西北工業大學

NORTHWESTERN POLYTECHNICAL UNIVERSITY

概率论与数理统计

徐爽

西北工业大学理学院概率统计教研室

第一节 多维随机变量及其分布(3)

- 八、随机变量的独立性
- 九、条件分布

八、随机变量的独立性

回忆：两事件 A,B独立的定义：

若 $P(AB)=P(A)P(B)$,

则称事件 A,B独立.

定义2.6 设X,Y是两个随机变量,若对任意实数 x,y,

事件$\{X\le x\},\{Y\le y\}$是相互独立的,即

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

则称 X,Y是相互独立的.

用分布函数表示，即是

设X,Y是两个的随机变量，若对任意的x, y有

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

则称X,Y相互独立.

它表明，两个随机变量(简记为*r.v*)相互独立时，它们的**联合分布函数**等于两个**边缘分布函数的乘积**.

对于离散型随机变量,上述独立性定义等价于:

对(X,Y)所有可能取值(x_i, y_j),有

$$P(X = x_i, Y = y_j) = P(X = x_i)P(Y = y_j)$$

即 $p_{ij} = p_{i\cdot}p_{\cdot j}$

则称X,Y相互独立.

注 如果 X 和 Y 相互独立,那么 它们的连续函数

$f(X)$ 和 $g(Y)$也相互独立.（证明略）

对于连续型随机变量,上述定义等价于:

对于任意的x, y有

$$p(x,y)=p_X(x)p_Y(y)$$

成立,则称X,Y相互独立.

其中$p(x,y)$是 X,Y的联合概率密度;

$p_X(x)$和$p_Y(y)$分别是 X和Y的边缘概率密度.

例1 已知 (X,Y) 的分布律为

(X,Y)	$(1,1)$	$(1,2)$	$(1,3)$	$(2,1)$	$(2,2)$	$(2,3)$
p_{ij}	$\frac{1}{6}$	$\frac{1}{9}$	$\frac{1}{18}$	$\frac{1}{3}$	α	β

(1) 求 α 与 β 应满足的条件;

(2) 若 X 与 Y 相互独立,求 α 与 β 的值.

解 将 (X,Y) 的分布律改写为

$X \backslash Y$	1	2	3	$p_{i\bullet} = P\{X = x_i\}$
1	$\frac{1}{6}$	$\frac{1}{9}$	$\frac{1}{18}$	$\frac{1}{3}$
2	$\frac{1}{3}$	α	β	$\frac{1}{3}+\alpha+\beta$
$p_{\bullet j} = P\{Y = y_j\}$	$\frac{1}{2}$	$\frac{1}{9}+\alpha$	$\frac{1}{18}+\beta$	$\frac{2}{3}+\alpha+\beta$

(1)由分布律的性质知 $\alpha \geq 0, \beta \geq 0, \frac{2}{3}+\alpha+\beta=1,$

故α与β应满足的条件是：$\alpha \geq 0, \beta \geq 0$ 且 $\alpha+\beta=\frac{1}{3}.$

(2) 因为 X 与 Y 相互独立, 所以有

$$p_{ij} = p_{i\bullet} \cdot p_{\bullet j}, \quad (i = 1,2;\ j = 1,2,3)$$

特别有

$$p_{12} = p_{1\bullet} \cdot p_{\bullet 2} \Rightarrow \frac{1}{9} = \frac{1}{3}\left(\frac{1}{9} + \alpha\right) \Rightarrow \alpha = \frac{2}{9},$$

又 $\alpha + \beta = \dfrac{1}{3}$, 得 $\beta = \dfrac{1}{9}$.

例2 假如每个人晚上是否去操场跑步都是独立的，因此去操场跑步的总人数N服从参数为λ的泊松分布。记来跑步的人是女生的概率为p，则是男生的概率为$1-p$。记来跑步的女生数量为N_1，男生数量为N_2。

求(1) N_1和N_2的联合分布律 $P\{N_1 = n, N_2 = m\}$。

(2) N_1和N_2是否独立？

解 根据全概率公式可知，

$$P\{N_1 = n, N_2 = m\} = \sum_{i=0}^{\infty} P\{N_1 = n, N_2 = m | N = i\} P\{N = i\}$$

当$i \neq m+n$时，

$$P\{N_1=n, N_2=m \mid N=i\}=0$$

所以，

$$P\{N_1=n, N_2=m\}=P\{N_1=n, N_2=m|N=n+m\}e^{-\lambda}\frac{\lambda^{n+m}}{(n+m)!}$$

第一项的含义：如果来了$n+m$个人跑步，n个是女生、m个是男生的概率。

这其实是做了$n+m$次伯努利试验，每次是女生的概率为p、男生的概率为$1-p$。

$$P\{N_1=n, N_2=m\}=\binom{n+m}{n}p^n(1-p)^m e^{-\lambda}\frac{\lambda^{n+m}}{(n+m)!}$$

整理可得，联合分布律：

$$
\begin{aligned}
P\{N_1=n,N_2=m\} &= \binom{n+m}{n} p^n(1-p)^m e^{-\lambda}\frac{\lambda^{n+m}}{(n+m)!}\\
&= \frac{(n+m)!}{n!m!}p^n(1-p)^m e^{-\lambda p}e^{-\lambda(1-p)}\frac{\lambda^n\lambda^m}{(n+m)!}\\
&= e^{-\lambda p}\frac{(\lambda p)^n}{n!}e^{-\lambda(1-p)}\frac{(\lambda(1-p))^m}{m!}
\end{aligned}
$$

前两项只和n有关，后两项只和m有关，直觉上两者是独立的。但我们仍需进一步验证！

$$
\begin{aligned}
P\{N_1=n\} &= \sum_{m=0}^{\infty} P\{N_1=n,\,N_2=m\}\\
&= e^{-\lambda p}\frac{(\lambda p)^n}{n!}\sum_{m=0}^{\infty} e^{-\lambda(1-p)}\frac{(\lambda(1-p))^m}{m!} = e^{-\lambda p}\frac{(\lambda p)^n}{n!}
\end{aligned}
$$

类似地，

$$P\{N_2 = m\} = e^{-\lambda(1-p)} \frac{(\lambda(1-p))^m}{m!}$$

因此是独立的！

例2 设随机变量(X,Y)的联合密度函数为

$$p(x,y)=\begin{cases}3x, & 0<x<1, 0<y<x,\\ 0, & \text{其他}.\end{cases}$$

试求(1)边际密度函数$p_X(x)$和$p_Y(y)$;

(2) X与Y是否独立?

解 (1)因为当$0<x<1$时，有

$$p_X(x)=\int_{-\infty}^{+\infty}p(x,y)\mathrm{d}y=\int_0^x 3x\mathrm{d}y=3x^2,$$

所以X的边际密度函数为

$$p_X(x)=\begin{cases}3x^2, & 0<x<1,\\ 0, & \text{其他}.\end{cases}$$

这是贝塔分布$Be(3,1)$.

又因为当$0<y<1$时，有

$$p_Y(y)=\int_{-\infty}^{+\infty}p(x,y)\mathrm{d}x=\int_y^1 3x\mathrm{d}x=\frac{3}{2}x^2\Big|_y^1=\frac{3}{2}(1-y^2),$$

所以Y的边际密度函数为

$$p_Y(y)=\begin{cases}\dfrac{3}{2}(1-y^2), & 0<y<1,\\ 0, & \text{其他}.\end{cases}$$

(2)因为$P(x,y)\neq P_X(x)P_Y(y)$,所以X与Y不独立.

例3 设$(X,Y)\sim N(\mu_1,\mu_2,\sigma_1^{\ 2},\sigma_2^{\ 2},\rho)$

证明X与Y相互独立的充要条件是$\rho=0$.

证 上节课已经给出的X和Y边缘概率密度分别为

$$p_X(x)=\frac{1}{\sqrt{2\pi}\sigma_1}\mathrm{e}^{-\frac{(x-\mu_1)^2}{2\sigma_1^2}},\quad -\infty<x<+\infty.$$

$$p_Y(y)=\frac{1}{\sqrt{2\pi}\sigma_2}\mathrm{e}^{-\frac{(y-\mu_2)^2}{2\sigma_2^2}},\quad -\infty<y<+\infty.$$

(1) 充分性

(注：$\exp x = e^x$)

若 $\rho = 0$，则

$$p(x,y) = \frac{1}{2\pi\sigma_1\sigma_2}\exp\{-\frac{1}{2}[\frac{(x-\mu_1)^2}{\sigma_1^2} + \frac{(y-\mu_2)^2}{\sigma_2^2}]\}$$

$$= p_X(x)p_Y(y)$$

说明 X 与 Y 相互独立.

（2）必要性,若 X 与 Y 相互独立,则 $\forall x, y$

$$p(x,y) = p_X(x)p_Y(y)$$

$$\therefore \frac{1}{2\pi\sigma_1\sigma_2\sqrt{1-\rho^2}}\exp\{-\frac{1}{2(1-\rho^2)}[\frac{(x-\mu_1)^2}{\sigma_1^2}-$$

$$2\rho\frac{(x-\mu_1)(y-\mu_2)}{\sigma_1\sigma_2}+\frac{(y-\mu_2)^2}{\sigma_2^2}]\}$$

$$=\frac{1}{\sqrt{2\pi}\sigma_1}\exp\{-\frac{(x-\mu_1)^2}{2\sigma_1^2}\}\times\frac{1}{\sqrt{2\pi}\sigma_2}\exp\{-\frac{(y-\mu_2)^2}{2\sigma_2^2}\}$$

令$x=\mu_1, y=\mu_2$,则上式变为

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

从而推出$\rho=0.$

例4 设 X_1, X_2 独立，都服从标准正态分布 $N(0,1)$。把点(X_1, X_2)的极坐标记为(R,Θ), $0 \le R < \infty, 0 \le \Theta < 2\pi$. 求证：$R$ 和Θ独立

证 X_1, X_2概率密度函数分别为

$$p_{X_1}(x_1) = \frac{1}{\sqrt{2\pi}} e^{-\frac{x_1^2}{2}} \qquad p_{X_2}(x_2) = \frac{1}{\sqrt{2\pi}} e^{-\frac{x_2^2}{2}}$$

(X_1, X_2)的联合概率密度函数为

$$p(x_1, x_2) = \frac{1}{2\pi} e^{-\frac{(x_1^2 + x_2^2)}{2}}$$

接下来考虑，(R,Θ)的联合分布函数

$$F(r_0,\theta_0)=P\{R\le r_0,\Theta\le\theta_0\}$$

$$=P\{0\le R\le r_0,0\le\Theta\le\theta_0\}$$

记区域$B=\{0\le R\le r_0,0\le\Theta\le\theta_0\}$

$$F(r_0,\theta_0)=P(B)=\iint_B\frac{1}{2\pi}e^{-\frac{(x_1^2+x_2^2)}{2}}\mathrm{d}x_1\mathrm{d}x_2$$

极坐标变换：$\begin{cases}x=r\cos\theta\\y=r\sin\theta\end{cases}$ $\quad\mathrm{d}x\mathrm{d}y=r\mathrm{d}r\mathrm{d}\theta$

$$F(r_0,\theta_0)=P(B)=\iint\limits_B \frac{1}{2\pi}e^{-\frac{(x_1^2+x_2^2)}{2}}\mathrm{d}x_1\mathrm{d}x_2$$

$$=(2\pi)^{-1}\int_0^{\theta_0}\int_0^{r_0}\mathrm{e}^{-r^2/2}\,r\mathrm{d}r\mathrm{d}\theta$$

所以，(R,Θ)的密度函数为

$$p(r,\theta)=\begin{cases}(2\pi)^{-1}\mathrm{e}^{-r^2/2}, 0\le r<\infty, 0\le\theta<2\pi\\ 0, \quad else\end{cases}$$

分别计算R和Θ的边缘密度函数

$$p_R(r)=\begin{cases}\mathrm{e}^{-r^2/2}r\,, r\geq 0\\ 0\,,\ r<0\end{cases}$$

$$p_\Theta(\theta)=\begin{cases}1/2\pi, 0\leq\theta<2\pi\\ 0\,,\text{其他}\end{cases}$$

所以独立。

九、条件分布

问题的提出: 从遗传学的角度看, 父亲的身高 X 会影响儿子的身高Y.这里父亲的身高 X与儿子的身高Y都是随机变量,都有自己的分布.那么两者之间关系如何呢?

一般的处理方法是将父亲的身高X 固定在一特定值X_0处,考察儿子身高Y 的分布情况.

这就是我们要讲的条件分布.

事件的条件概率：$P(A|B)=P(AB)/P(B)$

1.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

$$P\{X=x_i|Y=y_j\}=\frac{P\{X=x_i;Y=y_j\}}{P\{Y=y_j\}}=\frac{p_{ij}}{p_{\cdot j}}$$

其中$i=1,2,\cdots$

为事件$\{Y=y_j\}$发生的条件下随机变量 X的条件分布律.

同样，对于固定的 i,如果$P\{X = x_i\} > 0$,则称

$$P\{Y = y_j \mid X = x_i\} = \frac{P\{X = x_i, Y = y_j\}}{P\{X = x_i\}} = \frac{p_{ij}}{p_{i\cdot}}$$

其中 $j = 1,2,\cdots$

为在事件$\{X = x_i\}$发生的条件下随机变量Y的

条件分布律.

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

例4 在一汽车工厂中,一辆汽车有两道工序是由机器人完成的,其一是紧固3只螺栓,其二是焊接2处焊点.以X表示由机器人紧固螺栓紧固不良的数目,以Y表示由机器人焊接的不良焊点的数目.根据累计的资料知(X,Y)有如下分布律:

$Y \backslash X$	0	1	2	3	$P\{Y=j\}$
0	0.840	0.030	0.020	0.010	0.900
1	0.060	0.010	0.008	0.002	0.080
2	0.010	0.005	0.004	0.001	0.020
$P\{X=i\}$	0.910	0.045	0.032	0.013	1.000

(1) 求在 $X=1$ 的条件下,Y 的条件分布律;

(2) 求在 $Y=0$ 的条件下,X 的条件分布律.

解 由上述分布律的表格可得

$$P\{Y=0 \mid X=1\}=\frac{P\{X=1,Y=0\}}{P\{X=1\}}=\frac{0.030}{0.045},$$

$Y \backslash X$	0	1	2	3	$P\{Y=j\}$
0	0.840	0.030	0.020	0.010	0.900
1	0.060	0.010	0.008	0.002	0.080
2	0.010	0.005	0.004	0.001	0.020
$P\{X=i\}$	0.910	0.045	0.032	0.013	1.000

即在 $X=1$ 的条件下,Y 的条件分布律为

$Y=k$	0	1	2
$P\{Y=k \mid X=1\}$	$\frac{6}{9}$	$\frac{2}{9}$	$\frac{1}{9}$

同理可得在 $Y=0$ 的条件下,X 的条件分布律为

$X=k$	0	1	2	3
$P\{X=k \mid Y=0\}$	$\frac{84}{90}$	$\frac{3}{90}$	$\frac{2}{90}$	$\frac{1}{90}$

例 5 设随机变量 (X,Y) 的联合概率密度为

$$p(x,y)=\begin{cases}3x, & 0\le x<1, 0\le y<x,\\ 0, & \text{其它}.\end{cases}$$

求 $P\{Y\le \frac{1}{8} \mid X=\frac{1}{4}\}$? **判断下面的解法是否正确?**

解 因为 $P\{X=\frac{1}{4}\}=0$,

所以 $P\{Y\le \frac{1}{8} \mid X=\frac{1}{4}\}=\dfrac{P\{X=\frac{1}{4}, Y\le\frac{1}{8}\}}{P\{X=\frac{1}{4}\}}$ 不存在.

2. 连续型随机变量的条件分布

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

$$P\{X \le x \mid Y = y\} = \lim_{\Delta y \to 0} P\{X \le x \mid y - \Delta y < Y \le y + \Delta y\}$$

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

$$= \lim_{\Delta y \to 0} \frac{\int_{y-\Delta y}^{y+\Delta y} \int_{-\infty}^{x} p(u,v)\,du\,dv}{\int_{y-\Delta y}^{y+\Delta y} \int_{-\infty}^{+\infty} p(u,v)\,du\,dv}$$

$$f(y) = \int_0^{y+\Delta y} \int_{-\infty}^{x} p(u,v)\,du\,dv \qquad g(y) = \int_0^{y+\Delta y} \int_{-\infty}^{+\infty} p(u,v)\,du\,dv$$

$$f(y)=\int_0^y \boxed{\int_{-\infty}^{x} p(u,v)du}\,dv \qquad g(y)=\int_0^y \boxed{\int_{-\infty}^{+\infty} p(u,v)du}\,dv$$

$$P\{X\le x\,|\,Y=y\}=\lim_{\Delta y\to 0}\frac{\frac{1}{2\Delta y}[f(y+\Delta y)-f(y-\Delta y)]}{\frac{1}{2\Delta y}[g(y+\Delta y)-g(y-\Delta y)]}$$

$$=\frac{f'(y)}{g'(y)} \qquad =\frac{\int_{-\infty}^{x} p(u,y)du}{p_Y(y)}$$

$$=\int_{-\infty}^{x}\boxed{\frac{p(u,y)}{p_Y(y)}}\,du$$

2. 连续型随机变量的条件分布

定义 设二维随机变量 (X,Y) 的联合密度函数为 $p(x,y)$, (X,Y) 关于 Y 的边缘密度函数为 $p_Y(y)$. 若对于固定的 y, $p_Y(y)>0$, 则称 $\dfrac{p(x,y)}{p_Y(y)}$ 为在 $Y=y$ 的条件下 X 的条件密度函数, 记为

$$p_{X|Y}(x|y)=\frac{p(x,y)}{p_Y(y)}.$$

条件密度函数 = $\dfrac{\text{联合密度函数}}{\text{边缘密度函数}}$

称$\int_{-\infty}^{x} p_{X|Y}(x \mid y)\mathrm{d}x = \int_{-\infty}^{x} \frac{p(x,y)}{p_Y(y)}\mathrm{d}x$为在

$Y = y$的条件下,X的条件分布函数,记为

$$P\{X \le x \mid Y = y\} \text{ 或 } F_{X|Y}(x \mid y),$$

即$F_{X|Y}(x \mid y) = P\{X \le x \mid Y = y\} = \int_{-\infty}^{x} \frac{p(x,y)}{p_Y(y)}\mathrm{d}x$

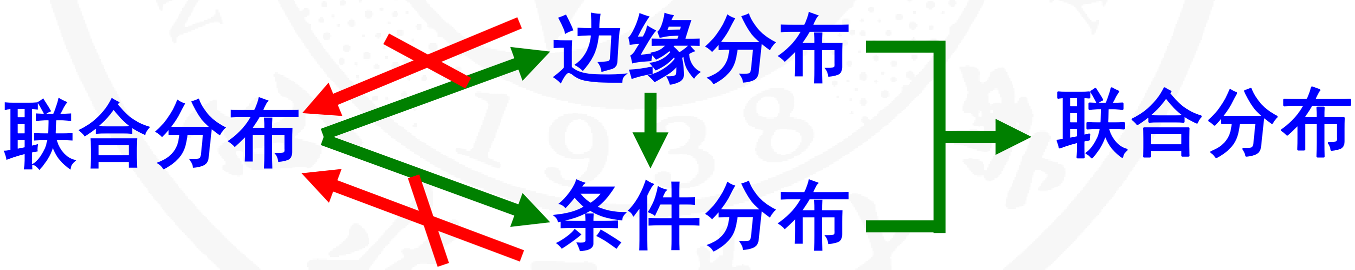
同理可定义在$X = x$的条件下,Y的条件分布函数,

$$F_{Y|X}(y \mid x) = P\{Y \le y \mid X = x\} = \int_{-\infty}^{y} \frac{p(x,y)}{p_X(x)}\mathrm{d}y.$$

说明

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

联合分布、边缘分布、条件分布的关系如下



例 5 设随机变量 (X,Y) 的联合概率密度为

$$p(x,y)=\begin{cases}3x, & 0\le x<1,0\le y<x,\\ 0, & \text{其它}.\end{cases}$$

求 $P\{Y\le \frac{1}{8}\big| X=\frac{1}{4}\}$? **判断下面的解法是否正确?**

解 因为 $P\{X=\frac{1}{4}\}=0,$

所以 $P\{Y\le \frac{1}{8}\big| X=\frac{1}{4}\}=\dfrac{P\{X=\frac{1}{4},Y\le\frac{1}{8}\}}{P\{X=\frac{1}{4}\}}$ 不存在.

上述解法不正确,正确解法应该为

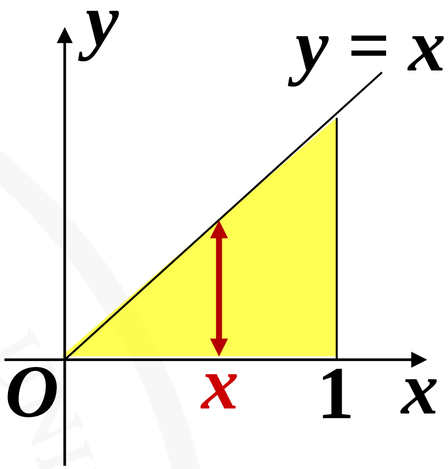
$$p_X(x)=\int_{-\infty}^{+\infty}p(x,y)\,\mathrm{d}\,y$$

$$=\begin{cases}\int_0^x 3x\,\mathrm{d}\,y, & 0\le x<1,\\ 0, & \text{其它}.\end{cases}$$

$$=\begin{cases}3x^2, & 0\le x<1,\\ 0, & \text{其它}.\end{cases}$$

因此 $p_{Y|X}(y|x)=\dfrac{p(x,y)}{p_X(x)}$

$$=\begin{cases}3x/3x^2=1/x, & 0\le y<x,\\ 0, & \text{其它}.\end{cases}$$



于是 $P\{Y\le\frac{1}{8}\,\big|\,X=\frac{1}{4}\}=\int_{-\infty}^{\frac{1}{8}}p_{Y|X}(y|\frac{1}{4})\,\mathrm{d}y$

$$=\int_{0}^{\frac{1}{8}}4\,\mathrm{d}y=\frac{1}{2}.$$

例6 设 $(X,Y)\sim N(\mu_1,\mu_2,\sigma_1^2,\sigma_2^2,\rho)$，求条件概率密度 $p(x\mid y),p(y\mid x)$.

解 由于上节已经求出了X和Y的边缘概率密度，

所以对于一切 $x,y\in(-\infty,+\infty)$，有

$$p(y\mid x)=\frac{p(x,y)}{p_X(x)}=\frac{1}{\sqrt{2\pi}\sigma_2\sqrt{1-\rho^2}}$$

$$\times\exp\left\{-\frac{\left[y-\mu_2-\rho\sigma_2(x-\mu_1)/\sigma_1\right]^2}{2\sigma_2^2(1-\rho^2)}\right\}$$

$$\sim N(\mu_2+\rho\sigma_2(x-\mu_1)/\sigma_1,\sqrt{1-\rho^2}\sigma_2)$$

$$p(x\,|\,y)=\frac{1}{\sqrt{2\pi}\sigma_1\sqrt{1-\rho^2}}$$

$$\times\exp\left\{-\frac{\left[x-\mu_1-\rho\sigma_1(y-\mu_2)/\sigma_2\right]^2}{2\sigma_1^2(1-\rho^2)}\right\}$$

$$\sim N\left(\mu_1+\rho\sigma_1(y-\mu_2)/\sigma_2,\sqrt{1-\rho^2}\sigma_1\right)$$

本题说明了对于二元正态分布,其条件分布仍为正态分布.

例7 设数 X 在区间(0,1)上等可能地随机取值，当观察到 $X = x\ (0 < x < 1)$时，数 Y在区间$(x,1)$上等可能地随机取值，求 Y 的密度函数.

解 由题意知，X 的密度函数为

$$p_X(x) = \begin{cases} 1, & 0 < x < 1, \\ 0, & \text{其它}. \end{cases}$$

对于任意给定的值$x(0<x<1)$,在$X=x$的条件下,Y的条件密度函数为

$$p_{Y|X}(y|x)=\begin{cases}\dfrac{1}{1-x}, & 0<x<y<1,\\ 0, & \text{其它}.\end{cases}$$

因此, X 和 Y的联合密度函数为

$$p(x,y)=p_{Y|X}(y|x)p_X(x)$$

$$=\begin{cases}\dfrac{1}{1-x}, & 0<x<y<1,\\ 0, & \text{其它}.\end{cases}$$

根据边缘概率密度的公式可得 Y 的边缘密度函数为

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

$$=\begin{cases}\displaystyle\int_0^y \frac{1}{1-x}\mathrm{d}x=-\ln(1-y), & 0<y<1,\\ 0, & \text{其它}.\end{cases}$$

内容小结

1. 独立性

(1) 若随机变量 (X,Y) 的联合分布函数为 $F(x,y)$,

边缘分布函数分别为 $F_X(x)$ 和 $F_Y(y)$, 则有 X

和 Y 相互独立 $\Leftrightarrow F(x,y)=F_X(x)F_Y(y).$

(2) 若离散型随机变量 (X,Y) 的联合分布律为

$$p\{X=i,Y=j\}=p_{ij},\ \ i,j=1,2,\cdots$$

则 X 与 Y 相互独立 $\Leftrightarrow$

$$P\{X=x_i,Y=y_j\}=P\{X=x_i\}P\{Y=y_j\}$$

(3) 设连续型随机变量(X,Y)的联合概率密度函数为

$p(x,y)$,边缘概率密度分别为 $p_X(x)$, $p_Y(y)$,则有

X与Y相互独立 $\Leftrightarrow p(x,y)=p_X(x)p_Y(y)$

(4) 若 X和 Y相互独立，则 $f(X)$与 $g(y)$也相互独立.

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

(1) 若离散型随机变量(X,Y)的联合分布律为

$$p\{X=i,Y=j\}=p_{ij},\ \ i,j=1,2,\cdots$$

在给定$Y=y_j$的条件下随机变量X的条件分布律为

$$P\{X=x_i|Y=y_j\}=\frac{P\{X=x_i;Y=y_j\}}{P\{Y=y_j\}}=\frac{p_{ij}}{p_{\cdot j}}$$

在给定$X=x_i$的条件下随机变量Y的条件分布律为

$$P\{Y=y_j|X=x_i\}=\frac{P\{X=x_i,Y=y_j\}}{P\{X=x_i\}}=\frac{p_{ij}}{p_{i\cdot}}$$

其中$i,j=1,2,\cdots$

(2) 设 (X,Y)是二维连续型随机变量,则有

在$Y=y$的条件下随机变量 X的条件密度函数为

$$p_{X|Y}(x\,|\,y)=p(x,y)/p_Y(y)$$

在$X=x$的条件下随机变量Y的条件密度函数为

$$p_{X|Y}(y\,|\,x)=p(x,y)/p_X(x)$$

$$F_{X|Y}(x\,|\,y)=\int_{-\infty}^{x}p_{X|Y}(x\,|\,y)\mathrm{d}x$$

$$F_{Y|X}(y\,|\,x)=\int_{-\infty}^{y}p_{Y|X}(y\,|\,x)\mathrm{d}y$$

2-2 多维随机变量

Thank You !

备用题

例2-1 设随机变量 X 和 Y 相互独立,并且 X 服从 $N(a,\sigma^2)$,Y 在 $[-b,b]$ 上服从均匀分布,求 (X,Y) 的联合概率密度.

解 由于 X 与 Y 相互独立,

所以 $$p(x,y)=p_X(x)\cdot p_Y(y)$$

又 $$p_X(x)=\frac{1}{\sqrt{2\pi}\sigma}e^{-\frac{(x-a)^2}{2\sigma^2}},\ -\infty<x<\infty;$$

$$p_Y(y)=\begin{cases}\dfrac{1}{2b}, & -b\le y\le b,\\ 0, & \text{其它}.\end{cases}$$

得 $$p(x,y)=\frac{1}{2b}\cdot\frac{1}{\sqrt{2\pi}\sigma}\mathrm{e}^{-\frac{(x-a)^2}{2\sigma^2}},$$

其中 $-\infty<x<\infty,\quad -b\le y\le b.$

当$|y|>b$时，$p(x,y)=0.$

例2-2 设(X,Y)的概率密度为

$$p(x,y)=\begin{cases} xe^{-(x+y)}, & x>0, y>0, \\ 0, & \text{其它}. \end{cases}$$

问 X与 Y是否独立?

解 $p_X(x)=\int_0^{+\infty} x\mathrm{e}^{-(x+y)}\mathrm{d}y = x\mathrm{e}^{-x},$

$$p_Y(y)=\int_0^{+\infty} x\mathrm{e}^{-(x+y)}\mathrm{d}x = \mathrm{e}^{-y}\int_0^{+\infty} x\mathrm{e}^{-x}\mathrm{d}x = \mathrm{e}^{-y},$$

即 $p_X(x)=\begin{cases} x\mathrm{e}^{-x}, & x>0, \\ 0, & \text{其它}. \end{cases}$ $p_Y(y)=\begin{cases} \mathrm{e}^{-y}, & y>0, \\ 0, & \text{其它}. \end{cases}$

$\therefore p(x,y)=p_X(x)p_Y(y)$ 所以 X 与Y独立.

例2-3 某电子仪器由两部分构成，其寿命(单位：千小时)X与Y的联合分布函数为：

$$F(x,y)=\begin{cases}1-e^{-0.5x}-e^{-0.5y}+e^{-0.5(x+y)}, & x\geq 0, y\geq 0,\\ 0, & \text{其他}.\end{cases}$$

问: (1) X与Y是否独立？

(2) 两部件的寿命都超过100小时的概率？

解 (1)当$x\geq 0, y\geq 0$时,

由$F(x,y)=1-e^{-0.5x}-e^{-0.5y}+e^{-0.5(x+y)}$

$$得F_X(x)=\lim_{y\to+\infty}F(x,y)$$

$$=\lim_{y\to+\infty}(1-\mathrm{e}^{-0.5x}-\mathrm{e}^{-0.5y}+\mathrm{e}^{-0.5(x+y)})$$

$$=1-\mathrm{e}^{-0.5x},x\geq 0.$$

$$F_Y(y)=\lim_{x\to+\infty}F(x,y)$$

$$=\lim_{x\to+\infty}(1-\mathrm{e}^{-0.5x}-\mathrm{e}^{-0.5y}+\mathrm{e}^{-0.5(x+y)})$$

$$=1-\mathrm{e}^{-0.5y},y\geq 0.$$

$$由F_X(x)\cdot F_Y(y)=(1-\mathrm{e}^{-0.5x})(1-\mathrm{e}^{-0.5y})$$

$$=1-\mathrm{e}^{-0.5x}-\mathrm{e}^{-0.5y}+\mathrm{e}^{-0.5(x+y)}=F(x,y)$$

可知X与Y相互独立.

$$(2)P(X>0.1,Y>0.1)=P(X>0.1)P(Y>0.1)$$

$$=[1-P(X\le 0.1)][1-P(Y\le 0.1)]$$

$$=[1-F_X(0.1)][1-F_Y(0.1)]$$

$$=\mathrm{e}^{-0.05}\cdot\mathrm{e}^{-0.05}=\mathrm{e}^{-0.1}$$

例3-1 设随机变量(X,Y)的两个分量X和Y相互独立，且服从同一分布.试证

$$P\{X \le Y\} = 1/2.$$

证 因为X,Y独立，所以 $p(x,y) = p_X(x)p_Y(y)$.

$$P\{X \le Y\} = \iint_{x \le y} p(x,y)\mathrm{d}x\mathrm{d}y = \iint_{x \le y} p_X(x)p_Y(y)\mathrm{d}x\mathrm{d}y$$

$$= \int_{-\infty}^{+\infty} [p_Y(y)\int_{-\infty}^{y} p_X(x)\mathrm{d}x]\mathrm{d}y$$

$$= \int_{-\infty}^{+\infty} [p_Y(y)F_Y(y)]\mathrm{d}y$$

$$= \int_{-\infty}^{+\infty} F_Y(y)\mathrm{d}F_Y(y) = F^2(y)/2\Big|_{-\infty}^{+\infty} = 1/2.$$

也可以利用对称性来证，因为X，Y独立同分布，所以有

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

而$P\{X \le Y\} + P\{Y \le X\} = 1$，故$P\{X \le Y\} = 1/2.$

例5 设(X,Y)在椭圆$\dfrac{x^2}{a^2}+\dfrac{y^2}{b^2}\le 1$上服从均匀分布，求条件分布密度 $p(x\mid y)$.

解 由题设知

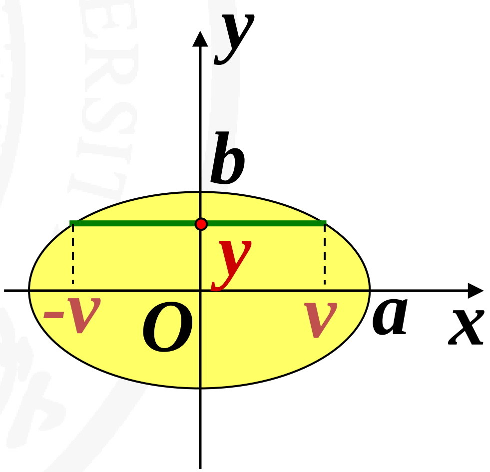
$$p(x,y)=\begin{cases}\dfrac{1}{\pi ab}, & \dfrac{x^2}{a^2}+\dfrac{y^2}{b^2}\le 1,\\ 0, & \dfrac{x^2}{a^2}+\dfrac{y^2}{b^2}> 1.\end{cases}$$

由联合概率密度我们可以得到关于Y的边缘概率密度.

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

$$=\begin{cases}\displaystyle\int_{-v}^{v}\frac{1}{\pi ab}\mathrm{d}x=\frac{2}{\pi b}\sqrt{1-y^2/b^2}, & |y|\le b,\\ 0, & |y|>b.\end{cases}$$

其中 $v=a\sqrt{1-\dfrac{y^2}{b^2}}$

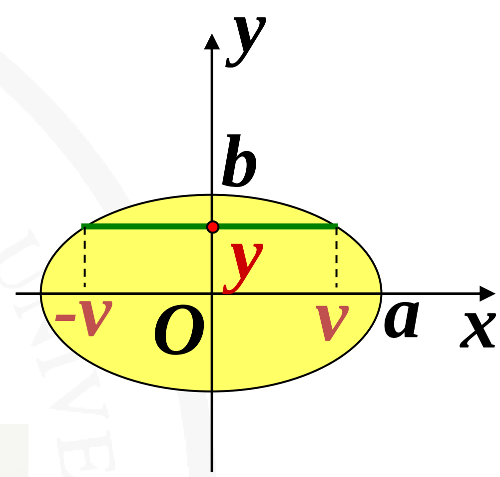


从而对于 $y\in(-b,b)$,有

$$p_{X|Y}(x\,|\,y)=\frac{p(x,y)}{p_Y(y)}$$

$$p_{X|Y}(x\,|\,y)=\frac{p(x,y)}{p_Y(y)}$$

$$=\begin{cases}\dfrac{1}{2a\sqrt{1-y^2/b^2}}, & |x|\le a\sqrt{1-y^2/b^2},\\ 0, & |x|>a\sqrt{1-y^2/b^2}\end{cases}$$



由上式可以看出，在 $(Y=y)$ 的条件下，X 在

区间$[-a\sqrt{1-y^2/b^2},a\sqrt{1-y^2/b^2}]$上服从均匀分布.