Advances in Education 教育进展, 2017, 7(6), 328-333
Published Online November 2017 in Hans. http://www.hanspub.org/journal/ae
https://doi.org/10.12677/ae.2017.76051

Hans汉斯

Some Famous Problems Solved by Full Probability Formula

Xiaohan Yang

School of Mathematics Science, Tongji University, Shanghai
Email: xiaohyang@tongji.edu.cn

Received: Oct. 19th, 2017; accepted: Nov. 1st, 2017; published: Nov. 8th, 2017

Abstract

Full probability formula is a basic subject of the theory of Probability. By presenting some interesting and famous problems that are applications of this subject instead of mathematics deduction, this paper attempts not only to illustrate how this extremely important formula comes into play but also to let individual feel it is fundamental and awesome to learn probability.

Keywords

Full Probability Formula, Monty Hall Problem, Simpson's Paradox, Sensitivity Analysis

全概率公式解释的经典问题

杨筱菡

同济大学数学科学学院，上海
Email: xiaohyang@tongji.edu.cn

收稿日期：2017年10月19日；录用日期：2017年11月1日；发布日期：2017年11月8日

摘　要

《概率论与数理统计》课程与实际问题联系非常密切，其重要性不言而喻。另一方面，不管是教科书还是学生，在教学和学习过程中都缺乏直接体会概率统计课程重要性的载体。本文尝试以课程中一个非常重要的公式——全概率公式为切入点，收集整理了用全概率公式解释的一些有趣的经典问题，并结合直观的树图讲解，使得学生在轻松掌握全概率公式这个知识点的同时，还有了利用概率统计方法解释现实中经典案例的直观体验，寓教于乐，提高学习积极性。

文章引用：杨筱菡. 全概率公式解释的经典问题[J]. 教育进展, 2017, 7(6): 328-333.
DOI: 10.12677/ae.2017.76051

关键词

全概率公式，蒙提霍尔问题，辛普森悖论，敏感性问题

1. 引言

众所周知，全概率公式是《概率论与数理统计》课程中一个非常重要的公式。在大多数的教科书[1] [2]上，我们能看到详细的关于全概率公式的介绍及公式的推导。纵观以往的文献，也不难发现很多关于完备事件组的分解注释、这个公式的推广及其应用[3] [4] [5] [6]，教案设计、教学方法研究[7]等等，但是很少有文献讨论关于这一知识点的例子选择和收集。我们在教与学的过程中通常都会借助一些例子来加强对数学概念或公式的理解和运用，例如疾病检测就是一个被经常选入教科书的典型例子，因为例子是最直接最有效的学习载体，也是理解知识点的最佳途径。笔者在多年的教学过程中，参考了多本教材，发现全概率公式这一知识点的例子都比较中规中矩，主要注重对全概率公式的讲解和运用，但是相对都比较沉闷，学生在学习过程中缺乏兴趣和动力，主动性不高。因此，在本文中，我们收集整理了三个和全概率公式相关的生动有趣的问题或例子，供学生学习和理解这两个公式时借鉴，同时也能了解一些流传的经典案例，提高学习概率统计的积极性。

为了后续内容介绍的连贯性，首先，我们还是先简单阐述一下全概率公式的定义。

完备事件组的定义：设 E 是随机试验，Ω 是相应的样本空间，$A_1, A_2, \cdots, A_n$ 为事件组，若 $A_1, A_2, \cdots, A_n$ 满足条件：(1) $A_i \cap A_j = \phi (i \neq j)$；(2) $A_1 \cup A_2 \cup \cdots \cup A_n = \Omega$，则称事件组 $A_1, A_2, \cdots, A_n$ 为样本空间的一个**完备事件组**。完备事件组完成了对样本空间的一个分割。同时也完成了对事件 B 的一个分割，见图 1 和图 2。

全概率公式：设 $A_1, A_2, \cdots, A_n$ 为完备事件组，且 $P(A_i) > 0 (i = 1, 2, \cdots, n)$，$B$ 为任一事件，则

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

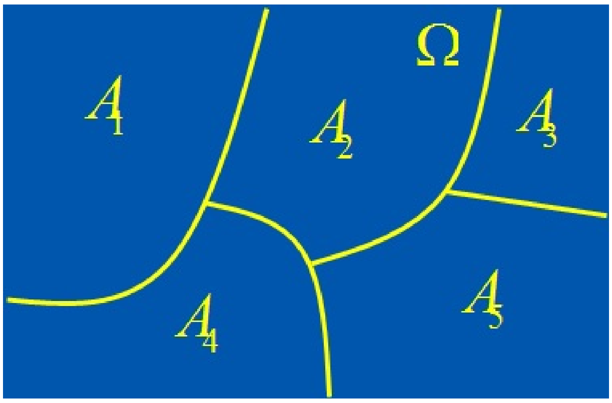

Figure 1. Partition of the sample space
图 1. 完备事件组

例如，当 $n=2$ 时，即为 $P(B)=P(A)P(B|A)+P(\bar{A})P(B|\bar{A})$。下面的树图(图 3)给出了全概率公式的分解。

2. 蒙提霍尔问题(Monty Hall Problem)

这是一个源自博弈论的数学游戏问题。这个概率问题也因为影片“决胜 21 点”中，主角班•坎贝尔(Ben Campbell)成功解开教授米基•罗沙(Mickey Rosa)在课上的提问而非常有名。影片中是这样描述的，有三扇关闭了的门 A、B 和 C，其中一扇门后是一辆汽车(寓意价值高，是奖品)，其他两扇门后各藏有一只山羊(寓意价值很低)，Ben 选了第一扇门 A，然后教授 Michey 把第三扇门 C 打开了，后面是一只山羊。这时候教授 Michey 问 Ben：“你换不换到第二扇门？”Ben 的回答是：换。因为如果不换，赢得汽车的概率是 $\frac{1}{3}$；如果换，赢得汽车的概率将是 $\frac{2}{3}$。

这样的回答似乎感觉上与我们的直观相悖，因为从直观上来说，既然已经知道 C 门后是羊，那么 A 门和 B 门一个后面是汽车，另一个后面是山羊，不管选 A 或 B，选到汽车山羊的概率都是 $\frac{1}{2}$。换句话说，这时候，换或不换，赢得汽车的概率都是 $\frac{1}{2}$。事实上，如果 Ben 先选中的 A 门后是山羊，换后百分之百赢；如果 A 门后是汽车，换后百分之百输。而 A 门后是山羊的概率是 $\frac{2}{3}$，A 门后是汽车的概率是 $\frac{1}{3}$。所以不管怎样都换，相对最初的赢得汽车仅为 $\frac{1}{3}$ 的机率来说，转换选择可以增加赢的机会。

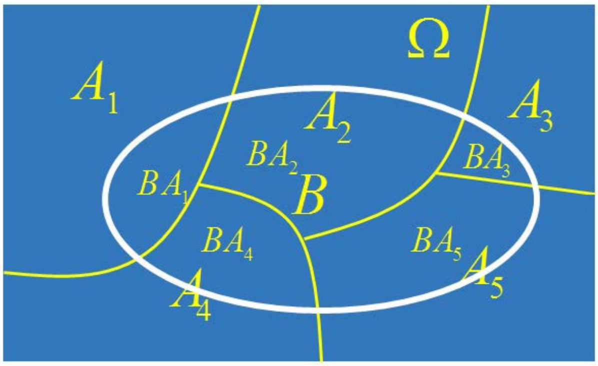

Figure 2. Partition of event B
图 2. 事件 B 的分割

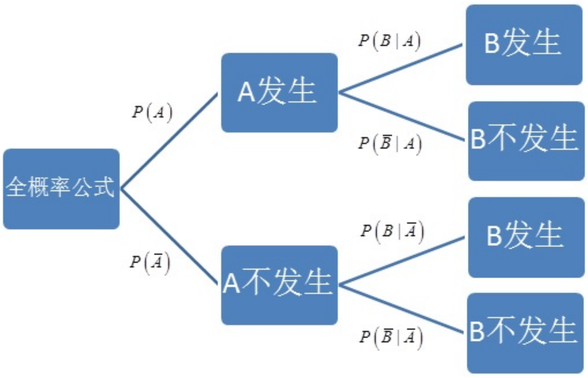

Figure 3. Tree diagram of full probability formula
图 3. 全概率公式的树图分解

关于这个问题，我们可以查询到很多种解释方法，而借助全概率公式的解释是比较容易理解的一种解释方式。首先可以用树图(图 4)来表示两个不同策略及其相应的概率值。

首先设 $A=$“最初选择的门后是汽车”，B 表示“最终赢得汽车”，则由已知条件知，实际情况中汽车在 A 门后的概率是 $\frac{1}{3}$，不在 A 门后的概率是 $\frac{2}{3}$，即

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

策略一：Ben 不换选择，即仍然选择 A 门，则 Ben 能最终赢得汽车的概率，即

$$P(B)=P(A)P(B|A)+P(\bar{A})P(B|\bar{A})=\frac{1}{3}\times 1+\frac{2}{3}\times 0=\frac{1}{3},$$

策略二：Ben 换选择，即换至未开启的 B 门，则 Ben 能最终赢得汽车的概率，即

$$P(B)=P(A)P(B|A)+P(\bar{A})P(B|\bar{A})=\frac{1}{3}\times 0+\frac{2}{3}\times 1=\frac{2}{3},$$

所以，显然，策略二即 Ben 换到未打开的 B 门，他能赢得汽车的概率将比不换增加一倍。

3. 辛普森悖论(Simpson's paradox)

例如，有两种治疗肾结石的方案：方案 1 和方案 2。在接受方案 1 治疗的所有患者中小结石患者占 23%，大结石患者占 77%，小结石患者的治愈率是 93%，大结石患者的治愈率是 73%。在接受方案 2 治疗的所有患者中小结石患者占 67%，大结石患者占 33%，小结石患者的治愈率是 87%，大结石患者的治愈率是 69%。如表 1 所示。

首先，我们发现不管是对小结石患者还是大结石患者，方案 1 的治愈率都要高于方案 2，那么我们能就此判断方案 1 要优于方案 2 吗？

同样设 $A=$“小结石患者”，$B=$“治愈”，

方案 1：由已知条件可知：

$$P(A)=0.23, P(\bar{A})=0.77, P(B|A)=0.93, P(B|\bar{A})=0.73,$$

则根据全概率公式，可得所有接受方案 1 的患者治愈率为：

$$P(B)=P(A)P(B|A)+P(\bar{A})P(B|\bar{A})=0.23\times 0.93+0.77\times 0.73=0.776$$

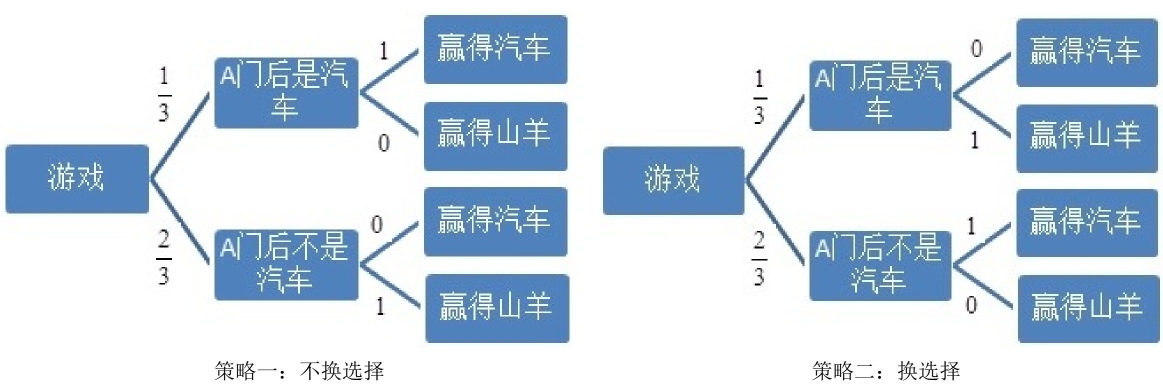

Figure 4. Tree diagram of Monty Hall Problem
图 4. 蒙提霍尔问题策略树图

Table 1. Two treatments for kidney stone
表 1. 两种治疗肾结石的方案

	方案 1		方案 2	
	患者比例	治愈率(B)	患者比例	治愈率(B)
小结石患者(A)	23%	93%	67%	87%
大结石患者($\bar{A}$)	77%	73%	33%	69%

方案 2：由已知条件可知：

$$P(A)=0.67,\ P(\bar{A})=0.33,\ P(B|A)=0.87,\ P(B|\bar{A})=0.69\text{，}$$

则所有接受方案 2 的患者治愈率为：

$$P(B)=P(A)P(B|A)+P(\bar{A})P(B|\bar{A})=0.67\times0.87+0.33\times0.69=0.8106\text{，}$$

所以，方案 2 的患者治愈率要比方案 1 高！这个结论大大出乎我们之前的直观结论。

究其原因，那是因为之前观察数据的时候，我们比较的是每种方案下，不同患者的治愈率，换句话说，我们比较的这些“治愈率”都是条件概率。

如果把不同患者定义成“原因”(A 和 $\bar{A}$)，治愈定义成“结果”(B)。也可以说，我们比较的是，在已知不同“原因”发生的条件下，“结果”发生的概率。而通过全概率公式的计算，最终我们只是比较“结果”发生概率的大小，这是综合了所有“原因”以后的一个结论。而各个“原因”在全概率公式计算中占有的权重直接影响了最终的概率结论，发生了所谓的“悖论”的出现！

4. 敏感性问题调查(sensitivity analysis)

对于考试作弊，赌博，偷税漏税，酒后驾车等一些涉及个人隐私或利害关系，不受被调查对象欢迎或感到尴尬的敏感问题，即使做无记名的直接调查，也很难消除被调查者的顾虑，他们极有可能拒绝应答或故意做出错误的回答，很难保证数据的真实性，使得调查的结果存在很大的误差。如何设计合理的调查方案，来提高应答率并降低不真实回答率呢？基于全概率公式的调查方案设计就能解决这个问题。

调查方案设计的基本思想是，让被调查者从

问题 1：你在考试中曾经作弊过吗？

问题 2：你生日的月份是奇数吗？(约定一年有 365 天)

这两个问题中，随机地选答其中一个，同时调查者并也不知情被调查者回答的是哪一个问题，从而保护被调查者的隐私，消除被调查者的顾虑，能够对自己所选的问题真实地回答。

调查者准备一套 13 张同一花色的扑克，在选答上述问题前，要求被调查的学生随机抽取一张，看后放回，调查者并不知道学生抽取的情况。约定如下：如果学生抽取的是不超过 10 的数则回答问题 1；反之，则回答问题 2。假定调查结果是收回 400 张有效答卷，其中有 80 个学生回答“是”，320 个学生回答“否”，求被调查的学生考试作弊的概率 p 。

以 A 表示选答问题 1， B 表示回答“是”， $P(B|A)=p$ ，则由已知条件知：

$$P(A)=\frac{10}{13},\ P(\bar{A})=\frac{3}{13},\ P(B|\bar{A})=\frac{184}{365}\text{，}$$

由全概率公式 $P(B)=P(A)P(B|A)+P(\bar{A})P(B|\bar{A})=\frac{10}{13}\times p+\frac{3}{13}\times\frac{184}{365}=\frac{1}{5}$ ，由此可算得 $p=\frac{397}{3650}\approx0.109$ 。

5. 结语

以上三个例子都是可以利用全概率公式来解决的著名经典问题，从全概率公式的讲解来看，简单易懂，相比目前的教材中多以盒子取球或掷骰子为背景的例题来说，趣味性大大增强，不失为课堂教学和活跃气氛的好例子，使得学生能轻松快速掌握全概率公式这个知识点，还有了利用概率统计方法解释现实中经典案例的直观体验，寓教于乐，提高学习积极性。

参考文献 (References)

[1] 茆诗松，程依明，濮晓龙. 概率论与数理统计教程[M]. 北京: 高等教育出版社, 2005.

[2] 李贤平，沈崇圣，陈子腾. 概率论与数理统计[M]. 上海: 复旦大学出版社, 2003.

[3] 张丽，闫善文，刘亚东. 全概率公式与贝叶斯公式的应用及推广[J]. 牡丹江师范学院学报(自然科学版), 2005(1): 15-16.

[4] 庄建红. 全概率公式、贝叶斯公式的推广及其应用[J]. 辽宁省交通高等专科学校学报: 自然科学版, 2003, 5(2): 48-50.

[5] 陈光曙，王新利. 全概率公式的推广及应用[J]. 高等数学研究, 2010, 13(4): 53-55.

[6] 朱凤娟. 全概率公式的启发式教学方法研究[J]. 渤海大学学报(自然科学版), 2008, 29(1): 66-68.

[7] 唐旭晖，李洹岸，段利霞. 全概率公式的推广与应用[J]. 高等数学研究, 2011, 14(4): 51:52.