5.3 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 $f(x,y)$, 对任意实数 x 和 y 有

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

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**.

容易发现联合密度函数 $f(x,y)$ 满足非负性 ($f(x,y)\geqslant 0$) 和规范性 ($\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty}f(x,y)dxdy=1$). 任何满足非负性和规范性的二元函数 $f(x,y)$ 可以成为某随机向量 (X,Y) 的联合密度函数.

若 G 为平面上的一个区域, 则点 (X,Y) 落入 G 的概率为

$$P((X,Y)\in G)=\iint\limits_{(x,y)\in G}f(x,y)dxdy,$$

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(a) 所示.

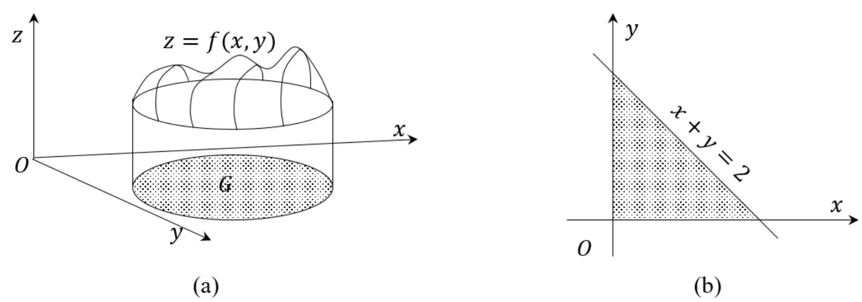

图 5.3 二维密度函数的几何意义和例 5.2 的积分区域

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

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

根据此性质、并利用多元泰勒展开式有

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

由此可知

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

密度函数 $f(x, y)$ 反映了二维随机向量 (X, Y) 落入 (x, y) 邻域内的概率.

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

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

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.

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 5.2 设二维随机变量 (X, Y) 的概率密度为

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

求联合分布函数 $F(x, y)$, X 和 Y 的边缘概率密度, 以及概率 $P(X + Y \leqslant 2)$.

解 根据密度函数的性质可知

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

求解出 $c = 12$. 当 $x > 0$ 和 $y > 0$ 时有

$$F(x, y) = \int_{-\infty}^{x} \int_{-\infty}^{y} f(u, v) du dv = \int_{0}^{x} \int_{0}^{y} 12e^{-(3u+4v)} du dv = (1 - e^{-3x})(1 - e^{-4y}),$$

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 5.3(b) 所示, 有

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

常用二维连续分布

均匀分布是一种常用的连续分布, 考虑在平面区域 G 上每一点发生的可能性相同.

定义 5.8 设 G 为平面上一个面积为 A_G 的有界区域, 若二维随机向量 (X,Y) 的联合密度函数

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

则称 (X,Y) 服从区域 G 上的 **二维均匀分布**.

二维均匀分布所描述的随机现象: 在平面区域 G 中随机投点, 如果该点坐标 (X,Y) 落在 G 中某子区域 D 中的概率只与 D 的面积成正比, 而与 D 的位置无关. 用密度函数可表述为

$$P((X,Y)\in D) = \iint_D f(x,y)dxdy = \iint_D \frac{1}{A_G}dxdy = \frac{D\text{ 的面积}}{G\text{ 的面积}}.$$

例 5.3 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) 的联合密度函数和边缘密度函数.

解 根据圆面积为 πR^2 可得随机向量 (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 $x \in [-R,R]$ 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.

二维正态分布是另一种常用的二维连续分布, 其定义如下:

定义 5.9 若随机向量 (X,Y) 的密度函数为

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

则称 (X,Y) 服从参数为 $\mu_x,\mu_y,\sigma_x^2,\sigma_y^2,\rho$ 的 **二维正态分布**, 记 $(X,Y)\sim\mathcal{N}(\mu_x,\mu_y,\sigma_x^2,\sigma_y^2,\rho)$. 这里常数 $\mu_x,\mu_y\in(-\infty,+\infty)$, $\sigma_x,\sigma_y\in(0,+\infty)$ 以及 $\rho\in(-1,1)$.

关于二维正态分布, 有如下性质.

性质 5.2 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 将 (X,Y) 的联合密度函数 $f(x,y)$ 分解为

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

不难发现联合密度函数可分解为两个一维正态分布 $\mathcal{N}(\mu_x,\sigma_x)$ 和 $\mathcal{N}(\mu_y+\rho\sigma_y(x-\mu_x)/\sigma_x,\sigma_y^2(1-\rho^2))$ 密度函数的乘积. 根据密度函数的规范性有

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

由此可得随机变量 X 的边缘密度函数为

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

同理可证随机变量 Y 的边缘密度函数.

性质 5.2 说明正态分布的边缘分布还是正态分布, 并给出了二维正态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相关程度, 将在后续章节介绍.

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, 已知两个边缘分布为 $\mathcal{N}(\mu_x,\sigma_x)$ 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) 还可以验证二维正太分布的规范性

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

这是因为二维正态分布的密度函数本质是两个 (一维) 正态分布的密度函数的乘积.

5.4 随机变量的独立性

第二章介绍了随机事件的独立性, 即相互独立的的随机事件 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的取值没有影响, 则称两个随机变量相互独立.

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**.

随机变量 X 与 Y 相互独立等价于任意实数 x 和 y 随机事件 $\{X\leqslant x\}$ 和 $\{Y\leqslant y\}$ 相互独立.

对于离散型随机向量, 可以通过分布列来刻画独立性,

定理 5.1 设二维离散型随机向量 (X,Y) 的分布列为

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

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 不妨假设 $x_1<x_2<\cdots<x_i<\cdots$ 和 $y_1<y_2<\cdots<y_j<\cdots$. 首先证明必要性, 根据独立性的定义有

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

其次证明充分性, 根据 $p_{ij}=p_{i\cdot}p_{\cdot j}$ 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 5.4 设随机变量 X 和 Y 相互独立且它们的取值均为 $\{1,2,3\}$, 已知

$$P(X=1,Y=3)=1/16,\quad P(X=1,Y=1)=P(X=2,Y=1)=1/8,\quad P(Y=1)=1/3,$$

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 5.1 有 $P(X=1)=P(X=2)=3/8$ 和 $P(X=3)=1/4$, 类似地计算其他概率, 最后得到联合分布列和边缘分布列为

$X \backslash Y$	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, 还可以通过密度函数来刻画独立性.

定理 5.2 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 X 和 Y 相互独立, 则有 $F(x,y)=F_X(x)F_Y(y)$, 即

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

上式两边分别求偏导有

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

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:

性质 5.3 若随机变量 X 和 Y 相互独立, 则对任意集合 $A,B\subseteq\mathbb{R}$, 事件 $\{X\in A\}$ 和事件 $\{Y\in B\}$ 相互独立.

证明 这里仅仅给出连续型随机变量的详细证明, 离散型随机变量可类似考虑. 根据独立性有 $f(x,y)=f_X(x)f_Y(y)$, 于是有

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