§4.2　随机变量的方差

一、方差的定义

数学期望反映了随机变量取值的平均结果，它仅是从一个角度反映了随机变量的特征, 有时我们还需要了解随机变量取值的离散程度. 例如：(1)学习成绩：我们不但要了解平均成绩, 也要了解成绩之间差异的大小. (2)收入状况：我们要知道平均收入的大小，也要了解收入差异的大小. 随机变量的离散程度如何来求？

随机变量的离散程度如何来衡量？

$$X-EX\quad\begin{cases}(1)\text{不能直接用(是随机变量);}\\(2)\text{数学期望}E(X-EX)=0\text{不能用;}\\(3)E|X-EX|\text{可行，但数学处理上困难;}\\(4)E(X-EX)^2\text{可行，便于处理.}\end{cases}$$

我们用$E(X-EX)^2$来表示随机变量对于数学期望的偏离程度，因此有下面的定义.

定义1 若$E(X-EX)^2$存在，则称它为随机变量X的**方差**(variance)，并记为DX，而$\sqrt{DX}$称为X的**标准差**(standard deviation)或根方差、均方差.

两者量纲相同

若X为离散型随机变量，概率分布为

$$P(X = x_k) = p_k, \qquad k = 1, 2, \cdots$$

则

$$D(X) = \sum_k (x_k - EX)^2 p_k .$$

若X为连续型随机变量，概率密度为$f(x)$. 则

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

计算方差的简化公式：

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

证明：$DX = E(X - EX)^2$

$$= E\left[X^2 - 2X \cdot EX + (EX)^2 \right]$$

$$= EX^2 - 2EX \cdot EX + (EX)^2 = EX^2 - (EX)^2 .$$

方差是描述随机变量X的取值与平均值的偏离程度的量. 由定义可知，方差实际上就是随机变量函数$g(X)=(X-EX)^2$的数学期望，因而方差的性质可根据数学期望的性质推出.

性质1 $DX \geq 0$，且

$$DX=0 \Leftrightarrow P\left\{X=EX\right\}=1.$$

特别地 常数的方差为零.

性质2 $D(aX+b)=a^2D(X).$

性质3 $D(X \pm Y)=D(X)+D(Y)$

$$\pm 2E[(X-EX)(Y-EY)].$$

特别地，若X,Y相互独立，则

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

性质4　若$c \neq EX$，则 $DX < E(X-c)^2$.

说明：随机变量与其数学期望的离散程度最小.

注：利用$\{X \neq E(X)\} = \bigcup_{n=1}^{\infty} \left\{ |X - E(X)| \geq \frac{1}{n} \right\}$可

证明 $DX = 0 \Leftrightarrow P\{X = EX\} = 1$.

练习题：

试计算 $D(\sum_{i=1}^{n} a_i X_i + c)$.

其中$a_1, a_2, \cdots\cdots, a_n$，$c$是常数.

$$D(\sum_{i=1}^{n} a_i X_i + c) = E\left[(\sum_{i=1}^{n} a_i X_i + c) - E(\sum_{i=1}^{n} a_i X_i + c)\right]^2$$

$$= E\left[\sum_{i=1}^{n} a_i (X_i - EX_i)\right]^2$$

$$= E\left[\sum_{i=1}^{n} a_i^2 (X_i - EX_i)^2 + \sum_{i \neq j}^{n} a_i a_j (X_i - EX_i)(X_j - EX_j)\right]$$

$$= (a_1, a_2, \cdots, a_n) A (a_1, a_2, \cdots, a_n)^T$$

其中$A = (E(X_i - EX_i)(X_j - EX_j))_{n \times n}$

二　方差的计算实例

【两点分布】

$$EX^2 = 1^2 \cdot p + 0^2 \cdot (1-p) = p,$$

$$DX = EX^2 - (EX)^2 = p - p^2 = pq.$$

【二项分布】　$X \sim B(n, p).$

$$EX^2 = \sum_{k=0}^{n} k^2 C_n^k p^k (1-p)^{n-k} = npq + (np)^2,$$

$$DX = EX^2 - (EX)^2 = npq + (np)^2 - (np)^2 = npq.$$

【泊松分布】　$X \sim P(\lambda).$

$$EX^2=\sum_{k=0}^{\infty}k^2\frac{\lambda^k}{k!}e^{-\lambda}=\lambda^2+\lambda,$$

$$DX=EX^2-(EX)^2=\lambda^2+\lambda-\lambda^2=\lambda.$$

【几何分布】

$$P\{X=k\}=q^{k-1}p,\qquad k=1,2,3,\cdots$$

$$EX=\sum_{k=1}^{\infty}kq^{k-1}p=p\sum_{k=1}^{\infty}(q^k)'_q=p(\sum_{k=1}^{\infty}q^k)'_q=p(\frac{q}{1-q})'=\frac{1}{p},$$

$$EX^2=\sum_{k=1}^{\infty}k^2q^{k-1}p=p\sum_{k=1}^{\infty}\left[k(k-1)\cdot q^{k-1}+kq^{k-1}\right]$$

$$=p\sum_{k=2}^{\infty}q\cdot(q^k)''_q+p\sum_{k=1}^{\infty}(q^k)'_q=\frac{2q}{p^2}+\frac{1}{p},$$

$$DX=EX^2-(EX)^2=\frac{2q}{p^2}+\frac{1}{p}-(\frac{1}{p})^2=\frac{q}{p^2}.$$

【均匀分布】 $X \sim U[a,b]$,

$$EX = \int_{-\infty}^{+\infty} xf(x)\mathrm{d}x = \int_a^b x \cdot \frac{1}{b-a}\mathrm{d}x = \frac{a+b}{2}$$

$$EX^2 = \int_{-\infty}^{+\infty} x^2 f(x)\mathrm{d}x = \int_a^b x^2 \cdot \frac{1}{b-a}\mathrm{d}x = \frac{a^2+ab+b^2}{3},$$

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

【正态分布】 $X \sim N(\mu, \sigma^2)$,

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

$$= \frac{\sigma^2}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} t^2 e^{-\frac{t^2}{2}}\mathrm{d}t = \frac{\sigma^2}{\sqrt{2\pi}} \left[\left(-2te^{-\frac{t^2}{2}}\right)\Big|_0^{\infty} + 2\int_0^{+\infty} e^{-\frac{t^2}{2}}\mathrm{d}t \right] = \sigma^2.$$

几种常用的随机变量的数学期望与方差

分布	概率分布或概率密度	数学期望	方差
0−1分布	$P\{X=k\}=p^k q^{1-k}$，$k=0,1$ $(0<p<1, q=1-p)$	p	pq
二项分布	$P\{X=k\}=C_n^k p^k q^{n-k}$ $k=0,1,2,\cdots,n$ $(0<p<1, q=1-p)$	np	npq
泊松分布	$P\{X=k\}=\dfrac{\lambda^k}{k!}\mathrm{e}^{-\lambda}$，$(\lambda>0)$ $k=0,1,2,\cdots$	λ	λ
均匀分布	$f(x)=\begin{cases}\dfrac{1}{b-a}, & a<x<b\\ 0, & \text{else}\end{cases}$	$\dfrac{a+b}{2}$	$\dfrac{(b-a)^2}{12}$
指数分布	$f(x)=\begin{cases}\lambda\mathrm{e}^{-\lambda x}, & x\geq 0\\ 0, & \text{else}\end{cases}$ $(\lambda>0)$	$\dfrac{1}{\lambda}$	$\dfrac{1}{\lambda^2}$
正态分布	$f(x)=\dfrac{1}{\sqrt{2\pi}\sigma}\mathrm{e}^{-\frac{(x-\mu)^2}{2\sigma^2}}$，$-\infty<x<\infty$ $(\sigma>0)$	μ	σ^2

正态分布的期望和方差分别为两个参数μ和σ^2. 而正态分布也完全由这两个参数决定.

练习题　设$X \sim \exp(\lambda)$，试求$E(X)$.

例1　已知$X\ ,Y$相互独立,且都服从$N(0,0.5)$，求$D(|X-Y|)$.

解　$X \sim N(0,0.5)$，$Y \sim N(0,0.5)$

$$E(X-Y)=0,\quad D(X-Y)=1$$

故

$$Z=X-Y \sim N(0,1)$$

$$E(|X-Y|)=\int_{-\infty}^{+\infty}|z|\frac{1}{\sqrt{2\pi}}e^{-z^2/2}\mathrm{d}z$$

$$=\frac{2}{\sqrt{2\pi}}\int_{0}^{+\infty}ze^{-z^2/2}\mathrm{d}z=\sqrt{\frac{2}{\pi}}$$

$$E(|X-Y|^2)=1,\quad D(|X-Y|)=1-2/\pi.$$

例2 设连续型随机变量 X 的概率密度为

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

求随机变量 $Y=X^2$ 的方差 $D(Y)$.

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

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

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

所以 $D(Y)=E(Y^2)-[E(Y)]^2=20-2\pi^2.$

例3 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： $D(2X^3+5)=D(2X^3)=4D(X^3)$

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

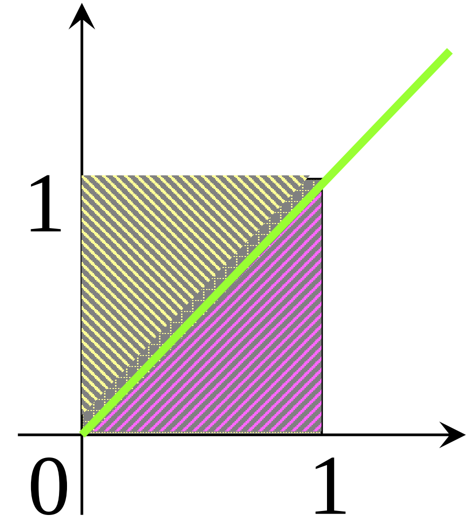
$$E(X^6)=(-2)^6\times\frac{1}{3}+0^6\times\frac{1}{2}+1^6\times\frac{1}{12}+3^6\times\frac{1}{12}=\frac{493}{6},$$

$$[E(X^3)]^2=\left[(-2)^3\times\frac{1}{3}+0^3\times\frac{1}{2}+1^3\times\frac{1}{12}+3^3\times\frac{1}{12}\right]^2=\frac{1}{9},$$

故 $D(2X^3+5)=4[E(X^6)-(E(X^3))^2]=\dfrac{2954}{9}.$

例4 在$[0,\ 1]$中随机地取两个数X,Y.

求$D(\min\{X,Y\})$.



解

$$f(x,y)=\begin{cases}1, & 0<x<1,0<y<1\\ 0, & \text{其它}\end{cases},$$

$$E(\min\{X,Y\})=\iint\limits_{\substack{0<x<1\\0<y<1}}\min\{x,y\}\mathrm{d}x\mathrm{d}y.$$

$$=\int_0^1\left(\int_x^1 x\mathrm{d}y\right)\mathrm{d}x+\int_0^1\left(\int_y^1 y\mathrm{d}x\right)\mathrm{d}y=1/3.$$

$$E(\min{}^2\{X,Y\})=\int_0^1\left(\int_x^1 x^2\mathrm{d}y\right)\mathrm{d}x+\int_0^1\left(\int_y^1 y^2\mathrm{d}x\right)\mathrm{d}y=1/6.$$

$$D(\min\{X,Y\})=E(\min{}^2\{X,Y\})-E^2(\min\{X,Y\})=1/18.$$

例5 已知X的密度函数为

$$f(x)=\begin{cases} Ax^2+Bx, & 0<x<1, \\ 0, & \text{其它} \end{cases}$$

其中A ,B是常数，且$E(X)=0.5.$

（1）求 A ,B.

（2）设$Y=X^2$,求$E(Y),D(Y)$

解 $$\int_{-\infty}^{+\infty} f(x)\mathrm{d}x=\int_0^1 (Ax^2+Bx)\mathrm{d}x=1$$

$$\int_{-\infty}^{+\infty} xf(x)\mathrm{d}x=\int_0^1 x(Ax^2+Bx)\mathrm{d}x=\frac{1}{2}$$

$$\frac{A}{3}+\frac{B}{2}=1,\frac{A}{4}+\frac{B}{3}=\frac{1}{2}.$$

$$\Rightarrow A=-6,\quad B=6.$$

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

$$=\int_0^1 x^2(-6x^2+6x)\mathrm{d}x=3/10,$$

$$E(Y^2)=E(X^4)=\int_{-\infty}^{+\infty} x^4 f(x)\mathrm{d}x$$

$$=\int_0^1 x^4(-6x^2+6x)\mathrm{d}x=1/7.$$

$$D(Y)=E(Y^2)-E^2(Y)=\frac{37}{700}.$$

说明：

1. 仅知随机变量的期望与方差并不能确定其分布.

例如：$X \sim P(1)$，$Y \sim EXP(1)$. $E(X) = E(Y) = 1$,

$D(X) = D(Y) = 1$. 但X、Y分布不同.

2. 在已知某些分布类型时，若知道其期望和方差，便常能确定分布——正态分布、指数分布等.

练习题：

已知X服从正态分布,$E(X) = 3$，$D(X) = 4$, 令

$$Y = 3 - 5X,$$

求Y的密度函数.