Annals of the Institute of Statistical Mathematics (2022) 74:1067–1083
https://doi.org/10.1007/s10463-022-00823-9

On comparing competing risks using the ratio of their cumulative incidence functions

Hammou El Barmi[1]

Received: 7 August 2021 / Revised: 8 March 2022 / Accepted: 11 March 2022 /
Published online: 14 May 2022

Abstract
For $1 \leq i \leq r$, let F_i be the cumulative incidence function (CIF) corresponding to the *ith* risk in an r-competing risks model. We assume a discrete or a grouped time framework and obtain the maximum likelihood estimators (m.l.e.) of these CIFs under the restriction that $F_i(t)/F_{i+1}(t)$ is nondecreasing, $1 \leq i \leq r-1$. We also derive the likelihood ratio tests for testing for and against this restriction and obtain their asymptotic distributions. The theory developed here can also be used to investigate the association between a failure time and a discretized or ordinal mark variable that is observed only at the time of failure. To illustrate the applicability of our results, we give examples in the competing risks and the mark variable settings.

Keywords Competing risks · Cumulative incidence function · Likelihood ratio test · Chi-bar squared distribution

1 Introduction

In a competing risks setting, a unit or a subject is exposed to several risks at the same time but their actual failure (or death) is due to exactly one of them. In a study with a mark variable, interest is in exploring the association between the failure time of a subject and the level of a mark variable that is measured only when the subject fails (or dies). What is observed is (T, δ), where T is the time of failure and δ is the cause of failure or the level of the mark variable at the time of failure.

In these situations, statistical inferences are typically based the CIFs, the sub-survival functions (SSF) or the cause specific hazard rates (CSHR) corresponding to these risks. The CIF due to the *ith* risk is a sub-distribution function that is defined as

✉ Hammou El Barmi
hammou.elbarmi@baruch.cuny.edu

[1] Paul H. Chook Department of Information Systems and Statistics, Baruch College, The City University of New York, One Baruch Way New York, New York, NY 10010, USA

$$F_i(t) = P[T \le t,\ \delta = i], i = 1, 2, \dots, r, \tag{1}$$

with $F(t) = \sum_{i=1}^{r} F_i(t)$ being the distribution function of T. Here, r is the number of competing risk or the number of the levels of the mark variable. Its corresponding SSF and CSHR are defined, respectively, as

$$S_i(t) = P[T \ge t,\ \delta = i], i = 1, 2, \dots, r,$$

and

$$\lambda_i(t) = \lim_{\Delta t \to 0} \frac{1}{\Delta t} \mathrm{P}\,[t \le T < t + \Delta t,\ \delta = i \mid T \ge t].$$

Note that $S(t) = \sum_{i=1}^{r} S_i(t)$ and $\lambda(t) = \sum_{i=1}^{r} \lambda_i(t)$ represent the survival function and hazard rate of T, respectively. In the continuous case, the *ith* CIF can be expressed in terms of the CSHR by the following relation:

$$F_i(t) = \int_0^t \lambda_i(u) S(u)\, du, i = 1, 2, \dots, r. \tag{2}$$

Similar relations hold also in the discrete case.

Many tests are available for comparing the CIFs corresponding to competing risks. For the case of $r = 2$, El Barmi et al. (2004) and Aly et al. (1994) used Kolmogorov-Smirnov type statistics for this situation in the continuous case while El Barmi and Kochar (2002) developed a LRT for the same problem in the discrete or grouped data situation. Extensions of these tests to the r-sample case ($r \ge 3$) have also been considered in El Barmi and Mukerjee (2006) and in El Barmi et al. (2006) in the general and the discrete/grouped data case, respectively.

Most of the methods that have been developed in the literature for comparing the CIFs are based on their differences. However, when interest is in the comparison of the conditional distributions of T given the different risks, these difference are not suitable. In this case, as pointed out in Dauxois and Kirmani (2003), interest should be in the temporal functions $\phi_i(t) \equiv F_i(t)/F_{i+1}(t), i = 1, 2, \dots, r-1$, since $\phi_i(t)$ is proportional to $h_i(t) \equiv P[T \le t | \delta = i]/P[T \le t | \delta = i+1]$ and it is nondecreasing if and only if $h_i(t)$ is also nondecreasing. This is the case if and only if the conditional distribution of T given $\delta = i$ is stochastically larger than the conditional distribution of T given $\delta = i + 1$ in the reversed hazard rate ordering. This implies, in particular, that the distribution T given $\delta = i$ is stochastically larger than the conditional distribution of T given $\delta = i + 1$. For more on the reversed hazard rate and stochastic orderings, see Shaked and Shanthikumar (2006).

The aim of this paper is to develop in the discrete/grouped data case the LRTs of H_0 against $H_1 - H_0$ and H_1 against $H_2 - H_1$ where

$$H_0 : \phi_i(t) \quad \text{is constant}, \ i = 1, 2, \dots, r-1,$$
$$H_1 : \phi_i(t) \quad \text{is nondecreasing in t}, \ i = 1, 2, \dots, r-1,$$

and H_2 imposes no constraints on these CIFs. We note that H_0 holds if and only if T and δ are independent. We will also consider testing $H_0^* : F_1 = F_2 = \ldots = F_r$ against $H_1 - H_0^*$.

Discrete failure times arise in competing risk and mark variable studies when the recorded times to failure are grouped in intervals. A discrete mark variable can result by grouping a continuous mark variable in intervals or observing an ordinal categorical variable at time of failure.

To our knowledge, very little attention has been given to the problem of testing for or against H_1. The only test that we are aware of is a nonparametric test, developed in the general case for H_0 against $H_1 - H_0$ when $r = 2$, in Dauxois and Kirmani (2003). Recently, Al-Kandari and El Barmi (2022) derived the nonparametric m.l.e.s of the CIFs under H_1 in the continuous case and showed that they are inconsistent. They also developed projection type estimators that are consistent, studied the weak convergence of the resulting process and proposed a test for H_0^* against $H_1 - H_0^*$.

Besides many applications in the health sciences, our procedure has potential applications in industrial accelerated life tests. For example, when comparing different brands of a component from two different suppliers, the components may be tested in series. The components are functioning in the same environment and their times to failure are generally dependent. The system in this case will fail as soon as one of the components fails. The theory we develop here will allow us to test whether these components are of the same quality against the ordered alternative, thus leading to early identification of weak components.

The rest of the paper is organized as follows. In Sect. 2, we obtain the m.l.e.s of the CIFs under H_0, H_1 and H_2. In Sect. 3, we derive the LRT statistics for testing H_0 against $H_1 - H_0$ and H_1 against $H_2 - H_1$ and study their asymptotic distributions. To illustrate the theory given in the previous sections, we shall present, in Sect. 4, two examples, one from a competing risks study using some data from Hoel (1972) and one from a clinical trial study designed to investigate the association between survival and a mark variable. Throughout we use $\xrightarrow{d}$ to denote convergence in distribution and we assume that 0/0=0.

2 Maximum likelihood estimation

Suppose that n individuals are exposed to r risks and that their times and causes of failure form a random sample from (T, δ). Assume that the failures can only occur at the discrete time points $t_1 < t_2 < \ldots < t_m$ and let p_{ij} and d_{ij} represent the probability of failure and the number of failures from the *ith* cause at time t_j, respectively. We also assume that the time point t_m is large enough that all individuals have died by this time.

Then

$$F_i(t_j) = P(T \le t_j, \delta = i) = \sum_{\ell=1}^{j} p_{i\ell}, i = 1, 2, \ldots, r, j = 1, 2, \ldots, m. \tag{3}$$

In this section, we derive the m.l.e.s of these CIFs under the hypotheses H_0 and H_1.

For $i = 1, 2, \dots, r$, let $\mathbf{p}_i = (p_{i1}, p_{i2}, \dots, p_{im})^T$. The likelihood function is given by

$$L(\mathbf{p}_1, \mathbf{p}_2, \dots, \mathbf{p}_r) \propto \prod_{i=1}^{r} \prod_{j=1}^{m} p_{ij}^{d_{ij}}$$

since, in this case, $\{d_{ij}, 1 \le i \le r, 1 \le j \le m\}$ has a multinomial distribution with parameters n and $\{p_{ij}, 1 \le i \le r, 1 \le j \le m\}$.

Clearly the unrestricted m.l.e. of p_{ij} is $\hat{p}_{ij} = \frac{d_{ij}}{n}$, the relative frequency of the event $\{T = t_j, \delta = i\}$, and the unconstrained m.l.e. of $F_i(t_j)$ is $\hat{F}_i(t_j) = \sum_{\ell=1}^{j} \hat{p}_{i\ell}$.

To find the m.l.e.s of the p_{ij}s under H_0 and H_1, we reparametrize the problem by letting

$$\theta_{ij} = \frac{F_i(t_j)}{F_i(t_{j+1})}, i = 1, 2, \dots, r,\ j = 1, 2, \dots, m-1, \quad \text{and} \quad \theta_{im} = F_i(t_{im}), i = 1, 2, \dots, r,$$

so that, for $i = 1, 2, \dots, r$,

$$p_{i1} = \prod_{\ell=1}^{r} \theta_{ij} \ \text{ and } \ p_{ij} = (1 - \theta_{i,j-1}) \prod_{\ell=j}^{m} \theta_{ij}, j = 2, 3, \dots, m. \tag{4}$$

In terms of this new parametrization, H_0 and H_1 are equivalent to

$$H_0: \quad \theta_{1j} = \theta_{2j} = \dots = \theta_{rj}, j = 1, 2, \dots, m-1, \sum_{i=1}^{r} \theta_{im} = 1,$$

$$H_0: \quad \theta_{1j} \le \theta_{2j} \le \dots \le \theta_{rj}, j = 1, 2, \dots, m-1, \sum_{i=1}^{r} \theta_{im} = 1,$$

and the likelihood function is

$$L \propto \prod_{j=1}^{m-1} \prod_{i=1}^{r} (1 - \theta_{ij})^{d_{i,j+1}} \theta_{ij}^{\sum_{\ell=1}^{j} d_{i\ell}} \times \prod_{i=1}^{r} \theta_{im}^{d_{i+}} \equiv \prod_{j=1}^{m} L_j(\boldsymbol{\theta}_j)$$

where $\boldsymbol{\theta}_j = (\theta_{1j}, \theta_{2j}, \dots, \theta_{rj})^T, j = 1, 2, \dots, m$,

$$L_j(\boldsymbol{\theta}_j) = \prod_{i=1}^{r} (1 - \theta_{ij})^{d_{i,j+1}} \theta_{ij}^{\sum_{\ell=1}^{j} d_{i\ell}}, j = 1, 2, \dots, m-1, L_m(\boldsymbol{\theta}_m) = \prod_{i=1}^{r} \theta_{im}^{d_{i+}},$$

and $d_{i+} = \sum_{\ell=1}^{m} d_{ij}, i = 1, 2, \ldots, r$. As such, L factors into m terms and a careful inspection of the constraints corresponding to the hypotheses H_0, H_1 and H_2 shows that maximizing L under these hypotheses is equivalent to maximizing each term separately from the remaining terms.

It straightforward to see that, under H_2, L_j is maximized at $\boldsymbol{\theta}_j^{(2)} = (\hat{\theta}_{1j}^{(2)}, \hat{\theta}_{2j}^{(2)}, \ldots, \hat{\theta}_{rj}^{(2)})^T$ where, for $1 \le i \le r$,

$$\hat{\theta}_{ij}^{(2)} = \begin{cases} \dfrac{\sum_{\ell=1}^{j} d_{i\ell}}{\sum_{\ell=1}^{j+1} d_{i\ell}} = \frac{\hat{F}_i(t_j)}{\hat{F}_i(t_{j+1})}, \ j = 1, 2, \ldots, m-1, \\ \frac{d_{i+}}{n} = \hat{F}_i(t_m), \qquad j = m. \end{cases}$$

Under H_0, L_j is maximized at $\boldsymbol{\theta}_j^{(0)} = (\hat{\theta}_{1j}^{(0)}, \hat{\theta}_{2j}^{(0)}, \ldots, \hat{\theta}_{rj}^{(0)})^T$ where ,for $1 \le i \le r$,

$$\hat{\theta}_{ij}^{(0)} = \begin{cases} \dfrac{\sum_{i=1}^{r} \sum_{\ell=1}^{j} d_{i\ell}}{\sum_{i=1}^{r} \sum_{\ell=1}^{j+1} d_{i\ell}} = \dfrac{\sum_{i=1}^{r} d_{i+1} \hat{F}_i(t_j)}{\sum_{i=1}^{r} d_{i+} \hat{F}_i(t_{j+1})}, \ j = 1, 2, \ldots, m-1, \\ \frac{d_{i+}}{n} = \hat{F}_i(t_m), \qquad j = m. \end{cases}$$

Under H_1, maximizing $L_j, j = 1, 2, \ldots, m-1$, is the classical bioassay problem, as discussed in Robertson et al. (1988), p. 32. Its solution is given by

$$\boldsymbol{\theta}_j^{(1)} = \left(\hat{\theta}_{1j}^{(1)}, \hat{\theta}_{2j}^{(1)}, \ldots, \hat{\theta}_{rj}^{(1)}\right)^T = E_{\mathbf{w}_j}\left[\hat{\boldsymbol{\theta}}_j^{(2)} | \mathcal{I}\right],$$

the least squares projection of $\hat{\boldsymbol{\theta}}_j^{(2)}$ onto the cone $\mathcal{I} = \{\mathbf{z} \in \mathbf{R}^r : z_1 \le z_2 \le \ldots \le z_r\}$ with weights $\mathbf{w}_j = \left(\hat{F}_1(t_{j+1}), \hat{F}_2(t_{j+1}), \ldots, \hat{F}_r(t_{j+1})\right)^T$. In addition, L_m is maximized by $\hat{\theta}_{im}^{(1)} = \frac{d_{i+1}}{n} = \hat{F}_i(t_m), i = 1, 2, \ldots, r$. In passing, we mention that several algorithms have been developed for computing the least squares projection $E_{\mathbf{w}_j}\left[\hat{\boldsymbol{\theta}}_j^{(2)} | \mathcal{I}\right]$. They include the pool-adjacent-violators algorithm (PAVA) and they are discussed at length in Robertson et al. (1988). We have the following theorem

Theorem 1 *The m.l.e. of* $(\mathbf{p}_1, \mathbf{p}_2, \ldots, \mathbf{p}_r)$ *under* $H_s, s = 0, 1,$ *is given by* $\left(\hat{\mathbf{p}}_1^{(s)}, \hat{\mathbf{p}}_2^{(s)}, \ldots, \hat{\mathbf{p}}_r^{(s)}\right)$ *where, for* $i = 1, 2, \ldots, r$,

$$\hat{p}_{i1}^{(s)} = \prod_{\ell=1}^{r} \theta_{ij}^{(s)} \quad \text{and} \quad p_{ij}^{(s)} = \left(1 - \theta_{i,j-1}^{(s)}\right) \prod_{\ell=j}^{m} \theta_{ij}^{(s)}, j = 2, 3, \ldots, m. \tag{5}$$

In addition, the m.l.e. of $F_i(t_j)$ *is* $\hat{F}_i^{(s)}(t_j) = \sum_{j=1}^{r} \hat{p}_{ij}^{(s)}$.

Proof The m.l .e.s in (5) are obtained by plugging the m.l.e.s of the θ_{ij}s under $H_s, s = 0, 1$, into (4). □

3 Likelihood ratio tests

In this section, we derive the LRTs for testing H_0 against $H_1 - H_0$ and H_1 against $H_2 - H_1$. We will also show that their limiting distributions are of chi-bar squared type and give the expression of the weighting values.

3.1 Testing H_0 against $H_1 - H_0$

We begin by testing H_0 against $H_1 - H_0$. The LRT in this case rejects H_0 for small values of

$$\Lambda_{01} = \frac{\prod_{j=1}^{m} L_j(\hat{\boldsymbol{\theta}}_j^{(0)})}{\prod_{j=1}^{m} L_j(\hat{\boldsymbol{\theta}}_j^{(1)})} = \prod_{j=1}^{m-1} \prod_{i=1}^{r} \left[\frac{\hat{\theta}_{ij}^{(0)}}{\hat{\theta}_{ij}^{(1)}}\right]^{\sum_{\ell=1}^{j} d_{ij}} \left[\frac{1 - \hat{\theta}_{ij}^{(0)}}{1 - \hat{\theta}_{ij}^{(1)}}\right]^{d_{i,j+1}}$$

because the m.l.e. of $\boldsymbol{\theta}_m$ is the same under both H_0 and H_1. Equivalently, it rejects H_0 for large values of the statistic

$$\begin{aligned} T_{01} &= -2\log(\Lambda_{01}) \\ &= 2\sum_{j=1}^{m}\sum_{i=1}^{r}\left[\left(\sum_{\ell=1}^{j} d_{ij}\right)\log\left\{\frac{\hat{\theta}_{ij}^{(1)}}{\hat{\theta}_{ij}^{(0)}}\right\} + d_{ij}\log\left\{\frac{1-\hat{\theta}_{ij}^{(1)}}{1-\hat{\theta}_{ij}^{(0)}}\right\}\right]. \end{aligned}$$

Expanding $\log(\hat{\theta}_{ij}^{(1)})$ and $\log(\hat{\theta}_{ij}^{(0)})$ about $\hat{\theta}_{ij}^{(2)}$ and $\log\left(1 - \hat{\theta}_{ij}^{(1)}\right)$ and $\log\left(1 - \hat{\theta}_{ij}^{(0)}\right)$ about $1 - \hat{\theta}_{ij}^{(2)}$ using a Taylor expansion with a second remainder, we get, using the properties of the isotonic regression (see Robertson et al. 1988),

$$T_{01} = \sum_{j=1}^{m-1} \sum_{i=1}^{r} \left[\left\{ \frac{\sum_{\ell=1}^{j} d_{i\ell}}{\alpha_{1,ij}^2} + \frac{d_{i,j+1}}{\alpha_{2,ij}^2} \right\} \left(\hat{\theta}_{ij}^{(2)} - \hat{\theta}_{ij}^{(0)} \right)^2 - \left\{ \frac{\sum_{\ell=1}^{j} d_{i\ell}}{\alpha_{3,ij}^2} + \frac{d_{i,j+1}}{\alpha_{4,ij}^2} \right\} \left(\hat{\theta}_{ij}^{(1)} - \hat{\theta}_{ij}^{(0)} \right)^2 \right]$$
$$= \sum_{j=1}^{m-1} \sum_{i=1}^{r} n\hat{F}_i(t_{j+1}) \left[\left\{ \frac{\hat{\theta}_{ij}^{(2)}}{\alpha_{1,ij}^2} + \frac{1-\hat{\theta}_{ij}^{(2)}}{\alpha_{2,ij}^2} \right\} \left(\hat{\theta}_{ij}^{(2)} - \hat{\theta}_{ij}^{(0)} \right)^2 - \left\{ \frac{\hat{\theta}_{ij}^{(2)}}{\alpha_{3,ij}^2} + \frac{1-\hat{\theta}_{ij}^{(2)}}{\alpha_{4,ij}^2} \right\} \left(\hat{\theta}_{ij}^{(1)} - \hat{\theta}_{ij}^{(0)} \right)^2 \right]$$

where

$$\max \left\{ |\alpha_{1,ij} - \hat{\theta}_{ij}^{(2)}|, |\alpha_{2,ij} - (1-\hat{\theta}_{ij}^{(2)})| \right\} \leq |\hat{\theta}_{ij}^{(0)} - \hat{\theta}_{ij}^{(2)}| \tag{6}$$

and

$$\max \left\{ |\alpha_{3,ij} - \hat{\theta}_{ij}^{(2)}|, |\alpha_{4,ij} - (1-\hat{\theta}_{ij}^{(2)})| \right\} \leq |\hat{\theta}_{ij}^{(1)} - \hat{\theta}_{ij}^{(2)}|. \tag{7}$$

When H_0 is true, the right hand sides of (6) and (7) go to zero almost surely. This implies that T_{01} is asymptotically equivalent to

$$\begin{aligned} &\sum_{j=1}^{m-1} \sum_{i=1}^{r} \frac{n\hat{F}_i(t_{j+1})}{\hat{\theta}_{ij}^{(0)}\left(1-\hat{\theta}_{ij}^{(0)}\right)} \left[\left(\hat{\theta}_{ij}^{(2)} - \hat{\theta}_{ij}^{(0)} \right)^2 - \left(\hat{\theta}_{ij}^{(1)} - \hat{\theta}_{ij}^{(0)} \right)^2 \right] \\ &= \sum_{j=1}^{m-1} \sum_{i=1}^{r} \frac{n\hat{F}_i(t_{j+1})}{\hat{\theta}_{ij}^{(0)}\left(1-\hat{\theta}_{ij}^{(0)}\right)} \left(\hat{\theta}_{ij}^{(1)} - \hat{\theta}_{ij}^{(0)} \right)^2 \end{aligned} \tag{8}$$

where the last equality holds because

$$\sum_{i=1}^{r} \hat{F}_i(t_{j+1}) \left(\hat{\theta}_{ij}^{(2)} - \hat{\theta}_{ij}^{(0)} \right)^2 = \sum_{i=1}^{r} \hat{F}_i(t_{j+1}) \left(\hat{\theta}_{ij}^{(2)} - \hat{\theta}_{ij}^{(1)} \right)^2 + \sum_{i=1}^{r} \hat{F}_i(t_{j+1}) \left(\hat{\theta}_{ij}^{(1)} - \hat{\theta}_{ij}^{(0)} \right)^2$$

by the properties of isotonic regression (see Robertson et al., 1988). The right hand side of (8) can be expressed as

$$\sum_{j=1}^{m-1} \sum_{i=1}^{r} \frac{\hat{F}_i(t_{j+1})}{\hat{\theta}_{ij}^{(0)}\left(1-\hat{\theta}_{ij}^{(0)}\right)} \left[\sqrt{n}\left(\hat{\theta}_{ij}^{(1)} - \theta_{ij} \right) - \sqrt{n}\left(\hat{\theta}_{ij}^{(0)} - \theta_{ij} \right) \right]^2.$$

Using the central limit theorem, we get

$$\sqrt{n}\left[\hat{\mathbf{p}}_1^T - \mathbf{p}_1^T, \hat{\mathbf{p}}_2^T - \mathbf{p}_2^T, \dots, \hat{\mathbf{p}}_r^T - \mathbf{p}_r^T\right]^T \xrightarrow{d} N(\mathbf{0}, \Sigma) \tag{9}$$

where $\Sigma = \left(\sigma_{i_1 j_1, i_2 j_2}\right)$ with $\sigma_{i_1 j_1, i_2 j_2} = p_{i_1 j_1}(\delta_{i_1 j_1, i_2 j_2} - p_{i_2 j_2})$. Here, $\delta_{i_1 j_1, i_2 j_2}$ is the Kronecker delta. i.e. $\delta_{i_1 j_1, i_2 j_2} = 1$ if $(i_1, j_1) = (i_2, j_2)$ and 0 otherwise.

Straightforward but tedious application of the delta method gives under $H_0 : \theta_{1j} = \theta_{2j} = \dots = \theta_{rj} \equiv \theta_j^{(0)}, j = 1, 2, \dots, m-1, \sum_{i=1}^{r} \theta_{im} = 1,$

$$\sqrt{n}\begin{bmatrix} \hat{\boldsymbol{\theta}}_1^{(2)} - \boldsymbol{\theta}_1^{(0)} \\ \hat{\boldsymbol{\theta}}_2^{(2)} - \boldsymbol{\theta}_2^{(0)} \\ \vdots \\ \hat{\boldsymbol{\theta}}_{m-1}^{(2)} - \boldsymbol{\theta}_{m-1}^{(0)} \end{bmatrix} \xrightarrow{d} N(\mathbf{0}, \Gamma) \tag{10}$$

where $\boldsymbol{\theta}_j^{(0)} = (\theta_j^{(0)}, \theta_j^{(0)}, \dots, \theta_j^{(0)})^T, \Gamma = \text{ diag } (\Gamma_1, \Gamma_2, \dots, \Gamma_{m-1})$ where

$$\Gamma_i = \text{ diag } (\gamma_{i,11}, \gamma_{i,22}, \dots, \gamma_{i,rr})$$

with $\gamma_{i,jj} = \theta_j^{(0)}(1 - \theta_j^{(0)})/F_i(t_{j+1}), j = 1, 2, \dots, m-1$. This implies that, if $X_{ij}, 1 \le i \le r, 1 \le j \le m-1$, are independent random variables such that X_{ij} has $N(0, \gamma_{i,jj}), \mathbf{X}_j = (X_{1j}, X_{2j}, \dots, X_{rj})^T$ and $\mathbf{w}_j = (F_1(t_{j+1}), F_2(t_{j+1}), \dots, F_r(t_{j+1}))^T$, $j = 1, 2, \dots, m-1$, then

$$\begin{aligned} &\sum_{j=1}^{m-1} \sum_{i=1}^{r} \frac{\hat{F}_i(t_{j+1})}{\hat{\theta}_{ij}^{(0)}(1 - \hat{\theta}_{ij}^{(0)})} \left[\sqrt{n}(\hat{\theta}_{ij}^{(1)} - \theta_{ij}) - \sqrt{n}(\hat{\theta}_{ij}^{(0)} - \theta_{ij})\right]^2 \\ &\xrightarrow{d} \sum_{j=1}^{m-1} \sum_{i=1}^{r} \frac{F_i(t_{j+1})}{\theta_{ij}^{(0)}(1 - \theta_{ij}^{(0)})} \left[E_{\mathbf{w}_j}\left[\mathbf{X}_j | \mathcal{I}\right]_i - \bar{X}_j\right]^2 \\ &\xrightarrow{d} \sum_{j=1}^{m-1} \sum_{i=1}^{r} \left[E_{\mathbf{w}_j}\left[\mathbf{X}_j | \mathcal{I}\right]_i - \bar{X}_j\right]^2 \left[\text{ Var } (X_{ij})\right]^{-1} \\ &\equiv \sum_{j=1}^{m-1} T_{01,j} \end{aligned}$$

where

$$\bar{X}_j = \frac{\sum_{i=1}^r w_{ij} X_{ij}}{\sum_{i=1}^r w_{ij}}$$

and

$$T_{01,j} = \sum_{i=1}^{r} \left[E_{\mathbf{w}_j}\left[\mathbf{X}_j | \mathcal{I}\right]_i - \bar{X}_j\right]^2 \left[\text{ Var } (X_{ij})\right]^{-1}, j = 1, 2, \dots, m-1.$$

Since, under H_0, the time until failure T and the cause of failure δ are independent, $\mathbf{w}_j = F(t_{j+1})\mathbf{q}$ where $\mathbf{q} = (P(\delta = 1), P(\delta = 2), \dots, P(\delta = r))^T$. Careful scrutiny then shows that

$$T_{01,j} = \sum_{i=1}^{r} q_i \Big[E_{\mathbf{q}_j}\big[\tilde{\mathbf{X}}_j|\mathcal{I}\big]_i - \bar{\tilde{X}}_j\Big]^2, j = 1, 2, \dots, m-1,$$

where $\tilde{X}_{ij}$ are independent zero mean normal variates with Var $(\tilde{X}_{ij}) = 1/q_i$ for all (i,j), $\tilde{\mathbf{X}}_j = (\tilde{X}_{1j}, \tilde{X}_{2j}, \dots, \tilde{X}_{rj})^T$ and

$$\bar{\tilde{X}}_j = \frac{\sum_{i=1}^{r} q_i \tilde{X}_{ij}}{\sum_{i=1}^{r} q_i}.$$

The $T_{01,j}$s are independent and the exact distribution of $T_{01,j}$ is given in Robertson et al. (1988). It is a chi-bar squared distribution, that is, a mixture of chi-squared distributions, mixed over their degrees of freedom. Specifically, for all $t > 0$,

$$P(T_{01,j} \geq t) = \sum_{\ell=1}^{r} P(\ell, r, \mathbf{q}) P(\chi^2_{\ell-1} \geq t) \tag{11}$$

where $\chi^2_0 \equiv 0$ and $P(\ell, r, \mathbf{q})$ is the probability that the least squares projection of $\tilde{\mathbf{X}}_j$ with weights $\mathbf{q}$ onto $\mathcal{I}$ has exactly ℓ levels. Putting all this together leads to the following theorem that shows that the asymptotic null distribution of T_{01} is a chi-bar-squared distribution. Its weights on the various chi-squared tail probabilities are obtained by convoluting the r sequences of level probabilities corresponding to the $T_{01,j}$s.

Theorem 2 *If* $\mathbf{F} = (F_1, F_2, \dots, F_r)$ *satisfies* H_0, *for any* $t > 0$,

$$\lim_{n\to\infty} P(T_{01} \geq t) = \sum_{\ell=m-1}^{r(m-1)} P^*(\ell, r, \mathbf{q}) P(\chi^2_{\ell-(m-1)} \geq t) \tag{12}$$

where

$$P^*(\ell, r, \mathbf{q}) = \sum_{1\leq \ell_1, \ell_2, \dots, \ell_r, \ell_1+\ell_2+\dots+\ell_r=\ell} \prod_{j=1}^{m-1} P(\ell, r, \mathbf{q})$$

and $\chi^2_0 \equiv 0$.

Clearly the null limiting distribution of T_{01} depends only on the distribution of δ through the weights $P(\ell, r, \mathbf{q})$. These weights, also known as a level probabilities, are sums of products of normal orthant probabilities. In general, they do not exist in a closed form. However, when $q_1 = q_2 = \dots = q_r$, they do not depend on $\mathbf{q}$ in which case it is omitted. In addition, they satisfy in this case the following recurrence relations:

$$P(\ell, r) = \frac{1}{r}P(\ell - 1, r - 1) + \frac{r-1}{r}P(\ell, r - 1)$$

where $P(0, r - 1) = P(r, r - 1) = 0$. For more on this, see Robertson et al. (1988).

Remark 1 When $r = 2$, it is easy to show that

$$\hat{\theta}_{1j}^{(1)} = \hat{\theta}_{1j}^{(2)} \wedge \frac{\hat{w}_{1j}\hat{\theta}_{1j}^{(2)} + \hat{w}_{2j}\hat{\theta}_{2j}^{(2)}}{\hat{w}_{1j} + \hat{w}_{2j}} = \frac{\hat{F}_1(t_{j+1})}{\hat{F}_2(t_{j+1})} \wedge \frac{\hat{F}(t_{j+1})}{\hat{F}(t_{j+1})} \tag{13}$$

and

$$\hat{\theta}_{1j}^{(1)} = \hat{\theta}_{2j}^{(2)} \vee \frac{\hat{w}_{1j}\hat{\theta}_{1j}^{(2)} + \hat{w}_{2j}\hat{\theta}_{2j}^{(2)}}{\hat{w}_{1j} + \hat{w}_{2j}} = \frac{\hat{F}_2(t_{j+1})}{\hat{F}_2(t_{j+1})} \vee \frac{\hat{F}(t_{j+1})}{\hat{F}(t_{j+1})} \tag{14}$$

where $\hat{F} = \hat{F}_1 + \hat{F}_2$ and $\wedge$ and $\vee$ are used to denote the maximum and the minimum, respectively. In this case $P(1, 2, \mathbf{q}) = P(1, 2, \mathbf{q}) = 1/2$, $P(T_{01,j} \geq t) = \frac{1}{2}P(\chi_1^2 \geq t)$, $j = 1, 2, \dots, m - 1$, and

$$\lim_{n\to\infty} P(T_{01} \geq t) = \sum_{\ell=0}^{m-1} \binom{m-1}{\ell} 2^{-m+1} P(\chi_\ell^2 \geq t).$$

This implies that the LRT statistic is asymptotically distribution free over H_0. In passing we mention that, in general, the test procedures that involve inequality constraints do not result in asymptotically distribution free (similar) tests. As a result they tend to be conservative over much of the null hypothesis region.

Remark 2 The weight $P^*(\ell, r, \mathbf{q})$ depends on the unknown distribution of δ. It can be approximated by $P^*(\ell, r, \hat{\mathbf{q}}^{(0)})$ where $\hat{\mathbf{q}}^{(0)} = (d_{1+}/n, d_{2+}/n, \dots, d_{r+}/n)^T$ is the m.l.e. of $\mathbf{q}$ under H_0. The expression $\sum_{\ell=m-1}^{r(m-1)} P^*(\ell, r, \hat{\mathbf{q}}^{(0)}) P(\chi^2_{\ell-(m-1)} \geq t)$ converges as n goes to infinity to the right hand side of (12) and it does provide in general a good approximation for it. Numerical simulations can be used to approximate its weights. For more discussion on this, see Robertson et al. (1988).

Consider now testing H_0^* against $H_1 - H_0^*$. In terms of the new parameterization H_0^* reduces to $\theta_{1j} = \theta_{2j} = \dots = \theta_{rj}$, $j = 1, 2, \dots, m$. The m.l.e of $\boldsymbol{\theta}_j$ under H_0^*, denoted by $\tilde{\boldsymbol{\theta}}^{(0)} = (\tilde{\theta}_{1j}^{(0)}, \tilde{\theta}_{2j}^{(0)}, \dots, \tilde{\theta}_{rj}^{(0)})^T$, is given by

$$\tilde{\theta}_{ij}^{(0)} = \begin{cases} \hat{\theta}_{ij}^{(0)}, & j = 1, 2, \dots, m - 1, \\ \frac{1}{r}, & j = m, \end{cases}$$

and $\hat{\theta}_{ij}^{(0)}$ is defined before. This implies that the LRT rejects H_0^* for large values of

$$T_{01}^* = T_{01} - 2\sum_{\ell=1}^{r} d_{i+} \log\left(\frac{1/r}{\hat{\theta}_{im}^{(2)}}\right),$$

where T_{01} is also defined before. When H_0^* is true, (8) and an application of a Taylor expansion with a second order remainder of $\log\left(\hat{\theta}_{im}^{(2)}\right)$ around $1/r$ show that asymptotically, T_{01}^* is equivalent to:

$$\sum_{j=1}^{m-1}\sum_{i=1}^{r} \frac{n\hat{F}_i(t_{j+1})}{\hat{\theta}_{ij}^{(0)}(1-\hat{\theta}_{ij}^{(0)})}\left(\hat{\theta}_{ij}^{(1)} - \hat{\theta}_{ij}^{(0)}\right)^2 + \sum_{i=1}^{r} \frac{1}{\frac{1}{r}\left(1-\frac{1}{r}\right)}\left(\hat{\theta}_{im}^{(2)} - \frac{1}{r}\right)^2. \tag{15}$$

In addition, if $\{X_{im}, i = 1, 2, \dots, r\}$ is a random sample from $N\left(0, \frac{1}{r}\left(1-\frac{1}{r}\right)\right)$, then

$$\sum_{i=1}^{r} \frac{1}{(1/r)(1-1/r)}\left(\hat{\theta}_{im}^{(2)} - 1/r\right)^2 \xrightarrow{d} T_{01,m} \equiv \sum_{i=1}^{r} \frac{1}{(1/r)(1-1/r)}\left(X_{im} - 1/r\right)^2 \sim \chi_{r-1}^2$$

and $T_{01,m}$ is independent of the $T_{01,j}, 1 \le j \le m-1$. Since, under $H_0^*, q_1 = q_2 = \dots = q_r = 1/r$, the weights in (11) do not depend on $\mathbf{q}$ and hence

$$P(T_{01,j} \ge t) = \sum_{\ell=1}^{r} P(\ell, r)P(\chi_{\ell-1}^2 \ge t).$$

This leads to the following theorem.

Theorem 3 *If* $\mathbf{F} = (F_1, F_2, \dots, F_r)$ *satisfies* H_0^*, *for any* $t > 0$,

$$\lim_{n\to\infty} P(\tilde{T}_{01} \ge t) = \sum_{\ell=m-1}^{r(m-1)} \tilde{P}(\ell, r)P(\chi_{\ell+r-m}^2 \ge t) \tag{16}$$

where

$$\tilde{P}(\ell, r) = \sum_{1\le\ell_1,\ell_2,\dots,\ell_r,\ell_1+\ell_2+\dots+\ell_r=\ell}\ \prod_{j=1}^{m-1} P(\ell, r)$$

and $\chi_0^2 \equiv 0$.

3.2 Testing H_1 against $H_2 - H_1$

Next, we consider the problem of testing H_1 as a null hypothesis against $H_2 - H_1$ where H_2 imposes no restrictions on the parameters. The likelihood ratio test rejects H_1 in this case for large values of

$$\Lambda_{01} = \frac{\prod_{j=1}^{m} L_j(\hat{\theta}_j^{(1)})}{\prod_{j=1}^{m} L_j(\hat{\theta}_j^{(2)})} = \prod_{j=1}^{m-1}\prod_{i=1}^{r}\left[\frac{\hat{\theta}_{ij}^{(1)}}{\hat{\theta}_{ij}^{(2)}}\right]^{\sum_{\ell=1}^{j} d_{ij}}\left[\frac{1-\hat{\theta}_{ij}^{(1)}}{1-\hat{\theta}_{ij}^{(2)}}\right]^{d_{i,j+1}}$$

because the m.l.e. of $\boldsymbol{\theta}_m$ is the same under both H_1 and H_2. Equivalently, it rejects H_0 for large values of of the statistic

$$\begin{aligned} T_{01} =& -2\log(\Lambda_{12}) \\ =& 2\sum_{j=1}^{m}\sum_{i=1}^{r}\left[\left(\sum_{\ell=1}^{j} d_{ij}\right)\log\left\{\frac{\hat{\theta}_{ij}^{(2)}}{\hat{\theta}_{ij}^{(1)}}\right\} + d_{ij}\log\left\{\frac{1-\hat{\theta}_{ij}^{(2)}}{1-\hat{\theta}_{ij}^{(1)}}\right\}\right]. \end{aligned}$$

Define on $D = \{1, 2, \dots, r\}$ the quasi-order $\preceq_{\boldsymbol{\theta}_i}$ which requires that $j_1 \preceq_{\boldsymbol{\theta}_i} j_2$ when $j_1 \le j_2$ and $\theta_{ij_1} = \theta_{ij_2}$. Define also $\mathcal{I}_{\boldsymbol{\theta}_i}$ to be the cone of isotonic functions on D with respect to the quasi-order $\preceq_{\boldsymbol{\theta}_i}$. In what follows we assume the X_{ij} has now $N(0, \theta_{ij}(1-\theta_{ij})/F_i(t_{j+1}))$ and define $P_{\boldsymbol{\theta}_i}(\ell, r, \mathbf{w}_j)$ to be the probability that $E\left[\mathbf{X}_j|\mathcal{I}_{\boldsymbol{\theta}_i}\right]$, the least squares projection of $\mathbf{X}_j$ onto $\mathcal{I}_{\boldsymbol{\theta}_i}$ with weights $\mathbf{w}_j$ has exactly ℓ levels.

For $\mathbf{F} = (F_1, F_2, \dots, F_r)$, let $P_{\mathbf{F}}(A)$ denote the probability of the event A when $F_1, F_2, \dots, F_r$ are the true population CIFs. We have the following distributional result.

Theorem 4 *If* $\mathbf{F} = (F_1, F_2, \dots, F_r)$ *satisfies* H_1, *for any* $t > 0$,

$$\lim_{n\to\infty} P_{\mathbf{F}}(T_{12} \ge t) = \sum_{\ell=m-1}^{r(m-1)} P_{\boldsymbol{\theta}(\mathbf{F})}(\ell, r, \mathbf{w}(\mathbf{F}))P(\chi^2_{r(m-1)-\ell} \ge t) \tag{17}$$

where $\theta(\mathbf{F})$ *and* $\mathbf{w}(\mathbf{F})$ *are the values of* $(\boldsymbol{\theta}_1^T, \boldsymbol{\theta}_2^T, \dots, \boldsymbol{\theta}_{m-1}^T)^T$ *and* $(\mathbf{w}_1^T, \mathbf{w}_2^T, \dots, \mathbf{w}_{m-1}^T)^T$ *corresponding to* $\mathbf{F}$,

$$P_{\boldsymbol{\theta}(\mathbf{F})}(\ell, r, \mathbf{w}) = \sum_{1\le \ell_1, \ell_2, \dots, \ell_r, \ell_1+\ell_2+\dots+\ell_r=\ell}\ \prod_{j=1}^{m-1} P_{\boldsymbol{\theta}_j(\mathbf{F})}(\ell, r, \mathbf{w}_j(\mathbf{F}))$$

and $\chi_0^2 \equiv 0$. *In addition*,

$$\lim_{n\to\infty} P_{\mathbf{F}}(T_{12} \ge t) \le \sum_{\ell=m-1}^{r(m-1)} P^*(\ell, r, \mathbf{F})P(\chi^2_{r(m-1)-\ell} \ge t) \tag{18}$$

where

$$P^*(\ell, r, \mathbf{F}) = \sum_{1\le \ell_1, \ell_2, \dots, \ell_r, \ell_1+\ell_2+\dots+\ell_r=\ell}\ \prod_{j=1}^{m-1} P(\ell, r, \mathbf{w}_j(\mathbf{F}))$$

and $P(\ell, r, \mathbf{w}_j(\mathbf{F}))$ *is the probability that the least squares projection of* $(\tilde{X}_{1j}, \tilde{X}_{2j}, \ldots, \tilde{X}_{rj})^T$, *a zero mean normal vector with dispersion* diag $(1/F_1(t_{j+1}), 1/F_2(t_{j+1}), \ldots, 1/F_r(t_{j+1}))$, *onto* $\mathcal{I}$ *with weights* $(F_1(t_{j+1}), F_2(t_{j+1}), \ldots, F_r(t_{j+1}))^T$ *has* ℓ *levels.*

Proof Expanding $\log(\hat{\theta}_{ij}^{(1)})$ about $\hat{\theta}_{ij}^{(2)}$ and $\log(1-\hat{\theta}_{ij}^{(1)})$ about $1-\hat{\theta}_{ij}^{(2)}$ using a Taylor expansion with a second remainder, we get, using the properties of the isotonic regression (see (Robertson et al. 1988)),

$$T_{12} = \sum_{j=1}^{m-1}\sum_{i=1}^{r}\left\{\frac{\sum_{\ell=1}^{j} d_{i\ell}}{\beta_{1,ij}^2} + \frac{d_{i,j+1}}{\beta_{2,ij}^2}\right\}\left(\hat{\theta}_{ij}^{(1)} - \hat{\theta}_{ij}^{(2)}\right)^2$$

where

$$\max\left\{|\beta_{1,ij} - \hat{\theta}_{ij}^{(2)}|, |\beta_{2,ij} - (1-\hat{\theta}_{ij}^{(2)})|\right\} \le |\hat{\theta}_{ij}^{(1)} - \hat{\theta}_{ij}^{(2)}|. \tag{19}$$

When H_1 is true, the right hand side of (19) goes to zero almost surely. This implies that T_{12} is asymptotically equivalent to

$$\sum_{j=1}^{m-1}\sum_{i=1}^{r}\frac{n\hat{F}_i(t_{j+1})}{\hat{\theta}_{ij}^{(2)}(1-\hat{\theta}_{ij}^{(2)})}\left(\hat{\theta}_{ij}^{(1)} - \hat{\theta}_{ij}^{(2)}\right)^2. \tag{20}$$

Arguing as in the proof of Theorem 5.2.1 in Robertson et al. (1988), we find that (20) is equal, for large n with probability one, to

$$\sum_{j=1}^{m-1}\sum_{i=1}^{r}\frac{\hat{F}_i(t_{j+1})}{\hat{\theta}_{ij}^{(2)}(1-\hat{\theta}_{ij}^{(2)})}\left(E_{\hat{\mathbf{w}}_j}\left[\sqrt{n}(\hat{\boldsymbol{\theta}}_j^{(1)} - \boldsymbol{\theta}_j)|\mathcal{I}_{\boldsymbol{\theta}_i}\right]_i - \sqrt{n}(\hat{\theta}_{ij}^{(2)} - \theta_{ij})\right)^2$$

and converges in distribution to

$$\sum_{j=1}^{m-1}\sum_{i=1}^{r}\frac{F_i(t_{j+1})}{\theta_{ij}(1-\theta_{ij})}\left(E_{\mathbf{w}_j}\left[\mathbf{X}_j|\mathcal{I}_{\boldsymbol{\theta}_i}\right]_i - X_{ij}\right)^2 \equiv \sum_{j=1}^{m-1} T_{12,j}$$

where, for $1 \le j \le m-1$,

$$T_{12,j} = \sum_{i=1}^{r}\frac{F_i(t_{j+1})}{\theta_{ij}(1-\theta_{ij})}\left(E_{\mathbf{w}_j}\left[\mathbf{X}_j|\mathcal{I}_{\boldsymbol{\theta}_i}\right]_i - X_{ij})\right)^2.$$

Now Lemma A on page 321 in Robertson et al. (1988) implies that

$$T_{12,j} = \sum_{i=1}^{r} F_i(t_{j+1})\left(E_{\mathbf{w}_j}\left[\tilde{\mathbf{X}}_j|\mathcal{I}_{\boldsymbol{\theta}_i}\right]_i - \tilde{X}_{ij})\right)^2, j = 1, 2, \ldots, m-1,$$

where $\tilde{\mathbf{X}} = (\tilde{X}_{1j}, \tilde{X}_{2j}, \ldots, X_{rj})^T$ and $\tilde{X}_{ij}$s are independent zero mean normal variates with Var $(\tilde{X}_{ij}) = 1/F_i(t_{j+1})$. Since $T_{12,j}$s are independent and

$$P_{\mathbf{F}}(T_{12,j} \geq t) = \sum_{i=1}^{r} P_{\theta_i(\mathbf{F})}(\ell, r, \mathbf{w}_j(\mathbf{F}))P(\chi^2_{\ell-1} \geq t)$$

for any $t > 0$ (see (Robertson et al., 1988)), (17) follows by a direct computation of $P_{\mathbf{F}}\left(\sum_{j=1}^{m-1} T_{12,j} > t\right)$.

To show (18), notice that $\mathcal{I}$ is a subset of $\mathcal{I}_{\theta_i}$ for all i.. Hence

$$\tilde{T}_{12,j} = \sum_{i=1}^{r} F_i(t_{j+1})\left(E_{\mathbf{w}_j}\left[\tilde{\mathbf{X}}_j|\mathcal{I}\right]_i - \tilde{X}_{ij})\right)^2 \leq \sum_{i=1}^{r} F_i(t_{j+1})\left(E_{\mathbf{w}_j}\left[\tilde{\mathbf{X}}_j|\mathcal{I}_{\theta_i}\right]_i - \tilde{X}_{ij})\right)^2 \equiv T_{12}.$$

Since

$$P_{\mathbf{F}}(\tilde{T}_{12,j} \geq t) = \sum_{i=1}^{r} P(\ell, r, \mathbf{w}_j(\mathbf{F}))P(\chi^2_{\ell-1} \geq t)$$

and the $\tilde{T}_{12,j}$ are independent, (18) follows also by computing $P_{\mathbf{F}}\left(\sum_{j=1}^{m-1} \tilde{T}_{12,j} > t\right)$. □

Corollary 1 :*When* $r = 2$, *if* $\mathbf{F}$ *satisfies* H_1, *then*

$$\lim_{n\to\infty} P_{\mathbf{F}}(T_{12} > t) = \sum_{\ell=0}^{M} \binom{M}{\ell} \frac{1}{2^M} P(\chi^2_\ell > t)$$

where $\chi^2_0 \equiv 0$ *and* $M =$ card $\{j, \theta_{1j} = \theta_{2j}\}$. *Moreover*

$$\lim_{n\to\infty} P_{\mathbf{F}}(T_{12} > t) \leq \sum_{\ell=0}^{m-1} \binom{m-1}{\ell} \frac{1}{2^{m-1}} P(\chi^2_\ell > t). \tag{21}$$

Proof Plugging (13) and (14) into (20) shows that T_{12} is equivalent to

$$\sum_{j=1}^{m-1} \tilde{w}_j \left[\sqrt{n}(\hat{\theta}^{(2)}_{2j} - \hat{\theta}^{(2)}_{1j}) \vee 0\right]^2 \tag{22}$$

where

$$\tilde{w}_j = \frac{\hat{F}_1(t_{j+1})\hat{F}_2(t_{j+1})}{\hat{F}(t_{j+1})}\left[\sum_{\ell=1}^{2} \frac{\hat{F}_2(t_{j+1})}{\hat{\theta}^{(2)}_{1j}(1-\hat{\theta}^{(2)}_{1j})} + \frac{\hat{F}_1(t_{j+1})}{\hat{\theta}^{(2)}_{2j}(1-\hat{\theta}^{(2)}_{2j})}\right], j = 1, 2, \ldots, m.$$

Using (9) and the delta method, it is easy to show that , when $\theta_{1j} = \theta_{2j} \equiv \theta_j$, then

$$\sqrt{n}[\hat{\theta}^{(2)}_{2j} - \hat{\theta}^{(2)}_{1j}] \xrightarrow{d} Y_j$$

where Y_j has a normal distribution with mean zero and variance $\dfrac{\theta_j(1-\theta_j)F(t_{j+1})}{F_1(t_{j+1})F_2(t_{j+1})}$. Otherwise it converges to $-\infty$. In addition the Y_js are independent. This implies that (22) converges in distribution to

$$\sum_{\{j;\theta_{1j}=\theta_{2j}\}} \frac{F_1(t_{j+1})F_2(t_{j+1})}{\theta_j(1-\theta_j)F(t_{j+1})}\left[Y_j \vee 0\right]^2 = \sum_{\{j;\theta_{1j}=\theta_{2j}\}} \left[Z_j \vee 0\right]^2$$

where $Z_j, j = 1, 2, \dots, m-1$, are independent standard normals. The first conclusion follows immediately by computing $P\left(\sum_{\{j;\theta_{1j}=\theta_{2j}\}} \left[Z_j \vee 0\right]^2 > t\right)$. To show (21), notice that, since $M \le m-1$,

$$\sum_{\{j;\theta_{1j}=\theta_{2j}\}} \left[Z_j \vee 0\right]^2 \le \sum_{i=1}^{m-1} \left[Z_j \vee 0\right]^2.$$

The result now follows since

$$P\left(\sum_{i=1}^{m-1} \left[Z_j \vee 0\right]^2 \ge t\right) = \sum_{\ell=0}^{m-1} \binom{m-1}{\ell} \frac{1}{2^{m-1}} P(\chi^2_\ell > t).$$

□

This corollary provides an upper bound on the Type I error and give a method for investigating the behavior of $\lim_{n\to\infty} P_{\mathbf{F}}(T_{12} \ge t)$ for various $\mathbf{F}$ that satisfy H_1. In shows in particular that, for $\mathbf{F}$ that satisfies $\theta_{1j} < \theta_{2j}$, for all j, $\lim_{n\to\infty} P_{\mathbf{F}}(T_{12} \ge t) = 0$.

4 Examples

In this section, we discuss two numerical examples that are designed to illustrate the theory developed in Sects. 2 and 3.

4.1 Example 1

For our first illustration, we consider the mortality data on RFM strain male mice as reported in Hoel (1972). We consider two risks with the second risk being cancer and take the first risk a combination of all other risks. The failure times are grouped into $m = 6$ categories. Thus, $r = 2$ since we have two risks and $m = 6$ time periods. The data as well as the m.l.e.s of the p_{ij}s under the different hypotheses are give in Table 1.

Since $r = 2$, the null limiting distribution of the LRT statistic, T_{01}, for testing H_0 against $H_1 - H_0$ is given in Remark 1. Its value is 1.619 corresponding-to a p-value of 0.63. To test H_1 against $H_2 - H_1$, the value of $T_{12} = 6.857$ corresponding to a

Table 1 Estimated probabilities under H_0, H_1 and H_2

No.	Interval	d_{1j}	d_{2j}	$\hat{p}_{1j}^{(2)}$	$\hat{p}_{2j}^{(2)}$	$\hat{p}_{1j}^{(0)}$	$\hat{p}_{2j}^{(0)}$	$\hat{p}_{1j}^{(1)}$	$\hat{p}_{2j}^{(1)}$
1	(0, 350]	15	18	0.1515	0.1818	0.1313	0.2020	0.1211	0.2158
2	(0, 350]	6	7	0.0606	0.0707	0.0517	0.0796	0.0485	0.0839
3	(0, 350]	6	4	0.0606	0.0404	0.0398	0.0612	0.0485	0.0480
4	(0, 350]	8	18	0.0808	0.1818	0.1035	0.1592	0.1012	0.1614
5	(0, 350]	2	12	0.0202	0.1212	0.0557	0.0857	0.0545	0.0869
6	(0, 350]	2	1	0.0202	0.0101	0.0119	0.0184	0.0202	0.0101

p-value of 0.0638 based on (21). Evidently, unless H_1 holds, one should not test H_0 against $H_1 - H_0$. We include here for illustration purposes.

4.2 Example 2

In our second example, we consider some data from a randomized study conducted by the Adult AIDS Clinical Trials Group (AACTG) to evaluate two combination antiretroviral treatments in terms of their ability to suppress HIV viral load. The failure time,T, in this case was defined as the time from randomization until plasma HIV levels rose above 1000 copies/ml. At failure a measure of acquired mutational distance during the trial was obtained. This distance is a measure of the accumulated HIV genetic resistance due to treatment exposure and is only obtained when a subject fails. Gilbert et al. (2004) normalize this distance so that it lies in the interval [0, 1]. For our purposes we discretize the normalized distance measure, call which we call V, and consider $r = 3$ groups. A subject is classified as belonging to group 1 if $V \in (0, 1/3]$, to group 2 if $V \in (1/3, 2/3]$ and to group 3 if $V \in (2/3, 1]$. We also consider $m = 3$ failure time intervals and $j = 1$ if $T \in (0, 5]$,$j = 2$ if $T \in (5, 20]$ and $j = 3$ if $T \in$ (20;50].

The data and the m.l.e.s of the p_{ij}s are given in the following two tables (Tables 2, 3).

Hence, we have $r = 3$ and $m = 3$, In this case the data satisfies H_1 since the m.l.e.s under H_1 and H_2 are equal. The value of the test statistic for testing H_0 against $H_1 - H_0$ is 0.8959. To compute the p-value in this case, we use simulations to estimate the weights in (12) after replacing $\mathbf{q}$ by $\hat{\mathbf{q}}^{(0)}$, its m.l.e. under H_0. The estimated weights are (0.1241, 0.3497, 0.3524, 0.1507, 0.0231) and the estimated p-value is 0.51.

Table 2 Estimated probabilities under H_2

Interval	d_{1j}	d_{2j}	d_{3j}	$\hat{p}_{1j}^{(2)}$	$\hat{p}_{2j}^{(2)}$	$\hat{p}_{3j}^{(2)}$
(0, 5]	5	7	7	0.1111	0.1556	0.1556
(5, 20]	6	5	4	0.1333	0.1111	0.0889
(20, 50]	4	4	3	0.0889	0.0889	0.0667

Table 3 Estimated probabilities under H_0 and H_1

Interval	d_{1j}	d_{2j}	d_{3j}	$\hat{p}_{1j}^{(0)}$	$\hat{p}_{2j}^{(0)}$	$\hat{p}_{3j}^{(0)}$	$\hat{p}_{1j}^{(1)}$	$\hat{p}_{2j}^{(1)}$	$\hat{p}_{3j}^{(1)}$
(0, 5]	5	7	7	0.1407	0.1501	0.1314	0.1111	0.1556	0.1556
(5, 20]	6	5	4	0.1111	0.1185	0.1037	0.1333	0.1111	0.0889
(20, 50]	4	4	3	0.0815	0.0869	0.0667	0.0889	0.0889	0.0667

Acknowledgements The author is grateful to the associate and a referee for helpful comments that led to a much improved paper. This work was supported by The City University of New York through a PSC-CUNY grant.

References

Al-Kandari, N., El Barmi, H. (2022). Restricted estimation of the cumulative incidence functions of two competing risks. *Journal of Statistical Planning and Inference, 217,* 122–140.

Aly, E. A. A., Kochar, S. C., McKeague, I. W. (1994). Some tests for comparing cumulative incidence functions and cause-specific hazard rates. *Journal of the American Statistical Association, 89,* 994–999.

Dauxois, J. Y., Kirmani, S. (2003). On testing for a survival ratio under random right censoring. *Journal of Nonparametric Statistics, 17,* 949–955.

El Barmi, H., Kochar, S. (2002). Inference for subsurvival functions under order restrictions. *Journal of the Indian Statistical Association, 40,* 85–103.

El Barmi, H., Mukerjee, H. (2006). *Restricted estimation of cumulative incidence functions corresponding to competing risks*. *Optimality, Second Lehmann Symposium, IMS-LNMS*, 241–252.

El Barmi, H., Kochar, S., Mukerjee, H., Samaniego, F. (2004). Estimation of cumulative incidence functions in competing risks studies under an order restriction. *Journal of Statistical Planning and Inference, 118,* 145–165.

El Barmi, H., Kochar, S., Tsimikas, J. (2006). Likelihood ratio test for and against ordering of cumulative incidence functions in multiple competing risks and discrete mark variable models. *Journal of Statistical Planning and Inference, 136,* 1588–1607.

Gilbert, P. B., Mckeague, I. W., Sun, Y. (2004). Tests for comparing mark-specific hazards and cumulative incidence functions. *Lifetime Data Analysis, 10,* 5–28.

Hoel, D. G. (1972). A representation of mortality data by competing risks. *Biometrics, 28,* 475–478.

Robertson, T., Wright, F. T., Dykstra, R. L. (1988). *Order restricted inference*. New York: Wiley.

Shaked, M., Shanthikumar, G. J. (2006). *Stochastic orders*. New York: Springer.