第三章 联合分布

§2 二维离散随机变量

二维随机变量的基本分类 $\begin{cases} \text{二维离散型 r.v} \\ \text{二维连续型 r.v} \end{cases}$

二维离散型随机变量

设r.v (X,Y)的所有可能的取值为

$$(x_i, y_j) \quad (i, j = 1, 2, \cdots)$$

取值的概率为

$$P\{X = x_i, Y = y_j\} = p(x_i, y_j) \triangleq p_{ij} \quad (i, j = 1, 2, \cdots)$$

称上式为二维离散型r.v (X,Y)的**频率函数**，或称为r.v X,Y的**联合频率函数(joint frequency function)**.

频率函数的基本性质

设r.v (X,Y)的频率函数为

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

则

① $p_{ij} \ge 0 \qquad (i,j=1,2,\cdots)$

② $\sum_{i=1}^{\infty}\sum_{j=1}^{\infty} p_{ij}=1$

离散型r.v频率函数的本质特征

频率函数的表格表示法

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

例 袋中装有2只白球及3只黑球，现进行无放回的摸球，定义随机变量如下：

$$X=\begin{cases}1, & \text{第一次摸出白球}\\ 0, & \text{第一次摸出黑球}\end{cases}\qquad Y=\begin{cases}1, & \text{第二次摸出白球}\\ 0, & \text{第二次摸出黑球}\end{cases}$$

求(X,Y)的频率函数？

解

$$P\{X=0,Y=0\}=P\{Y=0\mid X=0\}\cdot P\{X=0\}=(2/4)\cdot(3/5)$$

$$P\{X=0,Y=1\}=P\{Y=1\mid X=0\}\cdot P\{X=0\}=(2/4)\cdot(3/5)$$

$$P\{X=1,Y=0\}=P\{Y=0\mid X=1\}\cdot P\{X=1\}=(3/4)\cdot(2/5)$$

$$P\{X=1,Y=1\}=P\{Y=1\mid X=1\}\cdot P\{X=1\}=(1/4)\cdot(2/5)$$

例 有一个射击游戏, 参加游戏的人先掷一次骰子, 若出现点数为 X, 则射击 X次. 设某人击中目标概率为 $p=0.9$, 记击中目标的次数为 Y. 求(X,Y)的频率函数.

解 X的取值为$1,2,\cdots,6$,Y的取值为$0,1,2,\cdots,X$.

当 $X=i$时 ,$Y\sim b(i,p)$ $(i=1,2,3,4,5,6)$

由乘法公式求得

$$P\{X=i,Y=j\}=P\{Y=j\mid X=i\}\cdot P\{X=i\}$$

$$=\begin{cases}\dfrac{1}{6}C_i^j p^j(1-p)^{i-j}, 0\le j\le i, i=1,2,\cdots,6\\ 0\quad , \qquad\qquad \text{其它}\end{cases}$$

代入 $p=0.9$，求得 (X,Y) 的频率函数为

Y \ X	1	2	3	4	5	6
0	0.017	0.0017	0.00017	0.000017	0.0000017	0.00000017
1	0.15	0.03	0.0045	0.0006	0.000075	0.000009
2	0	0.14	0.0405	0.0081	0.00135	0.000203
3	0	0	0.1215	0.0486	0.01215	0.002430
4	0	0	0	0.1094	0.05468	0.016403
5	0	0	0	0	0.09842	0.059049
6	0	0	0	0	0	0.088573

问题 question

如果不掷骰子，直接射击一次，则

$$P\{Y=0\}=0.1,\quad P\{Y=1\}=0.9$$

为什么概率不一样？

二维离散型随机变量的边际频率函数

设 (X,Y) 的频率函数为

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

则 r.v X 的频率函数是

$$P\{X=x_i\}=P(\{X=x_i\}\cap\Omega)\triangleq p_{i\cdot} \qquad (i=1,2,\cdots)$$

$$=P(\{X=x_i\}\cap(\bigcup_{j=1}^{\infty}\{Y=y_j\}))$$

$$=P(\bigcup_{j=1}^{\infty}(\{X=x_i\}\cap\{Y=y_j\}))$$

$$=P(\bigcup_{j=1}^{\infty}\{X=x_i, Y=y_j\})$$

$$=\sum_{j=1}^{\infty}P\{X=x_i, Y=y_j\}$$

$$=\sum_{j=1}^{\infty}p_{ij}$$

二维离散型随机变量的边际频率函数

设(X,Y)的频率函数为

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

则r.v X的频率函数是

$$P\{X=x_i\}=\sum_{j=1}^{\infty} p_{ij} \triangleq p_{i\cdot} \qquad (i=1,2,\cdots)$$

同理 Y 的频率函数是

$$P\{Y=y_j\}=\sum_{i=1}^{\infty} p_{ij} \triangleq p_{\cdot j} \qquad (j=1,2,\cdots)$$

定义 称数列$\{p_{i\cdot}\}$为(X,Y)关于X的边际频率函数

称数列$\{p_{\cdot j}\}$为(X,Y)关于Y的边际频率函数

(marginal frequency function).

①它是一维r.v的频率函数

②它可通过二维r.v的频率函数计算得到

例 设 r.v X 从 1, 2, 3, 4中等可能取值, 又设 r.v Y从 $1\sim X$ 中等可能取值. 求X,Y的联合频率函数及边际频率函数.

解 X 取值为1, 2, 3, 4 ,而当 $X=i\ (i=1,2,3,4)$ 时, Y 的取值为$1\sim i$.由乘法公式有

$$P\{X=i,Y=j\}=P\{Y=j\mid X=i\}\cdot P\{X=i\}=\frac{1}{i}\cdot\frac{1}{4}\quad(1\le j\le i)$$

故 X 的联合频率函数为

$Y \backslash X$	1	2	3	4	$p_{\cdot j}=\sum_{i=1}^{4}p_{ij}$
1	1/4	1/8	1/12	1/16	25/48
2	0	1/8	1/12	1/16	13/48
3	0	0	1/12	1/16	7/48
4	0	0	0	1/16	3/48
$p_{i\cdot}=\sum_{j=1}^{4}p_{ij}$	$\frac{1}{4}$	$\frac{1}{4}$	$\frac{1}{4}$	$\frac{1}{4}$	

故边际频率函数为

X	1	2	3	4
$p_{i\cdot}$	1/4	1/4	1/4	1/4

Y	1	2	3	4
$p_{\cdot j}$	25/48	13/48	7/48	3/48

n维离散型随机变量的边际频率函数

设$X_1,\ldots,X_n$的联合频率函数为

$$P\{X_1=x_1,\ldots,X_n=x_n\}=p(x_1,\ldots,x_n)$$

则r.v X_1的边际频率函数是

$$p_{X_1}(x_1)=\sum_{x_2\ldots x_n}p(x_1,x_2,\ldots,x_n)$$

r.v X_1和 X_2的二维边际频率函数是

$$p_{X_1X_2}(x_1,x_2)=\sum_{x_3\ldots x_n}p(x_1,x_2,\ldots,x_n)$$

例 **多项(multinomial)分布**：二项分布的推广

假设进行 n 次独立试验，每次试验有 r 种可能的结果，各自出现的概率分别为 $p_1, p_2, \ldots, p_r$.

令N_i 是n次试验出现第i 种试验结果的所有次数，其中 $i = 1, \ldots, r$.

$N_1, N_2, \ldots, N_r$ 的联合频率函数是

$$p(n_1, \ldots, n_r) = \binom{n}{n_1 \cdots n_r} p_1^{n_1} p_2^{n_2} \cdots p_r^{n_r}$$

N_i 的边际频率函数的计算 【两种理解】：

① 联合频率函数关于其它的n_j 求和;

② N_i 可解释为n次试验中<u>成功的次数</u>，故 $N_i \sim \mathrm{b}(n, p_i)$.

$$p_{N_i}(n_i) = \binom{n}{n_i} p_i^{n_i} (1 - p_i)^{n - n_i}$$

§2 二维离散随机变量

例 箱子里装有4只白球和2只黑球，在其中随机地取两次，每次取一只。考虑两种试验：

（1）有放回抽样，（2）不放回抽样。

我们定义随机变量 X，Y 如下，写出X和Y的联合分布律和边缘分布律 。

$$X=\begin{cases}0,\text{若第一次取出的是黑球},\\1,\text{若第一次取出的是白球}.\end{cases}$$

$$Y=\begin{cases}0,\text{若第二次取出的是黑球},\\1,\text{若第二次取出的是白球}.\end{cases}$$

（1）有放回抽样

$X \backslash Y$	0	1	$p_{i\bullet}$
0	$\frac{1}{9}$	$\frac{2}{9}$	$\frac{1}{3}$
1	$\frac{2}{9}$	$\frac{4}{9}$	$\frac{2}{3}$
$p_{\bullet j}$	$\frac{1}{3}$	$\frac{2}{3}$	1

（2）不放回抽样

$X \backslash Y$	0	1	$p_{i\bullet}$
0	$\frac{1}{15}$	$\frac{4}{15}$	$\frac{1}{3}$
1	$\frac{4}{15}$	$\frac{6}{15}$	$\frac{2}{3}$
$p_{\bullet j}$	$\frac{1}{3}$	$\frac{2}{3}$	1

例5：袋中有1个红球，2个黑球，3个白球，现有放回地取两次，每次取一球，以X,Y,Z分别表示两次取球所得的红、黑、白球个数。求：
(1)$P(X=1|Z=0)$ (2) $P(X=1,Z=0)$(3)(X,Y)分布.

例5：袋中有1个红球，2个黑球，3个白球，现有放回地取两次，每次取一球，以X,Y,Z分别表示两次取球所得的红、黑、白球个数。求：
(1)$P(X=1|Z=0)$ (2) $P(X=1,Z=0)$(3)(X,Y)分布.

解：(1) $P(X=1\,|\,Z=0)=\underset{\text{红}}{\frac{1}{3}}\times\underset{\text{黑}}{\frac{2}{3}}+\underset{\text{黑}}{\frac{2}{3}}\times\underset{\text{红}}{\frac{1}{3}}=\frac{4}{9}$

例5：袋中有1个红球，2个黑球，3个白球，现有放回地取两次，每次取一球，以X,Y,Z分别表示两次取球所得的红、黑、白球个数。求：
(1)$P(X=1|Z=0)$ (2) $P(X=1,Z=0)$(3)(X,Y)分布.

解：(1) $P(X=1|Z=0)=\underset{红}{\frac{1}{3}}\times\underset{黑}{\frac{2}{3}}+\underset{黑}{\frac{2}{3}}\times\underset{红}{\frac{1}{3}}=\frac{4}{9}$

(2) $P(X=1,Z=0)=\underset{红}{\frac{1}{6}}\times\underset{黑}{\frac{2}{6}}+\underset{黑}{\frac{2}{6}}\times\underset{红}{\frac{1}{6}}=\frac{1}{9}$

注意两者的区别!

例5：袋中有1个红球，2个黑球，3个白球，现有放回地取两次，每次取一球，以X,Y,Z分别表示两次取球所得的红、黑、白球个数。求：(1)$P(X=1|Z=0)$ (2) $P(X=1,Z=0)$(3)(X,Y)分布.

解：(3) X,Y的取值范围均为0,1,2.

例5：袋中有1个红球，2个黑球，3个白球，现有放回地取两次，每次取一球，以X,Y,Z分别表示两次取球所得的红、黑、白球个数。求：(1)$P(X=1|Z=0)$ (2) $P(X=1,Z=0)$(3)(X,Y)分布.

解：(3) X,Y的取值范围均为0,1,2.

$$P(X=0,Y=0)=\frac{3}{6}\times\frac{3}{6}=\frac{1}{4}$$ 2球均为白球

$$P(X=0,Y=1)=\frac{2}{6}\times\frac{3}{6}\times 2=\frac{1}{3}$$ 黑白或白黑

$$P(X=1,Y=2)=0$$ 总数超2只，不可能！

$$P(X=2,Y=0)=\frac{1}{6}\times\frac{1}{6}=\frac{1}{36}$$ 2球均为红球

其余类似得到！

课后作业 P76：3，补充题

1、把一枚均匀硬币抛掷三次，设 X 为三次抛掷中正面出现的次数，而 Y 为正面出现次数与反面出现次数之差的绝对值，求 (X, Y) 的频率函数 .

2、设 X 的分布为 P(X = -1)= P(X=0)=P(X=1)=1/3. 令 $Y=X^2$， 求(X,Y)的联合频率函数及边缘频率函数。

3、设随机变量 Y 服从参数为 1 的指数分布，随机变量

$$X_k = \begin{cases} 0, & 若 Y \le k, \\ 1, & 若 Y > k, \end{cases} \quad k = 1,2$$

求二维随机变量(X_1,X_2)的联合频率函数及边缘频率函数。