第三章 多维随机变量及其分布

§1 二维随机变量
§2 边缘分布
§3 条件分布
§4 相互独立的随机变量

§3 条件分布(Conditional Distribution)

3.1 从条件概率到条件分布

第一章内容

设试验的基本事件总数为n，A所包含的基本事件数为m(m>0)，AB所包含的基本事件数为k，即有

$$P(B|A)=\frac{k}{m}=\frac{k/n}{m/n}=\frac{P(AB)}{P(A)}$$

条件概率定义：设A, B是两个事件，且P(A)>0，称

$$\mathrm{P(B|A)}=\frac{P(AB)}{P(A)}$$

为在事件A发生的条件下事件B发生的条件概率

$P(A)$与边缘分布

设(X,Y)是二维离散型随机变量，其分布律为

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

(X,Y)关于 X 和关于 Y 的边缘分布律分别为

$$P\{X=x_i\}=p_{i\cdot}=\sum_j p_{ij}, \qquad (i=1,2,\dots,)$$

$$P\{Y=y_j\}=p_{\cdot j}=\sum_i p_{ij}, \qquad (j=1,2,\dots,)$$

3.2 二维离散型随机变量的条件分布

定义：设(X,Y)是二维离散型随机变量，对于固定的j，若$P\{Y=y_j\}>0$, 则称

$$P\left(X=x_i \middle| Y=y_j\right)=\frac{P\{X=x_i,Y=y_j\}}{P\{Y=y_j\}}=\frac{p_{ij}}{p_{\cdot j}},$$

$$i=1,2,\ldots$$

为在$Y=y_j$ 条件下随机变量 X 的条件分布律。

条件分布律的性质

1. 非负性：

$P\{X = x_i | Y = y_j\} \geq 0$；

2. 规范性：

$$\sum_{i=1}^{\infty} P\{X = x_i | Y = y_j\} = \sum_{i=1}^{\infty} \frac{p_{ij}}{p_{\cdot j}} = \frac{1}{p_{\cdot j}} \sum_{i=1}^{\infty} p_{ij}$$

$$= \frac{p_{\cdot j}}{p_{\cdot j}} = 1$$

例1. (X,Y)的联合分布律及其边缘分布律为：

X \ Y	0	1	2	$p_{i\cdot}$
0	1/9	2/9	1/9	4/9=$p_{0\cdot}$
1	2/9	2/9	0	4/9=$p_{1\cdot}$
2	1/9	0	0	1/9=$p_{2\cdot}$
$p_{\cdot j}$	4/9=$p_{\cdot 0}$	4/9=$p_{\cdot 1}$	1/9=$p_{\cdot 2}$	**1**

在则在Y= 0条件下，随机变量 X 的条件分布律为

$$P\{X=0|Y=0\}=\frac{p_{00}}{p_{\cdot 0}}=\frac{1/9}{4/9}=\frac{1}{4}$$

$$P\{X=1|Y=0\}=\frac{p_{10}}{p_{\cdot 0}}=\frac{2/9}{4/9}=\frac{1}{2}$$

$$P\{X=2|Y=0\}=\frac{p_{20}}{p_{\cdot 0}}=\frac{1/9}{4/9}=\frac{1}{4}$$

例2. 一射手进行射击，击中目标的概率为 p ，射击到击中目标两次为止。设以X表示首次击中目标所进行的射击次数，以Y表示总共进行的射击次数，试求X和Y的联合分布律以及条件分布律。

解：Y的取值是$2,3,\ldots$， X的取值是$1,2,\ldots$，并且$X < Y$。

于是X,Y的联合分布律为

$P\{X=m,Y=n\}=$
$\{$第m次首次击中，且第n次第二次击中$\}$

由独立性有，

$$P\{X=m,Y=n\}=P\{X=m\}P\{Y=n\}$$
$$=(1-p)^{m-1}p(1-p)^{n-m-1}p=(1-p)^{n-2}p^2$$

X的边缘分布律为

$$P\{X=m\}=\sum_{n=m+1}^{\infty}P\{X=m,Y=n\}$$

$$=\sum_{n=m+1}^{\infty}(1-p)^{n-2}p^{2}$$

$$=p^{2}\sum_{n=m+1}^{\infty}(1-p)^{n-2}=p^{2}\frac{(1-p)^{(m+1)-2}}{1-(1-p)}$$

$$=p(1-p)^{m-1}$$

$$m=1,2,\ldots$$

于是，在$X=m$条件下随机变量Y的条件分布为：

$$P\{Y=n|X=m\}=\frac{P\{X=m,Y=n\}}{P\{X=m\}}$$

$$=\frac{p^2(1-p)^{n-2}}{p(1-p)^{m-1}}$$

$$=p(1-p)^{n-m-1}\quad ,n=m+1,m+2,\dots$$

直觉上的一致：后面n-m次射击中，恰好只有一次命中，而n-m-1次脱靶

类似地，Y的边缘分布律为

$$P\{Y=n\}=\sum_{m=1}^{n-1}P\{X=m,Y=n\}$$

$$=\sum_{m=1}^{n-1}(1-p)^{n-2}p^2=(n-1)(1-p)^{n-2}p^2,$$

$$n=2,3,\ldots$$

在$Y=n$条件下随机变量X的条件分布为：

$$P\{X=m|Y=n\}=\frac{P\{X=m,Y=n\}}{P\{Y=n\}}$$

$$=\frac{(1-p)^{n-2}p^2}{(n-1)(1-p)^{n-2}p^2}=\frac{1}{n-1}$$

$$(m=1,2,\ldots,n-1,n=2,3,\ldots)$$

直觉上的解释？

3.3 二维连续型随机变量的条件分布

注意：对连续型分布而言$P\{X = x_i\} = 0$，$P\{Y = y_i\} = 0$，因而不能直接代入条件概率公式，利用极限引入条件分布函数

定义：对给定的 y，设对与任意固定的正数$\epsilon > 0$，有

$$P\{y < Y \leq y + \epsilon\} > 0，$$

若对于任意实数x，下面极限存在，

$$\lim_{\epsilon \to 0} P\{X \leq x | y < Y \leq y + \epsilon\} = \lim_{\epsilon \to 0} \frac{P\{X \leq x, y < Y \leq y + \epsilon\}}{P\{y < Y \leq y + \epsilon\}}$$

则称上式为在条件$Y = y$下X的条件分布函数，写成$P\{X < x | Y = y\}$，或记为$F_{X|Y}(x|y)$

从条件分布函数到条件概率密度

$$F_{X|Y}(x\mid y)=\lim_{\varepsilon\to 0^+}\frac{P\{X\le x, y<Y\le y+\varepsilon\}}{P\{y<Y\le y+\varepsilon\}}$$

$$=\lim_{\varepsilon\to 0^+}\frac{F(x,y+\varepsilon)-F(x,y)}{F_Y(y+\varepsilon)-F_Y(y)}=\lim_{\varepsilon\to 0^+}\frac{\left[F(x,y+\varepsilon)-F(x,y)\right]/\varepsilon}{\left[F_Y(y+\varepsilon)-F_Y(y)\right]/\varepsilon}$$

$$=\frac{\frac{\partial}{\partial y}F(x,y)}{\frac{\partial}{\partial y}F_Y(y)}=\frac{\frac{\partial}{\partial y}\left(\int_{-\infty}^{y}\int_{-\infty}^{x}f(u,v)dudv\right)}{f_Y(y)}$$

$$=\frac{\int_{-\infty}^{x}f(u,y)du}{f_Y(y)}=\int_{-\infty}^{x}\frac{f(u,y)}{f_Y(y)}du$$

从条件分布函数到条件概率密度

注意$\frac{f(u,y)}{f_Y(y)}$的形式，引入条件概率密度

定义：设二维随机变量(X,Y)的概率密度为$f(x,y)$，关于Y的边缘概率密度为$f_Y(y)$，若对固定的y，$f_Y(y)>0$，则称

$$\frac{f(x,y)}{f_Y(y)}$$

思考：逻辑上的含义

为在$Y=y$的条件下X的条件概率密度，

记为

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

条件密度函数的性质

1. 非负性：
   对任意的x，有$f_{X|Y}(x|y) \geq 0$
2. 规范性(归一化特性的体现)：

$$\int_{-\infty}^{\infty} f_{X|Y}(x|y)dx = 1$$

保证了$f_{X|Y}(x|y)$满足密度函数的要求

上述内容和性质对$f_{Y|X}(y|x)$和$F_{Y|X}(y|x)$同样适用

边缘密度函数

例3. 设X在区间(0,1)上随机均匀地取值，当$X = x$ $(0 < x < 1)$时Y在区间$(x, 1)$上随机均匀地取值，求Y的概率密度$f_Y(y)$。

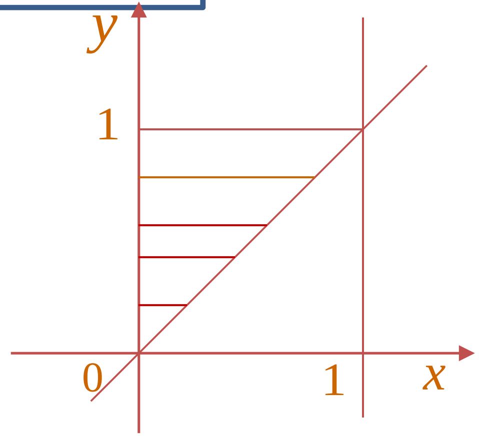
条件密度函数

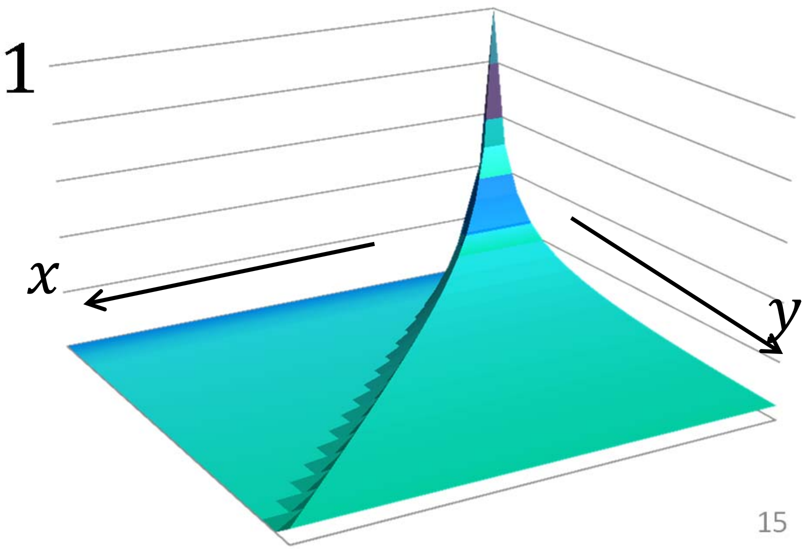
解：X的边缘概率密度函数为

$$f_X(x) = \begin{cases} 1 & 0 < x < 1 \\ 0 & \text{其他} \end{cases}$$

于是对于Y有相应的条件分布

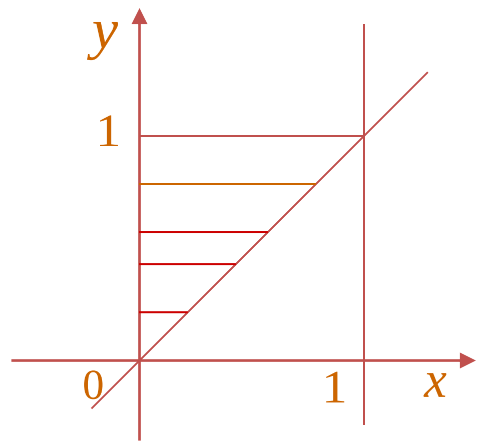
$$f_{Y|X}(y|x) = \begin{cases} \dfrac{1}{1-x} & x < y < 1 \\ 0 & \text{其他} \end{cases}$$





例3. 设X在区间(0,1)上随机均匀地取值，当$X=x$ $(0<x<1)$时Y在区间$(x,1)$上随机均匀地取值，求Y的概率密度$f_Y(y)$。

解(续)：



于是边缘分布为

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

$$=\begin{cases}\displaystyle\int_0^y\frac{1}{1-x}dx=-\ln(1-y) & 0<y<1\\ 0 & \text{其他}\end{cases}$$

例4. 设二维随机变量(X,Y)服从参数为$(\mu_1,\mu_2,\sigma_1^2,\sigma_2^2,\rho)$的正态分布，即密度函数为

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

给出X和Y的条件分布。

由本章第二节例10知

$$f_Y(y) = \frac{1}{\sqrt{2\pi}\sigma_2} e^{-\frac{(y-\mu_2)^2}{2\sigma_2^2}}$$

于是，条件概率密度

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

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

$$(-\infty < x < \infty)$$

类似地

$$f_X(x)=\frac{1}{\sqrt{2\pi}\sigma_1}e^{-\frac{(x-\mu_1)^2}{2\sigma_1^2}}$$

于是

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

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

$$(-\infty<y<\infty)$$

注意到只要$\rho\neq0$，$f_{X|Y}(x|y)$即与μ_2,σ_2有关，$f_{Y|X}(y|x)$即与μ_1,σ_1有关

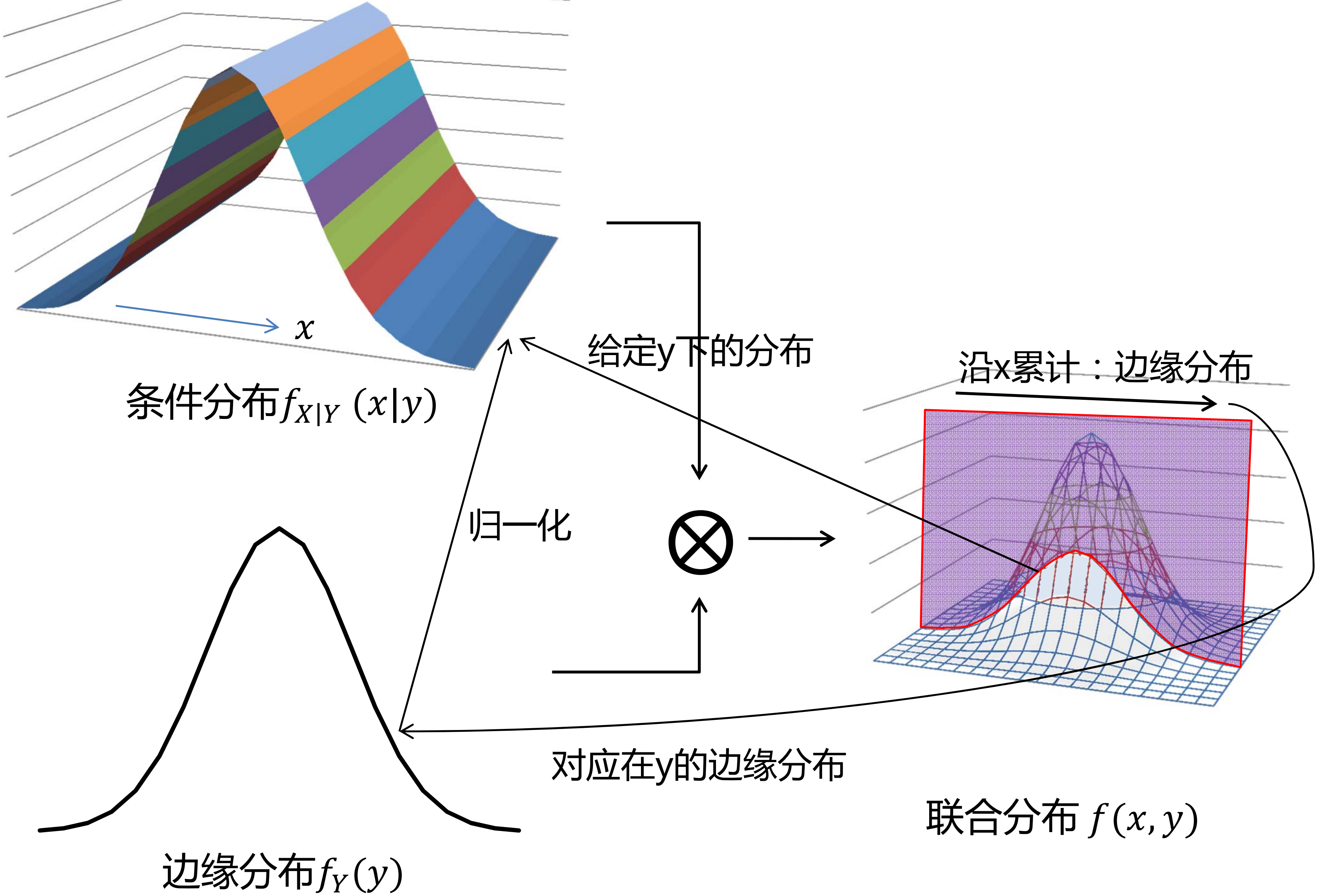
几何上的解释

1. 联合分布唯一确定边缘分布和条件分布
2. 边缘分布和条件分布个自都不能唯一确定联合分布
3. 一个条件分布和对应的边缘分布一起，能唯一确定联合分布，利用
$$f(x,y)=f_X(x)f_{Y|X}(y|x)=f_Y(y)f_{X|Y}(x|y)$$

可以看作是利用边缘分布对条件分布进行了调制

$$\begin{bmatrix} f(x_1,y_1) & f(x_1,y_2) & \cdots & f(x_1,y_n) \\ f(x_2,y_1) & f(x_2,y_2) & \cdots & f(x_2,y_n) \\ \vdots & \vdots & \ddots & \vdots \\ f(x_m,y_1) & f(x_m,y_2) & \cdots & f(x_m,y_n) \end{bmatrix} =$$

$$\begin{bmatrix} f_{X|Y}(x_1|y_1) & f_{X|Y}(x_1,y_2) & \cdots & f_{X|Y}(x_1,y_n) \\ f_{X|Y}(x_2|y_1) & f_{X|Y}(x_2,y_2) & \cdots & f_{X|Y}(x_2,y_n) \\ \vdots & \vdots & \ddots & \vdots \\ f_{X|Y}(x_m,y_1) & f_{X|Y}(x_m,y_2) & \cdots & f_{X|Y}(x_m,y_n) \end{bmatrix} \begin{bmatrix} f_Y(y_1) & 0 & \cdots & 0 \\ 0 & f_Y(y_2) & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & f_Y(y_n) \end{bmatrix}$$



x
条件分布$f_{X|Y}\ (x|y)$
给定y下的分布
沿x累计：边缘分布
归一化
$\otimes$
边缘分布$f_Y(y)$
对应在y的边缘分布
联合分布 $f(x,y)$

例5.设店主在每日开门营业时，放在柜台上的货物量为Y，当日销售量为X，假定一天中不再往柜台上补充货物，于是$X \leq Y$。根据历史资料,(X,Y)的概率密度函数为

$$f(x,y)=\begin{cases}\dfrac{1}{200} & 0 \leq x \leq y, 0 \leq y \leq 20 \\ 0 & \text{其他}\end{cases}$$

求:(1)给定$Y=y$条件下,X的条件概率密度.

(2)假定某日开门时,$Y=10$件,求这天顾客买走$X \leq 5$件的概率.

(3)如果$Y=20$件呢?

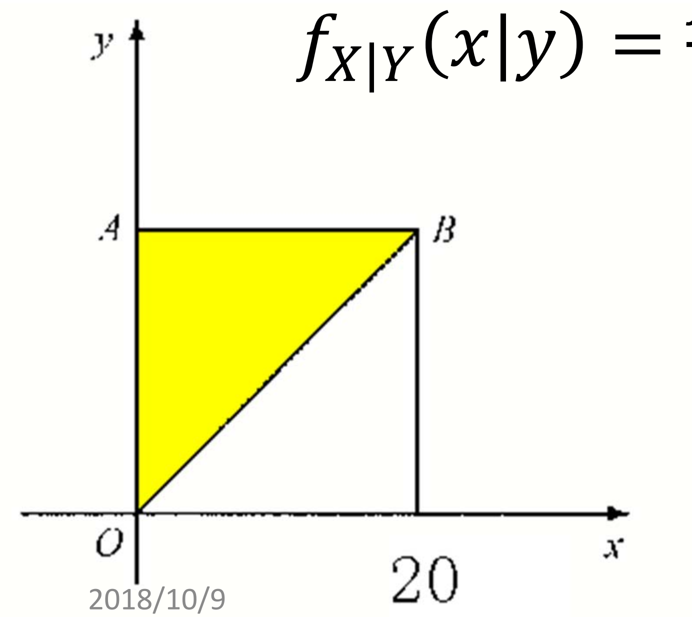
解: (1)

$$f_Y(y)=\int_{-\infty}^{+\infty} f(x,y)dx=\begin{cases}\int_0^y \frac{1}{200}dx & 0\le y\le 20\\ 0 & 其他\end{cases}$$

因此，$y\in(0,20]$时，$f_Y(y)>0$

故，

$$f_{X|Y}(x|y)=\frac{f(x,y)}{f_Y(y)}=\begin{cases}\frac{1}{y} & x\in[y\le 20]\\ 0 & x\notin(y,20\}\end{cases}$$



(2)Y=10时,顾客买走X≤5件的概率为

$$P\{X \le 5|Y = 10\} = F_{X|Y}(5|10)$$
$$= \int_{-\infty}^{5} f_{X|Y}(x|10)dx = \int_{0}^{5} \frac{1}{10}dx = 0.5$$

(3)Y=20时,顾客买走X≤5件的概率

$$P\{X \le 5|Y = 20\} = F_{X|Y}(5|20)$$
$$= \int_{-\infty}^{5} f_{X|Y}(x|20)dx = \int_{0}^{5} \frac{1}{20}dx = 0.25$$

$$\boxed{f_{X|Y}(x|y) = \frac{f(x,y)}{f_Y(y)} = \begin{cases} \frac{1}{y} & x \in [y \le 20] \\ 0 & x \notin (y, 20\} \end{cases}}$$

§4 相互独立的随机变量

第一章内容回顾——随机事件的独立

设 A、B 是两个随机事件，如果满足

$$P(AB)=P(A)P(B|A)=P(A)P(B)$$

则称 A 与 B 是相互独立的随机事件

4.1 随机变量的独立

定义：设$F(x,y)$及$F_X(x)$,$F_Y(y)$分别是二维随机变量(X,Y)的分布函数及边缘分布函数，若对所有x,y有

$$P\{X \le x, Y \le y\} = P\{X \le x\}P\{Y \le y\}$$

即 $F(x,y) = F_X(x)F_Y(y)$

独立性的含义

1. 所谓随机变量X与Y相互独立，形式上满足

$$P\{X \le x, Y \le y\} = P\{X \le x\}P\{Y \le y\}$$

本质上意味着两个事件：

$$\{X \le x\}\text{和}\{Y \le y\}$$

相互独立。

2. 注意：在X与Y相互独立时有

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

意味着，(X,Y)相互独立时，其联合分布可由其边缘分布$F_X(x)$和$F_Y(y)$唯一确定

4.2 相互独立的离散型随机变量

对于离散型随机变量(X, Y)，设其联合分布律为

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

又有随机变量X和Y的分布律分别为

$$p_{i\cdot} = P\{X = x_i\} \quad (i = 1,2, \dots)$$

和

$$p_{\cdot j} = P\{Y = y_j\} \quad (j = 1,2, \dots)$$

如果对任意的i, j，有

$$p_{ij} = p_{i\cdot} p_{\cdot j}$$

则称X和Y为相互独立的随机变量

联合分布律

X \ Y	y_1	y_2	…	y_j	…	$p_{i\cdot}$
x_1	p_{11}	p_{12}	…	p_{1j}	…	$p_{1\cdot}$
x_2	p_{21}	p_{22}	…	p_{2j}	…	$p_{2\cdot}$
$\vdots$	$\vdots$	$\vdots$	$\ddots$	$\vdots$	$\ddots$	$\vdots$
x_i	$\sum p_{i1}$	p_{i2}	…	p_{ij}	… $=$	$p_{i\cdot}$
$\vdots$	$\vdots$	$\vdots$	$\ddots$	$\vdots$	$\ddots$	$\vdots$
$p_{\cdot j}$	$p_{\cdot 1}$	$p_{\cdot 2}$	…	$p_{\cdot j}$	…	1

边缘分布

边缘分布

独立性条件：$p_{ij} = p_{i\cdot} p_{\cdot j}$

判断离散随机变量的相互独立

例6.

Y \ 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	1

显然

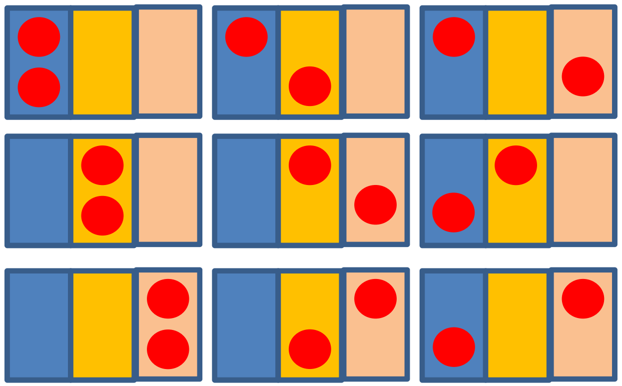
$$P\{X=0,Y=1\}=P\{X=0\}P\{Y=1\}$$
$$P\{X=0,Y=2\}=P\{X=0\}P\{Y=2\}$$
$$P\{X=1,Y=1\}=P\{X=1\}P\{Y=1\}$$
$$P\{X=1,Y=2\}=P\{X=1\}P\{Y=2\}$$

满足独立性条件

判断离散随机变量的相互独立

例7. 将两个球等可能地放入编号为1，2，3的盒子中，令，X：放在1号盒子中球的数量， Y：放在2号盒子中球的数量，判断X，Y的独立性。

解：列出所有可能的情况



X \ Y	0	1	2	$p_{i\cdot}$
0	1/9	2/9	1/9	4/9
1	2/9	2/9	0	4/9
2	1/9	0	0	1/9
$p_{\cdot j}$	4/9	4/9	1/9	1

显然$P\{X = 1, Y = 2\} \neq P\{X = 1\}P\{Y = 2\}$

于是不独立

4.3 相互独立的连续型随机变量

基本的独立性定义是从分布函数出发定义的

例如从上一课例7知，如二维随机变量(X,Y)的联合分布函数为

$$F(x,y)=\frac{1}{\pi^2}\left(\frac{\pi}{2}+\operatorname{atan}\frac{x}{2}\right)\left(\frac{\pi}{2}+\operatorname{atan}\frac{y}{5}\right),$$

$$(-\infty<x,y<+\infty)$$

则其边缘分布分别为：

$$F_X(x)=\frac{1}{\pi}\left(\frac{\pi}{2}+\operatorname{atan}\frac{x}{2}\right),F_Y(y)=\frac{1}{\pi}\left(\frac{\pi}{2}+\operatorname{atan}\frac{y}{5}\right)$$

即：$F(x,y)$=$F_X(x)F_Y(y)$，即相互独立

独立性与密度函数

由独立性

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

注意到

$$F(x,y)=\int_{-\infty}^{y}\int_{-\infty}^{x}f(x,y)dxdy$$

$$F_X(x)F_Y(y)=\int_{-\infty}^{y}\int_{-\infty}^{x}f_X(x)f_Y(y)dx\,dy$$

于是独立性等价于几乎处处满足

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

定义：设(X,Y)是二维连续型随机变量，

其联合概率密度函数为$f(x,y)$，$f_X(x)$和 $f_Y(y)$分别为(X,Y)的分布函数及边缘分布函数，

若对几乎所有x,y有

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

则称X，Y是相互独立的随机变量。

注意这里对于所有连续点都成立，不成立的点的面积为零

例8(考察独立性)：给定二维随机变量(X,Y)的密度函数为

$$f(x,y)=\begin{cases} x^2+\frac{1}{3}xy & 0\le x\le 1, 0\le y\le 2 \\ 0 & \text{其他} \end{cases}$$

试判断X与Y是否相互独立

解：当$0\le x\le 1$时，

$$f_X(x)=\int_{-\infty}^{\infty}f(x,y)dy=\int_0^2(x^2+\frac{1}{3}xy)dy$$

$$=2x^2+\frac{2}{3}x$$

解(续)： X的边缘分布为

$$f_X(x) = \begin{cases} 2x^2 + \frac{2}{3}x & 0 \le x \le 1 \\ 0 & \text{其他} \end{cases}$$

类似地，当$0 \le x \le 2$时，

$$f_Y(y) = \int_{-\infty}^{\infty} f(x,y)dx = \int_0^1 (x^2 + \frac{1}{3}xy)dx = \frac{1}{3} + \frac{1}{6}y$$

Y的边缘分布为

$$f_Y(y) = \begin{cases} \frac{1}{3} + \frac{1}{6}y & 0 \le y \le 2 \\ 0 & \text{其他} \end{cases}$$

注意到当$0 \le x \le 1, 0 \le y \le 2$时

$$f_X(x)f_Y(y) = \frac{1}{3}x^2y + \frac{1}{9}xy + \frac{2}{3}x^2 + \frac{2}{9}x \neq f(x,y)$$

于是随机变量X与Y不独立

例9 (正态随机变量的独立性). 设二维随机变量(X,Y)服从参数为$(\mu_1,\mu_2,\sigma_1^2,\sigma_2^2,\rho)$的正态分布，即密度函数为

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

由本章例10知X,Y的边缘密度函数分别为：

$$f_X(x)=\frac{1}{\sqrt{2\pi}\sigma_1}e^{-\frac{(x-\mu_1)^2}{2\sigma_1^2}}\quad(-\infty\leq x\leq\infty)$$

$$f_Y(y)=\frac{1}{\sqrt{2\pi}\sigma_2}e^{-\frac{(y-\mu_2)^2}{2\sigma_2^2}}\quad(-\infty\leq y\leq\infty)$$

于是有：

(1) 当$\rho = 0$时，$f(x, y) = f_X(x) f_Y(y)$，于是X与Y相互独立

(2) 当$\rho \neq 0$时，$f(x, y) \neq f_X(x) f_Y(y)$，于是X与Y不独立

后面会看到ρ是这两个随机变量相关性的度量——相关系数

例10. 设X与Y是两个相互独立的随机变量，其概率密度分别为：

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

其中$\lambda>0,\mu>0$。引入随机变量

$$Z=\begin{cases}1 & X\le Y\\ 0 & X>Y\end{cases}$$

求(1) 条件概率密度$f_{X|Y}(x|y)$;

(2) Z分布律和分布函数

解: (1)由独立性

$$f_{X|Y}(x|y)=f_X(x)=\begin{cases}\lambda e^{-\lambda x} & x>0\\ 0 & x\le 0\end{cases}$$

解(续)：(2) 由相互独立，联合分布密度函数为

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

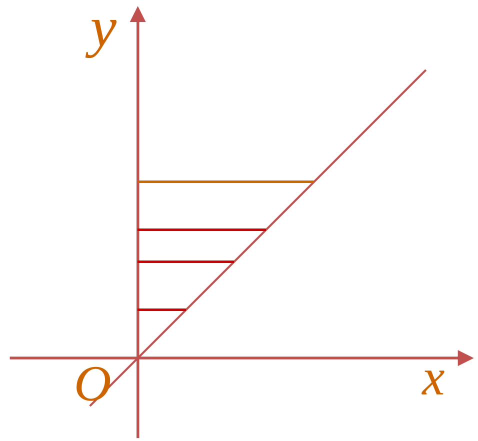
于是Z的分布律为：

$$P\{Z=1\} = P\{X \le Y\} = \lambda\mu \iint\limits_{0<x\le y} e^{-\lambda x-\mu y}dxdy$$

$$= \lambda\mu \int_0^{+\infty} e^{-\lambda x} \int_x^{+\infty} e^{-\mu y} dy\, dx = \lambda \int_0^{+\infty} e^{-(\lambda+\mu)x} dx$$

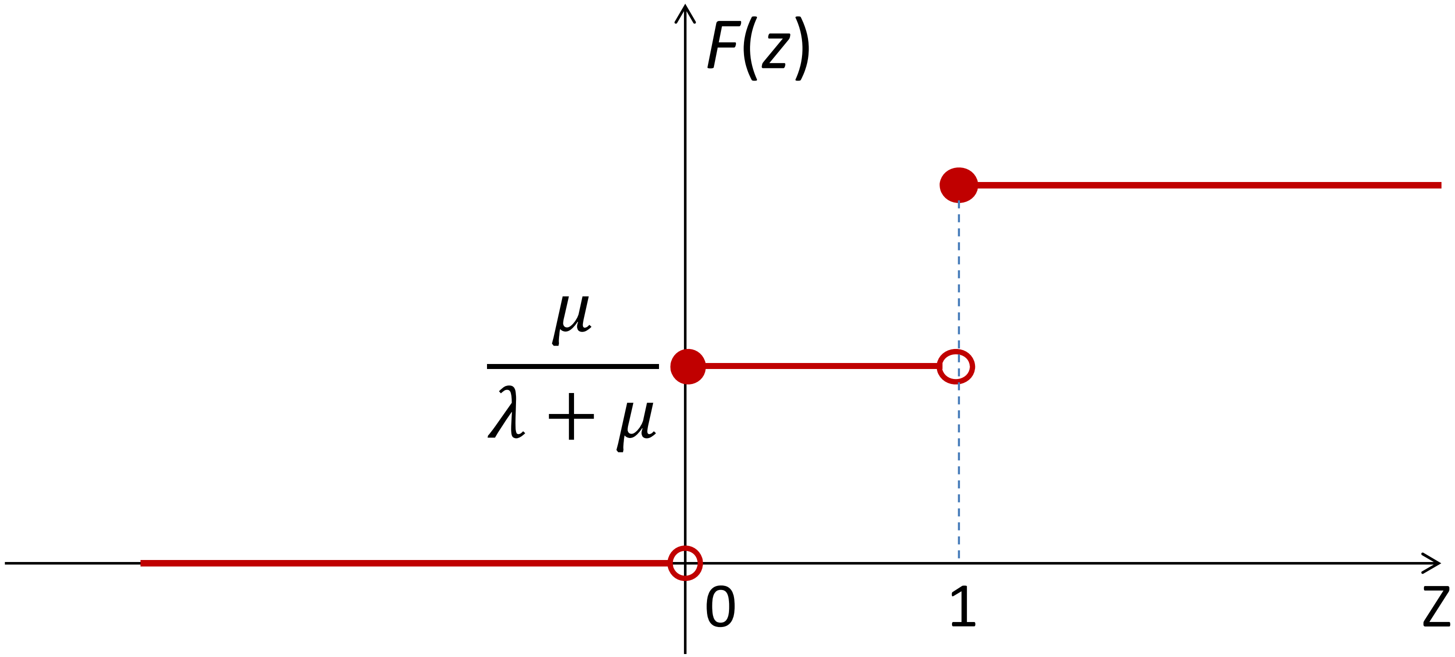
$$= \frac{\lambda}{\lambda+\mu}$$

$$P\{Z=0\} = 1 - P\{Z=1\} = \frac{\mu}{\lambda+\mu}$$



相应Z的分布函数为

$$F(z)=\begin{cases}0 & z<0\\ \dfrac{\mu}{\lambda+\mu} & 0\le z<1\\ 1 & z\ge 1\end{cases}$$



例 (约会迟到问题)：甲乙两人经常固定约定某地相会，假定甲的到达时间在10点到12点之间均匀分布，两人各自独立。乙为了保证在大多数情况(75%)下不迟到，并且同样是在某个固定时间段到达，考虑乙希望尽可能短的等待时间，试给出乙到达时间的最迟分布。

解：设甲到达时间为X，乙到达时间为Y，乙到达的最早时间不必早于10点，最晚时间为t(t>10)。

于是随机变量的边缘分布为

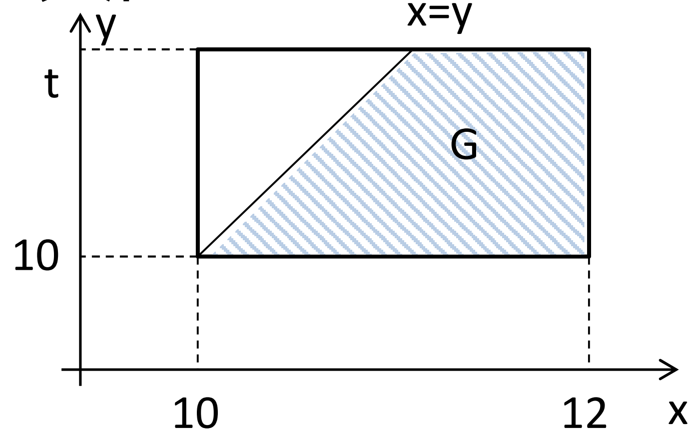
$$f_X(x)=\begin{cases}\frac{1}{2} & x\in[10,12]\\ 0 & \text{其他}\end{cases},\ f_Y(y)=\begin{cases}\frac{1}{t-10} & y\in[10,t]\\ 0 & \text{其他}\end{cases}$$

解(续)：联合概率密度为$f(x,y)=f_X(x)f_Y(y)$

$$f(x,y)=\begin{cases}\dfrac{1}{2(t-10)} & x\in[10,12],y\in[10,t]\\ 0 & \text{其他}\end{cases}$$

乙不迟于甲到达对应图中区域G

因此落在G中的概率为$P(y\le x)$



$$P(y\le x)=\iint_G \frac{1}{2(t-10)}dxdy$$

$$=\frac{1}{2(t-10)}\left[\frac{(t-10)^2}{2}+(12-t)(t-10)\right]$$

$$\ge 0.75$$

解得$t\le 11$，因此乙可以选择在[10，11]之间到达

4.4 n维随机变量的独立性

定义：设$(X_1, X_2, \ldots, X_n)$是n维随机变量，其联合分布函数为$F(x_1, x_2, \ldots, x_n)$，随机变量X_i的分布函数为$F_{X_i}(x_i)$，对所有的$x_1, x_2, \ldots, x_n$有

$$F(x_1, x_2, \ldots, x_n) = \prod_{i=1}^{n} F_{X_i}(x_i)$$

则称$X_1, X_2, \ldots, X_n$是相互独立的。

注意：

1. 若$X_1, X_2, \ldots, X_n$相互独立，则其中任意k个随机变量$X_{i_1}, X_{i_2}, \ldots, X_{i_n}$也相互独立
2. 若X, Y 独立，$f(x), g(y)$ 是连续函数，则$f(X), g(Y)$ 也独立

定义(分组独立性)

随机向量$(X_1, X_2, \ldots, X_m)$, $(Y_1, Y_2, \ldots, Y_n)$和$(X_1, X_2, \ldots, X_m, Y_1, Y_2, \ldots, Y_n)$的分布函数分别为$F_1(x_1, x_2, \ldots, x_m)$，$F_2(y_1, y_2, \ldots, y_n)$和$F(x_1, x_2, \ldots, x_m, y_1, y_2, \ldots, y_n)$，

并且对所有的$x_1, x_2, \ldots, x_m; y_1, y_2, \ldots, y_n$有

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

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

定理：设$(X_1, X_2, \dots, X_m)$与$(Y_1, Y_2, \dots, Y_n)$相互独立,
则$X_i(i = 1,2, \dots, m)$ 和$Y_j(j = 1,2, \dots, n)$ 相互独立。
若h,g是连续函数，则$h(X_1, X_2, \dots, X_m)$与
$g(Y_1, Y_2, \dots, Y_n)$相互独立

注意： $h(X)$与$g(Y)$相互独立， X与Y未必相互独立

作业

概率论与数理统计

- pp. 84-86, #11, #13, #15, #17