DOI:10.11721/cqnuj20220513

一种新的关于行动决策问题的后验概率的计算方法*

郑 海，王洪春
(重庆师范大学 数学科学学院，重庆 401331)

摘要:【目的】从一个实际的贝叶斯检验例子出发,为克服此类型问题中后验概率计算的复杂性,提出改进计算方法。【方法】通过数值实验引出改进后的计算方法,并在理论上证明了改进方法的可行性与准确性。【结果】在该方法的基础上加以推广使它适用更复杂的情况,最后的算例分析进一步表明该方法的推广是成立的。【结论】基于本文提出的后验概率的计算方法,不仅未使最后的计算结果发生改变,而且使得计算的步骤得到了极大的简化。

关键词:贝叶斯推断;先验分布;后验分布;行动决策
中图分类号:O212.8　　**文献标志码**:A　　**文章编号**:1672-6693(2022)05-0110-08

贝叶斯统计作为统计学派的一个分支,最早由英国学者贝叶斯提出,后被统计学者们发展为一种系统的统计推断方法。贝叶斯统计方法常被应用于分类[1-4]、预测[5-8]、决策[9-10]、控制工程[11]等。贝叶斯统计在考虑统计推断问题时除了利用抽样信息以外,还利用参数的先验信息,当先验信息足够多时,可获得先验分布[12]。贝叶斯统计在面对决策问题时,常常利用后验概率最大的类作为输出,如经典的朴素贝叶斯[13],它在解决决策问题时运算效率往往较高。

当贝叶斯理论用于解决包含复杂步骤的产品检验一类实际问题时,为提高产品的质量和决策的可信度并减少检验的偶然性,一般采取分步检验[14]。这类序贯分析的方法在控制工程中是常见的,因为使用这种方法可以在每步检验的过程中通过改进技术来增加产品的质量,从而提高经济效益。同时使用此种检验方法要求每一步检验的先验信息都基于前一步检验的后验信息,这是合理的;否则,若每次检验都基于最初的先验信息,那么所谓的先验信息将失去本质含义。

上述情形所面临的实际问题是在做最终决策时会涉及大量且复杂的运算,同时也会消耗较多的财力、物力和时间,而运算的复杂程度与检验步骤的数量是正相关的,这意味着可以通过减少检验步骤来提高做决策的效率,但此时必须承担一定的检验偶然性和低可信度的风险[15]。目前在现有文献中并未检索到合理解决上述问题的方法,基于此,本文鉴于贝叶斯检验过程的复杂步骤导致后验概率计算的复杂性,提出一种仅依赖于抽样总数和不满足要求产品数的贝叶斯推断方法,该方法不仅未影响决策的最终选取,还会大大简化计算过程,提高运算的效率。此外本文不仅在理论上给出了证明,而且将该方法推广至更一般的情形,并以一个实际的例子用于检验该方法的可行性。

本文主要内容安排如下:第 1 节介绍文中涉及的一些预备知识;第 2 节基于文献[16]中的具体实例提出问题,给出了两种不同情况的变式;第 3 节给出了合理解决所提问题的方法,并进一步探讨该方法,使它适用于更复杂情形;在第 4 节中提供一个例子,通过对比两种方法得到的结果,论述本文所提的方法的简洁性;第 5 节总结全文。本文所有实验数据均基于 R 软件计算得出,结果具有可靠性。

1 预备知识

本节主要介绍一些基本定义和贝叶斯推断的基本方法。

* **收稿日期**:2021-07-21　**修回日期**:2022-03-31　**网络出版时间**:2022-09-19 15:01
资助项目:国家社会科学基金(No. 13BTJ008);重庆市教育委员会人文社会科学重点项目(No. 22SKGH081);重庆市教育改革项目(No. 213139)
第一作者简介:郑海,男,研究方向为贝叶斯推断,E-mail:2431425624@qq.com;**通信作者**:王洪春,男,教授,博士,E-mail:wanghc@cqnu.edu.cn
网络出版地址:https://kns.cnki.net/kcms/detail/50.1165.n.20220916.1810.017.html

定义 1[12]　由参数 θ 的先验信息确定的概率分布称为先验分布，用 $\pi(\theta)$ 表示随机变量 θ 的概率函数。

定义 2[12]　在获得样本 X 之后，θ 的后验分布就是给定 $X=x$ 条件下 θ 的条件分布，记为 $\pi(\theta|x)$。它的计算公式为：

$$\pi(\theta|x)=\frac{h(x,\theta)}{m(x)}=\frac{f(x|\theta)\pi(\theta)}{\int_{\theta}f(x|\theta)\pi(\theta)\mathrm{d}\theta},\tag{1}$$

其中：$h(x,\theta)=f(x|\theta)\pi(\theta)$ 为 X 和 θ 的联合概率密度，$m(x)=\int_{\theta}h(x,\theta)\mathrm{d}\theta=\int_{\theta}f(x|\theta)\pi(\theta)\mathrm{d}\theta$ 为 X 的边缘概率分布，当 θ 为离散情形时：

$$\pi(\theta_i|x)=\frac{f(x|\theta_i)\pi(\theta)}{\sum_i f(x|\theta_i)\pi(\theta_i)},i=1,2,\cdots。\tag{2}$$

定义 3　将采取某种决策后所取得的最终效果称为决策效果。

定义 4　在采取某种决策后生产的产品中随机抽样，未达到该决策预期的产品称为不满足要求的产品。

2 问题提出

文献[16]中给出了一个实例：为提高产品质量，决定增加投资来改进设备，现有两种决策。θ_1：改进设备后，高质量产品占 90％；θ_2：改进设备后，高质量产品占 70％。由过去的经验认为 θ_1 的可信度有 40％，θ_2 的可信度有 60％，为规避主观经验的影响，前后做了两次试验 A，B。实验 A：试制 5 个产品，全是高质量。在实验 A 之后再做实验 B：试制 10 个产品，其中 9 个高质量。在本文中，给出上述例题的两种变式情况。

变式 1　将实验 A 和实验 B 合并为一个实验 C：试制 15 个产品，其中优等品 14 个，计算实验 C 情形下两种决策的后验概率 $\pi(\theta_1|\mathrm{C})$，$\pi(\theta_2|\mathrm{C})$并与没发生变化的情况下两种决策的后验概率 $\pi(\theta_1|\mathrm{B})$，$\pi(\theta_2|\mathrm{B})$的值进行比较。

解　根据实验 C 可得，两种决策的先验概率分别为：$\pi(\theta_1)=0.4$，$\pi(\theta_2)=0.6$。而两种决策效果分别为：$Q(\theta_1)=0.9$，$Q(\theta_2)=0.7$。于是，实验 C 在两种决策下的条件概率分别为：$P(\mathrm{C}|\theta_1)=15(0.9)^{14}(0.1)$，$P(\mathrm{C}|\theta_2)=15(0.7)^{14}(0.3)$。

由此可算得：$P(\mathrm{C})=P(\mathrm{C}|\theta_1)\pi(\theta_1)+P(\mathrm{C}|\theta_2)\pi(\theta_2)$。由(2)式算得两种决策的后验概率分别为：

$$\pi(\theta_i|\mathrm{C})=\frac{P(\mathrm{C}|\theta_i)\pi(\theta_i)}{P(\mathrm{C})},i=1,2。$$

在实验 C 的情形下，计算各决策的后验概率所需的步骤由原来的两步变成了一步，也即是原来的实验 B 是需要依赖实验 A 的，而变式之后仅依据总的抽样数和不满足要求产品数即可，将上述计算过程通过 R 软件实现，计算结果如表 1 所示。表 1 中将该结果与文献[16]中的结果进行了对比。

表 1　变式 1 的后验概率值

Tab. 1　The posterior probability of variant 1

不满足要求产品所属分类	$\pi(\theta_1\|\mathrm{B})$	$\pi(\theta_2\|\mathrm{B})$	$\pi(\theta_1\|\mathrm{C})$	$\pi(\theta_2\|\mathrm{C})$
(1,0)	0.882 292 9	0.117 707 1	0.882 292 9	0.117 707 1
(0,1)	0.882 292 9	0.117 707 1	0.882 292 9	0.117 707 1

表中的(1,0)表示在实验 A 中不满足要求产品数为 1，在实验 B 中不满足要求产品数为 0；(0,1)表示在实验 A 中不满足要求产品数为 0，在实验 B 中不满足要求产品数为 1。本文在后面小节用同样的方法表示不满足要求产品的归类。

通过表 1 的具体数据观察发现，将两个实验合成一个实验，抽样总数和不满足要求产品数等于这两次实验的总和，仅通过一步计算得来的后验概率值与之前分为两步进行计算的后验概率的值是一样的，这也就说明了将二者合为一个实验不会对所做的决策产生影响。但变式 1 的情况考虑较为简单，并未很好地体现本文所提出新方法计算的简洁性，且不满足要求产品数仅为 1 可能比较偶然，因此提出以下更复杂的变式 2。

变式 2　仍考虑实验 A 为试制 5 个产品，实验 B 为试制 10 个产品，所得不满足要求产品数如表 2 第 1 列分

类所示,将实验 A,B 合并为实验 C:试制 15 个产品,其中优等品 11 个,计算变式 2 情形下的后验概率 $\pi(\theta_1|\mathrm{C})$,$\pi(\theta_2|\mathrm{C})$并与没发生变化的情况下的后验概率 $\pi(\theta_1|\mathrm{B})$,$\pi(\theta_2|\mathrm{B})$的值进行比较。

未发生变化情形下的后验概率计算步骤仍如文献[16]中所示,不过相应的不满足要求产品的分类会发生改变,变式 2 的计算步骤类似于变式 1,同样相应的不满足要求产品的分类会发生改变,基于 R 程序显示的结果如表 2 所示。

表 2 变式 2 的后验概率值

Tab.2 The posterior probability of variant 2

不满足要求产品所属分类	$\pi(\theta_1\|\mathrm{B})$	$\pi(\theta_2\|\mathrm{B})$	$\pi(\theta_1\|\mathrm{C})$	$\pi(\theta_2\|\mathrm{C})$
(0,4)	0.115 530 3	0.884 469 7	0.115 530 3	0.884 469 7
(1,3)	0.115 530 3	0.884 469 7	0.115 530 3	0.884 469 7
(2,2)	0.115 530 3	0.884 469 7	0.115 530 3	0.884 469 7
(3,1)	0.115 530 3	0.884 469 7	0.115 530 3	0.884 469 7
(4,0)	0.115 530 3	0.884 469 7	0.115 530 3	0.884 469 7

变式 2 情形说明,即使不满足要求的产品数不为 1,分类情况更复杂,将两次实验合为一次实验,最后的后验概率值也没有发生改变。

由表 1 和表 2 的数据可以得出结论:若将实验分成两步进行,且后一步的先验信息为前一步的后验信息时,所计算的后验概率与将两者结合成一步计算所得结果是一致的,并且这里的后一步的先验信息基于前一步的后验信息是合理的。由此可以看出,基于先验信息、抽样总数和不满足要求产品的个数计算后验概率使得计算过程得到了大大的简化,并且结果并未发生改变。本文将在第 3 节给出理论证明,并对分步情况进行进一步推广,使得该方法在更复杂的情况下也适用。

3 理论证明并推广

本节首先将结果以定理或推论的形式给出,然后给出相应的证明。

定理 1 对于两种决策 θ_1,θ_2,设采取这两种决策所取得的决策效果分别是 $Q(\theta_1)$,$Q(\theta_2)$,它们的先验分布分别为 $\pi(\theta_1)$,$\pi(\theta_2)$,若实验 A 抽样检查了 k 件产品,其中有 p 件不满足要求,接着基于实验 A 进行实验 B 抽取 m 件产品,有 q 产品不满足要求,由实验 A 和实验 B 所算得的各决策的后验概率与将二者合成一个实验 C 所算得的后验概率值相等。

证明 根据实验 A 可得:

$$P(\mathrm{A}|\theta_1)=\mathrm{C}_k^p Q(\theta_1)^{k-p}(1-Q(\theta_1))^p,$$

$$P(\mathrm{A}|\theta_2)=\mathrm{C}_k^p Q(\theta_2)^{k-p}(1-Q(\theta_2))^p,$$

$$P(\mathrm{A})=P(\mathrm{A}|\theta_1)\pi(\theta_1)+P(\mathrm{A}|\theta_2)\pi(\theta_2),$$

$$P(\mathrm{A})=\mathrm{C}_k^p\{\pi(\theta_1)Q(\theta_1)^{k-p}(1-Q(\theta_1))^p+\pi(\theta_2)Q(\theta_2)^{k-p}(1-Q(\theta_2))^p\}。$$

从而由(2)式可得:

$$\pi(\theta_1|\mathrm{A})=\frac{\mathrm{C}_k^p Q(\theta_1)^{k-p}(1-Q(\theta_1))^p\pi(\theta_1)}{\mathrm{C}_k^p\{\pi(\theta_1)Q(\theta_1)^{k-p}(1-Q(\theta_1))^p+\pi(\theta_2)Q(\theta_2)^{k-p}(1-Q(\theta_2))^p\}}=$$

$$\frac{Q(\theta_1)^{k-p}(1-Q(\theta_1))^p\pi(\theta_1)}{\pi(\theta_1)Q(\theta_1)^{k-p}(1-Q(\theta_1))^p+\pi(\theta_2)Q(\theta_2)^{k-p}(1-Q(\theta_2))^p}。$$

同理可得:

$$\pi(\theta_2|\mathrm{A})=\frac{Q(\theta_2)^{k-p}(1-Q(\theta_2))^p\pi(\theta_1)}{\pi(\theta_1)Q(\theta_1)^{k-p}(1-Q(\theta_1))^p+\pi(\theta_2)Q(\theta_2)^{k-p}(1-Q(\theta_2))^p}。$$

实验 B 是在实验 A 之后,对于决策 θ_1,θ_2 有了相应的了解,从而不再根据最初始的先验分布,而依据实验 A 所算得的后验分布作为本次实验的先验分布:

$$P(\mathrm{B}|\theta_1)=\mathrm{C}_m^q Q(\theta_1)^{m-q}(1-Q(\theta_1))^q,$$

$$P(\mathrm{B}|\theta_2)=\mathrm{C}_m^q Q(\theta_2)^{m-q}(1-Q(\theta_2))^q,$$
$$P(\mathrm{B})=P(\mathrm{B}|\theta_1)\pi(\theta_1|\mathrm{A})+P(\mathrm{A}|\theta_2)\pi(\theta_2|\mathrm{A})。$$

从而由(2)式可得:

$$\pi(\theta_1|\mathrm{B})=\frac{P(\mathrm{B}|\theta_1)\pi(\theta_1|\mathrm{A})}{P(\mathrm{B})}, \tag{3}$$

$$\pi(\theta_2|\mathrm{B})=\frac{P(\mathrm{B}|\theta_2)\pi(\theta_2|\mathrm{A})}{P(\mathrm{B})}。 \tag{4}$$

其中:

$$P(\mathrm{B}|\theta_1)\pi(\theta_1|\mathrm{A})=\frac{\mathrm{C}_m^q Q(\theta_1)^{m-q}(1-Q(\theta_1))^q Q(\theta_1)^{k-p}(1-Q(\theta_1))^p\pi(\theta_1)}{\pi(\theta_1)Q(\theta_1)^{k-p}(1-Q(\theta_1))^p+\pi(\theta_2)Q(\theta_2)^{k-p}(1-Q(\theta_2))^p},$$

$$P(\mathrm{B}|\theta_2)\pi(\theta_2|\mathrm{A})=\frac{\mathrm{C}_m^q Q(\theta_2)^{m-q}(1-Q(\theta_2))^q P(\theta_2)^{k-p}(1-Q(\theta_2))^p\pi(\theta_2)}{\pi(\theta_1)Q(\theta_1)^{k-p}(1-Q(\theta_1))^p+\pi(\theta_2)Q(\theta_2)^{k-p}(1-Q(\theta_2))^p}。$$

于是由(3)和(4)式可得:

$$\pi(\theta_1|\mathrm{B})=$$
$$\frac{Q(\theta_1)^{m-q}(1-Q(\theta_1))^q Q(\theta_1)^{k-p}(1-Q(\theta_1))^p\pi(\theta_1)}{Q(\theta_1)^{m-q}(1-Q(\theta_1))^q Q(\theta_1)^{k-p}(1-Q(\theta_1))^p\pi(\theta_1)+Q(\theta_2)^{m-q}(1-Q(\theta_2))^q Q(\theta_2)^{k-p}(1-Q(\theta_2))^p\pi(\theta_2)}, \tag{5}$$

$$\pi(\theta_2|\mathrm{B})=$$
$$\frac{Q(\theta_2)^{m-q}(1-Q(\theta_2))^q Q(\theta_2)^{k-p}(1-Q(\theta_2))^p\pi(\theta_2)}{Q(\theta_1)^{m-q}(1-Q(\theta_1))^q Q(\theta_1)^{k-p}(1-Q(\theta_1))^p\pi(\theta_1)+Q(\theta_2)^{m-q}(1-Q(\theta_2))^q Q(\theta_2)^{k-p}(1-Q(\theta_2))^p\pi(\theta_2)}。 \tag{6}$$

实验 C 为两步实验 A,B 合成一个实验,那么抽样总数为 $k+m$,不满足要求的产品数为 $p+q$,为计算方便,令 $n=k+m$,$i=p+q$,则先验分布则分别为 $\pi(\theta_1)$,$\pi(\theta_2)$,于是:

$$P(\mathrm{C}|\theta_1)=\mathrm{C}_n^i Q(\theta_1)^{n-i}(1-Q(\theta_1))^i,$$
$$P(\mathrm{C}|\theta_2)=\mathrm{C}_n^i Q(\theta_2)^{n-i}(1-Q(\theta_2))^i,$$
$$P(\mathrm{C})=P(\mathrm{C}|\theta_1)\pi(\theta_1)+P(\mathrm{C}|\theta_2)\pi(\theta_2),$$
$$P(\mathrm{C})=\mathrm{C}_n^i\{\pi(\theta_1)Q(\theta_1)^{n-i}(1-Q(\theta_1))^i+\pi(\theta_2)Q(\theta_2)^{n-i}(1-Q(\theta_2))^i\}。$$

从而由(2)式可得:

$$\pi(\theta_1|\mathrm{C})=\frac{\mathrm{C}_n^i\pi(\theta_1)Q(\theta_1)^{n-i}(1-Q(\theta_1))^i}{\mathrm{C}_n^i\{\pi(\theta_1)Q(\theta_1)^{n-i}(1-Q(\theta_1))^i+\pi(\theta_2)Q(\theta_2)^{n-i}(1-Q(\theta_2))^i\}}=$$
$$\frac{\pi(\theta_1)Q(\theta_1)^{n-i}(1-Q(\theta_1))^i}{\pi(\theta_1)Q(\theta_1)^{n-i}(1-Q(\theta_1))^i+\pi(\theta_2)Q(\theta_2)^{n-i}(1-Q(\theta_2))^i}。$$

同理可得:$\pi(\theta_2|\mathrm{C})=\dfrac{\pi(\theta_2)Q(\theta_2)^{n-i}(1-Q(\theta_2))^i}{\pi(\theta_1)Q(\theta_1)^{n-i}(1-Q(\theta_1))^i+\pi(\theta_2)Q(\theta_2)^{n-i}(1-Q(\theta_2))^i}$。

将(5)和(6)式化简可得:

$$\pi(\theta_1|\mathrm{B})=\frac{Q(\theta_1)^{m+k-q-p}(1-Q(\theta_1))^{q+q}\pi(\theta_1)}{Q(\theta_1)^{m+k-q-p}(1-Q(\theta_1))^{q+q}\pi(\theta_1)+Q(\theta_2)^{m+k-q-q}(1-Q(\theta_2))^{q+p}\pi(\theta_2)}, \tag{7}$$

$$\pi(\theta_2|\mathrm{B})=\frac{Q(\theta_2)^{m+k-q-p}(1-Q(\theta_2))^{q+q}\pi(\theta_2)}{Q(\theta_1)^{m+k-q-p}(1-Q(\theta_1))^{q+q}\pi(\theta_1)+Q(\theta_2)^{m+k-q-q}(1-Q(\theta_2))^{q+p}\pi(\theta_2)}。 \tag{8}$$

又由 $n=k+m$,$i=p+q$,可得 $\pi(\theta_1|\mathrm{B})=\pi(\theta_1|\mathrm{C})$,$\pi(\theta_2|\mathrm{B})=\pi(\theta_2|\mathrm{C})$,从而定理 1 得证。　　证毕

事实上,定理 1 的逆命题也是正确的,即将实验 C 分成两步进行,保持两步之和的总数与总的不满足要求产品数不变,且第二步先验信息为第一步后验信息,所算得的后验概率值不发生改变,这从上述的证明过程可以看出。应当指出此逆命题仅具有理论意义不具有实际意义,因为当抽样检查完成之后再进行细分不符合实际。

推论 1　对于 $n(n>2)$ 种决策 $\theta_1,\theta_2,\cdots,\theta_n$,设采取这 n 种决策所取得的决策效果分别是 $Q(\theta_1),Q(\theta_2),\cdots,Q(\theta_n)$,它们的先验分布分别为 $\pi(\theta_1),\pi(\theta_2),\cdots,\pi(\theta_n)$,若实验 A 抽样检查了 k 件产品,其中有 p 件不满足要

求,接着基于实验A进行实验B抽取 m 件产品,有 q 产品不满足要求,由实验A和实验B所算得的各决策的后验概率与将二者合成一个实验C所算得的后验概率值相等。

证明 对于 n 种决策,根据实验A可得:

$$P(\mathrm{A}|\theta_i)=\mathrm{C}_k^P Q(\theta_i)^{k-P}(1-Q(\theta_i))^p,i=1,\cdots,n,$$

$$P(\mathrm{A})=\sum_{i=1}^{n}\mathrm{C}_k^P Q(\theta_i)^{k-P}(1-Q(\theta_i))^p\pi(\theta_i),$$

$$\pi(\theta_i|\mathrm{A})=\frac{P(\mathrm{A}|\theta_i)\pi(\theta_i)}{P(\mathrm{A})}=\frac{Q(\theta_i)^{k-P}(1-Q(\theta_i))^p\pi(\theta_i)}{\sum_{i=1}^{n}Q(\theta_i)^{k-P}(1-Q(\theta_i))^p\pi(\theta_i)}。\tag{9}$$

实验B是在实验A之后,对于决策 θ_i 有了相应的了解,从而不再根据最初始的先验分布,而依据实验A所算得的后验分布作为本次实验的先验分布:$P(\mathrm{B}|\theta_i)=\mathrm{C}_m^q Q(\theta_i)^{m-q}(1-Q(\theta_i))^q,i=1,\cdots,n$。

由实验 A 所算得的每次决策的后验概率可得:

$$P(\mathrm{B})=\sum_{l=1}^{n}\mathrm{C}_m^q Q(\theta_l)^{m-q}(1-Q(\theta_l))^q\ \frac{Q(\theta_l)^{k-P}(1-Q(\theta_l))^p\pi(\theta_l)}{\sum_{i=1}^{n}Q(\theta_i)^{k-P}(1-Q(\theta_i))^p\pi(\theta_i)}。$$

于是化简得:$P(\mathrm{B})=\dfrac{\sum_{l=1}^{n}\mathrm{C}_m^q Q(\theta_l)^{m+k-P-q}(1-Q(\theta_l))^{p+q}\pi(\theta_l)}{\sum_{i=1}^{n}Q(\theta_i)^{k-P}(1-Q(\theta_i))^p\pi(\theta_i)}$。

从而由(2)式可得:

$$\pi(\theta_i|\mathrm{B})=\frac{Q(\theta_i)^{m+k-p-q}(1-Q(\theta_i))^{p+q}\pi(\theta_i)}{\sum_{l=1}^{n}Q(\theta_l)^{m+k-P-q}(1-Q(\theta_l))^{p+q}\pi(\theta_l)},i=1,\cdots,n。\tag{10}$$

实验C为两步实验A,B合成一个实验,那么抽样总数为 $k+m$,不满足要求的产品数为 $p+q$,为计算方便,令 $a=k+m$,$r=p+q$,则先验分布则分别为 $\pi(\theta_1),\pi(\theta_2),\cdots,\pi(\theta_n)$,于是:$P(\mathrm{C}|\theta_i)=\mathrm{C}_a^r Q(\theta_i)^{a-r}(1-Q(\theta_i))^r,i=1,\cdots,n,P(\mathrm{C})=\sum_{l=1}^{n}\mathrm{C}_a^r Q(\theta_l)^{a-r}(1-Q(\theta_l))^r\pi(\theta_l)$。

从而由(2)式可得:

$$\pi(\theta_i|\mathrm{C})=\frac{Q(\theta_i)^{a-r}(1-Q(\theta_i))^r\pi(\theta_i)}{\sum_{l=1}^{n}Q(\theta_l)^{a-r}(1-Q(\theta_l))^r\pi(\theta_l)},i=1,\cdots,n。\tag{11}$$

由于 $a=k+m$,$r=p+q$,所以实验C与实验B的结果是一致的,从而推论1得证。 证毕

推论2 对于两种决策 θ_1,θ_2,设采取这两种决策所取得的决策效果分别是 $Q(\theta_1),Q(\theta_2)$,它们的先验分布分别为 $\pi(\theta_1),\pi(\theta_2)$,若进行了 g 次实验分别记为 $\mathrm{A}_1,\mathrm{A}_2,\cdots,\mathrm{A}_g$,其中每次实验抽取的产品数目分别为 $k_1,k_2,\cdots,k_g$,且每次实验不满足要求产品数为 $p_1,p_2,\cdots,p_g$,将前 h 次实验合成的一次实验记为 B_h,那么由前 g 次实验合成的实验 B_g 所算得的各决策的后验概率与由实验 $\mathrm{A}_1,\mathrm{A}_2,\cdots,\mathrm{A}_g$ 所算得的后验概率一致。

证明 下面采用数学归纳法给出证明。

由定理1可知,当 g 等于2时,结论是成立的,假设 $g-1$ 时,结论成立,即,当实验共有 $g-1$ 次时,最后一次实验 A_{g-1} 对于每种决策的后验概率满足推论2的结论,即 $\pi(\theta_i|\mathrm{A}_{g-1})=\pi(\theta_i|\mathrm{B}_{g-1})$。

于是:

$$\pi(\theta_i|\mathrm{A}_{g-1})=\frac{\pi(\theta_i)Q(\theta_i)^{\sum_{s=1}^{g-1}k_s-p_s}(1-Q(\theta_i))^{\sum_{s=1}^{g-1}p_s}}{\sum_{l=1}^{2}\pi(\theta_l)Q(\theta_l)^{\sum_{s=1}^{g-1}k_s-p_s}(1-Q(\theta_l))^{\sum_{s=1}^{g-1}p_s}},i=1,2。\tag{12}$$

对于第 g 次实验 A_g,此时实验对于各决策的先验概率为第 $g-1$ 次实验 A_{g-1} 对各决策的后验概率,于是有:$P(\mathrm{A}_g|\theta_i)=\mathrm{C}_{k_g}^{p_g}Q(\theta_i)^{k_g-p_g}(1-Q(\theta_i))^{p_g},i=1,2,P(\mathrm{A}_g)=\sum_{i=1}^{2}P(\mathrm{A}_g|\theta_i)\pi(\theta_i|\mathrm{A}_{g-1})$。

由(2)式可得：

$$\pi(\theta_i \mid \mathrm{A}_g)=\frac{P(\mathrm{A}_g \mid \theta_i)\pi(\theta_i \mid \mathrm{A}_{g-1})}{\sum_{l=1}^{2} P(\mathrm{A}_g \mid \theta_l)\pi(\theta_l \mid \mathrm{A}_{g-1})},i=1,2, \tag{13}$$

$$\pi(\theta_i \mid \mathrm{A}_g)=\frac{\mathrm{C}_{k_g}^{p_g} Q(\theta_i)^{k_g-p_g}(1-Q(\theta_i))^{p_g}\pi(\theta_i)Q(\theta_i)^{\sum_{s=1}^{g-1}k_s-p_s}(1-Q(\theta_i))^{\sum_{s=1}^{g-1}p_s}}{\sum_{l=1}^{2}\mathrm{C}_{k_g}^{p_g} Q(\theta_l)^{k_g-p_g}(1-Q(\theta_l))^{p_g}\pi(\theta_l)Q(\theta_l)^{\sum_{s=1}^{g-1}k_s-p_s}(1-Q(\theta_l))^{\sum_{s=1}^{g-1}p_s}},i=1,2。 \tag{14}$$

而对于前 g 次实验的合成实验 B_g 由(7)和(8)式易得：

$$P(\theta_i \mid \mathrm{B}_g)=\frac{\pi(\theta_i)Q(\theta_i)^{\sum_{s=1}^{g}k_s-p_s}(1-Q(\theta_i))^{\sum_{s=1}^{g}p_s}}{\sum_{l=1}^{2}\pi(\theta_l)Q(\theta_l)^{\sum_{s=1}^{g}k_s-p_s}(1-Q(\theta_l))^{\sum_{s=1}^{g}p_s}},i=1,2。 \tag{15}$$

将(14)式化简正是(15)式，于是推论 2 得证。　　证毕

推论 3　对于 $n(n>2)$ 种决策 $\theta_1,\theta_2,\cdots,\theta_n$，设采取这 n 种决策所取得的决策效果分别是 $Q(\theta_1),Q(\theta_2),\cdots,Q(\theta_n)$，它们的先验分布分别为 $\pi(\theta_1),\pi(\theta_2),\cdots,\pi(\theta_n)$，若进行了 g 次实验分别记为 $\mathrm{A}_1,\mathrm{A}_2,\cdots,\mathrm{A}_g$，其中每次实验抽取的产品数目分别为 $k_1,k_2,\cdots,k_g$，且每次实验不满足要求产品数为 $p_1,p_2,\cdots,p_g$，将前 h 次实验合成的一次实验记为 B_h，那么由前 g 次实验合成的实验 B_g 所算得的各决策的后验概率与由实验 $\mathrm{A}_1,\mathrm{A}_2,\cdots,\mathrm{A}_g$ 所算得的后验概率一致。

推论 3 的证明可结合推论 1,2 的证明过程得出，在下节中将举例说明并检验该推论。

4 算例分析

例 1　公司为提高产品的质量，采取 3 种决策 $\theta_1,\theta_2,\theta_3$，所取得的决策效果是高质量产品分别占 0.9,0.8,0.75，但根据以往的经验，这 3 种决策的可信度分别是 0.3,0.3,0.4。但产品经理不想用过去的经验来决策此事，想通过小规模的实验后再下结论，为此前后共做了 3 次实验记为 A,B,C，结果分别如下，实验 A：试制 5 个产品，均为高质量；为减少检验的偶然性再做实验 B：试制 10 个产品，有 9 个是高质量；得到此结果经理还是不太确认，接着做了实验 C：试制 15 个产品，有 13 个是高质量，现在需要根据实验结果对 3 种决策的选择做出一个判断。

在本例中，因为 3 次实验是先后进行的，从而每次实验完成之后，对于各决策的可信度会相应发生变化。从而有两种计算方法：一种是根据题目的步骤，依次计算后验概率，再根据最后一步后验概率的最大值做决策；另一种是根据本文提出的计算方法，将 3 次实验合在一起令为实验 D，仅依据最初始的先验概率 $\pi(\theta_1)=0.3$，$\pi(\theta_2)=0.3$，$\pi(\theta_3)=0.4$ 和抽样总数 30，高质量产品数 27 进行计算，再根据最后一步后验概率的最大值做决策。由 R 程序执行结果如下。根据题目要求的步骤算得各后验概率如表 3～5 所示。根据本文提出的方法算得的结果如表 6 所示。

表 3　实验 A 的后验概率值

Tab.3　The posterior probability of experiment A

$\pi(\theta_1 \mid \mathrm{A})$	$\pi(\theta_2 \mid \mathrm{A})$	$\pi(\theta_3 \mid \mathrm{A})$
0.478 293 7	0.265 419	0.256 287 3

表 4　实验 B 的后验概率值

Tab.4　The posterior probability of experiment B

$\pi(\theta_1 \mid \mathrm{B})$	$\pi(\theta_2 \mid \mathrm{B})$	$\pi(\theta_3 \mid \mathrm{B})$
0.608 228	0.233 862 8	0.157 909 3

表 5　实验 C 的后验概率值

Tab.5　The posterior probability of experiment C

$\pi(\theta_1 \mid \mathrm{C})$	$\pi(\theta_2 \mid \mathrm{C})$	$\pi(\theta_3 \mid \mathrm{C})$
0.673 720 1	0.224 104 9	0.102 175

表 6　实验 D 的后验概率值

Tab.6　The posterior probability of experiment D

$\pi(\theta_1 \mid \mathrm{D})$	$\pi(\theta_2 \mid \mathrm{D})$	$\pi(\theta_3 \mid \mathrm{D})$
0.673 720 1	0.224 104 9	0.102 175

表 3～5 依次表示每次实验的后验概率，且实验 B 的后验概率计算依赖实验 A，实验 C 的后验概率计算依赖

于实验 B,即基于分步检验的方法计算最终的后验概率需要3步完成;表6显示的是实验 D 的后验概率值,只需一步就可算的最终的结果,对比表5发现,尽管二者的步骤不同,但最终的值是一样的,即不影响决策者做最终的判断,并且可以看到计算的过程得到大大简化。

5 总结

本文从一个实际的贝叶斯检验例子出发,鉴于该例后验概率计算过程的复杂步骤而导致做决策时的低效率问题,提出一种新的计算方法,理论证明了该方法的可行性,并在此基础上进行了推广,使它适用于更复杂的检验情形,提供了复杂情形的例子来验证推广形式的可行性,从最后的实验结果可以看出,该方法在复杂情形下也适用。本文提出的后验概率的计算方法不仅没有改变最终的结果,即不影响最终决策,而且使得过程大大简化,提高了运算得效率,该方法在更复杂得情形也下适用,在包含复杂步骤的贝叶斯推断过程中利用该方法计算可节省大量的计算时间。

参考文献:

[1] 李思奇,吕王勇,邓柙,等.基于改进 PCA 的朴素贝叶斯分类算法[J].统计与决策,2022,38(1):34-37.
LI S Q,LÜ W Y,DENG X,et al. Naive bayes classification algorithm based on improved PCA[J]. Statistics and Decision,2022,38(1):34-37.

[2] 张婧,袁敏,刘妍岩.基于正态混合模型的贝叶斯分类方法及其应用[J].应用数学学报,2020,43(4):742-755.
ZHANG J,YUAN M,LIU Y Y. Bayesian classification method based on Gaussian mixture model and its application[J]. Journal of Applied Mathematics,2020,43(4):742-755.

[3] 葛继科,陈栋,王文和,等.基于改进朴素贝叶斯分类算法的火灾分类[J].安全与环境学报,2019,19(4):1122-1127.
GE J K,CHEN D,WANG W H,et al. Based on the naive Bayesian classification algorithm of fire classification[J]. Journal of Safety and the Environment,2019,19(4):1122-1127.

[4] 冯小荣,惠康华,柳振东.基于卷积特征和贝叶斯分类器的人脸识别[J].智能系统学报,2018,13(5):769-775.
FENG X R,HUI K H,LIU Z D. Characteristics and Bayesian classifier based on convolution face recognition[J]. Journal of Intelligent Systems,2018,13(5):769-775.

[5] 薛安荣,毛文渊,王孟頔,等.基于贝叶斯方法和变化表的恐怖行为预测算法[J].计算机科学,2016,43(12):130-134.
XUE A R,MAO W Y,WANG M D,et al. Based on Bayesian method and the change table of terrorism prediction algorithm[J]. Computer Science,2016,43(12):130-134.

[6] 付光杰,胡明哲.贝叶斯预测蜂群算法在无线传感器网络优化中的应用[J].重庆大学学报,2018,41(5):15-22.
FU G J,HU M Z. Bayesian forecasting colony algorithm in the application of wireless sensor network optimization[J]. Journal of Chongqing University,2018,41(5):15-22.

[7] 李艳玲.双边定数截尾场合双参数指数分布的贝叶斯预测[J].统计与决策,2014(11):68-70.
LI Y L. Bilateral destiny truncated double parameter exponential distribution of occasions Bayesian prediction[J]. Statistics and Decision,2014(11):68-70.

[8] 朱慧明.时间序列 ARFIMA 模型的贝叶斯预测分析[J].统计与决策,2006(4):4-6.
ZHU H M. The Bayesian prediction analysis of time series ARFIMA model[J]. Statistics and Decision,2006(4):4-6.

[9] 邱晓华,李敏,张丽琼,等.基于卷积特征和贝叶斯决策的双波段场景分类[J].激光与光电子学进展,2021,58(4):366-374.
QIU X H,LI M,ZHANG L Q,et al. Based on the characteristics of convolution and Bayesian decision dual-band scene classification [J]. Laser and Optoelectronics,2021,58(4):366-374.

[10] 安葳鹏,程小博,刘雨.Fleiss' Kappa 系数在贝叶斯决策树算法中的应用[J].计算机工程与应用,2020,56(7):137-140.
AN W P,CHENG X B,LIU Y. Fleiss' Kappa coefficient in the application of Bayesian decision tree algorithm[J]. Computer Engineering and Application,2020,56(7):137-140.

[11] 王波,郑晓东,李晓晔,等.用于癌症亚分型的生物医学大数据谱聚类技术研究[J].科学技术创新,2018(16):9-10.
WANG B,ZHENG X D,LI X H,et al. For cancer and the classification of biomedical data spectral clustering technology research[J]. Science and Technology Innovation,2018(16):9-10.

[12] 韦来生.贝叶斯统计[M].北京:高等教育出版社,2016.
WEI L S. Bayesian statistics[M]. Beijing:Higher Education Press,2016.

[13] 李航.统计学习方法[M].北京:清华大学出版社,2019.

LI H. Statistical learning methods[M]. Beijing: Tsinghua University Press, 2019.

[14] 王汉斌,陈晓怀,程真英,等.融合生产信息的产品检验不确定度评定[J].计量学报,2017,38(2):164-167.
WANG H B, CHEN X H, CHENG Z Y, et al. Production information integration of product inspection uncertainty evaluation [J]. Journal of Measurement, 2017, 38(2): 164-167.

[15] 刘雄,谭智勇,刘铎,等.带可信度特征的多级安全模型行为风险分析[J].清华大学学报(自然科学版),2010,50(1):67-70.
LIU X, TAN Z Y, LIU D, et al. With multilevel security model behavior characteristic of credibility risk analysis[J]. Journal of Tsinghua University (Natural Science Edition), 2010, 50(1): 67-70.

[16] 茆诗松,汤银才.贝叶斯统计[M].北京:中国统计出版社,2012.
MAO S S, TANG Y C. Bayesian statistics[M]. Beijing: China Statistics Press, 2012.

A New Method for Calculating the Posterior Probability of Action Decision Problems

ZHENG Hai, WANG Hongchun

(School of Mathematical Science, Chongqing Normal University, Chongqing 401331, China)

Abstract: [**Purposes**]Based on a practical example of Bayesian inference, an improved calculation method is proposed to overcome the complexity of posterior probability calculation for this type of problem. [**Methods**]The improved calculation method was derived by numerical experiment, and the accuracy of the improved method was proved theoretically. [**Findings**]On the basis of this method, it is extended to apply to more complex cases. Finally, the analysis of an example further shows that this method is valid. [**Conclusions**]The calculation method based on the posterior probability proposed not only does not change the final calculation result, but also greatly simplifies the calculation steps.

Keywords: Bayesian inference; prior distribution; posterior distribution; action decision making

(责任编辑 许 甲)