Tests for some statistical hypotheses for dependent competing risks - A review

Date :July, 2003

1 Introduction

The competing risks situation arises in life studies when a unit is subject to many, say k , modes of failure and the actual failure , when it occurs , can be ascribed to a unique mode. These k modes are also called the k risks to which the unit is exposed and as they all seemingly compete for the life of the unit, the term 'competing risks' is used to describe it. Suppose that the continuous positive valued random variable T represents the lifetime of the unit and δ taking values $1, 2, \ldots, k$ represent the risk which caused the failure of the unit.

The joint probability distribution of (T, δ) is specified by the set of k distribution functions

$$F(i, t) = P[T \leq t, \delta = i], \tag{1}$$

or equivalently by the subsurvival functions

$$S(i, t) = P[T > t, \delta = i], \; i = 1, 2, \ldots, k. \tag{2}$$

Let $H(t)$ and $S(t)$ denote , respectively the distribution function and the survivor function of T. Let $f(i, t)$ denote the sub-density function corresponding to ith risk. Then the density function of T is $h(t) = \sum_{i=1}^{k} f(i, t)$. It is easy to see that

$$\sum_{i=1}^{k} F(i, t) = H(t), \;\; \sum_{i=1}^{k} S(i, t) = S(t)$$

and $p_i = F(i, \infty)$ is the probability of failure due to the ith risk.

A commonly used description of the competing risks situation is the latent failure time model. Let $X_1, X_2, \ldots, X_k$ be the latent failure times of any unit exposed to k risks, where X_i represents the time to failure if cause i were the only cause of failure present in the situation. F_i denotes the marginal distribution of X_i and $F(x_1, \ldots, x_k)$ denotes their joint distribution. The observable random variables are still the time to failure T, where $T = \min(X_1, X_2, \ldots, X_k)$ and the cause of failure δ where $\delta = j$ if $X_j = \min(X_1, X_2, \ldots, X_k)$. If $X_1, X_2, \ldots, X_k$ are independent, then their marginal distributions carry all the probabilistic information regarding the k risks. It is easily seen that the marginal and hence the joint distribution is identifiable from the probability distribution of the observable random variables (T, δ). However, in general when the risks are not independent, neither the joint distribution of $X's$ nor their marginals are identifiable from the probability distribution of (T, δ) (Tsiatis (1975), Crowder (1991), (1993)). They have proved that a unique independent and infinitely many dependent probability distributions of $(X_1, X_2, \ldots, X_k)$ correspond to a single probability distribution of (T, δ). Hence , the independence or otherwise of the latent lifetimes $(X_1, X_2, \ldots, X_k)$ cannot be statistically tested from any data collected on (T, δ). The independence of $(X_1, X_2, \ldots, X_k)$ has to be assumed on the basis of a priori information, if any.

Also, the marginal distribution functions $F_i(x)$ may not represent the probability distribution of lifetimes in any practical situation. Elimination of jth risk may change the environment in such a way that $F_i(x)$ does not represent the lifetime of X_i in the changed scenario.

In view of the above considerations, unless one can assume independence, it is necessary to suggest appropriate models, develop methodology and carry out the data analysis in terms of the observable random variables (T, δ) alone.

Kalbfleisch and Prentice (1980) proposed methods for analysing competing risks data in terms of cause specific hazard rates $\lambda_i(t)$, where

$$\lambda(i,t) = \lim_{\Delta t \to 0} \frac{1}{\Delta t} P(t \leq T < t + \Delta t, \delta = i | T > t) = \frac{f(i,t)}{S(t)} \tag{3}$$

For recent results see Crowder (2001), Kalbfleisch and Prentice (2002). Dewan and Kulathinal (2003) have considered parametric models for sub-survival functions by assuming a suitable parametric form for cause specific hazard rates. Aly, Kochar and McKeague (1994), Kochar (1995) and Sun and Tiwari (1995) have considered the problem of testing for equality of cause-specific hazard functions. Deshpande (1990), Aras and Deshpande (1992), Deshpande and Dewan (2000) have considered the problem of analysing competing risks data by using the sub-distribution functions and sub-survival functions. Recently Dewan, Deshpande and Kulathinal (2003) have proposed tests for testing independence of T and δ.

2 Locally Most Powerful Rank Tests

Suppose $k = 2$, that is , a unit is exposed to two risks of failure denoted by 1 and 0. When n units are put to trial , the data consists of $(T_i, {\delta^*}_i), i = 1, \ldots, n$ where $\delta^* = 2 - \delta$.

Suppose we wish to test the hypothesis $H_0 : F(1,t) = F(2,t), \text{for all } \ t$. First we look at tests based on likelihood theory. Under the null hypothesis the two risks are equally effective. However one would expect that under the alternative hypothesis the two risks are not equally effective atleast at some ages.

The likelihood function is given by (see , Aras and Deshpande (1992))

$$L(\underline{T}, \underline{\delta^*}, f(1,t_i), f(2,t_i)) = \prod_{i=1}^{n} [f(1,t_i)]^{{\delta^*}_i} [f(2,t_i)]^{1-{\delta^*}_i} \tag{1}$$

where $\underline{T} = (T_1, \ldots, T_n), \underline{\delta^*} = (\delta_1, \ldots, \delta_n)$.

If $F(i,t)$ depends upon the parameter θ , then inference about it can be based on the above likelihood function.

When T and δ^* are independent then Deshpande (1990) proposed the model

$$F(1,t) = \theta H(t), \quad F(2,t) = (1-\theta) H(t). \tag{2}$$

Here $\theta = P[\delta^* = 1]$. Under this model the likelihood reduces to

$$L(\theta, H) = \theta^{\sum_{i=1}^{n} {\delta^*}_i} (1-\theta)^{\sum_{i=1}^{n} (1-{\delta^*}_i)} \prod_{i=1}^{n} h(t_i). \tag{3}$$

The hypothesis $F(1,t) = F(2,t)$ reduces to testing that $\theta = 1/2$. Then the obvious statistic is the sign statistic

$$U_1 = \frac{1}{n}\sum_{i=1}^{n} \delta^*{}_i. \tag{4}$$

nU_1 has $B(n,\theta)$ distribution and one can have optimal estimation and testing procedures based on it. However, if $F(1,t)$ and $F(2,t)$ depend on a parameter θ in a more complicated manner then one needs to look at locally most powerful rank tests.

Let $f(1,t) = f(t,\theta), f(2,t) = h(t) - f(t,\theta)$ where $h(t)$ and $f(t,\theta)$ are known density functions and incidence density such that $f(t,\theta_0) = \frac{1}{2}h(t)$.

Let $T_{(1)} \leq T_{(2)} \leq \ldots \leq T_{(n)}$ denote the ordered failure times. Let

$$W_i = \begin{cases} 1 & \text{if } T_{(i)} \text{ corresponds to first risk} \\ 0 & \text{otherwise.} \end{cases} \tag{5}$$

Let R_j be the rank of T_j among $T_1, \ldots, T_n$. Let

$$\underline{R}' = (R_1, R_2, \ldots, R_n), \ \ \underline{W}' = (W_1, W_2, \ldots, W_n)$$

denote the vector of ranks and indicator functions corresponding to ordered minima. The likelihood of ranks is given by

$$P(\underline{R}, \underline{W}, \theta) = \int \cdots \int_{0<t_1<\ldots<t_n<\infty} \prod_{i=1}^{n} [f(t_i,\theta)]^{w_i} [h(t_i) - f(t_i,\theta)]^{1-w_i} dt_i \tag{6}$$

Theorem : If $f'(t,\theta)$ is the derivative of $f(t,\theta)$ with respect to θ, then the locally most powerful rank test for $H_0 : \theta = \theta_0$ against $H_1 : \theta > \theta_0$ is given by: reject H_0 for large values of $L_c = \sum_{i=1}^{n} w_i a_i$, where

$$a_i = \int \cdots \int_{0<t_1<\ldots<t_n<\infty} \frac{f'(t_i,\theta_0)}{f(t_i,\theta_0)} \prod_{i=1}^{n} [f(t_i,\theta_0) dt_i]. \tag{7}$$

Special cases

(i) If the model (2) holds with $\theta_0 = 1/2$, then sign test is the LMPR test.

(ii) If $f(1,t) = \frac{1}{2}g(t,0)$ and $f(2,t) = \frac{1}{2}g(t,\theta), \theta > 0$ and $g(t,\theta)$ is the logistic density function $g(t,\theta) = \frac{e^{(x-\theta)}}{[1+e^{(x-\theta)}]^2}$, then the LMPR test is based on the statistic $W^+ = \sum_{i=1}^{n} W_i R_i$, which is the analogue of the Wilcoxon signed rank statistic for competing risks data.

(iii) In case of Lehmann type alternative defined by $F(1,t) = [\frac{H(t)}{2}]^\theta, \ \ F(2,t) = H(t) - [\frac{H(t)}{2}]^\theta$,, LMPR test is based on scores $a_i = E(E_{(j)})$ where $E_{(j)}$ is the jth order statistic from a random sample of size n from standard exponential distribution.

But for more complicated families of distributions, e.g., Gumbel (1960), the scores are complicated and need to be solved using numerical integration (see Aras and Deshpande (1992)).

3 Tests for bivariate symmetry

Assume that the latent failure times X and Y are dependent. Suppose their joint distribution is given by $F(x,y)$. On the basis of independent pairs $(T_i, {\delta^*}_i)$ we want to test whether the forces of two risks are equivalent against the alternative that the force of one risk is greater than that of the other. That is we test the null hypothesis of bivariate symmetry

$$H_0 : F(x,y) = F(y,x) \text{ for every } (x,y) \tag{1}$$

Before we formulate the alternative of interest let us consider the following theorem which is easy to prove.
Theorem 3.1: Under the null hypothesis of bivariate symmetry we have
(i)$F(1,t) = F(2,t)$ for all t,
(ii)$S(1,t) = S(2,t)$ for all t,
(iii) $\lambda(1,t) = \lambda(2,t)$ for all t,
(iv) $P[\delta^* = 1] = P[\delta^* = 0]$
(v) T and δ^* are independent.

In view of the above theorem, the following alternatives to the null hypothesis are worth considering.

$$\begin{aligned} H_1 : &\quad \lambda(1,t) < \lambda(2,t) \\ H_2 : &\quad F(1,t) < F(2,t) \\ H_3 : &\quad S(1,t) > S(2,t). \end{aligned} \tag{2}$$

All these alternatives say that risk II is more potent than risk I at all ages t in some stochastic sense.

Sen (1979) considered fixed sample and sequential tests for the null hypothesis of bivariate symmetry of the joint distribution of (X,Y).The alternatives are expressed in terms of $\pi_1(t) = Pr[\tilde{\delta} = 1|T = t]$, the conditional probability that the failure is due to first risk , given that failure occurs at time t. He derived optimal score statistics for such parametric situations. But the statistics cannot be used without the knowledge of the joint distribution $F(x,y)$.

We look at various distribution-free test procedures for testing H_0 against above alternatives.

For testing H_0 against H_1 consider

$$\begin{aligned} \psi(t) &= F(1,t) - F(2,t) \\ &= \int_0^t S(t)[\lambda(1,u) - \lambda(2,u)]du \end{aligned} \tag{3}$$

H_1 holds iff the above function is non-increasing on t.

Kochar and Dewan (2000) have suggested considering the following measure of deviation between H_0 and H_1,

$$\Delta = \int_{0<x<y<\infty} [\psi(x) - \psi(y)]\, dF(x)\, dF(y) \tag{4}$$

and its empirical estimator Δ_n as the test statistic where Δ_n is given by

$$\Delta_n = \int_{0<x<y<\infty} [\psi_n(x) - \psi_n(y)]\, dF_n(x)\, dF_n(y) \tag{5}$$

where

$$\begin{aligned} F_{1n}(t) &= \frac{1}{n}\sum_{j=1}^{n} I\{\delta_j = 1, T_j \le t\}, \\ F_n(t) &= \frac{1}{n}\sum_{j=1}^{n} I\{T_j \le t\}, \end{aligned}$$

and

$$\psi_n(t) = 2F_{1n}(t) - F_n(t). \tag{6}$$

are the empirical estimators of F_1, F and ψ, respectively.

Then

$$\begin{aligned} \Delta_n &= \int_{0<x<y<\infty} [\psi_n(x) - \psi_n(y)]\, dF_n(x)\, dF_n(y) \\ &= \frac{1}{n^3} \sum_{1\le i<j\le n} [j - i - 2\sum_{\ell=i+1}^{j} W_\ell] \\ &= \frac{1}{n^3}[\frac{n(n^2-1)}{6} - 2\sum_{i=1}^{n}(i-1)(n-i+1)W_i] \end{aligned} \tag{7}$$

where W_i is as defined in (2.5).

Under H_0, $\Delta = 0$, but under the alternative, $\Delta > 0$. Large values of the statistic are significant. Rejecting H_0 for large values of Δ_n is equivalent to rejecting it for small values of the statistic

$$U_2 = \sum_{i=1}^{n-1} i(n-i)W_{i+1} \tag{8}$$

Since T and δ are independent under H_0, $W_1, W_2, \ldots, W_n$ are i.i.d. Bernoulli random variables with $P\{W_i = 1\} = P\{W_i = 0\} = \frac{1}{2}$, $i = 1, 2, \ldots, n$. Hence, the moment generating function of the null distribution of U_2 is given by

$$M(t) = 2^{-n+1}\prod_{i=1}^{n-1}(1 + e^{a_i t}) \tag{9}$$

where $a_i = i(n-i), i = 1, 2, \cdots, n-1$.

Using (8), the exact null distribution of U_2 can be obtained following the approach of Hettmansperger (1984 pp 35). For $n = 5(1)20$, the 5% and 1% critical values of U_2 are given in Table 2.1.

TABLE 2.1

Critical points of U_2 and exact significance levels.

n	$\alpha \approx 0.01$	$\alpha \approx 0.05$
5	-	0, 0.0625
6	-	0, 0.03125
7	0, 0.015625	6, 0.046875
8	7, 0.015625	14, 0.046875
9	8, 0.011719	22, 0.054688
10	18, 0.011719	34, 0.054688
11	28, 0.013672	50, 0.050781
12	38, 0.010254	67, 0.051758
13	54, 0.010498	90, 0.051514
14	74, 0.01001	118, 0.052734
15	98, 0.010132	148, 0.050293
16	126, 0.01001	186, 0.0513
17	158, 0.01001	228, 0.050308
18	198, 0.010269	278, 0.050613
19	242, 0.010311	332, 0.049911
20	290, 0.01005	396, 0.050467

The mean and variance of U_2 under H_0 are given by

$$E[U_2] = \frac{n(n^2-1)}{12}, \quad Var[U_2] = \frac{n(n^4-1)}{120} \tag{10}$$

By the Central Limit Theorem, it can be shown that, under H_0,

$$n^{\frac{1}{2}}\{\frac{U_2}{n^3} - \frac{1}{12}\} \xrightarrow{\mathcal{L}} N(0, \frac{1}{120}). \tag{11}$$

Hence, for large n, the critical values of U_2 can be obtained by using the above normal approximation.

To give an idea about the accuracy of the normal approximation of U_2, we give the exact (respectively, approximate) significance levels for $n = 10$ and $n = 20$. For $n = 10$, $P[U_2 \leq 18] = 0.011719$ (resp. 0.0118) and $P[U_2 \leq 34] = 0.054688$ (resp. 0.0594). For $n = 20$, $P[U_2 \leq 290]$= 0.01005 (resp. 0.0111) and $P[U_2 \leq 396]$ = 0.050467 (resp. 0.049).

After change of integration, Δ can also be expressed as

$$\Delta = \int_0^\infty S^2(x)F(x)[\lambda_2(x) - \lambda_1(x)]\, dx. \tag{12}$$

It is seen from this representation that the test based on U_2 is equivalent to the test V of Yip and Lam (1992) proposed for the case of independent risks without censoring. They have not discussed its small sample exact null distribution.

Deshpande (1990) proposed two tests for testing H_0 versus H_2 on heuristic grounds.

The first test is the Wilcoxon signed rank type statistic

$$W^+ = \sum_{i=1}^{n}(1 - \delta^*{}_i)R_i. \tag{13}$$

It was felt that W^+ will be large when the alternative H_2 is true, there being a greater incidence of the second risk upto any fixed time t.

Another test is based on the U-statistic

$$U_3 = \frac{1}{\binom{n}{2}} \sum_{1 \le i < j \le n} \phi_3(T_i, \delta^*{}_i, T_j, \delta^*{}_j) \tag{14}$$

where ϕ_3 is given by

$$\phi_3(T_i, \delta^*{}_i, T_j, \delta^*{}_j) = \begin{cases} 1 & \text{if } \delta^*{}_j = 0, T_i > T_j \\ & \text{or } \ \delta^*{}_i = 0, T_i < T_j \\ 0 & \text{otherwise} \end{cases} \tag{15}$$

Here the kernel ϕ_3 takes value 1 if , and only if, a Y observation is the smallest among (X_i, X_j, Y_i, Y_j).

Note that

$$\binom{n}{2} U_3 = \sum_{i=1}^{n} (n - R_i + 1)\delta^*{}_i. \tag{16}$$

Note that $E(U_3) = 1/2$ under H_0 and strictly larger than $1/2$ under H_2. U_3 is same as the statistic proposed earlier to test for H_0 against the alternative H_{A1} of stochastic dominance of distribution functions of independent latent failure times (see, Deshpande and Dewan (2003)) is also consistent for testing bivariate symmetry against dominance of incidence functions.

For testing H_0 against H_2, one can consider the measure of deviation $F(2,t) - F(1,t)$, which is non-negative under H_2. Then

$$\int_0^\infty [F(2,t) - F(1,t)]dH(t) = P[\delta^*{}_1 = 0, T_1 \le T_2] - \frac{1}{2}. \tag{17}$$

A U-statistic estimator of this parameter is the statistic U_3 discussed above.

Similarly for testing H_0 against H_3 consider the measure of deviation $S(1,t) - S(2,t)$, which is non-negative under H_3. Then

$$\int_0^\infty [S(1,t) - S(2,t)]dH(t) = P[\delta^*{}_1 = 1, T_1 > T_2] - \frac{1}{2}. \tag{18}$$

Consider the kernel

$$\phi_4(T_i, \delta_i, T_j, \delta_j) = \begin{cases} 1 & \text{if } \delta^*{}_i = 1, T_i > T_j \\ & \text{or } \ \delta^*{}_j = 1, T_i < T_j \\ 0 & \text{otherwise} \end{cases} \tag{19}$$

Then the corresponding U-statistic is given by

$$U_4 = \frac{1}{\binom{n}{2}} \sum_{1 \le i < j \le n} \phi_4(T_i, \delta_i, T_j, \delta_j) \tag{20}$$

Then

$$\binom{n}{2} U_4 = \sum_{i=1}^{n} (R_i - 1)\delta^*{}_i. \tag{21}$$

This statistic was earlier proposed by Bagai, Deshpande and Kochar (1989) to test for equality of failure rates of independent latent competing risks.

Aly , Kochar and Mckeague (1994) proposed Kolmogrov-Smirnov type tests for testing the equality of two competing risks against the alternatives H_1 and H_2. Here we discuss their approach.

Consider

$$\begin{aligned} \psi^*(t) &= F(2,t) - F(1,t) \\ &= \int_0^t S(u)(\lambda(2,u) - \lambda(1,u))du. \end{aligned} \tag{22}$$

Under H_0, $\psi^*(t) = 0$. H_1 holds iff $\psi^*(t)$ is nondecreasing in t . Let $\psi_n^*(t)$ be its empirical estimator. Consider

$$D_{1n} = \sup_{0 \le s < t < \infty} \{\psi_n^*(t) - \psi_n^*(s)\}. \tag{23}$$

Large values of the statistic are significant.

The exact null distribution of D_{1n} is given by

$$\begin{aligned} P[nD_{1n} < t] &= \frac{2}{2t+1} \sum_{j=0}^{2t} (\cos \frac{j\pi}{2t+1})(\sin \frac{j\pi(t+1)}{2t+1}) \\ &\quad \times (1 + \cos \frac{j\pi}{2t+1})(\frac{1-(-1)^j}{2})/ \sin \frac{j\pi}{2t+1}. \end{aligned} \tag{24}$$

for $t = 1, \ldots, n+1$.

The asymptotic null distribution of D_{1n} is given by

$$\sqrt{n} D_{1n} \xrightarrow{\mathcal{L}} \sup_{0 \le x \le 1} |W(x)|, \tag{25}$$

where $\{W(t), t \ge 0\}$ is a standard Brownian motion. The asymptotic $0.90, 0.95, 0.99$ quantiles of $\sqrt{n}D_{1n}$ were found to be $1.96, 2.241, 2, 807$, respectively.

If one is interested in a general two sided alternative $F(1,t) \neq F(2,t)$ for some t or equivalently $\lambda_1(t) \neq \lambda_2(t)$ for some t , then one can use the Kolmogrov-Smirnov type statistic

$$D_n = \sup_{t \ge 0} |\psi_n^*(t)|. \tag{26}$$

Under H_0, $\sqrt{n}D_n$ converges in distribution to $\sup_{0 \le x \le 1} |W(x)|$. This test is consistent against arbitrary departures from H_0.

For testing H_0 against H_2, Aly , Kochar and McKeague (1994) proposed the statistic

$$D_{2n} = \sup_{0 \le t \infty} \psi_n(t). \tag{27}$$

Large values of D_{2n} are significant for testing H_0 against H_2. The exact nulltribution of D_{2n} is given by

$$P[nD_{2n} = k] = \frac{1}{2^n} \binom{n}{[\frac{n-k}{2}]}, \quad k = 0, 1, \ldots, n. \tag{28}$$

The asymptotic null distribution is given by

$$P[\sqrt{n}D_{2n} > x] \to 2(1 - \Phi(x)), \quad x \ge 0, \tag{29}$$

where Φ is the standard normal distribution function.

4 Censored data

Most of the above tests can be generalized to the case when the data is right censored. Let C be the censoring random variable independent of the latent failure times X and Y. Denote the survival function of C by S_c and assume that $S_C(t) > 0$ for all t.

Now the available information consists of $(\tilde{T}_i, \tilde{\delta}_i)$, $i = 1, 2, \cdots, n$ where $\tilde{T} = \min(T, C)$ and $\tilde{\delta} = \delta^* I(T \leq C)$. Aly et al (1994) generalised the function ψ^* so as to capture departures of H_0 from H_2 in case of censored data.

Let

$$\phi(t) = \int_0^t S(u-)(S_C(u-))^{\frac{1}{2}}(\lambda(2,u) - \lambda(1,u))du \tag{1}$$

which is the ψ^* function when there is no censoring.

The integrand $S_C(u-)^{1/2}$ is the function required to compensate for censoring in order that the D statistics remain asymptotically distribution-free. Under $H_0, \phi(t) = 0$. H_1 holds iff $\phi(t)$ is increasing in t.

The relevant statistic is

$$D_{3n} = \sup_{0 \leq s < t < \infty} \{\phi_n(t) - \phi_n(s)\}, \tag{2}$$

An obvious choice of ϕ_n is

$$\phi_n(t) = \int_0^t \hat{S}(u-)(\hat{S}_C(u-))^{1/2}\, d(\hat{\Lambda}_1 - \hat{\Lambda}_2)(u) \tag{3}$$

where $\hat{S}$ and $\hat{S}_C$ are the product limit estimators of $\bar{S}$ and S_C, and $\hat{\Lambda}_j$ is the Aalen estimator of the cumulative CSHR function $\Lambda_j(t) = \int_0^t \lambda_j(u)\, du$.

$$\hat{\Lambda}_j(t) = \sum_{i | \tilde{T}_i \leq t} \frac{I(\tilde{\delta}_i = j)}{R_i} \tag{4}$$

where $R_i = \#\{k : \tilde{T}_k \geq \tilde{T}_i\}$ is the risk set at time $\tilde{T}_i$. Large values of the statistic are significant.

Since $\phi(t) > 0$ for some t under H_2, a suitable test procedure is based on large values of

$$D_{4n} = \sup_{0 \leq t < \infty} \phi_n(t). \tag{5}$$

Aly et al (1994) showed that D_{3n} and D_{4n} are asymptotically distribution-free with the same limiting distributions as those obtained in the uncensored case.

A suitable modification of U_2 to censored data is given by the statistic

$$K_n = \int_{0<x<y<\infty} (\phi_n(x) - \phi_n(y))\, d\hat{S}(x)\, d\hat{S}(y) \tag{6}$$

where ϕ_n and $\hat{S}$ are as defined above.

K_n can be expressed as

$$K_n = \sum_{i | \tilde{\delta}_i \neq 0} \phi_n(i)(2\hat{S}(\tilde{T}_{(i-1)}) - 1)(\hat{S}(\tilde{T}_{(i-1)})\frac{1}{n-i+1} \tag{7}$$

Large values of K_n are significant for testing H_0 against H_1.

Aly et al (1994) showed that, under H_0,

$$n^{1/2}\phi_n \quad \stackrel{\mathcal{L}}{\to} \quad W(S(.)) \tag{8}$$

where W(.) is the standard Brownian Motion.

Using the continuous mapping theorem, we have

$$\begin{aligned}
n^{1/2}K_n &\stackrel{\mathcal{L}}{\to} \int_{0<s<t<1} (W(t)-W(s))\,ds\,dt \\
&= \int_0^1 (2t-1)W(t)\,dt \\
&= \int_0^1 W(t)\,d(t^2-t) \\
&\stackrel{\mathcal{L}}{\to} N(0,\sigma^2)
\end{aligned}$$

where

$$\begin{aligned}
\sigma^2 &= 2\int_{0\le s<t\le 1} s(2s-1)(2t-1)\,ds\,dt \\
&= \frac{1}{30}
\end{aligned} \tag{9}$$

Sun and Tiwari (1995) modified the statistic U_3 so that it can be used for censored data.

$$\begin{aligned}
V &= \int_0^\infty (F(2,t)-F(1,t))dH(t) \\
&= \int_0^\infty [S(t)]^2 d(\Lambda_2(t)-\Lambda_1(t))
\end{aligned} \tag{10}$$

A natural estimator of D is given by V_n where

$$V_n = \int_0^\infty [\hat{S}(t-)]^2 d(\hat{\Lambda}_2(t)-\hat{\Lambda}_1(t)) \tag{11}$$

where $\hat{S}(t)$ and $\hat{\Lambda}_j$ are as defined above. In the absence of censoring V_n reduces to the statistic U_3.

Then Sun and Tiwari (1995) proved the following theorem.

Theorem: Let $K(t) = 1 - S(t)S_C(t)$, $\tau_K = \sup\{t : K(t) < 1\}$. If $\int_0^{\tau_K} \frac{dH(t)}{S_C(t)} < \infty$, then

$$\sqrt{n}(V_n - V) \stackrel{\mathcal{L}}{\to} N(0,\sigma^2) \;\; asn \to \infty, \tag{12}$$

where

$$\begin{aligned}
\sigma^2 &= \int_0^\infty S^3(t)\frac{d(\Lambda_1(t)+\Lambda_2(t))}{S_C(t)} \\
&+ 4\int_0^\infty (\int_t^\infty S^2(u)d(\Lambda_2(u)-\Lambda_1(u)))^2\frac{d(\Lambda_1(t)+\Lambda_2(t))}{S(t)S_C(t)} \\
&- 4\int_0^\infty (\int_t^\infty S^2(u)d(\Lambda_2(u)-\Lambda_1(u)))\frac{d(\Lambda_1(t)+\Lambda_2(t))}{S_C(t)}
\end{aligned} \tag{13}$$

In particular, under H_0,

$$\sigma^2 = \int_0^\infty \frac{S^2(t)}{S_C(t)} dH(t). \tag{14}$$

A consistent estimator $\hat{\sigma}^2$ of σ^2 is obtained by replacing S ans S_C by their consistent Kaplan Meier-estimators.

When there is no censoring $\sigma^2 = \frac{1}{3}$. The test rejects for large values of the statistic . The approximate power of this test of size α is equal to $1 - \Phi(z_\alpha - \sqrt{n}V/\hat{\sigma})$, where z_α is the upper α percentile of standard normal distribution.

5 SIMULATION RESULTS

Given below are the results of a simulation study done for power comparisons of various tests for the uncensored case listed above.

Random samples were generated from absolutely continuous bivariate exponential (ACBVE) due to Block and Basu (1974) with density

$$f(x,y) = \begin{cases} \frac{\lambda\lambda_1(\lambda_2+\lambda_0)}{\lambda_1+\lambda_2} e^{-\lambda_1 x-(\lambda_2+\lambda_0)y} & \text{if} \quad x < y, \\ \frac{\lambda\lambda_2(\lambda_1+\lambda_0)}{\lambda_1+\lambda_2} e^{-\lambda_2 y-(\lambda_1+\lambda_0)x} & \text{if} \quad x > y, \end{cases}$$

where $(\lambda_0, \lambda_1, \lambda_2)$ are the parameters and $\lambda = \lambda_0 + \lambda_1 + \lambda_2$. The CSHR's are

$$\lambda_j(t) = \frac{\lambda_j \lambda}{\lambda_1 + \lambda_2}, \quad j = 1, 2.$$

Under $H_1 \quad \lambda_1 < \lambda_2$.

X and Y are independent if and only if $\lambda_0 = 0$. We set $\lambda_1 = 1$ and consider $\lambda_2 = 1.0, 1.4, 1.8, 2.2$ indicating larger and larger departures from H_0. The case $\lambda_2 = 1.0$ corresponds to the null hypothesis. $n = 100$ and there are 10000 replications.

	λ_2			
Test	1.0	1.4	1.8	2.2
D_1	3.76	41.98	82.53	96.83
D_2	4.85	47.71	86.98	98.14
U_2	5.09	44.60	83.96	96.92
U_3	4.79	43.06	80.54	95.42
U_4	4.99	42.67	80.81	95.77
Sign	4.39	49.50	88.29	98.66

Next we look at the censored case. The censoring distribution was exponential with parameters 1 and 3, respectively. We use asymptotic critical levels of 5 percent. Results are based on 5,000 replications.

TABLE 3.1

Observed levels and powers of K_n at an asymptotic level of 5 percent. The underlying distribution of (X, Y) is Block and Basu (1974) ACBVE with $\lambda_1 = 1$.

(a) CENSORED (EXP(1))

	n=50		n=100	
λ_2	$\lambda_0 = 0$	$\lambda_0 = 1$	$\lambda_0 = 0$	$\lambda_0 = 1$
1.0	.0218	.0312	.0360	.0376
1.5	.1864	.2192	.3732	.4302
2.0	.4482	.5080	.7862	.8342
2.5	.6928	.7414	.9546	.9704

(b) CENSORED (EXP(3))

	n=50		n=100	
λ_2	$\lambda_0 = 0$	$\lambda_0 = 1$	$\lambda_0 = 0$	$\lambda_0 = 1$
1.0	.0048	.0124	.0084	.0172
1.5	.0342	.0762	.1012	.1834
2.0	.1202	.1986	.3216	.4860
2.5	.2496	.3774	.5862	.7344

From the table it is clear that the asymptotic critical levels give conservative tests for the censored case, with the effect increasing as the censoring becomes more severe. There is slight effect on the levels or the power due to lack of independence of X and Y in the presence of censoring. The results are comparable with the test proposed by Aly et al (1994) for the lightly censored case.

Remarks

(i) The various tests are consistent against their intended alternatives.

(ii We can also use these tests for the hypothesis $\lambda_1(t) = \lambda_2(t)$ against the alternative that cause-specific hazards are ordered.

(iii) The tests are distribution-free under H_0. The null distribution of the tests U_3 and U_4 is same as in the case of independent latent failures (see, Deshpande and Dewan (2003)).

(iv) It is important to note that T and δ continue to be independent under the null hypothesis of bivariate symmetry . Hence the conclusions of Lemma 1 in the review paper hold under H_0.

(v) The statistic U_2 puts more weight on the middle observations and is less sensitive to the observations in the beginning and the end of the experiment. On the other hand, U_3 puts more weight to later observations and U_4 puts higher weight to observations in the beginning.

(vi) Deshpande and Dewan (2000) proposed tests for testing bivariate symmetry aginst dispersive asymmetry. Here the alternatives can be expressed in terms of ordering of sub-survival functions and ordering of sub-distributions of the maximum of observations and δ . The statistic is a linear combination of two statistics, the first one is a U-statistic based on minimum and δ and the other one is a

U-statistic based on maximum and δ. The one based on minimum and δ is the statistic U_4

(vii) The statistics U_2, U_3, U_4 are all linear combinations of the sign statistic and the Wilcoxon-signed rank type statistic.

(viii) Tests proposed by Aly, Kochar and Mckeague (1994) can be extended to the case of multiple risks in which any two of the cause-specific risks are to be compared. The statistic can be modified to test dominance of one risk over the other in a specified interval.

6 Test for independence of T and δ

The nature of dependence between T and δ is crucial and useful in modelling competing risks data via sub-distribution/subsurvival functions. If T and δ are independent then $S_i(t) = pr(\delta = i)S(t)$, allowing the study of the failure times and the causes (risks) of failure separately. The hypothesis of equality of incidence functions or that of cause-specific hazard rates reduces to testing whether $pr(\delta = 1) = pr(\delta = 0) = 1/2$. This simplifies the study of competing risks to a great extent.

Dewan, Deshpande and Kulathinal (2004) studied the properties of the conditional probability functions

$$\Phi_i(t) = pr(\delta = i \mid T \geq t) = S_i(t)/S(t),\ i = 1, 2$$

and

$$\Phi_i^*(t) = pr(\delta = i \mid T < t) = F_i(t)/H(t),\ i = 1, 2.$$

They observed

(i) T and δ are independent iff $\Phi_1(t) = P[\delta = 1]$ or $\Phi_2^*(t) = 1 - P[\delta = 1]$

(ii) T and δ are PQD iff $\Phi_1(t) \geq P[\delta = 1]$ or $\Phi_2^*(t) \geq 1 - P[\delta = 1]$

(iii) δ is Right Tail Increasing in T iff $\Phi_1(t)$ is increasing in t.

(iv) δ is Left Tail Decreasing in T iff $\Phi_2^*(t)$ is decreasing in t.

They considered the problem of testing $H_0 : T\ and\ \delta$ are independent which is equivalent to

$$H_0 : \Phi_1(t) \text{ is a constant}$$

against various alternative hypotheses which characterise the properties of $\Phi_1(t)$ and $\Phi_0^*(t)$:

$$\begin{aligned} H_1 &: \Phi_1(t) \text{ is not a constant} \\ H_2 &: \Phi_1(t) \geq P[\delta = 1] \text{ for all } t \text{ with strict inequality for some } t \\ H_3 &: \Phi_1(t) \text{ is a monotone nondecreasing function of } t \\ H_4 &: \Phi_0^*(t) \text{ is a monotone nonincreasing function of } t. \end{aligned}$$

A test based on the concept of concordance and discordance was proposed for testing H_0 against H_1. Actually a one-sided version of the test was seen to be consistent against H_2 . Two tests were proposed to test H_0 against H_2. A test using U-statistic was proposed for testing H_0 against H_3 and on the same lines a test was proposed for testing H_0 against H_4. Note that there is no relationship between H_3 and H_4 but both imply H_2. Some of the test statistics considered are already in the literature but in other contexts.

6.1 Testing H_0 against H_1

Kendall's τ is used as a test statistic for a very general alternative of non-independence. A pair (T_i, δ_i) and (T_j, δ_j) is a concordant pair if $T_i > T_j,\ \delta_i = 1, \delta_j = 0$ or $T_i < T_j,\ \delta_i = 0, \delta_j = 1$ and is a discordant pair if $T_i > T_j,\ \delta_i = 0, \delta_j = 1$ or $T_i < T_j,\ \delta_i = 1, \delta_j = 0$. Define the kernel

$$\psi_k(T_i, \delta_i, T_j, \delta_j) = \begin{cases} 1 & \text{if } T_i > T_j,\ \delta_i = 1, \delta_j = 0 \\ & \text{or } T_i < T_j,\ \delta_i = 0, \delta_j = 1 \\ -1 & \text{if } T_i > T_j,\ \delta_i = 0, \delta_j = 1 \\ & \text{or } T_i < T_j,\ \delta_i = 1, \delta_j = 0 \\ 0 & \text{otherwise.} \end{cases}$$

Note that when both δ_i and δ_j are 1 or 0, then $\delta_i - \delta_j = 0$. The corresponding U-statistic is given by

$$U_k = \binom{n}{2}^{-1} \sum_{1 \le i < j \le n} \psi_1(T_i, \delta_i, T_j, \delta_j).$$

It is seen that $E(U_k) \ge 0$ under H_2. Hence, a one-sided test based on U_k can be used to test $\Phi_1(t) \ge \phi_1(0)$ for all t also.

$\binom{n}{2} U_k$ can be written as

$$\binom{n}{2} U_k = \sum_{j=1}^{n} (2R_j - n - 1)\delta_j = \sum_{j=1}^{n} (2j - n - 1) W_j = \sum_{j=1}^{n} a_j W_j \qquad (1)$$

where $a_j = 2j - n - 1$.

This statistic was introduced for the first time in Deshpande and Sengupta (1995) for proportionality of cause specific hazard rates with independent competing risks. For details see Deshpande and Dewan (2003) and Dewan et al (2004).

6.2 Testing H_0 against H_2

$H_2 : \Phi_1(t) \ge \Phi_1(0)$ which is equivalent to $\Phi_0^*(t) \ge \Phi_0^*(0)$. **A. Test based on $\Phi_1(t)$**

Consider

$$\Delta_2(S_1, S) = \int_0^\infty [S_1(t) - \phi S(t)] dF(t) = pr(T_2 > T_1, \delta_2 = 1) - \phi/2.$$

Under H_0, $S_1(t)/S(t) = pr(\delta = 1)$. This implies that $\Delta_2(S_1, S) = 0$. Under H_2, $S_1(t) \ge \Phi_1(0) S(t)$ and hence $\Delta_2(S_1, S) \ge 0$. Define the symmetric kernel

$$\psi_2(T_i, \delta_i, T_j, \delta_j) = \begin{cases} 1 & \text{if } T_j > T_i,\ \delta_j = 1 \\ & \text{or if } T_i > T_j,\ \delta_i = 1 \\ 0 & \text{otherwise.} \end{cases}$$

The corresponding U-statistic estimator is given by

$$U_S = \binom{n}{2}^{-1} \sum_{1 \le i < j \le n} \psi_2(T_i, \delta_i, T_j, \delta_j). \qquad (2)$$

It can be shown that

$$\binom{n}{2} U_S = \sum_{j=1}^{n} (R_j - 1)\delta_j = \sum_{j=1}^{n} (j-1) W_j. \tag{3}$$

The above statistic is proposed in equation (2.6) by Bagai *et al.* (1989) for testing the equality of failure rates of two independent competing risks and is same as the statistic U_4.

We can derive another test for H_0 versus H_2 using the fact that $\Phi_0^*(t) \geq \Phi_0^*(0)$ under H_2.

6.3 Testing H_0 against H_3

Note that $\Phi_1(t) \uparrow t$ is equivalent to $\Phi_1(t_1) \leq \Phi_1(t_2)$, whenever $t_1 \leq t_2$. This gives $\gamma(t_1, t_2) = S_1(t_2)S(t_1) - S_1(t_1)S(t_2) \geq 0, t_1 \leq t_2$ with strict inequality for some (t_1, t_2). Define

$$\begin{aligned} \Delta_3(S_1, S) &= \int\int_{t_1 \leq t_2} \gamma(t_1, t_2) dF_1(t_1) dF_1(t_2) \\ &= \int_0^\infty [S_1^2(t) - \phi^2/2] S(t) dF_1(t). \end{aligned} \tag{4}$$

Under H_0, $S_1(t)/S(t) = \phi$. This implies that $\Delta_3(S_1, S) = 0$. Under H_3, $\Delta_3(S_1, S) \geq 0$.

Define the kernel

$$\psi_3^*(T_i, \delta_i, T_j, \delta_j, T_k, \delta_k, T_l, \delta_l) = \begin{cases} 1 & \text{if } T_k > T_j > T_l > T_i, \\ & \delta_i = \delta_j = \delta_k = 1, \delta_l = 0 \\ -1 & \text{if } T_l > T_j > T_k > T_i, \\ & \delta_i = \delta_j = \delta_k = 1, \delta_l = 0 \\ 0 & \text{otherwise.} \end{cases}$$

Then the U-statistic corresponding to $\Delta_3(S_1, S)$ is given by

$$U_R = \binom{n}{4}^{-1} \sum_{1 \leq i_1 < i_2 < i_3 < i_4 \leq n} \psi_3(T_{i_1}, \delta_{i_1}, T_{i_2}, \delta_{i_2}, T_{i_3}, \delta_{i_3}, T_{i_4}, \delta_{i_4}),$$

where ψ_3 is the symmetric version corresponding to ψ_3^*.

Note that $E(\psi_3^*(T_i, \delta_i, T_j, \delta_j, T_k, \delta_k, T_l, \delta_l)) = \Delta_3(S_1, S)$ and the expectation of the symmetric kernel is $24\Delta_3(S_1, S)$ due to the possible combinations required to obtain the symmetric kernel. Hence, $E(U_R) = 24\Delta_3(S_1, S)$. Under H_0, $E(U_3) = 0$ and under H_3, $E(U_R) \geq 0$.

Theorem 1 *As n tends to ∞, under H_0, $n^{1/2}U_R$ converges in distribution to $N(0, \sigma_3^2)$, where $\sigma_3^2 = (96/35)\phi^5(1-\phi)$.*

The null hypothesis is rejected for large values of $n^{1/2}U_R/\hat{\sigma_3}$ where $\hat{\sigma_3^2} = (96/35)\hat{\phi}^5(1-\hat{\phi})$.

Tests proposed above will help in discriminating between the constant or proportional warning-constant inspection and random signs censoring models due to

Cooke (1996) and also to determine whether the corresponding mode of failure becomes more likely with increasing age.

Based on similar considerations Dewan et al (2004) proposed a test for testing H_0 against H_4.

For modelling the competing risks data in terms of (T, δ), it is of prime importance to check whether T and δ are independent.The above tests are simple and perform satisfactorily in distinguishing between the hypotheses. All tests are typically consistent against larger alternatives than the one for which they are proposed . The tests are “almost” distribution free in the sense that their null distribution depends only on the parameter $P(\delta = 1)$ which can be estimated consistently. If the hypothesis of independence is accepted then one can simplify the model and study the failure time and cause of failure separately. If the hypothesis is rejected then a suitable model under specific dependence between T and δ in terms of the incidence functions is needed.

References

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

Aras, G. and Deshpande, J.V. (1992). Statistical analysis of dependent competing risks. *Stats and Decisions*, **10**, 323-336.

Block,H.W. and Basu, A.P. (1974). A continuous bivariate exponential distribution. *J. Amer. Statist. Assoc.*, **69**, 1031-1037.

Cooke, R. M. (1996). The design of reliability databases, part II. *Reliability engineering and system safety* **51**, 209-223.

Crowder, M.J. (1991) . On the identifiability crisis in competing risks analysis. *Scanadanavian J. Statist.*, **18**,223-233.

Crowder, M.J. (1993) . Identifiability crisis in competing risks analysis. *Int. Statist. Rev.* ,**62**, 379-391.

Crowder, M. J. (2001). *Classical competing risks.* Chapman and Hall/CRC, London.

Deshpande, J. V. (1990). A test for bivariate symmetry of dependent competing risks. *Biometrical Journal*, **32**, 736-746.

Deshpande, J.V. and Dewan, I. (2000). Testing bivariate symmetry against dispersive asymmetry. *Journal Indian Statist. Assoc.*, ,

Deshpande, J.V. and Dewan, I. (2003). A review of some statistical tests under independent competing risks model. *Preprint.*

Dewan, I. and Kulathinal, S.B. (2003). Parametric models for subsurvival functions. Preprint.

Dewan I., Deshpande J.V. and Kulathinal S.B. (2004). On testing dependence between time to failure and cause of failure via conditional probabilities. *Scandinavian J. Statistics* (to appear).

Dewan, I. and Kochar, S.C. (200)

Gumbel , E.J. (1960). Bivariate exponential distribution. *Journal Amer. Statist. Assoc.*, **55**, 698-707.

Hettmansperger , T.P. (1984). *Statistical Inference Based On Ranks* . John Wiley, New York.

Kalbfleisch, J.D. and Prentice, R.L. (1980). *The statistical analysis of failure time data.* John Wiley, New York.

Kalbfleisch, J.D. and Prentice, R.L. (2002). *The statistical analysis of failure time data.* Second Edition, John Wiley, New Jersey.

Kochar, S.C. (1995). A review of some distribution-free tests for the equality of cause specific hazard rates. *Analysis of Censored Data* , IMS Lecture Notes, **27**, 147 - 162.

Sen, P.K. (1979). Nonparametric tests for interchangeability under competing risks. In *Contributions to Statistics, Jaroslav Hajek Memorial Volume*, 211-228.Reidel, Dordrecht.

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 Lecture Notes, **27**, 255 - 270.

Tsiatis, A. (1975). A nonidentifiability aspect of the problem of competing risks. *Proc. Natl. Acad. Sci. U.S.A.*, **72**, 20-22.

Yip, P. and Lam, K.F. (1992). A class of nonparametric tests for the equality of failure rates in a competing risks model. *Commun. Statist. Theory Methods*, **21** , 2541-2556.