Studies on Some Inferential Aspects of Graybill-Deal Estimators

by

Keyu Nie
B.A. (University of Science and Technology of China) 2007
M.S. (University of Illinois at Chicago) 2011

Thesis submitted in partial fulfillment of the requirements
for the degree of Doctor of Philosophy in Mathematics
in the Graduate College of the
University of Illinois at Chicago, 2016

Chicago, Illinois

Defense Committee:
Samad Hedayat, Chair and Advisor
Cheng Ouyang
Jie Yang
Min Yang
Donald Hedeker, University of Chicago

To

Myself

Family Members

Friends

At

Chicago

the Earth.

ACKNOWLEDGMENTS

I want to express my greatest and sincerest appreciation to my advisor Professor Samad Hedayat, whose active help and tremendous support have not only made this dissertation possible, but also profoundly influenced my research, study, and my whole life. It is my honor to have my thesis being the 40^{th} Ph.D. dissertation directed by him.

I also would like to express my deepest gratitude to Professor B.K. Sinha, to whom I owe millions of thanks for his general supervision and numerous instructions. I am glad to be one of his 85+ collaborators, and the 6^{th} student who had worked for doctoral programs under Professor Samad Hedayat and his guidance.

I would like to thank Professor Jie Yang who encouraged me to apply to the Ph.D. program here. I would like to also thank Professor Jing Wang, Professor Cheng Ouyang, Professor Ryan Martin, Professor Junhui Wang, Professor Min Yang, Professor Lawrence Lin, and Professor T.E.S. Raghavan for teaching me statistical courses and providing helpful assistants on my research activities.

I would like to thank all my committee members, Professor Samad Hedayat, Professor Donald Hedeker, Professor Cheng Ouyang, Professor Jie Yang, and Professor Min Yang.

Finally, I owe a special thanks to all my family members for their love and support throughout my education.

This material is based upon work supported by the U.S. National Science Foundation under Grant DMS-1306394 of professor Samad Hedayat. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the authors) and do not necessarily reflect the views of the National Science Foundation.

PREFACE

Please note that all results in this dissertation were achieved by collaboratively working with professor B.K. Sinha and professor Samad Hedayat.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS

BLUE	Best Linear Unbiased Estimator
BIBD	Balanced Incomplete Block Design
GDE	Graybill-Deal Estimator
I.I.D	Identical Independent Distribution
K-dim	K-Dimensional
N.N.D	None-negative Definite
OLS	Ordinary Least Squares
P.D	Positive Definite
UMVUE	Uniformly Minimum-Variance Unbiased Estimator
W.P	With Probability

SUMMARY

This dissertation included two fields in the application of Statistical Inference, specialized in Meta Analysis and Estimation Theory viz. : 'Estimation of Common Mean' and 'function Estimation'.

Part A Estimation of Common Mean.

In Chapter 2, we extended Graybill-Deal Estimator (GDE) to the higher dimension: common parameter estimation in linear regression models. We found the same result continues to hold in situations wherein the p ($p \geq 2$) linear regression models involve k ($k > 1$) common estimable parameter(s) in the mean models. In this context, we used the criterion of 'Loewner Order Domination' of information or dispersion matrices.

Then Chapter 3, we studied GDE's properties under Pitman closeness criterion. Specifically, we compared a p-source based Graybill-Deal estimator against its q-sub-source based competitors for q ($< p$)-dimensional subsets of p-dimensional data.

Part B Function Estimation.

In Chapter 4, we presented a negative report about the estimation of reliability function by using a single observation from a mixture of two exponential distributions. We showed that there exists proper estimator on if we require negative weight on the distributions.

All the references cited in this thesis would be presented at the end.

CHAPTER 1

INTRODUCTION

1.1 Estimation of Common Mean of Normal Populations

The common mean estimation problem was first introduced by Cochran (14), while he was considering combining a series of similar experiments. The general setting for this kind of problem was: we had p independent groups of normal variables with sample size n_i, for the i^{th} group, having the sample mean $\bar{x}_i \sim N(\mu, \frac{\sigma_i^2}{n_i})$, where $i = 1, 2, \ldots, p$. The setup presupposed that there was a common unknown mean μ for the p populations and unknown but possible likely unequal variances. The problem considered was that of efficient unbiased estimation of μ based on the data from the p groups.

Yates (49; 50) initiated that this was related to balanced incomplete block design (BIBD) with fixed treatment effects and uncorrelated random block effects. In this set up, one had to combine inter-block and intra-block information to better estimate the treatment effects. Under this usual assumpation of indepence and normality, these two estimators are independent, following the same setting of common mean of normal populations. Nair (32) and Rao (38; 39) extended Yates' (49; 50) work, and hence attracted a lot of attention. Specially Bhattacharya (3; 4; 2; 5; 6) considered this within design problems in many details.

For $p = 2$, Cochran (14) suggested the unbiased estimator

$$\hat{\mu}_C = \frac{\left(\bar{x}_1 \frac{n_1}{\sigma_1^2} + \bar{x}_2 \frac{n_2}{\sigma_2^2}\right)}{\frac{n_1}{\sigma_1^2} + \frac{n_2}{\sigma_2^2}}. \tag{1.1}$$

This estimator is the best linear unbiased estimator (BLUE) for μ, assuming that the two variances are known [in fact, this is also the uniformly minmum variance unbiased estimator

(UMVUE)]. Motivated by Cochran's (14) work, Graybill and Deal (19) introduced their estimator $\hat{\mu}_{GD|2}$, known as Graybill-Deal Estimator (GDE), by replacing the true variances with their corresponding unbiased estimators:

$$\hat{\mu}_{GD|2} = \frac{\left(\bar{x}_1 \frac{n_1}{s_1^2} + \bar{x}_2 \frac{n_2}{s_2^2}\right)}{\frac{n_1}{s_1^2} + \frac{n_2}{s_2^2}}, \tag{1.2}$$

where $s_i^2 = \frac{1}{n_i - 1} \sum_{j=1}^{n_i} (x_{ij} - \bar{x}_i)^2$ for $i = 1, 2$. In view of distributional independence of $\bar{x}_i$ and s_i^2, $\hat{\mu}_{GD|2}$ is an unbiased estimator for μ. Furthermore, Graybill and Deal established that $\hat{\mu}_{GD|2}$ is uniformly superior to any single unbiased estimator of μ if and only if the following condition holds:

Either both n_1 and $n_2 > 10$ or $n_1 = 10$ ($n_2 = 10$) and $n_2 > 18$ (respectively $n_1 > 18$) .

Norwood and Hinkelmann (33) extended Graybill and Deal's (19) results to general p groups, and they established that $\hat{\mu}_{GD|p}$ is a uniformly better estimator of μ than each $\bar{x}_i$ if and only if either $n_i > 10$ for $i = 1, 2, \ldots, p$ or $n_i = 10$ for some i, and $n_j > 18$ ($i, j = 1, 2, \ldots, p$ for each $j \neq i$), where

$$\hat{\mu}_{GD|p} = \sum_{i=1}^{p} \frac{\bar{x}_i \frac{n_i}{s_i^2}}{\sum_{j=1}^{p} \frac{n_j}{s_j^2}}. \tag{1.3}$$

Shinozaki (44) extended $\mu_{GD|p}$ to a general form:

$$\hat{\mu}_S = \sum_{i=1}^{p} \frac{\bar{x}_i \frac{c_i}{s_i^{2*}}}{\sum_{j=1}^{p} \frac{c_j}{s_j^{2*}}}, \tag{1.4}$$

where $s_i^{2*} = \frac{s_i^2}{n_i}$. By a careful choice of $(c_1, c_2, \ldots, c_p)$, Shinozaki (44) presented a proof of the claim that $\hat{\mu}_{S,p}$ is an uniformly better estimator of μ than any $\hat{\mu}_{S,q}$ of combining q $(< p)$ components, if and only if $\frac{c_j}{c_i} \leq 2\frac{(n_i-1)(n_j-5)}{(n_i+1)(n_j-1)}$ for any $1 \leq i \neq j \leq p$ and for all $1 \leq q < p$. It is readily verified that when our choice of $(c_1, c_2, \ldots, c_p)$ corresponds to $(c_1 = c_2 = \cdots = c_p)$, the condition above simplifies to what is stated earlier involving the sample sizes only.

The research of such estimators (the GDE or similar types) have been widely studied in the literature. Pal and Sinha (1996) gave a very comprehensive and detailed review on GDE. They mentioned that generally there were two parts in GDE research:

1. comparison the GDE with other estimators;
2. properties of the GDE.

In particular we would like to emphasize the work of Zacks (51; 52), Khatri and Shah (23), Brown and Cohen (8), and Cohen and Sackrowitz (16). They all tried to improve the GDE by their own estimators, but most of them had little practical usage. On the other side, Meier (1953) first established an approximate estimatior of $Var(\hat{\mu}_{GD|p})$, Sinha (1985) also provided another first order approximate of $Var(\hat{\mu}_{GD|p})$ which is comparable to Meier (31). Although the admissibility of the GDE under usual square error loss function is still up in the air, Zacks (51; 52), Kubokawa (26), Sinha and Mouqadem (47) and Sinha (46) studied this area. Particularly, Sinha and Mouqadem (47) considered a special case with $p = 2$ and sample size $n_1 = n_2 = n$, and they defined four classes as follows:

$$C = \{\hat{\mu} : \hat{\mu} = \bar{x}_1 + D\phi_1, 0 \leq \phi_1(s_1^2, s_2^2, D^2) \leq 1\};$$

$$C_0 = \{\hat{\mu} : \hat{\mu} = \bar{x}_1 + D\phi_2, 0 \le \phi_2(s_1^2/s_2^2) \le 1\};$$

$$C_1 = \{\hat{\mu} : \hat{\mu} = \bar{x}_1 + D\phi_3, 0 \le \phi_3(s_1^2, s_2^2) \le 1\};$$

$$C_2 = \{\hat{\mu} : \hat{\mu} = \bar{x}_1 + D\phi_4, 0 \le \phi_4(s_1^2/D^2, s_2^2/D^2) \le 1\};$$

where $D = \bar{x}_2 - \bar{x}_1$. It is obvious that $C \supset C_1 \supset C_0$ and $C \supset C_2 \supset C_0$. C_1; and C are equivariant under location transformation while C_0 and C_2 are equivariant under affine transformation (such as $(\bar{x}_1, \bar{x}_2, s_1^2, s_2^2) \to (a\bar{x}_1 + b, a\bar{x}_2 + b, a^2 s_1^2, a^2 s_2^2), a > 0, b \in R$). Sinha and Mouqadem (47) showed that the GDE is admissible in C_0 and C_2, and it is extended admissbile in C. They also provided an estimator which is admissible in C_1. Sinha (46) found that if there is prior knowledge of the variances (the order of σ_i^2's), the GDE could be improved.

In all of these studies, the populations were assumed to be normally distributed with identical means but heterogeneous variability. Because in practice, populations came into examination were usually from the same or homogeneous sources, but different populations were exposed to different environmental conditions. For example, a soil laboratory theorized that samples collecting from several different locations had the same mean since these soils were formed from the same geological phenomenon but have different variances due to the fact that they were exposed to different meteorological and microbiological environments. For another example, a set of chemical products were divided into several pools and sent to different laboratories for analysis of the ingredients. The means should be the same as these chemical products were produced in the same batch. But the variances might differ from one to the other due the method, instrument and/or human bias.

So far in the literature related to GDE, the standard set-up of estimation of normal common mean has been investigated. Chiou and Cohen (12), Loh (29), Kubokawa (24), Tsukuma and Konno (48) investigated the multivariate normal counterpart in this problem. Chiou and Cohen (12) reported some negative results of GDE in higher dimension. In Chapter 2, we were primarily interested in a linear regression set-up with common parameter. We followed the general approach of formation of Graybill-Deal-type estimators in such set-ups and then examined conditions for their superiority over corresponding estimators based on partial exposure to the entire body of data. There were exceptions such as in Kubokawa (24). However, linear regression set-up has not been studied yet. In a view of practical application, often when different researchers analyzed different samples with the same linear model, the estimators generated from each sample portion may include more than one parameter (intercept and/or slopes). We would then face a challenge of combining these estimators efficiently.

In Chapter 3, we compared GDE under Pitman closeness criterion, which was introduced by Pitman (36): We say estimator $\hat{\mu}_1$ is better (closer) than $\hat{\mu}_2$ for the estimation of the parameter μ if and only if $P\{|\hat{\mu}_1 - \mu| \leq |\hat{\mu}_2 - \mu|\} \geq 1/2$. Kubokawa (25) and Sarkar (43) established that for $p = 2$, necessary and sufficient condition for

$$P\{|\hat{\mu}_{GD|2} - \mu| \leq |\bar{x}_i - \mu|\} \geq 1/2, \tag{1.5}$$

to hold uniformly in $(\mu, \sigma_1^2, \sigma_2^2)$ is that $m_i = n_i - 1 \geq 4$ for each $i = 1, 2$. Sarkar (43) also revealed that

$$P\{|\hat{\mu}_{GD|p} - \mu| \leq |\bar{x}_i - \mu|\} \geq 1/2, \tag{1.6}$$

for all $i = 1, 2, \ldots, p$ and uniformly in $(\mu, \sigma^2_{\bar{x}_1}, \sigma^2_{\bar{x}_2}, \ldots, \sigma^2_{\bar{x}_p})$, if and only if

$$2E\left\{(\sum_{j=1, j\neq i}^{p} \sigma_{\bar{x}_j}^{-2} Y_j^2)^{-1/2} (\sum_{j=1, j\neq i}^{p} \sigma_{\bar{x}_j}^{-2} Y_j)\right\} \leq E\left\{(\sum_{j=1, j\neq i}^{p} \sigma_{\bar{x}_j}^{-2} Y_j^2)^{1/2}\right\}, \tag{1.7}$$

holds for all $i = 1, 2, \ldots, p$ and uniformly in $(\sigma^2_{\bar{x}_1}, \sigma^2_{\bar{x}_2}, \ldots, \sigma^2_{\bar{x}_p})$, where Y_j's are independently distributed as $\frac{m_j}{\chi^2(m_j)}$ and $m_j = n_j - 1$, for $j = 1, 2, \ldots, p$. Sarkar (43) showed that Equation 1.7 holds for any i in $\{1, 2, \ldots, p\}$, if

$$1 - 8m_k^{-1} + 4(\sum_{j=1, j\neq i}^{p} m_j)^{-1} \geq 0, \tag{1.8}$$

for all k in $\{i = 1, 2, \ldots, i-1, i+1, \ldots, p\}$.

Base on the results of Kubokawa (25) and Sarkar (43), we examined $\hat{\mu}_{GD|p}$ with $\hat{\mu}_{GD|q}$, the GDE of any q $(< p)$ subgroups, in the sense of Pitman closeness criterion. Several sufficient or necessary conditions were provided.

1.2 Unbiased Estimation of Reliability Function From Mixture of Two Exponential Distributions

In the study of life testing and reliability analysis, an important approach is to find suitable estimate of an underlying 'life' distribution. For a practical reason, it is relevant to get an unbiased estimate of reliability function (survival function) $R(t) = e^{-t/\lambda}$ when the real 'life' distribution follows exponenital distribution as $\frac{1}{\lambda}e^{-t/\lambda}$. Pugh (37) and Basu (1) found the UMVUE of reliability function, $\hat{R}_{umvue}(t)$, under the following assumptions:

Let X_1, ...,X_n be n identical independent observations on X, which is following an exponential distribution with unknown mean λ (> 0).

A simple unbiased estimator of $R(t)$, called $\hat{R}_1(t)$, is:

$$\hat{R}_1(t) = \frac{1}{n}\sum_{i=1}^{n} I(X_i > t), \tag{1.9}$$

where $I(.)$ is the indicator function.

The variance of $\hat{R}_1(t)$ is :

$$Var(\hat{R}_1(t)) = \frac{R(t)(1 - R(t))}{n}. \tag{1.10}$$

The UMVUE given by Pugh (37) and Basu (1) is that:

$$\hat{R}_{umvue}(t) = \begin{cases} (1 - \frac{t}{W})^{n-1} & W > t \\ 0 & otherwise \end{cases} \tag{1.11}$$

where $W = \sum_{i=1}^{n} X_i$. The variance of $\hat{R}_{umvue}(t)$ is:

$$Var(\hat{R}_{umvue}(t)) = R(t)(\phi(t) - R(t)), \tag{1.12}$$

where $\phi(t) = \frac{1}{\Gamma(n)} \int_0^\infty e^{-\mu} \frac{\mu^{2n-2}}{(\mu+\frac{t}{\lambda})^{n-1}} d\mu$.

Due to the well known facts that the difference of two adjacent order statistics of exponential distributions also follows a exponential distribution, Basu (1) additionally provided the UMVUE of $R(t)$ from first r $(r < n)$order statistics:

$$\hat{R}_{os}(t) = \begin{cases} (1 - \frac{t}{W^*})^{r-1} & W^* > t \\ 0 & otherwise \end{cases} \tag{1.13}$$

where $W^* = \sum_{i=1}^{r} X_{(i)} + (n-r)X_{(r)}$, and $X_{(i)}$ is the i^{th} order statistics, $i = 1, 2, ..., n$.

Sinha, Sengupta and Mukhuti (45) extended to more general cases for order statistics situation. They mentioned that "in many practical issues, instead of a complete random sample of size n, only r $(1 \leq r < n)$ selected order statistics from it, such as $X_{(i_1)}$, $X_{(i_2)}$, ..., $X_{(i_r)}$, $1 \leq i_1 < i_2 < ... < i_r \leq n$." The key idea of solving it was how to unbiased estimate $R(t)$ when there was only a single order statistics. Sinha, Sengupta and Mukhuti (45) proved that the unique unbiased estimator of $R(t)$ based on a single i^{th} order statistic X_i $(1 \leq i \leq n)$, denoted by $h_i(Z_{(i)})$ and $Z_{(i)} = (n-i+1)X_{(i)}$, is given by:

$$h_i(Z_{(i)}) = \sum_{y_1=0}^{\infty} \sum_{y_2=0}^{\infty} \cdots \sum_{y_{i-1}=0}^{\infty} d_{y_1 y_2 \ldots y_{i-1}} I(Z_i > \alpha_1^{y_1} \alpha_2^{y_2} \ldots \alpha_{i-1}^{y_{i-1}} t), \tag{1.14}$$

where $\alpha_j = \frac{n-i+j}{n-i}$, $j = 1, 2, \ldots, i$, $d_{y_1 y_2 \ldots y_{i-1}} = \frac{(-1)^{\sum_j y_j}}{\binom{n}{i}} \frac{(y_1+y_2+\cdots+y_{i-1})!}{y_1! y_2! \ldots y_{i-1}!} \frac{\binom{i-1}{1}^{y_1}}{\alpha^{y_1}} \frac{\binom{i-1}{2}^{y_2}}{\alpha^{y_2}} \cdots \frac{\binom{i-1}{i-1}^{y_{i-1}}}{\alpha^{y_{i-1}}}$, $\sum_j$ means sum of all even suffixes of y and all y_j are integers, $j = 1, 2, \ldots, i-1$. For example, when $i = 2$, we reduced our expression as:

$$h_2(Z_{(2)}) = \sum_{y_1=0}^{\infty} \frac{1}{n\alpha^{y_1}} I(Z_{(2)} > \alpha^{y_1} t), \tag{1.15}$$

where $\alpha = \frac{n}{n-1}$.

Sinha, Sengupta and Mukhuti (45) also revealed that this unbiased estimator based on a point of single order statistics could achieve a smaller variance than UMVUE from a complete random size of n by applying ranked set sampling method, which was due to McIntyre (30). According to his procedure, one selected n independent random sets with size n for each set, then observed only the i^{th} order statistics (i^{th} smallest) at set i. Chiuv and Sinha (13) and Chen Bai and Sinha (11) gave a good detailed review on this procedure.

In Chapter 4, inspired by Sinha, Sengupta and Mukhuti (45), we tried to apply similar method to find the unbiased estimator of reliability function based on a single observation from a mixture exponential distributions. We reported some negative results: the unbiased estimator of $R(t)$ is a proper estimator (between 0 and 1) if only if the mixture probability is negative. Jevremovic (22) provided some example on mixed exponential distributions with negative weights.

CHAPTER 2

ON A GENERAL PATTERN OF DOMINATION USING THE GRAYBILL–DEAL ESTIMATOR

[1] In environmental pollution studies, in order to understand the environmental factors affecting the mean 'contamination/pollution level' of air/water/land, representative samples are sent to different laboratories for statistical analysis. This corresponds to what is technically addressed as 'Meta Analysis' problem. All studies in different laboratories have a common goal viz., estimation and assessment of global contamination level in the experimental region. At times, linear or quadratic or higher degree regression models are adequate with/without common intercept term and/or common slope. Of course, the laboratories are likely to have instruments with different precision levels. In such situations, we call for natural application of Graybill-Deal Type estimators. Our purpose in this section is to examine the effectiveness of such estimators.

In this chapter we were primarily interested in a linear regression set-up with common parameter. We found that the same requirement of sample sizes repeated to hold in situations wherein the p (≥ 2) linear regression models involve k (> 1) common estimable parameter(s)

---

[1] Part of this chapter is coming from a working paper "On a General Pattern of Domination Using the Graybill–Deal Estimator" from Nie, Sinha and Hedayat.

in the mean models. In this context, we used the criterion of ' Loewner Order Domination' of information or dispersion matrices.

2.1 Introduction

Let $X_i \sim N(\mu, \sigma_i^2)$, $i = 1, 2, \ldots, p$ be p independent normal distributions sharing common mean μ and unknown but possibly unequal variances σ_i^2. Let n_1, n_2, ..., n_p be the corresponding sample size for each population. We have the following notations for sample means and sample variances respectively:

$$\bar{x}_i = \frac{1}{n_i}\sum_{j=1}^{n_i} x_{ij},$$
$$s_i^2 = \frac{\sum_{j=1}^{n_i}(x_{ij} - \bar{x}_i)^2}{n_i - 1}. \tag{2.1}$$

Here x_{ij} represent the independent observations from i^{th} normal population. We notice that $\bar{x}_i$'s and s_i^2's are mutually independent with

$$\bar{x}_i \sim N(\mu, \frac{\sigma_i^2}{n_i}) \ and \ \frac{(n_i - 1)s_i^2}{\sigma_i^2} \sim \chi^2(n_i - 1), \ i = 1, \ 2, \ \ldots, \ p.$$

Furthermore, the sample means and sample variances are minimal sufficient statistics.

For $p = 2$ and assuming that the two true variances are known, the sufficient statistics of $\bar{x}_1 \frac{n_1}{\sigma_1^2} + \bar{x}_2 \frac{n_2}{\sigma_2^2}$ is also complete for μ. Cochran (14) suggested the unbiased estimator

$$\hat{\mu}_C = \frac{\left(\bar{x}_1 \frac{n_1}{\sigma_1^2} + \bar{x}_2 \frac{n_2}{\sigma_2^2}\right)}{\frac{n_1}{\sigma_1^2} + \frac{n_2}{\sigma_2^2}}. \tag{2.2}$$

This estimator is the BLUE for μ, which is also the UMVUE.

When σ_1^2 and σ_2^2 are unknown, but the ratio $\zeta = \frac{\sigma_1^2}{\sigma_2^2}$ is known. We shall still have that $n_1\bar{x}_1 + n_2\zeta\bar{x}_2$ is the complete sufficient statistics. Hence we still have the UMVUE as follows:

$$\hat{\mu}_{UMVUE} = \frac{n_1\bar{x}_1 + n_2\zeta\bar{x}_2}{n_1 + n_2\zeta}. \tag{2.3}$$

It is trivial to note that $\hat{\mu}_{UMVUE}$ can be directly derived from $\hat{\mu}_C$.

If $\zeta = \frac{\sigma_1^2}{\sigma_2^2}$ is unknown, we only have $(\bar{x}_1, \bar{x}_2, s_1^2, s_2^2)$ being the minimal sufficient statistics. But they are not complete, since $E(\bar{x}_1 - \bar{x}_2) = 0$. The following Theorem [Lehmann (27)] showed that UMVUE of μ does not exist when ζ is unknown.

Theorem 1. *The UMVUE of μ does not exist when ζ is unknown.*

Proof.

We prove this by contradiction. Suppose the UMVUE of μ exists, denote as $U = \phi\bar{x}_1 + (1-\phi)\bar{x}_2$. Clearly ϕ is in the range $[0, 1]$. For any ζ_0 belong to $[0, \infty)$, $\hat{\mu}_{\zeta_0} = \frac{n_1\bar{x}_1 + n_2\zeta_0\bar{x}_2}{n_1 + n_2\zeta_0}$ is the UMVUE of μ when $\zeta = \zeta_0$. Since the UMVUE is unique, there must exist a ζ_0, such that $U = \hat{\mu}_{\zeta_0}$ for any value of ζ. But this contradicts the assumption that ζ is unknown. $\square$

When ζ is unknown, Graybill and Deal (19) introduced their plug-in estimator, also known as the GDE in literature, $\hat{\mu}_{GD|2}$ for $p = 2$ groups,

$$\hat{\mu}_{GD|2} = \frac{\left(\bar{x}_1\frac{n_1}{s_1^2} + \bar{x}_2\frac{n_2}{s_2^2}\right)}{\frac{n_1}{s_1^2} + \frac{n_2}{s_2^2}}, \tag{2.4}$$

and showed that $\hat{\mu}_{GD|2}$ was preferable than both sample means in the criteria of mean square error, if and only if sample sizes n_1 and n_2 were moderately enough (≥ 11), which was corrected by Khatri and Shah (23) as ($n_1 \geq 11, n_2 \geq 11$), ($n_1 = 10, n_2 \geq 19$) or ($n_1 \geq 19, n_2 = 10$). Later Graybill and Deal's (19) results were generalized to p populations by Norwood and Hinkelmann (33). Furthermore, Shinozaki (44) gave a proof showing that p-source based GDE dominates its any q-sub-source based competitors, again if and only if the same conditions continue to hold. Specifically, we have the following theorem.

Theorem 2. *1. The estimator* $\hat{\mu}_{GD|p} = \sum_{i=1}^{p} \frac{\bar{x}_i \frac{n_i}{s_i^2}}{\sum_{j=1}^{p} \frac{n_j}{s_j^2}}$ *is unbiased for* μ.

2. $Var(\hat{\mu}_{GD|p}) < \frac{\sigma_i^2}{n_i}$ *for all values of* σ_i^2 *(*$i = 1, 2, \ldots, p$*) if and only if*

(a) $n_i > 10$, $i = 1, 2, \ldots, p$ *or*

(b) $n_i = 10$ *for some* i*, and* $n_j > 18$ *for* $i, j \in \{1, 2, \ldots, p\}$ *and each* $j \neq i$.

3. Furthermore, $Var(\hat{\mu}_{GD|p}) < Var(\hat{\mu}_{GD|q})$ *for all* q*-sub populations and all values of* σ_i^2 *(*$i = 1, 2, \ldots, p$*) if and only if the same conditions hold.*

Proof.

See Shinozaki (44). □

Chiou and Cohen (12) tried to extend the above results to multivariate normal case, but they reported some negative type results. To be more specific, "Let $\mathbb{X}_i$, $i = 1, 2, \ldots, n$ be a random sample of size n, from a k dimensional multivariate normal distribution with mean vector $\boldsymbol{\mu}$ and covariance matrix $\Sigma_{\mathbb{X}}$. Let $\mathbb{Y}_i$, $i = 1, 2, \ldots, n$ be a random sample of size n from a

k dimensional multivariate normal distribution $N(\boldsymbol{\mu}, \Sigma_{\mathbb{Y}})$. Assume the $\mathbb{X}$-sample and $\mathbb{Y}$-sample are independent." We had the following notations:

$$\begin{aligned}
\bar{\mathbb{X}} &= \frac{\sum_{i=1}^{n} \mathbb{X}_i}{n}, \\
\bar{\mathbb{Y}} &= \frac{\sum_{i=1}^{n} \mathbb{Y}_i}{n}, \\
S_{\mathbb{X}} &= \frac{\sum_{i=1}^{n} (\mathbb{X}_i - \bar{\mathbb{X}})(\mathbb{X}_i - \bar{\mathbb{X}})'}{(n-1)}, \\
S_{\mathbb{Y}} &= \frac{\sum_{i=1}^{n} (\mathbb{Y}_i - \bar{\mathbb{Y}})(\mathbb{Y}_i - \bar{\mathbb{Y}})'}{(n-1)}, \\
\Sigma_{\bar{\mathbb{X}}} &= \frac{\Sigma_{\mathbb{X}}}{n}, \\
\Sigma_{\bar{\mathbb{Y}}} &= \frac{\Sigma_{\mathbb{Y}}}{n}, \\
S_{\bar{\mathbb{X}}} &= \frac{S_{\mathbb{X}}}{n}, \\
S_{\bar{\mathbb{Y}}} &= \frac{S_{\mathbb{Y}}}{n}. \qquad (2.5)
\end{aligned}$$

Considering the problem of estimating the common mean vector $\boldsymbol{\mu}$ while $\Sigma_{\mathbb{X}}$ and $\Sigma_{\mathbb{Y}}$ are unknown, Chiou and Cohen (12) suggested a Graybill-Deal-type estimator:

$$\mathbb{T} = S_{\bar{\mathbb{Y}}}(S_{\bar{\mathbb{X}}} + S_{\bar{\mathbb{Y}}})^{-1}\bar{\mathbb{X}} + S_{\bar{\mathbb{X}}}(S_{\bar{\mathbb{X}}} + S_{\bar{\mathbb{Y}}})^{-1}\bar{\mathbb{Y}}. \qquad (2.6)$$

Surprisingly Chiou and Cohen (12) found that neither $\Sigma_{\bar{\mathbb{X}}} - \Sigma_{\mathbb{T}}$ nor $\Sigma_{\bar{\mathbb{Y}}} - \Sigma_{\mathbb{T}}$ was positive semi-definite for all ($\Sigma_{\mathbb{X}}$, $\Sigma_{\mathbb{Y}}$) for any n, where $\Sigma_{\mathbb{T}}$ was the covariance matrix.

Theorem 3. *The differences $\Sigma_{\bar{\mathbb{X}}} - \Sigma_{\mathbb{T}} \ngeq 0$ and $\Sigma_{\bar{\mathbb{Y}}} - \Sigma_{\mathbb{T}} \ngeq 0$ for all ($\Sigma_{\mathbb{X}}$, $\Sigma_{\mathbb{Y}}$).*

Proof.

See Chiou and Cohen (12). □

In the next section, we were trying to solve the similar question: estimating the common mean vector from multivariate normal population with a known covariance matrix up to a constant. Specially we considered the situation of regular linear regression models with common parameter estimation. However we had some positive type results to be reported.

2.2 Common Parameter Estimation in Regular linear regression models with Independent Normal Errors

Consider p independent linear regression models in matrix form, with sample size n_i, $i = 1, 2, \ldots, p$. At the i^{th} population, we have:

$$\boldsymbol{Y}_i = \boldsymbol{X}_i\boldsymbol{\theta} + \boldsymbol{Z}_i\boldsymbol{\tau}_i + \boldsymbol{\epsilon}_i.$$

Where $\boldsymbol{Y}_i = (y_{i1}, y_{i2}, \dots, y_{in_i})'$ is the response variable for i^{th} population. $\boldsymbol{X}_i$ and $\boldsymbol{Z}_i$ are corresponding design matrices with size $n_i \times k$ and $n_i \times t_i$ respectively.

$$\boldsymbol{X}_i = \begin{pmatrix} x_{i11} & x_{i21} & \dots & x_{ik1} \\ x_{i12} & x_{i22} & \dots & x_{ik2} \\ \dots & \dots & \dots & \dots \\ x_{i1n_i} & x_{i2n_i} & \dots & x_{ikn_i} \end{pmatrix},$$

$$\boldsymbol{Z}_i = \begin{pmatrix} z_{i11} & z_{i21} & \dots & z_{it_i1} \\ z_{i12} & z_{i22} & \dots & z_{it_i2} \\ \dots & \dots & \dots & \dots \\ z_{i1n_i} & z_{i2n_i} & \dots & z_{it_in_i} \end{pmatrix}.$$

The i^{th} linear model error term is $\boldsymbol{\epsilon}_i = (\epsilon_{i1}, \epsilon_{i2}, \dots, \epsilon_{in_i})$. Each ϵ_{ij} is identical and follows $N(0, \sigma_i^2)$ for all $j = 1, 2, \dots, n_i$. We can also denote as:

$$\boldsymbol{\epsilon}_i \sim N\left(\mathbf{0}_{(n_i \times 1)}, \sigma_i^2 \mathbf{I}_{(n_i \times n_i)}\right),$$

where $\mathbf{I}_{(n_i \times n_i)}$ is the identity matrix with size $n_i \times n_i$.

We assume that these p independent linear regression models share a k-dimensional (k-dim) common estimable parameter vector $\boldsymbol{\theta} = (\theta_1, \theta_2, \dots, \theta_k)'$, and an extra t_i-dimensional estimable parameter vector $\boldsymbol{\tau}_i = (\tau_{i1}, \tau_{i2}, \ldots, \tau_{it_i})'$ at i^{th} model, $i = 1, 2, \ldots, p$. All of them are fixed. From the normal equations, the OLS unbiased estimator,

$$(\hat{\theta}_1, \hat{\theta}_2, \ldots, \hat{\theta}_k, \hat{\tau}_{i1}, \hat{\tau}_{i2}, \ldots, \hat{\tau}_{it_i})' = \left(\left(\boldsymbol{X}_i \;\vdots\; \boldsymbol{Z}_i\right)^T \left(\boldsymbol{X}_i \;\vdots\; \boldsymbol{Z}_i\right)\right)^{-1} \left(\boldsymbol{X}_i \;\vdots\; \boldsymbol{Z}_i\right)^T \boldsymbol{Y}_i,$$

of $\begin{pmatrix} \boldsymbol{\theta} \\ \cdots \\ \boldsymbol{\tau}_i \end{pmatrix}$ based on the data arising out of the i^{th} model. [1]

Hence $\hat{\boldsymbol{\theta}}_i = (\hat{\theta}_1, \hat{\theta}_2, \ldots, \hat{\theta}_k)'$ has the following distribution

$$\hat{\boldsymbol{\theta}}_i \sim N\left(\boldsymbol{\theta}, \sigma_i^2 \boldsymbol{W}_{i(k\times k)}\right),$$

where $\boldsymbol{W}_{i(k\times k)}$ is the $k \times k$ upper submatrix of $\left(\left(\boldsymbol{X}_i \;\vdots\; \boldsymbol{Z}_i\right)^T \left(\boldsymbol{X}_i \;\vdots\; \boldsymbol{Z}_i\right)\right)^{-1}$, and $\boldsymbol{W}_{i(k\times k)}$'s are nonsingular matrices.

An unbiased estimator for σ_i^2 is the mean square residuals s_{ri}^2, which is the sum of square residuals divided by the degrees of freedom $\nu_i = n_i - k - t_i$:

$$s_{ri}^2 = \frac{\sum_{j=1}^{n_i} (y_{ij} - \hat{y}_{ij})^2}{n_i - k - t_i} = \frac{\sum_{j=1}^{n_i} e_{ij}^2}{n_i - k - t_i}, \tag{2.7}$$

---

[1]If A is a matrix of order $p \times q$ and B is another matrix of order $p \times r$, then $(\boldsymbol{A} \vdots \boldsymbol{B})$ represents a matrix of order $p \times (q + r)$, wherein the columns of A are preceded by the columns of B without any change of their relative positions.

for all $i = 1, 2, \ldots, p$, where $e_{ij} = y_{ij} - \hat{y}_{ij}$ is the residual for j^{th} observation at i^{th} model and $\hat{y}_{ij}$ is the predicted value respectively. Further, it is known that $\frac{\nu_i s_{ri}^2}{\sigma_i^2} \sim \chi^2(\nu_i)$. [1]

In the sequel, we shall deal with the estimation of the common $\boldsymbol{\theta} = (\theta_1, \theta_2, \ldots, \theta_k)'$ under three different scenarios: in Section 2.3, we shall study the case $k = 1$ for $p \geq 2$ groups; in Section 2.4, we shall study the case $k \geq 2$ for $p = 2$ groups; and in Section 2.5, we shall study the case $k \geq 2$ for $p > 2$ groups.

2.3 Single Common Parameter Involving p Groups

Suppose there is only one single common parameter α among these p independent linear regression models. As mentioned above, under certain conditions, the GDE will be efficient. Here we use a simple example to express our ideas by assuming that common α is the intercept in two linear regression models.

Example 1. *Consider two simple independent linear regression models involving unequal unknown variances. We have*

$$\mathbf{Y}_1 = \alpha_1 \boldsymbol{1}_{(n_1 \times 1)} + \beta_1 \mathbf{X}_{(n_1 \times 1)} + \boldsymbol{\epsilon}_1$$

$$\mathbf{Y}_2 = \alpha_2 \boldsymbol{1}_{(n_2 \times 1)} + \gamma_1 \mathbf{Z}_{(n_2 \times 1)} + \boldsymbol{\epsilon}_2$$

$$\boldsymbol{\epsilon}_1 \sim N(\boldsymbol{0}_{(n_1 \times 1)}, \sigma_1^2 \mathbf{I}_{(n_1 \times n_1)})$$

$$\boldsymbol{\epsilon}_2 \sim N(\boldsymbol{0}_{(n_2 \times 1)}, \sigma_2^2 \mathbf{I}_{(n_2 \times n_2)}),$$

---

[1]It is tacitly assumed that all nuisance parameters (τ_i's) are estimable.

[1] *where*

$$\mathbf{Y}_1 = (y_{11}, y_{12}, \ldots, y_{1n_1})',$$

$$\mathbf{Y}_2 = (y_{21}, y_{22}, \ldots, y_{2n_2})',$$

$$\boldsymbol{1}_{(n_1 \times 1)} = \underbrace{(1, 1, \ldots, 1)}_{n_1}',$$

$$\boldsymbol{1}_{(n_2 \times 1)} = \underbrace{(1, 1, \ldots, 1)}_{n_2},$$

$$\mathbf{X}_{(n_1 \times 1)} = (x_1, x_2, \ldots, x_{n_1})',$$

$$\mathbf{Z}_{(n_2 \times 1)} = (z_1, z_2, \ldots, z_{n_2})'.$$

The OLS estimators $(\hat{\alpha}_1, \hat{\beta}_1)$ *and* $(\hat{\alpha}_2, \hat{\gamma}_1)$*, respectively for* (α_1, β_1) *and* (α_2, γ_1)*, can be expressed as:*

$$\begin{pmatrix} \hat{\alpha}_1 \\ \hat{\beta}_1 \end{pmatrix} = \begin{pmatrix} n_1 & \sum_{i=1}^{n_1} x_i \\ \sum_{i=1}^{n_1} x_i & \sum_{i=1}^{n_1} x_i^2 \end{pmatrix}^{-1} \begin{pmatrix} 1 & 1 & \ldots & 1 \\ x_1 & x_2 & \ldots & x_{n_1} \end{pmatrix} \begin{pmatrix} y_{11} \\ y_{12} \\ \ldots \\ y_{1n_1} \end{pmatrix},$$

---

[1] *Here* $\mathbf{X}_{(n_1 \times 1)}$ *and* $\mathbf{Z}_{(n_2 \times 1)}$ *are different from previous section.*

$$\begin{pmatrix}\hat{\alpha}_2\\ \hat{\gamma}_1\end{pmatrix} = \begin{pmatrix} n_2 & \sum_{i=1}^{n_2} z_i \\ \sum_{i=1}^{n_2} z_i & \sum_{i=1}^{n_2} z_i^2\end{pmatrix}^{-1} \begin{pmatrix} 1 & 1 & \dots & 1 \\ & & & \\ z_1 & z_2 & \dots & z_{n_2}\end{pmatrix} \begin{pmatrix} y_{21} \\ y_{22} \\ \dots \\ y_{2n_2}\end{pmatrix};$$

and have the following variance-covariance matrices:

$$\mathbf{Var}\begin{pmatrix}\hat{\alpha}_1\\ \hat{\beta}_1\end{pmatrix} = \sigma_1^2 \begin{pmatrix} n_1 & \sum_{i=1}^{n_1} x_i \\ \sum_{i=1}^{n_1} x_i & \sum_{i=1}^{n_1} x_i^2\end{pmatrix}^{-1}$$

and

$$\mathbf{Var}\begin{pmatrix}\hat{\alpha}_2\\ \hat{\gamma}_1\end{pmatrix} = \sigma_2^2 \begin{pmatrix} n_2 & \sum_{i=1}^{n_2} z_i \\ \sum_{i=1}^{n_2} z_i & \sum_{i=1}^{n_2} z_i^2\end{pmatrix}^{-1}.$$

Suppose $\alpha_1 = \alpha_2 = \alpha$, *then it follows :*

$$E(\hat{\alpha}_1) = E(\hat{\alpha}_2) = \alpha,$$
$$Var(\hat{\alpha}_1) = \sigma_1^2(\frac{1}{n_1} + \frac{\bar{x}^2}{SS_x}),$$
$$Var(\hat{\alpha}_2) = \sigma_2^2(\frac{1}{n_2} + \frac{\bar{z}^2}{SS_z})$$

where

$$SS_x = \sum_{i=1}^{n_1}(x_i - \bar{x})^2,$$
$$SS_z = \sum_{i=1}^{n_2}(z_i - \bar{z})^2,$$

$$\bar{x} = \frac{1}{n_1} \sum_{i=1}^{n_1} x_i,$$
$$\bar{z} = \frac{1}{n_2} \sum_{i=1}^{n_2} z_i$$

Denote $\frac{1}{n_1} + \frac{\bar{x}^2}{SS_x}$ *by* w_1^{-1}, *and* $\frac{1}{n_2} + \frac{\bar{z}^2}{SS_z}$ *by* w_2^{-1}.

The GDE of α *combing these two estimates is*

$$\alpha_{GD|2} = \frac{\frac{\hat{\alpha}_1}{s_{r1}^2 w_1^{-1}} + \frac{\hat{\alpha}_2}{s_{r2}^2 w_2^{-1}}}{\frac{1}{s_{r1}^2 w_1^{-1}} + \frac{1}{s_{r2}^2 w_2^{-1}}}.$$

It is noticed that $\hat{\alpha}_{GD|2}$ defined here is little different from Graybill-Deal (19)'s definition. We replaced n_1 and n_2 by w_1 and w_2 respectively. Later we will show that this change does not affect the Graybill-Deal (19)'s claim. We can easily extend the formula for $\hat{\alpha}_{GD|2}$ to

$$\alpha_{GD|p} = \sum_{i=1}^{p} \frac{\hat{\alpha}_i \frac{w_i}{s_{ri}^2}}{\sum_{i=1}^{p} \frac{w_i}{s_{ri}^2}}$$

in case there are p such models to be combined. It is to be noted that $w_1, w_2, \ldots, w_p$ have similar algebraic expressions. At this stage, we will state and prove a general result on the property of $\hat{\alpha}_{GD|p}$. By a simple application of the results, from Norwood and Hinkelmann (33), and Shinozaki (44), we have the following Theorem 4 for p independent linear regression models sharing one single common intercept parameter α. We may note in passing that ν_i defined above in the beginning of Section 2 assumes the form $\nu_i = n_i - 2$ for the model in Example 1 being studied.

Theorem 4. *If p independent linear regression models share one single common intercept parameter α, $i = 1, 2, \ldots, p$, then we have the following results:*

1. *$\hat{\alpha}_{GD|p}$ is an unbiased estimator for α.*

2. *A necessary and sufficient condition for $\hat{\alpha}_{GD|p}$ to have a smaller variance than each $\hat{\alpha}_i$, for all values of σ_i^2, $w_i > 0$ $(i = 1, \ldots, p)$ is either*

   (a) $\nu_i > 9$, $i = 1, 2, \ldots, p$ or

   (b) $\nu_i = 9$ for some i, and $\nu_j > 17$ for $i, j \in \{1, 2, \ldots, p\}$ and each $j \neq i$.

Moreover, if either Condition (2a) or (2b) is satisfied, then $\hat{\alpha}_{GD|p}=$ $\phi_p(\hat{\alpha}_1, \ldots, \hat{\alpha}_p; s_1^2, \ldots, s_p^2)$ has a smaller variance than $\hat{\alpha}_{GD|q} = \phi_q(\hat{\alpha}_1, \ldots, \hat{\alpha}_q; s_1^2, \ldots, s_q^2)$ of any q $(< p)$ subgroups .

Proof.

Set

$$\sigma_i^{2\star} = \frac{\sigma_i^2}{w_i},$$
$$s_{ri}^{2\ \star} = \frac{s_{ri}^2}{w_i}$$

for $i = 1, 2, \ldots, p$. Then we have

$$\hat{\alpha}_i \sim N(\alpha, \sigma_i^{2\star})$$

and

$$\frac{\nu_i s_{ri}^{2\ \star}}{\sigma_i^{2\star}} = \frac{\nu_i s_{ri}^2}{\sigma_i^2} \sim \chi^2(\nu_i)$$

for $i = 1, 2, \ldots, p$.

Then the GDE

$$\hat{\alpha}_{GD|p} = \sum_{i=1}^{p} \frac{\hat{\alpha}_i \frac{w_i}{s_{ri}^2}}{\sum_{i=1}^{p} \frac{w_i}{s_{ri}^2}} = \sum_{i=1}^{p} \frac{\frac{\hat{\alpha}_i}{s_{ri}^{2\ \star}}}{\sum_{i=1}^{p} \frac{1}{s_{ri}^{2\ \star}}}.$$

This is the same setting as in Norwood and Hinkelmann (33). It is easy to show that $\hat{\alpha}_{GD|p}$ is an unbiased estimator for α, and

$$Var(\hat{\alpha}_{GD|p}) < \sigma_i^{2\star} = \frac{\sigma_i^2}{w_i} \quad \forall i$$

if and only if either Condition (2a) or (2b) holds.

Furthermore, we notice that our $\hat{\alpha}_{GD|p}$ is a special case of $\hat{\alpha}_S$ at $c_i = 1$ for $i = 1, \ldots, p$. From Shinozaki (44)'s results, we know that

$$Var(\hat{\alpha}_{GD|p}) < Var(\hat{\alpha}_{GD|q}), \quad q < p$$

if and only if

$$2\frac{\nu_i(\nu_j - 4)}{(\nu_i + 2)\nu_j} \geq \frac{c_j}{c_i} = 1$$

for any $i \neq j$. This is equivalent to condition (2a) or (2b) as stated above. □

Remark 1. *The statistical independence of the $\hat{\alpha}_i$ and s_{ri}^2 guarantees $\hat{\alpha}_{GD|p}$ to be an unbiased estimator of α.*

Remark 2. *The necessary and sufficient condition in (2a) or (2b) only concerns the degrees of freedom ν_i in group i, which is related to the sample size n_i. Theorem 4 indicates us that if the sample size is reasonable enough (subject to condition in (2a) or (2b)), then the GDE of a single common parameter utilizing p independent linear regression models always provides a more efficient unbiased estimator than any single group or any $q(< p)$ subgroups.*

2.4 k-dim Common Parameter Involving Two Groups

In this section, we consider $p = 2$ independent linear regression models. The k-dim GDE of the common estimable parameter vector $\boldsymbol{\theta}_{(k\times 1)}$ in matrix forms is:

$$\hat{\boldsymbol{\theta}}_{GD|2} = \left((s_{r1}^2 \boldsymbol{W_1})^{-1} + (s_{r2}^2 \boldsymbol{W_2})^{-1}\right)^{-1} \left((s_{r1}^2 \boldsymbol{W_1})^{-1}\hat{\boldsymbol{\theta}}_1 + (s_{r2}^2 \boldsymbol{W_2})^{-1}\hat{\boldsymbol{\theta}}_2\right).$$

It is easily determined that this k-dim GDE is an unbiased estimator of $\boldsymbol{\theta}$. Since

$$\begin{aligned}
&\boldsymbol{E}(\hat{\boldsymbol{\theta}}_{GD|2}) \\
&= \boldsymbol{E}\left(\boldsymbol{E}(\hat{\boldsymbol{\theta}}_{GD|2}|s_{r1}^2, s_{r2}^2)\right) \\
&= \boldsymbol{E}\left(\left((s_{r1}^2 \boldsymbol{W_1})^{-1} + (s_{r2}^2 \boldsymbol{W_2})^{-1}\right)^{-1} \left((s_{r1}^2 \boldsymbol{W_1})^{-1}\boldsymbol{E}(\hat{\boldsymbol{\theta}}_1|s_{r1}^2, s_{r2}^2) + (s_{r2}^2 \boldsymbol{W_2})^{-1}\boldsymbol{E}(\hat{\boldsymbol{\theta}}_2|s_{r1}^2, s_{r2}^2)\right)\right) \\
&= \boldsymbol{E}\left(\left((s_{r1}^2 \boldsymbol{W_1})^{-1} + (s_{r2}^2 \boldsymbol{W_2})^{-1}\right)^{-1} \left((s_{r1}^2 \boldsymbol{W_1})^{-1}\boldsymbol{\theta} + (s_{r2}^2 \boldsymbol{W_2})^{-1}\boldsymbol{\theta}\right)\right) \\
&= \boldsymbol{\theta}.
\end{aligned}$$

The dispersion matrix of $\hat{\boldsymbol{\theta}}_{GD|2}$ is:

$$\boldsymbol{D}(\hat{\boldsymbol{\theta}}_{GD|2}) \tag{2.8}$$

$$= \boldsymbol{E}\left(\boldsymbol{D}\left(\hat{\boldsymbol{\theta}}_{GD|2}|s_{r1}^2, s_{r2}^2\right)\right) + \boldsymbol{D}\left(\boldsymbol{E}\left(\hat{\boldsymbol{\theta}}_{GD|2}|s_{r1}^2, s_{r2}^2\right)\right) \tag{2.9}$$

$$= \boldsymbol{E}\left(\boldsymbol{D}\left(\hat{\boldsymbol{\theta}}_{GD|2}|s_{r1}^2, s_{r2}^2\right)\right) + \boldsymbol{D}\left(\boldsymbol{\theta}\right) \tag{2.10}$$

$$= \boldsymbol{E}\left(\boldsymbol{D}\left(\hat{\boldsymbol{\theta}}_{GD|2}|s_{r1}^2, s_{r2}^2\right)\right) + 0 \tag{2.11}$$

$$= \boldsymbol{E}\left(\left(s_{r1}^{-2}\boldsymbol{W_1}^{-1} + s_{r2}^{-2}\boldsymbol{W_2}^{-1}\right)^{-1}\left(\frac{\sigma_1^2}{s_{r1}^4}\boldsymbol{W_1}^{-1} + \frac{\sigma_2^2}{s_{r2}^4}\boldsymbol{W_2}^{-1}\right)\left(s_{r1}^{-2}\boldsymbol{W_1}^{-1} + s_{r2}^{-2}\boldsymbol{W_2}^{-1}\right)^{-1}\right).$$

In the above, we conditionally fixed s_{r1}^2 and s_{r2}^2. So at the end, we only need to compute expectation with respect to these variance estimates.

By way of notation, if a dispersion matrix $\boldsymbol{A}$ is non-negative definite (n.n.d.) we write $\boldsymbol{A} \geq 0$, if it is positive definite (p.d.) we write $\boldsymbol{A} > 0$. The Loewner order domination of a dispersion matrix $\boldsymbol{A}$ over $\boldsymbol{B}$ ($\boldsymbol{A} > \boldsymbol{B}$) means $\boldsymbol{A} - \boldsymbol{B} > 0$, the Loewner order of $\boldsymbol{A}$ below $\boldsymbol{B}$ ($\boldsymbol{A} < \boldsymbol{B}$) means $\boldsymbol{A} - \boldsymbol{B} < 0$.

Lemma 5. *If $\boldsymbol{W_1}$ and $\boldsymbol{W_2}$ are diagonal matrices, then $\boldsymbol{D}(\hat{\boldsymbol{\theta}}_{GD|2}) < min(\sigma_1^2\boldsymbol{W}_1, \sigma_2^2\boldsymbol{W}_2)$ if and only if either*

1. *$\nu_i > 9$, $i = 1, 2$ or*

2. *$\nu_i = 9$ for some i, and $\nu_j > 17$ for the other $j \neq i$.*

Proof.

Let

$$\boldsymbol{W_1} = \boldsymbol{Diag}(w_{1g}^{-1}) = \begin{pmatrix} w_{1g}^{-1} & 0 & \dots & 0 \\ 0 & w_{12}^{-1} & \dots & 0 \\ \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & w_{1k}^{-1} \end{pmatrix}$$

and

$$\boldsymbol{W_2} = \boldsymbol{Diag}(w_{2g}^{-1}) = \begin{pmatrix} w_{2g}^{-1} & 0 & \dots & 0 \\ 0 & w_{22}^{-1} & \dots & 0 \\ \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & w_{2k}^{-1} \end{pmatrix},$$

where w_{1g}^{-1} and w_{2g}^{-1} are the g-th diagonal entries of matrices $\boldsymbol{W_1}$ and $\boldsymbol{W_2}$ respectively, $g = 1, \dots, k$. To prove this Lemma, it is enough to show the g-th diagonal entry of the dispersion matrix

$$\boldsymbol{D}(\hat{\boldsymbol{\theta}}_{GD|2})_{gg} = Var\left(\frac{\frac{\hat{\theta}_{1g}}{s_{r1}^2 w_{1g}^{-1}} + \frac{\hat{\theta}_{2g}}{s_{r2}^2 w_{2g}^{-1}}}{\frac{1}{s_{r1}^2 w_{1g}^{-1}} + \frac{1}{s_{r2}^2 w_{2g}^{-1}}}\right) \le min(\frac{\sigma_1^2}{w_{1g}}, \frac{\sigma_2^2}{w_{2g}}),$$

for all $g = 1, 2, \dots, k$, and for all values of σ_1^2, σ_2^2.

From Theorem 4, it is known that:

$$Var\left(\frac{\frac{\hat{\theta}_{1g}}{s_{r1}^2 w_{1g}^{-1}} + \frac{\hat{\theta}_{2g}}{s_{r2}^2 w_{2g}^{-1}}}{\frac{1}{s_{r1}^2 w_{1g}^{-1}} + \frac{1}{s_{r2}^2 w_{2g}^{-1}}}\right) < min\left(\frac{\sigma_1^2}{w_{1g}}, \frac{\sigma_2^2}{w_{2g}}\right).$$

if and only if condition stated in (1) or (2) of Lemma 5 holds. $\square$

Remark 3. *When $\boldsymbol{W_1}$ and $\boldsymbol{W_2}$ are k-dim diagonal matrices, we can decompose the k-dim GDE into k simple single GDEs.*

It is pertinent to observe that the above result holds even without the two matrices being diagonal matrices. This is established below.

Theorem 6. *In two independent linear regression models, the Loewner order of $\boldsymbol{D}(\hat{\boldsymbol{\theta}}_{GD|2})$ is below $min(\sigma_1^2\boldsymbol{W}_1, \sigma_2^2\boldsymbol{W}_2)$, for all values of σ_1^2, σ_2^2, if and only if condition stated in (1) or (2) of Lemma 5 holds:*

1. $\nu_i > 9$, $i = 1, 2$ *or*
2. $\nu_i = 9$ *for some i, and* $\nu_j > 17$ *for the other* $j \neq i$.

Proof.

Notice that $\boldsymbol{W_1}^{-1}$ and $\boldsymbol{W_2}^{-1}$ are positive definite matrices. We need to show:

$$\begin{aligned}\boldsymbol{E}\Bigg(&\left(s_{r1}^{-2}\boldsymbol{W_1}^{-1} + s_{r2}^{-2}\boldsymbol{W_2}^{-1}\right)^{-1}\left(\frac{\sigma_1^2}{s_{r1}^4}\boldsymbol{W_1}^{-1} + \frac{\sigma_2^2}{s_{r2}^4}\boldsymbol{W_2}^{-1}\right)\\ &\left(s_{r1}^{-2}\boldsymbol{W_1}^{-1} + s_{r2}^{-2}\boldsymbol{W_2}^{-1}\right)^{-1}\Bigg) < \sigma_1^2\boldsymbol{W}_1,\end{aligned} \tag{2.12}$$

and

$$\boldsymbol{E}\Bigg(\left(s_{r1}^{-2}\boldsymbol{W_1}^{-1} + s_{r2}^{-2}\boldsymbol{W_2}^{-1}\right)^{-1}\left(\frac{\sigma_1^2}{s_{r1}^4}\boldsymbol{W_1}^{-1} + \frac{\sigma_2^2}{s_{r2}^4}\boldsymbol{W_2}^{-1}\right) \tag{2.13}$$

$$\left(s_{r1}^{-2}\boldsymbol{W_1}^{-1} + s_{r2}^{-2}\boldsymbol{W_2}^{-1}\right)^{-1}\Bigg) < \sigma_2^2\boldsymbol{W}_2. \tag{2.14}$$

For Equation 2.12, we can obtain

$$\boldsymbol{E}\left(\boldsymbol{W_1}^{-1/2}\left(s_{r1}^{-2}\boldsymbol{W_1}^{-1}+s_{r2}^{-2}\boldsymbol{W_2}^{-1}\right)^{-1}\boldsymbol{W_1}^{-1/2}\boldsymbol{W_1}^{1/2}\left(\frac{\sigma_1^2}{s_{r1}^4}\boldsymbol{W_1}^{-1}+\frac{\sigma_2^2}{s_{r2}^4}\boldsymbol{W_2}^{-1}\right)\right. \tag{2.15}$$

$$\left.\left(\boldsymbol{W_1}^{1/2}\boldsymbol{W_1}^{-1/2}\left(s_{r1}^{-2}\boldsymbol{W_1}^{-1}+s_{r2}^{-2}\boldsymbol{W_2}^{-1}\right)^{-1}\boldsymbol{W_1}^{-1/2}\right)\right)<\sigma_1^2\boldsymbol{W_1}^{-1/2}\boldsymbol{W}_1\boldsymbol{W_1}^{-1/2}. \tag{2.16}$$

Denote $\boldsymbol{W_1}^{-1/2}\boldsymbol{W}_2\boldsymbol{W_1}^{-1/2}$ by $\boldsymbol{A}$. Then Equation 2.15 can be rewritten as:

$$\boldsymbol{E}\left(\left(s_{r1}^{-2}\boldsymbol{I}^{-1}+s_{r2}^{-2}\boldsymbol{A}^{-1}\right)^{-1}\left(\frac{\sigma_1^2}{s_{r1}^4}\boldsymbol{I}^{-1}+\frac{\sigma_2^2}{s_{r2}^4}\boldsymbol{A}^{-1}\right)\left(s_{r1}^{-2}\boldsymbol{I}^{-1}+s_{r2}^{-2}\boldsymbol{A}^{-1}\right)^{-1}\right)<\sigma_1^2\boldsymbol{I}.$$

Due to the fact that $\boldsymbol{A}$ is symmetric, there exists an orthogonal matrix $\boldsymbol{P}$, while $\boldsymbol{P}^T\boldsymbol{P}=\boldsymbol{P}\boldsymbol{P}^T=I$, such that $\boldsymbol{P}^T\boldsymbol{A}\boldsymbol{P}=\boldsymbol{C}$, where $\boldsymbol{C}$ is a diagonal matrix.

Then we have:

$$\boldsymbol{E}\left(\boldsymbol{P}^T\left(s_{r1}^{-2}\boldsymbol{I}^{-1}+s_{r2}^{-2}\boldsymbol{A}^{-1}\right)^{-1}\boldsymbol{P}\boldsymbol{P}^T\left(\frac{\sigma_1^2}{s_{r1}^4}\boldsymbol{I}^{-1}+\frac{\sigma_2^2}{s_{r2}^4}\boldsymbol{A}^{-1}\right)\boldsymbol{P}\boldsymbol{P}^T\left(s_{r1}^{-2}\boldsymbol{I}^{-1}+s_{r2}^{-2}\boldsymbol{A}^{-1}\right)^{-1}\boldsymbol{P}\right) \tag{2.17}$$

$$<\sigma_1^2\boldsymbol{P}^T\boldsymbol{I}\boldsymbol{P}.$$

Upon simplification:

$$\boldsymbol{E}\left(\left(s_{r1}^{-2}\boldsymbol{I}^{-1}+s_{r2}^{-2}\boldsymbol{C}^{-1}\right)^{-1}\left(\frac{\sigma_1^2}{s_{r1}^4}\boldsymbol{I}^{-1}+\frac{\sigma_2^2}{s_{r2}^4}\boldsymbol{C}^{-1}\right)\right. \tag{2.18}$$

$$\left.\left(s_{r1}^{-2}\boldsymbol{I}^{-1}+s_{r2}^{-2}\boldsymbol{C}^{-1}\right)^{-1}\right)<\sigma_1^2\boldsymbol{I}. \tag{2.19}$$

Similarly, from Equation 2.13 we obtain:

$$\boldsymbol{E}\left(\left(s_{r1}^{-2}\boldsymbol{I}^{-1}+s_{r2}^{-2}\boldsymbol{C}^{-1}\right)^{-1}\left(\frac{\sigma_1^2}{s_{r1}^4}\boldsymbol{I}^{-1}+\frac{\sigma_2^2}{s_{r2}^4}\boldsymbol{C}^{-1}\right)\right. \tag{2.20}$$

$$\left.\left(s_{r1}^{-2}\boldsymbol{I}^{-1}+s_{r2}^{-2}\boldsymbol{C}^{-1}\right)^{-1}\right) < \sigma_2^2\boldsymbol{C}. \tag{2.21}$$

By combining Equation 2.18 and Equation 2.20, we have:

$$\boldsymbol{E}\left(\left(s_{r1}^{-2}\boldsymbol{I}^{-1}+s_{r2}^{-2}\boldsymbol{C}^{-1}\right)^{-1}\left(\frac{\sigma_1^2}{s_{r1}^4}\boldsymbol{I}^{-1}+\frac{\sigma_2^2}{s_{r2}^4}\boldsymbol{C}^{-1}\right)\right. \tag{2.22}$$

$$\left.\left(s_{r1}^{-2}\boldsymbol{I}^{-1}+s_{r2}^{-2}\boldsymbol{C}^{-1}\right)^{-1}\right) < min\left(\sigma_1^2\boldsymbol{I},\sigma_2^2\boldsymbol{C}\right). \tag{2.23}$$

Note that Equation 2.22 exhibits a pattern of the comparison of the GDE against individual estimators based on two diagonal matrices viz., identity matrix and the C matrix. This is exactly the same formulation as in Lemma 5 above. Hence the result follows by an application of Lemma 5. □

2.5 General k-dim Common Parameter Involving p Independent Groups

Next we consider the general case of p (> 2) independent groups of linear regression models sharing a k-dim common estimable parameter vector $\boldsymbol{\theta}$. The k-dim GDE of p groups is:

$$\hat{\boldsymbol{\theta}}_{GD|p} = \left(\sum_{i=1}^{p} s_{ri}^{-2}\boldsymbol{W_i}^{-1}\right)^{-1}\left(\sum_{i=1}^{p} s_{ri}^{-2}\boldsymbol{W_i}^{-1}\hat{\boldsymbol{\theta}}_i\right).$$

Again, it is easy to show $\hat{\boldsymbol{\theta}}_{GD|p}$ is an unbiased estimator of $\boldsymbol{\theta}$, as:

$$\boldsymbol{E}(\hat{\boldsymbol{\theta}}_{GD|p}) \tag{2.24}$$

$$= \boldsymbol{E}\left(\boldsymbol{E}\left(\hat{\boldsymbol{\theta}}_{GD|p}|s_{r1}^2, s_{r2}^2, \ldots, s_{rp}^2\right)\right) \tag{2.25}$$

$$= \boldsymbol{E}\left(\boldsymbol{E}\left(\left(\sum_{i=1}^{p} s_{ri}^{-2}\boldsymbol{W_i}^{-1}\right)^{-1}\left(\sum_{i=1}^{p} s_{ri}^{-2}\boldsymbol{W_i}^{-1}\hat{\boldsymbol{\theta}}_i\right)|s_{r1}^2, s_{r2}^2, \ldots, s_{rp}^2\right)\right) \tag{2.26}$$

$$= \boldsymbol{E}\left(\left(\sum_{i=1}^{p} s_{ri}^{-2}\boldsymbol{W_i}^{-1}\right)^{-1}\left(\sum_{i=1}^{p} s_{ri}^{-2}\boldsymbol{W_i}^{-1}\boldsymbol{E}(\hat{\boldsymbol{\theta}}_i|s_{r1}^2, s_{r2}^2, \ldots, s_{rp}^2)\right)\right) \tag{2.27}$$

$$= \boldsymbol{E}\left(\left(\sum_{i=1}^{p} s_{ri}^{-2}\boldsymbol{W_i}^{-1}\right)^{-1}\left(\sum_{i=1}^{p} s_{ri}^{-2}\boldsymbol{W_i}^{-1}\boldsymbol{\theta}\right)\right) \tag{2.28}$$

$$= \boldsymbol{\theta}.$$

It has the following dispersion matrix:

$$\boldsymbol{D}(\hat{\boldsymbol{\theta}}_{GD|p}) = \boldsymbol{E}\left(\boldsymbol{D}\left(\hat{\boldsymbol{\theta}}_{GD|p}|s_{r1}^2, s_{r2}^2, \ldots, s_{rp}^2\right)\right) + \boldsymbol{D}\left(\boldsymbol{E}\left(\hat{\boldsymbol{\theta}}_{GD|p}|s_{r1}^2, s_{r2}^2, \ldots, s_{rp}^2\right)\right) \tag{2.29}$$

$$= \boldsymbol{E}\left(\boldsymbol{D}\left(\hat{\boldsymbol{\theta}}_{GD|p}|s_{r1}^2, s_{r2}^2, \ldots, s_{rp}^2\right)\right) \tag{2.30}$$

$$= \boldsymbol{E}\left(\left(\sum_{i=1}^{p} s_{ri}^{-2}\boldsymbol{W_i}^{-1}\right)^{-1}\left(\sum_{i=1}^{p} \frac{\sigma_i^2}{s_{ri}^4}\boldsymbol{W_i}^{-1}\right)\left(\sum_{i=1}^{p} s_{ri}^{-2}\boldsymbol{W_i}^{-1}\right)^{-1}\right).$$

Lemma 7. *If $\boldsymbol{W_i}$'s are diagonal matrices $i = 1, 2, \ldots, p$, then $\boldsymbol{D}(\hat{\boldsymbol{\theta}}_{GD|p}) < \min_{i\in\{1,\ldots,p\}}\left(\sigma_i^2\boldsymbol{W_i}\right)$, for all values of σ_i^2, if and only if condition in (2a) or (2b) of Theorem 4 holds:*

1. *$\nu_i > 9$, $i = 1, 2, \ldots, p$ or*

2. $\nu_i = 9$ *for some* i, *and* $\nu_j > 17$ *for* $i, j \in \{1, 2, \ldots, p\}$ *and each* $j \neq i$.

Moreover, if condition in (2a) or (2b) of Theorem 4 is satisfied, the Loewner order of dispersion matrix, $\boldsymbol{D}(\hat{\boldsymbol{\theta}}_{GD|p})$ *is below* $\boldsymbol{D}(\hat{\boldsymbol{\theta}}_{GD|q})$ *of any* $q(< p)$ *subgroups.*

Proof.

This is the extension of Lemma 5. In case that $\boldsymbol{W_i}$ (i=1,2,...,p)'s are diagonal matrices, that can be decomposed into k single GDE of p groups. Our claim follows from an application of Theorem 4. □

If not all $\boldsymbol{W_i}$'s are diagonal matrices, $i \in \{1, 2, \ldots, p\}$, we have the following Theorem.

Theorem 8. *Suppose there exists a nonsingular matrix* $\boldsymbol{P}$, *such that* $\boldsymbol{W_i^{-1}} = \boldsymbol{P}\boldsymbol{C_i}^{-1}\boldsymbol{P}^T$, *where* $\boldsymbol{C_i}$*'s are diagonal matrices for all* $i = 1, 2, \ldots, p$. *Then the Loewner order of* $\boldsymbol{D}(\hat{\boldsymbol{\theta}}_{GD|p})$ *is below* $\boldsymbol{D}(\boldsymbol{\theta}_{GD|q})$ *of any* $q(< p)$ *subgroups, for all values of* σ_i^2, *if and only if condition in (2a) or (2b) of Theorem 4 holds:*

1. $\nu_i > 9$, $i = 1, 2, \ldots, p$ *or*

2. $\nu_i = 9$ *for some* i, *and* $\nu_j > 17$ *for* $i, j \in \{1, 2, \ldots, p\}$ *and each* $j \neq i$.

Proof.

We need to examine the validity of

$$\boldsymbol{D}(\hat{\boldsymbol{\theta}}_{GD|p}) < \boldsymbol{D}(\hat{\boldsymbol{\theta}}_{GD|q}), q < p.$$

Equivalently,

$$\boldsymbol{E}\left(\left(\sum_{i=1}^{p} s_{ri}^{-2}\boldsymbol{W_i}^{-1}\right)^{-1}\left(\sum_{i=1}^{p}\frac{\sigma_i^2}{s_{ri}^4}\boldsymbol{W_i}^{-1}\right)\left(\sum_{i=1}^{p} s_{ri}^{-2}\boldsymbol{W_i}^{-1}\right)^{-1}\right)$$
$$<\boldsymbol{E}\left(\left(\sum_{i=1}^{q} s_{ri}^{-2}\boldsymbol{W_i}^{-1}\right)^{-1}\left(\sum_{i=1}^{q}\frac{\sigma_i^2}{s_{ri}^4}\boldsymbol{W_i}^{-1}\right)\left(\sum_{i=1}^{q} s_{ri}^{-2}\boldsymbol{W_i}^{-1}\right)^{-1}\right).$$

Since $\boldsymbol{W_i^{-1}} = \boldsymbol{P}\boldsymbol{C_i^{-1}}\boldsymbol{P}^T$,then

$$\boldsymbol{E}\left(\left(\sum_{i=1}^{p} s_{ri}^{-2}\boldsymbol{P}\boldsymbol{C_i}^{-1}\boldsymbol{P}^T\right)^{-1}\left(\sum_{i=1}^{p}\frac{\sigma_i^2}{s_{ri}^4}\boldsymbol{P}\boldsymbol{C_i}^{-1}\boldsymbol{P}^T\right)\left(\sum_{i=1}^{p} s_{ri}^{-2}\boldsymbol{P}\boldsymbol{C_i}^{-1}\boldsymbol{P}^T\right)^{-1}\right)$$
$$<\boldsymbol{E}\left(\left(\sum_{i=1}^{q} s_{ri}^{-2}\boldsymbol{P}\boldsymbol{C_i}^{-1}\boldsymbol{P}^T\right)^{-1}\left(\sum_{i=1}^{q}\frac{\sigma_i^2}{s_{ri}^4}\boldsymbol{P}\boldsymbol{C_i}^{-1}\boldsymbol{P}^T\right)\left(\sum_{i=1}^{q} s_{ri}^{-2}\boldsymbol{P}\boldsymbol{C_i}^{-1}\boldsymbol{P}^T\right)^{-1}\right).$$

Upon simplification:

$$\boldsymbol{E}\left(\boldsymbol{P}^{T-1}\left(\sum_{i=1}^{p} s_{ri}^{-2}\boldsymbol{C_i}^{-1}\right)^{-1}\boldsymbol{P}^{-1}\boldsymbol{P}\left(\sum_{i=1}^{p}\frac{\sigma_i^2}{s_{ri}^4}\boldsymbol{C_i}^{-1}\right)\boldsymbol{P}^T\boldsymbol{P}^{T-1}\right.$$
$$\left.\left(\sum_{i=1}^{p} s_{ri}^{-2}\boldsymbol{C_i}^{-1}\right)^{-1}\boldsymbol{P}^{-1}\right)<\boldsymbol{E}\left(\boldsymbol{P}^{T-1}\left(\sum_{i=1}^{q} s_{ri}^{-2}\boldsymbol{C_i}^{-1}\right)^{-1}\boldsymbol{P}^{-1}\boldsymbol{P}\right.$$
$$\left.\left(\sum_{i=1}^{q}\frac{\sigma_i^2}{s_{ri}^4}\boldsymbol{C_i}^{-1}\right)\boldsymbol{P}^T\boldsymbol{P}^{T-1}\left(\sum_{i=1}^{q} s_{ri}^{-2}\boldsymbol{C_i}^{-1}\right)^{-1}\boldsymbol{P}^{-1}\right).$$

This reduces to:

$$\boldsymbol{E}\left(\boldsymbol{P}^{-1T}\left(\sum_{i=1}^{p} s_{ri}^{-2}\boldsymbol{C_i}^{-1}\right)^{-1}\left(\sum_{i=1}^{p}\frac{\sigma_i^2}{s_{ri}^4}\boldsymbol{C_i}^{-1}\right)\left(\sum_{i=1}^{p} s_{ri}^{-2}\boldsymbol{C_i}^{-1}\right)^{-1}\boldsymbol{P}^{-1}\right)$$
$$<\boldsymbol{E}\left(\boldsymbol{P}^{-1T}\left(\sum_{i=1}^{q} s_{ri}^{-2}\boldsymbol{C_i}^{-1}\right)^{-1}\left(\sum_{i=1}^{q}\frac{\sigma_i^2}{s_{ri}^4}\boldsymbol{C_i}^{-1}\right)\left(\sum_{i=1}^{q} s_{ri}^{-2}\boldsymbol{C_i}^{-1}\right)^{-1}\boldsymbol{P}^{-1}\right),$$

which requires

$$\boldsymbol{E}\left(\left(\sum_{i=1}^{p} s_{ri}^{-2}\boldsymbol{C_i}^{-1}\right)^{-1}\left(\sum_{i=1}^{p}\frac{\sigma_i^2}{s_{ri}^4}\boldsymbol{C_i}^{-1}\right)\left(\sum_{i=1}^{p} s_{ri}^{-2}\boldsymbol{C_i}^{-1}\right)^{-1}\right)$$
$$< \boldsymbol{E}\left(\left(\sum_{i=1}^{q} s_{ri}^{-2}\boldsymbol{C_i}^{-1}\right)^{-1}\left(\sum_{i=1}^{q}\frac{\sigma_i^2}{s_{ri}^4}\boldsymbol{C_i}^{-1}\right)\left(\sum_{i=1}^{q} s_{ri}^{-2}\boldsymbol{C_i}^{-1}\right)^{-1}\right).$$

Since $\boldsymbol{C_i}'s$ are diagonal matrices, the results follows by an application of Lemma 7. □

Remark 4. *Generally, the existence of such a nonsingular matrix $\boldsymbol{P}$ that diagonalizes all $\boldsymbol{W_i}$ simultaneously is not guaranteed. However, Corollary 9 below provides a special case.*

Suppose in the most general representation of the linear regression model described in the beginning of Section 2, $\boldsymbol{Z}_{i(n_i\times t_i)}$ does not exist, which indicates that all these p independent groups of linear regression models are following the same linear regression model. In such a case, $\boldsymbol{W}_{i(k\times k)} = (\boldsymbol{X}_i^T\boldsymbol{X}_i)^{-1}$.

Corollary 9. *Suppose $\boldsymbol{Z}_{i(n_i\times t_i)}$'s do not exist. For p independent groups of linear regression models sharing a 2-dim common estimable parameter $\boldsymbol{\theta} = \binom{\theta_1}{\theta_2}$, if $\frac{\left(\sum_{j=1}^{n_i} x_{i1j}x_{i2j}\right)}{\left(\sum_{j=1}^{n_i} x_{i1j}^2\right)}$=constant, $i = 1, 2, \ldots, p$, then a necessary and sufficient condition for the Loewner order of $\boldsymbol{D}(\hat{\boldsymbol{\theta}}_{GD|p})$ to be below $\boldsymbol{D}(\hat{\boldsymbol{\theta}}_{GD|q})$ for any $q(< p)$ subgroups, for all values of σ_i^2, is that the condition in (2a) or (2b) of Theorem 4 holds.*

Proof.

From the linear regression theory, we know:

$$\boldsymbol{W_i}^{-1} = \begin{pmatrix} \sum_{j=1}^{n_i} x_{i1j}^2 & \sum_{j=1}^{n_i} x_{i1j}x_{i2j} \\ \sum_{j=1}^{n_i} x_{i1j}x_{i2j} & \sum_{j=1}^{n_i} x_{i2j}^2 \end{pmatrix}$$

$$= \begin{pmatrix} 1 & 0 \\ \frac{\left(\sum_{j=1}^{n_i} x_{i1j}x_{i2j}\right)}{\left(\sum_{j=1}^{n_i} x_{i1j}^2\right)} & 1 \end{pmatrix} \begin{pmatrix} \sum_{j=1}^{n_i} x_{i1j}^2 & 0 \\ 0 & \sum_{j=1}^{n_i} x_{i2j}^2 - \frac{\left(\sum_{j=1}^{n_i} x_{i1j}x_{i2j}\right)^2}{\left(\sum_{j=1}^{n_i} x_{i1j}^2\right)} \end{pmatrix} \begin{pmatrix} 1 & \frac{\left(\sum_{j=1}^{n_i} x_{i1j}x_{i2j}\right)}{\left(\sum_{j=1}^{n_i} x_{i1j}^2\right)} \\ 0 & 1 \end{pmatrix}.$$

If $\frac{\left(\sum_{j=1}^{n_i} x_{i1j}x_{i2j}\right)}{\left(\sum_{j=1}^{n_i} x_{i1j}^2\right)} = b$, a constant, then let

$$\boldsymbol{P} = \begin{pmatrix} 1 & 0 \\ \frac{\left(\sum_{j=1}^{n_i} x_{i1j}x_{i2j}\right)}{\left(\sum_{j=1}^{n_i} x_{i1j}^2\right)} & 1 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ b & 1 \end{pmatrix}.$$

Hence from Theorem 8, the result follows. □

Remark 5. *When this 2-dim common estimable parameter of interest* $\boldsymbol{\theta} = \binom{\theta_1}{\theta_2}$ *contains intercept, viz.* θ_1 *, then it asks* $\sum_{j=1}^{n_i} x_{i1j}x_{i2j} / \sum_{j=1}^{n_i} x_{i1j}^2 = \bar{x}_{i2\cdot}$*, which is the sample mean, to be constant.*

Remark 6. *The ratio* $\left(\sum_{j=1}^{n_i} x_{i1j}x_{i2j}\right) / \left(\sum_{j=1}^{n_i} x_{i1j}^2\right) = \frac{\|\vec{x}_{i2\cdot} \cos\theta\|}{\|\vec{x}_{i1\cdot}\|}$*, where* θ *is the angle between two covariate variable vectors* $\vec{x}_{i1\cdot}$ *and* $\vec{x}_{i2\cdot}$*. This suggests that, typically in non-intercept linear models, if we can pre-select our* $\vec{x}_{i1\cdot}$ *and* $\vec{x}_{i2\cdot}$ *to make this ratio constant, then we can obtain a* *more efficient estimator of this common parameter of interest by utilizing all p models, as long as we can collect enough observations for each linear model.*

2.6 Conclusion and Discussion

We investigated the properties of GDE in higher dimension. We established that GDE is still an unbiased estimator for the vector parameter of interest. We also found that the condition (2a) or (2b) in Theorem 4 (condition (1) or (2) in Lemma 5 for groups of two) continued to hold when estimating a k-dim common parameter vector for p independent groups of linear regression models. Consequently the GDE of k-dim common parameters by combining these p groups provides a better and more efficient estimator.

In the linear regression model that we have studied here, we tacitly assumed that the regression parameters are fixed and unknown. In the literature there are studies on what are called 'random coefficient regression models', such as Carter and Yang (9) and Liski, Luoma and Sinha (28). We may postulate a model with fixed unknown α parameter (the intercept) but the coefficients are random. The problem of estimation of the common mean α in such scenarios is a rather routine exercise. We propose to examine the domination results in such scenarios in a subsequent communication.

CHAPTER 3

PERFORMANCE OF THE GRAYBILLDEAL ESTIMATOR VIA PITMAN CLOSENESS CRITERION

[1] Pitman closeness criterion is a coverage probability-based criterion to examine the relative performance of estimators. Commonly, the performance of the standard Graybill-Deal estimator of the common mean has been examined with respect to the mean squared error (variance). In this chapter we examined its performance with respect to the Pitman closeness criterion. Specifically, we compared a p-source based Graybill-Deal estimator against its q-source based competitors for q $(< p)$-dimensional subsets of p-dimensional data.

3.1 Introduction

In chapter 3, we still considered the problem of estimation of the common mean shared by several independent normal populations with unknown and most likely unequal variances. Generally, we had p independent sources with distributions $N(\mu, \sigma_i^2)$, $i = 1, 2, \ldots, p$. Also, let $\bar{x}_i$ and s_i^2 be sample mean and sample variance, respectively; and $s_{\bar{x}_i}^2 = s_i^2/n_i$, $\sigma_{\bar{x}_i}^2 = \sigma_i^2/n_i$, where n_i was the sample size available from the i^{th} group, $i = 1, 2, \ldots, p$. We have showed in chapter 2 that, $\hat{\mu}_{GD|p}$ was preferable to $\hat{\mu}_{GD|q}$ in the criteria of mean square error, if and only if

---

[1]Part of this chapter is coming from a working paper "Performance of the GraybillDeal Estimator via Pitman Closeness Criterion" from Nie, Sinha and Hedayat.

1. either $n_i > 10$ for $i = 1, 2, \ldots, p$,

2. or $n_i = 10$ for some i, and $n_j > 18$ ($i, j = 1, 2, \ldots, p$ for each $j \neq i$).

Here we intended to compare $\hat{\mu}_{GD|p}$ with $\hat{\mu}_{GD|q}$ by employing Pitman closeness criterion, which was introduced by Pitman (36):

We say estimator $\hat{\mu}_1$ is better (closer) than $\hat{\mu}_2$ for the estimation of the parameter μ if and only if $P\{|\hat{\mu}_1 - \mu| \leq |\hat{\mu}_2 - \mu|\} \geq 1/2$.

Although Robert, Hwang and Strawderman (42) pointed some drawbacks of Pitman closeness, especially its lack of transitivity. In other words, there exist three estimates t_1, t_2 and t_3 of μ, such that t_1 is Pitman closer than t_2 and t_2 is Pitman closer than t_3, but t_1 is not Pitman closer than t_3. Other views, supporting it or against it, were presented by Blyth (7), Casella and Wells (10), Ghosh, Keating and Sen (18), Peddada (35) and Rao (40). But as Rao (40) wrote: "There are indeed situations where any given method of estimation can produce counter intuitive conclusions", and he believed "that the performance of an estimator should be examined under different criteria to under stand the nature of the estimator and possibly to provide information to the decision maker."

Kubokawa (25) and Sarkar (43) established that :

Theorem 10. *The necessary and sufficient condition for* $P\{|\hat{\mu}_{GD|2} - \mu| \leq |\bar{x}_i - \mu|\} \geq 1/2$ *to hold uniformly in* $(\mu, \sigma^2_{\bar{x}_1}, \sigma^2_{\bar{x}_2})$ *is that* $m_i = n_i - 1 \geq 4$ *for each* $i = 1, 2$.

Proof.

See Kubokawa (25) and Sarkar (43). $\square$

Sarkar (43) further established that:

Theorem 11. *$P\{|\hat{\mu}_{GD|p} - \mu| \leq |\bar{x}_i - \mu|\} \geq 1/2$ for all $i = 1, 2, \ldots, p$ and uniformly in $(\mu, \sigma^2_{\bar{x}_1}, \sigma^2_{\bar{x}_2}, \ldots, \sigma^2_{\bar{x}_p})$, if and only if*

$$2E\left\{ (\sum_{j=1, j\neq i}^{p} \sigma_{\bar{x}_j}^{-2} Y_j^2)^{-1/2} (\sum_{j=1, j\neq i}^{p} \sigma_{\bar{x}_j}^{-2} Y_j) \right\} \leq E\left\{ (\sum_{j=1, j\neq i}^{p} \sigma_{\bar{x}_j}^{-2} Y_j^2)^{1/2} \right\}, \tag{3.1}$$

holds for any $i = 1, 2, \ldots, p$ and uniformly in $(\sigma^2_{\bar{x}_1}, \sigma^2_{\bar{x}_2}, \ldots, \sigma^2_{\bar{x}_p})$ for all $p - 1$ subgroups, where Y_j's are independently distributed as $\frac{m_j}{\chi^2(m_j)}$ and $m_j = n_j - 1$, for $j = 1, 2, \ldots, p$.

Proof.

See Sarkar (43). □

Sarkar (43) also discussed the sample size $n_i = m_i + 1$ requirement for *Equation* 3.1, and showed that:

Theorem 12. *Let $Y_1, \ldots, Y_s$ be independent random variables such that $Y_k \sim \frac{m_k}{\chi^2(m_k)}$, for $k = 1, 2, \ldots, s$, then for any $\theta_k > 0$ $(k = 1, 2, \ldots, s)$*

$$2E\left\{ (\sum_{j=1}^{s} \theta_j Y_j^2)^{-1/2} (\sum_{j=1}^{s} \theta_j Y_j) \right\} \leq E\left\{ (\sum_{j=1}^{s} \theta_j Y_j^2)^{1/2} \right\}, \tag{3.2}$$

if

$$1 - 8m_k^{-1} + 4(\sum_{j=1}^{s} m_j)^{-1} \geq 0, \tag{3.3}$$

for all k in $\{1, 2, \ldots, s\}$.

Proof.

See Sarkar (43). $\square$

In addition, Sarkar (43) observed the following

1. For $s = 1$, $m_1 \geq 4$ is also the necessary condition.
2. When $s = 2$, Condition (Equation 3.3) is true if either $m_k = 6$ for all k, or $m_k = 7$ for some k and $7 \leq m_j \leq 21$ for $j \neq k$, or $m_k \geq 8$ for all k in $\{1, 2\}$.
3. When $s = 3$, Condition (Equation 3.3) is true if either $m_k = 7$ for some k and $14 \leq \sum_{j=1,j\neq k}^{3} m_j \leq 21$ for $j \neq k$, or $m_k \geq 8$ for all k in $\{1, 2, s = 3\}$.
4. When $s = 4$, Condition (Equation 3.3) requires $m_k = 7$, or ≥ 8 for all k in $\{1, 2, \ldots, s = 4\}$.
5. When $s \geq 5$, Condition (Equation 3.3) demands $m_k \geq 8$ for all k in $\{1, 2, \ldots, s\}$.

In the following sections we intended to show that $\hat{\mu}_{GD|p}$ is a closer estimator than $\hat{\mu}_{GD|q}$, the GDE of any q $(< p)$ subgroups, in the sense of Pitman closeness criterion. Without loss of generality, our results were based on comparing with the initial q subgroups, and for any other q subgroups, we would obtain similar results. Several sufficient conditions [including the Equation 3.1 (suitably modified) which was to hold uniformly in $(\sigma^2_{\bar{x}_1}, \sigma^2_{\bar{x}_2}, \ldots, \sigma^2_{\bar{x}_p})$ for all q subgroups] were provided. A necessary condition was also provided.

3.2 Problem Settings and Lemmas

Before introducing our main results, we first introduce the following notations and lemmas.

We borrow the definition from Sarkar (43):

Let

$$Y_i = \sigma_{\bar{x}_i}^2 s_{\bar{x}_i}^{-2} \sim \frac{m_i}{\chi^2(m_i)},\ for\ i = 1, 2, \ldots, p,$$

where Y_i's are independently distributed and $m_i = n_i - 1$ is the degrees of freedom ($i = 1, 2, \ldots, p$). Obviously we have $E(Y_i^{-1}) = 1$.

Let $p = q + r$ and define:

$$U_q = (\sum_{i=1}^{q} \sigma_{\bar{x}_i}^2 s_{\bar{x}_i}^{-4})^{-1/2} (\sum_{i=1}^{q} s_{\bar{x}_i}^{-2}) = (\sum_{i=1}^{q} \sigma_{\bar{x}_i}^{-2} Y_i^2)^{-1/2} (\sum_{i=1}^{q} \sigma_{\bar{x}_i}^{-2} Y_i), \tag{3.4}$$

$$U_r = (\sum_{i=q+1}^{q+r} \sigma_{\bar{x}_i}^2 s_{\bar{x}_i}^{-4})^{-1/2} (\sum_{i=q+1}^{q+r} s_{\bar{x}_i}^{-2}) = (\sum_{i=q+1}^{q+r} \sigma_{\bar{x}_i}^{-2} Y_i^2)^{-1/2} (\sum_{i=q+1}^{q+r} \sigma_{\bar{x}_i}^{-2} Y_i), \tag{3.5}$$

$$V_q = (\sum_{i=1}^{q} \sigma_{\bar{x}_i}^2 s_{\bar{x}_i}^{-4})^{-1/2} = (\sum_{i=1}^{q} \sigma_{\bar{x}_i}^{-2} Y_i^2)^{-1/2}, \tag{3.6}$$

$$V_r = (\sum_{i=q+1}^{q+r} \sigma_{\bar{x}_i}^2 s_{\bar{x}_i}^{-4})^{-1/2} = (\sum_{i=q+1}^{q+r} \sigma_{\bar{x}_i}^{-2} Y_i^2)^{-1/2}. \tag{3.7}$$

(U_q, V_q) and (U_r, V_r) are mutually independent. As we mentioned before, U_q and V_q represented the initial q subgroups, and hence U_r and V_r represented the remaining subgroups.

Lemma 13.

$$\sum_{i=1}^{s} \sigma_{\bar{x}_i}^{-2} \geq U_s^2 \geq \min_{i=1,2,\ldots,s} \{\sigma_{\bar{x}_i}^{-2}\}$$

holds uniformly in $(\sigma_{\bar{x}_1}^2, \sigma_{\bar{x}_2}^2, \ldots, \sigma_{\bar{x}_s}^2)$.

Proof. Due to Cauchy-Schwarz inequality, we have:

$$\begin{aligned} U_s^2 &= (\sum_{i=1}^{s} \sigma_{\bar{x}_i}^{-2} Y_i^2)^{-1} (\sum_{i=1}^{s} \sigma_{\bar{x}_i}^{-2} Y_i)^2 \\ &\leq \sum_{i=1}^{s} \sigma_{\bar{x}_i}^{-2}. \end{aligned}$$

On the other side,

$$\begin{aligned} U_s^2 &= (\sum_{i=1}^{s} \sigma_{\bar{x}_i}^{-2} Y_i^2)^{-1} (\sum_{i=1}^{s} \sigma_{\bar{x}_i}^{-2} Y_i)^2 \\ &\geq (\sum_{i=1}^{s} \sigma_{\bar{x}_i}^{-2} Y_i^2)^{-1} (\sum_{i=1}^{s} \sigma_{\bar{x}_i}^{-4} Y_i^2) \\ &= \sum_{j=1}^{s} \frac{\sigma_{\bar{x}_j}^{-2} Y_j^2}{\sum_{i=1}^{s} \sigma_{\bar{x}_i}^{-2} Y_i^2} \sigma_{\bar{x}_i}^{-2} \\ &\geq \min_{i=1,2,\ldots,s} \{\sigma_{\bar{x}_i}^{-2}\}. \end{aligned}$$

□

Lemma 14. *The inequality* $E\{U_s^{-1}\} \geq (\sum_{i=1}^{s} \sigma_{\bar{x}_i}^{-2})^{-1/2} \geq E\{V_s\}$ *holds uniformly in* $(\sigma_{\bar{x}_1}^2, \sigma_{\bar{x}_2}^2, \ldots, \sigma_{\bar{x}_s}^2)$.

Proof. Due to the fact:

$$E\{U_s^{-1}\} = E\{(\sum_{i=1}^{s} \sigma_{\bar{x}_i}^{-2} Y_i^2)^{1/2} (\sum_{i=1}^{s} \sigma_{\bar{x}_i}^{-2} Y_i)^{-1}\} \tag{3.8}$$

$$\geq (\sum_{i=1}^{s} \sigma_{\bar{x}_i}^{-2})^{-1/2} \tag{3.9}$$

$$= (\sum_{i=1}^{s} \sigma_{\bar{x}_i}^{-2})^{-1/2} E\{\sum_{j=1}^{s} \frac{\sigma_{\bar{x}_j}^{-2}}{\sum_{i=1}^{s} \sigma_{\bar{x}_i}^{-2}} Y_j^{-1}\} \tag{3.10}$$

$$= (\sum_{i=1}^{s} \sigma_{\bar{x}_i}^{-2})^{-1/2} E\{\sum_{j=1}^{s} \frac{\sigma_{\bar{x}_j}^{-2}}{\sum_{i=1}^{s} \sigma_{\bar{x}_i}^{-2}} (Y_j^2)^{-1/2}\} \tag{3.11}$$

$$\geq (\sum_{i=1}^{s} \sigma_{\bar{x}_i}^{-2})^{-1/2} E\{\sum_{j=1}^{s} \frac{\sigma_{\bar{x}_j}^{-2}}{\sum_{i=1}^{s} \sigma_{\bar{x}_i}^{-2}} Y_j^2\}^{-1/2} \tag{3.12}$$

$$= E\{\sum_{j=1}^{s} \sigma_{\bar{x}_j}^{-2} Y_j^2\}^{-1/2} \tag{3.13}$$

$$= E\{V_s\}. \tag{3.14}$$

The Equation 3.9 is based on Lemma 13, and Equation 3.12 is due to Jensen's inequality. □

Lemma 15. *The inequality* $(\sum_{i=1}^{s} \sigma_{\bar{x}_i}^{-2}) E(U_s^{-1}) \leq E(V_s^{-1})$ *holds uniformly in* $(\sigma_{\bar{x}_1}^2, \sigma_{\bar{x}_2}^2, \ldots, \sigma_{\bar{x}_s}^2)$

Proof.

See Sarkar (43). □

Lemma 16. *The probability* $P\{|\hat{\mu}_{GD|p} - \mu| \leq |\hat{\mu}_{GD|q} - \mu|\}$ *is*

$1/\pi \left(E\{\arctan(U_r U_q^{-1})\} + E\{\arctan(U_r U_q^{-1} + 2 V_r V_q^{-1})\} \right)$.

Proof.

Without loss of generality, it is clear μ can be assumed to be $\mu = 0$.

We need to compute the probability:

$$P\{|\hat{\mu}_{GD|p}| \leq |\hat{\mu}_{GD|q}|\}. \tag{3.15}$$

By applying the fact that:

$$\hat{\mu}_{GD|p} = \frac{(\sum_{i=1}^{q} s_{\bar{x}_i}^{-2})\hat{\mu}_{GD|q} + (\sum_{i=q+1}^{q+r} s_{\bar{x}_i}^{-2})\hat{\mu}_{GD|r}}{\sum_{i=1}^{q+r} s_{\bar{x}_i}^{-2}}, \tag{3.16}$$

the probability in *Equation* 3.15 can be written as:

$$P\{|\hat{\mu}_{GD|p}| \leq |\hat{\mu}_{GD|q}|\}$$

$$= P\{(\hat{\mu}_{GD|p})^2 \leq (\hat{\mu}_{GD|q})^2\} \tag{3.17}$$

$$= P\{(\frac{(\sum_{i=1}^{q} s_{\bar{x}_i}^{-2})\hat{\mu}_{GD|q} + (\sum_{i=q+1}^{q+r} s_{\bar{x}_i}^{-2})\hat{\mu}_{GD|r}}{\sum_{i=1}^{q+r} s_{\bar{x}_i}^{-2}})^2 \leq (\hat{\mu}_{GD|q})^2\} \tag{3.18}$$

$$= P\{((\sum_{i=1}^{q} s_{\bar{x}_i}^{-2})\hat{\mu}_{GD|q} + (\sum_{i=q+1}^{q+r} s_{\bar{x}_i}^{-2})\hat{\mu}_{GD|r})^2 \leq ((\sum_{i=1}^{q+r} s_{\bar{x}_i}^{-2})\hat{\mu}_{GD|q})^2\} \tag{3.19}$$

$$= 2P\{\hat{\mu}_{GD|q}(1 + 2\frac{\sum_{i=1}^{q} s_{\bar{x}_i}^{-2}}{\sum_{i=q+1}^{q+r} s_{\bar{x}_i}^{-2}}) \leq \hat{\mu}_{GD|r} \leq \hat{\mu}_{GD|q}, \hat{\mu}_{GD|q} > 0\} \tag{3.20}$$

$$\begin{aligned} = P\Bigg\{ &-\left(\frac{(\sum_{i=1}^{q} \sigma_{\bar{x}_i}^2 s_{\bar{x}_i}^{-4})^{1/2}(\sum_{i=1}^{q} s_{\bar{x}_i}^{-2})^{-1}}{(\sum_{i=q+1}^{q+r} \sigma_{\bar{x}_i}^2 s_{\bar{x}_i}^{-4})^{1/2}(\sum_{i=q+1}^{q+r} s_{\bar{x}_i}^{-2})^{-1}} + 2\frac{(\sum_{i=1}^{q} \sigma_{\bar{x}_i}^2 s_{\bar{x}_i}^{-4})^{1/2}}{(\sum_{i=q+1}^{q+r} \sigma_{\bar{x}_i}^2 s_{\bar{x}_i}^{-4})^{1/2}}\right) \\ &\leq \left(\left(\frac{\hat{\mu}_{GD|r}}{(\sum_{i=q+1}^{q+r} \sigma_{\bar{x}_i}^2 s_{\bar{x}_i}^{-4})^{1/2}(\sum_{i=q+1}^{q+r} s_{\bar{x}_i}^{-2})^{-1}}\right) / \left(\frac{\hat{\mu}_{GD|q}}{(\sum_{i=1}^{q} \sigma_{\bar{x}_i}^2 s_{\bar{x}_i}^{-4})^{1/2}(\sum_{i=1}^{q} s_{\bar{x}_i}^{-2})^{-1}}\right)\right) \\ &\leq \frac{(\sum_{i=1}^{q} \sigma_{\bar{x}_i}^2 s_{\bar{x}_i}^{-4})^{1/2}(\sum_{i=1}^{q} s_{\bar{x}_i}^{-2})^{-1}}{(\sum_{i=q+1}^{q+r} \sigma_{\bar{x}_i}^2 s_{\bar{x}_i}^{-4})^{1/2}(\sum_{i=q+1}^{q+r} s_{\bar{x}_i}^{-2})^{-1}} \Bigg\} \end{aligned} \tag{3.21}$$

$$= P\left\{ -(U_r U_q^{-1} + 2V_r V_q^{-1}) \leq \frac{\frac{\hat{\mu}_{GD|r}}{\left(\sum_{i=q+1}^{q+r} \sigma_{\bar{x}_i}^2 s_{\bar{x}_i}^{-4}\right)^{1/2}\left(\sum_{i=q+1}^{q+r} s_{\bar{x}_i}^{-2}\right)^{-1}}}{\frac{\hat{\mu}_{GD|q}}{\left(\sum_{i=1}^{q} \sigma_{\bar{x}_i}^2 s_{\bar{x}_i}^{-4}\right)^{1/2}\left(\sum_{i=1}^{q} s_{\bar{x}_i}^{-2}\right)^{-1}}} \leq U_r U_q^{-1} \right\}. \tag{3.22}$$

Since conditionally given $s_{\bar{x}_i}^{-2}$ for $i = 1, \ldots, p$,

$$\frac{\dfrac{\hat{\mu}_{GD|r}}{\left(\sum_{i=q+1}^{q+r}\sigma_{\bar{x}_i}^2 s_{\bar{x}_i}^{-4}\right)^{1/2}\left(\sum_{i=q+1}^{q+r} s_{\bar{x}_i}^{-2}\right)^{-1}}}{\dfrac{\hat{\mu}_{GD|q}}{\left(\sum_{i=1}^{q}\sigma_{\bar{x}_i}^2 s_{\bar{x}_i}^{-4}\right)^{1/2}\left(\sum_{i=1}^{q} s_{\bar{x}_i}^{-2}\right)^{-1}}}$$

follows a Cauchy distribution. The probability in *Equation* 3.22, can be denoted as

$$\begin{aligned}
\gamma &= \gamma(\sigma_{\bar{x}_1}^2, \sigma_{\bar{x}_2}^2, \ldots, \sigma_{\bar{x}_p}^2) && (3.23)\\
&= 1/\pi \left(E\{\arctan(U_r U_q^{-1})\} - E\{\arctan(-U_r U_q^{-1} - 2V_r V_q^{-1})\}\right) && (3.24)\\
&= 1/\pi \left(E\{\arctan(U_r U_q^{-1})\} + E\{\arctan(U_r U_q^{-1} + 2V_r V_q^{-1})\}\right). && (3.25)
\end{aligned}$$

□

3.3 Preferable of Combining: A Necessary Condition

Theorem 17.

$$P\{|\hat{\mu}_{GD|p} - \mu| \leq |\hat{\mu}_{GD|q} - \mu|\} \geq 1/2 \tag{3.26}$$

holds uniformly in $(\mu, \sigma_{\bar{x}_1}^2, \sigma_{\bar{x}_2}^2, \ldots, \sigma_{\bar{x}_p}^2)$, *only if*

$$\begin{aligned}
&2E\left\{(\sum_{i=q+1}^{p}\sigma_{\bar{x}_i}^{-2}Y_i^2)^{-1/2}(\sum_{i=q+1}^{p}\sigma_{\bar{x}_i}^{-2}Y_i)\right\}E\left\{(\sum_{i=1}^{q}\sigma_{\bar{x}_i}^{-2}Y_i^2)^{1/2}(\sum_{i=1}^{q}\sigma_{\bar{x}_i}^{-2}Y_i)^{-1}\right\} \leq\\
&E\left\{(\sum_{i=q+1}^{p}\sigma_{\bar{x}_i}^{-2}Y_i^2)^{1/2}\right\}E\left\{(\sum_{i=1}^{q}\sigma_{\bar{x}_i}^{-2}Y_i^2)^{-1/2}\right\}, \qquad (3.27)
\end{aligned}$$

which is equivalent to stating:

$$2E\{U_r\}E\{U_q^{-1}\} \geq E\{V_q\}E\{V_r^{-1}\}, \tag{3.28}$$

holds uniformly in $(\sigma^2_{\bar{x}_1}, \sigma^2_{\bar{x}_2}, \ldots, \sigma^2_{\bar{x}_p})$.

Proof. According to Lemma 16, the probability $P\{|\hat{\mu}_{GD|p} - \mu| \leq |\hat{\mu}_{GD|q} - \mu|\}$ can be denoted as:

$$\begin{aligned} \gamma &= \gamma(\sigma^2_{\bar{x}_1}, \sigma^2_{\bar{x}_2}, \ldots, \sigma^2_{\bar{x}_p}) \\ &= 1/\pi \left(E\{\arctan(U_r U_q^{-1})\} + E\{\arctan(U_r U_q^{-1} + 2V_r V_q^{-1})\} \right). \end{aligned} \tag{3.29}$$

The expectations are taken with respect to the independent random variables Y_i's.

Let $\sigma^{-2}_{\bar{x}_i} = \sigma^{-2}_{\bar{x}_1}\tau_i$, for $i = 1, 2, \ldots, q$; and $\sigma^{-2}_{\bar{x}_i} = \sigma^2_{\bar{x}_1}\tau_i$, for $i = q+1, q+2, \ldots, q+r$. Here τ_i's are in the range of $[0, \infty)$, except $\tau_1 = 1$.[1]

So $U_q = \sigma^{-1}_{\bar{x}_1} U_q^*$, $V_q = \sigma_{\bar{x}_1} V_q^*$, $U_r = \sigma_{\bar{x}_1} U_r^*$, and $V_r = \sigma^{-1}_{\bar{x}_1} V_r^*$, where

$$\begin{aligned} U_q^* &= (\sum_{i=1}^{q} \tau_i Y_i^2)^{-1/2} (\sum_{i=1}^{q} \tau_i Y_i), \qquad V_q^* = (\sum_{i=1}^{q} \tau_i Y_i^2)^{-1/2}, \\ U_r^* &= (\sum_{i=q+1}^{q+r} \tau_i Y_i^2)^{-1/2} (\sum_{i=q+1}^{q+r} \tau_i Y_i), V_r^* = (\sum_{i=q+1}^{q+r} \tau_i Y_i^2)^{-1/2}. \end{aligned} \tag{3.30}$$

---

[1]Since in the inequality we are going to prove, $2E\{U_r\}E\{U_q^{-1}\} \geq E\{V_q\}E\{V_r^{-1}\}$, $U_\star$'s and $V_\star^{-1}$'s are proportional to $\sigma_\star^{-2}$.

The probability in Equation 3.29 will be

$$
\begin{aligned}
\gamma =& \gamma(\sigma^2_{\bar{x}_1}, \tau_1, \tau_2, \ldots, \tau_p) \\
=& 1/\pi \left(E\{\arctan(\sigma^2_{\bar{x}_1} U^*_r (U^*_q)^{-1})\} \right. \\
& \left. + E\{\arctan(\sigma^2_{\bar{x}_1} U^*_r (U^*_q)^{-1} + 2\sigma^{-2}_{\bar{x}_1} V^*_r (V^*_q)^{-1})\} \right). \qquad (3.31)
\end{aligned}
$$

Obviously, the dominated convergence theorem implies that:

$$
\gamma(\sigma^2_{\bar{x}_1}) \to 1/\pi(E\{\arctan 0\} + E\{\arctan \infty\}) = 1/2 \ as \ \sigma^2_{\bar{x}_1} \to 0.
$$

Next, we are going to show that $\gamma \geq 1/2$ uniformly in $(\sigma^2_{\bar{x}_1}, \sigma^2_{\bar{x}_2}, \ldots, \sigma^2_{\bar{x}_p})$, only if

$$
2E\{U_r\}E\{U_q^{-1}\} \geq E\{V_q\}E\{V_r^{-1}\} \qquad (3.32)
$$

holds uniformly in $(\sigma^2_{\bar{x}_1}, \sigma^2_{\bar{x}_2}, \ldots, \sigma^2_{\bar{x}_p})$.

The derivative of γ with respect to $\sigma^2_{\bar{x}_1}$, is given by

$$
\begin{aligned}
& \frac{\partial}{\partial \sigma^2_{\bar{x}_1}} \pi\gamma \\
&= E\left\{ \frac{U^*_r (U^*_q)^{-1} - 2\sigma^{-4}_{\bar{x}_1} V^*_r (V^*_q)^{-1}}{1 + \left(\sigma^2_{\bar{x}_1} U^*_r (U^*_q)^{-1} + 2\sigma^{-2}_{\bar{x}_1} V^*_r (V^*_q)^{-1}\right)^2} \right\} + E\left\{ \frac{U^*_r (U^*_q)^{-1}}{1 + \sigma^4_{\bar{x}_1} U^{*2}_r (U^*_q)^{-2}} \right\} \qquad (3.33) \\
&= E\left\{ \frac{U^*_r (U^*_q)^{-1} \sigma^4_{\bar{x}_1} - 2V^*_r (V^*_q)^{-1}}{\sigma^4_{\bar{x}_1} \left(1 + \sigma^4_{\bar{x}_1} U^{*2}_r (U^*_q)^{-2}\right) + 4V^*_r (V^*_q)^{-1} \left(\sigma^4_{\bar{x}_1} U^*_r (U^*_q)^{-1} + V^*_r (V^*_q)^{-1}\right)} \right\} \\
&\quad + E\left\{ \frac{U^*_r (U^*_q)^{-1}}{1 + \sigma^4_{\bar{x}_1} U^{*2}_r (U^*_q)^{-2}} \right\}. \qquad (3.34)
\end{aligned}
$$

To check whether the inequality:

$$E\left\{-\frac{1}{2}V_q^*(V_r^*)^{-1}+U_r^*(U_q^*)^{-1})\right\}\geq 0 \tag{3.35}$$

is a necessary condition, we assume that Equation 3.35 is not true for some $(\tau_2,\ldots,\tau_p)$. Then we notice that

$$\frac{\partial}{\partial\sigma^2_{\bar{x}_1}}\gamma|_{\sigma^2_{\bar{x}_1}=0}=1/\pi E\left\{-\frac{1}{2}V_q^*(V_r^*)^{-1}+U_r^*(U_q^*)^{-1})\right\}<0,$$

which leads to γ being non-decreasing and contradicts that $\gamma\geq 1/2$ for $\sigma^2_{\bar{x}_1}$ in a small region near 0.

It is easy to verify that Equation 3.35 is equivalent to Equation 3.32. Hence we prove the theorem. □

3.4 Preferable of Combining: Sufficient Conditions

Theorem 18. *If*

$$E\left\{\frac{-2^{-1}}{\sigma^4_{\bar{x}_1}U_r^*(U_q^*)^{-1}+V_r^*(V_q^*)^{-1}}\right\}+E\left\{\frac{U_r^*(U_q^*)^{-1}}{1+\sigma^4_{\bar{x}_1}U_r^{*2}(U_q^*)^{-2}}\right\}\geq 0, \tag{3.36}$$

then $P\{|\hat{\mu}_{GD|p}-\mu|\leq|\hat{\mu}_{GD|q}-\mu|\}\geq 1/2$ *holds uniformly in* $(\mu,\sigma^2_{\bar{x}_1},\sigma^2_{\bar{x}_2},\ldots,\sigma^2_{\bar{x}_p})$.

Proof.

Continue from *Equation* 3.34, we get the following:

$$\frac{\partial}{\partial\sigma^2_{\bar{x}_1}}\pi\gamma$$

$$
\begin{aligned}
&= E\left\{\frac{U_r^*(U_q^*)^{-1}\sigma_{\bar{x}_1}^4 - 2V_r^*(V_q^*)^{-1}}{\sigma_{\bar{x}_1}^4\left(1+\sigma_{\bar{x}_1}^4 U_r^{*2}(U_q^*)^{-2}\right) + 4V_r^*(V_q^*)^{-1}\left(\sigma_{\bar{x}_1}^4 U_r^*(U_q^*)^{-1} + V_r^*(V_q^*)^{-1}\right)}\right\} \\
&\quad + E\left\{\frac{U_r^*(U_q^*)^{-1}}{1+\sigma_{\bar{x}_1}^4 U_r^{*2}(U_q^*)^{-2}}\right\} \\
&\geq E\left\{\frac{-2V_r^*(V_q^*)^{-1}}{\sigma_{\bar{x}_1}^4\left(1+\sigma_{\bar{x}_1}^4 U_r^{*2}(U_q^*)^{-2}\right) + 4V_r^*(V_q^*)^{-1}\left(\sigma_{\bar{x}_1}^4 U_r^*(U_q^*)^{-1} + V_r^*(V_q^*)^{-1}\right)}\right\} \\
&\quad + E\left\{\frac{U_r^*(U_q^*)^{-1}}{1+\sigma_{\bar{x}_1}^4 U_r^{*2}(U_q^*)^{-2}}\right\} \qquad (3.37)\\
&\geq E\left\{\frac{-2V_r^*(V_q^*)^{-1}}{4V_r^*(V_q^*)^{-1}\left(\sigma_{\bar{x}_1}^4 U_r^*(U_q^*)^{-1} + V_r^*(V_q^*)^{-1}\right)}\right\} + E\left\{\frac{U_r^*(U_q^*)^{-1}}{1+\sigma_{\bar{x}_1}^4 U_r^{*2}(U_q^*)^{-2}}\right\} \\
&= E\left\{\frac{-2^{-1}}{\sigma_{\bar{x}_1}^4 U_r^*(U_q^*)^{-1} + V_r^*(V_q^*)^{-1}}\right\} + E\left\{\frac{U_r^*(U_q^*)^{-1}}{1+\sigma_{\bar{x}_1}^4 U_r^{*2}(U_q^*)^{-2}}\right\}.
\end{aligned}
$$

Hence if

$$
E\left\{\frac{-2^{-1}}{\sigma_{\bar{x}_1}^4 U_r^*(U_q^*)^{-1} + V_r^*(V_q^*)^{-1}}\right\} + E\left\{\frac{U_r^*(U_q^*)^{-1}}{1+\sigma_{\bar{x}_1}^4 U_r^{*2}(U_q^*)^{-2}}\right\} \geq 0,
$$

then $\frac{\partial}{\partial \sigma_{\bar{x}_1}^2}\pi\gamma \geq 0$, which leads to $\gamma(\sigma_{\bar{x}_1}^2, \tau_2, \ldots, \tau_p)$ is non-decreasing in $\sigma_{\bar{x}_1}^2$ for any of $(\tau_2, \ldots, \tau_p)$. The dominated convergence theorem implies that $\gamma(\sigma_{\bar{x}_1}^2) \to 1/\pi(E\{\arctan 0\} + E\{\arctan \infty\}) = 1/2$ as $\sigma_{\bar{x}_1}^2 \to 0$. For any other finite $(\tau_2, \ldots, \tau_p)$ and $\sigma_{\bar{x}_1}^2$, we have $\gamma(\sigma_{\bar{x}_1}^2, \tau_2, \ldots, \tau_p) \geq \gamma(0, \tau_2, \ldots, \tau_p) = 1/2$. At the boundary where one or some τ_i's go to ∞ or 0, $\gamma(\sigma_{\bar{x}_1}^2, \tau_2, \ldots, \tau_p) = 1$ *or* $1/2 \geq 1/2$. Hence $\gamma \geq 1/2$ holds uniformly in $(\sigma_{\bar{x}_1}^2, \sigma_{\bar{x}_2}^2, \ldots, \sigma_{\bar{x}_p}^2)$.

□

Based on Theorem 18, we derive several sufficient conditions in the sequel for $P\{|\hat{\mu}_{GD|p} - \mu| \leq |\hat{\mu}_{GD|q} - \mu|\} \geq 1/2$ to hold uniformly in $(\mu, \sigma_{\bar{x}_1}^2, \sigma_{\bar{x}_2}^2, \ldots, \sigma_{\bar{x}_p}^2)$.

Corollary 19.

$$2E(U_r)E(U_q^{-1}) \geq \frac{E\left(U_q\right)}{\min_{i=1,\ldots,q}(\sigma_{\bar{x}_i}^{-2})}E(V_r^{-1})$$

is a sufficient condition for $P\{|\hat{\mu}_{GD|p}-\mu| \leq |\hat{\mu}_{GD|q}-\mu|\} \geq 1/2$ *to hold uniformly in* $(\mu, \sigma_{\bar{x}_1}^2, \sigma_{\bar{x}_2}^2, \ldots, \sigma_{\bar{x}_p}^2)$.

Proof.

It is easy to state that

$$2E(U_r)E(U_q^{-1}) \geq \frac{E\left(U_q\right)}{\min_{i=1,\ldots,q}(\sigma_{\bar{x}_i}^{-2})}E(V_r^{-1})$$

is equivalent to

$$2E(U_r^*)E(U_q^{*-1}) \geq \frac{E\left(U_q^*\right)}{\min_{i=1,\ldots,q}(\tau_i)}E(V_r^{*-1}).$$

From Theorem 18, we only need to show the validity of Equation 3.36.

At the left side of Equation 3.36, we have the following inequality for the second term:

$$E\left\{\frac{U_r^*(U_q^*)^{-1})}{1+\sigma_{\bar{x}_1}^4 U_r^{*2}(U_q^*)^{-2}}\right\} \geq E\left\{\frac{U_r^*(U_q^*)^{-1})}{1+\sigma_{\bar{x}_1}^4(\sum_{i=q+1}^{q+r}\tau_i)(\min_{i=1,\ldots,q}(\tau_i))^{-1}}\right\}. \tag{3.38}$$

This is because of $U_r^{*2} \leq \sum_{i=q+1}^{q+r}\tau_i$ and $(U_q^*)^{-2} \leq (\min_{i=1,\ldots,q}(\tau_i))^{-1}$ following by Lemma 13.

We also have:

$$\left(1+\sigma_{\bar{x}_1}^4(\sum_{i=q+1}^{q+r}\tau_i)(\min_{i=1,\ldots,q}(\tau_i))^{-1}\right) E\left\{\{\sigma_{\bar{x}_1}^4 U_r^*(U_q^*)^{-1} + V_r^*(V_q^*)^{-1}\}^{-1}\right\}$$

$$\leq \left(1+\frac{\sigma_{\bar{x}_1}^4(\sum_{i=q+1}^{q+r}\tau_i)}{\min_{i=1,\ldots,q}(\tau_i)}\right)^{-1} E\left\{\frac{\sigma_{\bar{x}_1}^4(\sum_{i=q+1}^{q+r}\tau_i)^2}{(\min_{i=1,\ldots,q}(\tau_i))^2}(U_r^*)^{-1}(U_q^*)+V_q^*(V_r^*)^{-1}\right\} \tag{3.39}$$

$$\leq max\left\{\frac{E\left(U_q^*\right)}{\min_{i=1,\ldots,q}(\tau_i)}(\sum_{i=q+1}^{q+r}\tau_i)E(U_r^{*-1}), E(V_q^*)E(V_r^{*-1})\right\}. \tag{3.40}$$

The Equation 3.39 is from Jensen's inequality:

$$(1+a)E(aA+B)^{-1} \leq (1+a)^{-1}E(aA^{-1}+B^{-1}),$$

where $a=\frac{\sigma_{\bar{x}_1}^4(\sum_{i=q+1}^{q+r}\tau_i)}{\min_{i=1,\ldots,q}(\tau_i)}$, $A=\frac{\min_{i=1,\ldots,q}(\tau_i)}{(\sum_{i=q+1}^{q+r}\tau_i)}U_r^*(U_q^*)^{-1}$, and $B=V_r^*(V_q^*)^{-1}$.

From Lemma 15, we know

$$(\sum_{i=q+1}^{q+r}\tau_i)E(U_r^{*-1}) \leq E(V_r^{*-1}). \tag{3.41}$$

From Lemma 14 and the fact that $\frac{U_q^*}{\min_{i=1,\ldots,q}(\tau_i)} \geq U_q^{*-1}$, we note

$$\frac{E\left(U_q^*\right)}{\min_{i=1,\ldots,q}(\tau_i)} \geq E(U_q^{*-1}) \geq E(V_q^*). \tag{3.42}$$

So we have the following :

$$max\left\{\frac{E\left(U_q^*\right)}{\min_{i=1,\ldots,q}(\tau_i)}(\sum_{i=q+1}^{q+r}\tau_i)E(U_r^{*-1}), E(V_q^*)E(V_r^{*-1})\right\}$$
$$\leq \frac{E\left(U_q^*\right)}{\min_{i=1,\ldots,q}(\tau_i)}E(V_r^{*-1}).$$

Then on the left side of Equation 3.36, we have the following inequality for the first term:

$$
\begin{aligned}
&E\left\{\{\sigma^4_{\bar{x}_1}U^*_r(U^*_q)^{-1}+V^*_r(V^*_q)^{-1}\}^{-1}\right\} \leq \left(1+\frac{\sigma^4_{\bar{x}_1}(\sum_{i=q+1}^{q+r}\tau_i)}{\min_{i=1,\ldots,q}(\tau_i)}\right)^{-1} \\
&max\left\{\frac{E\left(U^*_q\right)}{\min_{i=1,\ldots,q}(\tau_i)}(\sum_{i=q+1}^{q+r}\tau_i)E(U^{*-1}_r), E(V^*_q)E(V^{*-1}_r)\right\} \\
&\leq \left(1+\frac{\sigma^4_{\bar{x}_1}(\sum_{i=q+1}^{q+r}\tau_i)}{\min_{i=1,\ldots,q}(\tau_i)}\right)^{-1}\frac{E\left(U^*_q\right)}{\min_{i=1,\ldots,q}(\tau_i)}E(V^{*-1}_r)
\end{aligned}
$$

Combining the above and Equation 3.38, we have

$$
\begin{aligned}
&\frac{\partial}{\partial\sigma^2_{\bar{x}_1}}\pi\gamma \geq E\left\{\frac{-2^{-1}}{\sigma^4_{\bar{x}_1}U^*_r(U^*_q)^{-1}+V^*_r(V^*_q)^{-1}}\right\}+E\left\{\frac{U^*_r(U^*_q)^{-1})}{1+\sigma^4_{\bar{x}_1}U^{*2}_r(U^*_q)^{-2}}\right\} \\
&\geq\left\{-2^{-1}\frac{E\left(U^*_q\right)}{\min_{i=1,\ldots,q}(\tau_i)}E(V^{*-1}_r)+E\left\{U^*_r(U^*_q)^{-1}\right\}\right\}\left(1+\frac{\sigma^4_{\bar{x}_1}(\sum_{i=q+1}^{q+r}\tau_i)}{\min_{i=1,\ldots,q}(\tau_i)}\right)^{-1}. \qquad (3.43)
\end{aligned}
$$

If

$$
-2^{-1}\frac{E\left(U^*_q\right)}{\min_{i=1,\ldots,q}(\tau_i)}E(V^{*-1}_r)+E\left\{U^*_r(U^*_q)^{-1}\right\}\geq 0,
$$

then

$$
\frac{\partial}{\partial\sigma^2_{\bar{x}_1}}\pi\gamma \geq 0.
$$

Therefore a sufficient condition for $P\{|\hat{\mu}_{GD|p}-\mu|\leq|\hat{\mu}_{GD|q}-\mu|\}\geq 1/2$ to hold uniformly in $(\mu,\sigma^2_{\bar{x}_1},\sigma^2_{\bar{x}_2},\ldots,\sigma^2_{\bar{x}_p})$ is:

$$
2E(U_r)E(U^{-1}_q)\geq\frac{E\left(U_q\right)}{\min_{i=1,\ldots,q}(\sigma^{-2}_{\bar{x}_i})}E(V^{-1}_r). \qquad (3.44)
$$

$\square$

Another corollary is given in the following:

Corollary 20.

$$2E(U_r)E(U_q) \geq \sqrt{\sum_{i=1}^{q} \sigma_{\bar{x}_i}^{-2}} E(V_r^{-1}),$$

or

$$2E(U_r)E(U_q) \geq E(V_q^{-1})E(V_r^{-1}).$$

is a sufficient condition for $P\{|\hat{\mu}_{GD|p} - \mu| \leq |\hat{\mu}_{GD|q} - \mu|\} \geq 1/2$ *to hold uniformly in* $(\mu, \sigma_{\bar{x}_1}^2, \sigma_{\bar{x}_2}^2, \ldots, \sigma_{\bar{x}_p}^2)$.

Proof.

We need to validate Equation 3.36. On the left side of Equation 3.36, we have the following inequality for the second term:

$$E\left\{\frac{U_r^*(U_q^*)^{-1}}{1 + \sigma_{\bar{x}_1}^4 U_r^{*2}(U_q^*)^{-2}}\right\} \geq E\left\{\frac{U_r^* U_q^*}{\sum_{i=1}^{q} \tau_i + \sigma_{\bar{x}_1}^4 (\sum_{i=q+1}^{q+r} \tau_i)}\right\}. \tag{3.45}$$

it is because of $U_r^{*2} \leq \sum_{i=q+1}^{q+r} \tau_i$ and $U_q^{*2} \leq \sum_{i=1}^{q} \tau_i$ following from Lemma 13.

We also have:

$$
\begin{aligned}
&(\sum_{i=1}^{q}\tau_i+\sigma_{\bar{x}_1}^4(\sum_{i=q+1}^{q+r}\tau_i))E\left\{\{\sigma_{\bar{x}_1}^4U_r^*(U_q^*)^{-1}+V_r^*(V_q^*)^{-1}\}^{-1}\right\}\\
\leq&(\sum_{i=1}^{q}\tau_i+\sigma_{\bar{x}_1}^4(\sum_{i=q+1}^{q+r}\tau_i))^{-1}E\left\{\sigma_{\bar{x}_1}^4(\sum_{i=q+1}^{q+r}\tau_i)^2(U_r^*)^{-1}(U_q^*)+(\sum_{i=1}^{q}\tau_i)^2V_q^*(V_r^*)^{-1}\right\} \quad (3.46)\\
\leq&max\left\{(\sum_{i=q+1}^{q+r}\tau_i)E(U_r^*)^{-1}E(U_q^*),(\sum_{i=1}^{q}\tau_i)E(V_q^*)E(V_r^*)^{-1}\right\}. \quad (3.47)
\end{aligned}
$$

The Equation 3.46 is due to Jensen's inequality.

So on the left side of Equation 3.36, we have the following inequality for the first term:

$$
\begin{aligned}
&E\left\{\{\sigma_{\bar{x}_1}^4U_r^*(U_q^*)^{-1}+V_r^*(V_q^*)^{-1}\}^{-1}\right\}\leq\\
&(\sum_{i=1}^{q}\tau_i+\sigma_{\bar{x}_1}^4(\sum_{i=q+1}^{q+r}\tau_i))^{-1}max\left\{(\sum_{i=q+1}^{q+r}\tau_i)E(U_r^*)^{-1}E(U_q^*),(\sum_{i=1}^{q}\tau_i)E(V_q^*)E(V_r^*)^{-1}\right\}. \quad (3.48)
\end{aligned}
$$

From Lemma 15 we have:

$$
(\sum_{i=q+1}^{q+r}\tau_i)E(U_r^{*-1})\leq E(V_r^{*-1}). \quad (3.49)
$$

The upper bound of $E(U_q^*)$ and $(\sum_{i=1}^{q}\tau_i)E(V_q^*)$ is $\sqrt{\sum_{i=1}^{q}\tau_i}$. This is due to Lemma 13 and Lemma 14, respectively.

Another upper bound of $(\sum_{i=1}^{q}\tau_i)E(V_q^*)$ is $E(V_q^{*-1})$. Since

$$(\sum_{i=1}^{q}\tau_i)E(V_q^*) = \frac{(\sum_{i=1}^{q}\tau_i)E(U_q^{*-1})E(V_q^*)}{E(U_q^{*-1})} \leq \frac{E(V_q^{*-1})E(V_q^*)}{E(U_q^{*-1})} \leq E(V_q^{*-1}), \tag{3.50}$$

then we have

$$\begin{aligned} &max\left\{(\sum_{i=q+1}^{q+r}\tau_i)E(U_r^*)^{-1}E(U_q^*),(\sum_{i=1}^{q}\tau_i)E(V_q^*)E(V_r^*)^{-1}\right\} \leq \\ &\sqrt{\sum_{i=1}^{q}\tau_i E(V_r^{*-1})}, \end{aligned} \tag{3.51}$$

or

$$\begin{aligned} &max\left\{(\sum_{i=q+1}^{q+r}\tau_i)E(U_r^*)^{-1}E(U_q^*),(\sum_{i=1}^{q}\tau_i)E(V_q^*)E(V_r^*)^{-1}\right\} \leq \\ &max\left\{E(V_q^{*-1}),E(U_q^*)\right\}E(V_r^{*-1}) = E(V_q^{*-1})E(V_r^{*-1}). \end{aligned} \tag{3.52}$$

Combining Equation 3.45, Equation 3.48 and the above, we have the following:

$$\begin{aligned} \frac{\partial}{\partial\sigma_{\bar{x}_1}^2}\pi\gamma &\geq E\left\{\frac{-2^{-1}}{\sigma_{\bar{x}_1}^4 U_r^*(U_q^*)^{-1}+V_r^*(V_q^*)^{-1}}\right\}+E\left\{\frac{U_r^*(U_q^*)^{-1})}{1+\sigma_{\bar{x}_1}^4 U_r^{*2}(U_q^*)^{-2}}\right\} \\ &\geq \left\{-2^{-1}\sqrt{\sum_{i=1}^{q}\tau_i E(V_r^{*-1})}+E\left(U_r^*\right)E(U_q^*)\right\}(\sum_{i=1}^{q}\tau_i+\sigma_{\bar{x}_1}^4(\sum_{i=q+1}^{q+r}\tau_i))^{-1}, \end{aligned} \tag{3.53}$$

or

$$\frac{\partial}{\partial \sigma_{\bar{x}_1}^2}\pi\gamma \geq E\left\{\frac{-2^{-1}}{\sigma_{\bar{x}_1}^4 U_r^*(U_q^*)^{-1} + V_r^*(V_q^*)^{-1}}\right\} + E\left\{\frac{U_r^*(U_q^*)^{-1})}{1+\sigma_{\bar{x}_1}^4 U_r^{*2}(U_q^*)^{-2}}\right\}$$
$$\geq \left\{-2^{-1}E(V_q^{*-1})E(V_r^{*-1}) + E\left(U_r^*)E(U_q^*\right)\right\}(\sum_{i=1}^{q}\tau_i + \sigma_{\bar{x}_1}^4(\sum_{i=q+1}^{q+r}\tau_i))^{-1}. \tag{3.54}$$

So if

$$-2^{-1}\sqrt{\sum_{i=1}^{q}\tau_i}E(V_r^{*-1}) + E\left(U_r^*)E(U_q^*\right) \geq 0,$$

or

$$-2^{-1}E(V_q^{*-1})E(V_r^{*-1}) + E\left(U_r^*)E(U_q^*\right) \geq 0,$$

then

$$\frac{\partial}{\partial \sigma_{\bar{x}_1}^2}\pi\gamma \geq 0$$

Hence it is equivalent saying that a sufficient condition for $P\{|\hat{\mu}_{GD|p} - \mu| \leq |\hat{\mu}_{GD|q} - \mu|\} \geq 1/2$ to hold uniformly in $(\mu, \sigma_{\bar{x}_1}^2, \sigma_{\bar{x}_2}^2, \ldots, \sigma_{\bar{x}_p}^2)$ is:

$$2E(U_r)E(U_q) \geq \sqrt{\sum_{i=1}^{q}\sigma_{\bar{x}_i}^{-2}}E(V_r^{-1}), \tag{3.55}$$

or

$$2E(U_r)E(U_q) \geq E(V_q^{-1})E(V_r^{-1}). \tag{3.56}$$

□

3.5 Sample Size Discussion

In this section we will discuss the sample size requirement based on the condition provided in Corollary 20:

$$2E(U_r)E(U_q) \geq E(V_q^{-1})E(V_r^{-1}),$$

Which is the same as:

$$2E\left\{(\sum_{i=1}^{q}\theta_i Y_i)(\sum_{i=1}^{q}\theta_i Y_i^2)^{-\frac{1}{2}}\right\}E\left\{(\sum_{i=q+1}^{q+r}\theta_i Y_i)(\sum_{i=q+1}^{q+r}\theta_i Y_i^2)^{-\frac{1}{2}}\right\}$$
$$\geq E\left(\sum_{i=1}^{q}\theta_i Y_i^2\right)^{\frac{1}{2}} E\left(\sum_{i=q+1}^{q+r}\theta_i Y_i^2\right)^{\frac{1}{2}},$$

$\theta_i = \sigma_{\bar{x}_i}^{-2}$, $i = 1, 2, \ldots, p$. We have the following theorem:

Theorem 21. *Let $Y_1, \ldots, Y_p$ be independent random variables such that $Y_i \sim \frac{m_i}{\chi^2(m_i)}$, for $i = 1, 2, \ldots, p$,and let $p = q + r$. Then for any $\theta_i > 0$ ($i = 1, 2, \ldots, p$),*

$$2E\left\{(\sum_{i=1}^{q}\theta_i Y_i)(\sum_{i=1}^{q}\theta_i Y_i^2)^{-\frac{1}{2}}\right\}E\left\{(\sum_{i=q+1}^{q+r}\theta_i Y_i)(\sum_{i=q+1}^{q+r}\theta_i Y_i^2)^{-\frac{1}{2}}\right\}$$
$$\geq E\left(\sum_{i=1}^{q}\theta_i Y_i^2\right)^{\frac{1}{2}} E\left(\sum_{i=q+1}^{q+r}\theta_i Y_i^2\right)^{\frac{1}{2}} \tag{3.57}$$

if there exists $1 < b < 2$, such that

$$m_i^{-1}(m_i - 4) + 2(\sum_{j=1}^{q} m_j)^{-1} \geq \frac{1}{b} \tag{3.58}$$

for all $i = 1, 2, \ldots, q$, and

$$m_i^{-1}(m_i - 4) + 2(\sum_{j=q+1}^{q+r} m_j)^{-1} \geq \frac{b}{2} \tag{3.59}$$

for all $i = q+1, q+2, \ldots, q+r = p$.

Proof.

Define:

$$g_1(Y_i) = E\left\{ Y_i(\sum_{j=1}^{q} \theta_j Y_j^2)^{-\frac{1}{2}} | Y_i \right\}, \tag{3.60}$$

for $i = 1, 2, \ldots, q$; and define:

$$g_2(Y_i) = E\left\{ Y_i(\sum_{j=q+1}^{q+r} \theta_j Y_j^2)^{-\frac{1}{2}} | Y_i \right\}, \tag{3.61}$$

for $i = q+1, q+2, \ldots, q+r = p$.

From Sarkar (43) and Haff (20), we know the following for $i = 1, 2, \ldots, q$:

$$m_i E(g_1(Y_i)) = (m_i - 2)E(Y_i g_1(Y_i)) - 2E(Y_i^2 {g'}_1(Y_i)), \tag{3.62}$$

and similarly for $i = q+1, q+2, \ldots, q+r = p$, we have:

$$m_i E(g_2(Y_i)) = (m_i - 2)E(Y_i g_2(Y_i)) - 2E(Y_i^2 {g'}_2(Y_i)). \tag{3.63}$$

Then,

$$g'_1(Y_i) = E\left\{(\sum_{j=1}^{q}\theta_j Y_j^2)^{-\frac{1}{2}}|Y_i\right\} - \theta_i E\left\{Y_i^2(\sum_{j=1}^{q}\theta_j Y_j^2)^{-\frac{3}{2}}|Y_i\right\}, \tag{3.64}$$

and

$$g'_2(Y_i) = E\left\{(\sum_{j=q+1}^{q+r}\theta_j Y_j^2)^{-\frac{1}{2}}|Y_i\right\} - \theta_i E\left\{Y_i^2(\sum_{j=q+1}^{q+r}\theta_j Y_j^2)^{-\frac{3}{2}}|Y_i\right\}. \tag{3.65}$$

For $g_1(Y_i)$, we have that:

$$\begin{aligned}
& m_i E\left\{Y_i(\sum_{j=1}^{q}\theta_j Y_j^2)^{-\frac{1}{2}}\right\} \\
=& m_i E(g_1(Y_i)) \\
=& (m_i - 2)E(Y_i g_1(Y_i)) - 2E(Y_i^2 {g'}_1(Y_i)) \\
=& (m_i - 4)E\left\{Y_i^2(\sum_{j=1}^{q}\theta_j Y_j^2)^{-\frac{1}{2}}\right\} + 2\theta_i E\left\{Y_i^4(\sum_{j=1}^{q}\theta_j Y_j^2)^{-\frac{3}{2}}\right\}.
\end{aligned} \tag{3.66}$$

Based on the above,we see the following

$$E\left\{(\sum_{i=1}^{q}\theta_i Y_i)(\sum_{i=1}^{q}\theta_i Y_i^2)^{-\frac{1}{2}}\right\}$$

$$
\begin{aligned}
&= \sum_{i=1}^{q} \theta_i E \left\{ Y_i (\sum_{j=1}^{q} \theta_j Y_j^2)^{-\frac{1}{2}} \right\} \\
&= E \left\{ (\sum_{i=1}^{q} m_i^{-1}(m_i - 4)\theta_i Y_i^2)(\sum_{i=1}^{q} \theta_i Y_i^2)^{-\frac{1}{2}} \right\} + \\
&\quad 2E \left\{ (\sum_{i=1}^{q} m_i^{-1} \theta_i^2 Y_i^4)(\sum_{i=1}^{q} \theta_i Y_i^2)^{-\frac{3}{2}} \right\}. \qquad (3.67)
\end{aligned}
$$

From Cauchy-Schwarz inequality, we can prove:

$$
\sum_{i=1}^{q} m_i^{-1} \theta_i^2 Y_i^4 \geq (\sum_{i=1}^{q} m_i)^{-1} (\sum_{i=1}^{q} \theta_i Y_i^2)^2. \qquad (3.68)
$$

Applying *Equation* 3.68 into Equation 3.67, we have the following:

$$
\begin{aligned}
&E \left\{ (\sum_{i=1}^{q} \theta_i Y_i)(\sum_{i=1}^{q} \theta_i Y_i^2)^{-\frac{1}{2}} \right\} \\
\geq &E \left\{ \left[\sum_{i=1}^{q} \left(m_i^{-1}(m_i - 4) + 2(\sum_{j=1}^{q} m_j)^{-1} \right) \theta_i Y_i^2 \right] (\sum_{i=1}^{q} \theta_i Y_i^2)^{-\frac{1}{2}} \right\}. \qquad (3.69)
\end{aligned}
$$

Similarly for $g_2(Y_i)$, we obtain:

$$
\begin{aligned}
&E \left\{ (\sum_{i=q+1}^{q+r} \theta_i Y_i)(\sum_{q+i=1}^{q+r} \theta_i Y_i^2)^{-\frac{1}{2}} \right\} \\
\geq &E \left\{ \left[\sum_{i=q+1}^{q+r} \left(m_i^{-1}(m_i - 4) + 2(\sum_{j=q+1}^{q+r} m_j)^{-1} \right) \theta_i Y_i^2 \right] (\sum_{i=q+1}^{q+r} \theta_i Y_i^2)^{-\frac{1}{2}} \right\}. \qquad (3.70)
\end{aligned}
$$

Thus, Equation 3.57 is true if there exists $1 < b < 2$, such that

$$m_i^{-1}(m_i - 4) + 2(\sum_{j=1}^{q} m_j)^{-1} \geq \frac{1}{b}$$

for all $i = 1, 2, \ldots, q$, and

$$m_i^{-1}(m_i - 4) + 2(\sum_{j=q+1}^{q+r} m_j)^{-1} \geq \frac{b}{2}$$

for all $i = q+1, q+2, \ldots, q+r = p$. $\square$

Remark 7. *Observe that Equation* 3.58 *and Equation* 3.59 *are respectively equivalent to:*

$$(1 - \frac{1}{b}) \left\{ m_{min} \left(m_{min} - \frac{2}{1 - \frac{1}{b}} \right) + (\sum_{j=1, j \neq min}^{q} m_j) \left(m_{min} - \frac{4}{1 - \frac{1}{b}} \right) \right\} \geq 0 \tag{3.71}$$

and

$$(1 - \frac{b}{2}) \left\{ m_{min} \left(m_{min} - \frac{2}{1 - \frac{b}{2}} \right) + (\sum_{j=q+1, j \neq min}^{q+r} m_j) \left(m_{min} - \frac{4}{1 - \frac{b}{2}} \right) \right\} \geq 0. \tag{3.72}$$

In order to have solutions for Equation 3.71 and Equation 3.72, we need $1 < b < 2$.

Remark 8. *One possible choice is to let* $b = \sqrt{2}$. *Check the Condition* (*Equation* 3.71),

1. *When* $q = 1$, *Condition (Equation* 3.71*) is true if* $m_1 \geq \dfrac{4}{2 - \sqrt{2}} \approx 7$.

2. *When* $q = 2$, *Condition (Equation* 3.71*) is true if either*

(a) $m_{min} = 11$, and the other $11 \leq m_{j,j\neq min} \leq 17$.

(b) $m_{min} = 12$, and the other $12 \leq m_{j,j\neq min} \leq 37$.

(c) $m_{min} = 13$, and the other $13 \leq m_{j,j\neq min} \leq 122$.

(d) $m_i \geq 14$ for $i = 1, 2$.

3. When $q = 3$, Condition (Equation 3.71) is true if either

   (a) $m_{min} = 12$, and $\sum_{j=1}^{3} m_{j,j\neq min} \leq 37$. So either all equal to 12 or some $m_i = 13$ and the other two equal to 12.

   (b) $m_{min} = 13$, and $\sum_{j=1}^{3} m_{j,j\neq min} \leq 122$.

   (c) $m_i \geq 14$ for all $i = 1, 2, 3$.

4. When $q = 4$, Condition (Equation 3.71) is true if either

   (a) $m_{min} = 13$, and $\sum_{j=1}^{4} m_{j,j\neq min} \leq 122$.

   (b) $m_i \geq 14$ for all $i = 1, 2, 3, 4$.

5. When $q = 5$, Condition (Equation 3.71) is true if either

   (a) $m_{min} = 13$, and $\sum_{j=1}^{5} m_{j,j\neq min} \leq 122$.

   (b) $m_i \geq 14$ for all $i = 1, 2, 3, 4, 5$.

6. When $q = 6$, Condition (Equation 3.71) is true if either

   (a) $m_{min} = 13$, and $\sum_{j=1}^{6} m_{j,j\neq min} \leq 122$.

   (b) $m_i \geq 14$ for all $i = 1, 2, 3, 4, 5, 6$.

7. *When* $q = 7$, *Condition (Equation* 3.71*) is true if either*

   (a) $m_{min} = 13$, *and* $\sum_{j=1}^{7} m_{j,j \neq min} \leq 122$.

   (b) $m_i \geq 14$ *for all* $i = 1, 2, 3, 4, 5, 6, 7$.

8. *When* $q = 8$, *Condition (Equation* 3.71*) is true if either*

   (a) $m_{min} = 13$, *and* $\sum_{j=1}^{8} m_{j,j \neq min} \leq 122$.

   (b) $m_i \geq 14$ *for all* $i = 1, 2, 3, 4, 5, 6, 7, 8$.

9. *When* $q = 9$, *Condition (Equation* 3.71*) is true if either*

   (a) $m_{min} = 13$, *and* $\sum_{j=1}^{9} m_{j,j \neq min} \leq 122$.

   (b) $m_i \geq 14$ *for all* $i = 1, 2, 3, 4, 5, 6, 7, 8, 9$.

10. *When* $q \geq 10$, *Condition (Equation* 3.71*) is true if* $m_i \geq 14$ *for all* $i = 1, 2, \ldots, q$.

Similar discussion can be applied for condition (*Equation* 3.72).

3.6 Conclusion

In this chapter we compared the p-source based GDE with its q-sub-source based competitors under Pitman closeness criterion. We established a necessary condition and several sufficient conditions for the p-source based GDE to be Pitman closer than its q-sub-source based GDE. We further discussed the sample size requirement corresponding to each source, and we found that,one sufficient condition is$n_i \geq 15$ for $i = 1, 2, \ldots, p$.

This sample size requirement is relatively close to the requirement based on mean square errors. Hence in our point of view, the p-source-based GDE dominates any other q-sub-source based GDE not only in terms of mean square error loss function but also in the sense of Pitman closeness criterion of probability coverage, when the sample size of each source is moderately large enough.

CHAPTER 4

UNBIASED ESTIMATION OF RELIABILITY FUNCTION FROM MIXTURE EXPONENTIAL DISTRIBUTIONS

[1] In this chapter we investigated an unbiased estimation of the reliability function, based on a single observation from a mixture of two exponential distributions with known mixing proportions. We tried to do so by closely extrapolating a similar method of Sinha et al (2006). We introduced several equivalent versions of an unbiased estimators, ans find that they gave 'proper' estimates only if negative weights on component mixture distributions were allowed.

4.1 Introduction

In this section, we briefly reviewed the results on reliability estimation based on exponential distribution. Let X_1, ..., X_n be n independent observation on X, which follows an exponential distribution with unknown parameter (mean) λ (> 0), written henceforth as $exp(\lambda)$. The problem is to unbiasedly estimate the reliability function (survival function) $R(t) = e^{-t/\lambda}$ for a specified $t > 0$. The following result is immediate.

Theorem 22.

$$\hat{R}_1(t) = \frac{1}{n}\sum_{i=1}^{n} I(X_i > t),$$

[1]Part of this chapter is coming from a working paper "Unbiased Estimation of Reliability Function from Mixture Exponential Distributions" from Nie, Sinha and Hedayat.

where $I(\dot{.})$ is the indicator function, is a simple unbiased estimator of $R(t)$. The variance of $\hat{R}_1(t)$ is :

$$Var(\hat{R}_1(t)) = \frac{R(t)(1 - R(t))}{n}. \tag{4.1}$$

Proof.

As for any $i = 1, \ldots, n$, it is easy to show that:

$$E(I(X_i > t)) = \exp(-t/\lambda) = R(t).$$

So

$$E(\hat{R}_1(t)) = \frac{1}{n} \sum_{i=1}^{n} E(I(X_i > t)) = \frac{1}{n} \sum_{i=1}^{n} R(t) = R(t).$$

We also note that:

$$\begin{aligned} E(\hat{R}_1^2(t)) &= \frac{1}{n^2} \left(\sum_{i=1}^{n} E(I(X_i > t)) + \sum_{i=1}^{n} \sum_{j=1, j \neq i}^{n} E(I(X_i > t) I(X_j > t)) \right) \\ &= \frac{1}{n^2} \left(\sum_{i=1}^{n} R(t) + \sum_{i=1}^{n} \sum_{j=1, j \neq i}^{n} R^2(t) \right) \\ &= \frac{1}{n} R(t) + \frac{n-1}{n} R^2(t). \end{aligned}$$

So the variance of $\hat{R}_1(t)$ is:

$$\begin{aligned} V(\hat{R}_1(t)) &= E(\hat{R}_1^2(t)) - E^2(\hat{R}_1(t)) \\ &= \frac{1}{n} R(t) + \frac{n-1}{n} R^2(t) - R^2(t) \end{aligned}$$

$$= \frac{1}{n} R(t)(1 - R(t)).$$

□

Pugh (37) and Basu (1) provided the UMVUE of $R(t)$ in the following:

Theorem 23. *The UMVUE of $R(t)$ based on $X_1, \ldots, X_n$ is*

$$\hat{R}_{umvue}(t) = \begin{cases} (1 - \frac{t}{W})^{n-1} & W > t \\ \\ 0 & otherwise \end{cases} \tag{4.2}$$

where $W = \sum_{i=1}^{n} X_i$. The variance of $\hat{R}_{umvue}(t)$ is:

$$Var(\hat{R}_{umvue}(t)) = R(t)(\phi(t) - R(t)), \tag{4.3}$$

where $\phi(t) = \frac{1}{\Gamma(n)} \int_0^\infty e^{-\mu} \frac{\mu^{2n-2}}{(\mu + \frac{t}{\lambda})^{n-1}} d\mu$.

Proof.

See Pugh (37) and Basu (1). □

Due to the general fact that the difference of two successive spacings involving order statistics of exponential distributions is again related to the same exponential distribution, Basu (1) also provided the UMVUE of $R(t)$ based on first r $(1 < r < n)$ order statistics:

Theorem 24. *The UMVUE of $R(t)$ from first r $(r < n)$ order statistics is*

$$\hat{R}_{os}(t) = \begin{cases} (1 - \frac{t}{W^*})^{r-1} & W^* > t \\ \\ 0 & otherwise \end{cases} \tag{4.4}$$

where $W^* = \sum_{i=1}^{r} X_{(i)} + (n-r)X_{(r)}$, and $X_{(i)}$ is the i^{th} order statistic, $i = 1, 2, ..., n$.

Proof.

See Basu (1). □

Sinha et al.(45) derived an explicit form of the unique unbiased estimator of $R(t)$ based on a single i^{th} order statistic X_i $(1 \le i \le n)$. The result is reproduced below.

Theorem 25. *The unique unbiased estimator of $R(t)$ based on a single i^{th} order statistic X_i $(1 \le i \le n)$, denoted as $h_i(Z_{(i)})$ where $Z_{(i)} = (n-i+1)X_{(i)}$, is given by:*

$$h_i(Z_{(i)}) = \sum_{y_1=0}^{\infty} \sum_{y_2=0}^{\infty} \dots \sum_{y_{i-1}=0}^{\infty} d_{y_1 y_2 \dots y_{i-1}} I(Z_i > \alpha_1^{y_1} \alpha_2^{y_2} \dots \alpha_{i-1}^{y_{i-1}} t), \tag{4.5}$$

where $\alpha_j = \frac{n-i+j}{n-i}$, $j = 1, 2, \dots, i$,
$d_{y_1 y_2 \dots y_{i-1}} = \dfrac{(-1)^{\sum_j y_j}}{\binom{n}{i}} \frac{(y_1+y_2+\dots+y_{i-1})!}{y_1! y_2! \dots y_{i-1}!} \frac{\binom{i-1}{1}^{y_1}}{\alpha^{y_1}} \frac{\binom{i-1}{2}^{y_2}}{\alpha^{y_2}} \dots \frac{\binom{i-1}{i-1}^{y_{i-1}}}{\alpha^{y_{i-1}}}$,
$\sum_j$ means sum of all even suffixes of y and all y_j are integers, $j = 1, 2, \dots, i-1$.

Proof.

See Sinha et al. (45). □

For example, when $i = 2$, we can simplify the above expression as:

$$h_2(Z_{(2)}) = \sum_{y_1=0}^{\infty} \frac{1}{n\alpha^{y_1}} I(Z_{(2)} > \alpha^{y_1} t), \tag{4.6}$$

where $\alpha = \frac{n}{n-1}$.

Based on the results above, we efficiently obtain the unbiased estimate of $R(t)$ where only k selected order statistics are available ($1 \leq k < n$). However, in practice we know that X most likely follows a mixture distribution. In section 2 we shall establish that any unbiased estimator of $R(t)$ based on a random observation from a mixture of two exponential distributions remains a proper estimator (between 0 and 1) if and only if one of the two weights of the mixture distributions is negative. In section 4.3, we proceed to provide unbiased estimates for the variance, when a proper estimator is available. We should mention that Jevremovic (22) provided some example on mixed exponential distributions with negative weights, using an auto-regressive process.

Example 2. *Define:*

$$Y_t = \begin{cases} \alpha\xi_t & w.p.\ p_0, \\ \beta\xi_t + Y_{t-1} & w.p.\ p_1, \\ Y_{t-1} & w.p.\ q_1, \end{cases} \tag{4.7}$$

where α, β, p_0, p_1, q_1 are between 0 and 1 , and $p_0 + p_1 + q_1 = 1$. The sequence ξ_t is an i.i.d sequence with exponential density function: $\frac{1}{\lambda}\exp(-\frac{y}{\lambda})$, and we assume also the independence of ξ_t and Y_s for $s < t$.

Then Jevremovic (22) showed that Y_t follows a mixture exponential distributions :

$$f_Y(y) = a_1 \frac{1}{\gamma_1} \exp(-\frac{x}{\gamma_1}) + a_2 \frac{1}{\gamma_2} \exp(-\frac{y}{\gamma_2}),$$

with

$$a_1 = p_0(\beta - \alpha)/((1 - q_1)\beta - p_0\alpha),$$

$$a_2 = 1 - a_1,$$

$$\gamma_1 = \alpha\lambda,$$

$$\gamma_1 = \frac{\beta\lambda(1 - q_1)}{p_0}.$$

It is easy to show that :

1. $p_0(\beta - \alpha) + p_1\beta < 0$ *implies* $a_1 > 1$,
2. $\beta < \alpha$ *and* $p_0(\beta - \alpha) + p_1\beta > 0$ *implies* $a_1 < 0$,
3. *in the other cases,* $0 < a_1 < 1$.

4.2 Negative Result: Improper Unbiased Estimates Based On A Single Observation

Consider a random variable X following a mixture of two exponential distributions with means λ and $\theta\lambda$ (θ, $\lambda > 0$), say $X \sim p \exp(\lambda) + q \exp(\theta\lambda)$, where $p + q = 1$, p, θ were known and λ was unknown. The problem is to unbiased estimate the function $R(t) = e^{-t/\lambda}$ for a given

$t > 0$, based on a single observation x. It is clear that $R(t)$ defined above can be identified as the reliability at t for a distribution following $\exp(\lambda)$.

Inspired by Sinha et al (45), we have the following theorem:

Theorem 26. *An unbiased estimator of $R(t)$, say $h(x;t)$ should satisfy $I(x > t) = ph(x;t) + qh(\theta x;t)$, where $I(x > t)$ is an indicator function.*

Proof.

$$\begin{aligned}
&\frac{1}{\lambda}\int_0^\infty I(x > t)e^{-\frac{x}{\lambda}}dx = R(t) = E\left(h(x;t)\right)\\
&= \frac{1}{\lambda}\int_0^\infty h(x;t)(pe^{-\frac{x}{\lambda}} + \frac{q}{\theta}e^{-\frac{x}{\theta\lambda}})dx \qquad (4.8)\\
&= \frac{1}{\lambda}\left(\int_0^\infty ph(x;t)e^{-\frac{x}{\lambda}} + \int_0^\infty qh(\theta x;t)e^{-\frac{x}{\lambda}}\right)dx
\end{aligned}$$

So $I(x > t) = ph(x;t) + qh(\theta x;t)$ follows from the completeness property of the exponential distribution.

□

There could be more than one unbiased estimator for $R(t)$. Here we offer two of them:

$$h_1(x;t) = \sum_{k=0}^{\infty}\frac{1}{p}(-\frac{q}{p})^k I(X > \theta^{-k}t), \qquad (4.9)$$

or

$$h_2(x;t) = \sum_{k=0}^{\infty}\frac{1}{q}(-\frac{p}{q})^k I(X > \theta^{k+1}t). \qquad (4.10)$$

Theorem 27. $E(h_1(x;t)) = E(h_2(x;t)) = R(t)$.

Proof.

Based on Theorem 26, it follows that:

$$E(I(X > t)) = pR(t) + qR(\theta^{-1}t). \tag{4.11}$$

Hence:

$$\begin{aligned} R(t) &= E\left(\frac{1}{p}I(X > t)\right) - \frac{q}{p}R(\theta^{-1}t) = E\left(\frac{1}{p}I(X > t)\right) - E\left(\frac{1}{p}\frac{q}{p}I(X > \theta^{-1}t)\right) + (\frac{q}{p})^2R(\theta^{-2}t) \\ &= ...... = E\left(\sum_{k=0}^{\infty}\frac{1}{p}(-\frac{q}{p})^kI(X > \theta^{-k}t)\right) = E\left(h_1(x;t)\right). \end{aligned} \tag{4.12}$$

Similarly, since

$$E(I(X > \theta t)) = pR(\theta t) + qR(t), \tag{4.13}$$

we have:

$$\begin{aligned} R(t) &= E\left(\frac{1}{q}I(X > \theta t)\right) - \frac{p}{q}R(\theta t) = E\left(\frac{1}{q}I(X > \theta t)\right) - E\left(\frac{1}{q} * \frac{p}{q}I(X > \theta^2 t)\right) + (\frac{p}{q})^2R(\theta^3 t) \\ &= ...... = E\left(\sum_{k=0}^{\infty}\frac{1}{q}(-\frac{p}{q})^kI(X > \theta^{k+1}t)\right) = E\left(h_2(x;t)\right). \end{aligned} \tag{4.14}$$

□

Notice that both $h_1(x;t)$ and $h_2(x;t)$ are infinite sums. We have to show their convergence based on $|\frac{q}{p}|$ and θ in Table I.

By carefully choosing p and q, we then can decide when $h_1(x;t)$ or $h_2(x;t)$ or both are available to unbiased estimate the reliability function $R(t)$. We have the following three situations:

TABLE I

CONVERGENCE OF $H_1(X;T)$ AND $H_2(X;T)$

	$\theta < 1$	$\theta > 1$
$\lvert\frac{q}{p}\rvert < 1$	$h_1(x;t)$	$h_1(x;t)$,$h_2(x;t)$
$\lvert\frac{q}{p}\rvert > 1$	$h_1(x;t)$,$h_2(x;t)$	$h_2(x;t)$
$\lvert\frac{q}{p}\rvert = 1$	$h_1(x;t)$	$h_2(x;t)$

1. For $p > 1$, hence $q < 0$. We require the density function $\frac{p}{\lambda}e^{-\frac{x}{\lambda}} + \frac{q}{\theta\lambda}e^{-\frac{x}{\theta\lambda}}$ to be non-negative, which means $\theta < 1$. This is because when $\theta > 1$:

$$\frac{p}{\lambda}e^{-\frac{x}{\lambda}} + \frac{q}{\theta\lambda}e^{-\frac{x}{\theta\lambda}} = e^{-\frac{x}{\lambda}}\left(\frac{p}{\lambda} + \frac{q}{\theta\lambda}e^{\frac{x}{\lambda}(1-\frac{1}{\theta})}\right) \xrightarrow{x\to\infty} -\infty \tag{4.15}$$

   Based on Table I, we have $h_1(x;t)$ as our only unbiased estimator. Since the common ratio $0 < -\frac{q}{p} < 1$, this geometric sum $h_1(x;t)$ is in the range $[0,1]$, which implies it is a proper estimator. We give an example in Table II ($\theta = 1/3$,$t = 1$,$p = 1.5$,$q = -0.5$).

2. For $q > 1$, we have $p < 0$ and $\theta > 1$. Similarly to the previous case, we have $h_2(x;t)$ as our proper estimator.

3. For $p, q > 0$, no proper unbiased estimator exists. As the common ratio $-\frac{q}{p}$ in $h_1(x;t)$ and $-\frac{p}{q}$ in $h_2(x;t)$ are negative. Further we proved that the linear combination $\phi h_1(x;t) + (1-\phi)h_2(x;t)$ is not a proper unbiased estimator either. We have an example at Table III

TABLE II

$\theta = 1/3$,$T = 1$,$P = 1.5$,$Q = -0.5$

Range of x	$h_1(x;t)$	$h_2(x;t)$
...	...	...
$(3)^{-3} < x < (3)^{-2}$	0	NA
$(3)^{-2} < x < (3)^{-1}$	0	NA
$(3)^{-1} < x < 1$	0	NA
$1 < x < 3$	$\frac{2}{3}$	NA
$3 < x < 3^2$	$\frac{2}{3} + \frac{2}{3} * \frac{1}{3} = \frac{8}{9}$	NA
$3^2 < x < 3^3$	$\frac{2}{3} + \frac{2}{3} * \frac{1}{3} + \frac{2}{3} * (\frac{1}{3})^2 = \frac{26}{27}$	NA
...	...	...

for $\theta = 1.5^{-1}$,$t = 1$,$p = \frac{1}{3}$,$q = \frac{2}{3}$.

Theorem 28. *When $p, q > 0$, there does not exist ϕ,such that $\phi h_1(x;t) + (1-\phi)h_2(x;t)$ is a proper unbiased estimator* $([0,1])$ *for $R(t)$.*

Proof.

It is easy to verify that $h_1(x;t)$ and $h_2(x;t)$ are improper estimators by themselves, due to the fact that the common ratios are negative.

Consider the case $\frac{q}{p} > 1$ and $\theta > 1$. When $x > 1$, then we have the following:

$$h_1(x;t) = \frac{\frac{1}{p}}{1 + \frac{q}{p}} = 1 \tag{4.16}$$

TABLE III

$\theta = 1.5^{-1}$,$T = 1$,$P = \frac{1}{3}$,$Q = \frac{2}{3}$

Range of x	$h_1(x;t)$	$h_2(x;t)$
...	...	...
$(1.5)^{-3} < x < (1.5)^{-2}$	0	$(\frac{p}{q})^2 = 0.25$
$(1.5)^{-2} < x < (1.5)^{-1}$	0	$-\frac{p}{q} = -0.5$
$(1.5)^{-1} < x < 1$	0	1
$1 < x < 1.5$	$\frac{1}{p} = 3$	1
$1.5 < x < 1.5^2$	$\frac{1}{p} - \frac{1}{p} * \frac{q}{p} = 3 - 3 * 2 = -3$	1
$1.5^2 < x < 1.5^3$	$3 - 3 * 2 + 3 * 2^2 = 9$	1
...	...	...

and

$$h_2(x;t) = \frac{\frac{1}{q}}{1 + \frac{q}{p}} \left(1 - (-\frac{q}{p})^k\right) = 1 - (-\frac{q}{p})^k \textit{ when } \theta^{k+2} > x > \theta^{k+1}. \tag{4.17}$$

For any given k, we note that:

$$\phi h_1(x;t) + (1-\phi)h_2(x;t) = \phi + (1-\phi)\left(1 - (-\frac{q}{p})^k\right) \tag{4.18}$$

$$= 1 - (1-\phi)(-\frac{q}{p})^k. \tag{4.19}$$

Since $\frac{q}{p} > 1$, Equation 4.19 lies in the range $[0, 1]$ if and only if $\phi = 1$. But $h_1(x;t)$ is not a proper estimator either.

We will have similar discussion in the case $|\frac{q}{p}| < 1$ and $\theta < 1$. Hence there is no ϕ such that $\phi h_1(x;t) + (1-\phi)h_2(x;t)$ is a proper estimator when $p, q > 0$. $\square$

In short, the cases that estimator of $R(t)$ is a proper estimator only happens at $p > 1,\ q < 0$ or $p < 0,\ q > 1$. Which implies a negative weight on mixture exponential distributions.

4.3 Unbiased Estimators of the Variances of $h_1(x;t)$ and $h_2(x;t)$

In this section, we provided unbiased estimators of the variances of $h_1(x;t)$ and $h_2(x;t)$. Based on Section 4.2, we only consider proper unbiased estimators. Thus we distinguished between two situations:

1. $\theta < 1$, and $p > 1$.
2. $\theta > 1$, and $q > 1$.

Theorem 29. *When $\theta < 1$, an unbiased estimator of the variance of $h_1(x;t)$ is*

$$\begin{aligned}
&\hat{V}(h_1(x;t)) = h_1^2(x;t) - h_1(x/2;t) \\
&= \frac{1}{p^2}\sum\nolimits_{k=1}^{\infty}\left(2p(-\frac{q}{p})^k + (1-2p)(\frac{q}{p})^{2k}\right) I(X > \theta^{-k}t) - \sum_{k=0}^{\infty}\frac{1}{p}(-\frac{q}{p})^k I(X > 2\theta^{-k}t). \qquad (4.20)
\end{aligned}$$

Proof.

It is straightforward to show the unbiasedness as:

$$\begin{aligned}
Var(h_1(x)) &= E\left(h_1^2(x)\right) - E^2\left(h_1(x)\right) \\
&= E\left(h_1^2(x)\right) - R(2t) \\
&= E\left(h_1^2(x)\right) - E\left(h_1(x/2)\right) \\
&= E\left(h_1^2(x) - h_1(x/2)\right).
\end{aligned}$$

We also note that:

$$\begin{aligned}
h_1^2(x;t) &= \frac{1}{p^2}\left\{\sum_{k=0}^{\infty}(\frac{q}{p})^{2k}I(X>\theta^{-k}t)+2\sum\sum\nolimits_{k>l=0}^{\infty}(-\frac{q}{p})^{l+k}I(X>\theta^{-k}t)\right\}\\
&= \frac{1}{p^2}\left\{\sum_{k=0}^{\infty}(\frac{q}{p})^{2k}I(X>\theta^{-k}t)+2\sum\nolimits_{k=1}^{\infty}(-\frac{q}{p})^{k}\frac{1-(-\frac{q}{p})^k}{1+\frac{q}{p}}I(X>\theta^{-k}t)\right\}\\
&= \frac{1}{p^2}\left\{\sum_{k=0}^{\infty}(\frac{q}{p})^{2k}I(X>\theta^{-k}t)+2\sum\nolimits_{k=0}^{\infty}(-\frac{q}{p})^{k}\frac{1-(-\frac{q}{p})^k}{1+\frac{q}{p}}I(X>\theta^{-k}t)\right\}\\
&= \frac{1}{p^2}\left\{\sum_{k=0}^{\infty}(\frac{q}{p})^{2k}I(X>\theta^{-k}t)+2p\sum\nolimits_{k=0}^{\infty}((-\frac{q}{p})^{k}-(\frac{q}{p})^{2k})I(X>\theta^{-k}t)\right\}\\
&= \frac{1}{p^2}\left\{\sum\nolimits_{k=0}^{\infty}\left(2p(-\frac{q}{p})^{k}+(1-2p)(\frac{q}{p})^{2k}\right)I(X>\theta^{-k}t)\right\}. \qquad (4.21)
\end{aligned}$$

So

$$\begin{aligned}
&\hat{V}(h_1(x;t)) = h_1^2(x;t)-h_1(x/2;t)\\
=&\frac{1}{p^2}\sum\nolimits_{k=0}^{\infty}\left(2p(-\frac{q}{p})^{k}+(1-2p)(\frac{q}{p})^{2k}\right)I(X>\theta^{-k}t)-\sum_{k=0}^{\infty}\frac{1}{p}(-\frac{q}{p})^{k}I(X>2\theta^{-k}t). \qquad (4.22)
\end{aligned}$$

□

Theorem 30. *When $\theta>1$, an unbiased estimator of the variance of $h_2(x;t)$ is*

$$\begin{aligned}
&\hat{V}(h_2(x;t)) = h_2^2(x;t)-h_2(x/2;t)\\
=&\frac{1}{q^2}\sum_{k=0}^{\infty}\left(2q(-\frac{p}{q})^{k}+(1-2q)(\frac{p}{q})^{2k}\right)I(X>\theta^{k+1}t)-\sum\nolimits_{k=0}^{\infty}\frac{1}{q}(-\frac{p}{q})^{k}I(X>2\theta^{k+1}t). \qquad (4.23)
\end{aligned}$$

Proof.

This proof can be achieved by switch the role of p and q from the previous proof. $\square$

We notice that $\hat{V}(h_1(x;t))$ and $\hat{V}(h_2(x;t))$ can be negative. There exist some special cases like $\theta = 0.5$ (or 2), we can check the sign of the proposed unbiased variance estimators. We have the following:

Corollary 31. *When $\theta = 0.5$ and $p > 1$, an unbiased estimator of the variance of $h_1(x;t)$ is*

$$\hat{V}(h_1(x;t)) = \frac{1}{p^2} I(X > t) + \frac{1}{p^2} \sum_{k=1}^{\infty} \left(-(1+q)(-\frac{q}{p})^{k-1} + (1-2p)(\frac{q}{p})^{2k} \right) I(X > 2^k t), \quad (4.24)$$

which is uniformly non-negative if and only if $p < 2$.

Proof.

Plug $\theta = 0.5$ in Theorem 29, then:

$$\begin{aligned}
\hat{V}(h_1(x;t)) =& \frac{1}{p^2} \sum_{k=0}^{\infty} \left(2p(-\frac{q}{p})^k + (1-2p)(\frac{q}{p})^{2k} \right) I(X > 2^k t) - \sum_{k=0}^{\infty} \frac{1}{p}(-\frac{q}{p})^k I(X > 2^{k+1} t) \\
& \frac{1}{p^2} \left\{ I(X > t) + \sum_{k=1}^{\infty} \left(2p(-\frac{q}{p})^k + (1-2p)(\frac{q}{p})^{2k} \right) I(X > 2^k t) \right\} - \\
& \sum_{k=0}^{\infty} \frac{1}{p}(-\frac{q}{p})^k I(X > 2^{k+1} t) \\
=& \frac{1}{p^2} I(X > t) + \sum_{k=1}^{\infty} \left\{ \frac{1}{p^2} \left(2p(-\frac{q}{p})^k + (1-2p)(\frac{q}{p})^{2k} \right) - \frac{1}{p}(-\frac{q}{p})^{k-1} \right\} I(X > 2^k t) \\
=& \frac{1}{p^2} I(X > t) + \frac{1}{p^2} \sum_{k=1}^{\infty} \left(-(p+2q)(-\frac{q}{p})^{k-1} + (1-2p)(\frac{q}{p})^{2k} \right) I(X > 2^k t) \\
=& \frac{1}{p^2} I(X > t) + \frac{1}{p^2} \sum_{k=1}^{\infty} \left(-(1+q)(-\frac{q}{p})^{k-1} + (1-2p)(\frac{q}{p})^{2k} \right) I(X > 2^k t). \quad (4.25)
\end{aligned}$$

When $2^{k+1}t > X > 2^k t$, $k = 0, 1, 2, \ldots, \infty$ we will have:

$$
\begin{aligned}
\hat{V}(h_1(x;t)) &= \frac{1}{p^2}\left\{1 - (1+q)\frac{1-(-\frac{q}{p})^k}{1+\frac{q}{p}} + (1-2p)(\frac{q}{p})^2\frac{1-(-\frac{q}{p})^{2k}}{1-(\frac{q}{p})^2}\right\} \\
&= \frac{1}{p^2}\left\{1 - p(1+q)\left(1-(-\frac{q}{p})^k\right) + (1-2p)q^2\frac{1-(-\frac{q}{p})^{2k}}{p^2-q^2}\right\}. \qquad (4.26)
\end{aligned}
$$

As $p + q = 1$, so it is easy to see that:

$$
\begin{aligned}
& p(1+q) = (1-q)(1+q) = 1 - q^2 \\
& p^2 - q^2 = (p+q)(p-q) = p - q = 2p - 1.
\end{aligned}
$$

Then Equation 4.26 can be rewritten as:

$$
\begin{aligned}
\hat{V}(h_1(x;t)) &= \frac{1}{p^2}\left\{1 - (1-q^2)\left(1-(-\frac{q}{p})^k\right) - q^2\left(1-(-\frac{q}{p})^{2k}\right)\right\} \\
&= \frac{1}{p^2}\left\{(1-q^2)(-\frac{q}{p})^k + q^2(-\frac{q}{p})^{2k}\right\} \\
&= \frac{1}{p^2}(-\frac{q}{p})^k\left\{(1-q^2) + q^2(-\frac{q}{p})^k\right\}. \qquad (4.27)
\end{aligned}
$$

When $1 < p < 2$, then $-1 < q < 0$ and $0 < -\frac{q}{p} < 1$. Hence $1 - q^2 > 0$. It is then easy to show that $\hat{V}(h_1(x;t)) \geq 0$. To check that it is also the necessary condition, note that *Equation* $4.27 \geq 0$ for any $k = 0, 1, \ldots, \infty$.

When $k \to \infty$, we have that:

$$\hat{V}(h_1(x;t)) \xrightarrow{k\to\infty} \frac{1}{p^2}\left(-\frac{q}{p}\right)^k (1-q^2) \geq 0.$$

For it to remain non-negative, we need to have:

$$(1-q^2) \geq 0$$

Since $p > 1$, we will need $-1 < q < 0$. □

Similarly by switch the role of p and q, we will have the following corollary.

Corollary 32. *When $\theta = 2$ and $p < 0$, an unbiased estimator of the variance of $h_2(x;t)$ is*

$$\hat{V}(h_2(x;t)) = \frac{1}{q^2}I(X > 2t) + \frac{1}{q^2}\sum\nolimits_{k=1}^{\infty}\left(-(1+p)(-\frac{p}{q})^{k-1} + (1-2q)(\frac{p}{q})^{2k}\right) I(X > 2^{k+1}t), \tag{4.28}$$

which is uniformly non-negative if and only if $1 < q < 2$.

4.4 Conclusion

In this chapter, we provided two unbiased estimators $h_1(x;t)$ and $h_2(x;t)$ of the reliability function $R(t) = e^{-t/\lambda}$ based on a single observation x from a mixture of two exponential distributions. Further we showed that there is no proper estimator with linear combination of $h_1(x;t)$ and $h_2(x;t)$, when the weights of the distributions p and q are non-negative. This negative result is due to the nature of mixing properties of the two exponential distributions.

There exist proper unbiased estimators when we accept negative weights on one of the two component distributions. We also investigated the form of unbiased estimators of the variances, and verified that under certain situations, unbiased variance estimators may turn out to be non-negative.

Based on our results, it might be interesting to further investigate the unbiased estimator of $R(t)$ based on a observation from a general mixture exponential distributions. The authors would also suggest to check the situation when an order statistic observation is available.

CITED LITERATURE

1. Basu, A.: Estimates of reliability for some distributions useful in life testing. Technometrics, 6(2):215–219, 1964.

2. Bhattacharya, C. G.: Estimation of a common mean and recovery of interblock information. The Annals of Statistics, 8(1):205–211, 1980.

3. Bhattacharya, C.: Yates type estimators of a common mean. Annals of the Institute of Statistical Mathematics, 30(1):407–414, 1978.

4. Bhattacharya, C.: A note on estimating the common mean of k-normal populations. Sankhyā: The Indian Journal of Statistics, Series B, 40(3/4):272–275, 1979.

5. Bhattacharya, C.: Estimation of a common location. Communications in Statistics-Theory and Methods, 10(10):955–961, 1981.

6. Bhattacharya, C.: On the cohen-sackrowitz estimator of a common mean. Statistics: A Journal of Theoretical and Applied Statistics, 19(4):493–501, 1988.

7. Blyth, C.: Is pitman closeness a reasonable criterion?:comment. Journal of the American Statistical Association, 88(421):72–74, 1993.

8. Brown, L. and Cohen, A.: Point and confidence estimation of a common mean and recovery of interblock information. The Annals of Statistics, 2(5):963–976, 1974.

9. Carter, R. L. and Yang, M. C.: Large sample inference in random coefficient regression models. Communications in Statistics-Theory and Methods, 15(8):2507–2525, 1986.

10. Casella, G. and Wells, M. T.: Is pitman closeness a reasonable criterion?: comment. Journal of the American Statistical Association, 88(421):70–71, 1993.

11. Chen, Z., Bai, Z., and Sinha, B. K.: Ranked set sampling: theory and applications. 2004.

12. Chiou, W.-J. and Cohen, A.: On estimating a common multivariate normal mean vector. Annals of the Institute of Statistical Mathematics, 37(1):499–506, 1985.

13. Chuiv, N. N. and Sinha, B. K.: On some aspects of ranked set sampling in parametric estimation. Handbook of Statistics, 17:337–377, 1998.

14. Cochran, W. G.: Problems arising in the analysis of a series of similar experiments. Supplement to the Journal of the Royal Statistical Society, 4(1):102–118, 1937.

15. Cochran, W. G. and Carroll, S. P.: A sampling investigation of the efficiency of weighting inversely as the estimated variance. Biometrics, 9(4):447–459, 1953.

16. Cohen, A. and Sackrowitz, H. B.: On estimating the common mean of two normal distributions. The Annals of Statistics, 2(6):1274–1282, 1974.

17. Ghosh, J. and Sinha, B. K.: A necessary and sufficient condition for second order admissibility with applications to berkson's bioassay problem. The Annals of Statistics, 9(6):1334–1338, 1981.

18. Ghosh, M., Keating, J., and Sen, P. K.: Is pitman closeness a reasonable criterion?: comment. Journal of the American Statistical Association, 88(421):63–66, 1993.

19. Graybill, F. A. and Deal, R.: Combining unbiased estimators. Biometrics, 15(4):543–550, 1959.

20. Haff, L.: Minimax estimators for a multinormal precision matrix. Journal of Multivariate Analysis, 7(3):374–385, 1977.

21. Hartung, J.: A note on combining dependent tests of significance. Biometrical Journal, 41(7):849–855, 1998.

22. Jevremovic, V.: A note on mixed exponential distribution with negative weights. Statistics and Probability Letters, 11(3):259–265, 1991.

23. Khatri, C. and Shah, K.: Estimation of location parameters from two linear models under normality. Communications in Statistics-Theory and Methods, 3(7):647–663, 1974.

24. Kubokawa, T.: Estimating common parameters of growth curve models under a quadratic loss. Communications in Statistics-Theory and Methods, 18(9):3149–3155, 1989.

25. Kubokawa, T.: Closer estimators of a common mean in the sense of pitman. Annals of the Institute of Statistical Mathematics, 41(3):477–484, 1989.

26. Kubokawa, T.: Minimax estimation of common coefficients of several regression models under quadratic loss. Journal of Statistical Planning and Inference, 24(3):337–345, 1990.

27. Lehnman, E. and Casella, G.: Theory of point estimation. 1983.

28. Liski, E. P., Luoma, A., and Sinha, B. K.: Optimal designs in random coefficient linear regression models. Bulletin of the Calcutta Statistical Association, 46(183):211–230, 1996.

29. Loh, W.-L.: Estimating the common mean of two multivariate normal distributions. The Annals of Statistics, 19(1):297–313, 1991.

30. McIntyre, G.: A method for unbiased selective sampling, using ranked sets. Crop and Pasture Science, 3(4):385–390, 1952.

31. Meier, P.: Variance of a weighted mean. Biometrics, 9(1):59–73, 1953.

32. Nair, K.: The recovery of inter-block information in incomplete block designs. Sankhyā: The Indian Journal of Statistics, 10(1):383–390, 1943.

33. Norwood Jr., T. E. and Hinkelmann, K.: Estimating the common mean of several normal populations. The Annals of Statistics, 5(5):1047–1050, 1977.

34. Pal, N. and Sinha, B. K.: Estimation of a common mean of several normal populations. Far East Journal of Mathematical Sciences, Special(I):97–110, 1996.

35. Peddada, S. D.: Is pitman closeness a reasonable criterion?: comment. Journal of the American Statistical Association, 88(421):67–69, 1993.

36. Pitman, E. J.: The closest estimates of statistical parameters. Mathematical Proceedings of the Cambridge Philosophical Society, 33(2):212–222, 1937.

37. Pugh, E. L.: The best estimate of reliability in the exponential case. Operations Research, 11(1):57–61, 1963.

38. Rao, C. R.: General methods of analysis for incomplete block designs. Journal of the American Statistical Association, 42(240):541–561, 1947.

39. Rao, C. R.: On the recovery of inter block information in varietal trials. Sankhyā: The Indian Journal of Statistics, 17(2):105–114, 1956.

40. Rao, C.: Is pitman closeness a reasonable criterion?:comment. Journal of the American Statistical Association, 88(421):69–70, 1993.

41. Rao, J. and Subrahmaniam, K.: Combining independent estimators and estimation in linear regression with unequal variances. Biometrics, 27(4):971–990, 1971.

42. Robert, C. P., Hwang, J. G., and Strawderman, W. E.: Is pitman closeness a reasonable criterion? Journal of the American Statistical Association, 88(421):57–63, 1993.

43. Sarkar, S. K.: On estimating the common mean of several normal populations under the pitman closeness criterion. Communications in Statistics-Theory and Methods, 20(11):3487–3498, 1991.

44. Shinozaki, N.: A note on estimating the common mean of k normal distributions and the stein problem. Communications in Statistics-Theory and Methods, 7(15):1421–1432, 1978.

45. Sinha, B. K., Sengupta, S., and Mukhuti, S.: Unbiased estimation of the distribution function of an exponential population using order statistics with application in ranked set sampling. Communications in Statistics-Theory and Methods, 35(9):1655–1670, 2006.

46. Sinha, B. K.: Unbiased estimation of the variance of the graybill-deal estimator of the common mean of several normal populations. Canadian Journal of Statistics, 13(3):243–247, 1985.

47. Sinha, B. K. and Mouqadem, O.: Estimation of the common mean of two univariate normal populations. Communications in Statistics-Theory and Methods, 11(14):1603–1614, 1982.

48. Tsukuma, H. and Konno, Y.: Modifying the graybill-deal estimator of the common. Research Institute for Mathematical Sciences, 1334:95–111, 2003.

49. Yates, F.: The recovery of inter-block information in variety trials arranged in three-dimensional lattices. Annals of Eugenics, 9(2):136–156, 1939.

50. Yates, F.: The recovery of inter-block information in balanced incomplete block designs. Annals of Eugenics, 10(1):317–325, 1940.

51. Zacks, S.: Unbiased estimation of the common mean of two normal distributions based on small samples of equal size. Journal of the American statistical Association, 61(314):467–476, 1966.

52. Zacks, S.: Bayes and fiducial equivariant estimators of the common mean of two normal distributions. The Annals of Mathematical Statistics, 41(1):59–69, 1970.

VITA

Education

- Ph.D. in Statistics, University of Illinois at Chicago, 2016.
- M.S. in Statistics and Physics, University of Illinois at Chicago, 2011.
- B.S. in Physics, University of Science and Technology of China, 2007.

Experience

- Quantitative Research Analyst, First Capital Securities Co., May 2012-Aug 2012.

Membership

- American Statistical Association (ASA)