

Ch05: 多维随机变量及其数字特征

# **Operations on Multi-dimensional Random Variables – 1**

November 14, 2024

# 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;

| 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.

# 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

| $X \backslash Y$ | 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  |
|-----|----------|-------------------|-----|----------------------------|-----|-------------------|-----|-----|-----|
| $P$ | $p_{11}$ | $p_{12} + p_{21}$ | ... | $p_{14} + p_{41} + p_{22}$ | ... | $p_{24} + p_{42}$ | ... | ... | ... |

## 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 0.107

例 0.107 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 0.107

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  $\chi^2$ -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 0.108

**例 0.108** 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 0.108

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

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

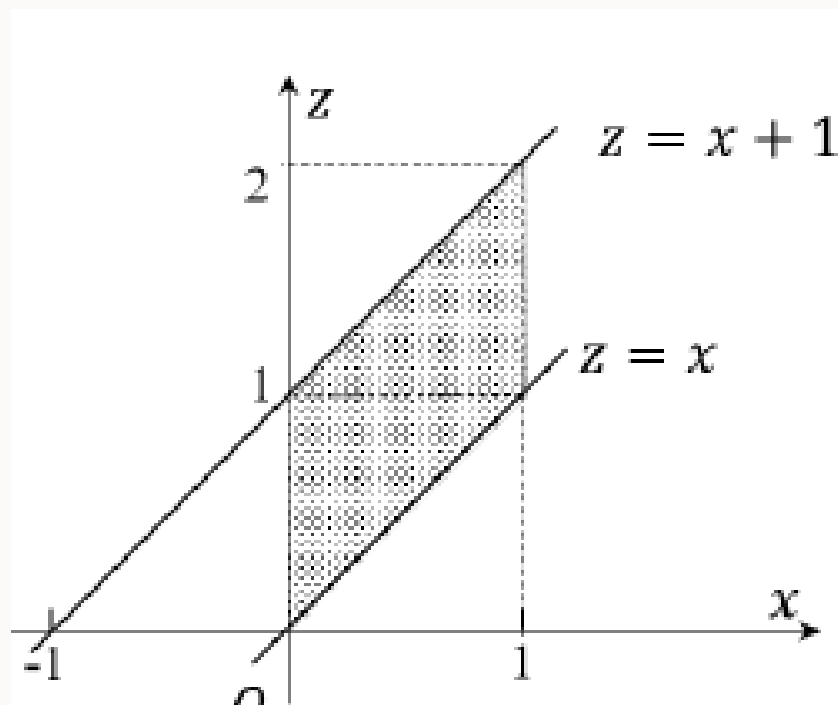
分析  $f_X(x)$ ,  $f_Y(z - x)$  的取值情况可知:

- 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 0.109

**例 0.109** 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 0.109

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 0.110** 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 0.111

例 0.111 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 0.112 指数分布

**例 0.112** 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 0.112

- 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 0.113

**例 0.113** 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  $\theta \sim U(0, 2\pi)$ .

## 解答：例 0.113

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  $\theta \sim U(0, 2\pi)$ .

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  $\Theta$  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  $\Theta$  相互独立.

# **Operations on Multi-dimensional Random Variables – 2**

November 14, 2024

# 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;

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

- 对相同的  $Z$  值合并, 对应的概率相加.

## 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 0.108.

- 指数分布: 例 0.109.

- 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

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