相依配对离散资料的边际分布一致性检验

罗明奎* 蔡昌启 (第三军医大学基础医学部数学教研室) 重庆,400038
史树贵 (第三军医大学附属西南医院神经内科) 重庆,400038

提 要 目的: 针对相依配对 r× r列联表资料建立一种合理的统计处理方法。方法: 将资料的实际处理目的进行正确的统计转化,利用极大似然估计法和拟合优度检验法进行参数估计和假设检验。结果: 给出了参数的具体估计量和假设的具体检验法则。结论: 资料及其它例算结果表明此统计处理方法具有合理性。

关键词 相依配对 r× r列联表; 极大似然估计法; 拟合优度检验法
中图法分类号 R195.1; R311

医学研究或医学教学管理中经常碰到要考察两种方式(或状态)下某离散指标间是否有显著差别。如考察治疗前和治疗后两种状态下病人的疾病情况(离散指标)是否有差异,或考察两种方式对人员的考核情况是否有差别等。这类情况往往对离散指标的测定都不具破坏性,实验设计应合理地采取配对试验收集资料,即先收集 n个受试对象,分别按两种方式配对地考察同一离散指标,记下各格发生的数目。由于试验是针对同一对象分两种方式去考察同一指标,因此所得计数资料称作相依配对计数资料。

对于如上相依配对计数资料,针对边际分布一致的检验问题,文献[1]的 χ^2 检验、文献[2]的秩和检验使用不当,文献[3]中 Stuart和 Maxwell提出的方法欠合理。本文参照文献[4]的理论,对文献[5]中的方法进行改进,以下提出一种较为合理的处理方法。

1 模型

假定对 n个对象进行考察,将每个对象按 A B两种方式配对观察同一离散指标,离散指标分 r个等级,可构成如下相依配对 r× r列联表。

$$n_{i+}=\sum_{j=1}^{r} n_{ij},\ n_{+j}=\sum_{i=1}^{r} n_{ij}$$

记方式 A B下的同一离散指标分别为 ζ,η,假定(ζ,η)具有如下联合分布,见表 1,2

表 1 相依配对 r× r列联表资料

Tab 1 The dependent paired r× r data

		Manner B						
		1	2	…	j	…	r	n_{i+}
Manner A	1	n_{11}	n_{12}	…	n_{1j}	…	n_{1r}	n_{1+}
	2	n_{21}	n_{22}	…	n_{2j}	…	n_{2r}	n_{2+}
	⋮	⋮	⋮		⋮		⋮	⋮
	i	n_{i1}	n_{i2}	…	n_{ij}	…	n_{ir}	n_{i+}
	⋮	⋮	⋮		⋮		⋮	⋮
	r	n_{r1}	n_{r2}	…	n_{rj}	…	n_{rr}	n_{r+}
	n_{+j}	n_{+1}	n_{+2}	…	n_{+j}	…	n_{+r}	n

表 2 (ζ,η)的联合分布表

Tab 2 Joint distribution of (ζ,η)

ζ	η						
	1	2	…	j	…	r	p_{i+}
1	p_{11}	p_{12}	…	p_{1j}	…	P_{1r}	p_{1+}
2	p_{21}	p_{22}	…	p_{2j}	…	p_{2r}	p_{2+}
⋮	⋮	⋮		⋮		⋮	⋮
i	p_{i1}	p_{i2}	…	p_{ij}	…	p_{ir}	p_{i+}
⋮	⋮	⋮		⋮		⋮	⋮
r	p_{r1}	p_{r2}	…	p_{rj}	…	p_{rr}	p_{r+}
p_{+j}	p_{+1}	p_{+2}	…	p_{+j}	…	p_{+r}	1

根据实际要求,要考察两种方式 A B下指标间是否有差异,就是要考察 ζ η间的边际分布是否对应相等,即体现为如下检验问题:

$$\begin{cases} H_0: P(Y=i)=P(Z=i), \text{对所有 } i \text{ 成立}, i=1,2,\cdots,r \\ H_1: P(Y=k)\neq P(Z=k), \text{对某固定 } k \text{ 成立} \end{cases} \quad (1)$$

2 统计处理方法

2.1 参数的估计问题

记

$$p_i=(p_{i1},p_{i2},\cdots,p_{ir}),\ i=1,2,\cdots,r \quad (2)$$

$$p_{i+}=P(Y=i),\ i=1,2,\cdots,r$$
$$p_{+j}=P(Z=j),\ j=1,2,\cdots,r \quad (3)$$

$$p=(p_1,p_2,\cdots,p_r) \quad (4)$$

假定 $(n_{11},n_{12},\cdots,n_{1r},\cdots,n_{r1},n_{r2},\cdots,n_{rr})$服从参数为 n和 $(p_1,\cdots,p_r)$的多项分布,则 p_i的极大似然估计为:

$$\hat{p_i}=\frac{1}{n}(n_{11},n_{12},\cdots,n_{1r}),\ i=1,2,\cdots,r \quad (5)$$

相应地 p_{i+} 和 p_{+j}的估计为:

$$\hat{p_{i+}}=\frac{n_{i+}}{n},\quad i=1,2,\cdots,r \quad (6)$$

$$\hat{p_{+j}}=\frac{n_{+j}}{n},\ j=1,2,\cdots,r \quad (7)$$

* 罗明奎,男,33岁,讲师,硕士

?1994-2016 China Academic Journal Electronic Publishing House. All rights reserved. http://www.cnki.net

则 $$\hat{p}=(\hat{p}_1,\hat{p}_2,\cdots,\hat{p}_r) \tag{8}$$

可以得到 $\hat{p}'$的协方差相容估计为:

$$V(\hat{p}')=\frac{1}{n}(diag(\hat{p}')-\hat{p}'\hat{p}) \tag{9}$$

其中 $diag(\hat{p}')$为 $r^2\times r^2$阶对角阵,主对角线上元素依次为 $\hat{p}_{11},\hat{p}_{12},\cdots,\hat{p}_{rr}$,$\hat{p}'$为 $\hat{p}$的转置。

我们可以通过比较两估计值 $(\hat{p}_{1+}\hat{p}_{2+},\cdots,\hat{p}_{r+})$与 $(\hat{p}_{+1},\hat{p}_{+2},\cdots,\hat{p}_{+r})$的大小,初步判断检验假设 (1)是否成立。

2.2 假设检验问题

令 B_i为 $r\times r$矩阵,其中只有第 i行元素全为 1,其它元素全为 0;E_1为 $r\times r$阶单位阵。

构造矩阵 A为 $2r\times r^2$阶矩阵如下:

$$A=\begin{bmatrix} B_1 & B_2 & \cdots & B_r \\ E_1 & E_1 & \cdots & E_1 \end{bmatrix} \tag{10}$$

令 $U=(p_{1+},p_{2+},\cdots,p_{r+},p_{+1},p_{+2},\cdots,p_{+r})$

则有

$$U=Ap' \tag{11}$$

再构造矩阵 C为 $(r-1)\times 2r$阶矩阵如下:

$$C=(E_2 \quad 0 \quad -E_2 \quad 0) \tag{12}$$

其中 E_2为 $(r-1)\times(r-1)$阶单矩阵,0为 $(r-1)$维零列向量(下同)

事实上,假设检设 (1)式等价为下式:

$$H_0:CU=0 \rightarrow H_1:CU\neq 0 \tag{13}$$

则针对检验假设 (13),采用拟合优度检验法可构造如下检验统计量:

$$i^2=(C\hat{U})'(CAV(\hat{p})'A'C')^{-1}(C\hat{U}) \overset{H_0}{\sim} i^2(r-1) \tag{14}$$

其中 $\hat{\beta}=(\hat{p}_{1+},\hat{p}_{2+},\cdots,\hat{p}_{r+},\hat{p}_{+1},\hat{p}_{+2},\cdots,\hat{p}_{+r})'$

将 (14)式整理可得:

$$n(C\hat{\beta})'=(n_{1+}-n_{+1},\cdots,n_{(r-1)+}-n_{+(r-1)})$$

$n^3(CAV(\hat{p})'A'C'$为 $(r-1)\times(r-1)$矩阵,其主对角线上元素为 $n(n_{i+}+n_{+i})-(n_{i+}-n_{+i})^2-2nn_{ii}$,$i=1,2,\cdots,(r-1)$,其第 i行第 j列元素为 $-(n_{i+}-n_{+j})(n_{j+}-n_{+j})-n(n_{ij}+n_{ji})$,$i\neq j$,$i,j=1,2,\cdots,(r-1)$。

显然如果给出了表 1的具体数据,我们可以根据 (14)式算得 χ^2统计量的观测值 χ_0^2,在给定显著水平 α下,若 $\chi_0^2\geq\chi_\alpha^2(r-1)$,则拒绝零假设 H_0,认为离散指标在两种方式下有显著差别;若 $\chi_0^2<\chi_\alpha^2(r-1)$,则接受零假设 H_0,认为离散指标在两种方式下无显著差别。

3 实例分析

以下我们以 4×4的相依配对资料对上述方法进行分析。

[例]西南医院神经内科提供数据,为考察肉毒毒素 A对治疗面肌痉挛的疗效。总共观测了 179例病人,将每个病人按治疗前和治疗后配对地观测病情状况 (分四级),等级越高表明病情越严重。数据分布情况见表 3

表 3 肉毒毒素 A对面肌痉挛的治疗对比

Tab 3 The treatment on hemifacico spasm with botuoinum toxin type A

		Post-treatment				
		0~1	2	3	4	Total
Pre-treatment	0~1	0	0	0	0	0
	2	21	0	0	0	21
	3	40	29	5	0	74
	4	37	37	8	2	84
	Total	98	66	13	2	179

问治疗对病情好转是否有效?

首先计算估计值:

$(\hat{p}_{1+},\hat{p}_{2+},\cdots,\hat{p}_{r+})=(0,21/179,74/179,84/179)$

$(\hat{p}_{+1},\hat{p}_{+2},\cdots,\hat{p}_{+r})=(98/179,66/179,13/179,2/179)$

从估计值可以看出治疗前高级别的病人偏多,而治疗后低级别的病人偏多,可以初步断定治疗对病情有疗效。

进一步利用 (14)式统计量可算出

$$(C\hat{\beta})'=\frac{1}{179}(-98,-45,61)$$

$$CAV(\hat{p})A'C'=\frac{1}{179^3}\begin{bmatrix} 7938 & -8169 & -1182 \\ -8169 & 13548 & -2446 \\ -1182 & -2446 & 10062 \end{bmatrix}$$

$$(CAV(\hat{p})A'C')^{-1}=\frac{179^3}{10^5}\begin{bmatrix} 43.8863 & 28.6502 & 12.12 \\ 28.6502 & 26.4236 & 9.78897 \\ 12.12 & 9.78897 & 13.7416 \end{bmatrix}$$

$$\chi_0^2\approx 1038.5>>\chi_{0.01}^2(3)=11.345$$

因此认为治疗对病情有显著影响,从估计值可以看出治疗对病情有显著改善。

4 讨论

本文所提供的统计方法可对任意 $r\times r$配对计数数据进行处理,且 (14)式中的统计量包含了 n_{ij}的信息。对 $r\times r$表资料的处理过程中涉及 $(r-1)\times(r-1)$阶逆矩阵的计算,当 r较大时可利用计算机进行计算。

比较 Staurt& Maxwell方法,其检验统计量的形式如下:

$$i_{sm}^2=d'V^{-1}d \overset{H_0}{\sim} i^2(r-1) \tag{15}$$

其中 $d'=(n_{1+}-n_{+1},\cdots,n_{(r-1)+}-n_{+(r-1)})$

$V=(g_{ij})$为 $(r-1)\times(r-1)$阶矩阵,$g_{ij}=n_{i+}+n_{+i}-2n_{ii}$;$g_{ij}=-(n_{ij}+n_{ji})$,$i\neq j$

显然本文方法 (14)式与方法 (15)式有所不同。事实上,(14)式统计量 i^2可写成如下形式:

$$i^2=d'(V-W)^{-1}d \tag{16}$$

其中 d,V定义同于 (15)式,$W=(k_{ij})$为 $(r-1)\times$

?1994-2016 China Academic Journal Electronic Publishing House. All rights reserved. http://www.cnki.net

$(r-1)$阶矩阵 , $k_{ij}=\frac{1}{n}(n_{i+}-n_{+i})(n_{j+}-n_{+j})$, $i,j=1,2,\cdots,r-1$

方法 (15)式统计量与主对角线上数据无关,容易错误地接受 H_0 而方法 (14)式统计量则与主对角线上数据有关。(14)式比 (15)式结果合理,特别是针对小样本情形,其合理性更能体现。如果我们同样以表 3资料为例利用 (15)式进行处理,其结果是: χ^2_{sm}统计量的观测值为 152.489 虽然仍大于临界值,但其拒绝 H_0的程度比利用 (14)式时要弱些。

事实上,我们很容易构造出主对角线数据不同而其它数据相同的两个表格数据,利用本研究提出的方法 (见 16式)处理较为符合实际,而利用 Staurt及 Maxwell方法 (见 15式)则不能符合实际

参考文献

1 汤晓芙,万新华. A型肉毒毒素治疗局限性肌张力障碍与面肌痉挛 .中华神经科杂志 , 1996, 29(2): 111

2 刘好文,李 娜,田成林,等. A型肉毒毒素治疗 118例面肌痉挛的疗效观察 .中华神经科杂志 , 1998, 31(4): 207

3 Everitt E S.列联表分析 .北京:科学出版社 , 1980. 117~ 118

4 张尧庭 .离散多元分析理论与实践 (译著).北京:中国统计出版社 , 1998. 556~ 625

5 熊林平,郭祖超 .重复测量分类数据的分析 . 中国卫生统计 , 1997, 14(1): 1

(收稿: 1999-01-20;修回: 1999-05-12)

(编辑 张大春)

Marginal distribution test for the disposal of dependent paired data

Luo Mingkui, Shi Shugui, Cai Changqi (Department of Mathematics, Third Military Medical University, Chongqing, 400038)

Abstract Objective: To establish a statistical method to dispose dependent paired $r\times r$ data. Methods: A practical problem should be converted with statistical procedures and the maximum likelihood estimation and the goodness of fit test were employed for estimation of the parameters and test of the assumptions. Results: The estimator and the test rules were constructed for the dependent paired $r\times r$ data. Conclusion: An example was shown to confirm that our statistical method could be employed for the disposal of dependent paired $r\times r$ data.

Key words dependent paired contingency table; maximum likelihood estimation; goodness of fit test

(上接 594页)

3 讨论

Na_2ATP可直接供给组织能量,改善缺血缺氧组织的能量供应, ATP也能抑制 Ca^{2+} 内流和钾外流,临床上常静脉注射 ATP治疗房室结折性心动过速[2]。另外,有报道 20 min缺血再灌注伊始给予 ATP-$MgCl_2$能明显减轻心肌“迟呆” (stunned)程度,保护缺血再灌注心功能[3]。

本实验结果显示: 三伤后立即给予 Na_2ATP能有效地防治三伤对心功能的损害。使放射伤后 8 h心功能保持正常,烧伤后 8 h心功能大部分恢复, Na_2ATP对复合伤心功能虽有一定程度的保护作用,但仍显著低于对照水平,说明复合伤时由于放射和烧伤的复合效应,使烧伤休克期持续时间更长、更重,单纯给予 Na_2ATP不能从根本上防治复合伤对心功能的损害。

以前的研究表明 1, 6-二磷酸果糖 (FDP)对三伤心功能的保护作用可能与其改善能量代谢有关[4]。烧伤早期,尤其是烧伤休克状态下,心肌严重缺血及缺氧造成的能量代谢障碍是心肌的收缩性降低的重要原因,而 FDP可在一定程度上使严重烧伤后心肌的能量代谢紊乱得到纠正, Na_2ATP对三伤心功能的保护作用可能也与其改善能量代谢有关。 Na_2ATP可直接供给组织能量,改善心肌能量代谢,从而改善心功能。

关键词 放射损伤;烧伤; Na_2ATP;心功能

中图法分类号 R818.74

参考文献

1 李 靖,肖家思 .小哺乳动物心脏顺行性离体灌流技术 .第三军医大学学报 , 1990, 12(5): 424

2 钱学贤 主编 .现代冠心病监护治疗学 .北京:人民军医出版社 , 1993. 46~ 48

3 郑 军,赵 明,祝景平,等 .三磷酸腺苷-氯化镁对“迟呆”心肌钙离子含量和舒缩功能恢复的影响 .第四军医大学学报 , 1992, 13(6): 414

4 李 敏,肖家思,万子兵 . 1, 6二磷酸果糖对放射伤、烧伤及放烧复合伤早期心功能的影响 .第三军医大学学报 , 1997, 19(3): 289

(收稿: 1998-08-16;修回: 1999-03-18)

(编辑 刘 鹏)

?1994-2016 China Academic Journal Electronic Publishing House. All rights reserved. http://www.cnki.net