0.1 多维随机变量的特征数

二维随机变量 (X,Y) 之间主要表现为独立与相互依存两类关系。在许多问题中，随机变量取值往往是彼此有影响的，条件分布是研究变量之间相依关系的一个有力工具。

对于二维随机变量 (X,Y) 而言，所谓随机变量 X 的条件分布，就是在给定 Y 取得某个值的条件下 X 的分布。根据两个变量连续与离散的不同，条件分布有四种不同情形，以下举例说明。

例 0.1 设随机地选取一名清华本科生 ω，定义与 ω 有关的四个随机变量：连续型随机变量 $X(\omega)\,Y(\omega)$ 分别表示所选取本科生的身高与体重，离散型随机变量 $U(\omega)\,V(\omega)$ 则分别表示该本科生的性别与学科专业，其中：

$$U(\omega)=\begin{cases}0, & \text{男生}\\ 1, & \text{女生}\end{cases}\quad V(\omega)=\begin{cases}0, & \text{理工科}\\ 1, & \text{文科}\end{cases}$$

下表中展示的四中条件分布组合代表了四种不同类型：

表 1: 条件分布的四种情况

分布种类	离散	连续
离散	$U\,\vert V=0$	$X\,\vert V=0$
连续	$V\,\vert Y=70kg$	$X\,\vert Y=70kg$

可见，根据主变量与条件变量是离散还是连续，条件分布可分为四种情形。几乎所有的概率论教材都只介绍全离散与全连续两种情形，而对其他两种不做讲述，但是后续课程中，另外两种情况却经常用到。

0.1.1 条件分布

首先，我们对变量的条件分布函数做出一般的定义：

$$\begin{aligned}&F_{X|Y}\left(x\,|y\right)\triangleq P\left(X\le x\,|Y=y\right)\\&=\begin{cases}\frac{P\left(X\le x|Y=y\right)}{P\left(Y=y\right)}, & \text{若}P\left(Y=y\right)>0\\ \lim\limits_{\varepsilon\to0^+}P\left(X\le x\,|y\le Y\le y+\varepsilon\right)=\lim\limits_{\varepsilon\to0^+}\frac{P\left(X\le x,y\le Y\le y+\varepsilon\right)}{P\left(y\le Y\le y+\varepsilon\right)}, & \text{若}P\left(Y=y\right)=0\end{cases}\end{aligned}\tag{1}$$

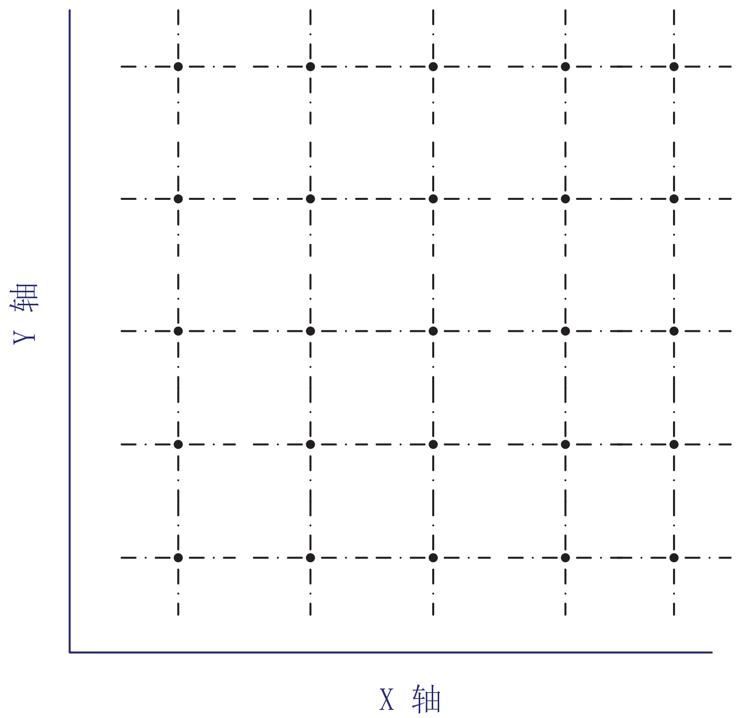
注 (1)： 若 $\exists\delta>0,\forall 0<\varepsilon<\delta,P\left(X\le x,y\le Y\le y+\varepsilon\right)>0,$，则 $P\left(X\le x\,|Y=y\right)\triangleq\lim\limits_{\varepsilon\to0^+}P\left(X\le x\,|y\le Y\le y+\varepsilon\right)$。

注 (2)： 上述定义在一元和多元情形下是统一的，也就是随机变量扩展到多维空间，$X\in R^n,Y\in R^m$ 上述定义仍然成立。

下面，我们分情形讨论条件分布。

全离散型

考虑二维随机变量 (X,Y)，其中 X,Y 均为离散型随机变量，设其支集为 $\mathscr{J}$。$\mathscr{J}$ 在第一维的投影即为 X 的支集 $\mathscr{X}$，在第二维的投影即为 Y 的支集 $\mathscr{Y}$。



(a) 全离散型示意图

设 X,Y 的联合分布与边缘分布分别为 $f_{XY}(x,y), f_X(x), f_Y(y)$，其中 $x\in\mathscr{X}, y\in\mathscr{Y}$，则：

$$P(X=x\,|Y=y)=\frac{P(X=x,Y=y)}{P(Y=y)}=\frac{f_{XY}(x,y)}{f_Y(y)}\triangleq f_{X|Y}(x\,|y) \tag{2}$$

$$F_{X|Y}(x|y)\triangleq P(X\le x\,|Y=y)=\sum_{u\le x}f_{X|Y}(u|y)$$

另设 $x\in\mathscr{X}$，可以得出以下结论。

乘法公式：

$$f_{XY}(x,y)=f_Y(y)\,f_{X|Y}(x|y)\,f_{XYZ}(x,y,z)=f_Z(z)\,f_{Y|Z}(y|z)\,f_{X|YZ}(x|yz)$$

全概率公式：

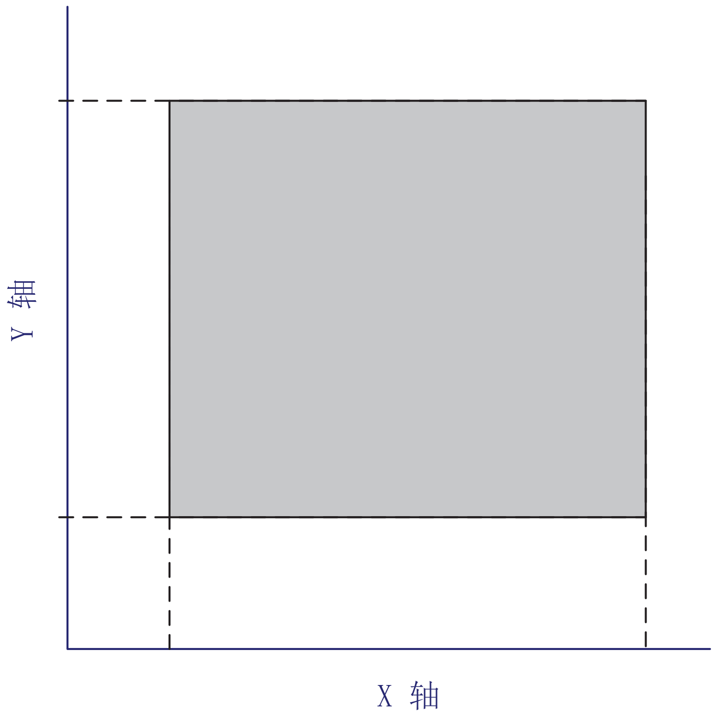
$$f_X(x)=\sum_{y\in Y}f_{XY}(x,y)=\sum_{y\in\mathscr{Y}}f_Y(y)f_{X|Y}(x|y)$$

贝叶斯公式：

$$f_{Y|X}(y\,|x)=\frac{f_{XY}(x,y)}{f_X(x)}=\frac{f_Y(y)\,f_{X|Y}(x|y)}{\sum\limits_{v\in\mathscr{Y}}f_Y(v)f_{X|Y}(x|v)}$$

全连续型

paragraph 考虑二维随机变量 (X,Y)，其中 X,Y 均为连续型随机变量，设其支集为 $\mathscr{J}$。$\mathscr{J}$ 在第一维的投影即为 X 的支集 $\mathscr{X}$，在第二维的投影即为 Y 的支集 $\mathscr{Y}$。



(b) 全连续型示意图

设 X,Y 均为连续型随机变量，具有联合分布与边缘分布分别 $f_{XY}(x,y), f_X(x), f_Y(y)$。

$$
\begin{aligned}
F_{X|Y}(x|y) &\triangleq \lim_{\varepsilon\to 0^+} \frac{P(X\le x, y\le Y\le y+\varepsilon)}{P(y\le Y\le y+\varepsilon)} \\
&= \lim_{\varepsilon\to 0^+} \frac{\int_{-\infty}^{x}\int_{y}^{y+\varepsilon} f_{XY}(u,v)dvdu}{\int_{y}^{y+\varepsilon} f_Y(v)dv} \\
&= \lim_{\varepsilon\to 0^+} \frac{\int_{-\infty}^{x}\left[\frac{1}{\varepsilon}\int_{y}^{y+\varepsilon} f_{XY}(u,v)dv\right]du}{\frac{1}{\varepsilon}\int_{y}^{y+\varepsilon} f_Y(v)dv} \\
&= \frac{\int_{-\infty}^{x} f_{XY}(u,y)du}{f_Y(y)} \\
&= \int_{-\infty}^{x} \frac{f_{XY}(u,y)}{f_Y(y)}du
\end{aligned}
$$

回想概率密度的定义：

回忆： 称 $f_X(\cdot)$ 为概率密度函数，如果 $F_X(x)=\int_{-\infty}^{x} f_X(u)\,du$。因此，利用以上推导的结果，我们可以定义条件分布密度：

$$
f_{X|Y}(x,y) = \frac{f_{XY}(x,y)}{f_Y(y)} \tag{3}
$$

利用上述定义，可以类似推得连续情形下得乘法公式、全概率公式以及贝叶斯公式，此处不作赘述。

例 0.2 设

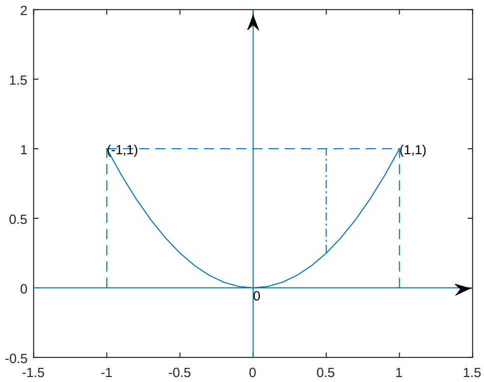
$$p(x,y)=\begin{cases}\frac{21}{4}x^2y, & x^2\le y\le 1\\ 0, & others\end{cases}$$

(1) 求 $P\left(Y\le\frac{3}{4}|X=0.5\right)$；*(2)* 求 $P\left(Y\le\frac{3}{4}|X\le 0.5\right)$。

解：(1) 计算边缘概率密度，

$$p(x)=\int_{-\infty}^{+\infty}p(x,y)\,dy=\int_{x^2}^{1}\frac{21}{4}x^2ydy=\frac{21}{8}x^2\left(1-x^4\right)$$

注：可以通过归一性验证结果的正确性。计算条件分布密度，当 $-1\le x\le 1, x^2\le y\le 1$ 时



(c) 概率密度函数区域示意图

有：

$$p(y|x)=\frac{p(x,y)}{p(x)}=\frac{\frac{21}{4}x^2y}{\frac{21}{8}x^2(1-x^4)}=\frac{2y}{1-x^4}$$

故，

$$P\left(Y\le\frac{3}{4}|X=0.5\right)=\int_{\frac{1}{4}}^{\frac{3}{4}}\frac{2y}{1-x^4}dy=\frac{8}{15}$$

(2) 计算条件概率，

$$P\left(Y\le\frac{3}{4}|X\le 0.5\right)=\frac{P\left(Y\le\frac{3}{4},X\le 0.5\right)}{P\left(X\le 0.5\right)}=\frac{\int_{-\frac{\sqrt{3}}{2}}^{\frac{1}{2}}\left(\int_{X^2}^{\frac{3}{4}}\frac{21}{4}x^2ydy\right)dx}{\int_{-1}^{0.5}\frac{21}{8}x^2\left(1-x^4\right)dx}$$

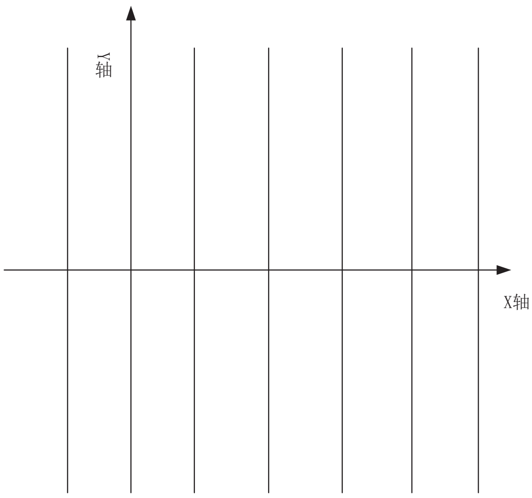
混合型

讨论 $X\in\mathscr{X}$ 为离散型随机变量，而 $y\in\mathscr{Y}$ 为连续型随机变量，集合 $\mathscr{J}=(\mathscr{X},\mathscr{Y})$。

联合分布函数：$F_{XY}(x,y)\triangleq P(X\le x,Y\le y)$

定义 0.1 若 $F_{XY}(x,y)\triangleq P(X\le x,Y\le y)$ 满足 $\forall(x,y)\in\mathscr{J}$, $\sum\limits_{u\le x}\int_{-\infty}^{y}f_{XY}(u,v)\,dv=F_{XY}(x,y)$，则称 $f_{XY}(\cdot,\cdot)$ 为混合概率密度。

可以推导相应的条件分布密度：



(d) 混合型示意图

连续变量 Y 作为条件变量

$$f_{X|Y}\left(x\,|y\right) = \begin{cases} \frac{f_{XY}(x,y)}{f_Y(y)}, & (x,y) \in \mathscr{J} \\ 0, & (x,y) \in \mathscr{J} \end{cases} \tag{4}$$

离散变量 X 作为条件变量

$$f_{Y|X}\left(y\,|x\right) = \frac{f_{XY}\left(x,y\right)}{f_X\left(x\right)}, X \in \mathscr{X} \tag{5}$$

又有 ，

$$f_{XY}\left(x,y\right) = f_X\left(x\right) f_{Y|X}\left(y\,|x\right) = f_Y\left(y\right) f_{X|Y}\left(x\,|y\right) \tag{6}$$

由此不难看出，可以类似推得混合情形下的乘法公式、全概率公式以及贝叶斯公式，此处不作赘述。