概率论与数理统计

第十五讲

边缘分布

(X, Y) 作为整体有联合分布 $F(x, y)$.

- **问题** X, Y 作为单个的随机变量, 其分布与联合分布有什么关系?

1. 边缘分布函数

$$(X, Y) \sim F(x, y) = P\{X \le x, Y \le y\}$$

若令 $y \to +\infty$, 则 { Y 取一切值 }是必然事件. 于是

$$\lim_{y \to +\infty} F(x, y) = P\{X \le x\} \stackrel{\Delta}{=} F_X(x)$$

正是 X 的分布函数, 称为 X 的**边缘**(marginal)分布函数.

同理, $\lim\limits_{x \to +\infty} F(x, y) = F_Y(y)$ 是 Y 的边缘分布函数.

2. 分布律

设(X, Y)为离散型二维随机变量, 其分布律为

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

对于固定的x_i, 考虑和式

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

$$= P[\{X = x_i\} \cap \underbrace{\bigcup_{j=1}^{+\infty}\{Y = y_j\}}_{\text{必然事件}}] = P\{X = x_i\} \overset{\Delta}{=} p_{i\bullet}$$

这是 X 的分布律.

同理 $\sum\limits_{i=1}^{+\infty} p_{ij} = P\{Y = y_j\} = p_{\cdot j} \quad j = 1, 2, \cdots$ 是 Y 的分布律.

将 X, Y 的联合分布律写成列表形式:

$Y \backslash X$	x_1	x_2	$\cdots$	x_i	$\cdots$	
y_1	p_{11}	p_{21}	$\cdots$	p_{i1}	$\cdots$	$p_{.1}$
y_2	p_{21}	p_{22}	$\cdots$	p_{2n}	$\cdots$	$p_{.2}$
$\vdots$	$\vdots$	$\vdots$		$\vdots$		$\vdots$
y_j	p_{1j}	p_{2j}	$\cdots$	p_{ij}	$\cdots$	$p_{.j}$
$\vdots$	$\vdots$	$\vdots$		$\vdots$		$\vdots$
	$p_{1.}$	$p_{2.}$	$\cdots$	$p_{i.}$	$\cdots$	1

$p_{i.}$, $p_{.j}$分别称为是 X, Y 的**边缘分布律**.

“边缘”由此得名.

例（掷双骰子）

$Y\backslash X$	2	3	4	5	6	7	8	9	10	11	12	
0	$\frac{1}{36}$	0	$\frac{1}{36}$	0	$\frac{1}{36}$	0	$\frac{1}{36}$	0	$\frac{1}{36}$	0	$\frac{1}{36}$	$\frac{6}{36}$
1	0	$\frac{2}{36}$	0	$\frac{2}{36}$	0	$\frac{2}{36}$	0	$\frac{2}{36}$	0	$\frac{2}{36}$	0	$\frac{10}{36}$
2	0	0	$\frac{2}{36}$	0	$\frac{2}{36}$	0	$\frac{2}{36}$	0	$\frac{2}{36}$	0	0	$\frac{8}{36}$
3	0	0	0	$\frac{2}{36}$	0	$\frac{2}{36}$	0	$\frac{2}{36}$	0	0	0	$\frac{6}{36}$
4	0	0	0	0	$\frac{2}{36}$	0	$\frac{2}{36}$	0	0	0	0	$\frac{4}{36}$
5	0	0	0	0	0	$\frac{2}{36}$	0	0	0	0	0	$\frac{2}{36}$
	$\frac{1}{36}$	$\frac{2}{36}$	$\frac{3}{36}$	$\frac{4}{36}$	$\frac{5}{36}$	$\frac{6}{36}$	$\frac{5}{36}$	$\frac{4}{36}$	$\frac{3}{36}$	$\frac{2}{36}$	$\frac{1}{36}$	

有时联合分布律是以乘法公式来定义的.

例（电游竞赛）

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

则 $P_{i\cdot}=P\{X=x_i\}$，并由全概率公式

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

即所有$P\{Y=y_j \mid X=x_i\}$关于 i 的加权平均.

$Y \backslash X$	0	1	2	3	4	5	
0	$\frac{1}{6\times 6}$	0	0	0	0	0	$\frac{1}{6\times 6}$
1	$\frac{1}{6\times 6}$	$\frac{1}{6\times 5}$	0	0	0	0	$\frac{1}{6}\left(\frac{1}{5}+\frac{1}{6}\right)$
2	$\frac{1}{6\times 6}$	$\frac{1}{6\times 5}$	$\frac{1}{6\times 4}$	0	0	0	$\frac{1}{6}\left(\frac{1}{4}+\frac{1}{5}+\frac{1}{6}\right)$
3	$\frac{1}{6\times 6}$	$\frac{1}{6\times 5}$	$\frac{1}{6\times 4}$	$\frac{1}{6\times 3}$	0	0	$\frac{1}{6}\left(\frac{1}{3}+\frac{1}{4}+\frac{1}{5}+\frac{1}{6}\right)$
4	$\frac{1}{6\times 6}$	$\frac{1}{6\times 5}$	$\frac{1}{6\times 4}$	$\frac{1}{6\times 3}$	$\frac{1}{6\times 2}$	0	$\frac{1}{6}\left(\frac{1}{2}+\frac{1}{3}+\frac{1}{4}+\frac{1}{5}+\frac{1}{6}\right)$
5	$\frac{1}{6\times 6}$	$\frac{1}{6\times 5}$	$\frac{1}{6\times 4}$	$\frac{1}{6\times 3}$	$\frac{1}{6\times 2}$	$\frac{1}{6\times 1}$	$\frac{1}{6}\left(1+\frac{1}{2}+\frac{1}{3}+\frac{1}{4}+\frac{1}{5}+\frac{1}{6}\right)$
	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	

例（昆虫产卵）设某种昆虫产卵数 $X\sim P(\lambda)$, 设卵的孵化率为 p, 孵化数记为Y, 求

a) X, Y的联合分布律;

b) X, Y的边缘分布律.

分析: a) 由题意知, 当产卵数 x 固定时, $Y\sim B(x, p)$, 故

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

$$=\binom{i}{j}p^j(1-p)^{i-j}\cdot e^{-\lambda}\frac{\lambda^i}{i!},\quad i\ge j,\ i=0,1,\ldots$$

b) $p_{i\bullet}=e^{-\lambda}\dfrac{\lambda^i}{i!},\quad i=0,1,\ldots \qquad p_{\bullet j}=e^{-\lambda p}\dfrac{(\lambda p)^j}{j!},\quad j=0,1,\cdots$

今后我们知道泊松分布中参数λ, 即“平均个数”的意思, 问 λp 的意义?

3. 边缘密度函数

若X, Y有联合函数$f(x, y)$, X, Y的密度函数如何？

从分布函数开始:

$$F(x,y)=\int_{-\infty}^{x}\int_{-\infty}^{y}f(u,v)\mathrm{d}u\mathrm{d}v$$

$$F_X(x)=\lim_{y\to\infty}F(x,y)=\int_{-\infty}^{x}\left(\int_{-\infty}^{+\infty}f(u,v)\mathrm{d}v\right)\mathrm{d}u$$

对 x 求导, 即得 X 的边缘密度函数.

$$f_X(x)=F_X'(x)=\int_{-\infty}^{+\infty}f(x,y)\mathrm{d}y$$

同理, 可得 Y 的边缘密度函数为:

$$f_Y(y)=\int_{-\infty}^{+\infty}f(x,y)\mathrm{d}x$$

注 联合密度函数往往是分片定义的, 在计算边缘密度函数积分时, 需要仔细确定积分限.

例 设D为xoy平面上由$x=0$, $x=1$, $x=y$, $x=y-1$围成的区域, 定义随机变量 X, Y 联合密度函数如下

$$f(x,y)=\begin{cases} c, & (x,y)\in D \\ 0, & \text{其他} \end{cases}$$

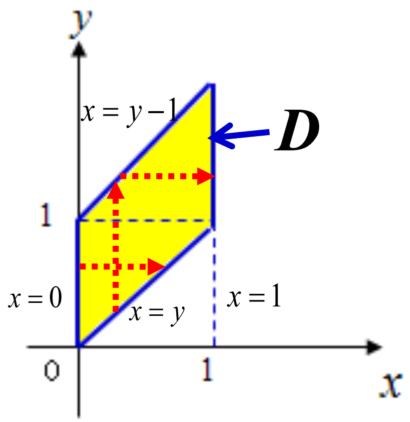
a) 确定常数c;

b) 求 X, Y 的边缘密度函数 $f_X(x)$, $f_Y(y)$.

解: 我们有

$$1=\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} f(x,y)\mathrm{d}x\mathrm{d}y=\iint_D c\mathrm{d}x\mathrm{d}y$$

$$=c$$



$$f_X(x)=\int_{-\infty}^{+\infty} f(x,y)\mathrm{d}y=\int_x^{x+1}\mathrm{d}y=1,\quad (0\le x\le 1)$$

$$f_Y(y)=\int_{-\infty}^{+\infty} f(x,y)\mathrm{d}x=\begin{cases}\int_0^y \mathrm{d}x=y, & 0\le y\le 1\\ \int_{y-1}^1 \mathrm{d}x=2-y, & 1<y\le 2\\ 0, & \text{其他}\end{cases}$$

再看一个例子.

例 设 X, Y 的联合密度函数为

$$f(x,y)=\begin{cases}\dfrac{21}{4}x^2 y, & x^2\le y\le 1\\ 0, & \text{其他}\end{cases}$$

求 X, Y 的边缘密度函数 $f_X(x), f_Y(y)$.

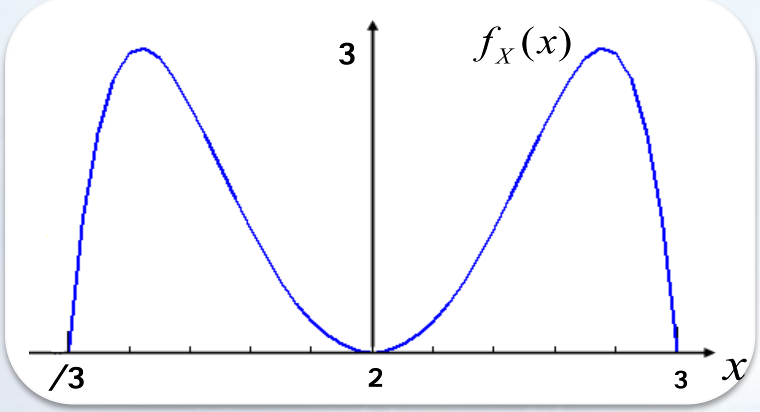
第15讲　边缘分布

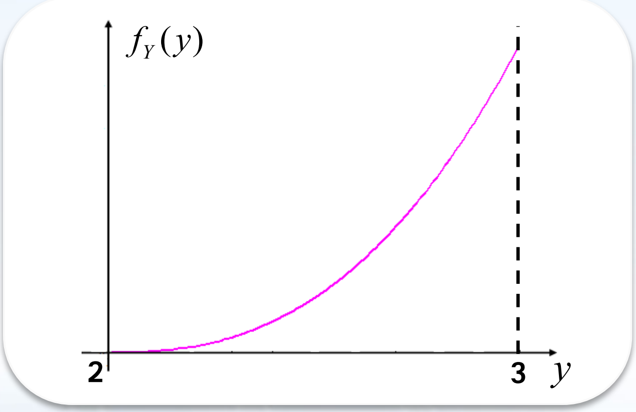
解:

$$f_X(x)=\int_{-\infty}^{+\infty} f(x,y)\mathrm{d}y \quad =\int_{x^2}^{1}\frac{21}{4}x^2y\mathrm{d}y$$

$$=\frac{21}{8}x^2(1-x^4),\qquad -1\le x\le 1$$

$$f_Y(y)=\int_{-\infty}^{+\infty} f(x,y)\mathrm{d}x=\int_{-\sqrt{y}}^{\sqrt{y}}\frac{21}{4}x^2y\mathrm{d}x=\frac{7}{2}y^{\frac{5}{2}},\quad 0\le y\le 1$$





4. 推广

二维联合分布的边缘分布的概念可以推广到 n维.

设$X_1, X_2, \ldots, X_n$的联合分布函数为 $F(x_1, x_2, \cdots, x_n)$,

则对于任意的$n_1, n_2, \ldots, n_k, 1 \le k < n$, 可类似定义$k$维

边缘分布函数 $F_{n_1,\cdots,n_k}(x_{n_1}, \cdots, x_{n_k})$.

可见这样的边缘分布函数(相应的分布律, 密度函数)共有 $\sum_{k=1}^{n-1}\binom{n}{k}=2^n-2$ 个