

MAT220 第6讲：概率论的进一步讨论

Contents

1 涉及计数的问题	2
1.1 形式体系的简要回顾	2
1.2 例1: 选择衬衫和领带	2
1.3 例2: 排列水果	5
1.4 例3: 皇家同花顺	8
2 超越古典概率	9
2.1 古典概率中的隐藏假设	9
2.2 给简单事件分配概率	9
2.3 熟悉的公式仍然成立	10
2.4 例: 不公平的骰子	12
2.5 例: 不公平的硬币	12
3 离散概率分布	13
3.1 从概率权重到分布	13
3.2 随机变量	14
3.3 离散概率分布	15
3.4 分布的形状	16

在第5讲中, 我们介绍了概率的基本词汇: 简单事件、样本空间、事件、补事件、交集、并集、条件概率和独立性。我们还看到, 古典概率特别方便, 因为所有结果都是等可能的, 所以可以通过我们的黄金公式用计数来计算概率。今天, 我们会把这个理论付诸实践, 研究一些与古典概率和第4讲的计数技巧有关的问题。之后, 我们会开始超越古典概率, 进入一个结果不一定等可能的一般情形。这自然会引向概率分布, 并最终回到描述统计。

1 涉及计数的问题

今天我们会先花一些时间复习古典概率中的练习题，然后转向结果不一定等可能的情形。在此之前，我们先简要回顾上一讲的理论。

1.1 形式体系的简要回顾

在第5讲中，我们介绍了概率论的基本形式体系。值得回忆的几个关键概念如下：

1. 结果是指一次实验或操作可能发生的事情。
2. 所有结果的集合称为样本空间。形式上，这是一个包含所有结果 ω 的集合 S 。
3. 事件可以宽泛地理解为我们能对系统作出的某种陈述。形式上，事件被定义为样本空间 S 的子集 A 。
4. 简单事件是只包含一个结果的特殊事件，也就是说，对于某个结果 $\omega \in S$ ，有 $A = \{\omega\}$ 。
5. 一个事件的概率记为 $P(A)$ ，它是 0 到 1 之间的一个数，其中 0 表示不可能，1 表示确定。

在第5讲中，我们讨论了古典概率。古典概率额外假设所有结果都是等可能的，并且样本空间 S 总共只包含有限多个结果。如果样本空间的大小为 $|S| = n$ ，那么每个结果的概率都是 $\frac{1}{n}$ 。事实上，我们可以用几种等价方式来写这件事。严格地说， P 是一个以事件为输入的函数，而不是以结果为输入的函数。因此，某个结果发生的概率应当用简单事件来写：

$$P(\{\omega\}) = \frac{1}{n}.$$

上一讲的核心公式是古典概率的黄金公式。我们先直观地引入它，然后又基于集合的性质仔细研究了它。黄金公式如下。

黄金公式

如果样本空间是有限的，并且所有结果都是等可能的，那么

$$P(A) = \frac{\text{有利结果的个数}}{\text{结果的总个数}} = \frac{|A|}{|S|}.$$

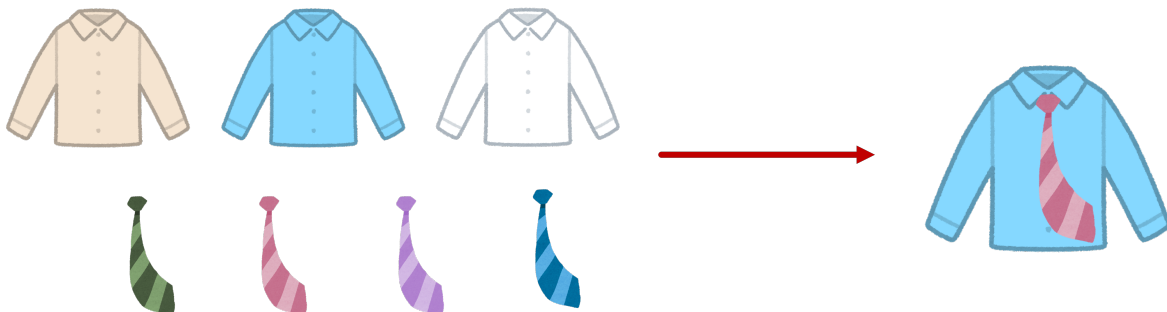
特别地，如果我们取一个简单事件 $A = \{\omega\}$ ，那么就可以应用黄金公式，得到每个结果的概率 $\frac{1}{n}$ 。

黄金公式只适用于古典概率，也就是说，它适用于样本空间 S 有限且 S 中所有结果都等可能的情形。很多时候，真正困难的并不是计算本身，而是把所有情况数清楚。接下来，我们会仔细看三个例子，说明如何在不同语境中应用黄金公式。这些例子基于第4讲的讨论：

- 从几个类别中作出选择，
- 将物体排列成不同顺序，
- 从一个大的集合中随机抽取。

1.2 例1：选择衬衫和领带

回忆第4讲中的例子：衣柜里有 3 件衬衫和 4 条领带，我们想各选一件来组成一套穿搭。例如，其中一种组合如下图所示。



在这个设定下，我们可以开始随机选择物品，因此也可以把穿搭问题变成概率问题。考虑下面三个问题。

练习 1

假设我们随机选择一件衬衫，然后随机选择一条领带，并且每件衬衫和每条领带被选中的机会相同。

- (a) 选择棕色衬衫的概率是多少？
- (b) 选择白色衬衫和红色领带的概率是多少？
- (c) 选择一套含有蓝色的穿搭的概率是多少？

为了回答这个问题，我们先建立一个样本空间。根据这个设定，结果是完整的穿搭选择。正如第4讲所见，它们可以排列成一个网格：



可以看到，一共有 $12 = 3 \times 4$ 种可能的穿搭。这意味着这个系统的样本空间 S 的大小是 12，即 $|S| = 12$ 。

- (a) 第一个问题可以用几种不同方式回答。首先，我们可以在非正式层面直接使用黄金公式进行计数。令 A 表示事件“选择棕色衬衫和任意领带”。则

$$P(A) = \frac{\text{含有棕色衬衫的穿搭数}}{\text{穿搭总数}} = \frac{4}{12} = \frac{1}{3}.$$

另一种更实际的看法是，衬衫的选择与领带的选择完全分离，因此我们可以把衬衫的选择单独看成一个概率系统。我们可以写出新的样本空间

$$S' = \{\text{白色衬衫}, \text{蓝色衬衫}, \text{棕色衬衫}\}.$$

共有三件衬衫，并且每件被选中的概率相同，所以选择棕色衬衫的概率就是 $\frac{1}{3}$ 。之后再选择领带并不会改变选择棕色衬衫的概率。

(b) 下一个问题要求我们求一个单独结果的概率：白色衬衫和红色领带。直接使用黄金公式可得

$$P(\text{白色衬衫和红色领带}) = \frac{\text{白色衬衫和红色领带的穿搭数}}{\text{穿搭总数}} = \frac{1}{12}.$$

另一种方法是，定义事件 A 为“选择白色衬衫”，事件 B 为“选择红色领带”。那么“选择白色衬衫和红色领带”可以重新解释为

A and B ：选择白色衬衫和红色领带。

由于这两个事件相互独立，我们可以使用之前的乘法公式，得到

$$P(A \text{ and } B) = P(A) \cdot P(B) = \frac{1}{3} \cdot \frac{1}{4} = \frac{1}{12}.$$

(c) 最后，我们计算选择某些蓝色物品的概率。这里同样有几种方法。也许最简单的方法是查看样本空间，并把所有含有蓝色的穿搭标出来：



可以看到，这样的穿搭有 6 种。因此，我们可以直接应用黄金公式得到

$$P(\text{某些蓝色物品}) = \frac{\text{含有蓝色的穿搭数}}{\text{穿搭总数}} = \frac{6}{12} = \frac{1}{2}.$$

另一种方法是定义两个事件

$B = \{\text{含有蓝色衬衫的穿搭}\}$ and $C = \{\text{含有蓝色领带的穿搭}\}.$

上图中，集合 B 用红色表示，集合 C 用蓝色表示。注意，这两个事件不是互斥的，因为存在一套穿搭同时包含蓝色衬衫和蓝色领带。这意味着计算 or 语句时，必须减去这个重叠部分：

$$P(B \text{ or } C) = P(B) + P(C) - P(B \text{ and } C) = \frac{4}{12} + \frac{3}{12} - \frac{1}{12} = \frac{6}{12} = \frac{1}{2}.$$

当然，所有这些问题也可以通过实际写出事件的符号形式来更正式地回答。我们可以把系统写成衬衫和领带选择所对应的有序对集合：



根据上图，棕色衬衫是第三件，对应 s_3 。我们可以把样本空间写成上方网格中所有有序对的集合：

$$S = \{(s_1, t_1), (s_1, t_2), \dots, (s_3, t_4)\},$$

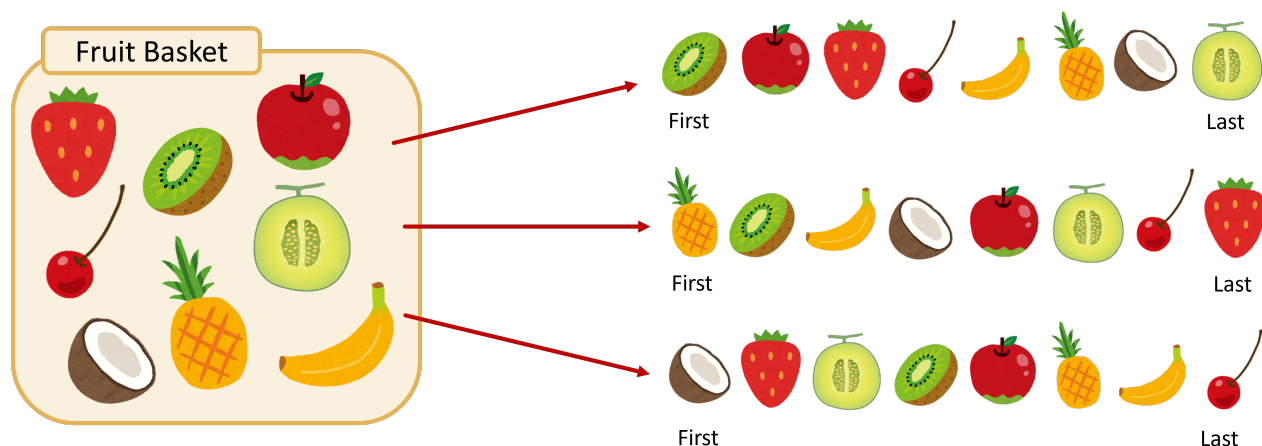
而所有含有棕色衬衫的穿搭组成的事件为

$$A = \{(s_3, t_1), (s_3, t_2), (s_3, t_3), (s_3, t_4)\}.$$

使用公式 $P(A) = \frac{|A|}{|S|}$ ，我们就得到与之前相同的结果 $P(A) = \frac{1}{3}$ 。

1.3 例2：排列水果

回忆第4讲中的水果篮例子。那里有 8 个不同的水果放在篮子里，我们逐个取出它们。正如我们所见，共有 $8! = 40320$ 种不同的取法，其中三种如下图所示：



练习 2

假设我们按顺序随机取出水果。换句话说，假设我们从篮子里的 8 个水果中随机选择一个排列。

- 第一个取出猕猴桃的概率是多少？
- 选择一个草莓排在第三位的排列的概率是多少？
- 选择一个甜瓜排在第二位且樱桃排在第三位的排列的概率是多少？

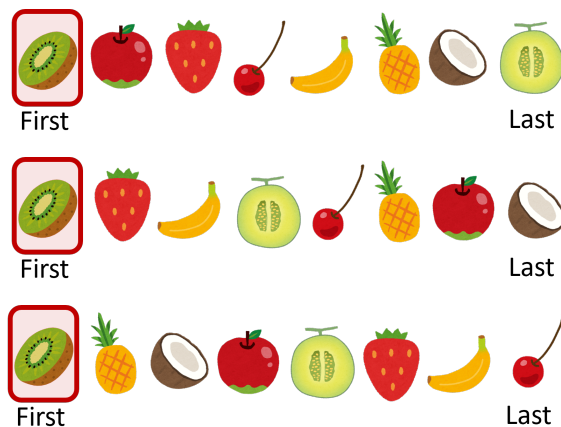
- (a) 我们用两种方法来说明。首先，最简单的方法也许是直接观察：对于排列中的第一个位置，篮子里的 8 个水果每个都有相同的概率被选中。共有 8 个水果，且只有一个是猕猴桃，因此第一个取出猕猴桃的概率应为

$$P(\text{猕猴桃排在第一位}) = \frac{\text{猕猴桃的个数}}{\text{水果总数}} = \frac{1}{8}.$$

如果我们在完整排列的层面上回答这个问题，也可以确认这一点。题目说明每个排列都是等可能的，所以我们也可以通过下面的比值计算第一个取出猕猴桃的概率：

$$P(\text{猕猴桃排在第一位}) = \frac{\text{猕猴桃排在第一位的排列数}}{\text{排列总数}}.$$

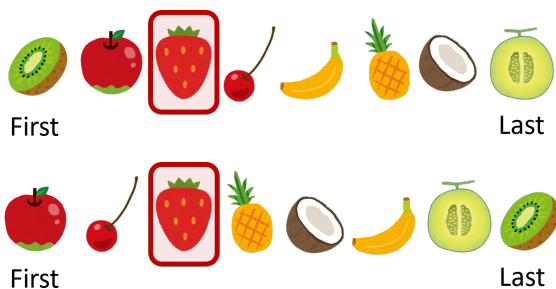
对于分母，第4讲中我们已经看到，排列 8 个水果共有 $8! = 40320$ 种不同方式。对于上面的分子，我们可以通过把猕猴桃固定在第一位来计算。例如，下面是三种猕猴桃固定在第一位的可能排列：

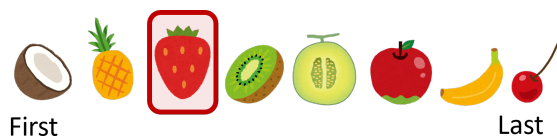


为了计算这些排列的总数，我们注意到还有 7 个水果需要排列，并且有 7 个位置可以放它们。因此，总的可能排列数是 $7!$ ，约分后得到与之前相同的概率：

$$\begin{aligned} P(\text{猕猴桃排在第一位}) &= \frac{\text{猕猴桃排在第一位的排列数}}{\text{排列总数}} \\ &= \frac{7!}{8!} = \frac{\cancel{7} \times \cancel{6} \times \cancel{5} \times \cancel{4} \times \cancel{3} \times \cancel{2} \times \cancel{1}}{8 \times \cancel{7} \times \cancel{6} \times \cancel{5} \times \cancel{4} \times \cancel{3} \times \cancel{2} \times \cancel{1}} = \frac{1}{8}. \end{aligned}$$

- (b) 按照同样的逻辑，我们可以把草莓固定在某个位置，然后排列它周围的所有水果。例如，三种这样的排列可能是：

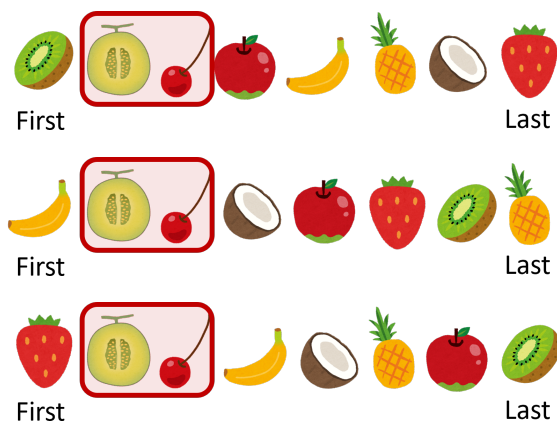




与上一问一样，草莓周围还有 7 个水果需要排列，并且有 7! 种方法。因此，概率再次化简为

$$P(\text{草莓排在第三位}) = \frac{\text{草莓排在第三位的排列数}}{\text{排列总数}} = \frac{7!}{8!} = \frac{7 \times \cancel{6} \times \cancel{5} \times \cancel{4} \times \cancel{3} \times \cancel{2} \times \cancel{1}}{8 \times \cancel{7} \times \cancel{6} \times \cancel{5} \times \cancel{4} \times \cancel{3} \times \cancel{2} \times \cancel{1}} = \frac{1}{8}.$$

- (c) 最后，我们要求的是选择一个甜瓜在第二位且樱桃在第三位的排列的概率。我们会作两种论证。首先，可以像之前一样进行纯计数。不过，与前两个问题不同，这里已经有两个位置被占据。例如，一些排列可能如下：



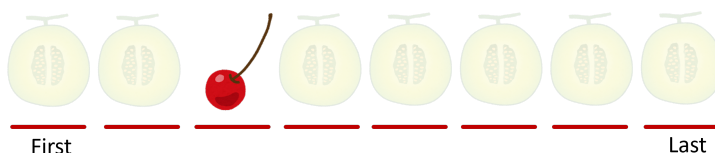
在这种情况下，我们实际上是在数剩下 6 个水果围绕甜瓜和樱桃重新排列的方法。由于有 6 个水果，总排列数是 6!，因此概率为

$$P(\text{甜瓜第二，樱桃第三}) = \frac{\text{甜瓜第二且樱桃第三的排列数}}{\text{排列总数}} = \frac{6!}{8!} = \frac{\cancel{6} \times \cancel{5} \times \cancel{4} \times \cancel{3} \times \cancel{2} \times \cancel{1}}{8 \times 7 \times \cancel{6} \times \cancel{5} \times \cancel{4} \times \cancel{3} \times \cancel{2} \times \cancel{1}} = \frac{1}{8 \times 7} = \frac{1}{56}.$$

另一种方法是把这个问题解释为复合事件。如果令

$$A = \{\text{甜瓜排在第二位}\}, \quad B = \{\text{樱桃排在第三位}\},$$

那么我们要求的概率是 $P(A \text{ and } B)$ 。根据 “and” 语句的公式，我们需要计算条件概率 $P(A|B)$ 。根据 (a) 和 (b) 的讨论，樱桃出现在第三位的概率是 $\frac{1}{8}$ ，因为有 8 个可能的位置，并且每个位置同样可能包含樱桃。相比之下，为了计算条件概率 $P(A|B)$ ，我们需要在已知樱桃会落在第三位的条件下，计算甜瓜被放在第二位的概率。既然我们知道樱桃在第三位，那么可用位置数就从 8 减少到 7：



甜瓜落在这些位置中的每一个位置都是等可能的。特别地，我们得到甜瓜出现在第二位的概率是 $\frac{1}{7}$ 。把这些放在一起，就可以用这个条件概率乘以 $P(B)$ 来计算复合概率：

$$P(A \text{ and } B) = P(A|B) \cdot P(B) = \frac{1}{7} \cdot \frac{1}{8} = \frac{1}{56}.$$

1.4 例3：皇家同花顺

扑克有几种不同玩法。一种简单的玩法叫作五张抽牌，顾名思义，每位玩家会得到 5 张手牌。玩家可以随意重新排列手牌，因此手牌中卡牌的顺序完全不重要。按照规则，最好的牌型是“皇家同花顺”，它由同一花色的最高 5 张牌组成：10、J、Q、K 和 A。

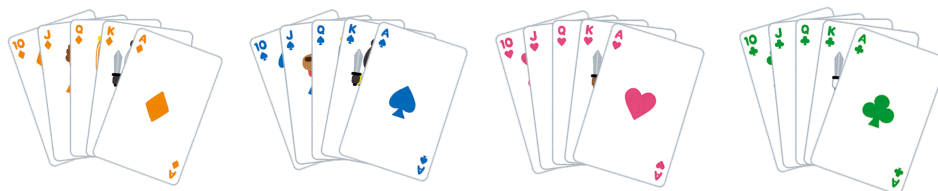
练习 3

玩家拿到皇家同花顺的概率是多少？

为了回答这个问题，我们会结合黄金公式和组合数学。概率应为

$$P(\text{皇家同花顺}) = \frac{\text{可能的皇家同花顺数}}{\text{可能的手牌数}}.$$

幸运的是，分子很容易计算。有 4 种花色，因此可能的皇家同花顺有 4 种：



为了计算分母，我们回忆一副标准扑克牌有 52 张牌。根据五张抽牌的规则，一副手牌由 5 张牌组成，并且手牌中牌的顺序并不重要。这意味着我们需要数从 52 张牌中选出 5 张，且不考虑顺序的方法数。为此，我们使用第4讲中的 $C_{n,r}$ 公式。这里 $n = 52$ ， $r = 5$ ：

$$C_{52,5} = \frac{52!}{5!47!} = 2,598,960.$$

有了这个结果，我们可以计算概率：

$$P(\text{皇家同花顺}) = \frac{\text{可能的皇家同花顺数}}{\text{可能的手牌数}} = \frac{4}{2,598,960} = \frac{1}{649,740}.$$

因此，拿到皇家同花顺的机会是 649,740 分之 1。这个例子很好地提醒我们，黄金公式本身非常简单。困难的部分通常在分母：在计算概率之前，我们需要充分理解样本空间的结构，才能把它数清楚。

2 超越古典概率

2.1 古典概率中的隐藏假设

到目前为止，我们几乎所有概率计算都使用了黄金公式：

$$P(A) = \frac{|A|}{|S|}.$$

在适用时，这个公式非常有用。但是，它的使用范围有限，因为它假设 n 个结果都以相同概率发生，即 $\frac{1}{n}$ 。这个假设通常可以由系统中的某种对称性来解释：公平硬币有两个相对的面，6 面骰子是一个立方体，等等。然而，现实中很多系统并不是这样完全对称的。在这些情况下，我们仍然可以讨论概率问题，但必须放松“每个结果都有相同机会发生”这一假设。

例如，假设一枚硬币落在正面或反面的概率不均匀，比如：

$$P(H) = 0.4, \quad P(T) = 0.6.$$

这里 $P(H)$ 的意思是 $P(\{H\})$ ， T 也是类似的。这里样本空间仍然是 $S = \{H, T\}$ ，但是概率不再都是 50%。换句话说，两个结果不再等可能。注意，在这里我们不能使用简单的黄金公式来确定事件的概率：

$$P(\text{正面朝上}) \neq \frac{S \text{ 中正面的个数}}{S \text{ 的大小}} = \frac{1}{2}.$$

这只是一个简单的例子，但它说明我们需要一种更一般的概率处理方式。接下来我们会详细讨论这一点。在本节中，我们仍然假设概率论的所有规则成立，但是我们不再假设每个结果都是等可能的。

2.2 给简单事件分配概率

令 S 是一个有限样本空间，例如：

$$S = \{\omega_1, \omega_2, \dots, \omega_n\}.$$

这里 ω_i 是这个系统的 n 个不同结果。我们不再假设每个 ω_i 都有固定概率 $\frac{1}{n}$ ，而是放松这个假设，把每个单独结果的概率简单写为

$$p(\omega_i),$$

这只是 $P(\{\omega_i\})$ 的简写。

为了与概率论的其他规则相容，我们仍然需要要求：

1. 每个概率 $p(\omega_i)$ 都是 0 到 1 之间的数；
2. 它们所有概率的总和等于 1：

$$\sum_{i=1}^n p(\omega_i) = p(\omega_1) + p(\omega_2) + \dots + p(\omega_n) = 1.$$

可以验证，我们的不公平硬币例子满足这些性质。使用这个设置，我们可以为更一般的情形建立一个新的“黄金公式”。

有限概率模型

令 $S = \{\omega_1, \omega_2, \dots, \omega_n\}$ 是一个大小为 n 的有限样本空间。一个概率模型为每个结果 ω_i 分配一个数 $p(\omega_i)$ ，使得

$$0 \leq p(\omega_i) \leq 1$$

并且

$$\sum_{\omega \in S} p(\omega) = 1.$$

那么，对任意事件 $A \subseteq S$ ，我们定义

$$P(A) = \sum_{\omega \in A} p(\omega).$$

最后这个公式就是黄金公式在一般情形下的替代品。它说明，为了计算事件 A 的概率，我们必须查看 A 中包含的所有结果 ω ，然后把它们的概率加起来。

事实上，古典概率的黄金公式只是这个更一般思想的一个特例。为此，假设事件 A 包含 m 个结果，也就是说

$$A = \{\omega_1, \dots, \omega_m\}.$$

如果我们暂时假设每个 ω_i 都带有相同概率 $\frac{1}{n}$ ，其中 n 是样本空间 S 的大小，那么简单的代数计算给出：

$$P(A) = \sum_{\omega \in A} p(\omega) = p(\omega_1) + \dots + p(\omega_m) = \underbrace{\frac{1}{n} + \dots + \frac{1}{n}}_{m \uparrow} = \frac{1 + \dots + 1}{n} = \frac{m}{n} = \frac{|A|}{|S|}.$$

因此，我们看到古典概率的黄金公式已经被包含在一个更大的框架之中。

2.3 熟悉的公式仍然成立

尽管我们已经超越了古典概率，第5讲中的主要概率公式仍然成立。为完整起见，我们再次列出它们。

概率公式

对于事件 A 和 B :

$$P(\text{not } A) = 1 - P(A).$$

对于 and 语句:

$$P(A \text{ and } B) = \begin{cases} P(A) \cdot P(B), & A \text{ 和 } B \text{ 独立时,} \\ P(A|B) \cdot P(B), & A \text{ 和 } B \text{ 相关时 (且 } P(B) \neq 0 \text{) .} \end{cases}$$

对于 or 语句:

$$P(A \text{ or } B) = \begin{cases} P(A) + P(B), & A \text{ 和 } B \text{ 互斥时,} \\ P(A) + P(B) - P(A \text{ and } B), & \text{一般情形.} \end{cases}$$

我们简要说明为什么这是合理的。

对于补事件，样本空间分成两个不重叠的部分:

$$S = A \cup A^c.$$

任何东西要么属于 A ，要么不属于 A 。因此

$$P(A) + P(A^c) = 1,$$

所以

$$P(A^c) = 1 - P(A).$$

对于 or 语句，问题在于重叠。如果我们把 $P(A)$ 和 $P(B)$ 相加，那么 $A \cap B$ 中的每个结果都会被数两次。因此我们减去一份:

$$P(A \text{ or } B) = P(A) + P(B) - P(A \text{ and } B).$$

对于 and 语句，一般规则来自条件概率。如果 $P(A) \neq 0$ ，那么

$$P(B|A) = \frac{P(A \text{ and } B)}{P(A)}.$$

整理得到

$$P(A \text{ and } B) = P(A) \cdot P(B|A).$$

如果 A 和 B 独立，那么知道 A 发生并不会改变 B 的概率。在这种情况下，

$$P(B|A) = P(B),$$

所以

$$P(A \text{ and } B) = P(A) \cdot P(B).$$

2.4 例：不公平的骰子

假设我们有一个不公平的骰子，其概率如下：

$$p(1) = 0.1, \quad p(2) = 0.1, \quad p(3) = 0.1, \quad p(4) = 0.1, \quad p(5) = 0.5, \quad p(6) = 0.1.$$

这不是一个公平骰子，因为掷出 5 的概率远高于其他每一个结果。然而，它仍然是一个完全有效的概率模型，因为每个单独的概率都在 0 和 1 之间，并且它们的总和正好等于 1：

$$0.1 + 0.1 + 0.1 + 0.1 + 0.5 + 0.1 = 1.$$

现在我们用新的公式来计算与这个系统有关的一些概率。

练习 4

使用上面的不公平骰子，计算下列概率：

- (a) 掷出不是 5 的数的概率是多少？
- (b) 掷出偶数的概率是多少？
- (c) 掷出小于 3 的数的概率是多少？

解答

- (a) 令 A 为事件“掷出 5”。那么事件“掷出不是 5 的数”就是 A 的补事件。因此

$$P(\text{不是 } 5) = 1 - p(5) = 1 - 0.5 = 0.5.$$

- (b) 偶数结果是 2、4 和 6。因此

$$P(\text{偶数}) = p(2) + p(4) + p(6) = 0.1 + 0.1 + 0.1 = 0.3.$$

- (c) 小于 3 的结果是 1 和 2。因此

$$P(\text{小于 } 3) = p(1) + p(2) = 0.1 + 0.1 = 0.2.$$

2.5 例：不公平的硬币

回到我们的不公平硬币例子，其中

$$P(H) = 0.4, \quad \text{and} \quad P(T) = 0.6.$$

练习 5

假设我们把这枚不公平硬币独立地连续抛掷若干次。

- (a) 连续两次正面朝上的概率是多少？
- (b) 一次正面、一次反面，以任意顺序出现的概率是多少？
- (c) 连续三次反面朝上的概率是多少？

解答

- (a) 令 A 为事件“第一次抛掷正面朝上”， B 为事件“第二次抛掷正面朝上”。这些抛掷相互独立，所以可以直接相乘：

$$P(HH) = P(H) \cdot P(H) = 0.4 \times 0.4 = 0.16.$$

- (b) 要得到一次正面和一次反面，并且顺序任意，有两种可能序列：

$$HT \quad \text{or} \quad TH.$$

这两个事件互斥，所以我们把它们的概率相加：

$$P(HT \text{ or } TH) = P(HT) + P(TH).$$

由于硬币抛掷相互独立，

$$P(HT) = 0.4 \times 0.6 = 0.24,$$

并且

$$P(TH) = 0.6 \times 0.4 = 0.24.$$

因此

$$P(HT \text{ or } TH) = 0.24 + 0.24 = 0.48.$$

- (c) 连续三次反面朝上的概率是

$$P(TTT) = P(T) \cdot P(T) \cdot P(T) = 0.6 \times 0.6 \times 0.6 = 0.216.$$

作为 (b) 的扩展，我们可以写出两次抛掷的完整样本空间：

$$S = \{HH, HT, TH, TT\}.$$

四个概率由乘法得到：

结果	HH	HT	TH	TT
概率	0.16	0.24	0.24	0.36

这些概率并不相等，并且总体上概率偏向于 TT ，远离 HH 。作为确认，这些概率确实应当加起来等于 1：

$$0.16 + 0.24 + 0.24 + 0.36 = 1.$$

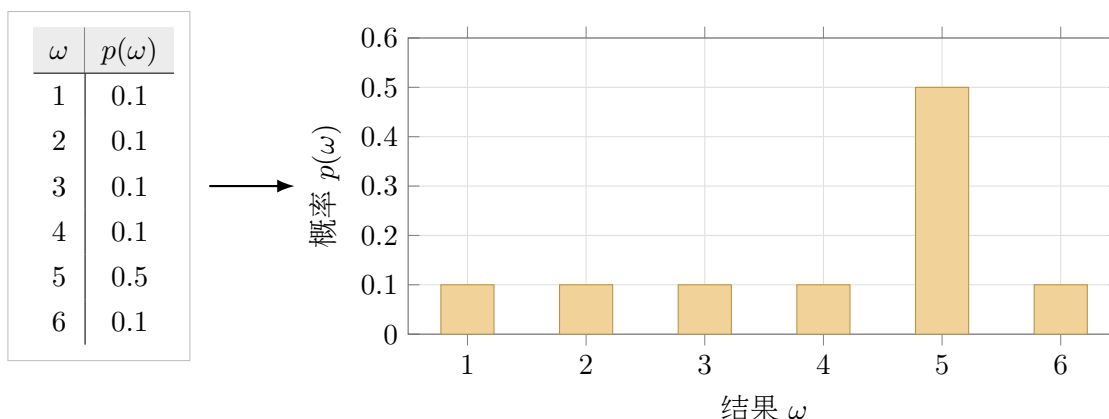
3 离散概率分布

3.1 从概率权重到分布

在第2.4节的不公平骰子例子中，我们给每个结果分配了概率：

$$p(1) = 0.1, \quad p(2) = 0.1, \quad p(3) = 0.1, \quad p(4) = 0.1, \quad p(5) = 0.5, \quad p(6) = 0.1.$$

现在我们要作一个概念上的跳跃：这一组信息可以像数据一样整理，也就是说，我们可以把它写成表格，并画成直方图：



这很像我们在第2讲中的讨论。不过，现在我们带有概率语境：柱子的高度不是单纯的计数，而是概率。因此，这个直方图告诉我们整个系统的概率如何分布在各个单独结果上。于是我们就得到了概率分布。

概率分布

概率分布告诉我们，概率如何分布在随机实验的可能结果或可能取值之间。

接下来，我们简要讨论概率分布横轴的含义。

3.2 随机变量

在上面的例子中，直方图的 x 轴由样本空间标记，也就是从 1 到 6 的数。在通常的数学中，我们可能用一个变量来标记这条轴，也就是一个神秘的字母 x ，它可以取实数轴上的值。普通图像通常通过让 x 变量沿着数轴移动，并计算相应输出而得到描述。对于概率分布，我们面对的不是简单地“滑动”变量，而是让变量随机地取样本空间中的某个值。因此，我们把 x -轴解释为显示随机变量的可能取值。我们通常用大写字母来表示随机变量，例如 X 。

我们会在后面的讲义中详细讨论随机变量，尤其是在第8讲处理二项分布时。现在，只需理解下面的定义就足够了。

定义：随机变量

随机变量是一个规则，它给随机实验的每个结果分配一个数值。

这里的思想是：给定的样本空间不一定由可以直接画成直方图的定量数据组成。因此，随机变量给了我们一种把结果转化为数字的方法，从而可以对得到的概率分布使用描述统计。

例如，假设我们有一枚公平硬币，并且抛掷两次。样本空间是定性的：

$$S = \{HH, HT, TH, TT\}$$

对应的概率为

$$P(HH) = P(HT) = P(TH) = P(TT) = \frac{1}{4}.$$

由此，我们可以通过计数来定义一个随机变量 X ，例如：

$$X = \text{正面的个数}.$$

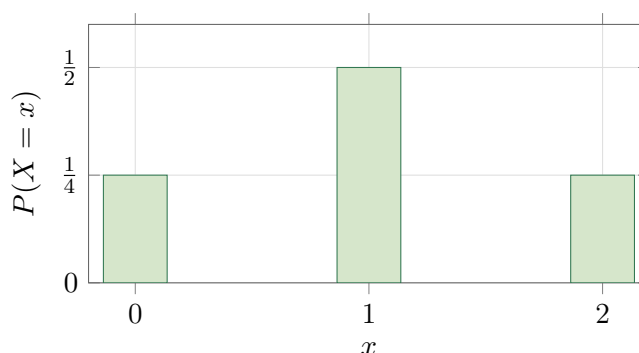
这使我们能够用数字来描述同一个系统:

$$X(HH) = 2, \quad X(HT) = 1, \quad X(TH) = 1, \quad X(TT) = 0.$$

从这里, 我们可以重新计算概率:

$$P(X = 0) = P(TT) = \frac{1}{4}, \quad P(X = 1) = P(HT) + P(TH) = \frac{1}{4} + \frac{1}{4} = \frac{1}{2}, \quad P(X = 2) = P(HH) = \frac{1}{4}.$$

由这些信息, 我们得到一个真正的概率分布:



3.3 离散概率分布

回忆一下, 普通数字, 例如 3, 既可以有离散的解释, 也可以有连续的解释: 我们可以把 3 看成某些东西的数量, 也可以把 3 看成一段连续区间的长度。在本讲中, 我们只处理离散随机变量, 因为我们的样本空间都是有限大小的。

定义: 离散概率分布

如果 X 是一个有限离散随机变量, 其可能取值为

$$x_1, x_2, \dots, x_n,$$

那么一个**离散概率分布**会给每个取值分配一个概率:

$$P(X = x_1), \quad P(X = x_2), \quad \dots, \quad P(X = x_n).$$

这些概率必须满足

$$0 \leq P(X = x_i) \leq 1$$

并且

$$P(X = x_1) + P(X = x_2) + \dots + P(X = x_n) = 1.$$

一旦我们有了分布, 概率问题就变成了加法问题。例如,

$$P(X \geq 1) = P(X = 1) + P(X = 2) = \frac{2}{4} + \frac{1}{4} = \frac{3}{4}.$$

用语言来说, 这就是至少得到一次正面的概率。

类似地,

$$P(X < 2) = P(X = 0) + P(X = 1) = \frac{1}{4} + \frac{2}{4} = \frac{3}{4}.$$

如何使用离散分布

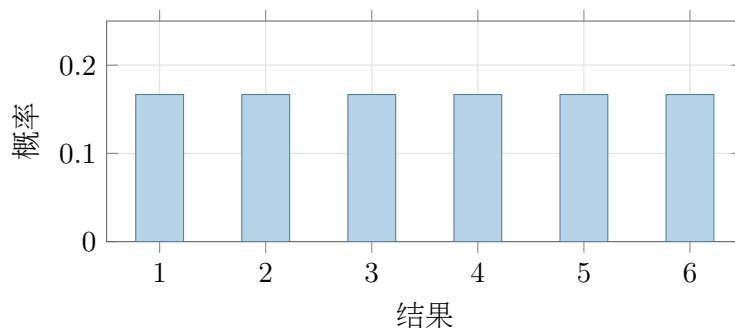
要计算涉及 X 的某个条件的概率, 就把所有满足该条件的取值 x 的概率加起来。

例如,

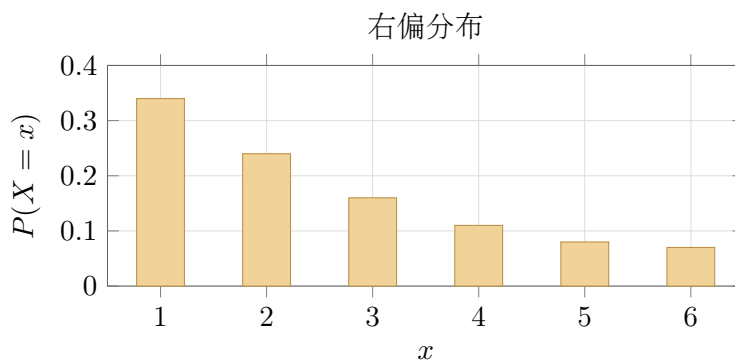
$$P(X \geq 1) = \sum_{x \geq 1} P(X = x).$$

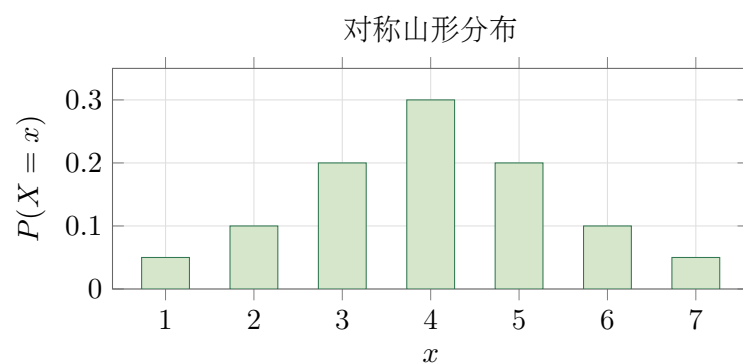
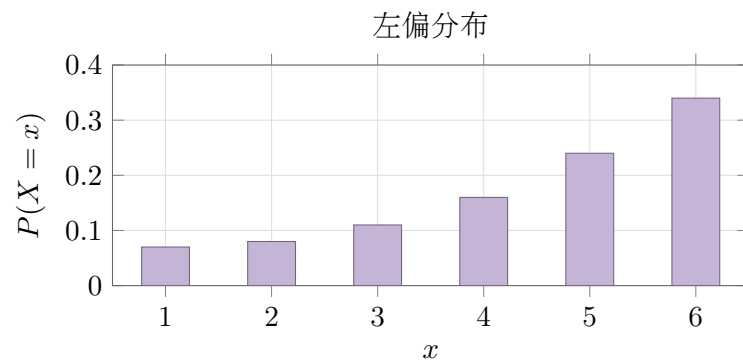
3.4 分布的形状

正如我们在第3讲中看到的, 直方图的形状可以有效地告诉我们一个分布的整体结构。同样的思想也适用于概率分布。也许最简单的情况就是古典概率。在这里, 每个结果都有完全相同的概率, 因此得到的分布是均匀的:



不过, 其他系统可能呈现不同的形状。例如, 概率可能向左偏斜、向右偏斜, 甚至呈对称的钟形或山形:





在第8讲和第9讲中，我们会详细讨论离散概率分布。正如我们将会看到的，二项分布是概率中一个特别有趣的设置。

哇！我没来得及把期望值那一部分打出来！我会在周末补上。