§1.2 概率

① 主观概率

某企业想投资一个新的项目，那么投资成功的可能性多大呢？A认为投资成功的概率为0.7, B认为投资成功的概率为0.5, 说明什么？

② 频率理论学派的客观概率

通过一枚均匀的硬币打赌, 为什么公平?为什么可能性各为50%？

③ 以测度为基础的概率公理化体系(重点)

概率：推理不确定性的数学

微积分：推理变化率的数学

§1.2 概率

一、概率

随机事件发生的可能性大小是一个客观存在的量。
(机遇的数量化)。

概率是刻划随机事件发生可能性大小的数量指标。

客观量度

事件*A*出现的概率 (*Probability*) 记为*P*(*A*)

注: 随机现象的统计规律:

(物理或试验)直观理解和针对的广泛性.

例如 抛骰子试验 抛骰子试验

例如 抛骰子试验

例1 抛一颗均匀的骰子，观察其出现的点数情况。

我们通过实践与分析可得：

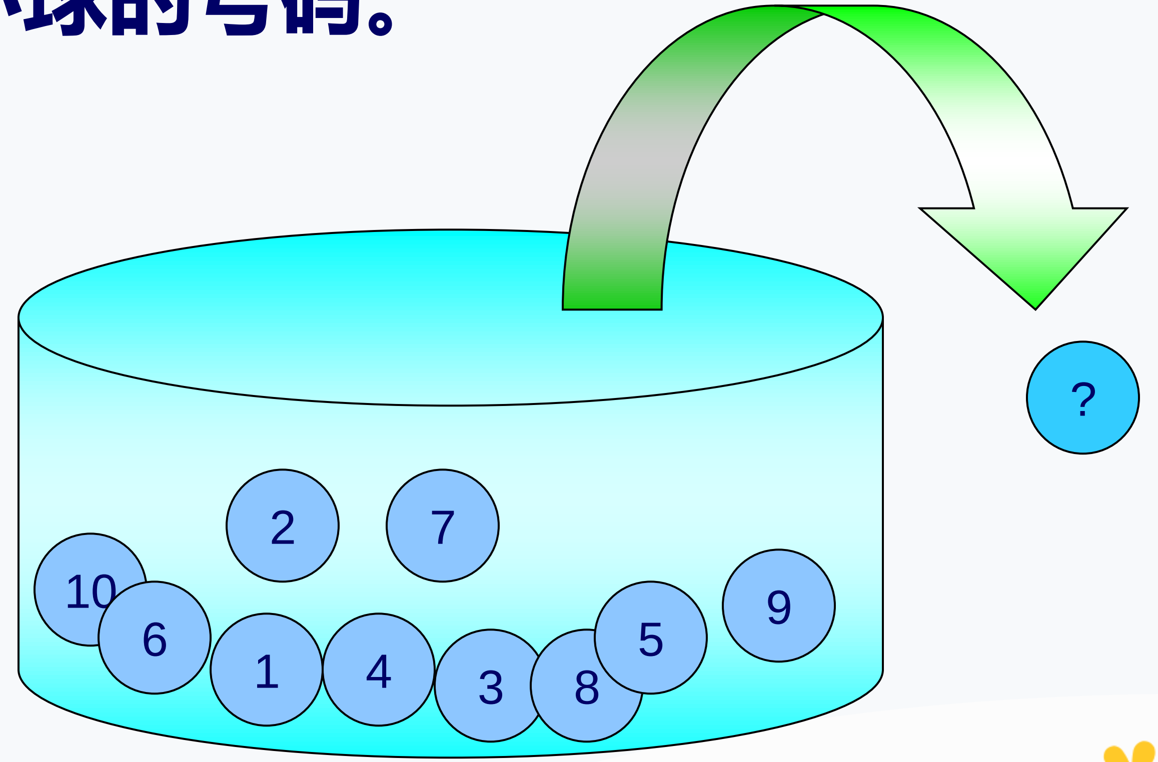
出现的点数为1, 2, 3, 4, 5, 6

的可能性都是相等的。

抛骰子试验

例2 从 10个标有号码 1, 2,…, 10 的小球中任取一个, 记录所得小球的号码。

我们可得：摸出任一号码的小球的可能性是相同的，这是客观存在的事实。



§1.2 概率

一、概率

随机事件发生的可能性大小是一个客观存在的量。
(机遇的数量化)

概率是刻划随机事件发生可能性大小的数量指标。

事件*A*出现的概率 (*Probability*) 记为*P*(*A*)

客观量度

注: 随机现象的统计规律, 针对的广泛性.

例题: 新开发的某药品治愈某疾病效果达到70%.

如何计算概率？怎样客观量度随机事件发生可能性大小？

二：频率

定义：在相同条件下，进行了n次试验，事件A发生了m次，称比值

$$f_n(A)=\frac{m}{n}$$

为事件A发生的频率。

频率从一定程度上反映了事件发生可能性的大小，可能因试验的次数、试验者的变化而改变。

例如：

抛硬币试验

π的小数频率

例3 抛一枚硬币，观察其出现正面H和反面T的情况。

我们通过实践与分析可得：硬币出现正面的可能性等于它出现反面的可能性。

历史上几位著名科学家亲自做了实验，并记录下来。

实验者	抛掷次数	出现正面次数	m/n
德.摩根	2048	1061	0.5181
蒲丰	4040	2048	0.5069
皮尔逊	24000	12012	0.5005
维尼	30000	14994	0.4998

例4 圆周率π的计算。

刘徽(公元263年，割圆术) π=3927/1250=3.1416。

祖冲之(429~500) $3.1415926 < \pi < 3.1415927$。

威廉.向克斯：用20年时间于1872年将π算到小数后707位。

法格逊怀疑向克斯的结果，用了一年的时间，发现向克斯π只有前527位是正确的。

法格逊猜想：在π的数值中各数码0,1,…9出现的可能性大小应当相等。

1937年，法国学者对π的前100万位小数中各数码的频率统计结果表明，尽管各数字出现也有起伏，但频率都稳定于1/10。

数码	出现次数	出现频率
0	99959	0.1000
1	99758	0.0998
2	100026	0.1000
3	100229	0.1002
4	100230	0.1002
5	100359	0.1003
6	99548	0.0995
7	99800	0.0998
8	99985	0.1000
9	100106	0.1001

2021年8月17日，美国趣味科学网站报道，瑞士研究人员使用一台超级计算机，历时108天，将著名数学常数圆周率π计算到小数点后62.8万亿位，创下该常数迄今最精确值记录。

二：频率

定义：在相同条件下，进行了n次试验，事件A发生了m次，称比值 $f_n(A)=\dfrac{m}{n}$ 为事件A发生的频率。

频率：从一定程度上反映了事件发生可能性的大小，为主观的量。

频率具有稳定性：在一定条件下，频率稳定于概率。但频率不是概率，只是概率的估计.

频率又称统计型概率。

频率具有稳定性：在一定条件下，频率稳定于概率。

频率稳定性的广泛应用

(1) 英语中字母出现频率具有稳定性, E,T,O频率高，J,Q,Z频率低；电脑键盘设计、信息的编码(常用字母用较短的码)；

字母	空格	E	T	O	A	N	I	R	S
频率	0.2	0.105	0.072	0.065 4	0.063	0.059	0.055	0.054	0.052
字母	H	D	L	C	F	U	M	P	Y
频率	0.047	0.035	0.029	0.023	0.022 5	0.022 5	0.021	0.017 5	0.012
字母	W	G	B	V	K	X	J	Q	Z
频率	0.012	0.011	0.010 5	0.008	0.003	0.002	0.001	0.001	0.001

英文字母使用频率统计表

(2) 产品抽样检查，某电视节目收视率的调查, 衣服和用具总在同样部位相似的方式破损等, 英国生物统计学家高尔顿设计的高尔顿板；

(3) Monte Carlo仿真。

频率具有稳定性：在一定条件下，频率稳定于概率。

频率稳定性的广泛应用

另一个验证频率稳定性的著名试验是由英国生物统计学家高尔顿(Galton)设计的. 它的试验模型如图 1.1.1 所示.

自上端放入一小球，任其自由下落，在下落过程中当小球碰到钉子时，从左边落下与从右边落下的机会相等. 碰到下一排钉子时又是如此. 最后落入底板中的某一格子. 因此，任意放入一球，则此球落入哪一个格子，预先难以确定. 但是实验证明，如放入大量小球，则其最后所呈现的曲线，几乎总是一样的. 也就是说，小球落入各个格子的频率十分稳定. 这个试验模型称为高尔顿板. 试验中呈现出来的规律性，在学习第五章极限定理之后，就会有更深刻的理解.

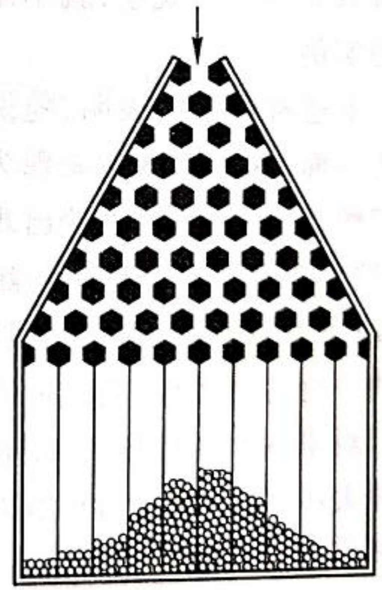


图 1.1.1 高尔顿板

· 5 ·

蒙特卡洛方法

Monte Carlo方法是计算机模拟的基础，它的名字来源于世界著名的赌城——摩纳哥的蒙特卡洛， 其历史起源于 1777 年法国科学家蒲丰提出的一种计算圆周π 的方法——随机投针法，即著名的蒲丰投针问题。

“在平面上画有一组间距为d的平行线，将一根长度为l(l<d)的针任意掷在这个平面上，求此针与平行线中任一条相交的概率。”

Monte Carlo方法的基本思想是首先建立一个概率模型，使所求问题的解正好是该模型的参数或其他有关的特征量. 然后通过模拟一统计试验， 即多次随机抽样试验 (确定 m和 n)，统计出某事件发生的百分比. 只要试验次数很大，该百分比便近似于事件发生的概率.

这实际上就是概率的统计定义.

利用建立的概率模型，求出要估计的参数.

蒙特卡洛方法适用范围很广泛，它既能求解确定性的问题，也能求解随机性的问题以及 科学研究中的理论问题.

例如利用蒙特卡洛方法可以近似地计算定积分，即产生数值积分问题.

蒲 丰 投 针 问 题

18世纪，法国数学家蒲丰和勒可莱尔提出的“投针问题”，记载于布丰1777年出版的著作中：“在平面上画有一组间距为d的平行线，将一根长度为l（l<d）的针任意掷在这个平面上，求此针与平行线中任一条相交的概率。”布丰本人证明了，这个概率是p=2l/（πd) π为圆周率.

利用这个公式可以用**概率的方法得到圆周率的近似值**。

实验者	年代	投掷次数	相交次数	圆周率估计值
沃尔夫	1850	5000	2531	3.1596
史密斯	1855	3204	1219	3.1554
德摩根	1680	600	383	3.137
福克斯	1884	1030	489	3.1595
拉泽里尼	1901	3408	1808	3.1415929
赖纳	1925	2520	859	3.1795

蒲丰投针实验是第一个用几何形式表达概率问题的例子，他首次使用随机实验处理确定性数学问题，为概率论的发展起到一定的推动作用。

蒲丰投针证明方法

下面就是一个简单而巧妙的证明。找一根铁丝弯成一个圆圈，使其直径恰恰等于平行线间的距离d。可以想象得到，对于这样的圆圈来说，不管怎么扔下，都将和平行线有两个交点。因此，如果圆圈扔下的次数为n次，那么相交的交点总数必为2n。现在设想把圆圈拉直，变成一条长为πd的铁丝。显然，这样的铁丝扔下时与平行线相交的情形要比圆圈复杂些，可能有4个交点，3个交点，2个交点，1个交点，甚至于都不相交。由于圆圈和直线的长度同为πd，根据机会均等的原理，当它们投掷次数较多，且相等时，两者与平行线组交点的总数期望也是一样的。这就是说，当长为πd的铁丝扔下n次时，与平行线相交的交点总数应大致为2n。

现在转而讨论铁丝长为l的情形。当投掷次数n增大的时候，这种铁丝跟平行线相交的交点总数m应当与长度l成正比，因而有：m=kl，式中k是比例系数。为了求出k来，只需注意到，对于l=πd的特殊情形，有m=2n。于是求得k=（2n)/（πd）。代入前式就有：m≈（2ln)/（πd）从而π≈（2ln)/(dm)

例6 故事发生在十七世纪中叶，法国贵族德·美黑热中于赌博，经常遇到赌资分配问题。他曾写信向当时法国的大数学家Pascal 请教问题：

假如一场比赛中胜 6 局才算赢，两个赌徒在一人胜五局，另一人胜两局的情况下中断赌博，如何分配赌金？

有两种方案： 5:2； 15:1；

Pascal 和当时第一流的数学家 Fermat 一起研究了此问题，得到正确的解答：15 : 1。

思考：如何得到？

三：古典概率

古典概率的起源 → 赌徒分赌金问题

定义：设E是一个随机试验,若它满足以下两个条件：

（1）仅有有限多个基本事件；（有限性）

(2)每个基本事件发生的可能性相等。（等可能性）

则称E <u>古典概型的试验</u>。

例如：→ 掷骰子试验　抛硬币试验等

例1 抛一颗均匀的骰子，观察其出现的点数情况。

因为该试验的基本事件有6个：

$\{\omega_i\}$={出现的点数为i }　i=1, 2, .., 6

而且基本事件$\{\omega_1\}, \{\omega_2\}, \ldots \{\omega_6\}$发生的可能性相等。

这是一个古典概型的随机试验。

E1: 抛一枚质量分布不均匀的硬币，观察其出现正面H和反面T的情况。

因为该试验的基本事件只有两个:

$\{\omega_1\}$={出现正面H}, $\{\omega_2\}$={出现反面T}。

但基本事件$\{\omega_1\}$、$\{\omega_2\}$发生的可能性不相等。

E2：仪器上某种型号的电子元件使用时间已达30小时，测该元件还能使用多少小时？

该试验不是古典概型的随机试验，因为它的样本空间有无数多个样本点。

E1, E2都不是一个古典概型的随机试验。

三：古典概率

设*E*是一个随机试验(重复性,明确性和不知预知性),

若它满足以下两个条件：

（1）仅有有限多个基本事件；（有限性）

(2)每个基本事件发生的可能性相等。（等可能性）

则称*E* 古典概型的试验。

定义：设试验E为古典概型试验，A是E的任一事件，则有：

$$P(A)=\frac{A\text{所含的基本事件个数}}{\text{基本事件总数}}$$

关键词：试验、可能

判断样本空间与事件分别所包含的基本事件数目.

$$=\frac{A\text{所含样本点的数目}}{\text{样本空间的样本点总数}}$$

所确定的概率称为事件A的古典概率。

注1：法国数学家拉普拉斯在1812年把它作为概率的一般定义，现在通常称为概率的古典定义。

注2：在古典概率的计算中常用到排列组合的知识，如乘法原理、加法原理等等。

古典概率的计算中常见的类型：

- (1) 随机取数：例1.2.2, 1.2.3, 1.2.8
- (2) 分(住)房，生日问题：例1.2.4
- (3) 抽球：例1.2.5(重点)

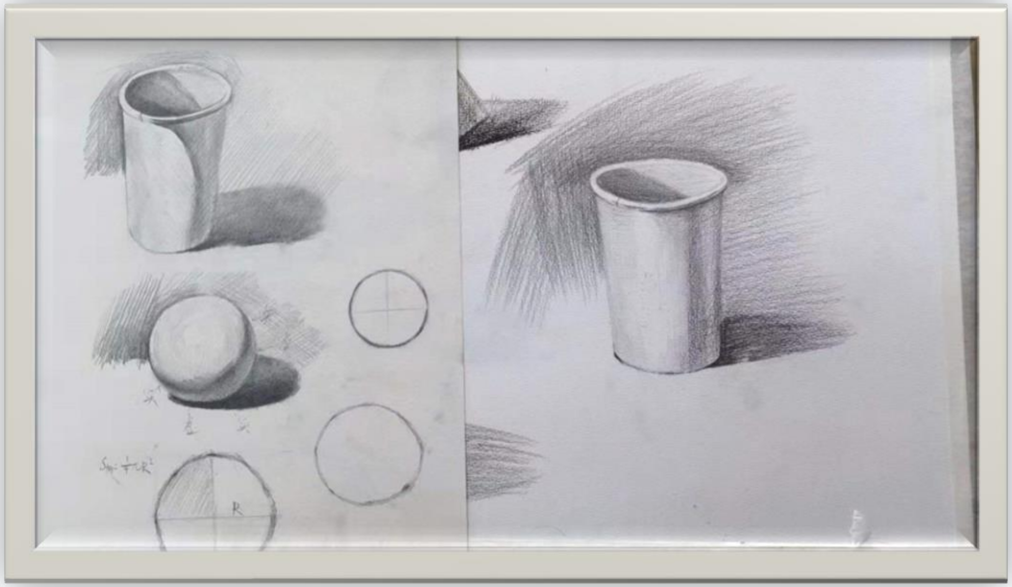
  注:古典概率的大部分问题都能形象化地用摸球模型来描述。
- (4) 配对(很少涉及)

注1: 种水稻地块的调查，某电视节目收视率的调查，某种疾病的抽查等可以用抽象化的摸球问题来描述。

注2：古典概率的计算在产品质量抽样检查等实际问题及理论物理的研究中都有重要应用。

用样本空间求概率　赌金分配　鸽笼问题　摸彩试验

本次课的重点是：事件关系的判断，事件运算律熟练运用，概率定义的理解，熟练古典概率的计算问题。



下次课内容：

要讲到 § 1.3 条件概率开头，其中包括剩下的古典概率计算，古典概率的性质，柯氏公里定义及性质，条件概率的定义