其中 $k_1, k_2, \cdots, k_n$ 是非负的整数且满足 $k_1 + k_2 + \cdots + k_n = m$, 则称随机向量 $(X_1, X_2, \cdots, X_n)$ 服从参数为 $m, p_1, p_2, \cdots, p_n$ 的 **多项分布** (multinomial distribution), 记为 $(X_1, X_2, \cdots, X_n) \sim M(m, p_1, p_2, \cdots, p_n)$.

很容易验证 $P(X_1 = k_1, X_2 = k_2, \cdots, X_n = k_n) \geqslant 0$ 以及

$$\begin{aligned}&\sum_{k_i \geqslant 0, k_1+k_2+\cdots+k_n=m} P(X_1 = k_1, X_2 = k_2, \cdots, X_n = k_n)\\ = &\sum_{k_i \geqslant 0, k_1+k_2+\cdots+k_n=m} \binom{m}{k_1, k_2, \cdots, k_n} p_1^{k_1} p_2^{k_2} \cdots p_n^{k_n} = (p_1 + p_2 + \cdots + p_n)^m = 1\ .\end{aligned}$$

当 $n = 2$ 时多项分布简化为二项分布.

引理 5.1　若多维随机向量 $(X_1, X_2, \cdots, X_n) \sim M(m, p_1, p_2, \cdots, p_n)$, 则每个随机变量 X_i 的边缘分布是二项分布 $B(m, p_i)$.

根据 X_i 的实际含义, 考虑事件 A_i 发生或不发生的伯努利试验, 则有 $X_i \sim (m, p_i)$. 另一种方法是通过多项分布的定义直接计算, 我们将其作为一个作业题.

5.3 二维连续型随机向量

定义 5.6　设二维随机向量 (X, Y) 的分布函数为 $F(x, y)$, 如果存在二元非负可积函数 $f(x, y)$, 使得对任意实数 x 和 y 有

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

则称 (X, Y) 为 **二维连续型随机向量**, 称 $f(x, y)$ 为二维随机向量 (X, Y) 的 **密度函数**, 或称随机变量 X 和 Y 的 **联合密度函数**.

联合密度函数 $f(x, y)$ 满足如下性质:

1) 非负性: 对任意实数 x 和 y 有 $f(x, y) \geqslant 0$.
2) 规范性: $\int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} f(x, y) dx dy = 1$.
   任何满足上面两条性质的二元函数 $f(x, y)$ 可以成为某随机向量 (X, Y) 的联合密度函数.
3) 若 G 为平面上的一个区域, 则点 (X, Y) 落入 G 的概率为

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

   在几何上可以看作是以 G 为底面, $z = f(x, y)$ 为顶面的柱体体积, 如图 5.3 所示.

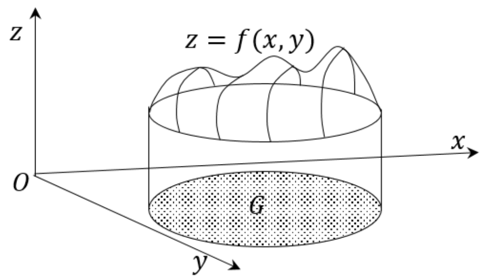


图 5.3 二维密度函数的几何意义

4) 若密度函数 $f(x,y)$ 在 (x,y) 连续, 则联合分布函数 $F(x,y)$ 和密度函数 $f(x,y)$ 满足

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

根据此性质、并利用多元泰勒展开式有

$$\begin{aligned}
&\lim_{\substack{\Delta x\to 0^+\\ \Delta y\to 0^+}}\frac{P(x<X\leqslant x+\Delta x, y<Y\leqslant y+\Delta y)}{\Delta x\Delta y}\\
=&\lim_{\substack{\Delta x\to 0^+\\ \Delta y\to 0^+}}\frac{F(x+\Delta x,y+\Delta y)-F(x+\Delta x,y)-F(x,y+\Delta y)+F(x,y)}{\Delta x\Delta y}\\
=&\frac{\partial^2 F(x,y)}{\partial x\partial y}=f(x,y)\ ,
\end{aligned}$$

由此可知

$$P(x<X\leqslant x+\Delta x, y<Y\leqslant y+\Delta y)\approx f(x,y)\Delta x\Delta y\ ,$$

概率 $f(x,y)$ 的值反映了二维随机向量 (X,Y) 落入 (x,y) 邻域内概率的大小.

根据 (X,Y) 的联合密度函数 $f(x,y)$, 还可以研究每个随机变量 X 和 Y 的密度函数 $f_X(x)$ 和 $f_Y(y)$. 首先考虑随机变量 X 的边缘分布

$$\begin{aligned}
F_X(x)&=P(X\leqslant x)=P(X\leqslant x,Y<\infty)=F(x,+\infty)\\
&=\int_{-\infty}^{x}\int_{-\infty}^{+\infty}f(t,y)dtdy=\int_{-\infty}^{x}\left(\int_{-\infty}^{+\infty}f(t,y)dy\right)dt\ ,
\end{aligned}$$

对上式两边求导得到 X 的边缘概率密度

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

同理分析随机变量 Y 的边缘分布, 于是得到边缘概率密度的严格定义.

定义 5.7 设二维随机向量 (X, Y) 的联合密度函数为 $f(x, y)$, 则随机变量 X 和 Y 的 **边缘密度函数** 分别为

$$f_X(x) = \int_{-\infty}^{+\infty} f(x, y)dy \qquad \text{和} \qquad f_Y(y) = \int_{-\infty}^{+\infty} f(x, y)dx \ .$$

例 5.3 设二维随机变量 (X, Y) 的概率密度为

$$f(x, y) = \begin{cases} ce^{-(3x+4y)} & x > 0, y > 0 \\ 0 & \text{其它} \ . \end{cases}$$

求: 1) 常数 c; 2) 联合分布函数 $F(x, y)$; 3) X 和 Y 的边缘概率密度; 4) 概率 $P(X + Y \leqslant 2)$.

解 根据密度函数的性质可知

$$1 = \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} f(x, y)dxdy = \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} ce^{-(3x+4y)}dxdy = \frac{c}{12},$$

求解出 $c = 12$. 当 $x > 0$ 和 $y > 0$ 时有

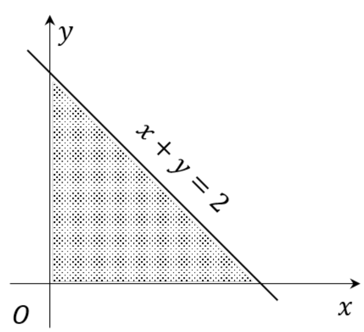
$$F(x, y) = \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} f(x, y)dxdy = \int_0^x\int_0^y 12e^{-(3x+4y)}dxdy = (1 - e^{-3x})(1 - e^{-4y}) \ ,$$

进一步根据边缘概率密度的定义有

$$f_X(x) = \int_{-\infty}^{+\infty} f(x, y)dy = \int_0^{+\infty} 12e^{-(3x+4y)}dy = 3e^{-3x} \ ,$$

同理可得 $f_Y(y) = 4e^{-4y}$. 最后计算概率 $P(X + Y \leqslant 2)$, 其积分区域如图 5.4(a) 所示, 有

$$P(X + Y \leqslant 2) = 12\int_0^2 dx\int_0^{2-x} e^{-(3x+4y)}dy = 3\int_0^2 e^{-3x}(1 - e^{-8+4x})dx = 1 - 4e^{-6} + 3e^{-8} \ .$$



(a)

图 5.4 例 5.3和xx的积分区域图

5.3.1 常用二维连续分布

下面介绍两种常用的二维连续分布: 均匀分布和正太分布.

定义 5.8 设 G 为平面上一个有界的区域, 其面积为 A_G, 若二维随机向量 (X,Y) 的联合密度函数为

$$f(x,y)=\begin{cases}1/A_G & (x,y)\in G\\ 0 & (x,y)\notin G,\end{cases}$$

则称 (X,Y) 服从区域 G 上的 **二维均匀分布**.

二维均匀分布在区域 G 上每一点等可能发生, 本质上就是 (平面) 几何概型的随机向量描述. 这里以圆的均匀分布为例, 可类似考虑三角形、椭圆等平面上一个有界区域的均匀分布.

例 5.4 在一个以坐标原点为中心、半径为 R 的圆内等可能随机投点. 用随机向量 (X,Y) 分别表示落点的横坐标和纵坐标, 求: 随机向量 (X,Y) 的联合密度函数, 边缘密度函数, 以及 (X,Y) 落入 $X^2+Y^2\leqslant r^2$ $(0<r\leqslant R)$ 的概率.

解 很容易得到圆的面积为 πR^2, 由此可知随机向量 (X,Y) 的联合密度函数

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

对于随机变量 X 的边缘密度函数, 当 $x^2\leqslant R^2$ 时有

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

同理可得随机变量 Y 的边缘密度函数. 最后所求概率

$$P(X^2+Y^2\leqslant r^2)=\iint\limits_{x^2+y^2\leqslant r^2}\frac{1}{\pi R^2}dxdy=\frac{r^2}{R^2}\ .$$

二维连续分布中最重要的是二维正太分布, 其定义如下:

定义 5.9 对任意实数 x,y, 若随机向量 (X,Y) 的密度函数为

$$f(x,y)=\frac{1}{2\pi\sqrt{1-\rho^2}\sigma_x\sigma_y}\exp\left(-\frac{1}{2(1-\rho^2)}\left[\frac{(x-\mu_x)^2}{\sigma_x^2}+\frac{(y-\mu_y)^2}{\sigma_y^2}-\frac{2\rho(x-\mu_x)(y-\mu_y)}{\sigma_x\sigma_y}\right]\right)\ ,$$

其中常数 $\mu_x,\mu_y\in(-\infty,+\infty)$, $\sigma_x,\sigma_y\in(0,+\infty)$ 以及 $\rho\in(-1,1)$, 则称 (X,Y) 服从 **二维正太分布**, 记 $(X,Y)\sim\mathcal{N}(\mu_x,\mu_y,\sigma_x^2,\sigma_y^2,\rho)$.

下面研究二维正态分布的性质:

定理 5.1 设二维随机向量 (X,Y) 服从正态分布 $\mathcal{N}(\mu_x,\mu_y,\sigma_x^2,\sigma_y^2,\rho)$, 则有随机变量 X 和 Y 的边缘分布分别为 $X\sim\mathcal{N}(\mu_x,\sigma_x^2)$ 和 $Y\sim\mathcal{N}(\mu_y,\sigma_y^2)$.

证明 这里将证明随机变量 X 的边缘密度函数, 可同理证明 Y 的边缘密度函数. 首先将二维随机向量 (X,Y) 的联合密度函数 $f(x,y)$ 分解为

$$f(x,y)=\frac{1}{\sqrt{2\pi}\sigma_x}\exp\left(-\frac{(x-\mu_x)^2}{2\sigma_x^2}\right)\times\frac{1}{\sqrt{2\pi}\sigma_y\sqrt{1-\rho^2}}\exp\left(-\frac{(y-\mu_y-\rho\sigma_y(x-\mu_x)/\sigma_x)^2}{2\sigma_y^2(1-\rho^2)}\right). \quad (5.1)$$

因此联合密度函数等于两个一维正太分布 $\mathcal{N}(\mu_x,\sigma_x)$ 和 $\mathcal{N}(\mu_y+\rho\sigma_y(x-\mu_x)/\sigma_x,\sigma_y^2(1-\rho^2))$ 的密度函数的乘积. 给定 $x,\mu_x\in(-\infty,+\infty),\sigma_x>0,\rho\in(-1,1)$ 则有

$$\int_{-\infty}^{+\infty}\frac{1}{\sqrt{2\pi}\sigma_y\sqrt{1-\rho^2}}\exp\Big(-\frac{(y-\mu_y-\rho\sigma_y(x-\mu_x)/\sigma_x)^2}{2\sigma_y^2(1-\rho^2)}\Big)dy=1\ ,$$

于是得到

$$f_X(x)=\int_{-\infty}^{+\infty}f(x,y)dy=\frac{1}{\sqrt{2\pi}\sigma_x}\exp\Big(-\frac{(x-\mu_x)^2}{2\sigma_x^2}\Big)\ .$$

由此完成证明.

定理 5.1 说明正太分布的边缘分布还是正太分布, 并给出了二维正太分布前四个参数的意义, 即随机变量 X 和 Y 的期望和方差, 第五个参数反应了两个随机变量的密切程度, 我们将在后面介绍.

二维联合分布可以唯一确定它们的边缘分布, 但反之不成立, 即使知道两个随机变量的边缘分布, 也不足以决定联合分布. 例如, 两个边缘分布为 $\mathcal{N}(\mu_x,\sigma_x)$ 和 $\mathcal{N}(\mu_y,\sigma_y)$, 因为不能确定 ρ 的值而不能确定它们的联合分布. 基于 (5.1), 我们还可以验证二维正太分布的规范性

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

以及二维正太分布的密度函数本质是两个 (一维) 正太分布的密度函数的乘积.

5.4 随机变量的独立性

前面第二章介绍了随机事件的独立性, 即独立的的随机事件 A 和 B 满足 $P(AB)=P(A)P(B)$. 本节介绍概率统计中另一个重要的概念: 随机变量的独立性. 考虑两个随机变量, 若一个随机变量的取值对另一个随机变量没有什么影响, 则称两个随机变量相互独立. 下面给出严格的数学定义:

定义 5.10 设二维随机向量 (X,Y) 的联合分布函数为 $F(x,y)$, 以及 X 和 Y 的边缘分布函数分别为 $F_X(x)$ 和 $F_Y(y)$, 若对任意的实数 x 和 y 有

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

则称随机变量 X 与 Y **相互独立**.

根据上面的定义可知, 随机变量 X 与 Y 相互独立等价于随机事件 $\{X \leqslant x\}$ 和 $\{Y \leqslant y\}$ 对任意实数 x 和 y 都相互独立; 容易发现常数 c 与任意随机变量相互独立.

对于离散型随机向量, 可以考虑通过分布列来刻画它的统计规律, 关于独立性有

定理 5.2 设二维离散型随机向量 (X,Y) 的分布列为 $p_{ij} = P(X = x_i, Y = y_j)$ $(i, j = 1, 2, \cdots)$, 以及 X 和 Y 的边缘分布列为 $p_{i\cdot} = P(X = x_i)$ 和 $p_{\cdot j} = P(Y = y_j)$, 则随机变量 X 和 Y 相互独立的充要条件是 $p_{ij} = p_{i\cdot}p_{\cdot j}$.

证明 首先证明必要性, 根据定义 5.10 分布函数的独立性有

$$\begin{aligned}
p_{i,j} &= F(x_i, y_j) - F(x_{i-1}, y_j) - F(x_i, y_{j-1}) + F(x_{i-1}, y_{j-1}) \\
&= F_X(x_i)F_Y(y_j) - F_X(x_{i-1})F_Y(y_j) - F_X(x_i)F_Y(y_{j-1}) + F_X(x_{i-1})F_Y(y_{j-1}) \\
&= (F_X(x_i) - F_X(x_{i-1}))F_Y(y_j) - (F_X(x_i) - F_X(x_{i-1}))F_Y(y_{j-1}) \\
&= p_{i\cdot}F_Y(y_j) - p_{i\cdot}F_Y(y_{j-1}) = p_{i\cdot}p_{\cdot j}\ .
\end{aligned}$$

其次证明充分性, 根据 $p_{ij} = p_{i\cdot}p_{\cdot j}$ $(i, j = 1, 2, \cdots)$ 有

$$F(x_m, y_n) = \sum_{i\leqslant m}\sum_{j\leqslant n} p_{ij} = \sum_{i\leqslant m}\sum_{j\leqslant m} p_{i\cdot}p_{\cdot j} = \sum_{i\leqslant m} p_{i\cdot} \times \sum_{j\leqslant n} p_{\cdot j} = F_X(x_m)F_Y(y_n)\ .$$

由此完成证明.

例 5.5 设离散型随机变量 X 和 Y 相互独立且它们的取值均为 $\{1, 2, 3\}$, 已知 $P(Y = 1) = 1/3$, $P(X = 1, Y = 1) = P(X = 2, Y = 1) = 1/8$ 和 $P(X = 1, Y = 3) = 1/16$, 求 X 和 Y 的联合分布列和边缘分布列.

解 根据边缘分布列的定义有

$$P(X = 3, Y = 1) = P(Y = 1) - P(X = 1, Y = 1) - P(X = 2, Y = 1) = 1/12\ ,$$

再根据定理 5.2 有 $P(X = 1) = P(X = 2) = 3/8$ 和 $P(X = 3) = 1/4$, 同理计算其它概率, 最后得到的分布列为

$X \backslash Y$	1	2	3	$p_{i\cdot}$
1	1/8	3/16	1/16	3/8
2	1/8	3/16	1/16	3/8
3	1/12	1/8	1/24	1/4
$p_{\cdot j}$	1/3	1/2	1/6	

对于连续型随机向量, 一般可以通过密度函数来进行刻画, 关于独立性有

定理 5.3 设二维随机向量 (X,Y) 的联合密度函数为 $f(x,y)$, 及 X 和 Y 的边缘密度函数分别为 $f_X(x)$ 和 $f_Y(y)$, 则随机变量 X 和 Y 相互独立的充要条件是 $f(x,y)=f_X(x)f_Y(y)$.

证明 首先证明必要性: 若二维连续随机变量满足 $F(x,y)=F_X(x)F_Y(y)$, 则有

$$\int_{-\infty}^{x}\int_{-\infty}^{y} f(u,v)dudv = \int_{-\infty}^{x} f_X(u)du \int_{-\infty}^{y} f_Y(v)dv\ ,$$

对上式两边同时求偏导有

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

其次证明充分性: 若 $f(x,y)=f_X(x)f_Y(y)$, 则有

$$\begin{aligned} F(x,y) &= \int_{-\infty}^{x}\int_{-\infty}^{y} f(u,v)dudv = \int_{-\infty}^{x}\int_{-\infty}^{y} f_X(u)f_Y(v)dudv \\ &= \int_{-\infty}^{x} f_X(u)du \int_{-\infty}^{y} f_Y(v)dv = F_X(x)F_Y(y)\ , \end{aligned}$$

由此完成证明.

下面介绍关于随机变量独立性的一些性质:

引理 5.2 若随机变量 X 和 Y 相互独立, 则对任意给定的集合 $A,B\subseteq\mathbb{R}$, 事件 $\{X\in A\}$ 和事件 $\{Y\in B\}$ 相互独立.

证明 该引理对离散型和连续型随机变量均成立, 这里我们详细证明连续随机变量情形. 根据独立性有 $f(x,y)=f_X(x)f_Y(y)$, 由此可得

$$\begin{aligned} P(X\in A, Y\in B) &= \iint_{x\in A, y\in B} f(x,y)dxdy \\ &= \iint_{x\in A, y\in B} f_X(x)f_Y(y)dxdy = \int_{x\in A} f_X(x)dx\int_{y\in B} f_Y(y)dy = P(X\in A)P(Y\in B)\ , \end{aligned}$$

引理得证.

引理 5.3 设随机变量 X 和 Y 相互独立, 以及 $f(x)$ 和 $g(y)$ 是连续或分段连续的函数, 则有 $f(X)$ 与 $g(Y)$ 相互独立.

该定理对离散型和连续型随机变量均成立, 这里没给出它的证明是因此其超出了本书的范围. 根据此引理, 若随机变量 X 与 Y 相互独立, 则 X^2 与 Y^3 相互独立, 以及 $\sin X$ 与 $\cos Y$ 也相互独立.