



Chapter 4: Expected Values (期望值)

- The Expected Value of a Random Variable (随机变量的期望)
- Variance and Standard Deviation (方差和标准差)
- Covariance and Correlation Coefficient (协方差和相关系数)
- Conditional Expectation (条件期望)



Chapter 4: Expected Values (期望值)

- The Expected Value of a Random Variable(随机变量的期望)
- **Variance and Standard Deviation(方差和标准差)**
- Covariance and Correlation Coefficient (协方差和相关系数)
- Conditional Expectation (条件期望)



Background

Example: there are two watches with daily accuracies shown as follows:

Daily Accuracy (s)	-3	-2	-1	0	1	2	3
Probability (Watch A)	0.10	0.15	0.15	0.20	0.15	0.15	0.10
Probability (Watch B)	0.05	0.05	0.10	0.60	0.10	0.05	0.05

Which watch has better quality?

Answer: Assume that the accuracy of these two watches are X and Y , then

$$E(X) = \sum_{k=1}^7 x_k P\{X = x_k\} = 0$$
$$E(Y) = \sum_{k=1}^7 y_k P\{Y = y_k\} = 0$$

Therefore, the average accuracies from the two watches are the same.

Are they of the same quality?





§ 2 方差和标准差

4

Thinking based on the deviation (偏差) from the average

The deviation (偏差) of r.v. X is

$$|X - E(X)|$$

If the **deviation** is small, the quality is stable

Calculation based on absolute value is **not** convenient

Consider the squared deviation (平方偏差)

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

Squared deviation is still a **r.v.**

Consider the average of squared deviation

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

Variance (方差) represents the average squared deviation between a r.v. and its average.

Definition: For r.v. X , if the following exists

$$\text{Var}(X) \triangleq D(X) \triangleq \underline{E(X - E(X))^2}$$

Then $D(X)$ is the **Variance (方差)** of r.v. X , $\sqrt{D(x)}$ is the **Standard Deviation (标准差)**



§ 2 方差和标准差

5

Example: Assume that the scores of shooters A and B are X and Y respectively. The frequency functions are:

x_k	8	9	10
p_k	0.15	0.40	0.45

y_k	8	9	10
p_k	0.35	0.10	0.55

Please evaluate these shooters.

Answer: First find out the **expected values**:

$$E(X) = 8 \times 0.15 + 9 \times 0.40 + 10 \times 0.45 = 9.3$$

$$E(Y) = 8 \times 0.35 + 9 \times 0.10 + 10 \times 0.55 = 9.2$$

And then find out the **variance** $\text{Var}(X) \triangleq D(X) \triangleq E(X - E(X))^2$

$$D(X) = (8 - 9.3)^2 \times 0.15 + (9 - 9.3)^2 \times 0.40 + (10 - 9.3)^2 \times 0.45 = 0.51$$

$$D(Y) = (8 - 9.2)^2 \times 0.35 + (9 - 9.2)^2 \times 0.10 + (10 - 9.2)^2 \times 0.55 = 0.86$$

Based on the results, shooter A is not only better than shooter B in terms of shooting level, but also the score is more stable.



§ 2 方差和标准差

6

Example: there are two watches with daily accuracies shown as follows:

Daily Accuracy (s)	-3	-2	-1	0	1	2	3
Probability (Watch A)	0.10	0.15	0.15	0.20	0.15	0.15	0.10
Probability (Watch B)	0.05	0.05	0.10	0.60	0.10	0.05	0.05

Which watch has better quality?

Answer: Assume that the accuracies of the two watches A, B are X and Y , respectively. then

$$E(X) = \sum_{k=1}^7 x_k P\{X = x_k\} = 0$$

$$E(Y) = \sum_{k=1}^7 y_k P\{Y = y_k\} = 0$$

$$D(X) = E(X - E(X))^2 = E(X^2) = 3.3$$

$$D(Y) = E(Y - E(Y))^2 = E(Y^2) = 1.5$$

As above, the average accuracies are the same, but their variances are different. So, Watch B is better than Watch A because it is more stable.



§ 2 方差和标准差

7



Explanation: **Expected value** ——— The average value of a r.v.

Variance ——— The average deviation between a r.v. and its average.



The calculation of variance

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

is the expected value of $g(X) = (X - E(X))^2$, then



Assume that the **frequency** function of X is

$$P\{X = x_i\} = p_k, k = 1, 2, \dots$$

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



Assume that the probability **density** of X is $f(x)$, then

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



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



Example: Assume that $X \sim P(\lambda)$, find $D(X)$

Answer: Based on the calculation from the last section, we have $E(X) = \lambda$

$$\begin{aligned} E(X^2) &= E[X(X-1)] + E(X) \\ &= \sum_{k=0}^{\infty} k(k-1) \frac{\lambda^k}{k!} e^{-\lambda} + \lambda \\ &= \lambda^2 e^{-\lambda} \sum_{k=2}^{\infty} \frac{\lambda^{k-2}}{(k-2)!} + \lambda \\ &= \lambda^2 + \lambda \end{aligned}$$

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



§ 2 方差和标准差

9

Example: Assume that $X \sim U(a, b)$, find $D(X)$

Answer: Based on the calculation from the last section, we have $E(X) = \frac{a+b}{2}$, the density of X is

$$f(x) = \begin{cases} \frac{1}{b-a}, & a < x < b \\ 0, & \text{otherwise} \end{cases}$$

$$\begin{aligned} \therefore E(X^2) &= \int_a^b \frac{x^2}{b-a} dx \\ &= \frac{b^3 - a^3}{3(b-a)} = \frac{a^2 + ab + b^2}{3} \end{aligned}$$

$$\begin{aligned} \therefore D(X) &= E(X^2) - [E(X)]^2 \\ &= \frac{a^2 + ab + b^2}{3} - \left(\frac{a+b}{2}\right)^2 \\ &= \frac{(b-a)^2}{12} \end{aligned}$$



Example: Assume that the density function of X is

$$f(x) = \begin{cases} \frac{1}{\theta} e^{-x/\theta}, & x > 0 \\ 0, & x \leq 0 \end{cases}$$

Find $D(X)$.

Answer: Since it is the exponential distribution, then $E(X) = \theta$, so:

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

$$= \int_0^{\infty} x^2 \frac{1}{\theta} e^{-x/\theta} dx - \theta^2$$

$$= -x^2 e^{-x/\theta} \Big|_0^{\infty} + 2 \int_0^{\infty} x e^{-x/\theta} dx - \theta^2$$

$$= 2\theta^2 - \theta^2 = \theta^2$$



Basic properties of variance (方差的基本性质)

- ① If $X =_{a.e.} c$ (constant), then $D(X) = 0$.
- ② If c is a constant, then $D(cX) = c^2 D(X)$.

Proof:

$$\begin{aligned} D(cX) &= E(cX - E(cX))^2 \\ &= E(cX - cE(X))^2 \\ &= c^2 E(X - E(X))^2 \\ &= c^2 D(X) \end{aligned}$$



Basic properties of variance (方差的基本性质)

① If $X =_{a.e.} c$ (constant), then $D(X) = 0$.

② If c is a constant, then $D(cX) = c^2 D(X)$.

③ For r.v.s X and Y :

$$D(X + Y) = D(X) + D(Y) + 2E[(X - E(X))(Y - E(Y))]$$

Proof: $D(X + Y) = E[(X + Y) - E(X + Y)]^2$

$$= E[(X - E(X)) + (Y - E(Y))]^2$$

$$= E[X - E(X)]^2 + E[Y - E(Y)]^2$$

$$+ 2E[(X - E(X))(Y - E(Y))]$$

$$= D(X) + D(Y) + 2E[(X - E(X))(Y - E(Y))]$$



Basic properties of variance (方差的基本性质)

① If $X =_{a.e.} c$ (constant), then $D(X) = 0$.

② If c is a constant, then $D(cX) = c^2 D(X)$.

③ For r.v.s X and Y :

$$D(X + Y) = D(X) + D(Y) + 2E[(X - E(X))(Y - E(Y))]$$

If X and Y are independent, then $D(X + Y) = D(X) + D(Y)$

Proof: $\because X$ and Y are independent

$\therefore X - E(X)$ and $Y - E(Y)$ are independent

$$\begin{aligned}\therefore E[(X - E(X))(Y - E(Y))] &= E[X - E(X)] \cdot E[Y - E(Y)] \\ &= [E(X) - E(X)] \cdot [E(Y) - E(Y)] = 0\end{aligned}$$



If X and Y are independent, is the following right?

$$D(X - Y) = D(X) \oplus D(Y) \quad ?$$



Example: Assume that $X \sim b(n, p)$, find $D(X)$

Answer: Since binomial distribution is from n -fold Bernoulli trials, then

$$X = X_1 + X_2 + \cdots + X_n$$

where

$$X_i = \begin{cases} 1, & \text{A happened in the } i\text{th Bernoulli trial} \\ 0, & \bar{A} \text{ happened in the } i\text{th Bernoulli trial} \end{cases} \quad (i = 1, 2, \dots, n)$$

And X_1, X_2, \dots, X_n are independent and identically distributed (**i.i.d.**) with frequency functions:

$$P\{X_i = 1\} = p, P\{X_i = 0\} = 1 - p.$$

$$\therefore E(X) = E(X_1) + E(X_2) + \cdots + E(X_n)$$

$$= nE(X_1) = np$$

$$D(X) = D(X_1) + D(X_2) + \cdots + D(X_n)$$

$$= nD(X_1) = n[(1 - p)^2 \cdot p + (0 - p)^2 \cdot (1 - p)]$$

$$= np(1 - p)$$



§ 2 方差和标准差

15

Example: Assume that $X \sim N(\mu, \sigma^2)$, find $D(X)$

Answer: Based on the last section $E(X) = \mu$, then

$$\begin{aligned} D(X) &= \int_{-\infty}^{\infty} (x - \mu)^2 f(x) dx \\ &= \frac{1}{\sqrt{2\pi}\sigma} \int_{-\infty}^{\infty} (x - \mu)^2 e^{-\frac{(x-\mu)^2}{2\sigma^2}} dx \\ &= \frac{\sigma^2}{\sqrt{2\pi}} \int_{-\infty}^{\infty} t^2 e^{-\frac{t^2}{2}} dt \\ &= \frac{\sigma^2}{\sqrt{2\pi}} \int_{-\infty}^{\infty} (-t) de^{-\frac{t^2}{2}} \\ &= \frac{\sigma^2}{\sqrt{2\pi}} \left[(-t)e^{-\frac{t^2}{2}} \Big|_{-\infty}^{\infty} + \int_{-\infty}^{\infty} e^{-\frac{t^2}{2}} dt \right] \\ &= \frac{\sigma^2}{\sqrt{2\pi}} \sqrt{2\pi} = \sigma^2 \end{aligned}$$

let $t = \frac{x-\mu}{\sigma}$



§ 2 方差和标准差

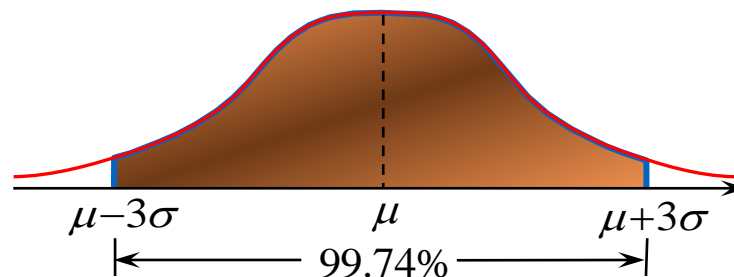
16

3 σ Principle: Most values of the normal r.v. are in the range $(\mu - 3\sigma, \mu + 3\sigma)$

If $X \sim N(\mu, \sigma^2)$, then

$$P\{|X - \mu| < 3\sigma\} = 0.9974$$

$$P\{|X - \mu| \geq 3\sigma\} = 0.0026$$



Generally, how to find the following probability of r.v. X :

$$P\{|X - \mu| \geq \varepsilon\}$$



where $\mu = E(X)$, $\varepsilon > 0$ are constants.

Theorem: Chebyshev's inequality (切比雪夫不等式)

If $\mu \triangleq E(X)$ and $\sigma^2 = D(X)$ both exist, then for $\forall \varepsilon > 0$, we have:

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



Theorem: Chebyshev's inequality (切比雪夫不等式)

If $\mu \triangleq E(X)$ and $\sigma^2 = D(X)$ both exist, then for $\forall \varepsilon > 0$, we have:

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

Proof: Only prove for continuous r.v.s, discrete r.v.s are similar. Let $f(x)$ be the density function of X , then:

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

Or we can let $Y = (X - \mu)^2$, then $E(Y) = \sigma^2$, applying the **Markov Inequality**:

$$P\{|X - \mu| \geq \varepsilon\} = P\{Y \geq \varepsilon^2\} \leq \frac{E(Y)}{\varepsilon^2} = \frac{\sigma^2}{\varepsilon^2}$$



Theorem: Chebyshev's inequality

If $\mu \triangleq E(X)$ and $\sigma^2 = D(X)$ both exist, then for $\forall \varepsilon > 0$, we have:

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

Equivalently, we have:

$$P\{|X - \mu| < \varepsilon\} \geq 1 - \frac{\sigma^2}{\varepsilon^2}$$

Let $\varepsilon = 3\sigma$ and 4σ , we have:

$$P\{|X - \mu| < 3\sigma\} \geq 1 - \frac{1}{9} = 88.90\%$$

$$P\{|X - \mu| < 4\sigma\} \geq 1 - \frac{1}{16} = 93.75\%$$

Inferencing result from Chebyshev's inequality:

$$\text{If } \sigma^2 = 0, \text{ then } P\{X = \mu\} = 1.$$

Even for ordinary r.v.s, the reliability of the 3 σ rule is close to 90%.

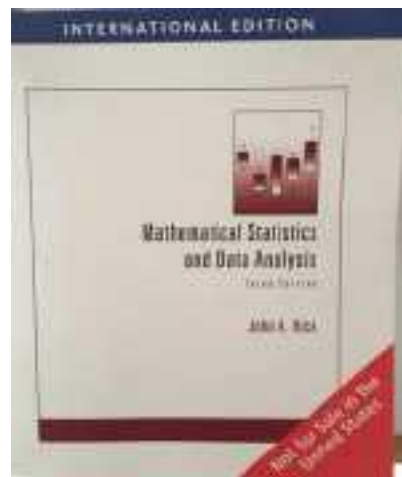


Summary Expected values and Variances

$X \sim$	$E(X)$	$D(X)$
$X \sim p(\lambda)$	λ	λ
$X \sim b(n, p)$	np	$np(1 - p)$
$X \sim U(a, b)$	$\frac{a + b}{2}$	$x = \frac{(b - a)^2}{12}$
$X \sim \exp(1/\theta)$	θ	θ^2
$X \sim N(\mu, \sigma^2)$	μ	σ^2



Homework



P170: 49、50、55



Supplementary Questions:

1. Suppose that X and Y are independent random variables. $E(X) = 3$, $E(Y) = 1$, $D(X) = 4$, $D(Y) = 9$. If $Z = 5X - 2Y + 15$, compute $E(Z)$, $D(Z)$.
2. Suppose that $X_i (i = 1, 2, 3, 4)$ are mutually independent to each other. $E(X_i) = 2i$, $D(X_i) = 5 - i$. If $Z = 2X_1 - X_2 + 3X_3 - 0.5X_4$, compute $E(Z)$ and $D(Z)$.