1

1.0

1.1

1.25

1.4

1.6

1.8

2.0

2.2

2.5

MICRO

National Library
of Canada

Bibliothèque nationale
du Canada

Canadian Theses Service

Service des thèses canadiennes

Ottawa, Canada
K1A 0N4

NOTICE

The quality of this microform is heavily dependent upon the quality of the original thesis submitted for microfilming. Every effort has been made to ensure the highest quality of reproduction possible.

If pages are missing, contact the university which granted the degree.

Some pages may have indistinct print especially if the original pages were typed with a poor typewriter ribbon or if the university sent us an inferior photocopy.

Previously copyrighted materials (journal articles, published tests, etc.) are not filmed.

Reproduction in full or in part of this microform is governed by the Canadian Copyright Act, R.S.C. 1970, c. C-30.

AVIS

La qualité de cette microforme dépend grandement de la qualité de la thèse soumise au microfilmage. Nous avons tout fait pour assurer une qualité supérieure de reproduction.

S'il manque des pages, veuillez communiquer avec l'université qui a conféré le grade.

La qualité d'impression de certaines pages peut laisser à désirer, surtout si les pages originales ont été dactylographiées à l'aide d'un ruban usé ou si l'université nous a fait parvenir une photocopie de qualité inférieure.

Les documents qui font déjà l'objet d'un droit d'auteur (articles de revue, tests publiés, etc.) ne sont pas microfilmés.

La reproduction, même partielle, de cette microforme est soumise à la Loi canadienne sur le droit d'auteur, S.R.C. 1970, c. C-30.

**ESTIMATION STRATEGIES
UNDER
UNCERTAIN PRIOR INFORMATION**

by

SYED EJAZ AHMED, B.S.(HONS.), M.S., M.Sc.

A thesis submitted to
the faculty of Graduate Studies and Research
in partial fulfillment of the requirements for the degree of

Doctor of Philosophy

Department of Mathematics and Statistics
Carleton University
Ottawa, Canada

August, 1987

©Copyright, 1987

S. E. Ahmed

Permission has been granted to the National Library of Canada to microfilm this thesis and to lend or sell copies of the film.

The author (copyright owner) has reserved other publication rights, and neither the thesis nor extensive extracts from it may be printed or otherwise reproduced without his/her written permission.

L'autorisation a été accordée à la Bibliothèque nationale du Canada de microfilmer cette thèse et de prêter ou de vendre des exemplaires du film.

L'auteur (titulaire du droit d'auteur) se réserve les autres droits de publication; ni la thèse ni de longs extraits de celle-ci ne doivent être imprimés ou autrement reproduits sans son autorisation écrite.

ISBN 0-315-39362-9

ESTIMATION STRATEGIES

UNDER

UNCERTAIN PRIOR INFORMATION

Syed Ejaz Ahmed

The undersigned recommend to the Faculty of Graduate Studies and Research acceptance of the thesis

Estimation strategies

under

Uncertain Prior Information

submitted by Syed Ejaz Ahmed, B.S.(HONS.), M.S., M.Sc., in partial fulfillment of the requirements for the degree of Doctor of Philosophy.

Dr. A. K. Md. E. Saleh, Thesis supervisor

Dr. P. K. Sen, External Examiner

*Dr. G. W. L. Garner, Chairman,
Department of Mathematics and Statistics*

ABSTRACT

The thesis deals with the problem of pooling means in correlated and uncorrelated cases together with the estimation of the intercept parameter in a simple linear model in the presence of uncertain prior information about the slope parameter. In particular, the estimation problems considered are:

- (i) the mean and the mean vector of a univariate and a multivariate normal distribution in a two-sample problem respectively
- (ii) the component mean and the component mean vector of a bivariate and a multivariate population respectively
- (iii) the intercept and the intercept vector of simple linear and multivariate normal models respectively.

In each univariate model, an estimator based on a preliminary test of significance is constructed, and the efficiency performance relative to a classical estimator is investigated. A max-min rule for the size of the preliminary test-of significance is presented.

For multivariate models, shrinkage estimators are constructed along with the preliminary test estimators.

Two measures of relative efficiency are considered; one based on the quadratic loss function, and the other on the determinant of the mean square error matrix. It is demonstrated that the shrinkage estimators dominates the classical estimator, whereas none of the shrinkage estimators and the preliminary test estimator dominate each other. The range in the parametric space where preliminary test estimator dominates shrinkage is investigated theoretically and computationally. It is found that the shrinkage estimator outperforms the preliminary test estimator except in a region around the null hypothesis. Moreover, for large values of α , the level of statistical significance, shrinkage-estimator dominates the preliminary test estimator uniformly.

This thesis is dedicated to my Mother and Father who have impressed upon me that an education is among the most valued gifts anyone can ever receive.

ACKNOWLEDGEMENTS

First of all, I wish to express my deep gratitude to almighty Allah, whose divine help enabled me to complete this thesis. In this regard I wish to extend my sincere gratitude to my thesis supervisor, Dr. A. K. Md. E. Saleh, who first introduced me to the subject of estimation theory under uncertain prior information with his characteristic enthusiasm, and inspired my interest in this area. He not only proposed problems of this thesis, but guided me to its successful completion, and provided invaluable supervision. Indeed, this job was not for a lesser man. It is a distinct pleasure to acknowledge my indebtedness to Dr. J. N. K. Rao, a member of my supervisory committee, for critical and constructive comments on the present work and for his encouragement in my academic pursuits at this university and elsewhere. I thank Dr. M. Csörgő, a member of my supervisory committee, for his moral support and inspiration to work even during religious holidays. I am also thankful to Dr. C. W. L. Garner, the chairman of the department for providing scholarships and research grants. Indeed, he is the one, who encouraged me to come to Carleton University for doctoral work. My special thanks go to Dr. J.J. Hubert from the University of Guelph, for proof-reading and improving the quality of this thesis. Thanks are also due to Faiz, Diane, Vera, Raj, Rod and Ali for many stimulating conversations. I wish to thank my best friend Hussain for his confident support over the years. My heartfelt thanks are extended to the Siddique family, from them I received the same love and warmth as my own family. I am also thankful to Maria Lavdas and Jackie Bell, for efficient and careful typesetting of perhaps a most difficult manuscript. The last, but not the least, is my gratitude due to my family, especially to my elder brother, Laique, for his encouragement over the years. His brotherly help made the completion of my doctoral work possible. A part of this research is supported by Natural Science and Engineering Research Council of Canada.

TABLE OF CONTENTS

Abstract	iii
Acknowledgements	iv
Table of Contents	v
List of Tables	vii
List of Figures	ix
Nomenclature	x
Notation	xi
 Chapter 1: INTRODUCTION	
 Chapter 2: ESTIMATION STRATEGY USING A PRELIMINARY TEST IN SOME UNIVARIATE NORMAL MODELS	
2.1. Introduction	12
2.2. Pooling Means	13
2.2.1. Estimation Strategies	14
2.2.2. Redefined Preliminary Test Estimator	17
2.2.3. Efficiency Analysis and Recommendations	20
2.3. Pooling Component-Means	26
2.3.1. Estimation Strategies	26
2.3.2. Redefined Preliminary Test Estimator	28
2.3.3. Efficiency Analysis and Recommendations	29
2.4. Simple Linear Model	32
2.4.1. Estimation Strategies	33
2.4.2. Redefined Preliminary Test Estimator	34
2.4.3. Efficiency Analysis and Recommendations	35
2.5. Concluding Remarks	46
 Chapter 3: ESTIMATION STRATEGIES FOR THE MEAN-VECTOR OF A MULTIVARIATE NORMAL DISTRIBUTION IN A TWO-SAMPLE PROBLEM	
3.1. Introduction	47
3.2. Estimation Strategies	48

3.3. Bias and Risks	49
3.3.1. Risk Analysis	55
3.3.2. Efficiency Analysis and Recommendations	57
3.4. Mean Square Error Matrices	81
3.4.1. Generalised Efficiency Analysis and Recommendations	84
3.5. Concluding Remarks	95
 Chapter 4: ESTIMATION STRATEGIES FOR THE COMPONENT MEAN-VECTOR IN A 2p-VARIATE NORMAL DISTRIBUTION	
4.1. Introduction	97
4.2. Estimation Strategies	98
4.3. Bias and Risks	100
4.3.1. Risk Efficiency Analysis and Recommendations	104
4.4. Mean Square Error Matrices	108
4.4.1. Generalised Efficiency Analysis	109
4.5. Concluding Remarks	110
 Chapter 5: ESTIMATION STRATEGIES FOR THE INTERCEPT VECTOR IN A SIMPLE LINEAR MULTIVARIATE NORMAL REGRESSION MODEL	
5.1. Introduction	111
5.2. Estimation Strategies	112
5.3. Bias and Risks	114
5.3.1. Risk Analysis	119
5.3.2. Risk Efficiency Analysis and Recommendations	120
5.4. Mean Square Matrices	163
5.4.1. Generalised Efficiency Analysis and Recommendations	167
5.5. Concluding Remarks	184
 Chapter 6: CONCLUSIONS AND AREAS OF FURTHER STUDY	 185
 APPENDIX A	 189
 APPENDIX B	 190
 REFERENCES	 195

LIST OF TABLES

Table	Description	page
1	Maximum and minimum guaranteed efficiencies	24
2	Maximum and minimum guaranteed efficiencies for $N=8(4)20$ and 28	41
3	Maximum and minimum guaranteed efficiencies for $p=2$, and various sample sizes	62
4	Maximum and minimum guaranteed efficiencies for $p=3$, and various sample sizes	64
5	Maximum and minimum guaranteed efficiencies for $p=4$, and various sample sizes	66
6	Maximum and minimum guaranteed efficiencies for $p=5$, and various sample sizes	68
7	Maximum relative efficiencies of PTMLE, RMLE, and SMLE and the intersecting efficiencies for PTMLE and SMLE for each α with corresponding Δ_α -values: $p=3$	75
8	Maximum relative efficiencies of PTMLE, RMLE, and SMLE and the intersecting efficiencies for PTMLE and SMLE for each α with corresponding Δ_α -values: $p=4$	77
9	Maximum relative efficiencies of PTMLE, RMLE, and SMLE and the intersecting efficiencies for PTMLE and SMLE for each α with corresponding Δ_α -values: $p=5$	79
10	Maximum and minimum guaranteed generalized efficiencies for $p=2$, and various sample sizes	87
11	Maximum and minimum guaranteed generalized efficiencies for $p=3$, and various sample sizes	89

12	Maximum and minimum guaranteed generalized efficiencies for $p=4$, and various sample sizes	91
13	Maximum and minimum guaranteed generalized efficiencies for $p=5$, and various sample sizes	93
14	Maximum and minimum guaranteed efficiencies for $p=2$, $N=8(4)20,28$ with selected values of $\bar{x}^2/Q$	125
15	Maximum and minimum guaranteed efficiencies for $p=3$, $N=8(4)20,28$ with selected values of $\bar{x}^2/Q$	130
16	Maximum and minimum guaranteed efficiencies for $p=4$, $N=8(4)20,28$ with selected values of $\bar{x}^2/Q$	135
17	Maximum and minimum guaranteed efficiencies for $p=5$, $N=8(4)20,28$ with selected values of $\bar{x}^2/Q$	140
18	Maximum relative efficiencies of PTMLE, RMLE, and SMLE and the intersecting efficiencies for PTMLE and SMLE for each α with corresponding Δ_α -values and each $\bar{x}^2/Q$ -value: $p=3$	150
19	Maximum relative efficiencies of PTMLE, RMLE, and SMLE and the intersecting efficiencies for PTMLE and SMLE for each α with corresponding Δ_α -values and each $\bar{x}^2/Q$ -value: $p=4$	155
20	Maximum relative efficiencies of PTMLE, RMLE, and SMLE and the intersecting efficiencies for PTMLE and SMLE for each α with corresponding Δ_α -values and each $\bar{x}^2/Q$ -value: $p=5$	159
21	Maximum and minimum guaranteed generalized efficiencies for $p=3$, $N=8(4)20,28$ with selected values of $\bar{x}^2/Q$	170
22	Maximum and minimum guaranteed generalized efficiencies for $p=4$, $N=8(4)20,28$ with selected values of $\bar{x}^2/Q$	174
23	Maximum and minimum guaranteed generalized efficiencies for $p=5$, $N=8(4)20,28$ with selected values of $\bar{x}^2/Q$	179

LIST OF FIGURES

Figure	Description	Page
1	Relative efficiency of PTMLE to UMLE for various sample sizes	22
2	Relative efficiency of PTMLE to UMLE for various sample sizes, and $\bar{x}^2/Q=0.5$	39
3	Relative efficiency of PTMLE to UMLE for various sample sizes, with $p=2$	59
4	Relative efficiency of PTMLE to UMLE for various sample sizes, with $p=4$	60
5	Dominance picture of UMLE, RMLE, PTMLE and SMLE for $p=3$, $N=8$, $N=12$, and various α values	72
6	Dominance picture of UMLE, RMLE, PTMLE and SMLE for $p=4$, $N=8$, $N=12$, and various α values	73
7	Dominance picture of UMLE, RMLE, PTMLE and SMLE for $p=5$, $N=8$, $N=12$, and various α values	74
8	Relative efficiency of PTMLE to UMLE for various sample sizes, with $p=2$ (and $\bar{x}^2/Q=0.5$)	122
9	Relative efficiency of PTMLE to UMLE for various sample sizes, with $p=4$ (and $\bar{x}^2/Q=0.5$)	123
10	Dominance picture of UMLE, RMLE, PTMLE and SMLE for $p=3$, $N=20$, (and $\bar{x}^2/Q=0.5$)	147
11	Dominance picture of UMLE, RMLE, PTMLE and SMLE for $p=4$, $N=20$, (and $\bar{x}^2/Q=0.5$)	148
12	Dominance picture of UMLE, RMLE, PTMLE and SMLE for $p=5$, $N=20$, (and $\bar{x}^2/Q=0.5$)	149

NOMENCLATURE

UE	Unrestricted Estimator
RE	Restricted Estimator
PTE	Preliminary Test Estimator
SE	Shrinkage Estimator
UMLE	Unrestricted Maximum Likelihood Estimator
RMLE	Restricted Maximum Likelihood Estimator
PTMLE	Preliminary Test Maximum Likelihood Estimator
SMLE	Shrinkage Maximum Likelihood Estimator
<u>Var</u>	Variance
MSE	Mean Square Error
R.E.	Relative(Risk) Efficiency
GRE	Generalized Relative Efficiency

NOTATION

μ, θ, β	parameter (scalar)
$\boldsymbol{\mu}, \boldsymbol{\theta}, \boldsymbol{\beta}$	parameter (vector)
$\tilde{\mu}$	unrestricted maximum likelihood estimator (UMLE) of μ
$\hat{\mu}$	restricted maximum likelihood estimator (RMLE) of μ
$\hat{\mu}^{PT}$	preliminary test maximum likelihood estimator (PTMLE) of μ
$\hat{\mu}^S$	shrinkage maximum likelihood estimator (SMLE) of μ
Δ	non-centrality parameter
$N(\mu, \sigma^2)$	normal distribution with mean μ and variance σ^2
$N_p(\boldsymbol{\mu}, \boldsymbol{\Sigma})$	p -variate normal distribution with mean vector $\boldsymbol{\mu}$ and covariance matrix $\boldsymbol{\Sigma}$
$\hat{\mu} > \tilde{\mu}$	means that $\hat{\mu}$ dominates $\tilde{\mu}$
$R.E.(\hat{\mu} : \tilde{\mu})$	risk efficiency of $\hat{\mu}$ relative to $\tilde{\mu}$
$GRE(\hat{\mu} : \tilde{\mu})$	generalized efficiency of $\hat{\mu}$ relative to $\tilde{\mu}$
$R(\hat{\mu}; \boldsymbol{\Sigma}^{-1})$	risk of $\hat{\mu}$
$\Gamma(\hat{\mu})$	mean square error matrix of $\hat{\mu}$

CHAPTER 1

INTRODUCTION

In problems of statistical inference (estimation and test of hypothesis), *prior information* on some of the parameters in a statistical model generally leads to an improved inference procedure for other parameters of interest. The known prior information is generally incorporated in the model in the form of a constraint, giving rise to restricted models. The analysis of such *restricted models* leads to an improved statistical procedure than in the *unrestricted model* case when such constraints hold. On the other hand, the validity and the efficiency of the restricted model analysis retain their validity over the restricted parameter space induced by the constraint, while unrestricted model analysis is valid over the entire parameter space. Therefore, the results of an analysis of the restricted and unrestricted models should be weighted against loss of efficiency and validity of the constraints in deciding a choice between two extreme inference techniques, when a full confidence may not be in the prior information.

It is not unusual to encounter problems of statistical inference with *uncertain prior information* in a statistical model. From outside considerations, some prior information may be imposed on the model but there may be reasons to suspect their validity so as to recommend restricted model analysis. Thus, the *uncertain prior information* in the form of a constraint may be regarded as a "nuisance parameter" in the statistical inference of the model. To alleviate this uncertainty of the prior information in the model, it is natural to follow Fisher's recipe of elimination of the nuisance parameter by a preliminary test on

the validity of the uncertain prior information in the form of parametric constraints and then choose between the *restricted* and *unrestricted* inference procedures depending on the outcome of the preliminary test. This brings about a compromised inference procedure between two extremes. This line of thought was first proposed by Bancroft (1944) when studying the consequences of incorporating a preliminary test in pooling several estimates of variances in an analysis of variance table. Such a preliminary test in the point estimation of a multi-parameter situation arose indirectly in the study of the admissibility question of the sample mean-vector in the p -variate normal distribution by Stein (1956) and James and Stein (1961). Generally, the James and Stein (1961) procedure may be regarded as a smooth-version of the Bancroft (1944) procedure. The parametric inference procedure incorporating preliminary tests dates back 40 years and always assumed normality; see for example, Bancroft and Han (1977) and Judge and Bock (1978). Non-normal theory and asymptotic theory of point-estimation including robustness questions and test of hypothesis have been developed by Saleh and Sen in the last decade. For an excellent account of these developments, based on a broad class of rank order statistics see Saleh and Sen (1982a, 1982b, 1983a, 1983b, 1985a, 1985b, 1986a and 1986b) and Sen and Saleh (1985). For recent developments in the finite sample, parametric theory we refer to Judge and Bock (1978), Berger (1980, 1985), Anderson (1984) and Arnold (1981), among others.

Let Y be a vector-valued random variable, with distribution function $F(y; \theta)$ where $\theta_{(p+q) \times 1} = (\theta_{(1 \times p)}^{(1)'}, \theta_{(1 \times q)}^{(2)'})'$, a $(p+q)$ dimensional vector. The basic problem is to estimate $\theta^{(1)}$ when it is apriori suspected that $\theta_{(q \times 1)}^{(2)} = \theta_{0(q \times 1)} = 0_{q \times 1}$ (without loss of generality). Thus we may define the following estimators of $\theta^{(1)}$:

- (i) Unrestricted estimator (UE): Let $\tilde{\theta}_n = (\tilde{\theta}_n^{(1)'}, \tilde{\theta}_n^{(2)'})'$ be an optimum estimator (say maximum likelihood estimator) from a sample of size n of $\theta = (\theta^{(1)'}, \theta^{(2)'})'$. Then $\tilde{\theta}_n^{(1)}$

is the corresponding estimator of $\theta^{(1)}$, called the *unrestricted estimator* of $\theta^{(1)}$ in the full model.

- (ii) **Restricted estimator (RE):** If $\theta^{(2)} = 0$, then $\theta = (\theta^{(1)'}, 0')'$ and let $\hat{\theta}_n^{(1)}$ be an optimum estimator of $\theta^{(1)}$ in the restricted model, and let $\tilde{\theta}_n^{(1)}$ be the restricted estimator of $\theta^{(1)}$.

The restricted estimator $\tilde{\theta}_n^{(1)}$ performs better than the unrestricted estimator $\hat{\theta}_n^{(1)}$ when $\theta^{(2)} = 0$, but as $\theta^{(2)}$ moves away from the pivot $\theta^{(2)} = 0$, $\tilde{\theta}_n^{(1)}$ may be considerably biased, inefficient and inconsistent, while the performance of $\hat{\theta}_n^{(1)}$ remains steady over such departures. For this reason, often it is desirable to develop estimators which are a compromise between $\tilde{\theta}_n^{(1)}$ and $\hat{\theta}_n^{(1)}$ under an *uncertain prior information* $\theta^{(2)} = 0$ and incorporating a preliminary test (PT) of the null hypothesis $H_0: \theta^{(2)} = 0$.

- (iii) **Preliminary test estimator (PTE):** Let L_n be an optimal test statistic for the null hypothesis $H_0: \theta^{(2)} = 0$ against $H_A: \theta^{(2)} \neq 0$ based on the sample size n , then the preliminary test estimator of $\theta^{(1)}$ is defined by.

$$\hat{\theta}_n^{(1)PT} = \hat{\theta}_n^{(1)} I(L_n < l_{n,\alpha}) + \tilde{\theta}_n^{(1)} I(L_n \geq l_{n,\alpha}) \quad (1.1)$$

$$= \tilde{\theta}_n^{(1)} + (\hat{\theta}_n^{(1)} - \tilde{\theta}_n^{(1)}) I(L_n < l_{n,\alpha}), \quad (1.2)$$

where $I(A)$ is the indicator function of the set A and $l_{n,\alpha}$ is the upper α -percentile of the distribution of L_n under H_0 . That is to say, $\hat{\theta}_n^{(1)PT} = \hat{\theta}_n^{(1)}$ if H_0 is tenable and equals $\tilde{\theta}_n^{(1)}$ if H_0 is untenable at a α -level of significance with critical value $l_{n,\alpha}$ ($0 < \alpha < 1$).

Finally, another estimator may be defined by incorporating the test-statistic L_n , which is known as the shrinkage estimator and it is also a compromise between $\hat{\theta}_n^{(1)}$ and $\tilde{\theta}_n^{(1)}$.

(iv) Shrinkage estimator (SE). A simple form of a shrinkage estimator of $\theta^{(1)}$ is defined by

$$\hat{\theta}_n^{(1)S} = \hat{\theta}_n^{(1)} + (1 - cL_n^{-1})(\hat{\theta}_n^{(1)} - \bar{\theta}_n^{(1)}) \quad (1.3)$$

$$= \bar{\theta}_n^{(1)} + c(\hat{\theta}_n^{(1)} - \bar{\theta}_n^{(1)})L_n^{-1}, \quad (1.4)$$

where c is a positive constant and is called the shrinkage factor and depends on the dimension (p) of the parameter $\theta^{(1)}$. This will be clearly defined in the specific problems we shall discuss later.

The shrinkage estimator $\hat{\theta}_n^{(1)S}$ is a smooth version of the preliminary test estimator $\hat{\theta}_n^{(1)PT}$. Equations (1.2) and (1.4) are similar looking where the test-statistic L_n adjusts for the departure of the parameter $\theta^{(2)}$ from the pivot 0. When L_n is large, PTE and SE approach the same value $\bar{\theta}_n^{(1)}$, while for small values of L_n there is a difference. Thus, we shall study critically the performance of the four estimators of the type defined above using the following two criteria:

- (i) Risk (expected value of the loss function) of the estimators under the loss function defined by the Mahalanobis distance of the estimator $\theta_n^{*(1)}$ from $\theta^{(1)}$:

$$L(\theta_n^{*(1)}, \theta^{(1)}) = (\theta_n^{*(1)} - \theta^{(1)})' V_{\bar{\theta}_n^{(1)}}^{-1} (\theta_n^{*(1)} - \theta^{(1)}) \quad (1.5)$$

where $V_{\bar{\theta}_n^{(1)}}$ is the covariance matrix of $\bar{\theta}_n^{(1)}$.

- (ii) Mean square error matrices of the estimators.

These two criteria lead to the definition of risk efficiency (R.E.) and the generalized relative efficiency (GRE) of one estimator against another:

Definition 1.1 Let R_1 and R_2 be the risks of the estimators $\hat{\theta}_1$ and $\hat{\theta}_2$ respectively.

The risk efficiency of $\hat{\theta}_2$ relative to $\hat{\theta}_1$ denoted by $R.E.(\hat{\theta}_2 : \hat{\theta}_1)$ is defined by the ratio

R_1/R_2 .

Definition 1.2 Let Γ_1 and Γ_2 be the mean square error matrices of $\hat{\theta}_1$ and $\hat{\theta}_2$ respectively. Then the generalized relative efficiency of $\hat{\theta}_2$ relative to $\hat{\theta}_1$ is defined by the ratio of the p^{th} root of the determinant of the mean square error matrices; i.e., $GRE(\hat{\theta}_2 : \hat{\theta}_1) = |\Gamma_1|^{1/p} / |\Gamma_2|^{1/p}$, where p is the dimension of the parameter θ which is being estimated.

The mean square error matrix is useful to identify the components which contributes most to the total risk of an estimator. Thus, we have used both risk and mean square error matrix analysis of the four estimators to discover relative merits of each other.

The main purpose of this thesis is to study the classical problem of "pooling means". This problem has been studied by Bancroft (1944) in pooling estimates of the variance in an analysis of variance table via a preliminary test estimation procedure. However, Mosteller (1948) presented the analysis of a univariate problem of "pooling means" to understand various aspects of the different procedures. Thus, apart from the preliminary test estimation procedure, Mosteller (1948) also put forward the Bayes and empirical Bayes procedures and the benefits of each method. Kitagawa (1951) obtained the distribution and moments of the preliminary test estimator in the two sample problem. Bennett (1952) derived the expression for the 'bias' and mean square error of the preliminary estimator in the equal sample size case. Huntsberger (1955) generalized the preliminary test estimator of a mean in terms of weight functions.

A bivariate analogue of pooling data was considered by Asano and Sato (1962). Sato (1962) extended this problem to a multivariate case. Asano (1960a, 1960b, 1961) and Asano and Sugimura (1961) provide some applications of the preliminary test procedure in biometrical and pharmaceutical fields. Han and Bancroft (1968) considered the univariate

two-sample problem with equal, unknown variance and unequal sample sizes and proposed a preliminary test estimator and also a Bayes estimator. They provided a criterion to determine the optimum level of significance for an efficient preliminary test estimator.

Various other authors (Srivastava and Bancroft, 1969; Srivastava, 1967; Arnold, 1970) have considered the problem of estimating the population correlation coefficient in the presence of some uncertain prior information. Mehta and Gurland (1969) have performed a preliminary test on the correlation coefficient before testing equality of means of two normal populations. The performance of the preliminary test estimator has also been discussed by Cochran (1970). A regression-type estimator based on a preliminary test of significance of a bivariate normal distribution has been proposed by Han (1973b).

Now, we focus on some developments in the regression situation. Bancroft (1944) was the first to consider this problem. His work was concerned with the bias introduced by an estimator of a regression coefficient, if a second coefficient is tested to be zero. Shirafuji (1959) and Kitagawa (1959) have introduced such techniques in response surface analysis. Larson and Bancroft (1963a) studied the effect of sequentially testing several regression coefficients to be zero (delete the variable if the hypothesis is not rejected) on the bias and mean square error of the resulting predictand. Later they (1963b) derived the expression for the bias of the preliminary test estimator in prediction problems under uncertain prior information among two sets of explanatory variables; one set of variables are 'doubtful' for prediction. Sclove (1968) proposed an estimator based on the positive part of the shrinkage estimator of the regression coefficient. Kennedy and Bancroft (1971) discussed the model building problem in regression based upon repeated significance tests and presented two different model building procedures; called 'forward selection' and 'sequential deletion', in

the presence of some uncertain information. Ahsanullah and Saleh (1972) extended the Han and Bancroft (1968) methodology of the estimation of the intercept parameter to a simple regression model. Rahim (1976) considered the preliminary test estimation of the intercept parameter in a linear model by giving weights to prior belief in the uncertain prior information. Nonparametric estimation of the intercept in a simple linear model using rank estimators has been studied by Saleh and Sen (1978) which encompasses robust procedures and asymptotic theory. Nonparametric two-sample problems have been studied by Tamura (1965a, 1965b, 1967a, 1967b, 1967c). Bock *et al.* (1973a, 1973b) and Judge *et al.* (1973) have developed estimators based upon the preliminary test of significance for linear models in the presence of a general linear hypothesis. Leamer and Chamberlian (1976) gave Bayesian interpretations of the preliminary test estimator in regression models.

Various other authors (Paull, 1950; Bechhofer, 1951; Bozivich *et al.*, 1956; Srivastava and Bozivich 1962; Srivastava, 1964 and 1972) have studied the effect of a preliminary test of significance on the size and the power of the subsequent test. See also, Bancroft (1968).

Now, we survey some of the developments in sampling theory. Sukhatme and Tang (1970, 1975), Tang and Sukhatme (1972a, 1972b) proposed a sample size allocation method to strata in the stratified sampling situation, subsequent to a preliminary test of significance on strata variances. They also investigated the efficiency of the proposed allocation with respect to proportional allocation and modified Neyman allocation. A regression type estimator based on a preliminary test of significance that chooses between the difference estimator and regression estimator has been proposed by Grimes and Sukhatme (1980). Han (1973a) and Sisodia and Srivastava (1982) investigated the properties of the preliminary test estimator in a two-stage sampling situation when some information is available on an auxiliary variable. See also, Al-Bayyati and Arnold (1972).

This thesis considers the univariate and multivariate two-sample problem of pooling means assuming normality in three situations: (i) two independent samples with the same but unknown variance/covariance matrix, (ii) two correlated samples of equal size and, (iii) simple linear model structure which leads to (i) with a special choice of the design matrix. For the preliminary test and shrinkage estimation we choose the F-test (univariate case) and Hotelling's T^2 -test (again F-test related statistics) for the univariate and multivariate analysis of the problem. Thus, this thesis provides a unified approach to the problem of estimation in the presence of uncertain prior information.

This thesis mainly deals with the estimation of *multi-parameter* in *multivariate normal models* while, Judge and Bock (1978) considers the *multi-parameter* estimation problem in *univariate normal models*. Moreover, we have simultaneously provided analytical and computational comparison of the four estimators together with a max-min rule for the size of the preliminary test. These types of comparative performances results are not available in Judge and Bock (1978) and elsewhere.

In Chapter 2, we consider the univariate problem in three situations. First, we consider two independent samples of different sizes (N_1 and N_2) from two normal distributions $N(\mu_i, \sigma^2)$, $i = 1, 2$ with the same but unknown variance with *uncertain prior information* that μ_1 may be equal to μ_2 . Secondly, we consider a sample of size N from a bivariate normal distribution $N_2(\mu_1, \mu_2, \sigma_1^2, \sigma_2^2, \rho)$ with the uncertain prior information that μ_1 may be equal to μ_2 . In both cases our primary interest is to estimate μ_1 . Finally, we consider the simple linear model $y_j = \theta + \beta x_j + e_j$, $j = 1, 2, \dots, N$, where e_j is distributed normally $N(0, \sigma^2)$, with σ^2 unknown and the primary interest is the estimation of θ when it is a priori

suspected that β may be 0. A special case of this model is a two-sample (independent) case. In all the above three situations, preliminary test estimators have been defined via an F-test to achieve complete analysis which was not available in the study of Han and Bancroft (1968) and others.

Chapters 3 to 5 contain the main contributions in this thesis — the analysis and discussion of multivariate problems. First, we consider two independent samples of different sizes (N_1 and N_2) from two p -variate normal distributions $N_p(\mu^{(1)}, \Sigma)$ and $N_p(\mu^{(2)}, \Sigma)$, where Σ is unknown with the uncertain prior information that $\mu^{(1)}$ may be equal to $\mu^{(2)}$. Secondly, we consider a single sample of size N from $2p$ -variate normal distribution $N_{2p}(\mu^{(1)}, \mu^{(2)}; \Sigma)$ where Σ is unknown with uncertain prior information that $\mu^{(1)}$ may be equal to $\mu^{(2)}$. In both of these problems, the primary interest is the estimation of $\mu^{(1)}$. Finally, we consider the simple linear model $y_j = \theta + \beta x_j + e_j$, where $e_j \sim N_p(0, \Sigma)$, (Σ unknown) and y_j , θ and β are p -dimensional vectors. The primary interest is the estimation of θ when it is apriori suspected that $\beta = 0$ may hold. A special case of this problem is the two-sample problem. In all the above three cases, preliminary test estimators together with their shrinkage versions have been defined via Hotelling's T^2 -test (F-test) to achieve an *analysis which is new in the literature*. In this regard we may mention that Bancroft and Han (1976) considered the two-sample problem when $\Sigma = I_p$ and studied the preliminary test estimator and Saleh (1973) considered the case when Σ is completely unknown. In both papers no shrinkage estimation was considered. A special case of this model is the two-sample problem dealt with by Han and Bancroft (1968); see also, Kale and Bancroft (1967).

In all the three cases, the preliminary test estimators have been defined via F-test to achieve complete analysis *which is not available in the study of the problem elsewhere*.

Chapter 3 discusses the *unrestricted* maximum likelihood estimator, *restricted* maximum likelihood estimator, *preliminary test* maximum likelihood estimator and the *shrinkage* maximum likelihood estimator of $\mu^{(1)}$ and compares the last three estimators relative to the *unrestricted* maximum likelihood estimator using risk analysis as well as, generalized relative efficiency analysis. Tables for optimum selection of the level of the significance are presented for chosen sample sizes for the use of the preliminary estimators. Further, tables are presented giving the maximum relative risk efficiencies of the *restricted*, *preliminary test* and *shrinkage* maximum likelihood estimator relative to the *unrestricted* maximum likelihood estimator along with risk efficiency at the intersection of the preliminary test maximum likelihood and the shrinkage maximum likelihood estimator together with the intersecting point. This gives the range of applicability of the preliminary test maximum likelihood estimator and of the shrinkage maximum likelihood estimator for the practitioners.

In Chapter 4, we consider the estimation of the component-mean vector $\mu^{(1)}$ in a $2p$ -dimensional normal distribution $N_{2p}(\mu, \Sigma)$ where

$$\mu_{2p \times 1} = \begin{pmatrix} \mu_{p \times 1}^{(1)} \\ \mu_{p \times 1}^{(2)} \end{pmatrix} \quad \text{and} \quad \Sigma_{2p \times 2p} = \begin{pmatrix} \Sigma_{11} & \Sigma_{12} \\ \Sigma_{21} & \Sigma_{22} \end{pmatrix}$$

and $\Sigma_{11} = \Sigma_{22}$ and $\Sigma_{12} = \Sigma_{21}$, based on a sample of size N . Here, again we consider the *unrestricted*, *restricted*, *preliminary test* and *shrinkage* maximum likelihood estimators of $\mu^{(1)}$ when it is suspected apriori that $\mu^{(1)} = \mu^{(2)}$ may hold. This chapter concentrates on the theoretical development and the analysis of the estimators using risk efficiency as well as mean square error matrices criteria. This problem has been addressed by Ahsanullah (1976) based on preliminary test estimator for $p = 1$. Our theoretical development

considers general p where Ahsanullah's (1971) study reduces to a special case. We also demonstrate that the shrinkage estimator of $\mu^{(1)}$ outperforms the preliminary test estimators except in a region around the null hypothesis.

Chapter 5 contains the estimation theory of the intercept vector θ in a linear model $y_j = \theta + \beta x_j + e_j$, where $e_j \sim N_p(0, \Sigma)$ with Σ unknown and y_j, θ and β are the p -vectors with the design matrix

$$X = \begin{pmatrix} 1 & x_1 \\ 1 & \vdots \\ \vdots & \vdots \\ 1 & x_N \end{pmatrix}$$

when it is apriori suspected that $\beta = 0$ may hold. The two-sample problem of Chapter 3 is a special case of this problem. Once again, we consider the *unrestricted*, *restricted*, *preliminary test* and *shrinkage* maximum likelihood estimators of θ . Asymptotic analysis of this problem using rank statistics has been considered by Sen and Saleh (1979, 1983). We consider the analysis based on normal theory which confirms the finding of Sen and Saleh (1979, 1983). Tables of optimum selection of level of significance are presented for selected sample sizes and selected coefficients of variation of the regressors. Finally, efficiency analysis is presented using generalized relative efficiency as well as risk efficiency criteria.

The final Chapter 6 includes a discussion of some open problems and summarizes the main results of the thesis.

CHAPTER 2

ESTIMATION STRATEGY USING A PRELIMINARY TEST IN SOME UNIVARIATE NORMAL MODELS

2.1. INTRODUCTION

This chapter is divided into five sections. In Section 2.2 we consider the classical problem concerning of pooling means of two independent samples taken from two normal populations. The parameter of interest is μ_1 , the population mean of the first population. It is assumed that the population variance for both samples are equal, but the mean μ_2 may differ from μ_1 . Various estimation strategies for the estimation of μ_1 in the presence of the uncertain information that $\mu_1 = \mu_2$ are studied. The bias and mean square error of related estimators are given and their efficiencies relative to a classical estimator are investigated. The optimum rule for the preliminary test maximum likelihood estimator is discussed and tables for an appropriate choice of the size of the test of the null hypothesis $H_0 : \mu_1 = \mu_2$ are provided.

In Section 2.3 the problem of estimation of the mean μ_1 , one of the components of a bivariate normal distribution with equal marginal variances, from a sample of size N is considered. Various estimation strategies are presented in the presence of the uncertain information that the mean μ_1 and μ_2 of the two components of the bivariate normal distribution are equal. The bias and mean square error of a related estimator are derived. The optimum rule for the preliminary test maximum likelihood estimator is discussed in terms of its efficiency relative to a classical estimator of μ_1 .

In Section 2.4 a simple linear normal model is considered. Our primary focus will be on the estimation of the intercept parameter θ when it is apriori suspected that the slope parameter $\beta = 0$. Various estimation strategies have been developed for the intercept parameter. Exact bias and mean square error of related estimators are derived and their efficiencies relative to a classical estimator are studied. The optimum rule for the preliminary test estimator is discussed and tables for the appropriate choice the significance levels of that test are provided.

Section 2.5 summarizes the findings.

2.2. POOLING MEANS

The main purpose of this section is to discuss the classical problem concerning pooling of means. Suppose we have a sample of size N_1 from $N(\mu_1, \sigma^2)$ and a second independent random sample of size N_2 from $N(\mu_2, \sigma^2)$ are available and it is apriori suspected that μ_1 and μ_2 are identical. How should we estimate μ_1 ? Should we merely take the mean of the first sample or should we attempt to combine the means of the two samples in the hope of improving the estimate of μ_1 in some sense?

A variation of this problem is frequently encountered in the analysis of variance where various estimates of errors are sometimes pooled after an *initial test of significance on the homogeneity of the variances*. See, for example Bancroft (1944) and (1964). Through the discussion of the pooling of means various estimation procedures and their implications will come to light to adopt the best strategy in selecting the optimum procedure.

If two sets of similar measures of performances of two products which have been made from materials on the same machine by the same machine operators are available, should

we estimate performances of the products by the percentage of defectives on each the products or can we obtain a better picture by combining the results of the sampling on both products? These and similar questions are concerned with the question of pooling of means.

We assume in the following that $x_{11}, x_{12}, \dots, x_{1N_1}$ and $x_{21}, x_{22}, \dots, x_{2N_2}$ are the two samples for $N(\mu_1, \sigma^2)$ and $N(\mu_2, \sigma^2)$ respectively where μ_1, μ_2 and σ^2 are unknown parameters and we are interested in the estimation of μ_1 when it is apriori suspected that $\mu_1 = \mu_2$ may be true.

2.2.1. ESTIMATION STRATEGIES

In this section we discuss various estimation strategies for the estimation of μ_1 in the presence of an uncertain information and provide bias and mean square error expressions for related estimators.

If $\mu_1 = \mu_2$ is really true, then the maximum likelihood estimator of μ_1 is given by

$$\hat{\mu}_1 = \frac{N_1 \bar{\mu}_1 + N_2 \bar{\mu}_2}{N_1 + N_2}, \quad \bar{\mu}_i = N_i^{-1} \sum_{j=1}^{N_i} x_{ij}, \quad i = 1, 2. \quad (2.2.1)$$

If $\mu_1 \neq \mu_2$, then the maximum likelihood estimator of parameter of interest μ_1 is $\bar{\mu}_1$

and

$$E(\bar{\mu}_i) = \mu_i, \quad Var(\bar{\mu}_i) = \frac{\sigma^2}{N_i}. \quad (2.2.2)$$

Also, $\bar{\mu}_1$ is a uniformly minimum variance unbiased estimator (UMVUE). Moreover,

$$E(\hat{\mu}_1) = \mu_1 + \frac{N_2}{N_1 + N_2}(\mu_2 - \mu_1), \quad (\text{biased for } \mu_1) \quad (2.2.3)$$

$$MSE(\hat{\mu}_1) = \frac{\sigma^2}{N_1 + N_2} + \frac{N_2 \sigma^2 \delta^2}{N_1(N_1 + N_2)}, \quad \delta^2 = \frac{N_1 N_2}{N_1 + N_2} \left(\frac{\mu_2 - \mu_1}{\sigma} \right)^2. \quad (\text{unbounded in } \delta^2) \quad (2.2.4)$$

On the other hand, if $\mu_1 = \mu_2$ then

$$E(\hat{\mu}_1) = \mu_1, \quad Var(\hat{\mu}_1) = \frac{\sigma^2}{N_1 + N_2} < \frac{\sigma^2}{N_1} = Var(\tilde{\mu}_1). \quad (2.2.5)$$

Considering the problem as a game with nature we then have the following outcome matrix:

Statistician's strategy	Nature's strategy	
	$\mu_1 \neq \mu_2$	$\mu_1 = \mu_2$
$\mu_1 \neq \mu_2$	$\tilde{\mu}_1$ is UMVUE	$\tilde{\mu}_1$ is unbiased and has variance larger than MVUE
$\mu_1 = \mu_2$	$\hat{\mu}_1$ is biased and has smaller variance than MVUE	$\hat{\mu}_1$ is UMVUE

Thus, each hypothesis ($\mu_1 \neq \mu_2$) or ($\mu_1 = \mu_2$) involves a tradeoff between variance and bias and we must attempt a compromise between $\tilde{\mu}_1$ and $\hat{\mu}_1$ so that we have a balance in the variance and bias. The uncertain prior information ($\mu_1 = \mu_2$) is a nuisance parameter in the estimation problem. Thus, according to Fisher we should test it out with the available sample. Therefore it is logical to introduce a preliminary test for the null hypothesis $H_0 : \mu_1 = \mu_2$ against the alternative $H_A : \mu_1 \neq \mu_2$ to decide between $\hat{\mu}_1$ and $\tilde{\mu}_1$ according as H_0 is tenable or untenable. Han and Bancroft (1968) proposed the following preliminary test estimator (PTE):

$$\hat{\mu}_1^{PT} = \tilde{\mu}_1 I(|t| < t_{\alpha/2}) + \hat{\mu}_1 I(|t| > t_{\alpha/2}) \quad (2.2.6)$$

$$= \tilde{\mu}_1 + (\hat{\mu}_1 - \tilde{\mu}_1) I(|t| > t_{\alpha/2}), \quad (2.2.7)$$

where

$$t = (\bar{\mu}_2 - \bar{\mu}_1) / s_p \left(\frac{1}{N_1} + \frac{1}{N_2} \right)^{\frac{1}{2}}, \quad (2.2.8)$$

$$(N_1 + N_2 - 2)s_p^2 = \sum_{j=1}^{N_1} (x_{1j} - \bar{\mu}_1)^2 + \sum_{j=1}^{N_2} (x_{2j} - \bar{\mu}_2)^2 \quad (2.2.9)$$

and $t_{\alpha/2}$ is the upper $\alpha/2$ percentile of a student's t -distribution with $N_1 + N_2 - 2$ degrees of freedom and $I(A)$ denotes the indicator function of the set A . Note that $\hat{\mu}_1^{PT}$ is a convex combination of $\bar{\mu}_1$ and $\bar{\mu}_2$ and the test statistic t adjusts for the departure from the null hypothesis $H_0 : \mu_1 = \mu_2$. The bias of $\hat{\mu}_1^{PT}$ is given by

$$\begin{aligned} B(\delta) &= \delta \sigma \sqrt{\frac{N_2}{N_1(N_1 + N_2)}} \sqrt{\frac{t_{\alpha/2}^2}{t_{\alpha/2}^2 + N_1 + N_2 - 2}} \\ &\quad \cdot e^{\frac{\delta^2}{2} \{ (N_1 + N_2 - 2) / (t_{\alpha/2}^2 + N_1 + N_2 - 2) \}} \\ &\quad \cdot \left[\sum_{r=0}^{N_1 + N_2 - 4/2} \left(\frac{N_1 + N_2 - 2}{2t_{\alpha/2}^2} \right)^r \frac{\mu'_{2r+1}}{r!} \right] \end{aligned} \quad (2.2.10)$$

the MSE is given by

$$\begin{aligned} MSE(\hat{\mu}_1^{PT}) &= \frac{\sigma^2}{N_1} + e^{-\delta^2/2 \{ (N_1 + N_2 - 2) / (t_{\alpha/2}^2 + N_1 + N_2 - 2) \}} \\ &\quad \cdot \frac{\sigma^2 N_2}{N_1(N_1 + N_2)} \sqrt{\frac{t_{\alpha/2}^2}{t_{\alpha/2}^2 + N_1 + N_2 - 2}} \\ &\quad \cdot \left[\sum_{r=0}^{N_1 + N_2 - 4/2} \left(\frac{N_1 + N_2 - 2}{2t_{\alpha/2}^2} \right)^r \frac{1}{r!} \left(2\delta \sqrt{\frac{N_1 N_2}{(N_1 + N_2)}} \mu'_{2r+1} - \mu'_{2r+2} \right) \right] \end{aligned} \quad (2.2.11)$$

for $N = N_1 + N_2$ even and greater than 4 and

$$\mu'_k = \int_{-\infty}^{+\infty} \frac{Z^k}{\sigma_z \sqrt{2\pi}} e^{-\frac{1}{2} \left(\frac{Z - \mu_z}{\sigma_z} \right)^2} d_z \quad (2.2.12)$$

where

$$\mu_z = \delta \frac{t_{\alpha/2}^2}{t_{\alpha/2}^2 + N_1 + N_2 - 2} \quad (2.2.13)$$

and

$$\sigma_z^2 = t_{\alpha/2}^2 / (t_{\alpha/2}^2 + N_1 + N_2 - 2). \quad (2.2.14)$$

μ'_k is the k^{th} moment of $N(\mu_z, \sigma_z^2)$ which can be written as $\mu'_k = \sum_{j=0}^K \binom{k}{j} \mu_z^j \mu_j^0$, where μ_j^0 is moment of $N(0, 1)$. A detailed derivation of these results are given in Han and Bancroft (1968).

In the following, we redefined preliminary test estimator using an F-test.

2.2.2. REDEFINED PRELIMINARY TEST ESTIMATOR

A redefined preliminary test estimator is as follows:

$$\hat{\mu}_1^{PT} = \hat{\mu}_1 I(t^2 < F_{1,\nu}(\alpha)) + \bar{\mu}_1 I(t^2 \geq F_{1,\nu}(\alpha)) \quad (2.2.15)$$

$$= \bar{\mu}_1 + (\hat{\mu}_1 - \bar{\mu}_1) I(t^2 < F_{1,\nu}(\alpha)), \quad (2.2.16)$$

where $\nu = N_1 + N_2 - 2$ and $F_{1,\nu}(\alpha)$ is the upper α -percentile of a central F-distribution with $(1, \nu)$ degrees of freedom. Then we have the following theorem.

Theorem 2.2.1. The bias and MSE of $\hat{\mu}_1^{PT}$ are given by

$$B_3 = \sigma \delta \sqrt{\frac{N_2}{N_1(N_1 + N_2)}} H_{3,\nu}(F_{1,\nu}(\alpha); \delta^2), \quad (2.2.17)$$

and

$$\begin{aligned} MSE(\hat{\mu}_1^{PT}) = \frac{\sigma^2}{N_1} + \frac{\sigma^2 N_2}{N_1(N_1 + N_2)} \left[\delta^2 \{ 2H_{3,\nu}(F_{1,\nu}(\alpha); \delta^2) \right. \\ \left. - H_{5,\nu}(F_{1,\nu}(\alpha); \delta^2) \} - H_{3,\nu}(F_{1,\nu}(\alpha); \delta^2) \right], \end{aligned} \quad (2.2.18)$$

where $H_{\nu_1, \nu_2}(\cdot; \delta^2)$ is the cumulative distribution function of a non-central F-distribution with (ν_1, ν_2) degrees of freedoms and non-centrality parameter δ^2 .

PROOF.

$$\begin{aligned} B_3 = E(\hat{\mu}_1^{PT} - \mu_1) &= \frac{N_2}{N_1 + N_2} E(\tilde{\mu}_2 - \tilde{\mu}_1) I(t^2 < F_{1,\nu}(\alpha)) \\ &= \frac{N_2 \sigma}{N_1 + N_2} \sqrt{\frac{N_1 N_2}{N_1 N_2}} E[(y) I(\frac{y^2}{s_p^2 / \sigma^2} < F_{1,\nu}(\alpha))], \end{aligned} \quad (2.2.19)$$

where

$$y = \sqrt{\frac{N_1 N_2}{N_1 + N_2}} \left(\frac{\tilde{\mu}_2 - \tilde{\mu}_1}{\sigma} \right).$$

By Theorem A.1 from Appendix A

$$E(\hat{\mu}_1^{PT} - \mu_1) = \sigma \delta \sqrt{\frac{N_2}{N_1(N_1 + N_2)}} H_{3,\nu}(F_{1,\nu}^{(\alpha)}; \delta^2), \quad (2.2.20)$$

where

$$\delta^2 = \frac{N_1 N_2}{N_1 + N_2} \left(\frac{\mu_2 - \mu_1}{\sigma} \right)^2$$

Now, the mean square error

$$\begin{aligned} MSE(\hat{\mu}_1^{PT}) &= E(\hat{\mu}_1^{PT} - \mu_1)^2 = E \left[(\tilde{\mu}_1 - \mu_1) + \frac{N_2}{N_1 + N_2} (\tilde{\mu}_2 - \tilde{\mu}_1) I(t^2 < F_{1,\nu}(\alpha)) \right]^2 \\ &= E(\tilde{\mu}_1 - \mu_1)^2 + \frac{N_2^2}{(N_1 + N_2)^2} E[(\tilde{\mu}_2 - \tilde{\mu}_1)^2 I(t^2 < F_{1,\nu}(\alpha))] \\ &\quad + \frac{2N_2}{N_1 + N_2} E[(\tilde{\mu}_1 - \mu_1)(\tilde{\mu}_2 - \tilde{\mu}_1) I(t^2 < F_{1,\nu}(\alpha))] \end{aligned} \quad (2.2.21)$$

Using the conditional expectation of $(\tilde{\mu}_1 - \mu_1)$ given $(\tilde{\mu}_2 - \tilde{\mu}_1)$ which is $\frac{N_2}{N_1 + N_2} [(\mu_2 - \mu_1) - (\tilde{\mu}_2 - \tilde{\mu}_1)]$, (2.2.21) becomes

$$\begin{aligned} &= E(\tilde{\mu}_1 - \mu_1)^2 + \frac{N_2^2}{(N_1 + N_2)^2} E[(\tilde{\mu}_2 - \tilde{\mu}_1)^2 I(t^2 < F_{1,\nu}(\alpha))] \\ &\quad + 2 \frac{N_2^2}{(N_1 + N_2)^2} E\{[(\mu_2 - \mu_1) - (\tilde{\mu}_2 - \tilde{\mu}_1)](\tilde{\mu}_2 - \tilde{\mu}_1) I(t^2 < F_{1,\nu}(\alpha))\} \\ &= E(\tilde{\mu}_1 - \mu_1)^2 - \frac{N_2^2}{(N_1 + N_2)^2} E[(\tilde{\mu}_2 - \tilde{\mu}_1)^2 I(t^2 < F_{1,\nu}(\alpha))] \\ &\quad + \frac{2N_2^2}{(N_1 + N_2)^2} (\mu_2 - \mu_1) E[(\tilde{\mu}_2 - \tilde{\mu}_1) I(t^2 < F_{1,\nu}(\alpha))], \end{aligned} \quad (2.2.22)$$

where the second term may be evaluated as

$$\begin{aligned}
 & \frac{N_2^2}{(N_1 + N_2)^2} E[(\tilde{\mu}_2 - \tilde{\mu}_1)^2 I(t^2 < F_{1,\nu}(\tilde{\alpha}))] \\
 &= \frac{N_2^2}{(N_1 + N_2)^2} \left(\frac{N_1 + N_2}{N_1 N_2} \right) E[(y^2) I(\frac{y^2}{s_p^2/\sigma^2} < F_{1,\nu}(\alpha))] \\
 &= \frac{N_2 \sigma^2}{N_1(N_1 + N_2)} \{H_{3,\nu}(F_{1,\nu}(\alpha) + \delta^2 H_{5,\nu}(F_{1,\nu}(\alpha); \delta^2)\}.
 \end{aligned}$$

By applying Theorem A.2 from Appendix A, (2.2.22) reduces to

$$\begin{aligned}
 &= \frac{\sigma^2}{N_1} - \frac{\sigma^2 N_2}{N_1(N_1 + N_2)} \{H_{3,\nu}(F_{1,\nu}(\alpha); \delta^2) + \delta^2 H_{5,\nu}(F_{1,\nu}(\alpha); \delta^2)\} \\
 &\quad + \frac{2N_2}{N_1(N_1 + N_2)} \sigma^2 \delta^2 H_{3,\nu}(F_{1,\nu}(\alpha); \delta^2) \quad (2.2.23)
 \end{aligned}$$

$$\begin{aligned}
 &= \frac{\sigma^2}{N_1} + \frac{\sigma^2 N_2 \delta^2}{N_1(N_1 + N_2)} \{2H_{3,\nu}(F_{1,\nu}(\alpha); \delta^2) - H_{5,\nu}(F_{1,\nu}(\alpha); \delta^2)\} \\
 &\quad - \frac{N_2 \sigma^2}{N_1(N_1 + N_2)} H_{3,\nu}(F_{1,\nu}(F_{1,\nu}; \delta^2). \quad (2.2.24)
 \end{aligned}$$

We may remark that Han and Bancroft (1968) results are difficult to extend to higher dimensions and puts a *restriction* on the *choice of sample sizes*. Also, Mosteller (1948) considers only the equal sample case. Similar limitations are noticed in the work of Bennett (1956) and Kitagawa (1963). The formulae of the Theorem 2.2.1 are easily computable for any (N_1, N_2) choice.

The following theorem shows the dominance picture of estimators under the null hypothesis.

Theorem 2.2.2. Under the null hypothesis, $H_0 : \mu_1 = \mu_2$, the three estimators of μ_1 , namely, $\tilde{\mu}_1$, $\hat{\mu}_1$ and $\hat{\mu}_1^{PT}$, may be ordered according to the magnitudes of the mean square error (MSE) as follows:

$$\hat{\mu}_1 > \hat{\mu}_1^{PT} > \tilde{\mu}_1. \quad (2.2.25)$$

PROOF.

$$(i) \text{MSE}(\bar{\mu}_1) - \text{MSE}(\hat{\mu}_1^{PT}) = \frac{N_2 \sigma^2}{N_1(N_1 + N_2)} H_{3,\nu}(F_{1,\nu}(\alpha); 0) > 0. \quad (2.2.26)$$

$$(ii) \text{MSE}(\hat{\mu}_1^{PT}) - \text{MSE}(\hat{\mu}_1) = \frac{\sigma^2}{N_1(N_1 + N_2)} \{1 - H_{3,\nu}(F_{1,\nu}(\alpha); 0)\} > 0. \quad (2.2.27)$$

2.2.3. EFFICIENCY ANALYSIS AND RECOMMENDATIONS

The relative efficiency of the restricted maximum likelihood estimator (RMLE) to the unrestricted maximum likelihood estimator (UMLE) is

$$R.E.(\hat{\mu}_1 : \bar{\mu}_1) = \left(\frac{N_1}{N_1 + N_2} + \frac{N_2 \delta^2}{N_1 + N_2} \right)^{-1} \quad (2.2.28)$$

Similarly, the relative efficiency of μ_1^{PT} to RMLE and UMLE are given by

$$R.E.(\hat{\mu}_1^{PT} : \hat{\mu}_1) = \left(\frac{N_1}{N_1 + N_2} + \frac{N_2 \delta^2}{N_1 + N_2} \right) (1 + g(\delta))^{-1}; \quad (2.2.29)$$

$$R.E.(\hat{\mu}_1^{PT} : \bar{\mu}_1) = (1 + g(\delta))^{-1}, \quad (2.2.30)$$

where

$$g(\delta) = \frac{N_2}{N_1 + N_2} \left[\delta^2 \{2H_{3,\nu}(F_{1,\nu}(\alpha); \delta^2) - H_{5,\nu}(F_{1,\nu}(\alpha); \delta^2)\} - H_{3,\nu}(F_{1,\nu}(\alpha); \delta^2) \right] \quad (2.2.31)$$

Now, we present the analysis of the relative efficiencies. First we consider $\hat{\mu}_1$ relative to $\bar{\mu}_1$. We note that it is a monotonically decreasing function of δ^2 . At $\delta = 0$, the maximum relative efficiency is $1 + \frac{N_2}{N_1}$. Thus $\hat{\mu}_1$ is more efficient at the null hypothesis; and as $\delta^2 \rightarrow \infty$, the efficiency drops to zero, crossing the line $R.E. = 1$ at $\delta^2 = 1$. The relative efficiency equals 1 whenever $\delta^2 = 1$. However, the relative efficiency is ≥ 1 according as

$\delta^2 \leq 1$. Thus if one knows that $\delta^2 \leq 1$, then $\hat{\mu}_1$ is the better choice, otherwise $\tilde{\mu}_1$ should be chosen. Next, we compare $\hat{\mu}_1^{PT}$ to $\hat{\mu}_1$. First we note that under H_0

$$R.E.(\hat{\mu}_1^{PT} : \hat{\mu}_1) = \frac{N_1}{N_1 + N_2} \left[1 - \frac{N_2}{N_1 + N_2} H_{3,\nu}(F_{1,\nu}(\alpha); 0) \right]^{-1} \left(\geq \frac{N_1}{N_1 + N_2} \right) \quad (2.2.32)$$

while the relative efficiency of $\hat{\mu}_1^{PT}$ to $\tilde{\mu}_1$ is

$$R.E.(\hat{\mu}_1^{PT} : \tilde{\mu}_1) = \left[1 - \frac{N_2}{N_1 + N_2} H_{3,\nu}(F_{1,\nu}(\alpha); 0) \right]^{-1} (\geq 1). \quad (2.2.33)$$

Thus,

$$\frac{N_1}{N_1 + N_2} \leq R.E.(\hat{\mu}_1^{PT} : \hat{\mu}_1) \leq 1 \leq R.E.(\hat{\mu}_1^{PT} : \tilde{\mu}_1). \quad (2.2.34)$$

On the other hand, as δ^2 moves away from 0

$$R.E.(\hat{\mu}_1^{PT} : \hat{\mu}_1) \leq 1 \quad \text{according as} \\ \delta^2 \leq 1 + \left(\frac{N_1 + N_2}{N_2} \right) g(\delta) = \delta_{\alpha}^{*2} \quad (\text{say}). \quad (2.2.35)$$

Finally, as $\delta^2 \rightarrow \infty$, $R.E.(\hat{\mu}_1^{PT} : \hat{\mu}_1) \rightarrow 0$. Thus, except when δ^2 is small, $\hat{\mu}_1^{PT}$ is relatively more efficient than $\hat{\mu}_1$.

Finally, we consider the relative efficiency of $\hat{\mu}_1^{PT}$ relative to $\tilde{\mu}_1$. It is a function of α ($0 < \alpha < 1$) and δ^2 for given (N_1, N_2) . Let this be denoted by $E(\alpha, \delta^2)$. The function $E(\alpha, \delta^2)$, for $\alpha \neq 0$, has its maximum at $\delta^2 = 0$ with value E^* given by

$$E^* = \left[1 - \frac{N_2}{N_1 + N_2} H_{3,\nu}(F_{1,\nu}(\alpha); 0) \right]^{-1} (\geq 1). \quad (2.2.36)$$

This function decreases, crossing the line $E(\alpha, \delta^2) = 1$, to a minimum $E_0(\alpha, \delta_0)$ at $\delta^2 = \delta_0^2$, then increases towards 1 as δ^2 tends to infinity. Now, for $\delta^2 = 0$, and α varying, we have $\max_{0 \leq \alpha \leq 1} E(\alpha, 0) = E(0, 0) = 1 + \frac{N_2}{N_1}$. The value of $E(\alpha, 0)$ decreases as α increases. On the other hand, let $\alpha = 0$ and δ^2 vary; then the curves $E(0, \delta^2)$ and $E(1, \delta^2) = 1$ intersect at $\delta^2 = 1$. For a general α , $E(\alpha, \delta^2)$ and $E(1, \delta^2)$ will always intersect in the interval $0 \leq \delta^2 \leq 1$. The value of δ^2 at the intersection decreases as α increases. Therefore, for two different α , $E(\alpha_1, \delta^2)$ and $E(\alpha_2, \delta^2)$ will never intersect above $E(\alpha, \delta^2)$ but intersect somewhere below 1. Figure 1 gives some examples of this function for various sample sizes and $\alpha = 0.05, 0.1, 0.2$, and 0.5 .

$N_1 = 8, N_2 = 4$ and $\alpha = 0.05, 0.1, 0.2, 0.5$

$N_1 = 10, N_2 = 4$ and $\alpha = 0.05, 0.1, 0.2, 0.5$

$N_1 = 10, N_2 = 12$ and $\alpha = 0.05, 0.1, 0.2, 0.5$

$N_1 = 10, N_2 = 4$ and $\alpha = 0.05, 0.1, 0.2, 0.5$

FIG. 1 Relative efficiency of PTMLE to UMLE for various sample sizes.

In order to choose an estimator with maximum relative efficiency (i.e., minimum MSE and minimum bias), we adopt the following rule: If it is known that $0 \leq \delta^2 \leq 1$, the estimator $\hat{\mu}_1$ is chosen since $E(0, \delta^2)$ is the largest for all δ^2 in this range. However, generally δ^2 is unknown (and also may not be in $[0, 1]$). In this case there is no way of choosing an uniformly best estimator. Thus, we preassign a value of E_0 of tolerable efficiency. Then we consider the set $A = \{\alpha \mid E(\alpha, \delta^2) \geq E_0\}$ and the estimator is chosen which maximizes $E(\alpha, \delta^2)$ over all $\alpha \in A$ and δ^2 . Thus, we solve for α such that

$$\max_{\alpha \in A} \min_{\delta^2} E(\alpha, \delta^2) = E_0. \quad (2.2.37)$$

Hence, we have a max-min rule for the optimum choice of the level of significance for the preliminary test of the null hypothesis $H_0 : \mu_1 = \mu_2$. Tables for the relative efficiency (maximum and minimum) and δ^2 -value for which minimum efficiency occurs are provided for various choices of sample size and α . These tabulated minimum efficiencies are the minimum guaranteed efficiencies for a given choice of α . Table 1 reported at the end of this section gives the values of E_0 for $N_1 = 4(2)10$ and $N_2 = 4(4)16$. Table of the relative efficiency for odd samples is provided in Appendix B. For given sample sizes N_1 and N_2 one enters the table and locates E_0 which is the minimum relative efficiency he wishes to accept. Then the recommended α^* level is selected. For example, if $N_1 = 8, N_2 = 16$ and a researcher wishes to have an estimator with a minimum efficiency no less than 0.66, then the recommended α^* is 0.15 from Table 1. This choice of α^* would yield an estimator with a maximum efficiency as high as 1.4442 and a minimum guaranteed efficiency 0.66.

Table 1 Maximum and minimum guaranteed efficiencies

	(4,12)	(5,12)	(6,12)	(7,12)	(8,12)	(9,12)	(10,12)	(11,12)
55 E*	1.2578	1.2518	1.2507	1.2505	1.2502	1.2501	1.2500	1.2500
E	1.1213	1.1207	1.1206	1.1205	1.1204	1.1203	1.1202	1.1201
Δ	5.5565	5.5561	5.5559	5.5558	5.5557	5.5556	5.5555	5.5554
60 E*	1.2500	1.2499	1.2498	1.2497	1.2496	1.2495	1.2494	1.2493
E	1.1200	1.1199	1.1198	1.1197	1.1196	1.1195	1.1194	1.1193
Δ	4.4707	4.4706	4.4705	4.4704	4.4703	4.4702	4.4701	4.4700
65 E*	1.2489	1.2488	1.2487	1.2486	1.2485	1.2484	1.2483	1.2482
E	1.1189	1.1188	1.1187	1.1186	1.1185	1.1184	1.1183	1.1182
Δ	3.9409	3.9408	3.9407	3.9406	3.9405	3.9404	3.9403	3.9402
70 E*	1.2478	1.2477	1.2476	1.2475	1.2474	1.2473	1.2472	1.2471
E	1.1178	1.1177	1.1176	1.1175	1.1174	1.1173	1.1172	1.1171
Δ	3.4109	3.4108	3.4107	3.4106	3.4105	3.4104	3.4103	3.4102
75 E*	1.2467	1.2466	1.2465	1.2464	1.2463	1.2462	1.2461	1.2460
E	1.1167	1.1166	1.1165	1.1164	1.1163	1.1162	1.1161	1.1160
Δ	3.3111	3.3110	3.3109	3.3108	3.3107	3.3106	3.3105	3.3104
80 E*	1.2456	1.2455	1.2454	1.2453	1.2452	1.2451	1.2450	1.2449
E	1.1156	1.1155	1.1154	1.1153	1.1152	1.1151	1.1150	1.1149
Δ	3.1924	3.1923	3.1922	3.1921	3.1920	3.1919	3.1918	3.1917
85 E*	1.2445	1.2444	1.2443	1.2442	1.2441	1.2440	1.2439	1.2438
E	1.1145	1.1144	1.1143	1.1142	1.1141	1.1140	1.1139	1.1138
Δ	3.0737	3.0736	3.0735	3.0734	3.0733	3.0732	3.0731	3.0730
90 E*	1.2434	1.2433	1.2432	1.2431	1.2430	1.2429	1.2428	1.2427
E	1.1134	1.1133	1.1132	1.1131	1.1130	1.1129	1.1128	1.1127
Δ	2.9550	2.9549	2.9548	2.9547	2.9546	2.9545	2.9544	2.9543
95 E*	1.2423	1.2422	1.2421	1.2420	1.2419	1.2418	1.2417	1.2416
E	1.1123	1.1122	1.1121	1.1120	1.1119	1.1118	1.1117	1.1116
Δ	2.8363	2.8362	2.8361	2.8360	2.8359	2.8358	2.8357	2.8356
100 E*	1.2412	1.2411	1.2410	1.2409	1.2408	1.2407	1.2406	1.2405
E	1.1112	1.1111	1.1110	1.1109	1.1108	1.1107	1.1106	1.1105
Δ	2.7176	2.7175	2.7174	2.7173	2.7172	2.7171	2.7170	2.7169

Table 1 Maximum and minimum guaranteed efficiencies

α	E_0	(4.4)	(5.4)	(6.4)	(7.4)	(8.4)	(9.4)	(10.4)	(11.4)
95	E^*	16445	14451	13400	12400	11400	10400	9400	8400
	E_0	4623	5477	6402	7400	8400	9400	10400	11400
	Δ_0	7000.4	6000.2	5000.0	4000.0	3000.0	2000.0	1000.0	0
10	E^*	14405	12405	10405	8405	6405	4405	2405	405
	E_0	5405	6405	7405	8405	9405	10405	11405	12405
	Δ_0	53280	49274	45087	40712	36287	31912	27537	23162
15	E^*	12415	10415	8415	6415	4415	2415	415	0
	E_0	6781	7410	8039	8668	9297	9926	10555	11184
	Δ_0	10438	12448	14458	16468	18478	20488	22498	24508
20	E^*	10478	8478	6478	4478	2478	478	0	0
	E_0	7478	7978	8478	8978	9478	9978	10478	10978
	Δ_0	10000	12000	14000	16000	18000	20000	22000	24000
25	E^*	8481	6481	4481	2481	481	0	0	0
	E_0	7978	8478	8978	9478	9978	10478	10978	11478
	Δ_0	10000	12000	14000	16000	18000	20000	22000	24000
30	E^*	6495	4495	2495	495	0	0	0	0
	E_0	8495	8995	9495	9995	10495	10995	11495	11995
	Δ_0	10000	12000	14000	16000	18000	20000	22000	24000
35	E^*	4495	2495	495	0	0	0	0	0
	E_0	8495	8995	9495	9995	10495	10995	11495	11995
	Δ_0	10000	12000	14000	16000	18000	20000	22000	24000
40	E^*	2495	495	0	0	0	0	0	0
	E_0	8495	8995	9495	9995	10495	10995	11495	11995
	Δ_0	10000	12000	14000	16000	18000	20000	22000	24000
45	E^*	495	0	0	0	0	0	0	0
	E_0	8495	8995	9495	9995	10495	10995	11495	11995
	Δ_0	10000	12000	14000	16000	18000	20000	22000	24000
50	E^*	0	0	0	0	0	0	0	0
	E_0	8495	8995	9495	9995	10495	10995	11495	11995
	Δ_0	10000	12000	14000	16000	18000	20000	22000	24000

2.3. POOLING COMPONENT - MEANS

This section considers the problem of estimating the mean μ_1 of one of the components of a bivariate normal distribution with equal component variance from a sample of size N .

Let $(x_1, x_2)(j = 1, 2, \dots, N)$ be a random sample of size N from a bivariate normal distribution with parameters μ_1, μ_2, σ^2 and correlation coefficient ρ .

We are primarily interested in the estimation of μ_1 when it is apriori suspected that μ_1 may be equal to μ_2 in the presence of unknown correlation. The reasons for introducing correlation is that most multistage interpenetrating samples have a series of common first stage, and secondary stage sampling units and the responses are correlated.

2.3.1. ESTIMATION STRATEGIES

In this section we discuss various estimation strategies for the estimation of μ_1 in the presence of an uncertain information that $\mu_1 = \mu_2$ may hold, and provide bias and mean square error expressions for related estimators.

We have the following three estimators for μ_1 :

(i) Unrestricted maximum likelihood estimator (UMLE):

$$\bar{\mu}_1 = N^{-1} \sum_{j=1}^N x_{1j}. \quad (2.3.1)$$

The bias and mean square error of $\bar{\mu}_1$ are given by:

$$B_1(\bar{\mu}_1) = E(\bar{\mu}_1 - \mu_1) = 0 \quad (2.3.2)$$

$$MSE(\bar{\mu}_1) = E(\bar{\mu}_1 - \mu_1)^2 = \frac{\sigma^2}{N} \quad (2.3.3)$$

(ii) Restricted maximum likelihood estimator (RMLE):

$$\hat{\mu}_1 = \frac{1}{2}(\bar{\mu}_1 + \bar{\mu}_2), \quad \bar{\mu}_i = N^{-1} \sum_{j=1}^N x_{ij} \quad (i = 1, 2). \quad (2.3.4)$$

The bias and mean square error for RMLE are:

$$B_2(\hat{\mu}_1) = E(\hat{\mu}_1 - \mu_1) = \frac{1}{2}(\mu_2 - \mu_1) \quad (2.3.5)$$

$$MSE(\hat{\mu}_1) = E(\hat{\mu}_1 - \mu_1)^2 = \frac{\sigma^2}{2N}(1 + \rho) + \frac{\sigma^2 \delta^2}{2N}(1 - \rho^2), \quad (2.3.6)$$

where

$$\delta^2 = \frac{N(\mu_2 - \mu_1)^2}{2\sigma^2(1 - \rho^2)}$$

(iii) Preliminary test maximum likelihood estimator (PTMLE):

$$\hat{\mu}_1^{PT} = \hat{\mu}_1 I(|t| < t_{\alpha/2}) + \bar{\mu}_1 I(|t| > t_{\alpha/2}), \quad (2.3.7)$$

where

$$t = (\bar{\mu}_2 - \bar{\mu}_1) \left\{ \frac{s_{11} + s_{22} - 2s_{12}}{N(N-1)} \right\}^{-\frac{1}{2}} \quad (2.3.8)$$

and

$$s_{ij} = \sum_{k=1}^N (x_{ik} - \bar{\mu}_i)(x_{jk} - \bar{\mu}_j) \quad (i = 1, 2, j = 1, 2)$$

and $t_{\alpha/2}$ is the upper $\alpha/2$ percentile of a student's t-distribution with $N - 1$ degrees of freedom. The bias and mean square error have been derived by Ahsanullah (1971) and has the limitation that the sample size N must be even and greater than 3.

In the following section, we redefined preliminary test estimator using an F-test.

2.3.2. REDEFINED PRELIMINARY TEST ESTIMATOR

The redefined preliminary test estimator is given by:

$$\hat{\mu}_1^{PT} = \hat{\mu}_1 I(t^2 < F_{1,\nu}(\alpha)) + \tilde{\mu}_1 I(t^2 \geq F_{1,\nu}(\alpha)), \quad (2.3.9)$$

$$= \tilde{\mu}_1 + (\hat{\mu}_1 - \tilde{\mu}_1) I(t^2 < F_{1,\nu}(\alpha)), \quad (2.3.10)$$

where $\nu = N - 1$ and $F_{1,\nu}(\alpha)$ is the upper α -percentile of a central F-distribution with $(1, \nu)$ degrees of freedom. This estimator has the following properties:

Theorem 2.3.1. The bias and mean square error of $\hat{\mu}_1^{PT}$ are given by

$$B_3 = E(\hat{\mu}_1^{PT} - \mu_1) = \frac{1}{2}(\mu_2 - \mu_1)H_{3,\nu}(F_{1,\nu}(\alpha); \delta^2), \quad (2.3.11)$$

$$MSE(\hat{\mu}_1^{PT}) = E(\hat{\mu}_1^{PT} - \mu_1)^2 = \frac{\sigma^2}{N} + \frac{\sigma^2}{4N}(1 - \rho) \left[\delta^2 \{ 2H_{3,\nu}(F_{1,\nu}(\alpha); \delta^2) - H_{5,\nu}(F_{1,\nu}(\alpha); \delta^2) \} - H_{3,\nu}(F_{1,\nu}(\alpha); \delta^2) \right], \quad (2.3.12)$$

where $H_{\nu_1, \nu_2}(\cdot; \delta^2)$ is the cumulative distribution function of a non-central F-distribution with (ν_1, ν_2) degrees of freedom and non-centrality parameter δ^2 .

PROOF. The evaluation of (2.3.11) is straightforward. We compute only $E(\hat{\mu}_1^{PT} - \mu_1)^2$.

$$\begin{aligned} E(\hat{\mu}_1^{PT} - \mu_1)^2 &= E(\tilde{\mu}_1 - \mu_1)^2 + E[(\hat{\mu}_1 - \tilde{\mu}_1)^2 I(t^2 < F_{1,\nu}(\alpha))] \\ &\quad + 2E[(\tilde{\mu}_1 - \mu_1)(\hat{\mu}_1 - \tilde{\mu}_1) I(t^2 < F_{1,\nu}(\alpha))] \\ &= E(\tilde{\mu}_1 - \mu_1)^2 + \frac{1}{4}E[(\tilde{\mu}_2 - \tilde{\mu}_1)^2 I(t^2 < F_{1,\nu}(\alpha))] \\ &\quad + E[(\tilde{\mu}_1 - \mu_1)(\tilde{\mu}_2 - \tilde{\mu}_1) I(t^2 < F_{1,\nu}(\alpha))] \\ &= E(\tilde{\mu}_1 - \mu_1)^2 - \frac{1}{4}E[(\tilde{\mu}_2 - \tilde{\mu}_1)^2 I(t^2 < F_{1,\nu}(\alpha))] \\ &\quad + \frac{1}{2}(\mu_2 - \mu_1)E[(\tilde{\mu}_2 - \tilde{\mu}_1) I(t^2 < F_{1,\nu}(\alpha))], \end{aligned}$$

using the conditional expectation of $(\tilde{\mu}_1 - \mu_1)$ given $(\tilde{\mu}_2 - \tilde{\mu}_1)$ and applying Theorem A.2 from Appendix A, (2.3.12) is established.

The following Theorem shows the dominance picture of estimators under the null hypothesis.

Theorem 2.3.2. Under the null hypothesis $H_0 : \mu_1 = \mu_2$, $\tilde{\mu}_1$, $\hat{\mu}_1$ and $\hat{\mu}_1^{PT}$ may be ordered according to the magnitude of their mean square error as follows:

$$\hat{\mu}_1 > \hat{\mu}_1^{PT} > \tilde{\mu}_1. \quad (2.3.13)$$

PROOF.

$$(i) \text{ } MSE(\tilde{\mu}_1) - MSE(\hat{\mu}_1^{PT}) = \frac{\sigma^2}{4N} (1 - \rho) H_{3,\nu}(F_{1,\nu}(\alpha); 0) \geq 0, \quad (2.3.14)$$

$$(ii) \text{ } MSE(\hat{\mu}_1^{PT}) - MSE(\hat{\mu}_1) = \frac{\sigma^2(1-\rho)}{4N} \{2 - H_{3,\nu}(F_{1,\nu}(\alpha); 0)\} \geq 0. \quad (2.3.15)$$

2.3.3. EFFICIENCY ANALYSIS AND RECOMMENDATIONS

The relative efficiencies of these three estimators with respect to each other are:

$$R.E.(\hat{\mu}_1 : \tilde{\mu}_1) = 2[(1 + \rho) + \delta^2(1 - \rho)]^{-1}, \quad (2.3.16)$$

$$R.E.(\hat{\mu}_1^{PT} : \tilde{\mu}_1) = \frac{1}{2}[(1 + \rho) + \delta^2(1 - \rho)][1 + g(\delta)]^{-1}, \quad (2.3.17)$$

$$R.E.(\hat{\mu}_1^{PT} : \hat{\mu}_1) = [1 + g(\delta)]^{-1}, \quad (2.3.18)$$

where

$$g(\delta) = \frac{1}{4}(1 - \rho)[\delta^2\{2H_{3,\nu}(F_{1,\nu}(\alpha); \delta^2) - H_{5,\nu}(F_{1,\nu}(\alpha); \delta^2)\} - H_{3,\nu}(F_{1,\nu}(\alpha); \delta^2)]. \quad (2.3.19)$$

The efficiency of $\hat{\mu}_1$ relative to $\tilde{\mu}_1$ is a monotone function in δ^2 with a maximum (at $\delta^2 = 0$)

$$2(1 + \rho)^{-1} \geq 1 \quad (2.3.20)$$

Thus $\hat{\mu}_1$ is more efficient than $\tilde{\mu}_1$ under the null hypothesis for fixed ρ ($-1 \leq \rho \leq 1$). The efficiency of $\hat{\mu}_1$ drops to zero as the hypothesis error grows (δ^2 increases), crossing the line $R.E. = 1$ at $\delta^2 = 1$, also $R.E.(\hat{\mu}_1 : \tilde{\mu}_1) \geq 1$ according as $\delta^2 \leq 1$. Thus, if $0 \leq \delta^2 < 1$ then $\hat{\mu}_1$ is the efficient estimator while $\tilde{\mu}_1$ performs efficiently outside this interval.

Under the null hypothesis the relative efficiency of $\hat{\mu}_1^{PT}$ to $\hat{\mu}_1$ is

$$R.E.(\hat{\mu}_1^{PT} : \hat{\mu}_1) = \frac{1}{2}(1 + \rho) \left[1 - \frac{1}{4}(1 - \rho) H_{3,\nu}(F_{1,\nu}(\alpha); 0) \right]^{-1} \geq \left(\frac{1}{2} + \frac{\rho}{2} \right), \quad (2.3.21)$$

while the relative efficiency of $\hat{\mu}_1^{PT}$ to $\tilde{\mu}_1$ is

$$R.E.(\hat{\mu}_1^{PT} : \tilde{\mu}_1) = \left[1 - \frac{1}{4}(1 - \rho) H_{3,\nu}(F_{1,\nu}(\alpha), 0) \right]^{-1} \geq 1. \quad (2.3.22)$$

Thus, under the null hypothesis we establish the following inequality:

$$\frac{1}{2} + \frac{\rho}{2} \leq R.E.(\hat{\mu}_1^{PT} : \hat{\mu}_1) \leq 1 \leq R.E.(\hat{\mu}_1^{PT} : \tilde{\mu}_1). \quad (2.3.23)$$

But as the hypothesis error grows, i.e. δ^2 moves away from the origin,

$$R.E.(\hat{\mu}_1^{PT} : \hat{\mu}_1) \gtrless 1 \text{ according as } \delta^2 \gtrless 1 + 2(1 - \rho)^{-1}g(\delta). \quad (2.3.24)$$

Also, as $\delta^2 \rightarrow \infty$, $R.E.(\hat{\mu}_1^{PT} : \hat{\mu}_1) \rightarrow 0$. Thus, except when δ^2 is small, $\hat{\mu}_1^{PT}$ is relatively more efficient than $\hat{\mu}_1$.

Finally, we consider the relative efficiency of $\hat{\mu}_1^{PT}$ to $\bar{\mu}_1$. For fixed ρ ($-1 < \rho < +1$) this R.E. is a function of α and δ^2 for a given sample size N . Let this be denoted by $E_\rho(\alpha, \delta^2)$. The function $E_\rho(\alpha, \delta^2)$, for $\alpha \neq 0$, has its maximum at $\delta^2 = 0$ with its value given by (2.3.22). This function decreases, crossing the line $E_\rho(\alpha, \delta^2) = 1$, to a minimum $E_\rho^0(\alpha, \delta_0^2)$ at $\delta^2 = \delta_0^2$; then increases towards 1 as $\delta^2 \rightarrow \infty$. For $\delta^2 = 0$ and varying α , we obtain

$$\max_{0 \leq \alpha \leq 1} E_\rho(\alpha, 0) = E_\rho(0, 0) = 4(3 + \rho)^{-1}.$$

The value of $E_\rho(\alpha, 0)$ decreases as α increases. On the other hand, if $\alpha = 0$ and δ^2 vary, then the curves $E_\rho(0, \delta)$ and $E_\rho(1, \delta^2)$ intersect at $\delta^2 = 1$. Also, for a general α , $E_\rho(\alpha, \delta^2)$ and $E_\rho(1, \delta^2)$ intersect in the interval $0 \leq \delta^2 \leq 1$ somewhere below the line $E_\rho(\alpha, \delta^2) = 1$.

In order to choose an estimator with optimum relative efficiency, we adopt the following rule for fixed ρ value: If $0 \leq \delta^2 \leq 1$, $\hat{\mu}_1$ is chosen since $E_\rho(0, \delta^2)$ is the largest in this interval. However, generally δ^2 is unknown and may not be in this interval. In this case there is no way of choosing uniformly best estimator. Thus, we must preassign a value of E_0 of tolerable efficiency (minimum efficiency), consider the set $A = \{\alpha \mid E_\rho(\alpha, \delta^2) \geq E_0 \text{ for all } \delta^2\}$; and choose the estimator which maximizes $E(\alpha, \delta^2)$ over all $\alpha \in A$ and δ^2 . Thus, we solve for α such that

$$\max_{\alpha \in A} \min_{\delta^2} E_\rho(\alpha, \delta^2) = E_0. \quad (2.3.25)$$

One may construct tables for maximum and minimum guaranteed efficiency with related δ^2 -value for various sample sizes and α -values. From the table one enters (N, ρ) , locate the E_0 -value, and then the corresponding α -values gives the required α^* , the optimal size of the test.

2.4. SIMPLE LINEAR MODEL

Let $(x_1, y_1), \dots, (x_N, y_N)$ be N independent observations such that

$$y_j = \theta + \beta x_j + e_j \quad (j = 1, 2, \dots, N) \quad (2.4.1)$$

where we assume that $e_1, \dots, e_N$ are independent and identically distributed normal random variables with mean 0 and constant (unknown) variance σ^2 . We are primarily interested in the estimation of the intercept parameter θ when it is apriori suspected that $\beta = 0$ (without loss of generality). The situation may arise in the following way: An experimenter may be interested in calibrating a straight-faced spring, i.e., he may wish to determine a functional relationship of the form $y = \theta + \beta x$ between the position y of the scale pointer and mass x placed repeatedly on the pan, the pointer reading varying somewhat. That is, the appropriate relationship may be $y = \theta + \beta x + e$, where $e \sim N(0, \sigma^2)$. As part of the calibration, the experimenter wishes that corresponds to an empty pan, which is to say, estimate θ in the model. The proposed estimation of θ depends upon the result of a preliminary test of significance of the hypothesis $H_0 : \beta = 0$. Another situation is: In studying the strength of a certain virus in different situations some virologists adopt a scoring system to score the sensitivity of the virus effect as well as the day of the death. At stages where compounds have been screened for protection of animals against a certain virus, there is usually prior knowledge from preliminary studies of the virus with no treatments (control). In such a situation, in order to determine the treatment effect one needs to carry out a preliminary test on dose effect before deciding the main effect.

2.4.1. ESTIMATION STRATEGIES

This section is devoted to a discussion of three estimation strategies for the estimation of the intercept parameter θ when $\beta = 0$ is suspected. The expressions for the bias and mean square errors for all three estimators are provided.

The three estimators of θ are:

(i) Unrestricted maximum likelihood estimator (UMLE):

$$\tilde{\theta} = \bar{y} - \tilde{\beta}\bar{x}, \quad (2.4.2)$$

where $\tilde{\beta}$ is the ordinary least square estimator of β . The bias and mean square error of $\tilde{\theta}$ are:

$$B_1(\tilde{\theta}) = E(\tilde{\theta} - \theta) = 0 \quad (2.4.3)$$

$$MSE(\tilde{\theta}) = E(\tilde{\theta} - \theta)^2 = \sigma^2\left(\frac{1}{N} + \frac{\bar{x}^2}{Q}\right), \quad (2.4.4)$$

where

$$Q = \sum_{j=1}^N (x_j - \bar{x})^2.$$

(ii) Restricted maximum likelihood estimator (RMLE):

$$\hat{\theta} = \bar{y} = N^{-1} \sum_{j=1}^N y_j, \quad (2.4.5)$$

with bias

$$B_2(\hat{\theta}) = E(\hat{\theta} - \theta) = \beta\bar{x}, \quad (2.4.6)$$

and mean square error

$$MSE(\hat{\theta}) = E(\hat{\theta} - \theta)^2 = \frac{\sigma^2}{N} + \beta^2\bar{x}^2 = \sigma^2\left(\frac{1}{N} + \frac{\bar{x}^2}{Q}\Delta\right), \quad (2.4.7)$$

where

$$\Delta = (\beta/\sigma)^2 Q.$$

(iii) Preliminary test maximum likelihood estimator (PTMLE):

$$\begin{aligned}\hat{\theta}^{PT} &= \hat{\theta}I(|t| < t_{\alpha/2}) + \bar{\theta}I(|t| \geq t_{\alpha/2}) \\ &= \bar{\theta} - (\bar{\theta} - \hat{\theta})I(|t| < t_{\alpha/2}) \\ &= (\bar{y} - \bar{\beta}\bar{x})I(|t| < t_{\alpha/2}),\end{aligned}\tag{2.4.8}$$

where

$$t = \frac{\bar{\beta}\sqrt{\sum_{j=1}^N (x_j - \bar{x})^2}}{S_e},\tag{2.4.9}$$

$$(N-2)S_e^2 = \sum_{j=1}^N [(y_j - \bar{y}) - \bar{\beta}(x_j - \bar{x})]^2.$$

The bias and mean square error of the preliminary test estimator of θ have been derived by Ahsanullah and Saleh (1972) and has the limitation that the sample size N must be even and greater than 4. In the present work, we use the F-test for the null hypothesis $H_0 : \beta = 0$ instead of the usual t-test.

2.4.2. REDEFINED PRELIMINARY TEST ESTIMATOR

The redefined preliminary test estimator is given by

$$\hat{\theta}^{PT} = \bar{y} + \bar{\beta}\bar{x}I(t^2 < F_{1,\nu}(\alpha)), \quad \nu = N - 2\tag{2.4.10}$$

where t is the usual test statistic for the null hypothesis and $F_{1,\nu}(\alpha)$ is upper α -percentile of a central F-distribution with $(1, \nu)$ degrees of freedom. The bias is

$$B_3 = E(\hat{\theta}^{PT} - \theta) = \bar{x}\beta H_{3,\nu}(F_{1,\nu}(\alpha); \Delta). \quad (2.4.11)$$

And the mean square error for $\hat{\theta}^{PT}$ is:

$$\begin{aligned} MSE(\hat{\theta}^{PT}) &= E(\hat{\theta}^{PT} - \theta)^2 = E\{(\tilde{\theta} - \theta) + \hat{\beta}\bar{x}(F < F_{1,\nu}(\alpha))\}^2 \\ &= E(\tilde{\theta} - \theta)^2 + \bar{x}^2 E\{\tilde{\beta}^2 I(t^2 < F_{1,\nu}(\alpha))\} \\ &\quad + 2\bar{x}E\{(\tilde{\theta} - \theta)\tilde{\beta}I(t^2 < F_{1,\nu}(\alpha))\}, \\ &= \sigma^2\left(\frac{1}{N} + \frac{\bar{x}^2}{Q}\right) + \frac{\sigma^2\bar{x}^2}{Q} \left[\Delta\{2H_{3,\nu}(F(\alpha); \Delta) - H_{5,\nu}(F(\alpha); \Delta)\} \right. \\ &\quad \left. - H_{3,\nu}(F_{1,\nu}(\alpha); \Delta) \right], \end{aligned} \quad (2.4.12)$$

where $H_{\nu_1, \nu_2}(\cdot; \Delta)$ is the cumulative distribution of a non-central F-distribution with degrees of freedom (ν_1, ν_2) and non-centrality parameter $\Delta = (\beta/\sigma)^2 Q$ and $Q = \sum_{j=1}^N (x_j - \bar{x})^2$, by applying Theorem A.2 from Appendix A.

2.4.3. EFFICIENCY ANALYSIS AND RECOMMENDATIONS

In this section we consider the analysis of the relative efficiencies of the estimators.

The R.E. of $\hat{\theta}$ versus $\tilde{\theta}$ is

$$R.E.(\hat{\theta} : \tilde{\theta}) = \left(\frac{1}{N} + \frac{\bar{x}^2}{Q}\right) \left(\frac{1}{N} + \frac{\bar{x}^2}{Q}\Delta\right)^{-1}. \quad (2.4.13)$$

From (2.4.13), $R.E.(\hat{\theta} : \tilde{\theta})$ is a decreasing function of Δ . At $\Delta = 0$, the value is

$$1 + \frac{N\bar{x}^2}{Q} (> 1) \quad (2.4.14)$$

and at $\Delta = 1$, $R.E.(\hat{\theta} : \tilde{\theta}) = 1$ and as $\Delta \rightarrow \infty$, $R.E.(\hat{\theta} : \tilde{\theta}) \rightarrow 0$. Therefore, $\tilde{\theta}$ inefficient.

Indeed,

$$R.E.(\hat{\theta} : \tilde{\theta}) \geq 1 \quad \text{according as } \Delta \leq 1.$$

Thus, if $\Delta \leq 1$, $\hat{\theta}$ is more efficient than $\tilde{\theta}$ otherwise $\tilde{\theta}$ is efficient.

The relative efficiency of $\hat{\theta}^{PT}$ to $\hat{\theta}$ and $\tilde{\theta}$ are given by (2.4.15) and (2.4.16) respectively as follows:

$$R.E.(\hat{\theta}^{PT} : \hat{\theta}) = \left[\frac{1}{N} + \frac{\bar{x}^2}{Q} \Delta \right] \left[1 + g(\Delta) \right]^{-1} \left[\frac{1}{N} + \frac{\bar{x}^2}{Q} \right]^{-1} \quad (2.4.15)$$

$$R.E.(\hat{\theta}^{PT} : \tilde{\theta}) = \left[1 + g(\Delta) \right]^{-1}, \quad (2.4.16)$$

where

$$g(\Delta) = \frac{\bar{x}^2}{Q} \left[\Delta \{ 2H_{3,N-2}(F(\alpha); \Delta) - H_{5,\nu}(F(\alpha); \Delta) \} - H_{3,\nu}(F(\alpha); \Delta) \right]. \quad (2.4.17)$$

First, we compare $\hat{\theta}^{PT}$ to $\hat{\theta}$. We note that under H_0

$$RE(\hat{\theta}^{PT} : \hat{\theta}) = \left(1 + \frac{N\bar{x}^2}{Q} \right)^{-1} \left[1 - \frac{\bar{x}^2}{Q} H_{3,\nu}(F_{1,\nu}(\alpha); 0) \right]^{-1} \quad (2.4.18)$$

while the relative efficiency of $\hat{\theta}^{PT}$ to $\tilde{\theta}$ is

$$R.E.(\hat{\theta}^{PT} : \tilde{\theta}) = \left[1 - \frac{\bar{x}^2}{Q} H_{3,\nu}(F_{1,\nu}(\alpha); 0) \right]^{-1} (\geq 1) \quad (2.4.19)$$

Thus, under H_0 , we obtain

$$\left(1 + \frac{\bar{x}^2}{Q} \right)^{-1} \leq R.E.(\hat{\theta}^{PT} : \hat{\theta}) \leq 1 \leq RE(\hat{\theta}^{PT} : \tilde{\theta}). \quad (2.4.20)$$

On the other hand as Δ moves away from 0,

$$RE(\hat{\theta}^{PT} : \hat{\theta}) \gtrless 1 \quad \text{according as}$$

$$\Delta \lessgtr 1 + \left(1 + \frac{Q}{N\bar{x}^2} \right) g(\Delta) \quad (2.4.21)$$

Finally, as $\Delta \rightarrow \infty$, the R.E. $(\hat{\theta}^{PT} : \hat{\theta})$ tends to 0. Thus, except when Δ is small, $\hat{\theta}^{PT}$ is relatively more efficient than $\hat{\theta}$. For $\bar{x} = 0$,

$$B_1 = B_2 = B_3 = 0 \quad (2.4.22)$$

$$MSE(\tilde{\theta}) = MSE(\hat{\theta}) = MSE(\hat{\theta}^{PT}) = \frac{\sigma^2}{N}. \quad (2.4.23)$$

Therefore, all estimators are unbiased and have constant variance. However, if $\bar{x} \neq 0$, then only the UMLE is unbiased and its mean square error is independent of Δ . Now, we have the following theorem under H_0 .

Theorem 2.4.1. Under $H_0 : \beta = 0$, provided $\bar{x} \neq 0$, the estimators are ranked according to mean square error, as follows:

$$\hat{\theta} > \hat{\theta}^{PT} > \tilde{\theta}. \quad (2.4.24)$$

PROOF.

$$MSE(\tilde{\theta}) - MSE(\hat{\theta}) = \left(\frac{\bar{x}^2}{Q}\right)\sigma^2 > 0 \quad (2.4.25)$$

$$MSE(\tilde{\theta}) - MSE(\hat{\theta}^{PT}) = \left(\frac{\bar{x}^2}{Q}\right)\sigma^2 H_{2,\nu}(F(\alpha); 0) > 0 \quad (2.4.26)$$

$$MSE(\hat{\theta}^{PT}) - MSE(\hat{\theta}) = \left(\frac{\bar{x}^2}{Q}\right)\sigma^2 \{1 - H_{3,\nu}(F(\alpha); 0)\} > 0. \quad (2.4.27)$$

Finally, we compare $\hat{\theta}^{PT}$ to $\tilde{\theta}$. We first note that the R.E. $(\hat{\theta}^{PT} : \tilde{\theta})$ is a function of α and Δ , say $E(\alpha, \Delta)$ for fixed design (x)-value. The function $E(\alpha, \Delta)$ has its maximum $\Delta = 0$ for $\alpha \neq 0$ with value

$$\left\{1 - \frac{\bar{x}^2}{Q} \left(\frac{1}{N} + \frac{\bar{x}^{2-1}}{Q}\right) H_{3,N-2}(F(\alpha); 0)\right\}^{-1} (\geq 1). \quad (2.4.28)$$

But as Δ moves away from the origin, $E(\alpha, \Delta)$ decreases to a minimum value

$$[1 + g(\Delta_0)]^{-1} \quad \text{at} \quad \Delta = \Delta_0, \quad (2.4.29)$$

crossing the line $E(\alpha, \Delta) = 1$; then increases towards 1 as Δ tends to infinity.

We note that

$$\max_{\alpha} E(\alpha, 0) = E(0, 0) = 1 + \frac{N\bar{x}^2}{Q}.$$

The value of $E(\alpha, 0)$ decreases as α increases. On the other hand, if $\alpha = 0$, and Δ vary, then the curve $E(0, \Delta)$ and $E(1, \Delta) = 1$ intersects at $\Delta = 1$. For a general α , $E(\alpha, \Delta)$ and $E(1, \Delta)$ will never intersect outside the interval $0 \leq \Delta \leq 1$; the value of Δ at the intersection decreases as α increases. Therefore, for two different α , $E(\alpha_1, \Delta)$ and $E(\alpha_2, \Delta)$ curve will never intersect above the line $E(\alpha, \Delta) = 1$ but somewhere below $E(\alpha, \Delta) = 1$. Figure 2 displays some examples of this function for $\alpha = 0.05, 0.1, 0.2$ and 0.5.

$N = 16$ and $\alpha = 0.05, 0.1, 0.2, 0.5$

$N = 20$ and $\alpha = 0.05, 0.1, 0.2, 0.5$

$N = 16$ and $\alpha = 0.05, 0.1, 0.2, 0.5$

$N = 20$ and $\alpha = 0.05, 0.1, 0.2, 0.5$

Fig.2 Relative efficiency of PTMLL to UMLL for various sample sizes.

the problem for known covariance matrix $\Sigma = I_p$. Saleh (1973) considered the case when Σ is unknown. Both papers considered only the preliminary test estimator of $\mu^{(1)}$ and compared it with the classical estimator.

3.2. ESTIMATION STRATEGIES

In this section we propose the following four estimators for the mean vector $\mu^{(1)}$ when it is apriori suspected that $\mu^{(1)} = \mu^{(2)}$ may hold.

i) Unrestricted maximum likelihood estimator (UMLE):

$$\tilde{\mu}^{(1)} = \bar{x}^{(1)} = N_1^{-1} \sum_{j=1}^{N_1} x_{1j}. \quad (3.2.1)$$

ii) Restricted maximum likelihood estimator (RMLE):

$$\hat{\mu}^{(1)} = \frac{N_1 \tilde{\mu}_1^{(1)} + N_2 \tilde{\mu}_2^{(2)}}{N_1 + N_2}, \quad (3.2.2)$$

where

$$\tilde{\mu}^{(i)} = \bar{x}^{(i)} = N_i^{-1} \sum_{j=1}^{N_i} x_{ij}, \quad i = 1, 2.$$

iii) Preliminary test maximum likelihood estimator (PTMLE):

$$\hat{\mu}^{(1)PT} = \tilde{\mu}^{(1)} I(T_N^2 \geq T_{N,\alpha}^2) + \hat{\mu}^{(1)} I(T_N^2 < T_{N,\alpha}^2) \quad (3.2.3)$$

$$= \tilde{\mu}^{(1)} + \frac{N_2}{N_1 + N_2} (\tilde{\mu}^{(2)} - \tilde{\mu}^{(1)}) I(T_N^2 < T_{N,\alpha}^2); \quad (3.2.4)$$

where T_N^2 is the Hotelling's T_N^2 -statistic for testing the null-hypothesis $H_0: \mu^{(1)} = \mu^{(2)}$:

$$T_N^2 = \frac{N_1 N_2}{N_1 + N_2} (\tilde{\mu}_2^{(2)} - \tilde{\mu}^{(1)})' S^{-1} (\tilde{\mu}_2^{(2)} - \tilde{\mu}^{(1)}) \quad (3.2.5)$$

$N = 12$.

[illegible]

2.5. Concluding Remarks

We have presented and discussed three estimators for estimating a parameter of interest in presence of uncertain prior information. The decision whether to use the UMLE or the RMLE or the PTMLE rests on the experimenter. We conclude with the following steps to select an estimator:

- 1) If the hypothesis is true, i.e., $\delta^2 = 0$, or $0 \leq \delta^2 \leq 1$, then select the RMLE because it has the largest efficiency as compared to the other two estimators.
- 2) If δ is unknown and the experimenter is willing to accept a minimum guaranteed efficiency (E_0) less than one, then he should use the PTMLE. He may attain a maximum efficiency (E^*) greater than one.
- 3) If the experimenter is not willing to accept a minimum efficiency and also has no knowledge about δ , and selects the UMLE, but this may not be the best choice because he may be ignoring other information.

To compute the relative efficiency, we have developed the program with the help of the sub-routine 1BPDF from NAG Library.

In forthcoming chapters we extend these results to the multivariate case and introduce another estimator.

CHAPTER 3

ESTIMATION STRATEGIES FOR THE MEAN-VECTOR OF A MULTIVARIATE NORMAL DISTRIBUTION IN A TWO-SAMPLE PROBLEM

3.1. INTRODUCTION

This chapter concerns the estimation of the mean vector $\mu^{(1)}$ of a p -variate normal distribution in the presence of an uncertain prior information. Four methods of estimation have been considered. Exact bias and risk of all of these estimators are derived. Two measures of relative efficiency are considered: one based on quadratic loss function, and the other based on the determinant of the covariance matrix which may be called the generalized mean square error. Tables for the appropriate choice of the significance level of the test are provided. Four estimation strategies are developed in Section 3.2; the derivation of bias and the risk (under quadratic loss function) are given in Section 3.3; risk efficiency analysis is discussed in Section 3.3.2; mean square error matrices are derived in Section 3.4 and the generalized efficiency is discussed in Section 3.4.1. Section 3.5 summarizes the findings.

Let $X_{11}, \dots, X_{1N_1}$ and $X_{21}, \dots, X_{2N_2}$ be two independent samples of size N_1 and N_2 from $N_p(\mu^{(1)}, \Sigma)$ and $N_p(\mu^{(2)}, \Sigma)$, respectively, where $\mu^{(i)}$, $i = 1, 2$ are $(p \times 1)$ vectors and Σ is a $(p \times p)$ positive semidefinite matrix. The primary objective is to estimate $\mu^{(1)}$ when it is apriori suspected that $\mu^{(1)} = \mu^{(2)}$ may hold. Bancroft and Han (1976) considered

the problem for known covariance matrix $\Sigma = I_p$. Saleh (1973) considered the case when Σ is unknown. Both papers considered only the preliminary test estimator of $\mu^{(1)}$ and compared it with the classical estimator.

3.2. ESTIMATION STRATEGIES

In this section we propose the following four estimators for the mean vector $\mu^{(1)}$ when it is apriori suspected that $\mu^{(1)} = \mu^{(2)}$ may hold.

i) Unrestricted maximum likelihood estimator (UMLE):

$$\tilde{\mu}^{(1)} = \bar{x}^{(1)} = N_1^{-1} \sum_{j=1}^{N_1} x_{1j}. \quad (3.2.1)$$

ii) Restricted maximum likelihood estimator (RMLE):

$$\hat{\mu}^{(1)} = \frac{N_1 \tilde{\mu}_1^{(1)} + N_2 \tilde{\mu}_2^{(2)}}{N_1 + N_2}, \quad (3.2.2)$$

where

$$\tilde{\mu}^{(i)} = \bar{x}^{(i)} = N_i^{-1} \sum_{j=1}^{N_i} x_{ij}, \quad i = 1, 2.$$

iii) Preliminary test maximum likelihood estimator (PTMLE):

$$\hat{\mu}^{(1)PT} = \tilde{\mu}^{(1)} I(T_N^2 \geq T_{N,\alpha}^2) + \hat{\mu}^{(1)} I(T_N^2 < T_{N,\alpha}^2) \quad (3.2.3)$$

$$= \tilde{\mu}^{(1)} + \frac{N_2}{N_1 + N_2} (\tilde{\mu}^{(2)} - \tilde{\mu}^{(1)}) I(T_N^2 < T_{N,\alpha}^2); \quad (3.2.4)$$

where T_N^2 is the Hotelling's T_N^2 -statistic for testing the null-hypothesis $H_0: \mu^{(1)} = \mu^{(2)}$:

$$T_N^2 = \frac{N_1 N_2}{N_1 + N_2} (\tilde{\mu}_2^{(2)} - \tilde{\mu}^{(1)})' S^{-1} (\tilde{\mu}_2^{(2)} - \tilde{\mu}^{(1)}) \quad (3.2.5)$$

$$(N_1 + N_2 - 2)S = \sum_{j=1}^{N_1} (x_{1j}^{(1)} - \tilde{\mu}^{(1)})(x_{1j}^{(1)} - \tilde{\mu}^{(1)})' + \sum_{j=1}^{N_2} (x_{2j}^{(2)} - \tilde{\mu}^{(2)})(x_{2j}^{(2)} - \tilde{\mu}^{(2)})', \quad (3.2.6)$$

$T_{N,\alpha}^2$ is the α -level critical value of T_N^2 -statistics and $I(A)$ is the indicator function of the set A .

iv) Shrinkage maximum likelihood estimator (SMLE):

$$\hat{\mu}^{(1)S} = \tilde{\mu}^{(1)} + \frac{N_2}{N_1 + N_2} nc T_N^{-2} (\tilde{\mu}^{(2)} - \tilde{\mu}^{(1)}); \quad (3.2.7)$$

where c is the shrinkage factor, $0 < c < \frac{2(p-2)}{(n-p+3)}$ and $n = N_1 + N_2 - 2$.

Note that the PTMLE and SMLE of $\mu^{(1)}$ are similar because SMLE is the smooth version of PTMLE. Both adjust for the departure from the null hypothesis $H_0 : \mu^{(1)} = \mu^{(2)}$ through the test statistic T_N^2 ; for large values of T_N^2 , both PTMLE and SMLE reduce to $\tilde{\mu}^{(1)}$, while for small values of T_N^2 , they behave in a different way. To study the properties of the estimators, first we provide a risk comparison in the light of the quadratic loss function (Mahalanobis distance)

$$L(\mu^{*(1)}, \mu^{(1)}) = N_1 (\mu^{*(1)} - \mu^{(1)})' \Sigma^{-1} (\mu^{*(1)} - \mu^{(1)}), \quad (3.2.8)$$

where $\mu^{*(1)}$ is any estimator of $\mu^{(1)}$. Then, the risk of the estimator $\mu^{*(1)}$ is

$$R(\mu^{*(1)}; \Sigma) = E\{N_1 (\mu^{*(1)} - \mu^{(1)})' \Sigma^{-1} (\mu^{*(1)} - \mu^{(1)})\}, \quad (3.2.9)$$

where E denotes the expectation operator.

3.3. BIAS AND RISKS

In this section we derive expressions for the bias and risks under Mahalanobis distance (3.2.8) of all four estimators. Theorem 3.3.1 and 3.3.2 provide such expressions respectively.

THEOREM 3.3.1. Bias of the UMLE, RMLE and SMLE are given by (3.3.1) - (3.3.4) respectively:

$$B_1 = E(\hat{\mu}^{(1)} - \mu^{(1)}) = 0 \quad (3.3.1)$$

$$B_2 = E(\hat{\mu}^{(1)} - \mu^{(1)}) = M\delta \quad (3.3.2)$$

$$B_3 = E(\hat{\mu}^{(1)PT} - \mu^{(1)}) = M\delta H_{p+2,m}(T_{N,\alpha}^{*2}; \Delta) \quad (3.3.3)$$

$$B_4 = E(\hat{\mu}^{(1)S} - \mu^{(1)}) = cmM\delta E(\chi_{p+2}^{-2}(\Delta)), \quad (3.3.4)$$

where

$$\delta = \mu^{(2)} - \mu^{(1)} \text{ and } \Delta = N_1 M \delta' \Sigma^{-1} \delta, \quad M = \frac{N_2}{N_1 + N_2},$$

$$T_{N,\alpha}^{*2} = \frac{p}{m} F_{p,m}(\alpha) \text{ with } m = N_1 + N_2 - p - 1.$$

Also $H_{\nu_1, \nu_2}(T_{N,\alpha}^{*2}; \Delta)$ is the cumulative distribution function of a non-central F -distribution with (ν_1, ν_2) degrees of freedom and non-centrality parameter Δ and $F_{\nu_1, \nu_2}(\alpha)$ is the upper α level critical value of a central F -distribution with (ν_1, ν_2) degrees of freedom. Further, $\chi_{\nu}^2(\Delta)$ is a non-central chi-square random variable with ν degrees of freedom and non-centrality parameter Δ .

PROOF.

$$B_1 = E(\hat{\mu}^{(1)} - \mu^{(1)}) = 0.$$

$$\begin{aligned} B_2 &= E(\hat{\mu}^{(1)} - \mu^{(1)}) + ME(\hat{\mu}^{(2)} - \hat{\mu}^{(1)}) \\ &= M\delta. \end{aligned}$$

From (3.2.4), and by definition

$$\begin{aligned} B_3 &= E(\tilde{\mu}^{(1)PT} - \mu^{(1)}) \\ &= ME \left\{ (\tilde{\mu}^{(2)} - \tilde{\mu}^{(1)}) I(T_N^2 < T_{N,\alpha}^2) \right\}. \end{aligned}$$

If we make the following transformations

$$\tilde{\mu}^{(2)} - \tilde{\mu}^{(1)} = (N_1 M)^{-\frac{1}{2}} y, \quad (3.3.5)$$

$$y^* = Dy,$$

where D is any $p \times p$ orthogonal matrix such that $D \Sigma D' = I_p$, then, from Anderson (1984)

$$\frac{T_N^2}{n} = y^{*'} y^* / \chi_m^2,$$

where $y^{*'} y^*$ follows a non-central chi-square distribution with p degrees of freedom and non-centrality parameter Δ and independent of the χ_m^2 variable. Thus, we obtain

$$\begin{aligned} B_3 &= (N_1 M)^{\frac{1}{2}} E[(y) I(y' S^{-1} y < T_{N,\alpha}^2)] \\ &= (N_1 M)^{\frac{1}{2}} D^{-1} E \{ (y^*) I(y^{*'} y^* / \chi_m^2 < T_{N,\alpha}^2) \} \\ &= \delta H_{p+2,m}(T_{N,\alpha}^2; \Delta) \end{aligned}$$

by Theorem A.1 from Appendix A, establishing (3.3.3).

Finally,

$$\begin{aligned} B_4 &= ncME \left\{ (\tilde{\mu}^{(2)} - \tilde{\mu}^{(1)}) T_N^{-2} \right\} \\ &= Mc(MN_1)^{\frac{1}{2}} E \{ (y) (n^{-1} y' S^{-1} y)^{-1} \} \\ &= Mc(MN_1)^{\frac{1}{2}} D^{-1} E \{ (y^*) (y^{*'} y^* / \chi_m^2)^{-1} \} \\ &= cmM\delta E(\chi_{p+2}^{-2}(\Delta)). \end{aligned}$$

by Theorem A.1 from Appendix A, establishing (3.3.4).

THEOREM 3.3.2. Risks of UMLE, RMLE, PTMLE and SMLE under (3.2.8) are given by (3.3.6)–(3.3.9) respectively:

$$R_1(\tilde{\mu}^{(1)}; \Sigma^{-1}) = p \quad (3.3.6)$$

$$R_2(\hat{\mu}^{(1)}; \Sigma^{-1}) = \frac{N_1}{N_1 + N_2} p + M\Delta \quad (3.3.7)$$

$$R_3(\hat{\mu}^{(1)PT}; \Sigma^{-1}) = p + M \left[\Delta \{ 2H_{p+2,m}(T_{N,\alpha}^{*2}; \Delta) - H_{p+4,m}(T_{N,\alpha}^{*2}; \Delta) \} - pH_{p+2,m}(T_{N,\alpha}^{*2}; \Delta) \right] \quad (3.3.8)$$

$$R_4(\hat{\mu}^{(1)S}; \Sigma^{-1}) = p - cmM \{ 2(p-2) - c(m+2) \} E\{\chi_p^{-2}(\Delta)\}, \quad (3.3.9)$$

where

$$c_{opt} = \frac{p-2}{m+2}$$

PROOF.

$$R_1(\tilde{\mu}_1^{(1)}; \Sigma^{-1}) = E \left\{ N_1 (\tilde{\mu}_1^{(1)} - \mu^{(1)})' \Sigma^{-1} (\tilde{\mu}_1^{(1)} - \mu^{(1)}) \right\} = p,$$

establishing (3.3.6).

$$\begin{aligned} R_2(\hat{\mu}_1^{(1)}; \Sigma^{-1}) &= E \left\{ N_1 (\hat{\mu}_1^{(1)} - \mu^{(1)})' \Sigma^{-1} (\hat{\mu}_1^{(1)} - \mu^{(1)}) \right\} \\ &= EN_1 \left\{ (\tilde{\mu}_1^{(1)} - \mu^{(1)}) + M(\tilde{\mu}^{(2)} - \tilde{\mu}^{(1)}) \right\}' \Sigma^{-1} \\ &\quad \left\{ (\tilde{\mu}_1^{(1)} - \mu^{(1)}) + M(\tilde{\mu}^{(2)} - \tilde{\mu}^{(1)}) \right\} \\ &= p + M^2 E \left\{ N_1 (\tilde{\mu}^{(2)} - \tilde{\mu}^{(1)})' \Sigma^{-1} (\tilde{\mu}^{(2)} - \tilde{\mu}^{(1)}) \right\} \\ &\quad + 2ME \left\{ (\tilde{\mu}_1^{(1)} - \mu^{(1)})' \Sigma^{-1} (\tilde{\mu}^{(1)} - \mu^{(1)}) \right\} \\ &= p - MN_1^2 \left\{ (N_1 M)^{-1} p + \delta' \Sigma^{-1} \delta \right\} + 2M\Delta \\ &= p - Mp - N_1 M^2 \delta' \Sigma^{-1} \delta + 2M\Delta \\ &= \frac{N_1}{N_1 + N_2} p + M\Delta, \end{aligned}$$

establishing (3.3.7).

$$\begin{aligned}
 R_3(\hat{\mu}^{(1)PT}; \Sigma^{-1}) &= E \left\{ N_1 (\hat{\mu}^{(1)PT} - \mu^{(1)})' \Sigma^{-1} (\hat{\mu}^{(1)PT} - \mu) \right\} \\
 &= EN_1 \left[\left\{ (\tilde{\mu}^{(1)} - \mu^{(1)}) + M(\tilde{\mu}^{(2)} - \mu^{(1)}) I(T_N^2 < T_{N,\alpha}^2) \right\}' \right. \\
 &\quad \left. \Sigma^{-1} \left\{ (\tilde{\mu}^{(1)} - \mu^{(1)}) + M(\tilde{\mu}^{(2)} - \mu^{(1)}) I(T_N^2 < T_{N,\alpha}^2) \right\} \right] \\
 &= p + M^2 N_1 E \left\{ (\tilde{\mu}^{(2)} - \tilde{\mu}^{(1)})' \Sigma^{-1} (\tilde{\mu}^{(2)} - \tilde{\mu}^{(1)}) I(T_N^2 < T_{N,\alpha}^2) \right\} \\
 &\quad + 2N_1 M E \left\{ (\tilde{\mu}^{(1)} - \mu^{(1)})' \Sigma^{-1} (\tilde{\mu}^{(2)} - \tilde{\mu}^{(1)}) I(T_N^2 < T_{N,\alpha}^2) \right\},
 \end{aligned}$$

by the transformation (3.3.5) and applying Theorem A.2 from Appendix A,

$$\begin{aligned}
 &= p - N_1 M^2 \left\{ (N_1 M)^{-1} p H_{p+2,m}(T_{N,\alpha}^{*2}; \Delta) + \delta' \Sigma^{-1} \delta H_{p+4,m}(T_{N,\alpha}^{*2}; \Delta) \right\} \\
 &\quad + 2M \Delta H_{p+2,m}(T_{N,\alpha}^{*2}; \Delta) \\
 &= p + \Delta M \left\{ 2H_{p+2,m}(T_{N,\alpha}^{*2}; \Delta) - H_{p+4,m}(T_{N,\alpha}^{*2}; \Delta) \right\} \\
 &\quad - p M H_{p+2,m}(T_{N,\alpha}^{*2}; \Delta),
 \end{aligned}$$

establishing (3.3.8).

Finally, the risk of $\hat{\mu}^{(1)S}$ is calculated as follows:

$$\begin{aligned}
 R_4(\hat{\mu}^{(1)S}; \Sigma^{-1}) &= N_1 E \left\{ (\hat{\mu}^{(1)S} - \mu^{(1)})' \Sigma^{-1} (\hat{\mu}^{(1)S} - \mu^{(1)}) \right\} \\
 &= N_1 E \left[\left\{ (\tilde{\mu}^{(1)} - \mu^{(1)}) + c\tilde{M}(\tilde{\mu}^{(2)} - \tilde{\mu}^{(1)})(nT_N^{-2}) \right\}' \Sigma^{-1} \right. \\
 &\quad \left. \left\{ (\tilde{\mu}^{(1)} - \mu^{(1)}) + cM(\tilde{\mu}^{(2)} - \tilde{\mu}^{(1)})(nT_N^{-2}) \right\} \right] \\
 &= p + N_1 M c^2 E \left\{ (\tilde{\mu}^{(2)} - \tilde{\mu}^{(1)})' \Sigma^{-1} (\tilde{\mu}^{(2)} - \tilde{\mu}^{(1)})(nT_N^{-2})^2 \right\} \\
 &\quad + 2N_1 M c E \left\{ (\tilde{\mu}^{(1)} - \mu^{(1)})' \Sigma^{-1} (\tilde{\mu}^{(2)} - \tilde{\mu}^{(1)})(nT_N^{-2}) \right\} \\
 &= p + M c^2 E \left\{ (y' \Sigma^{-1} y)(y' S^{-1} y)(n^2) \right\} \\
 &\quad + 2N_1 M^2 c \left[E \{ \delta' \Sigma^{-1} (\tilde{\mu}^{(2)} - \tilde{\mu}^{(1)})(nT_N^{-2}) \} \right. \\
 &\quad \left. - E \{ (\tilde{\mu}^{(2)} - \tilde{\mu}^{(1)})' \Sigma^{-1} (\tilde{\mu}^{(2)} - \tilde{\mu}^{(1)})(nT_N^{-2}) \} \right].
 \end{aligned}$$

By using transformation (3.3.5) and evaluating the expectation, we obtain

$$R_4(\hat{\mu}^{(1)S}; \Sigma^{-1}) = p + Mc^2m(m+2)E(\chi_p^{-2}(\Delta)) - 2Mcm \\ + 2Mcm\Delta E(\chi_{p+2}^{-2}(\Delta)),$$

by Theorem A.1 and A.2 from Appendix A.

$$\begin{aligned} R_4(\hat{\mu}^{(1)S}; \Sigma^{-1}) &= p + Mc^2m(m+2)E(\chi_p^{-2}(\Delta)) - 2mMc \{1 - \Delta E(\chi_{p+2}^{-2}(\Delta))\} \\ &= p + Mc^2m(m+2)E(\chi_p^{-2}(\Delta)) - 2Mmc(p-2)E(\chi_p^{-2}(\Delta)) \\ &= p - mMc \{2(p-2) - c(m+2)\} E(\chi_p^{-2}(\Delta)), \end{aligned}$$

using the identity $1 - (p-2)E(\chi_p^{-2}(\Delta)) = \Delta E(\chi_{p+2}^{-2}(\Delta))$, which establishes (3.3.9).

In order that the risk be smaller than p we must have

$$0 < c < \frac{2(p-2)}{m+2}$$

and the optimum value of c is $\frac{p-2}{m+2}$.

3.3.1. RISK ANALYSIS

In this section, we consider the risk analysis of the four estimators for $\mu^{(1)}$. First, we note that the risk of $\tilde{\mu}^{(1)}$ is a constant line with height p , while the risk for $\hat{\mu}^{(1)}$ is a straight line, function of Δ with slope M and intercept $\frac{N_1 p}{N_1 + N_2}$. The risk of $\tilde{\mu}^{(1)}$ and $\hat{\mu}^{(1)}$ intersect at $\Delta = p$. Further, $R_2(\hat{\mu}^{(1)}; \Sigma^{-1}) \leq R_1(\tilde{\mu}^{(1)}; \Sigma^{-1})$ accordingly as $\Delta \geq p$. Thus, the knowledge of Δ determines which estimator is to be preferred. Now, consider the risk of $\hat{\mu}^{(1)PT}$. First, we note that under H_0 the risks of $\tilde{\mu}^{(1)}$, $\hat{\mu}^{(1)}$ and $\hat{\mu}^{(1)PT}$ may be ordered according to the magnitude of the risks as follows:

$$\hat{\mu}^{(1)} > \hat{\mu}^{(1)PT} > \tilde{\mu}^{(1)}$$

since the risk of $\hat{\mu}^{(1)PT}$ under H_0 is given by

$$p - MpH_{p+2}(T_{N,\alpha}^{*2}; 0),$$

which is less than p .

For the non-null case

$$R_3(\hat{\mu}^{(1)PT}; \Sigma^{-1}) \leq p \text{ according as}$$

$$\Delta \geq p H_{p+2,m}(T_{N,\alpha}^{*2}; \Delta) \{2H_{p+2,m}(T_{N,\alpha}^{*2}; \Delta) - H_{p+4,m}(T_{N,\alpha}^{*2}; \Delta)\}.$$

Thus, we notice $\hat{\mu}^{(1)PT}$ dominates $\tilde{\mu}^{(1)}$ whenever

$$\Delta \leq p H_{p+2,m}(T_{N,\alpha}^{*2}; \Delta) \{2H_{p+2,m}(T_{N,\alpha}^{*2}; \Delta) - H_{p+4,m}(T_{N,\alpha}^{*2}; \Delta)\}.$$

They have the same risk when the equality sign holds.

Finally, we analyze the risk of $\hat{\mu}^{(1)S}$ relative to $\tilde{\mu}^{(1)}$ and $\hat{\mu}^{(1)PT}$. We notice at once that

$$R_4(\hat{\mu}^{(1)S}; \Sigma^{-1}) \leq p \text{ for all } \Delta,$$

and $R_4(\hat{\mu}^{(1)S} : \Sigma^{-1})$ is monotone in Δ , the smaller value being

$$p - \frac{mM(p-2)}{(m+2)} \quad \text{at } \Delta = 0,$$

and the larger is p . Thus, $\hat{\mu}^{(1)S}$ dominates $\tilde{\mu}^{(1)}$ and is an admissible estimator when compared with $\tilde{\mu}^{(1)}$. Secondly, let Δ_α be the point at which $R_3(\hat{\mu}^{(1)PT} : \Sigma^{-1})$ and $R_4(\hat{\mu}^{(1)S} : \Sigma^{-1})$ intersect for each α (level of significance), then for $\Delta \in [0, \Delta_\alpha]$, $\hat{\mu}^{(1)PT}$ dominates $\hat{\mu}^{(1)S}$ for a range of α -values while for $\Delta \in (\Delta_\alpha, \infty)$, $\hat{\mu}^{(1)S}$ dominates $\hat{\mu}^{(1)PT}$. Generally $\hat{\mu}^{(1)S}$ always dominates $\hat{\mu}^{(1)PT}$ for large values of α . Also, for large values of Δ , $\tilde{\mu}^{(1)}$ dominates $\hat{\mu}^{(1)}$. $\hat{\mu}^{(1)S}$ and $\hat{\mu}^{(1)PT}$ both estimators share a common property that as $\Delta \rightarrow \infty$ their risks converge to the common limit p . But the risk of $\hat{\mu}^{(1)S}$ always lies below this asymptotic value, while the risk of $\hat{\mu}^{(1)PT}$ is more than the risk of $\tilde{\mu}^{(1)}$ for $\Delta > p$ and remains above the asymptotic value, for some intermediate values of Δ though the difference may not be of significant amount depending on α .

The following theorem presents the dominance picture of the four estimators of $\mu^{(1)}$ under $H_0 : \mu^{(1)} = \mu^{(2)}$:

THEOREM 3.3.3. Under H_0 , the dominance picture of the estimators is:

$$\hat{\mu}^{(1)} > \hat{\mu}^{(1)PT} > \hat{\mu}^{(1)S} > \tilde{\mu}^{(1)}, \quad (3.3.10)$$

whenever

$$pH_{p+2,m}(T_{N,\alpha}^{*2}; 0) > m(m+2)^{-1}(p-2), \quad (3.3.11)$$

holds.

PROOF.

Consider the risk-differences under H_0

- (i) $R_1(\tilde{\mu}^{(1)}; \Sigma^{-1}) - R_4(\hat{\mu}^{(1)S}; \Sigma^{-1}) > 0$ for all Δ including 0.
- (ii) $R_4(\hat{\mu}^{(1)S}; \Sigma^{-1}) - R_3(\hat{\mu}^{(1)PT}; \Sigma^{-1}) = m \{ pH_{p+2,m}(T_{N,\alpha}^{*2}; 0) - m(m+2)^{-1}(p-2) \} > 0$, for a range of α -values satisfying (3.3.11).
- (iii) $R_3(\hat{\mu}^{(1)PT}; \Sigma^{-1}) - R_2(\hat{\mu}^{(1)}; \Sigma^{-1}) = Mp \{ 1 - H_{p+2,m}(T_{N,\alpha}^{*2}; 0) \} > 0$, for all α .

3.3.2. EFFICIENCY ANALYSIS AND RECOMMENDATIONS

In this section, we consider the risk efficiency of the four estimators for $\mu^{(1)}$ and provide a max-min rule for the optimum choice of the level of significance for the preliminary test of the null hypothesis $H_0 : \mu^{(1)} = \mu^{(2)}$. Tables for the relative efficiency (maximum and minimum) and Δ -value at which minimum efficiency of $\hat{\mu}^{(1)PT}$ occurs relative to $\tilde{\mu}^{(1)}$ are provided.

First, the relative efficiency of $\hat{\mu}^{(1)PT}$ compared to $\tilde{\mu}^{(1)}$, is given by

$$R.E.(\hat{\mu}^{(1)PT} : \tilde{\mu}^{(1)}) = \frac{R_1(\tilde{\mu}^{(1)}; \Sigma^{-1})}{R_2(\hat{\mu}^{(1)PT}; \Sigma^{-1})} = \frac{1}{1 + g(\Delta)} = E_3(\alpha, \Delta), \quad (3.3.12)$$

where

$$g(\Delta) = N_2 M [p^{-1} \Delta \{ 2H_{p+2,m}(T_{N,\alpha}^{*2}; \Delta) - H_{p+\mu,m}(T_{N,\alpha}^{*2}; \Delta) \} - H_{p+4,m}(T_{N,\alpha}^{*2}; \Delta)]. \quad (3.3.13)$$

For a given N_1 and N_2 , E_3 is a function of α and Δ . The function, $E_3(\alpha, \Delta)$, for $\alpha \neq 0$, has its maximum at $\Delta = 0$ with value $\{1 - MN_2 H_{p+2,m}(T_{N,\alpha}^{*2}; 0)\}^{-1} (> 1)$ and decreases crossing the line $E_3(\alpha, \Delta) = 1$ to a minimum $E_3(\alpha, \Delta_0) = E_0$ at $\Delta = \Delta_0$ as Δ increases,

then increases to a maximum towards 1 as Δ tends to infinity. In the case when $\Delta = 0$ and α -varies, we have $\max_{\alpha} E_3(\alpha, 0) = E_3(0, 0) = 1 + N_2 N_1^{-1}$. The value $E_3(\alpha, 0)$ decreases as α increases. On the other, hand if $\alpha = 0$ and Δ varies; then the curve $E_3(0, \Delta)$ and $E_3(1, \Delta) = 1$ intersect at $\Delta = p$. For a general α , $E_3(\alpha, \Delta)$ and $E_3(1, \Delta)$ will intersect in the interval $0 \leq \Delta \leq p$; the value of Δ at the intersection decreases as α increases. Therefore, for two different α , $E_3(\alpha_1, \Delta)$ and $E_3(\alpha_2, \Delta)$ will never intersect above $E_3(\alpha, \Delta) = 1$ but will intersect somewhere below 1. Thus, we have plotted $E_3(\alpha, \Delta)$ against Δ for $\alpha=0.05, 0.1, 0.2$ and 0.5 for various sample sizes and for $p=2, 4$. Figures 3 and 4 display such properties of the preliminary test estimator.

$N_1 = 8, N_2 = 4$ and $\alpha = 0.05, 0.1, 0.2, 0.5$

$N_1 = 8, N_2 = 8$ and $\alpha = 0.05, 0.1, 0.2, 0.5$

$N_1 = 10, N_2 = 12$ and $\alpha = 0.05, 0.1, 0.2, 0.5$

$N_1 = 10, N_2 = 4$ and $\alpha = 0.05, 0.1, 0.2, 0.5$

FIG.3 Relative efficiency of PTMLE to UMLE for various sample sizes, with $p=2$

$N_1 = 8, N_2 = 4$ and $\alpha = 0.05, 0.1, 0.2, 0.5$

$N_1 = 8, N_2 = 6$ and $\alpha = 0.05, 0.1, 0.2, 0.5$

$N_1 = 10, N_2 = 12$ and $\alpha = 0.05, 0.1, 0.2, 0.5$

$N_1 = 10, N_2 = 4$ and $\alpha = 0.05, 0.1, 0.2, 0.5$

FIG.4 Relative efficiency of PTMLE to UMLE for various sample sizes, with $p=4$

In order to choose an estimator with maximum relative efficiency we adopt the following procedure: If it is known that $0 \leq \Delta \leq p$, the estimator $\hat{\mu}^{(1)}$ is chosen since $E_3(0, \Delta)$ is the largest for all Δ in this range. However, if Δ is unknown, there is no way of choosing a uniformly best estimator. Thus, we preassign a value of the efficiency, say E_0 , that we are willing to accept then consider the set $A = \{\alpha \mid E_3(\alpha, \Delta) \geq E_0, \text{ for all } \Delta\}$ and the estimator is chosen which maximizes $E_3(\alpha, \Delta)$ over all $\alpha \in A$ and Δ . Thus, we solve for α such that $\max_{\alpha \in A} \min_{\Delta} E_3(\alpha, \Delta) = E_0$. Tables 3 to 6 provide the values of maximum relative efficiency E^* , minimum relative efficiency E_0 and the recommended corresponding size of the preliminary test α^* for $p = 2(1)5$ and for various values of $N_i (i = 1, 2)$. Tables of the relative efficiency for odd samples are provided in Appendix B. For example, if $N_1 = 10$, $N_2 = 16$ and $p = 2$, and the experimenter wishes to have an estimator with a minimum guaranteed efficiency of 0.7, then using Table 3 we recommend to him to select $\alpha^* = 0.10$, because such a choice of α^* would yield an estimator with a maximum efficiency of 1.7404 at $\Delta = 0$ and a minimum guaranteed efficiency of 0.7021.

Table 3 Maximum and minimum guaranteed efficiencies for $p=2$ (continued).

$a(N_1, N_2)$	(4,12)	(6,12)	(8,12)	(10,12)	(4,16)	(6,16)	(8,16)	(10,16)
.05 E^*	2.6489	2.2272	1.9774	1.8119	2.0326	2.4843	2.2049	2.0139
E_0	.5266	.5060	.5993	.6278	.5287	.5585	.5848	.6081
Δ_0	1.9713	1.7160	1.5185	1.4505	1.5183	1.3705	1.2533	1.1582
.10 E^*	2.1258	1.8791	1.7211	1.6115	2.0056	2.0285	1.8592	1.7464
E_0	.6333	.6679	.6965	.7205	.6325	.6591	.6821	.7021
Δ_0	6.5206	6.3059	6.2124	6.1475	6.2124	6.1475	6.0719	6.0105
.15 E^*	1.8261	1.6640	1.5551	1.4769	1.9182	1.7561	1.6477	1.5665
E_0	.7048	.7348	.7591	.7794	.7627	.7869	.8058	.8260
Δ_0	5.7064	5.6461	5.5560	5.4861	5.5709	5.4861	5.4302	5.3845
.20 E^*	1.6291	1.5150	1.4369	1.3799	1.6018	1.5784	1.5023	1.4439
E_0	.7589	.7846	.8053	.8221	.7864	.8064	.8235	.8381
Δ_0	5.2064	5.1670	5.0967	5.0412	5.0967	5.0429	4.9882	4.9422
.25 E^*	1.4490	1.3606	1.2979	1.2591	1.5218	1.4712	1.4058	1.3525
E_0	.8222	.8441	.8615	.8759	.8400	.8590	.8711	.8815
Δ_0	4.8887	4.8134	4.7508	4.7120	4.7508	4.7120	4.6772	4.6482
.30 E^*	1.3845	1.3291	1.2787	1.2450	1.4006	1.3557	1.3143	1.2810
E_0	.8377	.8502	.8609	.8688	.8439	.8596	.8719	.8773
Δ_0	4.5989	4.5365	4.4800	4.4437	4.4800	4.4537	4.4245	4.4005
.35 E^*	1.3038	1.2575	1.2234	1.1973	1.3218	1.2816	1.2503	1.2253
E_0	.8674	.8829	.8951	.9059	.8648	.8772	.8874	.8961
Δ_0	4.3621	4.3110	4.2724	4.2422	4.2724	4.2422	4.2178	4.1979
.40 E^*	1.2491	1.2048	1.1786	1.1583	1.2532	1.2229	1.1991	1.1798
E_0	.8925	.9052	.9153	.9234	.8901	.9064	.9089	.9169
Δ_0	4.1650	4.1223	4.0900	4.0646	4.0900	4.0646	4.0442	4.0275
.45 E^*	1.1989	1.1620	1.1418	1.1201	1.1985	1.1750	1.1574	1.1426
E_0	.9137	.9241	.9323	.9388	.9117	.9201	.9270	.9328
Δ_0	3.9967	3.9609	3.9338	3.9125	3.9338	3.9125	3.8954	3.8813
.50 E^*	1.1474	1.1269	1.1115	1.0994	1.1544	1.1371	1.1233	1.1120
E_0	.9317	.9400	.9465	.9518	.9369	.9467	.9423	.9469
Δ_0	3.8508	3.8208	3.7980	3.7801	3.7980	3.7801	3.7657	3.7538

Table 3 Maximum and minimum guaranteed efficiencies for $p=2$.

$q(W, N_2)$	(4,4)	(6,4)	(8,4)	(10,4)	(4,8)	(6,8)	(8,8)	(10,8)
.05 E^*	1.7580	1.5146	1.3853	1.3131	2.2749	1.9215	1.7094	1.5806
E_0	.5054	.6456	.6858	.7441	.5218	.5799	.6293	.6617
Δ_0	12.1306	9.9298	8.9239	8.3741	8.9209	8.3441	7.9713	7.7106
.10 E^*	1.5904	1.4135	1.3157	1.2553	1.9225	1.6857	1.5457	1.4532
E_0	.6341	.7227	.7763	.8123	.6343	.6839	.7215	.7511
Δ_0	8.9615	7.7106	7.1156	6.7021	7.1156	6.7621	6.5296	6.3650
.15 E^*	1.4821	1.3571	1.2707	1.2111	1.7606	1.5542	1.4316	1.3624
E_0	.7154	.7969	.8281	.8574	.7081	.7504	.7817	.8000
Δ_0	7.4691	6.6139	6.1899	5.7672	6.1899	5.9352	5.7664	5.6401
.20 E^*	1.3915	1.2768	1.2144	1.1745	1.6454	1.4238	1.3407	1.2933
E_0	.7739	.8420	.8643	.8841	.7690	.8091	.8253	.8453
Δ_0	6.5455	5.9187	5.5819	5.3119	5.5859	5.3999	5.2904	5.1670
.25 E^*	1.3192	1.2278	1.1771	1.1452	1.5439	1.3596	1.2803	1.2387
E_0	.8180	.8619	.8831	.9011	.8072	.8570	.8888	.9154
Δ_0	5.8900	5.4651	5.1482	4.9029	5.1482	4.9529	4.8855	4.8131
.30 E^*	1.2605	1.1873	1.1401	1.1101	1.4477	1.2733	1.2272	1.1945
E_0	.8549	.8925	.9147	.9311	.8427	.8870	.9056	.9293
Δ_0	5.4123	5.0189	4.7098	4.4814	4.7098	4.5804	4.5980	4.5305
.35 E^*	1.2121	1.1534	1.1204	1.0921	1.3736	1.2199	1.1839	1.1581
E_0	.8827	.9158	.9319	.9453	.8729	.9027	.9075	.9187
Δ_0	5.0284	4.7062	4.4370	4.2128	4.4370	4.3028	4.3621	4.3110
.40 E^*	1.1717	1.1249	1.0983	1.0819	1.3179	1.1764	1.1482	1.1278
E_0	.9062	.9312	.9455	.9543	.8966	.9155	.9257	.9348
Δ_0	4.7152	4.4506	4.2109	4.0259	4.2109	4.0239	4.1650	4.1223
.45 E^*	1.1379	1.1037	1.0791	1.0656	1.2725	1.1405	1.1185	1.1025
E_0	.9256	.9455	.9599	.9647	.9175	.9310	.9408	.9481
Δ_0	4.4734	4.2739	4.1141	3.9862	4.1141	4.0401	3.9967	3.9609
.50 E^*	1.1095	1.0802	1.0634	1.0524	1.2353	1.1107	1.0936	1.0812
E_0	.9417	.9573	.9665	.9721	.9347	.9456	.9534	.9592
Δ_0	4.2302	4.0495	3.9325	3.8321	3.9325	3.8921	3.8508	3.8208

Table 4 Maximum and minimum guaranteed efficiencies for $r = 3$ (continued)

		(4,12)	(6,12)	(8,12)	(10,12)	(1,16)	(6,16)	(8,16)	(10,16)
.05	E^*	2.8376	2.5435	2.0583	1.8737	3.1819	2.6437	2.3185	2.1007
	E_0	.5799	.0226	.6574	.6863	.5900	.0213	.6482	.6715
	Δ_0	10.5082	1.8744	9.5281	9.2821	9.5281	1.8744	9.0993	8.9364
.10	E^*	2.3123	2.0053	1.6147	1.5813	2.3916	2.1835	1.9810	1.8377
	E_0	.6833	.7155	.7467	.7608	.6886	.7150	.7272	.7502
	Δ_0	8.4215	8.1931	7.9812	7.8203	7.9842	7.8204	7.6925	7.5839
.15	E^*	1.9020	1.7837	1.6473	1.5514	2.1500	1.8908	1.7625	1.6616
	E_0	.7487	.7783	.8016	.8215	.7419	.7741	.7927	.8084
	Δ_0	7.5276	7.3662	7.1449	7.0201	7.1446	7.0201	6.9224	6.8475
.20	E^*	1.7731	1.6248	1.5233	1.4700	1.8452	1.7163	1.6167	1.5317
	E_0	.7967	.8214	.8498	.8563	.7901	.8172	.8327	.8458
	Δ_0	6.8798	6.7030	6.5749	6.4746	6.5749	6.4746	6.3751	6.3322
.25	E^*	1.6162	1.4715	1.4272	1.3700	1.6631	1.5307	1.4399	1.4125
	E_0	.8340	.8537	.8767	.8805	.8300	.8500	.8630	.8719
	Δ_0	6.3975	6.2727	6.1458	6.0335	6.1458	6.0335	5.9361	5.8451
.30	E^*	1.4951	1.4110	1.3593	1.3037	1.5200	1.4309	1.3551	1.2937
	E_0	.8711	.8832	.8947	.9000	.8632	.8773	.8884	.8975
	Δ_0	6.0157	5.8942	5.8043	5.7100	5.8043	5.7100	5.6199	5.5432
.35	E^*	1.3598	1.3345	1.2876	1.2323	1.4312	1.3473	1.3240	1.2807
	E_0	.8886	.8929	.9140	.9217	.8883	.8906	.8986	.9162
	Δ_0	5.7006	5.5975	5.5210	5.4619	5.5210	5.4619	5.4149	5.3700
.40	E^*	1.3231	1.2725	1.2350	1.2076	1.3410	1.2773	1.2636	1.2308
	E_0	.9002	.9010	.9309	.9372	.9008	.9080	.9255	.9317
	Δ_0	5.4130	5.3448	5.2792	5.2281	5.2792	5.2284	5.1869	5.1550
.45	E^*	1.2094	1.2210	1.1920	1.1698	1.2735	1.2399	1.2137	1.1929
	E_0	.9265	.9361	.9435	.9499	.9205	.9355	.9497	.9448
	Δ_0	5.2096	5.1217	5.0682	5.0244	5.0682	5.0244	4.9805	4.9611
.50	E^*	1.2085	1.1779	1.1552	1.1376	1.2182	1.1923	1.1719	1.1554
	E_0	.9411	.9458	.9516	.9595	.9406	.9467	.9516	.9557
	Δ_0	4.9949	4.9297	4.8809	4.8432	4.8809	4.8432	4.8130	4.7884

Table 4 Maximum and minimum guaranteed efficiencies for $p=3$.

(N, n, p)	(4,4)	(6,4)	(8,4)	(10,4)	(4,8)	(6,8)	(8,8)	(10,8)
5								
E^*	1.8121	1.5461	1.4112	1.3235	2.3979	1.9842	1.7597	1.6186
E_0	.4836	.6329	.7165	.7677	.5782	.6250	.6743	.7122
Δ_0	19.9230	14.2160	12.6801	10.9812	12.0161	10.9812	10.3082	9.8544
10								
E^*	1.6719	1.4500	1.3162	1.2787	2.0090	1.7729	1.6085	1.5020
E_0	.6249	.7415	.8028	.8405	.8706	.7249	.7637	.7929
Δ_0	11.0516	10.8908	9.6150	8.9200	9.6150	8.9200	8.4545	8.1901
15								
E^*	1.5609	1.3853	1.2907	1.2470	1.8417	1.6223	1.4973	1.4142
E_0	.7127	.8015	.8504	.8804	.7417	.7839	.8172	.8404
Δ_0	11.1227	10.272	8.7541	7.8189	8.3701	7.8180	7.5276	7.3002
20								
E^*	1.4705	1.3201	1.2498	1.2025	1.6059	1.5079	1.4106	1.3447
E_0	.7745	.8473	.8812	.9000	.7421	.8291	.8546	.8741
Δ_0	9.8502	8.2100	7.5244	7.1000	7.5244	7.1001	6.8798	6.7036
25								
E^*	1.3954	1.2701	1.2111	1.1717	1.5405	1.4176	1.3408	1.2879
E_0	.8207	.8788	.9084	.9270	.7647	.8019	.8228	.8392
Δ_0	8.7408	7.1100	6.5076	6.0000	6.5076	6.0017	6.3975	6.2527
30								
E^*	1.3320	1.2204	1.1600	1.1190	1.4398	1.3443	1.2813	1.2406
E_0	.8506	.9030	.9201	.9377	.7829	.8278	.8500	.8716
Δ_0	7.3314	6.0085	5.4157	5.0000	5.4159	5.0086	6.0157	5.8512
35								
E^*	1.2786	1.1845	1.1223	1.1011	1.3503	1.2812	1.2352	1.2005
E_0	.8851	.9222	.9400	.9522	.8084	.8355	.8529	.8731
Δ_0	7.2902	6.4585	6.0713	5.8170	6.0713	5.8470	5.7035	5.5975
40								
E^*	1.2317	1.1645	1.1279	1.1047	1.2982	1.2338	1.1945	1.1665
E_0	.9081	.9377	.9520	.9617	.8201	.8261	.8376	.8459
Δ_0	6.7000	6.0751	5.7185	5.5379	5.7185	5.5379	5.4330	5.3448
45								
E^*	1.1917	1.1303	1.1004	1.0870	1.2382	1.1912	1.1598	1.1372
E_0	.9269	.9503	.9621	.9695	.8270	.8405	.8497	.8565
Δ_0	6.5276	5.7475	5.4703	5.3074	5.4703	5.3076	5.2006	5.1247
50								
E^*	1.1570	1.1124	1.0876	1.0719	1.1929	1.1550	1.1300	1.1120
E_0	.9424	.9607	.9690	.9757	.8418	.8525	.8590	.8653
Δ_0	5.9505	5.4620	5.2261	5.0609	5.2261	5.0609	4.9049	4.9207

Table 5 Maximum and minimum guaranteed efficiencies for $p=4$ (continued).

(λ_1, λ_2)	(4,12)	(6,12)	(8,12)	(10,12)	(12,12)	(14,12)	(16,12)	(18,12)
0.05 E^*	236000	24166	24101	19112	19112	19112	19112	21541
E	6101	6767	6955	7131	7131	7131	7131	7131
Δ	12.7583	12.0311	11.5238	11.1577	11.5238	11.1577	10.8722	10.6161
0.10 E^*	24562	20888	18733	17518	20778	22106	26607	19014
E	7121	7494	7781	8001	8215	8511	8730	8913
Δ	10.5267	10.013	9.7114	9.4565	9.7114	9.1105	9.2046	9.1012
0.15 E^*	24167	18753	17088	16112	2277	20055	18398	17210
E	7751	8054	8285	8400	8515	8654	8712	8800
Δ	9.3104	8.6781	8.7104	8.5254	8.7104	8.5254	8.5745	8.2558
0.20 E^*	18802	17077	15822	14150	14150	17011	16755	15012
E	8190	8418	8600	8785	8911	8940	8980	8907
Δ	8.7445	8.2421	8.0132	7.8514	8.0132	7.8514	7.7022	7.0535
0.25 E^*	17100	15751	14820	13415	13415	15033	15030	14075
E	8541	8746	8899	9011	9111	9154	9155	9054
Δ	7.9400	7.0070	7.0122	7.0111	7.0111	7.0111	7.2003	7.2110
0.30 E^*	15700	14728	13907	12878	12878	13245	14560	13010
E	8811	8910	9011	9111	9211	9211	9211	9111
Δ	7.4700	7.2710	7.0210	7.0111	7.0111	7.0111	6.9202	6.8457
0.35 E^*	14712	13904	13334	12609	12609	12605	13764	13551
E	9031	9160	9271	9353	9411	9455	9457	9305
Δ	7.0843	6.9074	6.7780	6.6811	6.7780	6.6811	6.6044	6.5424
0.40 E^*	13855	13222	12768	12150	12150	12519	13103	12775
E	9212	9324	9409	9474	9527	9511	9479	9414
Δ	6.7485	6.5512	6.4835	6.3985	6.4835	6.3985	6.3315	6.2774
0.45 E^*	13147	12648	12287	11912	11912	12076	12549	12288
E	9562	9654	9721	9775	9811	9811	9811	9541
Δ	6.4515	6.3219	6.2365	6.1491	6.2365	6.1491	6.0915	6.0441
0.50 E^*	12554	12112	11875	11611	11611	11900	12079	11873
E	9818	9902	9967	10011	10011	10011	9967	9921
Δ	6.1921	6.0756	5.9903	5.9212	5.9903	5.9212	5.8737	5.8321

Table 5 Maximum and minimum guaranteed efficiencies for p=4

		(4.4)	(6.4)	(8.4)	(10.4)	(12.4)	(14.4)	(16.4)	(18.4)
5	E*	1.8158	1.8055	1.1244	1.887	2.1712	2.0201	1.7566	1.6415
	E	.4915	.6175	.7215	.7932	.8571	.9163	.9713	.9746
	Δ	32.7747	20.2073	17.8719	13.8815	13.8705	14.8815	12.7589	12.0344
10	E*	1.7209	1.4712	1.0511	1.112	1.1511	1.8298	1.6187	1.5326
	E	.5925	.7175	.8121	.871	.9163	.9459	.9777	.9877
	Δ	27.7629	15.6573	12.4877	11.2775	12.4706	11.2525	10.5267	10.3493
15	E*	1.6199	1.4175	1.0175	1.1312	1.021	1.6916	1.5404	1.4477
	E	.671	.821	.871	.91	.9711	.9879	.979	.9613
	Δ	17.119	12.212	11.7	12.94	12.9206	12.914	9.3404	8.9781
20	E*	1.5284	1.3199	1.077	1.235	1.1521	1.5059	1.4597	1.3767
	E	.7113	.871	.9121	.911	.9716	.9859	.9725	.9609
	Δ	15.119	11.119	11.77	12.987	12.97	12.987	8.7445	8.2528
25	E*	1.4396	1.21	1.002	1.111	1.0196	1.4725	1.3523	1.3211
	E	.776	.897	.931	.912	.9717	.9805	.9678	.95127
	Δ	13.119	10.119	10.935	11.11	12.988	12.9836	7.9459	7.5071
30	E*	1.3871	1.1195	1.0187	1.1171	1.0111	1.3951	1.3124	1.2723
	E	.812	.911	.911	.917	.9717	.9803	.9773	.9696
	Δ	11.879	10.119	10.911	11.115	12.981	12.9795	7.4719	7.2710
35	E*	1.3219	1.1181	1.0145	1.1119	1.0126	1.3098	1.2715	1.2503
	E	.8225	.9216	.9161	.9179	.9717	.98193	.9637	.9439
	Δ	10.509	9.509	10.91	11.112	12.971	12.9832	7.0843	6.9074
40	E*	1.2798	1.1141	1.1191	1.1113	1.0126	1.2769	1.2278	1.1939
	E	.8321	.919	.9161	.9173	.9713	.9847	.9460	.9519
	Δ	9.6691	7.9691	10.915	10.9719	12.985	12.9719	6.7485	6.5952
45	E*	1.2356	1.1139	1.1192	1.1127	1.01288	1.2296	1.1899	1.1623
	E	.9265	.9129	.9161	.9173	.9837	.9875	.9505	.9629
	Δ	8.9265	7.9265	10.911	10.912	12.989	12.9433	6.4515	6.3210
50	E*	1.1965	1.1171	1.1157	1.1161	1.01363	1.1885	1.1569	1.1314
	E	.9378	.9123	.9128	.9176	.9847	.9881	.9653	.9704
	Δ	8.2957	7.1129	10.971	10.967	12.971	12.9607	6.1921	6.0756

Table 6 Maximum and minimum guaranteed efficiencies for p = 5 (continued).

		(1,12)	(6,12)	(8,12)	(10,12)	(4,16)	(6,16)	(8,16)	(10,16)
05	E*	30488	24683	24442	17773	34112	28176	24112	24912
	E	6252	6770	7465	7475	6746	6894	7181	7417
	Δ	154553	113508	116649	17777	17777	17787	126696	123477
10	E*	26708	21501	19193	17773	17773	21509	24117	17454
	E	7281	7085	7585	8217	7484	7756	7756	8156
	Δ	127240	11790	114862	11777	114862	114194	103411	106295
15	E*	12016	10205	17341	17773	17773	21517	17772	1777
	E	7734	8217	8112	8112	8112	8112	8112	8112
	Δ	112815	107210	11777	17777	17777	17777	9827	9650
20	E*	19421	17777	10202	17773	17773	17773	17773	17773
	E	8330	8748	8748	8748	8748	8748	8748	8748
	Δ	106163	10777	10777	10777	10777	10777	10777	10777
25	E*	77225	7777	7777	7777	7777	7777	7777	7777
	E	8748	8777	8777	8777	8777	8777	8777	8777
	Δ	6777	7777	7777	7777	7777	7777	7777	7777
30	E*	17773	17773	17773	17773	17773	17773	17773	17773
	E	8777	8777	8777	8777	8777	8777	8777	8777
	Δ	7777	7777	7777	7777	7777	7777	7777	7777
35	E*	15285	14112	13020	17773	17773	14781	14176	17773
	E	9125	9125	9125	9125	9125	9125	9125	9125
	Δ	8777	8777	8777	8777	8777	8777	8777	8777
40	E*	14773	14112	13020	17773	17773	17773	17773	17773
	E	1292	1401	1486	9448	9448	9448	9448	9448
	Δ	81375	7777	7777	7777	7777	7777	7777	7777
45	E*	13536	12976	12777	12777	12777	12777	12777	12777
	E	1420	1520	1557	1557	1557	1557	1557	1557
	Δ	7777	7777	7777	7777	7777	7777	7777	7777
50	E*	12936	12468	12136	11874	11874	12665	12502	12121
	E	9543	9616	9668	9774	9774	9774	9774	9774
	Δ	74568	72677	71323	70705	71323	70705	69512	68576

Table 6 Maximum and minimum guaranteed efficiencies for $p=5$.

		(4,4)	(6,6)	(8,8)	(10,10)	(12,12)	(14,14)	(16,16)	(18,18)
05	E^*	1.8704	1.7779	1.6854	1.5929	1.5004	1.4079	1.3154	1.2229
	E_0	.8329	.5091	.1853	.7815	.4577	.1339	.8101	.4863
	Δ_0	39.1864	29.9114	20.6364	11.3614	2.0864	17.2307	15.4599	14.3508
10	E^*	1.8554	1.6654	1.4754	1.2854	1.0954	1.2214	1.0314	1.5512
	E_0	.4055	.7034	.8784	.9534	.9784	.9784	.9534	.8328
	Δ_0	50.9388	21.4742	10.0712	13.8264	10.0739	13.8984	12.7246	11.9896
15	E^*	1.6618	1.4418	1.2218	1.0018	.7818	1.7202	1.5720	1.4719
	E_0	.5433	.7815	.8787	.9537	.9787	.9537	.8407	.8742
	Δ_0	3.3243	17.5433	13.7559	12.1312	10.7532	11.1026	11.2615	10.7210
20	E^*	1.5704	1.3814	1.1914	1.0014	.8114	1.6104	1.4858	1.4037
	E_0	.6104	.8329	.9534	.9784	.9784	.9784	.8827	.9320
	Δ_0	27.8307	15.1114	12.7711	11.0114	10.2514	11.9215	17.3113	9.8625
25	E^*	1.5025	1.2925	1.0825	.8725	.6625	1.5149	1.4137	1.3456
	E_0	.7151	.8707	.9457	.9707	.9707	.9707	.9009	.9222
	Δ_0	22.9363	14.4753	11.3112	10.1511	11.2112	11.1811	9.5901	9.2163
30	E^*	1.4357	1.2057	1.0757	.9457	.8157	1.4352	1.3525	1.2901
	E_0	.7759	.8959	.9611	.9711	.9711	.9711	.9253	.9576
	Δ_0	19.1569	11.2296	10.0739	8.9182	10.8733	9.5186	9.0227	8.6984
35	E^*	1.3705	1.2557	1.1024	.9853	.8682	1.3676	1.2995	1.2529
	E_0	.8269	.9206	.9478	.9612	.9608	.9352	.9399	.9497
	Δ_0	10.9524	11.2420	9.7997	8.9725	9.7006	8.9725	8.5461	8.2654
40	E^*	1.3237	1.2184	1.1359	1.0441	.9523	1.3096	1.2537	1.2150
	E_0	.8649	.9309	.9735	.9796	.9794	.9399	.9516	.9595
	Δ_0	15.6422	10.8753	9.1643	8.5076	9.1449	8.5076	8.1375	7.8925
45	E^*	1.2763	1.1807	1.1029	1.0145	.9261	1.2594	1.2136	1.1816
	E_0	.8951	.9503	.9671	.9755	.9741	.9519	.9612	.9678
	Δ_0	13.5264	9.7580	8.0415	7.4117	8.0415	7.4117	7.7785	7.5636
50	E^*	1.2338	1.1549	1.0863	1.0074	.9285	1.2734	1.2157	1.1520
	E_0	.9192	.9611	.9741	.9805	.9795	.9317	.9691	.9741
	Δ_0	12.2849	9.1083	7.8213	7.4109	8.2123	7.4109	7.4508	7.2677

In order to compare $\hat{\mu}^{(1)S}$ with $\tilde{\mu}^{(1)}$, we at once observe that

$$E_4 = R.E.(\hat{\mu}^{(1)S} : \tilde{\mu}^{(1)}) = [1 - m(m+2)^{-1}p^{-1}N_2M(p-2)^2E(\chi_p^{-2}(\Delta))]^{-1} \geq 1 \quad (3.3.14)$$

The relative efficiency is a monotonically decreasing in Δ with maximum value

$$[1 - N_2Mp^{-1}(p-2)(m-2)^{-1}]^{-1}$$

at $\Delta = 0$ and as $\Delta \rightarrow \infty$, the relative efficiency equals 1. Thus, $\hat{\mu}^{(1)S}$ uniformly dominates $\tilde{\mu}^{(1)}$. We have also noticed that $\hat{\mu}^{(1)S}$ does not dominate $\hat{\mu}^{(1)PT}$ under H_0 for a range of α , the level of significance for the null hypothesis. However, except for a range of Δ near the origin, $\hat{\mu}^{(1)S}$ dominates $\hat{\mu}^{(1)PT}$ and both of the estimators share a common property that as $\Delta \rightarrow \infty$ their *R.E.* converge to the common limit 1. But the *R.E.* of $\hat{\mu}^{(1)S}$ always lies above this asymptotic value, while the *R.E.* of $\hat{\mu}^{(1)PT}$ may drop for some intermediate values of Δ though the drops may not be of significant amount depending on the α .

Thus, from the point of robust-efficiency both $\hat{\mu}^{(1)PT}$ and $\hat{\mu}^{(1)S}$ may be advocated leaning more towards $\hat{\mu}^{(1)S}$ since the size of Δ is generally unknown and unlikely to be small.

We have prepared Tables 7 to 9 for $p=3(1)5$ for the maximum relative efficiencies for the RMLE(E_1), PTMLE(E_3) and SMLE(E_4) relative to the UMLE and the intersecting relative efficiencies (E_{Δ_α}) of the PTMLE and the SMLE and the $\Delta\alpha$ -values value at which the intersection occurs. Table 7 to 9 show the range of values of Δ for which $\hat{\mu}^{(1)PT}$ dominates $\hat{\mu}^{(1)S}$ for selected values of α . For example, if $(N_1, N_2) = (6, 16)$ and $p = 3$ with $\alpha = 0.10$ and $0 \leq \Delta \leq 1.6328$, then from Table 7 PTMLE dominates SMLE. But, outside this interval SMLE dominates PTMLE. If $\Delta = 0$, then $\hat{\mu}^{(1)}$ has larger efficiency ($E_1 = 3.667$) than $\hat{\mu}^{(1)PT}$ and $\hat{\mu}^{(1)S}$ ($E_4 = 1.2791$) relative to $\tilde{\mu}^{(1)}$. Also, $E_3(0.10, \Delta)$ and

$E_4(0.10, \Delta)$ intersect at $\Delta_{0.10} = 1.6328$ and the intersecting efficiency of $\hat{\mu}^{(1)S}$ and $\hat{\mu}^{(1)PT}$ to $\hat{\mu}^{(1)}$ is 1.1631 ($E_{\Delta_{0.10}}$). It is interesting to note that for large α , $\hat{\mu}^{(1)S}$ dominates $\hat{\mu}^{(1)PT}$ uniformly. To examine this we have plotted the dominance picture of UMLE, RMLE, PTMLE and SMLE for $p = 3(1)5$, $\alpha = 0.05, 0.1, 0.2$, and 0.5 and various sample sizes. Figures 5 to 7 display such features.

Fig.5 Dominance picture of UMLE, RMLE, PTMLE and SMLE
for $p=3$, $N_1=8$, $N_2=12$, and various α values.

Fig. 6 Dominance picture of UMLE, RMLE, PTMLE and SMLE for $p=4$, $N_1=8$, $N_2=12$, and various α values.

$\alpha = 0.5$ $\alpha = 1.0$ $\alpha = 1.5$ $\alpha = 2.0$

Fig.7 Dominance picture of UMLE, RMLE, PTMLE and SMLE for $p=5$, $N_1=8$, $N_2=12$, and various α values.

Table 7 Maximum relative efficiencies of PTMLE, RMLE and SMLE and the intersecting efficiencies for PTMLE and SMLE for each α with corresponding Δ_{α} -values: $p=3$ (continued)

$\nu(\nu, N_2)$	(4,12)	(6,12)	(8,12)	(10,12)	(4,16)	(6,16)	(8,16)	(10,16)
5	E_1 4.0000 E_2 2.8376 E_3 1.2727 $E_{\Delta,11}$ 1.1294 Δ_{11} 2.0232	3.0000 2.3435 1.2414 1.1176 1.9810	2.5000 2.0589 1.2162 1.1075 1.9476	2.2000 1.8737 1.1957 1.0906 1.9206	5.0000 3.1819 1.3197 1.1497 1.9476	3.0067 2.6437 1.2791 1.1365 1.9206	3.0000 2.3185 1.2532 1.1260 1.8984	2.6000 2.1007 1.2316 1.1170 1.8797
10	E_1 4.0000 E_2 2.3123 E_3 1.2727 $E_{\Delta,11}$ 1.1417 Δ_{11} 1.7439	3.0000 2.0053 1.2414 1.1289 1.6978	2.5000 1.8147 1.2162 1.1179 1.6617	2.2000 1.6549 1.1957 1.1086 1.6328	5.0000 2.8913 1.3197 1.1637 1.6017	3.0067 2.1835 1.2791 1.1502 1.6328	3.0000 1.9810 1.2532 1.1387 1.6090	2.6000 1.8577 1.2316 1.1287 1.5892
15	E_1 4.0000 E_2 1.9929 E_3 1.2727 $E_{\Delta,11}$ 1.1534 Δ_{11} 1.5112	3.0000 1.7837 1.2414 1.1305 1.4523	2.5000 1.6473 1.2162 1.1277 1.4242	2.2000 1.5314 1.1957 1.1175 1.3937	5.0000 2.1906 1.3197 1.1778 1.4242	3.0067 1.9006 1.2791 1.1631 1.3937	3.0000 1.7625 1.2532 1.1505 1.3687	2.6000 1.6010 1.2316 1.1396 1.3478
20	E_1 4.0000 E_2 1.5751 E_3 1.2727 $E_{\Delta,20}$ 1.1656 Δ_{20} 1.2907	3.0000 1.6248 1.2414 1.1507 1.2890	2.5000 1.5234 1.2162 1.1379 1.1687	2.2000 1.4504 1.1957 1.1270 1.1664	5.0000 1.8452 1.3197 1.1928 1.1957	3.0067 1.7063 1.2791 1.1765 1.1664	3.0000 1.6067 1.2532 1.1631 1.1406	2.6000 1.5317 1.2316 1.1515 1.1180
25	E_1 4.0000 E_2 1.1162 E_3 1.2727 $E_{\Delta,25}$ 1.1794 Δ_{25} 1.0670	3.0000 1.5645 1.2414 1.1633 1.0118	2.5000 1.4272 1.2162 1.1496 0.9489	2.2000 1.3704 1.1957 1.1377 0.9346	5.0000 1.6641 1.3197 1.2059 0.9689	3.0067 1.5635 1.2791 1.1924 0.9346	3.0000 1.4693 1.2532 1.1775 0.9066	2.6000 1.4325 1.2316 1.1646 0.8832
30	E_1 4.0000 E_2 1.3951 E_3 1.2727 $E_{\Delta,30}$ 1.1959 Δ_{30} 0.8289	3.0000 1.4103 1.2414 1.1785 0.7698	2.5000 1.3504 1.2162 1.1635 0.7238	2.2000 1.3057 1.1957 1.1505 0.6871	5.0000 1.5289 1.3197 1.2305 0.7238	3.0067 1.4539 1.2791 1.2113 0.6871	3.0000 1.3975 1.2532 1.1949 0.6571	2.6000 1.3737 1.2316 1.1856 0.6321
35	E_1 4.0000 E_2 1.3908 E_3 1.2727 $E_{\Delta,35}$ 1.2168 Δ_{35} 0.6666	3.0000 1.3545 1.2414 1.1976 0.6027	2.5000 1.2876 1.2162 1.1609 0.4530	2.2000 1.2523 1.1957 1.1667 0.4134	5.0000 1.4242 1.3197 1.2567 0.4530	3.0067 1.3673 1.2791 1.2353 0.4134	3.0000 1.3239 1.2532 1.2169 0.3810	2.6000 1.2897 1.2316 1.2010 0.3540
40	E_1 4.0000 E_2 1.3231 E_3 1.2727 $E_{\Delta,40}$ 1.2439 Δ_{40} 0.2687	3.0000 1.2725 1.2414 1.2226 0.1991	2.5000 1.2356 1.2162 1.2040 0.1451	2.2000 1.2070 1.1957 1.1879 0.1020	5.0000 1.3410 1.3197 1.2918 0.1451	3.0067 1.2973 1.2791 1.2673 0.1020	3.0000 1.2636 1.2532 1.2462 0.0668	2.6000 1.2368 1.2316 1.2280 0.0376
45	E_1 4.0000 E_2 1.2604 E_3 1.2727 $E_{\Delta,45}$ 1.2604 Δ_{45} 0.0000	3.0000 1.2210 1.2414 1.2210 0.0000	2.5000 1.1920 1.2162 1.1920 0.0000	2.2000 1.1695 1.1957 1.1698 0.0000	5.0000 1.2735 1.3197 1.2735 0.0000	3.0067 1.2399 1.2791 1.2399 0.0000	3.0000 1.2137 1.2532 1.2137 0.0000	2.6000 1.1926 1.2316 1.1926 0.0000
50	E_1 4.0000 E_2 1.2085 E_3 1.2727 $E_{\Delta,50}$ 1.2083 Δ_{50} 0.0000	3.0000 1.1779 1.2414 1.1779 0.0000	2.5000 1.1552 1.2162 1.1552 0.0000	2.2000 1.1376 1.1957 1.1376 0.0000	5.0000 1.2182 1.3197 1.2182 0.0000	3.0067 1.1923 1.2791 1.1923 0.0000	3.0000 1.1719 1.2532 1.1719 0.0000	2.6000 1.1554 1.2316 1.1554 0.0000

Table 7 Maximum relative efficiencies of PTMLE, RMLE and SMLE and the intersecting efficiencies for PTMLE and SMLE for each α with corresponding Δ_{α} -values: $p=3$.

$\alpha \backslash (N_1, N_2)$	(4,4)	(6,6)	(8,4)	(10,4)	(8,8)	(6,8)	(8,8)	(10,8)
.05	E_1 2.0000 E_2 1.8121 E_3 1.1250 $E_{\Delta_{.05}}$ 1.0362 $\Delta_{.05}$ 2.4161	1.6667 1.5461 1.1111 1.0527 2.2586	1.5000 1.4112 1.0976 1.0450 2.1530	1.4000 1.3208 1.0862 1.0436 2.0701	3.0000 2.3970 1.2162 1.1038 2.1530	2.3333 1.9542 1.1887 1.0911 2.0784	2.0000 1.7597 1.1667 1.0827 2.0232	1.8000 1.6186 1.1489 1.0754 1.9810
.10	E_1 2.0000 E_2 1.6710 E_3 1.1250 $E_{\Delta_{.10}}$ 1.0598 $\Delta_{.10}$ 2.2034	1.6667 1.4560 1.1111 1.0567 2.0111	1.5000 1.3462 1.0976 1.0519 1.8887	1.4000 1.2787 1.0862 1.0473 1.8647	3.0000 2.0590 1.2162 1.1006 1.8887	2.3333 1.7729 1.1887 1.0993 1.8647	2.0000 1.6086 1.1667 1.0902 1.7439	1.8000 1.5020 1.1489 1.0824 1.6978
.15	E_1 2.0000 E_2 1.5000 E_3 1.1250 $E_{\Delta_{.15}}$ 1.0663 $\Delta_{.15}$ 2.0136	1.6667 1.3853 1.1111 1.0605 1.7903	1.5000 1.2937 1.0976 1.0556 1.6661	1.4000 1.2373 1.0862 1.0508 1.5700	3.0000 1.8375 1.2162 1.1178 1.6661	2.3333 1.6224 1.1887 1.1063 1.5760	2.0000 1.4973 1.1667 1.0973 1.5112	1.8000 1.4142 1.1489 1.0869 1.4623
.20	E_1 2.0000 E_2 1.4705 E_3 1.1250 $E_{\Delta_{.20}}$ 1.0679 $\Delta_{.20}$ 1.8313	1.6667 1.3261 1.1111 1.0644 1.5685	1.5000 1.2408 1.0976 1.0594 1.4557	1.4000 1.2025 1.0862 1.0548 1.3706	3.0000 1.6659 1.2162 1.1263 1.4707	2.3333 1.5079 1.1887 1.1149 1.3593	2.0000 1.4165 1.1667 1.1046 1.2907	1.8000 1.3447 1.1489 1.0957 1.2390
.25	E_1 2.0000 E_2 1.3954 E_3 1.1250 $E_{\Delta_{.25}}$ 1.0709 $\Delta_{.25}$ 1.6475	1.6667 1.2761 1.1111 1.0636 1.3762	1.5000 1.2124 1.0976 1.0636 1.2431	1.4000 1.1717 1.0862 1.0583 1.1464	3.0000 1.5795 1.2162 1.1558 1.2431	2.3333 1.4176 1.1887 1.1258 1.1464	2.0000 1.3408 1.1667 1.1129 1.0670	1.8000 1.2579 1.1489 1.1043 1.0118
.30	E_1 2.0000 E_2 1.3320 E_3 1.1250 $E_{\Delta_{.30}}$ 1.0754 $\Delta_{.30}$ 1.4548	1.6667 1.2354 1.1111 1.0735 1.1831	1.5000 1.1893 1.0976 1.0684 1.0182	1.4000 1.1469 1.0862 1.0629 0.9678	3.0000 1.4342 1.2162 1.1470 1.0162	2.3333 1.3445 1.1887 1.1343 0.9078	2.0000 1.2833 1.1667 1.1226 0.8269	1.8000 1.2406 1.1489 1.1123 0.7698
.35	E_1 2.0000 E_2 1.2780 E_3 1.1250 $E_{\Delta_{.35}}$ 1.0807 $\Delta_{.35}$ 1.2458	1.6667 1.1965 1.1111 1.0795 0.9503	1.5000 1.1524 1.0976 1.0743 0.7714	1.4000 1.1244 1.0862 1.0686 0.6519	3.0000 1.3593 1.2162 1.1666 0.7714	2.3333 1.2842 1.1887 1.1472 0.6519	2.0000 1.2352 1.1667 1.1347 0.5064	1.8000 1.2066 1.1489 1.1235 0.5027
.40	E_1 2.0000 E_2 1.2317 E_3 1.1250 $E_{\Delta_{.40}}$ 1.0872 $\Delta_{.40}$ 1.0120	1.6667 1.1645 1.1111 1.0869 0.6880	1.5000 1.1279 1.0976 1.0815 0.4923	1.4000 1.1047 1.0862 1.0758 0.3018	3.0000 1.2702 1.2162 1.1793 0.4923	2.3333 1.2338 1.1887 1.1640 0.3018	2.0000 1.1945 1.1667 1.1504 0.2657	1.8000 1.1665 1.1489 1.1381 0.1901
.45	E_1 2.0000 E_2 1.1917 E_3 1.1250 $E_{\Delta_{.45}}$ 1.0957 $\Delta_{.45}$ 0.7423	1.6667 1.1366 1.1111 1.0867 0.3836	1.5000 1.1064 1.0976 1.0913 0.1675	1.4000 1.0873 1.0862 1.0855 0.0238	3.0000 1.1582 1.2162 1.2022 0.1675	2.3333 1.1912 1.1887 1.1869 0.0238	2.0000 1.1598 1.1667 1.1598 0.0000	1.8000 1.1372 1.1489 1.1372 0.0000
.50	E_1 2.0000 E_2 1.1570 E_3 1.1250 $E_{\Delta_{.50}}$ 1.1071 $\Delta_{.50}$ 0.4219	1.6667 1.1088 1.1111 1.1103 0.0204	1.5000 1.0876 1.0976 1.0876 0.0000	1.4000 1.0719 1.0862 1.0719 0.0000	3.0000 1.1920 1.2162 1.1920 0.0000	2.3333 1.1550 1.1887 1.1550 0.0000	2.0000 1.1300 1.1667 1.1300 0.0000	1.8000 1.1120 1.1489 1.1120 0.0000

Table 8 Maximum relative efficiencies of PTMLE, RMLE and SMLE and the intersecting efficiencies for PTMLE and SMLE for each α with corresponding Δ_n -values: $p=4$ (continued)

$\alpha(N_1, N_2)$	(4,12)	(6,12)	(8,12)	(10,12)	(4,16)	(6,16)	(8,16)	(10,16)
.05	E_1 4.0000 E_2 2.9000 E_3 1.4048 $E_{A_{.05}}$ 1.2276 $\Delta_{.05}$ 2.3842	3.0000 2.4166 1.4062 1.2071 2.2218	2.5000 2.1031 1.3000 1.1895 2.2372	2.2000 1.9113 1.3228 1.1743 2.1852	5.0000 3.3462 1.5455 1.2697 2.2272	3.0667 2.7452 1.4823 1.2468 2.1852	3.0000 2.3894 1.4318 1.2271 2.1424	2.6000 2.1541 1.3907 1.2102 2.1067
.10	E_1 4.0000 E_2 2.4402 E_3 1.4018 $E_{A_{.10}}$ 1.2537 $\Delta_{.10}$ 1.9573	3.0000 2.0888 1.4062 1.2314 1.8639	2.5000 1.8753 1.3000 1.2119 1.7915	2.2000 1.7318 1.3228 1.1950 1.7338	5.0000 2.6478 1.5455 1.3040 1.7915	3.0667 2.2906 1.4823 1.2781 1.7338	3.0000 2.0607 1.4318 1.2559 1.6867	2.6000 1.9004 1.3907 1.2367 1.6477
.15	E_1 4.0000 E_2 2.1107 E_3 1.4018 $E_{A_{.15}}$ 1.2508 $\Delta_{.15}$ 1.5827	3.0000 1.8653 1.4062 1.2504 1.4816	2.5000 1.7088 1.3000 1.2349 1.4042	2.2000 1.6682 1.3228 1.2101 1.3427	5.0000 2.2373 1.5455 1.3397 1.4642	3.0667 2.0905 1.4823 1.3106 1.3427	3.0000 1.8398 1.4318 1.2855 1.2925	2.6000 1.7236 1.3907 1.2633 1.2515
.20	E_1 4.0000 E_2 1.8802 E_3 1.4018 $E_{A_{.20}}$ 1.3116 $\Delta_{.20}$ 1.2190	3.0000 1.7097 1.4062 1.2847 1.1120	2.5000 1.5822 1.3000 1.2603 1.0300	2.2000 1.4989 1.3228 1.2400 0.8713	5.0000 1.9031 1.5455 1.3809 1.0300	3.0667 1.7961 1.4823 1.3479 0.8651	3.0000 1.6785 1.4318 1.3195 0.9131	2.6000 1.5912 1.3907 1.2956 0.8699
.25	E_1 4.0000 E_2 1.7090 E_3 1.4048 $E_{A_{.25}}$ 1.3459 $\Delta_{.25}$.8474	3.0000 1.5738 1.4062 1.3180 0.7347	2.5000 1.4620 1.3000 1.2921 0.6480	2.2000 1.4177 1.3228 1.2669 0.5814	5.0000 1.7658 1.5455 1.4318 0.6480	3.0667 1.6433 1.4823 1.3534 0.5814	3.0000 1.5546 1.4318 1.3608 0.5270	2.6000 1.4575 1.3907 1.3527 0.4823
.30	E_1 4.0000 E_2 1.5503 E_3 1.4048 $E_{A_{.30}}$ 1.3503 $\Delta_{.30}$.4571	3.0000 1.3728 1.4062 1.3621 0.5370	2.5000 1.3067 1.3000 1.3315 0.2470	2.2000 1.3178 1.3228 1.3016 0.1776	5.0000 1.6167 1.5455 1.4569 0.2476	3.0667 1.5245 1.4823 1.4521 0.1776	3.0000 1.4563 1.4318 1.4137 0.1214	2.6000 1.4033 1.3907 1.3607 0.0753
.35	E_1 4.0000 E_2 1.4712 E_3 1.4048 $E_{A_{.35}}$ 1.4596 $\Delta_{.35}$.0306	3.0000 1.3904 1.4062 1.3904 0.0000	2.5000 1.3334 1.3600 1.3334 0.0000	2.2000 1.2509 1.3228 1.2600 0.0000	5.0000 1.5000 1.5455 1.5000 0.0000	3.0667 1.4295 1.4823 1.4295 0.0000	3.0000 1.3764 1.4318 1.3764 0.0000	2.6000 1.3351 1.3907 1.3351 0.0000
.40	E_1 4.0000 E_2 1.3855 E_3 1.4048 $E_{A_{.40}}$ 1.3855 $\Delta_{.40}$.0000	3.0000 1.3222 1.4062 1.3222 0.0000	2.5000 1.2768 1.3600 1.2768 0.0000	2.2000 1.2426 1.3228 1.2426 0.0000	5.0000 1.4065 1.5455 1.4065 0.0000	3.0667 1.3519 1.4823 1.3519 0.0000	3.0000 1.3103 1.4318 1.3103 0.0000	2.6000 1.2775 1.3907 1.2775 0.0000
.45	E_1 4.0000 E_2 1.3147 E_3 1.4048 $E_{A_{.45}}$ 1.3147 $\Delta_{.45}$.0000	3.0000 1.2648 1.4062 1.2648 0.0000	2.5000 1.2287 1.3600 1.2287 0.0000	2.2000 1.2012 1.3228 1.2012 0.0000	5.0000 1.3300 1.5455 1.3300 0.0000	3.0667 1.2876 1.4823 1.2876 0.0000	3.0000 1.2549 1.4318 1.2549 0.0000	2.6000 1.2288 1.3907 1.2288 0.0000
.50	E_1 4.0000 E_2 1.2554 E_3 1.4048 $E_{A_{.50}}$ 1.2554 $\Delta_{.50}$.0000	3.0000 1.2162 1.4062 1.2162 0.0000	2.5000 1.1875 1.3600 1.1875 0.0000	2.2000 1.1655 1.3228 1.1655 0.0000	5.0000 1.2667 1.5455 1.2667 0.0000	3.0667 1.2336 1.4823 1.2336 0.0000	3.0000 1.2079 1.4318 1.2079 0.0000	2.6000 1.1873 1.3907 1.1873 0.0000

Table 8 Maximum relative efficiencies of PTMLE, RMLE and SMLE and the intersecting efficiencies for PTMLE and SMLE for each α with corresponding Δ_{α} -values: $p=4$.

$\alpha(N_1, N_2)$	(4,4)	(6,4)	(8,4)	(10,4)	(4,8)	(6,8)	(8,8)	(10,8)
.05	E_1 2.0000 E_2 1.8453 E_3 1.1765 $E_{\Delta_{.05}}$ 1.0815 $\Delta_{.05}$ 3.1558	1.6667 1.5653 1.1667 1.0824 2.8484	1.5000 1.4244 1.1489 1.0776 2.6893	1.4000 1.3397 1.1324 1.0716 2.4922	3.0000 2.4713 1.3500 1.1632 2.6835	2.3333 2.0291 1.3051 1.1543 2.4322	2.0000 1.7903 1.2683 1.1410 2.3812	1.8000 1.6115 1.2385 1.1292 2.3018
.10	E_1 2.0000 E_2 1.7268 E_3 1.1765 $E_{\Delta_{.10}}$ 1.0857 $\Delta_{.10}$ 2.9145	1.6667 1.4848 1.1667 1.0885 2.5121	1.5000 1.3654 1.1489 1.0843 2.2553	1.4000 1.2932 1.1324 1.0783 2.0820	3.0000 2.1514 1.3500 1.1841 2.2558	2.3333 1.8238 1.3051 1.1700 2.0820	2.0000 1.6487 1.2683 1.1560 1.9573	1.8000 1.5326 1.2385 1.1432 1.8639
.15	E_1 2.0000 E_2 1.6466 E_3 1.1765 $E_{\Delta_{.15}}$ 1.0902 $\Delta_{.15}$ 2.6723	1.6667 1.4174 1.1667 1.0948 2.2605	1.5000 1.3159 1.1489 1.0811 1.9109	1.4000 1.2542 1.1324 1.0851 1.7186	3.0000 1.9234 1.3500 1.2034 1.9109	2.3333 1.6816 1.3051 1.1660 1.7186	2.0000 1.5491 1.2683 1.1712 1.5827	1.8000 1.4477 1.2385 1.1574 1.4518
.20	E_1 2.0000 E_2 1.5253 E_3 1.1765 $E_{\Delta_{.20}}$ 1.0952 $\Delta_{.20}$ 2.4232	1.6667 1.3599 1.1667 1.1018 1.8466	1.5000 1.2734 1.1489 1.0986 1.5724	1.4000 1.2205 1.1324 1.0926 1.5447	3.0000 1.7524 1.3500 1.2185 1.5724	2.3333 1.5659 1.3051 1.2040 1.3615	2.0000 1.4537 1.2683 1.1882 1.2186	1.8000 1.3787 1.2385 1.1739 1.1126
.25	E_1 2.0000 E_2 1.4526 E_3 1.1765 $E_{\Delta_{.25}}$ 1.1010 $\Delta_{.25}$ 2.1595	1.6667 1.3100 1.1667 1.1099 1.5690	1.5000 1.2303 1.1489 1.1074 1.2240	1.4000 1.1911 1.1324 1.1016 1.0016	3.0000 1.6189 1.3500 1.2126 1.2240	2.3333 1.4725 1.3051 1.2253 1.0016	2.0000 1.3823 1.2683 1.2084 .8474	1.8000 1.3211 1.2385 1.1921 .7347
.30	E_1 2.0000 E_2 1.3571 E_3 1.1765 $E_{\Delta_{.30}}$ 1.1079 $\Delta_{.30}$ 1.8739	1.6667 1.2685 1.1667 1.1136 1.2234	1.5000 1.2057 1.1489 1.1189 .8539	1.4000 1.1651 1.1324 1.1118 .6175	3.0000 1.5117 1.3500 1.2675 .8539	2.3333 1.3954 1.3051 1.2513 .6175	2.0000 1.3224 1.2683 1.2353 .4551	1.8000 1.2723 1.2385 1.2154 .3370
.35	E_1 2.0000 E_2 1.3299 E_3 1.1765 $E_{\Delta_{.35}}$ 1.1163 $\Delta_{.35}$ 1.5581	1.6667 1.2281 1.1667 1.1317 .8493	1.5000 1.1748 1.1489 1.1315 .4512	1.4000 1.1415 1.1324 1.1251 .2010	3.0000 1.4238 1.3500 1.3022 .4512	2.3333 1.3308 1.3051 1.2860 .2910	2.0000 1.2715 1.2683 1.2657 .0306	1.8000 1.2505 1.2385 1.2303 .0699
.40	E_1 2.0000 E_2 1.2708 E_3 1.1765 $E_{\Delta_{.40}}$ 1.1271 $\Delta_{.40}$ 1.2021	1.6667 1.1941 1.1667 1.1475 .4293	1.5000 1.1491 1.1489 1.1433 .0039	1.4000 1.1213 1.1324 1.1213 .0000	3.0000 1.3500 1.3500 1.3445 .0039	2.3333 1.2760 1.3051 1.2760 .0000	2.0000 1.2278 1.2683 1.2278 .0000	1.8000 1.1929 1.2385 1.1939 .0000
.45	E_1 2.0000 E_2 1.2356 E_3 1.1765 $E_{\Delta_{.45}}$ 1.1414 $\Delta_{.45}$.7925	1.6667 1.1639 1.1667 1.1639 .0000	1.5000 1.1202 1.1489 1.1202 .0000	1.4000 1.1027 1.1324 1.1027 .0000	3.0000 1.2828 1.3500 1.2828 .0000	2.3333 1.2229 1.3051 1.2229 .0000	2.0000 1.1899 1.2683 1.1899 .0000	1.8000 1.1622 1.2385 1.1622 .0000
.50	E_1 2.0000 E_2 1.1965 E_3 1.1765 $E_{\Delta_{.50}}$ 1.1613 $\Delta_{.50}$.3120	1.6667 1.1371 1.1667 1.1371 .0000	1.5000 1.1057 1.1489 1.1057 .0000	1.4000 1.081 1.1324 1.081 .0000	3.0000 1.2303 1.3500 1.2363 .0000	2.3333 1.1825 1.3051 1.1885 .0000	2.0000 1.1569 1.2683 1.1569 .0000	1.8000 1.1344 1.2385 1.1344 .0000

Table 9 Maximum relative efficiencies of PTMLE, RMLE and SMLE and the intersecting efficiencies for PTMLE and SMLE for each α with corresponding Δ_{α} -values: $p=5$ (continued)

$\alpha(N_1, N_2)$	(4,12)	(6,12)	(8,12)	(10,12)	(1,16)	(6,16)	(8,16)	(10,16)
.05	E_1 4.0000 E_2 3.0458 E_3 1.6000 $E_{\Delta_{.05}}$ 1.2985 $\Delta_{.05}$ 2.7563	3.6000 2.4683 1.5217 1.2734 2.6210	2.5000 2.1442 1.4599 1.2500 2.5158	2.2000 1.9573 1.4103 1.2312 2.4317	5.0000 3.4962 1.7241 1.3751 2.5158	3.6667 2.8176 1.6337 1.3340 2.4317	3.0000 2.4392 1.5625 1.3071 2.3631	2.6000 2.1912 1.5053 1.2839 2.3061
.10	E_1 4.0000 E_2 2.5368 E_3 1.6000 $E_{\Delta_{.10}}$ 1.3369 $\Delta_{.10}$ 2.1961	3.0000 2.1501 1.5217 1.3086 2.0103	2.5000 1.9199 1.4599 1.2840 1.9213	2.2000 1.7653 1.4103 1.2619 1.8274	5.0000 2.7605 1.7241 1.4182 1.9213	3.6667 2.3699 1.6337 1.3826 1.8274	3.0000 2.1187 1.5625 1.3516 1.7516	2.6000 1.9454 1.5053 1.3248 1.6892
.15	E_1 4.0000 E_2 2.2016 E_3 1.6000 $E_{\Delta_{.15}}$ 1.3709 $\Delta_{.15}$ 1.6937	3.0000 1.9265 1.5217 1.3471 1.5251	2.5000 1.7541 1.4599 1.3194 1.5081	2.2000 1.6658 1.4103 1.2848 1.2587	5.0000 2.3770 1.7241 1.4779 1.5921	3.6667 2.0757 1.6337 1.4356 1.2987	3.0000 1.8972 1.5625 1.3999 1.2191	2.6000 1.7696 1.5053 1.3689 1.1539
.20	E_1 4.0000 E_2 1.9624 E_3 1.6000 $E_{\Delta_{.20}}$ 1.4273 $\Delta_{.20}$ 1.2033	3.0000 1.7587 1.5217 1.3922 1.6266	2.5000 1.6262 1.4599 1.3606 1.6339	2.2000 1.5882 1.4103 1.3227 1.7000	5.0000 2.0552 1.7241 1.5471 1.8919	3.6667 1.8646 1.6337 1.4990 1.7910	3.0000 1.7324 1.5625 1.4572 1.7002	2.6000 1.6353 1.5053 1.4209 1.6425
.25	E_1 4.0000 E_2 1.7823 E_3 1.6000 $E_{\Delta_{.25}}$ 1.4859 $\Delta_{.25}$.7057	3.0000 1.6272 1.5217 1.4381 1.5221	2.5000 1.5256 1.4599 1.4124 1.5859	2.2000 1.4771 1.4103 1.3757 1.6111	5.0000 1.8458 1.7241 1.6375 1.3859	3.6667 1.7645 1.6337 1.5793 1.2511	3.0000 1.6042 1.5625 1.5295 1.582	2.6000 1.5288 1.5053 1.4866 1.316
.30	E_1 4.0000 E_2 1.6115 E_3 1.6000 $E_{\Delta_{.30}}$ 1.5000 $\Delta_{.30}$.1837	3.0000 1.5214 1.5217 1.5214 1.0000	2.5000 1.4392 1.4599 1.4392 1.0000	2.2000 1.3775 1.4103 1.3795 1.0000	5.0000 1.6861 1.7241 1.6801 1.0000	3.6667 1.5793 1.6337 1.5793 1.0000	3.0000 1.5014 1.5625 1.5014 1.0000	2.6000 1.4420 1.5053 1.4420 1.0000
.35	E_1 4.0000 E_2 1.5285 E_3 1.6000 $E_{\Delta_{.35}}$ 1.5285 $\Delta_{.35}$.0000	3.0000 1.4342 1.5217 1.4342 1.0000	2.5000 1.3686 1.4599 1.3686 1.0000	2.2000 1.3203 1.4103 1.3203 1.0000	5.0000 1.5004 1.7241 1.5603 1.0000	3.6667 1.4781 1.6337 1.4781 1.0000	3.0000 1.4170 1.5625 1.4170 1.0000	2.6000 1.3693 1.5053 1.3693 1.0000
.40	E_1 4.0000 E_2 1.4358 E_3 1.6000 $E_{\Delta_{.40}}$ 1.4358 $\Delta_{.40}$.0000	3.0000 1.3613 1.5217 1.3613 1.0000	2.5000 1.2687 1.4599 1.3087 1.0000	2.2000 1.2695 1.4103 1.2695 1.0000	5.0000 1.4588 1.7241 1.4588 1.0000	3.6667 1.3948 1.6337 1.3948 1.0000	3.0000 1.3466 1.5625 1.3466 1.0000	2.6000 1.3088 1.5053 1.3088 1.0000
.45	E_1 4.0000 E_2 1.3586 E_3 1.6000 $E_{\Delta_{.45}}$ 1.3586 $\Delta_{.45}$.0000	3.0000 1.2956 1.5217 1.2956 1.0000	2.5000 1.2573 1.4599 1.2573 1.0000	2.2000 1.2256 1.4103 1.2256 1.0000	5.0000 1.3753 1.7241 1.3753 1.0000	3.6667 1.3252 1.6337 1.3252 1.0000	3.0000 1.2570 1.5625 1.2570 1.0000	2.6000 1.2568 1.5053 1.2568 1.0000
.50	E_1 4.0000 E_2 1.2936 E_3 1.6000 $E_{\Delta_{.50}}$ 1.2936 $\Delta_{.50}$.0000	3.0000 1.2468 1.5217 1.2468 1.0000	2.5000 1.2130 1.4599 1.2130 1.0000	2.2000 1.1874 1.4103 1.1874 1.0000	5.0000 1.3057 1.7241 1.3057 1.0000	3.6667 1.2665 1.6337 1.2665 1.0000	3.0000 1.2362 1.5625 1.2362 1.0000	2.6000 1.2121 1.5053 1.2121 1.0000

Table 9 Maximum relative error E_{max} of PTMLE, TMLE¹ and SMLE and the intercepting efficiency for PTMLE, SMLE for each γ with $\alpha = 0.05$ and $\beta = 0.01$ (Table 5)

3.4. MEAN SQUARE ERROR MATRICES

In this section we derive the mean square error matrices of the four estimators. Let $\Gamma_1(\tilde{\mu}^{(1)})$, $\Gamma_2(\hat{\mu}^{(1)})$, $\Gamma_3(\hat{\mu}^{(1)PT})$, and $\Gamma_4(\hat{\mu}^{(1)S})$, denote mean square matrix of $\tilde{\mu}^{(1)}$, $\hat{\mu}^{(1)}$, $\hat{\mu}^{(1)PT}$, and $\hat{\mu}^{(1)S}$ respectively. Then,

$$\Gamma_1(\tilde{\mu}^{(1)}) = N_1^{-1}\Sigma, \quad (3.4.1)$$

$$\Gamma_2(\hat{\mu}^{(1)}) = (N_1 + N_2)^{-1}\Sigma + M^2\delta\delta', \quad (3.4.2)$$

$$\begin{aligned} \Gamma_3(\hat{\mu}^{(1)PT}) &= E(\hat{\mu}^{(1)PT} - \mu^{(1)})(\hat{\mu}^{(1)PT} - \mu^{(1)})' \\ &= E\left\{(\tilde{\mu}^{(1)} - \mu^{(1)}) + M_2(\tilde{\mu}^{(2)} - \tilde{\mu}^{(1)})I(T_N^2 \leq T_{N,\alpha}^2)\right\} \\ &\quad \left\{(\tilde{\mu}^{(1)} - \mu^{(1)}) + M_2(\tilde{\mu}^{(2)} - \tilde{\mu}^{(1)})I(T_N^2 \leq T_{N,\alpha}^2)\right\}' \end{aligned} \quad (3.4.3)$$

$$\begin{aligned} &= N^{-1}\Sigma + M^2 E(\tilde{\mu}^{(2)} - \tilde{\mu}^{(1)})(\tilde{\mu}^{(1)} - \tilde{\mu}^{(1)})' I(T_N^2 \leq T_{N,\alpha}^2) \\ &\quad + 2ME(\tilde{\mu}^{(1)} - \mu^{(1)})(\tilde{\mu}^{(2)} - \tilde{\mu}^{(1)})' I(T_N^2 \leq T_{N,\alpha}^2). \end{aligned} \quad (3.4.4)$$

Consider the second term of (3.4.4) and applying transformation (3.3.5)

$$E[(\tilde{\mu}^{(2)} - \tilde{\mu}^{(1)})(\tilde{\mu}^{(2)} - \tilde{\mu}^{(1)})' I(T_N^2 \leq T_{N,\alpha}^2)] = (N_1 M)^{-1} E[(yy') I(y'S^{-1}y < T_{N,\alpha}^2)],$$

retransforming y by $y^* = Dy$, then

$$\begin{aligned} &= (N_1 M)^{-1} E\left\{(D^{-1}y^*y^{*'}D^{-1'})I(y^{*'}y^*/\chi_m^2 < T_{N,\alpha}^{*2})\right\} \\ &= (N_1 M)^{-1} \left\{H_{p+2,m}(T_{N,\alpha}^2; \Delta)\Sigma + N_1 M\delta\delta'H_{p+4,m}(T_{N,\alpha}^{*2}; \Delta)\right\}, \end{aligned} \quad (3.4.5)$$

by Theorem A.3 from Appendix A. Now we consider the third term of (3.4.4) and using the identity

$$\tilde{\mu}^{(1)} - \mu^{(1)} = \hat{\mu}^{(1)} - \mu^{(1)} - M(\tilde{\mu}^{(2)} - \tilde{\mu}^{(1)}), \text{ then}$$

$$\begin{aligned}
E[(\tilde{\mu}^{(1)} - \mu^{(1)})(\tilde{\mu}^{(1)} - \mu^{(1)})' I(T_N^2 \leq T_{N,\alpha}^2)] &= E \left\{ \left(\tilde{\mu}^{(1)} - \frac{N_1 \mu^{(1)} + N_2 \mu^{(2)}}{N_1 + N_2} \right) (\tilde{\mu}^{(2)} - \tilde{\mu}^{(1)})' \right\} \\
&\quad + M \delta E \left\{ (\tilde{\mu}^{(2)} - \tilde{\mu}^{(1)})' I(T_N^2 \leq T_{N,\alpha}^2) \right\} \\
&\quad - M E \left\{ (\tilde{\mu}^{(2)} - \tilde{\mu}^{(1)})(\tilde{\mu}^{(2)} - \tilde{\mu}^{(1)})' \right. \\
&\quad \left. \cdot I(T_N^2 \leq T_{N,\alpha}^2) \right\}. \tag{3.4.6}
\end{aligned}$$

Combining all these terms, we obtain

$$\begin{aligned}
\Gamma_3(\tilde{\mu}^{(1)PT}) &= N_1^{-1} \Sigma - N_1^{-1} M \{ H_{p+2,m}(T_{N,\alpha}^{*2}; \Delta) + N_1 M \delta \delta' H_{p+4,m}(T_{N,\alpha}^{*2}; \Delta) \} \\
&\quad + 2M^2 \delta \delta' H_{p+2,m}(T_{N,\alpha}^{*2}; \Delta) \\
&= N_1^{-1} \{ \Sigma - N_1^{-1} M H_{p+2,m}(T_{N,\alpha}^{*2}; \Delta) \Sigma \\
&\quad + M^2 \delta \delta' \{ 2H_{p+2,m}(T_{N,\alpha}^{*2}; \Delta) - H_{p+4,m}(T_{N,\alpha}^{*2}; \Delta) \} \}. \tag{3.4.7}
\end{aligned}$$

Finally, consider the mean square error matrix of $\tilde{\mu}^{(1)S}$ $\diamond$

$$\begin{aligned}
\Gamma_4(\tilde{\mu}^{(1)S}) &= E(\tilde{\mu}^{(1)S} - \mu^{(1)})(\tilde{\mu}^{(1)S} - \mu^{(1)})' \\
&= E \left\{ (\tilde{\mu}^{(1)} - \mu^{(1)}) + M c (\tilde{\mu}^{(2)} - \tilde{\mu}^{(1)})(n T_N^{-2}) \right\} \\
&\quad \left\{ (\tilde{\mu}^{(1)} - \mu^{(1)}) + M c (\tilde{\mu}^{(2)} - \tilde{\mu}^{(1)})(n T_N^{-2}) \right\}' \\
&= E \left\{ (\tilde{\mu}^{(1)} - \mu^{(1)})(\tilde{\mu}^{(1)} - \mu^{(1)})' \right\} \\
&\quad + M c^2 E \left\{ (\tilde{\mu}^{(2)} - \tilde{\mu}^{(1)})(\tilde{\mu}^{(2)} - \tilde{\mu}^{(1)})'(n T_N^{-2})^2 \right\} \\
&\quad + 2M c E \left\{ (\tilde{\mu}^{(1)} - \mu^{(1)})(\tilde{\mu}^{(2)} - \tilde{\mu}^{(1)})'(n T_N^{-2}) \right\}. \tag{3.4.8}
\end{aligned}$$

Consider the second term of (3.4.8)

$$\begin{aligned}
E \left\{ (\tilde{\mu}^{(2)} - \tilde{\mu}^{(1)})(\tilde{\mu}^{(2)} - \tilde{\mu}^{(1)})'(n T_N^{-2})^2 \right\} &= (N_1 M)^{-1} E \{ (y y') (y' S^{-1} y / n)^2 \} \\
&= (N_1 M)^{-1} E \{ (D^{-1} y \cdot y \cdot D^{-1}) (y \cdot y \cdot / \chi_m^2) \} \\
&= (N_1 M)^{-1} m(m+2) \cdot \\
&\quad \{ \Sigma E(\chi_{p+2}^2(\Delta)) + N_1 M \delta \delta' E(\chi_{p+4}^{-2}(\Delta)) \} \tag{3.4.9}
\end{aligned}$$

2

by applying transformations and Theorem A.3 from Appendix A.

The third term of (3.4.8) yields to

$$mM \left[\delta\delta' E(\chi_{p+2}^{-2}(\Delta)) - \{ \Sigma E(\chi_{p+2}^{-2}(\Delta) + N_1 M \delta\delta' E(\chi_{p+4}^{-2}(\Delta)) \} \right], \quad (3.4.10)$$

yielding the mean square error matrix of $\hat{\mu}^{(1)S}$ as

$$\begin{aligned} \Gamma_4(\hat{\mu}^{(1)S}) &= N_1^{-1} \Sigma - N_1^{-1} c m M \{ 2E(\chi_{p+2}^{-2}(\Delta)) - c(m+2)E(\chi_{p+2}^{-4}(\Delta)) \} \Sigma \\ &\quad - \{ \{ c(m+2) + 4 \} N_1 M \delta\delta' E(\chi_{p+2}^{-4}(\Delta)) \}. \end{aligned}$$

Now, we establish the following theorem for the dominance picture of four estimators.

THEOREM 3.4.1. Under the null hypothesis $H_0 : \mu^{(1)} = \mu^{(2)}$,

$$\Gamma_1(\hat{\mu}^{(1)}) - \Gamma_2(\hat{\mu}^{(1)}), \quad \Gamma_1(\hat{\mu}^{(1)}) - \Gamma_3(\hat{\mu}^{(1)PT}), \quad \Gamma_3(\hat{\mu}^{(1)PT}) - \Gamma_2(\hat{\mu}^{(1)})$$

are positive definite and $\Gamma_4(\hat{\mu}^{(1)S}) - \Gamma_3(\hat{\mu}^{(1)PT})$ is positive definite, for a range of α .

PROOF.

Since Σ is positive definite and

$$\Gamma_1(\hat{\mu}^{(1)}) - \Gamma_2(\hat{\mu}^{(1)}) = N_1^{-1} - (N_1 + N_2)^{-1} \Sigma = N_1^{-1} M \Sigma \quad (3.4.11)$$

$$\Gamma_1(\hat{\mu}^{(1)}) - \Gamma_3(\hat{\mu}^{(1)PT}) = N_1^{-1} M H_{p+2,m}(T_{N,\alpha}^{*2}; 0) \Sigma \quad (3.4.12)$$

$$\Gamma_3(\hat{\mu}^{(1)PT}) - \Gamma_2(\hat{\mu}^{(1)}) = N_1^{-1} M \{ 1 - H_{p+2,m}(T_{N,\alpha}^{*2}; 0) \} \Sigma, \quad (3.4.13)$$

are positive definite,

$$\begin{aligned} \Gamma_4(\hat{\mu}^{(1)S}) - \Gamma_3(\hat{\mu}^{(1)PT}) &= N_1^{-1} M \{ H_{p+2,m}(T_{N,\alpha}^{*2}; 0) \\ &\quad - m p^{-1} (p-2)(m+2)^{-1} \} \Sigma, \end{aligned} \quad (3.4.14)$$

is positive definite if $H_{p+2,m}(T_{N,\alpha}^{*2}; 0) > mp^{-1}(p-2)(m+2)^{-1}$.

3.4.1. GENERALIZED EFFICIENCY ANALYSIS AND RECOMMENDATIONS

In this subsection we focus on the efficiencies of $\hat{\mu}^{(1)S}$, $\hat{\mu}^{(1)PT}$ and $\hat{\mu}^{(1)}$ relative to $\tilde{\mu}^{(1)}$ and provide a max-min rule for the optimum choice of the level of significance for the preliminary test.

We define the generalized relative efficiency (GRE) in terms of the generalized variances (p th root of the determinant of the mean square error matrices). First, consider the generalized relative efficiency of RMLE with respect to UMLE:

$$\begin{aligned} \text{GRE}_1(\hat{\mu}^{(1)} : \tilde{\mu}^{(1)}) &= \frac{|N_1^{-1}\Sigma|^{\frac{1}{p}}}{|(N_1 + N_2)\Sigma + M^2\delta\delta'|^{\frac{1}{p}}} \\ &= |N_1(N_1 + N_2)^{-1}I_p + N_1M^2\Sigma^{-1}\delta\delta'|^{-\frac{1}{p}} \\ &= \left\{ \prod_{j=1}^p N_1(N_1 + N_2)^{-1} + M\ell_j \right\}^{-\frac{1}{p}} \\ &= \{N_1(N_1 + N_2)^{-1} + M\Delta\}^{-\frac{1}{p}}, \end{aligned} \quad (3.4.15)$$

where $\ell_1 = \Delta$ and $\ell_2 = \dots = \ell_p = 0$. Since, GRE_1 is constant for fixed Δ -value and

$$\text{GRE}_1 = \left(1 + \frac{N_2}{N_1}\right)^{-\frac{1}{p}} \text{ for } \Delta = 0, \quad \lim_{\Delta \rightarrow \infty} \text{GRE}_1 = 0.$$

Then for large values of Δ , RMLE becomes inefficient and if $0 \leq \Delta \leq p$, then RMLE is more efficient than other estimators.

Next, we consider the GRE of PTMLE compared to UMLE:

$$\begin{aligned} \text{GRE}_2(\hat{\mu}^{(1)PT} : \tilde{\mu}^{(1)}) &= |N_1^{-1}\Sigma|^{\frac{1}{p}} / |N_1^{-1}\Sigma - N_1^{-1}MH_{p+2,m}(T_{N,\alpha}^{*2}; \Delta)\Sigma \\ &\quad + M^2\delta\delta' \{2H_{p+2,m}(T_{N,\alpha}^{*2}) - H_{p+4,m}(T_{N,\alpha}^{*2}; \Delta)\}|^{\frac{1}{p}} \\ &= \{1 + G_2(\Delta)\}^{-\frac{1}{p}}, \end{aligned} \quad (3.4.16)$$

where

$$G_2(\Delta) = M \left[\Delta \{ 2H_{p+2,m}(T_{N,\alpha}^{*2}; \Delta) - H_{p+4,m}(T_{N,\alpha}^{*2}; \Delta) \} - H_{p+2,m}(T_{N,\alpha}^{*2}; \Delta) \right]. \quad (3.4.17)$$

For a given (N_1, N_2) GRE_2 is a function of α and Δ . This function, for $\alpha \neq 0$, has its maximum under the null hypothesis with the value

$$GRE_2(\alpha, 0) = \{ 1 - M H_{p+2,m}(T_{N,\alpha}^{*2}; 0) \}^{-\frac{1}{p}} (> 1). \quad (3.4.18)$$

This function decreases, crossing the line $GRE_2(\alpha, \Delta) = 1$, to a minimum as Δ increases, then increases to a maximum towards 1 as Δ tends to infinity. If $\Delta \neq 0$ and α arbitrary, then

$$\max_{\alpha} GRE_2(\alpha, 0) = GRE_2(0, 0) = \left(1 + \frac{N_2}{N_1} \right)^{\frac{1}{p}}.$$

The value of $GRE_2(\alpha, 0)$ decreases as α increases. On the other hand, if $\alpha = 0$ and Δ vary; then the curve $GRE_2(\alpha, \Delta)$ and $GRE_2(1, \Delta)$ intersect at $\Delta = p$. For a general α , $GRE_2(\alpha, \Delta)$ and $GRE_2(1, \Delta)$ will intersect in the interval $0 \leq \Delta \leq p$. The value of Δ at the intersection decreases as α increases. Therefore, for two different α , $GRE_2(\alpha_1, \Delta)$ and $GRE_2(\alpha_2, \Delta)$ will never intersect above $GRE_2(\alpha, \Delta) = 1$ but will intersect somewhere below it.

In order to choose an estimator with the maximum generalized relative efficiency we adopt the following procedure: If it is known that $0 \leq \Delta \leq p$, the estimator $\hat{\mu}^{(1)}$ is chosen since $GRE_2(0, \Delta)$ is the largest for all Δ in this range. If $\Delta > p$, there is no way of choosing a uniformly best estimator. Thus, we preassign a value of the efficiency, E_0 , that we are willing to accept. Then if $A = \{ \alpha \mid GRE_2(\alpha, \Delta) \geq E_0, \text{ for all } \Delta \}$ and the estimator is

chosen which maximizes $\text{GRE}_3(\alpha, \Delta)$ over all $\alpha \in A$ and Δ . Then, we solve for α such that

$$\max_{\alpha \in A} \min_{\Delta} \text{GRE}_3(\alpha, \Delta) = E_0.$$

Tables 10 to 13 provide the values of the maximum generalized relative efficiency, E^* , minimum generalized relative efficiency, E_0 , and the recommended corresponding size of the preliminary test $\alpha^* = 0.5(0.05)0.50$ for $p = 2(1)5$ and for various values of $N_i (i = 1, 2)$. For example, if $N_1 = 10, N_2 = 16$ and $p = 2$ and if an estimator with a minimum guaranteed efficiency 0.71 is desired, then using Table 10 select $\alpha^* = 0.10$, because such a choice of α^* would yield an estimator with a maximum efficiency of 1.3191 at $\Delta = 0$.

Table 10 Maximum and minimum guaranteed generalized efficiencies for $p=2$ (continued).

α	(N_1, N_2)	(4,12)	(6,12)	(8,12)	(10,12)	(4,16)	(6,16)	(8,16)	(10,16)
.05	E^*	1.0276	1.4924	1.4362	1.3961	1.7125	1.5702	1.4819	1.4191
	E_0	.5756	.6055	.6312	.6535	.5761	.5788	.6190	.6372
	Δ_0	7.3565	7.0955	6.9076	6.7651	1.9076	1.7681	1.6487	1.5558
.10	E^*	1.4580	1.3705	1.3119	1.2694	1.7052	1.4225	1.3635	1.3193
	E_0	.6553	.6828	.7061	.7261	.6536	.6749	.6935	.7101
	Δ_0	5.9047	5.7453	5.6256	5.5347	5.6256	5.5347	5.4616	5.4022
.15	E^*	1.3513	1.2960	1.2470	1.2155	1.3814	1.3252	1.2856	1.2516
	E_0	.7113	.7364	.7574	.7752	.7087	.7282	.7452	.7602
	Δ_0	5.1451	5.0235	4.9430	4.8753	4.9119	4.8558	4.8221	4.7783
.20	E^*	1.2763	1.2311	1.1957	1.1744	1.2968	1.2504	1.2257	1.2016
	E_0	.7500	.7780	.7973	.8131	.7520	.7707	.7860	.8001
	Δ_0	4.6155	4.5503	4.4933	4.4371	4.5100	4.4471	4.3953	4.3411
.25	E^*	1.2203	1.1800	1.1410	1.1120	1.2100	1.2017	1.1814	1.1620
	E_0	.7925	.8137	.8302	.8447	.7903	.8063	.8199	.8318
	Δ_0	4.2511	4.2000	4.1563	4.1141	4.1503	4.1144	4.0808	4.0474
.30	E^*	1.1746	1.1398	1.1068	1.0755	1.1851	1.1615	1.1404	1.1211
	E_0	.8250	.8417	.8581	.8701	.8227	.8369	.8489	.8595
	Δ_0	3.9589	3.9408	3.9070	3.8763	3.9070	3.8628	3.8353	3.8127
.35	E^*	1.1419	1.1214	1.1061	1.0942	1.1497	1.1321	1.1182	1.1069
	E_0	.8542	.8696	.8821	.8921	.8512	.8636	.8740	.8820
	Δ_0	3.7713	3.7234	3.6871	3.6554	3.6871	3.6588	3.6360	3.6172
.40	E^*	1.1136	1.0976	1.0856	1.0763	1.1155	1.1058	1.0950	1.0862
	E_0	.8790	.8923	.9029	.9117	.8765	.8870	.8959	.9035
	Δ_0	3.5825	3.5426	3.5124	3.4888	3.5124	3.4888	3.4697	3.4511
.45	E^*	1.0904	1.0780	1.0686	1.0612	1.0948	1.0842	1.0758	1.0689
	E_0	.9009	.9121	.9210	.9282	.9045	.9075	.9150	.9214
	Δ_0	3.4224	3.3892	3.3639	3.3441	3.3639	3.3441	3.3282	3.3151
.50	E^*	1.0712	1.0616	1.0543	1.0485	1.0724	1.0663	1.0598	1.0545
	E_0	.9201	.9293	.9366	.9426	.9180	.9255	.9317	.9370
	Δ_0	3.2845	3.2567	3.2356	3.2191	3.2356	3.2191	3.2057	3.1947

Table 10 Maximum and minimum guaranteed generalized efficiencies for $p=2$.

$\phi(N_1, N_2)$		(4,4)	(6,4)	(8,4)	(10,4)	(4,6)	(6,6)	(8,6)	(10,6)
.05	E^*	1.3261	1.2397	1.1787	1.1459	1.3083	1.3226	1.3074	1.2572
	E_0	.5663	.6492	.7040	.7434	.5740	.6179	.6530	.6818
	Δ_0	11.4314	9.3093	8.2801	7.7153	8.2801	7.7153	7.3595	7.0955
.10	E^*	1.2647	1.1889	1.1470	1.1204	1.2965	1.2983	1.2433	1.2055
	E_0	.6020	.7322	.7772	.8089	.6578	.6973	.7282	.7531
	Δ_0	18.2684	7.0516	6.4729	6.1331	6.4729	6.1301	5.9047	5.7453
.15	E^*	1.2174	1.1563	1.1222	1.1034	1.2890	1.2385	1.1966	1.1672
	E_0	.7254	.7854	.8231	.8493	.7156	.7510	.7783	.8000
	Δ_0	6.7879	5.9711	5.5516	5.3071	5.5516	5.3074	5.1471	5.0203
.20	E^*	1.1796	1.1330	1.1019	1.0839	1.2831	1.1932	1.1505	1.1272
	E_0	.7735	.8248	.8567	.8715	.7615	.7927	.8165	.8354
	Δ_0	5.8703	5.2651	4.9565	4.7703	4.9565	4.7701	4.6335	4.5363
.25	E^*	1.1436	1.1081	1.0870	1.0701	1.2802	1.1574	1.1315	1.1139
	E_0	.8122	.8501	.8829	.9014	.7991	.8208	.8475	.8638
	Δ_0	5.2427	4.7702	4.5278	4.3853	4.5278	4.3803	4.2811	4.2009
.30	E^*	1.1227	1.0896	1.0707	1.0584	1.2809	1.1284	1.1078	1.0920
	E_0	.8444	.8816	.9041	.9193	.8314	.8555	.8735	.8874
	Δ_0	4.7691	4.3937	4.1948	4.0795	4.1988	4.0795	3.9589	3.8498
.35	E^*	1.1009	1.0740	1.0585	1.0484	1.2886	1.1045	1.0881	1.0761
	E_0	.8717	.9030	.9218	.9344	.8594	.8801	.8955	.9073
	Δ_0	4.3912	4.0916	3.9356	3.8370	3.9356	3.8376	3.7713	3.7234
.40	E^*	1.0825	1.0606	1.0480	1.0397	1.1036	1.0346	1.0715	1.0620
	E_0	.8950	.9210	.9365	.9469	.8838	.9014	.9143	.9242
	Δ_0	4.0976	3.8496	3.7186	3.6375	3.7186	3.6375	3.5825	3.5426
.45	E^*	1.0667	1.0491	1.0390	1.0323	1.0828	1.0679	1.0576	1.0530
	E_0	.9150	.9364	.9490	.9574	.9051	.9199	.9306	.9388
	Δ_0	3.8480	3.6441	3.5357	3.4623	3.5357	3.4683	3.4224	3.3892
.50	E^*	1.0533	1.0393	1.0312	1.0259	1.0655	1.0539	1.0458	1.0398
	E_0	.9322	.9494	.9596	.9673	.9237	.9358	.9446	.9512
	Δ_0	3.6366	3.4687	3.3787	3.3227	3.3787	3.3227	3.2845	3.2567

Table 11 Maximum and minimum guaranteed generalized efficiencies for $p=3$ (continued).

$\alpha \backslash (N_1, N_2)$		(4,12)	(6,12)	(8,12)	(10,12)	(4,16)	(6,16)	(8,16)	(10,16)
.05	E^*	1.4157	1.3243	1.2722	1.2328	1.4708	1.3827	1.3245	1.2807
	E_0	.6499	.6765	.6089	.7189	.6541	.6739	.6912	.7066
	Δ_0	9.0500	8.6113	8.2963	8.0791	8.2963	8.0591	7.8740	7.7257
.10	E^*	1.3224	1.2619	1.2198	1.1890	1.3556	1.2973	1.2559	1.2249
	E_0	.7126	.7368	.7569	.7741	.7145	.7327	.7465	.7624
	Δ_0	7.2184	6.9372	6.7329	6.5777	6.7329	6.5777	6.4558	6.3570
.15	E^*	1.2584	1.2128	1.1810	1.1577	1.2807	1.2387	1.2079	1.1843
	E_0	.7562	.7782	.7963	.8116	.7568	.7735	.7880	.8006
	Δ_0	6.2510	6.0429	5.8907	5.7745	5.8907	5.7745	5.6829	5.6088
.20	E^*	1.2108	1.1756	1.1506	1.1319	1.2265	1.1950	1.1712	1.1527
	E_0	.7908	.8108	.8272	.8408	.7907	.8060	.8191	.8305
	Δ_0	5.6036	5.4159	5.3256	5.2334	5.3256	5.2334	5.1699	5.1014
.25	E^*	1.1735	1.1459	1.1259	1.1115	1.1859	1.1606	1.1420	1.1273
	E_0	.8201	.8381	.8527	.8648	.8195	.8333	.8452	.8554
	Δ_0	5.1574	5.0042	4.9059	4.8404	4.9059	4.8354	4.7706	4.7120
.30	E^*	1.1435	1.1214	1.1053	1.0900	1.1526	1.1329	1.1180	1.1062
	E_0	.8454	.8616	.8746	.8853	.8446	.8571	.8676	.8767
	Δ_0	4.7076	4.6570	4.5751	4.5121	4.5751	4.5121	4.4621	4.4214
.35	E^*	1.1186	1.1016	1.0879	1.0779	1.1251	1.1099	1.0980	1.0885
	E_0	.8679	.8822	.8936	.9030	.8609	.8780	.8874	.8954
	Δ_0	4.4061	4.3730	4.3039	4.2506	4.3039	4.2506	4.2083	4.1736
.40	E^*	1.0978	1.0836	1.0731	1.0649	1.1027	1.0906	1.0811	1.0734
	E_0	.8879	.9005	.9104	.9185	.8868	.8966	.9048	.9118
	Δ_0	4.2128	4.1338	4.0750	4.0297	4.0750	4.0297	3.9936	3.9641
.45	E^*	1.0802	1.0688	1.0603	1.0537	1.0839	1.0743	1.0667	1.0605
	E_0	.9059	.9167	.9253	.9322	.9048	.9133	.9204	.9264
	Δ_0	3.9952	3.9279	3.8777	3.8382	3.8777	3.8389	3.8079	3.7827
.50	E^*	1.0652	1.0561	1.0493	1.0439	1.0680	1.0604	1.0543	1.0493
	E_0	.9220	.9312	.9384	.9443	.9210	.9272	.9343	.9393
	Δ_0	3.8049	3.7474	3.7045	3.6712	3.7045	3.6712	3.6447	3.6230

Table 11 Maximum and minimum guaranteed generalized efficiencies for $p=3$.

α	(N_1, N_2)	(4,4)	(6,4)	(8,4)	(10,4)	(4,8)	(6,8)	(8,8)	(10,8)
.05	E^*	1.2192	1.1563	1.1217	1.0997	1.3383	1.2566	1.2073	1.1741
	E_0	.6034	.6926	.7404	.7831	.6403	.6812	.7128	.7381
	Δ_0	18.4245	12.8437	10.7726	9.7022	10.7726	9.7022	9.0500	8.6113
.10	E^*	1.1869	1.1336	1.1042	1.0854	1.2722	1.2103	1.1717	1.1452
	E_0	.6869	.7619	.8060	.8357	.7078	.7441	.7718	.7937
	Δ_0	12.5595	9.5140	8.2389	7.6298	8.2389	7.6298	7.2184	6.9372
.15	E^*	1.1600	1.1148	1.0896	1.0736	1.2235	1.1750	1.1440	1.1225
	E_0	.7420	.8002	.8435	.8683	.7510	.7865	.8109	.8301
	Δ_0	9.9489	7.8948	7.0293	6.5527	7.0293	6.5527	6.2510	6.0429
.20	E^*	1.1372	1.0986	1.0772	1.0634	1.1954	1.1467	1.1215	1.1038
	E_0	.7839	.8393	.8711	.8929	.7993	.8193	.8410	.8579
	Δ_0	8.3876	6.8733	6.2143	5.8454	6.2143	5.8454	5.6096	5.4459
.25	E^*	1.1175	1.0847	1.0663	1.0543	1.1547	1.1234	1.1027	1.0880
	E_0	.8179	.8657	.8929	.9106	.8205	.8464	.8655	.8803
	Δ_0	7.3195	6.1475	5.6246	5.3282	5.6246	5.3282	5.1374	5.0042
.30	E^*	1.1003	1.0724	1.0508	1.0408	1.1292	1.1037	1.0867	1.0745
	E_0	.8464	.8875	.9107	.9257	.8466	.8695	.8863	.8991
	Δ_0	6.5294	5.5945	5.1690	4.9255	5.1690	4.9255	4.7676	4.6570
.35	E^*	1.0852	1.0616	1.0484	1.0399	1.1077	1.0870	1.0729	1.0628
	E_0	.8709	.9060	.9256	.9383	.8694	.8895	.9041	.9152
	Δ_0	5.9141	5.1536	4.6016	4.3954	4.6016	4.3954	4.2661	4.1730
.40	E^*	1.0719	1.0521	1.0409	1.0337	1.0895	1.0725	1.0610	1.0527
	E_0	.8921	.9219	.9384	.9490	.8896	.9071	.9197	.9292
	Δ_0	5.4109	4.7001	4.4059	4.3247	4.4959	4.3247	4.2128	4.1338
.45	E^*	1.0602	1.0436	1.0343	1.0283	1.0738	1.0601	1.0506	1.0438
	E_0	.9166	.9356	.9494	.9582	.9077	.9226	.9333	.9414
	Δ_0	5.0038	4.4829	4.2353	4.0904	4.2353	4.0904	3.9952	3.9279
.50	E^*	1.0498	1.0361	1.0284	1.0234	1.0603	1.0492	1.0416	1.0360
	E_0	.9268	.9475	.9588	.9661	.9237	.9363	.9453	.9520
	Δ_0	4.6529	4.2181	4.0091	3.8860	4.0091	3.8860	3.8049	3.7474

Table 12 Maximum and minimum guaranteed generalized efficiencies for $p=4$ (continued).

$\alpha \backslash (N_1, N_2)$		(4,12)	(6,12)	(8,12)	(10,12)	(4,16)	(6,16)	(8,16)	(10,16)
.05	E^*	1.3117	1.2468	1.2051	1.1756	1.3525	1.2872	1.2433	1.2115
	E_0	.6946	.7104	.7398	.7571	.7020	.7198	.7353	.7489
	Δ_0	10.8430	10.1451	9.6047	9.3087	9.0047	9.3087	9.0356	8.8196
.10	E^*	1.2496	1.2022	1.1702	1.1472	1.2750	1.2302	1.1981	1.1741
	E_0	.7475	.7698	.7850	.8034	.7522	.7685	.7825	.7948
	Δ_0	8.5755	8.1258	7.8081	7.5718	7.5031	7.5718	7.3891	7.2436
.15	E^*	1.2053	1.1680	1.1433	1.1247	1.2230	1.1893	1.1646	1.1458
	E_0	.7840	.8043	.8208	.8346	.7872	.8022	.8150	.8262
	Δ_0	7.3812	7.0454	6.8002	6.6271	6.5002	6.6271	6.4579	6.3707
.20	E^*	1.1710	1.1420	1.1215	1.1063	1.1837	1.1577	1.1382	1.1231
	E_0	.8130	.8315	.8464	.8588	.8152	.8290	.8407	.8509
	Δ_0	6.5890	6.3225	6.1515	5.9876	6.1015	5.9876	5.8753	5.7861
.25	E^*	1.1434	1.1201	1.1034	1.0902	1.1657	1.1322	1.1106	1.1011
	E_0	.8371	.8543	.8678	.8789	.8301	.8516	.8623	.8715
	Δ_0	6.0034	5.7857	5.6282	5.5094	5.6282	5.5094	5.4164	5.3418
.30	E^*	1.1205	1.1015	1.0879	1.0775	1.1470	1.1112	1.0985	1.0885
	E_0	.8540	.8742	.8862	.8961	.8472	.8714	.8811	.8903
	Δ_0	5.5453	5.3021	5.2290	5.1294	5.2290	5.1293	5.0508	4.9875
.35	E^*	1.1013	1.0859	1.0746	1.0659	1.1367	1.0934	1.0832	1.0749
	E_0	.8781	.8917	.9024	.9112	.8767	.8890	.8976	.9059
	Δ_0	5.1093	5.0137	4.9399	4.8832	4.9607	4.8152	4.7480	4.6938
.40	E^*	1.0849	1.0723	1.0639	1.0558	1.0899	1.0783	1.0699	1.0631
	E_0	.8954	.9074	.9168	.9245	.8957	.9048	.9125	.9190
	Δ_0	4.8521	4.7168	4.6219	4.5481	4.6219	4.5481	4.4902	4.4434
.45	E^*	1.0708	1.0605	1.0528	1.0469	1.0739	1.0652	1.0584	1.0529
	E_0	.9110	.9215	.9297	.9364	.9111	.9191	.9258	.9315
	Δ_0	4.5782	4.4534	4.3797	4.3159	4.3797	4.3159	4.2657	4.2251
.50	E^*	1.0585	1.0502	1.0439	1.0390	1.0609	1.0539	1.0484	1.0438
	E_0	.9251	.9342	.9413	.9470	.9251	.9321	.9379	.9427
	Δ_0	4.3372	4.2381	4.1656	4.1103	4.1656	4.1103	4.0667	4.0315

Table 12 Maximum and minimum guaranteed generalized efficiencies for $p=4$.

$\alpha \backslash (N_1, N_2)$		(4,4)	(6,4)	(8,4)	(10,4)	(4,8)	(6,8)	(8,8)	(10,8)
5	E^*	1.1056	1.1185	1.0925	1.0759	1.2512	1.1933	1.1567	1.1319
	E_0	.6051	.7115	.7685	.8051	.6779	.7181	.7460	.7716
	Δ	32.0761	18.0744	13.8337	11.9241	13.8007	11.9241	10.8130	10.1484
10	E^*	1.1453	1.1039	1.0710	1.0604	1.2111	1.1031	1.1331	1.1126
	E_0	.6942	.7741	.8209	.8506	.7302	.7717	.7978	.8181
	Δ	21.1035	12.8019	10.4336	9.2410	10.4336	9.2010	8.5755	8.1258
15	E^*	1.1270	1.0911	1.0710	1.0583	1.1777	1.1388	1.1141	1.0969
	E_0	.7372	.8111	.8541	.8754	.7460	.8077	.8309	.8498
	Δ	15.7778	10.4370	8.7359	7.8970	8.7359	7.8970	7.2912	7.0454
20	E^*	1.1110	1.0759	1.0623	1.0511	1.1504	1.1158	1.0980	1.0830
	E_0	.7727	.8446	.8787	.8900	.8072	.8557	.8804	.8922
	Δ	12.7000	8.9705	7.6459	6.9491	7.6459	6.9460	6.5490	6.3225
25	E^*	1.0978	1.0608	1.0545	1.0447	1.1211	1.1010	1.0840	1.0721
	E_0	.8109	.8689	.8982	.9100	.8334	.8589	.8773	.8913
	Δ	10.7927	7.8204	6.8003	6.2888	6.8003	6.3288	6.0444	5.7857
30	E^*	1.0852	1.0608	1.0474	1.0359	1.1084	1.0869	1.0724	1.0621
	E_0	.8389	.8891	.9143	.9268	.8507	.8788	.8951	.9074
	Δ	9.3747	7.0052	6.2546	5.8155	6.2546	5.8155	5.5453	5.3621
35	E^*	1.0739	1.0527	1.0411	1.0337	1.0925	1.0741	1.0619	1.0532
	E_0	.8631	.9064	.9279	.9411	.8701	.8963	.9100	.9213
	Δ	8.2900	6.4623	5.7061	5.3077	5.7061	5.3077	5.1093	5.0187
40	E^*	1.0630	1.0453	1.0354	1.0280	1.0786	1.0628	1.0526	1.0453
	E_0	.8844	.9214	.9386	.9508	.8910	.9117	.9242	.9335
	Δ	7.4429	5.9428	5.3591	5.0469	5.3591	5.0469	4.8521	4.7188
45	E^*	1.0543	1.0387	1.0302	1.0246	1.0655	1.0529	1.0444	1.0383
	E_0	.9031	.9344	.9498	.9592	.9161	.9255	.9363	.9442
	Δ	6.7443	5.5047	5.0117	4.7453	5.0117	4.7453	4.5782	4.4634
50	E^*	1.0459	1.0326	1.0254	1.0209	1.0545	1.0441	1.0371	1.0320
	E_0	.9197	.9459	.9587	.9664	.9247	.9379	.9470	.9537
	Δ	6.1588	5.1274	4.7090	4.4810	4.7090	4.4810	4.3372	4.2381

Table 13 Maximum and minimum guaranteed generalized efficiencies for $p=5$ (continued).

$\alpha \backslash (N_1, N_2)$		(4,12)	(6,12)	(8,12)	(10,12)	(4,16)	(6,16)	(8,16)	(10,16)
.05	E^*	1.2498	1.1981	1.1648	1.1414	1.2822	1.2302	1.1952	1.1699
	E_0	.7245	.7481	.7673	.7835	.7644	.7510	.7652	.7777
	Δ_0	12.8655	11.8041	11.0928	10.5820	11.0928	10.5820	10.1978	9.8983
.10	E^*	1.2016	1.1654	1.1392	1.1204	1.2257	1.1884	1.1620	1.1424
	E_0	.7706	.7919	.8091	.8234	.7778	.7929	.8058	.8171
	Δ_0	10.0717	9.3835	8.4438	8.5722	8.9138	8.5722	8.3128	8.1090
.15	E^*	1.1710	1.1491	1.1189	1.1084	1.1556	1.1573	1.1566	1.1209
	E_0	.8025	.8219	.8375	.8534	.8080	.8219	.8328	.8446
	Δ_0	8.6108	8.0856	7.7598	7.4794	7.7598	7.4793	7.2801	7.1272
.20	E^*	1.1443	1.1115	1.1021	1.0802	1.1333	1.1327	1.1102	1.1034
	E_0	.8279	.8456	.8598	.8714	.8722	.8470	.8559	.8632
	Δ_0	7.6448	7.2136	6.9491	6.7584	6.9491	6.7489	6.5770	6.4802
.25	E^*	1.1225	1.1023	1.0779	1.0771	1.1111	1.1120	1.0961	1.0886
	E_0	.8495	.8676	.8785	.8884	.8929	.8647	.8740	.8831
	Δ_0	6.9550	6.8951	6.8586	6.8556	6.8556	6.8435	6.8485	6.8421
.30	E^*	1.1012	1.0875	1.0755	1.0665	1.1131	1.0957	1.0847	1.0759
	E_0	.8684	.8831	.8946	.9020	.8712	.8819	.8921	.8986
	Δ_0	6.3741	6.2699	5.8899	5.7415	5.8899	5.7415	5.6209	5.5358
.35	E^*	1.0886	1.0748	1.0648	1.0571	1.1061	1.0813	1.0722	1.0649
	E_0	.8854	.8986	.9050	.9173	.8876	.8973	.9054	.9123
	Δ_0	5.9162	5.6737	5.5026	5.3550	5.5026	5.3750	5.2764	5.1977
.40	E^*	1.0750	1.0636	1.0553	1.0489	1.0784	1.0688	1.0613	1.0553
	E_0	.9008	.9126	.9218	.9292	.9025	.9112	.9185	.9246
	Δ_0	5.5295	5.3208	5.1728	5.0623	5.1728	5.0623	4.9708	4.9084
.45	E^*	1.0632	1.0538	1.0469	1.0415	1.0658	1.0579	1.0518	1.0468
	E_0	.9148	.9253	.9333	.9398	.9162	.9239	.9303	.9357
	Δ_0	5.1940	5.0143	4.8858	4.7896	4.8858	4.7896	4.7148	4.6551
.50	E^*	1.0528	1.0451	1.0394	1.0349	1.0548	1.0484	1.0433	1.0392
	E_0	.9277	.9368	.9438	.9494	.9287	.9355	.9411	.9457
	Δ_0	4.9000	4.7431	4.6312	4.5472	4.6312	4.5472	4.4819	4.4296

Table 13. Maximum and minimum guaranteed generalized efficiencies for $p=5$.

$\alpha \backslash (N_1, N_2)$		(4,4)	(6,4)	(8,4)	(10,4)	(4,8)	(6,8)	(8,8)	(10,8)
.05	E^*	1.1334	1.0357	1.0747	1.0613	1.2041	1.1556	1.1262	1.1063
	E_0	.0061	.7140	.7792	.8178	.6997	.7412	.7709	.7935
	Δ_0	38.3913	27.1561	18.0157	14.6130	18.0157	14.6130	12.8055	11.8044
.10	E^*	1.1195	1.0453	1.0664	1.0544	1.1731	1.1335	1.1020	1.0922
	E_0	.6554	.7742	.8278	.8591	.7525	.7889	.8147	.8343
	Δ_0	35.7571	18.3943	13.2338	11.1755	13.2338	11.1755	10.0717	9.3835
.15	E^*	1.1069	1.0759	1.0580	1.0484	1.1481	1.1154	1.0847	1.0804
	E_0	.7041	.8131	.8587	.8851	.7884	.8210	.8439	.8612
	Δ_0	31.0653	14.4312	10.9027	9.4246	10.9037	9.4246	8.0108	8.0956
.20	E^*	1.0954	1.0675	1.0524	1.0429	1.1273	1.0999	1.0824	1.0702
	E_0	.7474	.8125	.8517	.8643	.8100	.8460	.8665	.8819
	Δ_0	21.6034	12.9000	9.4239	8.2450	9.4239	8.2850	7.6448	7.2310
.25	E^*	1.0846	1.0708	1.0464	1.0350	1.1037	1.0586	1.0717	1.0612
	E_0	.7833	.8002	.8001	.8194	.8494	.8068	.8851	.8938
	Δ_0	19.3010	10.4156	8.4763	7.4562	8.3743	7.4562	6.9339	6.5951
.30	E^*	1.0730	1.0527	1.0409	1.0301	1.0914	1.0749	1.0622	1.0532
	E_0	.8142	.8501	.8173	.8320	.8911	.8848	.9011	.9132
	Δ_0	15.9256	9.2811	7.5703	6.8110	7.5703	6.8110	6.3741	6.0899
.35	E^*	1.0620	1.0463	1.0358	1.0293	1.0811	1.0646	1.0538	1.0431
	E_0	.8411	.9032	.9283	.9126	.8795	.9006	.9151	.9257
	Δ_0	13.4000	8.2900	6.9265	6.2865	6.9235	6.2865	5.9162	5.6737
.40	E^*	1.0577	1.0403	1.0312	1.0255	1.0493	1.0554	1.0463	1.0397
	E_0	.8648	.9181	.9396	.9517	.8860	.9147	.9274	.9367
	Δ_0	11.0062	7.5264	6.5867	5.8463	6.5867	5.8463	5.5295	5.3208
.45	E^*	1.0500	1.0348	1.0269	1.0220	1.0528	1.0472	1.0395	1.0339
	E_0	.8860	.9312	.9494	.9590	.9111	.9274	.9385	.9465
	Δ_0	10.2279	6.8863	5.9296	5.4679	5.9293	5.4679	5.1949	5.0143
.50	E^*	1.0429	1.0298	1.0230	1.0188	1.0495	1.0398	1.0334	1.0287
	E_0	.9048	.9428	.9550	.9606	.9248	.9389	.9484	.9552
	Δ_0	9.0647	6.3440	5.5321	5.1359	5.5321	5.1359	4.9000	4.7431

Finally, we consider the generalized-relative efficiency of SMLE relative to UMLE:

$$GRE_3(\hat{\mu}^{(1)S} : \hat{\mu}^{(1)}) = [1 - G_2(\Delta)]^{\frac{-1}{p}},$$

where

$$G_2(\Delta) = cmM \left[\{2E(\chi_{p+2,m}^{-2}(\Delta)) - c(m+2)E(\chi_{p+2,m}^{-2}(\Delta))\} - \{c(m+2) + 4\} \Delta E(\chi_{p+4}^{-2}(\Delta)) \right].$$

It is a decreasing function of Δ and at $\Delta = 0$, the value is

$$\left[1 - cmM \left\{ \frac{2(p-2) - c(m+2)}{2(p-2)} \right\} \right]^{-1} (> 1),$$

provided that $0 \leq c \leq \frac{2(p-2)}{(m+2)}$; and as $\Delta \rightarrow \infty$, the relative efficiency equals 1. Thus, $\hat{\mu}^{(1)S}$ uniformly dominates $\hat{\mu}^{(1)}$. Notice that $\hat{\mu}^{(1)S}$ does not dominate $\hat{\mu}^{(1)PT}$ under H_0 for a range of α , the level of significance for the null hypothesis. However, except for a range of Δ near the origin, $\hat{\mu}^{(1)S}$ dominates $\hat{\mu}^{(1)PT}$; also both estimators share a common property that as $\Delta \rightarrow \infty$ their GRE converge to the common limit 1. The GRE of $\hat{\mu}^{(1)S}$ always lies above this asymptotic value, while the GRE of $\hat{\mu}^{(1)PT}$ may drop, for some intermediate values of Δ , although the drop may not be significant depending on α .

Thus, from the point of robust-efficiency both $\hat{\mu}^{(1)PT}$ and $\hat{\mu}^{(1)S}$ may be advocated, leaning more towards $\hat{\mu}^{(1)S}$ since the size of Δ is generally unknown and unlikely to be small.

3.5. CONCLUDING REMARKS

We reach the same conclusions about the performance of the estimators when we use the risk criterion or the mean square error matrix measure. It should be remarked that for

$p \leq 2$, the only sensible choice left is to use preliminary test estimator with the decision rule provided for the choice of the size of the test. However, if Δ is unknown and $p \geq 3$, we have the table of maximum and minimum efficiencies at the intersection of $\hat{\mu}^{(1)PT}$ and $\hat{\mu}^{(1)S}$ to make the decision for an appropriate estimator. In general, the shrinkage maximum likelihood estimator is preferred to the preliminary test maximum likelihood estimator for $p \geq 3$.

CHAPTER 4

ESTIMATION STRATEGIES FOR THE COMPONENT MEAN-VECTORS IN A $2p$ -VARIATE NORMAL DISTRIBUTION

4.1. INTRODUCTION

In this chapter we discuss four estimation methods for the first component mean vector $\mu^{(1)}$ of a $2p$ -variate normal distribution when it is suspected that $\mu^{(1)} = \mu^{(2)}$, where $\mu^{(2)}$ is the second component mean vector. Exact bias, risks and generalized mean square error are derived for all of these estimators. Two measures of relative efficiency are presented: one based on risk using (Mahalanobis distance as a loss function), and the other based on the determinant of the mean square error matrix.

Four estimation strategies are developed in Section 4.2, bias and risk are derived in Section 4.3. Risk efficiency analysis and recommendations for selecting an estimator are discussed in Section 4.3.1. Derivation of the mean square matrices are given in Section 4.4 and generalized efficiency is discussed in Section 4.4.1. Section 4.5 summarizes the findings.

Suppose we are conducting studies on the p -fatality characteristics of AIDS victims. It is proper to collect data on the couples involved in partnership.

Let the data set of the p -factors be

$$\left\{ \begin{pmatrix} x_1^{(1)} \\ x_1^{(2)} \end{pmatrix}, \begin{pmatrix} x_2^{(1)} \\ x_2^{(2)} \end{pmatrix}, \dots, \begin{pmatrix} x_N^{(1)} \\ x_N^{(2)} \end{pmatrix} \right\}.$$

Suppose, this is a result of drawing a random sample of size N from a $2p$ -dimensional multivariate normal distribution with the mean vector μ and the covariance matrix Σ .

Let

$$\mu = \begin{pmatrix} \mu^{(1)} \\ \mu^{(2)} \end{pmatrix} \text{ and } \Sigma = \begin{pmatrix} \Sigma_{11} & \Sigma_{12} \\ \Sigma_{21} & \Sigma_{22} \end{pmatrix}$$

respectively.

Since the couple have similar basic characteristics, the p -fatality characteristics are related. It is our objective to estimate $\mu^{(1)}$, the first component mean vector when it is suspected that $\mu^{(1)} = \mu^{(2)}$ is plausible. For the estimation of $\mu^{(1)}$ we assume that

$$\Sigma_{11} = \Sigma_{22} \text{ and } \Sigma_{12} = \Sigma_{21}.$$

4.2. ESTIMATION STRATEGIES

We consider the following four estimators for $\mu^{(1)}$ when it is suspected that $\mu^{(1)} = \mu^{(2)}$ is plausible but not sure.

i) Unrestricted maximum likelihood estimator (UMLE):

$$\tilde{\mu}^{(i)} = \bar{x}^{(i)} = N_i^{-1} \sum_{j=1}^N x_{ij}, \quad i = 1, 2. \quad (4.2.1)$$

ii) Restricted maximum likelihood estimator:

$$\hat{\mu}^{(1)} = \frac{1}{2}(\tilde{\mu}^{(1)} + \tilde{\mu}^{(2)}), \quad (4.2.2)$$

iii) Preliminary test maximum likelihood estimator (PTMLE):

$$\hat{\mu}^{(1)PT} = \tilde{\mu}^{(1)} I(T_N^2 \geq T_{N,\alpha}^2) + \hat{\mu}^{(1)} I(T_N^2 < T_{N,\alpha}^2), \quad (4.2.3)$$

where $I_{(A)}$ is the indicator function of the set A and T_N^2 is the Hotellings T_N^2 -statistic defined by

$$T_N^2 = N(\bar{\mu}^{(2)} - \bar{\mu}^{(1)})' S^{-1} (\bar{\mu}^{(2)} - \bar{\mu}^{(1)}), \quad (4.2.4)$$

where

$$S = S_{22} - S_{21} - S_{12} + S_{11} \quad (4.2.5)$$

is an estimate of $\Sigma^* = 2(\Sigma_{22} - \Sigma_{21})$ and

$$(N-1)S_{ij} = \sum_{s=1}^N (x_s^{(i)} - \bar{x}^{(i)})(x_s^{(j)} - \bar{x}^{(j)})'. \quad (4.2.6)$$

$T_{N,\alpha}^2$ denotes the α -level critical values of the T_N^2 -statistics. $\bar{\mu}^{(1)PT}$ may be written as

$$\bar{\mu}^{(1)PT} = \bar{\mu}^{(1)} + \frac{1}{2}(\bar{\mu}^{(2)} - \bar{\mu}^{(1)})I(T_N^2 < T_{N,\alpha}^2). \quad (4.2.7)$$

iv) Shrinkage maximum likelihood estimator (SMLE):

$$\hat{\mu}^{(1)S} = \bar{\mu}^{(1)} + \frac{c}{2}(\bar{\mu}^{(2)} - \bar{\mu}^{(1)})(N-1)T_N^{-2}, \quad (4.2.8)$$

where c is the shrinkage factor such that

$$0 < c < \frac{2(p-2)}{m+2}, \quad m = N - p.$$

We notice that PTMLE and SMLE of $\bar{\mu}^{(1)}$ are similar looking since both depend on the statistic T_N^2 and adjust themselves according to the departure from the null hypothesis $H_0: \mu^{(1)} = \mu^{(2)}$. For large values of T_N^2 , both tend to $\bar{\mu}^{(1)}$, while for small values of T_N^2 , they behave in a different way. These properties along with others will be considered by studying the risk function under a quadratic loss as well as mean square error matrices.

4.3. BIAS AND RISKS

Theorem 4.3.1 and 4.3.2 provides expressions for the bias and the risk for four estimators.

THEOREM 4.3.1. Let B_1, B_2, B_3 and B_4 denote the bias vectors of the four estimators.

Then we have:

$$B_1 = E(\hat{\mu}^{(1)} - \mu^{(1)}) = 0 \quad (4.3.1)$$

$$B_2 = E(\hat{\mu}^{(1)} - \mu^{(1)}) = \frac{1}{2}\delta, \quad \delta = \mu^{(2)} - \mu^{(1)} \quad (4.3.2)$$

$$B_3 = E(\hat{\mu}^{(1)PT} - \mu^{(1)}) = \frac{1}{2}\delta H_{p+2,m}(T_{N,\alpha}^{*2}; \Delta), \quad (4.3.3)$$

where $m = N - p$ and $\Delta = N\delta'\Sigma^{*-1}\delta$ and $H_{\nu_1, \nu_2}(T_{N,\alpha}^{*2}; \Delta)$ is the cumulative distribution function of a non-central F -distribution with (ν_1, ν_2) degrees of freedom and non-centrality parameter Δ . Also,

$$T_{N,\alpha}^{*2} = \frac{p}{m} F_{p,m}(\alpha),$$

where $F_{\nu_1, \nu_2}(\alpha)$ is the upper α -level critical value of a central F -distribution with (ν_1, ν_2) degrees of freedom.

$$B_4 = E(\hat{\mu}^{(1)S} - \mu^{(1)}) = \frac{c}{2}\delta E(\chi_{p+2}^{-2}(\Delta)), \quad (4.3.4)$$

where $\chi_{\nu}^2(\Delta)$ is a non-central chi-square random variable with ν degrees of freedom and non-centrality parameter Δ and c is shrinkage constant.

PROOF. Analogous to Theorem 3.3.1 in Chapter 3.

To study the properties of the four estimators, we provide risk comparison in the light of the quadratic loss function is given by

$$L(\mu^{*(1)}; \mu^{(1)}) = N(\mu^{*(1)} - \mu^{(1)})' \Sigma_{11}^{-1} (\mu^{*(1)} - \mu^{(1)}), \quad (4.3.5)$$

where $\mu^{*(1)}$ is an estimator of the $\mu^{(1)}$. Then the risk of estimator $\mu^{*(1)}$ is

$$R(\mu^{*(1)}; \Sigma_{11}) = E \left\{ N(\mu^{*(1)} - \mu^{(1)})' \Sigma_{11}^{-1} (\mu^{*(1)} - \mu^{(1)}) \right\}, \quad (4.3.6)$$

where E denotes the usual expectation operator. Then we have the following theorem.

THEOREM 4.3.2. Risks of UMLE, RMLE, PTMLE and SMLE under (4.3.5) are given by (4.3.7) - (4.3.10) respectively:

$$R_1(\hat{\mu}^{(1)}; \Sigma_{11}^{-1}) = p \quad (4.3.7)$$

$$R_2(\hat{\mu}^{(1)}; \Sigma_{11}^{-1}) = \frac{1}{2}(p + \text{tr } \Sigma^x) + \frac{N}{4} \delta' \Sigma_{11}^{-1} \delta, \quad (4.3.8)$$

where $\text{tr } \Sigma^x$ is trace of matrix $\Sigma^x = \Sigma_{11}^{-1} \Sigma_{12}$.

$$\begin{aligned} R_3(\hat{\mu}^{(1)PT}; \Sigma_{11}^{-1}) &= p + \frac{\Delta^*}{4} \{ 2H_{p+2,m}(T_{N,\alpha}^{*2}; \Delta) - H_{p+4,m}(T_{N,\alpha}^{*2}; \Delta) \} \\ &\quad - \frac{1}{2}(p - \text{tr } \Sigma^x) H_{p+2,m}(T_{N,\alpha}^{*2}; \Delta), \end{aligned} \quad (4.3.9)$$

where $\Delta^* = N \delta' \Sigma_{11}^{-1} \delta$.

$$\begin{aligned} R_4(\hat{\mu}^{(1)S}; \Sigma_{11}^{-1}) &= p - \frac{cm}{4} (\text{tr } \Sigma^0) [\{ 2E(\chi_{p+2}^{-2}(\Delta)) - c(m+2)E(\chi_{p+2}^{-4}(\Delta)) \} \\ &\quad - \Delta^* \{ c(m+2) + 4 \} E(\chi_{p+4}^{-4}(\Delta))], \end{aligned} \quad (4.3.10)$$

where $\Sigma^0 = \Sigma_{11}^{-1} \Sigma^*$.

PROOF. We only prove $R_3(\hat{\mu}^{(1)PT}; \Sigma_{11}^{-1})$ and $R_4(\hat{\mu}^{(1)S}; \Sigma_{11}^{-1})$.

$$\begin{aligned}
 R_3(\hat{\mu}^{(1)PT}; \Sigma_{11}^{-1}) &= NE \left\{ (\hat{\mu}^{(1)PT} - \mu^{(1)})' \Sigma_{11}^{-1} (\hat{\mu}^{(1)PT} - \mu^{(1)}) \right\} \\
 &= NE \left\{ (\tilde{\mu}^{(1)} - \mu^{(1)})' \Sigma_{11}^{-1} (\tilde{\mu}^{(1)} - \mu^{(1)}) \right\} + \frac{N}{4} E \left\{ (\hat{\mu}^{(2)} - \tilde{\mu}^{(1)})' \Sigma_{11}^{-1} (\hat{\mu}^{(2)} - \tilde{\mu}^{(1)}) \right\} \\
 &\quad I(T^2 < T_{N,\alpha}^2) + NE \left\{ (\tilde{\mu}^{(1)} - \mu^{(1)})' \Sigma_{11}^{-1} (\tilde{\mu}^{(2)} - \tilde{\mu}^{(1)}) I(T^2 < T_{N,\alpha}^2) \right\} \\
 &= p - \frac{N}{4} E \left\{ (\tilde{\mu}^{(2)} - \tilde{\mu}^{(1)})' \Sigma_{11}^{-1} (\tilde{\mu}^{(2)} - \mu^{(1)}) I(T^2 < T_{N,\alpha}^2) \right\} \\
 &\quad + \frac{N}{2} \delta' \Sigma_{11}^{-1} E \left\{ (\tilde{\mu}^{(2)} - \tilde{\mu}^{(1)}) I(T^2 < T_{N,\alpha}^2) \right\} \\
 &= p - \frac{N}{4} E (\tilde{\mu}^{(2)} - \tilde{\mu}^{(1)})' \Sigma_{11}^{-1} (\tilde{\mu}^{(2)} - \tilde{\mu}^{(1)}) I(T^2 < T_{N,\alpha}^2) \\
 &\quad + \frac{\Delta^*}{2} H_{p+2,m}(T_{N,\alpha}^{*2}; \Delta). \tag{4.3.11}
 \end{aligned}$$

Consider the second term of (4.3.11)

$$NE \left\{ (\tilde{\mu}^{(2)} - \tilde{\mu}^{(1)})' \Sigma_{11}^{-1} (\tilde{\mu}^{(2)} - \tilde{\mu}^{(1)}) I(T^2 < T_{N,\alpha}^2) \right\} = E \left\{ (y' \Sigma_{11}^{-1} y) I(y' S^{-1} y < T_{N,\alpha}^2) \right\},$$

where $y = \sqrt{N}(\tilde{\mu}^{(2)} - \tilde{\mu}^{(1)})$.

$$= E[(y' D^{-1} \Sigma_{11}^{-1} D^{-1} y^*) I(y' y^* / \chi_m^2 < T_{N,\alpha}^{*2})], \tag{4.3.12}$$

where

$$y^* = Dy$$

with $D \Sigma^* D' = I_p$. Thus (4.3.12) equals

$$\text{tr} \Sigma^* H_{p+2,m}(T_{N,\alpha}^{*2}; \Delta) + \Delta^* H_{p+4,m}(T_{N,\alpha}^{*2}; \Delta),$$

using Theorem A.2 from Appendix A. Hence, we have

$$R_3(\hat{\mu}^{(1)PT}, \Sigma_{11}^{-1}) = p + \frac{\Delta^*}{4} \{2H_{p+2,m}(T_{N,\alpha}^{*2}; \Delta) - H_{p+4,m}(T_{N,\alpha}^{*2}; \Delta)\} \\ - \frac{1}{2}(p - \text{tr } \Sigma^x)H_{p+2,m}(T_{N,\alpha}^{*2}; \Delta).$$

Similarly,

$$R_4(\hat{\mu}^{(1)S}, \Sigma_{11}^{-1}) = NE \{(\hat{\mu}^{(1)S} - \mu^{(1)})' \Sigma_{11}^{-1} (\hat{\mu}^{(1)S} - \mu^{(1)})\} \\ = NE \{(\hat{\mu}^{(1)} - \mu^{(1)})' \Sigma_{11}^{-1} (\hat{\mu}^{(1)} - \mu^{(1)})\} \\ + \frac{c^2}{4} NE \{(\hat{\mu}^{(2)} - \hat{\mu}^{(1)})' \Sigma_{11}^{-1} (\hat{\mu}^{(2)} - \hat{\mu}^{(1)}) (N-1) T^{-4}\} \\ + c NE \{(\hat{\mu}^{(1)} - \mu^{(1)})' \Sigma_{11}^{-1} (\hat{\mu}^{(2)} - \mu^{(1)}) (N-1) T^{-2}\}.$$

Following usual transformations, we have

$$R_4(\hat{\mu}^{(1)S}, \Sigma_{11}^{-1}) = p + \frac{c^2}{4} E \{ (y \Sigma_{11}^{-1} y) (N-1)^2 (y' S^{-1} y)^{-2} \} \\ + \frac{c}{2} E \{ \sqrt{N} (\delta' \Sigma_{11}^{-1} \delta) (y' S^{-1} y)^{-1} (N-1) \} \\ - \frac{c}{2} m E \{ N (y' \Sigma_{11}^{-1} y) (y' S^{-1} y)^{-1} \} \\ = p + \frac{c^2}{4} E \{ (y^* D^{-1} \Sigma_{11}^{-1} D^{-1} y^*) (y^* y^* / \chi_m^2)^{-2} \} \\ + \frac{c}{2} m \Delta^* E(\chi_{p+2}^{-2}(\Delta)) - \frac{c}{2} m \{ \text{tr } \Sigma^0 E(\chi_{p+2}^{-2}(\Delta)) + \Delta^* E(\chi_{p+4}^{-2}(\Delta)) \} \\ = p + \frac{c^2}{4} m(m+2) \{ \text{tr } \Sigma^0 E(\chi_{p+2}^{-4}(\Delta)) + \Delta^* E(\chi_{p+4}^{-4}(\Delta)) \} \\ + \frac{c}{2} m \Delta^* E(\chi_{p+2}^{-2}(\Delta)) - \frac{c}{2} m \{ \text{tr } \Sigma^0 E(\chi_{p+2}^{-2}(\Delta)) + \Delta^* E(\chi_{p+4}^{-2}(\Delta)) \} \\ = p - \frac{cm}{4} [\text{tr } \Sigma^0 \{ 2E(\chi_{p+2}^{-2}(\Delta)) - c(m+2)E(\chi_{p+2}^{-4}(\Delta)) \} \\ - \Delta^* \{ c(m+2) + 4 \} E(\chi_{p+4}^{-4}(\Delta))],$$

where we use the relation

$$E(\chi_{p+2}^{-2}(\Delta)) - E(\chi_{p+4}^{-4}(\Delta)) = 2E(\chi_{p+4}^{-4}(\Delta)),$$

establishing (4.3.10).

4.3.1. RISK EFFICIENCY ANALYSIS AND RECOMMENDATIONS

In this section we consider the efficiency analysis of the four estimators of $\mu^{(1)}$ and discuss a max-min rule for the optimum choice of the level of significance for the preliminary test of the null hypothesis $H_0 : \mu^{(1)} = \mu^{(2)}$. First, we present the following dominance theorem under the null hypothesis.

THEOREM 4.3.3. Under H_0 , estimators are ordered according to the magnitude of the risks as follows:

$$\hat{\mu}^{(1)} > \hat{\mu}^{(1)PT} > \hat{\mu}^{(1)S} > \bar{\mu}^{(1)},$$

for a range of α .

PROOF.

$$R_1(\bar{\mu}^{(1)}; \Sigma_{11}^{-1}) - R_2(\hat{\mu}^{(1)}; \Sigma_{11}^{-1}) = \frac{1}{2}(p - \text{tr } \Sigma^x) > 0$$

$$R_1(\bar{\mu}^{(1)}; \Sigma_{11}^{-1}) - R_3(\hat{\mu}^{(1)PT}; \Sigma_{11}^{-1}) = \frac{1}{2}(p - \text{tr } \Sigma^x) H_{p+2,m}(T_{N,\alpha}^{*2}; 0) > 0$$

$$R_1(\bar{\mu}^{(1)}; \Sigma_{11}^{-1}) - R_4(\hat{\mu}^{(1)S}; \Sigma_{11}^{-1}) = \frac{c}{4}m(p - \text{tr } \Sigma^x) \left(\frac{2}{p} - \frac{c(m+2)}{p(p-2)} \right) > 0,$$

for $0 < c < \frac{2(p-2)}{(m+2)}$.

$$R_3(\hat{\mu}^{(1)PT}; \Sigma_{11}^{-1}) - R_2(\hat{\mu}^{(1)}; \Sigma_{11}^{-1}) = \frac{1}{2} \text{tr } \Sigma^x \{1 - H_{p+2,m}(T_{N,\alpha}^{*2}; 0)\} > 0,$$

$$R_4(\hat{\mu}^{(1)S}; \Sigma_{11}^{-1}) - R_3(\hat{\mu}^{(1)PT}; \Sigma_{11}^{-1}) = \frac{1}{2}(p - \text{tr } \Sigma^x) \{H_{p+2,m}(T_{N,\alpha}^{*2}; 0)\} \\ - cm \left\{ \frac{2}{p} - \frac{c(m+2)}{p(p-2)} \right\} > 0,$$

for some $\alpha > \alpha_0$ (generally $\alpha_0 > 0.5$). The last relation depends on α .

To make the risk analysis more manageable we consider the relative efficiencies of other estimators compared to $\tilde{\mu}^{(1)}$. Consider the relative efficiency of $\hat{\mu}^{(1)}$ to $\tilde{\mu}^{(1)}$. This is given by

$$\text{R.E.} (\hat{\mu}^{(1)} : \tilde{\mu}^{(1)}) = \left(\frac{1}{2} + \frac{\text{tr } \Sigma^x}{2p} + \frac{\Delta^*}{4p} \right)^{-1}.$$

The R.E. $(\hat{\mu}^{(1)} : \tilde{\mu}^{(1)})$ assumes a constant value on the ellipsoid $\Delta^* = k$ and equals 1 whenever $\Delta^* = 2(p - \text{tr } \Sigma^x)$. Also, at $\delta = 0$, i.e., $\Delta = 0$,

$$\text{R.E.} (\hat{\mu}^{(1)} : \tilde{\mu}^{(1)}) = \frac{1}{2} + \frac{1}{2p} \text{tr } \Sigma^x$$

and as δ moves away from 0, $\text{R.E.} (\hat{\mu}^{(1)} : \tilde{\mu}^{(1)}) \rightarrow 0$. Further,

$$\text{R.E.} (\hat{\mu}^{(1)} : \tilde{\mu}^{(1)}) \geq 1 \text{ according as } \Delta^* \geq 2(p - \text{tr } \Sigma^0).$$

Next, we consider the efficiency of $\hat{\mu}^{(1)PT}$ relative to $\tilde{\mu}^{(1)}$, which is given by

$$\text{R.E.} (\hat{\mu}^{(1)PT} : \tilde{\mu}^{(1)}) = \{1 + g(\Delta)\}^{-1},$$

where

$$g(\Delta) = \frac{\Delta^*}{4p} \{2H_{p+2,m}(T_{N,\alpha}^{*2}; \Delta) - H_{p+4,m}(T_{N,\alpha}^{*2}; \Delta)\} - \frac{1}{2p}(p - \text{tr } \Sigma^0)H_{p+2,m}(T_{N,\alpha}^{*2}; \Delta).$$

Also, we note that

$$g(\Delta) = \frac{\Delta^*}{4p} \{2H_{p+2,m}(T_{N,\alpha}^{*2}; \Delta) - H_{p+4,m}(T_{N,\alpha}^{*2})\} - \left(\frac{\text{tr } \Sigma^0}{4p} \right) H_{p+2,m}(T_{N,\alpha}^{*2}; \Delta).$$

Thus, it is easy to see that $\text{R.E.}_L \leq \text{R.E.} (\hat{\mu}^{(1)PT} : \tilde{\mu}^{(1)}) \leq \text{R.E.}_U$, where

$$\begin{aligned} \text{R.E.}_L &= \left[1 + \frac{\Delta}{4p} Ch_{\max}(\Sigma^0) \{2H_{p+2,m}(T_{N,\alpha}^{*2}; \Delta) - H_{p+4,m}(T_{N,\alpha}^{*2}; \Delta)\} \right. \\ &\quad \left. - \left(\frac{\text{tr } \Sigma^0}{4p} \right) H_{p+2,m}(T_{N,\alpha}^{*2}) \right]^{-1} \end{aligned}$$

and

$$\text{R.E. } U = \left[1 + \frac{\Delta}{4p} Ch_{\min}(\Sigma^0) \{ 2H_{p+2,m}(T_{N,\alpha}^{*2}; \Delta) - H_{p+4,m}(T_{N,\alpha}^{*2}; \Delta) \} - \left(\frac{\text{tr } \Sigma^0}{4p} \right) H_{p+2,m}(T_{N,\alpha}^{*2}) \right]^{-1},$$

where $Ch_{\max}(\Sigma^0)$ and $Ch_{\min}(\Sigma^0)$ represents the maximum and minimum characteristic root of the matrix Σ^0 , respectively. The relative efficiency bounds are equal with value

$$E^* = \left\{ 1 - \left(\frac{\text{tr } \Sigma^0}{4p} \right) H_{p+2,m}(T_{N,\alpha}^{*2}; 0) \right\}^{-1} \quad \text{at } \Delta = 0. \quad (4.3.13)$$

Further, both lower bound (R.E._L) and upper bound (R.E._U) decrease from E^* to a minimum L_0 and U_0 at $\Delta = \Delta_L$ and $\Delta = \Delta_U$ respectively, then increases and converges towards 1 as $\Delta \rightarrow \infty$. Thus, $\lim_{\Delta \rightarrow \infty} \text{R.E.}(\hat{\mu}^{(1)PT} : \tilde{\mu}^{(1)}) = 1$. Now, consider any of the bounds, R.E._L. It is a function $\text{R.E.}_L(\alpha, \Delta)$ of α and Δ , say $E_L(\alpha, \Delta)$ for fixed p and N . When $\Delta = 0$ and α varies, we have $\max_{\alpha} E_L(\alpha, 0) = E_L(0, 0) = \left(1 - \frac{\text{tr } \Sigma^0}{4p} \right)^{-1}$. The value of $E_L(\alpha, 0)$ decreases as α increases. On the other hand, let $\alpha = 0$ and Δ vary; then the curve $E_L(0, \Delta)$ and $E_L(1, \Delta) = 1$ intersect at $\Delta = \frac{\text{tr } \Sigma^0}{Ch_{\max}(\Sigma^0)}$ and the value of Δ decreases at the intersection of $E_L(\alpha, \Delta)$ and $E_L(1, \Delta) = 1$, as α increases. Therefore, for two different α , $E_L(\alpha_1, \Delta)$ and $E_L(\alpha_2, \Delta)$ can never intersect above the line $E_L(1, \Delta) = 1$.

In order to choose an estimator with minimum guaranteed relative efficiency we follow the rule: If it is known that $0 \leq \Delta \leq \left\{ \frac{\text{tr } \Sigma^0}{Ch_{\max}(\Sigma^0)} \right\}^{-1}$, then $\hat{\mu}^{(1)}$ is chosen, since E_L is the largest in this interval for all Δ . However, if Δ is unknown (which is generally the case) and we are willing to accept an estimator with relative efficiency at least E_0 , then we solve the equation $\max_{\alpha} \min_{\Delta} E_L(\alpha, \Delta) = E_{L0}$ is the minimum guaranteed efficiency of the preliminary test estimator with $\alpha = \alpha_L^*$ as the optimal choice for α for the preliminary test of $H_0 : \mu^{(1)} = \mu^{(2)}$. Similarly, if we consider the upper relative efficiency as a

function of (α, Δ) we have $E_U(\alpha, \Delta)$ to yield $\alpha = \alpha_U^*$ for the level of significance solving $\max_{\alpha} \min_{\Delta} E_U(\alpha, \Delta) = E_{U_0}$. Tables of R.E. $(\hat{\mu}^{(1)PT} : \bar{\mu}^{(1)})$ can be prepared for choices of α, p and N showing the minimum guaranteed efficiencies at each α (for fixed p and N).

Finally, we consider the comparison of $\hat{\mu}^{(1)S}$ and $\bar{\mu}^{(1)}$. The relative efficiency is given by:

$$\text{R.E. } (\hat{\mu}^{(1)S} : \bar{\mu}^{(1)}) = \{1 - g_*(\Delta)\}^{-1},$$

where

$$g_*(\Delta) = \frac{cm \operatorname{tr} \Sigma^0}{4p} \left[\{2E(\chi_{p+2}^{-2}(\Delta)) - c(m+2)E(\chi_{p+2}^{-4}(\Delta))\} - \Delta^* (\operatorname{tr} \Sigma^0)^{-1} \right. \\ \left. \{c(m+2) + 4\}E(\chi_{p+4}^{-4}(\Delta)) \right].$$

Now,

$$\text{R.E. } L^* \leq \text{R.E. } (\hat{\mu}^{(1)S} : \bar{\mu}^{(1)}) \leq \text{R.E. } U^*,$$

where

$$\text{R.E. } L^* = \left\{ 1 - \frac{cm}{4p} \operatorname{tr} \Sigma^0 \left[\{2E(\chi_{p+2}^{-2}(\Delta)) - c(m+2)E(\chi_{p+2}^{-4}(\Delta))\} \right. \right. \\ \left. \left. - \Delta (\operatorname{tr} \Sigma^0)^{-1} \{Ch_{\max}(\Sigma^0)\{c(m+2)\}E(\chi_{p+4}^{-4}(\Delta))\} \right] \right\}.$$

and

$$\text{R.E. } U^* = \left\{ 1 - \frac{cm}{4p} \operatorname{tr} \Sigma^0 \left[\{2E(\chi_{p+2}^{-2}(\Delta)) - c(m+2)E(\chi_{p+2}^{-4}(\Delta))\} \right. \right. \\ \left. \left. - \Delta (\operatorname{tr} \Sigma^0)^{-1} \{Ch_{\min}(\Sigma^0)\{c(m+2)\}E(\chi_{p+4}^{-4}(\Delta))\} \right] \right\}.$$

For, $\Delta = 0$, both upper and lower bounds are equal with value

$$\left[1 - \frac{cm}{4p} \operatorname{tr} \Sigma^0 \left\{ \frac{2}{p} - \frac{c(m+2)}{p(p-2)} \right\} \right]^{-1} > 1,$$

provided $0 < c < \frac{2(p-2)}{m+2}$. Also, as $\Delta \rightarrow \infty$, both bounds converge to 1. Finally, both bounds are monotonically decreasing functions of Δ with a maximum

$$\left[1 - \frac{cm \operatorname{tr} \Sigma^0}{4p^2(p-2)} \{2(p-2) - c(m+2)\} \right]^{-1}$$

and they decrease to a minimum of 1 as $\Delta \rightarrow \infty$. Thus, for large Δ , the performance of $\hat{\mu}^{(1)PT}$ and $\hat{\mu}^{(1)S}$ are similar. However, performances differ when Δ is small.

Consider the intersection of R.E.L and R.E.L⁻ at $\Delta = \Delta_{L,\alpha}$ for fixed α . (For large α there may not be any intersection, where $\hat{\mu}^{(1)PT}$ is inefficient than $\hat{\mu}^{(1)S}$.) For $\Delta \in (0, \Delta_{L,\alpha}]$, $\hat{\mu}^{(1)PT}$ is more efficient, while for $\Delta \in (\Delta_{L,\alpha}, \infty)$, $\hat{\mu}^{(1)S}$ is more efficient. Thus, neither of $\hat{\mu}^{(1)PT}$ nor $\hat{\mu}^{(1)S}$ claim victory over each other. A similar situation occurs with the upper bound. Thus, if $p \geq 3$ and Δ is unknown, $\hat{\mu}^{(1)S}$ is recommended for estimating $\mu^{(1)}$ while for $p \leq 2$, $\hat{\mu}^{(1)PT}$ may be a sensible choice.

4.4. MEAN SQUARE ERROR MATRICES

Let $\Gamma_1(\hat{\mu}^{(1)})$, $\Gamma_2(\hat{\mu}^{(1)})$, $\Gamma_3(\hat{\mu}^{(1)PT})$ and $\Gamma_4(\hat{\mu}^{(1)S})$ be mean square error matrices of $\hat{\mu}^{(1)}$, $\hat{\mu}^{(1)}$, $\hat{\mu}^{(1)PT}$ and $\hat{\mu}^{(1)S}$ respectively. Thus, we have the following theorem.

THEOREM 4.4.1. The mean square matrices are given by:

$$\Gamma_1(\hat{\mu}^{(1)}) = \frac{1}{N} \Sigma_{11} \quad (4.4.1)$$

$$\Gamma_2(\hat{\mu}^{(1)}) = \frac{1}{4N} \Sigma_{\cdot} + \frac{1}{4} \delta \delta', \Sigma_{\cdot} = 2(\Sigma_{11} + \Sigma_{12}) \quad (4.4.2)$$

$$\begin{aligned} \Gamma_3(\hat{\mu}^{(1)PT}) = & \frac{1}{N} \Sigma_{11} + \frac{1}{4N} \{2H_{p+2,m}(T_{N,\alpha}^{*2}; \Delta) - H_{p+4,m}(T_{N,\alpha}^{*2}; \Delta)\} \delta \delta' \\ & - \frac{1}{4N} H_{p+2,m}(T_{N,\alpha}^{*2}; \Delta) \Sigma_{\cdot} \end{aligned} \quad (4.4.3)$$

$$\begin{aligned} \Gamma_4(\hat{\mu}^{(1)S}) = & \frac{1}{N} \Sigma_{11} - \frac{cm}{4} \left[\frac{1}{N} \Sigma_{\cdot} \{2E(\chi_{p+2}^{-2}(\Delta)) - c(m+2)E(\chi_{p+2}^{-4}(\Delta))\} \right. \\ & \left. - \delta \delta' \{c(m+2) + 4\} E(\chi_{p+4}^{-4}(\Delta)) \right] \end{aligned} \quad (4.4.4)$$

PROOF. Follows in same way as in Section 3.4.

4.4.1. GENERALIZED EFFICIENCY ANALYSIS

In this section, we consider the mean square matrix analysis of the estimators and show that the conclusions are the same as in the previous section. First, we present the generalized relative efficiency of $\hat{\mu}^{(1)}$, $\hat{\mu}^{(1)PT}$ and $\hat{\mu}^{(1)S}$ with respect to $\tilde{\mu}^{(1)}$. They are

$$\begin{aligned} \text{GRE}(\hat{\mu}^{(1)} : \tilde{\mu}^{(1)}) &= |\Gamma_1(\tilde{\mu}^{(1)})|^{\frac{1}{p}} / |\Gamma_2(\hat{\mu}^{(1)})|^{\frac{1}{p}} \\ &= \left\{ \frac{1}{2}(p + \text{tr } \Sigma^x) + \frac{\Delta^*}{4} \right\}^{-\frac{1}{p}} \\ \text{GRE}(\hat{\mu}^{(1)PT} : \tilde{\mu}^{(1)}) &= |\Gamma_1(\tilde{\mu}^{(1)})|^{\frac{1}{p}} / |\Gamma_3(\hat{\mu}^{(1)PT})|^{\frac{1}{p}} \\ &= \{1 + G(\Delta)\}^{-\frac{1}{p}}, \end{aligned}$$

where

$$G(\Delta) = \frac{\Delta^*}{4} \{2H_{p+2,m}(T_{N,\alpha}^{*2}; \Delta) - H_{p+4,m}(T_{N,\alpha}^{*2}; \Delta)\} - \frac{\text{tr } \Sigma^0}{4} H_{p+2,m}(T_{N,\alpha}^{*2}; \Delta).$$

Finally,

$$\begin{aligned} \text{GRE}(\hat{\mu}^{(1)S} : \tilde{\mu}^{(1)}) &= |\Gamma_1(\tilde{\mu}^{(1)})|^{\frac{1}{p}} / |\Gamma_4(\hat{\mu}^{(1)S})|^{\frac{1}{p}} \\ &= \{1 - G_*(\Delta)\}^{-\frac{1}{p}}, \end{aligned}$$

where

$$\begin{aligned} G_*(\Delta) &= \frac{cm}{4} \text{tr } \Sigma^0 [\{2E(\chi_{p+2}^{-2}(\Delta)) - c(m+2)E(\chi_{p+2}^{-4}(\Delta)) \\ &\quad - \Delta^* (\text{tr } \Sigma^0)^{-1} \{c(m+2) + 4\} E(\chi_{p+4}^{-4}(\Delta))\}]. \end{aligned}$$

Now, note that the analysis of $\text{GRE}(\hat{\mu}^{(1)} : \tilde{\mu}^{(1)})$ is the same as in the risk efficiency analysis.

Next, we observe that

$$\text{GRE}_L^{(PT)} \leq \text{GRE}(\hat{\mu}^{(1)PT} : \tilde{\mu}^{(1)}) \leq \text{GRE}_U^{(PT)},$$

where

$$\begin{aligned} \text{GRE}_L^{(PT)} &= \left[1 + \frac{\Delta}{4} Ch_{\max}(\Sigma^0) \{2H_{p+2,m}(T_{N,\alpha}^{*2}; \Delta) - H_{p+4,m}(T_{N,\alpha}^{*2}; \Delta)\} \right. \\ &\quad \left. - \frac{1}{4} \text{tr } \Sigma^0 H_{p+2,m}(T_{N,\alpha}^{*2}; \Delta) \right]^{-\frac{1}{p}} \end{aligned}$$

and

$$\text{GRE}_U^{(PT)} = \left[1 + \frac{\Delta}{4} Ch_{\min}(\Sigma^0) \{2H_{p+2,m}(T_{N,\alpha}^{*2}; \Delta) - H_{p+4,m}(T_{N,\alpha}^{*2}; \Delta)\} - \frac{1}{4} \text{tr } \Sigma^0 H_{p+2,m}(T_{N,\alpha}^{*2}; \Delta) \right]^{-\frac{1}{p}}$$

Thus, the upper and lower $\text{GRE}(\hat{\mu}^{(1)PT} : \tilde{\mu}^{(1)})$ analysis leads to the same conclusion as in the risk efficiency analysis. Similarly,

$$\text{GRE}_L^{(S)} \leq \text{GRE}(\hat{\mu}^{(1)S} : \tilde{\mu}^{(1)}) \leq \text{GRE}_U^{(S)},$$

where

$$\text{GRE}_L^{(S)} = \left(1 - \frac{cm}{4} (\text{tr } \Sigma^0) [\{2E(\chi_{p+2}^{-2}(\Delta)) - c(m+2)E(\chi_{p+2}^{-4}(\Delta))\} - \Delta Ch_{\max}(\Sigma^0)(\text{tr } \Sigma^0)^{-1} \{c(m+2) + 4\}E(\chi_{p+4}^{-4}(\Delta))] \right)^{-\frac{1}{p}}$$

and

$$\text{GRE}_U^{(S)} = \left(1 - \frac{cm}{4} (\text{tr } \Sigma^0) [\{2E(\chi_{p+2}^{-2}(\Delta)) - c(m+2)E(\chi_{p+2}^{-4}(\Delta))\} - \Delta Ch_{\min}(\Sigma^0)(\text{tr } \Sigma^0)^{-1} \{c(m+2) + 4\}E(\chi_{p+4}^{-4}(\Delta))] \right)^{-\frac{1}{p}}$$

These expressions imply that we reach the same conclusions as before.

4.5. CONCLUDING REMARKS

We reach the same conclusion about the behaviors of estimators using risk criterion as well as mean square error matrix measure. It should be remarked that for $p \geq 3$ and Δ unknown, $\hat{\mu}^{(1)S}$ is recommended for estimating $\mu^{(1)}$ while for $p \leq 2$, $\hat{\mu}^{(1)PT}$ may be a sensible choice. It should be noted that, in a preliminary test estimation problem, one sets to test the null hypothesis first, before choosing the estimator, while in a shrinkage estimation, the test statistic is itself incorporated in the estimator to adjust for possible shift from the null hypothesis.

CHAPTER 5

ESTIMATION STRATEGIES FOR THE INTERCEPT VECTOR IN A SIMPLE LINEAR MULTIVARIATE NORMAL REGRESSION MODEL

5.1. INTRODUCTION

In the presence of uncertain prior information in the form of an initial estimate or realistic guessed value about the slope vector, four estimation strategies have been developed for the intercept parameter. In this situation, the estimates based on a preliminary test as well as on the Stein-rule are desirable. Exact bias and risks of all of these estimators are derived and their efficiencies relative to a classical estimators are studied. An optimum rule for the preliminary test estimator (PTE) is discussed and tables for the appropriate choice of the significance levels of that test are provided. It is shown that the Stein-rule estimator (SE) dominates the classical one, whereas none of the PTE and SE dominate each other.

Four estimation strategies are developed in Section 5.2; the derivation of the bias and risk (under quadratic loss function) are presented in Section 5.3; and risk efficiency analysis is discussed in Section 5.3.2, mean square matrices are derived in Section 5.4 and the generalized efficiency is discussed in Section 5.4.1. Section 5.5 summarizes the findings.

Suppose $y_1, y_2, \dots, y_N$ are a set of N observations from $N_p(\theta + \beta x_j, \Sigma)$; $x_1, \dots, x_N$ are fixed constants, Σ is a $(p \times p)$ matrix and the $(p \times 1)$ vectors θ and β are unknown. We assume that $N \geq p + 2$ and rank of a non-stochastic matrix X is 2, where

$$X = \begin{pmatrix} 1 & x_1 \\ 1 & \vdots \\ \vdots & \vdots \\ 1 & x_N' \end{pmatrix}$$

We are primarily interested in the estimation of θ when it is apriori suspected that $\beta = 0$ is plausible. The estimator of θ is different in the two situations where β is specified or not. Then, $\tilde{\theta} = \bar{y} - \tilde{\beta}\bar{x}_n$ and $\hat{\theta} = \bar{y}$ are the maximum likelihood estimators of θ when $\beta \neq 0$ and $\beta = 0$ respectively where $\tilde{\beta}$ is the usual maximum likelihood estimator of β . Usually $\tilde{\theta}$ has the larger dispersion than $\hat{\theta}$, while $\hat{\theta}$ may not be a desirable estimator when departure occurs from the null hypothesis, $H_0: \beta = 0$ and it may be even the worst i.e. inconsistent. Therefore a compromise between these two estimators is needed and we consider the preliminary test maximum likelihood estimator (PTMLE) which is obtained by performing a preliminary test on the uncertain portion of information, $H_0: \beta = 0$. If the null hypothesis is tenable, then $\hat{\theta}$ is used, otherwise $\tilde{\theta}$ is a sensible choice. Finally, we consider another estimator, which is a smooth version of PTMLE, is the Stein-rule estimator (SMLE) and is denoted by $\hat{\theta}^S$. Stein (1956) and James and Stein (1961) provided such an estimator for the mean-vector of a p -dimensional normal distribution when $p \geq 3$. The main object of this chapter is to study the performance of these four estimators under a quadratic loss function (Mahalanobis distance). Sen and Saleh (1979) have considered this problem for large samples.

5.2. ESTIMATION STRATEGIES

We consider four methods to estimate the intercept parameter θ in the presence of a uncertain prior information about the slope parameter β .

i) Unrestricted maximum likelihood estimator (UMLE):

The UMLE of θ and β are:

$$\tilde{\theta} = \bar{y} - \tilde{\beta}\bar{x}, \quad (5.2.1)$$

$$\tilde{\beta} = \sum_{j=1}^N (y_j - \bar{y})(x_j - \bar{x}) / \sum_{j=1}^N (x_j - \bar{x})^2, \quad (5.2.2)$$

where

$$\bar{y} = \frac{1}{N} \sum_{j=1}^N y_j, \quad \bar{x} = \frac{1}{N} \sum_{j=1}^N x_j.$$

ii) Restricted maximum likelihood estimator (RMLE):

The RMLE of θ is defined as:

$$\hat{\theta} = \bar{y} = \frac{1}{N} \sum_{j=1}^N y_j.$$

iii) Preliminary test maximum likelihood estimator (PTMLE):

Since PTMLE is a function of test-statistic, we consider the likelihood ratio test

$$L^* = \{ |S_e| / |S_e + Q\tilde{\beta}\tilde{\beta}'| \}^{N/2}, \quad (5.2.3)$$

where

$$Q = \sum_{j=1}^N (x_j - \bar{x})^2, \quad S_e = (N-2)S_e^*$$

and

$$S_e^* = (N-2)^{-1} \sum_{j=1}^N \{ (y_j - \bar{y}) - \tilde{\beta}(x_j - \bar{x}) \} \cdot \{ (y_j - \bar{y}) - \tilde{\beta}(x_j - \bar{x}) \}' \quad (5.2.4)$$

We observe that the ratio test is equivalent to the Hotelling T^2 -test defined as

$$T^2 = Q\tilde{\beta}' S_e^{*-1} \tilde{\beta}. \quad (5.2.5)$$

Now we are in a position to construct the PTMLE:

$$\hat{\theta}^{PT} = \begin{cases} \hat{\theta} & \text{if } T^2 < T_{N,\alpha}^2 \\ \tilde{\theta} & \text{if } T^2 \geq T_{N,\alpha}^2 \end{cases} \quad (5.2.6)$$

$$= \tilde{\theta} + \tilde{\beta}\bar{x}I(T^2 < T_{N,\alpha}^2), \quad (5.2.7)$$

where $T_{N,\alpha}^2 = \frac{(N-2)p}{(N-p-1)} F_{p,N-p-1}(\alpha)$ and $F_{p,N-p-1}(\alpha)$ is the upper α -percentile of a central F -distribution with $(p, N-p-1)$ degrees of freedom; $I(A)$ is the indicator function for a set A .

iv) Shrinkage maximum likelihood estimator (SMLE):

We construct a smooth version of PTMLE, shrinkage estimator as:

$$\hat{\theta}^S = \hat{\theta} + [1 - c(N-2)T^{-2}](\tilde{\theta} - \hat{\theta}) \quad (5.2.8)$$

$$= \tilde{\theta} + c\bar{x}(N-2)\tilde{\beta}T^{-2}, \quad (5.2.9)$$

where c is the shrinkage constant. The optimal value of c will be determined, while computing the risk for SML.

At this point we would like to make a remark, that in the construction of both PTMLE (5.2.7) and SML (5.2.9) we have used the test-statistic T^2 which adjusts the departure from the null hypothesis. For large T^2 values, $\hat{\theta}^{PT}$ and $\hat{\theta}^S$ reduce to the same value while for small T^2 -value they behave differently. These properties along with others will be studied using risk analysis and mean square error matrix analysis.

5.3. BIAS AND RISKS

We obtain expressions for the bias and risks of four estimators. Theorems 5.3.1 and 5.3.2 provide such expressions.

THEOREM 5.3.1.

$$B_1 = E(\tilde{\theta} - \theta) = 0, \quad (5.3.10)$$

$$B_2 = E(\hat{\theta} - \theta) = \beta\bar{x}, \quad (5.3.11)$$

$$B_3 = E(\hat{\theta}^{PT} - \theta) = \beta\bar{x}H_{p+2,N-p-1}(T_{N,\alpha}^2; \Delta), \quad (5.3.12)$$

where $H_{v_1, v_2}(T_{N, \alpha}^{*2}; \Delta)$ is the cumulative distribution function of a non-central F -distribution with (v_1, v_2) degrees of freedom and non-centrality parameter $\Delta = Q\beta'\Sigma^{-1}\beta$ and

$$T_{N, \alpha}^{*2} = \frac{p}{(N - p - 1)} F_{p, N-p-1}(\alpha).$$

$$B_4 = (\hat{\theta}^S - \theta) = c\bar{x}(N - p - 1)E(\chi_p^{-2}(\Delta)) \quad (5.3.13)$$

where $\chi_p^2(\Delta)$ is a non-central chi-square random variable with p degree of freedom and non-centrality parameter Δ , defined as above.

PROOF. B_1 and B_2 are straightforward. For B_3 and B_4 we note that

$$\begin{aligned} B_3 &= E(\hat{\theta}^{PT} - \theta) = E(\tilde{\theta} - \theta) + E\{\tilde{\beta}\bar{x}I(T^2 < T_{N, \alpha}^2)\} \\ &= \bar{x}E\{\tilde{\beta}I(T^2 < T_{N, \alpha}^2)\} \\ &= \beta\bar{x}H_{p+2, N-p-1}(T_{N, \alpha}^{*2}; \Delta) \end{aligned} \quad (5.3.14)$$

and

$$\begin{aligned} B_4 &= E(\hat{\theta} - \theta) = c\bar{x}\beta E\{\tilde{\beta}(N - 2)T^{-2}\} \\ &= c\bar{x}(N - p - 1)E(\chi_p^{-2}(\Delta))\beta. \end{aligned}$$

Both follow by using Theorem A.1 from Appendix A and transformation $\sqrt{Q}\tilde{\beta} = y$ and $y^* = Dy$ such that $D\bar{\Sigma}D' = I_p$.

For the risks of these estimators we consider the quadratic loss function which is really the Mahalanobis distance of θ^* from θ :

$$L(\theta^*, \theta) = (\theta^* - \theta)' \Sigma^{-1} (\theta^* - \theta). \quad (5.3.15)$$

THEOREM 5.3.2. Risks of UMLE, RMLE, PTMLE and SMLE under (5.3.15) are given by (5.3.16) - (5.3.19) respectively:

$$R_1(\tilde{\theta}; \Sigma^{-1}) = p\left(\frac{1}{N} + \frac{\bar{x}^2}{Q}\right) \quad (5.3.16)$$

$$R_2(\hat{\theta}; \Sigma^{-1}) = \frac{\bar{p}}{N} + \frac{\bar{x}^2}{Q} \Delta \quad (5.3.17)$$

$$R_3(\hat{\theta}^{PT}; \Sigma^{-1}) = p\left(\frac{1}{N} + \frac{\bar{x}^2}{Q}\right) - \frac{\bar{x}^2 p}{Q} H_{p+2, N-p-1}(T_{N, \alpha}^2; \Delta) \\ + \frac{\bar{x}^2}{Q} \Delta \{2H_{p+2, N-p-1}(T_{N, \alpha}^{2*}; \Delta) - H_{p+4, N-p-1}(T_{N, \alpha}^{2*}; \Delta)\} \quad (5.3.18)$$

$$R_4(\hat{\theta}^S; \Sigma^{-1}) = p\left(\frac{1}{N} + \frac{\bar{x}^2}{Q}\right) - \frac{\bar{x}^2 cm}{Q} \{2(p-2) - c(m+2)\} E(\chi_p^{-2}(\Delta)), \quad (5.3.19)$$

where $m = N - p + 1$ for $c: 0 < c < \frac{2(p-2)}{m+2}$. The optimal value of c from (5.3.19) is

$$c = \frac{p-2}{m+2}.$$

We only provide the proof of (5.3.18) and (5.3.19).

PROOF.

$$R_3(\hat{\theta}^{PT}; \Sigma^{-1}) = E(\hat{\theta}^{PT} - \theta)' \Sigma^{-1} (\hat{\theta}^{PT} - \theta) \\ = E \left[\{(\tilde{\theta} - \theta) + \bar{x} \tilde{\beta} I(T^2 < T_{N, \alpha}^2)\}' \Sigma^{-1} \cdot \right. \\ \left. \{(\tilde{\theta} - \theta) + \bar{x} \tilde{\beta} I(T^2 < T_{N, \alpha}^2)\} \right] \\ = E(\tilde{\theta} - \theta)' \Sigma^{-1} (\tilde{\theta} - \theta) + \bar{x}^2 E \left\{ \tilde{\beta}' \Sigma^{-1} \tilde{\beta} I(T^2 < T_{N, \alpha}^2) \right\} \\ + 2\bar{x} E \left\{ (\tilde{\theta} - \theta)' \Sigma^{-1} \tilde{\beta} I(T^2 < T_{N, \alpha}^2) \right\}. \quad (5.3.20)$$

Since,

$$E(\tilde{\theta} - \theta)(\tilde{\theta} - \theta)' = \left(\frac{1}{N} + \frac{\bar{x}^2}{Q} \right) \Sigma,$$

therefore

$$E(\tilde{\theta} - \theta)' \Sigma^{-1} (\tilde{\theta} - \theta) = p \left(\frac{1}{N} + \frac{\bar{x}^2}{Q} \right). \quad (5.3.21)$$

Next, we consider

$$\begin{aligned} E \{ (\tilde{\beta}' \Sigma^{-1} \tilde{\beta}) I(T^2 < T_{N,\alpha}^2) \} &= \frac{1}{Q} E \{ (Q \tilde{\beta}' \Sigma^{-1} \tilde{\beta}) I(T^2 < T_{N,\alpha}^2) \} \\ &= \frac{1}{Q} \{ p H_{p+2, N-p-1}(T_{N,\alpha}^{*2}; \Delta) + \Delta H_{p+4, N-p-1}(T_{N,\alpha}^{*2}; \Delta) \}, \quad (5.3.22) \end{aligned}$$

using Theorem A.2 from Appendix A.

Finally,

$$\begin{aligned} &E \{ (\tilde{\theta} - \theta)' \Sigma^{-1} \tilde{\beta} I(T^2 < T_{N,\alpha}^2) \} \\ &= E \{ [(\bar{y} - \tilde{\beta} \bar{x}) - (\theta + \beta \bar{x}) + \beta \bar{x}]' \Sigma^{-1} \tilde{\beta} I(T^2 < T_{N,\alpha}^2) \} \\ &= E \{ [(\bar{y} - \theta - \beta \bar{x}) - \bar{x}(\tilde{\beta} - \beta)]' \Sigma^{-1} \tilde{\beta} I(T^2 < T_{N,\alpha}^2) \} \\ &= E \{ \bar{x} \beta' \Sigma^{-1} \tilde{\beta} I(T^2 < T_{N,\alpha}^2) - \bar{x} \tilde{\beta}' \Sigma^{-1} \tilde{\beta} I(T^2 < T_{N,\alpha}^2) \} \\ &= \frac{\bar{x}}{Q} \Delta H_{p+2, N-p-1}(T_{N,\alpha}^{*2}; \Delta) - \frac{\bar{x}^2}{Q} \{ p H_{p+2, N-p-1}(T_{N,\alpha}^{*2}; \Delta) \\ &\quad + \Delta H_{p+4, N-p-1}(T_{N,\alpha}^{*2}; \Delta) \}. \quad (5.3.23) \end{aligned}$$

Substituting (5.3.21), (5.3.22) and (5.3.23) in (5.3.20) and after some simplification, we establish (5.3.18). Now, consider risk for $\hat{\theta}^S$

$$\begin{aligned} R_4(\hat{\theta}^S; \Sigma^{-1}) &= E(\hat{\theta}^S - \theta)' \Sigma^{-1} (\hat{\theta}^S - \theta) \\ &= E \{ [(\tilde{\theta} - \theta) + c \bar{x} \tilde{\beta} (N-2) T^{-2}]' \Sigma^{-1} [(\tilde{\theta} - \theta) + c \bar{x} \tilde{\beta} (N-2) T^{-2}] \} \\ &= E \{ (\tilde{\theta} - \theta)' \Sigma^{-1} (\tilde{\theta} - \theta) \} + c^2 \bar{x} E \{ \tilde{\beta}' \Sigma^{-1} \tilde{\beta} (N-2)^2 T^{-4} \} \\ &\quad + 2c \bar{x} E \{ (\tilde{\theta} - \theta)' \Sigma^{-1} \tilde{\beta} (N-2) T^{-2} \}, \end{aligned}$$

applying transformation $y = \sqrt{Q} \tilde{\beta}$

$$\begin{aligned} &= p \left(\frac{1}{N} + \frac{\bar{x}^2}{Q} \right) E \left\{ \frac{c^2 \bar{x}^2}{Q} (y' \Sigma^{-1} y) (y' S^{-1} y / (N-2))^{-2} \right\} \\ &\quad + 2c \frac{\bar{x}^2}{Q} E \left\{ (y' \Sigma^{-1} y) (y' S^{-1} y / (N-2))^{-1} \right\} \\ &\quad - 2c \frac{\bar{x}^2}{\sqrt{Q}} E \left\{ (\beta' \Sigma^{-1} y) (y' S^{-1} y / (N-2))^{-1} \right\}, \quad (5.3.24) \end{aligned}$$

making another transformation

$$y^* = Dy \text{ such that } DED' = I_p \text{ and}$$

$$E(y^*) = D(\sqrt{Q}\beta), \text{ Cov}(y^*) = I_p.$$

$$\begin{aligned} E\{\tilde{\beta}(N-2)T^{-2}\} &= \frac{1}{\sqrt{Q}}D^{-1}E\{y^{*'}(y^{*'}y^*/\chi_{N-p-1}^2)\} \\ &= \beta(N-p-1)E(\chi_{p+2}^{-2}(\Delta)), \end{aligned} \quad (5.3.25)$$

by Theorem A.1 from Appendix A. Similarly,

$$\begin{aligned} E\{(\tilde{\beta}\Sigma^{-1}\tilde{\beta})(N-2)^2T^{-4}\} &= \frac{1}{Q}E\{(y^{*'}y^*)(y^{*'}y^*/\chi_{N-p-1}^2)^{-2}\} \\ &= \frac{1}{Q}(N-p-1)(N-p+1)E(\chi_p^{-2}(\Delta)). \end{aligned} \quad (5.3.26)$$

Finally,

$$\begin{aligned} E\{(\tilde{\beta}'\Sigma^{-1}\tilde{\beta})(N-2)T^{-2}\} &= \frac{1}{Q}E\{(y^{*'}y^*)(y^{*'}y^*/\chi_{N-p-1}^2)^{-1}\} \\ &= \frac{1}{Q}(N-p-1). \end{aligned} \quad (5.3.27)$$

Combining (5.3.25), (5.3.26) and (5.3.27) in (5.3.24) and after some simplification

$$\begin{aligned} R_4(\hat{\theta}^S; \Sigma^{-1}) &= p\left(\frac{1}{N} + \frac{\bar{x}^2}{Q}\right) + \frac{c^2\bar{x}^2}{Q}(N-p-1)(N-p-1)E(\chi_p^{-2}(\Delta)) \\ &\quad + 2c(N-p-1)\frac{\bar{x}^2}{Q} - 2\frac{c\bar{x}^2}{Q}(N-p-1)\Delta E(\chi_{p+2}^{-2}(\Delta)) \\ &= p\left(\frac{1}{N} + \frac{\bar{x}^2}{Q}\right) - c(N-p-1)\frac{\bar{x}^2}{Q}[2\{\Delta E(\chi_{p+2}^{-2}(\Delta)) - 1\} \\ &\quad - c(N-p+1)E(\chi_p^{-2}(\Delta))] \\ &= p\left(\frac{1}{N} + \frac{\bar{x}^2}{Q}\right) - c(N-p-1)\frac{\bar{x}^2}{Q}\{2(p-2) - c(N-p+1)\}E(\chi_p^{-2}(\Delta)). \end{aligned}$$

Thus, the optimum value of c is

$$c_{opt} = \frac{p-2}{N-p+1}.$$

5.3.1. RISK ANALYSIS

In this section, we analyse the risks of four estimators for θ . First, we note that from (5.3.11) to (5.3.13) for $\bar{x} = 0$, it follows that

$$B_1 = B_2 = B_3 = B_4 = 0 \quad (5.3.28)$$

and

$$R_1(\bar{\theta}; \Sigma^{-1}) = R_2(\hat{\theta}; \Sigma^{-1}) = R_3(\hat{\theta}^{PT}; \Sigma^{-1}) = R_4(\hat{\theta}^S; \Sigma^{-1}) = \frac{p}{N}. \quad (5.3.29)$$

Therefore we conclude that all four estimators are unbiased having the same risk p/N , provided $\bar{x} = 0$. When this condition is removed, i.e., $\bar{x} \neq 0$, then only UMLE is an unbiased estimator and its risk is independent of Δ .

To visualize the risk under the null hypothesis, we establish the following dominance theorem.

THEOREM 5.3.3.

Under $H_0 : \beta = 0$ provided $\bar{x} \neq 0$, the dominance picture is:

$$\bar{\theta} > \hat{\theta}^{PT} > \hat{\theta}^S > \bar{\theta}. \quad (5.3.30)$$

PROOF. The proof follows immediately from Theorem 5.3.2, since

$$R_1(\bar{\theta}; \Sigma^{-1}) - R_2(\hat{\theta}; \Sigma^{-1}) = p \cdot \left(\frac{\bar{x}^2}{Q}\right) > 0, \quad (5.3.31)$$

$$R_1(\bar{\theta}; \Sigma^{-1}) - R_3(\hat{\theta}^{PT}; \Sigma^{-1}) = p \cdot \left(\frac{\bar{x}^2}{Q}\right) H_{p+2, N-p-1}(T_{N, \alpha}^{*2}; 0) > 0, \quad (5.3.32)$$

$$R_1(\bar{\theta}; \Sigma^{-1}) - R_4(\hat{\theta}^S; \Sigma^{-1}) = \frac{\bar{x}^2(p-2)(N-p-1)}{Q(N-p+1)} > 0, \quad (5.3.33)$$

$$R_3(\hat{\theta}^{PT}; \Sigma^{-1}) - R_2(\hat{\theta}; \Sigma^{-1}) = p \cdot \left(\frac{\bar{x}^2}{Q}\right) \{1 - H_{p+2, N-p-1}(T_{N, \alpha}^{*2}; 0)\} > 0, \quad (5.3.34)$$

$$R_4(\hat{\theta}^S; \Sigma^{-1}) - R_3(\hat{\theta}^{PT}; \Sigma^{-1}) = \left(\frac{\bar{x}^2}{Q}\right) \{p H_{p+2, N-p-1}(T_{N, \alpha}^{*2}; 0) - \left(\frac{N-p-1}{N-p+1}\right)(p-2)\} > 0 \quad (5.3.35)$$

(5.3.35) holds only if

$$pH_{p+2, N-p-1}(T_{N, \alpha}^{*2}; 0) > (p-2)\left(\frac{N-p-1}{N-p+1}\right). \quad (5.3.36)$$

Equation (5.3.36) indicates that there is a range for α ($0 < \alpha < 1$) for which (5.3.35) is satisfied.

To make risk analysis more manageable, the risk efficiency of different estimators relative to UMLE are considered.

5.3.2. RISK EFFICIENCY ANALYSIS AND RECOMMENDATIONS

In this section, we consider the risk efficiency of the four estimators for θ and provide a max-min rule for the optimum choice of the level of the significance for the preliminary test of the null hypothesis $H_0: \beta = 0$. Tables for the relative efficiency (maximum and minimum) and Δ -value at which minimum efficiency of $\hat{\theta}^{PT}$ occurs relative to $\tilde{\theta}$ are provided.

The relative efficiency of RMLE ($\hat{\theta}$) to UMLE ($\tilde{\theta}$) is:

$$R.E.(\hat{\theta} : \tilde{\theta}) = p\left(\frac{1}{N} + \frac{\bar{x}^2}{Q}\right)\left(\frac{p}{N} + \frac{\bar{x}^2}{Q}\Delta\right)^{-1}. \quad (5.3.37)$$

Thus (5.3.37) assumes a constant value 1 on the ellipsoid $\beta'\Sigma^{-1}\beta = \frac{p}{Q}$ and $R.E.(\hat{\theta} : \tilde{\theta})$ is a monotone decreasing function of Δ . The limits are

$$\lim_{\Delta \rightarrow 0} R.E.(\hat{\theta} : \tilde{\theta}) = \left(1 + \frac{N\bar{x}^2}{Q}\right) \quad \text{and} \quad \lim_{\Delta \rightarrow \infty} R.E.(\hat{\theta} : \tilde{\theta}) = 0.$$

Further,

$$R.E.(\hat{\theta} : \tilde{\theta}) \geq 1 \quad \text{according as} \quad \Delta \leq p. \quad (5.3.38)$$

Thus, for $\Delta > p$, $\hat{\theta}$ becomes inefficient.

Next, we compute the relative efficiency of $\hat{\theta}^{PT}$ to $\bar{\theta}$:

$$R.E. (\hat{\theta}^{PT} : \bar{\theta}) = [1 + g(\Delta)]^{-1}, \quad (5.3.39)$$

where

$$g(\Delta) = p^{-1} \Delta \frac{\bar{x}^2}{Q} \left(\frac{1}{N} + \frac{\bar{x}^2}{Q} \right)^{-1} \{ 2H_{p+2, N-p-1}(T_{N, \alpha}^{*2}; \Delta) - H_{p+4, N-p-1}(T_{N, \alpha}^{*2}; \Delta) \} - \frac{\bar{x}^2}{Q} \left(\frac{1}{N} + \frac{\bar{x}^2}{Q} \right)^{-1} H_{p+2, N-p-1}(T_{N, \alpha}^{*2}; \Delta) \quad (5.3.40)$$

For a given $\frac{\bar{x}^2}{Q}$ and N , the R.E. $(\hat{\theta}^{PT} : \bar{\theta})$ is a function of α and Δ , $E(\alpha, \Delta)$. The function $E(\alpha, \Delta)$ for $\alpha \neq 0$ has its maximum at $\Delta = 0$ with value

$$\left\{ 1 - \frac{\bar{x}^2}{Q} \left(\frac{1}{N} + \frac{\bar{x}^2}{Q} \right)^{-1} H_{p+2, N-p-1}(T_{N, \alpha}^{*2}; 0) \right\}^{-1},$$

which is strictly greater than one. But as Δ increases, the efficiency decreases, crossing the line at $E(\alpha, \Delta) = 1$, to a minimum, then increases to a maximum towards 1 as Δ tends to infinity.

For analysis purposes, first we consider the case where $\Delta = 0$ and α varies; we have $\max_{\alpha} E(\alpha, 0) = E(0, 0) = 1 + \frac{N\bar{x}^2}{Q}$. The value of $E(\alpha, 0)$ decreases as α increases. On the other hand, if $\alpha = 0$ and Δ vary; then the curve $E(0, \Delta)$ and $E(1, \Delta) = 1$ intersect at $\Delta = p$. For general α , $E(\alpha, \Delta)$ and $E(1, \Delta)$ will intersect in the interval $0 \leq \Delta \leq p$. The value of Δ at the intersection decreases as α increases. Therefore, for two different α , $E(\alpha_1, \Delta)$ and $E(\alpha_2, \Delta)$ will never intersect above the line $E(\alpha, \Delta) = 1$, but somewhere below the line $E(\alpha, \Delta) = 1$. Thus, we have plotted $E(\alpha, \Delta)$ versus Δ for $\alpha = 0.05, 0.1, 0.2, 0.5$ and various sample sizes (and $\bar{x}^2/Q = 0.5$) to observe such features. Figure 8 ($p=2$) and Figure 9 ($p=4$) display such properties of PTMLE.

$N = 8$ and $\alpha = 0.05, 0.1, 0.2, 0.5$

$N = 12$ and $\alpha = 0.05, 0.1, 0.2, 0.5$

$N = 16$ and $\alpha = 0.05, 0.1, 0.2, 0.5$

$N = 20$ and $\alpha = 0.05, 0.1, 0.2, 0.5$

Fig.8 Relative efficiency of PTMLE to UMLE for various sample sizes, with $p=2$ (and $\frac{\sigma^2}{\mu^2}=0.5$.)

$N=5$ and $\alpha = 0.05, 0.1, 0.2, 0.5$

$N=10$ and $\alpha = 0.05, 0.1, 0.2, 0.5$

$N=16$ and $\alpha = 0.05, 0.1, 0.2, 0.5$

$N=20$ and $\alpha = 0.05, 0.1, 0.2, 0.5$

Fig.9 Relative efficiency of PTMLE to UMLE for various sample size n with $p=1$ (and $\bar{X} = 0.5$)

In order to choose an estimator with minimum guaranteed relative efficiency we adopt the following procedure: If it is known that $0 \leq \Delta \leq p$, then $\hat{\theta}$ is chosen since $E(\alpha, \Delta)$ is the largest for all Δ for this range. However, if $\Delta > p$ and we are willing to accept an estimator with relative efficiency at least E_0 , then we solve the equation

$$\max_{\alpha} \min_{\Delta} E(\alpha, \Delta) = E_0,$$

for all α in the set

$$A = \{\alpha \mid E(\alpha, \Delta) \geq E_0 \text{ for all } \Delta\}$$

Tables from 14 to 17 for maximum (E^*) and minimum (E_0) efficiency with associated minimum Δ -value for $\alpha = 0.05, 0.10, 0.2$ and 0.5 have been prepared for selected values of $\frac{\Sigma^2}{Q}$. The purpose of preparing the table is two-fold; one is to examine the performances of estimators and the other is to determine the optimum choice of the significance levels of the preliminary tests, say α^* which provides the minimum guaranteed efficiency E_0 . For example, if $N = 28$, $\frac{\Sigma^2}{Q} = 0.90$ and $p = 5$, and the experimenter wishes to have an estimator with a minimum guaranteed efficiency 0.74, then using Table 17 we recommend him to select $\alpha^* = 0.10$. Because such a choice of α^* would yield an estimator with a maximum efficiency 4.1268 at $\Delta = 0$.

[illegible]

Table 16 Maximum and minimum guaranteed efficiencies for $j = 4$, $N = 8$

[illegible]

[illegible]

Table 17: Maximum and minimum guaranteed efficiencies for $r = 5$, $N = 8$.

Δ/Q	.10	.20	.30	.40	.50	.60	.70	.80	.90
.05 E^*	1.7055	2.3405	2.9152	3.4377	3.9149	4.3524	4.7549	5.1265	5.4707
E	.3586	.2876	.2604	.2459	.2370	.2309	.2265	.2232	.2206
Δ	39.1864	39.1864	39.1864	39.1864	39.1564	39.1864	39.1864	39.1864	39.1864
.10 E^*	1.6220	2.1322	2.5581	2.9190	3.2288	3.4976	3.7331	3.9410	4.1260
E	.4047	.3853	.3534	.3361	.3253	.3179	.3126	.3085	.3052
Δ	36.9388	36.9388	36.9388	36.9388	36.9388	36.9388	36.9388	36.9388	36.9388
.15 E^*	1.5480	1.9614	2.2844	2.5467	2.7565	2.9342	3.0849	3.2142	3.3265
E	.5723	.4915	.4573	.4384	.4264	.4181	.4121	.4075	.4038
Δ	33.3233	33.3233	33.3233	33.3233	33.3233	33.3233	33.3233	33.3233	33.3233
.20 E^*	1.4819	1.8191	2.0682	2.2598	2.4117	2.5351	2.6373	2.7234	2.7969
E	.6070	.5913	.5577	.5388	.5267	.5182	.5120	.5072	.5034
Δ	27.8307	27.8307	27.8307	27.8307	27.8307	27.8307	27.8307	27.8307	27.8307
.25 E^*	1.4227	1.6989	1.8995	2.0685	2.1995	2.2882	2.3505	2.3975	2.4310
E	.6743	.6742	.6184	.6257	.6142	.6061	.6002	.5956	.5919
Δ	22.9033	22.9033	22.9033	22.9033	22.9033	22.9033	22.9033	22.9033	22.9033
.30 E^*	1.3694	1.5962	1.7196	1.8092	1.8638	1.9002	1.9218	1.9409	1.9490
E	.7086	.7411	.7119	.6981	.6877	.6804	.6750	.6707	.6674
Δ	19.4560	19.4560	19.4560	19.4560	19.4560	19.4560	19.4560	19.4560	19.4560
.35 E^*	1.3213	1.5076	1.6292	1.7148	1.7784	1.8274	1.8664	1.8982	1.9246
E	.8431	.7951	.7719	.7582	.7491	.7427	.7379	.7342	.7312
Δ	16.9524	16.9524	16.9524	16.9524	16.9524	16.9524	16.9524	16.9524	16.9524
.40 E^*	1.2778	1.4366	1.5273	1.5940	1.6428	1.6801	1.7094	1.7332	1.7528
E	.8781	.8337	.8193	.8077	.8000	.7946	.7905	.7873	.7847
Δ	15.0422	15.0422	15.0422	15.0422	15.0422	15.0422	15.0422	15.0422	15.0422
.45 E^*	1.2383	1.3633	1.4402	1.4924	1.5300	1.5585	1.5808	1.5987	1.6135
E	.9057	.8740	.8581	.8485	.8421	.8376	.8341	.8315	.8294
Δ	13.5264	13.5264	13.5264	13.5264	13.5264	13.5264	13.5264	13.5264	13.5264
.50 E^*	1.2025	1.3041	1.3651	1.4059	1.4350	1.4569	1.4739	1.4875	1.4986
E	.9276	.9024	.8897	.8819	.8768	.8741	.8703	.8681	.8663
Δ	12.2849	12.2849	12.2849	12.2849	12.2849	12.2849	12.2349	12.2849	12.2849

The relative efficiency of $\hat{\theta}^{PT}$ to $\hat{\theta}$ is:

$$R.E. (\hat{\theta}^{PT} : \hat{\theta}) = (p/N + \frac{\bar{x}^2}{Q} \Delta) \left[p \left(\frac{1}{N} + \frac{\bar{x}^2}{Q} \right) - \frac{\bar{x}^2 p}{Q} \right. \\ \left. \cdot H_{p+2, N-p-1}(T_{n, \alpha}^{*2}; \Delta) + \frac{\bar{x}^2}{Q} \Delta \{ 2H_{p+2, N-p-1}(T_{N, \alpha}^{*2}; \Delta) \right. \\ \left. - H_{p+4, N-p-1}(T_{N, \alpha}^{*2}; \Delta) \} \right]^{-1} \quad (5.3.41)$$

This expression (5.3.41) achieves constant value on the ellipsoid $\beta' \Sigma^{-1} \beta = p/Q$ with value

$$\left[1 + \frac{\bar{x}^2}{Q} \left(\frac{1}{N} + \frac{\bar{x}^2}{Q} \right)^{-1} \{ H_{p+2, N-p-1}(T_{n, \alpha}^{*2}; p) - H_{p+4, N-p-1}(T_{N, \alpha}^{*2}; p) \} \right]^{-1} \quad (5.3.42)$$

Thus, the relative efficiency (5.3.41) is less than 1 on or inside the ellipsoid, therefore $\hat{\theta}$ is superior than $\hat{\theta}^{PT}$; but as β goes outside the ellipsoid, the situation is reversed, $\hat{\theta}^{PT}$ becomes more efficient than $\hat{\theta}$.

Finally, the relative efficiency of $\hat{\theta}^S$ to $\hat{\theta}$ is

$$R.E. (\hat{\theta}^S : \hat{\theta}) = \left\{ 1 - \frac{\bar{x}^2}{Q} \left(\frac{1}{N} + \frac{\bar{x}^2}{Q} \right)^{-1} \frac{(p-2)^2 (N-p-1)}{p(N-p+1)} E(\chi_p^2(\Delta)) \right\}^{-1} \geq 1, \quad (5.3.43)$$

for all $\Delta \geq 0$. Thus, $\hat{\theta}^S$ is uniformly more efficient than $\hat{\theta}$. The relative efficiency is a monotonically decreasing function in Δ with maximum value

$$\left\{ 1 - \frac{\bar{x}^2}{Q} \left(\frac{1}{N} + \frac{\bar{x}^2}{Q} \right)^{-1} \frac{(p-2)^2 (N-p-1)}{p(N-p+1)} \right\}^{-1} \quad (5.3.44)$$

at $\Delta = 0$. Also,

$$\lim_{\Delta \rightarrow \infty} R.E. (\hat{\theta}^S : \hat{\theta}) = 1.$$

Recall that, $\lim_{\Delta \rightarrow \infty} R.E. (\hat{\theta}^{PT} : \hat{\theta}) = 1$. Thus for large values of Δ the performance of $\hat{\theta}^S$ and $\hat{\theta}^{PT}$ are almost similar. A noticeable difference occurs for small value of Δ , where $\hat{\theta}^{PT}$ dominates $\hat{\theta}^S$.

Thus, from the point of robust efficiency, both $\hat{\theta}^{PT}$ and $\hat{\theta}^S$ may be advocated, leaning more towards $\hat{\theta}^S$ since the size of Δ is generally unknown and unlikely to be small. Since, $\hat{\theta}^{PT}$ dominates $\hat{\theta}^S$ for a range of α ($0 < \alpha < 1$) and for small values of Δ , thus we have plotted the efficiency curves of $\hat{\theta}$, $\hat{\theta}^{PT}$, $\hat{\theta}^S$, versus Δ relative to $\bar{\theta}$ for various values of α and fixed sample size. From Figures 10 to 12, for $p=3(1)5$, it is found that for larger values of α , $\hat{\theta}^S$ dominates $\hat{\theta}^{PT}$ uniformly whereas for small values of α and small Δ , $\hat{\theta}^{PT}$ performs better than $\hat{\theta}^S$. Therefore, we need to find the range of Δ where $\hat{\theta}^{PT}$ performs better than $\hat{\theta}^S$. Tables for the maximum relative efficiencies (at $\Delta=0$) for $\hat{\theta}^{PT}(E_3)$, $\hat{\theta}^S(E_4)$, and $\hat{\theta}(E_1)$ relative to $\bar{\theta}$ and the intersecting relative efficiencies ($E_{\Delta\alpha}$) of $\hat{\theta}^S$ and $\hat{\theta}^{PT}$ and the Δ_α -values at which the intersection occurs have been computed. Tables 18 to 20 provide these values for $p=3(1)5$, $\frac{\bar{x}^2}{Q} = 0.10(0.10)0.90$ and various sample sizes. For example, if $N=20$, $\frac{\bar{x}^2}{Q} = 0.90$ and $p=5$ with $\alpha=0.10$, then by using Table 19, if $\Delta \in [0, 1.7915]$, then $\hat{\theta}^{PT}$ dominates $\hat{\theta}^S$, but outside this the interval domination is reversed. On the other hand, if the hypothesis is correct, i.e., $\Delta = 0$, then $\hat{\theta}$ has the largest efficiency ($E_1 = 19.0000$) than $\hat{\theta}^{PT}(E_3 = 3.8018)$ and $\hat{\theta}^S(E_4 = 1.7181)$ relative to $\bar{\theta}$.

Fig.10 Dominance picture of UMLE¹, RMLE², PMLE³ and SMLE⁴ for $p=3$, $N=20$, (and $\bar{\Sigma}_0^2=0.5$).

$\alpha = .05$ $\alpha = .10$

148

$\alpha = .20$ $\alpha = .50$

Fig. 11 Distance picture of UMLE¹, RMLE², PTMLE³ and SMLE⁴
for $r=1$, $N=20$, (and $\bar{Y}=0.5$).

Fig12 Dominance picture of UNLE, RMLE, PTMLE and SMLE
for $p=5$, $N=20$, (and $\frac{\sigma}{\mu}=0.5$)

Table 9: Maximum relative efficiencies, of RMLE, PTMLE and SMLE and the intersecting efficiencies for PTMLE and SMLE for each α with corresponding Δ_n -values and each $\frac{\pi^2}{Q}$ -value: $p = 4$, $N = 16$.

$\alpha \sqrt{\frac{E_1}{E_2}}$	.10	.20	.30	.40	.50	.60	.70	.80	.90
.05 E_1	2.6000	4.2000	5.8000	7.4000	9.0000	10.6000	12.2000	13.8000	15.4000
E_2	2.1897	3.0350	3.7128	4.2296	4.6464	4.9897	5.2774	5.5219	5.7323
E_4	1.3520	1.4757	1.5388	1.5770	1.6027	1.6212	1.6351	1.6459	1.6545
$E_{A_{.05}}$	1.1794	1.2320	1.2571	1.2719	1.2816	1.2884	1.2935	1.2974	1.3006
$\Delta_{.05}$	2.3842	2.3842	2.3842	2.3842	2.3842	2.3842	2.3842	2.3842	2.3842
.10 E_1	2.6000	4.2000	5.8000	7.4000	9.0000	10.6000	12.2000	13.8000	15.4000
E_2	1.9390	2.4973	2.8674	3.1307	3.3277	3.4805	3.6026	3.7023	3.7854
E_4	1.3520	1.4757	1.5388	1.5770	1.6027	1.6212	1.6351	1.6459	1.6545
$E_{A_{.10}}$	1.1991	1.2588	1.2875	1.3044	1.3155	1.3234	1.3293	1.3338	1.3375
$\Delta_{.10}$	1.9573	1.9573	1.9573	1.9573	1.9573	1.9573	1.9573	1.9573	1.9573
.15 E_1	2.6000	4.2000	5.8000	7.4000	9.0000	10.6000	12.2000	13.8000	15.4000
E_2	1.7598	2.1480	2.3847	2.5433	2.6572	2.7440	2.8099	2.8635	2.9075
E_4	1.3520	1.4757	1.5388	1.5770	1.6027	1.6212	1.6351	1.6459	1.6545
$E_{A_{.15}}$	1.2133	1.2865	1.3191	1.3384	1.3511	1.3601	1.3668	1.3720	1.3762
$\Delta_{.15}$	1.5827	1.5827	1.5827	1.5827	1.5827	1.5827	1.5827	1.5827	1.5827
.20 E_1	2.6000	4.2000	5.8000	7.4000	9.0000	10.6000	12.2000	13.8000	15.4000
E_2	1.6237	1.9668	2.0685	2.1731	2.2403	2.2804	2.3420	2.3750	2.4018
E_4	1.3520	1.4757	1.5388	1.5770	1.6027	1.6212	1.6351	1.6459	1.6545
$E_{A_{.20}}$	1.2422	1.3182	1.3553	1.3774	1.3920	1.4023	1.4101	1.4161	1.4200
$\Delta_{.20}$	1.2190	1.2190	1.2190	1.2190	1.2190	1.2190	1.2190	1.2190	1.2190
.25 E_1	2.6000	4.2000	5.8000	7.4000	9.0000	10.6000	12.2000	13.8000	15.4000
E_2	1.5161	1.7284	1.8442	1.9171	1.9673	2.0048	2.0317	2.0536	2.0713
E_4	1.3520	1.4757	1.5388	1.5770	1.6027	1.6212	1.6351	1.6459	1.6545
$E_{A_{.25}}$	1.2694	1.3564	1.3994	1.4250	1.4421	1.4542	1.4633	1.4703	1.4760
$\Delta_{.25}$	.8474	.8474	.8474	.8474	.8474	.8474	.8474	.8474	.8474
.30 E_1	2.6000	4.2000	5.8000	7.4000	9.0000	10.6000	12.2000	13.8000	15.4000
E_2	1.4287	1.5912	1.6700	1.7293	1.7751	1.8069	1.8305	1.8458	1.8631
E_4	1.3520	1.4757	1.5388	1.5770	1.6027	1.6212	1.6351	1.6459	1.6545
$E_{A_{.30}}$	1.3035	1.4051	1.4559	1.4865	1.5068	1.5214	1.5323	1.5408	1.5476
$\Delta_{.30}$	.4551	.4551	.4551	.4551	.4551	.4551	.4551	.4551	.4551
.35 E_1	2.6000	4.2000	5.8000	7.4000	9.0000	10.6000	12.2000	13.8000	1

Finally, we have the following recommendations:

- 1) select $\hat{\theta}$ if $\Delta \in [0, p]$
- 2) select $\hat{\theta}^{PT}$ if $\Delta \in [0, \Delta_\alpha]$, otherwise select $\hat{\theta}^S$;

if no knowledge about Δ is available; then

- 1) select $\hat{\theta}^{PT}$ with minimum guaranteed efficiency to obtain larger efficiency than $\hat{\theta}^S$ for small α , otherwise,
- 2) select $\hat{\theta}^S$ which always has a larger efficiency than $\hat{\theta}$.

Note that to make use of $\hat{\theta}^S$, p should be greater than 2. Whenever $p \leq 2$, PTMLE is the only choice over UMLE. One may select PTMLE with minimum guaranteed efficiency and determine the size of the test with the help of the Tables from 14 to 17.

5.4. MEAN SQUARE ERROR MATRICES

This section is devoted to the derivation of the mean square error matrices of the four estimators. In the following sub-section properties of four estimators are studied via generalized mean square error.

Let $\Gamma_1(\hat{\theta})$, $\Gamma_2(\hat{\theta})$, $\Gamma_3(\hat{\theta}^{PT})$ and $\Gamma_4(\hat{\theta}^S)$ be the mean square matrices of $\hat{\theta}$, $\hat{\theta}$, $\hat{\theta}^{PT}$ and $\hat{\theta}^S$ respectively. Then the following theorem provides the mean square matrices.

THEOREM 5.4.1. The mean square error matrices of UMLE, RMLE, PTMLE, and SMLE are given by (5.3.45) – (5.3.48) respectively:

$$\Gamma_1(\hat{\theta}) = \left(\frac{1}{N} + \frac{\bar{x}^2}{Q} \right) \Sigma \quad (5.3.45)$$

$$\Gamma_2(\hat{\theta}) = \frac{1}{N} \Sigma + \bar{x}^2 \beta \beta' \quad (5.3.46)$$

$$\begin{aligned}\Gamma_3(\hat{\theta}^{PT}) &= \left(\frac{1}{N} + \frac{\bar{x}^2}{Q}\right) \Sigma - \frac{\bar{x}^2}{Q} H_{p+2,m}(T_{N,\alpha}^{*2}; \Delta) \Sigma \\ &\quad + \bar{x}^2 \beta \beta' \{2H_{p+2,m}(T_{N,\alpha}^{*2}; \Delta) - H_{p+4,m}(T_{N,\alpha}^{*2}; \Delta)\}\end{aligned}\quad (5.3.47)$$

$$\begin{aligned}\Gamma_4(\hat{\theta}^S) &= \left(\frac{1}{N} + \frac{\bar{x}^2}{Q}\right) \Sigma \\ &\quad - cm \frac{\bar{x}^2}{Q} [\{2E(\chi_{p+2}^{-2}(\Delta)) - c(m+2)E(\chi_{p+2}^{-4}(\Delta))\} \Sigma \\ &\quad - \{(m+2)c + 4\} Q \beta \beta' E(\chi_{p+4}^{-4}(\Delta))]\end{aligned}\quad (5.3.48)$$

PROOF. The $\Gamma_1(\hat{\theta})$ and $\Gamma_2(\hat{\theta})$ expressions are straightforward and we provide proof for the mean square error matrices for PTMLE and SMLE.

By definition,

$$\begin{aligned}\Gamma_3(\hat{\theta}^{PT}) &= E(\hat{\theta}^{PT} - \theta)(\hat{\theta}^{PT} - \theta)' \\ &= E\left[\{\bar{\theta} + \bar{\beta}\bar{x}I(T^2 < T_{N,\alpha}^2) - \theta\}\{\bar{\theta} + \bar{\beta}\bar{x}I(T^2 < T_{N,\alpha}^2) - \theta\}'\right] \\ &= E(\bar{\theta} - \theta)(\bar{\theta} - \theta)' + \bar{x}^2 E\{\bar{\beta}\bar{\beta}'I(T^2 < T_{N,\alpha}^2)\} \\ &\quad + 2\bar{x}E\{(\bar{\theta} - \theta)\bar{\beta}'I(T^2 < T_{N,\alpha}^2)\} \\ &= \left(\frac{1}{N} + \frac{\bar{x}^2}{Q}\right) \Sigma - \frac{\bar{x}^2}{Q} H_{p+2,m}(T_{N,\alpha}^{*2}; \Delta) \Sigma \\ &\quad - \bar{x}^2 \beta \beta' H_{p+4,m}(T_{N,\alpha}^{*2}; \Delta) + 2\bar{x}^2 \beta \beta' H_{p+2,m}(T_{N,\alpha}^{*2}; \Delta) \\ &= \left(\frac{1}{N} + \frac{\bar{x}^2}{Q}\right) \Sigma - \frac{\bar{x}^2}{Q} H_{p+2,m}(T_{N,\alpha}^{*2}; \Delta) \Sigma \\ &\quad + \bar{x}^2 \beta \beta' \{2H_{p+2,m}(T_{N,\alpha}^{*2}; \Delta) - H_{p+4,m}(T_{N,\alpha}^{*2}; \Delta)\}.\end{aligned}$$

Since

$$\begin{aligned}E\{(\bar{\beta}\bar{\beta}')I(T^2 < T_{N,\alpha}^2)\} \\ = \frac{1}{Q} \{H_{p+2,m}(T_{N,\alpha}^{*2}; \Delta) \Sigma + Q \beta \beta' H_{p+4,m}(T_{N,\alpha}^{*2}; \Delta)\}\end{aligned}$$

using Theorem A.3 from Appendix A, after transformation as in Section 5.3.

Also,

$$\begin{aligned} & E \{ (\bar{\theta} - \theta) \bar{\beta}' I(T^2 < T_{N,\alpha}^2) \} \\ &= E \{ \{ (\bar{y} - \theta - \beta \bar{x}) - \bar{x}(\bar{\beta} - \beta) \} \bar{\beta}' I(T^2 < T_{N,\alpha}^2) \} \\ &= -\bar{x} E \{ \{ \bar{\beta} \bar{\beta}' (T^2 < T_{N,\alpha}^2) \} - \beta \beta' I(T^2 < T_{N,\alpha}^2) \}, \end{aligned}$$

due to independence of $\bar{y}$ and $\bar{\beta}$ which simplifies the expression

$$\bar{x}^2 E \{ \bar{\beta} \bar{\beta}' I(T^2 < T_{N,\alpha}^2) \} + 2\bar{x} E \{ (\bar{\theta} - \theta) \bar{\beta}' I(T^2 < T_{N,\alpha}^2) \}$$

to

$$-\bar{x}^2 E \{ \bar{\beta} \bar{\beta}' I(T^2 < T_{N,\alpha}^2) \} + 2\bar{x} \beta E \{ \bar{\beta}' I(T^2 < T_{N,\alpha}^2) \}.$$

Hence, establishing (5.3.47). Now, consider

$$\begin{aligned} \Gamma_4(\hat{\theta}^S) &= E \{ (\hat{\theta}^S - \theta)(\hat{\theta}^S - \theta)' \} \\ &= E \{ \{ (\bar{\theta} - \theta) + c\bar{x}\bar{\beta}(N-2)T^{-2} \} \{ (\bar{\theta} - \theta) + c\bar{x}\bar{\beta}(N-2)T^{-2} \}' \} \\ &= E \{ (\bar{\theta} - \theta)(\bar{\theta} - \theta)' + c^2\bar{x}^2 E \{ (\bar{\beta} \bar{\beta}') (N-2)^2 T^{-4} \} \\ &\quad + 2c\bar{x} E \{ (\bar{\theta} - \theta) \bar{\beta}' (N-2) T^{-2} \} \} \\ &= \left(\frac{1}{N} + \frac{\bar{x}^2}{Q} \right) \Sigma + \frac{c^2\bar{x}^2}{Q} m(m+2) \{ E(\chi_{p+2}^{-4}(\Delta)) \Sigma + Q\beta\beta' E(\chi_{p+4}^{-4}(\Delta)) \} \\ &\quad - 2c\frac{\bar{x}^2}{Q} m \{ E(\chi_{p+2}^2(\Delta)) + Q\beta\beta' E(\chi_{p+4}^{-2}(\Delta)) \} \\ &\quad + 2c\frac{\bar{x}^2}{Q} Q\beta\beta' E(\chi_{p+2}^{-2}(\Delta)) \\ &= \left(\frac{1}{N} + \frac{\bar{x}^2}{Q} \right) \Sigma - cm\frac{\bar{x}^2}{Q} \{ 2E(\chi_{p+2}^{-2}(\Delta)) - c(m+2)E(\chi_{p+2}^{-4}(\Delta)) \} \Sigma \\ &\quad + c^2m(m+2)\bar{x}^2\beta\beta' E(\chi_{p+4}^{-4}(\Delta)) \\ &\quad + 4cm\bar{x}^2Q\beta\beta' E(\chi_{p+4}^{-4}(\Delta)) \end{aligned}$$

using the relation

$$E(\chi_q^{-2}(\Delta)) - E(\chi_{q+2}^{-4}(\Delta)) = 2E(\chi_{q+4}^{-4}(\Delta)).$$

Thus

$$\Gamma_4(\theta^S) = \left(\frac{1}{N} + \frac{\bar{x}^2}{Q} \right) \Sigma - cm \frac{\bar{x}^2}{Q} \left[\{ 2E(\chi_{p+2}^{-2}(\Delta)) - c(m+2)E(\chi_{p+2}^{-4}(\Delta)) \} \Sigma \right. \\ \left. - \{ c(m+2) + 4 \} Q\beta\beta'E(\chi_{p+4}^{-4}(\Delta)) \right],$$

since

$$E \{ \tilde{\beta}\tilde{\beta}'(N-2)T^{-4} \}, \text{ by transformation,} \\ = \frac{1}{Q} E \{ (D^{-1}y^*y^{*'}D^{-1})(y^{*'}y^*/\chi_m^2)^{-2} \} \\ = \frac{m}{Q} (m+2) \{ E(\chi_{p+2}^{-4}(\Delta))\Sigma + Q\beta\beta'E(\chi_{p+4}^{-4}(\Delta)) \},$$

by using Theorem A.3 from Appendix A.

Similarly,

$$\bar{x} E \{ (\tilde{\beta}\tilde{\beta}'(N-2)T^{-2})^4 \} \\ = \frac{\bar{x}}{2} E \{ (D^{-1}y^*y^{*'}D^{-1})(y^{*'}y^*/\chi_m^2)^{-1} \} \\ = \frac{\bar{x}^2}{Q} m \{ E(\chi_{p+2}^{-2}(\Delta))\Sigma + Q\beta\beta'E(\chi_{p+4}^{-2}(\Delta)) \}.$$

Also, using the fact that y is independent of $\tilde{\beta}$, we establish (5.3.48).

Now, we have established the following dominance picture of estimators under the null hypothesis.

THEOREM 5.4.2. Under the null hypothesis, $H_0: \beta = 0$,

$$\Gamma_1(\tilde{\theta}) - \Gamma_2(\tilde{\theta}), \quad \Gamma_1(\tilde{\theta}) - \Gamma_3(\tilde{\theta}^{PT}), \quad \Gamma_3(\tilde{\theta}^{PT}) - \Gamma_2(\tilde{\theta})$$

are positive definite, and

$$\Gamma_4(\tilde{\theta}^S) - \Gamma_3(\tilde{\theta}^{PT})$$

is positive definite, for a range of α .

PROOF.

$$\begin{aligned}\Gamma_1(\tilde{\theta}) - \Gamma_2(\tilde{\theta}) &= \frac{\bar{x}}{Q} \Sigma \\ \Gamma_1(\tilde{\theta}) - \Gamma_3(\tilde{\theta}^{PT}) &= \frac{\bar{x}^2}{Q} H_{p+2,m}(T_{N,\alpha}^{*2}; 0) \Sigma \\ \Gamma_3(\tilde{\theta}^{PT}) - \Gamma_2(\tilde{\theta}) &= \frac{\bar{x}^2}{Q} \{1 - H_{p+2,m}(T_{N,\alpha}^{*2}; 0)\} \Sigma. \\ \Gamma_4(\tilde{\theta}^S) - \Gamma_3(\tilde{\theta}^{PT}) &= \frac{\bar{x}^2}{Q} \left\{ H_{p+2,m}(T_{N,\alpha}^{*2}; 0) - \frac{m(p-2)}{p(m+2)} \right\} \Sigma,\end{aligned}$$

whenever $H_{p+2,m}(T_{N,\alpha}^{*2}; \Delta) > \frac{m(p-2)}{p(m+2)}$ depending upon α .

5.4.1. GENERALIZED EFFICIENCY ANALYSIS AND RECOMMENDATIONS

In this section we focus on the generalized efficiencies of $\hat{\theta}$, $\hat{\theta}^{PT}$ and $\hat{\theta}^S$ relative to $\tilde{\theta}$ and provide a max-min rule for the optimal choice of the level of significance for the preliminary test. The generalized relative efficiency (GRE) of $\hat{\theta}$ with respect to $\tilde{\theta}$ is

$$\begin{aligned}\text{GRE}_{11}(\hat{\theta} : \tilde{\theta}) &= \left| \left(\frac{1}{N} + \frac{\bar{x}^2}{Q} \right) \Sigma \right|^{\frac{1}{2}} / \left| \frac{1}{N} \Sigma + \bar{x}^2 \beta \beta' \right|^{\frac{1}{2}} \\ &= \left(1 + \frac{N\bar{x}^2}{Q} \right) \left(1 + \frac{N\bar{x}^2}{Q} \Delta \right)^{-\frac{1}{2}}.\end{aligned}$$

Thus, GRE_{11} is a constant for fixed Δ -value and at $\Delta = 0$, it equals $\left(1 + \frac{N\bar{x}^2}{Q} \right)$, while as $\Delta \rightarrow \infty$ it goes to 0, indicating that for large values of Δ GRE_{11} of $\hat{\theta}$ becomes low and $\tilde{\theta}$ becomes more efficient. In fact,

$\text{GRE}_{11}(\hat{\theta} : \tilde{\theta}) \geq 1$ according as

$$\Delta \geq \left\{ \left(1 + \frac{N\bar{x}^2}{Q} \right)^{\frac{1}{2}} - 1 \right\} \left(\frac{N\bar{x}^2}{Q} \right).$$

Next, we consider the GRE of $\hat{\theta}^{PT}$ relative to $\tilde{\theta}$ which is given by

$$\begin{aligned}\text{GRE}_{22}(\hat{\theta}^{PT} : \tilde{\theta}) &= \left| \left(\frac{1}{N} + \frac{\bar{x}^2}{Q} \right) \Sigma \right|^{\frac{1}{2}} / \left| \left(\frac{1}{N} + \frac{\bar{x}^2}{Q} \right) \Sigma - \frac{\bar{x}^2}{Q} H_{p+2,m}(T_{N,\alpha}^{*2}; \Delta) \right. \\ &\quad \left. + \bar{x} \beta \beta' \{ 2H_{p+2,m}(T_{N,\alpha}^{*2}; \Delta) - H_{p+4,m}(T_{N,\alpha}^{*2}; \Delta) \} \right|^{\frac{1}{2}} \\ &= \{ 1 + G_{22}(\Delta) \}^{-\frac{1}{2}},\end{aligned}$$

where

$$G_{22}(\Delta) = \Delta \frac{\bar{x}^2}{Q} \left(\frac{1}{N} + \frac{\bar{x}^2}{Q} \right)^{-1} \{ 2H_{p+2,m}(T_{N,\alpha}^{*2}; \Delta) - H_{p+4,m}(T_{N,\alpha}^{*2}; \Delta) \} \\ - \frac{\bar{x}^2}{Q} \left(\frac{1}{N} + \frac{\bar{x}^2}{Q} \right)^{-1} H_{p+2,m}(T_{N,\alpha}^{*2}; \Delta).$$

First, note that at $\Delta = 0$, and $\bar{x} \neq 0$ $\text{GRE}_{22}(\hat{\theta}^{PT}; \tilde{\theta})$ has its maximum (E^*).

$$E^* = \left\{ 1 - \frac{\bar{x}^2}{Q} \left(\frac{1}{N} + \frac{\bar{x}^2}{Q} \right)^{-1} H_{p+2,m}(T_{N,\alpha}^{*2}; 0) \right\}^{-\frac{1}{p}} (> 1)$$

if $\bar{x} \neq 0$ and $0 < \alpha < 1$. Also, as Δ goes to infinity, $\text{GRE}_{22}(\hat{\theta}^{PT}; \tilde{\theta})$ goes to 1. Further, we see that $\text{GRE}_{22}(\hat{\theta}^{PT}; \tilde{\theta})$ is a function of Δ (for fixed α) decreases to a minimum (E_0), crossing the efficiency 1, at $\Delta = \Delta_0$, then increases towards 1 as Δ increases to infinity. The efficiency is more than 1 if $0 \leq \Delta \leq p$, and for Δ outside this interval the efficiency decreases to a value at $\Delta = \Delta_0$, E_0 , then increases to 1 as $\Delta \rightarrow \infty$.

Now, consider GRE_{22} as a function of α and Δ for fixed N and $\frac{\bar{x}^2}{Q}$. The function $\text{GRE}_{22}(\alpha, \Delta)$, for $\alpha \neq 0$, has its maximum at $\Delta = 0$; it decreases as Δ increases, crossing the line $\text{GRE}_{22}(\alpha, \Delta) = 1$ to a minimum at $\Delta = \Delta_0$, then it increases to a maximum towards 1. When $\Delta = 0$ and α varies we have

$$\max_{\alpha} \text{GRE}_{22}(\alpha, 0) = \text{GRE}_{22}(0, 0) \\ = \left\{ 1 - \frac{\bar{x}^2}{Q} \left(\frac{1}{N} + \frac{\bar{x}^2}{Q} \right)^{-1} H_{p+2,m}(T_{N,\alpha}^{*2}; 0) \right\}^{-\frac{1}{p}}.$$

The value of $\text{GRE}_{22}(\alpha, 0)$ decreases as α increases. On the other hand, let $\alpha = 0$ and Δ vary; then the curve $\text{GRE}_{22}(0, \Delta)$ and $\text{GRE}_{22}(1, \Delta)$ will intersect in the interval $0 \leq \Delta \leq p$, and the value of Δ at the intersection decreases as Δ increases. Therefore, for two different α , $\text{GRE}_{22}(\alpha_1, \Delta)$ and $\text{GRE}_{22}(\alpha_2, \Delta)$ will never intersect above $\text{GRE}_{22}(\alpha, \Delta) = 1$; but they will intersect somewhere below $\text{GRE}_{22}(\alpha, \Delta) = 1$.

In order to choose an estimator with maximum relative efficiency we adopt the following procedure: If it is known that $0 \leq \Delta \leq p$, the estimator $\hat{\theta}$ is chosen since $\text{GRE}_{11}(\alpha, \Delta)$ is the largest for all Δ in this range. If $\Delta > p$, there is no way of choosing a uniformly best estimator. If we preassign a value of the efficiency, say E_0 , that we are willing to accept, then if $A = \{\alpha \mid \text{GRE}_{22}(\alpha, \Delta) \geq E_0, \text{ for all } \Delta\}$, then the estimator is chosen which maximizes $\text{GRE}_{22}(\alpha, \Delta)$ over all $\alpha \in A$ and Δ . Thus, we solve for α such that

$$\max_{\alpha \in A} \min_{\Delta} \text{GRE}_{22}(\alpha, \Delta) = E_0.$$

Tables 21 to 23 provide the values of maximum relative efficiency E^* , minimum relative efficiency, E_0 , and the recommended corresponding size of the preliminary test, α^* , for $p = 3(1)5$ and for various values of N . For example, if $N = 28$, $\bar{x}^2/Q = 0.70$, and $p = 3$ and the experimenter wishes to have an estimator with a minimum guaranteed efficiency of 0.70, then Table 21 suggests that $\alpha^* = 0.10$. Because such a choice of α^* would yield him to enjoy an estimator with a maximum efficiency of 1.4990 at $\Delta = 0$, and with a minimum (E_0) guaranteed efficiency of 0.7001.

3 of/de 3

Finally, the generalized relative efficiency of SMLE relative to UMLE is given by

$$\begin{aligned} \text{GRE}_{33}(\hat{\theta}^S : \tilde{\theta}) &= \left| \left(\frac{1}{N} + \frac{\bar{x}^2}{Q} \right) \Sigma \right|^{\frac{1}{p}} / \left| \left(\frac{1}{N} + \frac{\bar{x}^2}{Q} \right) \Sigma - cm \frac{\bar{x}^2}{Q} \{ 2E(\chi_{p+2}^{-2}(\Delta)) \right. \right. \\ &\quad \left. \left. - c(m+2)E(\chi_{p+2}^{-4}(\Delta)) \} \Sigma - \{c(m+2)+4\} \Delta E(\chi_{p+4}^{-4}(\Delta)) \right|^{\frac{1}{p}} \\ &= \{1 - G_{33}(\Delta)\}^{-\frac{1}{p}}, \end{aligned}$$

where

$$G_{33}(\Delta) = cm \frac{\bar{x}^2}{Q} \left(\frac{1}{N} + \frac{\bar{x}^2}{Q} \right)^{-1} \left[\{2E(\chi_{p+2}^{-2}(\Delta)) - c(m+2)E(\chi_{p+2}^{-4}(\Delta))\} - \{c(m+2)+4\} \right]$$

It may be noted that $\text{GRE}_{33}(\hat{\theta}^S : \tilde{\theta})$ is a decreasing function of Δ and at $\Delta = 0$ it assumes the value

$$\left[1 - \frac{c\bar{x}^2 m}{Q} \left(\frac{1}{N} + \frac{\bar{x}^2}{Q} \right)^{-1} \left\{ \frac{2}{p} - \frac{c(m+2)}{p(p-2)} \right\} \right]^{-\frac{1}{p}},$$

$0 < c < \frac{2(p-2)}{m+2}$, and as $\Delta \rightarrow \infty$, $\text{GRE}_{33}(\hat{\theta}^S, \tilde{\theta})$ tends to 1. Thus, $\hat{\theta}^S$ is uniformly efficient than $\tilde{\theta}$. Finally, consider the intersection of $\text{GRE}_{22}(\hat{\theta}^{PT} : \tilde{\theta})$ and $\text{GRE}_{33}(\hat{\theta}^S : \tilde{\theta})$. Except for large α -values, they will intersect at $\Delta = \Delta_\alpha$ for fixed N, p and $\frac{\bar{x}^2}{Q}$ and for $\Delta \in (0, \Delta_\alpha)$, $\hat{\theta}^{PT}$ is more efficient than $\hat{\theta}^S$, but outside this interval $\hat{\theta}^S$ dominates $\hat{\theta}^{PT}$. Thus, none of the estimators $\hat{\theta}^{PT}$ and $\hat{\theta}^S$ is efficient over each other in the interval $(0, \infty)$.

5.5. CONCLUDING REMARKS

Note that we reach the same conclusion about the performances of the estimators using risk analysis as well as mean square error matrices. If $p \leq 2$, the only choice left is the use of preliminary test estimator with decision rule provided for the choice of size of the test. However, if Δ is unknown and $p \geq 3$, we have the table of maximum and minimum efficiencies at the intersection $\Delta = \Delta_\alpha$ for the decision of an appropriate estimator ($\hat{\theta}^{PT}$ or $\hat{\theta}^S$). In general shrinkage maximum likelihood estimator is preferred to the preliminary test estimator for $p \geq 3$.

CHAPTER 6

CONCLUSIONS AND AREAS OF FURTHER STUDY

This thesis has studied the classical problem of pooling means in correlated and uncorrelated cases alongwith the estimation of the intercept parameter in the presence of uncertain prior information about the slope parameter, for univariate and multivariate normal models.

We have developed four estimation strategies for estimating a parameter in the presence of uncertain information as follows:

1. Shrinkage maximum likelihood estimator(SMLE)
2. Preliminary test maximum likelihood estimator(PTMLE)
3. Restricted maximum likelihood estimator(RMLE)
4. Unrestricted maximum likelihood estimator(UMLE)

Two measures of relative efficiency were considered, one based on the quadratic loss function, and the other based on the determinant of mean square error matrix. A max-min rule for the size of the preliminary test is presented. Tables for an appropriate choice of the level of the significance are provided. The range in the parameter space where PTMLE dominates SMLE is also investigated theoretically and computationally.

We make the following conclusions:

1. SMLE is more efficient than UMLE in the whole parameter space.
2. If the hypothesis error is correctly specified, i.e, $\Delta=0$, or in the neighbourhood of the null hypothesis(pivot), then the PTMLE is more efficient than the SMLE and

the UMLE, but for larger values of α , the level of the statistical significance, SMLE dominates PTMLE uniformly. Thus, we have plotted efficiency function of SMLE, PTMLE, and RMLE relative to UMLE versus Δ to observe such properties.

3. As the hypothesis error grows, the PTMLE becomes more inefficient than the UMLE. Indeed, the value of the efficiency function of PTMLE monotonically decreases, achieves a minimum after crossing the UMLE efficiency and then monotonically increases, to approach the UMLE efficiency. We have plotted the efficiency function of the PTMLE relative to the UMLE versus Δ for $\alpha=0.05, 0.1, 0.2$ and 0.5 and for various sample sizes to exhibit such features of the PTMLE.
4. The RMLE is more efficient than the other three estimators under the null hypothesis and for a range of the parameter space; but as the hypothesis departs from the null hypothesis, its efficiency converges to zero (unbounded risk).

The decision on whether to use the UMLE or the PTMLE or the SMLE rests on the experimenter. The following steps should be used to select an estimator:

1. If the hypothesis is true i.e., $\Delta=0$, then select the RMLE because it has the largest efficiency as compared to the other three estimators. Furthermore, if $\Delta \in [0, p]$ and $p < 3$, then also select the RMLE.
2. If $\Delta \in (p, \Delta_\alpha]$, then choose PTMLE because in this range PTMLE has the the largest efficiency as compared to the other three estimators. The plots of efficiency functions can determine the value of Δ_α .
3. If Δ is unknown and the experimenter is willing to accept a minimum guaranteed efficiency less than one, then the PTMLE should be used. A maximum efficiency is attained at $\Delta = 0$ and is greater than the SMLE and the UMLE.

4. If the experimenter is not willing to accept a minimum guaranteed efficiency and also has no knowledge about Δ , which is generally the case, then the SMLE should be used because it dominates the UMLE uniformly and the PTMLE for a range of α .

It should be noted that the SMLE can only be used if $p \geq 3$. Therefore, for $p \leq 3$, we recommend steps (1), (3) and (5) follow:

5. If the experimenter is not willing to accept a minimum guaranteed efficiency and has no knowledge about Δ , and selects the UMLE, then this may not be the best choice since other information may be ignored.

The PTMLE can be uniformly improved by using SMLE instead of UMLE. That is, a convex combination of RMLE and SMLE will improve the PTMLE. Also, SMLE can be improved uniformly by the positive-rule of the James-Stein rule. In the two-sample problem, a positive rule estimator may constructed as:

$$\hat{\mu}^{(1)S^+} = \hat{\mu}^{(1)} + \left(1 - \frac{2(p-2)}{(m+2)} nT^{-2}\right)^+ (\bar{\mu}^{(1)} - \hat{\mu}^{(1)})$$

where

$$Z^+ = \begin{cases} z, & \text{if } z > 0; \\ 0, & \text{otherwise.} \end{cases}$$

This will control the over shrinking problem inherent in the SMLE; the estimators under the null hypothesis may be ordered as follows:

$$\hat{\mu}^{(1)S^+} > \hat{\mu}^{(1)S} > \hat{\mu}^{(1)PT}, \quad \text{for } \Delta \in (\Delta_\alpha, \infty)$$

Thus, one may study the improved PTMLE and positive-rule shrinkage maximum likelihood estimator (PSMLE) and determine the level of significance for the use of the improved PTMLE.

A natural extension of pooling mean vector in two-sample problem can be worked out in multi-sample pooling problem as well. Similarly, the estimation of the location parameter vector in a simple linear multivariate model may be extended to a multi-sample problem. Then, one would be involved with test-statistic such as a Wilk's likelihood ratio test or Roy's largest root criterion for the SE and PTE. However, this will require asymptotic theory to obtain similar conclusions as in Chapter 3 and 5.

Chapter 4, containing material on the estimation of component-means, may be extended, in various ways. For example, one may try to remove the condition on the covariance matrices. Similarly, one may consider the estimation of $\mu^{(1)}$ when it is suspected that $\mu^{(2)} = \mu_0^{(2)}$.

The area of regression analysis is wide open; more investigation is needed. The application of the preliminary and shrinkage estimation methodology to survey sampling, design of experiments and categorical data analysis is another area of investigation. Some scattered results are available in the literature, but we need unification.

A difficult area which has not been investigated is the confidence interval problem, where PSMLE has the immediate prospect. Another application oriented problem is the estimation of the risk/covariance matrices of the proposed estimators.

APPENDIX A

THEOREM A.1. If the $(J \times 1)$ vector w is distributed normally with mean θ vector and covariance I then

$$E[\phi(w'w)w] = \theta E[\phi(\chi^2_{(J+2, \theta'\theta/2)})]. \quad (A.1.1)$$

THEOREM A.2. If w is a $(J \times 1)$ normally distributed random vector with mean θ vector and covariance I and A is any positive definite symmetric matrix, then

$$E[\phi(w'w)w'Aw] = E[\phi(\chi^2_{(J+2, \theta'\theta/2)})] \text{tr } A + E[\phi(\chi^2_{(J+4, \theta'\theta/2)})]\theta'\theta. \quad (A.1.2)$$

THEOREM A.3. If the $(J \times 1)$ vector w is distributed normally with mean θ vector and covariance matrix I then

$$E[\phi(w'w)ww'] = E[\phi(\chi^2_{(J+2, \lambda)})]I + E[\phi(\chi^2_{(J+4, \lambda)})]\theta\theta'$$

$$\text{where } \lambda = \theta\theta'. \quad (A.1.3)$$

One may refer to Judge and Bock (1978) for the proofs of all of these theorems.

APPENDIX B

Table: 1 Maximum and minimum guaranteed efficiencies.

$\alpha \backslash (N_1, N_2)$	(3, 4)	(7, 8)	(11, 12)	(13, 14)	(15, 16)
E^*					
E_0					
Δ_0					

E^*
 E_0
 Δ_0

Table: 3 Maximum and minimum guaranteed efficiencies for $p = 2$.

$\alpha \backslash (N_1, N_2)$		(3, 4)	(7, 8)	(11, 12)	(13, 14)	(15, 16)
.05	$\frac{E^*}{E_0}$	1.19123	1.19412	1.19727	1.19912	1.19998
	Δ_0	12.117	2.9213	2.9213	2.9213	2.9213
.10	$\frac{E^*}{E_0}$	1.19123	1.19412	1.19727	1.19912	1.19998
	Δ_0	12.117	2.9213	2.9213	2.9213	2.9213
.15	$\frac{E^*}{E_0}$	1.19123	1.19412	1.19727	1.19912	1.19998
	Δ_0	12.117	2.9213	2.9213	2.9213	2.9213
.20	$\frac{E^*}{E_0}$	1.19123	1.19412	1.19727	1.19912	1.19998
	Δ_0	12.117	2.9213	2.9213	2.9213	2.9213
.25	$\frac{E^*}{E_0}$	1.19123	1.19412	1.19727	1.19912	1.19998
	Δ_0	12.117	2.9213	2.9213	2.9213	2.9213
.30	$\frac{E^*}{E_0}$	1.19123	1.19412	1.19727	1.19912	1.19998
	Δ_0	12.117	2.9213	2.9213	2.9213	2.9213
.35	$\frac{E^*}{E_0}$	1.19123	1.19412	1.19727	1.19912	1.19998
	Δ_0	12.117	2.9213	2.9213	2.9213	2.9213
.40	$\frac{E^*}{E_0}$	1.19123	1.19412	1.19727	1.19912	1.19998
	Δ_0	12.117	2.9213	2.9213	2.9213	2.9213
.45	$\frac{E^*}{E_0}$	1.19123	1.19412	1.19727	1.19912	1.19998
	Δ_0	12.117	2.9213	2.9213	2.9213	2.9213
.50	$\frac{E^*}{E_0}$	1.19123	1.19412	1.19727	1.19912	1.19998
	Δ_0	12.117	2.9213	2.9213	2.9213	2.9213

Table: 4 Maximum and minimum guaranteed efficiencies for $p=3$.

$\alpha \backslash (N_1, N_2)$	(3, 4)	(7, 8)	(11, 12)	(13, 14)	(15, 16)
0.1	E^* 2.0499 E_0 1.4503 Δ_0 10.5259	1.3574 1.6600 10.3021	1.0000 0.7018 0.3018	1.0000 0.7018 0.3018	1.0000 0.7018 0.3018
1.	1.8494 0.1311 14.0516	1.6747 0.7519 8.4949	1.4334 0.8192 7.0000	1.0000 0.7018 0.3018	1.0000 0.7018 0.3018
15.	1.6967 0.6646 17.4027	1.5427 0.8774 7.0000	1.1111 0.7018 0.3018	1.0000 0.7018 0.3018	1.0000 0.7018 0.3018
20.	1.6707 0.6646 18.392	1.4707 0.8774 6.8392	1.1111 0.7018 0.3018	1.0000 0.7018 0.3018	1.0000 0.7018 0.3018
25.	1.4707 0.6646 18.392	1.4707 0.8774 6.8392	1.1111 0.7018 0.3018	1.0000 0.7018 0.3018	1.0000 0.7018 0.3018
30.	1.4707 0.6646 18.392	1.4707 0.8774 6.8392	1.1111 0.7018 0.3018	1.0000 0.7018 0.3018	1.0000 0.7018 0.3018
35.	1.3333 0.8774 5.7600	1.2547 0.8774 5.7600	1.1111 0.7018 0.3018	1.0000 0.7018 0.3018	1.0000 0.7018 0.3018
40.	1.2732 0.8774 6.7669	1.2100 0.8774 6.4340	1.1111 0.7018 0.3018	1.0000 0.7018 0.3018	1.0000 0.7018 0.3018
45.	1.2732 0.8774 6.7669	1.1700 0.8774 6.2000	1.1111 0.7018 0.3018	1.0000 0.7018 0.3018	1.0000 0.7018 0.3018
50.	1.1833 0.8774 6.9500	1.1300 0.8774 6.1500	1.1111 0.7018 0.3018	1.0000 0.7018 0.3018	1.0000 0.7018 0.3018

Table: 3 Maximum and minimum guaranteed efficiencies for $p = 4$.

$\alpha \backslash (N_1, N_2)$	(3, 4)	(7, 8)	(11, 12)	(13, 14)	(15, 16)
0.05	$\frac{E^*}{E_0}$ Δ_0 2.8924 1.3753 34.034	1.3896 1.3877 12.500	1.3911 1.3907 10.722	1.391 1.391 10.468	1.391 1.391 10.11
0.10	1.810 1.814 23.695	1.7775 1.7767 10.500	1.7744 1.7741 7.744	1.7747 1.7741 7.744	1.7748 1.7741 7.744
0.15	1.770 1.768 13.110	1.7676 1.7673 6.110	1.7652 1.7641 6.110	1.764 1.764 6.110	1.764 1.764 6.110
0.20	1.731 1.732 10.645	1.7320 1.7317 6.415	1.7327 1.7320 6.415	1.7317 1.7317 6.415	1.7317 1.7317 6.415
0.25	1.711 1.712 10.117	1.7117 1.7117 6.117	1.7117 1.7117 6.117	1.7117 1.7117 6.117	1.7117 1.7117 6.117
0.30	1.691 1.691 1.691	1.691 1.691 7.111	1.691 1.691 7.111	1.691 1.691 7.111	1.691 1.691 7.111
0.35	1.6707 1.6709 10.1309	1.6704 1.6701 7.0013	1.6707 1.6703 6.943	1.6707 1.6707 6.943	1.6707 1.6707 6.943
0.40	1.6531 1.653 7.6594	1.654 1.654 6.7495	1.6531 1.6531 6.531	1.6531 1.6531 6.531	1.6531 1.6531 6.531
0.45	1.6326 1.631 9.9743	1.6351 1.6330 6.4541	1.6326 1.6301 6.0009	1.6326 1.6317 6.0009	1.6326 1.6317 6.0009
0.50	1.6111 1.6095 8.2957	1.6191 1.6131 6.1921	1.615 1.6107 5.8747	1.6152 1.6107 5.7977	1.6152 1.6107 5.7977

Table: 6 Maximum and minimum guaranteed efficiencies for $p = 5$.

$\alpha \backslash (N_1, N_2)$	(3, 4)	(7, 8)	(11, 12)	(13, 14)	(15, 16)
E^*					
E_0	2.1100	1.9100	1.7100	1.5100	1.3100
Δ_0	0.0000	0.0000	0.0000	0.0000	0.0000
0.05	1.7100	1.5100	1.3100	1.1100	0.9100
0.10	1.5100	1.3100	1.1100	0.9100	0.7100
0.15	1.3100	1.1100	0.9100	0.7100	0.5100
0.20	1.1100	0.9100	0.7100	0.5100	0.3100
0.25	0.9100	0.7100	0.5100	0.3100	0.1100
0.30	0.7100	0.5100	0.3100	0.1100	0.0000
0.35	0.5100	0.3100	0.1100	0.0000	
0.40	0.3100	0.1100	0.0000		
0.45	0.1100	0.0000			
0.50	0.0000				
0.55					
0.60					
0.65					
0.70					
0.75					
0.80					
0.85					
0.90					
0.95					
1.00					

REFERENCES

The following abbreviations of journals are used:

ASA Proc. - American Statistical Association Proceedings.

Ann. Statist. - Annals of Statistics.

Ann. Inst. Stat. Math. - Annals of the Institute of Statistical Mathematics.

Ann. Math. Stat. - The Annals of Mathematical Statistics.

Bull. Math. Stat. - Bulletin of Mathematical Statistics.

Comm. Statist. - Communications in Statistics.

JASA - Journal of the American Statistical Association.

JRSS - Journal of the Royal Statistical Society.

Ahsanullah, M. (1971), "On the estimation of means in a bivariate normal distribution with equal marginal variances". *Biometrika*, 58, 230-233.

Ahsanullah, M. and A. K. Md. Ehsanes Saleh (1972), "Estimation of intercept in a linear regression model with one dependent variable after a preliminary test on the regression coefficient". *International Statistical Review*, 40, 139-145.

Al-Bayyati, H. A. and J. C. Arnold (1972), "On double-stage estimation in simple linear regression using prior knowledge". *Technometrics*, 14, 405-414.

Anderson, T. W. (1984), *An Introduction to Multivariate Analysis*. New York: Wiley.

Arnold, R. S. (1981), *The Theory of linear Models and Multivariate Analysis*. New York: Wiley.

Arnold, B. C. (1970), "An alternative derivation of a result due to Srivastava and Bancroft". *JRSS B*, 32, 265-267.

- Asano, C. (1960a), "Estimations after preliminary test of significance and their applications to biometrical researches". *Bull. Math. Stat.*, 9(2-3), 1-24.
- Asano, C. (1960b), "Tests due to pooling data through preliminary test on biological direct assay". *Bull. Math. Stat.*, 9(2-3), 25-39.
- Asano, C. (1961), "Some considerations on the combination of estimates from different biological assays". *Bull. Math. Stat.*, 10(1-2), 17-32.
- Asano, C. and M. Sugimura (1961), "Some considerations on estimation of population variance due to the use of pooling data". *Bull. Math. Stat.*, 10(1-2), 33-44.
- Asano, C. and S. Sato (1962), "A bivariate analogue of pooling of data". *Bull. Math. Stat.*, 10 (3-4), 39-59.
- Bancroft, T. A. (1944), "On biases in estimation due to the use of preliminary tests of significance." *Ann. Math. Stat.* 15, 190-204.
- Bancroft, T. A. (1964), "Analysis and inference for incompletely specified models involving the use of preliminary test(s) of significance". *Biometrics*, 20, 427-442.
- Bancroft, T. A. (1968), *Topics in Intermediate Statistical Methods*. Volume 1. Ames: Iowa State University Press.
- Bancroft, T. A., and C. P. Han (1976), "On pooling mean in multivariate normal distributions." *Essays in Probability and Statistics*, Tokyo, Japan: Ogawa. Volume (Ikeda et al editors). 353-366.
- Bancroft, T. A., and C. P. Han (1977), "Inference based on conditional specification. A note and bibliography". *Int. Inst. Statist. Reviews* 45, 117-127.
- Bechhofer, R. E. (1951), "The effect of preliminary tests of significance on the size and power of certain tests of univariate linear hypotheses". Unpubl. Ph.D. thesis, Columbia University.

Bennett, B. M. (1952), "Estimation of means on the basis of preliminary tests of significance." *Ann. Inst. Stat. Math.*, 4, 31-43.

Bennett, B. M. (1956), "On the use of preliminary tests in certain statistical procedures". *Ann. Inst. Stat. Math.*, 8, 45-57.

Berger, J. O. (1980), "A robust generalized Bayes estimation and confidence region for multivariate normal mean". *Ann. Statist.* 8, 716-761.

Berger, J. O. (1985), "Statistical Decision Theory". 2nd edition. New York: Springer-Verlag.

Bock, M.E., G.G. Judge and T.A. Yancey (1973a), "Some comments on estimation in regression after preliminary tests of significance". *Journal of Econometrics* 1, 191-200.

Bock, M.E., T.A. Yancey and G.G. Judge (1973b), "The statistical consequence of preliminary test estimators in regression". *JASA*, 68, 109-116.

Bozivich, H., T.A. Bancroft and H.O. Hartley (1956), "Power of analysis of variance test procedures for certain incompletely specified models, I". *Ann. Math. Stat.*, 27, 1017-1043.

Cochran, W.G. (1970), "Performance of a preliminary test of comparability in observational studies". Technical Report 29. Office of Naval Research, Statistics and Probability Program.

Grimes, J.E. and B.V. Sukhatme (1980), "A regression type estimator based on preliminary test of significance". *JASA*, 75, 957-962.

Han, C.-P. (1973a), "Double sampling with partial information on auxiliary variables". *JASA*, 68, 914-918.

Han, C.-P. (1973b), "Regression estimation for bivariate normal distributions". *Ann. Inst. Stat. Math.*, 25, 335-344.

- Han, C.-P., and T.A. Bancroft (1968), "On pooling means when variance is unknown". *JASA*, **63**, 1332-1342.
- Huntsberger, D.V. (1955), "A generalization of a preliminary test procedure for pooling data". *Ann. Math. Stat.*, **26**, 734-743.
- James, W., and G. Stein (1961), "Estimation with quadratic loss". *Proceedings of the Fourth Berkeley Symposium on Mathematical Statistics and probability* (UCLA Press, Berkeley, California), 361-379.
- Judge, G.G., T.A. Yancey and M.E. Bock (1973), "Properties of estimators after preliminary tests of significance when stochastic restrictions are used in regression". *Journal of Econometrics*, **1**, 29-47.
- Judge, G. G., and M. E. Bock (1978), "The statistical implication of pre-test and Stein-rule estimators in econometrics". New York: North-Holland.
- Kale, B.K. and T.A. Bancroft (1967), "Inference for some incompletely specified models involving normal approximations to discrete data". *Biometrics*, **23**, 335-348.
- Kennedy, W.J. and T.A. Bancroft (1971), "Model building for prediction in regression based upon repeated significance tests". *Ann. Math. Stat.*, **42**, 1273-1284.
- Kitagawa, T. (1951), "Successive process of statistical inferences". *Memoir of the Faculty of Science, Kyushu University, Ser. A*, **5**, 139-180.
- Kitagawa, T. (1959), "Successive process of statistical inferences applied to linear regression analysis and its specializations to response surface analysis". *Bull. Math. Stat.*, **8**(3-4), 80-114.
- Kitagawa, T. (1963), "Estimation after preliminary test of significance". *Univ. of California Publication in Statistics*, Vol. B, 147-186.

- Larson, H.J. and T.A. Bancroft (1963a), "Sequential model building for prediction in regression analysis, I". *Ann. Math. Stat.*, **34**, 462-479.
- Larson, H. J., and T. A. Bancroft (1963b); "Biases in prediction by regression for certain incompletely specified models". *Biometrika* **50**, 391-402.
- Leamer, E.E. and G. Chamberlain (1976), "A Bayesian interpretation of pretesting". *JRSS*, **38**, 85-94.
- Mehta, J.S. and J. Gurland (1969), "Testing equality of means in the presence of correlation". *Biometrika*, **56**, 119-126.
- Mosteller, F. (1948), "On pooling data". *JASA* **43**, 231-242.
- Paul, A.E. (1950), "On a preliminary test for pooling mean squares in the analysis of variance". *Ann. Math. Stat.*, **21**, 539-556.
- Rahim, M.A. (1976), "Pre-test estimation of simple linear regression model using prior knowledge of one of the parameters". *Technometrics*, **18**, 439-445.
- Saleh, A.K. Md. Ehsanes (1973), "Pretest estimator of means of multivariate normal distribution". *Carleton mathematical series*, No. 94.
- Saleh, A.K. Md. Ehsanes and Sen, P.K. (1978), "Non parametric estimation of location parameter after a preliminary test on regression". *Ann. Statist.*, **6**, 154-168.
- Saleh, A.K. Md. Ehsanes and Sen, P.K. (1982a), "Nonparametric tests for location after preliminary test on regression". *Comm. Statist.*, **11(6)**, 639-652.
- Saleh, A.K. Md. Ehsanes and Sen, P.K. (1982b), "Nonparametric estimation following a preliminary test on regression (a collection of four papers)". *Carleton Mathematical Lecture Notes* No. 41.

- Saleh, A.K. Md. Ehsanes and Sen, P.K. (1983a), "Nonparametric tests for location after a preliminary test on regression in the multivariate case". *Comm. Statist.*, 12(16), 1855-1872.
- Saleh, A.K. Md. Ehsanes and Sen, P.K. (1983b), "Asymptotic properties of tests of hypothesis following a preliminary test". *Statistics and Decisions*, 1, 455-477.
- Saleh, A.K. Md. Ehsanes and Sen, P.K. (1985a), "Nonparametric shrinkage estimation in a parallelism problem". *Sankhyā*, 47, Series A, 156-165.
- Saleh, A.K. Md. Ehsanes and Sen, P.K. (1985b), "On Shrinkage M-estimators of location parameters", *Comm. Statist.*, 14(10), 2313-2329.
- Saleh, A.K. Md. Ehsanes and Sen, P.K. (1986a), "On Shrinkage least square estimation in a parallelism problem". *Comm. Statist.*, 15(5), 1451-1466.
- Saleh, A.K. Md. Ehsanes and Sen, P.K. (1986b), "On Shrinkage R-estimation in a multiple regression model". *Comm. Statist.*, 15(7), 2229-2244.
- Sato, S. (1962), "A multivariate analogue of pooling of data". *Bull. Math. Stat.*, 10(3-4), 61-76.
- Sclove, S.L. (1968), "Improved estimators for coefficients in linear regression". *JASA*, 63, 596-606.
- Sen, P.K. and Saleh, A.K. Md. Ehsanes (1979), "Nonparametric estimation of location parameter after a preliminary test on regression in multivariate case". *Journal of Multivariate Analysis*, 9, 322-331.
- Sen, P.K. and Saleh, A.K. Md. Ehsanes (1983), "Nonparametric shrinkage estimators of location in a multivariate simple regression model". *Proceeding of 4th Pannonian symposium on Math. Stats., Austria*, 311-323.

- Sen, P.K. and Saleh, A.K. Md. Ehsanes (1985), "On some shrinkage estimators of multivariate location". *Ann. Statist.*, 13, 272-281.
- Shirafuji, M. (1959), "A two stage sequential design in response surface analysis". *Bull. Math. Stat.*, 8(3-4), 115-126.
- Sisodia, B.V.S. and A.K. Srivastava (1982), "Modifying regression estimator with a preliminary test in double sampling". *Sankhyā*, 44 Series B, 295-303.
- Srivastava, S.R. (1964), "Power of analysis of variance test procedure on the basis of preliminary tests of significance". *Journal of Indian Statistical Association*, 2, 137-153.
- Srivastava, S.R. (1969), "A note on the estimation of a correlation coefficient incorporating a preliminary test of significance". *Trabajos de Estadística*, 20(1), 113-115.
- Srivastava, S.R. (1972), "Pooling mean squares in ANOVA Model II". *JASA*, 67, 676-679.
- Srivastava, S.R. and H. Bozovich (1962), "Power of certain analysis of variance test procedures involving preliminary tests". *Bulletin of the International Statistical Institute, Proceedings of the 33rd Session*, 39(3), 133-143.
- Srivastava, S.R. and T.A. Bancroft (1967), "Inferences concerning a population correlation coefficient from one or possibly two samples subsequent to a preliminary test of significance". *JRSS B*, 29, 282-291.
- Stein, C. (1956), "Inadmissibility of the usual estimator for the mean of a multivariate normal distribution". *Proceedings of the Third Berkeley Symposium on Mathematical Statistics and Probability*, Berkeley and Los Angeles, Univ. of California Press, Vol 1, 197-206.
- Sukhatme, B.V. and V.K.T. Tang (1970), "Allocation in stratified sampling based on preliminary test of significance". *ASA Proc., Social Statistics Section*, 368-371.

- Sukhatme, B.V. and V.K.T. Tang (1975), "Allocation in stratified sampling subsequent to preliminary test of significance". *JASA*, 70, 175-179.
- Tamura, R. (1965a), "Nonparametric inferences with a preliminary test". *Bull. Math. Stat.*, 11(3-4), 39-61.
- Tamura, R. (1965b), "Some estimate procedures with a nonparametric preliminary test I". *Bull. Math. Stat.*, 11(3-4), 63-71.
- Tamura, R. (1967a), "Some estimate procedures with a nonparametric preliminary test II". *Bull. Math. Stat.*, 12(3-4), 47-59.
- Tamura, R. (1967b), "Some weighted estimates in the nonparametric case". *Bull. Math. Stat.*, 12(3-4), 61-73.
- Tamura, R. (1967c), "Small sample properties of a "sometimes pool estimate" in the nonparametric case". *Bull. Math. Stat.*, 12(3-4), 75-83.
- Tang, V.K.T. and B.V. Sukhatme (1972a), "Use of preliminary test of significance for allocation in stratified sampling". *Journal of Indian Statistical Association*, 10, 100-114.
- Tang, V.K.T. and B.V. Sukhatme (1972b), "Allocation in stratified sampling based on preliminary test of significance II". *ASA Proc., Social Statistics Section*, 442-445.

END

2006-88

FIN