

如何量化信息的价值

条件概率

小胖

目录

ONE 一个经典的问题

Monty Hall problem

TWO 条件概率

定量地描述信息的价值

THREE 贝叶斯定理

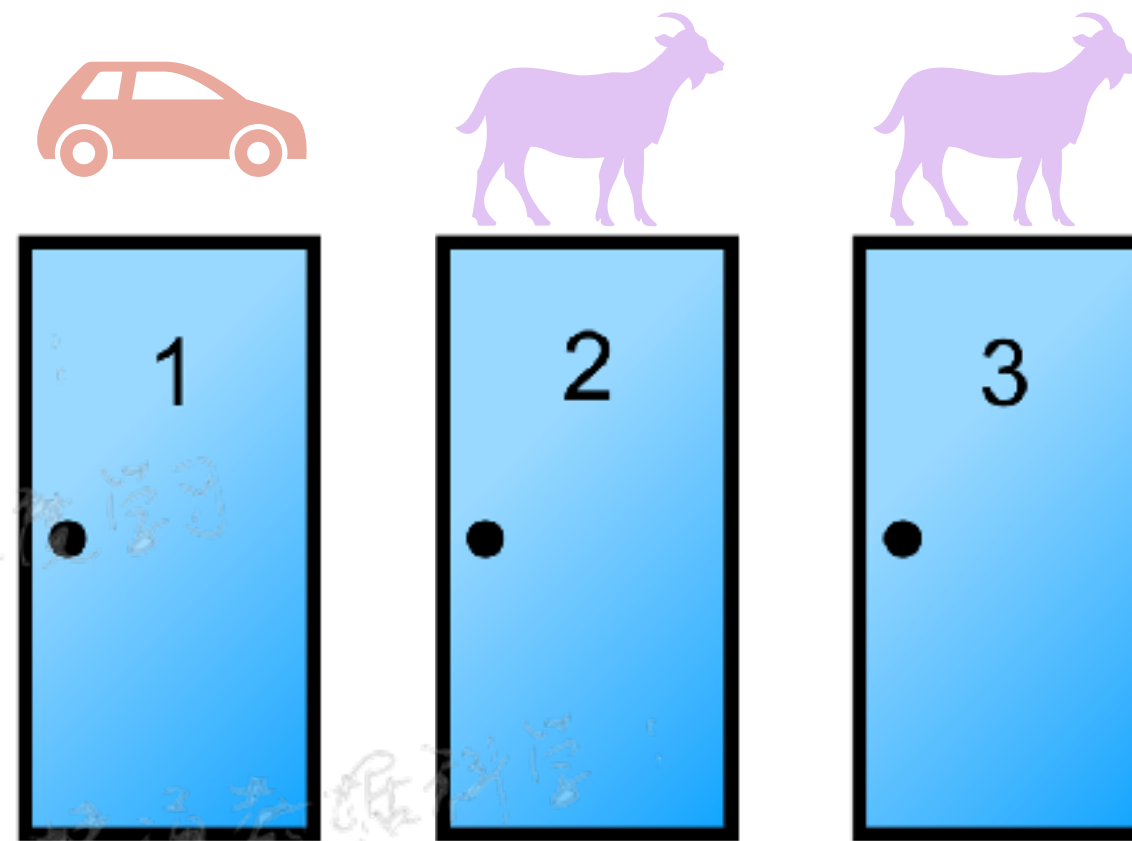
先验概率与后验概率

一个经典的问题

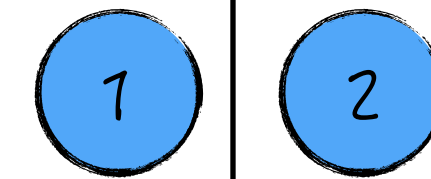
Monty Hall problem

Monty Hall problem (蒙提霍尔问题)：

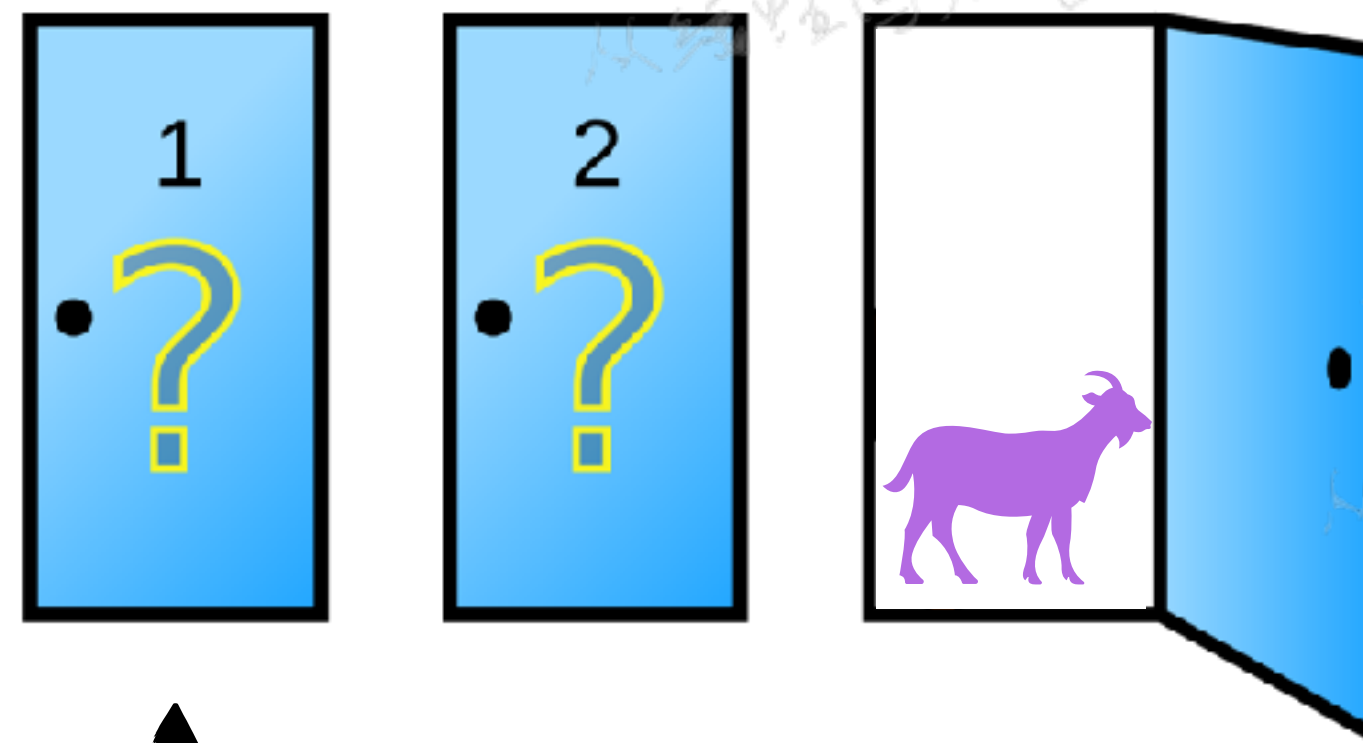
- 坚持最初选择还是更改选择？
- 绝大部分人的直觉是没必要更改选择，因为获奖的概率都是50%



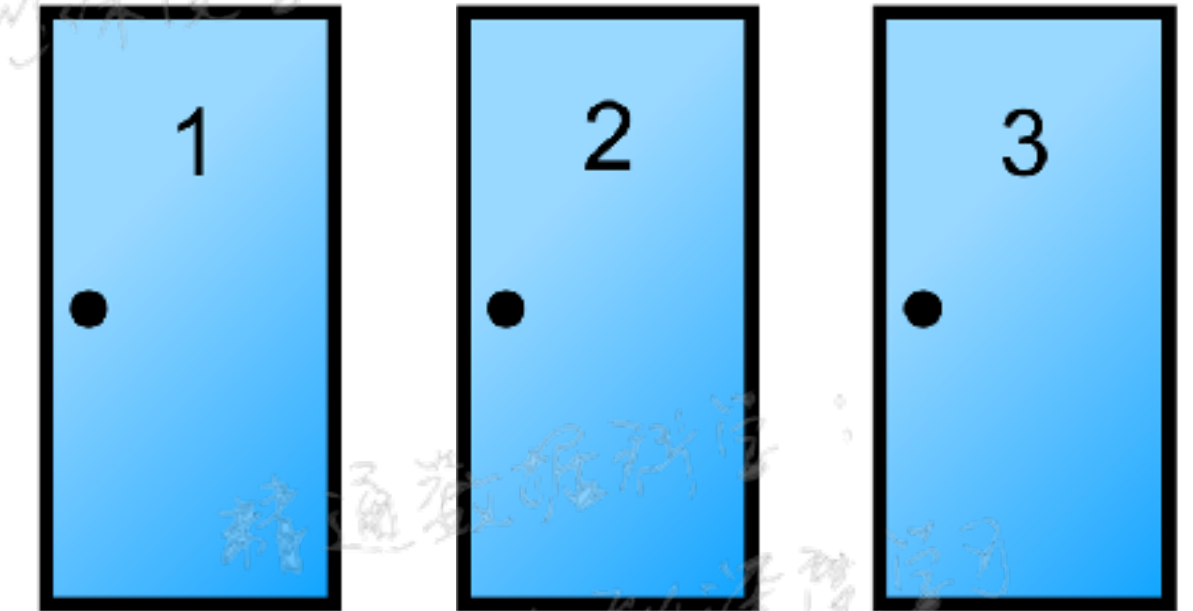
1辆汽车，2只山羊，3扇门



选手做出最初的选择



主持人打开一扇有山羊的门，坚持还是改变？

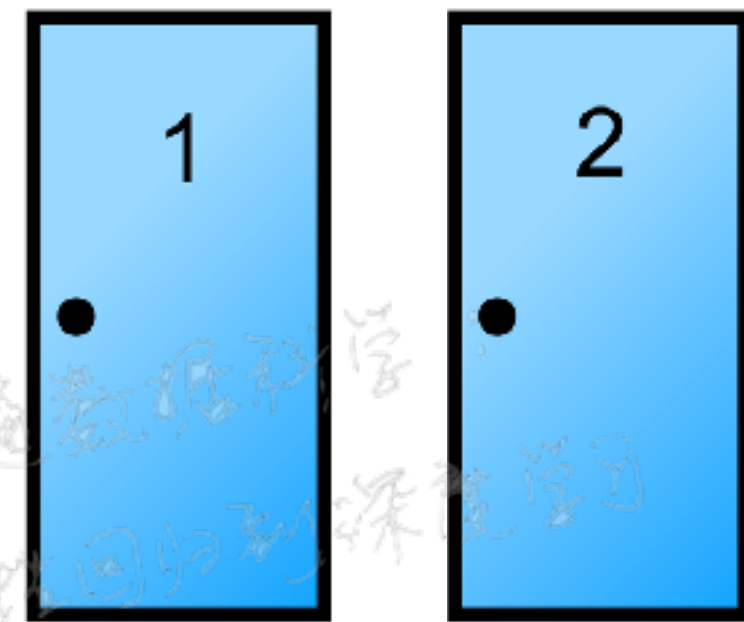


直觉上，两扇门没有任何区别
背后是汽车的概率都为50%
所以，无所谓坚持或改变

一个经典的问题

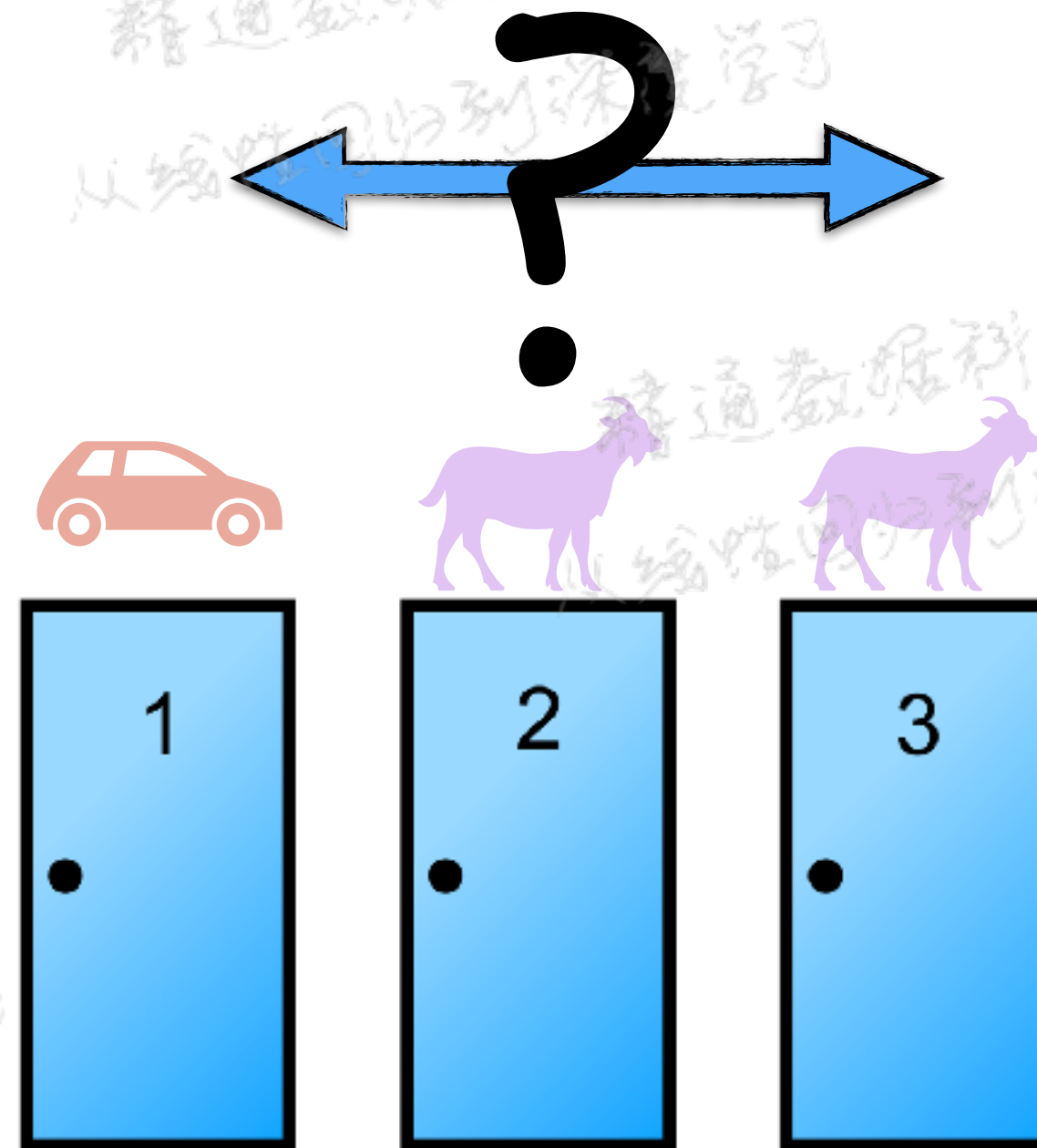
蒙特卡洛方法

直觉



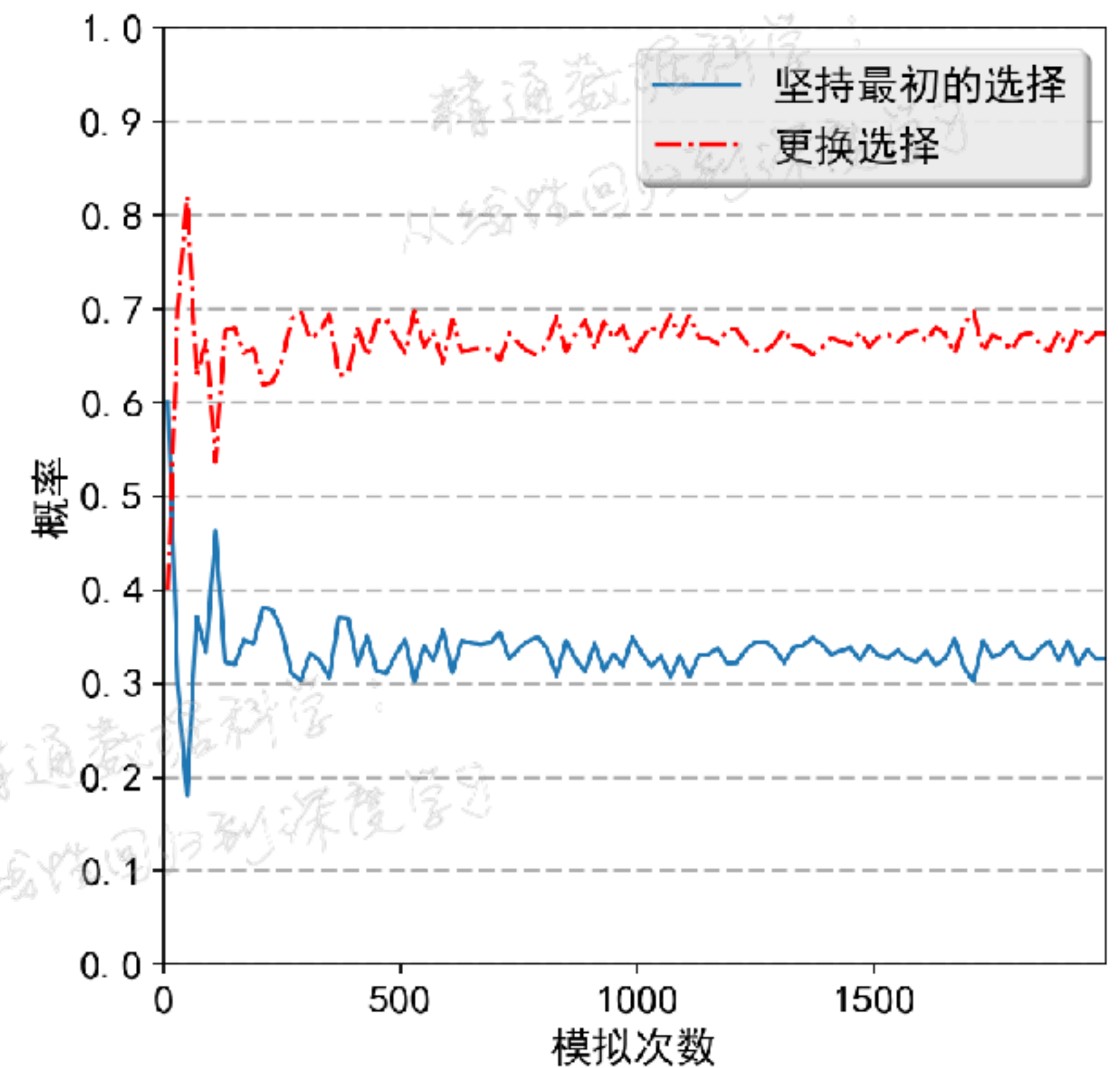
直觉上，两扇门没有任何区别
背后是汽车的概率都为50%
所以，无所谓坚持或改变

蒙特卡洛方法



通过计算机来随机模拟现实

实际情况



目录

ONE 一个经典的问题

Monty Hall problem

TWO 条件概率

定量地描述信息的价值

THREE 贝叶斯定理

先验概率与后验概率

条件概率

Monty Hall problem的数学证明

用A表示汽车所在的门牌号

$$P(A = 1) = P(A = 2) = P(A = 3) = \frac{1}{3}$$

用B表示主持人打开的门牌号

坚持原有选择的
获奖概率

$$P(A = 1 | B = 3)$$

更改选择的
获奖概率

$$P(A = 2 | B = 3)$$

根据贝叶斯定理

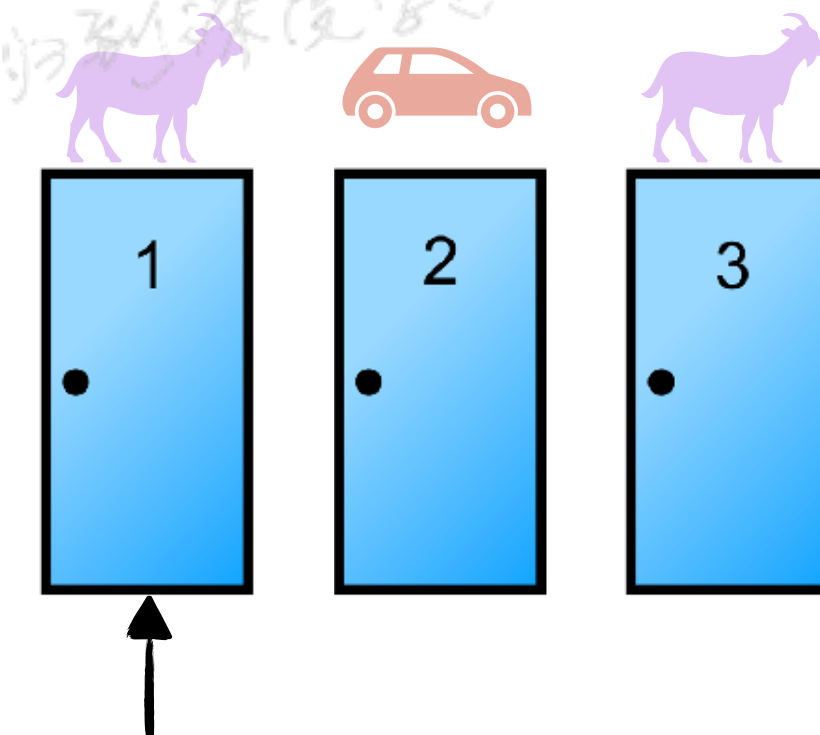
$$P(A = 2 | B = 3) = \frac{P(A = 2, B = 3)}{P(B = 3)}$$

计算汽车在2号门
且主持人打开3号门的概率

$$P(A = 2, B = 3) = P(B = 3 | A = 2) \times P(A = 2)$$

$$P(B = 3 | A = 2) = 1$$

$$P(A = 2, B = 3) = 1/3$$



条件概率

Monty Hall problem的数学证明

根据贝叶斯定理

$$P(A = 2 | B = 3) = \frac{P(A = 2, B = 3)}{P(B = 3)} \quad P(A = 2, B = 3) = 1/3$$

计算主持人打开3号门的概率

$$P(B = 3) = P(A = 1, B = 3) + P(A = 2, B = 3) + P(A = 3, B = 3)$$

$$P(A = 1, B = 3) = P(B = 3 | A = 1) \times P(A = 1)$$

$$P(B = 3 | A = 1) = 1/2$$

$$P(A = 1, B = 3) = 1/2 \times 1/3 = 1/6$$

$$P(B = 3) = 1/6 + 1/3 = 1/2$$

最后的计算结果

$$P(A = 2 | B = 3) = \frac{1/3}{1/2} = \frac{2}{3} \quad P(A = 1 | B = 3) = \frac{1}{3}$$

与模拟结果相符

更改选择的
获奖概率

坚持原有选择的
获奖概率

定量地描述信息的价值



目录

ONE 一个经典的问题

Monty Hall problem

TWO 条件概率

定量地描述信息的价值

THREE 贝叶斯定理

先验概率与后验概率

贝叶斯定理

先验概率与后验概率

贝叶斯定理

后验概率

建模目的

知果求因

知因求果

建模假设

先验概率

$$P(y | \mathbf{X}) = \frac{P(\mathbf{X} | y) \times P(y)}{P(\mathbf{X})}$$

$A \uparrow$

$B \downarrow$

已知汽车位置的情况下，主持人打开各号门的概率分布

生活常识、模型假设

透过表象，猜测原因

已知主持人开门位置的情况下，汽车所在位置的概率分布

THANK YOU

精通数据科学：
从线性回归到深度学习