

一种基于离散马尔可夫过程的诊断风险模型

李利杰, 姚娅川, 王 林, 徐增伟, 徐卫东, 赵 寅

(四川理工学院自动化与电子信息学院, 四川 自贡 643000)

摘 要:文章建立了一个基于贝叶斯公式和马尔可夫链的诊断模型,并根据中国现有的医疗管理体制进行假设检验,对假设结果进行评价。在分类过程中应用贝叶斯决策,将医疗诊断简单情况下的二值分类进行研究,同时依据分类平均风险最小的原则给出了分类的决策函数,并应用贝叶斯理论和马尔可夫过程进行讨论。实验证实了在我国建立强制医疗责任保险制度的正确性与必要性。

关键词:贝叶斯估计; 马尔可夫过程; 决策函数; 二值分类; 平均风险; 医疗责任保险制度

中图分类号: TH16

文献标识码: A

在实践中,造成医疗事故的因素很多,如医院的管理,医院片面追求高收益,将有可能出现重复检查,过量开药的现象,而在此过程中增加了医疗事故发生的概率。还有,医院的硬件设备,如果医院具有先进的诊断设备,则误诊、错诊的概率将大幅度降低。而医生的素质与经验也在诊断过程中起着至关重要的作用^[1]。如果医生具有扎实的理论基础和丰富的临床经验,对于某些病症则可以准确给与确定及诊治,无需其它无关检查过程,也尽可能的避免了在此无关紧要过程带来的误诊错诊。而本文正是以医生的素质和经验为考虑因素,在此基础上建立一个量化的数学模型,并以贝叶斯公式和马尔可夫链为依据,从我国现行的医疗体制出发,对此问题进行了有益的探索^[1-3]。

1 实例分析

为验证马尔可夫模型计算的可靠性,采用文献[4]分析数据为例进行对比计算。首先对1989年、1991年、1999年和2001年四期TM图像进行严格的几何精校正和基于相对辐射校正方法的辐射校正;其次,通过季相模拟消除季节差异对分类精度带来的影响;最后,根据研究区实际情况,将各期图象进行分类,并确定为农田、高密度草、中密度草、低密度草、水体、荒漠和盐渍地七大类,最后采用马尔可夫过程完成系统的计算。通

过计算结果总精度和Kappa系数来评价分类精度,总精度分别为:91.39、98.69、96.3和97.05;Kappa系数分别为:0.91、0.98、0.94和0.94;对本文应用来讲,精度均符合要求。

2 模型讨论

设“患者有某种病症”为事件A,引起事件A的病因为样本空间 Ω 。 B_1, B_2, \dots, B_n 为 Ω 的一个分划,即 $B_i \cap B_j = \Phi, i \neq j, \sum_{i=1}^n B_i = \Omega$,并假定 $P(B_i) > 0$ 。由贝叶斯公式^[5],计算某病因引起事件A的概率为:

$$P(B_i/A) = \frac{P(B_i)P(A/B_i)}{\sum_{i=1}^n P(B_i)P(A/B_i)} \quad (1)$$

2.1 医生方面

在诊疗过程中,医生要根据临床经验对各种病因 B_i 进行权衡。如果误诊,则有可能承担相应的医疗事故风险,相应的误诊可能承担风险而承担的赔偿费用为 λ_{ji} ,患者承担医生针对病因 B_i 开出的诊疗方案的费用为 M_i ,于是在一次诊治过程中,

风险损失函数:

$$\lambda_{ji} = \text{将 } X \text{ 归入 } \omega_i \text{ 的代价 } X \in \omega_j \quad (2)$$

条件风险函数:

$$l^i(x) = \sum_{j=1}^n \lambda_{ji} P\left(\frac{w_j}{x}\right) \quad (3)$$

平均风险:

$$r^i = \int_{\Omega_i} I^i(x) p(x) dx \quad (4)$$

总的期望风险^[3]:

$$r = \sum_{i=1}^n r^i = \sum_{i=1}^n \int_{\Omega_i} \sum_{j=1}^n \lambda_{ji} p(w_j/x) p(x) dx \quad (5)$$

病人平均承担费用:

$$E(A) = P(B_i/A) M_i \quad (6)$$

2.2 病人方面

假设当时病人不患有任何疾病。此时的心情指数为 i 每当一项检查完毕后病人的心情 $i+1$ 的概率为 p 心情 $i-1$ 的概率为 q 检查不明的即认为心情不变为 i 的概率为 r 且 X_{n-1} 与 X_n 之间彼此独立 其中 X_n 为病人 n 项检查后的心情 则 $(X_n, n \geq 0)$ 为齐次马尔可夫过程。其中 0 表示病人心情压抑 n 表示病人心情舒畅。

易知:

$$P_{00} = P(X_{n+1} = 0/X_n = 0) = 1 - P_m = P(X_{n+1} = 0/X_n = 0)$$

$$P(X_{n+1} = j/X_n = i) = \begin{cases} p & j = i + 1 \\ r & j = i \\ q & j = i - 1 \\ 0 & |j - i| \geq 2 \end{cases}$$

其转移概率矩阵为:

$$\begin{pmatrix} 1 & 0 & 0 & 0 & \cdots & 0 & 0 & 0 & 0 \\ q & r & p & 0 & \cdots & 0 & 0 & 0 & 0 \\ 0 & q & r & p & \cdots & 0 & 0 & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & 0 & \cdots & q & r & p \\ 0 & 0 & 0 & 0 & \cdots & 0 & q & r & p \\ 0 & 0 & 0 & 0 & \cdots & 0 & 0 & 0 & 1 \end{pmatrix}$$

为 $N+1$ 阶方阵。如果此时心情为 i 心情变好、变坏和不变的概率为 p, q 和 r 并且保证所有的检查都进行到底。则用 X_n 表示第 n 次检查后的心情。则 $(X_n, n \geq 0)$ 就是上述马尔可夫链^[5]。

由全概率公式得:

$$\begin{aligned} f_{j0} &= P(X_\tau = 0/X_0 = j) = \frac{P(X_\tau = 0, X_0 = j)}{P(X_0 = j)} = \\ &= \sum_{i=1}^N \frac{P(X_\tau = 0, X_1 = i, X_0 = j)}{P(X_0 = j)} = \\ &= \sum_{i=1}^N P(X_\tau = 0/X_0 = j, X_1 = i) P_{ji} = \\ &= \sum_{i=1}^N j_{j0} P_{ji} = p f_{j+1,0} + r f_{j-1,0} + r f_{j,0} \end{aligned} \quad (7)$$

即:

$$(p + q) f_{j0} = p f_{j+1,0} + r f_{j-1,0} \quad (8)$$

解方程(8) 令 $f_{j0} = \lambda^j$ (λ 为带定常数) 则:

$$p \lambda^2 - (p + q) \lambda + q = 0 \quad (9)$$

解代数方程(9)。

(1) 当 $p \neq q$ 时 得 $\lambda_1 = 1, \lambda_2 = \frac{q}{p}$ 得方程(9)

的解为: $f_{j0} = C_1 + C_2 \left(\frac{q}{p}\right)^j$ 由边界条件得 $f_{00} = 1$ 与 $f_{N0} = 0$

得到常数 $C_1 = -\frac{(q/p)^j}{1 - (q/p)^N}, C_2 = \frac{1}{1 - (q/p)^N}$ 故 $f_{j0} = 1 - \frac{1 - (q/p)^j}{1 - (q/p)^N}$

(2) 当 $p = q$ 时 得 $\lambda_1 = \lambda_2 = 1$ 故方程(9) 的通解

为 $f_{j0} = C_1 + C_2 j$ 再由边界条件得 $C_1 = 1, C_2 = -\frac{1}{N}$

故 $f_{j0} = 1 - \frac{j}{N}$ 所以

$$f_{j0} = \begin{cases} 1 - \frac{1 - (q/p)^j}{1 - (q/p)^N} & p = q \\ 1 - \frac{j}{N} & p \neq q \end{cases} \quad (10)$$

在此基础上做如下的检验:

$$X_i = \begin{cases} 1 & \text{若在 } i \text{ 次检查正确} \\ -1 & \text{其他情况} \end{cases}$$

对于一个确定的检查次数 N 检查在 N 次后停止 如果有 $X_1 + X_2 + \dots + X_N > 0$ 则判断病人心情没有受到伤害 如果有 $X_1 + X_2 + \dots + X_N < 0$ 则判断病人心情受到了打击。则此判断过程同样满足公式(10) 即有,

$$\begin{aligned} p(\text{检验断定病人心情受到打击}) &= \\ &= \begin{cases} 1 - \frac{1 - (q/p)^j}{1 - (q/p)^N} & p \neq q \\ 1 - \frac{j}{N} & p = q \end{cases} \end{aligned} \quad (11)$$

3 模型验证与仿真

《医疗事故条例》规定了 11 项医疗事故赔偿的项目 每项中规定了具体的计算标准。其中医疗事故赔偿包含有医疗费用和精神损失抚慰金。医疗费包括住院费、检查费、中西药费和护理费等。精神损害抚慰金是患者及其近亲属因医疗事故受到的精神损害给予的抚慰费用 按照事故发生地居民年平均生活费计算。另外 医疗事故赔偿费用的结算方式在《医疗事故条例》都有规定的标准和计算公式 并由承担医疗事故责任的医疗机构一次性支付。

设“发热症状”为事件 A_1 , “上腹疼痛症状”为事件 A_2 由表 1 和表 2 的数据(表 1 和表 2 的工资水平参照 2007 年 2 月湖北省第十届人民代表大会上的湖北省政治报告工作中的数据) 通过公式(5)、式(6) 计算得: $E(A_1) = 121, r(A_1) = 187165; E(A_2) = 265, r(A_2) =$

22232; $r(A1) >> E(A1)$, $r(A2) >> E(A2)$ 。

表1 发热症状诊断模型的相关数据

疾病 (Ω)	发生概率 (P(Bi A))	误诊率 (λ(ai wj))	赔偿费(元) (Δji)	医疗费 (Mi)
B1	0.3	0	0	0
B2 = 感冒	0.6	0	0	120
B3 = 非典	0.00001	0.4	196060	18000
B4	0.00001	0.4	186060	18000
B5	0.0002	0.15	202110	612
B5 = 其它	0.09978	0.15	0	490

注: B1 = 人体生理功能的正常表现, B4 = 某中类似非典的突发疾病, B5 = 心肺功能缺陷。

表2 上腹疼痛症状诊断模型的相关数据

疾病 (Ω)	发生概率 (P(Bi A))	误诊率 (λ(ai wj))	赔偿(元) (Δji)	医疗费 (Mi)
B1 = 胃炎	0.8	0.005	0	216
B2	0.1	0.005	0	432
B3 = 胃癌	0.00001	0.01	202110	720
B4	0.0001	0.1	202110	756
B5 = 其它	0.09989	0.88	0	490

注: B2 = 胃溃疡、十二指肠溃疡, B4 = 心、膈等器官有病变。

表明医生所承担的风险是远高于病人。设患者检查项数为 N, 此时患者的最高心情指数也为 N, 患者检查前心情指数为 $[N/2]$ (取整), 由式(11): 当医生诊断失误的概率为 0.01, $r = 0.09$, $j = 16$, $N = 32$ 时:

$$p(\text{检验断定病人心情受到打击}) = e^{-30}$$

医生诊断失误的概率为 0.1, $r = 0.09$, $j = 16$, $N = 32$ 时:

$$p(\text{检验断定病人心情受到打击}) = 5.4 \times e^{-16}$$

当医生诊断失误的概率为 0.1, $r = 0.09$, $j = 4$, $N = 8$ 时:

$$p(\text{检验断定病人心情受到打击}) = 0.2323$$

当医生诊断失误的概率为 q, $p = q$ 时:

$$p(\text{检验断定病人心情受到打击}) = 0.5$$

可见, 医生的诊断准确率高低, 无论是对医生自己还是对病人都有巨大的影响(在实际情况中 $j < [N/2]$), 所以

上面的计算的结果要比实际的要小^[6-8]。

4 结束语

通过以上的证明过程表明, 仅以医生的素质和经验为讨论对象, 说明医生准确率的提高, 无论对医生本人还是对病人而言, 都显得尤为重要, 也说明我国建立强制医疗责任保险制度的正确性和必要性。同时, 也说明了现阶段应该提高医生的素质, 加强学习。论文仅从造成误诊的一个方面入手讨论还不够深入, 希望在以后同其它方面的因素一并进行深入的探讨^[9]。

参考文献:

- [1] 工 则. 马尔科夫转移矩阵法[J]. 科技咨询导报 2005 (2): 60.
- [2] 罗 斐. 基于协同论的中国能源消费结构优化研究[D]. 北京: 中国矿业大学 2009.
- [3] 徐 岚, 赵 弈. 利用马尔柯夫过程预测东陵土地利用格局的变化[J]. 应用生态学报 1993 4(3): 272-277.
- [5] 刘小雄, 武 燕. 一种故障诊断的贝叶斯优化算法研究[J]. 计算机应用研究 2009 26(1): 129-131
- [6] 殷少美, 周寅康, 淮励杰, 等. 马尔科夫链在预测土地利用结构中的应用: 以湖南娄底万宝镇为例[J]. 经济地理 2006 26(1): 121-125.
- [7] 何选森. 随机过程[M]. 北京: 人民邮电出版社 2009.
- [8] 王 贤, 牛东晓. 基于马尔科夫链的电力能源竞争力预测[J]. 现代电力 2004 21(1): 94-97.
- [4] 塔西甫拉提·特依拜, 赵 睿. 干旱区绿洲 LUCC 马尔柯夫过程分析[J]. 干旱区地理 2006 29(4): 548-556.
- [9] 陈吕强, 杨平先. 基于图像增强的边缘检测方法[J]. 四川理工学院学报: 自然科学版 2007 20(2): 77-80.

A Model of Diagnosis Risk Based on Discrete-time Markovian Process

LI Li-jie, YAO Ya-chuan, WANG Lin, XU Zeng-wei, XU Wei-dong, ZHAO Yin

(School of Automation and Electronic Information, Sichuan University of Science & Engineering, Zigong 643000, China)

Abstract: This paper established a model for the diagnosis based on Bayesian formula and Markov chain, making hypothesis testing for the model according to China current medical management system, and the results of hypothesis testing were evaluated. Bayesian decision was applied in classification process, studying the two-value classification of medical diagnosis in simple case. While the classification decision function is given according to the principle of minimum average risk classification. And discussion was given by applying the Bayesian theory and Markov processes. The experiments confirmed that the establishment of compulsory medical liability insurance system was correct and necessary in our country.

Key words: Bayesian estimation; Markov process; decision function; two-value classification; average risk; medical liability insurance system