

第3章 多维随机变量及其分布

上一章讨论了一个随机变量的情况,但在实际问题中某些随机试验的结果需要同时用两个或两个以上的随机变量来描述.例如,研究某一地区的学龄前儿童的身体发育情况,需要同时考虑儿童的身高和体重;研究一个地区的财政收入情况,需要同时分析当地的国内生产总值、税收和其他收入等.因此,本章介绍多维随机变量及其分布.

3.1 多维随机变量及其分布函数

3.1.1 多维随机变量

定义 3.1.1 设 Ω 是随机试验的样本空间,对 Ω 中的每一个样本点 ω ,有 n 个实数 $X_1(\omega), X_2(\omega), \dots, X_n(\omega)$ 与之对应,称 (X_1, X_2, \dots, X_n) 为定义在 Ω 上的一个 n 维随机变量.

由于二维随机变量与 n 维随机变量没有本质的区别,为简单起见,下面着重讨论二维随机变量.

定义 3.1.2 设 Ω 是随机试验的样本空间,对 Ω 中的每一个样本点 ω ,有两个实数 $X(\omega), Y(\omega)$ 与之对应,称 (X, Y) 为定义在 Ω 上的一个二维随机变量.

3.1.2 联合分布函数

与一维随机变量类似,首先引入二维随机变量的分布函数.

定义 3.1.3 设 (X, Y) 是二维随机变量,对任意 $x, y \in \mathbf{R}$,称二元函数

$$F(x, y) = P\{X \leq x, Y \leq y\} \quad (3.1.1)$$

为二维随机变量 (X, Y) 的联合分布函数.

若将二维随机变量 (X, Y) 看作平面上的随机点,则联合分布函数 $F(x, y)$ 的几何意义为随机点 (X, Y) 落在点 (x, y) 左下方阴影部分内的概率,如图 3.1.1 所示.

与一维随机变量的分布函数类似,二维随机变量的联合分布函数 $F(x, y)$ 具有以下基本性质.

定理 3.1.1 设 $F(x, y)$ 为二维随机变量 (X, Y) 的联合分布函数,则 $F(x, y)$ 满足

(1) 单调性: $F(x, y)$ 分别关于 x, y 单调不降,即

当 $x_1 < x_2$ 时, $F(x_1, y) \leq F(x_2, y)$;

当 $y_1 < y_2$ 时, $F(x, y_1) \leq F(x, y_2)$.

(2) 有界性: $0 \leq F(x, y) \leq 1$, 且 $\lim_{x \rightarrow -\infty} F(x, y) = 0$, $\lim_{y \rightarrow -\infty} F(x, y) = 0$, $\lim_{\substack{x \rightarrow +\infty \\ y \rightarrow +\infty}} F(x, y) = 1$.

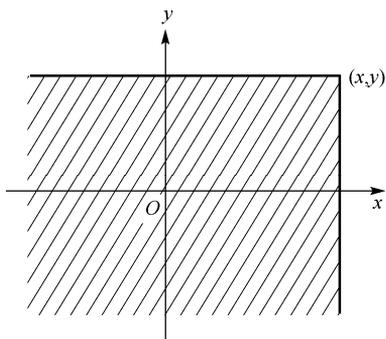


图 3.1.1 分布函数 $F(x, y)$ 的几何意义

(3) 右连续性: $F(x, y)$ 是右连续函数, 即

$$F(x_0 + 0, y) = \lim_{x \rightarrow x_0^+} F(x, y) = F(x_0, y),$$

$$F(x, y_0 + 0) = \lim_{y \rightarrow y_0^+} F(x, y) = F(x, y_0).$$

(4) 相容性: 对于任意实数 $x_1 < x_2, y_1 < y_2$,

$$\begin{aligned} & P\{x_1 < X \leq x_2, y_1 < Y \leq y_2\} \\ &= F(x_2, y_2) - F(x_2, y_1) - F(x_1, y_2) + F(x_1, y_1) \geq 0. \end{aligned}$$

二维概率计算如图 3.1.2 所示.

满足以上四条性质的二元函数 $F(x, y)$, 一定是某个二维随机变量的联合分布函数; 反之, 二维随机变量的联合分布函数一定满足以上四条性质.

例 3.1.1 设二维随机变量 (X, Y) 的联合分布函数为

$$F(x, y) = A(B + \arctan x)(C + \arctan y), \quad (x, y) \in \mathbf{R}^2$$

试求系数 A, B, C , 并求 $P\{0 \leq x < 1, 0 \leq y < 1\}$.

解 $\lim_{x \rightarrow -\infty} F(x, y) = \lim_{x \rightarrow -\infty} A(B + \arctan x)(C + \arctan y)$

$$= A \left(B - \frac{\pi}{2} \right) (C + \arctan y) = 0$$

$$\lim_{y \rightarrow -\infty} F(x, y) = \lim_{y \rightarrow -\infty} A(B + \arctan x)(C + \arctan y)$$

$$= A(B + \arctan x) \left(C - \frac{\pi}{2} \right) = 0$$

$$\lim_{\substack{x \rightarrow +\infty \\ y \rightarrow +\infty}} F(x, y) = \lim_{\substack{x \rightarrow +\infty \\ y \rightarrow +\infty}} A(B + \arctan x)(C + \arctan y)$$

$$= A \left(B + \frac{\pi}{2} \right) \left(C + \frac{\pi}{2} \right) = 1$$

由此可得 $A = \frac{1}{\pi^2}, B = C = \frac{\pi}{2}$

根据联合分布函数的相容性可知

$$P\{0 \leq X < 1, 0 \leq Y < 1\} = F(1, 1) - F(0, 1) - F(1, 0) + F(0, 0) = \frac{1}{16}$$

定义 3.1.4 设 (X_1, X_2, \dots, X_n) 为 Ω 上的一个 n 维随机变量, 对任意 $x_i \in \mathbf{R}, i = 1, 2, \dots, n$, 称 n 元函数

$$F(x_1, x_2, \dots, x_n) = P\{X_1 \leq x_1, X_2 \leq x_2, \dots, X_n \leq x_n\} \quad (3.1.2)$$

为 n 维随机变量 (X_1, X_2, \dots, X_n) 的联合分布函数.

式中, $\{X_1 \leq x_1, X_2 \leq x_2, \dots, X_n \leq x_n\} = \{X_1 \leq x_1\} \cap \{X_2 \leq x_2\} \cap \dots \cap \{X_n \leq x_n\}$.

3.1.3 联合分布律

定义 3.1.5 如果二维随机变量 (X, Y) 只取有限对或可列无穷多对值 $(x_i, y_j), i, j = 1, 2, \dots$,

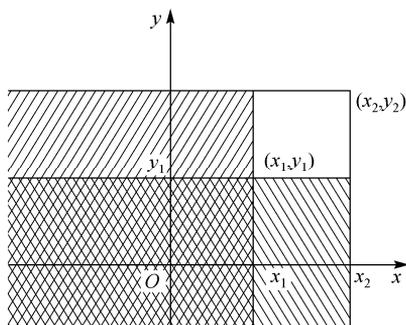


图 3.1.2 二维概率计算

则称 (X, Y) 为二维离散型随机变量, 称

$$P\{X = x_i, Y = y_j\} = p_{ij}, \quad (i, j = 1, 2, \dots) \quad (3.1.3)$$

为 (X, Y) 的联合分布律.

二维离散型随机变量的联合分布律也可表示为如表 3.1.1 所示的表格形式.

表 3.1.1 二维离散型随机变量的联合分布律

X	Y				
	y_1	y_2	\dots	y_j	\dots
x_1	p_{11}	p_{12}	\dots	p_{1j}	\dots
x_2	p_{21}	p_{22}	\dots	p_{2j}	\dots
\vdots	\vdots	\vdots		\vdots	
x_i	p_{i1}	p_{i2}	\dots	p_{ij}	\dots
\vdots	\vdots	\vdots		\vdots	

二维离散型随机变量的联合分布律满足下列两条基本性质.

(1) $p_{ij} \geq 0, \quad i, j = 1, 2, \dots;$

(2) $\sum_{i=1}^{\infty} \sum_{j=1}^{\infty} p_{ij} = 1.$

二维离散型随机变量的联合分布函数为

$$F(x, y) = \sum_{x_i \leq x} \sum_{y_j \leq y} p_{ij} \quad (3.1.4)$$

例 3.1.2 (二维两点分布) 用剪刀随机地去剪一条悬挂有小球的绳子. 剪中绳子的概率为 p ($0 < p < 1$), 设 X 表示剪中绳子的次数, Y 表示小球下落的次数. 求 (X, Y) 的联合分布律.

解 (X, Y) 的联合分布律如表 3.1.2 所示.

表 3.1.2 (X, Y) 的联合分布律

X	Y	
	0	1
0	$1-p$	0
1	0	p

3.1.4 联合概率密度函数

定义 3.1.6 设 $F(x, y)$ 是二维随机变量 (X, Y) 的联合分布函数, 如果存在非负可积函数 $f(x, y)$, 使得对于任意实数对 (x, y) , 有

$$F(x, y) = \int_{-\infty}^x \int_{-\infty}^y f(u, v) du dv \quad (3.1.5)$$

则称 (X, Y) 为二维连续型随机变量, $f(x, y)$ 为 (X, Y) 的联合概率密度函数.

二维连续型随机变量的联合概率密度函数满足下列两条基本性质.

(1) $f(x, y) \geq 0;$

$$(2) \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} f(x, y) dx dy = 1.$$

凡是满足上述两条性质的二元函数 $f(x, y)$ ，一定是某个二维连续型随机变量的概率密度函数；反之，某个二维连续型随机变量的概率密度函数一定满足上述两条性质。

由联合分布函数和联合概率密度函数的性质可以得到以下结论。

$$(1) \text{ 在 } f(x, y) \text{ 的连续点处, 有 } \frac{\partial^2 F(x, y)}{\partial x \partial y} = f(x, y);$$

$$(2) P\{(X, Y) \in G\} = \iint_G f(x, y) dx dy.$$

例 3.1.3 设二维随机变量 (X, Y) 的联合概率密度函数为

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

求：(1) 常数 c ；

(2) $P\{X \geq Y\}$ ；

(3) 联合分布函数 $F(x, y)$ 。

解 (1) $1 = \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} f(x, y) dx dy = \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} ce^{-(2x+y)} dx dy = c \int_0^{+\infty} e^{-2x} dx \int_0^{+\infty} e^{-y} dy = \frac{c}{2}$, 得 $c = 2$ 。

$$(2) P\{X \geq Y\} = \iint_{x \geq y} f(x, y) dx dy = \int_0^{+\infty} 2e^{-2x} dx \int_0^x e^{-y} dy = \int_0^{+\infty} 2e^{-2x}(1 - e^{-x}) dx = \frac{1}{3}$$

$$(3) F(x, y) = \int_{-\infty}^x \int_{-\infty}^y f(u, v) du dv = \begin{cases} \int_0^x \int_0^y 2e^{-2u} e^{-v} du dv, & x > 0, y > 0 \\ 0, & \text{其他} \\ (1 - e^{-2x})(1 - e^{-y}), & x > 0, y > 0 \\ 0, & \text{其他} \end{cases}$$

3.1.5 常见的二维连续型分布

1. 二维均匀分布（几何概率）

设 G 为平面上某个有界区域，其面积为 $S(G)$ ，如果二维连续型随机变量 (X, Y) 具有联合概率密度函数

$$f(x, y) = \begin{cases} \frac{1}{S(G)}, & (x, y) \in G \\ 0, & (x, y) \notin G \end{cases} \quad (3.1.6)$$

则称 (X, Y) 在区域 G 上服从**二维均匀分布**，记为 $(X, Y) \sim U(G)$ 。

若二维随机变量 $(X, Y) \sim U(G)$ ，则对任意区域 $D \subset G$ ，有

$$P\{(X, Y) \in D\} = \iint_D f(x, y) d\sigma = \frac{1}{S(G)} \iint_D d\sigma = \frac{S(D)}{S(G)} \quad (3.1.7)$$

二维均匀分布描述的随机现象是向平面区域 G 中随机投点, 该点坐标 (X, Y) 落在 G 的任何子区域 D 中的概率只与该子区域的面积有关, 而与其位置无关. 上一章讨论了一维随机变量的均匀分布, 即

若随机变量 $X \sim U(a, b)$, 则对任意区间 $(c, d) \subset (a, b)$, 有

$$P\{c < X \leq d\} = \frac{d-c}{b-a} = \frac{(c, d) \text{ 的长度}}{(a, b) \text{ 的长度}} \quad (3.1.8)$$

随机变量 X 落在 (a, b) 的任何子区间内的概率只与该子区间的长度有关, 而与其位置无关. 在上述结论中, 借助几何度量 (长度、面积、体积等) 来计算的概率, 即为第 1 章中提到的几何概型.

例 3.1.4 设二维随机变量 (X, Y) 在区域 $G = \{(x, y) | 0 < x < 1, |y| < x\}$ 上服从均匀分布, 求

- (1) (X, Y) 的联合概率密度函数;
- (2) $P\{X+Y \leq 1\}$.

解 (1) $S(G) = 1$, $f(x, y) = \begin{cases} 1, & 0 < x < 1, |y| < x \\ 0, & \text{其他} \end{cases}$

(2) 设 $D = \{(x, y) | x + y \leq 1\}$, 如图 3.1.3 所示, $S(D) = \frac{3}{4}$,

$$P\{X+Y \leq 1\} = \frac{S(D)}{S(G)} = \frac{3}{4}$$

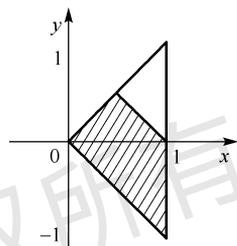


图 3.1.3 区域 D 示意图

例 3.1.5 (约会等待问题) 甲、乙两艘轮船驶向一个不能同时停泊两艘轮船的码头停泊, 它们在一昼夜内到达的时刻是等可能的. 如果甲的停泊时间为 1h, 乙的停泊时间为 2h, 求两艘轮船相遇的概率.

解 设 X 、 Y 分别表示甲、乙两船到达的时间,

则 $(X, Y) \sim U(G)$, $G = \{(x, y) | 0 \leq x \leq 24, 0 \leq y \leq 24\}$,

假设甲船先到, 则两船要相遇, 应满足 $X \leq Y \leq X+1$;

假设乙船先到, 则两船要相遇, 应满足 $Y \leq X \leq Y+2$.

所以两船要相遇, 即 (X, Y) 落在 $D = \{(x, y) | x-2 \leq y \leq x+1\}$ 内, 如图 3.1.4 所示,

$$p = \frac{S(D)}{S(G)} = \frac{24^2 - \frac{1}{2} \times 23^2 - \frac{1}{2} \times 22^2}{24^2} \approx 0.12$$

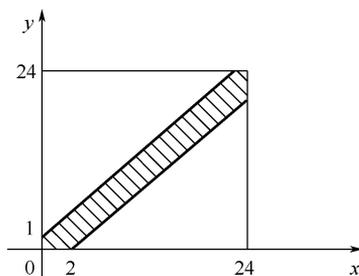


图 3.1.4 两船相遇示意图

2. 二维正态分布

设二维连续型随机变量 (X, Y) 具有联合概率密度函数

$$f(x, y) = \frac{1}{2\pi\sigma_1\sigma_2\sqrt{1-\rho^2}} \exp \left\{ \frac{-1}{2(1-\rho^2)} \left[\frac{(x-\mu_1)^2}{\sigma_1^2} - 2\rho \frac{(x-\mu_1)(y-\mu_2)}{\sigma_1\sigma_2} + \frac{(y-\mu_2)^2}{\sigma_2^2} \right] \right\} \quad (3.1.9)$$

$$x \in \mathbf{R}, \quad y \in \mathbf{R}$$

式中, $\mu_1, \mu_2, \sigma_1, \sigma_2, \rho$ 都是常数, 且 $\sigma_1 > 0, \sigma_2 > 0, -1 < \rho < 1$, 则称 (X, Y) 服从二维正态分

布, 记为 $(X, Y) \sim N(\mu_1, \sigma_1^2; \mu_2, \sigma_2^2; \rho)$.

3.2 边缘分布与独立性

3.2.1 边缘分布函数

X 和 Y 的分布函数 $F_X(x) = P\{X \leq x\}$ 、 $F_Y(y) = P\{Y \leq y\}$ 分别称为 (X, Y) 关于 X 和 Y 的**边缘分布函数**.

如果已知 (X, Y) 的联合分布函数 $F(x, y)$, 则由 $F(x, y)$ 可以导出 X 和 Y 的边缘分布函数为

$$F_X(x) = P\{X \leq x\} = P\{X \leq x, Y < +\infty\} = F(x, +\infty) = \lim_{y \rightarrow +\infty} F(x, y) \quad (3.2.1)$$

$$F_Y(y) = P\{Y \leq y\} = P\{X < +\infty, Y \leq y\} = F(+\infty, y) = \lim_{x \rightarrow +\infty} F(x, y) \quad (3.2.2)$$

联合分布函数 $F(x, y)$ 可以完全决定 X 和 Y 的边缘分布函数 $F_X(x)$ 、 $F_Y(y)$, 但反之不一定, X 和 Y 的边缘分布函数 $F_X(x)$ 、 $F_Y(y)$ 不能完全决定联合分布函数.

例 3.2.1 设二维随机变量 (X, Y) 的联合分布函数为

$$F(x, y) = \begin{cases} 1 - e^{-x} - e^{-y} + e^{-x-y-\lambda xy}, & x > 0, y > 0 \\ 0, & \text{其他} \end{cases}$$

求 X 和 Y 的边缘分布函数.

解 X 的边缘分布函数为

$$F_X(x) = \lim_{y \rightarrow +\infty} F(x, y) = \begin{cases} 1 - e^{-x}, & x > 0 \\ 0, & x \leq 0 \end{cases}$$

Y 的边缘分布函数为

$$F_Y(y) = \lim_{x \rightarrow +\infty} F(x, y) = \begin{cases} 1 - e^{-y}, & y > 0 \\ 0, & y \leq 0 \end{cases}$$

这两个边缘分布函数都是一维指数分布函数, 它们与 λ 无关. 当 λ 不同时, 对应的二维联合分布函数不同, 但它们的边缘分布函数却相同. 这说明, 仅利用边缘分布函数还不足以完全描述联合分布函数. 这是因为二维随机变量不仅与两个分量有关, 还与各分量间的联系有关.

X_i 的分布函数 $F_{X_i}(x_i) = P\{X_i \leq x_i\}$, $i = 1, 2, \dots, n$, 称为 (X_1, X_2, \dots, X_n) 关于 X_i 的**边缘分布函数**.

3.2.2 边缘分布律

对于二维离散型随机变量 (X, Y) , $P\{X = x_i\}$ 称为随机变量 X 的**边缘分布律**.

$$P\{X = x_i\} = p_{i\cdot} = \sum_{j=1}^{\infty} p_{ij}, \quad (i = 1, 2, \dots) \quad (3.2.3)$$

对于二维离散型随机变量 (X, Y) , $P\{Y = y_j\}$ 称为随机变量 Y 的**边缘分布律**.

$$P\{Y = y_j\} = p_{\cdot j} = \sum_{i=1}^{\infty} p_{ij}, \quad (j = 1, 2, \dots) \quad (3.2.4)$$

边缘分布律与联合分布律可以表示为如表 3.2.1 所示的表格形式.

表 3.2.1 边缘分布律与联合分布律

X	Y					
	y_1	y_2	\dots	y_j	\dots	$p_{i\cdot}$
x_1	p_{11}	p_{12}	\dots	p_{1j}	\dots	$p_{1\cdot}$
x_2	p_{21}	p_{22}	\dots	p_{2j}	\dots	$p_{2\cdot}$
\vdots	\vdots	\vdots		\vdots		\vdots
x_i	p_{i1}	p_{i2}	\dots	p_{ij}	\dots	$p_{i\cdot}$
\vdots	\vdots	\vdots		\vdots		\vdots
$p_{\cdot j}$	$p_{\cdot 1}$	$p_{\cdot 2}$	\dots	$p_{\cdot j}$	\dots	1

由表 3.2.1 可见, 将表格中各行的元素相加, 置于最后一列, 构成随机变量 X 的边缘分布律; 将表格中各列的元素相加, 置于最后一行, 构成随机变量 Y 的边缘分布律.

例 3.2.2 袋子中有 3 个球, 分别标有数字 1、2、3, 从中任取两次, 每次取 1 个球, 以 X 、 Y 分别表示第 1 次和第 2 次取到的球上的数字, 在无放回取球和有放回取球的两种情形下, 求 (X, Y) 的联合分布律和边缘分布律.

解 无放回取球的情形如表 3.2.2 所示.

表 3.2.2 无放回取球的情形

X	Y		
	1	2	$p_{i\cdot}$
1	0	$\frac{1}{3}$	$\frac{1}{3}$
2	$\frac{1}{3}$	$\frac{1}{3}$	$\frac{2}{3}$
$p_{\cdot j}$	$\frac{1}{3}$	$\frac{2}{3}$	1

有放回取球的情形如表 3.2.3 所示.

表 3.2.3 有放回取球的情形

X	Y		
	1	2	$p_{i\cdot}$
1	$\frac{1}{9}$	$\frac{2}{9}$	$\frac{1}{3}$
2	$\frac{2}{9}$	$\frac{4}{9}$	$\frac{2}{3}$
$p_{\cdot j}$	$\frac{1}{3}$	$\frac{2}{3}$	1

根据上例可以看出, 由 (X, Y) 的联合分布律可以完全确定 X 、 Y 的边缘分布律, 但由 X 、 Y 的边缘分布律不能完全确定 (X, Y) 的联合分布律. 在上例中 X 、 Y 的边缘分布律都相同, 但两种情况却有不同联合分布律. 所以二维随机变量的分布不仅与两个分量有关, 还与各分量间的联系有关.

3.2.3 边缘概率密度函数

X 和 Y 的概率密度函数 $f_X(x)$ 、 $f_Y(y)$ 分别称为 (X, Y) 关于 X 、 Y 的**边缘概率密度函数**. 随机变量 X 的边缘概率密度函数为

$$f_X(x) = \int_{-\infty}^{+\infty} f(x, y) dy, \quad x \in \mathbf{R} \quad (3.2.5)$$

随机变量 Y 的边缘概率密度函数为

$$f_Y(y) = \int_{-\infty}^{+\infty} f(x, y) dx, \quad y \in \mathbf{R} \quad (3.2.6)$$

因为

$$F_X(x) = F(x, +\infty) = \int_{-\infty}^x \left[\int_{-\infty}^{+\infty} f(u, v) dv \right] du$$

$$f_X(x) = F'_X(x) = \int_{-\infty}^{+\infty} f(x, v) dv, \quad \text{即 } f_X(x) = \int_{-\infty}^{+\infty} f(x, y) dy$$

同理

$$f_Y(y) = F'_Y(y) = \int_{-\infty}^y \left[\int_{-\infty}^{+\infty} f(u, v) du \right] dv = \int_{-\infty}^{+\infty} f(u, y) du = \int_{-\infty}^{+\infty} f(x, y) dx$$

例 3.2.3 设二维随机变量 (X, Y) 的联合概率密度函数为

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

求: (1) 边缘概率密度函数 $f_X(x)$;

(2) 边缘概率密度函数 $f_Y(y)$.

$$\text{解} \quad (1) \quad f_X(x) = \int_{-\infty}^{+\infty} f(x, y) dy = \begin{cases} \int_0^{+\infty} 2e^{-2x} e^{-y} dy, & x > 0 \\ 0, & x \leq 0 \end{cases} = \begin{cases} 2e^{-2x}, & x > 0 \\ 0, & x \leq 0 \end{cases}$$

$$(2) \quad f_Y(y) = \int_{-\infty}^{+\infty} f(x, y) dx = \begin{cases} \int_0^{+\infty} 2e^{-2x} e^{-y} dx, & y > 0 \\ 0, & y \leq 0 \end{cases} = \begin{cases} e^{-y}, & y > 0 \\ 0, & y \leq 0 \end{cases}$$

定理 3.2.1 若二维随机变量 $(X, Y) \sim N(\mu_1, \sigma_1^2; \mu_2, \sigma_2^2; \rho)$, 则 $X \sim N(\mu_1, \sigma_1^2)$, $Y \sim N(\mu_2, \sigma_2^2)$. 证略.

该定理说明随机变量 (X, Y) 服从二维正态分布, 则 X 、 Y 的边缘分布函数均为一维正态分布函数.

推论 3.2.1 若二维随机变量 $(X, Y) \sim N(0, 1; 0, 1; \rho)$, 则 $X \sim N(0, 1)$, $Y \sim N(0, 1)$.

例 3.2.4 已知二维随机变量 (X, Y) 在区域 $G = \{(x, y) | x^2 + y^2 \leq 1\}$ 上服从二维均匀分布, 试求关于 X 、 Y 的边缘概率密度函数.

解 (X, Y) 的联合概率密度函数为

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

X 的边缘概率密度函数为

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

Y 的边缘概率密度函数为

$$f_Y(y) = \int_{-\infty}^{+\infty} f(x, y) dx = \begin{cases} \int_{-\sqrt{1-y^2}}^{+\sqrt{1-y^2}} \frac{1}{\pi} dx, & -1 \leq y \leq 1 \\ 0, & \text{其他} \end{cases} = \begin{cases} \frac{2}{\pi} \sqrt{1-y^2}, & -1 \leq y \leq 1 \\ 0, & \text{其他} \end{cases}$$

此例说明二维均匀随机变量的边缘分布不一定是均匀分布; 联合分布可以确定边缘分布, 但边缘分布不一定能确定联合分布.

3.2.4 随机变量的独立性

在第 1 章中我们介绍过随机事件的独立性, 下面我们借助随机事件的独立性, 引入随机变量的独立性.

定义 3.2.1 设 (X, Y) 为二维随机变量, 若对任意实数对 (x, y) , 有

$$P\{X \leq x, Y \leq y\} = P\{X \leq x\}P\{Y \leq y\} \quad (3.2.7)$$

成立, 则称随机变量 X 与 Y 相互独立.

如果两个事件 A 和 B 相互独立, 则 $P(AB) = P(A)P(B)$, 把 $P\{X \leq x\}$ 和 $P\{Y \leq y\}$ 分别看作两个事件 A 、 B , 则根据事件独立性就可得出上述定义.

定理 3.2.2 设 (X, Y) 为二维随机变量, 其联合分布函数为 $F(x, y)$, X 和 Y 的边缘分布函数分别为 $F_X(x)$ 、 $F_Y(y)$, 则 X 与 Y 相互独立的充分必要条件是对一切实数对 (x, y) , 都有

$$F(x, y) = F_X(x) \cdot F_Y(y) \quad (3.2.8)$$

定理 3.2.3 设 (X, Y) 为二维离散型随机变量, 则 X 与 Y 相互独立的充分必要条件是对 (X, Y) 的任意一对取值 (x_i, y_j) , 都有

$$P\{X = x_i, Y = y_j\} = P\{X = x_i\} \cdot P\{Y = y_j\}, \quad i, j = 1, 2, \dots \quad (3.2.9)$$

定理 3.2.4 设 (X, Y) 为二维连续型随机变量, 其联合概率密度函数为 $f(x, y)$, X 和 Y 的边缘概率密度函数分别为 $f_X(x)$ 、 $f_Y(y)$, 则 X 与 Y 相互独立的充分必要条件是对 $f(x, y)$ 、 $f_X(x)$ 、 $f_Y(y)$ 的一切公共连续点, 都有

$$f(x, y) = f_X(x) \cdot f_Y(y) \quad (3.2.10)$$

随机变量的独立性具有与随机事件的独立性相同的性质.

(1) 若 X_1, X_2, \dots, X_n 相互独立, 则其中任意 m ($2 \leq m \leq n$) 个随机变量也相互独立, 反之不成立.

(2) 若 X 与 Y 相互独立, $h(\cdot)$ 和 $g(\cdot)$ 是连续函数, 则 $h(X)$ 与 $g(Y)$ 也相互独立.

(3) 若 (X_1, X_2, \dots, X_m) 与 $(X_{m+1}, X_{m+2}, \dots, X_n)$ 相互独立, $h(\cdot)$ 和 $g(\cdot)$ 是连续函数, 则 $h(X_1, X_2, \dots, X_m)$ 与 $g(X_{m+1}, X_{m+2}, \dots, X_n)$ 也相互独立.

例 3.2.5 在例 3.2.2 中, 在无放回取球和有放回取球的两种情形下讨论随机变量 X 与 Y 的独立性.

解 无放回取球的情形如表 3.2.4 所示.

表 3.2.4 无放回取球的情形

X	Y		
	1	2	$p_{i\cdot}$
1	0	$\frac{1}{3}$	$\frac{1}{3}$
2	$\frac{1}{3}$	$\frac{1}{3}$	$\frac{2}{3}$
$p_{\cdot j}$	$\frac{1}{3}$	$\frac{2}{3}$	1

$P\{X=1, Y=1\} \neq P\{X=1\} \cdot P\{Y=1\}$, X 与 Y 不相互独立.

有放回取球的情形如表 3.2.5 所示.

表 3.2.5 有放回取球的情形

X	Y		
	1	2	$p_{i\cdot}$
1	$\frac{1}{9}$	$\frac{2}{9}$	$\frac{1}{3}$
2	$\frac{2}{9}$	$\frac{4}{9}$	$\frac{2}{3}$
$p_{\cdot j}$	$\frac{1}{3}$	$\frac{2}{3}$	1

$P\{X=i, Y=j\} = P\{X=i\} \cdot P\{Y=j\}$, $i, j=1, 2$, X 与 Y 相互独立.

例 3.2.6 已知二维随机变量 (X, Y) 的联合概率密度函数为

$$f(x, y) = \begin{cases} 8xy, & 0 \leq x \leq y \leq 1 \\ 0, & \text{其他} \end{cases}$$

讨论 X 与 Y 的独立性.

$$\text{解} \quad f_X(x) = \int_{-\infty}^{+\infty} f(x, y) dy = \begin{cases} \int_x^1 8xy dy, & 0 \leq x \leq 1 \\ 0, & \text{其他} \end{cases} = \begin{cases} 4x(1-x^2), & 0 \leq x \leq 1 \\ 0, & \text{其他} \end{cases}$$

$$f_Y(y) = \int_{-\infty}^{+\infty} f(x, y) dx = \begin{cases} \int_0^y 8xy dx, & 0 \leq y \leq 1 \\ 0, & \text{其他} \end{cases} = \begin{cases} 4y^3, & 0 \leq y \leq 1 \\ 0, & \text{其他} \end{cases}$$

由 $f_X(x) \cdot f_Y(y) \neq f(x, y)$, 则 X 与 Y 不相互独立.

例 3.2.7 设甲、乙两种元件的寿命 X 、 Y 相互独立, 服从同一分布, 其概率密度函数为

$$f(x) = \begin{cases} \frac{1}{2} e^{-\frac{x}{2}}, & x > 0 \\ 0, & \text{其他} \end{cases}$$

求甲元件寿命不大于乙元件寿命 2 倍的概率.

解 由于 X 、 Y 服从同一分布且相互独立, 则 X 、 Y 的联合概率密度函数 $f(x, y)$ 为

$$f(x, y) = f_X(x)f_Y(y) = \begin{cases} \frac{1}{4} e^{-\frac{x+y}{2}}, & x > 0, y > 0 \\ 0, & \text{其他} \end{cases}$$

故所求概率为

$$P\{X \leq 2Y\} = \int_0^{+\infty} dx \int_{\frac{x}{2}}^{+\infty} \frac{1}{4} e^{-\frac{x+y}{2}} dy = \int_0^{+\infty} \frac{1}{2} e^{-\frac{x}{2}} e^{-\frac{x}{4}} dx = \int_0^{+\infty} \frac{1}{2} e^{-\frac{3x}{4}} dx = \frac{2}{3}$$

即甲元件寿命不大于乙元件寿命 2 倍的概率为 $\frac{2}{3}$.

例 3.2.8 证明: 已知二维随机变量 $(X, Y) \sim N(\mu_1, \sigma_1^2; \mu_2, \sigma_2^2; \rho)$, 则 X 与 Y 相互独立的充分必要条件是 $\rho = 0$.

证 由定理 3.2.1 可知, 若 $(X, Y) \sim N(\mu_1, \sigma_1^2; \mu_2, \sigma_2^2; \rho)$, 则

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

(X, Y) 的联合概率密度函数为

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

X 、 Y 的边缘概率密度函数分别为

$$f_X(x) = \frac{1}{\sqrt{2\pi}\sigma_1} \exp\left\{-\frac{1}{2} \left[\frac{(x-\mu_1)^2}{\sigma_1^2} \right]\right\}$$

$$f_Y(y) = \frac{1}{\sqrt{2\pi}\sigma_2} \exp\left\{-\frac{1}{2} \left[\frac{(y-\mu_2)^2}{\sigma_2^2} \right]\right\}$$

若 $\rho = 0$, 则

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

即 X 与 Y 相互独立.

反之, 若 X 与 Y 相互独立, $f(x, y) = f_X(x) \cdot f_Y(y)$ 对一切 $(x, y) \in \mathbf{R}^2$ 成立,

令 $x = \mu_1$, $y = \mu_2$, 有 $\frac{1}{\sqrt{2\pi}\sigma_1} \cdot \frac{1}{\sqrt{2\pi}\sigma_2} = \frac{1}{2\pi\sigma_1\sigma_2\sqrt{1-\rho^2}}$, 故 $\rho = 0$.

例 3.2.9 某公司老板到达办公室的时间均匀分布在 8:00~12:00, 他的秘书到达办公室的

时间均匀分布在 7:00~9:00, 假设两人到达的时间相互独立, 求两人到达办公室的时间相差不超过 5min 的概率.

解 设 X 、 Y 分别是老板和他的秘书到达办公室的时间, 由假设 X 和 Y 的概率密度函数分别为

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

因为 X 、 Y 相互独立, 故 (X, Y) 的联合概率密度函数为

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

两人到达办公室的时间相差不超过 5min, 即 $|X - Y| \leq \frac{5}{60}$,

如图 3.2.1 所示, 其概率为

$$P\left\{|X - Y| \leq \frac{5}{60}\right\} = \frac{S(D)}{S(G)} = \frac{\frac{1}{2} \times \left(\frac{13}{12}\right)^2 - \frac{1}{2} \times \left(\frac{11}{12}\right)^2}{4 \times 2} = \frac{1}{48}$$

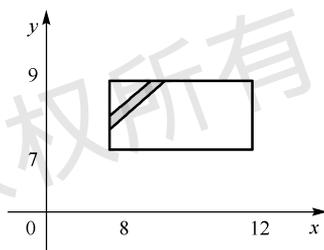


图 3.2.1 X 、 Y 取值范围

即老板和他的秘书到达办公室的时间相差不超过 5min 的概率为 $\frac{1}{48}$.

3.3 条件分布

3.3.1 条件分布律

定义 3.3.1 设二维离散型随机变量 (X, Y) 的联合分布律为

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

对固定的 j , 若 $P\{Y = y_j\} > 0$, 则称

$$P\{X = x_i | Y = y_j\} = \frac{P\{X = x_i, Y = y_j\}}{P\{Y = y_j\}} = \frac{p_{ij}}{p_{.j}} \quad (i = 1, 2, \dots) \quad (3.3.1)$$

为在 $Y = y_j$ 的条件下, 随机变量 X 的**条件分布律**.

对固定的 i , 若 $P\{X = x_i\} > 0$, 则称

$$P\{Y = y_j | X = x_i\} = \frac{P\{X = x_i, Y = y_j\}}{P\{X = x_i\}} = \frac{p_{ij}}{p_{i.}} \quad (j = 1, 2, \dots) \quad (3.3.2)$$

为在 $X = x_i$ 的条件下, 随机变量 Y 的**条件分布律**.

条件分布律 $P\{X = x_i | Y = y_j\}$ 满足分布律的性质.

$$(1) P\{X = x_i | Y = y_j\} > 0, (i = 1, 2, \dots);$$

$$(2) \sum_{i=1}^{\infty} P\{X = x_i | Y = y_j\} = \frac{1}{p_{.j}} \sum_{i=1}^{\infty} p_{ij} = 1.$$

条件分布律 $P\{Y = y_j | X = x_i\}$ 类似.

例 3.3.1 设随机变量 X 在 1、2、3、4 四个整数中等可能地取值, 另一随机变量 Y 在 $1 \sim X$ 中等可能地取值, 求

(1) (X, Y) 的联合分布律;

(2) X 、 Y 的边缘分布律;

(3) 判断 X 、 Y 的独立性.

解 (1) X 的分布律如表 3.3.1 所示.

表 3.3.1 X 的分布律

X	1	2	3	4
P	$\frac{1}{4}$	$\frac{1}{4}$	$\frac{1}{4}$	$\frac{1}{4}$

在 $X = i$ 的条件下, Y 的条件分布律为

$$P\{Y = j | X = i\} = \begin{cases} \frac{1}{i}, & j \leq i \\ 0, & j > i \end{cases} \quad (i, j = 1, 2, 3, 4)$$

(X, Y) 的联合分布律为

$$p_{ij} = P\{X = i, Y = j\} = P\{X = i\} P\{Y = j | X = i\} = \begin{cases} \frac{1}{4} \cdot \frac{1}{i}, & j \leq i \\ 0, & j > i \end{cases} \quad (i, j = 1, 2, 3, 4)$$

(X, Y) 的联合分布律用表格表示, 如表 3.3.2 所示.

表 3.3.2 (X, Y) 的联合分布律

X	Y				$P_{i.}$
	1	2	3	4	
1	$\frac{1}{4}$	0	0	0	$\frac{1}{4}$
2	$\frac{1}{8}$	$\frac{1}{8}$	0	0	$\frac{1}{4}$
3	$\frac{1}{12}$	$\frac{1}{12}$	$\frac{1}{12}$	0	$\frac{1}{4}$
4	$\frac{1}{16}$	$\frac{1}{16}$	$\frac{1}{16}$	$\frac{1}{16}$	$\frac{1}{4}$
$P_{.j}$	$\frac{25}{48}$	$\frac{13}{48}$	$\frac{7}{48}$	$\frac{3}{48}$	1

(2) Y 的分布律如表 3.3.3 所示.

表 3.3.3 Y 的分布律

Y	1	2	3	4
P	$\frac{25}{48}$	$\frac{13}{48}$	$\frac{7}{48}$	$\frac{3}{48}$

(3) 由于 $P\{X=i, Y=j\} \neq P\{X=i\} \cdot P\{Y=j\}$, 所以 X 与 Y 不相互独立.

3.3.2 条件概率密度函数

定义 3.3.2 设二维连续型随机变量 (X, Y) 的联合概率密度函数为 $f(x, y)$, 对一切使 $f_Y(y) > 0$ 的 y , 分别称

$$F_{X|Y}(x|y) = \int_{-\infty}^x \frac{f(u, y)}{f_Y(y)} du \quad (3.3.3)$$

$$f_{X|Y}(x|y) = \frac{f(x, y)}{f_Y(y)} \quad (3.3.4)$$

为在 $Y=y$ 的条件下, 随机变量 X 的**条件分布函数**和**条件概率密度函数**.

对一切使 $f_X(x) > 0$ 的 x , 分别称

$$F_{Y|X}(y|x) = \int_{-\infty}^y \frac{f(x, v)}{f_X(x)} dv \quad (3.3.5)$$

$$f_{Y|X}(y|x) = \frac{f(x, y)}{f_X(x)} \quad (3.3.6)$$

为在 $X=x$ 的条件下, 随机变量 Y 的**条件分布函数**和**条件概率密度函数**.

例 3.3.2 已知二维随机变量 (X, Y) 在区域 $G = \{(x, y) | x^2 + y^2 \leq 1\}$ 上服从二维均匀分布, 试求在 $Y=y$ 的条件下, X 的条件概率密度函数 $f_{X|Y}(x|y)$.

解 (X, Y) 的联合概率密度函数为

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

Y 的边缘概率密度函数为

$$f_Y(y) = \int_{-\infty}^{+\infty} f(x, y) dx = \begin{cases} \frac{2}{\pi} \sqrt{1-y^2}, & -1 \leq y \leq 1 \\ 0, & \text{其他} \end{cases}$$

当 $-1 < y < 1$ 时, 在 $Y=y$ 的条件下, X 的条件概率密度函数为

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

3.4 二维随机变量函数的分布

在第2章中讨论过一维随机变量函数 $Y = g(X)$ 的分布, 本节进一步讨论二维随机变量函数 $Z = f(X, Y)$ 的分布. 假设 (X, Y) 为二维随机变量, $Z = f(X, Y)$ 是 (X, Y) 的函数, Z 也是随机变量.

3.4.1 二维离散型随机变量函数的分布

设二维离散型随机变量 (X, Y) 的联合分布律为

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

若 (X, Y) 的函数 $Z = f(X, Y)$ 仍是离散型随机变量, 则其分布律为

$$P\{Z = z_k\} = P\{f(X, Y) = z_k\} = \sum_{(x_i, y_j) \in S_k} P\{X = x_i, Y = y_j\} \quad (3.4.1)$$

式中, $S_k = \{(x_i, y_j) : f(x_i, y_j) = z_k\}$.

例 3.4.1 设二维随机变量 (X, Y) 的联合分布律如表 3.4.1 所示.

表 3.4.1 (X, Y) 的联合分布律

X	Y		
	-1	1	2
-1	$\frac{1}{10}$	$\frac{2}{10}$	$\frac{3}{10}$
2	$\frac{2}{10}$	$\frac{1}{10}$	$\frac{1}{10}$

求: (1) $X+Y$ 的分布律; (2) XY 的分布律; (3) X/Y 的分布律; (4) $\max(X, Y)$ 的分布律.

解 由 (X, Y) 联合分布律, 得 (X, Y) 及函数的分布如表 3.4.2 所示.

表 3.4.2 (X, Y) 及函数的分布

P	$\frac{1}{10}$	$\frac{2}{10}$	$\frac{3}{10}$	$\frac{2}{10}$	$\frac{1}{10}$	$\frac{1}{10}$
(X, Y)	$(-1, -1)$	$(-1, 1)$	$(-1, 2)$	$(2, -1)$	$(2, 1)$	$(2, 2)$
$X+Y$	-2	0	1	1	3	4
XY	1	-1	-2	-2	2	4
X/Y	1	-1	-1/2	-2	2	1
$\max(X, Y)$	-1	1	2	2	2	2

(1) $X+Y$ 的分布律如表 3.4.3 所示.

表 3.4.3 $X+Y$ 的分布律

$X+Y$	-2	0	1	3	4
P	$\frac{1}{10}$	$\frac{2}{10}$	$\frac{5}{10}$	$\frac{1}{10}$	$\frac{1}{10}$

(2) XY 的分布律如表 3.4.4 所示.

表 3.4.4 XY 的分布律

XY	-2	-1	1	2	4
P	$\frac{5}{10}$	$\frac{2}{10}$	$\frac{1}{10}$	$\frac{1}{10}$	$\frac{1}{10}$

(3) X/Y 的分布律如表 3.4.5 所示.

表 3.4.5 X/Y 的分布律

X/Y	-2	-1	-1/2	1	2
P	$\frac{2}{10}$	$\frac{2}{10}$	$\frac{3}{10}$	$\frac{2}{10}$	$\frac{1}{10}$

(4) $\max(X, Y)$ 的分布律如表 3.4.6 所示.

表 3.4.6 $\max(X, Y)$ 的分布律

$\max(X, Y)$	-1	1	2
P	$\frac{1}{10}$	$\frac{2}{10}$	$\frac{7}{10}$

例 3.4.2 (泊松分布的可加性) 设随机变量 X 与 Y 相互独立, 且 $X \sim P(\lambda_1)$, $Y \sim P(\lambda_2)$, 求 $X+Y$ 的分布律.

$$\text{解 } P\{X=k\} = \frac{\lambda_1^k}{k!} e^{-\lambda_1}, \quad P\{Y=h\} = \frac{\lambda_2^h}{h!} e^{-\lambda_2}, \quad (k, h = 0, 1, 2, \dots)$$

$$P\{X+Y=n\} = P\{X=0, Y=n\} + P\{X=1, Y=n-1\} + \dots + P\{X=n, Y=0\}$$

$$\begin{aligned} &= \sum_{k=0}^n P\{X=k\} P\{Y=n-k\} = \sum_{k=0}^n \frac{\lambda_1^k}{k!} e^{-\lambda_1} \frac{\lambda_2^{n-k}}{(n-k)!} e^{-\lambda_2} \\ &= \frac{e^{-\lambda_1-\lambda_2}}{n!} \sum_{k=0}^n \frac{n!}{k!(n-k)!} \lambda_1^k \lambda_2^{n-k} = \frac{e^{-\lambda_1-\lambda_2}}{n!} \sum_{k=0}^n C_n^k \lambda_1^k \lambda_2^{n-k} \\ &= \frac{(\lambda_1 + \lambda_2)^n}{n!} e^{-(\lambda_1 + \lambda_2)} \quad (n = 0, 1, 2, \dots) \end{aligned}$$

从而 $X+Y \sim P(\lambda_1 + \lambda_2)$.

两个相互独立的泊松分布随机变量之和仍然服从泊松分布, 且参数为相应的参数之和, 称泊松分布具有**可加性**.

类似地, 二项分布具有可加性: $X \sim B(n_1, p)$, $Y \sim B(n_2, p)$, 且相互独立, 则它们的和 $X+Y \sim B(n_1 + n_2, p)$.

两点分布具有可加性: $X_i \sim B(1, p)$, X_1, X_2, \dots, X_n 相互独立同分布, 则它们的和 $X_1 + X_2 + \dots + X_n \sim B(n, p)$. 二项分布可以看作 n 个相互独立的两点分布随机变量的和.

3.4.2 二维连续型随机变量函数的分布

设二维连续型随机变量 (X, Y) 的联合概率密度函数为 $f(x, y)$, 若 (X, Y) 的函数 $Z = f(X, Y)$ 仍为连续型随机变量, 则其分布函数为

$$F_Z(z) = P\{Z \leq z\} = P\{f(X, Y) \leq z\} = \iint_{\{(x, y) | G(x, y) \leq z\}} f(x, y) dx dy \quad (3.4.2)$$

概率密度函数为

$$f_Z(z) = \begin{cases} F'_Z(z), & \text{在 } f_Z(z) \text{ 的连续点} \\ 0, & \text{其他} \end{cases} \quad (3.4.3)$$

此为求随机变量函数的分布的基本方法，即从求 Z 的分布函数出发，将求 Z 的分布函数转化为求 (X, Y) 的事件的概率，求出 Z 的分布函数，再求其概率密度函数。

例 3.4.3 设随机变量 X 与 Y 相互独立，且都服从 $N(0, 1)$ ，求 $Z = \sqrt{X^2 + Y^2}$ 的概率密度函数。

解 (X, Y) 的联合概率密度函数为 $f(x, y) = \frac{1}{2\pi} e^{-\frac{x^2+y^2}{2}}$

当 $z < 0$ 时， $F_Z(z) = 0$

$$\begin{aligned} \text{当 } z \geq 0 \text{ 时, } F_Z(z) &= P\{Z \leq z\} = P\{\sqrt{X^2 + Y^2} \leq z\} = \iint_{\sqrt{x^2+y^2} \leq z} \frac{1}{2\pi} e^{-\frac{x^2+y^2}{2}} d\sigma \\ &= \frac{1}{2\pi} \int_0^{2\pi} d\theta \int_0^z e^{-\frac{r^2}{2}} r dr = \int_0^z e^{-\frac{r^2}{2}} r dr \end{aligned}$$

综上， Z 的分布函数为

$$F_Z(z) = \begin{cases} \int_0^z e^{-\frac{r^2}{2}} r dr, & z \geq 0 \\ 0, & z < 0 \end{cases}$$

Z 的概率密度函数为

$$f_Z(z) = F'_Z(z) = \begin{cases} ze^{-\frac{z^2}{2}}, & z \geq 0 \\ 0, & z < 0 \end{cases}$$

例 3.4.4 已知 (X, Y) 的联合概率密度函数为

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

求 $Z = X + Y$ 的概率密度函数 $f_Z(z)$ 。

解 设 $Z = X + Y$ 的分布函数为 $F_Z(z)$

当 $z \leq 0$ 时， $F_Z(z) = 0$ ；当 $z \geq 2$ 时， $F_Z(z) = 1$

当 $0 < z \leq 1$ 时，如图 3.4.1 所示。

$$F_Z(z) = P\{Z \leq z\} = P\{X + Y \leq z\} = \iint_{x+y \leq z} 3x dx dy = \int_0^{z/2} dy \int_y^{z-y} 3x dx = \frac{3}{8} z^3$$

当 $1 < z < 2$ 时，如图 3.4.2 所示。

$$\begin{aligned}
 F_Z(z) &= P\{Z \leq z\} = P\{X+Y \leq z\} = 1 - P\{X+Y > z\} \\
 &= 1 - \iint_{x+y>z} 3xdxdy = 1 - \int_{z/2}^1 dx \int_{z-x}^x 3xdy = \frac{3}{2}z - \frac{z^3}{8} - 1
 \end{aligned}$$

综上, Z 的分布函数为

$$F_Z(z) = \begin{cases} 0, & z \leq 0 \\ \frac{3}{8}z^3, & 0 < z \leq 1 \\ \frac{3}{2}z - \frac{z^3}{8} - 1, & 1 < z < 2 \\ 1, & z \geq 2 \end{cases}$$

Z 的概率密度函数为

$$f_Z(z) = F'_Z(z) = \begin{cases} \frac{9}{8}z^2, & 0 < z \leq 1 \\ \frac{3}{2}\left(1 - \frac{z^2}{4}\right), & 1 < z < 2 \\ 0, & \text{其他} \end{cases}$$

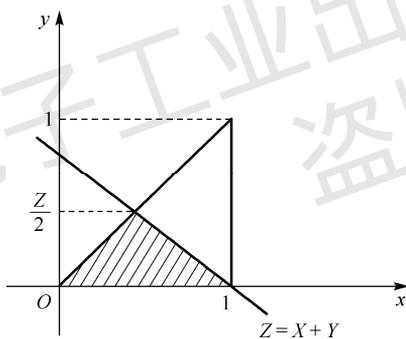


图 3.4.1 $0 < z \leq 1$ 时的积分区域

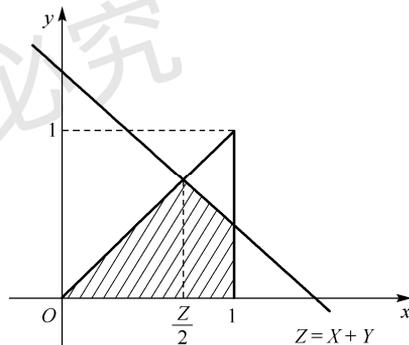


图 3.4.2 $1 < z < 2$ 时的积分区域

3.4.3 特殊函数的分布

下面讨论几种特殊函数的分布.

1. 和的分布 (卷积公式)

设二维随机变量 (X, Y) 的联合概率密度函数为 $f(x, y)$, X 和 Y 的边缘概率密度函数分别为 $f_X(x)$ 、 $f_Y(y)$, 则和 $Z = X + Y$ 的概率密度函数为

$$f_Z(z) = \int_{-\infty}^{+\infty} f(x, z-x)dx = \int_{-\infty}^{+\infty} f(z-y, y)dy \quad (3.4.4)$$

特别地, 当 X 与 Y 相互独立时

$$f_Z(z) = \int_{-\infty}^{+\infty} f_X(x)f_Y(z-x)dx = \int_{-\infty}^{+\infty} f_X(z-y)f_Y(y)dy \quad (3.4.5)$$

上述两个公式也称为卷积公式.

$$\begin{aligned} \text{证 } F_Z(z) &= P\{Z \leq z\} = P\{X+Y \leq z\} = \iint_{x+y \leq z} f(x,y)dxdy = \int_{-\infty}^{+\infty} \left[\int_{-\infty}^{z-y} f(x,y)dx \right] dy \\ &= \int_{-\infty}^{z-y} \left[\int_{-\infty}^{+\infty} f(x,y)dy \right] dx \end{aligned}$$

Z 的概率密度函数为

$$f_Z(z) = F'_Z(z) = \int_{-\infty}^{+\infty} f(z-y,y)dy$$

由 X 与 Y 的对称性, 得 $f_Z(z) = \int_{-\infty}^{+\infty} f(x,z-x)dx$.

例 3.4.5 例 3.4.4 也可用卷积公式求解.

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

由卷积公式 $f_Z(z) = \int_{-\infty}^{+\infty} f(x,z-x)dx$ 求解, 得

$$f(x,z-x) = \begin{cases} 3x, & 0 < x < 1, x < z < 2x \\ 0, & \text{其他} \end{cases}$$

积分区域如图 3.4.3 所示.

当 $z < 0$ 或 $z \geq 2$ 时, $f_Z(z) = 0$

$$\text{当 } 0 \leq z < 1 \text{ 时, } f_Z(z) = \int_{-\infty}^{+\infty} f(x,z-x)dx = \int_{z/2}^z 3xdx = \frac{9}{8}z^2$$

$$\text{当 } 1 \leq z < 2 \text{ 时, } f_Z(z) = \int_{-\infty}^{+\infty} f(x,z-x)dx = \int_{z/2}^1 3xdx = \frac{3}{2} \left(1 - \frac{z^2}{4} \right)$$

综上, Z 的概率密度函数为

$$f_Z(z) = \begin{cases} \frac{9}{8}z^2, & 0 \leq z < 1 \\ \frac{3}{2} \left(1 - \frac{z^2}{4} \right), & 1 \leq z < 2 \\ 0, & \text{其他} \end{cases}$$

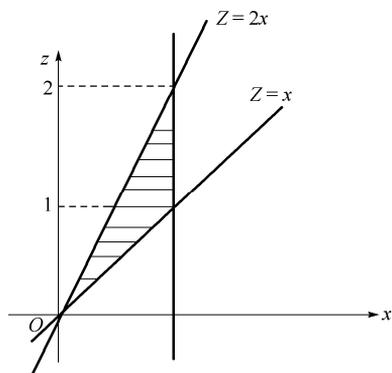


图 3.4.3 积分区域

例 3.4.6 设随机变量 X 与 Y 相互独立, $X \sim N(0,1)$, $Y \sim N(0,1)$, 求 $Z = X+Y$ 的概率密度函数.

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

令 $t = x - \frac{z}{2}$, 则

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

即 $Z \sim N(0, 2)$.

例 3.4.7 设二维随机变量 (X, Y) 在以点 $(0, 1)$ 、 $(1, 0)$ 、 $(1, 1)$ 为顶点的三角形区域上服从均匀分布, 求 $Z = X + Y$ 的概率密度函数.

解 (X, Y) 的联合概率密度函数为 $f(x, y) = \begin{cases} 2, & 0 \leq x \leq 1, 1-x \leq y \leq 1 \\ 0, & \text{其他} \end{cases}$

$$G = \{(x, z) | 0 \leq x \leq 1, 1-x \leq z-x \leq 1\} = \{(x, z) | 0 \leq x \leq 1, 1 \leq z \leq 1+x\}$$

当 $1 \leq z < 2$ 时, $f_Z(z) = \int_{-\infty}^{+\infty} f_X(x)f_Y(z-x)dx = \int_{z-1}^1 2dx = 2(2-z)$

当 $z < 1$ 或 $z \geq 2$ 时, $f_Z(z) = 0$

综上, Z 的概率密度函数为

$$f_Z(z) = \begin{cases} 2(2-z), & 1 \leq z < 2 \\ 0, & \text{其他} \end{cases}$$

2. 极值分布

设随机变量 X 与 Y 相互独立, 其分布函数分别为 $F_X(x)$ 、 $F_Y(y)$, 则最大值 $Z_1 = \max(X, Y)$ 和最小值 $Z_2 = \min(X, Y)$ 的分布函数分别为

$$\begin{aligned} F_{Z_1}(z) &= P\{\max(X, Y) \leq z\} = P\{X \leq z, Y \leq z\} \\ &= P\{X \leq z\}P\{Y \leq z\} = F_X(z)F_Y(z), \quad z \in \mathbf{R} \end{aligned} \quad (3.4.6)$$

$$\begin{aligned} F_{Z_2}(z) &= P\{\min(X, Y) \leq z\} = 1 - P\{\min(X, Y) > z\} \\ &= 1 - P\{X > z, Y > z\} = 1 - P\{X > z\}P\{Y > z\} \\ &= 1 - [1 - P\{X \leq z\}][1 - P\{Y \leq z\}] \\ &= 1 - [1 - F_X(z)][1 - F_Y(z)], \quad z \in \mathbf{R} \end{aligned} \quad (3.4.7)$$

3.5 应用实例

例 3.5.1 设某班车起点站上车人数 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 ($0 < p < 1$), 并且他们在中途下车与否是相互独立的, 用 Y 表示在中途下车的人数, 求二维随机变量 (X, Y) 的概率分布.

解 X 服从泊松分布 $P\{X = n\} = \frac{\lambda^n e^{-\lambda}}{n!}$, ($n = 0, 1, 2, \dots$)

在上车人数为 n 的条件下, 每位乘客下车与否看作一次伯努利试验, 下车人数 Y 服从二项分布

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

(X,Y) 的联合分布律

当 $m \leq n, n=1,2,\dots, m=0,1,2,\dots$ 时,

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

当 $n=m=0$ 时,

$$P(X=n, Y=m) = e^{-\lambda}$$

例 3.5.2 某足球队在任何长度为 t 的时间区间内得黄牌 (或红牌) 的次数 $N(t)$ 服从参数为 λt 的泊松分布, 记 X_i 为比赛进行 t_i min 后的得牌数, $i=1,2$ ($t_2 > t_1$), 试写出 (X_1, X_2) 的联合分布律.

解 $N(t)$ 服从参数为 λt 的泊松分布

$$P\{N(t)=k\} = \frac{e^{-\lambda t} (\lambda t)^k}{k!} \quad (k=0,1,2,\dots)$$

(X_1, X_2) 的联合分布律为

$$\begin{aligned} P_{ij} &= P\{X_1=i, X_2=j\} = P\{X_1=i\}P\{X_2=j|X_1=i\} \\ &= \frac{e^{-\lambda t_1} (\lambda t_1)^i}{i!} \cdot \frac{e^{-\lambda(t_2-t_1)} [\lambda(t_2-t_1)]^{j-i}}{(j-i)!} \quad (i=0,1,2,\dots, j=i,i+1,\dots) \end{aligned}$$

例 3.5.3 设系统 L 由两个相互独立的子系统 L_1, L_2 连接而成, 连接方式分别为 (1) 串联; (2) 并联; (3) 备用 (当系统 L_1 损坏时, 系统 L_2 开始工作), 如图 3.5.1 所示. 设 L_1, L_2 的寿命分别为 X, Y , 已知它们的概率密度函数分别为

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

试分别就以上三种连接方式写出 L 的寿命 Z 的概率密度函数.

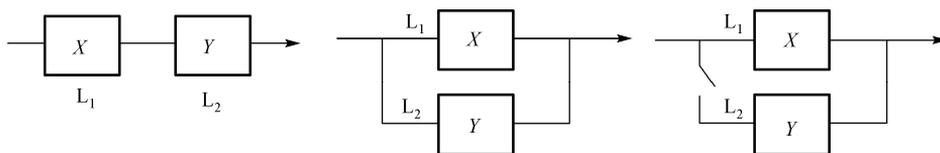


图 3.5.1 连接方式

解 (1) 串联的情况

由于当 L_1, L_2 中有一个损坏时, 系统 L 就停止工作, 所以 L 的寿命为 $Z = \min(X, Y)$, 而 X, Y 的分布函数分别为

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

故 Z 的分布函数为

$$F_{\min}(z) = 1 - [1 - F_X(z)][1 - F_Y(z)] = \begin{cases} 1 - e^{-(\alpha+\beta)z}, & z > 0 \\ 0, & z \leq 0 \end{cases}$$

Z 的概率密度函数为

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

即 Z 仍服从指数分布.

(2) 并联的情况

由于当且仅当 L_1 、 L_2 都损坏时, 系统 L 才停止工作, 所以 L 的寿命为 $Z = \max(X, Y)$, Z 的分布函数为

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

Z 的概率密度函数为

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

(3) 备用的情况

由于当系统 L_1 损坏时, 系统 L_2 才开始工作, 所以 L 的寿命 Z 是 L_1 、 L_2 的寿命之和, 即 $Z = X + Y$, 因此

当 $z \leq 0$ 时, $f_Z(z) = 0$

当 $z \geq 0$ 时,

$$\begin{aligned} f_Z(z) &= \int_{-\infty}^{+\infty} f_X(z-y)f_Y(y)dy = \int_0^z \alpha e^{-\alpha(z-y)} \beta e^{-\beta y} dy \\ &= \alpha \beta e^{-\alpha z} \int_0^z e^{-(\beta-\alpha)y} dy = \frac{\alpha \beta}{\beta - \alpha} [e^{-\alpha z} - e^{-\beta z}] \end{aligned}$$

即 Z 的概率密度函数为

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

例 3.5.4 设随机变量 X 、 Y 相互独立, X 的概率分布为 $P\{X=0\} = P\{x=2\} = \frac{1}{2}$, Y 的概率密度函数为

$$f_Y(y) = \begin{cases} 2y, & 0 < y < 1 \\ 0, & \text{其他} \end{cases}$$

求 $Z = X + Y$ 的概率密度函数.

解 设 $Z = X + Y$ 的分布函数为 $F_Z(z)$, 则

$$\begin{aligned}
 F_Z(z) &= P\{Z \leq z\} = P\{X+Y \leq z\} \\
 &= P\{X=0\} \cdot P\{X+Y \leq z | X=0\} + P\{X=2\} \cdot P\{X+Y \leq z | X=2\} \\
 &= \frac{1}{2} \cdot P\{Y \leq z | X=0\} + \frac{1}{2} \cdot P\{Y \leq z-2 | X=2\} \\
 &= \frac{1}{2} \cdot P\{Y \leq z\} + \frac{1}{2} \cdot P\{Y \leq z-2\} = \frac{1}{2} F_Y(z) + \frac{1}{2} F_Y(z-2)
 \end{aligned}$$

Z 的概率密度函数为 $f_Z(z) = F'_Z(z) = \frac{1}{2} f_Y(z) + \frac{1}{2} f_Y(z-2)$

$$f_Z(z) = \begin{cases} z, & 0 < z < 1 \\ z-2, & 2 < z < 3 \\ 0, & \text{其他} \end{cases}$$

习 题 3

1. 将两个元件并联组成一个电子部件, 两个元件的寿命分别为 X 与 Y (单位: h), 已知 (X, Y) 的联合分布函数为

$$F(x, y) = \begin{cases} 1 - e^{-0.01x} - e^{-0.01y} + e^{-0.01(x+y)}, & x \geq 0, y \geq 0 \\ 0, & \text{其他} \end{cases}$$

求: (1) X, Y 的边缘分布函数; (2) 此电子部件正常工作 120h 以上的概率.

2. 袋中有 5 个球, 分别标号 1、2、3、4、5, 现从中任取 3 个球, 设 X 和 Y 分别表示取出球的号码中的最大值和最小值, 求二维随机变量 (X, Y) 的联合分布律.

3. 将三封信投入 3 个编号分别为 1、2、3 的信箱, 用 X, Y 分别表示投入第 1、2 号信箱的信的数目. 求: (1) (X, Y) 的联合分布律; (2) X, Y 的边缘分布律; (3) 判断 X, Y 是否相互独立.

4. 设随机变量 $Z \sim U(-2, 2)$, 令 $X = \begin{cases} -1, & Z \leq -1 \\ 1, & Z > -1 \end{cases}$, $Y = \begin{cases} -1, & Z \leq 1 \\ 1, & Z > 1 \end{cases}$, 求二维随机变量 (X, Y) 的联合分布律.

5. 设二维随机变量 (X, Y) 的联合概率密度函数为

$$f(x, y) = \begin{cases} k(6-x-y), & 0 < x < 2, 2 < y < 4 \\ 0, & \text{其他} \end{cases}$$

求: (1) 常数 k ; (2) $P\{X < 1.5\}$; (3) $P\{X+Y \leq 4\}$.

6. 设二维随机变量 (X, Y) 的联合概率密度函数为

$$f(x, y) = \begin{cases} Cx^2y, & x^2 \leq y \leq 1 \\ 0, & \text{其他} \end{cases}$$

求: (1) 常数 C ; (2) X 的边缘概率密度函数 $f_X(x)$; (3) $P\{X \geq Y\}$.

7. 设二维随机变量 (X, Y) 的联合概率密度函数为

$$f(x,y) = \begin{cases} 12e^{-3x-4y}, & x > 0, y > 0 \\ 0, & \text{其他} \end{cases}$$

求: (1) (X,Y) 的联合分布函数 $F(x,y)$; (2) $P\{0 < X \leq 1, 0 < Y \leq 2\}$.

8. 设二元函数 $f(x,y)$ 为

$$f(x,y) = \begin{cases} \sin x \cos y, & 0 \leq x \leq \pi, C \leq y \leq \frac{\pi}{2} \\ 0, & \text{其他} \end{cases}$$

求当 C 取何值时, $f(x,y)$ 是二维随机变量的概率密度函数?

9. 在长为 a 的线段的中点两侧随机地各取一个点, 求两点间的距离小于 $\frac{a}{3}$ 的概率.

10. 设随机变量 X 与 Y 相互独立, X 在 $(0,1)$ 上服从均匀分布, Y 的概率密度函数为

$$f_Y(y) = \begin{cases} \frac{1}{2}e^{-\frac{y}{2}}, & y > 0 \\ 0, & \text{其他} \end{cases}$$

(1) 求 X 和 Y 的联合概率密度函数;

(2) 设关于 a 的二次方程 $a^2 + 2Xa + Y^2 = 0$, 求方程有实根的概率.

11. 设二维随机变量 (X,Y) 的联合分布律为

X	Y		
	-1	0	2
0	0.18	0.30	0.12
1	a	b	c

若 X 与 Y 相互独立, 求 a 、 b 、 c 的值.

12. 甲、乙两人各自独立地进行两次射击, 假设甲的命中率为 $1/5$, 乙的命中率为 $1/2$, 以 X 和 Y 分别表示甲和乙的命中次数, 求 $P\{X \leq Y\}$.

13. 设二维随机变量 (X,Y) 的联合概率密度函数为

$$f(x,y) = \begin{cases} \frac{1}{2}(x+y)e^{-(x+y)}, & x > 0, y > 0 \\ 0, & \text{其他} \end{cases}$$

问 X 与 Y 是否相互独立?

14. 设二维随机变量 (X,Y) 的联合概率密度函数为

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

问 X 与 Y 是否相互独立?

15. 设二维随机变量 (X,Y) 的联合分布律为

X	Y		
	1	2	3
1	0.1	0.3	0.2
2	0.2	0.05	0.15

求：(1) 在 $X=1$ 的条件下， Y 的条件分布律；

(2) 在 $Y=2$ 的条件下， X 的条件分布律.

16. 设二维随机变量 (X, Y) 的联合概率密度函数为

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

求条件概率密度函数 $f_{X|Y}(x|y)$ 、 $f_{Y|X}(y|x)$.

17. 设随机变量 X 与 Y 相互独立， $X \sim U(0, 2)$ ， $Y \sim \text{Exp}(1)$ ，求

(1) $P\{-1 < X < 1, 0 < Y < 2\}$ ；

(2) $P\{X+Y > 1\}$.

18. 设二维随机变量 (X, Y) 的联合分布律为

X	Y		
	1	2	3
1	$\frac{1}{4}$	$\frac{1}{4}$	$\frac{1}{8}$
2	$\frac{1}{8}$	0	0
3	$\frac{1}{8}$	$\frac{1}{8}$	0

求 $X+Y$ 、 $X-Y$ 、 $2X$ 、 XY 的分布律.

19. 设二维随机变量 (X, Y) 的联合概率密度函数为

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

求随机变量 $Z = X + 2Y$ 的分布函数和概率密度函数.

20. 设二维随机变量 (X, Y) 的联合概率密度函数为

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

求随机变量 $Z = X - Y$ 的概率密度函数.

21. 设二维随机变量 (X, Y) 的联合概率密度函数为

$$f(x, y) = \begin{cases} x+y, & 0 \leq x \leq 1, 0 \leq y \leq 1 \\ 0, & \text{其他} \end{cases}$$

求随机变量 $Z = X + Y$ 的概率密度函数.