作业三讲评

崔仲杰

2024 年 3 月 29 日

边缘分布及边缘密度函数

定义

记 (X, Y) 的联合概率分布为 $p_{ij} = P(X = x_i, Y = y_j)$，则 X, Y 的边缘分布为

$$P(X = x_i) = \sum_j P(X = x_i, Y = y_j) := p_{i\bullet}$$

$$P(Y = y_j) = \sum_i P(X = x_i, Y = y_j) := p_{\bullet j}$$

定义

记 (X, Y) 的概率密度为 $f(x, y)$，则 X, Y 的概率密度为

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

注意到

$$F_X(x) = F(x, \infty) = \int_{-\infty}^{x} \left[\int_{-\infty}^{\infty} f(x, y)dy\right] dx$$

同时我们有 $F_X(x) = \int_{-\infty}^{x} f_X(x)dx$，因此我们有

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

例

1. 离散型随机向量 (X, Y) 的分布列如下表.

$X \backslash Y$	-1	0	1
-1	a	0	0.2
0	0.2	b	0.1
1	0	0.1	c

已知 $P(X \geq 0) = 0.6$, $P(X \neq 0 \mid Y \neq 0) = 5/8$. 求:

(1) a, b, c 的值.

(2) X, Y 各自的边缘分布列.

(3) $X + Y$ 的分布列.

解.

(1) 由分布列的归一性，$a+b+c=0.4$. 由 $P(X\geq 0)=0.6$ 知 $b+c=0.2$. 由 $P(X\neq 0 \mid Y\neq 0)=5/8$ 知 $8P(X\neq 0, Y\neq 0)=5, P(Y\neq 0)$，即 $8(a+c+0.2)=5(a+c+0.5)$. 联立以上三式解得：$(a,b,c)=(0.2,0.1,0.1)$.

(2) X, Y 的边缘分布列如下：

k	-1	0	1
$P(X=k)$	0.4	0.4	0.2
$P(Y=k)$	0.4	0.2	0.4

(3) $X+Y$ 的分布列如下：

k	-2	-1	0	1	2
$P(X+Y=k)$	0.2	0.2	0.3	0.2	0.1

密度变换

定义

若 $\boldsymbol{X}$ 和 $\boldsymbol{Y}=\boldsymbol{g}(\boldsymbol{X})$ 都是 n 维随机向量，我们有密度变换公式:
设连续型随机向量 $\boldsymbol{X}=(X_1,X_2,\ldots,X_n)$ 的概率密度为 $f_{\boldsymbol{X}}(\boldsymbol{x})$，$\boldsymbol{g}:\mathbb{R}^n\to\mathbb{R}^n$ 有连续偏导数，且反函数 $\boldsymbol{h}(\boldsymbol{y})=\boldsymbol{g}^{-1}(\boldsymbol{y})$ 唯一，则随机向量 $\boldsymbol{Y}=\boldsymbol{g}(\boldsymbol{X})$ 的概率密度为 $f_{\boldsymbol{Y}}(\boldsymbol{y})=f_{\boldsymbol{X}}(\boldsymbol{h}(\boldsymbol{y}))|J(\boldsymbol{y})|$，其中 $J(\boldsymbol{y})=\det \boldsymbol{J}(\boldsymbol{y})$ 为 $\boldsymbol{h}$ 在 $\boldsymbol{y}$ 处的 Jacobi 行列式，这里有矩阵 $\boldsymbol{J}_{jk}=\partial h_j(\boldsymbol{y})/\partial y_k$
直观理解:

$$f_{\boldsymbol{Y}}(\boldsymbol{y})=f_{\boldsymbol{X}}(\boldsymbol{x})\left|\frac{\mathrm{d}\boldsymbol{x}}{\mathrm{d}\boldsymbol{y}}\right|=f_{\boldsymbol{X}}(\boldsymbol{h}(\boldsymbol{y}))|J(\boldsymbol{y})|$$

一元情形下，$\left|\frac{\mathrm{d}x}{\mathrm{d}y}\right|=|h'(y)|$，其中 h 是 g 的反函数，这就是第二章连续随机变量的函数分布定理.

例

2. 证明如下两个问题:

(1) 设随机向量 (X, Y, Z) 的密度函数为

$$f(x,y,z)=\begin{cases}\frac{1}{8\pi^3}(1-\sin x\sin y\sin z), & 0\le x,y,z\le 2\pi,\\ 0, & \text{其他}.\end{cases}$$

证明 X, Y, Z 两两独立但不相互独立.

(2) 设随机向量 (X, Y) 的密度函数为

$$f(x,y)=\begin{cases}\frac{1+xy}{4}, & |x|<1, |y|<1,\\ 0, & \text{其他}.\end{cases}$$

证明 X, Y 不独立但是 X^2, Y^2 是相互独立的.

解.

(1) $(X,Y),(X,Z),(Y,Z)$ 的联合分布均为

$$f(x,y)=\begin{cases}\frac{1}{4\pi^2}, & 0\le x,y\le 2\pi\\ 0, & \text{其他}.\end{cases}$$

而 X,Y,Z 的边缘分布相等且都等于

$$f(x)=\begin{cases}\frac{1}{2\pi}, & 0\le x\le 2\pi,\\ 0, & \text{其他}.\end{cases}$$

由此立得 X,Y,Z 两两独立. 但 $f_{X,Y,Z}(x,y,z)\neq f_X(x)f_Y(y)f_Z(z)$, 因此 X,Y,Z 不是相互独立的.

(2) 对联合密度函数积分得到:

$$f_X(x)=f_Y(x)=\begin{cases}1/2, & |x|<1,\\ 0, & \text{其他},\end{cases}$$

所以 $f(x,y)\neq f_X(x)f_Y(y)$, 即 X,Y 不独立. 而

$$f_{X^2,Y^2}(x,y)=\begin{cases}1/(4\sqrt{xy}), & 0<x,y<1,\\ 0, & \text{其他},\end{cases}$$

积分可得

$$f_{X^2}(x)=f_{Y^2}(x)=\begin{cases}1/(2\sqrt{x}), & 0<x,y<1,\\ 0, & \text{其他},\end{cases}$$

注

关于 f_{X^2,Y^2} 的计算.
为方便理解, 一般的, 令 $U=X^2, V=Y^2$, (U,V) 的密度函数为 $g(u,v)=f_{X^2,Y^2}(x,y)$, 下面应用概率元的思想:
只考虑 $|x|<1, |y|<1$ 的情况, $g(u,v)|dudv|=\Sigma_{(\pm x,\pm y)}f(x,y)|dxdy|$,
这里对 $(x,y),(x,-y),(-x,y),(-x,-y)$ 四个点进行求和,
带入 $du=2xdx, dv=2ydy$, 可得 $g(u,v)=1/|4xy|=1/(4\sqrt{uv})$.
联系理解雅可比行列式 (一一映射时)

$$J=\begin{vmatrix} \frac{\partial x}{\partial u} & \frac{\partial x}{\partial v} \\ \frac{\partial y}{\partial u} & \frac{\partial y}{\partial v} \end{vmatrix}=\left|\frac{dxdy}{dudv}\right|.$$

(上式写法不严谨, 仅提供理解)

例

3. 设随机向量 (X, Y) 有联合密度

$$f_{X,Y}(x,y) = \begin{cases} C(|xy| + xy/2), & |x| < 1, |y| < 1, \\ 0, & \text{其它}. \end{cases}$$

(1) 求 C 的值；计算 X, Y 各自的边缘密度.

(2) X, Y 是否相互独立？

(3) $X^2 + Y^2, X^2 - Y^2$ 的概率密度是什么？是否相互独立？

解.

(1) 因为 $\int_{-1}^{1}\int_{-1}^{1}(|xy|+xy/2)\mathrm{d}x\mathrm{d}y=\int_{-1}^{1}\int_{-1}^{1}|xy|\mathrm{d}x\mathrm{d}y=1$，所以由归一化条件知 $C=1$.

X 的边缘密度为 $f_X(x)=\int_{-1}^{1}f_{X,Y}(x,y)\mathrm{d}y=\begin{cases}|x|, & |x|<1,\\ 0, & |x|\geq 1.\end{cases}$.

同理，Y 的边缘密度 $f_Y(y)=\begin{cases}|y|, & |y|<1,\\ 0, & |y|\geq 1.\end{cases}$.

(2) $f_{X,Y}(x,y)\neq f_X(x)f_Y(y)$，所以 X,Y 不相互独立.

解.

(3) 复杂的换元使用密度变换公式求解.
首先求出 Jacobi 行列式

$$\left|\frac{\partial\left(x^2+y^2, x^2-y^2\right)}{\partial\left(x, y\right)}\right| = 8|xy|$$

因此得到 $f(x,y) = \frac{f_{X,Y}(x,y)}{8|xy|}$
但注意这一变换是 4 对 1 的，因此最终 $f_{X^2+Y^2,X^2-Y^2}$ 的结果应分为四支，即

$$f_{X^2+Y^2,X^2-Y^2}(k,l) = 4|xy|/8|xy| = \frac{1}{2}$$

考虑到取值范围，概率密度应为

$$f_{X^2+Y^2,X^2-Y^2}(k,l) = \frac{1}{2}1_{-l\le k\le 2-l}1_{l\le k\le 2+l}$$

注意虽然在 $1_{-l\le k\le 2-l}1_{l\le k\le 2+l}$ 并不能分解为独立两部分的乘积，因此 X^2+Y^2, X^2-Y^2 并不相互独立.
X^2+Y^2, X^2-Y^2 的密度函数如下:

$$f_{X^2+Y^2}(k) = \begin{cases} k, & 0<k<1, \\ 2-k, & 1\le k\le 2, \\ 0, & \text{else}. \end{cases} \qquad f_{X^2-Y^2}(l) = \begin{cases} 1-l, & 0<l<1, \\ l+1, & -1\le l\le 0, \\ 0, & \text{else}. \end{cases}$$

例

4. 设 Z 是取值为正整数的随机变量，在集合 $\{1, 2, \cdots, Z\}$ 中均匀随机取一个数 X，记 R 是集合 $\{X+1, X+2, \ldots, Z\}$ 的大小，即比 X 大的数的个数. 设随机变量 U 与 Z 独立，且 $U \sim U(0, 1)$，问 R 与 $[UZ]$ 是否同分布？($[x]$ 表示 x 的整数部分)

解. 当 $Z = z$ 时，R 为一均匀分布，即

$$P(R = r \mid Z = z) = \frac{1}{z} \quad r = 0, 1, \ldots, z-1$$

因此 R 的分布列为

$$P(R = r) = \sum_{z} P(R = r \mid Z = z) P(Z = z) = \sum_{z=r+1}^{+\infty} \frac{1}{z} P(Z = z)$$

同时 $[UZ]$ 的分布列为

$$P([UZ] = r) = \sum_{z} P(\frac{r}{z} \le U < \frac{r+1}{z}) P(Z = z) = \sum_{z=r+1}^{+\infty} \frac{1}{z} P(Z = z)$$

因此二者是同分布的.

例

5. 称取值为正整数的离散型随机变量 ξ 服从以 (n,p) 为参数的帕斯卡分布，如果它的分布列为

$$P(\xi=r)=\begin{cases}\binom{r-1}{n-1}p^nq^{r-n}, & r=n,n+1,\cdots,\\ 0, & \text{其它}.\end{cases}$$

设随机变量 X,Y 相互独立，且分别服从参数为 $(1,p)$ 和 $(2,p)$ 的帕斯卡分布，求证：$X+Y$ 服从参数为 $(3,p)$ 的帕斯卡分布.

证明.

由题意，$X+Y$ 的可能取值为所有不小于 3 的整数. 对 $r\in N^*$，$r\geq 3$，有:

$$P(X+Y=r)=\sum_{\ell=1}^{r-2}P(X=\ell)P(Y=r-\ell)=q^{r-3}p^3\sum_{\ell=1}^{r-2}(r-\ell-1)$$
$$=\frac{(r-1)(r-2)}{2}q^{r-3}p^3.$$

由定义知，$X+Y$ 服从参数为 $(3,p)$ 的帕斯卡分布. □

注

关于帕斯卡分布，我们有如下更一般的结论:

若 $X_1,\cdots,X_n$ 相互独立且都服从参数为 p 的几何分布，则 $X_1+\cdots+X_n$ 服从参数为 (n,p) 的帕斯卡分布. 这里，X 服从参数为 p 的几何分布指 $P(X=k)=pq^{k-1}$.

若随机变量 X,Y 相互独立，且分别服从参数为 (m,p) 和 (n,p) 的帕斯卡分布，则 $X+Y$ 服从参数为 $(m+n,p)$ 的帕斯卡分布.
第一个结论表明帕斯卡分布刻画的是在伯努利实验中需要多长时间才会出现第 n 次 “成功”；第二个结论是第一个结论的直接推论.

例

6. 已知随机变量 X, Y 相互独立且均服从参数为 p 的几何分布，令 $Z = \max\{X, Y\}$. 求：

(1) (X, Z) 的联合概率分布.

(2) X 关于 Z 的条件概率分布.

解. 记 $q = 1 - p$.

(1) X, Z 都只在 N 上取值. 设 $x, z \in N$. 分情况讨论如下：

- 当 $x < z$ 时，$P(X = x, Z = z) = P(X = x, Y = z) = pq^x \cdot pq^z$.
- 当 $x = z$ 时，$P(X = x, Z = z) = P(X = x, Y \le x) = pq^x(1 - q^{x+1})$.
- 当 $x > z$ 时，$P(X = x, Z = z) = 0$.

综上，(X, Z) 的联合概率分布为

$$P(X = x, Z = z) = \begin{cases} p^2 q^{x+z}, & x < z, \\ pq^x(1 - q^{x+1}), & x = z, \\ 0, & x > z. \end{cases}$$

解.

(2) 设 $z \in N$，则 $P(Z \le z) = P(X \le z, Y \le z) = (1 - q^{z+1})^2$，所以 Z 的边缘分布列为

$$P(Z = z) = P(Z \le z) - P(Z \le z - 1) = pq^z(2 - q^{z+1} - q^z). \qquad (z \in N)$$

X 关于 Z 的条件概率分布为 $(x, z \in N)$：

$$P(X = x \mid Z = z) = \frac{P(X = x, Z = z)}{P(Z = z)} = \begin{cases} pq^x/(2 - q^{z+1} - q^z), & x < z, \\ (1 - q^{x+1})/(2 - q^{z+1} - q^z), & x = z, \\ 0, & x > z. \end{cases}$$

注

求 Z 的边缘的时候用了一点小技巧：先求分布函数再求分布列.

例

7. 设随机变量 $X_1, \cdots, X_n \sim \mathrm{i.i.d.} U(0,1)$.

(1) 设 $Y_n = \min\limits_{1 \le i \le n} X_i$，$Y_n$ 的分布函数为 $F_n(x)$. 对一切 $x \in \mathbb{R}$，计算 $\lim\limits_{n \to +\infty} F_n(x)$.

(2) 令 $Z = \min\{n \in \mathbb{N}^* \mid X_1 + \cdots + X_n > 1\}$. 求证：$P(Z > 3) = 1/6$.

解.

(1) 当 $x \leq 0$ 时，$\lim_{n\to+\infty} F_n(x) = 0$；当 $x \geq 1$ 时，$\lim_{n\to+\infty} F_n(x) = 1$. 下设 $x \in (0,1)$.
注意到，$F_n(x) = P(Y_n \leq x) = 1 - P(Y_n > x) = 1 - P(X_1 > x, \cdots, X_n > x) = 1 - (1-x)^n$，所以 $\lim_{n\to+\infty} F_n(x) = 1$.
综上，$\lim_{n\to+\infty} F_n(x) = \begin{cases} 0 & x \leq 0 \\ 1 & x > 0 \end{cases}$.

(2) 注意到，事件 $\{Z > 3\}$ 等价于 $\{X_1 + X_2 + X_3 \leq 1\}$. 设区域 $D = \{(x_1, x_2, x_3) \in (0,1)^3 \mid x_1 + x_2 + x_3 \leq 1\}$，则

$$P(Z > 3) = P(X_1 + X_2 + X_3 \leq 1) = \int_D f_{X_1,X_2,X_3}(x_1,x_2,x_3)\mathrm{d}x_1\mathrm{d}x_2\mathrm{d}x_3$$
$$= \int_D \mathrm{d}x_1\mathrm{d}x_2\mathrm{d}x_3 = \frac{1}{6}.$$

注

此题意在练习用分布函数求随机变量函数的分布.
第 (1) 问表明这种方法在研究最值的分布时非常有效.
第 (2) 问也可以先用卷积公式求出 $X_1 + X_2 + X_3$ 的密度，再对密度积分得到分布函数，但直接求解显得更为简捷.
第 (2) 问的一个推广是：$P(Z > n) = 1/n!(n \in N^*)$.

例

8. 若 ξ 与 η 是相互独立的随机变量，均服从 $N(0,1)$，现在将 (ξ,η) 化为极坐标 (ρ,φ)，其中 $\xi=\rho\cos\varphi,\eta=\rho\sin\varphi$ $(\rho\geq 0,\varphi\in[0,2\pi])$. 试证 ρ,φ 是相互独立的.

证明.

假设 (ξ,η) 的密度函数为 $p(x,y)$，(ρ,φ) 的密度函数为 $q(r,\theta)$，采用极坐标，$x=r\cos\theta, y=r\sin\theta$，因为 (ξ,η) 的密度函数为

$$p(x,y)=\frac{1}{2\pi}\mathrm{e}^{-(x^2+y^2)/2}$$

而

$$J=\begin{vmatrix}\frac{\partial x}{\partial r} & \frac{\partial x}{\partial\theta}\\ \frac{\partial y}{\partial r} & \frac{\partial y}{\partial\theta}\end{vmatrix}=\begin{vmatrix}\cos\theta & -r\sin\theta\\ \sin\theta & r\cos\theta\end{vmatrix}=r$$

故 (ρ,φ) 的密度函数为

$$\begin{aligned}q(r,\theta)&=\frac{1}{2\pi}\mathrm{e}^{-(x^2+y^2)/2}\cdot r\\ &=\frac{1}{2\pi}\cdot r\mathrm{e}^{-r^2/2},\quad r\geqslant 0,\quad 0\leqslant\theta\leqslant 2\pi\end{aligned}$$

证明.

对 θ 积分得到 $\rho=\sqrt{\xi^2+\eta^2}$ 的密度函数为

$$R(r)=\begin{cases} re^{-r^2/2}, & r\geqslant 0\\ 0, & r<0\end{cases}$$

这个分布称为瑞利 (Rayleigh) 分布.
而联合密度函数与 θ 无关, 故 $\varphi=\operatorname{arctg}\frac{\eta}{\xi}$ 服从 $[0,2\pi]$ 中的均匀分布, 所以 ρ 与 φ 是独立的. □

注

这个结果常被用来产生服从正态分布的随机数. 做法如下: 产生相互独立的 $[0,1]$ 均匀分布的随机数 U_1, U_2, 令

$$\begin{cases} \xi=(-2\ln U_1)^{\frac{1}{2}}\cos 2\pi U_2\\ \eta=(-2\ln U_1)^{\frac{1}{2}}\sin 2\pi U_2\end{cases}$$

则 ξ 与 η 是相互独立的 $N(0,1)$ 随机数.

二元正态分布

定义

随机变量 (X, Y) 的概率密度为

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

其中 $\sigma_1, \sigma_2 > 0, |\rho| < 1$，并称 (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 $N\left(\mu_1, \mu_2, \sigma_1^2, \sigma_2^2, \rho\right)$ $N(0,0,1,1,0)$ 为二维独立（标准）正态分布

$f_X(x) = \frac{1}{\sqrt{2\pi}\sigma_1} e^{-\frac{(x-\mu_1)^2}{2\sigma_1^2}}$，也即是 $X \sim N\left(\mu_1, \sigma_1^2\right)$. 同理 $Y \sim N\left(\mu_2, \sigma_2^2\right)$

条件密度函数为 $f_{X|Y}(x \mid y) = \frac{1}{\sqrt{2\pi}\sigma_1\sqrt{1-\rho^2}} e^{\frac{-\left[x-\left(\mu_1+\rho\frac{\sigma_1}{\sigma_2}(y-\mu_2)\right)\right]^2}{2(1-\rho^2)\sigma_1^2}}$,

也即 $Y = y$ 条件下 $X \sim N\left(\mu_1 + \rho\frac{\sigma_1}{\sigma_2}(y-\mu_2), (1-\rho^2)\sigma_1^2\right)$,

同理在 $X = x$ 条件下 $Y \sim N\left(\mu_2 + \rho\frac{\sigma_2}{\sigma_1}(x-\mu_1), (1-\rho^2)\sigma_2^2\right)$

例

9. 设 (X, Y) 服从二元正态分布 $N\left(\mu_1, \mu_2, \sigma_1^2, \sigma_2^2, -1/\sqrt{2}\right)$. 已知 (X, Y) 的联合密度为

$$f_{X,Y}(x, y) = \frac{1}{2\pi} \exp\left(-\frac{1}{2}(2x^2 + y^2 + 2xy - 22x - 14y + 65)\right).$$

(1) 将 $f_{X,Y}(x, y)$ 写成二元正态分布定义中的密度的形式，指出 $\mu_1, \mu_2, \sigma_1^2, \sigma_2^2$.
(2) 写出边缘密度 $f_X(x)$.
(3) 写出条件密度 $f_{Y|X}(y \mid x)$.

解.

(1) 由条件得： $f_{X,Y}(x,y) = \frac{1}{2\pi\cdot 1\cdot\sqrt{2}\cdot\sqrt{1-\frac{1}{2}}}\cdot$
$\exp\left(-\frac{1}{2(1-\frac{1}{2})}\left((x-4)^2 - 2(-\frac{1}{\sqrt{2}})\frac{(x-4)(y-3)}{1\cdot\sqrt{2}} + \frac{(y-3)^2}{2}\right)\right).$
由上式可以看出： $(\mu_1, \mu_2, \sigma_1^2, \sigma_2^2) = (4, 3, 1, 2).$

(2) 由二元正态分布的性质， $X \sim N(\mu_1, \sigma_1^2)$. 所以密度 $f_X(x) = \frac{1}{\sqrt{2\pi}}\exp\left(-\frac{1}{2}(x-4)^2\right).$

(3) 条件密度 $f_{Y|X}(y \mid x) = \frac{1}{\sqrt{2\pi}}\exp\left(-\frac{1}{2}(y-(7-x))^2\right).$

注

有很多种方法可以写出条件密度 $f_{Y|X}(y \mid x)$.
其一是记住结论： 在 $X = x$ 的条件下， Y 的条件分布是 $N(\mu_2 + \rho\frac{\sigma_1}{\sigma_2}(x-\mu_1), (1-\rho^2)\sigma_2^2)$.
其二是用条件密度的定义： $f_{Y|X}(y \mid x) = f_{X,Y}(x,y)/f_X(x)$.

注

其三是注意 $f_{Y|X}(y \mid x)$ 与联合密度 $f_{X,Y}(x,y)$ 只差了一个与 y 无关的归一化常数. 将 y 作为主元改写联合密度 $f_{X,Y}(x,y)$，知 $f_{Y|X}(y \mid x)$ 一定具有形式

$$f_{Y|X}(y \mid x) = C\exp\left(-\frac{1}{2}(y^2 + (2x-14)y)\right),$$

其中 C 是只依赖 x 的归一化常数.

然后注意在 $X = x$ 的条件下，Y 的条件分布是**正态分布**，所以上式中指数内的小括号一定可以配凑成完全平方，即

$$f_{Y|X}(y \mid x) = \tilde{C}\exp\left(-\frac{1}{2}(y-(7-x))^2\right).$$

最后注意 $\tilde{C}$ 的作用只是让 $\int_{\mathbb{R}} f_{Y|X}(y \mid x)\,\mathrm{d}y = 1$，所以立即可以写出 $\tilde{C} = \frac{1}{\sqrt{2\pi}}$. 这就得到了要求的条件密度.

例

10. 对任意两个随机变量 X, Y，设其存在概率分布密度为 $f_{X,Y}(x, y)$，是否一定存在两个可逆函数 $Z = g(X, Y), W = h(X, Y)$ 使得 Z, W 相互独立？更一般地，是否能使 $Z, W \sim \mathrm{i.i.d.} U(0, 1)$? 为了简便起见本题假设 $f_{X,Y}(x, y) \neq 0$.

证明.

令

$$Z = \frac{\int_{-\infty}^{X} f_{X,Y}(x, Y)\mathrm{d}x}{\int_{-\infty}^{+\infty} f_{X,Y}(x, Y)\mathrm{d}x}$$

即 $Y = y$ 时 X 的条件分布函数. 注意 $\forall y$, Z 关于 X 均是严格单调上升的且值域为 $[0, 1]$, 因此变换 $(X, Y) \to (Z, Y)$ 是可逆的.
直接使用密度变换公式

$$f_{Z,Y}(z, y) = f_{X,Y}(x, y) \left| \frac{\partial(Z, Y)}{\partial(X, Y)} \right|^{-1} = f_{X,Y}(x, y) \frac{\int_{-\infty}^{+\infty} f_{X,Y}(x, y)\mathrm{d}x}{f_{X,Y}(x, y)}$$
$$= \int_{-\infty}^{+\infty} f_{X,Y}(x, y)\mathrm{d}x$$

可以发现 $f_{Z,Y}(z, y)$ 只与 y 相关，结合 Z 的值域可知 $Z \sim U(0, 1)$，因此二者是独立的. □

证明.

进一步的，还可以再次使用此方法，令

$$W = \frac{\int_{-\infty}^{Y} \mathrm{d}y \int_{-\infty}^{+\infty} \mathrm{d}x f_{X,Y}(x,y)}{\int_{-\infty}^{+\infty} \mathrm{d}y \int_{-\infty}^{+\infty} \mathrm{d}x f_{X,Y}(x,y)}$$

最终得到

$$f_{Z,W}(z,w) = f_{Z,Y}(z,y) \left| \frac{\partial(Z,W)}{\partial(Z,Y)} \right|^{-1} = 1$$

即 $Z, W \sim \text{i.i.d.} U(0,1)$. □

注

这一方法是 chapter2 所讲的使用 $U(0,1)$ 采样任意一维分布的高维推广，对于更高的维度方法仍然适用.

谢谢大家！