# 多维随机变量及其分布

王友乐

电话：18980082143

邮箱：**Eleconnection@126.com**

2015 年 4 月 10 日

# 二维随机变量及其分布函数

**定义**

设 $X$ 与 $Y$ 是定义在同一样本空间 $\Omega$ 上的两个随机变量，则称 $(X, Y)$ 为**二维随机变量**.

**定义**

设 $(X, Y)$ 是二维随机变量，对任意实数 $x, y$，称

$$F(x, y) = P(X \leq x, Y \leq y) \tag{1}$$

为二维随机变量 $(X, Y)$ 的**二维分布函数**，或称 $X$ 与 $Y$ 的**联合分布函数**.

## 二维随机变量及其分布函数

### 定理

设 $F(x,y)$ 为二维随机变量 $(X,Y)$ 的二维分布函数，则

*(1)* $F(x,y)$ 分别关于 $x$ 及 $y$ 单调不减，即当 $x_1 < x_2$ 时，有 $F(x_1,y) \le F(x_2,y)$；当 $y_1 < y_2$ 时，有 $F(x,y_1) \le F(x,y_2)$.

*(2)* $F(-\infty,-\infty) = F(-\infty,y) = F(x,-\infty) = 0, F(+\infty,+\infty) = 1.$

*(3)* $F(x,y)$ 关于 $x$ 及 $y$ 都右连续，即对任意实数 $x,y$，有

$$F(x+0,y) = F(x,y), \quad F(x,y+0) = F(x,y).$$

*(4)* 对任意 $x_1 < x_2, y_1 < y_2$，有

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

## 二维随机变量及其分布函数

如果二维随机变量 $(X, Y)$ 只取有限个或可数无穷个点对 $(x_i, y_i), i, j = 1, 2, \cdots$ , 则称 $(X, Y)$ 为**二维离散型随机变量**.

**定义**

设二维离散型随机变量 $(X, Y)$ 所有可能取值为 $(x_i, y_i), i, j = 1, 2, \cdots$ , 记

$$p_{ij} = P(X = x_i, Y = y_j), \qquad i, j, = 1, 2, \cdots , \tag{2}$$

则称上式为 $(X, Y)$ 的**二维概率分布**或**分布律**, 或称为 $X$ 与 $Y$ 的**联合概率分布**.

(X,Y) 的二维概率分布

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

$X$ 与 $Y$ 的联合概率分布有如下性质: $(1) p_{ij} \geq 0, i, j = 1, 2, \cdots$ ;
$(2) \sum_{ij} p_{ij} = 1$.

在 $n$ 重独立试验中, 若每次试验只有 $A_1, A_2, A_3$ 三个可能的结果, 且 $0 < p_i = P(A_i) < 1, i = 1, 2, 3$, 则 $p_1 + p_2 + p_3 = 1$. 令随机变量 $X$ 及 $Y$ 分别表示 $n$ 次试验中 $A_1$ 与 $A_2$ 发生的次数, 则知 $X$ 与 $Y$ 有联合概率分布

$$P(X = k_1, Y = k_2) = \frac{n!}{k_1!k_2!(n - k_1 - k_2)!} p_1^{k_1} p_2^{k_2} p_3^{n-k_1-k_2} \tag{3}$$

$k_1 + k_2 = 0, 1, \cdots, n, k_1, k_2 \geq 0$, 并称 $(X, Y)$ 服从参数为 $p_1, p_2, n$ 的**三项分布**, 记为 $(X, Y) \sim T(n; p_1, p_2)$.

## 二维随机变量及其分布函数

### 定义

对二维随机变量 $(X,Y)$, 如果存在二元非负函数 $f(x,y)$, 使得对任意实数 $x,y$, 有

$$F(x,y)=\int_{-\infty}^{x}\int_{-\infty}^{y}f(x,y)dxdy \tag{4}$$

则称 $(X,Y)$ 是**二维连续性随机变量**, 称 $f(x,y)$ 为 $(X,Y)$ 的**二维概率密度函数**, 或称为 $X$ 和 $Y$ 的**联合密度函数**, 简称为**密度函数**或**密度**.

二维密度函数具有以下性质:

(1)$f(x,y)\geq 0,(x,y)\in\mathbb{R}^2$;

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

**定理**

二维连续性随机变量 $(X,Y)$ 有密度函数 $f(x,y)$, 则 *(1)*$F(x,y)$ 是连续函数且在 $f(x,y)$ 的连续点 $(x,y)$, 有

$$f(x,y) = \frac{\partial^2 F(x,y)}{\partial x \partial y}; \tag{5}$$

*(2)* 对平面上任意区域 $G \subset \mathbb{R}^2$, 若 $f(x,y)$ 在 $G$ 上可积, 有

$$P((X,Y) \in G) = \iint_G f(x,y) dx dy; \tag{6}$$

*(3)* 对平面上任一条曲线 $L$, 有

$$P((X,Y) \in L) = 0.$$

**定义**

令 $G$ 是平面上一个有界区域, 若二维随机变量 $(X,Y)$ 有密度函数

$$f(x,y)=\begin{cases}\frac{1}{m(G)}, & (x,y)\in G,\\ 0, & \text{其它}\end{cases}\tag{7}$$

其中 $m(G)$ 为 $G$ 的面积, 则称 $(X,Y)$ 在 $G$ 上**均匀分布**, 记为 $(X,Y)\sim U(G)$.

**定理**

设 $F(x,y)$ 为二维随机变量 $(X,Y)$ 的二维分布函数，则 $X$ 及 $Y$ 的边缘分布函数 $F_X(x), F_Y(y)$ 有

$$\begin{aligned} F_X(x) &= F(x,+\infty), x \in \mathbb{R}; \\ F_Y(y) &= F(+\infty, y), y \in \mathbb{R} \end{aligned} \tag{8}$$

**定理**

$F(x,y)$ 是二维随机变量 $(X,Y)$ 的二维分布函数,$F_X(x), F_Y(y)$ 分别为 $X,Y$ 的边缘分布函数. 若对任意 $x,y$, 有

$$F(x,y) = F_X(x)F_Y(y)$$

则称 $X$ 与 $Y$ **相对独立**.

**定理**

设随机变量 $X$ 与 $Y$ 相互独立, 且 $g(x)$ 与 $h(y)$ 分别是 $x$ 与 $y$ 的连续函数, 则 $X_1 = g(X)$, 与 $Y_1 = h(Y)$ 也相互独立.

## 边缘分布及随机变量的独立性

设 $(X,Y)$ 是二维离散型随机变量, 有二维概率分布

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

$X,Y$ 都是一维离散型随机变量, 有各自的分布律

$$p_{i\cdot}=P(X=x_i),i=1,2,\cdots.$$
$$p_{\cdot j}=P(Y=y_j),j=1,2,\cdots.$$

相对于二维概率分布,$X$ 与 $Y$ 各自的分布叫做**边缘概率分布**, 简称**边缘分布**.

## 边缘分布及随机变量的独立性

**定理**

$X$ 与 $Y$ 的边缘分布可由二维概率分布求出, 即

$$
\begin{aligned}
p_{i\cdot} &= \textstyle\sum_j p_{ij}, i = 1, 2, \cdots ; \\
p_{\cdot j} &= \textstyle\sum_i p_{ij}, j = 1, 2, \cdots .
\end{aligned} \tag{9}
$$

**定理**

随机变量 $X$ 与 $Y$ 独立 $\Longleftrightarrow$

$$p_{ij} = p_{i\cdot}p_{\cdot j}, i, j = 1, 2, \cdot .$$

**定理**

$(X, Y)$ 的二维密度函数为 $f(x, y)$,$X$ 与 $Y$ 的边缘密度函数分别为 $f_X(x), f_Y(y)$, 则

$$\begin{aligned} f_X(x) &= \int_{-\infty}^{+\infty} f(x, y)dy; \\ f_Y(y) &= \int_{-\infty}^{+\infty} f(x, y)dx. \end{aligned} \tag{10}$$

**定理**

$X$ 与 $Y$ 相互独立 $\Longleftrightarrow$

$$f(x, y) = f_X(x)f_Y(y) \tag{11}$$

在三个密度函数的公共连续点上成立.

## 条件分布与条件密度

当 $P(Y=y)>0$ 时, 记

$$F_{X|Y}=P(X\leq x|Y=y), x\in\mathbb{R} \tag{12}$$

称为 $Y=y$ 条件下 $X$ 的**条件分布函数**.

**定义**

$y_j$ 任意给定，且 $P(Y=y_j)=p_{\cdot j}>0$，称

$$P(X=x_i|Y=y_j)=\frac{p_{ij}}{p_{\cdot j}}, i=1,2,\cdots \tag{13}$$

为 $Y=y_j$ 条件下 $X$ 的**条件概率分布**.

同理，也可定义 $P(X=x_i)=p_{i\cdot}>0$,

$$P(Y=y_j|X=x_i)=\frac{p_{ij}}{p_{\cdot j}} \tag{14}$$

为 $X=x_i$ 条件下 $Y$ 的**条件概率分布**.

# 条件概率分布

条件分布性质:

$$\begin{aligned}(1)&P(X=x_i|Y=y_j)\geq 0, i=1,2,\cdots;\\(2)&\sum_i P(X=x_i|Y=y_j)=\sum_i \frac{p_{ij}}{p_{\cdot j}}=1.\end{aligned}\tag{15}$$

## 条件概率分布

### 定义

设 $(X,Y)$ 为二维连续型随机变量，若对任意 $\varepsilon > 0$，有 $P(y-\varepsilon < Y \le y+\varepsilon) > 0$，且对每个实数 $x$，极限

$$\lim_{\varepsilon \to 0+} P(X \le x | y-\varepsilon < Y \le y+\varepsilon) \tag{16}$$

存在，则定义此极限为连续型随机变量 $Y=y$ 条件下 $X$ 的**条件分布函数**，记为 $F_{X|Y}(x|y) = P(X \le x | Y=y) = \lim_{\varepsilon \to 0+} P(X \le x | y-\varepsilon < Y \le y+\varepsilon)$

## 条件概率分布

**定理**

设 $(X,Y)$ 有二维密度 $f(x,y)$, 从而 $X$ 及 $Y$ 有边缘密度 $f_X(x), f_Y(y)$, 则

$$\begin{aligned} F_{X|Y}(x|y) = \int_{-\infty}^{x} \frac{f(x,y)}{f_Y(y)} dy, x \in \mathbb{R} \\ F_{Y|X}(y|x) = \int_{-\infty}^{y} \frac{f(x,y)}{f_X(x)} dy, y \in \mathbb{R} \end{aligned} \tag{17}$$

**推论**

条件密度为

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

# 二维随机变量函数的分布

求 $f_Z(z)$ 的一般方法:

(1) 确定 $Z$ 的值域 $R(Z)$.

(2) 对任意 $z \in R(Z)$, 求 $Z$ 的分布函数 $F_Z(z)$, 即

$$F_Z(x) = P(Z \leq z) = P(g(X,Y) \leq z) = P((X,Y) \in G(z))$$

$$= \iint_{G(z)} f(x,y)dxdy.$$

其中 $(X,Y) \in G(Z) \subset \mathbb{R}^2$ 由不等式 $g(X,Y) \leq z$ 等价解除.

(3)$f_z(z) = F'_Z(z)$.

若 $X_1, X_2, \cdots, X_n$ 为 $n$ 个定义在同一样本空间 $\Omega$ 上的随机变量, 则称 $(X_1, X_2, \cdots, X_n)$ 为 $n$ **维随机变量**.

令随机变量 $X_1, X_2, \cdots, X_n$ 相互独立, 且 $X_i$ 有分布函数 $F_i(x_i), i = 1, 2, \cdots, n$. 再令 $M = \max\{X_1, X_2, \cdots, X_n\}$, $N = \min\{X_1, X_2, \cdots, X_n\}$, 则 $M$ 与 $N$ 的分布函数如下:

$$
\begin{aligned}
F_M(x) =& P(\max\{X_1, X_2, \cdots, X_n\} \leq x) \\
=& P(X_1 \leq x, X_2 \leq x, \cdots, X_n \leq x) \\
=& P(X_1 \leq x) P(X_2 \leq x) \cdots P(X_n \leq x) \\
=& F_1(x) F_2(x) \cdots F_n(x)
\end{aligned}
$$

# 多维随机变量

$$
\begin{aligned}
F_N(x) &= P(\min\{X_1, X_2, \cdots, X_n \leq x\}) \\
&= 1 - P(\min\{X_1, X_2, \cdots, X_n\} > x) \\
&= 1 - P(X_1 > x, X_2 > x, \cdots, X_n > x) \\
&= 1 - P(X_1 > x)P(X_2 > 2) \cdots P(X_n > x) \\
&= 1 - [1 - F_1(x)][1 - F_2(x)] \cdots [1 - F_n(x)].
\end{aligned}
$$