EINDHOVEN UNIVERSITY OF TECHNOLOGY
Department of Mathematics and Computing Science

Master's Thesis

# Combining two classical approaches for statistical selection

by

**J.H.M. Verheijen**

April, 1996

Supervisors: Prof. dr. P. van der Laan (Eindhoven University of Technology)
Dr. ir. F.P.A. Coolen (University of Durham)

# Summary

The normal location problem, selecting the population with the largest mean, has been investigated from the 1950's on, initiated by Bechhofer and Gupta. They published two procedures, now considered as standard approaches. In this Master's Thesis we look again at the location problem, assuming normally indepent distributed populations with common known variances and propose a procedure that can be regarded as a generalization of the former two procedures. We combine both of them, by introducing a so-called preference threshold. In fact, the procedure we propose is a subset selection approach with an additional requirement, related to Bechhofer's approach. Whereas Bechhofer selects only the population corresponding to the largest sample sum, we determine a distance $c$ and sample size $n$ and select all populations in a selected subset of populations that have sample sums within this distance from the largest sample sum. We believe this is a procedure more in accordance with intuition: in Bechhofer's method we choose only the population with the largest sample sum, also if the difference to the second-largest is only very small, in our procedure we select a single population as the best one, only if its sample sum has a difference of at least $c$ to the second-largest. The values for $c$ and $n$ are determined using the two requirements; selecting the actual best population in the subset with at least a certain probability and selecting the actual best population alone with a certain probability. In the thesis we consider several aspects of the proposed procedure. We generalize the procedure for the selection of $t$ best populations, we consider the expected subset size and pay attention to the probability of a correct selection given that one population is selected. Also, departures from the assumption of common known variance is investigated and a Bayesian data analysis method is studied. Simulation studies illustrate features of the procedure and they show good results for the preference threshold procedure.

# Samenvatting

De selectie van de populatie met het grootste gemiddelde uit een aantal normaal verdeelde populaties is onderzocht sinds 1950, aangezet door het werk van Bechhofer en Gupta. Zij publiceerden twee procedures die nu bekend zijn als standaard aanpakken. In mijn afstudeerverslag bekijken we opnieuw dit zogenoemde locatie-probleem, onder de aanname van normaal onafhankelijk verdeelde populaties met dezelfde bekende variantie en we stellen een procedure voor die als een generalisatie van de twee standaard procedures kan worden beschouwd. We combineren beide door middel van de introductie van een zogenoemde voorkeursdrempel. In feite is de procedure die wij voorstellen een 'subset selection' benadering met een extra eis, die gerelateerd is aan Bechhofer's methode. Terwijl Bechhofer alleen de populatie selecteert die met de grootste steekproefsom overeenkomt, bepalen wij een afstand $c$ en steekproefgrootte $n$ en selecteren alle populaties in de geselecteerde deelverzameling die een steekproefsom hebben binnen deze afstand $c$ van de grootste steekproefsom. Volgens ons is deze procedure meer in overeenstemming met de intuïtie: in Bechhofer's procedure kiezen we alleen de populatie met de grootste steekproefsom, ook als het verschil met de tweede grootste maar heel klein is, in onze procedure selecteren we slechts een enkele populatie als de enige beste als zijn steekproefsom ten minste een verschil heeft van $c$ tot de tweede grootste steekproefsom. De waarden van $n$ en $c$ worden bepaald aan de hand van twee eisen; de selectie van de echte beste populatie in de geselecteerde deelverzameling met ten minste een bepaalde kans en de selectie van de echte beste populatie alleen met ten minste een bepaalde kans. In het afstudeerverslag bekijken we verschillende aspecten van de voorgestelde procedure. We generaliseren de procedure tot de selectie van $t$ beste populaties, we bekijken de verwachte grootte van de deelverzameling en we besteden aandacht aan de kans op een correcte selectie gegeven dat er één populatie is geselecteerd. Ook onderzoeken we afwijkingen van de aanname van dezelfde bekende variantie en we bestuderen een Bayesian data analyse methode. Simulaties illustreren bepaalde eigenschappen van de procedure en ze laten goede resultaten zien voor de 'preference threshold procedure'.

# Contents

# Chapter 1

# Introduction

Choosing among several alternatives is a problem which occurs often in practical situtations. Of course, we are usually interested in selecting the best alternative, in some well-defined sense of 'best'. For example, if we have $k$ different kinds of drugs to cure a certain disease, we want to choose the most effective one (e.g. the one that cures in the shortest period of time) and this can be measured by a random variable that can be modelled by some parameter $\theta_j$ for the $j$th drug. Generalized, we have $k$ ($k \geq 2$) populations that are modelled by some parameter $\theta$ and we want to choose the best one of them. An often used classical procedure is to test the null-hypothesis that all $k$ parameters have the same value (the homogeneity hypothesis). In practice, this is a very unrealistic situation. The alternative is usually that the $k$ populations do not all have equal $\theta$ values, which is very likely to be the actual situation. A researcher is usually not satisfied when rejecting $H_0$, but, for example, wants to know which population can be considered to be the best. For this purpose, ranking and selection procedures are specifically designed.

In this thesis we will regard the problem of selecting the best population out of $k$ independent normally distributed populations with unknown means and common known variance. The 'best' population is defined to be the one having the largest mean. Particularly Bechhofer [1] and Gupta [18] (before this publication already in his Ph.D. thesis [17]) have made valuable contributions to the development of procedures to tackle the selection problem. Their works can be considered as two basic approaches.

Bechhofer introduced the concept of an indifference and preference zone, arguing that if the best means of the $k$ populations are close to each other we are indifferent about which one we choose and if the difference between the best mean and second-best mean is at least a certain value $\delta^*(> 0)$ (a situation referred to as being in the preference zone), we want to make a **Correct Selection** (selecting the best population) with at probability at least $P^*$. This method is usually referred to as the **indifference zone selection** approach. In order to select the best population, Bechhofer's procedure consists of taking a sample of size $n$ of all $k$ populations, calculating the $k$ sample means and selecting the one with the largest sample mean. The observations are assumed to be distributed like $N(\mu_i, \sigma^2)$. This assumption is made to prevent the calculations from being too complex, however also different distributed observations have been studied. It should be noted here that the (common) sample size $n$ depends on $\delta^*$, $k$, $\sigma^2$ and $P^*$, where $\sigma^2$ is the common variance for the populations. The greater the value for $P^*$, the larger $n$ must be to satisfy the probability requirement. Bechhofer's approach explicitly aims at determining sample size $n$, so it is useful in the design stage of a selection experiment.

Gupta's procedure [18] is known as the **subset selection** approach. Gupta proposed to select a subset of the $k$ populations, as small as possible, that contains the best population

with probability of at least $P^*$. The procedure is to calculate the sample means and select all those populations in the subset that are within a certain distance from the maximum of the sample means. A major difference with Bechhofer's procedure is that Gupta can apply his selection criterion after the sampling has been done, there is no requirement for the sample size $n$, although you usually have to pay for small sample sizes with large selected subsets.

In this thesis we aim at a generalization of the existing procedures by integrating the two just introduced basic approaches. The design of the procedure should aim at specifying a minimum required sample size $n$ to achieve certain requirements, corresponding to the standard indifference zone approach. The assumption Bechhofer makes by dividing the parameter space into a preference and indifference zone seems acceptable, for if the $k$ populations all have the same value there isn't any selection problem, and so it is reasonable to assume a minimum difference $\delta^*(> 0)$ between the two largest means. For this reason we will hold on to this assumption and also assume this for the part of our procedure which corresponds to Gupta's approach although in the standard subset approach all means are allowed to be equal. The choice of a reasonable value for the minimum difference $\delta^*$ remains a problem. For this choice we have to rely on expert opinions and other information.

However, for the standard indifference zone procedure another unsatisfactory feature arises. Let us suppose we meet an owner of a chicken farm who wants us to select, between four different strains, the one with the largest egg production. We take samples according to a specified plan following Bechhofer's approach and we find values for the sample means like: the first strain results in 250.3 eggs per pullet housed, the second 254.7, the third 278.4 and the fourth 278.1. Now, using Bechhofer's procedure we would select population (strain) three as being the best one, ignoring the fact that population four produced only a slightly smaller sample mean. It will be a though job to convince the farmer that strain 3 is really better than strain 4. This is a decision where intuition may not approve with at a high confidence level. The data we used is based on a study by Becker [3] on selecting between chicken stocks. The data has been manipulated to show the point.

To deal with this unsatisfactory feature of Bechhofer's procedure we generalize the standard indifference zone approach by introducing a minimal distance $c$ ($c \geq 0$), the **preference threshold value**, which should be specified in the design stage. Coolen and Van der Laan [12] have studied this generalization providing a new foundational argument for a positive c. For the design of the experiment we need to specify two values, $n$ and $c$, so we also need two requirements. Coolen and Van der Laan use one requirement similar to Bechhofer's $P^*$ requirement, that is selecting the best one with at least probability $P^*$, and their second requirement is related to the probability of a False Selection, $P(FS)$, the selection of a single but non-best population. Their rule is to select the population with the largest sample sum if the difference to the other sample sums exceeds threshold $c$. This is what they call a strong preference. If the difference is smaller than $c$ they do not make a selection.

In this thesis we propose to select the population with the largest sample sum alone if the difference to the second-largest sample sum exceeds the threshold value $c$. We can call this a strong preference for this population. In case the difference between the largest and next-to-largest sample sums does not exceed $c$, we propose to select all populations that have a sample sum with distance $c$ or less to the maximum sample sum. The value of $c$ is dependent on the value of the sample size $n$. Bechhofer's and Gupta's basic approaches will be described in more detail in chapter 2, together with a brief description of some other approaches.

In this thesis several aspects of selection procedures will be discussed. The problem of selection has been introduced in this introduction and some basic approaches to deal with

selection problems will be considered in chapter 2. In chapter 3 the selection procedure that we propose is introduced and the basic features of the procedure are considered. Further analysis on our procedure is presented in chapter 4. The robustness of our procedure with regard to the assumption of known variances is considered in chapter 5. In chapter 6 we briefly consider a Bayesian data-analytic approach. To illustrate our procedure and to compare it to other procedures we report the results of a number of simulations in chapter 7. Finally, in chapter 8 some suggestions for future research will be discussed.

# Chapter 2

# Statistical Selection

## 2.1 Bechhofer's indifference zone approach

Bechhofer [1] introduces a procedure to select a single 'best' population out of $k$ populations, in some well-defined sense of 'best'. Throughout this thesis we will denote random variables in capitals and their realized observations in small letters. Hence, we denote the $j$th observation taken from population $i$ by $y_{ij}$ $(i = 1, \ldots, k)$ and we assume all populations to be independent normally distributed with unknown means and common known variance $\sigma^2 (> 0)$, so $Y_{ij} \sim N(\mu_i, \sigma^2)$. The sample sizes are all equal to $n$, where $n$ is to be determined. We can rank the means like

$$\mu_{[1]} \leq \mu_{[2]} \leq \cdots \leq \mu_{[k]}. \tag{2.1}$$

The goal is to select the population with mean $\mu_{[k]}$, the actual best mean. The sample means are denoted by $(i = 1, \ldots, k)$

$$\bar{y}_i = \frac{1}{n} \sum_{j=1}^{n} y_{ij} \tag{2.2}$$

and we have $\bar{Y}_i \sim N(\mu_i, \frac{\sigma^2}{n})$. The population with mean $\mu_{[i]}$ is denoted by $\pi_{(i)}$ with related sample mean, $\bar{Y}_{(i)}$ and sample sum, $\sum_{j=1}^{n} Y_{(i)j}$ which is distributed like $N(n\mu_{[i]}, n\sigma^2)$ $(i = 1, \ldots, k)$. The procedure is to take a sample of size $n$ from each of the $k$ populations, to calculate the $k$ sample sums and to assert that the one with the largest sample sum is the best population. This rule will be referred to as $R_B$. Bechhofer uses the sample means as response variable but that is essentially the same, because of the fixed sample size. The probability of making a Correct Selection in this approach $(CS_B)$, that is selecting the best

population, $\pi_{(k)}$, is (analogous to [1])

$$\begin{aligned} P(CS_B|\, R_B) &= P(\sum_{j=1}^{n} Y_{(k)j} > \max_{1\le i\le k-1} \sum_{j=1}^{n} Y_{(i)j}) \\ &= P(n\bar{Y}_{(i)} < n\bar{Y}_{(k)};\ i = 1,\ldots,k-1) \\ &= \int_{-\infty}^{\infty} P(n\bar{Y}_{(i)} < y;\ i = 1,\ldots,k-1) dP(n\bar{Y}_{(k)} \le y) \\ &= \int_{-\infty}^{\infty} \prod_{i=1}^{k-1} \left[\Phi(\frac{y - n\mu_{[i]}}{\sqrt{n}\sigma})\right] d\Phi(\frac{y - n\mu_{[k]}}{\sqrt{n}\sigma}) \\ &= \int_{-\infty}^{\infty} \prod_{i=1}^{k-1} \left[\Phi(z + \frac{n\mu_{[k]} - n\mu_{[i]}}{\sqrt{n}\sigma})\right] d\Phi(z). \end{aligned} \tag{2.3}$$

In this derivation and throughout the thesis we will denote the standard normal cumulative distribution by $\Phi(\cdot)$ and its derivative the standard normal density function by $\phi(\cdot)$. Bechhofer divides the complete parameter space, $\Omega$,

$$\Omega = \{\mu : \mu = (\mu_1,\ldots,\mu_k), \mu_i \in \mathbb{R},\ i = 1,\ldots,k\} \tag{2.4}$$

in the **indifference zone**, where we are indifferent about a correct selection because the difference between the largest and second-largest mean is less than a chosen value $\delta^*(> 0)$, and the **preference zone**, where this difference is at least $\delta^*$. Thus we have

$$\Omega(\delta^*) = \{\mu \in \Omega\, |\mu_{[k]} - \mu_{[k-1]} \ge \delta^*\} \tag{2.5}$$

is the preference zone and the indifference zone is the complement of this: $\mathbb{R}^k \setminus \Omega(\delta^*)$. The requirement imposed on the probability of a correct selection, $P(CS_B)$, for $\mu \in \Omega(\delta^*)$ is

$$\inf_{\Omega(\delta^*)} P(CS_B \mid R_B) \ge P^*, \tag{2.6}$$

where $\delta^* > 0$ and $P^* \in (\frac{1}{k}, 1)$ are to be specified, and we want to obtain the smallest common sample size $n$ for which the probability requirement is satisfied. Smaller sample sizes is usually more economical in time and money. Of course, $P^*$ should be larger than $1/k$, because you can achieve that probability by just picking one populationwithout further observations. The infimum of $P(CS_B \mid R_B)$ over $\Omega(\delta^*)$ is attained for the Least Favourable Configuration (LFC), the parameter configuration for which the value of the infimum over $\Omega(\delta^*)$ is actually attained. This LFC is [1] (and also more generally studied by Bofinger [7])

$$\mu_{[1]} = \cdots = \mu_{[k-1]} = \mu_{[k]} - \delta^* \tag{2.7}$$

and the configuration will be referred to as $\mu_{LFC(\delta^*)}$. The probability of a correct selection for this least favourable configuration is equal to (using result (2.3))

$$\begin{aligned} P(CS_B \mid R_B, LFC) &= \int_{-\infty}^{\infty} \prod_{i=1}^{k-1} \left[\Phi(z + \frac{n\mu_{[k]} - n\mu_{[i]}}{\sqrt{n}\sigma} \mid LFC)\right] d\Phi(z) \\ &= \int_{-\infty}^{\infty} \Phi^{k-1}(z + \frac{n\delta^*}{\sqrt{n}\sigma}) d\Phi(z) \end{aligned} \tag{2.8}$$

and this leads to the $P^*$ requirement

$$\int_{-\infty}^{\infty} \Phi^{k-1}(z+\tau_B)d\Phi(z) = P^* \tag{2.9}$$

where

$$\tau_B = \frac{\sqrt{n}\delta^*}{\sigma}. \tag{2.10}$$

Values for $\tau_B$ have been tabulated for different $k, \delta^*$ and $P^*$ values, in for example [1]. The table we provide in appendix C.1 also gives the appropriate values of $\tau_B$. Using these values we get the minimum required value for $n$ to satisfy (2.9), determined by

$$n = (\frac{\sigma\tau_B}{\delta^*})^2. \tag{2.11}$$

If this value is not an integer, the minimum required sample size $n$ is derived by taking the nearest higher integer value, for which the probability $P(CS_B)$ will not be less than $P^*$.

## 2.2 Gupta's subset selection

The essential difference between Bechhofer's indifference zone approach and the second basic approach by Gupta, that we introduce here, is that in Gupta's approach the sample size is assumed to be given and it is not considered an explicit goal to determine an optimal sample size. Gupta worked on this problem in 1956 as part of his Ph.D. thesis [17], but an article about it only appeared in 1965. This section is based on the latter article. Using the same notation as in section 2.1, Gupta's rule [18] can be described as: select the $i$th population if and only if

$$\sum_{j=1}^{n} y_{ij} \geq \max_{1\leq l\leq k} \sum_{j=1}^{n} y_{lj} - d, \qquad (d>0). \tag{2.12}$$

Note that we again use the sample sum as response variable, whereas Gupta used the sample means, but due to the fixed sample sizes this is in fact the same. We refer to this rule as $R_G$. A Correct Selection for this subset selection approach occurs when a subset is selected, that includes the best population $\pi_{(k)}$. To distinguish this correct selection from Bechhofer's definition of a correct selection, we denote Gupta's Correct Selection by $CS_G$. The goal in the design is to determine distance $d$ to achieve a certain probability requirement. The probability of a $CS_G$ is [18]

$$\begin{aligned}
P(CS_G \mid R_G) &= P(\sum_{j=1}^{n} Y_{(k)j} \geq \max_{1\leq i\leq k-1} \sum_{j=1}^{n} Y_{(i)j} - d) \\
&= P(n\bar{Y}_{(i)} \leq n\bar{Y}_{(k)} + d;\ i = 1,\ldots,k-1) \\
&= \int_{-\infty}^{\infty} P(n\bar{Y}_{(i)} \leq y + d;\ i = 1,\ldots,k-1)dP(n\bar{Y}_{(k)} \leq y) \\
&= \int_{-\infty}^{\infty} \prod_{i=1}^{k-1} \left[\Phi(\frac{y+d-n\mu_{[i]}}{\sqrt{n}\sigma})\right] d\Phi(\frac{y-n\mu_{[k]}}{\sqrt{n}\sigma}) \\
&= \int_{-\infty}^{\infty} \prod_{i=1}^{k-1} \left[\Phi(z + \frac{d+n\mu_{[k]}-n\mu_{[i]}}{\sqrt{n}\sigma})\right] d\Phi(z).
\end{aligned} \tag{2.13}$$

From (2.13), because of the normal distribution function being monotonously increasing and the ordering of the means, $\mu_{[1]} \leq \mu_{[2]} \leq \cdots \leq \mu_{[k]}$, we see that the Least Favourable Configuration is taken when $\mu_{[1]} = \mu_{[2]} = \cdots = \mu_{[k]}$ [18] and we have

$$\inf_{\Omega} P(CS_G \mid R_G) = \int_{-\infty}^{\infty} \Phi^{k-1}(z + \frac{d}{\sqrt{n}\sigma}) d\Phi(z), \tag{2.14}$$

where $\Omega$ is again the space defined by (2.4). Thus if we want to make a correct selection with probability $P^* \in (1/k, 1)$, the probability requirement becomes

$$\int_{-\infty}^{\infty} \Phi^{k-1}(z + \tau_G) d\Phi(z) = P^*, \tag{2.15}$$

where

$$\tau_G = \frac{d}{\sqrt{n}\sigma}. \tag{2.16}$$

## 2.3 Sequential selection

One can distinguish two different basic ideas of selection procedures. There are the procedures we have considered before, designing an experiment by calculating the required sample size or taking a fixed sample size, take the samples and make a statement. Another way is to take observations of each population one by one (or batchwise) and after each stage try to make a statement or continue taking observations. These procedures are called sequential procedures. The big advantage of a sequential selection procedure above the non-sequential ones, is that you can avoid taking samples from populations that are obviously worse than at least one of the other populations, by eliminating them in an early stage and hence you can economize on the total number of samples to be taken. In practice, this means saving resources, money or time. Sequential procedures are especially applicable in flow-production situations etc.. However, we must realize that this is not always possible, because in some situations you have to work with a design. For example, if we want to select the best variety of potatoes, we have to plant a certain number of plants and this has to be large enough in order to make the required statement after the experiment, which will typically take some months.

In this section we take a look at a sequential selection procedure as proposed by Coolen [11], which is an adaption of the plan by Hoel and Mazumdar [21], and we will compare the sequential procedure in chapter 7 to the preference threshold procedure. The problem is again to find the best population (the one with the largest mean) among $k$ independent normally distributed populations with common known variance $\sigma^2$ and unknown means $\mu_i$. Dividing the parameter space into the indifference and preference zone, the sequential procedure has to select the best population with probability of at least $1 - \alpha$ when the difference between the largest and second-largest mean is larger than $\delta^*$, so

$$\inf_{\Omega(\delta^*)} P(CS) \geq 1 - \alpha. \tag{2.17}$$

Of course, $\alpha$ is restricted to $\alpha \in (0, 1 - 1/k)$, for obvious reasons. If we take $P^* = 1 - \alpha$ we achieve the same probability requirement as for the standard indifference zone approach. The sequential selection rule, $R_{seq}$, is described as follows. First, we need a stopping boundary, which is defined as

$$g = \frac{\sigma^2}{\delta^*} \ln\left(\frac{k - 1 - \alpha}{1 - \alpha}\right). \tag{2.18}$$

To start, we take one observation from each population, obtaining observations $(y_{11}, \ldots, y_{k1})$. Then any population $\pi_i$ for which

$$y_{i1} < \max\{y_{11}, \ldots, y_{k1}\} - g$$

is eliminated. If only a single population is not eliminated, that population is selected as the best one and the experiment is terminated. If not, we take a second observation from all remaining populations. In general, in the $n$th stage of the experiment $(n = 2, 3, \ldots)$ we take one observation from each population not eliminated before or at stage $n-1$ and we eliminate any remaining population $\pi_i$ for which

$$\sum_{j=1}^{n} y_{ij} < \max_{r}\{\sum_{j=1}^{n} y_{rj}\} - g$$

where the maximum is taken over all populations that are still in the experiment. The experiment stops at the first moment that only a single population is left, which is then selected as the best one. This procedure is called an open sequential procedure, that is it has no upperbound for the maximum number of stages needed before selection of a single population. However, the sequential procedure terminates with probability one. Indeed this procedure selects the best population with probability of at least $(1-\alpha)$ as Coolen shows in [11] using the theory of Brownian Motion Processes. It is to be remarked that $(1-\alpha)$ is an actual lower bound of the probability of correct selection in this procedure and there are no non-trivial cases known in which this lower bound is actually achieved. (Different to the standard indifference zone and standard subset selection approaches, in which the lower bound is achieved for respectively least favourable configurations. Some simulations of this sequential procedure can be found in chapter 7.

## 2.4 Some other developments in selection procedures

Since the work of the two pioneers, Bechhofer and Gupta [1] and [18], the problem of selecting the best has been considered in a lot of studies. We mention two textbooks on the problem. One, more easily accessible and presenting mostly variations on the standard indifference zone approach, but also considering the subset selection approach, is by Gibbons *et al.* [16]. More comprehensive is the book by Gupta and Panchapakesan [19] that presents a survey of ranking and selection theory at that time. In these books we found the methods to tackle selection problems without assuming common known variances, the two-stage procedure by Bechhofer, Dunnet and Sobel [2], the sequential procedure by Robbins, Sobel and Starr [26] and the two-stage method by Dudewicz and Dalal [15]. We briefly describe them in section 5.1.

Driessen *et al.* [14] study the effect of a deviation from the assumption of common known variances and chapter 5, in which we regard the robustness of our procedure with respect to this, is inspired by this article. Bofinger [7] studied the least favourable configuration in more detail and formulated a general definition, however, each case has to be examined carefully. With regard to the problem of non-monotonicity of the probability of correct selection with respect to sample sizes and variances, Bofinger [8] gives conditions on the variances which ensure that the probability of correct selection increases when any sample size is increased and she concludes with "bigger samples are usually better". Driessen [13], however, also pays attention to this problem and indicates some incorrect proofs. Driessen shows that when the means of $k$ normal populations are close enough, the probability of the population with the largest mean is strictly increasing in $\sigma_k/n_k$.

Gutman and Maymin [20] use a selection rule based on a threshold between the largest and second-largest observation of the random variable you are interested in. However, they

do not look at the design of an experiment, but study the probability that the selected population is the actual best one, after taking the data. Hsu [22] regards confidence intervals for all distances from the "best" under the location model and discusses the relations for (amongst others) the standard indifference zone approach and the standard subset selection procedure.

A new kind of subset selection procedure for $k$ normal populations with a common known variance is proposed by Somerville [28], based on F-ratios. The procedure is compared to Gupta's standard subset selection procedure using Monte Carlo methods and from the results the conclusion is made that there is strong evidence that the procedure is at least as efficient as Gupta's in terms of expected subset size.

Van der Laan and Verdooren [23] present an overview of selection procedures with emphasis on normal, binomial, Poisson and multinomial populations, because of their importance for agricultural applications. Furthermore some generalizations, modifications and the selection procedure of Somerville are discussed.

A sequence of articles has appeared in the early sixties on an application of selection procedures on chicken stocks. Becker [3] presents a study on the effect of the difference $\delta$ between the best and second-best chicken stock, with 'best' related to the maximum of hen-housed egg production, using Bechhofer's procedure. He concludes that the probability of the best one winning the test was fairly low for a sample of size 50 to 100, and if the differences between stocks become smaller, the chance of a correct ranking becomes less. He also shows how to use a combination of several tests by using the so-called Control Method, utilizing the strains that appeared in all tests. The latter dealt with a quantitative trait, Becker [4] studied all-or-none traits, such as mortality. He shows that the ranking of samples in this case, is affected by the same factors (sample size, difference between stocks) in a similar way as is the case with a quantitative trait. Continuing on this, Becker [5] demonstrated that by artificial and economical selection the true differences between stocks ($\delta$) decreased and therefore the probability of correctly selecting the best one also decreased for both quantitative and all-or-none traits . Usually a least favourable configuration is used to calculate the probability of a correct ranking, but we have to keep in mind that this is an *underestimate*. Soller and Putter [27] realized this and compared the results found under assumption of a least favourable configuration and under values more likely to appear in practice, the so-called *average configuration*, based on the Becker's data. In the cases they consider, the probability of the best stock actually ranking best, are about twice as large under the "average" configuration as under the least favourable configuration.

# Chapter 3

# Preference threshold procedure

## 3.1 Introduction

The first standard approach, the indifference zone approach, has the disadvantage of completely neglecting the second-largest sample sum, even when its sample sum only has a very small difference to the largest sample sum. This has already been mentioned in the introduction in chapter 1. For example, we could take our samples and calculate the sample sums and find out that the largest one is 100.1 and the second-largest is equal to 100. Then it seems unsatisfactory to neglect the population corresponding to the second-largest sample sum. In the second approach, Gupta's subset selection procedure, this problem does not occur, but a disadvantage is that we cannot determine the sample size which is considered to be given. The ability of determining the sample size is especially important in the designing phase of an experiment. In the procedure we are proposing in this chapter, we combine the two advantages of the standard approaches by introducing a **preference threshold**, a distance from the maximum sample sum. We select all populations in a subset that are within this distance to the maximum sample sum. We can say that we have a *strong* preference for all populations in the selected subset with regard to the non-selected populations and that we have a weak preference for the population corresponding to the largest sample sum with regard to the selected other populations.

## 3.2 Selection rule

For $k$ ($k \geq 2$) independent normally distributed populations with unknown means and common known variance $\sigma^2 > 0$, we regard the problem of selecting the 'best' population, where the 'best' is the population with the largest population mean. (If we are interested in the population with the smallest mean a similar procedure can be used, due to the symmetry of the normal distribution. See appendix A for a derivation.) The $k$ populations are denoted by $\pi_i$, ($i = 1, 2, \ldots, k$). We will take samples of fixed size $n(> 0)$ of each population and we denote the $j$th observation taken from population $\pi_i$ by $y_{ij}$, so $Y_{ij} \sim N(\mu_i, \sigma^2)$, with $\mu_i$ the unknown mean of population $\pi_i$ and all $Y_{ij}$ mutual independent. A standard notation for the ordered means is

$$\mu_{[1]} \leq \mu_{[2]} \leq \cdots \leq \mu_{[k]}, \tag{3.1}$$

and we say that $\mu_{[i]}$ is the mean of population $\pi_{(i)}$. Our goal can now be stated as: Selecting population $\pi_{(k)}$, that is the one corresponding to $\mu_{[k]}$, the largest mean. The sample mean of population $i$ will be denoted as usual by $\bar{Y}_i$. As a response variable we take the sample sums, $\sum_{j=1}^{n} Y_{ij} = n\bar{Y}_i \sim N(n\mu_i, n\sigma^2)$ ($i = 1, \ldots, k$). For the maximum value of the response

variables of the $k$ populations we introduce

$$\sum_{j=1}^{n} Y_{Mj} = \max_{1 \le i \le k} \sum_{j=1}^{n} Y_{ij},$$

corresponding to population $\pi_M$. For the reasons mentioned in the introduction 1, we propose the following procedure:

## Procedure

- Take a sample of equal size $n$ from each of the $k$ populations.
- Calculate the $k$ sample sums.
- Select all populations $\pi_i$ satisfying:

$$\sum_{j=1}^{n} y_{ij} \ge \sum_{j=1}^{n} y_{Mj} - c, \qquad (c \ge 0) \tag{3.2}$$

The introduced $c$ in the procedure above is called the **preference threshold**, because the maximum sample sum of a population has to have a difference of at least this value $c$ to the other populations in order to be selected as the only best one. We refer to **selection rule** (3.2) as $R_{n,c}$. As Bechhofer did, we divide the complete parameter space in the indifference and preference zone, where the difference between the largest and second-largest mean is at least $\delta^*$. Assuming to be in the preference zone ($\mu_{[k]} - \mu_{[k-1]} \ge \delta^*$), we want our procedure to satisfy **two requirements**:

## Requirements

1. The probability of selecting the best population, $\pi_{(k)}$, *only* in the selected subset should be at least $P^* \in (1/k, 1)$.
2. The overall probability of selecting $\pi_{(k)}$ in the selected subset should be at least $Q^* \in (1/k, 1)$.

In fact, this procedure is a kind of subset selection procedure, with the probability requirement restricted to $\Omega(\delta^*)$, with an additional requirement, namely the probability that the selected subset includes the best population only should be at least $P^*$. In the design stage of a selection experiment, we have to determine the values for preference threshold $c$ and the smallest possible $n$ in order to guarantee that we satisfy the two requirements, for given values of $P^*, Q^*$ and $\delta^*$. The values of $P^*$ and $Q^*$ should be $\in (1/k, 1)$, because by just picking one population out of $k$ we have already a probability of $1/k$ of picking the best one, and of course every required probability should be less than 1.

This idea of a preference threshold is first developed by Coolen and Van der Laan in [12]. They used the same first requirement as we do (selecting *only* the best population with a certain probability), however, their second requirement is different than ours and is based on the probability of a false selection. Other requirements would also be interesting and we briefly discuss some possibilities in chapter 8.

## 3.3 Probability of a correct selection

As we concluded before, our procedure is in fact a subset selection procedure with an additional requirement. This is reflected in the probability of a correct selection. A correct selection for our selection procedure occurs, like in Gupta's approach, when the best population, $\pi_{(k)}$, is selected in the subset. This event, referred to as $CS_2$, should occur

with a probability of at least $Q^*$. According to the additional requirement, we want the probability of correctly selecting the best population $\pi_{(k)}$ *only*, to be at least $P^*$. We refer to this as $CS_1$ (subscript $_1$ is denoting the event that only 1 population is selected). The two requirements the procedure has to satisfy, as mentioned in section 3.2, are:

$$\inf_{\Omega(\delta^*)} P(CS_1 \mid R_{n,c}) \geq P^* \tag{3.3}$$

and

$$\inf_{\Omega(\delta^*)} P(CS_2 \mid R_{n,c}) \geq Q^*. \tag{3.4}$$

The probability of a correct selection of $\pi_{(k)}$ only, is

$$\begin{aligned}
P(CS_1|R_{n,c}) &= P(\sum_{j=1}^{n} Y_{(k)j} > \max_{1\leq i\leq k-1} \sum_{j=1}^{n} Y_{(i)j} + c) \\
&= P(n\bar{Y}_{(i)} + c < n\bar{Y}_{(k)};\ i = 1, \ldots, k-1) \\
&= \int_{-\infty}^{\infty} P(n\bar{Y}_{(i)} < y - c;\ i = 1, \ldots, k-1) dP(n\bar{Y}_{(k)} \leq y) \\
&= \int_{-\infty}^{\infty} \prod_{i=1}^{k-1} \left[\Phi(\frac{y - c - n\mu_{[i]}}{\sqrt{n}\sigma})\right] d\Phi(\frac{y - n\mu_{[k]}}{\sqrt{n}\sigma}) \\
&= \int_{-\infty}^{\infty} \prod_{i=1}^{k-1} \left[\Phi(z + \frac{n\mu_{[k]} - n\mu_{[i]} - c}{\sqrt{n}\sigma})\right] d\Phi(z).
\end{aligned} \tag{3.5}$$

The minimum of this probability over $\Omega(\delta^*)$ is attained at the Least Favourable Configuration (LFC), [1] and [7] and this configuration for the means is:

$$\mu_{[1]} = \cdots = \mu_{[k-1]} = \mu_{[k]} - \delta^*. \tag{3.6}$$

For ease of notation, we will refer to this configuration by $\mu_{LFC(\delta^*)}$. The probability of a correct selection for the LFC becomes,

$$\begin{aligned}
\inf_{\Omega(\delta^*)} P(CS_1 \mid R_{n,c}) &= P(CS_1|R_{n,c}, LFC) \\
&= \int_{-\infty}^{\infty} \Phi^{k-1}(z + \frac{n\delta^* - c}{\sqrt{n}\sigma}) d\Phi(z) \\
&= \int_{-\infty}^{\infty} \Phi^{k-1}(z + \tau_1) d\Phi(z),
\end{aligned} \tag{3.7}$$

where

$$\tau_1 = \frac{n\delta^* - c}{\sqrt{n}\sigma}. \tag{3.8}$$

The first probability requirement,

$$\inf_{\Omega(\delta^*)} P(CS_1 \mid R_{n,c}) \geq P^* \tag{3.9}$$

is satisfied if

$$\int_{-\infty}^{\infty} \Phi^{k-1}(z + \tau_1) d\Phi(z) = P^*. \tag{3.10}$$

Values for $\tau_1$ can be calculated for given values of $P^*$ and $k$ using (3.10), thus giving a **first relation** between $n$ and $c$,

$$c = n\delta^* - \tau_1\sqrt{n}\sigma. \tag{3.11}$$

In table C.1 these values for $\tau_1$ are tabulated for $k = 2, \ldots, 10, 15, 20, 30, 40, 50$ populations and $P^* = 0.6, 0.7, 0.75, 0.8, 0.85, 0.9, 0.95, 0.975, 0.99$. For the probability of selecting a correct subset using the distance $c$, the second requirement, we can derive the following:

$$\begin{aligned}
P(CS_2 \mid R_{n,c}) &= P(\sum_{j=1}^{n} Y_{(k)j} \geq \max_{1\leq l\leq k-1} \sum_{j=1}^{n} Y_{(l)j} - c) \\
&= P(n\bar{Y}_{(i)} \leq n\bar{Y}_{(k)} + c;\ i = 1, \ldots, k-1) \\
&= \int_{-\infty}^{\infty} P(n\bar{Y}_{(i)} \leq y + c;\ i = 1, \ldots, k-1) dP(n\bar{Y}_{(k)} \leq y) \\
&= \int_{-\infty}^{\infty} \prod_{i=1}^{k-1} \left[\Phi(\frac{y + c - n\mu_{[i]}}{\sqrt{n}\sigma})\right] d\Phi(\frac{y - n\mu_{[k]}}{\sqrt{n}\sigma}) \\
&= \int_{-\infty}^{\infty} \prod_{i=1}^{k-1} \left[\Phi(z + \frac{n\mu_{[k]} - n\mu_{[i]} + c}{\sqrt{n}\sigma}\right] d\Phi(z).
\end{aligned} \tag{3.12}$$

Using a similar argument like Gupta in [18], we see that the infimum for (3.12) over $\Omega(\delta^*)$ will be taken for the same LFC as for $P(CS_1)$, $\mu_{LFC(\delta^*)}$, because of the monotonously increasing normal distribution function, so we have

$$\begin{aligned}
\inf_{\Omega(\delta^*)} P(CS_2 \mid R_{n,c}) &= P(CS_2 \mid R_{n,c}, LFC) \\
&= \int_{-\infty}^{\infty} \Phi^{k-1}(z + \frac{c + n\delta^*}{\sqrt{n}\sigma}) d\Phi(z) \\
&= \int_{-\infty}^{\infty} \Phi^{k-1}(z + \tau_2) d\Phi(z),
\end{aligned} \tag{3.13}$$

where

$$\tau_2 = \frac{c + n\delta^*}{\sqrt{n}\sigma}. \tag{3.14}$$

From this we can conclude that the second probability requirement,

$$\inf_{\Omega(\delta^*)} P(CS_2 \mid R_{n,c}) \geq Q^* \tag{3.15}$$

will be satisfied if

$$\int_{-\infty}^{\infty} \Phi^{k-1}(z + \tau_2) d\Phi(z) = Q^*. \tag{3.16}$$

Note that (3.16) has exactly the same form as (3.10), so values of $\tau_2$ can be derived analogously and we can use the same table C.1 for the values of $\tau_2$. Using (3.14) we get a **second relation**

$$c = \tau_2\sqrt{n}\sigma - n\delta^*. \tag{3.17}$$

Together these two relations (3.11) and (3.17) can be solved for $n > 0$ and $c \geq 0$, resulting in:

$$n\delta^* - \sqrt{n}\sigma\tau_1 = \sqrt{n}\sigma\tau_2 - n\delta^*$$
$$n = \left(\frac{\sigma(\tau_1 + \tau_2)}{2\delta^*}\right)^2$$

and

$$c = \frac{\sigma^2}{4\delta^*}(\tau_2^2 - \tau_1^2).$$

So, we should apply selection rule $R_{n,c}$ for given $k, P^*, Q^*$ and $\delta^*$ with:

$$n = \left(\frac{\sigma(\tau_1 + \tau_2)}{2\delta^*}\right)^2 \quad \text{and} \quad c = \frac{\sigma^2}{4\delta^*}(\tau_2^2 - \tau_1^2). \tag{3.18}$$

To satisfy both requirements in practice, we have to take the nearest higher integer value for the sample size if $n$ is not an integer. We must remark here, that because of the rounding of $n$, the values for $\tau_1$ and $\tau_2$ also change, and therefore the achieved probabilities for the two requirements are slightly different. Sample size $n$ is always rounded upward, hence the achieved practical values for $\inf_{\Omega(\delta^*)} P(CS_1 \mid R_{n,c})$ and $\inf_{\Omega(\delta^*)} P(CS_2 \mid R_{n,c})$ will be slightly larger than $P^*$ and $Q^*$ respectively. This follows from the positive correlation between $\tau_1$ and $n$ in (3.8) and between $\tau_2$ and $n$ in (3.14). In the simulations in chapter 7 there are some examples.

## 3.4 Analysis on $P^*$ and $Q^*$

When considering the first and second requirement carefully we can remark that the requirement of selecting a subset *including* the best population is a weaker condition than the requirement of selecting *only* the best one in the selected subset. If the latter condition should be satisfied with at least a probability of $P^*$, the probability for the selection of a subset including the best one (at least $Q^*$) should be larger to make sense. Hence, we should have $Q^* > P^*$. This can also be made clear by putting the first and second requirements in the following slightly different way. Our conditions are:

1. $$P(\sum_{j=1}^{n} Y_{(k)j} - \sum_{j=1}^{n} Y_{(i)j} > c;\ \forall i \neq k) \geq P^* \tag{3.19}$$
2. $$P(\sum_{j=1}^{n} Y_{(k)j} - \sum_{j=1}^{n} Y_{(i)j} > -c;\ \forall i \neq k) \geq Q^*. \tag{3.20}$$

If the first requirement (3.19) is satisfied and $Q^* < P^*$, then the second (3.20) will surely be satisfied, because $c \geq 0$. So to determine both $n$ and $c$ we must have

$$Q^* \geq P^*. \tag{3.21}$$

As the probabilities of a correct selection derived in (3.10) and (3.16) have the same functional forms and $\Phi(\cdot)$ is increasing, there are some remarks about the features of our procedure to be made. From conditions (3.10) and (3.16) we can derive

**Corollary 1**

$$P^* = Q^* \quad \Leftrightarrow \quad \tau_1 = \tau_2 \tag{3.22}$$
$$Q^* > P^* \quad \Leftrightarrow \quad \tau_2 > \tau_1 \tag{3.23}$$

The preference threshold procedure can easily be reduced to the standard indifference zone approach. If we choose $P^* = Q^*$ then we apply in fact the standard Bechhofer approach. This can be seen from (3.22), because in case $P^* = Q^*$ we have $\tau_1 = \tau_2$, so using (3.18) we derive that $n$ and $c$ reduce to $n = (\frac{\sigma 2\tau_1}{2\delta^*})^2 = (\frac{\sigma\tau_1}{\delta^*})^2$ and $c = 0$, for which values $R_{n,c}$ equals $R_B$ (section 2.1).

Trying to derive the standard Gupta approach from our procedure, we can in effect delete the first requirement by choosing $P^* = 0$. By choosing $\delta^* = 0$, resulting in the parameter space Gupta is assuming, we get from (3.14) $\tau_2 = \frac{c}{\sqrt{n}\sigma}$ which is the relation we would find using Gupta's approach (2.16).

Furthermore, from (3.23) and (3.18) it follows that we should choose $Q^* > P^*$ to have a positive threshold $c$ as is intended in our procedure.

## 3.5 Analysis on $n$ and $c$

### 3.5.1 Intuitive explanation of functional form

For a given $k$, $P^*$, $Q^*$ and $\delta^*$ the final result for the minimal required sample size $n$ and the preference threshold $c$ are given by (3.18) $n = \left(\frac{\sigma(\tau_1+\tau_2)}{2\delta^*}\right)^2$ and $c = \frac{\sigma^2}{4\delta^*}(\tau_2^2 - \tau_1^2)$. Sample size $n$ is increasing in both $\tau_1$ and $\tau_2$, what is to be expected, because a larger value for $P^*$ (resp. $Q^*$), hence a larger $\tau_1$ ($\tau_2$) signifies that we want to select the one best population with a higher probability and thus we need to take more observations in order to achieve this probability. Remark that $n$ is increasing in $\sigma$ and decreasing in $\delta^*$. This is also easy to understand, for a larger variance compels us - for the sample sums would vary more and one of the $k-1$ non-best populations could coincidently produce a sample sum larger than the best population- to take more observations to stabilize the sample sums. A larger value of $\delta^*$ means that the indifference zone becomes larger, hence we do not have to be as accurate as before, so the sample size does not need to be that large.

The preference threshold value $c$ is increasing in $\tau_2$ (related to $Q^*$) and decreasing in $\tau_1$ (related to $P^*$). This is logically justifiable, for a larger value for $Q^*$ would mean a higher probability to include the best population in the selected subset, thus the threshold $c$, the difference from the maximum sample sum, should be enlarged. Moreover, when the value for $P^*$ is raised, we want to select only the one best population with a higher probability, therefore the difference $c$ should get smaller in order to increase the probability of selecting only the population with the largest sample sum. Like $n$, $c$ is increasing in $\sigma$ and decreasing in $\delta^*$. When the variance becomes larger, the value for $c$ should also become larger to ensure the same probability requirements. When $\delta^*$ is increased, the indifference zone is enlarged and thus in a larger area we are indifferent with regard to a selection and therefore the value for $c$ will be smaller for larger values of $\delta^*$.

### 3.5.2 Comparison of sample size with standard approaches

It is interesting to compare the value we get for the sample size $n$ to the value we would get using the standard indifference zone approach. If we would apply Bechhofer's procedure, (section 2.1), we would find under $\mu_{LFC(\delta^*)}$, the LFC,

$$P(CS|R_B, LFC) = \int_{-\infty}^{\infty} \prod_{i=1}^{k-1} \Phi^{k-1}(z + \tau_B) d\Phi(z), \tag{3.24}$$

where

$$\tau_B = \frac{\sqrt{n}\delta^*}{\sigma}, \tag{3.25}$$

according to (2.8) and (2.10). For the mimimum required sample size to satisfy probability requirement (2.9), with $P^*$ replaced by $P'$ and associated with $\tau_{P'}$, the relation following from (3.25) holds

$$n_{P'} = \left(\frac{\sigma\tau_{P'}}{\delta^*}\right)^2. \tag{3.26}$$

Remember that for our procedure $R_{n,c}$ sample size $n$ is $\left(\frac{\sigma(\tau_1+\tau_2)}{2\delta^*}\right)^2$. If we choose $P^*$ for our threshold selection procedure and we choose $P' = P^*$ for Bechhofer's procedure, we find $\tau_1 = \tau_{P'}$. Also, according to the restriction for $P^*$ and $Q^*$, we choose $Q^* > P^*$ leading to (using (3.23)) $\tau_2 > \tau_1$. Now we can derive

$$n = \left(\frac{\sigma(\tau_1+\tau_2)}{2\delta^*}\right)^2 > \left(\frac{\sigma(2\tau_1)}{2\delta^*}\right)^2 = \left(\frac{\sigma(\tau_{P'})}{\delta^*}\right)^2 = n_{P'}. \tag{3.27}$$

Analogously, if we choose $P' = Q^*$, with $P'$ for Bechhofer's procedure, we can derive

$$n = \left(\frac{\sigma(\tau_1+\tau_2)}{2\delta^*}\right)^2 < \left(\frac{\sigma(2\tau_2)}{2\delta^*}\right)^2 = \left(\frac{\sigma(\tau_{P'})}{\delta^*}\right)^2 = n_{P'}. \tag{3.28}$$

Therefore, assuming $Q^* > P^*$, we have

$$n^B_{P^*} < n < n^B_{Q^*}, \tag{3.29}$$

with $n^B_{P^*}$, denoting the minimum required sample size according to the standard Bechhofer approach with probability requirement $\inf_{\Omega(\delta^*)} P(CS) \geq P^*$. In other words, we conclude that our sample size $n$ will always be greater than the one we would require for the standard indifference zone approach for $P^*$, but $n$ will be always smaller than the one we would require for Bechhofer's approach with probability $Q^*$, assuming $Q^* > P^*$. Intuitively this is logical for the following arguments. We need at least a sample size as large as for the standard indifference zone approach with $P^*$, because we also need to satisfy the requirement that *only* the best population will be selected with probability $P^*$. However, we need a sample size smaller than the size required for the standard indifference zone approach with $Q^*$. For by taking samples of this size, we know that we select the one best population with this probability $Q^*$ and thus also our additional requirement that with probability $Q^*$ a subset of best populations is certainly satisfied.

## 3.6 Selection procedure with second probability requirement on $\Omega$

In section 3.3, we have defined a correct subset selection, referred to as $CS_2$, as having the actual best population, $\pi_{(k)}$, included in the selected subset. However, we assumed $\mu \in \Omega(\delta^*)$ and this is different from Gupta's subset selection approach, which allows configurations of $\mu$ to be in the complete parameter space $\Omega$. Concern about selection in case the actual configuration for the means is not in the preference zone, but in the indifference zone, could be a reason to consider this approach in our procedure. By not making any restriction on the parameter space for the second requirement, we have at least the guarantee that, although we are in the indifference zone and hence the first requirement is not

valid anymore, the second requirement is and thus we select a subset including the best population with at least probability $Q^*$. Let us consider what happens if we apply this to our preference threshold procedure. When we have no restrictions on the parameter space for the second requirement, we refer to the procedure as $R^{\Omega}_{n,c}$. The probability of a correct selection, $CS_2$ is analogous to (3.12),

$$P(CS_2|R^{\Omega}_{n,c}) = \int_{-\infty}^{\infty} \prod_{i=1}^{k-1} \left[ \Phi(z + \frac{n\mu_{[k]} - n\mu_{[i]} + c}{\sqrt{n}\sigma}) \right] d\Phi(z). \tag{3.30}$$

But $\mu \in \Omega$, so the infimum of this probability will be attained for the Least Favourable Configuration, $\mu_{[1]} = \cdots = \mu_{[k-1]} = \mu_{[k]}$, [18] and we get

$$\begin{aligned} \inf_{\Omega} P(CS_2 \mid R^{\Omega}_{n,c}) &= \int_{-\infty}^{\infty} \Phi^{k-1}(z + \frac{c}{\sqrt{n}\sigma}) d\Phi(z) \\ &= \int_{-\infty}^{\infty} \Phi^{k-1}(z + \tau_2^{\Omega}) d\Phi(z), \end{aligned} \tag{3.31}$$

where

$$\tau_2^{\Omega} = \frac{c}{\sqrt{n}\sigma}. \tag{3.32}$$

Our second probability requirement becomes

$$\inf_{\Omega} P(CS_2 \mid R^{\Omega}_{n,c}) \geq Q^*, \tag{3.33}$$

which will be satisfied if

$$\int_{-\infty}^{\infty} \Phi^{k-1}(z + \tau_2^{\Omega}) d\Phi(z) = Q^*. \tag{3.34}$$

Together with the first relation (3.11) we get two relations for $n$ and $c$;

$$c = \tau_2^{\Omega} \sqrt{n}\sigma \tag{3.35}$$

and

$$c = n\delta^* - \tau_1 \sqrt{n}\sigma, \tag{3.36}$$

so we can solve this for $n > 0$ and $c \geq 0$, resulting in:

$$\begin{aligned} n\delta^* - \tau_1\sqrt{n}\sigma &= \tau_2^{\Omega}\sqrt{n}\sigma \\ \sqrt{n}\{\sqrt{n}\delta^* - \sigma(\tau_1 + \tau_2^{G})\} &= 0 \\ \sqrt{n} &= \frac{\sigma(\tau_1 + \tau_2^{\Omega})}{\delta^*} \\ n &= \left( \frac{\sigma(\tau_1 + \tau_2^{\Omega})}{\delta^*} \right)^2 \end{aligned} \tag{3.37}$$

and

$$c = \frac{\sigma^2}{\delta^*}(\tau_1 + \tau_2^{\Omega})\tau_2^{\Omega}. \tag{3.38}$$

If we compare this to the sample size say (3.18) $n_{\delta^*} = (\frac{\sigma(\tau_1+\tau_2)}{2\delta^*})^2$ we found for the approach with the assumption of $\mu \in \Omega(\delta^*)$, we see that for the same values of $P^*$ and $Q^*$ the sample size in case of $\mu \in \Omega$ is four times the value of the sample size we would get under the assumption $\mu \in \Omega(\delta^*)$. This factor four can be explained by defining two $\delta$ values, one corresponding to each requirement. Hence, we impose $\delta_1^*$ on the first requirement and $\delta_2^*$ on the second. The Least Favourable Configurations are in this case defined by, for the first requirement

$$\mu_{[1]} = \cdots = \mu_{[k-1]} = \mu_{[k]} - \delta_1^* \tag{3.39}$$

and for the second requirement

$$\mu_{[1]} = \cdots = \mu_{[k-1]} = \mu_{[k]} - \delta_2^*. \tag{3.40}$$

Instead of relation (3.11) and (3.17) we now get respectively,

$$c = n\delta_1^* - \tau_1\sqrt{n}\sigma \tag{3.41}$$

and

$$c = \tau_2\sqrt{n}\sigma - n\delta_2^*. \tag{3.42}$$

We solve these two relations for $n$ and $c$, resulting in

$$n = \left(\frac{\sigma(\tau_1+\tau_2)}{\delta_1^*+\delta_2^*}\right)^2 \tag{3.43}$$

and

$$c = \frac{\sigma^2}{\delta_1^*+\delta_2^*}\left\{\frac{\delta_1^*}{\delta_1^*+\delta_2^*}(\tau_1+\tau_2)^2 - (\tau_1^2+\tau_1\tau_2)\right\}. \tag{3.44}$$

From (3.43) it is easy to see that $\delta_2^* = \delta_1^*$, as mostly used in this thesis, yields a four times smaller value for $n$ then for the case where $\delta_2^* = 0$, as suggested in this section. Of course, by imposing restrictions on the parameter space and by doing so reducing the parameter space, we will never need to take more observations then we would have to take when considering the complete parameter space $\Omega$.

Denoting the $c$ value for the approach with a restriction on the parameter space by $c_{\delta^*}$ (3.18), and the $c$ value for the approach without a restriction by $c^\Omega$ (3.38), we have

$$\begin{aligned} c_{\delta^*} &= \frac{\sigma^2}{4\delta^*}(\tau_2^2 - \tau_1^2) \\ &< \frac{\sigma^2}{\delta^*}(\tau_2^2 + \tau_1\tau_2) \\ &= c^\Omega, \end{aligned} \tag{3.45}$$

because for equal values of $P^*$ and $Q^*$ we have $\tau_2 = \tau_2^\Omega$. This is an intuitively logical result, for $\Omega(\delta^*)$ is a special case of $\Omega$, therefore the value of $c^\Omega$ will be larger than the value of $c_{\delta^*}$ to satisfy the same probability requirements.

The introduction of the two different $\delta_1^*$ and $\delta_2^*$ as in (3.39) and (3.40), raises the question if we are completely free in specifying different $\delta_1^*$ and $\delta_2^*$ values for our procedure. The answer is 'no', due to the fact that we need to have $c \geq 0$. From (3.44) we see that $c \geq 0$ implies

$$\frac{\delta_1^*}{\delta_1^*+\delta_2^*} \geq \frac{\tau_1^2+\tau_1\tau_2}{(\tau_1+\tau_2)^2}, \tag{3.46}$$

so

$$\delta_1^* \geq \frac{\tau_1^2 + \tau_1\tau_2}{\tau_2^2 + \tau_1\tau_2}\delta_2^*. \tag{3.47}$$

Remark that $P^* > 1/k$ and $Q^* > 1/k$ imply $\tau_1 > 0$ and $\tau_2 > 0$. It is easy to see that the fraction multiplying $\delta_2^*$ on the righthandside is $\in (0,1)$, because $\tau_1 \leq \tau_2$ (3.21). To be able to specify $n$ we need to have at least one of the $\delta$ values not equal to 0, this is also clear from (3.43). Therefore

$$\delta_1^* \neq 0, \tag{3.48}$$

because $\delta_1^* = 0$, (3.47) implies $\delta_2^* = 0$ by (3.47). If $\delta_2^* = \delta_1^*$, like in section 3.2 with restriction on the parameter space, (3.47) is always satisfied. If $\delta_2^* = 0$ and $\delta_1^*$ is not, then (3.47) is also satisfied, so the approach with no restrictions on the parameter space in this section also satisfies (3.47).

## 3.7 Selection of $t$ best populations

After considering the selection of one best population, we can generalize to selecting $t$ ($t \geq 1$) best populations from $k$ independent normally distributed populations with unknown means and common known variance $\sigma^2(> 0)$. Using the same notation as before, we need to introduce the ordered sample sums

$$\sum_{j=1}^{n} y_{[1]j} \leq \cdots \leq \sum_{j=1}^{n} y_{[k-t]j} \leq \sum_{j=1}^{n} y_{[k-t+1]j} \leq \cdots \leq \sum_{j=1}^{n} y_{[k]j}. \tag{3.49}$$

After taking a sample of size $n$ from each of the $k$ populations we apply the following **selection rule**: select all populations satisfying

$$\sum_{j=1}^{n} y_{ij} \geq \sum_{j=1}^{n} y_{[k-t+1]j} - c, \qquad (c \geq 0). \tag{3.50}$$

We will refer to this selection rule as $R_{n,c}^t$. A little example may illustrate the procedure. Suppose we have 5 different populations and we want to select the $t = 2$ best ones. We specify the values for $P^*$ and $Q^*$ and use a certain $n$ and $c$. Let us say $c = 1$. We take a sample of each of the populations and calculate their sample sums, resulting in for example:

$$\sum_{j=1}^{n} y_{1j} = 10, \quad \sum_{j=1}^{n} y_{2j} = 8.3, \quad \sum_{j=1}^{n} y_{3j} = 4.9, \quad \sum_{j=1}^{n} y_{4j} = 9.3, \quad \sum_{j=1}^{n} y_{5j} = 12.5.$$

Thus, $\sum_{j=1}^{n} y_{[k-t+1]j} = \sum_{j=1}^{n} y_{[4]j} = \sum_{j=1}^{n} y_{1j}$ and we select all populations with sample sums larger than or equal to $10 - 1 = 9$, leading to the selection of $\pi_5, \pi_1$ and $\pi_4$.

### 3.7.1 Second requirement on restricted space

If, again, we assume $\mu$ to be in the preference zone, now defined by

$$\mu_{[1]} \leq \cdots \leq \mu_{[k-t]} \leq \mu_{[k-t+1]} - \delta^*$$
$$\mu_{[k-t+1]} \leq \cdots \leq \mu_{[k]}$$

and denoted by $\Omega_t(\delta^*)$, we have to satisfy two requirements similar to those in section 3.2

1. The probability of selecting *only* the $t$ best populations, $\pi_{(k-t+1)}, \ldots, \pi_{(k)}$, in the selected subset should be at least $P^*$.

2. The overall probability of selecting $\pi_{(k-t+1)}, \ldots, \pi_{(k)}$ in the selected subset should be at least $Q^*$.

We denote the Correct Selection of only the $t$ best populations by $CS_t$ and the Correct Selection of a subset including the $t$ best populations is denoted by $CS_{ts}$. The first probability requirement is

$$\inf_{\Omega_t(\delta^*)} P(CS_t|R^t_{n,c}) \geq P^* \tag{3.51}$$

and the infimum will be attained for the means in the the Least Favourable Configuration, which is, see [1] and [7],

$$\begin{aligned} \mu_{[k]} &= \cdots = \mu_{[k-t+1]} \\ \mu_{[k-t+1]} - \mu_{[k-t]} &= \delta^* \\ \mu_{[k-t]} &= \cdots = \mu_{[1]}. \end{aligned} \tag{3.52}$$

Remark that in the LFC, the distributions for $\pi_{(k-t+1)}, \ldots, \pi_{(k)}$ are exactly the same. For the probability of selecting *only* $t$ populations, we have (analogously as in [1])

$$\begin{aligned} P(CS_t|R^t_{n,c}, LFC) &= P\Big(\min_{k-t+1\leq i\leq k}\sum_{j=1}^{n} Y_{(i)j} > \max_{1\leq i\leq k-t}\sum_{j=1}^{n} Y_{(i)j} + c\Big) \\ &= t\,P\Big(\max_{1\leq i\leq k-t}\sum_{j=1}^{n} Y_{(i)j} + c < \sum_{j=1}^{n} Y_{(k-t+1)j} \leq \min_{k-t+2\leq i\leq k}\sum_{j=1}^{n} Y_{(i)j}\Big) \\ &= t\int_{-\infty}^{\infty} P\Big(\max_{1\leq i\leq k-t} n\bar{Y}_{(i)} + c < y \leq \min_{k-t+2\leq i\leq k} n\bar{Y}_{(i)}\Big)\, dP\big(n\bar{Y}_{(k-t+1)} \leq y\big) \\ &= t\int_{-\infty}^{\infty} P\Big(\max_{1\leq i\leq k-t} n\bar{Y}_{(i)} < y - c\Big) P\Big(y \leq \min_{k-t+2\leq i\leq k} n\bar{Y}_{(i)}\Big)\, dP\big(n\bar{Y}_{(k-t+1)} \leq y\big) \\ &= t\int_{-\infty}^{\infty} \Phi^{k-t}\Big(\frac{y-c-n\mu_{[k-t]}}{\sqrt{n}\sigma}\Big)\Big[1-\Phi\Big(\frac{y-n\mu_{[k-t+1]}}{\sqrt{n}\sigma}\Big)\Big]^{t-1} d\Phi\Big(\frac{y-n\mu_{[k-t+1]}}{\sqrt{n}\sigma}\Big) \\ &= t\int_{-\infty}^{\infty} \Phi^{k-t}\Big(z + \frac{n\mu_{[k-t+1]} - n\mu_{[k-t]} - c}{\sqrt{n}\sigma}\Big)\,[1-\Phi(z)]^{t-1}\, d\Phi(z) \\ &= t\int_{-\infty}^{\infty} \Phi^{k-t}\Big(z + \frac{n\delta^* - c}{\sqrt{n}\sigma}\Big)\,[1-\Phi(z)]^{t-1}\, d\Phi(z). \end{aligned} \tag{3.53}$$

Our first requirement is satisfied when

$$t\int_{-\infty}^{\infty} \Phi^{k-t}(z+\tau_{t1})\,[1-\Phi(z)]^{t-1}\, d\Phi(z) = P^*, \tag{3.54}$$

where

$$\tau_{t1} = \frac{n\delta^* - c}{\sqrt{n}\sigma}. \tag{3.55}$$

The second requirement is

$$\inf_{\Omega_t(\delta^*)} P(CS_{ts}|R^t_{n,c}) \geq Q^*, \tag{3.56}$$

Using the selection rule, we get

$$P(CS_{ts}|R^t_{n,c}) = P\left( \min_{k-t+1\leq i\leq k} \sum_{j=1}^{n} Y_{(i)j} \geq \sum_{j=1}^{n} Y_{[k-t+1]j} - c \right). \tag{3.57}$$

For the first requirement it is possible to replace $\sum_{j=1}^{n} Y_{[k-t+1]j}$ with $\max_{1\leq i\leq k-t} \sum_{j=1}^{n} Y_{(i)j}$, because all sample sums corresponding to the $t$ best populations should be larger than all other sample sums, in order to select the $t$ best populations alone. For the second requirement, the probabilities are not equal, using the same replacement, since a correct subset can include a non-best population. However, if we do replace $\sum_{j=1}^{n} Y_{[k-t+1]j}$ with $\max_{1\leq i\leq k-t} \sum_{j=1}^{n} Y_{(i)j}$ we find at least a lower bound of $P(CS_{ts})$, for we have

$$\begin{aligned} P(CS_{ts}|R^t_{n,c}) &\geq P(\min_{k-t+1\leq i\leq k} \sum_{j=1}^{n} Y_{(i)j} \geq \max_{1\leq i\leq k-t} \sum_{j=1}^{n} Y_{(i)j} - c \mid LFC) \\ &= t\, P(\max_{1\leq i\leq k-t} \sum_{j=1}^{n} Y_{(i)j} - c \leq \sum_{j=1}^{n} Y_{(k-t+1)j} \leq \min_{k-t+2\leq i\leq k} \sum_{j=1}^{n} Y_{(i)j}) \\ &= t \int_{-\infty}^{\infty} \Phi^{k-t}(z + \frac{n\delta^* + c}{\sqrt{n}\sigma}) [1-\Phi(z)]^{t-1}\, d\Phi(z). \end{aligned} \tag{3.58}$$

Therefore, (3.56) is satisfied if

$$t \int_{-\infty}^{\infty} \Phi^{k-t}(z + \tau_{t2}) [1-\Phi(z)]^{t-1}\, d\Phi(z) = Q^*, \tag{3.59}$$

with

$$\tau_{t2} = \frac{n\delta^* + c}{\sqrt{n}\sigma}. \tag{3.60}$$

With the tables provided in appendix C.2 we can solve the relations for $n$ and $c$, given $k, P^*, Q^*$ and $\delta^*$, getting these results

$$n = \left( \frac{\sigma(\tau_{t1} + \tau_{t2})}{2\delta^*} \right)^2 \quad \text{and} \quad c = \frac{\sigma^2}{4\delta^*}(\tau^2_{t2} - \tau^2_{t1}). \tag{3.61}$$

We remark that it may be possible to find lower values for $n$ and $c$, when one does not have to use the lower bound for the second requirement to determine a relation for $n$ and $c$. To satisfy both requirements in practice, we have to take the nearest higher integer value for the sample size if $n$ is not an integer. Like before, because of the rounded value of $n$, $\tau_{t1}$ and $\tau_{t2}$ change and we get slightly higher practical values for $P^*$ and $Q^*$.

**Comment on tabled values**

We notice in the table in appendix C.2 that the values for $\tau_{t1}$ are the same for example for the selection of 2 populations out of 5 and the selection of 3 populations out of 5. In

general, the values are the same for the selection of $t$ populations out of $k$ populations and for the selection of $k-t$ populations out of $k$. The following analysis is the explanation for this feature. From (3.54) we see that in that case we should have an equality of the probabilities of correct selections, so

$$t \int_{-\infty}^{\infty} \Phi^{k-t}(z+\tau_{t1}) [1-\Phi(z)]^{t-1} \, d\Phi(z) = (k-t) \int_{-\infty}^{\infty} \Phi^{t}(z+\tau_{t1}) [1-\Phi(z)]^{k-t-1} \, d\Phi(z). \tag{3.62}$$

We can prove this to be true; starting from the left-handside we get (by using partial integration):

$$\begin{aligned}
& t \int_{-\infty}^{\infty} \Phi^{k-t}(z+\tau_{t1}) [\Phi(-z)]^{t-1} \, \phi(z) dz = \\
& = t \int_{-\infty}^{\infty} \Phi^{k-t}(z+\tau_{t1}) [\Phi(-z)]^{t-1} \, \phi(-z) dz \\
& = \int_{-\infty}^{\infty} \Phi^{k-t}(-x+\tau_{t1}) d\Phi^{t}(x) \\
& = - \int_{-\infty}^{\infty} \Phi^{t}(x) d\Phi^{k-t}(-x+\tau_{t1}) \\
& = -(k-t) \int_{-\infty}^{\infty} \Phi^{t}(x) \Phi^{k-t-1}(-x+\tau_{t1}) \phi(-x+\tau_{t1}) d(-x) \\
& = -(k-t) \int_{-\infty}^{\infty} \Phi^{t}(w+\tau_{t1}) \Phi^{k-t-1}(-w) \phi(-w) d(-w-\tau_{t1}) \\
& = (k-t) \int_{-\infty}^{\infty} \Phi^{t}(z+\tau_{t1}) \Phi^{k-t-1}(-z) \phi(z) d(z).
\end{aligned} \tag{3.63}$$

Hence, we have equality and it is proven that the expressions for the probability of a CS are exactly the same when selecting $t$ populations out of $k$ and selecting $k-t$ out of $k$ populations.

### 3.7.2 Second requirement on $\Omega$

The selection of the $t$ best populations without a restriction on the parameter space is based on similar analyses. The first probability requirement remains the same and is satisfied when (3.54) holds. For the second requirement, the parameter space $\Omega$ is now defined by

$$\mu_{[1]} \leq \mu_{[2]} \leq \cdots \leq \mu_{[k]}. \tag{3.64}$$

If we assume to be in the Least Favourable Configuration, see [1] and [7],

$$\mu_{[1]} = \mu_{[2]} = \cdots = \mu_{[k]} \tag{3.65}$$

so all the means as close as possible to each other, we get for the probability of correctly selecting the $t$ best populations in the subset

$$P(CS_{ts}|R^{t\Omega}_{n,c}) \geq P(\min_{k-t+1\leq i\leq k}\sum_{j=1}^{n} Y_{(i)j} \geq \max_{1\leq i\leq k-t}\sum_{j=1}^{n} Y_{(i)j} - c \mid LFC)$$

$$= t\int_{-\infty}^{\infty} \Phi^{k-t}(z + \frac{n\mu_{[k-t+1]} - n\mu_{[k-t]} + c}{\sqrt{n}\sigma})\,[1-\Phi(z)]^{t-1}\,d\Phi(z)$$

$$= t\int_{-\infty}^{\infty} \Phi^{k-t}(z + \frac{c}{\sqrt{n}\sigma})\,[1-\Phi(z)]^{t-1}\,d\Phi(z). \tag{3.66}$$

The minimal value of this is attained for the LFC, hence, the second probability requirement,

$$\inf_{\Omega} P(CS_{ts}|R^{t\Omega}_{n,c}) \geq Q^* \tag{3.67}$$

is satisfied when

$$t\int_{-\infty}^{\infty} \Phi^{k-t}(z + \tau^{\Omega}_{t2})\,[1-\Phi(z)]^{t-1}\,d\Phi(z) = Q^*, \tag{3.68}$$

where

$$\tau^{\Omega}_{t2} = \frac{c}{\sqrt{n}\sigma}. \tag{3.69}$$

With the relations we have now (3.54) and (3.68) we determine $n$ and $c$ by

$$n = (\frac{\sigma(\tau_{t1} + \tau^{\Omega}_{t2})}{\delta^*})^2 \quad \text{and} \quad c = \frac{\sigma^2}{\delta^*}(\tau_{t1} + \tau^{\Omega}_{t2})\tau^{\Omega}_{t2}. \tag{3.70}$$

The form of these functions for $n$ and $c$ are the same as for the selection of 1 best population with no restrictions on the parameter space, so the same as in (3.37) and (3.38). Therefore, the same analysis concerning the comparison of these values with the values when we have a restriction on the parameter space applies, as we have done in the end of section 3.6.

# Chapter 4

# Properties of preference threshold procedure $R_{n,c}$

## 4.1 Expected subset size

Besides our procedure $R_{n,c}$ satisfying the two probability requirements, we would like the size of the selected subset to be small. This size, $S$, is a random variable and $S$ can take the values $1, 2, \ldots, k$. It can be expected that our procedure will produce a smaller size of the selected subset than the standard Gupta subset selection approach, because of our additional requirement of selecting only the best population with at least probability $P^*$. We can express the expected subset size in the following way [18]

$$E(S) = \sum_{i=1}^{k} P(\text{Selecting } \pi_{(i)}). \tag{4.1}$$

For, if we introduce a variable $Z_i$, defined by

$$Z_i = \begin{cases} 0 \text{ if population } i \text{ is not selected }, \\ 1 \text{ if population } i \text{ is selected }. \end{cases} \tag{4.2}$$

then

$$E(S) = E(\sum_{i=1}^{k} Z_i) = \sum_{i=1}^{k} E(Z_i) = \sum_{i=1}^{k} P(\text{Selecting } \pi_{(i)}). \tag{4.3}$$

In order to express the expected size of the selected subset as in (4.1), we first derive the probability of selecting population $\pi_{(i)}$ for procedure $R_{n,c}$. This is

$$\begin{aligned}
P(\text{Selecting } \pi_{(i)}|R_{n,c}) &= P(\sum_{j=1}^{n} Y_{(i)j} \geq \max_{1\leq l\leq k} \sum_{j=1}^{n} Y_{(l)j} - c) \\
&= P(n\bar{Y}_{(l)} \leq n\bar{Y}_{(i)} + c;\ l = 1, \ldots, k; l \neq i) \\
&= \int_{-\infty}^{\infty} P(n\bar{Y}_{(l)} \leq y + c;\ l = 1, \ldots, k; l \neq i) dP(n\bar{Y}_{(i)} \leq y) \\
&= \int_{-\infty}^{\infty} \prod_{\substack{l=1 \\ l\neq i}}^{k} \left[\Phi(\frac{y + c - n\mu_{[l]}}{\sqrt{n}\sigma})\right] d\Phi(\frac{y - n\mu_{[i]}}{\sqrt{n}\sigma}) \\
&= \int_{-\infty}^{\infty} \prod_{\substack{l=1 \\ l\neq i}}^{k} \left[\Phi(z + \frac{n\mu_{[i]} - n\mu_{[l]} + c}{\sqrt{n}\sigma})\right] d\Phi(z). \qquad (4.4)
\end{aligned}$$

Hence the expected size of the selected subset becomes:

$$E(S|R_{n,c}) = \sum_{i=1}^{k} \int_{-\infty}^{\infty} \prod_{\substack{l=1 \\ l\neq i}}^{k} \left[\Phi(z + \frac{n\mu_{[i]} - n\mu_{[l]} + c}{\sqrt{n}\sigma})\right] d\Phi(z). \qquad (4.5)$$

Gupta shows that the maximum value of the expected size of the subset in the preference zone $\Omega(\delta^*)$, will be attained for the LFC (3.6), therefore we have

$$\begin{aligned}
\max_{\Omega(\delta^*)} E(S|R_{n,c}) = &\int_{-\infty}^{\infty} \Phi^{k-1}(z + \frac{n\delta^* + c}{\sqrt{n}\sigma}) d\Phi(z) \\
&+ (k-1) \int_{-\infty}^{\infty} \Phi^{k-2}(z + \frac{c}{\sqrt{n}\sigma}) \Phi(z + \frac{c - n\delta^*}{\sqrt{n}\sigma}) d\Phi(z). \qquad (4.6)
\end{aligned}$$

Due to the second probability requirement (with $n$ not rounded up to the nearest higher integer value) we can rewrite this to get

$$\max_{\Omega(\delta^*)} E(S|R_{n,c}) = Q^* + (k-1) \int_{-\infty}^{\infty} \Phi^{k-2}(z + \frac{c}{\sqrt{n}\sigma}) \Phi(z + \frac{c - n\delta^*}{\sqrt{n}\sigma}) d\Phi(z). \qquad (4.7)$$

It is easy to see that an upperbound for the maximum expected subset size in $\Omega(\delta^*)$ is $kQ^*$, when we impose $Q^*$ on our second probability requirement, because

$$\begin{aligned}
\max_{\Omega(\delta^*)} E(S) &= Q^* + (k-1) \int_{-\infty}^{\infty} \Phi^{k-2}(z + \frac{c}{\sqrt{n}\sigma}) \Phi(z + \frac{c - n\delta^*}{\sqrt{n}\sigma}) d\Phi(z) \\
&\leq Q^* + (k-1) \int_{-\infty}^{\infty} \Phi^{k-2}(z + \frac{n\delta^* + c}{\sqrt{n}\sigma}) \Phi(z + \frac{c + n\delta^*}{\sqrt{n}\sigma}) d\Phi(z) \\
&= kQ^*. \qquad (4.8)
\end{aligned}$$

The maximum of the expected subset size with **no restrictions** on the parameterspace $\Omega$, is attained for $\mu_{[1]} = \cdots = \mu_{[k]}$ [18] and given by

$$\max_{\Omega} E(S|R_{n,c}) = k \int_{-\infty}^{\infty} \Phi^{k-1}(z + \frac{c}{\sqrt{n}\sigma}) d\Phi(z) = kQ^*. \qquad (4.9)$$

We see that the former upperbound $kQ^*$ is now actually achieved.

The results Gupta presented [18] are analogous, with our $c$ threshold replaced by his distance $d$. He found that the maximum of his expected subset size over $\Omega$ using rule $R_G$ as described in section 2.2 is attained for $\mu_{[1]} = \mu_{[2]} = \cdots = \mu_{[k]}$ and

$$\max_{\Omega} E(S|R_G) = k \int_{-\infty}^{\infty} \Phi^{k-1}(z + \frac{d}{\sqrt{n}\sigma}) d\Phi(z) = kQ^*. \tag{4.10}$$

Gupta also derived the expected subset size assuming that we are in the preference zone, $\Omega(\delta^*)$. In this case the maximum of the expected subset size is attained at the LFC (3.6) and becomes

$$\max_{\Omega(\delta^*)} E(S|R_G) = \int_{-\infty}^{\infty} \Phi^{k-1}(z + \frac{d+n\delta^*}{\sqrt{n}\sigma}) d\Phi(z) + (k-1) \int_{-\infty}^{\infty} \Phi^{k-2}(z + \frac{d}{\sqrt{n}\sigma}) \Phi(z + \frac{d-n\delta^*}{\sqrt{n}\sigma}) d\Phi(z). \tag{4.11}$$

### 4.1.1 Comparison of expected subset sizes

It is interesting to compare the expected size of the selected subset for the standard subset selection approach to that of our preference threshold procedure. However, the only sensible comparison can be made when the sample size $n$ is equal for both procedures. Hence, we specify $P^*$ and $Q^*$ for respectively the first and second probability requirement of the preference threshold procedure, and we specify $\delta^*$. We can now calculate the minimum required sample size $n$ to satisfy both requirements for the preference threshold procedure $R_{n,c}$, because we assume to know the value of the common variance $\sigma^2$. For the standard subset selection procedure, we impose the same value as $Q^*$ for the probability requirement of including the best population in the subset, implying $\tau_2 = \tau_G$ (sections 3.3 and 2.2). We can calculate the distance $d$ according to Gupta's approach. Remember $1/k < P^* \leq Q^* < 1$ and $\tau_2 \geq \tau_1$. Since

$$\frac{c + n\delta^*}{\sqrt{n}\sigma} = \frac{d}{\sqrt{n}\sigma} \tag{4.12}$$

when taking the same value of $Q^*$ in the second requirement of our procedure $R_{n,c}$ as the value for the requirement in the standard subset selection approach, we have, taking the same sample size $n$ ($> 0$) for both procedures,

$$\begin{aligned} E(S|R_{n,c}) &= \sum_{i=1}^{k} \int_{-\infty}^{\infty} \prod_{\substack{l=1 \\ l \neq i}}^{k} \left[ \Phi(z + \frac{n\mu_{[i]} - n\mu_{[l]} + c}{\sqrt{n}\sigma}) \right] d\Phi(z) \\ &< \sum_{i=1}^{k} \int_{-\infty}^{\infty} \prod_{\substack{l=1 \\ l \neq i}}^{k} \left[ \Phi(z + \frac{n\mu_{[i]} - n\mu_{[l]} + d}{\sqrt{n}\sigma}) \right] d\Phi(z) \\ &= E(S|R_G). \end{aligned} \tag{4.13}$$

Of course, we are considering the same $\mu$ values. Obviously, the same holds for the maximum of the subset sizes, so

$$\max_{\Omega(\delta^*)} E(S|R_{n,c}) < \max_{\Omega} E(S|R_G). \tag{4.14}$$

Intuitively, this result can be justified by noting that $n\delta^* > 0$ (due to our assumption of being in the preference zone) implies $c < d$ according to (4.12) and hence the selection threshold in our approach is smaller than in Gupta's approach, so we select less or an equal number of populations.

If we assume no restrictions for the parameter space to satisfy the second requirement of procedure $R_{n,c}$, then, for the same value of $Q^*$ and the same sample sizes $n$ for both procedures,

$$\frac{c}{\sqrt{n}\sigma} = \frac{d}{\sqrt{n}\sigma} \tag{4.15}$$

holds, according to (2.15) and (3.34). This results in

$$E(S|R_{n,c}) = E(S|R_G). \tag{4.16}$$

Therefore, in case of no restriction on the parameter space to satisfy the second probability requirement of $R_{n,c}$ and equal $Q^*$ values, for the determined sample size $n$ both procedures yield exactly the same subset, thus the same expected subset size. However, it should be noted that using the preference threshold procedure, we can determine the sample size $n$, where this is given for the standard subset selection procedure.

## 4.2 Probability of correct selection given that one population has been selected

When we apply the preference threshold procedure, we only select a single population if its sample sum has 'enough' distance to all other sample sums. Due to our first requirement we design the experiment in such a way that the probability of correctly selecting a single population is at least $P^*$. However, after the experiment has been carried out and we have selected a single population, it is interesting to be able to make a statement about the probability of a correct selection *given* that only one population has been selected. Obviously this probability should be greater than $P^*$. We need to consider the following probability, where $S_1$ denotes the selection of one single population, $CS_1$ one correctly selected single population and $FS_1$ a false selection of a single population;

$$\begin{aligned} P(CS \mid S_1) &= \frac{P(CS \cap S_1)}{P(S_1)} \\ &= \frac{P(CS_1)}{P(S_1)} \end{aligned} \tag{4.17}$$

with

$$P(S_1) = P(CS_1) + P(FS_1). \tag{4.18}$$

Therefore, using $P_\mu(CS)$ with subscript $\mu$ to stress the dependence on $\mu$,

$$P(CS|S_1) = \frac{P_\mu(CS_1)}{P_\mu(CS_1) + P_\mu(FS_1)}. \tag{4.19}$$

We take products over an empty domain to be equal to 1 by definition and sums over an empty domain are obviously equal to zero. Remember the probability of a Correct Selection of a single population is expressed (3.5)

$$P_\mu(CS_1) = \int_{-\infty}^{\infty} \prod_{i=1}^{k-1} \left[\Phi(z + \frac{n\mu_{[k]} - n\mu_{[i]} - c}{\sqrt{n}\sigma})\right] d\Phi(z) \tag{4.20}$$

and for the probability of a False Selection of one population we get

$$P_\mu(FS_1) = \sum_{l=1}^{k-1} P(\sum_{j=1}^{n} Y_{(l)j} > \sum_{j=1}^{n} Y_{(i)j} + c,\ i = 1, \ldots, k,\ i \neq l)$$
$$= \sum_{l=1}^{k-1} \int_{-\infty}^{\infty} \prod_{\substack{i=1 \\ i \neq l}}^{k} \Phi(z + \frac{n\mu_{[l]} - n\mu_{[i]} - c}{\sqrt{n}\sigma}) d\Phi(z). \tag{4.21}$$

### 4.2.1 Local minimum in the Least Favourable Configuration

In this subsection we will prove that the probability of a CS given that only one population has been selected attains a local minimum in the LFC.
So, we want to prove:

$$P(CS \mid S_1) \textit{ attains a local minimum in the LFC.}$$

Let us define the differences between the actual means as

$$\begin{aligned}
&\mu_{[k]} \\
&\mu_{[k-1]} = \mu_{[k]} - \delta^* - \epsilon_{k-1} \\
&\mu_{[k-2]} = \mu_{[k-1]} - \epsilon_{k-2} \\
&\mu_{[k-3]} = \mu_{[k-1]} - \epsilon_{k-2} - \epsilon_{k-3} \\
&\quad \vdots \\
&\mu_{[1]} = \mu_{[k-1]} - \epsilon_{k-2} - \cdots - \epsilon_1, \quad \text{where } \epsilon_j \geq 0 \text{ for all } j = 1, \ldots, k-1,
\end{aligned} \tag{4.22}$$

so $\mu_{[i]} = \mu_{[k]} - \delta^* - \sum_{j=i}^{k-1} \epsilon_j$. We rewrite the expressions for $P_\mu(CS_1)$ and $P_\mu(FS_1)$ and take the partial derivative of these expressions to $\epsilon_j$. First, the probability of a correct selection of one population is

$$P_\mu(CS_1) = \int_{-\infty}^{\infty} \prod_{i=1}^{k-1} \left[ \Phi(z + \frac{n(\epsilon_i + \cdots + \epsilon_{k-1} + \delta^*) - c}{\sqrt{n}\sigma}) \right] \phi(z) dz \tag{4.23}$$

and its partial derivative with respect to $\epsilon_j$ is (see appendix B)

$$\frac{\partial P_\mu(CS_1)}{\partial \epsilon_j} =$$
$$\frac{\sqrt{n}}{\sigma} \int_{-\infty}^{\infty} \sum_{m=1}^{j} \left[ \phi(z + \frac{n(\epsilon_m + \cdots + \epsilon_{k-1} + \delta^*) - c}{\sqrt{n}\sigma}) \prod_{\substack{i=1 \\ i \neq m}}^{k-1} \Phi(z + \frac{n(\epsilon_i + \cdots + \epsilon_{k-1} + \delta^*) - c}{\sqrt{n}\sigma}) \right] \phi(z) dz. \tag{4.24}$$

For $\phi(\cdot)$ and $\Phi(\cdot)$ are positive, it is easy to see that we have

$$\frac{\partial P_\mu(CS_1)}{\partial \epsilon_j} > 0 \text{ for all } j = 1, \ldots, k-1. \tag{4.25}$$

This implies that the larger the distances between the means become, the larger the probability of a $CS_1$ becomes, which is logical. The probability of a False Selection

of one population (4.21) can be rewritten in this way

$$P_\mu(FS_1) = \sum_{l=1}^{k-1} \int_{-\infty}^{\infty} \prod_{i=1}^{l-1} \left[\Phi(z + \frac{n(\mu_{[l]} - \mu_{[i]}) - c}{\sqrt{n}\sigma})\right] \prod_{i=l+1}^{k-1} \left[\Phi(z + \frac{n(\mu_{[l]} - \mu_{[i]}) - c}{\sqrt{n}\sigma})\right] *$$
$$\Phi(z + \frac{n(\mu_{[l]} - \mu_{[k]}) - c}{\sqrt{n}\sigma})\phi(z)dz$$
$$= \sum_{l=1}^{k-1} \int_{-\infty}^{\infty} \prod_{i=1}^{l-1} \left[\Phi(z + \frac{n(\epsilon_i + \cdots + \epsilon_{l-1}) - c}{\sqrt{n}\sigma})\right] \prod_{i=l}^{k-2} \left[\Phi(z - \frac{n(\epsilon_l + \cdots + \epsilon_i) + c}{\sqrt{n}\sigma})\right] *$$
$$\Phi(z - \frac{n(\epsilon_l + \cdots + \epsilon_{k-1} + \delta^*) + c}{\sqrt{n}\sigma})\phi(z)dz \tag{4.26}$$

and the partial derivative of this expression to $\epsilon_j$ is (see appendix B)

$$\frac{\partial P_\mu(FS_1)}{\partial \epsilon_j} = -\frac{\sqrt{n}}{\sigma} \sum_{l=1}^{j} \left( \int_{-\infty}^{\infty} \prod_{i=1}^{l-1} \Phi(z + \frac{n(\epsilon_i + \cdots + \epsilon_{l-1}) - c}{\sqrt{n}\sigma}) * \prod_{i=l}^{j-1} \Phi(z - \frac{n(\epsilon_l + \cdots + \epsilon_i) + c}{\sqrt{n}\sigma}) * \right.$$
$$\left\{ \sum_{m=j}^{k-2} \left[ \phi(z - \frac{n(\epsilon_l + \cdots + \epsilon_m) + c}{\sqrt{n}\sigma}) \prod_{\substack{i=j \\ i \neq m}}^{k-2} \Phi(z - \frac{n(\epsilon_l + \cdots + \epsilon_i) + c}{\sqrt{n}\sigma}) \right] * \right.$$
$$\Phi(z - \frac{n(\epsilon_l + \cdots + \epsilon_{k-1} + \delta^*) + c}{\sqrt{n}\sigma}) +$$
$$\left. \left. \prod_{i=j}^{k-2} \Phi(z - \frac{n(\epsilon_l + \ldots \epsilon_i) + c}{\sqrt{n}\sigma}) \phi(z - \frac{n(\epsilon_l + \cdots + \epsilon_{k-1} + \delta^*) + c}{\sqrt{n}\sigma}) \right\} \phi(z)dz \right)$$
$$+$$
$$\frac{\sqrt{n}}{\sigma} \sum_{l=j+1}^{k-1} \left( \int_{-\infty}^{\infty} \sum_{m=1}^{j} \left[ \phi(z + \frac{n(\epsilon_m + \cdots + \epsilon_{l-1}) - c}{\sqrt{n}\sigma}) \prod_{\substack{i=1 \\ i \neq m}}^{j} \Phi(z + \frac{n(\epsilon_i + \cdots + \epsilon_{l-1}) - c}{\sqrt{n}\sigma}) \right] * \right.$$
$$\prod_{i=j+1}^{l-1} \Phi(z + \frac{n(\epsilon_i + \cdots + \epsilon_{l-1}) - c}{\sqrt{n}\sigma}) \prod_{i=l}^{k-2} \Phi(z - \frac{n(\epsilon_l + \cdots + \epsilon_i) + c}{\sqrt{n}\sigma}) *$$
$$\left. \Phi(z - \frac{n(\epsilon_l + \cdots + \epsilon_{k-1} + \delta^*) + c}{\sqrt{n}\sigma}) \phi(z)dz \right). \tag{4.27}$$

In the LFC, $\mu_{[1]} = \cdots = \mu_{[k-1]} = \mu_{[k]} - \delta^*$, we have $\epsilon_j = 0$, $j = 1, \ldots, k-1$, and the derivatives become

$$\frac{\partial P_\mu(FS_1)}{\partial \epsilon_j}|_{LFC} = -\frac{\sqrt{n}}{\sigma} j(k-j-1) \int_{-\infty}^{\infty} \phi(z - \frac{c}{\sqrt{n}\sigma}) \Phi^{k-3}(z - \frac{c}{\sqrt{n}\sigma}) \Phi(z - \frac{n\delta^* + c}{\sqrt{n}\sigma}) \phi(z)dz$$
$$- \frac{\sqrt{n}}{\sigma} j \int_{-\infty}^{\infty} \Phi^{k-2}(z - \frac{c}{\sqrt{n}\sigma}) \phi(z - \frac{n\delta^* + c}{\sqrt{n}\sigma}) \phi(z)dz$$
$$+ \frac{\sqrt{n}}{\sigma} j(k-j-1) \int_{-\infty}^{\infty} \phi(z - \frac{c}{\sqrt{n}\sigma}) \Phi^{k-3}(z - \frac{c}{\sqrt{n}\sigma}) \Phi(z - \frac{n\delta^* + c}{\sqrt{n}\sigma}) \phi(z)dz$$
$$= -\frac{\sqrt{n}}{\sigma} j \int_{-\infty}^{\infty} \Phi^{k-2}(z - \frac{c}{\sqrt{n}\sigma}) \phi(z - \frac{n\delta^* + c}{\sqrt{n}\sigma}) \phi(z)dz. \tag{4.28}$$

Therefore, the partial derivative of $P_\mu(FS_1)$ for each $\epsilon_j$ is strictly less than zero in the LFC ($\Phi(\cdot) > 0$ and $\phi(\cdot) > 0$):

$$\frac{\partial P_\mu(FS_1)}{\partial \epsilon_j}|_{LFC} < 0 \qquad \text{for } j = 1, \ldots, k-1. \tag{4.29}$$

This means that the gradient

$$\nabla P_{LFC}(FS_1) < 0 \qquad \text{(in every element of this vector)} \tag{4.30}$$

and this is a sufficient condition for a **local maximum** of $P_\mu(FS_1)$ in the LFC. From (4.25) we have that $P_\mu(CS_1)$ attains a global minimum in the LFC. The probability of a correct selection given that one population has been selected (4.19) is of the form $f(x,y) = \frac{x}{x+y}$ where $0 < x, y \leq 1$. For this function we have

$$\frac{\partial f(x,y)}{\partial x} = \frac{y}{(x+y)^2} > 0 \quad \text{and} \quad \frac{\partial f(x,y)}{\partial y} = \frac{-x}{(x+y)^2} < 0, \tag{4.31}$$

hence $f(x,y)$ is increasing in $x$ and decreasing in $y$. Substituting $P_\mu(CS_1)$ for $x$ and $P_\mu(FS_1)$ for $y$, we conclude that the probability of a Correct Selection given that one population has been selected (4.19) attains a **local minimum** in the LFC.

### 4.2.2 Necessary condition for global minimum

Intuitively, if we start in the LFC, and increase the difference between the largest and second-largest mean, by increasing $\epsilon_{k-1}$ and thus also increasing the difference between the best and the other means, we expect the probability of a False Selection to decrease, because it becomes 'harder' for the non-best populations to beat the best one. We will justify this in this subsection. The difference between the largest and second-largest mean in $\Omega(\delta^*)$ is given by $\epsilon_{k-1}$, see (4.22). Taking the derivative to $\epsilon_{k-1}$ of $P_\mu(FS_1)$ yields (4.27)

$$\frac{\partial P_\mu(FS_1)}{\partial \epsilon_{k-1}} = -\frac{\sqrt{n}}{\sigma} \sum_{l=1}^{k-1} \int_{-\infty}^{\infty} \prod_{\substack{i=1 \\ i \neq l}}^{l} \Phi(z + \frac{n(\epsilon_i + \cdots + \epsilon_{l-1}) - c}{\sqrt{n}\sigma}) *$$

$$\prod_{i=l+1}^{k-1} \Phi(z - \frac{n(\epsilon_l + \cdots + \epsilon_{i-1}) + c}{\sqrt{n}\sigma}) \phi(z - \frac{n(\epsilon_l + \cdots + \epsilon_{k-1} + \delta^*) + c}{\sqrt{n}\sigma}) \phi(z) dz. \tag{4.32}$$

It is easy to see ($\Phi(\cdot) > 0$ and $\phi(\cdot) > 0$) that

$$\frac{\partial P_\mu(FS_1)}{\partial \epsilon_{k-1}} < 0, \text{ for all } \epsilon_i \geq 0,\ i = 1, \ldots, k-1. \tag{4.33}$$

Therefore a necessary condition for a global maximum of $P_\mu(FS_1)$ is

$$\epsilon_{k-1} = 0 \tag{4.34}$$

and because $\epsilon_{k-1} = 0$ is also a necessary condition for a global minimum of $P_\mu(CS_1)$, according to (4.25), by using the same feature of $P(CS|S_1)$ as before (4.31), we see that $\epsilon_{k-1} = 0$ is a necessary condition for a *global minimum* of $P(CS|S_1)$.

Because of the probability of a false selection of one single population not having a nice monotonous behaviour ([12]), it is difficult to find the global minimum of $P(CS|S_1)$. This would be very interesting, we could use this as a alternative second requirement, by imposing a probability $Q^*$ on this global minimum. We have the conjecture that the *global* minimum of the $P(CS|S_1)$ is attained in the LFC, where $\epsilon_i = 0$, $(i = 1, \ldots, k-1)$, however this remained unproven and can be a topic for future research.

# Chapter 5

# Robustness

## 5.1 Varying variances

Both Bechhofer's and Gupta's approach assume equal known variances for all populations, but in practice this is usually not the actual situation. It has been shown by Driessen, Van der Laan and Van Putten [14] that departures from the assumption of a common known variance in the case of normal populations for the Indifference Zone approach and the Subset Selection approach, can cause an actual lower bound of the probability of a Correct Selection that is seriously lower than the pretended lower bound $P^*$ of the probability of CS based on the assumption of a common known variance.

Some procedures have been presented in the literature to deal with the problem of selecting the best population out of $k$ normal populations with unknown variances, both with and without the assumption of equality of these variances. When the variances of the $k$ populations are common but unknown, a two-stage procedure has been suggested by Bechhofer, Dunnett and Sobel in [2]. Also, a sequential procedure based on the work of Robbins, Sobel and Starr in [26] has been developed to tackle this problem. For unequal and unknown variances of the $k$ populations a two-stage procedure can be applied as proposed by Dudewicz and Dalal [15]. We will summarize the procedures as they are presented in [16]. The two-stage procedure for common but unknown variances consists of taking a first sample of size $n$ from each population and computing the pooled sample variance. After that, in the second stage, another sample of size $N - n$ from each population is taken, where N is depending on $n$, on the pooled variance as estimated from the first stage (that is treated like it were the actual common population variance),and on $k, \delta^*$ and $P^*$. Then the observations of the two stages are combined in a single sample and the sample with the largest sample mean is considered to be the best one.

Another possible procedure is the sequential procedure, where in each stage one observation of each population is taken and after each stage $r$, the pooled sample variance $s_r^2$ is computed until we reach a stage $N$ for which $s_N^2$ is smaller than a certain value, depending on $N, \delta^*, P^*$ and $k$. At that moment, the sample with the largest sample mean is asserted to be the best one.

Assuming nothing about the variances, so unequal and unknown variances, the two-stage procedure consists of taking a sample of size $n$ in the first stage, computing the $k$ sample means and sample variances $s_j^2$. In the second stage a second sample of size $N_j - n$ is taken from the $j$th population, where $N_j$ is depending on $n, s_j^2, \delta^*, P^*$ and $k$. Based on a weighted mean of the first and second stage means for each population, the largest one of these is asserted to correspond to the best population.

In this chapter, we show the results for the robustness of the probability of a CS against deviations from the assumption of a common known variance, in a similar manner as in [14]. We consider again our problem of selecting the best population out of $k$ independent normally distributed populations with unknown means $\mu_i$, but we allow the standard deviations $\sigma_1, \sigma_2, \ldots, \sigma_k$ of populations $\pi_1, \ldots, \pi_k$ to vary in the interval $[\gamma^{-1}\sigma_0, \gamma\sigma_0]$ $(\gamma \geq 1)$ where $\sigma_0$ is a chosen value. For ease of notation we write $\sigma = (\sigma_1, \sigma_2, \ldots, \sigma_k)$ and by $G = G(\gamma, \sigma_0)$ we denote the set $[\gamma^{-1}\sigma_0, \gamma\sigma_0]^k \in \mathbb{R}^k$, so $\sigma \in G$.

Using sample sums as response variables, the probability of correctly selecting one population related to our procedure $R_{n,c}$ (corresponding to the first requirement) is

$$\begin{aligned}
P(CS_1|R_{n,c},\sigma) &= P(\sum_{j=1}^{n} Y_{(k)j} > \sum_{j=1}^{n} Y_{(i)j} + c;\ i = 1, \ldots, k-1) \\
&= \int_{-\infty}^{\infty} P(n\bar{Y}_{(i)} < y - c;\ i = 1, \ldots, k-1) dP(n\bar{Y}_{(k)} \leq y) \\
&= \int_{-\infty}^{\infty} \prod_{i=1}^{k-1} \left[ \Phi(\frac{y - c - n\mu_{[i]}}{\sqrt{n}\sigma_i}) \right] d\Phi(\frac{y - n\mu_{[k]}}{\sqrt{n}\sigma_k}) \\
&= \int_{-\infty}^{\infty} \prod_{i=1}^{k-1} \left[ \Phi(\frac{y - n\mu_{[k]} + n\mu_{[k]} - n\mu_{[i]} - c}{\sqrt{n}\sigma_k} * \frac{\sigma_k}{\sigma_i}) \right] d\Phi(\frac{y - n\mu_{[k]}}{\sqrt{n}\sigma_k}) \\
&= \int_{-\infty}^{\infty} \prod_{i=1}^{k-1} \left[ \Phi(z\frac{\sigma_k}{\sigma_i} + \frac{n\mu_{[k]} - n\mu_{[i]} - c}{\sqrt{n}\sigma_i}) \right] d\Phi(z). \qquad (5.1)
\end{aligned}$$

We define the loss function as proposed by Driessen *et al.* [14] to measure the robustness under investigation. This loss is a function of $k, P^*, \gamma$ and $\sigma_0$ and defined by

$$\text{loss}(k, P^*, \gamma, \sigma_0) = \min_{\Omega(\delta^*)} P(CS_1 \mid (\sigma_0, \ldots, \sigma_0)) - \min_{\Omega(\delta^*), \sigma \in G} P(CS_1 \mid \sigma). \qquad (5.2)$$

In words, it is the loss of the lower bound of $P(CS_1)$ for $\mu \in \Omega(\delta^*)$ and $\sigma = (\sigma_0, \sigma_0, \ldots, \sigma_0)$ compared to the lower bound of $P(CS_1 \mid \sigma)$ for $\mu \in \Omega(\delta^*)$ and $\sigma \in G$. For fixed $\sigma \in G$, the same LFC, $\mu_{[1]} = \cdots = \mu_{[k-1]} = \mu_{[k]} - \delta^*$, as in the problem with common known variances applies, resulting in the minimum probability of $P(CS_1 \mid \sigma)$ (again using that $\Phi(\cdot)$ is increasing)

$$\begin{aligned}
\min_{\Omega(\delta^*), \sigma \in G} P(CS_1 \mid \sigma) &= \min_{\sigma \in G} P(CS_1 \mid \sigma, LFC) \\
&= \min_{\sigma \in G} \int_{-\infty}^{\infty} \prod_{i=1}^{k-1} \left[ \Phi(z\frac{\sigma_k}{\sigma_i} + \frac{n\delta^* - c}{\sqrt{n}\sigma_i}) \right] d\Phi(z) \qquad (5.3)
\end{aligned}$$

and because $P(CS_1 \mid \sigma, LFC)$ is a decreasing function of $\sigma_1, \sigma_2, \ldots, \sigma_{k-1}$ (Driessen [14] and Tong and Wetzell [29] [1])

$$\begin{aligned}
\min_{\sigma \in G} P(CS_1 \mid \sigma, LFC) &= \min_{\sigma_k \in [\gamma^{-1}\sigma_0, \gamma\sigma_0]} P(CS_1 \mid (\gamma\sigma_0, \ldots, \gamma\sigma_0, \sigma_k), LFC) \\
&= \min_{\sigma_k \in [\gamma^{-1}\sigma_0, \gamma\sigma_0]} \int_{-\infty}^{\infty} \Phi^{k-1}(\frac{z\sigma_k}{\gamma\sigma_0} + \frac{n\delta^* - c}{\sqrt{n}\gamma\sigma_0}) d\Phi(z) \qquad (5.4)
\end{aligned}$$

[1] Tong and Wetzell [29] show, by using an application of Slepian's inequality, that $P(CS_1 \mid \sigma)$ is decreasing for $\tau_j$, $j \neq k$, defined by $\tau_i^2 = \sigma_i^2/n_i$ ; $i = 1, \ldots, k$, when the other $\tau_i$, $i \neq j$ are kept constant. Hence in our case, with sample sizes $n$ fixed, we derive that $P(CS_1 \mid \sigma)$ is decreasing in $\sigma_i$, $i \neq k$.

The expression for $P(CS_1 \mid \sigma, LFC)$ is not a monotonous function for $\sigma_k$ as Driessen *et al.* point out in [14] and $P(CS_1 \mid (\gamma\sigma_0, \ldots, \gamma\sigma_0, \sigma_k), LFC)$ is difficult to analyze. For the defined loss function, these results lead to

$$\begin{aligned}
\text{loss}(k, P^*, \gamma, \sigma_0) &= \min_{\Omega(\delta^*)} P(CS_1 \mid (\sigma_0, \ldots, \sigma_0)) - \min_{\Omega(\delta^*), \sigma \in G} P(CS_1 \mid \sigma) \\
&= P(CS_1 \mid (\sigma_0, \ldots, \sigma_0), LFC) - \min_{\sigma_k \in [\gamma^{-1}\sigma_0, \gamma\sigma_0]} \int_{-\infty}^{\infty} \Phi^{k-1}(\frac{z\sigma_k}{\gamma\sigma_0} + \frac{n\delta^* - c}{\sqrt{n}\gamma\sigma_0}) d\Phi(z) \\
&= P^* - \min_{\sigma_k \in [\gamma^{-1}\sigma_0, \gamma\sigma_0]} \int_{-\infty}^{\infty} \Phi^{k-1}(\frac{z\sigma_k}{\gamma\sigma_0} + \frac{n\delta^* - c}{\sqrt{n}\gamma\sigma_0}) d\Phi(z) \\
&= P^* - \min_{\gamma^{-2} \le s \le 1} \int_{-\infty}^{\infty} \Phi^{k-1}(sz + \gamma^{-1}\frac{n\delta^* - c}{\sqrt{n}\sigma_0}) d\Phi(z) \\
&= P^* - \min_{\gamma^{-2} \le s \le 1} \int_{-\infty}^{\infty} \Phi^{k-1}(\gamma^{-1}\tau_{0,1} + sz) d\Phi(z) \qquad (5.5)
\end{aligned}$$

where

$$\tau_{0,1} = \frac{n\delta^* - c}{\sqrt{n}\sigma_0} \text{ and determined by } \int_{-\infty}^{\infty} \Phi^{k-1}(y + \tau_{0,1}) d\Phi(y) = P^*.$$

An interesting conclusion is that $\text{loss}(k, P^*, \gamma, \sigma_0)$ is not depending on $\sigma_0$. Thus we can say $\text{loss}(k, P^*, \gamma, \sigma_0) = \text{loss}(k, P^*, \gamma)$. Result (5.5) is of the same functional form as Driessen *et al.* found, so we can use the tables and figures they provided in [14]. These figures we present in appendix D. Like them, we will also consider the subset selection approach, particularly for our procedure, that is the probability of a correct selection by including the best population in the selected subset (we have referred to this event by $CS_2$). Remember our assumption $\mu_{[1]} \le \mu_{[2]} \le \cdots \le \mu_{[k]} - \delta^*$. If again we allow the variances to be in the interval $[\gamma^{-1}\sigma_0, \gamma\sigma_0]$, thus $\sigma \in G$, we find

$$\begin{aligned}
P(CS_2 | R_{n,c}, \sigma) &= P(\sum_{j=1}^{n} Y_{(k)j} \ge \sum_{j=1}^{n} Y_{(i)j} - c;\ i = 1, \ldots, k-1) \\
&= \int_{-\infty}^{\infty} P(n\bar{Y}_{(i)} \le y + c;\ i = 1, \ldots, k-1) dP(n\bar{Y}_{(k)} \le y) \\
&= \int_{-\infty}^{\infty} \prod_{i=1}^{k-1} \left[ \Phi(\frac{y + c - n\mu_{[i]}}{\sqrt{n}\sigma_i}) \right] d\Phi(\frac{y - n\mu_{[k]}}{\sqrt{n}\sigma_k}) \\
&= \int_{-\infty}^{\infty} \prod_{i=1}^{k-1} \left[ \Phi(\frac{y - n\mu_{[k]} + n\mu_{[k]} - n\mu_{[i]} + c}{\sqrt{n}\sigma_k} * \frac{\sigma_k}{\sigma_i}) \right] d\Phi(\frac{y - n\mu_{[k]}}{\sqrt{n}\sigma_k}) \\
&= \int_{-\infty}^{\infty} \prod_{i=1}^{k-1} \left[ \Phi(z\frac{\sigma_k}{\sigma_i} + \frac{n\mu_{[k]} - n\mu_{[i]} + c}{\sqrt{n}\sigma_i}) \right] d\Phi(z). \qquad (5.6)
\end{aligned}$$

The minimum probability of $P(CS_2 \mid \sigma)$ over $\Omega(\delta^*)$ is attained for $\mu$ in the LFC, because

$\Phi(\cdot)$ is increasing, resulting in

$$\min_{\Omega(\delta^*),\sigma\in G} P(CS_2 \mid \sigma) = \min_{\sigma\in G} P(CS_2 \mid \sigma, LFC)$$
$$= \min_{\sigma\in G} \int_{-\infty}^{\infty} \prod_{i=1}^{k-1} \left[\Phi(z\frac{\sigma_k}{\sigma_i} + \frac{n\delta^* + c}{\sqrt{n}\sigma_i})\right] d\Phi(z). \tag{5.7}$$

This function is decreasing for $\sigma_1, \ldots, \sigma_{k-1}$ as for (5.3) and therefore

$$\min_{\sigma\in G} P(CS_2 \mid \sigma, LFC) = \min_{\sigma_k\in[\gamma^{-1}\sigma_0,\gamma\sigma_0]} \int_{-\infty}^{\infty} \Phi^{k-1}(\frac{z\sigma_k}{\gamma\sigma_0} + \frac{n\delta^* + c}{\sqrt{n}\gamma\sigma_0}) d\Phi(z). \tag{5.8}$$

The definition of the loss function will be analogous to (5.2)

$$\text{loss}(k, Q^*, \gamma, \sigma_0) = \min_{\Omega(\delta^*)} P(CS_2 \mid (\sigma_0, \ldots, \sigma_0)) - \min_{\Omega(\delta^*),\sigma\in G} P(CS_2 \mid \sigma) \tag{5.9}$$

and using the derived results gives

$$\text{loss}(k, Q^*, \gamma, \sigma_0) = \min_{\Omega(\delta^*)} P(CS_2 \mid (\sigma_0, \ldots, \sigma_0)) - \min_{\Omega(\delta^*),\sigma\in G} P(CS_2 \mid \sigma)$$
$$= P(CS_2 \mid (\sigma_0, \ldots, \sigma_0), LFC) - \min_{\sigma_k\in[\gamma^{-1}\sigma_0,\gamma\sigma_0]} \int_{-\infty}^{\infty} \Phi^{k-1}(\frac{z\sigma_k}{\gamma\sigma_0} + \frac{n\delta^* + c}{\sqrt{n}\gamma\sigma_0}) d\Phi(z)$$
$$= Q^* - \min_{\gamma^{-2}\le s\le 1} \int_{-\infty}^{\infty} \Phi^{k-1}(\gamma^{-1}\tau_{0,2} + sz) d\Phi(z), \tag{5.10}$$

where

$$\tau_{0,2} = \frac{n\delta^* + c}{\sqrt{n}\sigma_0} \text{ and determined by } \int_{-\infty}^{\infty} \Phi^{k-1}(y + \tau_{0,2}) d\Phi(y) = Q^*.$$

The loss (5.10) has the same functional form as (5.5), so we can use the same figures provided in [14] and reproduced in appendix D. They show that the loss can be quite high, especially for a large number of population, such as $k = 50$ or $k = 100$.

## 5.2 Known variance $\sigma_k$, others in interval

As the minimum of the probability of a correct selection where all standard deviations are allowed to vary in the interval $[\gamma^{-1}\sigma_0, \gamma\sigma_0]$ is hard to be found numerically, Driessen *et al* used a discretization procedure. In both cases, the Indifference Zone approach and the Subset Selection approach, we ended up with the same form of probability of a Correct Selection ($CS_1$ or $CS_2$) in the LFC,

$$\min_{\sigma\in G} P(CS \mid \sigma) = \min_{\sigma_k\in[\gamma^{-1}\sigma_0,\gamma\sigma_0]} \int_{-\infty}^{\infty} \Phi^{k-1}(\frac{z\sigma_k}{\gamma\sigma_0} + \gamma^{-1}\tau_0) d\Phi(z), \tag{5.11}$$

where according to considering the Indifference Zone approach or the Subset Selection approach, respectively

$$\tau_0 = \tau_{0,1} = \frac{n\delta^* - c}{\sqrt{n}\sigma_0} \quad \text{or} \quad \tau_0 = \tau_{0,2} = \frac{n\delta^* + c}{\sqrt{n}\sigma_0}. \tag{5.12}$$

To calculate the value of this minimum to use it in the loss function, Driessen used a discretization procedure, and although this is not exact, it yielded a lower bound for the loss. If we assume the value of $\sigma_k$ to be known, which is still a considerable smaller requirement than the assumption of knowing all $\sigma$ values like we did before, we can solve the problem numerically. So, we assume $\sigma_k = \sigma_0$, that is all the other standard deviations are allowed to vary around $\sigma_k$. We can also regard at this model as by allowing all variances to vary around some kind of a pivot value $\sigma_0$ and only assume $\sigma_k = \sigma_0$. We did consider some instances (10 populations, $P^* = 0.6, 0.75, 0.9$) where *all* (including $\sigma_k$) standard deviations could vary in $[\gamma^{-1}\sigma_0, \gamma\sigma_0]$ and by a discretization procedure (at 20 points with equal distance from each other), we found that the largest value for $\tau_0$ with the righthandside of (5.11) equal to $P^*$ were always attained for $\sigma_k = \gamma\sigma_0$ and thus for all $\sigma_i = \gamma\sigma_0$, $i = 1, \ldots, k$. In that case, we can apply the normal preference threshold procedure and use table C.1 and just take common standard deviation $\gamma\sigma_0$ to calculate $n$ and $c$.

In *this* model, we assume $\sigma_k = \sigma_0$ and therefore (5.11) reduces to

$$P(CS \mid \sigma) = \int_{-\infty}^{\infty} \Phi^{k-1}(\gamma^{-1}z + \gamma^{-1}\tau_0)d\Phi(z). \tag{5.13}$$

So, if we want our procedure to satisfy the two requirements that we defined in section 3.2, we have to look for the values of $\tau_0$ satisfying

$$\int_{-\infty}^{\infty} \Phi^{k-1}(\gamma^{-1}z + \gamma^{-1}\tau_0)d\Phi(z) = P^* \tag{5.14}$$

where we replace $P^*$ by $Q^*$ to find the value of $\tau_0$ required to have a $CS_2$ with probability at least $Q^*$. In appendix C.3 we provide a table for the values of $\tau_0$ for different values of $P^*$ or $Q^*$ and $\gamma$. Remark that all entries are larger than their corresponding values in table C.1, where we assumed to have known and equal variances. The relations we find for $n$ and $c$ are the same as in section 3.3 and thus (3.18), where $\tau_1 := \tau_{0,1}$ and $\tau_2 := \tau_{0,2}$

$$n = \left(\frac{\sigma_0(\tau_{0,1} + \tau_{0,2})}{2\delta^*}\right)^2 \quad \text{and} \quad c = \frac{\sigma_0^2}{4\delta^*}({\tau_{0,2}}^2 - {\tau_{0,1}}^2). \tag{5.15}$$

We refer to the specification of $n$ and $c$ in this manner by a robust design. We must however keep in mind that the value of $\sigma_k$ is assumed to be known.

For the selection of $t$ ($t > 1$) best populations, the results are more difficult, because the probability of a correct selection of the $t$ best ones is not clearly monotonous for $\sigma_{[k-t+1]}, \ldots, \sigma_{[k]}$. This probability is namely (under the LFC (3.52))

$$\begin{aligned} P(CS_t|R^t_{n,c}) &= P(\min_{k-t+1\le i\le k}\sum_{j=1}^{n} Y_{(i)j} > \max_{1\le i\le k-t}\sum_{j=1}^{n} Y_{(i)j} + c) \\ &= \sum_{l=k-t+1}^{k} P(\max_{1\le i\le k-t}\sum_{j=1}^{n} Y_{(i)j} + c < \sum_{j=1}^{n} Y_{(l)j} \le \min_{\substack{k-t+1\le i\le k \\ i\neq l}}\sum_{j=1}^{n} Y_{(i)j}) \\ &= \sum_{l=k-t+1}^{k} \int_{-\infty}^{\infty} \prod_{i=1}^{k-t}\left[\Phi(\frac{y-c-n\mu_{[i]}}{\sqrt{n}\sigma_{[i]}})\right] \prod_{\substack{i=k-t+1 \\ i\neq l}}^{k}\left[1-\Phi(\frac{y-n\mu_{[i]}}{\sqrt{n}\sigma_{[i]}})\right] d\Phi(\frac{y-n\mu_{[l]}}{\sqrt{n}\sigma_{[l]}}). \end{aligned} \tag{5.16}$$

We could try to find to minimum for $P(CS_t|R^t_{n,c})$, when all standard deviations are allowed to be in $[\gamma^{-1}\sigma_0, \gamma\sigma_0]$, but as we can see, it is difficult to conclude whether $\sigma_{[i]}$ for

$i = k - t + 1, \ldots, k$ should be large or small in the minimum. To solve this, we should solve this $t$ dimensional optimization problem, which we do not consider here.

# Chapter 6

# Bayesian approach

## 6.1 Introduction

Having used frequentist statistics in the former chapters, we now take a look at our selection problem from a Bayesian viewpoint. A Bayesian approach to the selection problem seems to be promising, because it enables us to get rid of some of our assumptions, such as to know the variances. However, if we want to handle the selection problem in all its aspects in a Bayesian way, the selection problem should be regarded as a decision problem especially with respect to the determination of the sample sizes and it would take too much to analyze it in this thesis. Anyway, it is an interesting problem for future research and in section 6.4 we give some remarks on a full Bayesian approach. We restrict ourselves to the problem considered in the former chapters, the selection of the best population among $k$ independent normally distributed populationsand we do not consider determination of sample sizes, but only a Bayesian data analytic approach after the experiments have been performed. First we assume all variances to be equal and known, but we will consider later the problem of unknown variances.

## 6.2 Equal and known variances

To get the idea of what is going on we first look at a single population. In this section we use the results derived in [10, par. 2.2]. We assume the population is normally distributed with unknown mean $\mu$ and known variance $\sigma^2 > 0$. Like before, we denote the random variables by capital letters and the observations of this random variable by lowercase letters. We take a sample of $n$ (fixed) observations, $y_j$, $j = 1, \ldots, n$, assumed to be independent and equally distributed

$$Y_j \sim N(\mu, \sigma^2). \tag{6.1}$$

In the previous chapters, in particular in section 3.2, we developed the procedure in order to design the experiment in order to find the minimum sample size $n$ to take from each population and still satisfy the two conditions we imposed. This value for $n$ and for the preference threshold value $c$, to prevent us from selecting a single population that is only slightly better than another population, could be determined by specifying two probabilities for the two probability requirements and by dividing the parameter space into a preference and an indifference zone, by introducing a $\delta^*$ value for the difference between the largest and second-largest mean. In the Bayesian approach in this chapter, we cannot determine $n$ beforehand.

In Bayesian statistics the important formula is

$$p(\mu|(y_1, \ldots, y_n)) \propto l(\mu|(y_1, \ldots, y_n)) * p(\mu). \tag{6.2}$$

where $p(\mu|(y_1,\ldots,y_n))$ is the posterior probability density function of $\mu$ after collecting the data $y_1,\ldots,y_n$, $\propto$ means proportional to, $l(\mu|(y_1,\ldots,y_n))$ is the likelihood function of $\mu$ based on data $y_1,\ldots,y_n$ and $p(\mu)$ is the prior distribution of $\mu$. The likelihood function based on the model and data $y_1,\ldots,y_n$ is

$$l(\mu|(y_1,\ldots,y_n)) \propto \exp[-\frac{n}{2\sigma^2}(\mu-\bar{y})^2], \tag{6.3}$$

so $\bar{Y} \sim N(\mu, \frac{\sigma^2}{n})$ is a sufficient statistic for updating the probability distribution function for $\mu$. We take a so-called *non-informative* prior, meaning the form of the posterior is (almost) equal to the form of the likelihood function. For this location problem we take prior

$$p(\mu|\sigma^2) \propto c, \tag{6.4}$$

where $c > 0$ is a constant. Remark that this is not a proper probability distribution, because the integral over $(-\infty,\infty)$ does not exist, but it can be used as an approximation. We can assume that this prior is locally uniform and thus, approximately,

$$p(\mu|(y_1,\ldots,y_n)) \propto \exp[-\frac{n}{2\sigma^2}(\mu-\bar{y})^2]. \tag{6.5}$$

After calculating the normalizing constant to make the righthand-side integrate to one, we get

$$p(\mu|(y_1,\ldots,y_n)) = (2\pi\frac{\sigma^2}{n})^{-1/2}\exp[-\frac{n}{2\sigma^2}(\mu-\bar{y})^2], \quad -\infty < \mu < \infty. \tag{6.6}$$

Hence, the posterior distribution is $\mu \sim N(\bar{y}, \frac{\sigma^2}{n})$.

Now suppose that we have $k$ normally distributed populations and suppose

1. we take a sample of $n$ observations from each population
2. the variances of each population are all equal to a known $\sigma^2 > 0$
3. the observations within a sample are independent, conditional on having the same distribution and also the populations are independent of each other.

Like before, we denote the $j$th observation of the $i$th population by $Y_{ij}$ and we know that the sample means are distributed as

$$\bar{Y_i} = \frac{1}{n}\sum_{j=1}^{n} Y_{ij} \sim N(\mu_i, \frac{\sigma^2}{n}). \tag{6.7}$$

The posterior distribution for $\mu_i$ related to prior (6.4) is

$$\mu_i|(y_{i1}, y_{i2}, \ldots, y_{in}) \sim N(\bar{y_i}, \frac{\sigma^2}{n}). \tag{6.8}$$

Hence, given the data $y_{ij}$, we have $k$ posterior distributions for $\mu_i$, $i = 1,\ldots,k$. We are interested in the probability

$$p_i := P(\text{population } i \text{ is the best one}) = P(\mu_i = \max_{l=1,\ldots,k} \mu_l). \tag{6.9}$$

Remark that $\sum_{i=1}^{k} p_i = 1$ and this can be used as a control of the result. Using the posterior

distributions for $\mu_i$,

$$\begin{aligned} p_i &= P(\mu_i > \mu_l;\ l = 1, \ldots, k, l \neq i) \\ &= \int_{-\infty}^{\infty} \prod_{\substack{l=1 \\ l \neq i}}^{k} \Phi(\frac{m - \bar{y}_l}{\sigma/\sqrt{n}}) d\Phi(\frac{m - \bar{y}_i}{\sigma/\sqrt{n}}) \\ &= \int_{-\infty}^{\infty} \prod_{\substack{l=1 \\ l \neq i}}^{k} \Phi(z + \frac{\bar{y}_i - \bar{y}_l}{\sigma/\sqrt{n}}) d\Phi(z). \end{aligned} \tag{6.10}$$

After getting the data, these probabilities can be calculated. The values for the probabilities we find in this manner can be used to determine a subset of minimal size that counts for at least a specified amount of probability of containing the best population. The idea behind this is similar to the standard subset selection approach, where we selected a subset that should include the best one with at least a specified probability. We will show some results for some simulations in chapter 7.

## 6.3 Both unknown mean $\mu$ and unknown $\sigma^2$

In this section we consider the case of an unknown variance $\sigma^2$ for all populations. In a non-Bayesian approach, as in the previous part of this thesis, the calculations when $\sigma^2$ is assumed unknown and not necessarily equal for all populations tend to be very complicated. Some procedures have been developed to deal with this problem, we described some of them briefly in section 5.1. Two of the methods described are two-stage procedures and the other one is a sequential procedure. Here, from a Bayesian viewpoint, we first look at a single population and the observations of this population are denoted by $y_j$; $j = 1, \ldots, n$. This section is again based on Box & Tiao [10, par. 2.4]. There it is shown that the sample mean $\bar{Y}$ and the sample variance

$$s^2 = \nu^{-1} \sum_{j=1}^{n} (y_j - \bar{y})^2, \quad \nu = n - 1, \tag{6.11}$$

are jointly sufficient for $(\mu, \sigma^2)$. The likelihood function is (defined on $\mu \in \mathbb{R}, \sigma > 0$)

$$l(\mu, \sigma, (y_1, \ldots, y_n)) \propto \sigma^{-n} \exp\left\{-\frac{1}{2\sigma^2}[\nu s^2 + n(\mu - \bar{y})^2]\right\}. \tag{6.12}$$

We assume that $\mu$ and $\sigma$ are independent and therefore the prior $p(\mu, \sigma) = p((\mu)p(\sigma)$. Following Box & Tiao, we take the same *non-informative prior*

$$p(\mu) \propto c \tag{6.13}$$

and it turns out that

$$p(\sigma) \propto \sigma^{-1} \quad \text{for } \sigma > 0 \tag{6.14}$$

is a non-informative prior for $\sigma$. Now the posterior distribution for $-\infty < \mu < \infty$, $\sigma > 0$ becomes

$$\begin{aligned} &p(\mu, \sigma | (y_1, \ldots, y_n)) = \\ &\sqrt{\frac{n}{2\pi}} \left[\frac{1}{2}\Gamma(\frac{\nu}{2})\right]^{-1} \left(\frac{\nu s^2}{2}\right)^{\nu/2} \sigma^{-(n+1)} \exp\left\{-\frac{1}{2\sigma^2}[\nu s^2 + n(\mu - \bar{y})^2]\right\}. \end{aligned} \tag{6.15}$$

Our selection problem is concerned with the values of $\mu$, so we are interested in this and not in $\sigma$. Hence, we take the marginal posterior distribution of $\mu$

$$p(\mu|(y_1,\ldots,y_n)) = \int_0^\infty p(\mu,\sigma|(y_1,\ldots,y_n))d\sigma, \tag{6.16}$$

and, after defining

$$t = \frac{\mu - \bar{y}}{s/\sqrt{n}}, \tag{6.17}$$

we have for $-\infty < t < \infty$

$$p(t|(y_1,\ldots,y_n)) = \frac{1}{B(\frac{1}{2}\nu,\frac{1}{2})\sqrt{\nu}}\left(1+\frac{t^2}{\nu}\right)^{-\frac{1}{2}(\nu+1)}, \tag{6.18}$$

where $B(p,q)$ is the beta function $B(p,q) = \Gamma(p)\Gamma(q)/\Gamma(p+q)$. This function is Student's $t$-distribution with $\nu = n-1$ degrees of freedom. Note that $\mu$ is the random variable and $(\bar{y}, s^2)$ are known sample quantities. Summarizing, using the Bayesian approach, we find that the posterior distribution of $\frac{\mu-\bar{y}}{s/\sqrt{n}}$ is Student distributed with n-1 degrees of freedom.

Now, suppose we have $k$ populations, and

1. we take a sample of $n$ observations of each population
2. the means $\mu_i$ and the variances $\sigma_i^2$ are independent from each other
3. the observations within a sample are independent, conditionally on having the same distribution and also the populations are independent of each other.

Hence, we do not assume known variances. We are interested in the probability

$$\begin{aligned}
p_i :&= P(\mu_i = \max_{l=1,\ldots,k} \mu_l) \\
&= \int_{-\infty}^{\infty} \prod_{\substack{l=1 \\ l\neq i}}^{k} P(\mu_l \leq m) dP(\mu_i \leq m) \\
&= \int_{-\infty}^{\infty} \prod_{\substack{l=1 \\ l\neq i}}^{k} P(\frac{\mu_l - \bar{y}_l}{s_l/\sqrt{n}} \leq \frac{m-\bar{y}_l}{s_l/\sqrt{n}}) dP(\frac{\mu_i - \bar{y}_i}{s_i/\sqrt{n}} \leq \frac{m-\bar{y}_l}{s_l/\sqrt{n}}) \\
&= \int_{-\infty}^{\infty} \prod_{\substack{l=1 \\ l\neq i}}^{k} T_{n-1}(z * \frac{s_i}{s_l} + \frac{\bar{y}_i - \bar{y}_l}{s_l/\sqrt{n}}) dT_{n-1}(z),
\end{aligned} \tag{6.19}$$

where $T_{n-1}$ denotes the cumulative distribution function of the Student distribution with $n-1$ degrees of freedom and $s_i$ the sample variance of the $i$th population. In chapter 7 we discuss some simulations of the data analysis proposed in this section.

## 6.4 Remarks on full Bayesian approach

In a full Bayesian approach, this selection problem should be regarded as a *decision problem*. Doing this, we can handle this problem in a broader way, without having too many restrictions. The idea would be to use a utility function, in which all kind of aspects could be taken into account. This can be the costs per observation we take (this will have

a negative influence on the utility), and other aspects such as the time could be taken into account. Obviously, the utility function should depend on the value we attach to a correct selection. Perhaps there is a restriction on the maximum sample size to take, but certainly the sample size $n$ will play an important role. To design the procedure we would want to optimize the expected utility over this value $n$ and in that manner be able to determine $n$. In [24] Lindley describes the way decisions should be made based on a Bayesian approach under uncertainty. This book can be used as a first guideline to gain insight in Bayesian theory. Lindley and also Bernardo and Smith [6] strongly argue for maximizing the expected utility as the only sound criterion for decision making. The book by Bernardo and Smith describes how the Bayesian theory should be applied to approach this problem. An important role, next to the sample size $n$, will be played by the prior distribution, which should be used to model current beliefs. It is however quite likely that we can not delete a model assumption, because we have to decide about the design of the experiment before the data is taken. A full Bayesian approach seems to be very promising and can be a nice topic for future research.

# Chapter 7

# Simulation studies

In order to illustrate our **preference threshold procedure** and to be able to compare it to other approaches, we present a number of simulations of the selection problem in this chapter. For this aim, it is most useful to be in control of all input like the variance or the means. Therefore, we use Mathematica (Version 2.2, for the X Window System) to produce (pseudo-) random samples of the populations we want to consider. We should remark however, that we have some doubts about the generator producing realistic normal distributed observations. Although the sample variance estimators for a random sample still were in a 95% confidence interval, they were often close to the edge. Still, the worries are not too alarming, and we can refer to Bowman [9] who pays some attention to the performance of random generators.

## 7.1 Some cases

First we apply the different procedures on one instance to illustrate how they work. In the next section 7.2 the simulations will be repeated for at least 1000 times and with that we can draw some conclusions.

### 7.1.1 Preference threshold and standard procedures

Suppose we want to choose the best population out of $k = 10$ independent normally distributed populations with common variance $\sigma^2 = 1$ and we are indifferent about the selection if the difference between the largest and second-largest actual mean is less than $\delta^* = 0.5$. Furthermore, we want to satisfy the probability requirements of the preference threshold procedure with probability $P^* = 0.75$ for the first one (3.3) and $Q^* = 0.9$ for the second one (3.4). According to (3.18) and table C.1, we have to take a sample of size at least $(1 * \frac{2.26367+2.98293}{2*0.5})^2 = 27.5268$ and the value for $c$ becomes $c = 1.88683$. We have to round the value of the sample size upward to the nearest integer value, this means we have to take a sample of size $n = 28$ of each population. Because the value for $n$ is rounded upward, the probabilities with which we satisfy the two requirements have slightly changed. For $n = 28$, according to (3.8), the value for $\tau_1$ becomes $\tau_1 = 2.28917$ and therefore the probability for the first requirement becomes 0.7568 from (3.9), instead of the specified $P^* = 0.75$. Also for $n = 28$ we have $\tau_2 = 3.00233$ from (3.14) and the value associated with $Q^*$ becomes (3.13) 0.9028 instead of the specified $Q^* = 0.9$. We see that these changes are considerably small.

We choose the actual means to be in the least favourable configuration, for example we take $\mu = (19.5, \ldots, 19.5, 20)$ and we draw the related populations using the random generator for these means and variance $\sigma^2 = 1$. Of course, the actual means are unknown in practice. We take 100 observations of each of the 10 populations and take the first 28 of each

| Pop. | Sample sum | Rank |
|---|---|---|
| 1 | 543.04 | 1 |
| 2 | 551.69 | 9 |
| 3 | 547.71 | 7 |
| 4 | 546.61 | 6 |
| 5 | 547.85 | 8 |
| 6 | 544.24 | 3 |
| 7 | 544.51 | 5 |
| 8 | 544.50 | 4 |
| 9 | 543.48 | 2 |
| 10 | 554.35 | 10 |

Table 7.1: Preference threshold procedure

population as sample. This results in the following (rounded) sample sums in table 7.1. The largest sample sum is achieved by population 10 and the second-largest by population 2. The distance between these two is 2.66. The threshold value $c = 1.88683$, so we only select populations with sample sums larger than or equal to $554.35 - 1.88683 = 552.463$, thus only $\pi_{10}$. This is actually the best one, so this is a correct selection.

If, in this situation, we would apply **the standard indifference zone approach** as described in section 2.1, on the same data with $\delta^* = 0.5$ and $P^* = 0.75$, then according to (2.11), we would have to take samples of at least size $(\frac{1*2.26367}{0.5})^2 = 20.4968$, so rounded to the next larger integer we have $n = 21$. From the same population as before we now take instead of the first 28 observations, the first 21 observations of each population. This results in (rounded) sample sums, given in vector format:

$$(410.57, 418.87, 408.24, 414.32, 410.94, 406.60, 407.34, 408.77, 405.64, 417.99).$$

The maximum of the sample sums is achieved by $\pi_2$, hence according to Bechhofer's procedure we would select $\pi_2$, which turns out to be not the actual best one. Remark that this is just coincidence for this run, in 75 % of all runs it should select the actual best one.

If, instead of $P^* = 0.75$, we would have used $P^* = 0.9$ then the sample size would become $n = 36$ and by taking the first 36 observations of each population from the data as before, the sample sums become

$$(702.84, 709.82, 701.33, 706.20, 702.85, 699.57, 696.70, 701.78, 695.46, 711.06).$$

We select again the population with the largest sample sum, so now we select the actual best one, $\pi_{10}$. We can remark that, as stated in (3.29), the sample size for the preference threshold procedure with $P^*$ and $Q^*$ is between the values for the sample size we would need for the standard indifference zone approach with $P^*$ and $Q^*$ respectively, for we have $21 < 28 < 36$.

In section 4.1 we discussed the expected subsetsize of the preference threshold procedure. We showed that we can make a statement about the *maximum* expected subset size, although in this instance, where $\mu$ is in the LFC, the expected subset size is equal to the maximum expected subset size. This is, (4.6), for $n = 28$, $c = 1.88683, \delta^* = 0.5$ and $k = 10$

$$\max_{\Omega(\delta^*)} E(S|R_{n,c}) = \int_{-\infty}^{\infty} \Phi^{k-1}(z + \frac{n\delta^* + c}{\sqrt{n}\sigma}) d\Phi(z)$$

$$+ (k-1) \int_{-\infty}^{\infty} \Phi^{k-2}(z + \frac{c}{\sqrt{n}\sigma}) \Phi(z + \frac{c - n\delta^*}{\sqrt{n}\sigma}) d\Phi(z) = 1.21374 \quad (7.1)$$

The theoretical value $n = 27.5268$ yields a maximum expected subset size 1.22102.

If we apply the **standard subset selection approach** (section 2.2) on the samples of the first 28 (sample size we found in the designing phase of the preference threshold procedure) observations of each drawn population sizes we find the value of $d = 2.98293 * \sqrt{28} * 1 = 15.784$ (rounded) according to (2.15) and table C.1, for $P^* = 0.9$. Consequently we then select all populations with sample sum larger or equal to $554.35 - 15.784 = 538.566$, resulting in the selected subset of all populations. The maximum expected subset size with $d$ is

$$\max_{\Omega} E(S|R_G) = 10 * 0.9 = 9.$$

This result is achieved by applying (4.10). The subset size of 10 we found in this run is consequently not very surprising. We have to keep in mind that the former result is without restrictions on the parameter space, hence we allow $\mu$ to be in the entire space. With a restriction on the parameter space, so assuming we are in the preference zone, the maximum expected subset size can be calculated using (4.11) and this gives for $n = 28$

$$\max_{\Omega(\delta^*)} E(S|R_G) = 6.20737,$$

a considerable smaller size compared to the maximum expected subset size of 9, we found without restrictions on the parameter space.

### 7.1.2 Bayes, $\sigma^2$ known

As we explained in chapter 6, after getting the data we can establish the posterior distribution functions and from that make statements about the probability that a selected population is the actual best one. First, we look at the case where the variances for $k = 10$ populations are known and equal, $\sigma^2 = 1$. For the simulation the mean vector is chosen to be $\mu = (19.5, \ldots, 19.5, 20)$, so the first nine populations have the same mean value and the 10th is the best population. We use the same observations from the 10 populations as in 7.1.1, so we have the same sample sums with $n = 28$. After calculating the sample means, we can calculate the posterior probability that population $i$ is the best population by using (6.10). The results are given in table 7.2. Indeed, the sum of these probabilities

| Pop. | Sample mean | $p_i$ |
|---|---|---|
| 1 | 19.3942 | 0.0126 |
| 2 | 19.7032 | 0.2409 |
| 3 | 19.5612 | 0.0742 |
| 4 | 19.5219 | 0.0509 |
| 5 | 19.5662 | 0.0777 |
| 6 | 19.4372 | 0.0208 |
| 7 | 19.4468 | 0.0232 |
| 8 | 19.4466 | 0.0231 |
| 9 | 19.4100 | 0.0152 |
| 10 | 19.7984 | 0.4614 |

Table 7.2: Bayesian probs., $\sigma^2$ known

is 1. The actual best population $\pi_{10}$ has indeed the highest posterior probability to be the best one and if we want to have a minimum subset with at least a certain probability, let us say 0.9, to contain the best one, we select $\{\pi_{10}, \pi_2, \pi_5, \pi_3, \pi_4\}$ leading to total posterior probability 0.9051 that this subset contains the best population.

### 7.1.3 Bayes, $\sigma_i^2$ unknown

We can also not assume to know the actual variances and use the Bayesian analysis of the data as derived in section 6.3. Applying this method on the same data as in section 7.1.2 we get the following in table 7.3: Note that the probabilities sum up to 0.9999 (rounding

| Pop. | Sample mean | $p_i$ |
|---|---|---|
| 1 | 19.3942 | 0.0017 |
| 2 | 19.7032 | 0.2010 |
| 3 | 19.5612 | 0.0404 |
| 4 | 19.5219 | 0.1413 |
| 5 | 19.5662 | 0.0145 |
| 6 | 19.4372 | 0.0107 |
| 7 | 19.4468 | 0.1055 |
| 8 | 19.4466 | 0.0284 |
| 9 | 19.4100 | 0.0189 |
| 10 | 19.7984 | 0.4375 |

Table 7.3: Bayesian probs., $\sigma_i^2$ unknown

errors). The posterior probability $p_i$ is calculated from (6.19). Population $\pi_{10}$, indeed, has the largest posterior probability that it is the actual best populationand if we look for the minimum subset counting for at least 0.9 of the probability of including the best population in the selected subset, we get subset $\{\pi_{10}, \pi_2, \pi_4, \pi_7, \pi_3\}$. Together they count for total posterior probability 0.92575 that this subset includes the best population. The selected subset is only a bit different from the subset selected in section 7.1.2 and $\pi_{10}$ and $\pi_2$ have the largest probabilities in both subsets.

### 7.1.4 Sequential procedure

To the same observations from the $k$ populations as in the first simulation in section 7.1.1, we apply the sequential procedure as described in section 2.3. For the value of $\alpha$ we take $\alpha = 0.10$, giving a boundary value of $g = 8.97727$, according to (2.18). We start the procedure by taking the first observation of each population. This gives sample sums (of a single observation)

$$(19.94, 20.30, 20.33, 21.49, 19.09, 19.85, 18.68, 18.40, 18.64, 19.99)$$

and we do not eliminate any population because each sum is larger than $20.33 - 8.97727 = 11.3527$. We proceed by taking, at each stage, one more observation of each population and when $n = 13$ we have sample sums

$$(252.083, 261.369, 255.235, 256.343, 254.219, 254.116, 257.187, 254.539, 251.737, 256.738),$$

so we eliminate all populations having sample sums less than 252.392. Therefore, we eliminate populations $\pi_1$ and $\pi_9$. For $n = 15$ we eliminate $\pi_6$, and for $n = 19$ we eliminate $\pi_7$. In the 20th stage we eliminate populations $\pi_3$ and $\pi_8$, in the 30th stage we eliminate $\pi_5$, in stage 40 population $\pi_4$ is eliminated and finally in stage 45 we have sample sums

$$(882.046, 892.312),$$

for $\pi_2$ and $\pi_{10}$ respectively, because they are the only ones left to be considered. In this stage $\pi_2$ is eliminated and the process stops. The remaining population is $\pi_{10}$ and this is the one we assert to be the best. Hence, after a total number of 75 observations the sequential procedure results in a single remaining population, and it is the actual best one, indeed, in this simulation.

## 7.2 Long run simulations

In this section, the simulations consist of at least hundred runs, so we we can draw some conclusions. We will investigate the performance of the preference threshold procedure with $\mu$ in the LFC, so we can compare it with the theoretical values, we carry out some Bayesian data analysis and we investigate the effects of departures from the assumption of known standard deviations. For some simulations it is necessary to take observations from the populations one by one, for example in case of a sequential procedure, but for others we can draw the sample sum directly from the random generator which is much faster.

### 7.2.1 Means in least favourable configuration

**Preference threshold procedure $R_{n,c}$**

The first simulation we run is with $\delta^* = 0.5, P^* = 0.75$ and $Q^* = 0.9$ for $k = 10$ populations. We take $\mu$ to be in the LFC for the simulations, so for example

$$\mu = (19.5, \ldots, 19.5, 20) \tag{7.2}$$

and we take the variances all equal to one. In order to satisfy both requirements (3.3) and (3.4) on the $R_{n,c}$ we need to take samples of size $n = 28$ and the value for $c$ must be $c = 1.88683$. In this simulation of 1000 runs, in each run we take a new set of 10 populations, for each population we take observations distributed with mean $\mu_i$ and variance $\sigma^2$ and apply $R_{n,c}$. The detailed results for the simulation (for the first 100 runs) can be found in appendix E.1 and the results for several starting values of the random generator for 1000 runs are presented in appendix E.2. In E.1 we took observations one by one, but in E.2 we drew the sample sum directly, distributed $N(n\mu_i, n\sigma^2)$.

The theoretical value of the (maximum) expected subsetsize is equal to 1.2137, according to (4.6), in the LFC when the rounded value for $n$ is used and from the results in E.2 we see that the observed average subset size denoted by $\bar{s}$, is for every instance never far from this theoretical expected subset size. The average of the observed subset size over the nine runs is equal to 1.206.

In each of these simulations, the number correct selections of a single population ($CS_1$), are expected to be 756 and the actual observed values are close to this theoretical value. The average over the nine runs is 761.1. For the number of $CS_2$, correctly selecting a subset, we expect 902.8. In the nine runs, the values for $CS_2$ are indeed around this value. The average is equal to 901.78.

The number of correct selections in case only one population has been selected, as analysed in section 4.2, is for the first instance $CS_1/(CS_1+FS_1) = 768/846 = 0.9078$. Indeed, this value is larger than $P^* = 0.75$. We can calculate the theoretical value, using (4.20) and (4.21). For $P(CS_1|LFC)$ we have of course $P^* = 0.75$ and we calculate $P(FS_1|LFC) = 0.0751348$. Therefore, the theoretical value for

$$P_{LFC}(CS_1|S_1) = 0.908942. \tag{7.3}$$

The observed 0.9078 is very close to this. For the other eight runs the values are: 0.9197, 0.9090, 0.9164, 0.9187, 0.8970, 0.9110, 0.9036 and 0.9095. These results are also close to the theoretical value.

**Bayesian probabilities for known variances**

On exactly the same populations as used in the previous section 7.2.1 for the preference threshold procedure ($\mu$ as (7.2) and the variances of the $k = 10$ populations all equal to one), we calculate the Bayesian posterior probability for each population to be the best

one, after taking a sample of $n = 28$ using (6.10). The details are given in appendix E.3 and they conform with the results in the previous section 7.2.1. In every run where the Bayesian probability of $\pi_{10}$ is the maximum, it is selected by $R_{n,c}$ as being the only best one, with the exception of 10 cases where $\pi_{10}$ is selected in a subset of size larger than 1. Every time $R_{n,c}$ makes a False Selection the chosen population has a fairly higher value than $\pi_{10}$.

**Bayesian probabilities for unknown variances**

Assuming not to know the variances on exactly the same observations as before in subsection 7.2.1 for the preference threshold procedure, we must proceed as explained in section 6.3 and use (6.19) to calculate the Bayesian posterior probabilities. Therefore we have to calculate the sample means and sample variances. The results can be found in appendix E.4. For the greater part, these results conform with the Bayesian probabilities for known variances. Of course, the exact probabilities are not the same and the selected subset is sometimes a bit different than for known variances. In all runs, a subset not containing $\pi_{10}$ is never selected.

**Standard subset selection procedure**

To see how the standard subset selection procedure (section 2.2) performs and to see if the selected subset conforms with the subset found in E.3 (and E.4) we carry out a simulation on the same data as before. We have $\mu$ in the LFC (7.2), $k = 10$ populations, $\sigma^2 = 1$ (known) and we choose $P^* = 0.9$. We have $n = 28$ and from (2.16) we get $d = 15.784$. Because we assume to be in the preference zone, with $\delta^* = 0.5$, the maximum expected subset (now equal to expected subset, for $\mu$ is in the LFC) can be calculated from (4.11) and is equal to 6.20737 for this value of $n$ and $d$. In appendix E.5, where the results of this simulation is given, the actual subset size is 5.82, which is smaller.

Remark that for this instance we assume to be in the preference zone, but for the calculation of $d$ we take Gupta's LFC where *all* means are equal. Therefore, $d$ is larger than necessary and this is reflected in the results in E.5. Comparing to E.3 and E.4 we notice that the subset selected in the Bayesian approach is always contained in the subset selected by Gupta's approach.

**Sequential procedure**

On the same data for the first 100 runs, we apply the sequential procedure of section 2.3 with $\alpha = 0.1$, hence boundary $g = 8.97727$. The results are in appendix E.6. After 200 runs, the procedure selects 190 times the actual best population and it selects 8 times a non-best one. In the other 2 runs, the procedure needed more than 100 observations of some population to make a decision. We can say that the procedure selected 190 out of 200 times the right one (including two indecisive runs, when the stage was larger than 100), hence 95%. At first sight this result is higher than expected (since $\alpha = 0.1$ we expect 90% correct selections), but as we remarked in section 2.3 this 90 % is a lower bound. Furthermore, to come to this result, the sequential procedure needed a total number of 37,409 observations. Divided over 200 runs and 10 populations, this is equal to an average number of 18.705 observations per population per run. Compared to the preference threshold where we have to take 28 observations per population per run, this is more economical. However, once again, we remark that a sequential method is not applicable in every situation. Remarkable is that the largest number of observations taken is not in an indecisive run, with stage larger than 100, but in run 104 for only 66 stages. Of course, this is possible if in the first stages populations are 'close' to each other and we can not exclude them.

### 7.2.2 Simulation with $\mu$ not in LFC

**Preference threshold procedure $R_{n,c}$**

If we take $\delta^* = 0.5$ and we take

$$\mu = (19, \ldots, 19, 20) \tag{7.4}$$

then the actual $\mu$ is not in the Least Favourable Configuration and the performance of our procedure is increasing. For a run of 500 with $P^* = 0.75, Q^* = 0.9, k = 10, \delta^* = 0.5$ and all variances equal to one, our preference threshold procedure $R_{n,c}$ selects 500 times only the actual best population, hence no false selections at all. It shows that an increase in differences between the means increases the probability of the actual best population ranking best.

**Bayesian probabilities; $\sigma^2$ known**

Also for this configuration of $\mu$ (7.4) we calculate the probability for each population to be the best one, based on the sample of size $n = 28$. We observe that for the first 100 runs the probability assigned to $\pi_{10}$ is always larger than 0.90823, thus the probability of population 10 to be the actual best one is very high.

**Sequential procedure**

If we apply the sequential procedure of section 2.3 with $\alpha = 0.1$ to exactly the same populations of the first 200 runs, we see that it selects 200 times the actual best population, after taking in total 20715 observations. That is on average 10.3575 observations per run per population, compared to 28 observations per population per run needed for $R_{n,c}$, but that sample size had to be based on the least favourable configuration. This is considerably lower, however, it is not always possible to apply a sequential procedure, as we explained in section 2.3. Notice also that the number of observations per run per population with the means in this configuration is also lower than the average number of 18.7 observations needed for $\mu$ in the LFC (see section 7.2.1 for the sequential procedure). This is a strong feature of the sequential procedure, that it can 'decide' in an early stage that it is possible to stop taking observations if a population is clearly worse.

**Preference threshold procedure $R_{n,c}$**

Taking a $\mu$ configuration not as far as before from the least favourable configuration, we take for $\delta^* = 0.5$,

$$\mu = (19, 19, 19, 19.1, 19.2, 19.3, 19.3, 19.4, 19.5, 20). \tag{7.5}$$

The results of this simulations are in appendix E.7. From that we draw the conclusion that also now the preference threshold procedure performs better than it would do with $\mu$ in the least favourable configuration. The number of $CS_1$, correctly selecting only one population, is between 920 and 934, higher than the expected 756 for $\mu$ in the LFC. The same is true for the number of $CS_2$, which is between 972 and 982, where we expected 900.

### 7.2.3 Decreasing $P^*$ for fixed $Q^*$

What happens to the preference threshold procedure, if we decrease the value of $P^*$ and we keep $Q^*$ fixed? We will investigate this for the configuration

$$\mu = (19.5, \ldots, 19.5, 20).$$

We have the result for $P^* = 0.75$ and $Q^* = 0.9$ in appendix E.2, now take $P^* = 0.5$. Then we have to take samples of size $(\frac{1*(1.47483+2.98293)}{2*0.5})^2 = 19.8716$, that means rounded upward $n = 20$. The value of the preference threshold is $c = 3.36137$, so we will select all populations with sample sums within this distance from the maximum sample sum. The results are summarised in table 7.4. It is logical that the subset size has increased, because

| $\bar{s}$ | 1.812 | Seed | 1 |
|---|---|---|---|
| $CS_1$ | 496 | $FS_1$ | 57 |
| $CS_+$ | 404 | $FS_+$ | 43 |
| $CS_2$ | 900 | $FS_2$ | 100 |

Table 7.4: $P^* = 0.5$ and $Q^* = 0.9$

the requirement on selecting one population is less strong. The number of $CS_1$ (expected value 500) and $CS_2$ (expected value 900) are as expected. Now take $P^* = 0.3$, resulting in samples of size ($\tau_1 = 0.867933$) $n = 15$. The preference threshold $c$ becomes $c = c = 4.07227$. Taking the sample sums directly from the random generator we get table 7.5. Again, the

| $\bar{s}$ | 2.548 | Seed | 1 |
|---|---|---|---|
| $CS_1$ | 316 | $FS_1$ | 44 |
| $CS_+$ | 584 | $FS_+$ | 56 |
| $CS_2$ | 900 | $FS_2$ | 100 |

Table 7.5: $P^* = 0.3$ and $Q^* = 0.9$

subset size has increased, which is logical. The number of $CS_1$ is a bit higher than the expected 300, but the number of $CS_2$ fits perfectly with the expected 900. Finally, we take $P^* = 0.1$, then $n = 9$ and $c = 4.44893$, and the results are in 7.6.

The maximum of the expected subset size in this situation $n = 9$ and $c = 4.44893$ is

$$\max_{\Omega(\delta^*)} E(S|R_{n,c}) = 4.16122.$$

The observed 4.117 is close to this theoretical value, like the observed number of $CS_1$ and $CS_2$ conform to their expected values. It is very remarkable that in every instance the number of $CS_2$ is exactly equal to 900, but this is just 'coincidence'.

## 7.3 Selecting $t = 3$ best populations

As an example of selecting more than one best population, we simulate the selection of $t = 3$ best populations out of $k = 10$ populations(see section 3.7). For $\delta^* = 0.5$ the least favourable configuration can be taken as

$$\mu = (19.5, \ldots, 19.5, 20, 20, 20). \tag{7.6}$$

| $\bar{s}$ | 4.117 | Seed | 1 |
|---|---|---|---|
| $CS_1$ | 107 | $FS_1$ | 24 |
| $CS_+$ | 793 | $FS_+$ | 76 |
| $CS_2$ | 900 | $FS_2$ | 100 |

Table 7.6: $P^* = 0.1$ and $Q^* = 0.9$

We assume to know all variances equal, say $\sigma^2 = 1$, and we specify $P^* = 0.75$ and $Q^* = 0.9$. For these values, by using table C.2, we find $n = 40$ and $c = 1.98143$. We simulate the experiment by drawing for each population the sample sum based on $\mu_i$ and $\sigma$ and apply the preference threshold procedure for the selection of $t$ best populations, $R^t_{n,c}$. The results are in appendix E.8. We expect a number of about 750 for the value of $CS_t$, the selection of the $t$ best populations alone. The observed values are quite widely spread around this 750, but the average is 752 which is close. As we remarked in section 3.7, for the second requirement, including the $t$ best populations in the subset, we worked with a lower bound. Therefore, we expect the numbers of $CS_{t,s}$ to be larger than 900, and indeed the average of all observed values for $CS_{t,s}$ is 908 which is higher.

## 7.4 Varying variances

As we have seen in chapter 5, deviations from the assumption of equal and known variances can have serious effects on the actual lower bound of our probabilities of correct selection. In that chapter we allow the variances to be in an interval defined by $\gamma$ and $\sigma_0$. We run several simulations to study the effects. First we consider a rather extreme instance, taking $\gamma = 1.25$, signifying that all standard deviations can vary in the interval $[\gamma^{-1}\sigma_0, \gamma\sigma_0]$, where $\sigma_0$ is a known value. We *design* the experiment as if all standard deviations are equal to $\sigma_0 = 1$, but when we the samples are taken with the actual standard deviation equal to $\sigma = 1.25$ for all 10 populations, a situation that can appear for $\gamma = 1.25$. The configuration for $\mu$ is in the LFC, where $\delta^* = 0.5$, $P^* = 0.75$ and $Q^* = 0.9$. Therefore, assuming $\sigma = 1$, we take samples of size $n = 28$ and use preference threshold $c = 1.88683$. In appendix E.9 we find the results for this instance. For all five different starting values of the random generator (seed) the number of selecting only the best population ($CS_1$) is considerably lower than the specified 75% and also the number of selecting the best one in the subset ($CS_2$) is lower than the specified 90%. Of course, this is due to the larger variances. From the figures in appendix D, reproduced from Driessen *et al* [14], we find that the loss for $P^* = 0.75, k = 10$ and $\gamma = 1.25$ is about 0.14. Thus with the means in the LFC (as in this instance) and the worst configuration for the variances (likely in this instance), we expect a number of single correct selections of about $750 - 140 = 610$ in 1000 runs. The values observed in appendix E.9 are in this neighbourhood. For $Q^* = 0.9, k = 10$ and $\gamma = 1.25$ we find a loss of 0.12. Therefore the expected number for correct subsets would be in this instance around $900 - 120 = 780$. In appendix E.9 we observe values close to this.

To check, we ran for the same instance a simulation of the preference threshold procedure, but now with a sample size and preference threshold based on the actual standard deviations $\sigma = 1.25$. Then, for $P^* = 0.75$ and $Q^* = 0.9$, we have to take samples of size $n = 44$ and use preference threshold $c = 2.94818$. The results in appendix E.10 show that now the observed number of $CS_1$ and $CS_2$ are close to the specifications, hence $R_{n,c}$ performs as expected.

### 7.4.1 Robust design

In a robust design we assume to know the value of the standard deviation of the population corresponding with the largest mean, so $\sigma_k$ and we allow the other standard deviations to vary around this value. We now assume to know $\sigma_k = \sigma_0 = 1$ and we allow the other standard deviations to be in the interval $[\gamma^{-1}\sigma_0, \gamma\sigma_0]$ where $\gamma = 1.25$. We run a simulation for $\mu$ in the LFC, $\mu = (19.5, \ldots, 19.5, 20)$ for $k = 10$ populations, where $\delta^* = 0.5$ and we want to have $P^* = 0.75$ and $Q^* = 0.9$. Designing the experiment as if all variances are equal to one, we have to take samples of size $n = 28$ with preference threshold $c = 1.88683$. The

actual value of the standard deviations we take are

$$\sigma = (0.8, 0.9, 1, 1.1, 1.2, 1.25, 1.25, 1.25, 1.25, 1). \tag{7.7}$$

In appendix E.11 the results are presented. Of course, the procedure doesn't achieve the expected probability of a correct selection, because of the variance in the standard deviations. However, the result is not as bad as in the previous example (appendix E.9). There the number of $CS_1$ was around 630 and here it is around 700. This can be explained by the fact that there are populations in the instance of this section with standard deviations *less* than the assumed value $\sigma = 1$ and in the other instance all standard deviations were equal to 1.25.

However, if we assume to know that the standard deviations vary around the known $\sigma_k = \sigma_0 = 1$ value with $\gamma = 1.25$, we can use a robust design, and by using table C.3, we have to take samples of size $n = 38$ and use preference threshold $c = 2.39279$. In appendix E.12 the results of this simulation are presented. Indeed, the procedure performs better now, and even exceeds the expected number of correct selections. Again, this can be explained by the fact that there are populations with standard deviations *less* than the most extreme value ($\sigma_i = 1.25$), but that value is still taken into account in the robust approach. Remark as a support for this, that in the first 100 runs, where we have the details from, it seldom happens that population 1 or 2 is selected in the selected subset, the only occasion is in run 72. The ones with the largest standard deviations, populations 6,7,8,9 are more frequently selected in the subset.

This effect disappears if we take for the simulation the most extreme case for all standard deviations, meaning to take

$$\sigma = (1.25, \ldots, 1.25, 1). \tag{7.8}$$

In appendix E.13 we assumed all $\sigma$ equal to one in the designing phase and hence applied the ordinary preference threshold procedure, so we had to take $n = 28$ and $c = 1.88683$. As a result, we see that it really performs very bad and the number of $CS_1$ is around 650 instead of the expected 756. Comparing it to the instance before with $\sigma$ as in (7.7), it is indeed worse, due to the extreme configuration of $\sigma$. In appendix E.14 and E.16 we have calculated the corresponding Bayesian probabilities for $\sigma_i$ unknown as in section 6.3 for respectively appendix E.13 and E.15. For the greater part, the posterior probability for population 10 to be the actual best one is larger for the robust design in E.16 than in the ordinary design in E.14. Applying the robust design, we see from table C.3 that we have to take samples of size $n = 38$ and $c = 2.39279$. This results in the numbers we expected to get, see appendix E.15, for example the number of $CS_1$ is around 755 and we expect 756.

# Chapter 8

# Suggestions for future research

Selection procedures have been considered from the 1950's on, but still prove to be an interesting topic. To determine in our preference threshold procedure both sample size $n$ and threshold $c$, we had to define two requirements. The first one seems very natural, to select the single best population with at least a certain probability, there may however be other ideas for a second requirement. This would be interesting to discuss with people in the practical situation, however, it is disappointing to find out that, although the application of selection procedures is useful in some areas, they are not much used. Reasons could be that selection procedures are relatively young and not often mentioned in basic statistics textbooks. Secondly, one is often focussed on variance analysing methods for which numerous software packages have been developed. Consequently, we could not share the experiences from practice, but this is of course, still a useful and interesting topic to discuss.

Coolen and Van der Laan [12] chose a second requirement related to the probability of a false selection of a single population, others could be considered, such as one related to the probability of a correct selection given that one population has been selected as suggested in section 4.2. In that section we proved that this probability $P(CS_1 \mid S_1)$ attains a local minimum for the means in the Least Favourable Configuration. It would be nice if it could be proven that a global minimum is attained in the LFC (as we suspect), or if it could be proven that the area for which the local minimum in the LFC of $P(CS_1 \mid S_1)$ is a global minimum on that area, is such that for practical situations we can safely assume that the minimum of $P(CS_1 \mid S_1)$ is attained for the means in the LFC. The problem is the form of the probability of a false selection of a single population, and it seems to be a tedious analytical problem, but solved it would yield a nice result.

A promising future research topic in selection procedures is the selection problem defined in a Bayesian framework. In chapter 6 we presented a Bayesian way of data analysis, but in a full Bayesian approach we should regard the selection problem as a decision problem. In such an approach, a utility function and the modelling of prior information need to be considered, and one of the goals can be the determination of the sample size in the designing phase of the experiment, typically by maximizing the expected utility. Using a Bayesian approach enables us to get rid of some assumptions (such as assuming common known variance) and it will be interesting to see how a Bayesian method would behave compared to the frequentist approaches, such as our preference threshold procedure.

# Bibliography


[1] Bechhofer, R.E. (1954). A single-sample multiple decision procedure for ranking means of normal populations with known variances. *Ann. Math. Statist.* **25**, 1-39.

[2] Bechhofer, R.E., Dunnett, C.W. and Sobel, M. (1954). A two-sample multiple decision procedure for ranking means of normal populations with a common unknown variance. *Biometrika* **41**, 170-176.

[3] Becker, W.A. (1961). Comparing entries in random sample tests. *Poultry Science* **40**, 1507-1514.

[4] Becker, W.A. (1962). Ranking all-or-none traits in random sample tests. *Poultry Science* **41**, 1437-1438.

[5] Becker, W.A. (1964). Changes in performance of entries in random sample tests. *Poultry Science* **43**, 716-722.

[6] Bernardo, J.M. and Smith, A.F.M. (1994). *Bayesian theory.* Wiley, Chichester.

[7] Bofinger, E. (1983). Least favorable configuration. *Encyclopedia of statistical sciences* **4**, 585-588.

[8] Bofinger, E. (1985). Monotonicity of the probability of correct selection or are bigger samples better? *J. R. Statist. Soc. B* **47**, 84-89.

[9] Bowman, R.L. (1995). Evaluating pseudo-random number generators. *Computers and Graphics* **19**, 315-324.

[10] Box, G.E.P. and Tiao, G.C. (1973). *Bayesian inference in statistical analysis.* Wiley, New York.

[11] Coolen, F.P.A. (1995). On the stopping boundary for sequential selection of the normal population with the largest mean. *Manuscript University of Durham.* (submitted to *Statistica Neerlandica*)

[12] Coolen, F.P.A. and Van der Laan, P. (1995). On indifference zone selection with a preference threshold. *Memorandum COSOR 95-02.* (submitted to *Biometrical Journal*)

[13] Driessen, S.G.A.J. (1988). The probability of correct selection. *Biometrika* **75**, 618-619.

[14] Driessen, S., Van der Laan, P. and Van Putten, B. (1990). Robustness of the probability of correct selection against deviations from the assumption of a common known variance. *Biom. J.* **32**, 131-142.

[15] Dudewicz, E.J. and Dalal, S.R. (1975). Allocation of observations in ranking and selection with unequal variances. *Sankhya, B*, **37**, 28-78.

[16] Gibbons, J.D., Olkin, I. and Sobel, M. (1977). *Selecting and ordering populations. A new statistical methodology.* John Wiley & Sons, New York.

[17] Gupta, S.S. (1956). On a decision rule for a problem in ranking means. *Ph.D. Thesis. (Mimeo. Ser. No. 150).* Institute of Statistics, University of North Carolina, Chapel Hill.

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

[19] Gupta, S.S. and Panchapakesan, S. (1979). *Multiple decision procedures; Theory and methodology of selecting and ranking populations.* Wiley, New York.

[20] Gutmann, S. and Maymin, Z. (1987). Is the selected population the best? *Ann. Statist.* **15**, 456-461.

[21] Hoel, D.G. and Mazumdar, M. (1968). An extension of Paulson's selection procedure. *Annals of Mathematical Statistics* **39**, 2067-2074.

[22] Hsu, J.C. (1981). Simultaneous confidence intervals for all distances from the "best". *The annals of statistics* **9**, 1026-1034.

[23] Van der Laan, P. and Verdooren, L.R. (1988). Selection of populations: an overview and some recent results. *Technical note 88-01* Agricultural University Wageningen.

[24] Lindley, D.V. (1985). *Making decisions, second edition.* John Wiley & Sons, London.

[25] Pitt, H.R. (1985). *Measure and integration for use.* Oxford University Press, New York.

[26] Robbins, H., Sobel, M. and Starr, N. (1968). A sequential procedure for selecting the best of $k$ populations. *The annals of mathematical statistics,* **39**, 88-92.

[27] Soller, M. and Putter, J. (1964). On the probability that the best chicken stock will come out best in a single random sample test. *Poultry Science* **43**, 1425-1427.

[28] Somerville, P.N. (1985). A new subset selection method for normal populations. *J. Statist. Comput. Simul.* **22**, 27-50.

[29] Tong, Y.L. and Wetzell, D.E. (1979). On the behaviour of the probability function for selecting the best normal population. *Biometrika* **66**, 174-176.

# Appendix A

# Smallest mean

The selection of the population with the smallest mean out of $k$ populations using the preference threshold procedure is similar to the problem of selecting the largest mean, because of the symmetry of the normal distribution. This will be explained below. The **selection rule** is to select all populations corresponding to sample sums smaller than or equal to the minimum sample sum plus threshold c, that is, select $\pi_i$ if

$$\sum_{j=1}^{n} yij \leq \min_{1 \leq l \leq k} \sum_{j=1}^{n} y_{lj} + c. \tag{A.1}$$

We still rank the actual means as

$$\mu_{[1]} \leq \mu_{[2]} \leq \cdots \leq \mu_{[k]}, \tag{A.2}$$

where the goal is to select $\mu_{[1]}$, but the preference zone now becomes

$$\mu_{[1]} + \delta^* \leq \mu_{[2]} \leq \cdots \leq \mu_{[k]}. \tag{A.3}$$

For the probability of a correct selection we have

$$\begin{aligned}
P(CS_1|R_{n,c}) &= P(\sum_{j=1}^{n} Y_{(1)j} < \min_{2 \leq i \leq k} \sum_{j=1}^{n} Y_{(i)j} - c) \\
&= P(n\bar{Y}_{(i)} > n\bar{Y}_{(1)} + c;\ i = 2, \ldots, k) \\
&= \int_{-\infty}^{\infty} P(n\bar{Y}_{(i)} > y + c;\ i = 2, \ldots, k) dP(n\bar{Y}_{(1)} \leq y) \\
&= \int_{-\infty}^{\infty} \prod_{i=2}^{k} \left[1 - \Phi(\frac{y + c - n\mu_{[i]}}{\sqrt{n}\sigma})\right] d\Phi(\frac{y - n\mu_{[1]}}{\sqrt{n}\sigma}) \\
&= \int_{-\infty}^{\infty} \prod_{i=2}^{k} \left[1 - \Phi(z + \frac{n\mu_{[1]} - n\mu_{[i]} + c}{\sqrt{n}\sigma})\right] d\Phi(z).
\end{aligned} \tag{A.4}$$

The minimum of this probability over $\Omega(\delta^*)$ is attained at the Least Favourable Configuration (LFC), and in this case the least favourable configuration is

$$\mu_{[1]} + \delta^* = \mu_{[2]} = \cdots = \mu_{[k]}. \tag{A.5}$$

Then

$$P(CS_1|R_{n,c}, LFC) = \int_{-\infty}^{\infty} \left[1 - \Phi(z + \frac{-n\delta^* + c}{\sqrt{n}\sigma})\right]^{k-1} d\Phi(z)$$
$$= \int_{-\infty}^{\infty} \Phi^{k-1}(-z + \frac{n\delta^* - c}{\sqrt{n}\sigma}) d\Phi(z)$$
$$= \int_{-\infty}^{\infty} \Phi^{k-1}(w + \frac{n\delta^* - c}{\sqrt{n}\sigma}) \phi(-w) d(w)$$
$$= \int_{-\infty}^{\infty} \Phi^{k-1}(z + \frac{n\delta^* - c}{\sqrt{n}\sigma}) d\Phi(z). \quad \text{(A.6)}$$

This is the same as we found in derivation (3.7) and the this works analogously for $CS_2$, hence the problem of selecting the population with the smallest mean is similar to the problem of selecting the population with the largest mean.

# Appendix B

# Differentiation under integral

We consider the function of a correct selection of one population and we use the notation for $\epsilon_i$ as introduced in (4.22)

$$P_\mu(CS_1) = \int_{-\infty}^{\infty} \prod_{i=1}^{k-1} \left[\Phi(z + \frac{n(\epsilon_i + \cdots + \epsilon_{k-1} + \delta^*) - c}{\sqrt{n}\sigma}\right] \phi(z)dz. \tag{B.1}$$

We want to prove that differentiating $P_\mu(CS_1)$ with respect to $\epsilon_j$ is equal to taking the derivative with respect to $\epsilon_j$ first under the integral sign, hence we want to prove

$$\frac{\partial P_\mu(CS_1)}{\partial \epsilon_j} = \int_{-\infty}^{\infty} \left( \frac{\partial}{\partial \epsilon_j} \prod_{i=1}^{k-1} \left[\Phi(z + \frac{n(\epsilon_i + \cdots + \epsilon_{k-1} + \delta^*) - c}{\sqrt{n}\sigma}\right] \phi(z) \right) dz. \tag{B.2}$$

To do this, we first look at a general function and substitute $P_\mu(CS_1)$ later. We prove the following

**Corollary 2** *Define function*

$$F(\epsilon) = \int_{-\infty}^{\infty} f(z, \epsilon)dz, \qquad \textit{with } \epsilon > 0, \tag{B.3}$$

*then, assuming*

$$\int_0^{\epsilon} \int_{-\infty}^{\infty} \left(\frac{\partial}{\partial t} f(z,t)\right) dz\, dt < \infty \qquad \textit{with } f(z,t) \geq 0, \tag{B.4}$$

*we have*

$$\frac{\partial F(\epsilon)}{\partial \epsilon} = \int_{-\infty}^{\infty} \frac{\partial}{\partial \epsilon} f(z, \epsilon)dz. \tag{B.5}$$

Proof:

Assuming that the following expression is bounded, hence

$$\int_0^{\epsilon} \int_{-\infty}^{\infty} \left(\frac{\partial}{\partial t} f(z,t)\right) dz\, dt < \infty, \qquad \text{with } f(z,t) \geq 0, \tag{B.6}$$

we have, by an application of Fubini's theorem [25, p. 32] that states that the existence of

$$\int_0^{\epsilon}\int_{-\infty}^{\infty}\left|\frac{\partial}{\partial t}f(z,t)\right| dz\,dt \tag{B.7}$$

implies the existence and equality of

$$\int_{-\infty}^{\infty}\int_0^{\epsilon}\left(\frac{\partial}{\partial t}f(z,t)\right) dt\,dz, \tag{B.8}$$

that

$$\int_0^{\epsilon}\int_{-\infty}^{\infty}\left(\frac{\partial}{\partial t}f(z,t)\right) dz\,dt = \int_{-\infty}^{\infty}\int_0^{\epsilon}\left(\frac{\partial}{\partial t}f(z,t)\right) dt\,dz. \tag{B.9}$$

Therefore we can derive, with

$$G(t) = \int_{-\infty}^{\infty}\frac{\partial}{\partial t}f(z,t)dz, \tag{B.10}$$

that

$$\begin{aligned} F(0) + \int_0^{\epsilon} G(t)dt &= F(0) + \int_0^{\epsilon}\int_{-\infty}^{\infty}\left(\frac{\partial}{\partial t}f(z,t)\right) dz\,dt \\ &= F(0) + \int_{-\infty}^{\infty}\int_0^{\epsilon}\left(\frac{\partial}{\partial t}f(z,t)\right) dt\,dz \\ &= F(0) + \int_{-\infty}^{\infty}\{f(z,\epsilon) - f(z,0)\}\,dz \\ &= F(\epsilon). \end{aligned} \tag{B.11}$$

Now, according to the definition of integrability, which says that if

$$F(\epsilon) = F(0) + \int_0^{\epsilon} G(t)dt \tag{B.12}$$

and if $G$ is a continuous function over $t \in [0,\infty)$, then $F$ is differentiable and

$$F'(t) = G(t), \qquad \text{for all } t > 0. \tag{B.13}$$

Consequently (B.5) and the corollary is proved □.

To prove (B.2, the original problem, we define for ease of notation

$$\begin{aligned} c_i &= \frac{n(\epsilon_i + \cdots + \epsilon_{j-1} + t + \epsilon_{j+1} + \cdots + \epsilon_{k-1} + \delta^*) - c}{\sqrt{n}\sigma} && \text{for } i < j \\ c_j &= \frac{n(t + \epsilon_{j+1} + \cdots + \epsilon_{k-1} + \delta^*) - c}{\sqrt{n}\sigma} \\ c_i &= \frac{n(\epsilon_i + \cdots + \epsilon_{k-1} + \delta^*) - c}{\sqrt{n}\sigma} && \text{for } i > j \end{aligned} \tag{B.14}$$

and we define the function

$$h(z,t,j) = \prod_{i=1}^{k-1} [\Phi(z+c_i)]\,\phi(z) \tag{B.15}$$

for fixed values of $n, \delta^*$ and $\epsilon_i$; $i \neq j$ with $1 \leq j \leq k-1$. We have

$$\frac{\partial}{\partial t}h(z,t,j) = \sum_{m=1}^{j}\left[\phi(z+c_i)\prod_{\substack{i=1\\ i\neq m}}^{k-1}\Phi(z+c_i)\right]\phi(z), \tag{B.16}$$

and of course $\Phi(z) \leq 1$, so we can derive

$$\begin{aligned}\frac{\partial}{\partial t}h(z,t,j) &\leq \left[\sum_{m=1}^{j}\phi(z+c_i)\right]\phi(z)\\ &\leq j\frac{1}{\sqrt{2\pi}}e^{-z^2}\\ &\leq (k-1)\frac{1}{\sqrt{2\pi}}e^{-z^2}.\end{aligned} \tag{B.17}$$

With this result, we easily see that

$$\begin{aligned}\int_0^{\epsilon_j}\int_{-\infty}^{\infty}\left|\frac{\partial}{\partial t}h(z,t,j)\right|dz\,dt &\leq \int_0^{\epsilon_j}\int_{-\infty}^{\infty}(k-1)\frac{1}{\sqrt{2\pi}}e^{-z^2}dz\,dt\\ &\leq \int_0^{\epsilon_j}\frac{k-1}{\sqrt{2\pi}}dt\\ &= \frac{(k-1)\epsilon_j}{\sqrt{2\pi}} < \infty.\end{aligned} \tag{B.18}$$

Hence, by applying corollary with the substitutes of $F(\epsilon)$ by $P_\mu(CS_1)$ and $f(z,t)$ by $h(z,t,j)$, we have that

$$\frac{\partial P_\mu(CS_1)}{\partial \epsilon_j} = \int_{-\infty}^{\infty}\left(\frac{\partial}{\partial \epsilon_j}\prod_{i=1}^{k-1}\left[\Phi(z+\frac{n(\epsilon_i+\cdots+\epsilon_{k-1}+\delta^*)-c}{\sqrt{n}\sigma}\right]\phi(z)\right)dz \quad \square. \tag{B.19}$$

The proof for $P_\mu(FS_1)$ is analogous, the only difference is that for $P_\mu(FS_1)$ we have a sum of integral, but the derivation as before can be applied similarly.

# Appendix C

# Tables

## C.1 Table of $\tau_1$ and $\tau_2$ for selecting one best population

The tabled value $\tau$ is the value satisfying

$$\int_{-\infty}^{\infty} \Phi^{k-1}(z+\tau)d\Phi(z) = P^*.$$

| | $P^*$ or $Q^*$ | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| k | 0.6 | 0.7 | 0.75 | 0.8 | 0.85 | 0.9 | 0.95 | 0.975 | 0.99 |
| 2 | 0.358287 | 0.741615 | 0.953873 | 1.19023 | 1.46574 | 1.81238 | 2.32618 | 2.77181 | 3.28995 |
| 3 | 0.885173 | 1.23795 | 1.43383 | 1.65241 | 1.90783 | 2.2302 | 2.71011 | 3.12843 | 3.6173 |
| 4 | 1.15324 | 1.49325 | 1.6822 | 1.89317 | 2.13988 | 2.45157 | 2.91623 | 3.32195 | 3.79694 |
| 5 | 1.32874 | 1.66137 | 1.84628 | 2.0528 | 2.29438 | 2.59971 | 3.05517 | 3.45318 | 3.91958 |
| 6 | 1.45748 | 1.78518 | 1.96736 | 2.17087 | 2.40897 | 2.70996 | 3.15909 | 3.55175 | 4.01209 |
| 7 | 1.55831 | 1.88242 | 2.06261 | 2.26391 | 2.49945 | 2.79722 | 3.24165 | 3.63029 | 4.08605 |
| 8 | 1.64073 | 1.96207 | 2.14073 | 2.34032 | 2.57387 | 2.86914 | 3.30988 | 3.69535 | 4.14748 |
| 9 | 1.71016 | 2.02928 | 2.20671 | 2.40493 | 2.63686 | 2.93012 | 3.36786 | 3.75074 | 4.19989 |
| 10 | 1.76996 | 2.08725 | 2.26367 | 2.46075 | 2.69135 | 2.98293 | 3.41818 | 3.79889 | 4.24553 |
| 15 | 1.98332 | 2.29468 | 2.46777 | 2.66114 | 2.88738 | 3.17341 | 3.60037 | 3.97384 | 4.41205 |
| 20 | 2.1217 | 2.42967 | 2.60087 | 2.79209 | 3.0158 | 3.2986 | 3.72069 | 4.08985 | 4.52298 |
| 25 | 2.22314 | 2.52884 | 2.69875 | 2.88852 | 3.11052 | 3.39113 | 3.80989 | 4.17609 | 4.6057 |
| 30 | 2.30273 | 2.60676 | 2.77573 | 2.96444 | 3.18517 | 3.46417 | 3.88045 | 4.24444 | 4.6714 |
| 40 | 2.42309 | 2.72477 | 2.89242 | 3.07962 | 3.29858 | 3.57527 | 3.98804 | 4.34887 | 4.77203 |
| 50 | 2.51255 | 2.81263 | 2.97936 | 3.16554 | 3.38326 | 3.65836 | 4.06867 | 4.42729 | 4.84779 |

## C.2 Table of $\tau_{t1}$ and $\tau_{t2}$ for selecting $t$ best populations

The tabled value $\tau$ is the value satisfying

$$t \int_{-\infty}^{\infty} \Phi^{k-t}(z+\tau)[1-\Phi(z)]^{t-1} d\Phi(z) = P^*.$$

| | $P^*$ or $Q^*$, $t = 2$ best populations | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| k | 0.6 | 0.7 | 0.75 | 0.8 | 0.85 | 0.9 | 0.95 | 0.975 | 0.99 |
| 3 | 0.885173 | 1.23795 | 1.43383 | 1.65241 | 1.90783 | 2.2302 | 2.71011 | 3.12843 | 3.6173 |
| 4 | 1.40553 | 1.72534 | 1.90374 | 2.10354 | 2.33796 | 2.63527 | 3.08084 | 3.47202 | 3.93232 |
| 5 | 1.6706 | 1.97647 | 2.14739 | 2.33906 | 2.56429 | 2.85049 | 3.2805 | 3.65909 | 4.10576 |
| 6 | 1.84432 | 2.14206 | 2.30859 | 2.49548 | 2.71525 | 2.99479 | 3.41537 | 3.78622 | 4.22436 |
| 7 | 1.97185 | 2.26412 | 2.42769 | 2.61132 | 2.82737 | 3.10235 | 3.5164 | 3.88182 | 4.31402 |
| 8 | 2.07179 | 2.36006 | 2.52145 | 2.70269 | 2.916 | 3.18758 | 3.59677 | 3.95813 | 4.38577 |
| 9 | 2.15353 | 2.43871 | 2.59841 | 2.77778 | 2.98895 | 3.25789 | 3.66325 | 4.0214 | 4.44542 |
| 10 | 2.22241 | 2.50511 | 2.66345 | 2.84132 | 3.05076 | 3.31756 | 3.7198 | 4.07532 | 4.49637 |
| 15 | 2.45923 | 2.7342 | 2.88829 | 3.06146 | 3.26545 | 3.52546 | 3.91777 | 4.26482 | 4.67622 |
| 20 | 2.60734 | 2.87809 | 3.02984 | 3.2004 | 3.40137 | 3.65757 | 4.04427 | 4.38648 | 4.79227 |
| 25 | 2.71398 | 2.98194 | 3.13215 | 3.30099 | 3.49994 | 3.75361 | 4.13654 | 4.47548 | 4.87746 |
| 30 | 2.79677 | 3.06271 | 3.21179 | 3.37938 | 3.57686 | 3.82868 | 4.20884 | 4.54536 | 4.94455 |
| 40 | 2.92077 | 3.18391 | 3.33143 | 3.49726 | 3.69269 | 3.9419 | 4.31818 | 4.65126 | 5.04643 |
| 50 | 3.01223 | 3.27345 | 3.4199 | 3.58454 | 3.77856 | 4.02598 | 4.39956 | 4.73027 | 5.12262 |

| | $P^*$ or $Q^*$, $t = 3$ best populations | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| k | 0.6 | 0.7 | 0.75 | 0.8 | 0.85 | 0.9 | 0.95 | 0.975 | 0.99 |
| 4 | 1.15324 | 1.49325 | 1.6822 | 1.89317 | 2.13988 | 2.45157 | 2.91623 | 3.32195 | 3.79694 |
| 5 | 1.6706 | 1.97647 | 2.14739 | 2.33906 | 2.56429 | 2.85049 | 3.2805 | 3.65909 | 4.10576 |
| 6 | 1.93425 | 2.22557 | 2.38874 | 2.57203 | 2.78783 | 3.06267 | 3.47692 | 3.84286 | 4.27601 |
| 7 | 2.10708 | 2.38992 | 2.54852 | 2.72685 | 2.93703 | 3.20506 | 3.60973 | 3.9679 | 4.39257 |
| 8 | 2.23401 | 2.51111 | 2.66662 | 2.84157 | 3.04791 | 3.31125 | 3.7093 | 4.062 | 4.4807 |
| 9 | 2.33351 | 2.6064 | 2.75964 | 2.9321 | 3.13559 | 3.39545 | 3.78853 | 4.13713 | 4.55128 |
| 10 | 2.4149 | 2.68454 | 2.83601 | 3.00653 | 3.2078 | 3.46494 | 3.85411 | 4.19947 | 4.610 |
| 15 | 2.68189 | 2.94197 | 3.08821 | 3.25298 | 3.44765 | 3.69661 | 4.07397 | 4.40939 | 4.80872 |
| 20 | 2.84168 | 3.0968 | 3.24032 | 3.40207 | 3.59325 | 3.83786 | 4.20889 | 4.53888 | 4.93208 |
| 25 | 2.9544 | 3.20632 | 3.34808 | 3.50789 | 3.6968 | 3.93859 | 4.30545 | 4.63187 | 5.02096 |
| 30 | 3.04085 | 3.2905 | 3.431 | 3.58941 | 3.77669 | 4.01643 | 4.38028 | 4.7041 | 5.09017 |
| 40 | 3.16903 | 3.41555 | 3.55432 | 3.71079 | 3.89581 | 4.13272 | 4.49236 | 4.81254 | 5.19432 |
| 50 | 3.26276 | 3.50716 | 3.64476 | 3.79992 | 3.98341 | 4.21838 | 4.57514 | 4.89281 | 5.27166 |

| | $P^*$ or $Q^*$, $t = 4$ best populations | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| k | 0.6 | 0.7 | 0.75 | 0.8 | 0.85 | 0.9 | 0.95 | 0.975 | 0.99 |
| 5 | 1.32874 | 1.66137 | 1.84628 | 2.0528 | 2.29438 | 2.59971 | 3.05517 | 3.45318 | 3.91958 |
| 6 | 1.84432 | 2.14206 | 2.30859 | 2.49548 | 2.71525 | 2.99479 | 3.41537 | 3.78621 | 4.22436 |
| 7 | 2.10708 | 2.38992 | 2.54852 | 2.72685 | 2.93703 | 3.20506 | 3.60973 | 3.9679 | 4.39257 |
| 8 | 2.27938 | 2.55348 | 2.70741 | 2.88067 | 3.08511 | 3.34622 | 3.74122 | 4.09156 | 4.50778 |
| 9 | 2.40593 | 2.67413 | 2.82488 | 2.99468 | 3.1952 | 3.45155 | 3.83983 | 4.18468 | 4.59492 |
| 10 | 2.50515 | 2.76901 | 2.91742 | 3.08467 | 3.28228 | 3.53508 | 3.91834 | 4.25905 | 4.66473 |
| 15 | 2.81135 | 3.06339 | 3.20538 | 3.36558 | 3.55514 | 3.79803 | 4.16709 | 4.49597 | 4.88851 |
| 20 | 2.98499 | 3.23129 | 3.37014 | 3.52688 | 3.71246 | 3.9504 | 4.31231 | 4.63513 | 5.02082 |
| 25 | 3.10461 | 3.34734 | 3.48422 | 3.63879 | 3.82185 | 4.05667 | 4.414 | 4.73295 | 5.11418 |
| 30 | 3.19517 | 3.43538 | 3.57088 | 3.72391 | 3.90519 | 4.13778 | 4.49185 | 4.808 | 5.18599 |
| 40 | 3.32794 | 3.56473 | 3.69835 | 3.84928 | 4.02813 | 4.25767 | 4.60725 | 4.91952 | 5.29304 |
| 50 | 3.42413 | 3.65864 | 3.791 | 3.94053 | 4.11774 | 4.34522 | 4.69176 | 5.00138 | 5.37184 |

| | $P^*$ or $Q^*$, $t = 5$ best populations | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| k | 0.6 | 0.7 | 0.75 | 0.8 | 0.85 | 0.9 | 0.95 | 0.975 | 0.99 |
| 6 | 1.45748 | 1.78518 | 1.96736 | 2.17087 | 2.40897 | 2.70996 | 3.15909 | 3.55175 | 4.01209 |
| 7 | 1.97185 | 2.26412 | 2.42769 | 2.61132 | 2.82737 | 3.10235 | 3.5164 | 3.88182 | 4.31402 |
| 8 | 2.23401 | 2.51111 | 2.66662 | 2.84157 | 3.04791 | 3.31125 | 3.7093 | 4.062 | 4.4807 |
| 9 | 2.40593 | 2.67413 | 2.82488 | 2.99468 | 3.1952 | 3.45155 | 3.83983 | 4.18468 | 4.59492 |
| 10 | 2.53221 | 2.79438 | 2.9419 | 3.10819 | 3.30473 | 3.55625 | 3.93775 | 4.27709 | 4.68133 |
| 15 | 2.89101 | 3.13832 | 3.2778 | 3.4353 | 3.62184 | 3.86111 | 4.22522 | 4.55017 | 4.93857 |
| 20 | 3.08113 | 3.32183 | 3.45771 | 3.61125 | 3.79323 | 4.02688 | 4.38288 | 4.70102 | 5.08176 |
| 25 | 3.20863 | 3.44534 | 3.57903 | 3.73016 | 3.90936 | 4.13956 | 4.49054 | 4.80444 | 5.18033 |
| 30 | 3.30374 | 3.53769 | 3.66987 | 3.81932 | 3.99658 | 4.22437 | 4.57183 | 4.88273 | 5.2552 |
| 40 | 3.44149 | 3.67176 | 3.80191 | 3.94911 | 4.12376 | 4.34829 | 4.69097 | 4.99777 | 5.36549 |
| 50 | 3.54033 | 3.76817 | 3.89697 | 4.04268 | 4.2156 | 4.43795 | 4.77742 | 5.08146 | 5.44599 |

| | $P^*$ or $Q^*$, $t = 6$ best populations | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| k | 0.6 | 0.7 | 0.75 | 0.8 | 0.85 | 0.9 | 0.95 | 0.975 | 0.99 |
| 7 | 1.55831 | 1.88242 | 2.06261 | 2.26391 | 2.49945 | 2.79722 | 3.24165 | 3.63029 | 4.08605 |
| 8 | 2.07179 | 2.36006 | 2.52145 | 2.70269 | 2.916 | 3.18758 | 3.59677 | 3.95813 | 4.38577 |
| 9 | 2.33351 | 2.6064 | 2.75964 | 2.9321 | 3.13559 | 3.39545 | 3.78853 | 4.13713 | 4.55128 |
| 10 | 2.50515 | 2.76901 | 2.91742 | 3.08467 | 3.28228 | 3.53508 | 3.91834 | 4.25905 | 4.66473 |
| 15 | 2.93797 | 3.18258 | 3.32062 | 3.47657 | 3.66137 | 3.89855 | 4.25979 | 4.58245 | 4.96845 |
| 20 | 3.14799 | 3.38494 | 3.51882 | 3.6702 | 3.84976 | 4.08051 | 4.43249 | 4.74744 | 5.12476 |
| 25 | 3.28447 | 3.51698 | 3.64843 | 3.79714 | 3.97363 | 4.20056 | 4.54704 | 4.85735 | 5.22943 |
| 30 | 3.38462 | 3.61412 | 3.74393 | 3.89082 | 4.0652 | 4.28953 | 4.6322 | 4.93929 | 5.30771 |
| 40 | 3.52778 | 3.75333 | 3.88096 | 4.02544 | 4.19704 | 4.4179 | 4.75549 | 5.05823 | 5.42165 |
| 50 | 3.62944 | 3.85241 | 3.97862 | 4.12153 | 4.29129 | 4.50986 | 4.84408 | 5.14393 | 5.50404 |

| | $P^*$ or $Q^*$, $t = 7$ best populations | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| k | 0.6 | 0.7 | 0.75 | 0.8 | 0.85 | 0.9 | 0.95 | 0.975 | 0.99 |
| 8 | 1.64073 | 1.96207 | 2.14073 | 2.34032 | 2.57387 | 2.86914 | 3.30988 | 3.69535 | 4.14748 |
| 9 | 2.15353 | 2.43871 | 2.59841 | 2.77778 | 2.98895 | 3.25789 | 3.66325 | 4.0214 | 4.44542 |
| 10 | 2.4149 | 2.68454 | 2.83601 | 3.00653 | 3.2078 | 3.46494 | 3.85411 | 4.19947 | 4.61 |
| 15 | 2.95992 | 3.20327 | 3.34066 | 3.49589 | 3.67988 | 3.91611 | 4.27601 | 4.59762 | 4.98249 |
| 20 | 3.19434 | 3.42875 | 3.56129 | 3.71121 | 3.88912 | 4.1179 | 4.46715 | 4.77992 | 5.15492 |
| 25 | 3.34113 | 3.5706 | 3.70043 | 3.84738 | 4.02189 | 4.24645 | 4.58963 | 4.89729 | 5.26657 |
| 30 | 3.44689 | 3.67308 | 3.80112 | 3.9461 | 4.11833 | 4.34006 | 4.67914 | 4.98333 | 5.34866 |
| 40 | 3.59592 | 3.81788 | 3.9436 | 4.086 | 4.25526 | 4.4733 | 4.80698 | 5.10658 | 5.46668 |
| 50 | 3.70058 | 3.91982 | 4.04403 | 4.18478 | 4.3521 | 4.56774 | 4.89789 | 5.19448 | 5.55111 |

## C.3 Table of $\tau_1$ and $\tau_2$ for selecting one best population with different values of $\gamma$

The tabled value $\tau$ is the value satisfying

$$\int_{-\infty}^{\infty} \Phi^{k-1}(\gamma^{-1}z + \gamma^{-1}\tau_0)d\Phi(z) = P^*. \tag{C.1}$$

| | $P^*$ or $Q^*$, $\gamma = 1.1$ | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| k | 0.6 | 0.7 | 0.75 | 0.8 | 0.85 | 0.9 | 0.95 | 0.975 | 0.99 |
| 2 | 0.376628 | 0.779578 | 1.0027 | 1.25116 | 1.54077 | 1.90516 | 2.44525 | 2.9137 | 3.45835 |
| 3 | 0.953552 | 1.32126 | 1.52553 | 1.75356 | 2.02015 | 2.35681 | 2.85838 | 3.29599 | 3.80782 |
| 4 | 1.2473 | 1.60017 | 1.79639 | 2.01562 | 2.27215 | 2.5965 | 3.08059 | 3.50382 | 3.99998 |
| 5 | 1.43971 | 1.78397 | 1.97549 | 2.18953 | 2.4401 | 2.75708 | 3.23057 | 3.64498 | 4.13131 |
| 6 | 1.58091 | 1.91939 | 2.10773 | 2.31826 | 2.56477 | 2.87671 | 3.34287 | 3.75109 | 4.23048 |
| 7 | 1.69153 | 2.0258 | 2.21181 | 2.41976 | 2.66327 | 2.97146 | 3.43215 | 3.83572 | 4.30984 |
| 8 | 1.78197 | 2.11299 | 2.2972 | 2.50314 | 2.74432 | 3.04959 | 3.50598 | 3.90587 | 4.37579 |
| 9 | 1.85817 | 2.18658 | 2.36934 | 2.57367 | 2.81297 | 3.11587 | 3.56876 | 3.96564 | 4.43217 |
| 10 | 1.92382 | 2.25007 | 2.43164 | 2.63463 | 2.87236 | 3.1733 | 3.62326 | 4.01761 | 4.48125 |
| 15 | 2.15808 | 2.47733 | 2.65499 | 2.8536 | 3.08619 | 3.3806 | 3.82083 | 4.20671 | 4.66047 |
| 20 | 2.31008 | 2.62532 | 2.80072 | 2.99679 | 3.22639 | 3.51699 | 3.95147 | 4.33228 | 4.78008 |
| 25 | 2.42151 | 2.73406 | 2.90794 | 3.1023 | 3.32987 | 3.61787 | 4.0484 | 4.42571 | 4.86935 |
| 30 | 2.50897 | 2.81954 | 2.9923 | 3.18539 | 3.41146 | 3.69753 | 4.12513 | 4.49981 | 4.94033 |
| 40 | 2.64122 | 2.94901 | 3.12019 | 3.31151 | 3.53546 | 3.81879 | 4.2422 | 4.61312 | 5.04913 |
| 50 | 2.73954 | 3.04542 | 3.21552 | 3.4056 | 3.62808 | 3.90952 | 4.33001 | 4.6983 | 5.13111 |

| | $P^*$ or $Q^*$, $\gamma = 1.25$ | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| k | 0.6 | 0.7 | 0.75 | 0.8 | 0.85 | 0.9 | 0.95 | 0.975 | 0.99 |
| 2 | 0.405553 | 0.839446 | 1.07971 | 1.34725 | 1.6591 | 2.05148 | 2.63305 | 3.13747 | 3.72397 |
| 3 | 1.05721 | 1.44871 | 1.66636 | 1.90947 | 2.19387 | 2.55333 | 3.08948 | 3.55785 | 4.10635 |
| 4 | 1.38928 | 1.76276 | 1.97067 | 2.20314 | 2.47543 | 2.82011 | 3.33544 | 3.78684 | 4.31696 |
| 5 | 1.60693 | 1.9699 | 2.17207 | 2.39823 | 2.6633 | 2.99909 | 3.50169 | 3.94256 | 4.46109 |
| 6 | 1.76673 | 2.12261 | 2.32089 | 2.54277 | 2.80288 | 3.13255 | 3.62629 | 4.05976 | 4.5701 |
| 7 | 1.89197 | 2.24265 | 2.43808 | 2.65679 | 2.91325 | 3.23834 | 3.72546 | 4.15334 | 4.6574 |
| 8 | 1.99439 | 2.34107 | 2.53428 | 2.75052 | 3.00411 | 3.32564 | 3.80753 | 4.23097 | 4.73001 |
| 9 | 2.0807 | 2.42415 | 2.61558 | 2.82984 | 3.08112 | 3.39974 | 3.87737 | 4.29716 | 4.79207 |
| 10 | 2.15507 | 2.49586 | 2.68581 | 2.89842 | 3.14777 | 3.46398 | 3.93803 | 4.35475 | 4.84616 |
| 15 | 2.42058 | 2.75268 | 2.93777 | 3.14495 | 3.38794 | 3.6961 | 4.15817 | 4.56452 | 5.04397 |
| 20 | 2.59293 | 2.92003 | 3.10232 | 3.30635 | 3.54562 | 3.84904 | 4.30398 | 4.70407 | 5.17621 |
| 25 | 2.71933 | 3.04306 | 3.22346 | 3.42534 | 3.66208 | 3.96225 | 4.41229 | 4.80804 | 5.27505 |
| 30 | 2.81854 | 3.1398 | 3.31879 | 3.51909 | 3.75396 | 4.05172 | 4.49809 | 4.89057 | 5.35371 |
| 40 | 2.9686 | 3.28638 | 3.4634 | 3.66146 | 3.89367 | 4.18801 | 4.62912 | 5.01691 | 5.47448 |
| 50 | 3.08019 | 3.39557 | 3.57123 | 3.76774 | 3.9981 | 4.29005 | 4.72749 | 5.11197 | 5.56551 |

| | $P^*$ or $Q^*$, $\gamma = \sqrt{2}$ | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| k | 0.6 | 0.7 | 0.75 | 0.8 | 0.85 | 0.9 | 0.95 | 0.975 | 0.99 |
| 2 | 0.43881 | 0.908288 | 1.16825 | 1.45773 | 1.79515 | 2.21971 | 2.84897 | 3.39476 | 4.02934 |
| 3 | 1.17199 | 1.59117 | 1.82439 | 2.08504 | 2.39016 | 2.77614 | 3.35253 | 3.85669 | 4.44778 |
| 4 | 1.54581 | 1.94342 | 2.16501 | 2.41299 | 2.70374 | 3.07225 | 3.62418 | 4.10857 | 4.67845 |
| 5 | 1.79095 | 2.1759 | 2.3906 | 2.63103 | 2.91316 | 3.2711 | 3.80799 | 4.28005 | 4.83653 |
| 6 | 1.971 | 2.34739 | 2.5574 | 2.79267 | 3.06886 | 3.41949 | 3.94588 | 4.40922 | 4.95611 |
| 7 | 2.11216 | 2.48225 | 2.68881 | 2.92027 | 3.19204 | 3.5372 | 4.05568 | 4.5124 | 5.05194 |
| 8 | 2.22763 | 2.59285 | 2.79672 | 3.0252 | 3.29353 | 3.63438 | 4.14661 | 4.59806 | 5.13171 |
| 9 | 2.32496 | 2.68626 | 2.88796 | 3.11402 | 3.37955 | 3.71691 | 4.22402 | 4.67114 | 5.1999 |
| 10 | 2.40883 | 2.76689 | 2.96679 | 3.19085 | 3.45405 | 3.78848 | 4.29129 | 4.73475 | 5.25938 |
| 15 | 2.7084 | 3.05583 | 3.24981 | 3.46725 | 3.7227 | 4.04735 | 4.53569 | 4.9667 | 5.47708 |
| 20 | 2.90296 | 3.24425 | 3.43479 | 3.64835 | 3.89925 | 4.21812 | 4.6978 | 5.12123 | 5.6229 |
| 25 | 3.04568 | 3.38282 | 3.57103 | 3.78196 | 4.02975 | 4.34464 | 4.81832 | 5.23647 | 5.73197 |
| 30 | 3.15773 | 3.49181 | 3.67829 | 3.88728 | 4.13275 | 4.44469 | 4.91386 | 5.32803 | 5.81884 |
| 40 | 3.32724 | 3.65703 | 3.84107 | 4.04729 | 4.28947 | 4.59717 | 5.05989 | 5.4683 | 5.95224 |
| 50 | 3.45331 | 3.78013 | 3.96248 | 4.16679 | 4.40669 | 4.71143 | 5.16961 | 5.57394 | 6.05302 |

# Appendix D

# Figures of robustness

The figures in this appendix are reproduced from Driessen *et al.* as they were presented in [14].

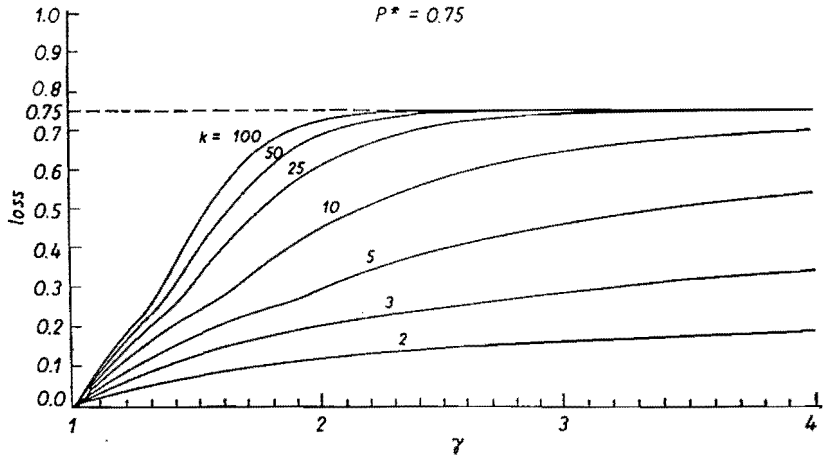


Figure D.1: Plots of loss $(k, 0.75, \gamma)$ as a function of $\gamma$ for the cases $k = 2, 3, 5, 10, 25, 50$ and 100

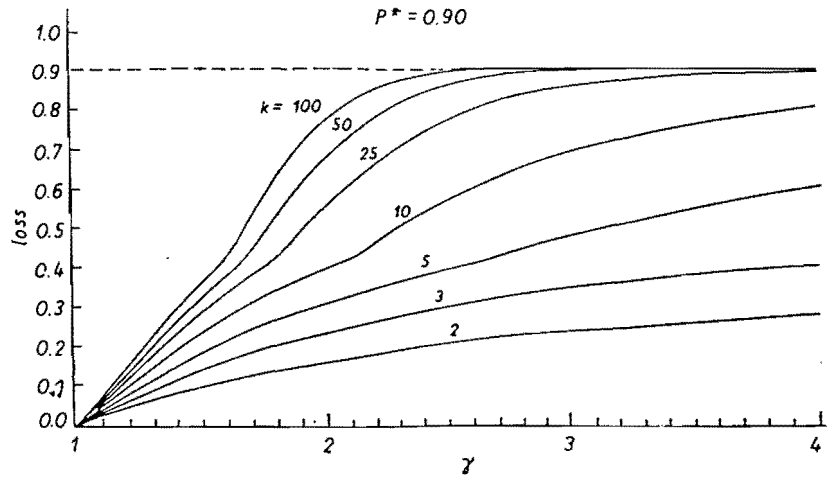


Figure D.2: Plots of loss $(k, 0.90, \gamma)$ as a function of $\gamma$ for the cases $k = 2, 3, 5, 10, 25, 50$ and $100$

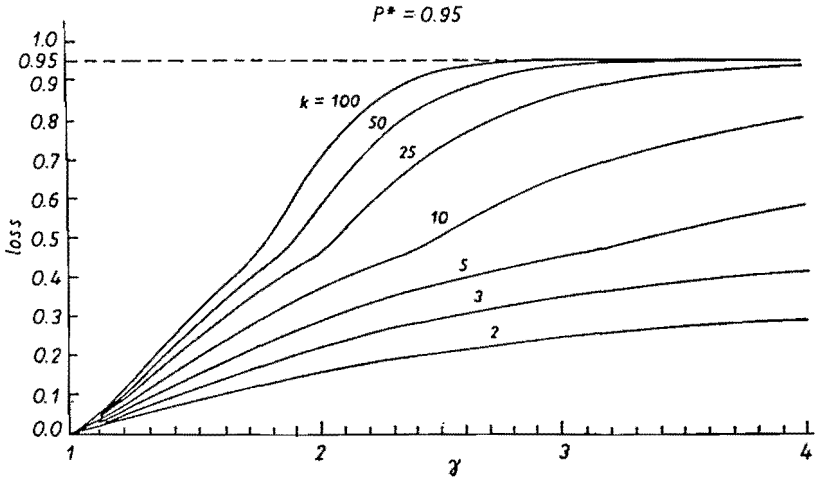


Figure D.3: Plots of loss $(k, 0.95, \gamma)$ as a function of $\gamma$ for the cases $k = 2, 3, 5, 10, 25, 50$ and $100$

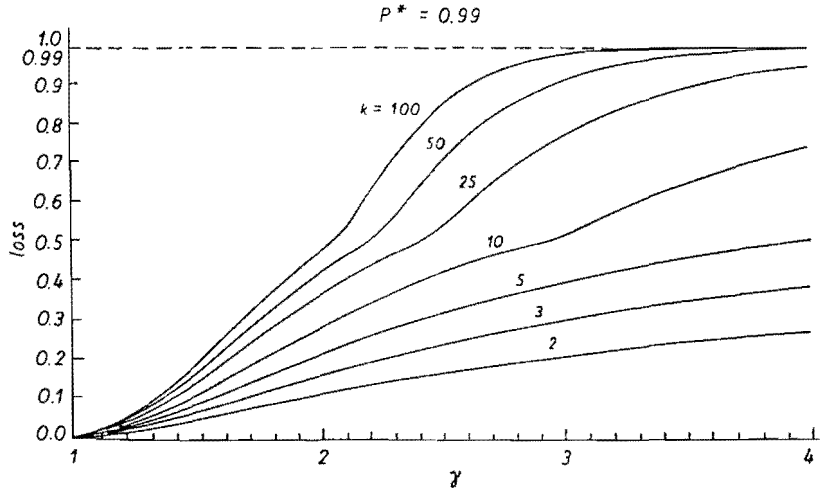


Figure D.4: Plots of loss $(k, 0.99, \gamma)$ as a function of $\gamma$ for the cases $k = 2, 3, 5, 10, 25, 50$ and $100$

# Appendix E

# Simulations

In this appendix E, we apply the preference threshold procedure and other procedures on artificial populations produced by the Mathematica random generator. We used the Mathematica, Version 2.2, for the X Window System. In the presentation of the results of the simulations we use some notation, we need to introduce. We define:

$CS_1$ : correct selection of one single best population.
$CS_+$ : correct selection of subset of size $> 1$ (including best population).
$CS_2$ : total of correct selections.
$FS_1$ : false selection of one single population.
$FS_+$ : false selection of subset of size $> 1$ (not including best population).
$FS_2$ : total of false selections.

Furthermore, in each run where a selection is done, we indicate the 'status' of the selection by a number. We use '1' to indicate a $CS_1$, '2' for $CS_+$, '3' for $FS_1$ and '4' to indicate a $FS_+$.

## E.1 Simulation with $\mu$ in LFC; $R_{n,c}$ (100 detailed runs)

We have taken $\mu$ to be in the LFC, for example, $\mu = (19.5, \ldots, 19.5, 20)$, where $\delta^* = 0.5$, $k = 10$ and the variances all equal to $\sigma^2 = 1$. We apply the **preference threshold procedure**, $R_{n,c}$ on this instance and choose $P^* = 0.75$, $Q^* = 0.9$. For these values, the sample size becomes $n = 28$ and the threshold is $c = 1.88683$. Because $n$ is rounded, the practical values for $P^*$ and $Q^*$ become $P^* = 0.7568$ and $Q^* = 0.9028$. We take from each population 100 observations and then take a sample of size 28. The results are:

| run | SubsetSize | Subset | Status | run | SubsetSize | Subset | Status |
|---|---|---|---|---|---|---|---|
| 1 | 1 | 10 | 1 | 51 | 1 | 10 | 1 |
| 2 | 1 | 10 | 1 | 52 | 1 | 10 | 1 |
| 3 | 1 | 10 | 1 | 53 | 1 | 10 | 1 |
| 4 | 1 | 10 | 1 | 54 | 2 | 7, 10 | 2 |
| 5 | 1 | 10 | 1 | 55 | 1 | 10 | 1 |
| 6 | 2 | 1, 10 | 2 | 56 | 1 | 10 | 1 |
| 7 | 1 | 3 | 3 | 57 | 1 | 10 | 1 |
| 8 | 1 | 10 | 1 | 58 | 1 | 10 | 1 |
| 9 | 2 | 4, 7 | 4 | 59 | 1 | 10 | 1 |
| 10 | 1 | 10 | 1 | 60 | 1 | 10 | 1 |
| 11 | 1 | 10 | 1 | 61 | 1 | 10 | 1 |
| 12 | 1 | 10 | 1 | 62 | 1 | 2 | 3 |
| 13 | 2 | 2, 10 | 2 | 63 | 1 | 10 | 1 |
| 14 | 3 | 7, 8, 10 | 2 | 64 | 1 | 10 | 1 |
| 15 | 1 | 10 | 1 | 65 | 1 | 10 | 1 |
| 16 | 2 | 9, 10 | 2 | 66 | 1 | 10 | 1 |
| 17 | 1 | 10 | 1 | 67 | 2 | 1, 10 | 2 |
| 18 | 2 | 7, 10 | 2 | 68 | 1 | 10 | 1 |
| 19 | 1 | 10 | 1 | 69 | 2 | 4, 10 | 2 |
| 20 | 1 | 10 | 1 | 70 | 1 | 10 | 1 |
| 21 | 1 | 10 | 1 | 71 | 1 | 10 | 1 |
| 22 | 1 | 10 | 1 | 72 | 1 | 10 | 1 |
| 23 | 1 | 10 | 1 | 73 | 1 | 10 | 1 |
| 24 | 1 | 10 | 1 | 74 | 1 | 10 | 1 |
| 25 | 1 | 10 | 1 | 75 | 1 | 4 | 3 |
| 26 | 1 | 10 | 1 | 76 | 4 | 5, 6, 7, 10 | 2 |
| 27 | 1 | 10 | 1 | 77 | 1 | 10 | 1 |
| 28 | 1 | 10 | 1 | 78 | 1 | 10 | 1 |
| 29 | 1 | 10 | 1 | 79 | 1 | 10 | 1 |
| 30 | 1 | 10 | 1 | 80 | 1 | 10 | 1 |
| 31 | 2 | 1, 10 | 2 | 81 | 1 | 10 | 1 |
| 32 | 1 | 3 | 3 | 82 | 1 | 10 | 1 |
| 33 | 1 | 10 | 1 | 83 | 1 | 10 | 1 |
| 34 | 3 | 4, 6, 10 | 2 | 84 | 1 | 10 | 1 |
| 35 | 1 | 10 | 1 | 85 | 1 | 10 | 1 |
| 36 | 1 | 10 | 1 | 86 | 2 | 2, 10 | 2 |
| 37 | 1 | 10 | 1 | 87 | 2 | 1, 10 | 2 |
| 38 | 1 | 10 | 1 | 88 | 1 | 10 | 1 |
| 39 | 1 | 10 | 1 | 89 | 2 | 8, 10 | 2 |
| 40 | 1 | 10 | 1 | 90 | 1 | 10 | 1 |
| 41 | 1 | 10 | 1 | 91 | 2 | 1, 10 | 2 |
| 42 | 1 | 10 | 1 | 92 | 1 | 10 | 1 |
| 43 | 1 | 10 | 1 | 93 | 1 | 10 | 1 |
| 44 | 2 | 6, 10 | 2 | 94 | 1 | 10 | 1 |
| 45 | 1 | 10 | 1 | 95 | 1 | 10 | 1 |
| 46 | 1 | 10 | 1 | 96 | 2 | 9, 10 | 2 |
| 47 | 1 | 10 | 1 | 97 | 2 | 4, 10 | 2 |
| 48 | 1 | 10 | 1 | 98 | 1 | 10 | 1 |
| 49 | 1 | 10 | 1 | 99 | 1 | 10 | 1 |
| 50 | 1 | 10 | 1 | 100 | 1 | 10 | 1 |

Subset size sum = 123, $CS_1$= 77, $FS_1$=4, $CS_+$= 18, $FS_+$= 1.

## E.2 Simulation with $\mu$ in LFC; $R_{n,c}$

We take $\mu$ to be in the LFC, for example, $\mu = (19.5, \ldots, 19.5, 20)$, $k = 10$ and the variances all equal to one. We apply the **preference threshold procedure**, $R_{n,c}$ on this instance and choose $P^* = 0.75$, $Q^* = 0.9$ and $\delta^* = 0.5$. For these values, the sample size becomes $n = 28$ and the threshold is $c = 1.88683$. Because $n$ is rounded, the practical values for $P^*$ and $Q^*$ become $P^* = 0.7568$ and $Q^* = 0.9028$. Here we take the sample sums for samples of size $n = 28$ directly from the random generator, distributed $N(28 * \mu_i, \sqrt{28})$. For different starting values of the random generator the results after each time 1000 runs are:

| $\bar{s}$ | 1.201 | Seed | 1 |
|---|---|---|---|
| $CS_1$ | 768 | $FS_1$ | 78 |
| $CS_+$ | 134 | $FS_+$ | 20 |
| $CS_2$ | 902 | $FS_2$ | 98 |

| $\bar{s}$ | 1.206 | Seed | 2 |
|---|---|---|---|
| $CS_1$ | 767 | $FS_1$ | 67 |
| $CS_+$ | 145 | $FS_+$ | 21 |
| $CS_2$ | 912 | $FS_2$ | 88 |

| $\bar{s}$ | 1.206 | Seed | 3 |
|---|---|---|---|
| $CS_1$ | 759 | $FS_1$ | 76 |
| $CS_+$ | 142 | $FS_+$ | 23 |
| $CS_2$ | 901 | $FS_2$ | 99 |

| $\bar{s}$ | 1.228 | Seed | 4 |
|---|---|---|---|
| $CS_1$ | 756 | $FS_1$ | 69 |
| $CS_+$ | 157 | $FS_+$ | 18 |
| $CS_2$ | 913 | $FS_2$ | 87 |

| $\bar{s}$ | 1.197 | Seed | 5 |
|---|---|---|---|
| $CS_1$ | 768 | $FS_1$ | 68 |
| $CS_+$ | 136 | $FS_+$ | 28 |
| $CS_2$ | 904 | $FS_2$ | 96 |

| $\bar{s}$ | 1.230 | Seed | 6 |
|---|---|---|---|
| $CS_1$ | 740 | $FS_1$ | 85 |
| $CS_+$ | 154 | $FS_+$ | 21 |
| $CS_2$ | 894 | $FS_2$ | 106 |

| $\bar{s}$ | 1.192 | Seed | 7 |
|---|---|---|---|
| $CS_1$ | 768 | $FS_1$ | 75 |
| $CS_+$ | 123 | $FS_+$ | 34 |
| $CS_2$ | 891 | $FS_2$ | 109 |

| $\bar{s}$ | 1.210 | Seed | 8 |
|---|---|---|---|
| $CS_1$ | 750 | $FS_1$ | 80 |
| $CS_+$ | 146 | $FS_+$ | 24 |
| $CS_2$ | 896 | $FS_2$ | 104 |

| $\bar{s}$ | 1.186 | Seed | 9 |
|---|---|---|---|
| $CS_1$ | 774 | $FS_1$ | 77 |
| $CS_+$ | 129 | $FS_+$ | 20 |
| $CS_2$ | 903 | $FS_2$ | 97 |

## E.3 Simulation with $\mu$ in LFC; Bayes, $\sigma^2$ known

For $k = 10$ populations with $\mu = (19.5, \ldots, 19.5, 20)$, all variances equal to one, calculated is in each of the 100 runs the minimum set of populations that counts for at least 0.9 of the probability of containing the best one. The populations we used are exactly the same as in E.1, meaning populations of in total 100 observations, with samples of size 28.

| Run | Subset | Probs |
|---|---|---|
| 1 | 10, 2, 5, 3, 4 | 0.4614, 0.241, 0.07771, 0.07419, 0.05091 |
| 2 | 10, 6, 4 | 0.7134, 0.1223, 0.09071 |
| 3 | 10, 8 | 0.872, 0.05525 |
| 4 | 10, 6, 5, 8 | 0.6167, 0.1473, 0.1328, 0.04269 |
| 5 | 10, 2, 5, 8 | 0.5247, 0.2116, 0.1081, 0.08172 |
| 6 | 10, 1, 3 | 0.4352, 0.357, 0.1155 |
| 7 | 3, 10 | 0.6728, 0.2516 |
| 8 | 10, 3, 2, 4 | 0.7854, 0.07284, 0.03873, 0.03166 |
| 9 | 7, 4, 10, 8 | 0.3901, 0.2868, 0.1723, 0.1177 |
| 10 | 10, 3 | 0.6769, 0.247 |
| 11 | 10, 8, 5 | 0.7198, 0.1257, 0.09206 |
| 12 | 10, 6, 8, 2, 1 | 0.3436, 0.186, 0.1592, 0.1445, 0.134 |
| 13 | 10, 2, 5, 6 | 0.3639, 0.3261, 0.1821, 0.04842 |
| 14 | 10, 7, 8, 5, 2, 1, 3 | 0.2706, 0.2414, 0.1655, 0.07555, 0.06433, 0.06298, 0.05237 |
| 15 | 10, 3, 5 | 0.8139, 0.08058, 0.06665 |
| 16 | 9, 10, 5, 1, 6 | 0.3335, 0.269, 0.1788, 0.08345, 0.05813 |
| 17 | 10, 3 | 0.8816, 0.05967 |
| 18 | 7, 10, 8 | 0.4303, 0.3752, 0.1068 |
| 19 | 10 | 0.9269 |
| 20 | 10, 9 | 0.8135, 0.1028 |
| 21 | 10, 9, 5, 3, 6 | 0.528, 0.1624, 0.0966, 0.06766, 0.06722 |
| 22 | 10 | 0.9977 |
| 23 | 10, 4, 7 | 0.62, 0.2482, 0.04157 |
| 24 | 10, 9 | 0.5966, 0.3134 |
| 25 | 10 | 0.9189 |
| 26 | 10, 5, 6, 8, 9 | 0.4209, 0.2358, 0.09713, 0.09277, 0.08495 |
| 27 | 10 | 0.9138 |
| 28 | 10 | 0.9943 |
| 29 | 10, 5 | 0.8949, 0.03602 |
| 30 | 10, 2, 7, 6, 8 | 0.6671, 0.09353, 0.06272, 0.04978, 0.04784 |
| 31 | 1, 10 | 0.5241, 0.4013 |
| 32 | 3, 1, 10 | 0.5169, 0.2607, 0.1352 |
| 33 | 10, 2, 5, 6 | 0.6384, 0.175, 0.05767, 0.03576 |
| 34 | 6, 10, 4, 8 | 0.3159, 0.268, 0.2637, 0.07704 |
| 35 | 10, 3, 1, 5 | 0.6252, 0.135, 0.1188, 0.06001 |
| 36 | 10, 6, 9 | 0.5051, 0.286, 0.116 |
| 37 | 10 | 0.9029 |
| 38 | 10, 2 | 0.8385, 0.06728 |
| 39 | 10 | 0.9893 |
| 40 | 10, 3, 9, 4, 1 | 0.625, 0.1271, 0.08347, 0.05382, 0.03867 |
| 41 | 10 | 0.9784 |
| 42 | 10, 3, 5 | 0.6856, 0.194, 0.0914 |
| 43 | 10, 9, 8 | 0.7374, 0.1302, 0.05585 |
| 44 | 10, 6, 5, 8 | 0.3743, 0.3342, 0.1573, 0.1021 |
| 45 | 10 | 0.9218 |
| 46 | 10, 3, 8 | 0.8367, 0.05865, 0.03564 |
| 47 | 10, 9, 3 | 0.5264, 0.2864, 0.108 |
| 48 | 10 | 0.9835 |
| 49 | 10 | 0.9191 |
| 50 | 10 | 0.9423 |

| Run | Subset | Probs |
|---|---|---|
| 51 | 10, 5 | 0.8929, 0.06985 |
| 52 | 10, 2 | 0.832, 0.08533 |
| 53 | 10 | 0.9991 |
| 54 | 10, 7, 8 | 0.4629, 0.378, 0.1184 |
| 55 | 10, 9, 4 | 0.6699, 0.1755, 0.1234 |
| 56 | 10, 3 | 0.8893, 0.0577 |
| 57 | 10 | 0.9583 |
| 58 | 10 | 0.9 |
| 59 | 10, 3, 5 | 0.6733, 0.1216, 0.1213 |
| 60 | 10, 8, 1, 5 | 0.8254, 0.03855, 0.03487, 0.02424 |
| 61 | 10 | 0.9825 |
| 62 | 2, 10, 6, 8 | 0.4554, 0.2659, 0.1407, 0.04736 |
| 63 | 10, 5 | 0.8768, 0.06164 |
| 64 | 10, 9, 6, 2 | 0.7284, 0.07167, 0.06289, 0.04326 |
| 65 | 10 | 0.9729 |
| 66 | 10, 5 | 0.6172, 0.3036 |
| 67 | 10, 1, 7 | 0.4176, 0.414, 0.1116 |
| 68 | 10, 8, 2, 3 | 0.6417, 0.1432, 0.09574, 0.04095 |
| 69 | 4, 10, 8 | 0.4549, 0.4412, 0.04776 |
| 70 | 10, 8, 3 | 0.7114, 0.1688, 0.08371 |
| 71 | 10, 8, 1, 4, 6 | 0.4345, 0.1829, 0.1456, 0.0734, 0.07008 |
| 72 | 10, 4, 5, 3, 8 | 0.4617, 0.1836, 0.1287, 0.1037, 0.06346 |
| 73 | 10 | 0.9054 |
| 74 | 10 | 0.9421 |
| 75 | 4, 10, 7 | 0.6616, 0.2052, 0.109 |
| 76 | 10, 5, 6, 7 | 0.2963, 0.2317, 0.2204, 0.183 |
| 77 | 10, 1, 7 | 0.5353, 0.3014, 0.1222 |
| 78 | 10 | 0.9726 |
| 79 | 10, 1 | 0.8408, 0.09501 |
| 80 | 10 | 0.9127 |
| 81 | 10 | 0.9915 |
| 82 | 10, 8 | 0.8886, 0.09949 |
| 83 | 10, 6, 2, 9 | 0.5221, 0.2491, 0.07829, 0.0723 |
| 84 | 10, 5 | 0.8957, 0.03553 |
| 85 | 10, 1, 4 | 0.5386, 0.3331, 0.05379 |
| 86 | 10, 2, 5, 3, 1 | 0.3066, 0.2237, 0.1737, 0.1336, 0.06474 |
| 87 | 1, 10, 9, 4 | 0.4052, 0.3185, 0.1405, 0.06126 |
| 88 | 10 | 0.9635 |
| 89 | 10, 8, 6, 2 | 0.3407, 0.2665, 0.1842, 0.1209 |
| 90 | 10, 4, 7 | 0.77, 0.07306, 0.05756 |
| 91 | 10, 1, 9, 7 | 0.4413, 0.3531, 0.08367, 0.05353 |
| 92 | 10, 4, 9, 6, 7 | 0.546, 0.123, 0.1109, 0.09193, 0.05038 |
| 93 | 10, 1 | 0.8101, 0.1689 |
| 94 | 10, 4 | 0.701, 0.2112 |
| 95 | 10, 3, 5 | 0.5587, 0.2669, 0.09852 |
| 96 | 10, 9 | 0.4787, 0.4346 |
| 97 | 4, 10, 1, 2 | 0.3816, 0.2708, 0.2308, 0.0546 |
| 98 | 10, 1, 8 | 0.7772, 0.099, 0.03418 |
| 99 | 10, 6, 4, 7 | 0.468, 0.1845, 0.1621, 0.09936 |
| 100 | 10, 5 | 0.6716, 0.2346 |

## E.4 Simulation with $\mu$ in LFC; Bayes, $\sigma_i^2$ unknown

For $k = 10$ populations with $\mu = (19.5, \ldots, 19.5, 20)$, all variances equal to one, calculated is in each of the 100 runs the minimum set of populations that counts for at least 0.9 of the probability of containing the best one. However, we assume not to know the variances and these are estimated to calculate the Bayesian probabilities. The populations we used are exactly the same as in E.1, meaning populations of in total 100 observations, with samples of size 28.

| Run | Subset | Probs |
|---|---|---|
| 1 | 10, 2, 4, 7, 3 | 0.4375, 0.201, 0.1413, 0.1056, 0.04035 |
| 2 | 10, 6, 5 | 0.777, 0.1033, 0.05553 |
| 3 | 10, 1 | 0.8148, 0.09335 |
| 4 | 10, 6, 5, 8, 7 | 0.5927, 0.12, 0.09995, 0.05094, 0.04055 |
| 5 | 10, 2, 8, 5, 1 | 0.4326, 0.1885, 0.1645, 0.09385, 0.07787 |
| 7 | 3, 10, 4 | 0.6539, 0.2058, 0.09183 |
| 8 | 10, 2, 1, 5 | 0.7019, 0.1196, 0.04642, 0.03539 |
| 9 | 7, 4, 10 | 0.3844, 0.2725, 0.254 |
| 10 | 10, 3 | 0.6995, 0.2237 |
| 11 | 10, 5, 8 | 0.7023, 0.1254, 0.09549 |
| 12 | 10, 6, 2, 1, 8 | 0.3192, 0.2448, 0.1734, 0.1019, 0.09407 |
| 13 | 10, 2, 5, 9 | 0.376, 0.3469, 0.1678, 0.03508 |
| 14 | 10, 8, 7, 1, 5, 3 | 0.281, 0.2197, 0.1756, 0.08384, 0.0759, 0.07544 |
| 15 | 10, 3, 5 | 0.7905, 0.08472, 0.0568 |
| 16 | 9, 10, 5, 8, 6, 1 | 0.3495, 0.2319, 0.2056, 0.05563, 0.04128, 0.03797 |
| 17 | 10, 3, 6 | 0.7754, 0.08447, 0.05302 |
| 18 | 7, 10, 3, 8 | 0.4215, 0.3614, 0.092, 0.0756 |
| 19 | 10, 3 | 0.8888, 0.05947 |
| 20 | 10, 2, 6 | 0.8363, 0.04935, 0.03967 |
| 21 | 10, 9, 6, 3 | 0.5796, 0.1903, 0.07992, 0.06835 |
| 22 | 10 | 0.9987 |
| 23 | 10, 4, 6 | 0.6481, 0.2199, 0.06126 |
| 24 | 10, 9, 2 | 0.5955, 0.2603, 0.05341 |
| 25 | 10, 8 | 0.8771, 0.1109 |
| 26 | 10, 5, 8, 9, 6 | 0.3794, 0.2388, 0.1442, 0.08288, 0.08018 |
| 27 | 10 | 0.9042 |
| 28 | 10 | 0.9973 |
| 29 | 10, 5, 6 | 0.8443, 0.05408, 0.03972 |
| 30 | 10, 8, 7, 1 | 0.6737, 0.08656, 0.08211, 0.06155 |
| 31 | 1, 10, 4 | 0.5007, 0.3719, 0.08036 |
| 32 | 3, 1, 10 | 0.5614, 0.2706, 0.08 |
| 33 | 10, 2, 4, 9 | 0.6963, 0.1289, 0.04237, 0.0391 |
| 34 | 10, 4, 6, 8 | 0.3009, 0.2711, 0.2587, 0.07874 |
| 35 | 10, 9, 1, 3, 5 | 0.6086, 0.09527, 0.09454, 0.09303, 0.08166 |
| 36 | 10, 6, 9 | 0.4637, 0.3611, 0.07543 |
| 37 | 10, 8 | 0.8745, 0.07758 |
| 38 | 10, 2, 4, 8 | 0.6546, 0.1633, 0.05122, 0.04595 |
| 39 | 10 | 0.9911 |
| 40 | 10, 9, 3, 4 | 0.6279, 0.1313, 0.09877, 0.06359 |
| 41 | 10 | 0.9479 |
| 42 | 10, 5, 3 | 0.6509, 0.1717, 0.1124 |
| 43 | 10, 9, 2 | 0.6511, 0.1709, 0.08804 |
| 44 | 10, 6, 8, 5 | 0.3775, 0.3464, 0.1584, 0.1001 |
| 45 | 10 | 0.9487 |
| 46 | 10, 3, 4, 8 | 0.7643, 0.08247, 0.05075, 0.03961 |
| 47 | 10, 9, 3 | 0.5115, 0.3106, 0.1119 |
| 48 | 10 | 0.9701 |
| 49 | 10 | 0.9284 |
| 50 | 10 | 0.9454 |

| Run | Subset | Probs |
|---|---|---|
| 51 | 10 | 0.9043 |
| 52 | 10, 1, 5, 2 | 0.7376, 0.09083, 0.06103, 0.04276 |
| 53 | 10 | 0.9989 |
| 54 | 10, 7, 8 | 0.4046, 0.3637, 0.1862 |
| 55 | 10, 9, 4 | 0.6911, 0.1511, 0.1142 |
| 56 | 10 | 0.9042 |
| 57 | 10 | 0.993 |
| 58 | 10, 3 | 0.8955, 0.06345 |
| 59 | 10, 5, 3 | 0.6387, 0.1645, 0.1385 |
| 60 | 10, 5, 8, 1 | 0.7302, 0.1068, 0.04074, 0.04017 |
| 61 | 10 | 0.9748 |
| 62 | 2, 10, 6, 4 | 0.4187, 0.2455, 0.162, 0.1052 |
| 63 | 10, 1, 4 | 0.8118, 0.06607, 0.05539 |
| 64 | 10, 6, 3, 9, 5, 7 | 0.6428, 0.09011, 0.05919, 0.05692, 0.04796, 0.03852 |
| 65 | 10 | 0.9659 |
| 66 | 10, 5 | 0.6486, 0.2694 |
| 67 | 1, 10 | 0.4654, 0.4418 |
| 68 | 10, 8, 2, 6 | 0.561, 0.1701, 0.1651, 0.06226 |
| 69 | 4, 10 | 0.4672, 0.4464 |
| 70 | 10, 8, 3 | 0.5993, 0.1789, 0.1435 |
| 71 | 10, 1, 8, 5, 6 | 0.4663, 0.1689, 0.1683, 0.06752, 0.06675 |
| 72 | 10, 4, 3, 5, 8 | 0.4348, 0.2491, 0.1058, 0.07312, 0.06042 |
| 73 | 10, 9, 2 | 0.822, 0.04441, 0.03784 |
| 74 | 10, 3 | 0.8894, 0.06821 |
| 75 | 4, 10 | 0.7294, 0.2106 |
| 76 | 10, 7, 5, 6, 2 | 0.2819, 0.2232, 0.2006, 0.1667, 0.0761 |
| 77 | 10, 1, 7, 2 | 0.5209, 0.2133, 0.132, 0.1212 |
| 78 | 10 | 0.957 |
| 79 | 10, 1 | 0.8882, 0.07205 |
| 80 | 10, 8 | 0.8607, 0.07183 |
| 81 | 10 | 0.99 |
| 82 | 10, 8 | 0.8669, 0.1106 |
| 83 | 10, 6, 3, 2 | 0.5654, 0.2072, 0.07577, 0.06966 |
| 84 | 10, 2 | 0.8833, 0.04739 |
| 85 | 10, 1, 4, 7 | 0.5052, 0.306, 0.05613, 0.04988 |
| 86 | 10, 2, 3, 5, 1, 6 | 0.3152, 0.1807, 0.1525, 0.1433, 0.07742, 0.06524 |
| 87 | 1, 10, 9, 7 | 0.372, 0.3283, 0.1623, 0.05653 |
| 88 | 10 | 0.9045 |
| 89 | 10, 8, 6, 3, 1, 2 | 0.3386, 0.1833, 0.1744, 0.1016, 0.07675, 0.06207 |
| 90 | 10, 4, 7, 2 | 0.6751, 0.1014, 0.08615, 0.07758 |
| 91 | 10, 1, 7, 3, 9 | 0.4322, 0.373, 0.05122, 0.03679, 0.03489 |
| 92 | 10, 9, 4, 3, 7 | 0.5602, 0.1476, 0.1058, 0.05957, 0.03506 |
| 93 | 10, 1 | 0.7745, 0.2018 |
| 94 | 10, 4 | 0.792, 0.1094 |
| 95 | 10, 3, 5 | 0.5613, 0.2455, 0.111 |
| 96 | 10, 9, 7 | 0.4562, 0.4085, 0.05601 |
| 97 | 4, 10, 1, 2 | 0.3718, 0.3153, 0.2089, 0.03537 |
| 98 | 10, 1, 8 | 0.7803, 0.09334, 0.04681 |
| 99 | 10, 4, 6, 7 | 0.4447, 0.2304, 0.1776, 0.09655 |
| 100 | 10, 5, 4 | 0.5787, 0.2968, 0.04369 |

## E.5 Standard subset selection; $\mu$ in LFC

We take $\mu$ to be in the LFC, for example, $\mu = (19.5, \ldots, 19.5, 20)$, where $\delta^* = 0.5$, $k = 10$ and the variances all equal to one. We apply the **standard subset selection procedure**, $R_G$ on this instance and choose $P^* = 0.9$. We take the same sample size $n = 28$ and the distance $d$ becomes 15.784.

| Run | SubsetSize | Subset | Status |
|---|---|---|---|
| 1 | 10 | 1, 2, 3, 4, 5, 6, 7, 8, 9, 10 | 2 |
| 2 | 6 | 2, 4, 5, 6, 9, 10 | 2 |
| 3 | 3 | 1, 8, 10 | 2 |
| 4 | 7 | 4, 5, 6, 7, 8, 9, 10 | 2 |
| 5 | 7 | 1, 2, 3, 5, 6, 8, 10 | 2 |
| 6 | 9 | 1, 2, 3, 4, 5, 7, 8, 9, 10 | 2 |
| 7 | 6 | 3, 4, 5, 6, 8, 10 | 2 |
| 8 | 7 | 1, 2, 3, 4, 5, 8, 10 | 2 |
| 9 | 7 | 1, 3, 4, 5, 7, 8, 10 | 2 |
| 10 | 6 | 2, 3, 5, 8, 9, 10 | 2 |
| 11 | 4 | 1, 5, 8, 10 | 2 |
| 12 | 10 | 1, 2, 3, 4, 5, 6, 7, 8, 9, 10 | 2 |
| 13 | 10 | 1, 2, 3, 4, 5, 6, 7, 8, 9, 10 | 2 |
| 14 | 10 | 1, 2, 3, 4, 5, 6, 7, 8, 9, 10 | 2 |
| 15 | 5 | 3, 4, 5, 9, 10 | 2 |
| 16 | 10 | 1, 2, 3, 4, 5, 6, 7, 8, 9, 10 | 2 |
| 17 | 3 | 3, 8, 10 | 2 |
| 18 | 7 | 3, 4, 6, 7, 8, 9, 10 | 2 |
| 19 | 4 | 3, 5, 9, 10 | 2 |
| 20 | 4 | 2, 6, 9, 10 | 2 |
| 21 | 10 | 1, 2, 3, 4, 5, 6, 7, 8, 9, 10 | 2 |
| 22 | 1 | 10 | 1 |
| 23 | 9 | 1, 2, 3, 4, 5, 6, 7, 9, 10 | 2 |
| 24 | 8 | 1, 2, 4, 5, 7, 8, 9, 10 | 2 |
| 25 | 2 | 8, 10 | 2 |
| 26 | 9 | 1, 2, 3, 4, 5, 6, 8, 9, 10 | 2 |
| 27 | 4 | 1, 5, 9, 10 | 2 |
| 28 | 1 | 10 | 1 |
| 29 | 5 | 3, 5, 6, 8, 10 | 2 |
| 30 | 9 | 1, 2, 3, 5, 6, 7, 8, 9, 10 | 2 |
| 31 | 5 | 1, 2, 4, 9, 10 | 2 |
| 32 | 8 | 1, 3, 4, 5, 6, 7, 8, 10 | 2 |
| 33 | 9 | 1, 2, 3, 4, 5, 6, 8, 9, 10 | 2 |
| 34 | 8 | 1, 3, 4, 6, 7, 8, 9, 10 | 2 |
| 35 | 10 | 1, 2, 3, 4, 5, 6, 7, 8, 9, 10 | 2 |
| 36 | 8 | 1, 3, 4, 5, 6, 7, 9, 10 | 2 |
| 37 | 3 | 8, 9, 10 | 2 |
| 38 | 5 | 2, 3, 4, 8, 10 | 2 |
| 39 | 1 | 10 | 1 |
| 40 | 9 | 1, 2, 3, 4, 6, 7, 8, 9, 10 | 2 |
| 41 | 1 | 10 | 1 |
| 42 | 5 | 3, 4, 5, 7, 10 | 2 |
| 43 | 7 | 2, 3, 4, 7, 8, 9, 10 | 2 |
| 44 | 8 | 2, 4, 5, 6, 7, 8, 9, 10 | 2 |
| 45 | 2 | 1, 10 | 2 |
| 46 | 6 | 3, 4, 6, 7, 8, 10 | 2 |
| 47 | 9 | 1, 2, 3, 4, 5, 6, 7, 9, 10 | 2 |
| 48 | 1 | 10 | 1 |
| 49 | 3 | 6, 8, 10 | 2 |
| 50 | 3 | 2, 6, 10 | 2 |

| Run | SubsetSize | Subset | Status |
|---|---|---|---|
| 51 | 3 | 5, 9, 10 | 2 |
| 52 | 4 | 1, 2, 4, 10 | 2 |
| 53 | 1 | 10 | 1 |
| 54 | 6 | 1, 4, 6, 7, 8, 10 | 2 |
| 55 | 6 | 1, 3, 4, 5, 9, 10 | 2 |
| 56 | 3 | 3, 7, 10 | 2 |
| 57 | 2 | 7, 10 | 2 |
| 58 | 5 | 3, 5, 6, 8, 10 | 2 |
| 59 | 6 | 2, 3, 4, 5, 9, 10 | 2 |
| 60 | 7 | 1, 3, 4, 5, 7, 8, 10 | 2 |
| 61 | 1 | 10 | 1 |
| 62 | 9 | 2, 3, 4, 5, 6, 7, 8, 9, 10 | 2 |
| 63 | 3 | 4, 5, 10 | 2 |
| 64 | 10 | 1, 2, 3, 4, 5, 6, 7, 8, 9, 10 | 2 |
| 65 | 2 | 8, 10 | 2 |
| 66 | 6 | 1, 3, 5, 6, 7, 10 | 2 |
| 67 | 5 | 1, 4, 7, 8, 10 | 2 |
| 68 | 8 | 1, 2, 3, 6, 7, 8, 9, 10 | 2 |
| 69 | 7 | 1, 2, 3, 4, 7, 8, 10 | 2 |
| 70 | 5 | 1, 3, 8, 9, 10 | 2 |
| 71 | 9 | 1, 2, 3, 4, 5, 6, 8, 9, 10 | 2 |
| 72 | 8 | 1, 2, 3, 4, 5, 6, 8, 10 | 2 |
| 73 | 4 | 1, 8, 9, 10 | 2 |
| 74 | 3 | 3, 5, 10 | 2 |
| 75 | 5 | 2, 4, 5, 7, 10 | 2 |
| 76 | 9 | 1, 2, 3, 5, 6, 7, 8, 9, 10 | 2 |
| 77 | 5 | 1, 2, 4, 7, 10 | 2 |
| 78 | 1 | 10 | 1 |
| 79 | 4 | 1, 7, 8, 10 | 2 |
| 80 | 4 | 1, 7, 8, 10 | 2 |
| 81 | 1 | 10 | 1 |
| 82 | 2 | 8, 10 | 2 |
| 83 | 8 | 1, 2, 3, 5, 6, 8, 9, 10 | 2 |
| 84 | 4 | 2, 4, 5, 10 | 2 |
| 85 | 8 | 1, 3, 4, 5, 6, 7, 9, 10 | 2 |
| 86 | 8 | 1, 2, 3, 4, 5, 6, 9, 10 | 2 |
| 87 | 8 | 1, 2, 3, 4, 6, 7, 9, 10 | 2 |
| 88 | 2 | 7, 10 | 2 |
| 89 | 9 | 1, 2, 3, 4, 5, 6, 7, 8, 10 | 2 |
| 90 | 5 | 1, 2, 4, 7, 10 | 2 |
| 91 | 9 | 1, 2, 3, 5, 6, 7, 8, 9, 10 | 2 |
| 92 | 9 | 1, 2, 3, 4, 6, 7, 8, 9, 10 | 2 |
| 93 | 3 | 1, 4, 10 | 2 |
| 94 | 6 | 1, 2, 3, 4, 7, 10 | 2 |
| 95 | 9 | 1, 2, 3, 4, 5, 6, 7, 9, 10 | 2 |
| 96 | 5 | 1, 4, 7, 9, 10 | 2 |
| 97 | 9 | 1, 2, 3, 4, 5, 6, 8, 9, 10 | 2 |
| 98 | 7 | 1, 2, 3, 4, 8, 9, 10 | 2 |
| 99 | 10 | 1, 2, 3, 4, 5, 6, 7, 8, 9, 10 | 2 |
| 100 | 5 | 3, 4, 5, 9, 10 | 2 |

Subset size sum = 582, $CS_1$= 9, $FS_1$=0, $CS_+$= 91, $FS_+$= 0.

## E.6 Simulation with $\mu$ in LFC; Sequential procedure

For $k = 10$ populations with $\mu = (19.5, \ldots, 19.5, 20)$; all variances equal to one, for $\alpha = 0.1$, hence boundary $g = 8.97727$, the result is, for exactly the same populations as in E.1 for the first 100 runs:

| run | obs | stage | selected | status | run | obs | stage | selected | status |
|---|---|---|---|---|---|---|---|---|---|
| 1 | 260 | 45 | 10 | 1 | 51 | 176 | 27 | 10 | 1 |
| 2 | 128 | 14 | 10 | 1 | 52 | 146 | 26 | 10 | 1 |
| 3 | 119 | 23 | 10 | 1 | 53 | 120 | 17 | 10 | 1 |
| 4 | 139 | 23 | 10 | 1 | 54 | 232 | 36 | 10 | 1 |
| 5 | 193 | 48 | 10 | 1 | 55 | 219 | 40 | 10 | 1 |
| 6 | 216 | 41 | 10 | 1 | 56 | 141 | 25 | 10 | 1 |
| 7 | 189 | 41 | 10 | 1 | 57 | 111 | 15 | 10 | 1 |
| 8 | 169 | 24 | 10 | 1 | 58 | 142 | 21 | 10 | 1 |
| 9 | 269 | 49 | 10 | 1 | 59 | 157 | 25 | 10 | 1 |
| 10 | 201 | 49 | 10 | 1 | 60 | 188 | 24 | 10 | 1 |
| 11 | 199 | 32 | 10 | 1 | 61 | 130 | 20 | 10 | 1 |
| 12 | 253 | 43 | 10 | 1 | 62 | 308 | 78 | 10 | 1 |
| 13 | 202 | 38 | 10 | 1 | 63 | 171 | 28 | 10 | 1 |
| 14 | 305 | 52 | 10 | 1 | 64 | 160 | 23 | 10 | 1 |
| 15 | 202 | 25 | 10 | 1 | 65 | 89 | 17 | 10 | 1 |
| 16 | 293 | 58 | 10 | 1 | 66 | 172 | 40 | 10 | 1 |
| 17 | 130 | 22 | 10 | 1 | 67 | 217 | 44 | 10 | 1 |
| 18 | 327 | 94 | 10 | 1 | 68 | 226 | 31 | 10 | 1 |
| 19 | 102 | 23 | 10 | 1 | 69 | 256 | 61 | 10 | 1 |
| 20 | 107 | 15 | 10 | 1 | 70 | 78 | 9 | 10 | 1 |
| 21 | 256 | 49 | 10 | 1 | 71 | 246 | 42 | 10 | 1 |
| 22 | 118 | 15 | 10 | 1 | 72 | 261 | 37 | 10 | 1 |
| 23 | 192 | 32 | 10 | 1 | 73 | 126 | 16 | 10 | 1 |
| 24 | 224 | 38 | 10 | 1 | 74 | 122 | 19 | 10 | 1 |
| 25 | 91 | 11 | 10 | 1 | 75 | 245 | 46 | 7 | 3 |
| 26 | 283 | 41 | 10 | 1 | 76 | 263 | 44 | 10 | 1 |
| 27 | 91 | 15 | 10 | 1 | 77 | 193 | 40 | 10 | 1 |
| 28 | 125 | 18 | 10 | 1 | 78 | 116 | 16 | 10 | 1 |
| 29 | 162 | 22 | 10 | 1 | 79 | 178 | 26 | 10 | 1 |
| 30 | 144 | 19 | 10 | 1 | 80 | 175 | 25 | 10 | 1 |
| 31 | 207 | 52 | 10 | 1 | 81 | 114 | 16 | 10 | 1 |
| 32 | 293 | 79 | 3 | 3 | 82 | 162 | 20 | 10 | 1 |
| 33 | 256 | 42 | 10 | 1 | 83 | 233 | 31 | 10 | 1 |
| 34 | 262 | 47 | 10 | 1 | 84 | 103 | 25 | 10 | 1 |
| 35 | 223 | 31 | 10 | 1 | 85 | 176 | 25 | 10 | 1 |
| 36 | 225 | 50 | 10 | 1 | 86 | 286 | 58 | 10 | 1 |
| 37 | 142 | 19 | 10 | 1 | 87 | 279 | 59 | 10 | 1 |
| 38 | 93 | 21 | 10 | 1 | 88 | 110 | 16 | 10 | 1 |
| 39 | 140 | 23 | 10 | 1 | 89 | 304 | 69 | 10 | 1 |
| 40 | 234 | 32 | 10 | 1 | 90 | 118 | 24 | 10 | 1 |
| 41 | 156 | 22 | 10 | 1 | 91 | 210 | 46 | 10 | 1 |
| 42 | 184 | 30 | 10 | 1 | 92 | 306 | 47 | 10 | 1 |
| 43 | 135 | 23 | 10 | 1 | 93 | 166 | 32 | 10 | 1 |
| 44 | 278 | 54 | 10 | 1 | 94 | 173 | 33 | 10 | 1 |
| 45 | 112 | 15 | 10 | 1 | 95 | 184 | 32 | 10 | 1 |
| 46 | 137 | 25 | 10 | 1 | 96 | 194 | 41 | 10 | 1 |
| 47 | 126 | 18 | 9 | 3 | 97 | 315 | 61 | 10 | 1 |
| 48 | 119 | 14 | 10 | 1 | 98 | 207 | 44 | 10 | 1 |
| 49 | 158 | 21 | 10 | 1 | 99 | 206 | 39 | 10 | 1 |
| 50 | 157 | 25 | 10 | 1 | 100 | 189 | 35 | 10 | 1 |

| run | obs | stage | selected | status | run | obs | stage | selected | status |
|---|---|---|---|---|---|---|---|---|---|
| 101 | 225 | 40 | 5 | 3 | 151 | 289 | 54 | 10 | 1 |
| 102 | 180 | 25 | 10 | 1 | 152 | 77 | 9 | 10 | 1 |
| 103 | 113 | 14 | 10 | 1 | 153 | 182 | 54 | 10 | 1 |
| 104 | 384 | 66 | 10 | 1 | 154 | 193 | 35 | 10 | 1 |
| 105 | 242 | 31 | 10 | 1 | 155 | 126 | 18 | 10 | 1 |
| 106 | 191 | 39 | 10 | 1 | 156 | 207 | 35 | 1 | 3 |
| 107 | 285 | 57 | 10 | 1 | 157 | 181 | 31 | 10 | 1 |
| 108 | 193 | 37 | 10 | 1 | 158 | 297 | 100 | 2 | 3 |
| 109 | 169 | 35 | 10 | 1 | 159 | 183 | 43 | 10 | 1 |
| 110 | 119 | 21 | 10 | 1 | 160 | 144 | 19 | 10 | 1 |
| 111 | 105 | 14 | 10 | 1 | 161 | 209 | 29 | 10 | 1 |
| 112 | 201 | 40 | 10 | 1 | 162 | 148 | 23 | 10 | 1 |
| 113 | 173 | 32 | 10 | 1 | 163 | 172 | 24 | 10 | 1 |
| 114 | 167 | 35 | 10 | 1 | 164 | 179 | 31 | 10 | 1 |
| 115 | 187 | 40 | 10 | 1 | 165 | 233 | 33 | 10 | 1 |
| 116 | 155 | 21 | 10 | 1 | 166 | 155 | 38 | 10 | 1 |
| 117 | 127 | 16 | 10 | 1 | 167 | 173 | 27 | 10 | 1 |
| 118 | 204 | 37 | 10 | 1 | 168 | 84 | 13 | 10 | 1 |
| 119 | 181 | 41 | 10 | 1 | 169 | 185 | 30 | 10 | 1 |
| 120 | 254 | 36 | 10 | 1 | 170 | 195 | 52 | 10 | 1 |
| 121 | 165 | 20 | 10 | 1 | 171 | 78 | 11 | 10 | 1 |
| 122 | 172 | 22 | 10 | 1 | 172 | 273 | 43 | 10 | 1 |
| 123 | 118 | 17 | 10 | 1 | 173 | 146 | 23 | 10 | 1 |
| 124 | 219 | 28 | 10 | 1 | 174 | 179 | 37 | 10 | 1 |
| 125 | 110 | 16 | 10 | 1 | 175 | 171 | 18 | 10 | 1 |
| 126 | 199 | 41 | 10 | 1 | 176 | 254 | 42 | 10 | 1 |
| 127 | 235 | 42 | 10 | 1 | 177 | 265 | 46 | 10 | 1 |
| 128 | 187 | 30 | 10 | 1 | 178 | 207 | 35 | 10 | 1 |
| 129 | 143 | 22 | 10 | 1 | 179 | 245 | 56 | 10 | 1 |
| 130 | 216 | 39 | 10 | 1 | 180 | 164 | 24 | 10 | 1 |
| 131 | 258 | 51 | 10 | 1 | 181 | 160 | 31 | 10 | 1 |
| 132 | 310 | 76 | 10 | 1 | 182 | 105 | 12 | 10 | 1 |
| 133 | 230 | 43 | 10 | 1 | 183 | 168 | 42 | 10 | 1 |
| 134 | 225 | 51 | 10 | 1 | 184 | 125 | 23 | 10 | 1 |
| 135 | 128 | 21 | 10 | 1 | 185 | 133 | 17 | 10 | 1 |
| 136 | 134 | 20 | 10 | 1 | 186 | 171 | 22 | 10 | 1 |
| 137 | 249 | 45 | 10 | 1 | 187 | 128 | 19 | 10 | 1 |
| 138 | 245 | 60 | 10 | 1 | 188 | 85 | 14 | 10 | 1 |
| 139 | 189 | 59 | 10 | 1 | 189 | 264 | 51 | 8 | 3 |
| 140 | 159 | 18 | 10 | 1 | 190 | 103 | 18 | 10 | 1 |
| 141 | 111 | 17 | 10 | 1 | 191 | 89 | 14 | 10 | 1 |
| 142 | 171 | 35 | 10 | 1 | 192 | 272 | 66 | 10 | 1 |
| 143 | 182 | 28 | 10 | 1 | 193 | 280 | 44 | 10 | 1 |
| 144 | 120 | 14 | 10 | 1 | 194 | 200 | 27 | 10 | 1 |
| 145 | 190 | 24 | 10 | 1 | 195 | 190 | 23 | 2 | 3 |
| 146 | 273 | 78 | 10 | 1 | 196 | 165 | 24 | 10 | 1 |
| 147 | 225 | 40 | 10 | 1 | 197 | 130 | 18 | 10 | 1 |
| 148 | 252 | 31 | 10 | 1 | 198 | 154 | 32 | 9 | 3 |
| 149 | 138 | 23 | 10 | 1 | 199 | 254 | 46 | 10 | 1 |
| 150 | 127 | 31 | 10 | 1 | 200 | 379 | 100 | 2 | 3 |

For this simulation, the total results are:

CS = 190

FS = 8

Runs without decision (stage > 100) = 2

Total number of observations is 37409

The maximum number of obs was 384

## E.7 Simulation with $\mu$ not in LFC; $R_{n,c}$

We take $\mu$ not in the LFC, but, $\mu = (19, 19, 19, 19.1, 19.2, 19.3, 19.3, 19.4, 19.5, 20)$, where $\delta^* = 0.5$, $k = 10$ and the variances all equal to one. We apply the **preference threshold procedure**, $R_{n,c}$ on this instance and choose $P^* = 0.75$ and $Q^* = 0.9$. For these values, the sample size becomes $n = 28$ and the threshold is $c = 1.88683$. Because $n$ is rounded, the practical values for $P^*$ and $Q^*$ become $P^* = 0.7568$ and $Q^* = 0.9028$. Here we take the sample sums for samples of size $n = 28$ directly from the random generator, distributed $N(28 * \mu_i, \sqrt{28})$. For different starting values of the random generator the results after 1000 runs are:

| $\bar{s}$ | 1.057 | Seed | 1 |
|---|---|---|---|
| $CS_1$ | 925 | $FS_1$ | 25 |
| $CS_+$ | 47 | $FS_+$ | 3 |
| $CS_2$ | 972 | $FS_2$ | 28 |

| $\bar{s}$ | 1.058 | Seed | 2 |
|---|---|---|---|
| $CS_1$ | 934 | $FS_1$ | 15 |
| $CS_+$ | 48 | $FS_+$ | 3 |
| $CS_2$ | 982 | $FS_2$ | 18 |

| $\bar{s}$ | 1.063 | Seed | 3 |
|---|---|---|---|
| $CS_1$ | 920 | $FS_1$ | 26 |
| $CS_+$ | 53 | $FS_+$ | 1 |
| $CS_2$ | 973 | $FS_2$ | 27 |

| $\bar{s}$ | 1.056 | Seed | 4 |
|---|---|---|---|
| $CS_1$ | 927 | $FS_1$ | 24 |
| $CS_+$ | 46 | $FS_+$ | 3 |
| $CS_2$ | 973 | $FS_2$ | 27 |

| $\bar{s}$ | 1.076 | Seed | 5 |
|---|---|---|---|
| $CS_1$ | 921 | $FS_1$ | 16 |
| $CS_+$ | 60 | $FS_+$ | 3 |
| $CS_2$ | 981 | $FS_2$ | 17 |

## E.8 Simulation selection $t = 3$ best populations with $\mu$ in LFC; $R^t_{n,c}$

We take $\mu$ to be in the LFC, for example, $\mu = (19.5, \ldots, 19.5, 20, 20, 20)$, $k = 10$ where we want to select the $t = 3$ best populations, and the variances all equal to one. We apply the **preference threshold procedure**, $R^t_{n,c}$ on this instance and choose $P^* = 0.75$, $Q^* = 0.9$ and $\delta^* = 0.5$. For these values, the sample size becomes $n = 40$ and the threshold is $c = 1.98143$. In this simulation we take the sample sums for samples of size $n = 40$ directly from the random generator, distributed $N(40 * \mu_i, \sqrt{40})$. For different starting values of the random generator the results after each time 1000 runs are:

| $\bar{s}$ | 3.223 | Seed | 1 |
|---|---|---|---|
| $CS_t$ | 732 | $FS_t$ | 89 |
| $CS_+$ | 161 | $FS_+$ | 18 |
| $CS_{ts}$ | 893 | $FS_{ts}$ | 107 |

| $\bar{s}$ | 3.189 | Seed | 2 |
|---|---|---|---|
| $CS_t$ | 781 | $FS_t$ | 61 |
| $CS_+$ | 146 | $FS_+$ | 12 |
| $CS_{ts}$ | 927 | $FS_{ts}$ | 73 |

| $\bar{s}$ | 3.193 | Seed | 3 |
|---|---|---|---|
| $CS_t$ | 746 | $FS_t$ | 90 |
| $CS_+$ | 145 | $FS_+$ | 19 |
| $CS_{ts}$ | 891 | $FS_{ts}$ | 109 |

| $\bar{s}$ | 3.201 | Seed | 4 |
|---|---|---|---|
| $CS_t$ | 766 | $FS_t$ | 63 |
| $CS_+$ | 153 | $FS_+$ | 18 |
| $CS_{ts}$ | 919 | $FS_{ts}$ | 81 |

| $\bar{s}$ | 3.203 | Seed | 5 |
|---|---|---|---|
| $CS_t$ | 762 | $FS_t$ | 66 |
| $CS_+$ | 154 | $FS_+$ | 18 |
| $CS_{ts}$ | 916 | $FS_{ts}$ | 84 |

| $\bar{s}$ | 3.217 | Seed | 6 |
|---|---|---|---|
| $CS_t$ | 741 | $FS_t$ | 79 |
| $CS_+$ | 158 | $FS_+$ | 22 |
| $CS_{ts}$ | 899 | $FS_{ts}$ | 101 |

| $\bar{s}$ | 3.243 | Seed | 7 |
|---|---|---|---|
| $CS_t$ | 736 | $FS_t$ | 75 |
| $CS_+$ | 180 | $FS_+$ | 9 |
| $CS_{ts}$ | 916 | $FS_{ts}$ | 84 |

## E.9 Robustness simulation in LFC

The $\mu$ we have chosen is in the LFC, $\mu = (19.5, \ldots, 19.5, 20)$, where $\delta^* = 0.5$ and we design the experiment as if $\sigma = 1$ for the $k = 10$ populations. Hence we take samples of size $n = 28$ and have $c = 1.88683$ for $P^* = 0.75$ and $Q^* = 0.9$. However, the real standard deviation we use to take the samples, is $\sigma = 1.25$ for all populations. This situation can appear when $\gamma = 1.25$. The sample sums are directly gained from the random generator with initializing seed value equal to 1.

| run | SubsetSize | Subset | Status | run | SubsetSize | Subset | Status |
|---|---|---|---|---|---|---|---|
| 1 | 1 | 10 | 1 | 51 | 1 | 10 | 1 |
| 2 | 1 | 10 | 1 | 52 | 2 | 1, 10 | 2 |
| 3 | 1 | 10 | 1 | 53 | 1 | 10 | 1 |
| 4 | 1 | 10 | 1 | 54 | 1 | 10 | 1 |
| 5 | 1 | 10 | 1 | 55 | 1 | 8 | 3 |
| 6 | 1 | 10 | 1 | 56 | 1 | 10 | 1 |
| 7 | 1 | 10 | 1 | 57 | 2 | 5, 10 | 2 |
| 8 | 1 | 10 | 1 | 58 | 1 | 10 | 1 |
| 9 | 1 | 10 | 1 | 59 | 1 | 10 | 1 |
| 10 | 1 | 10 | 1 | 60 | 1 | 10 | 1 |
| 11 | 2 | 3, 10 | 2 | 61 | 1 | 6 | 3 |
| 12 | 1 | 10 | 1 | 62 | 1 | 10 | 1 |
| 13 | 1 | 10 | 1 | 63 | 2 | 7, 8 | 4 |
| 14 | 3 | 2, 3, 10 | 2 | 64 | 1 | 10 | 1 |
| 15 | 1 | 10 | 1 | 65 | 1 | 10 | 1 |
| 16 | 1 | 10 | 1 | 66 | 1 | 8 | 3 |
| 17 | 3 | 2, 5, 6 | 4 | 67 | 1 | 10 | 1 |
| 18 | 1 | 10 | 1 | 68 | 1 | 5 | 3 |
| 19 | 1 | 10 | 1 | 69 | 1 | 10 | 1 |
| 20 | 1 | 10 | 1 | 70 | 1 | 10 | 1 |
| 21 | 1 | 8 | 3 | 71 | 2 | 4, 10 | 2 |
| 22 | 1 | 10 | 1 | 72 | 1 | 1 | 3 |
| 23 | 2 | 7, 10 | 2 | 73 | 1 | 10 | 1 |
| 24 | 1 | 10 | 1 | 74 | 1 | 10 | 1 |
| 25 | 1 | 10 | 1 | 75 | 1 | 10 | 1 |
| 26 | 1 | 10 | 1 | 76 | 1 | 10 | 1 |
| 27 | 2 | 4, 10 | 2 | 77 | 3 | 2, 3, 10 | 2 |
| 28 | 3 | 1, 8, 10 | 2 | 78 | 1 | 7 | 3 |
| 29 | 1 | 10 | 1 | 79 | 1 | 10 | 1 |
| 30 | 1 | 6 | 3 | 80 | 3 | 1, 3, 9 | 4 |
| 31 | 1 | 10 | 1 | 81 | 1 | 10 | 1 |
| 32 | 1 | 10 | 1 | 82 | 2 | 8, 10 | 2 |
| 33 | 1 | 10 | 1 | 83 | 1 | 10 | 1 |
| 34 | 4 | 1, 2, 9, 10 | 2 | 84 | 1 | 10 | 1 |
| 35 | 1 | 9 | 3 | 85 | 1 | 10 | 1 |
| 36 | 1 | 10 | 1 | 86 | 1 | 10 | 1 |
| 37 | 1 | 10 | 1 | 87 | 1 | 10 | 1 |
| 38 | 1 | 10 | 1 | 88 | 1 | 4 | 3 |
| 39 | 1 | 10 | 1 | 89 | 1 | 6 | 3 |
| 40 | 2 | 2, 7 | 4 | 90 | 1 | 10 | 1 |
| 41 | 1 | 10 | 1 | 91 | 1 | 10 | 1 |
| 42 | 1 | 10 | 1 | 92 | 2 | 1, 10 | 2 |
| 43 | 1 | 3 | 3 | 93 | 1 | 10 | 1 |
| 44 | 1 | 10 | 1 | 94 | 1 | 10 | 1 |
| 45 | 1 | 10 | 1 | 95 | 1 | 10 | 1 |
| 46 | 1 | 10 | 1 | 96 | 1 | 10 | 1 |
| 47 | 1 | 10 | 1 | 97 | 1 | 10 | 1 |
| 48 | 1 | 10 | 1 | 98 | 2 | 5, 8 | 4 |
| 49 | 1 | 10 | 1 | 99 | 2 | 6, 10 | 2 |
| 50 | 1 | 10 | 1 | 100 | 1 | 10 | 1 |

Subset size sum = 125, $CS_1$= 70, $FS_1$=12, $CS_+$= 13, $FS_+$= 5.

After a total of 1000 runs, we have:

| $\bar{s}$ | 1.278 | Seed | 1 |
|---|---|---|---|
| $CS_1$ | 629 | $FS_1$ | 158 |
| $CS_+$ | 162 | $FS_+$ | 51 |
| $CS_2$ | 791 | $FS_2$ | 209 |

| $\bar{s}$ | 1.276 | Seed | 2 |
|---|---|---|---|
| $CS_1$ | 619 | $FS_1$ | 160 |
| $CS_+$ | 181 | $FS_+$ | 40 |
| $CS_2$ | 800 | $FS_2$ | 200 |

| $\bar{s}$ | 1.244 | Seed | 3 |
|---|---|---|---|
| $CS_1$ | 639 | $FS_1$ | 165 |
| $CS_+$ | 151 | $FS_+$ | 45 |
| $CS_2$ | 790 | $FS_2$ | 210 |

| $\bar{s}$ | 1.268 | Seed | 4 |
|---|---|---|---|
| $CS_1$ | 611 | $FS_1$ | 174 |
| $CS_+$ | 164 | $FS_+$ | 51 |
| $CS_2$ | 775 | $FS_2$ | 225 |

| $\bar{s}$ | 1.246 | Seed | 5 |
|---|---|---|---|
| $CS_1$ | 627 | $FS_1$ | 156 |
| $CS_+$ | 172 | $FS_+$ | 45 |
| $CS_2$ | 799 | $FS_2$ | 201 |

## E.10 Preference threshold procedure with $\sigma = 1.25$

In this instance, $\mu = (19.5, \ldots, 19.5, 20)$ and the standard deviation is for all $k = 10$ populations equal to $\sigma = 1.25$. $P^* = 0.75, Q^* = 0.9$ and $\delta^* = 0.5$. Therefore, $c = 2.94818$ and $n = 43.0106$ rounded upward to $n = 44$. Therefore the practical $\tau_1$ becomes 2.297736 and the practical value for $P^*$ is $P^* = 0.759104$. The practical value for $\tau_2$ becomes 3.00886 and the realized $Q^*$ is $Q^* = 0.903761$. We take our the sample sums directly from the random generator with different seed values.

| $\bar{s}$ | 1.198 | Seed | 1 |
|---|---|---|---|
| $CS_1$ | 770 | $FS_1$ | 78 |
| $CS_+$ | 132 | $FS_+$ | 20 |
| $CS_2$ | 902 | $FS_2$ | 98 |

| $\bar{s}$ | 1.205 | Seed | 2 |
|---|---|---|---|
| $CS_1$ | 769 | $FS_1$ | 66 |
| $CS_+$ | 144 | $FS_+$ | 21 |
| $CS_2$ | 913 | $FS_2$ | 87 |

| $\bar{s}$ | 1.201 | Seed | 3 |
|---|---|---|---|
| $CS_1$ | 762 | $FS_1$ | 76 |
| $CS_+$ | 139 | $FS_+$ | 23 |
| $CS_2$ | 901 | $FS_2$ | 99 |

| $\bar{s}$ | 1.228 | Seed | 4 |
|---|---|---|---|
| $CS_1$ | 756 | $FS_1$ | 69 |
| $CS_+$ | 158 | $FS_+$ | 17 |
| $CS_2$ | 914 | $FS_2$ | 86 |

| $\bar{s}$ | 1.196 | Seed | 5 |
|---|---|---|---|
| $CS_1$ | 770 | $FS_1$ | 66 |
| $CS_+$ | 136 | $FS_+$ | 28 |
| $CS_2$ | 906 | $FS_2$ | 94 |

| $\bar{s}$ | 1.230 | Seed | 6 |
|---|---|---|---|
| $CS_1$ | 741 | $FS_1$ | 83 |
| $CS_+$ | 156 | $FS_+$ | 20 |
| $CS_2$ | 897 | $FS_2$ | 103 |

| $\bar{s}$ | 1.192 | Seed | 7 |
|---|---|---|---|
| $CS_1$ | 769 | $FS_1$ | 74 |
| $CS_+$ | 123 | $FS_+$ | 34 |
| $CS_2$ | 892 | $FS_2$ | 108 |

## E.11 Robustness; Ordinary procedure but $\gamma = 1.25$

We take $\mu$ in the LFC ($\mu = (19.5, \ldots, 19.5, 20)$), the standard deviations to be $\sigma = (0.8, 0.9, 1, 1.1, 1.2, 1.25, 1.25, 1.25, 1.25, 1)$, where $\delta^* = 0.5$, $P^* = 0.75$ and $Q^* = 0.9$. However, in our design we assume all variances to be equal to one, so the sample size should be $n = 28$ and $c = 1.88683$. This situation can appear when $\gamma = 1.25$. Because $n$ is rounded, the practical values for $P^*$ and $Q^*$ become $P^* = 0.756833$ and $Q^* = 0.902824$, because we now have $r_1 = 2.28917$ and $r_2 = 3.00233$ (if the variances were really all equal to one).

| Run | SubsetSize | Subset | Status | Run | SubsetSize | Subset | Status |
|---|---|---|---|---|---|---|---|
| 1 | 1 | 10 | 1 | 51 | 1 | 10 | 1 |
| 2 | 1 | 10 | 1 | 52 | 1 | 10 | 1 |
| 3 | 1 | 10 | 1 | 53 | 1 | 10 | 1 |
| 4 | 1 | 10 | 1 | 54 | 1 | 7 | 3 |
| 5 | 1 | 10 | 1 | 55 | 1 | 10 | 1 |
| 6 | 1 | 10 | 1 | 56 | 1 | 10 | 1 |
| 7 | 1 | 3 | 3 | 57 | 1 | 10 | 1 |
| 8 | 1 | 10 | 1 | 58 | 1 | 10 | 1 |
| 9 | 1 | 7 | 3 | 59 | 1 | 10 | 1 |
| 10 | 1 | 10 | 1 | 60 | 1 | 10 | 1 |
| 11 | 1 | 10 | 1 | 61 | 1 | 10 | 1 |
| 12 | 2 | 6, 10 | 2 | 62 | 3 | 2, 6, 10 | 2 |
| 13 | 3 | 2, 5, 10 | 2 | 63 | 1 | 10 | 1 |
| 14 | 3 | 7, 8, 10 | 2 | 64 | 1 | 10 | 1 |
| 15 | 1 | 10 | 1 | 65 | 1 | 10 | 1 |
| 16 | 1 | 9 | 3 | 66 | 2 | 5, 10 | 2 |
| 17 | 1 | 10 | 1 | 67 | 1 | 10 | 1 |
| 18 | 1 | 7 | 3 | 68 | 1 | 10 | 1 |
| 19 | 1 | 10 | 1 | 69 | 2 | 4, 10 | 2 |
| 20 | 1 | 10 | 1 | 70 | 1 | 10 | 1 |
| 21 | 1 | 10 | 1 | 71 | 2 | 8, 10 | 2 |
| 22 | 1 | 10 | 1 | 72 | 1 | 10 | 1 |
| 23 | 1 | 10 | 1 | 73 | 1 | 10 | 1 |
| 24 | 2 | 9, 10 | 2 | 74 | 1 | 10 | 1 |
| 25 | 1 | 10 | 1 | 75 | 1 | 4 | 3 |
| 26 | 2 | 5, 10 | 2 | 76 | 4 | 5, 6, 7, 10 | 2 |
| 27 | 1 | 10 | 1 | 77 | 1 | 10 | 1 |
| 28 | 1 | 10 | 1 | 78 | 1 | 10 | 1 |
| 29 | 1 | 10 | 1 | 79 | 1 | 10 | 1 |
| 30 | 1 | 10 | 1 | 80 | 1 | 10 | 1 |
| 31 | 2 | 1, 10 | 2 | 81 | 1 | 10 | 1 |
| 32 | 1 | 3 | 3 | 82 | 1 | 10 | 1 |
| 33 | 1 | 10 | 1 | 83 | 2 | 6, 10 | 2 |
| 34 | 1 | 6 | 3 | 84 | 1 | 10 | 1 |
| 35 | 1 | 10 | 1 | 85 | 1 | 10 | 1 |
| 36 | 2 | 6, 10 | 2 | 86 | 3 | 2, 5, 10 | 2 |
| 37 | 1 | 10 | 1 | 87 | 2 | 1, 10 | 2 |
| 38 | 1 | 10 | 1 | 88 | 1 | 10 | 1 |
| 39 | 1 | 10 | 1 | 89 | 3 | 6, 8, 10 | 2 |
| 40 | 1 | 10 | 1 | 90 | 1 | 10 | 1 |
| 41 | 1 | 10 | 1 | 91 | 1 | 10 | 1 |
| 42 | 1 | 10 | 1 | 92 | 1 | 10 | 1 |
| 43 | 1 | 10 | 1 | 93 | 1 | 10 | 1 |
| 44 | 1 | 6 | 3 | 94 | 1 | 10 | 1 |
| 45 | 1 | 10 | 1 | 95 | 1 | 10 | 1 |
| 46 | 1 | 10 | 1 | 96 | 1 | 9 | 3 |
| 47 | 2 | 9, 10 | 2 | 97 | 1 | 4 | 3 |
| 48 | 1 | 10 | 1 | 98 | 1 | 10 | 1 |
| 49 | 1 | 10 | 1 | 99 | 1 | 10 | 1 |
| 50 | 1 | 10 | 1 | 100 | 1 | 10 | 1 |

Subset size sum = 124, CS1= 72, FS1=11, CS2= 17, FS2= 0.

After 1000 runs we have for different seed values (when the sample sums are directly taken from the random generator):

| $\bar{s}$ | 1.213 | Seed | 1 |
|---|---|---|---|
| $CS_1$ | 702 | $FS_1$ | 121 |
| $CS_+$ | 156 | $FS_+$ | 21 |
| $CS_2$ | 858 | $FS_2$ | 142 |

| $\bar{s}$ | 1.237 | Seed | 2 |
|---|---|---|---|
| $CS_1$ | 696 | $FS_1$ | 115 |
| $CS_+$ | 174 | $FS_+$ | 15 |
| $CS_2$ | 870 | $FS_2$ | 130 |

| $\bar{s}$ | 1.206 | Seed | 3 |
|---|---|---|---|
| $CS_1$ | 706 | $FS_1$ | 129 |
| $CS_+$ | 145 | $FS_+$ | 20 |
| $CS_2$ | 851 | $FS_2$ | 149 |

| $\bar{s}$ | 1.249 | Seed | 4 |
|---|---|---|---|
| $CS_1$ | 694 | $FS_1$ | 108 |
| $CS_+$ | 172 | $FS_+$ | 26 |
| $CS_2$ | 866 | $FS_2$ | 134 |

| $\bar{s}$ | 1.198 | Seed | 5 |
|---|---|---|---|
| $CS_1$ | 717 | $FS_1$ | 118 |
| $CS_+$ | 142 | $FS_+$ | 23 |
| $CS_2$ | 859 | $FS_2$ | 141 |

## E.12 Robustness; Robust design with $\gamma = 1.25$

We take $\mu$ in the LFC ($\mu = (19.5, \ldots, 19.5, 20)$), the standard deviations to be $\sigma = (0.8, 0.9, 1, 1.1, 1.2, 1.25, 1.25, 1.25, 1.25, 1)$, where $\delta^* = 0.5$, $P^* = 0.75$ and $Q^* = 0.9$. In the robust design we have $\sigma_0 = 1$ and $\gamma = 1.25$, hence we need to take $n = 38$ and $c = 2.39279$. For this rounded value of $n$ the practical values for $P^*$ and $Q^*$ become $P^* = 0.75206$ and $Q^* = 0.900863$, because we now have $\tau_{0,1} = 2.69405$ and $\tau_{0,2} = 3.47037$.

| Run | SubsetSize | Subset | Status | Run | SubsetSize | Subset | Status |
|---|---|---|---|---|---|---|---|
| 1 | 1 | 10 | 1 | 51 | 1 | 10 | 1 |
| 2 | 1 | 10 | 1 | 52 | 1 | 10 | 1 |
| 3 | 1 | 10 | 1 | 53 | 1 | 10 | 1 |
| 4 | 1 | 10 | 1 | 54 | 1 | 10 | 1 |
| 5 | 2 | 5, 10 | 2 | 55 | 1 | 10 | 1 |
| 6 | 1 | 10 | 1 | 56 | 1 | 10 | 1 |
| 7 | 1 | 10 | 1 | 57 | 1 | 10 | 1 |
| 8 | 1 | 10 | 1 | 58 | 1 | 10 | 1 |
| 9 | 1 | 4 | 3 | 59 | 1 | 10 | 1 |
| 10 | 1 | 10 | 1 | 60 | 1 | 10 | 1 |
| 11 | 1 | 10 | 1 | 61 | 1 | 10 | 1 |
| 12 | 2 | 6, 10 | 2 | 62 | 3 | 2, 6, 10 | 2 |
| 13 | 1 | 10 | 1 | 63 | 1 | 10 | 1 |
| 14 | 1 | 5 | 3 | 64 | 1 | 10 | 1 |
| 15 | 1 | 10 | 1 | 65 | 1 | 10 | 1 |
| 16 | 1 | 9 | 3 | 66 | 1 | 10 | 1 |
| 17 | 1 | 10 | 1 | 67 | 2 | 7, 10 | 2 |
| 18 | 1 | 7 | 3 | 68 | 1 | 10 | 1 |
| 19 | 1 | 10 | 1 | 69 | 1 | 10 | 1 |
| 20 | 1 | 10 | 1 | 70 | 1 | 10 | 1 |
| 21 | 1 | 10 | 1 | 71 | 1 | 10 | 1 |
| 22 | 1 | 10 | 1 | 72 | 1 | 10 | 1 |
| 23 | 1 | 10 | 1 | 73 | 1 | 10 | 1 |
| 24 | 1 | 10 | 1 | 74 | 1 | 10 | 1 |
| 25 | 1 | 10 | 1 | 75 | 1 | 7 | 3 |
| 26 | 1 | 10 | 1 | 76 | 1 | 10 | 1 |
| 27 | 1 | 10 | 1 | 77 | 1 | 10 | 1 |
| 28 | 1 | 10 | 1 | 78 | 1 | 10 | 1 |
| 29 | 1 | 10 | 1 | 79 | 1 | 10 | 1 |
| 30 | 1 | 10 | 1 | 80 | 1 | 10 | 1 |
| 31 | 1 | 10 | 1 | 81 | 1 | 10 | 1 |
| 32 | 3 | 1, 3, 10 | 2 | 82 | 1 | 10 | 1 |
| 33 | 1 | 10 | 1 | 83 | 1 | 10 | 1 |
| 34 | 3 | 4, 8, 10 | 2 | 84 | 1 | 10 | 1 |
| 35 | 1 | 10 | 1 | 85 | 1 | 10 | 1 |
| 36 | 2 | 6, 10 | 2 | 86 | 1 | 10 | 1 |
| 37 | 1 | 10 | 1 | 87 | 2 | 9, 10 | 2 |
| 38 | 1 | 10 | 1 | 88 | 1 | 10 | 1 |
| 39 | 1 | 10 | 1 | 89 | 1 | 8 | 3 |
| 40 | 1 | 10 | 1 | 90 | 1 | 10 | 1 |
| 41 | 1 | 10 | 1 | 91 | 1 | 10 | 1 |
| 42 | 1 | 10 | 1 | 92 | 1 | 10 | 1 |
| 43 | 1 | 10 | 1 | 93 | 1 | 10 | 1 |
| 44 | 1 | 10 | 1 | 94 | 1 | 10 | 1 |
| 45 | 1 | 10 | 1 | 95 | 1 | 10 | 1 |
| 46 | 1 | 10 | 1 | 96 | 1 | 10 | 1 |
| 47 | 1 | 10 | 1 | 97 | 3 | 1, 4, 10 | 2 |
| 48 | 1 | 10 | 1 | 98 | 1 | 10 | 1 |
| 49 | 1 | 10 | 1 | 99 | 1 | 10 | 1 |
| 50 | 1 | 10 | 1 | 100 | 1 | 10 | 1 |

Subset size sum = 113, $CS_1 = 85$, $FS_1 = 6$, $CS_+ = 9$, $FS_+ = 0$.

After 1000 runs we have for different seed values (when we take the sample sums directly from the random generator):

| $\bar{s}$ | 1.141 | Seed | 1 |
|---|---|---|---|
| $CS_1$ | 807 | $FS_1$ | 66 |
| $CS_+$ | 112 | $FS_+$ | 15 |
| $CS_2$ | 919 | $FS_2$ | 81 |

| $\bar{s}$ | 1.166 | Seed | 2 |
|---|---|---|---|
| $CS_1$ | 801 | $FS_1$ | 56 |
| $CS_+$ | 135 | $FS_+$ | 8 |
| $CS_2$ | 936 | $FS_2$ | 64 |

| $\bar{s}$ | 1.150 | Seed | 3 |
|---|---|---|---|
| $CS_1$ | 799 | $FS_1$ | 66 |
| $CS_+$ | 126 | $FS_+$ | 9 |
| $CS_2$ | 925 | $FS_2$ | 75 |

| $\bar{s}$ | 1.191 | Seed | 4 |
|---|---|---|---|
| $CS_1$ | 792 | $FS_1$ | 55 |
| $CS_+$ | 142 | $FS_+$ | 11 |
| $CS_2$ | 934 | $FS_2$ | 66 |

| $\bar{s}$ | 1.145 | Seed | 5 |
|---|---|---|---|
| $CS_1$ | 810 | $FS_1$ | 65 |
| $CS_+$ | 113 | $FS_+$ | 12 |
| $CS_2$ | 923 | $FS_2$ | 77 |

## E.13 Ordinary design, but $\gamma = 1.25$; $\sigma$ extreme

We take $\mu$ in the LFC ($\mu = (19.5, \ldots, 19.5, 20)$), the standard deviations to be $\sigma = (1.25, \ldots, 1.25, 1)$ for $k = 10$ populations, where $\delta^* = 0.5$, $P^* = 0.75$ and $Q^* = 0.9$. However, in our design we assume all variances to be equal to one, so the sample size should be $n = 28$ and $c = 1.88683$. Because $n$ is rounded, the practical values for $P^*$ and $Q^*$ become $P^* = 0.756833$ and $Q^* = 0.902824$, because we now have $r_1 = 2.28917$ and $r_2 = 3.00233$ (if the variances were really all equal to one).

| Run | SubsetSize | Subset | Status | Run | SubsetSize | Subset | Status |
|---|---|---|---|---|---|---|---|
| 1 | 2 | 2, 10 | 2 | 51 | 1 | 10 | 1 |
| 2 | 1 | 10 | 1 | 52 | 1 | 10 | 1 |
| 3 | 1 | 10 | 1 | 53 | 1 | 10 | 1 |
| 4 | 1 | 10 | 1 | 54 | 1 | 7 | 3 |
| 5 | 1 | 10 | 1 | 55 | 1 | 10 | 1 |
| 6 | 2 | 1, 10 | 2 | 56 | 1 | 10 | 1 |
| 7 | 1 | 3 | 3 | 57 | 1 | 10 | 1 |
| 8 | 1 | 10 | 1 | 58 | 1 | 10 | 1 |
| 9 | 2 | 4, 7 | 4 | 59 | 1 | 10 | 1 |
| 10 | 1 | 10 | 1 | 60 | 1 | 10 | 1 |
| 11 | 1 | 10 | 1 | 61 | 1 | 10 | 1 |
| 12 | 2 | 6, 10 | 2 | 62 | 1 | 2 | 3 |
| 13 | 2 | 2, 10 | 2 | 63 | 1 | 10 | 1 |
| 14 | 3 | 7, 8, 10 | 2 | 64 | 1 | 10 | 1 |
| 15 | 1 | 10 | 1 | 65 | 1 | 10 | 1 |
| 16 | 1 | 9 | 3 | 66 | 2 | 5, 10 | 2 |
| 17 | 1 | 10 | 1 | 67 | 1 | 1 | 3 |
| 18 | 1 | 7 | 3 | 68 | 1 | 10 | 1 |
| 19 | 1 | 10 | 1 | 69 | 1 | 4 | 3 |
| 20 | 1 | 10 | 1 | 70 | 1 | 10 | 1 |
| 21 | 1 | 10 | 1 | 71 | 2 | 8, 10 | 2 |
| 22 | 1 | 10 | 1 | 72 | 1 | 10 | 1 |
| 23 | 2 | 4, 10 | 2 | 73 | 1 | 10 | 1 |
| 24 | 2 | 9, 10 | 2 | 74 | 1 | 10 | 1 |
| 25 | 1 | 10 | 1 | 75 | 1 | 4 | 3 |
| 26 | 2 | 5, 10 | 2 | 76 | 4 | 5, 6, 7, 10 | 2 |
| 27 | 1 | 10 | 1 | 77 | 2 | 1, 10 | 2 |
| 28 | 1 | 10 | 1 | 78 | 1 | 10 | 1 |
| 29 | 1 | 10 | 1 | 79 | 1 | 10 | 1 |
| 30 | 1 | 10 | 1 | 80 | 1 | 10 | 1 |
| 31 | 1 | 1 | 3 | 81 | 1 | 10 | 1 |
| 32 | 1 | 3 | 3 | 82 | 1 | 10 | 1 |
| 33 | 1 | 10 | 1 | 83 | 2 | 6, 10 | 2 |
| 34 | 2 | 4, 6 | 4 | 84 | 1 | 10 | 1 |
| 35 | 1 | 10 | 1 | 85 | 2 | 1, 10 | 2 |
| 36 | 2 | 6, 10 | 2 | 86 | 3 | 2, 5, 10 | 2 |
| 37 | 1 | 10 | 1 | 87 | 1 | 1 | 3 |
| 38 | 1 | 10 | 1 | 88 | 1 | 10 | 1 |
| 39 | 1 | 10 | 1 | 89 | 3 | 6, 8, 10 | 2 |
| 40 | 1 | 10 | 1 | 90 | 1 | 10 | 1 |
| 41 | 1 | 10 | 1 | 91 | 2 | 1, 10 | 2 |
| 42 | 1 | 10 | 1 | 92 | 1 | 10 | 1 |
| 43 | 1 | 10 | 1 | 93 | 1 | 10 | 1 |
| 44 | 1 | 6 | 3 | 94 | 1 | 10 | 1 |
| 45 | 1 | 10 | 1 | 95 | 2 | 3, 10 | 2 |
| 46 | 1 | 10 | 1 | 96 | 1 | 9 | 3 |
| 47 | 2 | 9, 10 | 2 | 97 | 1 | 4 | 3 |
| 48 | 1 | 10 | 1 | 98 | 1 | 10 | 1 |
| 49 | 1 | 10 | 1 | 99 | 1 | 10 | 1 |
| 50 | 1 | 10 | 1 | 100 | 1 | 10 | 1 |

Subset size sum = 127, $CS_1$= 64, $FS_1$=14, $CS_+$= 20, $FS_+$= 2.

After 1000 runs we have for different seed values (when we take the sample sums directly from the random generator):

| $\bar{s}$ | 1.280 | Seed | 1 |
|---|---|---|---|
| $CS_1$ | 651 | $FS_1$ | 134 |
| $CS_+$ | 181 | $FS_+$ | 34 |
| $CS_2$ | 832 | $FS_2$ | 168 |

| $\bar{s}$ | 1.271 | Seed | 2 |
|---|---|---|---|
| $CS_1$ | 637 | $FS_1$ | 137 |
| $CS_+$ | 195 | $FS_+$ | 31 |
| $CS_2$ | 832 | $FS_2$ | 168 |

| $\bar{s}$ | 1.240 | Seed | 3 |
|---|---|---|---|
| $CS_1$ | 659 | $FS_1$ | 146 |
| $CS_+$ | 164 | $FS_+$ | 31 |
| $CS_2$ | 823 | $FS_2$ | 177 |

| $\bar{s}$ | 1.273 | Seed | 4 |
|---|---|---|---|
| $CS_1$ | 634 | $FS_1$ | 146 |
| $CS_+$ | 186 | $FS_+$ | 34 |
| $CS_2$ | 820 | $FS_2$ | 180 |

| $\bar{s}$ | 1.218 | Seed | 5 |
|---|---|---|---|
| $CS_1$ | 669 | $FS_1$ | 138 |
| $CS_+$ | 156 | $FS_+$ | 37 |
| $CS_2$ | 825 | $FS_2$ | 175 |

## E.14 Bayes, $\sigma_i^2$ unknown, for ordinary design with $\gamma = 1.25$

We have for the $k = 10$ populations with $\mu = (19.5, \ldots, 19.5, 20)$ ($\delta^* = 0.5$), and standard deviations $\sigma = (1.25, \ldots, 1.25, 1)$ and $P^* = 0.75$ and $Q^* = 0.9$. For the first 100 runs of the simulation in E.13 we have calculated the minimum set of populations that counts for at least 0.9 of the probability of containing the best one. However, we assume not to know the variances and these are estimated to calculate the Bayesian probabilities.

| Run | Subset | Probs |
|---|---|---|
| 1 | 10, 2, 4, 7, 3, 8 | 0.2366, 0.2256, 0.1852, 0.149, 0.06245, 0.05144 |
| 2 | 10, 6, 5, 4 | 0.4886, 0.2004, 0.1251, 0.103 |
| 3 | 10, 1, 8, 4 | 0.6171, 0.1697, 0.08772, 0.07004 |
| 4 | 10, 6, 5, 8, 7, 9, 2 | 0.3412, 0.1598, 0.1365, 0.08565, 0.07237, 0.07096, 0.05292 |
| 5 | 10, 2, 8, 5, 1 | 0.233, 0.2232, 0.2094, 0.1282, 0.1171 |
| 6 | 1, 10, 7, 3, 6 | 0.2712, 0.2296, 0.2213, 0.1747, 0.04986 |
| 7 | 3, 4, 10, 6 | 0.6495, 0.1421, 0.07723, 0.04513 |
| 8 | 10, 2, 1, 5, 6, 3 | 0.4849, 0.1808, 0.08264, 0.06922, 0.05497, 0.04395 |
| 9 | 7, 4, 10, 8 | 0.4025, 0.3138, 0.1137, 0.09707 |
| 10 | 10, 3, 9, 6 | 0.4698, 0.3247, 0.07918, 0.04914 |
| 11 | 10, 5, 8, 1 | 0.4307, 0.2158, 0.1791, 0.142 |
| 12 | 6, 2, 10, 1, 8, 3 | 0.2679, 0.2004, 0.1654, 0.1245, 0.1114, 0.05084 |
| 13 | 2, 10, 5, 9, 6 | 0.3716, 0.2137, 0.2038, 0.06137, 0.05641 |
| 14 | 8, 7, 10, 1, 3, 5, 2 | 0.2372, 0.1737, 0.1445, 0.1079, 0.09988, 0.09682, 0.04645 |
| 15 | 10, 3, 5, 9, 4 | 0.5926, 0.1389, 0.09776, 0.05916, 0.04611 |
| 16 | 9, 5, 10, 8, 6, 7, 1 | 0.3496, 0.2227, 0.1123, 0.08158, 0.05748, 0.0551, 0.05016 |
| 17 | 10, 3, 6, 9, 8 | 0.6186, 0.1179, 0.09814, 0.0533, 0.03792 |
| 18 | 7, 10, 3, 8, 6 | 0.4397, 0.2054, 0.1389, 0.1094, 0.04638 |
| 19 | 10, 3, 8 | 0.696, 0.1411, 0.09095 |
| 20 | 10, 2, 6, 9, 4 | 0.627, 0.1045, 0.08888, 0.05739, 0.05158 |
| 21 | 10, 9, 6, 3, 5 | 0.3629, 0.245, 0.125, 0.1097, 0.08342 |
| 22 | 10 | 0.975 |
| 23 | 10, 4, 6, 7, 9, 1 | 0.3289, 0.3234, 0.1221, 0.06379, 0.05869, 0.04337 |
| 24 | 9, 10, 2, 7, 1, 5 | 0.3524, 0.3016, 0.1053, 0.06261, 0.0616, 0.05717 |
| 25 | 10, 8 | 0.7039, 0.2389 |
| 26 | 5, 8, 10, 9, 6, 4 | 0.2667, 0.1838, 0.1831, 0.116, 0.1116, 0.06487 |
| 27 | 10, 9, 7, 1, 5 | 0.6557, 0.08365, 0.08094, 0.07701, 0.04579 |
| 28 | 10 | 0.9665 |
| 29 | 10, 5, 6, 8 | 0.5939, 0.1245, 0.09965, 0.08511 |
| 30 | 10, 8, 7, 1, 2 | 0.4315, 0.1402, 0.1327, 0.1072, 0.09092 |
| 31 | 1, 10, 4, 6 | 0.5559, 0.1859, 0.1349, 0.0588 |
| 32 | 3, 1, 6, 10 | 0.5245, 0.2829, 0.09013, 0.0304 |
| 33 | 10, 2, 4, 9, 5, 6 | 0.4289, 0.1882, 0.08733, 0.07873, 0.06553, 0.05237 |
| 34 | 4, 6, 10, 8, 3 | 0.3004, 0.2784, 0.154, 0.1101, 0.07219 |
| 35 | 10, 9, 1, 3, 5 | 0.3778, 0.1503, 0.1377, 0.1325, 0.129 |
| 36 | 6, 10, 9, 1 | 0.4172, 0.2683, 0.1155, 0.1097 |
| 37 | 10, 8, 9 | 0.656, 0.1831, 0.1208 |
| 38 | 10, 2, 4, 8, 3 | 0.4719, 0.2198, 0.07716, 0.06925, 0.06889 |
| 39 | 10 | 0.9412 |
| 40 | 10, 9, 3, 4, 8 | 0.3754, 0.1947, 0.149, 0.1098, 0.07879 |
| 41 | 10, 4 | 0.8421, 0.06352 |
| 42 | 10, 5, 3, 9 | 0.404, 0.2573, 0.191, 0.1139 |
| 43 | 10, 9, 2, 8 | 0.4315, 0.2387, 0.1455, 0.1059 |
| 44 | 6, 8, 10, 5 | 0.398, 0.218, 0.176, 0.1507 |
| 45 | 10, 2, 1 | 0.8406, 0.04826, 0.03402 |
| 46 | 10, 3, 4, 8, 7 | 0.5717, 0.1314, 0.09256, 0.07019, 0.05693 |
| 47 | 9, 10, 3, 7 | 0.3642, 0.3348, 0.1585, 0.07058 |
| 48 | 10, 4 | 0.8712, 0.05971 |
| 49 | 10, 6, 8 | 0.7148, 0.1068, 0.09924 |
| 50 | 10, 1, 6 | 0.8016, 0.06753, 0.04615 |

| Run | Subset | Probs |
|---|---|---|
| 51 | 10, 5, 9, 4 | 0.7261, 0.09269, 0.07765, 0.0504 |
| 52 | 10, 1, 5, 4, 2 | 0.5073, 0.1504, 0.1133, 0.07186, 0.07105 |
| 53 | 10 | 0.9826 |
| 54 | 7, 8, 10, 4 | 0.4095, 0.244, 0.2261, 0.05181 |
| 55 | 10, 9, 4, 3 | 0.4322, 0.2444, 0.1976, 0.07959 |
| 56 | 10, 3, 7 | 0.6951, 0.1349, 0.07386 |
| 57 | 10 | 0.9133 |
| 58 | 10, 3, 5, 8 | 0.6545, 0.1517, 0.05496, 0.04573 |
| 59 | 10, 5, 3, 9 | 0.3849, 0.2456, 0.2154, 0.06513 |
| 60 | 10, 5, 8, 1, 3, 4 | 0.4857, 0.1712, 0.07777, 0.07739, 0.07367, 0.03639 |
| 61 | 10, 6 | 0.8651, 0.05052 |
| 62 | 2, 6, 4, 10, 5 | 0.411, 0.1981, 0.1453, 0.1086, 0.05195 |
| 63 | 10, 1, 4, 5 | 0.6067, 0.1322, 0.1127, 0.08829 |
| 64 | 10, 6, 3, 5, 9, 7, 2 | 0.4789, 0.1205, 0.09242, 0.07513, 0.0724, 0.05794, 0.05201 |
| 65 | 10, 8 | 0.8056, 0.1103 |
| 66 | 10, 5, 7, 1 | 0.3926, 0.3814, 0.09239, 0.04084 |
| 67 | 1, 10, 7, 4 | 0.5542, 0.2059, 0.1319, 0.05654 |
| 68 | 10, 8, 2, 6 | 0.359, 0.2226, 0.22, 0.1046 |
| 69 | 4, 10, 1, 8 | 0.5105, 0.2943, 0.06624, 0.03441 |
| 70 | 10, 8, 3, 1 | 0.433, 0.212, 0.1956, 0.06942 |
| 71 | 10, 1, 8, 5, 6, 4 | 0.2868, 0.2083, 0.1989, 0.1037, 0.09656, 0.05738 |
| 72 | 4, 10, 3, 5, 8, 1 | 0.2899, 0.2294, 0.1426, 0.09827, 0.09144, 0.06117 |
| 73 | 10, 9, 2, 7, 1 | 0.5843, 0.09493, 0.08825, 0.08671, 0.05796 |
| 74 | 10, 3, 4 | 0.7509, 0.1336, 0.05278 |
| 75 | 4, 10, 7 | 0.7526, 0.0971, 0.08029 |
| 76 | 7, 5, 6, 10, 2 | 0.2513, 0.2218, 0.1878, 0.135, 0.1088 |
| 77 | 10, 1, 7, 2 | 0.2927, 0.2803, 0.197, 0.1848 |
| 78 | 10, 4, 2 | 0.8371, 0.04076, 0.03339 |
| 79 | 10, 1, 7 | 0.6385, 0.1941, 0.09737 |
| 80 | 10, 8, 7, 1, 6 | 0.6085, 0.1569, 0.07972, 0.05279, 0.03481 |
| 81 | 10 | 0.9161 |
| 82 | 10, 8 | 0.6674, 0.2373 |
| 83 | 10, 6, 3, 2, 9 | 0.3039, 0.2727, 0.1298, 0.1195, 0.08805 |
| 84 | 10, 2, 5, 7 | 0.7172, 0.104, 0.07587, 0.03628 |
| 85 | 1, 10, 4, 7, 3 | 0.3421, 0.3196, 0.08823, 0.0878, 0.07309 |
| 86 | 2, 3, 5, 10, 1, 6, 9 | 0.1948, 0.1794, 0.1631, 0.159, 0.1047, 0.09384, 0.08028 |
| 87 | 1, 9, 10, 7, 4 | 0.3769, 0.1984, 0.1883, 0.09039, 0.06155 |
| 88 | 10, 3, 7 | 0.7438, 0.1061, 0.09559 |
| 89 | 6, 10, 8, 3, 1, 2, 7 | 0.1923, 0.1909, 0.1836, 0.1352, 0.1089, 0.0763, 0.06072 |
| 90 | 10, 4, 7, 2, 3 | 0.4371, 0.1586, 0.1399, 0.1281, 0.07114 |
| 91 | 1, 10, 7, 3, 9, 2, 5 | 0.4111, 0.1809, 0.09142, 0.07167, 0.0681, 0.04934, 0.04854 |
| 92 | 10, 9, 4, 3, 2, 7, 6 | 0.283, 0.2061, 0.1558, 0.1035, 0.06758, 0.06627, 0.05911 |
| 93 | 10, 1, 8 | 0.5555, 0.3375, 0.03164 |
| 94 | 10, 4, 3 | 0.4984, 0.2466, 0.1595 |
| 95 | 10, 3, 5, 2, 1 | 0.3399, 0.307, 0.1675, 0.08329, 0.04741 |
| 96 | 9, 10, 7, 1 | 0.4693, 0.2531, 0.1029, 0.08623 |
| 97 | 4, 1, 10, 2, 3 | 0.3922, 0.2379, 0.1781, 0.05556, 0.04849 |
| 98 | 10, 1, 8, 3, 6, 9 | 0.5013, 0.1692, 0.1011, 0.07387, 0.04225, 0.04117 |
| 99 | 4, 6, 10, 7, 8 | 0.2818, 0.2279, 0.2185, 0.1428, 0.04579 |
| 100 | 5, 10, 4, 9, 1, 8 | 0.3686, 0.3397, 0.08593, 0.06058, 0.04376, 0.03614 |

## E.15 Robust design with $\gamma = 1.25$; $\sigma$ extreme

We apply the **preference threshold procedure**, $R_{n,c}$, on the instance where we take $\mu$ in the LFC ($\mu = (19.5, \ldots, 19.5, 20)$), the standard deviations $\sigma = (1.25, \ldots, 1.25, 1)$, and $\delta^* = 0.5$, $P^* = 0.75$ and $Q^* = 0.9$. In the robust design with $\gamma = 1.25$, we need to take $n = 38$ and $c = 2.39279$. For this rounded value of $n$ the practical values for $P^*$ and $Q^*$ become $P^* = 0.75206$ and $Q^* = 0.900863$, because we now have $r_{0,1} = 2.69405$ and $r_{0,2} = 3.47037$. The sample sums, $N(38\mu_i, 38\sigma_i^2)$, are taken directly from the random generator.

| Run | SubsetSize | Subset | Status | Run | SubsetSize | Subset | Status |
|---|---|---|---|---|---|---|---|
| 1 | 1 | 10 | 1 | 51 | 1 | 10 | 1 |
| 2 | 1 | 10 | 1 | 52 | 1 | 10 | 1 |
| 3 | 1 | 10 | 1 | 53 | 1 | 10 | 1 |
| 4 | 1 | 10 | 1 | 54 | 1 | 10 | 1 |
| 5 | 1 | 5 | 3 | 55 | 1 | 10 | 1 |
| 6 | 2 | 1, 10 | 2 | 56 | 1 | 10 | 1 |
| 7 | 1 | 10 | 1 | 57 | 1 | 10 | 1 |
| 8 | 1 | 10 | 1 | 58 | 1 | 10 | 1 |
| 9 | 1 | 4 | 3 | 59 | 1 | 10 | 1 |
| 10 | 2 | 3, 10 | 2 | 60 | 1 | 10 | 1 |
| 11 | 1 | 10 | 1 | 61 | 1 | 10 | 1 |
| 12 | 2 | 6, 10 | 2 | 62 | 1 | 2 | 3 |
| 13 | 1 | 10 | 1 | 63 | 1 | 10 | 1 |
| 14 | 1 | 5 | 3 | 64 | 1 | 10 | 1 |
| 15 | 1 | 10 | 1 | 65 | 1 | 10 | 1 |
| 16 | 1 | 9 | 3 | 66 | 1 | 10 | 1 |
| 17 | 1 | 10 | 1 | 67 | 2 | 7, 10 | 2 |
| 18 | 1 | 7 | 3 | 68 | 1 | 10 | 1 |
| 19 | 1 | 10 | 1 | 69 | 2 | 4, 10 | 2 |
| 20 | 1 | 10 | 1 | 70 | 1 | 10 | 1 |
| 21 | 1 | 10 | 1 | 71 | 1 | 10 | 1 |
| 22 | 1 | 10 | 1 | 72 | 1 | 10 | 1 |
| 23 | 1 | 10 | 1 | 73 | 1 | 10 | 1 |
| 24 | 1 | 10 | 1 | 74 | 1 | 10 | 1 |
| 25 | 1 | 10 | 1 | 75 | 1 | 7 | 3 |
| 26 | 1 | 10 | 1 | 76 | 2 | 5, 10 | 2 |
| 27 | 1 | 10 | 1 | 77 | 2 | 1, 10 | 2 |
| 28 | 1 | 10 | 1 | 78 | 1 | 10 | 1 |
| 29 | 1 | 10 | 1 | 79 | 1 | 10 | 1 |
| 30 | 1 | 10 | 1 | 80 | 1 | 10 | 1 |
| 31 | 1 | 1 | 3 | 81 | 1 | 10 | 1 |
| 32 | 1 | 1 | 3 | 82 | 1 | 10 | 1 |
| 33 | 1 | 10 | 1 | 83 | 1 | 10 | 1 |
| 34 | 1 | 4 | 3 | 84 | 1 | 10 | 1 |
| 35 | 1 | 10 | 1 | 85 | 1 | 10 | 1 |
| 36 | 2 | 6, 10 | 2 | 86 | 1 | 10 | 1 |
| 37 | 1 | 10 | 1 | 87 | 2 | 9, 10 | 2 |
| 38 | 1 | 10 | 1 | 88 | 1 | 10 | 1 |
| 39 | 1 | 10 | 1 | 89 | 1 | 8 | 3 |
| 40 | 1 | 10 | 1 | 90 | 1 | 10 | 1 |
| 41 | 1 | 10 | 1 | 91 | 1 | 10 | 1 |
| 42 | 1 | 10 | 1 | 92 | 1 | 10 | 1 |
| 43 | 1 | 10 | 1 | 93 | 1 | 10 | 1 |
| 44 | 1 | 10 | 1 | 94 | 1 | 10 | 1 |
| 45 | 1 | 10 | 1 | 95 | 1 | 10 | 1 |
| 46 | 1 | 10 | 1 | 96 | 1 | 10 | 1 |
| 47 | 1 | 10 | 1 | 97 | 1 | 1 | 3 |
| 48 | 1 | 10 | 1 | 98 | 2 | 1, 10 | 2 |
| 49 | 1 | 10 | 1 | 99 | 1 | 10 | 1 |
| 50 | 1 | 10 | 1 | 100 | 1 | 10 | 1 |

Subset size sum = 110, $CS_1$= 78, $FS_1$ =12, $CS_+$ = 10, $FS_2$ = 0.

After 1000 runs we have for different seed values (when we take the sample sums directly from the random generator):

| $\bar{s}$ | 1.178 | Seed | 1 |
|---|---|---|---|
| $CS_1$ | 768 | $FS_1$ | 79 |
| $CS_+$ | 131 | $FS_+$ | 22 |
| $CS_2$ | 899 | $FS_2$ | 101 |

| $\bar{s}$ | 1.204 | Seed | 2 |
|---|---|---|---|
| $CS_1$ | 751 | $FS_1$ | 75 |
| $CS_+$ | 156 | $FS_+$ | 18 |
| $CS_2$ | 907 | $FS_2$ | 93 |

| $\bar{s}$ | 1.186 | Seed | 3 |
|---|---|---|---|
| $CS_1$ | 754 | $FS_1$ | 81 |
| $CS_+$ | 147 | $FS_+$ | 18 |
| $CS_2$ | 901 | $FS_2$ | 99 |

| $\bar{s}$ | 1.217 | Seed | 4 |
|---|---|---|---|
| $CS_1$ | 743 | $FS_1$ | 77 |
| $CS_+$ | 165 | $FS_+$ | 15 |
| $CS_2$ | 902 | $FS_2$ | 92 |

| $\bar{s}$ | 1.178 | Seed | 5 |
|---|---|---|---|
| $CS_1$ | 762 | $FS_1$ | 83 |
| $CS_+$ | 131 | $FS_+$ | 24 |
| $CS_2$ | 903 | $FS_2$ | 107 |

## E.16 Bayes, $\sigma_i^2$ unknown, for robust design with $\gamma = 1.25$

We have for the $k = 10$ populations with $\mu = (19.5, \ldots, 19.5, 20)$ $(\delta^* = 0)$ and standard deviations $\sigma = (1.25, \ldots, 1.25, 1)$ $P^* = 0.75$ and $Q^* = 0.9$. For the first 100 runs of the simulation in E.15 we have calculated the minimum set of populations that counts for at least 0.9 of the probability of containing the best one. However, we assume not to know the variances and these are estimated to calculate the Bayesian probabilities.

| Run | Subset | Probs |
|---|---|---|
| 1 | 10, 2, 4, 8, 7 | 0.3126, 0.2706, 0.2047, 0.06094, 0.05195 |
| 2 | 10, 4, 6 | 0.7241, 0.1285, 0.0879 |
| 3 | 10, 8, 4, 1, 3 | 0.652, 0.1253, 0.0597, 0.05213, 0.05075 |
| 4 | 10, 8, 5, 2, 6, 9, 1 | 0.3431, 0.2057, 0.09281, 0.08721, 0.0712, 0.0593, 0.05035 |
| 5 | 5, 2, 10, 8, 1 | 0.3294, 0.2371, 0.1568, 0.1424, 0.08631 |
| 6 | 10, 1, 7, 3, 6 | 0.2934, 0.2085, 0.1999, 0.18, 0.06989 |
| 7 | 10, 3, 2, 4, 6 | 0.3814, 0.255, 0.112, 0.111, 0.08136 |
| 8 | 10, 2, 3, 1, 6 | 0.5955, 0.1137, 0.08185, 0.06689, 0.05502 |
| 9 | 4, 10, 7, 8 | 0.6617, 0.1269, 0.1028, 0.06123 |
| 10 | 3, 10 | 0.5027, 0.3983 |
| 11 | 10, 5, 1 | 0.6665, 0.1941, 0.1095 |
| 12 | 6, 10, 4, 8, 3, 1 | 0.3729, 0.3051, 0.1223, 0.04483, 0.03471, 0.03334 |
| 13 | 10, 2, 6, 5, 1 | 0.5462, 0.1461, 0.1124, 0.07023, 0.04093 |
| 14 | 5, 8, 7, 4, 1 | 0.3646, 0.2132, 0.1598, 0.1179, 0.04738 |
| 15 | 10, 9, 8, 4 | 0.7327, 0.08427, 0.05688, 0.03612 |
| 16 | 9, 10, 5, 8, 7 | 0.4576, 0.1711, 0.1292, 0.08463, 0.07726 |
| 17 | 10, 6, 3 | 0.8253, 0.04459, 0.04345 |
| 18 | 7, 10, 3 | 0.5443, 0.3337, 0.05664 |
| 19 | 10, 3 | 0.8593, 0.06792 |
| 20 | 10, 1, 9, 2 | 0.6596, 0.1016, 0.09408, 0.07086 |
| 21 | 10, 9, 6, 3, 7 | 0.4503, 0.2716, 0.08008, 0.06639, 0.05131 |
| 22 | 10 | 0.9888 |
| 23 | 10, 4, 6 | 0.5354, 0.2541, 0.1122 |
| 24 | 10, 9, 8, 1, 7, 2 | 0.4798, 0.1692, 0.09672, 0.07628, 0.05759, 0.04016 |
| 25 | 10, 8 | 0.7353, 0.222 |
| 26 | 10, 5, 4, 8, 6, 2 | 0.2358, 0.169, 0.1583, 0.1463, 0.1378, 0.05691 |
| 27 | 10, 5, 7 | 0.8159, 0.0831, 0.03272 |
| 28 | 10 | 0.9742 |
| 29 | 10, 3, 5, 9 | 0.7718, 0.0608, 0.05743, 0.04913 |
| 30 | 10, 8, 2, 7, 1 | 0.6307, 0.1019, 0.09279, 0.05138, 0.04367 |
| 31 | 1, 10, 9, 4 | 0.5454, 0.2099, 0.1132, 0.07496 |
| 32 | 1, 3, 10 | 0.4969, 0.3782, 0.04767 |
| 33 | 10, 2, 4, 5 | 0.4731, 0.2567, 0.1066, 0.08286 |
| 34 | 4, 10, 3, 8, 6 | 0.3331, 0.199, 0.1727, 0.1269, 0.1053 |
| 35 | 10, 3, 9, 5, 1 | 0.521, 0.1192, 0.1147, 0.08789, 0.08422 |
| 36 | 6, 10, 9, 5, 1 | 0.4261, 0.2505, 0.1129, 0.09096, 0.04246 |
| 37 | 10, 8 | 0.7133, 0.22 |
| 38 | 10, 2, 8, 4, 9, 5 | 0.5151, 0.1157, 0.09439, 0.08555, 0.07626, 0.07617 |
| 39 | 10 | 0.9303 |
| 40 | 10, 9, 5, 3 | 0.4814, 0.2585, 0.1025, 0.07456 |
| 41 | 10, 3 | 0.8835, 0.0379 |
| 42 | 10, 5, 9 | 0.5435, 0.2179, 0.1415 |
| 43 | 10, 9, 1, 2 | 0.4932, 0.2926, 0.06862, 0.05434 |
| 44 | 8, 10, 5, 6, 4, 9 | 0.3034, 0.2925, 0.1348, 0.09773, 0.06864, 0.06104 |
| 45 | 10 | 0.9294 |
| 46 | 10, 8, 3, 5 | 0.5954, 0.2158, 0.08148, 0.04756 |
| 47 | 10, 9, 3, 5 | 0.539, 0.2735, 0.06187, 0.04249 |
| 48 | 10, 2 | 0.8635, 0.0512 |
| 49 | 10, 6, 8 | 0.7781, 0.09784, 0.04581 |
| 50 | 10, 6, 1, 2 | 0.5417, 0.2364, 0.1045, 0.05487 |

| Run | Subset | Probs |
|---|---|---|
| 51 | 10, 3, 5 | 0.8078, 0.09141, 0.05221 |
| 52 | 10, 1, 5, 2, 4 | 0.6706, 0.09481, 0.08276, 0.05177, 0.03439 |
| 53 | 10 | 0.9833 |
| 54 | 10, 8, 7, 4 | 0.4588, 0.225, 0.2112, 0.06167 |
| 55 | 10, 4, 9 | 0.6117, 0.205, 0.08717 |
| 56 | 10, 3, 6 | 0.6802, 0.2095, 0.05807 |
| 57 | 10 | 0.9744 |
| 58 | 10, 8, 3 | 0.8267, 0.05451, 0.05246 |
| 59 | 10, 5, 3, 9 | 0.5608, 0.2644, 0.0743, 0.05199 |
| 60 | 10, 5, 3, 4, 9 | 0.6911, 0.08323, 0.07466, 0.03801, 0.02992 |
| 61 | 10, 8 | 0.8914, 0.04645 |
| 62 | 2, 6, 10, 4, 8 | 0.3972, 0.1853, 0.139, 0.1384, 0.05156 |
| 63 | 10, 4, 1, 5 | 0.594, 0.2047, 0.08187, 0.06012 |
| 64 | 10, 2, 5, 7, 3 | 0.6555, 0.1138, 0.05018, 0.04871, 0.04678 |
| 65 | 10, 8 | 0.8328, 0.1042 |
| 66 | 10, 5, 7, 1, 3 | 0.5117, 0.2319, 0.08841, 0.04655, 0.0432 |
| 67 | 7, 1, 10, 6, 4 | 0.312, 0.2616, 0.2151, 0.05968, 0.05464 |
| 68 | 10, 2, 8 | 0.8096, 0.07367, 0.0402 |
| 69 | 4, 10, 8, 3 | 0.3954, 0.3616, 0.09053, 0.05662 |
| 70 | 10, 3, 1, 8, 9 | 0.4525, 0.238, 0.1129, 0.05838, 0.0422 |
| 71 | 10, 1, 6, 8, 4, 5 | 0.3923, 0.173, 0.1419, 0.1295, 0.06311, 0.05527 |
| 72 | 10, 4, 1, 2, 6, 8, 5 | 0.3757, 0.1337, 0.1268, 0.1029, 0.09097, 0.05561, 0.04866 |
| 73 | 10, 9, 1, 4 | 0.7735, 0.06436, 0.05455, 0.03825 |
| 74 | 10, 4, 3 | 0.7419, 0.09837, 0.07232 |
| 75 | 7, 4, 2 | 0.6374, 0.2357, 0.05549 |
| 76 | 10, 5, 7, 8, 2 | 0.3694, 0.2641, 0.1469, 0.09427, 0.07493 |
| 77 | 10, 1, 7, 2 | 0.4584, 0.3065, 0.1134, 0.1099 |
| 78 | 10, 5, 4, 3, 9 | 0.7017, 0.08387, 0.05912, 0.03831, 0.03059 |
| 79 | 10, 7, 1 | 0.5818, 0.3086, 0.08827 |
| 80 | 10, 7, 8 | 0.8466, 0.05031, 0.04704 |
| 81 | 10, 5 | 0.7825, 0.1195 |
| 82 | 10, 8 | 0.6999, 0.2261 |
| 83 | 10, 3, 9, 2 | 0.7354, 0.07391, 0.05421, 0.04557 |
| 84 | 10, 2 | 0.8554, 0.06778 |
| 85 | 10, 1, 3, 4, 7 | 0.457, 0.208, 0.09744, 0.07501, 0.06485 |
| 86 | 10, 3, 5, 2, 6, 1, 9 | 0.2412, 0.1899, 0.1297, 0.1229, 0.118, 0.0886, 0.0559 |
| 87 | 9, 10, 1, 7, 3, 4 | 0.3256, 0.2659, 0.1566, 0.09693, 0.05216, 0.04863 |
| 88 | 10, 7, 3 | 0.7274, 0.1363, 0.05057 |
| 89 | 8, 3, 6, 1, 10 | 0.2727, 0.1953, 0.1744, 0.1661, 0.1396 |
| 90 | 10, 2, 4, 7 | 0.4461, 0.3129, 0.111, 0.06974 |
| 91 | 1, 10, 7, 3, 9, 8, 4, 2 | 0.2077, 0.2064, 0.2007, 0.08508, 0.06373, 0.0535, 0.05161, 0.05038 |
| 92 | 10, 9, 6, 4, 3, 2 | 0.2942, 0.2633, 0.1375, 0.1127, 0.06828, 0.04339 |
| 93 | 10, 1 | 0.8, 0.1728 |
| 94 | 10, 4, 3 | 0.802, 0.08629, 0.05503 |
| 95 | 10, 5, 3, 2 | 0.5734, 0.2012, 0.09216, 0.04645 |
| 96 | 10, 9, 4, 7, 1 | 0.3812, 0.2918, 0.1126, 0.09201, 0.08137 |
| 97 | 1, 4, 10, 2 | 0.4386, 0.2142, 0.1933, 0.08027 |
| 98 | 1, 10, 8, 4, 6, 2 | 0.3705, 0.3484, 0.06885, 0.05766, 0.04938, 0.03354 |
| 99 | 10, 4, 6, 3, 9 | 0.3584, 0.2137, 0.1918, 0.1098, 0.04578 |
| 100 | 10, 5, 9, 8, 3 | 0.5182, 0.2026, 0.0714, 0.06765, 0.05129 |