西安電子科技大學
XIDIAN UNIVERSITY

第3章 多维随机变量及其分布

概率论与数理统计课程组

CHAPTER 3

多维随机变量及其分布

3.6 随机变量的独立性

问题的提出

定义了**二维随机变量**并获取其联合分布之后，有可能需要探讨两个变量之间的独立性，讨论二者的发生与否有无相互影响的关系

例如：

为研究某地区学龄儿童的发育情况，同时考察了每个儿童的身高和体重，即二维随机变量(X、Y)，为研究二者间的关系，可探讨两个变量X与Y之间的独立性

定义 ——分布函数

设$F(x, y)$及$F_X(x)$、 $F_Y(y)$是二维随机变量 (X、Y) 的联合分布函数及边缘分布函数，若对所有x、y有 $F(x,y)=F_X(x)F_Y(y)$

即 $P(X \le x, Y \le y) = P(X \le x)P(Y \le y)$ 称随机变量X和Y 相互独立

定理1 ——分布律

设二维离散型随机变量 (X、Y)的联合分布律是 $P(X = x_i, Y = y_j) = p_{ij}$， $i, j = 1, 2, \cdots$

边缘分布律分别是 $P(X = x_i) = p_{i\cdot}$， $P(Y = y_j) = p_{\cdot j}$ 则X和Y 相互独立等价于

$P(X = x_i, Y = y_j) = P(X = x_i)P(Y = y_j)$ 即 $p_{ij} = p_{i\cdot} \cdot p_{\cdot j}$ 对一切 i 和 j 都成立

定理2 ——概率密度

设 (X、Y)是连续型随机变量的联合概率密度$f(x, y)$及边缘概率密度$f_X(x)$、 $f_Y(y)$是除有限点外的连续函数，则X和Y 相互独立等价于 $f(x,y) = f_X(x) f_Y(y)$

例 (X,Y)具有联合分布律（下图）
求二者是否独立？

$Y \backslash X$	0	1	$P(Y=j)$
1	1/6	2/6	1/2
2	1/6	2/6	1/2
$P(X=i)$	1/3	2/3	

解

$$P(X=0,Y=1)=1/6=P(X=0)P(Y=1)$$

$$P(X=0,Y=2)=1/6=P(X=0)P(Y=2)$$

$$P(X=1,Y=1)=2/6=P(X=1)P(Y=1)$$

$$P(X=1,Y=2)=2/6=P(X=1)P(Y=2)$$

X、Y 相互独立

例 (X,Y)具有联合分布律（右图）

求二者是否独立？

$Y \backslash X$	0	1	$P(Y=j)$
1	1/6	2/6	1/2
2	2/6	1/6	1/2
$P(X=i)$	1/2	1/2	

解 $P(X=0,Y=1)=1/6$

$P(X=0)P(Y=1)$

$=\frac{1}{2}\times\frac{1}{2}=\frac{1}{4}\neq P(X=0,Y=1)$ X、Y不相互独立

例 X、Y是相互独立的随机变量，已知(X, Y)联合分布律，求表中剩余的概率值

解

$X \backslash Y$	0	1	2	$P(X=i)$
1	0.01	0.2	0.04	0.25
2	0.03	0.6	0.12	0.75
$P(Y=j)$	0.04	0.8	0.16	

例 (X,Y)具有联合概率密度 $f(x,y)=\begin{cases}6e^{-(2x+3y)}, & x>0, y>0\\ 0, & \text{其他}\end{cases}$

二者是否独立？

解 X 和 Y 的边缘概率密度分别为 $f_X(x)=\begin{cases}2e^{-2x}, & x>0\\ 0, & x\le 0\end{cases}$

$$f_Y(y)=\begin{cases}3e^{-3y}, & y>0\\ 0, & y\le 0\end{cases}$$

满足 $f(x,y)=f_X(x)f_Y(y)$ 故X、Y 相互独立

例 证明：对二维正态随机变量(X, Y)，X、Y相互独立的充要条件是参数 $\rho=0$

证 (X, Y)联合概率密度为

$$f(x,y)=\frac{1}{2\pi\sigma_1\sigma_2\sqrt{1-\rho^2}}\times \exp\left\{\frac{-1}{2(1-\rho^2)}\left[\frac{(x-\mu_1)^2}{\sigma_1^2}-2\rho\frac{(x-\mu_1)(y-\mu_2)}{\sigma_1\sigma_2}+\frac{(y-\mu_2)^2}{\sigma_2^2}\right]\right\}$$

由之前例题知，X、Y边缘概率密度为 $f_X(x)=\frac{1}{\sqrt{2\pi}\sigma_1}e^{-\frac{(x-\mu_1)^2}{2\sigma_1^2}}$ $\quad f_Y(y)=\frac{1}{\sqrt{2\pi}\sigma_2}e^{-\frac{(y-\mu_2)^2}{2\sigma_2^2}}$

$$f_X(x)f_Y(y)=\frac{1}{2\pi\sigma_1\sigma_2}\exp\left\{-\frac{1}{2}\left[\frac{(x-\mu_1)^2}{\sigma_1^2}+\frac{(y-\mu_2)^2}{\sigma_2^2}\right]\right\}$$

"$\Rightarrow$" 如果$\rho=0$，则对于所有x、y有 $f(x,y)=f_X(x)f_Y(y)$ 即X、Y相互独立

"$\Leftarrow$" 反之，若X、Y相互独立，由于$f(x, y)$及$f_X(x)$、$f_Y(y)$都是连续函数，故对所有x、y有 $f(x,y)=f_X(x)f_Y(y)$ 特别的有 $f(\mu_1,\mu_2)=f_X(\mu_1)f_Y(\mu_2)$

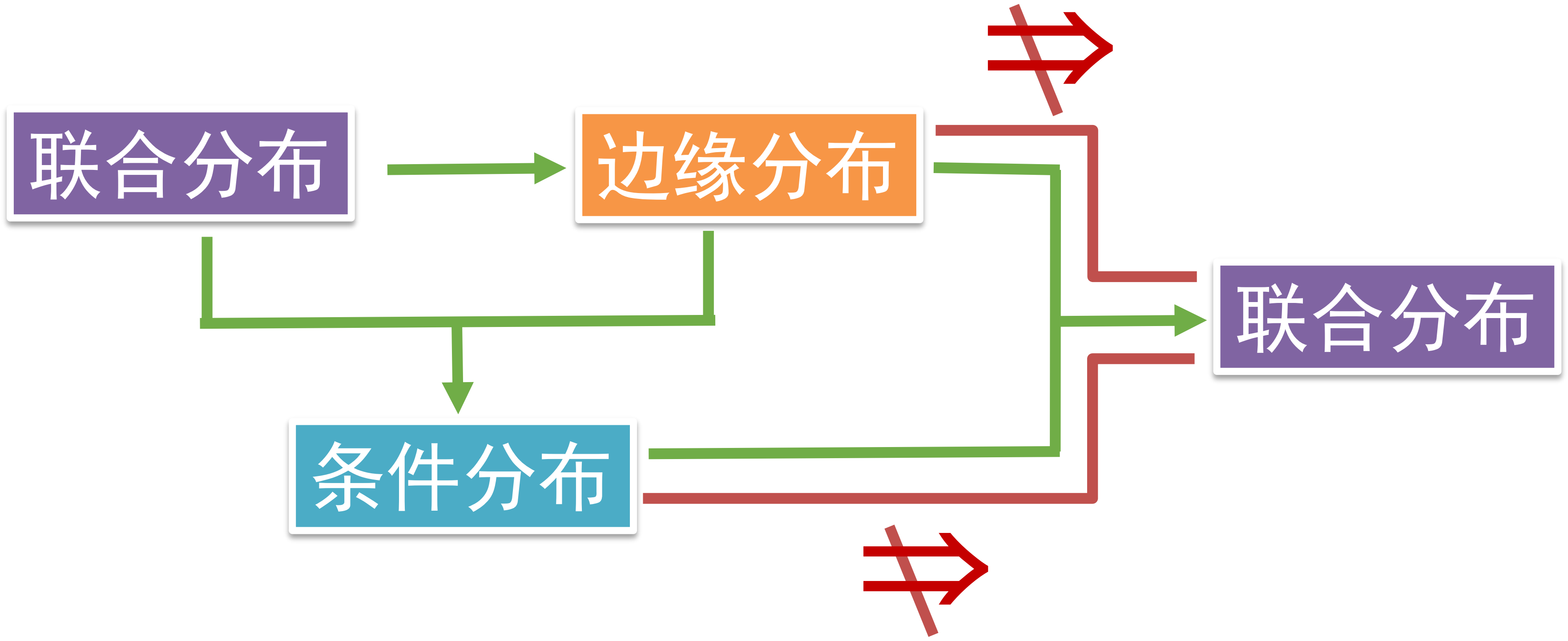
即 $\frac{1}{2\pi\sigma_1\sigma_2\sqrt{1-\rho^2}}=\frac{1}{2\pi\sigma_1\sigma_2}$ $\qquad \rho=0$

注意

如果只知道关于X和关于Y的边缘概率分布，一般不能推导出(X, Y)的联合概率分布

例如：若二维随机变量(X, Y)中的X和Y都分别服从参数已知的一维正态分布，但不知参数ρ值的情况下仍然无法获知二维正态随机变量的具体分布

联合分布、边缘分布、条件分布的关系图示



条件分布 = $\dfrac{\text{联合分布}}{\text{边缘分布}}$

联合分布 = 边缘分布 × 条件分布

例 设甲、乙两种元器件的寿命X、Y相互独立，服从同一分布，其概率密度为

$$f(x)=\begin{cases}\frac{1}{2}e^{-\frac{x}{2}}, & x>0\\ 0, & x\le 0\end{cases}$$

求甲寿命不大于乙寿命2倍的概率。

解 (X, Y)的联合概率密度为

$$f(x,y)=f_X(x)f_Y(y)=\begin{cases}\frac{1}{4}e^{-\frac{x+y}{2}} & x>0, y>0\\ 0 & 其他\end{cases}$$

$$P(X\le 2Y)=\int_0^{+\infty}dx\int_{x/2}^{+\infty}\frac{1}{4}e^{-\frac{x+y}{2}}dy=\int_0^{+\infty}\frac{1}{2}e^{-\frac{x}{2}}e^{-\frac{x}{4}}dx=\int_0^{+\infty}\frac{1}{2}e^{-\frac{3x}{4}}dx=\frac{2}{3}$$

边缘分布

$(X_1, X_2, \ldots, X_n)$的联合分布函数 $F(x_1, x_2, \ldots, x_n)$已知，则$(X_1, X_2, \ldots, X_n)$的$k\ (1\le k \le n)$维边缘分布函数就随之确定

例如边缘分布函数 $F_{X_1}(x_1)=F(x_1,+\infty,+\infty,\cdots,+\infty)$

$$F_{X_1,X_2}(x_1,x_2)=F(x_1,x_2,+\infty,\cdots,+\infty)$$

其边缘分布律 $P(X_1=x_{1i_1})=\sum\limits_{i_2,i_3,\cdots i_n} P(X_1=x_{1i_1},X_2=x_{2i_2},\cdots,X_n=x_{ni_n})$

$$P(X_1=x_{1i_1},X_2=x_{2i_2})=\sum_{i_3,i_4,\cdots i_n} P(X_1=x_{1i_1},X_2=x_{2i_2},\cdots,X_n=x_{ni_n})$$

其边缘概率密度 $f_{X_1}(x_1)=\int_{-\infty}^{+\infty}\cdots\int_{-\infty}^{+\infty} f(x_1,x_2,\cdots,x_n)dx_2dx_3\cdots dx_n$

$$f_{X_1,X_2}(x_1,x_2)=\int_{-\infty}^{+\infty}\cdots\int_{-\infty}^{+\infty} f(x_1,x_2,\cdots,x_n)dx_3dx_4\cdots dx_n$$

推广至n维随机变量 ($n\geq 2$)

设E是一个随机试验，样本空间$\Omega=\{e\}$；$X_1=X_1(e), X_2=X_2(e),\ldots, X_n=X_n(e)$是定义在$\Omega$上的随机变量，$n$维向量$(X_1, X_2, \ldots, X_n)$称$n$维随机变量或$n$维随机向量

n维随机变量的联合分布函数

对于任意实数n个实数$x_1, x_2, \ldots, x_n$有n元函数 $F(x_1,x_2,\ldots,x_n)=P(X_1\leq x_1,X_2\leq x_2,\ldots,X_n\leq x_n)$

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

n维离散型随机变量的联合分布律

设 $(X_1, X_2, \ldots, X_n)$所有可能取值为$(x_{1i_1},x_{2i_2},\cdots,x_{ni_n})$ $i_j=1, 2, \ldots$

$$P(X_1=x_{1i_1},X_2=x_{2i_2},\cdots,X_n=x_{ni_n})\quad j=1,2,\cdots n$$

称n维离散型随机变量 $(X_1, X_2, \ldots, X_n)$的联合分布律

n维连续型随机变量的联合概率密度

若存在非负函数 $f(x_1, x_2, \ldots, x_n)$，使得对于任意实数$x_1, x_2, \ldots, x_n$， 有

$$F(x_1,x_2,\cdots x_n)=\int_{-\infty}^{x_n}\int_{-\infty}^{x_{n-1}}\cdots\int_{-\infty}^{x_1}f(x_1,x_2,\cdots x_n)dx_1dx_2\cdots dx_n$$

称其为n维连续型随机变量 $(X_1, X_2, \ldots, X_n)$的联合概率密度

相互独立

若对于所有$x_1, x_2, \ldots, x_n$，有

$$F(x_1, x_2, \cdots, x_n) = F_{X_1}(x_1)F_{X_2}(x_2)\cdots F_{X_n}(x_n)$$

则称$X_1, X_2, \ldots, X_n$相互独立

若对于所有$x_1, x_2, \ldots, x_m$； $y_1, y_2, \ldots, y_n$有

$$F(x_1, x_2, \cdots, x_m, y_1, y_2, \cdots, y_n) = F_1(x_1, x_2, \cdots x_m)F_2(y_1, y_2, \cdots y_n)$$

其中F_1、F_2、F依次为随机变量$(X_1, X_2, \ldots, X_m)$、$(Y_1, Y_2, \ldots, Y_n)$、$(X_1, X_2, \ldots, X_m, Y_1, Y_2, \ldots, Y_n)$的联合分布函数

则称$(X_1, X_2, \ldots, X_m)$与$(Y_1, Y_2, \ldots, Y_n)$相互独立

定理1

设$(X_1, X_2, \ldots, X_m)$与$(Y_1, Y_2, \ldots, Y_n)$相互独立，则$X_i\,(i=1, 2, \ldots, m)$与$Y_j\,(j=1, 2, \ldots, n)$ 相互独立

设$h(x_1, x_2, \ldots, x_m)$与$g(y_1, y_2, \ldots, y_n)$是连续函数，则$h(X_1, X_2, \ldots, X_m)$与$g(Y_1, Y_2, \ldots, Y_n)$相互独立

定理2

设$X_1, X_2, \ldots, X_m$相互独立，将其分成任意k个没有相同随机变量的不同小组，并对每个小组的随机变量施以相应连续函数运算后，所得到的k个随机变量也相互独立

本节回顾

相互独立的判断

分布函数： 设$F(x, y)$及$F_X(x)$、 $F_Y(y)$是二维随机变量 $(X、Y)$ 的联合分布函数及边缘分布函数，若对所有x、y有 $F(x,y)=F_X(x)F_Y(y)$

即 $P(X\le x,Y\le y)=P(X\le x)P(Y\le y)$ 称随机变量X和Y 相互独立

分布律： 设二维离散型随机变量 $(X、Y)$的联合分布律是 $P(X=x_i,Y=y_j)=p_{ij},\ i,j=1,2,\cdots$

边缘分布律分别是 $P(X=x_i)=p_{i\bullet}$， $P(Y=y_j)=p_{\bullet j}$ 则X和Y 相互独立等价于

$P(X=x_i,Y=y_j)=P(X=x_i)P(Y=y_j)$ 即 $p_{ij}=p_{i\cdot}\cdot p_{\cdot j}$ 对一切 i 和 j 都成立

概率密度： 设 $(X、Y)$是连续型随机变量的联合概率密度$f(x, y)$及边缘概率密度$f_X(x)$、 $f_Y(y)$是除有限点外的连续函数，则X和Y 相互独立等价于 $f(x,y)=f_X(x)f_Y(y)$