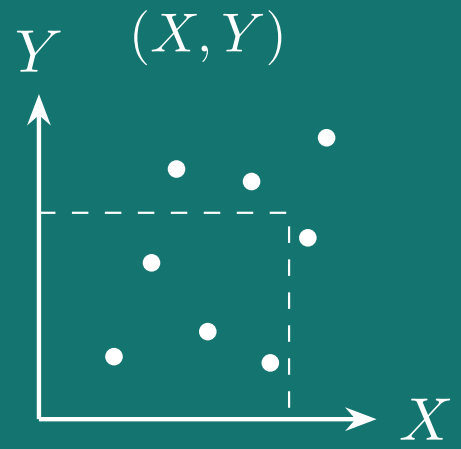
概率论与数理统计

第三章：多维随机变量及其分布

联合分布　边缘分布　条件分布　相互独立



叶文杰

yewenjie@fjnu.edu.cn

yewenjie.org

2026 年秋季学期

第三章学习内容

从一个随机变量到一组随机变量

除了各自的取值规律，还要描述随机变量之间的联系。本章以二维随机向量 (X, Y) 为主线。

1. **联合分布**：(X, Y) 同时取值的概率规律。
2. **边缘分布**：从联合分布得到 X 和 Y 各自的分布。
3. **条件分布**：已知一个变量的信息后，另一个变量如何分布。
4. **相互独立**：何时联合分布可以由边缘分布确定。

本章的核心问题

给定两个边缘分布，是否就能确定联合分布？一般不能；还需要知道变量之间的关系。

1

二维随机变量及其联合分布

第三章　多维随机变量及其分布

同一试验中的多个随机变量

一次试验的结果有时需要多个数据共同描述：打靶的命中点需要两个坐标，飞机在空中的位置需要三个坐标。

随机向量

设 X, Y 是定义在同一样本空间 S 上的随机变量，则 (X, Y) 称为**二维随机向量**或**二维随机变量**。

$$(X, Y) : S \longrightarrow \mathbb{R}^2, \qquad e \longmapsto \big(X(e), Y(e)\big).$$



一般地，$(X_1, \ldots, X_n)$ 称为 n 维随机向量。本章重点讨论二维情形。

需要描述共同取值的规律

(X, Y) 的分布不仅涉及 X, Y 各自的分布，还涉及它们之间的关系。

联合分布函数：左下方的概率

定义

二维随机变量 (X, Y) 的**联合分布函数**定义为

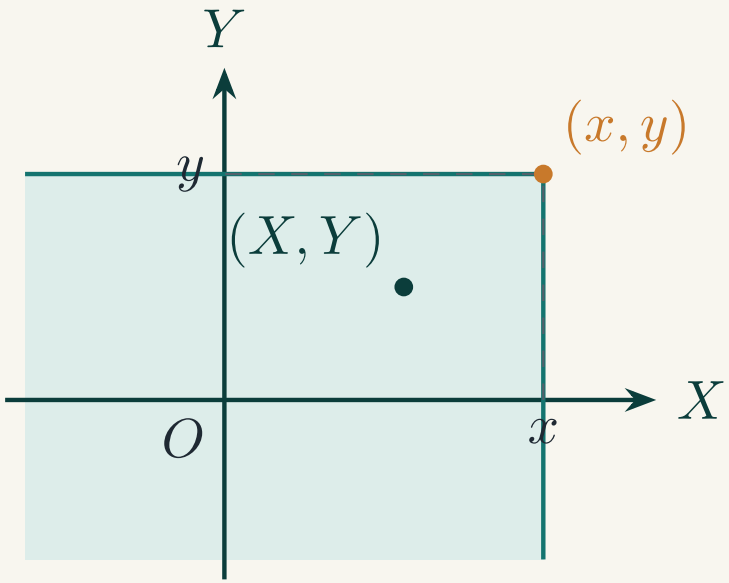
$$F(x, y) = \mathbb{P}(X \le x, Y \le y), \qquad (x, y) \in \mathbb{R}^2.$$

把 (X, Y) 看作平面上的随机点，
则 $F(x, y)$ 是该点落在矩形

$$(-\infty, x] \times (-\infty, y]$$

内的概率。

它同时限制两个坐标，是两个事件交集的概率。



一般不能直接相乘

一般情况下，$F(x, y) \ne \mathbb{P}(X \le x)\mathbb{P}(Y \le y)$。相乘需要额外条件。

用分布函数计算矩形概率

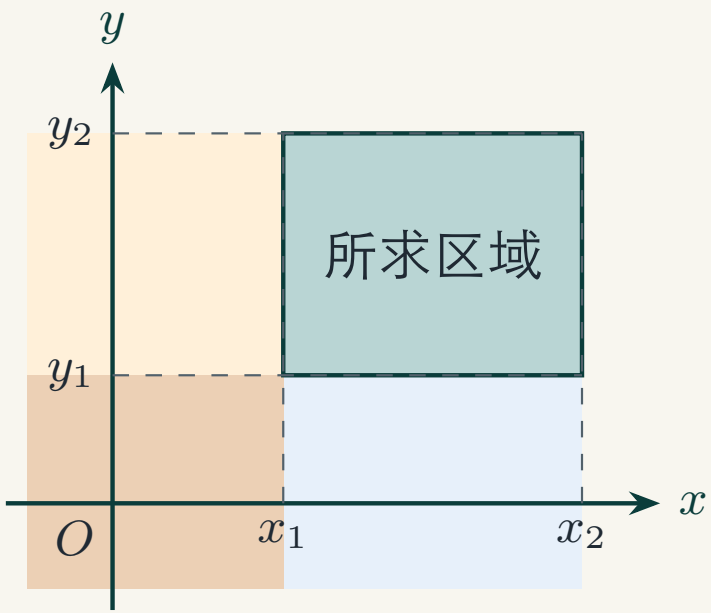
设 $x_1 < x_2$，$y_1 < y_2$。用容斥公式可得

$$\mathbb{P}(x_1 < X \le x_2,\, y_1 < Y \le y_2) = F(x_2, y_2) - F(x_1, y_2) - F(x_2, y_1) + F(x_1, y_1).$$

计算顺序

1. 从以 (x_2, y_2) 为顶点的左下方矩形出发；
2. 减去左侧和下侧多出的部分；
3. 加回被减去两次的左下角。

半开半闭区间对应分布函数中的“$\le$”，离散型情形尤其要留意端点。



联合分布函数的基本性质

1. **分别单调不减**：固定一个坐标，增大另一个坐标，概率不会减小。

$$x_1 < x_2 \Rightarrow F(x_1, y) \le F(x_2, y), \qquad y_1 < y_2 \Rightarrow F(x, y_1) \le F(x, y_2).$$

2. **取值与边界**：$0 \le F(x, y) \le 1$，并且

$$\lim_{x \to -\infty} F(x, y) = 0, \qquad \lim_{y \to -\infty} F(x, y) = 0, \qquad \lim_{t \to +\infty} F(t, t) = 1.$$

3. **分别右连续**：对每个固定的 x, y，

$$\lim_{h \downarrow 0} F(x + h, y) = F(x, y), \qquad \lim_{h \downarrow 0} F(x, y + h) = F(x, y).$$

为什么右连续?

当 $h \downarrow 0$ 时，事件 $\{X \le x + h, Y \le y\}$ 单调递减到 $\{X \le x, Y \le y\}$；由概率的上连续性得到结论。另一个坐标同理。

矩形概率必须非负

设 $x_1 < x_2$，$y_1 < y_2$。由上一页的矩形概率公式，

四个函数值的组合，就是矩形内的概率

$$\underbrace{F(x_2, y_2) - F(x_1, y_2) - F(x_2, y_1) + F(x_1, y_1)}_{\mathbb{P}(x_1 < X \le x_2,\ y_1 < Y \le y_2)} \ge 0.$$

概率不能为负，所以这个四项组合必须非负。

为什么分别单调不减还不够?

假设函数在四个角上的取值如下：

$F(x,y)$	y_1	y_2
x_1	0.1	0.5
x_2	0.5	0.6

每行、每列都递增，但矩形增量为

$$0.6 - 0.5 - 0.5 + 0.1 = -0.3 < 0.$$

这会产生负的矩形概率，因此这四个值不可能来自联合分布函数。

一维要保证区间概率非负；二维还要保证每个矩形的概率非负。

二维离散型随机变量与联合分布律

若 (X, Y) 以概率 1 取值于有限或可列的点集，则称它为**二维离散型随机变量**。

联合分布律

设 X 的可能取值为 x_i，Y 的可能取值为 y_j，记

$$p_{ij} = \mathbb{P}(X = x_i, Y = y_j), \qquad p_{ij} \geq 0, \qquad \sum_i \sum_j p_{ij} = 1.$$

$X \backslash Y$	y_1	y_2	$\cdots$	y_j
x_1	p_{11}	p_{12}	$\cdots$	p_{1j}
x_2	p_{21}	p_{22}	$\cdots$	p_{2j}
$\vdots$	$\vdots$	$\vdots$		$\vdots$
x_i	p_{i1}	p_{i2}	$\cdots$	p_{ij}

$$F(x, y) = \sum_{i:\, x_i \leq x} \sum_{j:\, y_j \leq y} p_{ij}.$$

表中可出现 0；它表示该组合不会发生。求概率就是把相应点上的概率相加。

例 1：三次抛硬币的两个统计量

将一枚均匀硬币独立抛掷三次。设 X 为正面出现的次数，Y 为正面与反面出现次数之差的绝对值，求 (X, Y) 的分布律。

先找两个变量之间的关系

反面出现 $3 - X$ 次，因此

$$Y = |X - (3 - X)| = |2X - 3|, \qquad X \sim \mathrm{Bin}(3, 1/2).$$

$$\mathbb{P}(X = k) = \binom{3}{k} \left(\frac{1}{2}\right)^3.$$

可能的数对只有

$$(0, 3),\ (1, 1),\ (2, 1),\ (3, 3).$$

$X \backslash Y$	1	3
0	0	1/8
1	3/8	0
2	3/8	0
3	0	1/8

例 2：第二次取值的上限由第一次决定

X 在 $1,2,3,4$ 中等可能取值；给定 $X=i$ 后，Y 在 $1,\dots,i$ 中等可能取值。求 (X,Y) 的分布律。

乘法公式

$$\mathbb{P}(X=i,Y=j)=\mathbb{P}(X=i)\mathbb{P}(Y=j\mid X=i)=\frac{1}{4i},\qquad 1\le j\le i\le 4.$$

其余数对的概率为 0。

$X\backslash Y$	1	2	3	4
1	1/4	0	0	0
2	1/8	1/8	0	0
3	1/12	1/12	1/12	0
4	1/16	1/16	1/16	1/16

检查：每一行的和为 $1/4$，全表的和为 1。

例 3：不放回抽取两个球

袋中三个球分别标有数字 $1, 2, 2$。依次等可能、不放回地抽取两个球，用 X, Y 记录第一次、第二次取到的数字。求联合分布律与分布函数。



$$\mathbb{P}(X=1, Y=2) = \frac{1}{3} \cdot 1 = \frac{1}{3}.$$

$$\mathbb{P}(X=2, Y=1) = \frac{2}{3} \cdot \frac{1}{2} = \frac{1}{3}.$$

$$\mathbb{P}(X=2, Y=2) = \frac{2}{3} \cdot \frac{1}{2} = \frac{1}{3}.$$

只有一个标有 1 的球，因此 $(1,1)$ 不可能出现。

$X\backslash Y$	1	2
1	0	1/3
2	1/3	1/3

二维连续型随机变量与联合密度

定义：联合概率密度

若存在非负可测函数 $f(x,y)$，使得对任意 $(x,y)\in\mathbb{R}^2$，

$$F(x,y)=\int_{-\infty}^{x}\int_{-\infty}^{y} f(u,v)\,\mathrm{d}v\,\mathrm{d}u,$$

则称 (X,Y) 为**二维连续型随机变量**，f 为其**联合概率密度**。

这里的“连续型”指分布具有联合密度，也称**绝对连续型**。

联合密度的基本性质

1. **非负性与归一性**：

$$f(x,y) \geq 0 \ \text{（几乎处处）}, \qquad \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f(x,y)\,\mathrm{d}y\,\mathrm{d}x = 1.$$

2. **小矩形的概率**：若 f 在 (x,y) 连续，则当 $\Delta x, \Delta y \downarrow 0$ 时，

$$\frac{\mathbb{P}(x < X \leq x+\Delta x,\, y < Y \leq y+\Delta y)}{\Delta x\,\Delta y} \longrightarrow f(x,y).$$

   因而对很小的矩形，可用 $f(x,y)\Delta x \Delta y$ 近似它的概率。

3. **由分布函数求密度**：若 F 二阶连续可微，则可取

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

判断一个函数能否成为联合密度，先检查非负性，再检查全平面积分是否为 1。

区域概率：对密度积分

对平面上的可测区域 G，

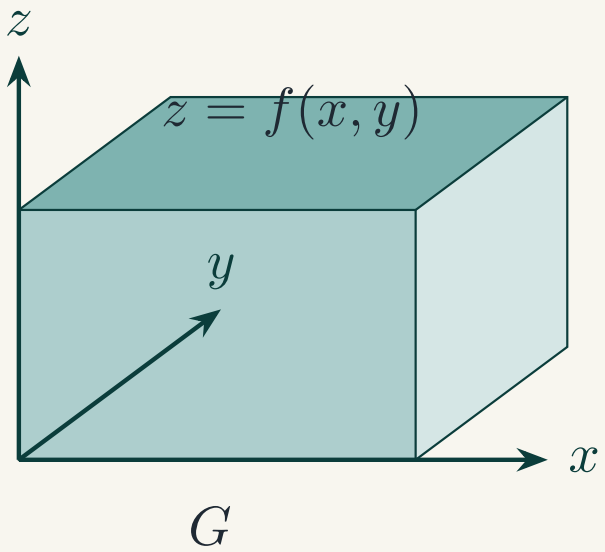
$$\mathbb{P}\big((X,Y)\in G\big)=\iint_G f(x,y)\,\mathrm{d}x\,\mathrm{d}y.$$

几何解释

曲面 $z=f(x,y)$ 下方、区域 G 上方的体积，就是随机点落入 G 的概率。

全平面对应的总体积为 1。

计算前先画出 G 与密度支撑集的交集，再决定积分上下限。



图中用常数密度表示“概率等于体积”。

例 4：指数型联合密度的分布函数

设 (X, Y) 的联合密度为

$$f(x,y) = \begin{cases} 2\mathrm{e}^{-(2x+y)}, & x > 0,\ y > 0, \\ 0, & \text{其他.} \end{cases}$$

求联合分布函数 $F(x,y)$，并求 $\mathbb{P}(Y \le X)$。

先求分布函数

当 $x > 0, y > 0$ 时，积分区域与支持集的交集为 $(0,x] \times (0,y]$，故

$$F(x,y) = \int_0^x \int_0^y 2\mathrm{e}^{-(2u+v)}\,\mathrm{d}v\,\mathrm{d}u = (1-\mathrm{e}^{-2x})(1-\mathrm{e}^{-y}).$$

当 $x \le 0$ 或 $y \le 0$ 时，交集为空，$F(x,y) = 0$。

$$F(x,y) = \begin{cases} (1-\mathrm{e}^{-2x})(1-\mathrm{e}^{-y}), & x > 0,\ y > 0, \\ 0, & \text{其他.} \end{cases}$$

例 4（续）：先确定事件对应的区域

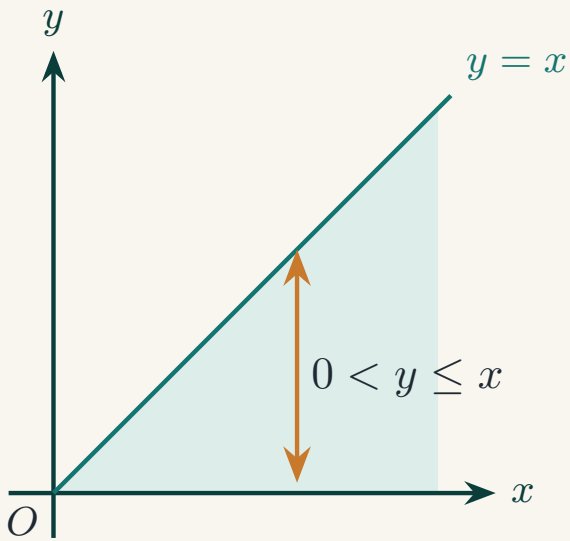
事件 $\{Y \le X\}$ 对应直线 $y = x$ 下方。与支撑集相交后，实际积分区域为

$$G = \{(x, y) : x > 0,\ 0 < y \le x\}.$$

$$\mathbb{P}(Y \le X) = \int_0^{\infty} \int_0^{x} 2\mathrm{e}^{-(2x+y)}\,\mathrm{d}y\,\mathrm{d}x.$$

$$\mathbb{P}(Y \le X) = 2\int_0^{\infty} (\mathrm{e}^{-2x} - \mathrm{e}^{-3x})\,\mathrm{d}x.$$

$$\mathbb{P}(Y \le X) = 2\left(\frac{1}{2} - \frac{1}{3}\right) = \frac{1}{3}.$$



积分上限来自图形

固定 $x > 0$，y 从 0 移动到 x。图中区域向右无限延伸。

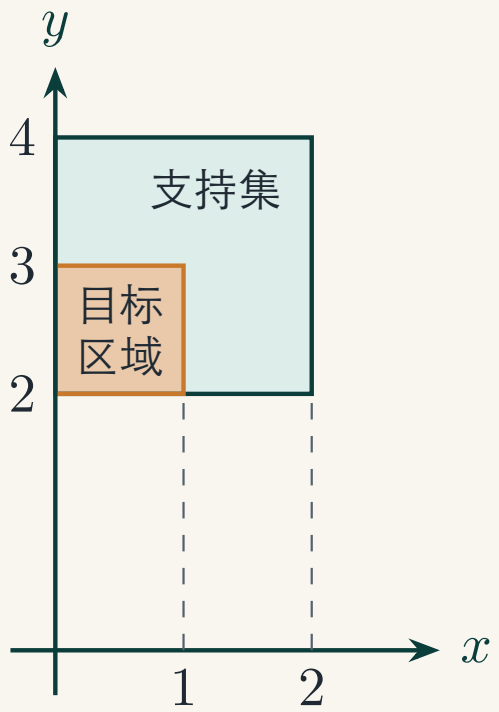
练习：确定常数并计算区域概率

设 (X, Y) 的联合密度为

$$f(x,y) = \begin{cases} k(6-x-y), & 0 < x < 2,\ 2 < y < 4, \\ 0, & \text{其他.} \end{cases}$$

1. 确定常数 k；
2. 求 $\mathbb{P}(X < 1, Y < 3)$。

先在整个支撑集上归一化，再在目标区域与支撑集的交集上积分。



练习解答:

(1) 确定常数

在支撑集内，$6-x-y>0$。由归一性，

$$1=k\int_0^2\int_2^4(6-x-y)\,\mathrm{d}y\,\mathrm{d}x=2k\int_0^2(3-x)\,\mathrm{d}x=8k.$$

故 $k=1/8$。

(2) 计算概率

目标区域与支撑集的交集为 $(0,1)\times(2,3)$，所以

$$\mathbb{P}(X<1,Y<3)=\frac{1}{8}\int_0^1\int_2^3(6-x-y)\,\mathrm{d}y\,\mathrm{d}x=\frac{1}{8}\int_0^1\left(\frac{7}{2}-x\right)\mathrm{d}x=\frac{3}{8}.$$

有联合密度时，直线边界的概率为 0；这里将严格不等号改为非严格不等号不会改变结果。

常用分布之一：平面区域上的均匀分布

设 $D \subset \mathbb{R}^2$ 是可测区域，面积满足 $0 < S < \infty$。若

$$f(x,y) = \begin{cases} 1/S, & (x,y) \in D, \\ 0, & (x,y) \notin D, \end{cases}$$

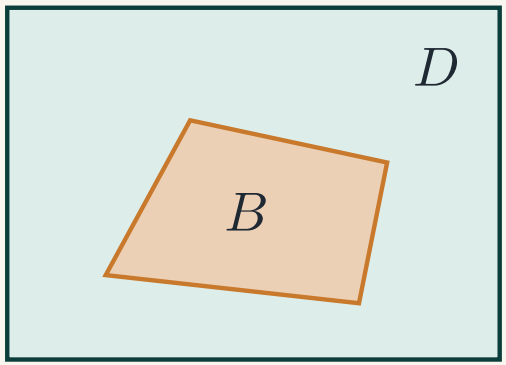
则称 (X,Y) 在 D 上服从**均匀分布**。

概率等于面积比例

对任意可测区域 B，

$$\mathbb{P}\big((X,Y) \in B\big) = \frac{\text{area}(B \cap D)}{\text{area}(D)}.$$

特别地，$B \subset D$ 时，概率仅由 B 的面积决定，与其形状或位置无关。



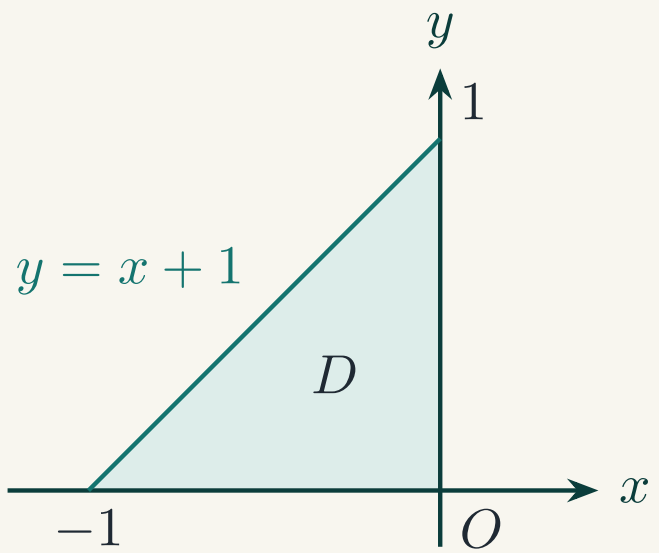
例 5：三角形上的均匀分布

设 (X, Y) 在由 x 轴、y 轴及直线 $y = x + 1$ 围成的三角形 D 上均匀分布，求联合密度与分布函数。

$$D = \{(x, y) : -1 \le x \le 0,\ 0 \le y \le x + 1\}.$$

三角形的面积为 $S = 1/2$，因此

$$f(x, y) = \begin{cases} 2, & (x, y) \in D, \\ 0, & (x, y) \notin D. \end{cases}$$



分布函数的计算思路

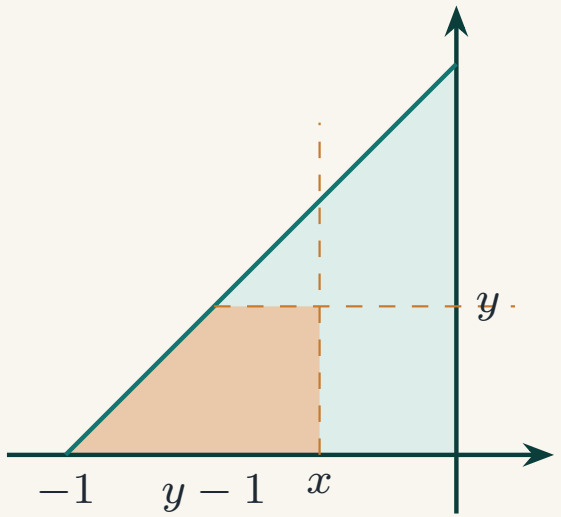
把 D 与左下方矩形 $(-\infty, x] \times (-\infty, y]$ 相交，再将交集面积乘以 2。

例 5（续）：横线截过三角形

若 $x < -1$ 或 $y < 0$，截出的区域面积为 0，所以 $F(x, y) = 0$。以下设 $-1 \le x < 0$。

(1) $0 \le y < x + 1$

$$F(x, y) = \int_{-1}^{y-1} 2(x+1)\,\mathrm{d}x + \int_{y-1}^{x} 2y\,\mathrm{d}u = (2x - y + 2)y.$$



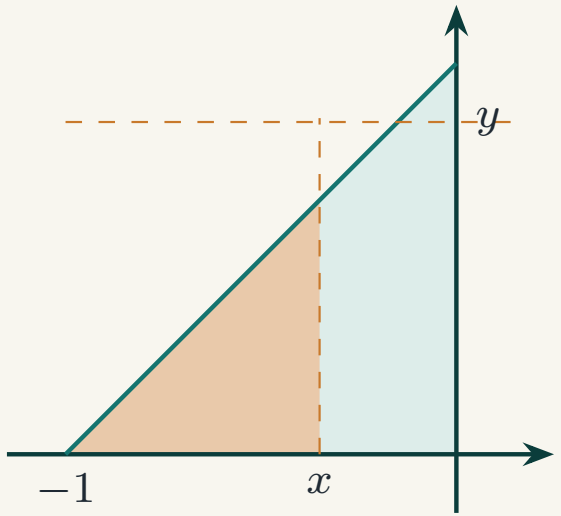
例 5（续）：纳入左侧小三角形

继续设 $-1 \le x < 0$。当 $y \ge x+1$ 时，横线位于左侧小三角形之上。

(2) $y \ge x+1$

左侧小三角形已全部纳入：

$$F(x,y) = \int_{-1}^{x} 2(u+1)\,\mathrm{d}u = (x+1)^2.$$



此时再增大 y 不会纳入新的点，因此分布函数只随 x 变化。

例 5（续）：当 $x \geq 0$ 时

此时 $X \leq x$ 已不再限制随机点，只需看横线 $v = y$ 截出多少面积。

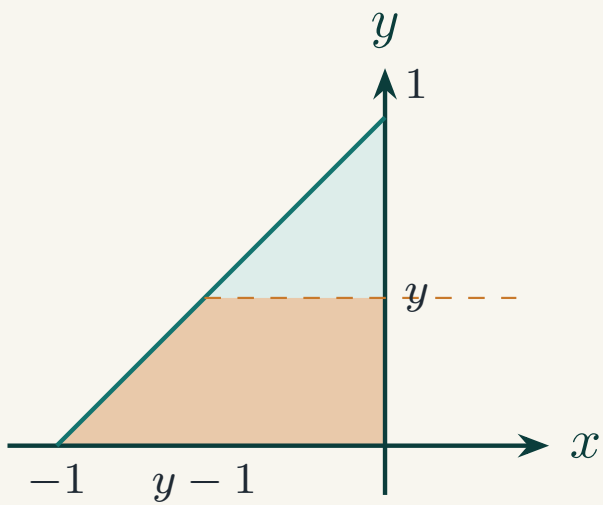
(3) $0 \leq y < 1$

$$F(x,y) = \int_{-1}^{y-1} 2(x+1)\,\mathrm{d}x + \int_{y-1}^{0} 2y\,\mathrm{d}u = (2-y)y.$$

(4) $y \geq 1$

整个三角形都被纳入：

$$F(x,y) = \int_{-1}^{0} 2(x+1)\,\mathrm{d}x = 1.$$



例 5（续）：完整的分布函数

综合各个区域，得到

$$F(x,y)=\begin{cases}0, & x<-1 \text{ 或} y<0,\\ (2x-y+2)y, & -1\le x<0,\ 0\le y<x+1,\\ (x+1)^2, & -1\le x<0,\ y\ge x+1,\\ (2-y)y, & x\ge 0,\ 0\le y<1,\\ 1, & x\ge 0,\ y\ge 1.\end{cases}$$

如何检查分段结果?

- 在分界线 $y=x+1$ 上，$(2x-y+2)y=(x+1)^2$；
- 在分界线 $x=0$ 上，$(2x-y+2)y=(2-y)y$；
- 在 $y=1$ 上，$(2-y)y=1$。

有密度的分布函数连续；相邻区域的公式应在共同边界上衔接。

常用分布之二：二维正态分布

设 $\mu_1, \mu_2 \in \mathbb{R}$，$\sigma_1, \sigma_2 > 0$，$-1 < \rho < 1$。记标准化坐标

$$u = \frac{x - \mu_1}{\sigma_1}, \qquad v = \frac{y - \mu_2}{\sigma_2}.$$

若 (X, Y) 的联合密度为

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

其中 $(x, y) \in \mathbb{R}^2$，则称其服从**二维正态分布**，记为

$$(X, Y) \sim \mathrm{N}(\mu_1, \mu_2, \sigma_1^2, \sigma_2^2, \rho).$$

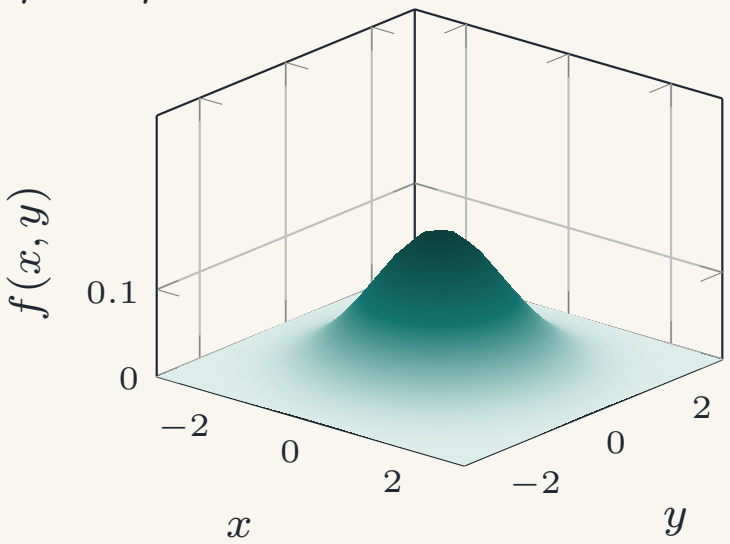
参数如何影响图形?

- μ_1, μ_2 决定密度中心的位置；
- σ_1, σ_2 决定两个坐标方向的尺度；
- ρ 决定等密度椭圆的形状和倾斜方向。

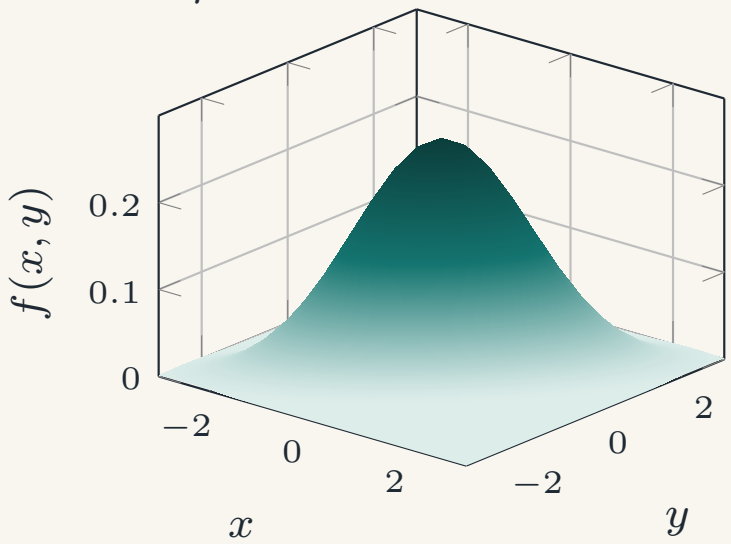
本节只讨论非退化情形（$|\rho| < 1$）；$\rho = \pm 1$ 时不适用此密度。

二维正态密度的图形

取 $\mu_1 = \mu_2 = 0$、$\sigma_1 = \sigma_2 = 1$。下面比较 $\rho = 0$ 与 $\rho = 0.8$。



$\rho = 0$：等密度线为圆。



$\rho = 0.8$：密度沿 $y = x$ 方向延伸。

两幅图采用相同坐标范围与高度范围；所有图形均由密度公式直接绘制。

本节小结：三种描述方式

联合分布函数适用于一般二维随机变量

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

离散型：把点概率相加

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

$$\mathbb{P}\big((X,Y)\in G\big)=\sum_{(x_i,y_j)\in G}p_{ij}.$$

非负，全部概率之和为 1。

有密度型：对区域积分

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

$$\mathbb{P}\big((X,Y)\in G\big)=\iint_{G}f(x,y)\,\mathrm{d}x\,\mathrm{d}y.$$

非负，全平面积分为 1。

共同的步骤：把事件写成平面区域，再依据分布类型求和或积分。

本节作业

课后习题

第五版教材 86 页：第 $1, 2, 3$ 题;
第六版教材 104 页：第 $1, 2, 3$ 题;

完成时留意

- 分布律是否遗漏了可能的数对? 全表之和是否为 1?
- 分布函数的分段是否覆盖整个平面? 端点是否正确?
- 积分区域是否已经与密度的支持集相交?

2

边缘分布

第三章　多维随机变量及其分布

从联合分布到边缘分布

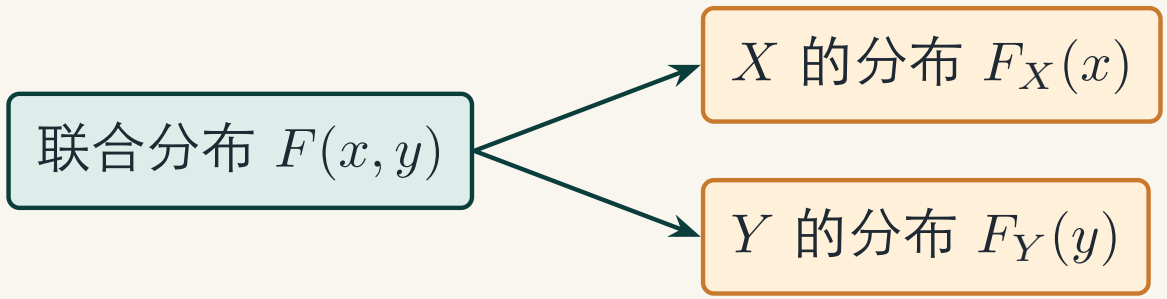
二维随机变量 (X, Y) 的联合分布描述两个变量的整体取值规律。只关心其中一个变量时，得到它的**边缘分布**。

边缘分布函数

设 $F(x, y) = \mathbb{P}(X \le x, Y \le y)$，则

$$F_X(x) = \mathbb{P}(X \le x) = \lim_{y \to +\infty} F(x, y) =: F(x, +\infty),$$

$$F_Y(y) = \mathbb{P}(Y \le y) = \lim_{x \to +\infty} F(x, y) =: F(+\infty, y).$$



这里的 $+\infty$ 表示取极限；从联合分布中消去不关心的那个变量。

离散型：沿另一个变量求和

设 (X, Y) 为二维离散型随机变量，联合分布律为

$$p_{ij} = \mathbb{P}(X = x_i, Y = y_j), \qquad i, j = 1, 2, \ldots.$$

边缘分布律

$$p_{i\cdot} = \mathbb{P}(X = x_i) = \sum_j p_{ij}, \qquad p_{\cdot j} = \mathbb{P}(Y = y_j) = \sum_i p_{ij}.$$

$Y \backslash X$	x_1	x_2	$\cdots$	行和
y_1	p_{11}	p_{21}	$\cdots$	$p_{\cdot 1}$
y_2	p_{12}	p_{22}	$\cdots$	$p_{\cdot 2}$
$\vdots$	$\vdots$	$\vdots$		$\vdots$
列和	$p_{1\cdot}$	$p_{2\cdot}$	$\cdots$	1

读表方法

每列求和得到 X 的边缘分布律；每行求和得到 Y 的边缘分布律。

$$F_X(x) = \sum_{x_i \le x} p_{i\cdot},$$

$$F_Y(y) = \sum_{y_j \le y} p_{\cdot j}.$$

例 1：由联合分布律求边缘分布律

已知 (X, Y) 的联合分布律如下。按列、按行分别求和即可得到两个边缘分布律。

$Y \backslash X$	0	1	$\mathbb{P}(Y = y)$
0	$\frac{16}{49}$	$\frac{12}{49}$	$\frac{4}{7}$
1	$\frac{12}{49}$	$\frac{9}{49}$	$\frac{3}{7}$
$\mathbb{P}(X = x)$	$\frac{4}{7}$	$\frac{3}{7}$	1

$$\mathbb{P}(X = 0) = \frac{16 + 12}{49} = \frac{4}{7}, \qquad \mathbb{P}(Y = 1) = \frac{12 + 9}{49} = \frac{3}{7}.$$

信息的方向

联合分布能确定两个边缘分布；两个边缘分布通常不能确定联合分布。

为什么边缘分布不能确定联合分布?

下面两张表都满足 $\mathbb{P}(X=0)=\mathbb{P}(Y=0)=4/7$，但它们的联合分布不同。

例 1 的联合分布

$Y\backslash X$	0	1	行和
0	$16/49$	$12/49$	$4/7$
1	$12/49$	$9/49$	$3/7$
列和	$4/7$	$3/7$	1

另一个联合分布

$Y\backslash X$	0	1	行和
0	$17/49$	$11/49$	$4/7$
1	$11/49$	$10/49$	$3/7$
列和	$4/7$	$3/7$	1

相同的行和、列和，可以对应不同的表格内部数值。

边缘分布只保留各变量自身的取值规律，无法完整保留它们之间的联系。

例 2：随机整数的约数个数

整数 N 等可能地取 $1, 2, \ldots, 10$ 中的一个值。令 D 为 N 的正约数个数，F 为 N 的不同素因子个数。求 (D, F) 的联合分布律及边缘分布律。

N	1	2	3	4	5	6	7	8	9	10
D	1	2	2	3	2	4	2	4	3	4
F	0	1	1	1	1	2	1	1	1	2

先逐个列举，再合并相同的数对

$N = 8$ 时，正约数为 $1, 2, 4, 8$，不同素因子只有 2，所以 $(D, F) = (4, 1)$。$N = 6$ 或 10 时，均有 $(D, F) = (4, 2)$。

每个整数出现的概率为 $1/10$；同一数对的概率等于对应整数的个数除以 10。

例 2：联合分布律与两个边缘分布律

$F\backslash D$	1	2	3	4	$\mathbb{P}(F=j)$
0	1/10	0	0	0	1/10
1	0	4/10	2/10	1/10	7/10
2	0	0	0	2/10	2/10
$\mathbb{P}(D=i)$	1/10	4/10	2/10	3/10	1

也可将边缘分布律单独列出：

D	1	2	3	4
概率	1/10	4/10	2/10	3/10

F	0	1	2
概率	1/10	7/10	2/10

连续型：沿另一个变量积分

设 (X, Y) 的联合概率密度为 $f(x, y)$。对任意 $x \in \mathbb{R}$，

$$F_X(x) = \mathbb{P}(X \le x) = \int_{-\infty}^{x} \left[\int_{-\infty}^{+\infty} f(u, v)\, \mathrm{d}v \right] \mathrm{d}u.$$

边缘概率密度

$$f_X(x) = \int_{-\infty}^{+\infty} f(x, y)\, \mathrm{d}y, \qquad f_Y(y) = \int_{-\infty}^{+\infty} f(x, y)\, \mathrm{d}x.$$

与离散型公式对应

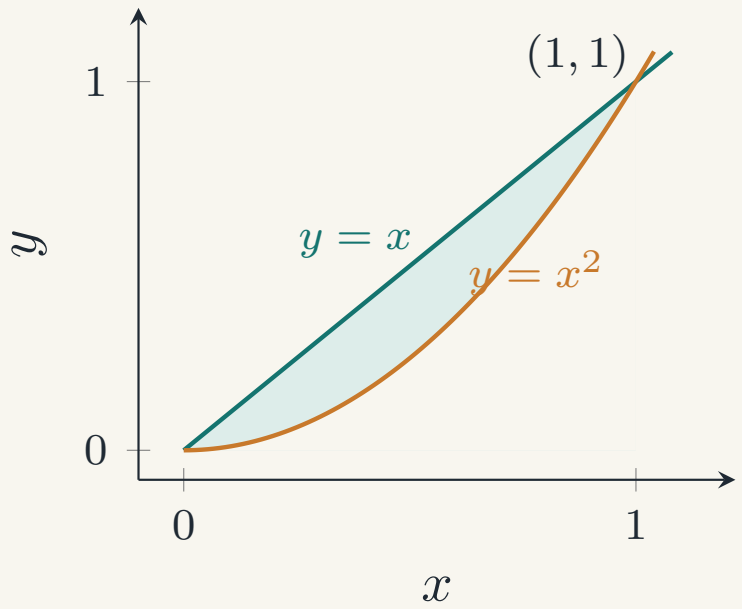
固定所关心的变量，沿另一个变量的全部可能取值求和或积分。

非负性保证可以交换积分次序。边缘密度公式在几乎处处的意义下成立；密度在零测集上的取值不影响分布。

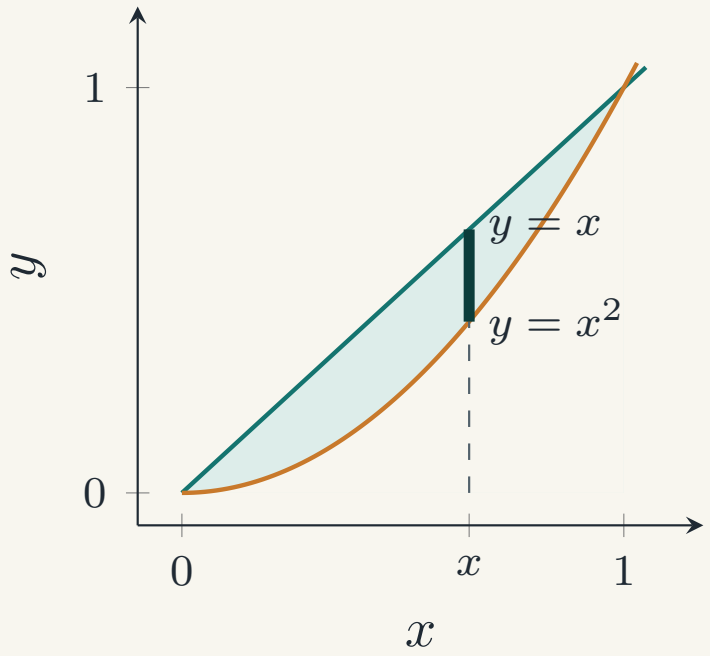
例 3：曲线之间的联合密度

设 (X, Y) 的联合密度为

$$f(x,y) = \begin{cases} 6, & x^2 \le y \le x, \\ 0, & \text{其他}. \end{cases}$$



例 3：固定 x，沿竖直方向积分



当 $0 \le x \le 1$ 时，竖直截线与非零区域相交于 $x^2 \le y \le x$，于是

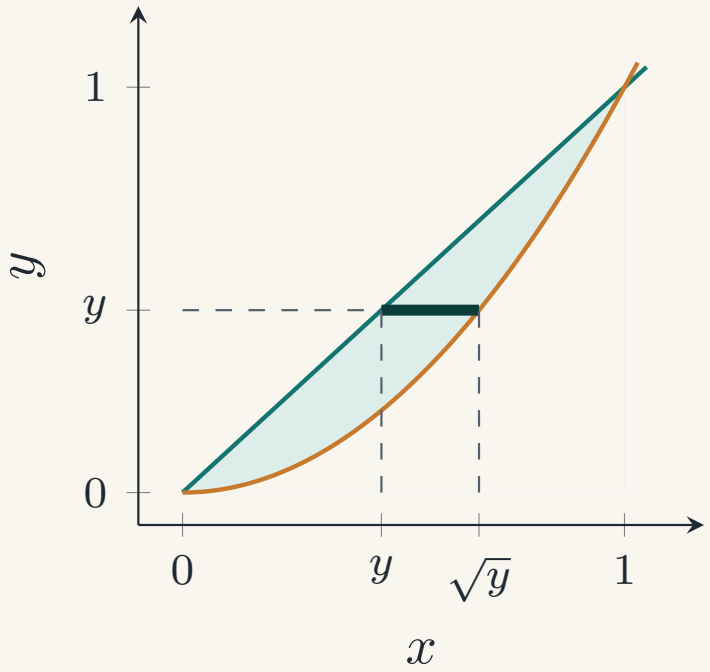
$$f_X(x) = \int_{x^2}^{x} 6 \, \mathrm{d}y = 6(x - x^2).$$

当 $x < 0$ 或 $x > 1$ 时，截线与该区域不相交。

结论

$$f_X(x) = \begin{cases} 6(x - x^2), & 0 \le x \le 1, \\ 0, & \text{其他}. \end{cases}$$

例 3：固定 y，沿水平方向积分



当 $0 \le y \le 1$ 时，条件 $x^2 \le y \le x$ 等价于

$$y \le x \le \sqrt{y}.$$

因此

$$f_Y(y) = \int_y^{\sqrt{y}} 6 \, \mathrm{d}x = 6(\sqrt{y} - y).$$

结论

$$f_Y(y) = \begin{cases} 6(\sqrt{y} - y), & 0 \le y \le 1, \\ 0, & \text{其他}. \end{cases}$$

例 4：二维正态分布的边缘分布

设 (X, Y) 的联合密度为

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

其中 $(x, y) \in \mathbb{R}^2$，$\mu_1, \mu_2 \in \mathbb{R}$，$\sigma_1, \sigma_2 > 0$，$|\rho| < 1$。

求两个边缘密度

$$f_X(x) = \int_{-\infty}^{+\infty} f(x,y)\,\mathrm{d}y, \qquad f_Y(y) = \int_{-\infty}^{+\infty} f(x,y)\,\mathrm{d}x.$$

计算方法

固定 x，对指数中关于 y 的二次式配方，再对 y 积分。求 f_Y 的方法相同。

例 4：对指数中的二次式配方

固定 x，把与 y 无关的项拆成两部分：

$$\frac{(x-\mu_1)^2}{\sigma_1^2} = (1-\rho^2)\frac{(x-\mu_1)^2}{\sigma_1^2} + \frac{\rho^2(x-\mu_1)^2}{\sigma_1^2}.$$

后一部分与原式中含 y 的两项组成完全平方：

$$\frac{\rho^2(x-\mu_1)^2}{\sigma_1^2} - \frac{2\rho(x-\mu_1)(y-\mu_2)}{\sigma_1\sigma_2} + \frac{(y-\mu_2)^2}{\sigma_2^2} = \left(\frac{y-\mu_2}{\sigma_2} - \rho\frac{x-\mu_1}{\sigma_1}\right)^2.$$

因此，原密度中的指数可写为

$$-\frac{(x-\mu_1)^2}{2\sigma_1^2} - \frac{\left(y-\mu_2-\rho\frac{\sigma_2}{\sigma_1}(x-\mu_1)\right)^2}{2\sigma_2^2(1-\rho^2)}.$$

例 4：对 y 积分求边缘分布

固定 x，由高斯积分公式，

$$\int_{-\infty}^{+\infty} \exp\left[-\frac{\left(y-\mu_2-\rho\frac{\sigma_2}{\sigma_1}(x-\mu_1)\right)^2}{2\sigma_2^2(1-\rho^2)}\right] \mathrm{d}y = \sqrt{2\pi}\sigma_2\sqrt{1-\rho^2}.$$

将与 y 无关的因子提出积分号，得到

$$\begin{aligned} f_X(x) &= \frac{\exp\left[-\frac{(x-\mu_1)^2}{2\sigma_1^2}\right]}{2\pi\sigma_1\sigma_2\sqrt{1-\rho^2}} \cdot \sqrt{2\pi}\sigma_2\sqrt{1-\rho^2} \\ &= \frac{1}{\sqrt{2\pi}\sigma_1}\exp\left[-\frac{(x-\mu_1)^2}{2\sigma_1^2}\right]. \end{aligned}$$

因此 $X \sim \mathrm{N}(\mu_1, \sigma_1^2)$；同理 $Y \sim \mathrm{N}(\mu_2, \sigma_2^2)$。

高斯积分：经典的二重积分法

记 $I=\int_{-\infty}^{+\infty}\mathrm{e}^{-x^2}\,\mathrm{d}x$。当 $|x|\geq 1$ 时，$\mathrm{e}^{-x^2}\leq \mathrm{e}^{-|x|}$，故 $0<I<\infty$。

第一步：将积分平方，化为二重积分

$$I^2=\left(\int_{\mathbb{R}}\mathrm{e}^{-x^2}\,\mathrm{d}x\right)\left(\int_{\mathbb{R}}\mathrm{e}^{-y^2}\,\mathrm{d}y\right)=\iint_{\mathbb{R}^2}\mathrm{e}^{-(x^2+y^2)}\,\mathrm{d}x\,\mathrm{d}y.$$

第二步：改用极坐标计算

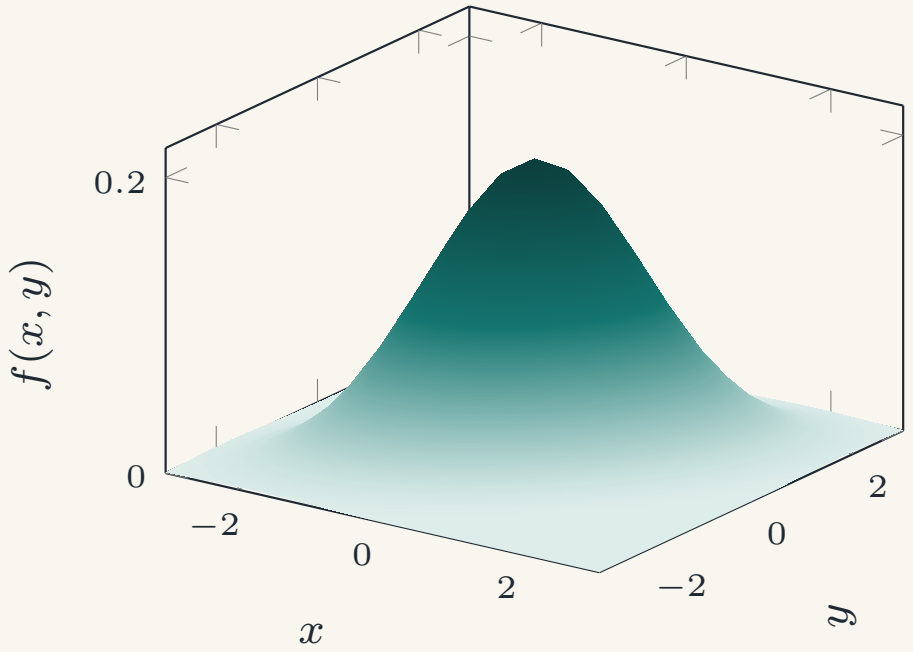
令 $x=r\cos\theta$，$y=r\sin\theta$，则 $\mathrm{d}x\,\mathrm{d}y=r\,\mathrm{d}r\,\mathrm{d}\theta$。

$$I^2=\int_0^{2\pi}\int_0^{+\infty}\mathrm{e}^{-r^2}r\,\mathrm{d}r\,\mathrm{d}\theta=2\pi\left[-\frac{1}{2}\mathrm{e}^{-r^2}\right]_0^{+\infty}=\pi.$$

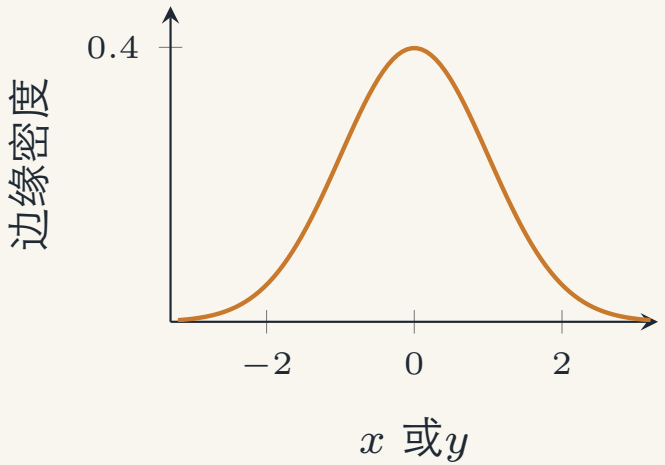
第三步：由于 $I>0$，取正的平方根

$$\boxed{\int_{-\infty}^{+\infty}\mathrm{e}^{-x^2}\,\mathrm{d}x=\sqrt{\pi}.}$$

正态联合密度与边缘密度：积分的图像



示意：$\mu_1 = \mu_2 = 0$，$\sigma_1 = \sigma_2 = 1$，$\rho = 0.6$。



$$f_X(x) = f_Y(x) = \frac{\mathrm{e}^{-x^2/2}}{\sqrt{2\pi}}.$$

边缘密度由联合密度沿一个方向积分得到，通常不等于它的某个截面。

思考：边缘正态能推出联合正态吗?

问题

若 X、Y 都服从一维正态分布，(X,Y) 一定服从二维正态分布吗?

答案是不一定。记标准正态密度为 $\varphi(x)=\mathrm{e}^{-x^2/2}/\sqrt{2\pi}$，考虑

$$f(x,y)=\varphi(x)\varphi(y)(1+\sin x\sin y),\qquad (x,y)\in\mathbb{R}^2.$$

因为 $1+\sin x\sin y\geq 0$，且 φ 为偶函数、$\sin$ 为奇函数，

$$\int_{-\infty}^{+\infty}\varphi(y)\sin y\,\mathrm{d}y=0.$$

于是

$$f_X(x)=\varphi(x),\qquad f_Y(y)=\varphi(y).$$

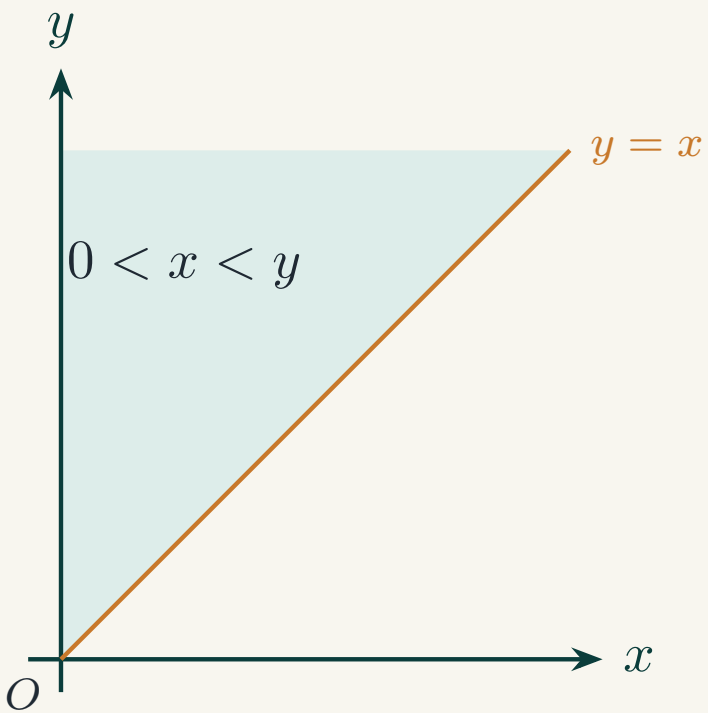
两个边缘密度都归一化，联合密度也归一化。因此 $X,Y\sim\mathrm{N}(0,1)$。

课堂练习：先画区域，再求边缘密度

设 (X, Y) 的联合密度为

$$f(x,y) = \begin{cases} \mathrm{e}^{-y}, & x > 0,\ y > x, \\ 0, & \text{其他.} \end{cases}$$

求 $f_X(x)$ 与 $f_Y(y)$。



确定积分范围

- 固定 $x > 0$，y 从 x 到 $+\infty$。
- 固定 $y > 0$，x 从 0 到 y。
- 非零区域无界，但积分仍可能有限。

课堂练习解答：指数密度与伽马密度

固定 $x > 0$，沿竖直方向积分：

$$f_X(x) = \int_x^{+\infty} \mathrm{e}^{-y}\,\mathrm{d}y = \mathrm{e}^{-x}.$$

固定 $y > 0$，沿水平方向积分：

$$f_Y(y) = \int_0^{y} \mathrm{e}^{-y}\,\mathrm{d}x = y\mathrm{e}^{-y}.$$

两个边缘密度

$$f_X(x) = \begin{cases} \mathrm{e}^{-x}, & x > 0, \\ 0, & \text{其他}, \end{cases} \qquad f_Y(y) = \begin{cases} y\mathrm{e}^{-y}, & y > 0, \\ 0, & \text{其他}. \end{cases}$$

检查：$\int_0^{+\infty} \mathrm{e}^{-x}\,\mathrm{d}x = 1$，$\int_0^{+\infty} y\mathrm{e}^{-y}\,\mathrm{d}y = 1$。

小结：求和与积分，消去另一个变量

分布函数：任何二维分布都适用

$$F_X(x) = F(x, +\infty), \qquad F_Y(y) = F(+\infty, y).$$

离散型：求和

$$p_{i\cdot} = \sum_j p_{ij}, \qquad p_{\cdot j} = \sum_i p_{ij}.$$

在分布律表中，沿列、沿行求和。

连续型：积分

$$f_X(x) = \int_{\mathbb{R}} f(x, y)\,\mathrm{d}y,$$

$$f_Y(y) = \int_{\mathbb{R}} f(x, y)\,\mathrm{d}x.$$

先画非零区域，再确定积分限。

联合分布与边缘分布关系

联合分布确定边缘分布，边缘分布通常不能确定联合分布；边缘正态也不能推出联合正态。

课后练习

原课件配套教材

第五版课后习题 86 页：第 1、2、3、8、9 题;
第六版课后习题 104 页：第 1、2、3、8、9 题;

完成时检查三个环节

1. 联合分布律的概率总和是否为 1，或联合密度是否归一化?
2. 求边缘分布时，是否对另一个变量的全部可能取值求和或积分?
3. 得到的两个边缘分布是否分别归一化?

3

条件分布

第三章　多维随机变量及其分布

条件分布：知道一部分信息以后，随机变量的分布

条件概率

若 $\mathbb{P}(B) > 0$，则事件 B 发生时，事件 A 的条件概率为

$$\mathbb{P}(A \mid B) = \frac{\mathbb{P}(A \cap B)}{\mathbb{P}(B)}.$$

把事件换成随机变量的取值信息

给定 Y 取某个值或落在某个集合中的信息，研究 X 的取值规律，得到 X 的**条件分布**。

- 离散型：给定 $Y = y_j$，求 X 的条件分布律。
- 连续型：给定 $Y = y$，用条件密度描述 X 的分布。

要点：分清哪个变量的值已给定以及我们要研究哪个变量？

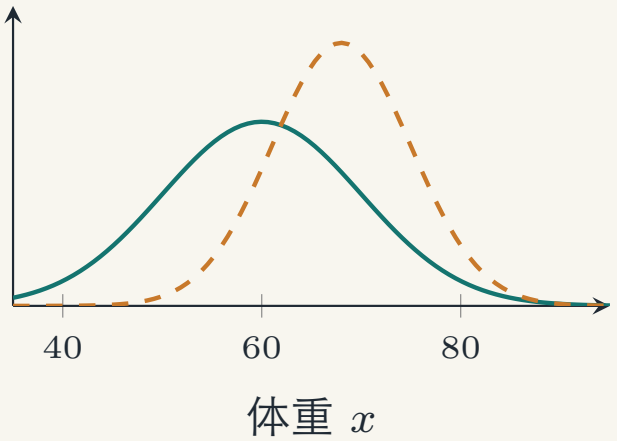
引例：身高范围内的体重分布

从某大学学生中随机抽取一人，令 X 为体重（千克），Y 为身高（米）。现在只考察身高在 1.7 米与 1.8 米之间的学生，并假设这一条件的概率为正。

$$\mathbb{P}(X \le x \mid 1.7 < Y < 1.8) = \frac{\mathbb{P}(X \le x,\ 1.7 < Y < 1.8)}{\mathbb{P}(1.7 < Y < 1.8)}.$$

条件改变了研究群体

先筛选身高，再研究被选中学生的体重。条件分布与全体学生的体重分布可能不同。



曲线仅示意一种可能的变化；体重怎样变化，需要由数据或模型确定。

离散型随机变量的条件分布律

设 (X, Y) 的联合分布律为 $p_{ij} = \mathbb{P}(X = x_i, Y = y_j)$，边缘分布律为 $p_{i\cdot}$ 与 $p_{\cdot j}$。

给定 $Y = y_j$，研究 X

对于固定的 j，若 $p_{\cdot j} > 0$，则

$$\mathbb{P}(X = x_i \mid Y = y_j) = \frac{p_{ij}}{p_{\cdot j}}, \qquad i = 1, 2, \ldots.$$

给定 $X = x_i$，研究 Y

对于固定的 i，若 $p_{i\cdot} > 0$，则

$$\mathbb{P}(Y = y_j \mid X = x_i) = \frac{p_{ij}}{p_{i\cdot}}, \qquad j = 1, 2, \ldots.$$

条件分布律 = 对应的联合概率 ÷ 条件事件的边缘概率。

条件分布律仍然是一个分布律

固定满足 $p_{\cdot j} > 0$ 的 j，记 $q_{i|j} = \mathbb{P}(X = x_i \mid Y = y_j)$。

两个基本性质

$$q_{i|j} \geq 0, \qquad \sum_i q_{i|j} = \frac{\sum_i p_{ij}}{p_{\cdot j}} = 1.$$

恢复联合分布与边缘分布

$$p_{ij} = q_{i|j} p_{\cdot j}, \qquad p_{i\cdot} = \sum_{j : p_{\cdot j} > 0} q_{i|j} p_{\cdot j}.$$

- 条件分布描述某一给定条件下的取值规律。
- 把各个条件下的分布按条件概率加权，可得到边缘分布。

例 1：从联合表看条件分布

令 $X=1$ 表示患肺癌，$X=0$ 表示未患肺癌；$Y=1$ 表示吸烟，$Y=0$ 表示不吸烟。假设联合分布如下：

	$X=0$	$X=1$	$\mathbb{P}(Y=j)$
$Y=0$	0.78	0.02	0.80
$Y=1$	0.15	0.05	0.20
$\mathbb{P}(X=i)$	0.93	0.07	1

问题

分别求吸烟与不吸烟条件下，X 的条件分布律。如何比较两组中的患病概率？

这里的数据是教学示例；条件概率反映关联，单凭这张表不能作因果推断。

例 1：按行归一化

吸烟：$Y = 1$

$$\mathbb{P}(X = 1 \mid Y = 1) = \frac{0.05}{0.20} = 0.25.$$

$$\mathbb{P}(X = 0 \mid Y = 1) = \frac{0.15}{0.20} = 0.75.$$

不吸烟：$Y = 0$

$$\mathbb{P}(X = 1 \mid Y = 0) = \frac{0.02}{0.80} = 0.025.$$

$$\mathbb{P}(X = 0 \mid Y = 0) = \frac{0.78}{0.80} = 0.975.$$

比较

在这个示例中，两个条件下的患病概率不同；每组的两个条件概率之和均为 1。全体患病概率可由加权求和恢复：

$$\mathbb{P}(X = 1) = 0.25 \times 0.20 + 0.025 \times 0.80 = 0.07.$$

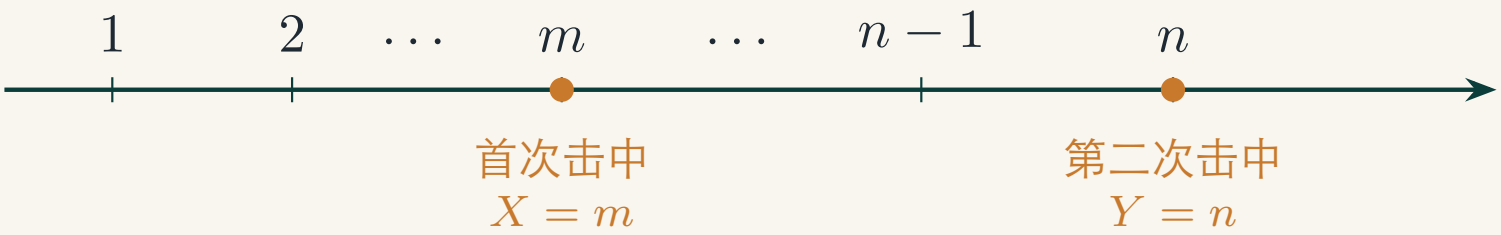
例 2：直到第二次击中为止

某射手连续射击，各次射击**相互独立**，每次击中目标的概率为 p，其中 $0 < p < 1$。射击到第二次击中目标为止，令

$$X = \text{首次击中时的射击次数}, \qquad Y = \text{第二次击中时的射击次数}.$$

计算分布律

求 (X, Y) 的联合分布律、边缘分布律，以及两个方向的条件分布律。



事件 $\{X = m, Y = n\}$ 要求 $1 \le m < n$，除了第 m 次，第 n 次其余射击均未击中。

例 2：联合分布律

记 $q = 1 - p$。对 $1 \le m < n$，事件 $\{X = m, Y = n\}$ 对应下面四段：

$1, \dots, m-1$ 全部未中 概率 q^{m-1}	第 m 次 击中 概率 p	$m+1, \dots, n-1$ 全部未中 概率 q^{n-m-1}	第 n 次 击中 概率 p

由各次射击的独立性，

$$\mathbb{P}(X = m, Y = n) = q^{m-1} p \, q^{n-m-1} p = p^2 q^{n-2}.$$

写清所有取值

$$\mathbb{P}(X = m, Y = n) = \begin{cases} p^2 q^{n-2}, & n = 2, 3, \dots, \quad m = 1, \dots, n-1, \\ 0, & \text{其他情形.} \end{cases}$$

固定 n 后，联合概率与首次击中的位置 m 无关。

例 2：边缘分布律

首次击中次数 X

对 $m = 1, 2, \ldots$，沿所有可能的第二次击中时刻求和：

$$\mathbb{P}(X = m) = \sum_{n=m+1}^{\infty} p^2 q^{n-2} = \frac{p^2 q^{m-1}}{1-q} = pq^{m-1}.$$

第二次击中次数 Y

对 $n = 2, 3, \ldots$，首次击中可能发生在 $1, \ldots, n-1$：

$$\mathbb{P}(Y = n) = \sum_{m=1}^{n-1} p^2 q^{n-2} = (n-1)p^2 q^{n-2}.$$

X 服从几何分布；Y 是获得第二次成功所需的试验次数。

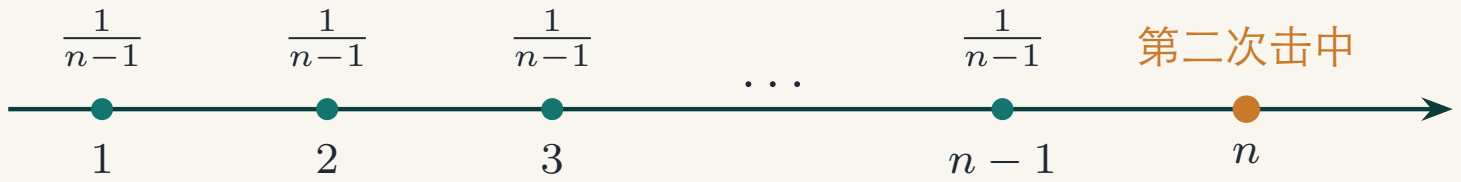
例 2：给定总次数，首次击中在哪里?

固定 $n \geq 2$，由联合分布律与 Y 的边缘分布律，

$$\mathbb{P}(X = m \mid Y = n) = \frac{p^2 q^{n-2}}{(n-1)p^2 q^{n-2}} = \frac{1}{n-1}, \qquad m = 1, \ldots, n-1.$$

结论

给定 $Y = n$，X 在 $\{1, 2, \ldots, n-1\}$ 上服从均匀分布。



条件已确定第 n 次是第二次击中；此前唯一一次击中，在各个位置等可能。

例 2：给定首次击中，还要等多久？

固定 $m \geq 1$，对 $n = m+1, m+2, \ldots$，

$$\mathbb{P}(Y = n \mid X = m) = \frac{p^2 q^{n-2}}{p q^{m-1}} = p q^{n-m-1}.$$

令 $K = Y - m$，则在 $X = m$ 的条件下，

$$\mathbb{P}(K = k \mid X = m) = p q^{k-1}, \qquad k = 1, 2, \ldots.$$

解释

第一次击中以后，为等到下一次击中所需的射击次数仍服从参数为 p 的几何分布；其分布不依赖首次击中的时刻 m。

同一个联合分布，两个方向的条件分布可以很不一样。