2. 方差

方差

例：有一批灯泡，知其平均寿命是$E(X)=1000$（小时）。仅由这一指标我们还不能判定这批灯泡的质量好坏。事实上，有可能其中绝大部分灯泡的寿命都在950～1050小时；也有可能其中约有一半是高质量的，它们的寿命大约有1300小时，另一半却是质量很差的，其寿命大约只有700小时，为要评定这批灯泡质量的好坏，还需进一步考察灯泡寿命X与其均值$E(X)=1000$的偏离程度。容易看到

$$E\{|X-E(X)|\}$$

能度量偏离程度，但由于绝对值运算不便，通常使用

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

◆ **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$Var(X)$，即

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

方差

在应用上还引入量$\sqrt{D(X)}$，记为$\sigma(X)$，称为标准差或均方差。随机变量X的方差表达了X的取值与其数学期望的偏离程度，若$D(X)$较小意味着X的取值比较集中在$E(X)$的附近，反之，若$D(X)$较大则表示X的取值较分散。由定义知，方差实际上就是随机变量X的函数$g(X)=(X-E(X))^2$的数学期望。

对于离散型随机变量，有

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

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

对于连续型随机变量，有

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

其中$f(x)$是X的概率密度

方差

随机变量X的方差可按下列公式计算

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

证： 由数学期望的性质

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

方差

例：设随机变量X具有数学期望$E(X)=\mu$，方差$D(X)=\sigma^2\neq 0$.

记
$$X^*=\frac{X-\mu}{\sigma}$$

则
$$E(X^*)=\frac{1}{\sigma}E(X-\mu)=\frac{1}{\sigma}[E(X)-\mu]=0;$$
$$D(X^*)=E(X^{*2})-[E(X^*)]^2=E\left[\left(\frac{X-\mu}{\sigma}\right)^2\right]$$
$$=\frac{1}{\sigma^2}E[(X-\mu)^2]=\frac{\sigma^2}{\sigma^2}=1$$

即$X^*=\frac{X-\mu}{\sigma}$ 的数学期望为0，方差为1。

X^*称为X的**标准化变量**。

方差

例：设随机变量X具有$(0-1)$分布，其分布律为

$P\{X=0\}=1-p, P\{X=1\}=p$。求$D(X)$。

解:

$$E(X)=0\cdot(1-p)+1\cdot p=p,$$

$$E(X^2)=0^2\cdot(1-p)+1^2\cdot p=p$$

$$D(X)=E(X^2)-[E(X)]^2=p-p^2=p(1-p).$$

方差

例：设随机变量$X \sim \pi(\lambda)$，求$D(X)$。

解：X的分布律为

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

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

所以方差$D(X)=E(X^2)-[E(X)]^2=\lambda$，泊松分布的数学期望和方差相等，都等于参数$\lambda$。

方差

例：设随机变量$X \sim U(a,b)$，求$D(X)$.

解：X的概率密度为$f(x)=\begin{cases}\frac{1}{b-a}, & a<x<b \\ 0 & , others\end{cases}$

已算得$E(X)=\frac{a+b}{2}$，方差为

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

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

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

方差

例：设随机变量X服从指数分布，其概率密度为

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

其中$\theta>0$，求$E(X)$，$D(X)$。

解：

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

$$=-x\mathrm{e}^{-x/\theta}|_0^\infty+\int_0^\infty \mathrm{e}^{-x/\theta}\mathrm{d}x=\theta$$

$$E(X^2)=\int_{-\infty}^{\infty}x^2f(x)\mathrm{d}x=\int_0^\infty x^2\frac{1}{\theta}\mathrm{e}^{-x/\theta}\mathrm{d}x=-x^2\mathrm{e}^{-x/\theta}|_0^\infty+\int_0^\infty 2x\mathrm{e}^{-x/\theta}\mathrm{d}x=2\theta^2$$

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

方差

◆ 方差性质

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

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

- 设X，Y是两个随机变量，则有

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

若X，Y相互**独立**，则有$D(X + Y) = D(X) + D(Y)$

可推广到任意有限多个相互独立随机变量之和

- $D(X) = 0$的充要条件是X以概率1取常数$E(X)$，即

$$P\{X = E(X)\} = 1$$

方差

证1：$D(C)=E\{[C-E(C)]^2\}=0$

证2：$D(CX)=E\{[CX-E(CX)]^2\}=C^2E\{[X-E(X)]^2\}=C^2D(X)$

证3：

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

上式右端第三项

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

若X，Y相互独立，可知上式为0

于是$D(X+Y)=D(X)+D(Y)$

方差

证4：充分性

设$P\{X=E(X)\}=1$，则有$P\{X^2=[E(X)]^2\}=1$，于是

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

必要性写在切比雪夫不等式证明后

例：设随机变量$X\sim b(n,p)$，求$E(X)$，$D(X)$.

解：随机变量X是n重伯努利试验中事件A发生次数，且每次试验中A发生概率为p，引入随机变量 $X_k=\begin{cases}1, A\text{在第k次试验发生}\\0, A\text{在第k次试验不发生}\end{cases}, k=1,2,\dots n$

易知$X=X_1+X_2+\cdots+X_n$，各次试验相互独立。

而X_k服从同一$(0-1)$分布

X_k	0	1
p_k	$1-p$	p

方差

以n，p为参数的二项分布变量，可以分解为n个相互独立且都服从p为参数的$(0-1)$分布的随机变量之和。

已知$E(X_k)=p, D(X_k)=p(1-p), k=1,2,\dots,n.$ 故知

$$E(X)=E(\sum_{k=1}^{n}X_k)=\sum_{k=1}^{n}E(X_k)=np$$

又由于各次试验相互独立，得

$$D(X)=D(\sum_{k=1}^{n}X_k)=\sum_{k=1}^{n}D(X_k)=np(1-p).$$

即$E(X)=np$， $D(X)=np(1-p)$

方差

例：设随机变量$X \sim N(\mu, \sigma^2)$，求$E(X), D(X)$.

解：先求标准正态变量$Z = \frac{X-\mu}{\sigma}$的期望和方差。Z的概率密度为

$$\varphi(t) = \frac{1}{\sqrt{2\pi}} \mathrm{e}^{-t^2/2}$$

于是$E(Z) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} t e^{-t^2/2} dt = \frac{-1}{\sqrt{2\pi}} e^{-t^2/2} |_{-\infty}^{\infty} = 0$，

$$D(Z) = E(Z^2) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} t^2 e^{-t^2/2}\, dt$$

$$= \frac{-1}{\sqrt{2\pi}} t e^{-t^2/2} |_{-\infty}^{\infty} + \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-t^2/2}\, dt = 1$$

令$X = \mu + \sigma Z$，得

$$E(X) = E(\mu + \sigma Z) = \mu$$

$$D(X) = D(\mu + \sigma Z) = D(\sigma Z) = \sigma^2 D(Z) = \sigma^2$$

方差

正态分布概率密度中的两个参数μ和σ分别就是该分布的数学期望和均方差，正态分布完全可由数学期望和方差所确定。

- 若$X_i \sim N(\mu, \sigma^2), i = 1,2,\dots,n$，且它们相互独立，则它们的线性组合$C_1X_1 + C_2X_2 + \cdots + C_nX_n$（系数不全为0）仍然服从正态分布，由数学期望和方差性质可知

$$C_1X_1 + C_2X_2 + \cdots + C_nX_n \sim N(\sum_{i=1}^{n} C_i\mu_i, \sum_{i=1}^{n} C_i^2\sigma_i^2)$$

方差

例：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：按题意$P\{X < Y\} = P\{X - Y < 0\}$。由于

$$X - Y \sim N(-0.10, 0.0025)$$

所以

$$P\{X < Y\} = P\{X - Y < 0\} = P\{\frac{(X-Y) - (-0.10)}{\sqrt{0.0025}} < \frac{0 - (-0.10)}{\sqrt{0.0025}}\} = \Phi(\frac{0.10}{0.05}) = \Phi(2) = 0.9772$$

方差

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

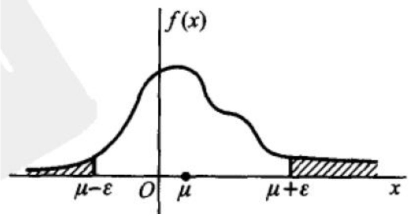
$$P\{|X-\mu|\geqslant\varepsilon\}\leqslant\frac{\sigma^2}{\varepsilon^2}$$

成立。称为切比雪夫不等式

证：就连续性随机变量证明。设X的概率密度为$f(x)$，则有

$$P\{|X-\mu|\geqslant\varepsilon\}=\int_{|x-\mu|\geqslant\varepsilon}f(x)\,\mathrm{d}x\leqslant\int_{|x-\mu|\geqslant\varepsilon}\frac{|x-\mu|^2}{\varepsilon^2}f(x)\mathrm{d}x$$

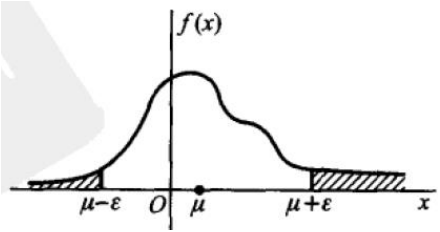
$$\leqslant\frac{1}{\varepsilon^2}\int_{-\infty}^{\infty}(x-\mu)^2f(x)\mathrm{d}x=\frac{\sigma^2}{\varepsilon^2}$$



方差

切比雪夫不等式也可以写成

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



给出了在随机变量分布未知，而只知道$E(X)$和$D(X)$的情况下估计概率$P\{|X-\mu|<\varepsilon\}$的界限。

方差

◆ 方差性质4

➢ $D(X)=0$的充要条件是X以概率1取常数$E(X)$，即

$$P\{X=E(X)\}=1$$

必要性证明：

设$D(X)=0$，要证$P\{X=E(X)\}=1$。

证：用反证法，假设$P\{X=E(X)\}<1$，则对于某一个数$\varepsilon>0$，有$P\{|X-\mu|\geqslant\varepsilon\}>0$，但由于切比雪夫不等式，对于任意$\varepsilon>0$，因$\sigma^2=0$，有$P\{|X-\mu|\geqslant\varepsilon\}=0$，矛盾，于是$P\{X=E(X)\}=1$

3. 协方差及相关系数

协方差及相关系数

如果两个随机变量X和Y是相互**独立**的，则

$$E\{[(X-E(X)][Y-E(Y)]\}=0$$

这意味着当$E\{[(X-E(X)][Y-E(Y)]\}\neq 0$时，$X$与$Y$不相互独立，而是存在着一定的关系。

定义：$E\{[(X-E(X)][Y-E(Y)]\}$称为随机变量X和Y的协方差。记作$Cov(X,Y)$，即

$$Cov(X,Y)=E\{[(X-E(X)][Y-E(Y)]\}$$

而

$$\rho_{XY}=\frac{\mathrm{Cov}(X,Y)}{\sqrt{D(X)}\sqrt{D(Y)}}$$

称为随机变量X与Y的相关系数

协方差及相关系数

由定义

$$Cov(X,Y) = Cov(Y,X), Cov(X,X) = D(X)$$

对于任意两个随机变量X和Y，下列等式成立：

$$D(X+Y) = D(X) + D(Y) + 2Cov(X,y)$$

将$Cov(X,Y)$定义展开，得

$$Cov(X,Y) = E(XY) - E(X)E(Y)$$

协方差具有下述性质

- $Cov(aX,bY) = abCov(X,Y), a,b$是常数
- $Cov(X_1+X_2,Y) = Cov(X_1,Y) + Cov(X_2,Y)$

协方差及相关系数

ρ_{XY}的重要性质以及含义

考虑以X的线性函数$a+bX$来近似表示Y，以均方误差

$$e = E[(Y-(a+bX))^2]$$

$$= E(Y^2) + b^2E(X^2) + a^2 - 2bE(XY) + 2abE(X) - 2aE(Y)$$

来衡量$a+bX$近似Y的好坏程度。将e分别关于a，b求偏导数，

并令它们等于零

$$\begin{cases} \dfrac{\partial e}{\partial a} = 2a + 2bE(X) - 2E(Y) = 0 \\ \dfrac{\partial e}{\partial b} = 2bE(X^2) - 2E(XY) + 2aE(X) = 0 \end{cases}$$

协方差及相关系数

解得

$$b_0 = \frac{\mathrm{Cov}(X,Y)}{D(X)}$$

$$a_0 = E(Y) - b_0 E(X) = E(Y) - E(X)\frac{\mathrm{Cov}(X,Y)}{D(X)}$$

代入原式得

$$\min_{a,b} E\{[Y-(a+bX)]^2\}$$

$$= E\{[Y-(a_0+b_0X)]^2\} = (1-\rho_{XY}^2)D(Y)$$

协方差及相关系数

定理：

- $|\rho_{XY}| \le 1$
- $|\rho_{XY}| = 1$的充要条件是，存在常数a，b使

$$P\{Y = a + bX\} = 1\}$$

证1：由$E[(Y-(a+bX))^2]$ 及$D(Y)$的非负性，得$1-\rho_{XY}^2 \ge 0$，即$|\rho_{XY}| \le 1$。

证2：充分性，若$|\rho_{XY}| = 1$，则$E\{[Y-(a_0+b_0X)]^2\} = 0$

从而$0 = E\{[Y-(a_0+b_0X)]^2\} = D[Y-(a_0+b_0X)] + [E(Y-(a_0+b_0X))]^2$，

故有$D[Y-(a_0+b_0X)] = 0$，　$[E(Y-(a_0+b_0X))]^2 = 0.$

协方差及相关系数

又由方差的性质4知

$$P\{Y-(a_0+b_0X)=0\}=1, \text{即} P\{Y=a_0+b_0X\}=1$$

必要性，若存在常数a^*，b^*使

$$P\{Y=a^*+b^*X\}=1，\text{即} P\{Y-(a^*+b^*X)=0\}=1$$

$$P\{[Y-(a^*+b^*X)]^2=0\}=1$$

$$E\{[Y-(a^*+b^*X)]^2\}=0$$

$$\text{故有}\quad 0=E\{[Y-(a^*+b^*X)]^2\}\geq \min_{a,b} E\{[Y-(a+bX)]^2\}$$

$$=E\{[Y-(a_0+b_0X))]^2\}=(1-\rho_{XY}^2)D(Y)$$

$$\text{即得} |\rho_{XY}|=1$$