

第三章 联合分布

- § 1 引言：联合累积分布函数
- § 2 (二维)离散随机变量
- § 3 (二维)连续随机变量
- § 4 独立随机变量
- § 5 条件分布
- § 6 联合分布随机变量函数
- § 7 极值和顺序统计量

二维随机变量的基本分类 { 二维离散型 r.v
二维连续型 r.v

二维离散型随机变量

设 r.v (X, Y) 的所有可能的取值为

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

取值的概率为

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

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)**.

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：

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

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

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

由乘法公式求得

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

$$= \begin{cases} \frac{1}{6} C_i^j p^j (1-p)^{i-j}, & 0 \leq j \leq i, i = 1, 2, \dots, 6 \\ 0, & \text{其它} \end{cases}$$

代入 $p = 0.9$, 求得 (X, Y) 的频率函数为

$Y \backslash X$	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

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?

问题
question

频率函数的基本性质

设 r.v (X, Y) 的频率函数为

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

则

$$① \quad p_{ij} \geq 0 \quad (i, j = 1, 2, \dots)$$

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

离散型 r.v 频率函数的本质特征

频率函数的表格表示法

$Y \backslash X$	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

二维 r.v 的整体概率特性: $(X, Y) \sim F(x, y)$

两个一维 r.v 的概率特性: $X \sim F_X(x), Y \sim F_Y(y)$

定义 称 $F_X(x)$ 为 (X, Y) 关于 X 的**边际分布(函数)**
称 $F_Y(y)$ 为 (X, Y) 关于 Y 的**边际分布(函数)**

 $F(x, y), F_X(x), F_Y(y)$ 之间有什么关系?

分析

$$F_X(x) = P\{X \leq x\}$$

$$= P\{X \leq x, Y < +\infty\}$$

$$= F(x, +\infty)$$

$$F_Y(y) = P\{Y \leq y\}$$

$$= P\{X < +\infty, Y \leq y\}$$

$$= F(+\infty, y)$$

结论

随机变量的边际分布完全由它们的联合分布确定

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

$$\begin{aligned} P\{X = x_i\} &= P(\{X = x_i\} \cap \Omega) \\ &= P(\{X = x_i\} \cap (\bigcup_{j=1}^{\infty} \{Y = y_j\})) \\ &= P(\bigcup_{j=1}^{\infty} (\{X = x_i\} \cap \{Y = y_j\})) \\ &= P(\bigcup_{j=1}^{\infty} \{X = x_i, Y = y_j\}) \\ &= \sum_{j=1}^{\infty} P\{X = x_i, Y = y_j\} \\ &= \sum_{j=1}^{\infty} p_{ij} \triangleq p_{i\cdot} \quad (i = 1, 2, \cdots) \end{aligned}$$

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

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

则 r.v X 的频率函数是

$$P\{X = x_i\} = \sum_{j=1}^{\infty} p_{ij} \triangleq p_{i.} \quad (i = 1, 2, \dots)$$

同理 Y 的频率函数是

$$P\{Y = y_j\} = \sum_{i=1}^{\infty} p_{ij} \triangleq p_{.j} \quad (j = 1, 2, \dots)$$

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

称数列 $\{p_{.j}\}$ 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

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

故 X, Y 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
$p_{i \cdot} = \sum_{j=1}^4 p_{ij}$	1/4	1/4	1/4	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, \dots, X_n 的联合频率函数为

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

则 r.v X_1 的边际频率函数是

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

r.v X_1 和 X_2 的二维边际频率函数是

$$p_{X_1 X_2}(x_1, x_2) = \sum_{x_3 \dots x_n} p(x_1, x_2, \dots, x_n)$$

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, \dots, 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 $i = 1, \dots, r$.

N_1, N_2, \dots, N_r 的联合频率函数是

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

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 成功的次数, 故 $N_i \sim b(n, p_i)$.

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



课后作业

P76: 3, 补充题

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

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

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。