多维随机变量及其分布

第 6 讲 · 概率论与数理统计

复旦大学经济学院　ECON130001

上一讲我们做了什么

概率密度与分布函数

连续型用 f 描述，$F(x)=\int_{-\infty}^{x} f$；三个分布：均匀、指数、正态。 对正态随机变量作标准化 $Z=\frac{X-\mu}{\sigma}$，即可使用同一张标准正态分布表。

也学会了求函数的分布

离散型按函数值合并概率；连续型**先求 G 再求导**。反函数法须乘以反函数导数的绝对值。

★ 留下的问题

到现在为止，我们一次只研究**一个**随机变量。 但现实中重要的往往是**几个量之间的关系**。

例如，分别知道两只股票的收益率分布，还不能确定**它们同时出现大幅亏损的概率**。这种依赖关系应如何描述?

为什么需要研究多个随机变量

背景　2006 年，Netflix 设立**100 万美元**奖金，要求参赛系统将已有推荐系统 Cinematch 的评分预测误差降低**10%**，并公开约**1 亿条**电影评分。

转折　只比较两部电影各自的评分分布，我们可能认为它们吸引相似的观众。但即使两部电影的高分比例相同，**给它们高分的也可能是不同的人**。总体评分不能直接回答某位用户是否会喜欢一部电影。

问题　**随机抽取一位同时评价过两部电影的用户，记其对甲片、乙片的评分为 X, Y。怎样描述同一用户的两项评分之间的关系？知道甲片评分后，如何判断乙片获得高分的概率？**

来源与说明

竞赛背景据 Netflix 作者的原始论文与技术博客。两部电影评分的例子是本讲的简化教学模型。Cinematch 本身已使用评分之间的关系作个性化预测，不能将其概括为「只使用边缘分布」。 Bennett & Lanning, *The Netflix Prize* (2007)； Netflix 技术博客 (2012)。核对日期：2026-10-25。

本讲学习目标

1. 掌握**联合分布**（多个变量共同取值的概率）、 **边缘分布**（单个变量的分布）和**条件分布**（给定部分变量取值后的分布）
2. 会判断两个随机变量是否**独立**，并使用独立性计算联合概率
3. 会求两个随机变量的**线性组合**（和、差）的分布

收入与受教育年限、个股收益与市场收益等问题，需要同时研究多个变量。**本讲重点讨论二元随机变量及其分布。**

二元随机变量

定义

设 X, Y 定义在同一个样本空间上，则有序数对 (X, Y) 称为**二元随机变量**。其**联合分布**描述 (X, Y) 落入平面各区域 C 的概率 $P[(X, Y) \in C]$。

例

相互独立地掷一枚均匀骰子、抛一枚公平硬币。X 表示点数，$Y = 1$ 表示正面、$Y = 0$ 表示反面。

(X, Y) 的取值集是 $\{(1,1), \ldots, (6,1), (1,0), \ldots, (6,0)\}$ 共 12 个， 每个概率 $\frac{1}{12}$。

联合分布将 (X, Y) 作为**一个整体**，描述两个变量同时满足某些条件的概率，例如 $P(X \geq 4, Y = 1)$。

离散型的联合概率函数

定义

若 X, Y 都是离散型，其**联合概率函数**为

$$f(x,y) = P(\{X = x\} \cap \{Y = y\})$$

满足 $f(x,y) \geq 0$，且 $\sum_{\text{所有}(x,y)} f(x,y) = 1$。

和第 4 讲对比

一维概率函数为 $f(x) = P(X = x)$；二维概率函数为 $f(x,y) = P(X = x, Y = y)$，可用**联合概率表**表示。

归一化要求**整张表中所有取值组合的概率之和为 1**，而非每一行或每一列的和均为 1。

例：一个小镇的汽车与电视机

	y			
x	1	2	3	4
1	0.1	0	0.1	0
2	0.3	0	0.1	0.2
3	0	0.2	0	0

$f(x,y)$：X 汽车数、Y 电视机数

题

随机抽一户，求 (1) 至少两辆车**且**至少两台电视的概率；(2) 恰有一辆车的概率。

解：对满足事件条件的单元格求和

(1) 对 $x \geq 2$ 且 $y \geq 2$ 的六个单元格求和：

$$P(X \geq 2, Y \geq 2) = 0.5$$

(2) 对 $x = 1$ 所在**整行的概率**求和：

$$P(X = 1) = 0.2$$

第 (2) 问通过**对 Y 的所有可能取值求和**，得到只涉及 X 的概率。第 15–16 页将介绍对应的边缘分布。

交互：联合分布 • 边缘 • 条件

用联合概率表描述两部电影的评分

若 X, Y 分别表示同一用户对甲片、乙片的评分， 则 $f(x, y)$ 表示两项评分同时为 x, y 的概率。 **$f(5,5)$ 表示两部电影均获5星的概率**；已知甲片获5星时乙片获5星的概率，要在第 29–31 页用条件分布计算。

观察联合概率与行、列求和

1. 表格最右列 $f_1(x)$ 等于**对应行的概率之和**
2. 表格全部单元格的概率之和为 1

点击一行可高亮该行；按住 Shift 点击可选择一列。 本页重点观察行、列求和与右上方的边缘分布图。



交互演示：点行/列查看边缘分布与条件分布

https://fin-tech.fun/courseware/probability-statistics-undergrad/toolkit/joint_dist.html

连续型的联合密度函数

定义

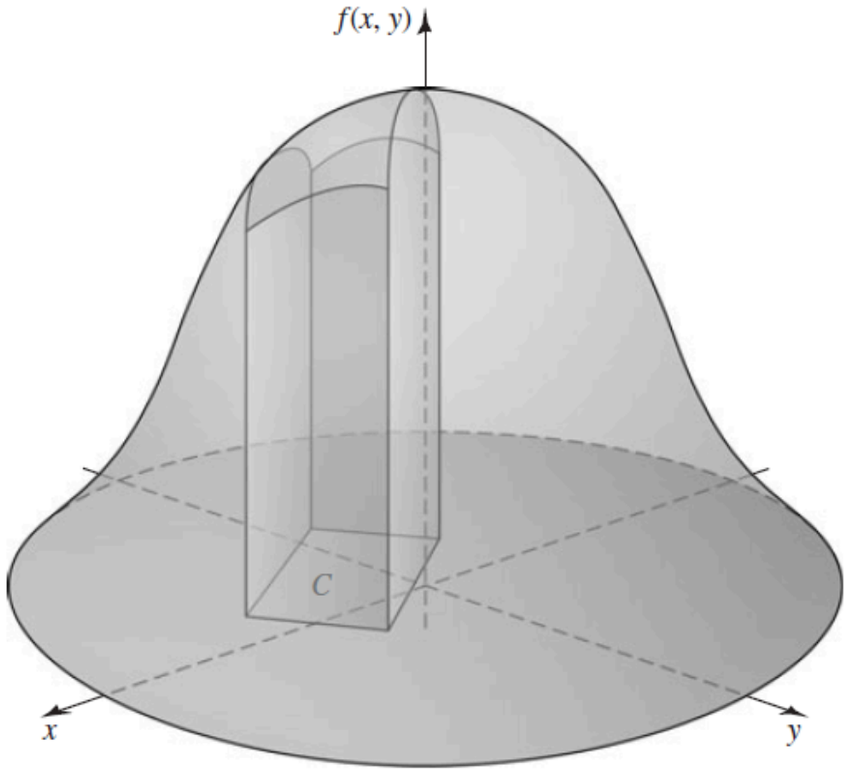
若存在非负函数 f，使 (X, Y) 落入平面区域 C 的概率可写为

$$P\{(X, Y) \in C\} = \iint_C f(x, y)\,\mathrm{d}x\,\mathrm{d}y$$

则称 (X, Y) 具有连续型联合分布，f 为**联合概率密度函数**。

非负性与归一化

$$f(x, y) \geq 0, \qquad \iint_{\mathbb{R}^2} f(x, y)\,\mathrm{d}x\,\mathrm{d}y = 1$$



联合密度可用曲面表示

区域 C 的概率等于**该区域上方、密度曲面下方的体积**。只有密度在 C 上为常数时，概率才等于「密度 × 面积」。

由联合密度的归一化确定常数

例

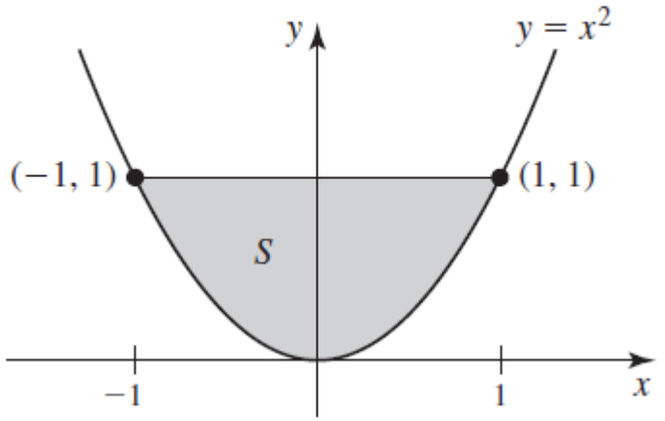
$f(x,y)=cx^2y$（$x^2\le y\le 1$），其余为 0。求 c。

解：由取值区域确定积分限

区域是抛物线 $y=x^2$ 与直线 $y=1$ 围成的，x 从 -1 到 1。**先对 y 积，y 从 x^2 到 1；再对 x 积：**

$$\int_{-1}^{1}\int_{x^2}^{1} cx^2y\,\mathrm{d}y\,\mathrm{d}x=\frac{4}{21}c=1$$

故 $c=\dfrac{21}{4}$。



$x^2\le y\le 1$

先画出**密度非零的区域**，再根据积分次序确定内、外层积分限。积分限应覆盖该区域且不重复。

在区域上求概率

例（续第 10 页）

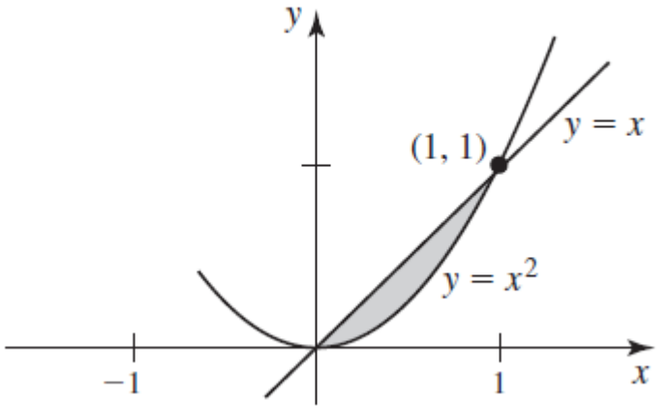
$f(x,y)=\frac{21}{4}x^2y$ $(x^2\le y\le 1)$ 。求 $P(X\ge Y)$。

解

$X\ge Y$ 即 $y\le x$。与原区域 $x^2\le y$ 求交： 需 $x^2\le y\le x$，这要求 $0\le x\le 1$。

$$\iint_{S_0} f=\int_0^1\int_{x^2}^{x}\frac{21}{4}x^2y\,\mathrm{d}y\,\mathrm{d}x=\frac{3}{20}$$

积分区域是**密度非零区域与事件对应区域的交集**。本题应同时满足 $x^2\le y\le 1$ 与 $y\le x$。



$x^2\le y\le x$

联合分布函数

定义

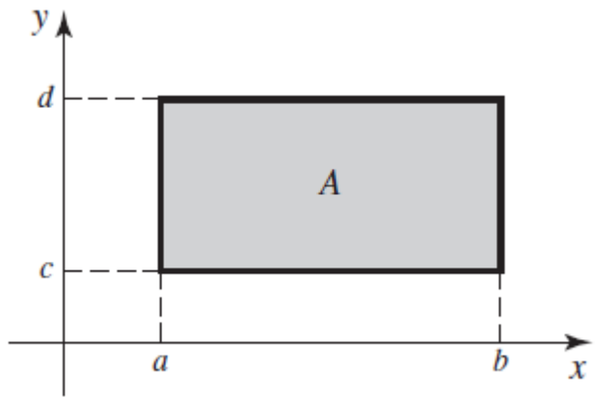
$$F(x,y)=P(X\le x,\ Y\le y)$$

X 的分布函数由 $F_1(x)=\lim\limits_{y\to+\infty}F(x,y)$ 得到，Y 的由 $F_2(y)=\lim\limits_{x\to+\infty}F(x,y)$ 得到。

矩形区域概率的计算

$$P(a<X\le b,\ c<Y\le d)$$
$$=F(b,d)-F(a,d)-F(b,c)+F(a,c)$$

从 $F(b,d)$ 对应的区域中减去左侧与下侧区域后，交叠部分被重复扣除，因此加回 $F(a,c)$。这是容斥原理的应用。



四项概率公式对应的区域

联合分布函数与联合密度的互换

若 X, Y 连续型且有联合密度 f，则

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

在 f 连续的点，可由**二阶混合偏导**求密度：

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

例

已知 X, Y 均在 $[0,2]$ 取值，且在 $0 \le x, y \le 2$ 时，$F(x,y) = \frac{1}{16}xy(x+y)$。求 $F_1(x)$ 与 $f(x,y)$。

F_1：固定 x 时，$y \ge 2$ 后 $F(x,y)$ 不再随 y 变化，因此 $F_1(x) = F(x,2)$：

$$F_1(x) = \tfrac{1}{8}x(x+2), \quad 0 \le x \le 2$$

f：$F = \frac{1}{16}(x^2y + xy^2)$，$\partial_x F = \frac{1}{16}(2xy + y^2)$，再对 y 求导：

$$f(x,y) = \tfrac{1}{8}(x+y), \quad 0 \le x, y \le 2$$

$x < 0$ 时 $F_1(x) = 0$，$x > 2$ 时 $F_1(x) = 1$；方形区域外的联合密度为0。

小结：联合分布

- **离散型**用联合概率函数 $f(x,y)=P(X=x,Y=y)$，所有取值组合的概率之和为 1
- **连续型**用联合密度 $f(x,y)$，二重积分等于 1；概率是曲面下的体积
- **联合分布函数** $F(x,y)=P(X\le x,Y\le y)$ 对两类通用； 有联合密度时，由二重积分求 F，在密度连续点由混合偏导求 f
- 计算区域概率时，先将**密度非零区域与事件对应区域求交**，再确定积分限

边缘分布：单个变量的分布

定义

二元随机变量 (X, Y) 中，单个变量 X 或 Y 的分布称为**边缘分布**。

记 X 的边缘分布函数为 $F_1(x)$、边缘概率（密度）函数为 $f_1(x)$；Y 的记为 $F_2(y)$、$f_2(y)$。

名字的来历

联合概率表通常将各行的概率之和写在**右侧**、各列的概率之和写在**下方**。这些位于表格边缘的概率给出各变量的边缘分布。

离散型的边缘分布

	y				
x	1	2	3	4	Total
1	0.1	0	0.1	0	0.2
2	0.3	0	0.1	0.2	0.6
3	0	0.2	0	0	0.2
Total	0.4	0.2	0.2	0.2	1.0

右侧为行和，下方为列和

公式

X 的边缘概率函数

$$f_1(x) = \sum_{\text{所有 } y} f(x, y)$$

Y 的边缘概率函数

$$f_2(y) = \sum_{\text{所有 } x} f(x, y)$$

求边缘概率的方法

固定一个变量，对另一个变量的**所有可能取值求和**。第 7 页第 (2) 问采用的就是这一方法。

边缘分布的应用：总体评分分布

核心思想

沿用第 8 页的评分变量 X, Y。对乙片评分 Y 求和，可得甲片评分 X 的**总体分布**：

$$f_1(x) = \sum_{y=1}^{5} f(x, y) \quad (\text{对 } Y \text{ 的所有评分求和})$$

推荐系统里的用途

- 对联合分布求和，得到某部电影的边缘分布 $f_1(x)$
- 它描述所选用户群体对甲片的**总体评分状况**
- 例如，可据此比较各电影获得4星或5星的概率，形成**统一的推荐依据**

这个分布没有用到该用户对乙片的评分，所以**推荐不会根据他的已知评分做调整**。要做个性化判断，还需要知道两项评分之间的关系。

连续型的边缘分布

公式：求和换成积分

$$f_1(x)=\int_{-\infty}^{\infty} f(x,y)\,\mathrm{d}y,\qquad f_2(y)=\int_{-\infty}^{\infty} f(x,y)\,\mathrm{d}x$$

边缘分布函数则由极限得到：$F_1(x)=\lim\limits_{y\to\infty}F(x,y)$，$F_2(y)=\lim\limits_{x\to\infty}F(x,y)$。

例（续第 10–11 页）

$f(x,y)=\frac{21}{4}x^2y$（$x^2\le y\le 1$）。求两个边缘密度。

$$f_1(x)=\int_{x^2}^{1}\frac{21}{4}x^2y\,\mathrm{d}y=\frac{21}{8}x^2(1-x^4),\quad -1\le x\le 1$$

$$f_2(y)=\int_{-\sqrt{y}}^{\sqrt{y}}\frac{21}{4}x^2y\,\mathrm{d}x=\frac{7}{2}y^{5/2},\quad 0\le y\le 1$$

求 f_1 时固定 x，对 $y\in[x^2,1]$ 积分；求 f_2 时固定 y，对 $x\in[-\sqrt{y},\sqrt{y}]$ 积分。**积分限由同一区域的不同截面确定。**上述区间外，边缘密度为0。

★ 边缘分布不能唯一确定联合分布

例

两枚公平硬币各掷十次，各次投掷相互独立。考虑两种情形：

1. X, Y 分别是**两枚**硬币的正面次数
2. X, Y **都是第一枚**硬币的正面次数

求 X, Y 的联合分布。

解：边缘分布相同，联合分布不同

两种情形下 X 与 Y 的**边缘分布都是** $B(10, 0.5)$。但

$$(1)\ p_{ij} = C_{10}^{i} C_{10}^{j}\, 0.5^{20}; \qquad (2)\ p_{ij} = C_{10}^{i}\, 0.5^{10}\ (i = j),\ \ 0\ (i \neq j)$$

边缘分布未确定变量间的依赖关系

情形 (2) 中 $Y = X$；情形 (1) 中两组计数相互独立。**相同的边缘分布可以对应不同的联合分布。**

交互：边缘一样，联合不同

怎么用

依次切换三个联合分布，比较联合概率表与**两个边缘分布的柱状图**。

切换时要注意的

三种情形的非零概率分别分布在**整个表格、主对角线和反对角线**， 而边缘分布**完全不变**。

三种情形分别对应：两组独立计数、$Y = X$、$Y = 3 - X$。



交互演示：三个联合分布，边缘完全相同

https://fin-tech.fun/courseware/probability-statistics-undergrad/toolkit/marginal_trap.html

从骨骼部件到完整骨架：还缺少什么？

回到硬币例题

两种情形中，每个变量各自的分布相同，但**两项计数之间的关系不同**，因此联合分布不同。

设想用骨骼化石重建一副骨架

知道完整骨架，就能了解各部分骨骼及其排列。反过来，即使各部分骨骼的资料齐全，若不知道**它们怎样连接**，仍不足以唯一确定整体结构。

这个类比强调：**各部分的信息，不等于各部分之间的关系信息。**

边缘分布描述各变量自身；要确定联合分布，还需要**变量之间的依赖关系**。

如果进一步知道变量**相互独立**，能否由边缘分布确定联合分布？

随机变量的独立性

定义

对任意区间 A, B，若事件 $\{X \in A\}$ 与 $\{Y \in B\}$ 独立，即

$$P(\{X \in A\} \cap \{Y \in B\}) = P(X \in A)\, P(Y \in B)$$

则称随机变量 X, Y **相互独立**。

独立性的等价判据

$$X, Y \text{ 独立} \iff F(x, y) = F_1(x) F_2(y) \quad \forall x, y$$

$$X, Y \text{ 独立} \iff f(x, y) = f_1(x) f_2(y)$$

离散型要求对所有取值成立；连续型可忽略零面积集合上的密度差异，因为这些差异不影响概率。

证明思路

定义取 $A = (-\infty, x]$、$B = (-\infty, y]$ 即得分布函数形式； 对连续型两边求二阶混合偏导得密度形式；离散型取单点集直接得到。

已知独立性时，可由两个边缘分布确定联合分布。若未给出独立性或其他依赖关系，仅凭边缘分布通常不能确定联合分布。

应用：独立性与评分信息

教学模型

设 X, Y 分别表示同一用户对甲片、乙片的评分。若两项评分**相互独立**，则

$$f(x, y) = f_1(x) f_2(y)$$

独立性的含义

在这个模型下，已知用户对甲片的评分，**不会改变对其乙片评分的概率判断**。仅凭甲片评分，不能提高对乙片偏好的预测能力。

应用时需要核实的前提

独立性是模型假设，应由数据与问题背景支持。**题材不同、分类不同，并不自动意味着评分独立。**

独立性例题（一）：由独立性确定联合密度

例

X, Y 相互独立，都服从 $g(x) = 2x$（$0 \le x \le 1$）。求 $P(X + Y \le 1)$。

独立 $\Rightarrow$ 联合是乘积：

$$f(x, y) = 4xy, \quad 0 \le x, y \le 1$$

$$P(X + Y \le 1) = \int_0^1 \int_0^{1-x} 4xy \, \mathrm{d}y \, \mathrm{d}x = \frac{1}{6}$$

例

$f(x, y) = ke^{-(x+2y)}$（$x, y \ge 0$）。X, Y 是否独立?

归一化给出 $k = 2$。两个边缘：

$$f_1(x) = e^{-x} \ (x \ge 0), \qquad f_2(y) = 2e^{-2y} \ (y \ge 0)$$

$f_1 f_2 = 2e^{-x-2y} = f$，**独立**。

本题的**非零区域为** $[0, \infty) \times [0, \infty)$，且密度等于两个边缘密度的乘积，故满足独立性判据。

独立性例题（二）：取值区域也需要检查

例

$f(x,y)=kx^2y^2$（$x^2+y^2\le 1$）。X,Y 是否独立？

区域内的表达式可分为 $kx^2\cdot y^2$，但还需考虑**圆盘约束**。

$X>0.8$ 与 $Y>0.8$ 各有正概率，但不能同时发生，因为 $0.8^2+0.8^2>1$。因此

$$P(X>0.8,Y>0.8)=0$$

小于 $P(X>0.8)P(Y>0.8)$，故**不独立**。

例

$f(x,y)=\frac{15}{4}x^2$（$0\le y\le 1-x^2$）。求边缘并判断独立性。

$$f_1(x)=\tfrac{15}{4}x^2(1-x^2),\quad -1\le x\le 1$$
$$f_2(y)=\tfrac{5}{2}(1-y)^{3/2},\quad 0\le y\le 1$$

其余位置边缘密度为 0。两个边缘密度的乘积与联合密度不一致，故**不独立**。

★ 同时检查表达式与取值区域

独立性要求**联合分布等于两个边缘分布的乘积**。只将区域内的密度表达式分解为两个因子，尚不能完成判定。

独立性例题（三）：离散与连续各一道

例（离散）

$f(x,y)=\frac{1}{30}(x+y)$，$x=0,1,2$，$y=0,1,2,3$。求边缘并判断独立性。

$$f_1(x)=\sum_{y=0}^{3}\frac{x+y}{30}=\frac{1}{15}(2x+3)$$

$$f_2(y)=\sum_{x=0}^{2}\frac{x+y}{30}=\frac{1}{10}(1+y)$$

$f_1f_2\neq f$，**不独立**。

例（连续）

$f(x,y)=\frac{3}{2}y^2$（$0\le x\le 2$，$0\le y\le 1$）。

$$f_1(x)=\int_0^1\frac{3}{2}y^2\mathrm{d}y=\frac{1}{2},\qquad f_2(y)=\int_0^2\frac{3}{2}y^2\mathrm{d}x=3y^2$$

$f_1f_2=\frac{3}{2}y^2=f$，**独立**。

进一步：事件 $\{X<1\}$ 与 $\{Y\ge\frac{1}{2}\}$ 也独立——**独立变量分别确定的事件 $\{X\in A\}$ 与 $\{Y\in B\}$ 相互独立**。

小结：边缘分布与独立性

- 求边缘分布：固定一个变量，对另一个变量的全部可能取值**求和或积分**
- 联合分布能确定边缘分布；仅有边缘分布通常不能确定**变量间的依赖关系**
- 独立性要求 $F(x,y) = F_1(x)F_2(y)$ 对所有 x, y 成立；概率函数或密度函数形式见**第 22 页**
- 用密度判断独立性时，应同时检查**表达式与非零区域**

条件分布：给定取值后的概率分布

	Brand Y					
Stolen X	1	2	3	4	5	Total
0	0.129	0.298	0.161	0.280	0.108	0.976
1	0.010	0.010	0.001	0.002	0.001	0.024
Total	0.139	0.308	0.162	0.282	0.109	1.000

$X = 1$ 表示失窃，Y 表示汽车品牌

问题

由这张表能算出 X 的边缘概率，也就是**总的失窃率**。 **已知汽车品牌后，失窃概率是多少?**

用第 3 讲的条件概率

$$P(X = 1 \mid Y = y) = \frac{P(X = 1, Y = y)}{P(Y = y)} = \frac{f(1, y)}{f_2(y)}$$

当 $P(Y = y) > 0$ 时，将品牌 y 所在列的各项概率除以该列总概率，即得条件概率。

离散型的条件分布

定义

设 X, Y 的联合概率函数为 f。对每个使 $f_2(y) > 0$ 的 y，定义

$$g_1(x \mid y) = \frac{f(x, y)}{f_2(y)}$$

称为给定 $Y = y$ 时 X 的**条件概率函数**。对称地，对 $f_1(x) > 0$ 的 x

$$g_2(y \mid x) = \frac{f(x, y)}{f_1(x)}$$

条件概率函数的归一化

$$\sum_x g_1(x \mid y) = \frac{\sum_x f(x, y)}{f_2(y)} = \frac{f_2(y)}{f_2(y)} = 1$$

用 $f_2(y)$ 归一化后，该列的条件概率之和为 1。 这对应第 3 讲给定事件后重新计算条件概率的方法。

交互：比较不同条件下的评分或计数分布

选择条件并观察条件分布

点击一行，右下方显示给定该 x 时的 $g_2(y \mid x)$，其**概率之和为 1**。

要做的对照

切到「两枚硬币（独立）」，点任意一行——**条件分布完全相同**，而且等于边缘分布。
再切到「同一枚硬币」，条件分布退化成一根柱子。

独立时，对所有满足 $f_1(x) > 0$ 的 x，都有$g_2(y \mid x) = f_2(y)$。



交互演示：点行/列查看条件分布的归一化

https://fin-tech.fun/courseware/probability-statistics-undergrad/toolkit/joint_dist.html

应用：根据已知评分计算高分概率

沿用第 8 页的评分变量

X, Y 分别表示同一用户对甲片、乙片的评分。已知甲片评分为5星，且 $f_1(5) > 0$，则

$$g_2(y \mid 5) = \frac{f(5, y)}{f_1(5)}$$

以4星或5星作为高分

$$P(Y \geq 4 \mid X = 5) = g_2(4 \mid 5) + g_2(5 \mid 5)$$

在这个教学模型中，可比较各候选电影在同一条件下获得高分的概率，选择概率较高的电影。

与总体评分分布的区别

边缘概率 $P(Y \geq 4)$ 说的是所选用户群体的总体情况；**条件概率则用上了「甲片获5星」这条已知信息**。两者是否相同，取决于两项评分的依赖关系。

连续型条件分布：从条件概率到条件密度

场景

加工分两步，第一步用时 Y，总用时 X，联合密度 $f(x,y)=e^{-x}$（$0\le y\le x<\infty$）。 给定第一步的用时 Y，求总用时 X 的条件分布。

困难

不能直接用两个单点概率的比值计算——因为 $P(Y=y)=0$，分母是 0。

条件密度的定义

对每个使 $f_2(y)>0$ 的 y，直接**定义**

$$g_1(x\mid y)=\frac{f(x,y)}{f_2(y)}$$

对 $f_2(y)=0$ 的点可任意定义，只要 g_1 仍是一个密度函数。

连续型使用**联合密度与边缘密度的比值**，而非两个单点概率的比值。对 x 积分可验证 $\int g_1(x\mid y)\,\mathrm{d}x=1$。

条件分布与联合分布的两个例子

例（续第 32 页）

$f(x,y)=e^{-x}$ $(0\le y\le x)$ 。求 $g_1(x\mid y)$。

先求边缘密度：

$$f_2(y)=\int_y^{\infty} e^{-x}\mathrm{d}x=e^{-y}$$

$$g_1(x\mid y)=\frac{e^{-x}}{e^{-y}}=e^{y-x},\quad x\ge y,\quad y\ge 0$$

给定第一步用时 $Y=y$，**剩余用时 $X-y$ 服从参数为1的指数分布**，且该分布不随 y 改变。$x<y$ 时条件密度为0。

例：由条件分布构造联合分布

将次品率建模为 $P\sim U[0,1]$。给定 $P=p$ 后，独立检验 n 件产品，记次品数为 X。求 (X,P) 的联合分布。

给定 $P=p$ 时 $X\sim B(n,p)$：

$$g_1(x\mid p)=C_n^x p^x(1-p)^{n-x}$$

由乘法法则 $f(x,p)=g_1(x\mid p)f_2(p)$，而 $f_2(p)=1$：

$$f(x,p)=C_n^x p^x(1-p)^{n-x}$$

这里 X 离散、P 连续：计算联合概率时，**对 $x=0,\dots,n$ 求和，对 $0\le p\le 1$ 积分**。

条件分布例题（一）

例：圆盘上的均匀分布

(X,Y) 在 $S=\{(x-1)^2+(y+2)^2\le 9\}$ 上均匀分布。 求 $g_2(y\mid x)$ 与 $P(Y>0\mid X=2)$。

$f=\frac{1}{9\pi}$。固定 x 后，对圆盘内的 y 积分得$f_1(x)=\frac{2}{9\pi}\sqrt{9-(x-1)^2}$，故

$$g_2(y\mid x)=\frac{1}{2}\left[9-(x-1)^2\right]^{-1/2}$$

其中 $-2<x<4$，$|y+2|\le\sqrt{9-(x-1)^2}$。这是该线段上的**均匀密度**，线段外为0。

$$P(Y>0\mid X=2)=\frac{2-\sqrt{2}}{4}\approx 0.146$$

例

$f(x,y)=c(x+y^2)$ $(0\le x,y\le 1)$ 。求 $g_1(x\mid y)$ 与 $P(X<\frac{1}{2}\mid Y=\frac{1}{2})$。

$f_2(y)=c(\frac{1}{2}+y^2)$，计算比值时 c **相消**：

$$g_1(x\mid y)=\frac{x+y^2}{\frac{1}{2}+y^2}$$

$$P(X<\tfrac{1}{2}\mid Y=\tfrac{1}{2})=\int_0^{1/2}\frac{x+\frac{1}{4}}{\frac{3}{4}}\mathrm{d}x=\frac{1}{3}$$

右题中联合密度与边缘密度含有相同的常数因子，计算条件密度时相消，因此可直接计算所需条件概率。

条件分布例题（二）

例

$f(x,y)=\frac{3}{16}(4-2x-y)$（$x,y>0$，$2x+y<4$）。求 $g_2(y\mid x)$ 与 $P(Y\geq 2\mid X=0.5)$。

$$f_1(x)=\int_0^{4-2x} f\,\mathrm{d}y=\tfrac{3}{8}(x-2)^2,\quad 0<x<2$$

$$g_2(y\mid x)=\frac{4-2x-y}{2(x-2)^2},\quad 0<y<4-2x$$

$$P(Y\geq 2\mid X=0.5)=\int_2^3 g_2(y\mid 0.5)\,\mathrm{d}y=\frac{1}{9}$$

例：泊松稀释

总产量 $Y\sim\mathrm{Poisson}(\lambda)$，每件独立地以概率 p 为次品。求次品数 X 的分布。

给定 $Y=y$ 时 $X\sim B(y,p)$。由乘法法则再对 y 求和：

$$P(X=x)=\sum_{y=x}^{\infty}C_y^x p^x(1-p)^{y-x}\frac{\lambda^y}{y!}e^{-\lambda}=\frac{(\lambda p)^x}{x!}e^{-\lambda p}$$

结果仍是泊松分布，参数变成 λp。 推导依次使用条件二项分布、全概率公式与求和。

两个随机变量的函数

联合分布描述共同取值，边缘分布描述单个变量，条件分布描述给定部分取值后的分布。

但还有一类问题没解决：两项投资的收益金额分别为 X 与 Y， 需要计算**总收益金额** $X+Y$。 它的分布怎么求?

离散型的线性组合

例

X_1 等可能取 $1,2,3,4$，X_2 等可能取 $1,2,3$，相互独立。求 $Y=X_1+X_2$ 的分布。

解：对具有相同和值的取值组合求和

Y 取 $2,\dots,7$，联合有 $4\times 3=12$ 种等可能组合：

y	2	3	4	5	6	7
P	1/12	2/12	3/12	3/12	2/12	1/12

比如 $Y=4$ 有 $(1,3),(2,2),(3,1)$ 三种。

一般地 $P(Y=y)=\sum_x f(x,\,y-x)$——**对直线 $x_1+x_2=y$ 上的联合概率求和**。X_1-X_2、$2X_1-3X_2$ 可类似处理，对相应直线上的取值组合求和。

独立二项变量与泊松变量之和

二项分布之和

若 $X_1, \ldots, X_k$ 相互独立，$X_i \sim B(n_i, p)$（p **必须相同**），则

$$X_1 + \cdots + X_k \sim B(n_1 + \cdots + n_k,\ p)$$

解释： $B(n, p)$ 就是 n 次独立伯努利试验的成功数。合并两组相互独立、成功概率同为 p 的试验，即有 $n_1 + n_2$ 次试验。

泊松分布之和

若 $X_1, \ldots, X_k$ 相互独立，$X_i \sim \text{Poisson}(\lambda_i)$，则

$$X_1 + \cdots + X_k \sim \text{Poisson}(\lambda_1 + \cdots + \lambda_k)$$

直观： 两个独立的泊松过程叠加，还是泊松过程，速率相加。 例如，两个独立泊松到达过程的速率分别为每小时10次与6次，合并后的速率为每小时16次。

p **相同这个条件不能省。**两个独立变量分别服从 $B(3, 0.5)$ 与 $B(3, 0.9)$ 时，**其和不服从二项分布**，因为两组试验的成功概率不同。

连续型随机变量之和与卷积

例：依次完成两项服务

服务时长 X_1, X_2 独立，均有密度 $f(x) = 2e^{-2x}$（$x > 0$）。求总时长 $Y = X_1 + X_2$ 的分布。

对 $y > 0$，先求分布函数：

$$\begin{aligned} G(y) &= P(X_1 + X_2 \le y) \\ &= \int_0^{y} \int_0^{y-x_2} 4e^{-2(x_1+x_2)} \, \mathrm{d}x_1 \mathrm{d}x_2 \\ &= 1 - (1+2y)e^{-2y} \end{aligned}$$

求导得 $g(y) = 4ye^{-2y}$。$y \le 0$ 时 $G(y) = g(y) = 0$。

和的密度与卷积公式

若 (X, Y) 有联合密度 f，令 $Z = X + Y$，则

$$f_Z(z) = \int_{-\infty}^{\infty} f(z-y, y) \, \mathrm{d}y$$

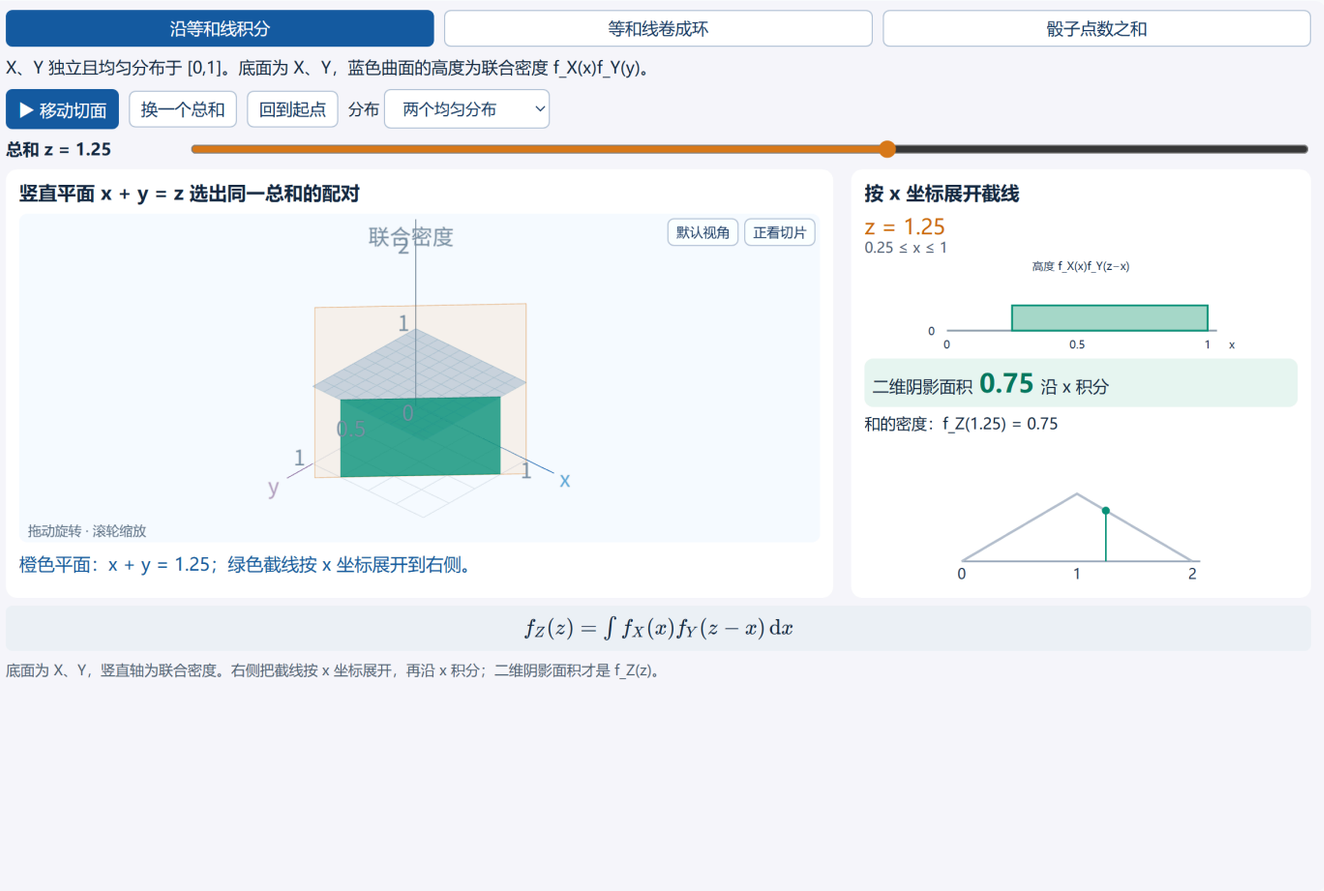
若 X, Y 进一步**相互独立**，则

$$f_Z(z) = \int_{-\infty}^{\infty} f_1(z-y) f_2(y) \, \mathrm{d}y$$

后一个积分称为两个密度函数的**卷积**。

第 37 页对满足 $x_1 + x_2 = y$ 的离散取值组合求和；本页用积分计算连续变量之和的分布。

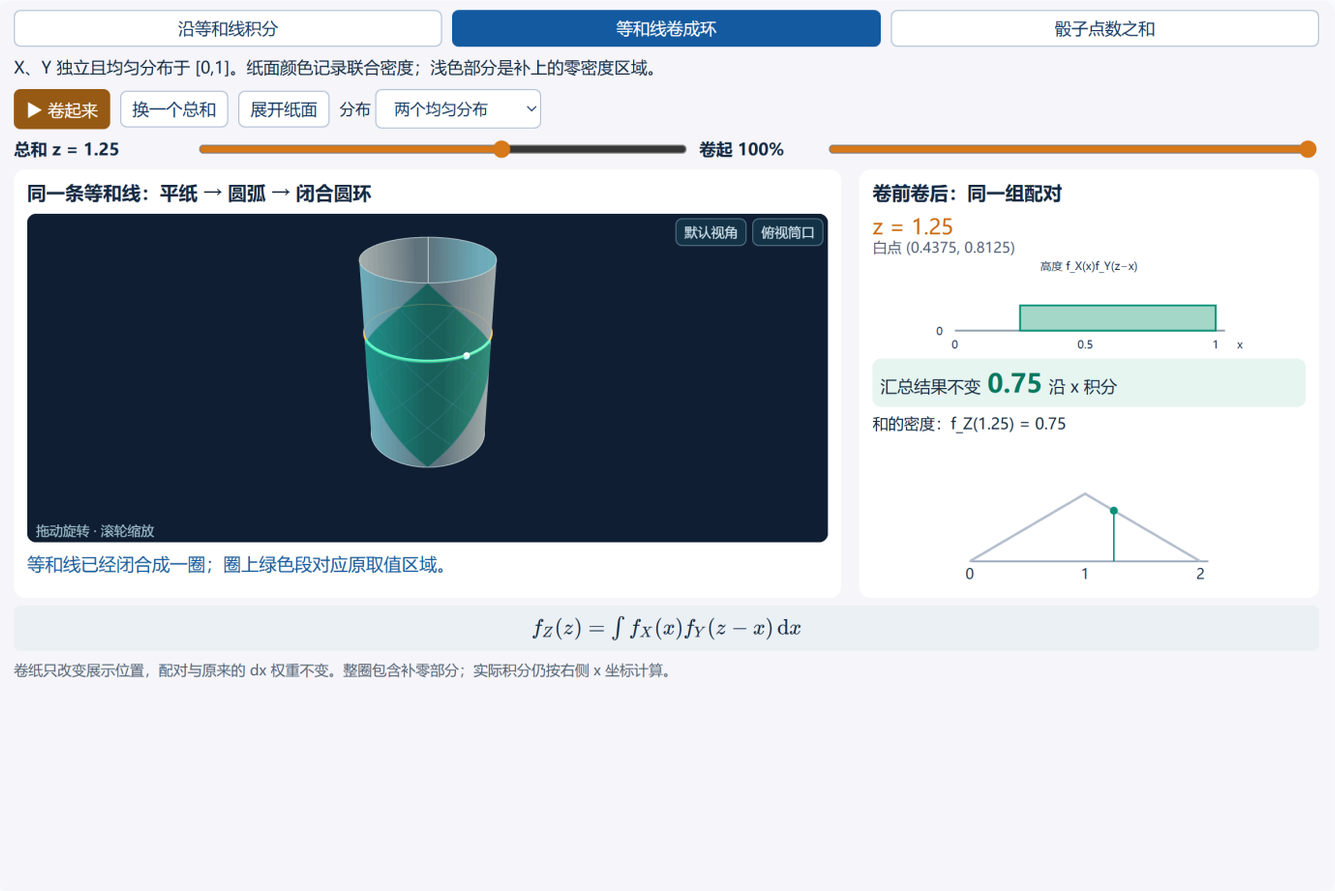
沿 X + Y = z 汇总：卷积的几何解释



底面为 X、Y；平面 x+y=z 选出配对，右侧按 x 坐标展开截线，二维阴影面积给出 f_Z(z)。

https://fin-tech.fun/courseware/probability-statistics-undergrad/toolkit/convolution.html

把等和线卷成一圈



卷纸只改变展示位置；同一圈对应相同总和，贡献仍按原来的 x 坐标汇总。[打开演示]

https://fin-tech.fun/courseware/probability-statistics-undergrad/toolkit/convolution.html

独立正态随机变量之和

定理

若 $X_1, \ldots, X_k$ 相互独立，$X_i \sim N(\mu_i, \sigma_i^2)$，则

$$X_1 + \cdots + X_k \sim N(\mu_1 + \cdots + \mu_k,\ \sigma_1^2 + \cdots + \sigma_k^2)$$

证明思路

对两个变量使用第 39 页的卷积公式。代入两个正态密度，配方后完成积分；多个变量的结论可由归纳得到。

参数的合并规则

新分布的两个参数分别是 $\sum_i \mu_i$ 与 $\sum_i \sigma_i^2$。对两个独立正态变量，和的参数满足：

$$\sigma_{X+Y} = \sqrt{\sigma_X^2 + \sigma_Y^2}$$

第 5 讲中 σ 表示曲线的宽窄；其标准差含义将在第 7 讲给出定义。

应用定理时必须核实**相互独立**这一前提。仅知道各变量分别服从正态分布，还不能保证它们的和服从正态分布。

例：两个均匀分布之和

例

投资者同时持股票与债券。年末股票价值 $X \sim U[1000, 4000]$， 债券价值 $Y \sim U[800, 1200]$，两者独立。求组合价值 $Z = X + Y$ 的分布。

解：由卷积求组合价值的密度

$$f_Z(z) = \begin{cases} \dfrac{z - 1800}{1\,200\,000}, & 1800 < z \le 2200 \\ \dfrac{1}{3000}, & 2200 < z \le 4800 \\ \dfrac{5200 - z}{1\,200\,000}, & 4800 < z \le 5200 \\ 0, & \text{其他} \end{cases}$$

密度函数的形状

本例两段取值区间长度不同，和的密度呈**梯形**；若两个独立均匀变量的区间等长，和的密度则呈三角形。

随机变量乘积 $S = XY$ 的分布

例

(X, Y) 在矩形 $G = \{0 \le x \le 2,\ 0 \le y \le 1\}$ 上均匀分布。 求以 X, Y 为边长的矩形**面积** $S = XY$ 的密度。

解：先求 F_S，再求密度

$P(XY \le s)$ 对应双曲线 $xy = s$ 下方的区域，积分后求导得

$$f(s) = \begin{cases} \dfrac{1}{2}(\ln 2 - \ln s), & 0 < s \le 2 \\ 0, & \text{其他} \end{cases}$$

当 $s \to 0^+$ 时，密度趋于无穷；但它**非负且积分为1**，仍满足概率密度的条件。

最大值与最小值的分布

定理

设 X, Y 独立，分布函数分别为 F_X, F_Y，则

$$F_{\max}(z) = F_X(z)F_Y(z)$$

$$F_{\min}(z) = 1 - \big[1 - F_X(z)\big]\big[1 - F_Y(z)\big]$$

证明思路

$\{\max \le z\} = \{X \le z\} \cap \{Y \le z\}$，由独立性概率相乘；$\{\min > z\} = \{X > z\} \cap \{Y > z\}$，同理相乘后取对立事件。

将最大值与最小值事件改写为交集

因为「最大值不超过 z」等价于「**每一个**都不超过 z」，「最小值超过 z」等价于「**每一个**都超过 z」——**先表示为事件的交集，再由独立性将概率写为乘积。**

第 5 讲第 17 页的 n 个灯泡例题（Y_1 服从 $\mathrm{Exp}(n\lambda)$），可由这里的 $F_{\min}$ 公式求得。

本讲总结

① 核心结论

- **联合分布**描述共同取值；**边缘分布**描述单个变量；**条件分布**描述给定部分变量取值后的分布
- 由联合分布求边缘分布：对另一个变量**求和或积分**；仅凭边缘分布通常不能确定联合分布
- **独立性：** $F(x,y)=F_1(x)F_2(y)$ 对所有 x,y 成立；概率函数或密度形式的适用条件见第 22 页
- **和的分布：** 离散型对相同和值的取值组合求和；连续型使用联合密度积分，独立时可用两个密度的卷积
- 独立二项变量之和的可加性要求成功概率相同；独立泊松变量、独立正态变量之和仍属于各自的分布族

本讲总结（续）

② 易错提醒

- 先画出密度非零区域，与事件对应区域**求交**，再确定**二重积分的积分限**
- 判断独立性应检查**完整的联合分布**，包括密度非零的区域
- 对**独立正态变量**求和：$\sigma_{X+Y}=\sqrt{\sigma_X^2+\sigma_Y^2}$，不能将 σ 直接相加
- 二项与泊松的可加性均要求**变量相互独立**；二项还要求 p 相同

③ 下一讲

本讲给出了独立正态变量求和时的参数合并规则。若变量不独立，如何描述**它们共同变化的程度**，并计算和的波动大小？

除了完整分布，也常需要用少数指标概括**平均水平、分散程度与变量间的关系**。

下一讲：随机变量的数字特征——期望、方差，以及描述共同变化的协方差。