On Testing Dependence between Time to Failure and Cause of Failure via Conditional Probabilities

ISHA DEWAN

Indian Statistical Institute

J. V. DESHPANDE

University of Pune

S. B. KULATHINAL

National Public Health Institute

ABSTRACT. Dependence structures between the failure time and the cause of failure are expressed in terms of the monotonicity properties of the conditional probabilities involving the cause of failure and the failure time. These properties of the conditional probabilities are used for testing four types of departures from the independence of the failure time and the cause of failure and tests based on *U*-statistics are proposed. In the process, a concept of concordance and discordance between a continuous and a binary variable is introduced to propose a statistical test. The proposed tests are applied to two illustrative applications.

Key words: competing risks, conditional probability, dependence structures, subsurvival functions, *U*-statistics

1. Introduction

Consider a situation where a unit can fail due to one of two competing causes. Let T_1 and T_2 denote the latent lifetimes of the unit under the two causes. The competing risk data available are the failure time T of the unit, which is the minimum of (T_1, T_2) and the cause of failure indicator δ, which is equal to 1 if $T = T_1$ and is 0 if $T = T_2$. These data are right censored data where each latent lifetime acts as a censoring variable for the other and, unlike in censoring, the interest lies in both the causes and hence in both the lifetimes. One concentrates on different aspects of the situation by assuming appropriate dependence structures (i) for the two latent lifetimes (T_1, T_2) and (ii) for the random variables (T, δ). The joint distribution of (T, δ) is defined here by the subsurvival functions, $S_i(t) = pr(T \geq t, \delta = i)$, $i = 0, 1$. The survival function of T is defined by $S(t) = pr(T \geq t) = S_0(t) + S_1(t)$. Throughout this paper, we assume that the subsurvival functions are continuous with $f_i(t)$, $i = 0, 1$, as the subdensity functions and $f(t) = f_0(t) + f_1(t)$ as the density of T. The cause-specific hazard rate for cause i is defined as $h_i(t) = f_i(t)/S(t)$ and the crude hazard rate for cause i is defined as $r_i(t) = f_i(t)/S_i(t)$. The hazard rate of T is $h(t) = f(t)/S(t) = h_1(t) + h_0(t)$.

The problem of identifiability in modelling the competing risks data in terms of the latent lifetimes is well known. The distributions of the latent lifetimes are identifiable under the assumption of independence of the competing causes and also under some weaker conditions of non-informative censoring, see Kalbfleisch & Prentice (2002). There has been an ongoing debate for many years about the use of the models in terms of latent lifetimes and the models in terms of (T, δ), see Prentice *et al.* (1978), Larson & Dinse (1985), Davis & Lawrance (1989), Deshpande (1990), Aras & Deshpande (1992), Gasbarra & Karia (2000), Crowder (2001), Kulathinal & Gasbarra (2002) and others. The problem of identifiability does not arise if the modelling of the competing risk data is done in terms of the subsurvival functions of (T, δ) or related quantities like cause-specific hazard rates and crude hazard rates. The nature of dependence between T and δ is crucial and useful in such modelling. 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.

In this paper, we study the properties of the conditional probability functions

$$\Phi_i(t) = pr(\delta = i \mid T \geq t) = \frac{S_i(t)}{S(t)}, \quad i = 0, 1$$

and

$$\Phi_i^*(t) = pr(\delta = i \mid T < t) = \frac{F_i(t)}{F(t)}, \quad i = 0, 1,$$

where $F_i(t) = pr(T < t, \delta = i)$, $i = 0, 1$ are the incidence functions or subdistribution functions and $F(t) = pr(T < t) = F_0(t) + F_1(t)$ is the distribution function of T.

We also study various kinds of dependence between T and δ via these probabilities. The motivation for studying these probabilities, partly, comes from Cooke (1996), who studied failure and preventive maintenance in a censoring setting with the interest in the distribution of the latent failure time which would have been observed in the absence of preventive maintenance. In the next section, the models considered in Cooke (1996) are reviewed and the properties of $\Phi_1(t)$ are stated for illustration. The results of this paper are especially of interest in reliability, but examples arise in many other fields where the conditional probabilities of the type $\Phi_i(t)$ and $\Phi_i^*(t)$ are of primary importance. In clinical trials carried out to study the performance of an intrauterine device where termination of the device could be due to several reasons such as pregnancy, expulsion, bleeding and pain, it is often of interest to know the chances of termination due to a specific reason given that the device was intact for some specified period. Also, in epidemiological follow-up studies the probability of occurrence of an event given that the age of a person is above a certain limit is of interest. In such situations, conditional probabilities are expected to vary with time. Hence the applicability of the results of this paper is quite wide.

In section 2, we define dependence structures between T and δ in terms of the shapes of the conditional probability functions $\Phi_1(t)$ and $\Phi_0^*(t)$. Also, their relation to ordering between the cause-specific hazard rates and crude hazard rates are studied. In section 3, we consider the problem of testing H_0 : T and δ are independent which is equivalent to

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

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

$$H_1 : \Phi_1(t) \text{ is not a constant}$$
$$H_2 : \Phi_1(t) \geq \phi \text{ for all } t \text{ with strict inequality for some } t$$
$$H_3 : \Phi_1(t) \text{ is a monotone non-decreasing function of } t$$
$$H_4 : \Phi_0^*(t) \text{ is a monotone non-increasing function of } t,$$

where $\phi = pr(\delta = 1) = 1 - pr(\delta = 0)$. The properties described in section 2 motivate the above alternative hypotheses. A test based on the concept of concordance and discordance is proposed for testing H_0 against H_1. Actually a one-sided version of the test is seen to be consistent against H_2. Two tests are proposed to test H_0 against H_2. A test using U-statistic is proposed for testing H_0 against H_3 and on the same lines a test is 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 here are already in the literature but in other contexts. In section 4, relative efficiencies of these tests are studied and in section 5 the tests are applied to two real data sets. To the best of our knowledge, tests based on the conditional probability functions of the type $\Phi_i(t)$ and $\Phi_i^*(t)$, which are useful in modelling the competing risks data in terms of (T, δ), are proposed and studied in detail here for the first time.

2. Properties of $\Phi_1(t)$ and $\Phi_0^*(t)$

It is obvious that the independence of T and δ is equivalent to constancy of $\Phi_1(t)$ and is also equivalent to constancy of $\Phi_0^*(t)$. Many popular bivariate parametric distributions used in survival analysis have constant $\Phi_1(t)$ and $\Phi_0^*(t)$, for example, Block & Basu (1974), Farlie–Gumbel–Morgenstern bivariate exponential distribution, Gumbel type A distribution. However, in many practical situations, this is not the case. We review the models considered in Cooke (1996) in the light of the conditional probability functions $\Phi_1(t)$ and $\Phi_0^*(t)$. It should be noted that the function $\Phi(t)$ used in Cooke (1996) is equivalent to $1 - \Phi_1(t)$ defined in this paper. In the following models, the two competing causes are the actual cause of failure of a unit and the censoring caused by the warning. The failure time of the unit is denoted as T_1 and the censoring variable defined according to the warning emitted by the unit before failure is denoted as T_2.

2.1. *Random signs censoring*

A random signs censoring, also known as an age-dependent censoring, is a model in which the lifetime of a unit T_1 is censored by $T_2 = T_1 - W\eta$, where $0 < W < T_1$ is a warning emitted by the unit before its failure, and η is a random variable taking values $\{-1,1\}$ and is independent of T_1. Hence $\eta = 1$ would lead to the censoring of the lifetime at $T_1 - W$ and $\eta = -1$ will lead to the observation of complete lifetime T_1. Assume that T_1 has exponential distribution with parameter λ. In this case, $P(T_2 > t, T_2 < T_1) = P(T_1 - W > t, \eta = 1)$ and $P(T_1 > t, T_1 < T_2) = P(T_1 > t)P(\eta = -1)$. This gives $\Phi_1(t) = P(T_1 > t, \eta = -1)/P(T_1 - W\eta > t, T_1 > t)$. When $W = aT_1$, $0 < a < 1$,

$$\Phi_1(t) = \left(1 + \frac{p}{1-p}\exp\{-\lambda t(a/(1-a))\}\right)^{-1}$$

where $p = P(\eta = 1) = 1 - P(\eta = -1)$, leading to the increasing nature of $\Phi_1(t)$ in t.

2.2. *Constant warning-constant inspection*

In a constant warning-constant inspection model, a warning is emitted at time $T_1 - d$ before the unit fails, where $d(<1)$ is a constant. Assume that T_1 has exponential distribution with parameter λ. Inspections are made at regular intervals I. Here, T_1 is censored by T_2, with $T_2 = I_{(i)}$ if $I_{(i-1)} < T_1 - d < I_{(i)} < T_1$ and T_1 is observed to fail at time T_1 if no inspection occurs in the interval $[T_1 - d, T_1]$. Take $I = 1$. In this case,

$$\Phi_1(i) = 1 - \frac{\exp\{\lambda d\} - 1}{\exp\{\lambda\} - 1},$$

which is independent of i.

2.3. Proportional warning-constant inspection

A proportional warning-constant inspection is similar to the constant warning-constant inspection model except that the warning is emitted at time T_1/η if the component fails at T_1; where η is a constant. In this case,

$$\Phi_1(i) = \frac{1 - \exp\{i\lambda(\eta - 1)\}}{(\exp\{\lambda\} - 1)} + \frac{1}{(\exp\{\lambda\eta\} - 1)}$$

which is clearly a decreasing function of i.

Thus, the monotonicity of $\Phi_1(t)$ helps in choosing the appropriate model.

2.4. Hazard rate ordering and ageing

The conditional probability, cause-specific hazard rate and crude hazard rate are functionally related by the identity $h_i(t) = \Phi_i(t)r_i(t)$, $i = 0, 1$.

Theorem 1

The conditional probability function $\Phi_1(t) \uparrow t$ is equivalent to $r_1(t) \leq h(t) \leq r_0(t)$ for all t, and also to $h_1(t) \leq \Phi_1(t)h(t)$ and $h_0(t) \geq (1 - \Phi_1(t))h(t)$.

The proof follows by using the fact that the derivative of $\Phi_1(t)$ is non-negative and the derivative of $1 - \Phi_1(t)$ is non-positive being decreasing function of t.

Thus, $\Phi_1(t)$ is increasing is equivalent to the fact that the overall failure rate is larger than the failure rate given that the failure is due to cause 1 and is smaller than the failure rate given that the failure is due to cause 2. It is also equivalent to saying that $S_1(t)/S_0(t)$ is non-decreasing in t. The above theorem also implies that $h_1(t)/h_0(t) \leq \Phi_1(t)/(1 - \Phi_1(t))$. This puts functional bounds on the relative rate of ageing of two risks, see Sengupta & Deshpande (1994) for definitions of relative ageing.

Another interesting result stated below connects the monotonicity of $\Phi_1(t)$ with the ordering between two survival functions.

Theorem 2

The conditional probability function $\Phi_1(t) \geq \phi$ for all t if and only if the survival function of T given $\delta = 1$ is larger than that of T given $\delta = 0$, that is, $S_1(t)/\phi \geq S_0(t)/(1 - \phi)$.

The proof follows by noting that $\Phi_1(t) \geq \phi$ is equivalent to $S_1(t)/\phi \geq S(t)$ and $S_0(t)/(1-\phi) \leq S(t)$.

It is important to note that the crude hazard rates $r_1(t)$ and $r_0(t)$ are the hazard rates of the distributions given by $S_1(t)/\phi$ and $S_0(t)/(1 - \phi)$, respectively. These distributions are called conditional subsurvival functions by Cooke (1996), and in fact theorem 2 gives the properties implied by the random signs censoring model of Cooke (1996).

Under the proportional hazards model, $h_1(t) = \phi h(t)$. This is equivalent to independence of T and δ and hence $\Phi_1(t) = \phi$, for all $t > 0$. It is easy to see that $h_1(t) \geq \phi h(t)$ implies $\Phi_1(t) \geq \phi$, for all t. Hence, the tests proposed in the next section can be used to test the proportionality of the two cause-specific hazard rates also. When $\phi \geq 1/2$, $S_1(t) \geq S_0(t)$ for all t and this means that there is stochastic dominance between the two incidence functions as well as the conditional distributions.

It is interesting and also useful to express the cause-specific hazard rate in terms of $\Phi_1(t)$. This enables one to study the ageing through the properties of $\Phi_1(t)$.

Theorem 3

If $\Phi_1(t)$ is monotone increasing and concave then $h_1(t)$ is an increasing function of t, provided $h(t)$ is increasing.

Proof. From the definitions of $\Phi_1(t)$ and $h_1(t)$, it is easy to note that $h_1(t) = -\Phi_1'(t) + \Phi_1(t)h(t)$, where $\Phi_1'(t)$ is the first derivative of $\Phi_1(t)$ with respect to t. Hence the result.

Further, let $r_i^*(t)$ and $h_i^*(t)$ denote crude and cause-specific reverse hazard rates, which are defined as

$$r_i^*(t) = \frac{f_i(t)}{F_i(t)} \text{ and } h_i^*(t) = \frac{f_i(t)}{F(t)}.$$

All the above results hold true between these reverse hazards and $\Phi_0^*(t)$. As the results are quite similar, the details are not given here. The above results bring out the fact that many important kinds of dependence between T and δ can be expressed in terms of various shapes of $\Phi_1(t)$ and $\Phi_0^*(t)$. Note that $\Phi_1(t)$ increasing in t does not necessarily imply that $\Phi_0^*(t)$ is decreasing in t and vice versa. These properties motivate various alternative hypotheses considered in the next section.

3. Test statistics and their distributions

Let (T_i, δ_i), $i = 1, 2, \ldots, n$, be the competing risk data obtained from n independent and identical units.

3.1. Testing H_0 against H_1

As defined earlier

$$H_0 : \Phi_1(t) \text{ is a constant}$$
$$H_1 : \Phi_1(t) \text{ is not a constant}$$

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

$$\psi_1(T_i, \delta_i, T_j, \delta_j) = \begin{cases} 1 & \text{if } T_i > T_j,\ \delta_i = 1,\ \delta_j = 0 \\ & \text{or } T_i < T_j,\ \delta_i = 0,\ \delta_j = 1 \\ -1 & \text{if } T_i > T_j,\ \delta_i = 0,\ \delta_j = 1 \\ & \text{or } T_i < T_j,\ \delta_i = 1,\ \delta_j = 0 \\ 0 & \text{otherwise.} \end{cases}$$

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

$$U_1 = \binom{n}{2}^{-1} \sum_{1 \le i < j \le n} \psi_1(T_i, \delta_i, T_j, \delta_j).$$

Note that

$$E(U_1) = E(\psi_1(T_i, \delta_i, T_j, \delta_j)) = 2\phi + 4 \int_0^\infty S(t) dS_1(t).$$

It is seen that $E(U_1) \geq 0$ under H_2. Hence, a one-sided test based on U_1 can be used to test $\Phi_1(t) \geq \phi$ for all t also.

It is easy to write the statistic U_1 as a function of ranks. Let R_j be the rank of T_j. Let $T_{(1)} < \cdots < T_{(n)}$ be the ordered T_i's. Let

$$W_j = \begin{cases} 1 & \text{if } T_{(j)} \text{ corresponds to } \delta = 1 \\ 0 & \text{otherwise.} \end{cases}$$

Then $V_1 = \binom{n}{2} U_1$ can be written as

$$V_1 = \sum_{j=1}^{n} (2R_j - n - 1)\delta_j = \sum_{j=1}^{n} (2j - n - 1) W_j = \sum_{j=1}^{n} a_j W_j \tag{1}$$

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

This statistic was introduced for the first time in Deshpande & Sengupta, (1995) for proportionality of cause-specific hazard rates with independent competing risks. The statistic given in equation (2.3) in Dykstra *et al.* (1996), page 214 in a different context, is $-U_1$ and the correct variance of V_n is $(1/3)n(n^2 - 1)\theta(1 - \theta)$ and not the one given on page 215. The null distribution of V_1 can be found from its moment generating function. Note that under H_0, $T_1, \ldots, T_n$ and $\delta_1, \ldots, \delta_n$ are independent. Hence, under H_0, $W_1, \ldots, W_n$ are independent and identically distributed with $pr(W_i = 1) = \phi$ and $pr(W_i = 0) = 1 - \phi$. From here we obtain that the moment generating function of V_1, under H_0, is given by

$$M(t) = \prod_{j=1}^{n} [\phi \, \exp\{t(2j - n - 1)\} + (1 - \phi)].$$

Hence the null distribution of V_1 depends on the unknown ϕ even under H_0. For large n, we can estimate ϕ consistently by $\hat{\phi} = n^{-1} \sum_{i=1}^{n} I(\delta_i = 1)$. Under H_0,

$$E(U_1) = 0 \text{ and } \mathrm{Var}(U_1) = \frac{4(n+1)}{3n(n-1)} \phi(1 - \phi).$$

Note that $E(U_1) \neq 0$ under H_1. From the results on U-statistics it follows that U_1 has an asymptotic normal distribution for large n (Serfling, 1980).

Theorem 4

As n tends to ∞, under H_0, $n^{1/2} U_1$ converges in distribution to $N(0, \sigma_1^2)$ where $\sigma_1^2 = (4/3)\phi(1 - \phi)$.

A consistent estimator of variance is $\hat{\sigma}_1^2 = (4/3)\hat{\phi}(1 - \hat{\phi})$. A test procedure for testing H_0 against H_1 is then: reject H_0 at $100\alpha\%$ level of significance if $|n^{1/2} U_1 / \hat{\sigma}_1|$ is larger than $z_{1-\alpha/2}$, the cut-off point of standard normal distribution.

3.2. Testing H_0 against H_2

Recall that $H_2 : \Phi_1(t) \geq \phi$ which is equivalent to $\Phi_0^*(t) \geq 1 - \phi$. It is clear that a one-sided test based on U_1 can be used for testing H_0 against H_2 as it is based on concordance and discordance principle and the number of concordances are expected to be larger than the number of discordances under H_2. A test procedure for testing H_0 against H_2 is then: reject H_0 at $100\alpha\%$ level of significance if $n^{1/2} U_1 / \hat{\sigma}_1$ is larger than $z_{1-\alpha}$, the cut-off point of standard normal distribution. Two more tests are given below using $\Phi_1(t)$ and $\Phi_0^*(t)$ for testing H_0 against H_2.

3.2.1. Test based on $\Phi_1(t)$
Consider

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

Under H_0, $S_1(t)/S(t) = \phi = pr(\delta = 1)$. This implies that $\Delta_2(S_1, S) = 0$. Under H_2, $S_1(t) \geq \phi S(t)$ and hence $\Delta_2(S_1, S) \geq 0$. 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

$$U_2 = \binom{n}{2}^{-1} \sum_{1 \leq i < j \leq n} \psi_2(T_i, \delta_i, T_j, \delta_j). \tag{2}$$

It can be shown that

$$\binom{n}{2} U_2 = \sum_{j=1}^{n} (R_j - 1)\delta_j = \sum_{j=1}^{n} (j-1) W_j. \tag{3}$$

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.

Theorem 5
As n tends to ∞, under H_0, $n^{1/2}(U_2 - \phi)$ converges in distribution to $N(0, \sigma_2^2)$, where $\sigma_2^2 = (4/3)\phi(1-\phi)$.

3.2.2. Test based on $\Phi_0^*(t)$
H_2 is also equivalent to $H_2 : \Phi_0^*(t) \geq \phi_0$ for all t with strict inequality for some t, where $\phi_0 = 1 - \phi$. As in the earlier section, we have

Theorem 6
As n tends to ∞, $n^{1/2}(U_2^* - \phi_0)$ converges in distribution to $N(0, \sigma_2^{*2})$, where

$$\binom{n}{2} U_2^* = \frac{n(n-1)}{2} - \sum_{j=1}^{n} (n-j) W_j \tag{4}$$

and $\sigma_2^{*2} = (4/3)\phi_0(1-\phi_0)$.

The consistent estimators of variances σ_2^2 and σ_2^{*2} can be found by replacing ϕ by $\hat{\phi}$. We reject the null hypothesis for large values of the standardised versions of the statistics. From equations (1), (3) and (4), it follows that $U_1 = U_2 + U_2^* - 1$.

3.3. Testing H_0 against H_3

Recall that $H_3 : \Phi_1(t) \uparrow t$. 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

$$\Delta_3(S_1, S) = \int \int_{t_1 \le t_2} \gamma(t_1, t_2) dF_1(t_1) dF_1(t_2)$$
$$= \int_0^\infty [S_1^2(t) - \phi^2/2] S(t) dF_1(t). \quad (5)$$

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

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

$$U_3 = \binom{n}{4}^{-1} \sum_{1 \le i_1 < i_2 < i_3 < i_4 \le n} \psi_3(T_{i_1}, \delta_{i_1}, T_{i_2}, \delta_{i_2}, T_{i_3}, \delta_{i_3}, T_{i_4}, \delta_{i_4}),$$

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, $E(U_3) = 24\Delta_3(S_1, S)$. Under H_0, $E(U_3) = 0$ and under H_3, $E(U_3) \ge 0$. Let T's corresponding to 1's be called X's and those corresponding to 0's be called Y's. Then the number of X's is $n_1 = \sum_{i=1}^n \delta_i$, and there are $n_2 = n - n_1 Y$'s. Let $R_{(i)}(S_{(j)})$ be the rank of $X_{(i)}(Y_{(j)})$ be the ith (jth) ordered statistic in the $X(Y)$ sample in the combined arrangement of $n_1 X$'s and $n_2 Y$'s (in fact $n T$'s). Hence

$$\binom{n}{4} U_3 = \sum_{j=1}^{n_2} (S_{(j)} - j) \binom{n_1 + j - S_{(j)}}{2} - \sum_{j=1}^{n_2} \binom{S_{(j)} - j}{3}.$$

It is interesting to note that in terms of X's and Y's the above statistic is the same as that proposed by Kochar (1979) for testing equality of failure rates, the only difference being that the number of X's and Y's is random.

Theorem 7

As n tends to ∞, under H_0, $n^{1/2} U_3$ 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 $n^{1/2} U_3 / \hat{\sigma}_3$ where $\hat{\sigma}_3^2 = (96/35)\hat{\phi}^5(1 - \hat{\phi})$.

Tests proposed in this section will help in discriminating between the constant or proportional warning-constant inspection and random signs censoring models and also to determine whether the corresponding mode of failure becomes more likely with increasing age.

3.4. *Testing H_0 against H_4*

Recall that $H_4 : \Phi_0^*(t) \downarrow t$. $\Phi_0^*(t) \downarrow t$ is equivalent to $\Phi_0^*(t_1) \ge \Phi_0^*(t_2)$, whenever $t_1 \le t_2$. This gives $F_0(t_1)F(t_2) - F_0(t_2)F(t_1) \ge 0$, $t_1 \le t_2$ with strict inequality for some (t_1, t_2). Define

$$\Delta_4(F, F_0) = \int \int_{t_1 \le t_2} [F_0(t_1)F(t_2) - F_0(t_2)F(t_1)] dF_0(t_1) dF_0(t_2)$$
$$= \int_0^\infty [F_0^2(t) - \phi_0^2/2] F(t) dF_0(t). \quad (6)$$

Under H_0, $F_0(t)/F(t) = \phi_0$. This implies that $\Delta_4(F, F_0) = 0$. Under H_4, $\Delta_4(F, F_0) \geq 0$.

Define the kernel

$$\psi_4^*(T_i, \delta_i, T_j, \delta_j, T_k, \delta_k, T_l, \delta_l) = \begin{cases} 1 & \text{if } T_k < T_j < T_l < T_i, \\ & \delta_i = \delta_j = \delta_k = 0, \delta_l = 1 \\ -1 & \text{if } T_l < T_j < T_k < T_i, \\ & \delta_i = \delta_j = \delta_k = 0, \delta_l = 1 \\ 0 & \text{otherwise.} \end{cases}$$

Then the U-statistic corresponding to $\Delta_4(F_0, F)$ is given by

$$U_4 = \binom{n}{4}^{-1} \sum_{1 \leq i_1 < i_2 < i_3 < i_4 \leq n} \psi_4(T_{i_1}, \delta_{i_1}, T_{i_2}, \delta_{i_2}, T_{i_3}, \delta_{i_3}, T_{i_4}, \delta_{i_4}),$$

where ψ_4 is the symmetric version corresponding to ψ_4^*. Note that $E(\psi_4^*(T_i, \delta_i, T_j, \delta_j, T_k, \delta_k, T_l, \delta_l)) = \Delta_4(F, F_0)$ and the expectation of the symmetric kernel is $24\Delta_4(S_1, S)$ due to the possible combinations required to obtain the symmetric kernel. Hence, $E(U_4) = 24\Delta_4(F, F_0)$. Under H_0, $E(U_4) = 0$ and under H_4,$E(U_4) \geq 0$. A rank representation of U_4 is

$$\binom{n}{4} U_4 = \sum_{j=1}^{n_1} \binom{R_{(j)} - j}{2} (n_2 + j - R_{(j)}) - \sum_{j=1}^{n_1} \binom{n_2 - R_{(j)} + j}{3}.$$

Theorem 8

As n tends to ∞, under $H_0, n^{1/2}U_4$ converges in distribution to $N(0, \sigma_4^2)$, where $\sigma_4^2 = (96/35)\phi_0^5(1-\phi_0) = (96/35)\phi(1-\phi)^5$.

We reject the null hypothesis for large values of $n^{1/2}U_4/\hat{\sigma}_4$, where $\hat{\sigma}_4^2 = (96/35)\hat{\phi}_0^5(1 - \hat{\phi}_0) = (96/35)\hat{\phi}(1 - \hat{\phi})^5$.

4. Asymptotic relative efficiency

To compare alternative tests proposed in this paper for testing H_0 against H_2, H_0 against H_3 and H_0 against H_4, we compute asymptotic relative efficiency of the tests for a semiparametric family of distributions proposed in Deshpande (1990). The semiparametric family considered here is $F_1(t) = pF^a(t)$, $F_0(t) = F(t) - pF^a(t)$, where $1 \leq a \leq 2$, $0 \leq p \leq 0.5$ and $F(t)$ is a proper distribution function. Note that $\phi = p$ and

$$\Phi_1(t) = \frac{p(1 - F^a(t))}{1 - F(t)}$$

which is an increasing function of t. Also,

$$\Phi_0^*(t) = 1 - pF^{a-1}(t)$$

which is a decreasing function of t. H_0 corresponds to $a = 1$, and other alternative hypotheses correspond to $1 < a \leq 2$. By the limiting theorem of U-statistics, all the U-statistics proposed here have asymptotic normal distributions under both null and the alternative hypothesis (see Serfling, 1980).

The asymptotic relative efficiency of test U with respect to test U_* is then defined as $eff(U_*, U) = e(U)/e(U_*)$ where $e(U) = \mu'^2(1)/\text{var}(U \mid H_0)$ and $\mu'(1)$ is the derivative of expected value of U with respect to a evaluated at $a = 1$, and $\text{var}(U \mid H_0)$ is the asymptotic variance of $n^{1/2}U$ under H_0.

The tests U_2 and U_2^* are equally efficient but the general test U_1 is four times more efficient compared with these tests for the alternatives considered. This indicates the superiority of U_1 as it is consistent for the alternative H_2. For this particular family of distributions, the other alternative tests are equally efficient. It need not be true in general.

5. Illustrations

We consider two real data sets here, one where the empirical $\Phi_1(t)$ is non-decreasing and the empirical $\Phi_0^*(t)$ is non-increasing. In the other example, both of these seem to be fairly constant.

Example 1. Consider the data on the times to failure, in millions of operations, and modes of failure of 37 switches, obtained from a reliability study conducted at AT&T, given in Nair (1993). There are two possible modes of failure, denoted by A ($\delta = 1$) and B ($\delta = 0$), for these switches.

Figure 1 shows the empirical estimates of the conditional probabilities corresponding to failure modes A and B, respectively. The empirical Φ_1 function corresponding to failure mode A is clearly increasing and the empirical Φ_0^* function corresponding to B is decreasing, indicating that the failure mode A becomes more likely with increase in the age of the switch. Table 1 gives the values of the test statistics. The value of Z corresponding to U_1 is 2.70 and hence we may conclude that the failure time and the type of failure are dependent. The non-constancy of the plot in Fig. 1 supports this conclusion. The one-sided test using U_1 for H_0

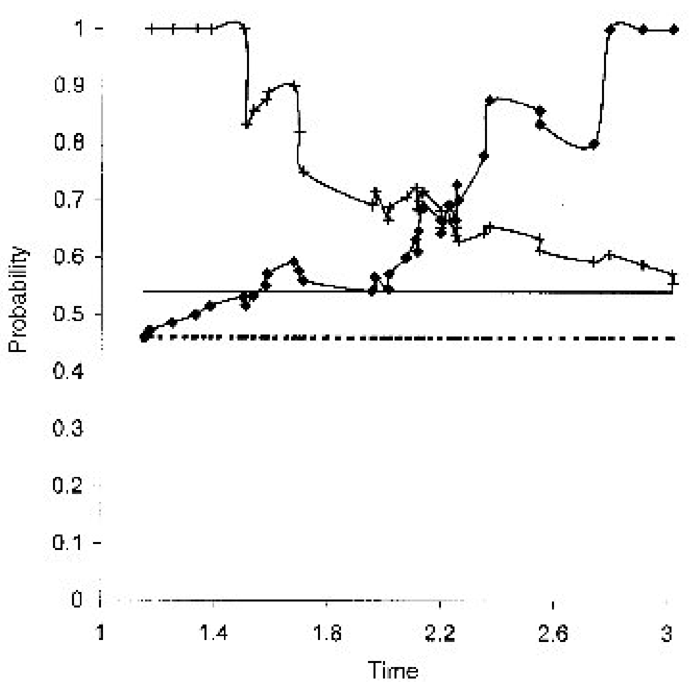

Fig. 1. Time versus empirical $\Phi_1(t)$, $\Phi_1(0)$, $\Phi_0^*(t)$ and $\Phi_0^*(\infty)$ for the data given in Nair (1993). Solid squares denote $\Phi_1(t)$, dotted line denotes $\Phi_1(0)$, pluses denote $\Phi_0^*(t)$ and solid line denotes $\Phi_0^*(\infty)$.

Table 1. *Values of the test statistics for Nair's (1993) data*

U-statistics	Expectation	Variance	Z	Conclusion
$U_1 = 0.26$	0	0.33	2.70	Reject H_0
$U_3 = 0.04$	0	0.03	1.45	Accept H_0
$U_4 = 0.06$	0	0.06	2.29	Reject H_0

against H_2 rejects the null hypothesis of independence of T and δ. It may be concluded that $\Phi_1(t) \geq \phi$ and also $\Phi_0^*(t) \geq \phi_0$ for all t. Further, the value of U_3 given in Table 1 is not statistically significant and hence it may be concluded that $\Phi_1(t)$ is not increasing in t. The test for checking whether $\Phi_0^*(t)$ is decreasing, rejects the null hypothesis and hence we may conclude that $\Phi_0^*(t)$ is a non-increasing function of t. A final conclusion after the application of proposed tests is that $\Phi_1(t) \geq \phi$ for all t and that $\Phi_0^*(t)$ is a non-increasing function of t.

Example 2. Consider the data set obtained from a laboratory experiment on male mice which had received a radiation dose of 300 rads at an age of 5–6 weeks given in Hoel (1972). The death occurred due to cancer ($\delta = 1$), or other causes ($\delta = 0$). In Hoel (1972), the main interest was in judging the equality of the survival functions of the independent latent lifetimes. We have brought out another aspect of the same data without going into the question of independence. Figure 2 shows the empirical conditional probabilities and in this case, the empirical conditional probability $\Phi_1(t)$ seems to be almost flat and the curve corresponding to $\Phi_0^*(t)$ is not so flat. Table 2 gives the values of the test statistics. Based on the application of the proposed tests, it may be concluded that the lifetime T and the cause of death δ are independent for mice living in conventional environment, allowing the analysis of lifetimes and the causes separately.

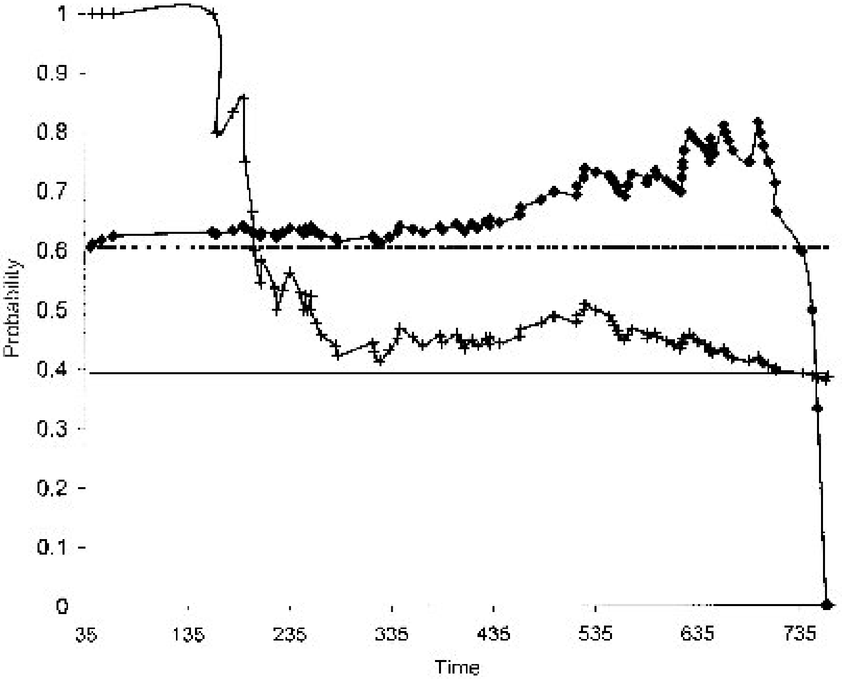

Fig. 2. Time versus empirical $\Phi_1(t)$, $\Phi_1(0)$, $\Phi_0^*(t)$ and $\Phi_0^*(\infty)$ for the data given in Hoel (1972). Solid squares denote $\Phi_1(t)$, dotted line denotes $\Phi_1(0)$, pluses denote $\Phi_0^*(t)$ and solid line denotes $\Phi_0^*(\infty)$.

Table 2. *Values of the test statistics for Hoel's (1972) data*

U-statistics	Expectation	Variance	Z	Conclusion
$U_1 = 0.11$	0	0.32	1.86	Accept H_0
$U_2 = 0.66$	0.61	0.32	0.93	Accept H_0
$U_2^* = 0.45$	0.39	0.32	0.53	Accept H_0
$U_3 = 0.04$	0	0.09	1.50	Accept H_0
$U_4 = 0.01$	0	0.02	1.14	Accept H_0

6. Concluding remarks

The (T, δ) data arise in several areas, viz., engineering studies, actuarial studies, unemployment registers as described in Crowder (2001). These tests would be applicable in all such situations. The tests can also be used to test for the departures from the independence in a more general case when one random variable is continuous and the other binary. A limitation of the proposed tests is that they can be applied only in the situations where there is no additional censoring imposed on to the competing risks. Currently, we are working towards the extension of the tests to incorporate censoring.

For modelling the competing risks data in terms of (T, δ), it is of prime importance to check whether T and δ are independent. We have proposed tests based on U-statistics to check whether T and δ are independent against four different kinds of alternative hypotheses representing various interesting departures from independence. These tests are simple and seem to be useful to distinguish between the possible types of dependence between the causes of failure. It is clear that the tests 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 $\phi = pr(\delta = 1)$ 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.

We suggest to use the test based on U_1 for the general dependence first. If the null hypothesis of independence of T and δ is rejected, then only other tests should be used. The choice of the test for further inference should be based on the plots of the empirical $\Phi_1(t)$ function against t and the empirical $\Phi_0^*(t)$ function against t. Because of hierarchy in the hypotheses H_2 and H_3, we recommend to use the one-sided test based on U_1 to test H_0 against H_2 first and if H_0 is rejected then carry out the test for testing H_0 against H_3 based on U_3. Similarly, the tests for checking the monotonicity of $\Phi_0^*(t)$ could be used.

Acknowledgements

We are thankful to the reviewers for their constructive and critical comments and also for helpful suggestions.

References

Aras, G. & Deshpande, J. V. (1992). Statistical analysis of dependent competing risks. *Statist. Decisions* **10**, 323–336.

Bagai, I., Deshpande, J. V. & Kochar, S. C. (1989). Distribution-free tests for stochastic ordering in the competing risks model. *Biometrika* **76**, 775–781.

Block, H. W. & 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. (2001). *Classical competing risks*. Chapman & Hall/CRC, London.

Davis, T. P. & Lawrance, A. J. (1989). The likelihood for competing risks. *Scand. J. Statist*. **16**, 23–28.

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

Deshpande, J. V. & Sengupta, D. (1995). Testing the hypothesis of proportional hazards in two populations. *Biometrika* **82**, 251–261.

Dykstra, R., Kochar, S. & Robertson, T. (1996). Testing whether one risk progresses faster than the other in a competing risks problem. *Statist. Decisions* **14**, 209–222.

Gasbarra, D. & Karia, S. R. (2000). Analysis of competing risks by using Bayesian smoothing. *Scand. J. Statist*. **27**, 605–617.

Hoel, D. G. (1972). A representation of mortality data by competing risks. *Biometrics* **28**, 475–488.

Kalbfleisch, J. D. & Prentice, R. L. (2002). *The statistical analysis of failure time data*, 2nd edn. Wiley, Hoboken, NJ.

Kochar, S. C. (1979). Distribution-free comparison of two probability distributions with reference to their hazard rates. *Biometrika* **66**, 437–442.

Kulathinal, S. B. & Gasbarra, D. (2002). Testing equality of cause-specific hazard rates corresponding to m competing risks among K groups. *Lifetime Data Anal*. **8**, 147–161.

Larson, M. G. & Dinse, G. E. (1985). A mixture model for the regression analysis of competing risks data. *Appl. Statist*. **34**, 201–211.

Nair, V. N. (1993). Bounds for reliability estimation under dependent censoring. *Int. Statist. Rev*. **61**, 169–182.

Prentice, R. L., Kalbfleisch, J. D., Peterson, A. V., Fluornoy, N., Farewell, V. S. & Breslow, N. E. (1978). The analysis of failure time in the presence of competing risks. *Biometrics* **34**, 541–554.

Sengupta, D. & Deshpande, J. V. (1994). Some results on the relative ageing of two life distributions. *J. Appl. Probab*. **31**, 991–1003.

Serfling, R. J. (1980). *Approximation theorems of mathematical statistics*. Wiley, New York.

Received May 2002, in final form May 2003

S. B. Kulathinal, Department of Epidemiology and Health Promotion, National Public Health Institute, Mannerheimintie 166, 00300 Helsinki, Finland.
E-mail: sangita.kulathinal@ktl.fi