第 3 章教学要求：

1. 了解多维随机变量的概念.

2. 理解二维离散型随机变量的分布律概念，理解二维连续型随机变量的概率密度概念.

3. 理解二维随机变量的边缘分布与条件分布.

4. 了解二维随机变量的分布函数.

5. 理解随机变量的相互独立概念.

6. 会求两个相互独立的随机变量的简单函数(和、极大、极小)的分布.

第 3 章 二维随机变量及其分布

同一随机试验的结果可以因实验目的的不同做多种形式的数量化，并因而产生多维随机变量. 本章研究多维随机变量，既要研究其中每一个随机分量的性质，还要研究随机变量之间的相关性质.

§3.1 多维随机变量的概念

一般来说，当试验的结果需要同时用两个或更多个随机变量来描述时就产生了多个随机变量，称为**多维随机变量**. 例如，某产品有若干指标，这些指标对应的随机变量就构成了描述该产品的多维随机变量；演习中炮弹的落点位置可由二维或三维随机变量表示.

描述多维随机变量取值规律的方式仍有概率函数、概率分布表和概率分布图等，通常根据随机变量的维数与类型选择使用.

本章主要介绍二随机变量，至于更多维情形可仿二维情形推导.

§3.2 二维离散型随机变量及其分布

如果(X,Y)的全部可能取值为可列对，(X,Y)就是二维离散型随机变量.

本节讨论二维离散型随机变量的联合分布律、边缘分布律、条件分布律.

3.2.1 二维离散型随机变量的联合分布律

二维离散型随机变量(X,Y)的概率分布又叫(X,Y)的分布律或随机变量 X 与 Y 的**联合分布律**. 其概率函数的一般形式为

$$f(x_i,y_j)=P\{X=x_i,Y=y_j\}=p_{ij},\quad i,j=1,2,\cdots. \tag{3.1}$$

这里 $\{X=x_i,Y=y_j\}$ 表示事件“$X=x_i$”与事件“$Y=y_j$”的积事件.

概率分布表有“条形”表和“矩阵”表两种. 条形概率分布表的一般形式为

(X,Y)	$(x_1,y_1)\cdots(x_1,y_j)\cdots(x_i,y_1)\cdots(x_i,y_j)\cdots$
P	$p_{11}\quad\cdots\quad p_{1j}\quad\cdots\quad p_{i1}\quad\cdots\quad p_{ij}\quad\cdots$

矩阵概率分布表的一般形式为

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}	$\cdots$
$\vdots$	$\vdots$	$\vdots$		$\vdots$	

规定： $f(x,y)=P\{X=x,Y=y\}=0$，如果(x,y)不是(X,Y)的可能取值.

注意：在矩形概率分布表中，当某些点不是(X,Y)的取值时，为与概率确实为 0 的取值点区分，有时在非取值点的概率位置不填 0 而是打上标记“–”.

概率分布图是三维散点图.

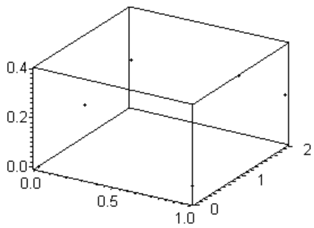


图 3. 1　二维离散型随机变量的概率分布示意图

联合分布律(3. 1)有如下性质：

(1) $p_{ij}\ge 0; i,j=1,2,\cdots$；

(2) $\sum_i\sum_j p_{ij}=1$.

根据联合分布律，可以计算任意事件的概率：

(1) (X,Y)取单点值(x,y)的概率为

$P\{X=x,Y=y\}=p_{ij}$，如果 $(x,y)=(x_i,y_j)$；

$P\{X=x,Y=y\}=0$，如果 $(x,y)\ne(x_i,y_j)(i,j=1,2,\cdots)$.

(2) (X,Y)的取值在区域 G 内的概率为

$$P\{(X,Y)\in G\}=\sum_{(x_i,y_j)\in G}p_{ij}\ . \tag{3.2}$$

这里 $\sum\limits_{(x_i,y_j)\in G}$ 表示对落在区域 G 内的(X,Y)取值的项求和.

例 3.1 (P_{76}) 已知二维随机变量(X,Y)只取四对值：(0,0),(-1,1),(-1,1/3)和(2,0)，且取它们的概率依次为 1/6, 1/3, 1/12 和 5/12，则(X,Y)的分布律

(1) 用概率函数可表示为

$P\{X=0,Y=0\}=1/6, P\{X=-1,Y=1\}=1/3$，

$P\{X=-1,Y=1/3\}=1/12, P\{X=2,Y=0\}=5/12$.

(2) 用条形概率分布表可表示为

(X,Y)	(0,0)	(−1,1)	(−1,1/3)	(2,0)
P	1/6	1/3	1/12	5/12

(3) 用矩形概率分布表可表示为

X\Y	0	1/3	1
−1	0	1/12	1/3
0	1/6	0	0
2	5/12	0	0

（或

X\Y	0	1/3	1
−1	–	1/12	1/3
0	1/6	–	–
2	5/12	–	–

）

例 3.2 (P_{77})

3.2.2 边缘分布律与条件分布律

1. 边缘分布律

二维离散型随机变量(X,Y)的分量 X 的和 Y 是一维离散型随机变量，X 的和 Y 的分布律称为(X,Y)关于 X 的和关于 Y 的**边缘分布律**，常用概率函数和概率分布表表示.

由(X,Y)的分布律(3.1)式可知，X 的可能取值为 $x_1,x_2,\cdots$，且

$$P\{X=x_i\}=P\{X=x_i,-\infty<Y<+\infty\}=\sum_j P\{X=x_i,Y=y_j\}=\sum_j p_{ij}, i=1,2,\cdots,$$

因此，

$$P\{X=x_i\}=\sum_j p_{ij}\ge 0,(i=1,2,\cdots),\quad \sum_i P\{X=x_i\}=\sum_i\sum_j p_{ij}=1.$$

所以

$$P\{X=x_i\}=\sum_j p_{ij}\overset{\Delta}{=}p_{i\cdot}, i=1,2,\cdots \tag{3.3}$$

是(X,Y)关于 X 的边缘分布律.

同理，(X,Y)关于 Y 的边缘分布律为

$$P\{Y=y_j\}=\sum_i p_{ij}\overset{\Delta}{=}p_{\cdot j},\ j=1,2,\cdots. \tag{3.4}$$

(3. 3) 式和 (3. 4) 式所对应的边缘概率分布表分别为

X	x_1	x_2	$\cdots$	x_i	$\cdots$
P	$p_{1\cdot}$	$p_{2\cdot}$	$\cdots$	$p_{i\cdot}$	$\cdots$

与

Y	y_1	y_2	$\cdots$	y_j	$\cdots$
P	$p_{\cdot 1}$	$p_{\cdot 2}$	$\cdots$	$p_{\cdot j}$	$\cdots$

显然，联合分布律决定唯一的边缘分布律.

做题时，我们可以在联合概率分布矩形表的右端加一列，在下端加一行，把两个边缘概率分布的概率值分别填入其中，得到带有边缘概率分布的综合概率分布表：

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

例 3. 3 (P_{79})

本题表明：虽然联合分布律决定边缘分布律，但边缘分布律却不能决定联合分布律.

2. 条件分布律

在第一章我们曾讨论过条件概率 P(B|A)，它表示在事件 A 已发生的条件下事件 B 再发生的概率. 把此概念运用到二维随机变量的研究中，就出现了随机变量的条件概率分布.

二维离散型随机变量中的一个随机变量的取值分布状态如果受另一个随机变量取值的影响，其取值的概率分布规律就称为**条件分布律**.

如果当 $Y=y_j$ 时 $p_{\cdot j}>0$，则考虑条件概率

$$P\{X=x_i|Y=y_j\}=\frac{P\{X=x_i,Y=y_j\}}{P\{Y=y_j\}}=\frac{p_{ij}}{p_{\cdot j}},\quad i=1,2,\cdots. \tag{3.5}$$

由于 $\frac{p_{ij}}{p_{\cdot j}}>0$ 且 $\sum_i \frac{p_{ij}}{p_{\cdot j}}=1$，因而(3.5)式描述了在 $Y=y_j$ 条件下随机变量 X 的概率分布规律，故称为在 $Y=y_j$ 条件下 X 的条件分布律.

同理，如果当 $X=x_i$ 时 $p_{i\cdot}>0$，则

$$P\{Y=y_j|X=x_i\}=\frac{P\{X=x_i,Y=y_j\}}{P\{X=x_i\}}=\frac{p_{ij}}{p_{i\cdot}},\quad j=1,2,\cdots \tag{3.6}$$

称为在 $X=x_i$ 条件下 Y 的条件分布律.

(3.5)式和(3.6)式所对应的条件概率分布表为

$X\|Y=y_j$	$x_1\quad x_2\cdots x_i\quad\cdots$
P	$\frac{p_{1j}}{p_{\cdot j}}\quad\frac{p_{2j}}{p_{\cdot j}}\cdots\frac{p_{ij}}{p_{\cdot j}}\cdots$

与

$Y\|X=x_i$	$y_1\quad y_2\cdots y_j\quad\cdots$
P	$\frac{p_{i1}}{p_{i\cdot}}\quad\frac{p_{i2}}{p_{i\cdot}}\cdots\frac{p_{ij}}{p_{i\cdot}}\cdots$

做题时，我们可以在综合概率分布表的右端与下端再分别加上一列与一行，把两个条件概率分布的概率值填入其中，便得到带有条件概率分布的大综合概率分布表：

X \ Y	$y_1\quad y_2\cdots y_j\quad\cdots$	$P\{X=x_i\}$	$P\{X=x_i\|Y=y_j\}$
x_1 x_2 $\vdots$ x_i $\vdots$	$p_{11}\quad p_{12}\cdots p_{1j}\cdots$ $p_{21}\quad p_{22}\cdots p_{2j}\cdots$ $\vdots\quad\vdots\quad\vdots$ $p_{i1}\quad p_{i2}\cdots p_{ij}\cdots$ $\vdots\quad\vdots\quad\vdots$	$p_{1\cdot}$ $p_{2\cdot}$ $\vdots$ $p_{i\cdot}$ $\vdots$	$p_{1j}/p_{\cdot j}$ $p_{2j}/p_{\cdot j}$ $\vdots$ $p_{ij}/p_{\cdot j}$ $\vdots$
$P\{Y=y_j\}$	$p_{\cdot 1}\quad p_{\cdot 2}\cdots p_{\cdot j}\cdots$	1	
$P\{Y=y_j\|X=x_i\}$	$\frac{p_{i1}}{p_{i\cdot}}\quad\frac{p_{i2}}{p_{i\cdot}}\cdots\frac{p_{ij}}{p_{i\cdot}}\cdots$		1

例 3.4(P_{80})

作业(P_{82})：1.-2.　4.　7.-11.　3.　6.

§3.3 二维连续型随机变量及其分布

以下讨论二维连续型随机变量的密度函数、边缘密度函数、条件密度函数与常见分布.

3.3.1 二维连续型随机变量的联合密度函数

如果存在实非负可积函数 $f(x,y)(-\infty < x,y < +\infty)$，使得对任意二维区域 G，随机变量(X,Y)的取值在 G 中的概率为

$$P\{(X,Y)\in D\} = \iint_G f(x,y)d\sigma, \tag{3.9}$$

则称(X,Y)是二维连续型随机变量. 这里的实非负可积函数 f(x,y)即是描述二维连续型随机变量(X,Y)的取值规律的概率函数，称为(X,Y)的**概率密度函数**，简称密度函数，并记 $(X,Y)\sim f(x,y)$.

(X,Y)的概率密度函数又叫随机变量 X 和 Y 的联合概率密度函数，简称联合密度函数.

从几何看，(X,Y)的取值在区域 G 中的概率等于以 G 为底、以曲面 $z=f(x,y)$ 为高的曲顶柱体的体积.

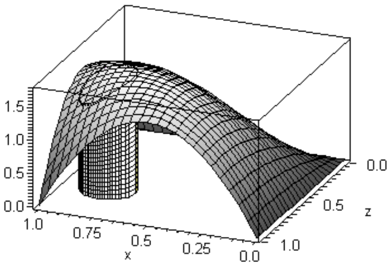


图 3.2 (3.9)式的几何解释

密度函数 f(x,y) 有如下性质：

(1) $f(x,y)\geq 0$；

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

规定： $f(x,y)=0$，如果(x,y)不是(X,Y)的可能取值.

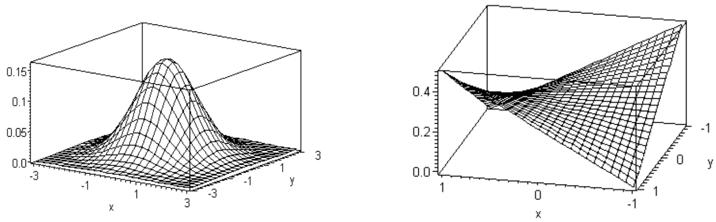


图 3.3 联合密度函数的性质示意图

在几何上，曲面 z=f(x,y)总位于 xOy 坐标面 z 轴正方向一侧，它与 xOy 坐标平面之间

的曲顶柱体体积恒为 1.

根据密度函数，可以计算任意事件的概率：

(1) (X,Y)取任意一对值(x,y)的概率为 0，即 $P\{X=x,Y=y\}=0$.

(2) (X,Y)的取值在任意一条曲线(包括直线)上的概率为 0.

(3) (X,Y)的取值在内部非空的区域 G 中的概率为

$$P\{(X,Y)\in G\}=\iint_G f(x,y)d\sigma. \tag{3.10}$$

特别地，当 $G=\{(x,y)\,|\,a<x\le b,c<y\le d\}$ 时，有

$$P\{a<X\le b,c<Y\le d\}=\int_c^d\int_a^b f(x,y)dxdy. \tag{3.11}$$

例 3.7 (P_{84}) 设随机变量(X,Y)的密度函数为

$$f(x,y)=\begin{cases} ce^{-(x+y)}, & 0<x<1,y>0, \\ 0, & \text{其它.} \end{cases}$$

确定常数 c，并求 $P\{Y<2\},P\{X-Y\le 0\}$.

解 由 $f(x,y)\ge 0$，可见 $c\ge 0$. 另由 $\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty}f(x,y)dxdy=1$，有

$$\int_0^1\int_0^{+\infty}ce^{-(x+y)}dydx=1,$$

$$c\int_0^1 e^{-x}dx\int_0^{+\infty}e^{-y}dy=1,$$

即 $c(1-e^{-1})=1$，得 $c=\dfrac{1}{1-e^{-1}}$.

$$P\{Y<2\}=\int_{-\infty}^{+\infty}\int_{-\infty}^{2}f(x,y)dydx$$

$$=\int_0^1\int_0^2\frac{1}{1-e^{-1}}e^{-(x+y)}dydx$$

$$=\frac{1}{1-e^{-1}}(1-e^{-2})(1-e^{-1})=1-e^{-2}.$$

$$P\{X-Y\le 0\}=\int_{-\infty}^{+\infty}\int_x^{+\infty}f(x,y)dydx=\int_0^1\int_x^{+\infty}\frac{1}{1-e^{-1}}e^{-(x+y)}dydx$$

$$=\frac{1}{1-e^{-1}}\int_0^1 e^{-2x}dx=\frac{1-e^{-2}}{2(1-e^{-1})}.$$

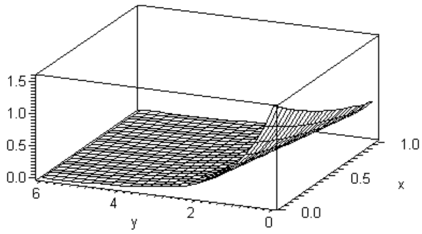


图 3.3　例 3.7 中的随机变量(X,Y)的密度函数图像

3.3.2 边缘密度函数与条件密度函数

1. 边缘概率密度函数

二维连续型随机变量(X,Y)的分量 X 和 Y 是一维连续型随机变量，它们的概率分布常用概率函数表示， 称为(X,Y)关于 X 的和关于 Y 的**边缘概率密度函数**，简称边缘密度函数.

设(X,Y)的密度函数为 f(x,y)，那么对任意实数 $a,b(a<b)$， 总有

$$P\{a<X\le b\}=P\{a<X\le b,-\infty<Y<+\infty\}=\int_a^b\left(\int_{-\infty}^{+\infty}f(x,y)dy\right)dx\,.$$

且由于

$$\int_{-\infty}^{+\infty}f(x,y)dy\ge 0\,,\quad \int_{-\infty}^{+\infty}\left(\int_{-\infty}^{+\infty}f(x,y)dy\right)dx=1\,,$$

所以

$$f_X(x)=\int_{-\infty}^{+\infty}f(x,y)dy \tag{3.12}$$

是(X,Y)关于 X 的边缘密度函数.

同理，(X,Y)关于 Y 的边缘密度函数为

$$f_Y(y)=\int_{-\infty}^{+\infty}f(x,y)dx\,. \tag{3.13}$$

显然，联合密度函数决定唯一的边缘密度函数.

从几何上看， $f_X(x)$ 在点 $\hat{x}$ 处的函数值是以 $z=f(\hat{x},y)$ 为曲边的曲边梯形的面积(图 3.4 中阴影部分)，而 $f_Y(y)$ 在点 $\hat{y}$ 处的函数值是以 $z=f(x,\hat{y})$ 为曲边的曲边梯形的面积(图 3.5 中阴影部分).

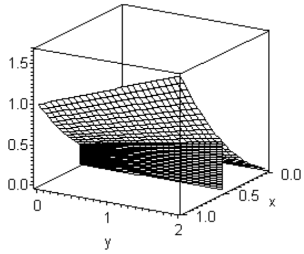


图 3. 4　关于 X 的边缘密度的函数值 $f_X(\hat{x})$ 的示意图

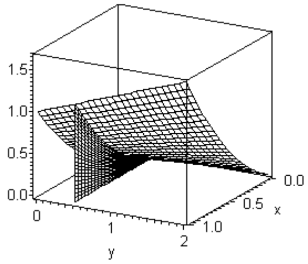


图 3. 5　关于 Y 的边缘密度的函数值 $f_Y(\hat{y})$ 的示意图

例 3. 8（P_{86}）　考虑例 3. 7 中的二维随机变量(X,Y)，求边缘密度函数 $f_X(x)$ 和 $f_Y(y)$.

解　(X,Y)的密度函数为

$$f(x,y)=\begin{cases}\dfrac{1}{1-e^{-1}}e^{-(x+y)}, & 0<x<1, y>0,\\ 0, & \text{其它.}\end{cases}$$

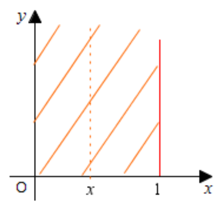
x=0 和 x=1 把 x 轴分为三个区间：(-∞,0],(0,1)和[1,+ ∞).

当 $-\infty<x\le 0$ 或 $1\le x<+\infty$ 时， $f(x,y)=0$ ， 此时 $f_X(x)=\int_{-\infty}^{+\infty}f(x,y)dy=0$ ；

当 $0<x<1$ 时，

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

$$=\frac{e^{-x}}{1-e^{-1}}\left[\int_{-\infty}^{0}0dy+\int_{0}^{+\infty}e^{-y}dy\right]=\frac{e^{-x}}{1-e^{-1}}.$$



附图　关于 X 的边缘密度函数的积分图示.

所以(X,Y)关于 X 的边缘密度函数为

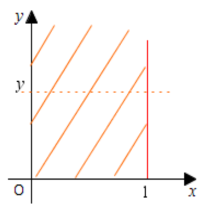
$$f_X(x)=\begin{cases}\dfrac{e^{-x}}{1-e^{-1}}, & 0<x<1,\\ 0, & \text{其它.}\end{cases}$$

又 y=0 把 y 轴分为两个区间：(-∞,0]和(0,+ ∞).

当 $y\le 0$ 时， $f(x,y)=0$ ， 此时 $f_Y(y)=\int_{-\infty}^{+\infty}f(x,y)dx=0$ ；

当 $y>0$ 时，

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



附图　关于 Y 的边缘密度函数的积分图示.

$$=\frac{e^{-y}}{1-e^{-1}}\left[\int_{-\infty}^{0}0dx+\int_{0}^{1}e^{-x}dy+\int_{1}^{+\infty}0dy\right]=e^{-y}.$$

所以(X,Y)关于 Y 的边缘密度函数为

$$f_Y(y)=\begin{cases}e^{-y}, & y>0,\\ 0, & y\le 0.\end{cases}$$

例 3.9（P_{87}）

2. 条件密度函数

设二维连续型随机变量(X,Y)的密度函数为 f(x,y)，那么在 Y=y 条件下 X 的概率分布与在 X=x 条件下 Y 的概率分布称为(X,Y)的条件概率分布. 条件概率分布常用概率函数表示，称为**条件概率密度函数**，简称条件密度函数.

将在 Y=y 条件下事件$\{a\le X\le b\}$发生的概率记作$P\{a<X\le b\,|\,Y=y\}$，在 Y=y 条件下 X 的条件密度函数记作$f_{X|Y}(x\,|\,y)$，那么对于任意实数 a,b(a<b) 及 y，只要$f_Y(y)>0$，总有

$$\begin{aligned}&P\{a<X\le b\,|\,Y=y\}\\ &=\lim_{\Delta y\to 0^+}P\{a<X\le b\,|\,y-\Delta y<Y\le y\}\\ &=\lim_{\Delta y\to 0^+}\frac{P\{a<X\le b,y-\Delta y<Y\le y\}}{P\{y-\Delta y<Y\le y\}}\\ &=\lim_{\Delta y\to 0^+}\frac{\int_a^b\int_{y-\Delta y}^{y}f(x,u)dudx}{\int_{-\infty}^{+\infty}\int_{y-\Delta y}^{y}f(x,u)dudx}\\ &=\frac{\int_a^b f(x,y)dx}{\int_{-\infty}^{+\infty}f(x,y)dx)}=\frac{\int_a^b f(x,y)dx}{f_Y(y)}=\int_a^b\frac{f(x,y)}{f_Y(y)}dx,\end{aligned}$$

而

$$\frac{f(x,y)}{f_Y(y)}\ge 0,\quad \int_{-\infty}^{+\infty}\frac{f(x,y)}{f_Y(y)}dx=1,$$

所以当$f_Y(y)>0$时，在 Y=y 条件下 X 的条件密度函数为

$$f_{X|Y}(x\,|\,y)=\frac{f(x,y)}{f_Y(y)}. \tag{3.14}$$

同理，当$f_X(x)>0$时， 在 X=x 条件下的 Y 的条件密度函数为

$$f_{Y|X}(y|x)=\frac{f(x,y)}{f_X(x)}. \tag{3.15}$$

例 3. 10（P_{88}） 考虑例 3. 9 中的二维随机变量(X,Y)，求条件密度函数 $f_{X|Y}(x|y)$ 和 $f_{Y|X}(y|x)$.

解 由例 3. 9 知，(X,Y)的密度函数及其条件密度函数分别为

$$f(x,y)=\begin{cases}24(1-x)y, & 0\le x\le 1, 0\le y\le x,\\ 0, & \text{其它.}\end{cases}$$

$$f_X(x)=\begin{cases}12x^2(1-x), & 0\le x\le 1,\\ 0, & \text{其它.}\end{cases}$$

和
$$f_Y(y)=\begin{cases}12y(1-y)^2, & 0\le y\le 1,\\ 0, & \text{其它.}\end{cases}$$

当 $f_Y(y)>0$，即当 0<y<1 而 $f_Y(y)=12y(1-y)^2$ 时，按公式（3. 14）得在 Y=y 条件下 X 的条件密度函数为

$$f_{X|Y}(x|y)=\frac{f(x,y)}{f_Y(y)}=\begin{cases}\dfrac{2(1-x)}{(1-y)^2}, & y\le x\le 1,\\ 0, & \text{其它.}\end{cases}$$

当 $f_X(x)>0$，即当 0<x<1 而 $f_X(x)=12x^2(1-x)$ 时，按公式（3. 15）得在 X=x 条件下 Y 的条件密度函数

$$f_{Y|X}(y|x)=\frac{f(x,y)}{f_X(x)}=\begin{cases}\dfrac{2y}{x^2}, & 0\le y\le x,\\ 0, & \text{其它.}\end{cases}$$

例 3. 11（P_{89}）

3. 3. 3 常见的二维连续型随机变量

1． 二维均匀变量及其分布

如果随机变量(X,Y)在有界平面区域 G 上取值，并在 G 内等面积的任意区域上取值的概率相等，则称(X,Y)是区域 G 上的二维均匀变量．(X,Y)的密度函数 f(x,y) 为

$$f(x,y)=\begin{cases}\dfrac{1}{A}, & (x,y)\in G,\\ 0, & \text{其它,}\end{cases} \tag{3.16}$$

称(X,Y)服从区域 G 上的**二维均匀分布**，并记 $(X,Y)\sim U(G)$.

例 3. 12（P_{84}） 设随机变量(X,Y)在由直线 $x=2, y=x/2$ 及 x 轴所围成的三角形区域上服

从均匀分布，求 $P\{X+Y<1\}$ 和 $P\{X\ge 1\}$.

解 直线 $x=2, y=x/2$ 及 x 轴所围成的三角形区域的面积为 1，所以(X,Y)的密度函数为

$$f(x,y)=\begin{cases}1, & 0<x<2, 0<y<\dfrac{x}{2},\\ 0, & \text{其它,}\end{cases}$$

按公式(3.10)计算可得

$$P\{X+Y<1\}=\iint_{x+y<1} f(x,y)d\sigma=\int_0^{\frac{1}{3}}\int_{2y}^{1-y}dxdy=\int_0^{\frac{1}{3}}(1-3y)dy=\frac{1}{6},$$

$$P\{X\ge 1\}=\iint_{x>1} f(x,y)d\sigma=1-\int_0^1\int_0^{\frac{x}{2}}dydx=1-\int_0^1\frac{x}{2}dx=\frac{3}{4}.$$

如果注意到均匀分布的特点，也可以按下面的方法计算这两个概率.

$$P\{X+Y<1\}=\frac{\frac{1}{2}\times 1\times\frac{1}{3}}{\frac{1}{2}\times 2\times 1}=\frac{1}{6},$$

$$P\{X\ge 1\}=1-P\{X<1\}=1-\frac{\frac{1}{2}\times 1\times\frac{1}{2}}{\frac{1}{2}\times 2\times 1}=\frac{3}{4}.$$

2. 二维正态变量及其分布

如果二维随机变量(X,Y)的密度函数为

$$f(x,y)=\frac{1}{2\pi\sigma_1\sigma_2\sqrt{1-\rho^2}}e^{-\frac{1}{2(1-\rho^2)}\left[\frac{(x-\mu_1)^2}{\sigma_1^2}-\frac{2\rho(x-\mu_1)(y-\mu_2)}{\sigma_1\sigma_2}+\frac{(y-\mu_2)^2}{\sigma_2^2}\right]}, \tag{3.17}$$

则 (X,Y) 为二维正态变量，称其服从参数为 $\mu_1,\mu_2,\sigma_1^2,\sigma_2^2,\rho$ 的**二维正态分布**，并记 $(X,Y)\sim N(\mu_1,\mu_2,\sigma_1^2,\sigma_2^2,\rho)$. 式中 $\mu_1,\mu_2,\sigma_1^2,\sigma_2^2,\rho$ 均为常量，且 $\sigma_1>0,\sigma_2>0,-1<\rho<1$.

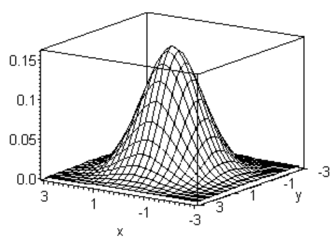


图 3.7　二维正态变量的密度函数图像

例 3.13(P_{91})

该例的结论表明：

(1) 二维正态变量的分量是一维正态变量，即边缘分布为一维正态分布.

(2) 边缘密度函数一般不能决定联合密度函数.

二维正态变量的一些性质：

①若 $(X,Y)\sim N(\mu_1,\mu_2,\sigma_1^2,\sigma_2^2,\rho)$，则 $X\sim N(\mu_1,\sigma_1^2),Y\sim N(\mu_2,\sigma_2^2)$；

②若 $(X,Y)\sim N(\mu_1,\mu_2,\sigma_1^2,\sigma_2^2,\rho)$，则 X 和 Y 相互独立的充要条件是 $\rho=0$；

③若 $(X,Y)\sim N(\mu_1,\mu_2,\sigma_1^2,\sigma_2^2,\rho)$，则

$$aX+bY\sim N(a\mu_1+b\mu_2,a^2\sigma_1^2+b^2\sigma_2^2+2ab\rho\sigma_1\sigma_2).$$

作业 (P_{92})：1.　3.　5.　6.　7.　　2.　4.

§3.4 二维随机变量的分布函数

本节讨论二维随机变量的联合分布函数、边缘分布函数、条件分布函数.

3.4.1 二维随机变量的联合分布函数

对于任意实数 x,y，二元函数

$$F(x,y)=P\{X\le x,Y\le y\} \tag{3.18}$$

称为二维随机变量(X,Y)的分布函数，或称为随机变量 X 与 Y 的**联合分布函数**.这里 $\{X\le x,Y\le y\}$表示 X 的取值不超过 x 且 X 的取值不超过 y 的积事件，而函数 F 在(x,y)处的函数值即是事件 $\{X\le x,Y\le y\}$ 发生的概率.

几何上，事件 $\{X\le x,Y\le y\}$ 的区域既是 xOy 平面上以(x,y)为右上顶点的无穷矩形区域.

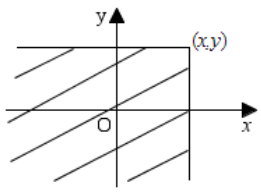


图 3.8　事件 $\{X\le x,Y\le y\}$ 的几何示意图

对于离散与连续这两种常见类型的二维随机变量，这里给出 F(x,y)的计算公式.

(1) 若(X,Y)是离散型随机变量，分布律为 (3.1) 式，则

$$F(x,y)=\sum_{x_i\le x}\sum_{y_j\le y}P\{X=x_i,Y=y_j\}=\sum_{x_i\le x}\sum_{y_j\le y}p_{ij}, \tag{3.19}$$

这里 $\sum_{x_i \le x}\sum_{y_j \le y}$ 表示对满足 $x_i \le x, y_j \le y$ 的(X,Y)取值的项求和.

(2) 若(X,Y)是连续型随机变量，密度函数为 f(x,y)，则

$$F(x,y)=\int_{-\infty}^{x}\int_{-\infty}^{y} f(u,v)dvdu . \tag{3.21}$$

反之，在密度函数 f(x,y)的任意连续点(x,y)处，有

$$f(x,y)=\frac{\partial F(x,y)}{\partial x \partial y} . \tag{3.22}$$

按定义可见，二维离散型随机变量的分布函数是关于两个变量右连续、单调不减的阶跃函数，连续型随机变量的分布函数则是关于两个变量连续的单调不减函数.

分布函数 F(x,y)有如下性质：

① $0 \le F(x,y) \le 1$；

②F(x,y)是关于 x 和关于 y 的单调不减函数；

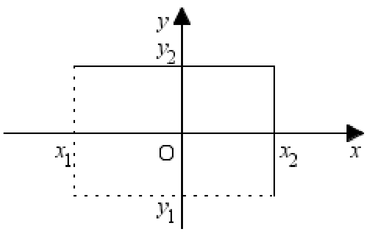
③ $F(-\infty,y)=0, F(x,-\infty)=0, F(-\infty,-\infty)=0, F(+\infty,+\infty)=1$；

④ $F(x^+,y)=F(x,y^+)=F(x,y)$，即 F(x,y)是关于 x 和关于 y 的右连续函数.

性质(1)、(2)表明曲面 z=F(x,y)是介于 z=0 与 z=1 之间的一张单调不减曲面；性质(3)、(4)则常用来确定分布函数中的待定常数.

分布函数 F(x,y)可以用来计算二维随机变量在左开右闭、下开上闭的矩形区域 $\{x_1 < X \le x_2, y_1 < Y \le y_2\}$ 上(见下图所示)取值的概率，即

$$P\{x_1 < X \le x_2, y_1 < Y \le y_2\} = F(x_2,y_2)-F(x_2,y_1)-F(x_1,y_2)+F(x_1,y_1) . \tag{3.24}$$



左开右闭、下开上闭矩形区域示意图

例 3. 14 (P_{94}) 已知二维随机变量(X,Y)的分布函数

$$F(x,y)=A\left(B+\arctan\frac{x}{2}\right)\left(C+\arctan\frac{y}{3}\right),$$

求常数 A,B,C 及 $P\{0 < X \le 2, 0 < Y \le 3\}, P\{Y < 3\}$.

3.4.2 边缘分布函数与条件分布函数

1．边缘分布函数

分量 X 和 Y 的分布函数称为(X,Y)关于 X 的和关于 Y 的**边缘分布函数**.

任意实数 x,y，总有

$$P\{X \le x\} = P\{X \le x, Y < +\infty\} = F(x, +\infty)\text{；}$$

$$P\{Y \le x\} = P\{X < +\infty, Y \le y\} = F(+\infty, y)\ .$$

所以(X,Y)关于 X 的边缘分布函数为

$$F_X(x) = F(x, +\infty)\text{，} \tag{3.24}$$

关于 Y 的边缘分布函数为

$$F_Y(y) = F(+\infty, y)\ . \tag{3.25}$$

显然，联合分布函数决定唯一的边缘分布函数.

例 求例 3.14 中的随机变量(X,Y)的边缘分布函数.

解

作业 (P_{96})：4. 5. 6.

§3.5 随机变量的独立性

在第 1 章我们曾讨论事件的独立性，事件 A,B 相互独立的定义为 P(AB)=P(A)P(B)，而且当 P(A)≠0 或 P(B)≠0 时，还可以表示为 P(B)=P(A|B)或 P(A)=P(A|B)，意为 A 发生不影响 B 发生或 B 发生不影响 A 发生，所以两个事件相互独立的本质是两个事件的发生互不影响. 将此概念运用到二维随机变量，就会产生两个随机变量相互独立的概念.

两个随机变量 X,Y 相互独立是指这两个变量之间没有任何关系，即关于 X 的任一事件与关于 Y 的任一事件皆相互独立，这可表示为对任意实数 x,y，总有

$$P\{X \le x, Y \le y\} = P\{X \le x\} \cdot P\{Y \le y\}\text{，} \tag{3.30}$$

即

$$F(x,y) = F_X(x)F_Y(y)\ . \tag{3.31}$$

以下是两个及两个以上变量相互独立的一些充分必要条件.

随机变量 X,Y 相互独立的充分必要条件是

$$F_{X|Y}(x\,|\,y) = F_X(x) \quad \text{（或 } F_{Y|X}(y\,|\,x) = F_Y(y)\text{）}. \tag{3.32}$$

离散型随机变量 X,Y 相互独立的充分必要条件是对任意 i,j=1,2,…，总有

$$P\{X=x_i, Y=y_j\}=P\{X=x_i\}\cdot P\{Y=y_j\}. \tag{3.33}$$

如果有一对下标 i,j 使(3. 33)式不成立，则表明离散型随机变量 X 与 Y 不独立.

连续型随机变量 X,Y 相互独立的充分必要条件是对任意实数 x,y，总有

$$f(x,y)=f_X(x)\cdot f_Y(y). \tag{3.34}$$

如果有一对实数 x,y 使(3. 34)式不成立，则表明连续型随机变量 X 与 Y 不独立.

总结：若 X,Y 是相互独立的连续型随机变量，则一定有 f(x,y)=g(x,y)h(x,y)，且 f(x,y)≠0 的各段定义域必为矩形区域.

连续型随机变量 X,Y 相互独立的充分必要条件是

$$f_{X|Y}(x|y)=f_X(x) \quad （或 f_{Y|X}(y|x)=f_Y(y)）. \tag{3.35}$$

如果有实数 x 使 $f_{X|Y}(x|y)\neq f_X(x)$（或有实数 y 使 $f_{Y|X}(y|x)\neq f_Y(y)$），则表明连续型随机变量 X 与 Y 不独立.

前面已经知道：联合分布决定边缘分布，但反之不成立. 现在又知道：当随机变量相互独立时，边缘分布也决定联合分布.

两个变量相互独立的意义：一是 X 的事件与 Y 的事件同时发生的概率可以简化为 X 事件发生的概率与 Y 事件发生的概率的乘积；二是 X,Y 的联合分布即为 X 的边缘分布与 Y 的边缘分布的乘积.

例 3. 15(P_{97})

例 3. 16(P_{98})

例 3. 17(P_{98})

该例表明，二维正态变量的分量相互独立的充分必要条件是 ρ=0.

在实际中，判断两个随机变量是否相互独立主要看两个随机变量是否相互有影响，没影响或影响微弱则可以认为它们相互独立，否则认为不独立.

例 3. 18(P_{99})

类似地，可以定义三个及三个以上随机变量的相互独立概念.

一般地，$n(n \ge 3)$ 个随机变量 $X_1,X_2,\cdots,X_n$ 相互独立，如果对任意实数 $x_1,x_2,\cdots,x_n$，总有

$$F(x_1,x_2,\cdots,x_n)=F_{X_1}(x_1)F_{X_2}(x_2)\cdots F_{X_n}(x_n).$$

离散型随机变量 $X_1,X_2,\cdots,X_n$ 相互独立的充分必要条件是，对 $(X_1,X_2,\cdots,X_n)$ 的任意一组取值 $(x_1,x_2,\cdots,x_n)$，总有

$$P\{X_1=x_1,X_2=x_2,\cdots,X_n=x_n\}=P\{X_1=x_1\}P\{X_2=x_2\}\cdots P\{X_n=x_n\}.$$

连续型随机变量 $X_1,X_2,\cdots,X_n$ 相互独立的充分必要条件是，对任意一组实数 $x_1,x_2,\cdots,x_n$，总有

$$f(x_1,x_2,\cdots,x_n)=f_{X_1}(x_1)f_{X_2}(x_2)\cdots f_{X_n}(x_n).$$

例 3.19 (P_{99})　设 $X_i \sim G(p)(i=1,2,\cdots,n)$，且相互独立，则 $X_1,X_2,\cdots,X_n$ 的联合分布律为

$$P\{X_1=k_1,X_2=k_2,\cdots,X_n=k_n\}=q^{k_1+k_2+\cdots+k_n}\left(\frac{p}{q}\right)^n,\quad k_1,k_2,\cdots,k_n=1,2,\cdots.$$

例 3.20 (P_{99})　若 $X_i \sim E(\lambda_i)(i=1,2,\cdots,n)$，且相互独立，则 $X_1,X_2,\cdots,X_n$ 的联合密度函数为

$$f(x_1,x_2,\cdots,x_n)=\begin{cases}\lambda_1\lambda_2\cdots\lambda_n e^{-(\lambda_1x_1+\lambda_2x_2+\cdots+\lambda_nx_n)}, & x_1>0,x_2>0,\cdots,x_n>0,\\ 0, & \text{其它}.\end{cases}$$

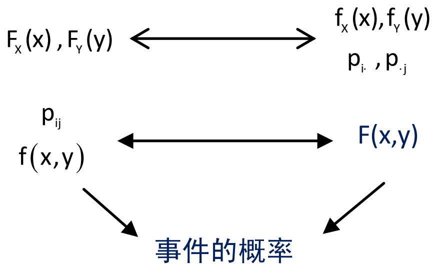
还可以定义两组随机变量相互独立的概念（见 P_{100}）.

这里不加证明地指出：若 $X_1,X_2,\cdots,X_m$ 与 $Y_1,Y_2,\cdots,Y_n$ 相互独立，则随机变量 $H=h(X_1,X_2,\cdots,X_m)$ 与 $G=g(Y_1,Y_2,\cdots,Y_n)$（h,g 为连续函数）也相互独立.

关系图：

$$\begin{array}{ccc} F(x,y) & \longleftrightarrow & \begin{array}{c} p_{ij} \\ f(x,y) \end{array} \\ \text{X,Y 独立} \ \uparrow\downarrow & & \uparrow\downarrow \ \text{X,Y 独立} \end{array}$$

关系图：



作业(P_{100})：　3.　4.　7.　8.　10.　11.　12.　1.-2.　5.-6.　9.　13.

§3.6 二维随机变量的函数的分布

二维随机变量的函数的分布问题的一般提法是：已知随机变量(X,Y)的概率分布，求随机变量 Z=g(X,Y)(g 一般为连续函数)的概率分布.

下面针对(X,Y)是离散型或连续型的，或 X,Y 一个为离散型的而另一个为连续型的二维随机变量，分别讨论 g 为连续函数时的 Z=g(X,Y)的概率分布.

1. (X,Y)为离散型的情形

如果(X,Y)是离散型随机变量，那么由连续函数 g 所产生的随机变量 Z=g(X,Y)也是离散型随机变量.

下面举例说明如何求二维离散型随机变量的函数的概率分布.

例 3.21 (P_{101})

例 3.22 (P_{102})

该例表明：相互独立的泊松变量之和仍为泊松变量.

2. (X,Y)为连续型的情形

如果(X,Y)是连续型随机变量，那么由连续函数 g 所产生的随机变量 Z=g(X,Y)也是连续型随机变量.

设(X,Y)的密度函数为 f(x,y)，则求函数 Z=g(X,Y)的概率分布的一般方式为：利用(X,Y)的密度函数 f(x,y)先求出 Z 的分布函数 $F_Z(z)$，然后对 $F_Z(z)$求导数得到 Z 的密度函数 $f_Z(z)$. 具体步骤如下：

(1) 求 Z 的分布函数

$$F_Z(z)=P\{Z\le z\}$$ （形式上写出 Z 的分布函数）

$$=P\{g(X,Y)\le z\}$$ （作变量代换，转化为关于(X,Y)的事件）

$$=\iint\limits_{g(x,y)\le z} f(x,y)d\sigma.$$ （计算关于(X,Y)的事件的概率）

(2) 求 Z 的密度函数

$$f_Z(z)=\frac{dF_Z(z)}{dz}.$$

按照以上步骤，我们可以推导出以下几个常用的二维随机变量的函数的分布.

(1) Z=X+Y 的分布

$$f_Z(z)=\int_{-\infty}^{+\infty}f(x,z-x)dx\text{，}\tag{3.36}$$

与

$$f_Z(z)=\int_{-\infty}^{+\infty}f(z-y,y)dy.\tag{3.37}$$

(2) Z=max{X,Y}的分布

$$F_Z(z)=F(z,z)\text{，}\tag{3.38}$$

$$f_Z(z)=\frac{dF(z,z)}{dz}.\tag{3.39}$$

当 X,Y 独立时，

$$F_Z(z)=F_X(z)\cdot F_Y(z)\text{，}$$

$$f_Z(z)=f_X(z)\cdot F_Y(z)+F_X(z)\cdot f_Y(z)\text{；}$$

(3) Z=min{X,Y}的分布

$$F_Z(z)=F_X(z)+F_Y(z)-F(z,z)\text{，}\tag{3.40}$$

$$f_Z(z)=f_X(z)+f_Y(z)-\frac{dF(z,z)}{dz}.\tag{3.41}$$

当 X,Y 独立时，将 $F(x,y)=F_X(x)F_Y(y)$ 带入以上四式，将得到独立条件下的结果.

例 3. 23 (P_{104})

结果表明：两个相互独立的正态变量的和变量仍为正态变量，且 $Z\sim N(\mu_1+\mu_2,\sigma_1^2+\sigma_2^2)$.

归纳推广：一般地，若 $X_1,X_2,\cdots,X_n$ 相互独立，且 $X_i\sim N(\mu_i,\sigma_i^2)(i=1,2,\cdots,n)$，则

$$\sum_{i=1}^n X_i\sim N\left(\sum_{i=1}^n\mu_i,\sum_{i=1}^n\sigma_i^2\right).$$

这被称为正态变量的可加性.

更一般地，若 $X_1,X_2,\cdots,X_n$ 相互独立，且 $X_i \sim N(\mu_i,\sigma_i^2)(i=1,2,\cdots,n)$，则

$$\sum_{i=1}^{n} a_i X_i \sim N\left(\sum_{i=1}^{n} a_i\mu_i, \sum_{i=1}^{n}(a_i\sigma_i)^2\right).$$

即相互独立的正态变量的线性组合仍然是正态变量.

例 3.24 (P_{105})

通过本例可理解概率中的串联、并联和备用的实际含义.

例 3.25 (P_{107}) 设二维随机变量(X,Y)的密度函数为

$$f(x,y)=\begin{cases}1, & 0<x<1, 0<y<2x,\\ 0, & \text{其它}.\end{cases}$$

求 Z=2X-Y 密度函数.

解 首先求 Z 的分布函数，

$$F_Z(z)=P\{Z\le z\}=P\{2X-Y\le z\}=\iint_{2x-y\le z} f(x,y)d\sigma.$$

注意到只有当 $0<x<1, 0<y<2x$ 时才有 $f(x,y)\ne 0$ （如图 3.13 所示）. 所以

当 $z\le 0$ 即 $2x-y\le 0$ 时， $f(x,y)=0$， 此时 $F_Z(z)=0$， 因而 $f_Z(z)=0$；

当 $0<z<2$ 时，

$$F_Z(z)=1-P\{2X-Y>z\}=1-\int_{z/2}^{1}\int_{0}^{2x-z}dydx=z-z^2/4,$$

此时 $f_Z(z)=1-z/2$；

当 $z\ge 2$ 时， $F_Z(z)=1$， 此时 $f_Z(z)=0$.

综上，Z 的密度函数为

$$f_Z(z)=\begin{cases}1-z/2, & 0<z<2,\\ 0, & \text{其它}.\end{cases}$$

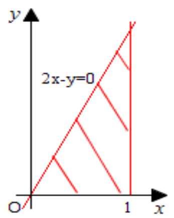


图 3.13 $f(x,y)\ne 0$ 区域图示

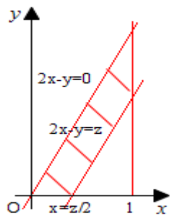


图 3.14 $0<z<2$ 时的积分区域图示

3*. X,Y 一个为离散型而另一个为连续型的情形

例 3.26 (P_{108})

例 3.27 (P_{108})

作业（P_{108}）：1． 4． 5． 9． 2． 8． 10.