河海大学

Ch4 随机变量的数字特征

• 数学期望(*Expectation*)

一、加权平均数

例 设某班40名学生的概率统计成绩及得分人数如下表所示：

分数	40	60	70	80	90	100
人数	1	6	9	15	7	2

则学生的平均成绩是总分÷总人数(分)。即

$$\frac{40\times1+60\times6+70\times9+80\times15+90\times7+100\times2}{1+6+9+15+7+2}=76.5(\text{分})$$

上式也可以写成：

$$\frac{1}{40}\times 40+\frac{6}{40}\times 60+\frac{9}{40}\times 70+\frac{15}{40}\times 80+\frac{7}{40}\times 90+\frac{2}{40}\times 100$$

$$=76.5(\text{分})$$

此即为40,60,70,80,90和100六个数的加权平均数。

一般地 $x_1, x_2, \cdots, x_n$的加权算术平均数为：$\sum_{i=1}^{n}\omega_i x_i$，

其中 $\omega_i \geq 0, \sum_{i=1}^{n}\omega_i = 1$, ω_i称为 x_i的权重。

引进 *r.v.* X表示学生得分，则X有分布律

X	40	60	70	80	90	100
P	1/40	6/40	9/40	15/40	7/40	2/40

于是上述平均数可以写成

$$\frac{1}{40}\times 40+\frac{6}{40}\times 60+\frac{9}{40}\times 70+\frac{15}{40}\times 80+\frac{7}{40}\times 90+\frac{2}{40}\times 100$$

$$=40\times P\{X=40\}+60\times P\{X=60\}+70\times P\{X=70\}+$$

$$+80\times P\{X=80\}+90\times P\{X=90\}+100\times P\{X=100\}$$

即取值乘取值的概率相加即得平均值。

这就是 *r.v.*的**数学期望**的概念

二、离散型随机变量的数学期望

定义： 离散型随机变量X，其分布律为：

$$P\{X = x_k\} = p_k, k = 1,2,\cdots\cdots$$

若级数$\sum_{k=1}^{\infty} x_k p_k$绝对收敛，则称级数$\sum_{k=1}^{\infty} x_k p_k$的和为随机变量$X$的**数学期望**，记为$E(X)$

即 $$E(X) = \sum_{k=1}^{\infty} x_k p_k$$

若级数$\sum_{k=1}^{\infty} |x_k| p_k$发散，则说$X$的数学期望不存在。

随机变量X的数学期望是其一切取值的平均值。

例 r.v. X的分布律为：

X	10	30	50	70	90
p	3/6	2/6	1/36	3/36	2/36

求 $E(X)$。

解 $E(X)=\sum x_k p_k =$

$$10\times\frac{3}{6}+30\times\frac{2}{6}+50\times\frac{1}{36}+70\times\frac{3}{36}+90\times\frac{2}{36}=27.22$$

例 单点分布（退化分布）

即 *r.v.X*的分布律为：$P\{X=c\}=1$

$\therefore E(X)=c\times 1=c$

即常数的数学期望为常数。$E(c)=c$

例 $X\sim$（0—1）分布

即 *r.v.X*的分布律为：

X	0	1
p	$1-p$	p

$E(X)=p$

$\therefore E(X)=0\times(1-p)+1\times p=p$

例 $X \sim B(n, p)$二项分布 $E(X) = np$

即 r.v.X的分布律为：

$$P\{X = k\} = C_n^k p^k (1-p)^{n-k}, k = 0,1,2,\cdots,n$$

例 $X \sim P(\lambda)$ Poisson分布

即 r.v.X的分布律为： $E(X) = \lambda$

$$P\{X = k\} = \frac{\lambda^k e^{-\lambda}}{k!}, k = 0,1,2,\cdots,(\lambda > 0)$$

例 $r.v.X$的取值为： $x_k=\dfrac{(-1)^k 2^k}{k}, k=1,2,\cdots$

对应的概率为： $p_k=\dfrac{1}{2^k}$

$$E(X)=\sum_{k=1}^{\infty} x_k p_k=\sum_{k=1}^{+\infty}\frac{(-1)^k 2^k}{k}\cdot\frac{1}{2^k}=\sum_{k=1}^{\infty}(-1)^k\frac{1}{k}=-\ln 2$$

但 $\sum_{k=1}^{\infty}|x_k|p_k=\sum_{k=1}^{\infty}\frac{1}{k}=\infty$ 所以$E(X)$不存在！

三、连续型随机变量的数学期望

定义 设连续型 $r.v.X$的概率密度函数为 $f(x)$,

若积分$\int_{-\infty}^{+\infty} xf(x)dx$绝对收敛，则称积分$\int_{-\infty}^{+\infty} xf(x)dx$

的值为连续型 $r.v.X$的数学期望，记为$E(X)$。

即
$$E(X)=\int_{-\infty}^{+\infty} xf(x)dx$$

若积分$\int_{-\infty}^{+\infty}|x|f(x)dx$ 发散时, X的数学期望不存在.

例 r.v.X的概率密度函数为：

$$f(x)=\begin{cases}1-e^{-x}, & 0\le x\le 1\\ (e-1)e^{-x}, & x\ge 1\\ 0, & others\end{cases}$$

求$E(X)$.

解 $$E(X)=\int_{-\infty}^{+\infty}xf(x)dx=\int_0^1 x(1-e^{-x})dx+\int_1^{+\infty}x(e-1)e^{-x}dx$$

$$=\int_0^1 xdx-\int_0^1 xe^{-x}dx+(e-1)\int_1^{+\infty}xe^{-x}dx=\frac{3}{2}$$

例 $X\sim U(a, b)$均匀分布

其概率密度函数为：

$$E(X)=\frac{a+b}{2}$$

$$f(x)=\begin{cases}\dfrac{1}{b-a}, & a<x<b\\ 0, & others\end{cases}$$

求$E(X)$.

解 $$E(X)=\int_{-\infty}^{+\infty}xf(x)dx=\int_a^b x\frac{1}{b-a}dx=\frac{a+b}{2}$$

例 指数分布$E(\lambda)$

X服从参数为λ的指数分布，其概率密度函数为：

$$f(x)=\begin{cases}\lambda e^{-\lambda x}, & x>0\\ 0, & others\end{cases}\quad(\lambda>0)\qquad E(X)=\frac{1}{\lambda}$$

求$E(X)$.

解 $$E(X)=\int_{-\infty}^{+\infty}xf(x)dx=\int_{0}^{+\infty}x\lambda e^{-\lambda x}dx=\frac{1}{\lambda}$$

例 正态分布 $X \sim N(\mu, \sigma^2)$

其概率密度函数为：　　　　$E(X)=\mu$

$$f(x)=\frac{1}{\sqrt{2\pi}\sigma}e^{-\frac{(x-\mu)^2}{2\sigma^2}},\quad \sigma>0, -\infty<x<+\infty$$

求$E(X)$.

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

$$\overset{t\overset{\Delta}{=}\frac{x-\mu}{\sigma}}{=}\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{+\infty}(\sigma t+\mu)e^{-\frac{t^2}{2}}dt=\frac{\sigma}{\sqrt{2\pi}}\int_{-\infty}^{+\infty}te^{-\frac{t^2}{2}}dt+\frac{\mu}{\sqrt{2\pi}}\int_{-\infty}^{+\infty}e^{-\frac{t^2}{2}}dt$$

$=\mu$

例 $r.v.X$的概率密度函数为：($Cauthy$分布)

$$f(x)=\frac{1}{\pi}\cdot\frac{1}{1+x^2},\qquad -\infty<x<+\infty$$

由于 $\displaystyle\int_{-\infty}^{+\infty}|x|f(x)dx=\int_{-\infty}^{+\infty}|x|\cdot\frac{1}{\pi}\cdot\frac{1}{1+x^2}dx$

$$=\int_{-\infty}^{0}-x\cdot\frac{1}{\pi}\cdot\frac{1}{1+x^2}dx+\int_{0}^{+\infty}x\cdot\frac{1}{\pi}\cdot\frac{1}{1+x^2}dx=\infty$$

所以$E(X)$不存在！

四、对于 $r.v.X$ 的函数的数学期望

- Y 为 $r.v.X$ 的函数，$Y=g(X)$，g 为连续函数

（i）X 是离散型随机变量，其分布律为

$$p_k = P\{X = x_k\}, \qquad k = 1,2,\cdots$$

若 $\sum_{k=1}^{\infty} g(x_k)p_k$ 绝对收敛，则有

$$E(Y) = E(g(X)) = \sum_{k=1}^{\infty} g(x_k)\cdot p_k$$

（ii）X是连续型随机变量，其概率密度函数为 $f(x)$

若 $\int_{-\infty}^{+\infty} g(x) f(x) dx$ 绝对收敛，则有

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

综上有：

若已知 X 的分布以及函数 $g(x)$,可以不必求出 $Y= g(x)$的分布,直接利用上面的公式求出Y的数学期望.

例 r.v.X的分布律为：

X	-2	0	5
p	0.4	0.1	0.5

求$E(2X^2+X-2)$ 和$E[X-E(X)]^2$.

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

$=[2\times(-2)^2+(-2)-2]\times 0.4+(2\times 0^2+0-2)\times 0.1+$

$+(2\times 5^2+5-2)\times 0.5=27.8$

例 $r.v.X$的概率密度函数为：

$$f(x)=\begin{cases} e^{-x}, & x>0 \\ 0, & x\le 0 \end{cases}$$

求$E(3X^2-7X+8)$和 $E[X-E(X)]^2$.

解 $E(3X^2-7X+8)=\int_{-\infty}^{+\infty}(3x^2-7x+8)\cdot f(x)dx$

$$=\int_0^{+\infty}(3x^2-7x+8)\cdot e^{-x}dx=3$$

o 二维 *r.v.*(X, Y), $Z=g(X, Y)$, g为连续函数

（i）(X, Y)是离散型随机变量，其分布律为

$$p_{ij} = P\{X = x_i, Y = y_j\}, \qquad i, j = 1, 2, \cdots$$

若 $\sum_{i=1}^{\infty}\sum_{j=1}^{\infty} g(x_i, y_j)\cdot p_{ij}$ 绝对收敛，则有

$$E(Z) = E(g(X, Y)) = \sum_{k=1}^{\infty}\sum_{j=1}^{\infty} g(x_i, y_j)\cdot p_{ij}$$

（ii）(X, Y)是连续型随机变量，其概密为 $f(x, y)$

若$\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} g(x,y)\cdot f(x,y)dxdy$ 绝对收敛，则有

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

例 *r.v.* (X, Y)的概率密度函数为：

$$f(x,y)=\begin{cases}12y^2, & 0\le y\le x\le 1\\ 0, & others\end{cases}$$

求$E(XY)$，$E(X^2+Y^2)$。

例 在长为 a 的线段上任取两点，求两点间距离的数学期望。

五、数学期望的性质

1. $E(c)=c$, c是常数；

2. $E(cX)=c\,E(X)$, c是常数；

3. $E(aX+b)=a\,E(X)+b$, a, b为常数；

4. $E(X+Y)=E(X)+E(Y)$;

$$E(X_1+X_2+\cdots+X_n)=E(X_1)+E(X_2)+\cdots+E(X_n);$$

$$E(c_1X_1+\cdots+c_nX_n)=c_1E(X_1)+\cdots+c_nE(X_n);$$

5. 若X与Y独立，则$E(XY)=E(X)E(Y)$.

若$X_1, X_2, \cdots, X_n$ 相互独立，则

$$E(X_1X_2\cdots X_n)=E(X_1)E(X_2)\cdots E(X_n)。$$

例 一个人有 n 把钥匙，其中只有一把能打开房门，他任取一把尝试，试过的钥匙拿开不再用，直到把门打开为止。试求平均试开次数。

- **方差(*Variance or Dispersion*)**

方差是衡量随机变量取值与其均值的偏离程度的一个数字特征。

1.定义 若$E(X)$存在，则称$E[X-E(X)]^2$为 r.v. X的方差，记为$D(X)$,或Var(X).

可见
$$D(X)=\begin{cases}\sum\limits_{k=1}^{\infty}[x_k-E(X)]^2P\{X=x_k\}, & \text{离散型情况}\\ \int_{-\infty}^{\infty}[x-E(X)]^2 f(x)dx, & \text{连续型情况}\end{cases}$$

2.推论 $D(X)=E(X^2)-[E(X)]^2$.

例 单点分布（退化分布）

即 $r.v.X$的分布律为：$P\{X=c\}=1$　$E(c)=c$

即常数的方差为零。$D(c)=0$

例 $X\sim$（0—1）分布

即 $r.v.X$的分布律为：

X	0	1
p	$1-p$	p

$E(X)=p$

$D(X)=p(1-p)$

例 $X \sim B(n, p)$二项分布

即 r.v.X的分布律为：

$$P\{X=k\}=C_n^k p^k (1-p)^{n-k}, k=0,1,2,\cdots,n$$

$E(X)=np$　　$D(X)=np(1-p)$

例 $X \sim P(\lambda)$ Poisson分布

即 $r.v.X$的分布律为：

$$P\{X=k\}=\frac{\lambda^k e^{-\lambda}}{k!},k=0,1,2,\cdots,(\lambda>0)$$

$E(X)=\lambda$ $D(X)=\lambda$

例 $X \sim U(a, b)$均匀分布

其概率密度函数为：

$$f(x)=\begin{cases}\dfrac{1}{b-a}, & a<x<b\\ 0, & others\end{cases}$$

$$E(X)=\frac{a+b}{2} \qquad D(X)=\frac{(b-a)^2}{12}$$

例 指数分布$E(\lambda)$

X服从参数为λ的指数分布，其概率密度函数为：

$$f(x)=\begin{cases}\lambda e^{-\lambda x}, & x>0\\ 0, & others\end{cases}\quad(\lambda>0)$$

$$E(X)=\frac{1}{\lambda}\qquad D(X)=\frac{1}{\lambda^2}$$

例 正态分布 $X \sim N(\mu, \sigma^2)$

其概率密度函数为：

$$f(x) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(x-\mu)^2}{2\sigma^2}}, \quad \sigma > 0, -\infty < x < +\infty$$

$$E(X) = \mu \qquad D(X) = \sigma^2$$

常见分布的期望和方差

分布名称	分 布 律	期 望	方 差
0-1分布	$\begin{pmatrix} 0 & 1 \\ 1-p & p \end{pmatrix}$	p	$p(1-p)$
二项分布 $B(n,\ p)$	$P\{X=k\}=C_n^k p^k (1-p)^{n-k}$ $k=0,1,2,\cdots,n$	np	$np(1-p)$
Poisson 分布 $P(\lambda)$	$P\{X=k\}=\dfrac{\lambda^k e^{-\lambda}}{k!}$ $k=0,1,2,\cdots$	λ	λ

分布名称	概率密度函数	期　望	方　差
均匀分布 U(a，b)	$f(x)=\begin{cases}\dfrac{1}{b-a}, & a<x<b\\ 0, & others\end{cases}$	$\dfrac{a+b}{2}$	$\dfrac{(b-a)^2}{12}$
指数分布 $E(\lambda)$	$f(x)=\begin{cases}\lambda e^{-\lambda x}, & x>0\\ 0, & others\end{cases}$	$\dfrac{1}{\lambda}$	$\dfrac{1}{\lambda^2}$
正态分布 $N(\mu,\sigma^2)$	$f(x)=\dfrac{1}{\sqrt{2\pi}\sigma}e^{-\frac{(x-\mu)^2}{2\sigma^2}}$	μ	σ^2

3. 方差的性质

(1) $D(C)=0$，C为常数；

(2) $D(cX)=c^2D(X)$，c为常数；

(3) 若 X，Y 独立，则 $D(X+Y)=D(X)+D(Y)$;

若 X，Y 独立，则 $D(X-Y)=D(X)+D(Y)$;

若$X_1, X_2, \cdots, X_n$ 相互独立，则

$$D(X_1+X_2+\cdots+X_n)=D(X_1)+D(X_2)+\cdots+D(X_n);$$

$$D(c_1X_1+\cdots+c_nX_n)=c_1^2D(X_1)+\cdots+c_n^2D(X_n);$$

(4) $D(X)=0 \Leftrightarrow$

存在常数C，使 $P\{X=C\}=1$，且$C=E(X)$。

X 记 n 重贝努里试验中A发生的次数, 则 $X \sim B(n, p)$

若记 $X_i = \begin{cases} 1, & \text{若在第 } i \text{ 次试验中 } A \text{ 发生;} \\ 0, & \text{若在第 } i \text{ 次试验中 } A \text{ 不发生。} \end{cases} \quad i = 1, \cdots, n$

$X_i \sim \begin{array}{c|cc} X & 0 & 1 \\ \hline P & 1-p & p \end{array}$ 且 $X_1, X_2, \cdots, X_n$ 相互独立

$\therefore E(X_i)=p$, $D(X_i)=p(1-p)$ $\quad i = 1, 2, \cdots, n$

又 $X=X_1+X_2+\cdots+X_n$,

$\therefore E(X)=E(X_1)+E(X_2)+\cdots+E(X_n)=np$,

$D(X)=D(X_1)+D(X_2)+\cdots+D(X_n)=np(1-p)$.

$\sqrt{D(X)}$ 称为X 的**均方差**或**标准差**，其量纲与X的一致。

若 $r.v.\ X$ 满足：$E(X)=\mu$，$D(X)=\sigma^2$

称 $X^* = \dfrac{X-E(X)}{\sqrt{D(X)}} = \dfrac{X-\mu}{\sigma}$ 为 X 的标准化.

易知 $E(X^*)=0$，$D(X^*)=1$.

称 $\delta_X = \dfrac{\sqrt{D(X)}}{E(X)}$ 为 X 的变异系数.

- **切比雪夫不等式**

若 $r.v.X$的期望和方差存在，则对任意$\varepsilon > 0$，有

$$P\{|X - E(X)| < \varepsilon\} \geq 1 - \frac{D(X)}{\varepsilon^2}.$$

这就是著名的切比雪夫(Chebyshev)不等式。

它有以下几种等价的形式：

$$P\{|X - E(X)| \geq \varepsilon\} \leq \frac{D(X)}{\varepsilon^2};$$

记 $\sigma^2 = D(X)$, $\mu = E(X)$, 则对$a > 0$，有

$$P\{|X - \mu| \geq a\sigma\} \leq \frac{1}{a^2} \text{或} P\{|X - \mu| < a\sigma\} \geq 1 - \frac{1}{a^2}.$$

定理 (*Cauchy-Schwarz*不等式)

若对任意的 *r.v.* X、Y，若$E(X^2)<+\infty$， $E(Y^2)<+\infty$，则

$$[E(XY)]^2 \le E(X^2)\cdot E(Y^2)$$

当且仅当$P\{Y=aX\}=1$时等号成立, 其中a为某常数.

- 协方差，相关系数

1.协方差定义(Co-variance)

若 r.v. X的期望 $E(X)$和Y 的期望 $E(Y)$存在, 则称 $E\{[X-E(X)][Y-E(Y)]\}$为X与Y的协方差,记为 $\mathrm{Cov}(X, Y)$.

即 $\mathrm{Cov}(X, Y)=E\{[X-E(X)][Y-E(Y)]\}$.

易有 $\mathrm{Cov}(X, Y)=E(XY)-E(X)E(Y)$。

$$\mathrm{Cov}(X, Y) = E\{[X-E(X)][Y-E(Y)]\}$$

2.协方差性质

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

(1) 若X, Y相互独立，则$\mathrm{Cov}(X, Y) = 0$；

(2) $\mathrm{Cov}(X, Y) = \mathrm{Cov}(Y, X)$;

(3) $\mathrm{Cov}(aX, bY) = ab\mathrm{Cov}(X, Y)$, 其中$a, b$为 常数；

(4) $\mathrm{Cov}(X+Y, Z) = \mathrm{Cov}(X, Z) + \mathrm{Cov}(Y, Z)$;

$$\mathrm{Cov}(\sum a_i X_i, \sum b_j Y_j) = \sum\sum a_i b_j \mathrm{Cov}(X_i, Y_j);$$

(5) $D(X \pm Y) = D(X) + D(Y) \pm 2\mathrm{Cov}(X, Y)$.

3. 相关系数

定义 若 *r.v.* X、Y的方差和协方差均存在，且 $D(X)>0, D(Y)>0$，则

$$\rho=\rho_{XY}=\frac{Cov(X,Y)}{\sqrt{DX}\sqrt{DY}}$$

称为X与Y的**相关系数**.

相关系数的性质 $|\rho_{XY}|\le 1$；　$\rho_{XY}=\dfrac{Cov(X,Y)}{\sqrt{DX}\sqrt{DY}}$

特别

1 若$\rho_{XY}=0$，则称 X与Y不相关，否则称 X与Y相关.

(1) X与Y不相关$\Leftrightarrow \rho_{XY}=0 \Leftrightarrow \mathrm{Cov}(X,Y)=0$

$\Leftrightarrow E(XY)=E(X)E(Y) \Leftrightarrow D(X\pm Y)=D(X)+D(Y)$。

(2) X与Y独立，则X与Y不相关，反之不然。

即　独立 $\underset{\overset{?}{\Leftarrow}}{\Rightarrow}$ 不相关

2 $|\rho_{XY}|=1 \Leftrightarrow$ 存在常数a, b 使$P\{X=aY+b\}=1$。

(1) 若$\rho_{XY}=1$, 称 X与Y 完全正相关。

即 存在常数$a>0, b$ 使$P\{X=aY+b\}=1$;

(2) 若$\rho_{XY}=-1$, 称 X与Y 完全负相关。

即 存在常数$a<0, b$ 使$P\{X=aY+b\}=1$。

例 设(X, Y)在 $D=\{(x, y)|\ x^2+y^2\le 1\}$上服从均匀分布，则$X$与$Y$不相关，但不是相互独立的。

例 已知二元 $r.v.(X,Y)\sim N(\mu_1,\mu_2,\sigma_1^2,\sigma_2^2,\rho)$

$$f(x,y)=\frac{1}{2\pi\sigma_1\sigma_2\sqrt{1-\rho^2}}\cdot$$

$$\cdot e^{-\frac{1}{2(1-\rho^2)}\left[\frac{(x-\mu_1)^2}{\sigma_1^2}-2\rho\frac{(x-\mu_1)(y-\mu_2)}{\sigma_1\sigma_2}+\frac{(y-\mu_2)^2}{\sigma_2^2}\right]},$$

求ρ_{XY}。

Notes 对于二元正态分布，其独立性和不相关性是等价的。**其它的分布没有这个性质**。

- 矩 (假设其中一切期望都存在)

1. k 阶原点矩

$$A_k=E(X^k),\quad k=1,\ 2,\ \ldots$$

2. k 阶中心矩

$$B_k=E[X-E(X)]^k,\quad k=1,\ 2,\ \ldots$$

易知 $E(X)=A_1$, $D(X)=B_2$.

3. $k+l$ 阶混合原点矩

$$E(X^k Y^l),\quad k, l=0, 1, 2, \ldots;$$

4. $k+l$ 阶混合中心矩

$$E\{[X-E(X)]^k [Y-E(Y)]^l\},\quad k, l=0, 1, 2, \ldots;$$

易知 $\mathrm{Cov}(X, Y)=E\{[X-E(X)][Y-E(Y)]\}$

是1+1阶混合中心矩。

• 多维随机变量的期望和协方差矩阵

数学期望

定义 设 $X_1,\cdots,X_n$ 为 n 个 r.v., 记

$$X=(X_1,\cdots,X_n)^T$$

称 $E(X)=(E(X_1),\cdots,E(X_n))^T$ 为 $X=(X_1,\cdots,X_n)^T$ 的**数学期望**。

协方差矩阵

1. 定义 设 $X_1, \cdots, X_n$为n个 r.v., 记 b_{ij}=Cov(X_i, X_j), i, j =1, 2, ⋯, n. 则称由b_{ij}组成的矩阵为 r.v. $X_1, \cdots, X_n$ 的协方差矩阵 B。即

$$B=(b_{ij})_{n\times n}=\begin{bmatrix} b_{11} & b_{12} & \cdots & b_{1n} \\ b_{21} & b_{22} & \cdots & b_{2n} \\ \cdots & \cdots & \cdots & \cdots \\ b_{n1} & b_{n2} & \cdots & b_{nn} \end{bmatrix}$$

2.协方差矩阵的性质

(1) $B=B^T$, 其中B^T为B的转置;

(2) B是非负定矩阵,

即对任意n维实向量$\alpha=(\alpha_1,\cdots,\alpha_n)^T$.有

$$\alpha^T B\alpha \ge 0.$$

例　设$(X_1,X_2)\sim N(\mu_1,\mu_2,\sigma_1^2,\sigma_2^2,\rho)$, 则

$$B=\begin{bmatrix}\sigma_1^2 & \rho\sigma_1\sigma_2\\ \rho\sigma_1\sigma_2 & \sigma_2^2\end{bmatrix},\quad |B|=\sigma_1^2\sigma_2^2(1-\rho^2),$$

$X=(X_1, X_2)^T$的概率密度为

$$f(x_1,x_2)=\frac{1}{2\pi|B|^{1/2}}\exp\{-\frac{1}{2}(x-\mu)^T B^{-1}(x-\mu)\},$$

其中$x=(x_1,x_2)^T, \mu=(\mu_1,\mu_2)^T,$

可记为 $X=(X_1,X_2)^T \sim N(\mu, B).$

一般地, 若$X=(X_1,\cdots,X_n)^T \sim N(\mu,\ B)$,则

$$f(x)=\frac{1}{(2\pi)^{n/2}|B|^{1/2}}\exp\{-\frac{1}{2}(x-\mu)^T B^{-1}(x-\mu)\},$$

其中 $x=(x_1,\cdots,x_n)^T$,

$\mu=(\mu_1,\cdots,\mu_n)^T$为$X=(X_1,\cdots,X_n)^T$的均值向量,

称X服从n维正态分布.