西安電子科技大學
XIDIAN UNIVERSITY

# 第三章 多维随机变量及其分布



**定义：** 设二维连续型随机变量$(X, Y)$的联合概率密度为 $f(x, y)$，

如果对于任意固定的$y$，有 $f_Y(y) > 0$，则称

$$f_{X|Y}(x\mid y)=\frac{f(x,y)}{f_Y(y)},\quad -\infty<x<\infty$$

为在$Y=y$条件下随机变量X的条件概率密度

如果对于任意固定的$x$，有 $f_X(x) > 0$，则称

$$f_{Y|X}(y\mid x)=\frac{f(x,y)}{f_X(x)},\quad -\infty<y<\infty$$

为在$X=x$条件下随机变量Y的条件概率密度。

**例3：** 设二维随机变量$(X, Y)$在区域G上服从均匀分布，其中G是由$x-y=0$，$x+y=2$与$y=0$所围成的三角形区域。

试求：(1)$(X, Y)$关于X的边缘概率密度$f_X(x)$；(2) 条件概率密度$f_{X|Y}(x|y)$。

**解：** (1)由题设知$(X, Y)$的联合概率密度为：

$$f(x,y)=\begin{cases}1, & (x,y)\in G\\ 0, & (x,y)\notin G\end{cases}\qquad f_X(x)=\int_{-\infty}^{+\infty}f(x,y)dy=\begin{cases}\int_0^x dy=x, & 0<x<1\\ \int_0^{2-x}dy=2-x, & 1\le x<2\\ 0, & Others\end{cases}$$

(2) $$f_Y(y)=\int_{-\infty}^{+\infty}f(x,y)dx=\begin{cases}\int_y^{2-y}dx=2(1-y), & 0<y<1\\ 0, & Others\end{cases}$$

$\forall\ 0<y<1$，$f_Y(y)>0$，X的条件概率密度为：

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

# 3.4 随机变量的独立性

$X$，$Y$独立是指与随机变量$X$有关的任一事件的发生与否和随机变量$Y$有关的任一事件发生无关。

**随机变量独立性的定义：**

设二维随机变量$(X, Y)$的联合分布函数为$F(x, y)$，边缘分布函数分别为$F_X(x)$和$F_Y(y)$，如果 $F(x, y)=F_X(x)F_Y(y)$，$x$，$y \in \mathrm{R}$，则称随机变量$X$和$Y$相互独立。

——分布函数

从定义可以看出，随机变量$X$和$Y$相互独立等价于：$\forall\, x$，$y \in \mathrm{R}$，随机事件$\{X \le x\}$与$\{Y \le y\}$相互独立，即：$P(X \le x, Y \le y) = P(X \le x)\, P(Y \le y)$，$x$，$y \in \mathrm{R}$。

## 随机变量独立性的定义：

**定理1：** 设二维连续型随机变量 $(X, Y)$ 的联合概率密度为 $f(x, y)$ 及边缘概率密度 $f_X(x)$ 和 $f_Y(y)$ 是除有限点之外的连续函数，则X与Y相互独立的**充要条件**是：

$$f(x, y) = f_X(x)\, f_Y(y), \quad x, y \in R$$

——概率密度

**定理2：** 设二维离散型随机变量 $(X, Y)$ 的联合分布律为 $P(X=x_i, Y=y_j) = p_{ij}\ (i, j=1,2,\ldots)$，边缘分布律分别为 $p_{i\bullet}$ 和 $p_{\bullet j}$，则X与Y相互独立的**充要条件**是：

$$p_{ij} = p_{i\bullet}\, p_{\bullet j}, \quad i, j=1, 2, \ldots$$

——分布律

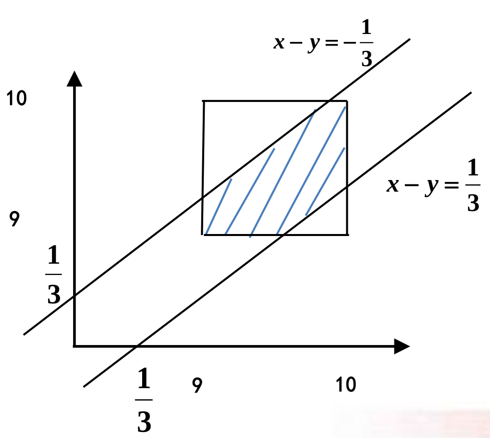
**例1：**两人约定9:00-10:00之间在公园门口见面，先到者等待二十分钟后离开，问两人能见面的概率。

**解：** 设 $X$: 第一个人到达公园门口的时间；$Y$: 第二个人到达公园门口的时间

则：$X \sim U(9,10)$ $Y \sim U(9,10)$ 且 $X$、$Y$ 独立。问题转化为求 $P(|X-Y| \le \frac{1}{3})$

$$f_X(x) = I_{(9,10)}(x) \quad f_Y(y) = I_{(9,10)}(y) \quad f(x,y) = f_X(x) f_Y(y)$$

$$P(|X-Y| \le \frac{1}{3}) = \iint\limits_{|X-Y| \le \frac{1}{3}} I_{(9,10)}(y) I_{(9,10)}(x) dxdy$$




即为求阴影部分面积： $= 1 - \left(\frac{2}{3}\right)^2 = \frac{5}{9}$

例2：$X$、$Y$ 相互独立, $X \sim N(0,1)$ ，Y的分布律为 $P(Y=0)=P(Y=1)=\frac{1}{2}$ ，求 $P(X+Y \le \frac{1}{2})$ 。

解：由全概率公式：

$$P(X+Y \le \frac{1}{2}) = P(X+Y \le \frac{1}{2} | Y=0)P(Y=0) + P(X+Y \le \frac{1}{2} | Y=1)P(Y=1)$$

$$= \frac{1}{2}P(X+Y \le \frac{1}{2} | Y=0) + \frac{1}{2}P(X+Y \le \frac{1}{2} | Y=1)$$

由于$X$与$Y$相互独立

原式 $= \frac{1}{2}P(X \le \frac{1}{2}) + \frac{1}{2}P(X \le -\frac{1}{2}) = \frac{1}{2}\left[\Phi(\frac{1}{2}) + \Phi(-\frac{1}{2})\right] = \frac{1}{2}$

若$X$ $Y$独立：$P(X \le x, Y \le y) = P(X \le x)\, P(Y \le y)$

$P(a \le X \le b, c \le Y \le d) = P(a \le X \le b)\, P(c \le Y \le d)$

$F(x, y)=F_X(x)F_Y(y)$ $\quad f(x, y) = f_X(x)\, f_Y(y)$

**独立性的判断：**

**一般步骤：**联合密度/分布律 ⟶ 边缘密度/分布律

判断：$f(x, y) \overset{?}{=} f_X(x)\, f_Y(y)$ $\quad p_{ij} \overset{?}{=} p_{i\bullet}\, p_{\bullet j}$

$(X,Y) \sim f(x,y)$，$X$ $Y$独立 $\Leftrightarrow f(x,y) = g_1(x) g_2(y)$

此时必存在实数$a$，$b$，使 $a g_1(x) = f_X(x), b g_2(y) = f_Y(y)$

即，**是否独立在于判断概率密度能否分离变量**。

## 独立性的判断:

若（$X,Y$）服从矩形（$(0,1)\times(0,1)$上的均匀分布，$X$，$Y$是否独立？

$$f(x,y)=I_{(0,1)\times(0,1)}(x,y)=I_{(0,1)}(x)I_{(0,1)}(y)$$ ，所以$X$，$Y$独立。

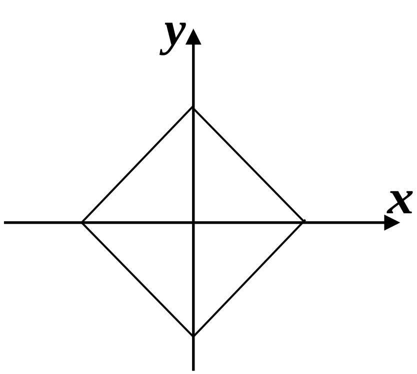
若 $(X,Y)\sim f(x,y)=\frac{1}{\pi}I_{x^2+y^2\le 1}(x,y)$（单位圆内的均匀分布），$X$，$Y$是否独立？

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

同理： $f_Y(y)=\frac{2}{\pi}\sqrt{1-y^2}\quad ,x\in(-1,1)$

$$f_X(x)f_Y(y)\neq f(x,y)$$

故$X$，$Y$不独立。




若（$X,Y$）满足如图正方形的均匀分布，$X$，$Y$是否独立？

## 独立性的判断：

**例1** 设二维离散型随机变量$X, Y$的分布律分别为

| $X$ | $-1$ | $0$ | $1$ |
|---|---|---|---|
| $p$ | $\frac{1}{4}$ | $\frac{1}{2}$ | $\frac{1}{4}$ |

| $Y$ | $1$ | $0$ |
|---|---|---|
| $p$ | $\frac{1}{2}$ | $\frac{1}{2}$ |

且$P(XY=0)=1$，问X，Y是否相互独立？

**解：** 由于$P(XY=0)=1$，因此 $P(X=-1,Y=1)=P(X=1,Y=1)=0$

从而，$(X, Y)$的联合分布律有如下结构：

| Y \ X | 0 | 1 | $p_{i\bullet}$ |
|---|---|---|---|
| $-1$ | $p_{11}$ | 0 | $\frac{1}{4}$ |
| 0 | $p_{21}$ | $p_{22}$ | $\frac{1}{2}$ |
| 1 | $p_{31}$ | 0 | $\frac{1}{4}$ |
| $p_{\bullet j}$ | $\frac{1}{2}$ | $\frac{1}{2}$ | 1 |

由联合分布律和边缘分布律关系得：

| $Y$ \ $X$ | 0 | 1 | $p_{i\bullet}$ |
|---|---|---|---|
| $-1$ | $\frac{1}{4}$ | 0 | $\frac{1}{4}$ |
| 0 | 0 | $\frac{1}{2}$ | $\frac{1}{2}$ |
| 1 | $\frac{1}{4}$ | 0 | $\frac{1}{4}$ |
| $p_{\bullet j}$ | $\frac{1}{2}$ | $\frac{1}{2}$ | 1 |

由于 $P(X=-1,Y=1)=0\neq\frac{1}{4}\times\frac{1}{2}=P(X=-1)P(Y=1)$，因此，$X, Y$不相互独立

**独立性的判断：**

**简单判断二维随机变量$X$与$Y$不独立的方法：**

若（$X,Y$）为离散型随机变量，当联合分布律中至少有一个$p_{ij}=0$ 时，$X,Y$肯定不独立。

若（$X,Y$）为连续型随机变量，G不为矩形形式$(a \le x \le b),(c \le y \le d)$，$X,Y$肯定不独立。

**独立性的判断：**

**定理** 设二维正态随机变量 $(X,Y)\sim N(\mu_1,\mu_2,\sigma_1^2,\sigma_2^2,\rho)$，则$X, Y$相互独立的充要条件是$\rho=0$。

**证明** 设 $(X,Y)\sim N(\mu_1,\mu_2,\sigma_1^2,\sigma_2^2,\rho)$, 则$(X, Y)$ 的联合概率密度及边缘概率密度分别为

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

$$f_X(x)=\frac{1}{\sqrt{2\pi}\sigma_1}e^{-\frac{(x-\mu_1)^2}{2\sigma_1^2}},\ -\infty<x<+\infty$$

$$f_Y(y)=\frac{1}{\sqrt{2\pi}\sigma_2}e^{-\frac{(y-\mu_2)^2}{2\sigma_2^2}},\ -\infty<y<+\infty$$

**独立性的判断:**

**充分性:** 设$\rho=0$，则$(X, Y)$的联合概率密度为

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

$$=\frac{1}{\sqrt{2\pi}\sigma_1}e^{-\frac{(x-\mu_1)^2}{2\sigma_1^2}}\frac{1}{\sqrt{2\pi}\sigma_2}e^{-\frac{(y-\mu_2)^2}{2\sigma_2^2}}=f_X(x)f_Y(y),\quad x,y\in R$$

故$X, Y$相互独立.

**必要性:** 设$X, Y$相互独立，则 $f(x,y)=f_X(x)f_Y(y),\qquad x,y\in R$

特别的，取$x=\mu_1$，$y=\mu_2$，得 $f(\mu_1,\mu_2)=f_X(\mu_1)f_Y(\mu_2)$

即 $$\frac{1}{2\pi\sigma_1\sigma_2\sqrt{1-\rho^2}}=\frac{1}{\sqrt{2\pi}\sigma_1}\frac{1}{\sqrt{2\pi}\sigma_2}$$

故$\rho=0$。

## 推广至$n$维随机变量 ($n\geq2$)

设$E$是一个随机试验，样本空间$\Omega=\{e\}$；$X_1=X_1(e), X_2=X_2(e),\ldots, X_n=X_n(e)$是定义在$\Omega$上的随机变量，$n$维向量$(X_1, X_2, \ldots, X_n)$称$n$维随机变量或$n$维随机向量

### $n$维随机变量的联合分布函数

对于任意实数$n$个实数$x_1, x_2, \ldots, x_n$有$n$元函数 $F(x_1,x_2,\ldots,x_n)=P(X_1\leq x_1,X_2\leq x_2,\ldots,X_n\leq x_n)$

称为$n$维随机变量 $(X_1, X_2, \ldots, X_n)$的联合分布函数

### $n$维离散型随机变量的联合分布律

设 $(X_1, X_2, \ldots, X_n)$所有可能取值为$i_j=1, 2, \ldots$

$$(x_{1i_1},x_{2i_2},\cdots,x_{ni_n})$$

$$P(X_1=x_{1i_1},X_2=x_{2i_2},\cdots,X_n=x_{ni_n})\quad j=1,2,\cdots n$$

称$n$维离散型随机变量 $(X_1, X_2, \ldots, X_n)$的联合分布律

### $n$维连续型随机变量的联合概率密度

若存在非负函数 $f(x_1, x_2, \ldots, x_n)$，使得对于任意实数$x_1, x_2, \ldots, x_n$，有

$$F(x_1,x_2,\cdots x_n)=\int_{-\infty}^{x_n}\int_{-\infty}^{x_{n-1}}\cdots\int_{-\infty}^{x_1}f(x_1,x_2,\cdots x_n)dx_1dx_2\cdots dx_n$$

称其为$n$维连续型随机变量 $(X_1, X_2, \ldots, X_n)$的联合概率密度

## 边缘分布

$(X_1, X_2, \ldots, X_n)$的联合分布函数 $F(x_1, x_2, \ldots, x_n)$已知，则$(X_1, X_2, \ldots, X_n)$的**$k$ $(1\le k \le n)$维边缘分布函数**就随之确定

例如边缘分布函数

$$F_{X_1}(x_1) = F(x_1, +\infty, +\infty, \cdots, +\infty)$$

$$F_{X_1,X_2}(x_1, x_2) = F(x_1, x_2, +\infty, \cdots, +\infty)$$

其边缘分布律

$$P(X_1 = x_{1i_1}) = \sum_{i_2, i_3, \cdots i_n} P(X_1 = x_{1i_1}, X_2 = x_{2i_2}, \cdots, X_n = x_{ni_n})$$

$$P(X_1 = x_{1i_1}, X_2 = x_{2i_2}) = \sum_{i_3, i_4, \cdots i_n} P(X_1 = x_{1i_1}, X_2 = x_{2i_2}, \cdots, X_n = x_{ni_n})$$

其边缘概率密度

$$f_{X_1}(x_1) = \int_{-\infty}^{+\infty} \cdots \int_{-\infty}^{+\infty} f(x_1, x_2, \cdots, x_n) dx_2 dx_3 \cdots dx_n$$

$$f_{X_1,X_2}(x_1, x_2) = \int_{-\infty}^{+\infty} \cdots \int_{-\infty}^{+\infty} f(x_1, x_2, \cdots, x_n) dx_3 dx_4 \cdots dx_n$$

## 相互独立

若对于所有$x_1, x_2, \ldots, x_n$，有

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

则称$X_1, X_2, \ldots, X_n$相互独立

若对于所有$x_1, x_2, \ldots, x_m$； $y_1, y_2, \ldots, y_n$有

$$F(x_1, x_2, \cdots, x_m, y_1, y_2, \cdots, y_n) = F_1(x_1, x_2, \cdots x_m)F_2(y_1, y_2, \cdots y_n)$$

其中$F_1$、$F_2$、$F$依次为随机变量$(X_1, X_2, \ldots, X_m)$、$(Y_1, Y_2, \ldots, Y_n)$、$(X_1, X_2, \ldots, X_m, Y_1, Y_2, \ldots, Y_n)$的联合分布函数

则称$(X_1, X_2, \ldots, X_m)$与$(Y_1, Y_2, \ldots, Y_n)$相互独立

### 定理1

设$(X_1, X_2, \ldots, X_m)$与$(Y_1, Y_2, \ldots, Y_n)$相互独立，则$X_i\,(i=1, 2, \ldots, m)$与$Y_j\,(j=1, 2, \ldots, n)$ 相互独立

设$h(x_1, x_2, \ldots, x_m)$与$g(y_1, y_2, \ldots, y_n)$是连续函数，则$h(X_1, X_2, \ldots, X_m)$与$g(Y_1, Y_2, \ldots, Y_n)$相互独立

### 定理2

设$X_1, X_2, \ldots, X_m$相互独立，将其分成任意$k$个没有相同随机变量的不同小组，并对每个小组的随机变量施以相应连续函数运算后，所得到的$k$个随机变量也相互独立

# 3.6 随机向量函数的分布

## 二维随机变量函数的定义：

设 $(X, Y)$是二维随机变量，$z=g(x, y)$为已知的连续函数，则称$Z=g(X, Y)$为二维随机变量$(X, Y)$的函数。显然，二维随机变量函数是一维随机变量。

## 1、二维离散型随机变量函数的分布

**例** 设二维离散型随机变量$(X, Y)$的联合分布律为

| Y \ X | 0 | 1 |
|---|---|---|
| 1 | $\frac{4}{25}$ | $\frac{6}{25}$ |
| 2 | $\frac{7}{25}$ | $\frac{8}{25}$ |

求Z=X+Y的分布律。

**解：** 由X和Y的取值易知Z=X+Y的取值为1, 2, 3

$$P(Z=1)=P(X=0,Y=1)=\frac{4}{25}$$

$$P(Z=2)=P(X=0,Y=2)+P(X=1,Y=1)=\frac{6}{25}+\frac{7}{25}=\frac{13}{25}$$

$$P(Z=3)=P(X=1,Y=2)=\frac{8}{25}$$

即Z=X+Y的分布律为：

| $Z$ | 1 | 2 | 3 |
|---|---|---|---|
| $P$ | $\frac{4}{25}$ | $\frac{13}{25}$ | $\frac{8}{25}$ |

## 1、二维离散型随机变量函数的分布

设随机变量$X, Y$相互独立，且$X\sim P(\lambda_1)$， $Y\sim P(\lambda_2)$，证明Z=X+Y $\sim P(\lambda_1+\lambda_2)$。

**证明：** 由于X和Y分别服从参数为$\lambda_1$，$\lambda_2$的泊松分布，因此其分布律分别为

$$P(X=k)=\frac{\lambda_1^k}{k!}e^{-\lambda_1},\ k=0,1,2,\cdots$$

$$P(Y=l)=\frac{\lambda_2^l}{l!}e^{-\lambda_2},\ l=0,1,2,\cdots$$

$$P(Z=m)=\sum_{k=0}^{m}P(X=k,Y=m-k)\ \xlongequal{X,Y\text{ 相互独立}}\ \sum_{k=0}^{m}P(X=k)P(Y=m-k)$$

$$=\sum_{k=0}^{m}\frac{\lambda_1^k}{k!}e^{-\lambda_1}\frac{\lambda_2^{m-k}}{(m-k)!}e^{-\lambda_2}=\frac{1}{m!}\sum_{k=0}^{m}\frac{m!}{k!(m-k)!}\lambda_1^k\lambda_2^{m-k}e^{-(\lambda_1+\lambda_2)}$$

$$=\frac{1}{m!}\left(\sum_{k=0}^{m}C_m^k\lambda_1^k\lambda_2^{m-k}\right)e^{-(\lambda_1+\lambda_2)}=\frac{(\lambda_1+\lambda_2)^m}{m!}e^{-(\lambda_1+\lambda_2)},m=0,1,2,\cdots$$

从而 Z=X+Y $\sim P(\lambda_1+\lambda_2)$

## 2、二维连续型随机变量函数的分布

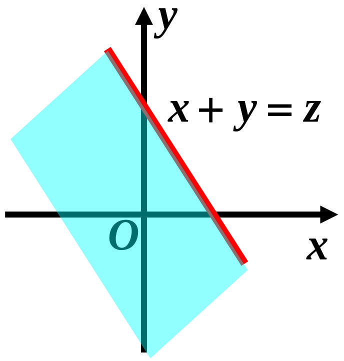
二维连续型随机变量$(X, Y)$的联合概率密度为$f(x, y)$,则Z的分布函数为：

$$F_z(z)=P(Z\le z)=P(g(X,Y)\le z)=\iint\limits_{g(x,y)\le z} f(x,y)dxdy$$

从而$Z=g(X, Y)$的概率密度为 $f_z(z)=F_z'(z)$

<1> 当Z=X+Y时， $F_z(z)=P(Z\le z)=P(X+Y\le z)=\iint\limits_{x+y\le z} f(x,y)dxdy$

将二重积分化成二次积分(积分区域如图所示)，得：




$$F_z(z)=\int_{-\infty}^{+\infty}\left[\int_{-\infty}^{z-y} f(x,y)dx\right]dy$$

$$=\int_{-\infty}^{z-y}\left[\int_{-\infty}^{+\infty} f(x,y)dy\right]dx$$

$$F_z(z)=\int_{-\infty}^{z-y}\left[\int_{-\infty}^{+\infty} f(x,y)dy\right]dx=\int_{-\infty}^{z-y}\left[\int_{-\infty}^{+\infty} f(z-y,y)dy\right]dx$$

由分布函数与概率密度的关系，得 $f_Z(z)=\int_{-\infty}^{+\infty} f(z-y,y)dy,\quad z\in R$

同理可得 $f_Z(z)=\int_{-\infty}^{+\infty} f(x,z-x)dx,\quad z\in R$

若X和Y相互独立，则 $f_Z(z)=\int_{-\infty}^{+\infty} f_X(z-y)f(y)dy,\quad z\in R$

$$=\int_{-\infty}^{+\infty} f_X(x)f_Y(z-x)dx,\quad z\in R$$

$$\triangleq f_X * f_Y$$ 卷积