Some Contributions to Meta Analysis

Quasi-Empirical Bayes Method of Estimation

by

Amal Ghania

A thesis submitted to
the Faculty of Graduate Studies and Postdoctoral Affairs
in partial fulfillment of the requirements for the degree of
Doctor of Philosophy

in
Probability and Statistics

Carleton University
Ottawa, Ontario

©Copyright
2011, Amal Ghania

Library and Archives
Canada

Published Heritage
Branch

395 Wellington Street
Ottawa ON K1A 0N4
Canada

Bibliothèque et
Archives Canada

Direction du
Patrimoine de l'édition

395, rue Wellington
Ottawa ON K1A 0N4
Canada

Your file *Votre référence*
ISBN: 978-0-494-83215-8
Our file *Notre référence*
ISBN: 978-0-494-83215-8

NOTICE:

The author has granted a non-exclusive license allowing Library and Archives Canada to reproduce, publish, archive, preserve, conserve, communicate to the public by telecommunication or on the Internet, loan, distribute and sell theses worldwide, for commercial or non-commercial purposes, in microform, paper, electronic and/or any other formats.

The author retains copyright ownership and moral rights in this thesis. Neither the thesis nor substantial extracts from it may be printed or otherwise reproduced without the author's permission.

AVIS:

L'auteur a accordé une licence non exclusive permettant à la Bibliothèque et Archives Canada de reproduire, publier, archiver, sauvegarder, conserver, transmettre au public par télécommunication ou par l'Internet, prêter, distribuer et vendre des thèses partout dans le monde, à des fins commerciales ou autres, sur support microforme, papier, électronique et/ou autres formats.

L'auteur conserve la propriété du droit d'auteur et des droits moraux qui protègent cette thèse. Ni la thèse ni des extraits substantiels de celle-ci ne doivent être imprimés ou autrement reproduits sans son autorisation.

In compliance with the Canadian Privacy Act some supporting forms may have been removed from this thesis.

While these forms may be included in the document page count, their removal does not represent any loss of content from the thesis.

Conformément à la loi canadienne sur la protection de la vie privée, quelques formulaires secondaires ont été enlevés de cette thèse.

Bien que ces formulaires aient inclus dans la pagination, il n'y aura aucun contenu manquant.

Canada

Abstract

This dissertation deals with the century old methodology of meta analysis which refers to the synthesis of results from a series of independent studies that attempt to find a common thread among the studies. In meta analysis, “effect size” is the measure of the magnitude of treatment effects, and is the basic parameter of interest and the primary objective is to identify common characteristics and combine the effect sizes accordingly across several independent studies. Naturally, appropriate estimation of this parameter is important and desirable to evaluate the studies successfully.

In this dissertation we provide a new dimension of statistical methodology in meta analysis by introducing the “quasi-empirical Bayes methods” of estimation of effect size, which yields efficient estimators than the existing ones. This method of estimation first formulates the homogeneity hypothesis and then combines the estimates of effect sizes in a suitable way to increase its efficiency than the standard ones. Asymptotic theory is provided to study the statistical properties of the estimators.

We first consider the “primary study” of the comparison of experimental and control groups with unequal sample sizes and define two new estimators of effect size besides the basic unbiased effect size estimator and study their finite sample properties. Secondly, we consider the estimation of effect size from a series of independent experiments and define three new competitive estimators besides the two usual estimators using quasi-empirical Bayes method. Properties of dominance characteristics of these estimators under quadratic loss function are studied using asymptotic the-

ory. We also extend the application of quasi-empirical Bayes method to multivariate models with several independent studies. Here, again we consider five estimators and their dominance properties using quadratic loss function, based on asymptotic theory.

Finally, we introduce and develop robust estimates of effect size for univariate and multivariate models based on “order statistics” and simple rank test, namely, sign test and signed rank test. This opens the door to enlarge the scope of robust methodology in meta analysis based on nonparametric methods.

Acknowledgements

I would like to express my special thanks to my co supervisors, Dr. A. K. Md.Ehsanes Saleh, Dr. M. Mojirsheibani and Dr. M. Ould Haye for their constant support, warm encouragement.

My deep gratitude goes to Professor Saleh who first introduced me to the subject of meta analysis and quasi empirical Bayes methods. His unique insight, enthusiasm and prudent guidance and advice have been truly inspirational.

I am grateful to my teachers for the contributions they have made to my background in Probability and Statistics. I will remember and appreciate them for the rest of my life. I would like to thank Mathematics and Statistics department administrative and assistant administrative of Carleton University for providing good environment and situation for my study.

My sincere thanks are due to the Ministry of Higher Education and State for Scientific Research in Egypt for their generous support.

I owe my loving thanks to my family for their patience and support during my study.

Contents

Abstract	i
Acknowledgements	iii
1 Introduction	1
2 Effect Size and Various Estimators	11
2.1 Introduction	11
2.2 Small Sample Properties of the Estimators	13
2.3 Large Sample Properties of the Unbiased Estimator of the Effect Size	17
3 Effect Size Estimation:Univariate Model	21
3.1 Introduction	21
3.2 One Study Problem	22
3.2.1 Case 1:Variance is Known	22
3.2.2 Confidence Intervals	28
3.2.3 Case 2:Variance is Unknown	36
3.2.4 Confidence Intervals	42
3.3 Empirical Bayes Estimators and Comparisons	45
3.4 Multiple-study Problem	47
3.4.1 Quasi-Empirical Bayes Methods of Meta Analysis	53
3.4.2 Asymptotic Distributional Properties of the Estimators	57

3.4.3	Analysis of Risk	65
3.4.4	Asymptotic Relative Efficiency of the Estimators	70
3.5	Variance-stabilizing Transformation	77
3.5.1	Proposed Estimators	79
3.5.2	Asymptotic Distributional Properties of the Estimators	79
3.6	Meta Analysis of Polls	83
3.6.1	Quasi-Empirical Bayes Methods of Meta Analysis of Polls . .	86
3.6.2	Asymptotic Distributional Properties of the Estimators . . .	86
3.6.3	ADMSE Matrix Analysis of the Estimators	90
4	Multivariate Models	98
4.1	Model and Notation	99
4.2	Multiple-study Problem	100
4.2.1	Quasi Empirical Bayes Estimators	104
4.2.2	Asymptotic Distributional Properties of the Estimators	104
4.2.3	Estimating A Vector of Effect Sizes	113
5	Robust Estimates of “Effect Size”	120
5.1	General Theory of Estimation of Effect Size Based on Order Statistics (L-estimates)	121
5.2	Robust Estimation of “Effect Size” Based on Sample Quantiles . . .	126
5.3	Nonparametric Methods of “Effect Size” Estimation	130
5.4	Robust Estimates of “Effect Sizes” in Multivariate Normal Models . .	132
6	Future Research	147
A	Appendix: Preliminaries	149
A.1	Normal Distribution	149
A.2	Chi-square Distribution and Properties	149

A.3 Some Definitions and Results On Matrices	153
--	-----

List of Tables

3.1	Maximum and minimum MSE-based relative efficiency of PTE for selected α and relative efficiency of SE at Δ_0^2	28
3.2	Coverage probability of 95% and 90% C.I. of the sets $C^{PT}(\hat{\delta}_n^{PT})$ and $C^s(\hat{\delta}_n^s)$	29
3.3	Efficiency of 95% C.I. of the sets $C^{PT}(\hat{\delta}_n^{*PT})$ and $C^s(\hat{\delta}_n^{*s})$	35
3.4	Efficiency of 90% C.I. of the sets $C^{PT}(\hat{\delta}_n^{*PT})$ and $C^s(\hat{\delta}_n^{*s})$	36
3.5	Maximum and minimum relative efficiency of PTE for $\alpha = 0.05(0.10)0.45$ for selected sample sizes	40
3.6	Maximum and Minimum relative efficiency of SE and efficiency of PTE for selected sample sizes	43
3.7	Computation of UE, PTE and SE using prostate cancer data	47
3.8	Maximum and Minimum Asymptotic Relative Efficiency of PTE for selected number of studies.	75
3.9	Maximum and Minimum Asymptotic Relative Efficiency of RE, JSE and PTE for $\alpha=0.05(0.10)0.45$ for selected number of studies.	76
4.1	Data layout	99
5.1	ARE for normal distribution	129
5.2	ARE and corresponding optimum spacings for Cauchy distribution	130

List of Figures

2.1	Exact MSE of four estimators of effect size as a function of δ^2	16
3.1	MSE-based relative efficiencies of SE and PTE for different values of α	26
3.2	Coverage probabilities of 95% C.I. based on PTE ($\alpha= 0.05, 0.1, 0.2$) and SE	30
3.3	Coverage probabilities of 90% C.I. based on PTE ($\alpha=0.05, 0.1, 0.2$) and SE	31
3.4	Graph of MSE-based relative efficiency of SE and PTE for $n^E =$ $9, n^C = 10, \alpha = 0.05$ and 0.15	41
3.5	Asymptotic relative efficiency of PTE for different values of $\alpha, p = 6$	73
3.6	Asymptotic relative efficiency of RE, JSE and PTE for different values of $\alpha, p = 6$	74

Chapter 1

Introduction

Meta analysis refers to statistical synthesis of results from a series of independent studies that attempt to answer similar questions. This methodology dates back to Pearson (1904) who used data from five small independent samples and computed the pooled estimate of correlation between mortality and inoculation in order to evaluate the efficacy of a vaccine. This pooled estimate reflects the efficacy of the vaccine if it passes the statistical test of homogeneity of the correlation coefficients from each sample which in turn shows the effectiveness of the vaccine across these five studies. Further, the five samples together increase the sample size to provide more precise estimator of the correlation than a single sample and provide narrower confidence interval with good coverage. Later, Tippet (1931) and Fisher (1932), Cochran (1937), Yates and Cochran (1938) proposed combining across agricultural experiments in order to obtain estimates of treatment effects and p-values for test of significance.

Meta analysis has grown out of intense effort of a small group of statisticians working on problems of research integration in education and psychotherapy with the initiative of Professor Glass in the early seventies. Glass has been a data-analyst with sophisticated approach. He called the first analysis of data in a research study as “ primary analysis ”. He perceived that two or more study results from similar experiments would yield efficient and reliable result on common characteristics inher-

ent in all the studies. This leads to the term “Meta analysis” to indicate analysis of an ensemble of analysis from individual studies for the purpose of integrating the findings from each of them. During 80’s Peto was instrumental in introducing meta analysis of clinical trials to assess the treatments effects. His idea of combination of studies was to increase the sample size and efficiency of estimation, - an idea which stems from Pearson (1904).

In the field of education, meta analysis is useful in combining studies about coaching to improve SAT scores in verbal and math, comparison of distance education with traditional class room learning, assessing the impact of schooling in developing countries and relationship between teacher credentials and student achievements. In Psychology, meta analysis has been applied to basic science as well as in support of evidence-based practice. Meta analysis of psychological intervention has been used to compare treatments for psychological problems, involving various problems such as obsessive-compulsive disorder, depression phobias, and panic disorder. Meta analysis is now widely used in medicine, education and throughout the social sciences.

In meta analysis the effect size is the unit of currency. Effect size is a name given to a family of indices that measure the magnitude of a treatment effect. There is a wide array of formulas used to measure effect size. The kind of results reported by the primary studies will lead to the choice of the effect size index, if the results are means and standard deviations in two groups, the effect size will be the standardized mean difference. If the results are binary outcomes, the appropriate effect size will be the risk ratio or odds ratio. If the data reports a correlation between two variables, then the correlation coefficient may be used as the effect size. The effect size is what we need to decide on a course of action. For instance, in educational research a researcher needs to make a decision about whether or not to implement an intervention to increase the test scores of students, the researcher wants to know if this intervention increases the scores of students by 10% or 20%, and this information is carried by the

effect size.

If different studies use different instruments (such as different psychological or educational tests) to assess the outcome, then the scale of measurement will differ from study to study. In such cases we can divide the mean difference in each study by that study standard deviation to create an index (the standardized mean difference) that would be comparable across studies.

One of the main goals in meta analysis is to present a summary effect. The summary effect is the weighted mean of the individual effects. Next, we present some preliminaries on various methods of combining effect sizes in meta analysis.

Methods of Combining Effect Sizes

We present here a simplified version of these procedures. Consider a sequence of k independent studies and let $\delta_1, \dots, \delta_k$ be the population effect sizes. Let $\tilde{\delta}_1, \dots, \tilde{\delta}_k$ be the unbiased effect size estimators based on samples of sizes $n_1, \dots, n_k$. Under the homogeneity hypothesis

$$H_0 : \delta_1 = \dots = \delta_k = \delta \text{ unknown,}$$

we obtain the restricted estimator, $\hat{\delta}$ of δ by minimizing $\sum_{i=1}^k \frac{1}{\sigma_i^2} (\tilde{\delta}_i - \delta)^2$ defined by

$$\hat{\delta} = \frac{\sum_{i=1}^k \frac{1}{\sigma_i^2} \tilde{\delta}_i}{\sum_{i=1}^k \frac{1}{\sigma_i^2}},$$

where $\sigma_i^2, i = 1, \dots, k$, is the within study variance. This model is known as **the fixed effect model**. The common estimate, $\hat{\delta}$ is a weighted mean where the weights are reciprocal of the variance of each estimate, $\tilde{\delta}_i$. Consequently,

$$\text{var}(\hat{\delta}) = \frac{1}{\sum_{i=1}^k \frac{1}{\sigma_i^2}}.$$

When σ_i^2 's are unknown replace σ_i^2 by its estimate to calculate $\hat{\delta}$.

Alternatively, we may have the following model

$$Y_{ij} = \delta + b_i + \epsilon_{ij}, \quad i = 1, \dots, k, \quad j = 1, \dots, n_i$$

where Y_{ij} denotes the observable variable, δ denotes the fixed but unknown grand mean and b_i is the unobservable random effect with mean zero and variance σ_b^2 and ϵ_{ij} is the sampling error with $E(\epsilon_{ij}) = 0$, and $E(\epsilon_{ij}^2) = \sigma_i^2$. The assumption is the random variables $b_1, \dots, b_k, \epsilon_{11}, \dots, \epsilon_{kn_k}$ are normally distributed and mutually stochastically independent. This model is known as **the random effects model**. The basic statistics for the above model are the sample means $\bar{Y}_i$ and sample sum of squares SS_i , defined by

$$\bar{Y}_i = \sum_{j=1}^{n_i} \frac{Y_{ij}}{n_i}, \quad SS_i = \sum_{j=1}^{n_i} (Y_{ij} - \bar{Y}_i)^2, \quad i = 1, \dots, k.$$

Then the grand mean and the two well-known sums of squares, namely, between sum of squares (BSS) and within sum of squares (WSS) are defined as

$$\begin{aligned} \bar{Y} &= \sum_{i=1}^k \frac{\bar{Y}_i}{n}, \quad n = \sum_{i=1}^k n_i. \\ \text{BSS} &= \sum_{i=1}^k n_i (\bar{Y}_i - \bar{Y})^2, \quad \text{WSS} = SS_1 + \dots + SS_k. \end{aligned}$$

Obviously, under the assumption of normality and independence, it holds that

$$\bar{Y} \sim N \left(\delta, \sum_{i=1}^k \frac{n_i^2 (\sigma_b^2 + \sigma_i^2 / n_i)}{n^2} \right).$$

Both σ_b^2 and σ_i^2 assumed to be known or estimated. Under this model the combined

estimate of δ is obtained by minimizing $\sum_{i=1}^k \frac{1}{\sigma_b^2 + \sigma_i^2} (Y_{ij} - \delta)^2$ w.r.t. δ as

$$\hat{\delta} = \frac{\sum_{i=1}^k \frac{1}{\sigma_i^2 + \sigma_b^2} \tilde{\delta}_i}{\sum_{i=1}^k \frac{1}{\sigma_i^2 + \sigma_b^2}}.$$

There are standard methods for estimation of σ_b^2 based on the BSS and WSS defined above. we refer the interested reader to the book by Hartung, Knapp and Sinha (2008) Chpter 7.

We can see that the weights placed in each of the above estimators are reciprocal of the associated variances which reflects the precisions of the estimators which depend on the assumptions about the errors. Under the fixed-effect model, the only source of uncertainty is the within-study (sampling or estimation) error. Under the random-effects model, there is the same source of uncertainty plus an additional source (between-studies variance). It follows that the variance, standard error, and confidence interval for the summary effect will always be larger (or wider) under the random-effects model than under the fixed-effect model.

Whether we have fixed effect model or random effects model, we can circumvent the problem, one may consider Bayesian approach where all parameters are considered as random. We shall now discuss Bayes and empirical Bayes approaches.

Bayesian approach

First serious attempt to give a rational approach to the process of statistical inference was put forward by Bayes (1763). We express briefly the Bayesian method applied to the effect sizes. Consider a sequence of k independent studies and let $\delta_1, \dots, \delta_k$ be the population effect sizes. Let $\tilde{\delta}_1, \dots, \tilde{\delta}_k$ be the sample effect sizes which are assumed to be independent with normal distribution with mean δ_i and variance σ_i^2 , ($i = 1, \dots, k$). That is,

$$\tilde{\delta}_i | \delta_i \sim N(\delta_i, \sigma_i^2), \text{ where } \sim \text{ stands for "distributed as".}$$

In the Bayesian method, the assumption is each effect size, δ_i , is normally distributed

with the same mean, δ , and variance σ_b^2 . That is

$$\delta_i \sim N(\delta, \sigma_b^2) \quad \forall i.$$

We note that σ_i^2 denotes within study variability while σ_b^2 denotes the between study variability. It is easy to find the marginal distribution of $\tilde{\delta}_i$ as $N(\delta, \sigma_i^2 + \sigma_b^2)$ for each $i = 1, \dots, k$. Hence, the Bayesian point estimator of δ is given by

$$\hat{\delta}^B = \frac{\sum_{i=1}^k \frac{1}{\sigma_i^2 + \sigma_b^2} \tilde{\delta}_i}{\sum_{i=1}^k \frac{1}{\sigma_i^2 + \sigma_b^2}}.$$

We assumed both σ_i^2 's and σ_b^2 are known or estimated, the variance of $\hat{\delta}^B$ is $(\sum_{i=1}^k \frac{1}{\sigma_i^2 + \sigma_b^2})^{-1}$. The results above hold good, even if we have prior $\pi(\delta, \sigma_b^2)$ which results in the posterior distribution which equals the likelihood function involving δ and σ_b^2 times the prior $\pi(\delta, \sigma_b^2)$. So the estimated common mean is obtained as weighted effect sizes of k studies and weights are proportional to the precision of each effect size.

James-Stein Estimator

Stein (1955, 1956) followed by James and Stein (1961), discovered a paradoxical statistics which undermined century and a half old work on estimation theory, going back to Gauss, Legendre, Fisher and Rao. The result is that the maximum likelihood estimators in Gaussian models are inadmissible beyond two dimensions. This was a shocking news for the statistical community and an era of new thought process. In 1961 James and Stein displayed their estimator as

$$\hat{\boldsymbol{\delta}}^{JS} = \left(1 - \frac{(k-2)}{S^2}\right) \tilde{\boldsymbol{\delta}},$$

where $\tilde{\boldsymbol{\delta}} = (\tilde{\delta}_1, \dots, \tilde{\delta}_k)'$ is a sample from $N_k(\boldsymbol{\delta}, \mathbf{I}_k)$, $\boldsymbol{\delta} = (\delta_1, \dots, \delta_k)'$ and $S^2 = \sum_{i=1}^k \tilde{\delta}_i^2$. For $k \geq 3$, James-Stein estimator everywhere dominates the MLE $\tilde{\boldsymbol{\delta}}$ in terms of mean-

squared error (MSE), that is

$$E\left\{\sum_{i=1}^k(\hat{\delta}_i^{JS} - \delta_i)^2\right\} \leq E\left\{\sum_{i=1}^k(\tilde{\delta}_i - \delta_i)^2\right\}$$

for every choice of $\boldsymbol{\delta} = (\delta_1, \dots, \delta_k)'$. James-Stein estimator offers frequent dominance properties. It does not matter what the prior belief may be about $\boldsymbol{\delta}$. It has been observed that the gain in using James-Stein estimator (JSE) could be enormous in favorable circumstances. James-Stein estimator gathers evidence from its ensemble which provides extra evidence to improve its performance.

Robbins (1956) proposed empirical Bayes method for statistical inference. Efron and Morris (1971 - 1973) initiated the formulation of James-Stein estimator using empirical Bayes approach to uncover the mysteries of this method, where prior may exist as a motivational tool. In this regard, Efron (1995) in his famous article “ The Statistical Century ” includes “ Empirical Bayes and James-Stein Estimation” as an important future topic of study among dozen of his favorite topics in applied statistics. As an evidence of this prediction see the recent book “ Large Scale Inference: Empirical Bayes Methods for Estimation, Testing and Prediction” by Efron (2010).

Empirical Bayes Approach

Now, we discuss the empirical Bayes estimation (EBE) of Efron and Morris (1971 - 1973). We assume that the conditional distribution of $\tilde{\boldsymbol{\delta}} = (\tilde{\delta}_1, \dots, \tilde{\delta}_k)'$ given $\boldsymbol{\delta} = (\delta_1, \dots, \delta_k)'$ is $N_k(\boldsymbol{\delta}, \sigma^2 \mathbf{I}_k)$. we may assume that the distribution of $\boldsymbol{\delta} = (\delta_1, \dots, \delta_k)'$ is $N_k(\delta_0 \mathbf{1}_k, \tau^2 \mathbf{I}_k)$ where τ^2 is the between studies variability, and $\mathbf{1}_k = (1, \dots, 1)'$ is a tuple of k one's and $\mathbf{I}_k$ is the identity matrix of dimension k . Then, the marginal distribution of $\tilde{\boldsymbol{\delta}}$ is given by $N_k(\delta_0 \mathbf{1}_k, (\sigma^2 + \tau^2) \mathbf{I}_k)$. Consequently, the conditional distribution of $\boldsymbol{\delta}$ given $\tilde{\boldsymbol{\delta}}$ is $N_k(\boldsymbol{\delta}^*, \frac{\tau^2 \sigma^2}{\sigma^2 + \tau^2} \mathbf{I}_k)$ where $\boldsymbol{\delta}^* = \sigma^2(\sigma^2 + \tau^2)^{-1} \delta_0 \mathbf{1}_k + \tau^2(\sigma^2 + \tau^2)^{-1} \tilde{\boldsymbol{\delta}}$ which is a convex combination of $\delta_0 \mathbf{1}_k$ and $\tilde{\boldsymbol{\delta}}$. Now, δ_0 is unknown and we can estimate it based on the marginal distribution of $\tilde{\boldsymbol{\delta}}$ as the mean of $\tilde{\delta}_1, \dots, \tilde{\delta}_k$, namely,

$\bar{\delta}_k = k^{-1}(\tilde{\delta}_1 + \cdots + \tilde{\delta}_k)$. Hence, the empirical Bayes estimator of δ is

$$\begin{aligned}\delta^{EB} &= \frac{\sigma^2}{\sigma^2 + \tau^2} \bar{\delta}_k \mathbf{1}_k + \frac{\tau^2}{\sigma^2 + \tau^2} \tilde{\delta} \\ &= \tilde{\delta} - \frac{\sigma^2}{\sigma^2 + \tau^2} (\tilde{\delta} - \bar{\delta}_k \mathbf{1}_k)\end{aligned}$$

We still have to estimate $\frac{\sigma^2}{\sigma^2 + \tau^2}$ since τ^2 is unknown but σ^2 is known or estimated. Let $S^2 = \sum_{i=1}^k (\tilde{\delta}_i - \bar{\delta}_k)^2$, then $S^2 \sim (\sigma^2 + \tau^2) \chi_{k-1}^2$ based on the marginal distribution of $\tilde{\delta} \sim N_k(\delta_0 \mathbf{1}_k, (\sigma^2 + \tau^2) I_k)$ so that $E(\frac{k-3}{S^2}) = \frac{1}{\sigma^2 + \tau^2}$. Substituting $E(\frac{k-3}{S^2})$ in the above we have the empirical Bayes estimator (EBE)

$$\delta^{EB} = \tilde{\delta} - \frac{(k-3)\sigma^2}{S^2} (\tilde{\delta} - \bar{\delta}_k \mathbf{1}_k)$$

which is the James-Stein (1961) estimator of δ . More generally, we may consider a function $g(S^2)$ of the statistics S^2 and write the estimator as

$$\hat{\delta}^g = \tilde{\delta} - g\left(\frac{S^2}{\sigma^2}\right) (\tilde{\delta} - \bar{\delta}_k \mathbf{1}_k).$$

Then, by choosing

$g(\frac{S^2}{\sigma^2}) = 0$, we obtain the estimator $\tilde{\delta}$.

$g(\frac{S^2}{\sigma^2}) = 1$, we obtain the estimator $\bar{\delta}_k \mathbf{1}_k$.

$g(\frac{S^2}{\sigma^2}) = I(\frac{S^2}{\sigma^2} < c_\alpha)$, we obtain the preliminary test estimator (PTE), namely,

$$\hat{\delta}^{PT} = \tilde{\delta} - (\tilde{\delta} - \bar{\delta}_k \mathbf{1}_k) I\left(\frac{S^2}{\sigma^2} < c_\alpha\right)$$

where $I_x(A)$ stands for the indicator function of the set A , c_α is the α -level critical value from the distribution of S^2 .

Similarly, by choosing $g(\frac{S^2}{\sigma^2}) = (k-3)\sigma^2 S^{-2}$, we obtain the James-Stein estimator

as

$$\hat{\boldsymbol{\delta}}^{JS} = \tilde{\boldsymbol{\delta}} - (k-3)\sigma^2(\tilde{\boldsymbol{\delta}} - \bar{\delta}_k \mathbf{1}_k)S^{-2}.$$

We may also choose

$$g\left(\frac{S^2}{\sigma^2}\right) = 1 - [1 - \sigma^2(k-3)S^{-2}]I\left(\frac{S^2}{\sigma^2} > k-3\right),$$

to obtain another preliminary test estimator, called the positive-rule James-Stein estimator, namely,

$$\begin{aligned} \hat{\boldsymbol{\delta}}^{S+} &= \bar{\delta}_k \mathbf{1}_k I\left(\frac{S^2}{\sigma^2} < k-3\right) + \hat{\boldsymbol{\delta}}^{JS} I\left(\frac{S^2}{\sigma^2} > k-3\right) \\ &= \hat{\boldsymbol{\delta}}^{JS} - (\tilde{\boldsymbol{\delta}} - \bar{\delta}_k \mathbf{1}_k)[1 - (k-3)S^{-2}\sigma^2]I\left(\frac{S^2}{\sigma^2} < k-3\right). \end{aligned}$$

Quasi Empirical Bayes Approach

Sometimes, one may face difficulty in assigning appropriate prior distribution for the effect size parameters $\boldsymbol{\delta} = (\delta_1, \dots, \delta_k)'$ but easily one can specify the common characteristic for the k studies instead of the prior distribution of the vector $\boldsymbol{\delta} = (\delta_1, \dots, \delta_k)'$, namely, one can specify the hypothesis H_0 as

$$H_0 : \delta_1 = \dots = \delta_k = \delta_0 \text{ unknown.}$$

We may now define a departure statistics, $\mathcal{D}$ from the hypothesis H_0 as

$$\mathcal{D} = (\tilde{\boldsymbol{\delta}} - \bar{\delta}_k \mathbf{1}_k)' \sigma^{-2} \mathbf{I}_k (\tilde{\boldsymbol{\delta}} - \bar{\delta}_k \mathbf{1}_k),$$

where $\bar{\delta}_k$ is the estimator of δ_0 . Then, we may define the estimator of $\boldsymbol{\delta} = (\delta_1, \dots, \delta_k)'$ as:

$$\hat{\boldsymbol{\delta}}^g = \tilde{\boldsymbol{\delta}} - g(\mathcal{D})(\tilde{\boldsymbol{\delta}} - \bar{\delta}_k \mathbf{1}_k).$$

For the choice of $g(\mathcal{D})$ as 0, 1, $I(\mathcal{D} < c_\alpha)$, $(k-3)\mathcal{D}^{-1}$ and $1 - [1 - (k-3)\mathcal{D}^{-1}]I(\mathcal{D} > k-3)$, we obtain the same estimators as obtained by the empirical Bayes method. We define these estimators as quasi-empirical Bayes estimators (QEBE). We approach the problem of combining several studies via “empirical Bayes” and “quasi-empirical Bayes” estimation. Initial attempts in this direction were made by Saleh et al. (2001), Saleh et al. (2006) and Saleh (2006), among others.

In this thesis, we are providing a new dimension of statistical methodology for meta analysis. The contents of the thesis are spread out as follows: Chapter 2 contains the notion of effect size based on means, the various estimators defined in the literature, and both small and large sample properties of the various estimators of the effect size. Chapter 3 introduces two new estimators of effect size in a single study with their finite sample properties. This chapter also introduces the “empirical Bayes ” and “quasi-empirical Bayes” methods of meta analysis for a series of independent studies. We also investigate the asymptotic properties of the proposed estimators and we present analysis of risk of the estimators. Tables and graphs of the asymptotic relative efficiency of the estimators are presented. Chapter 3 also contains the variance-stabilizing transformation and the estimators based on this transformation and their properties as well. In the last section of this chapter we study the effect size estimation in the discrete distribution and present the meta analysis of polls. In chapter 4, multivariate empirical Bayes and quasi-empirical Bayes methods have been applied for multiple-study. Properties of the estimators are investigated. In chapter 5, we introduce new effect size estimators based on order statistics and nonparametric, leading the subject towards robust estimation in both univariate and multivariate models. The concluding chapter contains open problems for future work that has come out of research in meta analysis. An appendix with all preliminaries that have been used throughout this thesis is provided.

Chapter 2

Effect Size and Various Estimators

2.1 Introduction

In this chapter we discuss the concept of “effect size” which is fundamental for the meta analysis of several studies of similar kinds. The notion of effect size naturally arises from the study of power of testing equality of a treatment group mean against a control group mean. Let $Y_1^E, \dots, Y_{n^E}^E$ be a sample of size n^E from the experimental group, let $Y_1^C, \dots, Y_{n^C}^C$ be a sample of size n^C from the control group and let $n = n^E + n^C$ be the total sample size. We assume that $Y_j^E \sim N(\mu^E, \sigma^2)$ and $Y_j^C \sim N(\mu^C, \sigma^2)$ respectively with different means and common variance σ^2 . Then the test of the null-hypothesis, $H_0 : \mu^E = \mu^C$ versus the alternative $H_A : \mu^E \neq \mu^C$ is obviously based on the famous Student’s test or the equivalent Fisher’s F-test given by

$$t_\nu = \frac{(\bar{Y}^E - \bar{Y}^C)}{(\frac{1}{n^E} + \frac{1}{n^C})^{\frac{1}{2}} s_p} \text{ or } F_{(1,\nu)} = \frac{(\bar{Y}^E - \bar{Y}^C)^2}{(\frac{1}{n^E} + \frac{1}{n^C}) s_p^2}$$

where

$$\bar{Y}^E = (n^E)^{-1} \left(\sum_{j=1}^{n^E} Y_j^E \right), \quad \bar{Y}^C = (n^C)^{-1} \left(\sum_{j=1}^{n^C} Y_j^C \right).$$

$$s_p^2 = \nu^{-1} \left\{ \sum_{j=1}^{n^E} (Y_j^E - \bar{Y}^E)^2 + \sum_{j=1}^{n^C} (Y_j^C - \bar{Y}^C)^2 \right\}, \nu = (n^E + n^C - 2).$$

The power function of the test based on t or F-statistics is given by $1 - G_{1,\nu}(F_\alpha : \delta^2)$ where

$$G_{1,\nu}(F_\alpha : \delta^2) = \sum_{r=0}^{\infty} \frac{e^{-\delta^2/2}}{r!} \left(\frac{\delta^2}{2}\right)^r I_{F_\alpha/(\nu+F_\alpha)}\left(\frac{1}{2} + r; \frac{\nu}{2}\right), \quad F_\alpha = t_{\frac{\alpha}{2}}^2$$

where $I_x(a, b)$ is the incomplete Beta function with $x = \frac{F_\alpha}{\nu+F_\alpha}$, F_α being the critical value for the test at α -level of significance. The quantity $\delta = \frac{\mu^E - \mu^C}{\sigma}$, the standardized mean difference, stands for the effect size. Its square δ^2 , is usually called the noncentrality parameter of the noncentral F-distribution with $(1, \nu)$ degrees of freedom D.F. It is the relative magnitude of the treatment effect with respect to the scale parameter σ . If δ is large, the power of the test is increased accordingly. Thus, the effect size δ determines the power of the test of H_0 . For quick assessment of the power of this test, one may estimate δ as

$$\tilde{\delta}_n = c(\nu) \left(\frac{\bar{Y}^E - \bar{Y}^C}{s_p} \right), \quad \text{where } c(\nu) = \frac{\Gamma(\nu/2)}{\sqrt{(\nu/2)} \Gamma(\frac{\nu-1}{2})}, \quad (2.1)$$

and determine the estimated power of the test using the expression given below

$$1 - G_{1,\nu}(F_\alpha : \delta^2) = 1 - \sum_{r=0}^{\infty} \frac{e^{-\tilde{\delta}_n^2/2}}{r!} \left(\frac{\tilde{\delta}_n^2}{2}\right)^r I_{F_\alpha/(\nu+F_\alpha)}\left(\frac{1}{2} + r; \frac{\nu}{2}\right).$$

The unbiased estimator, $\tilde{\delta}_n$ is one of the key elements in the synthesis of several studies. There are also biased estimators of the effect size, δ (see Hartung, Knapp

and Sinha (2008)) given below:

$$(a) \quad \text{MLE } \hat{\delta}_n^{(ML)} = \sqrt{\frac{\nu+2}{\nu}} \tilde{\delta}_n / c(\nu). \quad (2.2)$$

$$(b) \quad \text{Cohen's estimator, } \hat{\delta}_n^C = \frac{\bar{Y}^E - \bar{Y}^C}{s_p} = \frac{1}{c(\nu)} \tilde{\delta}_n. \quad (2.3)$$

$$(c) \quad \text{Minimum MSE estimator, } \delta_n^* = \frac{\nu-2}{\nu c^2(\nu)} \tilde{\delta}_n. \quad (2.4)$$

2.2 Small Sample Properties of the Estimators

First consider the unbiased estimator of the effect size

$$\tilde{\delta}_n = c(\nu) \frac{\bar{Y}^E - \bar{Y}^C}{s_p} = \left(\frac{1}{n^E} + \frac{1}{n^C} \right)^{\frac{1}{2}} c(\nu) \frac{\bar{Y}^E - \bar{Y}^C}{s_p \left(\frac{1}{n^E} + \frac{1}{n^C} \right)^{\frac{1}{2}}} = \left(\frac{1}{n^E} + \frac{1}{n^C} \right)^{\frac{1}{2}} c(\nu) t_\nu(.),$$

observe that

$$Z = \frac{\bar{Y}^E - \bar{Y}^C}{\sigma \left(\frac{1}{n^E} + \frac{1}{n^C} \right)^{\frac{1}{2}}} \sim N \left(\left(\frac{1}{n^E} + \frac{1}{n^C} \right)^{\frac{-1}{2}} \delta, 1 \right) \text{ and } u = \frac{\nu s_p^2}{\sigma^2} \sim \chi_\nu^2,$$

where s_p^2 is an estimate of σ^2 , independent of $\bar{Y}^E$ and $\bar{Y}^C$ implies that Z and u are independent. Hence,

$$t_\nu = \frac{Z}{\left(\frac{1}{\nu} u \right)^{\frac{1}{2}}} = \frac{\bar{Y}^E - \bar{Y}^C}{\left(\frac{1}{n^E} + \frac{1}{n^C} \right)^{\frac{1}{2}} s_p} \sim t_\nu \left(\left(\frac{1}{n^E} + \frac{1}{n^C} \right)^{\frac{-1}{2}} \delta \right),$$

then, $\frac{1}{c(\nu)} \left(\frac{1}{n^E} + \frac{1}{n^C} \right)^{\frac{-1}{2}} \tilde{\delta}_n \sim t_{(\nu)}(\Delta), \quad (2.5)$

$t_{(\nu)}(\Delta)$ represents a noncentral t-distribution with ν D.F. and noncentral parameter $\Delta = \left(\frac{1}{n^E} + \frac{1}{n^C} \right)^{\frac{-1}{2}} \delta$. Thus, mean and variance of this noncentral t-distribution are

given by

$$\begin{aligned} E(t_\nu) &= c^{-1}(\nu) \left(\frac{1}{n^E} + \frac{1}{n^C} \right)^{\frac{-1}{2}} \delta. \\ \text{var}(t_\nu) &= \frac{\nu}{(\nu-2)} + \left(\frac{1}{n^E} + \frac{1}{n^C} \right)^{-1} \delta^2 \left(\frac{\nu}{(\nu-2)} - \frac{1}{c^2(\nu)} \right). \end{aligned}$$

So that the mean of $\tilde{\delta}_n$ is given by

$$E(\tilde{\delta}_n) = c(\nu) \left(\frac{1}{n^E} + \frac{1}{n^C} \right)^{\frac{1}{2}} c^{-1}(\nu) \left(\frac{1}{n^E} + \frac{1}{n^C} \right)^{\frac{-1}{2}} \delta = \delta. \quad (2.6)$$

Hence, the bias of $\tilde{\delta}_n$, $b(\tilde{\delta}_n) = 0$ and the variance is given by

$$\text{var}(\tilde{\delta}_n) = \frac{\nu c^2(\nu)}{(\nu-2)} \left[\left(\frac{1}{n^E} + \frac{1}{n^C} \right) + \delta^2 \left(1 - \frac{\nu-2}{\nu c^2(\nu)} \right) \right] \quad (2.7)$$

$$= \left(\frac{1}{n^E} + \frac{1}{n^C} \right) \left[\frac{\nu c^2(\nu)}{(\nu-2)} (1 + \Delta^2) - \Delta^2 \right]. \quad (2.8)$$

Notice that $\tilde{\delta}_n$ is a consistent estimator of δ . The reason is that as $n = n^E + n^C \rightarrow \infty$, $\frac{\nu c^2(\nu)}{(\nu-2)} \rightarrow 1$ and

$$\text{var}(\tilde{\delta}_n) = \frac{1}{n^E} + \frac{1}{n^C}.$$

Let $\frac{n^E}{n} \rightarrow \pi^E$ and $\frac{n^C}{n} \rightarrow \pi^C$ where π^E, π^C remain fixed as $n \rightarrow \infty$, then

$$\text{var}(\tilde{\delta}_n) = \frac{1}{n} \left(\frac{1}{\pi^E} + \frac{1}{\pi^C} \right) \rightarrow 0 \text{ as } n \rightarrow \infty.$$

Next, we consider the MLE of δ , where

$$\begin{aligned} E(\hat{\delta}_n^{(ML)}) &= \sqrt{\frac{\nu+2}{\nu}} \frac{\delta}{c(\nu)}, \text{ bias}(\hat{\delta}_n^{(ML)}) = \delta \left[\sqrt{\frac{\nu+2}{\nu}} \frac{1}{c(\nu)} - 1 \right] \\ \text{var}(\hat{\delta}_n^{(ML)}) &= \frac{\nu+2}{\nu-2} \left[\left(\frac{1}{n^E} + \frac{1}{n^C} \right) + \delta^2 \left(1 - \frac{\nu-2}{\nu c^2(\nu)} \right) \right]. \end{aligned}$$

So that the MSE of $\hat{\delta}_n^{(ML)}$ is given by

$$M(\hat{\delta}_n^{(ML)}) = \frac{(\nu + 2)}{(\nu - 2)} \left[\left(\frac{1}{n^E} + \frac{1}{n^C} \right) + \delta^2 \left(1 - 2 \frac{(\nu - 2)}{\sqrt{\nu(\nu + 2)}} \frac{1}{c(\nu)} + \frac{\nu - 2}{\nu + 2} \right) \right].$$

Now, we consider the Cohen's estimator (1977),

$$\begin{aligned} E(\hat{\delta}_n^C) &= \frac{1}{c(\nu)} \delta, \text{ and } \text{bias}(\hat{\delta}_n^C) = \delta \left(\frac{1}{c(\nu)} - 1 \right) \\ \text{var}(\hat{\delta}_n^C) &= \frac{\nu}{(\nu - 2)} \left[\left(\frac{1}{n^E} + \frac{1}{n^C} \right) + \delta^2 \left(1 - \frac{(\nu - 2)}{\nu c^2(\nu)} \right) \right]. \\ M(\hat{\delta}_n^C) &= \frac{\nu}{(\nu - 2)} \left[\left(\frac{1}{n^E} + \frac{1}{n^C} \right) + \delta^2 \left(1 - \frac{2(\nu - 2)}{\nu c(\nu)} + \frac{(\nu - 2)}{\nu} \right) \right]. \end{aligned}$$

Similarly, consider the minimum MSE estimator, δ_n^* ,

$$\begin{aligned} E(\delta_n^*) &= \frac{(\nu - 2)}{\nu c^2(\nu)} \delta, \text{ and } \text{bias}(\delta_n^*) = \delta \left[\frac{(\nu - 2)}{\nu c^2(\nu)} - 1 \right] \\ \text{var}(\delta_n^*) &= \frac{(\nu - 2)}{\nu c^2(\nu)} \left[\left(\frac{1}{n^E} + \frac{1}{n^C} \right) + \delta^2 \left(1 - \frac{(\nu - 2)}{\nu c^2(\nu)} \right) \right]. \\ M(\delta_n^*) &= \frac{(\nu - 2)}{\nu c^2(\nu)} \left[\left(\frac{1}{n^E} + \frac{1}{n^C} \right) + \delta^2 \left(\frac{\nu c^2(\nu)}{(\nu - 2)} - 1 \right) \right]. \end{aligned}$$

Figure 2.1 presents a graphic illustration of the mean-squared error of each of the four estimators as a function of δ^2 for $n^E = 4, n^C = 5$. The figure illustrates that the minimum MSE estimator, δ_n^* dominates the other estimators. We conclude this section with the discussion of exact $100(1-\alpha)\%$ confidence interval for δ based on Cohen's estimate. We have seen in equation (2.5) that

$$\frac{1}{c(\nu)} \left(\frac{1}{n^E} + \frac{1}{n^C} \right)^{-\frac{1}{2}} \tilde{\delta}_n \sim t_\nu(\Delta),$$

consequently, for Cohen's estimate defined in (2.3), we obtain

$$\left(\frac{1}{n^E} + \frac{1}{n^C} \right)^{-\frac{1}{2}} \hat{\delta}_n^C \sim t_\nu \left(\left(\frac{1}{n^E} + \frac{1}{n^C} \right)^{-\frac{1}{2}} \delta \right).$$

Figure 2.1: Exact MSE of four estimators of effect size as a function of δ^2 .

That is, $\left(\frac{1}{n^E} + \frac{1}{n^C}\right)^{-\frac{1}{2}} \hat{\delta}_n^C$ has the noncentral t-distribution with ν D.F. and noncentrality parameter $\left(\frac{1}{n^E} + \frac{1}{n^C}\right)^{-\frac{1}{2}} \delta$. Now, to find the exact $100(1 - \alpha)\%$ confidence interval, (δ_L, δ_U) we solve for δ_L and δ_U from the two equations using the cdf of the noncentral F-distribution as

$$G_{(1,\nu)}(\sqrt{\tilde{n}} \hat{\delta}_n^C, \nu, \sqrt{\tilde{n}} \delta_L) = \frac{\alpha}{2}, \quad G_{(1,\nu)}(\sqrt{\tilde{n}} \hat{\delta}_n^C, \nu, \sqrt{\tilde{n}} \delta_U) = 1 - \frac{\alpha}{2}.$$

where,

$$\begin{aligned} G_{(1,\nu)}(\sqrt{\tilde{n}} \hat{\delta}_n^C, \nu, \sqrt{\tilde{n}} \delta) &= \sum_{r=0}^{\infty} \frac{e^{-\tilde{n}\delta^2/2}}{r!} \left(\frac{\tilde{n}\delta^2}{2}\right)^r I_{t_a^2/(\nu+t_a^2)}\left(\frac{1}{2} + r; \frac{\nu}{2}\right) \\ \tilde{n} &= \left(\frac{1}{n^E} + \frac{1}{n^C}\right)^{-1}. \end{aligned}$$

There is no formula that can give us these values of δ directly, and statistical software typically uses an iterative algorithmic search to find the values see Lecourt (2007) and Cumming and Finch among others.

2.3 Large Sample Properties of the Unbiased Estimator of the Effect Size

Hedges and Olkin (1981-1985) explained that although the four estimators of effect size in (2.1)-(2.4) differ in finite samples, the differences in performance of the estimators are small for 16 or more degrees of freedom. Indeed, the differences are appreciable only for very small degrees of freedom, which is unrealistic in most applications. Many applications involve sample sizes of at least 10 subjects per group, and in these cases the differences among the four estimators are likely to be negligible. We choose to work with the unbiased estimator $\tilde{\delta}_n$ since it is widely used as the effect size estimate in meta analysis. Now, we explain its large sample properties.

For large ν (δ being fixed), see Johnson and Kotz (1970), we obtain

$$\text{var}(t_\nu(\Delta)) \rightarrow 1 + \frac{1}{2} \left(\frac{1}{n^E} + \frac{1}{n^C} \right)^{-1} \delta^2 \nu^{-1}. \quad (2.9)$$

Using (2.5) and the fact that $c(\nu) \rightarrow 1$ as $n \rightarrow \infty$ then the large sample variance of $\tilde{\delta}_n$ is given by (see Hedges (1983))

$$\text{var}(\tilde{\delta}_n) = \left(\frac{1}{n^E} + \frac{1}{n^C} \right) + \frac{\delta^2}{2(n^E + n^C - 2)}. \quad (2.10)$$

Thus, asymptotic $100(1 - \alpha)\%$ confidence interval for δ is given by

$$\tilde{\delta}_n \pm z_{\alpha/2} \sqrt{\left(\frac{1}{n^E} + \frac{1}{n^C} \right) + \frac{\tilde{\delta}_n^2}{2(n^E + n^C - 2)}}, \quad (2.11)$$

where $z_{\alpha/2}$ is the two-tailed critical value of the standard normal distribution.

Example 2.3.1. Suppose that a study with $n^E = n^C = 20$ yields a sample effect size value of $\tilde{\delta}_n = 0.580$, then the 95% confidence interval for δ is given by

$$0.580 \pm 1.96 \sqrt{0.1 + \frac{(0.580)^2}{76}} = (-0.0533, 1.2133). \quad (2.12)$$

We have noticed that both finite and large sample variance of the estimators of effect size depend on the parameter itself. To avoid this nuisance parameter problem one takes recourse to the variance stabilizing transformation.

This variance-stabilizing transformation was studied by Hedges and Olkin (1983) in the balanced case $\frac{n^E}{n^C} = 1$, and is closely related to a variance-stabilizing transformation for the noncentral t-distribution that was studied by Laubscher (1960).

Fisher's transformation

Let r be the sample correlation and ρ be the population correlation, define

$$\theta = \frac{1}{2} \log \frac{1 + \rho}{1 - \rho}. \quad (2.13)$$

Then the Maximum likelihood estimator of θ is given by

$$\tilde{\theta} = \frac{1}{2} \log \frac{1 + r}{1 - r}. \quad (2.14)$$

The large sample distribution of $\tilde{\theta}$ is

$$\sqrt{n-3} (\tilde{\theta} - \theta) \sim N(0, 1). \quad (2.15)$$

For a proof of this result, see Rao (1973) pages 432-433 or Anderson (1984) pages 78-79 .

Consequently, the $100(1 - \alpha)\%$ confidence interval for θ is given by

$$\theta_L = \tilde{\theta} - (n-3)^{\frac{-1}{2}} z_{\alpha/2}, \quad \theta_U = \tilde{\theta} + (n-3)^{\frac{-1}{2}} z_{\alpha/2}, \quad (2.16)$$

Based on the confidence limit of θ given above, the confidence limit of ρ are obtained as

$$\rho_L = [\exp(2\theta_L) - 1] / [\exp(2\theta_L) + 1] \quad (2.17)$$

$$\rho_U = [\exp(2\theta_U) - 1] / [\exp(2\theta_U) + 1]. \quad (2.18)$$

Kraemer and Paik (1979) and Kraemer (1983) explained the fact that the distribution of the effect size estimate (the noncentral t-distribution) and the distribution

of the sample correlation coefficient are related. Kraemer showed that if

$$r = \frac{\tilde{\delta}_n}{\sqrt{\tilde{\delta}_n^2 + m}}, \quad (2.19)$$

$$\text{and } \rho = \frac{\delta}{\sqrt{\delta^2 + m}}, \quad m = \frac{n(n-2)}{n^E n^C}, \quad (2.20)$$

then $u = \frac{r-\rho}{1-r\rho}$ is asymptotically distributed according to the null distribution of the sample correlation coefficient, that is, $\frac{\nu^{\frac{1}{2}}u}{(1-u^2)^{\frac{1}{2}}} \xrightarrow{\mathcal{D}} t_\nu$, where $\xrightarrow{\mathcal{D}}$ stands for “converges in distribution”. Consequently, the confidence limit of the effect size δ are obtained as

$$\delta_L = \frac{m^{\frac{1}{2}}\rho_L}{(1-\rho_L^2)^{\frac{1}{2}}} \quad (2.21)$$

$$\delta_U = \frac{m^{\frac{1}{2}}\rho_U}{(1-\rho_U^2)^{\frac{1}{2}}}. \quad (2.22)$$

Example 2.3.2. Return to our example of a study with $n^E = n^C = 20$ yields a sample effect size value of $\tilde{\delta}_n = 0.580$. To calculate a 95% confidence interval for δ , we first calculate r using (2.19), we get $r = 0.27851$, then calculate $\tilde{\theta}$ using (2.14), $\tilde{\theta} = 0.1273$ then using (2.16) we obtain $\theta_L = -0.1949$ and $\theta_U = 0.4495$, using (2.17)-(2.18) we obtain $\rho_L = -0.1925$ and $\rho_U = -0.4215$ and the 95% confidence interval for δ as $(-0.3823, 0.9059)$.

In this chapter, we introduced the concept of effect size and its importance in the development of meta analysis. Side by side we presented various unrestricted estimators of the effect size and their MSE comparisons. Further, we discussed the confidence intervals and how to calculate using large sample theory as the small sample calculations are computer intensive. Also, we discussed the variance stabilizing statistics (estimators) and methods of confidence intervals particularly, consider correlation type statistics using Fisher’s transformations.

Chapter 3

Effect Size Estimation: Univariate Models

3.1 Introduction

We deal with the estimation problem of the effect size under uncertain prior information for both finite and large sample sizes. We estimate the effect size for each study, and then test the hypothesis of having a common effect size among studies . We start with estimation of effect size in one study problem, then we present estimation of effect size in multiple study problem. For one study problem we propose two new estimators in two cases (i) variance is known (ii) variance is unknown. We compare the estimators with respect to their bias and mean square errors. Coverage probability of the confidence interval based on each estimator is presented in both cases. We study Bayes and empirical Bayes estimators introduced by Efron (2010) and compare with the proposed estimators. In multiple study problem with unknown variance, we present five estimators, namely, the unrestricted estimator (UE), restricted estimator (RE), preliminary test estimator (PTE), James-Stein estimator (JS) and positive rule Stein estimator. The comparative properties of the estimators are studied.

Tables for optimum selection of level of the test are presented for each of those problems. Further tables and graphs are presented giving the relative efficiencies of the restricted, preliminary test, and James-Stein estimators relative to the unrestricted estimator. One study problem is considered in section 3.2. section 3.3 presents Bayes and empirical Bayes estimators and comparisons. In section 3.4 multiple-study problem is presented. In section 3.6 we introduce Variance-stabilizing Transformation. Finally in section 3.6 we study meta analysis of polls.

3.2 One Study Problem

Let $Y_1^E, \dots, Y_{n^E}^E$ be the experimental outcomes from $N(\mu^E, \sigma^2)$ and $Y_1^C, \dots, Y_{n^C}^C$ be the control outcome from $N(\mu^C, \sigma^2)$, in this section we propose two new estimators of $\delta = \frac{\mu^E - \mu^C}{\sigma}$, when it is priori suspected that δ may be equal to zero, based on the experimental and control outcomes given by the sample of unequal sizes, n^E and n^C respectively when (i) variance is known (ii) variance is unknown.

3.2.1 Case 1: Variance is Known

We propose two new estimators of $\delta = \frac{\mu^E - \mu^C}{\sigma}$ in addition to the well-known unbiased estimator, when one may suspect that δ may be equal to zero. Basically, the unbiased estimation of δ is complete when we have

$$\tilde{\delta}_n^* = \frac{\bar{Y}^E - \bar{Y}^C}{\sigma}$$

However, we formulate the hypothesis that δ may be equal to zero. That is to say, we suspect that the two means are equal, meaning that there is no effect of the experimental ingredients. To remove this vague hypothesis, we carry on a statistical

test of hypothesis:

$$H_0 : \mu^E = \mu^C \quad \text{Vs} \quad H_A : \mu^E \neq \mu^C$$

based on the test statistics

$$Z^2 = \frac{(\bar{Y}^E - \bar{Y}^C)^2}{\sigma^2(\frac{1}{n^E} + \frac{1}{n^C})}. \quad (3.1)$$

Now, we propose the following two estimators as in Saleh (2006) (Chapter 3 , Sections 3.2, 3.5) as follows

The Proposed Estimators

$$\hat{\delta}_n^{*PT} = \tilde{\delta}_n^* - \tilde{\delta}_n^* I(Z^2 \leq c_\alpha). \quad (3.2)$$

$$\hat{\delta}_n^{*s} = \tilde{\delta}_n^* - d|Z|^{-1}\tilde{\delta}_n^*, \quad d > 0. \quad (3.3)$$

where $I_x(A)$ stands for the indicator function of the set A , and c_α is the upper α -level critical value of the central chi-square distribution with one degree of freedom and d is called the shrinkage constant, free of any parameter of the distribution,

The estimator $\hat{\delta}_n^{*PT}$ is called “preliminary estimator” (PTE) due to Bancroft (1944), and $\hat{\delta}_n^{*s}$ is called “shrinkage type estimator”(SE)

Bias and Mean Square Error Expressions

In this section, we consider the properties of the estimators when σ^2 is known. Let us call $\tilde{\delta}_n^* = \frac{\bar{Y}^E - \bar{Y}^C}{\sigma}$, the unrestricted estimator (UE), then clearly, under the assumption of normal theory given above, $\tilde{\delta}_n^* \sim N\{\frac{\mu^E - \mu^C}{\sigma}, (\frac{1}{n^E} + \frac{1}{n^C})\}$ so that $E(\tilde{\delta}_n^*) = \delta$ giving the bias and the variance as

$$b_1(\tilde{\delta}_n^*) = 0 \quad \text{and the var}(\tilde{\delta}_n^*) = \left(\frac{1}{n^E} + \frac{1}{n^C} \right) = M_1(\tilde{\delta}_n^*).$$

$M_1(\tilde{\delta}_n^*)$ stands for the mean squared error (MSE) of $\tilde{\delta}_n^*$. For the test of H_0 , one uses the test-statistics Z^2 in (3.1) where Z^2 has chi-square distribution with one degree of freedom D.F. and noncentrality parameter $\frac{1}{2}\Delta^2$ where Δ^2 defined by $\Delta^2 = \left(\frac{1}{n^E} + \frac{1}{n^C}\right)^{-1} \delta^2$. Under H_0 , it follows a central chi-square distribution. Thus, at the α -level of significance we reject H_0 if $Z^2 \geq c_\alpha$.

Now, consider the PTE in (3.2) and obtain the expected value of $\hat{\delta}_n^{*PT}$ using (A.3) of Theorem A.2.5 as

$$E(\hat{\delta}_n^{*PT}) = \delta - \delta H_3(c_\alpha, \Delta^2),$$

where $H_3(., \Delta^2)$ is the cdf of a chi-square distribution with 3 D.F. and noncentrality parameter $\frac{1}{2}\Delta^2$. Hence the bias expression is given by

$$b_2(\hat{\delta}_n^{*PT}) = -\delta H_3(c_\alpha, \Delta^2). \quad (3.4)$$

Correspondingly, we obtain the MSE using Theorem A.2.5 as

$$M_2(\hat{\delta}_n^{*PT}) = \left(\frac{1}{n^E} + \frac{1}{n^C}\right) \{1 - H_3(c_\alpha, \Delta^2) + \Delta^2[2H_3(c_\alpha, \Delta^2) - H_5(c_\alpha, \Delta^2)]\}.$$

Similarly, for the shrinkage type estimator defined in (3.3), we obtain the bias using A.1.1 as

$$\begin{aligned} b_3(\hat{\delta}_n^{*s}) &= -\sqrt{\left(\frac{1}{n^E} + \frac{1}{n^C}\right)} dE\left[\frac{Z}{|Z|}\right], \\ &= -\sqrt{\left(\frac{1}{n^E} + \frac{1}{n^C}\right)} d[2\Phi(\Delta) - 1], \\ \text{where } Z &= \frac{\bar{Y}^E - \bar{Y}^C}{\sigma \sqrt{\left(\frac{1}{n^E} + \frac{1}{n^C}\right)}}. \end{aligned}$$

Here $\Phi(\cdot)$ is the cdf of $N(0,1)$. The MSE of $\hat{\delta}_n^{s*}$ may be obtained as

$$M_3(\hat{\delta}_n^{s*}) = \left(\frac{1}{n^E} + \frac{1}{n^C} \right) [1 + d^2 - 2d\sqrt{2/\pi} e^{-\Delta^2/2}].$$

The value of d which minimizes $M_3(\hat{\delta}_n^{s*})$ is $d = \sqrt{2/\pi} e^{-\Delta^2/2}$, which depends on Δ^2 . To make it free of the parameter, Δ^2 we choose $d^* = \sqrt{2/\pi}$. Hence, the resulting MSE becomes

$$M_3(\hat{\delta}_n^{s*}) = \left(\frac{1}{n^E} + \frac{1}{n^C} \right) \left\{ 1 - \frac{2}{\pi} [2e^{-\Delta^2/2} - 1] \right\}.$$

We summarize the previous results in the following theorem .

Theorem 3.2.1. *If σ^2 is known, then the bias and MSE expressions of the three estimators are given by*

$$\begin{aligned} b_1(\tilde{\delta}_n^*) &= 0, \\ M_1(\tilde{\delta}_n^*) &= \left(\frac{1}{n^E} + \frac{1}{n^C} \right), \\ b_2(\hat{\delta}_n^{*PT}) &= -\delta H_3(c_\alpha, \Delta^2), \\ M_2(\hat{\delta}_n^{*PT}) &= \left(\frac{1}{n^E} + \frac{1}{n^C} \right) \{1 - H_3(c_\alpha, \Delta^2) + \Delta^2[2H_3(c_\alpha, \Delta^2) - H_5(c_\alpha, \Delta^2)]\}, \\ b_3(\hat{\delta}_n^{s*}) &= -\sqrt{\left(\frac{1}{n^E} + \frac{1}{n^C} \right)} d[2\Phi(\Delta) - 1], \\ M_3(\hat{\delta}_n^{s*}) &= \left(\frac{1}{n^E} + \frac{1}{n^C} \right) \left\{ 1 - \frac{2}{\pi} [2e^{-\Delta^2/2} - 1] \right\}. \end{aligned}$$

MSE-based Relative Efficiency

We consider the unrestricted estimators as the basis for comparing the others. Thus, we may define the MSE-based relative efficiency (MRE) of $\hat{\delta}_n^{*PT}$ relative to $\tilde{\delta}_n^*$ as

$$\begin{aligned} MRE(\hat{\delta}_n^{*PT}, \tilde{\delta}_n^*) &= \frac{MSE(\tilde{\delta}_n^*)}{MSE(\hat{\delta}_n^{*PT})} \\ &= \{1 - H_3(c_\alpha, \Delta^2) + \Delta^2[2H_3(c_\alpha, \Delta^2) - H_5(c_\alpha, \Delta^2)]\}^{-1}. \end{aligned}$$

Figure 3.1: MSE-based relative efficiencies of SE and PTE for different values of α

$MRE(\hat{\delta}_n^{*PT}, \tilde{\delta}_n^*)$ may be written as a function of (α, Δ^2) . Figure 3.1 illustrates the graph of $MRE(\alpha, \Delta^2)$, as a function of Δ^2 . For fixed α , it decreases, crossing the 1-line to a minimum at $\Delta^2 = \Delta_{min}^2(\alpha)$, then increases towards the 1-line as $\Delta^2 \rightarrow \infty$. The maximum of $MRE(\alpha, \Delta^2)$ occurs at $\Delta^2 = 0$ with the value

$$MRE(\alpha, 0) = [1 - H_3(c_\alpha, 0)]^{-1} \quad (3.5)$$

for all $\alpha \in A$, the set of all possible values of α . The value of the max of $MRE(\alpha, \Delta^2)$ decreases as α increases. Intersection of the graph of $MRE(\alpha, \Delta^2)$ with the 1-line

occurs at Δ_1^2 . Thus if $0 \leq \Delta^2 \leq \Delta_1^2$, then $\hat{\delta}_n^{*PT}$ performs better than $\tilde{\delta}_n^*$, and if $\Delta^2 \geq \Delta_1^2$, then $\tilde{\delta}_n^*$ performs better. The cut-off point Δ_1^2 may be determined by

$$\Delta_1^2 = \frac{H_3(c_\alpha, \Delta^2)}{[2H_3(c_\alpha, \Delta^2) - H_5(c_\alpha, \Delta^2)]} \quad (3.6)$$

In order to obtain a PTE with minimum guaranteed RE, say E_0 , we adopt the following procedure : if $0 \leq \Delta^2 \leq \Delta_1^2$, we always choose $\tilde{\delta}_N$ since, $MRE(\alpha, \Delta^2) \geq 1$ in this interval. However, in general Δ^2 is unknown, and there is no way one can choose an estimator which is uniformly best. Thus, we look for an estimator with minimum guaranteed RE, say, E_0 , and look for a suitable α from the set $A = \{\alpha | MRE(\alpha, \Delta^2) \geq E_0\}$ which is well-defined and satisfies the equation

$$\max_{\alpha} \min_{\Delta^2} MRE(\alpha, \Delta^2) = E_0 \text{ for some } \alpha \in A.$$

Then solution α^* gives the optimum α -level satisfying the minimum guaranteed relative efficiency. Now, we consider the relative efficiency of $\hat{\delta}_n^{*s}$ versus $\tilde{\delta}_n^*$ as

$$MRE(\hat{\delta}_n^{*s}, \tilde{\delta}_n^*) = \{1 - \frac{2}{\pi}[2e^{-\frac{\Delta^2}{2}} - 1]\}^{-1}$$

If $\Delta^2 = 0$, $MRE(\hat{\delta}_n^{*s}, \tilde{\delta}_n^*) = \frac{\pi}{\pi-2}$ and if $\Delta^2 \rightarrow \infty$ the $MRE(\hat{\delta}_n^{*s}, \tilde{\delta}_n^*) = \frac{\pi}{\pi+2}$. The $MRE(\hat{\delta}_n^{*s}, \tilde{\delta}_n^*)$ as a function of Δ^2 crosses the 1-line at $\Delta^2 = \ln 4$ and drops to $\frac{\pi}{\pi+2}$. Thus, for $0 \leq \Delta^2 \leq \ln 4$, $\hat{\delta}_n^{*s}$ performs better than $\tilde{\delta}_n^*$, otherwise $\tilde{\delta}_n^*$ performs better outside this interval.

Table 3.1 gives the performance of $\hat{\delta}_n^{*s}$ and $\hat{\delta}_n^{*PT}$ (for each selected level of significance).

Columns of Table 3.1 contain the maximum relative efficiency of the PTE, E_{max} , the minimum relative efficiency, E_0 , of the PTE at $\Delta^2 = \Delta_0^2$ which has been recorded for each $\alpha=0.05(0.10)0.45$ with the corresponding efficiency ($E_{\Delta_0^2}$) of SE for $\Delta^2 = \Delta_0^2$.

Table 3.1: Maximum and minimum MSE-based relative efficiency of PTE for selected α and relative efficiency of SE at Δ_0^2

α	Δ_0^2	$E_{\max}$	E_0	$E_{\Delta_0^2}$
0.05	4.75	3.5829	0.4058	0.6185
0.15	3.65	1.7936	0.5950	0.6290
0.25	3.30	1.3820	0.7248	0.6388
0.35	2.95	1.2022	0.8220	0.6465
0.45	2.85	1.1073	0.8932	0.6525
0.50	2.75	1.0768	0.9202	0.7654

3.2.2 Confidence Intervals

In this section, we consider the fixed length $100(1 - \beta)\%$ confidence interval for δ when the variance is known, so that the confidence intervals based on the three estimators, namely, unrestricted estimator, preliminary test estimator and shrinkage type estimator can be compared appropriately. We may write the confidence interval using $\tilde{\delta}_n^*$ as the Neyman-Pearson solution

$$C^0(\tilde{\delta}_n^*) = \{\delta : \sqrt{\tilde{n}} |\delta - \tilde{\delta}_n^*| \leq z_{\beta/2}\}$$

where $z_{\beta/2}$ is the upper $\beta/2$ -level critical value from the standard normal distribution, $\tilde{n} = (\frac{1}{n^E} + \frac{1}{n^C})^{-1}$, where

$$P_0\{\sqrt{\tilde{n}} |\delta - \tilde{\delta}_n^*| \leq z_{\beta/2}\} = 1 - \beta = P_0\{\tilde{n}(\delta - \tilde{\delta}_n^*)^2 \leq z_{\beta/2}^2\}.$$

Next, we consider the $100(1 - \beta)\%$ confidence interval for δ using the PTE, $\hat{\delta}_n^{*PT}$ as the solution

$$\begin{aligned} C^{PT}(\hat{\delta}_n^{*PT}) &= \{\delta : \sqrt{\tilde{n}} |\delta - \hat{\delta}_n^{*PT}| \leq z_{\beta/2}\} \\ &= \{\sqrt{\tilde{n}} |\delta| < z_{\beta/2}; |Z| < z_{\alpha/2}\} \\ &\quad + \{\sqrt{\tilde{n}} |\delta - \tilde{\delta}_n^*| < z_{\beta/2}; |Z| > z_{\alpha/2}\} \end{aligned}$$

The computational formula of the coverage probability of the set $C^{*PT}(\hat{\delta}_n^{*PT})$ is given by

$$\begin{aligned}
P_{\Delta^2}\{\sqrt{\tilde{n}} |\delta - \hat{\delta}_n^{*PT}| \leq z_{\beta/2}\} &= H_1(\chi_{1,\alpha}^2; \Delta^2)I(\Delta^2 < z_{\beta/2}^2) + (1 - \beta) \\
&\quad - \Phi[\min\{z_{\beta/2}, z_{\alpha/2} - \Delta\}] \\
&\quad + \Phi[\max\{-z_{\beta/2}, -(z_{\alpha/2} + \Delta)\}] \quad (3.7)
\end{aligned}$$

Graphs of coverage probability of 95% and 90% C.I. and the test levels $\alpha = 0.05, 0.1$ and 0.2 of the set $C^{PT}(\hat{\delta}_n^{*PT})$ are shown in Figure 3.2 and 3.3 along with the coverage probability of set based on SE.

From the graphs one may see that from 0 to $z_{\beta/2}$, the coverage probability is greater than or equal to $1 - \beta$, and due to the discontinuity at $\Delta = z_{\beta/2}$ the coverage probability drops to its minimum (depending on α, β) then increases towards $1 - \beta$ at $\Delta = z_{\beta/2} + z_{\alpha/2}$ and eventually goes to $1 - \beta/2$ as $\Delta \rightarrow \infty$.

Tabular values of the coverage probability are presented in Table 3.2 for the 95% and 90% confidence sets based on PTE (and the test levels $\alpha = 0.05, 0.1, 0.2$) and SE.

Table 3.2: Coverage probability of 95% and 90% C.I. of the sets $C^{PT}(\hat{\delta}_n^{*PT})$ and $C^s(\hat{\delta}_n^s)$

$\Delta \backslash \alpha$	95%				90 %			
	0.05	0.1	0.2	SE	0.05	0.1	0.2	SE
0	0.9500	0.9500	0.9500	0.9942	0.9500	0.9000	0.9000	0.9854
1	0.9734	0.9703	0.9637	0.9942	0.9484	0.9459	0.9387	0.9456
2	0.4909	0.6137	0.7387	0.8943	0.4659	0.5887	0.7137	0.8096
3	0.8258	0.8873	0.9321	0.8729	0.8000	0.8623	0.9071	0.7882
4	0.9543	0.9657	0.9717	0.8716	0.9293	0.9407	0.9467	0.7869
5	0.9736	0.9746	0.9750	0.8716	0.9488	0.9496	0.9498	0.7869
6	0.9750	0.9750	0.9750	0.8716	0.9499	0.9499	0.9499	0.7869
7	0.9750	0.9750	0.9750	0.8716	0.9499	0.9499	0.9500	0.7869
8	0.9750	0.9750	0.9750	0.8716	0.9500	0.9500	0.9500	0.7869
9	0.9750	0.9750	0.9750	0.8716	0.9500	0.9500	0.9500	0.7869
10	0.9750	0.9750	0.9750	0.8716	0.9500	0.9500	0.9500	0.7869

Figure 3.2: Coverage probabilities of 95% C.I. based on PTE ($\alpha = 0.05, 0.1, 0.2$) and SE

Figure 3.3: Coverage probabilities of 90% C.I. based on PTE ($\alpha=0.05, 0.1, 0.2$) and SE

Now, we state the properties of the coverage probability of the set $C^{PT}(\hat{\delta}_n^{PT})$ in the following theorem.

Theorem 3.2.2.

- If
- (i) $0 \leq \Delta^2 \leq z_{\beta/2}^2$, $P\{C^{PT}(\hat{\delta}_n^{PT})\} \geq 1 - \beta$.
 - (ii) $z_{\beta/2}^2 \leq \Delta^2 < (z_{\beta/2} + z_{\alpha/2})^2$, $P\{C^{PT}(\hat{\delta}_n^{PT})\} \leq 1 - \beta$.
 - (iii) $\Delta^2 = (z_{\beta/2} + z_{\alpha/2})^2$, $P\{C^{PT}(\hat{\delta}_n^{PT})\} = 1 - \beta$.
 - (iv) $\Delta^2 = 0$ and $\min(z_{\beta/2}, z_{\alpha/2}) = z_{\alpha/2}$, $P\{C^{PT}(\hat{\delta}_n^{PT})\} = 1 - \beta$.
 - (v) $\Delta^2 \rightarrow \infty$, $P\{C^{PT}(\hat{\delta}_n^{PT})\} = 1 - \beta/2$.

Proof. Note that

$$\begin{aligned}
 (i) \quad P_{\Delta^2}\{\sqrt{\tilde{n}} |\delta - \hat{\delta}_n^{PT}| \leq z_{\beta/2}\} &= P_{\Delta^2}\{\tilde{n}(\delta - \tilde{\delta}_n^* I(Z^2 \geq \chi_{1,\alpha}^2))^2 \leq z_{\beta/2}^2\} \\
 &= P_{\Delta^2}\{\Delta^2 < z_{\beta/2}^2, Z^2 < \chi_{1,\alpha}^2\} \\
 &\quad + P\{\tilde{n}(\delta - \tilde{\delta}_n^*)^2 \leq z_{\beta/2}^2, Z^2 \geq \chi_{1,\alpha}^2\} \\
 &= H_1(\chi_{1,\alpha}^2; \Delta^2) I(\Delta^2 < z_{\beta/2}^2) \\
 &\quad + P\{\tilde{n}(\delta - \tilde{\delta}_n^*)^2 \leq z_{\beta/2}^2, Z^2 \geq \chi_{1,\alpha}^2\} \\
 &\geq P_{\Delta^2}\{\tilde{n}(\delta - \tilde{\delta}_n^*)^2 \leq z_{\beta/2}^2, Z^2 \leq \chi_{1,\alpha}^2\} \\
 &\quad + P\{\tilde{n}(\delta - \tilde{\delta}_n^*)^2 \leq z_{\beta/2}^2, Z^2 \geq \chi_{1,\alpha}^2\} \\
 &\geq P_{\Delta^2}\{\tilde{n}(\delta - \tilde{\delta}_n^*)^2 \leq z_{\beta/2}^2\} = 1 - \beta.
 \end{aligned}$$

Hence, we have proved that if $\Delta^2 < z_{\beta/2}^2$, then

$$P_{\Delta^2}\{\sqrt{\tilde{n}} |\delta - \hat{\delta}_n^{PT}| \leq z_{\beta/2}\} \geq 1 - \beta.$$

(ii) If $z_{\beta/2}^2 \leq \Delta^2 < (z_{\beta/2} + \chi_{1,\alpha})^2$, we have

$$\begin{aligned} P_{\Delta^2}\{\delta \in C^{PT}(\hat{\delta}_n^{*PT})\} &= P_{\Delta^2}\{\tilde{n}(\delta - \tilde{\delta}_n^*)^2 \leq z_{\beta/2}^2; Z^2 \geq \chi_{1,\alpha}^2\} \\ &\leq P_{\Delta^2}\{\tilde{n}(\delta - \tilde{\delta}_n^*)^2 \leq z_{\beta/2}^2\} = 1 - \beta. \end{aligned}$$

Hence, we have proved that if $z_{\beta/2}^2 \leq \Delta^2 < (z_{\beta/2} + \chi_{1,\alpha})^2$, then

$$P_{\Delta^2}\{\sqrt{\tilde{n}} |\delta - \hat{\delta}_n^{*PT}| \leq z_{\beta/2}\} \leq 1 - \beta.$$

(iii) If $\Delta^2 = (z_{\beta/2} + z_{\alpha/2})^2$, then

$$\begin{aligned} P_{\Delta^2}\{\sqrt{\tilde{n}} |\delta - \hat{\delta}_n^{*PT}| \leq z_{\beta/2}\} &= P_{\Delta^2}\{\tilde{n}(\delta - \tilde{\delta}_n^*)^2 \leq z_{\beta/2}^2, Z^2 \geq \chi_{1,\alpha}^2\} \\ &= P_{\Delta^2}\{\sqrt{\tilde{n}} |\delta - \tilde{\delta}_n^*| \leq z_{\beta/2}\} \\ &\quad - P_{\Delta^2}\{\sqrt{\tilde{n}} |\delta - \tilde{\delta}_n^*| \leq z_{\beta/2}, |Z| \leq z_{\alpha/2}\} \\ &= 1 - \beta - \Phi[\min\{z_{\beta/2}, z_{\alpha/2} - \Delta\}] \\ &\quad + \Phi[\max\{-z_{\beta/2}, -(z_{\alpha/2} + \Delta)\}]. \end{aligned} \tag{3.8}$$

Substituting the value of $\Delta = z_{\beta/2} + z_{\alpha/2}$, we obtain

$$P_{\Delta^2}\{\sqrt{\tilde{n}} |\delta - \hat{\delta}_n^{*PT}| \leq z_{\beta/2}\} = 1 - \beta.$$

(iv) If $\Delta^2 = 0$, and $\min(z_{\beta/2}, z_{\alpha/2}) = z_{\alpha/2}$, substituting the value of $\Delta = 0$ in (3.8)

we obtain $P_{\Delta^2}\{\sqrt{\tilde{n}} |\delta - \hat{\delta}_n^{*PT}| \leq z_{\beta/2}\} = 1 - \beta$.

(v) Further, if $\Delta^2 \rightarrow \infty$, from (3.8) we obtain

$$P_{\Delta^2}\{\sqrt{\tilde{n}} |\delta - \hat{\delta}_n^{*PT}| \leq z_{\beta/2}\} = 1 - \beta/2.$$

□

Next, we consider the coverage probability of confidence interval defined by $\hat{\delta}_n^{*s}$ as

$$C^s(\hat{\delta}_n^{*s}) = \{\delta : \sqrt{n} |\delta - \hat{\delta}_n^{*s}| \leq z_{\beta/2}\} \quad (3.9)$$

It may be shown that

$$\begin{aligned} P\{C^s(\hat{\delta}_n^{*s})\} &= P\left\{\sqrt{n} \left|\delta - \tilde{\delta}_n^* + \frac{d}{\sqrt{n}} \frac{\tilde{\delta}_n^*}{|\tilde{\delta}_n^*|}\right| \leq z_{\beta/2}\right\} \\ &= P\{(d - z_{\beta/2}) \leq (\sqrt{n} \tilde{\delta}_n^* - \Delta) \leq (d + z_{\beta/2}); \sqrt{n}\tilde{\delta}_n^* > 0\} \\ &\quad + P\{-(d + z_{\beta/2}) \leq (\sqrt{n} \tilde{\delta}_n^* - \Delta) \leq -(d - z_{\beta/2}); \sqrt{n}\tilde{\delta}_n^* < 0\} \\ &= 2\Phi(d + z_{\beta/2}) - 1 - \Phi\left(\max(d - z_{\beta/2}, -\Delta)\right) \\ &\quad + \Phi\left(\min(-(d - z_{\beta/2}), -\Delta)\right). \end{aligned}$$

Here, $d = \sqrt{\frac{2}{\pi}}$. For 95% and 90% C.I. of δ based on $\hat{\delta}_n^{*s}$, we obtain the coverage probability for variation of Δ as given in Table 3.2 . Thus, for 95% C.I. for δ based on $\hat{\delta}_n^{*s}$, we observe that the coverage probability as a function of Δ is decreasing up to a point say $\Delta = \Delta_{\min}$, then becomes constant 0.8716. Thus, eventual loss of efficiency of this C.I. is 8.25% and gain if Δ is near 0 is 4.65%. Similarly, for 90% C.I. of δ based on $\hat{\delta}_n^{*s}$, the eventual loss of efficiency is 12.56% and gain near the origin 9.5% . Graphs of coverage probabilities of 95% C.I. and 90% C.I. based on $\hat{\delta}_n^{*s}$ are shown in Figure 3.2 and Figure 3.3.

Next, we define the efficiency of the $100(1 - \beta)\%$ confidence interval as

$$EFF = \frac{\text{coverage probability obtained}}{1 - \beta}. \quad (3.10)$$

Tabular values of the efficiency of 95% and 90% confidence intervals based on PTE and SE for selected Δ values are presented in Table 3.3 and Table 3.4. The values of Δ are 0,1, $z_{\beta/2}$, $z_{\beta/2} + z_{\alpha/2}$ (for each $\alpha = 0.35(-0.05)0.1$), 4, 5, 6 and 7 .

The relative efficiency of the $100(1 - \beta)\%$ is more than one in the interval $[0, z_{\beta/2}]$,

Table 3.3: Efficiency of 95% C.I. of the sets $C^{PT}(\hat{\delta}_n^{*PT})$ and $C^s(\hat{\delta}_n^{*s})$

	α						
Δ	SE	0.1	0.15	0.20	0.25	0.30	0.35
0	1.046	1	1	1	1	1	1
1	1.046	1.022	1.018	1.014	1.009	1.004	1
1.96	0.9477	1.026	1.026	1.025	1.025	1.024	1.024
2.89	0.9197	0.9062	0.9429	0.9648	0.9793	0.9896	1
3	0.9188	0.9340	0.9638	0.9812	0.9924	1	1.006
3.11	0.9184	0.9496	0.9753	0.9900	0.9993	1.006	1.010
3.24	0.9180	0.9692	0.9893	1	1.007	1.012	1.015
3.40	0.9178	0.9846	1	1.008	1.013	1.017	1.019
3.60	0.9176	1.001	1.010	1.015	1.018	1.020	1.022
4	0.9174	1.020	1.020	1.023	1.024	1.024	1.025
5	0.9174	1.026	1.026	1.026	1.026	1.026	1.026
6	0.9174	1.026	1.026	1.026	1.026	1.026	1.026
7	0.9174	1.026	1.026	1.026	1.026	1.026	1.026

and due to the discontinuity of the indicator function , drops drastically to a minimum, then increases to one as Δ goes to $(z_{\beta/2} + z_{\alpha/2})$. Finally, the relative efficiency tends to $\frac{(1-\beta/2)}{1-\beta}$ as $\Delta \rightarrow \infty$. On the other hand, the relative efficiency of the confidence interval based on the shrinkage estimator is a continuous function of Δ and has value more than one in a shorter interval near the origin than the PTE, then decreases smoothly to a minimum guaranteed efficiency. The relative efficiency based on the SE near the origin is more than the efficiency of PTE based confidence interval . The discontinuity of the efficiency as a function of Δ for the PTE based confidence interval may have a damping effect on the practitioners and therefore it may be less attractive than the SE based confidence interval.

Table 3.4: Efficiency of 90% C.I. of the sets $C^{PT}(\hat{\delta}_n^{*PT})$ and $C^s(\hat{\delta}_n^{*s})$

	α						
Δ	SE	0.1	0.15	0.20	0.25	0.30	0.35
0	1.094	1	1	1	1	1	1
1	1.050	1.051	1.047	1.043	1.038	1.032	1.026
1.65	0.9293	1.055	1.054	1.053	1.052	1.051	1.049
2.58	0.8803	0.8597	0.9131	0.9467	0.9698	0.9866	0.9994
2.69	0.8782	0.8937	0.9403	0.9688	0.9882	1	1.012
2.80	0.8772	0.9178	0.9590	0.9839	1	1.012	1.021
2.93	0.8761	0.9489	0.9828	1.003	1.015	1.024	1.031
3.09	0.8755	0.9706	0.9989	1.015	1.025	1.032	1.037
3.30	0.8749	1	1.020	1.031	1.038	1.042	1.045
4	0.8744	1.046	1.049	1.052	1.053	1.054	1.054
5	0.8743	1.055	1.055	1.055	1.055	1.055	1.055
6	0.8743	1.055	1.055	1.055	1.055	1.055	1.055
7	0.8743	1.055	1.055	1.055	1.055	1.055	1.055

3.2.3 Case 2: Variance is Unknown

Basically, the unbiased estimation of δ is given by

$$\tilde{\delta}_n = c(\nu) \frac{\bar{Y}^E - \bar{Y}^C}{s_p} \quad (3.11)$$

$$\text{where } c(\nu) = \frac{\Gamma(\nu/2)}{\sqrt{(\nu/2)\Gamma(\frac{\nu-1}{2})}}, \nu = n^E + n^C - 2.$$

The unbiased estimator in (3.11) is due to Hedges (1981, 1982). Here, $\Gamma(\cdot)$ is the Gamma function, and the pooled within groups estimator of σ^2 is given by

$$s_p^2 = \nu^{-1} \left\{ \sum_{j=1}^{n^E} (Y_j^E - \bar{Y}^E)^2 + \sum_{j=1}^{n^C} (Y_j^C - \bar{Y}^C)^2 \right\}.$$

However, we formulate the hypothesis that δ may be equal to zero. That is to say, we suspect that the two means are equal, meaning that there is no effect of the experimental ingredients. To remove this vague hypothesis, we carry on a statistical

test of hypothesis:

$$H_0 : \mu^E = \mu^C \quad \text{Vs} \quad H_A : \mu^E \neq \mu^C \quad (3.12)$$

based on the test statistics

$$F_{1,\nu} = \frac{(\bar{Y}^E - \bar{Y}^C)^2}{s_p^2 \left(\frac{1}{n^E} + \frac{1}{n^C} \right)}. \quad (3.13)$$

The Proposed Estimators

Now, we consider the following estimators as in Saleh (2006)[Chapter 3, Sections 3.2, 3.5] .

$$\hat{\delta}_n^{PT} = \tilde{\delta}_n - \tilde{\delta}_n I(F_{1,\nu} \leq c_\alpha). \quad (3.14)$$

$$\hat{\delta}_n^s = \tilde{\delta}_n - d|(F_{1,\nu})^{1/2}|^{-1}\tilde{\delta}_n, \quad d > 0. \quad (3.15)$$

Bias and Mean Square Error Expressions

In this case, $\tilde{\delta}_n = c(\nu) \frac{\bar{Y}^E - \bar{Y}^C}{s_p}$, is the unrestricted estimator (UE) of δ , and $\frac{(\frac{1}{n^E} + \frac{1}{n^C})^{-\frac{1}{2}}}{c(\nu)} \tilde{\delta}_n$ follows a noncentral t-distribution with ν degrees of freedom D.F. , and noncentrality parameter $(\frac{1}{n^E} + \frac{1}{n^C})^{-\frac{1}{2}} \delta$. Thus, using the noncentral t-distribution, we have $E(\tilde{\delta}_n) = \delta$ and the bias and variance expressions are given by

$$b_1(\tilde{\delta}_n) = 0 \quad \text{and} \quad \text{var}(\tilde{\delta}_n) = \left(\frac{1}{n^E} + \frac{1}{n^C} \right) \left[\frac{\nu c^2(\nu)}{(\nu - 2)} (1 + \Delta^2) - \Delta^2 \right] = M_1(\tilde{\delta}_n).$$

For the test of $H_0 : \mu^E = \mu^C$, one uses the test-statistics $F_{1,\nu}$ given by (3.13) where $F_{1,\nu}$ follows a non central F-distribution with $(1, \nu)$ degrees of freedom D.F. and non centrality parameter $\frac{1}{2}\Delta^2$ where Δ^2 defined by $\Delta^2 = (\frac{1}{n^E} + \frac{1}{n^C})^{-1}\delta^2$ under the normal theory. Under H_0 , $F_{1,\nu}$ follows a central F-distribution. Thus, at the α -level of significance we reject H_0 if $F_{1,\nu} \geq c_\alpha$, where c_α is the upper α -level critical value of the central F-distribution with $(1, \nu)$ degrees of freedom.

Now, consider the PTE in (3.14). We may obtain the bias expression of $\hat{\delta}_n^{PT}$ following

Saleh (2006) (Chapter 4, Section 4.2) as

$$b_2(\hat{\delta}_n^{PT}) = -\delta G_{3,\nu-2}\left(\frac{\nu-2}{3\nu}c_\alpha, \Delta^2\right),$$

where $G_{3,\nu-2}(\cdot, \Delta^2)$, is the cdf of F-distribution with $(3, \nu-2)$ D.F. and noncentrality parameter $\frac{1}{2}\Delta^2$ and we may obtain the MSE expression for $\hat{\delta}_n^{PT}$ as

$$\begin{aligned} M(\hat{\delta}_n^{PT}) &= \left(\frac{1}{n^E} + \frac{1}{n^C}\right) \left\{ \frac{\nu c^2(\nu)}{\nu-2} \left[1 - G_{3,\nu-2}\left(\frac{\nu-2}{3\nu}c_\alpha, \Delta^2\right)\right] \right. \\ &\quad + \Delta^2 \left[\left(\frac{\nu c^2(\nu)}{\nu-2} - 1\right) + 2G_{3,\nu-2}\left(\frac{\nu-2}{3\nu}c_\alpha, \Delta^2\right) \right. \\ &\quad \left. \left. - \frac{\nu c^2(\nu)}{\nu-2} G_{5,\nu-2}\left(\frac{\nu-2}{5\nu}c_\alpha, \Delta^2\right) \right] \right\}. \end{aligned}$$

Next, we consider the shrinkage type estimator defined in (3.15), the bias expression may be shown to be

$$b_3(\hat{\delta}_n^s) = -d \, c(\nu) \, \sqrt{\left(\frac{1}{n^E} + \frac{1}{n^C}\right)} \, [1 - 2\Phi(-\Delta)].$$

We may calculate the MSE of $\hat{\delta}_n^s$ as follows

$$\begin{aligned} M_3(\hat{\delta}_n^s) &= \left(\frac{1}{n^E} + \frac{1}{n^C}\right) \left\{ \left[\frac{\nu c^2(\nu)}{(\nu-2)} (1 + \Delta^2) - \Delta^2 \right] \right. \\ &\quad \left. + d^2 c^2(\nu) - 2dc(\nu)[E|Z| - \Delta E\left(\frac{Z}{|Z|}\right)] \right\}. \end{aligned}$$

Minimizing w.r.t. d , we obtain the optimum value of d as $d = \frac{1}{c(\nu)} \sqrt{\frac{2}{\pi}} e^{-\frac{\Delta^2}{2}}$ which depends on Δ^2 . To make it free of Δ^2 , we choose $d^* = \frac{1}{c(\nu)} \sqrt{\frac{2}{\pi}}$. So that the MSE becomes

$$M_3(\hat{\delta}_n^s) = \left(\frac{1}{n^E} + \frac{1}{n^C}\right) \left\{ \left[\frac{\nu c^2(\nu)}{(\nu-2)} (1 + \Delta^2) - \Delta^2 \right] - \frac{2}{\pi} [2e^{-\frac{\Delta^2}{2}} - 1] \right\}.$$

We summarize the previous results in the following theorem.

Theorem 3.2.3. *If σ^2 is unknown, then the bias and MSE expressions of the three estimators are given by*

$$\begin{aligned}
b_1(\tilde{\delta}_n) &= 0, \\
M_1(\tilde{\delta}_n) &= \left(\frac{1}{n^E} + \frac{1}{n^C}\right) \left[\frac{\nu c^2(\nu)}{(\nu-2)} (1 + \Delta^2) - \Delta^2 \right], \\
b_2(\hat{\delta}_n^{PT}) &= -\delta G_{3,\nu-2} \left(\frac{\nu-2}{3\nu} c_\alpha, \Delta^2 \right), \\
M(\hat{\delta}_n^{PT}) &= \left(\frac{1}{n^E} + \frac{1}{n^C}\right) \left\{ \frac{\nu c^2(\nu)}{\nu-2} \left[1 - G_{3,\nu-2} \left(\frac{\nu-2}{3\nu} c_\alpha, \Delta^2 \right) \right] \right. \\
&\quad + \Delta^2 \left[\left(\frac{\nu c^2(\nu)}{\nu-2} - 1 \right) + 2G_{3,\nu-2} \left(\frac{\nu-2}{3\nu} c_\alpha, \Delta^2 \right) \right. \\
&\quad \left. \left. - \frac{\nu c^2(\nu)}{\nu-2} G_{5,\nu-2} \left(\frac{\nu-2}{5\nu} c_\alpha, \Delta^2 \right) \right] \right\}, \\
b_3(\hat{\delta}_n^s) &= -d c(\nu) \sqrt{\left(\frac{1}{n^E} + \frac{1}{n^C}\right)} [1 - 2\Phi(-\Delta)], \\
M_3(\hat{\delta}_n^s) &= \left(\frac{1}{n^E} + \frac{1}{n^C}\right) \left\{ \left[\frac{\nu c^2(\nu)}{(\nu-2)} (1 + \Delta^2) - \Delta^2 \right] - \frac{2}{\pi} [2e^{-\frac{\Delta^2}{2}} - 1] \right\}.
\end{aligned}$$

MSE-based Relative Efficiency

Here, the $\text{MRE}(\hat{\delta}_n^{PT}, \tilde{\delta}_n)$ is given by

$$\text{MRE}(\hat{\delta}_n^{PT}, \tilde{\delta}_n) = [1 + g_\alpha(\Delta^2)]^{-1}$$

where $g_\alpha(\Delta^2)$

$$\begin{aligned}
&= -\left(\frac{1}{n^E} + \frac{1}{n^C}\right) [\text{var}(\tilde{\delta}_n)]^{-1} \left(\frac{\nu c^2(\nu)}{\nu-2} G_{3,\nu-2} \left(\frac{\nu-2}{3\nu} c_\alpha, \Delta^2 \right) - \Delta^2 \right. \\
&\quad \left. \times \left[2G_{3,\nu-2} \left(\frac{\nu-2}{3\nu} c_\alpha, \Delta^2 \right) - \frac{\nu c^2(\nu)}{\nu-2} G_{5,\nu-2} \left(\frac{\nu-2}{5\nu} c_\alpha, \Delta^2 \right) \right] \right).
\end{aligned}$$

$\text{MRE}(\hat{\delta}_n^{PT}, \tilde{\delta}_n)$ may be written as a function of (α, Δ^2) . Figure 3.4 illustrates the graph of $\text{MRE}(\alpha, \Delta^2)$ as a function of Δ^2 . For fixed α , it decreases, crossing the 1-line to a minimum at $\Delta^2 = \Delta_{min}^2(\alpha)$, then increases towards the 1-line as $\Delta^2 \rightarrow \infty$. The

Table 3.5: Maximum and minimum relative efficiency of PTE for $\alpha = 0.05(0.10)0.45$ for selected sample sizes

$\alpha \backslash (n^E, n^C)$		(3, 4)	(5, 6)	(9, 10)	(13, 14)
0.05	E_{max}	2.4266	2.9413	3.2459	3.3546
	E_0	0.6286	0.5053	0.4522	0.4360
	Δ_0^2	4.7	4.95	4.7	4.7
0.15	E_{max}	1.4961	1.6343	1.7113	1.7381
	E_0	0.7843	0.6895	0.6417	0.6259
	Δ_0^2	3.4	3.5	3.55	3.55
0.25	E_{max}	1.2480	1.3110	1.3455	1.3575
	E_0	0.8674	0.8004	0.7633	0.7506
	Δ_0^2	2.95	3.05	3.1	3.15
0.35	E_{max}	1.1334	1.1660	1.1836	1.1897
	E_0	0.9201	0.8761	0.8502	0.8411
	Δ_0^2	2.7	2.75	2.85	2.9
0.45	E_{max}	1.0713	1.0884	1.0976	1.1008
	E_0	0.9544	0.9278	0.9115	0.9056
	Δ_0^2	2.55	2.65	2.7	2.75

maximum of $MRE(\alpha, \Delta^2)$ occurs at $\Delta^2 = 0$ with the value

$$MRE(\alpha, 0) = [1 - (\frac{1}{n^E} + \frac{1}{n^C})[\text{var}(\tilde{\delta}_n)]^{-1} \frac{\nu c^2(\nu)}{\nu - 2} G_{3, \nu-2}(\frac{\nu - 2}{3\nu} c_\alpha, 0)]^{-1}$$

for all $\alpha \in A$, the set of all possible values of α . The value of the max of $MRE(\alpha, \Delta^2)$ decreases as α increases. Intersection of the graph of $MRE(\alpha, \Delta^2)$ with the 1-line occurs at Δ_1^2 . Thus if $0 \leq \Delta^2 \leq \Delta_1^2$, then $\hat{\delta}_n^{PT}$ performs better than $\tilde{\delta}_n$, and if $\Delta^2 \geq \Delta_1^2$, then $\tilde{\delta}_n$ performs better. The cut-off point Δ_1^2 may be determined by

$$\Delta_1^2 = \frac{\frac{\nu c^2(\nu)}{\nu-2} [G_{3, \nu-2}(\frac{\nu-2}{3\nu} c_\alpha, \Delta^2)]}{2G_{3, \nu-2}(\frac{\nu-2}{3\nu} c_\alpha, \Delta^2) - \frac{\nu c^2(\nu)}{\nu-2} G_{5, \nu-2}(\frac{\nu-2}{5\nu} c_\alpha, \Delta^2)}$$

In order to obtain a PTE with minimum guaranteed RE, say E_0 , we follow the same procedure as in the known variance case. Table 3.5 may be used to determine the optimum level of significance for the PTE.

Figure 3.4: Graph of MSE-based relative efficiency of SE and PTE for $n^E = 9, n^C = 10, \alpha = 0.05$ and 0.15

Rows of Table 3.5 contain the maximum, minimum values of MRE of PTE and the value of Δ^2 at which the minimum holds. To explain how to use the table, for example, let $(n^E, n^C) = (3, 4)$ and assume that we wish to obtain a PTE with at least 86% efficiency, Table 3.5 gives $\alpha^* = 0.25$ at the intersection of $E_0=0.8674$ and $(n^E, n^C) = (3, 4)$. Hence the optimum level for the PTE in this case is $\alpha=0.25$ with a maximum possible efficiency 1.2480.

Now, we consider the MSE-based relative efficiency of $\hat{\delta}_n^s$ versus $\tilde{\delta}_n$ as

$$\text{MRE}(\hat{\delta}_n^s, \tilde{\delta}_n) = \left\{ 1 - \frac{2}{\pi} \left[\frac{\nu c^2(\nu)}{(\nu - 2)} (1 + \Delta^2) - \Delta^2 \right]^{-1} [2e^{-\frac{\Delta^2}{2}} - 1] \right\}^{-1}.$$

$$\text{If } \Delta^2 = 0, \text{MRE}(\hat{\delta}_n^s, \tilde{\delta}_n) = \frac{\nu \pi c^2(\nu)}{\nu \pi c^2(\nu) - 2(\nu - 2)}.$$

$$\text{and if } \Delta^2 \rightarrow \infty, \text{MRE}(\hat{\delta}_n^s, \tilde{\delta}_n) = \frac{\nu \pi c^2(\nu)}{\nu \pi c^2(\nu) + 2(\nu - 2)}.$$

Further, as $\nu \rightarrow \infty$ and $\Delta^2 = 0$, $\text{MRE}(\hat{\delta}_n^s, \tilde{\delta}_n) \rightarrow \frac{\pi}{\pi-2}$. Moreover, as $\nu \rightarrow \infty$ and $\Delta^2 \rightarrow \infty$, $\text{RE}(\hat{\delta}_n^s, \tilde{\delta}_n) \rightarrow \frac{\pi}{\pi+2}$.

Graphs of MSE-based relative efficiency of preliminary test estimators for $\alpha=0.05$ and 0.15 , and that of shrinkage type estimators are shown in Figure 3.4. Now, note that the $\text{MRE}(\hat{\delta}_n^s, \tilde{\delta}_n)$ as a function of Δ^2 crosses the 1-line at $\Delta^2 = \ln 4$. Thus, for $0 \leq \Delta^2 \leq \ln 4$, $\hat{\delta}_n^s$ performs better than $\tilde{\delta}_n$, otherwise $\tilde{\delta}_n$ performs better outside this interval. The first two rows of Table 3.6 contain the maximum and minimum relative efficiency of SE for selected sample sizes. The remaining rows of Table 3.6 contain the minimum relative efficiency, E_0 , of the PTE at $\Delta^2 = \Delta_0^2$, which have been recorded for each $\alpha=0.05(0.10)0.45$ with the corresponding efficiency ($E_{\Delta_0^2}$) of SE for $\Delta^2 = \Delta_0^2$. Figure 3.4 and Table 3.6, explain that when Δ^2 is near the origin, SE outperforms PTE for most α values. Also, at $\Delta^2 = 0$, the relative efficiency of SE is higher than that of PTE. Further, SE is independent of α . Thus, considering the overall performance of SE relative to PTE, SE is preferable to PTE since it produces interpolated estimators which are free from the level of significance, α .

3.2.4 Confidence Intervals

For the finite sample $100(1-\beta)\%$ confidence interval for δ one needs the computational procedures using the non central t-distribution or noncentral F-distribution, see for example, Cumming and Finch (2001) and Lecoutre (2007) among others. At the end,

Table 3.6: Maximum and Minimum relative efficiency of SE and efficiency of PTE for selected sample sizes

$\alpha \backslash (n^E, n^C)$		(3,4)	(5,6)	(9,10)	(13,14)
	E^{max}	2.1760	2.4560	2.6030	2.6520
	E^{min}	0.7952	0.7255	0.6798	0.6613
0.05	$E_{\Delta_0^2}$	0.7966	0.7338	0.6993	0.6870
	E_0	0.6286	0.5053	0.4522	0.4360
	Δ_0^2	4.7	4.95	4.7	4.7
0.15	$E_{\Delta_0^2}$	0.8168	0.7642	0.7350	0.7257
	E_0	0.7843	0.6895	0.6417	0.6259
	Δ_0^2	3.4	3.5	3.55	3.55
0.25	$E_{\Delta_0^2}$	0.8337	0.7860	0.7630	0.7484
	E_0	0.8674	0.8004	0.7633	0.7506
	Δ_0^2	2.95	3.05	3.1	3.15
0.35	$E_{\Delta_0^2}$	0.8469	0.8059	0.7777	0.7665
	E_0	0.9201	0.8761	0.8502	0.8411
	Δ_0^2	2.7	2.75	2.85	2.9
0.45	$E_{\Delta_0^2}$	0.8565	0.8137	0.7903	0.7791
	E_0	0.9544	0.9278	0.9115	0.9056
	Δ_0^2	2.55	2.65	2.7	2.75

they all proposed methods of approximations as it became difficult to specify simple methods. With this backdrop, here we consider the asymptotic distribution of $\tilde{\delta}_n$ as $N(\delta, \sigma^2(\delta))$, $\sigma^2(\delta) = \frac{1}{\tilde{n}} + \frac{\delta^2}{2(n-2)}$ and define the Neyman-Pearson solution as

$$C^0(\tilde{\delta}_n) = \{\delta : |\delta - \tilde{\delta}_n| \leq z_{\beta/2} \sqrt{\frac{1}{\tilde{n}} + \frac{\tilde{\delta}_n^2}{2(n-2)}}\}.$$

However, we notice that as $n \rightarrow \infty$

$$\tilde{\delta}_n = c(\nu) \frac{\bar{Y}^E - \bar{Y}^C}{\sigma} \frac{\sigma}{s_p} \xrightarrow{P} \frac{\bar{Y}^E - \bar{Y}^C}{\sigma} \quad \text{and as } \frac{s_p}{\sigma} \xrightarrow{a.s.} 1,$$

where $\xrightarrow{P}$ stands for “converges in probability” and $\xrightarrow{a.s.}$ stands for “converges almost surely”.

Hence,

$$\lim_{n \rightarrow \infty} P\{C^0(\tilde{\delta}_n)\} = P\{\sqrt{\tilde{n}} |\delta - \tilde{\delta}_n^*| \leq z_{\beta/2}\} = 1 - \beta.$$

Now, we consider 100(1- β)% confidence interval for δ using PTE, $\hat{\delta}_n^{PT}$ as the Neyman-Pearson solution

$$C^{PT}(\hat{\delta}_n^{PT}) = \{\delta : |\delta - \hat{\delta}_n^{PT}| \leq z_{\beta/2} \sqrt{\frac{1}{\tilde{n}} + \frac{\tilde{\delta}_n^2}{2(n-2)}}\}.$$

Let us now consider the limiting coverage probability of $C^{PT}(\hat{\delta}_n^{PT})$. First, note that if $\Delta^2 \leq z_{\beta/2}^2(1 + \frac{\tilde{n}\tilde{\delta}_n^2}{2(n-2)})$, we have

$$\begin{aligned} \lim_{n \rightarrow \infty} P\{C^{PT}(\hat{\delta}_n^{PT})\} &= \lim_{n \rightarrow \infty} P_{\Delta^2}\{\tilde{n}(\delta - \hat{\delta}_n^{PT})^2 \leq z_{\beta/2}^2(1 + \frac{\tilde{n}\tilde{\delta}_n^2}{2(n-2)})\} \\ &= H_1(\chi_{1,\alpha}^2; \Delta^2)I(\Delta^2 < z_{\beta/2}^2) \\ &\quad + P_{\Delta^2}\{-z_{\beta/2} \leq Z - \Delta \leq z_{\beta/2}, |Z| > z_{\alpha/2}\}. \end{aligned}$$

The R.H.S may be written as

$$\begin{aligned} &= H_1(\chi_{1,\alpha}^2; \Delta^2)I(\Delta^2 < z_{\beta/2}^2) + (1 - \beta) \\ &\quad - \Phi[\min\{z_{\beta/2}, z_{\alpha/2} - \Delta\}] \\ &\quad + \Phi[\max\{-z_{\beta/2}, -(z_{\alpha/2} + \Delta)\}]. \end{aligned}$$

For the 100(1- β)% confidence interval for δ based on SE, $\hat{\delta}_n^s$, we define the Neyman-Pearson solution as

$$\begin{aligned} C^s(\hat{\delta}_n^s) &= \{\delta : |\delta - \hat{\delta}_n^s| \leq z_{\beta/2} \sqrt{\frac{1}{\tilde{n}} + \frac{\tilde{\delta}_n^2}{2(n-2)}}\}. \\ &= \{\Delta : \sqrt{\tilde{n}}|\delta - \hat{\delta}_n^s| \leq z_{\beta/2} \sqrt{1 + \frac{\tilde{n}\tilde{\delta}_n^2}{2(n-2)}}\}. \end{aligned}$$

Then, the limiting coverage probability of $C^s(\hat{\delta}_n^s)$ is given by

$$\begin{aligned} \lim_{n \rightarrow \infty} P\{C^s(\hat{\delta}_n^s)\} &= \lim_{n \rightarrow \infty} P_{\Delta^2}\{\tilde{n}(\delta - \hat{\delta}_n^s)^2 \leq z_{\beta/2}^2(1 + \frac{\tilde{n}\tilde{\delta}_n^2}{2(n-2)})\} \\ &= 2\Phi(d + z_{\beta/2}) - 1 - \Phi\left(\max(d - z_{\beta/2}, -\Delta)\right) \\ &\quad + \Phi\left(\min(-(d - z_{\beta/2}), -\Delta)\right). \end{aligned}$$

The limiting coverage probabilities are the same as the known variance case for which the properties specified in section 3.2.2.

Example 3.2.4. Suppose that a study with $n^E = n^C = 15$ yields a sample effect size value of $\tilde{\delta}_n = 0.580$, $\hat{\delta}_n^{PT} = 0$, and $\hat{\delta}_n^s = 0.3284$ then the 90% confidence intervals for δ based on PTE and SE respectively are given by

$$\begin{aligned} \hat{\delta}_n^{PT} \pm z_{\beta/2} \sqrt{\frac{1}{\tilde{n}} + \frac{\tilde{\delta}_n^2}{2(n-2)}} &= \pm 1.65 \sqrt{\frac{1}{7.5} + \frac{(0.580)^2}{56}} = (-0.616, 0.616). \\ \hat{\delta}_n^s \pm z_{\beta/2} \sqrt{\frac{1}{\tilde{n}} + \frac{\tilde{\delta}_n^2}{2(n-2)}} &= 0.3284 \pm 1.65 \sqrt{\frac{1}{7.5} + \frac{(0.580)^2}{56}} = (-0.2875, 0.9443). \end{aligned}$$

3.3 Empirical Bayes Estimators and Comparisons

In this section, we discuss the empirical Bayes methods of effect size estimation due to Efron (2009, 2010). Let the effect size, δ , be distributed with density $g(\delta)$ and the transformed variable $z|\delta \sim N(\delta, 1)$ where $z = \Phi^{-1}(F_\nu(t))$, and $\Phi(\cdot)$ is the cdf of standard normal distribution, $F_\nu(\cdot)$ is the cdf of central t-distribution with ν D.F. $t = \frac{\bar{Y}^E - \bar{Y}^C}{s_p}$. Then, we have the following result:

Theorem 3.3.1. (Efron's Theorem) *Let $f(z)$ be the marginal density of z for the model described above, given by*

$$f(z) = \int_{-\infty}^{\infty} \phi(z - \delta) g(\delta) d\delta \quad \text{where} \quad \phi(z) = \frac{1}{\sqrt{2\pi}} \exp\left\{-\frac{z^2}{2}\right\}.$$

Then, the posterior density of δ given z is

$$g(\delta|z) = e^{z\delta - \psi(z)} [e^{-\delta^2/2} g(\delta)], \quad \text{with} \quad \psi(z) = \log(f(z)/\phi(z)).$$

As a consequence

$$E(\delta|z) = \psi'(z) \quad \text{and} \quad \text{var}(\delta|z) = \psi''(z),$$

where $\psi'(z)$ and $\psi''(z)$ are the first and second derivatives of $\psi(z)$. Let $\ell(z) = \log(f(z))$, then

$$E(\delta|z) = z + \ell'(z) \quad \text{and} \quad \text{var}(\delta|z) = 1 + \ell''(z).$$

Efron uses the R algorithm Ebay to obtain a smooth estimate $\hat{\ell}(z)$ of $\ell(z)$. We consider Table 11.2 of Efron (2010) page 220 which concerns a micro array example, the prostate data. Genetic expression level for 55 genes having the largest values of $|\hat{\delta}_i|$ were obtained for $n = 102$, $n^E = 52$ prostate cancer patients and $n^C = 50$ normal control subjects. The main goal of the study was to discover a small number of “interesting” genes, that is, genes whose expression levels differ between the prostate and normal subject. Table 3.7 gives the computational results corresponding to Efron’s Table 11.2.

Column 1 is the gene number, column 2 is the observed effect size value after transforming to Z value as explained above and column 3 is the empirical Bayes estimate . Column 4 is our unrestricted estimator it is equal to the preliminary test estimator in this case and both are based on converted z-values to adjusted t-values. Column 5 is our shrinkage estimator and is computed based on (3.15). Column 6 is a converted Efron’s estimator $\hat{\delta}$ to adjusted t-values for comparison. It is clear from the table that our estimators are always higher in magnitude than empirical

Table 3.7: Computation of UE, PTE and SE using prostate cancer data

step	z-value	$\hat{\delta}$	$\hat{\delta}_n = \hat{\delta}_n^{PT}$	$\hat{\delta}_n^s$	$\hat{\delta}(z)$	$\widehat{var}(\hat{\delta}_n^s)$	$\widehat{var}(\hat{\delta})$
1	5.29	4.11	5.6557	4.8578	4.119	0.2030	0.87
2	4.83	3.65	5.1013	4.3034	3.625	0.1725	0.89
3	-4.42	-3.57	-4.6225	-3.8246	-3.5399	0.1487	0.92
4	-4.33	-3.52	-4.5191	-3.7212	-3.4872	0.1439	0.92
5	-4.29	-3.47	-4.4734	-3.6755	-3.435	0.1418	0.93
6	-4.14	-3.30	-4.3029	-3.5050	-3.257	0.1341	0.97
7	4.47	3.24	4.6802	3.8823	3.1943	0.1515	0.91
8	4.40	3.16	4.5995	3.8016	3.1114	0.1476	0.92
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$
45	-3.38	-2.23	-3.4618	-2.6639	-2.1679	0.1007	1.18
46	3.57	2.22	3.6687	2.8709	2.1579	0.1083	0.97
47	3.56	2.20	3.6578	2.8599	2.1380	0.1079	0.97
48	3.56	2.20	3.6578	2.8599	2.1380	0.1079	0.97
49	-3.33	-2.15	-3.4076	-2.6097	-2.0883	0.0988	1.19
50	3.51	2.15	3.6032	2.8053	2.0883	0.1058	0.97
51	-3.33	-2.14	-3.4076	-2.6097	-2.0783	0.0988	1.19
52	-3.32	-2.12	-3.3968	-2.5989	-2.0585	0.0984	1.19
53	-3.32	-2.12	-3.3968	-2.5989	-2.0585	0.0984	1.19
54	3.47	2.09	3.5595	2.7617	2.0287	0.1042	0.97
55	3.46	2.09	3.5487	2.7508	2.0287	0.1038	0.97

Bayes estimators. The last two columns are the estimated variance of our shrinkage estimator and the variance of the empirical Bayes estimators as obtained by Efron, respectively. Our results belong to the case of estimating the noncentrality parameter, δ of a noncentral t-distribution with ν D.F.

It is evident that our estimators are comparable to the empirical Bayes estimators and are easy to compute because of our explicit formula (3.15).

3.4 Multiple-study Problem

In meta-analysis of research, effect size estimation is a primary concern of the analysts. Accordingly, Glass (1976), Cohen (1977) and Hedges (1981, 1982) among others, provided various estimators of effect size. In this section we discuss the extension

of the estimation of effect sizes from a series of p independent experiments due to Hedges (1982). In the process, we first postulate the hypothesis of homogeneity of effect sizes, and formulate a test-statistics for the test of the postulated hypothesis and subsequently define three new estimators of effect size besides the usual estimator and the pooled estimator. To discuss our problem, we have the following assumptions and model.

Suppose we have the data from a series of p independent studies, where each study compares the experimental group (E) with the control group (C). Let Y_{ij}^E and Y_{ij}^C be the j th scores on the i th experiment from the experimental and control groups, respectively. Assume that for fixed i , Y_{ij}^E and Y_{ij}^C are normally distributed with mean μ_i^E and μ_i^C and common variance, σ_i^2 . That is

$$Y_{ij}^E \sim N(\mu_i^E, \sigma_i^2), \quad j = 1, \dots, n_i^E, \quad i = 1, \dots, p.$$

$$Y_{ij}^C \sim N(\mu_i^C, \sigma_i^2), \quad j = 1, \dots, n_i^C, \quad i = 1, \dots, p.$$

The population effect size is defined by

$$\delta_i = \frac{\mu_i^E - \mu_i^C}{\sigma_i}, \quad i = 1, \dots, p.$$

It is well-known in the literature that the unbiased estimator of δ_i is given by Hedges (1982) as

$$\tilde{\delta}_i = c(\nu_i) \frac{\bar{Y}_i^E - \bar{Y}_i^C}{s_{i(p)}},$$

$$\text{where } c(\nu_i) = \frac{\Gamma(\nu_i/2)}{\sqrt{(\nu_i/2)} \Gamma(\frac{\nu_i-1}{2})}.$$

$$s_{i(p)}^2 = \nu_i^{-1} \left\{ \sum_{j=1}^{n_i^E} (Y_{ij}^E - \bar{Y}_i^E)^2 + \sum_{j=1}^{n_i^C} (Y_{ij}^C - \bar{Y}_i^C)^2 \right\},$$

and $\nu_i = n_i - 2$, $n_i = n_i^E + n_i^C$. The variance of $\tilde{\delta}_i$ is

$$v_i = \frac{\nu_i c^2(\nu_i)}{(\nu_i - 2)} \left[\left(\frac{1}{n_i^E} + \frac{1}{n_i^C} \right) + \delta_i^2 \left(1 - \frac{\nu_i - 2}{\nu_i c^2(\nu_i)} \right) \right], \quad i = 1, \dots, p.$$

which can be estimated by replacing δ_i by its unbiased and consistent estimator, namely, $\tilde{\delta}_i$ to obtain

$$\hat{v}_i = \frac{\nu_i c^2(\nu_i)}{(\nu_i - 2)} \left[\left(\frac{1}{n_i^E} + \frac{1}{n_i^C} \right) + \tilde{\delta}_i^2 \left(1 - \frac{\nu_i - 2}{\nu_i c^2(\nu_i)} \right) \right], \quad i = 1, \dots, p. \quad (3.16)$$

Let $\boldsymbol{\delta} = (\delta_1, \dots, \delta_p)'$ be the vector of the population effect sizes of p ($p \geq 3$) independent experiments and correspondingly, let $\tilde{\boldsymbol{\delta}}_n = (\tilde{\delta}_1, \dots, \tilde{\delta}_p)'$ be the vector of unbiased estimators of $\boldsymbol{\delta}$. We call $\tilde{\boldsymbol{\delta}}_n$ the unrestricted estimator of $\boldsymbol{\delta}$. Our objective is to provide methods for quantitative synthesis of p series of independent studies. First step in meta analysis is to formulate a plausible hypothesis to summarize the results of p experiments in the form that

$$H_0 : \delta_1 = \dots = \delta_p = \delta_0 \text{ unknown} \quad (3.17)$$

may hold. If H_0 holds then by generalized least square method we minimize

$$(\tilde{\boldsymbol{\delta}}_n - \delta_0 \mathbf{1}_p)' \hat{\boldsymbol{\Lambda}}_n^{-1} (\tilde{\boldsymbol{\delta}}_n - \delta_0 \mathbf{1}_p)$$

with respect to δ_0 to obtain

$$\hat{\delta}_{0n}^* = \frac{\mathbf{1}_p' \hat{\boldsymbol{\Lambda}}_n^{-1} \tilde{\boldsymbol{\delta}}_n}{\mathbf{1}_p' \hat{\boldsymbol{\Lambda}}_n^{-1} \mathbf{1}_p}, \quad \hat{\boldsymbol{\Lambda}}_n = \text{Diag}(\hat{v}_1, \dots, \hat{v}_p)$$

where $\mathbf{1}_p = (1, \dots, 1)'$ is a p -tuple of one's and $\hat{v}_i$ is the estimate of the variance, v_i as given in (3.16). We call $\hat{\delta}_{0n}^*$ the restricted estimator. If H_0 holds, then plausible estimate $\hat{\delta}_{0n}^*$ summarizes the findings of the synthesis of the p experiments.

In case H_0 is not tenable, then we calculate the departure statistics from the null hypothesis H_0 , as

$$\mathcal{D}_n = (\tilde{\boldsymbol{\delta}}_n - \hat{\boldsymbol{\delta}}_{0n}^* \mathbf{1}_p)' \hat{\boldsymbol{\Lambda}}_n^{-1} (\tilde{\boldsymbol{\delta}}_n - \hat{\boldsymbol{\delta}}_{0n}^* \mathbf{1}_p) \quad (3.18)$$

The “amount” of departure of $\tilde{\boldsymbol{\delta}}_n$ from $\hat{\boldsymbol{\delta}}_{0n}^* \mathbf{1}_p$ determines the acceptance of H_0 or not. The bigger the value of $\mathcal{D}_n$, the greater is the chance of rejecting the hypothesis H_0 and smaller values of $\mathcal{D}_n$ support the hypothesis.

Which estimator of $\boldsymbol{\delta} = (\delta_1, \dots, \delta_p)'$ is preferred? Is it $\tilde{\boldsymbol{\delta}}_n$ or $\hat{\boldsymbol{\delta}}_{0n}^* \mathbf{1}_p$? A combination of $\tilde{\boldsymbol{\delta}}_n$ and $\hat{\boldsymbol{\delta}}_{0n}^* \mathbf{1}_p$ may provide a better estimate of $\boldsymbol{\delta}$ against the extremes as a compromise between the two estimates. Hence, we may write the convex combination as

$$\begin{aligned} \hat{\boldsymbol{\delta}}_n^{(\tau)} &= \tau \hat{\boldsymbol{\delta}}_{0n}^* \mathbf{1}_p + (1 - \tau) \tilde{\boldsymbol{\delta}}_n \\ &= \tilde{\boldsymbol{\delta}}_n - \tau (\tilde{\boldsymbol{\delta}}_n - \hat{\boldsymbol{\delta}}_{0n}^* \mathbf{1}_p). \end{aligned}$$

If $\tau = 0$, we obtain $\tilde{\boldsymbol{\delta}}_n$ and if $\tau = 1$, we obtain $\hat{\boldsymbol{\delta}}_{0n}^*$. A simple form of τ may be taken as

$$\tau = I(\mathcal{D}_n \leq d_n) = \begin{cases} 0 & \text{if } \mathcal{D}_n > d_n, \\ 1 & \text{if } \mathcal{D}_n \leq d_n, \end{cases}$$

where d_n is a specific chosen number indicating the “amount of tolerable departure”.

Then, we obtain the estimator as

$$\hat{\boldsymbol{\delta}}_n^{PT} = \tilde{\boldsymbol{\delta}}_n - (\tilde{\boldsymbol{\delta}}_n - \hat{\boldsymbol{\delta}}_{0n}^* \mathbf{1}_p) I(\mathcal{D}_n \leq d_n).$$

We designate this estimator as the preliminary test estimator (PTE), where testing is based on the departure statistics, $\mathcal{D}_n$, if we follow this rule. Then we have extreme choices of the estimator. This allows us to think about some continuous function of

$\mathcal{D}_n$, say $g(\mathcal{D}_n)$ and write

$$\hat{\boldsymbol{\delta}}_n^g = \tilde{\boldsymbol{\delta}}_n - g(\mathcal{D}_n)(\tilde{\boldsymbol{\delta}}_n - \hat{\boldsymbol{\delta}}_{0n}^* \mathbf{1}_p).$$

For example, we may choose $g(\mathcal{D}_n) = c \mathcal{D}_n^{-1}$, $c > 0$, or $g(\mathcal{D}_n) = 1 - (1 - c \mathcal{D}_n^{-1})I(\mathcal{D}_n > c)$, $c > 0$. If $g(\mathcal{D}_n) = c \mathcal{D}_n^{-1}$, we rewrite

$$\hat{\boldsymbol{\delta}}_n^s = c \mathcal{D}_n^{-1} \hat{\boldsymbol{\delta}}_{0n}^* \mathbf{1}_p + (1 - c \mathcal{D}_n^{-1}) \tilde{\boldsymbol{\delta}}_n, c > 0.$$

However, if $\mathcal{D}_n < c$, $\hat{\boldsymbol{\delta}}_n^s$ will change signs of the estimators. To avoid this feature of the estimator $\hat{\boldsymbol{\delta}}_n^s$, we define the positive rule estimator as a preliminary test estimator given by

$$\hat{\boldsymbol{\delta}}_n^{s+} = \hat{\boldsymbol{\delta}}_{0n}^* \mathbf{1}_p I(\mathcal{D}_n \leq c) + \hat{\boldsymbol{\delta}}_n^s I(\mathcal{D}_n \geq c),$$

which is a convex combination of $\hat{\boldsymbol{\delta}}_n^s$ and $\hat{\boldsymbol{\delta}}_{0n}^* \mathbf{1}_p$.

The finite sample properties of the above three estimators are difficult to study analytically and would need huge set of simulation involving noncentral t-distribution. If sample sizes are large following Hedges (1982) and Hedges & Olkin (1985), we may obtain the following theorems:

Theorem 3.4.1. *The asymptotic distribution of $\tilde{\boldsymbol{\delta}}_n$ is given by*

$$\sqrt{n}(\tilde{\boldsymbol{\delta}}_n - \boldsymbol{\delta}) \sim N(0, \boldsymbol{\Sigma}_0)$$

$\boldsymbol{\Sigma}_0 = \lim_{n \rightarrow \infty} n \boldsymbol{\Sigma}_n$ where $\frac{n_i^E}{n_i} \rightarrow \pi_i^E$, $\frac{n_i^C}{n_i} \rightarrow \pi_i^C$, $\frac{n_i}{n} \rightarrow \pi_i$, π_i^E , π_i^C and π_i remain fixed as $\min_{1 \leq i \leq p} n_i \rightarrow \infty$ which implies $n \rightarrow \infty$,

$$\boldsymbol{\Sigma}_n = \text{Diag} \left(\left(\frac{1}{n_1^E} + \frac{1}{n_1^C} \right) + \frac{\delta_1^2}{2(n_1^E + n_1^C - 2)}, \dots, \left(\frac{1}{n_p^E} + \frac{1}{n_p^C} \right) + \frac{\delta_p^2}{2(n_p^E + n_p^C - 2)} \right). \quad (3.19)$$

That is

$$\Sigma_0 = \text{Diag} \left(\frac{1}{\pi_1} \left[\left(\frac{1}{\pi_1^E} + \frac{1}{\pi_1^C} \right) + \frac{\delta_1^2}{2(\pi_1^E + \pi_1^C)} \right], \dots, \frac{1}{\pi_p} \left[\left(\frac{1}{\pi_p^E} + \frac{1}{\pi_p^C} \right) + \frac{\delta_p^2}{2(\pi_p^E + \pi_p^C)} \right] \right). \quad (3.20)$$

The estimated variances are then obtained by replacing δ_i in (3.19) by the unbiased and consistent estimators of δ_i , namely, $\tilde{\delta}_i$, obtaining the consistent estimator of Σ_n as

$$\hat{\Sigma}_n = \text{Diag} \left(\left(\frac{1}{n_1^E} + \frac{1}{n_1^C} \right) + \frac{\tilde{\delta}_1^2}{2(n_1^E + n_1^C - 2)}, \dots, \left(\frac{1}{n_p^E} + \frac{1}{n_p^C} \right) + \frac{\tilde{\delta}_p^2}{2(n_p^E + n_p^C - 2)} \right). \quad (3.21)$$

Accordingly, the asymptotic $100(1 - \beta)\%$ confidence interval for each of δ_i is given by

$$\tilde{\delta}_i \pm z_{\beta/2} \sqrt{\left(\frac{1}{n_i^E} + \frac{1}{n_i^C} \right) + \frac{\tilde{\delta}_i^2}{2(n_i^E + n_i^C - 2)}}.$$

If H_0 in (3.17) is true, then the restricted estimator (RE) of δ , is given by

$$\hat{\delta}_{0n} = \frac{\mathbf{1}_p' \hat{\Sigma}_n^{-1} \tilde{\delta}_n}{\mathbf{1}_p' \hat{\Sigma}_n^{-1} \mathbf{1}_p}, \quad (3.22)$$

with the estimated (asymptotic) variance as

$$\hat{\sigma}^2(\hat{\delta}_{0n}) = (\mathbf{1}_p' \hat{\Sigma}_n^{-1} \mathbf{1}_p)^{-1}$$

Accordingly, the asymptotic $100(1 - \beta)\%$ confidence interval for δ_0 is given by

$$\hat{\delta}_{0n} \pm z_{\beta/2} \hat{\sigma}(\hat{\delta}_{0n}). \quad (3.23)$$

In general, H_0 in (3.17) is vague. We remove this vagueness based on a statistical test defined by

$$\mathcal{L}_n = n(\tilde{\delta}_n - \hat{\delta}_{0n} \mathbf{1}_p)' (n \hat{\Sigma}_n)^{-1} (\tilde{\delta}_n - \hat{\delta}_{0n} \mathbf{1}_p). \quad (3.24)$$

This test is known as Cochran's test and it is often used as the standard test for testing homogeneity in meta-analysis, see Hartung, Knapp and Sinha (2008).

Theorem 3.4.2. *If H_0 is true, and $\frac{n_i^E}{n_i} \rightarrow \pi_i^E$, $\frac{n_i^C}{n_i} \rightarrow \pi_i^C$, $\frac{n_i}{n} \rightarrow \pi_i$, π_i^E , π_i^C , and π_i remain fixed as $\min_{1 \leq i \leq p} n_i \rightarrow \infty$ which implies $n \rightarrow \infty$, then the asymptotic distribution of $\mathcal{L}_n$ is χ_{p-1}^2 (chi-square distribution with $p-1$ D.F.)*

We recall that we started with two estimators of $\boldsymbol{\delta}$, namely, the unrestricted estimator, $\tilde{\boldsymbol{\delta}}_n$ and the restricted estimator, $\hat{\boldsymbol{\delta}}_{0n}\mathbf{1}_p$. Meta analysis is complete if H_0 is accepted with a reasonable level of significance. Then, one reports the estimator $\hat{\boldsymbol{\delta}}_{0n}$ with its $100(1 - \beta)\%$ confidence interval in (3.23), otherwise, the combination of the effect sizes may be carried out in a random effects model, see Hartung, Knapp and Sinha (2008). However, there is more to it than what meets the eye. In the next section, we discuss the new aspects of meta analysis.

3.4.1 Quasi-Empirical Bayes Methods of Meta Analysis

Consider the effect size vector $\boldsymbol{\delta} = (\delta_1, \dots, \delta_p)'$ for p series of independent experiments with the unbiased estimator $\tilde{\boldsymbol{\delta}}_n = (\tilde{\delta}_1, \dots, \tilde{\delta}_p)'$. We suspect that $\boldsymbol{\delta}$ may be equal to $\delta_0\mathbf{1}_p$. We define the restricted estimator of $\boldsymbol{\delta}$ as we believe that $\boldsymbol{\delta} = \delta_0\mathbf{1}_p$, is truly valid, as

$$\hat{\boldsymbol{\delta}}_n = \hat{\boldsymbol{\delta}}_{0n}\mathbf{1}_p. \quad (3.25)$$

where $\hat{\boldsymbol{\delta}}_{0n}$ is given by (3.22). In most cases, however, it may not be known that the hypothesis $H_0 : \boldsymbol{\delta} = \delta_0\mathbf{1}_p$ is true, but only suspected to be true. Thus, it would be desirable to take into account of this suspected hypothesis in the estimation of $\boldsymbol{\delta}$.

A reasonable choice would be some convex combination of the two estimators $\tilde{\boldsymbol{\delta}}_n$ and $\hat{\boldsymbol{\delta}}_n$. Let $g(\mathcal{L}_n)$ be a real valued function of $\mathcal{L}_n$, then define the estimator as

$$\hat{\boldsymbol{\delta}}_n^g = \hat{\boldsymbol{\delta}}_n + [1 - g(\mathcal{L}_n)](\tilde{\boldsymbol{\delta}}_n - \hat{\boldsymbol{\delta}}_n).$$

We consider the following five choices of $g(\mathcal{L}_n)$:

$$(i) \quad g(\mathcal{L}_n) = 0, \quad \hat{\boldsymbol{\delta}}_n^{(0)} = \tilde{\boldsymbol{\delta}}_n, \text{ the unrestricted estimator,} \quad (3.26)$$

$$(ii) \quad g(\mathcal{L}_n) = 1, \quad \hat{\boldsymbol{\delta}}_n^{(1)} = \hat{\boldsymbol{\delta}}_n, \text{ the restricted estimator,} \quad (3.27)$$

$$(iii) \quad g(\mathcal{L}_n) = I(\mathcal{L}_n \leq \chi_{p-1, \alpha}^2), \quad \hat{\boldsymbol{\delta}}_n^{PT} = \tilde{\boldsymbol{\delta}}_n - (\tilde{\boldsymbol{\delta}}_n - \hat{\boldsymbol{\delta}}_n)I(\mathcal{L}_n \leq \chi_{p-1, \alpha}^2) \quad (3.28)$$

where $\chi_{p-1, \alpha}^2$ is the upper α -level critical value from the central chi-square distribution with $p - 1$ D.F. This estimator is called the “ preliminary test estimator” PTE.

$$(iv) \quad g(\mathcal{L}_n) = (p - 3)\mathcal{L}_n^{-1}, \quad \hat{\boldsymbol{\delta}}_n^{JS} = \tilde{\boldsymbol{\delta}}_n - (p - 3)(\tilde{\boldsymbol{\delta}}_n - \hat{\boldsymbol{\delta}}_n)\mathcal{L}_n^{-1}, \quad (3.29)$$

yields the James-Stein estimator (JS), which may rewritten as

$$\hat{\boldsymbol{\delta}}_n^{JS} = \hat{\boldsymbol{\delta}}_n + [1 - (p - 3)\mathcal{L}_n^{-1}](\tilde{\boldsymbol{\delta}}_n - \hat{\boldsymbol{\delta}}_n), \quad p > 3.$$

(v) $g(\mathcal{L}_n) = 1 - [1 - (p - 3)\mathcal{L}_n^{-1}]I(\mathcal{L}_n > (p - 3))$, yields the positive-rule Stein estimator (PRSE) given by

$$\hat{\boldsymbol{\delta}}_n^{s+} = \hat{\boldsymbol{\delta}}_n + [1 - (p - 3)\mathcal{L}_n^{-1}]I(\mathcal{L}_n > (p - 3))(\tilde{\boldsymbol{\delta}}_n - \hat{\boldsymbol{\delta}}_n), \quad p > 3. \quad (3.30)$$

The positive-rule Stein estimator is a PTE given by

$$\hat{\boldsymbol{\delta}}_n I(\mathcal{L}_n < (p - 3)) + \hat{\boldsymbol{\delta}}_n^{JS} I(\mathcal{L}_n > (p - 3)).$$

We call them “ quasi-empirical Bayes estimators ” (QEBE) as defined by Saleh (2006, Chapter 4, Section 4.4.3).

We can show that these estimators are similar to “empirical Bayes estimators ” (see Srivastava and Saleh (2005)) corresponding to the prior distribution of $\boldsymbol{\delta}$ given

by

$$\boldsymbol{\delta} \sim N_p(\delta_0 \mathbf{1}_p, \tau \boldsymbol{\Sigma}_n),$$

for some choice of $\tau \geq 0$. Since given $\boldsymbol{\delta}$, the asymptotic distribution of $\tilde{\boldsymbol{\delta}}_n$ is $N_p(\boldsymbol{\delta}, \boldsymbol{\Sigma}_n)$, where $\boldsymbol{\Sigma}_n$ is given by (3.19). It follows that the asymptotic marginal distribution of $\tilde{\boldsymbol{\delta}}_n$ is

$$N_p(\delta_0 \mathbf{1}_p, (1 + \tau) \boldsymbol{\Sigma}_n), \quad (3.31)$$

and the asymptotic conditional distribution of $\boldsymbol{\delta}$ given $\tilde{\boldsymbol{\delta}}_n$ is $N_p(\boldsymbol{\delta}^*, \frac{\tau}{(1+\tau)} \boldsymbol{\Sigma}_n)$, where $\boldsymbol{\delta}^* = (1 + \tau)^{-1} \delta_0 \mathbf{1}_p + \tau(1 + \tau)^{-1} \tilde{\boldsymbol{\delta}}_n$, a convex combination of $\delta_0 \mathbf{1}_p$ and $\tilde{\boldsymbol{\delta}}_n$.

Thus, an empirical Bayes estimator (EBE) of $\boldsymbol{\delta}$ can be obtained by using the estimator of $\delta_0 \mathbf{1}_p$ given by (3.25) for fixed τ which is a convex combination of $\tilde{\boldsymbol{\delta}}_n$ and $\hat{\boldsymbol{\delta}}_n$ for any given $\tau \geq 0$, namely,

$$\begin{aligned} \boldsymbol{\delta}_n^{EB}(\tau) &= (1 + \tau)^{-1} \hat{\boldsymbol{\delta}}_n + \tau(1 + \tau)^{-1} \tilde{\boldsymbol{\delta}}_n \\ &= \tilde{\boldsymbol{\delta}}_n - (1 + \tau)^{-1} (\tilde{\boldsymbol{\delta}}_n - \hat{\boldsymbol{\delta}}_n), \end{aligned} \quad (3.32)$$

where τ is arbitrary and unknown. If $\tau = \infty$, we get $\boldsymbol{\delta}_n^{EB}(\infty) = \tilde{\boldsymbol{\delta}}_n$ and if $\tau = 0$, we get $\boldsymbol{\delta}_n^{EB}(0) = \hat{\boldsymbol{\delta}}_n$. Thus, we need to estimate $(1 + \tau)^{-1}$. For this, we consider the marginal distribution of $\tilde{\boldsymbol{\delta}}_n$ as given in (3.31). Let $\mathbf{C}$ be a $p \times (p - 1)$ matrix of rank $p - 1$ such that $\mathbf{C}' \mathbf{1}_p = \mathbf{0}$. Then,

$$\mathbf{C}' \tilde{\boldsymbol{\delta}}_n \sim N_p\left(0, (1 + \tau)(\mathbf{C}' \boldsymbol{\Sigma}_n \mathbf{C})\right).$$

Hence,

$$n(1 + \tau)^{-1} \tilde{\boldsymbol{\delta}}_n' \mathbf{C} (\mathbf{C}' n \boldsymbol{\Sigma}_n \mathbf{C})^{-1} \mathbf{C}' \tilde{\boldsymbol{\delta}}_n \xrightarrow{\mathcal{D}} \chi_{p-1}^2.$$

Now by corollary A.3.1

$$\begin{aligned} \mathbf{C}(\mathbf{C}'n\boldsymbol{\Sigma}_n\mathbf{C})^{-1}\mathbf{C}' &= (n\boldsymbol{\Sigma}_n)^{-1} - \frac{(n\boldsymbol{\Sigma}_n)^{-1}\mathbf{1}_p\mathbf{1}_p'(n\boldsymbol{\Sigma}_n)^{-1}}{\mathbf{1}_p'(n\boldsymbol{\Sigma}_n)^{-1}\mathbf{1}_p} \\ &= (n\boldsymbol{\Sigma}_n)^{-1}\mathbf{J}_n, \end{aligned}$$

where

$$\mathbf{J}_n = \mathbf{I}_p - \frac{\mathbf{1}_p\mathbf{1}_p'(n\boldsymbol{\Sigma}_n)^{-1}}{\mathbf{1}_p'(n\boldsymbol{\Sigma}_n)^{-1}\mathbf{1}_p}$$

Now, we rewrite $\mathcal{L}_n$ in (3.24) as

$$\begin{aligned} \mathcal{L}_n &= n \tilde{\boldsymbol{\delta}}_n' \hat{\mathbf{J}}_n'(n\hat{\boldsymbol{\Sigma}}_n)^{-1} \hat{\mathbf{J}}_n \tilde{\boldsymbol{\delta}}_n \\ &= n \tilde{\boldsymbol{\delta}}_n' (n\hat{\boldsymbol{\Sigma}}_n)^{-1} \hat{\mathbf{J}}_n \tilde{\boldsymbol{\delta}}_n \\ &= n \tilde{\boldsymbol{\delta}}_n' \mathbf{C}(\mathbf{C}'n\hat{\boldsymbol{\Sigma}}_n\mathbf{C})^{-1}\mathbf{C}' \tilde{\boldsymbol{\delta}}_n, \end{aligned} \tag{3.33}$$

Hence,

$$(1 + \tau)^{-1} \mathcal{L}_n \xrightarrow{\mathcal{D}} \chi_{p-1}^2,$$

thus using Theorem A.2.2, the asymptotic distributional mean of $(1 + \tau)\mathcal{L}_n^{-1}$ is $\frac{1}{p-3}$, $p > 3$. Hence, $(1 + \tau)^{-1}$ can be estimated by a scalar multiple of $\mathcal{L}_n^{-1}$. More generally, let $g(\mathcal{L}_n)$, a real valued function of $\mathcal{L}_n$, be an estimate of $(1 + \tau)^{-1}$. Then, the estimator (3.32) becomes

$$\begin{aligned} \delta_n^{EB}(\hat{\tau}) &= \tilde{\boldsymbol{\delta}}_n - g(\mathcal{L}_n)(\tilde{\boldsymbol{\delta}}_n - \hat{\boldsymbol{\delta}}_n) \\ &= \hat{\boldsymbol{\delta}}_n + [1 - g(\mathcal{L}_n)](\tilde{\boldsymbol{\delta}}_n - \hat{\boldsymbol{\delta}}_n). \end{aligned}$$

Thus, our estimators in (3.26) - (3.30) can be obtained by considering the five choices in (i) - (v) of $g(\mathcal{L}_n)$.

3.4.2 Asymptotic Distributional Properties of the Estimators

We consider the sequence of local alternatives for the derivation of the asymptotic distribution of the estimators

$$\mathcal{K}_{(n)} : \boldsymbol{\delta}_{(n)} = \delta_0 \mathbf{1}_p + n^{-1/2} \boldsymbol{\xi}, \quad \boldsymbol{\xi} = (\xi_1, \dots, \xi_p)' \neq \mathbf{0}. \quad (3.34)$$

Note that for $\boldsymbol{\xi} = \mathbf{0}$, $\boldsymbol{\delta}_{(n)} = \delta_0 \mathbf{1}_p$, that is the null hypothesis in (3.17) is a particular case of (3.34). We arrive at the next theorem on the asymptotic distributions of the various statistics

Theorem 3.4.3. *Under $\{\mathcal{K}_{(n)}\}$ and as $\frac{n_i^E}{n_i} \rightarrow \pi_i^E$, $\frac{n_i^C}{n_i} \rightarrow \pi_i^C$, $\frac{n_i}{n} \rightarrow \pi_i$, π_i^E , π_i^C and π_i remain fixed as $\min_{1 \leq i \leq p} n_i \rightarrow \infty$ which implies $n \rightarrow \infty$, the following hold:*

- (i) $\sqrt{n}(\tilde{\boldsymbol{\delta}}_n - \boldsymbol{\delta}_{(n)}) \xrightarrow{\mathcal{D}} \mathbf{Z}_0 \sim N_p(\mathbf{0}, \boldsymbol{\Sigma}_0)$ and $\boldsymbol{\Sigma}_0$ given in (3.20).
- (ii) $\sqrt{n}(\tilde{\boldsymbol{\delta}}_n - \hat{\boldsymbol{\delta}}_n) \xrightarrow{\mathcal{D}} \mathbf{Z}_1, \mathbf{Z}_1 \sim N_p(\mathbf{J}\boldsymbol{\xi}, \boldsymbol{\Sigma}_0 \mathbf{J}'), \mathbf{J} = \mathbf{I}_p - \frac{\mathbf{1}_p \mathbf{1}_p' \boldsymbol{\Sigma}_0^{-1}}{\mathbf{1}_p' \boldsymbol{\Sigma}_0^{-1} \mathbf{1}_p}.$
- (iii) $\sqrt{n}(\hat{\boldsymbol{\delta}}_n - \boldsymbol{\delta}_{(n)}) \xrightarrow{\mathcal{D}} \mathbf{Z}_2 - \mathbf{J}\boldsymbol{\xi}, \mathbf{Z}_2 \sim N_p(\mathbf{0}, \mathbf{B}), \mathbf{B} = \frac{\mathbf{1}_p \mathbf{1}_p'}{\mathbf{1}_p' \boldsymbol{\Sigma}_0^{-1} \mathbf{1}_p}.$
- (iv) $\begin{pmatrix} \sqrt{n}(\tilde{\boldsymbol{\delta}}_n - \boldsymbol{\delta}_{(n)}) \\ \sqrt{n}(\tilde{\boldsymbol{\delta}}_n - \hat{\boldsymbol{\delta}}_n) \end{pmatrix} \xrightarrow{\mathcal{D}} \begin{pmatrix} \mathbf{Z}_0 \\ \mathbf{Z}_1 \end{pmatrix} \sim N_{2p} \left(\begin{bmatrix} \mathbf{0} \\ \mathbf{J}\boldsymbol{\xi} \end{bmatrix}, \begin{bmatrix} \boldsymbol{\Sigma}_0 & \boldsymbol{\Sigma}_0 \mathbf{J}' \\ \boldsymbol{\Sigma}_0 \mathbf{J}' & \boldsymbol{\Sigma}_0 \mathbf{J}' \end{bmatrix} \right).$
- (v) $\begin{pmatrix} \sqrt{n}(\hat{\boldsymbol{\delta}}_n - \boldsymbol{\delta}_{(n)}) \\ \sqrt{n}(\tilde{\boldsymbol{\delta}}_n - \hat{\boldsymbol{\delta}}_n) \end{pmatrix} \xrightarrow{\mathcal{D}} \begin{pmatrix} \mathbf{Z}_2 - \mathbf{J}\boldsymbol{\xi} \\ \mathbf{Z}_1 \end{pmatrix} \sim N_{2p} \left(\begin{bmatrix} -\mathbf{J}\boldsymbol{\xi} \\ \mathbf{J}\boldsymbol{\xi} \end{bmatrix}, \begin{bmatrix} \mathbf{B} & \mathbf{0} \\ \mathbf{0} & \boldsymbol{\Sigma}_0 \mathbf{J}' \end{bmatrix} \right).$
- (vi) $\lim_{n \rightarrow \infty} P\{\mathcal{L}_n \leq x | \mathcal{K}_{(n)}\} = H_{p-1}(x, \Delta^2)$ where $\Delta^2 = \boldsymbol{\eta}' \boldsymbol{\Sigma}_0^{-1} \boldsymbol{\eta}$, $\boldsymbol{\eta} = \mathbf{J}\boldsymbol{\xi}$.

where $H_{p-1}(\cdot, \Delta^2)$ is the cdf of noncentral chi-square distribution with noncentrality parameter $\frac{1}{2}\Delta^2$ and $p-1$ D.F.

Proof. The part (i) is a direct consequence of the large sample normal approximation

to the noncentral t-distribution. Parts (ii) – (v) of the theorem follow by applying part (i) of the theorem. Next, we proof the last part of the theorem.

By (ii) of the theorem, the asymptotic distribution of $\mathcal{L}_n$ is the same as that of $\mathbf{Z}'_1 \boldsymbol{\Sigma}_0^{-1} \mathbf{Z}_1$. This reduces the problem to the problem of finding the distribution of quadratic forms of multivariate normal random variables.

Apply Theorem A.2.1, then

$$\mathbf{Z}'_1 \boldsymbol{\Sigma}_0^{-1} \mathbf{Z}_1 \stackrel{\mathcal{D}}{=} \sum_{i=1}^p \lambda_i X_i^2,$$

where $X_1^2, \dots, X_p^2$ are independent chi square variables with one degree of freedom each and λ_i are the eigenvalues of

$$\mathbf{M} = \boldsymbol{\Sigma}_0^{-\frac{1}{2}} \boldsymbol{\Sigma}_0 \mathbf{J}' \boldsymbol{\Sigma}_0^{-\frac{1}{2}}.$$

Note that $\boldsymbol{\Sigma}_0 \mathbf{J}'$ can be written as $\boldsymbol{\Sigma}_0 \mathbf{J}' = \boldsymbol{\Sigma}_0 - \mathbf{B}$, therefore,

$$\mathbf{M} = \mathbf{I}_p - \boldsymbol{\Sigma}_0^{-\frac{1}{2}} \mathbf{B} \boldsymbol{\Sigma}_0^{-\frac{1}{2}}.$$

$\boldsymbol{\Sigma}_0^{-\frac{1}{2}} \mathbf{B} \boldsymbol{\Sigma}_0^{-\frac{1}{2}}$ can be written in the form of one matrix with trace=1. It is easily seen that $\mathbf{M}$ is idempotent matrix with rank, $r(\mathbf{M}) = tr(\mathbf{M}) = p - 1$, for definition of idempotent matrix and its rank see A.3.

Hence, we established that $\mathcal{L}_n$ has an asymptotic chi square distribution with $p - 1$ degrees of freedom and noncentrality parameter $\Delta^2 = (\mathbf{J}\boldsymbol{\xi})' \boldsymbol{\Sigma}_0^{-1} (\mathbf{J}\boldsymbol{\xi})$. $\square$

Based on Theorem 3.4.3 we obtain the following results.

$$\sqrt{n} (\hat{\boldsymbol{\delta}}_n^{PT} - \boldsymbol{\delta}_{(n)}) \xrightarrow{\mathcal{D}} \mathbf{Z}_0 - \mathbf{Z}_1 I[\mathbf{Z}'_1 \boldsymbol{\Sigma}_0^{-1} \mathbf{Z}_1 < \chi_{p-1, \alpha}^2] \quad (3.35)$$

$$\sqrt{n} (\hat{\boldsymbol{\delta}}_n^{JS} - \boldsymbol{\delta}_{(n)}) \xrightarrow{\mathcal{D}} \mathbf{Z}_0 - (p - 3) \mathbf{Z}_1 [\mathbf{Z}'_1 \boldsymbol{\Sigma}_0^{-1} \mathbf{Z}_1]^{-1} \quad (3.36)$$

$$\begin{aligned}
\sqrt{n} (\hat{\boldsymbol{\delta}}_n^{s+} - \boldsymbol{\delta}_{(n)}) &\xrightarrow{\mathcal{D}} \mathbf{Z}_0 - (p-3)\mathbf{Z}_1[\mathbf{Z}_1'\boldsymbol{\Sigma}_0^{-1}\mathbf{Z}_1]^{-1} \\
&\quad - \mathbf{Z}_1\{1 - (p-3)[\mathbf{Z}_1'\boldsymbol{\Sigma}_0^{-1}\mathbf{Z}_1]^{-1}\} \\
&\quad \times I[\mathbf{Z}_1'\boldsymbol{\Sigma}_0^{-1}\mathbf{Z}_1 < p-3].
\end{aligned} \tag{3.37}$$

Define the asymptotic distributional bias of the estimator $\boldsymbol{\delta}_n^*$, where $\boldsymbol{\delta}_n^*$ stands for $\tilde{\boldsymbol{\delta}}_n, \hat{\boldsymbol{\delta}}_n, \hat{\boldsymbol{\delta}}_n^{PT}, \hat{\boldsymbol{\delta}}_n^{JS}$, and $\hat{\boldsymbol{\delta}}_n^{s+}$ by

$$b(\boldsymbol{\delta}_n^*) = \lim_{n \rightarrow \infty} E[\sqrt{n} (\boldsymbol{\delta}_n^* - \boldsymbol{\delta}_{(n)})]. \tag{3.38}$$

Also, define the asymptotic distributional MSE matrices by

$$M(\boldsymbol{\delta}_n^*) = \lim_{n \rightarrow \infty} E[n (\boldsymbol{\delta}_n^* - \boldsymbol{\delta}_{(n)})(\boldsymbol{\delta}_n^* - \boldsymbol{\delta}_{(n)})']. \tag{3.39}$$

Theorem 3.4.4 and Theorem 3.4.5 provide expressions for the asymptotic distributional bias and asymptotic distributional MSE for the five estimators respectively.

Theorem 3.4.4. *Asymptotic distributional bias expressions of the UE, RE, PTE, JSE and PRSE under $\{\mathcal{K}_{(n)}\}$ in (3.34) are given by*

$$\begin{aligned}
b_1(\tilde{\boldsymbol{\delta}}_n) &= \mathbf{0}. \\
b_2(\hat{\boldsymbol{\delta}}_n) &= -\boldsymbol{\eta}. \\
b_3(\hat{\boldsymbol{\delta}}_n^{PT}) &= -\boldsymbol{\eta}H_{p+1}(\chi_{p-1}^2(\alpha), \Delta^2). \\
b_4(\hat{\boldsymbol{\delta}}_n^{JS}) &= -(p-3)\boldsymbol{\eta}E(\chi_{p+1}^{-2}(\Delta^2)). \\
b_5(\hat{\boldsymbol{\delta}}_n^{s+}) &= -(p-3)\boldsymbol{\eta}E[\chi_{p+1}^{-2}(\Delta^2)] - \boldsymbol{\eta}E[1 - (p-3)\chi_{p+1}^{-2}(\Delta^2)] \\
&\quad \times I[\chi_{p+1}^2(\Delta^2) \leq p-3].
\end{aligned}$$

Proof. Based on Theorem 3.4.3, we have

$$\begin{aligned} b_1(\tilde{\boldsymbol{\delta}}_n) &= E(\mathbf{Z}_0) = \mathbf{0}. \\ b_2(\hat{\boldsymbol{\delta}}_n) &= E(\mathbf{Z}_2 - \mathbf{J}\boldsymbol{\xi}) = -\mathbf{J}\boldsymbol{\xi} = -\boldsymbol{\eta}. \end{aligned}$$

For PTE using expression (3.35), we have

$$b_3(\hat{\boldsymbol{\delta}}_n^{PT}) = E\{\mathbf{Z}_0 - \mathbf{Z}_1 I[\mathbf{Z}_1' \boldsymbol{\Sigma}_0^{-1} \mathbf{Z}_1 < \chi_{p-1, \alpha}^2]\}$$

By Theorem A.2.6, we obtain

$$b_3(\hat{\boldsymbol{\delta}}_n^{PT}) = -\boldsymbol{\eta} H_{p+1}(\chi_{p-1}^2(\alpha), \Delta^2).$$

For JSE using expression (3.36), we have

$$b_4(\hat{\boldsymbol{\delta}}_n^{JS}) = E\{\mathbf{Z}_0 - (p-3)\mathbf{Z}_1[\mathbf{Z}_1' \boldsymbol{\Sigma}_0^{-1} \mathbf{Z}_1]^{-1}\}$$

By Theorem A.2.6, we obtain

$$b_4(\hat{\boldsymbol{\delta}}_n^{JS}) = -(p-3)\boldsymbol{\eta} E(\chi_{p+1}^{-2}(\Delta^2)).$$

For PRSE using expression (3.37), we have

$$\begin{aligned} b_5(\hat{\boldsymbol{\delta}}_n^{s+}) &= E\{\mathbf{Z}_0 - (p-3)\mathbf{Z}_1[\mathbf{Z}_1' \boldsymbol{\Sigma}_0^{-1} \mathbf{Z}_1]^{-1} \\ &\quad - \mathbf{Z}_1\{1 - (p-3)[\mathbf{Z}_1' \boldsymbol{\Sigma}_0^{-1} \mathbf{Z}_1]^{-1}\} \\ &\quad \times I[\mathbf{Z}_1' \boldsymbol{\Sigma}_0^{-1} \mathbf{Z}_1 < p-3]\} \end{aligned}$$

By Theorem A.2.6, we obtain

$$\begin{aligned} b_5(\hat{\delta}_n^{s+}) &= -(p-3)\boldsymbol{\eta}E[\chi_{p+1}^{-2}(\Delta^2)] - \boldsymbol{\eta}E[1 - (p-3)\chi_{p+1}^{-2}(\Delta^2)] \\ &\quad \times I[\chi_{p+1}^2(\Delta^2) \leq p-3]. \end{aligned}$$

□

Theorem 3.4.5. *Asymptotic distributional MSE expressions of the UE, RE, PTE, JSE and PRSE under $\{\mathcal{K}_{(n)}\}$ in (3.34) are given by*

$$M_1(\tilde{\delta}_n) = \Sigma_0. \quad (3.40)$$

$$M_2(\hat{\delta}_n) = \mathbf{B} + \boldsymbol{\eta}\boldsymbol{\eta}'. \quad (3.41)$$

$$\begin{aligned} M_3(\hat{\delta}_n^{PT}) &= \Sigma_0 - \Sigma_0 \mathbf{J}' H_{p+1}(\chi_{p-1}^2(\alpha), \Delta^2) \\ &\quad + \boldsymbol{\eta}\boldsymbol{\eta}' \{2H_{p+1}(\chi_{p-1}^2(\alpha), \Delta^2) - H_{p+3}(\chi_{p-1}^2(\alpha), \Delta^2)\}. \end{aligned} \quad (3.42)$$

$$\begin{aligned} M_4(\hat{\delta}_n^{JS}) &= \Sigma_0 - (p-3)\Sigma_0 \mathbf{J}' \{2E[\chi_{p+1}^{-2}(\Delta^2)] - (p-3)E[\chi_{p+1}^{-4}(\Delta^2)]\} \\ &\quad + (p-3)(p+1)\boldsymbol{\eta}\boldsymbol{\eta}' E[\chi_{p+3}^{-4}(\Delta^2)]. \end{aligned} \quad (3.43)$$

$$\begin{aligned} M_5(\hat{\delta}_n^{s+}) &= M_4(\hat{\delta}_n^{JS}) - (p-3)\Sigma_0 \mathbf{J}' \\ &\quad \times \{E[(1 - (p-3)\chi_{p+1}^{-2}(\Delta^2))^2 I(\chi_{p+1}^2(\Delta^2) \leq p-3)]\} \\ &\quad + \boldsymbol{\eta}\boldsymbol{\eta}' \{2E[(1 - (p-3)\chi_{p+1}^{-2}(\Delta^2)) I(\chi_{p+1}^2(\Delta^2) \leq p-3)] \\ &\quad - E[(1 - (p-3)\chi_{p+3}^{-2}(\Delta^2))^2 I(\chi_{p+3}^2(\Delta^2) \leq p-3)]\}. \end{aligned} \quad (3.44)$$

Proof. Based on Theorem 3.4.3, we have

$$M_1(\tilde{\delta}_n) = E(\mathbf{Z}_0 \mathbf{Z}_0') = \Sigma_0.$$

$$M_2(\hat{\delta}_n) = E[(\mathbf{Z}_2 - \mathbf{J}\boldsymbol{\xi})(\mathbf{Z}_2 - \mathbf{J}\boldsymbol{\xi})'] = \mathbf{B} + \boldsymbol{\eta}\boldsymbol{\eta}'.$$

For PTE, using expression (3.35), we have

$$\begin{aligned}
M_3(\hat{\boldsymbol{\delta}}_n^{PT}) &= E\left\{\left(\mathbf{Z}_0 - \mathbf{Z}_1 I[\mathbf{Z}'_1 \boldsymbol{\Sigma}_0^{-1} \mathbf{Z}_1 < \chi_{p-1, \alpha}^2]\right) \right. \\
&\quad \left. \times \left(\mathbf{Z}_0 - \mathbf{Z}_1 I[\mathbf{Z}'_1 \boldsymbol{\Sigma}_0^{-1} \mathbf{Z}_1 < \chi_{p-1, \alpha}^2]\right)' \right\} \\
&= E(\mathbf{Z}_0 \mathbf{Z}'_0) - 2E\{\mathbf{Z}_0 \mathbf{Z}'_1 I[\mathbf{Z}'_1 \boldsymbol{\Sigma}_0^{-1} \mathbf{Z}_1 < \chi_{p-1, \alpha}^2]\} \\
&\quad + E\{\mathbf{Z}_1 \mathbf{Z}'_1 I[\mathbf{Z}'_1 \boldsymbol{\Sigma}_0^{-1} \mathbf{Z}_1 < \chi_{p-1, \alpha}^2]\}
\end{aligned} \tag{3.45}$$

The second term of the above expression can be evaluated as

$$\begin{aligned}
&E\{\mathbf{Z}_0 \mathbf{Z}'_1 I[\mathbf{Z}'_1 \boldsymbol{\Sigma}_0^{-1} \mathbf{Z}_1 < \chi_{p-1, \alpha}^2]\} \\
&= E\{E[\mathbf{Z}_0 \mathbf{Z}'_1 I[\mathbf{Z}'_1 \boldsymbol{\Sigma}_0^{-1} \mathbf{Z}_1 < \chi_{p-1, \alpha}^2] | \mathbf{Z}_1]\} \\
&= E\{E(\mathbf{Z}_0 | \mathbf{Z}_1) \mathbf{Z}'_1 I[\mathbf{Z}'_1 \boldsymbol{\Sigma}_0^{-1} \mathbf{Z}_1 < \chi_{p-1, \alpha}^2]\}
\end{aligned} \tag{3.46}$$

Using Theorem 3.4.3, one gets

$$E(\mathbf{Z}_0 | \mathbf{Z}_1) = \mathbf{Z}_1 - \mathbf{J}\boldsymbol{\xi}. \tag{3.47}$$

Hence, $E\{\mathbf{Z}_0 \mathbf{Z}'_1 I[\mathbf{Z}'_1 \boldsymbol{\Sigma}_0^{-1} \mathbf{Z}_1 < \chi_{p-1, \alpha}^2]\}$

$$\begin{aligned}
&= E\{(\mathbf{Z}_1 - \mathbf{J}\boldsymbol{\xi}) \mathbf{Z}'_1 I[\mathbf{Z}'_1 \boldsymbol{\Sigma}_0^{-1} \mathbf{Z}_1 < \chi_{p-1, \alpha}^2]\} \\
&= E\{\mathbf{Z}_1 \mathbf{Z}'_1 I[\mathbf{Z}'_1 \boldsymbol{\Sigma}_0^{-1} \mathbf{Z}_1 < \chi_{p-1, \alpha}^2]\} \\
&\quad - \mathbf{J}\boldsymbol{\xi} E[\mathbf{Z}'_1 I[\mathbf{Z}'_1 \boldsymbol{\Sigma}_0^{-1} \mathbf{Z}_1 < \chi_{p-1, \alpha}^2]]
\end{aligned} \tag{3.48}$$

Substituting (3.48) in (3.45) and applying Theorem A.2.6 the result of (3.42) follows.

Next, we prove (3.43) using expression (3.36), we obtain

$$\begin{aligned}
M_4(\hat{\delta}_n^{JS}) &= E\left\{\left(\mathbf{Z}_0 - (p-3)\mathbf{Z}_1[\mathbf{Z}'_1\boldsymbol{\Sigma}_0^{-1}\mathbf{Z}_1]^{-1}\right)\right. \\
&\quad \left.\times \left(\mathbf{Z}_0 - (p-3)\mathbf{Z}_1[\mathbf{Z}'_1\boldsymbol{\Sigma}_0^{-1}\mathbf{Z}_1]^{-1}\right)'\right\} \\
&= E(\mathbf{Z}_0\mathbf{Z}'_0) - 2(p-3)E[\mathbf{Z}_0\mathbf{Z}_1[\mathbf{Z}'_1\boldsymbol{\Sigma}_0^{-1}\mathbf{Z}_1]^{-1}] \\
&\quad + (p-3)^2E[\mathbf{Z}_1\mathbf{Z}'_1[\mathbf{Z}'_1\boldsymbol{\Sigma}_0^{-1}\mathbf{Z}_1]^{-2}].
\end{aligned}$$

The second term of the above expression can be evaluated as

$$\begin{aligned}
E\{\mathbf{Z}_0\mathbf{Z}'_1[\mathbf{Z}'_1\boldsymbol{\Sigma}_0^{-1}\mathbf{Z}_1]^{-1}\} &= E\{E[\mathbf{Z}_0\mathbf{Z}'_1[\mathbf{Z}'_1\boldsymbol{\Sigma}_0^{-1}\mathbf{Z}_1]^{-1}|\mathbf{Z}_1]\} \\
&= E\{E(\mathbf{Z}_0|\mathbf{Z}_1)\mathbf{Z}'_1[\mathbf{Z}'_1\boldsymbol{\Sigma}_0^{-1}\mathbf{Z}_1]^{-1}\} \tag{3.49}
\end{aligned}$$

Then using (3.47), we get

$$\begin{aligned}
M_4(\hat{\delta}_n^{JS}) &= E(\mathbf{Z}_0\mathbf{Z}'_0) - 2(p-3)E[\mathbf{Z}_1\mathbf{Z}'_1[\mathbf{Z}'_1\boldsymbol{\Sigma}_0^{-1}\mathbf{Z}_1]^{-1}] \\
&\quad + 2(p-3)\mathbf{J}\boldsymbol{\xi}E[\mathbf{Z}_1\mathbf{Z}'_1[\mathbf{Z}'_1\boldsymbol{\Sigma}_0^{-1}\mathbf{Z}_1]^{-1}] \\
&\quad + (p-3)^2E[\mathbf{Z}_1\mathbf{Z}'_1[\mathbf{Z}'_1\boldsymbol{\Sigma}_0^{-1}\mathbf{Z}_1]^{-2}]
\end{aligned}$$

The result of (3.43) follows by applying Theorem A.2.6 and using the identity (A.5).

Next, for PRSE in (3.44) using expression (3.37), we obtain

$$\begin{aligned}
M_5(\hat{\boldsymbol{\delta}}_n^{s+}) &= E\{(\mathbf{Z}_0 - (p-3)\mathbf{Z}_1[\mathbf{Z}'_1\boldsymbol{\Sigma}_0^{-1}\mathbf{Z}_1]^{-1} \\
&\quad - \mathbf{Z}_1\{1 - (p-3)[\mathbf{Z}'_1\boldsymbol{\Sigma}_0^{-1}\mathbf{Z}_1]^{-1}\} \\
&\quad \times I[\mathbf{Z}'_1\boldsymbol{\Sigma}_0^{-1}\mathbf{Z}_1 < p-3]) \\
&\quad \times (\mathbf{Z}_0 - (p-3)\mathbf{Z}_1[\mathbf{Z}'_1\boldsymbol{\Sigma}_0^{-1}\mathbf{Z}_1]^{-1} \\
&\quad - \mathbf{Z}_1\{1 - (p-3)[\mathbf{Z}'_1\boldsymbol{\Sigma}_0^{-1}\mathbf{Z}_1]^{-1}\} \\
&\quad \times I[\mathbf{Z}'_1\boldsymbol{\Sigma}_0^{-1}\mathbf{Z}_1 < p-3])'\}.
\end{aligned}$$

Derivation of (3.44) is very similar to the derivation of (3.43). $\square$

We shall compare the five estimators using the weighted quadratic loss function defined by

$$L(\boldsymbol{\delta}_n^*, \boldsymbol{\delta}_{(n)}) = n(\boldsymbol{\delta}_n^* - \boldsymbol{\delta}_{(n)})' \mathbf{W} (\boldsymbol{\delta}_n^* - \boldsymbol{\delta}_{(n)}) = n \operatorname{tr}[\mathbf{W}(\boldsymbol{\delta}_n^* - \boldsymbol{\delta}_{(n)})(\boldsymbol{\delta}_n^* - \boldsymbol{\delta}_{(n)})']. \quad (3.50)$$

where $\mathbf{W}$ is a given positive semi-definite $(p \times p)$ matrix, and $\operatorname{tr}(\mathbf{A}) = \text{trace of } \mathbf{A}$.

The asymptotic distributional quadratic risk (ADQR) of $\boldsymbol{\delta}_n^*$ is given by

$$R(\boldsymbol{\delta}_n^*, \boldsymbol{\delta}_{(n)}) = EL(\boldsymbol{\delta}_n^*, \boldsymbol{\delta}_{(n)}) = \operatorname{tr}(\mathbf{W}M(\boldsymbol{\delta}_n^*)). \quad (3.51)$$

Expressions of the ADQR for the five estimators are as follows:

$$\begin{aligned}
R_1(\tilde{\boldsymbol{\delta}}_n, \mathbf{W}) &= \operatorname{tr}(\mathbf{W}\boldsymbol{\Sigma}_0). \\
R_2(\hat{\boldsymbol{\delta}}_n, \mathbf{W}) &= \operatorname{tr}(\mathbf{W}\mathbf{B}) + \boldsymbol{\eta}'\mathbf{W}\boldsymbol{\eta}. \\
R_3(\hat{\boldsymbol{\delta}}_n^{PT}, \mathbf{W}) &= \operatorname{tr}(\mathbf{W}\boldsymbol{\Sigma}_0) - \operatorname{tr}(\mathbf{W}\boldsymbol{\Sigma}_0\mathbf{J}')H_{p+1}(\chi_{p-1}^2(\alpha), \Delta^2) \\
&\quad + \boldsymbol{\eta}'\mathbf{W}\boldsymbol{\eta}\{2H_{p+1}(\chi_{p-1}^2(\alpha), \Delta^2) - H_{p+3}(\chi_{p-1}^2(\alpha), \Delta^2)\}.
\end{aligned}$$

$$\begin{aligned}
R_4(\hat{\boldsymbol{\delta}}_n^{JS}, \mathbf{W}) &= \text{tr}(\mathbf{W}\boldsymbol{\Sigma}_0) - (p-3)\text{tr}(\mathbf{W}\boldsymbol{\Sigma}_0\mathbf{J}')\{2E[\chi_{p+1}^{-2}(\Delta^2)] \\
&\quad - (p-3)E[\chi_{p+1}^{-4}(\Delta^2)]\} + (p-3)(p+1)\boldsymbol{\eta}'\mathbf{W}\boldsymbol{\eta}E[\chi_{p+3}^{-4}(\Delta^2)]. \\
R_5(\hat{\boldsymbol{\delta}}_n^{s+}, \mathbf{W}) &= R_4(\hat{\boldsymbol{\delta}}_n^{JS}, \mathbf{W}) - (p-3)\text{tr}(\mathbf{W}\boldsymbol{\Sigma}_0\mathbf{J}') \\
&\quad \times E[(1 - (p-3)\chi_{p+1}^{-2}(\Delta^2))^2 I(\chi_{p+1}^2(\Delta^2) \leq p-3)] \\
&\quad + \boldsymbol{\eta}'\mathbf{W}\boldsymbol{\eta}\{2E[(1 - (p-3)\chi_{p+1}^{-2}(\Delta^2))I(\chi_{p+1}^2(\Delta^2) \leq p-3)] \\
&\quad - E[(1 - (p-3)\chi_{p+3}^{-2}(\Delta^2))^2 I(\chi_{p+3}^2(\Delta^2) \leq p-3)]\}.
\end{aligned}$$

3.4.3 Analysis of Risk

We consider the ADQR analysis of the five estimators based on the results of the theorems below.

Theorem 3.4.6. *Let $p \geq 4$ and let the matrix $\mathbf{W}$ in the loss function,*

$$L(\boldsymbol{\delta}_n, \mathbf{W}) = n(\boldsymbol{\delta}_n - \boldsymbol{\delta})\mathbf{W}(\boldsymbol{\delta}_n - \boldsymbol{\delta})'$$

be positive semi-definite, satisfying

$$\{\mathbf{W} : \frac{\text{tr}(\mathbf{W}\boldsymbol{\Sigma}_0\mathbf{J}')}{ch_{\max}(\mathbf{W}\boldsymbol{\Sigma}_0)} \geq \frac{p+1}{2}\}. \quad (3.52)$$

where $ch_{\max}(\mathbf{W}\boldsymbol{\Sigma}_0)$ stands for the largest characteristic root of $\mathbf{W}\boldsymbol{\Sigma}_0$. Then, under $\{\mathcal{K}_{(n)}\}$, the ADQR of $\hat{\boldsymbol{\delta}}_n^{s+}$, $\hat{\boldsymbol{\delta}}_n^{JS}$ and $\tilde{\boldsymbol{\delta}}_n$ can be ordered as :

$$R_5(\hat{\boldsymbol{\delta}}_n^{s+}, \mathbf{W}) \leq R_4(\hat{\boldsymbol{\delta}}_n^{JS}, \mathbf{W}) \leq R_1(\tilde{\boldsymbol{\delta}}_n, \mathbf{W}).$$

uniformly in Δ^2 .

Proof. The ADQR expression for $\hat{\delta}_n^{JS}$ using (A.6) can be written as

$$\begin{aligned} R_4(\hat{\delta}_n^{JS}, \mathbf{W}) &= tr(\mathbf{W}\Sigma_0) - (p-3)tr[\mathbf{W}\Sigma_0\mathbf{J}']\{(p-3)E[\chi_{p+1}^{-4}(\Delta^2)] \\ &\quad + [1 - \frac{(p+1)\boldsymbol{\eta}'\mathbf{W}\boldsymbol{\eta}}{2\Delta^2 tr[\mathbf{W}\Sigma_0]}]2\Delta^2 E[\chi_{p+3}^{-4}(\Delta^2)]\}. \end{aligned}$$

The risk-difference $R_1(\tilde{\delta}_n, \mathbf{W}) - R_4(\hat{\delta}_n^{JS}, \mathbf{W})$ is nonnegative for all $\mathbf{W}$, satisfying (3.52), for all Δ^2 . Hence,

$$R_4(\hat{\delta}_n^{JS}, \mathbf{W}) \leq R_1(\tilde{\delta}_n, \mathbf{W}) \quad \forall (\Delta^2, \mathbf{W}).$$

Similarly, the risk-difference $R_5(\hat{\delta}_n^{s+}, \mathbf{W}) - R_4(\hat{\delta}_n^{JS}, \mathbf{W})$

$$\begin{aligned} &= -(p-3)tr(\mathbf{W}\Sigma_0\mathbf{J}') \\ &\quad \times E[(1 - (p-3)\chi_{p+1}^{-2}(\Delta^2))^2 I(\chi_{p+1}^2(\Delta^2) \leq p-3)] \\ &\quad + \boldsymbol{\eta}'\mathbf{W}\boldsymbol{\eta}\{2E[(1 - (p-3)\chi_{p+1}^{-2}(\Delta^2))I(\chi_{p+1}^2(\Delta^2) \leq p-3)]\} \\ &\quad - E[(1 - (p-3)\chi_{p+3}^{-2}(\Delta^2))^2 I(\chi_{p+3}^2(\Delta^2) \leq p-3)]. \end{aligned}$$

The expression above is non-positive, since

$$0 \leq \chi_{p+1}^2(\Delta^2) \leq p-3 \Rightarrow 1 - (p-3)\chi_{p+1}^{-2}(\Delta^2) \leq 0,$$

which implies $R_5(\hat{\delta}_n^{s+}, \mathbf{W}) - R_4(\hat{\delta}_n^{JS}, \mathbf{W}) \leq 0 \quad \forall \Delta^2$. Hence, the theorem holds. $\square$

Theorem 3.4.7. Under H_0 , the ADQR of estimators $\tilde{\delta}_n$, $\hat{\delta}_n$, $\hat{\delta}_n^{PT}$, $\hat{\delta}_n^{JS}$ and $\hat{\delta}_n^{s+}$ can be ordered as follows:

$$\begin{aligned} R_2(\hat{\delta}_n, \mathbf{W}) &\leq R_5(\hat{\delta}_n^{s+}, \mathbf{W}) \leq R_4(\hat{\delta}_n^{JS}, \mathbf{W}) \leq R_1(\tilde{\delta}_n, \mathbf{W}), \\ R_2(\hat{\delta}_n, \mathbf{W}) &\leq R_3(\hat{\delta}_n^{PT}, \mathbf{W}) \leq R_1(\tilde{\delta}_n, \mathbf{W}), \\ R_3(\hat{\delta}_n^{PT}, \mathbf{W}) &\leq R_5(\hat{\delta}_n^{s+}, \mathbf{W}) \leq R_4(\hat{\delta}_n^{JS}, \mathbf{W}) \leq R_1(\tilde{\delta}_n, \mathbf{W}), \end{aligned}$$

where the order depends on the level of significance, α for the PTE, $\hat{\delta}_n^{PT}$.

Proof. Under H_0 , using Theorem A.2.2, the ADQR of the estimators are

$$\begin{aligned}
R_1(\tilde{\delta}_n, \mathbf{W}) &= tr[\mathbf{W}\Sigma_0]. \\
R_2(\hat{\delta}_n, \mathbf{W}) &= tr[\mathbf{W}\mathbf{B}]. \\
R_3(\hat{\delta}_n^{PT}, \mathbf{W}) &= tr[\mathbf{W}\Sigma_0]\{1 - H_{p+1}(\chi_{p-1}^2(\alpha), 0)\} + tr[\mathbf{W}\mathbf{B}]H_{p+1}(\chi_{p-1}^2(\alpha), 0). \\
R_4(\hat{\delta}_n^{JS}, \mathbf{W}) &= \frac{2}{p-1}tr[\mathbf{W}\Sigma_0] + \frac{p-3}{p-1}tr[\mathbf{W}\mathbf{B}]. \\
R_5(\hat{\delta}_n^{s+}, \mathbf{W}) &= tr[\mathbf{W}\Sigma_0]\left\{\frac{2}{p-1} - (p-3) \int_0^{p-3} (1 - (p-3)\chi_{p+1}^{-2})^2 d H_{p+1}(x, 0)\right\} \\
&\quad + tr[\mathbf{W}\mathbf{B}]\left\{\frac{p-3}{p-1} + (p-3) \int_0^{p-3} (1 - (p-3)\chi_{p+1}^{-2})^2 d H_{p+1}(x, 0)\right\}.
\end{aligned}$$

The risk-difference between $\tilde{\delta}_n$ and $\hat{\delta}_n$ is

$$tr[\mathbf{W}(\Sigma_0 - \mathbf{B})] > 0.$$

Hence, $R_2(\hat{\delta}_n, \mathbf{W}) < R_1(\tilde{\delta}_n, \mathbf{W})$.

The risk-difference between $\tilde{\delta}_n$ and $\hat{\delta}_n^{PT}$, is

$$tr[\mathbf{W}(\Sigma_0 - \mathbf{B})]H_{p+1}(\chi_{p-1}^2(\alpha), 0) \geq 0 \quad \forall \alpha \in (0, 1),$$

and this implies $R_3(\hat{\delta}_n^{PT}, \mathbf{W}) \leq R_1(\tilde{\delta}_n, \mathbf{W})$.

The risk-difference between $\hat{\delta}_n$ and $\hat{\delta}_n^{PT}$, is

$$tr[\mathbf{W}(\Sigma_0 - \mathbf{B})]\{1 - H_{p+1}(\chi_{p-1}^2(\alpha), 0)\} \geq 0 \quad \forall \alpha \in (0, 1),$$

and this implies $R_2(\hat{\delta}_n, \mathbf{W}) \leq R_3(\hat{\delta}_n^{PT}, \mathbf{W}) \quad \forall \alpha \in (0, 1)$. Hence we have

$$R_2(\hat{\delta}_n, \mathbf{W}) \leq R_3(\hat{\delta}_n^{PT}, \mathbf{W}) \leq R_1(\tilde{\delta}_n, \mathbf{W}) \text{ under } H_0.$$

The risk-difference between $\hat{\delta}_n^{JS}$ and $\hat{\delta}_n^{PT}$ is

$$R_4(\hat{\delta}_n^{JS}, \mathbf{W}) - R_3(\hat{\delta}_n^{PT}, \mathbf{W}) = tr[\mathbf{W}(\Sigma_0 - \mathbf{B})]\{H_{p+1}(\chi_{p-1}^2(\alpha), 0) - \frac{p-3}{p-1}\}, p \geq 4.$$

Hence, $\hat{\delta}_n^{PT}$ has a smaller ADQR when

$$H_{p+1}(\chi_{p-1}^2(\alpha), 0) \geq \frac{p-3}{p-1}. \quad (3.53)$$

Otherwise, $\hat{\delta}_n^{JS}$ has a smaller ADQR. Also,

$$R_4(\hat{\delta}_n^{JS}, \mathbf{W}) = R_2(\hat{\delta}_n, \mathbf{W}) + \frac{2}{p-1}tr[\mathbf{W}(\Sigma_0 - \mathbf{B})] \geq R_2(\hat{\delta}_n, \mathbf{W}).$$

Hence,

$$R_2(\hat{\delta}_n, \mathbf{W}) \leq R_4(\hat{\delta}_n^{JS}, \mathbf{W})$$

Thus, whenever (3.53) holds, the order of the ADQR of $\tilde{\delta}_n, \hat{\theta}_n, \hat{\theta}_n^{PT}$, and $\hat{\delta}_n^{JS}$ under H_0 is given by :

$$R_2(\hat{\delta}_n, \mathbf{W}) \leq R_3(\hat{\delta}_n^{PT}, \mathbf{W}) \leq R_4(\hat{\delta}_n^{JS}, \mathbf{W}).$$

Also, note that $R_4(\hat{\delta}_n^{JS}, \mathbf{W}) - R_1(\tilde{\delta}_n, \mathbf{W}) = -\frac{p-3}{p-1}tr[\mathbf{W}(\Sigma_0 - \mathbf{B})] < 0$. Hence, $R_4(\hat{\delta}_n^{JS}, \mathbf{W}) \leq R_1(\tilde{\delta}_n, \mathbf{W})$ which implies that the ADQR order of $\tilde{\delta}_n, \hat{\delta}_n, \hat{\delta}_n^{PT}$ and $\hat{\delta}_n^{JS}$ under H_0 is given by:

$$R_2(\hat{\delta}_n, \mathbf{W}) \leq R_3(\hat{\delta}_n^{PT}, \mathbf{W}) \leq R_4(\hat{\delta}_n^{JS}, \mathbf{W}) \leq R_1(\tilde{\delta}_n, \mathbf{W}).$$

provided that (3.53) holds.

Now consider the ADQR comparison of $\hat{\delta}_n^{s+}$ and $\hat{\delta}_n^{JS}, \tilde{\delta}_n, \hat{\delta}_n$ and $\hat{\delta}_n^{PT}$ under H_0 . First,

compare $\hat{\boldsymbol{\delta}}_n^{JS}$, and $\hat{\boldsymbol{\delta}}_n^{s+}$. The risk difference $R_4(\hat{\boldsymbol{\delta}}_n^{JS}, \mathbf{W}) - R_5(\hat{\boldsymbol{\delta}}_n^{s+}, \mathbf{W})$

$$\begin{aligned} &= (p-3)tr[\mathbf{W}(\boldsymbol{\Sigma}_0 - \mathbf{B})] \int_0^{p-3} (1 - (p-3)\chi_{p+1}^{-2})^2 d H_{p+1}(x, 0) \\ &\leq \frac{2(p-3)}{p-1} tr[\mathbf{W}(\boldsymbol{\Sigma}_0 - \mathbf{B})] \geq 0. \end{aligned}$$

Hence, under H_0 ,

$$R_5(\hat{\boldsymbol{\delta}}_n^{s+}, \mathbf{W}) \leq R_4(\hat{\boldsymbol{\delta}}_n^{JS}, \mathbf{W}) \leq R_1(\tilde{\boldsymbol{\delta}}_n, \mathbf{W}).$$

Next, consider the comparison of $\hat{\boldsymbol{\delta}}_n^{PT}$ and $\hat{\boldsymbol{\delta}}_n^{s+}$. The risk difference is

$$\begin{aligned} R_5(\hat{\boldsymbol{\delta}}_n^{s+}, \mathbf{W}) - R_3(\hat{\boldsymbol{\delta}}_n^{PT}, \mathbf{W}) &= tr[\mathbf{W}(\boldsymbol{\Sigma}_0 - \mathbf{B})] \{ H_{p+1}(\chi_{p-1}^2(\alpha), 0) - \frac{p-3}{p-1} \\ &\quad + (p-3) \int_0^{p-3} (1 - (p-3)\chi_{p+1}^{-2})^2 d H_{p+1}(x, 0) \}. \end{aligned}$$

We know that $\hat{\boldsymbol{\delta}}_n^{PT}$ is better than $\hat{\boldsymbol{\delta}}_n^{JS}$, whenever (3.53) holds. Hence, the risk-difference above is nonnegative whenever (3.53) holds. Using equation (A.2) we can see that (3.53) is equivalent to the density inequality

$$h_{p+1}(\chi_{p-1}^2(\alpha), 0) \leq \frac{2}{p-1} - \alpha, \quad p \geq 4,$$

we note that for the range of α , $\hat{\boldsymbol{\delta}}_n^{PT}$ dominates $\hat{\boldsymbol{\delta}}_n^{s+}$ and satisfies the ordering as given below

$$R_2(\hat{\boldsymbol{\delta}}_n, \mathbf{W}) \leq R_3(\hat{\boldsymbol{\delta}}_n^{PT}, \mathbf{W}) \leq R_5(\hat{\boldsymbol{\delta}}_n^{s+}, \mathbf{W}) \leq R_4(\hat{\boldsymbol{\delta}}_n^{JS}, \mathbf{W}) \leq R_1(\tilde{\boldsymbol{\delta}}_n, \mathbf{W}). \quad \square$$

3.4.4 Asymptotic Relative Efficiency of the Estimators

We consider the asymptotic relative risk efficiency (ARRE) of the estimators compared to $\tilde{\delta}_n$.

- RE

$$\begin{aligned} \text{ARRE}(\hat{\delta}_n; \tilde{\delta}_n) &= R_1(\tilde{\delta}_n, \mathbf{W})[R_2(\hat{\delta}_n, \mathbf{W})]^{-1} \\ &= [\text{tr}[\mathbf{W}\mathbf{B}] + \boldsymbol{\eta}'\mathbf{W}\boldsymbol{\eta}]^{-1}\text{tr}(\mathbf{W}\boldsymbol{\Sigma}_0). \end{aligned} \quad (3.54)$$

- PTE

$$\text{ARRE}(\hat{\delta}_n^{PT}; \tilde{\delta}_n) = R_1(\tilde{\delta}_n, \mathbf{W})[R_3(\hat{\delta}_n^{PT}, \mathbf{W})]^{-1} = [1 + g_{pt}^*(\alpha, \Delta^2)]^{-1}, \quad (3.55)$$

where,

$$\begin{aligned} g_{pt}^*(\alpha, \Delta^2) &= -\text{tr}[\mathbf{W}\boldsymbol{\Sigma}_0\mathbf{J}'][\text{tr}(\mathbf{W}\boldsymbol{\Sigma}_0)]^{-1}H_{p+1}(\chi_{p-1}^2(\alpha); \Delta^2) \\ &\quad + \boldsymbol{\eta}'\mathbf{W}\boldsymbol{\eta}[\text{tr}(\mathbf{W}\boldsymbol{\Sigma}_0)]^{-1}[2H_{p+1}(\chi_{p-1}^2(\alpha), \Delta^2) \\ &\quad - H_{p+3}(\chi_{p-1}^2(\alpha), \Delta^2)]. \end{aligned} \quad (3.56)$$

- JS

$$\text{ARRE}(\hat{\delta}_n^{JS}; \tilde{\delta}_n) = R_1(\tilde{\delta}_n, \mathbf{W})[R_4(\hat{\delta}_n^{JS}, \mathbf{W})]^{-1} = [1 + g_{s*}(\Delta^2)]^{-1}, \quad (3.57)$$

where,

$$\begin{aligned} g_{s*}(\Delta^2) &= -(p-3)\text{tr}[\mathbf{W}\boldsymbol{\Sigma}_0\mathbf{J}'][\text{tr}(\mathbf{W}\boldsymbol{\Sigma}_0)]^{-1}\{2E[\chi_{p+1}^{-2}(\Delta^2)] \\ &\quad - (p-3)E[\chi_{p+1}^{-4}(\Delta^2)]\} + (p-3)(p+1)\boldsymbol{\eta}'\mathbf{W}\boldsymbol{\eta} \\ &\quad \times [\text{tr}(\mathbf{W}\boldsymbol{\Sigma}_0)]^{-1}E[\chi_{p+3}^{-4}(\Delta^2)]. \end{aligned} \quad (3.58)$$

• PRSE

$$\text{ARRE}(\hat{\delta}_n^{s+}; \tilde{\delta}_n) = R_1(\tilde{\delta}_n, \mathbf{W})[R_5(\hat{\delta}_n^{JS}, \mathbf{W})]^{-1} = [1 + g_{s^*}(\Delta^2) + g_{s+}^*(\Delta^2)]^{-1}, \quad (3.59)$$

where,

$$\begin{aligned} g_{s+}^*(\Delta^2) &= -(p-3)\text{tr}[\mathbf{W}\Sigma_0\mathbf{J}'][\text{tr}(\mathbf{W}\Sigma_0)]^{-1}E[(1-(p-3)\chi_{p+1}^{-2}(\Delta^2))^2 \\ &\quad \times I(\chi_{p+1}^{-2}(\Delta^2) < p-3)] + [\text{tr}(\mathbf{W}\Sigma_0)]^{-1}\boldsymbol{\eta}'\mathbf{W}\boldsymbol{\eta}\{2E[(1-(p-3)\chi_{p+1}^{-2}(\Delta^2)) \\ &\quad - 3)\chi_{p+1}^{-2}(\Delta^2))I(\chi_{p+1}^2(\Delta^2) < p-3) - E[(1-(p-3)\chi_{p+3}^{-2}(\Delta^2))^2 \\ &\quad \times I(\chi_{p+3}^2(\Delta^2) < p-3)]\} \end{aligned} \quad (3.60)$$

For some numerical work and graphs we consider the particular case $\mathbf{W} = \Sigma_0^{-1}$. Then the expressions in (3.54)-(3.60) reduce to

$$\begin{aligned} \text{ARRE}(\hat{\delta}_n; \tilde{\delta}_n) &= p[1 + \Delta^2]^{-1} \\ \text{ARRE}(\hat{\delta}_n^{PT}; \tilde{\delta}_n) &= \{1 + (-1 + \frac{1}{p})H_{p+1}(\chi_{p-1}^2(\alpha); \Delta^2) + p^{-1}\Delta^2[2H_{p+1}(\chi_{p-1}^2(\alpha), \Delta^2) \\ &\quad - H_{p+3}(\chi_{p-1}^2(\alpha), \Delta^2)]\}^{-1} \\ \text{ARRE}(\hat{\delta}_n^{JS}; \tilde{\delta}_n) &= \{1 - (1 - \frac{3}{p})(p-1)[2E[\chi_{p+1}^{-2}(\Delta^2)] - (p-3)E[\chi_{p+1}^{-4}(\Delta^2)]] \\ &\quad + (1 - \frac{3}{p})(p+1)\Delta^2E[\chi_{p+3}^{-4}(\Delta^2)]\}^{-1} \\ \text{ARRE}(\hat{\delta}_n^{s+}; \tilde{\delta}_n) &= \{1 - (1 - \frac{3}{p})(p-1)[2E[\chi_{p+1}^{-2}(\Delta^2)] - (p-3)E[\chi_{p+1}^{-4}(\Delta^2)]] \\ &\quad + (1 - \frac{3}{p})(p+1)\Delta^2E[\chi_{p+3}^{-4}(\Delta^2)] - (1 - \frac{3}{p})(p-1)E[(1-(p-3) \\ &\quad \times \chi_{p+1}^{-2}(\Delta^2))^2I(\chi_{p+1}^{-2}(\Delta^2) < p-3)] + p^{-1}\Delta^2[2E[(1-(p-3) \\ &\quad \times \chi_{p+1}^{-2}(\Delta^2))I(\chi_{p+1}^2(\Delta^2) < p-3) - E[(1-(p-3)\chi_{p+3}^{-2}(\Delta^2))^2 \\ &\quad \times I(\chi_{p+3}^2(\Delta^2) < p-3)]\}^{-1} \end{aligned}$$

Denoting the ARRE of PTE, as a function of α , Δ^2 by $ARRE(\alpha, \Delta^2)$. We can see that the graph of $ARRE(\alpha, \Delta^2)$ in Figure 3.5, as a function of Δ^2 for fixed α , is

decreasing, crossing the 1-line to a minimum at $\Delta^2 = \Delta_0^2$; then it increases towards the 1-line as $\Delta^2 \rightarrow \infty$. The maximum value of $ARRE(\alpha, \Delta^2)$ occurs at $\Delta^2 = 0$ with the value

$$\max_{\Delta^2} ARRE(\alpha, \Delta^2) = \{1 + (-1 + \frac{1}{p})H_{p+1}(\chi_{p-1}^2(\alpha); 0)\}^{-1} \geq 1. \quad (3.61)$$

for all $\alpha \in A$, the set of possible values of α .

Intersection of the graph of $ARRE(\alpha; \Delta^2)$ with the 1-line occurs at Δ_1^2 . Thus if $0 \leq \Delta^2 \leq \Delta_1^2$, then $\hat{\delta}_n^{PT}$ performs better than $\tilde{\delta}_n$ and if $\Delta^2 \geq \Delta_1^2$, then $\tilde{\delta}_n$ performs better. The cut-off point Δ_1^2 may be determined by

$$\Delta_1^2 = \frac{(p-1)H_{p+1}(\chi_{p-1}^2(\alpha), \Delta^2)}{[2H_{p+1}(\chi_{p-1}^2(\alpha), \Delta^2) - H_{p+3}(\chi_{p-1}^2(\alpha), \Delta^2)]}.$$

The efficiency expression allows one to determine the optimum level of significance for minimum guaranteed efficiency, E_0 , as the solution α^* of the equation $\min_{\Delta^2} ARRE(\alpha, \Delta^2) = E_0$ where E_0 is a preassigned efficiency number for the PTE. The efficiency based on the optimum level α^* may go up to the maximum in (3.61).

Table 3.8 may be used to determine the optimum level of significance for the PTE when the quadratic risk based criterion is used.

To explain how to use Table 3.8, for example, let $p = 12$ and assume that we wish to obtain a PTE with at least 96% efficiency, Table 3.8 gives $\alpha^* = 0.25$ at the intersection of $E_0 = 0.9631$ and $p = 12$. Hence the optimum level for the PTE in this case is $\alpha = 0.25$ with a maximum possible efficiency 2.2439. Now, consider $ARRE(\hat{\delta}_n^{JS}; \tilde{\delta}_n)$ as a function of Δ^2 . It attains its maximum at $\Delta^2 = 0$ with value $\frac{p}{3}$, then decreases from $\frac{p}{3}$ at $\Delta^2 = 0$ and then down to the unity at

$$\Delta_2^2 = \frac{(p-1)[2E(\chi_{p+3}^{-2}(\Delta^2)) - (p-3)E(\chi_{p+1}^{-4}(\Delta^2))]}{(p+1)E(\chi_{p+3}^{-4}(\Delta^2))}.$$

Figure 3.5: Asymptotic relative efficiency of PTE for different values of α , $p = 6$

thus if $0 \leq \Delta^2 \leq \Delta_2^2$, then $\hat{\delta}_n^{JS}$ performs better than $\tilde{\delta}_n$ and if $\Delta^2 \geq \Delta_2^2$, then $\hat{\delta}_n^{JS}$ performs as good as $\tilde{\delta}_n$. Table 3.9 gives the performance of $\hat{\delta}_n$, $\hat{\delta}_n^{JS}$, $\hat{\delta}_n^{PT}$ (for each selected level of significance) for different number of studies. Graph of the asymptotic relative risk efficiency of JSE against Δ^2 is plotted along with the graph of asymptotic relative efficiency of PTE and RE in Figure 3.6. Figure illustrates that ARRE of RE is higher than that of both PTE and JSE at $\Delta^2 = 0$ (under H_0) and in a small interval around it. outside this interval, PTE, JSE are better. When Δ^2 is near the origin, JSE outperforms PTE for most α values. Also, at $\Delta^2 = 0$, the asymptotic relative efficiency of JSE is higher than that of PTE except for small α . Further, JSE is

Figure 3.6: Asymptotic relative efficiency of RE, JSE and PTE for different values of $\alpha, p = 6$

Table 3.8: Maximum and Minimum Asymptotic Relative Efficiency of PTE for selected number of studies.

$\alpha \backslash p$		6	8	10	12
0.05	$E_{\max}$	3.5760	4.3485	5.0184	5.6082
	E_0	0.7591	0.8130	0.8521	0.8818
	Δ_0^2	9.755	11.85	13.70	15.00
0.15	$E_{\max}$	2.2965	2.6198	2.8783	3.0918
	E_0	0.8612	0.8971	0.9220	0.9398
	Δ_0^2	8.25	10.15	11.95	13.75
0.25	$E_{\max}$	1.7941	1.9817	2.1269	2.2439
	E_0	0.9089	0.9341	0.9512	0.9631
	Δ_0^2	7.45	9.3	11.10	12.80
0.35	$E_{\max}$	1.5182	1.6405	1.7335	1.8075
	E_0	0.9386	0.9562	0.9681	0.9763
	Δ_0^2	6.85	8.7	10.40	12.30
0.45	$E_{\max}$	1.3436	1.4272	1.4902	1.5399
	E_0	0.9589	0.9710	0.9791	0.9847
	Δ_0^2	6.40	8.25	9.95	11.60

independent of α . Thus, considering the overall performance of JSE relative to PTE, JSE is preferable to PTE since it produces interpolated estimator which is free from the level of significance, α .

Let us now discuss the utility of these estimators. The unrestricted estimator $\tilde{\delta}_n = (\tilde{\delta}_1, \dots, \tilde{\delta}_p)'$ is the usual estimator of p independent experiments while $\hat{\delta}_n$ is the restricted estimator and good for the use as a combined summary of the p experiments when H_0 is tenable. If $\mathcal{L}_n$ accepts the null hypothesis as truth, $\hat{\delta}_n$ is a legitimate statistics to summarize the results of the p experiments. However, the preliminary test estimator, $\hat{\delta}_n^{PT} = \tilde{\delta}_n - (\tilde{\delta}_n - \hat{\delta}_n)I(\mathcal{L}_n \leq \chi_{q,\alpha}^2)$ has some significant role. If $H_0 : \delta = \delta_0 \mathbf{1}_p$ is rejected, then $\hat{\delta}_n$ is not the statistically valid estimator for δ and does not summarize the p -findings. Then, one has to look back what is causing the rejection and maybe try to find a subset of δ that is homogeneous to summarize the p -studies for meta analysis. We have to mention that PTE is nothing but either the unrestricted estimator, $\tilde{\delta}_n$ or the restricted estimator, $\hat{\delta}_n$ based on accept—reject rule

Table 3.9: Maximum and Minimum Asymptotic Relative Efficiency of RE, JSE and PTE for $\alpha=0.05(0.10)0.45$ for selected number of studies.

$\alpha \backslash p$		6	8	10	12
	$E_2^{\max}$	6.000	8.000	10.000	12.000
	$E_2^{\min}$	0.2857	0.3810	0.4762	0.3871
	$E_4^{\max}$	2.000	2.667	3.333	4.000
	$E_4^{\min}$	0.99970	0.9988	0.9997	0.9998
0.05	$E_{2\Delta_0^2}$	0.555	0.6250	0.6825	0.7523
	E_0	0.7591	0.8130	0.8521	0.8818
	$E_{4\Delta_0^2}$	0.98670	0.9804	0.9705	0.9720
	Δ_0^2	9.755	11.85	13.70	15.00
0.15	$E_{2\Delta_0^2}$	0.6521	0.7207	0.7752	0.8163
	E_0	0.8612	0.8971	0.9220	0.9398
	$E_{4\Delta_0^2}$	1.018	0.9940	0.9791	0.9714
	Δ_0^2	8.25	10.15	11.95	13.75
0.25	$E_{2\Delta_0^2}$	0.7143	0.7805	0.8298	0.8450
	E_0	0.9089	0.9341	0.9512	0.9631
	$E_{4\Delta_0^2}$	1.044	1.013	0.9865	0.9725
	Δ_0^2	7.45	9.3	11.10	12.80
0.35	$E_{2\Delta_0^2}$	0.7692	0.8290	0.8810	0.9056
	E_0	0.9386	0.9562	0.9681	0.9763
	$E_{4\Delta_0^2}$	1.068	1.032	0.9989	0.9775
	Δ_0^2	6.85	8.7	10.40	12.30
0.45	$E_{2\Delta_0^2}$	0.8163	0.8669	0.9174	0.9561
	E_0	0.9589	0.9710	0.9791	0.9847
	$E_{4\Delta_0^2}$	1.090	1.050	1.009	0.9847
	Δ_0^2	6.40	8.25	9.95	11.60

which is the current rule in the classical methods of meta analysis. However, the role of James-Stein type estimator, $\hat{\delta}_n^{JS}$ is more intense than the PTE, the quasi-empirical Bayes techniques involve learning from the experience of others. Each experiment learning from the other $p - 1$ experiments, all of estimated effect sizes are pooled towards the restricted estimator, $\hat{\delta}_n$. The same holds for the positive-rule Stein-type estimator, $\hat{\delta}_n^{s+}$. The gain in using the quasi-empirical Bayes method of estimation is we have obtained all the estimators that currently widely used in meta analysis as well as we obtain James-Stein type estimators. As we have seen in the previous sections that the overall performance of James-Stein type estimators is preferable to the other estimators.

3.5 Variance-stabilizing Transformation

We have discussed in section 2.3, the relationship between the distribution of the effect size estimate and that of the sample correlation. Kraemer (1983) established the relationship between sample correlation and the effect size estimate $\tilde{\delta}_i$ for the balanced study (i.e $n_i^E = n_i^C = n_i$) as

$$r_i = \frac{\tilde{\delta}_i}{\sqrt{\tilde{\delta}_i^2 + 4}},$$

$$\text{and } \rho_i = \frac{\delta_i}{\sqrt{\delta_i^2 + 4}}.$$

Using Fisher's transformation define

$$\theta_i = \frac{1}{2} \log \frac{1 + \rho_i}{1 - \rho_i}.$$

Then the Maximum likelihood estimator of θ_i is given by

$$\tilde{\theta}_i = \frac{1}{2} \log \frac{1 + r_i}{1 - r_i}.$$

The large sample distribution of $\tilde{\theta}_i$ is

$$\sqrt{2n_i - 3} (\tilde{\theta}_i - \theta_i) \sim N(0, 1). \quad (3.62)$$

Because $\delta_1 = \dots = \delta_p$ is equivalent to $\theta_1 = \dots = \theta_p$, a test for hypothesis

$$H_0 : \delta_1 = \dots = \delta_p. \quad (3.63)$$

is equivalent to a test of homogeneity of θ 's, that is

$$H_0 : \theta_1 = \dots = \theta_p = \theta_0 \text{ unknown}. \quad (3.64)$$

Under model (3.64), the restricted estimator is given by

$$\hat{\boldsymbol{\theta}}_n = (\hat{\theta}_0, \dots, \hat{\theta}_0)' = \hat{\theta}_0 \mathbf{1}_p = \frac{\mathbf{1}_p \mathbf{1}_p' \boldsymbol{\Lambda}_n^{-1}}{\mathbf{1}_p' \boldsymbol{\Lambda}_n^{-1} \mathbf{1}_p} \tilde{\boldsymbol{\theta}}_n,$$

where

$$\begin{aligned} \hat{\theta}_0 &= \frac{\sum_{i=1}^p (2n_i - 3) \tilde{\theta}_i}{n - 3p}, \\ \tilde{\boldsymbol{\theta}}_n &= (\tilde{\theta}_1, \dots, \tilde{\theta}_p)' \\ \boldsymbol{\Lambda}_n &= \text{Diag}\left(\frac{1}{2n_1 - 3}, \dots, \frac{1}{2n_p - 3}\right) \end{aligned}$$

To test the hypothesis in (3.64) use the test-statistics

$$\mathcal{L}_n = n(\tilde{\boldsymbol{\theta}}_n - \hat{\boldsymbol{\theta}}_n)'(n\boldsymbol{\Lambda}_n)^{-1}(\tilde{\boldsymbol{\theta}}_n - \hat{\boldsymbol{\theta}}_n)$$

If H_0 is true, and $\frac{n_i}{n} \rightarrow \pi_i$ as $n \rightarrow \infty$, π_i remain fixed, then the asymptotic distribution of $\mathcal{L}_n$ is χ_{p-1}^2 .

3.5.1 Proposed Estimators

As before, one can define the following estimators

$$\hat{\boldsymbol{\theta}}_n^{PT} = \tilde{\boldsymbol{\theta}}_n - (\tilde{\boldsymbol{\theta}}_n - \hat{\boldsymbol{\theta}}_n)I(\mathcal{L}_n \leq \chi_{p-1}^2(\alpha)).$$

$$\hat{\boldsymbol{\theta}}_n^{JS} = \tilde{\boldsymbol{\theta}}_n - (p-3)(\tilde{\boldsymbol{\theta}}_n - \hat{\boldsymbol{\theta}}_n)\mathcal{L}_n^{-1}.$$

$$\hat{\boldsymbol{\theta}}_n^{s+} = \hat{\boldsymbol{\theta}}_n + [1 - (p-3)\mathcal{L}_n^{-1}]I(\mathcal{L}_n > (p-3))(\tilde{\boldsymbol{\theta}}_n - \hat{\boldsymbol{\theta}}_n).$$

3.5.2 Asymptotic Distributional Properties of the Estimators

We consider the sequence of local alternatives for the derivation of the asymptotic distribution of the estimators

$$\mathcal{K}_{(n)} : \boldsymbol{\theta}_{(n)} = \theta_0 \mathbf{1}_p + n^{-1/2} \boldsymbol{\xi}, \quad \boldsymbol{\xi} = (\xi_1, \dots, \xi_p)' \neq \mathbf{0}. \quad (3.65)$$

Note that for $\boldsymbol{\xi} = \mathbf{0}$, $\boldsymbol{\theta}_{(n)} = \theta_0 \mathbf{1}_p$, that is (3.64) is a particular case of (3.65).

We arrive at the next theorem on the asymptotic distributions of the various statistics

Theorem 3.5.1. *Under $\{\mathcal{K}_{(n)}\}$ and the assumed regularity conditions, as $n \rightarrow \infty$,*

the following hold:

$$\begin{aligned}
(i) \quad & \sqrt{n}(\tilde{\boldsymbol{\theta}}_n - \boldsymbol{\theta}_{(n)}) \xrightarrow{\mathcal{D}} \mathbf{Z}_0 \sim N_p(\mathbf{0}, \boldsymbol{\Lambda}_0), \text{ and } \boldsymbol{\Lambda}_0 = \text{Diag}\left(\frac{1}{2\pi_1}, \dots, \frac{1}{2\pi_p}\right). \\
(ii) \quad & \sqrt{n}(\tilde{\boldsymbol{\theta}}_n - \hat{\boldsymbol{\theta}}_n) \xrightarrow{\mathcal{D}} \mathbf{Z}_1, \mathbf{Z}_1 \sim N_p(\mathbf{J}\boldsymbol{\xi}, \boldsymbol{\Lambda}_0\mathbf{J}'), \mathbf{J} = \mathbf{I}_p - \frac{\mathbf{1}_p\mathbf{1}_p' \boldsymbol{\Lambda}_0^{-1}}{\mathbf{1}_p' \boldsymbol{\Lambda}_0^{-1} \mathbf{1}_p}. \\
(iii) \quad & \sqrt{n}(\hat{\boldsymbol{\theta}}_n - \boldsymbol{\theta}_{(n)}) \xrightarrow{\mathcal{D}} \mathbf{Z}_2 - \mathbf{J}\boldsymbol{\xi}, \mathbf{Z}_2 \sim N_p(\mathbf{0}, \mathbf{B}), \mathbf{B} = \frac{\mathbf{1}_p\mathbf{1}_p'}{\mathbf{1}_p' \boldsymbol{\Lambda}_0^{-1} \mathbf{1}_p}. \\
(iv) \quad & \begin{pmatrix} \sqrt{n}(\tilde{\boldsymbol{\theta}}_n - \boldsymbol{\theta}_{(n)}) \\ \sqrt{n}(\tilde{\boldsymbol{\theta}}_n - \hat{\boldsymbol{\theta}}_n) \end{pmatrix} \xrightarrow{\mathcal{D}} \begin{pmatrix} \mathbf{Z}_0 \\ \mathbf{Z}_1 \end{pmatrix} \sim N_{2p} \left(\begin{bmatrix} \mathbf{0} \\ \mathbf{J}\boldsymbol{\xi} \end{bmatrix}, \begin{bmatrix} \boldsymbol{\Lambda}_0 & \boldsymbol{\Lambda}_0\mathbf{J}' \\ \boldsymbol{\Lambda}_0\mathbf{J}' & \boldsymbol{\Lambda}_0\mathbf{J}' \end{bmatrix} \right). \\
(v) \quad & \begin{pmatrix} \sqrt{n}(\hat{\boldsymbol{\theta}}_n - \boldsymbol{\theta}_{(n)}) \\ \sqrt{n}(\tilde{\boldsymbol{\theta}}_n - \hat{\boldsymbol{\theta}}_n) \end{pmatrix} \xrightarrow{\mathcal{D}} \begin{pmatrix} \mathbf{Z}_2 - \mathbf{J}\boldsymbol{\xi} \\ \mathbf{Z}_1 \end{pmatrix} \sim N_{2p} \left(\begin{bmatrix} -\mathbf{J}\boldsymbol{\xi} \\ \mathbf{J}\boldsymbol{\xi} \end{bmatrix}, \begin{bmatrix} \mathbf{B} & \mathbf{0} \\ \mathbf{0} & \boldsymbol{\Lambda}_0\mathbf{J}' \end{bmatrix} \right). \\
(vi) \quad & \lim_{n \rightarrow \infty} P\{\mathcal{L}_n \leq x | \mathcal{K}_{(n)}\} = H_{p-1}(x, \Delta^2) \text{ where } \Delta^2 = \boldsymbol{\eta}' \boldsymbol{\Lambda}_0^{-1} \boldsymbol{\eta}, \boldsymbol{\eta} = \mathbf{J}\boldsymbol{\xi}.
\end{aligned}$$

where $H_{p-1}(\cdot, \Delta^2)$ is the cdf of noncentral chi-square distribution with noncentrality parameter $\frac{1}{2}\Delta^2$ and $p-1$ D.F.

Based on Theorem 3.5.1 we obtain the following results.

$$\sqrt{n}(\hat{\boldsymbol{\theta}}_n^{PT} - \boldsymbol{\theta}_{(n)}) \xrightarrow{\mathcal{D}} \mathbf{Z}_0 - \mathbf{Z}_1 I[\mathbf{Z}_1' \boldsymbol{\Lambda}_0^{-1} \mathbf{Z}_1 < \chi_{p-1, \alpha}^2] \quad (3.66)$$

$$\sqrt{n}(\hat{\boldsymbol{\theta}}_n^{JS} - \boldsymbol{\delta}_{(n)}) \xrightarrow{\mathcal{D}} \mathbf{Z}_0 - (p-3)\mathbf{Z}_1[\mathbf{Z}_1' \boldsymbol{\Lambda}_0^{-1} \mathbf{Z}_1]^{-1} \quad (3.67)$$

$$\begin{aligned}
\sqrt{n}(\hat{\boldsymbol{\theta}}_n^{s+} - \boldsymbol{\theta}_{(n)}) & \xrightarrow{\mathcal{D}} \mathbf{Z}_0 - (p-3)\mathbf{Z}_1[\mathbf{Z}_1' \boldsymbol{\Lambda}_0^{-1} \mathbf{Z}_1]^{-1} \\
& - \mathbf{Z}_1\{1 - (p-3)[\mathbf{Z}_1' \boldsymbol{\Lambda}_0^{-1} \mathbf{Z}_1]^{-1}\} \\
& \times I[\mathbf{Z}_1' \boldsymbol{\Lambda}_0^{-1} \mathbf{Z}_1 < p-3].
\end{aligned} \quad (3.68)$$

Define the asymptotic distributional bias (ADB), the asymptotic distributional MSE (ADMSE) and the asymptotic distributional risk (ADQR) of the estimator $\boldsymbol{\theta}_n^*$,

where $\boldsymbol{\theta}_n^*$ stands for $\tilde{\boldsymbol{\theta}}_n$, $\hat{\boldsymbol{\theta}}_n$, $\hat{\boldsymbol{\theta}}_n^{PT}$, $\hat{\boldsymbol{\theta}}_n^{JS}$, and $\hat{\boldsymbol{\theta}}_n^{s+}$ as in (3.38), (3.39) and (3.51). Based on Theorem 3.5.1 and the above results in (3.66) - (3.68) and using Theorem A.2.6 and A.2.7, we obtain the ADB, ADMSE and ADQR expressions respectively. Since the derivation of these expressions are very similar to the cases treated in Theorem 3.4.4 and Theorem 3.4.5, we omit these details of the derivation. Expressions of the ADB, ADMSE, and ADQR for the five estimators are as follows:

Unrestricted estimator (UE)

$$\begin{aligned} b_1(\tilde{\boldsymbol{\theta}}_n) &= \mathbf{0}. \\ M_1(\tilde{\boldsymbol{\theta}}_n) &= \boldsymbol{\Lambda}_0. \\ R_1(\tilde{\boldsymbol{\theta}}_n, \mathbf{W}) &= \text{tr}(\mathbf{W}\boldsymbol{\Lambda}_0). \end{aligned}$$

Restricted estimator (RE)

$$\begin{aligned} b_2(\hat{\boldsymbol{\theta}}_n) &= -\boldsymbol{\eta}. \\ M_2(\hat{\boldsymbol{\theta}}_n) &= \mathbf{B} + \boldsymbol{\eta}\boldsymbol{\eta}'. \\ R_2(\hat{\boldsymbol{\theta}}_n, \mathbf{W}) &= \text{tr}(\mathbf{W}\mathbf{B}) + \boldsymbol{\eta}'\mathbf{W}\boldsymbol{\eta}. \end{aligned}$$

Preliminary test estimator (PTE)

$$\begin{aligned} b_3(\hat{\boldsymbol{\theta}}_n^{PT}) &= -\boldsymbol{\eta}H_{p+1}(\chi_{p-1}^2(\alpha), \Delta^2). \\ M_3(\hat{\boldsymbol{\theta}}_n^{PT}) &= \boldsymbol{\Lambda}_0 - \boldsymbol{\Lambda}_0\mathbf{J}'H_{p+1}(\chi_{p-1}^2(\alpha), \Delta^2) \\ &\quad + \boldsymbol{\eta}\boldsymbol{\eta}'\{2H_{p+1}(\chi_{p-1}^2(\alpha), \Delta^2) - H_{p+3}(\chi_{p-1}^2(\alpha), \Delta^2)\}. \\ R_3(\hat{\boldsymbol{\theta}}_n^{PT}, \mathbf{W}) &= \text{tr}(\mathbf{W}\boldsymbol{\Lambda}_0) - \text{tr}(\mathbf{W}\boldsymbol{\Lambda}_0\mathbf{J}')H_{p+1}(\chi_{p-1}^2(\alpha), \Delta^2) \\ &\quad + \boldsymbol{\eta}'\mathbf{W}\boldsymbol{\eta}\{2H_{p+1}(\chi_{p-1}^2(\alpha), \Delta^2) - H_{p+3}(\chi_{p-1}^2(\alpha), \Delta^2)\}. \end{aligned}$$

James-Stein estimator (JSE)

$$\begin{aligned}
b_4(\hat{\boldsymbol{\theta}}_n^{JS}) &= -(p-3)\boldsymbol{\eta}E(\chi_{p+1}^{-2}(\Delta^2)). \\
M_4(\hat{\boldsymbol{\theta}}_n^{JS}) &= \boldsymbol{\Lambda}_0 - (p-3)\boldsymbol{\Lambda}_0\mathbf{J}'\{2E[\chi_{p+1}^{-2}(\Delta^2)] - (p-3)E[\chi_{p+1}^{-4}(\Delta^2)]\} \\
&\quad + (p-3)(p+1)\boldsymbol{\eta}\boldsymbol{\eta}'E[\chi_{p+3}^{-4}(\Delta^2)]. \\
R_4(\hat{\boldsymbol{\theta}}_n^{JS}, \mathbf{W}) &= \text{tr}(\mathbf{W}\boldsymbol{\Lambda}_0) - (p-3)\text{tr}(\mathbf{W}\boldsymbol{\Lambda}_0\mathbf{J}')\{2E[\chi_{p+1}^{-2}(\Delta^2)] \\
&\quad - (p-3)E[\chi_{p+1}^{-4}(\Delta^2)]\} + (p-3)(p+1)\boldsymbol{\eta}'\mathbf{W}\boldsymbol{\eta}E[\chi_{p+3}^{-4}(\Delta^2)].
\end{aligned}$$

Positive rule Stein estimator (PRSE)

$$\begin{aligned}
b_5(\hat{\boldsymbol{\theta}}_n^{s+}) &= (p-3)\boldsymbol{\eta}E[\chi_{p+1}^{-2}(\Delta^2)] - \boldsymbol{\eta}E[1 - (p-3)\chi_{p+1}^{-2}(\Delta^2)] \\
&\quad \times I[\chi_{p+1}^2(\Delta^2) \leq p-3]. \\
M_5(\hat{\boldsymbol{\theta}}_n^{s+}) &= M_4(\hat{\boldsymbol{\theta}}_n^{JS}) - (p-3)\boldsymbol{\Lambda}_0\mathbf{J}' \\
&\quad \times \{E[(1 - (p-3)\chi_{p+1}^{-2}(\Delta^2))^2 I(\chi_{p+1}^2(\Delta^2) \leq p-3)]\} \\
&\quad + \boldsymbol{\eta}\boldsymbol{\eta}'\{2E[(1 - (p-3)\chi_{p+1}^{-2}(\Delta^2))I(\chi_{p+1}^2(\Delta^2) \leq p-3)] \\
&\quad - E[(1 - (p-3)\chi_{p+3}^{-2}(\Delta^2))^2 I(\chi_{p+3}^2(\Delta^2) \leq p-3)]\}. \\
R_5(\hat{\boldsymbol{\theta}}_n^{s+}, \mathbf{W}) &= R_4(\hat{\boldsymbol{\theta}}_n^{JS}, \mathbf{W}) - (p-3)\text{tr}(\mathbf{W}\boldsymbol{\Lambda}_0\mathbf{J}') \\
&\quad \times E[(1 - (p-3)\chi_{p+1}^{-2}(\Delta^2))^2 I(\chi_{p+1}^2(\Delta^2) \leq p-3)] \\
&\quad + \boldsymbol{\eta}'\mathbf{W}\boldsymbol{\eta}\{2E[(1 - (p-3)\chi_{p+1}^{-2}(\Delta^2))I(\chi_{p+1}^2(\Delta^2) \leq p-3)] \\
&\quad - E[(1 - (p-3)\chi_{p+3}^{-2}(\Delta^2))^2 I(\chi_{p+3}^2(\Delta^2) \leq p-3)]\}.
\end{aligned}$$

The performance of the above estimators are exactly the same as the performance of the estimators in section 3.4, by observing the similarity of the risk expressions and the analysis that follows.

3.6 Meta Analysis of Polls

In this section, we consider meta analysis of various poll results conducted by several competing agencies, and meaningfully combine such results. The initial motivation is taken from Das and Sinha (2006) which is presented in Chapter 15 of the book by Hartung, Knapp and Sinha (2008).

Assume that p independent competing agencies conduct polls to study the effectiveness of k candidates of winning an election. The data is presented as:

Agency	Candidate				Total
	1	2	...	k	
1	x_{11}	x_{12}	...	x_{1k}	n_1
2	x_{21}	x_{22}	...	x_{2k}	n_2
...	...			...	...
p	x_{p1}	x_{p2}	...	x_{pk}	n_p

The frequencies x_{ij} are the number of votes received by candidate j in the i th agency polling of $n_i = \sum_{j=1}^k x_{ij}$, $i = 1, \dots, p$ and $n = n_1 + \dots + n_p$. The vector $(x_{i1}, \dots, x_{ik})$ follows the multinomial distribution

$$\frac{n_i!}{x_{i1}! \dots x_{ik}!} \prod_{j=1}^k (P_{ij})^{x_{ij}}, \quad \sum_{j=1}^k P_{ij} = 1 \text{ for all } i = 1, \dots, p.$$

Thus, we have

$$E(x_{ij}) = n_i P_{ij} \quad \text{and} \quad \text{cov}(x_{ij}, x_{ij'}) = -n_i P_{ij} P_{ij'}, \quad j \neq j'.$$

Although there are k candidates, generally one is interested in the two dominating candidates, say candidate 1 and 2. Effect size based on proportion is defined by the difference between two proportions P_{i1} and P_{i2} as $\theta_i = P_{i1} - P_{i2}$. The unbiased estimate of θ_i is $\tilde{\theta}_i = p_{i1} - p_{i2}$ where $p_{i1} = \frac{x_{i1}}{n_i}$ and $p_{i2} = \frac{x_{i2}}{n_i}$, $i = 1, \dots, p$. Obviously,

$\tilde{\theta}_1, \dots, \tilde{\theta}_p$ are independent but not identically distributed random variables. Clearly

$$E(\tilde{\theta}_i) = P_{i1} - P_{i2} \text{ and } \text{var}(\tilde{\theta}_i) = \frac{1}{n_i}(P_{i1} + P_{i2} - \theta_i^2), \quad i = 1, \dots, p.$$

Let $\boldsymbol{\theta} = (\theta_1, \dots, \theta_p)'$ be the vector of the population proportion difference between P_{i1} and $P_{i2}, i = 1, \dots, p$ and correspondingly, let $\tilde{\boldsymbol{\theta}}_n = (\tilde{\theta}_1, \dots, \tilde{\theta}_p)'$ be the vector of unbiased estimators of $\boldsymbol{\theta}$. We call $\tilde{\boldsymbol{\theta}}_n$ the unrestricted estimator of $\boldsymbol{\theta}$. If sample sizes are large, then we have the following results

Theorem 3.6.1. *The asymptotic distribution of $\tilde{\boldsymbol{\theta}}_n$ is given by*

$$\sqrt{n}(\tilde{\boldsymbol{\theta}}_n - \boldsymbol{\theta}) \sim N_p(0, \boldsymbol{\Sigma}_0)$$

$\boldsymbol{\Sigma}_0 = \lim_{n \rightarrow \infty} n\boldsymbol{\Sigma}_n$ where $\frac{n_i}{n} = \pi_i$, and π_i remains fixed as $\min_{1 \leq i \leq p} n_i \rightarrow \infty, n \rightarrow \infty$ and

$$\boldsymbol{\Sigma}_n = \text{Diag}\left(\frac{1}{n_1}(P_{11} + P_{12} - \theta_1^2), \dots, \frac{1}{n_p}(P_{p1} + P_{p2} - \theta_p^2)\right), \quad (3.69)$$

that is

$$\boldsymbol{\Sigma}_0 = \text{Diag}\left(\pi_1^{-1}(P_{11} + P_{12} - \theta_1^2), \dots, \pi_p^{-1}(P_{p1} + P_{p2} - \theta_p^2)\right). \quad (3.70)$$

The estimated variances are then obtained by replacing θ_i in (3.69) by the unbiased and consistent estimators of θ_i , namely, $\tilde{\theta}_i$, obtaining the estimator of the $\boldsymbol{\Sigma}_n$ as

$$\hat{\boldsymbol{\Sigma}}_n = \text{Diag}\left(\frac{1}{n_1}(p_{11} + p_{12} - \tilde{\theta}_1^2), \dots, \frac{1}{n_p}(p_{p1} + p_{p2} - \tilde{\theta}_p^2)\right), \quad (3.71)$$

Accordingly, the asymptotic $100(1 - \beta)\%$ confidence interval for each of θ_i is given by

$$\tilde{\theta}_i \pm z_{\beta/2} \sqrt{\frac{1}{n_i}(p_{i1} + p_{i2} - \tilde{\theta}_i^2)}.$$

First step in our approach is to formulate a plausible hypothesis to summarize the results of p polls. Assume that all agency results are similar, that is, $P_{1j} = \dots = P_{pj} = P_j$, $\theta_1 = \dots = \theta_p = \theta_0$ (say), that is to have the null hypothesis H_0 as

$$H_0 : \theta_1 = \dots = \theta_p = \theta_0 \text{ unknown} \quad (3.72)$$

may hold. If H_0 is true, then the restricted estimator (RE) of $\boldsymbol{\theta}$, is given by

$$\hat{\theta}_{0n} = \frac{\mathbf{1}_p' \hat{\Sigma}_n^{-1} \tilde{\boldsymbol{\theta}}_n}{\mathbf{1}_p' \hat{\Sigma}_n^{-1} \mathbf{1}_p}, \quad (3.73)$$

with the estimated asymptotic variance as

$$\hat{\sigma}^2(\hat{\theta}_{0n}) = (\mathbf{1}_p' \hat{\Sigma}_n^{-1} \mathbf{1}_p)^{-1}$$

Accordingly, the asymptotic $100(1 - \beta)\%$ confidence interval for θ_0 is given by

$$\hat{\theta}_{0n} \pm z_{\beta/2} \hat{\sigma}(\hat{\theta}_{0n}). \quad (3.74)$$

In general, H_0 is vague. We remove this vagueness based on a statistical test defined by

$$\mathcal{L}_n = n(\tilde{\boldsymbol{\theta}}_n - \hat{\theta}_{0n} \mathbf{1}_p)' (n \hat{\Sigma}_n)^{-1} (\tilde{\boldsymbol{\theta}}_n - \hat{\theta}_{0n} \mathbf{1}_p). \quad (3.75)$$

Theorem 3.6.2. *If H_0 in (3.72) is true, and $\frac{n_i}{n} \rightarrow \pi_i$, π_i remains fixed as $\min_{1 \leq i \leq p} n_i \rightarrow \infty$ which implies $n \rightarrow \infty$, then the asymptotic distribution of $\mathcal{L}_n$ is χ_{p-1}^2 (chi-square distribution with $p - 1$ D.F.)*

We recall that we started with two estimators of $\boldsymbol{\theta}$, namely, the unrestricted estimator, $\tilde{\boldsymbol{\theta}}_n$ and the restricted estimator, $\hat{\theta}_{0n} \mathbf{1}_p$ when H_0 holds.

Meta analysis is complete if H_0 is accepted with reasonable level of significance. Then, one reports the estimator $\hat{\theta}_{0n}$ in (3.73) with its $100(1 - \beta)\%$ confidence interval

in (3.74). In the next section, we discuss the new aspects of meta analysis.

3.6.1 Quasi-Empirical Bayes Methods of Meta Analysis of Polls

Based on what we have seen in section 3.4.1, we may define the following estimators

$$\begin{aligned}\hat{\boldsymbol{\theta}}_n^{PT} &= \tilde{\boldsymbol{\theta}}_n - (\tilde{\boldsymbol{\theta}}_n - \hat{\boldsymbol{\theta}}_n)I(\mathcal{L}_n \leq \chi_{p-1,\alpha}^2) \\ \hat{\boldsymbol{\theta}}_n^{JS} &= \hat{\boldsymbol{\theta}}_n + [1 - (p-3)\mathcal{L}_n^{-1}](\tilde{\boldsymbol{\theta}}_n - \hat{\boldsymbol{\theta}}_n). \\ \hat{\boldsymbol{\theta}}_n^{s+} &= \hat{\boldsymbol{\theta}}_n + [1 - (p-3)\mathcal{L}_n^{-1}]I(\mathcal{L}_n > (p-3))(\tilde{\boldsymbol{\theta}}_n - \hat{\boldsymbol{\theta}}_n).\end{aligned}$$

Thus, we have five possible estimators of $\boldsymbol{\theta}$. We provide their asymptotic properties in the next section.

3.6.2 Asymptotic Distributional Properties of the Estimators

We consider the sequence of local alternatives for the derivation of the asymptotic distribution of the estimators

$$\mathcal{K}_{(n)} : \boldsymbol{\theta}_{(n)} = \theta_0 \mathbf{1}_p + n^{-1/2} \boldsymbol{\xi}, \quad \boldsymbol{\xi} = (\xi_1, \dots, \xi_p)' \neq \mathbf{0}. \quad (3.76)$$

Note that for $\boldsymbol{\xi} = \mathbf{0}$, $\boldsymbol{\theta}_{(n)} = \theta_0 \mathbf{1}_p$, that is the null hypothesis is (3.72) a particular case of (3.76).

We arrive at the next theorem on the asymptotic distributions of the various statistics

Theorem 3.6.3. *Under $\{\mathcal{K}_{(n)}\}$ and $\frac{n_1}{n} \rightarrow \pi_i$, π_i remains fixed as $\min_{1 \leq i \leq p} n_i \rightarrow \infty$*

which implies $n \rightarrow \infty$, the following hold:

$$\begin{aligned}
(i) \quad & \sqrt{n}(\tilde{\boldsymbol{\theta}}_n - \boldsymbol{\theta}_{(n)}) \xrightarrow{\mathcal{D}} \mathbf{Z}_0 \sim N_p(\mathbf{0}, \boldsymbol{\Sigma}_0) \text{ and } \boldsymbol{\Sigma}_0 \text{ given in (3.70).} \\
(ii) \quad & \sqrt{n}(\tilde{\boldsymbol{\theta}}_n - \hat{\boldsymbol{\theta}}_n) \xrightarrow{\mathcal{D}} \mathbf{Z}_1, \mathbf{Z}_1 \sim N_p(\mathbf{J}\boldsymbol{\xi}, \boldsymbol{\Sigma}_0\mathbf{J}'), \mathbf{J} = \mathbf{I}_p - \frac{\mathbf{1}_p\mathbf{1}_p' \boldsymbol{\Sigma}_0^{-1}}{\mathbf{1}_p' \boldsymbol{\Sigma}_0^{-1} \mathbf{1}_p}. \\
(iii) \quad & \sqrt{n}(\hat{\boldsymbol{\theta}}_n - \boldsymbol{\theta}_{(n)}) \xrightarrow{\mathcal{D}} \mathbf{Z}_2 - \mathbf{J}\boldsymbol{\xi}, \mathbf{Z}_2 \sim N_p(\mathbf{0}, \mathbf{B}), \mathbf{B} = \frac{\mathbf{1}_p\mathbf{1}_p'}{\mathbf{1}_p' \boldsymbol{\Sigma}_0^{-1} \mathbf{1}_p}. \\
(iv) \quad & \begin{pmatrix} \sqrt{n}(\tilde{\boldsymbol{\theta}}_n - \boldsymbol{\theta}_{(n)}) \\ \sqrt{n}(\tilde{\boldsymbol{\theta}}_n - \hat{\boldsymbol{\theta}}_n) \end{pmatrix} \xrightarrow{\mathcal{D}} \begin{pmatrix} \mathbf{Z}_0 \\ \mathbf{Z}_1 \end{pmatrix} \sim N_{2p} \left(\begin{bmatrix} \mathbf{0} \\ \mathbf{J}\boldsymbol{\xi} \end{bmatrix}, \begin{bmatrix} \boldsymbol{\Sigma}_0 & \boldsymbol{\Sigma}_0\mathbf{J}' \\ \boldsymbol{\Sigma}_0\mathbf{J}' & \boldsymbol{\Sigma}_0\mathbf{J}' \end{bmatrix} \right). \\
(v) \quad & \begin{pmatrix} \sqrt{n}(\hat{\boldsymbol{\theta}}_n - \boldsymbol{\theta}_{(n)}) \\ \sqrt{n}(\tilde{\boldsymbol{\theta}}_n - \hat{\boldsymbol{\theta}}_n) \end{pmatrix} \xrightarrow{\mathcal{D}} \begin{pmatrix} \mathbf{Z}_2 - \mathbf{J}\boldsymbol{\xi} \\ \mathbf{Z}_1 \end{pmatrix} \sim N_{2p} \left(\begin{bmatrix} -\mathbf{J}\boldsymbol{\xi} \\ \mathbf{J}\boldsymbol{\xi} \end{bmatrix}, \begin{bmatrix} \mathbf{B} & \mathbf{0} \\ \mathbf{0} & \boldsymbol{\Sigma}_0\mathbf{J}' \end{bmatrix} \right). \\
(vi) \quad & \lim_{n \rightarrow \infty} P\{\mathcal{L}_n \leq x | \mathcal{K}_{(n)}\} = H_{p-1}(x, \Delta^2) \text{ where } \Delta^2 = \boldsymbol{\eta}' \boldsymbol{\Sigma}_0^{-1} \boldsymbol{\eta}, \boldsymbol{\eta} = \mathbf{J}\boldsymbol{\xi}.
\end{aligned}$$

where $H_{p-1}(\cdot, \Delta^2)$ is the cdf of noncentral chi-square distribution with noncentral-ity parameter $\frac{1}{2}\Delta^2$ and $p-1$ D.F.

Based on Theorem 3.6.3 we obtain the following results.

$$\sqrt{n}(\hat{\boldsymbol{\theta}}_n^{PT} - \boldsymbol{\theta}_{(n)}) \xrightarrow{\mathcal{D}} \mathbf{Z}_0 - \mathbf{Z}_1 I[\mathbf{Z}_1' \boldsymbol{\Sigma}_0^{-1} \mathbf{Z}_1 < \chi_{p-1, \alpha}^2] \quad (3.77)$$

$$\sqrt{n}(\hat{\boldsymbol{\theta}}_n^{JS} - \boldsymbol{\theta}_{(n)}) \xrightarrow{\mathcal{D}} \mathbf{Z}_0 - (p-3)\mathbf{Z}_1[\mathbf{Z}_1' \boldsymbol{\Sigma}_0^{-1} \mathbf{Z}_1]^{-1} \quad (3.78)$$

$$\begin{aligned}
\sqrt{n}(\hat{\boldsymbol{\theta}}_n^{s+} - \boldsymbol{\theta}_{(n)}) & \xrightarrow{\mathcal{D}} \mathbf{Z}_0 - (p-3)\mathbf{Z}_1[\mathbf{Z}_1' \boldsymbol{\Sigma}_0^{-1} \mathbf{Z}_1]^{-1} \\
& - \mathbf{Z}_1 \{1 - (p-3)[\mathbf{Z}_1' \boldsymbol{\Sigma}_0^{-1} \mathbf{Z}_1]^{-1}\} \\
& \times I[\mathbf{Z}_1' \boldsymbol{\Sigma}_0^{-1} \mathbf{Z}_1 < p-3].
\end{aligned} \quad (3.79)$$

Define the ADB, ADMSE and the ADQR of the estimator $\boldsymbol{\theta}_n^*$, where $\boldsymbol{\theta}_n^*$ stands for $\tilde{\boldsymbol{\theta}}_n, \hat{\boldsymbol{\theta}}_n, \hat{\boldsymbol{\theta}}_n^{PT}, \hat{\boldsymbol{\theta}}_n^{JS}$, and $\hat{\boldsymbol{\theta}}_n^{s+}$ as in (3.38), (3.39) and (3.51). Based on Theorem 3.6.3 and the above results in (3.77) - (3.79) and using Theorem A.2.6 and A.2.7, we obtain the

ADB, ADMSE and ADQR. Since the derivation of these expressions are very similar to the cases treated in Theorem 3.4.4 and Theorem 3.4.5, we omit these details of the derivation. Expressions of the ADB, ADMSE, and ADQR for the five estimators are as follows:

Unrestricted estimator (UE)

$$\begin{aligned} b_1(\tilde{\boldsymbol{\theta}}_n) &= \mathbf{0}. \\ M_1(\tilde{\boldsymbol{\theta}}_n) &= \boldsymbol{\Sigma}_0. \\ R_1(\tilde{\boldsymbol{\theta}}_n, \mathbf{W}) &= \text{tr}(\mathbf{W}\boldsymbol{\Sigma}_0). \end{aligned}$$

Restricted estimator (RE)

$$\begin{aligned} b_2(\hat{\boldsymbol{\theta}}_n) &= -\boldsymbol{\eta}. \\ M_2(\hat{\boldsymbol{\theta}}_n) &= \mathbf{B} + \boldsymbol{\eta}\boldsymbol{\eta}'. \\ R_2(\hat{\boldsymbol{\theta}}_n, \mathbf{W}) &= \text{tr}(\mathbf{W}\mathbf{B}) + \boldsymbol{\eta}'\mathbf{W}\boldsymbol{\eta}. \end{aligned}$$

Preliminary test estimator (PTE)

$$\begin{aligned} b_3(\hat{\boldsymbol{\theta}}_n^{PT}) &= -\boldsymbol{\eta}H_{p+1}(\chi_{p-1}^2(\alpha), \Delta^2). \\ M_3(\hat{\boldsymbol{\theta}}_n^{PT}) &= \boldsymbol{\Sigma}_0 - \boldsymbol{\Sigma}_0\mathbf{J}'H_{p+1}(\chi_{p-1}^2(\alpha), \Delta^2) \\ &\quad + \boldsymbol{\eta}\boldsymbol{\eta}'\{2H_{p+1}(\chi_{p-1}^2(\alpha), \Delta^2) - H_{p+3}(\chi_{p-1}^2(\alpha), \Delta^2)\}. \\ R_3(\hat{\boldsymbol{\theta}}_n^{PT}, \mathbf{W}) &= \text{tr}(\mathbf{W}\boldsymbol{\Sigma}_0) - \text{tr}(\mathbf{W}\boldsymbol{\Sigma}_0\mathbf{J}')H_{p+1}(\chi_{p-1}^2(\alpha), \Delta^2) \\ &\quad + \boldsymbol{\eta}'\mathbf{W}\boldsymbol{\eta}\{2H_{p+1}(\chi_{p-1}^2(\alpha), \Delta^2) - H_{p+3}(\chi_{p-1}^2(\alpha), \Delta^2)\}. \end{aligned}$$

James-Stein type estimator (JSE)

$$\begin{aligned}
b_4(\hat{\boldsymbol{\theta}}_n^{JS}) &= -(p-3)\boldsymbol{\eta}E(\chi_{p+1}^{-2}(\Delta^2)). \\
M_4(\hat{\boldsymbol{\theta}}_n^{JS}) &= \boldsymbol{\Sigma}_0 - (p-3)\boldsymbol{\Sigma}_0\mathbf{J}'\{2E[\chi_{p+1}^{-2}(\Delta^2)] - (p-3)E[\chi_{p+1}^{-4}(\Delta^2)]\} \\
&\quad + (p-3)(p+1)\boldsymbol{\eta}\boldsymbol{\eta}'E[\chi_{p+3}^{-4}(\Delta^2)]. \\
R_4(\hat{\boldsymbol{\theta}}_n^{JS}, \mathbf{W}) &= \text{tr}(\mathbf{W}\boldsymbol{\Sigma}_0) - (p-3)\text{tr}(\mathbf{W}\boldsymbol{\Sigma}_0\mathbf{J}')\{2E[\chi_{p+1}^{-2}(\Delta^2)] \\
&\quad - (p-3)E[\chi_{p+1}^{-4}(\Delta^2)]\} + (p-3)(p+1)\boldsymbol{\eta}'\mathbf{W}\boldsymbol{\eta}E[\chi_{p+3}^{-4}(\Delta^2)].
\end{aligned}$$

Positive rule shrinkage estimator (PRSE)

$$\begin{aligned}
b_5(\hat{\boldsymbol{\theta}}_n^{s+}) &= -(p-3)\boldsymbol{\eta}E[\chi_{p+1}^{-2}(\Delta^2)] - \boldsymbol{\eta}E[1 - (p-3)\chi_{p+1}^{-2}(\Delta^2)] \\
&\quad \times I[\chi_{p+1}^2(\Delta^2) \leq p-3]. \\
M_5(\hat{\boldsymbol{\theta}}_n^{s+}) &= M_4(\hat{\boldsymbol{\theta}}_n^{JS}) - (p-3)\boldsymbol{\Sigma}_0\mathbf{J}' \\
&\quad \times \{E[(1 - (p-3)\chi_{p+1}^{-2}(\Delta^2))^2 I(\chi_{p+1}^2(\Delta^2) \leq p-3)]\} \\
&\quad + \boldsymbol{\eta}\boldsymbol{\eta}'\{2E[(1 - (p-3)\chi_{p+1}^{-2}(\Delta^2))I(\chi_{p+1}^2(\Delta^2) \leq p-3)] \\
&\quad - E[(1 - (p-3)\chi_{p+3}^{-2}(\Delta^2))^2 I(\chi_{p+3}^2(\Delta^2) \leq p-3)]\}. \\
R_5(\hat{\boldsymbol{\theta}}_n^{s+}, \mathbf{W}) &= R_4(\hat{\boldsymbol{\theta}}_n^{JS}, \mathbf{W}) - (p-3)\text{tr}(\mathbf{W}\boldsymbol{\Sigma}_0\mathbf{J}') \\
&\quad \times E[(1 - (p-3)\chi_{p+1}^{-2}(\Delta^2))^2 I(\chi_{p+1}^2(\Delta^2) \leq p-3)] \\
&\quad + \boldsymbol{\eta}'\mathbf{W}\boldsymbol{\eta}\{2E[(1 - (p-3)\chi_{p+1}^{-2}(\Delta^2))I(\chi_{p+1}^2(\Delta^2) \leq p-3)] \\
&\quad - E[(1 - (p-3)\chi_{p+3}^{-2}(\Delta^2))^2 I(\chi_{p+3}^2(\Delta^2) \leq p-3)]\}.
\end{aligned}$$

The risk analysis is similar to the analysis based on Theorem 3.4.6 and Theorem 3.4.7. In the next section we study the ADMSE matrix analysis of the estimators.

3.6.3 ADMSE Matrix Analysis of the Estimators

In this section we study the ADMSE matrix analysis of the estimators and determine their dominance properties.

Comparison of $\hat{\boldsymbol{\theta}}_n$ and $\tilde{\boldsymbol{\theta}}_n$

The ADMSE (matrix) difference of $\tilde{\boldsymbol{\theta}}_n$ and $\hat{\boldsymbol{\theta}}_n$ is

$$M_1(\tilde{\boldsymbol{\theta}}_n) - M_2(\hat{\boldsymbol{\theta}}_n) = \boldsymbol{\Sigma}_0 - (\mathbf{B} + \boldsymbol{\eta}\boldsymbol{\eta}') \quad (3.80)$$

The ADMSE-difference is p.s.d. whenever for a non-zero vector $\boldsymbol{\ell} = (\ell_1, \dots, \ell_p)'$, we have

$$\boldsymbol{\ell}'[M_1(\tilde{\boldsymbol{\theta}}_n) - M_2(\hat{\boldsymbol{\theta}}_n)]\boldsymbol{\ell} \geq 0.$$

Thus,

$$\boldsymbol{\ell}'\boldsymbol{\eta}\boldsymbol{\eta}'\boldsymbol{\ell} \leq \boldsymbol{\ell}'(\boldsymbol{\Sigma}_0 - \mathbf{B})\boldsymbol{\ell}$$

divided by $\boldsymbol{\ell}'\boldsymbol{\Sigma}_0\boldsymbol{\ell}$ we obtain

$$\frac{\boldsymbol{\ell}'\boldsymbol{\eta}\boldsymbol{\eta}'\boldsymbol{\ell}}{\boldsymbol{\ell}'\boldsymbol{\Sigma}_0\boldsymbol{\ell}} \leq \frac{\boldsymbol{\ell}'(\boldsymbol{\Sigma}_0 - \mathbf{B})\boldsymbol{\ell}}{\boldsymbol{\ell}'\boldsymbol{\Sigma}_0\boldsymbol{\ell}} \quad \forall \boldsymbol{\ell}.$$

$$\max_{\boldsymbol{\ell}} \frac{\boldsymbol{\ell}'\boldsymbol{\eta}\boldsymbol{\eta}'\boldsymbol{\ell}}{\boldsymbol{\ell}'\boldsymbol{\Sigma}_0\boldsymbol{\ell}} \leq \max_{\boldsymbol{\ell}} \frac{\boldsymbol{\ell}'(\boldsymbol{\Sigma}_0 - \mathbf{B})\boldsymbol{\ell}}{\boldsymbol{\ell}'\boldsymbol{\Sigma}_0\boldsymbol{\ell}}.$$

Thus,

$$\max_{\boldsymbol{\ell}} \frac{\boldsymbol{\ell}'\boldsymbol{\eta}\boldsymbol{\eta}'\boldsymbol{\ell}}{\boldsymbol{\ell}'\boldsymbol{\Sigma}_0\boldsymbol{\ell}} = \boldsymbol{\eta}'\boldsymbol{\Sigma}_0^{-1}\boldsymbol{\eta} = \Delta^2.$$

Using Courant theorem (see Theorem A.3.2)

$$\max_{\boldsymbol{\ell}} \frac{\boldsymbol{\ell}'(\boldsymbol{\Sigma}_0 - \mathbf{B})\boldsymbol{\ell}}{\boldsymbol{\ell}'\boldsymbol{\Sigma}_0\boldsymbol{\ell}} \leq ch_{max}(\mathbf{I}_p - \boldsymbol{\Sigma}_0^{-1}\mathbf{B}).$$

Therefore, we get

$$\Delta^2 \leq ch_{max}(\mathbf{I}_p - \Sigma_0^{-1}\mathbf{B}),$$

since $(\mathbf{I}_p - \Sigma_0^{-1}\mathbf{B})$ is an idempotent matrix (see Appendix) with maximum characteristic value equal to one (see Theorem A.3.1). Thus, (3.80) is p.s.d. if and only if $\Delta^2 \leq 1$. Therefore, $\hat{\boldsymbol{\theta}}_n$ performs better than $\tilde{\boldsymbol{\theta}}_n$ when $\Delta^2 \leq 1$, otherwise $\tilde{\boldsymbol{\theta}}_n$ performs better than $\hat{\boldsymbol{\theta}}_n$.

Comparison of $\hat{\boldsymbol{\theta}}_n^{PT}$, $\tilde{\boldsymbol{\theta}}_n$ and $\hat{\boldsymbol{\theta}}_n$

The ADMSE-difference of $\tilde{\boldsymbol{\theta}}_n$ and $\hat{\boldsymbol{\theta}}_n^{PT}$ is given by

$$\begin{aligned} M_1(\tilde{\boldsymbol{\theta}}_n) - M_3(\hat{\boldsymbol{\theta}}_n^{PT}) &= \Sigma_0 \mathbf{J}' H_{p+1}(\chi_{p-1}^2(\alpha), \Delta^2) \\ &\quad - \boldsymbol{\eta} \boldsymbol{\eta}' [2H_{p+1}(\chi_{p-1}^2(\alpha), \Delta^2) - H_{p+3}(\chi_{p-1}^2(\alpha), \Delta^2)]. \end{aligned}$$

The ADMSE-difference is p.s.d. whenever for a non-zero vector $\boldsymbol{\ell}$, we have

$$\boldsymbol{\ell}' \Sigma_0 \mathbf{J}' \boldsymbol{\ell} H_{p+1}(\chi_{p-1}^2(\alpha), \Delta^2) - \boldsymbol{\ell}' \boldsymbol{\eta} \boldsymbol{\eta}' \boldsymbol{\ell} [2H_{p+1}(\chi_{p-1}^2(\alpha), \Delta^2) - H_{p+3}(\chi_{p-1}^2(\alpha), \Delta^2)] \geq 0,$$

and consequently, we may write

$$\begin{aligned} \frac{\boldsymbol{\ell}' \boldsymbol{\eta} \boldsymbol{\eta}' \boldsymbol{\ell}}{\boldsymbol{\ell}' \Sigma_0 \boldsymbol{\ell}} &\leq \frac{\boldsymbol{\ell}' \Sigma_0 \mathbf{J}' \boldsymbol{\ell}}{\boldsymbol{\ell}' \Sigma_0 \boldsymbol{\ell}} \frac{H_{p+1}(\chi_{p-1}^2(\alpha), \Delta^2)}{[2H_{p+1}(\chi_{p-1}^2(\alpha), \Delta^2) - H_{p+3}(\chi_{p-1}^2(\alpha), \Delta^2)]} \quad \forall \boldsymbol{\ell}. \\ \max_{\boldsymbol{\ell}} \frac{\boldsymbol{\ell}' \boldsymbol{\eta} \boldsymbol{\eta}' \boldsymbol{\ell}}{\boldsymbol{\ell}' \Sigma_0 \boldsymbol{\ell}} &\leq \max_{\boldsymbol{\ell}} \frac{\boldsymbol{\ell}' \Sigma_0 \mathbf{J}' \boldsymbol{\ell}}{\boldsymbol{\ell}' \Sigma_0 \boldsymbol{\ell}} \frac{H_{p+1}(\chi_{p-1}^2(\alpha), \Delta^2)}{[2H_{p+1}(\chi_{p-1}^2(\alpha), \Delta^2) - H_{p+3}(\chi_{p-1}^2(\alpha), \Delta^2)]}. \end{aligned}$$

Therefore, we get

$$\Delta^2 \leq ch_{max}(\mathbf{J}') \frac{H_{p+1}(\chi_{p-1}^2(\alpha), \Delta^2)}{[2H_{p+1}(\chi_{p-1}^2(\alpha), \Delta^2) - H_{p+3}(\chi_{p-1}^2(\alpha), \Delta^2)]},$$

since $\mathbf{J}'$ is an idempotent matrix with maximum characteristic value equal to 1. Thus, the difference is p.s.d. if and only if

$$\Delta^2 \leq \frac{H_{p+1}(\chi_{p-1}^2(\alpha), \Delta^2)}{[2H_{p+1}(\chi_{p-1}^2(\alpha), \Delta^2) - H_{p+3}(\chi_{p-1}^2(\alpha), \Delta^2)]}. \quad (3.81)$$

This means $\hat{\boldsymbol{\theta}}_n^{PT}$ is superior to $\tilde{\boldsymbol{\theta}}_n$ in the range of Δ^2 given by (3.81), otherwise $\tilde{\boldsymbol{\theta}}_n$ is superior to $\hat{\boldsymbol{\theta}}_n^{PT}$. To compare $\hat{\boldsymbol{\theta}}_n^{PT}$ and $\hat{\boldsymbol{\theta}}_n$, we consider the ADMSE-difference

$$\begin{aligned} M_2(\hat{\boldsymbol{\theta}}_n) - M_3(\hat{\boldsymbol{\theta}}_n^{PT}) &= -(\boldsymbol{\Sigma}_0 - \mathbf{B} - \boldsymbol{\eta}\boldsymbol{\eta}') + \boldsymbol{\Sigma}_0\mathbf{J}'H_{p+1}(\chi_{p-1}^2(\alpha), \Delta^2) \\ &\quad - \boldsymbol{\eta}\boldsymbol{\eta}'[2H_{p+1}(\chi_{p-1}^2(\alpha), \Delta^2) - H_{p+3}(\chi_{p-1}^2(\alpha), \Delta^2)], \end{aligned}$$

the ADMSE-difference is negative semi-definite whenever for a non-zero vector $\boldsymbol{\ell}$, we have

$$\begin{aligned} &- \boldsymbol{\ell}'(\boldsymbol{\Sigma}_0 - \mathbf{B} - \boldsymbol{\eta}\boldsymbol{\eta}')\boldsymbol{\ell} + \boldsymbol{\ell}'\boldsymbol{\Sigma}_0\mathbf{J}'\boldsymbol{\ell}H_{p+1}(\chi_{p-1}^2(\alpha), \Delta^2) \\ &- \boldsymbol{\ell}'\boldsymbol{\eta}\boldsymbol{\eta}'\boldsymbol{\ell}[2H_{p+1}(\chi_{p-1}^2(\alpha), \Delta^2) - H_{p+3}(\chi_{p-1}^2(\alpha), \Delta^2)] \leq 0, \end{aligned}$$

and consequently, we may write

$$\frac{\boldsymbol{\ell}'\boldsymbol{\eta}\boldsymbol{\eta}'\boldsymbol{\ell}}{\boldsymbol{\ell}'\boldsymbol{\Sigma}_0\boldsymbol{\ell}} \leq \frac{\boldsymbol{\ell}'\boldsymbol{\Sigma}_0\mathbf{J}'\boldsymbol{\ell}}{\boldsymbol{\ell}'\boldsymbol{\Sigma}_0\boldsymbol{\ell}} \frac{1 - H_{p+1}(\chi_{p-1}^2(\alpha), \Delta^2)}{[1 - 2H_{p+1}(\chi_{p-1}^2(\alpha), \Delta^2) + H_{p+3}(\chi_{p-1}^2(\alpha), \Delta^2)]} \quad \forall \boldsymbol{\ell}.$$

Therefore, we get

$$\max_{\boldsymbol{\ell}} \frac{\boldsymbol{\ell}'\boldsymbol{\xi}\boldsymbol{\xi}'\boldsymbol{\ell}}{\boldsymbol{\ell}'\boldsymbol{\Sigma}_0\boldsymbol{\ell}} \leq \max_{\boldsymbol{\ell}} \frac{\boldsymbol{\ell}'\boldsymbol{\Sigma}_0\mathbf{J}'\boldsymbol{\ell}}{\boldsymbol{\ell}'\boldsymbol{\Sigma}_0\boldsymbol{\ell}} \frac{1 - H_{p+1}(\chi_{p-1}^2(\alpha), \Delta^2)}{[1 - 2H_{p+1}(\chi_{p-1}^2(\alpha), \Delta^2) + H_{p+3}(\chi_{p-1}^2(\alpha), \Delta^2)]}$$

or

$$\begin{aligned}\Delta^2 &\leq ch_{max}(\mathbf{J}') \frac{1 - H_{p+1}(\chi_{p-1}^2(\alpha), \Delta^2)}{[1 - 2H_{p+1}(\chi_{p-1}^2(\alpha), \Delta^2) + H_{p+3}(\chi_{p-1}^2(\alpha), \Delta^2)]} \\ \Delta^2 &\leq \frac{1 - H_{p+1}(\chi_{p-1}^2(\alpha), \Delta^2)}{[1 - 2H_{p+1}(\chi_{p-1}^2(\alpha), \Delta^2) + H_{p+3}(\chi_{p-1}^2(\alpha), \Delta^2)]}.\end{aligned}$$

Thus, in this range, $\hat{\boldsymbol{\theta}}_n$ is superior to $\hat{\boldsymbol{\theta}}_n^{PT}$, otherwise $\hat{\boldsymbol{\theta}}_n^{PT}$ is superior.

Comparison of $\hat{\boldsymbol{\theta}}_n^{JS}$, $\tilde{\boldsymbol{\theta}}_n$ and $\hat{\boldsymbol{\theta}}_n$

First, we compare $\hat{\boldsymbol{\theta}}_n^{JS}$ and $\tilde{\boldsymbol{\theta}}_n$. The ADMSE-difference in this case is given by

$$\begin{aligned}M_1(\tilde{\boldsymbol{\theta}}_n) - M_4(\hat{\boldsymbol{\theta}}_n^{JS}) &= (p-3)\boldsymbol{\Sigma}_0\mathbf{J}'\{2E[\chi_{p+1}^{-2}(\Delta^2)] - (p-3)E[\chi_{p+1}^{-4}(\Delta^2)]\} \\ &\quad - \boldsymbol{\eta}\boldsymbol{\eta}'(p-3)(p+1)E[\chi_{p+3}^{-4}(\Delta^2)].\end{aligned}$$

The ADMSE-difference is p.s.d. whenever for a non-zero vector $\boldsymbol{\ell}$, we have

$$\boldsymbol{\ell}'\boldsymbol{\Sigma}_0\mathbf{J}'\boldsymbol{\ell}(p-3)\{2E[\chi_{p+1}^{-2}(\Delta^2)] - (p-3)E[\chi_{p+1}^{-4}(\Delta^2)]\} - \boldsymbol{\ell}'\boldsymbol{\eta}\boldsymbol{\eta}'\boldsymbol{\ell}(p-3)(p+1)E[\chi_{p+3}^{-4}(\Delta^2)] \geq 0,$$

this implies that (after simplification)

$$\begin{aligned}\Delta^2(p-3)(p+1)E[\chi_{p+3}^{-4}(\Delta^2)] &\leq (p-3)\{2E[\chi_{p+1}^{-2}(\Delta^2)] - (p-3) \\ &\quad \times E[\chi_{p+1}^{-4}(\Delta^2)]\},\end{aligned}$$

which does not hold for all Δ^2 . Thus, $\hat{\boldsymbol{\theta}}_n^{JS}$ is not uniformly better than $\tilde{\boldsymbol{\theta}}_n$ under the MSE criterion.

Similarly, the comparison of $\hat{\boldsymbol{\theta}}_n$ and $\hat{\boldsymbol{\theta}}_n^{JS}$ is obtained through the ADMSE-difference

$$\begin{aligned}M_2(\hat{\boldsymbol{\theta}}_n) - M_4(\hat{\boldsymbol{\theta}}_n^{JS}) &= \boldsymbol{\xi}\boldsymbol{\xi}' - \boldsymbol{\xi}\boldsymbol{\xi}'(p-3)(p+1)E[\chi_{p+3}^{-4}] + \boldsymbol{\Sigma}_0\mathbf{J}' \\ &\quad (p-3)\left(2E[\chi_{p+1}^{-2}(\Delta^2)] - (p-3)E[\chi_{p+1}^{-4}(\Delta^2)]\right) - \boldsymbol{\Sigma}_0\mathbf{J}'.\end{aligned}$$

The ADMSE-difference is negative semi-definite whenever for a non-zero vector $\boldsymbol{\ell}$, we have

$$\begin{aligned} - \boldsymbol{\ell}' \boldsymbol{\Sigma}_0 \mathbf{J}' \boldsymbol{\ell} \{ & 1 - (p-3) \left(2E[\chi_{p+1}^{-2}(\Delta^2)] - (p-3)E[\chi_{p+1}^{-4}(\Delta^2)] \right) \} \\ & + \boldsymbol{\ell}' \boldsymbol{\eta} \boldsymbol{\eta}' \boldsymbol{\ell} \left(1 - (p-3)(p+1)E[\chi_{p+3}^{-4}(\Delta^2)] \right) \} \leq 0. \end{aligned}$$

As before, after simplification, we get

$$\begin{aligned} \Delta^2 \{ & 1 - (p-3)(p+1)E[\chi_{p+3}^{-4}(\Delta^2)] \} \leq 1 - (p-3) \{ 2E[\chi_{p+1}^{-2}(\Delta^2)] - (p-3) \\ & \times E[\chi_{p+1}^{-4}(\Delta^2)] \}. \end{aligned}$$

Under H_0 , the ADMSE-difference is negative semi-definite. Therefore, $\hat{\boldsymbol{\theta}}_n$ performs better than $\hat{\boldsymbol{\theta}}_n^{JS}$ at $\Delta^2 = 0$. Except for small intervals around 0, $\hat{\boldsymbol{\theta}}_n^{JS}$ performs better than $\hat{\boldsymbol{\theta}}_n$.

Comparison of $\hat{\boldsymbol{\theta}}_n^{PT}$, $\hat{\boldsymbol{\theta}}_n^{JS}$, and $\hat{\boldsymbol{\theta}}_n^{s+}$

Consider the ADMSE-difference

$$\begin{aligned} M_4(\hat{\boldsymbol{\theta}}_n^{JS}) - M_3(\hat{\boldsymbol{\theta}}_n^{PT}) &= \boldsymbol{\xi} \boldsymbol{\xi}' \{ (p-3)(p+1)E[\chi_{p+3}^{-4}(\Delta^2)] - 2H_{p+1}(\chi_{p-1}^2(\alpha), \Delta^2) \\ &+ H_{p+3}(\chi_{p-1}^2(\alpha), \Delta^2) \} + \boldsymbol{\Sigma}_0 \mathbf{J}' \{ H_{p+1}(\chi_{p-1}^2(\alpha), \Delta^2) - (p \\ &- 3)[2E(\chi_{p+1}^{-2}(\Delta^2)) - (p-3)E(\chi_{p+1}^{-4}(\Delta^2))] \} \quad (3.82) \end{aligned}$$

and

$$\begin{aligned}
M_5(\hat{\boldsymbol{\theta}}_n^{s+}) - M_3(\hat{\boldsymbol{\theta}}_n^{PT}) &= \boldsymbol{\xi}\boldsymbol{\xi}'\{(p-3)(p+1)E[\chi_{p+3}^{-4}(\Delta^2)] \\
&\quad + (2E[(1-(p-3)\chi_{p+1}^{-2}(\Delta^2))I(\chi_{p+1}^2(\Delta^2) \leq p-3)] \\
&\quad - E[(1-(p-3)\chi_{p+3}^{-2}(\Delta^2))^2I(\chi_{p+3}^2(\Delta^2) \leq p-3)]) \\
&\quad - 2H_{p+1}(\chi_{p-1}^2(\alpha), \Delta^2) + H_{p+3}(\chi_{p-1}^2(\alpha), \Delta^2)\} \\
&\quad + \boldsymbol{\Sigma}_0\mathbf{J}'\{H_{p+1}(\chi_{p-1}^2(\alpha), \Delta^2) - (p-3)E[(1-(p \\
&\quad - 3)\chi_{p+1}^{-2}(\Delta^2))^2I(\chi_{p+1}^2(\Delta^2) \leq p-3)] - 2(p-3) \\
&\quad \times E(\chi_{p+1}^{-2}(\Delta^2)) + (p-3)^2E(\chi_{p+1}^{-4}(\Delta^2))\} \} \quad (3.83)
\end{aligned}$$

The expression in (3.82) is p.s.d. whenever

$$\Delta^2 \leq \frac{H_{p+1}(\chi_{p-1}^2(\alpha), \Delta^2) - 2(p-3)E(\chi_{p+1}^{-2}(\Delta^2)) + (p-3)^2E(\chi_{p+1}^{-4}(\Delta^2))}{2H_{p+1}(\chi_{p-1}^2(\alpha), \Delta^2) - H_{p+3}(\chi_{p-1}^2(\alpha), \Delta^2) - (p-3)(p+1)E(\chi_{p+3}^{-4}(\Delta^2))} \quad (3.84)$$

Thus, $\hat{\boldsymbol{\theta}}_n^{PT}$ performs better than $\hat{\boldsymbol{\theta}}_n^{JS}$ whenever Δ^2 satisfies (3.84); otherwise $\hat{\boldsymbol{\theta}}_n^{JS}$ performs better than $\hat{\boldsymbol{\theta}}_n^{PT}$. Neither estimators dominate the other.

Similarly, the expression in (3.84) is p.s.d. whenever

$$\begin{aligned}
\Delta^2 \leq & \{H_{p+1}(\chi_{p-1}^2(\alpha), \Delta^2) - (p-3)E[(1-(p-3)\chi_{p+1}^{-2}(\Delta^2))^2I(\chi_{p+1}^2(\Delta^2) \leq p-3)] \\
& - 2(p-3)E(\chi_{p+1}^{-2}(\Delta^2)) + (p-3)^2E(\chi_{p+1}^{-4}(\Delta^2))\} / \{2H_{p+1}(\chi_{p-1}^2(\alpha), \Delta^2) \\
& - H_{p+3}(\chi_{p-1}^2(\alpha), \Delta^2) - 2E[(1-(p-3)\chi_{p+1}^{-2}(\Delta^2))I(\chi_{p+1}^2(\Delta^2) \leq p-3)] \\
& + E[(1-(p-3)\chi_{p+3}^{-2}(\Delta^2))^2I(\chi_{p+3}^2(\Delta^2) \leq p-3)] - (p-3)(p+1) \\
& \times E(\chi_{p+3}^{-4}(\Delta^2))\}. \quad (3.85)
\end{aligned}$$

Thus, $\hat{\boldsymbol{\theta}}_n^{PT}$ performs better than $\hat{\boldsymbol{\theta}}_n^{s+}$ whenever Δ^2 satisfies (3.85); otherwise $\hat{\boldsymbol{\theta}}_n^{s+}$ performs better than $\hat{\boldsymbol{\theta}}_n^{PT}$. Neither estimators dominate the other.

Next, we consider the ADMSE-difference under H_0 for the comparison of $\hat{\boldsymbol{\theta}}_n^{s+}$ and $\hat{\boldsymbol{\theta}}_n$:

$$\begin{aligned}
M_5(\hat{\boldsymbol{\theta}}_n^{s+}) - M_2(\hat{\boldsymbol{\theta}}_n) &= (\boldsymbol{\Sigma}_0 - \mathbf{B}) - \boldsymbol{\Sigma}_0 \mathbf{J}' \{2(p-3)E(\chi_{p+1}^{-2}) - (p-3)^2 E(\chi_{p+1}^{-4}) \\
&\quad + (p-3)E[(1 - (p-3)\chi_{p+1}^{-2})^2 I(\chi_{p+1}^2 \leq (p-3))]\} \\
&\quad + \boldsymbol{\eta} \boldsymbol{\eta}' \{(p-3)(p+1)E(\chi_{p+3}^{-4}) + 2E[(1 - (p-3)\chi_{p+1}^{-2}) \\
&\quad \times I(\chi_{p+1}^2 \leq p-3)] - E[(1 - (p-3)\chi_{p+1}^{-2})^2 \\
&\quad \times I(\chi_{p+3}^2 \leq p-3)]\} - \boldsymbol{\eta} \boldsymbol{\eta}'
\end{aligned}$$

which is p.s.d. whenever

$$E[(1 - (p-3)\chi_{p+1}^{-2})^2 I(\chi_{p+1}^2 \leq p-3)] \leq \frac{1}{p-3} - \frac{1}{p-1}. \quad (3.86)$$

This implies that $\hat{\boldsymbol{\theta}}_n$ is superior to $\hat{\boldsymbol{\theta}}_n^{s+}$ under H_0 whenever (3.86) is satisfied.

A similar comparison of $\hat{\boldsymbol{\theta}}_n^{s+}$ with $\hat{\boldsymbol{\theta}}_n^{PT}$ under H_0 shows that $M_5(\hat{\boldsymbol{\theta}}_n^{s+}) - M_3(\hat{\boldsymbol{\theta}}_n^{PT})$ is p.s.d. whenever there exist a set $\mathcal{A}$ of α such that

$$\begin{aligned}
\mathcal{A} &= \{\alpha : H_{p+1}(\chi_{p-1}^2(\alpha), 0) \geq \frac{p-3}{p-1} + (p-3) \\
&\quad \times E[(1 - (p-3)\chi_{p+1}^{-2})^2 I(\chi_{p+1}^2 \leq p-3)]\}.
\end{aligned} \quad (3.87)$$

Thus, $\hat{\boldsymbol{\theta}}_n^{PT}$ is superior to $\hat{\boldsymbol{\theta}}_n^{s+}$ whenever (3.87) is satisfied; otherwise $\hat{\boldsymbol{\theta}}_n^{s+}$ is superior. Therefore, $\hat{\boldsymbol{\theta}}_n^{s+}$ does not always dominate $\hat{\boldsymbol{\theta}}_n^{PT}$, when the null hypothesis $H_0 : \boldsymbol{\theta} = \theta_0 \mathbf{1}_p$ holds. Hence, one can order the ADMSE matrices as

$$M_1(\tilde{\boldsymbol{\theta}}_n) \geq M_4(\hat{\boldsymbol{\theta}}_n^{JS}) \geq M_3(\hat{\boldsymbol{\theta}}_n^{PT}) \geq M_2(\hat{\boldsymbol{\theta}}_n)$$

The position of PTE may be changed depending on the size of the test to the order

$$M_1(\tilde{\boldsymbol{\theta}}_n) \geq M_3(\hat{\boldsymbol{\theta}}_n^{PT}) \geq M_4(\hat{\boldsymbol{\theta}}_n^{JS}) \geq M_5(\hat{\boldsymbol{\theta}}_n^{s+}) \geq M_2(\hat{\boldsymbol{\theta}}_n).$$

Comparison of $\hat{\boldsymbol{\theta}}_n^{s+}$, $\hat{\boldsymbol{\theta}}_n^{JS}$, and $\tilde{\boldsymbol{\theta}}_n$

Finally, we compare $\hat{\boldsymbol{\theta}}_n^{s+}$ and $\hat{\boldsymbol{\theta}}_n^{JS}$. In this case, the ADMSE-difference yields

$$\begin{aligned} M_4(\hat{\boldsymbol{\theta}}_n^{JS}) - M_5(\hat{\boldsymbol{\theta}}_n^{s+}) &= (p-3)\boldsymbol{\Sigma}_0\mathbf{J}'E[(1-(p-3)\chi_{p+1}^{-2}(\Delta^2))^2I(\chi_{p+1}^2(\Delta^2) \leq p-3)] \\ &\quad + \boldsymbol{\xi}\boldsymbol{\xi}'\{E[(1-(p-3)\chi_{p+3}^{-2}(\Delta^2))^2I(\chi_{p+3}^2(\Delta^2) \leq p-3)] \\ &\quad - 2E[(1-(p-3)\chi_{p+1}^{-2}(\Delta^2))I(\chi_{p+1}^2(\Delta^2) \leq p-3)]\} \geq 0. \end{aligned}$$

The ADMSE-difference is p.s.d. for all Δ^2 and hence $\hat{\boldsymbol{\theta}}_n^{s+}$ dominates $\hat{\boldsymbol{\theta}}_n^{JS}$ uniformly.

We can order the ADMSE of $\hat{\boldsymbol{\theta}}_n^{JS}$ and $\hat{\boldsymbol{\theta}}_n^{s+}$ as

$$M_4(\hat{\boldsymbol{\theta}}_n^{JS}) \geq M_5(\hat{\boldsymbol{\theta}}_n^{s+}).$$

Chapter 4

Effect Size Estimation: Multivariate Models

In this chapter, we consider multivariate models which are generated by more than one response on experimental as well as on control subjects and the analysis steps into studying these correlated effect sizes simultaneously. Dependent effect sizes occurs in several ways. When multiple treatments are compared against a common control, the effect sizes will be correlated. Studies of this type are called multiple treatment studies. In other studies there is a single treatment and a single control, but multiple measures are used as end-points for each subject. Studies of this type are called multiple end-point studies. Analyses for each type are important in meta analysis. We shall discuss the theory of effect size estimation in multiple end-point studies.

The notation that is used is given in section 4.1. In section 4.2 we present multiple-study problem. We present the proposed estimators. The asymptotic properties of the estimators under local alternatives are investigated. We obtain the asymptotic distributional bias, the asymptotic distributional mean square errors matrices and the asymptotic distributional risk expressions for the estimators.

4.1 Model and Notation

We consider the problem where there are p response variables for each subject in each of k independent studies. The data layout is given in Table 4.1 below. So that the

Table 4.1: Data layout

Study	<u>Experimental Group</u>		<u>Control Group</u>	
	Person	<u>Measure</u> $1 \dots p$	Person	<u>Measure</u> $1 \dots p$
1	1	$Y_{111}^E \dots Y_{1p1}^E$	1	$Y_{111}^C \dots Y_{1p1}^C$
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$
1	n_1^E	$Y_{11n_1^E}^E \dots Y_{1pn_1^E}^E$	n_1^C	$Y_{11n_1^C}^C \dots Y_{1pn_1^C}^C$
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$
k	1	$Y_{k11}^E \dots Y_{kp1}^E$	1	$Y_{k11}^C \dots Y_{kp1}^C$
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$
k	n_k^E	$Y_{k1n_k^E}^E \dots Y_{kpn_k^E}^E$	n_k^C	$Y_{k1n_k^C}^C \dots Y_{kpn_k^C}^C$

vectors $\mathbf{Y}_{ij}^E = (Y_{i1j}^E, \dots, Y_{ipj}^E)'$ and $\mathbf{Y}_{ij}^C = (Y_{i1j}^C, \dots, Y_{ipj}^C)'$ are the responses on the j -th subject (person) in the i th experimental and control group of sizes n_i^E, n_i^C respectively. Also, let $n_i = n_i^E + n_i^C, i = 1, \dots, k, n = \sum_{i=1}^k n_i$. Assume that for the i th study, the observations of the experimental and control groups, $\mathbf{Y}_{ij}^E$ and $\mathbf{Y}_{ij}^C$, are independently distributed with the p -variate normal distribution with mean vectors

$$\boldsymbol{\mu}_i^E = (\mu_{i1}^E, \dots, \mu_{ip}^E)', \boldsymbol{\mu}_i^C = (\mu_{i1}^C, \dots, \mu_{ip}^C)'$$

and with the common covariance matrix $\Sigma^{(i)}$. That is

$$\begin{aligned}\mathbf{Y}_{ij}^E &\sim N(\boldsymbol{\mu}_i^E, \Sigma^{(i)}), j = 1 \cdots, n_i^E, \\ \mathbf{Y}_{ij}^C &\sim N(\boldsymbol{\mu}_i^C, \Sigma^{(i)}), j = 1 \cdots, n_i^C, i = 1, \cdots, k.\end{aligned}$$

The vector $\boldsymbol{\delta}_i$ of effect sizes represents the population effect size for each component separately. In particular, the vector of i th population effect size is

$$\boldsymbol{\delta}_i = (\delta_{i1}, \cdots, \delta_{ip})',$$

$$\boldsymbol{\delta}_i = (\boldsymbol{\mu}_i^E - \boldsymbol{\mu}_i^C) D_{\Sigma^{(i)}}^{-1/2}$$

where $D_{\Sigma^{(i)}} = \text{diag}(\sigma_{11}^{(i)}, \cdots, \sigma_{pp}^{(i)})$. Our problem is the to obtain efficient estimates of the vector of effect sizes.

4.2 Multiple-Study Problem

In this section we discuss the extension of the estimation of effect sizes from a series of k independent experiments (studies) due to Hedges and Olkin (1985). We consider the problem where there are p response variables for each subject in each of these independent experiments. In the process, we first postulate the hypothesis of homogeneity of effect sizes, and formulate a test-statistics for the test of the postulated hypothesis and subsequently define three new estimators of effect size besides the usual estimator and pooled estimator. Let $\nu_i \mathbf{S}^{(p_i)} = (n_i^E - 1) \mathbf{S}^E + (n_i^C - 1) \mathbf{S}^C$ be the pooled within group estimate of the common population covariance matrix $\Sigma^{(i)}$

where $\nu_i = n_i^E + n_i^C - 2$,

$$\begin{aligned} (n_i^E - 1)\mathbf{S}^E &= \sum_{j=1}^{n_i^E} (\mathbf{Y}_{ij}^E - \bar{\mathbf{Y}}_i^E)(\mathbf{Y}_{ij}^E - \bar{\mathbf{Y}}_i^E)'. \\ (n_i^C - 1)\mathbf{S}^C &= \sum_{j=1}^{n_i^C} (\mathbf{Y}_{ij}^C - \bar{\mathbf{Y}}_i^C)(\mathbf{Y}_{ij}^C - \bar{\mathbf{Y}}_i^C)'. \\ \bar{\mathbf{Y}}_i^E &= (\bar{Y}_{i1}^E, \dots, \bar{Y}_{ip}^E)'. \\ \bar{\mathbf{Y}}_i^C &= (\bar{Y}_{i1}^C, \dots, \bar{Y}_{ip}^C)'. \end{aligned}$$

It is well-known in the literature that the unbiased estimator of $\boldsymbol{\delta}_i = (\delta_{i1}, \dots, \delta_{ip})'$ is given by Hedges and Olkin (1985) as $\tilde{\boldsymbol{\delta}}_{n_i} = (\tilde{\delta}_{i1}, \dots, \tilde{\delta}_{ip})'$ where

$$\tilde{\delta}_{il} = c(\nu_i) \frac{\bar{Y}_{il}^E - \bar{Y}_{il}^C}{\sqrt{s_{ll}^{(p_i)}}}, \quad l = 1, \dots, p.$$

In matrix notation

$$\tilde{\boldsymbol{\delta}}_{n_i} = c(\nu_i)(\bar{\mathbf{Y}}_i^E - \bar{\mathbf{Y}}_i^C)D_{\hat{\boldsymbol{\Sigma}}^{(i)}}^{-1}, \quad D_{\hat{\boldsymbol{\Sigma}}^{(i)}} = \text{Diag}(\sqrt{s_{11}^{(p_i)}}, \dots, \sqrt{s_{pp}^{(p_i)}}).$$

The vector of estimated effect sizes, $\tilde{\boldsymbol{\delta}}_{n_i}$ has a very complicated distribution in small samples related to multivariate noncentral t-distribution. However, if the sample sizes are large, we obtain the following theorems

Theorem 4.2.1. *The asymptotic distribution of $\tilde{\boldsymbol{\delta}}_{n_i}$ is given by*

$$\sqrt{n} (\tilde{\boldsymbol{\delta}}_{n_i} - \boldsymbol{\delta}_i) \sim N_p(0, \boldsymbol{\Psi}_{0i}). \quad (4.1)$$

$\boldsymbol{\Psi}_{0i} = \lim_{n \rightarrow \infty} n \boldsymbol{\Psi}_{n_i}$ such that $\frac{n_i^E}{n_i} \rightarrow \pi_i^E$, $\frac{n_i^C}{n_i} \rightarrow \pi_i^C$, $\frac{n_i}{n} \rightarrow \pi_i$, π_i^E , π_i^C and π_i remain

fixed as $\min_{1 \leq i \leq k} n_i \rightarrow \infty$, which implies $n \rightarrow \infty$ where

$$\Psi_{n_i} = ((\Psi_{lt n_i})) \text{ with} \quad (4.2)$$

$$\Psi_{ill} = \left(\frac{1}{n_i^E} + \frac{1}{n_i^C} \right) + \frac{\delta_{il}^2}{2(n_i^E + n_i^C - 2)}, \quad l = 1, \dots, p. \quad (4.3)$$

$$\Psi_{ilt} = \rho_{ilt} \sqrt{\Psi_{ill} \Psi_{itt}}, \quad l \neq t, \quad l, t = 1, \dots, p. \quad (4.4)$$

where ρ_{ilt} is the population correlation coefficient between the l and t components in the i th study.

That is

$$\Psi_{0i} = ((\Psi_{lt})) \text{ with} \quad (4.5)$$

$$\Psi_{ill} = \frac{1}{\pi_i} \left[\left(\frac{1}{\pi_i^E} + \frac{1}{\pi_i^C} \right) + \frac{\delta_{il}^2}{2(\pi_i^E + \pi_i^C)} \right], \quad l = 1, \dots, p, \quad i = 1, \dots, k. \quad (4.6)$$

$$\Psi_{ilt} = \rho_{ilt} \sqrt{\Psi_{ill} \Psi_{itt}}, \quad l \neq t, \quad l, t = 1, \dots, p. \quad (4.7)$$

We consider the homogeneity hypothesis defined by

$$H_{0i} : \delta_i = \delta_{0i} \mathbf{1}_p, \quad (\delta_{0i} \text{ unknown}), \quad i = 1, \dots, k. \quad (4.8)$$

That is, the effect sizes from different response variables are the same within each study. Under this hypothesis, the restricted estimator of δ_{0i} is obtained by minimizing

$$(\tilde{\delta}_{n_i} - \delta_{0i} \mathbf{1}_p)' \hat{\Psi}_{n_i}^{-1} (\tilde{\delta}_{n_i} - \delta_{0i} \mathbf{1}_p)$$

to obtain the estimator

$$\hat{\delta}_{0i} = \frac{\mathbf{1}_p' \hat{\Psi}_{n_i}^{-1} \tilde{\delta}_{n_i}}{\mathbf{1}_p' \hat{\Psi}_{n_i}^{-1} \mathbf{1}_p}, \quad (4.9)$$

where $\hat{\Psi}_{n_i}$ is the estimate of Ψ_{n_i} with elements

$$\hat{\Psi}_{lln_i} = \left(\frac{1}{n_i^E} + \frac{1}{n_i^C} \right) + \frac{\tilde{\delta}_{il}^2}{2(n_i^E + n_i^C - 2)}, \quad l = 1, \dots, p. \quad (4.10)$$

$$\hat{\Psi}_{lt n_i} = r_{ilt} \sqrt{\hat{\Psi}_{lln_i} \hat{\Psi}_{tt n_i}}, \quad l \neq t, \quad l, t = 1, \dots, p. \quad (4.11)$$

where r_{ilt} is the sample correlation coefficient in the i th experiment.

In general, H_{0i} , $i = 1, \dots, k$ is suspicious. We clear this suspicion following Berger and Bock (1976) for each $i = 1, \dots, k$ based on a statistical test defined by

$$\mathcal{L}_n = \mathcal{L}_{n_1} + \dots + \mathcal{L}_{n_k}.$$

$$\text{where, } \mathcal{L}_{n_i} = n(\tilde{\boldsymbol{\delta}}_{n_i} - \hat{\delta}_{0i} \mathbf{1}_p)' (n \hat{\Psi}_{n_i})^{-1} (\tilde{\boldsymbol{\delta}}_{n_i} - \hat{\delta}_{0i} \mathbf{1}_p), \quad i = 1, \dots, k.$$

Theorem 4.2.2. *Under H_{0i} , $i = 1, \dots, k$ in (4.8) as $\min_{1 \leq i \leq k} n_i \rightarrow \infty$ which implies $n \rightarrow \infty$ with $\frac{n_i^E}{n_i}$, $\frac{n_i^C}{n_i}$ and $\frac{n_i}{n}$ remain fixed for all $i = 1, \dots, k$, then $\mathcal{L}_n$ follows an asymptotic chi-square distribution with $p^* = k(p - 1)$ D.F.*

The vectors of the true effect sizes from the k studies in the meta-analysis can be combined into one large kp -dimensional vector. Denote this large kp -dimensional vector by $\text{vec } \boldsymbol{\delta}$ where $(\text{vec } \boldsymbol{\delta})' = (\boldsymbol{\delta}'_1, \dots, \boldsymbol{\delta}'_k)$ is a row vector obtained from $\boldsymbol{\delta}$ by arranging its rows in succession. The vectors of effect size estimates from the k studies can be combined into one large kp -dimensional vector of estimated effect sizes as well. Denote this large kp -dimensional vector by $\text{vec } \tilde{\boldsymbol{\delta}}_n$ where $(\text{vec } \tilde{\boldsymbol{\delta}}_n)' = (\tilde{\boldsymbol{\delta}}'_{n_1}, \dots, \tilde{\boldsymbol{\delta}}'_{n_k})$ is a row vector obtained from $\tilde{\boldsymbol{\delta}}_n$ by arranging its rows in succession, and the estimated covariance matrices will be diagonal blocks of a combined $kp \times kp$

estimated covariance matrix. Denote this block-diagonal matrix by $\hat{\Psi}_n$ where

$$\hat{\Psi}_n = \begin{bmatrix} \hat{\Psi}_{n_1} & \mathbf{0} & \cdots & \mathbf{0} \\ \vdots & & \ddots & \vdots \\ \mathbf{0} & \mathbf{0} & \cdots & \hat{\Psi}_{n_k} \end{bmatrix}. \quad (4.12)$$

In a similar fashion, the vectors of the restricted estimators from the k studies in the meta-analysis can be combined into one large kp -dimensional vector of estimated effect sizes. Denote this large kp -dimensional vector by $(\text{vec } \hat{\delta}_n)'$ where $(\text{vec } \hat{\delta}_n)' = (\hat{\delta}_{01}\mathbf{1}_p', \dots, \hat{\delta}_{0k}\mathbf{1}_p')$ is a row vector obtained from $\hat{\delta}_n$ by arranging its rows in succession.

4.2.1 Quasi Empirical Bayes Estimators

As before, we define a general class of estimators of $\text{vec } \delta$ by combining $\text{vec } \tilde{\delta}_n$ and $\text{vec } \hat{\delta}_n$. We may consider the following estimators

$$\begin{aligned} \text{vec } \hat{\delta}_n^{PT} &= \text{vec } \tilde{\delta}_n - (\text{vec } \tilde{\delta}_n - \text{vec } \hat{\delta}_n)I(\mathcal{L}_n \leq \chi_{p^*, \alpha}^2). \\ \text{vec } \hat{\delta}_n^{JS} &= \text{vec } \hat{\delta}_n + [1 - (p^* - 2)\mathcal{L}_n^{-1}](\text{vec } \tilde{\delta}_n - \text{vec } \hat{\delta}_n). \\ \text{vec } \hat{\delta}_n^{s+} &= \text{vec } \hat{\delta}_n + [1 - (p^* - 2)\mathcal{L}_n^{-1}]I(\mathcal{L}_n > (p^* - 2))(\text{vec } \tilde{\delta}_n - \text{vec } \hat{\delta}_n). \end{aligned}$$

4.2.2 Asymptotic Distributional Properties of the Estimators

We consider the sequence of local alternatives for the derivation of the asymptotic distribution of the estimators

$$\begin{aligned} \mathcal{K}_{(n)} : \text{vec } \delta_{(n)} &= \text{vec } \delta + n^{-1/2} \text{vec } \xi, \quad (\text{vec } \xi)' = (\xi_1', \dots, \xi_k') \neq \mathbf{0}, \\ \xi_i' &= (\xi_{i1}, \dots, \xi_{ip}), i = 1, \dots, k. \end{aligned}$$

We arrive at the next theorem on the asymptotic distributions of the various statistics

Theorem 4.2.3. *Under $\{\mathcal{K}_{(n)}\}$ and the assumed regularity conditions, as $n \rightarrow \infty$, the following hold:*

- (i) $\sqrt{n}(\text{vec } \tilde{\boldsymbol{\delta}}_n - \text{vec } \boldsymbol{\delta}_{(n)}) \xrightarrow{\mathcal{D}} \text{vec } \mathbf{Z}_0 \sim N_{kp}(\mathbf{0}, \boldsymbol{\Psi}_0), \boldsymbol{\Psi}_0 = \text{Diag}(\boldsymbol{\Psi}_{01}, \dots, \boldsymbol{\Psi}_{0k})$
 $, \boldsymbol{\Psi}_{0i} \text{ given in } (4.5 - 4.7).$
- (ii) $\sqrt{n}(\text{vec } \tilde{\boldsymbol{\delta}}_n - \text{vec } \hat{\boldsymbol{\delta}}_n) \xrightarrow{\mathcal{D}} \text{vec } \mathbf{Z}_1, \text{vec } \mathbf{Z}_1 \sim N_{kp}(\text{vec } \boldsymbol{\eta}, \boldsymbol{\Psi}_0 \mathbf{J}'),$
 $\mathbf{J} = \text{Diag}(\mathbf{J}_1, \dots, \mathbf{J}_k), \mathbf{J}_i = \mathbf{I}_p - \frac{\mathbf{1}_p \mathbf{1}_p' \boldsymbol{\Psi}_{0i}^{-1}}{\mathbf{1}_p' \boldsymbol{\Psi}_{0i}^{-1} \mathbf{1}_p}, (\text{vec } \boldsymbol{\eta})' = (\boldsymbol{\eta}'_1, \dots, \boldsymbol{\eta}'_k), \boldsymbol{\eta}_i = \mathbf{J}_i \boldsymbol{\xi}_i.$
- (iii) $\sqrt{n}(\text{vec } \hat{\boldsymbol{\delta}}_n - \text{vec } \boldsymbol{\delta}_{(n)}) \xrightarrow{\mathcal{D}} \text{vec } \mathbf{Z}_2 - \text{vec } \boldsymbol{\eta}, \text{vec } \mathbf{Z}_2 \sim N_{kp}(\mathbf{0}, \mathbf{B}),$
 $\mathbf{B} = \text{Diag}(\mathbf{B}_1, \dots, \mathbf{B}_k), \mathbf{B}_i = \frac{\mathbf{1}_p \mathbf{1}_p'}{\mathbf{1}_p' \boldsymbol{\Psi}_{0i}^{-1} \mathbf{1}_p}.$
- (iv) $\lim_{n \rightarrow \infty} P\{\mathcal{L}_n \leq x | \mathcal{K}_{(n)}\} = H_{p^*}(x, \Delta^2) \text{ where } \Delta^2 = (\text{vec } \boldsymbol{\eta})' \boldsymbol{\Psi}_0^{-1} \text{vec } \boldsymbol{\eta}.$

where $H_{p^*}(\cdot, \Delta^2)$ is the cdf of noncentral chi-square distribution with noncentrality parameter $\frac{1}{2}\Delta^2$ and $p^* = k(p - 1)$ D.F.

Based on Theorem 4.2.3 we obtain the following results.

$$\begin{aligned} \sqrt{n} (\text{vec } \hat{\boldsymbol{\delta}}_n^{PT} - \text{vec } \boldsymbol{\delta}_{(n)}) &\xrightarrow{\mathcal{D}} \text{vec } \mathbf{Z}_0 - \text{vec } \mathbf{Z}_1 \\ &\times I[(\text{vec } \mathbf{Z}_1)' \boldsymbol{\Psi}_0^{-1} \text{vec } \mathbf{Z}_1 < \chi_{p^*, \alpha}^2] \end{aligned} \quad (4.13)$$

$$\begin{aligned} \sqrt{n} (\text{vec } \hat{\boldsymbol{\delta}}_n^{JS} - \text{vec } \boldsymbol{\delta}_{(n)}) &\xrightarrow{\mathcal{D}} \text{vec } \mathbf{Z}_0 - (p^* - 2)\text{vec } \mathbf{Z}_1 \\ &\times [(\text{vec } \mathbf{Z}_1)' \boldsymbol{\Psi}_0^{-1} \text{vec } \mathbf{Z}_1]^{-1} \end{aligned} \quad (4.14)$$

$$\begin{aligned}
& \sqrt{n} \quad (vec \hat{\boldsymbol{\delta}}_n^{s+} - vec \boldsymbol{\delta}_{(n)}) \\
& \xrightarrow{D} \quad vec \mathbf{Z}_0 - (p^* - 2)vec \mathbf{Z}_1 [(vec \mathbf{Z}_1)' \boldsymbol{\Psi}_0^{-1} vec \mathbf{Z}_1]^{-1} \\
& \quad - vec \mathbf{Z}_1 \{1 - (p^* - 2)[(vec \mathbf{Z}_1)' \boldsymbol{\Psi}_0^{-1} vec \mathbf{Z}_1]^{-1}\} \\
& \quad \times I[(vec \mathbf{Z}_1)' \boldsymbol{\Psi}_0^{-1} vec \mathbf{Z}_1 < p^* - 2].
\end{aligned} \tag{4.15}$$

Define the asymptotic distributional bias (ADB) of the estimator $vec \boldsymbol{\delta}_n^*$, where $vec \boldsymbol{\delta}_n^*$ stands for $vec \tilde{\boldsymbol{\delta}}_n$, $vec \hat{\boldsymbol{\delta}}_n^*$, $vec \hat{\boldsymbol{\delta}}_n^{PT}$, $vec \hat{\boldsymbol{\delta}}_n^{JS}$, and $vec \hat{\boldsymbol{\delta}}_n^{s+}$ by

$$b(vec \boldsymbol{\delta}_n^*) = \lim_{n \rightarrow \infty} E[\sqrt{n} (vec \boldsymbol{\delta}_n^* - vec \boldsymbol{\delta}_{(n)})]. \tag{4.16}$$

Also, define the asymptotic distributional MSE (ADMSE) matrices by

$$M(vec \boldsymbol{\delta}_n^*) = \lim_{n \rightarrow \infty} E[n (vec \boldsymbol{\delta}_n^* - vec \boldsymbol{\delta}_{(n)})(vec \boldsymbol{\delta}_n^* - vec \boldsymbol{\delta}_{(n)})']. \tag{4.17}$$

We compare the five estimators using the weighted quadratic loss function defined by

$$\begin{aligned}
L(vec \boldsymbol{\delta}_n^*, vec \boldsymbol{\delta}_{(n)}) &= n(vec \boldsymbol{\delta}_n^* - vec \boldsymbol{\delta}_{(n)})' \mathbf{W} (vec \boldsymbol{\delta}_n^* - vec \boldsymbol{\delta}_{(n)}) \\
&= n \operatorname{tr}[\mathbf{W}(vec \boldsymbol{\delta}_n^* - vec \boldsymbol{\delta}_{(n)})(vec \boldsymbol{\delta}_n^* - vec \boldsymbol{\delta}_{(n)})'].
\end{aligned} \tag{4.18}$$

where $\mathbf{W}$ is a given positive semi-definite $(kp \times kp)$ block-diagonal matrix ,

$$\mathbf{W} = \begin{bmatrix} \mathbf{W}_1 & \mathbf{0} & \cdots & \mathbf{0} \\ \vdots & & \ddots & \vdots \\ \mathbf{0} & \mathbf{0} & \cdots & \mathbf{W}_k \end{bmatrix}. \tag{4.19}$$

where $\mathbf{W}_i, i = 1, \dots, k$ is $(p \times p)$ positive semi-definite matrix, and $\operatorname{tr}(\mathbf{A}) = \text{trace of } \mathbf{A}$.

Then the asymptotic distributional quadratic risk (ADQR) of $\boldsymbol{\delta}_n^*$ is given by

$$R(vec \boldsymbol{\delta}_n^*, vec \boldsymbol{\delta}_{(n)}) = EL(vec \boldsymbol{\delta}_n^*, vec \boldsymbol{\delta}_{(n)}) = \operatorname{tr}(\mathbf{W}M(vec \boldsymbol{\delta}_n^*)). \tag{4.20}$$

Based on Theorem 4.2.3 and the results in (4.13 - 4.15) and using Theorem A.2.6 and A.2.7, we obtain the ADB, ADMSE and ADQR expressions. Since the derivation of these expressions are very similar to the cases treated in Theorem 3.4.4 and Theorem 3.4.5, we omit these details of the derivation. Expressions of the ADB, ADMSE, and ADQR for the five estimators are as follows:

Unrestricted estimator (UE)

$$\begin{aligned} b_1(\text{vec } \tilde{\boldsymbol{\delta}}_n) &= \mathbf{0}. \\ M_1(\text{vec } \tilde{\boldsymbol{\delta}}_n) &= \boldsymbol{\Psi}_0. \\ R_1(\text{vec } \tilde{\boldsymbol{\delta}}_n, \mathbf{W}) &= \text{tr}(\mathbf{W}\boldsymbol{\Psi}_0). \end{aligned}$$

Restricted estimator (RE)

$$\begin{aligned} b_2(\text{vec } \hat{\boldsymbol{\delta}}_n) &= -\text{vec } \boldsymbol{\eta}. \\ M_2(\text{vec } \hat{\boldsymbol{\delta}}_n) &= \mathbf{B} + \text{vec } \boldsymbol{\eta}(\text{vec } \boldsymbol{\eta})'. \\ R_2(\text{vec } \hat{\boldsymbol{\delta}}_n, \mathbf{W}) &= \text{tr}(\mathbf{W}\mathbf{B}) + (\text{vec } \boldsymbol{\eta})' \mathbf{W} \text{vec } \boldsymbol{\eta}. \end{aligned}$$

Preliminary test estimator (PTE)

$$\begin{aligned} b_3(\text{vec } \hat{\boldsymbol{\delta}}_n^{PT}) &= -\text{vec } \boldsymbol{\eta} H_{p^*+2}(\chi_{p^*}^2(\alpha), \Delta^2). \\ M_3(\text{vec } \hat{\boldsymbol{\delta}}_n^{PT}) &= \boldsymbol{\Psi}_0 - \boldsymbol{\Psi}_0 \mathbf{J}' H_{p^*+2}(\chi_{p^*}^2(\alpha), \Delta^2) \\ &\quad + \text{vec } \boldsymbol{\eta} (\text{vec } \boldsymbol{\eta})' \{2H_{p^*+2}(\chi_{p^*}^2(\alpha), \Delta^2) - H_{p^*+4}(\chi_{p^*}^2(\alpha), \Delta^2)\}. \\ R_3(\text{vec } \hat{\boldsymbol{\delta}}_n^{PT}, \mathbf{W}) &= \text{tr}(\mathbf{W}\boldsymbol{\Psi}_0) - \text{tr}(\mathbf{W}\boldsymbol{\Psi}_0 \mathbf{J}') H_{p^*+2}(\chi_{p^*}^2(\alpha), \Delta^2) \\ &\quad + (\text{vec } \boldsymbol{\eta})' \mathbf{W} \text{vec } \boldsymbol{\eta} \{2H_{p^*+2}(\chi_{p^*}^2(\alpha), \Delta^2) - H_{p^*+4}(\chi_{p^*}^2(\alpha), \Delta^2)\}. \end{aligned}$$

James-Stein estimator (JSE)

$$\begin{aligned}
b_4(\text{vec } \hat{\boldsymbol{\delta}}_n^{JS}) &= -(p^* - 2) \text{vec } \boldsymbol{\eta} E(\chi_{p^*+2}^{-2}(\Delta^2)). \\
M_4(\text{vec } \hat{\boldsymbol{\delta}}_n^{JS}) &= \boldsymbol{\Psi}_0 - (p^* - 2)\boldsymbol{\Psi}_0\mathbf{J}' \{2E[\chi_{p^*+2}^{-2}(\Delta^2)] - (p^* - 2)E[\chi_{p^*+2}^{-4}(\Delta^2)]\} \\
&\quad + (p^* - 2)(p^* + 2) \text{vec } \boldsymbol{\eta} (\text{vec } \boldsymbol{\eta})' E[\chi_{p^*+4}^{-4}(\Delta^2)]. \\
R_4(\text{vec } \hat{\boldsymbol{\delta}}_n^{JS}, \mathbf{W}) &= \text{tr}(\mathbf{W}\boldsymbol{\Psi}_0) - (p^* - 2)\text{tr}(\mathbf{W}\boldsymbol{\Psi}_0\mathbf{J}') \{2E[\chi_{p^*+2}^{-2}(\Delta^2)] \\
&\quad - (p^* - 2)E[\chi_{p^*+2}^{-4}(\Delta^2)]\} + (p^* - 2)(p^* + 2) (\text{vec } \boldsymbol{\eta})' \mathbf{W} \text{vec } \boldsymbol{\eta} \\
&\quad \times E[\chi_{p^*+4}^{-4}(\Delta^2)].
\end{aligned}$$

Positive rule Stein estimator (PRSE)

$$\begin{aligned}
b_5(\text{vec } \hat{\boldsymbol{\delta}}_n^{s+}) &= -\text{vec } \boldsymbol{\eta} g(\Delta^2), \\
g(\Delta^2) &= (p^* - 2)E[\chi_{p^*+2}^{-2}(\Delta^2)] + E[1 - (p^* - 2)\chi_{p^*+2}^{-2}(\Delta^2)] \\
&\quad \times I[\chi_{p^*+2}^2(\Delta^2) \leq p^* - 2]. \\
M_5(\text{vec } \hat{\boldsymbol{\delta}}_n^{s+}) &= M_4(\text{vec } \hat{\boldsymbol{\delta}}_n^{JS}) - (p^* - 2)\boldsymbol{\Psi}_0\mathbf{J}' \\
&\quad \times \{E[(1 - (p^* - 2)\chi_{p^*+2}^{-2}(\Delta^2))^2 I(\chi_{p^*+2}^2(\Delta^2) \leq p^* - 2)] \\
&\quad + \text{vec } \boldsymbol{\eta} (\text{vec } \boldsymbol{\eta})' \{2E[(1 - (p^* - 2)\chi_{p^*+2}^{-2}(\Delta^2)) I(\chi_{p^*+2}^2(\Delta^2) \leq p^* - 2)] \\
&\quad - E[(1 - (p^* - 2)\chi_{p^*+4}^{-2}(\Delta^2))^2 I(\chi_{p^*+4}^2(\Delta^2) \leq p^* - 2)]\}. \\
R_5(\text{vec } \hat{\boldsymbol{\delta}}_n^{s+}, \mathbf{W}) &= R_4(\text{vec } \hat{\boldsymbol{\delta}}_n^{JS}, \mathbf{W}) - (p^* - 2)\text{tr}(\mathbf{W}\boldsymbol{\Psi}_0\mathbf{J}') \\
&\quad \times E[(1 - (p^* - 2)\chi_{p^*+2}^{-2}(\Delta^2))^2 I(\chi_{p^*+2}^2(\Delta^2) \leq p^* - 2)] \\
&\quad + (\text{vec } \boldsymbol{\eta})' \mathbf{W} \text{vec } \boldsymbol{\eta} \{2E[(1 - (p^* - 2)\chi_{p^*+2}^{-2}(\Delta^2)) I(\chi_{p^*+2}^2(\Delta^2) \leq p^* - 2)] \\
&\quad - E[(1 - (p^* - 2)\chi_{p^*+4}^{-2}(\Delta^2))^2 I(\chi_{p^*+4}^2(\Delta^2) \leq p^* - 2)]\}.
\end{aligned}$$

Alternatively, the analyst may begin with the hypothesis H_{0i} in (4.8), that is, $E(\tilde{\boldsymbol{\delta}}_{n_i}) =$

$\delta_{0i}\mathbf{1}_p$, $i = 1, \dots, k$ and further, the analyst may suspect that

$$H_{00} : \delta_{01} = \dots = \delta_{0k} = \delta_{00} \quad (4.21)$$

may hold. In this case the restricted estimator (RE) of δ_{00} is given by

$$\hat{\delta}_n^* = \hat{\delta}_{00}\mathbf{1}_k = \frac{\mathbf{1}_k\mathbf{1}_k' \hat{\mathbf{V}}_n^{-1}}{\mathbf{1}_k' \hat{\mathbf{V}}_n^{-1} \mathbf{1}_k} \hat{\delta}_n,$$

where

$$\hat{\delta}_n = (\hat{\delta}_{01}, \dots, \hat{\delta}_{0k})',$$

$\hat{\delta}_{0i}$, $i = 1, \dots, k$ is the common estimate of i th experiment defined in (4.9), and

$$\begin{aligned} \hat{\mathbf{V}}_n &= \text{Diag}(\hat{V}_{n_1}, \dots, \hat{V}_{n_k}), \\ \hat{V}_{n_i} &= \frac{1}{\mathbf{1}_p' \hat{\Psi}_{n_i}^{-1} \mathbf{1}_p}, \end{aligned}$$

and $\hat{\Psi}_{n_i}$ is the consistent estimator of Ψ_{n_i} defined in (4.10 - 4.11).

For the test of hypothesis H_{00} , we consider the statistics, $\mathcal{L}_n$ defined by

$$\mathcal{L}_n = n(\hat{\delta}_n - \hat{\delta}_n^*)'(n\hat{\mathbf{V}}_n)^{-1}(\hat{\delta}_n - \hat{\delta}_n^*)$$

Theorem 4.2.4. *Under H_{00} as $\min_{1 \leq i \leq k} n_i \rightarrow \infty$ and $n \rightarrow \infty$ with $\frac{n_i^E}{n_i}$, $\frac{n_i^C}{n_i}$ and $\frac{n_i}{n}$ remain fixed, then $\mathcal{L}_n$ has an asymptotic chi-square distribution with $k - 1$ D.F.*

Meta analysis is complete if H_{00} is accepted with reasonable level of significance.

Quasi-Empirical Bayes Methods of Meta Analysis

We consider the following estimators.

$$\hat{\delta}_n^{PT} = \hat{\delta}_n - (\hat{\delta}_n - \hat{\delta}_n^*)I(\mathcal{L}_n \leq \chi_{k-1}^2(\alpha)).$$

$$\hat{\boldsymbol{\delta}}_n^{JS} = \hat{\boldsymbol{\delta}}_n - (k-3) (\hat{\boldsymbol{\delta}}_n - \hat{\boldsymbol{\delta}}_n^*) \mathcal{L}_n^{-1}.$$

$$\hat{\boldsymbol{\delta}}_n^{s+} = \hat{\boldsymbol{\delta}}_n^* + (1 - (k-3) \mathcal{L}_n^{-1}) I(\mathcal{L}_n \geq k-3) (\hat{\boldsymbol{\delta}}_n - \hat{\boldsymbol{\delta}}_n^*).$$

Thus, we have defined five possible estimators of $\boldsymbol{\delta} = (\delta_{01}, \dots, \delta_{0k})'$. We study their asymptotic properties in the next section.

Asymptotic Distributional Properties of the Estimators

We consider the sequence of local alternatives for the derivation of the asymptotic distribution of the estimators

$$\mathcal{K}_{(n)} : \boldsymbol{\delta}_{(n)} = \delta_{00} \mathbf{1}_k + n^{-1/2} \boldsymbol{\mathcal{T}}, \quad \boldsymbol{\mathcal{T}} \in \mathcal{R}^k, \quad \boldsymbol{\mathcal{T}} \neq \mathbf{0}.$$

Theorem 4.2.5. *Under $\{\mathcal{K}_{(n)}\}$ and the assumed regularity conditions, as $n \rightarrow \infty$, the following hold:*

- (i) $\sqrt{n}(\hat{\boldsymbol{\delta}}_n - \boldsymbol{\delta}_{(n)}) \xrightarrow{\mathcal{D}} \mathbf{Z}_0 \sim N_k(\mathbf{0}, \mathbf{V}_0)$ and $\mathbf{V}_0 = \text{Diag}(V_{01}, \dots, V_{0k})$,
 $V_{0i} = (\mathbf{1}_p' \boldsymbol{\Psi}_{0i}^{-1} \mathbf{1}_p)^{-1}$ and $\boldsymbol{\Psi}_{0i}$ given in (4.5 – 4.7).
- (ii) $\sqrt{n}(\hat{\boldsymbol{\delta}}_n - \hat{\boldsymbol{\delta}}_n^*) \xrightarrow{\mathcal{D}} \mathbf{Z}_1, \mathbf{Z}_1 \sim N_k(\mathbf{J}\boldsymbol{\mathcal{T}}, \mathbf{V}_0 \mathbf{J}'), \mathbf{J} = \mathbf{I}_k - \frac{\mathbf{1}_k \mathbf{1}_k' \mathbf{V}_0^{-1}}{\mathbf{1}_k' \mathbf{V}_0^{-1} \mathbf{1}_k}.$
- (iii) $\sqrt{n}(\hat{\boldsymbol{\delta}}_n^* - \boldsymbol{\delta}_{(n)}) \xrightarrow{\mathcal{D}} \mathbf{Z}_2 - \mathbf{J}\boldsymbol{\mathcal{T}}, \mathbf{Z}_2 \sim N_k(\mathbf{0}, \mathbf{B}), \mathbf{B} = \frac{\mathbf{1}_k \mathbf{1}_k'}{\mathbf{1}_k' \mathbf{V}_0^{-1} \mathbf{1}_k}.$
- (iv) $\begin{pmatrix} \sqrt{n}(\hat{\boldsymbol{\delta}}_n - \boldsymbol{\delta}_{(n)}) \\ \sqrt{n}(\hat{\boldsymbol{\delta}}_n - \hat{\boldsymbol{\delta}}_n^*) \end{pmatrix} \xrightarrow{\mathcal{D}} \begin{pmatrix} \mathbf{Z}_0 \\ \mathbf{Z}_1 \end{pmatrix} \sim N_{2k} \left(\begin{bmatrix} \mathbf{0} \\ \mathbf{J}\boldsymbol{\mathcal{T}} \end{bmatrix}, \begin{bmatrix} \mathbf{V}_0 & \mathbf{V}_0 \mathbf{J}' \\ \mathbf{V}_0 \mathbf{J}' & \mathbf{V}_0 \mathbf{J}' \end{bmatrix} \right).$
- (v) $\begin{pmatrix} \sqrt{n}(\hat{\boldsymbol{\delta}}_n^* - \boldsymbol{\delta}_{(n)}) \\ \sqrt{n}(\hat{\boldsymbol{\delta}}_n - \hat{\boldsymbol{\delta}}_n^*) \end{pmatrix} \xrightarrow{\mathcal{D}} \begin{pmatrix} \mathbf{Z}_2 - \mathbf{J}\boldsymbol{\mathcal{T}} \\ \mathbf{Z}_1 \end{pmatrix} \sim N_{2k} \left(\begin{bmatrix} -\mathbf{J}\boldsymbol{\mathcal{T}} \\ \mathbf{J}\boldsymbol{\mathcal{T}} \end{bmatrix}, \begin{bmatrix} \mathbf{B} & \mathbf{0} \\ \mathbf{0} & \mathbf{V}_0 \mathbf{J}' \end{bmatrix} \right).$
- (vi) $\lim_{n \rightarrow \infty} P\{\mathcal{L}_n \leq x | \mathcal{K}_{(n)}\} = H_{k-1}(x, \Delta^2)$ where $\Delta^2 = \boldsymbol{\eta}' \mathbf{V}_0^{-1} \boldsymbol{\eta}$, $\boldsymbol{\eta} = \mathbf{J}\boldsymbol{\mathcal{T}}$.

where $H_{k-1}(\cdot, \Delta^2)$ is the cdf of noncentral chi-square distribution with noncentral-

ity parameter $\frac{1}{2}\Delta^2$ and $k - 1$ D.F.

Based on Theorem 4.2.5 we obtain the following results.

$$\sqrt{n}(\hat{\boldsymbol{\delta}}_n^{PT} - \boldsymbol{\delta}_{(n)}) \xrightarrow{\mathcal{D}} \mathbf{Z}_0 - \mathbf{Z}_1 I[\mathbf{Z}_1' \mathbf{V}_0^{-1} \mathbf{Z}_1 < \chi_{k-1, \alpha}^2] \quad (4.22)$$

$$\sqrt{n}(\hat{\boldsymbol{\delta}}_n^{JS} - \boldsymbol{\delta}_{(n)}) \xrightarrow{\mathcal{D}} \mathbf{Z}_0 - (k - 3)\mathbf{Z}_1[\mathbf{Z}_1' \mathbf{V}_0^{-1} \mathbf{Z}_1]^{-1} \quad (4.23)$$

$$\begin{aligned} \sqrt{n}(\hat{\boldsymbol{\delta}}_n^{s+} - \boldsymbol{\delta}_{(n)}) &\xrightarrow{\mathcal{D}} \mathbf{Z}_0 - (k - 3)\mathbf{Z}_1[\mathbf{Z}_1' \mathbf{V}_0^{-1} \mathbf{Z}_1]^{-1} \\ &\quad - \mathbf{Z}_1\{1 - (k - 3)[\mathbf{Z}_1' \mathbf{V}_0^{-1} \mathbf{Z}_1]^{-1}\} \\ &\quad \times I[\mathbf{Z}_1' \mathbf{V}_0^{-1} \mathbf{Z}_1 < k - 3]. \end{aligned} \quad (4.24)$$

Define the ADB, ADMSE, and ADQR of the estimators as in section 3.4.2 equations (3.38) - (3.51). Based on Theorem 4.2.5 and the results in (4.22) - (4.24) and using Theorem A.2.6 and A.2.7, we obtain the ADB, ADMSE and ADQR respectively. Expressions of the ADB, ADMSE, and ADQR for the five estimators are as follows:

Unrestricted estimator (UE)

$$b_1(\hat{\boldsymbol{\delta}}_n) = \mathbf{0}.$$

$$M_1(\hat{\boldsymbol{\delta}}_n) = \mathbf{V}_0.$$

$$R_1(\hat{\boldsymbol{\delta}}_n, \mathbf{W}) = \text{tr}(\mathbf{W}\mathbf{V}_0).$$

Restricted estimator (RE)

$$b_2(\hat{\boldsymbol{\delta}}_n^*) = -\boldsymbol{\eta}.$$

$$M_2(\hat{\boldsymbol{\delta}}_n^*) = \mathbf{B} + \boldsymbol{\eta}\boldsymbol{\eta}'.$$

$$R_2(\hat{\boldsymbol{\delta}}_n^*, \mathbf{W}) = \text{tr}(\mathbf{W}\mathbf{B}) + \boldsymbol{\eta}'\mathbf{W}\boldsymbol{\eta}.$$

Preliminary test estimator (PTE)

$$\begin{aligned}
b_3(\hat{\boldsymbol{\delta}}_n^{PT}) &= -\boldsymbol{\eta} H_{k+1}(\chi_{k-1}^2(\alpha), \Delta^2). \\
M_3(\hat{\boldsymbol{\delta}}_n^{PT}) &= \mathbf{V}_0 - \mathbf{V}_0 \mathbf{J}' H_{k+1}(\chi_{k-1}^2(\alpha), \Delta^2) \\
&\quad + \boldsymbol{\eta} \boldsymbol{\eta}' \{2H_{k+1}(\chi_{k-1}^2(\alpha), \Delta^2) - H_{k+3}(\chi_{k-1}^2(\alpha), \Delta^2)\}. \\
R_3(\hat{\boldsymbol{\delta}}_n^{PT}, \mathbf{W}) &= \text{tr}(\mathbf{W} \mathbf{V}_0) - \text{tr}(\mathbf{W} \mathbf{V}_0 \mathbf{J}') H_{k+1}(\chi_{k-1}^2(\alpha), \Delta^2) \\
&\quad + \boldsymbol{\eta}' \mathbf{W} \boldsymbol{\eta} \{2H_{k+1}(\chi_{k-1}^2(\alpha), \Delta^2) - H_{k+3}(\chi_{k-1}^2(\alpha), \Delta^2)\}.
\end{aligned}$$

James-Stein estimator (SE)

$$\begin{aligned}
b_4(\hat{\boldsymbol{\delta}}_n^{JS}) &= -(k-3) \boldsymbol{\eta} E(\chi_{k+1}^{-2}(\Delta^2)). \\
M_4(\hat{\boldsymbol{\delta}}_n^{JS}) &= \mathbf{V}_0 - (k-3) \mathbf{V}_0 \mathbf{J}' \{2E[\chi_{k+1}^{-2}(\Delta^2)] - (k-3)E[\chi_{k+1}^{-4}(\Delta^2)]\} \\
&\quad + (k-3)(k+1) \boldsymbol{\eta} \boldsymbol{\eta}' E[\chi_{k+3}^{-4}(\Delta^2)]. \\
R_4(\hat{\boldsymbol{\delta}}_n^{JS}, \mathbf{W}) &= \text{tr}(\mathbf{W} \mathbf{V}_0) - (k-3) \text{tr}(\mathbf{W} \mathbf{V}_0 \mathbf{J}') \{2E[\chi_{k+1}^{-2}(\Delta^2)] \\
&\quad - (k-3)E[\chi_{k+1}^{-4}(\Delta^2)]\} + (k-3)(k+1) \boldsymbol{\eta}' \mathbf{W} \boldsymbol{\eta} E[\chi_{k+3}^{-4}(\Delta^2)].
\end{aligned}$$

Positive rule Stein estimator (PRSE)

$$\begin{aligned}
b_5(\hat{\boldsymbol{\delta}}_n^{s+}) &= -(k-3) \boldsymbol{\eta} E[\chi_{k+1}^{-2}(\Delta^2)] - \boldsymbol{\eta} E[1 - (k-3) \chi_{k+1}^{-2}(\Delta^2)] \\
&\quad \times I[\chi_{k+1}^2(\Delta^2) \leq k-3]. \\
M_5(\hat{\boldsymbol{\delta}}_n^{s+}) &= M_4(\hat{\boldsymbol{\delta}}_n^{JS}) - (k-3) \mathbf{V}_0 \mathbf{J}' \\
&\quad \times \{E[(1 - (k-3) \chi_{k+1}^{-2}(\Delta^2))^2 I(\chi_{k+1}^2(\Delta^2) \leq k-3)]\} \\
&\quad + \boldsymbol{\eta} \boldsymbol{\eta}' \{2E[(1 - (k-3) \chi_{k+1}^{-2}(\Delta^2)) I(\chi_{k+1}^2(\Delta^2) \leq k-3)] \\
&\quad - E[(1 - (k-3) \chi_{k+3}^{-2}(\Delta^2))^2 I(\chi_{k+3}^2(\Delta^2) \leq k-3)]\}.
\end{aligned}$$

$$\begin{aligned}
R_5(\hat{\delta}_n^{s+}, \mathbf{W}) &= R_4(\hat{\delta}_n^{JS}, \mathbf{W}) - (k-3)tr(\mathbf{W}\mathbf{V}_0\mathbf{J}') \\
&\times E[(1 - (k-3)\chi_{k+1}^{-2}(\Delta^2))^2 I(\chi_{k+1}^2(\Delta^2) \leq k-3)] \\
&+ \boldsymbol{\eta}' \mathbf{W} \boldsymbol{\eta} \{2E[(1 - (k-3)\chi_{k+1}^{-2}(\Delta^2)) I(\chi_{k+1}^2(\Delta^2) \leq k-3)] \\
&- E[(1 - (k-3)\chi_{k+3}^{-2}(\Delta^2))^2 I(\chi_{k+3}^2(\Delta^2) \leq k-3)]\}.
\end{aligned}$$

Bias, MSE and risk expressions of these estimators are similar to these of the estimators given in section 3.4.2 . Therefore the analysis of properties of these estimators proceeds in the same way as the analysis of properties of the estimators given in section 3.4.2 Theorem 3.4.6 and Theorem 3.4.7.

4.2.3 Estimating A Vector of Effect Sizes

Consider the k independent studies giving the estimated effect size vectors $\tilde{\delta}_{n_1}, \dots, \tilde{\delta}_{n_k}$ and corresponding estimated large-sample covariance matrices $\hat{\Psi}_{n_i}$ given by (4.10) - (4.11). We assume that there are no missing observations means that we have data for all p measures (end-points) in each of the k studies. As before, the vectors of effect size estimates from the k studies in the meta-analysis can be combined into one large kp -dimensional vector, and the corresponding estimated covariance matrix, $\hat{\Psi}_n$ will be diagonal blocks of a combined $kp \times kp$ estimated covariance matrices where $\hat{\Psi}_n$ defined in (4.12). The meta-analyst may suspect that all the vectors of effect sizes are equal, that is

$$H_0 : \delta_1 = \dots = \delta_k = \delta \text{ unknown.}$$

Under H_0 , the estimate $\hat{\delta}_n$ of the vector of common effect size is

$$\hat{\delta}_n = (\hat{\Psi}_{n_1}^{-1} + \dots + \hat{\Psi}_{n_k}^{-1})^{-1} (\hat{\Psi}_{n_1}^{-1} \tilde{\delta}_{n_1} + \dots + \hat{\Psi}_{n_k}^{-1} \tilde{\delta}_{n_k}).$$

with the estimated covariance matrix of $\hat{\boldsymbol{\delta}}_n$ given by

$$\mathbf{B} = (\hat{\boldsymbol{\Psi}}_{n_1}^{-1} + \cdots + \hat{\boldsymbol{\Psi}}_{n_k}^{-1})^{-1} \quad (4.25)$$

We may write the large kp -dimensional vector of the common effect size vectors as

$$(vec \hat{\boldsymbol{\delta}}_n^*)' = (\hat{\boldsymbol{\delta}}_n', \dots, \hat{\boldsymbol{\delta}}_n')$$

We can write $vec \hat{\boldsymbol{\delta}}_n^*$ as

$$vec \hat{\boldsymbol{\delta}}_n^* = \left((\mathbf{1}_k \mathbf{1}_k') \otimes \mathbf{B} \right) \hat{\boldsymbol{\Psi}}_n^{-1} vec \tilde{\boldsymbol{\delta}}_n.$$

where $(\mathbf{1}_k \mathbf{1}_k') \otimes \mathbf{B}$ (see definition A.11) is the direct product of the matrix $(\mathbf{1}_k \mathbf{1}_k')$ and $\mathbf{B}$, where $\mathbf{B}$ is given by (4.25), defined as:

$$\mathcal{B} = (\mathbf{1}_k \mathbf{1}_k') \otimes \mathbf{B} = \begin{bmatrix} \mathbf{B} & \mathbf{B} & \cdots & \mathbf{B} \\ \vdots & \vdots & \vdots & \vdots \\ \mathbf{B} & \mathbf{B} & \cdots & \mathbf{B} \end{bmatrix}_{kp \times kp}.$$

For the test of hypothesis H_0 , we consider the statistics, $\mathcal{L}_n$ defined by

$$\mathcal{L}_n = \tilde{\boldsymbol{\delta}}_{n_1}' \hat{\boldsymbol{\Psi}}_{n_1}^{-1} \tilde{\boldsymbol{\delta}}_{n_1} + \cdots + \tilde{\boldsymbol{\delta}}_{n_k}' \hat{\boldsymbol{\Psi}}_{n_k}^{-1} \tilde{\boldsymbol{\delta}}_{n_k} - \hat{\boldsymbol{\delta}}_n' (\hat{\boldsymbol{\Psi}}_{n_1}^{-1} + \cdots + \hat{\boldsymbol{\Psi}}_{n_k}^{-1}) \hat{\boldsymbol{\delta}}_n.$$

which may be rewritten as

$$\mathcal{L}_n = n \left(vec (\tilde{\boldsymbol{\delta}}_n - \hat{\boldsymbol{\delta}}_n^*) \right)' (n \hat{\boldsymbol{\Psi}}_n)^{-1} vec (\tilde{\boldsymbol{\delta}}_n - \hat{\boldsymbol{\delta}}_n^*).$$

$\mathcal{L}_n$ approximately follows a central chi-square distribution with $p^* = p(k-1)$ D.F. under H_0 .

Quasi-Empirical Bayes Estimators

As before, we may define the estimators as

$$\begin{aligned} \text{vec } \hat{\boldsymbol{\delta}}_n^{PT} &= \text{vec } \tilde{\boldsymbol{\delta}}_n - (\text{vec } \tilde{\boldsymbol{\delta}}_n - \text{vec } \hat{\boldsymbol{\delta}}_n^*) I(\mathcal{L}_n \leq \chi_{p^*, \alpha}^2) \\ \text{vec } \hat{\boldsymbol{\delta}}_n^{JS} &= \text{vec } \tilde{\boldsymbol{\delta}}_n - (p^* - 2)(\text{vec } \tilde{\boldsymbol{\delta}}_n - \text{vec } \hat{\boldsymbol{\delta}}_n^*) \mathcal{L}_n^{-1}. \\ \text{vec } \hat{\boldsymbol{\delta}}_n^{s+} &= \text{vec } \hat{\boldsymbol{\delta}}_n^* + [1 - (p^* - 2)\mathcal{L}_n^{-1}] I(\mathcal{L}_n > (p^* - 2)) (\text{vec } \tilde{\boldsymbol{\delta}}_n - \text{vec } \hat{\boldsymbol{\delta}}_n^*). \end{aligned}$$

Asymptotic Distributional Properties of the Estimators

We consider the sequence of local alternatives for the derivation of the asymptotic distribution of the estimators

$$\begin{aligned} \mathcal{K}_{(n)} : \text{vec } \boldsymbol{\delta}_{(n)} &= \text{vec } \boldsymbol{\delta} + n^{-1/2} \text{vec } \boldsymbol{\xi}, \quad (\text{vec } \boldsymbol{\xi})' = (\boldsymbol{\xi}'_1, \dots, \boldsymbol{\xi}'_k) \neq \mathbf{0}, \\ \boldsymbol{\xi}'_i &= (\xi_{i1}, \dots, \xi_{ip}), i = 1, \dots, k. \end{aligned} \quad (4.26)$$

We arrive at the next theorem on the asymptotic distributions of the various statistics

Theorem 4.2.6. *Under $\{\mathcal{K}_{(n)}\}$ and the assumed regularity conditions, as $n \rightarrow \infty$, the following hold:*

- (i) $\sqrt{n}(\text{vec } \tilde{\boldsymbol{\delta}}_n - \text{vec } \boldsymbol{\delta}_{(n)}) \xrightarrow{\mathcal{D}} \text{vec } \mathbf{Z}_0 \sim N_{kp}(\mathbf{0}, \boldsymbol{\Psi}_0)$ and $\boldsymbol{\Psi}_0 = \text{Diag}(\boldsymbol{\Psi}_{01}, \dots, \boldsymbol{\Psi}_{0k})$,
 $\boldsymbol{\Psi}_{0i}, i = 1, \dots, k$, is given by (4.5-4.7).
- (ii) $\sqrt{n}(\text{vec } \tilde{\boldsymbol{\delta}}_n - \text{vec } \hat{\boldsymbol{\delta}}_n^*) \xrightarrow{\mathcal{D}} \text{vec } \mathbf{Z}_1, \text{vec } \mathbf{Z}_1 \sim N_{kp}(\text{vec } \boldsymbol{\eta}, \boldsymbol{\Psi}_0 \mathbf{J}')$,
 $\mathbf{J} = \mathbf{I}_{kp} - [(\mathbf{1}_k \mathbf{1}_k') \otimes \mathbf{B}] \boldsymbol{\Psi}_0^{-1}$, $\text{vec } \boldsymbol{\eta} = \mathbf{J} \text{vec } \boldsymbol{\xi}$.
- (iii) $\sqrt{n}(\text{vec } \hat{\boldsymbol{\delta}}_n^* - \text{vec } \boldsymbol{\delta}_{(n)}) \xrightarrow{\mathcal{D}} \text{vec } \mathbf{Z}_2 - \text{vec } \boldsymbol{\eta}, \text{vec } \mathbf{Z}_2 \sim N_{kp}(\mathbf{0}, (\mathbf{1}_k \mathbf{1}_k') \otimes \mathbf{B})$,
 $\mathbf{B}$ is given by (4.25).
- (iv) $\lim_{n \rightarrow \infty} P\{\mathcal{L}_n \leq x | \mathcal{K}_{(n)}\} = H_{p^*}(x, \Delta^2)$ where $\Delta^2 = (\text{vec } \boldsymbol{\eta})' \boldsymbol{\Psi}_0^{-1} \text{vec } \boldsymbol{\eta}$.

where $H_{p^*}(\cdot, \Delta^2)$ is the cdf of noncentral chi-squared distribution with noncentrality parameter $\frac{1}{2}\Delta^2$ and $p^* = p(k-1)$ D.F.

Based on Theorem 4.2.6 we obtain the following results.

$$\begin{aligned} \sqrt{n} \quad (vec \hat{\boldsymbol{\delta}}_n^{PT} - vec \boldsymbol{\delta}_{(n)}) &\xrightarrow{\mathcal{D}} vec \mathbf{Z}_0 - vec \mathbf{Z}_1 \\ &\times I[(vec \mathbf{Z}_1)' \boldsymbol{\Psi}_0^{-1} vec \mathbf{Z}_1 < \chi_{p^*, \alpha}^2]. \end{aligned} \quad (4.27)$$

$$\begin{aligned} \sqrt{n} \quad (vec \hat{\boldsymbol{\delta}}_n^{JS} - vec \boldsymbol{\delta}_{(n)}) &\xrightarrow{\mathcal{D}} vec \mathbf{Z}_0 - (p^* - 2)vec \mathbf{Z}_1 \\ &\times [(vec \mathbf{Z}_1)' \boldsymbol{\Psi}_0^{-1} vec \mathbf{Z}_1]^{-1}. \end{aligned} \quad (4.28)$$

$$\begin{aligned} \sqrt{n} \quad (vec \hat{\boldsymbol{\delta}}_n^{s+} - vec \boldsymbol{\delta}_{(n)}) &\xrightarrow{\mathcal{D}} vec \mathbf{Z}_0 - (p^* - 2)vec \mathbf{Z}_1 [(vec \mathbf{Z}_1)' \boldsymbol{\Psi}_0^{-1} vec \mathbf{Z}_1]^{-1} \\ &- vec \mathbf{Z}_1 \{1 - (p^* - 2)[(vec \mathbf{Z}_1)' \boldsymbol{\Psi}_0^{-1} vec \mathbf{Z}_1]^{-1}\} \\ &\times I[(vec \mathbf{Z}_1)' \boldsymbol{\Psi}_0^{-1} vec \mathbf{Z}_1 < p^* - 2]. \end{aligned} \quad (4.29)$$

Define the ADB, ADMSE, and ADQR of the estimators as in section 4.2 equations (4.16) - (4.20). Based on Theorem 4.2.6 and the results in (4.27) - (4.29) and using Theorem A.2.6 and A.2.7, we obtain ADB, ADMSE and ADQR expressions respectively. Expressions of the ADB, ADMSE, and ADQR for the five estimators are as follows:

Unrestricted estimator (UE)

$$b_1(vec \tilde{\boldsymbol{\delta}}_n) = \mathbf{0}.$$

$$M_1(vec \tilde{\boldsymbol{\delta}}_n) = \boldsymbol{\Psi}_0.$$

$$R_1(vec \tilde{\boldsymbol{\delta}}_n, \mathbf{W}) = tr(\mathbf{W} \boldsymbol{\Psi}_0).$$

Restricted estimator (RE)

$$\begin{aligned}
 b_2(\text{vec } \hat{\boldsymbol{\delta}}_n^*) &= - \text{vec } \boldsymbol{\eta}. \\
 M_2(\text{vec } \hat{\boldsymbol{\delta}}_n^*) &= \mathcal{B} + \text{vec } \boldsymbol{\eta} (\text{vec } \boldsymbol{\eta})'. \\
 R_2(\text{vec } \hat{\boldsymbol{\delta}}_n^*, \mathbf{W}) &= \text{tr}(\mathbf{W}\mathcal{B}) + (\text{vec } \boldsymbol{\eta})' \mathbf{W} \text{vec } \boldsymbol{\eta}.
 \end{aligned}$$

Preliminary test estimator (PTE)

$$\begin{aligned}
 b_3(\text{vec } \hat{\boldsymbol{\delta}}_n^{PT}) &= - \text{vec } \boldsymbol{\eta} H_{p^*+2}(\chi_{p^*}^2(\alpha), \Delta^2). \\
 M_3(\text{vec } \hat{\boldsymbol{\delta}}_n^{PT}) &= \boldsymbol{\Psi}_0 - \boldsymbol{\Psi}_0 \mathbf{J}' H_{p^*+2}(\chi_{p^*}^2(\alpha), \Delta^2) \\
 &\quad + \text{vec } \boldsymbol{\eta} (\text{vec } \boldsymbol{\eta})' \{2H_{p^*+2}(\chi_{p^*}^2(\alpha), \Delta^2) - H_{p^*+4}(\chi_{p^*}^2(\alpha), \Delta^2)\}. \\
 R_3(\text{vec } \hat{\boldsymbol{\delta}}_n^{PT}, \mathbf{W}) &= \text{tr}(\mathbf{W}\boldsymbol{\Psi}_0) - \text{tr}(\mathbf{W}\boldsymbol{\Psi}_0 \mathbf{J}') H_{p^*+2}(\chi_{p^*}^2(\alpha), \Delta^2) \\
 &\quad + (\text{vec } \boldsymbol{\eta})' \mathbf{W} \text{vec } \boldsymbol{\eta} \{2H_{p^*+2}(\chi_{p^*}^2(\alpha), \Delta^2) - H_{p^*+4}(\chi_{p^*}^2(\alpha), \Delta^2)\}.
 \end{aligned}$$

James-Stein estimator (JSE)

$$\begin{aligned}
 b_4(\text{vec } \hat{\boldsymbol{\delta}}_n^{JS}) &= -(p^* - 2) \text{vec } \boldsymbol{\eta} E(\chi_{p^*+2}^{-2}(\Delta^2)). \\
 M_4(\text{vec } \hat{\boldsymbol{\delta}}_n^{JS}) &= \boldsymbol{\Psi}_0 - (p^* - 2) \boldsymbol{\Psi}_0 \mathbf{J}' \{2E[\chi_{p^*+2}^{-2}(\Delta^2)] - (p^* - 2)E[\chi_{p^*+2}^{-4}(\Delta^2)]\} \\
 &\quad + (p^* - 2)(p^* + 2) \text{vec } \boldsymbol{\eta} (\text{vec } \boldsymbol{\eta})' E[\chi_{p^*+4}^{-4}(\Delta^2)]. \\
 R_4(\text{vec } \hat{\boldsymbol{\delta}}_n^{JS}, \mathbf{W}) &= \text{tr}(\mathbf{W}\boldsymbol{\Psi}_0) - (p^* - 2) \text{tr}(\mathbf{W}\boldsymbol{\Psi}_0 \mathbf{J}') \{2E[\chi_{p^*+2}^{-2}(\Delta^2)] \\
 &\quad - (p^* - 2)E[\chi_{p^*+2}^{-4}(\Delta^2)]\} + (p^* - 2)(p^* + 2) (\text{vec } \boldsymbol{\eta})' \mathbf{W} \text{vec } \boldsymbol{\eta} \\
 &\quad \times E[\chi_{p^*+4}^{-4}(\Delta^2)].
 \end{aligned}$$

Positive rule Stein estimator(PRSE)

$$\begin{aligned}
b_5(\text{vec } \hat{\boldsymbol{\delta}}_n^{s+}) &= -\text{vec } \boldsymbol{\eta} g(\Delta^2), \\
g(\Delta^2) &= (p^* - 2)E[\chi_{p^*+2}^{-2}(\Delta^2)] + E[1 - (p^* - 2)\chi_{p^*+2}^{-2}(\Delta^2)] \\
&\quad \times I[\chi_{p^*+2}^2(\Delta^2) \leq p^* - 2]. \\
M_5(\text{vec } \hat{\boldsymbol{\delta}}_n^{s+}) &= M_4(\text{vec } \hat{\boldsymbol{\delta}}_n^{JS}) - (p^* - 2)\boldsymbol{\Psi}_0\mathbf{J}' \\
&\quad \times \{E[(1 - (p^* - 2)\chi_{p^*+2}^{-2}(\Delta^2))^2 I(\chi_{p^*+2}^2(\Delta^2) \leq p^* - 2)]\} \\
&\quad + \text{vec } \boldsymbol{\eta}(\text{vec } \boldsymbol{\eta})' \{2E[(1 - (p^* - 2)\chi_{p^*+2}^{-2}(\Delta^2))I(\chi_{p^*+2}^2(\Delta^2) \leq p^* - 2)] \\
&\quad - E[(1 - (p^* - 2)\chi_{p^*+4}^{-2}(\Delta^2))^2 I(\chi_{p^*+4}^2(\Delta^2) \leq p^* - 2)]\}. \\
R_5(\text{vec } \hat{\boldsymbol{\delta}}_n^{s+}, \mathbf{W}) &= R_4(\text{vec } \hat{\boldsymbol{\delta}}_n^{JS}, \mathbf{W}) - (p^* - 2)\text{tr}(\mathbf{W}\boldsymbol{\Psi}_0\mathbf{J}') \\
&\quad \times E[(1 - (p^* - 2)\chi_{p^*+2}^{-2}(\Delta^2))^2 I(\chi_{p^*+2}^2(\Delta^2) \leq p^* - 2)] \\
&\quad + (\text{vec } \boldsymbol{\eta})' \mathbf{W} \text{vec } \boldsymbol{\eta} \{2E[(1 - (p^* - 2)\chi_{p^*+2}^{-2}(\Delta^2))I(\chi_{p^*+2}^2(\Delta^2) \leq p^* - 2)] \\
&\quad - E[(1 - (p^* - 2)\chi_{p^*+4}^{-2}(\Delta^2))^2 I(\chi_{p^*+4}^2(\Delta^2) \leq p^* - 2)]\}.
\end{aligned}$$

Analysis of Risk

We consider the ADQR analysis of the five estimators based on the results of the theorems below.

Theorem 4.2.7. *Let $p^* > 2$ and let the matrix $\mathbf{W}$ in the loss function,*

$$L(\text{vec } \boldsymbol{\delta}_n, \mathbf{W}) = n(\text{vec } \boldsymbol{\delta}_n - \text{vec } \boldsymbol{\delta})\mathbf{W}(\text{vec } \boldsymbol{\delta}_n - \text{vec } \boldsymbol{\delta})'$$

be positive semi-definite, satisfying

$$\{\mathbf{W} : \frac{\text{tr}(\mathbf{W}\boldsymbol{\Psi}_0\mathbf{J}')}{\text{ch}_{\max}(\mathbf{W}\boldsymbol{\Psi}_0)} \geq \frac{p^* + 2}{2}\}.$$

where $\text{ch}_{\max}(\mathbf{W}\boldsymbol{\Psi}_0)$ stands for the largest characteristic root of $\mathbf{W}\boldsymbol{\Psi}_0$. Then, under

$\{\mathcal{K}_{(n)}\}$, the ADQR of $\text{vec } \hat{\boldsymbol{\delta}}_n^{s+}$, $\text{vec } \hat{\boldsymbol{\delta}}_n^{JS}$ and $\text{vec } \tilde{\boldsymbol{\delta}}_n$ can be ordered as :

$$R_5(\text{vec } \hat{\boldsymbol{\delta}}_n^{s+}, \mathbf{W}) \leq R_4(\text{vec } \hat{\boldsymbol{\delta}}_n^{JS}, \mathbf{W}) \leq R_1(\text{vec } \tilde{\boldsymbol{\delta}}_n, \mathbf{W}).$$

uniformly in Δ^2 .

Theorem 4.2.8. Under H_0 , the positive semi definite matrix $\mathbf{W}$ in the loss function, the ADQR of estimators $\text{vec } \tilde{\boldsymbol{\delta}}_n$, $\text{vec } \hat{\boldsymbol{\delta}}_n$, $\text{vec } \hat{\boldsymbol{\delta}}_n^{PT}$, $\text{vec } \hat{\boldsymbol{\delta}}_n^{JS}$ and $\text{vec } \hat{\boldsymbol{\delta}}_n^{s+}$ can be ordered as follows:

$$\begin{aligned} R_2(\text{vec } \hat{\boldsymbol{\delta}}_n, \mathbf{W}) &\leq R_5(\text{vec } \hat{\boldsymbol{\delta}}_n^{s+}, \mathbf{W}) \leq R_4(\text{vec } \hat{\boldsymbol{\delta}}_n^{JS}, \mathbf{W}) \leq R_1(\text{vec } \tilde{\boldsymbol{\delta}}_n, \mathbf{W}), \\ R_2(\text{vec } \hat{\boldsymbol{\delta}}_n, \mathbf{W}) &\leq R_3(\text{vec } \hat{\boldsymbol{\delta}}_n^{PT}, \mathbf{W}) \leq R_1(\text{vec } \tilde{\boldsymbol{\delta}}_n, \mathbf{W}), \\ R_3(\text{vec } \hat{\boldsymbol{\delta}}_n^{PT}, \mathbf{W}) &\leq R_5(\text{vec } \hat{\boldsymbol{\delta}}_n^{s+}, \mathbf{W}) \leq R_4(\text{vec } \hat{\boldsymbol{\delta}}_n^{JS}, \mathbf{W}) \leq R_1(\text{vec } \tilde{\boldsymbol{\delta}}_n, \mathbf{W}), \end{aligned}$$

where the order depends on the level of significance, α for the PTE, $\hat{\boldsymbol{\delta}}_n^{PT}$. Proofs of these Theorems are very similar to the proofs of Theorem 3.4.6 and Theorem 3.4.7.

Chapter 5

Robust Estimates of “Effect Size”

In this chapter, we discuss the robust estimation procedures of effect size. It is well known that the sample mean and sample variance are non-robust estimators as they are very sensitive to outliers, therefore, the usual estimator of effect size, namely,

$$\tilde{\delta}_n = c(\nu) \frac{\bar{Y}^E - \bar{Y}^C}{s_p}.$$

is non-robust. However, there are a few nonparametric estimators of effect size in the literature, the first one is due to Kraemer and Andrews (1982) based on quantile-function which have been extended by Hedges and Olkin (1984). Further, Hedges and Olkin (1985) suggested the following estimator

$$\hat{\delta}_n^{(M)} = [\text{Med}(Y^E) - \text{Med}(Y^C)]/\tilde{\sigma},$$

where $\text{Med}(Y^E)$ and $\text{Med}(Y^C)$ are the medians of the observations from the experimental and control groups respectively, and $\tilde{\sigma}$ is the estimator of σ based on the trimmed estimator

$$\tilde{\sigma} = a_2 Y_{(2)}^C + \cdots + a_{(n^C-1)} Y_{(n^C-1)}^C,$$

where $Y_{(1)}^C \leq Y_{(2)}^C \leq \cdots \leq Y_{(n^C-1)}^C \leq Y_{(n^C)}^C$ are the order statistics of observations from the control group and $a_2, \dots, a_{(n^C-1)}$ are the optimal coefficients which minimize the variance $\tilde{\sigma}$ subject to $\tilde{\sigma} > 0$. Krauth (1983) followed up Kraemer and Andrews but did not discuss the distributional characteristics of this estimators. Latter, Zhang and Schoeps (1997) proposed a simplified estimator of effect size, namely,

$$\hat{\delta} = \frac{[\text{Med}(Y^E) - \text{Med}(Y^C)]}{Y_{(r_2)}^C - Y_{(r_1)}^C},$$

and $r_1 = [n^C \lambda_1] + 1$, $r_2 = [n^C \lambda_2] + 1$, $0 < \lambda_1 < \lambda_2 < 1$ where $Y_{(r_1)}^C, Y_{(r_2)}^C$ are the r_1^{th} and r_2^{th} order statistics from the control group. They studied the asymptotic properties of this estimator.

In this chapter, we shall discuss the procedures based on the theory of L-estimation and some R-estimators of effect size where one uses both the experimental and control groups scores measurements. First, we consider the general theory of estimation of effect size based on linear combination of order statistics (L-estimates) in section 5.1 and 5.2, and in section 5.3 we present the nonparametric methods based on sign test and rank theory (R-estimates). We conclude this chapter with the nonparametric theory for the multivariate normal models.

5.1 General Theory of Estimation of Effect Size Based on Order Statistics (L-estimates)

We consider the estimation of “effect size”, $\delta = \frac{\mu^E - \mu^C}{\sigma}$, when we have samples from the location-scale models defined by the continuous cdf $F(\frac{y - \mu^E}{\sigma})$ and $F(\frac{y - \mu^C}{\sigma})$ with the probability density functions given by $\frac{1}{\sigma} f(\frac{y - \mu^E}{\sigma})$ and $\frac{1}{\sigma} f(\frac{y - \mu^C}{\sigma})$ respectively.

To begin with, we consider method of estimation based on all the order statistics from one distribution. Let $Y_1, \dots, Y_n$ be a random sample from $F(\frac{y - \mu}{\sigma})$. Then, for the

order statistics $Y_{(1)} \leq Y_{(2)} \leq \cdots \leq Y_{(n)}$ of the sample, we have linear relation, $Y_{(j)} = \mu + \sigma Z_{(j)}$, $j = 1, \cdots, n$. where $Z_{(j)}$ is the j^{th} order statistics from the standardized population $F(z)$. Consequently, the expected value of the actual ordered observations $Y_{(j)}$ is given by

$$\begin{aligned} E(Y_{(j)}) &= \mu + \sigma E(Z_{(j)}), \\ &= \mu + \sigma \alpha_j \quad (\text{say}). \end{aligned}$$

Similarly, the covariances are given by

$$\text{Cov}(Y_{(i)}, Y_{(j)}) = \sigma^2 E(Z_{(i)} Z_{(j)})$$

The expectations, variance, and covariances of the $Z_{(j)}$ may be calculated from the distribution of the ordered observations on Z (standardized variate). Thus, we may define the covariance matrix of the order statistics as

$$\mathbf{\Omega}_n = \left(\left(E(Z_{(i)} Z_{(j)}) \right) \right)_{n \times n}.$$

In order to estimate μ and σ simultaneously we minimize the quadratic form:

$$Q(\mu, \sigma) = (\mathbf{Y}_{(.)} - \mu \mathbf{1}_n - \sigma \boldsymbol{\alpha}_n)' \mathbf{\Omega}_n^{-1} (\mathbf{Y}_{(.)} - \mu \mathbf{1}_n - \sigma \boldsymbol{\alpha}_n),$$

$$\mathbf{Y}_{(.)} = (Y_{(1)}, \cdots, Y_{(n)})', \quad \boldsymbol{\alpha}_n = (\alpha_1, \cdots, \alpha_n)' \text{ and } \mathbf{1}_n = (1, \cdots, 1)',$$

so that we have the normal equations

$$\mathbf{1}_n' \mathbf{\Omega}_n^{-1} \mathbf{Y}_{(.)} = \mu \mathbf{1}_n' \mathbf{\Omega}_n^{-1} \mathbf{1}_n + \sigma \mathbf{1}_n' \mathbf{\Omega}_n^{-1} \boldsymbol{\alpha}_n$$

$$\boldsymbol{\alpha}_n' \mathbf{\Omega}_n^{-1} \mathbf{Y}_{(.)} = \mu \boldsymbol{\alpha}_n' \mathbf{\Omega}_n^{-1} \mathbf{1}_n + \sigma \boldsymbol{\alpha}_n' \mathbf{\Omega}_n^{-1} \boldsymbol{\alpha}_n.$$

The “best linear unbiased estimator” (BLUE) of (μ, σ) is given by

$$\begin{bmatrix} \tilde{\mu} \\ \tilde{\sigma} \end{bmatrix} = \begin{bmatrix} K_1 & K_3 \\ K_3 & K_2 \end{bmatrix}^{-1} \begin{bmatrix} \mathbf{1}_n' \boldsymbol{\Omega}_n^{-1} \mathbf{Y}_{(.)} \\ \boldsymbol{\alpha}_n' \boldsymbol{\Omega}_n^{-1} \mathbf{Y}_{(.)} \end{bmatrix}$$

where $\tilde{\sigma} > 0$, $K_1 = \mathbf{1}_n' \boldsymbol{\Omega}_n^{-1} \mathbf{1}_n$, $K_2 = \boldsymbol{\alpha}_n' \boldsymbol{\Omega}_n^{-1} \boldsymbol{\alpha}_n$ and $K_3 = \mathbf{1}_n' \boldsymbol{\Omega}_n^{-1} \boldsymbol{\alpha}_n$.

Then the BLUE may be rewritten as

$$\begin{bmatrix} \tilde{\mu} \\ \tilde{\sigma} \end{bmatrix} = \frac{1}{K_1 K_2 - K_3^2} \begin{bmatrix} K_2 & -K_3 \\ -K_3 & K_1 \end{bmatrix} \begin{bmatrix} \mathbf{1}_n' \boldsymbol{\Omega}_n^{-1} \mathbf{Y}_{(.)} \\ \boldsymbol{\alpha}_n' \boldsymbol{\Omega}_n^{-1} \mathbf{Y}_{(.)} \end{bmatrix}.$$

The explicit expression for the estimators may be written as linear combination of the order statistics given by

$$\begin{bmatrix} \tilde{\mu} \\ \tilde{\sigma} \end{bmatrix} = \begin{bmatrix} a_1 Y_{(1)} + \cdots + a_n Y_{(n)} \\ b_1 Y_{(1)} + \cdots + b_n Y_{(n)} \end{bmatrix}$$

$$\text{where } \sum_{i=1}^n a_i = 1, \sum_{i=1}^n a_i \alpha_i = 0 \text{ and } \sum_{i=1}^n b_i = 0, \sum_{i=1}^n b_i \alpha_i = 1,$$

due to unbiasedness condition of the estimators. The coefficients may be determined by numerical computation and some are available in Sarhan and Greenberg (1962), and David (1981), among others.

Jung (1955) and Gastwirth, Chernoff and Jones (1967) with additional conditions

and squared integrable function, have shown that as $n \rightarrow \infty$

$$\begin{bmatrix} \sqrt{n} (\tilde{\mu} - \mu) \\ \sqrt{n} (\tilde{\sigma} - \sigma) \end{bmatrix} \sim N_2 \left(\begin{bmatrix} 0 \\ 0 \end{bmatrix}, \sigma^2 \begin{bmatrix} I_{11} & I_{12} \\ I_{21} & I_{22} \end{bmatrix}^{-1} \right)$$

$$\text{where } I_{11} = \lim_{n \rightarrow \infty} K_1 = \int_{-\infty}^{\infty} \left[-\frac{f'(x)}{f(x)} \right]^2 f(x) dx,$$

$$I_{22} = \lim_{n \rightarrow \infty} K_2 = \int_{-\infty}^{\infty} \left[-1 - x \frac{f'(x)}{f(x)} \right]^2 f(x) dx,$$

$$\text{and } I_{12} = I_{21} = \lim_{n \rightarrow \infty} K_3 = \int_{-\infty}^{\infty} \left[-\frac{f'(x)}{f(x)} \right] \left[-1 - x \frac{f'(x)}{f(x)} \right] f(x) dx$$

are the elements of the Fisher information matrix, $\begin{bmatrix} I_{11} & I_{12} \\ I_{21} & I_{22} \end{bmatrix}$. Also, $\tilde{\mu} \xrightarrow{P} \mu$ and $\tilde{\sigma} \xrightarrow{P} \sigma$ as $n \rightarrow \infty$ i.e. $\tilde{\mu}$ and $\tilde{\sigma}$ are consistent.

Now, we consider a sample of size n^E from the experimental group distribution, $F(\frac{y-\mu^E}{\sigma})$ resulting in the order statistics $Y_{(1)}^E \leq Y_{(2)}^E \leq \dots \leq Y_{(n^E)}^E$. Similarly, we consider a sample of size n^C from the control group distribution, $F(\frac{y-\mu^C}{\sigma})$ resulting in the order statistics $Y_{(1)}^C \leq Y_{(2)}^C \leq \dots \leq Y_{(n^C)}^C$. Then, following the steps above, we have the BLUE of μ^E , μ^C and σ as given by

$$\tilde{\mu}_{n^E}^E = a_1^E Y_{(1)}^E + \dots + a_{n^E}^E Y_{(n^E)}^E, \quad \tilde{\sigma}_{n^E}^E = b_1^E Y_{(1)}^E + \dots + b_{n^E}^E Y_{(n^E)}^E$$

$$\tilde{\mu}_{n^C}^C = a_1^C Y_{(1)}^C + \dots + a_{n^C}^C Y_{(n^C)}^C, \quad \tilde{\sigma}_{n^C}^C = b_1^C Y_{(1)}^C + \dots + b_{n^C}^C Y_{(n^C)}^C$$

and pooled estimate of σ is given by

$$\tilde{\sigma}_n^{(p)} = \frac{n^E}{n^E + n^C} (b_1^E Y_{(1)}^E + \dots + b_{n^E}^E Y_{(n^E)}^E) + \frac{n^C}{n^E + n^C} (b_1^C Y_{(1)}^C + \dots + b_{n^C}^C Y_{(n^C)}^C)$$

We may obtain the following results: As $n \rightarrow \infty$ such that $\frac{n^E}{n} \rightarrow \pi^E$ and $\frac{n^C}{n} \rightarrow \pi^C$, then,

$$(a) \quad \begin{bmatrix} \sqrt{n}(\tilde{\mu}_{n^E}^E - \mu^E) \\ \sqrt{n}(\tilde{\sigma}_{n^E}^E - \sigma) \end{bmatrix} \sim N_2 \left(\begin{bmatrix} 0 \\ 0 \end{bmatrix}, \frac{\sigma^2}{\pi^E \Delta} \begin{bmatrix} I_{22} & -I_{12} \\ -I_{21} & I_{11} \end{bmatrix} \right),$$

$$(b) \quad \begin{bmatrix} \sqrt{n}(\tilde{\mu}_{n^C}^C - \mu^C) \\ \sqrt{n}(\tilde{\sigma}_{n^C}^C - \sigma) \end{bmatrix} \sim N_2 \left(\begin{bmatrix} 0 \\ 0 \end{bmatrix}, \frac{\sigma^2}{\pi^C \Delta} \begin{bmatrix} I_{22} & -I_{12} \\ -I_{21} & I_{11} \end{bmatrix} \right),$$

$$(c) \quad \begin{bmatrix} \sqrt{n}(\tilde{\mu}_{n^E}^E - \mu^E) \\ \sqrt{n}(\tilde{\mu}_{n^C}^C - \mu^C) \\ \sqrt{n}(\tilde{\sigma}_n^{(p)} - \sigma) \end{bmatrix} \sim N_3 \left(\begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}, \frac{\sigma^2}{\Delta} \begin{bmatrix} \frac{I_{22}}{\pi^E} & 0 & -\frac{I_{12}}{\pi^E} \\ 0 & \frac{I_{22}}{\pi^C} & -\frac{I_{21}}{\pi^C} \\ -\frac{I_{21}}{\pi^E} & -\frac{I_{12}}{\pi^C} & I_{11} \end{bmatrix} \right),$$

where $\Delta = I_{11}I_{22} - I_{12}^2$.

Define the estimate of the effect size δ using these BLUE's of μ^E , μ^C and σ as:

$$\tilde{\delta}_n = \frac{\tilde{\mu}_{n^E}^E - \tilde{\mu}_{n^C}^C}{\tilde{\sigma}_n^{(p)}}. \quad (5.1)$$

Note that as $n \rightarrow \infty$, $\tilde{\sigma}_n^{(p)} \xrightarrow{P} \sigma$, then it is easy to see that

$$\left(\frac{n}{n^E} + \frac{n}{n^C} \right)^{-\frac{1}{2}} \sqrt{n} (\tilde{\delta}_n - \delta) \sim N \left(0, \frac{I_{22}}{I_{11}I_{22} - I_{12}^2} \right).$$

5.2 Robust Estimation of “Effect Size” Based on Sample Quantiles

It is well known that the estimators described in the previous section, are severely affected if the sample contains “outliers” as such some methods must be adopted to weed out the effects of the presence of “outliers” . One such method is to estimate the effect size based on a few selected order statistics. We describe this procedure now when we are sampling from the cdf $F\left(\frac{x-\mu}{\sigma}\right)$.

Let n be large and consider k spacings $0 < \lambda_1 < \cdots < \lambda_k < 1$ which define the ranks of the order statistics as $r_j = [n\lambda_j] + 1, j = 1 \cdots, k$. Then, as $n \rightarrow \infty$, the joint distribution of $Y_{(r_1)}, \cdots, Y_{(r_k)}$ converges in distribution to a k -variate normal distribution with mean $\mu \mathbf{1}_n + \sigma \boldsymbol{\alpha}_k$, $\boldsymbol{\alpha}_k = (F^{-1}(\lambda_1), \cdots, F^{-1}(\lambda_k))'$, where $F^{-1}(\lambda_j) = \inf\{x : F(x) \geq \lambda_j\}$ and covariance matrix $\frac{\sigma^2}{n} \boldsymbol{\Omega}_k$ where $\boldsymbol{\Omega}_k$ is given by

$$\left(\left(\frac{\lambda_i(1-\lambda_j)}{f(F^{-1}(\lambda_i)) f(F^{-1}(\lambda_j))} \right) \right)_{k \times k}.$$

Thus, using generalized least-squares methods we get the asymptotic BLUE (ABLUE) of μ, σ as

$$\begin{bmatrix} \hat{\mu}_k \\ \hat{\sigma}_k \end{bmatrix} = \begin{bmatrix} K_1^* & K_3^* \\ K_3^* & K_2^* \end{bmatrix}^{-1} \begin{bmatrix} \mathbf{1}_k' \boldsymbol{\Omega}_k^{-1} \mathbf{Y}_{(\cdot)}^{(k)} \\ \boldsymbol{\alpha}_k' \boldsymbol{\Omega}_k^{-1} \mathbf{Y}_{(\cdot)}^{(k)} \end{bmatrix}$$

where $K_1^* = \mathbf{1}_k' \boldsymbol{\Omega}_k^{-1} \mathbf{1}_k$, $K_2^* = \boldsymbol{\alpha}_k' \boldsymbol{\Omega}_k^{-1} \boldsymbol{\alpha}_k$ and $K_3^* = \mathbf{1}_k' \boldsymbol{\Omega}_k^{-1} \boldsymbol{\alpha}_k$,

$$\mathbf{Y}_{(\cdot)}^{(k)} = (Y_{(r_1)}, \cdots, Y_{(r_k)})', \quad \boldsymbol{\alpha}_k = (F^{-1}(\lambda_1), \cdots, F^{-1}(\lambda_k))$$

The covariance matrix of $(\hat{\mu}_k, \hat{\sigma}_k)$ is given by

$$\frac{\sigma^2}{n} \begin{bmatrix} K_1^* & K_3^* \\ K_3^* & K_2^* \end{bmatrix}^{-1}$$

$$K_1^* = \sum_{j=0}^k \frac{[f(F^{-1}(\lambda_{j+1})) - f(F^{-1}(\lambda_j))]^2}{\lambda_{j+1} - \lambda_j}$$

$$K_2^* = \sum_{j=0}^k \frac{[F^{-1}(\lambda_{j+1}))f(F^{-1}(\lambda_{j+1})) - F^{-1}(\lambda_j)f(F^{-1}(\lambda_j))]^2}{\lambda_{j+1} - \lambda_j}$$

$$K_3^* = \sum_{j=0}^k \frac{[f(F^{-1}(\lambda_{j+1})) - f(F^{-1}(\lambda_j))][F^{-1}(\lambda_{j+1})f(F^{-1}(\lambda_{j+1})) - F^{-1}(\lambda_j)f(F^{-1}(\lambda_j))]}{\lambda_{j+1} - \lambda_j},$$

assuming $\lambda_0 = 0, \lambda_{k+1} = 1$.

Now, consider the estimation of the “effect size” based on the k sample quantiles, $Y_{(r_1)}^E, \dots, Y_{(r_k)}^E$ and $Y_{(r_1)}^C, \dots, Y_{(r_k)}^C$ from the experimental and control distributions, given in section 5.1, where $r_j = [n^* \lambda_j] + 1$, with $n^* = \min(n^E, n^C)$, $j = 1, \dots, k$ and fixed spacings $0 < \lambda_1 < \dots < \lambda_k < 1$. Then, the asymptotically best linear unbiased estimators (ABLUE) of μ^E, μ^C and σ may be written as:

$$\hat{\mu}_k^E = a_1 Y_{(r_1)}^E + \dots + a_k Y_{(r_k)}^E, \quad \hat{\sigma}_k^E = b_1 Y_{(r_1)}^E + \dots + b_k Y_{(r_k)}^E$$

$$\hat{\mu}_k^C = a_1 Y_{(r_1)}^C + \dots + a_k Y_{(r_k)}^C, \quad \hat{\sigma}_k^C = b_1 Y_{(r_1)}^C + \dots + b_k Y_{(r_k)}^C$$

$$\hat{\sigma}_k^{(p)} = \frac{n^E}{n^E + n^C} \hat{\sigma}_k^E + \frac{n^C}{n^E + n^C} \hat{\sigma}_k^C.$$

We may obtain the following results: As $n \rightarrow \infty$ such that $\frac{n^E}{n} \rightarrow \pi^E$ and $\frac{n^C}{n} \rightarrow \pi^C$,

then,

$$(a) \quad \begin{bmatrix} \sqrt{n}(\hat{\mu}_k^E - \mu^E) \\ \sqrt{n}(\hat{\sigma}_k^E - \sigma) \end{bmatrix} \sim N_2 \left(\begin{bmatrix} 0 \\ 0 \end{bmatrix}, \frac{\sigma^2}{\pi^E \Delta} \begin{bmatrix} K_2^* & -K_3^* \\ -K_3^* & K_1^* \end{bmatrix} \right),$$

$$(b) \quad \begin{bmatrix} \sqrt{n}(\hat{\mu}_k^C - \mu^C) \\ \sqrt{n}(\hat{\sigma}_k^C - \sigma) \end{bmatrix} \sim N_2 \left(\begin{bmatrix} 0 \\ 0 \end{bmatrix}, \frac{\sigma^2}{\pi^C \Delta} \begin{bmatrix} K_2^* & -K_3^* \\ K_3^* & K_1^* \end{bmatrix} \right),$$

$$(c) \quad \begin{bmatrix} \sqrt{n}(\hat{\mu}_k^E - \mu^E) \\ \sqrt{n}(\hat{\mu}_k^C - \mu^C) \\ \sqrt{n}(\hat{\sigma}_k^{(p)} - \sigma) \end{bmatrix} \sim N_3 \left(\begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}, \frac{\sigma^2}{\Delta} \begin{bmatrix} \frac{K_2^*}{\pi^E} & 0 & -\frac{K_3^*}{\pi^E} \\ 0 & \frac{K_2^*}{\pi^C} & -\frac{K_3^*}{\pi^C} \\ -\frac{K_3^*}{\pi^E} & -\frac{K_3^*}{\pi^C} & K_1^* \end{bmatrix} \right),$$

where $\Delta = K_1^* K_2^* - K_3^{*2}$.

We define the estimator of $\delta = \frac{\mu^E - \mu^C}{\sigma}$ as

$$\hat{\delta}_n = \frac{\hat{\mu}_k^E - \hat{\mu}_k^C}{\hat{\sigma}_k^{(p)}}.$$

It is known, as $n \rightarrow \infty$ $\hat{\sigma}_k^{(p)} \xrightarrow{P} \sigma$. Then, as $n \rightarrow \infty$, we have

$$\left(\frac{n}{n^E} + \frac{n}{n^C} \right)^{-\frac{1}{2}} \sqrt{n} (\hat{\delta}_n - \delta) \sim N \left(0, \frac{K_2^*}{\Delta} \right).$$

Hence, ARE of $\hat{\delta}_n$ relative to “complete sample” estimator in (5.1) is given by

$$\text{ARE}(\hat{\delta}_n : \tilde{\delta}_n) = \left[\left(\frac{I_{22}}{I_{11}I_{22} - I_{12}^2} \right) / \left(\frac{K_2^*}{K_1^*K_2^* - K_3^{*2}} \right) \right].$$

If F is normal cdf, then $I_{11} = 1$, $I_{22} = 2$, and $I_{12} = I_{21} = 0$ and spacings are symmetric then the ARE is given by

$$K_N = \sum_{j=0}^k \frac{[\phi(\Phi^{-1}(\lambda_{j+1})) - \phi(\Phi^{-1}(\lambda_j))]^2}{\lambda_{j+1} - \lambda_j}, \quad \lambda_0 = 0, \quad \lambda_{k+1} = 1.$$

Some tabular values of ARE for $k = 1, \dots, 10$ are given in Sarhan and Greenberg page 280 and are displayed in the following table.

Table 5.1: ARE for normal distribution

k	1	2	3	4	5	6	7	8	9	10
ARE	$\frac{2}{\pi}$	0.8097	0.8800	0.9342	0.9420	0.9559	0.9654	0.9722	0.9771	0.9808

If F is Cauchy, then $I_{11} = I_{22} = \frac{1}{2}$, and $I_{12} = 0$ and ARE is $2K_c$, where

$$K_c = \sum_{j=0}^k \frac{[f_c(F_c^{-1}(\lambda_{j+1})) - f_c(F_c^{-1}(\lambda_j))]^2}{\lambda_{j+1} - \lambda_j}, \quad \lambda_0 = 0, \quad \lambda_{k+1} = 1.$$

where f_c and F_c are the pdf and cdf of Cauchy distribution. The following table is reproduced from Saleh et al (1985) and it gives the ARE for the corresponding optimum spacings

$$\lambda_j = \frac{j}{k+1}, j = 1, \dots, k$$

for $k = 1, 2, 3, 4, 5$.

Table 5.2: ARE and corresponding optimum spacings for Cauchy distribution

k	1	2	3	4	5
ARE	0.8106	0.9226	0.9582	0.9739	0.9821

5.3 Nonparametric Methods of “Effect Size” Estimation

Let $Y_1^E, \dots, Y_{n^E}^E$ be a sample of size n^E from the experimental distribution, $F(\frac{y-\mu^E}{\sigma})$ and $Y_1^C, \dots, Y_{n^C}^C$ be another sample of size n^C from the control distribution, $F(\frac{y-\mu^C}{\sigma})$.

We wish to estimate the effect size based on the two samples. We define the sign tests for the hypothesis

$$H_{01} : \mu^E = 0 \text{ Vs } H_{A1} : \mu^E \neq 0.$$

$$H_{02} : \mu^C = 0 \text{ Vs } H_{A2} : \mu^C \neq 0.$$

Assuming symmetry of the distributions about zero. The sign tests are defined by:

$$S_{n^E} = \sum_{j=1}^{n^E} \text{sign}(Y_j^E - \mu^E) \text{ and } S_{n^C} = \sum_{j=1}^{n^C} \text{sign}(Y_j^C - \mu^C).$$

Setting both statistics to zero, the estimators of μ^E and μ^C become the median of Y^E 's and median of Y^C 's respectively. Thus, defining the effect size estimate as

$$\hat{\delta}_n^{(M)} = \frac{[\text{Med}_{1 \leq j \leq n^E}(Y_j^E) - \text{Med}_{1 \leq j \leq n^C}(Y_j^C)]}{\text{MAD}_n},$$

where $\text{MAD}_n = \frac{1}{0.6745} \left[\frac{n^E}{n^E + n^C} \text{Med}_{1 \leq j \leq n^E} \{|(Y_j^E - \text{Med}(Y^E))|\} \right.$

$$\left. + \frac{n^C}{n^E + n^C} \text{Med}_{1 \leq j \leq n^C} \{|(Y_j^C - \text{Med}(Y^C))|\} \right]$$

is a consistent estimate of σ , (see Huber, 1981). Hence, one may show that as $n \rightarrow \infty$ such that $\frac{n^E}{n} \rightarrow \pi^E$ and $\frac{n^C}{n} \rightarrow \pi^C$

$$\left(\frac{n}{n^E} + \frac{n}{n^C}\right)^{\frac{-1}{2}} \sqrt{n} (\hat{\delta}_n^{(M)} - \delta) \sim N\left(0, \frac{1}{4f^2(0)}\right).$$

Similarly, for $\hat{\delta}_n^{(\text{mean})} = \frac{Y^E - \bar{Y}^C}{s_p}$, we obtain

$$\left(\frac{n}{n^E} + \frac{n}{n^C}\right)^{\frac{-1}{2}} \sqrt{n} (\hat{\delta}_n^{(\text{mean})} - \delta) \sim N(0, 1) \quad \text{as } n \rightarrow \infty.$$

$$\begin{aligned} \text{Therefore, the ARE}(\hat{\delta}_n^{(M)} : \hat{\delta}_n^{(\text{mean})}) = 4f^2(0) &= \frac{2}{\pi} \text{ if } f \text{ is normal density.} \\ &= \frac{4}{\pi^2} \text{ if } f \text{ is Cauchy density.} \\ &= 1 \text{ if } f \text{ is double-exponential density.} \end{aligned}$$

Now, consider the Wilcoxon signed rank test

$$T_{n^E} = \sum_{j=1}^{n^E} R_{n^E j}^+ \text{sign}(Y_j^E - \mu^E) \text{ and } T_{n^C} = \sum_{j=1}^{n^C} R_{n^C j}^+ \text{sign}(Y_j^C - \mu^C)$$

where $R_{n^E j}^+$ is the rank of $|Y_j^E - \mu^E|$ among $|Y_1^E - \mu^E|, \dots, |Y_{n^E}^E - \mu^E|$ and $R_{n^C j}^+$ is the rank of $|Y_j^C - \mu^C|$ among $|Y_1^C - \mu^C|, \dots, |Y_{n^C}^C - \mu^C|$. Setting $T_{n^E} = T_{n^C} = 0$, we obtain the Hodges and Lehmann (1963) solution given by

$$\mu_{HL}^E = \text{Med}_{1 \leq i \leq j \leq n^E} \frac{1}{2}(Y_i^E + Y_j^E) \text{ and } \mu_{HL}^C = \text{Med}_{1 \leq i \leq j \leq n^C} \frac{1}{2}(Y_i^C + Y_j^C)$$

Now, define the estimator of δ , R-estimate, as

$$\hat{\delta}_n^{HL} = \frac{\mu_{HL}^E - \mu_{HL}^C}{\text{MAD}_n}$$

as a robust estimation of δ . The asymptotic distribution of $\hat{\delta}_n^{HL}$ is given by

$$\left(\frac{n}{n^E} + \frac{n}{n^C}\right)^{-\frac{1}{2}} \sqrt{n} (\hat{\delta}_n^{(HL)} - \delta) \sim N\left(0, [12\left(\int_{-\infty}^{\infty} f^2(x)dx\right)^2]^{-1}\right).$$

$$\begin{aligned} \text{Hence, the ARE}(\hat{\delta}_n^{(HL)} : \tilde{\delta}_n^{(\text{mean})}) &= 12\left(\int_{-\infty}^{\infty} f^2(x)dx\right)^2. \\ &= \frac{3}{\pi} \approx 0.955 \text{ if } F \text{ is normal} \end{aligned}$$

5.4 Robust Estimates of “Effect Sizes” in Multivariate Normal Models

Let $\mathbf{Y}_j^E = (Y_{1j}^E, \dots, Y_{pj}^E)'$, $j = 1 \dots, n^E$ and $\mathbf{Y}_j^C = (Y_{1j}^C, \dots, Y_{pj}^C)'$, $j = 1 \dots, n^C$ be two independent samples of size n^E and n^C respectively from the p -variate normal distribution $N_p(\boldsymbol{\mu}^E, \boldsymbol{\Sigma})$ and $N_p(\boldsymbol{\mu}^C, \boldsymbol{\Sigma})$ respectively. We may consider the estimator of

$$\boldsymbol{\delta} = (\boldsymbol{\mu}^E - \boldsymbol{\mu}^C)D_{\boldsymbol{\Sigma}^{\frac{1}{2}}}^{-1}, \quad D_{\boldsymbol{\Sigma}^{\frac{1}{2}}}^{-1} = \text{Diag}(\sigma_1, \dots, \sigma_p).$$

as defined by

$$\tilde{\boldsymbol{\delta}}_n^{(\text{mean})} = (\bar{\mathbf{Y}}^E - \bar{\mathbf{Y}}^C)D_{\bar{\boldsymbol{\Sigma}}^{\frac{1}{2}}}^{-1}.$$

To obtain the asymptotic distribution of $\tilde{\boldsymbol{\delta}}_n^{(\text{mean})}$, we first state the following well-known result in the following theorem.

Theorem 5.4.1. *If $\mathbf{Y}_j = (Y_{1j}, \dots, Y_{pj})'$, $j = 1 \dots, n$ is a sample of size n from the p -variate normal distribution $N_p(\boldsymbol{\mu}, \boldsymbol{\Sigma})$, then*

$$\sqrt{n} (\bar{\mathbf{Y}} - \boldsymbol{\mu}) \sim N_p(\mathbf{0}, \boldsymbol{\Sigma}),$$

where $\bar{\mathbf{Y}} = (\bar{Y}_1, \dots, \bar{Y}_p)'$, $\Sigma = \left((\sigma_{ij}) \right)$ with

$$\begin{aligned}\sigma_{ii} &= \sigma_i^2 \\ \sigma_{ij} &= \sigma_i \sigma_j \rho_{ij}.\end{aligned}$$

As a consequence, the asymptotic distribution of $\sqrt{n} D_{\hat{\Sigma}^{\frac{1}{2}}}^{-1} (\bar{\mathbf{Y}} - \boldsymbol{\mu})$ is $N_p(\mathbf{0}, \boldsymbol{\rho})$, where $\boldsymbol{\rho} = D_{\Sigma^{\frac{1}{2}}}^{-1} \Sigma D_{\Sigma^{\frac{1}{2}}}^{-1}$, namely, $\boldsymbol{\rho} = \left((\rho_{ij}) \right)$.

Then, it is easy to see that

$$\left(\frac{n}{n^E} + \frac{n}{n^C} \right)^{-\frac{1}{2}} \sqrt{n} (\tilde{\boldsymbol{\delta}}_n^{(\text{mean})} - \boldsymbol{\delta}) \sim N_p(\mathbf{0}, \boldsymbol{\rho}) \text{ as } n \rightarrow \infty, \frac{n^E}{n} \rightarrow \pi^E \text{ and } \frac{n^C}{n} \rightarrow \pi^C.$$

$\tilde{\boldsymbol{\delta}}_n^{(\text{mean})}$ is affected by the extreme values of the samples. To overcome this problem, we study the asymptotic properties of the “scaled median difference vectors” and “scaled median of pair-wise average vectors” as estimators of effect size vector under p-variate normal distribution. As such, we consider the estimator of $\boldsymbol{\delta}$ based on the medians i.e.

$$\hat{\boldsymbol{\delta}}_n^{(M)} = (\text{Med}_{1 \leq j \leq n^E}(\mathbf{Y}_j^E) - \text{Med}_{1 \leq j \leq n^C}(\mathbf{Y}_j^C)) D_{\text{MAD}}^{-1}$$

where $\text{Med}_{1 \leq j \leq n^E}(\mathbf{Y}_j^E) = (\text{Med}(Y_{1j}^E), \dots, \text{Med}(Y_{pj}^E))'$

and $\text{Med}_{1 \leq j \leq n^C}(\mathbf{Y}_j^C) = (\text{Med}(Y_{1j}^C), \dots, \text{Med}(Y_{pj}^C))'$ respectively, and

D_{MAD} is the diagonal matrix with elements defined as follows

$$\begin{aligned}\text{and } \text{MAD}_n &= \frac{1}{0.6745} \left[\frac{n^E}{n^E + n^C} \text{Med}_{1 \leq j \leq n^E} \{ |Y_{lj}^E - \text{Med}(Y_l^E)| \} \right. \\ &\quad \left. + \frac{n^C}{n^E + n^C} \text{Med}_{1 \leq j \leq n^C} \{ |Y_{lj}^C - \text{Med}(Y_l^C)| \} \right], l = 1, \dots, p\end{aligned}$$

This estimator will not be effected by the outliers present in the sample. To obtain the asymptotic distribution of $\sqrt{n} (\hat{\boldsymbol{\delta}}_n^{(M)} - \boldsymbol{\delta})$ we first consider the second part of

Theorem 3.1 from Bickel (1964) under the p -variate normal distribution, $N_p(\boldsymbol{\mu}, \boldsymbol{\Sigma})$ together with the results from Puri and Sen (1971), pages 171 and 175, and we state them in the following theorem.

Theorem 5.4.2. *If $\mathbf{Y}_j = (Y_{1j}, \dots, Y_{pj})'$, $j = 1, \dots, n$ is a sample of size n from the p -variate normal distribution $N_p(\boldsymbol{\mu}, \boldsymbol{\Sigma})$, then as $n \rightarrow \infty$*

$$\sqrt{n} \left(\text{Med}(\mathbf{Y}_j) - \boldsymbol{\mu} \right) \sim N_p(\mathbf{0}, \boldsymbol{\tau}),$$

where $\boldsymbol{\tau} = \left(\left(\tau_{ij} \right) \right)$ with

$$\begin{aligned} \tau_{ii} &= \frac{\pi}{2} \sigma_i^2 \\ \tau_{ij} &= \sigma_i \sigma_j \sin^{-1}(\rho_{ij}), \quad i \neq j. \end{aligned}$$

As a consequence, the asymptotic distribution of $\sqrt{n} D_{MAD}^{-1} \left(\text{Med}(\mathbf{Y}_j) - \boldsymbol{\mu} \right)$ is $N_p(\mathbf{0}, \boldsymbol{\Psi}^{(M)})$, where $\boldsymbol{\Psi}^{(M)} = D_{\boldsymbol{\Sigma}^{\frac{1}{2}}}^{-1} \boldsymbol{\tau} D_{\boldsymbol{\Sigma}^{\frac{1}{2}}}^{-1}$, namely,

$$\boldsymbol{\Psi}^{(M)} = \begin{bmatrix} \frac{\pi}{2} & \sin^{-1}(\rho_{12}) & \cdots & \sin^{-1}(\rho_{1p}) \\ \sin^{-1}(\rho_{21}) & \frac{\pi}{2} & \cdots & \sin^{-1}(\rho_{2p}) \\ \vdots & \vdots & \vdots & \vdots \\ \sin^{-1}(\rho_{p1}) & \cdots & \cdots & \frac{\pi}{2} \end{bmatrix}.$$

Then, it is easy to see that

$$\left(\frac{n}{n^E} + \frac{n}{n^C} \right)^{-\frac{1}{2}} \sqrt{n} (\hat{\boldsymbol{\delta}}_n^{(M)} - \boldsymbol{\delta}) \sim N_p(\mathbf{0}, \boldsymbol{\Psi}^{(M)}) \text{ as } n = n^E + n^C \rightarrow \infty,$$

Hence, ARE of $\hat{\boldsymbol{\delta}}_n^{(M)}$ compared to $\tilde{\boldsymbol{\delta}}_n^{(\text{mean})}$ is given by

$$\begin{aligned} \text{ARE}(\hat{\boldsymbol{\delta}}_n^{(M)} : \tilde{\boldsymbol{\delta}}_n^{(\text{mean})}) &= \frac{2}{\pi} \left\{ \left| \left(\left(\rho_{ij} \right) \right) \right|^{\frac{1}{p}} \left| \left(\left(\frac{2}{\pi} \sin^{-1}(\rho_{ij}) \right) \right) \right|^{\frac{-1}{p}} \right\} \\ &\quad \text{using the determinants of the matrices } \boldsymbol{\rho}, \boldsymbol{\Psi}^{(M)}. \\ &= \frac{2}{\pi} \quad \text{using trace of the matrices } \boldsymbol{\rho}, \boldsymbol{\Psi}^{(M)}. \end{aligned}$$

Note that for the trace result, the ARE is the same as in the univariate case.

Next, we consider the estimator of $\boldsymbol{\delta}$ based on “scaled median of pair-wise average vectors” as

$$\begin{aligned} \hat{\boldsymbol{\delta}}_n^{(HL)} &= (\mathbf{Y}^{E(HL)} - \mathbf{Y}^{C(HL)}) D_{\text{MAD}}^{-1} \\ \text{where, } \mathbf{Y}^{E(HL)} &= (Y_1^{E(HL)}, \dots, Y_p^{E(HL)})' \\ \text{and, } \mathbf{Y}^{C(HL)} &= (Y_1^{C(HL)}, \dots, Y_p^{C(HL)})' \end{aligned}$$

respectively. Here (HL) stands for Hodges-Lehmann estimators, given by the pairwise averages

$$\begin{aligned} Y_i^{E(HL)} &= \text{Med}_{1 \leq j \leq k \leq n^E} \left[\frac{1}{2} (Y_{ij}^E + Y_{ik}^E) \right] \\ \text{and } Y_i^{C(HL)} &= \text{Med}_{1 \leq j \leq k \leq n^C} \left[\frac{1}{2} (Y_{ij}^C + Y_{ik}^C) \right], \quad i = 1, \dots, p. \end{aligned}$$

To obtain the asymptotic distribution of $\sqrt{n} (\hat{\boldsymbol{\delta}}_n^{(HL)} - \boldsymbol{\delta})$ we consider the third part of Theorem 3.1 from Bickel (1964) under the p-variate normal distribution, $N_p(\boldsymbol{\mu}, \boldsymbol{\Sigma})$ together with the results from Puri and Sen (1971), pages 175-176, and we state them in the following theorem.

Theorem 5.4.3. *If $\mathbf{Y}_j = (Y_{1j}, \dots, Y_{pj})'$, $j = 1, \dots, n$ is a sample of size n from the*

p -variate normal distribution $N_p(\boldsymbol{\mu}, \boldsymbol{\Sigma})$, then as $n \rightarrow \infty$

$$\sqrt{n} \left(\mathbf{Y}^{(HL)} - \boldsymbol{\mu} \right) \sim N_p(\mathbf{0}, \boldsymbol{\gamma}),$$

where $\boldsymbol{\gamma} = \left((\gamma_{ij}) \right)$ with

$$\begin{aligned} \gamma_{ii} &= \frac{\pi}{3} \sigma_i^2 \\ \gamma_{ij} &= 2\sigma_i \sigma_j \sin^{-1}\left(\frac{1}{2}\rho_{ij}\right), \quad i \neq j. \end{aligned}$$

As a consequence, the asymptotic distribution of $\sqrt{n} D_{MAD}^{-1} \left(\mathbf{Y}^{(HL)} - \boldsymbol{\mu} \right)$ is $N_p(\mathbf{0}, \boldsymbol{\Psi}^{(HL)})$, where $\boldsymbol{\Psi}^{(HL)} = D_{\boldsymbol{\Sigma}^{\frac{1}{2}}}^{-1} \boldsymbol{\gamma} D_{\boldsymbol{\Sigma}^{\frac{1}{2}}}^{-1}$, namely,

$$\boldsymbol{\Psi}^{(HL)} = \begin{bmatrix} \frac{\pi}{3} & 2 \sin^{-1}\left(\frac{1}{2}\rho_{12}\right) & \cdots & 2 \sin^{-1}\left(\frac{1}{2}\rho_{1p}\right) \\ 2 \sin^{-1}\left(\frac{1}{2}\rho_{21}\right) & \frac{\pi}{3} & \cdots & 2 \sin^{-1}\left(\frac{1}{2}\rho_{2p}\right) \\ \vdots & \vdots & \vdots & \vdots \\ 2 \sin^{-1}\left(\frac{1}{2}\rho_{p1}\right) & \cdots & \cdots & \frac{\pi}{3} \end{bmatrix}.$$

Then, it is easy to see that

$$\left(\frac{n}{n^E} + \frac{n}{n^C}\right)^{\frac{-1}{2}} \sqrt{n}(\hat{\boldsymbol{\delta}}_n^{(HL)} - \boldsymbol{\delta}) \sim N_p(\mathbf{0}, \boldsymbol{\Psi}^{(HL)})$$

Hence, the ARE of $\hat{\boldsymbol{\delta}}_n^{(HL)}$ relative to $\tilde{\boldsymbol{\delta}}_n^{(\text{mean})}$ is given by

$$\begin{aligned} \text{ARE}\left(\hat{\boldsymbol{\delta}}_n^{(HL)}; \tilde{\boldsymbol{\delta}}_n^{(\text{mean})}\right) &= \frac{3}{\pi} \left| \left((\rho_{ij}) \right) \right|^{\frac{1}{p}} \left| \left(\left(\frac{6}{\pi} \sin^{-1}\left(\frac{1}{2}\rho_{ij}\right) \right) \right) \right|^{\frac{-1}{p}} \\ &\quad \text{using the determinants of the matrices } \boldsymbol{\rho}, \boldsymbol{\Psi}^{(HL)}. \\ &= \frac{3}{\pi} \text{ using trace of the matrices } \boldsymbol{\rho}, \boldsymbol{\Psi}^{(HL)}. \end{aligned}$$

For $p = 2$, $0.91 \leq \text{ARE}(\hat{\delta}_n^{(HL)}; \tilde{\delta}_n^{(\text{mean})}) \leq 0.95$ where the lower bound is obtained when $\rho_{12} \rightarrow \pm 1$. For $\rho_{12} = 0$ i.e when the underlying population has pairwise independent coordinates, then the ARE is bounded below by $\frac{3}{\pi}$.

Next, we consider **test of hypothesis** $H_0 : \delta = 0$, one may use the test-statistic

$$\begin{aligned}\mathcal{L}_n^{(M)} &= n \hat{\delta}_n'^{(M)} \left(\hat{\Psi}^{*(M)} \right)^{-1} \hat{\delta}_n^{(M)} \\ \mathcal{L}_n^{(HL)} &= n \hat{\delta}_n'^{(HL)} \left(\hat{\Psi}^{*(HL)} \right)^{-1} \hat{\delta}_n^{(HL)} \\ \text{where } \hat{\Psi}^{*(M)} &= \left(\frac{n}{n^E} + \frac{n}{n^C} \right) \hat{\Psi}^{(M)} \text{ and } \hat{\Psi}^{*(HL)} = \left(\frac{n}{n^E} + \frac{n}{n^C} \right) \hat{\Psi}^{(HL)}\end{aligned}$$

As $n \rightarrow \infty$, $\mathcal{L}_n^{(M)} \xrightarrow{\mathcal{D}} \chi_p^2$, $\mathcal{L}_n^{(HL)} \xrightarrow{\mathcal{D}} \chi_p^2$.

Under local alternatives $\mathcal{K}_{(n)} : \delta = n^{-\frac{1}{2}} \xi$, $\mathcal{L}_n^{(M)} \xrightarrow{\mathcal{D}} \chi_p^2(\Delta^2)$, $\Delta^2 = \xi' \left(\Psi^{*(M)} \right)^{-1} \xi$. Similarly, $\mathcal{L}_n^{(HL)} \xrightarrow{\mathcal{D}} \chi_p^2(\Delta^2)$, $\Delta^2 = \xi' \left(\Psi^{*(HL)} \right)^{-1} \xi$.

Then, we may define three more estimators using $\hat{\delta}_n^{(M)}$ and $\hat{\delta}_n^{(HL)}$ respectively as follows.

$$\begin{aligned}\hat{\delta}_n^{PT(M)} &= \hat{\delta}_n^{(M)} - \hat{\delta}_n^{(M)} I(\mathcal{L}_n^{(M)} \leq \chi_{p,\alpha}^2). \\ \hat{\delta}_n^{PT(HL)} &= \hat{\delta}_n^{(HL)} - \hat{\delta}_n^{(HL)} I(\mathcal{L}_n^{(HL)} \leq \chi_{p,\alpha}^2). \\ \hat{\delta}_n^{s(M)} &= \hat{\delta}_n^{(M)} - (p-2)[\mathcal{L}_n^{(M)}]^{-1} \hat{\delta}_n^{(M)}. \\ \hat{\delta}_n^{s(HL)} &= \hat{\delta}_n^{(HL)} - (p-2)[\mathcal{L}_n^{(HL)}]^{-1} \hat{\delta}_n^{(HL)}. \\ \hat{\delta}_n^{s+(M)} &= \hat{\delta}_n^{(M)} \{1 - (p-2)[\mathcal{L}_n^{(M)}]^{-1}\} I(\mathcal{L}_n^{(M)} \geq p-2). \\ \hat{\delta}_n^{s+(HL)} &= \hat{\delta}_n^{(HL)} \{1 - (p-2)[\mathcal{L}_n^{(HL)}]^{-1}\} I(\mathcal{L}_n^{(HL)} \geq p-2).\end{aligned}$$

It may be shown following the results of chapter 4 that the ADB, ADMSE and ADQR expressions of the above estimators are given by

Preliminary test estimator (PTE)

$$\begin{aligned}
b_2(\hat{\boldsymbol{\delta}}_n^{PT(M)}) &= -\boldsymbol{\xi} H_{p+2}(\chi_{p,\alpha}^2, \Delta^2). \\
M_2(\hat{\boldsymbol{\delta}}_n^{PT(M)}) &= \boldsymbol{\Psi}^{*(M)} - \boldsymbol{\Psi}^{*(M)} H_{p+2}(\chi_{p,\alpha}^2; \Delta^2) \\
&\quad + \boldsymbol{\xi} \boldsymbol{\xi}' \{2H_{p+2}(\chi_{p,\alpha}^2; \Delta^2) - H_{p+4}(\chi_{p,\alpha}^2; \Delta^2)\}. \\
R_2(\hat{\boldsymbol{\delta}}_n^{PT(M)}, \mathbf{W}) &= \text{tr}(\mathbf{W} \boldsymbol{\Psi}^{*(M)}) \{1 - H_{p+2}(\chi_{p,\alpha}^2; \Delta^2)\} \\
&\quad + \boldsymbol{\xi}' \mathbf{W} \boldsymbol{\xi} \{2H_{p+2}(\chi_{p,\alpha}^2; \Delta^2) - H_{p+4}(\chi_{p,\alpha}^2; \Delta^2)\}.
\end{aligned}$$

James-Stein estimator (JSE)

$$\begin{aligned}
b_3(\hat{\boldsymbol{\delta}}_n^{s(M)}) &= -(p-2) \boldsymbol{\xi} E(\chi_{p+2}^{-2}(\Delta^2)). \\
M_3(\hat{\boldsymbol{\delta}}_n^{s(M)}) &= \boldsymbol{\Psi}^{*(M)} - (p-2) \boldsymbol{\Psi}^{*(M)} \{2E[\chi_{p+2}^{-2}(\Delta^2)] - (p-2)E[\chi_{p+2}^{-4}(\Delta^2)]\} \\
&\quad + (p^2-4) \boldsymbol{\xi} \boldsymbol{\xi}' E[\chi_{p+4}^{-4}(\Delta^2)]. \\
R_3(\hat{\boldsymbol{\delta}}_n^{s(M)}; \mathbf{W}) &= \text{tr}(\mathbf{W} \boldsymbol{\Psi}^{*(M)}) - (p-2) \text{tr}(\mathbf{W} \boldsymbol{\Psi}^{*(M)}) \{2E[\chi_{p+2}^{-2}(\Delta^2)] \\
&\quad - (p-2)E[\chi_{p+2}^{-4}(\Delta^2)]\} + (p^2-4) \boldsymbol{\xi}' \mathbf{W} \boldsymbol{\xi} E[\chi_{p+4}^{-4}(\Delta^2)].
\end{aligned}$$

Positive rule Stein estimator (PRSE)

$$\begin{aligned}
b_4(\hat{\delta}_n^{s+(M)}) &= -\xi\{(p-2)E[\chi_{p+2}^{-2}(\Delta^2)] - \xi E[1 - (p-2)\chi_{p+2}^{-2}(\Delta^2) I(\chi_{p+2}^2(\Delta^2) < (p-2))]\} \\
M_4(\hat{\delta}_N^{s+(M)}) &= M_3(\hat{\delta}_n^{s(M)}) - (p-2)\Psi^{*(M)} \\
&\quad \times \{E[(1 - (p-2)\chi_{p+2}^{-2}(\Delta^2))^2 I(\chi_{p+2}^2(\Delta^2) \leq p-2)]\} \\
&\quad + \xi\xi'\{2E[(1 - (p-2)\chi_{p+2}^{-2}(\Delta^2))I(\chi_{p+2}^2(\Delta^2) \leq p-2)] \\
&\quad - E[(1 - (p-2)\chi_{p+4}^{-2}(\Delta^2))^2 I(\chi_{p+4}^2(\Delta^2) \leq p-2)]\}. \\
R_4(\hat{\delta}_n^{s+(M)}; \mathbf{W}) &= R_3(\hat{\delta}_n^{s(M)}; \mathbf{W}) - (p-2)tr(\mathbf{W}\Psi^{*(M)}) \\
&\quad \times E[(1 - (p-2)\chi_{p+2}^{-2}(\Delta^2))^2 I(\chi_{p+2}^2(\Delta^2) \leq p-2)] \\
&\quad + \xi'\mathbf{W}\xi\{2E[(1 - (p-2)\chi_{p+2}^{-2}(\Delta^2))I(\chi_{p+2}^2(\Delta^2) \leq p-2)] \\
&\quad - E[(1 - (p-2)\chi_{p+4}^{-2}(\Delta^2))^2 I(\chi_{p+4}^2(\Delta^2) \leq p-2)]\}.
\end{aligned}$$

Similarly, the ADB, ADMSE and ADQR expressions of the HL estimators are given by

Preliminary test estimator (PTE)

$$\begin{aligned}
b_2(\hat{\delta}_n^{PT(HL)}) &= -\xi H_{p+2}(\chi_{p,\alpha}^2, \Delta^2). \\
M_2(\hat{\delta}_n^{PT(HL)}) &= \Psi^{*(HL)} - \Psi^{*(HL)} H_{p+2}(\chi_{p,\alpha}^2; \Delta^2) \\
&\quad + \xi\xi'\{2H_{p+2}(\chi_{p,\alpha}^2; \Delta^2) - H_{p+4}(\chi_{p,\alpha}^2; \Delta^2)\}. \\
R_2(\hat{\delta}_n^{PT(HL)}, \mathbf{W}) &= tr(\mathbf{W}\Psi^{*(HL)})\{1 - H_{p+2}(\chi_{p,\alpha}^2; \Delta^2)\} \\
&\quad + \xi'\mathbf{W}\xi\{2H_{p+2}(\chi_{p,\alpha}^2; \Delta^2) - H_{p+4}(\chi_{p,\alpha}^2; \Delta^2)\}.
\end{aligned}$$

James-Stein estimator (JSE)

$$\begin{aligned}
b_3(\hat{\boldsymbol{\delta}}_n^{s(HL)}) &= -(p-2)\boldsymbol{\xi}E(\chi_{p+2}^{-2}(\Delta^2)). \\
M_3(\hat{\boldsymbol{\delta}}_n^{s(HL)}) &= \boldsymbol{\Psi}^{*(HL)} - (p-2)\boldsymbol{\Psi}^{*(HL)}\{2E[\chi_{p+2}^{-2}(\Delta^2)] - (p-2)E[\chi_{p+2}^{-4}(\Delta^2)]\} \\
&\quad + (p^2-4)\boldsymbol{\xi}\boldsymbol{\xi}'E[\chi_{p+4}^{-4}(\Delta^2)]. \\
R_3(\hat{\boldsymbol{\delta}}_n^{s(HL)}; \mathbf{W}) &= tr(\mathbf{W}\boldsymbol{\Psi}^{*(HL)}) - (p-2)tr(\mathbf{W}\boldsymbol{\Psi}^{*(HL)})\{2E[\chi_{p+2}^{-2}(\Delta^2)] \\
&\quad - (p-2)E[\chi_{p+2}^{-4}(\Delta^2)]\} + (p^2-4)\boldsymbol{\xi}'\mathbf{W}\boldsymbol{\xi}E[\chi_{p+4}^{-4}(\Delta^2)].
\end{aligned}$$

Positive rule Stein estimator (PRSE)

$$\begin{aligned}
b_4(\hat{\boldsymbol{\delta}}_n^{s+(HL)}) &= -\boldsymbol{\xi}\{(p-2)E[\chi_{p+2}^{-2}(\Delta^2)] - \boldsymbol{\xi}E[1 - (p-2)\chi_{p+2}^{-2}(\Delta^2) I(\chi_{p+2}^2(\Delta^2) < (p-2))]\} \\
M_4(\hat{\boldsymbol{\delta}}_n^{s+(HL)}) &= M_3(\hat{\boldsymbol{\delta}}_n^{s(HL)}) - (p-2)\boldsymbol{\Psi}^{*(HL)} \\
&\quad \times \{E[(1 - (p-2)\chi_{p+2}^{-2}(\Delta^2))^2 I(\chi_{p+2}^2(\Delta^2) \leq p-2)]\} \\
&\quad + \boldsymbol{\xi}\boldsymbol{\xi}'\{2E[(1 - (p-2)\chi_{p+2}^{-2}(\Delta^2))I(\chi_{p+2}^2(\Delta^2) \leq p-2)] \\
&\quad - E[(1 - (p-2)\chi_{p+4}^{-2}(\Delta^2))^2 I(\chi_{p+4}^2(\Delta^2) \leq p-2)]\}. \\
R_4(\hat{\boldsymbol{\delta}}_n^{s+(HL)}; \mathbf{W}) &= R_3(\hat{\boldsymbol{\delta}}_n^{s(HL)}; \mathbf{W}) - (p-2)tr(\mathbf{W}\boldsymbol{\Psi}^{*(HL)}) \\
&\quad \times E[(1 - (p-2)\chi_{p+2}^{-2}(\Delta^2))^2 I(\chi_{p+2}^2(\Delta^2) \leq p-2)] \\
&\quad + \boldsymbol{\xi}'\mathbf{W}\boldsymbol{\xi}\{2E[(1 - (p-2)\chi_{p+2}^{-2}(\Delta^2))I(\chi_{p+2}^2(\Delta^2) \leq p-2)] \\
&\quad - E[(1 - (p-2)\chi_{p+4}^{-2}(\Delta^2))^2 I(\chi_{p+4}^2(\Delta^2) \leq p-2)]\}.
\end{aligned}$$

Next, we study the risk analysis and MSE analysis of the estimators based on medians and the analysis of the estimators based on HL solution is very similar.

Risk Analysis

We observe that $\hat{\delta}_n^{PT(M)}$ is superior to $\hat{\delta}_n^{(M)}$ whenever

$$\Delta^2 \leq \frac{tr(\mathbf{W}\Psi^{*(M)})H_{p+2}(\chi_{p,\alpha}^2; \Delta^2)}{ch_{max}(\mathbf{W}\Psi^{*(M)})[2H_{p+2}(\chi_{p,\alpha}^2; \Delta^2) - H_{p+4}(\chi_{p,\alpha}^2; \Delta^2)]}$$

Otherwise $\hat{\delta}_n^{(M)}$ is superior whenever

$$\Delta^2 \geq \frac{tr(\mathbf{W}\Psi^{*(M)})H_{p+2}(\chi_{p,\alpha}^2; \Delta^2)}{ch_{min}(\mathbf{W}\Psi^{*(M)})[2H_{p+2}(\chi_{p,\alpha}^2; \Delta^2) - H_{p+4}(\chi_{p,\alpha}^2; \Delta^2)]}$$

Next, we consider James-Stein type estimators, using (A.6) we can rewrite $R_3(\hat{\delta}_n^{s(M)}; \mathbf{W})$

as

$$\begin{aligned} R_3(\hat{\delta}_n^{s(M)}; \mathbf{W}) &= tr(\mathbf{W}\Psi^{*(M)}) - (p-2)tr(\mathbf{W}\Psi^{*(M)})\left\{ \left[1 - \frac{(p+2)\boldsymbol{\xi}'\mathbf{W}\boldsymbol{\xi}}{2\Delta^2 tr(\mathbf{W}\Psi^{*(M)})}\right] 2\Delta^2 E[\chi_{p+4}^{-4}(\Delta^2)] \right. \\ &\quad \left. + (p-2)E[\chi_{p+2}^{-4}(\Delta^2)] \right\} \end{aligned}$$

The risk-difference $R_1(\hat{\delta}_n^{(M)}; \mathbf{W}) - R_3(\hat{\delta}_n^{s(M)}; \mathbf{W})$ is non-negative for all Δ^2 for the matrix $\mathbf{W}$ belongs to the set satisfying

$$\frac{tr(\mathbf{W}\Psi^{*(M)})}{ch_{max}(\mathbf{W}\Psi^{*(M)})} \geq \frac{p+2}{2} \quad (5.2)$$

Further, the risk-difference

$$R_3(\hat{\delta}_n^{s(M)}; \mathbf{W}) - R_4(\hat{\delta}_n^{s+(M)}; \mathbf{W}) \geq 0 \quad \forall \Delta^2 \in \mathcal{R}_1^+ \quad (5.3)$$

since,

$$\begin{aligned}
R_3(\hat{\delta}_n^{s(M)}; \mathbf{W}) - R_4(\hat{\delta}_n^{s+(M)}; \mathbf{W}) &= (p-2)tr(\mathbf{W}\Psi^{*(M)})E[(1-(p-2)\chi_{p+2}^{-2}(\Delta^2))^2 \\
&\quad \times I(\chi_{p+2}^2(\Delta^2) \leq p-2)] - \boldsymbol{\xi}'\mathbf{W}\boldsymbol{\xi}\{2E[(1-(p-2)\chi_{p+2}^{-2}(\Delta^2))^2 \\
&\quad \times I(\chi_{p+2}^2(\Delta^2) \leq p-2)] - E[(1-(p-2)\chi_{p+4}^{-2}(\Delta^2))^2 \\
&\quad \times I(\chi_{p+4}^2(\Delta^2) \leq p-2)]\}.
\end{aligned} \tag{5.4}$$

we may note that $E[(1-(p-2)\chi_{p+2}^{-2}(\Delta^2))^2 I(\chi_{p+2}^2(\Delta^2) \leq p-2)]$ is positive, and $2E[(1-(p-2)\chi_{p+2}^{-2}(\Delta^2)) I(\chi_{p+2}^2(\Delta^2) \leq p-2)]$ is negative, while $E[(1-(p-2)\chi_{p+4}^{-2}(\Delta^2))^2 I(\chi_{p+4}^2(\Delta^2) \leq p-2)]$ is positive. Hence (5.3) holds. Bringing these results together, we obtain

$$R_1(\hat{\delta}_n^{(M)}; \mathbf{W}) \geq R_3(\hat{\delta}_n^{s(M)}; \mathbf{W}) \geq R_4(\hat{\delta}_n^{s+(M)}; \mathbf{W}) \quad \forall (\Delta^2; \mathbf{W})$$

where $\mathbf{W}$ belongs to the set (5.2).

MSE Matrix Analysis of the Estimators

We study the MSE matrix analysis of the estimators and determine their dominance properties.

- Comparison of $\hat{\delta}_n^{PT(M)}$ and $\hat{\delta}_n^{(M)}$

The MSE difference of $\hat{\delta}_n^{(M)}$ and $\hat{\delta}_n^{PT(M)}$ is

$$\begin{aligned}
M_1(\tilde{\delta}_n^{(M)}) - M_2(\hat{\delta}_n^{PT(M)}) &= \Psi^{*(M)} H_{p+2}(\chi_{p,\alpha}^2; \Delta^2) \\
&\quad - \boldsymbol{\xi}\boldsymbol{\xi}' \{2H_{p+2}(\chi_{p,\alpha}^2; \Delta^2) - H_{p+4}(\chi_{p,\alpha}^2; \Delta^2)\}.
\end{aligned}$$

The MSE-difference is (p.s.d.) whenever for a non-zero vector $\boldsymbol{\ell} = (\ell_1, \dots, \ell_p)'$, we

have

$$\ell' [M_1(\hat{\delta}_n^{(M)}) - M_2(\hat{\delta}_n^{PT(M)})] \ell \geq 0.$$

Thus,

$$\ell' \xi \xi' \ell \{2H_{p+2}(\chi_{p,\alpha}^2; \Delta^2) - H_{p+4}(\chi_{p,\alpha}^2; \Delta^2)\} \leq \ell' \Psi^{*(M)} \ell H_{p+2}(\chi_{p,\alpha}^2; \Delta^2)$$

dividing by $\ell' \Psi^{*(M)} \ell$ we obtain

$$\frac{\ell' \xi \xi' \ell}{\ell' \Psi^{*(M)} \ell} \leq \frac{H_{p+2}(\chi_{p,\alpha}^2; \Delta^2)}{\{2H_{p+2}(\chi_{p,\alpha}^2; \Delta^2) - H_{p+4}(\chi_{p,\alpha}^2; \Delta^2)\}}$$

using theorem (A.3.3), we get

$$\max_{\ell} \frac{\ell' \xi \xi' \ell}{\ell' \Psi^{*(M)} \ell} = \xi' [\Psi^{*(M)}]^{-1} \xi = \Delta^2$$

Therefore, we get

$$\Delta^2 \leq \frac{H_{p+2}(\chi_{p,\alpha}^2; \Delta^2)}{\{2H_{p+2}(\chi_{p,\alpha}^2; \Delta^2) - H_{p+4}(\chi_{p,\alpha}^2; \Delta^2)\}} \quad (5.5)$$

Therefore, $\hat{\delta}_n^{PT(M)}$ performs better than $\hat{\delta}_n^{(M)}$ in the range of Δ^2 given by (5.5) in the MSE sense.

- Comparison of $\hat{\delta}_n^{s(M)}$ and $\hat{\delta}_n$

The MSE-difference of $\hat{\delta}_n$ and $\hat{\delta}_n^{s(M)}$ is given by

$$\begin{aligned} M_1(\hat{\delta}_n^{(M)}) - M_3(\hat{\delta}_n^{s(M)}) &= \Psi^{*(M)}(p-2)\{2E[\chi_{p+2}^{-2}(\Delta^2)] - (p-2)E[\chi_{p+2}^{-4}(\Delta^2)]\} \\ &\quad - (p^2 - 4)\delta\delta' E[\chi_{p+4}^{-4}(\Delta^2)]. \end{aligned}$$

As above, the MSE-difference is p.s.d. whenever we have

$$\ell' [M_1(\hat{\delta}_n^{(M)}) - M_3(\hat{\theta}_n^{s(M)})] \ell \geq 0,$$

and consequently, we may write

$$\frac{\ell' \xi \xi' \ell}{\ell' \Psi^{*(M)} \ell} \leq \frac{2E[\chi_{p+2}^{-2}(\Delta^2)] - (p-2)E[\chi_{p+2}^{-4}(\Delta^2)]}{(p+2)E[\chi_{p+4}^{-4}(\Delta^2)]}.$$

Therefore, we get

$$\Delta^2 \leq \frac{2E[\chi_{p+2}^{-2}(\Delta^2)] - (p-2)E[\chi_{p+2}^{-4}(\Delta^2)]}{(p+2)E[\chi_{p+4}^{-4}(\Delta^2)]}. \quad (5.6)$$

This means $\hat{\delta}_n^{s(M)}$ is superior to $\hat{\delta}_n^{(M)}$ in the range of Δ^2 given by (5.6), otherwise $\hat{\delta}_n$ is superior to $\hat{\delta}_n^{s(M)}$.

- Comparison of $\hat{\delta}_n^{PT(M)}$, $\hat{\delta}_n^{s(M)}$, and $\hat{\delta}_n^{s+(M)}$

Consider the MSE-difference

$$\begin{aligned} M_3(\hat{\delta}_n^{s(M)}) - M_2(\hat{\delta}_n^{PT(M)}) &= \xi \xi' \{ (p^2 - 4)E[\chi_{p+4}^{-4}(\Delta^2)] - 2H_{p+2}(\chi_p^2(\alpha); \Delta^2) + H_{p+4}(\chi_p^2(\alpha); \Delta^2) \} \\ &\quad + \Psi^{*(M)} \{ H_{p+2}(\chi_p^2(\alpha), \Delta^2) - 2(p-2)E[\chi_{p+4}^{-2}(\Delta^2)] \\ &\quad + (p-2)^2 E[\chi_{p+2}^{-4}(\Delta^2)] \} \end{aligned} \quad (5.7)$$

and

$$\begin{aligned}
M_4(\hat{\delta}_n^{s+(M)}) - M_2(\hat{\delta}_n^{PT(M)}) &= -(p-2)\Psi^{*(M)}\{2E[\chi_{p+2}^{-2}(\Delta^2)] - (p-2)E[\chi_{p+2}^{-4}(\Delta^2)] \\
&\quad + E[(1 - (p-2)\chi_{p+2}^{-2}(\Delta^2))^2 I(\chi_{p+2}^2(\Delta^2) \leq p-2)] \\
&\quad - (p-2)^{-1}H_{p+2}(\chi_p^2(\alpha), \Delta^2)\} + \xi\xi'\{(p^2-4)E[\chi_{p+4}^{-4}(\Delta^2)] \\
&\quad + 2E[(1 - (p-2)\chi_{p+2}^{-2}(\Delta^2))I(\chi_{p+2}^2(\Delta^2) \leq p-2)] \\
&\quad - E[(1 - (p-2)\chi_{p+4}^{-2}(\Delta^2))^2 I(\chi_{p+4}^2(\Delta^2) \leq p-2)] \\
&\quad - 2H_{p+2}(\chi_p^2(\alpha), \Delta^2) + H_{p+4}(\chi_p^2(\alpha), \Delta^2)\} \quad (5.8)
\end{aligned}$$

The expression in (5.7) is p.s.d. whenever

$$\Delta^2 \leq \frac{H_{p+2}(\chi_p^2(\alpha), \Delta^2) - 2(p-2)E[\chi_{p+4}^{-2}(\Delta^2)] + (p-2)^2E[\chi_{p+2}^{-4}(\Delta^2)]}{2H_{p+2}(\chi_p^2(\alpha); \Delta^2) - H_{p+4}(\chi_p^2(\alpha); \Delta^2) - (p^2-4)E[\chi_{p+4}^{-4}(\Delta^2)]} \quad (5.9)$$

Thus, $\hat{\delta}_n^{PT(M)}$ performs better than $\hat{\delta}_n^{s(M)}$ whenever Δ^2 satisfies (5.9); otherwise $\hat{\delta}_n^{s(M)}$ performs better than $\hat{\delta}_n^{PT(M)}$. Neither estimators dominate the other.

Similarly the expression in (5.8) is p.s.d. whenever

$$\begin{aligned}
\Delta^2 \leq & \{(p-2)^2E[\chi_{p+2}^{-4}(\Delta^2)] - 2(p-2)E[\chi_{p+2}^{-2}(\Delta^2)] - (p-2)E[(1 - (p-2)\chi_{p+2}^{-2}(\Delta^2))^2 \\
& \times I(\chi_{p+2}^2(\Delta^2) \leq p-2)] + H_{p+2}(\chi_p^2(\alpha), \Delta^2)\} / \{E[(1 - (p-2)\chi_{p+4}^{-2}(\Delta^2))^2 \\
& \times I(\chi_{p+4}^2(\Delta^2) \leq p-2)] + 2H_{p+2}(\chi_p^2(\alpha), \Delta^2) - H_{p+4}(\chi_p^2(\alpha), \Delta^2) \\
& - (p^2-4)E[\chi_{p+4}^{-4}(\Delta^2)] - 2E[(1 - (p-2)\chi_{p+2}^{-2}(\Delta^2))I(\chi_{p+2}^2(\Delta^2) \leq p-2)]\} \quad (5.10)
\end{aligned}$$

Thus, $\hat{\delta}_n^{PT(M)}$ performs better than $\hat{\delta}_n^{s+(M)}$ whenever Δ^2 satisfies (5.10); otherwise $\hat{\delta}_n^{s+(M)}$ performs better than $\hat{\delta}_n^{PT(M)}$. Neither estimators dominate the other.

- Comparison of $\hat{\delta}_n^{s+(M)}$ and $\hat{\delta}_n^{s(M)}$

Finally, we compare $\hat{\boldsymbol{\delta}}_n^{s+(M)}$ and $\hat{\boldsymbol{\delta}}_n^{s(M)}$. In this case, MSE-difference yields

$$\begin{aligned}
 M_3(\hat{\boldsymbol{\delta}}_n^{s(M)}) - M_4(\hat{\boldsymbol{\delta}}_n^{s+(M)}) &= (p-2)\boldsymbol{\Psi}^{*(M)}\{E[(1-(p-2)\chi_{p+2}^{-2}(\Delta^2))^2 I(\chi_{p+2}^2(\Delta^2) \leq p-2)]\} \\
 &\quad + \boldsymbol{\xi}\boldsymbol{\xi}'\{2E[(1-(p-2)\chi_{p+2}^{-2}(\Delta^2))I(\chi_{p+2}^2(\Delta^2) \leq p-2)] \\
 &\quad - E[(1-(p-2)\chi_{p+4}^{-2}(\Delta^2))^2 I(\chi_{p+4}^2(\Delta^2) \leq p-2)]\} \\
 &\geq 0.
 \end{aligned}$$

The MSE-difference is p.s.d. for all Δ^2 and hence $\hat{\boldsymbol{\delta}}_n^{s+(M)}$ dominates $\hat{\boldsymbol{\delta}}_n^s$ uniformly.

We can order MSE of $\hat{\boldsymbol{\delta}}_n^{s+(M)}$ and $\hat{\boldsymbol{\delta}}_n^{s(M)}$ as

$$M_3(\hat{\boldsymbol{\delta}}_n^{s(M)}) \geq M_4(\hat{\boldsymbol{\delta}}_n^{s+(M)}).$$

Chapter 6

Future Research

In this thesis, we provided a new dimension of statistical methodology for meta analysis and approached the problem of combining several studies via quasi-empirical Bayes method. The method have been used to improve the point estimates of the effect sizes under several set ups. Next, immediate problem is to study how to combine these improved estimators to get a new value which can serve as the new common effect size. We also can study the problem of meta-regression, when we wish to estimate the effect sizes and it is suspected that δ belongs to the subspace defined by $\delta = \mathbf{B}\eta$ where $\mathbf{B}$ is a known $p \times q$ matrix and η is the regression vector of q -dimension. This approach is effective to deal with missing data problem as well as to capture the effect of auxiliary variables when the hypothesis of homogeneity is rejected, to combine the effect sizes by “ borrowing strength “ from components of regression.

In our study we have presented the new estimates for each problem discussed with their MSE and risk expressions but not the “ unbiased estimators “ of MSE or covariance and risk. One needs, particularly, the unbiased estimators of component wise variances to construct suitable confidence intervals. We like to pursue this idea of research to complement their point estimation methodology by confidence

intervals.

In meta analysis there are two general ways of combining studies, namely, combining points estimates and combining P-values from independent investigations. Hedges and Olkin (1985) and Hartung, Knapp and Sinha (2008) have given complete summary of the results available for the P-values combinations due to Tippett (1931), Fisher (1932) and Pearson (1933) with illustrations. Among all these methods Fisher's method is quite popular. The method is based on the fact that $-\ln(1 - P)$ is distributed as a chi square variable with two degrees of freedom under the null hypothesis of homogeneity when the P-values are themselves uniformly distributed. Thus, for k independent tests of the same hypothesis, $-2 \sum_{i=1}^k \ln(1 - P_i)$ is a chi square variable with $2k$ degrees of freedom. Hence, the null hypothesis is rejected when $-2 \sum_{i=1}^k \ln(1 - P_i)$ exceeds the $100(1 - \alpha)\%$ critical value of χ_{2k}^2 . This test suggest that at least one population has the effect size is not agreeing with the null hypothesis. Fisher's methods received criticism that it did not address the power and other conventional criteria. Recently, Fraser, Saleh and Ji (2010) used first order likelihood theory to derive a simple modification, namely, the P-values are converted to likelihood values and likelihood values to observed likelihood functions and in turn to a new composite P-value. We like to pursue this methodology for the meta analysis. Also, study the properties of preliminary test estimator defined with the P-values instead of the usual one presented in this thesis. In chapter 5, we have presented the preliminary steps towards robust methods. Our next problem of research would be develop general rank, M and L-methods for the combination of effect sizes to enlarge the scope of meta analysis.

Appendix A

Appendix: Preliminaries

In this appendix we summarize some of the core results on Normal distribution and Chi-square distribution which are referred to throughout the thesis . For proof of these results, the reader is referred to Judge and Bock (1978) and Saleh (2006). We also, provide some matrix definitions and theorems.

A.1 Normal Distribution

Theorem A.1.1. *If $Z \sim N(\Delta, 1)$, then*

$$\begin{aligned} E(|Z|) &= \sqrt{\frac{2}{\pi}} e^{-\frac{\Delta^2}{2}} + \Delta[2\Phi(\Delta) - 1]. \\ E\left[\frac{Z}{|Z|}\right] &= 1 - 2\Phi(-\Delta), \end{aligned}$$

where Φ is the cdf of the standard normal distribution.

A.2 Chi-square Distribution and Properties

If X is $N(0, 1)$, then X^2 is distributed as a central chi-square variable with one degree of freedom (D.F.) denoted as χ_1^2 . However, if X is $N(\mu, \sigma^2)$, then X^2 is

distributed as a non-central chi-square variable $\chi_1^2(\Delta^2)$ with one degree of freedom (D.F) and non-centrality parameter $\frac{\mu^2}{2\sigma^2} = \frac{1}{2}\Delta^2$. The pdf of a non-central chi-square variable with ν (D.F.) and non-centrality parameter $\frac{1}{2}\Delta^2$ may be written as

$$h_\nu(\chi^2; \Delta^2) = \exp\left\{-\frac{\Delta^2}{2}\right\} \sum_{r \geq 0} \frac{\left(\frac{-\Delta^2}{2}\right)^r}{r!} h_{\nu+2r}(\chi^2; 0) \quad (\text{A.1})$$

where $h_{\nu+2r}(\chi^2; 0)$ is the pdf of the central chi-square distribution with $\nu + 2r$ (D.F.). The pdf of a central chi-square distribution with ν (D.F.) may be written as

$$h_\nu(\chi^2; 0) = \frac{(\chi^2)^{\frac{\nu}{2}-1} e^{-\frac{\chi^2}{2}}}{2^{\frac{\nu}{2}} \Gamma(\frac{\nu}{2})}.$$

We denote by $H_\nu(x, 0)$ the cdf of a central chi square distribution with ν (D.F.).

$H_\nu(x, 0)$ may be written as

$$H_\nu(x, 0) = H_{\nu+2}(x, 0) + h_{\nu+2}(x, 0). \quad (\text{A.2})$$

We denote by $H_\nu(x, \Delta^2)$ the cdf of a non-central chi square distribution with ν (D.F.) and non-centrality parameter $\frac{1}{2}\Delta^2$. Some of the results involving chi-square distribution are given below in theorem/corollary form.

Theorem A.2.1. *If $\mathbf{Z}$ has the multivariate normal distribution $N_p(\boldsymbol{\mu}, \boldsymbol{\Sigma})$ and $Y = \mathbf{Z}' \mathbf{A} \mathbf{Z}$ for some symmetric matrix $\mathbf{A}$, then $Y \stackrel{\mathcal{D}}{=} \sum_{i=1}^p \lambda_i X_i^2$, where $X_1^2, \dots, X_p^2$ are independent chi square variables with one degree of freedom each and non-centrality parameter θ_i (say) and $\lambda_1, \dots, \lambda_p$ are the eigenvalues of $\mathbf{A}^{\frac{1}{2}} \boldsymbol{\Sigma} (\mathbf{A}^{\frac{1}{2}})'$.*

For the proof of this theorem see Bishop, Fienberg and Holland (1975) pages 473-474.

Theorem A.2.2. *The reciprocal of the central chi-square variable with ν degrees of*

freedom has expected value

$$E[(\chi_\nu^2)^{-1}] = \frac{1}{\nu - 2}, \text{ for } \nu \geq 3.$$

And the square of the reciprocal of the central chi-square variable with ν degrees of freedom has expected value

$$E[(\chi_\nu^2)^{-2}] = \frac{1}{(\nu - 2)(\nu - 4)}, \text{ for } \nu > 4.$$

Theorem A.2.3. *Let χ_ν^2 be a central chi-square random variable with ν degrees of freedom and let $\phi(\chi_\nu^2)$ be an integrable function of χ_ν^2 . Then*

$$E[\chi_\nu^2 \phi(\chi_\nu^2)] = \nu E[\phi(\chi_{\nu+2}^2)].$$

Theorem A.2.4. *Let $\chi_\nu^2(\Delta^2)$ be a non-central chi-square variable with ν degrees of freedom and non-centrality parameter $\frac{1}{2}\Delta^2$, and let $\phi(\chi_\nu^2(\Delta^2))$ be an integrable function of $\chi_\nu^2(\Delta^2)$. Then*

$$E[\chi_\nu^2(\Delta^2) \phi(\chi_\nu^2(\Delta^2))] = \nu E[\phi(\chi_{\nu+2}^2(\Delta^2))] + \Delta^2 E[\phi(\chi_{\nu+4}^2(\Delta^2))].$$

Theorem A.2.5. *If $Z \sim N(\Delta, 1)$ and $\phi(Z^2)$ is an integrable function of Z^2 , then*

$$E[Z\phi(Z^2)] = \Delta E[\phi(\chi_3^2(\Delta^2))]. \quad (\text{A.3})$$

$$E[Z^2\phi(Z^2)] = E[\phi(\chi_3^2(\Delta^2))] + \Delta^2 E[\phi(\chi_5^2(\Delta^2))]. \quad (\text{A.4})$$

Theorem A.2.6. *Let $\mathbf{Z} = (Z_1, \dots, Z_p)'$ be a p -dimensional vector distributed as*

$N_p(\boldsymbol{\mu}, \mathbf{I}_p)$, then for an integrable function $\phi(\mathbf{Z}'\mathbf{Z})$ of $\mathbf{Z}'\mathbf{Z}$ we have

$$\begin{aligned} E[\mathbf{Z} \phi(\mathbf{Z}'\mathbf{Z})] &= \boldsymbol{\mu} E[\phi(\chi_{p+2}^2(\Delta^2))]. \\ E[\mathbf{Z}\mathbf{Z}' \phi(\mathbf{Z}'\mathbf{Z})] &= \mathbf{I}_p E[\phi(\chi_{p+2}^2(\Delta^2))] + \boldsymbol{\mu}\boldsymbol{\mu}' E[\phi(\chi_{p+4}^2(\Delta^2))]. \end{aligned}$$

Theorem A.2.7. Let $\mathbf{Z} = (Z_1, \dots, Z_p)'$ be a p -dimensional vector distributed as $N_p(\boldsymbol{\mu}, \mathbf{I}_p)$, and let $\mathbf{W}$ be any positive definite symmetric matrix, then for any integrable function $\phi(\mathbf{Z}'\mathbf{Z})$, we have

$$E[\mathbf{Z}'\mathbf{W}\mathbf{Z} \phi(\mathbf{Z}'\mathbf{Z})] = \text{tr}(\mathbf{W}) E[\phi(\chi_{p+2}^2(\Delta^2))] + E[\phi(\chi_{p+4}^2(\Delta^2))] \boldsymbol{\mu}'\mathbf{W}\boldsymbol{\mu}.$$

We may obtain the following identities:

$$2E[\chi_{p+3}^{-4}(\Delta^2)] = E[\chi_{p+1}^{-2}(\Delta^2)] - E[\chi_{p+3}^{-2}(\Delta^2)]. \quad (\text{A.5})$$

$$\Delta^2 E[\chi_{p+4}^{-4}(\Delta^2)] = E[\chi_{p+2}^{-2}(\Delta^2)] - (p-2)E[\chi_{p+2}^{-4}(\Delta^2)], \quad (\text{A.6})$$

$$\begin{aligned} E[\chi_{p+2}^{-2}(\Delta^2)] &= \exp\left\{-\frac{\Delta^2}{2}\right\} \sum_{r \geq 0} \frac{\left(\frac{-\Delta^2}{2}\right)^r}{r!} \frac{1}{p+2r} \\ &= E_r(p+2r)^{-1}. \end{aligned} \quad (\text{A.7})$$

$$\begin{aligned} E[\chi_{p+2}^{-4}(\Delta^2)] &= \exp\left\{-\frac{\Delta^2}{2}\right\} \sum_{r \geq 0} \frac{\left(\frac{-\Delta^2}{2}\right)^r}{r!} \frac{1}{(p+2r)(p+2r-2)} \\ &= E_r[(p+2r)(p+2r-2)]^{-1}. \end{aligned} \quad (\text{A.8})$$

$$E[\chi_{p+2}^{-2} I(\chi_{p+2}^2(\Delta^2) < c)] = E_r(p+2r)^{-1} H_{p+2r}(c; 0). \quad (\text{A.9})$$

$$E[\chi_{p+2}^{-4} I(\chi_{p+2}^2(\Delta^2) < c)] = E_r(p+2r)^{-1} (p-2+2r)^{-1} H_{p-2+2r}(c; 0). \quad (\text{A.10})$$

where E_r stands for the expectation with respect to a Poisson variable r with mean $\frac{1}{2}\Delta^2$.

A.3 Some Definitions and Results On Matrices

We summarize some of the well-known definitions and theorems of matrix.

- **Rank of a Matrix**

The rank $r(\mathbf{A})$ of a matrix $\mathbf{A}$ is the maximum number of independent rows (or, the maximum number of independent columns).

- **Matrix Trace**

The trace of an $n \times n$ square matrix $\mathbf{A}$ is defined to be $tr\mathbf{A} = \sum_{i=1}^n a_{ii}$, the sum of the diagonal elements.

- **Idempotent Matrix**

A matrix $\mathbf{A}$ is said to be idempotent if $\mathbf{A}^2 = \mathbf{A}$. If $\mathbf{A}$ is idempotent matrix, then $tr\mathbf{A} = r(\mathbf{A})$.

- **Positive (Semi) Definite Matrix**

A matrix $\mathbf{A}$ is said to be positive (semi) definite iff $\mathbf{x}'\mathbf{A}\mathbf{x} > (\geq) 0$ for all non-null vectors $\mathbf{x}$.

- **Characteristic Roots of A Matrix**

The characteristic roots of a $p \times p$ matrix $\mathbf{A}$ are the p roots of the polynomial $|\mathbf{A} - \lambda\mathbf{I}|$, where λ is a scalar.

- **Direct Product**

Let $\mathbf{A} = (a_{ij})$ be a $p \times p$ matrix and $\mathbf{B} = (b_{\alpha\beta})$ be a $q \times q$ matrix. The $pq \times pq$ matrix with $a_{ij}b_{\alpha\beta}$ as the element in the i, α th row and the j, β th column is called

the direct product of $\mathbf{A}$ and $\mathbf{B}$ and is denoted by $\mathbf{A} \otimes \mathbf{B}$; that is

$$\mathbf{A} \otimes \mathbf{B} = \begin{bmatrix} a_{11}\mathbf{B} & a_{12}\mathbf{B} & \cdots & a_{1p}\mathbf{B} \\ \vdots & \vdots & \vdots & \vdots \\ a_{p1}\mathbf{B} & a_{p2}\mathbf{B} & \cdots & a_{pp}\mathbf{B} \end{bmatrix} \quad (\text{A.11})$$

Theorem A.3.1. *Given an $(n \times n)$ idempotent matrix $\mathbf{A}$, then if $\mathbf{A}$ is of rank r , $\mathbf{A}$ has r characteristic roots equal to 1 and $(n - r)$ roots equal to zero.*

Theorem A.3.2. Courant Theorem

Let $\lambda_1, \dots, \lambda_p$ be the characteristic roots of a $p \times p$ matrix $\mathbf{A}$ such that $\min \lambda_i = \lambda_1$, $\max \lambda_i = \lambda_p$, and let $\mathbf{x}$ be a non-null vector. Then $\sup \left(\frac{\mathbf{x}' \mathbf{A} \mathbf{x}}{\mathbf{x} \mathbf{x}'} \right) = \lambda_p$ and $\inf \left(\frac{\mathbf{x}' \mathbf{A} \mathbf{x}}{\mathbf{x} \mathbf{x}'} \right) = \lambda_1$. Hence,

$$ch_{\min}(\mathbf{A}) \leq \frac{\mathbf{x}' \mathbf{A} \mathbf{x}}{\mathbf{x} \mathbf{x}'} \leq ch_{\max}(\mathbf{A}) \quad (\text{A.12})$$

where $ch_{\min}(\mathbf{A}) = \min_i \lambda_i$ and $ch_{\max}(\mathbf{A}) = \max_i \lambda_i$

Theorem A.3.3. *If Ψ is $(n \times n)$ non-singular matrix and δ, ℓ are $(n \times 1)$ column vectors, then*

$$\max_{\ell} \frac{\ell' \delta \delta' \ell}{\ell' \Psi \ell} = \delta' \Psi^{-1} \delta. \quad (\text{A.13})$$

Corollary A.3.1. *Let $n\Sigma_n$ be a $p \times p$ positive definite matrix, and $\mathbf{C}$ be a $p \times (p - 1)$ matrix and $\mathbf{1}_p = (1, \dots, 1)'$ such that $r(\mathbf{C}) = p - 1$ and $\mathbf{C}' \mathbf{1}_p = 0$. Then*

$$\begin{aligned} \mathbf{C}(\mathbf{C}' n\Sigma_n \mathbf{C})^{-1} \mathbf{C}' &= (n\Sigma_n)^{-1} - \frac{(n\Sigma_n)^{-1} \mathbf{1}_p \mathbf{1}_p' (n\Sigma_n)^{-1}}{\mathbf{1}_p' (n\Sigma_n)^{-1} \mathbf{1}_p} \\ &= (n\Sigma_n)^{-1} \left(\mathbf{I}_p - \frac{\mathbf{1}_p \mathbf{1}_p' (n\Sigma_n)^{-1}}{\mathbf{1}_p' (n\Sigma_n)^{-1} \mathbf{1}_p} \right) \end{aligned} \quad (\text{A.14})$$

is a positive semi definite matrix of rank $p - 1$.

For proof see Srivastava and Khatri page 19.

Bibliography

- [1] Anderson, T. A. (1984). An Introduction to Multivariate Statistical Analysis. John Wiley & Sons, Inc.
- [2] Bancroft, T. A. (1944). On biases in estimation due to the use of preliminary test of significance. The Annals of Mathematical Statistics, **15**(2), 190–204.
- [3] Berger, J. O. and Bock, M. E. (1976). Combining independent normal mean estimation problems with unknown variances. The Annals of Statistics, **4**(3), 642–648.
- [4] Bickel, P.J. (1964). On some alternative estimates for shift in the p-variate one sample probleml, Ann. Math. Stat. , **35**, 1079–1090.
- [5] Bickel, P.J. (1965). On some robust estimators of location, Ann. Math. Stat. , **36**, 847–858.
- [6] Bishop, Y., Fienberg, S. and Holland, P. (1975). Discrete Multivariate Analysis: Theory and Practice. MIT Press, Cambridge, Massachusetts.
- [7] Borenstein, M., Hedges, L. V., Higgins, J. P. T., and Rothstein, H. R. (2009). Introduction to Meta-Analysis. John Wiley& Sons, Ltd.
- [8] Cochran, W. G. (1937). Problems arising in the analysis of a series of similar experiments. Journal of the Royal Statistical Society (Supplement), 4, 102-118.

- [9] Cohen, J. (1977). Statistical Power Analysis for the Behavioral Science. Academic press. New York.
- [10] Cumming, G. and Finch, S. (2001). A primer on the understanding, use, and calculation of confidence intervals that are based on central and noncentral distributions. Educational and Psychological Measurement, **61** (4), 532-574.
- [11] Efron, B. (2009). Empirical Bayes estimates for large-scale prediction problems, Journal of the American Statistical Association, Theory and Methods, **104**(487), 1015–1028.
- [12] Efron, B. (2010). Large-scale Inference: Empirical Bayes Methods for Estimation, Testing, and Prediction. Cambridge University press.
- [13] Efron, B. and Morris, C. (1971). Limiting the risk of Bayes and empirical Bayes estimators I. The Bayes case. Journal of the American Statistical Association, **66**, 807–815.
- [14] Efron, B. and Morris, C. (1972). Limiting the risk of Bayes and empirical Bayes estimators II. The empirical Bayes case. Journal of the American Statistical Association, **67**, 130–139.
- [15] Efron, B. and Morris, C. (1973). Stein’s estimation rule and its competitors - An empirical Bayes approach. Journal of the American Statistical Association, **68**, 117–130.
- [16] Fisher, R. A. (1932). Statistical Methods for Research Workers (4th ed.) London: Oliver and Boyd. (1,3, 5, 14).
- [17] Fraser, D.A.S, Saleh, A. K. Md. Ehsanes. and Ji, K. (2010). Combining P-values: a definitive process. Journal of Statistical Research, **44** (1), 15–29.

- [18] Gastwirth, J. L, Chernoff, H. and Jones, M. V. (1967). Asymptotic distribution of linear combinations of functions of order statistics with applications to estimation. *The Annals of Mathematical Statistics*, **38**(1), 52–72.
- [19] Ghosh, M., Saleh, A.K.Md. Ehsanes and Sen, P.K. (1989). Empirical Bayes subset estimation in regression models. *Statistics & Decisions*, **7**, 15–35.
- [20] Gibbons, D. J. (1971). *Nonparametric Statistical Inference*. McGraw-Hill Book Company, New York .
- [21] Glass, G. V. (1976). Primary, secondary, and met-analysis research. *Educational Researcher*, **5**, 3–8.
- [22] Glass, G. V. (1977). Integrating findings: The Meta-analysis of research. *Review of Research in Education*, **5**(1), 351–379.
- [23] Gleser, L. J. and Olkin, I. (1994). Stochastically dependent effect sizes. *The handbook of research synthesis*. New York Russell Sage Foundation. pp 339–355 .
- [24] Gleser, L. J. and Olkin, I. (2007). Stochastically dependent effect sizes Department of Statistics Stanford University. Stanford, California 94305-4065.
- [25] Han, C. P., and Bancroft, T. A. (1968). On pooling means when variance is unknown. *J.Amer.Stat.Assoc.*, **63**, 1333-1342.
- [26] Hartung, J. , Knapp G. and Sinha, B. K. (2008). *Statistical Meta-Analysis with Application*. John Wiley & Sons.Inc
- [27] Hedges, L. V. (1981). Distribution theory for Glass’s estimator of effect size and related estimators. *Journal of Education Statistics*, **6**, 107–128.

- [28] Hedges. L.V. (1982). Estimation of effect size from a series of independent experiments. *Psychological Bulletin*, **92**(2), 490–499 .
- [29] Hedges, L. V. and Olkin, I. (1983). Clustering estimates of effect magnitude from independent studies. *Psychological Bulletin*, **93**, 563-573.
- [30] Hedges. L. V. and Olkin, I. (1983). Regression models in research synthesis. *The American Statistician*, **73**, 137-140.
- [31] Hedges. L. V. and Olkin, I. (1984). Nonparametric estimators of effect size in Meta-analysis . *Psychological Bulletin*, **46**, 573-580.
- [32] Hedges. L. V. and Olkin, I. (1985). *Statistical Methods for Meta-analysis* . Boston: Academic Press.
- [33] Hodges, J. L. and Lehmann, E. L. (1963). Estimates of location based on rank tests. *The Annals of Mathematical Statistics*, **34** , 598–611 .
- [34] Huber, P. J. (1964). *Robust Estimation of a Location Parameter*, AMS, **35**, 73-101.
- [35] Huber, P. J. (1981). *Robust Statistics*. McGraw-Hill Book Company. New York.
- [36] James, W. and Stein, C. (1961). Estimation with quadratic loss. *Proceeding of the fourth Berkeley symposium on Mathematical Statistics and probability*.(Ulcla press, Berkeley, California), 361-379.
- [37] Johnson, N. L. and Kotz, S. (1970). *Continuous univariate distributions-2*. A Wiley-inter-science publication.
- [38] Judge, G. G. and Bock, M. E. (1978). *Statistical implications of pre-test and Stein-rule estimators in Econometrics*. North-Holland Publishing Company. Amsterdam.

- [39] Kraemer, H. C. and Paik, M. (1979). A central t approximation to the noncentral t distribution. *Technometrics*, **21**(3), 357–360.
- [40] Kraemer, H. C. and Andrews, G. (1982). A Nonparametric Technique for Meta Analysis Effect Size Calculation. *Psychological Bulletin*, **91** (2), 404–412 .
- [41] Kraemer, H. C. (1983). Theory of estimation and testing of effect sizes: use in meta-analysis. *Journal of Educational Statistics*, **8** (2), 93–101.
- [42] Krauth, J. (1983). Nonparametric effect size estimation: a comment on Kraemer and Andrews. *Psychological Bulletin*, **94** (1), 190–192.
- [43] Laubscher, N.F. (1960). Normalizing the noncentral t and F distribution. *The Annals of Mathematical Statistics*, **31**, 107–116.
- [44] Lecoutre, B. (2007). Another look at confidence intervals for the noncentral t distribution., *Journal of Modern Applied Statistical Methods*, **6** (1), 1105–1112.
- [45] Pearson, K. (1904). Report on certain enteric fever inoculation statistics. *British Medical Journal*, **2**, 1243-1246.
- [46] Peto, R. (1987). Why do we need systematic overview of randomized trials? *Stat. Med.* , **6**, 233-240.
- [47] Puri, M.L. and Sen, P. K. (1971). *Nonparametric Methods in Multivariate Analysis*. John Wiley & Sons, Inc.
- [48] Puri, M.L. and Sen, P. K. (1986). *Nonparametric Methods in General Linear Models*. Wiley& Sons, Inc.
- [49] Randales, C. and Wolfe, M. (1979). *Introduction to the Theory of Nonparametric Statistics*. Wiley& Sons Inc.

- [50] Rao, C. R. (1973). Linear Statistical Inference and Its Applications. Wiley, New York.
- [51] Robbins, H.(1956). An Empirical Bayes Approach to Statistics. Proceedings of the Third Berkeley Symposium on Mathematical Statistics and Probability, **1** : Contributions to the Theory of Statistics: 157-163.
- [52] Saleh, A. K. Md. Ehsanes, Hassanein, K. M. & Edward, F. B. (2001). Stability analysis of pharmaceutical products in multiple batches: preliminary test and shrinkage approach. Applied Statistical Science **V**, 21–37.
- [53] Saleh, A. K. Md. Ehsanes. (2006). Theory of Preliminary Test and Stein-Type Estimation with Applications. John Wiley & Sons, Inc., publication.
- [54] Saleh, A. K. Md. Ehsanes; Hassanein, K. M. & Brown, E. F. (1985). Optimum spacings for the joint estimation and tests of hypothesis of location and scale parameters of Cauchy distribution. Communications in Statistics- Theory and Methods, **14**(1), 247–254.
- [55] Saleh, A. K. Md. Ehsanes, Hassanein, K. M., Hassanein, R. S. & Kim, H. M. (2006). Quasi-empirical bayes methodology for improving meta-analysis. Journal of Biopharmaceutical Statistics, **16**, 77-90.
- [56] Sarhan, E. Ahmed and Greenberg, G. Bernard .(1962). Contributions to Order Statistics. John Wiley & Sons, Inc.
- [57] Srivastava, M.S. (1979). An Introduction to Multivariate Statistics. Elsevier North Holland, Inc.
- [58] Srivastava, M.S. and Saleh, A. K. Md. Ehsanes. (2005). Estimation of the mean vector of a multivariate normal distribution: subspace hypothesis. Journal of Multivariate Analysis, **96**, 55–72.

- [59] Stein, C. (1955). Inadmissibility of the usual estimator for the mean of a multivariate normal distribution. In Proc, Third Berkeley Symp. Math. Statist. Probab. bf1, 197-206. University of California Press, Berkeley.
- [60] Stein, C. (1981). Estimation of the mean of a multivariate distribution. Ann. Statist., **9**, 1135-1151.
- [61] Tippett, L. H. C. (1931). The Methods of Statistics. London: Williams & Norgate.
- [62] Yates, F. and Cochran, W. G. (1938). The analysis of groups of experiments. Journal of Agricultural Science, **28**(1), 556-580.
- [63] Zhang, Z. and Schoeps, N. (1997). On robust estimation of effect size under semi-parametric models. Psychometrika, **62** (2), 201-214 .