§4.2 方差

引例 甲、乙两射手各打了6发子弹,每发子弹击中的环数分别为:

甲	10, 7, 9, 8, 10, 6,
乙	8, 7, 10, 9, 8, 8,

有五个不同数

有四个不同数

问哪一个射手的技术较好?

解 首先比较平均环数

$\overline{甲} = 8.3, \quad \overline{乙} = 8.3$

再比较稳定程度

甲：$2\times(10-8.3)^2+(9-8.3)^2+(8-8.3)^2$

$+(7-8.3)^2+(6-8.3)^2=13.34$

乙：$(10-8.3)^2+(9-8.3)^2+3\times(8-8.3)^2$

$+(7-8.3)^2=5.34$

乙比甲技术稳定，故乙技术较好.

进一步比较平均偏离平均值的程度

甲 $$\frac{1}{6}[2\times(10-8.3)^2+(9-8.3)^2+(8-8.3)^2$$

$$+(7-8.3)^2+(6-8.3)^2]$$

$$=13.34/6=2.22 \triangleq \sum_{k=1}^{5}\left(x_k-E(X)\right)^2 p_k$$

乙 $$\frac{1}{6}[(10-8.3)^2+(9-8.3)^2$$

$$+3\times(8-8.3)^2+(7-8.3)^2]$$

$$=5.34/6=0.89 \triangleq \sum_{k=1}^{4}\left(x_k-E(X)\right)^2 p_k$$

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

方差概念

定义 若$E[X-E(X)]^2$存在，则称其为随机变量 X 的**方差**，记为$D(X)$ 或 $Var(X)$

即 $D(X)=E[X-E(X)]^2$

称 $\sqrt{D(X)}$ 为 X 的**均方差**或**标准差**.

两者量纲相同

$D(X)$ —— 描述 r.v. X 的取值偏离平均值的平均偏离程度 —— 数

若 X 为离散型 **r.v.**，分布律为

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

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

若 X 为连续型**r.v.**，概率密度为 $f(x)$

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

计算方差的常用公式：

$$\boxed{D(X) = E(X^2) - E^2(X)}$$

方差的性质

- $D(C)=0$
- $D(aX)=a^2D(X)$

$$\left.\begin{array}{l} D(C)=0 \\ D(aX)=a^2D(X) \end{array}\right\} D(aX+b)=a^2D(X)$$

- $$\boxed{\begin{aligned} D(X\pm Y)&=D(X)+D(Y)\\ &\pm 2E\big((X-E(X))(Y-E(Y))\big)\end{aligned}}$$

特别地，若X,Y相互独立，则

$$\boxed{D(X\pm Y)=D(X)+D(Y)}$$

若 $X_1,\cdots,X_n$ 相互独立，$a_1,a_2,\cdots,a_n,b$ 为常数

则 $$D\left(\sum_{i=1}^{n} a_i X_i + b\right) = \sum_{i=1}^{n} a_i^2 D(X_i)$$

若X,Y相互独立 $\overset{\Rightarrow}{\not\Leftarrow}$ $D(X \pm Y) = D(X) + D(Y)$

$\Longleftrightarrow$ $E(XY) = E(X)E(Y)$

- 对任意常数C, $D(X) \le E(X-C)^2$，当且仅当$C = E(X)$时等号成立
- $D(X) = 0 \Longleftrightarrow P(X = E(X)) = 1$ 称为X依概率1等于常数$E(X)$

性质 1 的证明:

$$D(C) = E(C - E(C))^2 = 0$$

性质 2 的证明:

$$D(aX+b) = E((aX+b) - E(aX+b))^2$$

$$= E(a(X - E(X)) + (b - E(b)))^2$$

$$= E\left(a^2 (X - E(X))^2\right)$$

$$= a^2 D(X)$$

性质 3 的证明:

$$D(X\pm Y)=E\big((X\pm Y)-E(X\pm Y)\big)^2$$

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

$$\pm 2E\big((X-E(X))(Y-E(Y))\big)$$

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

$$\pm 2E\big((X-E(X))(Y-E(Y))\big)$$

注意到, $E\big((X-E(X))(Y-E(Y))\big)$

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

当 X,Y 相互独立时,

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

性质 4 的证明：

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

$$= E(X-E(X))^2 + (C-E(X))^2$$

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

当 $C = E(X)$时，显然等号成立；

当 $C \neq E(X)$时，$(C-E(X))^2 > 0$

$$E(X-C)^2 > D(X)$$

方差的计算

例1 设$X \sim P(\lambda)$, 求$D(X)$.

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

$$\left.\begin{aligned}&E(X^2)=E(X(X-1))+E(X)\\&E(X(X-1))=\sum_{k=0}^{+\infty}k(k-1)\cdot\frac{\lambda^k e^{-\lambda}}{k!}\\&\qquad=\lambda^2 e^{-\lambda}\sum_{k=2}^{+\infty}\frac{\lambda^{k-2}}{(k-2)!}=\lambda^2\end{aligned}\right\}\Rightarrow E(X^2)=\lambda^2+\lambda$$

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

例2 设$X \sim B(n, p)$，求$D(X)$.

解一 仿照上例求$D(X)$.

解二 引入随机变量 $X_1, X_2, \cdots, X_n$

$$X_i = \begin{cases} 1, & \text{第}i\text{次试验事件}A\text{发生} \\ 0, & \text{第}i\text{次试验事件}\overline{A}\text{发生} \end{cases}$$

$$D(X_i) = p(1-p) \qquad i = 1, 2, \cdots, n$$

$X_1, X_2, \cdots, X_n$ 相互独立， $X = \sum_{i=1}^{n} X_i$

故 $D(X) = \sum_{i=1}^{n} D(X_i) = np(1-p)$

例3 设 $X \sim N(\mu, \sigma^2)$, 求 $D(X)$

解 $D(X) = \int_{-\infty}^{+\infty} (x-\mu)^2 \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(x-\mu)^2}{2\sigma^2}} dx$

$$\overset{\text{令}\frac{x-\mu}{\sigma}=t}{=} \int_{-\infty}^{+\infty} \sigma^2 t^2 \frac{1}{\sqrt{2\pi}} e^{-\frac{t^2}{2}} dt$$

$$= \sigma^2$$

常见随机变量的方差(P.159)

分布	概率分布	方差
参数为p 的 0-1分布	$P(X=1)=p$ $P(X=0)=1-p$	$p(1\text{-}p)$
$B(n,p)$	$P(X=k)=C_n^k p^k(1-p)^{n-k}$ $k=0,1,2,\cdots,n$	$np(1\text{-}p)$
$P(\lambda)$	$P(X=k)=\dfrac{\lambda^k e^{-\lambda}}{k!}$ $k=0,1,2,\cdots$	λ

分布	概率密度	方差
区间(a,b)上的均匀分布	$f(x)=\begin{cases}\dfrac{1}{b-a}, & a<x<b,\\ 0, & 其它\end{cases}$	$\dfrac{(b-a)^2}{12}$
$E(\lambda)$	$f(x)=\begin{cases}\lambda e^{-\lambda x}, & x>0,\\ 0, & 其它\end{cases}$	$\dfrac{1}{\lambda^2}$
$N(\mu,\sigma^2)$	$f(x)=\dfrac{1}{\sqrt{2\pi}\sigma}e^{-\frac{(x-\mu)^2}{2\sigma^2}}$	σ^2

例4 已知X ,Y相互独立，且都服从 $N(0,0.5)$， 求 $E(|X-Y|)$.

解 $X \sim N(0,0.5), Y \sim N(0,0.5)$

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

故 $X-Y \sim N(0,1)$

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

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

例5 设X表示独立射击直到击中目标n次为止所需射击的次数，已知每次射击中靶的概率为p，求$E(X)$, $D(X)$.

解 令X_i表示击中目标$i-1$次后到第i次击中目标所需射击的次数，$i=1,2,\ldots,n$

$X_1, X_2, \cdots, X_n$ 相互独立，且 $X=\sum_{i=1}^{n} X_i$

$$P(X_i=k)=pq^{k-1}, \quad k=1,2,\cdots$$

$$p+q=1$$

$$E(X_i)=\sum_{k=1}^{+\infty} kpq^{k-1}=p\sum_{k=1}^{+\infty} kq^{k-1}=p\frac{1}{(1-q)^2}=\frac{1}{p}$$

$$E(X_i^2)=\sum_{k=1}^{+\infty}k(k-1)pq^{k-1}+\sum_{k=1}^{+\infty}kpq^{k-1}$$

$$=pq\sum_{k=2}^{+\infty}k(k-1)q^{k-2}+\frac{1}{p}$$

$$=pq\frac{d^2}{dx^2}\left(\sum_{k=0}^{+\infty}x^k\right)\Bigg|_{x=q}+\frac{1}{p}$$

$$=pq\frac{2}{(1-x)^3}\Bigg|_{x=q}+\frac{1}{p}=\frac{2-p}{p^2}$$

$$D(X_i)=\frac{2-p}{p^2}-\frac{1}{p^2}=\frac{1-p}{p^2}$$

故 $$E(X)=\sum_{i=1}^{n}E(X_i)=\frac{n}{p}$$

$$D(X)=\sum_{i=1}^{n}D(X_i)=\frac{n(1-p)}{p^2}$$

本例给出了几何分布与巴斯卡分布的期望与方差

例6 将编号分别为 $1 \sim n$ 的 n 个球随机地放入编号分别为 $1 \sim n$ 的 n 只盒子中，每盒一球. 若球的号码与盒子的号码一致，则称为一个配对. 求配对个数 X 的期望与方差.

解 $X_i = \begin{cases} 1, & i\text{号球放入}i\text{号盒} \\ 0, & \text{其它} \end{cases} \qquad i = 1, 2, \cdots, n$

则 $X = \sum_{i=1}^{n} X_i$

但 $X_1, X_2, \cdots, X_n$ 不相互独立，

X_i	1	0
P	$\frac{1}{n}$	$1-\frac{1}{n}$

$$i = 1, 2, \cdots, n$$

$$E(X) = \sum_{i=1}^{n} E(X_i) = n \cdot \frac{1}{n} = 1$$

$$E(X^2) = E\left(\sum_{i=1}^{n} X_i\right)^2 = E\left(\sum_{i=1}^{n} X_i^2 + 2\sum_{1\le i<j\le n}^{n} X_i X_j\right)$$

$$= \sum_{i=1}^{n} E(X_i^2) + 2\sum_{1\le i<j\le n}^{n} E(X_i X_j)$$

X_i^2	1	0
P	$\frac{1}{n}$	$1-\frac{1}{n}$

$$E(X_i^2)=\frac{1}{n} \qquad i=1,2,\cdots,n$$

X_iX_j	1	0
P	$\frac{1}{n(n-1)}$	$1-\frac{1}{n(n-1)}$

$$E(X_iX_j)=\frac{1}{n(n-1)} \qquad i,j=1,2,\cdots,n$$

$$E(X^2)=\sum_{i=1}^{n}E(X_i^2)+2\sum_{1\le i<j\le n}^{n}E(X_iX_j)$$

$$=\sum_{i=1}^{n}\frac{1}{n}+2\sum_{1\le i<j\le n}^{n}\frac{1}{n(n-1)}$$

$$=n\cdot\frac{1}{n}+2\cdot C_n^2\cdot\frac{1}{n(n-1)}$$

$$=2$$

$$D(X)=E(X^2)-E^2(X)=1$$

标准化随机变量

设随机变量 X 的期望$E(X)$、方差$D(X)$都存在, 且$D(X)\neq 0$, 则称

$$X^* = \frac{X - E(X)}{\sqrt{D(X)}}$$

为 X 的标准化随机变量. 显然,

$$E(X^*) = 0, \quad D(X^*) = 1$$

仅知 r.v. 的期望与方差并不能确定其分布

例如

X	-1	0	1
P	0.1	0.8	0.1

$$E(X)=0, D(X)=0.2$$

与

Y	-2	0	2
P	0.025	0.95	0.025

$$E(Y)=0, D(Y)=0.2$$

有相同的期望方差但是分布却不相同

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.

例7 已知 X 服从正态分布, $E(X)=1.7$, $D(X)=3$, $Y=1-2X$, 求Y的密度函数.

解 $E(Y)=1-2\times1.7=-2.4,$

$D(Y)=4\times3=12$

$$f_Y(y)=\frac{1}{2\sqrt{6\pi}}e^{-\frac{(y+2.4)^2}{24}},\quad -\infty<y<+\infty$$

作业 P.170 习题三

9 11 16

17 19 21

附例 在 [0, 1] 中随机地取两个数 X, Y，求 $D(\min\{X, Y\})$

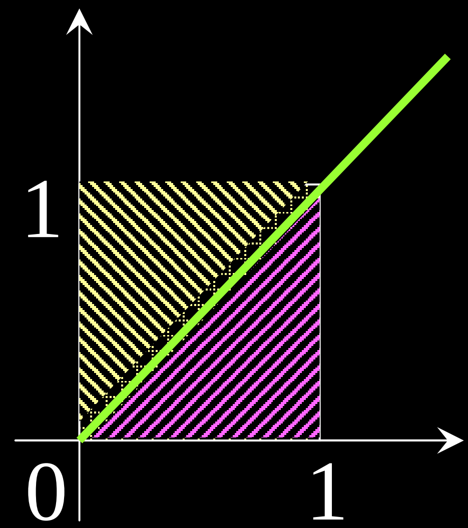
解

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

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

$$=\int_0^1\left(\int_x^1 xdy\right)dx+\int_0^1\left(\int_y^1 ydx\right)dy$$

$$=1/3.$$



$$E(\min{}^2\{X,Y\})$$

$$= \int_0^1 \left(\int_x^1 x^2 dy \right) dx + \int_0^1 \left(\int_y^1 y^2 dx \right) dy$$

$$= 1/6.$$

$$D(\min\{X,Y\})$$

$$= E\left(\min{}^2\{X,Y\}\right) - E^2(\min\{X,Y\})$$

$$= 1/18.$$

例8 已知 X 的 d.f.为

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

其中 A,B 是常数，且 $E(X)=0.5.$

(1) 求 $A,B.$

(1) 设 $Y=X^2$, 求 $E(Y),D(Y)$

解 (1)

$$\left.\begin{aligned}&\int_{-\infty}^{+\infty} f(x)dx=\int_0^1 (Ax^2+Bx)dx=1\\&\int_{-\infty}^{+\infty} xf(x)dx=\int_0^1 x(Ax^2+Bx)dx=\frac{1}{2}\end{aligned}\right\}\Rightarrow$$

$$\left.\begin{aligned}&\frac{A}{3}+\frac{B}{2}=1\\&\frac{A}{4}+\frac{B}{3}=\frac{1}{2}\end{aligned}\right\}\Rightarrow\quad \begin{aligned}&A=-6,\\&B=6\end{aligned}$$

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

$$=\int_0^1 x^2(-6x^2+6x)dx=3/10.$$

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

$$=\int_0^1 x^4(-6x^2+6x)dx=1/7.$$

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