0.1 条件分布与条件期望

0.1.1 条件分布

首先，我们再看几个条件分布的例子。

例 0.1 设 $(X,Y)\sim N(\mu_1,\mu_2,\sigma_1,\sigma_2,\rho)$，且 $\sigma_1,\sigma_2>0,-1<\rho<1$。求 $f_{X|Y}(x|y)$。

解： 令 $\Sigma=\begin{pmatrix}\sigma_1^2 & \rho\sigma_1\sigma_2\\ \rho\sigma_1\sigma_2 & \sigma_2^2\end{pmatrix}$

$$
\begin{aligned}
f_{X|Y}(x|y)&=\frac{f_{XY}(x,y)}{f_Y(y)}\\
&=\frac{\frac{1}{2\pi|\Sigma|^{1/2}}\exp\left\{-\frac{1}{2}\begin{pmatrix}x-\mu_1\\ y-\mu_2\end{pmatrix}^T\Sigma^{-1}\begin{pmatrix}x-\mu_1\\ y-\mu_2\end{pmatrix}\right\}}{\frac{1}{\sqrt{2\pi\sigma_2^2}}\exp\left\{-\frac{(y-\mu_2)^2}{2\sigma_2^2}\right\}}\\
&=\frac{1}{\sqrt{2\pi\sigma^2}}\exp\left\{-\frac{(x-\mu)^2}{2\sigma^2}\right\}
\end{aligned}
$$

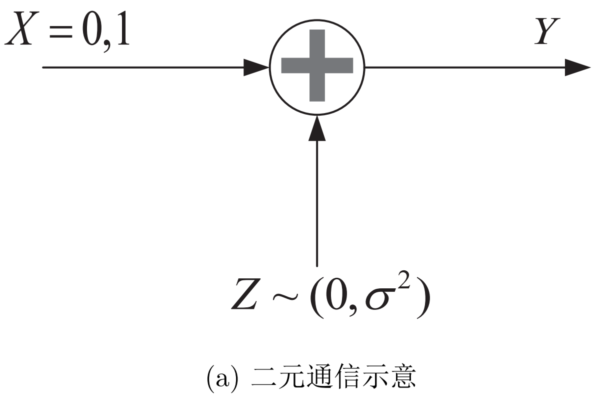
其中，

$$
\begin{aligned}
&\mu{=}E(X)-\mathrm{cov}(X,Y)\,\mathrm{cov}(Y)^{-1}[y-E(Y)]=\mu_1-\rho\tfrac{\sigma_1}{\sigma_2}(y-\mu_2)\\
&\sigma=\mathrm{cov}(X)-\mathrm{cov}(X,Y)\,\mathrm{cov}(Y)^{-1}\mathrm{cov}(X,Y)=\sigma_1^2-\rho^2\sigma_1^2
\end{aligned}
$$

即有 $E[X|Y=y]=\mu,\mathrm{var}[X|Y=y]=\sigma$

例 0.2 二元通信模型

设 X 服从等概 $0-1$ 分布，$Z\sim N(0,\sigma^2)$,X 与 Z 相互独立，设 $Y=X+Z$，求 $f_{X|Y}(x|y)$。



(a) 二元通信示意

解： 根据条件密度的定义

$$
\begin{aligned}
&f_{X|Y}(x|y)=\tfrac{f_{XY}(x,y)}{\sum\limits_u f_{XY}(u,y)}\\
&=\tfrac{f_X(x)f_{Y|X}(y|x)}{\sum\limits_u f_X(u)f_{Y|X}(y|u)}
\end{aligned}
$$

因此，

$$f_{X|Y}\left(x=0|y\right)=\frac{0.5\times\frac{1}{\sqrt{2\pi\sigma^2}}\exp\left\{-\frac{y^2}{2\sigma^2}\right\}}{0.5\times\frac{1}{\sqrt{2\pi\sigma^2}}\exp\left\{-\frac{y^2}{2\sigma^2}\right\}+0.5\times\frac{1}{\sqrt{2\pi\sigma^2}}\exp\left\{-\frac{(y-1)^2}{2\sigma^2}\right\}}=\frac{1}{1+\exp\left\{\frac{2y-1}{2\sigma^2}\right\}}$$

$$f_{X|Y}\left(x=1|y\right)=1-f_{X|Y}\left(x=0|y\right)=\frac{\exp\left\{\frac{2y-1}{2\sigma^2}\right\}}{1+\exp\left\{\frac{2y-1}{2\sigma^2}\right\}}$$

例 0.3 设 $Y=X+Z$，其联合概率密度 $f_{XZ}(x,z)$，求 $f_{Y|X}(y|x)$。

解：

$$f_{Y|X}\left(y|x\right)=f_{Z|X}\left(z|x\right)\Big|_{z=y-x}\overset{X,Z indepentent}{=}f_Z\left(z\right)|_{z=y-x}$$

分析一：如下来分析：$X=x$ 时，$Y=x+Z$，$f_Y\left(y\right)=\left[f_Z\left(z\right)\left|\frac{dz}{dy}\right|\right]_{z=y-x}=f_Z\left(z\right)|_{z=y-x}$ 特别值得注意的是，上述分析都应带着条件才可以，$f_{Y|X}\left(y|x\right)=\left[f_{Z|X}\left(z|x\right)\left|\frac{dz}{dy}\right|\right]_{z=y-x}=f_{Z|X}\left(z|x\right)|_{z=y-x}$。

分析二：事实上，根据定义我们可以得到更清晰的认识，

$$\begin{aligned}F_{Y|X}\left(y|x\right)&\triangleq\lim_{\varepsilon\to0^+}\frac{P\left\{Y\le y,x\le X\le x+\varepsilon\right\}}{P\left\{x\le X\le x+\varepsilon\right\}}\\&=\lim_{\varepsilon\to0^+}\frac{P\left\{X+Z\le y,x\le X\le x+\varepsilon\right\}}{P\left\{x\le X\le x+\varepsilon\right\}}\\&=\lim_{\varepsilon\to0^+}\frac{\frac{1}{\varepsilon}\int_x^{x+\varepsilon}\int_{-\infty}^{y-u}f_{XZ}\left(u,z\right)dzdu}{\frac{1}{\varepsilon}\int_x^{x+\varepsilon}f_X\left(u\right)du}\\&=\frac{\int_{-\infty}^{y-x}f_{XZ}\left(x,z\right)dz}{f_X\left(x\right)}\\&=\int_{-\infty}^{y-x}\frac{f_{XZ}\left(x,z\right)}{f_X\left(x\right)}dz\end{aligned}$$

因此，求导可得概率密度函数，

$$f_{Y|X}\left(y|x\right)=f_{Z|X}\left(z|x\right)|_{z=y-x}$$

独立性的再认识：

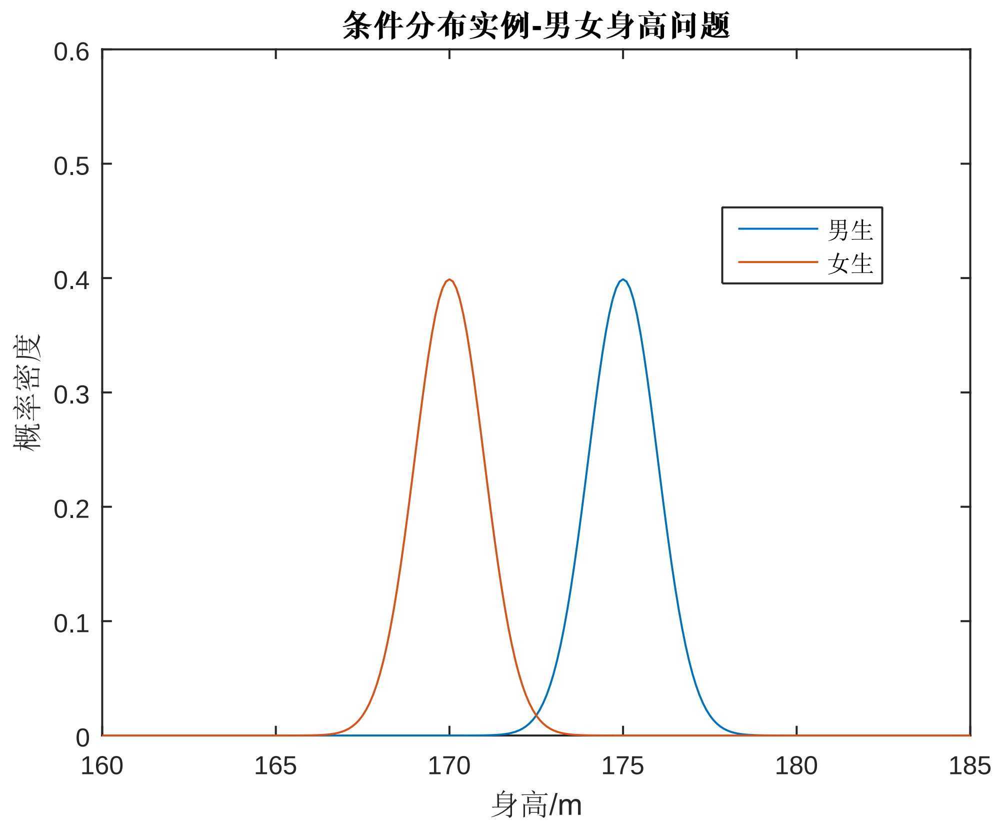
$$\begin{aligned}X_1,X_2,...,X_n\text{相互独立}&\Leftrightarrow F_{X_1X_2\ldots X_n}\left(x_1,x_2,...,x_n\right)=\prod_{i=1}^{n}F_{X_i}\left(x_i\right)\\&\Leftrightarrow f_{X_1X_2\ldots X_n}\left(x_1,x_2,...,x_n\right)=\prod_{i=1}^{n}f_{X_i}\left(x_i\right)\end{aligned}$$

(1) 讨论独立时不一定全离散或者全连续，上面结论对混合情形依然成立。

(2)

$$\begin{aligned}X,Y\text{相互独立}&\Leftrightarrow f_{XY}\left(x,y\right)=f_X\left(x\right)f_Y\left(y\right)\qquad\left(x,y\right)\in\mathscr{J}\\&\Leftrightarrow f_{X|Y}\left(x|y\right)=f_X\left(x\right),y\in\mathscr{Y}\\&\Rightarrow E\left(X|Y=y\right)=E\left(X\right)\end{aligned}$$

例 随机选取清华学生 ω，其身高与性别分别表示为 $H(\omega), G(\omega)$，则 $f_{H|G}\left(h|g=\text{男生}\right), f_{H|G}\left(h|g=\text{女生}\right)$ 有如下图的关系：



(a) 男女身高问题

0.1.2 条件期望

定义 0.1 条件期望 条件分布的数学期望称为条件数学期望。

$$E\left(X|Y=y\right)=\begin{cases}\int_{-\infty}^{+\infty} x f_{X|Y}\left(x|y\right)dx & X\text{为连续型随机变量}\\ \sum\limits_{x_i} x_i f_{X|Y}\left(x_i|y\right) & X\text{为离散型随机变量}\end{cases}$$

注 1： $E\left(X\right)$ 与 $E\left(X|Y=y\right)$ 不同，其对应的概率密度函数不同。

注 2：

$$E\left[h\left(X\right)|Y=y\right]=\int h\left(x\right)f_{X|Y}\left(x|y\right)dx$$

$$E\left[aU\left(X\right)+bV\left(Z\right)|Y\right]=aE\left[U\left(X\right)|Y\right]+bE\left[V\left(Z\right)|Y\right]$$

$$E\left[Y|Y=y\right]=y$$

$$E\left[\varsigma\left(X\right)\eta\left(Y\right)h\left(X,Y\right)|Y=y\right]=\eta\left(y\right)E\left[\varsigma\left(X\right)h\left(X,y\right)|Y=y\right]$$

注 3： 考虑函数 $g(y) \triangleq E(X|Y=y), y \in \mathscr{Y}$。由此，我们可定义随机变量 Y 的函数

$$g(Y) = E(X|Y)$$

$g(Y)$ 的定义域为 Y 的支集 $\mathscr{Y}$

$$\begin{aligned} g &: \mathscr{Y} \to \mathscr{R} \\ Y &: \Omega \to \mathscr{Y} \end{aligned}$$

这样，经过复合，

$$g(Y) \triangleq E(X|Y) : \Omega \to \mathscr{R}$$

显然，$g(Y)$ 是随机变量，我们可以计算其期望值，

$$\begin{aligned} E[g(Y)] &= \int_{y\in\mathscr{Y}} g(y) f_Y(y) dy \\ &= \int_{y\in\mathscr{Y}} E(X|Y=y) f_Y(y)\, dy \\ &= \int_{y\in\mathscr{Y}} \left[\int_{-\infty}^{+\infty} x f_{X|Y}(x|y) dx\right] f_Y(y)\, dy \\ &= \int_{-\infty}^{+\infty} x \left[\int_{y\in\mathscr{Y}} f_{X|Y}(x|y) f_Y(y)\, dy\right] dx \\ &= \int_{-\infty}^{+\infty} x \int_{y\in\mathscr{Y}} f_{XY}(x,y)\, dy dx \\ &= \int_{-\infty}^{+\infty} x f_X(x) dx \\ &= E(X) \end{aligned}$$

这样，我们得出一个重要结论。

重期望公式：

$$E[E(X|Y)] = E(X)$$

背后的道理：分情况讨论 为求 $E(X)$，有时 $E(X)$ 不易求得，可以考虑引入一个条件变量 Y，在条件变量 Y 取值于具体值 y 之下，$E(X|Y=y)$ 易于求出，则可考虑利用重期望公式求 $E(X)$。这有点全概率公式的味道。