

贝叶斯公式的教学实践

李宝娣

(陆军特种作战学院 广东 广州 510502)

[摘要] 贝叶斯公式是概率论中的一个重要公式,它是贝叶斯统计推断的基础,其应用十分广泛。在实际教学中,一些老师一般只介绍了贝叶斯公式“执果索因”的内涵,而忽视了其“概率修正,统计推断”的另一个内涵。此处的教学实践,从实例的角度介绍了贝叶斯公式两个方面的内涵,且从贝叶斯公式的引入,贝叶斯公式的理论推导以及理解记忆,到贝叶斯公式的应用,层层递进展开教学,力求体现创新性、高阶性和挑战度。

[关键词] 贝叶斯公式; 执果索因; 概率修正; 教学实践

【DOI】 10.12252/j.issn.2096-6288.2021.12.098

《概率论与数理统计》是高等院校理工类、经管类的重要课程之一。贝叶斯公式是《概率论与数理统计》课程中概率论部分里的一个重要公式。在贝叶斯公式的教学中,学生的痛点在于:一是,感觉公式太复杂,不好记;二是,对贝叶斯公式的内涵理解不透彻,不太会用贝叶斯公式解决实际问题。针对学生的特点,课程教学中采用先回顾上堂课全概率公式的内容,由旧知过渡到新知,学生易于接受。然后分析贝叶斯公式的特点,并进行推导,帮助学生理解记忆公式,并且掌握贝叶斯公式“执果索因”的求法。接下来,从信息的角度,介绍先验概率和后验概率的概念,于实例中讲清贝叶斯公式的另一个内涵“概率修正,统计推断”。最后布置的拓展题,需要查找资料才能解决,具有一定的挑战度。具体的课堂实践如下:

一、温故而知新,贝叶斯公式的引入

上节课,我们学习了全概率公式,一起回顾一下:

设实验E的样本空间为S, A为E的事件, B_1, B_2, \dots, B_n 为S的一个划分,且 $P(B_i) > 0 (i=1, 2, \dots, n)$, 则

$$P(A) = P(A|B_1)P(B_1) + P(A|B_2)P(B_2) + \dots + P(A|B_n)P(B_n). \quad (1)$$

(1)式称为全概率公式。它蕴含了“化整为零,聚零为整”的思想,其概率内涵是“由因索果”,即事件A的概率不好求时,找出引起事件A发生的所有可能原因作为一个划分,由全概率公式求出事件A的概率。

善于思考的同学会反过来想,如果事件A已经发生,那么引起事件A发生的各原因的可能性大小又是多少?也就是人们常说的“执果索因”。

例如,某电子设备厂所用的元件是由三家元件制造厂提供的,根据以往的记录有如下的数据。

元件制造厂	次品率	提供元件的份额
1	0.02	0.15
2	0.01	0.80
3	0.03	0.05

设这三家工厂的产品在仓库中是均匀混合的,且无区别的标志。在仓库中随机地区一只元件,求它是次品的概率。我们用全概率公式可以算出次品的概率。设事件A表示“取到的是一只次品”, $B_i (i=1, 2, 3)$ 表示“所取到的产品是由第i家工厂提供的”。易知, B_1, B_2, B_3 是样本空间为S的一个划分,利用全概率公式可以算出

$$P(A) = P(A|B_1)P(B_1) + P(A|B_2)P(B_2) + \dots + P(A|B_n)P(B_n) = 0.0125.$$

现在,在仓库中随机地取一只元件,若已取到的是次品,要分析此次品出自何厂,也就是求出此次品分别由三家

工厂生产的概率是多少。

分析:“取到次品”这件事A已经发生,要求此次品出自何厂,即求 $P(B_1|A)$, $P(B_2|A)$, $P(B_3|A)$ 。

根据前面所学的知识知道,所求的是条件概率,由条

件概率公式有, $P(B_i|A) = \frac{P(AB_i)}{P(A)}$, 对分子运用乘法公式

$P(AB_i) = P(A|B_i)P(B_i)$, 即有

$$P(B_1|A) = \frac{P(AB_1)}{P(A)} = \frac{P(A|B_1)P(B_1)}{P(A)} \quad (*)$$

$$= \frac{0.02 \times 0.15}{0.0125} = 0.24.$$

同理可得 $P(B_2|A) = 0.64$, $P(B_3|A) = 0.12$. 结果表明,这只次品来自第2家工厂的可能性最大。

事实上,上面的(*)式就是我们今天要学习的贝叶斯公式。

二、贝叶斯公式的理论及其特点

定理: 设实验E的样本空间为S, A为E的事件, B_1, B_2, \dots, B_n 为S的一个划分,且 $P(A) > 0, P(B_i) > 0 (i=1, 2, \dots, n)$,

$$P(B_i|A) = \frac{P(A|B_i)P(B_i)}{\sum_{j=1}^n P(A|B_j)P(B_j)} \quad (2)$$

(2)式称为贝叶斯公式。

公式看起来比较复杂,不好记。事实上,(2)式的左端是一个条件概率,由条件概率的定义有

$$P(B_i|A) = \frac{P(AB_i)}{P(A)} \quad (**)$$

对(**)式右端的分母 $P(A)$ 运用全概率公式(注意到 B_1, B_2, \dots, B_n 为S的一个划分),就有

$$P(A) = P(AB_1) + P(AB_2) + \dots + P(AB_n)$$

$$= P(A|B_1)P(B_1) + P(A|B_2)P(B_2) + \dots + P(A|B_n)P(B_n)$$

对(**)式右端的分子运用乘法公式有 $P(AB_i) = P(A|B_i)P(B_i)$

于是得到贝叶斯公式:

$$P(B_i|A) = \frac{P(A|B_i)P(B_i)}{P(A|B_1)P(B_1) + P(A|B_2)P(B_2) + \dots + P(A|B_n)P(B_n)}$$

$$= \frac{P(A|B_i)P(B_i)}{\sum_{j=1}^n P(A|B_j)P(B_j)}$$

进一步观察贝叶斯公式的右端，发现分子分母的形式是相同的，分子为分母的n项和中的第i项。这个概率是一个比例关系。分母代表A发生的概率，有n个原因都可能导致A发生。贝叶斯公式反映了第i个原因所占的份额，即对A发生的“贡献”大小，因此贝叶斯公式又称为逆概率公式。经过这样的推导与分析，同学们基本上记住了贝叶斯公式，解决了第一个痛点问题。

现在需要解决贝叶斯公式怎么用的问题，那就需要理清贝叶斯公式的内涵。

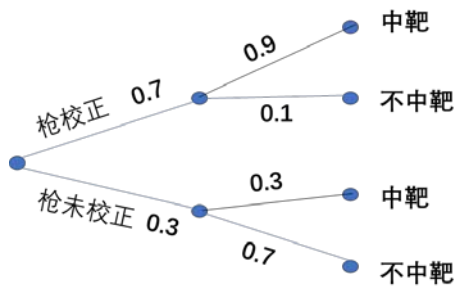
三、贝叶斯公式的内涵及其应用

内涵之一，可以由前面的介绍发现贝叶斯公式的思想是“执果索因”。事件A已经发生（或者假设事件A已经发生），寻找引起事件A发生的各原因的可能性大小。

例1 某军事训练科目中，设射击用枪共10支，其中7支枪经过试射校正，3支枪未经试射校正。一射手用校正过的枪射击，中靶率为0.9，用未校正过的枪射击，中靶率为0.3。若该射手任取一支枪射击，结果未中靶，那么其原因为该枪未校正的概率是多少？

分析：设事件B表示“枪已校正”，事件 \bar{B} 表示“枪未校正”，事件A表示“射击中靶”，事件 \bar{A} 表示“射击未中靶”。事件 \bar{A} “射击未中靶”已经发生，求 $P(\bar{B}|\bar{A})$ 。

概率树如下图所示



易知，事件B和 \bar{B} 是样本空间的一个划分。

依题意有 $P(B)=0.7, P(\bar{B})=0.3,$
 $P(A|B)=0.9, P(A|\bar{B})=0.3,$ 易得

$P(\bar{A}|B)=0.1, P(\bar{A}|\bar{B})=0.7,$ 由贝叶斯公式有

$$P(\bar{B}|\bar{A}) = \frac{P(\bar{A}|\bar{B})P(\bar{B})}{P(\bar{A}|\bar{B})P(\bar{B}) + P(\bar{A}|B)P(B)}$$

$$= \frac{0.7 \times 0.3}{0.7 \times 0.3 + 0.1 \times 0.7} = 0.75$$

结果表明，本次射击没有中靶，有75%的可能是因为枪未校正的原因。所以我们看事情不能只看结果，还得分析事情发生的原因，才能客观公正地处理问题。

内涵之二，从信息的角度来理解贝叶斯公式，即贝叶斯公式“概率修正，统计推断”的思想。通常把贝叶斯公式中 $P(B_i)(i=1,2,\dots,n)$ 称为先验概率，它一般是由以往经验和历史数据经分析得到的概率值，带有主观性。把 $P(B_i|A)$ 称为后验概率，它是基于新的信息，修正原来的先验概率后所获得的概率估计，其结果更接近于实际情况。先验概率和后验概率是相对的，如果以后还有新的信息引入，更新了现在所谓的后验概率，得到了新的概率值，那么这个新的概率值被称为

后验概率。

从信息角度来说，贝叶斯公式把先验信息和样本信息很好地结合在后验概率里面，使得统计推断更加站得住脚。贝叶斯公式为利用搜集到的信息对原有的判断进行修正提供了有效手段。下面看一个实例。

例2（贝叶斯公式在行军路线选择中的应用）

设想在一次战役中，某炮兵营接到上级命令，需在指定时间内运动到某指定位置以达到有效的毁伤打击效果。指挥员根据实地考察结果结合以往经验，迅速绘制出了甲、乙、丙三条可行的线路，其中甲路线是最优方案，乙和丙路线为备选方案。如果用数字来衡量，那么甲、乙、丙三条路线的选择倾向分别为46%，33%，21%。

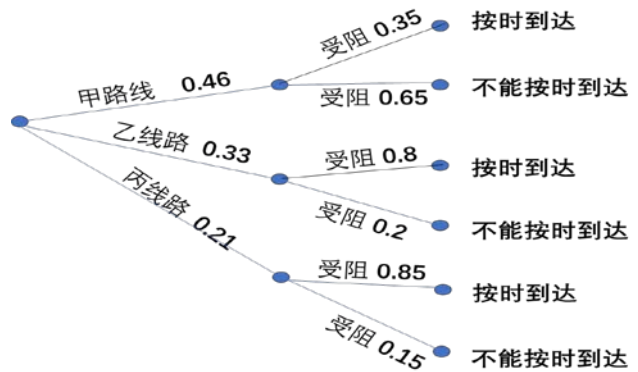
我们知道行军路线的选择受到任务性质、作战人员状态、武器装备性能、道路通行条件、敌情威胁等多个因素的影响。现根据最新情报，经综合研判，指挥员发现在甲、乙、丙三条线路中受到敌人的阻扰而不能按时到达的可能性大小分别为65%，20%，15%，那么现在，应如何决策？

分析：是继续选择甲路线？还是选择受敌阻扰最小的丙路线？

考虑：我们的目的是按时到达。当情况有变时，分别计算按时到达的前提下，选哪条路线的可行性最大。

设事件A表示“按时到达终点”，事件 \bar{A} 表示“不能按时到达终点”，事件 B_1, B_2, B_3 分别表示“选择甲、乙、丙三条路线”，求 $P(B_i|A), i=1,2,3.$ 根据 $\max_i P(B_i|A)$ 来选择路线。

概率树如下图所示



由题意知

$$P(B_1)=0.46, P(B_2)=0.33, P(B_3)=0.21.$$

$$\text{且 } P(\bar{A}|B_1)=0.65, P(\bar{A}|B_2)=0.20, P(\bar{A}|B_3)=0.15.$$

由上面这些概率值可以算出：

$$P(A|B_1)=1-P(\bar{A}|B_1)=1-0.65=0.35,$$

$$P(A|B_2)=0.80, P(A|B_3)=0.85$$

由贝叶斯公式有

$$P(B_1|A) = \frac{P(B_1)P(A|B_1)}{P(B_1)P(A|B_1) + P(B_2)P(A|B_2) + P(B_3)P(A|B_3)}$$

$$= \frac{0.46 \times 0.35}{0.46 \times 0.35 + 0.33 \times 0.8 + 0.21 \times 0.85} = \frac{0.161}{0.6035} = 26.7\%$$

$$P(B_2|A) = \frac{P(B_2)P(A|B_2)}{P(B_1)P(A|B_1) + P(B_2)P(A|B_2) + P(B_3)P(A|B_3)}$$

$$= \frac{0.33 \times 0.8}{0.46 \times 0.35 + 0.33 \times 0.8 + 0.21 \times 0.85} = \frac{0.264}{0.6035} = 43.7\%$$

$$P(B_3|A) = \frac{P(B_3)P(A|B_3)}{P(B_1)P(A|B_1) + P(B_2)P(A|B_2) + P(B_3)P(A|B_3)}$$

$$= \frac{0.21 \times 0.85}{0.46 \times 0.35 + 0.33 \times 0.8 + 0.21 \times 0.85} = \frac{0.1785}{0.6035} = 29.6\%$$

通过比较, 知道选择乙路线更加合适。

在本例中, 由于对事件A的发生提供了新的信息, 通过贝叶斯公式对先验概率 $P(B_1), P(B_2), P(B_3)$ 进行了一种修正, 得到了更加符合实际的后验概率 $P(B_1|A), P(B_2|A), P(B_3|A)$ 。随着新信息的不断增加, 可以应用贝叶斯公式作进一步调整, 就会得到更加精确的判断。

四、小结与拓展

小结: 贝叶斯公式本质上是一个条件概率, 可以由乘法公式和全概率公式推导而得到。贝叶斯公式可以从两个方面来理解, 一是“执果索因”, 二是“增加信息, 修正概率”。利用贝叶斯公式计算概率, 可以先画出概率树, 然后再计算。

同时, 贝叶斯公式的应用十分广泛, 它在查找故障源、

疾病诊断、人工智能、大数据、海难搜救、垃圾邮件过滤等方面都有着应用。

拓展题: 请同学们查找相关资料, 说明贝叶斯公式在海难、空难搜救中是如何应用的?

五、教学反思

贝叶斯公式的教学实践表明, 本次课解决了学生的痛点, 同学们记住了贝叶斯公式, 并且对于贝叶斯公式的内涵有着比较深刻的认识, 知道怎么运用贝叶斯公式解决实际问题。教学中的例子与学生的专业比较贴合, 学生学习兴趣比较浓厚。同时在授课过程中无声地融入了思政点: 一是客观公正地看待事物; 二是新的知识的获得可以由已知的知识来推导迁移, 训练了科学思维; 三是, 贝叶斯公式先验概率与后验概率的概念中蕴含了科学研究中的“大胆假设, 小心求证”的思想(如拓展题中的海难、空难搜救)。

参考文献

- [1] 盛骤, 谢式千, 潘承毅. 概率论与数理统计(第五版)[M]. 高等教育出版社, 2019. 12.
- [2] 李自勇. 基于贝叶斯公式及应用教学的认识与实践[J]. 数学教学研究, 2014, 3(3): 63-65.
- [3] 任芳玲, 刘瑞. 全概率公式和贝叶斯公式教学新探[J]. 西昌学院学报. 自然科学版, 2015, 3(1): 14-16.
- [4] 徐鹏, 柳福祥. 基于HPM视角下的贝叶斯公式教学设计[J]. 科教导刊(电子版), 2020, 12下(36): 198-199.
- [5] 方红燕, 王蕊, 杨文志, 胡夫涛. 贝叶斯公式实例的深度挖掘[J]. 曲阜师范大学学报, 2018, 10(4): 1-4.

(上接第60页)

6. 充分利用现代互联网技术。计算机设备的广泛应用, 在对于企业的发展发挥了巨大推进的同时, 也在某种程度上激发了诸多企业之间的市场竞争矛盾, 当然人力资源的竞争亦在其中, 现代企业必须要充分的利用好现代互联网技术, 依托于上述基础之上建立一个全面的人力资源数据库, 数据库需涵盖企业员工信息以及编外的精英人才信息, 企业在进行人才选拔任用的时候可以借助于上述系统数据进行辅助决策。

7. 重视人事档案管理工作。人事档案管理工作做得好, 人力资源管理的各项业务活动的顺利开展才有保障。人力资源管理即有“全面服务”的性质, 那么在档案收集、档案整理、档案保存等方面下足功夫, 就可以说是企业现代化建设与多元化服务模式成功构建的关键所在了。人事档案管理工作人员应该深刻认知到人事档案管理的重要意义, 扎实做好人事数据的整理、核对、归档等工作, 提升人事档案管理水

平。

总之, 实现信息化人力资源管理的重要性不言而喻, 我们付出更多的努力, 成功构建信息化的人力资源管理模式, 对于企业的长期发展、综合竞争力的提升来说大有帮助。组织内部也有凝聚力, 真正价值创造出来后, 继续执行组织管理、人才管理、绩效管理等任务, 相关理论不断加强, 实践方面就会有新的突破。

参考文献

- [1] 朱江萍. 浅析信息化人力资源管理工作的研究进展. 2020.
- [2] 刘兰凤. 关于信息化人力资源管理研究进展探析. 2021.
- [3] 王艳玲. 信息化人力资源管理研究进展分析[J]. 现代营销(下旬刊), 2019(2): 181-182.