Distribution Free Tests for Stochastic Ordering in the Competing Risks Model

Isha Bagai; Jayant V. Deshpande; Subhash C. Kochar

Biometrika, Vol. 76, No. 4 (Dec., 1989), 775-781.

Stable URL:
http://links.jstor.org/sici?sici=0006-3444%28198912%2976%3A4%3C775%3ADFTFSO%3E2.0.CO%3B2-N

Biometrika is currently published by Biometrika Trust.

Biometrika (1989), **76**, 4, *pp.* 775–81
Printed in Great Britain

Distribution free tests for stochastic ordering in the competing risks model

BY ISHA BAGAI
Department of Statistics, Panjab University, Chandigarh, 160014, India
JAYANT V. DESHPANDÉ
Department of Statistics, University of Poona, Pune, 411007, India
AND SUBHASH C. KOCHAR
Department of Statistics and Actuarial Science, University of Iowa, Iowa City, Iowa 52242, U.S.A.

SUMMARY

The competing risks set-up is considered where an individual is subject to failure due to two independent competing risks. The available data consist of observed times to failure and the causes of failure. On the basis of this information, distribution-free tests are proposed for testing the equality of the two failure distributions against location, scale and general stochastic ordering alternatives. Locally most powerful rank tests are derived and a generalization of the Wilcoxon test has been proposed. Exact critical points are provided for the newly proposed tests, and Pitman efficiency comparisons made.

Some key words: Asymptotic relative efficiency; Locally most powerful rank test; Rank test; U-statistic.

1. INTRODUCTION

Consider the competing risks set-up where a unit is subject to failure due to one of two risks. Let the notional lifetimes of a unit under these two risks be X and Y. The observations are $T = \min(X, Y)$, the time at which the unit fails and $\delta = I(X > Y)$, the failure type, $I(A)$ being the indicator function of the event A. Such data also arise for two components with lifetimes X and Y when they are arranged in series. Cox (1959) has discussed some situations when the above type of data arise and considered some inferential procedures for this model.

We assume that X and Y are independent and absolutely continuous random variables. The assumption of independence of the risks is not always appropriate and cannot be tested from observations (T, δ) alone, due to nonidentifiability problems.

Ordinarily, being lifetimes, X and Y would be positive, but we do not need this restriction. Let F and G be the distribution functions, $\bar{F}$ and $\bar{G}$ the survival functions and f and g the probability density functions of the two risks X and Y, respectively. Let $X_1, \ldots, X_n$ and $Y_1, \ldots, Y_n$ be two independent random samples from F and G, respectively, denoting the hypothetical times to failure of the n individuals in the sample under the two risks. However, because of the competing risks set-up, we observe only $(T_1, \delta_1), \ldots, (T_n, \delta_n)$, where $T_i = \min(X_i, Y_i)$ denotes the time to failure and $\delta_i = I(X_i > Y_i)$ indicates the cause of failure of the ith unit. On the basis of this data, we wish to test the null hypothesis, H_0: $F(x) = G(x)$, for every x, against the alternative, H_A: $F(x) \leqslant G(x)$, for every x, and with a strict inequality for some x.

For uncensored samples, many tests are available for this problem. The Wilcoxon and the Savage tests perform very well under different situations (Kochar, 1978; Deshpandé & Kochar, 1982). In § 2, we derive locally most powerful rank tests for the parametric version of the alternative H_A in the competing risks problem and specialize it to location and scale alternatives. In § 3, we propose two heuristic distribution-free tests based on certain U-statistics. One of these is analogous to the Wilcoxon-signed rank test. In § 4, the exact and the asymptotic distributions of these statistics are discussed. Tables of critical points are provided for small samples. The last section is devoted to asymptotic relative efficiency comparisons.

2. LOCALLY MOST POWERFUL RANK TESTS

Consider the null hypothesis H_0: $F(x)=G(x)$ for all x, against alternative H_1: $F(x)=F_0(x)$ and $G(x)=F_\theta(x)$, for $\theta>0$. In the competing risks set-up the likelihood function is

$$L(t, \delta, \theta)=L(t_1, \dots, t_n; \delta_1, \dots, \delta_n, \theta)=\prod_{i=1}^{n} \{f_\theta(t_i)\bar{F}(t_i)\}^{\delta_i}\{f(t_i)\bar{F}_\theta(t_i)\}^{1-\delta_i}.$$

See, for example, Miller (1981, pp. 16–7).

Let $T_{(1)}, \dots, T_{(n)}$ be the ordered life times and let R_j be the rank of T_j among these. Also let

$$W_j=\begin{cases}1 & \text{if } T_{(j)} \text{ corresponds to a } Y\text{-observation,}\\ 0 & \text{otherwise.}\end{cases}$$

Then the likelihood of ranks $P(R, w, \theta)$ is obtained by integrating $L(t, w, \theta)$ over the range $-\infty<t_{(1)}<\dots<t_{(n)}<\infty$. It may be written as

$$P(R, w, \theta)=\int \prod_{i=1}^{n} [\{f_\theta(t_i)\bar{F}(t_i)\}^{w_i}\{f(t_i)\bar{F}_\theta(t_i)\}^{1-w_i}]\, dt_i.$$

Here and later the integral is over the region $t_1<\dots<t_n$. In particular under H_0, $P(R, w, \theta)=(n!2^n)^{-1}$.

Let us define

$$F^*(t)=[\partial F_\theta(x)/\partial\theta]_{\theta=0}, \quad f^*(t)=[\partial f_\theta(x)/\partial\theta]_{\theta=0},$$
$$h_1(t)=f^*(t)/f(t), \quad h_2(t)=F^*(t)/\bar{F}(t),$$
$$a_j=n!2^n \int h_1(t_j) \prod_{i=1}^{n} \{f(t_i)\bar{F}(t_i)\}\, dt_i, \quad b_j=n!2^n \int h_2(t_j) \prod_{i=1}^{n} \{f(t_i)\bar{F}(t_i)\}\, dt_i.$$

THEOREM 2·1. *Under suitable regularity conditions (Puri & Sen, 1971, pp. 108–9), the locally most powerful rank test for testing H_0 against H_1 is based on the statistic* $\Sigma\{W_j a_j-(1-W_j)b_j\}$.

COROLLARY 2·2. *The locally most powerful rank test for testing H_0 against the logistic location alternative, that is*

$$H_1\text{: } F(x)=\exp(x)/\{1+\exp(x)\}, \quad F_\theta(x)=\exp(x+\theta)/\{1+\exp(x+\theta)\} \quad (\theta>0)$$

is based on the statistic

$$V=\sum_{j=1}^{n} (1-b_j)\, W_j,$$

where

$$b_j = \frac{1}{2n+1} + \sum_{k=2}^{j} \frac{2n(2n-2)\dots(2n-2k+4)}{(2n+1)(2n-1)\dots(2n-2k+3)} \quad (j = 1, \dots, n).$$

COROLLARY 2·3. *The locally most powerful rank test for testing H_0 against the exponential scale alternative, that is,*

$$H_1\colon \bar{F}(x) = \exp(-x), \quad \bar{F}_\theta(x) = \exp\{-(1+\theta)x\} \quad (\theta > 0)$$

is based on the sign statistic

$$S_1 = \sum_{i=1}^{n} W_i = \sum_{i=1}^{n} \delta_i.$$

Note that for uncensored samples, the corresponding locally powerful rank tests are the Wilcoxon test and the Savage exponential scores tests.

3. TESTS BASED ON U-STATISTICS

The locally most powerful rank tests derived in § 2 are difficult to use for most distributions because of lack of precise information about the alternatives and analytic difficulties. Hence we propose below simple distribution-free tests based on certain U-statistics.

Gehan's (1965) and Prentice's (1978) generalizations of the Wilcoxon test are useful when the two samples undergo independent random censoring. However, in the competing risks set-up the two random samples censor each other and these tests are not applicable. The sign statistic

$$U_1 = n^{-1} \sum \delta_i = n^{-1} S_1, \tag{3·1}$$

the proportion of deaths due to second cause, is easily constructed and has attractive properties. Under the alternative H_A, a greater number of individuals or units are expected to die from Risk II than from Risk I. Hence large values of U_1 are significant for testing H_0 against H_A. However, some additional information can be utilized for improving the efficiency for some alternatives.

Let us simultaneously look at the pairs (T_i, δ_i) and (T_j, δ_j). Table 1 is a description of all mutually exclusive and exhaustive arrangements of the pairs (T_i, δ_i) and (T_j, δ_j), in terms of failure times. Consider the kernel

$$\phi_2(T_i, \delta_i; T_j, \delta_j) = \begin{cases} 1 & (\delta_i = 1,\ T_i < T_j), \\ 1 & (\delta_j = 1,\ T_j < T_i), \\ 0 & \text{otherwise}. \end{cases} \tag{3·2}$$

Table 1. *Information regarding (X_i, Y_i) and (X_j, Y_j) available in (T_i, δ_i) and (T_j, δ_j)*

	$\delta_i = 1, \delta_j = 1$	$\delta_i = 1, \delta_j = 0$	$\delta_i = 0, \delta_j = 1$	$\delta_i = 0, \delta_j = 0$
$T_i > T_j$	$X_i > Y_i, X_j > Y_j$ $X_i > Y_j$	$X_i > Y_i, Y_j \geqslant X_j$ $Y_i > X_j$	$Y_i \geqslant X_i, X_j > Y_j$ $X_i > Y_j$	$Y_i \geqslant X_i, Y_j \geqslant X_j$ $Y_i > X_j$
$T_i < T_j$	$X_i > Y_i, X_j > Y_j$ $X_j > Y_i$	$X_i > Y_i, Y_j \geqslant X_j$ $X_j > Y_i$	$Y_i \geqslant X_i, X_j > Y_j$ $Y_j > X_i$	$Y_i \geqslant X_i, Y_j \geqslant X_j$ $Y_j > X_i$

In each of the four events for which $\phi(T_i, \delta_i; T_j, \delta_j)$ takes the value 1 out of the total of eight events, we find that $\min(Y_i, Y_j) < \min(X_i, X_j)$. In the other four events when the kernel takes the value 0, the minimum of X's is smaller than the minimum of Y's.

Construct the U-statistic

$$U_2^* = \left\{\binom{n}{2}\right\}^{-1} \sum_{1 \le i < j \le n} \phi_2(T_i, \delta_i; T_j, \delta_j) \tag{3·3}$$

$$= \left\{\binom{n}{2}\right\}^{-1} \sum_{i=1}^{n} (n - R_i)\delta_i, \tag{3·4}$$

where R_i is the rank of T_i among $T_1, \ldots, T_n$. Since $T_i = \min(X_i, Y_i)$ and $\delta_i = 1$ if and only if $Y_i < X_i$, the quantity in the parentheses in (3·4) is the number of X_j's greater than the observed Y_i, for $i \neq j$. However, since $\delta_i = 1$, $Y_i < X_i$. But the statistic U_2^* does not count this. Hence we modify U_2^* and suggest the asymptotically equivalent version U_2 defined by

$$U_2 = \left\{\binom{n}{2}\right\}^{-1} \sum_{i=1}^{n} (n - R_i + 1)\delta_i. \tag{3·5}$$

Here U_2 is the total number of times an X observation is greater than an observed Y observation in the combined arrangement of n X's and n Y's. In a sense this is analogous to the Wilcoxon signed rank statistic adapted to the competing risk set-up. Large values of U_2 are significant for testing H_0 against H_A.

We propose another kernel

$$\phi_3(T_i, \delta_i; T_j, \delta_j) = \begin{cases} 3 & (\delta_i = 1, \delta_j = 1), \\ 1 & (T_i > T_j, \delta_i = 0, \delta_j = 1), \\ 1 & (T_j > T_i, \delta_i = 1, \delta_j = 0), \\ -1 & (T_i > T_j, \delta_i = 0, \delta_j = 0), \\ -1 & (T_j > T_i, \delta_i = 0, \delta_j = 1), \\ -3 & (\delta_i = 0, \delta_j = 0). \end{cases} \tag{3·6}$$

The rationale for (3·6) is as follows.

In the pairs (X_i, Y_i) and (X_j, Y_j) we assign a score 1 if an X-observation is known to be greater than a Y-observation and a score -1, otherwise. Thus $\delta_i = 1$, $\delta_j = 1$, $T_i > T_j$ gives us the information that $X_i > Y_i$, $X_j > Y_j$, $X_i > Y_j$ but no information about (X_j, Y_i). Thus a total score of 3 is assigned to this arrangement. The assignment $\delta_i = 0$, $\delta_j = 1$, $T_i > T_j$ provides the information that $Y_i > X_i$, $X_i > Y_j$, $X_j > Y_j$, only so we have two favourable and one unfavourable comparison giving a total score of 1. Similarly, scores are attached to the other six arrangements.

We construct the U-statistic corresponding to the above kernel, namely

$$U_3 = \left\{\binom{n}{2}\right\}^{-1} \sum_{1 \le i < j \le n} \phi_3(T_i, \delta_i, T_j, \delta_j) \tag{3·7}$$

$$= \left\{\binom{n}{2}\right\}^{-1} \left\{2 \sum_{i=1}^{n} (2n - 1 - R_i)\delta_i - \frac{3n(n-1)}{2}\right\}. \tag{3·8}$$

Thus, U_3 is a linear combination of the sign statistic and the modified Wilcoxon statistic $\Sigma\, R_i\delta_i$. Again, large values of U_3 are significant for testing H_0 against H_A.

4. Distributions of the test statistic

Let

$$S_2 = \binom{n}{2} U_2, \quad S_3 = \binom{n}{2} U_3. \tag{4·1}$$

We first find the moment generating functions of S_2 and S_3 under H_0 using the following results.

The pair (X, Y) follows the proportional hazards model $\bar{G}(x) = \{\bar{F}(x)\}^{\beta}$ for a positive constant β if and only if $T = \min(X, Y)$ and $\delta = I(X > Y)$ are independent (Armitage, 1959; Allen, 1963). Also, under H_0, $W_1, \ldots, W_n$ are independent and identically distributed with $\text{pr}(W_i = 1) = \text{pr}(W_i = 0) = \frac{1}{2}$ $(i = 1, \ldots, n)$.

With the help of these results, we can prove the following theorem.

Theorem 4·1. *Under H_0, the moment generating functions of S_2 and S_3 are*

$$M_2(t) = 2^{-n} \prod_{j=1}^{n} \{1 + \exp(jt)\}, \tag{4·2}$$

$$M_3(t) = 2^{-n} \exp\left\{\frac{-3n(n-1)t}{2}\right\} \prod_{j=2}^{n+1} [1 - \exp\{2t(2n-j)\}]. \tag{4·3}$$

Proof. From (3·5) we know that

$$S_2 = \sum_{j=1}^{n} (n - R_j + 1)\delta_j = \sum_{j=1}^{n} (n - j + 1) W_j,$$

from which (4·2) follows immediately. Similarly, we find $M_3(t)$, the moment generating function of S_3. □

Now $M_2(t)$ is the same as the moment generating function of the Wilcoxon signed rank statistic under H_0 (Hettmansperger, 1984, p. 35). Hence S_2 will have the same null distribution as the Wilcoxon signed rank statistic. In particular, the tables of critical points of the Wilcoxon signed rank test can be used for performing the test S_2. These are given, for example, by Hollander & Wolfe (1973).

Following Hettmansperger (1984, pp. 35–6), we find the probability distribution of S_3, under H_0, by using the expansion

$$M_3(t) = a_0 + a_1 \exp(t) + a_2 \exp(2t) + \ldots$$

of the moment generating function and observing that $\text{pr}(S_3 = j) = a_j$. Table 2 gives the upper critical points of S_3. It is also clear from (4·3) that the null distribution of S_3 is symmetric about zero. From the moment generating functions, we easily find that under H_0,

$$E(S_2) = n(n+1)/4, \quad \text{var}(S_2) = n(n+1)(2n+1)/24,$$

$$E(S_3) = 0, \quad \text{var}(S_3) = n(n-1)(14n-13)/6.$$

From (3·3) and (3·5) we find that

$$n^{\frac{1}{2}}\{U_2 - E(U_2)\} = n^{\frac{1}{2}}\{U_2^* - E(U_2^*)\} + n^{\frac{1}{2}}\{2(n-1)\}^{-1}\{U_1 - E(U_1)\}.$$

Now U_1 and U_2^* are U-statistics and $n^{\frac{1}{2}}\{U_1 - E(U_1)\}$ converges to a normal distribution with mean zero. Hence asymptotically, U_2 and U_2^* are equivalent. The proof of the following theorem easily follows (Puri & Sen, 1971, p. 51).

Table 2. *Critical points for S_3, exact significance level and level based on normal approximation*

n	1%	5%
4	—	—
5	—	22; 0·0625, 0·0552
6	—	31; 0·0625, 0·0499
7	51; 0·0156, 0·0183	41; 0·0547, 0·0463
8	68; 0·0117, 0·0126	50; 0·0508, 0·0499
9	84; 0·0117, 0·0113	62; 0·0508, 0·0461
10	99; 0·0107, 0·0117	73; 0·0508, 0·0472
11	115; 0·0102, 0·0118	83; 0·0527, 0·0513
12	134; 0·0105, 0·0186	98; 0·0500, 0·0467
13	152; 0·0102, 0·0109	108; 0·0528, 0·0516
14	169; 0·0106, 0·0116	123; 0·0511, 0·0494
15	191; 0·0103, 0·0107	137; 0·0516, 0·0495
16	210; 0·0105, 0·0111	150; 0·0523, 0·0512
17	232; 0·0100, 0·0181	166; 0·0519, 0·0500
18	255; 0·0100, 0·0105	181; 0·0516, 0·0506
19	275; 0·0104, 0·0110	197; 0·0516, 0·0505
20	298; 0·0104, 0·0109	214; 0·0511, 0·0499

THEOREM 4·2. *The asymptotic distribution of $n^{\frac{1}{2}}\{U_i - E(U_i)\}$ is normal with mean zero and variance $\sigma^2_{U_i} = 4\xi_{i12}$ for $i = 2, 3$, where*

$$\xi_{i12} = E\{\psi_i^2(X_1, Y_1)\} - E^2(U_i), \quad \psi_i(x_1, y_1) = E\{\psi_i(x_1, y_1; X_2, Y_2)\}.$$

Under H_0,

$$E(U_2) = \tfrac{1}{2}, \quad \sigma^2_{U_2} = \tfrac{1}{3}, \quad E(U_3) = 0, \quad \sigma^2_{U_3} = \tfrac{28}{3}.$$

The normal approximation to the null distribution of U_3 or S_3 is fairly good for $n > 20$. In Table 2 we provide for $n \leqslant 20$ the critical points of the statistic S_3 for levels of significance near 0·01 and 0·05, and associated exact and approximate significance levels. Now $E_{H_0}(U_i) < E_{H_A}(U_i)$ $(i = 1, 2, 3)$; hence tests based on these statistics are consistent for the entire alternative hypothesis H_A.

5. ASYMPTOTIC RELATIVE EFFICIENCIES

To compare the Pitman asymptotic relative efficiencies, we parameterize the problem in the following way. Let $F(x) = F_0(x)$ and $G(x) = F_\theta(x)$, where $\theta \geqslant 0$ and such that $F_0(x) \leqslant F_\theta(x)$ for all x and with a strict inequality over a set of nonzero probability for every $\theta > 0$. The following alternatives have been considered for efficiency comparisons:

location alternative, H_1: $F_\theta(x) = F(x + \theta)$;

scale alternative, H_2: $F_\theta(x) = F\{(\theta + 1)x\}$ $(x \geqslant 0)$;

Makeham distribution, H_3: $\bar{F}_\theta(x) = \exp[-\{x + \theta(x + e^{-x} - 1)\}]$;

$$H_4: \bar{F}_\theta(x) = \bar{F}(x)\left[1 + \theta\left\{\sum_{i=1}^{k} F^i(x)\right\}\right] \quad (k \geqslant 1, 0 < \theta < 1/k);$$

$$H_5: \bar{F}_\theta(x) = (1 - \theta)\bar{F}(x) + \theta F^{k+1}(x) \quad (k > 0);$$

proportional hazards model, H_6: $\bar{F}_\theta(x) = \{\bar{F}(x)\}^{1+\theta}$.

All these alternatives indicate the desired dominance. Table 3 gives the asymptotic relative efficiencies of the U_2 and U_3 tests with respect to the sign test U_1 for the above alternatives (Puri & Sen, 1971, p. 116). The tests based on U_2 and U_3 do generally better than the sign test based on U_1 for location alternatives. The sign test seems to be better for alternatives similar to the exponential scales, for which it is optimal. In many other cases of stochastic dominance the tests based on U_2 and U_3 have larger efficiency. Hence a choice between U_1, U_2 and U_3 should be made accordingly. All three tests are consistent and unbiased for testing H_0 against H_A.

Table 3. *Asymptotic relative efficiencies of the U_2 and U_3 tests with respect to the sign test U_1*

	Alternatives	U_2	U_3		Alternatives	U_2	U_3
H_1	Uniform; H_5: $k=2$	1·333	1·190	H_2	Exponential, H_6	0·750	0·964
H_1	Normal	1·109	1·108	H_3		0·270	0·724
H_1	Logistic; H_4: $k=1$; H_5: $k=1$	1·080	1·097	H_4	$k=2$	1·422	1·008
H_1	Double exponential	1·021	1·074	H_4	$k=3$	0·499	0·958
H_1	Exponential	3·000	1·714	H_5	$k=3$	1·531	1·259

Acknowledgements

The authors are grateful to the referees and editor for their helpful comments and suggestions which resulted in an improved version of the paper. This research was completed while J. V. Deshpandé was visiting the University of California, Santa Barbara and S. C. Kochar was visiting the University of Iowa, Iowa City.

References

Allen, W. R. (1963). A note on the conditional probabilities of failure when hazards are proportional. *Oper. Res.* **11**, 658-9.

Armitage, P. (1959). The comparison of survival curves (with discussion). *J.R. Statist. Soc.* A **122**, 279-300.

Cox, D. R. (1959). The analysis of exponentially distributed lifetimes with two types of failures. *J.R. Statist. Soc.* B **21**, 411-21.

Deshpandé, J. V. & Kochar, S. C. (1982). Some competitors of the Wilcoxon-Mann-Whitney test for the location alternative. *J. Ind. Statist. Assoc.* **20**, 9-18.

Gehan, E. A. (1965). A generalized Wilcoxon test for comparing arbitrarily singly censored samples. *Biometrika* **52**, 203-23.

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

Hollander, M. & Wolfe, D. A. (1973). *Nonparametric Statistical Methods.* New York: Wiley.

Kochar, S. C. (1978). A class of distribution-free tests for the two sample slippage problem. *Comm. Statist.* A **7**, 1243-52.

Miller, R. G. (1981). *Survival Analysis.* New York: Wiley.

Prentice, R. L. (1978). Linear rank tests with right censored data. *Biometrika* **65**, 167-79.

Puri, M. L. & Sen, P. K. (1971). *Nonparametric Methods in Multivariate Analysis.* New York: Wiley.

[*Received December* 1987. *Revised February* 1989]