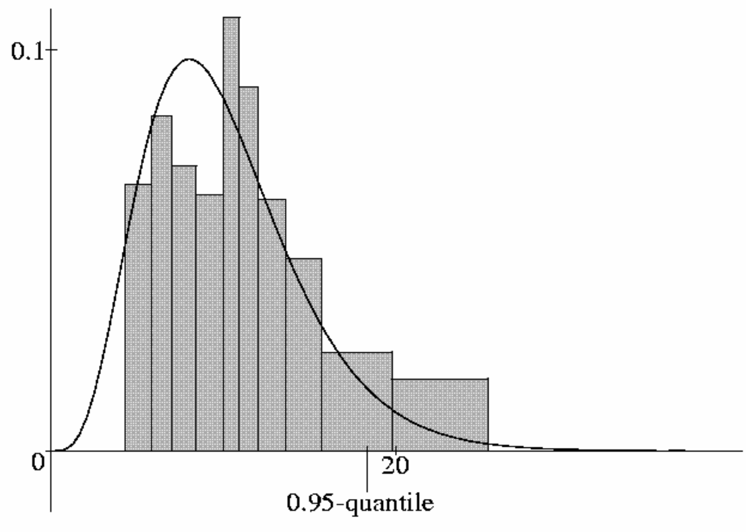
UNIVERSITÄT POTSDAM
Institut für Mathematik

Tests for homogeneity of survival distributions against non-location alternatives and analysis of the gastric cancer data

Vilijandas B. Bagdonavičius
Ruta Levuliene
Mikhail S. Nikulin
Olga Zdorova-Cheminade



Mathematische Statistik und Wahrscheinlichkeitstheorie

Universität Potsdam – Institut für Mathematik

Mathematische Statistik und Wahrscheinlichkeitstheorie

Tests for homogeneity of survival distributions against non-location alternatives and analysis of the gastric cancer data

Vilijandas B. Bagdonavičius
e-mail: algirdasbag@techas.lt

Ruta Levuliene
e-mail: Ruta.Levuliene@maf.vu.lt

Department of Statistics, University of Vilnius, Lithuania

Mikhail S. Nikulin
e-mail: nikou@sm.u-bordeaux2.fr

Olga Zdorova-Cheminade

Mathematical Statistics, University of Bordeaux-2, 33076 Bordeaux, France

Preprint 2004/03

Januar 2004

Impressum

© Institut für Mathematik Potsdam, Januar 2004

Herausgeber: Mathematische Statistik und Wahrscheinlichkeitstheorie
am Institut für Mathematik

Adresse: Universität Potsdam
PF 60 15 53
14415 Potsdam

Telefon:
Fax: +49-331-977 1500
E-mail: +49-331-977 1578
neisse@math.uni-potsdam.de

ISSN 1613-3307

Tests for homogeneity of survival distributions against non-location alternatives [1]

VILIJANDAS B. BAGDONAVIČIUS algirdasbag@techas.lt
RUTA LEVULIENE Ruta.Levuliene@maf.vu.lt
Department of Statistics, University of Vilnius, Lithuania

MIKHAIL S. NIKULIN nikou@sm.u-bordeaux2.fr
OLGA ZDOROVA-CHEMINADE
UFR MI2S, University of Bordeaux-2, 33076, Bordeaux, France

Abstract. The two and k-sample tests of equality of the survival distributions against the alternatives including cross-effects of survival functions, proportional and monotone hazard ratios, are given for the right censored data. The asymptotic power against approaching alternatives is investigated. The tests are applied to the well known chemio and radio therapy data of the Gastrointestinal Tumor Study Group. The P-values for both proposed tests are much smaller then in the case of other known tests. Differently from the test of Stablein and Koutrouvelis the new tests can be applied not only for singly but also to randomly censored data.

Keywords: Censoring; Cross-effects; k-tests; Kolmogorov-Smirnov type tests; Logrank test; Non-proportional hazards; Proportional hazards; Two-sample tests.

1. Introduction

There exists a number of two-sample tests for the hypothesis of the equality of the survival distributions when samples are censored. One group of tests includes various generalisations of the classical Cramer-von-Mises, Kolmogorov-Smirnov statistics to the censored case (Gill (1980), Fleming et al (1980), Koziol (1978), Schumacher (1984), Fleming et al (1987), see the surveys in Andersen et al (1993), p.p.392-395, Klein and Moeschberger (1997), p.p. 209-221).

Another group of tests are the weighted logrank tests. These tests are based on the weighted integrals with respect to the difference of the Nelson-Aalen estimators of the cumulative hazards (Gehan (1965), Peto and Peto (1972), Aalen (1978), Tarone and Ware (1977), Prentice (1978), Kalbfleisch and Prentice (1980), Gill (1980), Fleming and Harrington (1981), Harrington and Fleming (1982), see the surveys in Fleming and Harrington (1991), p.p. 255-277, Andersen et al (1993), p.p. 348-379, Klein and Moeschberger (1997), p.p. 191-194.

Brookmeyer and Crowley (1982) proposed a censored-data version of the median test. A generalization of the classical t-test, based on the Kaplan-Meier estimatots, is given in Klein and Moeschberger (1997). Tests based on the intensity ratio estimates was given by Andersen (1983).

Generalization of the logrank type tests to the case of k-sample situation can be found in Breslow (1970), Peto and Peto (1972), Tarone and Ware (1977), Pren-

[1]This work is done with the support of the Commission Permanente du Conseil Régional d'Aquitaine: Dossier # 20030304001A

tice (1978), Kalfleisch and Prentice (1980, Chapter 6), Andersen et al (1982), Hjort (1984), Fleming and Harrington (1991), Andersen et al (1993), Klein and Moeschberger (1997).

The classical weighted logrank tests have small power under the alternative of crossing survival functions because early differences in favor of one group are negated by late survival advantage of another group.

Cramer-von-Mises and Kolmogorov-Smirnov type statistics have greater but reduced power versus desirable alternatives because they are omnibus tests, not directed against these alternatives.

Stablein and Koutrouvelis (1985) proposed a two-sample test oriented to the cross-effects alternatives which can be applied only to singly censored data.

We propose two and k-sample tests of equality of the survival distributions against the alternatives including cross-effects of survival functions, proportional and monotone hazard ratios. These tests can be used in the case of the right censored data. One group of tests is a modified version of the score test to the semiparametric situation. It appears that in the two-sample situation the modified score functions are two specified weighted logrank statistics. As it was expected, each of these statistics separately does badly against the cross-effect alternative but simultaneously they do very well. In the k-sample case the modified score function is $2(k-1)$-dimensional. In the two-sample case we also propose a test statistic which is a combination of the modified score statistics done in such a manner that negation of the early and late survival differences is avoided under the cross-effect of survival functions alternative. This test also does better than the known tests.

2. Alternative

Let us consider the hypothesis of the equality of the survival distributions :

$$H_0 : S_1(t) = S_0(t)$$

against the alternative written in terms of the hazard rates

$$H_A : \lambda_1(t) = e^{\beta} \left\{1 + e^{\beta+\gamma}\Lambda_0(t)\right\}^{e^{-\gamma}-1} \lambda_0(t);$$

we note

$$\Lambda_i(t) = \int_0^t \lambda_i(u)du \quad (i = 0, 1)$$

the cumulative hazards. Note that $\Lambda_1(0) = \Lambda_0(0) = 0$.

Under the alternative the differences $\lambda_1 - \lambda_0$ and $\Lambda_1 - \Lambda_0$ have the following properties:

1) If $\gamma > 0$ and $\beta > 0$ then there exists $t_0 \in (0, \infty)$ such that $\lambda_1 - \lambda_0 > 0$ in the interval $(0, t_0)$, $\lambda_1(t_0) - \lambda_0(t_0) = 0$, and $\lambda_1 - \lambda_0 < 0$ in the interval (t_0, ∞). So the hazard rate functions intersect once in the interval $(0, \infty)$.

Moreover, there exists $t_1 > t_0$ such that $\Lambda_1 - \Lambda_0 > 0$ in the interval $(0, t_1)$, $\Lambda_1(t_1) - \Lambda_0(t_1) = 0$, and $\Lambda_1 - \Lambda_0 < 0$ in the interval (t_1, ∞). So the cumulative hazards and consequently the survival functions intersect once in the interval $(0, \infty)$.

2) If $\gamma < 0$ and $\beta < 0$ then there exists $t_0 \in (0, \infty)$ such that $\lambda_1 - \lambda_0 < 0$ in the interval $(0, t_0)$, $\lambda_1(t_0) - \lambda_0(t_0) = 0$, and $\lambda_1 - \lambda_0 > 0$ in the interval (t_0, ∞). So the hazard rate functions intersect once in the interval $(0, \infty)$.

Moreover, there exists $t_1 > t_0$ such that $\Lambda_1 - \Lambda_0 < 0$ in the interval $(0, t_1)$, $\Lambda_1(t_1) - \Lambda_0(t_1) = 0$, and $\Lambda_1 - \Lambda_0 > 0$ in the interval (t_1, ∞). So the cumulative hazards and consequently the survival functions intersect once in the interval $(0, \infty)$.

3) If $\gamma > 0$ and $\beta \leq 0$ then $\lambda_1 - \lambda_0 < 0$ in the interval $(0, \infty)$ and the ratio λ_1/λ_0 decreases from $e^\beta < 1$ to 0. So the hazard rate function λ_1 is smaller than the hazard function λ_0 in the interval $(0, \infty)$ and their ratio decreases.

Moreover, the difference $\Lambda_1 - \Lambda_0$ is negative and decreasing in the interval $(0, \infty)$. So the cumulative hazards go away one from another.

4) If $\gamma < 0$ and $\beta \geq 0$ then $\lambda_1 - \lambda_0 > 0$ in the interval $(0, \infty)$ and the ratio λ_1/λ_0 increases from $e^\beta > 1$ to ∞. So the hazard rate function λ_1 is greater than the hazard rate function λ_0 in the interval $(0, \infty)$ and their ratio increases.

Moreover, the difference $\Lambda_1 - \Lambda_0$ is positive and increasing in the interval $(0, \infty)$. So the cumulative hazards go away one from another.

5) If $\gamma = 0$ and $\beta \neq 0$ then $\lambda_1(t) - \lambda_0(t) = e^\beta - 1 = const.$

We seek a test which is powerful against this composite alternative. This alternative includes not only intersection of the survival functions as a partial case but also the possibilities of monotone and constant hazards ratios.

3. Modified score test

Suppose that n_0 objects of the group zero and n_1 objects of the group one are observed. Denote by T_{ij} and C_{ij} the failure and censoring times for the jth patient of the ith group, and set

$$X_{ij} = \min(T_{ij}, C_{ij}), \quad \delta_i = \mathbf{1}_{\{T_{ij} \leq C_{ij}\}},$$

$$N_{ij}(t) = \mathbf{1}_{\{T_{ij} \leq t, \delta_{ij} = 1\}}, \quad Y_{ij}(t) = \mathbf{1}_{\{X_{ij} \geq t\}},$$

where $\mathbf{1}_A$ denotes the indicator of the event A.

Set $N_i(t) = \sum_{j=1}^{n_i} N_{ij}(t)$ and $Y_i(t) = \sum_{j=1}^{n_i} Y_{ij}(t)$, $N(t) = N_0(t) + N_1(t)$, $Y(t) = Y_0(t) + Y_1(t)$.

The score functions (cf. Bagdonavičius and Nikulin (2002)) for the parameters β and γ estimation are

$$U_1(\beta, \gamma) = \int_0^\infty \left\{ 1 + (e^\beta - e^{\beta+\gamma}) \frac{\tilde{\Lambda}_0(t-, \beta, \gamma)}{1 + e^{\beta+\gamma}\tilde{\Lambda}_0(t-, \beta, \gamma)} \right\}$$

$$\left\{ dN_1(t) - Y_1(t) e^\beta (1 + e^{\beta+\gamma}\tilde{\Lambda}_0(t-, \beta, \gamma))^{e^{-\gamma}-1} d\tilde{\Lambda}_0(t, \beta, \gamma) \right\},$$

$$U_2(\beta, \gamma) = \int_0^\infty \left\{ -e^{-\gamma} \ln(1 + e^{\beta+\gamma}\tilde{\Lambda}_0(t-, \beta, \gamma)) + \frac{(e^\beta - e^{\beta+\gamma})\tilde{\Lambda}_0(t-, \beta, \gamma)}{1 + e^{\beta+\gamma}\tilde{\Lambda}_0(t-, \beta, \gamma)} \right\}$$

$$\left\{ dN_1(t) - Y_1(t) e^\beta (1 + e^{\beta+\gamma}\tilde{\Lambda}_0(t-, \beta, \gamma))^{e^{-\gamma}-1} d\tilde{\Lambda}_0(t, \beta, \gamma) \right\},$$

where $\tilde{\Lambda}_0$ is defined recurrently from the equation :

$$\tilde{\Lambda}_0(t,\beta,\gamma) = \int_0^t \frac{dN(u)}{S^{(0)}(u-,\tilde{\Lambda}_0,\beta,\gamma)},$$

$$S^{(0)}(v,\tilde{\Lambda}_0,\theta) = Y_0(v) + Y_1(v)e^{\beta}\left\{1 + e^{\beta+\gamma}\tilde{\Lambda}_0(v,\beta,\gamma)\right\}^{e^{-\gamma}-1}.$$

Under H_0 we have

$$U_1(0,0) = \int_0^\infty \left\{dN_1(t) - Y_1(t)d\tilde{\Lambda}_0(t,0,0)\right\},$$

$$U_2(0,0) = -\int_0^\infty \ln(1+\tilde{\Lambda}_0(t-,0,0))\left\{dN_1(t) - Y_1(t)d\tilde{\Lambda}_0(t,0,0)\right\},$$

where

$$\tilde{\Lambda}_0(t,0,0) = \int_0^t \frac{dN(u)}{Y(u)}.$$

Set

$$\hat{U}_1 := U_1(0,0), \quad \hat{U}_2 := U_2(0,0) \quad \tilde{\Lambda}_0(t) := \tilde{\Lambda}_0(t,0,0).$$

Both statistics $\hat{U}_1$ and $\hat{U}_2$ are logrank-type:

$$\hat{U}_1 = \int_0^\infty \frac{Y_0(t)Y_1(t)}{Y(t)} d\{\hat{\Lambda}_1(t) - \hat{\Lambda}_0(t)\},$$

$$\hat{U}_2 = -\int_0^\infty \frac{Y_0(t)Y_1(t)}{Y(t)} \ln(1+\tilde{\Lambda}_0(t-)) d\{\hat{\Lambda}_1(t) - \hat{\Lambda}_0(t)\},$$

where

$$\hat{\Lambda}_i(t) = \int_0^t \frac{dN_i(u)}{Y_i(u)}$$

is the Nelson-Aalen estimator of the cummulative hazard from the i sample ($i = 0,1$).

At first let us construct the modified score test. Denote by M_i the counting process N_i martingale. Under the null hypothesis

$$\hat{U}_1 = \int_0^\infty \frac{Y_0(t)Y_1(t)}{Y(t)} d\left\{\frac{M_1(t)}{Y_1(t)} - \frac{M_0(t)}{Y_0(t)}\right\},$$

$$\hat{U}_2 = -\int_0^\infty \frac{Y_0(t)Y_1(t)}{Y(t)} \ln(1+\tilde{\Lambda}_0(t-)) d\left\{\frac{M_1(t)}{Y_1(t)} - \frac{M_0(t)}{Y_0(t)}\right\},$$

and the predictable variations and the predictable covariation of the score statistics are

$$<\hat{U}_1>(t) = \int_0^t \frac{Y_0(u)Y_1(u)}{Y^2(u)} d\Lambda_0(u),$$

$$<\hat{U}_2>(t) = \int_0^t \frac{Y_0(u)Y_1(u)}{Y^2(u)} \ln^2(1+\tilde{\Lambda}_0(u-)) d\Lambda_0(u),$$

$$<\hat{U}_1,\hat{U}_2>(t) = -\int_0^\infty \frac{Y_0(u)Y_1(u)}{Y^2(u)} \ln(1+\tilde{\Lambda}_0(u-)) d\Lambda_0(u).$$

It implies that under standard conditions the limit distribution (as $n = n_0 + n_1 \to \infty,\ n_i/n \to l_i \in (0,1)$) of the statistic

$$X^2 = (\hat{U}_1, \hat{U}_2)\hat{\Sigma}^{-1}(\hat{U}_1, \hat{U}_2)^T$$

is the chi-square distribution with two degrees of freedom; here

$$\hat{\Sigma} = \begin{pmatrix} \int_0^\infty \frac{Y_0(t)Y_1(t)}{Y^2(t)}\, dN(t) & -\int_0^\infty \frac{Y_0(t)Y_1(t)}{Y^2(t)} \ln(1+\tilde{\Lambda}_0(t-))dN(t) \\ -\int_0^\infty \frac{Y_0(t)Y_1(t)}{Y^2(t)} \ln(1+\tilde{\Lambda}_0(t-))dN(t) & \int_0^\infty \frac{Y_0(t)Y_1(t)}{Y^2(t)} \ln^2(1+\tilde{\Lambda}_0(t-))dN(t) \end{pmatrix}.$$

The null hypothesis is rejected with the significance level α if $X^2 > \chi^2_{1-\alpha}(2)$, where $\chi^2_{1-\alpha}(2)$ is the $(1-\alpha)$-quantile of the chi-square distribution with 2 degrees of freedom.

4. Second test

We have seen that under the alternative when the parameters β and γ have the same sign, the difference $\lambda_1 - \lambda_0$ has one sign in the interval $(0, t_0)$ and another sign in the interval (t_0, ∞), where

$$t_0 = \Lambda_0^{-1}\left\{ e^{-\beta-\gamma}\left[\exp\{\frac{\beta}{1-e^{-\gamma}}\} - 1 \right] \right\}.$$

Consider the following test statistic, which is a combination of the two score statistics:

$$W = \int_0^\infty K(t)\, d\{\hat{\Lambda}_1(t) - \hat{\Lambda}_0(t)\},$$

where

$$K(t) = \frac{Y_0(t)Y_1(t)}{\sqrt{n}Y(t)}\{\ln(1+\tilde{\Lambda}_0(\hat{t}_0, \hat{\beta}, \hat{\gamma})) - \ln(1+\tilde{\Lambda}_0(t-, \hat{\beta}, \hat{\gamma}))\},$$

$$\hat{t}_0 = \begin{cases} \hat{\Lambda}_0^{-1}\left\{ e^{-\hat{\beta}-\hat{\gamma}}\left[\exp\{\frac{\hat{\beta}}{1-e^{-\hat{\gamma}}}\} - 1 \right] \right\}, & \text{if } \hat{\beta} > 0, \hat{\gamma} > 0 \text{ or } \hat{\beta} < 0, \hat{\gamma} < 0, \\ 0, & \text{otherwise}, \end{cases}$$

$\hat{\beta}$ and $\hat{\gamma}$ are the modified partial maximum likelihood estimators of the parameters β and γ maximising the logarithm of the modified partial likelihood function partial likelihood function

$$\ln L(\beta, \gamma) = \int_0^\infty \left\{ \beta + (e^{-\gamma} - 1)\ln\left(1 + e^{\beta+\gamma}\tilde{\Lambda}_0(v-, \beta, \gamma)\right)\right\} dN_1(v) -$$

$$\int_0^\infty \ln\left(Y_0(v) + Y_1(v)e^{\beta}\left(1 + e^{\beta+\gamma}\tilde{\Lambda}_0(v-, \beta, \gamma)\right)^{e^{-\gamma}-1}\right) dN(v).$$

For fixed $\theta = (\beta, \gamma)$ the function $\tilde{\Lambda}_0(t, \theta)$ can be found recurrently. Indeed, let $T_1^* < ... < T_r^*$ be observed and ordered distinct failure times for the unified sample, $r \le n$. Note by d_i the number of failures at the moment T_i. Then

$$\tilde{\Lambda}_0(0; \theta) = 0, \quad \tilde{\Lambda}_0(T_1^*; \theta) = \frac{d_1}{S^{(0)}(0, \tilde{\Lambda}_0, \theta)} = \frac{d_1}{n_0 + n_1 e^{\beta}},$$

$$\tilde{\Lambda}_0(T^*_{j+1};\theta) = \tilde{\Lambda}_0(T^*_j;\theta) + \frac{d_{j+1}}{S^{(0)}(T^*_j, \tilde{\Lambda}_0, \theta)} \quad (j = 1, ..., r-1).$$

The statistic W is asymptotically equivalent to the random variable

$$W_0 = \int_0^\infty \frac{Y_0(t)Y_1(t)}{\sqrt{n}Y(t)} \{\ln(1 + \Lambda_0(t_0, \beta, \gamma)) - \ln(1 + \Lambda_0(t-, \beta, \gamma))\}\, d\{\hat{\Lambda}_1(t) - \hat{\Lambda}_0(t)\},$$

Under H_0 the random variable W is asymptotically normal with zero mean and the variance σ^2 which can be consistently estimated by the statistic

$$\hat{\sigma}^2 = \int_0^\infty \frac{Y_0(t)Y_1(t)}{nY^2(t)} \{\ln(1 + \tilde{\Lambda}_0(\hat{t}_0, \hat{\beta}, \hat{\gamma})) - \ln(1 + \tilde{\Lambda}_0(t-, \hat{\beta}, \hat{\gamma}))\}^2 dN(t).$$

So the asymptotic distribution of the test statistic

$$T = W/\hat{\sigma}$$

is standard normal and the hypothesis H_0 is rejected with approximate significance level α if $\mid T \mid > z_{1-\alpha/2}$, where $z_{1-\alpha/2}$ is the $(1-\alpha/2)$-quantile of the standard normal distribution.

5. Power of the tests

Consider first the power of the second test under the sequence of the approaching alternatives with cross-effects of the survival functions:

$$H_n: \ \lambda_1(t) = e^{\frac{c_1}{\sqrt{n}}} \left\{1 + e^{\frac{c_1+c_2}{\sqrt{n}}} \Lambda_0(t)\right\}^{e^{-\frac{c_2}{\sqrt{n}}}-1} \lambda_0(t);$$

here $c_1 > 0$ and $c_2 > 0$ or $c_1 < 0$ and $c_2 < 0$.

Write the test statistic W as follows:

$$W = \int_0^\infty \frac{K(t)}{\sqrt{n}}\, d\sqrt{n}\{\hat{\Lambda}_1(t) - \Lambda_1(t)\} - \int_0^\infty \frac{K(t)}{\sqrt{n}}\, d\sqrt{n}\{\hat{\Lambda}_0(t) - \Lambda_0(t)\}$$

$$+ \int_0^\infty \frac{K(t)}{\sqrt{n}}\, d\sqrt{n}\{\Lambda_1(t) - \Lambda_0(t)\}.$$

Note that under the sequence of the alternatives H_n we have

$$\sqrt{n}(\lambda_1(t) - \lambda_0(t)) \to \{c_1 - c_2 \ln(1 + \Lambda_0(t))\}\lambda_0(t).$$

If $Y_i/n \xrightarrow{\mathcal{P}} y_i$, $y = y_1 + y_2$, then under standard conditions (see Harington and Fleming, Ch. 7) we have the convergence

$$W \xrightarrow{\mathcal{D}} N(\mu, \sigma^2),$$

where

$$\mu = \frac{1}{c_2} \int_0^\infty \frac{y_0(t)y_1(t)}{y(t)} \{c_1 - c_2 \ln(1 + \Lambda_0(t))\}^2 d\Lambda_0(t).$$

So

$$T = W/\hat{\sigma} \xrightarrow{\mathcal{D}} N(a, 1), \quad T^2 \xrightarrow{\mathcal{D}} \chi^2(1, a)$$

where $a = \mu/\sigma$, and $\chi^2(1, a)$ denotes the chi-square distribution with one degree of freedom and the non-centrality parameter a. Note that under the cross-effects alternative the function under the integral does not change the sign.

The asymptotic power function of the test is

$$\beta = \lim_{n\to\infty} \mathbf{P}\left\{\left(\frac{T}{\hat\sigma}\right)^2 > \chi^2_{1-\alpha}(1) \mid H_n\right\} = \mathbf{P}\left\{\chi^2(1, a) > \chi^2_{1-\alpha}(1)\right\}.$$

The non-centrality parameter a can be estimated by

$$\hat a = \frac{\hat\gamma}{\hat\sigma}\int_0^\infty \frac{Y_0(t)Y_1(t)}{Y(t)}\{\hat\beta - \hat\gamma\ln(1+\tilde\Lambda_0(t-))\}^2 d\tilde\Lambda_0(t).$$

In the case of the modified score test the limit distribution of the statistic

$$X^2 = (\hat U_1, \hat U_2)\hat\Sigma^{-1}(\hat U_1, \hat U_2)^T$$

under the sequence of approaching alternatives is the non-central chi-square distribution with two degrees of freedom and the non-centrality parameter a estimated by

$$\hat a = (\hat\mu_1, \hat\mu_2)\hat\Sigma^{-1}(\hat\mu_1, \hat\mu_2)^T,$$

where

$$\hat\mu_1 = \int_0^\infty \frac{Y_0(t)Y_1(t)}{Y(t)}\{\hat\beta - \hat\gamma\ln(1+\tilde\Lambda_0(t-))\}d\tilde\Lambda_0(t),$$

$$\hat\mu_2 = -\int_0^\infty \frac{Y_0(t)Y_1(t)}{Y(t)}\ln(1+\tilde\Lambda_0(t-))\{\hat\beta - \hat\gamma\ln(1+\tilde\Lambda_0(t-))\}d\tilde\Lambda_0(t).$$

6. k-sample tests

Let us generalize the problem and consider the hypothesis

$$H_0 : S_0 = S_1 = \ldots = S_{k-1}$$

of the equality of k survival distributions.

Suppose that n_i objects of the ith group ($i = 0, \ldots, k-1$) are observed. Denote by T_{ij} and C_{ij} the failure and censoring times for the jth patient of the ith group, and set

$$X_{ij} = \min(T_{ij}, C_{ij}),\ \ \delta_i = \mathbf{1}_{\{T_{ij}\le C_{ij}\}},$$

$$N_{ij}(t) = \mathbf{1}_{\{T_{ij}\le t, \delta_{ij}=1\}},\ \ Y_{ij}(t) = \mathbf{1}_{\{X_{ij}\ge t\}},$$

$$N_i(t) = \sum_{j=1}^{n_i} N_{ij}(t), \quad Y_i(t) = \sum_{j=1}^{n_i} Y_{ij}(t), \quad N(t) = \sum_{i=0}^{k-1} N_i(t), \quad Y(t) = \sum_{i=0}^{k-1} Y_i(t).$$

Consider the following alternatives:

$$H_A : \lambda_i(t) = e^{\beta_i}\left\{1 + e^{\beta_i+\gamma_i}\Lambda_0(t)\right\}^{e^{-\gamma_i}-1}\lambda_0(t) \quad (i = 1, \ldots, k-1).$$

Denote by $\beta = (\beta_1, \ldots, \beta_{k-1})^T$, $\gamma = (\gamma_1, \ldots, \gamma_k)$ the vectors of the unknown parameters, an define the $(k-1)$-dimensional vector $x^{(i)} = (0, \ldots, 0, 1, 0, \ldots, 0)^T$ such that the unit stays in the i-th place. Then the alternative can be written in the form

$$H_A : \lambda_i(t) = e^{\beta^T x^{(i)}} \left\{1 + e^{(\beta+\gamma)^T x^{(i)}} \Lambda_0(t)\right\}^{e^{-\gamma^T x^{(i)}}-1} \lambda_0(t) \quad (i = 1, \dots, k-1).$$

This alternative contains the cases when some hazard rates or cumulative hazard rates intersect, some go away one from another or are proportional. Under the alternative the $(k-1)$-dimensional score functions $U_1(\beta,\gamma)$ and $U_2(\beta,\gamma)$ for the parameters β and γ estimation are

$$U_1(\beta,\gamma) = \sum_{i=1}^{k-1} \int_0^\infty x^{(i)} \left\{1 + (e^{\beta^T x^{(i)}} - e^{(\beta+\gamma)^T x^{(i)}}) \frac{\tilde\Lambda_0(t-,\beta,\gamma)}{1 + e^{(\beta+\gamma) x^{(i)}} \tilde\Lambda_0(t-,\beta,\gamma)}\right\}$$

$$\left\{dN_i(t) - Y_i(t) e^{\beta^T x^{(i)}} (1 + e^{(\beta+\gamma)^T x^{(i)}} \tilde\Lambda_0(t-,\beta,\gamma))^{e^{-\gamma^T x^{(i)}}-1} d\tilde\Lambda_0(t,\beta,\gamma)\right\},$$

$$U_2(\beta,\gamma) = \sum_{i=1}^{k-1} \int_0^\infty \left\{-e^{-\gamma^T x^{(i)}} \ln(1 + e^{(\beta+\gamma)^T x^{(i)}} \tilde\Lambda_0(t-,\beta,\gamma)) + (e^{\beta^T x^{(i)}} - e^{(\beta+\gamma)^T x^{(i)}})\right.$$

$$\left.\frac{\tilde\Lambda_0(t-,\beta,\gamma)}{1 + e^{(\beta+\gamma)^T x^{(i)}} \tilde\Lambda_0(t-,\beta,\gamma)}\right\}$$

$$\left\{dN_i(t) - Y_i(t) e^{\beta^T x^{(i)}} (1 + e^{(\beta+\gamma)^T x^{(i)}} \tilde\Lambda_0(t-,\beta,\gamma))^{e^{-\gamma^T x^{(i)}}-1} d\tilde\Lambda_0(t,\beta,\gamma)\right\},$$

where $\tilde\Lambda_0$ is defined recurrently from the equation :

$$\tilde\Lambda_0(t,\beta,\gamma) = \int_0^t \frac{dN(u)}{S^{(0)}(u-,\tilde\Lambda_0,\beta,\gamma)},$$

$$S^{(0)}(v,\tilde\Lambda_0,\theta) = \sum_{i=0}^{k-1} Y_i(v) e^{\beta^T x^{(i)}} \left\{1 + e^{(\beta+\gamma)^T x^{(i)}} \tilde\Lambda_0(v,\beta,\gamma)\right\}^{e^{-\gamma^T x^{(i)}}-1}.$$

Under H_0 we have

$$\hat U_1 := U_1(0,0) = \sum_{i=1}^{k-1} \int_0^\infty x^{(i)} \left\{dN_i(t) - Y_i(t) d\tilde\Lambda_0(t)\right\},$$

$$\hat U_2 := U_2(0,0) = -\sum_{i=1}^{k-1} \int_0^\infty x^{(i)} \ln(1 + \tilde\Lambda_0(t-)) \left\{dN_i(t) - Y_i(t) d\tilde\Lambda_0(t)\right\},$$

where

$$\tilde\Lambda_0(t) = \int_0^t \frac{dN(u)}{Y(u)}.$$

These statistics can be written as follows

$$\hat U_1 = \sum_{i=1}^{k-1} \int_0^\infty x^{(i)} \{dM_i(t) - Y_i \frac{dM(t)}{Y(t)}\},$$

$$\hat U_2 = -\sum_{i=1}^{k-1} \int_0^\infty x^{(i)} \ln(1 + \tilde\Lambda_0(t-)) \{dM_i(t) - Y_i \frac{dM(t)}{Y(t)}\},$$

where $M = \sum_{i=0}^{k-1} M_i$ and M_i are the martingales of the counting processes N_i.

Similarly as in the two-sample case, under standard conditions the limit distribution (as $n = \sum_{i=0}^{k-1} n_i \to \infty, n_i/n \to l_i \in (0,1)$) of the statistic

$$X^2 = ((\hat{U}_1)^T, (\hat{U}_2)^T)\hat{\Sigma}^{-1}((\hat{U}_1)^T, (\hat{U}_2)^T)^T$$

is the chi-square distribution with $2(k-1)$ degrees of freedom; here

$$\hat{\Sigma} = \begin{pmatrix} \hat{\Sigma}_{11} & \hat{\Sigma}_{12} \\ \hat{\Sigma}_{21} & \hat{\Sigma}_{22} \end{pmatrix},$$

where $\hat{\Sigma}_{21} = \hat{\Sigma}_{12}^T$,

$$\hat{\Sigma}_{11} = \sum_{i=1}^{k-1} \int_0^\infty \{(x^{(i)} - \bar{x})^{\otimes 2} Y_i + \bar{x}^{\otimes 2} Y_0\} \frac{dN(t)}{Y(t)}$$

$$\hat{\Sigma}_{12} = -\sum_{i=1}^{k-1} \int_0^\infty \ln(1 + \tilde{\Lambda}_0(t-))\{(x^{(i)} - \bar{x})^{\otimes 2} Y_i + \bar{x}^{\otimes 2} Y_0\} \frac{dN(t)}{Y(t)}$$

$$\hat{\Sigma}_{22} = -\sum_{i=1}^{k-1} \int_0^\infty \ln^2(1 + \tilde{\Lambda}_0(t-))\{(x^{(i)} - \bar{x})^{\otimes 2} Y_i + \bar{x}^{\otimes 2} Y_0\} \frac{dN(t)}{Y(t)},$$

$$\bar{x} = \frac{\sum_{j=1}^{k-1} x^{(j)} Y_j(t)}{Y(t)},$$

and $A^{\otimes 2} = AA^T$ for any column A.

7. Finite-sample null distribution and simulated power results

As in Stablein and Koutrouvelis (1985), simulations were performed using the unit exponential to examine the small-sample characteristics of the modified score test statistic under the null hypothesis. We calculated the significance level from 5000 replications of complete samples of given $n_0 = n_1$. The results are given in Table 1. It shows that with n_i increasing the significance level converges to 0.05. Note that in the case of Stablein and Koutrouvelis the rate of convergence is much slower.

Table 1. Significance level.

n_i	25	50	100	200	500
α	0.065	0.059	0.057	0.053	0.

We compared the power of the modified score statistics with simulated power results of Stablein and Koutrouvelis (1985). As in their study, we calculated power from 1000 replications for sample sizes of 25 and 50 uncensored observations per group. The significance level $\alpha = 0.05$. The results are given in Table 2 (simulation results for the Stablein and Koutrouvelis, Modified Smirnov, Koziol-Petkau and logrank statistics are taken from Stablein and Koutrouvelis (1985), simulation results for the modified score and second test are done by us). The following situations were simulated:

1) $\lambda_0(t) = 2, \quad \lambda_1(t) = 1.$

The hazard rates are constant and do not intersect. The survival functions do not intersect, too. We have the case of proportional hazards. Naturally, the logrank statistic performs best and the test of Stablein-Koutrouvelis performs worst. Other statistics have similar powers.

2) $S_0(t) = e^{-2t}$, $S_1(t) = e^{-t^5}$.

3) $S_0(t) = e^{-2t}$, $S_1(t) = e^{-t^5}$.

In situations 2) and 3) the hazard rates and the survival functions cross. The modified score and the second test do considerably better than all other tests.

4)

$$\lambda_0(t) = \begin{cases} 1 & \text{if} \quad 0 \le t < 0.8, \\ 2 & \text{if} \quad t \ge 0.8, \end{cases} \qquad \lambda_1(t) = \begin{cases} 1 & \text{if} \quad 0 \le t < 0.8, \\ 0.2 & \text{if} \quad t \ge 0.8, \end{cases}$$

It is the situation of late difference, where no crossing of the hazard functions occurs. The modified score test does considerably better than all other tests.

5)

$$\lambda_0(t) = \begin{cases} 2 & \text{if} \quad 0 \le t < 0.1, \\ 3 & \text{if} \quad 0.1 \le t < 0.4, \\ 0.75 & \text{if} \quad t \ge 0.4, \end{cases} \qquad \lambda_1(t) = \begin{cases} 2 & \text{if} \quad 0 \le t < 0.2, \\ 0.75 & \text{if} \quad 0.2 \le t < 0.4, \\ 3 & \text{if} \quad t \ge 0.4, \end{cases}$$

It is not very realistic situation. The hazard rates cross with a jump at the point 0.4 and after crossing are constant. The survival functions also cross. In this situation the Stablein-Koutrouvelis, second, and modified score statistics do best.

In summary, in natural situations with crossing of the survival functions and some other situations the modified score statistic and the second statistic have considerably better power than other statistics.

Table 2. Power simulation.

Nr	n_1	$B_{n,r}$	*Modif. Smirn*	*Koziol Petkau*	*Log rank*	*Modif. score*	*Second statist.*
1	25	0.319	0.555	0.588	0.658	0.570	
	50	0.762	0.831	0.908	0.928	0.871	
2	25	0.737	0.572	0.339	0.469	0.999	0.660
	50	0.987	0.899	0.683	0.811	1.000	0.802
3	25	0.561	0.279	0.086	0.087	0.998	0.992
	50	0.908	0.554	0.209	0.148	1.000	1.000
4	25	0.571	0.802	0.429	0.596	0.949	
	50	0.954	0.985	0.782	0.883	0.999	
5	25	0.730	0.530	0.266	0.065	0.666	0.758
	50	0.986	0.868	0.508	0.053	0.950	0.968

8. Real data analysis

Stablein and Koutrouvelis (1985) studied the well known two-sample data of the Gastrointestinal Tumor Study Group concerning effects of chemotherapy and chemotherapy plus radiotherapy on the survival times of gastric cancer patients. The number of patients is 90. Survival times of chemotherapy (group 0 of size 45) and chemiotherapy plus radiotherapy (group 1 of size 45) patients are as follows:

Chemotherapy:

1 63 105 129 182 216 250 262 301 301 342 354 356 358 380 383 383 388 394 408 460 489 499 523 524 535 562 569 675 676 748 778 786 797 955 968 1000 1245 1271 1420 1551 1694 2363 2754* 2950*;

Chemotherapy plus Radiotherapy

17 42 44 48 60 72 74 95 103 108 122 144 167 170 183 185 193 195 197 208 234 235 254 307 315 401 445 464 484 528 542 567 577 580 795 855 1366 1577 2060 2412* 2486* 2796* 2802* 2934* 2988*.

* means censoring.

By plotting the two Kaplan-Meier estimators of survival functions pertaining to the both treatment groups, a crossing-effect phenomenon is clearly manifest. The resulting inference indicates that the radiotherapy would first be detrimental to a patient's survival but becomes beneficial later on.

To confirm graphic results we use test statistics for checking the hypothesis of the equality of survival distributions versus the cross-effects alternative.

The classical logrank statistic (it appears that it is based on the first modified score statistic $\hat{U}_1$) does not reject the null hypothesis (P-value $P = 0.64$, the value of the statistic is 0.23). The second weighted logrank test based on the second modified score statistic $\hat{U}_2$ also does not reject the null hypothesis (P-value $P = 0.21$).

The Renyi type statistics rejects but not very strongly H_0 (P-value $P = 0,053$, the value of the statistics Q is 2,20.

The test of Stablein et Koutrouvelis rejects the null hypothesis (the value of the test statistic is 3,78 and the critical value at the significance level 0,01 is 3,41).

The modified score statistic rejects H_0 very strongly (P-value $P = 0,00111$, the value of the statistic X^2 is 13,61). Note that

$$\hat{\Sigma} = \begin{pmatrix} 19.884 & -9.875 \\ -9.875 & 6.988 \end{pmatrix}, \quad \hat{\Sigma}^{-1} = \begin{pmatrix} 0.1687 & 0.2384 \\ 0.2384 & 0.4800 \end{pmatrix}.$$

The second test rejects H_0 with even smaller P-value ($P = 0,00081$, the value of the statistic T is 3,323). The estimators of the parameters β and γ are $\hat{\beta} = 1.8945$, $\hat{\gamma} = 1.3844$, the estimator of the intersection point of the hazard rates is $\hat{t}_0 = 382.9$.

So both proposed statistics reject the null hypothesis with smaller P-value than other tests. And both can be used not only for singly but also for randomly censored samples.

1)$S_0(t) = e^{-2t}$, $S_1(t) = e^{-t}$

1000 iterations: n=50 : puis=554 (nombre de cas quand Chi carr ¿=5.991)

n=100 : puis=871

2000 iterations :

n=50 : puis=1140/2000=0.570.

n=100 : puis=1738/2000=0,869.

REFERENCES

Aalen, O. (1978) Nonparametric inference for the family of counting processes. *Ann. Statist.*, **6**, 701-726.
Andersen, P.K., Borgan, O., Gill, R.D.& Keiding, N. (1982). Linear non-parametric tests for comparison of counting processes, with application to (with discussion). *International Statistical Review*, **50**, 219-258. Amendment: **52**, 225 (1984).
Andersen, P.K. (1983). Comparing survival distributions with hazard ratio estimates. *Scand.J. Statist.*, **10**, 77-85.
Andersen, P.K., Borgan, O., Gill, R.D.& Keiding, N. (1993). *Statistical Models Based on Counting Processes.* New York: Springer.
Breslow, N.E. (1970). A generalized Kruskal-Wallis test for comparing K samples subject to unequel patterns of censorship. *Biometrika*, **57**, 579-594.
Brookmeyer, R. and Crowley, J.J. (1982). A k-sample median test for censored data. *J.Amer.Statist.Assoc.*, **77**, 433-440.
Gill, R.D. *Censoring and Stochastic Integrals* (1980). Amsterdam: Mathematical Centre Tracts 124, Mathematisch Centrum.
Bagdonavičius, V. and Nikulin, M. (2002). *Accelerated Life Models: Modeling and Statistical Analysis.* Boca Raton: Chapman and Hall/CRC.
Fleming, T.R., O'Fallon, J.R., O'Brien, P.C., Harrington, D.F. (1980). Modified Kolmogorov-Smirnov Test Procedures with Application to Arbitrary Right Censored Data. *Biometrics*, **36**, 607-626.
Fleming, T.R.and Harrington, D.P. (1981). A class of hypothesis tests for one and two samples of censored survival data. *Comm. Statist.*, **10**, 763-794.
Fleming, T.R., Harrington, D.P., and O'Sullivan, M. (1987). Supremum versions of the log-rank and generalised Wilcoxon statistics. *J. Amer. Statist. Assoc.*, **82**, 312-320.
Gehan, E.A. (1965). A generalized Wilcoxon test for comparing arbitrary singly censored samples. *Biometrika*, **52**, 203-223.
Harrington, D.P. and Fleming, T.R.(1982). A class of rank test procedures for censored survival data. *Biometrika*, **69**, 133-143.
Hjort, T.R.(1984). Local asymptotical power of linear nonparametric tests for comparison of counting processes. Report 763, Norwegian Coputing Center.
Kalbfleisch, J.D. and Prentice, R.L. (1980) *The Statistical Analysis of Failure Time Data.* New York: John Wiley and Sons.
Klein and Moeschenberger (1997). *Statistics for Biology and Health.* New York: Springer.
Koziol, J.A. (1978). A Two-Sample Cramer-Von Mises Test for Randomly Censored Data. *Biometrical Journal*, **20**, 603-608.
Peto, R. and Peto, J. (1972). Asymptotically efficient rank invariant test procedures (with discussion). *J. Roy. Statist. Soc.*, A **135**, 185-206.
Prentice, R.L. (1978). Linear rank tests with right censored data. *Biometrika*, **65**, 167-179.
Schumacher, M. (1984) Two-sample tests of Cramer-von Mises and Kolmogorov-Smirnov type for random censored data. *Internat. Statist. Rev.*, **52**, 263-281.
Stablein, D. M., Koutrouvelis, I. A. (1985). A two sample test sensitive to crossing hazards in uncensored and singly censored data. *Biometrics* **41**, 643-652.
Tarone, R.E. and Ware, J.H. (1977). On distribution-free tests for equality for survival distributions. *Biometrika*, **64**, 156-160.