Nonparametric Predictive Inference for Ranking and Selection

Fatimah I. Alshihry[a], Tahani Coolen-Maturi[b], Frank P. A. Coolen[b]

[a]King Abdulaziz University, Jeddah, Saudi Arabia

[b]Department of Mathematical Sciences, Durham University, Durham, DH1 3LE, UK

Abstract

Selection and ranking of multiple groups are fundamental problems in many scientific and practical applications. This paper develops a Nonparametric Predictive Inference (NPI) framework for selection and ranking based on a single future observation from each group. The focus is on deriving NPI lower and upper probabilities for events concerning the identification and ordering of the best groups. Selection within subsets of independent groups is first studied, where 'better' is defined as the event that all future observations from the selected groups exceed those from the remaining groups. Exact NPI lower and upper probabilities are derived for selecting the two best groups, and approximate probabilities are obtained for selecting the three best groups. The methodology is subsequently extended to ranking groups organised into buckets, each containing one or more independent groups. Data-driven allocation of groups to buckets and the choice of the number of buckets are investigated, where the optimal configuration is defined as the one that optimises the relevant NPI lower and upper probabilities. An NPI-Bootstrap approach is also employed to approximate the probability of the event of interest. Simulated data and examples from the literature illustrate the proposed methods. The proposed framework provides a flexible and assumption-lean approach to selection and ranking, offering informative probability bounds that explicitly quantify predictive uncertainty.

Keywords: Nonparametric predictive inference (NPI); NPI-Bootstrap; selection; ranking; imprecise probabilities

1 Introduction

When comparing multiple populations, it is often necessary to determine which performs better. For example, in clinical trials involving several treatments, decision makers may wish to identify the most effective treatment, the safest option, or the treatment associated with the fewest side effects.

Selection problems aim to identify the best population, a subset of the best populations, or a subset that contains the best population. In other situations, the objective extends beyond selection to ranking populations according to a performance measure. For instance, in medical studies with multiple treatments, it may be important not only to determine the best-performing treatment, but also to rank treatments by effectiveness in order to prioritise further investigation.

The statistical theory of selection and ranking problems dates back to the 1950s, particularly in areas such as agriculture and clinical research, where testing homogeneity across multiple populations was common [37]. Homogeneity typically refers to equality of means, variances, or other characteristics. When statistical tests indicate differences—such as rejection of homogeneity of means—natural questions arise: how should uncertainty about being the best be quantified, and which population should be selected? For example, when comparing several drugs, a homogeneity test may reveal that they differ, but it does not indicate which drug is best. Here, "best" may refer to being most effective, safest, or having the fewest adverse effects. Thus, homogeneity testing alone is insufficient to address practical decision problems. To overcome this limitation, specialised selection and ranking procedures have been developed [20].

The meaning of "best" depends on the context. In some applications, it may correspond to a smaller number of fatalities in a medical study; in others, it may refer to a greater reduction in tumour size. More generally, selection and ranking procedures address questions such as: Which drug produces the strongest response? Which treatments outperform a placebo? Which advertising media reach the largest proportion of potential buyers? Which learning technique yields the greatest comprehension? How do political candidates rank in popularity at a given time? Which type of ski binding has the lowest accident rate?. Applications of selection and ranking procedures span diverse fields, including poultry science, drug development, advertising, stock investment, and supplier evaluation, highlighting their broad practical relevance [20].

In this paper, we develop Nonparametric Predictive Inference (NPI) methods for selection and ranking problems. NPI is a predictive statistical framework based on Hill's assumption [28,29], in which uncertainty about future observations is quantified through lower and upper probabilities. Due to its minimal modelling assumptions and intuitive predictive interpretation, NPI has been applied in various areas of statistics and reliability [6,9,14].

In classical statistics, selection and ranking problems are usually framed in terms of populations described by underlying probability distributions. In contrast, NPI does not assume a parametric model but bases inference directly on the observed samples. Therefore, throughout this paper, we use the term *group* to denote each set of observations corresponding to a treatment or population under comparison.

For selection problems, Coolen [5] introduced an NPI method for comparing two independent groups, aimed at identifying the group most likely to produce the largest future observation. Coolen and van der

Laan [13] extended this approach to multiple independent groups, considering both the selection of the best group and subsets containing the best groups. However, their work did not address ranking within a selected subset. The methodology developed in this paper extends NPI to settings where, in addition to identifying the best groups, their relative ordering within a selected subset is also determined.

The remainder of the paper is organised as follows. Section 2 reviews classical selection and ranking methods, and Section 3 introduces the main principles of NPI. Section 4 develops selection and ranking within the NPI framework. The general methodology for selecting a ranked subset is presented in Section 4.1, with Sections 4.2 and 4.3 examining in detail the cases of two and three selected groups, respectively. Section 4.4 discusses partial ranking, and Section 4.5 compares the proposed approach with alternative methods.

The paper then extends ranking to settings where groups are organised into buckets in Section 5. Section 5.1 derives NPI lower and upper probabilities for general bucket-ranking events, Section 5.2 considers data-driven bucket allocation, and Section 5.3 applies the NPI-bootstrap method to estimate ranking probabilities. Section 5.4 addresses the choice of the number of buckets. Finally, Section 6 concludes the paper.

2 Selection and ranking procedures

Over the past seventy years, substantial research has been devoted to selection and ranking procedures. A pioneering contribution is due to Bechhofer [4], who developed early methods for comparing multiple Normal populations and selecting the best one based on ranked sample means under known and common variances. Subsequent work shifted toward subset selection procedures, initiated by Gupta [23], where the aim is to select several good populations rather than a single best one. Gupta studied procedures for selecting a subset that contains the best population under parametric assumptions, and this line of work has been extended to various distributions, including Normal and Binomial models [22,24,26]. For Normal populations, the best population is typically defined as the one with the largest mean; for Binomial populations, it is the one with the highest success probability.

An important aspect of subset selection is the size of the selected subset, which may be fixed or random. In some applications, selecting only one population is not desirable, particularly when several populations perform similarly well. Procedures with random subset size determine the number of selected populations from the observed data [22,24,26], whereas fixed-size procedures ensure that the number of selected populations remains constant [17]. The choice depends on the experimenter's objectives. In this paper, the subset size is assumed to be fixed.

For Normal populations with common known variance, Gupta [23] proposed selecting a (possibly random-sized) subset that contains the best population with probability at least P^*. Related procedures for unknown

common variance were studied by Gupta and Sobel [25], and subset selection for Binomial populations was addressed in [26]. In contrast, Desu and Sobel [17] investigated selection of a fixed-size subset for Normal populations, building on earlier work by Mahamunulu [33]. For Normal populations with unknown and possibly unequal variances, Dudewicz and Dalal [19] proposed a two-stage procedure to select the population with the largest mean.

Beyond parametric settings, nonparametric and sequential selection procedures have also been developed. Sobel [40,41] studied selection problems based on distribution functions and quantiles, while Rizvi and Sobel [38] and Desu and Sobel [18] developed related nonparametric procedures. More recently, Kumar and Grover [32] proposed subset selection procedures based on U-statistics, and Kumar [31] introduced a general class of procedures applicable to both location and scale parameters.

Many classical procedures are formulated within an experimental design framework, where key characteristics such as indifference parameters and confidence levels are specified in advance. This includes Bechhofer's indifference-zone approach [4] and Tong's treatment–control procedures [42]. Although such design-based perspectives differ from the NPI framework adopted in this paper, they provide important context for selection and ranking methodology.

In the indifference-zone approach of Bechhofer [4], the objective is to select the best population among k independent Normal populations, where "best" refers to the largest mean. Assuming a common known variance σ^2, the ordered means are written as $\mu_{[1]} \leq \mu_{[2]} \leq \cdots \leq \mu_{[k]}$, and inference is based on the ordered sample means

$$\bar{x}_{[1]} \leq \bar{x}_{[2]} \leq \cdots \leq \bar{x}_{[k]}. \tag{1}$$

The experiment is designed such that the probability of correct selection and ranking, $P(CSR)$, is at least a prespecified value P^*. Under the least favourable configuration,

$$\mu_{[k]} - \mu_{[k-1]} \geq \delta^*, \tag{2}$$

the required common sample size per population is the smallest integer greater than or equal to

$$n = \left(\frac{d\sigma}{\delta^*} \right)^2, \tag{3}$$

where d is a multivariate Normal quantile constant corresponding to P^*.

Gupta's subset selection approach [23] similarly determines a common sample size for given δ^*, P^*, and

k. The minimum required sample size is the smallest integer greater than or equal to

$$n = \left(\frac{\lambda\sigma}{\delta^*}\right)^2, \tag{4}$$

where λ is a critical constant ensuring that the selected subset contains the best population with probability at least P^*. With ordered sample means $\bar{x}_{[1]} \leq \cdots \leq \bar{x}_{[k]}$, the ith population is included in the selected subset if and only if

$$\bar{x}_{[i]} \geq \bar{x}_{[k]} - \frac{\sigma}{\sqrt{n}}\lambda. \tag{5}$$

While classical methods focus on precise probability guarantees under parametric assumptions, the present paper considers ranking within a selected subset using the NPI framework, which does not rely on such assumptions.

3 Nonparametric Predictive Inference (NPI)

Nonparametric Predictive Inference (NPI) [2,8] is a statistical framework based on Hill's assumption $A_{(n)}$ [28, 29], which provides direct probabilities for a future observation given n observed related random quantities. Inferences derived from this assumption are predictive and nonparametric. They are particularly suitable when little prior knowledge is available about the random quantities of interest, or when one prefers not to impose additional modelling assumptions, for example when studying the effects of such assumptions in other statistical methods.

To introduce assumption $A_{(n)}$, suppose that $X_1, X_2, \ldots, X_{n+1}$ are continuous and exchangeable random quantities. Let the ordered observations of $X_1, \ldots, X_n$ be $x_1 < x_2 < \cdots < x_n$, and define $x_0 = -\infty$ and $x_{n+1} = \infty$. Note that x_{n+1} does not represent an observed value of X_{n+1}. The n observations divide the real line into $n+1$ intervals $I_j = (x_{j-1}, x_j)$ for $j = 1, \ldots, n+1$. Assumption $A_{(n)}$ states that the probability of the future observation X_{n+1} falling within the open interval I_j is equal for all I_j, which is

$$P(X_{n+1} \in I_j) = \frac{1}{n+1}, \quad j = 1, \ldots, n+1. \tag{6}$$

Throughout this paper, ties are assumed not to occur. If ties are present, one may assign probability mass to closed intervals $[x_{j-1}, x_j]$ instead of open intervals, or treat tied observations as differing by an arbitrarily small amount [30].

Assumption $A_{(n)}$ makes no further distributional assumptions and can be viewed as a post-data assumption related to exchangeability [16]. It does not yield precise probabilities for most events of interest [29].

Instead, it provides probability bounds via an application of de Finetti's fundamental theorem of probability [16]. These bounds are called lower and upper probabilities in imprecise probability theory [43] and interval probability theory [44]. For an event A, the NPI lower and upper probabilities are denoted by $\underline{P}(A)$ and $\overline{P}(A)$, respectively. The lower probability can be interpreted as a maximal lower bound for the precise probability of A, and the upper probability as a minimal upper bound [8]. Hence, uncertainty is quantified by the interval $[\underline{P}(A), \overline{P}(A)]$. Augustin and Coolen [2] defined the NPI lower and upper probabilities for the event $X_{n+1} \in B$, where $B \subset \mathbb{R}$, as

$$\underline{P}(X_{n+1} \in B) = \frac{1}{n+1} \sum_{j=1}^{n+1} \mathbf{1}\{I_j \subseteq B\}, \tag{7}$$

$$\overline{P}(X_{n+1} \in B) = \frac{1}{n+1} \sum_{j=1}^{n+1} \mathbf{1}\{I_j \cap B \neq \emptyset\}, \tag{8}$$

where $\mathbf{1}\{A\}$ is the indicator function. The lower probability sums the probability masses of intervals entirely contained in B, while the upper probability sums the probability masses of intervals that intersect B. These probabilities satisfy the conjugacy property $\overline{P}(A) = 1 - \underline{P}(A^c)$, and $0 \leq \underline{P}(A) \leq \overline{P}(A) \leq 1$. If $\underline{P}(A) = \overline{P}(A)$, the probability is precise. The quantity $\triangle(A) = \overline{P}(A) - \underline{P}(A)$ is called imprecision [7].

NPI procedures for selection and ranking problems have been developed extensively in the literature. NPI for pairwise comparison of real-valued observations was introduced by Coolen [5]. In this setting, two independent groups are compared based on a single future observation from each group, and selection is performed using the NPI lower and upper probabilities for the event that one group yields the larger future observation. The NPI framework was subsequently extended to multiple comparisons for $k \geq 2$ independent groups and to different data types. For real-valued observations, Coolen [13] considered three events of interest based on one future observation per group: selection of the best group, selection of the subset of best groups, and selection of a subset that contains the best group.

Further developments have extended NPI to other data structures and inference problems. Maturi et al. [35] studied comparisons of future observations when data may be tail-terminated. Coolen and Coolen-Schrijner [12] developed NPI methods for selecting the optimal group in Bernoulli trials, while Baker [3] considered selection of the best category for multinomial data. Maturi [34] extended multiple comparisons to right-censored data. Coolen et al. [11] further developed NPI methods for pairwise and multiple comparisons focusing on future order statistics. In all these studies, multiple comparisons for $k \geq 2$ populations were considered with respect to the three events described above.

Beyond selection, ranking of multiple groups is often of interest. For example, in treatment evaluation one may wish to determine whether groups follow a specific order of effectiveness. Ranking within the

NPI framework has been investigated by Coolen-Maturi [15], who developed methods for ranking future observations from three or more real-valued groups.

For three groups, deriving NPI lower and upper probabilities for ordered events such as $X_{n_x+1} < Y_{n_y+1} < Z_{n_z+1}$ requires assigning the probability masses for the extreme groups (X and Z) to interval endpoints that minimise the lower probability and maximise the upper probability. For the middle group (Y), an optimisation over the allocation of probability mass within each interval is required.

For four or more groups, the complexity increases because probability masses for multiple middle groups must be considered simultaneously. Minimising the lower probability and maximising the upper probability require evaluating all feasible allocations of probability mass across these groups. This becomes computationally demanding, particularly for large and overlapping datasets. Consequently, bounds have been introduced in the literature to avoid this computational burden.

4 NPI for selecting a ranked subset

In many applications, the objective is not only to identify the best groups but also to rank them within a selected subset. For example, in clinical trials involving multiple treatments, it is important to determine which treatments perform best and to order them according to effectiveness, thereby supporting prioritisation for further investigation. Similar challenges arise in medicine, industry, and economics.

Coolen and van der Laan [13] proposed an NPI approach for selecting a subset containing multiple best groups, but without addressing ranking within the selected subset. The methodology developed here extends this idea by explicitly incorporating the ranking of groups within the subset.

This section presents a framework for selecting a ranked subset of best groups based on one future observation from each group. We first consider the general case in which multiple groups are included in the subset and introduce the required notation and probability expressions. Two important special cases are then examined, where the subset contains two and three groups, respectively. These cases illustrate how the derivation of NPI lower and upper probabilities depends on the subset size. Finally, partial ranking within a selected subset is briefly discussed.

4.1 Selecting a ranked subset of the best groups

This section introduces the NPI-based methodology for selecting a ranked subset of the best groups, based on one future observation from each group. It presents the general formulation of the ranked subset selection event, which forms the foundation for the special cases with two and three groups developed in the subsequent sections. The notation is first introduced, followed by the derivation of the NPI lower and upper probabilities

for the event of interest. Since the computation of these probabilities can be demanding, particularly when observations from different groups overlap, lower and upper bounds are also derived.

For simplicity, the next future observation from each group is denoted by X^f instead of X_{n_x+1}. Suppose there are k independent groups, where independence means that information about the random quantities for one group does not provide information about the random quantities for another group. Consider a subset consisting of w groups, with $1 \leq w \leq k-1$. Let $S = \{\ell_1, \ell_2, \ldots, \ell_w\} \subset \{1, 2, \ldots, k\}$ denote the selected subset of size w, and let $NS = \{1, 2, \ldots, k\} \setminus S$ denote the remaining $k - w$ non-selected groups.

The event of interest is $\max\limits_{j_z \in NS} X^f_{j_z} < X^f_{\ell_w} < \cdots < X^f_{\ell_2} < X^f_{\ell_1}$, which states that all non-selected groups have smaller future observations than the selected groups and that the selected groups are strictly ranked. The objective is therefore to select the ranked subset S of the best w groups.

Lower and upper probabilities

Coolen and van der Laan [13] introduced NPI lower and upper probabilities for the event $\max\limits_{j \in NS} X^f_j < \min\limits_{\ell \in S} X^f_\ell$, that is, the event that all future observations from groups in S exceed those from groups in NS. We now extend this methodology to the ranked event $\max\limits_{j_z \in NS} X^f_{j_z} < X^f_{\ell_w} < \ldots < X^f_{\ell_2} < X^f_{\ell_1}$, which strengthens the earlier condition by additionally imposing an ordering within the selected subset. Throughout, we assume that there are no tied observations.

To derive the NPI lower and upper probabilities for this event, Hill's assumption $A_{(n)}$ is applied separately to each group. For a selected group $\ell_i \in S$, let the ordered observations be $x_{\ell_i,1} < x_{\ell_i,2} < \ldots < x_{\ell_i,n_{\ell_i}}$, with $x_{\ell_i,0} = -\infty$ and $x_{\ell_i,n_{\ell_i}+1} = \infty$. The n_{ℓ_i} observations divide the real line into $n_{\ell_i} + 1$ intervals, denoted by $\{I_1, \ldots, I_{n_{\ell_i}+1}\}$. Under $A_{(n_{\ell_i})}$,

$$P\Big(X^f_{\ell_i} \in (x_{\ell_i,u_{\ell_i}-1}, x_{\ell_i,u_{\ell_i}})\Big) = \frac{1}{n_{\ell_i}+1}, \quad u_{\ell_i} = 1, \ldots, n_{\ell_i}+1.$$

Similarly, for each non-selected group $j_z \in NS$, let the ordered observations be $x_{j_z,1} < x_{j_z,2} < \ldots < x_{j_z,n_{j_z}}$, with $x_{j_z,0} = -\infty$ and $x_{j_z,n_{j_z}+1} = \infty$. These define $n_{j_z} + 1$ intervals, and under $A_{(n_{j_z})}$,

$$P\Big(X^f_{j_z} \in (x_{j_z,v_{j_z}-1}, x_{j_z,v_{j_z}})\Big) = \frac{1}{n_{j_z}+1}, \quad v_{j_z} = 1, \ldots, n_{j_z}+1.$$

For the NPI lower probability, the probability mass of group ℓ_1 is assigned to the left endpoint of each interval $(x_{\ell_1,u_{\ell_1}-1}, x_{\ell_1,u_{\ell_1}})$, while for all $j_z \in NS$ it is assigned to the right endpoint of $(x_{j_z,v_{j_z}-1}, x_{j_z,v_{j_z}})$.

For the NPI upper probability, these assignments are reversed. Hence,

$$\underline{P}\left(\max_{j_z \in NS} X^f_{j_z} < X^f_{\ell_w} < \ldots < X^f_{\ell_2} < X^f_{\ell_1}\right) = \frac{1}{\prod_{t=1}^{k}(n_t+1)} \sum_{u_{\ell_1}=1}^{n_{\ell_1}+1} \sum_{\substack{v_{j_z}=1 \\ j_z \in NS}}^{n_{j_z}+1} P\left(\max_{j_z \in NS} x_{j_z,v_{j_z}} < X^f_{\ell_w} < \ldots < X^f_{\ell_2} < x_{\ell_1,u_{\ell_1}-1} \,\middle|\, D\right), \tag{9}$$

where $D = \{X^f_{\ell_i} \in (x_{\ell_i,u_{\ell_i}-1}, x_{\ell_i,u_{\ell_i}}), i = 2,\ldots,w\}$ and the notation $\sum_{\substack{v_{j_z}=1 \\ j_z \in NS}}^{n_{j_z}+1}$ is used for $k-w$ sums $\sum_{v_{j_1}=1}^{n_{j_1}+1} \sum_{v_{j_2}=1}^{n_{j_2}+1} \ldots \sum_{v_{j_{k-w}}=1}^{n_{j_{k-w}}+1}$. Similarly, the NPI upper probability is

$$\overline{P}\left(\max_{j_z \in NS} X^f_{j_z} < X^f_{\ell_w} < \ldots < X^f_{\ell_2} < X^f_{\ell_1}\right) = \frac{1}{\prod_{t=1}^{k}(n_t+1)} \sum_{u_{\ell_1}=1}^{n_{\ell_1}+1} \sum_{\substack{v_{j_z}=1 \\ j_z \in NS}}^{n_{j_z}+1} P\left(\max_{j_z \in NS} x_{j_z,v_{j_z}-1} < X^f_{\ell_w} < \ldots < X^f_{\ell_2} < x_{\ell_1,u_{\ell_1}} \,\middle|\, D\right). \tag{10}$$

We have shown above how the probability masses are assigned for the largest selected group ℓ_1 and for the non-selected groups $j_z \in NS$. For the intermediate selected groups $\ell_2,\ldots,\ell_w$, however, the allocation of probability mass within each interval is not immediate when deriving the NPI lower and upper probabilities. Following Coolen-Maturi [15], the optimisation is carried out sequentially, one middle group at a time. When optimising with respect to a group ℓ_i, $i = 2,\ldots,w$, the remaining middle groups are assigned to either the left or right endpoints of their intervals $(x_{\ell_i,u_{\ell_i}-1}, x_{\ell_i,u_{\ell_i}})$. This yields the extremal configuration required for the NPI lower and upper probabilities. The procedure is illustrated in Sections 4.2 and 4.3.

Bounds for the NPI lower and upper probabilities

Calculating the NPI lower and upper probabilities involves optimisation over different sub-intervals to determine how the probability mass of each group is distributed. This process can become cumbersome. Therefore, we consider lower and upper bounds for the NPI lower and upper probabilities, which are easier to compute.

First, consider the lower bound for the NPI lower probability of the event $\max_{j_z \in NS} X^f_{j_z} < X^f_{\ell_w} < \ldots < X^f_{\ell_2} < X^f_{\ell_1}$, denoted by $\underline{P}^L$. This bound is derived by requiring total separation of the intervals $(x_{\ell_i,u_{\ell_i}-1}, x_{\ell_i,u_{\ell_i}})$, for $i = 1,2,\ldots,w$ and $u_{\ell_i} = 1,2,\ldots,n_{\ell_i}+1$, and $(x_{j_z,v_{j_z}-1}, x_{j_z,v_{j_z}})$, for $z = 1,2,\ldots,k-w$ and $v_{j_z} = 1,2,\ldots,n_{j_z}+1$. In other words, the lower bound counts only those configurations in which the required ordering is guaranteed by the complete separation of the corresponding intervals. Let $I\{A\}$ denote the

indicator function, equal to one if the event A occurs and zero otherwise. The lower bound for the NPI lower probability is

$$\underline{P}^L\left(\max_{j_z\in NS} X^f_{j_z} < X^f_{\ell_w} < \ldots < X^f_{\ell_2} < X^f_{\ell_1}\right) = \frac{1}{\prod_{t=1}^k (n_t+1)} \sum_{\substack{u_{\ell_i}=1 \\ \ell_i\in S}}^{n_{\ell_i}+1} \sum_{\substack{v_{j_z}=1 \\ j_z\in NS}}^{n_{j_z}+1}$$

$$I\left\{\left(\max_{j_z\in NS} x_{j_z,v_{j_z}} < x_{\ell_w,u_{\ell_w}-1}\right) \wedge \bigcap_{i=1}^{w-1}\left\{x_{\ell_{w-i+1},\,u_{\ell_{w-i+1}}} < x_{\ell_{w-i},\,u_{\ell_{w-i}}-1}\right)\right\} \tag{11}$$

where $\sum_{\substack{u_{\ell_i}=1 \\ \ell_i\in S}}^{n_{\ell_i}+1}$ denotes the w nested sums $\sum_{u_{\ell_1}=1}^{n_{\ell_1}+1}\sum_{u_{\ell_2}=1}^{n_{\ell_2}+1}\cdots\sum_{u_{\ell_w}=1}^{n_{\ell_w}+1}$. The lower bound $\underline{P}^L$ is obtained by successive conditioning on the interval events under $A_{(n)}$ and retaining only those configurations that satisfy the strict ordering implied by total separation.

The upper bound for the NPI lower probability is derived by assigning the probability mass for group ℓ_1 to the left endpoint of each interval and the probability masses for the non-selected groups j_z to the right endpoints. The middle groups ℓ_i, $i = 2,\ldots,w$, may be assigned to either endpoint, as their placement does not affect the value of the bound. This is because assigning their masses to the left endpoints introduces $-\infty$ as a lower bound, while assigning them to the right endpoints introduces ∞ as an upper bound; neither affects the required ordering relative to group ℓ_1. The upper bound for the NPI lower probability is

$$\underline{P}^U\left(\max_{j_z\in NS} X^f_{j_z} < X^f_{\ell_w} < \ldots < X^f_{\ell_2} < X^f_{\ell_1}\right) = \quad \frac{1}{\prod_{t=1}^k (n_t+1)} \sum_{\substack{u_{\ell_i}=1 \\ \ell_i\in S}}^{n_{\ell_i}+1} \sum_{\substack{v_{j_z}=1 \\ j_z\in NS}}^{n_{j_z}+1}$$

$$I\left(\max_{j_z\in NS} x_{j_z,v_{j_z}} < x_{\ell_w,u_{\ell_w}} < x_{\ell_{w-1},u_{\ell_{w-1}}} < \ldots < x_{\ell_1,u_{\ell_1}-1}\right) \tag{12}$$

The lower bound for the NPI upper probability is derived similarly, but with a different assignment of probability mass. The probability mass for group ℓ_1 is assigned to the right endpoints of the intervals, while the probability masses for the non-selected groups j_z are assigned to the left endpoints. Thus, the lower bound for the NPI upper probability is

$$\overline{P}^L\left(\max_{j_z\in NS} X^f_{j_z} < X^f_{\ell_w} < \ldots < X^f_{\ell_2} < X^f_{\ell_1}\right) = \frac{1}{\prod_{t=1}^k (n_t+1)} \sum_{\substack{u_{\ell_i}=1 \\ \ell_i\in S}}^{n_{\ell_i}+1} \sum_{\substack{v_{j_z}=1 \\ j_z\in NS}}^{n_{j_z}+1}$$

$$I\left(\max_{j_z\in NS} x_{j_z,v_{j_z}-1} < x_{\ell_w,u_{\ell_w}} < x_{\ell_{w-1},u_{\ell_{w-1}}} < \ldots < x_{\ell_1,u_{\ell_1}}\right) \tag{13}$$

Finally, the upper bound for the NPI upper probability is obtained by considering all possible combinations of the intervals $(x_{\ell_i,u_{\ell_i}-1}, x_{\ell_i,u_{\ell_i}})$ and $(x_{j_z,v_{j_z}-1}, x_{j_z,v_{j_z}})$, and counting all configurations that are compatible

with the required ordering. Equivalently, this corresponds to identifying all interval configurations whose intersections allow the strict ordering to hold. The upper bound for the NPI upper probability is

$$\overline{P}^{U}\left(\max_{j_z\in NS} X^f_{j_z} < X^f_{\ell_w} < \ldots < X^f_{\ell_2} < X^f_{\ell_1}\right) = \frac{1}{\prod_{t=1}^{k}(n_t+1)} \sum_{\substack{u_{\ell_i}=1 \\ \ell_i\in S}}^{n_{\ell_i}+1} \sum_{\substack{v_{j_z}=1 \\ j_z\in NS}}^{n_{j_z}+1}$$

$$\bigcap_{i=1}^{w} I\Big\{\max_{j_z\in NS} x_{j_z,v_{j_z}-1} < x_{\ell_i,u_{\ell_i}}\Big\} \wedge \bigcap_{i=2}^{w}\bigcap_{a=1}^{i-1} I\Big\{x_{\ell_i,u_{\ell_i}-1} < x_{\ell_a,u_{\ell_a}}\Big\} \tag{14}$$

Special case for perfectly ordered data

A special scenario arises when the data are perfectly ordered, with no overlap between observations from different groups; that is, the groups are completely separated. Specifically, all observations in group ℓ_1 exceed all observations in the second-largest group ℓ_2, all observations in ℓ_2 exceed those in ℓ_3, and so on.

To indicate this special case, the subscript s is added to the bounds. Under perfect separation, Equations (11)–(14) simplify to

$$\underline{P}^{L}_{s}\left(\max_{j_z\in NS} X^f_{j_z} < X^f_{\ell_w} < \ldots < X^f_{\ell_1}\right) = \frac{1}{\prod_{t=1}^{k}(n_t+1)} \prod_{z=1}^{k-w} n_{j_z}\, n_{\ell_1} \prod_{i=2}^{w}(n_{\ell_i}-1) \tag{15}$$

$$\underline{P}^{U}_{s}\left(\max_{j_z\in NS} X^f_{j_z} < X^f_{\ell_w} < \ldots < X^f_{\ell_1}\right) = \frac{1}{\prod_{t=1}^{k}(n_t+1)} \prod_{z=1}^{k-w} n_{j_z} \prod_{i=1}^{w} n_{\ell_i} \tag{16}$$

$$\overline{P}^{L}_{s}\left(\max_{j_z\in NS} X^f_{j_z} < X^f_{\ell_w} < \ldots < X^f_{\ell_1}\right) = \frac{1}{\prod_{t=1}^{k}(n_t+1)} \prod_{z=1}^{k-w} (n_{j_z}+1)(n_{\ell_1}+1) \prod_{i=2}^{w} n_{\ell_i} \tag{17}$$

$$\overline{P}^{U}_{s}\left(\max_{j_z\in NS} X^f_{j_z} < X^f_{\ell_w} < \ldots < X^f_{\ell_1}\right) = \frac{1}{\prod_{t=1}^{k}(n_t+1)} \prod_{z=1}^{k-w} (n_{j_z}+1) \prod_{i=1}^{w}(n_{\ell_i}+1) \tag{18}$$

These expressions provide a direct way to compute the bounds in the absence of overlap. Although this represents an extreme and rarely encountered situation, it is useful for illustrating how the bounds behave when the data are perfectly ordered and fully separated.

The next two sections present special cases in which the selected subset contains two and three groups, respectively. These cases demonstrate how the NPI lower and upper probabilities depend on the number of groups included in the subset.

4.2 Selecting a ranked subset of two best groups

This section presents the NPI method for selecting a ranked subset of the two best groups based on a single future observation from each group. In Section 4.1, the general event of interest was $\max_{j_z\in NS} X^f_{j_z} < X^f_{\ell_w} <$

$\ldots < X^f_{\ell_2} < X^f_{\ell_1}$. When the subset size is $w = 2$, this reduces to $\max_{j_z \in NS} X^f_{j_z} < X^f_{\ell_2} < X^f_{\ell_1}$. The case $w = 1$ has already been studied in the literature [15], so the focus here is on $w = 2$. Hill's assumption $A_{(n)}$ is applied to each group, as described in Section 4.1, to derive the NPI lower and upper probabilities for the event $\max_{j_z \in NS} X^f_{j_z} < X^f_{\ell_2} < X^f_{\ell_1}$, where $S = \{\ell_1, \ell_2\} \subset \{1, \ldots, k\}$.

As discussed in Section 4.1, for the NPI lower (upper) probability the probability mass of group ℓ_1 is assigned to the left (right) endpoint of each interval $(x_{\ell_1, u_{\ell_1}-1}, x_{\ell_1, u_{\ell_1}})$, while the masses of all $j_z \in NS$ are assigned to the right (left) endpoints of $(x_{j_z, v_{j_z}-1}, x_{j_z, v_{j_z}})$. The remaining issue is how to allocate the probability mass $\frac{1}{n_{\ell_2}+1}$ of the middle group ℓ_2 within each interval $(x_{\ell_2, u_{\ell_2}-1}, x_{\ell_2, u_{\ell_2}})$, where $u_{\ell_2} = 1, \ldots, n_{\ell_2}+1$. This is addressed through the following optimisation procedure.

To obtain the exact NPI lower probability:

1. For each interval $(x_{\ell_2, u_{\ell_2}-1}, x_{\ell_2, u_{\ell_2}})$, determine the sub-intervals induced by the observations of group ℓ_1 and of all $j_z \in NS$, $z = 1, \ldots, k-2$. The number of such sub-intervals equals $n_{\ell_1} + \sum_{z=1}^{k-2} n_{j_z} + 1$.
2. Let $S^{t^{u_{\ell_2}}_{\min}}_{x_{\ell_1}}$ denote the number of probability masses $\frac{1}{n_{\ell_1}+1}$ assigned to left endpoints, and let $S^{t^{u_{\ell_2}}_{\min}}_{x_{j_z}}$ denote the number of probability masses $\frac{1}{n_{j_z}+1}$ assigned to right endpoints. The probability mass of ℓ_2 is allocated to a point $t^{u_{\ell_2}}$ (taken as the midpoint) within each sub-interval of $(x_{\ell_2, u_{\ell_2}-1}, x_{\ell_2, u_{\ell_2}})$.
3. The NPI lower probability is obtained by minimising $S^{t^{u_{\ell_2}}_{\min}}_{x_{\ell_1}} \times S^{t^{u_{\ell_2}}_{\min}}_{x_{j_z}}$ over all sub-intervals.

For the exact NPI upper probability:

1. Similarly, determine the sub-intervals induced by the observations of group ℓ_1 and of all $j_z \in NS$, $z = 1, \ldots, k-2$, within $(x_{\ell_2, u_{\ell_2}-1}, x_{\ell_2, u_{\ell_2}})$.
2. Let $S^{t^{u_{\ell_2}}_{\max}}_{x_{\ell_1}}$ denote the number of probability masses $\frac{1}{n_{\ell_1}+1}$ assigned to right endpoints, and let $S^{t^{u_{\ell_2}}_{\max}}_{x_{j_z}}$ denote the number of probability masses $\frac{1}{n_{j_z}+1}$ assigned to left endpoints. The probability mass of ℓ_2 is again allocated to a point $t^{u_{\ell_2}}$ (taken as the midpoint) within each sub-interval of $(x_{\ell_2, u_{\ell_2}-1}, x_{\ell_2, u_{\ell_2}})$.
3. The NPI upper probability is obtained by maximising $S^{t^{u_{\ell_2}}_{\max}}_{x_{\ell_1}} \times S^{t^{u_{\ell_2}}_{\max}}_{x_{j_z}}$ over all sub-intervals.

The resulting exact NPI lower and upper probabilities are

$$\underline{P}\left(\max_{j_z \in NS} X^f_{j_z} < X^f_{\ell_2} < X^f_{\ell_1}\right) = \frac{1}{\prod_{t=1}^{k}(n_t+1)} \sum_{u_{\ell_2}=1}^{n_{\ell_2}+1} S^{t^{u_{\ell_2}}_{\min}}_{x_{\ell_1}} \times S^{t^{u_{\ell_2}}_{\min}}_{x_{j_z}}, \tag{19}$$

$$\overline{P}\left(\max_{j_z \in NS} X^f_{j_z} < X^f_{\ell_2} < X^f_{\ell_1}\right) = \frac{1}{\prod_{t=1}^{k}(n_t+1)} \sum_{u_{\ell_2}=1}^{n_{\ell_2}+1} S^{t^{u_{\ell_2}}_{\max}}_{x_{\ell_1}} \times S^{t^{u_{\ell_2}}_{\max}}_{x_{j_z}}. \tag{20}$$

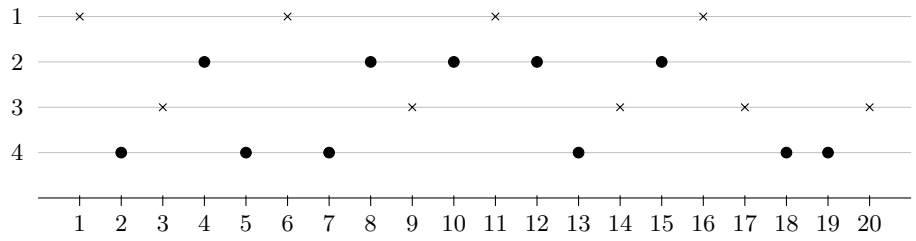

Figure 1: Four-group dataset used in Example 1.

The optimisation becomes increasingly demanding as n_{ℓ_2} grows, since each interval $(x_{\ell_2,u_{\ell_2}-1}, x_{\ell_2,u_{\ell_2}})$ may contain many sub-intervals requiring separate evaluation. In such cases, the lower and upper bounds given in Equations (11)–(14) with $w = 2$ provide a computationally simpler alternative.

Example 1. We consider the dataset from [15], consisting of $k = 4$ groups with sample sizes $n_1 = 4$, $n_2 = 5$, $n_3 = 5$, and $n_4 = 6$ (see Figure 1). Let $S = \{\ell_1, \ell_2\}$ denote the selected subset and $NS = \{j_1, j_2\}$ the non-selected groups. The event of interest is $\max_{j_z \in NS} X^f_{j_z} < X^f_{\ell_2} < X^f_{\ell_1}$. There are 12 possible ordered selections of two best groups out of four, listed in Table 1. We begin with $(\ell_1, \ell_2) = (1, 2)$, so that groups 3 and 4 belong to NS.

To derive the exact NPI lower probability for $\max_{j_z \in NS} X^f_{j_z} < X^f_{\ell_2} < X^f_{\ell_1}$, the probability mass of group 1 is assigned to the left endpoint of each interval, while the probability masses of groups 3 and 4 are assigned to the right endpoints.

For group 2, the probability mass can be placed at any point $t^{u_{\ell_2}}$, $u_{\ell_2} = 1, \ldots, 6$, within suitable sub-intervals of (x_{2,u_2-1}, x_{2,u_2}) in order to minimise the probability. Group 2 has six intervals: $(-\infty, 4)$, $(4, 8)$, $(8, 10)$, $(10, 12)$, $(12, 15)$ and $(15, \infty)$. In each interval, we select the sub-interval that minimises the product $S^t_{(3,4)} \times S^t_1$.

For $(-\infty, 4)$, the minimising sub-interval is $(-\infty, 3)$, yielding $S^t_{(3,4)} = 0$ and $S^t_1 = 4$, so the contribution is 0. For $(4, 8)$, the sub-interval $(4, 5)$ is optimal, giving $S^t_{(3,4)} = 1$ and $S^t_1 = 3$, hence a contribution of 3. For $(8, 10)$, the sub-interval $(8, 9)$ minimises the probability and yields a contribution of 6. The same reasoning applies to the next two intervals, $(10, 12)$ and $(12, 15)$, where the first sub-interval in each case is optimal and also gives a contribution of 6. Finally, for $(15, \infty)$, the sub-interval $(16, \infty)$ yields zero contribution since no probability mass from group 1 lies to its right. Summing the contributions gives

$$\underline{P}\Big(\max(X^f_3, X^f_4) < X^f_2 < X^f_1\Big) = \frac{1}{\prod_{t=1}^{4}(n_t + 1)} \sum_{u_{\ell_2}=1}^{6} S^{t^{u_{\ell_2}}_{\min}}_{(3,4)} \times S^{t^{u_{\ell_2}}_{\min}}_{1} = \frac{21}{1260} = 0.0167.$$

For the exact NPI upper probability, the sub-intervals are chosen to maximise the same product. The optimal

ℓ_1	ℓ_2	$\underline{P}$	$\overline{P}$
1	2	0.0167	0.1539
2	1	0.0063	0.1317
1	3	0.0071	0.1889
3	1	0.0278	0.2302
1	4	0.0032	0.1484
4	1	0.0151	0.1913
2	3	0.0103	0.1921
3	2	0.0508	0.2881
2	4	0.0079	0.1500
4	2	0.0269	0.2071
3	4	0.0413	0.2794
4	3	0.0286	0.2698

Table 1: Exact NPI lower and upper probabilities for all ordered selections for $\max_{j_z \in NS} X^f_{j_z} < X^f_{\ell_2} < X^f_{\ell_1}$ in Example 1.

sub-intervals are $(3,4)$, $(7,8)$, $(9,10)$, $(10,11)$, $(14,15)$ and $(20,\infty)$, with corresponding contributions 16, 24, 36, 36, 40 and 41. Hence,

$$\overline{P}\Big(\max(X^f_3, X^f_4) < X^f_2 < X^f_1\Big) = \frac{194}{1260} = 0.1539.$$

Using Equations (11)–(14) for $w = 2$, the bounds for the NPI lower probability are

$$[\underline{P}^L, \underline{P}^U] = [0.0158,\ 0.0309],$$

and for the NPI upper probability,

$$[\overline{P}^L, \overline{P}^U] = [0.1111,\ 0.1754].$$

The exact NPI lower probability 0.0167 lies between its bounds and is close to the lower bound, while the exact NPI upper probability 0.1539 lies between its bounds and is closer to the upper bound.

Table 1 presents the exact NPI lower and upper probabilities for all 12 ordered selections. The ordering $(3,2)$ yields the highest lower and upper probabilities and is therefore most supported. The ordering $(3,4)$ gives the second highest values. In contrast, selecting $(1,4)$ yields the smallest exact lower probability, whereas $(2,1)$ yields the smallest exact upper probability. Overall, the probabilities remain small, with imprecision ranging from 0.1254 to 0.2381.

Example 2. This example investigates the NPI method for selecting a ranked subset of the two best groups among four groups through simulation. Data are generated from Normal distributions: $X_1 \sim N(0,1), X_2 \sim N(0.1,1), X_3 \sim N(0.2,1)$, and $X_4 \sim N(0.3,1)$, with a common sample size $n = 25$ for all groups.

Table 2 presents the exact NPI lower and upper probabilities for all 12 ordered selections of two groups.

		$h=1$		$h=1000$		$h=10000$	
ℓ_1	ℓ_2	$\underline{P}$	$\overline{P}$	$\underline{P}$	$\overline{P}$	$\underline{P}$	$\overline{P}$
1	2	0.0323	0.0653	0.0456	0.0814	0.0454	0.0812
2	1	0.0333	0.0663	0.0474	0.0838	0.0488	0.0857
1	3	0.0345	0.0657	0.0521	0.0911	0.0504	0.0889
3	1	0.0369	0.0693	0.0589	0.1002	0.0580	0.0987
1	4	0.0550	0.0969	0.0565	0.0981	0.0557	0.0970
4	1	0.0666	0.1116	0.0697	0.1151	0.0695	0.1147
2	3	0.0591	0.1015	0.0585	0.1004	0.0600	0.1023
3	2	0.0520	0.0904	0.0651	0.1087	0.0645	0.1081
2	4	0.0851	0.1406	0.0645	0.1091	0.0665	0.1119
4	2	0.1079	0.1657	0.0765	0.1244	0.0773	0.1257
3	4	0.0854	0.1393	0.0798	0.1299	0.0789	0.1287
4	3	0.1097	0.1669	0.0859	0.1378	0.0856	0.1366

Table 2: Exact NPI lower and upper probabilities for all ordered selections of two groups for the event $\max_{j_z \in NS} X^f_{j_z} < X^f_{\ell_2} < X^f_{\ell_1}$ in Example 2, for different numbers of simulation replications.

The results are shown for different numbers of simulation replications, $h \in \{1, 1000, 10000\}$, in order to examine the effect of increasing h on the selection of the two best groups.

The results demonstrate that increasing the number of simulation replications from 1 to 1000 and 10000 leads to consistent conclusions. Selecting group 4 as the best and group 3 as the second best is supported by the largest exact NPI lower and upper probabilities across all replication levels. In contrast, selecting group 1 as the best and group 2 as the second best yields among the smallest probabilities, making it the least favourable ordered selection.

The differences between the exact NPI lower and upper probabilities are small, particularly when the number of replications increases to 1000 and 10000. Most combinations differ only in the third decimal place. Some slightly larger variations occur for combinations such as selecting group 2 as best and group 3 as second best (lower probability), or selecting group 1 as best and group 3 as second best (upper probability). Overall, increasing h from 1 to 1000 and then to 10000 results in only minor numerical changes. Moreover, the imprecision—defined as the difference between the exact NPI upper and lower probabilities—remains stable as the number of replications increases.

4.3 Selecting a ranked subset of three best groups

Section 4.2 presented the NPI lower and upper probabilities for a ranked subset of the two best groups. This section extends the methodology to a ranked subset of the three best groups. The event of interest is $\max_{j_z \in NS} X^f_{j_z} < X^f_{\ell_3} < X^f_{\ell_2} < X^f_{\ell_1}$, which means that the future observation of each group in $S = \{\ell_1, \ell_2, \ell_3\}$ exceeds that of each group in NS, and the observations within S are ordered from smallest ($X^f_{\ell_3}$) to largest ($X^f_{\ell_1}$).

With two middle groups, ℓ_3 and ℓ_2, deriving the exact NPI lower and upper probabilities becomes computationally demanding, particularly for large data sets with substantial overlap between groups. Following [15], an optimisation procedure is applied separately to each middle group. The aim is to determine optimal sub-interval allocations for each group and then approximate the NPI lower and upper probabilities based on these optimisations.

Hill's assumption $A_{(n)}$ is applied to each group. As discussed in Section 4.1, the probability masses for the best group ℓ_1 and the groups in NS are assigned to the appropriate interval endpoints. The remaining task is to determine how to allocate the probability masses $\frac{1}{n_{\ell_3}+1}$ and $\frac{1}{n_{\ell_2}+1}$ within the intervals $(x_{\ell_3,u_{\ell_3}-1}, x_{\ell_3,u_{\ell_3}})$ and $(x_{\ell_2,u_{\ell_2}-1}, x_{\ell_2,u_{\ell_2}})$, respectively. Approximate NPI lower and upper probabilities are obtained by applying the optimisation in Equations (9) and (10) with $w = 3$.

Optimisation over ℓ_3. First, the optimisation is carried out for group ℓ_3. At this stage, the probability mass of ℓ_2 is assigned to either endpoint of its intervals, since this does not affect the optimisation over ℓ_3. The lower and upper probabilities are

$$\underline{P}^{\min \ell_3}\left(\max_{j_z \in NS} X^f_{j_z} < X^f_{\ell_3} < X^f_{\ell_2} < X^f_{\ell_1}\right) = \frac{1}{\prod_{t=1}^{k}(n_t+1)} \sum_{\substack{u_{\ell_i}=1 \\ \ell_i \in S}}^{n_{\ell_i}+1} \sum_{\substack{v_{j_z}=1 \\ j_z \in NS}}^{n_{j_z}+1} I\big(x_{\ell_2,u_{\ell_2}} < x_{\ell_1,u_{\ell_1}-1}\big) \times P\left(\max_{j_z \in NS} x_{j_z,v_{j_z}} < X^f_{\ell_3} < x_{\ell_2,u_{\ell_2}} \,\middle|\, X^f_{\ell_3} \in (x_{\ell_3,u_{\ell_3}-1}, x_{\ell_3,u_{\ell_3}})\right), \quad (21)$$

$$\overline{P}^{\max \ell_3}\left(\max_{j_z \in NS} X^f_{j_z} < X^f_{\ell_3} < X^f_{\ell_2} < X^f_{\ell_1}\right) = \frac{1}{\prod_{t=1}^{k}(n_t+1)} \sum_{\substack{u_{\ell_i}=1 \\ \ell_i \in S}}^{n_{\ell_i}+1} \sum_{\substack{v_{j_z}=1 \\ j_z \in NS}}^{n_{j_z}+1} I\big(x_{\ell_2,u_{\ell_2}} < x_{\ell_1,u_{\ell_1}}\big) \times P\left(\max_{j_z \in NS} x_{j_z,v_{j_z}-1} < X^f_{\ell_3} < x_{\ell_2,u_{\ell_2}} \,\middle|\, X^f_{\ell_3} \in (x_{\ell_3,u_{\ell_3}-1}, x_{\ell_3,u_{\ell_3}})\right). \quad (22)$$

For the lower probability, the mass $\frac{1}{n_{\ell_3}+1}$ is assigned to the sub-interval that minimises the probability within each interval $(x_{\ell_3,u_{\ell_3}-1}, x_{\ell_3,u_{\ell_3}})$; the combination yielding the smallest value is selected. For the upper probability, the assignment is chosen to maximise the corresponding expression.

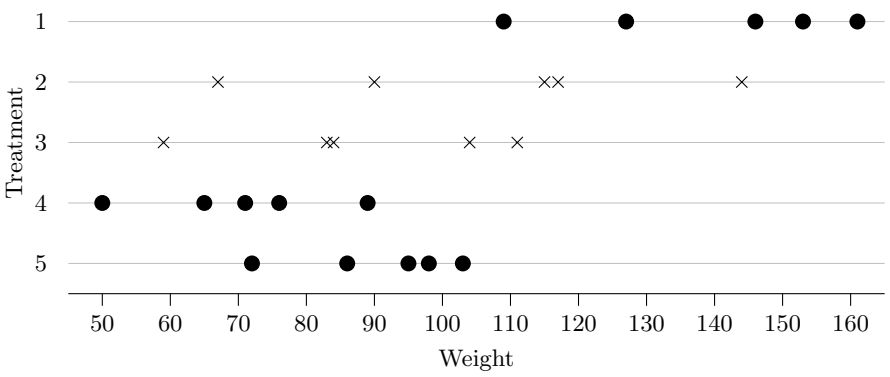

Figure 2: Fish weights across treatment levels for Example 3.

Optimisation over ℓ_2. Next, the same procedure is applied to group ℓ_2, now using the intervals $(x_{\ell_2,u_{\ell_2}-1}, x_{\ell_2,u_{\ell_2}})$. Hence, the lower and upper probabilities are

$$\underline{P}^{\min \ell_2}\left(\max_{j_z \in NS} X^f_{j_z} < X^f_{\ell_3} < X^f_{\ell_2} < X^f_{\ell_1}\right) = \frac{1}{\prod_{t=1}^{k}(n_t+1)} \sum_{\substack{u_{\ell_i}=1 \\ \ell_i \in S}}^{n_{\ell_i}+1} \sum_{\substack{v_{j_z}=1 \\ j_z \in NS}}^{n_{j_z}+1} I\left(\max_{j_z \in NS} x_{j_z,v_{j_z}} < x_{\ell_3,u_{\ell_3}}\right) \times$$
$$P\left(x_{\ell_3,u_{\ell_3}} < X^f_{\ell_2} < x_{\ell_1,u_{\ell_1}-1} \,\middle|\, X^f_{\ell_2} \in (x_{\ell_2,u_{\ell_2}-1}, x_{\ell_2,u_{\ell_2}})\right), \quad (23)$$

$$\overline{P}^{\max \ell_2}\left(\max_{j_z \in NS} X^f_{j_z} < X^f_{\ell_3} < X^f_{\ell_2} < X^f_{\ell_1}\right) = \frac{1}{\prod_{t=1}^{k}(n_t+1)} \sum_{\substack{u_{\ell_i}=1 \\ \ell_i \in S}}^{n_{\ell_i}+1} \sum_{\substack{v_{j_z}=1 \\ j_z \in NS}}^{n_{j_z}+1} I\left(\max_{j_z \in NS} x_{j_z,v_{j_z}-1} < x_{\ell_3,u_{\ell_3}}\right) \times$$
$$P\left(x_{\ell_3,u_{\ell_3}} < X^f_{\ell_2} < x_{\ell_1,u_{\ell_1}} \,\middle|\, X^f_{\ell_2} \in (x_{\ell_2,u_{\ell_2}-1}, x_{\ell_2,u_{\ell_2}})\right). \quad (24)$$

Finally, the approximate NPI lower probability is obtained by selecting the smaller of $\underline{P}^{\min \ell_2}$ and $\underline{P}^{\min \ell_3}$. Similarly, the approximate NPI upper probability is obtained by selecting the larger of $\overline{P}^{\max \ell_2}$ and $\overline{P}^{\max \ell_3}$. As in the general case, the lower and upper bounds given in Equations (11)–(14) with $w = 3$ provide a simpler computational alternative.

Example 3. We use a small subset of the `lengthWeight` data from the `StatCharme` package in R, consisting of fish weights recorded under five treatment levels. The observations in groups 1, 2, and 3 are $\{109, 127, 146, 153, 161\}$, $\{67, 90, 115, 117, 144\}$, and $\{59, 83, 84, 104, 111\}$, respectively. The remaining groups are group 4 with observations $\{50, 65, 71, 76, 89\}$ and group 5 with observations $\{72, 86, 95, 98, 103\}$. The data are shown in Figure 2.

The optimisation is carried out over the two middle groups, 2 and 3. We first optimise with respect to group 2 to obtain the approximate NPI lower probability. During this step, the probability mass of group 3 can be assigned to either endpoint of each interval, while the probability mass of group 2 must be placed within suitable sub-intervals of (x_{2,u_2-1}, x_{2,u_2}) so as to minimise Equation (23). For group 2, the six intervals

are $(-\infty, 67)$, $(67, 90)$, $(90, 115)$, $(115, 117)$, $(117, 144)$ and $(144, \infty)$.

In the first two intervals, no configurations satisfy the indicator condition, so the probability mass $1/(n_2+1)$ can be placed anywhere; we select representative points 66 and 75. For $(90, 115)$, the sub-interval $(90, 104)$ minimises the number of satisfying configurations (40), and the mass is placed at 97. For $(115, 117)$, 232 configurations satisfy the condition, and the mass is placed at 116. For $(117, 144)$, the sub-interval $(127, 144)$ yields the minimum (174 configurations), and the mass is placed at 135.5. In the final interval, again no configurations satisfy the condition in $(161, \infty)$, so the mass is placed at 170. This yields

$$\underline{P}^{\min 2}\Big(\max(X_4^f, X_5^f) < X_3^f < X_2^f < X_1^f\Big) = 0.0574.$$

To obtain $\overline{P}^{\max 2}$, the mass of group 2 is assigned to sub-intervals that maximise Equation (24), while the mass of group 3 remains at interval endpoints. The optimal sub-intervals are $(60, 67)$, $(85, 90)$, $(112, 115)$, $(115, 117)$, $(117, 126)$ and $(144, 145)$, with 12, 132, 470, 470, 470 and 376 satisfying configurations, respectively. Hence,

$$\overline{P}^{\max 2}\Big(\max(X_4^f, X_5^f) < X_3^f < X_2^f < X_1^f\Big) = 0.2482.$$

The same procedure is then applied to group 3, assigning its probability mass to sub-intervals that minimise and maximise Equations (21) and (22), respectively, while the mass of group 2 is placed at interval endpoints. This gives $\underline{P}^{\min 3} = 0.0518$ and $\overline{P}^{\max 3} = 0.2526$ for the same event. The heuristic approximation combines these results by taking $\underline{P} = \min\{\underline{P}^{\min 2}, \underline{P}^{\min 3}\} = 0.0518$ and $\overline{P} = \max\{\overline{P}^{\max 2}, \overline{P}^{\max 3}\} = 0.2526$.

A clear pattern emerges: sub-intervals minimising the lower probability tend to lie near the beginning of each interval (except for the last interval, where they lie near the end), whereas those maximising the upper probability tend to lie near the end (with the opposite behaviour in the final interval). Using the bounds from Section 4.1, we obtain $\underline{P}^L = 0.0318$, $\underline{P}^U = 0.0872$, $\overline{P}^L = 0.1876$ and $\overline{P}^U = 0.3645$, so that $0.0318 < 0.0518 < 0.0872$ and $0.1876 < 0.2526 < 0.3645$, as expected.

Repeating the same calculations for all possible ordered selections would lead to 60 cases. To reduce computation, we fix $NS = \{4, 5\}$ and consider only permutations of $S = \{1, 2, 3\}$, resulting in six ordered selections. The corresponding approximate NPI lower and upper probabilities are shown in Table 3. The ordering $(1, 2, 3)$ yields the largest approximate lower and upper probabilities (0.0518 and 0.2526) and is therefore most supported, whereas $(2, 3, 1)$ yields the smallest values (0.0000 and 0.0417). This is consistent with Figure 2, where groups 1 and 2 tend to have larger observations than the remaining groups.

ℓ_1	ℓ_2	ℓ_3	$\underline{P}^{\min \ell_2}$	$\overline{P}^{\max \ell_2}$	$\underline{P}^{\min \ell_3}$	$\overline{P}^{\max \ell_3}$	$\underline{P}$	$\overline{P}$
1	2	3	0.0574	0.2482	0.0518	0.2526	0.0518	0.2526
2	1	3	0.0075	0.1386	0.0073	0.1227	0.0073	0.1386
2	3	1	0.0000	0.0417	0.0000	0.0309	0.0000	0.0417
3	1	2	0.0000	0.0864	0.0000	0.0949	0.0000	0.0949
1	3	2	0.0051	0.0968	0.0000	0.0809	0.0000	0.0968
3	2	1	0.0000	0.0463	0.0000	0.0324	0.0000	0.0463

Table 3: Approximate NPI lower and upper probabilities for the event $\max(X_4^f, X_5^f) < X_{\ell_3}^f < X_{\ell_2}^f < X_{\ell_1}^f$ in Example 3.

4.4 Selecting a partially ranked subset

Selecting a partially ranked subset is relevant in situations where a complete ranking of all groups is neither meaningful nor necessary. Instead, interest may lie in identifying a subset of groups and ranking only a portion of them. For example, in drug development, researchers may select a subset of the most effective treatments and rank only the top treatments within that subset, while leaving the remaining selected treatments unranked.

This section introduces the notation and formulas for this setting. Let S denote a selected subset of groups, and let $S^* \subseteq S$ represent the partially ranked subset. Let $w = |S|$ be the total number of groups in S, and let $w^* = |S^*|$ be the number of ranked groups, where $1 \leq w^* \leq w$. The future observations corresponding to the ranked groups are denoted by $X_{\ell_{i^*}}^f$.

The objective is to compute the NPI lower and upper probabilities for the event

$$A^* = \left(\max_{j_z \in NS} X_{j_z}^f < \min_{\ell_i \in S \setminus S^*} X_{\ell_i}^f \right) \wedge \left(\max_{\ell_i \in S \setminus S^*} X_{\ell_i}^f < X_{\ell_{w^*}}^f < \ldots < X_{\ell_{1^*}}^f \right),$$

where $i^* \subset \{1, 2, \ldots, w^*\}$. Direct computation of the NPI lower and upper probabilities for this event is computationally demanding, as discussed in [15]. Therefore, bounds for the NPI lower and upper probabilities are considered, following the derivations presented in Section 4.1.

The lower bound for the NPI lower probability is

$$\begin{aligned}
\underline{P}^L(A^*) = \frac{1}{\prod_{t=1}^k (n_t + 1)} & \sum_{\substack{u_{\ell_{i^*}} = 1 \\ \ell_{i^*} \in S^*}}^{n_{\ell_{i^*}}+1} \sum_{\substack{u_{\ell_i} = 1 \\ \ell_i \in S \setminus S^*}}^{n_{\ell_i}+1} \sum_{\substack{v_{j_z} = 1 \\ j_z \in NS}}^{n_{j_z}+1} \\
& I\Bigg(\left(\max_{j_z \in NS} x_{j_z, v_{j_z}} < \min_{\ell_i \in S \setminus S^*} x_{\ell_i, u_{\ell_i} - 1} \right) \wedge \left(\max_{\ell_i \in S \setminus S^*} x_{\ell_i, u_{\ell_i}} < x_{\ell_{w^*}, u_{\ell_{w^*}} - 1} \right) \\
& \wedge \bigcap_{i=1}^{w^*-1} \left\{ x_{\ell_{w^*-i+1}, u_{\ell_{w^*-i+1}}} < x_{\ell_{w^*-i}, u_{\ell_{w^*-i}} - 1} \right\} \Bigg).
\end{aligned} \tag{25}$$

	$\delta^* = 0.2$	$\delta^* = 0.4$	$\delta^* = 0.6$
μ_1	0.0	0.0	0.0
μ_2	0.2	0.4	0.6
μ_3	0.4	0.8	1.2
μ_4	0.6	1.2	1.8

Table 4: Simulated mean values for four Normal populations corresponding to $\delta^* = 0.2$, 0.4, and 0.6.

The upper bound for the NPI upper probability is

$$\overline{P}^U(A^*) = \frac{1}{\prod_{t=1}^{k}(n_t+1)} \sum_{\substack{u_{\ell_{i^*}}=1 \\ \ell_{i^*}\in S^*}}^{n_{\ell_{i^*}}+1} \sum_{\substack{u_{\ell_i}=1 \\ \ell_i \in S\setminus S^*}}^{n_{\ell_i}+1} \sum_{\substack{v_{j_z}=1 \\ j_z \in NS}}^{n_{j_z}+1}$$
$$I\Bigg(\left(\max_{j_z\in NS} x_{j_z,v_{j_z}-1} < \min_{\ell_i\in S\setminus S^*} x_{\ell_i,u_{\ell_i}}\right) \wedge \bigcap_{i^*=1}^{w^*}\left(\max_{j_z\in NS} x_{j_z,v_{j_z}-1} < x_{\ell_{i^*},u_{\ell_{i^*}}}\right)$$
$$\wedge \bigcap_{i^*=1}^{w^*}\left(\min_{\ell_i\in S\setminus S^*} x_{\ell_i,u_{\ell_i}-1} < x_{\ell_{i^*},u_{\ell_{i^*}}}\right) \wedge \bigcap_{i^*=2}^{w^*}\bigcap_{a=1}^{i^*-1}\left(x_{\ell_{i^*},u_{\ell_{i^*}}-1} < x_{\ell_a,u_{\ell_a}}\right)\Bigg). \tag{26}$$

4.5 Comparison of selection methods

This section compares three methods for selecting a ranked subset of the two best populations: Bechhofer's indifference-zone procedure, Gupta's subset selection procedure, and the NPI method. The aim is to assess their behaviour for predictive inference and to examine whether they lead to the same selected subset and the same ranking within that subset.

A simulation study is conducted with $k = 4$ Normal populations and common standard deviation $\sigma = 1$. Three values of $P^* \in \{0.95, 0.90, 0.75\}$ are considered, together with three values of $\delta^* \in \{0.2, 0.4, 0.6\}$, where δ^* represents the minimum difference between the means of the best and second best populations under the least favourable configuration. The population means used in the simulation are given in Table 4. These values were chosen to ensure that the distance between the largest and second largest means equals the specified δ^*.

Using Bechhofer's sample size formula in (3), nine common sample sizes n are obtained and are reported in Table 5, together with the corresponding simulated sample means. For each case, the sample means are ranked as in (1), and the populations corresponding to $\bar{x}_{[4]}$ and $\bar{x}_{[3]}$ are taken as the best and second best, respectively. For example, in Case 1 ($P^* = 0.95$, $\delta^* = 0.2$), the required sample size is $n = 237$ and the sample means are $\bar{x}_1 = 0.1947$, $\bar{x}_2 = 0.2492$, $\bar{x}_3 = 0.3387$ and $\bar{x}_4 = 0.5208$. Hence populations 4 and 3 are selected as best and second best. Across the nine cases, the ranking is not fully consistent: populations 4 and 3 appear as the best two in Cases 1, 3, 4, 6, and 9, while populations 3 and 4 appear as the best two in Cases 2, 5, 7, and 8. This variation is due to Monte Carlo sampling variability, which is more pronounced

δ^*	P^*	n	$\bar{x}_1$	$\bar{x}_2$	$\bar{x}_3$	$\bar{x}_4$
0.2	0.95	237	0.1947	0.2492	0.3387	0.5208
0.2	0.90	174	0.2155	0.2899	0.5458	0.4013
0.2	0.75	90	0.1737	0.1180	0.2311	0.5620
0.4	0.95	59	0.02373	0.1064	0.4203	0.5523
0.4	0.90	43	−0.0770	−0.0784	0.4915	0.4367
0.4	0.75	22	−0.0618	0.0437	0.4417	0.5927
0.6	0.95	26	0.1086	0.2106	1.0626	1.0225
0.6	0.90	19	0.5011	0.1545	1.2626	0.9839
0.6	0.75	10	0.0952	−0.3054	0.7606	1.1078

Table 5: Sample sizes and corresponding sample means for four populations, based on simulated data from Normal distributions with standard deviation $\sigma = 1$, for different values of δ^* and P^* using Bechhofer's method.

δ^*	P^*	n	$\bar{x}_1$	$\bar{x}_2$	$\bar{x}_3$	$\bar{x}_4$
0.2	0.95	98	0.1574	0.1362	0.3554	0.7858
0.2	0.90	59	−0.0446	0.5234	0.1905	0.9711
0.4	0.95	25	0.0494	0.2908	1.1190	1.4147
0.4	0.90	15	−0.1052	0.4655	0.5221	0.5386
0.6	0.95	11	0.2052	0.5526	1.2051	1.7232
0.6	0.90	7	−0.2244	0.6951	0.4595	1.6051

Table 6: Sample sizes and corresponding sample means for four populations, based on simulated data from Normal distributions with standard deviation $\sigma = 1$, for different values of δ^* and P^* using Gupta's method.

for smaller sample sizes.

Gupta [23] proposed a procedure for selecting a subset that contains the best population, without ranking the selected populations. For the same parameter settings (P^*, δ^*, σ), the sample sizes obtained from (4) and the corresponding simulated sample means are reported in Table 6. The critical values λ are taken from [20]; these tables provide $\lambda = 2.9162$ for $P^* = 0.95$ and $\lambda = 2.4516$ for $P^* = 0.90$, while no value is available for $P^* = 0.75$. Thus, the selected subset is determined by applying inequality (5).

For $P^* = 0.95$, the selected subset is $\{4\}$ when $n = 98$, whereas $\{3, 4\}$ is selected for $n = 25$ and $n = 11$. For $P^* = 0.90$, the procedure selects $\{4\}$ when $n = 59$ and $n = 7$, and $\{2, 3, 4\}$ when $n = 15$. These results show that Gupta's procedure does not necessarily lead to the same selected subset as Bechhofer's method.

Next, the NPI method for selecting a ranked subset of the two best populations is applied to the same nine cases in Table 5, using the mean settings in Table 4. Tables 7–9 report the NPI lower and upper probabilities for all 12 possible ordered selections of two populations out of four for the event $\max_{j_z \in NS} X^f_{j_z} < X^f_{\ell_2} < X^f_{\ell_1}$.

Overall, the NPI probabilities are small, and the imprecision (the gap between lower and upper probabilities) increases as the sample size decreases. For example, when $n = 237$ (Case 1), the smallest lower probability is 0.0555 (selecting populations 1 then 2), and the largest upper probability is 0.1207 (selecting populations 4 then 3). When $n = 10$ (Case 9), the smallest lower probability decreases to 0.0022, while the largest upper probability increases to 0.2457, indicating greater imprecision for smaller samples. Across all

		Case 1		Case 2		Case 3	
ℓ_1	ℓ_2	$\underline{P}$	$\overline{P}$	$\underline{P}$	$\overline{P}$	$\underline{P}$	$\overline{P}$
1	2	0.0555	0.0577	0.0590	0.0621	0.0542	0.0598
2	1	0.0592	0.0615	0.0563	0.0593	0.0515	0.0571
1	3	0.0611	0.0635	0.0739	0.0778	0.0569	0.0630
3	1	0.0677	0.0701	0.0893	0.0930	0.0647	0.0710
1	4	0.0705	0.0733	0.0695	0.0730	0.0750	0.0830
4	1	0.0916	0.0945	0.0706	0.0740	0.1158	0.1242
2	3	0.0759	0.0786	0.0783	0.0823	0.0542	0.0602
3	2	0.0786	0.0812	0.0989	0.1030	0.0617	0.0677
2	4	0.0876	0.0906	0.0729	0.0765	0.0712	0.0790
4	2	0.1067	0.1098	0.0791	0.0828	0.1118	0.1199
3	4	0.0998	0.1031	0.1152	0.1198	0.0868	0.0952
4	3	0.1174	0.1207	0.0985	0.1029	0.1232	0.1320

Table 7: NPI lower and upper probabilities for all ordered selections of two populations for the event $\max_{j_z \in NS} X^f_{j_z} < X^f_{\ell_2} < X^f_{\ell_1}$, based on simulated data from Normal distributions with $\delta^* = 0.2$ and $\sigma = 1$. Case 1: $P^* = 0.95$, $n = 237$; Case 2: $P^* = 0.90$, $n = 174$; Case 3: $P^* = 0.75$, $n = 90$.

cases, the ordered selections $(\ell_1, \ell_2) = (4, 3)$ and $(3, 4)$ consistently achieve the largest (or second largest) NPI lower and upper probabilities, while selections involving populations 1 and 2 tend to yield the smallest probabilities. This aligns with the simulation design in Table 4, where populations 1 and 2 have smaller means than populations 3 and 4.

A key conceptual difference is that Bechhofer's procedure requires specifying P^* in advance and uses it to determine the design constant d and hence the required sample size. In contrast, NPI does not require a prespecified P^*; instead, it provides lower and upper probabilities for the event of interest directly from the observed data, thereby quantifying the uncertainty through imprecision.

Finally, Tables 10–12 report frequencies of the selected ordered pairs over 1000 simulation repetitions for each setting of (P^*, δ^*). These results confirm that populations 4 and 3 dominate as the best two populations across repetitions. For example, in Case 1 of Table 10, the ordered selection $(4, 3)$ occurs 859 times for the NPI lower probability and 850 times for the NPI upper probability. Moreover, variability in the selected combinations decreases as the sample size increases: smaller-sample cases show more dispersion across possible selections, whereas larger-sample cases concentrate more heavily on $(4, 3)$ (and, to a lesser extent, $(3, 4)$), with most other combinations occurring rarely or not at all.

In conclusion, for these simulated settings, Bechhofer's indifference-zone approach and the NPI method tend to select the same two populations and provide the same ranking within the selected subset. Gupta's method, which targets subset inclusion rather than ranking, does not always lead to the same selected subset as either Bechhofer's method or NPI.

		Case 4		Case 5		Case 6	
ℓ_1	ℓ_2	$\underline{P}$	$\overline{P}$	$\underline{P}$	$\overline{P}$	$\underline{P}$	$\overline{P}$
1	2	0.0526	0.0603	0.0228	0.0306	0.0259	0.0536
2	1	0.0285	0.0353	0.0337	0.0417	0.0368	0.0537
1	3	0.0640	0.0739	0.0431	0.0567	0.0326	0.0546
3	1	0.0541	0.0629	0.0817	0.0959	0.0642	0.0873
1	4	0.0602	0.0708	0.0391	0.0516	0.0479	0.0523
4	1	0.0730	0.0829	0.0672	0.0802	0.0949	0.1223
2	3	0.0687	0.0794	0.0666	0.0820	0.0459	0.0708
3	2	0.0847	0.0955	0.0915	0.1064	0.0623	0.0861
2	4	0.0606	0.0718	0.0621	0.0766	0.0403	0.0671
4	2	0.1045	0.1164	0.0674	0.0805	0.0970	0.1261
3	4	0.1058	0.1198	0.1457	0.1675	0.0786	0.1120
4	3	0.1343	0.1489	0.1319	0.1532	0.1154	0.1514

Table 8: NPI lower and upper probabilities for all ordered selections of two populations for the event $\max_{j_z \in NS} X^f_{j_z} < X^f_{\ell_2} < X^f_{\ell_1}$, based on simulated data from Normal distributions with $\delta^* = 0.4$ and $\sigma = 1$. Case 4: $P^* = 0.95$, $n = 59$; Case 5: $P^* = 0.90$, $n = 43$; Case 6: $P^* = 0.75$, $n = 22$.

5 NPI for generalised selection and ranking event

Ranking independent groups is extended here to the setting where groups are organised into *buckets*, by which we mean subsets of groups treated as indistinguishable in rank. Bucketing refers to the allocation of all groups into such subsets. This is useful when a complete strict ranking is unnecessary or impractical. At one extreme, assigning one bucket per group yields a full ranking [15]; at the other, placing all groups in a single bucket removes any ranking between them. The smallest meaningful intermediate case consists of two buckets, as studied by Coolen and van der Laan [13]. In this paper, we consider a fixed number of buckets chosen to suit the application. We first derive NPI lower and upper probabilities for general bucket-ranking events in Section 5.1. Section 5.2 discusses data-driven allocation of groups to buckets, for example based on median values. Section 5.3 applies the NPI-bootstrap method to estimate the probabilities of the ranking events, illustrating that these estimates typically lie between the corresponding NPI lower and upper probabilities. Finally, Section 5.4 addresses the choice of the number of buckets by evaluating possible allocations and selecting the configuration that maximises the NPI lower and upper probabilities of the event of interest. This provides a flexible and practical NPI framework for ranking independent groups organised into buckets.

5.1 Generalising the event of interest with buckets

Suppose there are k independent groups $X_1, X_2, \ldots, X_k$ with index set $K = \{1, 2, \ldots, k\}$. Let the number of buckets s be fixed. The groups are partitioned into s non-empty subsets $K_1, K_2, \ldots, K_s$ such that $K_i \cap K_j =$

		Case 7		Case 8		Case 9	
ℓ_1	ℓ_2	$\underline{P}$	$\overline{P}$	$\underline{P}$	$\overline{P}$	$\underline{P}$	$\overline{P}$
1	2	0.0109	0.0184	0.0089	0.0189	0.0022	0.0155
2	1	0.0088	0.0162	0.0079	0.0198	0.0064	0.0228
1	3	0.0370	0.0588	0.0546	0.0881	0.0265	0.0669
3	1	0.0536	0.0722	0.0750	0.1093	0.0265	0.0663
1	4	0.0343	0.0547	0.0370	0.0636	0.0212	0.0771
4	1	0.0504	0.0690	0.0425	0.0697	0.0766	0.1359
2	3	0.0238	0.0445	0.0269	0.0541	0.0141	0.0455
3	2	0.0783	0.1016	0.0606	0.0872	0.0152	0.0445
2	4	0.0316	0.0522	0.0205	0.0419	0.0113	0.0589
4	2	0.0660	0.0882	0.0283	0.0482	0.0367	0.0751
3	4	0.2026	0.2463	0.1779	0.2309	0.1073	0.1948
4	3	0.1726	0.2135	0.1580	0.2104	0.1546	0.2457

Table 9: NPI lower and upper probabilities for all ordered selections of two populations for the event $\max_{j_z \in NS} X^f_{j_z} < X^f_{\ell_2} < X^f_{\ell_1}$, based on simulated data from Normal distributions with $\delta^* = 0.6$ and $\sigma = 1$. Case 7: $P^* = 0.95$, $n = 26$; Case 8: $P^* = 0.90$, $n = 19$; Case 9: $P^* = 0.75$, $n = 10$.

$\emptyset$ for $i \neq j, \bigcup_{i=1}^{s} K_i = K, k_i = |K_i|$, with $\sum_{i=1}^{s} k_i = k$. Ignoring the ranking within buckets, the number of possible allocations equals

$$\frac{k!}{k_1!k_2!\ldots k_s!}.$$

Let $X^f_1, \ldots, X^f_k$ denote the next future observations from the k groups. The generalised ranking event of interest is

$$\{X^f_t\}_{t\in K_1} < \{X^f_t\}_{t\in K_2} < \cdots < \{X^f_t\}_{t\in K_s},$$

meaning that all future observations in bucket 1 are smaller than those in bucket 2, and so on. Equivalently,

$$A = \bigcap_{i=1}^{s-1} \left\{ \max_{t\in K_i} X^f_t < \min_{t\in K_{i+1}} X^f_t \right\}. \tag{27}$$

Deriving exact NPI lower and upper probabilities for A is computationally demanding. As noted by Coolen-Maturi [15] for the special case where each bucket contains a single group, exact calculation quickly becomes complex. Therefore, we derive a lower bound for the NPI lower probability and an upper bound for the NPI upper probability.

For each group $t = 1, \ldots, k$, let the observed data be ordered as $x_{t,1} < x_{t,2} < \cdots < x_{t,n_t}$, with $x_{t,0} = -\infty$ and $x_{t,n_t+1} = \infty$ for notational convenience. Inference is based on Hill's assumption $A_{(n_t)}$ for each group. The lower bound for the NPI lower probability and the upper bound for the NPI upper probability of event

		Case 1		Case 2		Case 3	
ℓ_1	ℓ_2	$\#\underline{P}$	$\#\overline{P}$	$\#\underline{P}$	$\#\overline{P}$	$\#\underline{P}$	$\#\overline{P}$
1	2	0	0	0	0	0	0
2	1	0	0	0	0	0	0
1	3	0	0	0	0	0	0
3	1	0	0	0	0	0	0
1	4	0	0	0	0	0	0
4	1	0	0	0	2	9	12
2	3	0	0	0	0	0	0
3	2	0	0	0	0	1	4
2	4	0	0	0	0	8	4
4	2	39	37	72	85	127	116
3	4	102	113	133	129	176	170
4	3	859	850	795	784	679	694

Table 10: Frequencies of the NPI lower and upper probabilities for all ordered selections of two populations for the event $\max_{j_z \in NS} X^f_{j_z} < X^f_{\ell_2} < X^f_{\ell_1}$, based on 1000 simulation repetitions with $\delta^* = 0.2$ and $\sigma = 1$. Case 1: $P^* = 0.95$, $n = 237$; Case 2: $P^* = 0.90$, $n = 174$; Case 3: $P^* = 0.75$, $n = 90$.

A are given by

$$\underline{P}^L(A) = \frac{1}{\prod_{t=1}^{k}(n_t+1)} \sum_{l_1=1}^{n_1+1} \cdots \sum_{l_k=1}^{n_k+1} \prod_{i=1}^{s-1} I\left(\max_{t\in K_i} x_{t,l_t} < \min_{t\in K_{i+1}} x_{t,l_t-1}\right), \tag{28}$$

$$\overline{P}^U(A) = \frac{1}{\prod_{t=1}^{k}(n_t+1)} \sum_{l_1=1}^{n_1+1} \cdots \sum_{l_k=1}^{n_k+1} \prod_{i=1}^{s-1} \prod_{a=i}^{s} I\left(\max_{t\in K_i} x_{t,l_t-1} < \min_{t\in K_{a+1}} x_{t,l_t}\right). \tag{29}$$

For the lower bound, complete separation between buckets is required: only configurations are counted in which all intervals corresponding to bucket i lie entirely to the left of those in bucket $i+1$ for all $i = 1, \ldots, s-1$. For the upper bound, configurations are counted whenever the interval arrangement is compatible with the required ordering across buckets.

These expressions provide computationally feasible approximations to the NPI lower and upper probabilities for the general bucket-ranking event.

5.2 Assigning groups to buckets for defining events of interest

To address the allocation of groups to buckets and the definition of the corresponding event of interest, data-driven summary statistics can be used to rank the groups. A natural approach is to order the groups according to a location measure, such as the median, and assign them to buckets accordingly. For instance, groups with the smallest medians may be placed in the first bucket, those with the next-smallest medians in the second bucket, and so on.

Let m_t denote the median of group t, for $t = 1, \ldots, k$, and let $m_{(1)} < m_{(2)} < \cdots < m_{(k)}$ denote the ordered medians. Under this ordering, the group corresponding to the largest median $m_{(k)}$ may be regarded

		Case 4		Case 5		Case 6	
ℓ_1	ℓ_2	$\#\underline{P}$	$\#\overline{P}$	$\#\underline{P}$	$\#\overline{P}$	$\#\underline{P}$	$\#\overline{P}$
1	2	0	0	0	0	0	0
2	1	0	0	0	0	0	0
1	3	0	0	0	0	0	0
3	1	0	0	0	0	0	0
1	4	0	0	0	0	0	1
4	1	0	0	3	0	2	5
2	3	0	0	0	0	0	0
3	2	0	0	0	0	0	0
2	4	0	0	1	0	0	1
4	2	37	36	62	54	99	102
3	4	59	72	102	119	155	176
4	3	904	892	832	827	744	715

Table 11: Frequencies of the NPI lower and upper probabilities for all ordered selections of two populations for the event $\max_{j_z \in NS} X^f_{j_z} < X^f_{\ell_2} < X^f_{\ell_1}$, based on 1000 simulation repetitions with $\delta^* = 0.4$ and $\sigma = 1$. Case 4: $P^* = 0.95$, $n = 59$; Case 5: $P^* = 0.90$, $n = 43$; Case 6: $P^* = 0.75$, $n = 22$.

as the best group, the group corresponding to $m_{(k-1)}$ as the second-best, and so on. Buckets are then formed by grouping consecutive ordered medians according to the user's objective and the chosen number of buckets. Although intuitive and simple, allocation based on medians does not necessarily maximise the lower and upper bounds for the NPI lower and upper probabilities of the event of interest. The following example illustrates this point.

Example 4. The data from Hansen [27] concern fasting serum growth hormone levels (ng/ml) after exercise in individuals with maturity-onset diabetes. Two groups are considered: middle-aged individuals (X) and elderly individuals (Y). The medians are $m_X = 1.1$ and $m_Y = 1.4$, so $m_X < m_Y$. Based on this ordering, a natural event of interest is $X^f < Y^f$, representing a smaller future hormone level for a middle-aged individual than for an elderly individual.

However, the NPI analysis shows that the complementary event $X^f \geq Y^f$ yields slightly larger lower and upper probabilities. Specifically, $\underline{P}^L(X^f \geq Y^f) = 0.4692$ and $\overline{P}^U(X^f \geq Y^f) = 0.5151$, compared to $\underline{P}^L(X^f < Y^f) = 0.4599$ and $\overline{P}^U(X^f < Y^f) = 0.5058$. Although the differences are small, and the boxplots suggest somewhat higher central tendency for group Y, the results illustrate that defining the event of interest solely according to median ordering does not necessarily correspond to the event with the largest NPI lower and upper probabilities.

5.3 NPI-Bootstrap for estimating the probability of the event

As discussed in Section 5.1, deriving exact NPI lower and upper probabilities for the general event of interest is computationally demanding, and bounds were therefore introduced as practical approximations. In this

		Case 7		Case 8		Case 9	
ℓ_1	ℓ_2	$\#\underline{P}$	$\#\overline{P}$	$\#\underline{P}$	$\#\overline{P}$	$\#\underline{P}$	$\#\overline{P}$
1	2	0	0	0	0	0	0
2	1	0	0	0	0	0	0
1	3	0	0	0	0	0	0
3	1	0	0	0	0	0	0
1	4	0	0	0	0	0	0
4	1	0	1	0	0	4	4
2	3	0	0	0	0	0	0
3	2	0	0	0	0	1	0
2	4	0	0	0	1	5	5
4	2	26	23	49	46	95	81
3	4	44	44	66	71	128	124
4	3	930	932	885	882	767	786

Table 12: Frequencies of the NPI lower and upper probabilities for all ordered selections of two populations for the event $\max_{j_z \in NS} X^f_{j_z} < X^f_{\ell_2} < X^f_{\ell_1}$, based on 1000 simulation repetitions with $\delta^* = 0.6$ and $\sigma = 1$. Case 7: $P^* = 0.95$, $n = 26$; Case 8: $P^* = 0.90$, $n = 19$; Case 9: $P^* = 0.75$, $n = 10$.

section, we consider a different approach based on the NPI-Bootstrap (NPI-B) method proposed by Coolen and BinHimd [10]. Rather than approximating the NPI lower and upper probabilities, this method is used to estimate the probability of the event itself.

The NPI-B method generates future observations by sampling from the $n+1$ intervals induced by the ordered data $x_1 < \cdots < x_n$. These intervals are $(-\infty, x_1), (x_1, x_2), \ldots, (x_{n-1}, x_n), (x_n, \infty)$. One interval is selected at random with probability $1/(n+1)$. If a middle interval is selected, a future observation is sampled uniformly from that interval. If the first or last interval is selected, sampling is performed from the normal tails, with mean $\mu = \frac{x_1 + x_n}{2}$ and standard deviation $\sigma = \frac{x_n - \mu}{\Phi^{-1}\left(\frac{n}{n+1}\right)}$, where Φ denotes the standard Normal cumulative distribution function.

In this paper, only one future observation per group is considered. In this case, all possible orderings of the new observation among the n past observations are equally likely, each with probability $1/(n+1)$. The above procedure is repeated independently for each of the k groups, and the resulting bootstrap observations are used to evaluate the event

$$A = \bigcap_{i=1}^{s-1} \left\{ \max_{t \in K_i} X^f_t < \min_{t \in K_{i+1}} X^f_t \right\}.$$

Let $x^*_{t,l_t} \in (x_{t,l_t-1}, x_{t,l_t})$ denote the bootstrap-sampled observation for group t, where sampling across the (n_t+1) intervals is performed with equal probability $1/(n_t+1)$. The NPI-B estimate of the probability of A is then given by

$$P_B(A) = \frac{1}{\prod_{t=1}^{k}(n_t+1)} \sum_{l_1=1}^{n_1+1} \cdots \sum_{l_k=1}^{n_k+1} \prod_{i=1}^{s-1} I\left(\max_{t \in K_i} x^*_{t,l_t} < \min_{t \in K_{i+1}} x^*_{t,l_t} \right). \tag{30}$$

In practice, the procedure is repeated B times to obtain bootstrap replications; in this paper, $B = 1000$ is used, which is common in the literature.

In the following example, the NPI-B estimate $P_B(A)$ is compared with the lower and upper bounds derived in Equations (28) and (29). As will be illustrated, the bootstrap estimate typically lies between the corresponding NPI lower and upper bounds.

Example 5. The dataset is taken from Simpson and Margolin [39]. It records revertant-colony counts from three replicate Ames assays on Salmonella TA98 at six dose levels of Acid Red 114. For each dose level, nine visible revertant colonies were observed.

The six doses are first ordered according to their medians: $m_1 = 22,\ m_2 = 25,\ m_3 = 35,\ m_4 = 43,\ m_5 = 28,\ m_6 = 16$, which gives the ranked order $m_6 < m_1 < m_2 < m_5 < m_3 < m_4$. We consider the case $k = 6$ doses divided into $s = 3$ buckets with $k_1 = k_2 = k_3 = 2$, allocating the two smallest medians to bucket 1, the next two to bucket 2, and the largest two to bucket 3. The event of interest is therefore

$$A = \left\{\max(X_6^f, X_1^f) < \min(X_2^f, X_5^f)\right\} \wedge \left\{\max(X_2^f, X_5^f) < \min(X_3^f, X_4^f)\right\},$$

where X_t^f denotes the future observation for dose t.

Using Equations (28) and (29), the lower and upper bounds for the NPI lower and upper probabilities are $\underline{P}^L(A) = 0.1213$ and $\overline{P}^U(A) = 0.2168$. To estimate the probability of A using the NPI-bootstrap method, one interval is randomly selected for each dose, and a future observation is sampled either uniformly from the interior intervals or from Normal tails for $(-\infty, 8)$ and $(98, \infty)$, where 8 and 98 are the observed data endpoints. Repeating this procedure 1000 times yields the bootstrap estimate $P_B(A) = 0.1500$. As expected, the bootstrap estimate lies between the lower and upper bounds, namely $0.1213 \leq 0.1500 \leq 0.2168$.

5.4 Selecting the number of buckets

In Section 5.1, the number of buckets was assumed to be fixed. In practice, however, determining the number of buckets is itself an important decision. Different allocations of groups to buckets lead to different ranking events and, consequently, different NPI lower and upper probabilities.

One possible approach is to evaluate all feasible group-to-bucket allocations for a given number of buckets and select the configuration that maximises the lower and upper bounds of the NPI lower and upper probabilities for the event of interest. Alternatively, data-driven summaries such as medians may be used to guide the allocation, for example, by placing groups with smaller medians in lower-ranked buckets and those with larger medians in higher-ranked buckets. The final choice of allocation may depend on the specific objective and application.

Groups	Data									
1	49	50	53	57	59	70	71	78	82	85
2	67	75	90	94	97	114	127	130	140	151
3	60	63	66	81	98	110	128	133	144	157
4	89	91	100	104	118	120	123	137	147	149

Table 13: Observed values for four groups with ten observations each, Example 6.

s	Best supported event	$\underline{P}^L$	$\overline{P}^U$	P_B
2	$X_1^f < \min(X_2^f, X_3^f, X_4^f)$	0.5314	0.8144	0.6870
3	$(X_1^f < X_3^f) \wedge (X_3^f < \min(X_2^f, X_4^f))$	0.2056	0.3080	0.2620
4	$X_1^f < X_2^f < X_4^f < X_3^f$	0.1063	0.1578	0.1320

Table 14: Most strongly supported event for each number of buckets s, Example 6.

When the number of groups and buckets is large, examining all possible allocations quickly becomes computationally demanding. However, for a small number of groups, a complete evaluation of all possible bucket configurations is feasible and provides insight into how the number of buckets and the distribution of groups across buckets affect the resulting ranking probabilities.

To illustrate this, the next example considers four groups. In this case, it is manageable to evaluate all possible bucket configurations with $s = 2, 3$, and 4, allowing investigation of how the number of buckets and the allocation of groups within them influence the selection and ranking results.

Example 6. Neuhauser et al. [36] present four groups with ten observations each; the data are shown in Table 13. Inspection of the data reveals a clear separation between groups 1 and 4, as all observations in group 1 are smaller than those in group 4. For $k = 4$ groups, the number of ordered partitions into s non-empty buckets equals $s!\,S(4, s)$, where $S(4, s)$ denotes the Stirling number of the second kind [21]. This yields 14 ordered partitions for $s = 2$, 36 for $s = 3$, and 24 for $s = 4$.

The complete enumeration of all bucket configurations and corresponding NPI lower and upper probabilities is provided in the first author's PhD thesis [1]. Here we report only the most strongly supported event for each value of s. For every event considered, the bootstrap estimate P_B lies between (or coincides with) the corresponding NPI lower and upper bounds, confirming the expected ordering $\underline{P}^L \leq P_B \leq \overline{P}^U$.

Table 14 shows that the strongest support occurs for $s = 2$, separating group 1 from the remaining groups. Increasing the number of buckets reduces both the lower and upper probabilities substantially. In particular, under $s = 3$ and $s = 4$, even the most favourable events receive markedly smaller support.

These results indicate that, for this dataset, a two-bucket structure provides the most coherent and strongly supported ranking according to the NPI lower and upper probabilities.

6 Conclusion

This paper presented Nonparametric Predictive Inference methods for selection and ranking based on future observations, leading to exact and approximate NPI lower and upper probabilities. For the selection of a ranked subset of the two best groups, an optimisation procedure was developed, and exact NPI lower and upper probabilities were derived. Because only one group lies between the two extremes, the optimisation problem is relatively simple, and exact probabilities can be obtained even for large and overlapping datasets. Lower and upper bounds were also derived to reduce computational effort. A real-valued dataset illustrated that the exact NPI lower and upper probabilities lie within their respective bounds, with the exact lower probability close to its lower bound and the exact upper probability close to its upper bound.

A simulation study further examined the method under ordered means. The results support selecting the group with the largest mean as best and those with the smallest means as worst. In addition, structural properties were identified that help reduce computational burden. Extreme sub-intervals often determine the minimum NPI lower probability and the maximum NPI upper probability, while for intermediate intervals, the extremal sub-intervals typically dominate. These findings streamline the optimisation process.

When the ranked subset contains three or more best groups, simultaneous optimisation over all groups becomes computationally demanding. A heuristic approach was therefore proposed, where optimisation is carried out separately for each middle group, yielding approximate NPI lower and upper probabilities. The approximate lower probability is defined as the minimum over the middle groups, and the approximate upper probability as the corresponding maximum. Bounds were again derived to further simplify computation.

The methodology was extended to a general selection and ranking event involving buckets of independent groups. Each group is assigned to exactly one bucket, and the focus is on ranking the buckets rather than the groups within them. For a predetermined number of buckets, both estimated probabilities and NPI lower and upper probabilities were obtained.

Two key questions were addressed: how to allocate groups to buckets and how many buckets to use. Allocation strategies based on medians were investigated, motivated by their common role in nonparametric analysis, but ordering groups by median does not necessarily maximise the NPI lower and upper probabilities. To determine the optimal number of buckets, all possible allocations can, in principle, be evaluated to identify the one that maximises the NPI lower and upper probabilities. Although exhaustive evaluation becomes challenging for large problems, it is feasible for moderate numbers of groups. In the illustrative example, the optimal allocation consisted of two buckets, with one group in the first and three in the second, yielding the highest NPI lower and upper probabilities. A bootstrap approach was used to obtain an estimated

probability for the general ranking event. In all examples, this estimate lay between the NPI lower and upper probabilities, as expected.

The NPI approach for ranking and selection is an attractive alternative to classical approaches as the results are easy to interpret and fully in line with corresponding results based on empirical distributions, in the sense that the NPI lower and upper probabilities for an event are bounds for the corresponding empirical probability. The explicit focus on a future observation from each group is also likely to be intuitively more logical than focus on characteristics of assumed underlying populations, as is typically done in established approaches. In addition, this facilitates the use of loss functions directly related to future observations, to generalize the presented approach [1]. The difference between the corresponding NPI upper and lower probabilities reflects the amount of data available, which is an attractive feature from the perspective of robustness, in particular for relatively small data sets. Implementing the approach is conceptually straightforward, as it is based on relatively straightforward combinatorics and optimisation. In case the number of groups or the number of observations per group becomes large, approximations can be achieved by applying the NPI bootstrap method [10].

Overall, the results demonstrate that NPI provides a coherent and flexible fully nonparametric framework for selection and ranking problems, offering informative lower and upper probability assessments for both subset selection and bucket ranking.

Acknowledgement(s)

We dedicate this paper to our friend Professor Sat Gupta, in appreciation of his exemplary activities in support of the international academic Statistics community.
This research was conducted as part of the first author's PhD at Durham University, supported by a studentship from the University of Jeddah, Saudi Arabia, and the Saudi Cultural Bureau in London.

Disclosure Statement

The authors declare that there are no conflicts of interest regarding the publication of this paper.

Funding

No funding was received.

References

[1] Alshihry, F. I. (2025). *Nonparametric Predictive Inference for Selection and Ranking*. Phd thesis, Durham University.

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

[3] Baker, R. (2010). *Multinomial Nonparametric predictive inference: Selection, Classification and Subcategory Data*. PhD thesis, Durham University. UK. `http://etheses.dur.ac.uk/257/`.

[4] Bechhofer, R.E. (1954). A single-sample multiple decision procedure for ranking means of Normal populations with known variances. *The Annals of Mathematical Statistics*, 25: 16–39.

[5] Coolen, F.P.A. (1996). Comparing two populations based on low stochastic structure assumptions. *Statistics and Probability Letters*, 29: 297–305.

[6] Coolen, F.P.A. (1998). Low structure imprecise predictive inference for Bayes' problem. *Statistics and Probability Letters*, 36: 349–357.

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

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

[9] Coolen, F.P.A. and Augustin, T. (2005). Learning from multinomial data: a nonparametric predictive alternative to the Imprecise Dirichlet Model. In *International Symposium on Imprecise Probabilities and Their Applications*, volume 5, pages 125–134.

[10] Coolen, F.P.A. and BinHimd, S. (2020). Nonparametric predictive inference bootstrap with application to reproducibility of the two-sample Kolmogorov-Smirnov test. *Journal of Statistical Theory and Practice*, 14.

[11] Coolen, F.P.A. and Coolen-Maturi, T. and Alqifari, H.N. (2018). Nonparametric predictive inference for future order statistics. *Communications in Statistics: Theory and methods*, 47(10): 2527–2548.

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

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

[14] 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.

[15] Coolen-Maturi, T. (2022). The ordering of future observations from multiple groups. *Communications in Statistics: Simulation and Computation*, 51: 7526–7543.

[16] De Finetti, B. (1974). *Theory of Probability*. Wiley, London.

[17] Desu, M.M. and Sobel, M. (1968). A fixed subset-size approach to the selection problem. *Biometrika*, 55: 401–410.

[18] Desu, M.M. and Sobel, M. (1971). Nonparametric procedures for selecting fixed-size subsets. In Gupta, S. and Yackel, J., editors, *Statistical Decision Theory and Related Topics*, pages 359–376. Academic Press, New York.

[19] Dudewicz, E.J. and Dalal, S.R. (1975). Allocation of observations in ranking and selection with unequal variances. *Sankhyā: The Indian Journal of Statistics, Series B*, 37: 28–78.

[20] Gibbons, J.D. and Olkin, I. and Sobel, M. (1999). *Selecting and Ordering Populations: A New Statistical Methodology*. Society for Industrial and Applied Mathematics, Philadelphia.

[21] Graham, R., Knuth, D., and Patashnik, O. (1994). *Concrete Mathematics: A Foundation for Computer Science*. Addison-Wesley, Reading, Massachusetts.

[22] Gupta, S.S. (1956). *On a Decision Rule for a Problem in Ranking Means*. PhD thesis, University of North Carolina, USA.

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

[24] Gupta, S.S. and Huyett, M.J. and Sobel, M. (1957). Selection and ranking with Binomial populations. *Transactions of American Society for Quality Control*, 9: 635–644.

[25] Gupta, S.S. and Sobel, M. (1957). On a statistic which arises in selection and ranking problems. *The Annals of Mathematical Statistics*, 28: 957–967.

[26] Gupta, S.S. and Sobel, M. (1960). Selecting a subset containing the best of several binomial populations. In Olkin, I., Hotelling, H., and Wolfowitz, J., editors, *Contributions to Probability and Statistics*, pages 389–413. Stanford University Press, Stanford, CA.

[27] Hansen, A.P. (1973). Abnormal serum growth hormone response to exercise in maturity-onset diabetics. *Diabetes*, 22: 619–628.

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

[29] Hill, B.M. (1988). De Finetti's theorem, induction, and A_n or Bayesian nonparametric predictive inference (with discussion). In Bernardo, J., DeGroot, M., Lindley, D., and Smith, A., editors, *Bayesian Statistics*, volume 3, pages 211–241. Oxford University Press.

[30] Hill, B.M. (1993). Parametric models for A_n: Splitting processes and mixtures. *Journal of the Royal Statistical Society*, 55: 423–433.

[31] Kumar, N. (2011a). A general class of subset selection procedures. *Journal of Combinatorics, Information and System Sciences*, 36: 135–158.

[32] Kumar, N. (2011b). Nonparametric subset selection procedure for location parameters. *Model Assisted Statistics and Applications*, 6: 337–347.

[33] Mahamunulu, D.M. (1967). Some fixed-sample ranking and selection problems. *The Annals of Mathematical Statistics*, 38: 1079–1091.

[34] Maturi, T. (2010). *Nonparametric predictive inference for Multiple Comparisons*. PhD thesis, Durham University, UK. `http://etheses.dur.ac.uk/230/`.

[35] Maturi, T. and Coolen-Schrijner, P. and Coolen, F.P.A. (2009). Nonparametric predictive pairwise comparison for real-valued data with terminated tails. *International Journal of Approximate Reasoning*, 51: 141–150.

[36] Neuhauser, M. and Liu, P.Y. and Hothorn, L. A. (1998). Nonparametric tests for trend: Jonckheere's test, a modification and a maximum test. *Biometrical Journal*, 40: 899–909.

[37] Paulson, E. (1949). A multiple decision procedure for certain problems in the analysis of variance. *The Annals of Mathematical Statistics*, 20: 95–98.

[38] Rizvi, M.H. and Sobel, M. (1967). Nonparametric procedures for selecting a subset containing the population with the largest q-quantile. *The Annals of Mathematical Statistics*, 38: 1788–1803.

[39] Simpson, D.G. and Margolin, B.H. (1986). Recursive nonparametric testing for dose-response relaflonships subject to downturns at high doses. *Biometrika*, 73: 589–596.

[40] Sobel, M. (1967). Nonparametric procedures for selecting the t population with the largest q-quantiles. *The Annals of Mathematical Statistics*, 38: 1804–1816.

[41] Sobel, M. (1969). Selecting a subset containing at least one of the t best populations. In Krishnaiah, P. R., editor, *Multivariate Analysis II*, pages 411–428. Academic Press, New York.

[42] Tong, Y.L. (1969). On partitioning a set of Normal populations by their locations with respect to a control using a single-stage, a two-stage and a sequential procedure. *The Annals of Mathematical Statistics*, 40: 1300–1324.

[43] Walley, P. (1991). *Statistical Reasoning with Imprecise Probabilities.* Chapman and Hall, London.

[44] Weichselberger, K. (2000). The theory of interval-probability as a unifying concept for uncertainty. *International Journal of Approximate Reasoning*, 24: 149–170.