数学期望的运算性质

性质 1：任意常数 c 的数学期望等于 c

证明：随机变量 X 服从

$$P(X) = \begin{cases} 1, & X = c \\ 0, & \text{else} \end{cases}$$

所以其期望就是

$$E[X] = c \times 1 + 0 = c$$

数学期望的运算性质

性质 2 (线性): 设随机变量 X 和 Y 的期望都存在, a, b 都是常数, 那么

$$E[aX + bY] = aE[X] + bE[Y]$$

数学期望的运算性质

证明：只对连续型的随机变量证明。
先证明随机变量 $Z = aX + bY$ 的期望存在。假设 X, Y 的联合密度函数为 $f_{X,Y}(x, y)$，判断下面积分的收敛性

$$
\begin{aligned}
&\int_{\mathbb{R}^2} ax + byf_{X,Y}(x,y)dxdy \\
&\leq a\int_{\mathbb{R}^2} xf_{X,Y}(x,y)dxdy + b\int_{\mathbb{R}^2} yf_{X,Y}(x,y)dxdy \\
&= a\int_{-\infty}^{\infty} xf_X(x)dx + b\int_{-\infty}^{\infty} yf_Y(y)dy \\
&< \infty
\end{aligned}
$$

数学期望的运算性质

上面的式子表明 $E[Z]$ 存在，根据积分运算的线性性

$$
\begin{aligned}
E[aX + bY] &= \int_{2} (ax + by) f_{X,Y}(x,y)\,dxdy \\
&= a \int_{2} x f_{X,Y}(x,y)\,dxdy + b \int_{2} y f_{X,Y}(x,y)\,dxdy \\
&= a \int_{-\infty}^{\infty} x f_X(x)\,dx + b \int_{-\infty}^{\infty} y f_Y(y)\,dy \\
&= aE[X] + bE[Y]
\end{aligned}
$$

数学期望的运算性质

性质 3: 若随机变量 X 与 Y 独立，且 X 与 Y 的期望都存在，那么

$$E[XY] = E[X]E[Y].$$

方差

随机变量的均值只能反映随机变量的平均预期，但却无法从中看出其取值的波动性大小，即随机试验的结果偏离平均值的程度。

方差的例子

巧克力加工厂 A，B 分别包装一带巧克力的重量
$X_A \sim U[50-0.5, 50+0.5]$，
$X_B \sim U[50-1.5, 50+1.5]$，那么

$$E[X_A] = E[X_B] = 50,$$

从均值来看，这两个加工厂的包装水平是一样的，但是

$$P(X_A \in [50-1.5, 50-0.5] \cup [50+0.5, 50+1.5]) = 0$$

$$P(X_B \in [50-1.5, 50-0.5] \cup [50+0.5, 50+1.5]) = \frac{2}{103}$$

方差的定义

我们用 方差 来描述一个随机变量的波动性。

定义 1.4.3: 设随机变量 X 有有限的数学期望，如果 $E[(X-E[X])^2]<\infty$，则称

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

为 X 的 方差 ，而称 $\sqrt{Var[X]}$ 为 X 的 标准差 ，记为 $\sigma[X]$.

方差的计算

- 对于离散型的随机变量:

$$Var[X] = \sum_{i=1}^{\infty}(x_i - E[X])^2 p_i$$

- 对于连续型的随机变量:

$$Var[X] = \int_{-\infty}^{\infty}(x - E[X])^2 f_X(x)dx$$

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

例 4.1.9

$X \sim B(1, p)$ 有 $Var[X] = p(1-p)$

$P(X=k) = C_n^k \cdot p^k \cdot (1-p)^{n-k}$.

$E[X^2] = \sum_{k=0} k^2 \cdot P(X=k) = n(n-1)p^2 + np$

$\underline{\underline{q=1-p}} \sum_{k=0} k^2 \frac{n!}{k!(n-k)!} p^k \cdot q^{n-k}$

$= \sum_{k=1} k \frac{n!}{(k-1)!(n-k)!} p^k \cdot q^{n-k}$

$= \sum_{k=1} [(k-1)+1] \frac{\cdots}{(k-1)! \cdots} \cdots$

$= \sum_{k=1} (k-1) \frac{\cdots}{(k-1)! \cdots} p \cdots + \sum_{k=1} \frac{\cdots}{(k-1)!} \cdots$

$= \sum_{k=2} \frac{n!}{(k-2)!(n-k)!} p^k \cdot q^{n-k} + \sum_{k=1} \frac{n!}{(k-1)!(n-k)!} p^k \cdot q^{n-k}$

$= n(n-1)p^2 \underbrace{\sum_{k=2}^{n} \frac{(n-2)!}{(k-2)!((n-2)-(k-2))!} p^{k-2} \cdot q^{(n-2)-(k-2)}}_{\underline{l=k-2}\ (p+q)^{n-2} = 1} + np \underbrace{\sum_{k=1} \frac{(n-1)!}{(k-1)!((n-1)-(k-1))!} p^{k-1} \cdot q^{n-1-(k-1)}}_{=1}$

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

$= n(n-1)p^2 + np - n^2p^2$

$= np - np^2$

$= np(1-p)$.

#

例 4.1.10

$X \sim U[a, b]$ 有 $Var[X] = \frac{1}{12}(b-a)^2$

$$f_X(x) = \begin{cases} \frac{1}{b-a}, & a \le x \le b \\ 0 & \text{otherwise} \end{cases}$$

$$Y = X - \frac{b-a}{2}$$

$$Y \sim U[-c, c] \quad c = \frac{b-a}{2}$$

$$\mathbb{E}[X] = \frac{a+b}{2}$$

$$\mathbb{E}[X^2] = \int_a^b x^2 \frac{1}{b-a} dx$$

$$= \frac{1}{3(b-a)} x^3 \Big|_a^b$$

$$= \frac{1}{3}(a^2+ab+b^2)$$

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

$$= \frac{1}{3}(a^2+ab+b^2) - \frac{1}{4}(a^2+2ab+b^2)$$

$$= \frac{1}{12}(4a^2+4ab+4b^2-3a^2-6ab-3b^2)$$

$$= \frac{1}{12}(a^2-2ab+b^2) = \frac{1}{12}(a-b)^2$$

例 4.1.11

$X \sim \mathrm{Exp}(\lambda)$ 有 $Var[X] = \frac{1}{\lambda^2}$

$$f_X(x) = \begin{cases} \lambda \cdot e^{-\lambda x} & x \geq 0 \\ 0 & x < 0 \end{cases}$$

$$\mathbb{E}[X] = \frac{1}{\lambda}.$$

$$\mathbb{E}[X^2] = \int_0^{+\infty} x^2 \cdot \lambda e^{-\lambda x} dx$$

$$= \frac{1}{\lambda^2} \int_0^{+\infty} (\lambda x)^2 e^{-\lambda x} d\lambda x$$

$$= \frac{1}{\lambda^2} \int_0^{+\infty} y^{3-1} e^{-y x} dx.$$

$$= \frac{1}{\lambda^2} \Gamma(3) = \frac{1}{\lambda^2} 2!$$

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

$$= \frac{2}{\lambda^2} - \left(\frac{1}{\lambda}\right)^2 = \frac{1}{\lambda^2} \qquad \#$$

例 4.1.12

$X \sim \mathrm{Pois}(\lambda)$ 有 $Var[X] = \lambda$

$P(X=k) = \frac{1}{k!}\lambda^k \cdot e^{-\lambda} \quad (k \geq 0)$

$\mathbb{E}[X] = \lambda$

$\mathbb{E}[X^2] = \sum_{k=0}^{+\infty} k^2 \cdot \frac{1}{k!}\lambda^k \cdot e^{-\lambda}$

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

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

$= \cdots$

例

$X \sim \mathrm{Geo}(p)$ 有 $Var[X] = \frac{1-p}{p^2}$

$P(X=k) = (1-p)^{k-1} \cdot p \quad (k \geq 0)$

$E[X] = \frac{1}{p}$

$E[X^2] = \sum_{k=0}^{+\infty} k^2 \cdot (1-p)^{k-1} \cdot p$

$= (1-q) \cdot \sum_{k=1} k^2 q^{k-1}$

$= \sum_{k=1} k^2 q^{k-1} - \sum_{k=1} k^2 q^k$

$= 1 + \sum_{k=1} (k+1)^2 q^k - \sum_{k=1} k^2 q^k$

$= 1 + \sum_{k=1} (2k+1) q^k$

$= 1 + 2\sum_{1} k q^k + \sum_{0} q^{k+1} = 1 + 2 \cdot \sum_{1} k q^k + \frac{q}{1-q}$

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

$E[X] = \sum_{k=0} k q^{k-1} p = \frac{1}{p}$

$\sum_{k=1} k q^k = \frac{q}{p^2}$

$\therefore Var[X] = E[X^2] - E[X]^2$

$= \frac{1}{p^2}(2-p) - \left(\frac{1}{p}\right)^2$

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

#

例 4.1.13

$X \sim N(\mu, \sigma^2)$ 有 $Var[X] = \sigma^2$

矩

$\int_{-\infty}^{+\infty} f_X(x) \cdot |X - E[X]| \, dx = E_X[\underbrace{|X - E[X]|}]$

定义 4.1.4: 设 X 为随机变量，c 为常数，k 为正整数，如果 $E[(X-c)^k] < \infty$，那么称

$$E[(X-c)^k]$$

为 X 关于 c 的 k 阶矩。

- $c = 0$，称 $E[X^k]$ 为 X 的 k 阶原点矩，例如均值 (1 阶原点矩)。
- $c = E[X]$，$E[(X - E[X])^k]$ 称为 X 的 k 阶中心矩

  $k=0$
  - $k = 1$ $E[X - EX] = EX - EX = 0$
  - $k = 2$ $E[(X-EX)^2] = \mathrm{Var}X$

“标准化”随机变量

对随机变量 X，如果 $Var[X] > 0$, 我们考虑其标准化

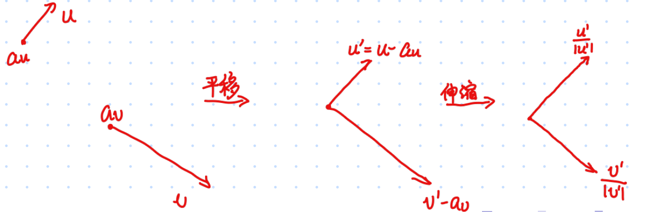
$$\bar{X} := \frac{1}{Var[X]^{\frac{1}{2}}}(X - E[X]). = \frac{X-\mu_X}{\sigma_X}$$

可见，

$$E[\bar{X}] = E[\frac{1}{\sigma_X}(X-\mu_X)] = \frac{1}{\sigma_X}E[X-\mu_X] = 0$$

- $E[\bar{X}] = 0$
- $E[\bar{X}]^2 = 1$ $\quad Var[\bar{X}] = \frac{1}{\sigma_X^2}Var[(X-\mu_X)] = \frac{1}{\sigma_X^2}Var[X] = \frac{1}{\sigma_X^2}\cdot\sigma_X^2 = 1.$

我们希望进一步刻画 $\bar{X}$ 的数字特征。



高阶矩

对于三阶和四阶中心矩，概率学上有明确的定义

定义 4.1.5：设 X 为随机变量，如果 $E[X^4] < \infty$，则称

$$E[\bar{X}^3] = \frac{E[(X - E[X])^3]}{(Var[X])^{\frac{3}{2}}}$$

为 X 的 偏度 ， 偏度刻画 X 的分布的偏斜程度。称

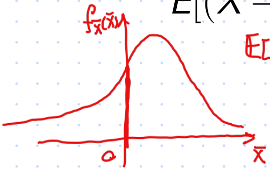
$$E[\bar{X}^4] = \frac{E[(X - E[X])^4]}{(Var[X])^2}$$

为 X 的 峰度 ， 峰度刻画随机变量的分布在其均值附近的陡峭程度。

高阶矩

当密度函数 $f_X(x)$ 关于 $E[X]$ 对称时，偏度为零。

$E[\bar{X}^3]=\int_{-\infty}^{+\infty}\bar{x}^3\cdot f_{\bar{X}}(\bar{x})d\bar{x}=\int_{-\infty}^{0}+\int_{0}^{+\infty}=\int_{0}^{+\infty}(-\bar{x})^3 f_{\bar{X}}(-\bar{x})dx+\int_{0}^{+\infty}\bar{x}^3\cdot f_{\bar{X}}(\bar{x})d\bar{x}=0$



$E[\bar{X}^3] \begin{smallmatrix} > \\ = \\ < \end{smallmatrix} 0$

$$
\begin{aligned}
E[(X-E[X])^3] &= \int_{-\infty}^{\infty}(x-c)^3 f_X(x)dx \\
&= \int_{-\infty}^{\infty} x^3 f_X(x+c)dx
\end{aligned}
$$

记 $g(x)=f_X(x+c)$，根据 $f_X(x+c)=f_X(c-x)$，知 $g(x)=g(-x)$，即 $g(x)$ 是偶函数，且 $E[X^3]<\infty$

$$
E[(X-E[X])^3]=\left(\int_{-\infty}^{0}+\int_{0}^{\infty}\right)x^3 g(x)dx=0
$$

计算偏度和峰度

$$U(a, b)$$

计算偏度和峰度

$$Exp(\lambda)$$

计算偏度和峰度

$$N(\mu, \sigma)$$

Outline

数字特征

对于二维的随机变量 (X,Y)，假设 $E[g(X,Y)]$ 存在，那么

- 设离散型随机变量 (X,Y) 有概率分布 $P(X=x_i,Y=y_i)=p_{ij}$，$i,j=1,2,\cdots$，那么
$$E[g(X,Y)]=\sum_{i=1}^{\infty}\sum_{j=1}^{\infty}g(x_i,y_j)p_{ij},$$
- 设连续型随机变量 (X,Y) 有分布密度函数 $f_{X,Y}$，那么
$$E[g(X,Y)]=\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}g(x,y)f_{X,Y}(x,y)dxdy$$

练习

假设 (X,Y) 的联合分布列为

$$P(X=0,Y=0)=P(X=1,Y=1)$$
$$=P(X=0,Y=1)=P(X=1,Y=0)=0.25$$

求 $E\left(\sin\left(\frac{\pi}{2}(X+Y)\right)\right)$

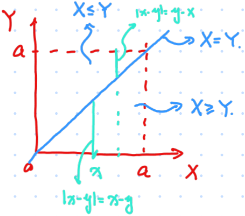
$$X+Y \sim \begin{pmatrix} 0 & 1 & 2 \\ \frac{1}{4} & \frac{1}{2} & \frac{1}{4} \end{pmatrix}$$

$$\sin\left(\tfrac{\pi}{2}(X+Y)\right) \sim \begin{pmatrix} 0 & 1 \\ \frac{1}{2} & \frac{1}{2} \end{pmatrix}$$

$$E[\ldots] = \frac{1}{2}$$

练习

在长度为 a 的线段上任取两个点 X 与 Y，求这两个点之间的平均长度。



$X, Y \sim U[0,a] \Rightarrow f_X = f_Y = \frac{1}{a}$

$Z = |X-Y|$

$f_{X,Y}(x,y) = f_X(x)\cdot f_Y(y) = \frac{1}{a}\times\frac{1}{a} = \frac{1}{a^2}$

$E[Z] = \int_{\mathbb{R}^2} z \cdot f_{X,Y}(x,y)\,dx\,dy$

$= \int_0^a \int_0^a |x-y| \frac{1}{a^2}\,dx\,dy$

$= \frac{1}{a^2}\int_0^a \int_0^x (x-y)\,dx\,dy + \frac{1}{a^2}\int_0^a \int_x^a (y-x)\,dx\,dy$

$= \cdots$

协方差与相关系数

定义 4.2.1: 设 (X, Y) 为二维随机向量，且 $Var[X] < \infty$， $Var[Y] < \infty$，

- 称
$$Cov(X, Y) = E[(X - E[X])(Y - E[Y])]$$
为 X 与 Y 的 协方差 。

$Cov(X,X) = E[(X-EX)^2] = Var[X]$

$r(X,X) = \frac{Cov(X,X)}{Var[X]} = 1.$

- 称
$$r(X, Y) = \frac{Cov(X, Y)}{\sqrt{Var[X]}\sqrt{Var[Y]}}$$
为 X 与 Y 的 相关系数 。

- 若 $r(X, Y) = 0$，称 X 与 Y 不相关 。

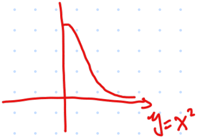
协方差的意义

- $Cov(X, Y) > 0$，则 X 与 Y 有同增同减的倾向。
- $Cov(X, Y) < 0$，则 X 与 Y 有此消彼涨的倾向。
- $Cov(X, Y) = 0$，则要么 X 与 Y 独立，要么 X 与 Y 有某种非线性关系 (见例题 4.2.1)。

$y = f(x) = \underline{a_0} \quad \underbrace{+ a_2x^2 + \cdots}$

例 4.2.1

考虑 $X \sim N(0,1)$ 和 $Y = X^2$。
可见 X 与 Y 有非线性关系。
求 $Cov(X, Y)$?



$$Cov(X,Y) = E[(X-EX)(Y-EY)]$$
$$= E[XY - X\cdot EY - EX\cdot Y + EX\, EY]$$
$$= E[XY] - EX\cdot EY$$
$$= E[XY] = \iint_{\mathbb{R}^2} x^3 \cdot f_{XY}(x,y)\,dx\,dy$$
$$= \int_{-\infty}^{+\infty} x^3 \underline{\underline{\left(\int_{-\infty}^{+\infty} f_{XY}(x,y)\,dy\right)}} dx$$

$e^{\cdots}$

$$\int_1^{+\infty} \frac{1}{x^2}\,dx$$

$$\int x^3 \cdot \frac{1}{x^2}\,dx$$

协方差的性质

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

即

$$E[XY] = E[X]E[Y] \Leftrightarrow Cov(X,Y) = 0$$

协方差的性质

独立与相关性: X 与 Y 相互独立，那么 X 与 Y 一定不相关，反之不成立。

不是常数，不依赖于X.

$Y = f(X) = a_0 + a_1 X^1 + a_2 X^2 + \cdots$

独立 $\Leftrightarrow$ $P(X,Y) = P(X) \cdot P(Y)$ $\Leftrightarrow$ $a_i = 0 \quad \forall i = 1, 2, \ldots$

相关性为零 $\Leftrightarrow$ $a_1 = 0$.

$Y = X^2$.

协方差的性质

对称性: 设随机变量 X 和 Y 的方差都存在，那么

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

协方差的性质

双线性性: 设随机变量 X, Y 和 Z 的方差都存在, a 和 b 为常数, 那么

$$Cov(aX + bY, Z) = aCov(X, Z) + bCov(Y, Z)$$

以及

$$Cov(Z, aX + bY) = aCov(Z, X) + bCov(Z, Y)$$

方差的运算性质

性质 1: 任意常数 c 的方差为 0

$$Var[c] = E[(c - E[c])^2] = 0$$

性质 2: 若随机变量 X 与 Y 相互独立 或者 不相关， 并且 X 与 Y 的方差都存在， a， b 为常数，那么

$$Var[aX + bY] = a^2 Var[X] + b^2 Var[Y]$$

对于任意的随机向量 (X, Y)

$$Var(aX \pm bY) = a^2 Var(X) + b^2 Var(Y) \pm 2abCov(X, Y)$$

均值以及方差性质的运用

若 X_1 与 X_2 相互独立，且 $X_1, X_2 \sim B(1, p)$，则 $X_1 + X_2 \sim B(2, p)$。

进一步的，若 $X_1 \sim B(m, p)$，$X_2 \sim B(n, p)$，则 $X_1 + X_2 \sim B(m + n, p)$。

均值以及方差性质的运用

若 $X \sim B(n,p)$，$X_i \sim B(1,p)$，$i = 1, 2, \cdots, n$，相互独立，那么

$$X = X_1 + X_2 + \cdots + X_n$$

根据期望的线性性，

$$E[X] = E[X_1] + E[X_2] + \cdots + E[X_n]$$

$E[X_i] = p$，故

$$E[X] = np$$

再由方差的性质

$$\begin{aligned} Var[X] &= Var[X_1 + X_2 + \cdots + X_n] \\ &= Var[X_1] + Var[X_2] + \cdots + Var[X_n] \\ &= np(1-p) \end{aligned}$$

数字特征的性质

设 (X, Y) 为二维随机向量，且 $Var[X] < \infty$，$Var[Y] < \infty$。
若 X 与 Y 独立，则 $r(X, Y) = 0$，即 X 与 Y 不相关。

相关性

设 (X, Y) 为二维随机向量，且 $Var[X] < \infty$，$Var[Y] < \infty$。则

$$r(X, Y) \leq 1$$

证明：考虑下面这三个随机变量

$$\bar{X} = \frac{X - E[X]}{\sqrt{Var[X]}}, \quad \bar{Y} = \frac{Y - E[Y]}{\sqrt{Var[Y]}}, \quad Z = \bar{X} \pm \bar{Y}$$

相关性

首先

$$Var[aX] = a^2 Var[X]$$

那么

$$Var[\bar{X}] = Var[\bar{Y}] = 1$$

其次

$$Cov(aX, bY) = E[abXY] - E[aX]E[bY] = abCov(X, Y)$$

计算下

$$\begin{aligned} Var[\bar{X} \pm \bar{Y}] &= Var[\bar{X}] + Var[\bar{Y}] \pm 2Cov(\bar{X}, \bar{Y}) \\ &= 2 \pm 2r(X, Y) \geq 0 \end{aligned}$$

所以 $r(X, Y) \leq 1$。

相关性

正相关: 若 $r(X, Y) = 1$, 则存在常数 $a > 0$ 和 b, 使得

$$P(Y = aX + b) = 1$$

负相关: 若 $r(X, Y) = -1$, 则存在常数 $a < 0$ 和 b, 使得

$$P(Y = aX + b) = 1$$

相关性

只讨论第一种情况，第二种类似讨论
若 $r(X, Y) = 1$ 时，根据上面第二条，

$$Var[\bar{Y} - \bar{X}] = 0,$$

$Var[\bar{Y}-\bar{X}] = Var[\bar{Y}] + Var[\bar{X}]$
$-2\cdot Cov(\bar{Y},\bar{X})$
$= 2-2\cdot r(\bar{Y},\bar{X})$
$= 2-2=0$

即

$$\bar{Y} - \bar{X} = c, \quad a.e.$$

其中 c 是一个常数。
变换上式得

$$Y = \sqrt{\frac{Var[Y]}{Var[X]}}X + c\sqrt{Var[Y]} - \sqrt{\frac{Var[Y]}{Var[X]}}\sqrt{Var[Y]}E[X] + E[Y]$$

相关性

令

$$a = \sqrt{\frac{Var[Y]}{Var[X]}} > 0,$$

$$b = c\sqrt{Var[Y]} - \sqrt{\frac{Var[Y]}{Var[X]}}\sqrt{E[X]} + E[Y]$$

相关系数只能反映 X 与 Y 的线性相关程度，而不能刻画 X 与 Y 的非线性关系。

二维正态分布的相关性

对于 $(X, Y) \sim N(\mu_1, \mu_2, \sigma_1^2, \sigma_2^2, \rho)$，$r(X, Y) = \rho$。
对于二维联合正态分布，

$$X, Y \text{ 独立} \Leftrightarrow \rho = 0 \Leftrightarrow X, Y \text{ 不相关}$$

即 X 与 Y 独立与 X 与 Y 不相关是等价的，这个性质是联合正态分布特有的。

例 4.2.2

设 (X, Y) 的联合分布密度函数为

$$f_{X,Y}(x,y) = \begin{cases} 8xy, & 0 \le y \le x, 0 \le x \le 1 \\ 0, & \text{其他} \end{cases}$$

求 $E[X]$, $E[Y]$, $Var[X]$, $Var[Y]$, $Cov(X, Y)$ 和 $r(X, Y)$

多维随机变量的数字特征

定义 4.2.2: 设随机向量 $X = (X_1, X_2, \cdots, X_n)'$ 的每个分量都有有限的方差，定义

$$E[\mathbb{X}] = (E[X_1], E[X_2], \cdots, E[X_n])'$$

和

$$Var[\mathbb{X}] = E[(\mathbb{X} - E[\mathbb{X}])(\mathbb{X} - E[\mathbb{X}])'] = (Cov(X_i, X_j))_{n \times n}$$

条件数学期望

条件数学期望是在条件分布的数学期望。

- 离散型：设 (X, Y) 是二维离散型随机向量，有有限的数学期望，在 $\{Y = b_j\}$ 发生的条件下，X 的 条件数学期望（条件期望），就是对条件分布列

  $$P(X = a_i Y = b_j), \ i = 1, 2, \cdots$$

  求数学期望，即

  $$E[XY = b_j] = \sum_{i=1}^{\infty} a_i P(X = a_i Y = b_j)$$

  同理，在 $X = a_i$ 发生的条件下，Y 的数学期望，就是在条件分布列

  $$P(Y = b_j X = a_i), \ j = 1, 2, \cdots$$

  下求 Y 的数学期望

  $$E[YX = a_i] = \sum^{\infty} b_j P(Y = b_j X = a_i)$$

条件数学期望

- 连续型: 若 (X, Y) 是连续型的随机向量，并且有有限的数学期望。在 $\{Y = y\}$ 发生的情况下，X 的条件数学期望，就是基于条件分布密度函数 $f_{X|Y}(x|y)$ 求 X 数学期望

$$E[X|Y = y] = \int_{-\infty}^{\infty} x f_{X|Y}(x|y)\,dx$$

在 $\{X = x\}$ 发生的情况下，Y 的条件数学期望，就是在条件分布密度函数 $f_{Y|X}(y|x)$ 下求 Y 的数学期望

$$E[Y|X = x] = \int_{-\infty}^{\infty} y f_{Y|X}(y|x)\,dy$$

条件期望的性质

- 条件期望是条件的函数。比如 $E[X|Y=y]$ 就是 y 的函数。
- 当 X 和 Y 相互独立的时候，由于

$$f_{X,Y}(x,y)=f_X(x)f_Y(y)$$

于是 $f_{X|Y}(x|y)=f_X(x)$，于是

$$E[X|Y]=E[X]$$

同理

$$E[Y|X]=E[Y]$$

- 重期望公式

$$E[E[X|Y]]=E[X],\ E[E[Y|X]]=E[Y]$$

重期望公式

- 若 Y 是离散型随机变量，则

$$E[X] = \sum_i E[X|Y = y_i] P(Y = y_i)$$

- 若 Y 是连续型随机变量，则

$$E[X] = \int_{\mathbb{R}} E[X|Y = y] f_Y(y) dy$$

条件期望的运用

假设 $E[X_i]=\mu$，$i=1,\cdots,N$，N 是随机变量，且与 X_i 都独立，求 $E[\sum_{k=1}^{N}X_k]$。

解:

$$
\begin{aligned}
E[\sum_{k=1}^{N}X_k] &= E[E[\sum_{k=1}^{N}X_k|N]] = \sum_{n=1}^{\infty}E[\sum_{k=1}^{N}X_k|N=n]\cdot P(N=n) \\
&= \sum_{n=1}^{\infty}E[\sum_{k=1}^{n}X_k]\cdot P(N=n) \\
&= \sum_{n=1}^{\infty}n\mu\cdot P(N=n) = \mu\sum_{n=1}^{\infty}nP(N=n) = \mu E[N]
\end{aligned}
$$

条件期望的运用

若 (X, Y) 服从二维正态分布
$(X, Y) \sim N(\mu_1, \mu_2, \sigma_1^2, \sigma_2^2, \rho)$

$$E_{X|Y}(x|y) = \mu_1 + \rho\frac{\sigma_1}{\sigma_2}(y - \mu_2)$$

$$E_{Y|X}(y|x) = \mu_2 + \rho\frac{\sigma_2}{\sigma_1}(x - \mu_1)$$

即 $E_{X|Y}$ 和 $E_{Y|X}$ 分别是 Y 和 X 的线性函数，这是正态分布的特征。