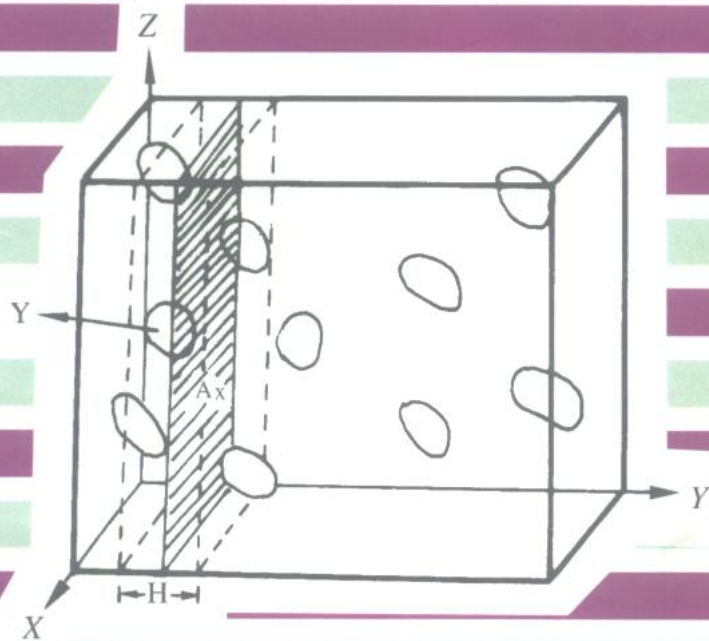


临床计量医学

LINCHUANG JILANG YIXUE

周怀梧 孙伟民 编著



上海医科大学出版社

主编：周怀梧 孙伟民

临床计量医学

上海医科大学出版社

责任编辑 冯 晓
版面设计 丁 玮
责任校对 赵 霞

临床计量医学

主编 周怀悟 孙伟民

上海医科大学出版社出版发行

上海市医学院路 138 号

邮政编码 200032

新华书店上海发行所经销

昆山亭林印刷总厂印刷

开本 787×1092 1/16 印张 11 字数 267 000

1999 年 2 月第 1 版 1999 年 2 月第 1 次印刷

印数 1—5 000

ISBN 7-5627-0445-7/R·421

定价: 17.00 元

如遇印、装质量问题,请直接与印刷厂联系调换

(地址:昆山市中山路 293 号 邮编:215300)

主 编 周怀梧(浙江医科大学,教授)

孙伟民(上海医科大学,教授)

副主编 张选群(湖北医科大学,教授)

张 仲(哈尔滨医科大学,教授)

曾照芳(重庆医科大学,教授)

李 泉(徐州医学院,副教授)

编 委 赵耐青(上海医科大学,副教授)

景荣荣(浙江医科大学,副教授)

陈文金(福建省漳平市卫生防疫站主管医师)

王慕洁(哈尔滨医科大学,副教授)

蓝 荣(青岛大学医学院,副教授)

前 言

回顾本世纪医学科学的发展,不难看出这样一个显著的特点:各种数学、统计学方法及计算机技术在医学科学研究中日益广泛、深入地应用,促使现代医学大踏步地由定性描述向定量分析发展。无论是基础医学、预防医学还是临床医学,概不例外。我们把借用数学、统计学方法及计算机技术研究医学的问题,揭示医学中的数量规律性所形成的医学新学科称为数理医学(mathematical medicine)或计量医学(quantitative medicine)。应用于临床工作的计量医学理论和方法便是临床计量医学。就本人所知,在临床计量医学研究中,已经或正在作系统研究并已获得有效应用的领域,主要有:临床计量诊断,临床决策分析,临床流行病学,临床计量病理学,以及临床药物动力学。本书对这些方面都作了简明、通俗的介绍。

本书以临床医生和临床科研人员、医学生及医科研究生为主要读者对象,也可供医学数学、医学统计及计算机医学应用工作者参考。各章相对独立,但都是着重阐述用数学、统计学研究并解决临床医学问题的思路和方法,强调临床应用的意义和步骤,使得读者能学以致用,增强对临床医学问题作定量分析、处理的能力,提高临床科研水平,促进临床工作进一步精确化、科学化。

本书是中国医药数学会组织编写的。由于缺乏经验,时间又比较仓促,若有不当之处,敬请读者指正。

周怀梧

1998年10月28日

目 录

1 临床计量诊断	(1)
1.1 什么是计量诊断	(1)
1.2 Bayes 公式法	(1)
1.3 最大似然法	(7)
1.4 多元回归分析法	(11)
1.4.1 回归分析的概念	(11)
1.4.2 多元线性回归分析法	(11)
1.4.3 逐步回归分析法	(17)
1.5 判别分析法	(21)
1.5.1 判别分析的概念	(21)
1.5.2 Fisher 意义下两类间的判别分析	(23)
1.5.3 Bayes 意义下多类间的判别分析	(26)
1.5.4 逐步判别分析	(28)
2 临床决策分析	(32)
2.1 决策分析的基本概念	(32)
2.1.1 决策分析的特点	(32)
2.1.2 常用的决策分析方法	(32)
2.1.3 临床决策分析的基本思想	(33)
2.2 矩阵决策法	(34)
2.2.1 矩阵决策的基本要素	(34)
2.2.2 矩阵决策法的应用	(36)
2.3 决策树法	(38)
2.3.1 决策树模型的构造	(38)
2.3.2 决策树法的应用	(39)
2.4 临床检验的决策分析	(40)
2.4.1 检验的诊断价值评判	(40)
2.4.2 检验的阳性预测值和阴性预测值	(42)
2.4.3 临床检验截断点的确定	(43)
2.5 代价——效益分析	(45)
3 临床模糊分类和诊断	(51)
3.1 模糊数学的基本概念	(51)

3.1.1	经典集合及特征函数	(51)
3.1.2	模糊集合及隶属函数	(51)
3.1.3	模糊集合与普通集合的转化	(54)
3.1.4	确定隶属函数的方法	(55)
3.2	临床模糊分类	(57)
3.2.1	模糊关系及模糊矩阵	(57)
3.2.2	模糊聚类分析	(60)
3.2.3	模糊聚类法在预报上的应用	(64)
3.3	临床模糊模式识别与诊断	(65)
3.4	模糊控制器及临床应用	(68)
3.4.1	模糊控制的基本概念	(68)
3.4.2	肿瘤微波热疗的模糊自控系统	(69)
4	临床药物动力学	(73)
4.1	药物动力学基本概念	(73)
4.1.1	药物的体内过程	(73)
4.1.2	最适治疗血药浓度	(75)
4.1.3	药动学模型和参数	(78)
4.2	室分析与矩分析	(81)
4.2.1	室分析	(81)
4.2.2	矩分析	(85)
4.3	给药方案的拟订	(88)
4.3.1	给药方案及简便拟订规则	(88)
4.3.2	血管内给药	(90)
4.3.3	血管外给药	(92)
4.4	给药方案的调整	(94)
4.4.1	肾功能不全病人的给药	(94)
4.4.2	老年病人的给药	(96)
4.4.3	儿童病人的给药	(97)
4.5	给药方案的个体化	(98)
4.5.1	个体化给药的概念	(98)
4.5.2	非线性药物的给药个体化	(100)
4.5.3	血药浓度监测	(102)
4.6	多峰药动学	(103)
4.6.1	双部位吸收药动学	(103)
4.6.2	肠肝循环药动学	(104)
5	临床流行病学研究的设计与分析	(108)
5.1	临床流行病学的基本概念	(108)
5.2	临床科研中几种常见的设计类型	(111)
5.2.1	横断面研究	(111)
5.2.2	病例对照研究	(112)
5.2.3	队列研究	(119)
5.2.4	随机对照试验研究	(122)
5.2.5	Meta 分析法	(123)

5.3	研究结果统计分析	(127)
5.4	评价	(131)
5.5	最小样本量计算	(131)
6	计量病理学方法及临床应用	(134)
6.1	病理形态计量方法	(134)
6.1.1	形态计量学的基本概念	(134)
6.1.2	平面图像的测量	(135)
6.1.3	体积密度的估计	(138)
6.1.4	表面积密度的估计	(139)
6.1.5	粒子数密度的估计	(141)
6.1.6	形状的计量分析	(142)
6.2	病理计量方法的临床应用	(145)
6.2.1	病理数值分类	(145)
6.2.2	病理数值诊断	(148)
6.2.3	影响预后的因素分析	(149)
6.3	炎症过程的数学模拟	(153)
6.3.1	炎症过程初始阶段的数学模型	(153)
6.3.2	无坏死区的慢性炎症过程	(155)
6.3.3	有坏死区的慢性炎症过程	(157)
6.3.4	炎症过程的数学模型与抗菌治疗的最优化	(158)
6.4	肝功能恢复过程的数学模拟	(159)
6.4.1	肝功能状态的定量评估指标——生化指数	(159)
6.4.2	肝功能恢复过程的数学模型	(160)
6.4.3	相对生化指数及参数估计	(161)

1 临床计量诊断

1.1 什么是计量诊断

诊断(diagnosis)就是把问诊、体格检查、实验室检查及其他各项检查所得的资料,经过分析、综合、推理和判断,作出合乎客观实际的结论。形成正确的诊断,不仅要有丰富的医学专业知识,而且要有正确的思维方法。

计量诊断(quantitative diagnosis)就是在传统医学知识的基础上,利用精确的数学方法,建立征候与疾病之间的某种数学模型,通过一定的运算程序作出诊断的方法。

本世纪 50 年代初,英国一位医生 Nash FA 制成一种机械装置,通过对比 82 个症状和体征的各种组合,使医生能从 337 种可能发生的疾病中选出最似然的诊断。可以说,这是计量诊断的初步实验。1959 年, Ledley RS 和 Lusted LB 首先系统地从理论上分析医生作出临床诊断的逻辑推理过程,认为:“医学诊断和处理的推理基础,可用一定的数学方法精确地加以研究和描述。”他们提出了应用数理逻辑和概率论的模型作出诊断的原理和方法,并在罗彻斯特大学医学院作了一组肺癌病例的分析。1961 年, Warner 等报道了 Bayes 定理可用于先天性心脏病的诊断,从此计量诊断开始兴起。进入 70 年代以后,随着电子计算机应用的日益广泛,计量诊断在欧美、日本等国有了较大的发展,所用的数学方法除早期的 Bayes 方法,最大似然法及数理逻辑方法,以解决疾病的鉴别诊断外,又将各种多元分析方法引入到解决疾病诊断及预后等更广泛的医学问题中。

在临床工作中,计量诊断的实现一般都以电子计算机作为主要手段,作出的结论供医生参考,所以常被称为**计算机辅助诊断**(computer aid diagnosis)。众所周知,医生作出诊断结论的可靠性取决于其本人的专业知识和经验。而计量诊断不同,它综合分析的依据不是某一个医生的个人经验,而是以往名家的丰富经验和医院历年积累的大量确诊病例。这就排除了医生在诊断过程中的主观因素,使诊断更全面、更客观,显著地提高了诊断的准确率。

1.2 Bayes 公式法

Bayes 公式法是以概率论中的 Bayes 公式为数学模型的。Warner HR 等在 1961 年首先把它成功地用于鉴别诊断先天性心脏病。他们利用 50 个征候,鉴别诊断 33 种先天性心脏病,借用电子计算机运算,共试验 36 例,结果由计算机得出的可能性最大(即条件概率最大)的诊断,与三位有经验的心脏病专家通过生理学研究 and 外科检查作出的最可能的诊断完全一致。

怎样用 Bayes 公式法作出计量诊断呢?

假设由 m 种疾病组成的“疾病集”

$$D = \{D_1, D_2, \dots, D_m\}$$

中, m 种疾病互斥(即同一个病人只患其中的一种病,不会同时患上其中两种以上的病),而且每种病发生的事前概率

$$P[D_1], P[D_2], \dots, P[D_m]$$

及在每种病发生的条件下,各重要征候的具体表现所组成的征候表现集 S 出现的概率

$$P[S|D_1], P[S|D_2], \dots, P[S|D_m]$$

都是已知的,则在一个确实的征候表现集 S 已经出现的条件下,各种疾病发生的事后概率 $P[D_i|S]$ ($i=1, 2, \dots, m$),可按下列公式计算:

$$P[D_i|S] = \frac{P[D_i]P[S|D_i]}{\sum_{j=1}^m P[D_j]P[S|D_j]} \quad (1.1)$$

此即 Bayes 条件概率模型。在用(1.1)式计算出各种疾病发生的事后概率后,便比较它们的大小,若其中某种疾病 D_k 的事后概率 $P[D_k|S]$ 显著大于其他各种疾病的事后概率 $P[D_i|S]$ ($i \neq k$),则可作出病人患疾病 D_k 的诊断。

下面先以一个简单例子说明(1.1)式的用法。

设有疾病集

$$D = [D_1, D_2, D_3, D_4]$$

其中 D_1 表示胃痛, D_2 表示胃良性病变, D_3 表示其他病变, D_4 表示正常。每种病的事前概率分别为

$$\begin{aligned} P[D_1] &= 0.0869 & P[D_2] &= 0.6836 \\ P[D_3] &= 0.1308 & P[D_4] &= 0.0987 \end{aligned}$$

征候表现集

$$S = \{S_{11}, S_{12}, S_{21}, S_{22}, S_{31}, S_{32}, S_{41}, S_{42}\}$$

其中 S_{11} 表示无上腹痛, S_{12} 表示有上腹痛; S_{21} 表示无黑粪, S_{22} 表示有黑粪; S_{31} 表示无腹块, S_{32} 表示有腹块; S_{41} 表示无发热, S_{42} 表示有发热。根据大量病例,可以得到在各种胃病的条件下征候表现集 S 中各重要征候的具体表现出现的条件概率如表 1-1 所示。

假设某患者有上腹痛,无黑粪,有腹块且发热,则该患者的征候表现集

$$S = [S_{12}, S_{21}, S_{32}, S_{42}]$$

为使用公式(1.1),我们需按表 1-1 计算 $P[S|D_i]$ ($i=1, 2, 3, 4$):

$$\begin{aligned} P[S|D_1] &= P[S_{12}|D_1] \cdot P[S_{21}|D_1] \cdot P[S_{32}|D_1] \cdot P[S_{42}|D_1] \\ &= 0.8182 \times 0.4546 \times 0.4893 \times 0.5719 = 0.1041 \end{aligned}$$

同理可得

$$P[S|D_2] = 0.8621 \times 0.6477 \times 0.1320 \times 0.0160 = 0.0012$$

$$P[S|D_3] = 0.0050$$

$$P[S|D_4] = 0$$

代入式(1.1),便得所求概率

$$P[D_1|S] = \frac{0.1041 \times 0.0869}{0.1041 \times 0.0869 + 0.0012 \times 0.6836 + 0.0050 \times 0.1308 + 0 \times 0.0987} \\ = \frac{0.0090}{0.0105} = 0.8571$$

同理可得

$$P[D_2|S] = \frac{0.0008}{0.0105} = 0.0762$$

$$P[D_3|S] = \frac{0.0007}{0.0105} = 0.0667$$

$$P[D_4|S] = 0$$

表 1-1 胃病鉴别诊断的 Bayes 计算表

征候及表现		胃癌(D_1)	胃良变(D_2)	其他变(D_3)	正常(D_4)	D_i
		0.0869	0.6836	0.1308	0.0987	$P(D_i)$
上腹痛 S_1	无(S_{11})	0.1818	0.1379	0.2670	0.4290	$P[S_{11} D_i]$
	有(S_{12})	0.8182	0.8621	0.7330	0.5710	$P[S_{12} D_i]$
黑粪 S_2	无(S_{21})	0.4546	0.6477	0.3332	0.5720	$P[S_{21} D_i]$
	有(S_{22})	0.5454	0.3523	0.6668	0.4280	$P[S_{22} D_i]$
腹块 S_3	无(S_{31})	0.5107	0.8680	0.8570	1	$P[S_{31} D_i]$
	有(S_{32})	0.4893	0.1320	0.1430	0	$P[S_{32} D_i]$
发热 S_4	无(S_{41})	0.4281	0.9840	0.8570	0.9570	$P[S_{41} D_i]$
	有(S_{42})	0.5719	0.0160	0.1430	0.0430	$P[S_{42} D_i]$
		0.1041	0.0012	0.0050	0	$P[S D_i]$
		0.0090	0.0008	0.0007	0	$P[D_i] \cdot P[S D_i]$
		0.8571	0.0762	0.0667	0	$P[D_i S]$
		D_1				诊断

由上可知,该病患者患胃癌的可能性最大,故诊断为胃癌。

在利用公式(1.1)计算各种疾病出现的事后概率时,分母是相同的,因而实际上只需计算和比较分子的大小即可。把相关计算和最后诊断记在表 1-1 的最后四行里,即可方便地用 Bayes 方法作出疾病的鉴别诊断。

周怀梧等研究急性肠梗阻的鉴别诊断,即在对单纯性肠梗阻(D_1)和绞窄性肠梗阻(D_2)中作出鉴别时,根据医生的临床经验,确定考虑 17 个征候($S_1, S_2, \dots, S_{17}$)。选取浙江医科大学附属一、二院急性肠梗阻住院病例 546 例(除外腹股沟疝、外伤所致肠梗阻,腹部手术未愈而继发肠梗阻,入院诊断疑为肠套叠或肠扭转但未施行手术而缓解出院者),其中单纯性 264 例,绞窄性 282 例。全部绞窄性病例均经手术确诊。而在单纯性病例中,采用手术疗法 89

例,其余 175 例均为采用中西医结合非手术疗法治愈或好转出院。对 546 例住院病例入院当天的 17 种征候加以统计,结果见表 1-2。对 17 个征候的原始资料,利用 χ^2 法筛去 3 个征候,得出鉴别诊断价值较大的 14 个征候具体表现的条件概率,制成计算表 1-3。在已知两型肠梗阻的事前概率为

表 1-2 546 例肠梗阻原始统计资料

征候及其表现			单纯性		绞窄性	
			例数	小计	例数	小计
发病(S_1)	急	(S_{11})	69	263	174	282
	缓	(S_{12})	194		108	
腹痛性质(S_2)	阵发性	(S_{21})	180	240	144	273
	持续性	(S_{22})	60		129	
腹痛强度(S_3)	剧	(S_{31})	83	258	169	279
	非剧	(S_{32})	175		110	
初始呕吐(S_4)	有	(S_{41})	151	224	197	255
	无	(S_{42})	73		58	
腹胀(S_5)	有	(S_{51})	173	238	208	264
	无	(S_{52})	65		56	
既往开腹史(S_6)	有	(S_{61})	181	264	134	282
	无	(S_{62})	83		148	
肠鸣音(S_7)	亢进	(S_{71})	131	213	98	239
	不亢进	(S_{72})	56		44	
	减弱或消失	(S_{73})	26		97	
腹部压痛(S_8)	重或中	(S_{81})	85	229	181	259
	轻或无	(S_{82})	144		78	
肌紧张(S_9)	有	(S_{91})	35	190	98	194
	无	(S_{92})	155		96	
反跳痛(S_{10})	有	($S_{10,1}$)	38	131	98	185
	无	($S_{10,2}$)	93		87	
腹部肿块(S_{11})	有	($S_{11,1}$)	20	129	45	130
	无	($S_{11,2}$)	109		85	
脉搏(次/min)(S_{12})	≤ 85	($S_{12,1}$)	115	177	75	177
	85~	($S_{12,2}$)	44		58	
	105~	($S_{12,3}$)	18		44	
血压(kPa)(S_{13})	≤ 12	($S_{13,1}$)	5	157	25	230
	12~	($S_{13,2}$)	103		147	
	16~	($S_{13,3}$)	49		58	

(续表)

征候及其表现			单纯性		绞窄性	
			例数	小计	例数	小计
体温(°C)(S_{14})	<37	($S_{14,1}$)	95	217	45	120
	37~	($S_{14,2}$)	122		75	
白细胞总数($\times 10^9/L$)(S_{15})	≤ 9	($S_{15,1}$)	118	171	64	165
	9~	($S_{15,2}$)	37		48	
	13~	($S_{15,3}$)	16		53	
中性粒细胞(S_{16})	<0.65	($S_{16,1}$)	26	166	8	161
	0.65~	($S_{16,2}$)	133		121	
	0.90~	($S_{16,3}$)	7		32	
肠内液平面(S_{17})	有	($S_{17,1}$)	80	103	74	89
	无	($S_{17,2}$)	23		15	

$$P[D_1] = \frac{264}{546} = 0.4835; P[D_2] = \frac{282}{546} = 0.5165$$

的基础上,就可以利用表 1-3 进行急性肠梗阻的鉴别诊断。

表 1-3 肠梗阻鉴别诊断的 Bayes 计算表

征候及其表现			单纯性(D_1)	绞窄性(D_2)	D_i
			0.4835	0.5165	$P[D_i]$
发病 (S_1)	急	(S_{11})	0.2624	0.6170	$P[S_{11} D_i]$
	缓	(S_{12})	0.7376	0.3830	$P[S_{12} D_i]$
腹痛性质(S_2)	阵发性	(S_{21})	0.7500	0.5275	$P[S_{21} D_i]$
	持续性	(S_{22})	0.2500	0.4725	$P[S_{22} D_i]$
腹痛强度(S_3)	剧	(S_{31})	0.3217	0.6057	$P[S_{31} D_i]$
	非剧	(S_{32})	0.6783	0.3943	$P[S_{32} D_i]$
初始呕吐(S_4)	有	(S_{41})	0.6741	0.7725	$P[S_{41} D_i]$
	无	(S_{42})	0.3259	0.2275	$P[S_{42} D_i]$
既往开腹史(S_6)	有	(S_{61})	0.6856	0.4752	$P[S_{61} D_i]$
	无	(S_{62})	0.3144	0.5248	$P[S_{62} D_i]$
肠鸣音(S_7)	亢进	(S_{71})	0.6150	0.4100	$P[S_{71} D_i]$
	不亢进	(S_{72})	0.2629	0.1841	$P[S_{72} D_i]$
	减弱或消失	(S_{73})	0.1221	0.4059	$P[S_{73} D_i]$

(续表)

征候及其表现			单纯性(D_1)	绞窄性(D_2)	D_i
			0.4835	0.5165	$P[D_i]$
腹部压痛(S_8)	重或中	(S_{81})	0.3712	0.6988	$P[S_{81} D_i]$
	轻或无	(S_{82})	0.6288	0.3012	$P[S_{82} D_i]$
肌紧张(S_9)	有	(S_{91})	0.1842	0.5052	$P[S_{91} D_i]$
	无	(S_{92})	0.8158	0.4948	$P[S_{92} D_i]$
反跳痛(S_{10})	有	($S_{10,1}$)	0.2901	0.5297	$P[S_{10,1} D_i]$
	无	($S_{10,2}$)	0.7099	0.4703	$P[S_{10,2} D_i]$
腹部肿块(S_{11})	有	($S_{11,1}$)	0.1550	0.3462	$P[S_{11,1} D_i]$
	无	($S_{11,2}$)	0.8450	0.6538	$P[S_{11,2} D_i]$
脉搏(次/min)(S_{12})	≤ 85	($S_{12,1}$)	0.6497	0.4237	$P[S_{12,1} D_i]$
	85~	($S_{12,2}$)	0.2486	0.3277	$P[S_{12,2} D_i]$
	105~	($S_{12,3}$)	0.1017	0.2486	$P[S_{12,3} D_i]$
血压(kPa)(S_{13})	≤ 12	($S_{13,1}$)	0.0318	0.1087	$P[S_{13,1} D_i]$
	12~	($S_{13,2}$)	0.6561	0.6391	$P[S_{13,2} D_i]$
	16~	($S_{13,3}$)	0.3121	0.2522	$P[S_{13,3} D_i]$
白细胞总数($\times 10^9/L$)(S_{15})	≤ 9	($S_{15,1}$)	0.6901	0.3879	$P[S_{15,1} D_i]$
	9~	($S_{15,2}$)	0.2164	0.2909	$P[S_{15,2} D_i]$
	13~	($S_{15,3}$)	0.0935	0.3212	$P[S_{15,3} D_i]$
中性粒细胞(%)(S_{16})	≤ 0.65	($S_{16,1}$)	0.1566	0.0497	$P[S_{16,1} D_i]$
	0.65~	($S_{16,2}$)	0.8012	0.7516	$P[S_{16,2} D_i]$
	0.90~	($S_{16,3}$)	0.0422	0.1987	$P[S_{16,3} D_i]$
			2.2988×10^{-4}	4.7505×10^{-6}	$P[S D_i]$
			111.15×10^{-6}	2.4534×10^{-6}	$P[D_i] \cdot P[S D_i]$
			D_1	诊断	

例如,病人徐 $\times\times$,住院当天零时左右开始腹痛,疼痛限于脐周,为阵发性,尚可忍受,下午九时许呕吐一次。既往无开腹史。血压 14.7/10.7kPa(110/80mmHg),脉搏 70 次/min,腹肌尚软,脐左下有轻度压痛,反跳痛不明显,未扪及肿块,肠鸣音不亢进。白细胞 $10 \times 10^9/L$,中性 0.80。

该病人出现的征候表现集为

$$S = \{S_{12}, S_{21}, S_{32}, S_{42}, S_{62}, S_{72}, S_{82}, S_{92}, S_{10,2}, S_{11,2}, S_{12,1}, S_{13,2}, S_{15,2}, S_{16,2}\}$$

将表 1-3 中相应的条件概率相乘,可得

$$P[S|D_1] = 2.2988 \times 10^{-4}$$

$$P[S|D_2] = 4.7505 \times 10^{-6}$$

分别乘以事前概率 $P[D_1]=0.4835$ 和 $P[D_2]=0.5165$, 便得

$$P[D_1] \cdot P[S|D_1] = 111.15 \times 10^{-6}$$

$$P[D_2] \cdot P[S|D_2] = 2.4534 \times 10^{-6}$$

前者显著大于后者, 故诊断为单纯性肠梗阻 (D_1) (表 1-3), 与该病人最后诊断相符。

在使用 Bayes 公式法时应注意以下几点:

(1) 在 Bayes 模型中, 所有 m 种疾病的事前概率之和等于 1, 即

$$P[D_1] + P[D_2] + \cdots + P[D_m] = 1 \quad (1.2)$$

$P[D_i] (i=1, 2, \cdots, m)$ 可以从既往所积累的临床档案资料或文献中疾病出现的相对比率作出近似估计, 通常与地理、流行病学以及某些社会因素等有关。因此将文献报告的结果应用到本地区、本单位时, 需对各事前概率作必要的检验和修正。

(2) 公式 (1.1) 中出现的条件概率 $P[S|D_i]$ 表示在确诊为疾病 D_i 的条件下, 出现征候表现集 S 的理论频率, 反映了征候表现与疾病之间的内在联系, 通常为常数。实际应用时, 一般均在各个征候表现在统计学上彼此独立的假设下进行计算。也许从临床角度观察, 各个征候表现多少是相互关联的, 不会完全符合“独立性”要求。然而, 根据文献报道和我们的体会, 除了对那些密切相互依赖、相互制约的征候, 应当只取其一, 摒弃其他外, 对于依赖关系不明显的征候, 毋须考虑这一点。

(3) 在比较 m 种疾病的事后概率 $P[D_i|S] (i=1, 2, \cdots, m)$ 时, 其中某个 $P[D_k|S]$ 需大到何种程度才认为是显著大于其他各个事后概率? 即诊断的阈值应多大才可靠? 假设最大的事后概率 $P[D_k|S]=0.5$, 而其他各个事后概率 $P[D_j|S] \leq 0.3 (j \neq k)$, 此时如给病人作出诊断 D_k , 可靠性显然不大。通常预定作出诊断的阈值不小于 0.9, 即认为只有当 $P[D_k|S] \geq 0.9$ 时, 作出诊断 D_k 才是可靠的。否则, 只能供医生作出判断时参考。

1.3 最大似然法

把 Bayes 模型实施简化处理后, 便导出最大似然 (maximum likelihood) 法, 它是比 Bayes 模型法更为简便、实用的方法。

假设要在疾病集

$$D = \{D_1, D_2, \cdots, D_m\}$$

中作出鉴别, 这 m 种疾病假定是互斥的; 重要征候集为

$$S = [S_1, S_2, \cdots, S_n]$$

这 n 个征候假定是互相独立的。由于每个征候都可能若干个具体表现, 因而在一个特定病人身上, 出现的是一个具体的征候表现集, 记为

$$S = \{S_{1u}, S_{2v}, \cdots, S_{nw}\}$$

如果在各种疾病发生的条件下, 已经获得各征候表现的概率, 则根据独立事件的概率乘法公式, 一个具体的征候表现集 S 关于各种疾病的概率应为

$$\begin{cases} P[S|D_1] = P[S_{1u}|D_1] \cdot P[S_{2v}|D_1] \cdots P[S_{nw}|D_1] \\ P[S|D_2] = P[S_{1u}|D_2] \cdot P[S_{2v}|D_2] \cdots P[S_{nw}|D_2] \\ \cdots \cdots \\ P[S|D_m] = P[S_{1u}|D_m] \cdot P[S_{2v}|D_m] \cdots P[S_{nw}|D_m] \end{cases} \quad (1.3)$$

这 m 个条件概率 $P[S|D_1], P[S|D_2], \dots, P[S|D_m]$ 中的最大者假设为 $P[S|D_k]$, 则意味着若病人患疾病 D_k , 则这个征候表现集 S 最有可能出现。我们就将 D_k 定为最大似然诊断。

比较(1.3)式与 Bayes 模型(1.1), 可以看出, 当各种疾病的事前概率 $P[D_i] (i=1, 2, \dots, m)$ 相等, 即

$$P[D_i] = \frac{1}{m} \quad (i=1, 2, \dots, m) \quad (1.4)$$

按 Bayes 模型(1.1)计算各种疾病发生的事后概率 $P[D_i|S]$ 的相对大小, 完全取决于 $P[S|D_i]$ 。从这个意义上, 我们可以把最大似然法看成是 Bayes 方法的简化结果。

实际应用时, 还可将(1.3)式两边取对数, 将求积变换为求和:

$$\left\{ \begin{array}{l} \lg P[S|D_1] = \lg P[S_{1u}|D_1] + \lg P[S_{2v}|D_1] + \dots \\ \quad + \lg P[S_{nw}|D_1] = \sum_{j=1}^n \lg P[S_{jl}|D_1] \\ \lg P[S|D_2] = \lg P[S_{1u}|D_2] + \lg P[S_{2v}|D_2] + \dots \\ \quad + \lg P[S_{nw}|D_2] = \sum_{j=1}^n \lg P[S_{jl}|D_2] \\ \dots\dots\dots \\ \lg P[S|D_m] = \lg P[S_{1u}|D_m] + \lg P[S_{2v}|D_m] + \dots \\ \quad + \lg P[S_{nw}|D_m] = \sum_{j=1}^n \lg P[S_{jl}|D_m] \end{array} \right. \quad (1.5)$$

($l=u, v, \dots, w$)

利用对数函数的单调递增性, 比较各个 $\lg P[S|D_i]$ 的大小, 也可以作出相同的最大似然诊断。

为使计算过程进一步简化, 把(1.5)式适当变形, 按下式计算各个 L_i 值:

$$L_i = \sum_{j=1}^n [\lg P[S_{jl}|D_i] + 1] \times 10 \quad (i=1, 2, \dots, m) \quad (1.6)$$

比较各个 L_i 的大小, 同样可以作出相同的最大似然诊断。(1.6)式中的各项, 称为征候表现 S_{jl} 对疾病 D_i 的诊断指数(indices of diagnosis), 记为

$$I_{ijl} = [\lg P[S_{jl}|D_i] + 1] \times 10 \quad (1.7)$$

通常仅取整数值, 便于计算。

在(1.7)式中, 若令 $P[S_{jl}|D_i]$ 只取两位小数, 最小值设为 0.01, 则有

$$\begin{aligned} -2 &\leq \lg P[S_{jl}|D_i] \leq 0 \\ -1 &\leq \lg P[S_{jl}|D_i] + 1 \leq 1 \\ -10 &\leq \{\lg P[S_{jl}|D_i] + 1\} \times 10 \leq 10 \end{aligned}$$

根据四舍五入原则, 保留(1.7)的整数部分, 则可把频率——概率的近似估计值换算成诊断指数如表 1-4 所示。

下面以中风部位的鉴别诊断为例说明如何用最大似然法进行鉴别诊断。在对因中风 24h 后仍处于昏迷状态的病人检查时, 需要尽快地确定病因, 究竟是脑出血还是脑栓塞? 是小脑出血还是蛛网膜下腔出血? ……以便对症治疗。日本福井 洸在因中风造成死亡的病例

中选择发作后 24 小时仍处于昏迷状态的 47 例病人为对象(年龄:62~87 岁)。因为中风即时的病例因素复杂,所以不用;中风后又引起第二次出血者也除外;病变扩散到脑内许多部位的对象,由于症状与病因难以对应处理,也不作考虑。病理诊断由尸检确定。

表 1-4 征候表现的频率及其相应的诊断指数

频率(%)	指数	频率(%)	指数	频率(%)	指数
0~	-10	4.5~	-3	23.0~	4
1.1~	-9	6.0~	-2	29.0~	5
1.4~	-8	7.0~	-1	36.0~	6
1.8~	-7	9.0~	0	45.0~	7
2.2~	-6	11.0~	1	57.0~	8
2.8~	-5	14.0~	2	71.0~	9
3.5~	-4	18.0~	3	89.0~	10

在上述情况下,假设要在疾病集

$$D = \{D_1, D_2, D_3\}$$

中作出鉴别,其中

D_1 :大脑前、中动脉支配区域出血与下丘脑出血

D_2 :大脑出血与蛛网膜下腔出血

D_3 :大脑中动脉支配区域栓塞

(对于其他区域的病变,由于病例少,不予考虑)。

重要征候集

$$S = \{S_1, S_2, S_3, S_4, S_5, S_6\}$$

其中

S_1 :呕吐

S_2 :陈施氏呼吸

S_3 :发作时血压上升到 26.7kPa(200mmHg)以上

S_4 :单侧麻痹

S_5 :对光反射减弱或消失

S_6 :心房颤动

对 47 例病人,按 D_1 (24 例), D_2 (6 例), D_3 (17 例)三类分组,根据每一种征候出现和未出现的例数,算出每一征候表现的频率(%),再对照表 1-4,求得相应的诊断指数列于表 1-5,即可据此进行最大似然诊断。注意:发现病例数为 0 时,征候频率(%)记为 0;发现病例数等于总例数时,频率(%)记为 100。

例如,一位因中风 24h 后仍处于昏迷的病病人的征候表现集为

$$S = \{S_{11}, S_{22}, S_{31}, S_{41}, S_{51}, S_{62}\}$$