第三章：多元随机变量及其分布[1]

Xiao Yuan[1]

[1]Center on Frontiers of Computing Studies, Peking University, Beijing 100871, China
xiaoyuan@pku.edu.cn

March 19, 2024

[1]讲义内容基于三本参考教材以及网络素材

目录

PEKING UNIVERSITY

目录

刻画同一概率空间上不同性质

- 例 1：研究某一地区学龄儿童的发育情况。仅研究身高 H 的分布或仅研究体重 W 的分布是不够的。需要同时考察每个儿童的身高和体重值，研究身高和体重之间的关系，这就要引入定义在同一样本空间的两个随机变量。

刻画同一概率空间上不同性质

- 例 1：研究某一地区学龄儿童的发育情况。仅研究身高 H 的分布或仅研究体重 W 的分布是不够的。需要同时考察每个儿童的身高和体重值，研究身高和体重之间的关系，这就要引入定义在同一样本空间的两个随机变量。
- 例 2：研究某种型号炮弹的弹着点分布。每枚炮弹的弹着点位置需要由横坐标和纵坐标来确定，而它们是定义在同一样本空间的两个随机变量。

刻画同一概率空间上不同性质

- 例 1：研究某一地区学龄儿童的发育情况。仅研究身高 H 的分布或仅研究体重 W 的分布是不够的。需要同时考察每个儿童的身高和体重值，研究身高和体重之间的关系，这就要引入定义在同一样本空间的两个随机变量。
- 例 2：研究某种型号炮弹的弹着点分布。每枚炮弹的弹着点位置需要由横坐标和纵坐标来确定，而它们是定义在同一样本空间的两个随机变量。

随机向量

对于样本空间 $\mathcal{S}=\{e\}$ 和样本点 e，$\boldsymbol{X}=(X_1,X_2,\ldots,X_n)$ 为随机向量，其中 $X_i=X_i(e)$ 为定义在 $\mathcal{S}$ 上的实、单值函数

离散型随机向量

离散型随机向量

若 X_i 均为离散型随机变量，则 $\boldsymbol{X} = (X_1, X_2, \dots, X_n)$ 为**离散型随机向量**

离散型随机向量

离散型随机向量

若 X_i 均为离散型随机变量，则 $\boldsymbol{X} = (X_1, X_2, \ldots, X_n)$ 为离散型随机向量

- 考虑二元离散型随机向量 (X, Y)，假设 X 的取值为 $\{x_i\}$，Y 的取值为 $\{Y_i\}$，则 $p_{ij} = P(X = x_i, Y = y_j)$ 为 (X, Y) 的联合概率分布

离散型随机向量

若 X_i 均为离散型随机变量，则 $\boldsymbol{X}=(X_1,X_2,\ldots,X_n)$ 为离散型随机向量

- 考虑二元离散型随机向量 (X,Y)，假设 X 的取值为 $\{x_i\}$，Y 的取值为 $\{Y_i\}$，则 $p_{ij}=P(X=x_i,Y=y_j)$ 为 (X,Y) 的联合概率分布
- p_{ij} 满足非负性和归一性，也即是 $p_{ij}\geq 0$ 和 $\sum_{i,j}p_{ij}=1$

离散型随机向量

PEKING UNIVERSITY

离散型随机向量

若 X_i 均为离散型随机变量，则 $\boldsymbol{X}=(X_1,X_2,\ldots,X_n)$ 为离散型随机向量

- 考虑二元离散型随机向量 (X,Y)，假设 X 的取值为 $\{x_i\}$，Y 的取值为 $\{Y_i\}$，则 $p_{ij}=P(X=x_i,Y=y_j)$ 为 (X,Y) 的联合概率分布
- p_{ij} 满足非负性和归一性，也即是 $p_{ij}\geq 0$ 和 $\sum_{i,j}p_{ij}=1$
- 例：设随机变量 X 为 1、2、3、4 四个整数中等可能地取一个值，另一个随机变量 Y 为 $1\sim X$ 中等可能地取一整数值，试求 (X,Y) 的联合概率分布。

边缘分布

边缘分布

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

$$\begin{aligned} P(X = x_i) &= \sum_j P(X = x_i, Y = y_j) := p_{i\bullet}, \\ P(Y = y_j) &= \sum_i P(X = x_i, Y = y_j) := p_{\bullet j} \end{aligned} \tag{1}$$

- 例：设不吸烟、少量吸烟、大量吸烟比例为 $80\%, 15\%, 5\%$，分别患呼吸道疾病概率为 $5\%, 25\%, 70\%$。记$X = \begin{cases} 0 & \text{不吸烟} \\ 1 & \text{少量吸烟} \\ 2 & \text{大量吸烟} \end{cases}$，$Y = \begin{cases} 1 & \text{患病} \\ 0 & \text{不患病} \end{cases}$
  求联合概率分布和边缘分布

边缘分布

PEKING UNIVERSITY

- 例：设不吸烟、少量吸烟、大量吸烟比例为 $80\%, 15\%, 5\%$，分别患呼吸道疾病概率为 $5\%, 25\%, 70\%$。

  记$X = \begin{cases} 0 & \text{不吸烟} \\ 1 & \text{少量吸烟} \\ 2 & \text{大量吸烟} \end{cases}$，$Y = \begin{cases} 1 & \text{患病} \\ 0 & \text{不患病} \end{cases}$

  求联合概率分布和边缘分布

X\Y	Y=0	Y=1	P(X=i)
X=0	0.76	0.04	0.8
X=1	0.1125	0.0375	0.15
X=2	0.015	0.035	0.05
P(Y=j)	0.8875	0.1125	-

条件分布

记 (X, Y) 的联合概率分布为 $p_{ij} = P(X = x_i, Y = y_j)$，假设 $p_{\bullet j} = P(Y = y_j) > 0$ 或 $p_{i\bullet} = P(X = x_i) > 0$，则

$$P(X = x_i | Y = y_j) = \frac{P(X = x_i, Y = y_j)}{P(Y = y_j)} = \frac{p_{ij}}{p_{\bullet j}},$$

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

分别为在 $Y = y_j$ 或 $X = x_i$ 条件下的条件概率分布

例

- 接上题

X\Y	Y=0	Y=1	P(X=i)
X=0	0.76	0.04	0.8
X=1	0.1125	0.0375	0.15
X=2	0.015	0.035	0.05
P(Y=j)	0.8875	0.1125	-

求患病人中吸烟的概率，也即是

$P(X \neq 0 | Y = 1) = (0.0375 + 0.035)/0.1125 = 0.6444$

例

- 接上题

X\Y	Y=0	Y=1	P(X=i)
X=0	0.76	0.04	0.8
X=1	0.1125	0.0375	0.15
X=2	0.015	0.035	0.05
P(Y=j)	0.8875	0.1125	-

求患病人中吸烟的概率，也即是
$P(X \neq 0 | Y = 1) = (0.0375 + 0.035)/0.1125 = 0.6444$

- 假设在时间 t 内球队获得黄牌的次数服从泊松分布 $\pi(\lambda t)$。记 X_i 为在 $t_i (i = 1, 2)$ 时间内得到黄牌数 $t_2 > t_1$，求 X_1, X_2 的联合概率分布

- 接上题

X\Y	Y=0	Y=1	P(X=i)
X=0	0.76	0.04	0.8
X=1	0.1125	0.0375	0.15
X=2	0.015	0.035	0.05
P(Y=j)	0.8875	0.1125	-

求患病人中吸烟的概率，也即是
$P(X \neq 0|Y = 1) = (0.0375 + 0.035)/0.1125 = 0.6444$

- 假设在时间 t 内球队获得黄牌的次数服从泊松分布 $\pi(\lambda t)$。记 X_i 为在 $t_i (i = 1, 2)$ 时间内得到黄牌数 $t_2 > t_1$，求 X_1, X_2 的联合概率分布
$P(X_1 = i, X_2 = j) = P(X_2 = j|X_1 = i)P(X_1 = i)$,
$P(X_1 = i) = e^{-\lambda t_1}(\lambda t_1)^i/i!$,
$P(X_2 = j|X_1 = i) = e^{-\lambda(t_2-t_1)}[\lambda(t_2 - t_1)]^{j-i}/(j - i)!$

- 例：一射手进行射击，击中目标的概率为 $p \in (0,1)$，射击直中目标两次为止，设以 X 表示首次击中目标所进行的射击次数，以 Y 表示总共进行的射击次数，试求 X 和 Y 的联合分布律和条件分布律。

例

- 例：一射手进行射击，击中目标的概率为 $p \in (0,1)$，射击直中目标两次为止，设以 X 表示首次击中目标所进行的射击次数，以 Y 表示总共进行的射击次数，试求 X 和 Y 的联合分布律和条件分布律。
  $P(X=m, Y=n) = p^2q^{n-2}$, $q = 1-p$,
  $m = 1, 2, \ldots, n-1$, $n = 2, 3, \ldots$
  $P(X=m) = pq^{m-1}$
  $P(Y=n) = (n-1)p^2q^{n-2}$
  $P(X=m|Y=n) = 1/(n-1)$,
  $P(Y=n|X=m) = pq^{n-m-1}$

例

- 例：已知随机变量 X 和 Y 取值为 ± 1，且满足 $P(X=1)=1/2$, $P(Y=1|X=1)=P(Y=-1|X=-1)=1/3$ （1）求 (X,Y) 的联合分布；（2）求方程 $t^2+Xt+Y=0$ 有实根的概率

- 例：已知随机变量 X 和 Y 取值为 ± 1，且满足 $P(X=1)=1/2$, $P(Y=1|X=1)=P(Y=-1|X=-1)=1/3$ （1）求 (X,Y) 的联合分布；（2）求方程 $t^2+Xt+Y=0$ 有实根的概率
  根据题目条件，有 $P(X=-1)=1/2$，因此有
  $P(X=1,Y=1)=1/6$, $P(X=1,Y=-1)=1/3$,
  $P(X=-1,Y=-1)=1/6$, $P(X=-1,Y=1)=1/3$
  方程 $t^2+Xt+Y=0$ 有实根条件为 $X^2-4Y\geq 0$，也即是对应 $Y=-1$ 的情况，因此概率为 $P(Y=-1)=1/2$.

分布函数

分布函数

考虑随机向量 $\boldsymbol{X}=(X_1,X_2,\ldots,X_n)$ 和实变量 $\boldsymbol{x}=(x_1,x_2,\ldots,x_n)$，函数 $F(\boldsymbol{x})=P(\boldsymbol{X}\leq\boldsymbol{x})$ 为 $\boldsymbol{x}$ 的概率**分布函数**

分布函数

考虑随机向量 $\boldsymbol{X} = (X_1, X_2, \ldots, X_n)$ 和实变量 $\boldsymbol{x} = (x_1, x_2, \ldots, x_n)$，函数 $F(\boldsymbol{x}) = P(\boldsymbol{X} \leq \boldsymbol{x})$ 为 $\boldsymbol{x}$ 的概率**分布函数**

- 非负性：$F(\boldsymbol{x}) \in [0, 1]$ 且满足 $F(-\infty) = 0$ 和 $F(\infty) = 1$

分布函数

分布函数

考虑随机向量 $\boldsymbol{X} = (X_1, X_2, \ldots, X_n)$ 和实变量 $\boldsymbol{x} = (x_1, x_2, \ldots, x_n)$，函数 $F(\boldsymbol{x}) = P(\boldsymbol{X} \le \boldsymbol{x})$ 为 $\boldsymbol{x}$ 的概率**分布函数**

- 非负性：$F(\boldsymbol{x}) \in [0, 1]$ 且满足 $F(-\infty) = 0$ 和 $F(\infty) = 1$
- 单调性：$F(\boldsymbol{x})$ 单调不减，也即是 $F(\boldsymbol{x}) \le F(\boldsymbol{x}')$ 当 $\boldsymbol{x} \le \boldsymbol{x}'$

分布函数

考虑随机向量 $\boldsymbol{X} = (X_1, X_2, \ldots, X_n)$ 和实变量 $\boldsymbol{x} = (x_1, x_2, \ldots, x_n)$，函数 $F(\boldsymbol{x}) = P(\boldsymbol{X} \le \boldsymbol{x})$ 为 $\boldsymbol{x}$ 的概率**分布函数**

- 非负性：$F(\boldsymbol{x}) \in [0, 1]$ 且满足 $F(-\infty) = 0$ 和 $F(\infty) = 1$
- 单调性：$F(\boldsymbol{x})$ 单调不减，也即是 $F(\boldsymbol{x}) \le F(\boldsymbol{x}')$ 当 $\boldsymbol{x} \le \boldsymbol{x}'$
- 右连续：$\lim_{\varepsilon \to 0^+} F(\boldsymbol{x} + \varepsilon) = F(\boldsymbol{x})$

分布函数

考虑随机向量 $\boldsymbol{X} = (X_1, X_2, \ldots, X_n)$ 和实变量 $\boldsymbol{x} = (x_1, x_2, \ldots, x_n)$，函数 $F(\boldsymbol{x}) = P(\boldsymbol{X} \leq \boldsymbol{x})$ 为 $\boldsymbol{x}$ 的概率**分布函数**

- 非负性：$F(\boldsymbol{x}) \in [0, 1]$ 且满足 $F(-\infty) = 0$ 和 $F(\infty) = 1$
- 单调性：$F(\boldsymbol{x})$ 单调不减，也即是 $F(\boldsymbol{x}) \leq F(\boldsymbol{x}')$ 当 $\boldsymbol{x} \leq \boldsymbol{x}'$
- 右连续：$\lim_{\varepsilon \to 0^+} F(\boldsymbol{x} + \varepsilon) = F(\boldsymbol{x})$
- 给定 $x_1 \leq x_1', x_2 < x_2'$，我们有
  $F(x_1, x_2) - F(x_1', x_2) - F(x_1, x_2') + F(x_1', x_2') \geq 0$
  注意到 $F(x_1, x_2) - F(x_1', x_2) - F(x_1, x_2') + F(x_1', x_2') =$
  $P(\boldsymbol{x} \leq \boldsymbol{X} \leq \boldsymbol{x}') \geq 0$

边缘分布与条件分布

边缘分布函数

记 (X, Y) 的分布函数为 $F(x, y)$，则 X, Y 的边缘分布函数为

$$F_X(x) = P(X \leq x) = F(x, \infty), \quad F_Y(y) = P(Y \leq y) = F(\infty, y)$$

边缘分布与条件分布

边缘分布函数

记 (X, Y) 的分布函数为 $F(x, y)$，则 X, Y 的**边缘分布函数**为

$$F_X(x) = P(X \leq x) = F(x, \infty),\ F_Y(y) = P(Y \leq y) = F(\infty, y)$$

条件分布函数

记 (X, Y) 的分布函数为 $F(x, y)$，当 $P(y < Y \leq y + \varepsilon) > 0, \forall \varepsilon > 0$，则 X 相对 $Y = y$ 的**条件分布函数**为

$$\begin{aligned} F_{X|Y}(x|y) &= \lim_{\varepsilon \to 0+} P(X \leq x | y < Y \leq y + \varepsilon) \\ &= \lim_{\varepsilon \to 0+} \frac{P(X \leq x, y < Y \leq y + \varepsilon)}{P(y < Y \leq y + \varepsilon)} \end{aligned}$$

记为 $P(X \leq x | Y = y)$

联合概率密度

随机向量 $\boldsymbol{X} = (X_1, X_2, \ldots, X_n)$ 的概率分布函数 $F(\boldsymbol{x})$ 满足 $F(\boldsymbol{x}) = \int_{-\infty}^{\boldsymbol{x}} f(\boldsymbol{x}) d\boldsymbol{x}$，且 f 非负归一，则称 $\boldsymbol{X}$ 为连续型随机变量，称 $f(\boldsymbol{x})$ 为联合概率密度

联合概率密度

联合概率密度

随机向量 $\boldsymbol{X} = (X_1, X_2, \ldots, X_n)$ 的概率分布函数 $F(\boldsymbol{x})$ 满足 $F(\boldsymbol{x}) = \int_{-\infty}^{\boldsymbol{x}} f(\boldsymbol{x}) d\boldsymbol{x}$，且 f 非负归一，则称 $\boldsymbol{X}$ 为连续型随机变量，称 $f(\boldsymbol{x})$ 为联合概率密度

- 非负性：$f(\boldsymbol{x}) \geq 0$
- 归一性：$\int_{-\infty}^{\infty} f(\boldsymbol{x}) d\boldsymbol{x} = 1$

联合概率密度

随机向量 $\boldsymbol{X}=(X_1,X_2,\ldots,X_n)$ 的概率分布函数 $F(\boldsymbol{x})$ 满足 $F(\boldsymbol{x})=\int_{-\infty}^{\boldsymbol{x}} f(\boldsymbol{x})d\boldsymbol{x}$，且 f 非负归一，则称 $\boldsymbol{X}$ 为连续型随机变量，称 $f(\boldsymbol{x})$ 为联合概率密度

- 非负性：$f(\boldsymbol{x})\geq 0$
- 归一性：$\int_{-\infty}^{\infty} f(\boldsymbol{x})d\boldsymbol{x}=1$
- 在 $f(\boldsymbol{x})$ 的连续点，我们有 $\frac{\partial^n F(\boldsymbol{x})}{\partial x_1\partial x_2\ldots\partial x_n}=f(\boldsymbol{x})$

联合概率密度

联合概率密度

随机向量 $\boldsymbol{X}=(X_1,X_2,\ldots,X_n)$ 的概率分布函数 $F(\boldsymbol{x})$ 满足 $F(\boldsymbol{x})=\int_{-\infty}^{\boldsymbol{x}} f(\boldsymbol{x})d\boldsymbol{x}$，且 f 非负归一，则称 $\boldsymbol{X}$ 为连续型随机变量，称 $f(\boldsymbol{x})$ 为联合概率密度

- 非负性：$f(\boldsymbol{x})\geq 0$
- 归一性：$\int_{-\infty}^{\infty} f(\boldsymbol{x})d\boldsymbol{x}=1$
- 在 $f(\boldsymbol{x})$ 的连续点，我们有 $\frac{\partial^n F(\boldsymbol{x})}{\partial x_1\partial x_2\ldots\partial x_n}=f(\boldsymbol{x})$
- 假设 G 是 n 维向量空间的区域，则 $P(\boldsymbol{x}\in G)=\int_{\boldsymbol{x}\in G} f(\boldsymbol{x})d\boldsymbol{x}$

例

- 例：设二元随机变量 (X,Y) 具有概率密度

  $f(x,y)=\begin{cases} cy & x\in(0,1), y\in(x^2,x) \\ 0 & else \end{cases}$ (1) 求 c

  (2) 求 (X,Y) 的联合分布函数

  (3) 求 $P(X>0.5)$

- 例：设二元随机变量 (X, Y) 具有概率密度

$$f(x,y) = \begin{cases} cy & x \in (0,1), y \in (x^2, x) \\ 0 & else \end{cases}$$ (1) 求 c

(2) 求 (X, Y) 的联合分布函数

(3) 求 $P(X > 0.5)$

(1) $c = 15$

$$(2)\ F(x,y) = \begin{cases} 0 & x \leq 0 \text{ or } y \leq 0 \\ 1 & x \geq 1 \text{ and } y \geq 1 \\ \frac{15}{2}(xy^2 - 2y^3/3 - x^5/5) & 0 < x^2 < y \leq x < 1 \\ \frac{15}{2}(-2y^3/3 + 4y^{2.5}/5) & y \leq x^2 \text{ and } 0 < y \geq 1 \\ \frac{15}{2}(x^3/3 - x^5/5) & x \leq y \text{ and } 0 < x \geq 1 \end{cases}$$

(3) $P(X > 0.5) = 47/64$

边缘概率密度

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

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

边缘概率密度

边缘概率密度

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

$$F_X(x) = F(x, \infty) = \int_{-\infty}^{x} \left[\int_{-\infty}^{\infty} f(x, y)dy\right] dx \tag{2}$$

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

$$f_X(x) = \int_{-\infty}^{\infty} f(x, y)dy. \tag{3}$$

例

- 例：设二元随机变量 (X, Y) 具有概率密度
  $$f(x,y)=\begin{cases}15y & x\in(0,1), y\in(x^2,x)\\ 0 & else\end{cases}$$
  求 (X, Y) 的边缘概率密度

例

- 例：设二元随机变量 (X, Y) 具有概率密度
  $$f(x,y) = \begin{cases} 15y & x \in (0,1), y \in (x^2, x) \\ 0 & else \end{cases}$$
  求 (X, Y) 的边缘概率密度

$$f_X(x) = \begin{cases} 15(x^2 - x^4)/2 & x \in (0,1) \\ 0 & else \end{cases}$$

$$f_Y(y) = \begin{cases} 15(y^{3/2} - y^2) & y \in (0,1) \\ 0 & else \end{cases}$$

条件概率密度

PEKING UNIVERSITY

条件概率密度

记 (X, Y) 的概率密度为 $f(x, y)$，若 $f_Y(y) > 0$ 或若 $f_X(x) > 0$，则 X 相对 $Y = y$ 的条件概率密度为

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

条件概率密度

条件概率密度

记 (X,Y) 的概率密度为 $f(x,y)$，若 $f_Y(y)>0$ 或若 $f_X(x)>0$，则 X 相对 $Y=y$ 的条件概率密度为

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

$$\begin{aligned}
F_{X|Y}(x|y) &= \lim_{\varepsilon\to 0^+}\frac{P(X\le x, y<Y\le y+\varepsilon)}{P(y<Y\le y+\varepsilon)}\\
&= \lim_{\varepsilon\to 0^+}\frac{F(x,y+\varepsilon)-F(x,y)}{F_Y(y+\varepsilon)-F_Y(y)}\quad (\text{divide by } \varepsilon)\\
&= \frac{\partial F(x,y)/\partial y}{\partial F_Y(y)/\partial y}=\frac{\partial\int_{-\infty}^{x}\int_{-\infty}^{y} f(x,y)dxdy/\partial y}{\partial\int_{-\infty}^{y} f_Y(y)dy/\partial y}\\
&= \int_{-\infty}^{x} f(x,y)/f_Y(y)dx
\end{aligned}$$

例子

- 例：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) 的联合概率密度以及条件概率密度；(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。

例子

PEKING UNIVERSITY

- 例：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) 的联合概率密度以及条件概率密度；(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。

$$f(x) = \begin{cases} 1/30 & x \in (1,30) \\ 0 & else \end{cases} \quad f_{Y|X}(y|x) = \begin{cases} 1/(30-x) & y \in (x,30) \\ 0 & else \end{cases}$$

$$f(x,y) = \begin{cases} 1/[30(30-x)] & x \in (0,30), y \in (x,30) \\ 0 & else \end{cases}$$

$$f_Y(y) = \int_{-\infty}^{y} f(x,y)dx = \begin{cases} \ln(30/(30-y))/30 & y \in (0,30) \\ 0 & else \end{cases}$$

$$f_{X|Y}(x|y) = f(x,y)/f_Y(y) = \begin{cases} 1/[(30-x)\ln(30/(30-y))] & y \in (x,30) \\ 0 & else \end{cases}$$

$$P(X \le 10 | Y = 25) = \int_{-\infty}^{10} f_{X|Y}(x|25)dx = \ln(3/2)/\ln 6$$

二元正态分布

二元正态分布

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 $N(\mu_1, \mu_2, \sigma_1^2, \sigma_2^2, \rho)$

二元正态分布

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 $N(\mu_1, \mu_2, \sigma_1^2, \sigma_2^2, \rho)$

- $N(0, 0, 1, 1, 0)$ 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 $X \sim N(\mu_1, \sigma_1^2)$. 同理 $Y \sim N(\mu_2, \sigma_2^2)$
- $f_{X|Y}(x|y) = \frac{1}{\sqrt{2\pi}\sigma_1\sqrt{1-\rho^2}} e^{\frac{-\left[x-(\mu_1+\rho\frac{\sigma_1}{\sigma_2}(y-\mu_2))\right]^2}{2(1-\rho^2)\sigma_1^2}}$，也即是在 $Y = y$ 条件下 $X \sim N(\mu_1 + \rho\frac{\sigma_1}{\sigma_2}(y - \mu_2), (1 - \rho^2)\sigma_1^2)$，同理在 $X = x$ 条件下 $Y \sim N(\mu_2 + \rho\frac{\sigma_2}{\sigma_1}(x - \mu_1), (1 - \rho^2)\sigma_2^2)$

目录

二元随机变量的独立性

二元随机变量的独立性

记 (X, Y) 的分布函数为 $F(x, y)$，边缘分布分别为 $F_X(x)$ 和 $F_Y(y)$，若对所有 x, y 有 $F(x, y) = F_X(x) \cdot F_Y(y)$，则称 X, Y 相互独立

二元随机变量的独立性

记 (X, Y) 的分布函数为 $F(x, y)$，边缘分布分别为 $F_X(x)$ 和 $F_Y(y)$，若对所有 x, y 有 $F(x, y) = F_X(x) \cdot F_Y(y)$，则称 X, Y 相互独立

- 离散型：离散型随机变量 (X, Y) 相互独立等价于概率 $P(X = x, Y = y) = P_X(X = x) \cdot P_Y(Y = y)$ 对于所有 x, y 成立

二元随机变量的独立性

记 (X, Y) 的分布函数为 $F(x, y)$，边缘分布分别为 $F_X(x)$ 和 $F_Y(y)$，若对所有 x, y 有 $F(x, y) = F_X(x) \cdot F_Y(y)$，则称 X, Y 相互独立

- 离散型：离散型随机变量 (X, Y) 相互独立等价于概率 $P(X = x, Y = y) = P_X(X = x) \cdot P_Y(Y = y)$ 对于所有 x, y 成立
- 连续型：连续型随机变量 (X, Y) 相互独立等价于概率密度 $f(X = x, Y = y) = f_X(X = x) \cdot f_Y(Y = y)$ 对于（几乎）所有 x, y 成立（除去面积为零的区域都成立）

例子

- 给定概率密度 $f(x,y) = \begin{cases} 6e^{-2x-3y} & x, y > 0 \\ 0 & else \end{cases}$ 问随机变量是否独立?

例子

- 给定概率密度 $f(x,y) = \begin{cases} 6e^{-2x-3y} & x,y>0 \\ 0 & else \end{cases}$ 问随机变量是否独立?
- 已知随机变量 X, Y 相互独立，且服从同一分布 $f(x) = \begin{cases} \frac{1}{2}e^{-2x} & x>0 \\ 0 & else \end{cases}$，问 $P(X \le 2Y)$?

例子

- 给定概率密度 $f(x,y)=\begin{cases}6e^{-2x-3y} & x,y>0\\ 0 & else\end{cases}$ 问随机变量是否独立?
- 已知随机变量 X,Y 相互独立，且服从同一分布 $f(x)=\begin{cases}\frac{1}{2}e^{-2x} & x>0\\ 0 & else\end{cases}$，问 $P(X\le 2Y)$?

$$f(x,y)=f_X(x)f_Y(Y)=\begin{cases}\frac{1}{4}e^{-2(x+y)} & x,y>0\\ 0 & else\end{cases}$$

$$P(X\le 2Y)=\int_{-\infty}^{\infty}dx\int_{x/2}^{\infty}f(x,y)dy=2/3$$

例子

- 给定概率密度 $f(x,y)=\begin{cases}6e^{-2x-3y} & x,y>0\\ 0 & else\end{cases}$ 问随机变量是否独立?
- 已知随机变量 X,Y 相互独立，且服从同一分布 $f(x)=\begin{cases}\frac{1}{2}e^{-2x} & x>0\\ 0 & else\end{cases}$，问 $P(X\le 2Y)$?

  $f(x,y)=f_X(x)f_Y(Y)=\begin{cases}\frac{1}{4}e^{-2(x+y)} & x,y>0\\ 0 & else\end{cases}$

  $P(X\le 2Y)=\int_{-\infty}^{\infty}dx\int_{x/2}^{\infty}f(x,y)dy=2/3$
- 二元正态分布 $N(\mu_1,\mu_2,\sigma_1^2,\sigma_2^2,\rho)$ 相互独立的充要条件是 $\rho=0$

例子

- 某条蚕的产卵数 X 服从泊松分布 $X \sim \pi(\lambda)$，每个卵变为成虫的概率为 p，各卵是否变为成虫彼此独立. 设成虫数和死卵数分别为 Y, Z，Y, Z 是否相互独立.

- 某条蚕的产卵数 X 服从泊松分布 $X \sim \pi(\lambda)$，每个卵变为成虫的概率为 p，各卵是否变为成虫彼此独立. 设成虫数和死卵数分别为 Y, Z，Y, Z 是否相互独立.
- 考虑 Y, Z 的联合分布：

例子

- 某条蚕的产卵数 X 服从泊松分布 $X \sim \pi(\lambda)$，每个卵变为成虫的概率为 p，各卵是否变为成虫彼此独立. 设成虫数和死卵数分别为 Y, Z，Y, Z 是否相互独立.
- 考虑 Y, Z 的联合分布：

$$\begin{aligned} P(Y = k, Z = l) &= P(Y = k, X = k + l) \\ &= P(X = k + l) P(Y = k \mid X = k + l) \\ &= \frac{\lambda^{k+l}}{(k + l)!} e^{-\lambda} \cdot \frac{(k + l)!}{k!\, l!} p^k (1 - p)^l \\ &= \frac{(\lambda p)^k}{k!} e^{-\lambda p} \cdot \frac{(\lambda(1 - p))^l}{l!} e^{-\lambda(1-p)} \end{aligned}$$

联合分布等于边缘分布乘积，所以 Y, Z 相互独立.

多元随机变量的独立性

记 $\boldsymbol{X} = (X_1, X_2, \ldots, X_n)$ 的分布函数为 $F(\boldsymbol{x})$，$\boldsymbol{x} = (x_1, x_2, \ldots, x_n)$，边缘分布分别为 $F_{X_i}(x_i)$，若对所有 $\boldsymbol{x}$ 有 $F(\boldsymbol{X}) = F_{X_1}(x_1) \cdot F_{X_2}(x_2) \cdot \cdots \cdot F_{X_n}(x_n)$，则称 $X_1, X_2, \ldots, X_n$ **相互独立**

多元随机变量的独立性

记 $\boldsymbol{X} = (X_1, X_2, \ldots, X_n)$ 的分布函数为 $F(\boldsymbol{x})$，
$\boldsymbol{x} = (x_1, x_2, \ldots, x_n)$，边缘分布分别为 $F_{X_i}(x_i)$，若对所有 $\boldsymbol{x}$ 有
$F(\boldsymbol{X}) = F_{X_1}(x_1) \cdot F_{X_2}(x_2) \cdot \cdots \cdot F_{X_n}(x_n)$，则称 $X_1, X_2, \ldots, X_n$
相互独立
若 $\boldsymbol{X}$ 为连续性随机向量，则相互独立等价于
$f(\boldsymbol{X}) = f_{X_1}(x_1) \cdot f_{X_2}(x_2) \cdot \cdots \cdot f_{X_n}(x_n)$

多元随机变量的独立性

记 $\boldsymbol{X}=(X_1,X_2,\ldots,X_n)$ 的分布函数为 $F(\boldsymbol{x})$，$\boldsymbol{x}=(x_1,x_2,\ldots,x_n)$，边缘分布分别为 $F_{X_i}(x_i)$，若对所有 $\boldsymbol{x}$ 有 $F(\boldsymbol{X})=F_{X_1}(x_1)\cdot F_{X_2}(x_2)\cdot\cdots\cdot F_{X_n}(x_n)$，则称 $X_1,X_2,\ldots,X_n$ **相互独立**

若 $\boldsymbol{X}$ 为连续性随机向量，则相互独立等价于
$f(\boldsymbol{X})=f_{X_1}(x_1)\cdot f_{X_2}(x_2)\cdot\cdots\cdot f_{X_n}(x_n)$

- 若 $X_1,X_2,\ldots,X_n$ 两两独立，是否有 $X_1,X_2,\ldots,X_n$ 相互独立?

多元随机变量的独立性

记 $\boldsymbol{X} = (X_1, X_2, \ldots, X_n)$ 的分布函数为 $F(\boldsymbol{x})$，
$\boldsymbol{x} = (x_1, x_2, \ldots, x_n)$，边缘分布分别为 $F_{X_i}(x_i)$，若对所有 $\boldsymbol{x}$ 有
$F(\boldsymbol{X}) = F_{X_1}(x_1) \cdot F_{X_2}(x_2) \cdot \cdots \cdot F_{X_n}(x_n)$，则称 $X_1, X_2, \ldots, X_n$
相互独立
若 $\boldsymbol{X}$ 为连续性随机向量，则相互独立等价于
$f(\boldsymbol{X}) = f_{X_1}(x_1) \cdot f_{X_2}(x_2) \cdot \cdots \cdot f_{X_n}(x_n)$

- 若 $X_1, X_2, \ldots, X_n$ 两两独立，是否有 $X_1, X_2, \ldots, X_n$ 相互独立?
- 考虑 0-1 随机变量且 $p(X_1 = x_1, X_2 = x_2, X_3 = x_3) = p_{X_1}(x_1)p_{X_2}(x_2)p_{X_3}(x_3)(1 + f(x_1)f(x_2)f(x_3))$，同时满足 $f(0) + f(1) = 0$，$p_{X_{1,2,3}}(x_{1,2,3}) = 1/2$

多元随机变量的独立性

PEKING UNIVERSITY

多元随机变量的独立性

记 $\boldsymbol{X}=(X_1,X_2,\ldots,X_n)$ 的分布函数为 $F(\boldsymbol{x})$，
$\boldsymbol{x}=(x_1,x_2,\ldots,x_n)$，边缘分布分别为 $F_{X_i}(x_i)$，若对所有 $\boldsymbol{x}$ 有
$F(\boldsymbol{X})=F_{X_1}(x_1)\cdot F_{X_2}(x_2)\cdot\cdots\cdot F_{X_n}(x_n)$，则称 $X_1,X_2,\ldots,X_n$ **相互独立**
若 $\boldsymbol{X}$ 为连续性随机向量，则相互独立等价于
$f(\boldsymbol{X})=f_{X_1}(x_1)\cdot f_{X_2}(x_2)\cdot\cdots\cdot f_{X_n}(x_n)$

- 若 $X_1,X_2,\ldots,X_n$ 两两独立，是否有 $X_1,X_2,\ldots,X_n$ 相互独立？
- 考虑 0-1 随机变量且 $p(X_1=x_1,X_2=x_2,X_3=x_3)=p_{X_1}(x_1)p_{X_2}(x_2)p_{X_3}(x_3)(1+f(x_1)f(x_2)f(x_3))$，同时满足 $f(0)+f(1)=0$，$p_{X_{1,2,3}}(x_{1,2,3})=1/2$
- 若 $X_1,X_2,\ldots,X_n$ 相互独立，是否有 $X_1,X_2,\ldots,X_n$ 两两独立？

多元随机变量的独立性

记 $\boldsymbol{X} = (X_1, X_2, \ldots, X_n)$ 的分布函数为 $F(\boldsymbol{x})$，$\boldsymbol{x} = (x_1, x_2, \ldots, x_n)$，边缘分布分别为 $F_{X_i}(x_i)$，若对所有 $\boldsymbol{x}$ 有 $F(\boldsymbol{X}) = F_{X_1}(x_1) \cdot F_{X_2}(x_2) \cdot \cdots \cdot F_{X_n}(x_n)$，则称 $X_1, X_2, \ldots, X_n$ **相互独立**

若 $\boldsymbol{X}$ 为连续性随机向量，则相互独立等价于
$f(\boldsymbol{X}) = f_{X_1}(x_1) \cdot f_{X_2}(x_2) \cdot \cdots \cdot f_{X_n}(x_n)$

- 若 $X_1, X_2, \ldots, X_n$ 两两独立，是否有 $X_1, X_2, \ldots, X_n$ 相互独立?
- 考虑 0-1 随机变量且 $p(X_1 = x_1, X_2 = x_2, X_3 = x_3) = p_{X_1}(x_1)p_{X_2}(x_2)p_{X_3}(x_3)(1 + f(x_1)f(x_2)f(x_3))$，同时满足 $f(0) + f(1) = 0$，$p_{X_{1,2,3}}(x_{1,2,3}) = 1/2$
- 若 $X_1, X_2, \ldots, X_n$ 相互独立，是否有 $X_1, X_2, \ldots, X_n$ 两两独立?
- 注意与事件独立的区别

随机变量集合的独立性

记 $\boldsymbol{X}=(X_1,X_2,\ldots,X_n)$ 的分布函数为 $F_{\boldsymbol{X}}(\boldsymbol{x})$，$\boldsymbol{x}=(x_1,x_2,\ldots,x_n)$，边缘分布分别为 $F_{X_i}(x_i)$，记 $\boldsymbol{Y}=(Y_1,Y_2,\ldots,Y_m)$ 的分布函数为 $F_{\boldsymbol{Y}}(\boldsymbol{y})$，$\boldsymbol{y}=(y_1,y_2,\ldots,y_m)$，边缘分布分别为 $F_{Y_i}(y_i)$，若对所有 $\boldsymbol{x},\boldsymbol{y}$ 有 $F(\boldsymbol{x},\boldsymbol{y})=F_{\boldsymbol{X}}(\boldsymbol{x})F_{\boldsymbol{Y}}(\boldsymbol{y})$，则称 $\boldsymbol{X}$ 与 $\boldsymbol{Y}$ 相互独立

随机变量集合的独立性

记 $\boldsymbol{X}=(X_1,X_2,\ldots,X_n)$ 的分布函数为 $F_{\boldsymbol{X}}(\boldsymbol{x})$，$\boldsymbol{x}=(x_1,x_2,\ldots,x_n)$，边缘分布分别为 $F_{X_i}(x_i)$，记 $\boldsymbol{Y}=(Y_1,Y_2,\ldots,Y_m)$ 的分布函数为 $F_{\boldsymbol{Y}}(\boldsymbol{y})$，$\boldsymbol{y}=(y_1,y_2,\ldots,y_m)$，边缘分布分别为 $F_{Y_i}(y_i)$，若对所有 $\boldsymbol{x},\boldsymbol{y}$ 有 $F(\boldsymbol{x},\boldsymbol{y})=F_{\boldsymbol{X}}(\boldsymbol{x})F_{\boldsymbol{Y}}(\boldsymbol{y})$，则称 $\boldsymbol{X}$ 与 $\boldsymbol{Y}$ 相互独立

- 若 $\boldsymbol{X}$ 与 $\boldsymbol{Y}$ 相互独立，则 X_i 和 Y_j 相互独立

随机变量集合的独立性

随机变量集合的独立性

记 $\boldsymbol{X} = (X_1, X_2, \ldots, X_n)$ 的分布函数为 $F_{\boldsymbol{X}}(\boldsymbol{x})$，
$\boldsymbol{x} = (x_1, x_2, \ldots, x_n)$，边缘分布分别为 $F_{X_i}(x_i)$，记
$\boldsymbol{Y} = (Y_1, Y_2, \ldots, Y_m)$ 的分布函数为 $F_{\boldsymbol{Y}}(\boldsymbol{y})$，
$\boldsymbol{y} = (y_1, y_2, \ldots, y_m)$，边缘分布分别为 $F_{Y_i}(y_i)$，若对所有 $\boldsymbol{x}, \boldsymbol{y}$
有 $F(\boldsymbol{x}, \boldsymbol{y}) = F_{\boldsymbol{X}}(\boldsymbol{x})F_{\boldsymbol{Y}}(\boldsymbol{y})$，则称 $\boldsymbol{X}$ 与 $\boldsymbol{Y}$ 相互独立

- 若 $\boldsymbol{X}$ 与 $\boldsymbol{Y}$ 相互独立，则 X_i 和 Y_j 相互独立
- 若任意 X_i 和 Y_j 相互独立，$\boldsymbol{X}$ 与 $\boldsymbol{Y}$ 不一定相互独立

随机变量集合的独立性

记 $\boldsymbol{X} = (X_1, X_2, \ldots, X_n)$ 的分布函数为 $F_{\boldsymbol{X}}(\boldsymbol{x})$，$\boldsymbol{x} = (x_1, x_2, \ldots, x_n)$，边缘分布分别为 $F_{X_i}(x_i)$，记 $\boldsymbol{Y} = (Y_1, Y_2, \ldots, Y_m)$ 的分布函数为 $F_{\boldsymbol{Y}}(\boldsymbol{y})$，$\boldsymbol{y} = (y_1, y_2, \ldots, y_m)$，边缘分布分别为 $F_{Y_i}(y_i)$，若对所有 $\boldsymbol{x}, \boldsymbol{y}$ 有 $F(\boldsymbol{x}, \boldsymbol{y}) = F_{\boldsymbol{X}}(\boldsymbol{x})F_{\boldsymbol{Y}}(\boldsymbol{y})$，则称 $\boldsymbol{X}$ 与 $\boldsymbol{Y}$ 相互独立

- 若 $\boldsymbol{X}$ 与 $\boldsymbol{Y}$ 相互独立，则 X_i 和 Y_j 相互独立
- 若任意 X_i 和 Y_j 相互独立，$\boldsymbol{X}$ 与 $\boldsymbol{Y}$ 不一定相互独立
- 若 $\boldsymbol{X}$ 与 $\boldsymbol{Y}$ 相互独立，则 $h(\boldsymbol{X})$ 和 $g(\boldsymbol{Y})$ 相互独立（其中 h, g 为连续函数）

目录

多元随机变量的离散型函数

问题

问题：已知多元随机变量 $\boldsymbol{X} = (X_1, X_2, \ldots, X_n)$ 的概率分布和离散型函数 $Y = g(\boldsymbol{X})$，求 Y 的概率分布

多元随机变量的离散型函数

问题

问题：已知多元随机变量 $\boldsymbol{X}=(X_1,X_2,\ldots,X_n)$ 的概率分布和离散型函数 $Y=g(\boldsymbol{X})$，求 Y 的概率分布

多元离散随机变量函数的概率分布

对于多元随机变量 $\boldsymbol{X}=(X_1,X_2,\ldots,X_n)$，若 $Y=g(\boldsymbol{X})$ 为离散随机变量，则可以首先找出 Y 的所有可能值，再找出每个值对应的等价事件来求出概率

- 已知随机变量 X, Y 的概率分布

X\Y	Y=0	Y=1
X=0	0.2	0.1
X=1	0.3	0.4

令 $U = X + Y$, $V = \max\{X, Y\}$, 求 (U, V) 的联合概率分布

例子

PEKING UNIVERSITY

- 已知随机变量 X, Y 的概率分布

X\Y	Y=0	Y=1
X=0	0.2	0.1
X=1	0.3	0.4

令 $U = X + Y$, $V = \max\{X, Y\}$, 求 (U, V) 的联合概率分布

U\V	V=0	V=1
U=0	0.2	0
U=1	0	0.4
U=2	0	0.4

例子

- 已知随机变量 X 的概率密度为 $f_X(x) = \begin{cases} e^{-x} & x > 0 \\ 0 & x \leq 0 \end{cases}$ 令 $U = \begin{cases} 1 & X > 1 \\ 0 & X \leq 1 \end{cases}$ 和 $V = \begin{cases} 1 & X > 2 \\ 0 & X \leq 2 \end{cases}$ 求 (U, V) 的联合概率

- 已知随机变量 X 的概率密度为 $f_X(x) = \begin{cases} e^{-x} & x > 0 \\ 0 & x \leq 0 \end{cases}$ 令 $U = \begin{cases} 1 & X > 1 \\ 0 & X \leq 1 \end{cases}$ 和 $V = \begin{cases} 1 & X > 2 \\ 0 & X \leq 2 \end{cases}$ 求 (U, V) 的联合概率

U\V	V=0	V=1
U=0	$1 - e^{-1}$	0
U=1	$e^{-1} - e^{-2}$	e^{-2}

例子

- 已知 $X \sim \pi(\lambda_x)$, $Y \sim \pi(\lambda_y)$, 求 $Z = X + Y$ 的概率分布

- 已知 $X \sim \pi(\lambda_x)$, $Y \sim \pi(\lambda_y)$, 求 $Z = X + Y$ 的概率分布
  $P(X = i) = e^{-\lambda_x}\lambda_x^i/i!$, $P(Y = j) = e^{-\lambda_y}\lambda_y^j/j!$, 因此

$$
\begin{aligned}
P(Z = k) &= \sum_{i=0}^{k} P(X = i)P(Y = k - i) \\
&= \sum_{i=0}^{k} \frac{e^{-\lambda_x}\lambda_x^i}{i!}\frac{e^{-\lambda_y}\lambda_y^{k-i}}{(k-i)!}, \\
&= \frac{e^{-(\lambda_x+\lambda_y)}}{k!}\sum_{i=0}^{k}\frac{k!}{i!(k-i)!}\lambda_x^i\lambda_y^{k-i}, \\
&= \frac{e^{-(\lambda_x+\lambda_y)}(\lambda_x+\lambda_y)^k}{k!}
\end{aligned}
\tag{4}
$$

也即是 $Z \sim \pi(\lambda_x + \lambda_y)$

和的分布

对于多元随机变量 $\boldsymbol{X}=(X_1,X_2,\ldots,X_n)$ 的连续型函数 $Y=g(\boldsymbol{X})$，我们可以用分布函数来分析 Y 的概率分布

对于多元随机变量 $\boldsymbol{X} = (X_1, X_2, \ldots, X_n)$ 的连续型函数 $Y = g(\boldsymbol{X})$，我们可以用分布函数来分析 Y 的概率分布

$Z = X + Y$

已知 (X, Y) 的概率密度为 $f(x, y)$，$Z = X + Y$ 的概率密度为
$f_Z(z) = \int_{-\infty}^{\infty} f(x, z - x)dx = \int_{-\infty}^{\infty} f(z - y, y)dy$

对于多元随机变量 $\boldsymbol{X}=(X_1,X_2,\ldots,X_n)$ 的连续型函数 $Y=g(\boldsymbol{X})$，我们可以用分布函数来分析 Y 的概率分布

$Z=X+Y$

已知 (X,Y) 的概率密度为 $f(x,y)$，$Z=X+Y$ 的概率密度为
$f_Z(z)=\int_{-\infty}^{\infty}f(x,z-x)dx=\int_{-\infty}^{\infty}f(z-y,y)dy$

$$
\begin{aligned}
F_Z(z)=P(Z\le z)&=\int_{x+y\le z} f(x,y)dxdy,\\
&=\int_{-\infty}^{\infty}dx\int_{-\infty}^{z-x}f(x,y)dy,\\
&=\int_{-\infty}^{\infty}dx\int_{-\infty}^{z}f(x,\tilde{y}-x)d\tilde{y},\ \ (\tilde{y}=x+y)\\
&=\int_{-\infty}^{z}d\tilde{y}\int_{-\infty}^{\infty}f(x,\tilde{y}-x)dx
\end{aligned}
$$

因此 $f_Z(z)=\int_{-\infty}^{\infty}f(x,z-x)dx$

卷积公式

已知 (X, Y) 相互独立，概率密度分别为 $f_X(x), f_Y(y)$，
$Z = X + Y$ 的概率密度为
$f_Z(z) = f_X * f_Z = \int_{-\infty}^{\infty} f_X(x) f_Y(z-x)dx = \int_{-\infty}^{\infty} f_X(z-y) f_Y(y)dy$

卷积公式

卷积公式

已知 (X,Y) 相互独立，概率密度分别为 $f_X(x), f_Y(y)$，
$Z = X + Y$ 的概率密度为
$f_Z(z) = f_X * f_Z = \int_{-\infty}^{\infty} f_X(x) f_Y(z-x)dx = \int_{-\infty}^{\infty} f_X(z-y) f_Y(y)dy$

- 例：已知 (X,Y) 的概率密度为 $f(x,y)$，求 $Z = X - Y$ 的概率密度

卷积公式

已知 (X, Y) 相互独立，概率密度分别为 $f_X(x), f_Y(y)$，
$Z = X + Y$ 的概率密度为
$f_Z(z) = f_X * f_Z = \int_{-\infty}^{\infty} f_X(x) f_Y(z-x)dx = \int_{-\infty}^{\infty} f_X(z-y) f_Y(y)dy$

- 例：已知 (X, Y) 的概率密度为 $f(x, y)$，求 $Z = X - Y$ 的概率密度
  $(X, -Y)$ 的概率密度为 $f(x, -y)$，因此 $Z = X - Y$ 的概率密度为 $f_Z(z) = \int_{-\infty}^{\infty} f(x, x-z)dx = \int_{-\infty}^{\infty} f(z+y, y)dy$

卷积公式

卷积公式

已知 (X, Y) 相互独立，概率密度分别为 $f_X(x), f_Y(y)$，
$Z = X + Y$ 的概率密度为
$f_Z(z) = f_X * f_Z = \int_{-\infty}^{\infty} f_X(x) f_Y(z-x)dx = \int_{-\infty}^{\infty} f_X(z-y) f_Y(y)dy$

- 例：已知 X, Y 相互独立且都服从标准正态分布 $N(0, 1)$，求 $Z = X + Y$ 的概率密度.

卷积公式

卷积公式

已知 (X, Y) 相互独立，概率密度分别为 $f_X(x), f_Y(y)$，
$Z = X + Y$ 的概率密度为
$f_Z(z) = f_X * f_Z = \int_{-\infty}^{\infty} f_X(x) f_Y(z-x)dx = \int_{-\infty}^{\infty} f_X(z-y) f_Y(y)dy$

- 例：已知 X, Y 相互独立且都服从标准正态分布 $N(0,1)$，求 $Z = X + Y$ 的概率密度.
  由卷积公式：

$$f_Z(z) = \int_{-\infty}^{\infty} f_X(x) f_Y(z-x)dx = \frac{1}{2\pi}\int_{-\infty}^{\infty} e^{-x^2/2} e^{-(z-x)^2/2}dx$$

$$= \frac{1}{2\sqrt{\pi}} e^{-z^2/4}.$$

卷积公式

.

卷积公式

已知 (X, Y) 相互独立，概率密度分别为 $f_X(x), f_Y(y)$，
$Z = X + Y$ 的概率密度为

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

- 例：已知 X, Y 相互独立且都服从参数为 1 的指数分布，求 $Z = X + Y$ 的概率密度.

卷积公式

.

卷积公式

已知 (X, Y) 相互独立，概率密度分别为 $f_X(x), f_Y(y)$，
$Z = X + Y$ 的概率密度为
$f_Z(z) = f_X * f_Z = \int_{-\infty}^{\infty} f_X(x) f_Y(z-x) dx = \int_{-\infty}^{\infty} f_X(z-y) f_Y(y) dy$

- 例：已知 X, Y 相互独立且都服从参数为 1 的指数分布，求 $Z = X + Y$ 的概率密度.
  注意 X, Y 的取值都是正数. 当 $z \leq 0$ 时显然有 $f_Z(z) = 0$.
  当 $z > 0$ 时，由卷积公式:

$$f_Z(z) = \int_0^z f_X(x) f_Y(z-x) dx = \int_0^z e^{-z} dx = ze^{-z}.$$

卷积公式

卷积公式常用结论

已知 (X, Y) 相互独立

- 二项分布：若 $X \sim B(n, p)$，$Y \sim B(m, p)$，则 $X + Y \sim B(m + n, p)$
- 泊松分布：若 $X \sim \pi(\lambda_1)$，$Y \sim \pi(\lambda_2)$，则 $X + Y \sim \pi(\lambda_1 + \lambda_2)$
- 负二项分布：若 $X \sim NB(n, p)$，$Y \sim NB(m, p)$，则 $X + Y \sim NB(m + n, p)$
- Γ 分布：若 $X \sim \Gamma(\alpha_1, \lambda)$，$Y \sim \Gamma(\alpha_2, \lambda)$，则 $X + Y \sim \Gamma(\alpha_1 + \alpha_2, \lambda)$
- 正态分布：若 $X \sim N(\mu_1, \sigma_1^2)$，$Y \sim N(\mu_2, \sigma_2^2)$，则 $X + Y \sim N(\mu_1 + \mu_2, \sigma_1^2 + \sigma_2^2)$
- 考虑 0-1 分布、几何分布、指数分布的可加性
- 考虑多个独立随机变量的可加性

χ^2 分布

χ^2 分布

概率密度为 $f(x) = \begin{cases} \frac{1}{2^{n/2}\Gamma(n/2)} x^{n/2-1} e^{-x/2} & x \geq 0 \\ 0 & else \end{cases}$ 随机变量 X 被称作 χ^2 分布，记作 $\chi^2(n)$

这里 $\Gamma(\alpha) = \int_0^\infty x^{\alpha-1} e^{-x} dx, \alpha > 0$

- 注意到 $\chi^2(n) = \Gamma(n/2, 1/2)$，也即是伽马分布的特殊情况
- 若随机变量 $X \sim N(0,1)$，则 $X^2 \sim \chi^2(1)$
- 考虑 n 个独立随机变量 $X_i \sim N(0,1)$，则 $\sum_i X_i^2 \sim \chi^2(n)$
- 若 $X \sim \chi^2(n)$, $Y \sim \chi^2(m)$, 则 $X + Y \sim \chi^2(n+m)$
- 一个常用的公式为 $\int_0^\infty \frac{1}{2^{n/2}\Gamma(n/2)} x^{n/2-1} e^{-x/2} = 1.$

商的分布

已知 (X, Y) 的概率密度为 $f(x, y)$，$Z = X/Y$ 的概率密度为 $f_Z(z) = \int_{-\infty}^{\infty} |y| f(yz, y) dy$

商的分布

商的分布

已知 (X, Y) 的概率密度为 $f(x, y)$，$Z = X/Y$ 的概率密度为 $f_Z(z) = \int_{-\infty}^{\infty} |y| f(yz, y) dy$

$$
\begin{aligned}
F_Z(z) = P(Z \leq z) &= \int\limits_{x/y \leq z} f(x, y) dxdy, \\
&= \int\limits_{x/y \leq z, y>0} f(x, y) dxdy + \int\limits_{x/y \leq z, y<0} f(x, y) dxdy \\
&= \int_0^{\infty} dy \int_{-\infty}^{yz} dx f(x, y) + \int_{-\infty}^{0} dy \int_{yz}^{\infty} dx f(x, y)
\end{aligned}
$$

因此 $f_Z(z) = F_Z'(z) = \int_0^{\infty} dy y f(yz, y) - \int_{-\infty}^{0} dy y f(yzz, y) = \int_{-\infty}^{\infty} |y| f(yz, y) dy$

例子

- 已知 $X \sim N(0,1)$, $Y \sim \chi^2(n)$, 且 X, Y 相互独立，求 $Z = \dfrac{X}{\sqrt{Y/n}}$

- 已知 $X \sim N(0,1)$, $Y \sim \chi^2(n)$, 且 X, Y 相互独立，求 $Z = \frac{X}{\sqrt{Y/n}}$

首先我们有 $Y' = \sqrt{Y/n}$ 的概率密度

$$f_{Y'}(y) = \begin{cases} \frac{2(n/2)^{n/2}}{\Gamma(n/2)} y^{n-1} e^{-ny^2/2} & x \geq 0 \\ 0 & else \end{cases}$$

X 与 Y' 也独立，因此 $Z = X/Y'$ 的概率密度为

$f_Z(z) = \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}} \frac{2(n/2)^{n/2}}{\Gamma(n/2)} e^{-y^2(z^2+n)/2} y^n dy$ 令 $t = y^2(z^2+n)$, 则

$f_Z(z) = \frac{1}{\sqrt{2\pi}} \frac{2(n/2)^{n/2}}{\Gamma(n/2)} (z^2+n)^{-(n+1)/2} \int_{-\infty}^{\infty} e^{-t/2} t^{(n+1)/2-1} dy$ 注意到 $\frac{1}{2^{(n+1)/2}\Gamma((n+1)/2)} \int_{-\infty}^{\infty} e^{-t/2} t^{(n+1)/2-1} dy = 1$ 我们有

$$f_Z(z) = \frac{\Gamma((n+1)/2)}{\sqrt{n\pi}\Gamma(n/2)} (1 + y^2/n)^{-(n+1)/2}$$

min, max 分布

已知 (X, Y) 独立且服从 $U(0,1)$，$Z_1 = \min\{X, Y\}$ 和 $Z_2 = \max\{X, Y\}$ 的概率密度为 $f_{Z_1}(z) = 2(1-z)$ 和 $f_{Z_2}(z) = 2z$，其中 x 非零取值范围为 $(0,1)$

min, max 分布

已知 (X,Y) 独立且服从 $U(0,1)$，$Z_1=\min\{X,Y\}$ 和 $Z_2=\max\{X,Y\}$ 的概率密度为 $f_{Z_1}(z)=2(1-z)$ 和 $f_{Z_2}(z)=2z$，其中 x 非零取值范围为 $(0,1)$

考虑 $z\in(0,1)$，有分布函数
$F_{Z_1}(z)=P(\min\{X,Y\}\leq z)=1-P(\min\{X,Y\}\geq z)=1-P(X\geq z\,\&\,Y\geq z)=1-P(X\geq z)^2=1-(1-z)^2$，因此
$f_{Z_1}(z)=2(1-z)$

min, max 分布

已知 (X, Y) 独立且服从 $U(0,1)$，$Z_1 = \min\{X, Y\}$ 和 $Z_2 = \max\{X, Y\}$ 的概率密度为 $f_{Z_1}(z) = 2(1-z)$ 和 $f_{Z_2}(z) = 2z$，其中 x 非零取值范围为 $(0,1)$

考虑 $z \in (0,1)$，有分布函数
$F_{Z_1}(z) = P(\min\{X, Y\} \leq z) = 1 - P(\min\{X, Y\} \geq z) = 1 - P(X \geq z \,\&\, Y \geq z) = 1 - P(X \geq z)^2 = 1 - (1-z)^2$，因此 $f_{Z_1}(z) = 2(1-z)$
考虑 $z \in (0,1)$，有分布函数 $F_{Z_2}(z) = P(\max\{X, Y\} \leq z) = P(X \leq z \,\&\, Y \leq z) = P(X \leq z)^2 = z^2$，因此 $f_{Z_2}(z) = 2z$

min, max 分布

已知 $\boldsymbol{X} = (X_1, X_2, \ldots, X_n)$ 相互独立且分布函数为 $F(\boldsymbol{X}) = F_{X_1}(x_1) \cdot F_{X_2}(x_2) \cdot \cdots \cdot F_{X_n}(x_n)$，$Z_1 = \min\{X_i\}$ 和 $Z_2 = \max\{X_i\}$ 的分布函数为

$F_{Z_1}(z) = 1 - (1 - F_{X_1}(z))(1 - F_{X_2}(z)) \ldots (1 - F_{X_n}(z))$ 和

$F_{Z_2}(z) = F_{X_1}(z) F_{X_2}(z) \ldots F_{X_n}(z)$

min, max 分布

已知 $\boldsymbol{X} = (X_1, X_2, \ldots, X_n)$ 相互独立且分布函数为 $F(\boldsymbol{X}) = F_{X_1}(x_1) \cdot F_{X_2}(x_2) \cdot \cdots \cdot F_{X_n}(x_n)$，$Z_1 = \min\{X_i\}$ 和 $Z_2 = \max\{X_i\}$ 的分布函数为

$F_{Z_1}(z) = 1 - (1 - F_{X_1}(z))(1 - F_{X_2}(z)) \ldots (1 - F_{X_n}(z))$ 和

$F_{Z_2}(z) = F_{X_1}(z) F_{X_2}(z) \ldots F_{X_n}(z)$

特别地，当 $F_{X_i}(x) = F_{X_j}(x)$，则有 $F_{Z_1}(z) = 1 - (1 - F_X(z))^n$ 和 $F_{Z_2}(z) = F_X(z)^n$

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 $\boldsymbol{J}_{jk}=\partial h_j(\boldsymbol{y})/\partial y_k$

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：

设 $\boldsymbol{y}=\boldsymbol{g}(\boldsymbol{x})$，则 $P(\boldsymbol{x}\le\boldsymbol{X}\le\boldsymbol{x}+\mathrm{d}\boldsymbol{x})=P(\boldsymbol{y}\le\boldsymbol{Y}\le\boldsymbol{y}+\mathrm{d}\boldsymbol{y})$.

所以 $f_{\boldsymbol{X}}(\boldsymbol{x})\mathrm{d}\boldsymbol{x}=f_{\boldsymbol{Y}}(\boldsymbol{y})\mathrm{d}\boldsymbol{y}$.

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

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：

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

一元情形下，$\frac{\mathrm{dx}}{\mathrm{dy}} = |h'(y)|$，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.

例子

- 例 1：已知 X, Y 相互独立且都服从 $N(0,1)$ 分布，令 $X = R\cos W$，$Y = R\sin W$，其中 $R > 0, W \in [0, 2\pi)$. 求 (R, W) 的概率密度.

例子

- 例 1：已知 X, Y 相互独立且都服从 $N(0,1)$ 分布，令 $X = R\cos W$，$Y = R\sin W$，其中 $R > 0, W \in [0, 2\pi)$. 求 (R, W) 的概率密度.
  考虑映射 $h : (r, w) \mapsto (x, y)$，$x = r\cos w$，$y = r\sin w$.
  Jacobi 行列式

$$\left|\frac{\partial(x,y)}{\partial(r,w)}\right| = \begin{vmatrix} \frac{\partial x}{\partial r} & \frac{\partial x}{\partial w} \\ \frac{\partial y}{\partial r} & \frac{\partial y}{\partial w} \end{vmatrix} = r.$$

由 $f_{X,Y}(x,y)\mathrm{d}x\mathrm{d}y = f_{R,W}(r,w)\mathrm{d}r\mathrm{d}w$ 知

$$f_{R,W}(r, w) = f_{X,Y}(x,y)\left|\frac{\partial(x,y)}{\partial(r,w)}\right| = \frac{1}{2\pi} re^{-\frac{r^2}{2}}.$$

其中，$r \in (0, +\infty)$，$w \in [0, 2\pi)$.

- 例 1：已知 X, Y 相互独立且都服从 $N(0,1)$ 分布，令 $X = R\cos W$，$Y = R\sin W$，其中 $R > 0, W \in [0, 2\pi)$. 求 (R, W) 的概率密度.

$$f_{R,W}(r, w) = \begin{cases} \frac{1}{2\pi} r e^{-\frac{r^2}{2}} & r \in (0, +\infty), w \in [0, 2\pi) \\ 0 & else \end{cases}$$

可以看出：

- R, W 相互独立（联合密度恰好是边缘密度的乘积）.

例子

- 例 1：已知 X, Y 相互独立且都服从 $N(0,1)$ 分布，令 $X = R\cos W$, $Y = R\sin W$，其中 $R > 0, W \in [0, 2\pi)$. 求 (R, W) 的概率密度.

$$f_{R,W}(r, w) = \begin{cases} \frac{1}{2\pi} re^{-\frac{r^2}{2}} & r \in (0, +\infty), w \in [0, 2\pi) \\ 0 & else \end{cases}$$

可以看出：

- R, W 相互独立（联合密度恰好是边缘密度的乘积）.
- $W \sim U(0, 2\pi)$，$V = R^2 = X^2 + Y^2 \sim Exp(\frac{1}{2})$.

$$p_V(v) = p_R(r)\left|\frac{\mathrm{d}r}{\mathrm{d}v}\right| = re^{-\frac{r^2}{2}} \cdot \frac{1}{2r} = \frac{1}{2}e^{-\frac{v}{2}}$$

例子

- 例 1：已知 X, Y 相互独立且都服从 $N(0,1)$ 分布，令 $X = R\cos W$，$Y = R\sin W$，其中 $R > 0, W \in [0, 2\pi)$. 求 (R, W) 的概率密度.

$$f_{R,W}(r, w) = \begin{cases} \frac{1}{2\pi} re^{-\frac{r^2}{2}} & r \in (0, +\infty), w \in [0, 2\pi) \\ 0 & else \end{cases}$$

可以看出：

- R, W 相互独立（联合密度恰好是边缘密度的乘积）.
- $W \sim U(0, 2\pi)$，$V = R^2 = X^2 + Y^2 \sim Exp(\frac{1}{2})$.

$$p_V(v) = p_R(r)\left|\frac{\mathrm{d}r}{\mathrm{d}v}\right| = re^{-\frac{r^2}{2}} \cdot \frac{1}{2r} = \frac{1}{2}e^{-\frac{v}{2}}$$

- 这一例子刻画了二维正态向量在极坐标表示下的行为. 能否将结论推广到更高维?

例子

- 例 2：已知 X, Y 相互独立且都服从 $N(0,1)$ 分布. 求 $Z = X/Y$ 的分布.

例子

- 例 2: 已知 X, Y 相互独立且都服从 $N(0,1)$ 分布. 求 $Z = X/Y$ 的分布.
  定义映射 $g : (x, y) \mapsto (z, w)$, $z = x/y$, $w = y$. Jacobi 行列式

$$\left|\frac{\partial(z,w)}{\partial(x,y)}\right| = \begin{vmatrix} \frac{\partial z}{\partial x} & \frac{\partial z}{\partial y} \\ \frac{\partial w}{\partial x} & \frac{\partial w}{\partial y} \end{vmatrix} = \begin{vmatrix} \frac{1}{y} & -\frac{x}{y^2} \\ 0 & 1 \end{vmatrix} = \frac{1}{y}.$$

由 $f_{X,Y}(x,y)|\mathrm{d}x\mathrm{d}y| = f_{Z,W}(z,w)|\mathrm{d}z\mathrm{d}w|$ 知

$$f_{Z,W}(z,w) = f_{X,Y}(x,y)\left|\frac{\partial(z,w)}{\partial(x,y)}\right|^{-1} = \frac{1}{2\pi}e^{-\frac{x^2+y^2}{2}} \cdot |y| = \frac{|w|}{2\pi}e^{-\frac{z^2w^2+w^2}{2}}.$$

例子

- 例 2：已知 X, Y 相互独立且都服从 $N(0,1)$ 分布．求 $Z = X/Y$ 的分布．
  定义映射 $g : (x, y) \mapsto (z, w)$，$z = x/y$，$w = y$. Jacobi 行列式

$$\left|\frac{\partial(z,w)}{\partial(x,y)}\right| = \begin{vmatrix} \frac{\partial z}{\partial x} & \frac{\partial z}{\partial y} \\ \frac{\partial w}{\partial x} & \frac{\partial w}{\partial y} \end{vmatrix} = \begin{vmatrix} \frac{1}{y} & -\frac{x}{y^2} \\ 0 & 1 \end{vmatrix} = \frac{1}{y}.$$

由 $f_{X,Y}(x,y)|\mathrm{d}x\mathrm{d}y| = f_{Z,W}(z,w)|\mathrm{d}z\mathrm{d}w|$ 知

$$f_{Z,W}(z,w) = f_{X,Y}(x,y)\left|\frac{\partial(z,w)}{\partial(x,y)}\right|^{-1} = \frac{1}{2\pi}e^{-\frac{x^2+y^2}{2}} \cdot |y| = \frac{|w|}{2\pi}e^{-\frac{z^2w^2+w^2}{2}}.$$

$$\Rightarrow f_Z(z) = \int_{\mathbb{R}} f_{Z,W}(z,w)\mathrm{d}w = \frac{1}{\pi}\int_0^{+\infty} we^{-\frac{(z^2+1)w^2}{2}}\mathrm{d}w = \frac{1}{\pi(z^2+1)}.$$

Z 的分布称为柯西分布．

二元正态分布

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 $N(\mu_1, \mu_2, \sigma_1^2, \sigma_2^2, \rho)$

二元正态分布

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 $N(\mu_1, \mu_2, \sigma_1^2, \sigma_2^2, \rho)$

- 若 $Z_1, Z_2 \sim N(0,1)$，则 $(X_1 = aZ_1 + bZ_2 + \mu_1, X_2 = cZ_1 + dZ_2 + \mu_2) \sim N(\mu_1, \mu_2, \sigma_1^2, \sigma_2^2, \rho)$，其中 $\sigma_1 = \sqrt{a^2 + b^2}$，$\sigma_2 = \sqrt{c^2 + d^2}$，$\rho = \frac{ac+bd}{\sigma_1\sigma_2}$，$ad - bc \neq 0$.

二元正态分布

二元正态分布

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 $N(\mu_1, \mu_2, \sigma_1^2, \sigma_2^2, \rho)$

- 若 $Z_1, Z_2 \sim N(0, 1)$，则 $(X_1 = aZ_1 + bZ_2 + \mu_1, X_2 = cZ_1 + dZ_2 + \mu_2) \sim N(\mu_1, \mu_2, \sigma_1^2, \sigma_2^2, \rho)$，其中 $\sigma_1 = \sqrt{a^2 + b^2}$，$\sigma_2 = \sqrt{c^2 + d^2}$，$\rho = \frac{ac+bd}{\sigma_1\sigma_2}$，$ad - bc \neq 0$.
- 若 X, Y 服从二元正态分布，则 $X' = aX + bY + e, Y' = cX + dY + f$（其中 $ad - bc \neq 0$），也服从二元正态分布

小结

给定 $\boldsymbol{X}$ 的分布和映射 $\boldsymbol{g}:\mathbb{R}^n\to\mathbb{R}^m$，计算随机变量函数 $\boldsymbol{Y}=\boldsymbol{g}(\boldsymbol{X})$ 的分布的方法：

- 离散型：找出 Y 的所有可能取值，逐个计算概率.

小结

给定 $\boldsymbol{X}$ 的分布和映射 $\boldsymbol{g}:\mathbb{R}^n\to\mathbb{R}^m$，计算随机变量函数 $\boldsymbol{Y}=\boldsymbol{g}(\boldsymbol{X})$ 的分布的方法：

- 离散型：找出 Y 的所有可能取值，逐个计算概率.
- 连续型：
  - $m=1$：考虑卷积公式，或先求 Y 的分布函数再求导数.
  - $m=n$：使用密度变换（例 1）.
  - $m<n$：找到合适的 $n-m$ 个随机变量，将 $\boldsymbol{g}$ 补足为 n 维向量值函数再使用密度变换（例 2）.

小结

给定 $\boldsymbol{X}$ 的分布和映射 $\boldsymbol{g}:\mathbb{R}^n\to\mathbb{R}^m$，计算随机变量函数 $\boldsymbol{Y}=\boldsymbol{g}(\boldsymbol{X})$ 的分布的方法：

- 离散型：找出 Y 的所有可能取值，逐个计算概率.
- 连续型：
  - $m=1$：考虑卷积公式，或先求 Y 的分布函数再求导数.
  - $m=n$：使用密度变换（例 1）.
  - $m<n$：找到合适的 $n-m$ 个随机变量，将 $\boldsymbol{g}$ 补足为 n 维向量值函数再使用密度变换（例 2）.

练习：利用密度变换公式分别求和、商、积的概率密度公式

利用均匀分布生成目标分布

问题

问题：如何利用均匀分布随机变量 $X \sim U(0,1)$ 生成连续随机变量 Y（概率密度 $f_Y(y)$）

给定 $Y = g(X)$, $X = h(Y)$，则有 $f_Y(y) = f_X(h(y))|h'(y)|$，因此 $h(y) = F_Y(y)$

利用均匀分布生成目标分布

PEKING UNIVERSITY

问题

问题：如何利用均匀分布随机变量 $X \sim U(0,1)$ 生成连续随机变量 Y（概率密度 $f_Y(y)$）

给定 $Y = g(X)$, $X = h(Y)$，则有 $f_Y(y) = f_X(h(y))|h'(y)|$, 因此 $h(y) = F_Y(y)$

利用均匀分布生成目标分布

假设 Y 的分布函数为 $F_Y(y)$, 已知均匀分布 $X \sim U(0,1)$，则有 $Y = F_Y^{-1}(X)$

- 实际过程中，我们可以利用均匀的随机数发生器来产生任意分布的随机数

例子

例：产生单位圆内均匀分布的点

例子

例：产生单位圆内均匀分布的点

- 策略 1：产生独立均匀分布 (X, Y)，并留下满足 $X^2 + Y^2 \leq 1$ 的点

例子

例：产生单位圆内均匀分布的点

- 策略 1：产生独立均匀分布 (X, Y)，并留下满足 $X^2 + Y^2 \leq 1$ 的点
- 缺点：当考虑高维球时，满足条件的概率指数变小

例子

例：产生单位圆内均匀分布的点

- 策略 1：产生独立均匀分布 (X, Y)，并留下满足 $X^2 + Y^2 \leq 1$ 的点
- 缺点：当考虑高维球时，满足条件的概率指数变小
- 策略 2：产生独立均匀分布 (X, Y)，再归一化为 $(X/R, Y/R)$，其中 $R = \sqrt{X^2 + Y^2}$?

例：产生单位圆内均匀分布的点

- 策略 1：产生独立均匀分布 (X, Y)，并留下满足 $X^2 + Y^2 \leq 1$ 的点
- 缺点：当考虑高维球时，满足条件的概率指数变小
- 策略 2：产生独立均匀分布 (X, Y)，再归一化为 $(X/R, Y/R)$，其中 $R = \sqrt{X^2 + Y^2}$?
- 策略 3：考虑极坐标 (R, W)，
  由 $f_{X,Y}(x, y)\mathrm{d}x\mathrm{d}y = f_{R,W}(r, w)\mathrm{d}r\mathrm{d}w$ 知 (R, W) 的概率密度为

$$f_{R,W}(r, w) = f_{X,Y}(x, y)\left|\frac{\partial(x, y)}{\partial(r, w)}\right| = \frac{r}{\pi}.$$

因此，$F_R(r) = r^2$，也即是对于 $X \sim U(0, 1)$，可以通过 $\sqrt{X}$ 得到 R 的分布

例子

例：产生单位球面内均匀分布的点

例子

例：产生单位球面内均匀分布的点

- 策略 1：产生独立均匀分布 (X, Y, Z)，并留下满足 $X^2 + Y^2 + Z^2 = 1$ 的点？

例子

例：产生单位球面内均匀分布的点

- 策略 1：产生独立均匀分布 (X, Y, Z)，并留下满足 $X^2+Y^2+Z^2=1$ 的点？
- 策略 2：产生独立均匀分布 (X, Y, Z)，再归一化为 $(X/R, Y/R, Z/R)$，其中 $R=\sqrt{X^2+Y^2+Z^2}$？

例子

PEKING UNIVERSITY

例：产生单位球面内均匀分布的点

- 策略 1：产生独立均匀分布 (X, Y, Z)，并留下满足 $X^2+Y^2+Z^2=1$ 的点?
- 策略 2：产生独立均匀分布 (X, Y, Z)，再归一化为 $(X/R, Y/R, Z/R)$，其中 $R=\sqrt{X^2+Y^2+Z^2}$?
- 策略 3：产生独立均匀分布 (X, Y, Z)，并留下满足 $X^2+Y^2+Z^2\leq 1$ 的点，再归一化为 $(X/R, Y/R, Z/R)$，其中 $R=\sqrt{X^2+Y^2+Z^2}$?

例子

例：产生单位球面内均匀分布的点

- 策略 1：产生独立均匀分布 (X, Y, Z)，并留下满足 $X^2 + Y^2 + Z^2 = 1$ 的点？
- 策略 2：产生独立均匀分布 (X, Y, Z)，再归一化为 $(X/R, Y/R, Z/R)$，其中 $R = \sqrt{X^2 + Y^2 + Z^2}$？
- 策略 3：产生独立均匀分布 (X, Y, Z)，并留下满足 $X^2 + Y^2 + Z^2 \leq 1$ 的点，再归一化为 $(X/R, Y/R, Z/R)$，其中 $R = \sqrt{X^2 + Y^2 + Z^2}$？
- 策略 4：产生单位球内均匀分布的点，再归一化？

例子

例：产生单位球面内均匀分布的点

- 策略 1：产生独立均匀分布 (X, Y, Z)，并留下满足 $X^2 + Y^2 + Z^2 = 1$ 的点？
- 策略 2：产生独立均匀分布 (X, Y, Z)，再归一化为 $(X/R, Y/R, Z/R)$，其中 $R = \sqrt{X^2 + Y^2 + Z^2}$？
- 策略 3：产生独立均匀分布 (X, Y, Z)，并留下满足 $X^2 + Y^2 + Z^2 \leq 1$ 的点，再归一化为 $(X/R, Y/R, Z/R)$，其中 $R = \sqrt{X^2 + Y^2 + Z^2}$？
- 策略 4：产生单位球内均匀分布的点，再归一化？
- 策略 5：考虑球坐标 (R, θ, ϕ)，我们有

$$\frac{1}{4\pi} dA = f(\theta, \phi) d\theta d\phi$$

因此 $f(\theta, \phi) = \sin(\phi)/4\pi$，也即是 $f(\theta) = 1/2\pi$，$f(\phi) = \sin(\phi)/2$。

例子

例：产生 d 维球内内或球面上均匀分布的点

例子

例：产生 d 维球内内或球面上均匀分布的点

- 球面：考虑独立同分布随机变量 X_i，概率密度为 $f(x_i) \sim N(0,1)$，则

$$f(\vec{x}) = f(x_1)f(x_2)\cdots f(x_d) = \frac{1}{(2\pi)^{d/2}} e^{-\frac{1}{2}|\vec{x}|^2}$$

因此 $f(\vec{x})$ 只与 $|\vec{x}|$ 相关与角度无关，因此归一化的 $Y_i = X_i/|\vec{X}|$ 为 d 维球面上均匀分布的点。

例子

例：产生 d 维球内内或球面上均匀分布的点

- 球面：考虑独立同分布随机变量 X_i，概率密度为 $f(x_i) \sim N(0,1)$，则

$$f(\vec{x}) = f(x_1)f(x_2)\cdots f(x_d) = \frac{1}{(2\pi)^{d/2}} e^{-\frac{1}{2}|\vec{x}|^2}$$

因此 $f(\vec{x})$ 只与 $|\vec{x}|$ 相关与角度无关，因此归一化的 $Y_i = X_i/|\vec{X}|$ 为 d 维球面上均匀分布的点。

- 球内：记求内均匀的点为 Z_i，则 $f_{\vec{Z}}(\vec{z}) \propto dV_d$。注意到 $dV_d = R^{d-1}dRdS_d$。因此，我们可以用 $U(0,1)^{1/d}$ 生成 R，同时有 $Z_i = Y_iR = U(0,1)^{1/d} \times X_i/|\vec{X}|$.