

# AI 中的数学

## 第二十讲

方聪，概率统计部分参考章复熹和张原老师课件

2024 年秋季

- ① 第三章作业
- ② 第二部分
- ③ 多元正态分布

- 1 第三章作业
- 2 第二部分
- 3 多元正态分布

### 3. 设二维随机向量 $(X, Y)$ 的联合密度为

$$p(x, y) = \begin{cases} c(R - \sqrt{x^2 + y^2}), & x^2 + y^2 \leq R^2, \\ 0, & x^2 + y^2 > R^2, \end{cases}$$

求：(1) 系数  $c$ ；(2) 向量  $(X, Y)$  落入圆  $x^2 + y^2 \leq r^2 (0 < r < R)$  的概率。

### 3. 设二维随机向量 $(X, Y)$ 的联合密度为

$$p(x, y) = \begin{cases} c(R - \sqrt{x^2 + y^2}), & x^2 + y^2 \leq R^2, \\ 0, & x^2 + y^2 > R^2, \end{cases}$$

求：(1) 系数  $c$ ；(2) 向量  $(X, Y)$  落入圆  $x^2 + y^2 \leq r^2 (0 < r < R)$  的概率。

解：(1)

$$\begin{aligned} \int_{\mathbb{R}^2} p(x, y) dx dy &= \int_{\{x^2 + y^2 \leq R^2\}} c(R - \sqrt{x^2 + y^2}) dx dy \\ &= \int_{0 \leq r \leq R} \int_{0 \leq \theta \leq 2\pi} c(R - r) r d\theta dr \\ &= 2\pi c \int_0^R (R - r) r dr = c\pi R^2 = 1 \end{aligned}$$

因此  $c = \frac{1}{\pi R^2}$ .

(2)

$$\begin{aligned} P(x^2 + y^2 \leq r^2) &= \int_{\{x^2 + y^2 \leq r^2\}} c(R - \sqrt{x^2 + y^2}) dx dy \\ &= \int_{0 \leq r' \leq r} \int_{0 \leq \theta \leq 2\pi} c(R - r') r' d\theta dr' \\ &= 2\pi c \int_0^r (R - r') r' dr' \\ &= 2\pi c \left( Rr - \frac{r^2}{2} \right) = \frac{2r}{R} - \frac{r^2}{R^2}. \end{aligned}$$

4. 设二维随机向量  $(X, Y)$  服从区域

$$D = \left\{ (x, y) : \frac{(x+y)^2}{2a^2} + \frac{(x-y)^2}{2b^2} \leq 1 \right\} \quad (a, b > 0)$$

上的均匀分布, 求  $(X, Y)$  的联合密度。

4. 设二维随机向量  $(X, Y)$  服从区域

$$D = \left\{ (x, y) : \frac{(x+y)^2}{2a^2} + \frac{(x-y)^2}{2b^2} \leq 1 \right\} \quad (a, b > 0)$$

上的均匀分布, 求  $(X, Y)$  的联合密度。

解: 设  $I_D(x, y)$  为示性函数,  $p(x, y) = cI_D(x, y)$ 。令

$u = \frac{x+y}{2}, v = \frac{x-y}{2}$  得

$$\int_{\mathbb{R}^2} p(x, y) dx dy = \int_D c dx dy = c \int_{\left\{ \frac{2u^2}{a^2} + \frac{2v^2}{b^2} \leq 1 \right\}} 2 du dv = c\pi ab = 1$$

解得  $c = \frac{1}{\pi ab}$ , 联合密度为  $p(x, y) = \frac{1}{\pi ab} I_D(x, y)$ 。



8. 设随机变量  $X$  与  $Y$  相互独立。分别服从自由度为  $m, n$  的  $\chi^2$  分布, 即

$$p_X(x) = \begin{cases} \frac{1}{2^{m/2}\Gamma(\frac{m}{2})} x^{m/2-1} e^{-x/2}, & x > 0, \\ 0, & x \leq 0, \end{cases}$$

$$p_Y(y) = \begin{cases} \frac{1}{2^{n/2}\Gamma(\frac{n}{2})} y^{n/2-1} e^{-y/2}, & y > 0, \\ 0, & y \leq 0, \end{cases}$$

试证明,  $X + Y$  服从  $\chi^2$  分布, 其自由度为  $m + n$ .

有  $x \sim \Gamma(m, 1/2)$ ,  $y \sim \Gamma(n, 1/2)$ . 由独立性, 可得结论。

若  $X$  与  $Y$  独立,  $X \sim \Gamma(r, \lambda)$ ,  $Y \sim \Gamma(s, \lambda)$ . 则

$$X + Y \sim \Gamma(r + s, \lambda).$$

- 密度:  $p_X(x) = \frac{\lambda^r}{\Gamma(r)} x^{r-1} e^{-\lambda x}, \quad x > 0.$
- $Z = X + Y : p_Z(z) = \int p_X(x) p_Y(z-x) dx. \forall z > 0,$

$$\begin{aligned}
 p_Z(z) &= C \int_0^z x^{r-1} e^{-\lambda x} \cdot (z-x)^{s-1} e^{-\lambda(z-x)} dx \\
 &= C e^{-\lambda z} \int_0^1 (tz)^{r-1} ((1-t)z)^{s-1} d(tz) = \hat{C} z^{r+s-1} e^{-\lambda z}.
 \end{aligned}$$

10. 设二维随机向量  $(X, Y)$  的联合密度为

$$p(x, y) = \begin{cases} 4xy \exp\{-(x^2 + y^2)\}, & x > 0, y > 0, \\ 0, & \text{其他,} \end{cases}$$

求  $Z = \sqrt{X^2 + Y^2}$  的均值。

10. 设二维随机向量  $(X, Y)$  的联合密度为

$$p(x, y) = \begin{cases} 4xy \exp\{-(x^2 + y^2)\}, & x > 0, y > 0, \\ 0, & \text{其他}, \end{cases}$$

求  $Z = \sqrt{X^2 + Y^2}$  的均值。

解：作极坐标变换  $x = r \cos \theta, y = r \sin \theta$  得

$$\begin{aligned} E(Z) &= \int_0^\infty \int_0^\infty \sqrt{x^2 + y^2} 4xy \exp\{-(x^2 + y^2)\} dx dy \\ &= \int_0^\infty \int_0^{\frac{\pi}{2}} r \cdot 4r^2 \cos \theta \sin \theta e^{-r^2} r d\theta dr \\ &= \left( \int_0^\infty 4r^4 e^{-r^2} dr \right) \left( \int_0^{\frac{\pi}{2}} \cos \theta \sin \theta d\theta \right) \\ &= \frac{3\sqrt{\pi}}{2} \cdot \frac{1}{2} = \frac{3\sqrt{\pi}}{4}. \end{aligned}$$

11. 设二维随机向量  $(X, Y)$  服从区域

$$D = \{(x, y) : 0 < x < 1, 0 < y < x\}$$

上的均匀分布, 求  $E(X)$ ,  $\text{var}(X)$  及  $X$  与  $Y$  的相关系数。

# 11. 设二维随机向量 $(X, Y)$ 服从区域

$$D = \{(x, y) : 0 < x < 1, 0 < y < x\}$$

上的均匀分布, 求  $E(X)$ ,  $\text{var}(X)$  及  $X$  与  $Y$  的相关系数。

解: 设  $I_D(x, y)$  为示性函数,  $p(x, y) = cI_D(x, y)$ 。

由于  $\int_0^1 \int_0^x c dx dy = 1$ , 解得  $c = 2$ 。

$$E(X) = \int_{\mathbb{R}^2} x \cdot 2I_D(x, y) dx dy = \int_0^1 2x^2 dx = \frac{2}{3}.$$

$$E(X^2) = \int_{\mathbb{R}^2} x^2 \cdot 2I_D(x, y) dx dy = \int_0^1 2x^3 dx = \frac{1}{2}.$$

$$\text{var}(X) = E(X^2) - (E(X))^2 = \frac{1}{18}.$$

$$E(XY) = \int_{\mathbb{R}^2} xy \cdot 2I_D(x, y) dx dy = \int_0^1 \int_0^x 2xy dy dx = \frac{1}{4}.$$

$$\rho(X, Y) = \frac{E(XY) - E(X)E(Y)}{\sqrt{\text{var}(X)}\sqrt{\text{var}(Y)}} = \frac{1}{2}.$$



**12.** 设随机变量  $X \sim N(0, 1)$ ,  $Y = X^n$  ( $n$  是正整数), 求  $X$  与  $Y$  的相关系数.

12. 设随机变量  $X \sim N(0, 1)$ ,  $Y = X^n$  ( $n$  是正整数), 求  $X$  与  $Y$  的相关系数.

解: 若随机变量  $X \sim N(0, 1)$ , 则对一切正整数  $k$ , 可以由递推得到

$$E(X^{2k-1}) = 0, \quad E(X^{2k}) = \prod_{0 \leq i < k} (2k - 1 - 2i)$$

若  $n$  为奇数,

$$E(Y) = E(X^n) = 0$$

$$E(XY) = E(X^{n+1}) = \prod_{0 \leq i < \frac{n+1}{2}} (n - 2i)$$

$$E(Y^2) = E(X^{2n}) = \prod_{0 \leq i < n} (2n - 1 - 2i)$$



13. 设随机变量  $X_1$  与  $X_2$  相互独立, 分布密度分别为

$$p_1(x) = \begin{cases} 2x, & 0 \leq x \leq 1, \\ 0, & \text{其他}, \end{cases}$$

$$p_2(x) = \begin{cases} e^{-(x-5)}, & x > 5, \\ 0, & \text{其他}. \end{cases}$$

求  $E(X_1 X_2)$ .

13. 设随机变量  $X_1$  与  $X_2$  相互独立, 分布密度分别为

$$p_1(x) = \begin{cases} 2x, & 0 \leq x \leq 1, \\ 0, & \text{其他.} \end{cases}$$

$$p_2(x) = \begin{cases} e^{-(x-5)}, & x > 5, \\ 0, & \text{其他.} \end{cases}$$

求  $E(X_1 X_2)$ .

解:

$$E(X_1) = \int_{-\infty}^{+\infty} x p_1(x) dx = \int_0^1 2x^2 dx = \frac{2}{3},$$

$$E(X_2) = \int_{-\infty}^{+\infty} x p_2(x) dx = \int_5^{\infty} x e^{-(x-5)} dx = \int_0^{\infty} (u+5) e^u du = 6$$

由于  $X_1$  和  $X_2$  相互独立, 因此

$$E(X_1 X_2) = E(X_1) E(X_2) = 4.$$

**14.** 设  $X$  和  $Y$  是随机变量,  $\text{var}(X) = 25$ ,  $\text{var}(Y) = 36$ , 相关系数  $\rho_{XY} = 0.4$ , 求  $\text{var}(X + Y)$  及  $\text{var}(X - Y)$ .

解：不妨假设  $X, Y$  零均值。

$$\text{var}(X + Y) = \text{var}(X) + \text{var}(Y) + 2\text{cov}(X, Y) = 85.$$

$$\text{var}(X - Y) = \text{var}(X) + \text{var}(Y) - 2\text{cov}(X, Y) = 37.$$

$$D = \left\{ (x, y) : \frac{x^2}{a^2} + \frac{y^2}{b^2} \leq k^2 \right\} \quad (k > 0)$$

的概率。



15. 设二维随机向量  $(X, Y)$  服从二维正态分布,  
 $E(X) = E(Y) = 0$ ,  $\text{var}(X) = a^2$ ,  $\text{var}(Y) = b^2$ ,  $\rho_{XY} = 0$ . 试求  
 $(X, Y)$  落入区域

$$D = \left\{ (x, y) : \frac{x^2}{a^2} + \frac{y^2}{b^2} \leq k^2 \right\} \quad (k > 0)$$

的概率。

解：由题意得，联合密度函数

$$f_{X,Y}(x, y) = \frac{1}{ab2\pi} e^{-\frac{x^2}{2a^2} - \frac{y^2}{2b^2}}$$

$(X, Y)$  落入区域  $D$  的概率为

$$P((X, Y) \in D) = \iint_D \frac{1}{ab2\pi} e^{-\frac{x^2}{2a^2} - \frac{y^2}{2b^2}} dx dy$$

使用极坐标变换： $x = ar \cos \theta$ ,  $y = br \sin \theta$ .

在这种变换下，雅可比行列式为：

$$J = \left| \frac{\partial(x, y)}{\partial(r, \theta)} \right| = abr$$

区域  $D$  变换后为：

$$\frac{(ar \cos \theta)^2}{a^2} + \frac{(br \sin \theta)^2}{b^2} = r^2(\cos^2 \theta + \sin^2 \theta) = r^2 \leq k^2$$

因此，积分可表示为

$$\begin{aligned} P((X, Y) \in D) &= \int_0^{2\pi} \int_0^k \frac{1}{ab2\pi} e^{-\frac{(ar \cos \theta)^2}{2a^2} - \frac{(br \sin \theta)^2}{2b^2}} \cdot abr \, dr \, d\theta \\ &= \int_0^{2\pi} \int_0^k \frac{1}{2\pi} e^{-\frac{r^2}{2}} r \, dr \, d\theta \end{aligned}$$

令  $u = \frac{r^2}{2}$ , 则

$$\begin{aligned} P((X, Y) \in D) &= \int_0^{2\pi} \int_0^{\frac{k^2}{2}} \frac{1}{2\pi} e^{-u} du d\theta \\ &= \int_0^{2\pi} \frac{1}{2\pi} (1 - e^{-\frac{k^2}{2}}) d\theta \end{aligned}$$

**16.** 设三维随机向量  $(X, Y, Z)$  的联合密度为

$$p(x, y, z) = \begin{cases} e^{-(x+y+z)}, & x > 0, y > 0, z > 0, \\ 0, & \text{其他,} \end{cases}$$

试分别求出随机变量  $X, Y, Z$  的分布密度, 又问:  $X, Y, Z$  相互独立吗?

解：

$$f_X(x) = \int_0^\infty \int_0^\infty p(x, y, z) dy dz = \int_0^\infty \int_0^\infty e^{-(x+y+z)} dy dz = \int_0^\infty e^{-x}$$

因此,  $X$  的边缘分布密度为:

$$f_X(x) = \begin{cases} e^{-x}, & x > 0, \\ 0, & \text{其他.} \end{cases}$$

$Y$  和  $Z$  相同。由

$$f_X(x) \cdot f_Y(y) \cdot f_Z(z) = e^{-x} \cdot e^{-y} \cdot e^{-z} = e^{-(x+y+z)} = p(x, y, z)$$

因此,  $X, Y, Z$  是相互独立的。

$\xi \triangleq \sqrt{X^2 + Y^2 + Z^2}$  的概率分布.



由变换公式,  $\xi$  的概率密度函数为:

$$\begin{aligned} f_{\xi}(\xi) &= f_W(g^{-1}(\xi)) \left| \frac{d}{d\xi} g^{-1}(\xi) \right| = f_W(\xi^2) \left| \frac{d}{d\xi} (\xi^2) \right| \\ &= \frac{(\xi^2)^{0.5} e^{-(\xi^2)/2}}{\sqrt{2\pi}} \cdot 2\xi = \frac{\xi e^{-(\xi^2)/2}}{\sqrt{2\pi}} \cdot 2\xi = \frac{2\xi^2 e^{-(\xi^2)/2}}{\sqrt{2\pi}} \end{aligned}$$

因此,  $\xi$  的概率密度函数为:

$$f_{\xi}(\xi) = \frac{2\xi^2 e^{-(\xi^2)/2}}{\sqrt{2\pi}}, \quad (\xi \geq 0)$$



18. 设随机变量  $X_1, \dots, X_n$  独立同分布，共同的分布是威布尔分布，即共同的分布函数是

$$F(x) = \begin{cases} 1 - e^{-\left(\frac{x}{\eta}\right)^m}, & x > 0, \\ 0, & x \leq 0, \end{cases} \quad (m > 0, \eta > 0),$$

试证明  $\xi \triangleq \min\{X_1, \dots, X_n\}$  仍服从威布尔分布。

18. 设随机变量  $X_1, \dots, X_n$  独立同分布，共同的分布是威布尔分布，即共同的分布函数是

$$F(x) = \begin{cases} 1 - e^{-\left(\frac{x}{\eta}\right)^m}, & x > 0, \\ 0, & x \leq 0, \end{cases} \quad (m > 0, \eta > 0),$$

试证明  $\xi \triangleq \min\{X_1, \dots, X_n\}$  仍服从威布尔分布。

证明：由于  $\xi = \min\{X_1, X_2, \dots, X_n\}$ ，所以  $\xi > x$  当且仅当所有  $X_i > x$  ( $i = 1, 2, \dots, n$ )。因为  $X_i$  是独立的，我们有：

$$\begin{aligned} P(\xi > x) &= P(X_1 > x, X_2 > x, \dots, X_n > x) = \prod_{i=1}^n P(X_i > x) \\ &= \prod_{i=1}^n (1 - F(x)) = e^{-n\left(\frac{x}{\eta}\right)^m} \end{aligned}$$

所以：

$$F_{\xi}(x) = 1 - P(\xi > x) = 1 - e^{-n\left(\frac{x}{\eta}\right)^m}$$

设  $\eta' = \eta \left(\frac{1}{n}\right)^{1/m}$ ，则：

$$F_{\xi}(x) = 1 - e^{-\left(\frac{x}{\eta'}\right)^m}$$

因此， $\xi = \min\{X_1, X_2, \dots, X_n\}$  仍然服从威布尔分布，其参数为：

$$\eta' = \eta \left(\frac{1}{n}\right)^{1/m}.$$

19. 对于随机变量  $X, Y, Z$ , 已知

$$\begin{aligned}E(X) &= E(Y) = 1, & E(Z) &= -1, \\ \text{var}(X) &= \text{var}(Y) = \text{var}(Z) = 1, \\ \rho_{XY} &= 0, & \rho_{XZ} &= 1/2, & \rho_{YZ} &= -1/2,\end{aligned}$$

试求  $E(X + Y + Z)$  及  $\text{var}(X + Y + Z)$ .

解:

$$E(X + Y + Z) = E(X) + E(Y) + E(Z) = 1 + 1 - 1 = 1$$

$$\begin{aligned}& \text{var}(X + Y + Z) \\ &= \text{var}(X) + \text{var}(Y) + \text{var}(Z) + 2(\text{cov}(X, Y) + \text{cov}(X, Z) + \text{cov}(Y, Z))\end{aligned}$$

计算协方差:

$$\text{cov}(X, Y) = \rho_{XY} \sqrt{\text{var}(X)\text{var}(Y)} = 0 \cdot \sqrt{1 \cdot 1} = 0$$

$$\text{cov}(X, Z) = \rho_{XZ} \sqrt{\text{var}(X) \text{var}(Z)} = \frac{1}{2} \cdot \sqrt{1 \cdot 1} = \frac{1}{2}$$

$$\text{cov}(Y, Z) = \rho_{YZ} \sqrt{\text{var}(Y) \text{var}(Z)} = -\frac{1}{2} \cdot \sqrt{1 \cdot 1} = -\frac{1}{2}$$

代入方差公式：

$$\text{var}(X + Y + Z) = 1 + 1 + 1 + 2 \left( 0 + \frac{1}{2} - \frac{1}{2} \right) = 3$$

因此,  $E(X + Y + Z) = 1$ ,  $\text{var}(X + Y + Z) = 3$ 。

解: 设变换:

$$\begin{pmatrix} U \\ V \end{pmatrix} = \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix} \begin{pmatrix} X \\ Y \end{pmatrix}$$

变换矩阵  $\mathbf{A} = \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}$ , 逆变换矩阵

$$\mathbf{A}^{-1} = \frac{1}{\det(\mathbf{A})} \begin{pmatrix} -1 & -1 \\ -1 & 1 \end{pmatrix} = \frac{1}{2} \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}.$$

由于  $X$  和  $Y$  相互独立且都服从标准正态分布, 它们的联合密度函数为:

$$f_{X,Y}(x,y) = \frac{1}{2\pi} e^{-\frac{x^2+y^2}{2}}$$

代入逆变换:

$$x(u, v) = \frac{u+v}{2}, \quad y(u, v) = \frac{u-v}{2}, \quad |\det(\mathbf{A}^{-1})| = \left| \frac{1}{2} \right| = \frac{1}{2}$$

所以,  $U, V$  的联合密度为:

$$\begin{aligned} f_{U,V}(u,v) &= f_{X,Y}(x(u,v), y(u,v)) \cdot |\det(\mathbf{A}^{-1})| \\ &= \frac{1}{2\pi} e^{-\frac{(\frac{u+v}{2})^2 + (\frac{u-v}{2})^2}{2}} \cdot \frac{1}{2} = \frac{1}{2\pi} e^{-\frac{u^2+v^2}{4}} \cdot \frac{1}{2} = \frac{1}{4\pi} e^{-\frac{u^2+v^2}{4}}. \end{aligned}$$

证:  $U = X^2 + Y^2$  与  $V = X/Y$  是相互独立的.





① 的边缘密度函数:

$$f_{\Theta}(\theta) = \int_0^{\infty} f_{R,\Theta}(r, \theta) dr = \int_0^{\infty} \frac{r}{2\pi} e^{-\frac{r^2}{2}} dr = \frac{1}{2\pi}$$

注意到  $f_{R,\Theta}(r, \theta) = f_R(r) \cdot f_\Theta(\theta)$ , 这表明  $R$  和  $\Theta$  是相互独立的。  
又由于

$$U = X^2 + Y^2 = R^2, \quad V = \frac{X}{Y} = \frac{\cos(\Theta)}{\sin(\Theta)} = \cot(\Theta)$$

因为  $R$  和  $\Theta$  是相互独立的，而  $U$  只依赖于  $R$ ， $V$  只依赖于  $\Theta$ ，所以  $U$  和  $V$  也是相互独立的。



解：

$$E(X) = \frac{0+1}{2} = \frac{1}{2}, \quad E(Y) = \frac{1+3}{2} = 2$$

由于  $X$  和  $Y$  相互独立,

$$E(XY) = E(X) \cdot E(Y) = \frac{1}{2} \cdot 2 = 1$$

$$E(X^2) = \int_0^1 x^2 dx = \frac{1}{3}, \quad E(Y^2) = \int_1^3 \frac{1}{2} y^2 dy = \frac{27}{6} - \frac{1}{6} = \frac{13}{3}$$

同样,  $X^2$  和  $Y^2$  相互独立,

$$E((XY)^2) = E(X^2) \cdot E(Y^2) = \frac{1}{3} \cdot \frac{13}{3} = \frac{13}{9}$$

因此

$$\text{var}(XY) = E((XY)^2) - [E(XY)]^2 = \frac{13}{9} - 1^2 = \frac{4}{9}$$

**23.** 设随机变量  $X$  与  $Y$  相互独立, 且  $\text{var}(X)$  和  $\text{var}(Y)$  存在, 试证:

$$\text{var}(XY) \geq \text{var}(X) \text{var}(Y).$$

23. 设随机变量  $X$  与  $Y$  相互独立, 且  $\text{var}(X)$  和  $\text{var}(Y)$  存在, 试证:

$$\text{var}(XY) \geq \text{var}(X) \text{var}(Y).$$

证明: 由于  $X$  和  $Y$  相互独立, 有:

$$E((XY)^2) = E(X^2 Y^2) = E(X^2)E(Y^2), \quad E(XY) = E(X)E(Y)$$

计算方差:

$$\begin{aligned} \text{var}(X) \text{var}(Y) &= (E(X^2) - [E(X)]^2)(E(Y^2) - [E(Y)]^2) \\ &= E(X^2)E(Y^2) - E(X^2)[E(Y)]^2 - [E(X)]^2 E(Y^2) + [E(X)]^2 [E(Y)]^2 \end{aligned}$$

$$\text{var}(XY) = E(X^2)E(Y^2) - [E(X)E(Y)]^2$$

由于:

$$E(X^2) \geq [E(X)]^2, \quad E(Y^2) \geq [E(Y)]^2$$

因此

$$E(X^2)[E(Y)]^2 + [E(X)]^2 E(Y^2) \geq 2[E(X)]^2 [E(Y)]^2$$

$$\begin{aligned} & \text{var}(X) \text{var}(Y) - \text{var}(XY) \\ &= 2[E(X)]^2 [E(Y)]^2 - E(X^2)[E(Y)]^2 + [E(X)]^2 E(Y^2) \leq 0 \end{aligned}$$

即

$$\text{var}(XY) \geq \text{var}(X) \text{var}(Y)$$



- 1 第三章作业
- 2 第二部分
- 3 多元正态分布

24. 设一城市有  $n$  个区, 其中住有  $x_j$  个居民的区共有  $n_j$  个 ( $\sum n_j = n$ ). 令

$$m = \sum_j \frac{n_j x_j}{n}, \quad \sigma^2 = \sum_j \frac{n_j x_j^2}{n} - m^2.$$

( $m$  是每个区的居民数的平均均数). 现在随机选取  $r$  个区, 并数出其中每个区中的居民数, 设  $X_1, \dots, X_r$  分别为这  $r$  个区的居民数, 试证:

$$E(X_1 + \dots + X_r) = mr, \quad \text{var}(X_1 + \dots + X_r) = \frac{\sigma^2 r(n-r)}{n-1}.$$

$$m = \sum_j \frac{n_j x_j}{n}, \quad \sigma^2 = \sum_j \frac{n_j x_j^2}{n} - m^2.$$
$$E(X_1 + \cdots + X_r) = mr, \quad \text{var}(X_1 + \cdots + X_r) = \frac{\sigma^2 r(n-r)}{n-1}.$$
$$E(X_i) = \sum_j x_j \cdot P(X_i = x_j) = \sum_j x_j \cdot \frac{n_j}{n} = \frac{1}{n} \sum_j n_j x_j = m$$

$$E(X_i^2) = \sum_j x_j^2 \cdot P(X_i = x_j) = \sum_j x_j^2 \cdot \frac{n_j}{n} = \frac{1}{n} \sum_j n_j x_j^2$$

因此：

$$\text{var}(X_i) = \frac{1}{n} \sum_j n_j x_j^2 - m^2 = \sigma^2$$

$$E(X_1 + X_2 + \cdots + X_r) = E(X_1) + E(X_2) + \cdots + E(X_r) = r \cdot E(X_i) = r \cdot m$$

对于任意两个变量  $X_i$  和  $X_j$  ,  $E(X_i X_j)$  可以表示为：

$$\begin{aligned} E(X_i X_j) &= \frac{1}{n(n-1)} \sum_{i \neq j} x_i x_j \\ &= \frac{1}{n(n-1)} \left[ \left( \sum_j x_j \right)^2 - \sum_j x_j^2 \right] = \frac{n^2 m^2 - n \sum_j \frac{n_j x_j^2}{n}}{n(n-1)} \\ &= \frac{1}{n(n-1)} (n^2 m^2 - n(n\sigma^2 + nm^2)) = \frac{-n^2 \sigma^2}{n(n-1)} \end{aligned}$$

协方差为

$$\text{cov}(X_i, X_j) = \frac{-n\sigma^2}{n-1} - m^2 = \frac{-n\sigma^2 - (n-1)m^2}{n-1} = \frac{-\sigma^2}{n-1}$$

$$\begin{aligned} & \text{var}(X_1 + X_2 + \cdots + X_r) \\ &= \sum_{i=1}^r \text{var}(X_i) + 2 \sum_{1 \leq i < j \leq r} \text{cov}(X_i, X_j) = r\sigma^2 - \frac{r(r-1)\sigma^2}{n-1} \\ &= \sigma^2 \left( \frac{r(n-r)}{n-1} \right) \end{aligned}$$

25. 设  $X_1, \dots, X_n$  是独立同分布的正值随机变列, 试证:

$$E \left( \frac{X_1 + \dots + X_k}{X_1 + \dots + X_n} \right) = \frac{k}{n} \quad (k = 1, \dots, n).$$

25. 设  $X_1, \dots, X_n$  是独立同分布的正值随机变列, 试证:

$$E\left(\frac{X_1 + \dots + X_k}{X_1 + \dots + X_n}\right) = \frac{k}{n} \quad (k = 1, \dots, n).$$

证明: 设  $S_n = X_1 + X_2 + \dots + X_n$  和  $S_k = X_1 + X_2 + \dots + X_k$ 。由于  $X_1, X_2, \dots, X_n$  是独立同分布的正值随机变量, 对于任意  $i$  和  $j$ ,  $\frac{X_i}{S_n}$  和  $\frac{X_j}{S_n}$  的期望值相同。因此, 我们可以考虑所有  $\frac{X_i}{S_n}$  的和:

$$E\left(\sum_{i=1}^n \frac{X_i}{S_n}\right) = E(1) = 1$$

由于  $\frac{X_i}{S_n}$  的期望值相同, 因此:

$$E\left(\frac{X_i}{S_n}\right) = \frac{1}{n} \quad \text{对于所有 } i = 1, 2, \dots, n$$

$$\begin{aligned} E\left(\frac{S_k}{S_n}\right) &= E\left(\frac{X_1 + X_2 + \dots + X_k}{S_n}\right) \\ &= E\left(\sum_{i=1}^k \frac{X_i}{S_n}\right) = \sum_{i=1}^k E\left(\frac{X_i}{S_n}\right) = k \cdot \frac{1}{n} = \frac{k}{n} \end{aligned}$$



26. 设随机变量  $X_1, \dots, X_n$  独立同分布,  $X_i \sim N(\mu, \sigma^2)$ , 记

$$\xi = \sum_{i=1}^m X_i, \quad \eta = \sum_{i=1}^n X_i \quad (1 \leq m < n),$$

试求  $(\xi, \eta)$  的联合密度.

26. 设随机变量  $X_1, \dots, X_n$  独立同分布,  $X_i \sim N(\mu, \sigma^2)$ , 记

$$\xi = \sum_{i=1}^m X_i, \quad \eta = \sum_{i=1}^n X_i \quad (1 \leq m < n),$$

试求  $(\xi, \eta)$  的联合密度.

解:  $\xi, \eta$  是  $m$  个独立正态随机变量的和, 因此  $\xi, \eta$  也服从正态分布:

$$\xi \sim N(m\mu, m\sigma^2), \quad \eta \sim N(n\mu, n\sigma^2)$$

由于  $X_i$  独立同分布, 只有当  $i=j$  时,  $\text{cov}(X_i, X_j) = \sigma^2$ , 否则为 0。因此  $\xi$  和  $\eta$  的协方差为:

$$\text{cov}(\xi, \eta) = \text{cov} \left( \sum_{i=1}^m X_i, \sum_{j=1}^n X_j \right) = \sum_{i=1}^m \sum_{j=1}^n \text{cov}(X_i, X_j) = m\sigma^2$$

$$\mu = \begin{pmatrix} m\mu \\ n\mu \end{pmatrix}, \quad \Sigma = \begin{pmatrix} m\sigma^2 & m\sigma^2 \\ m\sigma^2 & n\sigma^2 \end{pmatrix}$$

$$\mu = \begin{pmatrix} m\mu \\ n\mu \end{pmatrix}, \quad \Sigma = \begin{pmatrix} m\sigma^2 & m\sigma^2 \\ m\sigma^2 & n\sigma^2 \end{pmatrix}$$

$$f_{\xi,\eta}(\mathbf{x}, y) = \frac{1}{2\pi\sqrt{\det(\boldsymbol{\Sigma})}} \exp\left(-\frac{1}{2}(\mathbf{z} - \boldsymbol{\mu})^\top \boldsymbol{\Sigma}^{-1}(\mathbf{z} - \boldsymbol{\mu})\right)$$

27. 设随机变量  $X$  与  $Y$  独立同分布,  $X \sim N(\mu, \sigma^2)$ ,  $\alpha, \beta$  是两个实数 (全不为 0).

(1) 求  $\alpha X + \beta Y$  与  $\alpha X - \beta Y$  的相关系数和联合密度;

(2) 证明:  $E(\max\{X, Y\}) = \mu + \frac{\sigma}{\sqrt{\pi}}$ .

27. 设随机变量  $X$  与  $Y$  独立同分布,  $X \sim N(\mu, \sigma^2)$ ,  $\alpha, \beta$  是两个实数 (全不为 0).

(1) 求  $\alpha X + \beta Y$  与  $\alpha X - \beta Y$  的相关系数和联合密度;

(2) 证明:  $E(\max\{X, Y\}) = \mu + \frac{\sigma}{\sqrt{\pi}}$ .

解: (1) 设  $U = \alpha X + \beta Y$ ,  $V = \alpha X - \beta Y$ , 期望为

$$E(U) = E(\alpha X + \beta Y) = \alpha E(X) + \beta E(Y) = \alpha\mu + \beta\mu = (\alpha + \beta)\mu$$

$$E(V) = E(\alpha X - \beta Y) = \alpha E(X) - \beta E(Y) = \alpha\mu - \beta\mu = (\alpha - \beta)\mu$$

由于  $X$  和  $Y$  独立同分布:

$$\text{Var}(U) = \text{Var}(\alpha X + \beta Y) = \alpha^2 \text{Var}(X) + \beta^2 \text{Var}(Y) = (\alpha^2 + \beta^2)\sigma^2$$

$$\text{Var}(V) = \text{Var}(\alpha X - \beta Y) = \alpha^2 \text{Var}(X) + \beta^2 \text{Var}(Y) = (\alpha^2 + \beta^2)\sigma^2$$

$U, V$  的协方差为

$$\begin{aligned}\text{Cov}(U, V) &= \text{Cov}(\alpha X + \beta Y, \alpha X - \beta Y) \\ &= \alpha^2 \text{Cov}(X, X) - \beta^2 \text{Cov}(Y, Y) = \alpha^2 \sigma^2 - \beta^2 \sigma^2 = (\alpha^2 - \beta^2) \sigma^2\end{aligned}$$

相关系数

$$\rho_{UV} = \frac{\text{Cov}(U, V)}{\sqrt{\text{Var}(U) \text{Var}(V)}} = \frac{(\alpha^2 - \beta^2) \sigma^2}{\sqrt{(\alpha^2 + \beta^2) \sigma^2 (\alpha^2 + \beta^2) \sigma^2}} = \frac{\alpha^2 - \beta^2}{\alpha^2 + \beta^2}$$

由于  $X$  和  $Y$  独立同分布且均为正态分布,  $U$  和  $V$  也是正态分布的线性组合, 因此  $(U, V)$  也是二维正态分布, 参数为

$$\boldsymbol{\mu} = \begin{pmatrix} (\alpha + \beta) \mu \\ (\alpha - \beta) \mu \end{pmatrix}, \quad \boldsymbol{\Sigma} = \begin{pmatrix} (\alpha^2 + \beta^2) \sigma^2 & (\alpha^2 - \beta^2) \sigma^2 \\ (\alpha^2 - \beta^2) \sigma^2 & (\alpha^2 + \beta^2) \sigma^2 \end{pmatrix}$$

$$f_{\xi, \eta}(u, v) = \frac{1}{2\pi \sqrt{\det(\boldsymbol{\Sigma})}} \exp \left( -\frac{1}{2} ((\mathbf{u}; \mathbf{v}) - \boldsymbol{\mu})^\top \boldsymbol{\Sigma}^{-1} ((\mathbf{u}; \mathbf{v}) - \boldsymbol{\mu}) \right)$$

(2) 令  $X_1 = \frac{X-\mu}{\sigma}, Y_1 = \frac{Y-\mu}{\sigma}$ ,  $X_1, Y_1$  均服从标准正态分布,  $\max\{X, Y\} = \mu + \sigma \max\{X_1, Y_1\}$ . 注意到

$$\max\{X_1, Y_1\} = \frac{1}{2}(X_1 + Y_1 + |X_1 - Y_1|)$$

由于  $X_1 - Y_1 \sim N(0, 2)$ ,

$$E|X_1 - Y_1| = \frac{1}{2\pi} \int_{-\infty}^{+\infty} |x| \exp\left\{-\frac{x^2}{4}\right\} dx = \frac{2}{\sqrt{\pi}}$$

因此  $E(\max\{X, Y\}) = \mu + \frac{\sigma}{\sqrt{\pi}}$ .

28. 考虑  $m(m \geq 2)$  个独立试验. 每个试验具有  $r(r \geq 2)$  个可能的试验结果, 相应出现的概率分别为  $p_1, \dots, p_r$   $\left(\sum_{i=1}^r p_i = 1\right)$ .

用  $X_i$  表示  $m$  个试验中结果  $i(i = 1, \dots, r)$  出现的次数. 试求出:

(1)  $r$  维随机向量  $(X_1, \dots, X_r)$  的概率分布;

(2)  $X_i$  与  $X_j(i \neq j)$  的协方差。



28. 考虑  $m(m \geq 2)$  个独立试验. 每个试验具有  $r(r \geq 2)$  个可能的试验结果, 相应出现的概率分别为  $p_1, \dots, p_r$   $\left(\sum_{i=1}^r p_i = 1\right)$ .

用  $X_i$  表示  $m$  个试验中结果  $i(i = 1, \dots, r)$  出现的次数. 试求出:

(1)  $r$  维随机向量  $(X_1, \dots, X_r)$  的概率分布;

(2)  $X_i$  与  $X_j(i \neq j)$  的协方差。

解: (1) 设  $X_i$  是  $m$  次试验中结果  $i$  出现的次数, 则  $r$  维随机向量  $(X_1, X_2, \dots, X_r)$  的概率分布是一个多项式分布, 其分布可以表示为:

$$P(X_1 = x_1, X_2 = x_2, \dots, X_r = x_r) = \frac{m!}{x_1! x_2! \dots x_r!} p_1^{x_1} p_2^{x_2} \dots p_r^{x_r}$$

(2) 由于  $X_i$  服从二项分布，其期望和方差为：

$$E(X_i) = mp_i, \quad \text{Var}(X_i) = mp_i(1-p_i), \quad E(X_i^2) = (mp_i)^2 + mp_i(1-p_i)$$

对于任意的  $X_i, X_j$ ,  $X_i + X_j \sim B(m, p_i + p_j)$ , 我们有：

$$E(X_i + X_j)^2 = (m(p_i + p_j))^2 + m(p_i + p_j)(1 - p_i - p_j)$$

$$\begin{aligned} E(X_i X_j) &= \frac{1}{2}(E(X_i + X_j)^2 - E(X_i^2) - E(X_j^2)) \\ &= m^2 p_i p_j - m p_i p_j \end{aligned}$$

因此：

$$\text{Cov}(X_i, X_j) = E(X_i X_j) - E(X_i)E(X_j) = -m p_i p_j$$

29. 设  $X_1, \dots, X_n (n \geq 2)$  是独立同分布的随机变量列, 且

$E[(X_1 - E(X_1))^3] = 0$ , 试证: 随机变量  $\xi = \frac{1}{n} \sum_{i=1}^n X_i$  与

$\eta = \sum_{i=1}^n (X_i - \xi)^2$  是不相关的。

29. 设  $X_1, \dots, X_n (n \geq 2)$  是独立同分布的随机变量列, 且

$$E[(X_1 - E(X_1))^3] = 0, \text{ 试证: 随机变量 } \xi = \frac{1}{n} \sum_{i=1}^n X_i \text{ 与}$$

$$\eta = \sum_{i=1}^n (X_i - \xi)^2 \text{ 是不相关的。}$$

证明: 证明: 设  $X_1, X_2, \dots, X_n$  中每个  $X_i$  的期望为  $\mu$ , 方差为  $\sigma^2$ 。由独立同分布:

$$E(X_i) = \mu, \quad \text{Var}(X_i) = \sigma^2, \quad E(X_i X_j) = \begin{cases} \mu^2, & i \neq j, \\ \mu^2 + \sigma^2, & i = j. \end{cases}$$

$$E(\xi) = E\left(\frac{1}{n} \sum_{i=1}^n X_i\right) = \frac{1}{n} \sum_{i=1}^n E(X_i) = \frac{1}{n} \sum_{i=1}^n \mu = \mu$$

$$\begin{aligned}
 E(\eta) &= E\left(\sum_{i=1}^n (X_i - \xi)^2\right) = \sum_{i=1}^n E((X_i - \xi)^2) \\
 &= \sum_{i=1}^n \left[ E(X_i^2) - \frac{2}{n} \sum_{j=1}^n E(X_i X_j) + \left(\frac{1}{n} \sum_{j=1}^n X_j\right)^2 \right] \\
 &= \sum_{i=1}^n \left[ \mu^2 + \sigma^2 - \frac{2}{n} ((n-1)\mu^2 + \mu^2 + \sigma^2) + \frac{1}{n^2} (n(\mu^2 + \sigma^2) + n(n-1)\mu^2) \right] \\
 &= \sum_{i=1}^n \frac{(n-1)\sigma^2}{n} = (n-1)\sigma^2
 \end{aligned}$$

$$E[(X_1 - E(X_1))^3] = E(X_i^3) - 3\mu E(X_i^2) + 3\mu^2 E(X_i) - \mu^3 = E(X_i^3) - \mu^3 - 3\mu\sigma^2$$

因此由题意  $E(X_i^3) = \mu^3 + 3\mu\sigma^2$ ,

$$\begin{aligned} E(\xi\eta) &= E\left(\xi \sum_{i=1}^n (X_i - \xi)^2\right) = \sum_{i=1}^n E(\xi(X_i - \xi)^2) \\ &= \sum_{i=1}^n E\left(\frac{1}{n} \sum_{j=1}^n X_j (X_i - \xi)^2\right) \end{aligned}$$

其中

$$E(X_j(X_i - \xi)^2) = E(X_i^2 X_j) - \frac{2}{n} \sum_{k=1}^n E(X_i X_j X_k) + \frac{1}{n^2} \sum_{k=1}^n \sum_{l=1}^n E(X_j X_k X_l)$$

若  $i = j$ ,

$$\begin{aligned} E(X_j(X_i - \xi)^2) &= \mu^3 + 3\mu\sigma^2 - \frac{2}{n}((n-1)\mu(\mu^2 + \sigma^2) + \mu^3 + 3\mu\sigma^2) \\ &+ \frac{1}{n^2}(\mu^3 + 3\mu\sigma^2 + (n^2 - 3n + 2)\mu^3 + (3n - 3)(\mu^2 + \sigma^2)\mu) = \frac{3n-3}{n^2}\mu\sigma^2 \end{aligned}$$

若  $i \neq j$ ,

$$\begin{aligned} E(X_j(X_i - \xi)^2) &= \mu(\mu^2 + \sigma^2) - \frac{2}{n}(2\mu(\mu^2 + \sigma^2) + (n-2)\mu^3) \\ &+ \frac{1}{n^2}(\mu^3 + 3\mu\sigma^2 + (n^2 - 3n + 2)\mu^3 + (3n - 3)(\mu^2 + \sigma^2)\mu) = \frac{n-3}{n^2}\mu\sigma^2 \end{aligned}$$

因此

$$\begin{aligned} E(\xi\eta) &= \frac{1}{n} \sum_{i=1}^n \sum_{j=1}^n E(X_j(X_i - \xi)^2) \\ &= \frac{1}{n} \left( \frac{3n-3}{n^2} n\mu\sigma^2 + \frac{n-3}{n^2} (n-1)n\mu\sigma^2 \right) = (n-1)\mu\sigma^2 \end{aligned}$$

协方差为

$$\begin{aligned} \text{Cov}(\xi, \eta) &= E(\xi\eta) - E(\xi)E(\eta) = \mu\sigma^2(n-1) - \mu \cdot (n-1)\sigma^2 \\ &= \mu\sigma^2(n-1 - (n-1)) = 0 \end{aligned}$$

故  $\xi$  和  $\eta$  不相关



30. 若  $X$  的分布密度是偶函数, 且  $E(X^2)$  存在, 试证:  $|X|$  与  $X$  不相关, 但它们不相互独立. 若  $X_1, \dots, X_n (n \geq 2)$  相互独立,  $\text{var}(X_i) = \sigma_i^2 (i = 1, \dots, n)$ , 试找“权”

$a_1, \dots, a_n \left( a_i \geq 0, \sum_{i=1}^n a_i = 1 \right)$ , 使得  $\sum_{i=1}^n a_i X_i$  的方差最小。

30. 若  $X$  的分布密度是偶函数, 且  $E(X^2)$  存在, 试证:  $|X|$  与  $X$  不相关, 但它们不相互独立. 若  $X_1, \dots, X_n (n \geq 2)$  相互独立,  $\text{var}(X_i) = \sigma_i^2 (i = 1, \dots, n)$ , 试找“权”

$a_1, \dots, a_n \left( a_i \geq 0, \sum_{i=1}^n a_i = 1 \right)$ , 使得  $\sum_{i=1}^n a_i X_i$  的方差最小。

解: 由于  $f(x)$  是偶函数,  $X$  的期望  $E(X) = \int_{-\infty}^{\infty} xf(x) dx = 0$ ,

$$E(|X|X) = \int_{-\infty}^{\infty} |x|xf(x) dx = \int_{-\infty}^0 (-x)xf(x) dx + \int_0^{\infty} xxf(x) dx = 0$$

$$\text{Cov}(|X|, X) = E(|X|X) - E(|X|)E(X) = 0 - 0 = 0$$

因此,  $|X|$  和  $X$  不相关。

考虑  $X$  的具体分布, 例如  $X \sim N(0, 1)$ 。在这种情况下,  $|X|$  的分布为半正态分布。显然,  $P(|X| = x, X = -x) = 0$ , 而  $P(|X| = x)P(X = -x) > 0$ , 这说明  $|X|$  和  $X$  不独立。

因此,  $|X|$  和  $X$  不相关但不独立。

考虑线性组合  $Y = \sum_{i=1}^n a_i X_i$ , 其方差为:

$$\text{Var}(Y) = \text{Var} \left( \sum_{i=1}^n a_i X_i \right) = \sum_{i=1}^n a_i^2 \text{Var}(X_i) = \sum_{i=1}^n a_i^2 \sigma_i^2$$

其中, 若有方差  $\sigma_{i_0}^2 = 0$ , 则由约束条件  $\sum_{i=1}^n a_i = 1$  知, 应取  $a_{i_0} = 1, a_i = 0 (i \neq i_0)$ 。

若所有方差  $\sigma_i^2 > 0$ , 则为了使  $\text{Var}(Y)$  最小, 我们需要最小化  $\sum_{i=1}^n a_i^2 \sigma_i^2$ 。这是一个带约束的优化问题, 可以使用拉格朗日乘数法解决。定义拉格朗日函数:

$$L(a_1, a_2, \dots, a_n, \lambda) = \sum_{i=1}^n a_i^2 \sigma_i^2 + \lambda \left( 1 - \sum_{i=1}^n a_i \right)$$

对  $a_i$  求导并令导数为零:

$$\frac{\partial L}{\partial a_i} = 2a_i \sigma_i^2 - \lambda = 0 \implies a_i = \frac{\lambda}{2\sigma_i^2}$$

利用  $\sum_{i=1}^n a_i = 1$ :

$$\sum_{i=1}^n \frac{\lambda}{2\sigma_i^2} = 1 \implies \lambda \sum_{i=1}^n \frac{1}{2\sigma_i^2} = 1 \implies \lambda = \frac{2}{\sum_{i=1}^n \frac{1}{\sigma_i^2}}$$

代入  $\lambda$  得到最优权重:

$$a_i = \frac{1}{\sigma_i^2 \sum_{j=1}^n \frac{1}{\sigma_j^2}}$$

31. 若  $X$  与  $Y$  都是只取两个值的随机变量，试证明：若  $X$  与  $Y$  不相关，则  $X$  与  $Y$  相互独立。

31. 若  $X$  与  $Y$  都是只取两个值的随机变量，试证明：若  $X$  与  $Y$  不相关，则  $X$  与  $Y$  相互独立。

证明：不妨设  $X, Y$  分别取  $x_1, x_2$  和  $y_1, y_2$ ,  $P(X = x_1) = q_1$ ,  $P(X = x_2) = 1 - q_1$ ;  $P(Y = y_1) = q_2$ ,  $P(Y = y_2) = 1 - q_2$ ;  
 $P(X = x_1, Y = y_1) = p_1$ ,  $P(X = x_1, Y = y_2) = p_2$ ,  
 $P(X = x_2, Y = y_1) = p_3$ ,  $P(X = x_2, Y = y_2) = p_4$ .  
 由于不相关，

$$\begin{aligned} \text{cov}(X, Y) \\ = x_1 y_1 p_1 + x_1 y_2 p_2 + x_2 y_1 p_3 + x_2 y_2 p_4 \\ - (x_1 q_1 + x_2(1 - q_1))(y_1 q_2 + y_2(1 - q_2)) = 0 \end{aligned}$$

其中  $q_1 = p_1 + p_2$ ,  $q_2 = p_1 + p_3$ ,  $p_1 + p_2 + p_3 + p_4 = 1$   
 假设  $q_1, q_2$  已知，求解方程组得

$$p_1 = q_1 q_2, \quad p_2 = q_1(1 - q_2), \quad p_3 = (1 - q_1)q_2, \quad p_4 = (1 - q_1)(1 - q_2).$$

故  $X$  与  $Y$  相互独立

*(continued)*

1. *Journal of the American Medical Association*, 1997; 278: 1039-1044.

1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26 27 28 29 30 31 32 33 34 35 36 37 38 39 40 41 42 43 44 45 46 47 48 49 50 51 52 53 54 55 56 57 58 59 60 61 62 63 64 65 66 67 68 69 70 71 72 73 74 75 76 77 78 79 80 81 82 83 84 85 86 87 88 89 90 91 92 93 94 95 96 97 98 99 100

\_\_\_\_\_

[illegible]

原书(付)1部、分、出、卖、年、市、和、改、元、大、師、撰、付

因此，如果  $X$  和  $Y$  相互独立，则

$E(\delta(a - X)\delta(b - Y)) = E(\delta(a - X))E(\delta(b - Y))$  对于所有  $a$  和  $b$  成立。

必要性：

假设  $E(\delta(a - X)\delta(b - Y)) = E(\delta(a - X))E(\delta(b - Y))$  对于所有  $a$  和  $b$  成立。

设  $A = \{X \leq a\}$  和  $B = \{Y \leq b\}$ 。

$$P(A \cap B) = P(X \leq a, Y \leq b) = E(\delta(a - X)\delta(b - Y))$$

根据假设  $E(\delta(a - X)\delta(b - Y)) = E(\delta(a - X))E(\delta(b - Y))$ ，我们有：

$$P(X \leq a, Y \leq b) = P(X \leq a)P(Y \leq b)$$

对于任意的  $a$  和  $b$ ，我们都有

$P(X \leq a, Y \leq b) = P(X \leq a)P(Y \leq b)$ 。这意味着对于任意的  $A$  和  $B$ ，有  $P(A \cap B) = P(A)P(B)$ 。因此， $X$  和  $Y$  相互独立。





34. 一辆交通车送 25 名乘客到 7 个站, 假设每一个乘客等可能地在任一站下车, 且他们行动独立, 交通车只在有人下车时才停站, 问: 该交通车停站的期望次数是多少?

解: 设  $X_i$  是一个指示变量, 表示第  $i$  站是否有乘客下车 (如果有乘客下车则  $X_i = 1$ , 否则  $X_i = 0$ )。交通车停站的总次数  $X$  可以表示为  $X = X_1 + X_2 + \cdots + X_7$ , 期望为

$$E[X] = E[X_1 + X_2 + \cdots + X_7] = E[X_1] + E[X_2] + \cdots + E[X_7]$$

第  $i$  站至少有一名乘客下车的概率为:

$$P(\text{第 } i \text{ 站至少有一名乘客下车}) = 1 - \left(\frac{6}{7}\right)^{25}$$

因此,

$$E[X_i] = 1 - \left(\frac{6}{7}\right)^{25}, \quad E[X] = 7 \times \left(1 - \left(\frac{6}{7}\right)^{25}\right)$$



35. 50 个人排队作肺部透视, 假设他们中有 4 个阳性患者, 问: 在出现第一个阳性患者之前, 阴性反应者的人数平均是多少?

解: 设  $X$  为第一位阳性患者出现的位置, 则

$$P(X \geq k) = \frac{C_{51-k}^4}{C_{50}^4}$$

因此所求平均值为

$$\begin{aligned} E(X) - 1 &= \sum_{k=1}^{47} P(X \geq k) - 1 = \sum_{k=1}^{47} \frac{C_{51-k}^4}{C_{50}^4} - 1 \\ &= \frac{C_{51}^5}{C_{50}^4} - 1 = \frac{46}{5} \end{aligned}$$

38. 若对于随机变量  $X$ ,  $E(e^{aX})$  存在 ( $a$  是正常数), 试证:

$$P(X \geq \varepsilon) \leq e^{-a\varepsilon} E(e^{aX}) \quad (\text{一切 } \varepsilon > 0).$$

38. 若对于随机变量  $X$ ,  $E(e^{aX})$  存在 ( $a$  是正常数), 试证:

$$P(X \geq \varepsilon) \leq e^{-a\varepsilon} E(e^{aX}) \quad (\text{一切 } \varepsilon > 0).$$

证明: 设随机变量  $Y = e^{aX}$ , 其中  $a > 0$ 。显然,  $Y$  是非负的。  
由马尔可夫不等式, 得:

$$P(X \geq \varepsilon) = P(Y \geq e^{a\varepsilon}) \leq \frac{E(Y)}{e^{a\varepsilon}}.$$

因此:

$$P(X \geq \varepsilon) \leq e^{-a\varepsilon} E(e^{aX}).$$

39. 设随机变量  $X$  与  $Y$  相互独立,

$X \sim N(\mu_1, \sigma_1^2)$ ,  $Y \sim N(\mu_2, \sigma_2^2)$ , 已知  $\mu_1, \mu_2, \sigma_1^2, \sigma_2^2$ , 试求

$P(X > Y)$ .

39. 设随机变量  $X$  与  $Y$  相互独立,

$X \sim N(\mu_1, \sigma_1^2)$ ,  $Y \sim N(\mu_2, \sigma_2^2)$ , 已知  $\mu_1, \mu_2, \sigma_1^2, \sigma_2^2$ , 试求  $P(X > Y)$ .

解: 设  $Z = X - Y$ 。由于  $X$  和  $Y$  是相互独立的,  $Z$  也是一个正态随机变量,  $Z \sim N(\mu_1 - \mu_2, \sigma_1^2 + \sigma_2^2)$ 。

令  $W = \frac{Z - (\mu_1 - \mu_2)}{\sqrt{\sigma_1^2 + \sigma_2^2}}$ , 则  $W$  服从标准正态分布  $N(0, 1)$ 。因此:

$$P(Z > 0) = P\left(W > \frac{\mu_2 - \mu_1}{\sqrt{\sigma_1^2 + \sigma_2^2}}\right)$$

设  $d = \frac{\mu_1 - \mu_2}{\sqrt{\sigma_1^2 + \sigma_2^2}}$ , 则:

$$P(Z > 0) = P(W > -d) = 1 - P(W \leq -d) = P(W \leq d).$$

$$P(X > Y) = \Phi\left(\frac{\mu_1 - \mu_2}{\sqrt{\sigma_1^2 + \sigma_2^2}}\right)$$



给定二维随机向量  $(X, Y)$  的概率分布, 求

$$P(X = 2), P(Y \geq 2), P(X = Y), P(X \leq 2, Y \leq 2), P(X > Y).$$

给定二维随机向量  $(X, Y)$  的概率分布, 求

$P(X = 2), P(Y \geq 2), P(X = Y), P(X \leq 2, Y \leq 2), P(X > Y)$ 。

解:

$$P(X = 2) = \sum_{i=0}^4 P(X = 2, Y = i) = 0.27$$

$$P(Y \geq 2) = \sum_{i=2}^4 \sum_{j=0}^3 P(X = j, Y = i) = 0.69$$

$$P(X = Y) = \sum_{i=0}^3 P(X = i, Y = i) = 0.3$$

$$P(X \leq 2, Y \leq 2) = \sum_{i=0}^2 \sum_{j=0}^2 P(X = i, Y = j) = 0.69$$

$$P(X > Y) = \sum_{i=1}^3 \sum_{j=0}^{i-1} P(X = i, Y = j) = 0.25$$

$P(X = k) = pq^k \quad (k = 0, 1, \cdots; q = 1 - p)$ , 试证: (1)

 $\min\{X, Y\}$  与  $X - Y$  相互独立; (2)  $Z = \min\{X, Y\}$  与

$W = \max\{X, Y\} - Z$  相互独立。

**43.** 设随机变量  $X$  与  $Y$  独立同分布，共同分布是几何分布，即  $P(X = k) = pq^k$  ( $k = 0, 1, \dots; q = 1 - p$ )，试证：(1)  $\min\{X, Y\}$  与  $X - Y$  相互独立；(2)  $Z = \min\{X, Y\}$  与  $W = \max\{X, Y\} - Z$  相互独立。

证明：(1)

$$\begin{aligned} P(\min\{X, Y\} = k) &= P(X = k, Y \geq k) + P(Y = k, X \geq k) - P(X = k, Y = k) \\ &= 2pq^k \sum_{j=k}^{\infty} pq^j - (pq^k)^2 \\ &= 2pq^k \cdot \frac{pq^k}{p} - (pq^k)^2 = pq^{2k}(2 - p) \end{aligned}$$

考虑  $X - Y$  的概率分布：  $n > 0$ ,

$$P(X - Y = n) = \sum_{k=0}^{\infty} P(X = k+n, Y = k) = p^2 q^n \sum_{k=0}^{\infty} q^{2k} = p^2 q^n \cdot \frac{1}{1 - q^2}$$

对于  $n \leq 0$ ,

$$P(X-Y=n) = \sum_{k=-n}^{\infty} P(X=k+n, Y=k) = p^2 q^n \sum_{k=-n}^{\infty} q^{2k} = \frac{p^2 q^{-n}}{1-q^2}.$$

计算联合概率:

$$P(\min\{X, Y\} = k, X-Y=n) = \begin{cases} P(X=k+n, Y=k) = p^2 q^{2k+n}, & n > 0 \\ P(X=k, Y=k-n) = p^2 q^{2k-n}, & n \leq 0 \end{cases}$$

因此,

$$P(\min\{X, Y\} = k)P(X-Y=n) = P(\min\{X, Y\} = k, X-Y=n).$$

$\min\{X, Y\}$  与  $X-Y$  相互独立。

(2) 考虑  $X$  和  $Y$  的所有可能组合，计算  $W$  的概率分布

$$\begin{aligned} P(W = n) &= \sum_{k=0}^{\infty} P(X = k + n, Y = k) + P(X = k, Y = k + n) \\ &= \sum_{k=0}^{\infty} (p^2 q^{2k+n} + p^2 q^{2k+n}) = 2p^2 q^n \sum_{k=0}^{\infty} q^{2k} = 2p^2 q^n \cdot \frac{1}{1 - q^2} = \frac{2p^2 q^n}{1 - q^2}. \end{aligned}$$

联合概率：

$$P(Z = k, W = n) = P(X = k+n, Y = k) + P(X = k, Y = k+n) = 2p^2 q^{2k+n}$$

另一方面：

$$P(Z = k)P(W = n) = q^{2k}(2 - p^2) \cdot \frac{2p^2 q^n}{1 - q^2} = 2p^2 q^{2k+n}.$$

因此， $Z$  和  $W$  相互独立。



44. 设  $a$  是区间  $[0, 1]$  中的一个定点, 随机变量  $X$  服从  $[0, 1]$  上的均匀分布,  $Y = |X - a|$ . 问:  $a$  取何值时,  $X$  与  $Y$  不相关。

解:

$$E(X) = \int_0^1 x \, dx = \left[ \frac{x^2}{2} \right]_0^1 = \frac{1}{2}.$$

$$E(Y) = E(|X - a|) = \int_0^1 |x - a| \, dx = \int_0^a (a - x) \, dx + \int_a^1 (x - a) \, dx = \frac{a^2}{2} + \frac{1}{2}$$

$$\begin{aligned} E(XY) &= E(X|X - a|) = \int_0^1 x|x - a| \, dx \\ &= \int_0^a x(a - x) \, dx + \int_a^1 x(x - a) \, dx \\ &= \frac{a^3}{6} + \frac{1}{3} - \frac{a}{2} + \frac{a^3}{6} = \frac{a^3}{3} + \frac{1}{3} - \frac{a}{2}. \end{aligned}$$



协方差为

$$\text{Cov}(X, Y) = \left( \frac{a^3}{3} + \frac{1}{3} - \frac{a}{2} \right) - \left( \frac{1}{2} \right) \left( a^2 - a + \frac{1}{2} \right) = \frac{a^3}{3} - \frac{a^2}{2} + \frac{1}{12}.$$

为了使  $X$  和  $Y$  不相关, 需要:

$$\frac{a^3}{3} - \frac{a^2}{2} + \frac{1}{12} = 0.$$

解方程可得  $a = \frac{1}{2}$  是方程的一个根。

45. 设二维随机向量  $(X, Y)$  有联合密度

$$p(x, y) = \begin{cases} 3x, & 0 < x < 1 \text{ 且 } 0 < y < x, \\ 0, & \text{其他,} \end{cases}$$

试求条件分布密度

$p_{Y|X}(y|x)$  和  $p_{X|Y}(x|y)$ 。

45. 设二维随机向量  $(X, Y)$  有联合密度

$$p(x, y) = \begin{cases} 3x, & 0 < x < 1 \text{ 且 } 0 < y < x, \\ 0, & \text{其他,} \end{cases} \quad \text{试求条件分布密度}$$

$p_{Y|X}(y|x)$  和  $p_{X|Y}(x|y)$ 。

解：对于  $0 < x < 1$ ,  $0 < y < 1$ ,

$$p_X(x) = \int_{-\infty}^{\infty} p(x, y) dy = \int_0^x 3x dy = 3x \int_0^x 1 dy = 3x \cdot x = 3x^2.$$

$$p_Y(y) = \int_{-\infty}^{\infty} p(x, y) dx = \int_y^1 3x dx = 3 \int_y^1 x dx = \frac{3}{2}(1 - y^2).$$

条件分布密度为

$$p_{Y|X}(y|x) = \frac{p(x, y)}{p_X(x)} = \frac{3x}{3x^2} = \frac{1}{x} \quad (0 < y < x).$$

$$p_{X|Y}(x|y) = \frac{p(x, y)}{p_Y(y)} = \frac{3x}{\frac{3}{2}(1 - y^2)} = \frac{2x}{1 - y^2} \quad (y < x < 1).$$

46. 已知二维随机向量  $(X, Y)$  的联合密度为

$$p(x, y) = \begin{cases} \frac{21}{4}x^2y, & x^2 \leq y \leq 1, \\ 0, & \text{其他,} \end{cases}$$

求条件概率  $P(Y \geq 0.75 \mid X = 0.5)$ .

46. 已知二维随机向量  $(X, Y)$  的联合密度为

$$p(x, y) = \begin{cases} \frac{21}{4}x^2y, & x^2 \leq y \leq 1, \\ 0, & \text{其他,} \end{cases}$$

求条件概率  $P(Y \geq 0.75 \mid X = 0.5)$ .

解:

$$p_X(x) = \int_{-\infty}^{\infty} p(x, y) dy = \int_{x^2}^1 \frac{21}{4}x^2y dy = \frac{21}{4}x^2 \cdot \frac{1-x^4}{2} = \frac{21}{8}x^2(1-x^4).$$

$$p_{Y|X}(y|x) = \frac{p(x, y)}{p_X(x)} = \frac{\frac{21}{4}x^2y}{\frac{21}{8}x^2(1-x^4)} = \frac{2y}{1-x^4} \quad (x^2 \leq y \leq 1).$$

因此,  $P(Y \geq 0.75 \mid X = 0.5) = \int_{0.75}^1 p_{Y|X}(y \mid 0.5) dy$ .

代入  $x = 0.5$ :  $P(Y \geq 0.75 \mid X = 0.5) = \int_{0.75}^1 \frac{32y}{15} dy = \frac{7}{15}$ .

47. 设二维随机向量  $(X, Y)$  有联合密度

$$p(x, y) = \begin{cases} 24(1-x)y, & 0 < y < x < 1, \\ 0, & \text{其他,} \end{cases}$$

当  $0 < y < 1$  时, 试求条件期望  $E(X | Y = y)$ .

47. 设二维随机向量  $(X, Y)$  有联合密度

$$p(x, y) = \begin{cases} 24(1-x)y, & 0 < y < x < 1, \\ 0, & \text{其他,} \end{cases}$$

当  $0 < y < 1$  时, 试求条件期望  $E(X | Y = y)$ .

解:

$$p_Y(y) = \int_{-\infty}^{\infty} p(x, y) dx = \int_y^1 24(1-x)y dx = 12y(1-y)^2.$$

$$p_{X|Y}(x|y) = \frac{p(x, y)}{p_Y(y)} = \frac{24(1-x)y}{12y(1-y)^2} = \frac{2(1-x)}{(1-y)^2} \quad (y < x < 1).$$

条件期望为

$$\begin{aligned}
 E(X | Y = y) &= \int_{-\infty}^{\infty} x p_{X|Y}(x|y) dx = \int_y^1 x \cdot \frac{2(1-x)}{(1-y)^2} dx \\
 &= \frac{2}{(1-y)^2} \int_y^1 x(1-x) dx = \frac{2}{(1-y)^2} \cdot \frac{1-3y^2+2y^3}{6} = \frac{2y+1}{3}. (0 < y < 1)
 \end{aligned}$$



48. 设随机变量  $X$  与  $Y$  独立同分布, 都服从参数为  $\lambda$  的指数分布, 令

$$Z = \begin{cases} 3X + 1, & X \geq Y, \\ 6Y, & X < Y, \end{cases}$$

试求  $E(Z)$ .

48. 设随机变量  $X$  与  $Y$  独立同分布, 都服从参数为  $\lambda$  的指数分布, 令

$$Z = \begin{cases} 3X + 1, & X \geq Y, \\ 6Y, & X < Y, \end{cases}$$

试求  $E(Z)$ .

$$\begin{aligned} E(Z) &= E(3X + 1)I(X \geq Y) + E(6Y)I(X < Y) \\ &= \int_{x \geq y > 0} (3x + 1)\lambda^2 e^{-\lambda x} e^{-\lambda y} dx dy + \int_{0 < x < y} 6y\lambda^2 e^{-\lambda x} e^{-\lambda y} dx dy \\ &= \int_0^{+\infty} (3x + 1)\lambda^2 e^{-\lambda x} \int_0^x e^{-\lambda y} dy dx + \int_0^{+\infty} 6y\lambda^2 e^{-\lambda y} \int_0^y e^{-\lambda x} dx dy \\ &= \int_0^{+\infty} (3x + 1)\lambda e^{-\lambda x} (1 - e^{-\lambda x}) dx + \int_0^{+\infty} 6y\lambda e^{-\lambda y} (1 - e^{-\lambda y}) dy \\ &= \int_0^{+\infty} (9x + 1)\lambda e^{-\lambda x} (1 - e^{-\lambda x}) dx = \frac{27}{4\lambda} + \frac{1}{2}. \end{aligned}$$

49. 设随机变量  $X_1, \dots, X_n (n \geq 2)$  都是随机变量, 它们的期望、方差都存在, 试证协方差矩阵:

$$\Sigma = (\sigma_{ij})_{n \times n}, \quad \sigma_{ij} \triangleq \text{cov}(X_i, X_j)$$

是非负定的, 即对一切实数  $t_1, \dots, t_n$ , 均有  $\sum_{i,j} \sigma_{ij} t_i t_j \geq 0$ .

49. 设随机变量  $X_1, \dots, X_n (n \geq 2)$  都是随机变量, 它们的期望、方差都存在, 试证协方差矩阵:

$$\Sigma = (\sigma_{ij})_{n \times n}, \quad \sigma_{ij} \triangleq \text{cov}(X_i, X_j)$$

是非负定的, 即对一切实数  $t_1, \dots, t_n$  均有  $\sum_{i,j} \sigma_{ij} t_i t_j \geq 0$ .

$$\begin{aligned} \sum_{i,j} \sigma_{ij} t_i t_j &= \sum_{i,j} (E(X_i X_j) - E(X_i)E(X_j)) t_i t_j \\ &= \sum_{i,j} (E(t_i X_i t_j X_j) - E(t_i X_i)E(t_j X_j)) \end{aligned}$$

$$= E\left(\sum_{i,j} t_i X_i t_j X_j\right) - [E(\sum_i t_i X_i)]^2$$

$$= E\left(\left(\sum_i t_i X_i\right)^2\right) - [E(\sum_i t_i X_i)]^2 = \text{var}(\sum_i t_i X_i) \geq 0$$

51. 设随机变量  $X$  与  $U$  相互独立,  $X$  的分布密度是  $p_x(x)$ ,  $U$  服从区间  $[0, 1]$  上的均匀分布, 又函数  $q(x)$  满足:

①  $q(x) \geq 0$ , 且  $\int_{-\infty}^{+\infty} q(x) dx = 1$ ;

② 存在  $a > 0$ , 使得  $p_x(x)/q(x) \geq a$  (当  $q(x) > 0$  时);

令  $r(x) = a \frac{q(x)}{p_x(x)}$  (当  $p_x(x) = 0$  时, 规定  $r(x) = 0$ ),

$$M = \{U \leq r(X)\},$$

试证明:  $P(X \leq z | M) = \int_{-\infty}^z q(x) dx$ ,

即在  $M$  发生的条件下  $X$  的条件分布密度恰好是  $q(x)$ .

证明：

$$\begin{aligned}
 P(X \leq z|M) &= \frac{P(\{x \leq z, U \leq r(x)\})}{P(\{U \leq r(x)\})} \\
 &= \frac{\int_{-\infty}^{+\infty} P(X \leq z, U \leq r(X)|X = x)p(x)dx}{\int_{-\infty}^{+\infty} P(U \leq r(X)|X = x)p(x)dx} \\
 &= \frac{\int_{-\infty}^z P(U \leq r(x))p(x)dx}{\int_{-\infty}^{+\infty} P(U \leq r(x))p(x)dx} \\
 &= \frac{\int_{-\infty}^z r(x)p(x)dx}{\int_{-\infty}^{+\infty} r(x)p(x)dx} \\
 &= \frac{\int_{-\infty}^z aq(x)dx}{\int_{-\infty}^{+\infty} aq(x)dx} \\
 &= \int_{-\infty}^z q(x)dx
 \end{aligned}$$

- 1 第三章作业
- 2 第二部分
- 3 多元正态分布

$$f_X(x) = \frac{1}{\sqrt{2\pi\sigma_X^2}} \exp\left(-\frac{(x - \mu_X)^2}{2\sigma_X^2}\right) = \frac{1}{2\sqrt{2\pi}} \exp\left(-\frac{(x - 1)^2}{8}\right).$$



### (2)(3) 相关系数

$$\rho_{XY} = \frac{\text{Cov}(X, Y)}{\sigma_X \sigma_Y} = \frac{3}{2 \cdot 3} = 0.5.$$

由协方差矩阵  $\Sigma$  ,  $\sigma_X^2 = 4$  ,  $\sigma_Y^2 = 9$  , 条件期望

$$E(Y|X = x) = \mu_Y + \rho \frac{\sigma_Y}{\sigma_X}(x - \mu_X) = 2 + \rho \frac{3}{2}(x - 1) = \frac{3}{4}x + \frac{5}{4}.$$



$$\begin{aligned}
& \int p(\mathbf{x})(\mathbf{x} - \mu_{\mathbf{X}})^{\top} \Sigma_{\mathbf{X}}^{-1} (\mathbf{x} - \mu_{\mathbf{X}}) d\mathbf{x} \\
&= E_{\mathbf{X}} \begin{bmatrix} x_1 - 1 \\ x_2 \end{bmatrix}^{\top} \begin{bmatrix} 5 & -2 \\ -2 & 1 \end{bmatrix} \begin{bmatrix} x_1 - 1 \\ x_2 \end{bmatrix} \\
&= 5E_{\mathbf{X}}((x_1 - 1)^2) - 4E_{\mathbf{X}}((x_1 - 1)x_2) + E_{\mathbf{X}}(x_2^2) = 2
\end{aligned}$$

$$\begin{aligned}
& \int p(\mathbf{x})(\mathbf{x} - \mu_{\mathbf{Y}})^{\top} \Sigma_{\mathbf{Y}}^{-1} (\mathbf{x} - \mu_{\mathbf{Y}}) d\mathbf{x} \\
&= E_{\mathbf{X}} \begin{bmatrix} x_1 \\ x_2 - 1 \end{bmatrix}^{\top} \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 - 1 \end{bmatrix} \\
&= E_{\mathbf{X}}(x_1^2) + E_{\mathbf{X}}((x_2 - 1)^2) = 8
\end{aligned}$$

将上述结果代入并计算：

$$E_{\mathbf{X}} \log \frac{p(\mathbf{X})}{q(\mathbf{X})} = 0 - \frac{1}{2} \cdot 2 + \frac{1}{2} \cdot 8 = 3$$

3. 若随机向量  $\begin{bmatrix} \xi \\ \eta \end{bmatrix} \sim N\left(\begin{bmatrix} a \\ b \end{bmatrix}, \begin{bmatrix} 1 & \rho \\ \rho & 1 \end{bmatrix}\right)$ , 试求  $P(\xi \geq a, \eta \geq b)$ 。

解: 设  $X = \xi - a$ ,  $Y = \eta - b$ , 则

$$\begin{bmatrix} X \\ Y \end{bmatrix} \sim N \left( \begin{bmatrix} 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 & \rho \\ \rho & 1 \end{bmatrix} \right),$$

$$P(\xi \geq a, \eta \geq b) = \int_0^{+\infty} \int_0^{+\infty} \frac{1}{2\pi\sqrt{1-\rho^2}} \exp \left\{ -\frac{1}{2(1-\rho^2)}(x^2 + y^2 - 2\rho xy) \right\} dx dy.$$

作极坐标变换  $x = r \sin \theta$ ,  $y = r \cos \theta$ , 雅可比矩阵

$$J = \left| \frac{\partial(x, y)}{\partial(r, \theta)} \right| = r$$

$$\begin{aligned}
& P(\xi \geq a, \eta \geq b) \\
&= \int_0^{+\infty} \int_0^{\frac{\pi}{2}} \frac{1}{2\pi\sqrt{1-\rho^2}} r \exp \left\{ -\frac{1}{2(1-\rho^2)} (r^2 - 2\rho r^2 \sin \theta \cos \theta) \right\} d\theta dr \\
&= \frac{1}{4\pi\sqrt{1-\rho^2}} \int_0^{\frac{\pi}{2}} \int_0^{+\infty} \exp \left\{ -\frac{1}{2(1-\rho^2)} r^2 (1 - \rho \sin 2\theta) \right\} dr^2 d\theta \\
&= \frac{1}{4\pi\sqrt{1-\rho^2}} \int_0^{\frac{\pi}{2}} -\frac{2(1-\rho^2)}{1-\rho \sin 2\theta} d\theta \\
&= \frac{\sqrt{1-\rho^2}}{2\pi} \int_0^{\frac{\pi}{2}} -\frac{1}{\sin^2 \theta + \cos^2 \theta - \rho \sin 2\theta} d\theta \\
&= \frac{\sqrt{1-\rho^2}}{2\pi} \int_0^{\frac{\pi}{2}} -\frac{1}{\tan^2 \theta + 1 - 2\rho \tan \theta} d \tan \theta
\end{aligned}$$

$$\begin{aligned} &= \frac{\sqrt{1-\rho^2}}{2\pi} \int_0^{\frac{\pi}{2}} -\frac{1}{1-\rho^2} \frac{1}{\left(\frac{\tan\theta-\rho}{\sqrt{1-\rho^2}}\right)^2+1} d\tan\theta \\ &= \frac{\sqrt{1-\rho^2}}{2\pi} \int_{-\frac{\rho}{\sqrt{1-\rho^2}}}^{+\infty} -\frac{1}{1-\rho^2} \frac{1}{t^2+1} \sqrt{1-\rho^2} dt \\ &= \frac{1}{2\pi} \left( \frac{\pi}{2} + \arctan\left(\frac{\rho}{\sqrt{1-\rho^2}}\right) \right) \end{aligned}$$

4. 已知  $X \sim N(\mathbf{0}, I_2)$ , 向量  $\mathbf{u}, \mathbf{v} \in \mathbb{R}^2$  满足  $\|\mathbf{u}\| = \|\mathbf{v}\| = 1$ , 求:

(1)  $\text{cov}(\langle \mathbf{u}, \mathbf{x} \rangle, \langle \mathbf{v}, \mathbf{x} \rangle)$ ;

(2) (选做)  $E \text{sign}(\langle \mathbf{u}, \mathbf{x} \rangle) \text{sign}(\langle \mathbf{v}, \mathbf{x} \rangle)$ , 这里符号函数满足  $\text{sign}$

$$(X) = \begin{cases} 1, & x \geq 0, \\ -1, & x < 0. \end{cases}$$

解: (1)  $E[\langle \mathbf{u}, \mathbf{x} \rangle] = E[\mathbf{u}^\top \mathbf{x}] = \mathbf{u}^\top E[\mathbf{x}] = \mathbf{u}^\top \mathbf{0} = 0$ , 同理

$$E[\langle \mathbf{v}, \mathbf{x} \rangle] = 0.$$

所以,

$$\text{cov}(\langle \mathbf{u}, \mathbf{x} \rangle, \langle \mathbf{v}, \mathbf{x} \rangle) = E[\langle \mathbf{u}, \mathbf{x} \rangle \langle \mathbf{v}, \mathbf{x} \rangle] = E[(\mathbf{u}^\top \mathbf{x})(\mathbf{v}^\top \mathbf{x})] = E[\mathbf{u}^\top (\mathbf{x} \mathbf{x}^\top) \mathbf{v}]$$

由于  $\mathbf{x} \sim N(\mathbf{0}, I_2)$ , 我们有  $E[\mathbf{x} \mathbf{x}^\top] = I_2$ , 因此:

$$\text{cov}(\langle \mathbf{u}, \mathbf{x} \rangle, \langle \mathbf{v}, \mathbf{x} \rangle) = \mathbf{u}^\top I_2 \mathbf{v} = \mathbf{u}^\top \mathbf{v} = \mathbf{u} \cdot \mathbf{v}$$

(2) 设  $\xi = \langle \mathbf{u}, \mathbf{x} \rangle$ ,  $\eta = \langle \mathbf{v}, \mathbf{x} \rangle$ , 则方差  $\text{var}(\xi) = \mathbf{u}^\top I_2 \mathbf{u} = \|\mathbf{u}\| = 1$ ,  
 $\text{var}(\eta) = \mathbf{v}^\top I_2 \mathbf{v} = \|\mathbf{v}\| = 1$ , 相关系数  
 $\rho = \text{cov}(\langle \mathbf{u}, \mathbf{x} \rangle, \langle \mathbf{v}, \mathbf{x} \rangle) = \mathbf{u} \cdot \mathbf{v}$ , 因此

$$\begin{bmatrix} \xi \\ \eta \end{bmatrix} \sim N\left(\begin{bmatrix} 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 & \rho \\ \rho & 1 \end{bmatrix}\right)$$

利用第 3 题的结果, 有

$$P(\xi \geq 0, \eta \geq 0) = P(\xi < 0, \eta < 0) = \frac{1}{2\pi} \left( \frac{\pi}{2} + \arctan \left( \frac{\rho}{\sqrt{1-\rho^2}} \right) \right)$$

类似的,

$$P(\xi < 0, \eta \geq 0) = P(\xi \geq 0, \eta < 0) = \frac{1}{2\pi} \left( \frac{\pi}{2} - \arctan \left( \frac{\rho}{\sqrt{1-\rho^2}} \right) \right)$$



因此

$$\begin{aligned}
 & E\text{sign}(\langle \mathbf{u}, \mathbf{x} \rangle) \text{sign}(\langle \mathbf{v}, \mathbf{x} \rangle) \\
 &= P(\xi \geq 0, \eta \geq 0) + P(\xi < 0, \eta < 0) - P(\xi < 0, \eta \geq 0) - P(\xi \geq 0, \eta < 0) \\
 &= \frac{2}{\pi} \arctan \left( \frac{\rho}{\sqrt{1 - \rho^2}} \right) = \frac{2}{\pi} \arctan \left( \frac{\mathbf{u} \cdot \mathbf{v}}{\sqrt{1 - (\mathbf{u} \cdot \mathbf{v})^2}} \right)
 \end{aligned}$$

其中若  $\mathbf{u} \cdot \mathbf{v} = 1$ ，投影方向完全相同，

$E\text{sign}(\langle \mathbf{u}, \mathbf{x} \rangle) \text{sign}(\langle \mathbf{v}, \mathbf{x} \rangle) = 1$ ，若  $\mathbf{u} \cdot \mathbf{v} = -1$ ，投影方向完全相反，

$E\text{sign}(\langle \mathbf{u}, \mathbf{x} \rangle) \text{sign}(\langle \mathbf{v}, \mathbf{x} \rangle) = -1$ 。