随机模型

浙江大学 谈之奕

随机性

- 必然性与偶然性：揭示客观事物发生、发展和灭亡趋势的一对范畴
  - 必然性（necessity）：客观事物联系和发展中一定要发生的、确定不移的趋势
  - 偶然性（contingency）：客观事物联系发展中并非确定发生的，可以这样出现也可以那样出现的不确定趋势
- 随机性（randomness）：偶然性的一种形式
  - 可重复性：事件可以在基本相同的条件下重复进行。只有单一的偶然过程而无法判定它的可重复性则不称为随机事件
  - 多样性：在基本相同条件下某事件可能以多种方式表现出来，事先不能确定它以何种特定方式发生。只有唯一可能性的过程不是随机事件
  - 概率性：事先可以预见该事件以各种方式出现的所有可能性，预见它以某种特定方式出现的概率。在重复发生时没有确定概率的现象不是同一过程的随机事件

概率论

- 概率论（probability theory）
  - 研究随机现象数量规律的数学分支

Gerolamo Cardano
（1501 –1576）
意大利数学家

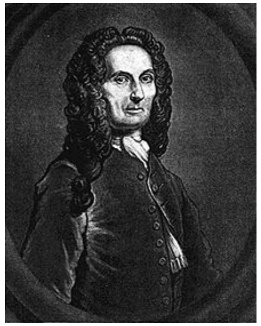
Blaise Pascal
（1623 –1662）
法国数学家、物理学家

Pierre de Fermat
(1607-1665)
法国数学家

Christiaan Huygens
（1629 –1695）
荷兰数学家、物理学家、天文学家

博彩问题 ⟶ 组合计算 ⟶ 分析方法 ⟶ 公理化体系

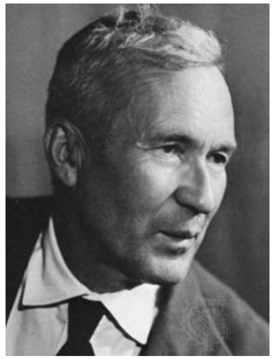
Jacob Bernoulli
（1655-1705）
瑞士数学家



Abraham de Moivre
（1667 –1754）
法国数学家

Pierre-Simon Laplace
（1749 –1827）
法国数学家

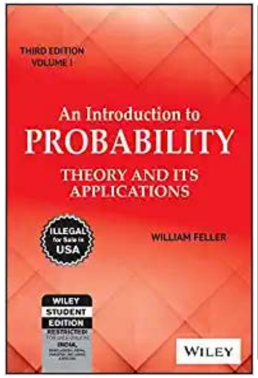
Chebyshev PL
Markov AA
Lyapunov AM
Khinchin AY
俄罗斯数学家

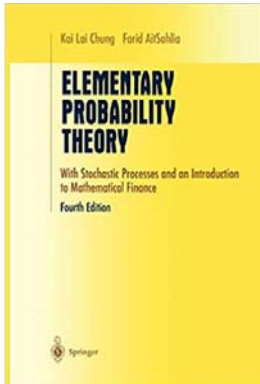


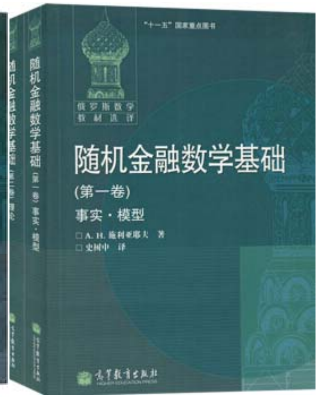
Andrey Nikolaevich Kolmogorov
（1903 –1987）
苏联数学家

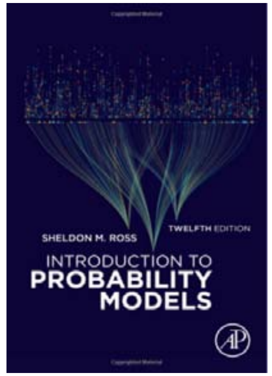
随机模型

数学
建模
MATH T









William Feller
(1906-1970)
克罗地亚裔美籍数学家

Feller W, *An Introduction to Probability Theory and Its Applications*, Wiley, 1958 (中译本：概率论及其应用（全两卷），胡迪鹤、郑元禄译，人民邮电出版社，2021年)

Chung KL, AitSahlia F, *Elementary Probability Theory: With Stochastic Processes and an Introduction to Mathematical Finance* (4th), Springer, 2003

A. H. 施利亚耶夫，概率（全两卷），周概容译，高等教育出版社，2008

A. H. 施利亚耶夫，随机金融数学基础（全两卷），史树中译，高等教育出版社，2013

Ross SM, *Introduction to Probability Models* (12th), Academic Press, 2019. (中译本：应用随机过程——概率模型导论（第11版），龚光鲁译，人民邮电出版社，**2016**年)

事件与概率

数学
建模
MATH T

- 随机试验
  - 随机现象的实现和对它的观察称为**随机试验**（random experiment）
  - 随机试验所有可能结果的集合称为该试验的**样本空间**（sample space）
    - 样本空间中的每一个元素称为**基本事件**，样本空间的子集称为**事件**（event）
  - 事件的**概率**（probability）是衡量该事件发生的可能性大小的度量
- **古典概型**（classical models of probability）
  - 样本空间中只有有限个基本事件，每个基本事件发生的可能性相等
    - 若样本空间 S 和事件 A 中分别包含 $|S|$ 个和 $|A|$ 个基本事件，事件 A 发生的概率为 $P(A)=\dfrac{|A|}{|S|}$
- 条件概率
  - 设 A,B 为事件，其中 $P(A)>0$，则 $P(B\mid A)=\dfrac{P(AB)}{P(A)}$ 称为在已知 A 发生的情况下，B 发生的**条件概率**（Conditional probability）
    - **概率的乘法原理** $P(AB)=P(A)\cdot P(B\mid A)=P(B)\cdot P(A\mid B)$

依次抛掷两枚硬币

$$S=\left\{\begin{matrix}(\mathrm{H},\mathrm{H}),(\mathrm{H},\mathrm{T}),\\(\mathrm{T},\mathrm{H}),(\mathrm{T},\mathrm{T})\end{matrix}\right\}$$

有一枚硬币出现正面

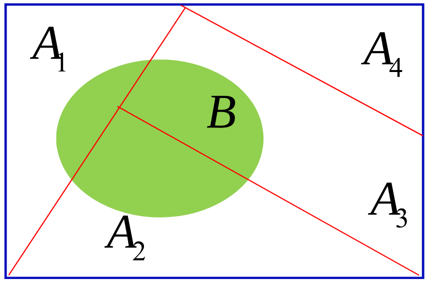
$$A=\{(\mathrm{H},\mathrm{T}),(\mathrm{T},\mathrm{H})\}$$

$$P(A)=\frac{1}{2}$$

全概率公式与Bayes公式

数学建模
MATH T

- 分划
  - 事件 $A_1, A_2, \cdots, A_n$ 称为样本空间 S 的一个分划，若对任意 $i \neq j$，$A_i \cap A_j = \varnothing$ 且 $\bigcup_{i=1}^{n} A_i = S$



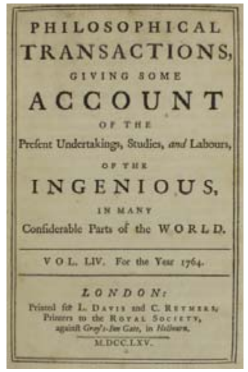
- 全概率公式（law of total probability）
  - 设事件 $A_1, A_2, \cdots, A_n$ 是样本空间的一个分划，且对任意 i，$P(A_i) > 0$，则对任意事件 B，

$$P(B) = \sum_{i=1}^{n} P(BA_i) = \sum_{i=1}^{n} P(B|A_i)P(A_i)$$

- Bayes公式
  - 设事件 $A_1, A_2, \cdots, A_n$ 是样本空间的一个分划，且对任意 i，$P(A_i) > 0$，则对任意事件 B，只要 $P(B) > 0$

$$P(A_i|B) = \frac{P(A_iB)}{P(B)} = \frac{P(B|A_i)P(A_i)}{\sum_{i=1}^{n} P(B|A_i)P(A_i)}$$

Thomas Bayes
（1701 –1761）
英国统计学家



PHILOSOPHICAL
TRANSACTIONS,
GIVING SOME
ACCOUNT
OF THE
Prefent Undertakings, Studies, and Labours,
OF THE
INGENIOUS,
IN MANY
Confiderable Parts of the WORLD.
VOL. LIV. For the Year 1764.
LONDON:
MDCCLXV.

Bayes T, An essay towards solving a problem in the doctrine of chances. *Philosophical Transactions of the Royal Society of London*, 53, 370–418, 1764.

疾病检测

数学
建模
MATH T

- 疾病检测
  - 疾病检测方法的性能指标
    - 灵敏度（sensitivity） $p=P(B|A)$：患病者被检测为阳性（positive）的概率
    - 特异度（specificity） $q=P(\overline{B}|\overline{A})$：未患病者被检测为阴性（negative）的概率
  - 被检测为阳性的情况下患病的概率
    - 记 A 为患病，B 为检测结果为阳性
    - 设疾病的发病率为 r

$$P(A|B)=\frac{P(B|A)P(A)}{P(B|A)P(A)+P(B|\overline{A})P(\overline{A})}$$

$$=\frac{pr}{pr+(1-q)(1-r)}$$

$$r=0.005,\ p=0.95,\ q=0.99$$

$$P(A|B)=\frac{95}{294}\approx 0.323$$

Antigen-detection in the diagnosis of SARS-CoV-2 infection (Interim guidance), WHO, 2021.10.6

WHO recommends the use of Ag-RDTs that meet minimum performance requirements of $\geq$ 80% sensitivity and $\geq$ 97% specificity. Ag-RDTs are less sensitive than NAAT, particularly in asymptomatic populations, but careful selection of cohorts for testing can mitigate this limitation.

International Journal of Infectious Diseases 109 (2021) 118–122

Contents lists available at ScienceDirect

International Journal of Infectious Diseases

journal homepage: www.elsevier.com/locate/ijid

Original Article

Diagnostic accuracy of a SARS-CoV-2 rapid antigen test in real-life clinical settings

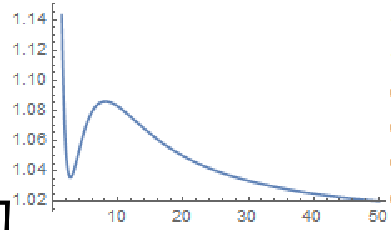
Sabrina Jegerlehner, MD, MSc[1,*], Franziska Suter-Riniker, PhD[2,*], Philipp Jent, MD[3], Pascal Bittel, PhD[2], Michael Nagler, MD, PhD, MSc[4,**]

The overall sensitivity of the Roche/SD Biosensor rapid antigen test was 65.3% (95% confidence interval [CI] 56.8–73.1); the specificity was 99.9 (95% CI 99.5–100.0). The number of false-

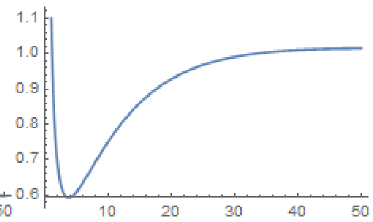
概率群试

- 两阶段群试
  - 将 n 人的样本混合后检测。若结果为阴性，说明这 n 人均未感染。若结果为阳性，说明这 n 人中至少有一人已感染。此时逐个检测每个人样本
  - 数量为 N 的人群中感染的概率为 p，寻找最优的检测方案
- 检测次数
  - 混合样本阴性概率为 $(1-p)^n$，总检测次数为 1
  - 混合样本阳性概率为 $1-(1-p)^n$，总检测次数为 $n+1$
  - 检测次数的数学期望为 $1\cdot(1-p)^n+(n+1)\left(1-(1-p)^n\right)$
  - 平均检测次数的数学期望为

$$\frac{1}{n}\left(1\cdot(1-p)^n+(n+1)\left(1-(1-p)^n\right)\right)$$



$p=\frac{1}{3}$



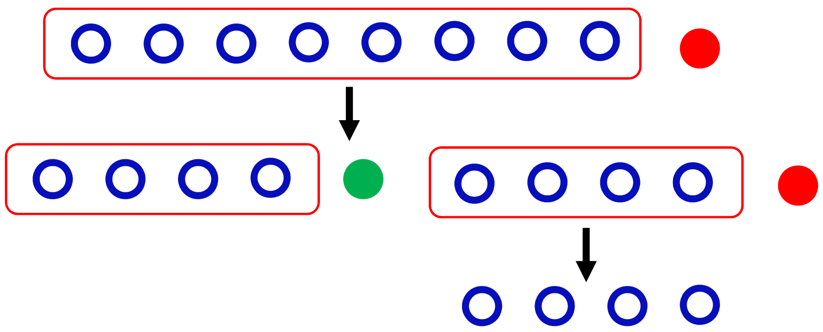
$p=\frac{1}{10}$

n	**1**	**3**	**4**
$1-p$	**0.6934**	**0.8761**	**0.9344**
n	**5**	**6**	**10**
$1-p$	**0.9589**	**0.9719**	**0.9899**
n	**20**	**50**	**100**
$1-p$	**0.9975**	**0.9975**	**0.9999**

概率群试

数学建模
MATH T

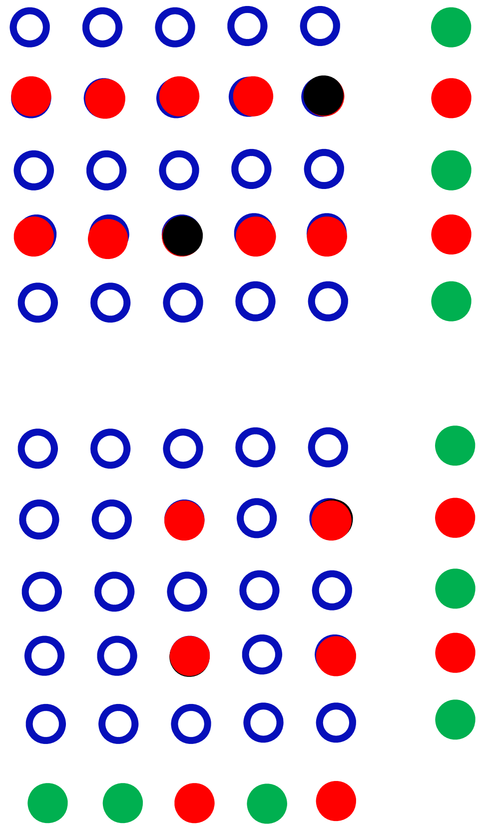
- 概率群试的改进方案
  - 矩阵检测方案
  - 三阶段检测方案
  - 二分检测方案
- 群试方案的选择
  - 平均检测次数
  - 检测阶段数
  - 每人最大检测次数
  - 每组最多样本数
  - 方案的可操作性
  - 检测的灵敏度与特异度



$$(1-p)^n\cdot 1+(1-p)^{\frac{n}{2}}\left(1-(1-p)^{\frac{n}{2}}\right)\left(\frac{n}{2}+2\right)+\left(1-(1-p)^{\frac{n}{2}}\right)(1-p)^{\frac{n}{2}}\left(\frac{n}{2}+3\right)+\left(1-(1-p)^{\frac{n}{2}}\right)^2(n+3)$$

$$1+\frac{3}{n}-\left(1+\frac{1}{n}\right)(1-p)^{\frac{n}{2}}-\frac{1}{n}(1-p)^n$$

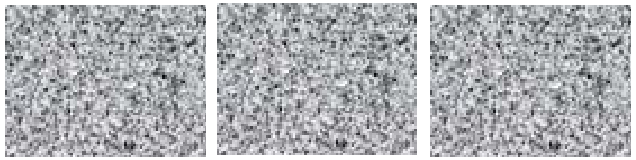
$p=0.01$	分组大小	平均次数
两阶段	11	0.1956
三阶段	15	0.1579

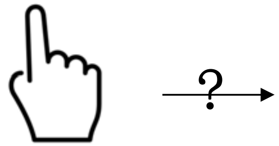


Monty Hall问题

数学
建模
MATH T

- The Monty Hall Problem
  - 舞台上有三扇道具门，其中一扇门后置有一辆汽车，另两扇门后各置有一头山羊。竞猜者可任选其中一扇门并获赠门后物品
  - 竞猜者选择了其中一扇门后，主持人打开了另两扇门中的一扇，门后面是一头山羊
    - 主持人知道汽车所在位置。他打开的门既不是竞猜者选择的，也不是后置汽车的。若有两扇门符合以上要求，他以相同概率选择其中一扇
  - 主持人允许竞猜者改变之前的选择，竞猜者为增加获得汽车的可能性，是否应该改变当前的选择





Monty Hall（1921 –2017），美国电视节目主持人、制片人，长期担任综艺节目Let’s Make a Deal主持人

Monty Hall问题，最早由美国统计学家Steve Selvin以通讯形式在1975年的The American Statistician期刊上提出并解决。1990年，专栏作家Marilyn vos Savant在美国Parade杂志专栏“Ask Marilyn”中回答了读者提出的这一问题，引起广泛讨论，成为概率论中的著名问题

Monty Hall问题

- The Monty Hall Problem
  - 假设竞猜者初次选择1号门，汽车位于1、2、3号门后的概率相同
  - 若竞猜者不改变选择，则获得汽车的概率为 $\frac{1}{3}$
  - 若竞猜者改变选择，则获得汽车的概率为 $\frac{2}{3}$

<table>
<tr><th rowspan="2">竞猜者初次选择的门</th><th colspan="2">汽车位置</th><th rowspan="2">主持人打开的门</th><th rowspan="2">竞猜者再次选择的门</th><th rowspan="2">获赠物品</th><th rowspan="2">竞猜者再次选择的门</th><th rowspan="2">获赠物品</th></tr>
<tr><th>门</th><th>概率</th></tr>
<tr><td rowspan="4">1</td><td rowspan="2">1</td><td rowspan="2">1/3</td><td>2</td><td rowspan="2">1</td><td rowspan="2">汽车</td><td>3</td><td>山羊</td></tr>
<tr><td>3</td><td>2</td><td>山羊</td></tr>
<tr><td>2</td><td>1/3</td><td>3</td><td>1</td><td>山羊</td><td>2</td><td>汽车</td></tr>
<tr><td>3</td><td>1/3</td><td>2</td><td>1</td><td>山羊</td><td>3</td><td>汽车</td></tr>
</table>

Monty Hall问题

数学
建模
MATH T

- The Monty Hall Problem
  - 假设竞猜者初次选择1号门
  - 记 C_i 为事件“汽车位于 i 号门后”，$P(C_1)=P(C_2)=P(C_3)=\frac{1}{3}$
  - 假设主持人打开2号门。记 M 为事件“主持人打开2号门”
    - $P(M|C_1)=\frac{1}{2}$，$P(M|C_2)=0$，$P(M|C_3)=1$
  - 若竞猜者不改变选择，获得汽车的概率为

$$P(C_1|M)=\frac{P(M|C_1)P(C_1)}{P(M|C_1)P(C_1)+P(M|C_2)P(C_2)+P(M|C_3)P(C_3)}=\frac{\frac{1}{2}\cdot\frac{1}{3}}{\frac{1}{2}\cdot\frac{1}{3}+0\cdot\frac{1}{3}+1\cdot\frac{1}{3}}=\frac{1}{3}$$

  - 若竞猜者改变选择，选择3号门，获得汽车的概率为

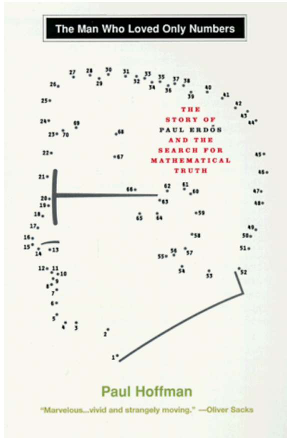
$$P(C_3|M)=\frac{P(M|C_3)P(C_3)}{P(M|C_1)P(C_1)+P(M|C_2)P(C_2)+P(M|C_3)P(C_3)}=\frac{1\cdot\frac{1}{3}}{\frac{1}{2}\cdot\frac{1}{3}+0\cdot\frac{1}{3}+1\cdot\frac{1}{3}}=\frac{2}{3}$$

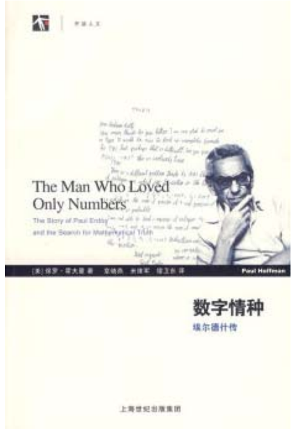
Monty Hall问题

6 GETTING THE GOAT 得了一只羊

瓦兹索尼把有关蒙迪·霍尔难题的情况告诉了埃尔德什。“我告诉埃尔德什答案是换一扇门，”瓦兹索尼说道，“然后满以为可以转到下一个话题。可令我吃惊的是，埃尔德什说，‘不，那不可能。换不换门应该没什么不同。’这时我对提出这个问题感到后悔，因为我有这样的经历，即人们会因这个答案而变得很激动，从而使得我们的谈话不欢而散。在根本就不可能从容了结此事的情况下，我给他画了一个我在大学的《量化管理技术》上学到的树形分析结果。”瓦兹索尼画出分析树，与萨万特曾经制作的那张罗列可能结果的表格不无相似之处，但这并没有说服他。“真是没办法，”瓦兹索尼说道。“我

瓦兹索尼在他的个人电脑上用蒙特卡罗法模拟蒙迪·霍尔难题。从来都不怎么用电脑的埃尔德什，现在却看着电脑随机地在“换门”和“不换门”之间作出选择。数百次的试验结果证明，换门时赢的概率是不换门的2倍，这时埃尔德什才不得不承认自己错了。但这种模拟并不比用计算机证明四色定理更让人满意：它没有揭示为什么换一下门会更好一些。埃尔德什对瓦兹索尼的解释还是不够满意。他准备走了。





Hoffman P, *The Man Who Loved Only Numbers: The Story of Paul Erdos and the Search for Mathematical Truth*, Hyperion, 1999 (中译本：数字情种——埃尔德什传，章晓燕、米绪军、缪卫东译，上海世纪出版集团，2009)

Paul Erdős
（1913 –1996）
匈牙利数学家
1983年Wolf奖
得主

Monty Hall问题

- 被蒙在鼓里的主持人
  - 主持人不知道汽车所在位置。他在竞猜者选择后以相同概率打开另两扇门中的一扇，后面是一头山羊
  - 假设竞猜者初次选择1号门，主持人打开2号门
    - 记 M 为事件“主持人打开2号门，门后是一头山羊”
    - $P(M|C_1)=\frac{1}{2}$, $P(M|C_2)=0$, $P(M|C_3)=\frac{1}{2}$
  - 若竞猜者不改变选择，获得汽车的概率为

$$P(C_1|M)=\frac{P(M|C_1)P(C_1)}{P(M|C_1)P(C_1)+P(M|C_2)P(C_2)+P(M|C_3)P(C_3)}=\frac{\frac{1}{2}\cdot\frac{1}{3}}{\frac{1}{2}\cdot\frac{1}{3}+0\cdot\frac{1}{3}+\frac{1}{2}\cdot\frac{1}{3}}=\frac{1}{2}$$

  - 若竞猜者改变选择，选择3号门，获得汽车的概率为

$$P(C_3|M)=\frac{P(M|C_3)P(C_3)}{P(M|C_1)P(C_1)+P(M|C_2)P(C_2)+P(M|C_3)P(C_3)}=\frac{\frac{1}{2}\cdot\frac{1}{3}}{\frac{1}{2}\cdot\frac{1}{3}+0\cdot\frac{1}{3}+\frac{1}{2}\cdot\frac{1}{3}}=\frac{1}{2}$$

Monty Hall问题

数学
建模
MATH T

- Monty Hall问题的变形
  - 多扇门的Monty Hall问题
    - 有 n 扇道具门，其中一扇门后置有一辆汽车，其他 $n-1$扇门后各置有一头山羊
    - 当至少有三扇门还未打开时，竞猜者选择其中一扇未打开的门，主持人以相同概率打开竞猜者未选择且后面是山羊的门中的任意一扇，并允许竞猜者改变之前的选择。继续上述过程直至只有两扇门还未打开
    - 在游戏进行过程中每次允许竞猜者改变选择时，竞猜者应采取怎样的策略
  - 多扇门、多辆车的Monty Hall问题
    - 有 n 扇道具门，其中 k 扇门后各置有一辆汽车，其他 $n-k$扇门后各置有一头山羊
    - 竞猜者选择其中一扇门后，主持人以相同概率打开了其他 $n-1$扇门中的 m扇，$1\le m\le n-2$，其中 j 扇门后各有一辆汽车，$m-j$ 扇门后各有一头山羊
    - 主持人允许竞猜者改变之前的选择，竞猜者是否应该改变当前的选择
  - 多扇门、多种奖品的Monty Hall问题
    - 有 n 扇道具门，其中 n_i 扇门后各置有价值为 v_i 的奖品，$i=1,\cdots,m$
    - 竞猜者选择其中一扇门后，主持人以相同概率打开了其他 $n-1$ 扇门中的一扇，并允许竞猜者改变之前的选择。竞猜者看到打开的门后的奖品后，是否应该改变当前的选择

随机变量

数学
建模
MATH T

- 随机变量（random variable）
  - 定义在样本空间上的实值函数。是随机试验结果的量的表示
  - 只能取有限个或可数个数值的随机变量称为离散型随机变量。可能取值为一个区间内所有实数的随机变量称为连续型随机变量
- 离散型分布
  - 离散型随机变量取值的概率规律称为离散型分布
  - 设离散型随机变量 X 所有可能取值为 $x_1, x_2, \cdots$，对任意的 i，$P(X = x_i) = p_i$，其中 $\{p_i\}$ 满足
    - $p_i \geq 0, i = 1, 2, \cdots$
    - $\sum_{i=1}^{\infty} p_i = 1$

离散分布列 $\begin{pmatrix} x_1 & x_2 & x_3 & \cdots \\ p_1 & p_2 & p_3 & \cdots \end{pmatrix}$

抛掷两枚硬币

$$S = \begin{Bmatrix} (\mathrm{H},\mathrm{H}), (\mathrm{H},\mathrm{T}), \\ (\mathrm{T},\mathrm{H}), (\mathrm{T},\mathrm{T}) \end{Bmatrix}$$

X：出现正面的硬币枚数

X	**0**	**1**	**2**
事件	**(T,T)**	**(H,T)** **(T,H)**	**(H,H)**
分布	$\frac{1}{4}$	$\frac{1}{2}$	$\frac{1}{4}$

离散型分布

数学
建模
MATH T

- Bernoulli试验
  - 独立、重复地多次进行同一项随机试验，每次试验的结果只有两种可能
  - 设一次随机试验的结果有“成功”和“失败”两种，“成功”的概率为 p

二项分布 (binomial distribution)	n 次Bernoulli试验中，成功的次数	$P(X=i)=\binom{n}{i}p^i(1-p)^{n-i}, i=0,1,\cdots,n$
几何分布 (geometric distribution)	进行Bernoulli试验，首次出现成功所需的试验次数	$P(X=i)=(1-p)^{i-1}p, i=1,2,\cdots$
负二项分布 (negative binomial distribution)	进行Bernoulli试验，累计出现 r 次成功所需的试验次数	$P(X=i)=\binom{i-1}{r-1}p^r(1-p)^{i-r}, i=r,r+1,\cdots$
超几何分布 (hypergeometric distribution)	已知 N 件产品中有 M 件不合格品。随机抽取 s 件产品中的不合格数	$P(X=i)=\dfrac{\binom{M}{i}\binom{N-M}{s-i}}{\binom{N}{s}}, i=0,1,\cdots,s$

数学期望

数学
建模
MATH T

- 几何分布的数学期望
  - $P(X=i)=(1-p)^{i-1}p, i=1,2,\cdots$
  - 利用离散分布列
    - $E(X)=\sum\limits_{i=1}^{\infty} i\cdot P(X=i)=\sum\limits_{i=1}^{\infty} i(1-p)^{i-1}p$ $\qquad i=1+(i-1)$

$$=\sum_{i=1}^{\infty}(1-p)^{i-1}p+\sum_{i=1}^{\infty}(i-1)(1-p)^{i-1}p$$

$$=1+(1-p)\sum_{i=2}^{\infty}(i-1)(1-p)^{i-2}p$$

$$=1+(1-p)\sum_{l=1}^{\infty}l(1-p)^{l-1}p=1+(1-p)E(X)\Rightarrow E(X)=\frac{1}{p}$$

  - 利用条件期望
    - 定义随机变量 $Y=\begin{cases}1 & \text{第一次试验成功}\\ 0 & \text{第一次试验失败}\end{cases}$, $P(Y=1)=p$
    - $E(X|Y=1)=1, E(X|Y=0)=1+E(X)$
    - $E(X)=E(X|Y=1)P(Y=1)+E(X|Y=0)P(Y=0)$

$$=p\cdot 1+(1-p)(1+E(X))=1+(1-p)E(X)$$

对随机变量 Y 的任一确定值 y，随机变量 X 的所有可能取值 $x_1,x_2,\cdots$，连同条件概率 $P(X=x_i|Y=y)$ 构成离散分布列

在 $Y=y$ 条件下，X 的条件期望为

$$E(X|Y=y)=\sum_{i=1}^{\infty}x_iP(X=x_i|Y=y)$$

当 Y 的取值变动时，$E(X|Y)$ 也为一随机变量，且 $E(E(X|Y))=E(X)$

$$E(E(X|Y))=\sum_{j=1}^{\infty}E(X|Y=y_j)P(Y=y_j)$$

$$=\sum_{j=1}^{\infty}\sum_{i=1}^{\infty}x_iP(X=x_i|Y=y_j)P(Y=y_j)$$

$$=\sum_{j=1}^{\infty}\sum_{i=1}^{\infty}x_iP(X=x_i,Y=y_j)=\sum_{i=1}^{\infty}x_iP(X=x_i)=E(X)$$

赠券收集问题

数学
建模
MATH T

- 赠券收集问题（Coupon collector's problem）
  - 一套赠券共有 N 种，商家在每件商品中随机放入一张赠券。集齐全套赠券平均需购买多少件商品
    - 假设每件商品中放入各种赠券的概率相同
  - 定义随机变量 X 为“集齐全套赠券需购买的商品件数”，$E(X)=\sum_{i=1}^{\infty} i\cdot P(X=i)$
    - 记 B_i 为事件“购买 i 件商品后集齐全套赠券”
      - 记 A_i^j 为事件“购买 i 件商品后收集到第 j 种赠券”

$$P\left(A_i^j\right)=1-P\left(\overline{A_i^j}\right)=1-\left(1-\frac{1}{N}\right)^i$$

$$P\left(B_i\right)=P\left(A_i^1 A_i^2 \cdots A_i^N\right)$$

西湖十景（图片来自网络）

苏堤春晓	花港观鱼	双峰插云	雷峰夕照	南屏晚钟
柳浪闻莺	三潭印月	曲院风荷	平湖秋月	断桥残雪

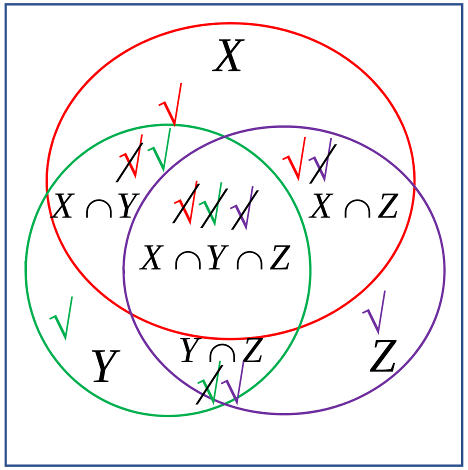
概率的加法原理

数学建模

MATH T

- De Morgan定律
  - 设 $A_1, A_2, \cdots, A_n$ 为集合 S 的 n 个子集，则
    - $\overline{A_1 \cup A_2 \cup \cdots \cup A_n} = \overline{A_1} \cap \overline{A_2} \cap \cdots \cap \overline{A_n}$
    - $\overline{A_1 \cap A_2 \cap \cdots \cap A_n} = \overline{A_1} \cup \overline{A_2} \cup \cdots \cup \overline{A_n}$
- 容斥原理（inclusion–exclusion principle）
  - 设 $A_1, A_2, \cdots, A_n$ 为集合 S 的 n 个有限子集，则

$$\left|A_1 \cup A_2 \cup \cdots \cup A_n\right| = \sum_{i=1}^{n} |A_i| - \sum_{1 \le i < j \le n} \left|A_i \cap A_j\right| + \sum_{1 \le i < j < k \le n} \left|A_i \cap A_j \cap A_k\right| + \cdots + (-1)^{n-1} \left|A_1 \cap A_2 \cap \cdots \cap A_n\right|$$



$$\begin{aligned}|X \cup Y \cup Z| &= (|X| + |Y| + |Z|) \\ &\quad - (|X \cap Y| + |Y \cap Z| + |X \cap Z|) \\ &\quad + |X \cap Y \cap Z|\end{aligned}$$

- 概率的加法原理
  - 设 $A_1, A_2, \cdots, A_n$ 为 n 个事件，则

$$P\left(A_1 \cup A_2 \cup \cdots \cup A_n\right) = \sum_{i=1}^{n} P(A_i) - \sum_{1 \le i < j \le n} P\left(A_i \cap A_j\right) + \sum_{1 \le i < j < k \le n} P\left(A_i \cap A_j \cap A_k\right)$$

$$+ \cdots + (-1)^{n-1} P\left(A_1 \cap A_2 \cap \cdots \cap A_n\right)$$

赠券收集问题

数学
建模
MATH T

- 赠券收集问题
  - $P\left(B_i\right)=P\left(A_i^1A_i^2\cdots A_i^N\right)=1-P\left(\overline{A_i^1A_i^2\cdots A_i^N}\right)=1-P\left(\overline{A_i^1}\cup\overline{A_i^2}\cup\cdots\cup\overline{A_i^N}\right)$
  - $P\left(\overline{A_i^1}\cup\overline{A_i^2}\cup\cdots\cup\overline{A_i^N}\right)=\sum_{j=1}^{N}P\left(\overline{A_i^j}\right)-\sum_{1\le j<k\le N}P\left(\overline{A_i^j}\cap\overline{A_i^k}\right)$ $\binom{N}{2}$

    $+\cdots+\sum_{1\le j_1<j_2<\cdots<j_l\le N}P\left(\overline{A_i^{j_1}}\cap\overline{A_i^{j_2}}\cap\cdots\cap\overline{A_i^{j_l}}\right)$ $\binom{N}{l}$

    $+\cdots+(-1)^{N-1}P\left(\overline{A_i^1}\cap\overline{A_i^2}\cap\cdots\cap\overline{A_i^N}\right)$
    - 对任意 $1\le j<k\le N$，$P\left(\overline{A_i^j}\cap\overline{A_i^k}\right)=\left(1-\frac{2}{N}\right)^i$
    - 对任意 $1\le l\le N$，$1\le j_1<j_2<\cdots<j_l\le N$，$P\left(\overline{A_i^{j_1}}\cap\overline{A_i^{j_2}}\cap\cdots\cap\overline{A_i^{j_l}}\right)=\left(1-\frac{l}{N}\right)^i$
  - $P\left(B_i\right)=1-\sum_{l=1}^{N}(-1)^{l-1}\binom{N}{l}\left(1-\frac{l}{N}\right)^i=\sum_{l=0}^{N}(-1)^l\binom{N}{l}\left(1-\frac{l}{N}\right)^i$

B_i：购买 i 件商品后收集到所有赠券

A_i^j：购买 i 件商品后收集到第 j 种赠券

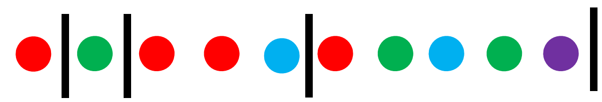
$\overline{A_i^j}$：购买 i 件商品后未收集到第 j 种赠券

$\overline{A_i^j}\cap\overline{A_i^k}$：购买 i 件商品后未收集到第 j 种和第 k 种赠券

赠券收集问题

- 赠券收集问题
  - 随机变量 X 为“集齐全套赠券购买的商品件数”，定义随机变量 Y_k 为“从收集到 $k-1$ 种赠券到 k 种赠券购买的商品件数”
  - $X = Y_1 + Y_2 + \cdots + Y_n$
  - $E(X) = E(Y_1) + E(Y_2) + \cdots + E(Y_n) = \sum_{k=1}^{N} \frac{N}{N-k+1} = N \sum_{k=1}^{N} \frac{1}{k}$

	B_{20}	B_{40}	B_{50}	B_{100}	B_{200}	$E(X)$
西湖十景	0.2147	0.8580	0.9491	0.9997	1.0000	29.2897
十二生肖	0.0513	0.6732	0.8523	0.9980	1.0000	37.2385
二十四节气		0.0018	0.0262	0.7025	0.9952	90.6230
五十六个民族				2.40×10^{-6}	0.1960	258.242
水浒一百单八将					8.99×10^{-11}	568.509



$X = 10$

$Y_1 = 1 \quad Y_2 = 1 \quad Y_3 = 3 \quad Y_4 = 5$

$Y_k = j$：先购买的 $j-1$ 件商品中的赠券均为已收集到的 $k-1$ 种中的一种，第 j 件商品中有未收集到的 $N-k+1$ 种赠券中的一种

几何分布

$$p = \frac{N-k+1}{N} \quad E(Y_k) = \frac{N}{N-k+1}$$

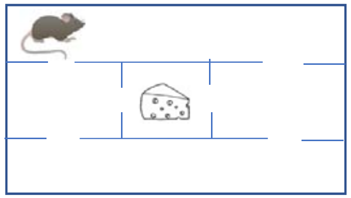
$$\sum_{i=1}^{N} i \sum_{l=0}^{N} (-1)^l \binom{N}{l} \left(1-\frac{l}{N}\right)^i = N \sum_{k=1}^{N} \frac{1}{k}$$

随机过程

数学建模
MATH T

- 随机过程（stochastic process）
  - 描述随机现象随时间推移而演化的一类数学模型
  - 一族随机变量 $\{X(t), t \in T\}$，其中 t 是参数，T 为参数集
    - T 为整数集的随机过程称为随机序列
- Markov过程
  - 在已知目前的状态（现在）的条件下，它未来的演变（将来）不依赖于它以往的演变（过去）
  - 随机序列 $\{X_n, n = 0,1,2,\cdots\}$，$X_n$ 只能取有限个或可数个数值
    - $P\{X_{n+1} = i_{n+1}\}$ 只与 X_n 有关，而与 $X_{n-1}, X_{n-2}, \cdots, X_0$ 无关
    - 对任意 $n \geq 0$ 和一列状态 $i_0, i_1, \cdots, i_{n-1}, i_n, i_{n+1}$，

$$P\{X_{n+1} = i_{n+1} \mid X_0 = i_0, X_1 = i_1, \cdots, X_{n-1} = i_{n-1}, X_n = i_n\} = P\{X_{n+1} = i_{n+1} \mid X_n = i_n\}$$



Andrey Andreyevich Markov
（1856 –1922）
俄罗斯数学家

赌徒破产问题

- 赌徒破产问题（Gambler's Ruin）
  - 一个赌徒在初始时拥有 h 个单位财富。在每局赌博中以概率 p 赢一个单位财富，以概率 $q=1-p$ 输一个单位财富。各局赌博结果独立
  - 赌徒在财富达到 N 个单位或 0 个单位（破产）时停止赌博
  - 求赌徒破产的概率

Blaise Pascal
（1623 –1662）
法国数学家、物理学家

A,B两人玩一种游戏，初始时两人各有12分。一次掷三枚骰子，若点数恰为11则A胜B负，点数恰为14则B胜A负，其他点数不分胜负。对出现胜负的一次投掷，胜一方增1分，负一方减1分。分数先为0分的一方失败。求双方获胜的概率之比

$$150094635296999121:129746337890625=3^{36}:3^{12}5^{12}$$

Huygens C, Oeuvres complètes. Tome I: Correspondance 1638-1656 Nijhoff, 1888.
No 342, Christiaan Huygens à P. de Carcavy. 12 octobre 1656.

递推关系

数学建模

MATH T

- 递推关系
  - $\begin{cases} A_h - A_{h+1} = r\left(A_{h-1} - A_h\right), 1 \le h \le N-1 \\ A_0, A_N \text{ 已知} \end{cases}$
    - 当 $r \neq 1$ 时，$A_h = \dfrac{r^h - r^N}{1-r^N} A_0 + \dfrac{1-r^h}{1-r^N} A_N$
      - 若 $A_0 = 0, A_N = 1$，$A_h = \dfrac{1-r^h}{1-r^N}$
    - 当 $r = 1$ 时，$A_h = \dfrac{N-h}{N} A_0 + \dfrac{h}{N} A_N$
      - 若 $A_0 = 0, A_N = 1$，$A_h = \dfrac{h}{N}$

$$A_h - A_N = (N-h)\left(A_0 - A_1\right)$$
$$A_0 - A_N = N\left(A_0 - A_1\right)$$
$$\frac{A_h - A_N}{A_0 - A_N} = \frac{N-h}{N} \Rightarrow A_h = \frac{N-h}{N} A_0 + \frac{h}{N} A_N$$

$$A_0 - A_1 = r^0\left(A_0 - A_1\right)$$
$$A_1 - A_2 = r\left(A_0 - A_1\right)$$
$$A_2 - A_3 = r\left(A_1 - A_2\right) = r^2\left(A_0 - A_1\right)$$
$$\cdots\cdots$$
$$A_h - A_{h+1} = r\left(A_{h-1} - A_h\right) = r^h\left(A_0 - A_1\right)$$
$$\cdots\cdots$$
$$A_{N-1} - A_N = r\left(A_{N-2} - A_{N-1}\right) = r^{N-1}\left(A_0 - A_1\right)$$

$$A_h - A_N = \left(r^h + \cdots + r^{N-1}\right)\left(A_0 - A_1\right)$$
$$= \frac{r^h\left(1 - r^{N-h}\right)}{1-r}\left(A_0 - A_1\right) = \frac{r^h - r^N}{1-r}\left(A_0 - A_1\right)$$
$$A_0 - A_N = \frac{1-r^N}{1-r}\left(A_0 - A_1\right)$$
$$\frac{A_h - A_N}{A_0 - A_N} = \frac{r^h - r^N}{1-r^N} \Rightarrow A_h = \frac{r^h - r^N}{1-r^N} A_0 + \frac{1-r^h}{1-r^N} A_N$$

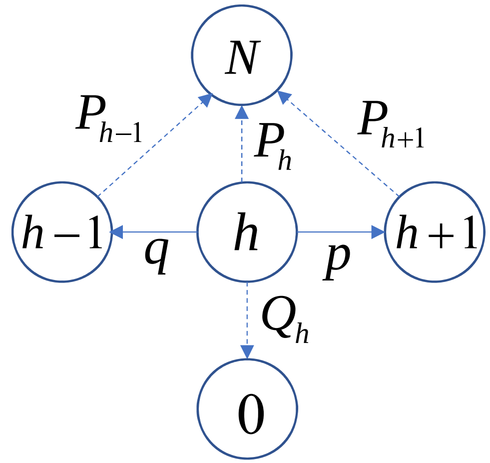
赌徒破产问题

数学
建模
MATH T

- 赌徒破产问题
  - 记 P_h 和 Q_h 分别为赌徒初始财富为 h 个单位时，财富最终达到 N 个单位的和 0 个单位的概率
    - $P_N = 1, P_0 = 0, Q_N = 0, Q_0 = 1$
  - $P_h = pP_{h+1} + qP_{h-1}$, $1 \le h \le N-1$
    - $P_h - P_{h+1} = \dfrac{q}{p}\left(P_{h-1} - P_h\right)$
    - $P_h = \begin{cases} \dfrac{\left(\dfrac{q}{p}\right)^h - 1}{\left(\dfrac{q}{p}\right)^N - 1} & p \ne q \\ \dfrac{h}{N} & p = q \end{cases}$

$$\begin{cases} A_h - A_{h+1} = r\left(A_{h-1} - A_h\right), 1 \le h \le N-1 \\ A_0 = 0, A_N = 1 \end{cases}$$

$$A_h = \begin{cases} \dfrac{1-r^h}{1-r^N} & r \ne 1 \\ \dfrac{h}{N} & r = 1 \end{cases}$$



赌徒破产问题

数学
建模
MATH T

- 赌徒破产问题
  - 设想赌徒与另一位在初始时拥有 $N-h$ 个单位财富的虚拟赌徒赌博。在每局赌博中虚拟赌徒以概率 q 赢一个单位财富，以概率 p 输一个单位财富。赌徒破产时，虚拟赌徒财富达到 N 个单位
  - $Q_h+P_h=1$，赌博不会永不终止
  - 当 N 充分大时，$Q_h \to \begin{cases} \left(\dfrac{q}{p}\right)^h, & p>q \\ 1, & p \le q \end{cases}$。即使赌博是“公平”的，赌徒终将破产

$$P_h=\begin{cases} \dfrac{\left(\dfrac{q}{p}\right)^h-1}{\left(\dfrac{q}{p}\right)^N-1} & p\neq q \\ \dfrac{h}{N} & p=q \end{cases}$$

$$Q_h=\begin{cases} \dfrac{\left(\dfrac{p}{q}\right)^h-1}{\left(\dfrac{p}{q}\right)^N-1}=\dfrac{\left(\dfrac{q}{p}\right)^h-\left(\dfrac{q}{p}\right)^N}{1-\left(\dfrac{q}{p}\right)^N} & p\neq q \\ \dfrac{N-h}{N} & p=q \end{cases}$$

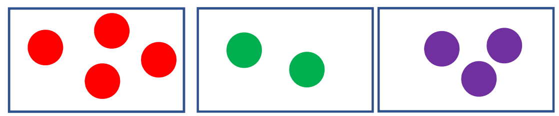
p	0.5	0.5	0.45	0.45	0.45	0.4	0.4
h	10	10	10	20	10	10	10
N	20	100	20	40	15	15	11
Q_h	0.5	0.9	0.8815	0.9822	0.6662	0.8703	0.3372

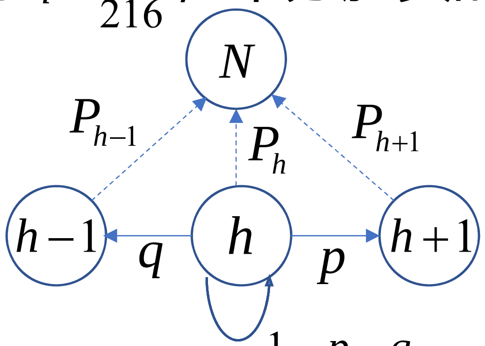
赌徒破产问题

数学建模

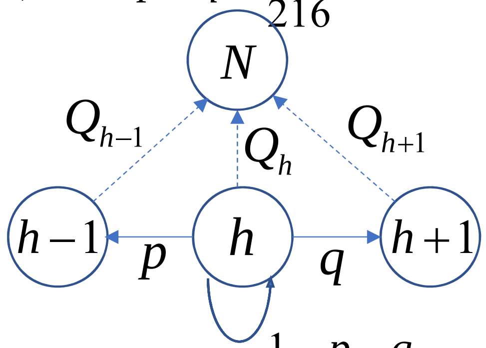
MATH T

- **Pascal**问题
  - 掷三枚骰子，点数恰为 j 的概率
    - $$\left(x+x^2+x^3+x^4+x^5+x^6\right)\cdot\left(x+x^2+x^3+x^4+x^5+x^6\right)\cdot\left(x+x^2+x^3+x^4+x^5+x^6\right)$$
      $$=x^3+3x^4+6x^5+10x^6+15x^7+21x^8+25x^9+27x^{10}$$
      $$+27x^{11}+25x^{12}+21x^{13}+15x^{14}+10x^{15}+6x^{16}+3x^{17}+x^{18}$$
  - 一次投掷，A胜的概率为 $p=\dfrac{27}{216}$， B胜的概率为 $q=\dfrac{15}{216}$，不分胜负的概率为 $1-p-q=\dfrac{174}{216}$

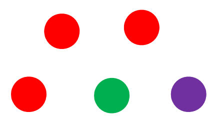




$$P_h=pP_{h+1}+qP_{h-1}+\left(1-p-q\right)P_h$$



$$Q_h=qQ_{h+1}+pQ_{h-1}+\left(1-p-q\right)Q_h$$



$$5=\begin{cases}3+1+1\\2+1+2\\2+2+1\\1+1+3\\1+2+2\end{cases}$$

$$\left(x+x^2+x^3+x^4\right)\cdot\left(x+x^2\right)\cdot\left(x+x^2+x^3\right)$$
$$=x^3+3x^4+5x^5+6x^6+5x^7+3x^8+x^9$$

赌徒破产问题

数学建模

MATH T

- **Pascal问题**
  - 记 P_h 和 Q_h 分别为A,B初始得分为 h 时，最终获胜的概率
    - $P_{24}=1, P_0=0, Q_{24}=1, Q_0=0$
  - $P_h = pP_{h+1}+qP_{h-1}+(1-p-q)P_h \;,1\le h\le 23$
    - $P_h-P_{h+1}=\dfrac{q}{p}(P_{h-1}-P_h)$
  - $Q_h = qQ_{h+1}+pQ_{h-1}+(1-p-q)Q_h, 1\le h\le 23$
    - $Q_h-Q_{h+1}=\dfrac{p}{q}(Q_{h-1}-Q_h)$
  - $P_{12}=\dfrac{\left(\frac{15}{27}\right)^{12}-1}{\left(\frac{15}{27}\right)^{24}-1}=\dfrac{282429536481}{282673677106} \quad Q_{12}=\dfrac{\left(\frac{27}{15}\right)^{12}-1}{\left(\frac{27}{15}\right)^{24}-1}=\dfrac{244140625}{282673677106}$

$$\frac{150094635296999121}{129746337890625}=\frac{282429536481}{244140625}=\frac{3^{24}}{5^{12}}$$

$$P_h=\frac{\left(\frac{q}{p}\right)^h-1}{\left(\frac{q}{p}\right)^{24}-1} \qquad Q_h=\frac{\left(\frac{p}{q}\right)^h-1}{\left(\frac{p}{q}\right)^{24}-1}$$

$$\begin{cases} A_h-A_{h+1}=r\left(A_{h-1}-A_h\right), \\ \qquad\qquad 1\le h\le N-1 \\ A_0=0, A_N=1 \end{cases}$$

$$A_h=\begin{cases} \dfrac{1-r^h}{1-r^N} & r\ne 1 \\ \dfrac{h}{N} & r=1 \end{cases}$$


谢 谢