《初等概率论》第 6 讲

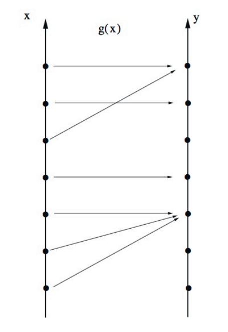
邓 婉 璐

清华大学
统计学研究中心

October 12, 2018

♣ 离散型随机变量的函数

对 $Y = g(X)$ 的每一个可能值，计算分布。（Y 也是离散的，故需算其分布列）



一、随机变量函数的分布

例 1.1

设 X 有如下的概率分布

X	-2	-1	0	1	3
$\mathbb{P}$	0.1	0.2	0.3	0.2	0.2

求 $Y = X^2$ 的分布.

解. Y 可能的取值为 $0, 1, 4, 9$，而且

$$\mathbb{P}(Y=0) = \mathbb{P}(X=0) = 0.3;$$
$$\mathbb{P}(Y=1) = \mathbb{P}(|X|=1) = \mathbb{P}(X=-1) + \mathbb{P}(X=1) = 0.4;$$
$$\mathbb{P}(Y=4) = \mathbb{P}(X=-2) = 0.1;$$
$$\mathbb{P}(Y=9) = \mathbb{P}(X=3) = 0.2.$$

于是 Y 的分布列为：

Y	0	1	4	9
$\mathbb{P}$	0.3	0.4	0.1	0.2

一、随机变量函数的分布

♣ 连续型随机变量的函数

考察 $Y=g(X)$ 的分布。若 $g(x)$ 是连续函数，则可通过下面两步实现：

1. 得到 Y 的 CDF: $F_Y(y)=P(Y\leq y)$
2. 求导得到：$f_Y(y)=\frac{dF_Y(y)}{dy}$.

一、随机变量函数的分布

例 1.2

设 $X \sim \mathcal{N}(\mu, \sigma^2)$, 则 $Y = (X-\mu)/\sigma \sim \mathcal{N}(0,1)$.

解. 显然

$$\begin{aligned}\mathbb{P}(Y \leq y) &= \mathbb{P}(X \leq \mu + \sigma y) \\ &= \int_{-\infty}^{\mu+\sigma y} \frac{1}{\sigma\sqrt{2\pi}} \exp\left\{-\frac{(x-\mu)^2}{2\sigma^2}\right\} dx \\ &= \int_{-\infty}^{y} \frac{1}{\sqrt{2\pi}} \exp\left\{-\frac{t^2}{2}\right\} dt \\ &= \int_{-\infty}^{y} \varphi(t)\, dt.\end{aligned}$$

因此 $Y \sim \mathcal{N}(0,1)$.

例 1.3

设 $X \sim \mathcal{N}(0,1)$，求 $Y = X^2$ 的概率密度.

解. 由于 $Y \geq 0$，对于 $y > 0$，分布函数

$$\begin{aligned} F_Y(y) = \mathbb{P}(Y \leq y) &= \mathbb{P}(-\sqrt{y} \leq X \leq \sqrt{y}) \\ &= \Phi(\sqrt{y}) - \Phi(-\sqrt{y}) \\ &= 2\Phi(\sqrt{y}) - 1. \end{aligned}$$

是连续函数，所以 Y 的密度为

$$f_Y(y) = F'_Y(y) = 2\varphi(\sqrt{y})\frac{1}{2\sqrt{y}} = \frac{1}{\sqrt{2\pi y}}\mathrm{e}^{-y/2}, \quad y > 0.$$

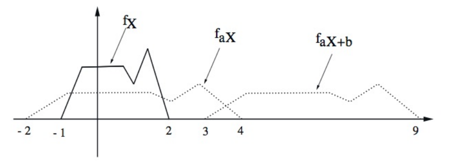
即 $Y \sim \Gamma(1/2, 1/2) := \chi^2_1$.

一、随机变量函数的分布

一些特殊情形可以从 $f_X(x)$ 直接得到 $f_Y(y)$，无需再计算 CDF。

♣ 形如 $Y = aX + b$.

例： $Y = 2X + 5$

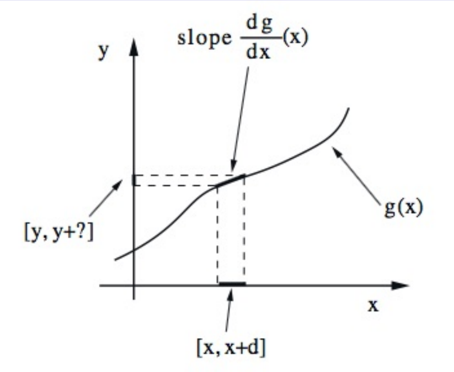


$$f_Y(y) = \frac{1}{|a|} f_X\left(\frac{y-b}{a}\right).$$

可以利用该结论验证：若 X 是服从正态分布的随机变量，则 Y 也是正态的。

一、随机变量函数的分布

♣ 扩展：$Y = g(X)$, $g(x)$ 是严格单调函数.



事件 $\{x < X \leq x+\delta\} = \{g(x) < Y \leq g(x+\delta)\}$ 可近似为 $\{g(x) < Y \leq g(x) + \delta|\frac{dg}{dx}(x)|\}$. 因此，$\delta f_X(x) = \delta f_Y(g(x))|g'(x)|$. 即 $f_Y(y) = f_X(h(y))|h'(y)|$, $h(y)$ 是 $g(x)$ 的逆映射.

一、随机变量函数的分布

更一般地，有如下定理：

定理 1.1

设 X 的概率密度 $f(x)$, $D \subset R$, $Y = g(X)$, $\mathbb{P}(Y \in D) = 1$, 如果

1. 对 $y \in D$, $\{Y = y\} = \cup_{i=1}^{n} \{X = h_i(y)\}$;
2. 每个 $h_i(y)$ 是 D 到其值域 D_i 的可逆映射，在 D 内有连续的导数；
3. $D_1, D_2, ..., D_n$ 互不相交，

则 Y 的密度函数为

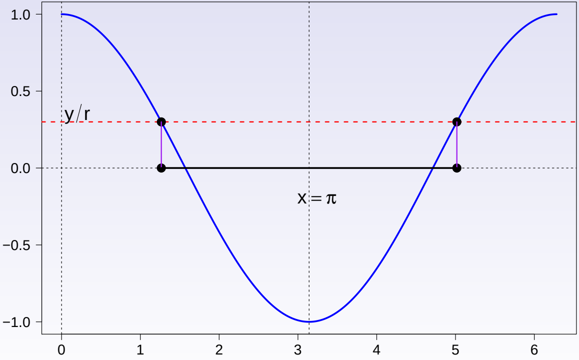
$$f_Y(y) = \begin{cases} \sum_{i=1}^{n} f(h_i(y))|h_i'(y)|, & y \in D, \\ 0, & y \in D^c. \end{cases}$$

一、随机变量函数的分布

例 1.4

设 r 是正常数，$X \sim \mathcal{U}(0, 2\pi)$，求 $Y = r\cos X$ 的概率密度.

解.



一、随机变量函数的分布

方法一：

设 $D=(-r,r)$，则 $\mathbb{P}(Y\in D)=1$. 对于 $y\in D$，有

$$
\begin{aligned}
\{Y=y\} &= \{\cos X = y/r\} \\
&= \{\cos X = y/r, X\in(0,\pi)\}\cup\{\cos X = y/r, X\in(\pi,2\pi)\} \\
&= \{\cos X = y/r, X\in(0,\pi)\}\cup\{\cos(2\pi - X) = y/r, X\in(\pi,2\pi)\} \\
&= \{X=\arccos(y/r)\}\cup\{X=2\pi-\arccos(y/r)\},
\end{aligned}
$$

$h_1(x)=\arccos(x/r): D\to(0,\pi)$, $h_2(x)=2\pi-\arccos(x/r): D\to(\pi,2\pi)$ 都是可逆的、连续可微的函数. 利用 $F_X(x)=1/(2\pi)$, $x\in(0,2\pi)$, 可得 Y 的密度函数

$$
\begin{aligned}
f_Y(y) &= \frac{1}{2\pi}\left|\frac{d}{dy}\arccos(y/r)\right| + \frac{1}{2\pi}\left|\frac{d}{dy}[2\pi-\arccos(y/r)]\right| \\
&= \frac{1}{\pi\sqrt{r^2-y^2}}, \quad y\in(-r,r).
\end{aligned}
$$

一、随机变量函数的分布

方法二：

对 $y \in (-r, r)$,

$$\begin{aligned}\mathbb{P}(Y \leq y) &= \mathbb{P}(\cos X \leq y/r) \\ &= \int_{\{x:\cos x \leq y/r\}} \frac{1}{2\pi} dx \\ &= \frac{1}{2\pi}\{2\pi - 2\arccos(y/r)\}.\end{aligned}$$

根据

$$[\arccos(x)]' = -\frac{1}{\sqrt{1-x^2}},$$

可得

$$f_Y(y) = \frac{1}{\pi\sqrt{r^2 - y^2}}, \quad y \in (-r, r).$$

二、随机向量及其分布

♣ 1. 随机向量及其联合分布

定义 2.1 (n 维随机向量)

如果 $X_1, ..., X_n$ 都是概率空间 $(\Omega, \mathscr{F}, \mathbb{P})$ 上的随机变量，称 $\mathbf{X} = (X_1, ..., X_n)$ 是概率空间 $(\Omega, \mathscr{F}, \mathbb{P})$ 上的 n **维随机向量**，简称**随机向量**.

定义 2.2 (联合概率分布函数、分布函数、联合分布)

设 $\mathbf{X} = (X_1, ..., X_n)$ 是随机向量，称 R^n 上的 n 元函数

$$F(x_1, x_2, ..., x_n) = \mathbb{P}(X_1 \le x_1, X_2 \le x_2, ..., X_n \le x_n)$$

为 $\mathbf{X} = (X_1, ..., X_n)$ 的**联合概率分布函数**，简称为**分布函数**或者**联合分布**.

♣ 联合分布函数 $F(x_1, x_2, ..., x_n)$ 是 R^n 上的右连续函数，关于每个自变元 x_j 单调非降.

♣ $F(\mathbf{x}) = \mathbb{P}(\mathbf{X} \le \mathbf{x})$，其中 $\mathbf{X} \le \mathbf{x} \iff X_j \le x_j,\ j = 1, ..., n.$

二、随机向量及其分布

定义 2.3 (边缘分布、边际分布 (marginal distribution))

对 $1 \leq k < n$, $S_k = \{i_1, i_2, ..., i_k\}$ 为指标集，称 $\mathbf{X}_{S_k} := (X_{i_1}, X_{i_2}, ..., X_{i_k})$ 的联合分布

$$
\begin{aligned}
&F_k(x_{i_1}, x_{i_2}, ..., x_{i_k}) \\
&= \mathbb{P}(X_{i_1} \leq x_{i_1}, X_{i_2} \leq x_{i_2}, ..., X_{i_n} \leq x_{i_n}) \\
&= \mathbb{P}(X_{i_j} \leq x_{i_j}, j \in S_k, X_{i_t} \leq \infty, t \in S_k^c)
\end{aligned}
$$

为 $F(x_1, x_2, ..., x_n)$ 的k-维边缘分布.

♣ $F(x_1, x_2, ..., x_n)$ 一共有 $(2^n - 2)$ 个边缘分布.

♣ 边缘分布由联合分布 $F(\mathrm{x})$ 唯一决定，反之不成立.（例如后面的二元正态分布的例子）

二、随机向量及其分布

例 2.1

设 $F(x, y)$ 是 (X, Y) 的联合分布，则 X, Y 分别有概率分布

$$F_X(x) = \mathbb{P}(X \le x, Y \le \infty) = F(x, \infty);$$
$$F_Y(y) = \mathbb{P}(X \le \infty, Y \le y) = F(\infty, y).$$

进一步，对矩形 $D = \{a < x \le b, c < y \le d\}$，有

$$\begin{aligned}\mathbb{P}((X, Y) \in D) &= \mathbb{P}(a < X \le b, c < Y \le d) \\ &= F(b, d) - F(b, c) - F(a, d) + F(a, c).\end{aligned}$$

因此

$$F(b, d) - F(b, c) - F(a, d) + F(a, c) \ge 0, \quad a < b, c < d.$$

二、随机向量及其分布

例 2.2

如果在 (x,y) 的领域内，

$$\frac{\partial F(x,y)}{\partial x}, \frac{\partial F(x,y)}{\partial y}, \frac{\partial^2 F(x,y)}{\partial x \partial y}, \frac{\partial^2 F(x,y)}{\partial y \partial x}$$

连续，由上式可得

$$\begin{aligned}
&\lim_{\substack{\triangle x \downarrow 0 \\ \triangle y \downarrow 0}} \frac{\mathbb{P}(x-\triangle x < X \le x, y-\triangle y < Y \le y)}{\triangle x \triangle y} \\
&= \frac{\partial^2 F(x,y)}{\partial x \partial y} = \frac{\partial^2 F(x,y)}{\partial y \partial x} \\
&:= f(x,y).
\end{aligned}$$

后面将介绍，$f(x,y)$ 称为 (X,Y) 的联合密度函数.

二、随机向量及其分布

♣ 独立性定义回顾

设 $(X_1, X_2, ..., X_n)$ 有联合分布 $F(x_1, x_2, ..., x_n)$，X_i 有边缘分布 $F_i(x_i)$. 则 $X_1, ..., X_n$ 相互独立当且仅当

$$F(x_1, x_2, ..., x_n) = F_1(x_1)F_2(x_2)\cdots F_n(x_n).$$

二、随机向量及其分布

♣ 2. 离散型随机向量

A. 离散型随机向量及其联合分布

定义 2.4 (离散型随机向量)

如果 $X_1, X_2, ..., X_n$ 都是离散型随机变量，则称 $\mathbf{X} = (X_1, X_2, ..., X_n)$ 为离散型随机向量. 如果 $\mathbf{X}$ 所有的不同取值是

$$\mathbf{x}(j_1, j_2, ..., j_n) = (x_1(j_1), x_2(j_2), ..., x_n(j_n)), \quad j_1, j_2, ..., j_n \geq 1,$$

则称

$$p_{j_1, j_2, ..., j_n} = \mathbb{P}(\mathbf{X} = \mathbf{x}(j_1, j_2, ..., j_n)), \quad j_1, j_2, ..., j_n \geq 1,$$

是 $\mathbf{X}$ 的联合分布列.

二、随机向量及其分布

♣ 概率分布的性质

1. $p_{j_1,j_2,\ldots,j_n} \geq 0$;
2. $\displaystyle\sum_{j_1,j_2,\ldots,j_n} p_{j_1,j_2,\ldots,j_n} = 1$;
3. 对任何整数 $k \in (0,n)$, $(X_1, X_2, \ldots, X_k)$ 有联合分布列

$$\mathbb{P}(X_1 = x_1(j_1), X_2 = x_2(j_2), \ldots, X_k = x_k(j_k)) = \sum_{j_{k+1},j_{k+2},\ldots,j_n} p_{j_1,j_2,\ldots,j_n}, \quad j_1, j_2, \ldots, j_k \geq 1.$$

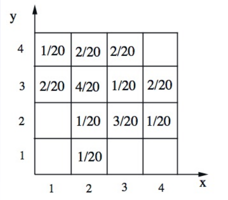
此称为 $\mathbf{X}$ 的边缘分布列.

二、随机向量及其分布

当 (X, Y) 的联合分布列的规律性不强时，还可以用表格的形式表达：

p_{ij}	y_1	y_2	y_3	$\cdots$	y_n	$\cdots$	$\{p_i\}$
x_1	p_{11}	p_{12}	p_{13}	$\cdots$	p_{1n}	$\cdots$	p_1
x_2	p_{21}	p_{22}	p_{23}	$\cdots$	p_{2n}	$\cdots$	p_2
x_3	p_{31}	p_{32}	p_{33}	$\cdots$	p_{3n}	$\cdots$	p_3
$\vdots$	$\vdots$	$\vdots$	$\vdots$		$\vdots$	$\cdots$	$\vdots$
$\{q_j\}$	q_1	q_2	q_3	$\cdots$	q_n	$\cdots$	$\Sigma = 1$

其中 p_i 是其所在行中 p_{ij} 之和，q_j 是其所在列的和.



二、随机向量及其分布

例 2.3

设 (X, Y) 有联合分布列

$$p_{ij} = \mathbb{P}(X = x_i, Y = y_j), \quad i, j \geq 1.$$

则 X 和 Y 分别有概率分布列

$$p_i \overset{\text{def}}{=} \mathbb{P}(X = x_i) = \sum_{j=1}^{\infty} \mathbb{P}(X = x_i, Y = y_j) = \sum_{j=1}^{\infty} p_{ij}, \quad i \geq 1$$

$$q_j \overset{\text{def}}{=} \mathbb{P}(Y = y_j) = \sum_{i=1}^{\infty} \mathbb{P}(X = x_i, Y = y_j) = \sum_{i=1}^{\infty} p_{ij}, \quad j \geq 1.$$

二、随机向量及其分布

例 2.4 (多项分布)

设 $A_1, A_2, ..., A_r$ 是试验 S 的完备事件组. 对试验 S 进行 n 次独立重复试验时, 用 X_i 表示 A_i 发生的次数, (用归纳法可得) 则 $(X_1, ..., X_r)$ 的概率分布是

$$\mathbb{P}(X_1 = k_1, ..., X_r = k_r) = \frac{n!}{k_1! \cdots k_r!} p_1^{k_1} \cdots p_r^{k_r},$$

其中 $k_i \geq 0$, $p_i = \mathbb{P}(A_i)$, $\sum_{i=1}^r k_i = n$.

对 $1 \leq m < n$, $(X_{j_1}, X_{j_2}, ..., X_{j_m})$ 的边缘分布为

$$\mathbb{P}(X_{j_1} = k_1, ..., X_{j_m} = k_m) = \frac{n!}{k_1! \cdots k_m! k!} p_{j_1}^{k_1} \cdots p_{j_m}^{k_m} q^k,$$

其中 $0 \leq k_i \leq n$, $k = n - \sum_{i=1}^m k_i \geq 0$, $q = 1 - \sum_{i=1}^m p_{j_i}$. 特别地,

$$\mathbb{P}(X_j = k) = \binom{n}{k} p_j^k (1 - p_j)^{n-k}, \quad k \geq 0.$$

二、随机向量及其分布

B. 离散型随机向量的独立性

定理 2.1

设离散型随机向量 $\mathbf{X} = (X_1, X_2, ..., X_n)$ 有概率分布

$$p_{j_1,j_2,...,j_n} = \mathbb{P}(\mathbf{X} = \mathbf{x}(j_1, j_2, ..., j_n)), \quad j_1, j_2, ..., j_n \geq 1.$$

则 $X_1, ..., X_n$ 相互独立的充分必要条件是：对任意的 $(x_1(j_1), x_2(j_2), ..., x_n(j_n))$，有

$$\mathbb{P}(X_1 = x_1(j_1), ..., X_n = x_n(j_n)) = \prod_{i=1}^{n} \mathbb{P}(X_i = x_i(j_i)).$$

推论 2.1

设离散型随机向量 (X, Y) 的所有不同取值是

$$(x_i, y_j), \quad i, j \geq 1,$$

则 X, Y 相互独立的充分必要条件是：对任意的 (x_i, y_j)，有

$$\mathbb{P}(X = x_i, Y = y_j) = \mathbb{P}(X = x_i)\mathbb{P}(Y = y_j).$$

证明. $(\Longleftarrow)$ 对任意的 $x, y \in R$，利用概率的可列可加性得

$$\mathbb{P}(X \leq x, Y \leq y) = \mathbb{P}\Big(\bigcup_{i:x_i \leq x} \{X = x_i\}, \bigcup_{j:y_j \leq x} \{Y = y_j\}\Big)$$

$$= \sum_{i:x_i \leq x} \sum_{j:y_j \leq x} \mathbb{P}(X = x_i, Y = y_j)$$

二、随机向量及其分布

$$
\begin{aligned}
&= \sum_{i: x_i \le x} \sum_{j: y_j \le x} \mathbb{P}(X = x_i)\mathbb{P}(Y = y_j) \\
&= \sum_{i: x_i \le x} \mathbb{P}(X = x_i) \sum_{j: y_j \le x} \mathbb{P}(Y = y_j) \\
&= \mathbb{P}(X \le x)\mathbb{P}(Y \le y).
\end{aligned}
$$

$(\Longrightarrow)$ 由于 $A = \{x_i\}$ 和 $B = \{y_j\}$ 都是 Borel 集，故 $\{X = x_i\} = \{X \in A\}$ 与 $\{Y = y_j\} = \{Y \in B\}$ 独立，因此

$$\mathbb{P}(X = x_i, Y = y_j) = \mathbb{P}(X = x_i)\mathbb{P}(Y = y_j).$$

♣ 3. 连续型随机向量及其联合密度

A. 联合概率密度

定义 2.5 (连续型随机向量)

设 $\mathbf{X} = (X_1, X_2, ..., X_n)$ 是随机向量. 如果存在 R^n 上的非负可积函数 $f(x_1, ..., x_n)$ (or $f(\mathbf{x})$), 使得对 R^n 的任何子立方体

$$D = \{(x_1, x_2, ..., x_n) | a_i < x_i \le b_i, 1 \le i \le n\},$$

有

$$\begin{aligned}\mathbb{P}(\mathbf{X} \in D) &= \int_D f(x_1, x_2, ..., x_n)\, dx_1\, dx_2 \cdots dx_n \\ &:= \int_D f(\mathbf{x})\, d\mathbf{x},\end{aligned}$$

就称 $\mathbf{X}$ 是连续型随机向量, 并称 $f(x_1, x_2, ..., x_n)$ (or $f(\mathbf{x})$) 是 $\mathbf{X}$ 的联合概率密度函数, 简称为联合密度或概率密度.

二、随机向量及其分布

定理 2.2

对任意 n 维 Borel 集 B 有:

$$\mathbb{P}(\mathbf{X} \in B) = \int_B f(\mathbf{x})\, d\mathbf{x}.$$

♣ 随机向量 $\mathbf{X}$ 的联合概率密度不必唯一.

♣ 由于 R^n 是 Borel 集，有

$$\int_{R^n} f(\mathbf{x})\, d\mathbf{x} = \mathbb{P}(\mathbf{X} \in R^n) = 1.$$

二、随机向量及其分布

定理 2.3 (Fubini)

设 D 是 R^n 上的子区域，$\varphi(\mathbf{x})$ 是 D 上的非负函数或绝对可积函数（指取绝对值后积分有限），则对区域 D 上的 n 重积分

$$\int_D \varphi(\mathbf{x})\, d\mathbf{x}$$

可以进行累次积分计算，且积分的次序可以交换.

B. 边缘密度

设 $f(\mathbf{x})$ 是 $\mathbf{X}$ 的概率密度. 对 $1 \le k < n$, 令 R^k 中的子立方体

$$D_k = \{(x_1, x_2, ..., x_k) | a_i < x_i \le b_i, 1 \le i \le k\},$$

定义

$$D_k \times R^{n-k}$$
$$= \{(x_1, x_2, ..., x_n) | (x_1, ..., x_k) \in D_k, (x_{k+1}, ..., x_n) \in R^{n-k}\},$$

则

$$\mathbb{P}((X_1, .., X_k) \in D_k) = \mathbb{P}(\mathbf{X} \in D_k \times R^{n-k}) = \int_{D_k \times R^{n-k}} f(\mathbf{x}) d\mathbf{x}$$

$$= \int_{D_k} \left\{ \int_{R^{n-k}} f(x_1, ..., x_n) dx_{k+1} \cdots dx_n \right\} dx_1 \cdots dx_k.$$

二、随机向量及其分布

于是有下列定义:

定义 2.6 (边缘密度、边际密度)

被积函数

$$f_k(x_1, ..., x_k) = \int_{R^{n-k}} f(x_1, ..., x_n)\, dx_{k+1} \cdots dx_n$$

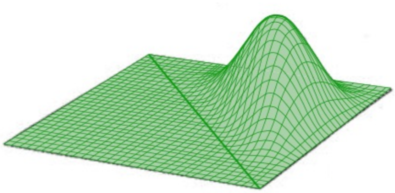
是 $(X_1, ..., X_k)$ 的联合密度，被称为 $\mathbf{X}$ 的**边缘概率密度函数**，简称为**边缘密度**.

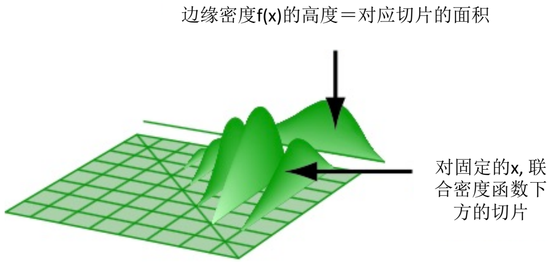
例 2.5

设 $f(\mathbf{x})$ 是 $\mathbf{X}$ 的概率密度. $(X_{j_1}, X_{j_2}, ..., X_{j_k})$ 的边缘密度为:

$$g_k(x_{j_1}, x_{j_2}, ..., x_{j_k}) = \int_{R^{n-k}} f(x_1, ..., x_n)\, dx_{j_{k+1}} \cdots dx_{j_n}.$$

二、随机向量及其分布





二、随机向量及其分布

例 2.6

设 (X, Y) 有联合密度 $f(x, y)$. 则 X 与 Y 分别有概率密度

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

例 2.7

设 (X, Y) 在单位圆 $D = \{(x, y) | x^2 + y^2 \leq 1\}$ 内均匀分布，求 X, Y 的概率密度.

解. (X, Y) 的联合密度为 $f(x, y) = \frac{1}{\pi} I_D$. X 只在 $[-1, 1]$ 中取值，则

$$f_X(x) = \int_{-\infty}^{\infty} f(x, y)\, dy = \frac{1}{\pi} \int_{-\infty}^{\infty} I_{x^2+y^2 \leq 1}\, dy$$

$$= \frac{1}{\pi} \int_{-\infty}^{\infty} I_{|y| \leq \sqrt{1-x^2}}\, dy = \frac{2}{\pi} \sqrt{1 - x^2}, \quad |x| \leq 1.$$

同理可得 Y 的密度.

二、随机向量及其分布

C. 联合分布与联合密度

设 $f(x, y)$ 是 (X, Y) 的联合密度，则有

$$F(x, y) = \int_{-\infty}^{x} \int_{-\infty}^{y} f(u, v)\, du dv.$$

当 $f(x, y)$ 连续时，有

$$f(x, y) = \frac{\partial^2 F(x, y)}{\partial x \partial y}.$$

二、随机向量及其分布

定理 2.4

设 (X, Y) 有连续的分布函数 $F(x, y)$，定义

$$f(x,y)=\begin{cases}\dfrac{\partial^2 F(x,y)}{\partial x\partial y}, & \text{当该混合偏导数存在,}\\[2ex] 0, & \text{其它.}\end{cases}$$

如果

$$\int_{R^2} f(x,y)\,dxdy=1,$$

则 $f(x, y)$ 是 (X, Y) 的联合密度.

例 2.8

如果 X, Y 的分布函数连续，则它们的联合分布 $F(x, y)$ 连续.

证明. 对 $\forall(x_0, y_0)$ 和 $\forall\varepsilon > 0$, $\exists\delta > 0$，使得 $|x - x_0| < \delta$, $|y - y_0| < \delta$ 时，有

$$|F_X(x) - F_X(x_0)| < \varepsilon/2, \quad |F_Y(y) - F_Y(y_0)| < \varepsilon/2.$$

于是对满足 $|x - x_0| < \delta$, $|y - y_0| < \delta$ 的 (x, y) 有

$$\begin{aligned}
&|F(x, y) - F(x_0, y_0)| \\
&\le |F(x, y) - F(x_0, y)| + |F(x_0, y) - F(x_0, y_0)| \\
&\le \mathbb{P}(x \wedge x_0 < X \le x \vee x_0, Y \le y) \\
&\quad + \mathbb{P}(X \le x_0, y \wedge y_0 < Y \le y \vee y_0) \\
&\le \mathbb{P}(x \wedge x_0 < X \le x \vee x_0) + \mathbb{P}(y \wedge y_0 < Y \le y \vee y_0) \\
&= |F_X(x) - F_X(x_0)| + |F_Y(y) - F_Y(y_0)| \\
&< \varepsilon.
\end{aligned}$$

例 2.9

存在随机向量 (X, Y)，它有连续的联合分布，但不是连续型随机向量.

证明. 设 $X \sim \mathcal{U}(0,1)$， $Y = X$， 则 (X, Y) 有连续的联合分布函数

$$F(x,y) = \mathbb{P}(X \le \min(x,y)) = \begin{cases} 0, & \min\{x,y\} \le 0, \\ \min\{x,y\}, & \min\{x,y\} \in (0,1], \\ 1, & \min\{x,y\} > 1. \end{cases}$$

如果 (X, Y) 有联合密度 $f(x,y)$， 用 D 表示直线 $y = x$ 上点的全体， 则

$$1 = \mathbb{P}(X = Y) = \mathbb{P}((X, Y) \in D) = \int_D f(x,y)\,dxdy = 0.$$

矛盾！因此 (X, Y) 没有联合密度，从而不是连续型随机向量.

二、随机向量及其分布

♣ 在例子2.9中，$F(x,y)$ 连续，除去在有限条直线外，偏导数 $\frac{\partial^2 F(x,y)}{\partial x \partial y}$ 存在且连续.

♣ 例子2.9说明，$F(x,y)$ 连续，除去在有限条直线外，偏导数 $\frac{\partial^2 F(x,y)}{\partial x \partial y}$ 存在且连续还是不能保证 (X,Y) 有联合密度.

二、随机向量及其分布

D. 连续型随机变量的独立性

定理 2.5

设对每个 $i(1 \leq i \leq n)$，随机变量 X_i 有概率密度 $f_i(x_i)$，则 $X_1, X_2, ..., X_n$ 相互独立的充分必要条件是随机向量 $\mathbf{X} = (X_1, X_2, ..., X_n)$ 有联合密度

$$f_1(x_1)f_2(x_2)\cdots f_n(x_n), \quad (x_1, x_2, ..., x_n) \in R^n.$$

二、随机向量及其分布

证明. $(\Longleftarrow)$ 对任意的 $(x_1, ..., x_n) \in R^n$, 有

$$\begin{aligned}\mathbb{P}(X_1 \le x_1, ..., X_n \le x_n) &= \int_{-\infty}^{x_1} \cdots \int_{-\infty}^{x_n} \prod_{i=1}^{n} f_i(t_i)\,dt_1 \cdots dt_n \\ &= \prod_{i=1}^{n} \int_{-\infty}^{x_i} f_i(t_i)\,dt_i \\ &= \prod_{i=1}^{n} \mathbb{P}(X_i \le x_i).\end{aligned}$$

故 $X_1, X_2, ..., X_n$ 相互独立.

$(\Longrightarrow)$ 如果 $X_1, ..., X_n$ 相互独立，对 R^n 的任何子立方体

$$D = \{(x_1, x_2, ..., x_n) | a_i < x_i \le b_i, 1 \le i \le n\},$$

利用 Fubini 定理，得

二、随机向量及其分布

$$
\begin{aligned}
\mathbb{P}(\mathbf{X} \in D) &= \mathbb{P}(a_1 < X_1 \le b_1, \ldots, a_n < X_n \le b_n) \\
&= \mathbb{P}(a_1 < X_1 \le b_1) \cdots \mathbb{P}(a_n < X_n \le b_n) \\
&= \int_{a_1}^{b_1} f_1(t_1)\, dt_1 \cdots \int_{a_n}^{b_n} f_n(t_n)\, dt_n \\
&= \int_D f_1(t_1) \cdots f_n(t_n)\, dt_1 \cdots dt_n.
\end{aligned}
$$

因此 $f_1(x_1)\cdots f_n(x_n)$ 是 $\mathbf{X}$ 的联合密度.

定义 2.7 (二元正态分布)

设 μ_1, μ_2 是常数，σ_1, σ_2 是正常数，$\rho \in (-1, 1)$ 中的常数. 如果随机向量 (X, Y) 有概率密度

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

则称 (X, Y) 服从二元正态分布，记作 $(X, Y) \sim \mathcal{N}(\mu_1, \mu_2, \sigma_1^2, \sigma_2^2, \rho)$.

二、随机向量及其分布

♣ 简写形式：

$$f(\mathbf{x}) = \frac{1}{\sqrt{(2\pi)^2|\Sigma|}} \exp\{-\frac{1}{2}(\mathbf{x}-\mu)'\Sigma^{-1}(\mathbf{x}-\mu)\},$$

其中 $\mathbf{x}=(x,y)'$, $\mu=(\mu_1,\mu_2)'$, $\Sigma=\begin{pmatrix}\sigma_1^2 & \rho\sigma_1\sigma_2\\ \rho\sigma_1\sigma_2 & \sigma_2^2\end{pmatrix}$,

$|\Sigma|$ 是 Σ 的行列式.

二、随机向量及其分布

定理 2.6

设 $(X, Y) \sim \mathcal{N}(\mu_1, \mu_2, \sigma_1^2, \sigma_2^2, \rho)$. 则 X, Y 独立的充分必要条件是 $\rho = 0$.

证明. 不失一般性，假设 $(X, Y) \sim \mathcal{N}(0, 0, 1, 1, \rho)$. 则

$$f_X(x) = \int_R f(x, y)\,dy = \frac{1}{\sqrt{2\pi}} \exp(-x^2/2),$$
$$f_Y(y) = \int_R f(x, y)\,dx = \frac{1}{\sqrt{2\pi}} \exp(-y^2/2).$$

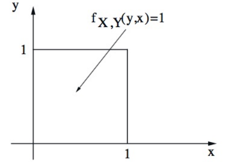
因此 X, Y 独立 $\Longleftrightarrow f_X(x) f_Y(y) = f(x, y)$.
当 $\rho = 0$ 时，有 $f_X(x) f_Y(y) = f(x, y)$, 故 X, Y 独立.
当 X, Y 独立时，取 $(x, y) = (0, 0)$，得

$$\sqrt{1 - \rho^2} = 1, \quad \Longrightarrow \rho = 0.$$

(i) 一般情形，直接求分布，再根据需要求 PDF 或 PMF.

例 3.1

设 (X, Y) 为单位正方形上的均匀分布，有联合密度 $f(x, y)$. 可验证 X, Y 独立. 求 $Z = g(X, Y) = Y/X$ 的概率密度函数。



$$F_Z(z) = P(Y/X \le z) = \frac{z}{2}, \quad 0 \le z \le 1.$$

$$F_Z(z) = P(Y/X \le z) = 1 - \frac{1}{2z}, \quad z > 1.$$

$$F_Z(z) = 0, \quad z < 0.$$

例 3.2

设 X, Y 独立，且服从 $\mathcal{N}(0,1)$. 求 $Z=\sqrt{X^2+Y^2}$ 的分布.

解. (X, Y) 有联合密度

$$f(x,y)=\frac{1}{2\pi}\exp\left(-\frac{x^2+y^2}{2}\right).$$

对 $z\geq 0$, 有

$$\begin{aligned}
F_Z(z) &= \mathbb{P}(\sqrt{X^2+Y^2}\leq z)\\
&= \int_{\{\sqrt{x^2+y^2}\leq z\}}\frac{1}{2\pi}\exp\left(-\frac{x^2+y^2}{2}\right)dxdy\\
&\xlongequal[y=r\sin\theta]{x=r\cos\theta}\frac{1}{2\pi}\int_0^{2\pi}d\theta\int_0^z re^{-r^2/2}\,dr\\
&= \int_0^z re^{-r^2/2}\,dr.
\end{aligned}$$

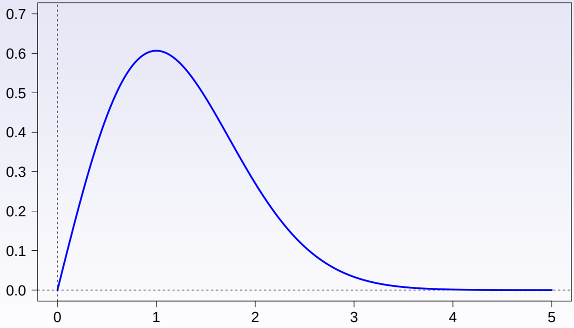
三、随机向量函数的分布

求导可得 Z 的密度函数

$$f_Z(z) = z\mathrm{e}^{-z^2/2}, \quad z \geq 0.$$

此分布称为 Rayleigh 分布.

♣ 向坐标 $(0,0)$ 射击时，弹落点 (X, Y) 服从二元正态分布.
$R = \sqrt{X^2 + Y^2}$ 为弹落点到目标的距离，被称为脱靶量.



三、随机向量函数的分布

(ii) 特殊情形 $Z = X + Y$

例 3.3

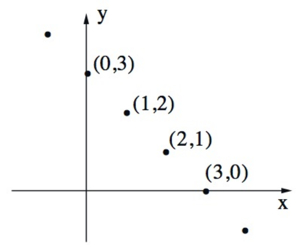
设 (X, Y) 有联合分布列，求 $Z = X + Y$ 的分布列.

$$P_Z(z) = P(X + Y = z) = \sum_x P(X = x, Y = z - x).$$

特别地，当 X, Y 独立时，$Z = X + Y$ 有分布列

$$P_Z(z) = \sum_x P(X = x)P(Y = z - x) = \sum_x P_X(x)P_Y(z - x).$$

此时 Z 的分布列为 X 和 Y 的分布列的卷积.



三、随机向量函数的分布

例 3.4

设 (X, Y) 有联合密度 $f(x, y)$，则 $Z = X + Y$ 有密度函数

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

特别地，当 X, Y 独立时，$Z = X + Y$ 有密度函数

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

证明. 对任意的 $a < b$，有

$$\mathbb{P}(a < Z \leq b) = \mathbb{P}(a < X + Y \leq b) = \int_{R^2} I_{\{a<x+y\leq b\}} f(x, y)\, dxdy$$

$$= \int_R \left(\int_{a-x}^{b-x} f(x, y)\, dy \right) dx$$

三、随机向量函数的分布

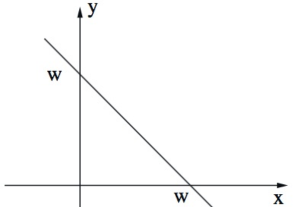
$$\overset{y=z-x}{=\!=\!=\!=} \int_R \left(\int_a^b f(x, z-x)\, dz \right) dx$$

$$= \int_a^b \left(\int_R f(x, z-x)\, dx \right) dz.$$

故，$Z = X + Y$ 的密度函数为

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

易见，Z 的密度函数也可以写为：$f_Z(z) = \int_{-\infty}^{\infty} f(z-y, y)\, dy$.



例 3.5

设 X, Y 独立，且服从 $\mathcal{U}(0,1)$. 求 $Z = X + Y$ 的概率密度.

证明. 显然 $Z \in (0,2)$，且

$$f_X(x) = f_Y(x) = I_{\{0<x<1\}}.$$

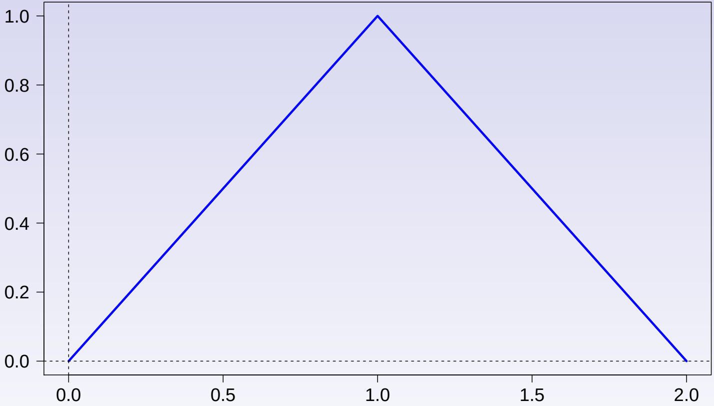
于是

$$f_Z(z) = \int_{-\infty}^{\infty} f_X(x) f_Y(z-x)\,dx = \int_0^1 I_{\{0<z-x<1\}}\,dx$$

$$\overset{z-x=t}{=\!=\!=} \int_{z-1}^{z} I_{\{0<t<1\}}\,dt$$

$$= \begin{cases} z, & \text{if } z \in [0,1), \\ 2-z, & \text{if } z \in [1,2], \\ 0, & \text{others.} \end{cases}$$

三、随机向量函数的分布



三、随机向量函数的分布

例 3.6

设 $X_1, X_2, \ldots, X_n$ 独立同分布，且服从 $\mathcal{U}[-1,1]$. 则

(1). X_1+X_2 的密度函数为：

$$f_{X_1+X_2}(x)=\begin{cases}\frac{2-|x|}{4}, & \text{if } |x|\le 2,\\ 0, & \text{if } |x|>2.\end{cases}$$

(2). $X_1+X_2+X_3$ 的密度函数为：

$$f_{X_1+X_2+X_3}(x)=\begin{cases}\frac{(3-|x|)^2}{16}, & \text{if } 1\le |x|\le 3,\\ \frac{3-x^2}{8}, & \text{if } 0\le |x|\le 1,\\ 0, & \text{if } |x|>3.\end{cases}$$

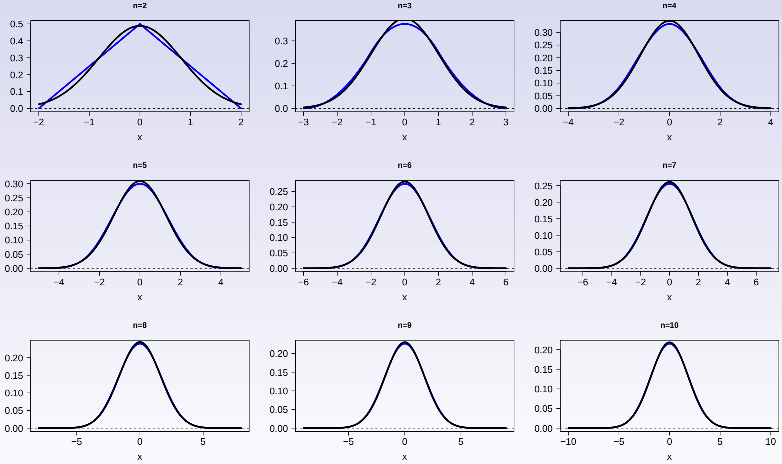
三、随机向量函数的分布

(3). $S_n = X_1 + X_2 + \ldots + X_n$ 的密度函数为:

$$f_{S_n}(x) = \begin{cases} \frac{1}{2^n \cdot (n-1)!} \sum_{k=0}^{[\frac{n+x}{2}]} (-1)^k \binom{n}{k} (n+x-2k)^{n-1}, & \text{if } |x| \leq n, \\ 0, & \text{if } |x| > n. \end{cases}$$

此处 $[a]$ 表示 a 的整数部分.

三、随机向量函数的分布



三、随机向量函数的分布

♣ 类似情形 $V = X - Y$.

例 3.7

设 (X, Y) 有联合密度 $f(x, y)$，则 $V = X - Y$ 有密度函数

$$f_V(v) = \int_{-\infty}^{\infty} f(x, x - v)\, dx.$$

特别地，当 X, Y 独立时，$V = X - Y$ 有密度函数

$$f_V(v) = \int_{-\infty}^{\infty} f_X(x) f_Y(x - v)\, dx.$$

三、随机向量函数的分布

(iii) 多个函数的联合密度

设 $\vec{X}=(X_1,\dots,X_n)$ 有联合密度 $f(\vec{x})$, $Y_1=g_1(\vec{X}),\ \dots\ ,Y_n=g_n(\vec{X})$ 是 n 个 $\vec{X}$ 的函数. $\vec{Y}=(Y_1,\dots,Y_n)$ 的联合密度?

回忆关于"随机变量函数的分布"的定理:

定理 2.1

设 X 的概率密度 $f(x)$, $D\subset R$, $Y=g(X)$, $\mathbb{P}(Y\in D)=1$, 如果

1. 对 $y\in D$, $\{Y=y\}=\cup_{i=1}^{n}\{X=h_i(y)\}$;
2. 每个 $h_i(y)$ 是 D 到其值域 D_i 的可逆映射, 在 D 内有连续的导数;
3. $D_1,D_2,\dots,D_n$ 互不相交,

则 Y 得密度函数为

$$f_Y(y)=\begin{cases}\sum_{i=1}^{n} f(h_i(y))|h_i'(y)|, & y\in D,\\ 0, & y\in D^c.\end{cases}$$

三、随机向量函数的分布

定理 3.1

设 (X, Y) 有联合密度 $f(x, y)$，$U = u(X, Y), V = v(X, Y)$ 是 (X, Y) 的函数，D 是平面上的区域使得 $\mathbb{P}((U, V) \in D) = 1$. 如果存在 D 上的函数 $x_i = x_i(u, v)$, $y_i = y_i(u, v)$, $(i = 1, 2, \dots, n)$, 使得

1. 对 $(u, v) \in D$, 有 $\{U = u, V = v\} = \bigcup_{i=1}^n \{X = x_i, Y = y_i\}$;
2. $\triangle_i : x_i = x_i(u, v), y_i = y_i(u, v)$ 是 D 到其值域 D_i 的可逆映射，有连续的偏导数，并且雅克比行列式的绝对值 $|\frac{\partial(x_i,y_i)}{\partial(u,v)}| \neq 0$, $(u, v) \in D$, $i = 1, \dots, n$.
3. $D_1, \dots, D_n$ 互不相交.

则 (U, V) 有联合密度

$$g(u, v) = \begin{cases} \sum_{i=1}^n f(x_i(u, v), y_i(u, v))|\frac{\partial(x_i,y_i)}{\partial(u,v)}|, & (u, v) \in D, \\ 0, & (u, v) \in D^c. \end{cases}$$

例 3.8

设 X, Y 独立，$\sim \mathcal{N}(0,1)$，(R,Θ) 由极坐标变换

$$\begin{cases} X = R\cos\Theta, \\ Y = R\sin\Theta \end{cases}$$

决定，求 (R,Θ) 的联合密度.

解. (X, Y) 有联合密度

$$f(x,y) = \frac{1}{2\pi}\exp\left(-\frac{x^2+y^2}{2}\right).$$

定义

$$D = \{(r,\theta)|r>0, \theta \in [0,2\pi)\}$$

则 $\mathbb{P}((R,\Theta) \in D) = 1.$

三、随机向量函数的分布

$$\triangle = \begin{cases} x = r\cos\theta, \\ y = r\sin\theta \end{cases}$$

建立了 D 到 $D_1 = \{(x,y)|(x,y) \neq 0\}$ 的可逆变换. 对 $(r,\theta) \in D$, 利用

$$\{R = r, \Theta = \theta\} = \{X = x, Y = y\}, \quad \frac{\partial(x,y)}{\partial(r,\theta)} = r > 0,$$

得 (R,Θ) 的联合密度

$$g(r,\theta) = f(r\cos\theta, r\sin\theta)r = \frac{1}{2\pi} \cdot r\exp(-r^2/2), \quad (r,\theta) \in D.$$

从中可以看出 R 和 Θ 独立，分别有概率密度

$$g_R(r) = r\exp(-r^2/2), \quad r \geq 0; \qquad g_\Theta(\theta) = \frac{1}{2\pi} I_{[0,2\pi)}.$$

即 R 是 Rayleigh 分布，$\Theta \sim \mathcal{U}[0,2\pi)$.

三、随机向量函数的分布

定理 3.2

设 2 个 n 维随机向量 $\mathbf{X}$ 和 $\mathbf{Y}$ 有相同的分布，$\mathbf{g}: R^n \hookrightarrow R^m$ 是可测的，则 $\mathbf{g}(\mathbf{X})$ 与 $\mathbf{g}(\mathbf{Y})$ 有相同的分布.

例 3.9

设 U, V 独立，$\sim \mathcal{U}(0,1)$，定义

$$\begin{cases} X = \sqrt{-2\log U_1}\ \cos(2\pi V), \\ Y = \sqrt{-2\log U_1}\ \sin(2\pi V). \end{cases}$$

则 X, Y 独立，且服从 $\sim \mathcal{N}(0,1)$. 【用来产生正态随机变量】

证明. $R = \sqrt{-2\log U_1}$ ~Rayleigh 分布，$\Theta = 2\pi V \sim \mathcal{U}[0, 2\pi)$，且 R 与 Θ 独立. 由定理3.2易得结论成立.

三、随机向量函数的分布

例 3.10

设 X, Y 独立，$\sim \mathcal{N}(0,1)$，求 $U = X/Y$ 和 $V = X^2 + Y^2$ 的联合密度.

解.

$$g(u,v) = \frac{1}{\pi(1+u^2)} \cdot \frac{1}{2}\mathrm{e}^{-v/2}, \quad v \geq 0, u \in R.$$

即：U, V 独立，且 U 服从 Cauchy 分布，V 服从指数分布.

♣ 定理可以直接从 2 维推广到 n 维.

小结

知识点

- 随机变量函数的分布
- 随机向量：联合分布、边缘分布、独立性的判定；离散型：多项分布、连续型：多元正态分布
- 随机向量函数的分布

技巧

- 2 套工具间的转换：一套是 CDF；一套是 PMF/PDF.
- 特别地，利用变量间的关系，便捷求解随机变（向）量函数的分布，例如卷积法、求导法
- 识别分布 (PMF/PDF) 的核心部分，根据归一化原则确认 PMF/PDF 的完整表达式.

作业

教材第二章：14，15，24(a);
第三章：6,7(b),15(i)(ii)(不求 $E[Y]$),23(a)(b)(c);
第四章：1, 3, 4.
随机向量函数的分布：第四章：5, 8, 10, 14.

祝大家期中考试顺利！