性质 5.4 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.

该性质对离散型和连续型随机变量均成立, 其证明超出了本书的范畴. 根据该性质可知: 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.

下面再给出一种判断两个随机变量独立性的方法.

性质 5.5 如果存在两个函数 $h(x)$ 和 $g(y)$, 使得对任意实数 x 和 y, 随机变量 X 和 Y 的联合密度函数 $f(x,y)$ 可分解为

$$f(x,y) = h(x)g(y),$$

则随机变量 X 和 Y 相互独立.

证明 不妨设

$$C_1 = \int_{-\infty}^{+\infty} h(x)dx \qquad \text{和} \qquad C_2 = \int_{-\infty}^{+\infty} g(y)dy.$$

根据密度函数的规范性有

$$1 = \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} f(x,y)dxdy = \int_{-\infty}^{+\infty} h(x)dx \int_{-\infty}^{+\infty} g(y)dy = C_1C_2.$$

随机变量 X 和 Y 的边缘密度函数分别为

$$f_X(x) = \int_{-\infty}^{+\infty} f(x,y)dy = h(x)C_2 \quad \text{和} \quad f_Y(y) = \int_{-\infty}^{+\infty} f(x,y)dx = C_1g(y).$$

于是得到

$$f_X(x)f_Y(y) = C_1C_2h(x)g(y) = h(x)g(y) = f(x,y),$$

由此完成证明.

定理 5.3 设二维随机向量 $(X,Y) \sim \mathcal{N}(\mu_x,\mu_y,\sigma_x^2,\sigma_y^2,\rho)$, 则 X 与 Y 独立的充要条件为 $\rho = 0$.

证明 首先证明必要性, 若随机变量 X 与 Y 相互独立, 则有 $f(x,y) = f_X(x)f_Y(y)$, 即

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

上式对任意实数 x 和 y 均成立, 取 $x = \mu_x$ 和 $y = \mu_y$ 求解出 $\rho = 0$.

其次证明充分性, 若随机向量 $(X,Y)\sim\mathcal{N}(\mu_x,\mu_y,\sigma_x^2,\sigma_y^2,\rho)$, 则 X 和 Y 的边缘分布分别为 $\mathcal{N}(\mu_x,\sigma_x^2)$ 和 $\mathcal{N}(\mu_y,\sigma_y^2)$. 当 $\rho=0$ 时, 根据二维正态分布的定义有

$$f(x,y)=\frac{1}{2\pi\sigma_x\sigma_y}\exp\left(-\frac{(x-\mu_x)^2}{2\sigma_x^2}-\frac{(y-\mu_y)^2}{2\sigma_y^2}\right)=f_X(x)f_Y(y).$$

由此完成证明。

例 5.5 设二维随机向量 (X,Y) 的密度函数

$$f(x,y)=\begin{cases}cxe^{-y} & 0<x<y<+\infty\\ 0 & \text{其它},\end{cases}$$

讨论随机变量 X 与 Y 的独立性.

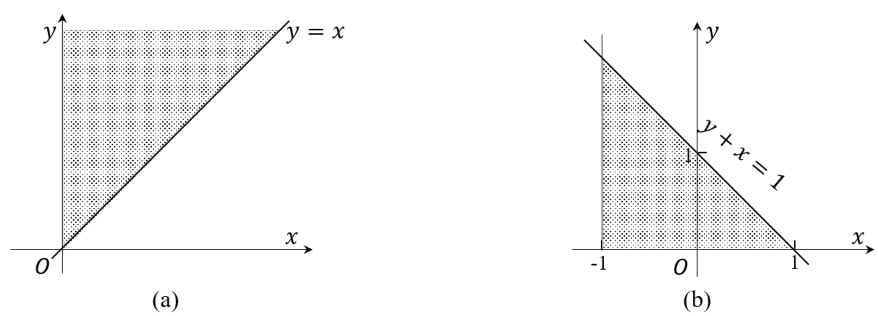

图 5.4 例 5.5 和 5.6 的积分区域

解 如图 5.4(a) 所示, 根据密度函数的规范性和分部积分有

$$1=\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty}f(x,y)dxdy=c\int_0^{+\infty}dx\int_x^{+\infty}xe^{-y}dy=c\int_0^{+\infty}xe^{-x}dx=c.$$

当 $x>0$ 时, 随机变量 X 的边缘概率密度为

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

同理当 $y>0$ 时, 随机变量 Y 的边缘概率密度为

$$f_Y(y)=\int_{-\infty}^{+\infty}f(x,y)dx=\int_0^{y}xe^{-y}dx=\frac{1}{2}y^2e^{-y}.$$

当 $x>0$ 和 $y>0$ 时有 $f(x,y)\neq f_X(x)f_Y(y)$, 由此可知随机变量 X 与 Y 不独立.

例 5.6 设随机变量 X 与 Y 相互独立, 且 X 服从 $[-1,1]$ 均匀分布, Y 服从参数为 $\lambda=2$ 的指数分布, 求 $P(X+Y\leqslant 1)$.

解 随机变量 X 与 Y 的边缘概率密度分别为

$$f_X(x)=\begin{cases}1/2 & x\in[-1,1]\\ 0 & \text{其它}\end{cases} \qquad \text{和} \qquad f_Y(y)=\begin{cases}2e^{-2y} & y\geqslant 0\\ 0 & \text{其它}.\end{cases}$$

根据独立性可知随机变量 X 与 Y 的联合概率密度

$$f(x,y)=\begin{cases}e^{-2y} & -1\leqslant x\leqslant 1, y\geqslant 0\\ 0 & \text{其它}.\end{cases}$$

所求积分区域如图 5.4(b) 所示, 最后得到

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

5.5 条件分布

前面介绍了随机事件的条件概率 $P(A|B)=P(AB)/P(B)$, 即在随机事件 B 发生的条件下随机事件 A 发生的概率. 本节介绍随机变量的条件分布, 在随机变量 Y 取值确定的条件下随机变量 X 的概率分布, 分离散和连续两种情况进行讨论.

5.5.1 离散型随机变量的条件概率

定义 5.11 设随机变量 (X,Y) 的联合分布列为 $p_{ij}=P(X=x_i,Y=y_j)$ $i,j=1,2,\cdots$, 随机变量 Y 的边缘概率 $p_{\cdot j}=P(Y=y_j)=\sum_{i=1}^{+\infty}p_{ij}>0$,

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

称为在 $Y=y_j$ 的条件下随机变量 X 的 **条件分布列** (conditional probability distribution). 类似地可定义在 $X=x_i$ 的条件下随机变量 Y 的条件分布列.

条件分布本质上也是一种概率分布, 具有分布的性质. 例如,

- **非负性**: $P(X=x_i|Y=y_j)\geqslant 0$;
- **规范性**:
$$\sum_{i=1}^{\infty}P(X=x_i|Y=y_j)=\sum_{i=1}^{+\infty}\frac{P(X=x_i,Y=y_j)}{P(Y=y_j)}=\sum_{i=1}^{+\infty}\frac{p_{ij}}{p_{\cdot j}}=1.$$

- 若离散随机变量 X 和 Y 相互独立, 则有

$$P(X = x_i|Y = y_j) = P(X = x_i) \quad \text{和} \quad P(Y = y_j|X = x_i) = P(Y = y_j).$$

在后面的章节中, 若出现条件概率 $P(X = x_i|Y = y_j)$, 一般都默认 $P(Y = y_j) > 0$.

例 5.7 一个选手击中目标的概率为 p, 射中两次目标为止, 用 X 表示首次击中目标所进行的射击次数, 用 Y 表示第二次射中目标所进行的射击次数, 求 X 和 Y 的联合分布列和条件分布列.

解 随机变量 $X = m$ 表示在第 m 次射击时首次击中目标射击, 随机变量 $Y = n$ 表示在第 n 次射击时第二次击中目标, 则 X 和 Y 的联合分布列为

$$P(X = m, Y = n) = \begin{cases} p^2(1-p)^{n-2} & 1 \leqslant m < n < \infty \\ 0 & \text{其它.} \end{cases}$$

由此可得随机变量 X 的边缘分布列

$$P(X = m) = \sum_{n=m+1}^{\infty} P\{X = m, Y = n\} = \sum_{n=m+1}^{\infty} p^2(1-p)^{n-2} = p(1-p)^{m-1} \qquad (1 \leqslant m),$$

以及随机变量 Y 的边缘分布列

$$P(Y = n) = \sum_{m=1}^{n-1} P\{X = m, Y = n\} = \sum_{m=1}^{n-1} p^2(1-p)^{n-2} = (n-1)p^2(1-p)^{n-2} \qquad (2 \leqslant n).$$

在 $Y = n$ 的条件下随机变量 X 的分布列为

$$P(X = m|Y = n) = \frac{P(X = m, Y = n)}{P(Y = n)} = \frac{p^2(1-p)^{n-2}}{(n-1)p^2(1-p)^{n-2}} = \frac{1}{n-1} \qquad (1 \leqslant m < n),$$

以及在 $X = m$ 的条件下随机变量 Y 的分布列

$$P(Y = n|X = m) = \frac{P(X = m, Y = n)}{P(X = m)} = \frac{p^2(1-p)^{n-2}}{p(1-p)^{m-1}} = p(1-p)^{n-m-1} \qquad (1 \leqslant m < n).$$

5.5.2 连续型随机向量的条件分布

对连续型随机向量 (X, Y), 不能直接利用条件概率公式来研究条件分布, 因为连续随机变量有 $P(X = x) = P(Y = y) = 0$. 下面引入条件概率函数来研究连续随机变量的条件分布.

定义 5.12 设连续型随机变量 (X, Y) 的联合概率密度为 $f(x, y)$, 以及 X 和 Y 的边缘概率密度分别为 $f_X(x)$ 和 $f_Y(y)$. 若 $f_Y(y) > 0$, 则称

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

为在 $Y=y$ 条件下随机变量 X 的 **条件密度函数** (conditional probability density function), 以及

$$F_{X|Y}(x|y)=P(X\leqslant x|Y=y)=\int_{-\infty}^{x} f_{X|Y}(u|y)du$$

为在 $Y=y$ 条件下 X 的 **条件分布函数** (conditional cumulative distribution function). 若边缘密度函数 $f_X(x)>0$, 可同理定义在 $X=x$ 条件下 Y 的条件密度函数和条件分布函数分别为

$$f_{Y|X}(y|x)=f(x,y)/f_X(x) \quad 和 \quad F_{Y|X}(y|x)=\int_{-\infty}^{y} f_{Y|X}(v|x)dv.$$

特别值得注意的是: 对于连续型随机变量 (X,Y), 条件分布

$$F_{X|Y}(x|y)=P(X\leqslant x|Y=y)\neq\frac{P(X\leqslant x,Y=y)}{P(Y=y)}.$$

但可以利用极限的观点给出条件分布函数和密度函数定义的一种解释

$$\begin{aligned}
F_{X|Y}(x|y) &= \lim_{\epsilon\to 0^+} P(X\leqslant x|y\leqslant Y\leqslant y+\epsilon)=\lim_{\epsilon\to 0^+}\frac{P(X\leqslant x,y\leqslant Y\leqslant y+\epsilon)}{P(y\leqslant Y\leqslant y+\epsilon)} \\
&= \lim_{\epsilon\to 0^+}\frac{\int_{-\infty}^{x}\int_{y}^{y+\epsilon} f(u,v)dudv}{\int_{y}^{y+\epsilon} f_Y(v)dv}=\lim_{\epsilon\to 0^+}\frac{\epsilon\int_{-\infty}^{x} f(u,y+\theta_1\epsilon)du}{\epsilon f_Y(y+\theta_2\epsilon)} \qquad \theta_1,\theta_2\in(0,1) \\
&= \frac{\int_{-\infty}^{x} f(u,y)du}{f_Y(y)}=\int_{-\infty}^{x}\frac{f(u,y)}{f_Y(y)}du \qquad (这里利用了积分中值定理).
\end{aligned}$$

进一步得到条件密度函数 $f_{X|Y}(x|y)=f(x,y)/f_Y(y)$. 在后续章节中, 若出现条件密度函数, 一般都默认分母 $f_Y(y)>0$ 或 $f_X(x)>0$. 条件密度函数本质上是密度函数, 具有以下性质:

- **非负性**: $f_{Y|X}(y|x)\geqslant 0$.
- **规范性**: $\int_{-\infty}^{+\infty} f_{X|Y}(x|y)dx=\int_{-\infty}^{+\infty} f(x,y)/f_Y(y)dx=\int_{-\infty}^{+\infty} f(x,y)dx/f_Y(y)=1$.
- **乘法公式**: $f(x,y)=f_X(x)f_{Y|X}(y|x)=f_Y(y)f_{X|Y}(x|y)$.
- **密度函数的贝叶斯公式**: 根据条件概率的乘法公式有

  $$f_{X|Y}(x|y)=\frac{f(x,y)}{f_Y(y)}=\frac{f_{Y|X}(y|x)f_X(x)}{f_Y(y)}=\frac{f_{Y|X}(y|x)f_X(x)}{\int_{-\infty}^{+\infty} f(x,y)dx}=\frac{f_{Y|X}(y|x)f_X(x)}{\int_{-\infty}^{+\infty} f_{Y|X}(y|x)f_X(x)dx},$$

  即在 $Y=y$ 的条件下 X 的条件密度函数 $f_{X|Y}(x|y)$ 可以通过在 $X=x$ 的条件下 Y 的条件密度函数 $f_{Y|X}(y|x)$ 来计算.
- 若随机变量 X 和 Y 相互独立, 则有 $f_{Y|X}(y|x)=f_Y(y)$ 和 $f_{X|Y}(x|y)=f_X(x)$.