Chpt.3 Multi-Dimensional Random Variables and Their Distributions

第三章 多维随机变量及其分布

上节回顾

- 随机向量（多维随机变量）
- 联合分布
- 边缘分布

二元离散型随机变量	二元连续型随机变量
(X,Y)联合分布律 $P(X=x_i,Y=y_j)=p_{ij},i,j=1,2,\ldots$	(X,Y)联合概率密度 $f(x,y),\quad (x,y)\in D$
X的边际分布律 $P(X=x_i)=\sum_{j=1}^{\infty}p_{ij}=p_{i\cdot},i=1,2,\ldots$	X的边际概率密度 $f_X(x)=\int_{-\infty}^{\infty}f(x,y)dy$

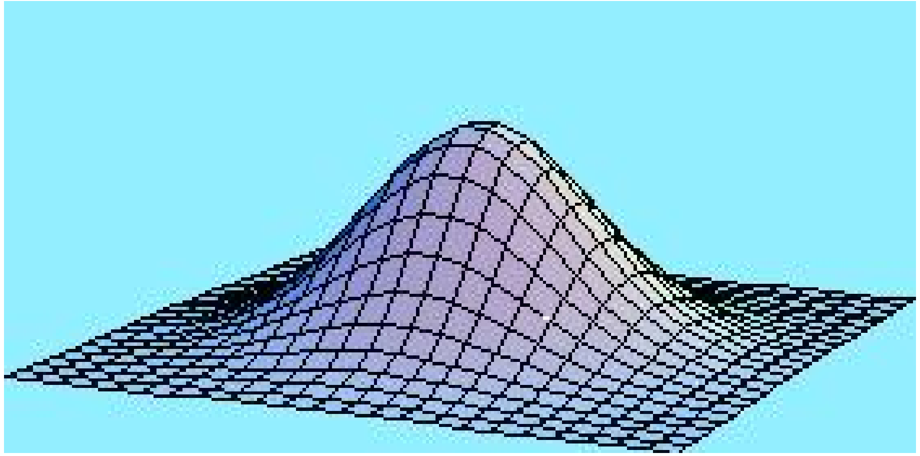
Example 设二维随机变量（X,Y）的分布概率密度是

$$f(x,y)=\frac{1}{2\pi\sqrt{1-\rho^2}\sigma_1\sigma_2}$$

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

称(X,Y)为服从参数 $\mu_1,\mu_2,\sigma_1,\sigma_2,\rho$ 的二元正态分布，

记作 $(X,Y)\sim N(\mu_1,\mu_2,\sigma_1^2,\sigma_2^2,\rho)$



求其边缘概率密度。

[解]：在上式的指数上对y配方，

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

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

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

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

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

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

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

$$f_x(x)=\int_{-\infty}^{\infty} f(x,y)dy$$

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

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

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

$\Downarrow$ $$t=\frac{1}{\sqrt{1-\rho^2}}\left(\frac{y-\mu_2}{\sigma_2}-\rho\frac{x-\mu_1}{\sigma_1}\right),\text{则}dy=\sigma_2\sqrt{1-\rho^2}dt$$

$$=\frac{1}{2\pi\sigma_1}\exp\left[-\frac{(x-\mu_1)^2}{2\sigma_1^2}\right]\int_{-\infty}^{\infty}e^{-\frac{t^2}{2}}dt$$

$$=\frac{1}{\sqrt{2\pi}\sigma_1}\exp\left[-\frac{(x-\mu_1)^2}{2\sigma_1^2}\right]$$

由标准正态分布可知 $\int_{-\infty}^{\infty}e^{-\frac{t^2}{2}}dt=\sqrt{2\pi}$

$$f_x(x)=\int_{-\infty}^{\infty}f(x,y)dy=\frac{1}{\sqrt{2\pi}\sigma_1}\exp\left\{-\frac{(x-\mu_1)^2}{2\sigma_1^2}\right\}$$

$$f_y(y)=\int_{-\infty}^{\infty} f(x,y)dx=\frac{1}{\sqrt{2\pi}\sigma_2}\exp\left\{-\frac{(y-\mu_2)^2}{2\sigma_2^2}\right\}$$

$X \sim N(\mu_1, \sigma_1^2)$

$Y \sim N(\mu_2, \sigma_2^2)$

上述结论是说，二元正态分布的边际分布仍是正态分布，并且与ρ无关.

但反过来不正确，即若(X,Y)的边际分布都是正态分布，其联合分布却未必是二元正态分布.

思考： **设 (X,Y)的联合密度为**

$$f(x,y)=\frac{1}{2\pi}e^{-\frac{x^2+y^2}{2}}(1+\sin x\sin y) \qquad -\infty<x,y<+\infty$$

求边际分布.

设 (X,Y)的联合密度为

$$f(x,y)=\frac{1}{2\pi}e^{-\frac{x^2+y^2}{2}}(1+\sin x\sin y)\qquad -\infty<x,y<+\infty$$

求边际分布.

$$f_x(x)=\int_{-\infty}^{\infty}f(x,y)dy$$

$$=\frac{1}{\sqrt{2\pi}}e^{-\frac{x^2}{2}}\int_{-\infty}^{\infty}\frac{1}{\sqrt{2\pi}}e^{-\frac{y^2}{2}}dy+\frac{1}{\sqrt{2\pi}}e^{-\frac{x^2}{2}}\sin x\int_{-\infty}^{\infty}\frac{1}{\sqrt{2\pi}}\boxed{\sin ye^{-\frac{y^2}{2}}}dy$$

奇函数在对称区间上的积分为0

$$=\frac{1}{\sqrt{2\pi}}e^{-x^2/2}$$

$$f_y(y)=\frac{1}{\sqrt{2\pi}}e^{-y^2/2}$$

因此X,Y都服从标准正态分布N(0,1)，但联合分布不是正态的.

问 题

1. 联合分布可以完全确定边缘分布。反过来，边缘分布能确定联合分布吗?
2. 边缘分布不能反映联合分布，也就是说边缘分布少了点什么。那么，什么可以与边缘分布一起来确定联合分布呢?
3. 能否在一些特殊条件下，边缘分布可以完全确定联合分布呢?

（以下可以仅仅从事件的角度而非随机变量的角度看）

$$F(x_1,x_2)=P\{\{X_1 \le x_1\}\cap\{X_2\le x_2\}\}$$
$$=P\{X_2\le x_2 \mid X_1\le x_1\}P\{X_1\le x_1\}$$
$$=P\{X_2\le x_2 \mid X_1\le x_1\}F_{X_1}(x_1)$$

当 $\{X_1\le x_1\}$独立于$\{X_2\le x_2\}$

$$P\{X_2\le x_2 \mid X_1\le x_1\}=P\{X_2\le x_2\}$$
$$=F_{X_2}(x_2)$$

$$F(x_1,x_2)=F_{X_1}(x_1)F_{X_2}(x_2)$$

此时，边缘分布函数完全界定了联合分布函数。

3.3 相互独立的随机变量-2维

[定义] 随机变量 X_1, X_2， $F(x,y), F_X(x), F_Y(y)$ 分别是联合分布函数和边缘分布函数，如果对于任意的 (x_1, x_2) 满足

$$F(x_1, x_2) = F_{X_1}(x_1) F_{X_2}(x_2)$$

则称随机变量 X_1, X_2 是相互独立的。

[等价定义] 当 (X_1, X_2) 是连续型随机向量时，X_1, X_2 联合概率密度为 $f(x_1, x_2)$，边缘概率密度分别是 $f_{X_1}(x_1), f_{X_2}(x_2)$，则 X_1, X_2 独立的条件等价于：

对于任意的 (x_1, x_2)满足

$$f(x_1, x_2) = f_{X_1}(x_1) f_{X_2}(x_2)$$

3.3 相互独立的随机变量-2维

Proof： 对 $F(x_1,x_2)=F_{X_1}(x_1)F_{X_2}(x_2)$ 进行展开

$$\int_{-\infty}^{x_1}\int_{-\infty}^{x_2} f(u,v)dudv=\int_{-\infty}^{x_1} f_{X_1}(u)du\int_{-\infty}^{x_2} f_{X_2}(v)dv=\int_{-\infty}^{x_1}\int_{-\infty}^{x_2} f_{X_1}(u)f_{X_2}(v)dudv$$

由 x_1,x_2 的任意性知道，必须有 $f(u,v)=f_{X_1}(u)f_{X_2}(v)$

[例]：上例中的二维正态随机分布 $N(\mu_1,\mu_2,\sigma_1,\sigma_2,\rho)$ 的边缘分布概率密度是 $f_X(x),f_Y(y)$，如果要使得 $f(x,y)=f_X(x)f_Y(y)$，必有 $\rho=0$。反之，很容易看到成立。即

二维正态随机变量 (X,Y) 中 X,Y 相互独立的充分必要条件是 $\rho=0$.

[等价定义] 当 (X, Y) 是离散型随机向量时，设(X, Y)的联合分布列为

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

X,Y对应的边缘分布列为

$$P(X=x_i)=p_{i\cdot} \quad P(Y=y_j)=p_{\cdot j}$$

则 X, Y 独立的条件等价于:

对于任意的(x_i, y_j) 都满足 $p_{ij} = p_{i\cdot} * p_{\cdot j}$

Example (pp.74) 一负责人到办公室的时间是8~12点，服从均匀分布X~U(8,12), 其秘书到达办公室的时间是7~9点，服从均匀分布Y~U(7,9)，二人到达的时间X,Y独立。

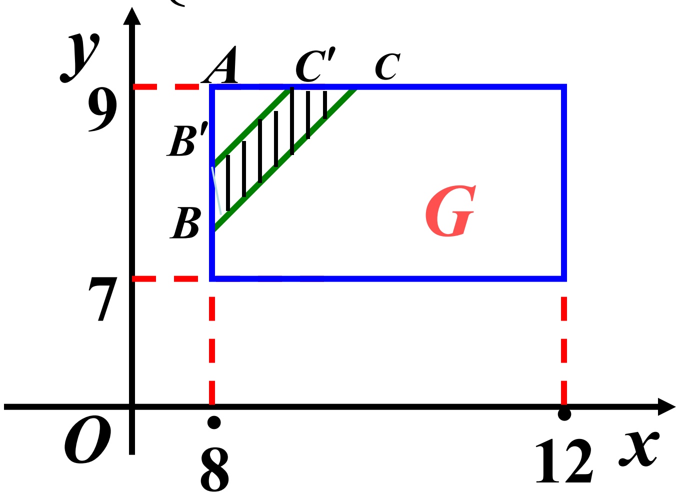
求：他们达到办公室时间相差不超过5分钟的概率。

$$G=\{(x,y)\mid 8<x<12, 7<y<9\}$$

$$f_x(x)=\begin{cases}1/4 & 8<x<12\\ 0 & elsewhere\end{cases}\qquad f_y(y)=\begin{cases}1/2 & 7<y<9\\ 0 & elsewhere\end{cases}$$

$$f(x,y)=f_x(x)f_y(y)$$

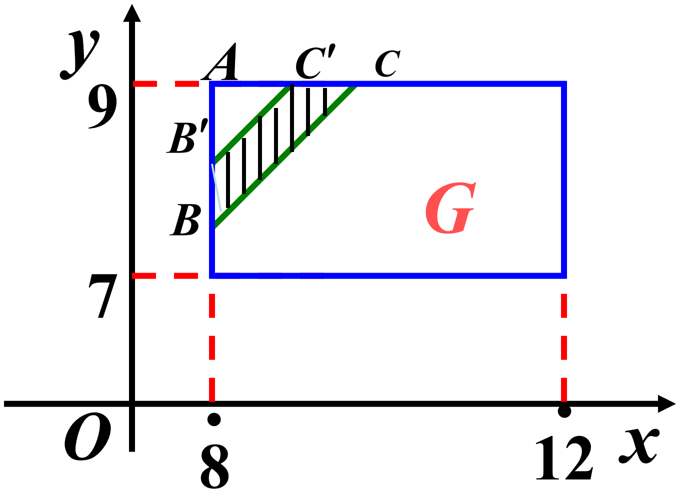
$$=\begin{cases}1/8 & (x,y)\in G\\ 0 & elsewhere\end{cases}$$



$$P\{|X-Y|<\frac{1}{12}\}=\iint\limits_{\substack{|x-y|<\frac{1}{12}\\(x,y)\in G}} f(x,y)dxdy$$

$$=\frac{1}{8}\{|x-y|\le\frac{1}{12}\text{的面积},(x,y)\in G\}$$

$$=1/48$$



3.3 相互独立的随机变量-n维

[定义] 关于多维随机变量 $X_1, X_2, \ldots, X_n$ 其联合分布 $F(x_1, x_2, \ldots\ldots, x_m)$，边缘分布为 $F_{X_i}(x_i)(i=1,2,\cdots,n)$ ，如果对**任意的** $x_1, x_2, \ldots, x_n$

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

则 $X_1, X_2, \ldots, X_n$ 是**相互独立的**。

[定义] 假设 $X_1, X_2, \ldots, X_m$ 有联合分布 $F_X(x_1,x_2,\ldots\ldots,x_m)$

$Y_1, Y_2, \ldots, Y_n$ 有联合分布 $F_Y(y_1,y_2,\ldots\ldots,y_n)$

$X_1, X_2, \ldots, X_m$， $Y_1, Y_2, \ldots, Y_n$ 的联合分布为 $F(x_1,\ldots, x_m, y_1, \ldots, y_n)$

如果 $F(x_1,\ldots, x_m, y_1, \ldots, y_n) \equiv F_X(x_1,x_2,\ldots\ldots,x_m)\, F_Y(y_1,y_2,\ldots\ldots,y_n)$

则称 $X_1, X_2, \ldots, X_m$ 与 $Y_1, Y_2, \ldots, Y_n$ 是**独立的**。

这两种独立性有什么区别与联系?

3.3 相互独立的随机变量

[定理] 设 $X_1,X_2,\cdots,X_m$ 与 $Y_1,Y_2,\cdots,Y_n$ 相互独立，则：

（1） $X_i(i=1,2,\cdots,m)$ 与 $Y_j(j=1,2,\cdots,n)$ 相互独立；

（2）如果 h,g 是连续函数，则 $H=h(X_1,X_2,\cdots,X_m)$ 与 $G=g(Y_1,Y_2,\cdots,Y_n)$ 独立。

$$F_{ij}(x_i,y_j)=F(\infty,\cdots,\infty,x_i,\infty,\cdots,\infty,\infty,\cdots\infty,y_j,\infty,\cdots\infty)$$

$$=F_x(\infty,\cdots,\infty,x_i,\infty,\cdots,\infty)F_y(\infty,\cdots\infty,y_j,\infty,\cdots\infty)$$

$$=F_{x_i}(x_i)F_{y_j}(y_j)$$

$$F(h,g)=P\{H\le h,G\le g\}$$

$$=P\{h(X_1,X_2,\cdots,X_m)\le h,g(Y_1,Y_2,\cdots,Y_n)\le g\}$$

$$=P\{h(X_1,X_2,\cdots,X_m)\le h\}P\{g(Y_1,Y_2,\cdots,Y_n)\le g\}$$

$$=F_H(h)F_G(g)$$

3.4 条件分布(Conditional Distribution)

当多个随机变量不满足独立性，他们之间存在着关联性，一变量的存在影响到另一个变量的存在。可类似条件概率定义条件分布。

3.4.1 离散随机变量的条件分布

设二维离散随机变量 (X, Y) 的联合分布列为

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

设X,Y对应的边缘分布列为 $P(X=x_i)=p_{i\cdot}$ ，$P(Y=y_j)=p_{\cdot j}$，且

$$p_{i\bullet}>0, \qquad p_{\bullet j}>0$$

考虑在事件 $\{Y=y_j\}$ 发生的条件下 $\{X=x_i\}$ 发生的概率，即事件 $\{X=x_i|Y=y_j\}$ 的概率。

3.4.1 离散随机变量的条件分布

$$P\left(X = x_i \mid Y = y_j\right) = \frac{P\left(X = x_i, Y = y_j\right)}{P(Y = y_j)} = \frac{p_{ij}}{p_{\bullet j}} \qquad (i = 1, 2, \cdots)$$

容易知道上述满足：

$$P(X = x_i \mid Y = y_j) \geq 0$$

$$\sum_{i=1}^{\infty} P(X = x_i \mid Y = y_j) = 1$$

3.4.1 离散随机变量的条件分布

[定义] 对于固定的j，若$P(Y=y_j)>0$，则称

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

为在$Y=y_j$条件下随机变量 X 的条件分布律。

对于固定的i，若$P(X=x_i)>0$，则称

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

为在$X=x_i$条件下随机变量 Y 的条件分布律。

此处我们没有使用定义分布函数的方法，而是直接给出了**分布律**

Example 例1的联合分布、边际分布与条件分布.

Y \ X	0	1	$p_{\cdot j}$
0	$\frac{3}{5}\times\frac{3}{5}$	$\frac{2}{5}\times\frac{3}{5}$	$\frac{3}{5}$
1	$\frac{3}{5}\times\frac{2}{5}$	$\frac{2}{5}\times\frac{2}{5}$	$\frac{2}{5}$
$p_{i\cdot}$	$\frac{3}{5}$	$\frac{2}{5}$	

Y \ X	0	1	$p_{\cdot j}$
0	$\frac{3}{5}\times\frac{2}{4}$	$\frac{2}{5}\times\frac{3}{4}$	$\frac{3}{5}$
1	$\frac{3}{5}\times\frac{2}{4}$	$\frac{2}{5}\times\frac{1}{4}$	$\frac{2}{5}$
$p_{i\cdot}$	$\frac{3}{5}$	$\frac{2}{5}$	

在$Y=0$条件下X的**条件分布律**为

$P(X=0|Y=0)=\frac{3}{5}$　　$P(X=1|Y=0)=\frac{2}{5}$　　--有放回

$P(X=0|Y=0)=\frac{2}{4}$　　$P(X=1|Y=0)=\frac{2}{4}$　　--无放回

Example (pp.69，例2) 一射击手单发击中目标的概率为p (p>0), 射击到击中两次目标为止，设 X 为首次射中目标所进行的射击次数，Y 为总的射击次数，求（X,Y）的联合分布与条件分布。

解：$$P(X=m,\ Y=n)=p^2(1-p)^{n-2}$$

此处为化简结果

联合分布

$$n=2,3,\cdots;\quad m=1,2,\cdots,n-1$$

$$P(X=m)=\sum_{n>m}P(X=m,Y=n)$$

边缘分布

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

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

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

注意：

我们事实上可以很简单地计算得到P{X＝m}

$$P(Y=n)=\sum_{m<n}P(X=m,Y=n)$$

边缘分布

$$=\sum_{m=1}^{n-1}p^2(1-p)^{n-2}=(n-1)p^2(1-p)^{n-2} \qquad (\mathrm{n}\geq 2)$$

$$P(X=m\mid Y=n)=\frac{P(X=m,Y=n)}{P(Y=n)}$$

条件分布

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

对求得的条件分布能给出直观解释吗？

$$P(Y=n\mid X=m)=\frac{P(X=m,Y=n)}{P(X=m)}$$

条件分布

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

以上二式子都可以进行直接的解释：

- 第一式 $P(X=m|Y=n)=\dfrac{1}{n-1}$

  在知道第n次为射中的话，前n－1次里射中一次，第m次射中的概率为 1/(n-1)；

- 第二式 $P(Y=n|X=m)=p(1-p)^{n-m-1}$

  在第m次射击时为第一次中，随后的第n次射击才射中的概率为 $p(1-p)^{n-m-1}$ (m+1,m+2, …,n-1为不中，n为中)

3.4.2 连续型随机变量的条件分布

注意：是在$Y=y$的条件下，而不是在$Y\le y$的条件下。

对于二维连续型随机变量，我们考察$P(X \le x|Y = y)$

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

对任何y, P(Y=y)=0，故不能采用上述公式计算，改用极限的方式定义：

$$P(X \le x|Y = y) = \lim_{\varepsilon \to 0+} P\left(X \le x|y - \varepsilon \le Y \le y + \varepsilon\right)$$

3.4.2 连续型随机变量的条件分布

$$P(X \le x | Y = y) = \lim_{\varepsilon \to 0+} P(X \le x | y - \varepsilon \le Y \le y + \varepsilon)$$

$$= \lim_{\varepsilon \to 0+} \frac{P(X \le x, y - \varepsilon \le Y \le y + \varepsilon)}{P(y - \varepsilon \le Y \le y + \varepsilon)}$$

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

$$= \lim_{\varepsilon \to 0+} \frac{\dfrac{F(x, y + \varepsilon) - F(x, y - \varepsilon)}{2\varepsilon}}{\dfrac{F_Y(y + \varepsilon) - F_Y(y - \varepsilon)}{2\varepsilon}}$$

$$= \frac{\dfrac{\partial F(x, y)}{\partial y}}{\dfrac{dF_Y(y)}{dy}}$$

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

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

定义Y=y下的条件密度函数为

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

定义X=x下的条件密度函数为

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

此处与之前不同，之前都是分布函数“开路”，求导得到密度函数，此处直接定义密度函数

条件分布函数与条件密度函数的关系

[定义] 条件分布函数为

$$F_{X|Y}(x|y)=\int_{-\infty}^{x} f_{X|Y}(x|y)\,\mathrm{d}\,x=\int_{-\infty}^{x}\frac{f(x,y)}{f_Y(y)}dx$$

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

联合分布、边缘分布、条件分布的关系如下

Example 设 $(X,Y)\sim N(\mu_1,\mu_2,\sigma_1^2,\sigma_2^2,\rho)$，求条件密度 $f_{Y|X}(y\,|\,x)$.

解 $F_X(x)$ 为正态分布 $N(\mu_1,\sigma_1^2)$

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

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

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

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

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

$$=\frac{1}{\sqrt{2\pi}\sqrt{1-\rho^2}\sigma_2}\exp\left\{-\frac{1}{2(1-\rho^2)}\right.$$

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

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

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

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

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

$$\boxed{\begin{aligned}\sigma &= \sqrt{1-\rho^2}\sigma_2\\ \mu &= \mu_2+\frac{\sigma_2}{\sigma_1}\rho(x-\mu_1)\end{aligned}}$$

$$=\frac{1}{\sqrt{2\pi}\sigma}e^{-\frac{(y-\mu)^2}{2\sigma^2}}$$

3.4.2 连续型随机变量的条件分布

前面求得条件密度

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

它表明：

已知X= x条件下，二维正态分布的对于Y的条件分布是正态分布

$$N\left(\mu_2+\frac{\sigma_2}{\sigma_1}\rho(x-\mu_1),\quad \left(\sqrt{1-\rho^2}\sigma_2\right)^2\right)$$

其中第一参数 $\mu_2+\frac{\sigma_2}{\sigma_1}\rho(x-\mu_1)$是$x$的线性函数，第二参数与$x$无关.

此结论在一些统计问题中很重要.

思考题

设有一件工作需要甲乙两人接力完成，完成时间不能超过30分钟。设甲先干了X分钟，再由乙完成，加起来共用Y分钟。若

$X \sim U(0, 30)$，在$X=x$条件下，$Y \sim U(x, 30)$。

(1) 求(X, Y)的联合概率密度及条件概率密度 $f_{X|Y}(x|y)$;

(2) 当已知两人共花了25分钟完成工作时，求甲的工作时间不超过10分钟的概率。