§3 条件分布

➢ 回忆：条件概率定义

$$P(A|B)=\frac{P(AB)}{P(B)},\quad P(B)>0$$

➢ **例** 在1, 2, 3, 4 中随机取出一数X ,再随机地从$1\sim X$中取一数Y, 求(X,Y)的联合分布律。

思考：如何确定Y的分布律？

$$p_{ij} = P\{X = i, Y = j\} \boxed{= P\{X = i\} P\{Y = j | X = i\}}$$

$$= \begin{cases} 0, & j > i \\ \dfrac{1}{4i}, & j \le i \end{cases} \qquad i, j = 1,2,3,4$$

X \ Y	1	2	3	4	
1	1/4	0	0	0	
2	1/8	1/8	0	0	
3	1/12	1/12	1/12	0	
4	1/16	1/16	1/16	1/16	
					1

思路：

确定联合分布律后，得到Y的边缘分布律

——要确定联合分布律，需先确定条件分布.

一、条件分布律

定义：设(X,Y)的联合分布律为:

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

若$P\{Y=y_j\}>0$，则在事件$\{Y=y_j\}$发生的条件下,事件$\{X=x_i\}$, $i=1,2,\cdots$ 发生的条件概率为

$$P\{X=x_i \mid Y=y_j\}=\frac{p_{ij}}{p_{\cdot j}} \quad i=1,2,\ldots. \qquad (*)$$

此概率数列具有分布律的性质:

1) $P\{X = x_i | Y = y_j\} \geq 0 \quad i = 1,2,....$

2) $\sum_{i=1}^{+\infty} P\{X = x_i | Y = y_j\} = 1$

称(*)为在$Y = y_j$的条件下,随机变量X的条件分布律.

条件分布律

婴儿数目的分布律

如何判断两个离散型随机变量X,Y相互独立?

如何判断两个离散型随机变量X,Y相互独立?

1) $F(x,y)=F_X(x)F_Y(y)$

2) $p_{ij}=p_{i\cdot}\,p_{\cdot j}$

3) $P\{X=i|Y=j\}=P\{X=i\}$

4) $P\{Y=j|X=i\}=P\{Y=j\}$

$i,j=1,2,\cdots$

对比：事件的独立性

二、条件概率密度

与条件分布律的定义几乎相同，只是用概率密度函数代替分布律即可。

定义： 设(X,Y)的联合概率密度函数为$f(x,y)$，

若边缘概率密度函数$f_X(x)>0$，则在$\{X=x\}$的条件下，Y的**条件概率密度函数**为

$$f_{Y|X}(y|x) == \frac{f(x,y)}{f_X(x)}$$

条件概率密度例一　　条件概率密度例二

如何判断两个连续型随机变量X,Y相互独立?

1) $F(x,y)=F_X(x)F_Y(y)$

2) $f(x,y)=f_X(x)f_Y(y)$

3) $f_{Y|X}(y|x)=f_Y(y)$

4) $f_{X|Y}(x|y)=f_X(x)$

$x,y\in R$

三、条件分布函数

条件分布函数$F_{X|Y}(x,y)$应如何定义?

$P\{X \le x|Y \le y\}$? $P\{X = x|Y = y\}$? $P\{X = x|Y \le y\}$?

结合实例思考: 警察通过调查犯罪现场的脚印，确定罪犯的身高?

$$F_{X|Y}(x|y_0) = P\{X \le x|Y = y_0\}$$

由于不能保证$P\{Y = y_0\} > 0$. 所以在一般情况下，就不能用条件概率的定义来直接定义条件分布函数.

这时需采用极限的方法来定义条件分布函数.

定义：给定$y_0 \in R$，对任意$\Delta y > 0$有$P\{y_0 < Y \le y_0 + \Delta y\} > 0$，且对任意$x \in R$，若极限

$$\lim_{\Delta y \to 0^+} P\left\{X \le x \middle| y_0 < Y \le y_0 + \Delta y\right\}$$

存在，称此极限函数为在$Y = y_0$的条件下，随机变量的**条件分布函数**。记作$F_{X|Y}(x|y_0)$

➢ 设(X,Y)是连续型随机变量，且满足$f(x,y), f_Y(y)$在(x,y_0)附近连续，且$f_Y(y_0) > 0$，则有

$$F_{X|Y}\left(x \middle| y_0\right) = \int_{-\infty}^{x} \frac{f\left(u, y_0\right)}{f_Y\left(y_0\right)} du$$

$$F_{X|Y}(x|y_0)=\int_{-\infty}^{x}\frac{f(u,y_0)}{f_Y(y_0)}du$$

证明：

$$F_{X|Y}(x|y)=\lim_{\Delta y\to 0^+}P\{X\le x|y_0<Y\le y_0+\Delta y\}$$

$$=\lim_{\Delta y\to 0^+}\frac{P\{X\le x,y_0<Y\le y_0+\Delta y\}}{P\{y_0<Y\le y_0+\Delta y\}}$$

$$=\lim_{\Delta y\to 0^+}\frac{F(x,y_0+\Delta y)-F(x,y_0)}{F_Y(y_0+\Delta y)-F_Y(y_0)}$$

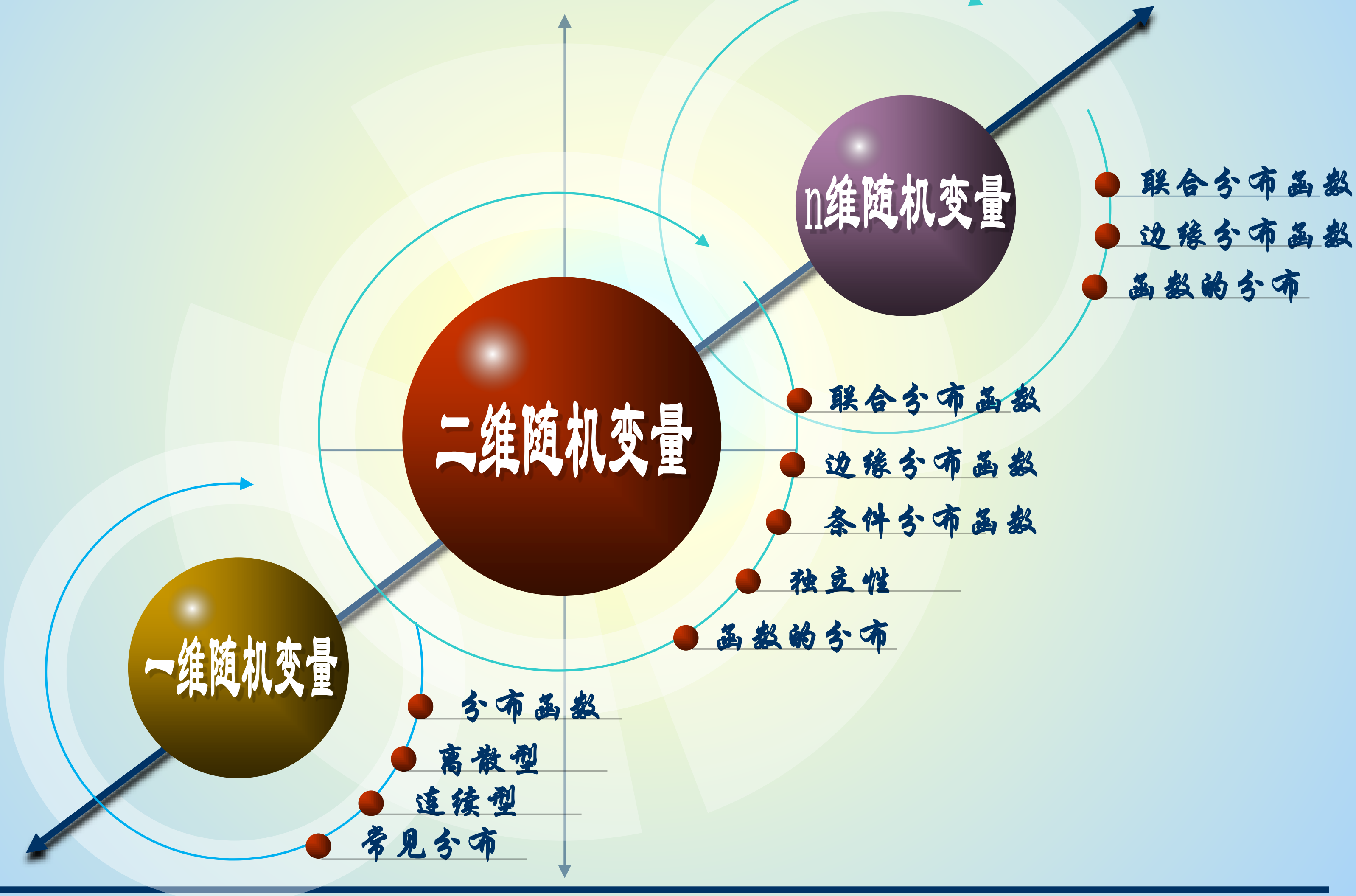
$$= \lim_{\Delta y \to 0^+} \frac{\dfrac{F(x, y_0 + \Delta y) - F(x, y_0)}{\Delta y}}{\dfrac{F_Y(y_0 + \Delta y) - F_Y(y_0)}{\Delta y}}$$

$$= \frac{\left.\dfrac{\partial F(x, y)}{\partial y}\right|_{y=y_0}}{f_Y(y_0)} = \int_{-\infty}^{x} \frac{f(u, y_0)}{f_Y(y_0)} du = F_{X|Y}(x|y_0)$$

$$f_{X|Y}(x|y_0) = F'_{X|Y}(x|y_0) = \frac{f(x, y_0)}{f_Y(y_0)}$$

为在$Y = y_0$的条件下随机变量X的**条件概率密度**.

□ 多维随机变量



n维随机变量
联合分布函数
边缘分布函数
函数的分布
二维随机变量
联合分布函数
边缘分布函数
条件分布函数
独立性
函数的分布
一维随机变量
分布函数
离散型
连续型
常见分布

□二维随机变量

尝试：补充相关公式！