《初等概率论》第 1 讲

邓 婉 璐

清华大学
统计学研究中心

September 17, 2018

目录

一、预备知识
——集合论

1. 概念：

1. 集合：将一些研究对象放在一起，形成集合. 通常用字母 Ω 或 S 等表示；
2. 元素：集合里面的对象. 通常用小写字母表示，如 x, y, a, b, c 等.
3. 空集：不包含任何元素的集合，记作：$\emptyset$.

2. 集合的表示方法

1. 列举法：$S=\{x_1, x_2, ..., x_n, ...\}$, 如果 S 含有至多可列个元素；如 $S=\{H, T\}$ 或 $S=\{1, 2, ..., 6\}$ 或 $S=\{1, 2, ...\}$.
2. 描述法：以 x 具有某种性质 P 为条件来刻画一个集合，记作 $\{x|x \text{ 满足性质 } P\}$, 如 $\{x|0 \le x \le 1\}$.

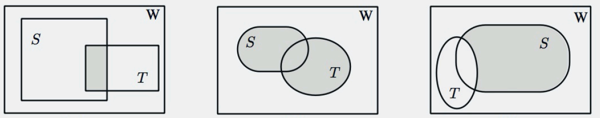
3. 关系：

1. 元素与集合的关系：元素 x 与集合 S 的隶属关系，要么 $x \in S$，要么 $x \notin S$.
2. 集合与集合的关系：
   - 子集：若集合 S 的所有元素均为集合 T 的元素，称 S 为 T 的子集，记作 $S \subset T$，或者 $T \supset S$. 特别地，$\emptyset \subset S$.
   - 相等：若 $S \subset T$ 且 $T \subset S$，则 $S = T$.

在以后的课程中，将我们感兴趣的所有元素放在一起，形成一个集合，称此集合为空间，记作 Ω. 当 Ω 确定以后，我们所讨论的集合 S 都是 Ω 的子集.

4. 运算：

1. 补集 (Complement)：称集合 $\{x\in\Omega|x\notin S\}$ 为集合 S 的补集，记作 S^c. 特别地 $\Omega^c=\emptyset$.
2. 并集 (Union)：由属于 S 或属于 T 的元素组成的集合，称为 S 和 T 的并，记作 $S\cup T$.
3. 交集 (Intersection)：由属于 S 且属于 T 的元素组成的集合，称为 S 和 T 的交，记作 $S\cap T$.
4. 差集 (difference)：由属于 S 但不属于 T 的元素组成的集合，称为 S 与 T 的差集，记作 $S\setminus T$. 显然 $S\setminus T=S\cap T^c$.
5. 对称差集 (symmetric difference)：记作 $S\triangle T$. 显然 $S\triangle T=(S\cap T^c)\cup(T\cap S^c)$.



(a) (b) (c)

推广情形：

1. 可列并或者不可列并

$$\bigcup_{\gamma\in\Gamma} S_\gamma = \{x|x\in S_\gamma \text{对某个}\ \gamma\ \text{成立}\}$$

2. 可列交或者不可列交

$$\bigcap_{\gamma\in\Gamma} S_\gamma = \{x|x\in S_\gamma \text{对所有}\ \gamma\ \text{成立}\}$$

5. 不相交、分割：

1. 称 S 和 T 是不相交的，如果集合 $S\cap T=\emptyset$。
2. 一组集合称为互不相交的，如果其中任何两个集合都没有公共元素。
3. 集合 S 的分割：如果一组集合是互不相交的，且它们的并为 S。

6. 集合的代数：

集合运算的若干性质：

1. $S\cup T = T\cup S$; (交换律)
2. $S\cup(T\cup U) = (S\cup T)\cup U$; (结合律)
3. $S\cap(T\cup U) = (S\cap T)\cup(S\cap U)$,
   $S\cup(T\cap U) = (S\cup T)\cap(S\cup U)$, (分配律)
4. $(S^c)^c = S, \quad S\cap S^c = \emptyset$,
5. $S\cup\Omega = \Omega, \quad S\cap\Omega = S$.

De Morgan's laws (德摩根定律或者德摩根法则)

$$\left(\bigcup_{\gamma\in\Gamma} S_\gamma\right)^c = \bigcap_{\gamma\in\Gamma} S_\gamma^c, \qquad \left(\bigcap_{\gamma\in\Gamma} S_\gamma\right)^c = \bigcup_{\gamma\in\Gamma} S_\gamma^c.$$

二、概率模型

概率模型的基本构成

- 样本空间 Ω: ?
- 概率 (probability): ?

概率模型

1. 样本空间

1. 样本空间是一个集合：每一个概率模型都关联着一个试验，这个试验将产生一个试验结果．该试验的所有可能结果形成样本空间，记作 Ω.
2. 样本空间的试验结果必须满足：互斥，并且完整.
3. 样本空间的试验结果可能是有限，也可能是无限个试验结果组成.
4. 一个试验由什么构成，并没有什么限制，然而我们所讨论的概率模型的问题中，只涉及一个试验，所以连续抛三次硬币的试验，只能作为一次试验，不能认为是三次试验.

space

概率模型

2. 选择样本空间的艺术

1. 即便对同一个试验，根据我们的兴趣也可以确定不同的模型.
2. 在确定 Ω 时，既要有足够多的细节，又要避免不必要的繁琐.

例子 1

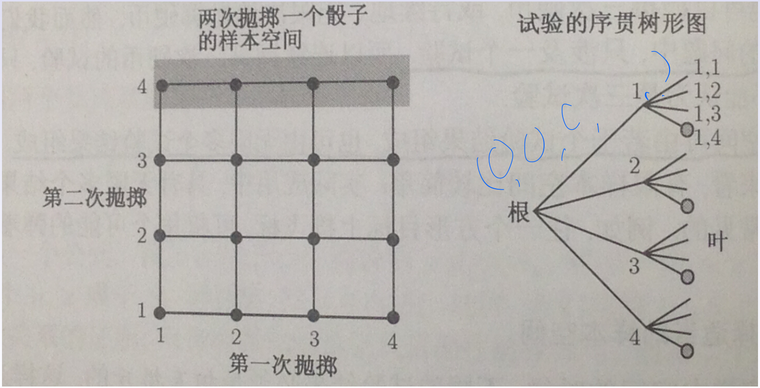
考虑两个不同的游戏，它们都涉及连续抛掷 10 次硬币.

1. 每次抛掷硬币的时候，只要出现正面向上，甲就赢 1 元钱；（只与 10 次抛掷中正面向上的次数有关）
2. 直到出现第一次正面向上（包括这一次），甲就赢 1 元钱；继续抛掷，直到第二次出现正面向上，甲就赢 2 元钱；每次抛掷得到正面向上的时候，以后每次抛掷硬币所赢的钱数比以前每次抛掷硬币所赢的钱数加倍.（不仅跟次数有关，还跟出现的顺序有关）

《初等概率论》
第 1 讲
邓婉璐
目录
预备知识
概率模型
随机抽样与随机分配
小结
作业

3. 样本空间－离散例子（序贯模型）

很多实验本身具有序贯的特征，比如连续抛掷一枚硬币，一共抛三次；或者连续观测一只股票，共观测五天。常用**序贯树形图**来刻画样本空间中的试验结果。

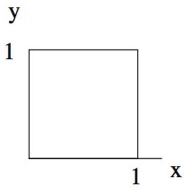


图：序贯树形图示例

3. 样本空间 – 连续例子

甲和乙约定在某时刻见面，而每个人到达约会地点的时间都会有延迟，延迟时间在 $0 \sim 1$ 小时．考虑直角坐标系的单位正方形 $\Omega = [0,1] \times [0,1]$．正方形中的每一个点的两个坐标分别表示他们可能的延迟时间．

$$\Omega = \{(x,y) \mid 0 \le x, y \le 1\}$$



图：连续样本空间示例

《初等概率论》
第 1 讲

邓婉璐

目录
预备知识
概率模型
随机抽样与随机分配
小结
作业

4. 概率

事件：样本空间的子集，即某些试验结果的集合.

概率：分配到事件上。

概率公理

1. (非负性) 对一切事件 A，满足 $\mathbf{P}(A) \geq 0$.
2. (归一化) $\mathbf{P}(\Omega) = 1$.
3. (可加性) 若 $A \cap B = \emptyset$，那么 $\mathbf{P}(A \cup B) = \mathbf{P}(A) + \mathbf{P}(B)$.

- $\mathbf{P}(\{s_1, s_2, ..., s_k\}) = \mathbf{P}(\{s_1\}) + ... + \mathbf{P}(\{s_k\}) = \mathbf{P}(s_1) + ... + \mathbf{P}(s_k)$.
- 公理 3 需要加强。
- 任意奇怪的集合也都有概率么？

measure

$\int f(x)\,dx$

Lebesgue.

《初等概率论》
第 1 讲

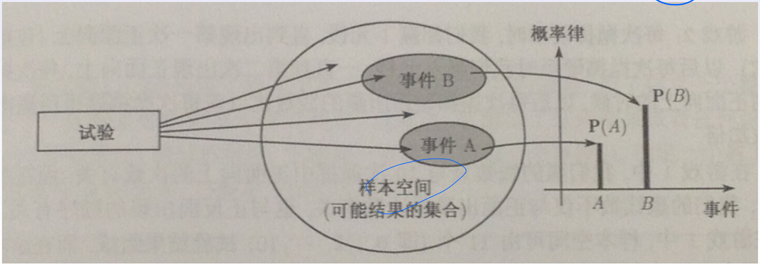
邓婉璐

目录
预备知识
概率模型
随机抽样与随机分配
小结
作业

4. 概率

概率模型的基本构成

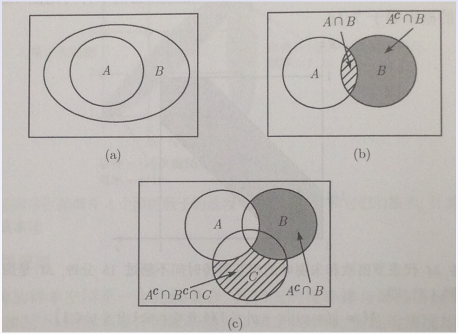
- 样本空间 Ω: 一个试验的所有可能结果的集合;
- 概率: 概率就是为试验结果的集合 A(称之为事件) 确定一个非负数 $\mathbf{P}(A)$ (称为事件 A 的概率). 此非负数刻画了我们对事件 A 的认识或所产生的信念程度.



概率模型

概率的若干性质：

1. 若 $A \subset B$, 则 $\mathbf{P}(A) \leq \mathbf{P}(B)$.
2. $\mathbf{P}(A \cup B) = \mathbf{P}(A) + \mathbf{P}(B) - \mathbf{P}(A \cap B)$.
3. $\mathbf{P}(A \cup B) \leq \mathbf{P}(A) + \mathbf{P}(B)$.
4. $\mathbf{P}(A \cup B \cup C) = \mathbf{P}(A) + \mathbf{P}(A^c \cap B) + \mathbf{P}(A^c \cap B^c \cap C)$.



图：利用维恩图直观验证概率的性质

5. 离散模型

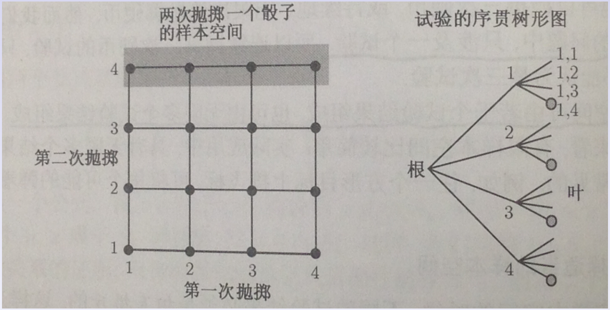


图: 序贯树形图示例

假设骰子是均匀的，每一面出现的机会是相同的。

$\mathbf{P}((X,Y)\ \ is\ \ (1,1)\ \ or\ \ (1,2)) = ?,\quad \mathbf{P}(\{X=1\}) = ?,$

$\mathbf{P}(X+Y\ \ is\ \ odd) = ?,\quad \mathbf{P}(min(X,Y)=2) = ?.$

例子 3

依次抛掷 3 枚硬币，试验结果是由正面和反面组成的长度为 3 的序列. 样本空间为

$$\Omega = \{HHH, HHT, HTH, HTT, THH, THT, TTH, TTT\}.$$

假定上述 8 种结果的可能性是相同的，即每个结果的概率为 $1/8$.

计算事件 $A = \{$ 两个正面向上，一个反面向上 $\}$ 的概率.

$A = \{HHT, HTH, THH\}$, 因此

$$\begin{aligned}\mathbf{P}(A) &= P(\{HHT, HTH, THH\}) \\ &= \mathbf{P}(\{HHT\}) + \mathbf{P}(\{HTH\}) + \mathbf{P}(\{THH\}) \\ &= 1/8 + 1/8 + 1/8 = 3/8.\end{aligned}$$

概率模型

古典概型（离散均匀概率）

设样本空间 Ω 由 n 个等可能性的试验结果组成，因此每个试验结果组成的事件（称为基本事件）的概率是相等的．由此得到

$$\mathbf{P}(A) = \frac{\text{含于事件 } A \text{ 的试验结果数}}{n}.$$

6. 连续模型

例子 4

甲和乙约定在某时刻见面，而每个人到达约会地点的时间都会有延迟，延迟时间在 0 ∼ 1 小时. 第一个到达约会地点的人会在那儿等待 15 分钟，等了 15 分钟后若对方还没有出现，先到者会离开约会地点. 问他们能够相会的概率有多大?

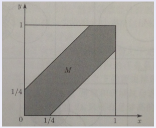


图: 甲乙相互等待时间不超过 15 分钟

将 Ω 的子集出现的概率定义为这个子集的面积. 此定义符合概率的三条公理. 令 $M = \{(x,y) \| |x-y| \le 1/4, 0 \le x \le 1, 0 \le y \le 1\}$. 则所求的概率即为 M 的面积，为 $7/16$.

概率模型

r- dim

用 R^r 表示 $r-$ 维向量空间，

$$R^r = \{(x_1, x_2, ..., x_r) | x_i \in (-\infty, \infty), 1 \le i \le r\}.$$

对于 R^r 的子集 A，用 $m(A) = \int_A dx_1 dx_2 ... dx_r$ 表示 A 的体积。

H

几何概率模型（连续均匀概率）

设样本空间 $\Omega \subset R^r$ 的体积 $m(\Omega)$ 是正数，且 Ω 中的每个试验结果发生的可能性相同，则对于事件 $A \subset \Omega$，其发生的概率为

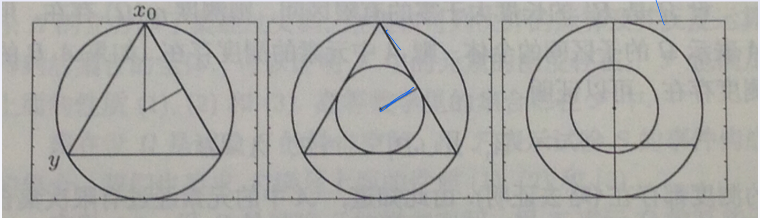
$$\mathbf{P}(A) = \frac{m(A)}{m(\Omega)}.$$

7. 模型与现实

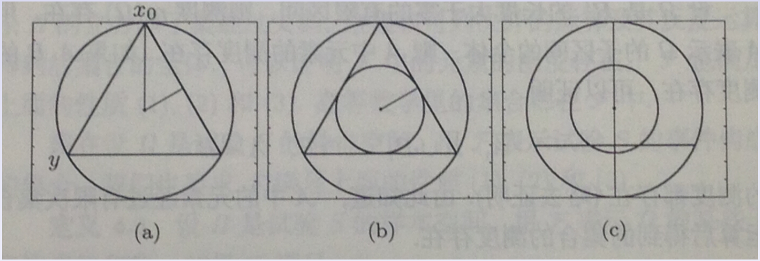
例子 5 (Bertrand 悖论)

在半径为 1 的圆内任取一条弦，求弦长大于等于 $\sqrt{3}$ 的概率.

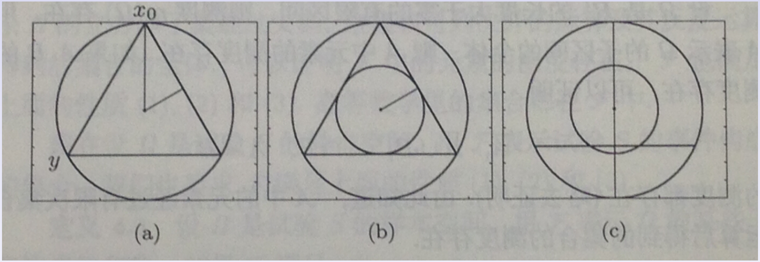
圆的内接等边三角形的边长为 $\sqrt{3}$. 用 Ω 表示试验的样本空间，用 A 表示得到的弦长大于等于 $\sqrt{3}$.



(a) (b) (c)



(1). 如果认为弦的端点等可能的落在圆周上，则点 x_0 确定后另一点 y 也等可能的落在圆周上 (图 (a)). 因为 $\Omega = \{y|y \in [0, 2\pi)\}$, $A = \{y|2\pi/3 \le y \le 4\pi/3\}$. 于是 $\mathbf{P}(A) = 1/3$.



(a) (b) (c)

(2). 如果认为弦的中点等可能的落在圆内 (图 (b)), 则 Ω 是大圆盘, A 是小圆盘. 因此 $\mathbf{P}(A) = 1/4$.

《初等概率论》
第 1 讲

邓婉璐

目录

预备知识

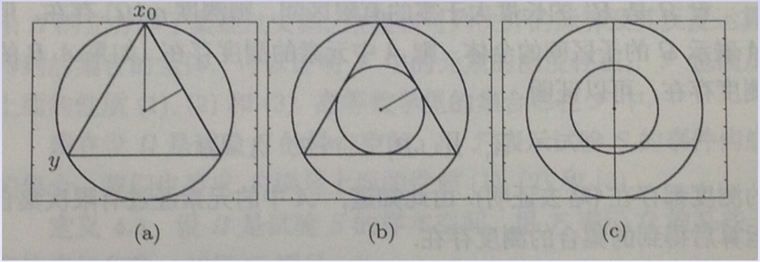
概率模型

随机抽样与随机分配

小结

作业

概率模型

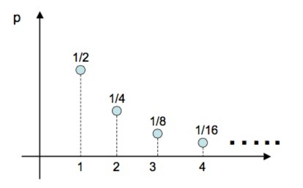


(3). 如果认为弦的中点等可能的落在与之垂直的直径上 (图 (c))：当且仅当其中点与圆心的距离小于 1/2 时，它的弦长才大于 $\sqrt{3}$. 因此 $\mathbf{P}(A) = 1/2$.

从本例可以看出，不同的等可能假设可以导致不同的计算结果.

8. 概率公理 3: 加强为可列可加性

- 样本空间: $\{1, 2, \ldots\}$
- $\mathbf{P}(n) = 2^{-n}, \quad n = 1, 2, \ldots$



- 求 $\mathbf{P}(outcome \;\; is \;\; even)$.
- $\mathbf{P}(\{2, 4, 6, \ldots\}) = \mathbf{P}(2) + \mathbf{P}(4) + \ldots = \frac{1}{2^2} + \frac{1}{2^4} + \ldots = \frac{1}{3}$.

可列可加性:

若 $A_1, A_2, \ldots$ 是互不相交的事件, 那么

$$\mathbf{P}(A_1 \cup A_2 \cup \ldots) = \mathbf{P}(A_1) + \mathbf{P}(A_2) + \ldots$$

三、随机抽样与随机分配

♣ 排列与组合

1. n 个对象的排列数：$n!$;
2. n 个对象中取 k 个对象的排列数：$P_n^k = n!/(n-k)!$;
3. n 个对象中取 k 个对象的组合数：$\binom{n}{k} = \frac{n!}{k!(n-k)!}$;

♣ 讨论在含有 M 个不同球的箱子中抽取 n 个球的不同方法.

1. 放回抽样

如果每次将抽到的球在下一次抽球前放回箱子中，则试验称做放回抽样. 这时，由 n 个球形成的每一个样本可以表示为向量 $(a_1, ..., a_n)$，其中，$a_i(i = 1, ..., n)$ 是第 i 次抽到的球的编号. 易见，对于放回抽样，每个 a_i 可以是 $\{1, 2, ..., M\}$ 中的任何一个数. 样本空间的描述，本质上与如下情形有关：诸如 $(4, 1, 2, 1)$ 和 $(1, 4, 2, 1)$ 是认为是 '不同的基本事件'，还是 '同一基本事件'. 因此，习惯上区分两种情形：有序抽样和无序抽样.

随机抽样与随机分配

对有序抽样，由相同元素组成的两个样本，只要其中元素的前后顺序有所不同，就视为‘不同’的样本. 对无序抽样，不管元素的顺序，只要由相同元素组成的样本，都视为‘同一个’样本. 为强调具体的样本属于哪一种，对有序样本，使用记号 $(a_1, ..., a_n)$，而无序样本则记作 $[a_1, ..., a_n]$.

(1.1). 放回有序抽样.

样本空间 Ω 具有如下构造：

$$\Omega = \{\omega : \omega = (a_1, ..., a_n), \quad a_i \in \{1, 2, ..., M\}\},$$

且 $\#(\Omega) = M^n$.

(1.2). 放回无序抽样.

样本空间 Ω 具有如下构造：

$$\Omega = \{\omega : \omega = [a_1, ..., a_n], \quad a_i \in \{1, 2, ..., M\}\},$$

且 $\#(\Omega) = \binom{M+n-1}{n}$.

2. 无放回抽样

假设 $n \le M$，且凡是抽到的球都不再放回. 这里也考虑两种可能的抽法：有序抽样和无序抽样.

(2.1). 不放回有序抽样.

样本空间 Ω 具有如下构造：

$$\Omega = \{\omega : \omega = (a_1, ..., a_n), a_k \neq a_l, k \neq l, a_i \in \{1, 2, ..., M\}\},$$

且 $\#(\Omega) = M(M-1)\cdots(M-n+1) = P_M^n$.

(2.2). 不放回无序抽样.

样本空间 Ω 具有如下构造：

$$\Omega = \{\omega : \omega = [a_1, ..., a_n], a_k \neq a_l, k \neq l, a_i \in \{1, 2, ..., M\}\},$$

且 $\#(\Omega) = \binom{M}{n}$.

随机抽样与随机分配

表: 汇总表

	抽样方式		不同抽样的总数
自含有 M 个球箱子中的 n 次抽样	放回	有序	M^n
		无序	$\binom{M+n-1}{n}$
	不放回	有序	P_M^n
		无序	$\binom{M}{n}$

表: 对 $M=3, n=2$，相应基本空间的结构列表.

	抽样方式		基本事件
自含有 M 个球箱子中的 n 次抽样	放回	有序	$(i,j), i,j=1,2,3$
		无序	$[1,1][2,2][3,3][1,2][1,3][2,3]$
	不放回	有序	$(1,2)(1,3)(2,1)(2,3)(3,1)(3,2$
		无序	$[1,2][1,3][2,3]$

♣ 按箱分配质点 (选修)

考虑按箱分配质点问题：将 n 个质点 (小球) 分配到 M 个箱子中去，并研究质点分配问题中样本空间的构造. 此问题产生于统计物理，比如，在研究 n 个粒子 (如质子、电子、...) 按 M 种状态的分布.

假设 M 个箱子一一编号为 $1, 2, ..., M$，而 n 个质点可以辨别 (两两不同)，且分别编号 $1, 2, ..., n$. 那么，n 个质点分配到 M 个箱子中，完全可以由数组 $(a_1, ..., a_n)$ 描绘，其中 a_i 表示第 i 个质点分配到编号为 a_i 的箱中；假如 n 个质点不可辨别 (完全相同)，则它们分配到 M 个箱中完全可以由数组 $[a_1, ..., a_n]$ 描绘.

对应关系：

$$(\text{有序抽样}) \Longleftrightarrow (\text{质点可以辨}),$$

$$(\text{无序抽样}) \Longleftrightarrow (\text{质点不可辨}).$$

类似地，对应关系：

$$(\text{放回抽样}) \Longleftrightarrow (\text{箱中可容纳任意多质点}),$$

$$(\text{不放回抽样}) \Longleftrightarrow (\text{箱中最多可容纳一个质点}).$$

	分配方式		基本事件
自含 M 个球箱子 Ω 中的 n 次抽样	有抑制	可辨	[●][ ][○] [○][ ][●] [○][●][ ] [●][○][ ] [ ][○][●] [ ][●][○]
		不可辨	[+][ ][+] [ ][+][+] [+][+][ ]
	无抑制	可辨	[○][ ][●] [●][○][ ] [ ][○][●] [ ][ ][○●] [ ][○●][ ] [●][ ][○] [○][●][ ] [ ][●][○] [○●][ ][ ]
		不可辨	[++][ ][ ] [ ][++][ ] [ ][ ][++] [+][+][ ] [+][ ][+] [ ][+][+]

上图演示了 3 个小球分配到 2 个箱子中的情形，其中‘○’和‘●’分别代表白球和黑球，‘+’表示分配到箱子中的球不可辨.

“有抑制”的箱子 (最多容纳一个质点的箱子),
“无抑制”的箱子 (可容纳任意个质点的箱子).

关于古典概型

古典概型问题小技巧:

- 不要失去常识、也不要过分依赖
- 用最简单情形验证
- 将物体编号. 例如, 10 个人分成两组, (4,6), (5,5)
- 情景证明 (story proof)

生日匹配问题－直观

小结

- 公理化的定义：样本空间、概率
- 古典概型

- 注意定义的明确性 (e.g. 样本空间、等可能)
- 一些小的技巧，直观化、简单化

作业

课后习题：2，5，6，7，8，9，10. 其中第 8 题的中文翻译不准确，第四句应为“必须连续赢两场才算赢”.

Thank you!