

Monty Hall 问题的数值模拟设计与结果*

统计 91 董晟渤, 2193510853

西安交通大学数学与统计学院

日期: 2022 年 3 月 9 日

摘 要

针对 Monty Hall 问题, 本篇报告在一定假设下, 设计了数值模拟, 计算得到参赛者选择不换门和选择换门时选中正确的门的概率分别为 $1/3$ 和 $2/3$, 并对该结果进行简单分析.

关键词: 条件概率, Monty Hall 问题, 数值模拟.

目录

1 概述: 条件概率与 Monty Hall 问题	2
2 数值模拟的设计与结果	2
2.1 基本假设	2
2.2 模拟设计	2
2.3 试验结果	3
附录: 数值模拟的 MATLAB 实现	4

*2021-2022 学年第二学期, 课程: 时间序列与金融统计, 指导老师: 惠永昌.

1 概述: 条件概率与 Monty Hall 问题

设 $(\Omega, \mathcal{F}, \mathbb{P})$ 是概率空间, 对于事件 $A, B \in \mathcal{F}$, 定义在事件 A 发生的条件下, 事件 B 的条件概率为

$$\mathbb{P}(B|A) := \frac{\mathbb{P}(AB)}{\mathbb{P}(A)}.$$

初等概率论对条件概率的理解更多停留在直观上, 这也导致了面对一些问题时, 会出现模棱两可的情况. 为了体现这一点, 本篇报告考虑的是著名的 **Monty Hall 问题**, 又称三门问题.

Monty Hall 问题来自美国的电视游戏节目 *Let's Make a Deal*, 参赛者会看见三扇关闭了的门, 其中一扇的后面有一辆汽车, 选中后面有车的那扇门可赢得该汽车, 另外两扇门后面则各藏有一只山羊. 当参赛者选定了一扇门, 但未去开启它的时候, 节目主持人开启剩下两扇门的其中一扇, 露出其中一只山羊. 主持人其后会问参赛者要不要换另一扇仍然关上的门. 具体可见 [Monty Hall Problem](#).

该节目的观众们对该问题进行了热烈的讨论. 本篇报告针对该问题, 在一定假设下, 设计了数值模拟, 并分析了结果.

2 数值模拟的设计与结果

2.1 基本假设

为了编写代码, 笔者用 1, 2, 3 来表示三扇门, 并提出如下的假设:

- 假设每扇门背后的物品在电视节目目前已经通过完全随机的方式, 被事先确定, 这与现实的情况是相符的. 从而, 每次模拟的正确答案都是来自 $\{1, 2, 3\}$ 的随机数;
- 假设最开始参加节目的参赛者并不知道每扇门背后的物品, 他或她的选择是完全随机的. 否则, 如果参赛者知道比赛的结果的话, 无论怎样都能选出正确的门. 从而, 每次模拟时第一次选择的门也是来自 $\{1, 2, 3\}$ 的随机数;
- 假设当参赛者选定好一扇门后, 主持人所打开的后面是山羊的门是号数最小的门. 这看起来会对我们的“随机性”产生影响, 但是后面我们将会解释为什么这个假设是合理的.

2.2 模拟设计

本篇报告使用的编程语言是MATLAB. 考虑到其进行矩阵计算的优势, 我们使用 $1 \times n$ 、且取值在 $\{1, 2, 3\}$ 中的随机向量来表示正确的答案 (记为 `answer`). 而对于第 i 次模拟, 具体的

流程如下:

步骤 1 作出第一次选择 $result1(i)$, 并判断与正确答案 $answer(i)$ 是否相等;
步骤 2 主持人打开一扇与 $result1(i)$ 和 $answer(i)$ 都不同的门 $wrong(i)$, 具体地说:
步骤 2.1 若 $result1(i) == answer(i)$:
• 若 $answer(i) == 1$, 则 $wrong(i) = 2$;
• 否则, $wrong(i) = 1$;
步骤 2.2 否则, 打开与 $result1(i)$ 和 $answer(i)$ 相异的门 $wrong(i)$.
步骤 3 此时参赛者有两种选择:
步骤 3.1 若选择不换门, 则最终的选择为 $result1(i)$;
步骤 3.2 若选择换门, 则最终的选择 $result2(i)$ 为与 $result1(i)$ 和 $wrong(i)$ 相异的门.

在经过 n 次模拟之后, 可以计算出选择不换门和选择换门的各自的成功次数 n_1 和 n_2 , 从而估计出两种选择的概率分别为

$$\hat{p}_1 = \frac{n_1}{n}, \quad \hat{p}_2 = \frac{n_2}{n}.$$

回想起在做假设时, 我们认为主持人所打开的门是号数最小的, 这是否会对模拟的结果产生影响呢? 该假设的影响主要体现在**步骤 2.1** 中, 而该步骤中, $result1(i) == answer(i)$, 也即参赛者已经选出了正确的门, 从而有以下两种可能:

- 选择不换门, 则该次模拟选出的答案一定是正确答案, 这与主持人的选择无关;
- 选择换门, 则该次模拟选出的答案一定是错误答案, 这也与主持人的选择无关: 无论主持人选择打开的是哪扇门, 这么做的后果都是选择了错误答案.

从而, 我们说明了假设的合理性.

2.3 试验结果

共进行 $n = 100,000,000$ 次试验, 其中不经过换门且结果正确的有 $n_1 = 33,334,064$ 次试验, 经过换门且结果正确的有 $n_2 = 66,665,936$ 次试验, 从而两种策略的正确率

$$\hat{p}_1 \approx 0.3333, \quad \hat{p}_2 \approx 0.6667.$$

这个结果说明了: 选择不换门的话, 答案正确的概率是 $\frac{1}{3}$; 而选择换门的话, 答案正确的概率是 $\frac{2}{3}$, 并且它们的和恰好为 1.

模拟的结果与直觉有一定矛盾, 但是从数值模拟的角度, 可以简单解释. 考虑到主持人选择的门一定是错误的答案, 我们可以得到

$$n_1 + n_2 = n \implies \hat{p}_1 + \hat{p}_2 = \frac{n_1 + n_2}{n} = 1;$$

另外, 在作出第一次选择的时候, 三个选项中只有一个是正确答案, 从而 $\hat{p}_1 = \frac{1}{3}$, 据此解得 $\hat{p}_2 = \frac{2}{3}$.

附录: 数值模拟的 MATLAB 实现

```
%% 初始化

clc;
clear;
close;

%% 初始化

n = 100000000; % 模拟次数
answer = randi(3, 1, n); % 正确答案

%% 第一次选择

result1 = randi(3, 1, n); % 第一次选择
count1 = 0; % 正确次数
for i = 1 : n
    if answer(i) == result1(i)
        count1 = count1 + 1;
    end
end
p1 = count1 / n

%% 公开错误答案

wrong = zeros(1, n);
for i = 1 : n
    if answer(i) == result1(i)
        if answer(i) == 1
            wrong(i) = 2;
        else
            wrong(i) = 1;
        end
    end
end
else
```

```
wrong(i) = 6 - answer(i) - result1(i);  
end  
end  
  
%% 第二次选择  
  
result2 = 6 - result1 - wrong; % 第二次选择  
count2 = 0;  
for i = 1 : n  
    if answer(i) == result2(i)  
        count2 = count2 + 1;  
    end  
end  
p2 = count2 / n
```