Nonparametric predictive selection with early experiment termination

Tahani Coolen-Maturi, Pauline Coolen-Schrijner[1], Frank P.A. Coolen*

Department of Mathematical Sciences, Durham University, Durham, DH1 3LE, 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 using lower and upper probabilities. In this paper, units from several groups are placed simultaneously on a lifetime experiment and times-to-failure are observed. The experiment may be ended before all units have failed. Depending on the available data and few assumptions, we present lower and upper probabilities for selecting the best group, the subset of best groups and the subset including the best group. We also compare our approach of selecting the best group with some classical precedence selection methods. Throughout, examples are provided to demonstrate our method.

Key words: Early experiment termination, Precedence tests, Lower and upper probabilities, Nonparametric predictive inference, Selection methods.

1. Introduction

Comparison of lifetimes of units from different groups is a common problem. In this paper, we consider the situation where units from $k \geq 2$ groups are simultaneously placed on a life-testing experiment, and decisions may

*Corresponding author.

Email addresses: tahani.maturi@durham.ac.uk (Tahani Coolen-Maturi), frank.coolen@durham.ac.uk (Frank P.A. Coolen)

[1]Pauline died in April 2008, when the research reported in this paper was at an advanced stage.

be needed before all units have failed due to cost or time considerations, so the data consist of both observed lifetimes and observations which are right-censored at the moment the experiment was terminated. In classical precedence testing, the experiment is terminated at a certain time or after a certain number of failures (for a particular group). Epstein (1955) first presented precedence testing, Nelson (1963) proposed it as an efficient life-test procedure that may enable decisions after relatively few lifetimes are observed. Balakrishnan and Ng (2006) present an excellent overview, and describe several nonparametric precedence tests based on the hypothesis of equal lifetime distributions. As an alternative, we propose nonparametric predictive precedence testing for $k \geq 2$ groups in order to select the best group, the subset of best groups and the subset including the best group.

In Section 2 we briefly describe some classical methods for precedence testing and selection. Section 3 is a short overview of nonparametric predictive inference (NPI). In Sections 4, 5 and 6 we present NPI for precedence testing for $k \geq 2$ groups in order to select the best group, the subset of best groups and the subset including the best group, respectively. Examples are provided throughout to illustrate our method and to compare it with the classical methods. Section 7 contains some concluding remarks.

2. Classical methods for precedence testing and selection

When the null-hypothesis of the equality (homogeneity) of two (or more) populations (e.g. processes, treatments) is rejected, one may want to identify which of these populations is the best. Balakrishnan and Ng (2006) introduced several nonparametric tests for this selection problem when an early decision is required (called precedence testing). Below we briefly describe these precedence selection methods following Balakrishnan and Ng (2006) in notation and definitions.

Suppose that we have independent random samples from $k \geq 2$ different populations. Let X_{j,i_j} $(i_j = 1, \ldots, n_j)$ be the lifetime of the i_jth component of a random sample from population π_j with distribution function F_j $(j = 1, \ldots, k)$. We have $N = \sum_{j=1}^{k} n_j$ units placed simultaneously on a lifetime-testing experiment. The question of interest is to test whether these populations are homogeneous, i.e. $H_0 : F_1 = F_2 = ... = F_k$, against the alternative that population π_i is the best (longer life), that is, $H_{Ai} : F_i < F_j$ for all $j \neq i$ and $j = 1, \ldots, k$. That is, it can be concluded that X_i, a random quantity representing the lifetime of a unit of population i, is stochastically larger than X_j (i.e. $X_i \succ_{st} X_j$) if and only if $F_i(x) \le F_j(x)$ for all $x \ge 0$ with strict inequality for at least one x, which we denote by $F_i < F_j$.

In precedence testing the aim is to reach a decision before all units have failed. So the experiment is terminated as soon as the $\bar{r}_i$th failure from group i is observed, where $\bar{r}_i = \lfloor n_i q \rfloor$ for $i = 1, 2, \ldots, k$ and $0 < q < 1$. Consequently, the stopping time T_0 can be defined as $T_0 = \min_{1 \le i \le k} X_{i,(\bar{r}_i)}$, where $X_{i,(\bar{r}_i)}$ is the $\bar{r}_i$th order statistic of sample i.

Suppose that the experiment is terminated at sample i, i.e. $T_0 = X_{i,(\bar{r}_i)}$, then the *Ordinary precedence statistic* (Balakrishnan et al., 2006) is defined as

$$Q^{*(i)} = \min_{\substack{1 \le j \le k \\ j \ne i}} \left(Q_j^{(i)} / n_j \right) \tag{1}$$

where $Q_j^{(i)}$ is the number of failures observed before $X_{i,(\bar{r})}$ from the remaining sample j ($j = 1, \ldots, k$, $j \ne i$). Small values of $Q^{*(i)}$ will lead to rejection of the null-hypothesis H_0, in this case one can choose H_{Aj} (π_j is the best) if and only if $(Q_j^{(i)}/n_j) = Q^{*(i)}$ for $j \ne i$ and $T_0 = X_{i,(\bar{r}_i)}$. If for two or more samples the statistic $(Q_j^{(i)}/n_j)$ is equal to $Q^{*(i)}$ then one of the corresponding populations is randomly selected as the best.

Now let $D_{j,s}^{(i)}$ be the number of failures of sample j that occur between the $(s-1)$th and sth failure of group i, $s = 1, \ldots, \bar{r}_i$. Let $W_j^{(i)}$ ($j = 1, \ldots, k$, $j \ne i$) be a random quantity defined by

$$W_j^{(i)} = \frac{1}{2} n_j (n_j + 2\bar{r}_i + 1) - (\bar{r}_i + 1) \sum_{s=1}^{\bar{r}_i} D_{j,s}^{(i)} + \sum_{s=1}^{\bar{r}_i} s \, D_{j,s}^{(i)} \tag{2}$$

Then the *Minimal Wilcoxon rank-sum statistic* (Ng et al., 2007) is given by

$$W^{*(i)} = \max_{\substack{1 \le j \le k \\ j \ne i}} \left(\frac{W_j^{(i)} - E[W_j^{(i)} | H_0]}{\sqrt{Var[W_j^{(i)} | H_0]}} \right) \tag{3}$$

where $E[W_j^{(i)}|H_0]$ and $Var[W_j^{(i)}|H_0]$ are the expected value and variance of the statistic $W_j^{(i)}$ under H_0. Large values of $W^{*(i)}$ will lead to rejection of the null-hypothesis, in which case one can choose the alternative hypothesis H_{Aj} (π_j is the best) if and only if $T_0 = X_{i,(\bar{r}_i)}$ and $(W_j^{(i)} - E[W_j^{(i)}|H_0])(Var[W_j^{(i)}|H_0])^{-1/2} = W^{*(i)}$, for $j \ne i$.

In this paper we will focus on the balanced-sample case only ($n_j = n$ for all j) when we compare our method with the classical precedence selection procedures, since these classical procedures may not be effective when the sample sizes vary much (Balakrishnan and Ng, 2006, pp.226, 265). For balanced-sample case, the statistics in (1) and (3) reduce to $Q^{*(i)} = \min Q_j^{(i)}$ and $W^{*(i)} = \max W_j^{(i)}$, over all $j = 1, \ldots, k$ and $j \neq i$, respectively. For more details we refer to Balakrishnan and Ng (2006). It should be emphasized that the NPI method presented in this paper is equally straightforward to implement for balanced-sample and unbalanced-sample cases.

A somewhat separate, yet strongly related, branch of statistical research is so-called 'selection methods', which also have the explicit target to select a single 'best' group or population or a subset of groups, along the same lines considered in this paper. The two main classical approaches in this field are indifference zone selection (Bechhofer, 1954; Bechhofer et al., 1995) and subset selection (Gupta, 1965), which were combined by Verheijen et al. (1997). Coolen and van der Laan (2001) presented NPI methods for selection, in this paper we follow the same approach with the generalization to allow early termination of the experiments, hence linking to the classical concepts of precedence testing.

3. Nonparametric predictive inference

Nonparametric predictive inference (NPI) is a statistical method based on Hill's assumption $A_{(n)}$ (Hill, 1968), which gives direct probabilities for a future observable random quantity, given observed values of related random quantities (Augustin and Coolen, 2004; Coolen, 2006).

Let $X_1, \ldots, X_n, X_{n+1}$ be positive, continuous and exchangeable random quantities representing event times. Now suppose we have observed the values of $X_1, \ldots, X_n$ and the corresponding ordered observed values are denoted by $0 < x_1 < \ldots < x_n < \infty$, we define $x_0 = 0$ for ease of notation. It is assumed that among the observed values no ties occur. However, in a similar way as described in (Hill, 1993), the results can be generalised to allow ties. For the random quantity X_{n+1} representing a future observation, based on n observations, the assumption $A_{(n)}$ (Hill, 1968) is

$$P(X_{n+1} \in (x_{j-1}, x_j)) = \frac{1}{n+1}, \; j = 1, \ldots, n, \text{ and } P(X_{n+1} \in (x_n, \infty)) = \frac{1}{n+1}$$

$A_{(n)}$ does not assume anything else and is a post-data assumption related to exchangeability (De Finetti, 1974). Inferences based on $A_{(n)}$ are predictive, nonparametric and from frequentist perspective exactly calibrated (Lawless and Fredette, 2005), and they 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; Augustin and Coolen, 2004). These NPI lower and upper probabilities have strong consistency properties in theory of interval probability (Augustin and Coolen, 2004). Hill (1993) showed that $A_{(n)}$ is coherent in a Bayesian framework under finite additivity, but we advocate NPI particularly as a frequentist statistical method that explicitly does not require prior probabilities to be specified (Coolen, 2006).

The assumption $A_{(n)}$ cannot be applied directly if the data contain censored observations. Right-censoring in specific is common in data in reliability and survival analysis. Coolen and Yan (2004) presented a generalization of $A_{(n)}$, called rc-$A_{(n)}$, suitable for right-censored data. 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. Recently, Maturi et al. (2010) presented the application of rc-$A_{(n)}$ to data resulting from progressive censoring.

4. Selecting the best group

In precedence testing, units of all groups are placed simultaneously on a life-testing experiment, and failures are observed as they arise during the experiment. The experiment is terminated as soon as a certain stop criterion has been reached, so the lifetimes of some units are typically right-censored. 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 as long as T_0 contains no further information on the residual event times beyond T_0 for right-censored units. 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. Here, 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, i.e. M-function (Coolen and Yan, 2004) which is very similar to the concept of Shafers 'basic probability assignment' (Shafer, 1976). The next definition gives the assumption rc-$A_{(n)}$ for precedence testing.

Definition 1. rc-$A_{(n)}$ for precedence testing
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 observed event times and $n - r$ right-censored observations, is partially specified by the following M-function values:

$$M_{X_{n+1}}(x_{i-1}, x_i) = \frac{1}{n+1} \quad ; \; i = 1, \ldots, r$$

$$M_{X_{n+1}}(x_r, \infty) = \frac{1}{n+1} \; \text{ and } \; M_{X_{n+1}}(T_0, \infty) = \frac{n-r}{n+1}.$$

Coolen-Schrijner et al. (2009) introduced NPI for precedence testing for two groups. They derived the lower and upper probabilities for the event that a future observation from one group is less than a future observation from the second group, based on the observations of both groups, stopping time T_0, and the assumption rc-$A_{(n)}$ per group. In this paper we extend NPI for precedence testing for $k \geq 2$ groups in order to select the best group, the subset of best groups, and the subset including the best group.

Suppose we have $k \geq 2$ groups and $n_j + 1$ random quantities from group j, denoted by X_{j,i_j} where $i_j = 1, 2, \ldots, n_j, n_j + 1$, $j = 1, 2, \ldots, k$. For each group j, n_j units are put on a lifetime experiment and we are interested in the behaviour of the future random variable X_{j,n_j+1}. Therefore, we have $N = \sum_{j=1}^{k} n_j$ units under the lifetime experiment and the expert may want to terminate the experiment at certain time T_0. Let $0 = x_{j,0} < x_{j,1} < x_{j,2} < \ldots < x_{j,r_j} \leq T_0 < \infty$ be the ordered observed values (failures) from group j, $j = 1, 2, \ldots, k$.

These observed values from group j produce r_j+2 intervals, where the first r_j intervals are defined by $I^j_{i_j} = (x_{j,i_j-1}, x_{j,i_j})$, $i_j = 1, 2, \ldots, r_j$, $j = 1, 2, \ldots, k$, and the remaining intervals are defined by $I^j_{r_j+1} = (x_{j,r_j}, \infty)$, $I^j_{r_j+2} = (T_0, \infty)$.

Let $L(I^j_{i_j})$ and $U(I^j_{i_j})$ be the lower and the upper bounds for the interval $I^j_{i_j}$, $i_j = 1, 2, \dots, r_j + 2$, $j = 1, 2, \dots, k$. That is, $L(I^j_{i_j}) = x_{j,i_j-1}$ for $i_j = 1, \dots, r_j + 1$ and $L(I^j_{r_j+2}) = T_0$. Similar for the upper bound, $U(I^j_{i_j}) = x_{j,i_j}$ for $i_j = 1, \dots, r_j$, and $U(I^j_{r_j+1}) = U(I^j_{r_j+2}) = \infty$. Here the intervals $I^j_{i_j}$ are open intervals, but in future when we mention the (left or right) end points we actually mean the limit end points which are not included in these open intervals.

Our inference depends on assumption rc-$A_{(n_j)}$ (Coolen and Yan, 2004). It does not assume any underlying distribution. Beyond the data, our method requires the exchangeability assumption of the random variables per group to be met. We also assume that groups are independent. We will specify partially the probability distribution for a future quantity, X_{j,n_j+1}, $j = 1, 2, \dots, k$, using M-functions (see Definition 1). We should mention here that rc-$A_{(n_j)}$ does not allow ties between failures to occur. Dealing with ties is straightforward (Coolen and Yan, 2004) by assuming that these tied observations differ by small amounts which tend to zero.

The lower and upper probability that the lifetime of the next observation from one group, say l, is greater than the lifetime of the next observation from each other group, that is

$$\underline{P}^{(l)} = \underline{P}\left(X_{l,n_l+1} = \max_{1\le j\le k} X_{j,n_j+1}\right) \quad \text{and} \quad \overline{P}^{(l)} = \overline{P}\left(X_{l,n_l+1} = \max_{1\le j\le k} X_{j,n_j+1}\right),$$

are given in Theorem 1. The indicator function $1\{E\}$ is equal to 1 if event E occurs and 0 else.

Theorem 1. *The lower and upper probabilities that the lifetime of the next observation from group l is greater than the lifetime of the next observation from each other group are given by*

$$\underline{P}^{(l)} = \frac{1}{\prod_{j=1}^{k}(n_j+1)} \left\{ \sum_{i_l=1}^{r_l} \prod_{\substack{j=1\\ j\neq l}}^{k} \sum_{i_j=1}^{r_j} 1\{x_{j,i_j} < x_{l,i_l}\} + (n_l - r_l) \prod_{\substack{j=1\\ j\neq l}}^{k} r_j \right\} \tag{4}$$

$$\overline{P}^{(l)} = \frac{1}{\prod_{j=1}^{k}(n_j+1)} \sum_{i_l=1}^{r_l} \prod_{\substack{j=1\\ j\neq l}}^{k} \left(1 + \sum_{i_j=1}^{r_j} 1\{x_{j,i_j} < x_{l,i_l}\}\right) + \frac{n_l - r_l + 1}{n_l + 1} \tag{5}$$

Proof. The proof is given in the appendix.

In the rest of this section, we discuss some properties and special cases of these lower and upper probabilities, which are easily verified from (4) and (5).

1. If $r_l \geq 0$ and $r_j \geq 0$, and there exists at least one $j \neq l$ for which $r_j = 0$, then the lower probability is $\underline{P}^{(l)} = 0$, since we have not seen any failure from group $j \neq l$. Hence, we cannot exclude the possibility that we will never see a failure from group(s) $j \neq l$. Further, if $r_l = 0$ then the upper probability $\overline{P}^{(l)}$ will be equal to one.

2. If $r_l = 0$ and $r_j > 0$, for all $j \neq l$, then the upper probability $\overline{P}^{(l)}$ is one, as we cannot exclude the possibility that we will never see a failure of group l. The corresponding lower probability is

$$\underline{P}^{(l)} = \frac{n_l}{\prod_{j=1}^{k}(n_j+1)} \prod_{\substack{j=1 \\ j \neq l}}^{k} r_j$$

Further, if $r_l = 0$ and $r_j = n_j$ for all $j \neq l$, that is we have observed all units from each group $j \neq l$ and the experiment is ended before we observe any failure from group l, then the lower probability is

$$\underline{P}^{(l)} = \prod_{j=1}^{k} \frac{n_j}{n_j+1}$$

3. If $r_l > 0$ and $r_j = 0$, for all $j \neq l$, so we have not seen any failure for all groups $j \neq l$, we cannot exclude the possibility that we will never see a failure of these groups and consequently $\underline{P}^{(l)} = 0$. The corresponding upper probability is

$$\overline{P}^{(l)} = \frac{r_l}{\prod_{j=1}^{k}(n_j+1)} + \frac{n_l - r_l + 1}{n_l + 1}$$

Further, if $r_l = n_l$ and $r_j = 0$, that is we have observed all units from group l and the experiment is ended before we observe any failure from all other groups, then the upper probability is

$$\overline{P}^{(l)} = \frac{n_l}{\prod_{j=1}^{k}(n_j+1)} + \frac{1}{n_l+1}$$

4. If $r_l > 0$, $r_j > 0$ and $x_{j,r_j} < x_{l,1}$, for all $j \neq l$, then the lower and upper probabilities are

$$\underline{P}^{(l)} = \frac{n_l}{\prod_{j=1}^{k}(n_j+1)} \prod_{\substack{j=1\\ j\neq l}}^{k} r_j, \quad \overline{P}^{(l)} = \frac{r_l}{\prod_{j=1}^{k}(n_j+1)} \prod_{\substack{j=1\\ j\neq l}}^{k} (r_j+1) + \frac{n_l - r_l + 1}{n_l+1}$$

But if $x_{j,1} > x_{l,r_l}$, for all $j \neq l$, then the lower and upper probabilities are

$$\underline{P}^{(l)} = \frac{(n_l - r_l)}{\prod_{j=1}^{k}(n_j+1)} \prod_{\substack{j=1\\ j\neq l}}^{k} r_j, \quad \overline{P}^{(l)} = \frac{r_l}{\prod_{j=1}^{k}(n_j+1)} + \frac{n_l - r_l + 1}{n_l+1}$$

Now, we study the effect upon the lower and upper probabilities when the stopping time is increased from T_0 to $T_0 + \epsilon$, for small $\epsilon > 0$ just to allow one failure from one group to occur.

Theorem 2. *(i) If a failure occurs from group l then the lower probability $\underline{P}^{(l)}$ remains constant. However the upper probability $\overline{P}^{(l)}$ decreases by*

$$\frac{1}{n_l+1} + \frac{1}{\prod_{j=1}^{k}(n_j+1)} \prod_{\substack{j=1\\ j\neq l}}^{k} (r_j+1)$$

except when $r_j = n_j$, for all $j \neq l$, in which case the upper probability remains constant.

(ii) If a failure occurs for group j^*, where $j^* \in \{1, \ldots, k\} \setminus \{l\}$, then the upper probability $\overline{P}^{(l)}$ remains constant. However, the lower probability increases by

$$\frac{n_l - r_l}{\prod_{j=1}^{k}(n_j+1)} \prod_{\substack{j=1\\ j\neq \{l,j^*\}}}^{k} r_j$$

except when $r_l = n_l$, or when at least one $r_j = 0$ for a $j \neq \{j^*, l\}$, in which cases the lower probability remains constant.

Proof. For case i (ii), replace r_l (r^*_j) by $r_l + 1$ ($r^*_j + 1$) in the formula (4) and (5), then this follows by basic analysis of the lower and upper probabilities of Theorem 1.

Theorem 2 shows that the lower probability for a certain event quantifies the amount of information in favour of the event while the upper probability quantifies the amount of information against the event. If r_l is increased while leaving all other r_j the same, then, when considering the event $X_{l,n_l+1} = \max_{1\leq j\leq k} X_{j,n_j+1}$, the amount of information in favour of this event remains the same but the amount of information against this event increases except when $r_j = n_j$ for all $j \neq l$. Consequently, $\underline{P}^{(l)}$ does not change but $\overline{P}^{(l)}$ may decrease. For the same event, when r_j for a $j \neq l$ increases while all other r_i, $i \neq j$, remain constant, the amount of information in favour of the event increases except when $r_l = n_l$ or when there exists a $j \neq \{l, j^*\}$ for which $r_j = 0$, while the amount of information against the event remains the same. Consequently, $\underline{P}^{(l)}$ may increase but $\overline{P}^{(l)}$ does not change.

With regard to making a decision using the NPI method, one can claim to have a strong indication that group l is the best group if $\overline{P}^{(j)} < \underline{P}^{(l)}$ for all $j \neq l$, and a weak indication that group l is the best group if $\underline{P}^{(j)} < \underline{P}^{(l)}$ and $\overline{P}^{(j)} < \overline{P}^{(l)}$ for all $j \neq l$. In discussions in the examples in this paper, we will call one group 'better' than another, or 'best', if the first of these conditions is satisfied, of course the use of 'better' and 'best' must be interpreted with care as these judgements are just based on direct comparison of one next observation for each group according to the NPI method.

We illustrate our method for selecting the best group via examples. We also show that the NPI method and the classical precedence tests do not necessarily lead to the same conclusions, but it is difficult to compare these two due to the different inferential goals and the different basic underlying assumptions. Hence, we do not see these as competing methods for the same problems, but more as complementary methods that can provide further insight into specific applications, and which may be more or less suitable depending on the explicit inferential goal.

Example 1. To illustrate our method for selecting the best group among k other groups, we use the data from Coolen and van der Laan (2001) as presented in Table 1.

This data set consists of four groups and is used by Coolen and van der

Group										
1	5.01	5.04	5.60	5.78	6.43	6.53	6.96	7.00	7.21	7.58
	8.12	8.26	8.27	8.34	8.62	8.66	8.91	8.94	9.05	9.16
2	4.50	4.86	5.10	5.15	5.17	5.34	5.99	6.18	6.72	7.39
	7.44	7.46	7.47	7.76	8.38	8.42	8.52	8.81		
3	6.84	6.91	7.22	7.24	7.25	7.35	7.55	7.62	7.69	7.98
	7.99	8.04	8.08	8.18	8.97					
4	4.71	8.20	9.03							

Table 1: Coolen and van der Laan (2001) Data Set

Laan (2001) in order to demonstrate the NPI method for selection of the best source and a subset to include the best source for complete data (no censoring). We interpret this data set as the lifetimes of units from 4 different groups. The size of the groups are $n_1 = 20$, $n_2 = 18$, $n_3 = 15$ and $n_4 = 3$, and X_{j,i_j} $(i_j = 1, \ldots, n_j)$ represents the lifetime of unit i_j in group j.

A group is considered as the 'best' when the lifetime of a future unit from this group is larger than the lifetime of a future unit from all other groups. Our inference depends on the data, the rc-$A_{(n_j)}$ assumptions $(j = 1, 2, 3, 4)$ and stopping time T_0. Table 2 presents the lower and upper probabilities for the event that the lifetime of a future unit of group l $(l = 1, 2, 3, 4)$ is larger than the lifetimes of a future unit of all other groups, as given by (4) and (5), for stopping time T_0 in several intervals. We denote these lower and upper probabilities by $\underline{P}^{(l)}$ and $\overline{P}^{(l)}$, respectively.

Let us consider the situation when we terminate the experiment at $T_0 = 5$. Until this point we observed only two failures from group 2 and one failure from group 4, however, we have not yet observed any failures from groups 1 and 3. Here all lower probabilities are equal to zero since for each l, there exists a group $j \neq l$ for which we have not observed a failure yet. Moreover, while the upper probabilities for the first and third groups are equal to 1, those for groups 2 and 4 are less than 1, being equal 0.895 and 0.750, respectively. As no failure is observed from group 1 and 3, we cannot exclude the possibility that we will never observe any failure from these groups and consequently $\overline{P}^{(1)} = \overline{P}^{(3)} = 1$.

At $T_0 = 6$, we still have not observe any failure from group 3, so we cannot exclude the possibility that we will never observe any failure from this group and consequently $\overline{P}^{(3)} = 1$. However, the lower probability for this group is now positive as there is no other group for which we have not observed a failure. For all other groups the lower probability is still zero as we have not

T_0	r_1	r_2	r_3	r_4	$\underline{P}^{(1)}$	$\overline{P}^{(1)}$	$\underline{P}^{(2)}$	$\overline{P}^{(2)}$	$\underline{P}^{(3)}$	$\overline{P}^{(3)}$	$\underline{P}^{(4)}$	$\overline{P}^{(4)}$
$[4.86, 5.01)$	0	2	0	1	0	1	0	0.895	0	1	0	0.750
$[5.99, 6.18)$	4	7	0	1	0	0.811	0	0.633	0.016	1	0	0.750
$[7.00, 7.21)$	8	9	2	1	0.010	0.627	0.006	0.529	0.041	0.886	0.011	0.750
$[7.21, 7.22)$	9	9	3	1	0.014	0.582	0.010	0.529	0.046	0.831	0.019	0.750
$[7.22, 7.24)$	9	9	4	1	0.018	0.582	0.013	0.529	0.046	0.777	0.025	0.750
$[7.46, 7.47)$	9	12	6	1	0.033	0.582	0.019	0.387	0.055	0.667	0.051	0.750
$[7.47, 7.55)$	9	13	6	1	0.036	0.582	0.019	0.340	0.058	0.667	0.055	0.750
$[7.55, 7.58)$	9	13	7	1	0.041	0.582	0.021	0.340	0.058	0.616	0.064	0.750
$[7.99, 8.04)$	10	14	11	1	0.066	0.543	0.029	0.296	0.065	0.416	0.121	0.750
$[8.42, 8.52)$	14	16	14	2	0.164	0.448	0.073	0.244	0.077	0.268	0.207	0.606
$[8.62, 8.66)$	15	17	14	2	0.171	0.432	0.075	0.218	0.080	0.268	0.224	0.606
$[8.66, 8.81)$	16	17	14	2	0.171	0.416	0.076	0.218	0.081	0.268	0.234	0.606
$[8.81, 8.91)$	16	18	14	2	0.175	0.416	0.076	0.195	0.082	0.268	0.242	0.606
$[8.91, 8.94)$	17	18	14	2	0.175	0.402	0.076	0.195	0.084	0.268	0.252	0.606
$[8.94, 8.97)$	18	18	15	2	0.178	0.388	0.076	0.195	0.085	0.248	0.275	0.606
$[8.97, 9.03)$	18	18	15	2	0.178	0.388	0.076	0.195	0.085	0.248	0.275	0.606
$[9.03, 9.05)$	18	18	15	3	0.199	0.388	0.076	0.195	0.085	0.2481	0.275	0.582
$[9.05, 9.16)$	19	18	15	3	0.199	0.388	0.076	0.195	0.085	0.248	0.275	0.582
$[9.16, \infty)$	20	18	15	3	0.199	0.388	0.076	0.195	0.085	0.248	0.275	0.582

Table 2: The best group: lower and upper probabilities (Example 1)

seen any failure yet from group 3.

From Theorem 2 we know that the lower probability never decreases and the upper probability never increases. For example, consider the situation where the stopping time T_0 is increased from 7.50 to 7.55. At $T_0 = 7.55$, a failure of group 3 occurs. We want to calculate the lower and upper probabilities for the event $X_{3,n_3+1} = \max_{1\leq j\leq 4} X_{j,n_j+1}$. For this case, the lower probability remains constant ($\underline{P}^{(3)} = 0.058$), but the upper probability decreases from 0.667 at $T_0 = 7.50$ to 0.616 at $T_0 = 7.55$, which illustrates Theorem 2(i). However, for the event $X_{1,n_1+1} = \max_{1\leq j\leq 4} X_{j,n_j+1}$ the upper probability remains constant ($\overline{P}^{(1)} = 0.582$) but the lower probability increases from 0.036 at $T_0 = 7.50$ to 0.041 at $T_0 = 7.55$, which illustrates Theorem 2(ii).

There are some special cases when all lower and upper probabilities remain constant when a failure occurs from any group. For example, at $T_0 = 9.03$ we have observed all units from all groups except the first group which still has two units have not failed. Let $l = 1$ and assume we will allow for an extra failure to occur. Here of course the failure must be from the first group ($T_0 = 9.05$). In this case all lower and upper probabilities remain as they were at $T_0 = 9.03$, as the amount of information in favour and against the event does not change. In fact, all lower and upper probabilities do not change anymore after 9.03.

At $T_0 = 8.81$, we have observed failure times of all units from the second group. Now consider $l = 2$ and let the stopping time increase to 8.91, so that we observe an extra failure of group 1. In this case the lower and upper probabilities, remain constant ($\underline{P}^{(2)} = 0.076$ and $\overline{P}^{(2)} = 0.195$). In fact, any failure from other groups after we have observed failures of all units from group 2 will not affect the lower and upper probabilities $\underline{P}^{(2)}$ and $\overline{P}^{(2)}$.

From $T_0 = 7.55$ on, the fourth group has the greatest lower and upper probabilities. However, from the beginning of the experiment till $T_0 = 7.22$ the third group has the greatest lower and upper probabilities. Which means that at $T_0 \leq 7.22$, there is some weak indication that group 3 is best, since $\underline{P}^{(j)} < \underline{P}^{(3)} < \overline{P}^{(j)} < \overline{P}^{(3)}$ for $j = 1, 2, 4$, however, there is some weak indication that group 4 is best for $T_0 \geq 7.55$, since $\underline{P}^{(j)} < \underline{P}^{(4)} < \overline{P}^{(j)} < \overline{P}^{(4)}$ for $j = 1, 3, 4$. Also we can note that $\overline{P}^{(3)}$ remains equal to 1 for quite a long time, since the first failure from the third group occurs relatively late.

The difference between corresponding lower and upper probabilities, called imprecision, decreases as the stopping time T_0 increases, which reflects the amount of information we have (Table 2). For example, we can see that the fourth group has larger imprecision as there are only a few observations in this group.

A crucial question is how to make decisions using these NPI lower and upper probabilities. If we observe all units from groups 2, 3, and 4, so for $T_0 \geq 9.03$, we see that $\overline{P}^{(2)} < \underline{P}^{(1)}$ and $\overline{P}^{(2)} < \underline{P}^{(4)}$ implying that group 1 and 4 are certainly better than group 2. Also $\overline{P}^{(3)} < \underline{P}^{(4)}$ implying that group 4 is better than group 3. It is still difficult to distinguish between group 1 and 4. As $\underline{P}^{(1)} < \underline{P}^{(4)} < \overline{P}^{(1)} < \overline{P}^{(4)}$ there is a weak preference for group 4. For $T_0 \geq 8.62$, group 4 is better than group 2 ($\overline{P}^{(2)} < \underline{P}^{(4)}$), and for $T_0 \geq 8.97$, group 4 is better than group 2 and 3 ($\overline{P}^{(2)} < \underline{P}^{(4)}$ and $\overline{P}^{(3)} < \underline{P}^{(4)}$). However, we have to be careful as group 4 only has 3 observations and its imprecision is large. Therefore, we will now exclude the fourth group from the comparison and we will recompute the lower and upper probabilities to study the effect of the fourth group on the comparison.

Table 3 presents lower and upper probabilities (4) and (5) after we have excluded the fourth group from this comparison, to study the effect of this group on our inferences. For example, at $T_0 = 8.42$ we observed 14, 16, and 14 failures from group 1, 2 and 3, respectively. Here $\overline{P}^{(2)} < \underline{P}^{(1)}$ which indicates that the first group is better than the second group. The second

T_0	r_1	r_2	r_3	$\underline{P}^{(1)}$	$\overline{P}^{(1)}$	$\underline{P}^{(2)}$	$\overline{P}^{(2)}$	$\underline{P}^{(3)}$	$\overline{P}^{(3)}$
$[4.86, 5.01)$	0	2	0	0	1	0	0.895	0	1
$[5.99, 6.18)$	4	7	0	0	0.813	0	0.635	0.066	1
$[7.00, 7.21)$	8	9	2	0.040	0.634	0.023	0.531	0.164	0.897
$[7.47, 7.55)$	9	13	6	0.143	0.592	0.076	0.365	0.233	0.710
$[7.55, 7.58)$	9	13	7	0.165	0.592	0.083	0.365	0.233	0.669
$[7.99, 8.04)$	10	14	11	0.264	0.561	0.117	0.329	0.258	0.519
$[8.42, 8.52)$	14	16	14	0.354	0.510	0.171	0.294	0.274	0.411
$[8.62, 8.66)$	15	17	14	0.367	0.504	0.173	0.277	0.279	0.411
$[8.66, 8.81)$	16	17	14	0.367	0.499	0.175	0.277	0.281	0.411
$[8.81, 8.91)$	16	18	14	0.376	0.499	0.175	0.264	0.284	0.411
$[8.91, 8.94)$	17	18	14	0.376	0.496	0.175	0.264	0.287	0.411
$[8.94, 8.97)$	18	18	14	0.376	0.493	0.175	0.264	0.289	0.411
$[8.97, 9.05)$	18	18	15	0.381	0.493	0.175	0.264	0.289	0.405
$[9.05, 9.16)$	19	18	15	0.381	0.493	0.175	0.264	0.289	0.405
$[9.16, \infty)$	20	18	15	0.381	0.493	0.175	0.264	0.289	0.405

Table 3: The best group: lower and upper probabilities (Example 1, without the fourth group)

group would be the worst group for $T_0 \geq 8.62$ since then $\overline{P}^{(2)} < \underline{P}^{(1)}$ and $\overline{P}^{(2)} < \underline{P}^{(3)}$. When observing all units from all groups, there exists a weak preference for group 1 compared to group 3 as $\underline{P}^{(3)} < \underline{P}^{(1)} < \overline{P}^{(3)} < \overline{P}^{(1)}$. In addition, the imprecision is slightly larger for group 3 than for group 1.

Furthermore, as we can see from Tables 2 and 3, dropping group 4 leads to significance increases in the NPI lower and upper probabilities for both the first and the third group when we exclude the fourth group, with slight increases in the lower and upper probabilities for the second group. However, it is still not possible to make a clear decision on which group will have the largest next observation. Removing the fourth group has an influence not only on improving the lower and upper probabilities but also on reducing the imprecision for other groups.

Example 2. In this example, we compare our method with the classical precedence selection methods in order to select the best group. Table 4 shows the natural logarithm of times to breakdown of an insulating fluid at three voltage levels (30kv, 35kv and 40kv), as given by Nelson (1982, p.278). We will refer to these voltage levels as groups $j = 1, 2, 3$, respectively. Here we have a balanced-sample case where $n_1 = n_2 = n_3 = 12$. Let X_{j,i_j} represent the natural logarithm of time to breakdown from the i_jth unit of voltage level j, $i_j = 1, \ldots, 12$ and $j = 1, 2, 3$. Balakrishnan et al. (2006) and Ng et al. (2007) considered the last two values at level 30kv as real failures although

Group		Data									
1	30kv	3.912	4.898	5.231	6.782	7.279	7.293	7.736	7.983	8.338	9.668
		10.282^+	11.363^+								
2	35kv	3.401	3.497	3.715	4.466	4.533	4.585	4.754	5.553	6.133	7.073
		7.208	7.313								
3	40kv	0.000	0.000	0.693	1.099	2.485	3.219	3.829	4.025	4.220	4.691
		5.778	6.033								

Table 4: Times (ln) to breakdown of an insulating fluid Nelson (1982, p.278)

they are in fact censored observations. We follow their approach, although as these values are larger than all observations for the other groups, it makes no difference to our approach for any $T_0 < 10.282$.

The classical precedence selection procedures normally test the homogeneity of the lifetime distributions against the alternative that one distribution stochastically dominates the other distributions (so one population is the best) in terms of their reliability (longer life). That means, the classical selection procedures are designed to test $H_0 : F_1 = F_2 = F_3$ in favour of the alternative $H_{Ai} : F_i < F_j$ (for all $j \neq i,\ j = 1, 2, 3$) when the stop criterion is expressed in terms of group i ($i = 1, 2, 3$).

Here we have $k = 3$ lifetime samples with equal sample sizes. We will stop the experiment as soon as the 8th failure ($\bar{r} = 8$) from any sample has occurred following (Ng et al., 2007). So the experiment is terminated at $T_0 = 4.025$, when the 8th breakdown time of group 3 is observed. The test statistic for the ordinary precedence test, calculated from (1), is $Q^{*(3)} = \min\{1, 3\} = 1$ and the p-value of this test is 0.02557. The minimal Wilcoxon rank-sum precedence test statistics (3) equals $W^{*(3)} = \max\{173, 168\} = 173$ and the p-value is 0.00662.

In this case, at significance level 5%, we reject the null-hypothesis for both test statistics $Q^{*(3)}$ and $W^{*(3)}$, and therefore we will select the first population (30kv) as the best, i.e. we reject H_0 in favour of H_{A1}. We would get a different decision at significance level 1%, for which both the minimal Wilcoxon rank-sum precedence selection method and the ordinary precedence selection method do not lead to rejection of the null-hypothesis. In such a situation it is a good idea to apply our method to the data to see whether our method leads to a 'best' or 'worst' group. Table 5 contains the lower and upper probabilities that the lifetime of the next observation of group l ($l = 1, 2, 3$) is larger than the lifetime of the next observation of each other group for certain values of T_0.

Table 5 shows that, after we have only observed three failures from group 3

T_0	r_1	r_2	r_3	$\underline{P}^{(1)}$	$\overline{P}^{(1)}$	$\underline{P}^{(2)}$	$\overline{P}^{(2)}$	$\underline{P}^{(3)}$	$\overline{P}^{(3)}$
0.693	0	0	3	0	1	0	1	0	0.771
3.401	0	1	6	0.033	1	0	0.926	0	0.541
4.025	1	3	8	0.130	0.938	0.033	0.779	0.007	0.393
4.691	1	6	10	0.310	0.938	0.040	0.575	0.011	0.249
4.898	2	7	10	0.360	0.901	0.062	0.508	0.018	0.249
6.033	3	8	12	0.467	0.864	0.096	0.452	0.027	0.128
6.133	3	9	12	0.516	0.864	0.096	0.398	0.027	0.128
7.293	6	11	12	0.603	0.834	0.123	0.304	0.027	0.128
11.363	12	12	12	0.636	0.834	0.123	0.268	0.027	0.128

Table 5: The best group: lower and upper probabilities (Example 2)

(40kv, $T_0 = 0.693$) we cannot make any reasonable decision ($\underline{P}^{(1)} = \underline{P}^{(2)} = 0$ and $\overline{P}^{(1)} = \overline{P}^{(2)} = 1$) on whether the first or the second group is the best, since we have not yet observed any failures from both groups.

However, at $T_0 = 4.025$, when we have observed the 8th failure from group 3 we have $\underline{P}^{(3)} < \underline{P}^{(2)} < \underline{P}^{(1)}$ and $\overline{P}^{(3)} < \overline{P}^{(2)} < \overline{P}^{(1)}$ but $\overline{P}^{(3)} \nleq \underline{P}^{(1)}$ or $\overline{P}^{(3)} \nleq \underline{P}^{(2)}$. So, there is no strong indication to select group 3 as the best group, which is in agreement with the ordinary precedence test at significance level 1%, that one particular group is the worst. However, there is some weak indication that group 3 is the worst, but this does not follow from the classical methods.

Here, when we have observed all failures, we have a strong indication that group 1 is the best as $\overline{P}^{(2)} < \underline{P}^{(1)}$ and $\overline{P}^{(3)} < \underline{P}^{(1)}$. In fact this holds already at $T_0 = 6.033$. At $T_0 = 4.691$ we can conclude already that group 1 is better than group 3 as from that moment on we have $\overline{P}^{(3)} < \underline{P}^{(1)}$. Then at $T_0 = 6.133$, we also have in addition $\overline{P}^{(2)} < \underline{P}^{(1)}$ and consequently we can conclude that group 1 is the best. So at $T_0 = 4.025$, our method supports the minimal Wilcoxon rank-sum precedence selection method.

It is worth to mention here that it is difficult to compare the different methods, since the classical precedence selection methods are testing the equality of some distributions depending on the data at hand to select the best population. Our method depends on the data and the rc-$A_{(n)}$ assumptions per group to compare a future (predicted) observation from all groups in the data set in order to derive lower and upper probabilities that the lifetime of the next observation from one group is greater than the lifetime of the next observation of each other group.

5. Selecting the subset of best groups

Suppose that the experiment is terminated at time T_0 and our interest is to select a subset of groups such that all the groups in this subset are 'better' than all not selected groups, that is the lifetime of the next observation of each group in the subset will be larger than the lifetime of the next observation of all groups not in the subset. Let $S = \{l_1, l_2, ..., l_m\}$ be a subset of m groups ($1 \leq m \leq k-1$) from k independent groups, and let NS be the complementary set of S which contains the remaining $k - m$ groups.

We are interested in the lower and upper probabilities for the event that the next observation of each group in S has longer lifetime than the next observation of each group in NS. These probabilities are denoted by

$$\underline{P}^S = \underline{P}\left(\min_{l \in S} X_{l,n_l+1} > \max_{j \in NS} X_{j,n_j+1}\right) \text{ and } \overline{P}^S = \overline{P}\left(\min_{l \in S} X_{l,n_l+1} > \max_{j \in NS} X_{j,n_j+1}\right)$$

These NPI lower and upper probabilities are given in Theorem 3, where the following notation is used

$$\sum_{\substack{i_l=1\\ l\in S}}^{r_l+2} = \sum_{i_{l_1}=1}^{r_{l_1}+2} \cdots \sum_{i_{l_m}=1}^{r_{l_m}+2} \tag{6}$$

Theorem 3. *The lower and upper probabilities for the event that the next observation of each group in S has longer lifetime than the next observation of each group in NS are given by*

$$\underline{P}^S = \sum_{\substack{i_l=1\\ l\in S}}^{r_l+2} \prod_{j\in NS} \left[\frac{\sum_{i_j=1}^{r_j} 1\{x_{j,i_j} < \min_{l\in S}\{L(I_{i_l}^l)\}\}}{n_j+1}\right] \cdot \prod_{l\in S} M_{X_{l,n_l+1}}(I_{i_l}^l) \tag{7}$$

$$\overline{P}^S = \sum_{\substack{i_l=1\\ l\in S}}^{r_l+2} \prod_{j\in NS} \left[\frac{1+\sum_{i_j=1}^{r_j} 1\{x_{j,i_j} < \min_{l\in S}\{U(I_{i_l}^l)\}\}}{n_j+1} + \frac{(n_j-r_j)1\{T_0 < \min_{l\in S}\{U(I_{i_l}^l)\}\}}{n_j+1}\right] \cdot \prod_{l\in S} M_{X_{l,n_l+1}}(I_{i_l}^l) \tag{8}$$

Proof. The proof is given in the appendix.

We now present some properties and special cases of these lower and upper probabilities (7) and (8).

1. If $r_l = 0$ for all $l \in S$ and $r_j \geq 0$, for all $j \in NS$, then the lower probability is
$$\underline{P}^S = \prod_{j \in NS} \frac{r_j}{n_j + 1} \prod_{l \in S} \frac{n_l}{n_l + 1}$$
and $\underline{P}^S = 0$ if there exists at least one $j \in NS$ for which $r_j = 0$. Since we have not seen any failure from any group in S, this means that we cannot exclude the situation that we will never see a failure from any group in S, consequently $\overline{P}^S = 1$.

2. If $r_j = 0$ for at least one $j \in NS$ and $r_l \geq 0$, for all $l \in S$, then the lower probability $\underline{P}^S = 0$ since there exists a group in NS for which we have not seen any failure. This means that we cannot exclude the situation that we will never see a failure from this group. Further, if $r_j = 0$ for all $j \in NS$ then the upper probability is
$$\overline{P}^S = \prod_{l \in S} \frac{n_l - r_l + 1}{n_l + 1} \left(1 - \prod_{j \in NS} \frac{1}{n_j + 1} \right) + \prod_{j \in NS} \frac{1}{n_j + 1}$$

Now, we study the effect upon the lower and upper probabilities (7) and (8) when the stopping time is increased from T_0 to $T_0 + \epsilon$, for small $\epsilon > 0$ just to allow one failure from one group to occur.

Theorem 4. *(i) If a failure from group $l^* \in S$ occurs in the interval $(T_0, T_0 + \epsilon)$, then the lower probability $\underline{P}^S$ remains constant. However, the upper probability $\overline{P}^S$ decreases by*
$$\frac{1}{n_{l^*} + 1} \prod_{l \in S \setminus \{l^*\}} \frac{n_l - r_l + 1}{n_l + 1} . \left(1 - \prod_{j \in NS} \frac{r_j + 1}{n_j + 1} \right)$$
except when $r_j = n_j$, for all $j \in NS$, in which case the upper probability remains constant.

(ii) If a failure from group $j^* \in NS$ occurs in the interval $(T_0, T_0 + \epsilon)$, then the upper probability $\overline{P}^S$ remains constant. However, the lower probability $\underline{P}^S$ increases by

$$\frac{1}{n_{j^*}+1} \prod_{j \in NS \setminus \{j^*\}} \frac{r_j}{n_j+1} \prod_{l \in S} \frac{n_l - r_l}{n_l+1}$$

except when $r_l = n_l$ for at least one $l \in S$ or when there exists a $j \in NS \setminus \{j^*\}$ for which $r_j = 0$, in which cases the lower probability remains constant.

Proof. For case i (ii), replace r_l (r_j^*) by $r_l + 1$ ($r_j^* + 1$) in the formula (7) and (8), then this follows by basic analysis of the lower and upper probabilities of Theorem 3.

Example 3. We use the data set of Example 1 to illustrate our method for selecting the subset of best groups. A subset S is considered as the 'best' when the lifetime of a future unit from each group in S is larger than the lifetime of a future unit from each group outside this set, so in NS. Our inference depends on the data, the rc-$A_{(n_j)}$ assumption for group j ($j = 1, 2, 3, 4$) and stopping time T_0. To begin, we compute the lower and upper probabilities from (7) and (8) for all possible subsets that contain only two groups, so S equal to $\{1,2\}$, $\{1,3\}$, $\{1,4\}$, $\{2,3\}$, $\{2,4\}$ and $\{3,4\}$, the results for several ranges of values of T_0 are presented in Table 6.

For example, at $T_0 = 4.50$ we observe the first failure for group 2. Here all lower probabilities are equal to zero since there exists at least one $j \notin S$ for which $r_j = 0$, which means that we cannot exclude the possibility that we will never observe any failure from this group. However, the upper probabilities for these events are not all the same. For example, the upper probability for S equal to $\{1,3\}$ is equal to one, since in this situation we have not observed any failure from any group in S, so there is a possibility that we would never observe any failure from any group in S, therefore this upper probability is equal to one. However, for all S including group 2, the upper probabilities are less than one), since in this case a failure from a group belonging to S has occurred which is indication against our event of interest.

We can also study the behaviour of the lower and upper probabilities when a failure from any group occurs. To this end, we will assume that we terminate the experiment at $T_0 = 7$. In this situation we have observed 8, 9, 2 and 1 failures from group 1, 2, 3 and 4, respectively. Suppose we

T_0	r_1	r_2	r_3	r_4	$\underline{P}^{\{1,2\}}$	$\overline{P}^{\{1,2\}}$	$\underline{P}^{\{1,3\}}$	$\overline{P}^{\{1,3\}}$	$\underline{P}^{\{1,4\}}$	$\overline{P}^{\{1,4\}}$
$[4.50, 4.71)$	0	1	0	0	0	0.948	0	1	0	1
$[4.71, 4.86)$	0	1	0	1	0	0.948	0.012	1	0	0.752
$[4.86, 5.01)$	0	2	0	1	0	0.897	0.024	1	0	0.752
$[5.60, 5.78)$	3	6	0	1	0	0.599	0.066	0.873	0	0.646
$[6.72, 6.84)$	6	9	0	1	0	0.395	0.093	0.762	0	0.542
$[7.00, 7.21)$	8	9	2	1	0.010	0.349	0.093	0.635	0.020	0.477
$[7.21, 7.22)$	9	9	2	1	0.010	0.327	0.093	0.604	0.020	0.445
$[7.47, 7.55)$	9	13	6	1	0.025	0.233	0.108	0.499	0.071	0.445
$[7.55, 7.58)$	9	13	7	1	0.028	0.233	0.108	0.476	0.083	0.445
$[8.04, 8.08)$	10	14	12	1	0.036	0.202	0.111	0.359	0.143	0.423
$[8.08, 8.12)$	10	14	13	1	0.038	0.202	0.111	0.339	0.154	0.423
$[8.42, 8.52)$	14	16	14	2	0.059	0.174	0.117	0.308	0.170	0.362
$[8.52, 8.62)$	14	17	14	2	0.059	0.168	0.117	0.308	0.174	0.362
$[8.66, 8.81)$	16	17	14	2	0.059	0.166	0.117	0.305	0.174	0.357
$[8.81, 8.91)$	16	18	14	2	0.059	0.162	0.117	0.305	0.176	0.357
$[8.91, 8.94)$	17	18	14	2	0.059	0.161	0.117	0.303	0.176	0.355
$[8.94, 8.97)$	18	18	14	2	0.059	0.160	0.117	0.302	0.176	0.354
$[8.97, 9.03)$	18	18	15	2	0.059	0.160	0.117	0.299	0.177	0.354
$[9.03, 9.05)$	18	18	15	3	0.059	0.160	0.117	0.299	0.177	0.354
$[9.05, 9.16)$	19	18	15	3	0.059	0.160	0.117	0.299	0.177	0.354
$[9.16, \infty)$	20	18	15	3	0.059	0.160	0.117	0.299	0.177	0.354

T_0	r_1	r_2	r_3	r_4	$\underline{P}^{\{2,3\}}$	$\overline{P}^{\{2,3\}}$	$\underline{P}^{\{2,4\}}$	$\overline{P}^{\{2,4\}}$	$\underline{P}^{\{3,4\}}$	$\overline{P}^{\{3,4\}}$
$[4.50, 4.71)$	0	1	0	0	0	0.948	0	0.948	0	1
$[4.71, 4.86)$	0	1	0	1	0	0.948	0	0.711	0	0.751
$[4.86, 5.01)$	0	2	0	1	0	0.897	0	0.672	0	0.751
$[5.60, 5.78)$	3	6	0	1	0.026	0.701	0	0.516	0.021	0.751
$[6.72, 6.84)$	6	9	0	1	0.045	0.565	0	0.399	0.063	0.751
$[7.00, 7.21)$	8	9	2	1	0.054	0.510	0.011	0.399	0.082	0.674
$[7.21, 7.22)$	9	9	2	1	0.058	0.510	0.013	0.399	0.091	0.674
$[7.47, 7.55)$	9	13	6	1	0.058	0.309	0.038	0.274	0.116	0.533
$[7.55, 7.58)$	9	13	7	1	0.058	0.294	0.042	0.274	0.116	0.503
$[8.04, 8.08)$	10	14	12	1	0.060	0.212	0.062	0.248	0.129	0.363
$[8.08, 8.12)$	10	14	13	1	0.060	0.200	0.065	0.248	0.129	0.336
$[8.42, 8.52)$	14	16	14	2	0.063	0.182	0.079	0.200	0.134	0.293
$[8.52, 8.62)$	14	17	14	2	0.063	0.179	0.079	0.191	0.135	0.293
$[8.66, 8.81)$	16	17	14	2	0.063	0.179	0.080	0.191	0.136	0.293
$[8.81, 8.91)$	16	18	14	2	0.063	0.176	0.080	0.185	0.137	0.293
$[8.91, 8.94)$	17	18	14	2	0.063	0.176	0.080	0.185	0.137	0.293
$[8.94, 8.97)$	18	18	14	2	0.063	0.176	0.080	0.185	0.138	0.293
$[8.97, 9.03)$	18	18	15	2	0.063	0.175	0.080	0.185	0.138	0.290
$[9.03, 9.05)$	18	18	15	3	0.063	0.175	0.080	0.183	0.138	0.288
$[9.05, 9.16)$	19	18	15	3	0.063	0.175	0.080	0.183	0.138	0.288
$[9.16, \infty)$	20	18	15	3	0.063	0.175	0.080	0.183	0.138	0.288

Table 6: The subset of best groups: lower and upper probabilities (Example 3)

are interested in the subset $S = \{1,2\}$. In this case, $\underline{P}^{\{1,2\}} = 0.010$ and $\overline{P}^{\{1,2\}} = 0.349$. Suppose now that the stopping time is increased from 7 to 7.21. In this case a failure occurs from the first group at time 7.21. We see that, while the lower probability remains constant, the upper probability decreases from 0.349 to 0.327, which illustrates Theorem 4(i). However, when increasing the stopping time from 8.04 to 8.08, so that an extra failure of group 3 occurs, the upper probability that $S = \{1,2\}$ is the subset with the best groups remains constant, but the lower probability increases from 0.036 to 0.038, which illustrates Theorem 4(ii).

Suppose now that we stop the experiment at $T_0 = 8.97$ and we are interested in $S = \{2,3\}$. Then we have observed all units from all groups in S, but there are 3 units that still have not failed in groups in NS. For $T_0 \geq 8.97$, the lower and upper probabilities that $S = \{2,3\}$ is the subset with the two best groups will not change ($\underline{P}^{\{2,3\}} = 0.063$ and $\overline{P}^{\{2,3\}} = 0.175$), which illustrates the special case of Theorem 4(ii). If we change attention to $S = \{1,4\}$, also for $T_0 \geq 8.97$, then the lower and upper probabilities again remain constant ($\underline{P}^{\{1,4\}} = 0.177$ and $\overline{P}^{\{1,4\}} = 0.354$) since we have observed all units from NS, which illustrates the special case of Theorem 4(i).

To carry out the comparison to select the best set, we notice that if we terminate the experiment at $T_0 = 8.52$, $\overline{P}^{\{1,2\}} < \underline{P}^{\{1,4\}}$ which indicates that we can exclude $\{1,2\}$ from being the best. In addition, at $T_0 = 8.97$ we can exclude the set $\{2,3\}$ from being the subset with the best groups as $\overline{P}^{\{2,3\}} < \underline{P}^{\{1,4\}}$. This may be due to the fact that the second group is included in these sets, since the second group was the worse group as found in Example 1. However, this does not hold for $\{2,4\}$ as $\overline{P}^{\{2,4\}} \not> \underline{P}^{\{1,4\}}$. This happens because this set consists of the best and the worse group (see the results of Example 1). So, we only have strong indication that $\{1,2\}$ and $\{2,3\}$ are not the best subsets. As $\underline{P}^{\{2,4\}} < \underline{P}^{\{1,3\}} < \underline{P}^{\{3,4\}} < \underline{P}^{\{1,4\}}$ and $\overline{P}^{\{2,4\}} < \overline{P}^{\{3,4\}} < \overline{P}^{\{1,3\}} < \overline{P}^{\{1,4\}}$ there is a weak indication that $S = \{1,4\}$ is the best subset.

6. Selecting the subset including the best group

In this section we consider a similar scenario as in Section 5, with the experiment terminated at time T_0 but now our objective is to select a subset of groups such that the group that provides the largest future lifetime is included in this subset. Again, let $S = \{l_1, l_2, ..., l_m\}$ be a selected subset of m groups ($1 \leq m \leq k-1$) from k independent groups and let NS be the complementary set of S which contains the $k-m$ nonselected groups. We are interested in the lower and upper probabilities for the event that the next observation from at least one of the selected groups in S is greater than the next observation from each group in NS. These lower and upper probabilities are denoted by

$$\underline{P}^{\tilde{S}} = \underline{P}\left(\max_{l \in S} X_{l,n_l+1} > \max_{j \in NS} X_{j,n_j+1}\right) \text{ and } \overline{P}^{\tilde{S}} = \overline{P}\left(\max_{l \in S} X_{l,n_l+1} > \max_{j \in NS} X_{j,n_j+1}\right)$$

These lower and upper probabilities are given in Theorem 5, using the notation (6) as before.

Theorem 5. *The lower and upper probabilities for the event that the next observation of at least one group in S is greater than the next observation of each group in NS are given by*

$$\underline{P}^{\tilde{S}} = \sum_{\substack{i_l=1 \\ l \in S}}^{r_l+2} \prod_{j \in NS} \left[\frac{\sum_{i_j=1}^{r_j} 1\{x_{j,i_j} < \max_{l \in S}\{L(I_{i_l}^l)\}\}}{n_j+1}\right] \cdot \prod_{l \in S} M_{X_{l,n_l+1}}(I_{i_l}^l) \quad (9)$$

$$\overline{P}^{\tilde{S}} = \sum_{\substack{i_l=1 \\ l \in S}}^{r_l+2} \prod_{j \in NS} \left[\frac{1 + \sum_{i_j=1}^{r_j} 1\{x_{j,i_j} < \max_{l \in S}\{U(I_{i_l}^l)\}\}}{n_j+1} + \frac{(n_j - r_j) 1\{T_0 < \max_{l \in S}\{U(I_{i_l}^l)\}\}}{n_j+1}\right] \cdot \prod_{l \in S} M_{X_{l,n_l+1}}(I_{i_l}^l) \quad (10)$$

Proof. The proof is similar to the proof of Theorem 3, given in the appendix, but with 'min' replaced by 'max'.

Below we present some properties and special cases of the lower and upper probabilities (9) and (10).

1. If $r_l = 0$ for at least one $l \in S$ and $r_j \geq 0$, for all $j \in NS$, then the upper probability $\overline{P}^{\tilde{S}} = 1$, since we have not seen any failure from at least one group in S. This means that we cannot exclude the situation

that we will never see a failure from such a group in S. Further, if $r_l = 0$ for all $l \in S$, then the lower probability is

$$\underline{P}^{\tilde{S}} = \prod_{j \in NS} \frac{r_j}{n_j + 1} \left(1 - \prod_{l \in S} \frac{1}{n_l + 1}\right)$$

this lower probability is equal to zero if there exists at least one $j \in NS$ for which $r_j = 0$.

2. If $r_j = 0$ for at least one $j \in NS$ and $r_l > 0$, for all $l \in S$, then the lower probability $\underline{P}^{\tilde{S}} = 0$ since there exists a group in NS for which we have not seen any failure. This means that we cannot exclude the situation that we will never see a failure from this group. Further, if $r_j = 0$ for all $j \in NS$, then the upper probability is

$$\overline{P}^{\tilde{S}} = 1 - \prod_{l \in S} \frac{r_l}{n_l + 1} \left[1 - \prod_{j \in NS} \frac{1}{n_j + 1}\right]$$

so if $r_l = 0$ for at least one $l \in S$, then $\overline{P}^{\tilde{S}} = 1$.

Now, we study the effect upon the lower and upper probabilities when the stopping time is increased from T_0 to $T_0 + \epsilon$, for small $\epsilon > 0$ just to allow one failure from one group to occur.

Theorem 6. *(i) If a failure from group $l^* \in S$ occurs in the interval $(T_0, T_0 + \epsilon)$, then the lower probability $\underline{P}^{\tilde{S}}$ remains constant, and the upper probability $\overline{P}^{\tilde{S}}$ decreases by*

$$\frac{1}{n_{l^*} + 1} \prod_{l \in S \setminus \{l^*\}} \frac{r_l}{n_l + 1} \left(1 - \prod_{j \in NS} \frac{r_j + 1}{n_j + 1}\right)$$

except when $r_j = n_j$, for all $j \in NS$ or when there exists a $l \in S \setminus \{l^*\}$ for which $r_l = 0$, in which cases the upper probability remains constant.

(ii) If a failure from group $j^* \in NS$ occurs in the interval $(T_0, T_0 + \epsilon)$, then the upper probability $\overline{P}^{\tilde{S}}$ remains constant, and the lower probability $\underline{P}^{\tilde{S}}$

increases by

$$\frac{1}{n_{j^*}+1}\prod_{j\in NS\setminus\{j^*\}}\frac{r_j}{n_j+1}\left(1-\prod_{l\in S}\frac{r_l+1}{n_l+1}\right)$$

except when $r_l = n_l$ for all $l \in S$ or when there exists a $j \in NS \setminus \{j^*\}$ for which $r_j = 0$, in which cases the lower probability remains constant.

Proof. For case i (ii), replace r_l (r_j^*) by $r_l + 1$ ($r_j^* + 1$) in the formula (9) and (10), then this follows by basic analysis of the lower and upper probabilities of Theorem 5.

It can easily be shown that the NPI lower and upper probabilities for selecting the subset of best groups, given by (7) and (8), cannot exceed those for selecting the subset including the best group, given by (9) and (10). This follows from $1\{x_{j,i_j} < \min_{l\in S}\{\bullet\}\} \leq 1\{x_{j,i_j} < \max_{l\in S}\{\bullet\}\}$ and $1\{T_0 < \min_{l\in S}\{\bullet\}\} \leq 1\{T_0 < \max_{l\in S}\{\bullet\}\}$, where '$\bullet$' refers to $L(I^l_{i_l})$ or $U(I^l_{i_l})$.

Example 4. Consider again the data set from Example 1, which we also used in Example 3. The lower and upper probabilities for the event that the lifetime of the next observation from at least one group in S is greater than the lifetime of the next observation of each group in NS, are calculated from (9) and (10) at different stopping times T_0 for all possible subsets containing 2 groups and are presented in Table 7.

At $T_0 = 4.5$, which is the moment when we observe the first failure (group 2), all lower probabilities are zero and all upper probabilities are one, which is different from the case when we select the subset of 2 best groups (Example 3), since for that case there are some upper probabilities which are less than one. This is because at $T_0 = 4.5$, whichever subset of 2 groups we consider, this subset will always contain at least one group for which we have not seen any failure, so there is no evidence against the probability that this subset can still contain the best group.

For example, for $S = \{1, 3\}$, the lower probability at $T_0 = 4.71$ is 0.013, while the corresponding upper probability is one. At $T_0 = 4.71$ we have seen failures from groups 2 and 4. Therefore, we cannot exclude the possibility that we will not observe any failure from any group in NS. In fact, the upper probabilities for the sets that contain group 3, i.e. $\{1, 3\}$, $\{2, 3\}$ and $\{3, 4\}$, will be one until T_0 is equal to the time at which we observe the first failure

T_0	r_1	r_2	r_3	r_4	$\underline{P}^{\{\widetilde{1,2}\}}$	$\overline{P}^{\{\widetilde{1,2}\}}$	$\underline{P}^{\{\widetilde{1,3}\}}$	$\overline{P}^{\{\widetilde{1,3}\}}$	$\underline{P}^{\{\widetilde{1,4}\}}$	$\overline{P}^{\{\widetilde{1,4}\}}$
$[4.50, 4.71)$	0	1	0	0	0	1	0	1	0	1
$[4.71, 4.86)$	0	1	0	1	0	1	0.013	1	0	1
$[4.86, 5.01)$	0	2	0	1	0	1	0.026	1	0	1
$[5.60, 5.78)$	3	6	0	1	0	0.956	0.078	1	0	0.965
$[6.72, 6.84)$	6	9	0	1	0	0.869	0.117	1	0	0.930
$[6.84, 6.91)$	6	9	1	1	0.013	0.869	0.117	0.987	0.025	0.930
$[7.00, 7.21)$	8	9	2	1	0.026	0.828	0.117	0.965	0.049	0.909
$[7.21, 7.22)$	9	9	2	1	0.026	0.808	0.117	0.961	0.049	0.898
$[7.47, 7.55)$	9	13	6	1	0.073	0.737	0.159	0.882	0.200	0.898
$[7.55, 7.58)$	9	13	7	1	0.083	0.737	0.159	0.865	0.232	0.898
$[8.04, 8.08)$	10	14	12	1	0.129	0.695	0.168	0.760	0.419	0.890
$[8.08, 8.12)$	10	14	13	1	0.139	0.695	0.168	0.742	0.453	0.890
$[8.42, 8.52)$	14	16	14	2	0.267	0.624	0.271	0.648	0.529	0.834
$[8.52, 8.62)$	14	17	14	2	0.267	0.613	0.279	0.648	0.550	0.834
$[8.66, 8.81)$	16	17	14	2	0.267	0.588	0.279	0.624	0.550	0.829
$[8.81, 8.91)$	16	18	14	2	0.267	0.576	0.286	0.624	0.568	0.829
$[8.91, 8.94)$	17	18	14	2	0.267	0.563	0.286	0.613	0.568	0.827
$[8.94, 8.97)$	18	18	14	2	0.267	0.549	0.286	0.603	0.568	0.826
$[8.97, 9.03)$	18	18	15	2	0.270	0.549	0.286	0.589	0.587	0.826
$[9.03, 9.05)$	18	18	15	3	0.293	0.549	0.308	0.589	0.587	0.826
$[9.05, 9.16)$	19	18	15	3	0.293	0.549	0.308	0.589	0.587	0.826
$[9.16, \infty)$	20	18	15	3	0.293	0.549	0.308	0.589	0.587	0.826

T_0	r_1	r_2	r_3	r_4	$\underline{P}^{\{\widetilde{2,3}\}}$	$\overline{P}^{\{\widetilde{2,3}\}}$	$\underline{P}^{\{\widetilde{2,4}\}}$	$\overline{P}^{\{\widetilde{2,4}\}}$	$\underline{P}^{\{\widetilde{3,4}\}}$	$\overline{P}^{\{\widetilde{3,4}\}}$
$[4.50, 4.71)$	0	1	0	0	0	1	0	1	0	1
$[4.71, 4.86)$	0	1	0	1	0	1	0	0.987	0	1
$[4.86, 5.01)$	0	2	0	1	0	1	0	0.974	0	1
$[5.60, 5.78)$	3	6	0	1	0.035	1	0	0.922	0.044	1
$[6.72, 6.84)$	6	9	0	1	0.070	1	0	0.883	0.131	1
$[6.84, 6.91)$	6	9	1	1	0.070	0.975	0.013	0.883	0.131	0.987
$[7.00, 7.21)$	8	9	2	1	0.091	0.951	0.035	0.883	0.172	0.974
$[7.21, 7.22)$	9	9	2	1	0.102	0.951	0.040	0.883	0.192	0.974
$[7.47, 7.55)$	9	13	6	1	0.102	0.800	0.118	0.841	0.263	0.927
$[7.55, 7.58)$	9	13	7	1	0.102	0.768	0.135	0.841	0.263	0.917
$[8.04, 8.08)$	10	14	12	1	0.110	0.581	0.240	0.832	0.305	0.871
$[8.08, 8.12)$	10	14	13	1	0.110	0.547	0.258	0.832	0.305	0.861
$[8.42, 8.52)$	14	16	14	2	0.166	0.471	0.352	0.729	0.376	0.733
$[8.52, 8.62)$	14	17	14	2	0.166	0.450	0.352	0.721	0.387	0.733
$[8.66, 8.81)$	16	17	14	2	0.171	0.450	0.377	0.721	0.412	0.733
$[8.81, 8.91)$	16	18	14	2	0.171	0.432	0.377	0.714	0.424	0.733
$[8.91, 8.94)$	17	18	14	2	0.173	0.432	0.387	0.714	0.437	0.733
$[8.94, 8.97)$	18	18	14	2	0.174	0.432	0.397	0.714	0.451	0.733
$[8.97, 9.03)$	18	18	15	2	0.174	0.413	0.411	0.714	0.451	0.730
$[9.03, 9.05)$	18	18	15	3	0.174	0.413	0.411	0.692	0.451	0.708
$[9.05, 9.16)$	19	18	15	3	0.174	0.413	0.411	0.692	0.451	0.708
$[9.16, \infty)$	20	18	15	3	0.174	0.413	0.411	0.692	0.451	0.708

Table 7: The set including the best group: lower and upper probabilities (Example 4)

from group 3 at $T_0 = 6.84$. The lower probabilities for the sets that do not include the third group, i.e. $\{1, 2\}$, $\{1, 4\}$ and $\{2, 4\}$, are zero until T_0 is equal to the time at which we observe the first failure from the third group.

To study the behaviour of these lower and upper probabilities, let us consider the situation when the stopping time T_0 is increased from 7 to 7.21, and let $S = \{1, 2\}$ be the set of interest. At time 7.21 a failure of group 1 is observed. Here, the lower probability remains constant as the amount of information in favour of our event remains the same. However, the upper probability is decreasing from 0.828 to 0.808 as the amount of information against our event has increased (Theorem 6(i)). When we change interest, from $S = \{1, 2\}$ to $S = \{2, 3\}$, when T_0 increases from 7 to 7.21, the upper probability remains constant, but, the lower probability increases from 0.091 to 0.102, as the amount of information in favour of this event now increases (Theorem 6(ii)).

At $T_0 = 8.97$ we have observed failures of all units from groups 2 and 3. If the set of interest is $S = \{2, 3\}$, we see that the lower and upper probabilities remain constant for $T_0 \geq 8.97$ since we have observed all units of all groups in S (special case Theorem 6(ii)). Also, if the set of interest is $S = \{1, 4\}$, the lower and upper probabilities remain constant since we have observed all units from all groups in NS (special case Theorem 6(i)).

At the time when we have observed all units from all groups, i.e. $T_0 = 9.16$, we can conclude that the set $\{1, 4\}$ is better than $\{1, 2\}$ and $\{2, 3\}$, in the sense that it is more likely that $\{1, 4\}$ contains the best group since

$$0.549 = \overline{P}^{\widetilde{\{1,2\}}} < \underline{P}^{\widetilde{\{1,4\}}} = 0.587 \text{ and } 0.413 = \overline{P}^{\widetilde{\{2,3\}}} < \underline{P}^{\widetilde{\{1,4\}}} = 0.587$$

In fact the set $\{1, 4\}$ has the highest lower and upper probability when the failure time have been observed for all units from all groups. However, we have a weak indication that $\{1, 4\}$ is the set which contains the best group since we have

$$\underline{P}^{\widetilde{\{1,3\}}} < \underline{P}^{\widetilde{\{2,4\}}} < \underline{P}^{\widetilde{\{3,4\}}} < \underline{P}^{\widetilde{\{1,4\}}} < \overline{P}^{\widetilde{\{1,3\}}} < \overline{P}^{\widetilde{\{2,4\}}} < \overline{P}^{\widetilde{\{3,4\}}} < \overline{P}^{\widetilde{\{1,4\}}}$$

However, we can exclude the set $\{2, 3\}$ from the comparison early from $T_0 = 8.42$ onwards since $\overline{P}^{\widetilde{\{2,3\}}} < \underline{P}^{\widetilde{\{1,4\}}}$. Also the set $\{1, 2\}$ can be excluded from $T_0 = 8.91$ onwards since $\overline{P}^{\widetilde{\{1,2\}}} < \underline{P}^{\widetilde{\{1,4\}}}$.

Example 5. In this example all NPI procedures that have been introduced in this paper are illustrated, so we consider selecting the best group (Section 4), selecting the subset of best groups (Section 5) and selecting the subset including the best group (Section 6). We consider subsets that contain three groups as well as subsets that contain one or two groups. Due to space limitations we summarize the results, as we will see later, such as the stopping times at which we exclude a subset or a group from the comparison are reported.

Group	breakdown times										
1	7.74	17.05	20.46	21.02	22.66	43.40	47.30	139.07	144.12	175.88	194.90
2	0.27	0.40	0.69	0.79	2.75	3.91	9.88	13.95	15.93	27.80	53.24
	82.85	89.29	100.58	215.10							
3	0.19	0.78	0.96	1.31	2.78	3.16	4.15	4.67	4.85	6.50	7.35
	8.01	8.27	12.06	31.75	32.52	33.91	36.71	72.89			
4	0.35	0.59	0.96	0.99	1.69	1.97	2.07	2.58	2.71	2.90	3.67
	3.99	5.35	13.77	25.50							
5	0.09	0.39	0.47	0.73	0.74	1.13	1.40	2.38			

Table 8: The times to breakdown (in minutes) at five voltage levels

We use a data set also used by Lawless (2003, p.3), which consists of the times to breakdown (in minutes) of electrical insulating fluids at seven voltage levels. These data were originally studied by Nelson (1972), who particularly studied the accelerated life testing nature of the data. We do not attempt to model the explicit effect of accelerated life testing but just use these data to illustrate the NPI methods presented in this paper. We will use only five out of these seven voltage levels in this example, see Table 8, to illustrate our method. More precisely, we exclude the first two voltage levels from the original data set since they contain relatively few units compared with the other voltage levels. Again X_{j,i_j} represents the time to breakdown for unit i_j at voltage level j which we refer to as group j, $i_j = 1, \ldots, n_j$, where $j = 1, \ldots, 5$ represents voltage level 30, 32, 34, 36 and 38, respectively. The corresponding sample sizes are 11, 15, 19, 15 and 8, respectively. In this data set, the range of times to breakdown vary from 0.09 at voltage level 5 to 215.10 at voltage level 2.

Let us consider selection of a subset of 3 groups, so we have 10 different possible subsets: $\{1,2,3\}$, $\{1,2,4\}$, $\{1,2,5\}$, $\{1,3,4\}$, $\{1,3,5\}$, $\{1,4,5\}$, $\{2,3,4\}$, $\{2,3,5\}$, $\{2,4,5\}$ and $\{3,4,5\}$. Of course, here are also 10 possible subsets contains 2 different voltage levels: $\{1,2\}$, $\{1,3\}$, $\{1,4\}$, $\{1,5\}$, $\{2,3\}$,

$\{2,4\}$, $\{2,5\}$, $\{3,4\}$, $\{3,5\}$ and $\{4,5\}$.

Let us consider selection of the subset of 3 best groups. When we have observed all units from all groups we will select the subset $\{1,2,3\}$ as the subset of best groups with lower and upper probabilities $\underline{P}^{\{1,2,3\}} = 0.337$ and $\overline{P}^{\{1,2,3\}} = 0.535$, respectively. Actually we can conclude the same result (i.e. $\{1,2,3\}$ as the best subset) at an early stage. In fact at $T_0 = 6.5$ we will select the set $\{1,2,3\}$ as the best subset among all ten subsets, see Table 9. Note that, at this point we still have not observed any failure from the first group while we have observed already all units from group 5. Table 9 explains how we can establish an early decision from the beginning. For example, at $T_0 = 2.38$, we have observed all breakdown times for units from group 5 and we have not observed any breakdowns from the first group, and therefore we will exclude any set that contains group 5 from being the best. Moreover, at $T_0 = 2.90$ we can exclude 7 of the 10 subsets from comparison of becoming the best subset. At this time, the subsets $\{1,2,3\}$, $\{1,2,4\}$ and $\{1,3,4\}$ remain in the comparison process. Figure 1 shows a pairwise comparison between the best subset $\{1,2,3\}$ and the two second best subsets $\{1,2,4\}$ and $\{1,3,4\}$, for different T_0. We can see that at $T_0 = 3.99$, $\overline{P}^{\{1,2,4\}} < \underline{P}^{\{1,2,3\}}$ and at $T_0 = 6.50$, $\overline{P}^{\{1,3,4\}} < \underline{P}^{\{1,2,3\}}$.

T_0	r_1	r_2	r_3	r_4	r_5	Set(s) out	Pairwise comparison with respect to $\{1,2,3\}$
1.69	0	4	4	5	7	$\{2,4,5\}$	$0.123 = \overline{P}^{\{2,4,5\}} < \underline{P}^{\{1,2,3\}} = 0.131$
1.97	0	4	4	6	7	$\{3,4,5\}$	$0.126 = \overline{P}^{\{3,4,5\}} < \underline{P}^{\{1,2,3\}} = 0.154$
						$\{2,3,5\}$	$0.145 = \overline{P}^{\{2,3,5\}} < \underline{P}^{\{1,2,3\}} = 0.154$
2.07	0	4	4	7	7	$\{1,4,5\}$	$0.153 = \overline{P}^{\{1,4,5\}} < \underline{P}^{\{1,2,3\}} = 0.177$
2.38	0	4	4	7	8	$\{1,2,5\}$	$0.114 = \overline{P}^{\{1,2,5\}} < \underline{P}^{\{1,2,3\}} = 0.200$
						$\{1,3,5\}$	$0.139 = \overline{P}^{\{1,3,5\}} < \underline{P}^{\{1,2,3\}} = 0.200$
2.90	0	5	5	10	8	$\{2,3,4\}$	$0.235 = \overline{P}^{\{2,3,4\}} < \underline{P}^{\{1,2,3\}} = 0.275$
3.99	0	6	6	12	8	$\{1,2,4\}$	$0.292 = \overline{P}^{\{1,2,4\}} < \underline{P}^{\{1,2,3\}} = 0.314$
6.50	0	6	10	13	8	$\{1,3,4\}$	$0.326 = \overline{P}^{\{1,3,4\}} < \underline{P}^{\{1,2,3\}} = 0.328$

Table 9: The subset of best groups: pairwise comparison with respect to $\{1,2,3\}$

Consequently, if we terminate the experiment at $T_0 = 6.5$ we get the same decision as when we would have observed all units from all groups. By doing that, beside a much shorter testing time, we can keep 9 units out of 15 from group 2, 9 out of 19 from group 3, 2 out of 15 from group 4 and all units from group 1 to be possible used for other purposes.

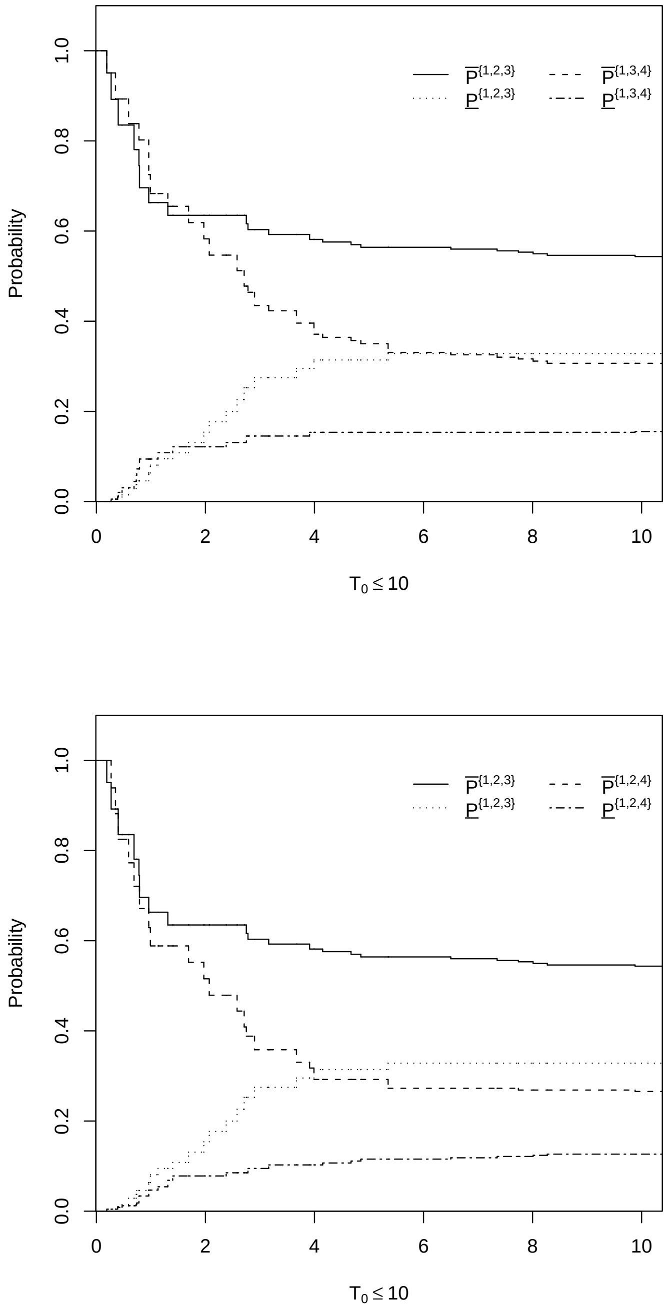

Figure 1: The subset of best groups: lower and upper probabilities for $T_0 \leq 10$

Now we consider the case of selecting the subset of 3 groups that includes the best group. Table 10 shows that 4 out of 10 subsets could be excluded from the comparison just at $T_0 = 13.77$, and a fifth subsets at $T_0 = 25.5$. Unlike the case of selecting the subset of best groups, there are three sets $\{1,2,3\}$, $\{1,2,4\}$ and $\{1,2,5\}$ which cannot be excluded from any time onwards. Consequently, we do not have enough indication to select one of these three sets as the set that is most likely to include the best group as even when we have observed all units, we have $\overline{P}^{\widetilde{\{1,2,4\}}} \not< \underline{P}^{\widetilde{\{1,2,3\}}}$ and $\overline{P}^{\widetilde{\{1,2,5\}}} \not< \underline{P}^{\widetilde{\{1,2,3\}}}$.

In Table 11, the stopping times at which we exclude a group from being the best group are reported, where the lower and upper probabilities are calculated from (4) and (5). As we can see from Table 11, we can exclude group 5 from being the best group already at $T_0 = 6.5$, as which time $0.112 = \overline{P}^{(5)} < \underline{P}^{(1)} = 0.124$. Group 4 and group 3 can be excluded at $T_0 = 12.06$ and $T_0 = 31.75$, respectively, as then we have $0.193 = \overline{P}^{(4)} < \underline{P}^{(1)} = 0.197$ and $0.288 = \overline{P}^{(3)} < \underline{P}^{(1)} = 0.310$. In addition, when we have observed breakdown times of all units from all groups (or even before, i.e. at $T_0 = 82.85$) we can conclude that the first group is the best since $\underline{P}^{(1)} > \overline{P}^{(l)}$ for $l = 2,3,4,5$. At $T_0 = 82.85$ we can exclude the second group from being the best group (where $0.378 = \overline{P}^{(2)} < \underline{P}^{(1)} = 0.391$) which may explain the situation of being $\{1,2,3\}$, $\{1,2,4\}$ and $\{1,2,5\}$ to be the subset that contains the best group since these contain the best group and the second best group (i.e. group 2).

Suppose, for example, that we terminate the experiment at $T_0 = 25.5$. From Table 10 it follows that in this case we can exclude 5 of the 10 subsets from being the subset that includes the best group, therefore, the sets $\{1,2,3\}$, $\{1,2,4\}$, $\{1,2,5\}$, $\{1,3,5\}$ and $\{1,3,4\}$ will remain under comparison. At the same time (at $T_0 = 25.5$), see Table 11, we can exclude groups 4 and 5 from being the best group, therefore we still have groups 1, 2 and 3 under comparison. This explains why any set that contains two of these groups is still under consideration for being the set that includes the best group. However, this is not the case for $\{2,3,5\}$ and $\{2,3,4\}$ since the first group (the best) is not included in both of them.

Now let us consider the case of selecting the subset of 2 best groups. From $T_0 = 8.27$ onwards, see Table 12, all subsets except $\{1,2\}$ and $\{1,3\}$ are excluded from being the subset of 2 best groups. Here we have strong indication to exclude these subsets since their corresponding upper probabil-

T_0	r_1	r_2	r_3	r_4	r_5	Set(s) out	Pairwise comparison with respect to $\{1,2,3\}$
5.35	0	6	9	13	8	$\{3,4,5\}$	$0.685 = \overline{P}^{\,\widetilde{\{3,4,5\}}} < \underline{P}^{\,\widetilde{\{1,2,3\}}} = 0.717$
9.88	1	7	13	13	8	$\{2,4,5\}$	$0.696 = \overline{P}^{\,\widetilde{\{2,4,5\}}} < \underline{P}^{\,\widetilde{\{1,2,3\}}} = 0.717$
13.77	1	7	14	14	8	$\{2,3,5\}$	$0.751 = \overline{P}^{\,\widetilde{\{2,3,5\}}} < \underline{P}^{\,\widetilde{\{1,2,3\}}} = 0.769$
						$\{2,3,4\}$	$0.762 = \overline{P}^{\,\widetilde{\{2,3,4\}}} < \underline{P}^{\,\widetilde{\{1,2,3\}}} = 0.769$
25.50	5	9	14	15	8	$\{1,4,5\}$	$0.803 = \overline{P}^{\,\widetilde{\{1,4,5\}}} < \underline{P}^{\,\widetilde{\{1,2,3\}}} = 0.811$
72.89	7	11	19	15	8	$\{1,3,5\}$	$0.811 = \overline{P}^{\,\widetilde{\{1,3,5\}}} < \underline{P}^{\,\widetilde{\{1,2,3\}}} = 0.811$
139.07	8	14	19	15	8	$\{1,3,4\}$	$0.809 = \overline{P}^{\,\widetilde{\{1,3,4\}}} < \underline{P}^{\,\widetilde{\{1,2,3\}}} = 0.811$

Table 10: The subset including the best group: pairwise comparison with respect to $\{1,2,3\}$

T_0	r_1	r_2	r_3	r_4	r_5	group out	Pairwise comparison with respect to group 1
6.50	0	6	10	13	8	group 5	$0.112 = \overline{P}^{\,(5)} < \underline{P}^{\,(1)} = 0.124$
12.06	1	7	14	13	8	group 4	$0.193 = \overline{P}^{\,(4)} < \underline{P}^{\,(1)} = 0.197$
31.75	5	10	15	15	8	group 3	$0.288 = \overline{P}^{\,(3)} < \underline{P}^{\,(1)} = 0.310$
82.85	7	12	19	15	8	group 2	$0.378 = \overline{P}^{\,(2)} < \underline{P}^{\,(1)} = 0.391$

Table 11: The best group: pairwise comparison with respect to group 1

ities are less than $\underline{P}^{\,\{1,2\}}$, but there is only a weak indication that $\{1,2\}$ is the best subset of 2 best groups since $\underline{P}^{\,\{1,3\}} < \underline{P}^{\,\{1,2\}} < \overline{P}^{\,\{1,3\}} < \overline{P}^{\,\{1,2\}}$.

From the one-group comparison, see Table 11, at $T_0 = 8.27$ the fifth group can be excluded (in fact this can be concluded already for $T_0 \geq 6.5$) as $\overline{P}^{\,(5)} < \underline{P}^{\,(1)}$. In addition, we exclude the fourth and the third group at $T_0 = 12.06$ and $T_0 = 31.75$, respectively. That may explain why, when considering subsets consisting of two groups (Table 12) the subsets that contain the fifth group are excluded from early time (at $T_0 = 4.15$). However, by the end of the experiment we do not have strong indication to choose between the subsets $\{1,2\}$ and $\{1,3\}$ for selecting the subset of best groups.

With regard to selection of the subset of 2 groups that includes the best group, and from $T_0 = 32.52$ onwards (Table 13) the subsets $\{2,3\}$, $\{2,4\}$, $\{2,5\}$, $\{3,4\}$, $\{3,5\}$ and $\{4,5\}$ are excluded from being the subset including the best group. Here we have strong indication that these subsets can be excluded since their corresponding upper probabilities are less than $\underline{P}^{\,\widetilde{\{1,2\}}}$, but there is only a weak indication to select $\{1,2\}$ as the subset including the

T_0	r_1	r_2	r_3	r_4	r_5	Set(s) out	Pairwise comparison with respect to $\{1,2\}$
2.71	0	4	4	9	8	$\{4,5\}$	$0.051 = \overline{P}^{\{4,5\}} < \underline{P}^{\{1,2\}} = 0.064$
2.90	0	5	5	10	8	$\{2,5\}$	$0.079 = \overline{P}^{\{2,5\}} < \underline{P}^{\{1,2\}} = 0.086$
3.16	0	5	6	10	8	$\{3,5\}$	$0.082 = \overline{P}^{\{3,5\}} < \underline{P}^{\{1,2\}} = 0.102$
4.15	0	6	7	12	8	$\{1,5\}$	$0.122 = \overline{P}^{\{1,5\}} < \underline{P}^{\{1,2\}} = 0.137$
4.85	0	6	9	12	8	$\{3,4\}$	$0.155 = \overline{P}^{\{3,4\}} < \underline{P}^{\{1,2\}} = 0.172$
						$\{2,4\}$	$0.168 = \overline{P}^{\{2,4\}} < \underline{P}^{\{1,2\}} = 0.172$
8.27	1	6	13	13	8	$\{1,4\}$	$0.243 = \overline{P}^{\{1,4\}} < \underline{P}^{\{1,2\}} = 0.256$
						$\{2,3\}$	$0.251 = \overline{P}^{\{2,3\}} < \underline{P}^{\{1,2\}} = 0.256$

Table 12: The subset of best groups: pairwise comparison with respect to $\{1,2\}$

best group since $\underline{P}^{\{\widetilde{1,4}\}} < \underline{P}^{\{\widetilde{1,5}\}} < \underline{P}^{\{\widetilde{1,3}\}} < \underline{P}^{\{\widetilde{1,2}\}}$ and $\overline{P}^{\{\widetilde{1,5}\}} < \overline{P}^{\{\widetilde{1,4}\}} < \overline{P}^{\{\widetilde{1,3}\}} < \overline{P}^{\{\widetilde{1,2}\}}$.

T_0	r_1	r_2	r_3	r_4	r_5	Set(s) out	Pairwise comparison with respect to $\{1,2\}$
5.35	0	6	9	13	8	$\{4,5\}$	$0.283 = \overline{P}^{\{\widetilde{4,5}\}} < \underline{P}^{\{\widetilde{1,2}\}} = 0.315$
8.27	1	6	13	13	8	$\{3,5\}$	$0.438 = \overline{P}^{\{\widetilde{3,5}\}} < \underline{P}^{\{\widetilde{1,2}\}} = 0.451$
12.06	1	7	14	13	8	$\{3,4\}$	$0.455 = \overline{P}^{\{\widetilde{3,4}\}} < \underline{P}^{\{\widetilde{1,2}\}} = 0.484$
25.50	5	9	14	15	8	$\{2,4\}$	$0.516 = \overline{P}^{\{\widetilde{2,4}\}} < \underline{P}^{\{\widetilde{1,2}\}} = 0.547$
						$\{2,5\}$	$0.522 = \overline{P}^{\{\widetilde{2,5}\}} < \underline{P}^{\{\widetilde{1,2}\}} = 0.547$
32.52	5	10	16	15	8	$\{2,3\}$	$0.592 = \overline{P}^{\{\widetilde{2,3}\}} < \underline{P}^{\{\widetilde{1,2}\}} = 0.601$

Table 13: The subset including the best group: pairwise comparison with respect to $\{1,2\}$

On the other hand, at $T_0 = 8.27$ we can exclude only the subsets $\{3,5\}$ and $\{4,5\}$ from being the subset including the best group. In the one-group comparison (Table 11), we exclude group 5 from being the best group already at $T_0 = 6.5$. However, while $\{2,5\}$ contains the fifth group, it is still under comparison until $T_0 = 25.5$. With respect to $\{1,5\}$, and by the end of the experiment, there is no strong indication to exclude this subset from being the subset including the best group although this subset includes the best (group 1) and the worse group (group 5). Note that this subset was excluded at $T_0 = 4.15$ from being the subset of 2 best groups (Table 12). Actually we exclude all subsets that contain the fifth group from being the subset of best groups early at $T_0 = 4.15$, which is even before we decided to exclude the fifth group from being the best group (one-group comparison, Table 11) at

$T_0 = 6.5$.

At $T_0 = 6.5$, we see from Table 9 that we can select $\{1, 2, 3\}$ as the subset of best groups containing 3 groups. However, at this time, we see from Table 12 that we do not have enough indication to select one subset for being the subset of best groups among $\{1, 2\}$, $\{1, 3\}$, $\{1, 4\}$ and $\{2, 3\}$, although $\{1, 2\}$, $\{1, 3\}$ and $\{2, 3\}$ are subsets of $\{1, 2, 3\}$, but the subset $\{1, 4\}$ contains only one group from $\{1, 2, 3\}$. Later, at $T_0 = 8.27$ we exclude the subset $\{2, 3\}$ ($\subset \{1, 2, 3\}$) from being the subset of best groups.

At $T_0 = 25.5$, see Table 10, we do not have strong indication to select one of $\{1, 2, 3\}$, $\{1, 2, 4\}$, $\{1, 2, 5\}$, $\{1, 3, 4\}$ and $\{1, 3, 5\}$, from being the subset including the best group. Here all possible 3-group subsets which include the best group 1 are still under comparison except for $\{1, 4, 5\}$, which contains the two worse groups (4 and 5) according to Table 11. However, we exclude the subset $\{1, 4, 5\}$, just at $T_0 = 25.5$ from being the subset including the best group. However at this time (Table 13) we do not have a strong indication in favor of selecting one of $\{1, 2\}$, $\{1, 3\}$,$\{1, 4\}$,$\{1, 5\}$ and $\{2, 3\}$ for being the subset including the best group. At $T_0 = 25.5$, the subset $\{2, 3\}$ is also still under comparison for being the subset including the best group while group 1 (the best group) is not included in this subset. However, at least all subsets that include group 1 (the best group) are still under comparison of being the subset including the best group until the end of the experiment.

7. Concluding remarks

In this paper we have presented NPI for comparison of several groups through experiments that may be terminated before the event of interest has been observed for all units. This work generalizes the NPI approach for selection presented by Coolen and van der Laan (2001), with close links to methods for precedence testing which are explicitly developed to deal with such early termination.

Our method has the advantage that the comparison is not based on testing the hypothesis of equal lifetime distributions, which, although a well established approach in classical statistics, is a somewhat surprising starting point as the reasons for making a comparison of lifetimes of units from different groups may make it very unlikely that units from all groups would actually have identical lifetime distributions. In addition, in both cases of rejection or not of such a hypothesis, it is not clear what such a conclusion implies for the lifetime of a future unit. Application of our method leads to lower and upper probabilities for certain events of interest, which enables conservative decisions by basing these on the worst possible situation for the event of interest.

We note that when the stopping time increases from T_0 to $T_0+\epsilon$ such that only one extra failure occurs from one particular group, the lower probability remains either constant or increases and the upper probability remains constant or decreases. Hence, when T_0 increases to $T_0+\epsilon$, the imprecision remains constant or decreases. It is shown, clearly, that the lower probability for a certain event quantifies the amount of information in favour of the event while the upper probability quantifies the amount of information against the event.

The NPI method presented here is not considered to be a competitor for established classical methods for precedence testing and selection, but it provides an interesting alternative that has strong frequentist properties and which may be suitable particularly in cases where interest is explicitly in a future observation from one or more selected groups. It may well be the case that these different methods lead to quite different conclusions, so care must be taken about the actual inferential conclusions. As always, applying a variety of suitable statistical methods to a practical problem might give valuable insights into the problem and the different methods, where differences typically occur due to different underlying assumptions and explicitly different inferential goals of the methods.

A. Appendix

We prove Theorem 3 in detail. The proof of Theorem 5 is similar by replacing the 'min' by 'max' in the formulae. For Theorem 1, the proof is straightforward (special case) where S contains only one group, say l, which simplifies the proof below. The derivation of the lower and upper probabilities of an event of interest require the following lemma presented and proven by Coolen and Yan (2003).

Lemma 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.

Proof of Theorem 3. First, let us consider the lower probability

$$P\left(\min_{l\in S} X_{l,n_l+1} > \max_{j\in NS} X_{j,n_j+1}\right) = P\left(\bigcap_{j\in NS}\{X_{j,n_j+1} < \min_{l\in S} X_{l,n_l+1}\}\right)$$

$$= \sum_{\substack{i_l=1\\ l\in S}}^{r_l+2} P\left(\bigcap_{j\in NS}\{X_{j,n_j+1} < \min_{l\in S} X_{l,n_l+1}\}|X_{l,n_l+1} \in I^l_{i_l}, l \in S\right) \cdot \prod_{l\in S} M_{X_{l,n_l+1}}(I^l_{i_l})$$

$$\geq \sum_{\substack{i_l=1\\ l\in S}}^{r_l+2} P\left(\bigcap_{j\in NS}\{X_{j,n_j+1} < \min_{l\in S}\{L(I^l_{i_l})\}\}\right) \cdot \prod_{l\in S} M_{X_{l,n_l+1}}(I^l_{i_l})$$

$$\geq \sum_{\substack{i_l=1\\ l\in S}}^{r_l+2} \prod_{j\in NS} \left[\frac{\sum_{i_j=1}^{r_j} 1\{x_{j,i_j} < \min_{l\in S}\{L(I^l_{i_l})\}\}}{n_j+1}\right] \cdot \prod_{l\in S} M_{X_{l,n_l+1}}(I^l_{i_l})$$

The first inequality follows by putting all probability mass for X_{l,n_l+1} $(l \in S)$ assigned to the intervals $I^l_{i_l} = (x_{l,i_l-1}, x_{l,i_l})$ $(i_l = 1,\ldots,r_l)$, (x_{l,r_l},∞) and (T_0,∞) in the left end points of these intervals, and by using Lemma 1 for the nested intervals (x_{l,r_l},∞) and (T_0,∞). The second inequality follows by putting all probability mass for X_{j,n_j+1} $(j \in NS)$ assigned to the intervals $I^j_{i_j} = (x_{j,i_j-1}, x_{j,i_j})$ $(i_j = 1,\ldots,r_j)$, (x_{j,r_j},∞) and (T_0,∞) in the right end points of these intervals.

Similarly, we derive the upper probability

$$P\left(\min_{l\in S} X_{l,n_l+1} > \max_{j\in NS} X_{j,n_j+1}\right) = P\left(\bigcap_{j\in NS}\{X_{j,n_j+1} < \min_{l\in S} X_{l,n_l+1}\}\right)$$

$$= \sum_{\substack{i_l=1\\ l\in S}}^{r_l+2} P\left(\bigcap_{j\in NS}\{X_{j,n_j+1} < \min_{l\in S} X_{l,n_l+1}\} | X_{l,n_l+1}\in I^l_{i_l}, l\in S\right) . \prod_{l\in S} M_{X_{l,n_l+1}}(I^l_{i_l})$$

$$\leq \sum_{\substack{i_l=1\\ l\in S}}^{r_l+2} P\left(\bigcap_{j\in NS}\{X_{j,n_j+1} < \min_{l\in S}\{U(I^l_{i_l})\}\}\right) . \prod_{l\in S} M_{X_{l,n_l+1}}(I^l_{i_l})$$

$$\leq \sum_{\substack{i_l=1\\ l\in S}}^{r_l+2} \prod_{j\in NS}\left[\sum_{i_j=1}^{r_j+2} 1\{L(I^j_{i_j}) < \min_{l\in S}\{U(I^l_{i_l})\}\}.\, M_{X_{j,n_j+1}}(I^j_{i_j})\right] \prod_{l\in S} M_{X_{l,n_l+1}}(I^l_{i_l})$$

$$= \sum_{\substack{i_l=1\\ l\in S}}^{r_l+2} \prod_{j\in NS}\left[\frac{1+\sum_{i_j=1}^{r_j} 1\{x_{j,i_j} < \min_{l\in S}\{U(I^l_{i_l})\}\}}{n_j+1} + \right.$$

$$\left. \frac{(n_j-r_j)1\{T_0 < \min_{l\in S}\{U(I^l_{i_l})\}\}}{n_j+1}\right] . \prod_{l\in S} M_{X_{l,n_l+1}}(I^l_{i_l})$$

The first inequality follows by putting all probability mass for X_{l,n_l+1} $(l\in S)$ assigned to the intervals $I^l_{i_l} = (x_{l,i_l-1}, x_{l,i_l})$ $(i_l = 1,\ldots,r_l)$, (x_{l,r_l},∞) and (T_0,∞) in the right end points of these intervals, and by using Lemma 1 for the nested intervals (x_{l,r_l},∞) and (T_0,∞). The second inequality follows by putting all probability mass for X_{j,n_j+1} $(j\in NS)$ assigned to the intervals $I^j_{i_j} = (x_{j,i_j-1}, x_{j,i_j})$ $(i_j = 1,\ldots,r_j)$, (x_{j,r_j},∞) and (T_0,∞) in the left end points of these intervals.

References

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

Balakrishnan, N., Ng, H. K. T., 2006. Precedence-type tests and applications. Wiley, Hoboken, N.J. ; Chichester.

Balakrishnan, N., Ng, H. K. T., Panchapakesan, S., 2006. A nonparametric procedure based on early failures for selecting the best population using a test for equality. Journal of Statistical Planning and Inference 136, 2087–2111.

Bechhofer, R. E., 1954. A single-sample multiple decision procedure for ranking means of normal populations with known variances. Annals of Mathematical Statistics 25, 16–39.

Bechhofer, R. E., Santner, T. J., Goldsman, D. M., 1995. Design and Analysis of Experiments for Statistical Selection, Screening, and Multiple Comparisons. Wiley, New York.

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., van der Laan, P., 2001. Imprecise predictive selection based on low structure assumptions. Journal of Statistical Planning and Inference 98, 259–277.

Coolen, F. P. A., Yan, K. J., 2003. Comparing two groups of lifetime data. In: ISIPTA'03: Proceedings of the Third International Symposium on Imprecise Probabilities and their Applications. pp. 148–161.

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

Coolen-Schrijner, P., Maturi, T. A., Coolen, F. P. A., 2009. Nonparametric predictive precedence testing for two groups. Journal of Statistical Theory and Practice 3, 273–287.

De Finetti, B., 1974. Theory of probability: a critical introductory treatment. Wiley, London.

Epstein, B., 1955. Comparison of some non-parametric tests against normal alternatives with an application to life testing. Journal of the American Statistical Association 50, 894–900.

Gupta, S. S., 1965. On some multiple decision (selection and ranking) rules. Technometrics 7, 225–245.

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., 1993. Parametric models for A_n: Splitting processes and mixtures. Journal of the Royal Statistical Society. Series B (Methodological) 55, 423–433.

Lawless, J. F., 2003. Statistical models and methods for lifetime data. Wiley, New Jersey ; Hoboken, Wiley.

Lawless, J. F., Fredette, M., 2005. Frequentist prediction intervals and predictive distributions. Biometrika 92, 529–542.

Maturi, T. A, Coolen-Schrijner, P., Coolen, F.P.A., 2010. Nonparametric predictive comparison of lifetime data under progressive censoring. Journal of Statistical Planning and Inference 140, 515–525.

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

Nelson, W. B., 1972. Graphical analysis of accelerated life test data with the inverse power law model. IEEE Transactions on Reliability 21, 2-11.

Nelson, W. B., 1982. Applied life data analysis. Wiley, New York ; Chichester.

Ng, H. K. T., Balakrishnan, N., Panchapakesan, S., 2007. Selecting the best population using a test for equality based on minimal wilcoxon rank-sum precedence statistic. Methodology and Computing in Applied Probability 9, 263–305.

Shafer, G., 1976. A Mathematical Theory of Evidence. Princeton University Press, Princeton (US).

Verheijen, J. H. M., Coolen, F. P. A., van der Laan, P., 1997. Combining two classical approaches for statistical selection. Communications in Statistics: Theory and Methods 26, 1291–1312.

Walley, P., 1991. Statistical reasoning with imprecise probabilities. Chapman and Hall, London.

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