多维随机变量及其分布 01

1、多维随机变量与联合分布函数

（一）多维随机变量

如果$X_1(\omega), X_2(\omega), \dots, X_n(\omega)$是定义在同一样本空间$\Omega=\{\omega\}$上的$n$个随机变量，则称$X(\omega)=(X_1(\omega), X_2(\omega), \dots, X_n(\omega))$为$n$维(或$n$元)随机变量或随机向量，例如：在研究四岁至六岁儿童的生长发育情况时，我们感兴趣于每个儿童(样本点ω)的身高$X_1(\omega)$和体重$X_2(\omega)$，这里X_2，X_2是一个二维随机变量。

（二）联合分布函数

对任意的 n 个实数$x_1, x_2, \dots, x_n$，则n个事件$\{X_1 \le x_1, X_2 \le x_2, \dots, X_n \le x_n\}$同时发生的概率：

$$F(x_1, x_2, \dots, x_n) = P(X_1 \le x_1, X_2 \le x_2, \dots, X_n \le x_n)$$

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

一般情况下，主要研究二维联合分布函数$F(x,y)=P(X\le x, Y\le y)$

2、多维随机变量与联合分布函数的定理

定理 1：任一二维联合分布函数$F(x,y)$必具有如下四条基本性质：

① 单调性　$F(x,y)$分别关于x和y单调非减的

② 有界性　对于任一的 x 或 y，有$0\le F(x,y)\le 1$，且$F(-\infty,y)=\lim\limits_{x\to-\infty}F(x,y)=0$　$F(x,-\infty)=\lim\limits_{y\to-\infty}F(x,y)=0$　$F(\infty,\infty)=\lim\limits_{x,y\to-\infty}F(x,y)=1$

③ 右连续性　$F(x,y)$分别关于x和y右连续

④ 非负性　对任意$a<b, c<d$有$P(a<X\le b, c<Y\le d)=$

$$F(b,d)-F(a,d)-F(b,c)+F(a,c)\geq 0$$

3、联合分布列

如果二维随机变量(X,Y)只有有限个或可列个数对(x_i,y_i)，则称(X,Y)为二维离散型随机变量，称$p_{ij}=P(X=x_i,Y=y_i),i,j=1,2,\ldots$为$(X,Y)$的联合分布列，用表格形式表示为：

$Y \backslash X$	x_1	$\cdots$	x_i	$\cdots$	$p_{\cdot j}$
y_1	p_{11}	$\cdots$	p_{i1}	$\cdots$	$p_{\cdot 1}$
$\vdots$	$\vdots$	$\cdots$	$\vdots$	$\cdots$	$\vdots$
y_j	p_{1j}	$\cdots$	p_{ij}	$\cdots$	$p_{\cdot j}$
$\vdots$	$\vdots$	$\cdots$	$\vdots$	$\cdots$	$\vdots$
$p_{i\cdot}$	$p_{1\cdot}$	$\cdots$	p_i	$\cdots$	1

联合分布列的基本性质：
(1) 非负性$p_{ij}\geq 0$
(2) (2)正则性$\sum_{i=1}^{\infty}\sum_{j=1}^{\infty}p_{ij}=1$

确定联合分布列的方法：
(1) 确定随机变量(X,Y)的所有取值数对
(2) 计算取每个数值对的概率
(3) 列出表格

4、联合密度函数

如果存在二元非负函数$p(x,y)$，使得二维随机变量(X,Y)的分布函数$F(x,y)$可表示为

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

则称(X,Y)为二位连续随机变量，称$p(u,v)$为(X,Y)的联合密度函数。

在$F(x,y)$偏导数存在的点有$p(x,y)=\frac{\partial^2}{\partial x\partial y}F(x,y)$

联合密度函数的基本性质：

（1）非负性　$p(x,y)\geq 0$

（2）正则性　$\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}p(x,y)dydx=1$

若G是平面上的一个区域，则事件$\{(X,Y)\in G\}$的概率可表示为G上对$p(x,y)$的二重积分$P\big((X,Y)\in G\big)=\iint p(x,y)dxdy$

5、常用的多维分布

（一）多项分布

进行n次独立重复试验，如果每次实验有r个互不相容结果，$A_1, A_2, \ldots A_r$之一发生，且每次试验中A_i发生的概率为$p_i = P(A_i), i = 1,2, \ldots, r$，且$p_1 + p_2 + \cdots + p_r = 1$. 记$X_i$为 n 次独立重复实验中$A_i$出现的次数，$i = 1,2, \ldots, r$，则$(X_1, X_2, \ldots X_r)$取值$(n_1, n_2, \ldots n_r)$的概率，即$A_1$出现$n_1$，$A_2$出现$n_2$次，…，$A_r$出现$n_r$次的概率为

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

其中，$n = n_1 + n_2 + \cdots + n_r$，这个联合分布列称为$r$项分布，又称多项分布，记为$M(n, p_1, p_2, \ldots, p_r)$

例题：一批产品中共有 100 件，其中一等品 60 件，二等品 30 件，三等品 10 件，从这批产品中有放回的任取 3 件，以X和Y分别表示取出 3 件产品中一等品，二等品的件数，求二维随机变量(X, Y)的联合分布列。

解：因为X和Y的可能取值都是 0，1，2，3，所以记(X, Y)的联合分布列为$P(X = i, Y = j) = p_{ij}, i, j = 0,1,2,3$

当$i + j > 3$时，有$p_{ij} = 0$；当$i + j \le 3$时，有$p_{ij} = \frac{3!}{i! j! (3-i-j)!} (\frac{6}{10})^i (\frac{3}{10})^j (\frac{1}{10})^{3-i-j}$

（二）多维超几何分布

袋中有N个球，其中有N_i个i号球，$i = 1,2, \ldots, r$，且$N = N_1 + N_2 + \cdots + N_r$，从中任意取出$n$个，若记$X_i$为取出的$n$个球中$i$号球的个数，$i = 1,2, \ldots, r$，则$P(X_1 = n_1, X_2 = n_2, \ldots, X_r = n_r) = \frac{C_{N_1}^{n_1} C_{N_2}^{n_2} \ldots C_{N_r}^{n_r}}{C_N^n}$，其中$n = n_1 + n_2 + \cdots + n_r$

例题：一批产品中共有 100 件，其中一等品 60 件，二等品 30 件，三等品 10 件，从这批产品中无放回的任取 3 件，以X和Y分别表示取出 3 件产品中一等品，二等品的件数，求二维随机变量(X, Y)的联合分布列。

解：因为X和Y的可能取值都是 0，1，2，3，所以记(X, Y)的联合分布列为$P(X = i, Y = j) = p_{ij}, i, j = 0,1,2,3$

当$i + j > 3$时，有$p_{ij} = 0$；当$i + j \le 3$时，有$p_{ij} = \frac{C_{60}^i C_{30}^j C_{10}^{3-i-j}}{C_{100}^3}$

(三) 多维均匀分布

设 D 为R^n中的一个有界区域，其度量(平面的为面积，空间的为体积等) 为S_D如果多维随机变量$(X_1,X_2,\dots,X_n)$的联合密度函数为

$$P(x_1,x_2,\dots x_n)=\begin{cases}\dfrac{1}{S_D},(x_1,x_2,\dots,x_n\in D)\\ 0,\text{其他}\end{cases}$$

则称$(X_1,X_2,\dots,X_n)$服从 D 上的多维均匀分布。记$(X_1,X_2,\dots,X_n)\sim U(D)$