

# 北京邮电大学

## 概率论期中论文



题目：贝叶斯定理与推断

姓 名 耿英盛

学 院 信息与通信工程学院

专 业 通信工程

班 级 2020211121

学 号 2020210477

班内序号 25

指导教师 闫祥伟

2022 年 5 月

# 目 录

0.1	人类推断的代表性偏差 .....	1
0.2	医学推断 .....	2
0.3	机器推断中的朴素贝叶斯算法 .....	4
0.4	总结与思考 .....	5

## 0.1 人类推断的代表性偏差

心理学家 Amos Tversky 和 Daniel Kahneman 曾经做过这样一个实验,他们向实验者介绍了一个叫做 Steve 的人,<sup>[2]</sup> 他是一个害羞,性格孤僻,虽然乐于助人,但却对周围的人或现实世界不太感兴趣的人。他温顺且井井有条。他做什么事都有所调理,结构清晰,而且他也热衷于钻研细节。那么 Steve 是一个图书管理员还是一个农民?

实验者大多选择了 Steve 是一个图书管理员。

但让我们从概率论的角度来看一下这个问题。假设上述特征为  $P$ , 设事件  $E$ : Steve 具有特征  $P$ 。事件  $H_2$ : Steve 是一个农民,  $H_1$ : Steve 是一个图书管理员。我们所求的即为  $P(H_1|E)$  与  $P(H_2|E)$  的较大值

假定在人群中,农民与图书管理员的人数比为 20 : 1, 40% 的图书馆管理员符合特征  $P$ , 而只有 10% 的农民具有特征  $P$ 。

根据上述信息我们可知

$$P(H_1) = \frac{1}{21}, P(H_2) = \frac{20}{21}$$

$$P(E|H_1) = 0.4, P(E|H_2) = 0.1$$

分别代入贝叶斯公式

$$\begin{aligned} P(H_1|E) &= \frac{P(E|H_1) \cdot P(H_1)}{P(E)} \\ &= \frac{P(E|H_1) \cdot P(H_1)}{P(E|H_1) \cdot P(H_1) + P(E|H_2) \cdot P(H_2)} \\ &= \frac{0.4 \times \frac{1}{21}}{0.4 \times \frac{1}{21} + 0.1 \times \frac{20}{21}} \\ &= \frac{1}{6} \approx 0.17 \\ P(H_2|E) &= 1 - P(H_1|E) = \frac{5}{6} \approx 0.83 \end{aligned}$$

如果按照我们的假设来看,实际上这个人时农民的概率反而更高。即使我们认为符合这个描述的人是一个图书馆管理员的可能性是是一个农民的 4 倍,也抵不过农民的数量很多。

这个问题的重点不在于人们是否判断正确以及人们是否知道关于两种人群的比例信息. 而是人们完全没有考虑利用先前信息这件事,即人类判断非理性的一种表现。

这种判断非理性的表现被称为认知的代表性偏差。即我们总是由当前证据的代表性特征而决定我们的判断,而忽视了其他信息。具体来说就是人们在看到一定的证据时,往往只是只是比较了条件与他们认识中类别的代表性特征与证据之间相似的概率,也就是我们假设中的农民与图书管理员符合特征  $P$  的概率,而漏掉了类别中之间的比例信息。而条件概率和贝叶斯定律说明的就是,我们在获得新信息时,这个信息的作用有限

制了我们做判断的范围，新信息不能决定你的看法，而是应该结合过去的经验更新你的判断，

## 0.2 医学推断

在医学推断中也有这样一类现象。假设你最近去看了医生，并决定检查一下自己有没有得一种罕见的疾病。如果你很不幸地收到了阳性的结果，你可能最想知道的是“已知这个检查结果，我真的得了这种病的概率是多少？”

假设该疾病为  $D$ ，一个人患有疾病  $D$  的概率为  $\frac{1}{10000}$ 。在医学中我们往往使用特异度 (Specificity) 和灵敏度 (Sensitivity) 来反映一种检测方式的好坏。

灵敏度指的是如果你患有疾病，那么经过检测后，结果显示为阳性的概率。

特异度指的是如果你不患有疾病，那么经过检测后，结果现实为阴性的概率。

假设你所接受的检测方式中灵敏度为 99%，特异度为 99%。

设事件  $E$ ：你的诊断结果为阳性。事件  $H$ ：你患有疾病  $D$ 。

根据上述信息我们可知

$$P(H) = 10^{-4}$$

$$P(E|H) = 0.99$$

$$P(\bar{E}|\bar{H}) = 0.99$$

$$P(E|\bar{H}) = 1 - P(\bar{E}|\bar{H}) = 1 - 0.99 = 0.01$$

$$P(\bar{H}) = 1 - P(H) = 1 - 10^{-4}$$

代入贝叶斯公式

$$\begin{aligned} P(H|E) &= \frac{P(E|H) \cdot P(H)}{P(E)} \\ &= \frac{P(E|H) \cdot P(H)}{P(E|H) \cdot P(H) + P(E|\bar{H}) \cdot P(\bar{H})} \\ &= \frac{0.99 \times 10^{-4}}{0.99 \times 10^{-4} + 0.01 \times (1 - 10^{-4})} \\ &= \frac{1}{102} \approx 0.0098 \end{aligned}$$

可见这个概率相当的低。

因此即使你使用检测精度十分高的检测方法检测某种疾病，并不幸被检测为阳性，也不必过分担心，在考虑人群比例后，我们才可以清晰的做出判断。

我们可以看到，在上述计算中，我们出现了多次对立事件的概率运算，那么在医学中，我们为什么要使用这样的数值呢？

实际上, 如果将  $P(E|H)$  改变为  $\frac{P(H|E)}{P(\bar{H}|E)}$ , 即已知检测阳性后患病与不患病的概率比  $p(H)$  改变为  $\frac{P(H)}{P(\bar{H})}$  即人群中患病与不患病的人数比。

我们可以将贝叶斯公式改写为比值的形式, 即

$$\frac{P(H|E)}{P(\bar{H}|E)} = \frac{\frac{P(HE)}{P(E)}}{\frac{P(\bar{H}E)}{P(E)}} = \frac{P(HE)}{P(\bar{H}E)} = \frac{P(H)}{P(\bar{H})} \cdot \frac{P(E|H)}{P(E|\bar{H})}$$

若令

$$O(H|E) = \frac{P(H|E)}{P(\bar{H}|E)}, O(D) = \frac{P(H)}{P(\bar{H})}$$

$$Bayesfactor = \frac{P(E|H)}{P(E|\bar{H})} = \frac{\text{灵敏度}}{\text{特异度}}$$

则可写为

$$O(H|E) = O(D) \times Bayesfactor$$

使用这种形式, 我们就可以根据人群中的患病比例简单的计算阳性检测者患病的概率。我们先使用该公式来计算一下第一次诊断阳性后, 患病与不患病的比率。

$$O(D) = \frac{P(H)}{P(\bar{H})} = \frac{10^{-4}}{1 - 10^{-4}} = \frac{1}{9999}$$

$$Bayesfactor = \frac{P(E|H)}{P(E|\bar{H})} = \frac{\text{灵敏度}}{1 - \text{特异度}} = \frac{0.99}{1 - 0.99} = 99$$

$$O(H|E) = \frac{P(H|E)}{P(\bar{H}|E)} = O(D) \times Bayesfactor$$

$$= \frac{1}{9999} \times 99 = \frac{1}{101}$$

或者我们可以换算为概率

$$P(H|E) = \frac{1}{1 + 101} = \frac{1}{102} \approx 0.0098, \text{与前述计算一致}$$

我们还可以使用该公式来计算一下复检 (第二次诊断) 阳性后, 患病与不患病的比率我们可以换一种特异度更高的检测方式。这种检测方式中灵敏度为 99%, 特异度为 99.9%。

$$O(D') = O(H|E)$$

$$Bayesfactor = \frac{P(E|H)}{P(E|\bar{H})} = \frac{\text{灵敏度}}{1 - \text{特异度}} = \frac{0.99}{1 - 0.999} = 990$$

$$O(H|E') = O(D') \times Bayesfactor = \frac{1}{101} \times 990 = \frac{990}{101} \approx 9.8$$

可见在第二次这个比值已经大幅度上传，此时患病的比例已经接近一半。显然如果再次复检，这个比值将不断增大，即患病的概率不断增加，我们也就越可以相信我们是否患病。

这种写法除了计算上的优势外，可以分开贝叶斯公式中先验的部分与后验的部分。即  $O(D)$  代表我们预先知道的信息，即该病的患病率， $Bayesfactor$  则代表后验信息，即我们在使用了某种检测方式后所带来的信息。增加后验信息就是在乘上一个又一个贝叶斯因子。

这也可以再次更加清晰地反映贝叶斯公式的思想，即新信息的作用是缩小我们的判断的范围，新信息不能决定你的看法，而是应该结合过去的经验更新你的判断。

### 0.3 机器推断中的朴素贝叶斯算法

目前 AI 的热潮正在引领时代，在早期的 AI 算法中，我们使用朴素贝叶斯算法 (Naive Bayes Algorithm) 来充当垃圾邮件分类器，即根据邮件所含有的信息将邮件分为垃圾邮件与非垃圾邮件两个类别。

将这个问题抽象化，即我们希望根据数据的信息将数据分辨为不同的类别，实际上是在做已知数据的特征，利用这些特征来获取数据在各个类别下出现的概率，选择最大的概率作为该数据的类别。假设类变量为  $y$ ，特征向量为  $x_i (i = 1, 2, \dots, n)$  则我们的目标为  $\max[p(y|x_1, x_2, \dots, x_n)]$  利用贝叶斯公式，我们可以把他转换为

$$P(y | x_1, \dots, x_n) = \frac{P(y)P(x_1, \dots, x_n | y)}{P(x_1, \dots, x_n)} \quad \text{式 (0-1)}$$

由于我们往往难以获得  $P(x_1, \dots, x_n|y)$  我们引入特征向量  $x_i|y (i = 1, 2, \dots, n)$  之间独立同分布的假设来化简公式 0-1 结果如 0-2

$$y = f(x) = \arg \max_{y_k} \frac{P(y) \prod_i P(x_i | y)}{P(y) \prod_i P(x_i | y)} \quad \text{式 (0-2)}$$

同时由于之间同分布，则对于  $y$  的每一个取值，每个输入的数据中  $P(x_1, \dots, x_n) =$

$P(y) \prod_i P(x_i | y)$  为相同值则我们的目标为

$$\begin{aligned} P(y | x_1, \dots, x_n) &\propto P(y) \prod_{i=1}^n P(x_i | y) \\ &\Downarrow \\ \hat{y} &= \arg \max_y P(y) \prod_{i=1}^n P(x_i | y) \end{aligned} \quad \text{式 (0-3)}$$

其中我们可以使用最大似然法求得参数。最后实现信息的分类判别。我们不对其参数做出详细说明，我们关注的重点在于实现的过程，在这个过程中我们同样将这样一个判断问题转化为由先验概率  $P(y)$  与后验概率类别的分布比例以及新信息的情况来进行判断。在垃圾邮件的分类问题中，我们将每一个单独的词作为数据特征，虽然在这种情况下，每个特征独立同分布的假设不太成立，因为我们往往通过词语或语句的组合来表达含义，但即使是这样，在具有一定量训练数据的情况下，我们也可以取得很高的准确度，可见这样一种推断方法确实反映了一定现实世界的规律。

## 0.4 总结与思考

由前面的叙述可知，贝叶斯公式不仅给予了我们求解条件概率的一种方法，更带给我们的是一种思维方式上的改变，即获得的证据不能决定我们的判断，而是要与我们先前的经验结合来更新我们的判断，这种方法无论从我们的生活学习，还是从科学发展，人类进步上都具有不可忽视的意义。贝叶斯定理确实反映了人类的判断偏差，帮助人们更好的认识自己，并做出更好的决策。

但从中我们也可以看到当前社会的一些问题我们进行贝叶斯式的决策，其中无论是先验信息还是后验信息，都脱离不了信息本身的数量，准确度以及分布，就像我们先前所作地种种假设。

在互联网时代，商家为了使得用户感到愉悦，利用基于贝叶斯定理的推荐系统，根据用户的喜好，经常推送用户经常观看或是使用户愉悦信息，并随着用户的点击逐渐更新参数。根据贝叶斯定理可以看出，平台根据用户的先验喜好推荐物品，用户的满意度又决定了后验概率，这两部分又决定了下一次的推荐，这样显然构成了一个正反馈，即随着平台的推荐，用户的满意度越来越高，即后验概率越来越高，平台先验的用户喜好叶越来越高，最后导致用户接收到的信息几乎构成了一个封闭的“回音室，”这一现象被称作信息茧房。

这样的推荐机制大大压缩了人们接触到不同信息的机会，即人们获取信息的分布逐渐缩小，

虽然人们不断接受着丰富的信息，觉得自己随时随地了解着世界。但实际上这不过是因为自己总能看到感兴趣的东西，从而获得了舒适的错觉，在这种情况下使用者喜欢的信息被得到不断重复，个人既有的价值取向和认知也在不断被强化这使得个人所获得的数据分布越来越窄，人们对于某一问题之间的看法也会有着巨大差异，使得人与人之间的隔阂越来越大。

因此优化推荐系统机制，尽可能地让人全方位地接触信息，形成正确地人生观，世界观，价值观需要提各大互联网企业上议程。我们每个人也要对信息产生警惕，对于一件事不要妄下判断。我们需要时刻告诫自己，有信息被我们遗漏了，我们看到的只是事物的一个侧面。