

第三章 多维随机变量及其分布

修贤超

<https://xianchaoxiu.github.io>

- 3.1 二维随机变量
- 3.2 边缘分布
- 3.3 条件分布
- 3.4 相互独立的随机变量
- 3.5 两个随机变量的函数的分布

二维随机变量

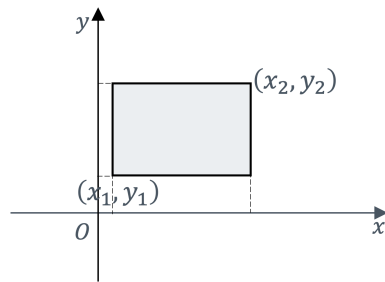
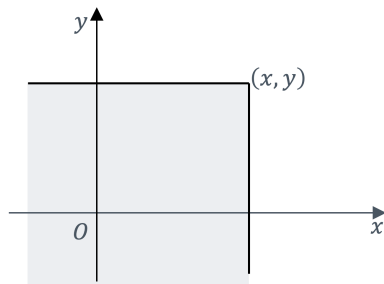
- **定义:** 随机试验 E 的样本空间为 $S = \{e\}$. 设 $X = X(e)$ 和 $Y = Y(e)$ 是定义在样本空间 S 上的随机变量, 则称向量 (X, Y) 为 **二维随机变量**
- 例如
 - $S = \{\text{社区中的全部家庭}\}$
 $X(e)$ 为每个家庭中男孩数量, $Y(e)$ 为每个家庭中女孩数量
 - $S = \{\text{同时投掷的两个骰子点数}\}$
 $X(e)$ 为骰子 1 点数, $Y(e)$ 为骰子 2 点数
 - $S = \{\text{炮弹弹着点}\}$
 $X(e)$ 为弹着点横坐标, $Y(e)$ 为弹着点纵坐标

二维随机变量

■ **定义:** 设 (X, Y) 是二维随机变量, 对于任意实数 x, y , 二元函数

$$F(x, y) = P\{(X \leq x) \cap (Y \leq y)\} = P\{X \leq x, Y \leq y\}$$

称为二维随机变量 (X, Y) 的**分布函数**, 或称为随机变量 X 和 Y 的**联合分布函数** (Joint cumulative probability distribution function)



基本性质

- $F(x, y)$ 是变量 x 和 y 的不减函数
 - 对于任意固定的 y , 如果 $x_2 > x_1$, 则 $F(x_2, y) \geq F(x_1, y)$
 - 对于任意固定的 x , 如果 $y_2 > y_1$, 则 $F(x, y_2) \geq F(x, y_1)$
- $0 \leq F(x, y) \leq 1$ 且
 - 对于任意固定的 y , 有 $F(-\infty, y) = 0$
 - 对于任意固定的 x , 有 $F(x, -\infty) = 0$
 - $F(-\infty, -\infty) = 0, F(\infty, \infty) = 1$
- $F(x + 0, y) = F(x, y), F(x, y + 0) = F(x, y)$, 即 $F(x, y)$ 是 x 和 y 右连续
- 对于任意 (x_1, y_1) 和 (x_2, y_2) , 满足 $x_1 < x_2$ 和 $y_1 < y_2$, 有

$$F(x_2, y_2) + F(x_1, y_1) - F(x_1, y_2) - F(x_2, y_1) \geq 0$$

离散型的随机变量

- **定义:** 如果二维随机变量 (X, Y) 的全部可能取值是有限对或可列无限多对, 则称 (X, Y) 为**离散型的随机变量**
- **定义:** 设 (X, Y) 的可能取值为 (x_i, y_j) , $i, j = 1, 2, \dots$. 记

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

为 (X, Y) 的**分布律**, 或 X 和 Y 的**联合分布律**, 满足

$$p_{ij} \geq 0, \quad \sum_{i=1}^{\infty} \sum_{j=1}^{\infty} p_{ij} = 1$$

例 1

- 设随机变量 X 在 1, 2, 3, 4 四个数中随机取值, 随机变量 Y 在 $1 \sim X$ 中随机取值. 试求 (X, Y) 的分布律

解答 由乘法公式容易求得, 对于 $i = 1, 2, 3, 4, j \leq i$ 有

$$P\{X = i, Y = j\} = P\{Y = j | X = i\}P\{X = i\} = \frac{1}{i} \cdot \frac{1}{4}$$

于是 (X, Y) 的分布律为

	$X = 1$	$X = 2$	$X = 3$	$X = 4$
$Y = 1$	$1/4$	$1/8$	$1/12$	$1/16$
$Y = 2$	0	$1/8$	$1/12$	$1/16$
$Y = 3$	0	0	$1/12$	$1/16$
$Y = 4$	0	0	0	$1/16$

连续型的随机变量

- **定义:** 对于二维随机变量 (X, Y) 的分布函数 $F(X, Y)$, 如果存在非负可积函数 $f(x, y)$ 对于任意 x, y 有

$$F(X, Y) = \int_{-\infty}^y \int_{-\infty}^x f(u, v) du dv$$

则称 (X, Y) 为**连续型的随机变量**

- 函数 $f(x, y)$ 称为二维连续型随机变量 (X, Y) 的**概率密度**, 或称为随机变量 X 和 Y 的**联合概率密度**
- 与离散型的随机变量有什么区别

基本性质

- $f(x, y) \geq 0$
- $\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f(x, y) dx dy = F(\infty, \infty) = 1$
- 设 G 是 xOy 平面上的区域, 则 (X, Y) 落在 G 内的概率是

$$P\{(X, Y) \in G\} = \iint_G f(x, y) dx dy$$

- 若 $f(x, y)$ 在 (x, y) 连续, 则有

$$\frac{\partial^2 F(x, y)}{\partial x \partial y} = f(x, y)$$

例 2

- 设随机变量 (X, Y) 具有概率密度

$$f(x, y) = \begin{cases} 2e^{-(2x+y)}, & x > 0, y > 0 \\ 0, & \text{其他} \end{cases}$$

(1) 求分布函数 $F(x, y)$

解答 (1) 根据定义有

$$\begin{aligned} F(x, y) &= \int_{-\infty}^x \int_{-\infty}^y f(x, y) dx dy \\ &= \begin{cases} 2 \int_0^x e^{-2x} \int_0^y e^{-y} dy dx, & x > 0, y > 0 \\ 0, & \text{其他} \end{cases} \\ &= \begin{cases} (1 - e^{-2x})(1 - e^{-y}), & x > 0, y > 0 \\ 0, & \text{其他} \end{cases} \end{aligned}$$

例 2

- 设随机变量 (X, Y) 具有概率密度

$$f(x, y) = \begin{cases} 2e^{-(2x+y)}, & x > 0, y > 0 \\ 0, & \text{其他} \end{cases}$$

(2) 求概率 $P\{Y \leq X\}$

解答 (2) 根据定义有

$$\begin{aligned} P\{Y \leq X\} &= \int_{-\infty}^{\infty} \int_{-\infty}^x f(x, y) dy dx \\ &= 2 \int_0^{\infty} e^{-2x} \int_0^x e^{-y} dy dx \\ &= 2 \int_0^{\infty} (1 - e^{-x}) e^{-2x} dx = \frac{1}{3} \end{aligned}$$

n 维随机变量

- **定义:** 随机试验 E 的样本空间为 $S = \{e\}$. 设

$$X_1 = X_1(e), X_2 = X_2(e), \cdots, X_n = X_n(e)$$

是定义在样本空间 S 上的随机变量, 则称向量 (X_1, X_2, \cdots, X_n) 为 n 维随机变量或 n 维随机向量

- 对于任意实数 x_1, x_2, \cdots, x_n , n 元函数

$$F(x_1, x_2, \cdots, x_n) = P\{X_1 \leq x_1, X_2 \leq x_2, \cdots, X_n \leq x_n\}$$

称为 n 维随机变量 (X_1, X_2, \cdots, X_n) 的分布函数, 或称为随机变量 X_1, X_2, \cdots 和 X_n 的联合分布函数

- $F(x_1, x_2, \cdots, x_n)$ 的性质与 $F(x, y)$ 相似

- 3.1 二维随机变量
- 3.2 边缘分布
- 3.3 条件分布
- 3.4 相互独立的随机变量
- 3.5 两个随机变量的函数的分布

边缘分布

- **定义:** 二维随机变量 (X, Y) 具有分布函数 $F(x, y)$, 则

$$F_X(x) = P\{X \leq x\} = P\{X \leq x, Y < \infty\} = F(x, \infty)$$

$$F_Y(y) = P\{Y \leq y\} = P\{X < \infty, Y \leq y\} = F(\infty, y)$$

分别称为二维随机变量 (X, Y) 的关于 X 和 Y 的**边缘分布函数**

- **定义:** 离散型的随机变量 X 和 Y 的联合分布函数为

$$F(x, y) = \sum_{x_i \leq x} \sum_{y_j \leq y} p_{ij}$$

关于 X 和 Y 的边缘分布律为

$$p_{i\cdot} = P\{X \leq x_i\} = \sum_{y_j=1}^{\infty} p_{ij}, \quad i = 1, 2, \dots$$

$$p_{\cdot j} = P\{Y \leq y_j\} = \sum_{x_i=1}^{\infty} p_{ij}, \quad j = 1, 2, \dots$$

例 1

- 一整数 N 等可能地在 $1, 2, \dots, 10$ 十个值中取一个值. 设 $D = D(N)$ 是能整除 N 的正整数的个数, $F = F(N)$ 是能整除 N 的素数的个数. 试写出 D 和 F 的联合分布律和边缘分布律

解答

	1	2	3	4	5	6	7	8	9	10
D	1	2	2	3	2	4	2	4	3	4
F	0	1	1	1	1	2	1	1	1	2

	$D = 1$	$D = 2$	$D = 3$	$D = 4$	$P\{F = j\}$
$F = 0$	$1/10$	0	0	0	$1/10$
$F = 1$	0	$4/10$	$2/10$	$1/10$	$7/10$
$F = 2$	0	0	0	$2/10$	$2/10$
$P\{D = i\}$	$1/10$	$4/10$	$2/10$	$3/10$	1

课堂练习 1

- 某社区 15% 家庭无孩子, 20% 家庭有 1 个孩子, 35% 家庭有 2 个孩子, 30% 家庭有 3 个孩子. 随机选取 1 个家庭, 写出男孩数量 B 与女孩数量 G 的边缘分布律

课堂练习 1

- 某社区 15% 家庭无孩子, 20% 家庭有 1 个孩子, 35% 家庭有 2 个孩子, 30% 家庭有 3 个孩子. 随机选取 1 个家庭, 写出男孩数量 B 与女孩数量 G 的边缘分布律

解答 已知 $B + G$ 取值为 1, 2, 3, 容易得到

	$G = 0$	$G = 1$	$G = 2$	$G = 3$	$P\{B = j\}$
$B = 0$	0.15	0.1	0.0875	0.0375	0.3750
$B = 1$	0.1	0.175	0.1125	0	0.3875
$B = 2$	0.0875	0.1125	0	0	0.2000
$B = 3$	0.0375	0	0	0	0.0375
$P\{G = i\}$	0.3750	0.3875	0.2000	0.0375	1

边缘分布

- **定义:** 二维随机变量 (X, Y) 的概率密度为 $f(x, y)$, 则 (X, Y) 关于 X 和 Y 的边缘分布函数分别为

$$F_X(x) = P\{X \leq x\} = P\{X \leq x, Y < \infty\} = F(x, \infty) = \int_{-\infty}^x \int_{-\infty}^{\infty} f(x, y) dy dx$$

$$F_Y(y) = P\{Y \leq y\} = P\{X < \infty, Y \leq y\} = F(\infty, y) = \int_{-\infty}^y \int_{-\infty}^{\infty} f(x, y) dx dy$$

- **定义:** 连续型的随机变量 X 和 Y 的边缘概率密度为

$$f_X(x) = \int_{-\infty}^{\infty} f(x, y) dy$$

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

例 2

- 设随机变量 X 和 Y 具有联合概率密度

$$f(x, y) = \begin{cases} 6, & x^2 \leq y \leq x \\ 0, & \text{其他} \end{cases}$$

求边缘概率密度 $f_X(x)$ 和 $f_Y(y)$

解答

$$f_X(x) = \int_{-\infty}^{\infty} f(x, y) dy = \begin{cases} \int_{x^2}^x 6 dy = 6(x - x^2), & 0 \leq x \leq 1 \\ 0, & \text{其他} \end{cases}$$

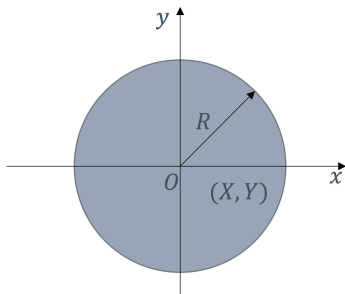
$$f_Y(y) = \int_{-\infty}^{\infty} f(x, y) dx = \begin{cases} \int_y^{\sqrt{y}} 6 dx = 6(\sqrt{y} - y), & 0 \leq y \leq 1 \\ 0, & \text{其他} \end{cases}$$

课堂练习 2

- 设一个圆的半径为 R , 圆心在 O , 其上点为均匀分布, 点的坐标为随机变量 (X, Y) , 其联合概率密度为

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

- (1) 确定 c , (2) 求边缘概率密度 $f_X(x)$ 和 $f_Y(y)$



课堂练习 2

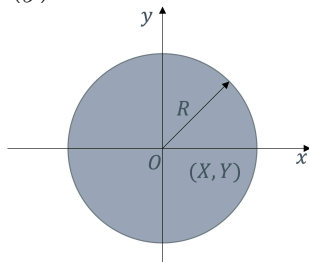
- 设一个圆的半径为 R , 圆心在 O , 其上点为均匀分布, 点的坐标为随机变量 (X, Y) , 其联合概率密度为

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

(1) 确定 c , (2) 求边缘概率密度 $f_X(x)$ 和 $f_Y(y)$

解答 (1) 确定 c

$$\begin{aligned} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f(x, y) dx dy &= \pi R^2 c = 1 \\ \implies c &= \frac{1}{\pi R^2} \end{aligned}$$

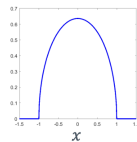
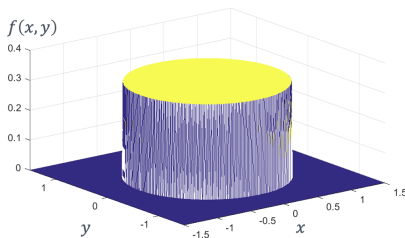


课堂练习 2

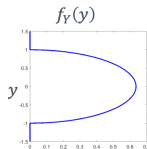
解答 求边缘概率密度 $f_X(x)$ 和 $f_Y(y)$

$$f_X(x) = \int_{-\infty}^{\infty} f(x, y) dy = \int_{-\sqrt{R^2-x^2}}^{\sqrt{R^2-x^2}} \frac{1}{\pi R^2} = \frac{2\sqrt{R^2-x^2}}{\pi R^2}, \quad x^2 \leq R^2$$

$$f_X(x) = \begin{cases} \frac{2\sqrt{R^2-x^2}}{\pi R^2}, & x^2 \leq R^2 \\ 0, & x^2 > R^2 \end{cases} \quad f_Y(y) = \begin{cases} \frac{2\sqrt{R^2-y^2}}{\pi R^2}, & y^2 \leq R^2 \\ 0, & y^2 > R^2 \end{cases}$$



$f_X(x)$



例 3

- 设二维随机变量 X 和 Y 的联合概率密度为

$$f(x, y) = \frac{1}{2\pi\sigma_1\sigma_2\sqrt{1-\rho^2}} \exp \left\{ \frac{-1}{2(1-\rho^2)} \left[\frac{(x-\mu_1)^2}{\sigma_1^2} - 2\rho \frac{(x-\mu_1)(y-\mu_2)}{\sigma_1\sigma_2} + \frac{(y-\mu_2)^2}{\sigma_2^2} \right] \right\}$$

求边缘概率密度 $f_X(x)$ 和 $f_Y(y)$

解答 由于

$$\frac{(y-\mu_2)^2}{\sigma_2^2} - 2\rho \frac{(x-\mu_1)(y-\mu_2)}{\sigma_1\sigma_2} = \left(\frac{y-\mu_2}{\sigma_2} - \rho \frac{x-\mu_1}{\sigma_1} \right)^2 - \rho^2 \frac{(x-\mu_1)^2}{\sigma_1^2}$$

于是

$$\begin{aligned} f_X(x) &= \int_{-\infty}^{\infty} f(x, y) dy \\ &= \frac{1}{2\pi\sigma_1\sigma_2\sqrt{1-\rho^2}} e^{-\frac{(x-\mu_1)^2}{2\sigma_1^2}} \int_{-\infty}^{\infty} e^{-\frac{-1}{2(1-\rho^2)} \left(\frac{y-\mu_2}{\sigma_2} - \rho \frac{x-\mu_1}{\sigma_1} \right)^2} dy \end{aligned}$$

例 3

解答 令

$$t = \frac{1}{\sqrt{1-\rho^2}} \left(\frac{y-\mu_2}{\sigma_2} - \rho \frac{x-\mu_1}{\sigma_1} \right)$$

则有

$$f_X(x) = \frac{1}{2\pi\sigma_1} e^{-\frac{(x-\mu_1)^2}{2\sigma_1^2}} \int_{-\infty}^{\infty} e^{-t^2/2} dt = \frac{1}{\sqrt{2\pi}\sigma_1} e^{-\frac{(x-\mu_1)^2}{2\sigma_1^2}}$$

同理

$$f_Y(y) = \frac{1}{\sqrt{2\pi}\sigma_2} e^{-\frac{(y-\mu_2)^2}{2\sigma_2^2}}$$

- 称 (X, Y) 服从 **二维正态分布**
- 二维正态分布的两个边缘分布都是一维正态分布, 且与 ρ 无关
- 一般来说, 由边缘分布不能确定随机变量的联合分布

- 3.1 二维随机变量
- 3.2 边缘分布
- 3.3 条件分布
- 3.4 相互独立的随机变量
- 3.5 两个随机变量的函数的分布

条件分布

- **定义:** 二维离散型随机变量 (X, Y) , 若对于固定的 j 有 $P\{Y = y_j\} > 0$, 则称

$$P\{X = x_i | Y = y_j\} = \frac{P\{X = x_i, Y = y_j\}}{P\{Y = y_j\}} = \frac{p_{ij}}{p_{\cdot j}}, \quad i = 1, 2, \dots$$

为在 $Y = y_j$ 条件下随机变量 X 的**条件分布律**

- 同理, 若对于固定的 i 有 $P\{X = x_i\} > 0$, 则称

$$P\{Y = y_j | X = x_i\} = \frac{P\{X = x_i, Y = y_j\}}{P\{X = x_i\}} = \frac{p_{ij}}{p_{i\cdot}}, \quad j = 1, 2, \dots$$

为在 $X = x_i$ 条件下随机变量 Y 的**条件分布律**

例 1

- 机器人完成两道工序——紧固 3 只螺栓, 焊接 2 处焊点, 随机变量 X, Y 分别为其完成结果不良的数量, (X, Y) 的联合分布律如下, 求 $P\{Y = j|X = 1\}$

	$X = 0$	$X = 1$	$X = 2$	$X = 3$	$P\{Y = j\}$
$Y = 0$	0.840	0.030	0.020	0.010	0.900
$Y = 1$	0.060	0.010	0.008	0.002	0.080
$Y = 2$	0.010	0.005	0.004	0.001	0.020
$P\{X = i\}$	0.910	0.045	0.032	0.013	1

解答 由 $P\{Y = j|X = 1\} = \frac{P\{X=1, Y=j\}}{P\{X=1\}}$, 计算得

$Y = j$	0	1	2
$P\{Y = j X = 1\}$	6/9	2/9	1/9

课堂练习 1

- 机器人完成两道工序——紧固 3 只螺栓, 焊接 2 处焊点, 随机变量 X, Y 分别为其完成结果不良的数量, (X, Y) 的联合分布律如下, 求 $P\{X = i|Y = 0\}$

	$X = 0$	$X = 1$	$X = 2$	$X = 3$	$P\{Y = j\}$
$Y = 0$	0.840	0.030	0.020	0.010	0.900
$Y = 1$	0.060	0.010	0.008	0.002	0.080
$Y = 2$	0.010	0.005	0.004	0.001	0.020
$P\{X = i\}$	0.910	0.045	0.032	0.013	1

例 2

- 已知射击一次命中概率是 p ($0 < p < 1$), 射击直至击中目标两次为止. X 表示第一次命中时的射击次数, Y 表示第二次命中时的射击次数. 计算 X 和 Y 的联合分布律和条件分布律

解答 X 和 Y 的联合分布律为

$$P\{X = m, Y = n\} = p^2(1 - p)^{n-2}, \quad n = 2, 3, \dots, m = 1, 2, \dots$$

又

$$\begin{aligned} P\{X = m\} &= \sum_{n=m+1}^{\infty} P\{X = m, Y = n\} \\ &= \sum_{n=m+1}^{\infty} p^2(1 - p)^{n-2} = (1 - p)^{m-1}p \end{aligned}$$

例 2

解答 同理有

$$\begin{aligned} P\{Y = n\} &= \sum_{m=1}^{n-1} P\{X = m, Y = n\} \\ &= \sum_{m=1}^{n-1} p^2(1-p)^{n-2} = (n-1)p^2(1-p)^{n-2} \end{aligned}$$

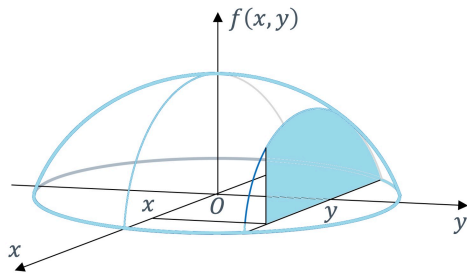
于是

$$\begin{aligned} P\{Y = n|X = m\} &= \frac{P\{X = m, Y = n\}}{P\{X = m\}} = \frac{p^2(1-p)^{n-2}}{(1-p)^{m-1}p} = p(1-p)^{n-m-1} \\ P\{X = m|Y = n\} &= \frac{P\{X = m, Y = n\}}{P\{Y = n\}} = \frac{p^2(1-p)^{n-2}}{(n-1)p^2(1-p)^{n-2}} = \frac{1}{n-1} \end{aligned}$$

条件分布

- 对于二维连续型随机变量 (X, Y) , 有

$$\begin{aligned} P\{X \leq x | Y = y\} &= \frac{P\{X \leq x, Y = y\}}{P\{Y = y\}} \\ &= \frac{\varepsilon \int_{-\infty}^x f(x, y) dx}{\varepsilon f_Y(y)} = \frac{\int_{-\infty}^x f(x, y) dx}{f_Y(y)} \end{aligned}$$



条件分布

- **定义:** 二维连续型随机变量 (X, Y) 有概率密度 $f(x, y)$, 及关于 Y 的边缘概率密度 $f_Y(y)$, 称

$$f_{X|Y}(x|y) = \frac{f(x, y)}{f_Y(y)}$$

为在 $Y = y$ 条件下, X 的**条件概率密度**, 称

$$F_{X|Y}(x|y) = P\{X \leq x|Y = y\} = \int_{-\infty}^x \frac{f(x, y)}{f_Y(y)} dx$$

为在 $Y = y$ 条件下, X 的**条件分布函数**

- 同理有

$$f_{Y|X}(y|x) = \frac{f(x, y)}{f_X(x)}$$

$$F_{Y|X}(y|x) = P\{Y \leq y|X = x\} = \int_{-\infty}^y \frac{f(y, x)}{f_X(x)} dy$$

例 4

- 设数 X 在区间 $(0, 1)$ 等可能随机取值, 当观察到 $X = x$, $(0 < x < 1)$ 时, Y 在区间 $(x, 1)$ 等可能随机取值, 求 $f_Y(y)$

解答 根据题意知 X 具有概率密度

$$f_X(x) = \begin{cases} 1, & 0 < x < 1 \\ 0, & \text{其他} \end{cases}$$

对于任意给定 $x \in (0, 1)$, 在 $X = x$ 的条件下 Y 的条件概率密度为

$$f_{Y|X}(y|x) = \begin{cases} \frac{1}{1-x}, & x < y < 1 \\ 0, & \text{其他} \end{cases}$$

例 4

解答 于是 X 和 Y 的联合概率密度为

$$\begin{aligned} f(x, y) &= f_{Y|X}(y|x)f_X(x) \\ &= \begin{cases} \frac{1}{1-x}, & 0 < x < y < 1 \\ 0, & \text{其他} \end{cases} \end{aligned}$$

得到关于 Y 的边缘概率密度为

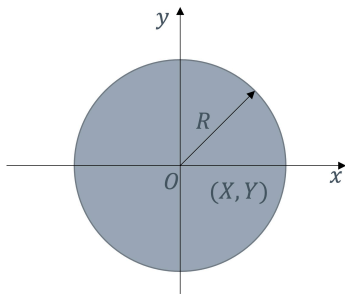
$$\begin{aligned} f_Y(y) &= \int_{-\infty}^{\infty} f(x, y)dx \\ &= \begin{cases} \int_0^y \frac{1}{1-x}dx = \ln(1-y), & 0 < y < 1 \\ 0, & \text{其他} \end{cases} \end{aligned}$$

课堂练习 2

- 设一个圆的半径为 R , 圆心在 O , 其上点为均匀分布, 点的坐标为随机变量 (X, Y) , 其联合概率密度为

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

(3) 求条件概率密度 $f_{X|Y}(x|y)$ 和 $f_{Y|X}(y|x)$



课堂练习 2

- 设一个圆的半径为 R , 圆心在 O , 其上点为均匀分布, 点的坐标为随机变量 (X, Y) , 其联合概率密度为

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

(3) 求条件概率密度 $f_{X|Y}(x|y)$ 和 $f_{Y|X}(y|x)$

解答 (3) 根据 (1) 和 (2) 知

$$f(x, y) = \begin{cases} \frac{1}{\pi R^2}, & x^2 + y^2 \leq R^2 \\ 0, & x^2 + y^2 > R^2 \end{cases}$$

$$f_X(x) = \begin{cases} \frac{2\sqrt{R^2-x^2}}{\pi R^2}, & x^2 \leq R^2 \\ 0, & x^2 > R^2 \end{cases} \quad f_Y(y) = \begin{cases} \frac{2\sqrt{R^2-y^2}}{\pi R^2}, & y^2 \leq R^2 \\ 0, & y^2 > R^2 \end{cases}$$

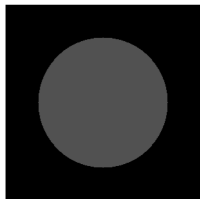
课堂练习 2

解答 于是

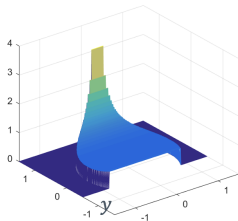
$$f_{X|Y}(x|y) = \frac{f(x, y)}{f_Y(y)} = \begin{cases} \frac{1}{2\sqrt{R^2 - y^2}}, & x^2 \leq R^2 - y^2 \\ 0, & \text{其他} \end{cases}$$

$$f_{Y|X}(y|x) = \frac{f(x, y)}{f_X(x)} = \begin{cases} \frac{1}{2\sqrt{R^2 - x^2}}, & y^2 \leq R^2 - x^2 \\ 0, & \text{其他} \end{cases}$$

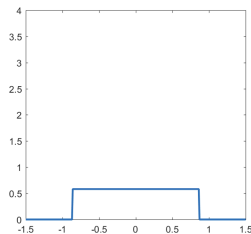
$f(x, y)$



$f_{X|Y}(x|y)$



$f_{X|Y}(x|y)$



- 3.1 二维随机变量
- 3.2 边缘分布
- 3.3 条件分布
- 3.4 相互独立的随机变量
- 3.5 两个随机变量的函数的分布

相互独立的随机变量

- **定义:** 如果对任意两个实数集 A 和 B , 满足

$$P\{X \in A, Y \in B\} = P\{X \in A\}P\{Y \in B\}$$

则称 A 和 B 是相互独立的

- 如果 X 和 Y 是相互独立的, 则

$$P\{X \leq x, Y \leq y\} = P\{X \leq x\}P\{Y \leq y\}$$

- 当 X 和 Y 是离散随机变量时, 有

$$p_{ij} = p_{i\cdot}p_{\cdot j}$$

- 当 X 和 Y 是连续随机变量时, 有

$$f(x, y) = f_X(x)f_Y(y)$$

例 1

- 一负责人到达办公室的时间均匀分布在 $8 \sim 12$ 时, 他的秘书到达办公室的时间均匀分布在 $7 \sim 9$ 时, 设他们两人到达的时间相互独立, 求他们到达办公室的时间相差不超过 5 分钟的概率

解答 设 X 和 Y 分别是两人到达办公室的时间, 由假设知概率密度分别为

$$f_X(x) = \begin{cases} \frac{1}{4}, & 8 < x < 12 \\ 0, & \text{其他} \end{cases} \quad f_Y(y) = \begin{cases} \frac{1}{2}, & 7 < y < 9 \\ 0, & \text{其他} \end{cases}$$

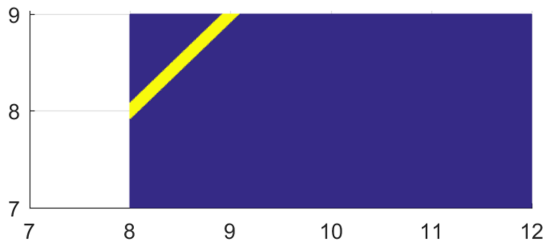
因为 X 和 Y 相互独立, 故 (X, Y) 的概率密度为

$$f(x, y) = f_X(x)f_Y(y) = \begin{cases} \frac{1}{8}, & 8 < x < 12, 7 < y < 9 \\ 0, & \text{其他} \end{cases}$$

例 1

解答 所求的概率为

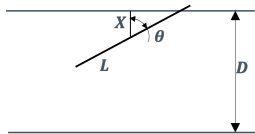
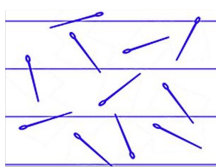
$$\begin{aligned} P\left\{|X - Y| \leq \frac{1}{12}\right\} &= \iint_G f(x, y) dx dy = \frac{1}{8} \times (G \text{的面积}) \\ &= \frac{1}{8} \times \left(\frac{1}{2} \left(\frac{13}{12} \right)^2 - \frac{1}{2} \left(\frac{11}{12} \right)^2 \right) = \frac{1}{8} \times \frac{1}{6} = \frac{1}{48} \end{aligned}$$



- 18 世纪, Buffon 提出以下问题: 设我们有一个以平行且等距木纹铺成的地板, 随意抛一支长度比木纹之间距离小的针, 求针和其中一条木纹相交的概率

解答 设线间距离 D , 针长 L , 针的角度 $\theta \sim U(0, \pi/2)$, 针中心至线的距离 $X \sim U(0, D/2)$, 于是

$$\begin{aligned}
 P\left\{X < \frac{L}{2}\cos\theta\right\} &= \iint_{x < \frac{L}{2}\cos\theta} f(x, y) dx dy \\
 &= \iint_{x < \frac{L}{2}\cos\theta} f_X(x) f_\theta(y) dx dy \\
 &= \int_0^{\pi/2} \int_0^{\frac{L}{2}\cos\theta} \frac{1}{\pi/2} \frac{1}{D/2} dx dy \\
 &= \int_0^{\pi/2} \frac{2L\cos y}{\pi D} dy = \frac{2L}{\pi D}
 \end{aligned}$$



蒙特卡洛方法

- 由 20 世纪 40 年代美国在第二次世界大战中研制原子弹的“曼哈顿计划”计划的成员乌拉姆和冯·诺伊曼首先提出,以驰名世界的赌城—摩纳哥的 Monte Carlo 来命名这种方法,为它蒙上了一层神秘色彩. 1777 年,法国数学家布丰提出用投针实验的方法求圆周率 π , 这被认为是蒙特卡罗方法的起源
- 针对 Buffon 问题, 取 $L = D/2$, 概率为 $1/\pi$

试验者	时间	投掷次数	相交次数	圆周率的估计值
Wolf	1850 年	5000	2532	3.1596
Smith	1855 年	3204	1218.5	3.1554
C.De Morgan	1860 年	600	382.5	3.137
Fox	1884 年	1030	489	3.1595
Lazzerini	1901 年	3408	1808	3.1415929

n 维随机变量

■ n 维随机变量 (X_1, \dots, X_n) 的分布函数

$$F(x_1, \dots, x_n) = P\{X_1 \leq x_1, \dots, X_n \leq x_n\}$$

■ 概率密度函数

$$F(x_1, x_2, \dots, x_n) = \int_{-\infty}^{x_n} \cdots \int_{-\infty}^1 f(x_1, \dots, x_n) dx_1 \cdots dx_n$$

■ 边缘分布函数

$$F_{X_k}(x_k) = F(\infty, \dots, \infty, x_k, \infty, \dots, \infty)$$

$$F_{X_k, X_l}(x_k, x_l) = F(\infty, \dots, \infty, x_k, \infty, \dots, \infty, x_l, \infty, \dots, \infty)$$

■ 边缘概率密度

$$f_{X_k}(x_k) = \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} f(x_1, \dots, x_n) dx_1 \cdots dx_{k-1} dx_{k+1} \cdots dx_n$$

$$f_{X_k, X_l}(x_k, x_l) = \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} f(x_1, \dots, x_n) dx_1 \cdots dx_{k-1} dx_{k+1} \cdots dx_{l-1} dx_{l+1} \cdots dx_n$$

- **定义:** 若对于所有的 x_1, \dots, x_n 有

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

则称 (X_1, \dots, X_n) **相互独立**

- **定义:** 若对于所有的 x_1, \dots, x_n 和 y_1, \dots, y_n 有

$$F(x_1, \dots, x_n, y_1, \dots, y_n) = F(x_1, \dots, x_n)F(y_1, \dots, y_n)$$

则称 (X_1, \dots, X_n) 和 (Y_1, \dots, Y_n) **相互独立**

- **定理:** 设 (X_1, \dots, X_n) 和 (Y_1, \dots, Y_n) 相互独立, 则 X_i ($i = 1, \dots, m$) 和 Y_i ($i = 1, \dots, n$) 相互独立. 又若 h 和 g 是连续函数, 则 $h(X_1, \dots, X_m)$ 和 $g(Y_1, \dots, Y_n)$ 相互独立

- 3.1 二维随机变量
- 3.2 边缘分布
- 3.3 条件分布
- 3.4 相互独立的随机变量
- 3.5 两个随机变量的函数的分布

$Z = X + Y$ 的分布

- 设 X 和 Y 是二维连续型随机变量, 具有概率密度 $f(x, y)$, 则 $Z = X + Y$ 仍为连续型随机变量, 其概率密度分别为

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

- 若 X 和 Y 相互独立, 设 (X, Y) 关于 X 和 Y 的边缘概率密度分别为 $f_X(x)$ 和 $f_Y(y)$, 则

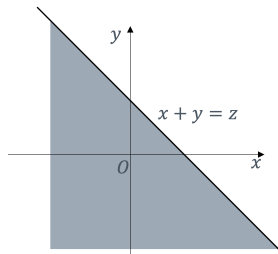
$$f_{X+Y}(z) = \int_{-\infty}^{\infty} f_X(z - y) f_Y(y) dy$$
$$f_{X+Y}(z) = \int_{-\infty}^{\infty} f_X(x) f_Y(z - x) dx$$

这两个公式称为 f_X 和 f_Y 的卷积公式, 记为 $f_X * f_Y$

$Z = X + Y$ 的分布

■ 讨论 $Z = X + Y$ 的分布函数 $F_Z(z)$

$$\begin{aligned} F_Z(z) &= P\{Z \leq z\} = P\{X + Y \leq z\} \\ &= P\{X \leq z - Y\} = F_X(z - Y) \\ &= \int_{-\infty}^{\infty} dy \int_{-\infty}^{z-y} f(x, y) dx \\ &= \int_{-\infty}^{\infty} dy \int_{-\infty}^{z-y} f(x, y) d(x + y) \\ &= \int_{-\infty}^{\infty} dy \int_{-\infty}^z f(u - y, y) du \\ &= \int_{-\infty}^z du \int_{-\infty}^{\infty} f(u - y, y) dy \end{aligned}$$



例 1

- 设 X 和 Y 是两个独立的随机变量, 且 $X \sim N(0, 1)$, $Y \sim N(0, 1)$, 求 $Z = X + Y$ 的概率密度

解答 由标准正态分布知

$$\begin{aligned} f_Z(z) &= \int_{-\infty}^{\infty} f_X(x) f_Y(z-x) dx \\ &= \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{-\frac{x^2}{2}} \cdot e^{-\frac{(z-x)^2}{2}} dx = \frac{1}{2\pi} e^{-\frac{z^2}{4}} \int_{-\infty}^{\infty} e^{-(x-\frac{z}{2})^2} dx \end{aligned}$$

令 $t = x - \frac{z}{2}$, 得

$$f_Z(z) = \frac{1}{2\pi} e^{-\frac{z^2}{4}} \int_{-\infty}^{\infty} e^{-t^2} dx = \frac{1}{2\pi} e^{-\frac{z^2}{4}} \sqrt{\pi} = \frac{1}{2\sqrt{\pi}} e^{-\frac{z^2}{4}}$$

则 Z 服从 $N(0, 2)$ 分布

- 设 X 和 Y 是两个独立的随机变量, 且 $X \sim N(0, 1)$, $Y \sim N(0, 1)$, 则 $Z = X + Y$ 满足

$$Z \sim N(0, 2)$$

- 设 X 和 Y 是两个独立的随机变量, 且 $X \sim N(\mu_1, \sigma_1^2)$, $Y \sim N(\mu_2, \sigma_2^2)$, 则 $Z = X + Y$ 满足

$$Z \sim N(\mu_1 + \mu_2, \sigma_1^2 + \sigma_2^2)$$

- 设 X_1, X_2, \dots, X_n 是 n 个独立的随机变量, 且 $X_i \sim N(\mu_i, \sigma_i^2)$, 则 $Z = X_1 + X_2 + \dots + X_n$ 满足

$$Z \sim N(\mu_1 + \mu_2 + \dots + \mu_n, \sigma_1^2 + \sigma_2^2 + \dots + \sigma_n^2)$$

- 有限个相互独立的正态随机变量的线性组合仍然服从正态分布

例 2

- 在一简单电路中, 两个串联电阻 R_1 和 R_2 相互独立, 概率密度均为

$$f_X(x) = \begin{cases} \frac{10-x}{50}, & 0 \leq x \leq 10 \\ 0, & \text{其他} \end{cases}$$

求 $R = R_1 + R_2$ 的概率密度

解答 R 的概率密度为

$$\begin{aligned} f_R(z) &= \int_{-\infty}^{\infty} f(x)f(z-x)dx \\ &= \begin{cases} \int_0^z \frac{10-x}{50} \frac{10-(z-x)}{50} dx = \frac{600z-60z^2+z^3}{15000}, & 0 \leq z < 10 \\ \int_{z-10}^{10} \frac{10-x}{50} \frac{10-(z-x)}{50} dx = \frac{(20-z)^2}{15000}, & 10 \leq z \leq 20 \end{cases} \end{aligned}$$

$Z = Y/X$ 的分布、 $Z = XY$ 的分布

- 设 X 和 Y 是二维连续型随机变量, 具有概率密度 $f(x, y)$, 则 $Z = Y/X$ 、 $Z = XY$ 仍为连续型随机变量, 其概率密度分别为

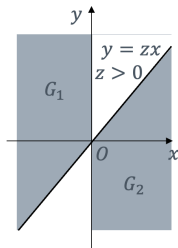
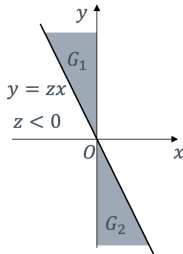
$$f_{Y/X}(z) = \int_{-\infty}^{\infty} |x| f(x, xz) dx$$
$$f_{XY}(z) = \int_{-\infty}^{\infty} 1/|x| f(x, z/x) dx$$

- 若 X 和 Y 相互独立, 设 (X, Y) 关于 X 和 Y 的边缘概率密度分别为 $f_X(x)$ 和 $f_Y(y)$, 则

$$f_{Y/X}(z) = \int_{-\infty}^{\infty} |x| f_X(x) f_Y(xz) dx$$
$$f_{XY}(z) = \int_{-\infty}^{\infty} 1/|x| f_X(x) f_Y(z/x) dx$$

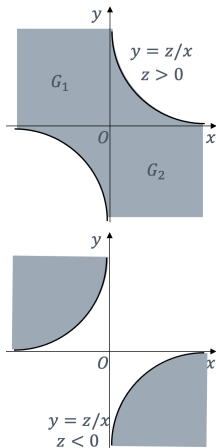
$Z = Y/X$ 的分布

$$\begin{aligned} F_Z(z) &= P\{Z \leq z\} = P\{Y/X \leq z\} = \iint_{G_1 \cup G_2} f(x, y) dx dy \\ &= \int_{-\infty}^0 dx \int_{zx}^{\infty} f(x, y) dy + \int_0^{\infty} dx \int_{-\infty}^{zx} f(x, y) dy \\ &= \int_{-\infty}^0 dx \int_{zx}^{\infty} f(x, xu) d(xu) + \int_0^{\infty} dx \int_{-\infty}^{zx} f(x, xu) d(xu) \\ &= \int_{-\infty}^0 dx \int_z^{-\infty} x f(x, u) du + \int_0^{\infty} dx \int_{-\infty}^z x f(x, xu) du \\ &= \int_{-\infty}^0 dx \int_{-\infty}^z (-x) f(x, xu) du + \int_0^{\infty} dx \int_{-\infty}^z x f(x, xu) du \\ &= \int_{-\infty}^z du \int_{-\infty}^0 (-x) f(x, xu) du + \int_{-\infty}^z du \int_0^{\infty} x f(x, xu) dx \\ &= \int_{-\infty}^z du \int_{-\infty}^{\infty} |x| f(x, xu) dx \end{aligned}$$



$Z = XY$ 的分布

$$\begin{aligned}F_Z(z) &= P\{Z \leq z\} = P\{XY \leq z\} = \iint_{G_1 \cup G_2} f(x, y) dx dy \\&= \int_{-\infty}^0 dx \int_{z/x}^{\infty} f(x, y) dy + \int_0^{\infty} dx \int_{-\infty}^{z/x} f(x, y) dy \\&= \int_{-\infty}^0 dx \int_{z/x}^{\infty} f(x, u/x) d(u/x) + \int_0^{\infty} dx \int_{-\infty}^{z/x} f(x, u/x) d(u/x) \\&= \int_{-\infty}^0 dx \int_z^{-\infty} 1/x f(x, u/x) du + \int_0^{\infty} dx \int_{-\infty}^z 1/x f(x, u/x) du \\&= \int_{-\infty}^0 dx \int_{-\infty}^z (-1/x) f(x, u/x) du + \int_0^{\infty} dx \int_{-\infty}^z 1/x f(x, u/x) du \\&= \int_{-\infty}^z du \int_{-\infty}^0 (-1/x) f(x, u/x) dx + \int_{-\infty}^z du \int_0^{\infty} -(1/x) f(x, u/x) dx \\&= \int_{-\infty}^z du \int_{-\infty}^{\infty} 1/|x| f(x, u/x) dx\end{aligned}$$



例 4

- 某公司提供一种地震保险, 保险费率 Y 和保险赔付 X 的概率密度分别为

$$f(y) = \begin{cases} \frac{y}{25} e^{-\frac{y}{5}}, & y > 0 \\ 0, & \text{其他} \end{cases} \quad g(x) = \begin{cases} \frac{1}{5} e^{-\frac{x}{5}}, & x > 0 \\ 0, & \text{其他} \end{cases}$$

设 X 与 Y 相互独立, 求 $Z = Y/X$ 的概率密度

解答 Z 的概率密度为

$$\begin{aligned} f_Z(z) &= \int_{-\infty}^{\infty} |x| g(x) f(xz) dx \\ &= \begin{cases} \int_{-\infty}^{\infty} x \frac{1}{5} e^{-\frac{x}{5}} \frac{xz}{25} e^{-\frac{xz}{5}} dx = \frac{2z}{(1+z)^3}, & z > 0 \\ 0, & \text{其他} \end{cases} \end{aligned}$$

$M = \max\{X, Y\}$ 的分布、 $N = \min\{X, Y\}$ 的分布

- 若 X 和 Y 是二维连续型随机变量, 其分布函数分别为 $F_X(x)$ 和 $F_Y(y)$, 则 $M = \max\{X, Y\}$ 、 $N = \min\{X, Y\}$ 的分布函数分别为

$$P\{M \leq z\} = P\{X \leq z, Y \leq z\} = P\{X \leq z\}P\{Y \leq z\}$$

$$\Rightarrow F_{\max}(z) = F_X(z)F_Y(z)$$

$$P\{N \leq z\} = 1 - P\{N > z\} = 1 - P\{X > z\}P\{Y > z\}$$

$$\Rightarrow F_{\min}(z) = 1 - [1 - F_X(z)][1 - F_Y(z)]$$

- 若随机变量 X_1, \dots, X_n 相互独立, 其分布函数分别为 $F_{X_1}(x), \dots, F_{X_n}(x)$, 则 $M = \max\{X_1, \dots, X_n\}$ 、 $N = \min\{X_1, \dots, X_n\}$ 的分布函数分别为

$$F_{\max}(z) = F_{X_1}(z)F_{X_2}(z) \cdots F_{X_n}(z)$$

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

例 5

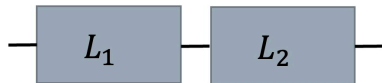
- 某公司提系统 L 由相互独立的子系统 L_1 和 L_2 组成, 两子系统寿命分别为 X 和 Y , 其概率密度分别为

$$f_X(x) = \begin{cases} \alpha e^{-\alpha x}, & x > 0 \\ 0, & \text{其他} \end{cases} \quad f_Y(y) = \begin{cases} \beta e^{-\beta y}, & y > 0 \\ 0, & \text{其他} \end{cases}$$

其中 $\alpha > 0, \beta > 0, \alpha \neq \beta$, 求在不同连接情况下, L 寿命 Z 的概率密度

解答 (1) 串联

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



例 5

解答 由于 $N = \min\{X, Y\}$, 有

$$\begin{aligned} F_Z(z) &= 1 - [1 - F_X(z)][1 - F_Y(z)] \\ &= \begin{cases} 1 - e^{-(\alpha+\beta)z}, & z > 0 \\ 0, & \text{其他} \end{cases} \end{aligned}$$

于是

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

例 5

- 某公司提系统 L 由相互独立的子系统 L_1 和 L_2 组成, 两子系统寿命分别为 X 和 Y , 其概率密度分别为

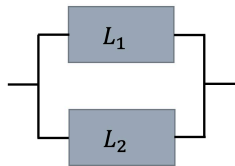
$$f_X(x) = \begin{cases} \alpha e^{-\alpha x}, & x > 0 \\ 0, & \text{其他} \end{cases} \quad f_Y(y) = \begin{cases} \beta e^{-\beta y}, & y > 0 \\ 0, & \text{其他} \end{cases}$$

其中 $\alpha > 0, \beta > 0, \alpha \neq \beta$, 求在不同连接情况下, L 寿命 Z 的概率密度

解答 (2) 并联 $Z = \max\{X, Y\}$

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

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



例 5

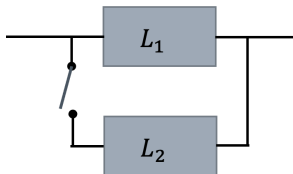
- 某公司提系统 L 由相互独立的子系统 L_1 和 L_2 组成, 两子系统寿命分别为 X 和 Y , 其概率密度分别为

$$f_X(x) = \begin{cases} \alpha e^{-\alpha x}, & x > 0 \\ 0, & \text{其他} \end{cases} \quad f_Y(y) = \begin{cases} \beta e^{-\beta y}, & y > 0 \\ 0, & \text{其他} \end{cases}$$

其中 $\alpha > 0, \beta > 0, \alpha \neq \beta$, 求在不同连接情况下, L 寿命 Z 的概率密度

解答 (3) 备用电路 $Z = X + Y$

$$\begin{aligned} f_Z(z) &= f_X(x) * f_Y(y) = \int_{-\infty}^{\infty} f_X(z-y)f_Y(y)dy \\ &= \begin{cases} \frac{\alpha\beta}{\beta-\alpha}(e^{-\alpha z} - e^{-\beta z}), & z > 0 \\ 0, & \text{其他} \end{cases} \end{aligned}$$



■ 考试内容

- 多维随机变量及其分布、二维离散型随机变量的概率分布、边缘分布和条件分布、二维连续型随机变量的概率密度、边缘概率密度和条件密度、随机变量的独立性和不相关性、常用二维随机变量的分布、两个及两个以上随机变量简单函数的分布

■ 考试要求

- 理解多维随机变量的概念, 理解多维随机变量的分布的概念和性质. 理解二维离散型随机变量的概率分布、边缘分布和条件分布, 理解二维连续型随机变量的概率密度、边缘密度和条件密度
- 理解随机变量的独立性及不相关性的概念, 掌握随机变量相互独立的条件
- 掌握二维均匀分布, 了解二维正态分布的概率密度, 理解参数的概率意义
- 会求两个随机变量函数的分布, 会求多个相互独立随机变量函数的分布

Q&A

Thank you!

感谢您的聆听和反馈