

工程问题建模与实践

蒙特卡罗方法和布丰投针试验



上海交通大学

电子工程系

2022年4月

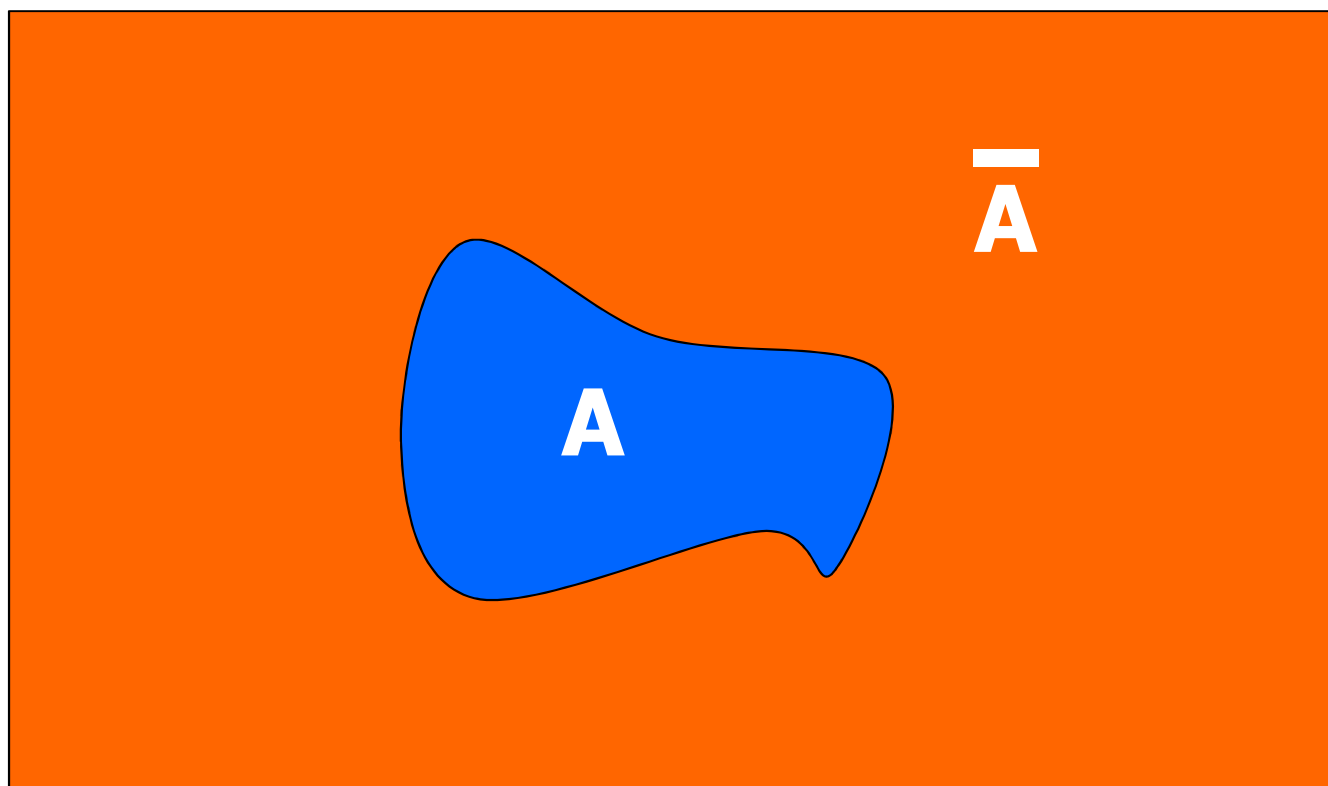


- 蒙特卡罗 (Monte Carlo) 方法概念和名词由来
- 布丰投针试验及其计算机模拟
 - 关于布丰
 - 关于投针试验
 - 投针试验的计算机模拟
- 在流行病学研究中的应用



□ 蒙特卡罗方法

蒙特·卡罗方法 (Monte Carlo method)，一种以概率统计理论为指导的非常有效的数值计算方法，也称统计（随机）试验方法。



$$S_A = \frac{\text{A的雨水量}}{\text{A的雨水量} + \bar{\text{A}}\text{的雨水量}} \cdot S_{\text{长方形}}$$



□ 蒙特卡罗方法

蒙特·卡罗方法 (Monte Carlo method)，一种以概率统计理论为指导的非常有效的数值计算方法，也称统计（随机）试验方法。

20世纪40年代第二次世界大战中，美国研制原子弹的“曼哈顿计划”成员S. M. 乌拉姆和J. 冯·诺伊曼提出使用随机试验法求解一些工程问题。

冯·诺伊曼用闻名于世的赌城——摩纳哥的Monte Carlo来命名这种方法，为它蒙上了一层神秘色彩。

1777年，法国科学家布丰 (Georges Louis Leclerc de Buffon, 1707—1788) 提出用投针试验的方法求圆周率 π 的数值。这被认为是蒙特卡罗方法的起源。

- 蒙特卡罗 (Monte Carlo) 方法概念和名词由来
- 布丰投针试验及其计算机模拟
 - 关于布丰
 - 关于投针试验
 - 投针试验的计算机模拟
- 在流行病学研究中的应用

□ 布丰



布丰，又译作布封、布冯、蒲丰

Georges Louis Leclerc de Buffon
(1707—1788)

法国博物学家、作家，人文主义的杰出代表，普及科学的先锋。

1707年生于法国勃艮第地方贵族世家。

青年时代就才华横溢。二十八岁受法王路易十五表彰，授以皇家植物园总监之职。

著有三十六卷的《自然史》，以及数十种重要论文，涉及各门科学。路易十五封他为布丰伯爵，叶卡捷琳娜二世女皇赐他荣誉。被选为伦敦、柏林和圣彼得堡各地的科学院院士。

《自然史》中名篇《松鼠》、《马》等曾入选我国中小学教材。

布丰投针试验

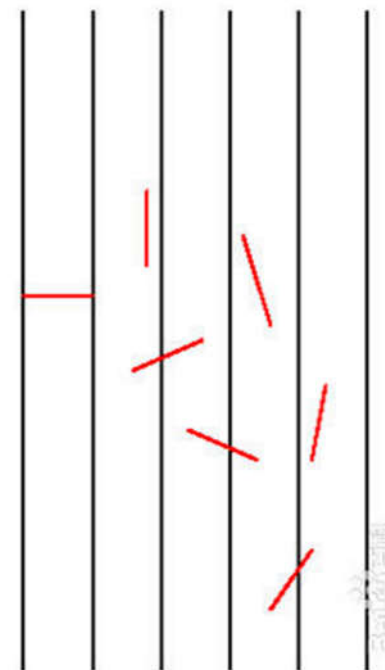
1777年，布丰提出以下问题：设在一片以平行且等距木纹铺成的地板上（如右图），随意抛一根长度不超过木纹间距的针，求针和其中任意一条木纹相交的概率表达式。

$$P = \frac{2l}{a\pi} \quad \text{式中, } l \text{ 为针的长度, } a \text{ 为木纹间距。}$$

以此表达式，布丰提出的一种估算圆周率数值的方法——随机投针法。

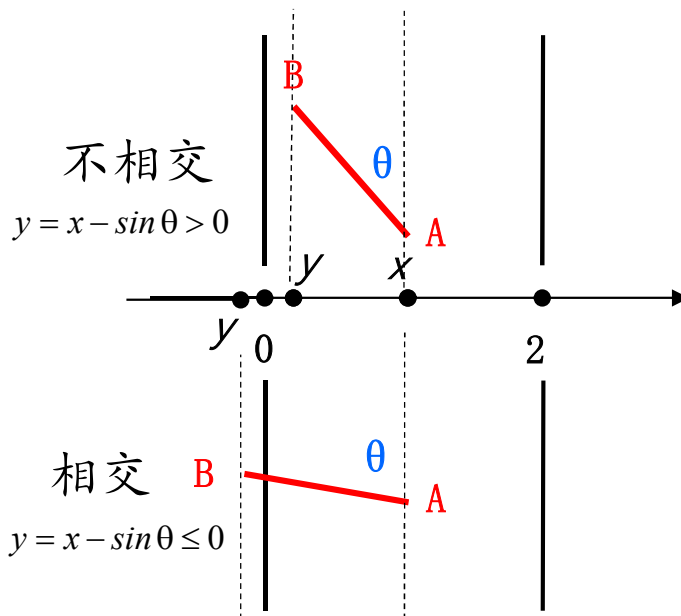
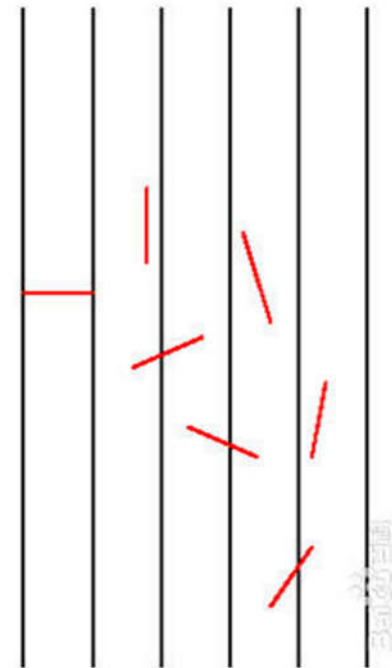
$$P = \frac{2l}{a\pi} \approx \frac{n}{N} \quad \Rightarrow \quad \pi \approx \frac{2lN}{an}$$

式中， n 是相交次数， N 是总的抛投次数。

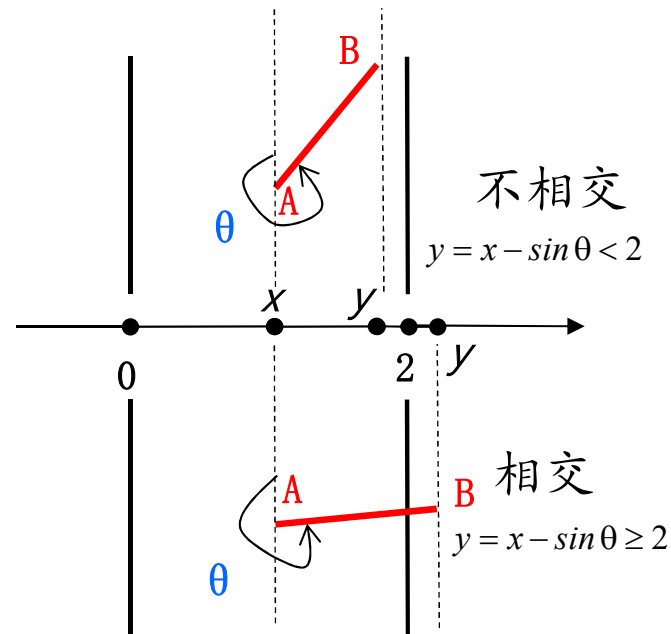


布丰投针试验

为简便起见，取针的长度为1，木纹间距为2。显然， x 为 $[0,2)$ 区间均匀分布随机数， θ 为 $[0,2\pi)$ 均匀分布随机数。



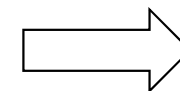
B在A左侧



B在A右侧

$$P = \Pr\{y \leq 0 \mid 0 \leq \theta < \pi\} + \Pr\{y \geq 2 \mid \pi \leq \theta < 2\pi\}$$

$$= \int_0^\pi \frac{1}{2\pi} d\theta \int_0^{\sin \theta} \frac{1}{2} dx + \int_\pi^{2\pi} \frac{1}{2\pi} d\theta \int_{2+\sin \theta}^2 \frac{1}{2} dx = \frac{1}{\pi}$$



$$\pi \approx \frac{N}{n}$$

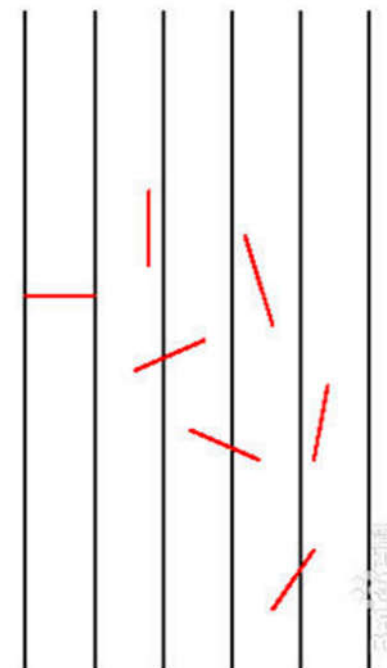
布丰投针试验

$$\pi \approx \frac{2lN}{an}$$

式中， l 为针的长度， a 为木纹间距；
 n 是相交次数， N 是总的抛投次数。

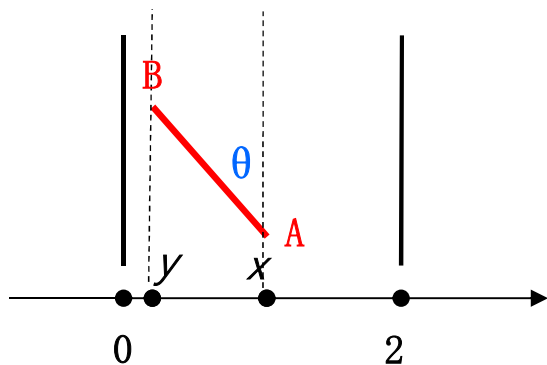
历史上真实开展投针试验的情况的部分报道：

试验者	时间	投掷次数	相交次数	圆周率估计值
Wolf	1850年	5000	2532	3.1596
Smith	1855年	3204	1218.5	3.1554
C.De Morgan	1860年	600	382.5	3.137
Fox	1884年	1030	489	3.1595
Lazzerini	1901年	3408	1808	3.1415929
Reina	1925年	2520	859	3.1795



布丰投针试验的计算机模拟

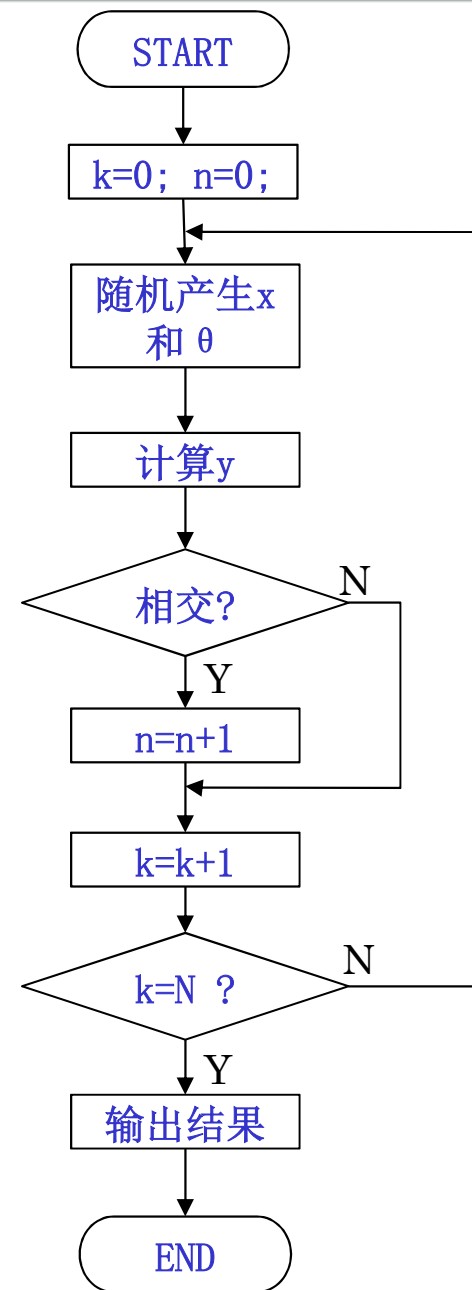
为简便起见，取针的长度为1，木纹间距为2。显然， x 为 $[0,2)$ 区间均匀分布随机数， θ 为 $[0,2\pi)$ 均匀分布随机数。



$$y = x - \sin \theta$$

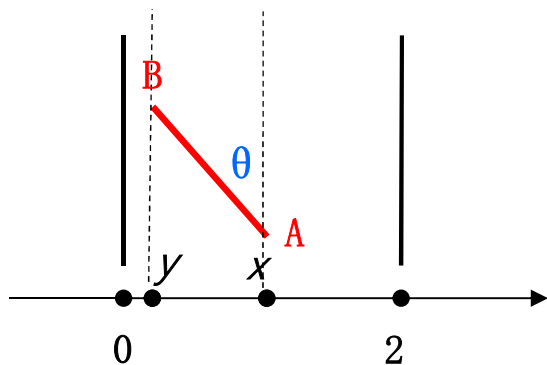
$$\pi \approx \frac{N}{n}$$

在电子计算机普及后，用计算机程序产生随机数（或更常见的伪随机数）来模拟随机事件变得十分高效，蒙特卡罗方法得以非常方便地用于仿真实验，求解各类数值问题。



布丰投针试验的计算机模拟

为简便起见，取针的长度为1，木纹间距为2。显然， x 为 $[0,2)$ 区间均匀分布随机数， θ 为 $[0,2\pi)$ 均匀分布随机数。



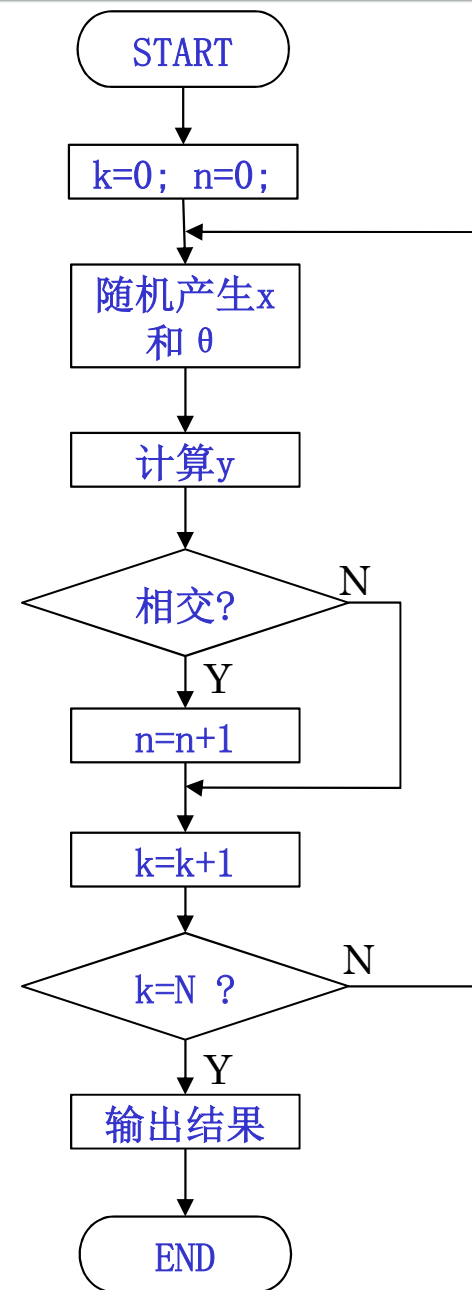
$$y = x - \sin \theta$$

$$\pi \approx \frac{N}{n}$$

不过，计算机仿真实验归根到底是对现实世界问题的一种近似模仿。程序产生随机数总有精度位数限制。

在本例中表现为 x 和 θ 的可取数值实际上是离散的，不可能做到在连续区间上的任意取值。变量对应的二进制位数越长，精度越高，仿真近似效果就越好，但程序运行效率会相应降低。

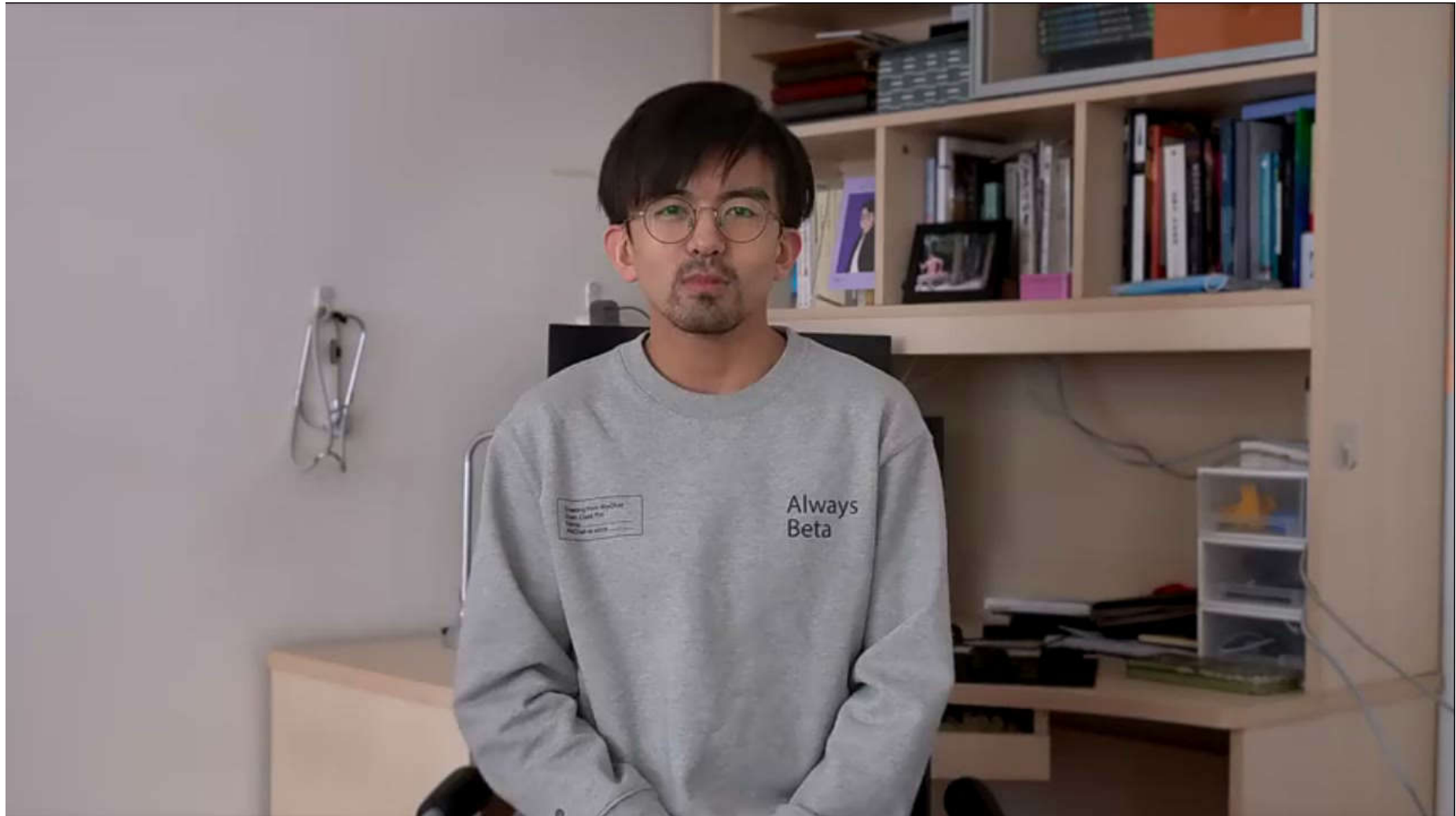
这其中包含的就是所谓 **仿真颗粒度** 的概念。



- 蒙特卡罗 (Monte Carlo) 方法概念和名词由来
- 布丰投针试验及其计算机模拟
 - 关于布丰
 - 关于投针试验
 - 投针试验的计算机模拟
- 在流行病学研究中的应用



基于传染病传播的常用数学模型, 利用蒙特卡罗方法完成仿真实验, 通过动画实现实验数据可视化。



基于SEIR等模型，组合使用蒙特卡罗方法和马尔科夫链，预测传染病疫情。

Forecasting the potential domestic and intern...



Figures

We used data from Dec 31, 2019, to Jan 28, 2020, on the number of cases exported from Wuhan internationally (known days of symptom onset from Dec 25, 2019, to Jan 19, 2020) to infer the number of infections in Wuhan from Dec 1, 2019, to Jan 25, 2020. Cases exported domestically were then estimated. We forecasted the national and global spread of 2019-nCoV, accounting for the effect of the metropolitan-wide quarantine of Wuhan and surrounding cities, which began Jan 23–24, 2020. We used data on monthly flight bookings from the Official Aviation Guide and data on human mobility across more than 100 prefecture-level cities in mainland China from the Tencent database. Data on confirmed cases were obtained from the reports published by the Chinese Center for Disease Control and Prevention. Serial interval estimates were based on previous studies of severe acute respiratory syndrome coronavirus (SARS-CoV). A susceptible-exposed-infectious-recovered metapopulation model was used to simulate the epidemics across all major cities in China. The basic reproductive number was estimated using Markov Chain Monte Carlo methods and presented using the resulting posterior mean and 95% credible interval (CrI).

MCMC

Findings

采用了SEIR模型，模拟了我国各主要城市的传染病流行情况。

用马尔科夫链蒙特卡罗方法估计基本繁殖数，然后用得到的后验均值和95%可信区间表示。

In our baseline scenario, we estimated that the basic reproductive number for 2019-nCoV was 2.68 (95% CrI 2.47–2.86) and that 75 815 individuals (95% CrI 37 304–130 330) have been infected in Wuhan as of Jan 25, 2020. The epidemic doubling time was 6.4 days (95% CrI 5.8–7.1). We estimated that in the baseline scenario, Chongqing, Beijing, Shanghai, Guangzhou, and Shenzhen had imported 461 (95% CrI 227–805), 113 (95% CrI 55–153), 98 (95% CrI 46–119), 11 (95% CrI 3–19), and 1 (95% CrI 0–1) cases from Wuhan, respectively. If the transmissibility of 2019-nCoV were similar everywhere domestically and over time, we inferred that epidemics are already growing exponentially in multiple major cities of China with a lag time behind the Wuhan outbreak of about 1–2 weeks.

在其中重点考虑了人口流动的影响



本讲结束，谢谢！

