§4.2 方差

引例 甲、乙两射手各打了10发子弹，每发子弹击中的环数分别为：

甲	10, 6, 7, 10, 8, 9, 9, 10, 5, 10
乙	8, 7, 9, 10, 9, 8, 7, 9, 8, 9

有六个不同数据

仅有四个不同数据

问哪一个射手的技术较好？

解 首先比较平均环数

$$\overline{甲} = 8.4, \qquad \overline{乙} = 8.4$$

再比较稳定程度

$$\sum_{i=1}^{10}(x_i-\overline{x})$$

$$\sum_{i=1}^{10}|x_i-\overline{x}|$$

$$\sum_{i=1}^{10}(x_i-\overline{x})^2$$

$$\sum_{i=1}^{10}(x_i-\bar{x})^2$$

甲：

$$4\times(10-8.4)^2+2\times(9-8.4)^2+(8-8.4)^2$$
$$+(7-8.4)^2+(6-8.4)^2+(5-8.4)^2$$
$$=30.4$$

乙：

$$(10-8.4)^2+4\times(9-8.4)^2$$
$$+3\times(8-8.4)^2+2\times(7-8.4)^2$$
$$=6.44$$

乙比甲技术稳定

进一步比较平均偏离平均值的程度

甲 $$\frac{1}{10}\{4\times(10-8.4)^2+2\times(9-8.4)^2+(8-8.4)^2$$
$$+(7-8.4)^2+(6-8.4)^2+(5-8.4)^2\}$$
$$=3.04 \triangleq \sum_{k=1}^{6}\left(x_k-E(X)\right)^2 p_k$$

乙 $$\frac{1}{10}\{(10-8.4)^2+4\times(9-8.4)^2$$
$$+3\times(8-8.4)^2+2\times(7-8.4)^2\}$$
$$=0.644 \triangleq \sum_{k=1}^{4}\left(x_k-E(X)\right)^2 p_k$$

方差的概念

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

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

称 $\sqrt{D(X)}$ 为X的**均方差**.

$(X-E(X))^2$ —— 随机变量X的取值偏离平均值的情况，是X的函数，也是随机变量

$E(X-E(X))^2$ —— 随机变量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.**，概率分布为

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

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

若 X 为连续型，概率密度为 $f(x)$

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

常用的计算方差的公式：

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

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

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

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

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

方差的计算

例1 设$X \sim P(\lambda)$, 求$D(X)$.

解

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

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

$$=\lambda e^{-\lambda} \sum_{m=0}^{+\infty} \frac{\lambda^{m}}{m!} \quad =\lambda e^{-\lambda} \cdot e^{\lambda}$$

$$=\lambda$$

$$E(X^2) = E(X(X-1)) + \overbrace{E(X)}^{=\lambda}$$

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

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

$$\Longrightarrow \quad E(X^2) = \lambda^2 + \lambda$$

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

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

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

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

$$= \sigma^2 \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} y^2 e^{-\frac{y^2}{2}} dy = \sigma^2 \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} y^2 \frac{1}{-y} de^{-\frac{y^2}{2}}$$

$$= -\sigma^2 \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} y de^{-\frac{y^2}{2}} = -\sigma^2 \frac{1}{\sqrt{2\pi}} \left[ye^{-\frac{y^2}{2}} \Big|_{-\infty}^{+\infty} - \int_{-\infty}^{+\infty} e^{-\frac{y^2}{2}} dy \right]$$

$$= \sigma^2 \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} e^{-\frac{y^2}{2}} dy \quad = \sigma^2$$

常见随机变量的方差

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

方差的性质

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

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

性质 1 ：$D(C)=0$

证明： $D(C)=E(C-E(C))^2=0$

性质 2 ： $D(aX)=a^2D(X)$

证明：$D(aX+b)=E((aX+b)-E(aX+b))^2$

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

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

性质 3 $:D(X \pm Y) = D(X) + D(Y)$

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

证明：

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

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

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

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

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

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

注意到， $E\big((X - E(X))(Y - E(Y))\big)$

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

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

$$= D(X) + D(Y)$$
$$\pm 2\big(E(XY) - EX \cdot EY\big)$$

特别地，若X,Y相互独立，则

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

若X,Y相互独立

$$\overset{\Longrightarrow}{\not\Longleftarrow} \quad D(X \pm Y) = D(X) + D(Y)$$

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

若 $X_1, X_2, \cdots, X_n$ 相互独立，$a_1, a_2, \cdots, a_n, b$ 为常数

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

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

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

$$= \sum_{i=1}^{n} a_i^2 D(X_i)$$

❑ 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

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$C \ne E(X)$时，$(C-E(X))^2 > 0$

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

❑ $D(X) = 0 \Longleftrightarrow P(X = E(X)) = 1$

称为X依概率1等于常数$E(X)$

例 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 $X = \sum_{i=1}^{n} X_i$

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

$$p + q = 1$$

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

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

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

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

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

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

故

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

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

例 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	$\frac{1}{n}$	$1-\frac{1}{n}$

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

$$E(X_i)=1\cdot\frac{1}{n}+0\cdot(1-\frac{1}{n}) \quad =\frac{1}{n}$$

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

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

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

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

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

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

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

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

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

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

$$=1+1=2$$

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

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，

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

仅知随机变量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

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

与

Y	-2	0	2
P	0.025	0.95	0.025

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

它们有相同的期望、方差但是分布却不同

但若已知分布的类型，及期望和方差，常能确定分布.

例 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 $$E(Y)=1-2\times 1.7 \quad =-2.4,$$

$$D(Y)=2^2\times 3 \quad =12$$

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

$$-\infty<y<+\infty$$