

· 科研方法专题 ·

如何正确运用 χ^2 检验——生存资料的似然比检验和六种非参数检验与 SAS 实现

胡纯严¹, 胡良平^{1,2*}

(1. 军事科学院研究生院, 北京 100850;

2. 世界中医药学会联合会临床科研统计学专业委员会, 北京 100029

*通信作者: 胡良平, E-mail: lphu927@163.com)

【摘要】 本文目的是介绍生存资料的似然比检验与六种非参数检验以及 SAS 实现。基于生存资料服从指数分布的假定, 推导出似然比检验方法; 而六种非参数检验方法的主要区别在于它们具有不同的权函数。在未分层和分层的条件下, 都可以运用前述提及的七种生存资料假设检验方法, 它们的共同点是其检验统计量都服从 χ^2 分布。本文通过两个实例, 并借助 SAS 软件, 实现两组或多组生存资料的各种假设检验, 对输出结果作出解释, 并给出统计结论和专业结论。

【关键词】 生存资料; 似然比检验; 非参数检验; 对数秩检验; χ^2 分布

中图分类号: R195.1

文献标识码: A

doi: 10.11886/scjsws20210915001

How to use χ^2 test correctly——the likelihood ratio test and the six nonparametric tests of the survival data and the implementation of the SAS software

Hu Chunyan¹, Hu Liangping^{1,2*}

(1. Graduate School, Academy of Military Sciences PLA China, Beijing 100850, China;

2. Specialty Committee of Clinical Scientific Research Statistics of World Federation of Chinese Medicine Societies, Beijing 100029, China

*Corresponding author: Hu Liangping, E-mail: lphu927@163.com)

【Abstract】 The purpose of this article was to introduce the likelihood ratio test, six nonparametric tests, and the SAS implementation of the survival data. Based on the assumption that the survival data had the exponential distribution, the likelihood ratio test method was derived, the main difference between six nonparametric test methods was that they had different weight functions. Under the conditions of non-stratification and stratification, the seven survival data hypothesis testing methods mentioned above could be used, and their common point was that their test statistics all followed the χ^2 distribution. Through two examples and by means of the SAS software, the article realized the various hypothesis tests for two or more groups of survival data, outputted and explained SAS calculation results, and made statistical and professional conclusions.

【Keywords】 Survival data; Likelihood ratio test; Nonparametric test; Log-rank test; χ^2 distribution

在临床研究中, 生存资料出现的频率很高^[1-7]。生存时间 t 是生存资料中的结果变量, 它具有两个明显的特点: ①生存时间一般不服从正态分布; ②生存资料中常包含删失数据或称为不完全数据。因此, 对生存资料进行统计分析的方法有别于常规的统计分析方法。本文将结合实例, 介绍单因素和多因素生存资料的差异性检验方法, 包括似然比检验方法和六种非参数检验方法。

1 生存函数假设检验的基础

1.1 生存函数的定义

生存分布函数(SDF)也叫生存函数, 常用符号

$S(t)$ 表示, 它描述感兴趣的总体(群体)的寿命时间^[8]。在 t 时刻 SDF 的估计值是来自总体的试验单元(或受试对象)的寿命时间超过 t 的概率, 即有下式:

$$S(t) = P_r(T > t) \quad (1)$$

式(1)中的 $S(t)$ 为生存函数, T 是随机选择的一个试验单元(或受试对象)的寿命时间。

1.2 基本变量的定义

让 $t_1 < t_2 < \dots < t_D$ 代表不同的事件时间。对每一个 i ($i=1, 2, \dots, D$), 让 Y_i 代表 t_i 时刻之前的生存单元数目(即风险集的大小), 让 d_i 代表 t_i 时刻的失效或死亡的数目。

1.3 生存函数的估计

1.3.1 生存函数的 Breslow 估计

$$\hat{S}(t_i) = \exp\left(-\sum_{j=1}^i \frac{d_j}{Y_j}\right) \quad (2)$$

注意: Breslow 估计是累计危险函数的负 Nelson-Aalen 估计的指数。

1.3.2 生存函数的 Fleming-Harrington 估计

$$\hat{S}(t_i) = \exp\left(-\sum_{k=1}^i \sum_{j=0}^{d_k-1} \frac{1}{Y_{k-j}}\right) \quad (3)$$

如果频数值不是整数,就不能计算 Fleming-Harrington 估计。

1.3.3 生存函数的 Kaplan-Meier 估计

在 t_i 时刻的生存函数的 Kaplan-Meier 估计(乘积-极限)是累计乘积:

$$\hat{S}(t_i) = \prod_{j=1}^i \left(1 - \frac{d_j}{Y_j}\right) \quad (4)$$

注意:式(2)、式(3)、式(4)定义的生存函数估计量都是右连续的,也就是说,在 t_i 时刻发生的事件应该被包含在 $S(t_i)$ 的估计之中。

1.4 生存函数的标准误

由 Kalbfleisch 和 Prentice 于 1980 年提出生存函数标准误的格林伍德(Greenwood)估计公式如下:

$$\hat{\sigma}[\hat{S}(t_i)] = \hat{S}(t_i) \sqrt{\sum_{j=1}^i \frac{d_j}{Y_j(Y_j - d_j)}} \quad (5)$$

1.5 两个或多个生存函数相等的检验假设

让 K 代表组数, $S_k(t)$ 代表第 k 组中的生存函数($k=1, 2, \dots, K$)。于是,被检验的无效假设和备择假设分别如下。

$H_0: S_1(t) = S_2(t) = \dots = S_K(t)$, 对于所有的 $t \leq \tau$ 都成立;

H_1 : 至少有一个 $S_k(t)$ 是不同的, 对于某些 $t \leq \tau$ 成立。

2 未分层生存函数相等的假设检验计算公式

2.1 似然比检验

Lawless 于 1982 年基于各组生存资料服从指数分布的假定,推导出检验前述“检验假设”的似然比统计量^[8],见下式:

$$\chi^2 = 2N \log\left(\frac{T}{N}\right) - 2 \sum_{k=1}^K N_k \log\left(\frac{T_k}{N_k}\right), df = K - 1 \quad (6)$$

在式(6)中, χ^2 是服从自由度为 $df=K-1$ 的 χ^2 分布的随机变量(此处也叫做“检验统计量”); N_k 为第 k 组中事件总数(在生存分析中为“死亡”人数); $N = \sum_{k=1}^K N_k$; T_k 为第 k 组中参与检验的全部生存时间之和; $T = \sum_{k=1}^K T_k$ 。

2.2 六种非参数检验

2.2.1 六种非参数检验的名称

在 SAS/STAT 的“LIFETEST”过程中,有六种用于比较两组或多组生存资料的生存函数是否相等的假设检验方法^[8],分别是对数秩(Log-rank)检验、威尔科克森(Wilcoxon)检验、Tarone-Ware 检验、Peto-Peto 检验、修正的 Peto-Peto 检验和 Harrington-Fleming(p, q) 检验。

2.2.2 基本变量的定义

让 $(T_i, \delta_i, X_i), i = 1, 2, \dots, n$, 代表具有右删失生存资料的一个独立样本, T_i 是一个可能的右删失的时间; δ_i 是一个删失的指示变量($\delta_i=0$ 代表 T_i 是删失时间, $\delta_i=1$ 代表 T_i 不是删失时间,即事件发生的时间); $X_i=1, 2, \dots, K$ 代表 K 个不同的组; 让 $t_1 < t_2 < \dots < t_D$ 代表样本中不同的事件时间。在时刻 T_j , 让 $W(t_j)$ 代表一个正权重函数(简称“权函数”), 让 Y_{jk} 和 d_{jk} 分别代表第 k 组中的风险集的大小和事件的数目, 其中, $Y_{jk} = \sum_{i: T_i \geq t_j} I(X_i = k), d_{jk} = \sum_{i: T_i = t_j} \delta_i I(X_i = k)$ 。再分别让 $Y_j = \sum_{k=1}^K Y_{jk}, d_j = \sum_{k=1}^K d_{jk} (j=1, 2, \dots, D)$ 。

2.2.3 六种非参数检验对应的权函数

六种非参数检验的检验统计量无论是在本质上还是在表达形式上都是一样的,所不同的仅仅是权函数,见表 1。

表 1 生存资料六种非参数检验对应的权函数

检验方法名称	权函数 $W(t_j)$
Log-rank	1.0
Wilcoxon	Y_j
Tarone-Ware	$\sqrt{Y_j}$
Peto-Peto	$\bar{S}(t_j)$
修正的 Peto-Peto	$\bar{S}(t_j) \frac{Y_j}{Y_j + 1}$
Harrington-Fleming(p, q)	$[\hat{S}(t_{j-1})]^p [1 - \hat{S}(t_{j-1})]^q, p \geq 0, q \geq 0$

2.2.4 六种非参数检验的共同检验统计量

六种非参数检验的共同检验统计量见式(7):

$$\chi^2 = v'V^{-1}v, df = \text{矩阵}V\text{的秩} \quad (7)$$

在式(7)中, χ^2 服从自由度为 df =矩阵 V 的秩的 χ^2 分布; v' 是 v 的转置向量; V^{-1} 是矩阵 V 的广义逆矩阵。式(7)中等号右边各变量的具体含义如下:

$$v = (v_1, v_2, \dots, v_K)' \quad (8)$$

$$V = (V_{kh}) \quad (9)$$

在式(8)中, 各分量的计算公式见式(10):

$$v_k = \sum_{j=1}^D \left[W(t_j) \left(d_{jk} - Y_{jk} \frac{d_j}{Y_j} \right) \right] \quad (10)$$

在式(10)中, $W(t_j)$ 为权函数, 其定义参见前面的表1。在式(9)中, V 是由 v_k 的方差和 v_k 与 v_h 之间的协方差组成的矩阵, 其各元素的定义如下:

$$V_{kk} = \sum_{j=1}^D \left[W^2(t_j) \frac{d_j(Y_j - d_j)Y_{jk}(Y_j - Y_{jk})}{Y_j^2(Y_j - 1)} \right], \quad 1 \leq k \leq K \quad (11)$$

$$V_{kh} = -\sum_{j=1}^D \left[W^2(t_j) \frac{d_j(Y_j - d_j)Y_{jk}Y_{jh}}{Y_j^2(Y_j - 1)} \right], \quad 1 \leq k \neq h \leq K \quad (12)$$

2.3 校正的对数秩检验

校正的对数秩检验的检验统计量见式(13):

$$\chi^2 = (v_1, v_2, \dots, v_K) V^{-1} (v_1, v_2, \dots, v_K)' \sim \chi^2_{(K-1)} \quad (13)$$

在式(13)中, χ^2 服从自由度为 $K-1$ 的 χ^2 分布, 它是对加权的 K 个样本的生存函数进行检验的检验统计量。式(13)中等号右边各变量的含义和定义如下:

让 $(T_i, \delta_i, X_i, w_i), i = 1, 2, \dots, n$, 代表具有右删失的生存资料的一个独立样本, T_i 是一个可能的右删失的时间; δ_i 是一个删失的指示变量 ($\delta_i=0$ 代表 T_i 是删失时间, $\delta_i=1$ 代表 T_i 不是删失时间, 即事件发生的时间); $X_i=1, 2, \dots, K$ 代表 K 个不同的组; w_i 是 LIFETEST 过程的“WEIGHT”语句中指定变量在各观测上的取值; 让 $t_1 < t_2 < \dots < t_D$ 代表样本中不同的事件时间, 在时刻 $T_j (j=1, 2, \dots, D)$, 并且, 对于每一个 $1 \leq k \leq K$, 则有下式:

$$d_{jk} = \sum_{i: T_i=t_j} I(X_i=k) d_{jk}^w = \sum_{i: T_i=t_j} w_i I(X_i=k) \quad (14)$$

$$Y_{jk} = \sum_{i: T_i \geq t_j} I(X_i=k) Y_{jk}^w = \sum_{i: T_i \geq t_j} w_i I(X_i=k) \quad (15)$$

在时刻 t_j 合并的样本中, 让 $Y_j = \sum_{k=1}^K Y_{jk}$ 和 $d_j = \sum_{k=1}^K d_{jk}$ 分别代表风险集的大小和事件的数目, 让 $Y_j^w = \sum_{k=1}^K Y_{jk}^w$ 与 $d_j^w = \sum_{k=1}^K d_{jk}^w$ 分别代表风险集中权重的数目和事件的权重数目。于是, 在式(13)中, 向量 v 中的各分量和矩阵 $V=(V_{kh})$ 中的各元素(即方差与协方差)分别见式(16)、式(17)、式(18):

$$v_k = \sum_{j=1}^D \left(d_{jk}^w - Y_{jk}^w \frac{d_j^w}{Y_j^w} \right), k = 1, 2, \dots, K \quad (16)$$

$$V_{kk} = \sum_{j=1}^D \left\{ A \sum_{i=1}^{Y_j} [B + C] \right\}, 1 \leq k \leq K \quad (17)$$

$$V_{kh} = \left\{ A \sum_{i=1}^{Y_j} [D - E - F] \right\}, 1 \leq k \neq h \leq K \quad (18)$$

在式(17)中, A、B和C的内容分别如下:

$$A = \frac{d_j(Y_j - d_j)}{Y_j(Y_j - 1)} \quad (19)$$

$$B = \left(\frac{Y_{jk}^w}{Y_j^w} \right)^2 w_i^2 I(X_i \neq k) \quad (20)$$

$$C = \left(\frac{Y_{jk}^w - Y_{jk}^w}{Y_j^w} \right)^2 w_i^2 I(X_i \neq k) \quad (21)$$

在式(18)中, A的内容与式(19)相同; D、E和F的内容分别如下:

$$D = \frac{Y_{jk}^w Y_{jh}^w}{(Y_j^w)^2} w_i^2 I(X_i \neq k, h) \quad (22)$$

$$E = \frac{(Y_j^w - Y_{jk}^w) Y_{jh}^w}{(Y_j^w)^2} w_i^2 I(X_i = k) \quad (23)$$

$$F = \frac{(Y_j^w - Y_{jh}^w) Y_{jk}^w}{(Y_j^w)^2} w_i^2 I(X_i = h) \quad (24)$$

3 分层生存函数相等的假设检验的计算公式

3.1 何为分层生存函数相等的假设检验问题

假定在生存资料中, 有一个具有 M 水平的分层因素, 在其各水平下均有一个具有 K 水平的重要试验因素。于是, 就需要考虑将 M 层生存函数按那个重要试验因素分别合并起来, 再进行合并后的 K 个生存函数是否相等的检验, 这就是分层生存函数相等的假设检验问题。

3.2 分层生存函数相等的检验统计量

进行分层生存函数相等的假设检验, Klein 和 Moeschberger 于 1997 年提出了如下方法: 设分层因素有 M 个水平、重要试验因素有 K 个水平、第 s ($s=1, 2, \dots, M$) 层中的检验统计量为 v_s , 于是, 对合并后的生存函数进行假设检验的检验统计量见下式:

$$\chi^2 = v' V^{-1} v \quad (25)$$

式(25)中的 χ^2 服从自由度为 $df=$ 各层中协方差矩阵 V 的秩的 χ^2 分布; v' 是 v 的转置向量; V^{-1} 是协方差矩阵 V 的广义逆矩阵。式(25)中等号右边各变量的具体含义如下:

$$v = \sum_{s=1}^M v_s \quad (26)$$

$$V = \sum_{s=1}^M V_s \quad (27)$$

4 实例及 SAS 实现

4.1 问题与数据

【例1】(未分层生存资料)将26例白血病患者分为两组,分别采用化疗、化疗联合中药两种方法,治疗后患者存活数(月)如下,其中带“+”号的为终检值。试估计两组患者的生存率,并比较两种方法治疗白血病患者生存率差异是否有统计学意义^[9]。化疗组(A):2⁺,13,7⁺,11⁺,6,1,11,3,17,7;化疗联合中药组(B):10,2⁺,12⁺,13,18,6⁺,19⁺,26,9⁺,8⁺,6⁺,43⁺,9,4,31,24。

【例2】(分层生存资料)某医院肿瘤科提供的一份关于肺癌患者的失效时间资料^[8],因变量为患者治疗后的生存时间 $t(d)$,当 t 为删失数据时,在前面加一个负号表示;考察的协变量(即危险因素或预后因素)如下:①癌细胞的类型,它有4个水平,即腺癌细胞、鳞癌细胞、小细胞肺癌和大细胞肺癌;②治疗类型,它有2个水平,即标准的方法和试验的方法;③疗前处理,它有2个水平,即采取了疗前处理和未采取疗前处理;④患者年龄(岁);⑤从诊断到治疗的等待时间;⑥患者的行动状态用 Karnofsky 率来度量,其取值用 KPS 表示, $0 \leq KPS \leq 10$ 表明患者完全靠医院护理, $40 \leq KPS \leq 60$ 表明患者的行动部分受到限制, $70 \leq KPS \leq 90$ 表明患者的行动可以自理。前3个变量被当作分类变量,后3个变量被当作连续性变量。资料的形式为:各组患者的治疗方法、癌细胞类型、同一组中的样本含量、生存时间、KPS

值、等待时间、年龄、与疗前处理对应的指示变量 PR 值(注:PR=0 等价于令 PRIOR='YES',即表示采取了疗前处理;PR=10 等价于令 PRIOR='NO',即表示未采取疗前处理)。(说明:因数据量很大,各变量取值从略,可查阅参考文献[8])

试按下面两个要求分析此生存资料:其一,仅按“治疗方法”一个因素分层,比较4种癌细胞类型患者的生存曲线差异是否有统计学意义;其二,同时按“治疗方法”和“有无疗前处理”两个因素分层,比较4种癌细胞类型患者的生存曲线差异是否有统计学意义。

4.2 基于 SAS 软件分析例1资料

【分析与解答】设所需要的 SAS 程序如下:

```
data abc;
input group $ n;
do i=1 to n;
input sj @@;
if sj<0 then shengcun=1;
else shengcun =0;
t=abs(sj);
output;
end;
cards;
A 10
-2 13 -7 -11 6 1 11 3 17 7
B 16
10 -2 -12 13 18 -6 -19 26 -9 -8 -6 -43 9 4 31 24
;
```

```
run;
proc lifetest data=abc method=pl plots=(s,ls,lls);
time t* shengcun (1);
strata group/test=all fleming(1,4) LR;
run;
```

【程序说明】在“strata”语句中,选项“test=all”代表选择六种非参数假设检验方法,其中, fleming(ρ_1, ρ_2) 自动设置为: fleming(1), 即 $\rho_1=1, \rho_2=0$; 当特别指定 ρ_1, ρ_2 为非负数且不为0时,对应的计算结果会随着这两个参数的具体取值而发生改变,例如:给定 fleming(1,4) 与 fleming(2,5) 时,对应的 χ^2 值和 P 值是不同的;如果要求输出“似然比检验”结果,必须加上选项“LR”。

【SAS 输出结果及解释】

层间等效检验

检验	卡方	自由度	Pr > 卡方
对数秩	6.5792	1	0.0103
Wilcoxon	5.3348	1	0.0209
Tarone	5.9252	1	0.0149
Peto	5.6198	1	0.0178
Peto 修正	5.5159	1	0.0188
Fleming(1,4)	7.8139	1	0.0052
-2Log(LR)	3.4497	1	0.0633

以上输出结果中,前6行是六种非参数检验结果,最后一行是似然比检验结果。前6行检验结果都得到“ $P < 0.05$ ”的结果;而最后一行的结果为“ $P > 0.05$ ”。当生存资料服从“指数分布”时,应选取“似然比检验”结果;当生存资料服从“威布尔分布”时,应选取“对数秩检验”结果;否则,可以考虑选取其他非参数检验结果。通过查看图1和图2,可大致判断出本例生存时间资料服从何种分布。

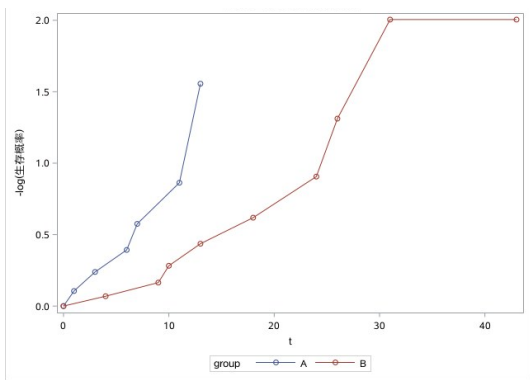


图1 用于判断生存资料是否服从指数分布的图形

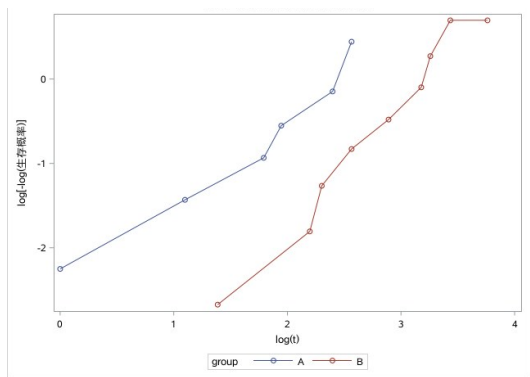


图2 用于判断生存资料是否服从威布尔分布的图形

在图1中,两条折线都不呈直线,说明本例两组生存资料都不服从指数分布;而在图2中,两条折线都近似呈直线,说明本例两组生存资料都近似服从威布尔分布。由此可知,本例的假设检验结果以选择“对数秩检验”结果为宜。为了便于直观判断哪一种治疗方法所对应的患者生存时间更长,需要呈现出他们的生存率曲线,见图3。

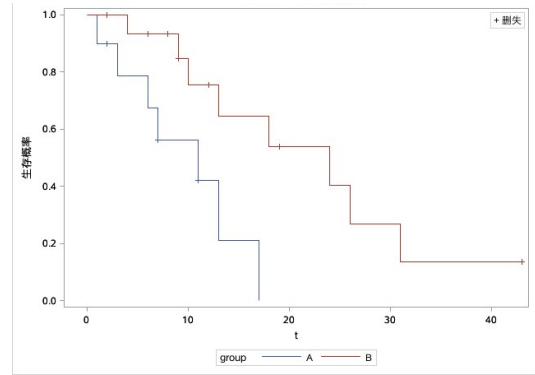


图3 两组生存资料的生存率曲线图

在图3中,左边的生存曲线为“A组(化疗组)”;右边的生存曲线为“B组(化疗联合中药组)”。

【统计结论与专业结论】因 $\chi^2=6.5792, df=1, P=0.0103 < 0.05$,说明两条生存曲线之间的差异有统计学意义;由于B组(化疗联合中药组)的生存曲线位于A组(化疗组)生存曲线的右边(见图3),说明“化疗联合中药”治疗的效果优于“单纯化疗”的效果。

4.3 基于SAS软件分析例2资料

【分析与解答】设所需要的SAS程序如下:

```

DATA abc;
RETAIN THERAPY CELL;
LENGTH PRIOR$ 3;
INPUT THERAPY$ CELL$ N;
DO I=1 TO N;
INPUT T KPS DIAGTIME AGE PR @@;
CENSOR=(T<0);
T=ABS(T);
IF PR=10 THEN PRIOR='YES';
ELSE PRIOR='NO';
OUTPUT;
END;
cards;
72 60 7 69 0 411 70 5 64 10 228 60 3 38 0 126
60 9 63 10
118 70 11 65 10 10 20 5 49 0 82 40 10 69 10
110 80 29 68 0
314 50 18 43 0 -100 70 6 70 0 42 60 4 81 0 8 40
58 63 10
144 30 4 63 0 -25 80 9 52 10 11 70 11 48 10
(还有与上面类似的7组数据从略)
;

```

```

run;
/*按治疗方法分层,比较四种癌细胞类型患者
生存曲线之间的差别*/
proc sort data=abc;
by therapy;
run;
PROC LIFETEST data=abc method=pl plots=(s,
ls,lls);
time T*CENSOR(1);
by therapy;
strata cell/test=all lr;
run;
/*按治疗方法和有无预处理分层,比较四种癌
细胞类型患者生存曲线之间的差别*/
proc sort data=abc;
by therapy prior;
run;
PROC LIFETEST data=abc method=pl plots=(s,
ls,lls);
time T*CENSOR(1);
by therapy prior;
strata cell/test=all lr fleming(1,2);
run;

```

因篇幅所限,第1个过程步的主要输出结果此处从略,仅给出概括性结论。第一层:标准治疗组。标准治疗的4种癌细胞类型的生存曲线之间的差别所对应的 P 值均小于0.05,大细胞肺癌患者的生存曲线最长。第二层:试验治疗组。试验治疗的4种癌细胞类型的生存曲线之间的差别所对应的 P 值均小于0.05,肺鳞癌细胞患者的生存曲线最长。

因篇幅所限,第2个过程步的主要输出结果此处从略,仅给出概括性结论。第一层:标准治疗且未接受预处理,结果显示,三种方法对应的结果为 $P>0.05$,有4种方法对应的 $P<0.05$ 。第二层:标准治疗但接受预处理,结果显示,七种方法对应的结果均为 $P>0.05$ 。第三层:试验治疗且未接受预处理,结果显示,七种方法对应的结果均为 $P<0.05$ 。第四层:试验治疗但接受预处理,结果显示,前六种方法对应的结果均为 $P>0.05$,而最后的似然比检验结果为 $P<0.05$ 。

【统计结论与专业结论】本例资料按“治疗方法”与“是否接受预处理”所形成的四层来分别比较4种癌细胞的生存曲线,因分析结果不尽相同,

故统计结论与专业结论都需要在各层的条件下进行具体描述,因篇幅所限,此处从略。

5 讨论与小结

5.1 讨论

在SAS/STAT的LIFETEST过程中,介绍了比较两组或多组生存资料的差异性检验方法,包括似然比检验和六种非参数检验。一般来说,在构建似然比检验统计量时,需要假定所研究的变量服从某种概率分布;而非参数检验,顾名思义,不假定所研究的资料服从何种概率分布。当各种非参数检验用于同一个资料产生出不同的检验结果时,如何选取最终的结果就成了一个棘手的问题。通常情况下,可以选择 P 值最小的非参数检验法给出的计算结果。例如,在例1的计算结果中,对数秩检验给出的 $P=0.0103$ 、 $\chi^2=6.5792$,而fleming(1,4)的 $P=0.0052$ 、 $\chi^2=7.8139$;若进一步尝试,还可得到fleming(3,6)的 $P=0.0051$ 、 $\chi^2=7.8382$ 。在例2的分析中,因篇幅所限,未分析定量变量(即协变量KPS DIAGTIME AGE)对生存时间的影响。

有一个易于误解之处:在SAS输出差异性检验结果时,呈现的标题为“层间等效检验”,似乎表明:其下方给出的假设检验结果是针对“分层因素”各水平计算出来的。其实,标题中的“层间”是指“试验因素各水平间”。在例2中,试验因素是“4种癌细胞类型”,第一种情况下的分层因素为“治疗方法”,第二种情况下有两个分层因素,即“治疗方法”与“是否接受预处理”。

5.2 小结

本文分别介绍了未分层与分层条件下,两组或多组生存资料比较的似然比检验和六种非参数检验方法,通过两个实例并借助SAS软件,进行单因素和多因素生存资料的差异性分析,对SAS输出结果作出解释,给出统计结论和专业结论。

参考文献

- [1] 姚婷婷,刘媛媛,李长平,等.生存资料回归模型分析:生存资料Cox比例风险回归模型分析[J].四川精神卫生,2020,33(1):27-32.
- [2] 张甜甜,刘红伟,刘媛媛,等.生存资料回归模型分析:生存资料参数回归模型分析SAS实现[J].四川精神卫生,2020,33(1):39-43.

(下转第410页)