第二节　方　差

一、随机变量方差的概念及性质

二、重要概率分布的方差

三、例题讲解

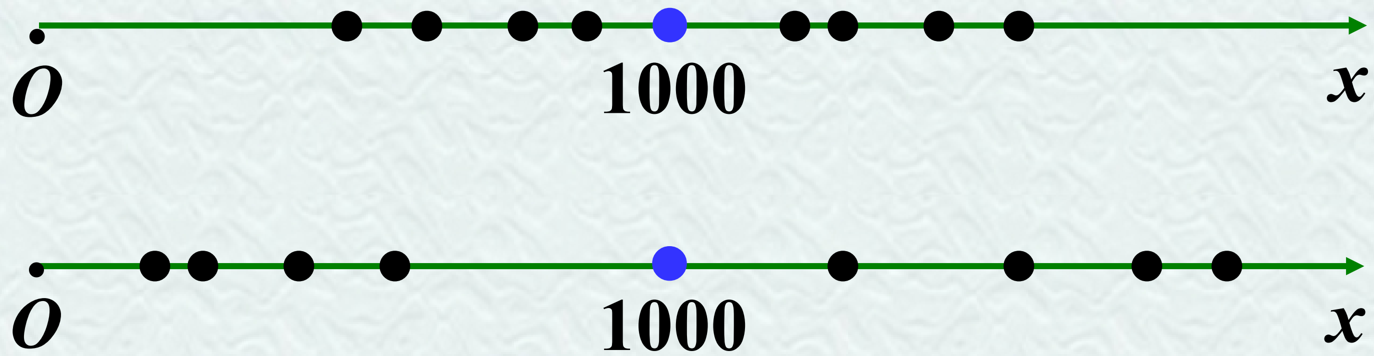
四、小结

一、随机变量方差的概念及性质

1. 概念的引入

方差是一个常用来体现随机变量取值分散程度的量.

实例 有两批灯泡,其平均寿命都是 $E(X)=1000$小时.



2. 方差的定义

设 X 是一个随机变量,若 $E\{[X-E(X)]^2\}$ 存在,则称 $E\{[X-E(X)]^2\}$ 为 X 的方差, 记为 $D(X)$ 或 $\mathrm{Var}(X)$, 即$\cdots$

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

称 $\sqrt{D(X)}$ 为标准差或均方差 , 记为 $\sigma(X)$.

3. 方差的意义

方差是一个常用来体现随机变量 X 取值分散程度的量. 如果 $D(X)$ 值大, 表示 X 取值分散程度大, $E(X)$ 的代表性差; 而如果 $D(X)$ 值小, 则表示X 的取值比较集中, 以 $E(X)$ 作为随机变量的代表性好.

4. 随机变量方差的计算

(1) 利用定义计算

离散型随机变量的方差

$$D(X)=\sum_{k=1}^{+\infty}[x_k-E(X)]^2 p_k,$$

其中 $P\{X=x_k\}=p_k, k=1,2,\cdots$ 是 X 的分布律.

连续型随机变量的方差

$$D(X)=\int_{-\infty}^{+\infty}[x-E(X)]^2 f(x)\,\mathrm{d}x,$$

其中 $f(x)$ 为X的概率密度.

(2) 利用公式计算

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

证明

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

5. 方差的性质

(1) 设 C 是常数, 则有 $D(C)=0.$

证明 $D(C)=E(C^2)-[E(C)]^2=C^2-C^2=0.$

(2) 设 X 是一个随机变量, C 是常数, 则有

$$D(CX)=C^2D(X).$$

证明 $D(CX)=E\{[CX-E(CX)]^2\}$

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

$$=C^2D(X).$$

(3) 设 X, Y 相互独立, $D(X), D(Y)$ 存在, 则

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

证明

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

$$= E\{[X - E(X)] \pm [Y - E(Y)]\}^2$$

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

$$\pm 2E\{[X - E(X)][Y - E(Y)]\}$$

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

推广　若 $X_1, X_2, \cdots, X_n$ 相互独立,则有

$$D(X_1 \pm X_2 \pm \cdots \pm X_n) = D(X_1) + D(X_2) + \cdots + D(X_n).$$

(4) $D(X) = 0$ 的充要条件是 X 以概率 1 取常数 C,即

$$P\{X = C\} = 1.$$

二、重要概率分布的方差

1. 两点分布

已知随机变量 X 的分布律为

X	1	0
p	p	$1-p$

则有 $E(X)=1\cdot p+0\cdot q=p,$

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

$$=1^2\cdot p+0^2\cdot(1-p)-p^2=pq.$$

2. 二项分布

设随机变量 X 服从参数为 n, p 二项分布，其分布律为

$$P\{X=k\}=\binom{n}{k}p^k(1-p)^{n-k},(k=0,1,2,\cdots,n),$$

$$0<p<1.$$

则有

$$E(X)=\sum_{k=0}^{n}k\cdot P\{X=k\}$$

$$=\sum_{k=0}^{n}k\binom{n}{k}p^k(1-p)^{n-k}$$

$$= \sum_{k=0}^{n} \frac{kn!}{k!(n-k)!} p^k (1-p)^{n-k}$$

$$= \sum_{k=1}^{n} \frac{np(n-1)!}{(k-1)![(n-1)-(k-1)]!} p^{k-1} (1-p)^{(n-1)-(k-1)}$$

$$= np \sum_{k=1}^{n} \frac{(n-1)!}{(k-1)![(n-1)-(k-1)]!} p^{k-1} (1-p)^{(n-1)-(k-1)}$$

$$= np[p + (1-p)]^{n-1}$$

$$= np.$$

$$E(X^2) = E[X(X-1)+X]$$

$$= E[X(X-1)] + E(X)$$

$$= \sum_{k=0}^{n} k(k-1)\binom{k}{n} p^k (1-p)^{n-k} + np$$

$$= \sum_{k=0}^{n} \frac{k(k-1)n!}{k!(n-k)!} p^k (1-p)^{n-k} + np$$

$$= n(n-1)p^2\sum_{k=2}^{n}\frac{(n-2)!}{(n-k)!(k-2)!}p^{k-2}(1-p)^{(n-2)-(k-2)}$$

$$+ np$$

$$= n(n-1)p^2[p+(1-p)]^{n-2} + np$$

$$= (n^2-n)p^2 + np.$$

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

$$= (n^2-n)p^2 + np - (np)^2$$

$$= np(1-p).$$

3. 泊松分布

设 $X \sim \pi(\lambda)$, 且分布律为

$$P\{X=k\}=\frac{\lambda^k}{k!}e^{-\lambda},\quad k=0,1,2,\cdots,\quad \lambda>0.$$

则有

$$E(X)=\sum_{k=0}^{\infty}k\cdot\frac{\lambda^k}{k!}e^{-\lambda}=e^{-\lambda}\sum_{k=1}^{\infty}\frac{\lambda^{k-1}}{(k-1)!}\cdot\lambda$$

$$=\lambda e^{-\lambda}\cdot e^{\lambda}=\lambda.$$

$$E(X^2) = E[X(X-1)+X]$$

$$= E[X(X-1)] + E(X)$$

$$= \sum_{k=0}^{+\infty} k(k-1) \cdot \frac{\lambda^k}{k!} \mathrm{e}^{-\lambda} + \lambda$$

$$= \lambda^2 \mathrm{e}^{-\lambda} \sum_{k=2}^{+\infty} \cdot \frac{\lambda^{k-2}}{(k-2)!} + \lambda = \lambda^2 \mathrm{e}^{-\lambda} \mathrm{e}^{\lambda} + \lambda = \lambda^2 + \lambda.$$

所以 $D(X) = E(X^2) - [E(X)]^2 = \lambda^2 + \lambda - \lambda^2 = \lambda$.

泊松分布的期望和方差都等于参数 λ.

4. 均匀分布

设 $X \sim U(a,b)$, 其概率密度为

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

则有 $E(X) = \int_{-\infty}^{\infty} x f(x) \mathrm{d}x = \int_a^b \dfrac{1}{b-a} x \, \mathrm{d}x$

$$= \frac{1}{2}(a+b).$$

结论 均匀分布的数学期望位于区间的中点.

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

$$=\int_a^b x^2\frac{1}{b-a}\mathrm{d}x-\left(\frac{a+b}{2}\right)^2$$

$$=\frac{(b-a)^2}{12}.$$

5. 指数分布

设随机变量 X 服从指数分布,其概率密度为

$$f(x)=\begin{cases}\dfrac{1}{\theta}\mathrm{e}^{-x/\theta}, & x>0,\\ 0, & x\le 0.\end{cases}\quad \text{其中 } \theta>0.$$

则有

$$E(X)=\int_{-\infty}^{+\infty}xf(x)\,\mathrm{d}x=\int_{0}^{+\infty}x\cdot\frac{1}{\theta}\mathrm{e}^{-x/\theta}\,\mathrm{d}x$$

$$=-x\mathrm{e}^{-x/\theta}\Big|_{0}^{+\infty}+\int_{0}^{+\infty}\mathrm{e}^{-x/\theta}\,\mathrm{d}x=\theta.$$

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

$$=\int_0^{+\infty}x^2\cdot\frac{1}{\theta}\mathrm{e}^{-x/\theta}\,\mathrm{d}x-\theta^2$$

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

$$=\theta^2.$$

指数分布的期望和方差分别为 θ 和 θ^2.

6. 正态分布

设 $X \sim N(\mu, \sigma^2)$, 其概率密度为

$$f(x) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(x-\mu)^2}{2\sigma^2}}, \quad \sigma > 0, \quad -\infty < x < +\infty.$$

则有 $$E(X) = \int_{-\infty}^{+\infty} x f(x) \mathrm{d}x$$

$$= \int_{-\infty}^{+\infty} x \cdot \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(x-\mu)^2}{2\sigma^2}} \mathrm{d}x.$$

令 $\dfrac{x-\mu}{\sigma} = t \Rightarrow x = \mu + \sigma t,$

所以 $$E(X)=\int_{-\infty}^{+\infty}x\cdot\frac{1}{\sqrt{2\pi}\sigma}\mathrm{e}^{-\frac{(x-\mu)^2}{2\sigma^2}}\,\mathrm{d}x$$

$$=\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{+\infty}(\mu+\sigma t)\mathrm{e}^{-\frac{t^2}{2}}\,\mathrm{d}t$$

$$=\mu\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{+\infty}\mathrm{e}^{-\frac{t^2}{2}}\,\mathrm{d}t+\frac{\sigma}{\sqrt{2\pi}}\int_{-\infty}^{+\infty}t\mathrm{e}^{-\frac{t^2}{2}}\,\mathrm{d}t$$

$$=\mu.$$

$$D(X)=\int_{-\infty}^{+\infty}(x-\mu)^2 f(x)\,\mathrm{d}x$$

$$=\int_{-\infty}^{+\infty}(x-\mu)^2\cdot\frac{1}{\sqrt{2\pi}\sigma}\mathrm{e}^{-\frac{(x-\mu)^2}{2\sigma^2}}\,\mathrm{d}x.$$

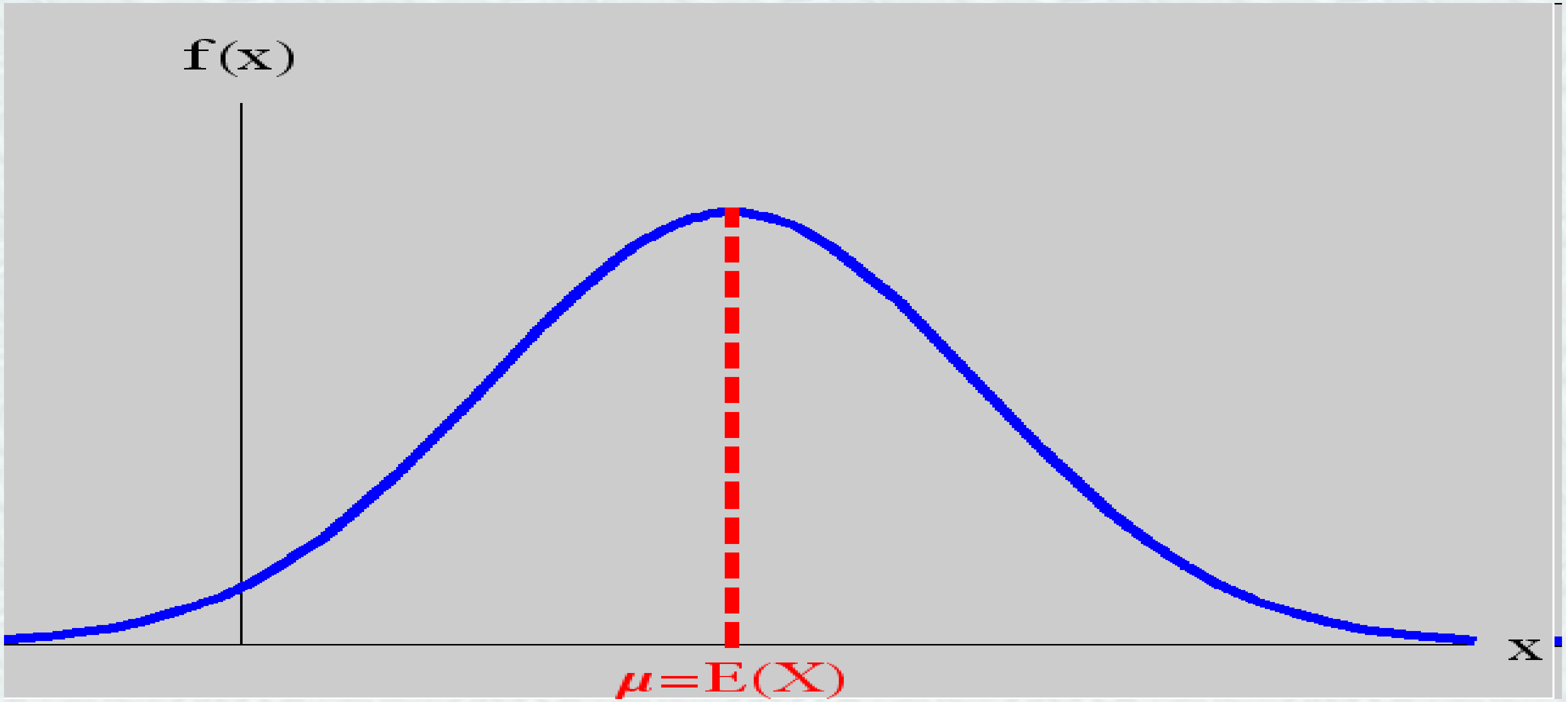
令 $\dfrac{x-\mu}{\sigma}=t$，得

$$D(X)=\frac{\sigma^2}{\sqrt{2\pi}}\int_{-\infty}^{+\infty}t^2\mathrm{e}^{-\frac{t^2}{2}}\,\mathrm{d}t$$

$$=\frac{\sigma^2}{\sqrt{2\pi}}\left(-t\mathrm{e}^{-\frac{t^2}{2}}\Big|_{-\infty}^{+\infty}+\int_{-\infty}^{+\infty}\mathrm{e}^{-\frac{t^2}{2}}\,\mathrm{d}t\right)$$

$$= 0 + \frac{\sigma^2}{\sqrt{2\pi}}\sqrt{2\pi} = \sigma^2.$$

正态分布的期望和方差分别为两个参数 μ 和 σ^2.



分　　布	参数	数学期望	方差
两点分布	$0<p<1$	p	$p(1-p)$
二项分布	$n\geq 1,$ $0<p<1$	np	$np(1-p)$
泊松分布	$\lambda>0$	λ	λ
均匀分布	$a<b$	$(a+b)/2$	$(b-a)^2/12$
指数分布	$\theta>0$	θ	θ^2
正态分布	$\mu,\sigma>0$	μ	σ^2

三、例题讲解

例1 设随机变量 X 具有概率密度

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

求 $D(X)$.

解 $E(X)=\int_{-1}^{0}x(1+x)\,\mathrm{d}x+\int_{0}^{1}x(1-x)\,\mathrm{d}x$

$$=0,$$

$$E(X^2)=\int_{-1}^{0}x^2(1+x)\,\mathrm{d}\,x+\int_{0}^{1}x^2(1-x)\,\mathrm{d}\,x$$

$$=\frac{1}{6},$$

于是

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

$$=\frac{1}{6}-0^2=\frac{1}{6}.$$

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

解 因为 $X \sim N(22.40, 0.03^2)$，$Y \sim N(22.50, 0.04^2)$，

所以 $X - Y \sim N(-0.10, 0.0025)$，

故有 $P\{X < Y\} = P\{X - Y < 0\}$

$$= P\left\{\frac{(X-Y)-(-0.10)}{\sqrt{0.0025}} < \frac{0-(-0.10)}{\sqrt{0.0025}}\right\} = \Phi(2) = 0.9772.$$

例3　设连续型随机变量 X 的概率密度为

$$f(x)=\begin{cases}\cos x, & 0\le x\le \dfrac{\pi}{2},\\ 0, & \text{其他}.\end{cases}$$

求随机变量 $Y=X^2$ 的方差 $D(Y)$.

解 $E(X^2)=\int_{-\infty}^{+\infty}x^2 f(x)\,\mathrm{d}x$

$$=\int_0^{\frac{\pi}{2}}x^2\cos x\,\mathrm{d}x=\frac{\pi^2}{4}-2,$$

$$E(X^4)=\int_{-\infty}^{+\infty}x^4 f(x)\,\mathrm{d}x=\int_0^{\frac{\pi}{2}}x^4\cos x\,\mathrm{d}x$$

$$= \frac{\pi^4}{16} - 3\pi^2 + 24,$$

因为 $D(X) = E(X^2) - [E(X)]^2,$

所以 $D(X^2) = E(X^4) - [E(X^2)]^2$

$$= \frac{\pi^4}{16} - 3\pi^2 + 24 - \left(\frac{\pi^2}{4} - 2\right)^2$$

$$= 20 - 2\pi^2.$$

例4 设 $X \sim \begin{pmatrix} -2 & 0 & 1 & 3 \\ \frac{1}{3} & \frac{1}{2} & \frac{1}{12} & \frac{1}{12} \end{pmatrix}$，求 $D(2X^3+5)$.

解 $$D(2X^3+5) = D(2X^3) + D(5)$$

$$= 4D(X^3)$$

$$= 4[E(X^6) - (E(X^3))^2]$$

$$E(X^6) = (-2)^6 \times \frac{1}{3} + 0^6 \times \frac{1}{2} + 1^6 \times \frac{1}{12} + 3^6 \times \frac{1}{12} = \frac{493}{6},$$

$$[E(X^3)]^2 = \left[(-2)^3 \times \frac{1}{3} + 0^3 \times \frac{1}{2} + 1^3 \times \frac{1}{12} + 3^3 \times \frac{1}{12}\right]^2$$

$$= \frac{1}{9},$$

故 $D(2X^3 + 5) = 4[E(X^6) - (E(X^3))^2]$

$$= \frac{2954}{9}.$$

契比雪夫不等式

定理 设随机变量 X 具有数学期望 $E(X)=\mu$, 方差 $D(X)=\sigma^2$,则对于任意正数 ε, 不等式

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

成立.

切比雪夫不等式

证明 取连续型随机变量的情况来证明.

设 X 的概率密度为 $f(x)$,则有

$$P\{|X-\mu|\geq\varepsilon\}=\int_{|x-\mu|\geq\varepsilon}f(x)\,\mathrm{d}x$$

$$\leq\int_{|x-\mu|\geq\varepsilon}\frac{|x-\mu|^2}{\varepsilon^2}f(x)\,\mathrm{d}x$$

$$\leq\frac{1}{\varepsilon^2}\int_{-\infty}^{+\infty}(x-\mu)^2f(x)\,\mathrm{d}x=\frac{1}{\varepsilon^2}\sigma^2.$$

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

$$P\{|X-\mu|\geq\varepsilon\}\leq\frac{\sigma^2}{\varepsilon^2}\Leftrightarrow P\{|X-\mu|<\varepsilon\}\geq1-\frac{\sigma^2}{\varepsilon^2}.$$

四、小结

1. 方差是一个常用来体现随机变量 X 取值分散程度的量. 如果 $D(X)$ 值大, 表示 X 取值分散程度大, $E(X)$ 的代表性差; 而如果 $D(X)$ 值小, 则表示 X 的取值比较集中, 以 $E(X)$ 作为随机变量的代表性好.

2. 方差的计算公式

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

$$D(X)=\sum_{k=1}^{+\infty}[x_k-E(X)]^2 p_k,$$

$$D(X)=\int_{-\infty}^{+\infty}[x-E(X)]^2 f(x)\,\mathrm{d}x.$$

3. 方差的性质

1° $D(C)=0$;

2° $D(CX)=C^{2}D(X)$;

3° $D(X\pm Y)=D(X)+D(Y)$.

4. 契比雪夫不等式

$$P\{|X-\mu|\geq\varepsilon\}\leq\frac{\sigma^{2}}{\varepsilon^{2}}\Leftrightarrow P\{|X-\mu|<\varepsilon\}\geq1-\frac{\sigma^{2}}{\varepsilon^{2}}.$$