Nonparametric predictive precedence testing for two groups

Pauline Coolen-Schrijner, Tahani Maturi and Frank Coolen
Dept of Mathematical Sciences, Durham University
Durham DH1 3LE, UK
pauline.schrijner@durham.ac.uk, tahani.maturi@durham.ac.uk
frank.coolen@durham.ac.uk

Abstract

Nonparametric predictive inference (NPI) is a statistical approach based on few assumptions about probability distributions, with inferences based on data. NPI assumes exchangeability of random quantities, both related to observed data and future observations, and uncertainty is quantified via lower and upper probabilities. In this paper, lifetimes of units from groups X and Y are compared, based on observed lifetimes from an experiment that may have ended before all units had failed. We present upper and lower probabilities for the event that the lifetime of a future unit from X is less than the lifetime of a future unit from Y, and we compare this approach with traditional precedence testing.

1 Introduction

Comparison of lifetimes of units from different groups is a common problem. Units from different groups are simultaneously placed on a life-testing experiment, and decisions may be needed before all units have failed due to cost or time considerations. In precedence testing, the experiment is terminated at a certain time or after a certain number of failures (for a particular group), so the data consist of both observed lifetimes and right-censored observations. Epstein (1955) presented precedence testing, Nelson (1963) proposed it as an efficient life-test procedure that enables decisions after relatively few lifetimes are observed. Balakrishnan and Ng (2006) describe several nonparametric precedence tests based on the hypothesis of equal lifetime distributions. As an alternative, we propose nonparametric predictive precedence testing for two groups, with lower and upper probabilities for the event that a future observation from one group is less than a future observation from another group.

Section 2 is a short overview of nonparametric predictive inference (NPI), in Section 3 we present our NPI approach to precedence testing and derive the main results. In Section 4 we briefly describe some established nonparametric precedence tests, and we compare our method with these tests via an example. Section 5 contains some concluding remarks.

2 Nonparametric predictive inference

Nonparametric predictive inference (NPI) is a statistical method based on Hill's assumption $A_{(n)}$ (Hill 1968), which gives a direct conditional probability for a future observable random quantity, conditional on observed values of related random quantities (Augustin and Coolen 2004, Coolen 2006). Suppose that $X_1, \ldots, X_n, X_{n+1}$ are positive, continuous and exchangeable random quantities representing lifetimes. Let the ordered observed values of $X_1, \ldots, X_n$ be denoted by $x_{1:n} < x_{2:n} < \ldots < x_{n:n} < \infty$, and let $x_{0:n} = 0$ for ease of notation. We assume that no ties occur, our results can be generalised to allow ties (Hill 1993). For positive X_{n+1}, representing a future observation, based on n observations, $A_{(n)}$ (Hill 1968) is

$$P(X_{n+1} \in (x_{j-1:n}, x_{j:n})) = \frac{1}{n+1},\ j = 1, 2, \ldots, n,\ \text{and}\ P(X_{n+1} \in (x_{n:n}, \infty)) = \frac{1}{n+1}. \tag{2.1}$$

$A_{(n)}$ does not assume anything else, and is a post-data assumption related to exchangeability (De Finetti 1974). Hill (1988) discusses $A_{(n)}$ in detail. Inferences based on $A_{(n)}$ are predictive and nonparametric, and can be considered suitable if there is hardly any knowledge about the random quantity of interest, other than the n observations, or if one does not want to use such information, e.g. to study effects of additional assumptions underlying statistical models. $A_{(n)}$ is not sufficient to derive precise probabilities for many events of interest, but it provides bounds for probabilities via the 'fundamental theorem of probability' (De Finetti 1974), which are lower and upper probabilities in interval probability theory (Walley 1991, Weichselberger 2001).

In precedence testing for two groups, units of both groups are placed simultaneously on a life-testing experiment, and failures are observed as they arise during the experiment, which is terminated as soon as a certain stop criterion has been reached, so the lifetimes of some units are typically right-censored. Coolen and Yan (2004) presented a generalization of $A_{(n)}$, called rc-$A_{(n)}$, suitable for right-censored data. In this paper, all right-censored observations are the same which simplifies the use of rc-$A_{(n)}$. To formulate the required form of rc-$A_{(n)}$, we need notation for probability mass assigned to intervals without further restrictions on the spread within the intervals. Such a partial specification of a probability distribution is called an M-function (Coolen and Yan 2004).

Definition 2.1 A partial specification of a probability distribution for a real-valued random quantity X can be provided via probability masses assigned to intervals, without any further restriction on the spread of the probability mass within each interval. A probability mass assigned, in such a way, to an interval (a, b) is denoted by $M_X(a, b)$, and referred to as M-function value for X on (a, b).

In precedence testing the experiment is terminated as soon as a certain stop criterion has been reached. We assume that this stop criterion is expressed in terms of a stopping time T_0, but if instead a number of failures were used as stop criterion then this would not affect our method, as it is of no relevance in NPI how T_0 is determined. When considering a single group of units, let r denote the number of observations of $X_1, \ldots, X_n$ that occur before the stopping time T_0, so $n - r$ observations are right-censored at T_0. The next definition provides the M-functions required for precedence testing, which follow from rc-$A_{(n)}$ (Coolen and Yan, 2004).

Definition 2.2 For nonparametric predictive precedence testing with stopping time T_0, the assumption rc-$A_{(n)}$ implies that the probability distribution for a nonnegative random quantity X_{n+1} on the basis of data including r real and $n - r$ right-censored observations, is partially specified by the following M-function values:

$$M_{X_{n+1}}(x_{j-1:n}, x_{j:n}) = \frac{1}{n+1}, \quad j = 1, \ldots, r,$$
$$M_{X_{n+1}}(x_{r:n}, \infty) = \frac{1}{n+1} \text{ and } M_{X_{n+1}}(T_0, \infty) = \frac{n-r}{n+1}. \tag{2.2}$$

In comparison to $A_{(n)}$, rc-$A_{(n)}$ uses the extra assumption that, at the moment of censoring, the residual lifetime of a right-censored unit is exchangeable with the residual lifetimes of all other units that have not yet failed or been censored. Further details of rc-$A_{(n)}$ are given in Coolen and Yan (2004).

In this paper we consider nonparametric predictive precedence testing for two groups, say X and Y, and we are interested in the lower and upper probabilities that a future observation X_{n_x+1} of group X is less than a future observation Y_{n_y+1} of group Y, based on n_x and n_y observations of group X and Y, stopping time T_0, and the assumptions rc-$A_{(n_x)}$ and rc-$A_{(n_y)}$. The derivation of these lower and upper probabilities, given in the next section, require the following lemma from Coolen and Yan (2003).

Lemma 2.1 For $s \geq 2$, let $J_l = (j_l, r)$, with $j_1 < j_2 < \ldots < j_s < r$, so we have nested intervals $J_1 \supset J_2 \supset \ldots \supset J_s$ with the same right end-point r (which may be infinity). We consider two independent real-valued random quantities, say X and Y. Let the probability distribution for X be partially specified via M-function values, with all probability mass $P(X \in J_1)$ described by the s M-function values $M_X(J_l)$, so $\sum_{l=1}^{s} M_X(J_l) = P(X \in J_1)$. Then, without additional assumptions, $\sum_{l=1}^{s} P(Y < j_l) M_X(J_l) \leq P(Y < X, X \in J_1) \leq P(Y < r) P(X \in J_1)$, and these bounds are the maximum lower and minimum upper bounds that generally hold.

3 Precedence testing

In this section NPI precedence testing is presented. Subsection 3.1 presents the NPI lower and upper probabilities for the event that a future observation of group X is less than a future observation of group Y, some of their properties are discussed in Subsection 3.2.

3.1 Upper and lower probabilities

To compare two groups of lifetime data by precedence testing, we use the notation as introduced above, but we add an index x or y corresponding to the groups X and Y. So, n_x and n_y units of groups X and Y are placed simultaneously on a life-testing experiment and r_x and r_y lifetimes of groups X and Y are observed before the experiment is terminated at time T_0. So $n_x - r_x$ and $n_y - r_y$ lifetimes of groups X and Y are right-censored at T_0. Throughout we assume that information on units from one group does not hold any information about units from the other group, so X_{n_x+1} and Y_{n_y+1} are independent and data from group X contain no information on Y_{n_y+1}, and vice versa.

Bounds for the probability of $X_{n_x+1} < Y_{n_y+1}$, given the data and stopping time T_0 and based on rc-$A_{(n_x)}$ and rc-$A_{(n_y)}$, are presented in Theorem 3.1. Throughout, conditioning on the data is left out of the notation. As these bounds are optimal, without any additional assumptions, they are lower and upper probabilities (Walley 1991), which we denote by $\underline{P}(X_{n_x+1} < Y_{n_y+1})$ and $\overline{P}(X_{n_x+1} < Y_{n_y+1})$, respectively. The indicator function 1_A is equal to 1 if event A occurs and 0 else.

Theorem 3.1 For the above scenario, the lower and upper probabilities for the event $X_{n_x+1} < Y_{n_y+1}$ are

$$\underline{P}(X_{n_x+1} < Y_{n_y+1}) = \frac{1}{(n_x+1)(n_y+1)} \left[\sum_{j=1}^{r_y} \sum_{i=1}^{r_x} 1_{\{x_{i:n_x} < y_{j:n_y}\}} + r_x(n_y - r_y) \right], \tag{3.3}$$

$$\overline{P}(X_{n_x+1} < Y_{n_y+1}) = \frac{1}{(n_x+1)(n_y+1)} \left[\sum_{j=1}^{r_y} \sum_{i=1}^{r_x} 1_{\{x_{i:n_x} < y_{j:n_y}\}} + r_y + (n_x+1)(n_y - r_y + 1) \right]. \tag{3.4}$$

Proof The lower probability for the event $X_{n_x+1} < Y_{n_y+1}$ given the data and T_0 is derived as follows:

$$\begin{aligned}
P(X_{n_x+1} < Y_{n_y+1}) =& \sum_{j=1}^{r_y} P(X_{n_x+1} < Y_{n_y+1}, Y_{n_y+1} \in (y_{j-1:n_y}, y_{j:n_y})) + \\
& P(X_{n_x+1} < Y_{n_y+1}, Y_{n_y+1} \in (y_{r_y:n_y}, \infty)) \\
\geq & \sum_{j=1}^{r_y} P(X_{n_x+1} < y_{j-1:n_y}) \; M_{Y_{n_y+1}}(y_{j-1:n_y}, y_{j:n_y}) + \\
& P(X_{n_x+1} < y_{r_y:n_y}) \; M_{Y_{n_y+1}}(y_{r_y:n_y}, \infty) + P(X_{n_x+1} < T_0) \; M_{Y_{n_y+1}}(T_0, \infty) \\
=& \frac{1}{n_y+1} \sum_{j=1}^{r_y} P(X_{n_x+1} < y_{j-1:n_y}) + \frac{1}{n_y+1} P(X_{n_x+1} < y_{r_y:n_y}) + \frac{n_y - r_y}{n_y+1} P(X_{n_x+1} < T_0) \\
\geq & \frac{1}{(n_x+1)(n_y+1)} \left[\sum_{j=1}^{r_y} \sum_{i=1}^{r_x} 1_{\{x_{i:n_x} < y_{j-1:n_y}\}} + \sum_{i=1}^{r_x} 1_{\{x_{i:n_x} < y_{r_y:n_y}\}} + (n_y - r_y) \sum_{i=1}^{r_x} 1_{\{x_{i:n_x} < T_0\}} \right] \\
=& \frac{1}{(n_x+1)(n_y+1)} \left[\sum_{j=1}^{r_y} \sum_{i=1}^{r_x} 1_{\{x_{i:n_x} < y_{j:n_y}\}} + r_x(n_y - r_y) \right].
\end{aligned}$$

The first inequality follows by putting all mass of Y_{n_y+1} corresponding to the intervals $(y_{j-1:n_y}, y_{j:n_y})$ $(j = 1, \ldots, r_y)$, $(y_{r_y:n_y}, \infty)$ and (T_0, ∞) in the left end points of these intervals, and by using Lemma

2.1 for the nested intervals $(y_{r_y:n_y}, \infty)$ and (T_0, ∞). The second inequality follows by putting all mass of X_{n_x+1} corresponding to the intervals $(x_{i-1:n_x}, x_{i:n_x})$ $(i = 1, \ldots, r_x)$, $(x_{r_x:n_x}, \infty)$ and (T_0, ∞) in the right end points of these intervals.

The corresponding upper probability for the event $X_{n_x+1} < Y_{n_y+1}$ is derived as follows:

$$\begin{aligned}
\overline{P}(X_{n_x+1} < Y_{n_y+1}) =& \sum_{j=1}^{r_y} P(X_{n_x+1} < Y_{n_y+1}, Y_{n_y+1} \in (y_{j-1:n_y}, y_{j:n_y})) + \\
& P(X_{n_x+1} < Y_{n_y+1}, Y_{n_y+1} \in (y_{r_y:n_y}, \infty)) \\
\leq& \sum_{j=1}^{r_y} P(X_{n_x+1} < y_{j:n_y})\ M_{Y_{n_y+1}}(y_{j-1:n_y}, y_{j:n_y}) + \\
& P(X_{n_x+1} < \infty)\ M_{Y_{n_y+1}}(y_{r_y:n_y}, \infty) + P(X_{n_x+1} < \infty)\ M_{Y_{n_y+1}}(T_0, \infty) \\
=& \frac{1}{n_y+1} \sum_{j=1}^{r_y} P(X_{n_x+1} < y_{j:n_y}) + \frac{1}{n_y+1} P(X_{n_x+1} < \infty) + \frac{n_y - r_y}{n_y+1} P(X_{n_x+1} < \infty) \\
\leq& \frac{1}{(n_x+1)(n_y+1)} \sum_{j=1}^{r_y} \sum_{i=1}^{r_x+1} 1_{\{x_{i-1:n_x} < y_{j:n_y}\}} + \frac{1}{n_y+1} + \frac{n_y - r_y}{n_y+1} \\
=& \frac{1}{(n_x+1)(n_y+1)} \left[\sum_{j=1}^{r_y} \sum_{i=1}^{r_x} 1_{\{x_{i:n_x} < y_{j:n_y}\}} + r_y + (n_y - r_y + 1)(n_x + 1) \right].
\end{aligned}$$

The first inequality follows by putting all mass of Y_{n_y+1} corresponding to the intervals $(y_{j-1:n_y}, y_{j:n_y})$ $(j = 1, \ldots, r_y)$, $(y_{r_y:n_y}, \infty)$ and (T_0, ∞) in the right end points of these intervals, using Lemma 2.1 for the nested intervals $(y_{r_y:n_y}, \infty)$ and (T_0, ∞). The second inequality follows by putting all mass of X_{n_x+1} corresponding to the intervals $(x_{i-1:n_x}, x_{i:n_x})$ $(i = 1, \ldots, r_x)$, $(x_{r_x:n_x}, \infty)$ and (T_0, ∞) in the left end points of these intervals. □

These lower and upper probabilities are based only on $x_{i:n_x}$ $(i = 1 \ldots, r_x)$, $y_{j:n_y}$ $(j = 1, \ldots, r_y)$ and T_0, further information on location as contained in the observations is not used. As such, this approach can be regarded as a fully predictive alternative to standard rank-based methods (Lehmann 1975). As these lower and upper probabilities are F-probability in the theory of interval probability (Augustin and Coolen 2004, Weichselberger 2001), the conjugacy property holds, that is for an event A and its complementary event A^c, $\underline{P}(A) = 1 - \overline{P}(A^c)$.

From Theorem 3.1 it follows that if $r_x = 0$ and $r_y \in \{0, 1, \ldots, n_y\}$, that is, the experiment is terminated before the first observation of group X is observed, we have

$$\underline{P}(X_{n_x+1} < Y_{n_y+1}) = 0 \quad \text{and} \quad \overline{P}(X_{n_x+1} < Y_{n_y+1}) = 1 - \frac{n_x r_y}{(n_x+1)(n_y+1)}. \tag{3.5}$$

This lower probability is zero, reflecting that on the basis of the data one cannot exclude the possibility that the X observations will always exceed all Y observations. If $r_y = 0$ and $r_x \in \{0, 1, \ldots, n_x\}$, that is, the experiment is terminated before the first observation of group Y is observed, we have

$$\underline{P}(X_{n_x+1} < Y_{n_y+1}) = \frac{r_x n_y}{(n_x+1)(n_y+1)} \quad \text{and} \quad \overline{P}(X_{n_x+1} < Y_{n_y+1}) = 1. \tag{3.6}$$

This upper probability is one, reflecting that one cannot exclude the possibility that the X observations will always be less than all Y observations. The lower and upper probabilities in (3.6) can also be obtained from (3.5) by using the conjugacy property.

If all units of group Y are observed before the first observation of group X is observed, that is, $y_{n_y:n_y} < x_{1:n_x}$, and the experiment is terminated after the last unit of group Y is observed $(T_0 > y_{n_y:n_y})$ then, independent of the number of units of group X observed, we have

$$\underline{P}(X_{n_x+1} < Y_{n_y+1}) = 0 \quad \text{and} \quad \overline{P}(X_{n_x+1} < Y_{n_y+1}) = 1 - \frac{n_x n_y}{(n_x+1)(n_y+1)}. \tag{3.7}$$

Similarly, if all units of group X are observed before the first observation of group Y is observed, that is $x_{n_x:n_x} < y_{1:n_y}$, and the experiment is terminated after the last unit of group X is observed ($T_0 > x_{n_x:n_x}$) then, independent of the number of units of group Y observed, we have

$$\underline{P}(X_{n_x+1} < Y_{n_y+1}) = \frac{n_x n_y}{(n_x+1)(n_y+1)} \quad \text{and} \quad \overline{P}(X_{n_x+1} < Y_{n_y+1}) = 1. \tag{3.8}$$

Again the lower and upper probabilities in (3.8) can be obtained from (3.7) in exactly the same way as the probabilities in (3.6) can be obtained from (3.5) as described before.

3.2 Analysis of upper and lower probabilities

Suppose that the stopping time is increased from T_0 to T_0^*, and denote by r_x^* and r_y^* the number of lifetimes of group X and Y, respectively, observed before T_0^*. The lower and upper probabilities for the event $X_{n_x+1} < Y_{n_y+1}$, based on the data, T_0, rc-$A_{(n_x)}$ and rc-$A_{(n_y)}$, are denoted by $\underline{P}(X_{n_x+1} < Y_{n_y+1})$ and $\overline{P}(X_{n_x+1} < Y_{n_y+1})$, while the corresponding lower and upper probabilities for T_0^* are denoted by $\underline{P}^*(X_{n_x+1} < Y_{n_y+1})$ and $\overline{P}^*(X_{n_x+1} < Y_{n_y+1})$. We can write $r_x^* = r_x + a$ and $r_y^* = r_y + b$ with a, b nonnegative integers. Using (3.3) the lower probability $\underline{P}^*(X_{n_x+1} < Y_{n_y+1})$ can be written as follows:

$$\begin{aligned}
\underline{P}^*(X_{n_x+1} < Y_{n_y+1}) &= \frac{1}{(n_x+1)(n_y+1)}\left[\sum_{j=1}^{r_y+b}\sum_{i=1}^{r_x+a} 1_{\{x_{i:n_x} < y_{j:n_y}\}} + (r_x+a)(n_y - r_y - b)\right] \\
&= \underline{P}(X_{n_x+1} < Y_{n_y+1}) + \frac{1}{(n_x+1)(n_y+1)}\left[\sum_{j=1}^{r_y}\sum_{i=r_x+1}^{r_x+a} 1_{\{x_{i:n_x} < y_{j:n_y}\}} + \right. \\
&\quad \left. \sum_{j=r_y+1}^{r_y+b}\sum_{i=1}^{r_x+a} 1_{\{x_{i:n_x} < y_{j:n_y}\}} + a(n_y - r_y - b) - b r_x\right].
\end{aligned} \tag{3.9}$$

In a similar way, using (3.4), the upper probability $\overline{P}^*(X_{n_x+1} < Y_{n_y+1})$ can be written as follows:

$$\begin{aligned}
\overline{P}^*(X_{n_x+1} < Y_{n_y+1}) &= \frac{1}{(n_x+1)(n_y+1)}\left[\sum_{j=1}^{r_y+b}\sum_{i=1}^{r_x+a} 1_{\{x_{i:n_x} < y_{j:n_y}\}} + r_y + b + (n_x+1)(n_y - r_y - b + 1)\right] \\
&= \overline{P}(X_{n_x+1} < Y_{n_y+1}) + \frac{1}{(n_x+1)(n_y+1)}\left[\sum_{j=1}^{r_y}\sum_{i=r_x+1}^{r_x+a} 1_{\{x_{i:n_x} < y_{j:n_y}\}} + \right. \\
&\quad \left. \sum_{j=r_y+1}^{r_y+b}\sum_{i=1}^{r_x+a} 1_{\{x_{i:n_x} < y_{j:n_y}\}} - b n_x\right].
\end{aligned} \tag{3.10}$$

Theorem 3.2 follows from (3.9) and (3.10).

Theorem 3.2

a. Consider the situation that r_x is increasing while r_y is kept constant. Then **(i)** the lower probability $\underline{P}(X_{n_x+1} < Y_{n_y+1})$ is strictly increasing in r_x, except if $x_{r_x+1:n_x} > y_{n_y:n_y}$ in which case the lower probability remains constant, and **(ii)** the upper probability $\overline{P}(X_{n_x+1} < Y_{n_y+1})$ remains constant.

b. Consider the situation that r_y is increasing while r_x is kept constant. Then **(i)** the lower probability $\underline{P}(X_{n_x+1} < Y_{n_y+1})$ remains constant, and **(ii)** the upper probability $\overline{P}(X_{n_x+1} < Y_{n_y+1})$ is strictly decreasing in r_y, except if $x_{n_x:n_x} < y_{r_y+1:n_y}$ in which case the upper probability remains constant.

Proof We prove part a, the proof of part b is similar. To prove **(i)**, increasing r_x while keeping r_y constant implies that a is a positive integer and $b = 0$. Substituting $b = 0$ into (3.9) yields

$$\underline{P}^*(X_{n_x+1}<Y_{n_y+1}) = \underline{P}(X_{n_x+1}<Y_{n_y+1}) + \frac{1}{(n_x+1)(n_y+1)}\left[\sum_{j=1}^{r_y}\sum_{i=r_x+1}^{r_x+a} 1_{\{x_{i:n_x}<y_{j:n_y}\}} + a(n_y - r_y)\right].$$

From this it follows that the lower probability is strictly increasing in r_x unless $n_y = r_y$ and the double sum equals zero, that is, if $n_y = r_y$ and all $x_{i:n_x}$, $i = r_x+1, \ldots, r_x+a$, are larger than $y_{r_y:n_y}$. These two conditions hold when $x_{r_x+1:n_x} > y_{n_y:n_y}$. To prove **(ii)**, substituting $b = 0$ into (3.10) yields

$$\overline{P}^*(X_{n_x+1} < Y_{n_y+1}) = \overline{P}(X_{n_x+1} < Y_{n_y+1}) + \frac{1}{(n_x+1)(n_y+1)}\left[\sum_{j=1}^{r_y}\sum_{i=r_x+1}^{r_x+a} 1_{\{x_{i:n_x}<y_{j:n_y}\}}\right].$$

From this it follows that the upper probability is strictly increasing in r_x unless the double sum equals zero, that is, if $x_{r_x+1:n_x} > y_{r_y:n_y}$. However, $x_{r_x+1:n_x}$ is by definition larger than $y_{r_y:n_y}$ and consequently the upper probability always remains constant in this case. □

The following example has been created to illustrate some of the special situations from Theorem 3.2.

Example 3.1 Six units each of group X and group Y are placed simultaneously on a life-testing experiment and their lifetimes are 1, 2, 3, 10, 11, 12 for X, and 4, 5, 6, 7, 8, 9 for Y, so all 6 observations of group Y are between the 3^{rd} and 4^{th} observations of group X. Suppose now that we would have terminated the experiment at stopping time T_0. We calculate the lower and upper probabilities that the lifetime of a future unit of group X is less than the lifetime of a future unit of group Y, given the observed lifetimes before T_0 for both groups and based on rc-$A_{(6)}$ for both groups. Table 3.1 and Figure 3.1 show the lower probabilities (3.3) and upper probabilities (3.4) when T_0 increases from 0 to ∞. As the lower and upper probabilities may only change when a lifetime of either group is observed, we only have to consider a finite number of time-intervals.

Table 3.1: Lower and upper probabilities for the event $X_7 < Y_7$

T_0	r_x	r_y	$\underline{P}(X_7 < Y_7)$	$\overline{P}(X_7 < Y_7)$	T_0	r_x	r_y	$\underline{P}(X_7 < Y_7)$	$\overline{P}(X_7 < Y_7)$
$[0,1)$	0	0	0	1	$[7,8)$	3	4	0.3673	0.7551
$[1,2)$	1	0	0.1224	1	$[8,9)$	3	5	0.3673	0.6939
$[2,3)$	2	0	0.2449	1	$[9,10)$	3	6	0.3673	0.6327
$[3,4)$	3	0	0.3673	1	$[10,11)$	4	6	0.3673	0.6327
$[4,5)$	3	1	0.3673	0.9388	$[11,12)$	5	6	0.3673	0.6327
$[5,6)$	3	2	0.3673	0.8776	$[12,\rightarrow)$	6	6	0.3673	0.6327
$[6,7)$	3	3	0.3673	0.8163					

From Table 3.1 we see that, when increasing r_x while keeping r_y constant, the lower probability is stepwise increasing, except for $T_0 \geq 9$ as then $x_{r_x+1:n_x} > y_{n_y:n_y}$. When increasing r_y while keeping r_x constant, the upper probability is stepwise decreasing. All this is in agreement with Theorem 3.2.

For each T_0, $\frac{1}{2} \in [\underline{P}(X_7 < Y_7), \overline{P}(X_7 < Y_7)]$ which implies that there is not much evidence that $X_7 < Y_7$. Table 3.1 also shows that the imprecision (difference between the lower and upper probability) decreases as the number of observations (or T_0) increases. The interval $[0.3673, 0.6327]$ is symmetric around $\frac{1}{2}$, due to the fact that our data are ‘symmetric’ in the order the observations occur: first 3 lifetimes of group X, followed by 6 lifetimes of group Y and then again 3 lifetimes of group X. This interval $[0.3673, 0.6327]$ has been reached already at $T_0 = 9$ as at that moment all units of group Y are observed, implying that the 3 remaining lifetimes of group X must be larger than the largest lifetime of group Y. For our method only the order is important, not the magnitude.

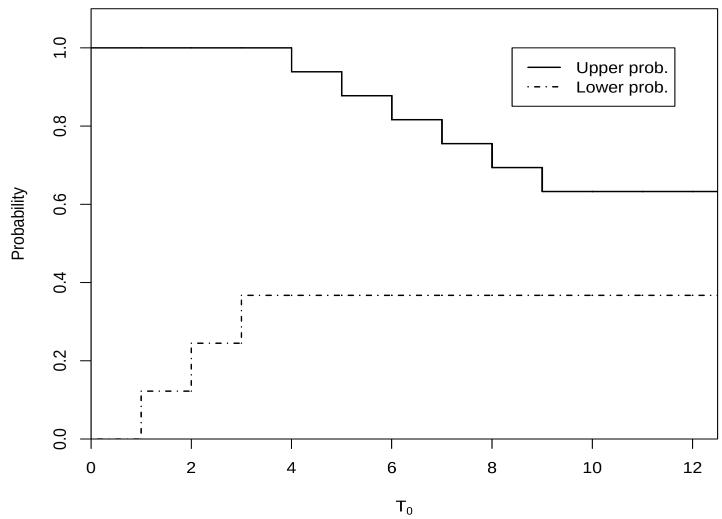

Figure 3.1: Lower and upper probabilities for the event $X_7 < Y_7$ (Ex. 3.1)

We should emphasize that the NPI lower (upper) probability for the event $X_{n_x+1} < Y_{n_y+1}$ never decreases (increases) if T_0 increases. This is in line with intuition, as all possible orderings of all lifetimes which are right-censored at T_0 are taken into account, and also with the general idea behind NPI, which is to explore what can be inferred from data with only few assumptions added.

4 Comparison with other nonparametric precedence tests

In Subsection 4.1 we briefly discuss several nonparametric precedence tests from the literature, and in Subsection 4.2 we compare our NPI approach with these more established methods via an example.

4.1 Some nonparametric precedence tests

We briefly discuss several nonparametric precedence tests, for more details we refer to Balakrishnan and Ng (2006). We compare the lifetimes of units from groups X and Y. Their lifetime distributions are denoted by F_X and F_Y, respectively, and n_x and n_y are the number of units of group X and Y that are placed simultaneously on a life-testing experiment. We assume that the experiment is terminated as soon as the r_y^{th} failure of group Y is observed (for these tests, the actual stop criterion employed is of relevance). The number of failures of group X observed before the r_y^{th} failure of group Y is denoted by r_x.

The Classical precedence test was introduced by Nelson (1963). One is interested in testing the null hypothesis H_0 that $F_X(x) = F_Y(x)$ for all $x \geq 0$. Let D_i be the random quantity representing the number of observed lifetimes of group X that are between the $(i-1)^{th}$ and i^{th} observed lifetime of group Y, for $i = 1, \ldots, r_y$, and denote their observed values by d_i. The precedence test statistic $Q_{(r_y)}$ is the number of lifetimes of group X that precede the r_y^{th} lifetime from group Y, so $Q_{(r_y)} = \sum_{i=1}^{r_y} D_i$. Under H_0, the distribution of $Q_{(r_y)}$ is

$$P(Q_{(r_y)} = j|H_0) = \frac{\binom{j+r_y-1}{j}\binom{n_x+n_y-j-r_y}{n_x-j}}{\binom{n_x+n_y}{n_y}}, \quad j = 0, \ldots, n_x. \tag{4.11}$$

The Maximal precedence test was proposed by Balakrishnan and Frattina (2000) to avoid the possible masking effect of the classical precedence test that the null hypothesis is not rejected for a certain value of r_y whilst there may exist a value less than this r_y such that the null hypothesis is rejected, at the same level of significance. The test statistic $U_{(r_y)}$ is the maximum number of lifetimes from group X that occur between the $(i-1)^{th}$ and i^{th} failure of group Y, for $i = 1, \ldots, r_y$, so $U_{(r_y)} = \max_{i=1,\ldots,r_y} D_i$. Under the null hypothesis H_0 that both lifetime distributions are the same, the cumulative distribution function of $U_{(r_y)}$ is given by

$$P(U_{(r_y)} \leq d|H_0) = P(D_1 \leq d, D_2 \leq d, \ldots, D_{r_y} \leq d|H_0) = \sum_{\substack{d_i\,(i=1,\ldots,r_y)=0 \\ \sum_{i=1}^{r_y} d_i \leq n_x}}^{d} \frac{\binom{n_x + n_y - \sum_{i=1}^{r_y} d_i - r_y}{n_y - r_y}}{\binom{n_x + n_y}{n_y}}. \quad (4.12)$$

Wilcoxon's minimal rank-sum precedence test was introduced by Ng and Balakrishnan (2004), together with Wilcoxon's maximal and expected rank-sum precedence tests (see below). We introduce $S_{r_y} = \sum_{i=1}^{r_y} D_i$ and $S^*_{r_y} = \sum_{i=1}^{r_y} iD_i$, and we denote their realisations by s_{r_y} and $s^*_{r_y}$. Let W_{r_y} be the rank-sum of the observed lifetimes of group X that occurred before the r_y^{th} observed lifetime of group Y. When all remaining $(n_x - s_{r_y})$ observations of group X occur between the r_y^{th} and $(r_y+1)^{th}$ observation of group Y, then Wilcoxon's test statistic will be minimal. The test statistic in this case, called the minimal rank-sum statistic, is

$$W_{\min,r_y} = W_{r_y} + (S_{r_y} + r_y + 1) + (S_{r_y} + r_y + 2) + \ldots + (n_x + r_y) = \frac{1}{2} n_x (n_x + 2r_y + 1) - (r_y + 1) S_{r_y} + S^*_{r_y}. \quad (4.13)$$

Alternatively, one can use Wilcoxon's maximal rank-sum precedence test if all remaining $(n_x - s_{r_y})$ observations of group X occur after the n_y^{th} observation of group Y, Wilcoxon's test statistic will be maximal. The test statistic in this case, called the maximal rank-sum statistic, is

$$W_{\max,r_y} = W_{r_y} + (S_{r_y} + n_y + 1) + (S_{r_y} + n_y + 2) + \ldots + (n_x + n_y) = \frac{1}{2} n_x (n_x + 2n_y + 1) - (n_y + 1) S_{r_y} + S^*_{r_y}. \quad (4.14)$$

Wilcoxon's expected rank-sum precedence test uses the expected rank sums of the lifetimes from group X between the r_y^{th} and $(r_y+1)^{th}$, ..., $(n_y-1)^{th}$ and n_y^{th} and after the n_y^{th} observation of group Y. This leads to the expected rank-sum statistic W_{E,r_y}. It can be shown that W_{E,r_y} is the average of $W_{\min,r_y}$ and $W_{\max,r_y}$, and hence the expected rank-sum statistic is given by

$$W_{E,r_y} = \frac{1}{2} n_x (n_x + n_y + r_y + 1) - (\frac{1}{2}(n_y + r_y) + 1) S_{r_y} + S^*_{r_y}. \quad (4.15)$$

Under the null hypothesis H_0 that the lifetime distributions of groups X and Y are the same, the distributions of $W_{\min,r_y}$, $W_{\max,r_y}$ and W_{E,r_y} are given by

$$P(W_{a,r_y} = w|H_0) = \sum_{\substack{m_i\,(i=1,\ldots,r_y)=0,\, s_{r_y} \leq n_x \\ \mathcal{B}}}^{n_x} \frac{\binom{n_x + n_y - s_{r_y} - r_y}{n_y - r_y}}{\binom{n_x + n_y}{n_y}}. \quad (4.16)$$

where for each a, the condition $\mathcal{B}$ is given by

a	$\mathcal{B}$
min	$\frac{1}{2} n_x (n_x + 2r_y + 1) - (r_y + 1) s_{r_y} + s^*_{r_y} = w$
max	$\frac{1}{2} n_x (n_x + 2n_y + 1) - (n_y + 1) s_{r_y} + s^*_{r_y} = w$
E	$\frac{1}{2} n_x (n_x + n_y + r_y + 1) - (\frac{1}{2}(n_y + r_y) + 1) s_{r_y} + s^*_{r_y} = w.$

4.2 Example

Via this example we compare our NPI precedence test approach with the nonparametric precedence tests mentioned in Subsection 4.1, using a subset of Nelson's dataset (1982) on breakdown times (in minutes) of an insulating fluid that is subject to high voltage stress. The data are given in Table 4.1.

Table 4.1: Lifetimes of two samples of an insulating fluid

Group	Lifetimes									
X	0.49	0.64	0.82	0.93	1.08	1.99	2.06	2.15	2.57	4.75
Y	1.34	1.49	1.56	2.10	2.12	3.83	3.97	5.13	7.21	8.71

We compare the lifetimes of units from groups X and Y, by calculating the lower and upper probabilities for the event $X_{11} < Y_{11}$, given the stopping time T_0, the observed lifetimes of both groups before T_0, and assuming rc-$A_{(10)}$ for both groups. These lower and upper probabilities are given in Table 4.2 and Figure 4.1.

Table 4.2: Lower and upper probabilities for the event $X_{11} < Y_{11}$

T_0	r_x	r_y	$\underline{P}(X_{11}<Y_{11})$	$\overline{P}(X_{11}<Y_{11})$	T_0	r_x	r_y	$\underline{P}(X_{11}<Y_{11})$	$\overline{P}(X_{11}<Y_{11})$
$[0, 0.49)$	0	0	0	1	$[2.10, 2.12)$	7	4	0.5289	0.8512
$[0.49, 0.64)$	1	0	0.0826	1	$[2.12, 2.15)$	7	5	0.5289	0.8264
$[0.64, 0.82)$	2	0	0.1653	1	$[2.15, 2.57)$	8	5	0.5702	0.8264
$[0.82, 0.93)$	3	0	0.2479	1	$[2.57, 3.83)$	9	5	0.6116	0.8264
$[0.93, 1.08)$	4	0	0.3306	1	$[3.83, 3.97)$	9	6	0.6116	0.8182
$[1.08, 1.34)$	5	0	0.4132	1	$[3.97, 4.75)$	9	7	0.6116	0.8099
$[1.34, 1.49)$	5	1	0.4132	0.9587	$[4.75, 5.13)$	10	7	0.6364	0.8099
$[1.49, 1.56)$	5	2	0.4132	0.9174	$[5.13, 7.21)$	10	8	0.6364	0.8099
$[1.56, 1.99)$	5	3	0.4132	0.8760	$[7.21, 8.71)$	10	9	0.6364	0.8099
$[1.99, 2.06)$	6	3	0.4711	0.8760	$[8.71, \infty)$	10	10	0.6364	0.8099
$[2.06, 2.10)$	7	3	0.5289	0.8760					

Table 4.2 and Figure 4.1 show that the lower probability is increasing when r_x is increasing and remains constant when r_y is increasing. The upper probability remains constant when r_x is increasing and is decreasing when r_y is increasing, except for $r_y \geq 7$ when it remains constant as then $x_{n_x:n_x} < y_{r_y+1:n_y}$. This is in agreement with Theorem 3.2. The imprecision is decreasing when more lifetimes are observed. However, if $T_0 \geq 4.75$, increasing r_y while keeping r_x constant does not lead to less imprecision due to the fact that at that time we have observed already all lifetimes of group X ($x_{n_x:n_x} < y_{r_y+1:n_y}$) and consequently increasing r_y will not give us more information about the lifetimes of group X in as far as relevant for the nonparametric comparison with group Y.

One could interpret $\underline{P}(X_{11} < Y_{11}) > \frac{1}{2}$ as quite strong evidence that indeed $X_{11} < Y_{11}$. From Table 4.2 we see that if we had stopped the experiment at $T_0 = 2.06$ or later, then indeed $\underline{P}(X_{11} < Y_{11}) > \frac{1}{2}$. Had the experiment be stopped earlier, then the lower and upper probabilities would not suggest a strong preference between the groups.

To compare this with the nonparametric precedence tests of the previous subsection, we test H_0 : $F_X = F_Y$ against the alternative hypothesis that $F_X(x) \geq F_Y(x)$ for $x \geq 0$, with strict inequality for some x. For the classical precedence test, this implies that the p-value of the observed test statistic is $P(Q_{(r_y)} \geq \sum_{i=1}^{r_y} d_i | H_0)$ where the distribution of $Q_{(r_y)}$ is given by (4.11). For the maximal precedence test, the p-value of the observed test statistic is given by $P(U_{(r_y)} \geq d | H_0)$ where d is the observed value of $U_{(r_y)}$ and the cumulative distribution of $U_{(r_y)}$ is given by (4.12). For the Wilcoxon's minimal, maximal

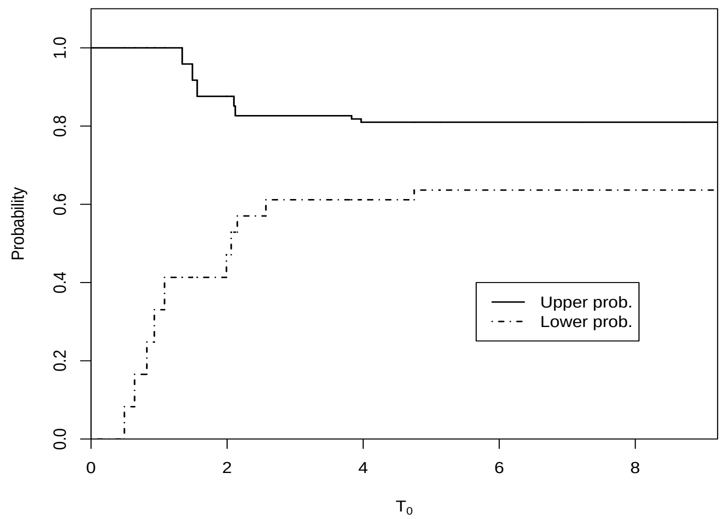

Figure 4.1: Lower and upper probabilities for the event $X_{11} < Y_{11}$

and expected rank-sum precedence test, the p-value of the test statistic is given by $P(W_{a,r_y} \leq w_a|H_0)$ $(a = \min,\ \max,\ E)$ where w_a is the observed value of the test statistic and where the distribution of the test statistic under H_0 is given by (4.16). Table 4.3 gives, for $r_y = 1, \ldots, 6$, the value of the test statistics and the corresponding p-values.

Table 4.3: Several nonparametric precedence tests

r_y	$Q_{(r_y)}$	p-value	$U_{(r_y)}$	p-value	$W_{\min,r_y}$	p-value	$W_{\max,r_y}$	p-value	W_{E,r_y}	p-value
1	5	0.0163	5	0.0163	60	0.0163	105	0.0163	82.5	0.0163
2	5	0.0704	5	0.0325	65	0.0217	105	0.0356	85	0.0356
3	5	0.1749	5	0.0487	70	0.0332	105	0.0856	87.5	0.0671
4	7	0.0894	5	0.0650	73	0.0305	91	0.0406	82	0.0348
5	7	0.1849	5	0.0812	76	0.0352	91	0.0645	83.5	0.0484
6	9	0.0704	5	0.0974	77	0.0293	81	0.0235	79	0.0247

Table 4.3 shows that the classical precedence test will not reject the null hypothesis of equal distributions at 5% significance level except when the experiment is terminated after the first lifetime of group Y. The maximal precedence test will reject the null hypothesis if the experiment is terminated after at most 3 lifetimes of group Y. Intuitively, this is logical as we have first observed 5 lifetimes of group X before the first observation of group Y and no observed lifetimes of group X between the first and third observation of group Y. Wilcoxon's minimal rank-sum precedence test always reject the null hypothesis at 5% significance level. Wilcoxon's maximal and expected rank-sum precedence tests reject the null hypothesis when the experiment is terminated early or late in the process. So, we see that although according to our nonparametric predictive precedence test there is evidence that $X_{11} < Y_{11}$ when the experiment is terminated after $T_0 = 2.06$, that is, when at least 3 lifetimes of group Y are observed, this is not supported by the classical and maximal (unless $r_y = 3$) precedence tests but the Wilcoxon's type precedence tests will support this more or less. As the NPI approach is fundamentally different to these hypothesis tests, studying the results of both might provide useful insights in practical problems.

5 Concluding remarks

The lower and upper probabilities for predictive precedence testing for two groups, presented in this paper, fit in the NPI framework and as such they have strong consistency properties in theory of interval probability (Augustin and Coolen 2004). This approach provides an attractive alternative to the more established methods for nonparametric precedence testing (Balakrishnan and Ng 2006), as instead of testing a null hypothesis the inference directly considers a comparison of the next observations from the groups considered. In classical tests, the starting point is usually the hypothesis that units of both groups have the same lifetime distributions, which is often unrealistic. So, our approach does not require a particular alternative hypothesis to be formulated, but uses a direct approach involving only future observations, enabling a natural manner of comparison that is particularly well suited if a decision must be made about e.g. the best treatment for the next unit or individual.

When considering the lower and upper probability for the event $X_{n_x+1} < Y_{n_y+1}$ as function of the stopping time T_0, we showed that these probabilities can only change at observed lifetimes for groups X or Y. In particular, we showed that, except for one special case, the lower probability is strictly increasing in r_x while keeping r_y constant, and the upper probability is strictly decreasing in r_y while keeping r_x constant. As a consequence of this, the imprecision, that is, the difference between the lower and upper probability, is decreasing as function of time and hence decreasing as function of the number of observed lifetimes.

An important issue in statistics is guidance on required design of experiments, in this situation the numbers of units to be used for both groups and choice of the stopping time for the experiment. Due to the rather minimal assumptions underlying our NPI approach, with the inferences largely based on observed data, it does not offer a satisfactory solution to this important question. However, once an experiment is underway, one can monitor the lower and upper probabilities as presented in this paper, and one can stop the experiment if one judges these to indicate strongly enough a preference between the two groups considered. Of course, before any data become available, one can study some design issues, e.g. the minimum required number of observations to possibly get a lower probability greater than a half, but as these would be based on most or least favourable configurations of the not yet observed data, indications from such studies might be of little practical value.

If the stopping time T_0 in the precedence tests, as considered in this paper, does not affect the experiment, in the sense that all units tested actually fail during the test, then the results in this paper are identical to those in NPI for pairwise comparisons presented by Coolen (1996), and a special case of the NPI results for multiple comparisons presented by Coolen and van der Laan (2001). This latter work is the basis of precedence testing for multiple groups of data, methods for which are currently in development. NPI has far more possible applications in reliability, see Coolen, Coolen-Schrijner and Yan (2002) for an early overview, other applications have since been presented. For example, Coolen and Coolen-Schrijner (2007) have recently presented NPI methods for multiple comparisons of proportions data, which are particularly of interest in reliability studies with few or zero failures in one or more groups (Coolen-Schrijner and Coolen 2007).

References

Augustin, T. and Coolen, F.P.A. (2004). Nonparametric predictive inference and interval probability. *Journal of Statistical Planning and Inference 124*, 251–272.

Balakrishnan, N. and Frattina, R. (2000). Precedence test and maximal precedence test. In N. Limnios and M. Nikulin (eds.), *Recent advances in reliability theory: Methodology, practice, and inference,* pp. 355–378. Boston: Birkhäuser.

Balakrishnan, N. and Ng, H.K.T. (2006). *Precedence-Type Tests and Applications.* New Jersey: Wiley.

Coolen, F.P.A. (1996). Comparing two populations based on low stochastic structure assumptions. *Statistical Probability Letters 29*, 297–305.

Coolen, F.P.A. (2004). On the use of imprecise probabilities in reliability. *Quality and Reliability Engineering International 20*, 193–202.

Coolen, F.P.A. (2006). On nonparametric predictive inference and objective Bayesianism. *Journal of Logic, Language and Information 15*, 21–47.

Coolen, F.P.A. and Coolen-Schrijner, P. (2007). Nonparametric predictive comparison of proportions. *Journal of Statistical Planning and Inference*, *137*, 23–33.

Coolen, F.P.A., Coolen-Schrijner, P. and Yan, K.J. (2002). Nonparametric predictive inference in reliability. *Reliability Engineering and System Safety*, *78*, 185–193.

Coolen, F.P.A. and van der Laan, P. (2001). Imprecise predictive selection based on low structure assumptions. *Journal of Statistical Planning & Inference 98*, 259–277.

Coolen, F.P.A. and Yan, K.J. (2003). Comparing two groups of lifetime data. In Bernard, et al. (Eds.), ISIPTA'03: *Proceedings of the Third International Symposium on Imprecise Probabilities and their Applications,* pp. 148-161. University of Lugano.

Coolen, F.P.A. and Yan, K.J. (2004). Nonparametric predictive inference with right-censored data. *Journal of Statistical Planning and Inference 126*, 25–54.

Coolen-Schrijner, P. and Coolen, F.P.A (2007). Nonparametric predictive comparison of success-failure data in reliability. Proceedings ARTS 2007, to appear.

De Finetti, B.(1974). *Theory of Probability* (2 volumes). London: Wiley.

Epstein, B. (1955). Comparison of some nonparametric tests against normal alternatives with an application to life testing. *Journal of the American Statistical Association 50*, 894–900.

Hill, B.M. (1968). Posterior distribution of percentiles: Bayes' theorem for sampling from a population. *Journal of the American Statistical Association 63*, 677–691.

Hill, B.M. (1988). De Finetti's theorem, induction, and $A_{(n)}$, or Bayesian nonparametric predictive inference (with discussion). In J.M. Bernando, M.H. DeGroot, D.V. Lindley and A.F.M. Smith (Eds.), *Bayesian Statistics 3,* pp 211–241. Oxford University Press.

Hill, B.M. (1993). Parametric models for $A_{(n)}$: splitting processes and mixtures. *Journal of the Royal Statistical Society B 55*, 423–433.

Lehmann, E.L. (1975). *Nonparametrics: Statistical methods based on ranks.* San Francisco: Holden-Day.

Nelson, L .(1963). Tables for a precedence life test. *Technometrics 5*, 491–499.

Nelson, L. (1982). *Applied life data analysis.* New York: John Wiley & Sons.

Ng, H.K.T. and Balakrishnan, N. (2004). Wilcoxon-type rank-sum precedence tests. *Australia and New Zealand Journal of Statistics 46*, 631–648.

Walley, P (1991). *Statistical Reasoning with Imprecise Probabilities.* London: Chapman & Hall.

Weichselberger, K (2001). *Elementare Grundbegriffe einer allgemeineren Wahrscheinlichkeitsrechnung I. Intervalwahrscheinlichkeit als umfassendes Konzept.* Heidelberg: Physika.