概率统计第七讲：多维随机变量

史灵生　清华数学系

1、本讲提要

1. 多维随机变量
   - 联合分布函数
   - 二维随机变量
   - 独立性
2. 随机变量函数的分布
   - 离散型
   - 连续型

2、联合分布函数

定义

- 若X,Y是同一概率空间上的随机变量，则称(X,Y)为二维随机变量。
- 称$F:\mathbb{R}^2\to\mathbb{R}$，$F(x,y)=P(X\le x,Y\le y)$为$(X,Y)$的联合分布函数（简称联合分布）。
- X的边际分布函数（简称边际分布）为

$$F_X(x):=P(X\le x)=F(x,\infty).$$

- Y的边际分布函数（简称边际分布）为

$$F_Y(y):=P(Y\le y)=F(\infty,y).$$

3、分布函数

定理

1. 单调性：当$x_1<x_2$时，$F(x_1,y)\le F(x_2,y)$；
   当$y_1<y_2$时，$F(x,y_1)\le F(x,y_2)$。
2. 有界性：
   $F(-\infty,-\infty)=0\le F(x,y)\le 1=F(\infty,\infty)=\lim\limits_{x,y\to\infty}F(x,y)$
   $F(-\infty,y)=\lim\limits_{x\to-\infty}F(x,y)=0$，
   $F(x,-\infty)=\lim\limits_{y\to-\infty}F(x,y)=0$。
3. 右连续性：$F(x+0,y)=F(x,y)$，$F(x,y+0)=F(x,y)$。
4. 非负性：对任意的$a<b,c<d$有

$$P(a<X\le b,c<Y\le d)=F(b,d)-F(a,d)-F(b,c)+F(a,c)\ge 0.$$

4、二维离散型随机变量

- 二维离散型随机变量(X,Y)的联合分布列为:

$$P(\{\omega : X(\omega)=x_i, Y(\omega)=y_j\}) = p_{i,j}.$$

- 我们通常用矩形表来表示其联合分布：

$X\backslash Y$	y_1	y_2	$\cdots$	y_j	$\cdots$	P_X
x_1	$p_{1,1}$	$p_{1,2}$	$\cdots$	$p_{1,j}$	$\cdots$	p_1^X
x_2	$p_{2,1}$	$p_{2,2}$	$\cdots$	$p_{2,j}$	$\cdots$	p_2^X
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$
x_i	$p_{i,1}$	$p_{i,2}$	$\cdots$	$p_{i,j}$	$\cdots$	p_i^X
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$
P_Y	p_1^Y	p_2^Y	$\cdots$	p_j^Y	$\cdots$	1

- 令$p_i^X = P(X=x_i)$，$p_j^Y = P(Y=y_j)$，并分别称之为(X,Y)关于X和Y的边际分布。
- 易见，$p_i^X = \sum_j p_{i,j}$，$p_j^Y = \sum_i p_{i,j}$。

5、二维连续型随机变量

定义

- 称二维随机变量(X,Y)是连续型的，如果存在非负可积函数$p:\mathbb{R}^2\to\mathbb{R}$使得

$$F(x,y) = \int_{-\infty}^{x}\int_{-\infty}^{y} p(u,v)dvdu。$$

- 这时称p为(X,Y)的一个联合密度函数（简称联合密度）。
- 易见X的概率密度是 $p_X(x) = \int_{-\infty}^{\infty} p(x,y)dy$，称为$X$的边际密度函数（或边际密度）。

注： 在p的连续点(x,y)处有

$$p(x,y) = \frac{\partial^2 F}{\partial x\partial y} = \lim_{\Delta x,\Delta y\to 0}\frac{P((X,Y)\in(x,x+\Delta x]\times(y,y+\Delta y])}{\Delta x\Delta y}.$$

6、联合分布的基本性质

联合分布列的基本性质：

1. 非负性：$p_{ij}\geq 0$；
2. 正则性：$\sum_{i,j} p_{ij} = 1$。

联合密度函数的基本性质：

1. 非负性：$p(x,y)\geq 0$；
2. 正则性：$\iint_{\mathbb{R}^2} p(x,y)dxdy = 1$。

与X,Y有关事件的概率计算：若$G\subset\mathbb{R}^2$，则

1. 离散型：$P[(X,Y)\in G] = \sum_{(x_i,y_j)\in G} p_{ij}$；
2. 连续型：$P[(X,Y)\in G] = \iint_G p(x,y)dxdy$。

7、多项分布

定义

- 用r种颜色对n个球进行独立随机染色，每个球被染成第i种颜色的概率为p_i（$p_1+\cdots+p_r=1$）。
- 记X_i是第i种颜色的球的个数，$i=1,2,\ldots,r$。
- 则对$n_1+n_2+\cdots+n_r=n$，

$$P(X_1=n_1,X_2=n_2,\ldots,X_r=n_r) = \frac{n!}{n_1!n_2!\cdots n_r!}p_1^{n_1}p_2^{n_2}\cdots p_r^{n_r}$$

- 我们称$(X_1,X_2,\ldots,X_r)$服从多项分布（*polynomial distribution*），记为$M(n,p_1,p_2,\ldots,p_r)$。
- 这是因为上述概率值是$(p_1z_1+\cdots+p_rz_r)^n$的展开式中$z_1^{n_1}\cdots z_r^{n_r}$的系数。

8、多项边际分布

由$(X_1,\ldots,X_n)$的联合分布求X_1的边际分布：

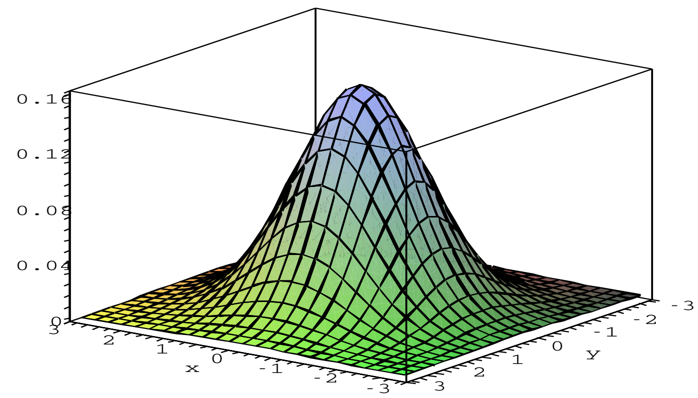
$$
\begin{aligned}
P(X_1=n_1) &= \sum_{n_2+\cdots+n_r=n-n_1} P(X_1=n_1,X_2=n_2,\ldots,X_r=n_r)\\
&= \sum_{n_2+\cdots+n_r=n-n_1} \frac{n!}{n_1!n_2!\cdots n_r!}p_1^{n_1}p_2^{n_2}\cdots p_r^{n_r}\\
&= \frac{n!}{n_1!(n-n_1)!}p_1^{n_1}\sum_{n_2+\cdots+n_r=n-n_1}\frac{(n-n_1)!}{n_2!\cdots n_r!}p_2^{n_2}\cdots p_r^{n_r}\\
&= \frac{n!}{n_1!(n-n_1)!}p_1^{n_1}(p_2+\cdots+p_r)^{n-n_1}\\
&= \frac{n!}{n_1!(n-n_1)!}p_1^{n_1}(1-p_1)^{n-n_1}.
\end{aligned}
$$

因此X_1服从二项分布$b(n,p_1)$。

9、二元标准正态分布

定义

设(X,Y)有联合概率密度函数 $p(x,y)=\frac{1}{2\pi\sqrt{1-\rho^2}}e^{-\frac{x^2-2\rho xy+y^2}{2(1-\rho^2)}}$，其中$|\rho|<1$是常数，我们称$(X,Y)$服从二元标准正态分布，记为$N(0,0,1,1,\rho)$。



10、正态边际分布

$$
\begin{aligned}
p(x,y) &= \frac{1}{2\pi\sqrt{1-\rho^2}}e^{-\frac{x^2-2\rho xy+y^2}{2(1-\rho^2)}} = \frac{1}{2\pi\sqrt{1-\rho^2}}e^{-\frac{(x-\rho y)^2}{2(1-\rho^2)}-y^2/2}\\
&= \frac{1}{\sqrt{2\pi}}e^{-y^2/2}g(x,y),\\
g(x,y) &= \frac{1}{\sqrt{2\pi(1-\rho^2)}}e^{-\frac{(x-\rho y)^2}{2(1-\rho^2)}}; \qquad (\sim N(\rho y,1-\rho^2))\\
p_Y(y) &= \int_{-\infty}^{\infty}p(x,y)dx=\int_{-\infty}^{\infty}\frac{1}{\sqrt{2\pi}}e^{-y^2/2}g(x,y)dx\\
&= \frac{1}{\sqrt{2\pi}}e^{-y^2/2}\int_{-\infty}^{\infty}g(x,y)dx=\frac{1}{\sqrt{2\pi}}e^{-y^2/2}\\
\Rightarrow &\quad Y\sim N(0,1) \quad \text{同理} \quad X\sim N(0,1).
\end{aligned}
$$

11、二元正态分布

定义

- 服从二元正态分布的随机变量(X,Y)的联合密度函数为

$$p(x,y)=\frac{1}{2\pi\sigma_1\sigma_2\sqrt{1-\rho^2}}\exp\left(-\frac{Q(x,y)}{2}\right),$$

- 其中$\sigma_1,\sigma_2>0, |\rho|<1$且

$$Q(x,y)=\frac{1}{1-\rho^2}\left[\left(\frac{x-\mu_1}{\sigma_1}\right)^2-2\rho\left(\frac{x-\mu_1}{\sigma_1}\right)\left(\frac{y-\mu_2}{\sigma_2}\right)+\left(\frac{y-\mu_2}{\sigma_2}\right)^2\right]$$

  是正定二次型（即对任意x,y总有$Q(x,y)\geq 0$，并且$Q(x,y)=0$有唯一解$x=\mu_1, y=\mu_2$）。我们也记这个二元正态分布为$N(\mu_1,\mu_2,\sigma_1^2,\sigma_2^2,\rho)$。

12、正态边际分布

- 作标准化：令$X^* = (X-\mu_1)/\sigma_1$，$Y^* = (Y-\mu_2)/\sigma_2$，
- 则$(X^*, Y^*) \sim N(0,0,1,1,\rho)$。（见第八讲）
- $X^*, Y^* \sim N(0,1)$
- $X \sim N(\mu_1, \sigma_1^2)$，$Y \sim N(\mu_2, \sigma_2^2)$。

边际正态分布$\not\Rightarrow$联合正态分布

- 设(X,Y)的联合分布为$F(x,y) = \min\{\Phi(x), \Phi(y)\}$，
- 则X的边际分布为

$$F_X(x) = F(x,\infty) = \min\{\Phi(x), 1\} = \Phi(x).$$

$\Rightarrow$ $X \sim N(0,1)$　同理　$Y \sim N(0,1)$。

- 但当$x \neq y$，$\frac{\partial^2 F}{\partial x \partial y} = 0$。故$X, Y$不是联合正态分布。

13、独立性

定义

称随机变量X, Y相互独立，如果其联合分布函数是其边际分布函数的乘积，即　$F(x,y) = F_X(x)F_Y(y)$，　$\forall x, y \in \mathbb{R}$。

注：

- 若X, Y是离散型的，则X与Y相互独立当且仅当$\forall x_i, y_j$，$P(X = x_i, Y = y_j) = P(X = x_i)P(Y = y_j)$。
- 若X, Y是连续型的，则X与Y相互独立当且仅当$p(x,y) = p_X(x)p_Y(y)$，　$\forall x, y \in \mathbb{R}$。

定理

若X与Y互相独立，g和h为两实（Borel）函数，则$g(X)$与$h(Y)$也互相独立。

14、独立性

注：

- 两个离散型随机变量X, Y相互独立当且仅当(X,Y)的联合分布中每个位置的概率值恰是所在行和列的两个边际概率值的乘积，
- 因此如果知道两个独立的离散型随机变量各自的分布列，就可以用通常的矩阵计算的方式得到它们的联合分布，

$$\begin{pmatrix} p_1^X \\ p_2^X \\ \vdots \\ p_m^X \end{pmatrix} (p_1^Y, p_2^Y, \dots, p_n^Y) = \begin{pmatrix} p_1^X p_1^Y & p_1^X p_2^Y & \cdots & p_1^X p_n^Y \\ p_2^X p_1^Y & p_2^X p_2^Y & \cdots & p_2^X p_n^Y \\ \vdots & \vdots & \vdots & \vdots \\ p_m^X p_1^Y & p_m^X p_2^Y & \cdots & p_m^X p_n^Y \end{pmatrix}$$

- 这在非独立情形是做不到的。

15、独立变量和的分布（卷积）

例：已知X和Y相互独立，

$P(X=i) = p_i$，$P(Y=j) = q_j$，$i, j \in \mathbb{Z}$，则$Z = X + Y$分布为：

$$P(Z=k) = \sum_{i\in\mathbb{Z}} p_i q_{k-i}.$$

证明：

$$P(Z=k) = P(X+Y=k) = \sum_{i\in\mathbb{Z}} P(X=i, Y=k-i)$$
$$= \sum_{i\in\mathbb{Z}} P(X=i)P(Y=k-i) = \sum_{i\in\mathbb{Z}} p_i q_{k-i}。$$
$$\left(f * g = \int_{-\infty}^{\infty} f(x)g(y-x)\mathrm{d}x \right)$$

16、Poisson分布的可加性

$P(\lambda_1) * P(\lambda_2) = P(\lambda_1 + \lambda_2)$

设相互独立随机变量X, Y服从参数分别为λ_1, λ_2的Poisson分布，则$X + Y \sim P(\lambda_1 + \lambda_2)$。

$$\begin{aligned}
P(X+Y=k) &= \sum_{i=0}^{k} P(X=i)P(Y=k-i) \\
&= \sum_{i=0}^{k} \frac{\lambda_1^i e^{-\lambda_1}}{i!} \frac{\lambda_2^{k-i} e^{-\lambda_2}}{(k-i)!} \\
&= \frac{e^{-(\lambda_1+\lambda_2)}}{k!} \sum_{i=0}^{k} \frac{k!}{i!(k-i)!} \lambda_1^i \lambda_2^{k-i} \\
&= \frac{e^{-(\lambda_1+\lambda_2)}}{k!} (\lambda_1+\lambda_2)^k\text{。}
\end{aligned}$$

17、二项分布的可加性

$b(n,p) * b(m,p) = b(m+n,p)$

若$X \sim b(n,p)$与$Y \sim b(m,p)$独立，则$X + Y \sim b(m+n,p)$。

记$a = \max\{0, k-m\}$，$b = \min\{n, k\}$，则

$$\begin{aligned}
P(X+Y=k) &= \sum_{i=0}^{k} P(X=i)P(Y=k-i) \\
&= \sum_{i=a}^{b} \binom{n}{i} p^i (1-p)^{n-i} \binom{m}{k-i} p^{k-i} (1-p)^{m-(k-i)} \\
&= p^k (1-p)^{m+n-k} \sum_{i=a}^{b} \binom{n}{i} \binom{m}{k-i} \\
&= \binom{m+n}{k} p^k (1-p)^{m+n-k}.
\end{aligned}$$

18、独立变量极值的分布

例

设X和Y相互独立，$M = \max\{X, Y\}$, $N = \min\{X, Y\}$，则$F_M(z) = F_X(z)F_Y(z)$，$F_N(z) = 1 - [1 - F_X(z)][1 - F_Y(z)]$。

证明：

$$\begin{aligned}
F_M(z) &= P(M \le z) = P(X \le z, Y \le z) = P(X \le z)P(Y \le z) \\
&= F_X(z)F_Y(z), \\
F_N(z) &= P(N \le z) = 1 - P(N > z) = 1 - P(X > z, Y > z) \\
&= 1 - P(X > z)P(Y > z) \\
&= 1 - [1 - P(X \le z)][1 - P(Y \le z)] \\
&= 1 - [1 - F_X(z)][1 - F_Y(z)].
\end{aligned}$$

19、和的分布

定理

设(X, Y)的联合概率密度为$p(x, y)$，则$Z = X + Y$的概率密度为$p_Z(z) = \int_{-\infty}^{\infty} p(z-y, y)dy$。

证明：

$$\begin{aligned}
F_Z(z) &= P(Z \le z) = P(X+Y \le z) = \iint_{x+y \le z} p(x,y)dxdy \\
&= \int_{-\infty}^{\infty} \int_{-\infty}^{z-y} p(x,y)dxdy = \int_{-\infty}^{\infty} \int_{-\infty}^{z} p(t-y,y)dtdy \\
&= \int_{-\infty}^{z} \int_{-\infty}^{\infty} p(t-y,y)dydt; \\
p_Z(z) &= F_Z' = \int_{-\infty}^{\infty} p(z-y,y)dy.
\end{aligned}$$

20、独立变量和的分布（卷积）

推论

若X, Y独立，则

$$p_{X+Y}(z)=\int_{-\infty}^{\infty} p_X(z-y)p_Y(y)dy=\int_{-\infty}^{\infty} p_X(x)p_Y(z-x)dx。$$

显然，p_{X+Y}为p_X和p_Y的卷积，记为$p_{X+Y}=p_X*p_Y=p_Y*p_X$.

例：X和Y是两独立的标准正态随机变量,求$Z=X+Y$的分布.

$$\begin{aligned}
p_Z(z) &= \int_{-\infty}^{\infty} p_X(x)p_Y(z-x)dx=\frac{1}{2\pi}\int_{-\infty}^{\infty} e^{-x^2/2}e^{-(z-x)^2/2}dx\\
&= \frac{1}{2\pi}e^{-z^2/4}\int_{-\infty}^{\infty} e^{-(x-z/2)^2}dx \overset{t=x-\frac{z}{2}}{=} \frac{1}{2\pi}e^{-\frac{z^2}{4}}\int_{-\infty}^{\infty} e^{-t^2}dt\\
&= \frac{1}{2\sqrt{\pi}}e^{-z^2/4}\int_{-\infty}^{\infty}\frac{1}{\sqrt{2\pi}}e^{-t^2}d\left(\sqrt{2}t\right)=\frac{1}{2\sqrt{\pi}}e^{-z^2/4}\\
\Rightarrow\quad & Z\sim N(0,2).
\end{aligned}$$

21、Γ分布的可加性

$\Gamma(\alpha_1,\lambda)*\Gamma(\alpha_2,\lambda)=\Gamma(\alpha_1+\alpha_2,\lambda)$

设$X\sim\Gamma(\alpha_1,\lambda), Y\sim\Gamma(\alpha_2,\lambda)$独立，则$X+Y\sim\Gamma(\alpha_1+\alpha_2,\lambda)$。

$$\begin{aligned}
p_{X+Y}(z) &= \frac{\lambda^{\alpha_1+\alpha_2}}{\Gamma(\alpha_1)\Gamma(\alpha_2)}\int_0^z (z-y)^{\alpha_1-1}e^{-\lambda(z-y)}y^{\alpha_2-1}e^{-\lambda y}dy\\
&= \frac{\lambda^{\alpha_1+\alpha_2}e^{-\lambda z}}{\Gamma(\alpha_1)\Gamma(\alpha_2)}\int_0^z (z-y)^{\alpha_1-1}y^{\alpha_2-1}dy\\
&\overset{y=zt}{=} \frac{\lambda^{\alpha_1+\alpha_2}e^{-\lambda z}}{\Gamma(\alpha_1)\Gamma(\alpha_2)}z^{\alpha_1+\alpha_2-1}\int_0^1 (1-t)^{\alpha_1-1}t^{\alpha_2-1}dt\\
&= \frac{\lambda^{\alpha_1+\alpha_2}}{\Gamma(\alpha_1+\alpha_2)}z^{\alpha_1+\alpha_2-1}e^{-\lambda z}.
\end{aligned}$$

1. $Exp(\lambda)*Exp(\lambda)=\Gamma(2,\lambda)$。
2. $\chi^2(m)*\chi^2(n)=\chi^2(m+n)$。
3. 设$X_i \overset{i.i.d.}{\sim} N(0,1), i=1,...,n$，则$\sum_{i=1}^n X_i^2\sim\chi^2(n)$。

22、次序统计量的分布

定理

设随机变量$X_1, X_2, \ldots, X_n$独立同分布，分布函数为F，密度为p.
一般地，我们将$X_1, X_2, \ldots, X_n$从小到大排序后得到

$$X_{(1)}\le X_{(2)}\le\cdots\le X_{(n)}。$$

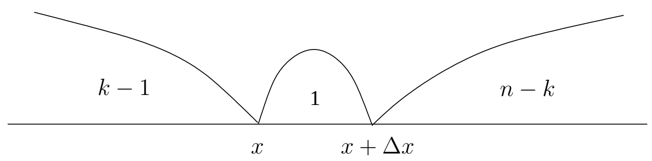
1. 严格地说，第k个次序统计量$X_{(k)}$满足
$$X_{(k)}=\min_{1\le i_1<i_2<\cdots<i_k\le n}\max\{X_{i_1}, X_{i_2}, \ldots, X_{i_k}\}.$$
且密度函数为$p_k(x)=\frac{n!}{(k-1)!(n-k)!}F(x)^{k-1}[1-F(x)]^{n-k}p(x)$.
2. 次序统计量$(X_{(i)}, X_{(j)})(i<j)$的联合密度函数为：当$x\le y$，
$$p_{ij}(x,y)=\frac{n!F(x)^{i-1}[F(y)-F(x)]^{j-i-1}[1-F(y)]^{n-j}p(x)p(y)}{(i-1)!(j-i-1)!(n-j)!}。$$

23、次序统计量的分布

证明（1）：注意到

- $p_k(x)=F_k'(x)=\lim_{\Delta x\to 0}\frac{1}{\Delta x}[F_k(x+\Delta x)-F_k(x)]$，
- $F_k(x+\Delta x)-F_k(x)=P\left(X_{(k)}\in(x,x+\Delta x]\right)$，

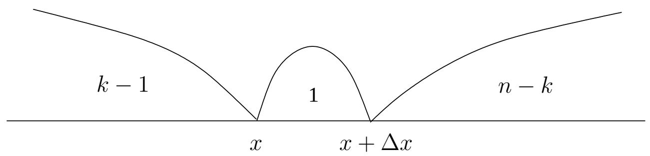
故　我们需要考虑事件$X_{(k)}\in(x,x+\Delta x]$，



$$\begin{aligned}
&F_k(x+\Delta x)-F_k(x)=P\left(X_{(k)}\in(x,x+\Delta x]\right)\\
&\approx\frac{n!F(x)^{k-1}[F(x+\Delta x)-F(x)][1-F(x+\Delta x)]^{n-k}}{(k-1)!(n-k)!},
\end{aligned}$$

24、次序统计量的分布

如图考虑事件$X_{(k)} \in (x, x+\Delta x]$：



$$F_k(x+\Delta x) - F_k(x) = P\left(X_{(k)} \in (x, x+\Delta x]\right)$$

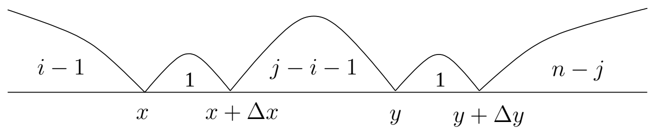
$$\approx \frac{n! F(x)^{k-1}[F(x+\Delta x) - F(x)][1 - F(x+\Delta x)]^{n-k}}{(k-1)!(n-k)!},$$

$$p_k(x) = \lim_{\Delta x \to 0} \frac{1}{\Delta x}[F_k(x+\Delta x) - F_k(x)]$$

$$= \frac{n!}{(k-1)!(n-k)!} F(x)^{k-1}[1-F(x)]^{n-k} p(x)。$$

25、次序统计量的分布

证明(2)：如图考虑事件$(X_{(i)}, X_{(j)}) \in (x, x+\Delta x] \times (y, y+\Delta y]$：



$$P\left((X_{(i)}, X_{(j)}) \in (x, x+\Delta x] \times (y, y+\Delta y]\right)$$

$$\approx \frac{n!}{(i-1)!(j-i-1)!(n-j)!} F(x)^{i-1} p(x) \Delta x \times [F(y) - F(x+\Delta x)]^{j-i-1} p(y) \Delta y [1 - F(y+\Delta y)]^{n-j},$$

$$p_{ij}(x,y) = \lim_{\Delta x, \Delta y \to 0} \frac{1}{\Delta x \Delta y} P((X_{(i)}, X_{(j)}) \in (x, x+\Delta x] \times (y, y+\Delta y])$$

$$= \frac{n! F(x)^{i-1}[F(y) - F(x)]^{j-i-1}[1-F(y)]^{n-j} p(x) p(y)}{(i-1)!(j-i-1)!(n-j)!}.$$