(X)$ 存在的条件下, 估计

$$P\{|X - E(X)| < \epsilon\} \geq 1 - \frac{\sigma^2}{\epsilon^2}.$$

切比雪夫不等式 (Chebyshev)

证:

$$\begin{aligned} P\{|X - \mu| \geq \epsilon\} &= \int_{|x - \mu| \geq \epsilon} f(x) dx \\ &\leq \int_{|x - \mu| \geq \epsilon} \frac{|x - \mu|^2}{\epsilon^2} f(x) dx \\ &\leq \frac{1}{\epsilon^2} \int_{-\infty}^{+\infty} (x - \mu)^2 f(x) dx = \frac{\sigma^2}{\epsilon^2}. \end{aligned}$$

□

方差性质 4 的证明

性质

设 $D(X) = 0$, 则 $P\{X = E(X)\} = 1$.

证明: (反证法) 假设 $P\{X = E(X)\} < 1$, 则存在 $\epsilon > 0$, 有

$$P\{|X - E(X)| \geq \epsilon\} > 0.$$

但由切比雪夫不等式, 对 $\forall \epsilon > 0$, 有

$$P\{|X - E(X)| \geq \epsilon\} = 0.$$

矛盾. 故 $P\{X = E(X)\} = 1$.

常见离散型随机变量的方差

- 两点分布 $X \sim 0-1(p)$

$$E(X) = p, D(X) = p(1 - p).$$

- 二项分布 $X \sim b(n, p)$

$$E(X) = np, D(X) = np(1 - p).$$

- 泊松分布 $X \sim \pi(\lambda)$

$$E(X) = \lambda, D(X) = \lambda.$$

- 几何分布 $X \sim Geom(p)$

$$E(X) = \frac{1}{p}, D(X) = \frac{p - 1}{p^2}.$$

常见连续型随机变量的方差

- 均匀分布 $X \sim U(a, b)$

$$E(X) = \frac{a+b}{2}, D(X) = \frac{(b-a)^2}{12}.$$

- 指数分布 $X \sim E(\theta)$

$$E(X) = \theta, D(X) = \theta^2.$$

- 正态分布 $X \sim N(\mu, \sigma^2)$

$$E(X) = \mu, D(X) = \sigma^2.$$