Inference for Sub-survival Functions Under Order Restrictions

Hammou El Barmi[1] and Subhash Kochar [2]

[1]Department of Statistics, Kansas State University, Manhattan, KS 66506

[2]Indian Statistical Institute, 7, S.J.S. Sansanwal Marg, New Delhi-110016, India

ABSTRACT

We consider the competing risks problem with two risks and when the data are grouped or discrete. We firstly obtain nonparametric maximum likelihood estimates of the sub-survival functions corresponding to the two risks under the restriction that they are uniformly ordered and then use them to derive the likelihood ratio statistic for testing the null hypothesis of equality of the two sub-survival functions against ordered alternatives. The asymptotic null distribution of the test statistic is seen to be of the chi-bar square ($\bar{\chi}^2$) type. A simulation study has been performed to compare the power of the new test with an existing one.

AMS 1980 Subject Classification: Primary: 62G10, 62G05; Secondary: 62E20, 62N06

Some key words and phrases: Isotonic regression, cause specific hazard rates, competing risks, two-sample problem.

1 Introduction

In the standard competing risks model, a unit or subject is exposed to several risks at the same time, but the actual failure (or death) is attributed to exactly one cause. In this paper we assume that there are only two risks. Let the notional (or latent) lifetimes of the unit under these two risks be denoted by X and Y. In general, X and Y are dependent. We only observe (T, δ), where $T = \min(X, Y)$ is called the *time of failure* and $\delta = 2 - I_{(X \leq Y)}$ is the *cause of failure*. Here I_A is the indicator function of the event A. We assume that $pr(X = Y) = 0$. Thus, the observed data are in the form of (T, δ) for each item.

On the basis of the competing risks data it is often of interest to know whether the two risks are equal or one risk is greater than the other. Such comparisons can be made in terms of *sub-survival functions*,

$$S_i(t) = pr[T \geq t, \delta = i],$$

or in terms of *cumulative incidence* (sub- distribution) functions,

$$F_i(t) = pr[T \leq t, \delta = i],$$

corresponding to each cause i. Note that $S_1(t) + S_2(t) = S_T(t)$ and $F_1(t) + F_2(t) = F_T(t)$ where S_T and F_T are the survival function and the distribution function of T, respectively.

Another approach would be to compare their *cause specific hazard rates* (CSHR). The cause specific hazard rate corresponding to the i^{th} cause is defined by

$$h_i(t) = \lim_{\Delta t \to 0} \frac{1}{\Delta t} \, pr[t \leq T < t + \Delta t, \ \delta = i | T \geq t]$$

$i = 1, 2$. If T is discrete, the ith cause specific hazard rate is given by $pr(T = t, \delta = i | T \geq t)$. In either case the overall hazard rate for time to failure is given by $h(t) = h_1(t) + h_2(t)$. In models where the various causes of failure are independent, $h_i(t)$ reduces to the (ordinary) hazard rate corresponding to the marginal distribution of failure from the i^{th} cause. In the continuous case the sub-survival functions and the cumulative incidence functions can be expressed in terms of the cause specific hazard rates by the relations,

$$S_i(t) = \int_t^\infty h_i(u) S_T(u)\, du, \quad F_i(t) = \int_0^t h_i(u) S_T(u)\, du, \tag{1.1}$$

for $i = 1, 2$. Similar relations can be established for the discrete case.

In this paper we consider the problem of testing the null hypothesis,

$$H_0 : S_1(t) = S_2(t) \quad \text{for } t \geq 0, \tag{1.2}$$

against the alternative,

$$H_1 : S_1(t) \leq S_2(t), \quad \text{for } t \geq 0, \tag{1.3}$$

and with strict inequality for some t.

Note that (1.3) can be equivalently expressed as

$$pr[\delta = 1 | T \geq t] \leq pr[\delta = 2 | T \geq t] \quad \text{for } t \geq 0,$$

and with strict inequality for some t. In this form it has the interpretation that given that a unit has survived up to time t, the conditional probability of its failing in the future from cause 2 is *uniformly* greater than that from cause 1.

Also note that H_0 is equivalent to $H_0' : h_1(t) = h_2(t)$ for all t as well as to $H_0'' : F_1(t) = F_2(t)$ for all t. H_1 is implied by the more stringent alternative

$$H_A : h_1(t) \leq h_2(t) \quad \text{for } t \geq 0, \tag{1.4}$$

with strict inequality for some t.

Caution should be exercised in the interpretation of sub- survival functions. Unlike in the case of ordinary two-sample problem, H_1 may not imply that the failures from cause 1 occur earlier than those from cause 2.

Several tests are available in the literature for testing the equality of competing risks and they have been referenced in Aly, Kochar & McKeague (1994) and in the review paper by Kochar (1995). In the case of continuous random variables, Deshpandé (1990), Aly, Kochar and McKeague (1994), Deshpandé and Karia (1995), and Sun and Tiwari (1998) among others considered the problem of testing the null hypothesis H_0 against the alternatives H_A and

$$H_A^{'} : F_1(t) \leq F_2(t),\ t \geq 0\,, \tag{1.5}$$

and with strict inequality for some t. The only test designed specifically for the problem of testing H_0 against H_1, that we are aware of, is by Carriere and Kochar (2000). Their test is distribution-free and is suitable when the random variables are of continuous type.

Note that $H_A^{'}$ is also implied by H_A, but, in general, H_1 and $H_A^{'}$ are not equivalent. It is plausible that in some cases the cumulative incidence functions cross each other but their sub-survival functions are ordered and vice versa.

In many practical problems, the data collected is in the form of groups or intervals. For this kind of data, Dykstra, Kochar and Robertson (1995a) obtained the nonparametric maximum likelihood estimates (MLEs) of the cause specific hazard rates under the alternative that they are ordered. They used these estimators to derive the likelihood ratio statistic for testing the null hypothesis of the equality of the cause specific hazard rates against ordered alternative of the type H_A.

In this paper we assume a discrete time framework. In Section 2 we obtain maximum likelihood estimators of the sub-survival functions S_1 and S_2 under H_0 as well as under H_1. In Section 3 we derive the likelihood ratio test for testing H_0 versus $H_1 - H_0$ and obtain its asymptotic null distribution. We also consider testing H_1 as a null hypotheses versus $H_2 - H_1$ where H_2 imposes no constraints on S_1 and S_2 and obtain the least favorable configuration corresponding to this case. Finally, in the last section we perform some simulation studies to compare the power of our test with that of Dykstra, Kochar and Robertson (1995a) for alternatives belonging to H_1.

2 Maximum Likelihood Estimation

Suppose that we have n individuals exposed to two risks and assume the times and causes of failure represent a random sample on (T, δ). We make no assumptions about the independence of notional lifetimes associated with the two risks.

In this section we obtain nonparametric maximum likelihood estimates of the sub-survival functions S_1 and S_2 under the restriction that $S_1 \leq S_2$. Peterson (1977) has derived the unrestricted generalized nonparametric m.l.e.'s of the sub-survival function S_1 and S_2. It is clear from his discussion or otherwise, that the generalized NPMLE's of the sub-survival functions put their weights on the set of observations.

We assume that failures occur on the times $t_1 < t_2 < \ldots < t_m$ ($t_0 = 0$ and $t_{m+1} = \infty$). For $i = 1, 2$ and $j = 1, \ldots, m$, let $p_{i,j}$ be the probability of failure from cause i at time t_j, n be the total number of items on test (sample size) and $d_{i,j}$ be the number of failures from cause i at time t_j. Then

$$\begin{aligned} S_i(t_j) &= pr(T \geq t_j, \delta = i) \\ &= \sum_{l=j}^{m} p_{i,l}, \end{aligned} \tag{2.1}$$

$i = 1, 2; j = 1, 2, \ldots, m$. The likelihood function is

$$\mathcal{L} = \left(\prod_{j=1}^{m} p_{1,j}^{d_{1,j}} \right) \left(\prod_{j=1}^{m} p_{2,j}^{d_{2,j}} \right).$$

Let $\mathbf{p}$ be the 2m-dimensional vector

$$\mathbf{p} = (p_{2,m}, \ldots, p_{2,1}, p_{1,1}, \ldots, p_{1,m})$$

and let

$$d_i = \begin{cases} d_{2,m-i+1}, & i = 1, 2, \ldots, m \\ d_{1,i-m}, & i = m+1, m+2, \ldots, 2m. \end{cases}$$

Now maximizing $\mathcal{L}$ is equivalent to maximizing L^2 which can be written as

$$\mathcal{L}^2 = \left(\prod_{j=1}^{2m} p_j^{d_j} \right) \left(\prod_{j=1}^{2m} p_{2m-j+1}^{d_{2m-j+1}} \right) \tag{2.2}$$

where

$$p_i = \begin{cases} p_{2,m-i+1}, & i = 1, 2, \ldots, m \\ p_{1,i-m}, & i = m+1, m+2, \ldots, 2m. \end{cases}$$

Formulated in this way, the problem reduces to the one considered by Dykstra, Kochar and Robertson (1995b) where they obtain nonparametric maximum likelihood estimators of a distribution function under the constraint that the distribution is symmetric about the origin as well as under the constraint that it is positively biased.

The unrestricted m.l.e. of $\mathbf{p}$ is $\hat{p}_i = \dfrac{d_i}{n}, i = 1, 2, \ldots, 2m$, and the m.l.e. of $\mathbf{p}$ under the restriction $S_1 = S_2$ is $\hat{p}_i^{(0)} = \hat{p}_{2m-i+1}^{(0)} = \dfrac{d_i + d_{2m-i+1}}{2n}, i = 1, 2, \ldots, m.$

Now maximizing $\mathcal{L}$ subject to the constraint $S_1 \leq S_2$ is equivalent to maximizing (2.2) subject to the constraint

$$\mathbf{p} \overset{st}{\leq} \mathbf{p}' \tag{2.3}$$

where "$\overset{st}{\leq}$" denotes stochastic ordering. Note here that $\mathbf{p}'$ denotes the reversed $2m$ dimensional vector $(p_{1,m}, \ldots, p_{1,1}, p_{2,1}, \ldots, p_{2,m})$. This is essentially the two-sample problem of estimating $\mathbf{p}$ and $\mathbf{p}'$ under the stochastic ordering constraint $\mathbf{p} \overset{st}{\leq} \mathbf{p}'$. The two-sample problem has been studied by Brunk et. al. (1966) and Barlow and Brunk (1972) and its solution can be used here. Throughout the rest of the paper, we assume multiplication and division of vectors are done coordinate-wise.

Theorem **2.1** *If $p_i > 0$ for $i = 1, 2, \ldots, 2m$, the m.l.e. of $\mathbf{p}$ subject to the constraint (2.3) is given by*

$$\hat{\mathbf{p}}^{(1)} = \hat{\mathbf{p}} E_{\hat{\mathbf{p}}} \left(\frac{\hat{\mathbf{p}} + \hat{\mathbf{p}}'}{2\hat{\mathbf{p}}} | \mathcal{A} \right) \tag{2.4}$$

where $E_{\mathbf{w}}(\mathbf{x}|\mathcal{A})$ denotes the least squares projection with weights $\mathbf{w}$ of the vector $\mathbf{x}$ onto the cone $\mathcal{A} = \{\mathbf{x}, \quad x_1 \geq x_2 \ldots \geq x_{2m}\}$, of nonincreasing vectors.

Proof: The proof is similar to that of Theorem 2.1 in Dykstra, Kochar and Robertson (1995b) and is omitted.

There are several algorithms available in the literature for computing $E_{\mathbf{w}}(\mathbf{x}|\mathcal{A})$. The easiest to implement is the pool adjacent violators algorithm (PAVA) and is discussed in Section 1.2 of Robertson, Wright and Dykstra (1988).

Using (2.4) and (2.1) we can obtain the MLE's of the sub- survival functions under the restriction that $S_1 \leq S_2$.

3 Hypothesis Testing

We now consider the problem of testing the null hypothesis H_0 against the alternative H_1. In our asymptotic theory, the number, m, of support points for T is fixed and the sample size n increases to ∞.

Let T_{01} be the log-likelihood ratio test statistic for this problem. We show that the limiting distribution of T_{01} is of chi-bar square (a mixture of independent chi-squares) type and provide the expression for the associated weights. In order to derive the limiting distribution of the test statistic, we make use of the following lemma.

LEMMA 3.1 *Let* $\mathbf{x} = (x_1, x_2, \ldots, x_{2m})$ *be such that* $x_i = -x_{2m-i+1}, \forall i$ *and let* $\mathbf{w} = (w_1, w_2, \ldots, w_{2m})$ *be such that* $w_i = w_{2m-i+1}$. *If* $\mathcal{A} = \{(u_1, u_2, \ldots, u_{2m})^T,$ $u_1 \geq u_2 \geq \cdots \geq u_{2m}\}$ *and* $\mathcal{I} = \{(u_1, u_2, \ldots, u_m), \quad 0 \geq u_1 \geq u_2 \geq \ldots \geq u_m\}$ *then*

$$E_{\mathbf{w}}(\mathbf{x}|\mathcal{A})_i = \begin{cases} E_{\mathbf{w}_r}(\mathbf{x}_r|\mathcal{I})_i & i = m+1, m+2, \ldots, 2m \\ -E_{\mathbf{w}_r}(\mathbf{x}_r|\mathcal{I})_i & i = 1, 2, \ldots, m \end{cases}$$

where $\mathbf{x}_r(\mathbf{w}_r)$ *is the restriction of* $\mathbf{x}(\mathbf{w})$ *to* $\{m+1, m+2, \ldots, 2m\}$.

The proof of this lemma follows immediately by verifying that $E_{\mathbf{w}}(\mathbf{x}|\mathcal{A})$ as given above satisfies the following sufficient conditions (see Theorem 1.3.2 of Robertson, Wright and Dykstra (1988)),

$$\sum_{i=1}^{2m} (x_i - E_{\mathbf{w}}(\mathbf{x}|\mathcal{A})_i) E_{\mathbf{w}}(\mathbf{x}|\mathcal{A})_i w_i = 0$$

and

$$\sum_{i=1}^{2m}(x_i - E_{\mathbf{w}}(\mathbf{x}|\mathcal{A})_i)y_i w_i \leq 0$$

for all $\mathbf{y} \in \mathcal{A}$.

Next we give a key distributional result. Expanding $\ln \hat{p}_i^{(0)}$ and $\ln \hat{p}_i^{(1)}$ about $\hat{p}_i$ with a second degree remainder term and using the fact that $\sum_{i=1}^{2m}\hat{p}_i^{(1)} = \sum_{i=1}^{2m}\hat{p}_i^{(0)} = 1$, it follows that

$$\begin{aligned} T_{01} &= 2n\left\{\sum_{i=1}^{2m}\hat{p}_i \ln \hat{p}_i^{(1)} - \sum_{i=1}^{2m}\hat{p}_i \ln \hat{p}_i^{(0)}\right\} \\ &= n\left\{\sum_{i=1}^{2m}\frac{\hat{p}_i}{\beta_i^2}\left(\hat{p}_i - \hat{p}_i^{(0)}\right)^2 - \sum_{i=1}^{2m}\frac{\hat{p}_i}{\alpha_i^2}\left(\hat{p}_i E_{\hat{\mathbf{p}}}\left(\frac{\hat{\mathbf{p}}+\hat{\mathbf{p}}'}{2\hat{\mathbf{p}}}|\mathcal{A}\right)_i - \hat{p}_i\right)^2\right\} \\ &= n\left\{\sum_{i=1}^{2m}\frac{\hat{p}_i}{\beta_i^2}\left(\frac{\hat{p}_i' - \hat{p}_i}{2}\right)^2 - \sum_{i=1}^{2m}\frac{\hat{p}_i^3}{\alpha_i^2}\left(E_{\hat{\mathbf{p}}}\left(\frac{\hat{\mathbf{p}}-\hat{\mathbf{p}}'}{2\hat{\mathbf{p}}}|\mathcal{A}\right)_i\right)^2\right\} \end{aligned}$$

where for $i = 1, 2, \ldots, 2m$, β_i (α_i) is between $\hat{p}_i$ and $\hat{p}_i^{(0)}$ ($\hat{p}_i$ and $\hat{p}_i^{(1)}$) and converges almost surely to p_i under H_0.

Let $\boldsymbol{\psi} = \sqrt{n}(\hat{\mathbf{p}}' - \hat{\mathbf{p}})/\hat{\mathbf{p}}$. By the multivariate central limit theorem the random vector $\sqrt{n}(\hat{\mathbf{p}} - \mathbf{p})$ converges in distribution to $\mathbf{p}'(\mathbf{U} - \overline{U}\mathbf{E})$ where $U_1, U_2, \ldots, U_{2m}$ are independent normal random variables with mean zero and respective variances $p_1^{-1}, p_2^{-1}, \ldots, p_{2m}^{-1}$, $\overline{U} = \sum_{i=1}^{2m} p_i U_i$ and $\mathbf{E} = (1, 1, \ldots, 1)^T$. Therefore under H_0

$$\begin{aligned} \boldsymbol{\psi} &= \frac{\sqrt{n}[(\hat{\mathbf{p}} - \mathbf{p})' - (\hat{\mathbf{p}} - \mathbf{p})]}{\hat{\mathbf{p}}} \\ &\xrightarrow{\mathcal{L}} \frac{\mathbf{p}'(\mathbf{U} - \overline{U}\mathbf{E})' - \mathbf{p}(\mathbf{U} - \overline{U}\mathbf{E})}{\mathbf{p}} \\ &= \mathbf{U}' - \mathbf{U} = \mathbf{V}. \end{aligned}$$

Hence

$$\begin{aligned} T_{01} &\xrightarrow{\mathcal{L}} \frac{1}{4}\sum_{i=1}^{2m} p_i \left(V_i - E_{\mathbf{p}}(\mathbf{V}|\mathcal{A})_i\right)^2 \\ &= \sum_{i=m+1}^{2m} \left(V_i - E_{\mathbf{p}_r}(\mathbf{V}_r|\mathcal{I})_i\right)^2 \left(\frac{p_i}{2}\right) \end{aligned}$$

where the last equality follows from Lemma 3.1 and $\mathbf{p}_r$ and $\mathbf{V}_r$ are the restrictions of $\mathbf{p}$ and $\mathbf{V}$ to $m+1, m+2, \ldots, 2m$. As shown in the next theorem, the limiting distribution of T_{01} is chi-bar square which depends on the unknown values of $\mathbf{p}$ through the level probabilities. The least favorable distribution can be found using the theory derived in Section 3.4 of Robertson, Wright and Dykstra (1988). These results are summarized in the following theorem.

THEOREM **3.1** *If $\mathbf{p}$ satisfies H_0 and $p_i > 0, i = 1, 2, \ldots, 2m$, then for any real number t*

$$\lim_{n\to\infty} pr_{\mathbf{p}}(T_{01} \geq t) = \sum_{j=0}^{m} p(j, m, \mathbf{p}_r) pr(\chi^2_{m-j} \geq t)$$

where $p(0, m, \mathbf{p}_r)$ is the probability that $E_{\mathbf{p}_r}(\mathbf{V}_r|\mathcal{I})$ is identically zero and $p(j, m, \mathbf{p}_r)$ for $j = 1, 2, \ldots, m$ is the probability that $E_{\mathbf{p}_r}(\mathbf{V}_r|\mathcal{I})$ has j distinct values. Furthermore,

$$\sup_{\mathbf{p}} \lim_{n\to\infty} pr_{\mathbf{p}}(T_{01} \geq t) = \frac{1}{2} pr(\chi^2_{m-1} \geq t) + \frac{1}{2} pr(\chi^2_m \geq t).$$

A test based upon the least favorable distribution given above is likely to be conservative. There is considerable evidence that if the values of $p_{m+1}, p_{m+2}, \ldots, p_{2m}$ do not vary too much then a test based on equal weights critical value will have a significance level reasonably close to the reported value. These equal weights level probabilities are discussed in Section 3.3 in Robertson, Wright and Dykstra (1988) and are tabulated in A.12 of their book. Since we have 0 as an upper bound in the cone of interest here, the value of m should be increased by 1 to account for it (this is like having $m+1$ normal means indexed by $0, 1, 2, \ldots, m$ with the weight associated with the variable indexed by 0 being ∞.)

Another alternative is to approximate $pr_{\mathbf{p}}(T_{01} \geq t)$ by $\sum_{j=0}^{m} p(j, m, \hat{\mathbf{p}}_r^{(0)}) pr(\chi^2_{m-j} \geq t)$ where $\hat{\mathbf{p}}^{(0)}$ is the estimator of $\mathbf{p}$ under the null hypothesis. This has the same asymptotic distribution as T_{01} and provides a good approximation.

Next we consider testing H_1 as a null hypothesis against the alternative H_2 of no restriction on S_1 and S_2. Let T_{12} be the log-likelihood ratio test statistic corresponding to this situation then

$$
\begin{aligned}
T_{12} &= 2n\left\{\sum_{i=1}^{2m}\hat{p}_i\ln\hat{p}_i - \sum_{i=1}^{2m}\hat{p}_i\ln\hat{p}_i^{(1)}\right\} \\
&= n\sum_{i=1}^{2m}\frac{\hat{p}_i}{\tau_i^2}\left(\hat{p}_i E_{\hat{\mathbf{p}}}\left(\frac{\hat{\mathbf{p}}+\hat{\mathbf{p}}'}{2\hat{\mathbf{p}}}|\mathcal{A}\right)_i - \hat{p}_i\right)^2 \\
&= n\sum_{i=1}^{2m}\frac{\hat{p}_i^3}{4\tau_i^2}\left(E_{\hat{\mathbf{p}}}\left(\frac{\hat{\mathbf{p}}'-\hat{\mathbf{p}}}{\hat{\mathbf{p}}}|\mathcal{A}\right)_i\right)^2
\end{aligned}
$$

where τ_i is between $\hat{p}_i$ and $\hat{p}_i^{(1)}$ and converges almost surely to p_i under H_1.

We have the following theorem.

Theorem **3.2** *For any real number t*

$$\sup_{\mathbf{p}\overset{st}{\leq}\mathbf{p}'}\lim_{n\to\infty} pr_{\mathbf{p}}(T_{12}\geq t) = \sup_{\mathbf{p}=\mathbf{p}'}\lim_{n\to\infty} pr_{\mathbf{p}}(T_{12}\geq t)$$

and

$$\sup_{\mathbf{p}=\mathbf{p}'}\lim_{n\to\infty} pr_{\mathbf{p}}(T_{12}\geq t) = \sum_{l=1}^{k+1}\binom{k}{l-1}2^{-k}pr(\chi^2_{l-1}\geq t).$$

Proof: The proof of this theorem is similar to that of Theorem 4.2 of Robertson and Wright (1981). We only give the main idea.

Let $\eta_0 = 0 < \eta_1 < \ldots < \eta_A = 2m$ be such that $p_1+\cdots+p_{\eta_i} = p_1'+\cdots+p_{\eta_i}', i = 1,2,\ldots,A$ and $p_1+\cdots+p_i > p_1'+\cdots+p_i'$ for $i\neq\eta_j, j = 1,2,\ldots,A$. We note that A is even only when $P(\delta=1) = P(\delta=2)$.

Assume that $P(\delta=1) < P(\delta=2)$, then $A = 2k+1$ for some k. Let

$$\mathcal{D} = \{\mathbf{x}\in\mathcal{A},\ x_{\eta_{i-1}+1} = x_{\eta_{i-1}+2} = \ldots = x_{\eta_i},\ i\neq k+1,\ x_{\eta_k+1} = x_{\eta_k+2} = \ldots = x_{\eta_{k+1}} = 0\}.$$

It follows from Robertson and Wright (1981) that for all w and n sufficiently large

$$E_{\hat{\mathbf{p}}}\left(\frac{\hat{\mathbf{p}}'-\hat{\mathbf{p}}}{\hat{\mathbf{p}}}|\mathcal{A}\right) = E_{\hat{\mathbf{p}}}\left(\frac{\hat{\mathbf{p}}'-\hat{\mathbf{p}}}{\hat{\mathbf{p}}}|\mathcal{D}\right)$$

and

$$
\begin{aligned}
T_{12} &= n\sum_{i=1}^{2m}\frac{\hat{p}_i^3}{4\tau_i^2}\left(E_{\hat{\mathbf{p}}}\left(\frac{\hat{\mathbf{p}}'-\hat{\mathbf{p}}}{\hat{\mathbf{p}}}|\mathcal{A}\right)_i\right)^2 \\
&= \sum_{i=1}^{2m}\frac{\hat{p}_i^3}{4\tau_i^2}\left(E_{\hat{\mathbf{p}}}\left(\frac{\sqrt{n}(\hat{\mathbf{p}}'-\hat{\mathbf{p}})}{\hat{\mathbf{p}}}|\mathcal{D}\right)_i\right)^2 \\
&\xrightarrow{\mathcal{L}} \frac{1}{4}\sum_{i=1}^{2m} p_i\left(E_{\mathbf{p}}\left(V|\mathcal{D}\right)_i\right)^2
\end{aligned}
$$

where $\mathbf{V} = (V_1, V_2, \ldots, V_{2m})^T$ is as defined in the proof of the previous theorem.

Let $\mathcal{A}_{\mathbf{p}}^{(k)} = \{(u_1, u_2, \ldots, u_{2k})^T, \quad u_1 \geq u_2 \geq \ldots \geq u_k \geq 0 \geq u_{k+1} \geq \ldots \geq u_{2k}\}$ and let

$$
w_i = \sum_{j=\eta_{i-1}+1}^{\eta_i} p_j, \quad i = 1, 2, \ldots, 2k+1, i \neq k+1
$$

Define $\mathbf{V}^{(k)} = (V_1^{(k)}, V_2^{(k)}, \ldots, V_{2k}^{(k)})$ where $V_i^{(k)} = V_{2k-i+1}^{(k)}, i = 1, 2, \ldots, k$, and $V_i^{(k)}, i = 1, 2, \ldots, k$, are independent and normally distributed with means equal zero and variances equal to $1/w_i, i = 1, 2, \ldots, k$. Careful inspection of the least squares projection onto $\mathcal{D}$ shows that

$$
\begin{aligned}
T_{12} &\xrightarrow{\mathcal{L}} \frac{1}{4}\sum_{i=1}^{2m} p_i\left(E_{\mathbf{p}}\left(\mathbf{V}|\mathcal{D}\right)_i\right)^2 \\
&= \frac{1}{4}\sum_{i=1}^{2k} w_i\left(E_{\mathbf{w}}\left(\mathbf{V}^{(k)}|\mathcal{A}_{\mathbf{p}}^{(k)}\right)_i\right)^2 \\
&= \sum_{i=1}^{k}\frac{w_i}{2}\left(E_{w_r}\left(V_r^{(k)}|\mathcal{I}_r^{(k)}\right)\right)^2
\end{aligned}
$$

where $\mathcal{I}_r^{(k)} = \{(u_1, \ldots, u_k)^T, \quad u_1 \geq u_2 \geq \ldots \geq u_k \geq 0\}, \mathbf{V}_r^{(k)} = (V_1, V_2, \ldots, V_k)^T, w_r = (w_1, w_2, \ldots, w_k)^T$ and the last equality follows from Lemma 3.1.

Since

$$
\sum_{i=1}^{2m} V_i^2 p_i = \sum_{i=1}^{2m}\left(V_i - E_{\mathbf{p}}\left(\mathbf{V}|\mathcal{D}\right)_i\right)^2 p_i + \sum_{i=1}^{2m}\left(E_{\mathbf{p}}\left(\mathbf{V}|\mathcal{D}\right)_i\right)^2 p_i
$$

and $\mathcal{D} \subset \mathcal{A}$ with equality only when $\mathbf{p} = \mathbf{p}'$, we have

$$
\sum_{i=1}^{2m}\left(E_{\mathbf{p}}\left(\mathbf{V}|\mathcal{D}\right)_i\right)^2 p_i
$$

is largest when $\mathbf{p} = \mathbf{p}^{'}$ and hence the first conclusion of the theorem. If it is the case that $P(\delta_1 = 1) = P(\delta_1 = 2)$, then $A = 2k$ for some k and the proof is the same as above with $\mathcal{D} = \{\mathbf{x} \in \mathcal{A},\ x_{\eta_{i-1}+1} = x_{\eta_{i-1}+2} = \ldots = x_{\eta_i},\ i = 1, 2, \ldots, 2k\}$ and $\mathcal{A}_{\mathbf{p}}^{(k)} = \{(u_1, u_2, \ldots, u_{2k})^T,\ u_1 \geq u_2 \geq \ldots \geq u_{2k}\}$. The second conclusion follows from Theorem 3.6.1 of Robertson, Wright and Dykstra (1988).

4 Simulation Studies

In this section we compare the power of our test T_{01} with that of the likelihood ratio test T_{01}^{*} described in Dykstra, Kochar and Robertson (1995a) for testing the equality of cause specific hazard rates against certain ordered alternatives belonging to H_1.

In the first study we assume that X and Y are independent exponential random variables with parameters λ_1 and λ_2, respectively and the available data are only in the form of (T, δ) for each pair of observations. We group the data into 4 cells obtained by dividing the real line into 4 intervals using the quartiles of $T = \min(X, Y)$. The sample size used in our study is 50 and the number of replications is 5000. For testing H_0 against H_1 at 5% level, the cutoff point of T_{01} is 8.191 and is obtained by simulating the weights in the limiting distributions in Theorem 3.1. The 5% cutoff point for T_{01}^{*} is 6.498 and the results are reported in Table 4.1. It is not surprising that in this case the T_{01}^{*} test is more powerful. For this alternative, the cause specific hazard rates are ordered and the T_{01}^{*} is the likelihood ratio statistic for this alternative.

Table 4.1

POWERS OF THE TESTS

λ_1	λ_2	T_{01}	T_{01}^*
40	40	0.0544	0.0568
40	45	0.0874	0.1052
40	50	0.1260	0.1718
40	55	0.1892	0.2540
40	60	0.2608	0.3492
40	65	0.3212	0.4356
40	70	0.4092	0.5190
40	75	0.4744	0.5962
40	80	0.5616	0.6784
40	85	0.6298	0.7446
40	90	0.6954	0.7994
40	95	0.7446	0.8340
40	100	0.7928	0.8768

In the next simulation study we assume that X and Y are independent and that under the null hypothesis each of them has exponential distribution with mean equal to 2. We assume that under the alternative hypothesis X has an exponential distribution with mean equal to 2 and Y has a Weibull distribution with scale parameter 1/2 and shape parameter 2. Note that in this case the cause specific hazard rates are the same as the ordinary hazard rates and are given by $h_X(t) = 1/2$ and $h_Y(t) = t$. Although in this case the cause specific hazard rates cross each other, it can be seen that their sub-survival functions are ordered (i.e. $S_Y(t) \geq S_X(t), \forall t$). For testing H_0 against H_1 at level 0.05, the same cutoff points as in the previous example are used and results of the power study are reported in Table 4.2.

Table 4.2

POWERS OF THE TESTS

Sample size	T_{01}	T_{01}^*
30	0.4990	0.3838
40	0.6052	0.4650
50	0.6996	0.5502
60	0.7816	0.6314
70	0.8322	0.7034
80	0.8744	0.7522
90	0.9170	0.8120
100	0.9412	0.8412
110	0.9610	0.8774
120	0.9762	0.9038

It is clear from the above tables that the likelihood ratio test T_{01} is quite powerful.

4.1 Example

We consider the mortality data on mice reported in Hoel (1972). The data were obtained from the laboratory experiment on RFM strain male mice which had received a radiation dose of 300r at an age of 5-6 weeks and were then kept in a conventional environment. We consider only two major risks of death- the second risk is cancer and the first risk is the combination of all other risks. This data set has also been analyzed previously by Dykstra, Kochar and Robertson (1995a) where they test for the equality of the cause specific hazard rates of the two risks. The grouped data along with the various estimates is given below in Table 4.3.

Table 4.3

No.	Interval	d_{1j}	d_{2j}	$\hat{p}_{1j}$	$\hat{p}_{2j}$	$p_{1j}^{(1)}$	$p_{2j}^{(1)}$
1	< 350.0	15	18	0.1515	0.1818	0.1515	0.1818
2	[350,450)	6	7	0.0606	0.0707	0.0606	0.0707
3	[450,550)	6	4	0.0606	0.0404	0.0606	0.0404
4	[550,650)	8	18	0.0808	0.1818	0.0808	0.1818
5	[650,750)	2	12	0.0202	0.1212	0.0202	0.1212
6	[750,850)	2	1	0.0202	0.0101	0.0152	0.0152

For this example the value of $T_{01} = 12.6247$ and the value of $T_{12} = 0.3397$. The simulated level probabilities using $\hat{\mathbf{p}}_r^{(0)}$ are $p(0, 6, \hat{\mathbf{p}}_r^{(0)}) = 0.22405, p(1, 6, \hat{\mathbf{p}}_r^{(0)}) = 0.08370, p(2, 6, \hat{\mathbf{p}}_r^{(0)}) = 0.43755, p(3, 6, \hat{\mathbf{p}}_r^{(0)}) = 0.20570, p(4, 6, \hat{\mathbf{p}}_r^{(0)}) = 0.04375, p(5, 6, \hat{\mathbf{p}}_r^{(0)}) = 0.00515, p(6, 6, \hat{\mathbf{p}}_r^{(0)}) = 0.00010$ and the p-value for testing H_0 versus $H_1 - H_0$ using these weights is 0.0204 and is 0.0383 when the least favorable distribution is used instead. For testing H_1 versus $H_2 - H1$, the p-value based on the least favorable distribution given in Theorem 3.2 is 0.8883.

This data would appear to support the conclusion that the sub-survival is larger for the risk of cancer than for all other risks when that hypothesis is compared to the equality of the two sub-survival functions. However, the value of the test statistic as developed in Dykstra et al. (1995a) for testing equality of the cause specific hazard rates against the alternative that they are ordered is 12.2222 and the p-value using Table 1 in Dykstra, Kochar and Robertson (1995a) is approximately 0.01.

Acknowledgments

The authors are grateful to the referee for carefully reading the paper.

References

1. Aly, E., Kochar, S. and McKeague, I. (1994). Some tests for comparing cause specific hazard rates. *J. Amer. Statist. Assoc.*, **89**, 994-999.

2. Barlow, R. E. and Brunk, H. D. (1972). The isotonic regression problem and its dual. *J. Amer. Statist. Assoc.*, **67**, 140-147.

3. Brunk, H. D., Franck, W. E. Hanson, D. L. and Hogg, R. V. (1966). Maximum likelihood estimation of the distributions of two stochastically ordered random variables. *J. Amer. Statist. Assoc.*, **61**, 1067-1080.

4. Carriere, K, C. and Kochar, S. C. (2000). Comparing sub-survival functions in a competing risks model. *Lifetime Data Analysis*, **6**, 85-97.

5. Deshpandé, J. V. (1990). A test for bivariate symmetry of dependent competing risks. *Biometric Journal* **32**, 737-746.

6. Deshpandé, J. V. and Karia, S. R. (1995). Modelling competing risks data via the cumulative incidence functions. *J. IAPQR* **20**, 79-91.

7. Dykstra, R., Kochar, S. and Robertson, T. (1995a). Likelihood based inference for cause specific hazard rates under order restrictions. *J. Mult. Anal.*, **54**, 163-174.

8. Dykstra, R., Kochar, S. and Robertson, T. (1995b). Likelihood ratio tests for symmetry against one-sided alternatives. *Annals of the Institute of Statistical Mathematics*, **47**, 719- 730.

9. Hoel, D. G. (1972). A representation of mortality data by competing risks. *Biometrics* , **65**, 475-488.

10. Kochar, S. C. (1995). A review of some distribution-free tests for the equality of cause specific hazard rates. *Analysis of Censored Data.* IMS-LNMS **27**, 147-162. Editors: Hira L. Koul and J.V. Desphandé.

11. Peterson, A. (1977). Expressing the Kaplan-Meier estimator as a function of empirical sub-survival functions. *J. Amer. Statist. Assoc.*, **72**, 854-858.

12. Sun, Y. and Tiwari, R. C. (1995). Comparing cause-specific hazard rates of a competing risks model with censored data. *Analysis of Censored Data.* IMS-LNMS **27**, 255-270. Editors: Hira L. Koul and J.V. Desphandé.

13. Robertson, T., Wright, T. (1981). Likelihood Ratio Tests For and Against a Stochastic Ordering Between Multinomial Populations. *Annals of Statistics*, **9**, 1248-1257.

14. Robertson, T., Wright, T. and Dykstra, R. (1988). *Order Restricted Statistical Inference.* Wiley, New York. nonidentifiability in 235-239.