4 随机变量的数字特征

数学期望、方差、矩、协方差、相关系数。

4.0 数学期望

4.0.0 离散型数学期望

> **定义1：** 设离散型随机变量 X 的概率分布为
>
> | X | x_1 | x_2 | $\dots$ | x_n | $\dots$ |
> |---|---|---|---|---|---|
> | P | p_1 | p_2 | $\dots$ | p_n | $\dots$ |
>
> 其中，$p_i \geqslant 0, \sum_{i=1}^{\infty} p_i = 1$ 。若级数 $\sum_{i=1}^{\infty} x_i p_i$ 绝对收敛，则称级数 $\sum_{i=1}^{\infty} x_i p_i$ 为 X 的数学期望（简称均值或期望），记为 $E(X)$，即 $E(X) = \sum_{i=1}^{\infty} x_i p_i$ 。
>
> **注：** 若级数 $\sum_{i=1}^{\infty} x_i p_i$ 不是绝对收敛，则 X 的数学期望不存在。

4.0.0.0 两点分布

X	0	1
P	$1-p$	p

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

4.0.0.1 二项分布

设有二项分布 $X \sim B(n,p)$，其概率分布为

$$P\{X=i\} = C_n^i p^i (1-p)^{n-i}, \quad (i = 0, 1, 2, \dots)\,.$$

$$\begin{aligned}
E(X) &= \sum_{i=0}^{n} x_i p_i = \sum_{i=0}^{n} i C_n^i p^i (1-p)^{n-i} \\
&= np \sum_{i=1}^{n} \frac{(n-1)!}{(i-1)![(n-1)-(i-1)]!} p^{i-1} (1-p)^{(n-1)-(i-1)} \\
&\overset{k=i-1}{=} np \sum_{k=0}^{n-1} C_{n-1}^k p^k (1-p)^{(n-1)-k} \\
&= np[p + (1-p)]^{n-1} \\
&= np\,.
\end{aligned}$$

4.0.0.2 泊松分布

设有泊松分布 $X \sim P(\lambda)$，其概率分布为

$$P\{X=i\} = \frac{\lambda^i e^{-\lambda}}{i!}, \quad (i = 0, 1, 2, \dots)\,.$$

$$\begin{aligned}
E(X) &= \sum_{i=0}^{\infty} x_i p_i = \sum_{i=0}^{\infty} i \frac{\lambda^i e^{-\lambda}}{i!} \\
&= \lambda e^{-\lambda} \sum_{i=1}^{\infty} \frac{\lambda^{i-1}}{(i-1)!} = \lambda e^{-\lambda} e^{\lambda} = \lambda\,.
\end{aligned}$$

4.0.1 连续型数学期望

> **定义2：** 设连续型随机变量 X 的概率密度为 $f(x)$ ，若广义积分 $\int_{-\infty}^{+\infty} xf(x)dx$ 绝对收敛，则称积分 $\int_{-\infty}^{+\infty} xf(x)dx$ 为 X 的数学期望，记为 $E(X)$ ，即
>
> $$E(X)=\int_{-\infty}^{+\infty} xf(x)dx\ .$$
>
> **注：** 若广义积分 $\int_{-\infty}^{+\infty} xf(x)dx$ 不是绝对收敛，则 X 的数学期望不存在。

4.0.1.0 均匀分布

设随机变量 X 服从 $[a,b]$ 上的均匀分布，其概率密度为

$$f(x)=\begin{cases}\frac{1}{b-a}, & a\leqslant x\leqslant b\ ;\\ 0, & \text{其 他}\ .\end{cases}$$

$$\begin{aligned}E(X)&=\int_{-\infty}^{+\infty} xf(x)=\int_a^b x\frac{1}{b-a}dx\\ &=\frac{1}{b-a}\cdot\frac{1}{2}(b^2-a^2)=\frac{a+b}{2}\ .\end{aligned}$$

4.0.1.1 指数分布

设随机变量 X 服从参数为 $\lambda>0$ 的指数分布，其概率密度为

$$f(x)=\begin{cases}\lambda e^{-\lambda x}, & x\geqslant 0\ ;\\ 0, & x<0.\end{cases}$$

$$E(X)=\int_{-\infty}^{+\infty} xf(x)dx=\int_0^{+\infty} x\lambda e^{-\lambda x}dx=\frac{1}{\lambda}\ .$$

4.0.1.2 正态分布

设随机变量 $X\sim N(\mu,\sigma^2)$ ，其概率密度为

$$f(x)=\frac{1}{\sqrt{2\pi}\cdot\sigma}e^{-\frac{(x-\mu)^2}{2\sigma^2}}\ ,$$

$$E(X)=\mu\ .$$

注： $\int_{-\infty}^{+\infty} e^{-x^2}dx=\sqrt{\pi}$ 。

4.0.2 二维随机变量的数学期望

1. 设二维离散型随机变量 (X,Y) 的联合概率分布为 $p_{ij}=P(X=x_i,Y=y_j),i,j=1,2,\ldots$ 则随机变量 X,Y 的数学期望分别定义如下:

$$E(X)=\sum_i x_iP(X=x_i)=\sum_i\sum_j x_ip_{ij}\ ;$$
$$E(Y)=\sum_j y_jP(Y=y_j)=\sum_j\sum_i y_jp_{ij}\ .$$

   上述公式与一维离散型随机变量的数学期望的定义一致，公式中的和式可以是有限项的和，也可以是级数的和，假定级数是绝对收敛的。

2. 设二维连续型随机变量 (X,Y) 的联合概率密度为 $f(x,y)$ ，则

$$E(X)=\int_{-\infty}^{+\infty} x f_X(x)dx=\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} x f(x,y)dxdy\ ,$$
$$E(Y)=\int_{-\infty}^{+\infty} y f_Y(y)dy=\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} y f(x,y)dxdy\ .$$

上述公式与一维连续型随机变量的数学期望的定义一致，假定广义积分是绝对收敛的。

4.0.3 随机变量函数的数学期望

> **定理1：** 设 Y 是随机变量 X 的函数：$Y=g(X)(g\text{是 连 续 函 数})$
>
> 1. 设离散型随机变量 X 的概率分布为
>
> $$P(X=x_k)=p_k\ ,\ \ k=1,2,\ldots$$
>
> 若无穷级数 $\sum_{k=1}^{\infty} g(x_k)p_k$ 绝对收敛，则有
>
> $$E(Y)=E\big[g(X)\big]=\sum_{k=1}^{\infty} g(x_k)p_k$$
>
> 2. 设连续型随机变量 X 的概率密度为 $f(x)$，若广义积分 $\int_{-\infty}^{+\infty} g(x)f(x)dx$ 绝对收敛，则有
>
> $$E(Y)=E\big[g(X)\big]=\int_{-\infty}^{+\infty} g(x)f(x)dx$$

> **定理2：** 设 $Z=g(X,Y)$ 是二维随机变量 (X,Y) 的函数（g 是连续函数）
>
> 1. 设离散型随机变量 (X,Y) 的联合概率密度分布为
>
> $$p_{ij}=P(X=x_i,Y=y_j),i,j=1,2,\ldots$$
>
> 若级数 $\sum\limits_i\sum\limits_j g(x_i,y_i)p_{ij}$ 绝对收敛，则有
>
> $$E\big[g(X,Y)\big]=\sum_i\sum_j g(x_i,y_j)p_{ij}$$

2. 设连续型随机变量 (X,Y) 的联合概率密度为 $f(x,y)$ ，若广义积分 $\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} g(x,y)f(x,y)dxdy$ 绝对收敛，则

$$E[g(X,Y)] = \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} g(x,y)f(x,y)dxdy\,.$$

4.0.4 数学期望的性质

性质1： 设 C 为一常数，则 $E(C)=C$ 。

性质2： 设 X 为随机变量，C 为常数，则 $E(CX)=CE(X)$ 。

性质3： 设 X,Y 为任意两个随机变量，则有 $E(X+Y)=E(X)+E(Y)$ 。

性质4： 设随机变量 X,Y 相互独立，则 $E(XY)=E(X)E(Y)$ 。

4.1 方差和矩

4.1.0 方差的定义

定义3： 设 X 是随机变量，若 $E[X-E(X)]^2$ 存在，则称 $E[X-E(X)]^2$ 为随机变量 X 的方差，记作 $D(X)$ 即

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

称 $\sqrt{D(X)}$ 为 X 的标准差或均方差，标准差与 X 具有相同的量纲。

4.1.1 方差的性质

性质1： $D(C)=0$

性质2： $D(X+C)=D(X)$

性质3： $D(CX)=C^2D(X)$

性质4： 若随机变量 X,Y 相互独立，则 $D(X\pm Y)=D(X)+D(Y)$ 。（也可以推广到 n 个随机变量）

性质5： $D(X)=0$ 的充要条件是 $P(X=E(X))=1$ 。

4.1.2 常见分布的方差及方差性质的应用

- 两点分布：$E(X)=p$, $D(X)=p(1-p)$
- 二项分布：对 $X\sim B(n,p)$ ，$E(X)=np$, $D(X)=np(1-p)$
- 泊松分布：对 $X\sim P(\lambda)$ ，$E(X)=\lambda$, $D(X)=\lambda$
- 均匀分布：对 $X\sim U[a,b]$ ，$E(X)=\frac{a+b}{2}$, $D(X)=\frac{(b-a)^2}{12}$
- 指数分布：对 $X\sim e(\lambda)$ ，$E(X)=\frac{1}{\lambda}$, $D(X)=\frac{1}{\lambda^2}$
- 正太分布：对 $X\sim N(\mu,\sigma^2)$ ，$E(X)=\mu$, $D(X)=\sigma^2$

已知随机变量 X 的数学期望为 $E(X)$ ，方差为 $D(X)>0$ ，设

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

有

> $$E(X^*)=E\Big[\frac{X-E(X)}{\sqrt{D(X)}}\Big]=\frac{E[X-E(X)]}{\sqrt{D(X)}}=\frac{E(X)-E(X)}{\sqrt{D(X)}}=0\,;$$
> $$D(X^*)=D\Big[\frac{X-E(X)}{\sqrt{D(X)}}\Big]=\frac{D[X-E(X)]}{\sqrt{D(X)}}=\frac{D(X)}{D(X)}=1\,.$$
>
> 这里 $X^*=\frac{X-E(X)}{\sqrt{D(X)}}$ 叫做 X 的标准化随机变量，特别地，若 $X\sim N(\mu,\sigma^2)$ ，则
>
> $$X^*=\frac{X-\mu}{\sigma}\sim N(0,1)\,.$$

4.1.3 原点矩和中心矩

> **定义4：** 设 X 为随机变量，若 $E(|X|^k)<\infty,k=1,2,\ldots$ 则称 $v_k=E(X^k)$ 为 X 的 k 阶原点矩，称 $\alpha_k=E(|X|^k)$ 为 X 的 k 阶原点绝对矩 $(k=1,2,\ldots)$ 。

> **定义5：** 设 X 为随机变量，若 $E(|X-E(X)|^k)<\infty$ 则称 $\mu_k=E(X-E(X))^k$ 为 X 的 k 阶中心矩，称 $\beta_k=E(|X-E(X)|^k)$ 为 X 的 k 阶中心绝对矩 $(k=1,2,\ldots)$ 。

显然 $v_1=E(X)$ ，$\mu_2=D(X)$ ，即随机变量 X 的一阶原点矩就是其数学期望，二阶中心矩就是其方差。

4.2 协方差与相关系数

4.2.0 协方差

> **定义6：** 设二维随机变量 (X,Y) ，若 $E\{[X-E(X)][Y-E(Y)]\}$ 存在，则称它为随机变量 X 与 Y 的协方差，记作 $Cov(X,Y)$ ，即
>
> $$Cov(X,Y)=E\{[X-E(X)][Y-E(Y)]\}=E(XY)-E(X)E(Y)\,.$$

由协方差的定义可得下面的性质：

1. $Cov(X,Y)=Cov(Y,X)$;
2. $Cov(aX,bY)=abCov(X,Y)$ ，其中 a,b 为常数；
3. $Cov(X_1+X_2,Y)=Cov(X_1,Y)+Cov(X_2,Y)$ ；
4. 若 X 与 Y 独立，则 $Cov(X,Y)=0$ ，反之不成立；
5. $Cov(X,X)=E\{[X-E(X)][X-E(X)]\}=D(X)$ 。
6. $D(X\pm Y)=D(X)+D(Y)\pm 2Cov(X,Y)$ ，若 X 与 Y 独立，则 $Cov(X,Y)=0$ ，从而 $D(X\pm Y)=D(X)+D(Y)$ 。

4.2.1 相关系数

协方差的带有量纲的量，它的量纲等于随机变量 X 与 Y 的量纲的乘积。而为了更好地反映 X 与 Y 之间的相关性，对 X,Y 分别标准化后，再求它们的协方差，并称之为 X 和 Y 的相关系数。

设 $X^*=\frac{X-E(X)}{\sqrt{D(X)}}$ ，$Y^*=\frac{Y-E(Y)}{\sqrt{D(Y)}}$ ，则 $E(X^*)=E(Y^*)=0$ ，从而有

$$Cov(X^*,Y^*)=E\{[X^*-E(X^*)][Y^*-E(Y^*)]\}=\cdots=\frac{Cov(X,Y)}{\sqrt{D(X)D(Y)}}$$

> **定义7：** 若 $D(X)>0,D(Y)>0$ ，称 $\frac{Cov(X,Y)}{\sqrt{D(X)D(Y)}}$ 为 X 与 Y 的相关系数，记为 ρ_{XY} ，即
>
> $$\rho_{XY}=\frac{Cov(X,Y)}{\sqrt{D(X)D(Y)}}\,.$$

相关系数的性质如下：

1. $|\rho_{XY}|\leqslant 1$ ；

2. $|\rho_{XY}|=1$ 的充要条件是存在常数 $a,b(a\neq 0)$，使得 $P\{Y=aX+b\}=1$；
3. 若 X 与 Y 相互独立，则 $Cov(X,Y)=0$，从而 $\rho_{XY}=0$。

> **定义8：** 若相关系数 $\rho_{XY}=0$，则称 X 与 Y 不相关，即线性无关。