Precedence tests for right-censored data: an overview and some results
Chakraborti, S.; van der Laan, P.

Published: 01/01/1996

Document Version
Publisher’s PDF, also known as Version of Record (includes final page, issue and volume numbers)

Please check the document version of this publication:
• A submitted manuscript is the author's version of the article upon submission and before peer-review. There can be important differences between the submitted version and the official published version of record. People interested in the research are advised to contact the author for the final version of the publication, or visit the DOI to the publisher's website.
• The final author version and the galley proof are versions of the publication after peer review.
• The final published version features the final layout of the paper including the volume, issue and page numbers.

Link to publication

Citation for published version (APA):

General rights
Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

• Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
• You may not further distribute the material or use it for any profit-making activity or commercial gain
• You may freely distribute the URL identifying the publication in the public portal

Take down policy
If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

Download date: 11. Dec. 2018
Memorandum COSOR 96-01

Precedence Tests for Right-Censored Data: An Overview and Some Results

S. Chakraborti
P. van der Laan

Eindhoven, January 1996
The Netherlands
Precedence Tests for Right-Censored Data: An Overview and Some Results

S. Chakraborti
Department of Management Science and Statistics
University of Alabama
P.O. Box 870226
Tuscaloosa, AL 35487
U.S.A.

and

P. van der Laan
Department of Mathematics and Computing Science
Eindhoven University of Technology
P.O. Box 513, HG
5600 MB Eindhoven
The Netherlands

Summary

Precedence tests are simple yet robust nonparametric procedures useful for comparing two or more distributions. In this paper precedence type tests are considered when the data contain some right-censored observations. Generalizing the precedence statistic for uncensored data, the precedence tests for censored data are based on the Kaplan-Meier estimators of the respective distribution functions and the corresponding quantile functions. An overview of the literature is given for the two-sample as well as some multi-sample problems. Some further problems are indicated.


Key words and Phrases: Right-censored data, Kaplan-Meier estimator, Two-sample problems, K-sample problems, Restricted alternatives, Precedence tests, Asymptotic relative efficiency.
1 Introduction

The purpose of this paper is to review a class of nonparametric tests based on what are called precedence statistics, when the data to be analyzed contain some right-censored observations. These tests, called precedence tests, provide a simple yet robust comparison of the underlying distribution functions. Precedence tests have been of interest in a variety of applications, especially in the context of life-testing or similar experiments where the observations become available in a time ordered manner. For a comprehensive overview of precedence tests for complete or uncensored data the reader may consult the recent paper by Chakraborti and van der Laan (hereafter referred to as CV) (1994). In this paper we assume random right-censorship and consider some two-sample problems as well as some \( K \)-sample problems. Precedence type tests are introduced along with a brief motivation from the complete data case. The literature is reviewed and some open problems are noted. We begin with the two-sample problem.

2 Two-Sample Problem

Let \( X_1, X_2, \ldots, X_{n_1} \) and \( Y_1, Y_2, \ldots, Y_{n_2} \) be two independent random samples from absolutely continuous distribution functions \( F_1 \) and \( F_2 \), respectively. In the applications of interest, the \( X \)'s and the \( Y \)'s are positive valued random variables so that \( F_1(0) = F_2(0) = 0 \). Let \( X(1) < X(2) < \ldots < X(n_1) \) and \( Y(1) < Y(2) < \ldots < Y(n_2) \) denote the order statistics of the \( X \)- and the \( Y \)-sample, respectively. For a specified value of \( r (1 \leq r \leq n_1) \), let \( v \) denote the number of \( Y \)-observations that precede (are less than) \( X(r) \), \( 0 \leq v \leq n_2 \). The random variable \( v \) is called a precedence statistic and a test based on \( v \) is referred to as a precedence test. The quantity \( r \) is often fixed in relation with a specified quantile of \( F_1 \). For example, to compare the \( Y \)'s with respect to the median of the \( X \)'s, one could take \( r = [n_1/2] + 1 \), where \([a]\) denotes the largest integer not exceeding \( a \). In general, one could take \( r = [n_1p] + 1 \), when the interest is focused on the \( p \)-quantile \( \xi_i(p) \) of \( F_i \), where the \( p \)-quantile of the \( ith \) population is defined as \( \xi_i(p) = F_i^{-1}(p) = \inf\{t \geq 0 : F_i(t) \geq p\} \), \( p \in (0,1) \), \( i = 1,2 \). In some applications one of the distributions would correspond to a "control" population and the other to an "experimental" population and it might be natural to identify \( F_1 \) with the control. One of the early precedence tests, called the control median test, proposed by Mathisen (1943), used such a formulation.

Consider the problem of testing \( H_0 : F_1(x) = F_2(x) \) for all \( x \), against the one-sided alternative \( H_1 : F_2(x) \leq F_1(x) \), with strict inequality for at least one \( x \). The precedence test for uncensored data rejects \( H_0 \) in favour of \( H_1 \) if \( v < v \), where \( v \) is determined so that the size of the test is \( \alpha \). It was noted in CV that a test based on \( v \) is statistically equivalent to a test based on a comparison of two order statistics, one from each sample. This follows easily, since, \( v < v \) if and only if \( X(r) < Y(v) \), so that a precedence test in fact involves a comparison of two sample quantiles, one from each sample. In the literature both forms of precedence tests can be found, some using the counting form.
and others using the order statistics form.

In some practical situations the data are subject to random right-censorship. This is fairly common in clinical trials, reliability studies and similar experiments. In the presence of randomly right-censored data, the precedence test statistic can be adapted using the Kaplan-Meier estimates of the quantiles. To define the Kaplan-Meier (KM) product-limit estimator, note that due to the presence of censorship one may not observe the responses (say lifetimes) \( X_i \), but pairs of random variables \((Z_i, \delta_i)\), where \( Z_i = \min(X_i, C_i) \) and \( C_i \) are some "censoring" variables, \( \delta_i = I(X_i \leq C_i) \), \( i = 1, 2, ..., n_1 \), \( I(\cdot) \) being the usual indicator function. The \( \delta_i \) are random variables that indicate whether \( X_i \) is censored (\( \delta_i = 0 \)) or uncensored (\( \delta_i = 1 \)). Similarly, for the \( Y\)-sample one observes \((Z^*_j, \xi_j)\), where \( Z^*_j = \min(Y_j, D_j) \), the \( D_j \) are the censoring variables and \( \xi_j = I(Y_j \leq D_j) \), \( j = 1, 2, ..., n_2 \). The censoring variables \( C_1, C_2, ..., C_{n_1} \) and \( D_1, D_2, ..., D_{n_2} \) are assumed to be independent continuous random variables with cumulative distribution functions \( G_1 \) and \( G_2 \), respectively. Also for each sample, the censoring variables are assumed to be independent of the lifetimes.

Let \( Z_{(1)} < Z_{(2)} < ... < Z_{(n_1)} \), be the ordered \( Z\)'s and let \( \delta_{[i]} \) be the value of \( \delta \) associated with \( Z_{(i)} \), \( i = 1, 2, ..., n_1 \). Further let \( Z'_{(1)} < Z'_{(2)} < ... < Z'_{(t)} \), \( 1 \leq t \leq n_1 \), be the distinct ordered values of the \( Z_i \) and let \( \delta_{[h]} \) be the delta value associated with \( Z'_{(h)} \), \( h = 1, 2, ..., t \). The KM estimator of the survival function \( S_1(x) = 1 - F_1(x) \) is given by

\[
\hat{S}_1(x) = \prod_{j: Z'_{(j)} \leq x} (1 - \frac{d_j}{R_j})^{\delta_{[j]}},
\]

where \( R_j \) denotes the number of units at risk (not failed) at time just prior to \( Z'_{(j)} \) and \( d_j \) denotes the number of units that failed at time \( Z'_{(j)} \). The KM estimator of \( F_1(x) \) is obviously \( \hat{F}_1(x) = 1 - \hat{S}_1(x) \). The KM estimator of \( F_2(y) \) is similarly defined. Also, for \( u \in (0, 1) \), let \( \hat{F}_i^{-1}(u) = \inf\{t \geq 0 : \hat{F}_i \geq u\} \) be the KM empirical quantile function corresponding to \( \hat{F}_i \). It may be noted that in the absence of censorship, \( \hat{F}_i \) and \( \hat{F}_i^{-1} \) reduces to the usual empirical distribution function \( F_{ni} \) and the empirical quantile function \( F_{ni}^{-1} \) of the sample, respectively, \( i = 1, 2 \).

Let \( \hat{F}_i^{-1}(p) \) be the KM estimator of the \( p \)-quantile \( \xi_i(p) \), \( i = 1, 2 \). Note that for uncensored data the precedence statistic \( V_r \) can be expressed as \( n_2 F_{n_2} F_{n_1}^{-1}(p) \), where \( p = \frac{r}{n} \).

A natural generalization of \( V_r \) to the case of censored data is based on \( \hat{V}_p = n_2 \hat{F}_2 \hat{F}_1^{-1}(p) \), where \( p \in (0, 1) \) is specified beforehand. The precedence test for censored data is to reject \( H_0 \) in favour of the alternative \( H_1 \) if \( \hat{V}_p < v \), where \( v \) is to be determined so that the size of the test is \( \alpha \).

With uncensored data the null distribution of the precedence statistic \( V_r \) can be determined exactly (see, for example, CV) and hence one can find the exact critical value or the exact \( P \)-value for the test. However, in the presence of censored data, the null distribution of \( \hat{V}_p \) is complicated and we settle for the asymptotic critical value (or the \( P \)-value). Towards this end, the following result plays a key role.
Theorem 2.1 Let $N = n_1 + n_2$ and $n_1, n_2 \to \infty$ such that $n_1/N \to \lambda_1$ and $n_2/N \to \lambda_2$, $0 < \lambda_i < 1$, $i = 1, 2$. Let $\nu_p = F_2(\xi_1(p))$ and let $f_i = F_i'$ exist, for $i = 1, 2$. Further let $0 < φ = \frac{f_1(\xi_1(p))}{f_2(\xi_1(p))} < \infty$. The asymptotic distribution of $N^{-1/2}(\bar{V}_p - n_2\nu_p)$ is normal with mean 0 and variance

$$σ^2 = \frac{λ_2}{λ_1}(λ_1 I_2 + φ^2 λ_2 I_1),$$

where

$$I_2 = (1 - ν_p)^2 \int_0^{ξ_1(p)} \frac{dF_2}{(1 - F_2)^2(1 - G_2)},$$

and

$$I_1 = (1 - p)^2 \int_0^{ξ_1(p)} \frac{dF_1}{(1 - F_1)^2(1 - G_1)}.$$

A proof of Theorem 1 can be found in Chakraborti (1984). As noted in CV, a special case of a precedence test is the control quantile test where the $X$-population is a "control" and the $Y$-population is some treatment population. The control quantile test is an extension of the control median test proposed by Mathisen (1943). Chakraborti (1984), in an unpublished dissertation, considered a generalization of the control quantile test to the case of randomly right-censored data. One of his main results is the above theorem, which was proved using results in Cheng (1984). Motivation behind Chakraborti’s work was the work of Brookmeyer and Crowley (1982), where an extension of Mood’s median test to the case of randomly right-censored data was studied. A part of Chakraborti’s work overlapped with the work of Brookmeyer (1983), where the main focus is prediction. Gastwirth and Wang (1988), using results of Lo and Singh (1985), extended the control quantile test to the case of randomly right-censored data and also obtained, in particular, the above result. Part of their work (mainly distributional) also overlapped with that of Chakraborti (1984). As we shall see later, Theorem 2.1 can be extended to the case of more than two groups and that will provide the basis for some $K$-sample test procedures.

Remark 1 The result of Theorem 2.1 may be restated as follows. Under the conditions of the theorem, $n_2^{-1/2}(\bar{V}_p - n_2\nu_p)$ has an asymptotically normal distribution with mean 0 and variance $I_2 + φ^2 \lambda_2 I_1$. In the absence of censorship ($G_i = 0$, $i = 1, 2$) one has $I_2 = ν_p(1 - ν_p)$ and $I_1 = p(1 - p)$, so the variance reduces to $ν_p(1 - ν_p) + φ^2 \lambda_2 p(1 - p)$. This agrees with the corresponding result given in CV.

Now, under the null hypothesis $φ = 1$ and $σ^2$ reduces to

$$σ_0^2 = \frac{λ_2}{λ_1}(λ_1 I_2^0 + λ_2 I_1^0),$$

where

$$I_i^0 = (1 - p)^2 \int_0^{ξ_1(p)} \frac{dF}{(1 - F)^2(1 - G_i)}, \quad i = 1, 2,$$

where $F$ is the common but unknown c.d.f. under $H_0$ and $ξ(p) = F^{-1}(p)$. From Theorem 2.1 it follows that under $H_0$, the asymptotic distribution of $N^{-1/2}(\bar{V}_p - n_2p)$
is normal with mean 0 and variance $\sigma^2_0$. The quantity $I^0_i$ can be consistently estimated using a Greenwood estimator (see for example, Miller, 1981), given by,

$$\hat{I}^0_i = (1 - p)^2 \sum_{j : Z_{i(j)} \leq \hat{F}^{-1}(p)} \frac{n_i d_{ij}}{R_{ij}(R_{ij} - d_{ij})},$$

(2)

where $d_{ij}$ and $R_{ij}$ denote the number of failures and the number at risk, respectively, at $Z_{i(j)}$, the distinct $j$th largest failure time in the $i$th sample, $i = 1, 2$, and $\hat{F}^{-1}(p)$ is some consistent estimator of $\xi(p)$. The question regarding how $\hat{F}$ is to be calculated needs to be addressed and some suggestions will be made later. Hence a consistent estimator of the asymptotic null variance $\sigma^2_0$ is

$$\hat{\sigma}^2_0 = \frac{\lambda_{2,N}}{\lambda_{1,N}}(\lambda_{1,N} \hat{I}^0_2 + \lambda_{2,N} \hat{I}^0_1),$$

(3)

where $\lambda_{i,N} = n_i/N$, $i = 1, 2$. The approximately size $\alpha$ precedence test for the two-sample problem with randomly right-censored data is to reject $H_0$ in favour of $H_1$ if

$$\frac{N^{-1/2}(\hat{\nu}_p - n_2 p)}{\hat{\sigma}_0} < -z_\alpha,$$

(4)

where $z_\alpha$ is the upper 100$\alpha$-percentile of the standard normal distribution. It may be noted that the test in (4) is valid whether or not the two groups are subject to the same pattern of random right-censorship.

Remark 2 Experience suggests that it is better to use a linearly interpolated version of the KM estimator while computing the precedence test statistic. This seems to be particularly important with smaller sample sizes and/or heavier censorship. The linear interpolation does not alter the asymptotic theory.

Remark 3 The critical region in (4) can be rewritten as

$$\hat{\nu}_p < n_2 p - N^{1/2} \hat{\sigma}_0 z_\alpha,$$

which can be further rewritten as

$$\hat{F}^{-1}_1(p) < \hat{F}_2^{-1}\{p - z_\alpha(N \lambda_{1,N} \lambda_{2,N})^{-1/2} (\lambda_{1,N} \hat{I}^0_2 + \lambda_{2,N} \hat{I}^0_1)^{1/2}\}. $$

(5)

Thus, as noted before, a two-sample precedence test for censored data involves a comparison of two KM quantiles, one from each sample. It is interesting to note that in the uncensored case $I^0_i = p(1 - p)$, and (5) reduces to
where
\[ u_\alpha = z_\alpha \{ p(1 - p) \}^{1/2} (N\lambda_1 N\lambda_2 N)^{-1/2}. \]

In other words, the approximately size \( \alpha \) two-sample precedence test for uncensored data compares the \( p \)-quantile of the \( X \)-sample with the \((p - u_\alpha)\)-quantile of the \( Y \)-sample.

Slud (1992) studied precedence tests for randomly right-censored data. He observed that "such a test at first appears wasteful of information in looking at only one section of the survival curve, but also allows a clear and robust interpretation for all families of stochastically ordered alternatives." One of Slud's objectives was to find the "best" precedence test and hence "to shed some light on the types of two-sample data with non-proportional hazards for which such a procedure can perform respectfully well compared to the logrank." He concluded that "although the logrank dominates this (precedence) statistic in many cases of practical interest, it is remarkable that such a simple testing strategy as the "best precedence tests" leads to asymptotic relative efficiencies against the logrank which range from about 2/3 to values much larger than 1 for a variety of reasonable local alternatives." The main result is given by the following theorem (stated below using our notation). Let

\[ \tau^2(r) = \int_0^{F_1^{-1}(r)} \left( \frac{1}{G_2(x)} + \frac{\lambda_2}{\lambda_1 G_1(x)} \right) \frac{dH_1(x)}{(1 - F_1(x))}, \]

where \( H_1(t) \) is the cumulative hazard function of population 1, given by,

\[ H_1(t) = \int_0^t \frac{dF_1(x)}{1 - F_1(x)}. \]

**Theorem 2.2** The test of \( H_0 : F_1 = F_2 \) versus \( H_1 : F_1(t) \leq F_2(t) \) for all \( t \), which rejects when \( F_2^{-1}(r) < F_1^{-1}(s) \), where

\[ s = r - z_\alpha (1 - r) \tau(r), \]  

and \( r^* \) is chosen to maximize

\[ \frac{ln^2(1 - r)}{\tau^2(r)} \]

is called the best (ideal) precedence test, and has the following asymptotic properties as \( n_1, n_2 \to \infty \) in such a way that \( \lambda_{2,N} \) tends to the limit \( \lambda_2 \) between 0 and 1.

(i) The asymptotic significance level is \( \alpha \), and the precedence test with \( r \) and \( s \) related by (8) is asymptotically equivalent to the size-\( \alpha \) test based on \( F_2(F_1^{-1}(r)) - F_1(F_1^{-1}(r)) \).
(ii) Among all test of this form for various values of r and s which have asymptotic size α, this test has greatest asymptotic power against contiguous Lehmann alternatives.

(iii) Against alternatives

\[ H_{1,n} : H_2(t) = \int_0^t \left( 1 + \frac{c(s)}{\sqrt{n}} \right) dH_1(s) \]

(for bounded c(.)) under which only \( H_2(t) \) and not \( H_1(t) \) or \( G_i(t) \) depend on \( n_1, n_2 \), as \( n \) goes to \( \infty \), this test has power \( 1 - \Phi(z_\alpha - \text{Eff}) \), with \( \text{Eff} \) given by the formula

\[ \int_0^{-\ln(1-r^*)} \frac{du}{c(H_1^{-1}(u)) \tau(r^*)}. \]

In order that the contiguous alternatives \( H_{1,n} \) satisfy the stochastic-ordering property \( F_2(t) \geq F_1(t) \) for all \( t \), it suffices to assume that \( \int_0^t c(H_1^{-1}(u)) du \geq 0 \) for all \( t \).

(iv) When there is no censoring, the unique value \( r^* \) maximizing (9) is the solution of \( \ln(1 - r) = -2r \) and is equal to .797. The asymptotic relative efficiency of the Best Precedence test to the logrank for Lehmann alternatives (\( H_{1,n} \) with constant \( c \)) is 0.65.

2.1 Testing Equality of Quantiles In some situations one may wish to test only the equality of some quantiles from the two distributions which can be formulated as testing \( H_{00} : \xi_1(p) = \xi_2(p) \), for a given \( p \in (0,1) \). Here a precedence type test based on \( \hat{V}_p \) can be used, however, one needs to estimate the quantity \( \phi = \frac{f_2(\xi_2(p))}{f_1(\xi_1(p))} \), under the null hypothesis. Although this can be done using more sophisticated density estimation techniques, one can construct a simple yet consistent estimator by proceeding as follows.

Note that under the null hypothesis \( \phi \) can be written as \( \phi^* = \frac{f_2(\xi_2(p))}{f_1(\xi_1(p))} \), so that one can construct an estimator of \( \phi \) by estimating the numerator and the denominator separately. Chakraborti (1988) considered a simple estimator of \( f_i^{-1}(\xi_i(p)) \), using the length of an approximately \( 100(1 - \alpha) \) confidence interval for \( \xi_i(p) \). To describe this estimator let

\[ Q_i(+) = \hat{I}_i^{-1}(p_i(+)) \quad \text{and} \quad Q_i(-) = \hat{I}_i^{-1}(p_i(-)), \quad (10) \]

where

\[ p_i(+) = p + z_{\alpha/2} \left( \hat{I}_i/n_i \right)^{1/2} \quad \text{and} \quad p_i(-) = p - z_{\alpha/2} \left( \hat{I}_i/n_i \right)^{1/2}, \]

and

\[ \hat{I}_i = (1 - p)^2 \sum_{j : Z_i \leq \hat{I}_i^{-1}(p)} \frac{n_i d_{ij}}{R_{ij}(R_{ij} - d_{ij})}, \quad (11) \]

\( i = 1, 2 \). Using results in Breslow and Crowley (1974) and Cheng (1984) it can be shown that
It follows that a consistent estimator of $\phi^{*}$ is

$$h = \frac{\sqrt{n_1 I_1^2(Q_1(+) - Q_1(-))}}{\sqrt{n_2 I_2^2(Q_2(+) - Q_2(-))}}$$

provided the same $\alpha$ is used in both confidence intervals. An approximately size $\alpha$ test of $H_{00}$ against, say, $H_{01}: \xi_2(p) > \xi_1(p)$ is to reject $H_{00}$ in favour of $H_{01}$ if

$$\frac{N^{-1/2}(\hat{V}_p - n_2 p)}{\hat{\sigma}_{00}} < -z_{\alpha},$$

where

$$\hat{\sigma}_{00}^2 = \frac{\lambda_{2,N}}{\lambda_{1,N}}(\lambda_{1,N} I_0^2 + h^2 \lambda_{2,N} I_0^2).$$

2.2 Confidence Estimation of $\nu_p$ A related problem is the estimation of the quantity $\nu_p$, the probability that a $Y$-observation will be less than or equal to the selected $p$-quantile of the $X$-population. Such a quantity may be important in judging the efficacy of one treatment over another. For example, if $\nu_p$ is greater than, say .5, one may conclude that treatment 2 (corresponding to $F_2$) is better than treatment 1 (corresponding to $F_1$). To understand this better, let $F_1(x)$ and $F_2(y)$ correspond to exponential distributions with means $\theta_1$ and $\theta_2$, respectively. Then $\nu_p = 1 - (1 - p)^{\theta_1/\theta_2}$, so that for uncensored data one can find the "best" confidence interval for $\nu_p$, working with the respective sample means, which are the UMVUE's of the populations means, $\theta$'s. In the present case, however, our approach is nonparametric, that is we are working with some continuous distribution for the lifetimes, moreover, the observations are subject to possibly unequal right-censorship.

Chakraborti and Mukerjee (1989) proposed and studied a confidence interval for this problem with uncensored or complete data. Chakraborti (1984), in his unpublished dissertation, considered a large sample confidence interval for $\nu_p$ in the presence of right-censored data. Later simulations, however, suggested that the proposed interval may be too liberal.

As noted in CV, with complete data, the quantity $\nu_p$ also arises in the study of two-sample P-P plots, a well known graphical tool useful for assessing, for example, the equality of two distributions. In an analogous manner, a study of the present problem is related to a two-sample P-P plot for randomly right-censored data. In this context it may be noted that for complete data, the stochastic process $E_{n_1,n_2}(p) =$
\[
\left(\frac{\min n_1, n_2}{n_1 + n_2}\right)^{1/2} \left\{ F_{n_2}F_{n_1}^{-1}(p) - F_2F_1^{-1}(p) \right\}, \text{ with } p \in [0,1],
\]
is called the two-sample (empirical) P-P process. Large sample properties of this process have been studied in the empirical (stochastic) process literature, which has seen a tremendous growth over the last several years. For example, it is well known that when \( F_1 = F_2 \), and \( n_1, n_2 \to \infty \), the process \( E_{n_1,n_2}(p) \) converges (weakly) to a "Brownian Bridge". Thus, asymptotically, under \( H_0 \), the mean and the variance of \( E_{n_1,n_2}(p) \), for some \( p \in [0,1] \), is \( 0 \) and \( p(1-p) \), respectively. It follows immediately that under \( H_0 \), the large sample distribution of \( n_2^{-1/2}(n_2F_{n_2}F_{n_1}^{-1}(p) - n_2F_2F_1^{-1}(p)) \) is normal with mean 0 and variance \( p(1-p)(1 + \frac{2}{n_2}) \), a result previously noted in Remark 1. Thus the empirical process approach provides an (easier and) alternative way of deriving the large sample distributional results. For a comprehensive account of many of the developments in the area of empirical processes, the reader is referred to the book by Shorack and Wellner (1986).

In the presence of random right-censorship, it would be natural to study the analogous two-sample KM-P-P process \( \left(\frac{\min n_1, n_2}{n_1 + n_2}\right)^{1/2} \left\{ \hat{F}_2 \hat{F}_1^{-1}(p) - F_2F_1^{-1}(p) \right\} \). The results obtained from such a study would enable us to obtain a deeper understanding of the properties of precedence tests and related confidence intervals. For example, it would be possible to obtain a large sample confidence interval for \( \nu_p \) for a given \( p \), and to obtain a confidence band about \( F_2F_1^{-1}(p) \), which would provide more information about the nature of the differences (if any) between \( F_1 \) and \( F_2 \).

### 3 K-Sample Problems

Suppose that independent random samples \( X_{11}, X_{12}, ..., X_{1n_1}, ..., X_{K1}, X_{K2}, ..., X_{Kn_K} \), of observations (lifetimes) are available from absolutely continuous distribution functions \( F_1, F_2, ..., F_K \), respectively. Let \( G_1, G_2, ..., G_K \), be the censoring distributions corresponding to the \( K \) distributions. As in the two-sample case assume that for each sample, the censoring variables are independently distributed of the lifetimes. In our applications of interest, the random variables are positive valued, so that, \( F_i(0) = 0, i = 1, 2, ..., K \). Let \( \hat{F}_i(x) \) denote the KM estimator of \( F_i(x) \) and let \( \hat{F}_i^{-1}(u) \) be the (K-M) quantile function corresponding to \( \hat{F}_i \). For a given \( p \in (0,1) \), let \( V_{ip} = n_i \hat{F}_i \hat{F}_1^{-1}(p) \) be the \( \hat{V}_p \) statistic for sample \( i = 2, 3, ..., K \). The various procedures discussed below are based on these statistics.

In many \( K \)-sample problems the null hypothesis of interest is the homogeneity of the distributions,

\[
H_0 : F_1(x) = F_2(x) = ... = F_K(x),
\]
for all \( x \). Several alternatives to homogeneity can be considered. Among these we have the global alternative

\[
H_1 : F_i(x) \neq F_j(x),
\]
for at least one pair of \( i \) and \( j \) and some \( x \). As has been noted before, in some applications one of the distributions (say \( F_1 \)) may correspond to a control population and the question arises if any of the populations corresponding to \( F_2, F_3, ..., F_K \), are better than
the control. Assuming larger means better, this one-sided alternative can be expressed as

\[ H_2 : F_i(x) \leq F_1(x), \]

with strict inequality for at least one \( i = 2, 3, \ldots, K \), and some \( x \).

As noted earlier, sometimes it might be the situation that one is not interested in a comparison of the entire distribution functions but only at some specific point. This leads to a comparison of \( p \)-quantiles from the \( K \) populations and the null hypothesis reduces to

\[ H_0^* : \xi_i(p) = \xi_2(p) = \ldots = \xi_K(p), \]

where \( \xi_i \) is the \( p \)-quantile of \( F_i \). In this case the alternatives \( H_1 \) and \( H_2 \), reduce, respectively, to

\[ H_1^* : \xi_i(p) \neq \xi_j(p), \]

for at least one pair of \( i \) and \( j \), and

\[ H_2^* : \xi_i(p) \geq \xi_1(p), \]

with strict inequality for at least for at least one \( i = 2, 3, \ldots, K \). We shall discuss applications of precedence type tests for each of these problems.

We begin with a generalization of Theorem 2.1 to the case of several groups. For a proof of this result the reader might refer to Chakraborti (1984).

Suppose that \( n_1, n_2, \ldots, n_K \to \infty \), such that \( n_i/N \to \lambda_i, 0 < \lambda_i < 1, i = 1, 2, 3, \ldots, K \). Further assume that \( f_i = F_i' \) exists and that \( 0 < \phi_i = \frac{f_i(\xi_i(p))}{f_i(\xi_i(p))} < \infty \), for \( i = 2, 3, \ldots, K \). Under certain regularity conditions and the random censorship model, the asymptotic distribution of \( N^{-1/2} \tilde{U} \) is a \( K - 1 \) dimensional normal distribution with mean vector \( \bar{0} \) and dispersion matrix \( \Gamma = (\gamma_{ij}) \), where

\[
\gamma_{ij} = \sigma_{i1} \phi_{i+1} \phi_{j+1} \lambda_{i+1} \lambda_{j+1}^{-1} + \delta_{ij} \sigma_{i+1,i+1} \lambda_{i+1}^{-1}, \quad i, j = 1, 2, ..., K - 1, \]

\[
\sigma_{ii} = [1 - F_i(\xi_i(p))]^2 \int_0^{\xi_i(p)} \frac{dF_i}{(1 - F_i)^2(1 - G_i)}, \]

and \( \delta_{ij} \) is the kronecker's delta. Under \( H_0 \), we have \( \phi_i = 1 \) and

\[
\sigma_{ii} = \sigma_{ii}^0 = (1 - p)^2 \int_0^{\xi_i(p)} \frac{dF}{(1 - F)^2(1 - G_i)}, \]

where \( \xi_i(p) = F_i^{-1}(p) \) and \( F \) is the common but unknown distribution function under \( H_0 \). Therefore, under \( H_0 \)

\[ \gamma_{ij} = \gamma_{ij}^0 = \sigma_{i1}^0 \lambda_{i+1} \lambda_{j+1} \lambda_{i+1}^{-1} + \delta_{ij} \sigma_{i+1,i+1}^0 \lambda_{i+1}. \]

Let \( \Gamma_0 = ((\gamma_{ij}^0)) \) denote the dispersion matrix \( \Gamma \) under \( H_0 \). Further let \( \tilde{U}_0 = (V_2 - n_2 p, V_3 - n_3 p, \ldots, V_K - n_K p)' \). The following result follows directly from Theorem 3.1.

10
Corollary 3.1.1  Under $H_0$ and the assumptions of Theorem 3.1, the asymptotic distribution of $N^{-1/2} \mathcal{U}_0$ is a $K-1$ dimensional normal with mean $\mathbf{0}$ and dispersion matrix $\Gamma_0$.  

The null dispersion matrix can be consistently estimated by estimating $\sigma^2_{ii}$ by the Greenwood estimator $\hat{\sigma}^2_{ii}$, given in (2). It follows that a consistent estimator of $\Gamma_0$ is $\hat{\Gamma}_0 = ((\hat{\sigma}^2_{ij}))$, where

$$\hat{\sigma}^2_{ij} = \frac{1}{N-1} \frac{1}{\lambda_{i+1,N} \lambda_j + \lambda_{j+1,N} \lambda_i + \delta_{ij} \frac{1}{\lambda_{i+1,N} \lambda_j}}.$$  

(16)

Chakraborti (1984), in his unpublished dissertation, proposed a test of $H_0$ against $H_1$ based on $Q_N = N^{-1} \mathcal{U}_0^\top \hat{\Gamma}_0^{-1} \mathcal{U}_0$. It can be shown that, under $H_0$, the asymptotic distribution of $Q_N$ is a chi-square distribution with $K-1$ degrees of freedom. Thus $H_0$ is rejected in favour of $H_1$ at approximately size $\alpha$ if

$$Q_N > \chi^2(\alpha, K-1),$$

(17)

where $\chi^2(\alpha, K-1)$ is the upper $100\alpha$-percentile of the chi-square distribution with $K-1$ degrees of freedom. The same test is also mentioned in Chakraborti and Desu (1990), where some related remarks can be found.

Next consider testing $H_0$ against the one-sided alternative $H_2$. This problem has been studied in Chakraborti (1984) and in Chakraborti and Desu (1990). A generalized Mathisen test is based on $W_N = \sum_{i=2}^{K} (V_{ip} - \lambda_{i,N})$. From Theorem 3.1 it follows that under $H_0$, the distribution of $N^{-1/2} \mathcal{W}$ can be approximated by a normal distribution with mean 0 and variance

$$\sigma^2_{i} = \frac{(1-\lambda_1)^2}{\lambda_1} \sigma^2_{11} + \sum_{i=2}^{K} \lambda_i \sigma^2_{ii}.$$  

(18)

Again, the asymptotic null variance $\sigma^2_{i}$ can be consistently estimated from the data by

$$\hat{\sigma}^2_{i} = \frac{(1-\lambda_{1,N})^2}{\lambda_{1,N}} \hat{\sigma}^2_{11} + \sum_{i=2}^{K} \lambda_i \hat{\sigma}^2_{ii}.$$  

(19)

Hence an approximately size $\alpha$ test is to reject $H_0$ in favor of $H_2$ if

$$W^* = \frac{N^{-1/2} \mathcal{W}}{\hat{\sigma}_1} < -z_\alpha.$$  

(20)

Remark 4  In some applications it might be reasonable to assume that $G_1 = G_2 = \ldots = G_K = G$, say, that is, the censoring variables for the different groups are identically distributed. This situation is often referred to as the "equal censorship" (EC) case in the literature. In the EC case, many of the formulas given above simplify. To this end note that in the EC case under $H_0$,
The following corollary to Theorem 3.1 will be useful.

**Corollary 3.1.2** Under the null hypothesis and equal censorship, the asymptotic distribution of \( N^{-1/2} \mathbf{U}_0 \) is a \( K - 1 \) dimensional normal distribution with mean \( \mathbf{0} \) and dispersion matrix

\[
\sigma^2 \mathbf{\lambda}_{i+1} (\mathbf{\lambda}_{j+1} \mathbf{\lambda}_i^{-1} + \mathbf{\delta}_{ij}).
\]

For the problem of testing \( H_0 \) against \( H_1 \) with uncensored data, Slivka (1970) proposed a multiple comparisons type test based on the union-intersection principle. Chakraborti (1984, 1990a) considered an extension of Slivka’s test using the minimum of \( V_{2p}, V_{3p}, \ldots, V_{Kp} \) as the test statistic and provided an expression for the approximate P-value of the test for the equal censorship case. When there is no censorship, Chakraborti’s test reduces to Slivka’s test. We now provide the details about the calculation of the approximate P-value. Let \( v_0 \) be the observed value of the minimum (the test statistic) and let \( n_1 = n \) and \( n_i = n_s \), for \( i = 2, 3, \ldots, K - 1 \).

\[
\begin{align*}
P\text{-value} & = P[ \min(V_{2p}, V_{3p}, \ldots, V_{Kp}) < v_0 | H_0] \\
& = P[\max(- (V_{2p} - nsp), -(V_{3p} - nsp), \ldots, -(V_{Kp} - nsp)) < v_0 | H_0] \\
& = P[\max(- \frac{N^{-1/2} U_{01}}{\tau^{1/2}}, - \frac{N^{-1/2} U_{02}}{\tau^{1/2}}, \ldots, - \frac{N^{-1/2} U_{0,K-1}}{\tau^{1/2}}) > v_0^* | H_0],
\end{align*}
\]

say, where \( v_0^* = -N^{-1/2}(\nu_0 - nsp)/\tau^{1/2} \), \( \tau = \frac{s(s+1)}{1+4s(K-1)} \sigma^2 \) and \( \sigma^2 \) is given in (21). Using Corollary 3.1.2 it follows that, under \( H_0 \) and EC, the random variables \( -\frac{N^{-1/2} U_{01}}{\tau^{1/2}}, -\frac{N^{-1/2} U_{02}}{\tau^{1/2}}, \ldots, -\frac{N^{-1/2} U_{0,K-1}}{\tau^{1/2}} \) are approximately normally distributed with mean 0, variance 1, and common correlation \( \rho = s(s + 1)^{-1} \). Thus the P-value can be approximated by using a consistent estimator of \( \tau \), say \( \tilde{\tau} \), (this involves consistent estimation of \( \sigma^2 \) by an appropriate Greenwood estimator; see remark 5 below) and evaluating the c.d.f. of the maximum of \( K - 1 \) equicorrelated standard normal random variables at \( \tilde{v}_0^* = -N^{-1/2}(\nu_0 - nsp)/\tilde{\tau}^{1/2} \). Gupta (1963) has tabulated the probability that the maximum of \( K \) equicorrelated standard normal random variables doesn’t exceed \( H \) for \( H = -3.5 \) for \( K = 1(1)12 \), and some 17 values of \( \rho \). For values outside the range in this table one could use the programs developed by Dunnett (1989,1993). In some applications it might be more convenient to have the critical value \( C_\alpha \), such that the size of the test is \( \alpha \), fixed in advance. Gupta (1963) has also tabulated the upper 100\( \alpha \)-percentage point of the distribution of the maximum when the common correlation equals 0.5. If the sample sizes are all equal (\( s = 1 \)) we have

\[
C_\alpha \approx np - (2N/K)^{1/2} \hat{\nu} G_{\alpha,K-1}, \tag{23}
\]

where \( G_{\alpha,K-1} \) is the quantity tabulated by Gupta for \( \alpha = 0.01, 0.025, 0.05, 0.010, 0.25 \) and \( k = 2(1)51 \). Chakraborti (1990a) reported the results of some simulation studies.
about the power of the proposed test. It was seen that the test performs rather well under moderate censorship, for sample sizes around 20. The reader is referred to his paper for details.

**Remark 5** Some comments about estimating $\sigma^2$ are in order. The problem is how to incorporate the null hypothesis in the estimation. First note that the quantity $\xi(p)$ is the common (but unknown) $p$-quantile of the $K$ populations under the null hypothesis. Clearly, there can be more than one consistent estimator of $\xi(p)$, and the question is whether or not one should use information from the $i$th sample only, or from some, or all the samples. For example, one can use $\hat{F}_i^{-1}(p)$, or $\tilde{F}_i^{-1}(p)$ or perhaps an average of $\hat{F}_i^{-1}(p), \ldots, \tilde{F}_K^{-1}(p)$. Each of these can be shown to be a consistent estimator of $\xi(p)$, under the null hypothesis. Secondly, there is the related question of how to estimate $F$. Intuitively one would like to utilize information from all the samples but the presence of unequal censorship $(G_i)$ makes things difficult. Of course, when there is equal censorship it would seem reasonable to pool information from all of the samples to estimate both $F$ and $\xi(p)$. Chakraborti (1990b) used the estimator in (2) with $\hat{F}_i^{-1}(p)$ for $\hat{F}^{-1}(p)$. His simulation results indicate that under moderate censorship and for equal sample sizes this is quite reasonable.

### 3.2 Testing Equality of $K$ Quantiles

Now consider testing the homogeneity of $K$ $p$-quantiles in the presence of random right-censorship. This is an extension of the work in section 2.1. Chakraborti (1990b) considered a class of tests which may be viewed as a generalization of tests considered by Sen (1962). The proposed tests can be used when the censoring distributions aren't necessarily equal and do not require the scale or the shape parameters of the underlying distributions to be either known or equal. The key here is to find a consistent estimator of the dispersion matrix $\Gamma$, under the null hypothesis $H_0$. Evidently, this requires consistent estimators of $\phi_{ii}$ (since this does not equal 1 as in the case of testing $H_0$) in addition to $\sigma_{ii}$ under $H_0$, say,

$$\sigma_{ii}^* = (1 - p)^2 \int_0^{\xi(p)} \frac{dF_i}{(1 - F_i)^2(1 - G_i)}.$$

One could use a Greenwood type estimator to estimate $\sigma_{ii}^*$. Chakraborti (1990b) suggested using

$$\hat{\sigma}_{ii}^* = (1 - \hat{F}_i(\hat{F}_i^{-1}(p)))^2 \sum_{j:Z_{ii} = \hat{F}_i^{-1}(p)} \frac{n_id_{ij}}{R_{ij}(R_{ij} - d_{ij})},$$

which is slightly different from the Greenwood estimator given in (2), since it uses a leading term different from $(1 - p)^2$. In large samples, however, the difference is likely to be negligible.

The main problem then is to find a consistent estimator for the quantity $\phi_i$ under $H_0$,
\[ \phi_i = \phi_i^* = \frac{f_i(\xi(p))}{f_i(\xi(p))}, \quad i = 2, 3, ..., K, \]  

(26)

where \( \xi(p) \) is the common value of the \( p \)-quantiles under \( H^*_0 \). As noted in section 2.1, Chakraborti (1988) proposed a consistent estimator for \( \phi_i^* \) given by

\[ h_i = \frac{\sqrt{n_i} \hat{\delta}_i (Q_i (+) - Q_i (-))}{\sqrt{n_i} \hat{\delta}_i (Q_i (+) - Q_i (-))}, \quad i = 2, 3, ..., K, \]

(27)

based on the lengths of 100(1 - \( \alpha \)) large sample confidence intervals for \( \xi_i(p) \) and \( \xi_i(p) \), respectively. The quantities \( Q_i (+) \), \( Q_i (-) \) and \( \hat{\delta}_i \) are given in (10) and (11), respectively. It follows that a consistent estimator of \( \Gamma = \Gamma^* = ((\gamma^*_{ij})) \), under \( H^*_0 \) is \( \hat{\Gamma}^* = ((\hat{\gamma}^*_{ij})) \), where

\[ \hat{\gamma}^*_{ij} = \hat{V}_{ij}^0 h_{i+1} h_{j+1} \lambda_{i+1,N} \lambda_{j+1,N}^{-1} + \delta_{ij} \hat{V}_{i+1,i+1}^0. \]

(28)

Chakraborti (1990b) proposed a test of \( H^*_0 \) against \( H^*_1 \) based on \( T_N = N^{-1} \tilde{U}_0^T \hat{\Gamma}^* \tilde{U}_0 \). Using corollary 3.1.1 it can be shown that under \( H^*_0 \) the asymptotic distribution of \( T_N \) is a chi-square distribution with \( K - 1 \) degrees of freedom. It can also be shown that the inverse of the matrix \( \hat{\Gamma}^* \) exists if \( \hat{V}_{ii}^0 \) is positive for \( i = 2, 3, ..., K \). Thus the test is to reject \( H^*_0 \) if

\[ T_N > \chi^2(\alpha, K - 1). \]

(29)

Because of the special structure of \( \hat{\Gamma}^* \) it's inverse can be explicitly obtained. After some routine calculations, the test statistic \( T_N \) can be expressed as

\[ T_N = \sum_{i=1}^{K} \frac{(U_i - n_i p)^2}{n_i \hat{V}_{ii}^0} - \frac{(\sum_{i=2}^{K} (U_i - n_i p) h_i)}{\sum_{i=1}^{K} n_i h_i^2 \hat{V}_{ii}^0}, \]

(30)

which is convenient for computing purposes. Note that if the null hypothesis is \( H_0 \), the homogeneity of the distributions, then \( h_i = 1 \) and the test reduces to the test in (17).

4 Acknowledgement

Support for this research was provided in part by NATO collaborative research grant no. CRG 920287. Research partially supported by European Union HCM grant ERB CHRX-CT 940693.
5 References


Mathisen, H. C. (1943): A method for testing the hypothesis that two samples are from the same population. *Annals of Mathematical Statistics* 14, 188-196.


