

§2.7 随机变量的方差及其他数字特征

- 定义7.1 & 7.2.. 假设 EX 存在, 且 $E(X - EX)^2$ 也存在. 则称 $E(X - EX)^2$ 为 X 的方差, 记为 $\text{var}(X)$ 或 $D(X)$.

称 $\sqrt{\text{var}(X)}$ 为标准差.

- 定理7.1 (切比雪夫不等式) 假设 $\text{var}(X)$ 都存在, 则 $\forall \varepsilon > 0$, 有

$$P(|X - EX| \geq \varepsilon) \leq \frac{1}{\varepsilon^2} \text{var}(X).$$

- 证: $\{|X - EX| \geq \varepsilon\} = \{(X - EX)^2 \geq \varepsilon^2\}$,

对 $Y = (X - EX)^2$ 用马尔可夫不等式.

- 推论7.1. 若 $\text{var}(X) = 0$, 则 X 退化.

- 证: $Y \geq 0$ 且 $EY = 0$, 故 $Y \equiv 0$, 即 $X \equiv c = EX$.

• 定理7.2. $\text{var}(X) = EX^2 - (EX)^2$.

• 证:

$$\text{var}(X) = E(X^2 - 2X \cdot EX + (EX)^2) = EX^2 - (EX)^2.$$

• 具体地, 离散型或连续型的公式如下:

$$\text{var}(X) = \sum_k x_k^2 p_k - (EX)^2,$$

$$\text{var}(X) = \int_{-\infty}^{\infty} x^2 p(x) dx - (EX)^2.$$

• X 的线性变换的方差:

$$\text{var}(aX + b) = a^2 \text{var}(X).$$

(3) 泊松分布.

$$P(X = k) = \frac{\lambda^k}{k!} e^{-\lambda}, \quad k = 0, 1, 2, \dots$$

- $EX = \lambda$, 且 $\forall k \geq 1, kp_k = \lambda p_{k-1}$. 因此,

$$k^2 p_k = \lambda k p_{k-1} = \lambda p_{k-1} + \lambda(k-1)p_{k-1} = \lambda^2 p_{k-2}, \quad \forall k \geq 2.$$

- 或者, $\forall k \geq 2$,

$$k(k-1)p_k = \lambda(k-1)p_{k-1} = \lambda^2 p_{k-2}.$$

- 于是, $EX(X-1) = \lambda^2$, 从而

$$\text{var}(X) = EX^2 - (EX)^2 = EX(X-1) + EX - (EX)^2 = \lambda.$$

(2) 二项分布.

$$P(X = k) = C_n^k p^k q^{n-k} =: b(n; k), \quad k = 0, 1, \dots, n, (q = 1 - p).$$

• $EX = np$, 且 $\forall 1 \leq k \leq n$,

$$k \cdot b(n; k) = np \cdot b(n-1, k-1).$$

• $\forall 2 \leq k \leq n$,

$$\begin{aligned} k(k-1) \cdot b(n; k) &= np \cdot (k-1) \cdot b(n-1, k-1) \\ &= np \cdot (n-1)p \cdot b(n-2, k-2) \end{aligned}$$

• 于是, $EX(X-1) = np(n-1)p = (np)^2 - np^2$, 从而

$$D(X) = EX^2 - (EX)^2 = EX(X-1) + EX - (EX)^2 = npq.$$

(6) 正态分布.

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

- 若 $\mu = EX = 0$, $\sigma^2 = 1$, 则,

$$\begin{aligned} \text{var}(X) &= EX^2 = \int_{-\infty}^{\infty} x^2 \cdot \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}} dx \\ &= -\frac{2}{\sqrt{2\pi}} \int_0^{\infty} x de^{-\frac{x^2}{2}} = \frac{2}{\sqrt{2\pi}} \int_0^{\infty} e^{-\frac{x^2}{2}} dx = 1. \end{aligned}$$

- 一般情形, $Y = \frac{X-\mu}{\sigma} \sim N(0, 1)$. 则

$$\begin{aligned} X - EX &= (\mu + \sigma Y) - (\mu + \sigma EY) = \sigma(Y - EY) \\ \Rightarrow \text{var}(X) &= E(\sigma(Y - EY))^2 = \sigma^2 \text{var}(Y) = \sigma^2. \end{aligned}$$

- 一般地, 若 X 的方差存在, 且 $\text{var}(X) > 0$, 则

$$X^* = \frac{X - EX}{\sqrt{\text{var}(X)}}$$

满足 $EX^* = 0$, $\text{var}(X^*) = 1$. 称 X^* 为 X 的标准化.

- 定义7.3. k 阶(原点)矩指 EX^k .

定义7.4. k 阶中心矩指 $E(X - EX)^k$.

- 定义7.5. 若

$$P(X < a) \leq p \leq P(X \leq a),$$

则称 a 为 X 的一个 p 分位数.

$p = 0.5$ 时, 也称 a 为一个中位数.