THE UNIVERSITY OF TEXAS AT SAN ANTONIO, COLLEGE OF BUSINESS

Working Paper SERIES

May 30, 2008

Wp# 0054MSS-213-2008

Precedence-type Test based on Progressively Censored Samples

N. Balakrishnan
Department of Mathematics and Statistics, McMaster University, Hamilton, Ontario, Canada

R. C. Tripathi
Department of Management Science and Statistics, University of Texas at San Antonio, One
University Circle, San Antonio, Texas 78249, USA

N. Kannan
Department of Management Science and Statistics, University of Texas at San Antonio, One
University Circle, San Antonio, Texas 78249, USA

H. K. T. Ng
Department of Statistical Science, Southern Methodist University, Dallas, Texas 75275-0332,
USA

Department of Management Science & Statistics,
University of Texas at San Antonio,
San Antonio, TX 78249, U.S.A

Precedence-type Test based on Progressively Censored Samples

N. Balakrishnan*, R. C. Tripathi**, N. Kannan**, H. K. T. Ng†

Abstract

In this paper, we introduce precedence-type tests for testing the hypothesis that two distribution functions are equal, which is an extension of the precedence life-test first proposed by Nelson (1963), when the two samples are progressively Type-II censored. The null distributions of the test statistics are derived. Critical values for some combination of sample sizes and censoring schemes for the proposed tests are presented. Then, we present the exact power functions under the Lehmann alternative, and compare the exact power as well as simulated power (under location-shift) of the proposed precedence test based on nonparametric estimates of CDF with other precedence-type tests. We then examine the power properties of the proposed test procedures through Monte Carlo simulations. Two examples are presented to illustrate all the test procedures discussed here. Finally, we make some concluding remarks.

Keywords: Precedence test; Product-limit estimator; Type-II progressive censoring; life-testing; level of significance; power; Lehmann alternative; Monte Carlo simulations.

JEL Code: C14

AMS Subject Classification. Primary: 62G10, 62N05; Secondary: 62P30

*Postal address: Department of Mathematics and Statistics, McMaster University, Hamilton, Ontario, Canada L8S 4K1

**Postal address: Department of Management Science and Statistics, University of Texas at San Antonio, One University Circle, San Antonio, Texas 78249, USA

†Postal address: Department of Statistical Science, Southern Methodist University, Dallas, Texas 75275-0332, USA

Acknowledgements: This research was supported by the College of Business Summer Research Grant for 2007, University of Texas at San Antonio, San Antonio, Texas.

1 Introduction

In many reliability and survival analysis studies, it is common to compare two or more populations. For example, while comparing a treatment with the control, one may be interested in assessing whether the population corresponding to the treatment has a longer life than the control population. Similarly, in reliability studies, one may be interested in inferring whether the components manufactured under a new design last longer than those manufactured under the standard design. In studies where experimental units are expensive, it is desirable to make decisions based on early failures and use the remaining units for some other purpose. The precedence test, first proposed by Nelson (1963), is a test for comparing two populations based on the order of early failures. It is a distribution-free test which allows a simple and robust comparison of two distribution functions. Suppose there are two failure time distributions F_X and F_Y and that we are interested in testing

$$H_0 : F_X = F_Y \quad \text{against} \quad H_1 : F_X > F_Y. \tag{1.1}$$

Note that some specific alternatives such as the location-shift alternative and the Lehmann alternative are subclasses of the general alternative considered in (1.1).

Various precedence type tests such as weighted precedence and maximal precedence tests have been developed in the literature. For a detailed discussion, see Balakrishnan and Ng (2006). These tests are developed for the situation when one of the samples, say, the Y-sample is progressively censored. Balakrishnan, Tripathi and Kannan (2007) developed a precedence test for the above hypothesis when both the samples are progressively censored. In this paper, they derived the exact null-distribution of the proposed test statistic and provided tables giving critical values and the corresponding significance levels for certain combination of sample sizes and censoring schemes.

In this paper we present two new precedence type tests when both the samples are progressively censored. The first one is a Wilcoxon-type Rank-sum precedence test, and the other is a precedence test based on the Kaplan-Meier estimator of the survival function. In section 2, we discuss the progressive censoring and placement statistics. In section 3, we derive the joint probability mass function (pmf) of the placement statistics under the null hypothesis. In section 4, we present the precedence statistic and its null distribution as in Balakrishnan, Tripathi and

Kannan (2007) along with two new precedence type statistics and give their null distributions. We also present a table which gives the critical values with significance levels close to 5% under various sampling schemes. We also derive the exact power function of the three tests. In section 5, we derive the joint pmf of the placements under the Lehmann alternatives and use it to compute exact power. We also compute the power of the three tests based on Monte Carlo simulation and compare them under various sampling and censoring schemes. Finally, in section 6 we draw some conclusions.

2 Progressive Type-II Right Censoring and Placement Statistic

Assume that a random sample of size n_1 is from distribution F_X, another independent sample of size n_2 is from distribution F_Y, and that all these sample units are placed simultaneously on a life-testing experiment. We use $X_1, X_2, \ldots, X_{n_1}$ to denote the sample from F_X, and $Y_1, Y_2, \ldots, Y_{n_2}$ to denote the sample from F_Y. A natural null hypothesis of interest is that the two failure time distributions are equal, and we are generally concerned with the alternative models where in one distribution is stochastically larger than the other; for example, the alternative that F_Y is stochastically larger than F_X.

In life-testing experiments, we may not always obtain complete information on failure times for all experimental units. Data obtained from such experiments are called *censored data*. The most common censoring schemes are Type-I and Type-II censoring, but the conventional Type-I and Type-II censoring schemes do not have the flexibility of allowing removal of units at points other than the terminal point of the experiment. For this reason, we consider a more general censoring scheme called *progressive Type-II right censoring* which can be described as follows: consider an experiment in which n units are placed on a life-test. At the time of the first failure, R_1 units are randomly removed from the remaining $n-1$ surviving units. At the second failure, R_2 units from the remaining $n-2-R_1$ units are randomly removed. The test continues until the m^{th} failure at which time, all remaining $R_m = n-m-R_1-R_2-\ldots-R_{m-1}$ units are removed. The $R_i^{'}s$ are fixed prior to the study. For more details about the theory and applications of progressive censoring, one can refer to Balakrishnan and Aggarwala (2000) and

Balakrishnan (2007).

In the two-sample problem, we consider the case when the X- and Y-samples are progressively Type-II censored samples with censoring scheme $\mathbf{R} = (R_1, R_2, \ldots, R_{m_1})$ and $\mathbf{S} = (S_1, S_2, \ldots, S_{m_2})$. We denote the progressively Type-II censored order statistics from the X-sample and the Y-sample by $X^{(\mathbf{R})}_{1:m_1:n_1} \leq X^{(\mathbf{R})}_{2:m_1:n_1} \leq \ldots \leq X^{(\mathbf{R})}_{m_1:m_1:n_1}$ and $Y^{(\mathbf{S})}_{1:m_2:n_2} \leq Y^{(\mathbf{S})}_{2:m_2:n_2} \leq \ldots \leq Y^{(\mathbf{S})}_{m_2:m_2:n_2}$, respectively.

The i-th placement from the X-sample is denoted by U_i, which is the number of observed X-failures that fall between the $(i-1)$-th and the i-th observed Y-failures, $i = 1, \ldots, m_2 + 1$. That is, for a fixed value of i, $i = 1, \ldots, m_2 + 1$, U_i = number of $X^{(\mathbf{R})}_{j:m_1:n_1}$ such that $Y^{(\mathbf{S})}_{i-1:m_2:n_2} < X^{(\mathbf{R})}_{j:m_1:n_1} < Y^{(\mathbf{S})}_{i:m_2:n_2}$ with $Y^{(\mathbf{S})}_{0:m_2:n_2} \equiv 0$ and $Y^{(\mathbf{S})}_{m_2+1:m_2:n_2} \equiv +\infty$.

For notational convenience, we further denote the total number of observed X-failures before $Y^{(\mathbf{S})}_{l:m_2:n_2}$ as a partial sum $U_{(l)} = \sum_{i=1}^{l} U_i$ and the total number of failed and censored items from the X-sample between the $(l-1)$-th and the l-th observed Y-failures as $W_l = \sum_{i=U_{(l-1)}+1}^{U_{(l)}} (R_i + 1)$. Then the total number of observed failures and censored items from both X- and Y-samples right after $Y^{(\mathbf{S})}_{l:m_2:n_2}$ is $V_l = \sum_{k=1}^{l} (W_k + S_k + 1)$. A schematic representation of a precedence life-test with progressive censoring is presented in Figure 1. The quantities defined above will be used in developing the three statistics in the next section.

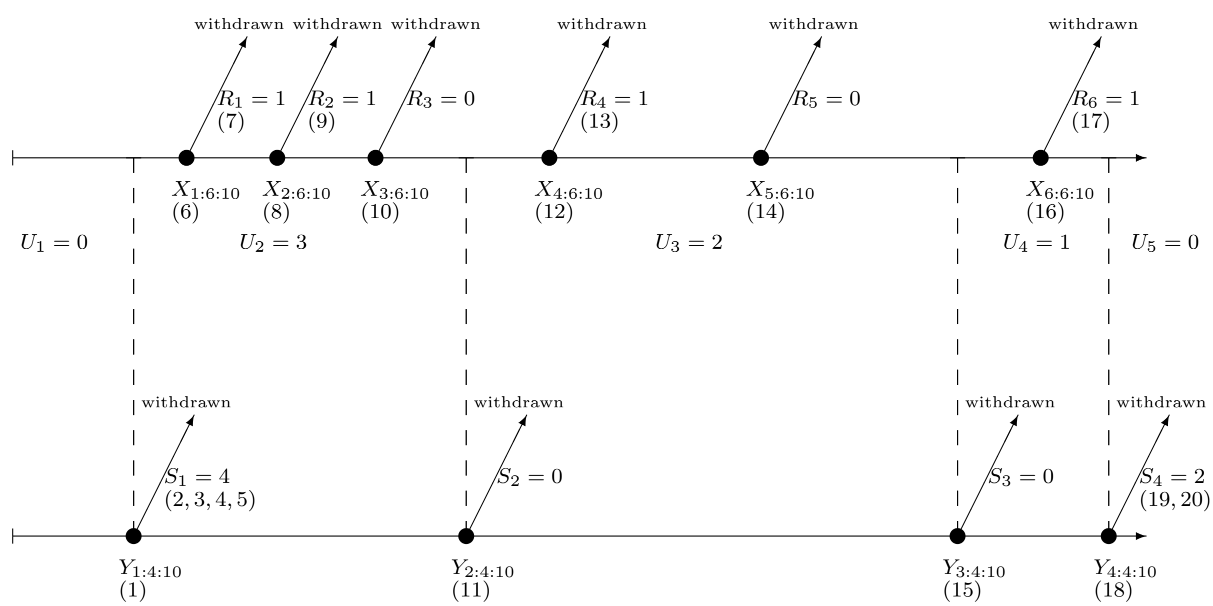

Figure 1: Schematic representation of a precedence life-test with progressive Type-II censoring

3 Probability Mass Functions of the Placement Statistics

Here, we present the joint pmf of the placement statistics $U_1, U_2, \cdots, U_l$ under the null hypothesis.

From Balakrishnan, Tripathi and Kannan (2007), the joint probability mass function of (U_1, U_2, ..., U_{m_2}) is given by

$$\Pr(U_1 = u_1, \ldots, U_{m_2} = u_{m_2}) = AB \sum_{i_1=0}^{u_1} \cdots \sum_{i_{m_2}=0}^{u_{m_2}} \frac{\prod_{j=1}^{m_2+1} \gamma_{i_j,u_j}(R_{u_{(j-1)}+1}, \ldots, R_{u_{(j)}})}{\prod_{j=0}^{m_2-1}(N - T_{(j)} - j)}$$

where

$$\begin{aligned}
\gamma_{i_j,u_j}(R_{u_{(j-1)}+1}, \ldots, R_{u_{(j)}}) &= \frac{(-1)^{i_j}}{\left[\prod_{g=1}^{i_j} \sum_{k=u_j-i_j+1}^{u_j-i_j+g} (R_{u_{(j-1)}+k}+1)\right] \left[\prod_{g=1}^{u_j-i_j} \sum_{k=g}^{u_j-i_j} (R_{u_{(j-1)}+k}+1)\right]}, \\
T_{(0)} &= 0, T_{(j)} = \sum_{k=1}^{j} T_k, j = 1, 2, \ldots, \\
T_k &= S_k + \sum_{j_k=u_{(k)}-i_k+1}^{u_{(k+1)}-i_{k+1}} R_{j_k} + (i_k + u_{k+1} - i_{k+1}), k = 1, \ldots, m_2, \\
N &= m_2 + \sum_{i=1}^{m_2} T_i, \\
A &= n_1(n_1 + R_1 - 1) \cdots (n_1 - R_1 - \cdots - R_{m_1-1} - m_1 + 1), \\
B &= n_2(n_2 + S_1 - 1) \cdots (n_2 - S_1 - \cdots - S_{m_2-1} - m_2 + 1).
\end{aligned}$$

We now use this joint pmf in developing the three precedence type tests.

4 Proposed Precedence-type Tests

4.1 Precedence Test

The precedence test statistic $P_{(m_2)}$ is simply defined as the number of observed failures from the X-sample that precede the r-th observed failure from the Y-sample, i.e., $P_{(m_2)} = U_{(m_2)} = \sum_{i=1}^{m_2} U_i$. Large values of $P_{(m_2)}$ lead to the rejection of H_0 and in favor of H_1 in (1.1). The probability mass function of the precedence test statistic $P_{(m_2)}$ under the null hypothesis in (1.1) is

$$\Pr(P_{(m_2)} = p | F_X = F_Y) = \sum_{\substack{u_i (i=1,2,\ldots,m_2)=0 \\ u_{(m_2)}=p}}^{m_1} \Pr(U_1 = u_1, U_2 = U_2, \ldots, U_{m_2} = u_{m_2} | F_X = F_Y). \quad (4.1)$$

The p-value of the test can be computed from this formula. For example, from Figure 1, with $n_1 = n_2 = 10$, $m_1 = 6$, $m_2 = 4$, $\mathbf{R} = (1, 1, 0, 1, 0, 1)$, $\mathbf{S} = (4, 0, 0, 2)$ and $U_1 = 0$, $U_2 = 3$, $U_3 = 2$, $U_4 = 1$, the precedence test statistic takes on the value $P_{(4)} = \sum_{i=1}^{4} U_i = 0 + 3 + 2 + 1 = 6$ with p-value 0.24255.

For a fixed level of significance α, the critical region for the precedence test will be $\{p, p + 1, \ldots, m_1\}$, where

$$\alpha = \Pr(P_{(m_2)} \geq p | F_X = F_Y). \quad (4.2)$$

The critical values s and the exact level of significance α (as close as possible to 5%) for different choices of the sample sizes n_1 and n_2, effective sample sizes m_1 and m_2 and censoring schemes $\mathbf{R}$ and $\mathbf{S}$ are presented in Table 1.

4.2 Wilcoxon-type Rank-sum Precedence Test

The Wilcoxon rank-sum test is a well-known nonparametric procedure for testing the hypotheses in (1.1) based on complete samples. For testing the hypotheses in (1.1), if complete samples of size n_1 and n_2 were available from F_X and F_Y, respectively, one can use the standard Wilcoxon's rank-sum statistic, proposed by Wilcoxon (1945), which is simply the sum of ranks of X-observations in the combined sample.

Ng and Balakrishnan (2002) proposed the Wilcoxon-type rank-sum precedence tests for testing the hypotheses in (1.1) when the Y-sample is Type-II censored. This test is a variation of the precedence test and a generalization of the Wilcoxon rank-sum test. In order to test the hypotheses in (1.1), one could use the sum of the ranks of those failures. The Wilcoxon's rank-sum test statistic is computed under the assumption that all the censored items are failed instantaneously after the censoring occurs. For example, suppose $Y^{(\mathbf{S})}_{i-1:m_2:n_2} < X^{(\mathbf{R})}_{j:m_1:n_1} < Y^{(\mathbf{S})}_{i:m_2:n_2}$, the R_j censored items are assumed to fail between $X^{(\mathbf{R})}_{j:m_1:n_1}$ and $Y^{(\mathbf{S})}_{i:m_2:n_2}$. The test statistic in this case would be the sum of ranks of X-observations in the combined sample

Table 1: Near 5% critical values (c.v) and exact levels of significance (l.o.s.) for $P_{(m_2)}$, T_{W,m_2} and $\bar{Q}_{(m_2)}$

							$P_{(m_2)}$		T_{W,m_2}		$\bar{Q}_{(m_2)}$	
Setting	n_1	n_2	m_1	m_2	**R**	**S**	c.v.	l.o.s.	c.v.	l.o.s.	c.v.	l.o.s.
1	10	10	5	5	(5, 0, 0, 0, 0)	(5, 0, 0, 0, 0)	5	0.500	75	0.045	5	0.136
2	10	10	5	5	(3, 2, 0, 0, 0)	(3, 2, 0, 0, 0)	5	0.500	73	0.053	5	0.134
3	10	10	5	5	(1, 1, 1, 1, 1)	(1, 1, 1, 1, 1)	5	0.500	71	0.048	5	0.167
4	10	10	5	5	(0, 3, 0, 0, 2)	(0, 3, 0, 0, 2)	5	0.500	74	0.047	5	0.178
5	10	10	5	5	(3, 1, 1, 0, 0)	(3, 2, 0, 0, 0)	5	0.523	73	0.049	5	0.075
6	10	10	5	5	(3, 0, 0, 0, 2)	(1, 1, 1, 0, 2)	5	0.545	73	0.047	5	0.081
7	10	10	5	5	(1, 1, 1, 0, 2)	(3, 0, 0, 0, 2)	5	0.455	78	0.050	5	0.148
8	10	10	5	5	(0, 3, 0, 0, 2)	(1, 1, 1, 1, 1)	5	0.566	73	0.053	5	0.092
9	10	10	5	3	(3, 1, 1, 0, 0)	(3, 1, 3)	5	0.050	75	0.061	5	0.030
10	10	10	5	3	(0, 3, 0, 0, 2)	(3, 2, 2)	5	0.241	83	0.051	5	0.073
11	10	10	5	3	(3, 0, 0, 0, 2)	(2, 2, 3)	5	0.150	79	0.046	5	0.068
12	10	10	5	3	(5, 0, 0, 0, 0)	(5, 0, 2)	5	0.075	76	0.046	5	0.041
13	15	10	5	5	(10, 0, 0, 0, 0)	(5, 0, 0, 0, 0)	5	0.510	140	0.055	4	0.047
14	15	10	5	5	(0, 0, 10, 0, 0)	(0, 0, 5, 0, 0)	5	0.539	129	0.042	4	0.066
15	15	10	5	5	(2, 8, 0, 0, 0)	(1, 4, 0, 0, 0)	5	0.522	135	0.049	4	0.062
16	15	10	5	5	(2, 2, 2, 2, 2)	(1, 1, 1, 1, 1)	5	0.707	132	0.048	4	0.137
17	15	10	7	5	(8, 0, 0, 0, 0, 0, 0)	(5, 0, 0, 0, 0)	7	0.409	151	0.048	6	0.037
18	15	10	7	5	(2, 2, 2, 2, 0, 0, 0)	(3, 1, 1, 0, 0)	7	0.430	146	0.049	6	0.037
19	15	10	7	5	(2, 0, 2, 0, 2, 0, 2)	(2, 0, 2, 0, 1)	7	0.407	147	0.048	6	0.082
20	15	10	7	5	(2, 1, 1, 1, 1, 1, 1)	(1, 1, 1, 1, 1)	7	0.419	148	0.053	6	0.073
21	15	15	7	5	(8, 0, 0, 0, 0, 0, 0)	(10, 0, 0, 0, 0)	7	0.400	180	0.051	7	0.027
22	15	15	7	5	(2, 2, 2, 2, 0, 0, 0)	(6, 2, 2, 0, 0)	7	0.394	185	0.050	7	0.029
23	15	15	7	5	(2, 0, 2, 0, 2, 0, 2)	(4, 0, 4, 0, 2)	7	0.347	194	0.051	7	0.043
24	15	15	7	5	(2, 1, 1, 1, 1, 1, 1)	(2, 2, 2, 2, 2)	7	0.216	183	0.046	7	0.053

which is given by

$$T_{W,m_2} = \frac{1}{2}\sum_{k=1}^{m_2+1} W_k(W_k+1) + \sum_{k=2}^{m_2+1} W_k V_{k-1}.$$

Small values of T_{W,m_2} lead to the rejection of H_0 and in favor of H_1 in (1.1). The probability mass function of the Wilcoxon-type rank-sum precedence test statistic T_{W,m_2} under the null hypothesis in (1.1) is

$$\Pr(T_{W,m_2} = w | F_X = F_Y) = \sum_{\substack{u_i(i=1,2,\ldots,m_2)=0 \\ T_{W,m_2}=w}}^{m_1} \Pr(U_1 = u_1, U_2 = U_2, \ldots, U_{m_2} = u_{m_2} | F_X = F_Y). \quad (4.3)$$

The p-value of the test can be computed from this formula. For instance, in Figure 1, we have $U_{(1)} = 0$, $U_{(2)} = 3$, $U_{(3)} = 5$, $U_{(4)} = 6$, $U_{(5)} = 6$, $W_1 = 0$, $W_2 = 5$, $W_3 = 3$, $W_4 = 2$, $W_5 = 0$, $V_1 = 5$, $V_2 = 11$, $V_3 = 15$, $V_4 = 20$, the ranks of the observed failures and the censored items in the combined sample are in the parenthesis, the test statistic is given by

$$T_{W,4} = 6 + 7 + 8 + 9 + 10 + 12 + 13 + 14 + 16 + 17$$

$$
\begin{aligned}
&= \frac{1}{2}\left[0(0+1)+5(5+1)+3(3+1)+2(2+1)+0(0+1)\right] \\
&\quad +\left[5(5)+3(11)+2(15)+0(20)\right] \\
&= \frac{48}{2}+88=112,
\end{aligned}
$$

and the p-value is 0.3846.

For a fixed level of significance α, the critical region for the Wilcoxon-type rank-sum precedence test will be $\{n_1(n_1+1)/2, \ldots, w\}$, where

$$\alpha = \Pr(T_{W,m_2} \leq w | F_X = F_Y). \tag{4.4}$$

The critical values w and the exact level of significance α (as close as possible to 5%) for different choices of the sample sizes n_1 and n_2, effective sample sizes m_1 and m_2 and censoring schemes $\mathbf{R}$ and $\mathbf{S}$ are presented in Table 1.

4.3 Precedence Test Based on the Kaplan-Meier Estimator

The proposed precedence-type test is based on the Kaplan-Meier nonparametric estimator (Kaplan and Meier, 1958) of CDF for data with observations reported as exact failure times. First, we will review the Kaplan-Meier nonparametric estimator of CDF based on a Type-II progressively censored sample and a conventional Type-II censored sample as a special case.

Refer to the Type-II progressive censoring experimental scheme on the X-sample, we observed exact failures at $x_{1:m_1:n_1}$, $x_{2:m_1:n_1}$, $\ldots$, $x_{m_1:m_1:n_1}$. The Kaplan-Meier nonparametric estimate (also called product-limit estimates) of $F_X(x_{j:m_1:n_1})$ is given by

$$\hat{F}_X(x_{j:m_1:n_1}) = 1 - \prod_{k=1}^{j}\left(1 - \frac{1}{n^*_{xk}}\right), \tag{4.5}$$

$j = 1, \ldots, m_1$, where n^*_{xj} is the risk set at $x_{j:m_1:n_1}$ with $n^*_{xj} = n_1 - j + 1 - \sum_{k=0}^{j-1} R_k$. Similarly, the Kaplan-Meier nonparametric estimate of $F_Y(y_{i:m_2:n_2})$ is given by

$$\hat{F}_Y(y_{i:m_2:n_2}) = 1 - \prod_{k=1}^{i}\left(1 - \frac{1}{n^*_{yk}}\right), \tag{4.6}$$

$i = 1, \ldots, m_2$, where n^*_{yi} is the risk set at $y_{i:m_1:n_1}$ with $n^*_{yi} = n_2 - i + 1 - \sum_{k=0}^{i-1} S_k$.

Following the same idea of the precedence-type test procedures, let U_1 denote the number of observed X-failures before $Y_{1:m_2:n_2}$, U_i the number of observed X-failures between $Y_{i-1:m_2:n_2}$ and $Y_{i:m_2:n_2}$ for $i = 2, \cdots, m_1$, and Q_i the number of observed X-failures among the U_i that are between $Y_{i-1:m_2:n_2}$ and $Y_{i:m_2:n_2}$ for which $\hat{F}_X(x_{j:m_1:n_1}) > \hat{F}_Y(y_{i:m_2:n_2})$ for $i = 1, \cdots, m_2$. If the information after the termination of the experiment at $Y_{m_2:m_2:n_2}$ is not taken into account, it would be reasonable to consider the statistic $Q_{(m_2)} = \sum_{i=1}^{m_2} Q_i$ which can be expressed in terms of $(U_1, \cdots, U_{m_2})$ as

$$Q_{(m_2)}(\mathbf{U}) = \sum_{i=1}^{m_2} \sum_{j=U_{(i-1)}+1}^{U_{(i)}} I[\hat{F}_X(x_{j:m_1:n_1}) > \hat{F}_Y(y_{i:m_2:n_2})],$$

By assuming that all the remaining unobserved X-failures will fail before the censored items from the Y-sample at $Y_{m_2:m_2:n_2}$, we obtain the statistic

$$Q^*_{(m_2)}(\mathbf{U}) = \sum_{i=1}^{m_2+1} \sum_{j=U_{(i-1)}+1}^{U_{(i)}} I[\hat{F}_X(x_{j:m_1:n_1}) > \hat{F}_Y(y_{i:m_2:n_2})],$$

where $y_{m_2+1:m_2:n_2}$ is taken as the $(m_2 + 1)$-th progressively Type-II censored order statistic $y_{m_2+1:m_2+1:n_2}$ with progressive censoring scheme $\left(S_1, \cdots, S_{m_2-1}, 0, n_2 - m_2 - 1 - \sum_{i=1}^{m_2-1} S_i\right)$. Then, the test statistic we propose is the average of the two statistics given by

$$\begin{aligned}\bar{Q}_{(m_2)}(\mathbf{U}) &= \sum_{i=1}^{m_2} \sum_{j=U_{(i-1)}+1}^{U_{(i)}} I[\hat{F}_X(x_{j:m_1:n_1}) > \hat{F}_Y(y_{i:m_2:n_2})] \\ &\quad + \frac{1}{2} \sum_{j=U_{(m_2)}+1}^{m_1} I[\hat{F}_X(x_{j:m_1:n_1}) > \hat{F}_Y(y_{m_2+1:m_2:n_2})],\end{aligned}$$

with large values of $\bar{Q}_{(m_2)}$ leading to the rejection of H_0 and in favor of H_1 in (1.1). The probability mass function of $\bar{Q}_{(m_2)}(\mathbf{U})$ under the null hypothesis in (1.1) is

$$\Pr(\bar{Q}_{(m_2)}(\mathbf{U}) = q | F_X = F_Y) = \sum_{\substack{u_i(i=1,2,\ldots,m_2)=0 \\ \bar{Q}_{(m_2)}(\mathbf{U})=q}}^{m_1} \Pr(U_1 = u_1, U_2 = U_2, \ldots, U_{m_2} = u_{m_2} | F_X = F_Y) \quad (4.7)$$

The p-value of the test can be computed from this formula. For example, from Figure 1, the Kaplan-Meier estimates of the CDF based on the progressively Type-II censored X- and Y-samples are presented in Table 2, from which we observe $Q_1 = 0, Q_2 = 1, Q_3 = 2, Q_4 = 1, Q_5 = 0$ with which we obtain $Q_{(4)} = Q^*_{(4)} = 4$ and the proposed test statistic $\bar{Q}_{(4)} = 4$. The corresponding p-value is 0.2756.

Table 2: Kaplan–Meier Estimates of the CDF based on the progressively Type-II censored X- and Y-samples in Figure 1.

t_j	n^*_{xj}	$1/n^*_{xj}$	$1-(1/n^*_{xj})$	$\hat{F}_X(t_j)$
$X_{1:6:10}$	10	0.100	0.900	0.10000
$X_{2:6:10}$	8	0.125	0.875	0.21250
$X_{3:6:10}$	6	0.167	0.833	0.34375
$X_{4:6:10}$	5	0.200	0.800	0.47500
$X_{5:6:10}$	3	0.333	0.667	0.35000
$X_{6:6:10}$	2	0.500	0.500	0.82500
t_i	n^*_i	$1/n^*_i$	$1-(1/n^*_i)$	$\hat{F}_Y(t_i)$
$Y_{1:4:10}$	10	0.100	0.900	0.10000
$Y_{2:4:10}$	5	0.200	0.800	0.28000
$Y_{3:4:10}$	4	0.250	0.750	0.46000
$Y_{4:4:10}$	3	0.333	0.667	0.64000
$Y_{5:4:10}$	2	0.500	0.500	0.82000

For a fixed level of significance α, the critical region for the precedence test based on will be $\{q, q+1 \ldots, m_1\}$, where

$$\alpha = \Pr(\bar{Q}_{(m_2)} \geq q | F_X = F_Y). \tag{4.8}$$

The critical values q and the exact level of significance α (as close as possible to 5%) for different choices of the sample sizes n_1 and n_2, effective sample sizes m_1 and m_2 and censoring schemes $\mathbf{R}$ and $\mathbf{S}$ are presented in Table 1.

It can be seen from Table 1 that for the schemes selected in the table, the support of the distribution of the test statistic $P_{(m_2)}$ is small and hence there is a limited choice for the level of significance, and the values are much larger than the nominal level of significance 0.05. Of the three statistics considered, the test based on T_{w,m_2} has the closest agreement with the nominal level of 0.05. followed by the test based on Q_{m_2}.

5 Exact Power Under Lehmann Alternative

There are two ways to define the Lehmann alternative.

1. The Lehmann alternative $H_1 : (1-F_X)^\delta = (1-F_Y)$ for some δ, which was first proposed by Lehmann (1953), is a subclass of the alternative $H_1 : F_X > F_Y$ when $\delta \in (0, 1)$ (see Gibbons and Chakraborti, 2003, Sect. 6.1).

Table 3: Power comparison under Lehmann alternative for $P_{(m_2)}$, T_{W,m_2} and $\bar{Q}_{(m_2)}$ with $1/\delta = 2(1)5$

	Exact l.o.s.			$\delta = 1/2$			$\delta = 1/3$			$\delta = 1/4$			$\delta = 1/5$		
Setting	$P_{(m_2)}$	T_{W,m_2}	$\bar{Q}_{m_2}$	$P_{(m_2)}$	T_{W,m_2}	$\bar{Q}_{m_2}$	$P_{(m_2)}$	T_{W,m_2}	$\bar{Q}_{m_2}$	$P_{(m_2)}$	T_{W,m_2}	$\bar{Q}_{m_2}$	$P_{(m_2)}$	T_{W,m_2}	$\bar{Q}_{m_2}$
1	0.500	0.045	0.136	0.814	0.168	0.402	0.921	0.293	0.578	0.962	0.396	0.685	0.835	0.773	0.790
2	0.500	0.053	0.134	0.808	0.194	0.400	0.916	0.328	0.580	0.958	0.435	0.688	0.837	0.819	0.799
3	0.500	0.048	0.167	0.823	0.209	0.452	0.928	0.371	0.623	0.967	0.497	0.721	0.943	0.851	0.932
4	0.500	0.047	0.178	0.837	0.208	0.464	0.939	0.372	0.631	0.973	0.501	0.725	0.967	0.838	0.955
5	0.523	0.049	0.075	0.821	0.169	0.278	0.922	0.285	0.451	0.962	0.383	0.574	0.851	0.818	0.755
6	0.545	0.047	0.081	0.811	0.212	0.422	0.927	0.378	0.595	0.968	0.507	0.698	0.958	0.825	0.940
7	0.455	0.050	0.148	0.865	0.189	0.277	0.952	0.316	0.433	0.980	0.414	0.540	0.974	0.870	0.907
8	0.566	0.053	0.092	0.869	0.227	0.306	0.952	0.395	0.469	0.980	0.523	0.576	0.973	0.854	0.905
9	0.050	0.061	0.030	0.208	0.196	0.150	0.372	0.327	0.292	0.505	0.434	0.416	0.474	0.855	0.473
10	0.241	0.051	0.073	0.560	0.205	0.258	0.736	0.352	0.414	0.832	0.466	0.526	0.895	0.876	0.852
11	0.150	0.046	0.068	0.441	0.204	0.264	0.639	0.362	0.438	0.759	0.487	0.563	0.848	0.815	0.834
12	0.075	0.046	0.041	0.274	0.169	0.188	0.455	0.294	0.344	0.589	0.399	0.472	0.490	0.766	0.487
13	0.510	0.055	0.047	0.819	0.189	0.189	0.923	0.320	0.324	0.963	0.424	0.430	0.840	0.794	0.696
14	0.539	0.042	0.066	0.808	0.176	0.239	0.907	0.315	0.392	0.950	0.430	0.505	0.862	0.725	0.788
15	0.522	0.049	0.062	0.816	0.147	0.237	0.918	0.269	0.396	0.959	0.374	0.514	0.849	0.763	0.744
16	0.707	0.048	0.137	0.928	0.192	0.385	0.976	0.332	0.548	0.991	0.442	0.649	0.991	0.885	0.961
17	0.409	0.048	0.037	0.764	0.194	0.192	0.896	0.334	0.348	0.949	0.444	0.467	0.778	0.849	0.667
18	0.430	0.049	0.037	0.764	0.188	0.193	0.891	0.319	0.350	0.945	0.424	0.469	0.799	0.890	0.723
19	0.407	0.048	0.082	0.873	0.169	0.218	0.955	0.318	0.378	0.981	0.437	0.496	0.973	0.872	0.879
20	0.419	0.053	0.073	0.780	0.254	0.299	0.908	0.441	0.479	0.957	0.575	0.597	0.923	0.926	0.899
21	0.400	0.051	0.027	0.754	0.192	0.152	0.889	0.327	0.289	0.944	0.432	0.399	0.771	0.902	0.665
22	0.394	0.050	0.029	0.726	0.184	0.172	0.864	0.312	0.332	0.927	0.415	0.459	0.776	0.949	0.695
23	0.347	0.051	0.043	0.732	0.246	0.219	0.884	0.432	0.387	0.944	0.568	0.509	0.943	0.967	0.915
24	0.216	0.046	0.053	0.581	0.248	0.268	0.780	0.454	0.466	0.879	0.604	0.601	0.858	0.959	0.852

2. The Lehmann alternative $H_1 : [F_X]^\gamma = F_Y$ for some γ, which was first proposed by Lehmann (1953). We can see that $H_1 : [F_X]^\gamma = F_Y$ is a subclass of the alternative $H_1 : F_X > F_Y$ when $\gamma > 1$.

Joint Non-null Distribution of the Placement Statistics Under the Lehmann Alternative (1):

The joint pmf of the placement statistics $U_1, U_2, \cdots, U_l$ under the Lehmann alternative hypothesis $H_1 : 1 - F_Y(x) = (1 - F_X(x))^\delta, \delta \leq 1$ is given by

$$\Pr(U_1 = u_1, \ldots, U_l = u_l) = \frac{AB}{C} \sum_{i_1=0}^{u_1} \cdots \sum_{i_l=0}^{u_l} \prod_{j=1}^{l+1} \gamma_{i_j, u_j}(R_{u_{(j-1)}+1}, \ldots, R_{u_{(j)}})$$

where

$$C = \prod_{j=0}^{m_2-1} (N^* - T_{(j)} - j)$$

with $T_{(j)} = 0$ if $j = 0$, and $T_{(j)} = T_1 + T_2 + \cdots + T_j$ otherwise, and $N^* = T_1 + T_2 + \cdots + T_{m_2} + m_2$.

The joint non-null pmf of $U_1, U_2, \cdots, U_l$ under the Lehmann alternative (**2**): $H_1 : [F_X]^\gamma = F_Y$ for $\gamma > 1$, can be derived similarly.

Table 4: Power comparison under Lehmann alternative for $P_{(m_2)}$, T_{W,m_2} and $\bar{Q}_{(m_2)}$ with $\gamma = 2(1)5$

	Exact l.o.s.			$\gamma = 2$			$\gamma = 3$			$\gamma = 4$			$\gamma = 5$		
Setting	$P_{(m_2)}$	T_{W,m_2}	$\bar{Q}_{m_2}$	$P_{(m_2)}$	T_{W,m_2}	$\bar{Q}_{m_2}$	$P_{(m_2)}$	T_{W,m_2}	$\bar{Q}_{m_2}$	$P_{(m_2)}$	T_{W,m_2}	$\bar{Q}_{m_2}$	$P_{(m_2)}$	T_{W,m_2}	$\bar{Q}_{m_2}$
1	0.500	0.045	0.136	0.668	0.283	0.456	0.751	0.514	0.638	0.802	0.672	0.734	0.835	0.773	0.790
2	0.500	0.053	0.134	0.670	0.326	0.465	0.754	0.572	0.649	0.804	0.727	0.744	0.837	0.819	0.799
3	0.500	0.048	0.167	0.764	0.334	0.596	0.868	0.600	0.805	0.917	0.761	0.891	0.943	0.851	0.932
4	0.500	0.047	0.178	0.796	0.321	0.627	0.901	0.582	0.837	0.946	0.744	0.919	0.967	0.838	0.955
5	0.523	0.049	0.075	0.691	0.308	0.351	0.773	0.557	0.557	0.820	0.720	0.680	0.851	0.818	0.755
6	0.545	0.047	0.081	0.762	0.317	0.572	0.881	0.571	0.797	0.933	0.731	0.895	0.958	0.825	0.940
7	0.455	0.050	0.148	0.825	0.342	0.440	0.919	0.615	0.704	0.956	0.779	0.839	0.974	0.870	0.907
8	0.566	0.053	0.092	0.829	0.349	0.461	0.919	0.612	0.715	0.956	0.767	0.842	0.973	0.854	0.905
9	0.050	0.061	0.030	0.180	0.352	0.163	0.301	0.609	0.294	0.398	0.765	0.395	0.474	0.855	0.473
10	0.241	0.051	0.073	0.562	0.341	0.404	0.743	0.604	0.661	0.840	0.763	0.801	0.895	0.876	0.852
11	0.150	0.046	0.068	0.451	0.311	0.367	0.655	0.561	0.608	0.775	0.720	0.751	0.848	0.815	0.834
12	0.075	0.046	0.041	0.211	0.279	0.188	0.328	0.508	0.317	0.419	0.664	0.414	0.490	0.766	0.487
13	0.510	0.055	0.047	0.676	0.313	0.269	0.758	0.547	0.471	0.807	0.699	0.607	0.840	0.794	0.696
14	0.539	0.042	0.066	0.708	0.281	0.369	0.787	0.502	0.596	0.832	0.640	0.719	0.862	0.725	0.788
15	0.522	0.049	0.062	0.689	0.302	0.332	0.770	0.529	0.543	0.817	0.673	0.668	0.849	0.763	0.744
16	0.707	0.048	0.137	0.915	0.353	0.596	0.966	0.636	0.834	0.983	0.799	0.924	0.991	0.885	0.961
17	0.409	0.048	0.037	0.582	0.324	0.241	0.677	0.589	0.441	0.736	0.754	0.577	0.778	0.849	0.667
18	0.430	0.049	0.037	0.608	0.352	0.273	0.702	0.636	0.499	0.760	0.802	0.639	0.799	0.890	0.723
19	0.407	0.048	0.082	0.824	0.295	0.353	0.916	0.589	0.637	0.955	0.772	0.796	0.973	0.872	0.879
20	0.419	0.053	0.073	0.702	0.400	0.441	0.827	0.700	0.707	0.889	0.854	0.836	0.923	0.926	0.899
21	0.400	0.051	0.027	0.573	0.381	0.228	0.668	0.667	0.435	0.729	0.823	0.575	0.771	0.902	0.665
22	0.394	0.050	0.029	0.573	0.428	0.260	0.671	0.739	0.478	0.733	0.887	0.613	0.776	0.949	0.695
23	0.347	0.051	0.043	0.685	0.466	0.402	0.837	0.786	0.703	0.907	0.918	0.848	0.943	0.967	0.915
24	0.216	0.046	0.053	0.522	0.445	0.399	0.700	0.766	0.655	0.800	0.905	0.784	0.858	0.959	0.852

6 Discussion

Table 3 provides a comparison of simulated power for the three tests P, T, and $\bar{Q}$ each with the level of significance listed under the column "Exact l. o. s". The power is computed under the Lehmann alternative $(1 - F_X)^\delta = 1 - F_X$ for $\delta = 1/2, 1/3, 1/4$, and $1/5$ for the schemes described in Table 1. We generated 10,000,000 sets of data in order to obtain the estimated rejection rates under Lehmann alternatives with $1/\delta = 2(1)5$. It can be seen that the power of each test increases as the value of δ gets smaller. As δ gets smaller, the shapes of the cdf under the null and the alternative hypotheses deviate substantially from each other, and the three tests can distinguish well between the null and the alternative cdf's. The test based on P has high power even for large δ. The powers of T and $\bar{Q}$ tests are small for large δ and increase as δ decreases. For smaller sample sizes n_1 and n_2, the power of the P test dominates the powers of the other two tests, it may be because this test has much larger l.o.s. and hence a larger rejection region as compared to the other two tests. For n_1, m_1 and n_2, m_2 both large and smaller δ, the power of T test is higher than those of the other two tests.

Table 4 provides simulated power of the three tests P, T, and $\bar{Q}$ under the same sampling and censoring schemes as in Table 3 for the Lehmann alternatives $H_1 : (F_X)^\gamma = F_Y$ for various values of γ. We generated 10,000,000 sets of data in order to obtain the estimated rejection rates under Lehmann alternatives with $\gamma = 2(1)5$. A similar behavior of the power function is seen from this table. The P test has higher power for smaller sample sizes and for smaller values of γ. As the sample sizes increase, the power of the T test dominates the power of the other two tests.

References

Balakrishnan, N., 2007. Progressive Censoring Methodology: An Appraisal. Test, 16, 211-296 (with discussion).

Balakrishnan, N., Aggarwala, R., 2000. Progressive Censoring: Theory, Methods and Applications. Birkhäuser, Boston.

Balakrishnan, N., Tripathi, R. C., Kannan, N., 2007. On the joint distribution of placement statistics under progressive censoring and applications to precedence test. Journal of Statistical Planning and Inference, 138, 1314-1324.

Gibbons, J. D., Chakraborti, S., 2003. Nonparametric Statistical Inference, Fourth edition. Marcel Dekker, New York.

Kaplan, E. L., Meier, P., 1958. Nonparametric estimation from incomplete observations. Journal of the American Statistical Association, 53, 457-481.

Lehmann, E. L., 1953. The power of rank tests. Annals of Mathematical Statistics, 24, 23-42.

Nelson, L. S., 1963. Tables of a precedence life test. Technometrics, 5, 491-499.

Ng, H. K. T., Balakrishnan, N., 2002. Wilcoxon-type rank-sum precedence tests: Large-sample approximation and evaluation. Applied Stochastic Models in Business and Industry, 18, 271-286.