

Ch05: 多维随机变量及其数字特征

Operations on Multi-dimensional Random Variables – 1

October 10, 2026

引言

二维随机变量的函数的分布:

- 即已知二维随机向量 (X, Y) 的概率分布, 求随机变量 (X, Y) 的函数 $Z = g(X, Y)$ 的概率分布
- 分离散和连续两种情况讨论

离散型随机变量函数的分布

已知离散型随机向量 (X, Y) 的联合分布列, 求随机变量 $Z = g(X, Y)$ 的分布列:

- 根据 X, Y 的各种取值, 计算随机变量 Z 的取值;

$\begin{array}{c} \text{Y} \\ \diagdown \\ \text{X} \end{array}$	y_1	y_2	$\dots$	y_n
x_1	$X = x_1, Y = y_1$	$X = x_1, Y = y_2$	$\dots$	$X = x_1, Y = y_n$
x_2	$X = x_2, Y = y_1$	$X = x_2, Y = y_2$	$\dots$	$X = x_2, Y = y_n$
$\vdots$	$\vdots$	$\vdots$	$\ddots$	$\vdots$
x_m	$X = x_m, Y = y_1$	$X = x_m, Y = y_2$	$\dots$	$X = x_m, Y = y_n$

- 对相同的 Z 值合并, 对应的概率相加.

举例：离散型随机变量函数的分布

求随机变量 $Z = g(X, Y) = X * Y$ 的分布列:

- X, Y 的分布列

$X \backslash Y$	Y			
	1	2	3	4
1	p_{11}	p_{12}	p_{13}	p_{41}
2	p_{21}	p_{22}	p_{23}	p_{42}
3	p_{31}	p_{23}	p_{33}	p_{43}
4	p_{41}	p_{24}	p_{34}	p_{44}

- 对相同的 Z 值合并, 对应的概率相加

Z	1	2	3	4	6	8	9	12	16
P	p_{11}	$p_{12} + p_{21}$	$\dots$	$p_{14} + p_{41} + p_{22}$	$\dots$	$p_{24} + p_{42}$	$\dots$	$\dots$	$\dots$

二维离散型随机变量和 (和函数 $g = X + Y$) 的分布

定理 0.18 [卷积公式 – 和函数] 若 X 与 Y 为离散随机变量 X 与 Y , 其取值 $(i, j \in \mathbb{N}^+$, 则随机变量 $Z = X + Y$ 的分布列为

$$P(Z = X + Y = k) = \sum_{i=1}^{k-1} P(X = i, Y = k - i) .$$

Remarks:

- 将 $k = i + j$ 的项加在一起.
- $P(Z = X + Y = k) = P(X = 1, Y = k - 1) + \cdots + P(X = k - 1, Y = 1)$
- 比如: $k = 3$, 有 $P(Z = 3) = P(X = 1)P(Y = 2) + P(X = 2)P(Y = 1) = a_1b_2 + a_2b_1$.

二维离散型随机变量和 (和函数 $g = X + Y$) 的分布

定理 0.19 [卷积公式 – 和函数、独立性] 若离散随机变量 X 与 Y 独立, 且其分布列为

$$a_i = P(X = i) \quad \text{和} \quad b_j = P(Y = j) \quad (i, j \in \mathbb{N}^+)$$

则随机变量 $Z = X + Y$ 的分布列为

$$P(Z = X + Y = k) = \sum_{i=1}^{k-1} a_i b_{k-i} .$$

Remarks:

- 独立性: $P(X = i, Y = j) = P(X = i)P(Y = j)$.

命题一：二项分布之和是二项分布

定理 0.20 若随机变量 $X \sim B(n_1, p)$ 和 $Y \sim B(n_2, p)$ 独立, 则

$$Z = X + Y \sim B(n_1 + n_2, p)$$

Remarks of Proof: 二项分布 $P(X = i) = \binom{n_1}{i} p^i (1 - p)^{n_1 - i} (i \in \mathbb{N})$.

$$\begin{aligned} P(Z = k) &= \sum_{i=0}^k P(X = i) P(Y = k - i) \\ &= \sum_{i=0}^k \binom{n_1}{i} p^i (1 - p)^{n_1 - i} \binom{n_2}{k - i} p^{k - i} (1 - p)^{n_2 - (k - i)} \\ &= \left[\sum_{i=0}^k \binom{n_1}{i} \binom{n_2}{k - i} \right] p^k (1 - p)^{n_1 + n_2 - k} \\ &= \binom{n_1 + n_2}{k} p^k (1 - p)^{n_1 + n_2 - k} \end{aligned}$$

命题一：二项分布之和是二项分布

推论 (多变量推论: 二项分布) 若相互独立的随机变量 $X_i \sim B(m_i, p)$ ($i \in \mathbb{N}^+$), 则

$$X = \sum_{i=1}^n X_i \sim B\left(\sum_{i=1}^n m_i, p\right)$$

推论 (多变量推论: 二项分布特例) 若相互独立的随机变量 $X_i \sim B(1, p)$ ($i \in \mathbb{N}^+$), 则

$$X = \sum_{i=1}^n X_i \sim B(n, p)$$

命题二：泊松分布之和是泊松分布

定理 **0.21** 若随机变量 $X \sim P(\lambda_1)$ 和 $Y \sim P(\lambda_2)$ 独立, 则

$$Z = X + Y \sim P(\lambda_1 + \lambda_2)$$

Remarks of Proof: 泊松分布 $P(X = k) = \frac{\lambda^k}{k!} e^{-\lambda}$ ($k \in \mathbb{N}$).

$$\begin{aligned} P(Z = k) &= \sum_{i=0}^k P(X = i)P(Y = k - i) = \sum_{i=0}^k \binom{n_1}{i} \frac{\lambda_1^i}{i!} e^{-\lambda_1} \frac{\lambda_2^{k-i}}{(k-i)!} e^{-\lambda_2} \\ &= \left[\frac{1}{k!} \sum_{i=0}^k \binom{k}{i} \lambda_1^i \lambda_2^{k-i} \right] e^{-\lambda_1 - \lambda_2} \\ &= \frac{e^{-(\lambda_1 + \lambda_2)}}{k!} (\lambda_1 + \lambda_2)^k \end{aligned}$$

连续型随机向量函数的分布

设二维连续型随机向量 (X, Y) 的联合概率密度为 $f(x, y)$, 求随机变量 $Z = g(X, Y)$ 的概率密度:

- 先计算分布函数 (积分区域)

$$F_Z(z) = P(Z \leq z) = P(g(x, y) \leq z) = \int \int_{g(x, y) \leq z} f(x, y) \, dx \, dy$$

- 对分布函数 $F_Z(z)$ 求导得到密度函数

$$f_Z(z) = F'_Z(z)$$

连续型随机向量函数的分布：例 0.108

例 0.108 设随机变量 X 与 Y 相互独立、且服从标准正态分布, 求下列随机变量的密度函数.

- $Z_1 = \sqrt{X^2 + Y^2}$

- $Z_2 = X^2 + Y^2$

解答：例 0.108

题目：设随机变量 X 与 Y 相互独立、且服从标准正态分布，求随机变量 $Z_1 = \sqrt{X^2 + Y^2}$ 和 $Z_2 = X^2 + Y^2$ 的密度函数。

解答：

- 由独立性有 (X, Y) 的联合概率密度为

$$f(x, y) = f_X(x)f_Y(y) = e^{-(x^2+y^2)/2}/2\pi \quad (x, y \in \mathbb{R})$$

- 当 $z_1 \leq 0$ 时有 $F_{Z_1}(z_1) = 0$ ；当 $z_1 > 0$ 时有

$$F_{Z_1}(z_1) = P(Z_1 \leq z_1) = P(\sqrt{X^2 + Y^2} \leq z_1) = \int \int_{X^2+Y^2 \leq z_1^2} e^{-(x^2+y^2)/2}/2\pi \, dx \, dy$$

利用极坐标积分变换 $x = r \cos \theta$ 和 $y = r \sin \theta$ 有

$$F_{Z_1}(z_1) = \int_0^{2\pi} \int_0^{z_1} \frac{r}{2\pi} e^{-r^2/2} \, d\theta \, dr = \int_0^{z_1} r e^{-r^2/2} \, dr = 1 - e^{-z_1^2/2}$$

- 对分布函数 $F_{Z_1}(z_1)$ 求导得到密度函数

$$f_{Z_1}(z_1) = \begin{cases} z_1 e^{-z_1^2/2}, & z_1 > 0 \\ 0, & z_1 \leq 0 \end{cases}$$

上述分布称为瑞利分布 (Rayleigh distribution), 该分布常用于通信等领域.

- 同理可证随机变量 $Z_2 \sim E(\frac{1}{2})$, 即

$$f_{Z_2}(z_2) = \begin{cases} \frac{1}{2} e^{-z_2/2}, & z_2 > 0 \\ 0, & z_2 \leq 0 \end{cases}$$

上述分布称为卡方分布 (chi-square distribution, OR χ^2 -distribution).

二维连续型随机向量和 (和函数 $g = X + Y$) 的分布

定理 0.22 [求和基本公式] 设随机向量 (X, Y) 的联合密度为 $f(x, y)$, 则随机变量 $Z = X + Y$ 的分布函数为

$$\begin{aligned} F_Z(z) &= \iint_{x+y \leq z} f(x, y) \, dx \, dy \\ &= \int_{-\infty}^{+\infty} \left[\int_{-\infty}^{z-x} f(x, y) \, dy \right] dx \quad (\text{积分区域}) \\ &= \int_{-\infty}^{+\infty} \left[\int_{-\infty}^z f(x, u-y) \, du \right] dx \quad (\text{变量}) \\ &= \int_{-\infty}^z \left[\int_{-\infty}^{+\infty} f(x, u-y) \, dx \right] du \end{aligned}$$

因此, 概率密度为

$$f_Z(z) = \int_{-\infty}^{+\infty} f(x, z-x) \, dx = \int_{-\infty}^{+\infty} f(z-y, y) \, dy$$

二维连续型随机向量和的分布

定理 0.23 [卷积公式－和函数] 若连续随机变量 X 与 Y 相互独立, 其概率密度函数分别为 $f_X(x)$ 和 $f_Y(y)$, 则随机变量 $Z = X + Y$ 的密度函数为

$$f_Z(z) = \int_{-\infty}^{+\infty} f_X(x) f_Y(z - x) \, dx = \int_{-\infty}^{+\infty} f_X(z - y) f_Y(y) \, dy$$

推论: 更一般的情况, 对于同定义域上的函数 $f(x), g(x)$, 其卷积为

$$f * g(y) = \int_{-\infty}^{+\infty} f(x) g(y - x) \, dx$$

二维连续随机向量函数和的分布 – 均匀分布: 例 0.109

例 0.109 设随机变量 $X \sim U(0, 1)$ 和 $Y \sim U(0, 1)$ 相互独立, 求 $Z = X + Y$ 的概率密度.

解答：例 0.109

题目：设随机变量 $X \sim U(0, 1)$ 和 $Y \sim U(0, 1)$ 相互独立, 求 $Z = X + Y$ 的概率密度.

解答:

- 根据卷积公式有

$$f_Z(z) = \int_{-\infty}^{+\infty} f_X(x) f_Y(z - x) dx$$

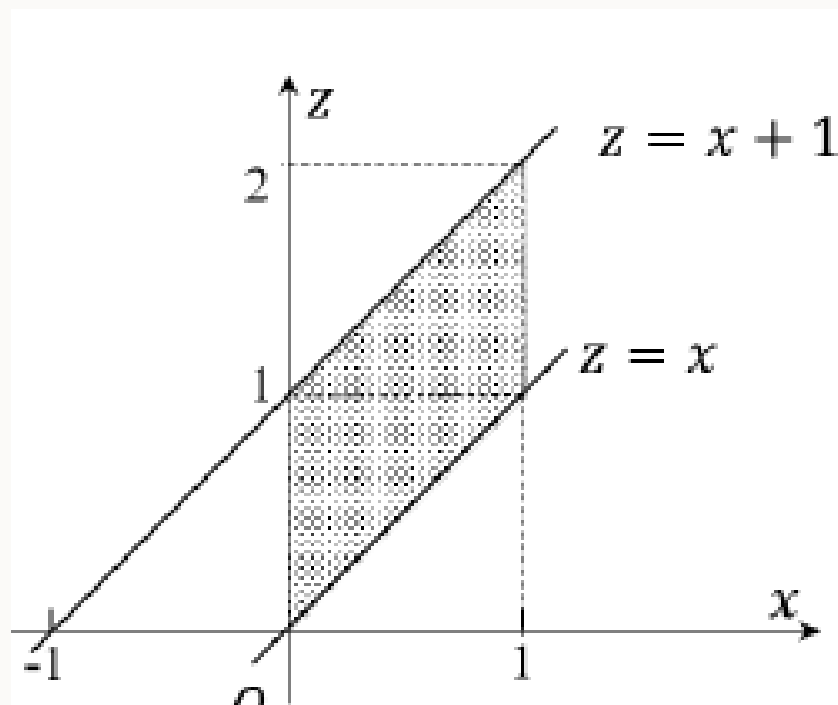
分析 $f_X(x)$, $f_Y(z - x)$ 的取值情况可知:

- 当 $x \in [0, 1]$ 时, 有 $f_X(x) = 1$
- 当 $z - x \in [0, 1]$ 时, 有 $f_Y(z - x) = 1$

因此, 积分区域为 $\{x \in [0, 1], z - x \in [0, 1]\}$.

- 由下图所示:

- 当 $z \leq 0$ 或 $z \geq 2$ 时, 有 $f_Z(z) = 0$
- 当 $z \in (0, 1)$ 时, 有 $f_Z(z) = \int_0^z 1 dz = z$
- 当 $z \in [1, 2)$ 时, 有 $f_Z(z) = \int_{z-1}^1 1 dx = 2 - z$



• $Z = X + Y$ 的概率密度为

$$f_Z(z) = \begin{cases} z, & z \in [0, 1] \\ 2 - z, & z \in [1, 2] \\ 0, & \text{其它} \end{cases}$$

二维连续随机向量函数和的分布 – 指数分布：例 0.110

例 0.110 设随机变量 $X \sim e(\lambda)$ 和 $Y \sim e(\lambda)$ 相互独立, 求 $Z = X + Y$ 的分布函数和概率密度.

解答：例 0.110

题目：设随机变量 $X \sim e(\lambda)$ 和 $Y \sim e(\lambda)$ 相互独立, 求 $Z = X + Y$ 的分布函数和概率密度.

解答:

- 根据卷积公式有

$$f_Z(z) = \int_{-\infty}^{+\infty} f_X(x) f_Y(z-x) dx$$

分析 $f_X(x)$, $f_Y(z-x)$ 的取值情况可知,

- 当 $x \geq 0$ 时, 有 $f_X(x) = \lambda \exp(-\lambda x)$;
- 当 $z-x \geq 0$ 时, 有 $f_Y(z-x) = \lambda \exp(-\lambda(z-x))$,

因此积分区域为 $\{x \in [0, +\infty], z-x \in [0, +\infty]\}$.

- 当 $z \geq 0$ 时有

$$f_Z(z) = \lambda^2 \int_0^z \exp(-\lambda x) \exp(-\lambda(z-x)) dx = \lambda^2 z \exp(-\lambda z)$$

命题三：正态分布之和是正态分布

定理 0.24 若随机变量 $X \sim \mathcal{N}(\mu_1, \sigma_1^2)$ 和 $Y \sim \mathcal{N}(\mu_2, \sigma_2^2)$ 相互独立, 则

$$X + Y \sim \mathcal{N}(\mu_1 + \mu_2, \sigma_1^2 + \sigma_2^2)$$

推论 若相互独立的随机变量 $X_i \sim \mathcal{N}(\mu_i, \sigma_i^2)$ ($i \in [n]$), 则

$$X = \sum_{i=1}^n X_i \sim \mathcal{N}\left(\sum_{i=1}^n \mu_i, \sum_{i=1}^n \sigma_i^2\right)$$

证明 若随机变量 $X \sim \mathcal{N}(\mu_x, \sigma_x^2)$ 和 $Y \sim \mathcal{N}(\mu_y, \sigma_y^2)$, 则根据正太分布的性质有

$$X' = X - \mu_x \sim \mathcal{N}(0, \sigma_x^2) \quad \text{和} \quad Y' = Y - \mu_y \sim \mathcal{N}(0, \sigma_y^2) .$$

因此只需证明 $Z = X' + Y' \sim \mathcal{N}(0, \sigma_1^2 + \sigma_2^2)$. 根据卷积公式有

$$\begin{aligned} f_Z(z) &= \int_{-\infty}^{+\infty} f_X(x) f_Y(z-x) dx = \frac{1}{2\pi\sigma_1\sigma_2} \int_{-\infty}^{+\infty} \exp\left(-\frac{x^2}{2\sigma_1^2} - \frac{(z-x)^2}{2\sigma_2^2}\right) dx \\ &= \frac{1}{2\pi\sigma_1\sigma_2} \int_{-\infty}^{+\infty} \exp\left(-\frac{\sigma_1^2 + \sigma_2^2}{2\sigma_1^2\sigma_2^2} \left(x - \frac{\sigma_1^2 z}{\sigma_1^2 + \sigma_2^2}\right)^2 - \frac{z^2}{2(\sigma_1^2 + \sigma_2^2)}\right) dx \\ &= \frac{\exp\left(-\frac{z^2}{2(\sigma_1^2 + \sigma_2^2)}\right)}{\sqrt{2\pi}\sqrt{\sigma_1^2 + \sigma_2^2}} \times \frac{\sqrt{\sigma_1^2 + \sigma_2^2}}{\sqrt{2\pi}\sigma_1\sigma_2} \int_{-\infty}^{+\infty} \exp\left(-\frac{\sigma_1^2 + \sigma_2^2}{2\sigma_1^2\sigma_2^2} \left(x - \frac{\sigma_1^2 z}{\sigma_1^2 + \sigma_2^2}\right)^2\right) dx \\ &= \frac{1}{\sqrt{2\pi}\sqrt{\sigma_1^2 + \sigma_2^2}} \exp\left(-\frac{z^2}{2(\sigma_1^2 + \sigma_2^2)}\right) , \end{aligned}$$

最后一个等式成立是因为正太分布的规范性.

举例：连续型随机向量函数的分布

例 0.111 设随机变量 $X \sim \mathcal{N}(0, 1)$ 与 $Y \sim \mathcal{N}(1, 1)$ 相互独立, 求下列随机变量的密度函数.

- $P(X + Y \leq 1)$

解答:

- 先求和函数 $Z = X + Y$ 的分布, 可以断言 Z 也服从正态分布, 并通过积分运算求得 $Z \sim \mathcal{N}(1, 2)$
- 求 $P(X + Y \leq 1) = P(Z \leq 1) = \Phi(\frac{1-1}{\sqrt{2}}) = \Phi(0) = \frac{1}{2}$

随机变量的乘/除法分布

前面提及了连续型随机变量的加法, 减法是同样的计算. 接下来我们介绍关于乘法/除法的计算.

定理 0.25 设随机向量 (X, Y) 的联合密度为 $f(x, y)$, 则随机变量 $Z = XY$ 的概率密度为

$$f_{XY}(z) = \int_{-\infty}^{+\infty} \frac{1}{|x|} f\left(x, \frac{z}{x}\right) dx$$

随机变量 $Z = Y/X$ 的概率密度为

$$f_{Y/X}(z) = \int_{-\infty}^{+\infty} |x| f(x, xz) dx$$

随机变量的乘/除法分布：证明

求证：

$$f_{XY}(z) = \int_{-\infty}^{+\infty} \frac{1}{|x|} f\left(x, \frac{z}{x}\right) dx$$

证明：对 $F_{XY}(z)$ 关于 z 求导即可，

$$\begin{aligned} F_{XY}(z) &= P(XY \leq z) \\ &= \int \int_{xy \leq z} f(x, y) \, dx \, dy = \int_{x < 0} \int_{y \geq z/x} f(x, y) \, dx \, dy + \int_{x \geq 0} \int_{y \leq z/x} f(x, y) \, dx \, dy \\ &= \int_{-\infty}^0 \left[\int_{z/x}^{+\infty} f(x, y) \, dy \right] dx + \int_0^{+\infty} \left[\int_{-\infty}^{z/x} f(x, y) \, dy \right] dx \\ &= \int_{-\infty}^0 \left[\int_z^{-\infty} \frac{1}{x} f\left(x, \frac{t}{x}\right) dt \right] dx + \int_0^{+\infty} \left[\int_{-\infty}^z \frac{1}{x} f\left(x, \frac{t}{x}\right) dt \right] dx \\ &= \int_{-\infty}^0 \left[\int_{-\infty}^z -\frac{1}{x} f\left(x, \frac{t}{x}\right) dt \right] dx + \int_0^{+\infty} \left[\int_{-\infty}^z \frac{1}{x} f\left(x, \frac{t}{x}\right) dt \right] dx \\ &= \int_{-\infty}^z \left[\int_{-\infty}^0 -\frac{1}{x} f\left(x, \frac{t}{x}\right) dx + \int_0^{+\infty} \frac{1}{x} f\left(x, \frac{t}{x}\right) dx \right] dt = \int_{-\infty}^z \left[\int_{-\infty}^{+\infty} \frac{1}{|x|} f\left(x, \frac{t}{x}\right) dx \right] dt \end{aligned}$$

命题四：连续随机向量函数的分布：例 0.112

例 0.112 若标准正态分布的随机变量 X 与 Y 相互独立, 证明: 则随机变量 $Z = Y/X$ 服从柯西分布.

Proof: 根据独立性和定理0.25有

$$\begin{aligned} f_Z(z) &= \int_{-\infty}^{+\infty} |x| f(x, xz) dx = \int_{-\infty}^{+\infty} |x| f_X(x) f_Y(xz) dx \\ &= \frac{1}{2\pi} \int_{-\infty}^{+\infty} |x| e^{-x^2(1+z^2)/2} dx = \frac{1}{\pi} \int_0^{+\infty} x e^{-x^2(1+z^2)/2} dx \\ &= \frac{1}{\pi} \left[-\frac{e^{-x^2(1+z^2)/2}}{1+z^2} \right]_0^{+\infty} = \frac{1}{\pi(1+z^2)} \end{aligned}$$

证毕.

最大值和最小值的分布

定理 0.26 设 $X_1, \dots, X_n$ 相互独立、分布函数为 $F_{X_1}(x_1), \dots, F_{X_n}(x_n)$, 则

- 随机变量 $Y = \max(X_1, \dots, X_n)$ 的分布函数为

$$F_Y(y) = F_{X_1}(y) F_{X_2}(y) \dots F_{X_n}(y)$$

- 随机变量 $Z = \min(X_1, \dots, X_n)$ 的概率密度为

$$F_Z(z) = 1 - [1 - F_{X_1}(z)] [1 - F_{X_2}(z)] \dots [1 - F_{X_n}(z)]$$

Proof Sketches:

- For max, $P(Y \leq y) = P(\max(X_1, \dots, X_n) \leq y) = P(X_1 \leq y, X_2 \leq y, \dots, X_n \leq y)$
- For min, $P(Z \leq z) = P(\min(X_1, \dots, X_n) \leq z) = 1 - P(\min(X_1, \dots, X_n) > z) = 1 - P(X_1 > z, X_2 > z, \dots, X_n > z)$

最大值和最小值的分布: 证明

- 随机变量 $Y = \max(X_1, \dots, X_n) \leq y$, 则需要每一个 $X_i \leq y$

$$\begin{aligned} F_Y(y) &= P(Y \leq y) = P(\max(X_1, \dots, X_n) \leq y) \\ &= P(X_1 \leq y, X_2 \leq y, \dots, X_n \leq y) \\ &= P(X_1 \leq y)P(X_2 \leq y) \dots P(X_n \leq y) \\ &= F_{X_1}(y) F_{X_2}(y) \dots F_{X_n}(y) \end{aligned}$$

- 随机变量 $Z = \min(X_1, \dots, X_n) \leq z$, 则需要每一个 $1 - X_i \geq y$

$$\begin{aligned} F_Z(z) &= P(Z \leq z) = P(\min(X_1, \dots, X_n) \leq z) = 1 - P(\min(X_1, \dots, X_n) > z) \\ &= 1 - P(X_1 > z, X_2 > z, \dots, X_n > z) \\ &= 1 - P(X_1 > z)P(X_2 > z) \dots P(X_n > z) \\ &= 1 - [1 - P(X_1 \leq z)] [1 - P(X_2 \leq z)] \dots [1 - P(X_n \leq z)] \\ &= 1 - [1 - F_{X_1}(z)] [1 - F_{X_2}(z)] \dots [1 - F_{X_n}(z)] \end{aligned}$$

最大值和最小值的分布: 推论

推论 若相互独立的随机变量 $X_1, \dots, X_n$ 其分布函数和密度函数分别为 $F(x)$ 和 $f(x)$, 即

$$F_{X_1}(x) = F_{X_2}(x) = \cdots = F_{X_n}(x) = F(x)$$

$$f_{X_1}(x) = f_{X_2}(x) = \cdots = f_{X_n}(x) = f(x)$$

则随机变量 $Y = \max(X_1, \dots, X_n)$ 的分布函数和密度函数分别为

$$F_Y(y) = (F(y))^n \quad \text{和} \quad f_Y(y) = n(F(y))^{n-1}f(y)$$

随机变量 $Z = \min(X_1, \dots, X_n)$ 的分布函数和密度函数分别为

$$F_Z(z) = 1 - (1 - F(z))^n \quad \text{和} \quad f_Z(z) = n(1 - F(z))^{n-1}f(z)$$

连续随机向量函数的分布：例 0.113 指数分布

例 0.113 若随机变量 X 与 Y 相互独立, 且 $X \sim e(\alpha)$ 和 $Y \sim e(\beta)$, 求随机变量 $Z_1 = \max(X, Y)$ 和 $Z_2 = \min(X, Y)$ 的概率密度.

解答：例 0.113

- 根据定理0.26有, 随机变量 Z_1 的分布函数为

$$F_{Z_1}(z_1) = F_X(z_1)F_Y(z_1) = \int_{-\infty}^{z_1} f_X(t) dt \int_{-\infty}^{z_1} f_Y(t) dt$$

当 $z_1 \leq 0$ 时有 $F_{Z_1}(z_1) = 0$; 当 $z_1 > 0$ 时

$$F_{Z_1}(z_1) = \int_{-\infty}^{z_1} f_X(t) dt \int_{-\infty}^{z_1} f_Y(t) dt = \int_{-\infty}^{z_1} \alpha e^{-\alpha t} dt \int_{-\infty}^{z_1} \beta e^{-\beta t} dt = (1 - e^{-\alpha z_1})(1 - e^{-\beta z_1})$$

等式两边对 z_1 求导可得其密度函数为

$$f_{Z_1}(z_1) = \begin{cases} \alpha e^{-\alpha z_1} + \beta e^{-\beta z_1} - (\alpha + \beta)e^{-(\alpha+\beta)z_1}, & z_1 > 0 \\ 0, & z_1 \leq 0 \end{cases}$$

- 同理可得, 随机变量 Z_2 的分布函数和密度函数分别为

$$F_{Z_2}(z_2) = \begin{cases} 1 - e^{-(\alpha+\beta)z_2}, & z_2 > 0 \\ 0, & z_2 \leq 0 \end{cases} \quad \text{和} \quad f_{Z_2}(z_2) = \begin{cases} (\alpha + \beta)e^{-(\alpha+\beta)z_2}, & z_2 > 0 \\ 0, & z_2 \leq 0 \end{cases}$$

连续型随机变量复合函数的联合分布函数

问题: 随机变量 (X, Y) 的联合密度为 $f(x, y)$, 设 (X, Y) 的函数

$$U = u(x, y), \quad V = v(x, y).$$

如何求 (U, V) 的联合分布? 这里二元函数 $u(\cdot, \cdot)$ 和 $v(\cdot, \cdot)$ 具有连续的偏导, 并满足

$$\begin{cases} u = u(x, y) \\ v = v(x, y) \end{cases} \quad \text{存在 (不要求唯一的) 反函数} \quad \begin{cases} x = x(u, v) \\ y = y(u, v) \end{cases}$$

连续型随机变量复合函数的联合分布函数

定理 0.27 设 $U = u(X, Y)$ 和 $V = v(X, Y)$ 有连续偏导, 反函数 $X = x(U, V)$ 和 $Y = y(U, V)$. 若 (X, Y) 的联合概率密度为 $f(x, y)$, 则 (U, V) 的联合密度为

$$f_{UV}(u, v) = f_{XY}(x(u, v), y(u, v))|J|,$$

其中 J 为变换的雅可比行列式, 即

$$J = \left| \frac{\partial(x, y)}{\partial(u, v)} \right| = \left| \frac{\partial(u, v)}{\partial(x, y)} \right|^{-1} = \begin{vmatrix} \frac{\partial u}{\partial x} & \frac{\partial u}{\partial y} \\ \frac{\partial v}{\partial x} & \frac{\partial v}{\partial y} \end{vmatrix}^{-1},$$

上述结论可推广到一般的 n 维随机变量.

连续型随机向量复合函数的分布：例 0.114

例 0.114 设 X 与 Y 是相互独立的标准正态分布随机变量. 证明: 有随机变量

$$R = X^2 + Y^2 \quad \text{与} \quad \theta = \arctan(Y/X)$$

相互独立, 且有 $R \sim e(1/2)$ 以及 $\theta \sim U(-\pi, \pi)$.

解答：例 0.114

题目：设 X 与 Y 是相互独立的标准正太分布随机变量, 证明: 有随机变量 $R^2 = X^2 + Y^2$ 与 $\theta = \arctan(Y/X)$ 相互独立, 且有 $R \sim e(1/2)$ 以及 $\theta \sim U(-\pi, \pi)$.

解答:

- 根据定理0.27列出雅可比行列式

$$J = \begin{vmatrix} \frac{\partial u}{\partial x} & \frac{\partial u}{\partial y} \\ \frac{\partial v}{\partial x} & \frac{\partial v}{\partial y} \end{vmatrix}^{-1} = \begin{vmatrix} 2x & 2y \\ \frac{-y}{x^2+y^2} & \frac{x}{x^2+y^2} \end{vmatrix}^{-1} = \frac{1}{2}.$$

- 由此可得 R 与 Θ 的联合分布为

$$f_{R \times \Theta}(r, \theta) = f_{X \times Y}(\sqrt{r} \cos \theta, \sqrt{r} \sin \theta) |J| = \frac{1}{4\pi} \exp(-r/2) = \frac{1}{2} \exp(-r/2) \times \frac{1}{2\pi}$$

由此可发现 $R \sim E(1/2)$ 以及 $\theta \sim U(0, 2\pi)$, R 与 Θ 相互独立.

Ch05: 多维随机变量及其数字特征

Operations on Multi-dimensional Random Variables – 2

October 10, 2026

引言

已知二维随机向量 (X, Y) 的概率分布, 求随机变量 (X, Y) 的函数 $Z = g(X, Y)$ 的概率分布, 分离散和连续两种情况讨论.

离散型随机变量函数的分布

已知离散型随机向量 (X, Y) 的联合分布列, 求随机变量 $Z = g(X, Y)$ 的分布列:

- 根据 X, Y 的各种取值, 计算随机变量 Z 的取值;

X \ Y	Y			
	y_1	y_2	$\dots$	y_n
x_1	$X = x_1, Y = y_1$	$X = x_1, Y = y_2$	$\dots$	$X = x_1, Y = y_n$
x_2	$X = x_2, Y = y_1$	$X = x_2, Y = y_2$	$\dots$	$X = x_2, Y = y_n$
$\vdots$	$\vdots$	$\vdots$	$\ddots$	$\vdots$
x_m	$X = x_m, Y = y_1$	$X = x_m, Y = y_2$	$\dots$	$X = x_m, Y = y_n$

- 对相同的 Z 值合并, 对应的概率相加.

离散型随机变量函数的分布——重要性质

- < 和函数 > $Z = X + Y$ 的分布列为 $P(Z = X + Y = k) = \sum_{i=1}^k a_i b_{k-i}$.
- 二项分布之和是二项分布, 即 $Z = X + Y \sim B(n_1 + n_2, p)$.
- 泊松分布之和是泊松分布, 即 $Z = X + Y \sim P(\lambda_1 + \lambda_2)$.

连续型随机向量函数的分布

设二维连续型随机向量 (X, Y) 的联合概率密度为 $f(x, y)$, 求随机变量 $Z = g(X, Y)$ 的概率密度:

- 先计算分布函数 (积分区域)

$$F_Z(z) = P(Z \leq z) = P(g(x, y) \leq z) = \int \int_{g(x, y) \leq z} f(x, y) \, dx \, dy$$

- 对分布函数 $F_Z(z)$ 求导得到密度函数

$$f_Z(z) = F'_Z(z)$$

连续型随机变量函数的分布——重要性质

- < 求和基本公式 > $Z = X + Y$ 的密度函数为

$$f_Z(z) = \int_{-\infty}^{+\infty} f(x, z-x) dx = \int_{-\infty}^{+\infty} f(z-y, y) dy$$

- < 和函数 + 独立性 > $Z = X + Y$ 的密度函数为

$$f_Z(z) = \int_{-\infty}^{+\infty} f_X(x)f_Y(z-x) dx = \int_{-\infty}^{+\infty} f_X(z-y)f_Y(y) dy$$

- 均匀分布: 例 0.109.

- 指数分布: 例 0.110.

- 正态分布之和是正态分布, 即 $X + Y \sim \mathcal{N}(\mu_1 + \mu_2, \sigma_1^2 + \sigma_2^2)$

- < 乘/除法 > $Z = XY$ 和 $Z = Y/X$ 的概率密度分别为

$$f_{XY}(z) = \int_{-\infty}^{+\infty} \frac{1}{|x|} f(x, \frac{z}{x}) dx \quad f_{Y/X}(z) = \int_{-\infty}^{+\infty} |x| f(x, xz) dx$$

- < 最大值/最小值 > $Y = \max(X_1, \dots, X_n)$ 和 $Z = \min(X_1, \dots, X_n)$ 的分布函数分别为

$$\begin{cases} F_Y(y) = F_{X_1}(y) F_{X_2}(y) \dots F_{X_n}(y) \\ F_Z(z) = 1 - [1 - F_{X_1}(z)] [1 - F_{X_2}(z)] \dots [1 - F_{X_n}(z)] \end{cases}$$

- < 复合函数 > $(u, v) = (u(x, y), v(x, y))$ **OR** $(x, y) = (x(u, v), y(u, v))$, 则有

$$f_{UV}(u, v) = f_{XY}(x(u, v), y(u, v))|J|,$$

其中, J 为变换的雅可比行列式, 即

$$J = \left| \frac{\partial(x, y)}{\partial(u, v)} \right| = \left| \frac{\partial(u, v)}{\partial(x, y)} \right|^{-1} = \begin{vmatrix} \frac{\partial u}{\partial x} & \frac{\partial u}{\partial y} \\ \frac{\partial v}{\partial x} & \frac{\partial v}{\partial y} \end{vmatrix}^{-1}.$$

例题补充：例 0.115

例 0.115 设 X_1, X_2, X_3 相互独立, 并且有相同的概率分布

$$P(X_i = 1) = p, \quad P(X_i = 0) = 1 - p, \quad i = 1, 2, 3$$

考虑变量

$$Y_1 = \begin{cases} 1, & \text{若 } X_1 + X_2 \text{ 为奇数,} \\ 0, & \text{若 } X_1 + X_2 \text{ 为偶数,} \end{cases}$$
$$Y_2 = \begin{cases} 1, & \text{若 } X_2 + X_3 \text{ 为奇数,} \\ 0, & \text{若 } X_2 + X_3 \text{ 为偶数,} \end{cases}$$

求乘积变量 $Y_1 Y_2$ 的概率分布.

解答：例 0.115

题目：如上所述.

解答：

• $Y_1, Y_2 \in \{0, 1\} \Rightarrow Y_1 Y_2 \in \{0, 1\}$, 原问题转化为“求 $P(Y_1 Y_2 = 1)$ 和 $P(Y_1 Y_2 = 0)$ ”.

• For the case of $Y_1 Y_2 = 1$,

$$\begin{aligned} P(Y_1 Y_2 = 1) &= P(Y_1 = 1, Y_2 = 1) \\ &= P(\text{若 } X_1 + X_2 \text{ 为奇数, 若 } X_2 + X_3 \text{ 为奇数}) \\ &= P(\{X_1 = 0, X_2 = 1, X_3 = 0\} \cup \{X_1 = 1, X_2 = 0, X_3 = 1\}) \quad \text{事件不相容} \\ &= P(\{X_1 = 0, X_2 = 1, X_3 = 0\}) + P(\{X_1 = 1, X_2 = 0, X_3 = 1\}) \\ &= P(X_1 = 0)P(X_2 = 1)P(X_3 = 0) + P(X_1 = 1)P(X_2 = 0)P(X_3 = 1) \\ &= p(1 - p)^2 + p^2(1 - p) . \end{aligned}$$

• For the case of $Y_1 Y_2 = 0$,

$$P(Y_1 Y_2 = 1) = 1 - P(Y_1 Y_2 = 0) = 1 - p(1 - p) .$$

例题补充：例 0.116

例 0.116 设系统 L 由两个相互独立的子系统 L_1 和 L_2 连接而成, 其连接的方式分别为 (1) 串联; (2) 并联. 设 L_1 和 L_2 的寿命分别为 X 和 Y , 已知它们的密度函数分别为

$$f_X(x) = \begin{cases} \alpha e^{-\alpha x}, & x > 0 \\ 0, & x \leq 0 \end{cases}$$

$$f_Y(y) = \begin{cases} \beta e^{-\beta y}, & y > 0 \\ 0, & y \leq 0 \end{cases}$$

其中 $\alpha > 0, \beta > 0$. 试分别就以上两种连接方式求系统 L 的寿命 Z 的密度函数.

解答：例 0.116

题目：如上所述.

解答：建模要点 – 串联和并联各代表何种运算？

- L_1 和 L_2 串联时, L 的寿命 $Z = \min(X, Y)$, 而 X 和 Y 的分布函数为

$$F_X(x) = \begin{cases} 1 - e^{-\alpha x}, & x > 0 \\ 0, & x \leq 0 \end{cases} \quad F_Y(y) = \begin{cases} 1 - e^{-\beta y}, & y > 0 \\ 0, & y \leq 0 \end{cases}$$

由此 Z 的分布函数为

$$F_Z(z) = 1 - (1 - F_X(x))(1 - F_Y(y)) = \begin{cases} 1 - e^{-(\alpha+\beta)z}, & z > 0 \\ 0, & z \leq 0 \end{cases}$$

于是 Z 的密度函数为

$$f_Z(z) = \begin{cases} (\alpha + \beta) e^{-(\alpha+\beta)z}, & z > 0 \\ 0, & z \leq 0 \end{cases}$$

- L_1 和 L_2 并联时, L 的寿命 $Z = \max(X, Y)$, 由此 Z 的分布函数为

$$F_Z(z) = F_X(x)F_Y(y) = \begin{cases} (1 - e^{-\alpha z})(1 - e^{-\beta z}), & z > 0 \\ 0, & z \leq 0 \end{cases}$$

于是 Z 的密度函数为

$$f_Z(z) = \begin{cases} \alpha e^{-\alpha z} + \beta e^{-\beta z} + (\alpha + \beta) e^{-(\alpha+\beta)z}, & z > 0 \\ 0, & z \leq 0 \end{cases}$$

例题补充：例 0.117

例 0.117 假设一电路装有 3 个同种电气元件, 其工作状态相互独立, 且无故障工作时间都服从参数为 $\lambda > 0$ 的指数分布 $e(\lambda)$. 当 3 个元件都无故障时, 电路正常工作, 否则整个电路都不能正常工作. 试求电路正常工作的时间 T 的概率分布.

解答：例 0.117

题目：如上所述.

解答：

- 解法一，以 $X_i (i = 1, 2, 3)$ 表示第 i 个电气元件无故障工作的时间，则 X_1, X_2, X_3 相互独立且同分布，其分布函数为

$$F(x) = \begin{cases} 1 - e^{-\lambda x}, & x > 0 \\ 0, & x \leq 0 \end{cases}$$

设 $G(t)$ 是 T 的分布函数, $T = \min\{X_1, X_2, X_3\}$. 当 $t \leq 0$ 时, $G(t) = 0$. 当 $t > 0$, 有

$$\begin{aligned} G(t) &= P\{T \leq t\} = 1 - \{T > t\} = 1 - P\{X_1 > t, X_2 > t, X_3 > t\} \\ &= 1 - P\{X_1 > t\}P\{X_2 > t\}P\{X_3 > t\} = 1 - [1 - F(t)]^3 \\ &= 1 - e^{-3\lambda x} \end{aligned}$$

得知 T 服从参数为 3λ 的指数分布:

$$G(t) = \begin{cases} 1 - e^{-3\lambda t}, & t > 0 \\ 0, & t \leq 0 \end{cases}$$

- 解法二, 本题也可以直接利用公式计算, 因为 X_1, X_2, X_3 相互独立且同分布, 而 $T = \min(X_1, X_2, X_3)$, 故

$$G(t) = 1 - [1 - F(t)]^3 = \begin{cases} 1 - e^{-3\lambda t}, & t > 0 \\ 0, & t \leq 0 \end{cases}$$

例题补充：例 0.118

例 0.118 假设一设备开机后无故障工作时间 X 服从指数分布, 平均无故障工作的时间 $\mathbb{E}(X)$ 为 5 小时. 设备定时开机. 出现故障时自动关机, 而在无故障的情况下工作 2 小时便关机. 试求该设备每次开机无故障工作的时间 Y 的分布函数 $F(y)$.

解答：例 0.118

题目：如上所述.

解答：

- 设 X 的分布参数为 λ . 由于 $\mathbb{E}(X) = \frac{1}{\lambda} = 5$, 可见 $\lambda = \frac{1}{5}$, 显然

$$Y = \min(X, 2)$$

对于 $y < 0$, $F(y) = 0$. 对于 $y \geq 2$, $F(y) = 1$. 设 $0 \leq y < 2$, 有

$$F(y) = P\{Y \leq y\} = P\{\min(X, 2) \leq y\}$$

于是 Y 的分布函数为

$$F(y) = \begin{cases} 0, & y < 0 \\ 1 - e^{-\frac{y}{5}}, & 0 \leq y < 2 \\ 1, & y \geq 2 \end{cases}$$

例题补充：例 0.119

例 0.119 设某班车起点站上客人数 X 服从参数为 λ ($\lambda > 0$) 的泊松分布, 每位乘客在中途下车的概率为 p ($0 < p < 1$), 且中途下车与否相互独立. 以 Y 表示在中途下车的人数. 求:

- (1) 在发车时有 n 个乘客的条件下, 中途有 m 人下车的概率;
- (2) 二维随机变量 (X, Y) 的概率分布.

解答：例 0.119

题目：如上所述.

解答：这是一个关于条件分布的例题.

- 对于第 1 小题, 有

$$P\{Y = m \mid X = n\} = C_n^m p^m (1 - p)^{n-m}, \quad 0 \leq m \leq n, n = 0, 1, 2, \dots$$

- 对于第 2 小题, 有

$$P\{X = n, Y = m\} = P\{Y = m \mid X = n\}P\{X = n\} = C_n^m p^m (1 - p)^{n-m} \cdot \frac{e^{-\lambda} \lambda^n}{n!}$$

其中, $0 \leq m \leq n, n = 0, 1, 2, \dots$.

例题补充：例 0.120

例 0.120 [作业加强题] 设某种型号的电子元件的寿命 (以小时计) 近似地服从 $\mathcal{N}(160, 20^2)$, 随机地选择 4 只. 求其中没有一只寿命小于 180 小时的概率.

解答：例 0.120

题目：如上所述.

解答：

- 随机取 4 只, 记其寿命分别为: X_1, X_2, X_3, X_4 . 这四个变量独立同分布, 即

$$X_i \sim \mathcal{N}(160, 20^2) \quad i = 1, 2, 3, 4$$

- 记 $X = \min\{X_1, X_2, X_3, X_4\}$.
- 所求问题可以转化为 $P(X \geq 180)$, 则有

$$\begin{aligned} P(X \geq 180) &= [1 - F(180)]^4 \\ &= \left[1 - \Phi\left(\frac{180 - 160}{20}\right)\right]^4 \\ &= (1 - 0.8413)^4 \\ &\approx 0.000634 \end{aligned}$$

多维随机变量的分布函数和密度函数

将二维随机向量及分布推广到多维随机向量, 二维与多维随机变量没有本质性的区别, 只是相关概念和结论的扩展.

定义 0.50 设 $(X_1, X_2, \dots, X_n)$ 为 n 维随机向量, 对任意实数 $x_1, x_2, \dots, x_n$, 函数

$$F(x_1, x_2, \dots, x_n) = P(X_1 \leq x_1, X_2 \leq x_2, \dots, X_n \leq x_n)$$

称为 n 维随机向量 $(X_1, X_2, \dots, X_n)$ 的分布函数, 或随机变量 $X_1, X_2, \dots, X_n$ 的联合分布函数. 若存在可积函数 $f(x_1, x_2, \dots, x_n)$, 使得对任意实数 $x_1, x_2, \dots, x_n$ 有

$$F(x_1, x_2, \dots, x_n) = \int_{-\infty}^{x_1} \int_{-\infty}^{x_2} \cdots \int_{-\infty}^{x_n} f(u_1, u_2, \dots, u_n) \mathrm{d}u_1 \mathrm{d}u_2 \cdots \mathrm{d}u_n,$$

则称 $(X_1, X_2, \dots, X_n)$ 为连续型随机向量, 称 $f(x_1, x_2, \dots, x_n)$ 为 n 维联合密度函数.

n 维联合密度函数的性质

- 非负性. 对任意实数 $x_1, x_2, \dots, x_n$ 有 $f(x_1, x_2, \dots, x_n) \geq 0$;
- 规范性.

$$\int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} \cdots \int_{-\infty}^{+\infty} f(u_1, u_2, \dots, u_n) \, du_1 \, du_2 \cdots du_n = 1$$

- 连续性. 若 $f(x_1, x_2, \dots, x_n)$ 在点 $(x_1, x_2, \dots, x_n)$ 处连续, 则有

$$\frac{\partial F(x_1, x_2, \dots, x_n)}{\partial x_1 \partial x_2 \cdots \partial x_n} = f(x_1, x_2, \dots, x_n)$$

- 有界区域可积. 设 G 是 n 维空间的一片区域, 则有

$$P((X_1, X_2, \dots, X_n) \in G) = \int \cdots \int_G f(u_1, u_2, \dots, u_n) \, du_1 \, du_2 \cdots du_n$$

多维随机变量的边缘分布函数和边缘密度函数

定义 0.51 n 维随机向量 $(X_1, X_2, \dots, X_n)$ 中任意 k ($k \leq n$) 个分量所构成的随机向量, 它的分布函数和密度函数被称为 k 维边缘分布函数和 k 维边缘密度函数.

例如, 随机向量 $(X_1, X_2, \dots, X_n)$ 前 k 维随机向量的边缘分布函数和边缘密度函数分布为

$$\begin{aligned} F_{X_1, X_2, \dots, X_k}(x_1, x_2, \dots, x_k) &= P(X_1 \leq x_1, X_2 \leq x_2, \dots, X_k \leq x_k) \\ &= \lim_{\substack{x_{k+1} \rightarrow +\infty \\ x_n \rightarrow +\infty}} F(x_1, x_2, \dots, x_k) \end{aligned}$$

$$f_{X_1, X_2, \dots, X_k}(x_1, x_2, \dots, x_k) = \int_{-\infty}^{+\infty} \cdots \int_{-\infty}^{+\infty} f(u_1, \dots, u_k, u_{k+1}, \dots, u_n) \, du_{k+1} \cdots du_n$$

多维随机变量的独立性

将二维随机向量及分布推广到多维随机向量, 二维与多维随机变量没有本质性的区别, 只是相关概念和结论的扩展.

定义 0.52 若随机向量 $(X_1, X_2, \dots, X_n)$ 的联合分布函数 $F(x_1, x_2, \dots, x_n)$ 满足

$$F(x_1, x_2, \dots, x_n) = F_{X_1}(x_1)F_{X_2}(x_2) \dots F_{X_n}(x_n)$$

则称 $X_1, X_2, \dots, X_n$ 相互独立. 若随机向量 $X = (X_1, X_2, \dots, X_m)$ 和 $Y = (Y_1, Y_2, \dots, Y_n)$ 的联合分布函数 $F(x_1, \dots, x_m, y_1, \dots, y_n)$ 满足

$$F(x_1, \dots, x_m, y_1, \dots, y_n) = F_X(x_1, \dots, x_m)F_Y(y_1, \dots, y_n)$$

则称随机向量 X 和 Y 相互独立.

多维正态分布

多维随机向量中最重要的常用是多维正态分布.

定义 0.53 给定一个向量 $\boldsymbol{\mu} \in \mathbb{R}^n$ 和正定矩阵 $\boldsymbol{\Sigma} \in \mathbb{R}^{n \times n}$, 对任意实数向量 $\boldsymbol{x} = (x_1, x_2, \dots, x_n)^\top$, 若随机向量 $X = (X_1, X_2, \dots, X_n)$ 的密度函数为

$$f(\boldsymbol{x}) = (2\pi)^{-n/2} |\boldsymbol{\Sigma}|^{-1/2} \exp \left(-(\boldsymbol{x} - \boldsymbol{\mu})^\top \boldsymbol{\Sigma}^{-1} (\boldsymbol{x} - \boldsymbol{\mu}) / 2 \right)$$

则称随机向量 X 服从参数为 $\boldsymbol{\mu}$ 和 $\boldsymbol{\Sigma}$ 的**多维正态分布 (multivariate normal distribution)**, 记为 $X = (X_1, X_2, \dots, X_n) \sim \mathcal{N}(\boldsymbol{\mu}, \boldsymbol{\Sigma})$. 特别地, 当 $n = 2$ 时, 二维随机变量 $(X, Y) \sim \mathcal{N}(\mu_x, \mu_y, \sigma_x^2, \sigma_y^2, \rho)$ 可以写成矩阵形式

$$\boldsymbol{\mu} = \begin{pmatrix} \mu_x \\ \mu_y \end{pmatrix} \quad \text{和} \quad \boldsymbol{\Sigma} = \begin{pmatrix} \sigma_x^2 & \rho\sigma_x\sigma_y \\ \rho\sigma_x\sigma_y & \sigma_y^2 \end{pmatrix}$$

多维正态分布

当 $n = 2$ 时, 三维随机变量 $(X_1, X_2, X_3) \sim \mathcal{N}(\mu_1, \mu_2, \mu_3, \sigma_1^2, \sigma_2^2, \sigma_3^2, \boldsymbol{\rho})$ 可以写成矩阵形式

$$\boldsymbol{\mu} = \begin{pmatrix} \mu_1 \\ \mu_2 \\ \mu_3 \end{pmatrix}$$

和

$$\boldsymbol{\Sigma} = \begin{pmatrix} \text{Var}(X_1) & \text{Cov}(X_1, X_2) & \text{Cov}(X_1, X_3) \\ \text{Cov}(X_2, X_1) & \text{Var}(X_2) & \text{Cov}(X_2, X_3) \\ \text{Cov}(X_3, X_1) & \text{Cov}(X_3, X_2) & \text{Var}(X_3) \end{pmatrix} = \begin{pmatrix} \sigma_1^2 & \rho_{12}\sigma_1\sigma_2 & \rho_{13}\sigma_1\sigma_3 \\ \rho_{12}\sigma_1\sigma_2 & \sigma_2^2 & \rho_{23}\sigma_2\sigma_3 \\ \rho_{13}\sigma_1\sigma_2 & \rho_{23}\sigma_1\sigma_2 & \sigma_3^2 \end{pmatrix}$$

协方差矩阵是一个对称的矩阵, 因此是有实特征值的, 进而是非半正定的。

多维标准正态分布及标准化

回顾: 一维随机变量 $X \sim \mathcal{N}(\mu, \sigma^2)$, 则

$$Y = (X - \mu)/\sigma \sim \mathcal{N}(0, 1) .$$

即正态变量都可以通过一个线性变换(标准化)化成标准状态变量.

定义 0.54 当 $\mu = \mathbf{0}_n$ (全为零的 n 维向量), 以及 $\Sigma = \mathbf{I}_n$ ($n \times n$ 单位阵) 时, 正态分布 $\mathcal{N}(\mathbf{0}_n, \mathbf{I}_n)$ 被称为 n 维标准正态分布.

定理 0.28 设 n 维随机向量 $X = (X_1, X_2, \dots, X_n) \sim \mathcal{N}(\mu, \Sigma)$, 以及正定矩阵 Σ 的特征值分解 $\Sigma = \mathbf{U}^\top \Lambda \mathbf{U}$, 则随机向量

$$Y = \Lambda^{-1/2} \mathbf{U} (X - \mu) \sim \mathcal{N}(\mathbf{0}_n, \mathbf{I}_n) .$$

Proof of Theorem 0.28

证明 根据 $Y = \Lambda^{-1/2}U(X - \mu)$ 可得 $X = U^T\Lambda^{1/2}Y + \mu$, 已知 X 的概率密度函数为

$$f_X(x) = (2\pi)^{-n/2} |\Sigma|^{-1/2} \exp \left(-(x - \mu)^T \Sigma^{-1} (x - \mu) / 2 \right) .$$

根据 n 维随机变量函数 (定理 5.10 的多维情况) 的概率密度公式有

$$f_Y(y) = f_X \left(U^T \Lambda^{1/2} y + \mu \right) \left| U^T \Lambda^{1/2} \right| ,$$

其中 $y = (y_1, y_2, \dots, y_n)^T$. 根据特征值分解 $\Sigma = U^T \Lambda U$ 有

$$\left| U^T \Lambda^{1/2} \right| = |\Sigma|^{1/2} ,$$

以及将 $x = U^T \Lambda^{1/2} y + \mu$ 代入有

$$(x - \mu)^T \Sigma^{-1} (x - \mu) = y^T y .$$

由此可得随机向量 $Y = \Lambda^{-1/2}U(X - \mu)$ 的密度函数为

$$f_Y(y) = (2\pi)^{-n/2} \exp \left(-y^T y / 2 \right) ,$$

Proof of Theorem 0.28

- 随机向量 $X = (X_1, X_2, \dots, X_n)$ 的密度函数为

$$f(\mathbf{x}) = (2\pi)^{-n/2} |\Sigma|^{-1/2} \exp \left(-(\mathbf{x} - \boldsymbol{\mu})^\top \Sigma^{-1} (\mathbf{x} - \boldsymbol{\mu}) / 2 \right)$$

- 根据

$$Y = \Lambda^{-1/2} \mathbf{U} (X - \boldsymbol{\mu}) = \Sigma^{-1/2} (X - \boldsymbol{\mu})$$

则有

$$\mathbf{y}^\top \mathbf{y} = (\mathbf{x} - \boldsymbol{\mu})^\top \Sigma^{-1} (\mathbf{x} - \boldsymbol{\mu})$$

- 带入, 则有

$$f_Y(\mathbf{y}) = (2\pi)^{-n/2} \exp \left(-\mathbf{y}^\top \mathbf{y} / 2 \right)$$

多维正态分布的可加性

定理 **0.29** [线性性] 随机向量 $X = (X_1, X_2, \dots, X_n) \sim \mathcal{N}(\boldsymbol{\mu}, \boldsymbol{\Sigma})$, 则有

$$Y = \mathbf{A}X + \mathbf{b} \sim \mathcal{N}(\mathbf{A}\boldsymbol{\mu} + \mathbf{b}, \mathbf{A}\boldsymbol{\Sigma}\mathbf{A}^\top)$$

其中 $|\mathbf{A}| \neq 0$, $\mathbf{A} \in \mathbb{R}^{n \times n}$ 和 $\mathbf{b} \in \mathbb{R}^{n \times 1}$.

多维正态分布其他性质 – 该证明留作思考题

定理 0.30 设随机向量 $X = (X_1, X_2, \dots, X_n)^\top$ 和 $Y = (Y_1, Y_2, \dots, Y_m)^\top$, 以及

$$\begin{pmatrix} X \\ Y \end{pmatrix} \sim \mathcal{N} \left(\begin{pmatrix} \mu_x \\ \mu_y \end{pmatrix}, \begin{pmatrix} \Sigma_{xx} & \Sigma_{xy} \\ \Sigma_{yx} & \Sigma_{yy} \end{pmatrix} \right),$$

则有

- 随机向量 X 和 Y 的边缘分布为 $X \sim \mathcal{N}(\mu_x, \Sigma_{xx})$ 和 $Y \sim \mathcal{N}(\mu_y, \Sigma_{yy})$
- 随机向量 X 和 Y 相互独立的充要条件是 $\Sigma_{xy} = \mathbf{0}_{m \times n}$ (元素全为零的 $m \times n$ 矩阵)
- 在 $X = x$ 的条件下, 随机向量 Y 的分布

$$Y \mid X = x \sim \mathcal{N}(\mu_y + \Sigma_{yx} \Sigma_{xx}^{-1}(x - \mu_x), \Sigma_{yy} - \Sigma_{yx} \Sigma_{xx}^{-1} \Sigma_{xy})$$

- 在 $Y = y$ 的条件下, 随机向量 X 的分布

$$X \mid Y = y \sim \mathcal{N}(\mu_x + \Sigma_{xy} \Sigma_{yy}^{-1}(y - \mu_y), \Sigma_{xx} - \Sigma_{xy} \Sigma_{yy}^{-1} \Sigma_{yx})$$