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 \ 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:

性质 5.1 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.

性质 5.2 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.

下面给出一种判断两个随机变量独立性的简单方法:

性质 5.3 如果存在两个函数 $h(x)$ 和 $g(y)$, 使得 X 和 Y 的联合密度函数 $f(x,y)$ 对任意实数 x 和 y 均有

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

则随机变量 X 和 Y 相互独立.

证明 不妨假设

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

根据密度函数的规范性有

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

随机变量 X 和 Y 的边缘分布分别为

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

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 5.4 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, 根据二维正太分布的定义有

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

其次证明充分性. 若 X 与 Y 相互独立, 则对任意实数 x 和 y 均有 $f(x,y) = f_X(x)f_Y(y)$ 成立, 即

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

取 $x = \mu_x$ 和 $y = \mu_y$ 求解出 $\rho = 0$, 由此完成证明。

例 5.6 设二维随机向量的密度函数

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

问随机变量 X 与 Y 是否相互独立?

解 如图 5.4(a) 所示, 利用密度函数的规范性和分部积分有

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

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

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

因为 $f(x,y)\neq f_X(x)f_Y(y)$ 可得随机变量 X 与 Y 不独立.

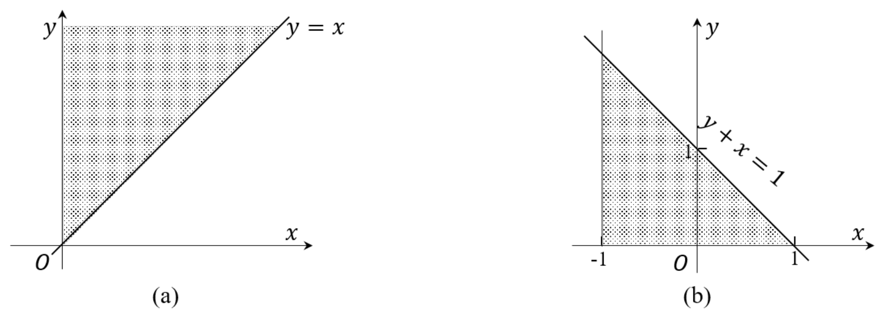

图 5.4 例 5.6 和 5.7 的积分区域

例 5.7 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 首先有随机变量 X 与 Y 的边缘概率密度分别为

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

根据独立性可得随机变量 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

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

5.5 条件分布

前面第二章讨论了随机事件的条件概率, 即在事件 B 发生的条件下事件 A 发生的条件概率 $P(A|B) = P(AB)/P(B)$. 可同理考虑随机变量的条件分布, 在给定随机变量 Y 具体取值的条件下, 考虑随机变量 X 的概率分布. 下面分离散和连续两种情形进行讨论.

5.5.1 离散型随机变量的条件概率

定义 5.11 设离散型随机变量 (X,Y) 的分布列为 $p_{ij} = P(X = x_i, Y = y_j)$ $(i, j = 1, 2, \cdots)$, 对于给定的边缘概率 $p_{\cdot j} = P(Y = y_j) = \sum_{i=1}^{+\infty} p_{ij} > 0$,

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

称为在 $Y = y_j$ 条件下随机变量 X 的 **条件分布列** (conditional probability distribution). 可以类似定义在 $X = x_i$ 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**: 对任意整数 $i \geqslant 1$ 有 $P(X = x_i|Y = y_j) \geqslant 0$;
- **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
$$P(X = x_i|Y = y_j) = P(X = x_i) \quad 和 \quad P(Y = y_j|X = x_i) = P(Y = y_j)\ .$$

在后面的章节中, 若出现条件概率 $P(X = x_i|Y = y_j)$, 一般默认概率 $P(Y = y_j) > 0$. 条件分布列也可以通过下面的表格给出:

X	x_1	x_2	$\cdots$	x_n	$\cdots$
$P(X = x_i\|Y = y_j)$	$p_{1j}/p_{\cdot j}$	$p_{2j}/p_{\cdot j}$	$\cdots$	$p_{nj}/p_{\cdot j}$	$\cdots$

例 5.8 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 的联合分布和条件分布.

解 随机变量 $X = m$ 表示首次击中目标射击了 m 次, $Y = n$ 表示第二次次击中目标射击了 n 次, 则 X 和 Y 的联合分布列为

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

于是得到随机变量 X 的边缘分布列

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

以及随机变量 Y 的边缘分布列

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

当 $n\geqslant 2$ 时, 随机变量 X 在 $Y=n$ 条件下的分布列

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

当 $m\geqslant 1$ 时, 随机变量 Y 在 $X=m$ 条件下的分布列

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

5.5.2 连续型随机向量的条件分布

连续型随机向量 (X,Y) 对任意实数 x,y 都有 $P(X=x)=0$ 和 $P(Y=y)=0$, 因此不能用条件概率的公式直接推导连续型随机向量的条件分布, 但可以通过下面的极限方式来考虑.

当 $P(y\leqslant Y\leqslant y+\epsilon)>0$ 时, 利用积分中值定理来求解条件分布函数

$$\begin{aligned}F_{X|Y}(x|y)&=\lim_{\epsilon\to 0^+}P\{X\leqslant x|y\leqslant Y\leqslant y+\epsilon\}=\lim_{\epsilon\to 0^+}\frac{P\{X\leqslant x,y\leqslant Y\leqslant y+\epsilon\}}{P\{y\leqslant Y\leqslant y+\epsilon\}}\\&=\lim_{\epsilon\to 0^+}\frac{\int_{-\infty}^{x}\int_{y}^{y+\epsilon}f(u,v)dudv}{\int_{y}^{y+\epsilon}f_Y(u)dv}=\lim_{\epsilon\to 0^+}\frac{\epsilon\int_{-\infty}^{x}f(u,y+\theta_1\epsilon)du}{\epsilon f_Y(y+\theta_2\epsilon)}\qquad\theta_1,\theta_2\in(0,1)\\&=\frac{\int_{-\infty}^{x}f(u,y)du}{f_Y(y)}=\int_{-\infty}^{x}\frac{f(u,y)}{f_Y(y)}du\ ,\end{aligned}$$

进一步得到条件密度函数 $f_{X|Y}(x|y)=f(x,y)/f_Y(y)$. 在后面的章节中, 若出现条件密度函数 $f_{X|Y}(x|y)$ (或 $f_{Y|X}(y|x)$), 一般都默认 $f_Y(y)>0$ (或 $f_X(x)>0$). 下面给详细的定义:

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 $f_Y(y)$. 对任意给定的 $f_Y(y)>0$, 称 $f(x,y)/f_Y(y)$ 为在 $Y=y$ 条件下随机变量 X 的 **条件密度函数** (conditional probability density function), 记为

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

以及在 $Y=y$ 条件下 X 的 **条件分布函数** (conditional cumulative distribution function) 为

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

对任意给定的 $f_X(x)>0$, 可类似定义在 $X=x$ 条件下 Y 的条件密度函数和条件分布函数分别为

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

条件密度函数本质上是密度函数, 具有以下性质:

- **非负性**: 对任意实数 x,y 有 $f_{Y|X}(y|x) \geqslant 0$.
- **规范性**: 对任意实数 y: $f_Y(y)>0$ 有

$$\int_{-\infty}^{+\infty} f_{X|Y}(x|y)dx = \int_{-\infty}^{+\infty} \frac{f(x,y)}{f_Y(y)}dx = \frac{1}{f_Y(y)}\int_{-\infty}^{+\infty} f(x,y)dx = 1 \ .$$

- **乘法公式**: $f(x,y) = f_X(x)f_{Y|X}(y|x) = f_Y(y)f_{X|Y}(x|y)$.
- 若随机变量 X 和 Y 相互独立, 则有

$$f_{Y|X}(y|x) = f_Y(y) \quad \text{和} \quad f_{X|Y}(x|y) = f_X(x) \ .$$

根据条件概率的乘法公式有

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

可以将其看作 **密度函数的贝叶斯公式**. 目前可能有三种途径来构造二维随机向量的联合分布函数:

i) 根据实际问题或实际数据归纳为 $f(x,y)$;

ii) 根据随机变量的独立性有 $f(x,y) = f_X(x)f_Y(y)$;

iii) 根据乘法公式 $f(x,y) = f_X(x)f_{Y|X}(y|x)$.

关于二维正太分布有

定理 5.5 若随机向量 $(X,Y) \sim \mathcal{N}(\mu_x,\mu_y,\sigma_x^2,\sigma_y^2,\rho)$, 则在 $Y=y$ 的条件下随机变量 X 服从正太分布 $\mathcal{N}(\mu_x-\rho\sigma_x(y-\mu_y)/\sigma_y,(1-\rho^2)\sigma_x^2)$, 以及在 $X=x$ 的条件下随机变量 Y 服从正太分布 $\mathcal{N}(\mu_y-\rho\sigma_y(x-\mu_x)/\sigma_x,(1-\rho^2)\sigma_y^2)$.

证明 若随机向量 $(X,Y) \sim \mathcal{N}(\mu_x,\mu_y,\sigma_x^2,\sigma_y^2,\rho)$, 则随机变量 X 的边缘分布为 $\mathcal{N}(\mu_x,\sigma_x^2)$, 即

$$f_X(x) = \frac{1}{\sqrt{2\pi}\sigma_x}\exp\Big(-\frac{(x-\mu_x)^2}{2\sigma_x^2}\Big) \ ,$$

此外可以将二维随机向量 (X,Y) 的联合密度函数 $f(x,y)$ 分解为

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

根据乘法公式 $f(x,y)=f_X(x)f_{Y|X}(y|x)$ 可得

$$f_{Y|X}(y|x)=\frac{1}{\sqrt{2\pi}\sigma_y\sqrt{1-\rho^2}}\exp\Big(-\frac{(y-\mu_y-\rho\sigma_y(x-\mu_x)/\sigma_x)^2}{2\sigma_y^2(1-\rho^2)}\Big)\ ,$$

即正太分布 $\mathcal{N}(\mu_y-\rho\sigma_y(x-\mu_x)/\sigma_x,(1-\rho^2)\sigma_y^2)$. 同理可证在 $Y=y$ 的条件下随机变量 X 的条件分布为 $\mathcal{N}(\mu_x-\rho\sigma_x(y-\mu_y)/\sigma_y,(1-\rho^2)\sigma_x^2)$.

例 5.9 设二维随机向量 (X,Y) 的密度函数

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

求 $P(X>1|Y=y)$.

解 积分区域如图 5.5(a) 所示, 求解随机变量 Y 的边缘分布为

$$f_Y(y)=\int_0^{+\infty}e^{-x/y}e^{-y}/ydx=e^{-y}\left[-e^{-x/y}\right]_0^{+\infty}=e^{-y}\qquad(y>0)\ .$$

进而得到在 $Y=y$ 条件下 X 的条件概率密度为

$$f_{X|Y}(x|y)=e^{-x/y}/y\ .$$

最后求解得到

$$P(X>1|Y=y)=\int_1^{\infty}e^{-x/y}/ydx=-\left[e^{-x/y}\right]_1^{\infty}=e^{-1/y}\ .$$

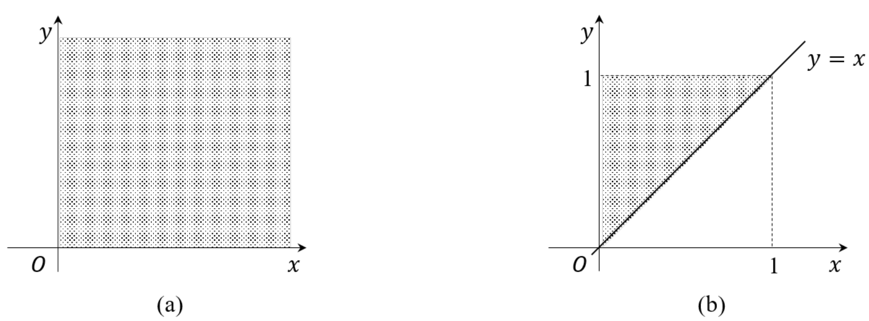

图 5.5 例 5.9 和 5.10 的积分区域

例 5.10 设随机变量 $X \sim U(0,1)$, 在观察到 $X=x$ 的条件下随机变量 $Y \sim U(x,1)$, 求随机变量 Y 的概率密度.

解 随机变量 $X \sim U(0,1)$, 在随机变量 $X=x$ 的条件下 $Y \sim U(x,1)$, 于是当 $x>0$ 时有

$$f_{Y|X}(y|x) = 1/(1-x) \ .$$

根据条件概率乘积公式有

$$f(x,y) = f_X(x) f_{Y|X}(y|x) = \begin{cases} 1/(1-x) & 0<x<y<1, \\ 0 & \text{其它}. \end{cases}$$

积分区域如图 5.5(b) 所示, 当 $y>0$ 时随机变量 Y 的边缘分布

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

5.6 多维随机变量函数的分布

本节研究若已知二维随机向量 (X,Y) 的概率分布, 求一维随机变量 $Z=g(X,Y)$ 的概率分布. 下面分别对离散型和连续型随机变量两种情况进行讨论.

5.6.1 二维离散型随机向量函数

若已知二维离散型随机向量 (X,Y) 的联合分布列, 求随机变量 $Z=g(X,Y)$ 的分布列相对较为简单. 首先对 X,Y 的各种取值, 计算随机变量 Z 的取值, 然后对相同的 Z 值合并, 对应的概率相加. 下面研究两个相互独立的离散型随机变量的求和, 即离散型随机变量的卷积公式:

定理 5.6 设离散型随机变量 X 与 Y 相互独立、且它们的分布列分别为 $a_i = P(X=i)$ 和 $b_j = P(Y=j)$ $(i,j=0,1,\cdots)$, 则随机变量 $Z=X+Y$ 的分布列为

$$P(Z=k) = \sum_{i=0}^{k} a_i b_{k-i} \ .$$

证明 对任意非负整数 i 和 j, 根据独立性可知

$$P(X=i, Y=j) = P(X=i)P(Y=j) = a_i b_j \ .$$

因此随机变量 Z 的分布列为

$$\begin{aligned} P(Z=k) &= P(X+Y=k) \\ &= \sum_{i=0}^{k} P(X=i, Y=k-i) = \sum_{i=0}^{k} P(X=i)P(Y=k-i) = \sum_{i=0}^{k} a_i b_{k-i} \ , \end{aligned}$$