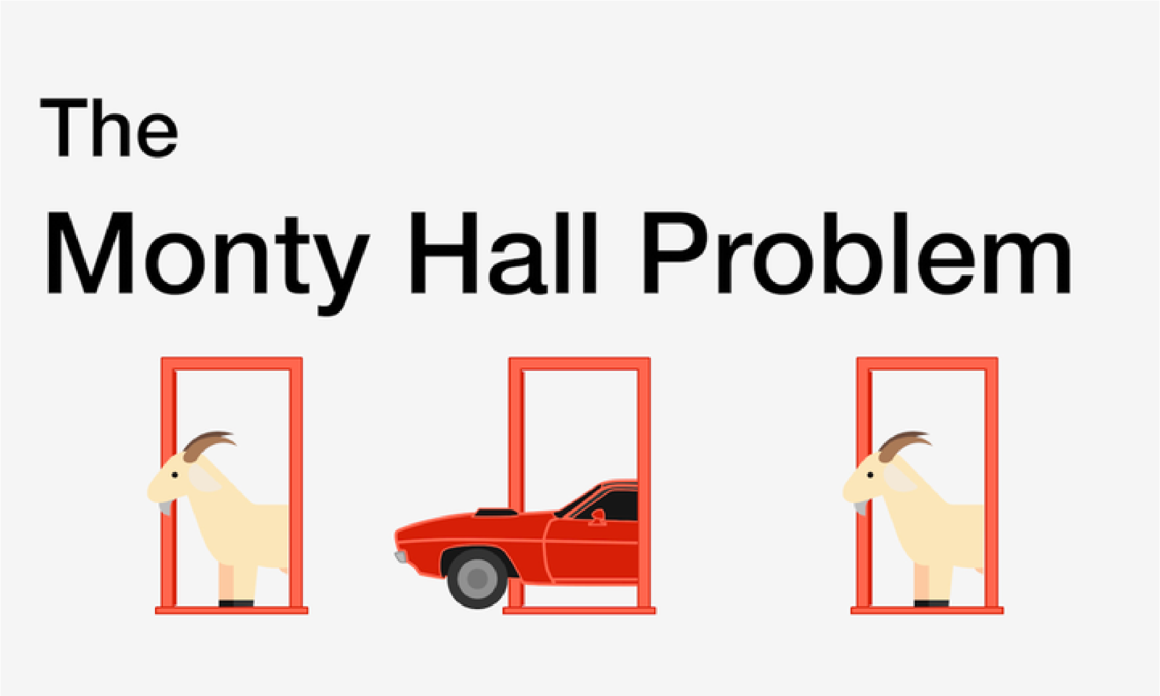
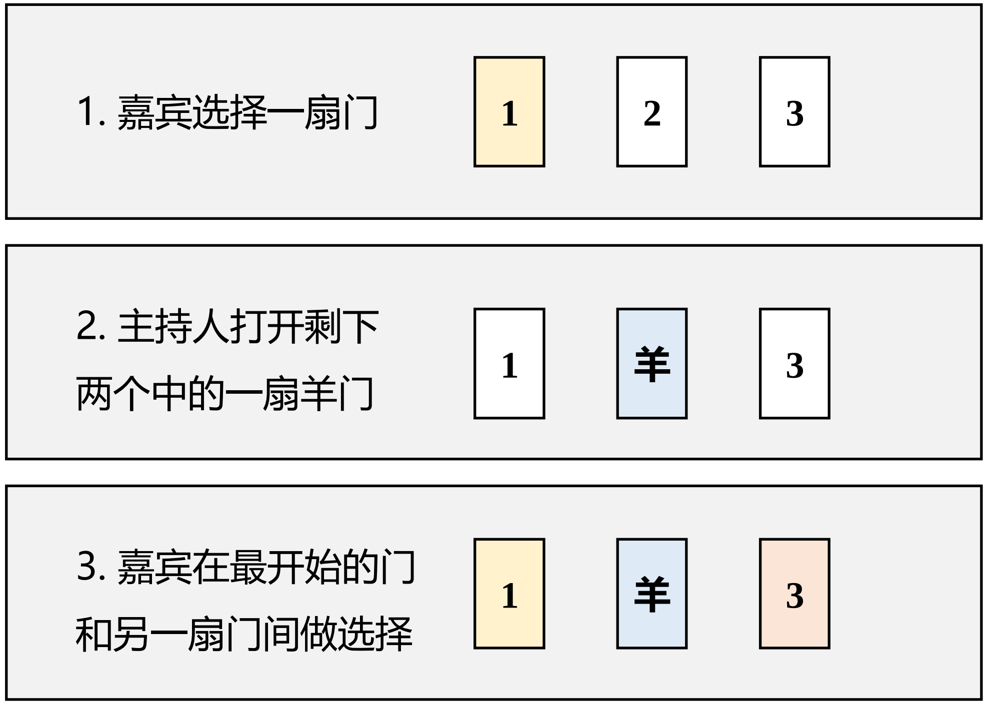
从蒙提霍尔悖论(三门问题)理解贝叶斯模型

蒙提霍尔问题

在一档电视节目中，主持人蒙提霍尔邀请一位嘉宾选择3扇门中的其中一扇，并告知嘉宾三扇门中有一扇后面放的是汽车，其余两扇后面放的是山羊。主持人实际上知道每扇门后面放的是什么，但不告诉嘉宾。

- 嘉宾首先选择三扇门中的一扇门
- 当嘉宾做出初选择后，为了让节目更加有趣，主持人会先打开另外2扇门的其中一扇背后是山羊的门
- 主持人会给嘉宾一次选择的机会——是坚持原来的选择还是选择另一扇门?





从蒙提霍尔悖论(三门问题)理解贝叶斯模型

贝叶斯基础

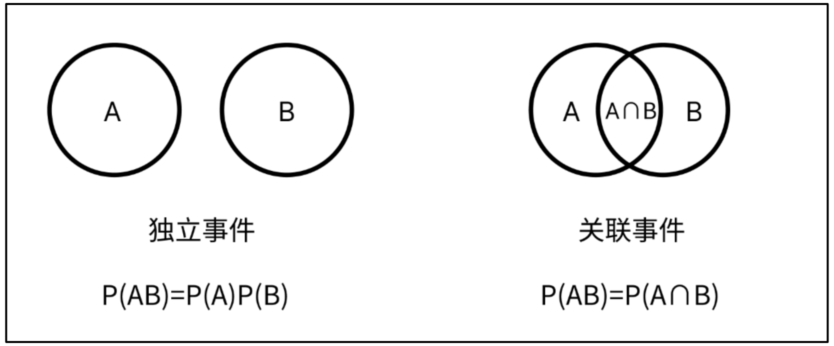
1. 独立事件与相关事件

- 假设样本空间中存在事件A、事件B
- 如果两个事件之间没有任何关联——独立事件。例如：看电视和喝牛奶
- 如果两个事件间存在关联，相互影响——相关事件。例如：吃早餐和喝牛奶

2. 条件概率

- 对于两个独立事件：事件同时发生概率P(AB)=P(A)P(B)
- 对于两个相关事件：事件同时发生概率P(AB)≠P(A)P(B)
- 那么，相关事件的全概率应该怎么算呢？ 答：我们需要引入条件概率
  - 条件概率：概率论将其定义为某件事的前提下另一件事发生的概率，写作P(B|A)
  - 相关事件的全概率就可以理解为：①在B事件发生的前提下②A事件也发生了
  - 因此，其概率写作

$$P(AB) = P(A)P(B|A) = P(B)P(A|B)$$ ———— 公式①



从蒙提霍尔悖论(三门问题)理解贝叶斯模型

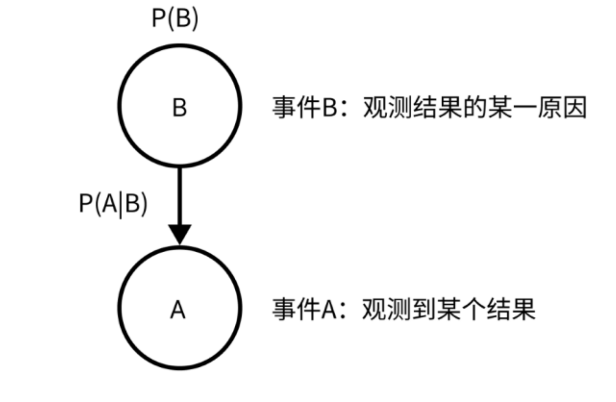
贝叶斯基础

3. 贝叶斯定理

- 基于公式①，我们可以得到典型的贝叶斯公式：

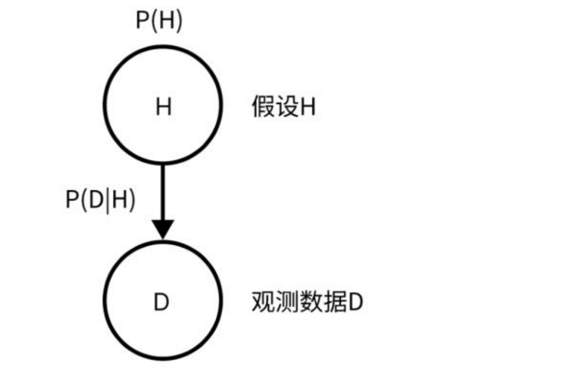
$$P(B|A)=\frac{P(B)P(A|B)}{P(A)} \qquad \text{——— 公式②}$$

- 概率论理论中，典型的贝叶斯定理用于描述两个相关事件间条件概率的关系。然而，贝叶斯定理的伟大之处在于其对现实问题的引申价值。其体现在两方面：
  - **寻果溯因**方面：贝叶斯定理可以表示为：事件A和事件B间存在因果关系。事件A是我们能够直观观测到的事件，例如：警报响了；事件B是事件A发生背后的原因，例如：发生火灾。因此，可以推理出：警报响了，有可能有地方失火了



两个具备因果关联的相关事件

  - **假设检验**方面：观测数据D和假设H之间的关系。例如，我假设一个病人可能患有某种疾病，观测数据D为该病人的检查结果。那么，可以通过贝叶斯定理在观测数据D的基础上，推理出该病人患病的概率。



两个假设检验关联的相关事件

贝叶斯基础

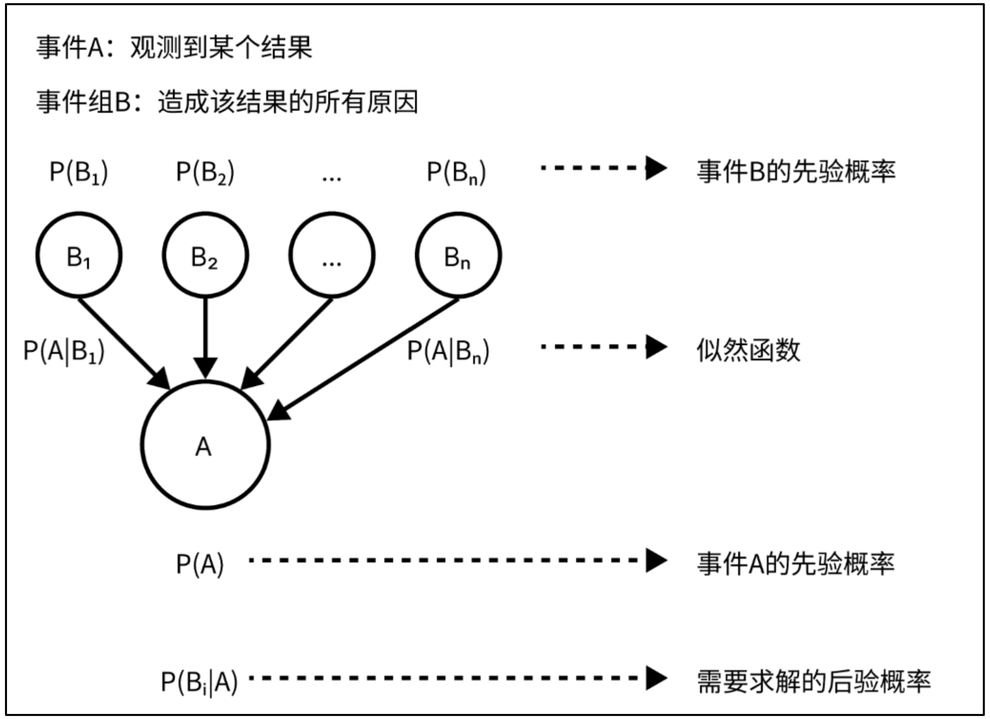
4. 基于全概率公式的贝叶斯定理

- 事实上，对于一个观测到的事件，其可能原因有很多。经典贝叶斯定理仅解释了两个事件间的条件概率关系，那么如果涉及到更多事件，贝叶斯定理该如何表示呢?
  - 我们可以将与事件A相关的所有事件放入一组，称为B组，共包含n个事件，我们将B组定义为完备事件组。那么，可引入全概率公式表示事件A发生的概率：

$$P(A)=\sum_{\mathrm{i}=1}^{n}P(A|B_i)P(B_i) \qquad \text{——— 公式③}$$

  - 此时，贝叶斯定理描述为：事件A发生的情况下是由事件Bi造成的概率：

$$P(B_i|A)=\frac{P(A|B_i)P(B_i)}{\sum_{\mathrm{i}=1}^{n}P(A|B_i)P(B_i)} \qquad \text{——— 公式④}$$

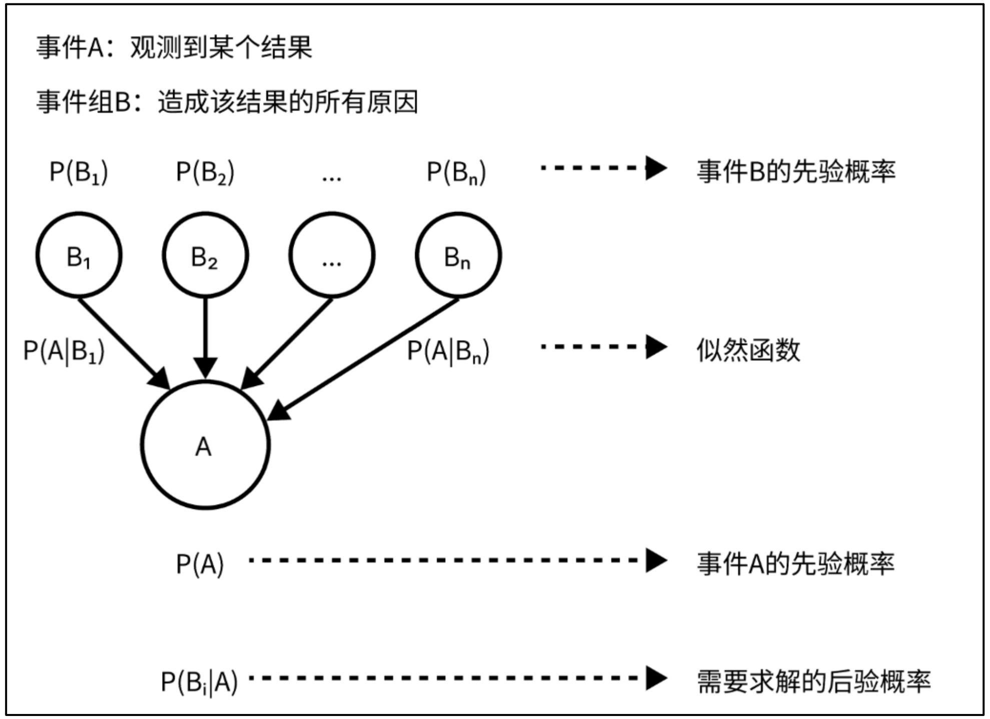


从蒙提霍尔悖论(三门问题)理解贝叶斯模型

贝叶斯基础

5. 先验概率、后验概率、最大似然概率

- 统计学中，我们通常将直接观测到事件的概率称为先验概率；在先验概率基础上总结某个原因造成某个事件发生的概率为似然概率或似然函数；推理某一事件发生背后的原因称为后验概率，也往往是需要求解的量。
  - 先验概率：强调“由因求果”问题中“因”出现的概率，这是根据以往经验或统计分析得到的
  - 后验概率：强调“执果溯因”问题，指某个事件发生的基础上是由某个因素发生导致的概率
  - 似然概率：强调“由因求果”问题中根据“因”估计“果”发生的概率分布



从蒙提霍尔悖论(三门问题)理解贝叶斯模型

蒙提霍尔悖论的贝叶斯论证

1. 嘉宾选择一扇门　1　2　3

2. 主持人打开剩下两个中的一扇羊门　1　羊　3

3. 嘉宾在最开始的门和另一扇门间做选择　1　羊　3

为什么选1号门中奖概率比选择3号门低呢?

从蒙提霍尔悖论(三门问题)理解贝叶斯模型

蒙提霍尔悖论的贝叶斯论证

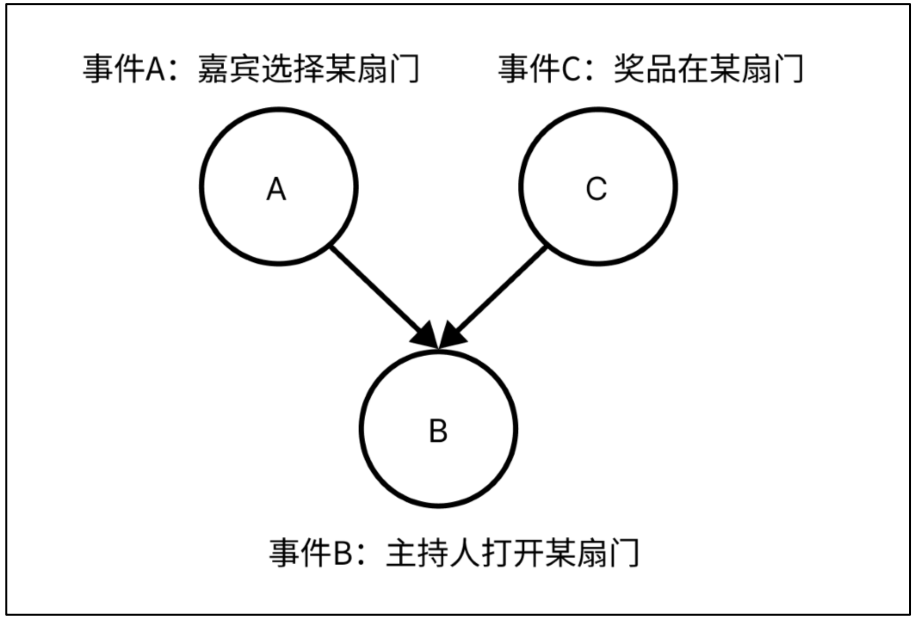
- 简单理解：
- 蒙提霍尔问题可以进行扩大，即：假设我去买彩票，彩票池里一共由100张彩票，我先选择一张，老板帮我剔除了剩下99张中98张没有奖的彩票，这是我是选择最开始的那张还是剩下的一张呢？显而易见，我最开始选择的那张中奖概率为1%，需要注意，**在同一样本空间中，已经发生的事件是不会受到后续事件而影响其发生概率。**而奖池中剩下的那张彩票代表的是99张的中奖概率，老板的剔除并没有改变这99张彩票的总体中奖概率，因此剩下的那张彩票中奖概率为99%。

- 信息熵角度理解：
- “主持人在剩下的两扇门中打开一扇无奖的门”和“主持人随机打开了一扇无奖的门”所包含的**信息量是完全不同的。**前者的信息熵并未影响最开始选择门A的中奖概率，但打破了门B和门C之间的概率均匀分布；而后者的信息熵则打破门A、B、C三者的概率均匀分布。因此，新的信息熵必然会导致概率变化，而概率变化状态与变化程度由信息熵所传递的信息量所决定

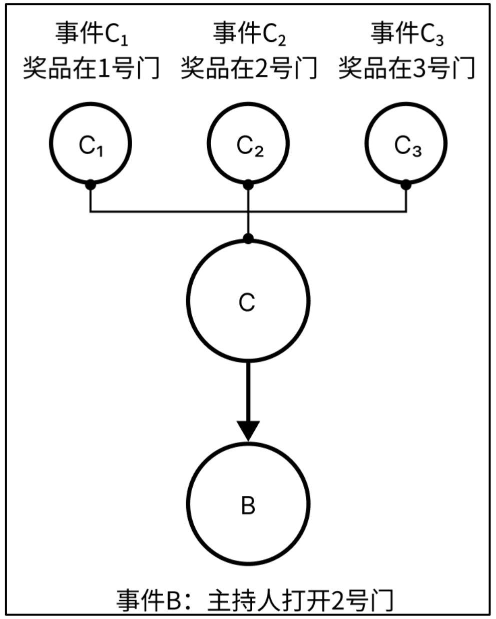
从蒙提霍尔悖论(三门问题)理解贝叶斯模型

蒙提霍尔悖论的贝叶斯论证

- 贝叶斯模型将蒙提霍尔悖论拆解：
- 主持人的选择(事件B)受到嘉宾最开始选择的门(事件A)和奖品所在的门(事件C)共同影响的，而嘉宾选择某扇门与奖品在某扇门之间彼此独立，其因果关系的描述如下所示：



- 为了简化模型，考虑嘉宾选择门完全随机，我们设定嘉宾一定首先选择第一扇门。基于此，该模型可以更简化为下图。其中，C1为奖品在1号门，C2为奖品在2号门，C3为奖品在3号门，三者共同组成完备事件组C；事件B为主持人打开2号门，为观测事件。



从蒙提霍尔悖论(三门问题)理解贝叶斯模型

蒙提霍尔悖论的贝叶斯论证

- 根据全概率公式：

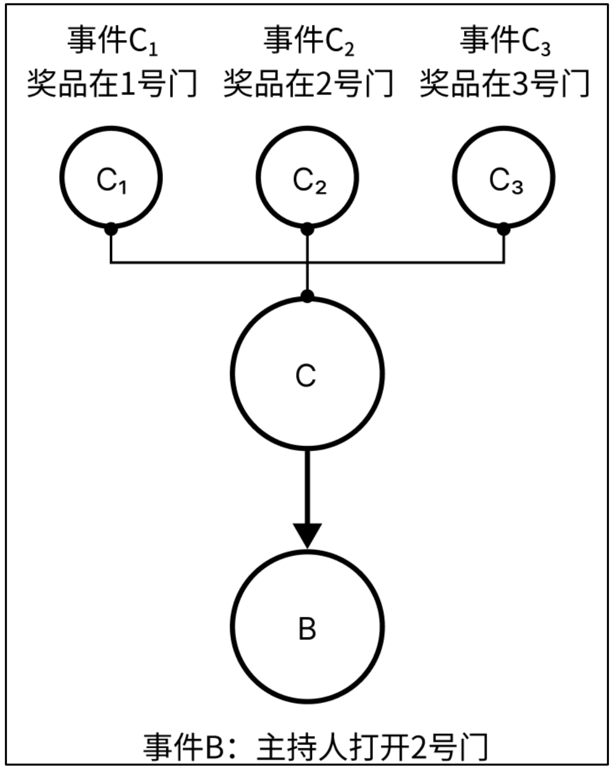
$$P(B)=\sum_{\mathrm{i}=1}^{n}P(B|C_i)P(C_i)=\frac{1}{2}*\frac{1}{3}+0*\frac{1}{3}+1*\frac{1}{3}=\frac{1}{2}$$

- 根据贝叶斯定理“由果溯因”可得：

$$P(C_1|B)=\frac{P(B|C_1)P(C_1)}{P(B)}=\frac{\frac{1}{2}*\frac{1}{3}}{\frac{1}{2}}=\frac{1}{3}$$

$$P(C_2|B)=\frac{P(B|C_2)P(C_2)}{P(B)}=\frac{0*\frac{1}{3}}{\frac{1}{2}}=0$$

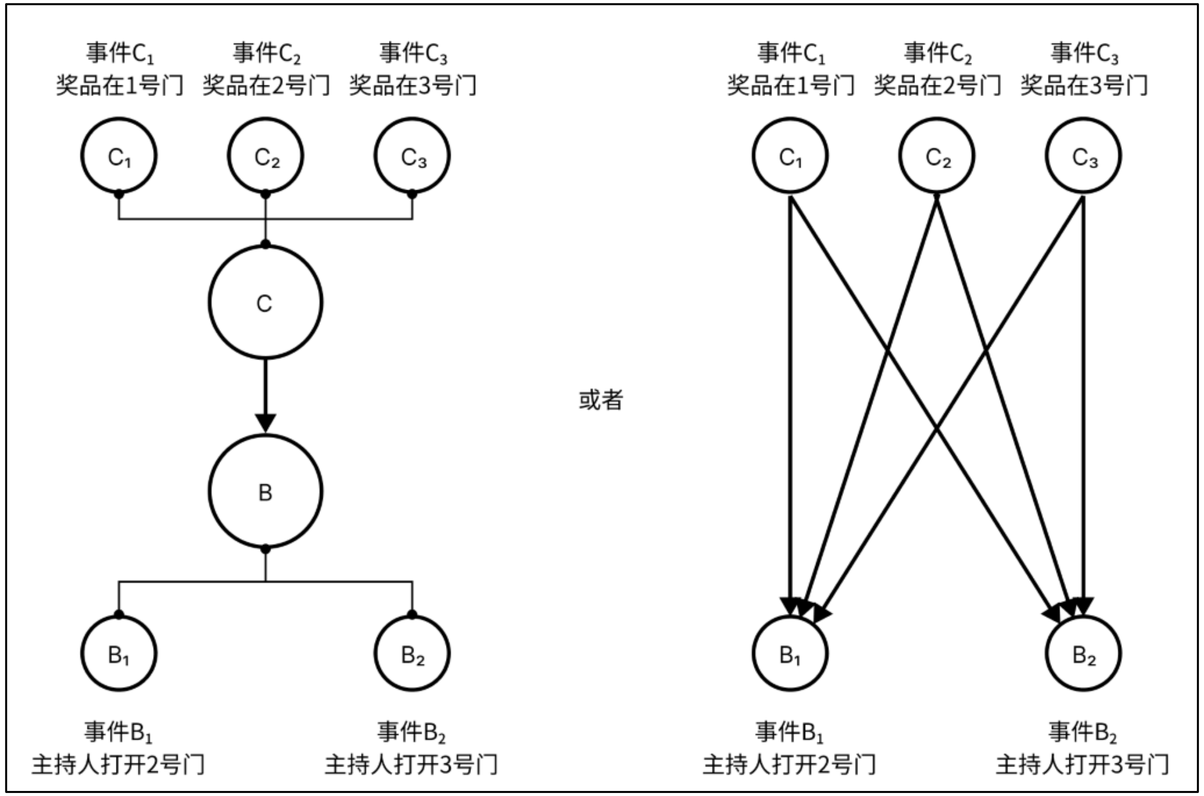
$$P(C_3|B)=\frac{P(B|C_3)P(C_3)}{P(B)}=\frac{1*\frac{1}{3}}{\frac{1}{2}}=\frac{2}{3}$$



从蒙提霍尔悖论(三门问题)理解贝叶斯模型

蒙提霍尔悖论的贝叶斯论证

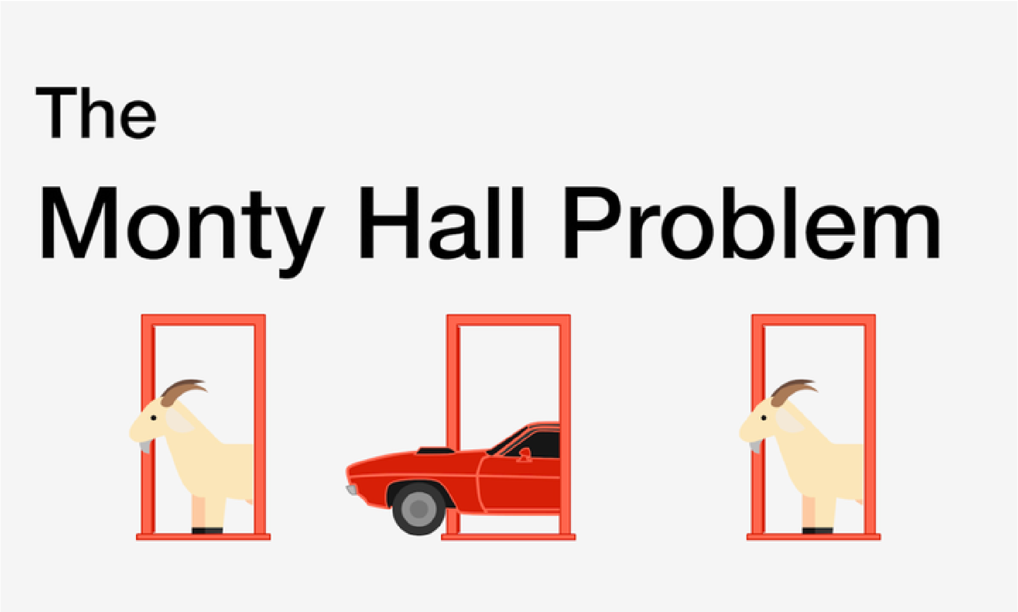
- 理论上这个模型需要更完整



从蒙提霍尔悖论(三门问题)理解贝叶斯模型

启发

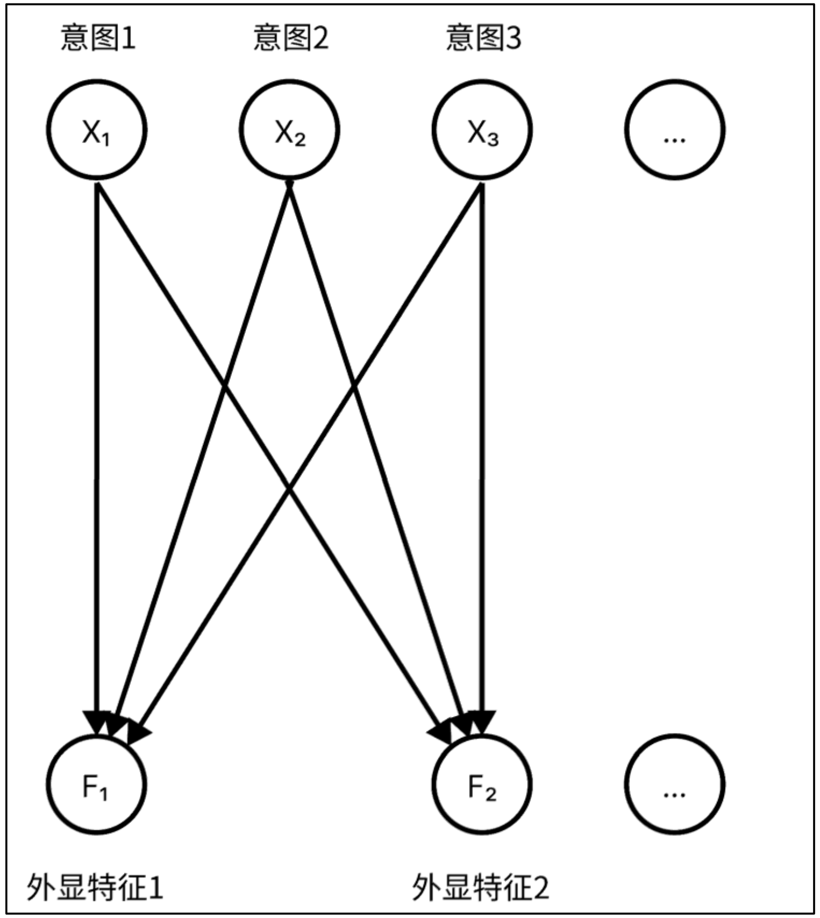
- 蒙提霍尔悖论的局限性：
  - 蒙提霍尔问题单纯从事物发生概率出发，**忽略了人的主观能动性。**反应在三门问题上，则是对于冒险精神的人来说，更倾向于选择低概率事件而实现更高的收益；对于保守的人来说，更倾向于保持客观规律来做选择。
  - 蒙提霍尔问题忽略了**事物对人的影响。**例如我正在制定一份行军路线，有三条路可以选择，其中两条是绝路，即便排除掉一条绝路， 我们也很难仅仅凭借1/3或2/3的概率来做出行动决策，因为行动意味着生或死。因此，依据事物或行动对人产生的影响，**确定是否做出该行动的具体阈值**是及其重要的。



从蒙提霍尔悖论(三门问题)理解贝叶斯模型

启发

- 对意图识别的启示:
  - 蒙提霍尔问题实质上为我们提供了构建贝叶斯网络的最基本模型。对于意图识别问题，实质上便是探究产生某一外显特征所传递的意图是哪一种。我们可以类比三门问题构建贝叶斯网络，如下所示：意图为我们所要探究的“因”，某一意图外显特征（例如按下某个按键、眼睛看向某个控件）为观测值，即“果”。



从蒙提霍尔悖论(三门问题)理解贝叶斯模型

启发

- 对意图识别的启示：
  - 意图之间也并非彼此独立的，意图1和意图2之间或许存在某些时序顺序，它们之间不再是独立事件，而是相关事件，因此我们又可以将上图进一步构建为右图
  - 在基于贝叶斯网络的机器学习模型训练求解过程中，我们已知输入给模型的外显特征、输入给模型的意图、在该外显特征上属于某个意图的概率，即已知事件组A和事件组B的先验概率和实验所收集到的后验概率，**机器学习的训练过程实际上便是求解似然概率的过程。在已知经过机器学习训练求解的似然概率、已知先验知识的基础上，我们便可以将模型应用与预测后验概率中，这便是意图预测/意图识别/意图推理的过程。**