

The copyright © of this thesis belongs to its rightful author and/or other copyright owner. Copies can be accessed and downloaded for non-commercial or learning purposes without any charge and permission. The thesis cannot be reproduced or quoted as a whole without the permission from its rightful owner. No alteration or changes in format is allowed without permission from its rightful owner.



**ROBUST HIGH DIMENSIONAL M-TEST USING
REGULARIZED GEOMETRIC MEDIAN COVARIANCE**

ALO OLUSEGUN KEHINDE



**DOCTOR OF PHILOSOPHY
UNIVERSITI UTARA MALAYSIA
2018**



Awang Had Salleh
Graduate School
of Arts And Sciences

Universiti Utara Malaysia

PERAKUAN KERJA TESIS / DISERTASI
(Certification of thesis / dissertation)

Kami, yang bertandatangan, memperakukan bahawa
(We, the undersigned, certify that)

ALO OLUSEGUN KEHINDE

calon untuk Ijazah
(candidate for the degree of)

PhD

telah mengemukakan tesis / disertasi yang bertajuk:
(has presented his/her thesis / dissertation of the following title):

**"ROBUST HIGH DIMENSIONAL M-TEST USING REGULARIZED GEOMETRIC MEDIAN
COVARIANCE"**

seperti yang tercatat di muka surat tajuk dan kulit tesis / disertasi.
(as it appears on the title page and front cover of the thesis / dissertation).

Bahawa tesis/disertasi tersebut boleh diterima dari segi bentuk serta kandungan dan meliputi bidang ilmu dengan memuaskan, sebagaimana yang ditunjukkan oleh calon dalam ujian lisan yang diadakan pada : **04 November 2018.**

That the said thesis/dissertation is acceptable in form and content and displays a satisfactory knowledge of the field of study as demonstrated by the candidate through an oral examination held on:
November 04, 2018.

Pengerusi Viva:
(Chairman for VIVA)

**Prof. Dr. Engku Muhammad Nazri
Engku Abu Bakar**

Tandatangan
(Signature)

Pemeriksa Luar:
(External Examiner)

Prof. Dr. Azme Khamis

Tandatangan
(Signature)

Pemeriksa Dalam:
(Internal Examiner)

Dr. Nor Aishah Ahad

Tandatangan
(Signature)

Nama Penyelia/Penyelia-penyelia:
(Name of Supervisor/Supervisors)

Dr. Shamshuritawati Sharif

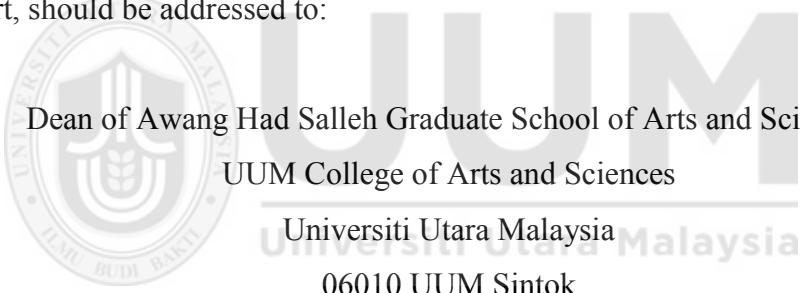
Tandatangan
(Signature)

Tarikh:
(Date) **November 04, 2018**

Permission to Use

In presenting this thesis in fulfilment of the requirements for a postgraduate degree from Universiti Utara Malaysia, I agree that the Universiti Library may take it freely available for inspection. I further agree that permission for the copying of this thesis in any manner, in whole or in part, for scholarly purpose may be granted by my supervisor(s) or, in their absence, by the Dean of Awang Had Salleh Graduate School of Arts and Sciences. It is understood that any copying or publication or use of this thesis or parts thereof for financial gain shall not be allowed without my written permission. It is also understood that due recognition shall be given to me and to Universiti Utara Malaysia for any scholarly use which may be made of any material from my thesis.

Requests for permission to copy or to make other use of materials in this thesis, in whole or in part, should be addressed to:



Dean of Awang Had Salleh Graduate School of Arts and Sciences
UUM College of Arts and Sciences
Universiti Utara Malaysia
06010 UUM Sintok

Abstrak

Ujian-M yang digunakan untuk menguji kesamaan beberapa sampel bebas matriks kovarians adalah dibangunkan berdasarkan ujian nisbah kebolehjadian dengan andaian bahawa tertabur secara multivariat normal adalah sensitif terhadap kehadiran data pencilan. Ujian digunakan dalam banyak bidang ekonomi dan pasaran kewangan. Dalam banyak aplikasi, saiz sampel, n , adalah lebih besar berbanding bilangan pembolehubah, p , yang dikenali sebagai data berdimensi tinggi $p > n$. Apabila data telah ditetapkan dalam keadaan multivariat, sampel bebas matriks kovarians yang diperlukan bukan lagi varians. Oleh itu, anggaran nilai sampel vektor min dan sampel matriks kovarians akan menjadi tidak tepat dengan adanya masalah pencilan data dan kesingularan. Kegagalan untuk memenuhi andaian dapat memutarbelitkan kadar ralat Jenis I dan kuasa ujian. Oleh itu, untuk mengatasi kelemahan ini, kami mengubahsuai ujian-M dengan mengintegrasikan kovarian median geometri untuk menghasilkan ujian kesamaan kovarians teguh yang baru yang boleh menangani masalah pencilan data. Seterusnya, kami menggabungkan penganggar teguh dengan *banding* dan *tapering* untuk menyelesaikan masalah singulariti. Kajian simulasi dijalankan dengan mempertimbangkan bilangan pembolehubah, saiz sampel dan perubahan kovarians yang berbeza-beza bagi menilai prestasi ujian teguh yang baru. Keputusan menunjukkan bahawa kadar ralat Jenis I dan kuasa bagi nilai ujian- M_{BGMed} dan ujian- M_B adalah yang terbaik sama ada dalam keadaan sampel saiz besar mahupun kecil. Selain itu, dengan menggunakan aplikasi data sebenar berkaitan struktur mata wang Eropah, Amerika dan Asia semasa krisis kewangan Subprima US, pengesahan menunjukkan ujian- M_{BGMed} adalah lebih sensitif kepada perubahan nilai kovarians. Kesimpulannya, keputusan berdasarkan aspek teori dan aplikasi sebenar menyokong bahawa M_{BGMed} dapat menampung pencilan data dan juga data berdimensi tinggi.

Kata kunci: Matrik kovarians, Berdimensi Tinggi, Ujian-M, Kaedah regularisasi, Pencilan data.

Abstract

The original M-test used for testing equality of several independent samples covariance matrices is developed based on likelihood ratio test under assumption of multivariate normality distribution, is sensitive to presence of outliers in the data. The test is fast achieving significant distinction in many areas of economics and financial market. In many practical applications, sample size, n , is smaller than number of variables, p , which is known as high dimension data $p > n$. When the data is considered as multivariate setting, sample covariance matrices independent of each other is required instead of variance. Consequently, estimation of sample mean vector and sample covariance matrix will be inaccurate in presence of outliers and singularity problems. Failure to fulfil assumption can distort Type I error rates and power of test values. Thus, in order to overcome these drawbacks, modified M-test by integrating geometric median covariance for producing new robust covariance equality testing which can handle outlier. Furthermore, hybrid that robust estimator with banding and tapering to solve the singularity problems. Simulation study was presented by considering different number of variables, sample sizes and covariance shifts to evaluate the performance of the new robust tests. Results highlighted, Type I error rates and power of test values of M-test banding with geometric median (M_{BGMed} -test) and M-test banding (M_B -test) performed best in both cases (either n small or large sample size), respectively. Additionally, real application on European, American and Asian currencies structure during the US subprime financial crisis, the validation shown that M_{BGMed} -test is more robust. In conclusion, the theoretical and real application results established and supported M_{BGMed} -test can accommodate outlier and high dimension data.

Keywords: Covariance matrix, High dimension, M-test, Regularization methods, Outliers.

Acknowledgements

I wish to express my sincere appreciation to those who have contributed to this thesis and supported me in one way or the other during this wonderful journey.

Firstly, my sincere appreciations to my supervisor Dr. Shamshuritawati Sharif without whose guidance, support, patience and encouragement, this study could not have been possible. I am indeed deeply indebted to her. I would also like to thank Universiti Utara Malaysia (UUM) for the support on the publications.

Thanks to Dr. Ahmed Egbako, the Rector Niger State Polytechnic, Zungeru, for his guidance and financial support. To all of my friends, especially Engr. Dr. Olusola Emmanuel who had directly or indirectly lend me their friendship, moral support and endless encouragement during my study, thank you from the bottom of my heart.

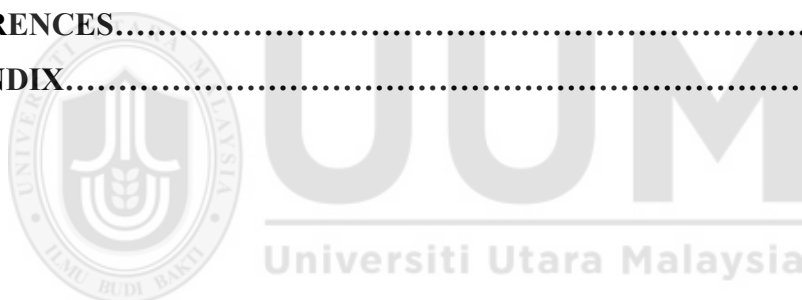
I am deeply grateful to my wife Ritta Favour Alo for her personal support, prayers and for being the good listener I could ever wish for and above all is her great patience at all time. Words cannot express the feelings I have for my mother Christiana and late father Joseph Alo and my siblings for emotionally constant support. Finally, to my beloved children Favour, Samuel and Blessing Alo that have been a constant source of strength and inspiration. I would not have been here if it is not for you all.

Table of Contents

Permission to Use	i
Abstrak	ii
Abstract	iii
Acknowledgements	iiiv
Table of Contents	v
List of Tables	viii
List of Figures	xv
List of Symbols	xvi
List of Appendices	xviii
CHAPTER ONE INTRODUCTION	1
1.1. Introduction to the Chapter	1
1.2. Background of the Study	10
1.3. Problem Statement	14
1.4. Research Questions	14
1.5. Research Objectives	15
1.6. Research Significance	16
1.7. Research Scope	16
CHAPTER TWO LITERATURE REVIEW	18
2.1. Introduction	18
2.1.1. Overview of M-test	18
2.1.2. Covariance and its Estimation	19
2.1.3. Multivariate Outliers	31
2.1.4. Robust Estimators	32
2.1.5. The Robust Geometric Median Covariance	36
2.1.6. Mathematical Definition of Geometric Median Covariance	36
2.1.7. Robust Properties of the Geometric Median	38
2.2. Regularizing the Sample Covariance Matrices	39
2.2.1. High Dimensional Datasets	43
2.2.2. Banding Method	44
2.2.3. Tapering Method	46

2.2.4. Thresholding Method	48
2.3. Summary Review of the Literature	51
CHAPTER THREE RESEARCH METHODOLOGY	52
3.1. Introduction to the Chapter	52
3.2. The Research Procedure	52
3.3. The Overall Research Framework with Flow Chart.....	54
3.4. The Procedure of New Robust Tests of Covariance Equality	54
3.4.1. Mean, Median and Geometric Median Covariance	55
3.4.2. Robust M-Test Based on Geometric Median Covariance.....	62
3.4.2.1 Integration of M-Test Based on Geometric Median Covariance with Banding Method	64
3.4.2.2 Integration of M-Test Based on Geometric Median Covariance with Tapering Method	66
3.5. Evaluation of the Modified M-Test	68
3.5.1. Type I Error Rate.....	68
3.5.2. Power of Test ($1 - \beta$)	69
3.5.3. Significance Level.....	73
3.6. Illustrative of the Impact of US Subprime Financial Crisis On the FOREX Using Covariance Matrices Structure	73
3.6.1 Illustrative Application on Financial Data Sets.....	76
3.6.2 Data Checking For Outliers in High Dimensional Data Sets.....	81
3.6.3 Implementation and Algorithms of Data Checking for Outliers in High Dimensional Data Sets	82
3.6.4 Testing the Currencies Covariance Matrices of Subprime Financial Crisis	86
3.6.5 Control Chart	87
CHAPTER FOUR RESULTS AND FINDINGS	89
4.1. Introduction	89
4.2. Type I Error Rates	89
4.2.1 Type I Error for Small Number of Variables ($p = 3, 5$ and 8)	90
4.2.2 Type I Error for Medium Number of Variables ($p = 10, 15$ and 20) ..	94
4.2.3 Type I Error for Large Number of Variables ($p = 25, 30$ and 50).....	98

4.3	Power of Test Values.....	103
4.3.1	Power of Test for Small Number of Variables ($p = 3, 5$ and 8).....	104
4.3.2	Power of Test for Medium Number of Variables ($p = 10, 15$ and 20).....	108
4.3.3	Power of Test for Large Number of Variables ($p = 25, 30$ and 50).....	112
4.3.4	Conclusion of the Type I Error Rates and Power of Test Values.....	116
4.4	Real Application Example.....	116
4.4.1	Checking for the Outlier(s) of Largest Difference from the Mean	117
4.4.2	Testing the Equality of Two Covariance Matrices.....	119
4.4.3	Testing Several Covariance Matrices Using Control Chart.....	120
4.4.3.1	M_{BGMed} -test Control Chart.....	121
4.4.3.2	M_B -test Control Chart.....	124
CHAPTER FIVE DISCUSSION AND RECOMMENDATION.....		128
5.1	Discussion and Conclusions	128
5.2	Future Research	133
REFERENCES.....		134
APPENDIX.....		156



List of Tables

Table 3.1 List Countries and their Symbols.....	79
Table K.1 Type I error for $p = 3, n = 2$	174
Table K.2 Type I error for $p = 3, n = 5$	175
Table K.3 Type I error for $p = 3, n = 10$	176
Table K.4 Type I error for $p = 3, n = 15$	177
Table K.5 Type I error for $p = 3, n = 20$	178
Table K.6 Type I error for $p = 3, n = 25$	179
Table K.7 Type I error for $p = 3, n = 50$	180
Table K.8 Type I error for $p = 5, n = 2$	181
Table K.9 Type I error for $p = 5, n = 5$	182
Table K.10 Type I error for $p = 5, n = 10$	183
Table K.11 Type I error for $p = 5, n = 15$	184
Table K.12 Type I error for $p = 5, n = 20$	185
Table K.13 Type I error for $p = 5, n = 25$	186
Table K.14 Type I error for $p = 5, n = 50$	187
Table K.15 Type I error for $p = 8, n = 2$	188
Table K.16 Type I error for $p = 8, n = 5$	189
Table K.17 Type I error for $p = 8, n = 10$	190

Table K.18 Type I error for $p = 8, n = 15$	191
Table K.19 Type I error for $p = 8, n = 20$	192
Table K.20 Type I error for $p = 8, n = 25$	193
Table K.21 Type I error for $p = 8, n = 50$	194
Table L.22 Type I error for $p = 10, n = 5$	195
Table L.23 Type I error for $p = 10, n = 10$	196
Table L.24 Type I error for $p = 10, n = 15$	197
Table L.25 Type I error for $p = 10, n = 20$	198
Table L.26 Type I error for $p = 10, n = 25$	199
Table L.27 Type I error for $p = 10, n = 50$	200
Table L.28 Type I error for $p = 15, n = 5$	201
Table L.29 Type I error for $p = 15, n = 10$	202
Table L.30 Type I error for $p = 15, n = 15$	203
Table L.31 Type I error for $p = 15, n = 20$	204
Table L.32 Type I error for $p = 15, n = 25$	205
Table L.33 Type I error for $p = 15, n = 50$	206
Table L.34 Type I error for $p = 20, n = 5$	207
Table L.35 Type I error for $p = 20, n = 10$	208
Table L.36 Type I error for $p = 20, n = 15$	209

Table L.37 Type I error for $p = 20, n = 20$	210
Table L.38 Type I error for $p = 20, n = 25$	211
Table L.39 Type I error for $p = 20, n = 50$	212
Table M.40 Type I error for $p = 25, n = 5$	213
Table M.41 Type I error for $p = 25, n = 10$	214
Table M.42 Type I error for $p = 25, n = 15$	215
Table M.43 Type I error for $p = 25, n = 20$	216
Table M.44 Type I error for $p = 25, n = 25$	217
Table M.45 Type I error for $p = 25, n = 50$	218
Table M.46 Type I error for $p = 30, n = 5$	219
Table M.47 Type I error for $p = 30, n = 10$	220
Table M.48 Type I error for $p = 30, n = 15$	221
Table M.49 Type I error for $p = 30, n = 20$	222
Table M.50 Type I error for $p = 30, n = 25$	223
Table M.51 Type I error for $p = 30, n = 50$	224
Table M.52 Type I error for $p = 50, n = 5$	225
Table M.53 Type I error for $p = 50, n = 10$	226
Table M.54 Type I error for $p = 50, n = 15$	227
Table M.55 Type I error for $p = 50, n = 20$	228

Table M.56 Type I error for $p = 50, n = 25$	229
Table M.57 Type I error for $p = 50, n = 50$	230
Table N.58 Power of test for $p = 3, n = 2$	231
Table N.59 Power of test for $p = 3, n = 5$	232
Table N.60 Power of test for $p = 3, n = 10$	233
Table N.61 Power of test for $p = 3, n = 15$	234
Table N.62 Power of test for $p = 3, n = 20$	235
Table N.63 Power of test for $p = 3, n = 25$	236
Table N.64 Power of test for $p = 3, n = 50$	237
Table N.65 Power of test for $p = 5, n = 2$	238
Table N.66 Power of test for $p = 5, n = 5$	239
Table N.67 Power of test for $p = 5, n = 10$	240
Table N.68 Power of test for $p = 5, n = 15$	241
Table N.69 Power of test for $p = 5, n = 20$	242
Table N.70 Power of test for $p = 5, n = 25$	243
Table N.71 Power of test for $p = 5, n = 50$	244
Table N.72 Power of test for $p = 8, n = 2$	245
Table N.73 Power of test for $p = 8, n = 5$	246
Table N.74 Power of test for $p = 8, n = 10$	247

Table N.75 Power of test for $p = 8, n = 15$	248
Table N.76 Power of test for $p = 8, n = 20$	249
Table N.77 Power of test for $p = 8, n = 25$	250
Table N.78 Power of test for $p = 8, n = 50$	251
Table O.79 Power of test for $p = 10, n = 5$	252
Table O.80 Power of test for $p = 10, n = 10$	253
Table O.81 Power of test for $p = 10, n = 15$	254
Table O.82 Power of test for $p = 10, n = 20$	255
Table O.83 Power of test for $p = 10, n = 25$	256
Table O.84 Power of test for $p = 10, n = 50$	257
Table O.85 Power of test for $p = 15, n = 5$	258
Table O.86 Power of test for $p = 15, n = 10$	259
Table O.87 Power of test for $p = 15, n = 15$	260
Table O.88 Power of test for $p = 15, n = 20$	261
Table O.89 Power of test for $p = 15, n = 25$	262
Table O.90 Power of test for $p = 15, n = 50$	263
Table O.91 Power of test for $p = 20, n = 5$	264
Table O.92 Power of test for $p = 20, n = 10$	265
Table O.93 Power of test for $p = 20, n = 15$	266

Table O.94 Power of test for $p = 20, n = 20$	267
Table O.95 Power of test for $p = 20, n = 25$	268
Table O.96 Power of test for $p = 20, n = 50$	269
Table P.97 Power of test for $p = 25, n = 5$	270
Table P.98 Power of test for $p = 25, n = 10$	271
Table P.99 Power of test for $p = 25, n = 15$	272
Table P.100 Power of test for $p = 25, n = 20$	273
Table P.101 Power of test for $p = 25, n = 25$	274
Table P.102 Power of test for $p = 25, n = 50$	275
Table P.103 Power of test for $p = 30, n = 5$	276
Table P.104 Power of test for $p = 30, n = 10$	277
Table P.105 Power of test for $p = 30, n = 15$	278
Table P.106 Power of test for $p = 30, n = 20$	279
Table P.107 Power of test for $p = 30, n = 25$	280
Table P.108 Power of test for $p = 30, n = 50$	281
Table P.109 Power of test for $p = 50, n = 5$	282
Table P.110 Power of test for $p = 50, n = 10$	283
Table P.111 Power of test for $p = 50, n = 15$	284
Table P.112 Power of test for $p = 50, n = 20$	285

Table P.113 Power of test for $p = 50, n = 25$	286
Table P.114 Power of test for $p = 50, n = 50$	287



List of Figures

Figure 3.1: Overall Research Framework with Flow Chart.....	54
Figure 3.2: Statistical test with corresponding M-Test and Several Modified M-Test.....	55
Figure 3.3: The flowchart showing the algorithms for outliers data check.....	84
Figure 4.1: Scatter plot of the largest different between sample mean values which is the outlier.....	118
Figure 4.2: FOREX Control chart for M_{BGMed}	123
Figure 4.3: FOREX Control Chart for M_B	126



List of Symbols

Σ_m	Elements of population covariance matrix of m
σ_i^2	Elements of population variance of i
$\forall_i$	For all values of i
χ^2	Chi-square distribution
$ \mathbf{S} $	Improved generalized variance
$\mathbf{S}^{-1}$	Inversion of sample covariance
$\ M\ _{op}$	Operator norm
$\ M\ _F$	Frobenius norm
$\ \cdot\ _2$	Euclidean norm for vectors
$\underset{\lambda}{\operatorname{argmin}}$	Minimizes the λ value
$\mathbb{R}^d$	Real values of d
$x_i \in \mathbb{R}$	All values of x_i belong to real numbers
$1 - \beta$	Power of a test
R_n	The number of experiments that fall in the rejection region
U	Is the total number of repetitions in the simulation experiments
I_p	Identity matrix
$P_j(t)$	Is the time price of currencies at the time t
$\nabla f(Y)$	Derivative in order functional f of Y
$M_{p \times p}$	Given a matrix of order $p \times p$
$(*)$	Schur coordinate-wise matrix multiplication
$\mathbf{S}$	Sample covariance matrix
$T_\lambda(\mathbf{S})$	The thresholding operator λ of a sample covariance matrix

$\check{Y}$	Sample median
S_{11}	Sample variance and covariance
$M(\mathbf{X})$	The median of set $\mathbf{X}$ is defined as a point m
α	Type I error rate
ρ	Population correlation coefficient
p	The number of variables
n	Sample sizes
epsilon	Percentages of outliers
kappa	Probability of the used shift of the mean
R	The price return logarithm exchange rate



List of Appendices

Appendix A The Algorithm of M_{BGMed} -Test and M_{TGMed} -Test using Regularization Methods.....	156
Appendix B The Steps for Power of Test Computation For Performance.....	158
Appendix C The Steps for Computing Parameter k by Cross Validation Method.....	159
Appendix D Computation on Banded Parameter k	163
Appendix E Banded Method Mathematical Formula Validation.....	164
Appendix F Simple Banding Sample Covariance Matrix with k -Parameter (R Software Code).....	165
Appendix G Simple Tapering Sample Covariance Matrix with l - Parameter (R Software Code).....	166
Appendix H Mean, Median and Geometric Median Covariance.....	167
Appendix I The Sample Median Covariance.....	169
Appendix J The Computation of Geometric Median Covariance	171

CHAPTER ONE

INTRODUCTION

1.1 Introduction

This chapter begins with background of the study in section 1.2. It aims to understand testing the equality of several independent covariance matrices in multivariate statistical analysis. This section is followed by the details of problem statement in section 1.3 with research questions in section 1.4 and the research objectives in section 1.5. Later on, in the sixth and seventh sections the research significance and research scope of this study will be highlighted, respectively.

1.2 Background of the Study

Carrying out the homogeneity test of variance is one of the critical assumptions in most statistical procedure (Abdi, 2007). There are several number of statistics available for testing the homogeneity of variance, such as, Bartlett (1937), Hartley (1940), Bartlett and Kendall (1946), and Tang and Gupta (1987). The hypothesis to be tested for several independent samples of variance is, $H_0: \sigma_1^2 = \sigma_2^2 = \dots = \sigma_m^2$ versus $H_1: \sigma_i^2 \neq \sigma_j^2$ for at least one (i, j) is different.

In many practical application, there are many interrelated or multidimensional variables under study because multivariate analysis methods permit the estimate of all variables simultaneously therefore, calling for multiple implementations of univariate methods.

Therefore, the data is considered as multivariate setting and covariance matrix is required instead of variance.

Covariance matrix is vital in performing most of the multivariate statistical analysis. For example, principal component analysis (Bai & Shi, 2011), discriminant analysis (Box & Watson, 1962; and Rencher, 2003), factor analysis (Martin & Eaves, 1977), multivariate analysis of variance (Carey, 1998; Grice & Iwasaki, 2007) and canonical correlation (Liu & Calhoun, 2014) can be used after checking on the homogeneity of covariance matrix assumption.

In details, the null and alternative hypotheses for examining the homogeneity (or equality) test of several covariance matrices of m independent samples are as follows, $H_0: \Sigma_1 = \Sigma_2 = \dots = \Sigma_m$ versus $H_1: \Sigma_i \neq \Sigma_j$ for at least one pair (i, j) where, $i, j = 1, 2, \dots, m$ (Sharif, Ruslan & Atiany, 2018).

The development of M-test started in 1928 by Neyman and Pearson. They proposed a general technique, called the likelihood ratio test. It is followed by Generalized Variance test (Wilks, 1932), Box M-test (Box, 1949), Jennrich test (Jennrich, 1970) and Vector Variance test (Djauhari, 2005). However, the development is not limited to the theoretical aspects but it has been studied extensively in applied research, especially in the area of quality control (Benneyan, 1998) and financial market (Marek Klimczak, 2008).

In quality control area, the test on equality of covariance matrices can be used to monitor the stability or instability for each of subgroup samples. For instance, Ghute and Shirke (2008) combine the traditional common sample variance chart and the conforming run length chart to monitor the stability of covariance matrix. Additionally, Machado and Costa (2012) used MRMAX control chart for the case of bivariate processes. While Hamed (2014) implement the improved generalized variance $|S|$ control chart for examining whether two or more related and joint variables are in control.

Additionally, in area of financial market, covariance matrix equality test may provide the information on the shift in the structure of $p \times p$ interrelationship when economic crisis arise, such as the occurrence of US subprime crisis. Several research works in financial market have been conducted by Tang (1998), Beirlant et al., (2005), and Annaert, Claes and De Ceuster (2006) to examine covariance matrix for identifying the intertemporal stability of stock returns, for determining the intertemporal stability of credit spread changes in the corporate bond market, and for implementing macro econometric, optimal portfolio, actuarial science and risk management, respectively.

The M-test is used to test the equality of several covariance matrices. M-test is among the well-known statistical test especially to the applied researchers. It is a parametric test that can be used to compare variation in multivariate data sets. M-test is improved Likelihood Ratio Test (LRT) which developed under the multivariate normality distribution (Box, 1949).

This test is also restricted under the assumption that the sample covariance matrices are independent and sample size, n must be larger than the number of variables, p (Sharif, 2013; Sharif, Ruslan & Atiany, 2018).

Mathematically, the original test (Box, 1949) is given as follows,

$$M = \mathbf{Y} \sum_{i=1}^m (n_i - 1) \log |\mathbf{S}_i^{-1} \mathbf{S}_{pooled}| \quad (1.1)$$

where:

$$\mathbf{Y} = 1 - \frac{2p^2 - 3p - 1}{6(p+1)(m-1)} \left(\sum_{i=1}^m \frac{1}{n_i - 1} - \frac{1}{n - m} \right)$$

$$\mathbf{S}_{pooled} = \frac{\sum_{i=1}^m (n_i - 1) \mathbf{S}_i}{n - m}$$

is pooled sample covariance matrix

$\mathbf{S}_i$ is the sample covariance matrix unit calculated from sample i based on mean.

m is the number of subgroup

$n = n_1 + n_2 + \dots + n_m$, n_i , i -th sample size.

M-test shows that the distribution can be approximated by χ^2 distribution if p and n are not greater than five and 20, respectively and F-distribution for p and n are greater than five. Under H_0 the statistical test is rejected if $\chi^2 \geq \chi_{(1-\alpha), \frac{1}{2}(m-1)p(p-1)}^2$ (Mardia, Kent & Bibby, 1979; Aitchison, 2003; Yusoff, Harun, Yahaya & Abdullah, 2013). The χ^2 approximation will be sufficient to be used in any practical purposes.

The same χ^2 distribution is used to test the hypothesis for the modified M-test based on the median (Cordeiro & de Paula Ferrari, 1991). The degrees of freedom is calculated by $\frac{p(p-1)}{2}$, where p is the number of variables. The Chi-squared distribution with $(1 - \alpha)^{th}$ location. Based on Equation 1.1 it is clearly define that the computation of M-test contain with the determinant, $|S|$ and inversion of sample covariance matrix, S^{-1} .

Basically, when it comes down to real application, data that can be used in working with normality assumption is hard to find. Failure to fulfil the normality assumption can distort the Type I error rates (Yusof, Abdullah, Yahaya & Othman, 2011). This make distributional performance totally fails if the normality assumption is not completely met (Aslam & Rocke, 2005).

The M-test is highly sensitive to the presence of outliers based on the use of the conventional estimate mean covariance in the formula. Consequently, the standard recommendation to this problem is to use a nonparametric test that is distribution-free, or performing a simple data transformation. However, non-parametric test is not so powerful compared to parametric test. This test usual call for large sample size to reject the null hypothesis, and the computation is tiresome and challenging (Daniel, 1990). In addition, simple transformation is another way to overwhelm the existence of outliers. The simple transformation method failed to treat outliers with efficiency (Wilcox, 2005).

Besides, when working with multivariate data, the probability that outliers exist in the dataset is high. The presence of outliers can usually cause and harm the normality assumption (Ali, 2013). Consequently, outliers still remain and equally reduced the performance of the statistical power when applying simple transformation.

Moreover, there are two other different procedures that can be used to minimize the effect of outliers. The first procedure is to calculate the original estimator i.e., M-test after identifying and removing the outliers from the data, then use the classical estimator the mean covariance estimate on the remaining data points.

In the second procedure, robust estimator is used to replace the classical estimator in reducing the influence of outliers (Yuan & Bentler, 2001; Yahaya, Ali & Omar, 2011). The aim of the robust procedure is to produce a reliable parameter estimator, good performance of the Type I error rates, and make it outliers free, even though data follow a given distribution correctly (Duan, Xu, Raffield, Chang, Wu, Lange & Li, 2017).

Besides, the robust procedure of geometric median is an important tools in analyzing data that contain contaminated observation (Muthukrishnan & Radha, 2010). The geometric median can be used to detect outliers, identify outliers and provide robust results when there is presence of outliers in the data sets. Therefore, the geometric median estimator attempts to provide a good results, hence, would be used in this research work (Diakonikolas, Kamath, Kane, Li Moitra & Stewart, 2017).

Moreover, the normality assumption required, M-test can only accommodate when sample size, n , is larger than the number of variables, p (Sharif, Djauhari & Djauhari, 2012). In many real application, not all data were able to meet the criteria set by M-test. For example, medical images and genetic arrays typically generate data with more variables, p than sample size, n (Rothman, Levina & Zhu, 2010).

Furthermore, in many sciences i.e., spectroscopy, functional magnetic resonance imaging, text retrieval and gene arrays high dimensional data is becoming the standard (Williamson, 2014). The introduction of new computing methods measured more variables than the sample size, using standard covariance matrices becomes challenging to estimate the relationship between variables, p (Fan, Liao & Liu, 2015).

High dimensional of data set in statistics denotes when the number of variable is larger than the number of sample size $p > n$ (Sen & Singer, 2017). For instance, a study on the relationship between several genes in humans. 10,000 genes were measured, with 200 observations in the study (Williamson, 2014). The data matrix is $200 \times 10,000$, but wish to study all pairwise relationship, covariance matrix computed is $10,000 \times 10,000$, resulted into covariance matrix not positive definite and may not give a good estimate of the population of covariance matrix.

While, data in economics and finance such as currencies and stock exchanges, commonly this kind of data do not have a problem to gain large sample size. However, when researchers want to observe the short-term effects of a global financial crisis such as US Subprime Crisis (2007-2010), Venezuelan banking crisis (2009-2010), Russian financial crisis (2014), Chinese stock market crash (2015) and many more, the data is usually small and the variable is large. Consequently to this problem, carrying out M-test is quite cumbersome to compute the covariance matrices with high dimensional data sets, since it involve the estimate of the determinant and inversion of a matrix, this make the matrix singular (Fan, Liao & Liu, 2015).

High dimensional dataset means that the number of variable p , is larger than the sample size, n (Pouramadi, 2013). It is also supported by several researchers that high dimensional data set could be the number of variables exceed the number of observations, for example, Tang, Liu, Zhang and Mei (2016) and Chatfield (2018). High dimensional also can be identified when the calculations become extremely difficult.

For the reason that in statistics, you will never have deterministic answer when $p > n$ unless you introduce your own assumption (Maronna, Martin, Yohai & Salibián-Barrera, 2018). The main problem when dealing with high dimensional data is the volume of the space increases so fast that the available data become sparse and the estimation of the determinant become zero and inversion of the matrices is difficult.

In multivariate setting, testing covariance matrices with high dimensional data sets becomes an important problems, that can be seen in many fields, ranging from finance and economics, social networks, biology and health (Fan, Han & Liu, 2014). Covariance matrix is sparse and estimation of covariance is more challenging.

In recent years, researchers have suggested various regularization methods to reliably estimate covariance matrices with high dimensional data set (Bickel & Levina, 2008a; Lam & Fan, 2009; El Karoui, 2010; Rigollet & Tsybakov, 2012). The method is particularly remarkable when the distribution of the data generating process is non-Gaussian and heavy-tailed (Wegkamp & Zhao, 2013; Mitra & Zhang, 2014; Fan, Liao & Liu, 2015).

The idea of regularizing the lower triangular matrix of the Cholesky decomposition of inversion of population covariance (Σ^{-1}) has been considered in Wu and Pourahmadi (2003), Huang, Liu, Pourahmadi and Liu, (2006) and Pourahmadi (2013) which is effective in guaranteeing the positive definiteness of the estimated covariance matrix. The convex banding estimator can be used to improve the performance of discriminant analysis for its classification (Bien, Bunea & Xiao, 2016). In order to address an effective robust M-test in low and high dimensional data set, this study attempt to enhance the regularization method in estimating robust covariance estimator.

The two problems of outlier and singularity mentioned above have inspired us to further explore the best method for modifying M-test. To solve these problems, it is essential to provide consistent performance results in the presence of outliers, however, the conventional M-test estimate covariance matrix based on mean is not robust with the formula, and also to resolve the singularity problem faced by M-test when involve high dimension data set.

1.3 Problem Statement

In this section, M-test has two main problems when we use the conventional mean covariance estimate. Firstly, problem to be solved is M-test fails to incorporate outliers in the data sets and secondly, singularity problems. The estimation of sample mean vector and sample covariance matrix results will not be accurate in the presence of outliers and singularity problems. In multivariate analysis settings, data collected are contaminated by errors, with failure to fulfil the normality assumption can distort the Type I error rates. This make the distributional performance totally fails if the normality assumption is not completely met. One problem that contribute to non-normality distribution is outliers. Outliers or contaminants are observations that are inconsistent in a data set.

The presence of outliers usually can cause or harm the normality assumption. M-test rely on assumption that data is normally distributed. If the assumption is violated, then the statistical result may not be appropriate. M-test is highly sensitive to the presence of outliers which will make the results inaccurate.

It has damaging effects on statistical analysis, increasing the variance of error and reducing the statistical test power (Osborne & Overbay, 2012; Dan & Ogu, 2013).

Although, outliers are contaminant to a set of data, as can be found in the field of natural sciences such as plant breeding, survey processing and geochemical exploration (Croux & Ruiz-Gazen 2005; Estaghvirou, Ogutu & Piepho, 2014; Chambers, Chandra, Salvati & Tzavidis, 2014). The method in which covariance matrix is estimated with the presence of outliers is key to multivariate analysis.

Furthermore, outliers can have a dramatic impact on the results of common multivariate statistical analyses i.e., correlation coefficients can be influenced (Marascuilo & Serlin, 1988; Osborne & Overbay, 2004), create difficulties in regression analysis, leading to the presence of collinearity among the set of predictor variables in multiple regression (Pedhazur, 1997). Outliers falsely influence the degree of linearity present between a pair of variables and correlation may in turn lead to bias sample estimates (Osborne et al., 2004).

In addition, the procedures that are based on the correlation coefficient for example, factor analysis and structural equation modeling are also negatively influenced by the presence of outliers in data sets (Brown, 2006). Cluster analysis is mostly sensitive to outliers with a shift of cluster results when outliers are the center or starting point of the analysis (Kaufman & Rousseeuw, 2005). Also, outliers can themselves form a cluster, which is not truly representative of the larger collection of values in the data sets. Outliers have shown to damagingly influence testing for mean differences using ANOVA through biasing group means where they are present (Osborne et al., 2004).

However, when data are analysed, outlying caused problems since they may strongly influence the result. Therefore, robust statistics aims at detecting the outliers by searching for the test fitted by the majority of the data. Robust statistics are analysis with good performance for data with outliers, drawn from a wide range of probability distributions with normality assumption problems.

Moreover, the striking characteristics of an ideal robust estimator is its ability to high breakdown points, computational and statistical efficiencies, equivariance, equal good behaviour with clean data sets, and test the dispersion matrix in multivariate analysis even in cases were the assumptions of the distribution are met (Chatfield & Collins, 2013).

Conversely, it becomes even more challenging if the normality assumptions of the distributions are not met and there are cases of suspected outliers inherent in the data. However, despite the use of several methods to minimize outliers in the data set, outliers still have serious bias and influence on good estimate (Osborne et al., 2004). Consequently, it poses a dilemma to the researchers in terms of explanation for either omitting or retaining them in the analysis. Therefore, in order to overcome the drawbacks of outliers and singularity in a multivariate distributions, this research propose to introduce a robust geometric median covariance estimation in M-test that able to incorporate the presence of outliers.

Secondly, M-test is developed based on likelihood ratio test for testing equality of several covariance matrices (Box, 1949). Theoretically, this test can be used for arbitrarily value of p . Additionally, in genomics, biomedical, imaging, functional magnetic resonance imaging (fMRI), tomography, and finance, its dimensionality p is usually large compared to the size of the sample n , ($p > n$) (Pourahmadi, 2013). Hence, the existing test cannot deal with high dimensional case. It relies on the determinant and inversion of covariance matrix which may be singular and not well-conditioned (Djauhari & Herdiani, 2008; Pamukcu, Bozdogan & Calik, 2015; Ullah, 2015).

Generally, the determinant of covariance matrix will be singular in a situation where the entire row of matrix is zero, or if the components of two rows or columns are the same, or the rows or columns elements are constant multiple of another rows or columns (Sharif, Yusoff, Omar & Ismail, 2014). Therefore, the singularity problem in covariance matrix leads to an issue with the calculation of the likelihood ratio test (Ayyildiz, Gazi & Wit, 2012).

In this study, the need to modify the original M-test in order to solve the outlier and singularity problems. A new statistical test that able to solve the problems is required because not all kind of data are competent to handle the inflexible of M-test assumptions. The following research questions and objectives are set to be answered and achieved by handling the limitations of the original test.

1.4 Research Questions

The following research questions for this study is addresses in line with problem statement.

- i. How can M-test be modified to resolve outlier and singularity problems in high dimensional case?
- ii. What are the improvement of modified M-test in terms of Power of test values and Type I error rates in simulation study?
- iii. What are the performance applications of real data with outlier and high dimensional data set in the presentation of the new statistical tests?

1.5 Research Objectives

The ultimate goal of this research is to introduce new statistical test that can improve the performance of the M-test in relations to outlier and singularity problems. The following are the sub objectives to achieve in this study.

- i. To proposed the new robust tests of covariance equality integrating the geometric median covariance in generating a new standard robust test of covariance equality testing in the presence of outliers.
- ii. To investigate the new robust test of covariance equality with regularization method to resolve singularity problem in high dimensional case using the simulation study.

- iii. To measure the performances of M-test and the new robust tests of covariance equality using Type I error rates and the power of test with 5,000 iterations on variables (p) and sample sizes (n) and other conditions with simulation study.
- iv. To evaluate the performances of M-test and the new robust tests of covariance equality on real currencies exchange rate data for the cases of before and after, control charts US subprime financial crisis with application on real data with simulations.

1.6 Research Significance

The focus of this study is to improve the knowledge development in multivariate hypothesis testing specifically in modifying covariance test when dealing with outlier and high dimensional data sets by proposing regularized geometric median estimator. For example, financial and economic analysts could use this suggested test to validate a universal financial and economic crisis within a short period of time while involving a large number of variables in their analysis.

With regards to robust estimation, some improvements were made on M-test in terms of its performance followed by the development of a modified robust tests known as: Geometric Median M-test (M_{GMed}), banding M-test (M_B), tapering M-test (M_T), banded Geometric Median M-test (M_{BGMed}) and tapered Geometric Median M-test (M_{TGMed}) tests and its algorithms.

The modified estimators when used in M-test can improve its performance in terms of power of test and Type I error rate in monitoring the financial and economic crisis covariance shift within a short period of time. This study might also lead to many possible research questions that remain for consideration by academicians for future contribution to the body of knowledge.

1.7 Research Scope

This study focused on the performance of M-test and several modified M-tests in testing the equality of several independence covariance matrices. The several modified M-tests are used for solving outliers and singularity problems in high dimensional cases, where performance of the test was measured using Type I error rate and power of test values with the following simulation study (Todros, 2019).

- i. Simulation study to measure the performance of Type I error rate between the M-test and new robust test in terms of Type I error rate and power of test, by considering different type number of variables which are small ($p = 3, 5$ and 8), medium ($p = 10, 15$ and 20), and large ($p = 25, 30$ and 50) (Vanbrabant, 2018 and Powers, Qian, Jung, Schuler, Shah, Hastie & Tibshirani, 2018).
- ii. Introduce the new robust test that can improve the performance of the M-test in relations to outlier and singularity problems.
- iii. Apply the new robust test into real data by using currency exchange rate in testing the equality of several independent samples of covariance matrices.

Despite M-test's feasibility in handling all types of variables simultaneously, the number of variables must be restricted, as otherwise estimation results becomes challenging and can possibly be inaccurate. The problem was highlighted by Sadrzadeh, Baskin and Kline (2017) and Hendry (2018). Selection of variables is suggested for choosing only useful variables to be included in the simulation rule.



CHAPTER TWO

LITERATURE REVIEW

2.1 Introduction

Based on the study in chapter one, suggestion of a new statistical test that is able to resolve the multivariate outlier and singularity problems. This test is expected to provide reliable performance results for both situations faced by M-test. Therefore, in this chapter, all related reviews of the study is provided. This chapter begin with the introduction in Section 2.1. Then, followed by overview of M-test in Section 2.1.1. Section 2.1.2 covariance and its estimation is discussed to highlight the causes of the problems. Subsequently, Section 2.1.3 discussed in details about the multivariate outliers then, robust estimators in Section 2.1.4 and thereafter the geometric median covariance as a proposed solution to outlier problems. The last Section 2.2 presented regularizing the sample covariance matrices. Integrating two methods which are robust methods and regularization methods are desired to develop a new robust high dimensional M-test using regularized geometric median covariance.

2.1.1 Overview of M-test

M-test is very sensitive to non-normality assumption, whereby a significant value will determine either equal or unequal covariance matrices. However, it is essential to establish the multivariate normality for M-test before it is used.

Some studies for example Kaplanis (1988), Meric and Metric (1989), Wahab and Lashgari (1993) have used this statistical test to examine the intertemporal stability in international stock market relationships through testing the hypothesis of equal covariance matrices. The test also used the correlation matrices for two standardized random variables when the correlation is the same as the covariance. For example, the hypothesis testing of equal correlation matrices of stock returns, Kaplanis (1988). Also, more study, Tang (1995) used an extension of M-test for same hypothesis. Tang (1995) applied M-test directly on the equality of correlation matrices by transforming first the raw data into standard score for each time period before applying the M-test procedures.

There are some advantages of M-test in multivariate analysis.

- i. M-test is powerful with small sample shift in covariance structure.
- ii. M-test perform better compare to J-test in power of test with large sample.

The disadvantages of M-test are as followings.

- i. M-test is very sensitive to the departures of normality distribution
- ii. M-test has a less power for small sample covariance data.

2.1.2 Covariance and its Estimation

Covariance is a measure of association between two random variables (Bentler & Dudgeon, 1996). Let $x_1, x_2, \dots, x_n$, be n independent observations of $\mathbf{X}$ and $y_1, y_2, \dots, y_n$, be n independent observations of $\mathbf{Y}$.

Therefore, the sample covariance of two random variables is defined as,

$$\text{Cov}(\mathbf{X}, \mathbf{Y}) = \sum_{i=1}^n \frac{(x_i - \bar{x})(y_i - \bar{y})}{n-1} = s_{xy} \quad (2.1)$$

where $\mu_X = \{\mathbf{X}\}$ and $\mu_Y = \{\mathbf{Y}\}$ are the respective mean. Likewise, in the special case of $\mathbf{Y} = \mathbf{X}$, then,

$$\text{Cov}(\mathbf{X}, \mathbf{X}) = \{X^2\} - \{X\}^2 = s_{xx} = s_x^2 \quad (2.2)$$

so the covariance of $\mathbf{X}$ random variable is equivalent to variance. Commonly, covariance matrix carries information about the pairwise dependence among the components of a random vector. It is characterized by the nonnegative definiteness that is a constraint on all the entries of the matrix (Grote, 2007). The estimation of covariance matrix is one of the basic tools in multivariate statistics.

Let $\mathbf{Y} = (Y_1, Y_2, \dots, Y_p)$ be a p dimensional random vector with the mean vector $\boldsymbol{\mu} = E(\mathbf{Y}) = (\mu_1, \mu_2, \dots, \mu_p)'$, where $\mu_i = E(Y_i)$ is the mean of the i -th component. Then, its covariance matrix is defined as the $p \times p$ matrix.

$$\boldsymbol{\Sigma} = E(\mathbf{Y} - \boldsymbol{\mu})(\mathbf{Y} - \boldsymbol{\mu})' = (\sigma_{ij}) \quad (2.3)$$

where $\sigma_{ij} = E(Y_i - \mu_i)(Y_j - \mu_j)$ is the covariance between the two random variables Y_i and Y_j for $i \neq j$, and σ_{ii} is simply the variance of the i -th component of $\mathbf{Y}$.

Therefore, the population covariance matrix is given as $\Sigma = \begin{pmatrix} \sigma_{11} & \sigma_{12} & \dots & \sigma_{1p} \\ \sigma_{21} & \sigma_{22} & \dots & \sigma_{2p} \\ \vdots & \vdots & \ddots & \vdots \\ \sigma_{p1} & \sigma_{p2} & \dots & \sigma_{pp} \end{pmatrix}$, and

its sample version is $\mathbf{S} = \begin{pmatrix} s_{11} & s_{12} & \dots & s_{1p} \\ s_{21} & s_{22} & \dots & s_{2p} \\ \vdots & \vdots & \ddots & \vdots \\ s_{p1} & s_{p2} & \dots & s_{pp} \end{pmatrix}$.

The basic properties of the mean and covariance $p \times 1$ random vector $\mathbf{Y}$, with covariance matrix $\text{Cov}(\mathbf{Y})$, and non-random matrices A , B , C and D , then we will have,

$$E(A\mathbf{Y} + B) = AE(\mathbf{Y}) + B$$

$$\text{Cov}(A\mathbf{Y} + B) = A\text{cov}(\mathbf{Y})A'$$

$$\text{Cov}(A\mathbf{Y} + B, C\mathbf{Y} + D) = A\text{cov}(\mathbf{Y})C'$$

The second property suggest more ways to simplify covariance matrices from the complex conventional matrices to simpler ones. For example, starting from a $\mathbf{Y}$ with the identity covariance matrix, it follows that AA' is a genuine covariance matrix for any matrix A . This supports the view that, just like the variance of a random variable, the covariance matrix of a random vector is nonnegative in the sense that for any vector $(c_1, c_2, \dots, c_p)'$ the variance of the linear combination $c'\mathbf{Y}$ is nonnegative.

More precisely,

$$c' \Sigma c = \sum_{i=1}^p \sum_{j=1}^p c_i c_j \sigma_{ij} = E[c'(Y - \mu)]^2 = \text{var}[c'(Y - \mu)] \geq 0 \quad (2.4)$$

Therefore a symmetric matrix Σ is nonnegative definite if it satisfies (2.4). It is said to be positive definite if the inequality (2.4) holds for a nonzero vector $c = c_0 \neq 0$, so

$$c_0'(Y - \mu) = \sum_{i=1}^p c_{0i}(Y_i - \mu_i) = 0 \quad (2.5)$$

So that the random variables $Y_1, Y_2, \dots, Y_p$ are linearly dependent. To avoid this kind of redundancies, in the outcome we often work with the positive definite covariance matrices. In Pourahmadi (2013), for a $p \times p$ symmetric matrix $\Sigma = (\sigma_{ij})$.

The following are equivalent,

Σ is nonnegative definite.

all its leading principal minors are nonnegative definite, that is the $i \times i$ matrices are nonnegative definite

$$\Sigma_{ii} = \begin{pmatrix} \sigma_{11} & \dots & \sigma_{1i} \\ \vdots & \ddots & \vdots \\ \sigma_{i1} & \dots & \sigma_{ii} \end{pmatrix}, i = 1, 2, \dots, p.$$

all eigenvalues of Σ are nonnegative.

there exists a matrix A such that $\Sigma = AA'$,

there exists a lower triangular matrix L such that $\Sigma = LL'$

there exist vectors $u_1, \dots, u_p$ in R^p such that $\sigma_{ij} = u_i' u_j$.

In multivariate statistics, sample covariance matrix and mean vector are important and are used to describe the population distribution. The mean vector involve means of each variable. Whereas, the covariance matrix i.e., the diagonal elements correspond to the variances of each variables, and the off-diagonal elements represent the covariance between each pair of variables.

In particular, actual data sets contains outliers that are unusually large or small values when related with others in the data set. As a result, the presence of outliers will affect the estimation of mean vector and covariance matrix. This is supported by Hubert and Engelen (2004) where they found that the classical method based on the mean vector and covariance matrix of the data sets is very sensitive to outlier, this make the covariance matrix give not good results in the presence of outlying estimations.

Moreover, if the majority in the data sets do not follow the pattern of the observations or otherwise the existence of outlier, it might caused poor quality of the covariance estimates (Wang & Raftery, 2002; Sequira, Tsourdos & Lazarus, 2011; Harris, Brunsdon, Charlton, Juggins & Clarke, 2014).

Furthermore, the sample means and covariance matrix will be inefficient to estimates its population if the data sets presented is moderately heavy tailed (Yuan & Bentler, 2007; Yuan, Chan & Tian, 2015). In detail, heavy tails also caused the presence of outliers in the data sets (Janus, Koopman & Lucas, 2014). Heavy tailed and outliers distribution caused the standard error of the sample mean inflated until power can be relatively low when groups are compared (Wilcox, 2005).

The preferred method to solve the outlier occurrence is by using robust estimator instead of the classical covariance estimator. In section 2.1.3, the details about multivariate outliers is discussed.

2.1.3 Multivariate Outliers

The study of outliers is as important for multivariate data as it is for univariate samples (Barnett & Lewis, 1994). Nevertheless, it is more difficult to deal with outliers in multivariate than univariate data. There are various definitions given for outliers. An exact definition of an outlier often depends on the hidden assumptions regarding the data sets and the applied robust estimation method. Yet, some definitions are regarded generally enough to cope with various types of data and methods. Hawkins (1980) defines outlier as an observation that deviates so much from other observations as to arouse suspicion that it was generated by a different mechanism.

Similarly, an outlying observation, or outlier, is one that appears to deviate markedly from other members of the sample in which it occurs (Barnet & Lewis, 1994). In the same way, outlier is an observation in a data set which appears to be inconsistent with the remainder of that set of data (Johnson & Wichern, 1992). Meanwhile, Beckman and Cook (1983) interpret outlier as a collective to refer to either a discordant or contaminate observation. Discordant observation is any observation that appears surprisingly to an investigator.

Contaminant is any observation that is not a realization from the target distribution. Besides an outlier is define to be a contaminating data (Rousseeuw & van Zomeren 1991). Multivariate outliers or contaminated data can be defined as an observation(s) that is distant from other observations which is inconsistent in a data sets (Barnett & Lewis, 1994; Yuan *et al.*, 2001, Sharif, Yusoff, Omar & Ismail, 2014). No matter how it is defined, in general, outliers refer to a point(s) that is surprisingly different from the rest of the data. An immediate consequence of the presence of outliers is that they may cause apparent non-normality.

One of the main problem that contribute to non-normality distribution is an outlier (Auer, Reiner & Leal, 2016). If the normality assumption for the data set being used seems to be not valid, then, it is obvious the presence of an outlier. The probability of Type I error will be less than significant level and also reduce power of test, when there is presence of outliers in the data sets. The statistical result may not be suitable when the assumption of normality is violated. In order to attain this, the shape of distribution either normal or non-normal can be evaluated by using either a statistical test or graphical method, such as, histogram and normality plot.

The presence of outliers can hardly be removed using naked eye when the number of dimension is more than two (Sajesh & Srinivasan, 2013). This is the danger researchers have to be cautious about when working with large datasets of high dimensions. Thus, a reliable method is needed to deal with outliers especially for this sort of data sets.

In an effort to change the random vectors to be random variables so that outliers could be seen more noticeably, Beckman et al., (1983) suggested a transformation method using the Mahalanobis distance. The works on Mahalanobis distance could be found in almost any literature on multivariate analysis, including outliers' studies. The classical way to remove the multivariate outlier from data set is by calculating Mahalanobis distance, D_i^2 (Mahalanobis, 1936). D_i^2 is a standard test to measure the distance between pairs of group centroid in a distribution which takes into account the covariance matrix. This distance use to estimate the values that far away from the main bulk of data. The equation of D_i^2 by considering the vectors x_i and μ with respect to the matrix Σ for $i = 1, 2, \dots, n$ (Maronna, Martin, Yohai, 2006) is as follows:

$$D_i^2 = (x_i - \hat{\mu})^t \hat{\Sigma}^{-1} (x_i - \hat{\mu}) \quad (2.6)$$

where $\hat{\mu}$ is estimated mean vector and $\hat{\Sigma}$ is the sample covariance matrix. This measure is approximately distributed as a chi-square distribution with p degree of freedom, χ_p^2 . The observation is considered as outlier(s) when the probability value of D^2 is less than or equal to 0.001 (Finch, 2012). After deleting the outliers, the statistical method can be employed by using the remaining data.

However, this method has several shortcomings. Dealing with outliers in high dimensional multivariate setting is important when multiple outliers occur. In this condition, D_i^2 suffers when the number of outliers in the sample is unknown, two occurrences happens that is masking and swamping effects can occur (Rousseeuw & van Zomeren, 1991).

Hence, by removing the first outlier, the second outlier acted as an outlier. Consequently, D_i^2 does not identify the atypical observation. This is due to the fact that mean vector and covariance matrix are usually estimated in a non-robust manner, which is also affected by the outliers. Therefore, to estimate D_i^2 , a robust procedure is required to enhance an accurate measure for removing the outliers. Under those circumstances, robust estimator is needed to exchange mean vector and covariance matrix in equation 2.6.

In this study, outlier identification is an important feature in any statistical modification and estimation. For multivariate data, given $Y = \{y_1, y_2, \dots, y_n\} \subset \mathbb{R}^p$ as independent and identically distributed p -dimensional random vectors with mean $\boldsymbol{\mu} = (\mu_1, \mu_2, \dots, \mu_n)^T$ and a positive definite covariance matrix $\boldsymbol{\Sigma}$ whose entries are $(\sigma_{jk})_{j,k=1,2,\dots,p}$. Conventional outlier identification methods often rely on distance measure to describe how far a particular data point is from the center of data. Moreover, it is important to obtain stable estimates of $\boldsymbol{\mu}$ and $\boldsymbol{\Sigma}$, as well as to determine the starting point for $(\boldsymbol{\mu}, \boldsymbol{\Sigma})$ so as to know whether an observation is an outlier (Cerioli, Riani, Atkinson & Corbellini, 2018).

Consequently, outlier will lead to model misspecification, biased parameter estimation, and poor predictions. Thus, there are some suggestions to use simple transformation and non-parametric method in handling presence of outliers. Moreover, when dealing with contaminated data, simple transformation might be reliable solution. Nevertheless, transformation of data failed to handle outliers effectively even as they can adjust distribution to make the data more symmetrical (Wilcox, 2003).

Also, presence of outliers lead to reduce statistical power when using simple transformation and contaminants still remain in the data sets. This is because simple transformations do not remove effect of outliers. Thus, value produced by simple transformations had made some issues in the interpretation of the data since it involves placing the data on another scale (Lix, Keselman & Keselman, 1996).

Although simple transformation is a choice to solve for outliers and non-normal distribution, still, it was not completely recommended (Tabachnick & Fidell, 2007). The transformation analysis of variable is difficult to interpret. The second method that is usually used by researchers is the non-parametric procedures. This method involves minimal assumption about the form of distribution of the population, and proper for simple design and small data sets. This method do not require the parameter. Therefore, it seems challenging to describe about the difference between the actual and estimated parameter.

Furthermore, the computation is founded on ranks which study the order of the data but discard the real value. For example, the non-parametric test, such as, Mann-Whitney U test, Wilcoxon Rank Sum test, Wilcoxon Signed Rank test, Kruskal Wallis test, are generally less powerful than parametric method because those tests involve larger sample sizes to reject false hypothesis (Yahaya, Othman & Keselman, 2006). Though, the tests listed above are univariate, also the multivariate non-parametric test faced the same circumstance.

Also, using the two methods transformation of data and non-parameteric methods, researchers however, discovered other methods to eliminate the problem of outliers. The method is applying the statistical method after removing the outlier from the data (Karrila, Lee & Tucker-Kellogg, 2011) and incorporate a robust estimator into the original method (Huber, 1994, Rousseeuw, 1985).

As the dimensions of the data increase, Beckman and Cook (1983) had discussed in detail on the need to study on outliers and their effect on linear models. They stated that the existing outliers in the data will affect the estimation of a population parameter, hence, causing the inability of the model to provide an adequate fit or statistical explanation. Subsequently, alternative methods in reducing the influence of outliers is robust methods (Yuan & Bentler, 2000; Yahaya et al., 2011).

In multivariate, the work started by Huber which is related to robust methods have been successfully applied to regression, repeated measures, principal components, canonical correlation, discriminant analysis, dimension reduction and correspondence analysis (Yuan & Bentler, 2001). In addition, the improvement of robust method can be seen especially in Hotelling T-squared statistic that applied various robust estimator such as Minimum Vector Variance (Yahaya et al., 2011), and Fast Minimum Covariance Determinant (Salleh & Djauhari, 2011).

In literatures, robust statistical procedures give more emphasis on estimation and testing the hypotheses. By comparing with the previous methods, there are several advantages from robust methods. First, the estimation on the parameter interest provided more reliable (Valero & Moreno, 2005). Afterward, robust methods can fit well in the bulk of the data, even in high dimensional multivariate data (Maronna, Martin & Yohai, 2006).

More importantly, this method performs well with data drawn from non-normal distribution (Wang & Zwilling, 2015). Furthermore, it also gives approximately the same results as the classical method when the absence or presence of small proportion of outliers in data (Wang & Zwilling, 2015). Recently, many research had been done by adopting robustness statistical estimators (Campbell, Lopuhaa & Rousseeuw, 1998); Pena & Prieto, 2001; Aslam & Rocke, 2005; Maronna et al., 2006; Yao, Deng, Zhang & Ai, 2014).

Based on the advantages and disadvantages of all approaches to overcome the drawback of outliers, in this study, we will explore the robust estimator since performance of this method is good compared to other approaches (Susanti & Pratiwi, 2014; Zuo & Slotine, 2005; Pena & Prieto, 2001; Rousseeuw & van Driessen, 1999).

2.1.4 Robust Estimators

The normality assumption required in the estimation of classical methods is commonly not satisfied, Huber (1964) proposed the robust estimators. One of the aims using robust estimator as stated by Hampel (1985), is to identify the deviation of data, or contaminants. Robust estimation can be used in two different methods. The first method is to identify and remove outliers, then use classical estimators on the remaining good data points. In the second methods, the robust estimators are used in place of classical estimators (Beckman & Cook, 1983). In searching for highly robust estimators of location and scatter, there are some properties the estimators should possess.

There are some properties that can be used to determine the usefulness of a robust multivariate estimator (Jensen, Birch & Woodall, 2007). The first is the breakdown point (BP), where it was introduced by Hampel (1971, 1974) as a measure for the robustness of an estimator against outliers. The breakdown point of an estimator is the proportion of incorrect observations i.e., arbitrarily large observations an estimator can handle before giving an incorrect arbitrarily large result (Roy, König, Dvorkin & Kumar, 2015).

Furthermore, the second property to consider is affine-equivariant, which is an important and often desirable property of statistical estimates. When an estimator is affine-equivariant, changing the measurement scale or affine transformations should not affect the properties of the estimator.

The third property of the efficient estimator, this property concerns on how well the estimator makes use of all the good data available. Efficiency is always a very important performance measure for any statistical procedure (Zuo, 2006). Finally, an additional important and required feature of an estimator is the computational efficient, for easy and fast computation. Estimating robust estimators of location and scatter with high dimension is one of the primary problems encountered in multivariate sets (Leys, Klein, Dominicy & Ley, 2018).

2.1.5 The Robust Geometric Median Covariance

The most popular measure of location is arithmetic mean (AM) because it has lots of good and interesting mathematical properties (Das & Imon, 2014). It is generally known as mean. One of the most attractive property is the sample AM having the most efficient population mean estimator (Maronna, Martin & Yohai, 2006). However, the usefulness of the mean is distorted by the presence of outliers in the data set. AM is contaminated, even a single outlier exist (Tukey, 1960).

The AM is a measurement which is highly prone to outliers, with just an outlier can pull the value of the mean away from the majority of the data set, which is of a concern, when dealing on very large data sets (Wilkinson, 2018). The main drawback of the mean is that it is sensitive to the presence of outliers, even when the sample size is small (Leys, Ley, Klein, Bernard & Licata, 2013). Therefore, the mean is not a suitable measure of location to use with data sets containing outliers or contaminants.

The geometric mean (GM) is more appropriate than the AM in measuring the multivariate data sets especially when dealing with geometric series. GM is defined as a measure of location for a set of numbers by using the product of their values. It is contrary to the AM which use the sum of their values (Seixas, Robins & Moulton, 1988).

When dealing with the geometric series data type, the GM is more useful than the AM for describing proportional growth. Priority is given to GM when comparing different items when each items has multiple properties that have different range of numeric, for example, data on the number of surviving bacteria (units of 100) (Das & Imon, 2014).

Therefore, GM has rising and extensive applications in business (Wu, 2010); biology (Byrne, Knief, Keating & Bonnier, 2016); genetics (Petoukhov, 2016); social science (Paltoglou & Thelwall, 2017); population science (Soykan, Sauer, Schuetz, LeBaron, Dale & Langham, 2016), and in many other areas (Bargiela & Pedrycz, 2016).

In addition, the GM perform better when the observation distribution is skewed compare to the AM. In contrast to AM, the GM is measured in logarithmic scale. Consequently, the GM is not very sensitive to the extreme observations or the presence of outliers. The GM robustness property examination is really necessary because the presence of 1-10% outliers in a data sets can distort the results (Hampel, Ronchetti, Rousseeuw & Stahel, 1986).

In estimating the measure of dispersion and precision matrices, covariance play an important role. Generally, mean covariance is used in multivariate statistical method. However, the mean covariance is not robust compare to median covariance (Bruce & Vershynin, 2011).

In 2009, Fletcher, Venkatasubramanian and Joshi, introduce the Geometric median (GMed) covariance as a robust alternative to the mean covariance. The GMed is defined as a discrete set of sample items in a Euclidean space, it can includes, two-dimensional; three-dimensional space of Euclidean in a geometric series and certain other spaces, is the item minimizing the sum of distances to the sample items (Thom, 2017, Hegde, 2017). It is also called l-median (Cormode & McGregor, 2008; Zhang, 2017); spatial median (Drezner & Hamacher, 2001; Auerbach, & Dix, 2017); Euclidean minisum (Hildenbrand, 2006; Tan, Hohage, Kalentev, Joseph, Wang, Voit & Frahm, 2017) or Torricelli point (Kupitz & Martini, 1994; Durocher, Leblanc & Skala, 2017).

In multivariate analysis, the GMed is an essential notion where the location of the data set is necessary even with the presence of outliers (Efteliogu, 2015). GMed is found to be robust alternative of mean covariance. It show that the estimator is more stable when dealing with the presence of outliers in rainfall data (Das & Imon, 2014).

Also, the estimation of GMed is a direct generalization of real median in metric spaces. It has good robustness properties (Cardot, Cenac & Zitt, 2013). GMed has been used by Fletcher, et al, (2009) as minimizer of sum of geodesic distances to the data points and results showed that GMed is unique for any non-positively curved manifold, under certain conditions for positively curved manifolds (Hua, Cheng, Wang, Qin & Li, 2017).

Based on literature review, the mean is number which minimizes mean error squared, μ and median, M is the number, which minimizes the mean error (i.e. the middle of a sorted list of numbers),

$$M = \underset{M \in \mathbb{R}}{\operatorname{argmin}} \sum_{i=1}^n |Y_i - M| \quad (2.7)$$

$$\mu = \underset{\mu \in \mathbb{R}}{\operatorname{argmin}} \sum_{i=1}^n |Y_i - \mu|^2 \quad (2.8)$$

Using this definition of median, it becomes easy to generalize median to higher dimensions as follows:

Definition 1: The median of points $Y_1, Y_2, \dots, Y_n \in \mathbb{R}^n$ is define as:

$$\text{GMed}(Y_1, Y_2, \dots, Y_n) = \underset{M \in \mathbb{R}}{\operatorname{argmin}} \sum_{i=1}^n \|Y_i - M\|$$

The generalized median called geometric median (GMed) is good as the median in one dimension with breakdown point of 50% (Bruce & Vershynin, 2011). However, there is need to discuss the properties of the GMed in Euclidean spaces and trying to quantitatively describe the robustness.

2.1.6 Mathematical Definition of Geometric Median Covariance

The geometric median of a discrete set of sample points in a Euclidean space is the point minimizing the sum of distances to the sample points. This generalizes the median, which has the property of minimizing the sum of distances for one-dimensional data, and provides a central tendency in higher dimensions.

Now, theorem and proofs.

Theorem 1. (Minsker, 2015). Let a set of points, $Y_1, Y_2, \dots, Y_n \in \mathbb{R}^n$, define a functional f by:

$$f(Y) = \sum_{i=1}^n \|Y_i - M\|_2.$$

Then, the derivative of f is:

$$\nabla f(Y) = - \sum_{i=1}^n \frac{(Y_i - M)}{\|Y_i - M\|_2}$$

Using the definition of the Euclidean norm the equation is as follows,

$$\frac{\partial}{\partial Y_j} \sum_{i=1}^n \|Y_i - M\|_2 = \frac{\partial}{\partial Y_j} \sum_{i=1}^n [\sum_{j=1}^n (Y_{ij} - M_j)^2]^{1/2} \quad (2.9)$$

which is equal to:

$$\sum_{i=1}^n \frac{-(Y_i - M_j)}{(\sum_{j=1}^n (Y_{ij} - M_j)^2)^{1/2}} = \sum_{i=1}^n \frac{-(Y_i - M)}{\|Y_i - M\|_2} \quad (2.10)$$

So that, the derivative of f in (2.10) is equal to (2.11)

$$\nabla f(Y) = -\sum_{i=1}^n \frac{(Y_i - M)}{\|Y_i - M\|_2}$$

Now by definition of GMed (Definition 1.1) we know that median M of $Y_1, Y_2, \dots, Y_n$ minimizes the function, f , we find an alternative characterization of the median.

Corollary 1.1. (Minsker, 2015). Given a set of points, $Y_1, Y_2, \dots, Y_n \in \mathbb{R}^n$ then $M \in \mathbb{R}^n$ is the median for this data set if and only if:

$$\sum_{i=1}^n \frac{(Y_i - M)}{\|Y_i - M\|_2} = 0$$

Assume M is the median for this data set then by definition M minimizes, $f(Y)$, and therefore we know that, $\nabla f(M) = 0$. So from Theorem 1, it found that $\nabla f(M) = 0$, then:

$$\sum_{i=1}^n \frac{(Y_i - M)}{\|Y_i - M\|_2} = 0 \quad (2.11)$$

Again by Theorem 1, we can conclude that, $\nabla f(M) = 0$. Since f is a convex function we know then that M minimizes, f , and therefore by definition M is the median of, $Y_1, Y_2, \dots, Y_n$. This characterization of GMed points to the interesting fact that median is based on not necessarily on how far apart points distributed, but instead on how they are distributed (Bruce & Vershynin, 2011).

2.1.7 Robust Properties of the Geometric Median

The concept of robust estimators validates the idea that no contaminant should affect the measures of location. The measure of robustness of an estimator is the breakdown point. Clearly, the GMed has a breakdown point of 50% which make the GMed as one of the best estimation of location which means that half of the data sets needs to be contaminated in order to contaminate this estimator (Efteliogu, 2015).

Additionally, the existence and uniqueness of the GMed in $\mathbb{R}^d$ follows directly from the convexity of the distance function. In one dimensional data set the GMed is the point that divides the point set into equal half on either side if n is odd and is any point on the line segment connecting the two middle points if n is even.

In order to generalize the median in one dimension to achieve a good performance and robust estimator based on the use of GMed, there is a need to consider the following properties (Minsker, 2015):

- i. For the one dimensional case, the GMed corresponds with the median. This is because the univariate median also minimizes the sum of distances from the point (Haldane 1948).
- ii. The GMed is unique whenever the points are not collinear i.e., three points are said to be collinear if they lie on the same line, this points usually determine a plane, but in the case of collinearity this does not happen (Vardi & Zhang 2000).
- iii. The GMed is equivariant for Euclidean similarity transformations, as well as translation and rotation (Haldane, 1948; Lopuhaa & Rousseeuw, 1991).

- iv. If it means that one would get the same result either by transforming the GMed, or applying the same transformation to the sample data and finding the GMed of the transformed data set.
- v. The next property is GMed has a breakdown point of 50% (Lopuhaa & Rousseeuw, 1991). That is, up to half of the sample data may be arbitrarily contaminated, and the median of the samples will still provide a robust estimator for the location of the uncontaminated data sets.

The literature review is on estimating a covariance matrix using regularization of the sample covariance matrices to resolved the singularity problems faced by M-test.

2.2 Regularizing the Sample Covariance Matrices

In the covariance matrix estimation, regularizing the covariance or computing the sparse of the covariance is very vital. It usually lead to computational challenging problems (Allen, 2012) and the use could result misleading conclusion (Fan, Liao & Mincheva, 2013).

Therefore, the banding, tapering and thresholding the sample covariance matrix can enhance their performance, statistical effectiveness and easy computation (Bickel & Levina, 2008b; Engel, Buydens & Blanchet, 2017). Besides, regularized estimation of covariance matrices have been extremely successful in high dimensional analysis in the justification of unpleasant and poorly modelled inverse problems (Runcie & Mukherjee, 2013).

A problem of statement is poorly presented if the number of observation to be compared is relatively larger than the sample size. In this case, the parameter estimate can be highly unstable, giving rise to high variance.

Using regularization method is attempt to improve and normalize the estimate of the covariance matrix, biasing them away from their sample based standard toward values that are considered to be more truly reasonable. Regularization of the estimation of large covariance matrices are applicable in the following areas, as example, used the principal components and eigenvalues to develop an understanding for low-dimensional data illustration (Pourahmadi, 2011); principal component analysis was used to regularized the estimation of large covariance matrices (Croux & Haesbroeck, 2000); construction of linear discriminant functions for classification of normal data (Hu, Wu, Yin, Gu, Liu, Xiao, & Yu, 2018); establishing independence and conditional independence relations between components using exploratory data analysis and testing (Cheng, Levina, Wang & Zhu, 2014) and setting confidence intervals on linear functions of the means of the components (Xie, Xue & Strickler, 2018).

In other words, poor performance of the sample covariance matrix in high dimensional situation and its eigenvalues call for how to deal with the singularity problem (Pourahmadi, 2013). Furthermore, covariance matrix play a significant role in high dimensional data analysis, including the principal component analysis (Pourahmadi, 2013), discriminant analysis (Bilgrau, Eriksen, Dybkaer & Bøgsted, 2015), clustering analysis (Sheikholeslami, Razavi, Gupta, Becker & Haghnegahdar, 2018), and regression analysis (Lan, Luo, Tsai, Wang & Yang, 2015).

Consequently, estimating a high dimensional covariance matrix and its inverse is a crucial problem in many applications. Working on the fundamental high dimensional covariance structure, we may face with the singularity problem if the spreading of the data have much more zero or small entries. The linear relationship between columns or rows of the covariance matrix, resulting in zero determinant in its calculation, and inverse of the underlying matrix become intractable, which lead to a computational problem in the likelihood (Ayyildiz, Gazi & Wit, 2012).

Additionally, various methods for improving high dimensional covariance estimators have been suggested by assuming the basic knowledge about the covariance matrix such as sparsity (Kurtz, Mueller, Miraldi, Littman, Blaser & Bonneau, 2015), graphical (Tarr, Müller & Weber, 2015), factor model (Pati, Bhattacharya, Pillai, & Dunson, 2014), and order among the underlying variables (Cai, Ren & Zhou, 2014). Estimating the covariance matrix involved the eigenstructure of the matrix.

There are a number of methods to unravel the fundamental challenge. The most well-known methods that can deal with singularity of covariance matrix determinant and inversion (Pourahmadi, 2013) will be discussed in section 2.2.1. The methods include the banding method (Bickel & Levina, 2008a; Wu & Pourahmadi, 2009; Bien, Bunea & Xiao, 2016), tapering method (Furrer & Bengtsson, 2007; Cai, Zhang & Zhou, 2010; Yi & Zou, 2013; Xue & Zou, 2014); thresholding method (Bickel & Levina, 2008b; Cai & Liu, 2011a, Fryzlewicz, 2013; Cao, Lin & Li, 2018).

However, the performance of any estimator depends heavily on the quality of tuning parameter selection. There are matrix norms which can be used to measure the accuracy of the estimation these are Operator norm and Frobenius norm (Fang, Wang & Feng, 2016).

Given any matrix $M_{p \times p} = [m_{i \times j}]$, its Operator norm and Frobenius norm can be defined as:

$$\|M\|_{op} = \sup\{\|Mx\|_2 : \|x\|_2 = 1\}$$

and

$$\|M\|_F = \sqrt{\sum_{i=1}^p \sum_{j=1}^p m_{ij}^2} \quad (2.17)$$

respectively, where $\|\cdot\|_2$ is the Euclidean norm for vectors.

If the estimate Σ were known, for any estimator, $\hat{\Sigma}(\lambda)$, we would select λ as:

$$\lambda_{op}^{oracle} = \underset{\lambda}{\operatorname{argmin}} \| \hat{\Sigma}(\lambda) - \Sigma \|_{op} \text{ or } \lambda_F^{oracle} = \underset{\lambda}{\operatorname{argmin}} \| \hat{\Sigma}(\lambda) - \Sigma \|_F \quad (2.18)$$

In this study we shall use the operator norm to express the measure accuracy in the analysis. In practice, we attempt to estimate the operator risk first,

$$R_{op}(\lambda) = E \| \hat{\Sigma}(\lambda) - \Sigma \|_{op}^2 \quad (2.19)$$

And then select a value for, λ . The operator norm also reduce the computational complexity (Bickel & Levina, 2008a, Yu & Yiquan, 2018). Bickel and Levina (2008a) obtained the estimates which are consistent at various rates in the operator norm. This infers that maximal and minimal eigenvalues of the estimates and Σ_p are close.

2.2.1 High Dimensional Datasets

The collection of developmental data such as financial, tools generation and media stored data has not simply increase the large number of data however, driven to tense rise of dimensional of the sample (Ma, Wu, Wang, Huang, Ranjan, Zomaya & Jie, 2015). The instance of new spaces like bioinformatics, WWW, ecommerce, health care industry, computer vision etc., has increase the high dimensional data (Ramanathan & Bangali, 2016). In the meantime, high dimension data sets is defined as when there are large variable numbers than the sample size number (Breiman, 2017).

In statistical study, the field of high-dimensional data whose dimension is larger than sample size is considered in classical multivariate analysis. In many applications, the dimension of the datasets may be larger than the sample size. Dimensionality in statistics refers to how many features a dataset has. For example, healthcare data is notorious for having huge amounts of variables e.g. blood pressure, weight, cholesterol level. i.e. Dimensionality means something slightly different in other areas of mathematics and science. In this study three methods of regularizing the high dimensional data sets will be discussed.

2.2.2 Banding Method

The classical statistical techniques cannot resolve high dimension $\mathbf{Y}$ since both rely on inverse of $\mathbf{Y}'\mathbf{Y}$ which could be singular or not well conditioned (Bajwa, Duarte & Calderbank, 2014). Many entries of the sample covariance matrix $\mathbf{S} = s_{ij}$ could be small or unstable in the small n , large p case. However, the most extreme case of this occur in time series analysis (Rydell, 2013), longitudinal (Allen, 2015), and spectroscopic data (James & Sugar, 2003). Banding the sample covariance matrix is a method by which we can obtain a well-conditioned and non-singular covariance estimator (Pourahmadi, 2013). It start by estimating the covariance matrix by a diagonal matrix and then successively add the other sub-diagonals, by estimating the first, second,..., k -th sub-diagonals.

The key question in implementing banding is how to choose the tuning or banding parameter k . Furthermore, the general use of the sample covariance matrix in the conventional multivariate procedures is largely due to sample covariance matrix being a consistent estimator of covariance matrix when p is relatively large as compared to the small sample size.

However, for high dimensional data such that, $\frac{p}{n} \rightarrow c \in (0, \infty)$, it is known that the eigenvalues of the sample covariance matrix are no longer consistent to their population counterpart, as demonstrated (Berthet & Rigollet, 2013; Deng & Tsui, 2013; Soloveychik & Wiesel, 2014; Qiu & Chen, 2014; Dobriban, 2015; and Danaher, Paul & Wang, 2015). It mean that the sample covariance is no longer reliable to the covariance matrix Σ that hindered applications of many conventional multivariate statistical procedures for high dimensional data.

Therefore, constructing covariance estimators by banding the sample covariance matrix, can help to overcome the problem. Wu and Pourahmadi (2003) considered banding the Cholesky factor matrix through the kernel smoothing estimation that was further developed by Rothman, Levina and Zhu (2010). Bickel and Levina, (2008a) proposed banding the sample covariance matrix directly for estimating covariance matrix and banding the Cholesky factor matrix for estimating inverse of the covariance matrix. It confirmed that both estimators are reliable to Σ and Σ^{-1} , respectively, for some bandable classes of covariance matrices.

Banding the sample covariance matrix does not perform better numerically in terms of positive definiteness compared to banding the Cholesky factor because of its good property and low computational cost. The high dimensional case, when $p > n$ still remains an open problem.

It show either banding sample covariance matrix or estimating a banded version of the inverse population covariance matrix, the problem of positive definiteness still remains (Bickel and Levina, 2008a). However, Kaufman, Schervish and Nychka (2008) pointed out that positive definiteness can be preserved by tapering the covariance matrix.

In the earlier work of Xue and Zou (2012) it was concluded that banding is a very useful regularization method for covariance estimation when the variables have a natural order and off-diagonal entries of the target covariance matrix decay to zero as they move away from the diagonal that is strongly supported. On the contrary, considered the case where the degree of the entries of Σ_0 decay at a polynomial rate of their distance from the diagonal and the result showed that banding the sample covariance matrix or $\mathbf{S}$, leads to well behaved estimates. Tapering the sample covariance matrix is another method of regularization in the estimation of covariance matrices.

2.2.3 Tapering Method

Tapering the sample covariance matrix has a long history in time series analysis. It has been used recently to improve the performance of linear discriminant analysis (Bickel, and Levina, 2004; and Pourahmadi, 2013) and Kalman filter bands (Furrer et al, 2007). The basic idea of tapering the covariance matrices is a popular method for reducing the computational cost of spatial prediction and parameter estimation for normal process models.

Given a tapered estimator of the sample covariance matrix for a tapering matrix $K = (k_{ij})$ replaces $\mathbf{S}$ by S_k , where S_k is the tapered sample covariance matrix with parameter, K . This will purposefully introduce zeros into the covariance matrix to make it sparse that must be maintained positive definiteness.

$$S_k = \mathbf{S} * K = (s_{ij}k_{ij}), \quad (2.20)$$

Where, $(*)$ denotes the Schur coordinate-wise matrix multiplication. When $K = (k_{ij})$ is a positive definite symmetric matrix, then S_k as the Schur product of two positive definite matrices. More importantly, Kaufman, Schervish & Nychka, (2008) pointed out that positive definiteness can be well maintained by tapering the covariance matrix.

Furthermore, Cai, Zhang and Zhou, (2010) suggested a tapering estimator that can be viewed as a soft banding on the sample covariance that was designed to improve the banding estimator of Bickel and Levina, (2008b). It demonstrated that the tapering estimator attains the optimal minimax rates of convergence for estimating the covariance matrix. Wagaman and Levina (2009) developed a method for discovering meaningful orderings of variables such that tapering can be applied.

Subsequently, Cai, Zhang and Zhou, (2010) established the optimal rates of convergence for estimating the covariance matrix under both the operator norm and Frobenius norm. It is shown that optimal procedures under the two norms are different, therefore matrix estimation under the operator norm is fundamentally different from vector estimation.

The minimax upper bound is obtained by constructing a special class of tapering estimators and by studying their risk properties as a key step in obtaining the optimal rate of convergence is the derivation of the minimax lower bound.

The tapered estimator of sample covariance matrix has good numerical performance and nearly outperform the banding estimator, but does not improve the bound under Frobenius norm, so that under this norm banding performs as good as tapering (Pourahmadi, 2013). The last method for regularizing the covariance matrix that will be discussed is thresholding procedures.

2.2.4 Thresholding Method

The technique of thresholding, initially developed in nonparametric function estimation, has been used in the estimation of large covariance matrices Bickel and Levina, 2008b; EL Karoui, 2010; Rothman, Levina and Zhu, 2009; and Pourahmadi, 2013). Unlike banding and tapering, thresholding does not required the variable to be ordered so that the estimator is invariant to permutation of the variables.

For a sample covariance matrix, $\mathbf{S} = (s_{ij})$, the thresholding operator λ for $\lambda \geq 0$ defined by,

$$T_{\lambda}(\mathbf{S}) = [s_{ij}I(|s_{ij}| \geq \lambda)], \quad (2.21)$$

The threshold $\mathbf{S}$ at λ will replaced all zero entries with absolute value less than λ . T_λ , now preserves symmetry, and is invariant under permutations of variable labels, but does not necessarily preserve positive definiteness. Its biggest advantage is the simplicity. It carries no major computational burden compared to its competitors such as penalized likelihood with the *Lasso* penalty (Huang, Liu, Pourahmadi & Liu, 2006; Rothman, Levina, Zhu, 2009; and Pourahmadi, 2013).

Furthermore, Cai and Liu, (2011) considered estimation of sparse covariance matrices and suggested a thresholding procedure that is adaptive to the variability of individual entries. The encouragement of the proposed procedure was based on works so far that belong to the class of El Karoui (2010) and Bickel and Levina, (2008a) that proposed thresholding of the sample covariance matrix Σ_n and the rates of convergence. Also, Rothman et al, (2009) considered thresholding of the sample covariance matrix with more general thresholding functions.

Subsequently, Xue and Zou (2012) and Steffens, Pesavento and Pfetsch (2018). established the minimax rates of convergence under the matrix operating norm and the spectral norm. Then, Wang and Zou (2010) considered the estimation of volatility matrices that is based on high dimensional situation.

Bickel and Levina, (2008b) pioneered work on thresholding the sample covariance and suggest regularization by thresholding the estimation of large covariance matrices. But, one of its major advantages is its simplicity, hard thresholding carries no computational burden, unlike many other methods for covariance regularization.

A theoretical justification for the cross-validation approach to selecting the threshold is provided. Furthermore, it shows that the results are uniform over relations of covariance matrices that satisfied a fairly natural notion of sparsity.

Moreover, Zhong, Chen and Xu (2013) considered two alternative tests of Donoho and Jin (2004) for high dimensional means with sparse nonzero means for a sub-normally distributed data, and unknown column-wise dependence. The analysis based on L_1 and L_2 statistics as an alternative formulations to the Higher Criticism (HC) test shows is more powerful when the pair is away from the detection boundary, corresponds with that of the HC test that was earlier discovered in Donoho and Jin (2004).

The tests shows that the powers were insensitive to level of dependence p , this later affirmed the finding that thresholding largely removed the dependence values. Among several component wise regularization methods for estimating sparse or approximately sparse covariance matrix, thresholding is the most common method.

In general, threshold estimators have a very low computational cost and good rates of convergence under high dimensional asymptotic, but does not guarantee positive definite eigenvalues in finite samples, since positive definite in finite samples is desirable.

Therefore, this study will use the banding and tapering methods of regularization to estimate the covariance matrices to resolve singularity problem. Both methods enhance consistent performance, computational efficiency, and better statistical interpretability of the covariance estimation results (DuVal, Vanderbilt & M'Gonigle, 2018 and Kang & Deng, 2018). Therefore, the detail about the steps for computing parameter k by cross validation method is discussed in Appendix C.

2.3 Summary Review of the Literature

This literature review, appraised work based on testing the equality of several covariance matrices using the M-test to solve the outlier and the singularity problems in high dimensional datasets. Based on the limitations and weakness of the original statistical test, further suggest to solve these problem by using a robust statistical estimator of the geometric median covariance method and intergrating the banding and tapering methods. The major contribution to knowledge in this study is to replace the mean covariance with the robust geometric median covariance, so as to solve the two main problems in high dimensional data sets.

CHAPTER THREE

RESEARCH METHODOLOGY

3.1 Introduction to the Chapter

The purpose of the study is to develop a new robust high dimensional M-test to be used in testing the equality of several independent sample of covariance matrices. At the beginning of this chapter, the detail of research procedure carried out is presented. It is followed by the overall research framework implemented. Then, Section 3.4.1 until 3.4.2 reviewed the procedure of new robust tests of covariance equality. Section 3.5 present the performance evaluation of the modified M-test whereas Section 3.6 illustrates the application of US subprime financial crisis on the forex using covariance equality test.

3.2 The Research Procedure

The M-test mainly used for testing the equality of several covariance matrices. But the test is highly sensitive to outliers and faced with the singularity problems in high dimensional datasets. In order to resolve these two problems, firstly, a robust estimator attempt to overcome the outliers problems using the geometric median covariance. Secondly, banded and tapered geometric median are used to overcome the singularity problems. Henceforth, the new statistical test is applied into real data by using currency exchange rate for the case of US subprime crisis. In achieving this goal, the following research procedure need to be accomplished.

In general, there are five (5) stages of research procedures carried out in the study:

Stage 1 Integrate M-test into robust location estimators of mean, geometric median by using M-test and M_{GMed} -test.

Stage 2 Examine the new statistical robust covariance equality test using regularization method by resolving the singularity problems in high dimensional data sets.

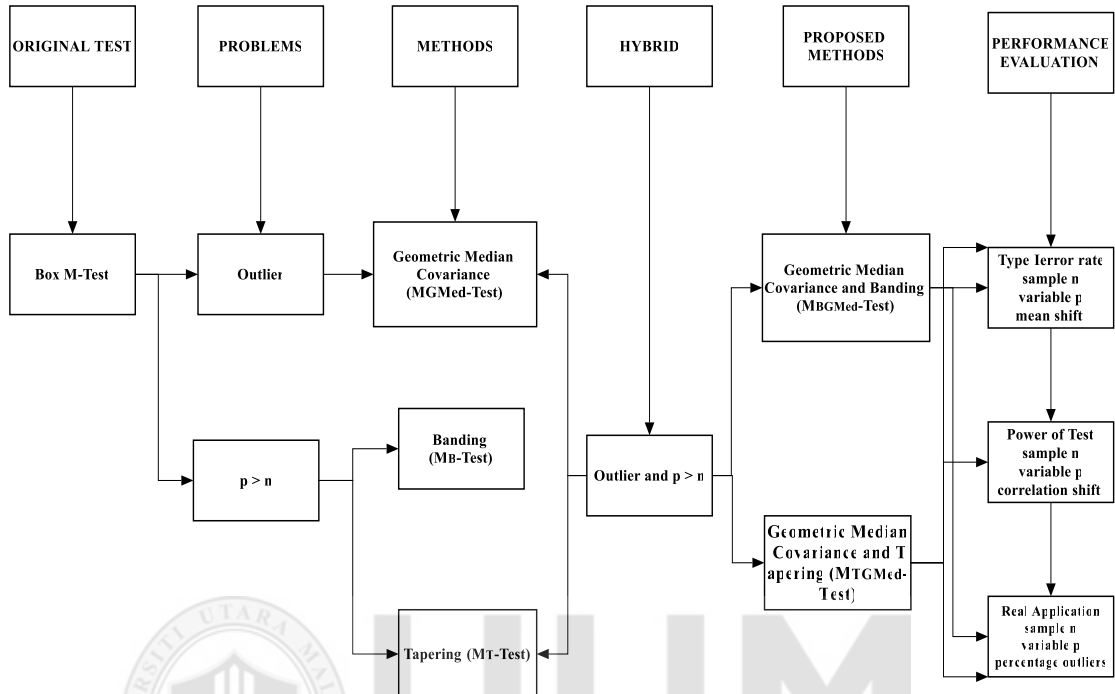
Stage 3 Compare the performance of M-test, M_{GMed} -test, M_B -test, M_T -test, M_{BGMed} -test, and M_{TGMed} -test based on Type I error rates.

Stage 4 Compare sensitivity of covariance shift performance of M-test, M_{GMed} -test, M_B -test, M_T -test, M_{BGMed} -test, and M_{TGMed} -test based on the power of test values.

Stage 5 Apply the new statistical test into real data by using currency exchange rate using two different methods in testing the equality of covariance matrices:

- i. To test two independent samples of covariance matrices; comparing before and after US subprime financial crisis.
- ii. To test several independent samples of covariance matrices using a control chart.

3.3 The Overall Research Framework with Flow Chart



3.4 The Procedure of New Robust Tests of Covariance Equality

The first step of this study is focused on modifying M-test. The M-test is modified using geometric median covariance to the corresponding robust estimators as displayed in Figure 3.1.

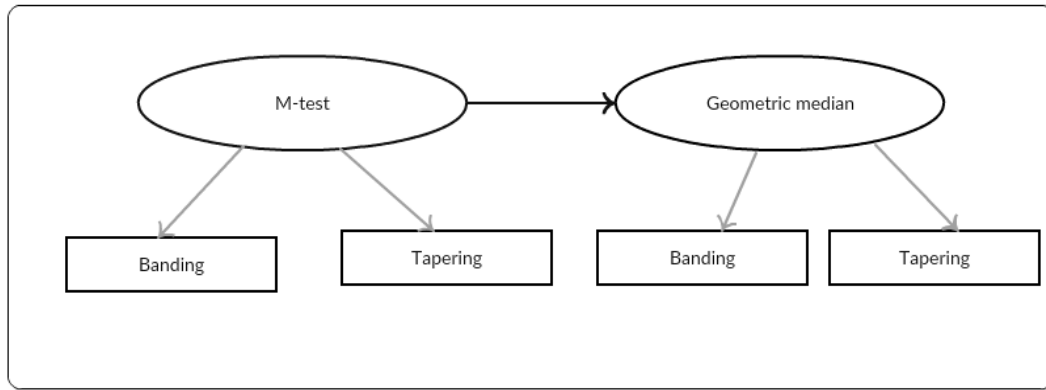


Figure 3. 1. Statistical test with corresponding M-test and modified M-tests

From Figure 3.1, it show that the process of M-test modification. Firstly, to resolve the problems of singularity, M-test is modified using banding and tapering methods are called M_B -test and M_T -test, respectively. Secondly, geometric median covariance is used to replace the mean covariance in handling the outlier problems. Then, M-test is modified using the robust estimator and is called M_{GMed} -test. Furthermore, intergrate the robust estimator with banding and tapering methods to solve the occurrence of that two problems simultanously. The new robust tests of covariance equality are produced called M_{BGMed} -test and M_{TGMed} -test.

3.4.1 Mean, Median and Geometric Median Covariance

Most multivariate statistical methods greatly depend on estimators which are measure of location and dispersion such as sample mean vector and sample covariance. The sample mean vector and sample covariance are the most commonly used estimators when the data is multivariate normal distribution.

However, the measure of location and dispersion is affected by outlier(s) and leads to incorrect interpretation of the data sets. For proper understanding of various estimators of location. Show the computations of the covariance matrix:

Let X be the data matrix

$$X = \begin{bmatrix} X_{1i} & X_{1j} \\ X_{2i} & X_{2j} \\ X_{3i} & X_{3j} \\ X_{4i} & X_{4j} \end{bmatrix}$$

The mean vector is,

$$\bar{X} = \begin{bmatrix} \bar{X}_i \\ \bar{X}_j \end{bmatrix} \text{ and,}$$

The sample covariance is,

$$S^2_k = S_{kk} = \frac{1}{n} \sum_{j=1}^n (x_{ik} - \bar{x}_k)^2, \text{ where } k = 1, 2, \dots, p$$

where $k = 1, 2, \dots, p$

$$S_{ik} = \frac{1}{n} \sum_{j=1}^n (x_{ji} - \bar{x}_i) (x_{jk} - \bar{x}_k), \quad i = 1, 2, \dots, p, \quad \text{where } k = 1, 2, \dots, p, \quad \text{can be}$$

written in the matrix form as,

$$\begin{bmatrix} S_{ii} & S_{ij} & \cdots & S_{ip} \\ S_{ji} & S_{jj} & \cdots & S_{jp} \\ \vdots & \vdots & \ddots & \vdots \\ S_{pi} & S_{pj} & \cdots & S_{pp} \end{bmatrix}$$

To show the computation of the sample median covariance matrix

Let Y be the data matrix

$$Y = \begin{bmatrix} Y_{1l} & Y_{1j} \\ Y_{2l} & Y_{2j} \\ Y_{3l} & Y_{3j} \\ Y_{4l} & Y_{4j} \end{bmatrix}$$

The sample median covariance

$S_{ll}^2 = S_{ll} = \frac{1}{n} \sum_{m=1}^n (y_{ml} - \tilde{y}_l)^2$, where $l = 1, 2, \dots, p$ and $\tilde{y} = \text{median}$ (i.e. the middle of a sorted list of numbers).

Therefore, $S_{jl} = \frac{1}{n} \sum_{m=1}^n (y_{mj} - \tilde{y}_j) (y_{ml} - \tilde{y}_l)$, $j = 1, 2 \dots p$, $l = 1, 2, \dots p$

$$S_{jl} = \frac{1}{n} \sum_{m=1}^n (y_{mj} - \tilde{y}_j) (y_{ml} - \tilde{y}_l)$$

$$\text{sample medians } \tilde{Y} = \begin{bmatrix} \tilde{y}_j \\ \tilde{y}_l \\ \vdots \\ \tilde{y}_n \end{bmatrix}$$

sample variances and covariances

$$S = \begin{bmatrix} S_{jj} & S_{jl} & \cdots & S_{jp} \\ S_{lj} & S_{ll} & \cdots & S_{lp} \\ \vdots & \vdots & \ddots & \vdots \\ S_{pj} & S_{pl} & \cdots & S_{pp} \end{bmatrix}$$

$$\tilde{Y} = \begin{bmatrix} \tilde{y}_j \\ \tilde{y}_l \end{bmatrix}$$

The sample variance and covariance of median

$$\begin{aligned} S_{jj} &= \frac{1}{n} \sum_{m=1}^n (y_{mj} - \tilde{y}_j)^2 \\ &= \frac{1}{n} ((y_{mj1} - \tilde{y}_{j1})^2 + (y_{mj2} - \tilde{y}_{j2})^2 + \cdots + (y_{mjn} - \tilde{y}_{jn})^2) \end{aligned}$$

$$\begin{aligned} S_{ll} &= \frac{1}{n} \sum_{m=1}^n (y_{ml} - \tilde{y}_l)^2 \\ &= \frac{1}{n} ((y_{ml1} - \tilde{y}_{l1})^2 + (y_{ml2} - \tilde{y}_{l2})^2 + \cdots + (y_{mln} - \tilde{y}_{ln})^2) \end{aligned}$$

$$S_{jj,jl} = \frac{1}{n} \sum_{m=1}^n (y_{mj} - \bar{y}_j) (y_{ml} - \bar{y}_l)$$

$$S_{jj,jl} = \frac{1}{n} ((y_{mj1} - \tilde{y}_{j1})(y_{ml1} - \tilde{y}_{l1}) + (y_{mj2} - \tilde{y}_{j2})(y_{ml2} - \tilde{y}_{l2}) \\ + \dots + (y_{mjn} - \tilde{y}_{jn})(y_{mln} - \tilde{y}_{ln}))$$

Then compare,

$$S_{lj} = S_{jl} \text{ and } S_n$$

The GMed will be computed based on its properties which is one of the possible extensions of a univariate median to higher dimension (Minsker, 2015). Moreover, Kemperman (1987), defined geometric median as a uniquely random variable that is not concentrated on a straight line, with the Hilbert space definitions. The computation of geometric median covariance are stated below:

Let X be the data matrix

$$X = \begin{bmatrix} X_{11i} & X_{21i} \\ X_{12i} & X_{22i} \\ X_{13i} & X_{23i} \\ X_{14i} & X_{24i} \end{bmatrix}$$

The geometric median vector consist of the geometric median of each variable and the variance-covariance matrix consist of the variances of the variables along the main diagonal and the covariances between each pair of variables in the other matrix positions (Eftelioglu, 2015).

The GMed is computed from: Given data matrix X with X_1 and X_2 ,

$M(X_1) = \arg \min_{x \in \mathbb{R}^d} \sum_{i=1}^m \|m - x_{1i}\|$ for points $x_{11}, x_{12}, \dots, x_{1i} \in \mathbb{R}$ is $\mathbb{R}^d$ the median of set X_1 is defined as a point m so.

$M(X_2) = \arg \min_{y \in \mathbb{R}^d} \sum_{i=1}^m \|m - x_{2i}\|$ for points $x_{21}, x_{22}, \dots, x_{2i} \in \mathbb{R}$ is $\mathbb{R}^d$ the median of set X_2 is defined as a point m .

The formula for computing the covariance of the variables is:

$$\text{geometric median sample} = \left[\begin{array}{c} \sum_{i=1}^n \frac{(x_{1i} - m(x_{11}))(x_{2i} - m(x_{21}))}{n-1} \\ \sum_{i=1}^n \frac{(x_{2i} - m(x_{21}))(x_{1i} - m(x_{11}))}{n-1} \end{array} \right]$$

Let X_1 be

$$X_1 = \sum_{i=1}^n \frac{(X_{1i} - m(x_{11}))(X_{2i} - m(x_{21}))}{n-1}$$

$$X_{1i} = \begin{bmatrix} X_{11i} \\ X_{12i} \\ X_{13i} \\ X_{14i} \end{bmatrix}$$

Geometric Median = $\arg \min_{x \in \mathbb{R}^d} \sum_{i=1}^m \|X_{1i} - m((x_{11}))\|$, where, $m((x_{11})) = \text{median}$

(i.e. the middle of a sorted list of numbers of variable, X_{1i} . Then,

$$m(x_{11}) = \frac{X_{13i} + X_{12i}}{2} = m(x_{11})$$

$$\sum_{i=1}^m \| X_{1i} - m(x_{11}) \| =$$

$$\| X_{11i} - m(x_{11}) \| + \| X_{12i} - m(x_{11}) \| + \| X_{13i} - m(x_{11}) \| + \| X_{14i} - m(x_{11}) \|$$

Let $\mathbf{X}_2$ be the second data matrix,

$$\mathbf{X}_2 = \sum_{i=1}^n \frac{(X_{2i} - m(x_{21}))(X_{1p} - m(x_{11}))}{n - 1}$$

$$\mathbf{X}_{2i} = \begin{bmatrix} X_{21i} \\ X_{22i} \\ X_{23i} \\ X_{24i} \end{bmatrix}$$

$\arg \min_{x \in \mathbb{R}^d} \sum_{i=1}^m \| X_{2i} - m(x_{21}) \|$, where, $m(x_{21}) =$ median (i.e. the middle of a sorted list of numbers of variable $\mathbf{X}_{2i}$, then,

$$m(x_{21}) = \frac{X_{23i} + X_{22i}}{2} = m(x_{21})$$

$$\sum_{i=1}^m \| X_{2i} - m(x_{21}) \| =$$

$$\| X_{21i} - m(x_{21}) \| + \| X_{22i} - m(x_{21}) \| + \| X_{23i} - m(x_{21}) \| + \| X_{24i} - m(x_{21}) \|^2$$

$$\sum_{i=1}^n \frac{(X_{2i} - m(x_{21}))(X_{1i} - m(x_{11}))}{n - 1}$$

$\sum_{i=1}^m \| X_{2p} - m(x_{21}) \|^2$ sample geometric median covariance

$$S^2_{X_1} = S_{X_1 X_1} = \sum_{i=1}^n \frac{((X_i - m(x_{11}))(X_{2p} - m(x_{21})))^2}{n-1},$$

$$S^2_{X_2} = S_{X_2 X_1} = \sum_{i=1}^n \frac{((X_{2p} - m(x_{21}))(X_{1i} - m(x_{11})))^2}{n-1}, \text{ where } x_{11} = 1, 2, \dots, i \text{ and}$$

$$x_{21} = 1, 2, \dots, p$$

Therefore,

$$\text{geometric median sample} = \left[\frac{\sum_{i=1}^n \frac{((X_{1i} - m(x_{11}))(X_{2p} - m(x_{21})))^2}{n-1}}{\sum_{i=1}^n \frac{((X_{2p} - m(x_{21}))(X_{1i} - m(x_{11})))^2}{n-1}} \right]$$

geometric median sample variances and covariances

$$S = \begin{bmatrix} S_{ii} & S_{iX_2} & \cdots & S_{ip} \\ S_{iX_2} & S_{X_2 X_2} & \cdots & S_{X_2 p} \\ \vdots & \vdots & \ddots & \vdots \\ S_{pi} & S_{pX_2} & \cdots & S_{pp} \end{bmatrix}$$

$$\sum_{i=1}^n \frac{((X_{1i} - m(x_{11}))(X_{2p} - m(x_{21})))^2}{n-1}$$

$$= \frac{1}{n-1} [(X_{11i} - m(x_{11}))^2 + (X_{12i} - m(x_{11}))^2 + (X_{13i} - m(x_{11}))^2 + (X_{14i} - m(x_{11}))^2]$$

$$\sum_{i=1}^n \frac{((X_{2i} - m(x_{21}))(X_{1p} - m(x_{11})))^2}{n-1}$$

$$= \frac{1}{n-1} [(X_{2i} - m(x_{21}))^2 + (X_{i22} - m(x_{21}))^2 + (X_{i23} - m(x_{21}))^2 + (X_{i24} - m(x_{21}))^2]$$

$$\sum_{i=1}^n \frac{\left((X_{2p} - m(x_{21})) (X_{1i} - m(x_{11})) \right)^2}{n-1}$$

$$S_{X_1 X_1, X_2 X_2} = \left[\sum_{i=1}^n \frac{((X_{1i} - m(x_{11}))(X_{2p} - m(x_{21})))^2 + ((X_{2p} - m(x_{21}))(X_{1i} - m(x_{11})))^2}{n-1} \right]$$

Then compare,

$S_{X_2 X_1} = S_{X_1 X_2}$ and S_n . The computational details are shown in Appendix N, O and P respectively.

3.4.2 Robust M-Test Based on Geometric Median Covariance

The GMed was introduced by Fermat-Weber in 1929 for three points in two dimensional Euclidean space (Fletcher, Venkatasubramanian & Joshi, 2009; Minsker, 2015 and Zhang, Shi, Sun & Cheng, 2018) suggested the need for determining the maximum, minimum and tangents to curved lines in 1643 (Eftelioglu, 2015). The three points in the plane, find fourth point such that the sum of its distances to the three given points is a minimum. However, the problem was solved by Torricelli in the seventeenth century (De Freitas, 2018).

The solution is known to be the Fermat and Torricelli point of a triangle formed by three points. On the other hand, it is known that other mathematicians also worked on the same problem at the same time (Brazil, Graham, Thomas & Zachariasen, 2014).

The early robust estimator for geometric median vector and sample covariance matrix was long presented in 1937. Weiszfeld introduced three proofs concerning the uniqueness of the geometric median in which one of the proofs provide an algorithm for its computations which is based on a strict convexity argument.

Recently, Eftelioglu (2015) introduced the geometric median as a robust location estimator in statistics that can be applied to key domains such as, facility location problem in strategic planning which aims to minimize the cost of transportation, in spatial economics where the land values are determined by the bid rent theory, and in biology and health domains where the outliers in high dimensional data are needed to be eliminated.

The robustness properties of GMed are summarized under three important features, breakdown point, uniqueness and equivariance (Eftelioglu, 2015). Moreover, GMed is an important concept in statistics where the location of the data sets is required regardless of the outliers present with 50% breakdown point, it is equivariant under Euclidean similarity transformations and it is unique under the condition that the points are not collinear, and the number of points in the set is odd (Haldane, 1948).

Let M_{GMed} denotes the robust estimator based on geometric median. The robust variance-covariance of the GMed estimator, $S_{M_{GMed(i)}}$, where $i = 1, 2, \dots, m$ is substituted into Equation 1.1. in page 3.

Thus, the test is now turn into the following equation,

$$M = Y \sum_{i=1}^m (n_i - 1) \log | S_{M_{GMed(i)}}^{-1} S_{M_{GMed}} | \quad (3.1)$$

where:

$$Y = 1 - \frac{2p^2 - 3p - 1}{6(p+1)(m-1)} \left(\sum_{i=1}^m \frac{1}{n_i - 1} - \frac{1}{n - m} \right)$$

$$S_{M_{GMed}} = \frac{\sum_{i=1}^m (n_i - 1) S_{M_{GMed(i)}}}{n - m} \text{ is pooled sample covariance matrix for geometric median}$$

$S_{M_{GMed(i)}}$ is the sample covariance matrix unit calculated from sample i

m is the number of subgroup for geometric median

$$n = n_1 + n_2 + \dots + n_m ; n_i, i\text{-th sample size}$$

In the section 3.4.2.1, the integration of orginal test into several modified tests are carried out based on geometric median covariance with banding and tapering methods of regularization to solve the outliers and singularity problems in high dimensional data sets.

3.4.2.1 Integration of M-Test Based on Geometric Median Covarinace with Banding Method

Banding is a simple and systematic way of estimating large covariance matrices, it starts estimating the covariance matrix by a diagonal and successively adds the other sub diagonals by estimating the first, second, ..., k th sub diagonals. This suggests that using a diagonal covariance matrix or ignoring the off-diagonal the off-diagonal entries will lead to better performance of the test than estimating the whole covariance matrix (Pourahmadi, 2013). In this study, we integrate the M-test based on banding method.

Also, to develop a regularized M-test with banding denoted by M_B -test, the variance-covariance of banded $S_{M_B(i)}$ where $i = 1, 2, \dots, m$ is substituted into equation 1.1.

$$M = Y \sum_{i=1}^m (n_i - 1) \log | S_{M_B(i)}^{-1} S_{M_B} | \quad (3.2)$$

where:

$$Y = 1 - \frac{2p^2 - 3p - 1}{6(p+1)(m-1)} \left(\sum_{i=1}^m \frac{1}{n_i - 1} - \frac{1}{n - m} \right)$$

$$S_{M_B} = \frac{\sum_{i=1}^m (n_i - 1) S_{M_B(i)}^{-1}}{n - m} \text{ is pooled sample covariance matrix for banding method}$$

$S_{M_B(i)}^{-1}$ is the sample covariance matrix unit calculated from sample i

m is the number of subgroup for Banding method.

$n = n_1 + n_2 + \dots + n_m$; n_i , i -th sample size.

To develop a regularized robust GMed-test with banding denoted by M_{BGMed} , the variance-covariance of banded GMed $S_{M_{BGMed}(i)}$ where $i = 1, 2, \dots, m$ is substituted into Equation 1.1

$$M = Y \sum_{i=1}^m (n_i - 1) \log | S_{M_{BGMed}(i)}^{-1} S_{M_{BGMed}} | \sim \chi^2 \quad (3.3)$$

where:

$$Y = 1 - \frac{2p^2 - 3p - 1}{6(p+1)(m-1)} \left(\sum_{i=1}^m \frac{1}{n_i - 1} - \frac{1}{n - m} \right)$$

$$S_{M_{BGMed}} = \frac{\sum_{i=1}^m (n_i - 1) S_{M_{BGMed}(i)}^{-1}}{n - m} \text{ is pooled sample covariance matrix for banding}$$

geometric median

$S_{M_{BGMed}(i)}^{-1}$ is the sample covariance matrix unit calculated from sample i

m is the number of subgroup for banding geometric median

$n = n_1 + n_2 + \dots + n_m$; n_i , i -th sample size

Therefore, the detail about computation on banded parameter k , banded method mathematical formula validation and simple banding sample covariance matrix with k parameter (R software code) are discussed in Appendix D, E and F respectively.

3.4.2.2 Integration of M-Test Based on Geometric Median Covariance With Tapering Method

Tapering is a smooth version of banding where it gradually shrinks the off-diagonal entries toward zero. However, the banding and tapering estimators are not necessarily positive semi-definite (Choi, Lim, Roy & Park, 2016). A practical approach to avoid this is to assign the estimators to the space of positive semi-definite matrices (Chatterjee, 2015).

The resulting estimator is then positive semi-definite. The regularization method of M-test using tapering is denoted by,

$$M = Y \sum_{i=1}^m (n_i - 1) \log | S_{M_T(i)}^{-1} S_{M_T} | \quad (3.4)$$

where:

$$Y = 1 - \frac{2p^2 - 3p - 1}{6(p+1)(m-1)} \left(\sum_{i=1}^m \frac{1}{n_i - 1} - \frac{1}{n - m} \right)$$

$$S_{M_T} = \frac{\sum_{i=1}^m (n_i - 1) S_{M_T(i)}^{-1}}{n - m} \text{ is pooled sample covariance matrix for tapering method}$$

$S_{M_T(i)}^{-1}$ is the sample covariance matrix unit calculated from sample i

m is the number of subgroup for tapering method

$n = n_1 + n_2 + \dots + n_m$; n_i , i -th sample size

To develop a regularized robust GMed-test with tapering denoted by M_{TGMed} , the variance-covariance of tapered GMed, $S_{M_{TGMed}(i)}$ where $i = 1, 2, \dots, m$ is substituted into equation 1.1

$$M = Y \sum_{i=1}^m (n_i - 1) \log | S_{M_{TGMed}(i)}^{-1} S_{M_{TGMed}} | \quad (3.5)$$

where:

$$Y = 1 - \frac{2p^2 - 3p - 1}{6(p+1)(m-1)} \left(\sum_{i=1}^m \frac{1}{n_i - 1} - \frac{1}{n - m} \right)$$

$S_{M_{TGMed}} = \frac{\sum_{i=1}^m (n_i - 1) S_{M_{TGMed}(i)}^{-1}}{n - m}$ is pooled sample covariance matrix for tapering geometric median

$S_{M_{TGMed}(i)}^{-1}$ is the sample covariance matrix unit calculated from sample i

m is the number of subgroup for tapering geometric median

$n = n_1 + n_2 + \dots + n_m$; n_i , i -th sample size.

Therefore, the detail about the algorithm of M_{BGMed} -test and M_{TGMed} -test by regularization methods, simple tapering sample covariance matrix with l - parameter (R software code) and Mean, Median and Geometric Median covariance are discussed in Appendix A, G, H, I and J respectively.

In order to evaluate the performance of M-test and several modified M-tests, we carried out simulation study on Type I error rate and power of test value, based on variables, p and sample size, n , with 5,000.

3.5 Evaluation of the Modified M-Test

There are two measures of evaluation performed on M-test and several modified M-test which are: the Type I error rate and power of test. Based on all equations presented in section 3.3 which are Equation. 1.1, 3.2, 3.3, 3.4, and 3.5, are evaluated by using Type I error (α) rate and power test using simulation study. Generally, the multivariate random data is generated using computer programming to compute the covariance matrices for producing the statistical tests. The process is replicated for 5,000 times to compute the Type I error rate and power of test by considering several manipulated variables.

3.5.1 Type I Error Rate

In order to determine the robustness, we considered the Bradley's criterion as a reference for the purpose of comparison and checking on the robustness level. There are three criteria for robustness which are stringent, moderate, and liberal defined as $\alpha \pm 0.1\alpha$, $\alpha \pm 0.2\alpha$, and $\alpha \pm 0.5\alpha$, respectively (Bradley, 1978; Nordstokke & Colp, 2014). We considered a robust test if the Type I error rates lies in either one of the criteria. Moreover, the closer the value to, α , the more robust the test (Maurer, Jones & Chen, 2018).

We considered a robust test if the Type I error rates lies in either one of the criteria. Moreover, the closer the value to, α , the more robust is the test, if the value is less than 0.05, we reject the null hypothesis that there is no difference between the means and conclude that a significant difference does exist (Gilleland, Hering, Fowler & Brown, 2018).

The study used the Bradley's (1978) of robustness to assess the performance of a statistical test to control its probability of Type I error rate. According to the Bradley's liberal criterion of robustness, a test is considered robust when the empirical rate of Type I error rate fall between 0.5α and 1.5α (Ahad, Othman & Yahaya, 2012). Thus, when the significance level is set at, $\alpha = 0.05$, a test is robust when the Type I error rate values lies between 0.025 and 0.075. Otherwise, a test is considered to be non-robust. However, if $\alpha = 0.05$, then the confidence level is 0.95. if we increase α , both will increase the probability of incorrectly rejecting the null hypothesis and decrease the confidence level (Bougioukas & Haidich, 2018).

3.5.2 Power of Test ($1 - \beta$)

The concept of power is used to make a comparison between different statistical testing procedures where the most powerful test will have the higher number of rejection of the null hypothesis (Mittelhammer & Mittelhammer, 1996). However, this study conducting a comparison between original M-test and several modified M-test (M_{GMed} -test, M_B -test, M_T -test, M_{BGMed} -test and M_{TGMed} -test).

Generally, the power of a test or known as $1 - \beta$ is the probability of correctly rejecting the null hypothesis when it is false. It is actually referred to the sensitivity of a statistical test and the ability to detect a true.

This is because the high power is very commendable. More importantly, highly powered research is often able to detect small effect as statistically significant (Sawyer & Sawyer, 1982). The significance level (α), sample size (n) and the effect size of population (N) were essential for statistical power (Cohen, 1992). In general, the power that is greater than or equal to 0.8 is acceptable. This means that the chance of results is statistically significantly difference would be at least 80% or greater. According to Yue, Pilon and Cavadias, (2002), the power can be estimated by

$$\text{Power} = \frac{R_n}{U} \quad (3.6)$$

where,

R_n , is the number of experiments that fall in the rejection region

U , is the total number of replication in the simulation experiments

The simulation study is done based on various conditions which are number of variables (p), sample sizes (n), percentages of outliers, epsilon (ϵ), scale amount of contamination Kappa (k) and mean shift (μ) (Ramey, 2013). Executed on 5,000 replications to obtain an accuracy of a power approximation (Sharif, Ismail & Omar, 2017).

Let $\sigma^2 = (\sigma_1, \sigma_2, \dots, \sigma_p)$ represent the population standard deviation vector of the p variables where $\sigma_i = \sqrt{\sigma_{ii}}$ for $i = 1, 2, \dots, p$ and σ_{ii} is the corresponding variance of i -th variable.

The study considered in each data set, it consists of different number of variables such as: $p = 3, 5, 8, 10, 15, 20, 25, 30$ and 50 as well as different size of sample, $n = 2, 5, 10, 15, 20, 25$ and 50 (He, Zhang, Avery & Lin, 2016), variable, $p = 50$ is large compare to small sample size. The simulation study is conducted using the R-package coding. Several codes were written to run the test based on the conditions provided for all pairs of variables and sample size, this study uses same covariance shift ρ , where $\rho = 0, 0.5, 0.7$ and 0.9 (Mason, Chou & Young, 2009). These ρ values are representatives of no covariance, moderate covariance and high covariance respectively. The three types of covariance shifts are adopted from Reynolds and Gho (2006).

Where,

- i. all p , components in σ increase with the same quantity,
- ii. only one component in σ increases while other $(p - 1)$ components remain the same,
- iii. $\frac{p}{2}$, components increase and other $\frac{p}{2}$ components decrease.

For example, size of shift is set at $t = 1.01$ (Mason, Chou & Young, 2009), then hypothesis testing is: $H_0: \Sigma_m = \Sigma_0$ versus $H_1: \Sigma_m = \Sigma_1$ where, $\Sigma_0 = I_p$ and Σ_1 are as follows,

$$\text{i.} \quad \Sigma_1 = \begin{bmatrix} 1.01 & 0 & \cdots & 0 \\ 0 & 1.01 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 1.01 \end{bmatrix} \text{ for } \rho = 0$$

$$\text{ii.} \quad \Sigma_1 = \begin{bmatrix} 1.01 & 0.2 * 1.01 & \cdots & 0.2 * 1.01 \\ 0.2 * 1.01 & 1.01 & \cdots & 0.2 * 1.01 \\ \vdots & \vdots & \ddots & \vdots \\ 0.2 * 1.01 & 0.2 * 1.01 & \cdots & 1.01 \end{bmatrix} \text{ for } \rho = 0.2$$

$$\text{iii.} \quad \Sigma_1 = \begin{bmatrix} 1.01 & 0.5 * 1.01 & \cdots & 0.5 * 1.01 \\ 0.5 * 1.01 & 1.01 & \cdots & 0.5 * 1.01 \\ \vdots & \vdots & \ddots & \vdots \\ 0.5 * 1.01 & 0.5 * 1.01 & \cdots & 1.01 \end{bmatrix} \text{ for } \rho = 0.5$$

$$\text{iv.} \quad \Sigma_1 = \begin{bmatrix} 1.01 & 0.7 * 1.01 & \cdots & 0.7 * 1.01 \\ 0.7 * 1.01 & 1.01 & \cdots & 0.7 * 1.01 \\ \vdots & \vdots & \ddots & \vdots \\ 0.7 * 1.01 & 0.7 * 1.01 & \cdots & 1.01 \end{bmatrix} \text{ for } \rho = 0.7$$

$$\text{v.} \quad \Sigma_1 = \begin{bmatrix} 1.01 & 0.9 * 1.01 & \cdots & 0.9 * 1.01 \\ 0.9 * 1.01 & 1.01 & \cdots & 0.9 * 1.01 \\ \vdots & \vdots & \ddots & \vdots \\ 0.9 * 1.01 & 0.9 * 1.01 & \cdots & 1.01 \end{bmatrix} \text{ for } \rho = 0.9$$

From power of test, sensitivity level of M-test and several modified M-test to shift in covariance structure is determined.

3.5.3 Significance Level

In hypothesis testing, the level of significance of the test carried out is denoted by α , is the criterion used for rejecting the null hypothesis and is defined as the probability of rejecting the null hypothesis when it is true. When the significance level 0.05 this indicate a 5% risk of concluding that a difference happened when no actual difference exists. To check the performance of the statistical tests, significance level $\alpha = 0.05$ was used.

3.6 Illustrative Application of the Impact of US Subprime Financial Crisis on the Forex Using Covariance Matrices Structure

Many real application data sets particularly in finance are high dimensional, heavy or long tailed such as the data sets of price or stock returns (Loretan & Phillips, 1994; Franco-Pedroso et al., 2018). It is known that such data sets are liable to have outliers (Ghoting, Parthasarathy & Otey, 2008; Pang, Cao, Chen, Lian & Liu, 2018).

In some situations, real data sets may contain hundreds or thousands of dimensions. In high dimensional space, the data sets becomes sparse, and the true outliers become masked by the contaminant effects of multiple irrelevant dimensions, when analyzed in full dimensionality (Aggarwal, 2017).

In reporting financial data foreign currency exchange (FOREX) accounting has been a topic of interest because of increased foreign exchange rate instability and multiple exchange rates (Kumhof, 2018). FOREX measurement is the process by which an entity expresses transactions whose terms are denominated in a foreign currency in its functional currency.

The United States Subprime mortgage financial crisis was a nationwide banking emergency. It was caused by a large failure in home prices after the collapse of a housing bubble, leading to mortgage crimes, foreclosures and the devaluation of housing related securities (Hamilton, 2015).

Failures in residential investment led the recession and were followed by reductions in household spending and then business investment. Spending reductions were more significant in areas with a combination of high household debt and large housing price drops (Hasbini, Homes & Satterfield, 2018).

The dramatic run-up and succession collapse of subprime financial crisis started from November 2006 to February 2008 (Tang, Wang, Huang, Wang & Wang, 2008). It sprang a great impact to US economy and on Foreign Exchange Market (FOREX) (Melvin & Taylor, 2009). The actual manifestation of the subprime mortgage crisis was in February to March 2007 where more than 25 subprime lenders declared insolvent and recorded significant lost (Singh, 2010).

According to Dymski and Dymski (2010) what has been most remarkable about this crisis is its resistance to policy cures. Several distinct incidents of near-meltdown in markets have occurred in August 2007, March 2008 and September to October 2008. Also wide range of policy actions to have proposed failed, this singular circumstance, caused a lot of damage to US and the rest of the World financial stability thereby caused crisis to FOREX. This made the subprime crisis in US to be one of the hottest topics of discussion around the globe, because of its effect on global FOREX and the economy.

Jaffee and Quigley (2008) labelled it as the beginning of the second great depression, the crisis ranks among the most serious economic event affecting the US since great depression of the 1930s. The crisis impacts on the economy activities, as example, mortgage (Tang et al., 2008; Acharya, Philippon, Richardson & Roubini, 2009; Bruecker, Calem & Nakamura, 2012), and financial market (Longstaff, 2010).

On the other hand, its impact on the foreign exchange market has been much less discussed (Melvin & Taylor, 2009). The illustration of the method is important to determine the method fit for the suggested test and to ensure the results are acceptable (Nickisch, Nockemann, Tillack, Murphy & Sturges, 1997). Consequently, our attention is focused on the FOREX since it is important in the measurement of financial data.

In order to elaborate the impact of subprime financial crisis, American, European and Asian continents currencies were employed to analyse the differences in covariance matrices before and after Subprime financial crisis incident. In this research, the evaluation on the robust performance M-test started with the application on real data.

3.6.1 Illustrative Application on Financial Data Sets

In general, an important issue for financial data sets is the level to which currency exchange rate changes affect the price index of the FOREX. Theoretically, in case of US subprime financial crisis a change in the value of the US dollar should raise or lower the price index of the FOREX market, thus decreasing or increasing US demand for FOREX.

Conversely, many financial studies have shown that when the US dollar fluctuates against the 25 currencies in this study, US FOREX prices index tend to show still less change (Jabara, 2009). Also, many financial studies have shown heavy-tailed data sets everywhere in the modern statistical analysis problems. They are naturally structured in heavy-tailed high dimensional data sets often corrupted by heavy-tailed outliers due to inadequate dimension, however, small sample are always used as estimates for a short time predictions (Fan, Wang & Zhu, 2016).

Moreover, even in the $p < n$ situation, as p increases, the traditional procedures for outlier identification based on Mahalanobis distance may become degenerate, and the contamination preference, which grows fast with p , could make the geometric median covariance determinant unreliable for large p (Alqallaf, Van Aelst, Yohai & Zamar, 2009 and Yu, Zou & Wang, 2012). This singularity drawback has been shown by some high dimensional location test for the mean vector of independent and identically distributed multivariate normal random vectors (Srivastava & Du, 2008; Chen & Qin, 2010).

Filzmoser, Maronna & Werner (2008) suggested a method using the properties of principal components analysis to identify outliers in a transformed space. Fritsch, Varoquaux, Thyreau, Poline & Thirion (2011) modified the minimum covariance determinant method by adding a regularization term to ensure that the estimation is well posed in high dimensional data sets.

On the other hand, there is no distributional result in Fritsch, Varoquaux, Thyreau, Poline & Thirion, (2011), so in practice it is not easy to find suitable cut off values to achieve an ideal robust rate. In order to overcome the problems of outlier and singularity in high dimensional data sets, suggested a high breakdown point method of finding the value with largest difference between it and sample mean, which can be an outlier (Ro, et al., 2015).

Let X be a data sample, vector in most cases. If argument is a dataframe, then outlier is calculated for each column. The same method is applied by when the matrix is given. Moreover, if the argument is true, then the opposite value (if largest value has maximum difference from the sample mean, it gives smallest and so on). On the other hand, if given vector of logical value and possible outlier position is marked by true. A vector of values with largest difference from the sample mean is identified (Komsta, 2011; Komsta, 2013).

The data set on daily price of subprime financial crisis currencies for two different periods of time; before (1st October 2006 until 31st October 2006) and after subprime financial crisis (1st March 2008 until 31st March 2008), is implemented to illustrate the application of the Geometric median covariance and its diagnosis analysis.

Thus, we focus on the performance evaluation of subprime financial crisis currencies. The banding geometric median (M_{BGMed} -test) and banding (M_B -test) are employed to highlight the impact of the financial crisis towards the stability of currencies structures. In order to analyse, the impact of subprime financial crisis and the performance of the new robust test, used daily data from 25 major currencies comprise of America, European and Asia countries. The data is retrieved from October 1st, 2006 to October 31st, 2006 and March 1st 2008 to March 31st 2008, through <http://fx.sauder.ubc.ca/data.html> on the sample size of 16 and variable is 24 on daily basis. The currencies symbols and the details are presented in Table 3.1.



Table 3.1

List of Countries and their currencies symbols

European		Asian		American	
Country	Symbol	Country	Symbol	Country	Symbol
United Kingdom	GBP	Indian	INR	Argentina	ARS
Czech Republic	CZK	China	CNY	United States	USD
Hungary	HUF	Indonesia	Rupiah	Brazil	BRL
Iceland	ISK	Japan	JPY	Colombia	COP
Poland	PLN	Malaysia	MYR	Chile	CLP
Russia	RUB	Philippine	PHP		
Sweden	SEK	Singapore	SGD		
Switzerland	CHF	Thailand	THB		
Romanian	RON	Taiwan	Dollar		
Norway	NOK	Pakistan	Rupee		

In general, there are 5 trading days per week, and about 18 to 21 days for month excluding holiday and weekend. The precious metals such as gold, platinum and silver can be considered as the base currency and US dollar as the base currency for comparison (Mizuno, Takayasu & Takayasu, 2006) because precious metals have been proven to offer a safe haven and are utilized to preserve fortunes and add safety to otherwise uncertain financial futures. Therefore, the daily exchange rate of each currency for gold is used as the base currency. The exchange rate of a currency for dollar can be converted into the exchange rate for gold by multiplying the exchange rate for gold by dollar (Wang, Xie, Chen & Chen, 2013, Olusegun & Sharif, 2016).

In this study, data are selected for 20 complete days, in a month where n_1 and n_2 , equals 20 and $p = 25$. For testing the equality of several independent sample covariance matrices using the M_{BGMed} -test and M_B -test involved in the hypothesis. Where $p = 25$ and $n = 16$.

The outlier identification problem has important applications in the field of financial data sets which are mostly dominated by hundreds of dimensions. In real application data sets there is need to identify outliers to show if the modified M-test can be applied in real application by carrying out data checking to identify the percentages of outliers and subsequently proceed with the test.

3.6.2 Data Checking For Outliers in High Dimensional Data Sets

The outlier identification and checking problem has vital applications in the field of financial data sets, however, most such applications are high dimensional fields in which the data comprises hundreds of dimensions. In order to illustrate if the new statistical test will be applicable for identifying outliers in high dimensional data sets, we need to check if the data sets contains certain percentages of outliers. Many algorithms use the concepts of closeness in order to find outliers based on their relationship to the rest of the data.

Thus, in high dimensional space, the data is sparse and the idea of closeness fails to maintain its importance. In fact, the sparsity of high dimensional data implies that every point is an almost equally good outlier from the view of closeness based descriptions. As a result, for high dimensional data, the idea of finding significant outliers becomes considerably more complex and not clear.

To carrying out data checking for outliers in multivariate high dimension, the classical method of Mahalanobis distances are only effective for very small number of variables (Ortner, Filzmoser, Zaharieva, Brodinova & Breiteneder, 2017). We check and compute the percentages of outliers occurrence in the data sets and simulation study. However, Mahalanobis distances cannot be used to find outliers in data. Since the Mahalanobis distance themselves are sensitive to outliers as they will always by construction sum to $n/(p + 1)$, the product of the dimensions of the data sets (Aggarwal & Sathe, 2017).

Also, Mahalanobis distance required the inversion of covariance matrices. Mahalanobis distance can not perform when $p > n$. From literature, in robust statistics, estimation of the multivariate location parameter μ and covariance matrix Σ is challenging, as many Mahalanobis distance method breakdown in high dimensional data set in the presence $n/(p + 1)$ outliers (Silva, Filzmoser & Brito, 2017 and Rousseeuw & Hubert, 2018).

In high dimension data sets, classical procedures based on the Mahalanobis distance are usually not appropriate, when the dimension p of the data is larger than the sample size, n , the Mahalanobis distance method is not possible (Filzmoser & Hron, 2008; Ro, Zou, Wang & Yin, 2015). Therefore, finding the values of the largest difference between it and the sample mean to identify outliers is practically used (Komsta, 2015).

3.6.3 Implementation and Algorithms of Data Checking for Outliers in High Dimensional Data Sets

The implementation and algorithm of data checking for outliers in high dimensional data sets contains the following three steps (Tukey, 1960 and Huang, Wang, Xie, Zhao & Lüers, 2016):

- i. Data Check: the data points will be marked as outliers if the biggest data points has maximum difference from the mean, each column with an outlier, it gives smallest and inversely, if the data point is True, gives vector of reasonable data points and possible outlier position is marked by True.

- ii. Error Check: Usually heavy-tailed data points that are interrelated to each other, especially in high dimensional data, the fluctuations and features within a time segment, hence the unexpected change of the FOREX price index value or the difference between two constant data points are possible to identify an outlier. Outlier identification compute the percentage maximum difference from the mean of each data point as the difference between its value and the mean value of the end-to-end points within the percentage of outlier 1 = 24, equals 24 divided by total sample number 480 multiply 100 giving as 5% and percentage of outlier 2 = 24, equals 24 divided by total sample number 479 multiply 100 giving as 5.01%. The data point is identified as the percentage of outliers.
- iii. Mean Deviation Check: This step implemented a further identification giving to the mean deviation starting point. First, take the data set column wise, compute the mean and subtract each data point, marked the largest value as an outlier in that way. This is given by:
- Mean Deviation = $\text{Max} (X_i - \bar{X})$. The figure 3.2, flow chart is used to explain the implementation and how the algorithm of data checking for outliers in high dimensional data sets is applied.

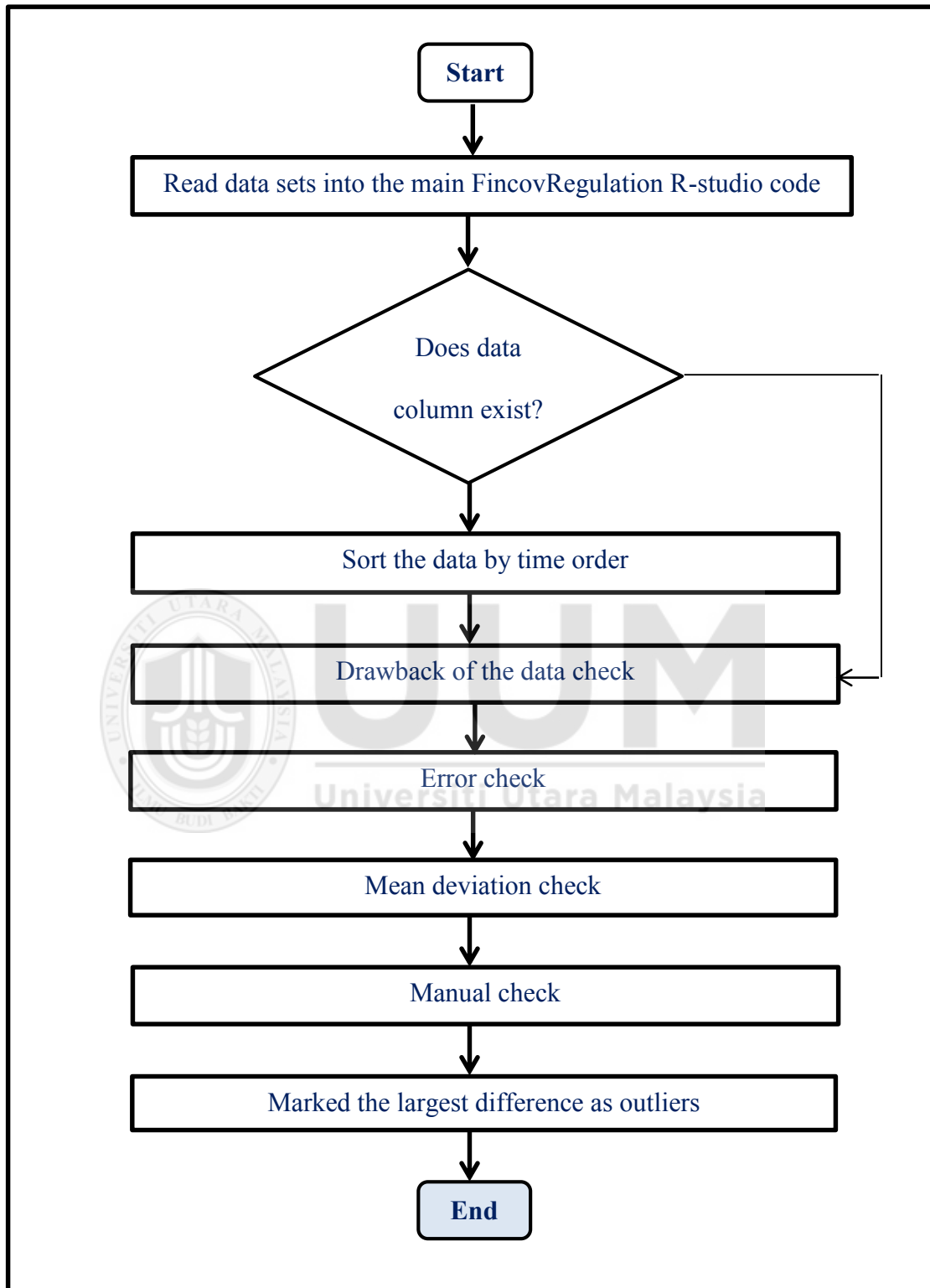


Figure 3.2: The flowchart of the algorithm for checking outlier (Tukey, 1960 and Huang, Wang, Xie, Zhao & Lüers, 2016).

For the purpose of this study, data sets on daily price of subprime financial crisis currencies for two periods of time; before and after subprime financial crisis, is retrieved. The change of data to the logarithm price of the currencies as in equation 3.7 makes that data independent, the ability to make changes to data characteristics without making changes to the programs that access the data (Fama, 1965 and Blau, 2018). Importantly, the data sets are transformed into the price return.

The use of the price return logarithm exchange rate R_i to evaluate the tests on the real application data to calculate the variables of interest (Mantegna & Stanley, 2000). Using non-stationary times series data sets in financial data produces unreliable and inaccurate results and leads to poor understanding of the analysis. The solution to the problem is to transform the financial data so that it becomes stationary (Serinaldi, Kilsby & Lombardo, 2018).

Let, P_i is the i -th currencies under study, the price return logarithm exchange rate R is computed as follows,

where,

$$R_i = \ln P_i(t + 1) - \ln P_i(t), i = 1, 2, \dots, p. \quad (3.7)$$

R_i is the daily logarithmic return of exchange rate of the price i .

$P_i(t)$ is the time price of currencies at the time t .

p is the number of variables (currencies).

3.6.4 Testing the Currencies Covariance Matrices of Subprime Financial Crisis

This section presents the illustrations of the suggested test and M-test by using two different methods. This is as follows:

- i. Testing two independent samples of covariance matrices; comparing before and after subprime financial crisis. Specifically, 1st October 2006 until 31st October 2006 represent the periods before and 1st March 2008 until 31st March 2008 after the subprime financial crisis respectively; and
- ii. Testing several independent samples of covariance matrices using a control chart; comparing i -th sample and reference sample, $i = 1, 2, \dots, m$ are the several covariance, where $i = 1, 2, \dots, 16$ and horizontal axis is 16 months. Thus, i represent the 16 months.

The hypothesis statement of the first method is for testing the equality of two covariance matrices for before and after the US subprime crisis datasets,

$$H_0: \Sigma_{\text{before}} = \Sigma_{\text{after}}$$

versus

$$H_1: \Sigma_{\text{before}} \neq \Sigma_{\text{after}}$$

where,

Σ_{before} covariance matrix of before US subprime financial crisis.

Σ_{after} covariance matrix of after US subprime financial crisis.

$n = 16$ and $p =$ varies numbers from 18, 19, 20 and 21.

The hypothesis statement of the second method is to test the equality of several covariance matrices using the control chart,

$$H_0: \Sigma_i = \Sigma_0$$

versus

$$H_1: \Sigma_i \neq \Sigma_0, \text{ where } i = 1, 2, \dots, m$$

where,

Σ_i sample covariance matrices from $i = 1, 2, \dots, m$

Σ_0 population covariance matrix

$n = 16$ and $p = 24$

3.6.5 Control Chart

Control charts are used usually in industries and finance to monitor or control the practices. This is a chart used to monitor covariance structure changes over time. Data are plotted in time order. Generally, control chart comprises of a central line that is the mean for in-control process. With two main horizontal lines the upper control limit (UCL) and the lower control limit (LCL). The control limits are selected so that almost all data points will fall within these limits such that the process remains in-control.

The most used control chart methods are the 3 sigma control limits with ± 3 standard deviations from the mean. Any sample statistic that falls between *UCL* and *LCL* will show that the process is displaying common causes of variation.

In this study, the concept of multivariate statistical process control (MSPC) can be used to monitor the stability of covariance structure. This concept is standard procedure in industry process quality control. Based on the concept introduced by Montgomery (2001) which is represented again in Montgomery (2005), monitoring the stability of covariance structure is equivalent to repeated tests of significance of the hypothesis that the process covariance matrix is equal to a particular matrix of constants (Wierda, 1994; Sullivan *et al.*, 1996; Woodall *et al.*, 1999; Alt *et al.*, 1988; Khoo *et al.*, 2003 & Djauhari, 2005). Therefore, the UCL and LCL are 315.75 and 0 respectively since the distribution of M_{BGMed} -test and M_B -test are Chi-square distribution. At the end of the analysis M_{BGMed} -test control chart and M_B -test control chart will be delivered to help the user for monitoring the stability of covariance structure from November 2006 until Februray 2008.

CHAPTER FOUR

RESULTS AND FINDINGS

4.1 Introduction

This chapter present two major findings in order to achieve objectives of the study. The first major finding is on the new robust covariance equality test in the presence of outliers in high dimensional cases using the Type I error rate and secondly to solve the singularity problem with the power of test values. We compared original M-test, M_{GMed} -test, M_B -test, M_T -test, M_{BGMed} -test and M_{TGMed} -test in terms of Type I error rate and Power of test value.

4.2 Type I Error Rate

For the Type I error rate, considered the 4 types of data contamination to examine the strength and weakness of the tests. Besides that, all these tests have been exposed to various conditions which are number of variables (p), sample sizes (n), percentages of outliers, epsilon (ϵ), scale amount of contamination Kappa (k) and mean shift (μ) (Ramey, 2013).

The comparisons are summarized in form of tables. The first column in each table shows the percentage of outliers (ϵ), followed by amount scale of contamination (k) and the mean vectors (μ). The following six columns recorded the Type I error rates of M-test, M_{GMed} -test, M_B -test, M_T -test, M_{BGMed} -test and M_{TGMed} -test which are investigated in the study. This conditions is repeated for different sample sizes. Under all level of data contaminations (5%, 10%, 20% and 50%).

The significance level $\alpha = 0.05$, a statistic is robust when the Type I error rates lies between 0.025 and 0.075 (Ahad, Othman & Yahaya, 2012) , otherwise the statistic is considered non-robust. The values that are closest to the significance level and within 0.025 and 0.075 are shaded in the tables. The significance level was choosing since there are some expected losses from Type I error rates and power of test values, often incur losses which affect people's lives, such as economic recession, ill health e.t.c (Ziliar and McCloskey, 2008).

4.2.1 Type I Error Rates for Small Number of Variables ($p = 3, 5, \text{ and } 8$)

Table 4.1-4.3 show the Type I error rates for small number of variables which are $p = 3$, $p = 5$ and $p = 8$. There are 175 conditions involved where the number of sample size multiply by the number of conditions, (7×25) in evaluating the robustness of each four different statistical tests. Specifically, Table 4.1 display the Type I error rates of $p = 3$, followed by Table 4.2 display the Type I error rates of $p = 5$, and Table 4.3 display the Type I error rates of $p = 8$. Overall, when $p = 3, n = 2$, $p = 5, n = 2$ and $5, p = 8, n = 2$ and 5 , the Type I error rates of M-test and M_{GMed} -test cannot be computed because both tests can only perform when the sample size is large than number of variables $n > p$. The singularity problems occur when the determinant is 0 and the inversion process is not possible. The details of Type I error rates for small number of variables are presented in Appendix K.1 to K.21.

In summary, Table 4.1 show that there are 175 conditions each involved in evaluating the robustness of each tests, under all level of data contaminations (5%, 10%, 20% and 50%). For $p = 3$ and $n = 2, 5, 10, 15, 20, 25$ and 50, there are 46 conditions of M-test, 98 conditions of M_{GMed} -test, 63 conditions of M_B -test, 64 conditions of M_T -test, 134 conditions of M_{BGMed} -test and 122 conditions of M_{TGMed} -test fall within the robust interval. Therefore, when $p = 3$, we conclude that M_{BGMed} -test is the best robust tests in both cases (either n is small or large).

Table 4.1

Summary of Type I error rates for small variables, $p = 3$

Statistical Tests	No of Robust conditions	Statistical Tests	No of Robust conditions
M-test	46*	M_{GMed} -test	98
M_B -test	63	M_{BGMed} -test	134
M_T -test	64	M_{TGMed} -test	122

* Based on 133 conditions since M-test cannot be computed when $n = 2$

In summary, Table 4.2 show that there are 175 conditions involved in evaluating the robustness of each tests for $p = 5$ and $n = 2, 5, 10, 15, 20, 25$ and 50. There are 31 out of 175 conditions of M-test, 66 conditions of M_{GMed} -test, 60 conditions of M_B -test, 61 conditions of M_T -test, 122 conditions of M_{BGMed} -test and 121 conditions of M_{TGMed} -test fall within the robust interval. Therefore, when $p = 5$, we conclude that M_{BGMed} -test as the best robust test in both cases (either n is small or large).

Table 4.2

Summary of Type I error rates for small variables, $p = 5$

Statistical Tests	No of Robust conditions	Statistical Tests	No of Robust conditions
M-test	31	M_{GMed} -test	66
M_B -test	60	M_{BGMed} -test	122
M_T -test	61	M_{TGMed} -test	121

In summary, Table 4.3 show that there are 175 conditions each involved in evaluating the robustness of test for $p = 8$ and $n = 2, 5, 10, 15, 20, 25$ and 50. There are 44 out of 175 conditions of M-test, 56 conditions of M_{GMed} -test, 55 conditions of M_B -test, 47 conditions of M_T -test, 119 conditions of M_{BGMed} -test and 118 conditions of M_{TGMed} -test fall within the robust interval. Therefore, when $p = 8$, we conclude that M_{BGMed} -test is the best robust test in both cases (either n is small or large).

Table 4.3

Summary of Type I error rates for small variables, $p = 8$

Statistical Tests	No of Robust conditions	Statistical Tests	No of Robust conditions
M-test	44	M_{GMed} -test	56
M_B -test	55	M_{BGMed} -test	119
M_T -test	47	M_{TGMed} -test	118

Table 4.4 show the summary of Type I error that fall under robustness criterion for small number of variables ($p = 3, 5$ and 8). The best test is M_{BGMed} -test, followed by M_{TGMed} -test and M_{GMed} -test when $p = 3$. However, M_B -test and M_T -test results are 63 and 64 conditions, respectively. M_{BGMed} -test is the best test for $p = 5$, next M_{TGMed} -test, but M_{GMed} -test 66, M_B -test 60 and M_T -test 61 conditions. The results of M-test is not good when compared with the previous results with 31 conditions. Lastly, M_{BGMed} -test is the best when $p = 8$, next M_{TGMed} -test, but M_{GMed} -test 56, M_B -test 55 and M_T -test 47 conditions.

In the summary for small variables ($p = 3, 5$ and 8), the conditions of M-test is 121, 220 M_{GMed} -test, 178 M_B -test, 172 M_T -test, 375 M_{BGMed} -test and 361 M_{TGMed} -test that fall within the robust interval. Hence, for small number of variables, M_{BGMed} -test is the best robust test for both cases (either n is small or large).

Table 4.4

Summary of Type I error rates for small variables ($p = 3, 5$, and 8).

Statistical Tests	No of Robust Conditions			Total
	$p = 3$	$p = 5$	$p = 8$	
M-test	46	31	44	121
M_B -test	63	60	55	178
M_T -test	64	61	47	172
M_{GMed} -test	98	66	56	220
M_{BGMed} -test	134	122	119	375
M_{TGMed} -test	122	121	118	361

4.2.2 Type I error for Medium Number of Variables ($p = 10, 15$ and 20)

Table 4.5–4.7 show the Type I error rates for medium number of variables which are $p = 10$, $p = 15$ and $p = 20$. There are 150 conditions involved where the number of sample size multiply by the number of conditions, 6×25 in evaluating the robustness of each six different statistical tests. Specifically, Table 4.5 display the Type I error rates of $p = 10$, followed by Table 4.6 display the Type I error rates of $p = 15$, and Table 4.7 display the Type I error rates of $p = 20$.

Overall, when $p = 10$, $n = 5$ and 10 , the Type I error rates of M-test and M_{GMed} -test cannot be computed for $p > n$, because both tests can only perform when the sample size is larger than number of variables. The details of Type I error rates for medium number of variables are presented in Appendix L.22 to L.39.

In summary, Table 4.5 show that there are 150 conditions each involved in evaluating the robustness of each tests. For $p = 10$ and $n = 5, 10, 15, 20, 25$ and 50 , there are 24 conditions of M-test, 43 conditions of M_{GMed} -test, 34 conditions of M_B -test, 32 conditions of M_T -test, 94 conditions of M_{BGMed} -test and 87 conditions of M_{TGMed} -test fall within the robust interval. Therefore, when $p = 10$, we conclude that M_{BGMed} -test is the best robust tests in both cases (either n is small or large).

Table 4.5

Summary of Type I error rates for medium variables, $p = 10$

Statistical Tests	No of Robust Conditions	Statistical Tests	No of Robust Conditions
M-test	24	M _{GMed} -test	43
M _B -test	34	M _{BGMed} -test	94
M _T -test	32	M _{TGMed} -test	87

In summary, Table 4.6, show that there are 150 conditions involved in evaluating the robustness of each test for $p = 15$ and $n = 5, 10, 15, 20, 25$ and 50 . There are 15 out of 150 conditions of M-test, 33 conditions of M_{GMed}-test, 36 conditions of M_B-test, 34 conditions of M_T-test, 79 conditions of M_{BGMed}-test and 77 conditions of M_{TGMed}-test fall within the robust interval. Therefore, when $p = 15$, we conclude that M_{BGMed}-test is the best robust test in both cases (either n is small or large).

Table 4.6

Summary of Type I error rates for medium variables, $p = 15$

Statistical Tests	No of Robust Conditions	Statistical Tests	No of Robust Conditions
M-test	15	M _{GMed} -test	33
M _B -test	36	M _{BGMed} -test	79
M _T -test	34	M _{TGMed} -test	77

In summary, Table 4.7 show that there are 150 conditions each involved in evaluating the robustness of test for $p = 20$ and $n = 5, 10, 15, 20, 25$ and 50 . There are 10 out of 114 conditions of M-test, 24 conditions of M_{GMed} -test, 26 conditions of M_B -test, 27 conditions of M_T -test, 74 conditions of M_{BGMed} -test and 66 conditions of M_{TGMed} -test fall within the robust interval. Therefore, when $p = 20$, we conclude that M_{BGMed} -test is the best robust test in both cases (either n is small or large).

Table 4.7

Summary of Type I error rates for medium variables, $p = 20$

Statistical Tests	No of Robust Conditions	Statistical Tests	No of Robust Conditions
M-test	10	M_{GMed} -test	24
M_B -test	26	M_{BGMed} -test	74
M_T -test	27	M_{TGMed} -test	66

In Table 4.8 the medium number of variables ($p = 10, 15$ and 20) display the summary of Type I error rate results. The best test is M_{BGMed} -test when $p = 10$, next are the M_{TGMed} -test, M_{GMed} -test, M_B -test and M_T -test condition results. The original test M-test result is the worst among the other. However, for $p = 15$, best test is M_{BGMed} -test, followed by M_{TGMed} -test, but M_{GMed} -test 33, M_B -test 36, and M_T -test 34, conditions. But the result of M-test is bad compare to other tests. Similarly, M_{BGMed} -test is the best test for $p = 20$, then M_{TGMed} -test, M_{GMed} -test, M_B -test and M_T -test are 24, 26 and 27 conditions respectively.

In conclusion, for medium variables ($p = 10, 15$ and 20), there are 49 conditions of M-test, 100 conditions of M_{GMed} -test, 96 conditions of M_B -test, 93 conditions of M_T -test, 247 conditions of M_{BGMed} -test and 230 conditions of M_{TGMed} -test that fall within the robust interval. Therefore, M_{BGMed} -test can be concluded as the best robust test. Moreover, M_{BGMed} -test also performed best in both cases (either n is small or large).

Table 4.8

Summary of Type I error rates for medium variables ($p = 10, 15$ and 20).

Statistical Tests	No of Robust Conditions			Total
	$p = 10$	$p = 15$	$p = 20$	
M-test	24	15	10	49
M_B -test	34	36	26	96
M_T -test	32	34	27	93
M_{GMed} -test	43	33	24	100
M_{BGMed} -test	94	79	74	247
M_{TGMed} -test	87	77	66	230

4.2.3 Type I error rate for Large Number of Variables ($p = 25, 30$ and 50)

Table 4.9–4.11 show the Type I error rates for large number of variables which are = 25, $p = 30$ and $p = 50$. There are 150 conditions involved where the number of sample size multiply by numbers of conditions (6×25) in evaluating the robustness of each six different statistical tests.

Specifically, Table 4.9 display the Type I error rates of $p = 25$, followed by Table 4.10 display the Type I error rates of $p = 30$, and Table 4.11 display the Type I error rates of $p = 50$. Overall, when $p = 25$ and 30 , $n = 5, 10, 15, 20$ and 25 , $p = 50, n = 5, 10, 15, 20, 25$ and 50 the Type I error rates of M-test and M_{GMed} -test cannot be computed because both tests can only perform when the sample size is larger than number of variables $n > p$.

When $p > n$ singularity problems occur since the determinant will be 0 and the inversion is not possible to be computed. The details of Type I error rates for small number of variables are presented in Appendix M.40 to M.57.

In summary, from Table 4.9, show that there are 150 conditions each involved in evaluating the robustness of each tests. For $p = 25$ and $n = 5, 10, 15, 20, 25$ and 50 , there are 3 conditions of M-test, 9 conditions of M_{GMed} -test, 19 conditions of M_B -test, 26 conditions of M_T -test, 67 conditions of M_{BGMed} -test and 53 conditions of M_{TGMEd} -test fall within the robust interval. Therefore, when $p = 25$, we conclude that M_{BGMed} -test is the best robust test in both cases (either n is small or large).

Table 4.9

Summary Type I error rates for large variables, $p = 25$

Statistical Tests	No of Robust Conditions	Statistical Tests	No of Robust Conditions
M-test	3	M_{GMed} -test	9
M_B -test	19	M_{BGMed} -test	67
M_T -test	26	M_{TGMed} -test	53

* Based on 150 conditions since M-test cannot be computed when $n = 5, 10, 15$ and 20

In summary, from Table 4.10, show that there are 150 conditions involved in evaluating the robustness of each tests for $p = 30$ and $n = 5, 10, 15, 20, 25$ and 50 . There are 3 out of 150 conditions of M-test, 9 conditions of M_{GMed} -test, 24 conditions of M_B -test, 26 conditions of M_T -test, 65 conditions of M_{BGMed} -test and 61 conditions of M_{TGMed} -test fall within the robust interval. Therefore, when $p = 30$, we conclude that M_{BGMed} -test is best robust test in both cases (either n is small or large).

Table 4.10

Summary of Type I error rates for large variables, $p = 30$

Statistical Tests	No of Robust Conditions	Statistical Tests	No of Robust Conditions
M-test	3	M_{GMed} -test	9
M_B -test	24	M_{BGMed} -test	65
M_T -test	26	M_{TGMed} -test	61

* Based on 150 conditions since M-test cannot be computed when $n = 5, 10, 15, 20$ and 25 .

In summary, from Table 4.11, there are 150 conditions involved in evaluating the robustness of each test for $p = 50$ and $n = 5, 10, 15, 20, 25$ and 50 . There are 0 out of 150 conditions of M-test, 0 condition of M_{GMed} -test, 16 conditions of M_B -test, 15 conditions of M_T -test, 58 conditions of M_{BGMed} -test and 48 conditions of M_{TGMed} -test fall within the robust interval. Therefore, when $p = 50$, we conclude that M_{BGMed} -test is the best robust test in both cases (either n is small or large).

Table 4.11

Summary of Type I error rates for large variables, $p = 50$

Statistical Tests	No of Robust Conditions	Statistical Tests	No of Robust Conditions
M-test	0	M_{GMed} -test	0
M_B -test	16	M_{BGMed} -test	58
M_T -test	15	M_{TGMed} -test	48

* Based on 150 conditions since M-test cannot be computed when $n = 5, 10, 15, 20, 25$ and 50 .

The summary of Table 4.12, for large number of variables ($p = 25, 30$ and 50) show the Type I error rate results. M_{BGMed} -test is the best test for $p = 25$, followed by M_{TGMed} -test, M_{GMed} -test 9, M_B -test 19 and M_T -test 26 conditions respectively. The M-test results is not good among the other. In $p = 30$, the best test is M_{BGMed} -test, then the next is M_{TGMed} -test, then M_{GMed} -test, M_B -test and M_T -test are 9, 24 and 26 conditions respectively. M-test result is not good enough compared to others.

Finally, for $p = 50$, the result of M_{BGMed} -test is the best, then M_{TGMed} -test, but M_{GMed} -test is 0, M_B -test 16 and M_T -test 15 conditions. M-test is bad among the others.

As a conclusion, for large variables ($p = 25, 30$, and 50), there are 6 conditions of M-test, 18 conditions of M_{GMed} -test, 59 conditions of M_B -test, 67 conditions of M_T -test, 190 conditions of M_{BGMed} -test and 162 conditions of M_{TGMed} -test that fall within the robust interval. Therefore, M_{BGMed} -test can be concluded as the best robust test, furthermore, M_{BGMed} -test also performed best in both cases (either n is small or large).

Table 4.12

Summary of Type I error rates for large variables ($p = 25, 30$ and 50)

Statistical Tests	No of Robust Conditions			Total
	$p = 25$	$p = 30$	$p = 50$	
M-test	3	3	0	6
M_B -test	19	24	16	59
M_T -test	26	26	15	67
M_{GMed} -test	9	9	0	18
M_{BGMed} -test	67	65	58	190
M_{TGMed} -test	53	61	48	162

4.3 Power of Test

The investigation on the robustness of the six statistical tests has been displayed in the previous section. The results show that M_{BGMed} -test performed well compared to other statistical tests in terms of Type I error rates. Then, continue the investigation by evaluating the sensitivity analysis between existing M-test and several modified M_{GMed} -test, M_B -test, M_T -test, M_{BGMed} -test, and M_{TGMed} -test based on power of test.

The power of test is defined as probability of correctly rejecting null hypothesis when it is false (Cohen, 1992). Generally, the power value is between 0 and 1. In this analysis, the minimum accepted value of power is greater than 0.5 (Lakens, 2013). If the power value is smaller than 0.5, it show the power to be unusual (Greenland, Senn, Rothman, Carlin, Poole, Goodman, & Altman, 2016). Thus, in the next section, power values that are 0.5 and above are shaded in all the following tables. If the test is able to detect small effect, i.e., small shift in covariance rather than large effect, the test show to be more sensitive to the others.

The simulation results are presented in Table N.58-P.114. Each table represents the ascending number of variables, i.e. small ($p = 3, 5$ and 8), medium ($p = 10, 15$ and 20) and large number variable ($p = 25, 30$ and 50) with $\alpha = 0.05$, presented in Appendix N.58 to P.114.

4.3.1 Power of Test for Small Number of Variables ($p = 3, 5$ and 8)

Table N.58-N.78 show the power of test for small number of variables which are $p = 3$, $p = 5$ and $p = 8$. There are 245 conditions involved in evaluating the sensitivity of test conditions for number of sample size multiply by number of conditions (7×35) in evaluating the sensitivity of test.

Specifically, Table N.58-N.64 displays the power of test for $p = 3$, followed by Table N.65-N.71 for $p = 5$, and Table N.72-N.78 for $p = 8$. However, when $p = 3$, $n = 2$ power of test conditions of M-test and M_{GMed} -test cannot be computed because both tests can only perform when the sample size is larger than number of variables $n > p$. The Table presentation is displayed in Appendix N.58 to N.78.

The summary for Table 4.13, there are 245 conditions involved in evaluating sensitivity of the test in small number of variables $p = 3$ and sample sizes $n = 2, 5, 10, 15, 20, 25$ and 50 . Only 91 out of 245 conditions of M-test, 57 of M_{GMed} -test, 52 of M_B -test, 51 of M_T -test, 1 of M_{BGMed} -test and M_{TGMed} -test conditions respectively that fall within the power interval. In $p = 3$, the conclusion is M-test as the best test when n large.

Table 4.13

Power of Test summary for small variables, $p = 3$

Statistical Tests	No of Power conditions	Statistical Tests	No of Power conditions
M-test	91	M_{GMed} -test	57
M_B -test	52	M_{BGMed} -test	1
M_T -test	51	M_{TGMed} -test	1

Table N.65-N.66 display the power of test rates of $p = 5, n = 2$ and $n = 5$ in Appendix 4, the power of test rates of M-test and M_{GMed} -test cannot be computed because both tests can only perform when the sample size is larger than number of variables ($n > p$).

As summary, Table 4.14, there are 245 conditions each involved in evaluating the power of each test for $p = 5$ and $n = 2, 5, 10, 15, 20, 25$ and 50. There are 97 out of 245 conditions of M-test, 78 conditions of M_{GMed} -test, 70 conditions of M_B -test, 68 conditions of M_T -test, 9 conditions of M_{BGMed} -test and 12 conditions of M_{TGMed} -test fall within the power interval. Therefore, when $p = 5$, we conclude that M-test is the best test in cases n is large compared with variable size.

Table 4.14

Power of Test summary for small variables, $p = 5$

Statistical Tests	No of Power conditions	Statistical Tests	No of Power conditions
M-test	97	M_{GMed} -test	78
M_B -test	70	M_{BGMed} -test	9
M_T -test	68	M_{TGMed} -test	12

Table N.72-N.73 display the power of test rates of $p = 8$. Overall, when $p = 8$ and $n = 2$ and $n = 5$ the power of test rates of M-test and M_{GMed} -test cannot be computed because both tests can only perform when the sample size is larger than number of variables ($n > p$).

In summary, Table 4.16, there are 245 conditions each involved in evaluating the power of each test for $p = 8$ and $n = 2, 5, 10, 15, 20, 25$ and 50. There are 87 out of 245 conditions of M-test, 81 conditions of M_{GMed} -test, 70 conditions of M_B -test, 74 conditions of M_T -test, 18 conditions of M_{BGMed} -test and 22 conditions of M_{TGMed} -test fall within the power interval. Therefore, when $p = 8$, we conclude that M-test is the best test in cases n is large compared with variable size.

Table 4.15

Power of Test summary for small variables, $p = 8$

Statistical Tests	No of Power conditions	Statistical Tests	No of Power conditions
M-test	87	M_{GMed} -test	81
M_B -test	70	M_{BGMed} -test	18
M_T -test	74	M_{TGMed} -test	22

Table 4.16 show the summary of power of test for small number of variables ($p = 3, 5$ and 8). For $p = 3$, M-test is the best, followed by M_{GMed} -test, then M_B -test and M_T -test, while M_{BGMed} -test and M_{TGMed} -test performed poorly both has 1 result each. For $p = 5$, M-test is the best, followed by M_{GMed} -test, then M_B -test and M_T -test, besides, M_{BGMed} -test and M_{TGMed} -test results has an increase compared to when $p = 3$. Lastly, in small number of variables $p = 8$, M-test is still the best then M_{GMed} -test and M_B -test and M_T -test, but M_{BGMed} -test and M_{TGMed} -test results with some increases as compared.

The overall summary for small variables $p = 3, 5$ and 8 , there are 275 conditions of M-test, 216 conditions of M_{GMed} -test, 191 conditions of M_B -test, 193 conditions of M_T -test, 28 conditions of M_{BGMed} -test and 35 conditions of M_{TGMed} -test that fall within the power interval. Therefore, M-test can be concluded as the best that fall within the power of test values. The power of test values for M-test dominate the other tests especially from $n > p$ cases, but, M_T -test perform better when $p > n$.

Table 4.16

Summary of Power of Test for small variables ($p = 3, 5$ and 8).

Statistical Tests	No of Power Conditions			Total
	$p = 3$	$p = 5$	$p = 8$	
M-test	91	97	87	275
M_B -test	51	70	70	191
M_T -test	50	69	74	193
M_{GMed} -test	57	78	81	216
M_{BGMed} -test	1	9	18	28
M_{TGMed} -test	1	12	22	35

4.3.2 Power of Test for Medium Number of Variable ($p = 10, 15$ and 20)

Table O.79-O.96 display power of test for medium number of variables which are $p = 10, p = 15$ and $p = 20$. There are 210 conditions number of sample size multiply by number of conditions (6×35) each involved in evaluating the power of test values of six different statistical tests. In detail, Table O.79-O.84 displays the power of test for $p = 10$, followed by Table O.85-O.90 for $p = 15$, and Table O.91-O.96 for $p = 20$. However, when $p = 10, n = 5$ and 10 , the power of test values of M-test and M_{GMed} -test cannot be computed because both tests can only perform when the sample size is larger than number of variables ($n > p$). The Table will be presented in Appendix O.79 to O.96.

In summary, Table 4.18, there are 210 conditions each involved in evaluating the power of each test for $p = 10$ and $n = 5, 10, 15, 20, 25$ and 50 . There are 80 out of 210 conditions of M-test, 75 conditions of M_{GMed} -test, 80 conditions of M_B -test, 83 conditions of M_T -test, 29 conditions of M_{BGMed} -test and 28 conditions of M_{TGMed} -test fall within the power interval. Therefore, when $p = 10$, we conclude that M_T -test is the best test when (either n is small or large).

Table 4.17

Power of Test summary for medium variables, $p = 10$

Statistical Tests	No of Power conditions	Statistical Tests	No of Power conditions
M-test	80	M_{GMed} -test	75
M_B -test	80	M_{BGMed} -test	29
M_T -test	83	M_{TGMed} -test	28

Table O.85-O.87 show the power of test for $p = 15$. Overall, when $p = 15$, $n = 5, 10$ and 15 power of test values of M-test and M_{GMed} -test cannot be computed because both tests can only perform when the sample size is larger than number of variables ($n > p$).

As summary, Table 4.18, there are 210 conditions each involved in evaluating the power of each test for $p = 15$ and $n = 5, 10, 15, 20, 25$ and 50 . There are 57 out of 210 conditions of M-test, 60 conditions of M_{GMed} -test, 92 conditions of M_B -test, 79 conditions of M_T -test, 43 conditions of M_{BGMed} -test and 64 conditions of M_{TGMed} -test fall within the power interval. Therefore, when $p = 15$, we conclude that M_B -test is the best robust test in both cases (either n is small or large).

Table 4.18

Power of Test summary for medium variables, $p = 15$

Statistical Tests	No of Power conditions	Statistical Tests	No of Power conditions
M-test	57	M _{GMed} -test	60
M _B -test	92	M _{BGMed} -test	43
M _T -test	79	M _{TGMed} -test	64

In summary, Table 4.19, there are 210 conditions each involved in evaluating the power of each test for $p = 20$ and $n = 5, 10, 15, 20, 25$ and 50. There are 43 out of 210 conditions of M-test, 43 conditions of M_{GMed}-test, 99 conditions of M_B-test, 100 conditions of M_T-test, 44 conditions of M_{BGMed}-test and 44 conditions of M_{TGMed}-test fall within the power interval. Therefore, when $p = 20$, we conclude that M_T-test is the powerful test in both cases (either when n is small or large).

Table 4.19

Power of Test summary for medium variables, $p = 20$

Statistical Tests	No of Power conditions	Statistical Tests	No of Power conditions
M-test	43	M _{GMed} -test	43
M _B -test	99	M _{BGMed} -test	44
M _T -test	100	M _{TGMed} -test	44

Table 4.20 display the summary of power of test for medium number of variables ($p = 10, 15$ and 20). For $p = 10$, the best test is M_T -test, followed by M -test and M_B -test same results then M_{GMed} -test, but M_{BGMed} -test and M_{TGMed} -test results still poor compared. Also, $p = 15$, M_B -test is the best test, then M_T -test, M_{TGMed} -test, M -test, M_{GMed} -test and M_{BGMed} -test. For $p = 20$, M_T -test is the best test followed by M_B -test, but M_{BGMed} -test, M_{TGMed} -test, M_{GMed} -test and M -test share results closely that fall within the power interval.

As summary for medium variable numbers $p = 10, 15$ and 20 , there are 180 conditions of M -test, 178 conditions of M_{GMed} -test, 271 conditions of M_B -test, 262 conditions of M_T -test, 116 conditions of M_{BGMed} -test and 136 conditions of M_{TGMed} -test that fall within the power interval. Therefore, when medium variables ($p = 10, 15$ and 20), we conclude that M_B -test is the best test in both cases (either n is small or large). The power value for M_B -test dominate the other tests especially when $p > n$. But, M -test is performed better when $n = 50$.

Table 4.20

Summary of Power of Test for Medium variables ($p = 10, 15$ and 20).

Statistical Tests	No of Power Conditions			Total
	$p = 10$	$p = 15$	$p = 20$	
M-test	80	57	43	180
M_B -test	80	92	99	271
M_T -test	83	79	100	262
M_{GMed} -test	75	60	43	178
M_{BGMed} -test	29	43	44	116
M_{TGMed} -test	28	64	44	136

4.3.3 Power of Test for Large Number of Variables ($p = 25, 30$ and 50)

Table P.97-P.114 display power of test for large number of variables which are $p = 25, p = 30$ and $p = 50$. There are 210 conditions number of sample size multiply number of conditions (6×35) involved in evaluating the power of test values of six different statistical tests. Specially, Table P.97-P.102 shows the power of test for $p = 25$, followed by Table P.103-P.108 for $p = 30$, and Table P.109-P.114 for $p = 50$. However, when $p = 25, n = 5, 10, 15, 20$ and 25 , the power of test values of M-test and M_{GMed} -test cannot be computed because both tests can only perform when the sample size is larger than number of variables ($n > p$). The tables are presented in Appendix P.97 to P.114.

As summary, Table 4.21, there are 210 conditions each involved in evaluating the power of test for $p = 25$ and $n = 5, 10, 15, 20, 25$ and 50 . There are 28 out of 210 conditions of M-test, 28 conditions of M_{GMed} -test, 106 conditions of M_B -test, 105 conditions of M_T -test, 52 conditions of M_{BGMed} -test and 52 conditions of M_{TGMed} -test fall within the power interval. Therefore, when $p = 25$, we conclude that M_B -test is the best test in both cases (either n is small or large).

Table 4.21

Power of Test summary for large variables, $p = 25$

Statistical Tests	No of Power conditions	Statistical Tests	No of Power conditions
M-test	28	M_{GMed} -test	28
M_B -test	106	M_{BGMed} -test	52
M_T -test	105	M_{TGMed} -test	52

In conclusion, from Table 4.22, there are 210 conditions each involved in evaluating the power of each test for $p = 30$ and $n = 5, 10, 15, 20, 25$ and 50 . There are 0 out of 210 conditions of M-test, 0 condition of M_{GMed} -test, 106 conditions of M_B -test, 108 conditions of M_T -test, 56 conditions of M_{BGMed} -test and 57 conditions of M_{TGMed} -test fall within the power interval. Therefore, when $p = 30$, we conclude that M_T -test is the best robust test in both cases (either n is small or large).

Table 4.22

Power of Test summary for large variables, $p = 30$

Statistical Tests	No of Power conditions	Statistical Tests	No of Power conditions
M-test	0	M _{GMed} -test	0
M _B -test	106	M _{BGMed} -test	56
M _T -test	108	M _{TGMed} -test	57

In summary, from Table 4.23, there are 210 conditions each involved in evaluating the power of each test for $p = 50$ and $n = 5, 10, 15, 20, 25$ and 50 . There are 0 condition of M-test, 0 condition of M_{GMed}-test, 116 conditions of M_B-test, 116 conditions of M_T-test, 68 conditions of M_{BGMed}-test and 67 conditions of M_{TGMed}-test fall within the power interval. Therefore, when $p = 50$, we conclude that M_B-test and M_T-test are the best powerful test in both cases (either n is small or large), especially when $p > n$.

Table 4.23

Power of Test summary for large variables, $p = 50$

Statistical Tests	No of Power conditions	Statistical Tests	No of Power conditions
M-test	0	M _{GMed} -test	0
M _B -test	116	M _{BGMed} -test	66
M _T -test	116	M _{TGMed} -test	65

As summary for $p = 25, 30$, and 50 , there are 28 conditions of M-test, 28 conditions of M_{GMed} -test, 328 conditions of M_B -test, 329 conditions of M_T -test, 174 conditions of M_{BGMed} -test and 174 conditions of M_{TGMed} -test that fall within the power interval. Therefore, when large variables ($p = 25, 30$, and 50), we conclude that M_T -test is the best test in both cases (either n is small or large).

Table 4.24

Summary of Power of Test for large variables ($p = 25, 30$ and 50).

Statistical Tests	No of Power Conditions			Total
	$p = 25$	$p = 30$	$p = 50$	
M-test	28	0	0	28
M_B -test	106	106	116	328
M_T -test	105	108	116	329
M_{GMed} -test	28	0	0	28
M_{BGMed} -test	52	56	66	175
M_{TGMed} -test	52	57	65	175

Therefore, the detail about the steps for power of test computation for performance is discussed in Appendix P.

4.3.4 Conclusion of Type I Error Rate and Power of Test Values

In this chapter, we test the performances of M-test and several modified M-test by integrating the geometric median covariance for producing a new robust covariance equality testing in the presence of outliers using the Type I error rate values and power of test conditions.

In terms of performance for Type I error rates, we conclude that with robust estimator M_{BGMed} -test is the best test in both cases (either n is small or large). Even though the test show the good properties such as high breakdown point of 50%, affine equivariant and better computational efficiency, this test is low in statistical power of test. In terms of power of test M-test and M_{GMed} -test performed well only when sample size is large than the variables. The best result in power of test is M_B -test in both cases (either n is small or large).

4.4 Real Application Example

In this study, the investigation on several robust M-test is continued with the application on real data. The US subprime financial crisis currencies are used to validate the new robust statistical M_{BGMed} -test and M_B -test since, the two tests are the best for Type I error rate and power of test respectively. So in real application there is need to validation which is the best among the two test.

4.4.1 Checking for the Outlier(s) of Largest Difference from the Mean

In this section, mean deviation (Tukey, 1960) is presented to show the presence of outliers by using R-package introduced by Komsta (2013). The real data set is transformed into the price return using Equation 3.9. In Figure 4.1, all the values above the black lines are outliers. The percentages of outlier presence before and after US subprime crisis are 5% and 5.01%, respectively.

The computation of the percentages of outliers are done by the following process.

Total data for before 480

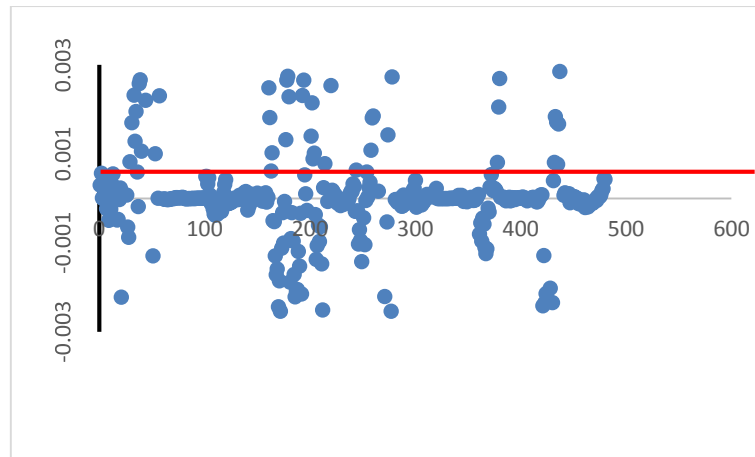
Outlier numbers 24

Percentage of outliers for before is $\frac{24}{480} * 100 = 5\%$

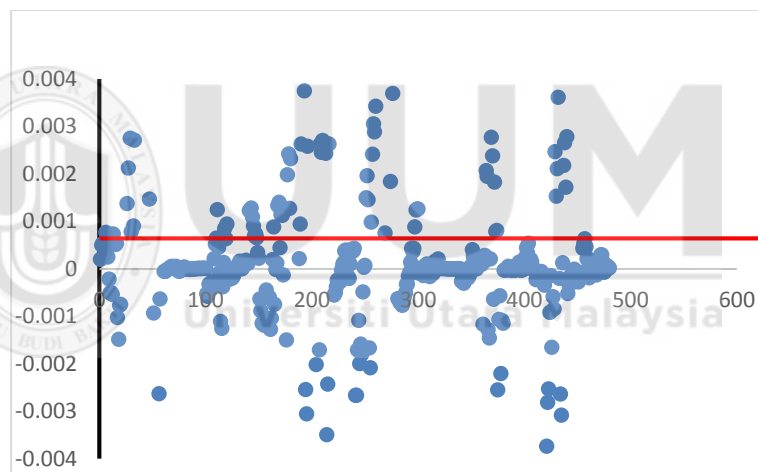
Total data for after 479

Outlier numbers 24

Percentage of outliers for after is $\frac{24}{479} * 100 = 5.01\%$



a) Before US subprime crisis



b) After US subprime crisis

Figure 4.1 Scatter plot of the sample mean deviation method (Tukey, 1960 and Maronna, Martin, Yohai & Salibián-Barrera, 2018).

Next, we continue the analysis by validating the new statistical test by using two approaches. Firstly, perform the equality test of two independent samples of covariance matrices and secondly, by the equality test of several independent samples of covariance matrices using control chart. However, we can now perform M-test due to the occurrence of singularity and outlier problems.

4.4.1 Testing the Equality of Two Covariance Matrices

In this section, we present the results of US subprime financial crisis using M_{BGMed} -test and M_{B} -test, for the purpose of performance investigation. The currencies from three continents countries were employed to analyse the differences in covariance matrices between (1st October 2006 until 31st October 2006) and (1st March 2008 until 31st March 2008), where it was one month before and one month after the US subprime financial crisis. There are 24 currencies from the America, Europe and Asia countries and sample size of two samples is equal to 19.

$$H_0: \Sigma_{\text{before}} = \Sigma_{\text{after}}$$

$$H_1: \Sigma_{\text{before}} \neq \Sigma_{\text{after}}$$

Furthermore, the result of Equation 3.3, M_{BGMed} -test is 388.86. Based on Chi-square distribution with $\frac{p(p-1)}{2} = \frac{24(24-1)}{2} = 276$ degree of freedom, $\chi^2_{(0.05, 276)} = 315.75$. The null hypothesis is rejected since the value of M_{BGMed} -test is greater than the critical value. Therefore, we conclude that there is a difference between two structures of covariance matrices. It shows that there is a shift in covariance when the subprime crisis occurred.

Secondly, the M_B -test is presented. Based on Chi-square distribution, we will not reject the null hypothesis because $M_B\text{-test} = 60.87$ is less than the critical value. Therefore, we conclude that there is no different between two structures of covariance matrices. It shows that there is no shift in covariance when the subprime crisis occurred.

The results of real application analysis using $M_{BGMed}\text{-test}$ and $M_B\text{-test}$ give two different results. The $M_{BGMed}\text{-test}$ is shown to be robust when compared with $M_B\text{-test}$.

4.4.2 Testing Several Covariance Matrices Using Control Chart

In this example, we test for several independent samples of covariance matrices using control chart. By utilizing this approach, the stability is the same to test the hypothesis of the similarity of two covariance matrices is done repeatedly.

$$H_0: \Sigma_i = \Sigma_0$$

$$H_1: \Sigma_i \neq \Sigma_0$$

where $i = 1, 2, \dots, m$, and Σ_0 is the reference sample.

In this example, we retrieved daily data of 25 major currencies from three continents which are America, European and Asian currencies. There are 16 months corresponding covariance matrices, with different sample size. Generally, there are 5 trading days per week, and about 18 to 21 days for a month excluding holiday and weekend. In Table 4.25, the sample size of the FOREX rate data.

Table 4.25

The Sample Size of the FOREX Rate Data

No.	n	size	No.	n	size
Nov 06	n_1	20	Jul 07	n_9	20
Dec 06	n_2	18	Aug 07	n_{10}	21
Jan 07	n_3	21	Sep 07	n_{11}	18
Feb 07	n_4	19	Oct 07	n_{12}	21
Mar 07	n_5	21	Nov 07	n_{13}	20
Apr 07	n_6	19	Dec 07	n_{14}	18
May 07	n_7	21	Jan 08	n_{15}	21
Jun 07	n_8	20	Feb 08	n_{16}	20

4.4.2.1 M_{BGMed} -test Control Chart

To construct the corresponding control chart for M_{BGMed} -test, we define $S_{M_{BGMed}(i)}$ is the sample covariance matrix unit calculated from sample i . Then, we compute pooled covariance matrix, $S_{M_{BGMed}}$. The M_{BGMed} -test is computed to every 16 months of samples and the results are shown in Table 4.26.

Table 4.26

The values of $M_{\text{BGMed-test}}$

i	M_{BGMed}	$\chi^2_{0.05,276}$	<i>Results</i>
1	1028.70	315.75	Reject H_0
2	901.36	315.75	Reject H_0
3	999.18	315.75	Reject H_0
4	1558.76	315.75	Reject H_0
5	979.42	315.75	Reject H_0
6	1373.36	315.75	Reject H_0
7	1395.67	315.75	Reject H_0
8	1152.39	315.75	Reject H_0
9	985.10	315.75	Reject H_0
10	773.72	315.75	Reject H_0
11	908.45	315.75	Reject H_0
12	977.78	315.75	Reject H_0
13	1106.56	315.75	Reject H_0
14	720.79	315.75	Reject H_0
15	1562.66	315.75	Reject H_0
16	729.78	315.75	Reject H_0

Table 4.26 present the values of M_{BGMed} -test, the first column represent 16 samples from November 2006 until February 2008. It is followed by the values of M_{BGMed} -test for 16 samples while, the third column represent the tabulated values for $\chi^2_{(0.05,276)}$, and the last column is the results.

In Figure 4.2, we present the M_{BGMed} control chart. At the significance level $\alpha = 0.05$, the upper limits of control chart $UCL = \chi^2_v = 315.75$ with degree of freedom, $v = \frac{p(p-1)}{2} = \frac{24(24-1)}{2} = 276$. All the 16 months in Table 4.141 are plotted in blue colour, UCL and LCL are in red colour. The vertical axis is the M_{BGMed} -test, where $i = 1, 2, \dots, 16$ and horizontal axis is 16 months.

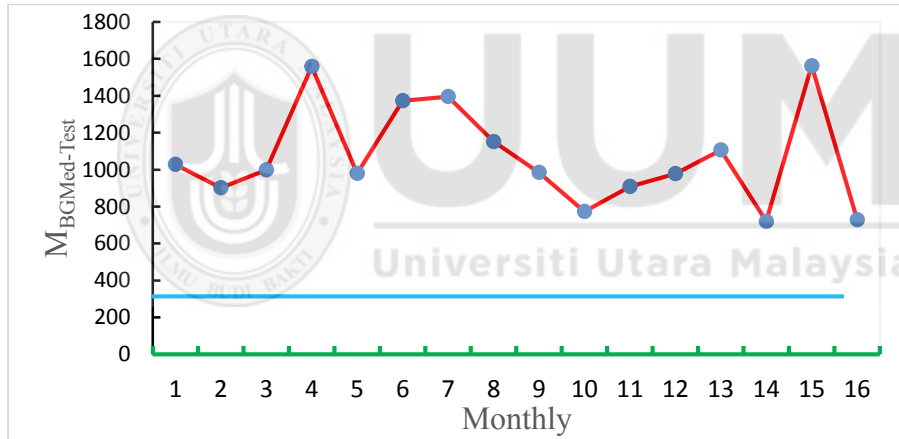


Figure 4.2. FOREX Control Chart for M_{BGMed}

Therefore, according to Table 4.26 and the Figure 4.2, we learn that all months give a signal of out of control. All months show that we reject the null hypothesis. This signal of out of control shows there is change in covariance structure at that particular months. We conclude from the result, that all the samples are not in control during the crisis.

In summary, Table 4.26 and Figure 4.2 presented the overall results and control chart for M_{BGMed} -test. There are 16 sample size conditions involved in evaluating the values of M_{BGMed} -test. All the results showed the signal of out of control compare to the Chi-square critical value. Therefore, we reject the null hypothesis and conclude that there are difference in the covariance structure during the US subprime financial crisis. Also, in real application M_{BGMed} -test is the best test, since the value is greater than the Chi-square critical value in cases of outlier and singularity problems in high dimensional data set.

4.4.2.2 M_B -test Control Chart

To construct the corresponding control chart for M_B -test, we define Σ_i . Then, we compute M_B -test pooled covariance matrix, $S_{M_B} = \frac{\sum_{i=1}^m (n_i - 1) S_{MB(i)}^{-1}}{n - m}$, where $S_{MB(i)}^{-1}$ is the sample covariance matrix unit calculated from sample $i = 1, 2, \dots, 16$. The M_B -test is computed to every 16 months of samples and the results are shown in Table 4.27.

Table 4.27 present M_B -test values, the first column represent the 16 samples i , from November 2006 to February 2008, for M_B -test values, then, next column represent the Chi-square values computed by $\chi^2_{(0.05, 276)} = 315.75$ and the last column is the results.

In Figure 4.3, we present the Chi-square control chart, with significance level of $\alpha = 0.05$ and the upper limits is $UCL = \chi^2_v = 315.75$ using degree of freedom, $v = \frac{p(p-1)}{2} = \frac{24(24-1)}{2} = 276$. All the 16 months in Table 4.27 are plotted in blue colour, UCL and LCL are in red colour. The vertical axis is the M_{BGMed} -test, where $i = 1, 2, \dots, 16$ and horizontal axis is the 16 months.

Table 4.27

The values of M_B -test

<i>i</i>	<i>M_B</i>	$\chi^2_{0.05,276}$	<i>Results</i>
1	471.49	315.75	Reject H_0
2	427.67	315.75	Reject H_0
3	354.53	315.75	Reject H_0
4	501.51	315.75	Reject H_0
5	293.25	315.75	Fail to Reject H_0
6	728.56	315.75	Reject H_0
7	509.20	315.75	Reject H_0
8	366.61	315.75	Reject H_0
9	399.07	315.75	Reject H_0
10	465.09	315.75	Reject H_0
11	282.41	315.75	Fail to Reject H_0
12	264.79	315.75	Fail to Reject H_0
13	340.16	315.75	Reject H_0
14	260.00	315.75	Fail to Reject H_0
15	544.02	315.75	Reject H_0
16	540.20	315.75	Reject H_0

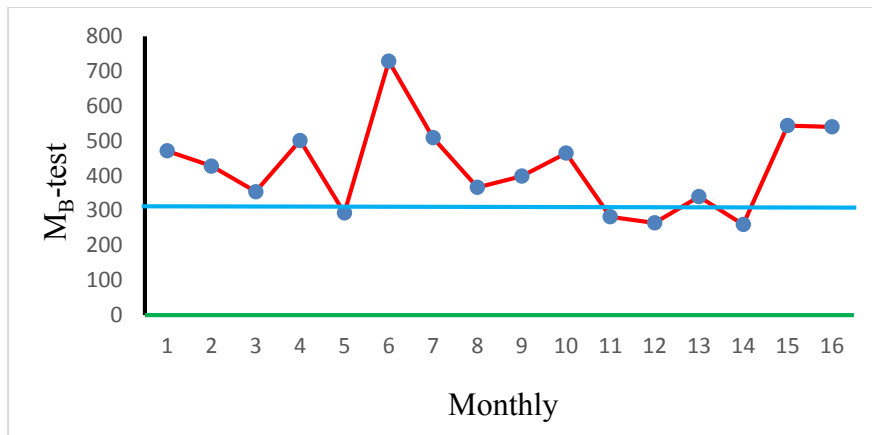


Figure 4.3. FOREX Control Chart for M_B -test

Therefore, according to Table 4.27 and the Figure 4.3, we learn that all months give a signal of out of control except on 5th (March 2007), 11th (September 2007), 12th (October 2007) and 14th (December 2007) sample. The four months show that there is failed to reject the null hypothesis. This signal of out of control shows there is change in covariance structure at that particular months. We conclude from the result, these four samples are stable and the others are not stable during the crisis.

In summary, Table 4.27 and Figure 4.3 presented the overall results and control chart for M_B -test. There are 16 sample size conditions involved in evaluating the values of M_B -test. The results showed out of control compare to the Chi-square critical value for four months is in control. Thus, we fail to reject the null hypothesis and conclude there are no difference in the covariance structure during the US subprime financial crisis. Also, in real application M_B -test value is less than the Chi-square critical value in cases of outlier and singularity problems when, $p > n$.

Testing the hypothesis repeatedly is more advantageous since we can investigate all the corresponding covariance matrices independently. The results of real application analysis using M_{BGMed} and M_B charts also give two different results. The M_{BGMed} chart is dominant compared to M_B chart since that chart is more sensitive. The M_B control chart cannot detect the covariance shift in 5th, 11th, 12th and 14th, sample covariance matrices. The two examples show M_{BGMed} -test as the best that resolve the outliers and singularity problems in high dimensional dataset.



CHAPTER FIVE

DISCUSSION AND RECOMMENDATION

5.1 Findings and Conclusions

The ultimate goal of this study is to search for a new robust covariance equality testing in the presence of outliers and resolving singularity problems of M-test. In achieving this goal, the modification on M-test based on mean estimator (M-test) and geometric median estimator (M_{GMed} -test) are presented in Chapter Four. Both with and without robust estimators are used to handle the problems of outlier and singularity.

The Type I error rates were measured and compared to investigate the performance of M-test, M_{GMed} -test, M_B -test, M_T -test, M_{BGMed} -test, and M_{TGMed} -test using simulated data. The robustness of M-test and M_{GMed} -test differs among small, medium and large number of variables. The Type I error for small number of variables ($p = 3, 5$ and 8) show that, six tests were investigated from $p = 3$ with $n = 2, 5, 10, 15, 20, 25$ and 50 , M_{BGMed} -test, dominate the other tests, except when $n = 15$ and 50 where M_{TGMed} -test and M_{GMed} -test dominate the other tests respectively. Therefore, under small number of variable $p = 3$, we conclude that M_{BGMed} -test is the best robust tests for both cases (either n is small or large).

Also, when $p = 5$ with $n = 2, 5, 10, 15, 20, 25$ and 50 , seven tests were investigated, M_{BGMed} -test, dominate the other tests, except when $n = 15$ and 50 where M_{TGMed} -test dominate the other tests. Therefore, when $p = 5$, we conclude that M_{BGMed} -test as the best robust test for both cases (either n is small or large).

Similarly, when $p = 8$ with $n = 2, 5, 10, 15, 20, 25$ and 50 , six tests investigated, M_{BGMed} -test dominate the other tests except when $n = 20$. Therefore, when $p = 8$, we conclude that M_{BGMed} -test is the best robust test for both cases (either n is small or large). Therefore, under small number of variables ($p = 3, 5$ and 8) we conclude that with robust estimator, M_{BGMed} -test is the best robust test for both cases (either n is small or large).

In detail, for medium number of variables when ($p = 10, 15$ and 20), for $n = 5, 10, 15, 20, 25$ and 50 . Six tests investigated when $p = 10$, we conclude that M_{BGMed} -test is the best robust test for both cases (either n is small or large). With $p = 15$, we conclude that M_{BGMed} -test is the best robust test for both cases (either n is small or large) and finally on medium variable when $p = 20$, we conclude that M_{BGMed} -test is the best robust test for both cases (either n is small or large).

Likewise, when there are large number of variables ($p = 25, 30$ and 50), for $p = 25$, we conclude that M_{BGMed} -test is the best robust test for both cases (either n is small or large). Also, when $p = 30$, we conclude that M_{BGMed} -test is best robust test for both cases (either n is small or large). Lastly, when $p = 50$, we conclude that M_{BGMed} -test is the best robust test for both cases (either n is small or large). In the Type I error rate test investigated M_{BGMed} -test is the best in all the conditions.

The power of the test for small number of variables ($p = 3, 5$ and 8) show that there are 275 conditions of M -test, 216 conditions of M_{GMed} -test, 191 conditions of M_B -test, 192 conditions of M_T -test, 28 conditions of M_{BGMed} -test and 35 conditions of M_{TGMed} -test that fall within the power interval.

For $p = 3$, we conclude that M-test is the best test when n large. Also, under $p = 5$, we conclude that M-test is the best test in cases n is large. Under $p = 8$, we conclude that M-test is the best test in cases n is large. Therefore, under small number of variables, we conclude that M-test is the best in cases n is large.

The power of test for medium number of variables ($p = 10, 15$ and 20) show that there are 180 conditions of M-test, 178 conditions of M_{GMed} -test, 271 conditions of M_B -test, 262 conditions of M_T -test, 114 conditions of M_{BGMed} -test and 134 conditions of M_{TGMed} -test that fall within the power interval. While under medium number of variables without robust estimator, M_B -test is the best test in both cases (either n is small or large). Its power values dominate the other tests especially when $p > n$. For when $p = 10$ we conclude that M_T -test is the best test when n is large. Also, when $p = 15$, we conclude that M_B -test is the best robust test in both cases (either n is small or large). And when $p = 20$, we conclude that M_T -test is the powerful test in both cases (either when n is small or large). Therefore, under medium number of variables ($p = 10, 15$ and 20), we conclude that M_B -test is the best test in both (either n is small or large). But M-test performed better when $n = 50$.

The power of test under large number of variables ($p = 25, 30$ and 50) show that there are 28 conditions of M-test, 28 conditions of M_{GMed} -test, 321 conditions of M_B -test, 322 conditions of M_T -test, 175 conditions of M_{BGMed} -test and 175 conditions of M_{TGMed} -test that fall within the power interval. For when $p = 25$, we conclude that M_B -test is the best robust test in both cases (either n is small or large). Also, when $p = 30$, we conclude that M_T -test is the best robust test in both cases (either n is small or large).

And lastly, when $p = 50$ and $n = 50$, we conclude that, M_B -test and M_T -test are the best tests both (either n is small or large). Therefore, under large number of variables ($p = 25, 30$ and 50), we conclude that M_T -test is the best test in both cases (either n is small or large). Its power values dominate the other tests especially when $p > n$. Furthermore, the power of test without robust estimator show that M -test performed worst when $p = 30$ and 50 in all condotions of the sample size, n and usually less than with robust M_{BGMed} -test and M_{TGMed} -test when n is small.

In general, the power of test with robust estimator show that M_{BGMed} -test and M_{TGMed} -test also performed well in both cases (either n is small or large). It power values also, performed well as the sample sizes increases, especially when $p > n$. Moreover, to validate the new robust statistical tests M_{BGMed} -test and M_B -test, the currencies from three continents such as Europe, America and Asia were used to analyze the difference in the covariance matrices before and after US subprime financial crisis. These statistical tests are applied to test hypothesis in high dimensional data set with presence of outlier and singularity problems by using two approaches.

Firstly, on testing the equality of two covariance matrices with before and after US subprime financial crisis data. The result of Equation 3.3, M_{BGMed} -test is 388.86. Based on Chi-square distribution with 276 degree of freedom, $\chi^2_{(0.05,276)} = 315.75$. The null hypothesis is rejected since the value of M_{BGMed} -test is greater than the critical value. And for the M_B -test, we conclude that there is a different between the two structures of covariance matrices. It shows that there is a shift in covariance when the subprime crisis occurred.

The M_B -test is presented, based on Chi-square distribution, we failed to reject the null hypothesis since the value of $M_B\text{-test} = 60.87$ is less than the critical value. Therefore, we conclude that there is no different between the two structures of covariance matrices. It shows that there is no shift in covariance when the subprime crisis occurred.

Second approach on testing several covariance matrices using control chart: firstly, the M_{BGMed} -test control chart, according to Table 4.26 and the Figure 4.2, we learn that at all months give a signal of out of control. Thus, we reject the null hypothesis. This signal of out of control shows there is changes in covariance structure at that particular months, and all the samples are not stable during the crisis. Therefore, reject the null hypothesis and conclude that there are difference in the covariance structure during the US subprime financial crisis. Also, in real application M_{BGMed} -test is the best test, since the value is greater than the chi square critical value in cases of outlier and singularity problems when, $p > n$.

Also, for M_B -test Control chart, according to Table 4.27 and the Figure 4.3, reveals that all months give a signal of out of control except on 5th (March 2007), 11th (September 2007), 12th (October 2007) and 14th (December 2007) sample. The four months show that there is failed to reject the null hypothesis. This signal of out of control shows there is changes in covariance structure at that particular months. Conclude from the result, these four samples are stable and the others are not stable during the crisis.

In conclusion, M_{BGMed} -test is the best compared to M_B -test in terms of stability of the covariance structure during the US subprime financial crisis in two different tests, and can handled both cases (either n is small or large). M_{BGMed} -test is more suitable in handling real application data set.

5.2 Future Research

In summary, the results of this thesis are expected to provide additional explanations and approaches in testing and analyzing of M-test. However, there are always rooms for improvement.

- i. We would like to suggest another alternative estimator, i.e Weiszfeld's algorithm for M-test that will resolve the outliers and singularity problems in high dimensional data sets.
- ii. Additionally, there are still some unanswered questions related to the positive definiteness in high dimensional cases with outliers. Banding and Tapering estimators does not necessarily guarantee positive definiteness in high dimensional data sets.



REFERENCES

- Abdi, H. (2007). O'Brien test for homogeneity of variance. *Encyclopedia of Measurement and Statistics*, 2, 701-704.
- Acharya, V., Philippon, T., Richardson, M., & Roubini, N. (2009). The financial crisis of 2007-2009: Causes and remedies. *Financial markets, institutions & instruments*, 18(2), 89-137.
- Aggarwal, C. C., & Sathe, S. (2017). Which Outlier Detection Algorithm Should I Use?. In *Outlier Ensembles* (pp. 207-274). Springer, Cham.
- Aggarwal, C. C. (2017). *Outlier Analysis*, 2nd edition. Springer Publishing Company, Incorporated.
- Ahad, N. A., Othman, A. R., Syed Yahaya, S.S. (2012). Performance of two-samples pseudo-median procedure. *Sains Malaysiana*, 41(9), 1149-1154.
- Aitchison, J. (2003). *The Statistical Analysis of Compositional Data*. The Blackburn Press: Cladwell, New Jersey.
- Ali, H. (2013), "Efficient and highly robust Hotelling T^2 Control chart using Reweighted minimum vector variance", "A Ph.D thesis submitted to Universiti Utara Malaysia".
- Allen, G. I. (2012). Regularized tensor factorizations and higher-order principal components analysis. *arXiv preprint arXiv:1202.2476*.
- Allen, J. (2015). A Bayesian Hierarchical Selection Model for Academic Growth with Missing Data. WP-2015-04.ACT working papers document preliminary research. The papers are intended to promote discussion and feedback before formal publication. The research does not necessarily reflect the views of ACT.
- Alqallaf, F., Van Aelst, S., Yohai, V. J. & Zamar, R. H. (2009). Propagation of outliers in multivariate data. *Ann. Statist.* 37, 311–31.
- Annaert, J., Claes, A. G., & De Ceuster, M. J. (2006). Intertemporal stability of the European credit spread co-movement structure1. *The European Journal of Finance*, 12(1), 23-32.
- Arnold, S. J., Bürger, R., Hohenlohe, P. A., Ajie, B. C., & Jones, A. G. (2008). Understanding the evolution and stability of the G-matrix. *Evolution: International Journal of Organic Evolution*, 62(10), 2451-2461.
- Aslam, S., & Rocke, D. M. (2005). A robust testing procedure for the equality of covariance matrices. *Journal for Computational Statistics & Data Analysis*, 49(3), 863-874.

- Auerbach, S., & Dix, R. (2017). Competition and spatial efficiency. Ayyildiz, E., Gazi, V. P., & Wit, E. (2012). A short note on resolving singularity problems in covariance matrices. *International Journal of Statistics and Probability*, 1(2), p113.
- Auer, P. L., Reiner, A. P., & Leal, S. M. (2016). The effect of phenotypic outliers and non-normality on rare-variant association testing. *European Journal of Human Genetics*, 24(8), 1188.
- Alt, F. B. and Smith, N. D. (1988). Multivariate Process Control. *Handbook of Statistical*. 7, 333-351.
- Ayyildiz, E., Gazi, V. P., & Wit, E. (2012). A short note on resolving singularity problems in covariance matrices. *International Journal of Statistics and Probability*, 1(2), 113.
- Bai, J., & Shi, S. (2011). Estimating high dimensional covariance matrices and its Applications. *Discussion Paper No.: 1112-03* <http://hdl.handle.net/10022/AC:P:11064>. Columbia University New York, NY 1002.
- Bajwa, W. U., Duarte, M. F., & Calderbank, R. (2014). Average case analysis of high dimensional block-sparse recovery and regression for arbitrary designs. In *Proceedings of the Seventeenth International Conference on Artificial Intelligence and Statistics* (pp. 57-67).
- Bargiela, A., & Pedrycz, W. (2016). Granular computing. In *Handbook On Computational Intelligence: Volume 1: Fuzzy Logic, Systems, Artificial Neural Networks, and Learning Systems* (pp. 43-66).
- Barnett, V., & Lewis, T. (1994). *Outliers in statistical data* (Vol. 3, No. 1). New York: Wiley.
- Bartlett, M. S. (1937). Properties of sufficiency and statistical tests. *Proceedings of the Royal Society of London, A*, 160, 268-282.
- Bartlett, M. S., & Kendall, D. G. (1946). The statistical analysis of variance-heterogeneity. and the logarithmic transformation. *J. Roy. Statistics. Sociology., Supplement*. 8, 128-138. based social space and labor flow network communities. *Computational and Mathematical Organization Theory*, 24(3), 378-400.
- Breiman, L. (2017). Classification and regression trees. *Routledge*.
- Beckman, R. J., & Cook, R. D. (1983). Outliers. *Technometrics*, 25(2), 119-149.
- Beirlant, J., Claeskens, G., Croux, C., Degryse, H., Dewachter, H., Dhaene, G., & Willekens, M. (2005). Managing Uncertainty: Financial, Actuarial and Statistical Modeling. *Tijdschrift voor Economie en Management*, 50(1), 23-49.

- Benneyan, J. C. (1998). Statistical quality control methods in infection control and hospital epidemiology, Part I Introduction and Basic Theory. *Infection Control*, 19(03), 194-214.
- Bentler, P. M., & Dudgeon, P. (1996). Covariance structure analysis: Statistical practice, theory, and directions. *Annual review of psychology*, 47(1), 563-592.
- Berthet, Q., & Rigollet, P. (2013). Optimal detection of sparse principal components in high dimension. *The Annals of Statistics*, 41(4), 1780-1815.
- Bickel, P. J. & Levina, E. (2004). Some theory of Fisher's linear discriminant function, 'naive Bayes,' and some alternatives when there are many more variables than observations. *Bernoulli* 10 989-1010. MR2108040.
- Bickel, P. J. & Levina, E. (2008a). Regularized estimation of large covariance matrices. *Ann. Statist.* 36 199-227. MR2387969.
- Bickel, P. J. & Levina, E. (2008b). Covariance regularization by thresholding. *Ann. Statist.* 36 2577-2604. MR2485008.
- Bien, J., Bunea, F., & Xiao, L. (2016). Convex banding of the covariance matrix. *Journal of the American Statistical Association*, 111(514), 834-845.
- Blau, B. M. (2018). Exchange rate volatility and the stability of stock prices. *International Review of Economics & Finance*.
- Bougioukas, K. I., & Haidich, A. B. (2019). Medical Biostatistics: Basic Concepts. In *Management of Hypertension*(pp. 19-53). Springer, Cham.
- Boudt, K., Rousseeuw, P., Vanduffel, S., & Verdonck, T. (2018). The minimum regularized covariance determinant estimator.
- Box, G. E., & Watson, G. S. (1962). "Robustness to non-normality of regression tests". *Biometrika*, 49(1-2), 93-106.
- Box, G., E. (1949). A general distribution theory for a class of likelihood criteria. *Biometrika*, 36: 317-346.
- Bradley, J.V. (1978). Robustness? *British Journal of Mathematical and Statistical Psychology*, 31, 144-152.
- Brazil, M., Graham, R. L., Thomas, D. A., & Zachariasen, M. (2014). On the history of the Euclidean Steiner tree problem. *Archive for history of exact sciences*, 68(3), 327-354.
- Brown T. A. (2006). *Confirmatory Factor Analysis for Applied Research*. New York: The Guilford Press.

- Bruce, D., & Vershynin, R. (2011). A multivariate median in banach spaces and applications to robust pca. *Kättesaadav: <http://www-personal.umich.edu/~romanv/students/bruce-REU.pdf>* (26.04. 15).
- Brueckner, J. K., Calem, P. S., & Nakamura, L. I. (2012). Subprime mortgages and the housing bubble. *Journal of Urban Economics*, 71(2), 230-243.
- Byrne, H. J., Knief, P., Keating, M. E., & Bonnier, F. (2016). Spectral pre and post processing for infrared and Raman spectroscopy of biological tissues and cells. *Chemical Society Reviews*, 45(7), 1865-1878.
- Cai, C. Zhang, & Zhou, H. (2010). Optimal rates of convergence for covariance matrix estimation. *Annals of Statistics*, 38:2118-2144.
- Cai, T. T., Ren, Z., & Zhou, H. H. (2014). Estimating structured high-dimensional covariance and precision matrices: Optimal rates and adaptive estimation. *The Annals of Statistics*, 38, 2118-2144.
- Cai, T., & Liu, W. (2011). Adaptive thresholding for sparse covariance matrix estimation. *Journal of the American Statistical Association*, 106(494), 672-684.
- Campbell, N., Lopuhaä, H., & Rousseeuw, P. (1998). On the calculation of a robust S-Estimator of a covariance matrix. *Journal of Statistics in Medicine*, 17(23).
- Cao, Y., Lin, W., & Li, H. (2018). Large covariance estimation for compositional data via composition-adjusted thresholding. *Journal of the American Statistical Association*, (just-accepted), 1-45.
- Cardot, H., Cénac, P., & Zitt, P. A. (2013). Efficient and fast estimation of the geometric median in Hilbert spaces with an averaged stochastic gradient algorithm. *Bernoulli*, 19(1), 18-43.
- Cerioni, A., Riani, M., Atkinson, A. C., & Corbellini, A. (2018). The power of monitoring: how to make the most of a contaminated multivariate sample. *Statistical Methods & Applications*, 1-29.
- Chambers, R., Chandra, H., Salvati, N., & Tzavidis, N. (2014). Outlier robust small area estimation. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 76(1), 47-69.
- Chatfield, C., & Collins, A. J. (2013). *Introduction to multivariate analysis*. Springer.
- Chatfield, C. (2018). *Introduction to multivariate analysis*. Routledge.
- Chatterjee, S. (2015). Matrix estimation by universal singular value thresholding. *The Annals of statistics*, 43(1), 177-214.
- Chen, S. X. & Qin, Y. L. (2010). A two-sample test for high-dimensional data with applications to gene-set testing. *Ann. Statist.* 38, 808–35.

- Cheng, J., Levina, E., Wang, P., & Zhu, J. (2014). A sparse Ising model with covariates. *Biometrics*, 70(4), 943-953.
- Cordeiro, G. M., & de Paula Ferrari, S. L. (1991). A modified score test statistic having chi-squared distribution to order $n-1$. *Biometrika*, 78(3), 573-582.
- Choi, Y., Lim, J., Roy, A., & Park, J. (2016). Positive-definite modification of covariance matrix estimators via linear shrinkage. https://www.researchgate.net/.../305881430_Positivedefinite_modification_of_covariance.
- Cohen, J. (1992). A power primer. *Psychological Bulletin*, 112, 155-159.
- Cormode, G., & McGregor, A. (2008). Approximation algorithms for clustering uncertain data. In *Proceedings of the twenty-seventh ACM SIGMOD-SIGACT-SIGART symposium on Principles of database systems* (pp. 191-200). ACM.
- Croux, C., & Haesbroeck, G. (2000). Principal component analysis based on robust estimators of the covariance or correlation matrix: influence functions and efficiencies. *Biometrika*, 87(3), 603-618.
- Croux, C., & Ruiz-Gazen, A. (2005). High breakdown estimators for principal components: the projection-pursuit approach revisited. *Journal of Multivariate Analysis*, 95(1), 206-226.
- Dan, E., D. & Ogu, A., I. (2013). Statistical analysis/methods of detecting outliers in a univariate data in a regression analysis model. *International Journal of Education and Research Vol. 1 No. 5 ISSN: 2201-6333 (Print) ISSN: 2201-6740 (Online) www.ijern.com*.
- Danaher, P., Paul, D., & Wang, P. (2015). Covariance-based analyses of biological pathways. *Biometrika*, 102(3), 533-544.
- Daniel, W. W. (1990). Procedures that utilize data from two independent samples. *Applied Nonparametric Statistic. 2nd ed. Boston, Massachusetts: PWS-Kent Publishing Co*, 82-143.
- Das, K. R., & Imon, A. R. (2014). Geometric median and its application in the identification of multiple outliers. *Journal of Applied Statistics*, 41(4), 817-831.
- De Freitas, L. (2018). The mathematical continuum: A haunting problematic. *The Mathematics Enthusiast*, 15(1), 148-158.
- Deng, X., & Tsui, K. W. (2013). Penalized covariance matrix estimation using a matrix-logarithm transformation. *Journal of Computational and Graphical Statistics*, 22(2), 494-512.

- Diakonikolas, I., Kamath, G., Kane, D. M., Li, J., Moitra, A., & Stewart, A. (2017). Being robust (in high dimensions) can be practical. *arXiv preprint arXiv:1703.00893*.
- Djauhari, M. A. (2005). Improved Monitoring of Multivariate Process Variability. *Journal of Quality Technology*, 37(1), 32-39.
- Djauhari, M. A., & Herdiani, E. T. (2008). Monitoring the stability of correlation structure in drive rib production process: an MSPC approach. *Open Industrial & Manufacturing Engineering Journal*, 1, 8-18.
- Djauhari M. A., (2005). Improved monitoring of multivariate process variability. *Journal of Quality Technology*, 37(1): 32-39.
- Djauhari, M. A. (2005). Improved monitoring of multivariate process variability. *Journal of Quality Technology*, 37(1), 32-39.
- Dobriban, E. (2015). Efficient Computation of Limit Spectra of Sample Covariance Matrices. *arXiv preprint arXiv:1507.01649*.
- Donoho, D. & Jin, J. (2004). Higher criticism for detecting sparse heterogeneous mixtures. *Ann. Statist.* 32 962–994. MR2065195.
- Drezner, Z. & Hamacher, H. W. (2001). *Facility location: applications and theory*. Springer Science & Business Media.
- Duan, Q., Xu, Z., Raffield, L. M., Chang, S., Wu, D., Lange, E. M., & Li, Y. (2018). A robust and powerful two-step testing procedure for local ancestry adjusted allelic association analysis in admixed populations. *Genetic epidemiology*, 42(3), 288-302.
- Durocher, S., Leblanc, A., & Skala, M. (2017). The projection median as a weighted average. *Journal of Computational Geometry*, 8(1), 78-104.
- DuVal, E. H., Vanderbilt, C. C., & M'Gonigle, L. K. (2018). The spatial dynamics of female choice in an exploded lek generate benefits of aggregation for experienced males. *Animal Behaviour*.
- Dymski, G., & Dymski, G. A. (2010). *Understanding the subprime crisis: institutional evolution and theoretical views*. Department of Economics University of California, Riverside Riverside CA 92521 Email: gary.dymski@ucr.edu.
- Eftelioglu, E. (2015). Geometric median. Computer science and Engineering, university of Minnesota, Minneapolis, MN, USA. Springer International Publishing Switzerland. S. Shekhar et al.(eds.), *Encyclopedia of GIS*, DOI 10.1007/978-3-319-23519-6_1555-1.
- El Karoui, N. (2010), "High-Dimensionality Effects in the Markowitz Problem and Other Quadratic Programs with Linear Constraints", Risk Underestimation1. *The Annals of Statistics* 2010, Vol. 38, No. 6, 3487–3566 DOI: 10.1214/10-AOS795 © Institute of Mathematical Statistics.

- Engel, J., Buydens, L., & Blanchet, L. (2017). An overview of large-dimensional covariance and precision matrix estimators with applications in chemometrics. *Journal of Chemometrics*, 31(4).
- Estaghvirou, S. B. O., Ogutu, J. O., & Piepho, H. P. (2014). Influence of outliers on accuracy estimation in genomic prediction in plant breeding. *G3: Genes|Genomes|Genetics*, 4(12), 2317-2328.
- Fama, E. F. (1965). The behavior of stock-market prices. *The journal of Business*, 38(1), 34-105.
- Fan, J., Han, F., & Liu, H. (2014). Challenges of big data analysis. *National science review*, 1(2), 293-314.
- Fan, J., Liao, Y., & Mincheva, M. (2013). Large covariance estimation by thresholding principal orthogonal complements. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 75(4), 603-680.
- Fan, J., Liao, Y., & Liu, H. (2015), "An Overview on the Estimation of Large Covariance and Precision Matrices", Department of ORFE, Sherrerd Hall, Princeton University, Princeton, NJ 08544, USA, e-mail: jqfan@princeton.edu, yuanliao@umd.edu, hanliu@princeton.edu. Fan's research was supported by National Science Foundation grants DMS-1206464 and DMS-1308566.
- Fan, J., Wang, W., & Zhu, Z. (2016). A Shrinkage Principle for Heavy-Tailed Data: High-Dimensional Robust Low-Rank Matrix Recovery. *arXiv preprint arXiv:1603.08315*.
- Fang, Y., Wang, B., & Feng, Y. (2016). Tuning-parameter selection in regularized estimations of large covariance matrices. *Journal of Statistical Computation and Simulation*, 86(3), 494-509.
- Filzmoser, P., & Hron, K. (2008). Outlier detection for compositional data using robust methods. *Mathematical Geosciences*, 40(3), 233-248.
- Filzmoser, P., Maronna, R. & Werner, M. (2008). Outlier identification in high dimensions. *Comp. Statist. Data Anal.* 52, 1694–711.
- Finch, H. (2012). Distribution of variables by method of outlier detection. *Frontiers in psychology*, 3, 211.
- Fletcher, P. T., Venkatasubramanian, S., & Joshi, S. (2009). The geometric median on Riemannian manifolds with application to robust atlas estimation. *NeuroImage*, 45(1), S143-S152.
- Franco-Pedroso, J., Gonzalez-Rodriguez, J., Cubero, J., Planas, M., Cobo, R., & Pablos, F. (2018). Generating virtual scenarios of multivariate financial data for quantitative trading applications. *arXiv preprint arXiv:1802.01861*.

- Fritsch, V., Varoquaux, G., Thyreau, B., Poline, J. B. & Thirion, B. (2011). Detecting outlying subjects in high-dimensional neuroimaging datasets with regularized minimum covariance determinant. In *Medical Image Computing and Computer Assisted Intervention, Part III*, Eds. G. Fichtinger, A. Martel and T. Peters, pp. 264–71. Heidelberg: Springer.
- Fryzlewicz, P. (2013). High-dimensional volatility matrix estimation via wavelets and thresholding. *Biometrika*, 100(4), 921-938.
- Furrer, R., & Bengtsson, T. (2007). Estimation of high-dimensional prior and posterior covariance matrices in Kalman filter variants. *Journal of Multivariate Analysis*, 98(2), 227-255.
- Gilleland, E., Hering, A. S., Fowler, T. L., & Brown, B. G. (2018). Testing the Tests: What Are the Impacts of Incorrect Assumptions When Applying Confidence Intervals or Hypothesis Tests to Compare Competing Forecasts?. *Monthly Weather Review*, 146(6), 1685-1703.
- Ghoting, A., Parthasarathy, S., & Otey, M. E. (2008). Fast mining of distance-based outliers in high-dimensional datasets. *Data Mining and Knowledge Discovery*, 16(3), 349-364.
- Ghute, V. B., & Shirke, D. T. (2008). A multivariate synthetic control chart for process dispersion. *Quality Technology & Quantitative Management*, 5(3), 271-288.
- Greenland, S., Senn, S. J., Rothman, K. J., Carlin, J. B., Poole, C., Goodman, S. N., & Altman, D. G. (2016). Statistical tests, P values, confidence intervals, and power: a guide to misinterpretations. *European journal of epidemiology*, 31(4), 337-350.
- Garson, G. D. (2012). Testing statistical assumptions. *Asheboro, NC: Statistical Associates Publishing*.
- Grice, J. W., & Iwasaki, M. (2007). A truly multivariate approach to MANOVA. *Applied Multivariate Research*, 12(3), 199-226.
- Grote, M. N. (2007). A covariance structure model for the admixture of binary genetic variation. *Genetics*, 176(4), 2405-2420.
- Haldane, J. B. S. (1948). Note on the median of a multivariate distribution. *Biometrika* 35, 414.
- Hamed, M. S. (2014). Generalized variance chart for multivariate quality control process procedure with application. *Applied Mathematical Sciences*, Vol. 8, 2014, no. 163, 8137 - 8151 HIKARI Ltd, www.mhikari.com/http://dx.doi.org/10.12988/ams.2014.49734.

- Hamilton, S. T. (2015). Deregulation and The 2007-2008 Housing/Debt Crisis Analysis of the Housing/Debt Crisis of 2007-2008 and its impact on the Financial Strength and Vulnerability of the United States and Global Economy(Doctoral dissertation).
- Hampel, F. R. (1971). A general qualitative definition of robustness. *The Annals of Mathematical Statistics*, 1887-1896.
- Hampel, F. R. (1974). The influence curve and its role in robust estimation. *Journal of the American Statistical Association*, 69(346), 383-393.
- Hampel, F. R. (1985). The breakdown points of the mean combined with some rejection rules. *Technometrics*, 27(2), 95–107.
- Hampel, F. R., Ronchetti, E. M., Rousseeuw, P. J., & Stahel, W. A. (1986). Linear models: robust estimation. *Robust Statistics: The Approach Based on Influence Functions*, 307-341.
- Harris, P., Brunsdon, C., Charlton, M., Juggins, S., & Clarke, A. (2014). Multivariate spatial outlier detection using robust geographically weighted methods. *Mathematical Geosciences*, 46(1), 1-31.
- Hartley, H. O. (1940). Testing the Homogeneity of a set of variances, *Biometrika*, 31, 249-255.
- Hasbini, A., Homes, S., & Satterfield, R. K. (2018). Navigating the Turbulent Economic Waters of the 2007 Great Recession and the Crash of the Housing Market. *Industry Analysis*. Vol. 2, Number 1.
- Hawkins, D. M. (1980). *Identification of outliers* (Vol. 11). London: Chapman and Hall.
- He, Q., Zhang, H. H., Avery, C. L., & Lin, D. Y. (2016). Sparse meta-analysis with high-dimensional data. *Biostatistics*, 17(2), 205-220.
- Hendry, D. F. (2018). Deciding between alternative approaches in macroeconomics. High Dimension Data Set. *International Journal of Applied Engineering Research*, 12(16), 6241-6249.
- Hildenbrand, D. (2006). *Geometric computing in computer graphics and robotics using conformal geometric algebra* (Doctoral dissertation, Hildenbrand, Dietmar).
- Hu, Y., Wu, H. L., Yin, X. L., Gu, H. W., Liu, Z., Xiao, R., & Yu, R. Q. (2018). A flexible and novel strategy of alternating trilinear decomposition method coupled with two-dimensional linear discriminant analysis for three-way chemical data analysis: Characterization and classification. *Analytica chimica acta*, 1021, 28-40.

- Hua, X., Cheng, Y., Wang, H., Qin, Y., & Li, Y. (2017). Geometric means and medians with applications to target detection. *IET Signal Processing*, 11(6), 711-720.
- Huang, J., Liu, N., Pourahmadi, M. & Liu, L. (2006). Covariance matrix selection and estimation via penalised normal likelihood. *Biometrika* 93, 85–98.
- Huang, S., Wang, Y., Xie, Y., Zhao, P., & Lüers, J. (2016). OutlierFlag: A Tool for Scientific Data Quality Control by Outlier Data Flagging. *Journal of open Research Software*, 4(1).pg 2.
- Huber, P. J. (1964). Robust estimation of a location parameter. *The Annals of Mathematical Statistics*, 35(1), 73-101.
- Hubert, M., & Engelen, S., (2004). Robust PCA and classification in biosciences. *Bioinformatics*, 20(11), 1728-1736.
- Jabara, C. (2009). How do Exchange Rates Affect Import Prices? Recent Economic Literature and Data Analysis, Office of Industries Working Paper No. ID 21. Washinton, DC: US International Trade Commission.
- Jaffee, D. M. (2009). The US subprime mortgage crisis: Issues raised and lessons learned. *Urbanization and growth*, 197.
- James, G.M., & Sugar, C. A.(2003). Clustering for sparsely sampled functional data *Journal of the America Statistical Association*, 98(462), 397-408.
- Janus, P., Koopman S. J. & Lucas A. (2014). Long memory dynamics for multivariate dependence under heavy tails, *Journal of Empirical Finance* 29, 187-206
- Jennrich, R. I. (1970). “Test for the equality of two correlation matrices”. *Journal of the American Statistical Association*, Vol. 65, No. 330, pp. 904-912.
- Jensen, W.A., Birch, J.B. & Woodall, W.H. (2007). High breakdown estimation methods for Phase I multivariate control charts, *Qual. Reliab. Eng. Int.* 23, pp. 615–629. doi: 10.1002/qre.837, <http://dx.doi.org/10.1002/qre.837>.
- Johnson, R. A., & Wichern, D. W. (1992). Applied multivariate statistical analysis. Prentice Hall, Englewood Cliffs, NJ. *Applied multivariate statistical analysis. Prentice-Hall, Englewood Cliffs, NJ.*
- Kang, X., & Deng, X. (2018). Ensemble Estimation of Large Sparse Covariance Matrix Based on Modified Cholesky Decomposition. *arXiv preprint arXiv:1801.00380*.
- Karoui, N., E. (2008). Spectrum estimation for large dimensional covariance matrices using random matrix theory. *The Annals of Statistics*, 2757-2790.

- Karrila, S., Lee, J. H. E., & Tucker-Kellogg, G. (2011). A comparison of methods for data-driven cancer outlier discovery, and an application scheme to semisupervised predictive biomarker discovery. *Cancer Informatics*, 10, CIN-S6868.
- Kaufman L. & Rousseeuw P. (2005). *Finding Groups in Data: An Introduction to Cluster Analysis*. Hoboken, NJ: John Wiley & Sons, Inc.
- Kaufman, C. G., Schervish, M. J., & Nychka, D. W. (2008). Covariance tapering for likelihood-based estimation in large spatial data sets. *Journal of the American Statistical Association*, 103(484), 1545-1555.
- Kemperman, J. H. B. (1987). The median of a finite measure on a Banach space. In *Statistical Data Analysis Based on the L1-norm and Related Methods* Neuchâtel, 1987) 217–230. Amsterdam: NorthHolland. MR0949228.
- Komsta, L. (2013). Outliers: Tests for outliers, 2011. URL <https://cran.r-project.org/web/packages/outliers>. *R package version 0.14*. [p 374].
- Komsta, L., & Komsta, M. L. (2011). Package ‘outliers’. *Lublin: Medical University of Lublin*.
- Komsta, L., & Novomestky, F. (2015). Moments, cumulants, skewness, kurtosis and related tests. *R package version 0.14*.
- Khoo, M. B. C. and Quah, S. H. (2003). Multivariate Control Chart for Process Dispersion Based on Individual Observations. *Quality Engineering* 15(4), 639-642.
- Kumhof, M. (2018). On the theory of international currency portfolios. *European Economic Review*, 101, 376-396.
- Kupek, E. (2002). Bias and heteroscedastic memory error in self-reported health behavior: an investigation using covariance structure analysis. *BMC medical research methodology*, 2(1), 14.
- Kupitz, Y. S., & Martini, H. (1994). The Fermat-Torricelli point and isosceles tetrahedra. *Journal of Geometry*, 49(1), 150-162.
- Kurtz, Z. D., Müller, C. L., Miraldi, E. R., Littman, D. R., Blaser, M. J., & Bonneau, R. A. (2015). Sparse and compositionally robust inference of microbial ecological networks. *PLoS computational biology*, 11(5), e1004226.
- Lakens, D. (2013). Calculating and reporting effect sizes to facilitate cumulative science: a practical primer for t-tests and ANOVAs. *Frontiers in psychology*, 4, 863.
- Lam, C., & Fan, J. (2009). “Sparsistency and rates of convergence in large covariance matrix Estimation”, *The Annals of Statistics* 2009, Vol. 37, No. 6B, 4254–4278 DOI: 10.1214/09-AOS720 © Institute of Mathematical Statistics.

- Lan, W., Luo, R., Tsai, C. L., Wang, H., & Yang, Y. (2015). Testing the diagonality of a large covariance matrix in a regression setting. *Journal of Business & Economic Statistics*, 33(1), 76-86.
- Leys, C., Klein, O., Dominicy, Y., & Ley, C. (2018). Detecting multivariate outliers: Use a robust variant of the Mahalanobis distance. *Journal of Experimental Social Psychology*, 74, 150-156.
- Leys, C., Ley, C., Klein, O., Bernard, P., & Licata, L. (2013). Detecting outliers: Do not use standard deviation around the mean, use absolute deviation around the median. *Journal of Experimental Social Psychology*, 49(4), 764-766.
- Liu, J., & Calhoun, V. D. (2014). A review of multivariate analyses in imaging genetics. *Frontiers in neuroinformatics*, 8, 29.
- Lix, L. M., Keselman, J. C., & Keselman, H. J. (1996). Consequences of assumption violations revisited: A quantitative review of alternatives to the one-way analysis of variance F test. *Review of educational research*, 66(4), 579-619.
- Longstaff, F. A. (2010). The subprime credit crisis and contagion in financial markets. *Journal of financial economics*, 97(3), 436-450.
- Lopuhaa, H. & Rousseeuw, P. (1991). Breakdown points of affine equivariant estimators of multivariate location and covariance matrices. *The Annals of Statistics* 19, 229–248.
- Lopuhaa, H. P. and Rousseeuw, P. J. (1991). Breakdown points of affine equivariant estimators of multivariate location and covariance matrices. *Ann. Statist.* 19 229-248.
- Loretan, M., & Phillips, P. C. (1994). Testing the covariance stationary of heavy-tailed time series: An overview of the theory with applications to several financial datasets. *Journal of empirical finance*, 1(2), 211-248.
- Machado, M. A., & Costa, A. F. (2012). Control charts with supplementary runs rules for monitoring bivariate processes. *International Journal of Advanced Manufacturing Technology*, 45, 772-785.
- Mahalanobis, P. C. (1936). On the generalised distance in statistics. *Proceedings of the National Institute of Sciences of India*, 1936, 49-55.
- Mantegna, R. N., & Stanley, H. E. (2000). *Introduction to econophysics: correlations and complexity in finance*. Cambridge university press.
- Marascuilo, L. A., & Serlin, R. C. (1988). *Statistical Methods for the Social and Behavioral Sciences*. New York: W. H. Freeman.
- Mardia, K., V., Kent, J., T., & Bibby, J., M., (1979). Multivariate analysis. Academic press Inc. London Ltd. *Subsidiary of Harcourt brace Jovanovich, publishers*. United States edition. New York: Academic Press.

- Marek Klimczak, K. (2008). Corporate hedging and risk management theory: evidence from polish listed companies. *The Journal of Risk Finance*, 9(1), 20-39.
- Ma, J., Hegde, V. I., Munira, K., Xie, Y., Keshavarz, S., Mildebrath, D. T., & Butler, W. H. (2017). Computational investigation of half-Heusler compounds for spintronics applications. *Physical Review B*, 95(2), 024411.
- Ma, Y., Wu, H., Wang, L., Huang, B., Ranjan, R., Zomaya, A., & Jie, W. (2015). Remote sensing big data computing: Challenges and opportunities. *Future Generation Computer Systems*, 51, 47-60.
- Maronna, R. A., Martin, R. D., Yohai, V. J., & Salibián-Barrera, M. (2018). *Robust Statistics: Theory and Methods (with R)*. Wiley.
- Maronna, R., Martin, D., & Yohai, V. (2006). *Robust Statistics: Theory and Methods*, John Wiley & Sons, New York.,
- Maurer, W., Jones, B., & Chen, Y. (2018). Controlling the type I error rate in two-stage sequential adaptive designs when testing for average bioequivalence. *Statistics in medicine*, 37(10), 1587-1607.
- Martin, N. G., & Eaves, L. J. (1977). The genetical analysis of covariance structure. *Heredity*, 38(1), 79-95.
- Mason, R. L., Chou, Y. M., & Young, J. C. (2009). Monitoring variation in a multivariate process when the dimension is large relative to the sample size. *Communications in Statistics—Theory and Methods*, 38(6), 939-951.
- Mathematical Statistics, 1887-1896.
- Matrix Based on Modified Cholesky Decomposition. *arXiv preprint arXiv:1801.00380*.
- Melvin, M., & Taylor, M. P. (2009). The crisis in the foreign exchange market. *Journal of International Money and Finance*, 28(8), 1317-1330.
- Minsker, S. (2015). Geometric median and robust estimation in Banach spaces. *Bernoulli*, 21(4), 2308-2335.
- Mitra, R., & Zhang, C. H. (2016). The benefit of group sparsity in group inference with de-biased scaled group Lasso. *Electronic Journal of Statistics*, 10(2), 1829-1873.
- Mittelhammer, R. C., & Mittelhammer, R. C. (1996). *Mathematical statistics for economics and business* (Vol. 78). New York, NY, USA: Springer.
- Mizuno, T., H. Takayasu, H. & Takayasu, M. (2006). Correlation networks among currencies. *Physica A*, 364: 336-342.
- Mondani, H. (2018). The underlying geometry of organizational dynamics: similarity-based social space and labor flow network communities. *Computational and Mathematical Organization Theory*, 24(3), 378-400.

- Monotgomery, D. C. (2005). *Introduction to Statistical Quality Control*. (5th ed.). New York: John Wiley & Sons, Inc.
- Monotgomery, D. C. (2001). *Introduction to Statistical Quality Control*. (4th ed.). New York: John Wiley & Sons, Inc.
- Muthukrishnan, R., & Radha, M., (2010). “M-estimator in regression models”, *Journal of Mathematics Research*, 2(4), 23-27.
- Nickisch, S., Nockemann, C., Tillack, G. R., Murphy, J., & Sturges, D. (1997). Alternative approaches to the validation of nondestructive testing methods. In *Review of Progress in Quantitative Nondestructive Evaluation* (pp. 2037-2044). Springer US.
- Nordstokke, D. W., & Colp, S. M. (2014). Investigating the robustness of the non parametric Levene test with more than two groups. *Psicológica*, 35(2).
- Olusegun, A., & Sharif, S. (2016). Research Article Covariance Stability Test for Exploring the Impact of Subprime Financial Crisis on the Forex. *Res. J. Appl. Sci. Eng. Technol.*, 12(7): 696-699.
- Ortner, T., Filzmoser, P., Zaharieva, M., Brodinova, S., & Breiteneder, C. (2017). Local projections for high-dimensional outlier detection. *arXiv preprint arXiv:1708.01550*.
- Osborne, J. W., & Overbay, A. (2004). The power of outliers (and why researchers should always check for them). *Practical assessment, research & evaluation*, 9(6), 1-12.
- Osborne, J. W. (2013). *Best practices in data cleaning: A complete guide to everything you need to do before and after collecting your data*. Sage.
- Paltoglou, G., & Thelwall, M. (2017). Sensing Social Media: A Range of Approaches for Sentiment Analysis. In *Cyberemotions* (pp. 97-117). Springer International Publishing.
- Pamukçu, E., Bozdoğan, H., & Çalık, S. (2015). A novel hybrid dimension reduction technique for undersized high dimensional gene expression data sets using information complexity criterion for cancer classification. *Computational and mathematical methods in medicine*.
- Pang, G., Cao, L., Chen, L., Lian, D., & Liu, H. (2018). Sparse Modeling-based Sequential Ensemble Learning for Effective Outlier Detection in High-dimensional Numeric Data. AAAI. *Researchgate.net*.
- Pati, D., Bhattacharya, A., Pillai, N. S., & Dunson, D. (2014). Posterior contraction in sparse Bayesian factor models for massive covariance matrices. *The Annals of Statistics*, 42(3), 1102-1130.

- Pedhazur, E. J. (1997). *Multiple regression in behavioral research* (3rd ed.). Ft. Worth, TX: Harcourt Brace.
- Peña, D., & Prieto, F. J. (2001). Multivariate outlier detection and robust covariance matrix estimation. *Technometrics*, 43(3), 286-310.
- Pereira-Maxwell, F. (2018). *Medical Statistics: An AZ Companion*. CRC Press.
- Petoukhov, S. V. (2016). The system-resonance approach in modeling genetic structures. *Biosystems*, 139, 1-11.
- Pourahmadi, M. (2011). Covariance estimation: The GLM and regularization perspectives. *Statistical Science*, 369-387.
- Pourahmadi, M. (2013). *High-dimensional covariance estimation: with high-dimensional data* (Vol. 882). John Wiley & Sons.
- Powers, S., Qian, J., Jung, K., Schuler, A., Shah, N. H., Hastie, T., & Tibshirani, R. (2018). Some methods for heterogeneous treatment effect estimation in high dimensions. *Statistics in medicine*, 37(11), 1767-1787.
- Piet, E. (1996). The stability of the covariances of international property share returns. *Journal of Real Estate Research*, 11(2), 149-158.
- Qiu, Y., & Chen, S. X. (2014). Band Width Selection for High Dimensional Covariance Matrix Estimation. *Journal of the American Statistical Association*.
- Ramanathan, V., & Bangali, S. S. (2016). Big and high dimensional data in the digital world: statistical, computational and privacy challenges. *International Journal of Automation and Logistics*, 2(4), 369-387.
- Rencher, A. C. (2003). *Methods of multivariate analysis* (Vol. 492). John Wiley & Sons. Requirements for the degree of Doctor of Philosophy in Statistics at Massey University, Albany, New Zealand (Doctoral dissertation, the author).
- Reynolds, M. R. & Gho, G. Y. (2006). Multivariate control charts for monitoring the mean vector and covariance matrix. *Journal of Quality Technology*, 38, 230-253.
- Rigollet, P., & Tsybakov, A. B. (2012). "Sparse estimation by exponential weighting", *Statistical Science*, 558-575.
- Ro, K., Zou, C., Wang, Z., & Yin, G. (2015). Outlier detection for high-dimensional data. *Biometrika*, 102(3), 589-599.
- Rothman, A. J., Levina, E., & Zhu, J. (2009). Generalized thresholding of large covariance matrices. *Journal of the American Statistical Association*, 104(485), 177-186.

- Rothman, A. J., Levina, E., & Zhu, J. (2010). A new approach to Cholesky-based covariance regularization in high dimensions. *Biometrika*, 97(3), 539-550.
- Rousseeuw, P. J., & Hubert, M. (2018). Anomaly detection by robust statistics. *Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery*, 8(2).
- Rousseeuw, P. J., & van Driessen, K. V. (1999). A fast algorithm for the minimum covariance determinant estimator. *Technometrics*, 41(3), 212-223.
- Rousseeuw, P. J., & van Zomeren, B. C. (1991). Robust distances: Simulation and cutoff values. *Directions in Robust Statistics and Diagnostics, Part II. Springer-Verlag, New York*.
- Rousseeuw, P. J. (1985). Multivariate estimation with high breakdown point. *Mathematical Statistics and Applications*, 8, 283-297.
- Roy, S., König, A. C., Dvorkin, I., & Kumar, M. (2015). Perfaugur: Robust diagnostics for performance anomalies in cloud services. In *Data Engineering (ICDE), 2015 IEEE 31st International Conference on* (pp. 1167-1178). IEEE.
- Runcie, D. E., & Mukherjee, S. (2013). Dissecting high-dimensional phenotypes with Bayesian sparse factor analysis of genetic covariance matrices. *Genetics*, genetics-113.
- Rydell, S. (2013). The use of extreme value theory and time series analysis to estimate risk measures for extreme events. Master's Thesis in Engineering Physics, Department of Physics, Umea University.
- Sadrzadeh, H., Baskin, L., & Kline, G. (2017). Variables affecting endocrine tests safety culture and worker's behaviour and safety culture: an experience in manufacturing industry. *International Journal of Basic & Applied Sciences*, 12(5), 1-6.
- Sajesh, T. A., & Srinivasan, M. R. (2013). An overview of multiple outliers in multidimensional data. *Sri Lankan Journal of Applied Statistics*, 14(2).
- Salleh, R. M., & Djauhari, M. A. (2011, December). Robust monitoring of process mean vector in female shrouded connector production: An experience in Malaysia. In *Industrial Engineering and Engineering Management (IEEM), 2011 IEEE International Conference on* (pp. 809-813). IEEE.
- Sawyer, W. & Sawyer, J. (1982). An assessment of preschool education options for rural and small village areas. *Journal of Research in Rural Education* 1(1), 7-10. Retrieved from http://jrre.vmhost.psu.edu/wpcontent/uploads/2014/02/1-1_1.pdf.
- Sen, P. K., & Singer, J. M. (2017). *Large Sample Methods in Statistics (1994): An Introduction with Applications*. CRC Press.

- Seixas, N. S., Robins, T. G., & Moulton, L. H. (1988). The use of geometric and arithmetic mean exposures in occupational epidemiology. *American Journal of Industrial Medicine*, 14(4), 465-477.
- Sequeira, J., Tsourdos, A., & Lazarus, S. B. (2011). Robust covariance estimation for data fusion from multiple sensors. *IEEE Transactions on Instrumentation and Measurement*, 60(12), 3833-3844.
- Serinaldi, F., Kilsby, C. G., & Lombardo, F. (2018). Untenable nonstationarity: An assessment of the fitness for purpose of trend tests in hydrology. *Advances in Water Resources*, 111, 132-15
- Sharif, S., Djauhari, M. A., & Djauhari, H. (2012). Correlation network analysis on worker's behavior and safety culture: an experience in manufacturing industry. *International Journal of Basic & Applied Sciences*, 12(5), 1-6.
- Sharif, S., Ismail, S., & Omar, Z. (2017). New Statistical Test for Quality Control in High Dimension Data Set. *International Journal of Applied Engineering Research*, 12(16), 6241-6249.
- Sharif, S., Ruslan, N & Atiany, A.M. (2018). Type I Error Rate Comparison between Classical and modified Box M- statistics. *Journal of Engineering and Applied sciences* 13(5): 1246-1252, ISSN: 1816-949X. Medwell Journals.
- Sharif, S., Yusoff, W. N. S. W., Omar, Z., & Ismail, S. (2014). Computational efficiency of generalized variance and vector variance. In International conference on quantitative sciences and its applications (ICOQSIA 2014): *Proceedings of the 3rd International Conference on Quantitative Sciences and Its Applications* (Vol. 1635, pp. 906-911). AIP Publishing.
- Sheikholeslami, R., Razavi, S., Gupta, H. V., Becker, W., & Haghnegahdar, A. (2018). Global sensitivity analysis for high-dimensional problems: How to objectively group factors and measure robustness and convergence while reducing computational cost. *Environmental Modelling & Software*.
- Silva, A. P. D., Filzmoser, P., & Brito, P. (2017). Outlier detection in interval data. *Advances in Data Analysis and Classification*, 1-38.
- Singh, S. K. (2010). Robust parameter estimation in gauged and ungauged basins. (Mitteilungen / Institut für Wasserbau, Universität Stuttgart: H. 198) Zugl.: Stuttgart, Univ., Diss., 2010 ISBN 978-3-942036-02-3.
- Soloveychik, I., & Wiesel, A. (2014). Non-asymptotic error analysis of Tyler's scatter estimator. *arXiv preprint arXiv, 1401*.
- Soykan, C. U., Sauer, J., Schuetz, J. G., LeBaron, G. S., Dale, K., & Langham, G. M. (2016). Population trends for North American winter birds based on hierarchical models. *Ecosphere*, 7(5), e01351.

- Stephens, M. A. (2017). Tests based on regression and correlation. In *Goodness-of-Fit-Techniques* (pp. 195-234). Routledge.
- Srivastava, M. S. & Du, M. (2008). A test for the mean vector with fewer observations than the dimension. *J. Mult. Anal.* 99, 386–402.
- Steffens, C., Pesavento, M., & Pfetsch, M. (2018). A Compact Formulation for the $\ell_{2,1}$ Mixed-Norm Minimization Problem. *IEEE Transactions on Signal Processing*.
- Susanti, Y., & Pratiwi, H. (2014). M estimation, S estimation, and MM estimation in robust regression. *International Journal of Pure and Applied Mathematics*, 91(3), 349-360.
- Sullivan, J. H., Stoumbos, Z. G., Mason, R. L. and Young, J. C. (2007). Step-Down Analysis for Changes in the Covariance Matrix and Other Parameters. *Journal of Quality Technology*. 39(1), 66-84.
- Tabachnick, B. G., & Fidell, L. S. (1996). *Using multivariate statistics* (3rd ed.). New York: Harper-Collins.
- Tabachnick, B. G., & Fidell, L. S. (2007). Multivariate analysis of variance and covariance. *Using multivariate statistics*, 3, 402-407.
- Tan, Z., Hohage, T., Kalentev, O., Joseph, A. A., Wang, X., Voit, D., & Frahm, J. (2017). An eigenvalue approach for the automatic scaling of unknowns in model-based reconstructions: Application to real-time phase-contrast flow MRI. *NMR in Biomedicine*, 30(12).
- Tang, G. Y. (1998). The intertemporal stability of the covariance and correlation matrices of Hong Kong stock returns. *Applied Financial Economics*, 8(4), 359-365.
- Tang, J. & Gupta, A. K. (1987). On testing homogeneity of variances for Gaussian models. *J. Statist. Comput. Simul.*, 27, 155-173.
- Tang, J., Liu, J., Zhang, M., & Mei, Q. (2016). Visualizing large-scale and high-dimensional data. In *Proceedings of the 25th International Conference on World Wide Web* (pp. 287-297). International World Wide Web Conferences Steering Committee.
- Tang, J. Wang, J., Huang, C., Wang, G. & Wang, X. (2008). Subprime mortgage crisis Detection in U.S. foreign exchange rate market by multifractal analysis. Proceeding of the 9th International Conference for Young Computer Scientists, 978-0-7695-3398-8/08\$25.002008 IEEE, DOI: 10.1109/ICYCS.2008.454.
- Tarr, G., Müller, S., & Weber, N. C. (2015). Robust estimation of precision matrices under cellwise contamination. *Computational Statistics & Data Analysis*, 93, 404-420.

- Thom, M. (2017). *Investigations on two classes of covering problems* (Doctoral dissertation, University of Lethbridge (Canada)).
- Todros, K. (2019). Robust composite binary hypothesis testing via measure-transformed quasi score test. *Signal Processing*, 155, 202-217.
- Tukey, J., W. (1960). A survey of sampling from contaminated distributions. *Contributions to probability and statistics*, 448-485.
- Ullah, I. (2015). Contributions to high-dimensional data analysis: some applications of the regularized covariance matrices: a thesis submitted in partial fulfilment of the requirements for the degree of Doctor of Philosophy in Statistics at Massey University, Albany, New Zealand (Doctoral dissertation, Massey University).V.
- Valero, J. L. B., & Moreno, S. B. (2005). Robust estimation in geodetic networks. *Física de la Tierra*, 17, 7.
- Vanbrabant, L. (2018). *Reduction in sample size by order restrictions* (Doctoral dissertation, Ghent University).
- Vardi, Y., & Zhang, C. H. (2000). The multivariate L1-median and associated data depth. *Proceedings of the National Academy of Sciences*, 97(4), 1423-1426.
- Wagaman, A. S. & Levina, E. (2009). Discovering Sparse Covariance Structures With the Isomap. *Journal of Computational and Graphical Statistics* 18 551-572.
- Wang, G. J., C. Xie, C., Chen Y.J. & Chen, S. (2013). Statistical properties of the foreign exchange network at Different time scales: Evidence from detrended cross-correlation coefficient and minimum spanning tree. *Entropy*, 15: 1643-1662.
- Wang, M. Y., & Zwilling, C. E. (2015). Multivariate computing and robust estimating for outlier and novelty in data and imaging sciences. In *Advances in Bioengineering*. InTech.
- Wang, N., & Raftery, A. E. (2002). Nearest-neighbor variance estimation (NNVE) robust covariance estimation via nearest-neighbor cleaning. *Journal of the American Statistical Association*, 97(460), 994-1019.
- Wang, Y. & Zou, J. (2010). Vast volatility matrix estimation for high-frequency financial data. *Annals of Statistics* 38, 943-978.
- Wegkamp, M., & Zhao, Y., (2013), "Adaptive estimation of the copula correlation matrix for semi parametric elliptical copulas", *Bernoulli* 22(2), 2016, 1184–1226 DOI: 10.3150/14-BEJ690.
- Wierda, S. J. (1994). Multivariate Statistical Process Control- Recent Results and Directions for Future Research. *Statistica Neerlandica*. 46(2), 147-168.

- Woodall, W.H. and Monotgomery, D. C. (1999). Research Issues and Ideas in Statistical Process Control. *Journal of Quality Technology*, 3(4), 376-386.
- Wilcox, R. R. (2003). *Applying contemporary statistical techniques*. Elsevier.
- Wilcox, R. R. (2005). "Introduction to robust estimation and hypothesis testing", (2nded). San Diego, CA: Academic Press.
- Wilkinson, L. (2018). Visualizing Big Data Outliers through Distributed Aggregation. *IEEE transactions on visualization and computer graphics*, 24(1), 256-266.
- Wilks, S. S. (1932). Certain generalizations in the analysis of variance. *Biometrika* 24, 471-94.
- Williamson, B., (2014). "Shrinkage Estimators for High-Dimensional Covariance Matrices", (Doctoral dissertation, Pomona College).
- Wu, W. B., & Pourahmadi, M. (2003). Nonparametric estimation of large covariance matrices of longitudinal data. *Biometrika*, 90(4), 831-844.
- Wu, W. B., & Pourahmadi, M. (2009). Banding sample auto covariance matrices of stationary processes. *Statistica Sinica*, 19(4), 1755.
- Wu, W. W. (2010). Beyond business failure prediction. *Expert systems with applications*, 37(3), 2371-2376.
- Xie, X., Xue, X., & Strickler, H. D. (2018). Generalized linear mixed model for binary outcomes when covariates are subject to measurement errors and detection limits. *Statistics in Medicine*, 37(1), 119-136.
- Xue, L., & Zou, H. (2012). Regularized rank-based estimation of high-dimensional nonparanormal graphical models. *The Annals of Statistics*, 40(5), 2541-2571.
- Xue, L., & Zou, H. (2014). Rank-based tapering estimation of bandable correlation matrices. *Statistica Sinica*, 24, 83-100.
- Yahaya, S. S., Othman, A. R., & Keselman, H. J. (2006). Comparing the "typical score" across independent groups based on different criteria for trimming. *Metodoloski zvezki*, 3(1), 49.
- Yahaya, S.S., Ali, H., & Omar, Z. (2011). An alternative Hotelling T^2 control chart based on minimum vector variance (MVV). *Modern Applied Science*, 5(4), 132-151. doi.10.5539/mas.v5n4p132.
- Yao, G. B., Deng, K. Z., Zhang, L., & Ai, H. B. (2014). Robust perspective invariant quasidense matching across large oblique images. *Journal of Electronic Imaging*, 23(4), 043005-043005.

- Yi, F., & Zou, H. (2013). SURE-tuned tapering estimation of large covariance matrices. *Computational Statistics & Data Analysis*, 58, 339-351.
- Yu, G., Zou, C. & Wang, Z. (2012). Outlier detection in the functional observations with applications to profile monitoring. *Technometrics* 54, 308–18.
- Yu, S., & Yiquan, W. (2018). Subspace clustering based on latent low rank representation with Frobenius norm minimization. *Neurocomputing*, 275, 2479-2489.
- Yuan, K. H., & Bentler, P. M. (2000). Robust mean and covariance structure analysis through iteratively reweighted least squares. *Psychometrika*, 65(1), 43-58.
- Yuan, K. H., & Bentler, P. M. (2001). Effect of outliers on estimators and tests in covariance structure analysis. *The British Journal of Mathematical and Statistical Psychology*, 54, 161-175. doi:10.1348/000711001159366.
- Yuan, K. H., & Bentler, P. M. (2007). Multilevel covariance structure analysis by fitting multiple single-level models. *Sociological methodology*, 37(1), 53-82.
- Yuan, K. H., Chan, W., & Tian, Y. (2015). Expectation-robust algorithm and estimating equations for means and dispersion matrix with missing data. Advance online publication. doi: 10.1007/s10463-014-0498-1.
- Yue, S., Pilon, P. & Cavadias, G. (2002). Power of the Mann-Kendall and Spearman's rho tests for detecting monotonic trends in hydrological series. *J. Hydrol.* 259, 254–271.
- Yusof, Z. M., Abdullah, S., Yahaya, S. S. S., & Othman, A. R. (2011). Type I error rates of Ft statistic with different trimming strategies for two groups case. *Modern Applied Science*, 5(4), 236.
- Yusof, Z. M., Harua, N. H., Syed Yahaya, S. S. & Abdullah, S. (2013). *Modified parametric bootstrap: A robust alternative to classical test*. In: World Conference on Integration of Knowledge: Business, Economy, Management, Cultural, Societal, Theological and Scientific (WCIK 2013), 25 & 26 November 2013, Langkawi, Malaysia.
- Zhang, J. (2017). *Geometric Facility Location Problems on Uncertain Data* (Doctoral dissertation, Utah State University).
- Zhang, X., Shi, X., Sun, Y., & Cheng, L. (2018). Multivariate Regression with Gross Errors on Manifold-valued Data. *IEEE Transactions on Pattern Analysis and Machine Intelligence*.
- Zhong, P. S., Chen, S. X., & Xu, M. (2013). Tests alternative to higher criticism for high-dimensional means under sparsity and column-wise dependence. *The Annals of Statistics*, 41(6), 2820-2851.

- Zuo, L., Slotine, J. J., & Nayfeh, S. A. (2005). Model reaching adaptive control for vibration isolation. *IEEE Transactions on Control Systems Technology*, 13(4), 611-617.
- Zuo, Y. (2006). Robust location and scatter estimators in multivariate analysis. In *Frontiers In Statistics* (pp. 467-490).



APPENDIX A.

The Algorithm of M_{BGMed} -Test and M_{TGMed} -Test by Regularization Methods

Step 1: Data preparation

- i. Prepare the dataset.
- ii. Split the data (into the required number of folds or samples).
- iii. Compute the covariance (for each of the splits data), i.e., target sample and second sample.
- iv. Compute the determinant of covariance matrix to satisfy $p > n$, as zero.

Step 2: Banding and tapering sample covariance matrix

- i. Estimate risk (k) where optimal k is choice of the banding parameter.
- ii. Determine value for α , n and p (are critical in the estimation of k).
- iii. Compute eigenvalues and eigenvectors, where α = minimum eigenvalues from the classes, n = sample size and p = dimension.

Step 3: Cross validation method (use cross validation method to choose optimal k as banding parameter).

- i. Compute the Risk (k) for both Target sample and Second sample
- ii. Choose the minimum of $R(k)$ as the optimal value k .
- iii. Compute and determine the cross validation by the Banding formula.

Step 4: Diagonalised the matrix and test for performance of banded k -diagonal estimator

- i. Estimate k -diagonal matrix (multiply eigenvalues and transpose of eigenvectors).
- ii. Multiply k with diagonal values gives k -diagonal matrix.
- iii. Compute banded covariance matrix with, k , then (k -diagonal estimator or matrix) M-test inversion problem solved and replaced with banded covariance matrix.
- iv. Use banded matrix to compute M-test, M_{GMed} -test, M_B -test, M_T -test, M_{BGMed} -test and M_{TGMed} -test with simulation study.
- v. Compare the performance M-test, M_B -test, M_T -test, M_{GMed} -test, M_{BGMed} -test and M_{TGMed} -test for Type I error rate values and Power of test.
- vi. Test for performance with statistical test, under condition n , p and α , then covariance shift, simulation study and real data application.

The computation process by simulations to carry out the power of test for performance with 5,000 iterations ways are as follows:

APPENDIX B.

The Steps for Power of Test Computation for Performance

Step 1. Generate multivariate random variable: Data I

- i. Mean vector = zero
- ii. Covariance matrices: I_p -identity matrix

Covariance 3×3

when, $p = 3, n = 10$,

Step 2. Generate multivariate random variable: Data II

- i. Mean vector = zero
- ii. Covariance = a size of $p \times p$ with shift (Mason and Young, 2009).

Step 3. Compute M-test $\Rightarrow \begin{cases} S_1 & \text{refer to Data I} \\ S_2 & \text{refer to Data II} \end{cases} M_1$
 where, $S_{pooled} = S_1 \text{ and } S_2$

Step 4. Repeat (10) times (we have 10 M-test values)

Step 5. Compute chi-square value (Cv_n).

Step 6. We compare (3) and (5).

Step 7. To reject /fail to reject the hypothesis. Decision: $M_n > Cv_n$

Table:

Shown results

Results	Chi – Square	Comparison
M_1	Cv_1	Reject / fail to reject
M_2	Cv_2	Reject / fail to reject
$\vdots$	$\vdots$	$\vdots$
M_{10}	Cv_{10}	Reject / fail to reject

APPENDIX C.

The Steps for Computing Parameter k by Cross Validation Method

Step 1: Data Preparation

- i. Prepare the data matrix set of $n \times p$. e.g. Given,

$$f = \text{matrix}(\text{c}(5,10,5,2,6,8,10,12,5,4,10,12,8,6,4,0,5,6,3,2,1,5,2,8), \\ \text{nrow} = 4, \text{by row} = \text{T}).$$

Then, $p > n$ matrix form,

f = display 4×6 matrix

$$\begin{bmatrix} 5 & 10 & 5 & 2 & 6 & 8 \\ 10 & 12 & 5 & 4 & 10 & 12 \\ 8 & 6 & 4 & 0 & 5 & 6 \\ 3 & 2 & 1 & 5 & 2 & 8 \end{bmatrix}$$

Step 2. We split the data matrix into the required number of folds or samples. i.e., two samples of n_1 and n_2 , where,

$$n_1 = n - n/\log(n)$$

$$n_2 = n - n_1$$

Step 3. Then, compute the values of n_1 and n_2 , i.e. $4 - \frac{4}{\log(4)} = 1.11461$ (arbitrarily

2), the number of split equal to 2. The matrix is split into two mat1 and mat2:

$$\text{mat1} = \text{matrix}(\text{c}(5,10,5,2,6,8,3,2,1,5,2,8), \text{nrow} = 2, \text{byrow} = \text{T})$$

i. e. row 1 and row 4 of the original matrix.

$$\text{mat2} = \text{matrix}(\text{c}(10,12,5,4,10,12,8,6,4,0,5,6), \text{nrow} = 2,$$

byrow = T). i. e., row 2 and row 3 of the original matrix.

To display mat 1 and mat 2 as follows:

mat 1 =

$$\begin{bmatrix} 5 & 10 & 5 & 2 & 6 & 8 \\ 3 & 2 & 1 & 5 & 2 & 8 \end{bmatrix}$$

mat 2 =

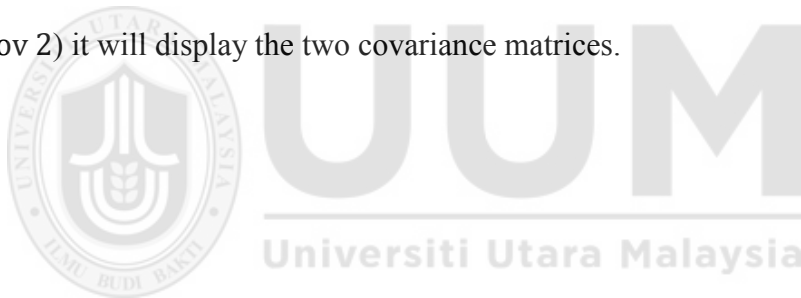
$$\begin{bmatrix} 10 & 12 & 5 & 4 & 10 & 12 \\ 8 & 6 & 4 & 0 & 5 & 6 \end{bmatrix}$$

Step 4. Compute the covariance for mat 1 and mat 2, then, cov 1 and cov 2 respectively as,

cov 1 = cov(mat 1)

cov 2 = cov(mat 2)

These two covariance matrices will be displayed by viewing: View(cov 1) and View(cov 2) it will display the two covariance matrices.



Banding condition for cross validation with 10 iterations.

Iterations	Samples Two Splits	
	n_1	n_2
1	1 & 4	2 & 3
2	4 & 3	1 & 2
3	1 & 3	2 & 4
4	3 & 2	4 & 1
5	3 & 4	2 & 1
6	2 & 4	3 & 1
7	4 & 1	2 & 3
8	2 & 3	1 & 4
9	3 & 1	4 & 2
10	2 & 1	4 & 3

The original matrix

$$\begin{bmatrix} 5 & 10 & 5 & 2 & 6 & 8 \\ 10 & 12 & 5 & 4 & 10 & 12 \\ 8 & 6 & 4 & 0 & 5 & 6 \\ 3 & 2 & 1 & 5 & 2 & 8 \end{bmatrix}$$

$\mathcal{S} = (S_{ij})$, any integer $k, 0 \leq k < p$, from the example since $p = 6$,

Possible values of $k = 0, 1, 2, 3, 4$ and 5 . i.e., select the optimal k , to be determined by cross validation methods.

We calculate the matrices difference by:

$\text{mat. diff} < -\text{banding}(\text{cov}(\text{mat1}), 0) - \text{cov}(\text{mat2})$, for all the values of

$k = 0, 1, 2, 3, 4$ and 5 .

To obtain the risk, the risk is the minimum value of k . $\text{svd}(\text{mat. diff})$. Alternatively,

$\text{O. norm2}(\text{mat. diff})$ i.e. operator norm will compute the (v_n) values.

Practical Example:

k	Table showing the repeated k tuning parameters										Mean values
	v_1	v_2	v_3	v_4	v_5	v_6	v_7	v_8	v_9	v_{10}	
0	2436.52	648.39	1781.10	1508.76	1955.34	648.39	14074.79	1508.76	14074.79	1655.46	4029.22
1	2529.15	794.16	1869.44	1586.87	5921.41	794.16	14699.96	1586.87	14699.96	1513.35	4599.52
2	3003.82	929.00	2366.37	1576.34	6715.20	929.00	14872.74	1576.34	14872.74	1513.35	4835.48
3	2464.16	1442.76	1816.51	1695.66	10240.43	1442.76	14475.09	1695.66	14475.09	1513.35	5126.14
4	2412.53	1622.13	2331.59	1540.13	13836.09	1622.13	14361.05	1540.13	14361.05	1513.35	5514.01
5	2412.53	1725.44	2412.53	1725.44	14525.14	1725.44	14525.14	1725.44	14525.14	1513.35	5681.55

The above table is computed applying the norm (operator norm), using the singular value decomposition method (SVD) of the matrices difference. Alternative, we compute the $O.\text{norm2}(\text{mat.diff})$ in estimating all the v_n values. Then, we add all the v_n along each k and take the averages, take the minimum of them all. In this example, the minimum risk is 4029.22 at $k = 0$. Then, the largest value of k is 2436.51 along, $k = 0$. The manually computed banded k -parameter is as followings

APPENDIX D.

Computation on Banded parameter k

Given,

2×2 Matrix Problem

$d = \text{matrix}(c(1, 2, 3, 4), \text{nrow} = 2, \text{byrow} = T)$ ### when $k = 0$ an arbitrary
between. (0,1)

$$d = \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix}$$

$\text{banding}(d, 0)$

$$\begin{bmatrix} 1 & 0 \\ 0 & 4 \end{bmatrix}$$

$\text{cova} = \text{cov}(d)$

cova

$$\begin{bmatrix} 2 & 2 \\ 2 & 2 \end{bmatrix}$$

$\text{banding}(\text{cova}, 0)$

$$\begin{bmatrix} 2 & 0 \\ 0 & 2 \end{bmatrix}$$

APPENDIX E.

Banded Method Mathematical formula Validation

Banding Formula

If $\text{abs}(i - j) > 0$

Then, $\text{sigma}(i, j) = 0$

$k = 0$, when $i = 1$ and $j = 1$

$\text{Sigma}(1, 1) = 2$

$\text{abs}(1-1) > k$

$\text{abs}(1 - 1) = 0$

since zero is not $> k$

Then, $\text{sigma}(1, 1) = 2$

Next element is $\text{sigma}(1, 2) = 2$

$\text{abs}(1 - 2) = 1$

$> k$ then 0

$\text{Sigma}(2, 1) = 2$

$\text{abs}(2 - 1) = 1$

$> k = 0$

$\text{Sigma}(2, 2) = 2$

$\text{abs}(2 - 2) = 0$

$< k = 2$

APPENDIX F.

Simple Banding Sample Covariance Matrix with k -parameter

(R software Code)

Data used (n_1 , n_2 and overall)

Install. Packages (“FincovRegularization”)

library(“FincovRegularization”)

h=read.csv (file.choose (), header =TRUE, sep = “,”)

View (h)

tuningparameter = banding.cv (h, n.cv =10, norm =”o”, seed =142857)

tuningparameter

display result [5] ## k=5

covar = cov (h)

View (cov)

det (covar)

display result [0] ## det=0

f = banding (covar, k=5) #### k =optimal banded tuning parameter

View (f)

det (f)

display result [1.510214e-169] #### banded determinant

write. csv (f, “n1 banded covariance mat.csv”)

J = solve (f) #### solve the inverse of the matrix

View (j)

write. csv (j,”inverse of n1 banded covariance mat.csv”)

APPENDIX G.

Simple Tapering Sample Covariance Matrix with l -parameter

(R software Code)

Data used (n_1 , n_2 and overall)

Install. Packages (“FincovRegularization”)

library(“FincovRegularization”)

h=read.csv (file.choose (), header =TRUE, sep = “,”)

View (h)

tuningparameter = tapering.cv (h, n.cv =10, norm =”o”, seed =142857)

tuningparameter

display result [5] ## $l = 5$

covar = cov (h)

View (cov)

det (covar)

display result [0] ## det=0

g= tapering (covar, $l = 5$, $h = 1/2$) #### l =optimal tapered tuning parameter

View (g)

det (g)

display result [-2.635138e-165] #### tapered determinant

write. csv (g, “n1 tapered covariance mat.csv”)

J = solve (g) #### solve the inverse of the matrix

View (g)

write. csv (t,”inverse of n1 tapered covariance mat.csv”)

APPENDIX H.

Mean, Median and Geometric Median Covariance

Next, we show the computations of the covariance matrix:

Let X be the data matrix

$$X = \begin{bmatrix} 42 & 4 \\ 52 & 5 \\ 48 & 4 \\ 58 & 3 \end{bmatrix}$$

The sample mean is,

$$\bar{X} = \begin{bmatrix} \bar{X}_1 \\ \bar{X}_2 \end{bmatrix} = \begin{bmatrix} 50 \\ 4 \end{bmatrix} \text{ and,}$$

The sample variance-covariance is,

$$S^2_K = S_{kk} = \frac{1}{n} \sum_{j=1}^n (x_{jk} - \bar{x}_k)^2, \text{ where } k = 1, 2, \dots, p$$

The sample variance is,

$$S_{ik} = \frac{1}{n} \sum_{j=1}^n (x_{ji} - \bar{x}_i) (x_{jk} - \bar{x}_k), i = 1, 2, \dots, p, \text{ where } k = 1, 2, \dots, p, \text{ can be written}$$

in the matrix form as,

$$\begin{bmatrix} S_{11} & S_{12} & \cdots & S_{1p} \\ S_{21} & S_{22} & \cdots & S_{2p} \\ \vdots & \vdots & \ddots & \vdots \\ S_{p1} & S_{p2} & \cdots & S_{pp} \end{bmatrix} = \begin{bmatrix} 34 & -1.5 \\ 1.5 & 0.5 \end{bmatrix}$$

The following we show the computation of the median covariance matrix

Let Y be the data matrix

$$Y = \begin{bmatrix} 42 & 4 \\ 52 & 5 \\ 48 & 4 \\ 58 & 3 \end{bmatrix}$$

$$Y_{11} = 42, \quad Y_{21} = 52, \quad Y_{31} = 48, \quad Y_{41} = 58$$

$$Y_{12} = 4, \quad Y_{22} = 5, \quad Y_{32} = 4, \quad Y_{42} = 3$$



Appendix I

The Sample Median Covariance

$S^2_l = S_{ll} = \frac{1}{n} \sum_{m=1}^n (y_{ml} - \check{y}_l)^2$, where $l = 1, 2, \dots, p$ and $\check{y} = \text{median}$ (i.e. the middle of a sorted list of numbers).

Therefore,

$$S_{jl} = \frac{1}{n} \sum_{m=1}^n (y_{mj} - \check{y}_j) (y_{ml} - \check{y}_l), j = 1, 2 \dots p, l = 1, 2, \dots p$$

$$S_{12} = \frac{1}{n} \sum_{m=1}^n (y_{m1} - \check{y}_1) (y_{m2} - \check{y}_2)$$

sample medians $\check{Y} = \begin{bmatrix} \check{y}_1 \\ \check{y}_1 \\ \vdots \\ \check{y}_1 \end{bmatrix}$

sample variances and covariances

$$S = \begin{bmatrix} S_{11} & S_{12} & \dots & S_{1p} \\ S_{21} & S_{22} & \dots & S_{2p} \\ \vdots & \vdots & \ddots & \vdots \\ S_{p1} & S_{p2} & \dots & S_{pp} \end{bmatrix}$$

$$\check{y}_1 = 52 + 48 = 100$$

$$\frac{100}{2} = 50$$

$$\check{y}_2 = 5 + 4 = 9$$

$$\frac{9}{2} = 4.5$$

$$\check{Y} = \begin{bmatrix} \check{y}_1 \\ \check{y}_2 \end{bmatrix} = \begin{bmatrix} 50 \\ 4.5 \end{bmatrix}$$

The sample variance and covariance

$$S_{11} = \frac{1}{4} \sum_{m=1}^4 (y_{m1} - \check{y}_1)^2$$

$$\frac{1}{4}((42 - 50)^2 + (52 - 50)^2 + (48 - 50)^2 + (58 - 50)^2) = 34$$

$$S_{22} = \frac{1}{4} \sum_{m=1}^4 (y_{m2} - \check{y}_2)^2$$

$$\frac{1}{4}((4 - 4.5)^2 + (5 - 4.5)^2 + (4 - 4.5)^2 + (3 - 4.5)^2) = 0.75$$

$$S_{11,12} = \frac{1}{4} \sum_{m=1}^4 (x_{m1} - \bar{x}_1)(x_{m2} - \bar{x}_2)$$

$$\begin{aligned} \frac{1}{4}((42 - 50)(4 - 4.5) + (52 - 50)(5 - 4.5) + (48 - 50)(4 - 4.5) \\ + (58 - 50)(3 - 4.5)) = 1.5 \end{aligned}$$

Then compare,

$$S_{21} = S_{12}$$

and

$$S_n = \begin{bmatrix} 34 & 1.5 \\ 1.5 & 0.75 \end{bmatrix}$$

Appendix J

The GMed will be computed based on its properties which is one of the possible extensions of a univariate median to higher dimension (Minsker, 2015). Moreover, Kemperman, 1987, defined geometric median as a uniquely random variable that is not concentrated on a straight line, with the Hilbert space definitions. The following we show the computation of geometric median covariance

Given the following:

Variable 1(dollar Sales)	42	52	48	58
Variable 2 (No of Books)	4	5	4	3

Let Z be the data matrix

$$Z = \begin{bmatrix} 42 & 4 \\ 52 & 5 \\ 48 & 4 \\ 58 & 3 \end{bmatrix}$$

The geometric median vector consist of the geometric median of each variable and the variance-covariance matrix consist of the variances of the variables along the main diagonal and the covariances between each pair of variables in the other matrix positions (Efteliogu, 2015).

The GMed is computed from: Given data matrix Z with X and Y ,

$M(X) = \arg \min_{x \in \mathbb{R}^d} \sum_{i=1}^m \| m - x_i \|$ for points $x_1, x_2, \dots, x_i \in \mathbb{R}^d$ the median of set

X is defined as a point m so, the formula for computing the covariance of the variables is:

$M(Y) = \arg \min_{y \in \mathbb{R}^d} \sum_{i=1}^m \|m - y_i\|$ for points $y_1, y_2, \dots, y_i \in \mathbb{R}$ is $\mathbb{R}^d$ the median of set

Y is defined as a point m .

The formula for computing the covariance of the variables is:

$$\text{geometric median sample} = \left[\frac{\sum_{i=1}^n \frac{(X_i - m(x))(Y_i - m(y))}{n-1}}{\sum_{i=1}^n \frac{(Y_i - m(y))(X_i - m(x))}{n-1}} \right]$$

Let X be the first data matrix:

$$X = \sum_{i=1}^n \frac{(X_i - m(x))(Y_p - m(y))}{n-1}$$

Let $X_i = X_{11} = 42, X_{21} = 52, X_{31} = 48, X_{41} = 58$

Geometric Median = $\arg \min_{x \in \mathbb{R}^d} \sum_{i=1}^m \|X_i - m(x)\|$, where, $m(x)$ = median (i.e. the middle of a sorted list of numbers of variable, X).

$$m(x) = 48 + 52 = \frac{100}{2} = 50$$

Then,

$$\begin{aligned} \sum_{i=1}^m \|X_i - m(x)\| &= \|42 - 50\| + \|52 - 50\| + \|48 - 50\| + \|58 - 50\| = \\ &= 8 + 2 + 2 + 8 = 20 \end{aligned}$$

Let $\mathbf{Y}$ be the second data matrix,

$\arg \min_{x \in \mathbb{R}^d} \sum_{i=1}^m \|Y_i - m(y)\|$, where, $m(y)$ = median (i.e. the middle of a sorted list

of numbers of variable $\mathbf{Y}$, then,

$$\sum_{i=1}^n \frac{(Y_i - m(y))(X_i - m(x))}{n-1}$$

Let $Y_p = Y_{11} = 4, Y_{21} = 5, Y_{31} = 4, Y_{41} = 3$

$$m(y) = 4 + 4 = \frac{8}{2} = 4$$

$$\sum_{i=1}^m \|Y_p - m(y)\| = \|4 - 4\| + \|5 - 4\| + \|4 - 4\| + \|3 - 4\| = 0 + 1 + 0 + 1 = 2$$

sample geometric median covariance

$$S^2_X = S_{XY} = \sum_{i=1}^n \frac{((X_i - m(x))(Y_p - m(y)))^2}{n-1},$$

$$S^2_Y = S_{YX} = \sum_{i=1}^n \frac{((Y_p - m(y))(X_i - m(x)))^2}{n-1}, \text{ where } x = 1, 2, \dots, i \text{ and } y = 1, 2, \dots, p$$

Therefore,

$$\text{geometric median sample} = \left[\frac{\sum_{i=1}^n \frac{((X_i - m(x))(Y_p - m(y)))^2}{n-1}}{\sum_{i=1}^n \frac{((Y_p - m(y))(X_i - m(x)))^2}{n-1}} \right] = \left[\frac{20}{2} \right]$$

geometric median sample variances and covariances

$$\mathbf{S} = \begin{pmatrix} s_{11} & s_{12} & \cdots & s_{1p} \\ s_{21} & s_{22} & \cdots & s_{2p} \\ \vdots & \vdots & \ddots & \vdots \\ s_{p1} & s_{p2} & \cdots & s_{pp} \end{pmatrix}$$

$$\begin{aligned} \sum_{i=1}^n \frac{((X_i - m(x))(Y_p - m(y)))^2}{n-1} \\ = \frac{1}{4-1} [(42-20)^2 + (52-20)^2 + (48-20)^2(58-20)^2] \\ = \frac{3736}{3} = 1245.3 \end{aligned}$$

$$\begin{aligned} \sum_{i=1}^n \frac{((Y_p - m(y))(X_i - m(x)))^2}{n-1} \\ = \frac{1}{4-1} [(4-2)^2 + (5-2)^2 + (4-2)^2 + (3-2)^2] = \frac{18}{3} = 6 \end{aligned}$$

$$S_{XX,YY} = \left[\sum_{i=1}^n \frac{((X_i - m(x))(Y_p - m(y)))^2 + ((Y_p - m(y))(X_i - m(x)))^2}{n-1} \right]$$

$$\begin{aligned} \frac{1}{4-1} [((42-20)(4-2)) + ((52-20)(5-2)) + ((48-20)(4-2)) \\ + ((58-20)(3-2))] = \frac{234}{3} = 78 \end{aligned}$$

Then compare,

$$S_{YX} = S_{XY}$$

and

$$S_n = \begin{bmatrix} 1245.3 & 78 \\ 78 & 6 \end{bmatrix}$$

Appendix K.

Table K.1

Type I error rates for variable, $p = 3$ and $n = 2$

ε	κ	μ	M_B	M_T	M_{BGMed}	M_{TGMed}
0	1	0	0.055	0.051	0.047	0.047
0.05	2	3	0.051	0.050	0.049	0.049
		5	0.053	0.051	0.051	0.049
	5	3	0.052	0.053	0.048	0.051
		5	0.049	0.052	0.040	0.047
	10	3	0.050	0.051	0.044	0.041
		5	0.053	0.050	0.050	0.044
0.1	2	3	0.047	0.050	0.042	0.047
		5	0.048	0.047	0.046	0.045
	5	3	0.053	0.051	0.045	0.045
		5	0.051	0.053	0.049	0.046
	10	3	0.061	0.064	0.046	0.048
		5	0.054	0.053	0.045	0.046
0.2	2	3	0.053	0.046	0.049	0.046
		5	0.050	0.053	0.044	0.051
	5	3	0.054	0.052	0.047	0.047
		5	0.059	0.054	0.047	0.046
	10	3	0.066	0.065	0.045	0.047
		5	0.072	0.064	0.049	0.045
0.5	2	3	0.094	0.085	0.053	0.057
		5	0.067	0.080	0.064	0.067
	5	3	0.078	0.070	0.059	0.093
		5	0.081	0.076	0.073	0.082
	10	3	0.088	0.076	0.080	0.098
		5	0.132	0.127	0.079	0.077
No. of Shaded region			19	19	23	21

Shaded region indicate Type I error within [0.025 – 0.075]

Table K.2

Type I error rates for variable, $p = 3$ and $n = 5$

ε	κ	μ	M	M_{GMed}	M_B	M_T	M_{BGMed}	M_{TGMed}
0	1	0	0.044	0.052	0.063	0.044	0.052	0.051
0.05	2	3	0.050	0.055	0.056	0.046	0.051	0.043
		5	0.049	0.055	0.061	0.042	0.047	0.048
	5	3	0.056	0.054	0.071	0.060	0.052	0.046
		5	0.048	0.050	0.065	0.058	0.046	0.050
	10	3	0.056	0.055	0.100	0.090	0.054	0.051
		5	0.061	0.055	0.094	0.085	0.053	0.057
0.1	2	3	0.045	0.047	0.057	0.050	0.044	0.047
		5	0.045	0.047	0.058	0.050	0.051	0.048
	5	3	0.059	0.053	0.081	0.076	0.053	0.052
		5	0.054	0.060	0.088	0.076	0.050	0.048
	10	3	0.071	0.058	0.132	0.127	0.059	0.067
		5	0.066	0.062	0.147	0.140	0.064	0.061
0.2	2	3	0.053	0.053	0.060	0.049	0.042	0.045
		5	0.048	0.048	0.067	0.056	0.048	0.047
	5	3	0.075	0.058	0.113	0.109	0.057	0.050
		5	0.065	0.056	0.121	0.107	0.057	0.061
	10	3	0.088	0.070	0.231	0.209	0.081	0.078
		5	0.087	0.066	0.235	0.214	0.085	0.077
0.5	2	3	0.127	0.078	0.052	0.057	0.063	0.072
		5	0.112	0.080	0.054	0.058	0.063	0.065
	5	3	0.081	0.067	0.152	0.157	0.074	0.082
		5	0.074	0.077	0.154	0.158	0.075	0.095
	10	3	0.088	0.091	0.122	0.125	0.094	0.081
		5	0.092	1.000	0.122	0.132	0.095	0.087
No. of Shaded region			17	20	11	11	20	19

Shaded region indicate Type I error within [0.025 – 0.075]

Table K.3

Type I error rates for variable, $p = 3$ and $n = 10$

ε	κ	μ	M	M_{GMed}	M_B	M_T	M_{BGMed}	M_{TGMed}
0	1	0	0.046	0.055	0.045	0.050	0.046	0.041
0.05	2	3	0.050	0.051	0.050	0.055	0.045	0.046
		5	0.054	0.056	0.046	0.052	0.044	0.041
	5	3	0.068	0.057	0.078	0.088	0.044	0.051
		5	0.079	0.063	0.086	0.086	0.051	0.051
	10	3	0.127	0.073	0.146	0.151	0.063	0.061
		5	0.112	0.070	0.147	0.153	0.063	0.065
0.1	2	3	0.061	0.057	0.052	0.057	0.044	0.042
		5	0.054	0.057	0.054	0.058	0.041	0.045
	5	3	0.088	0.071	0.122	0.125	0.054	0.061
		5	0.092	0.071	0.122	0.132	0.055	0.060
	10	3	0.192	0.090	0.256	0.275	0.089	0.078
		5	0.192	0.098	0.276	0.259	0.084	0.080
0.2	2	3	0.064	0.057	0.064	0.061	0.042	0.039
		5	0.061	0.048	0.066	0.070	0.040	0.042
	5	3	0.143	0.085	0.221	0.229	0.073	0.076
		5	0.158	0.083	0.220	0.228	0.072	0.075
	10	3	0.340	0.137	0.469	0.475	0.132	0.136
		5	0.339	0.135	0.468	0.478	0.142	0.128
0.5	2	3	0.090	0.063	0.098	0.097	0.049	0.057
		5	0.162	0.074	0.210	0.202	0.064	0.070
	5	3	0.172	0.081	0.203	0.100	0.060	0.065
		5	0.079	0.084	0.060	0.056	0.052	0.050
	10	3	0.085	0.077	0.058	0.053	0.061	0.046
		5	0.129	0.099	0.173	0.163	0.058	0.069
No. of Shaded region			8	15	9	9	21	20
Shaded region indicate Type I error within [0.025 – 0.075]								

Shaded region indicate Type I error within [0.025 – 0.075]

Table K.4

Type I error rates for variable, $p = 3$ and $n = 15$

ϵ	κ	μ	M	M_{GMed}	M_B	M_T	M_{BGMed}	M_{TGMed}
0	1	0	0.050	0.056	0.056	0.046	0.062	0.047
0.05	2	3	0.049	0.052	0.055	0.048	0.048	0.045
		5	0.050	0.059	0.054	0.050	0.049	0.047
	5	3	0.092	0.059	0.100	0.100	0.062	0.055
		5	0.090	0.063	0.098	0.097	0.057	0.049
	10	3	0.162	0.074	0.210	0.202	0.076	0.064
		5	0.172	0.081	0.203	0.100	0.065	0.060
0.1	2	3	0.059	0.054	0.060	0.056	0.050	0.042
		5	0.055	0.057	0.058	0.053	0.046	0.041
	5	3	0.129	0.069	0.173	0.163	0.069	0.058
		5	0.126	0.072	0.173	0.168	0.067	0.051
	10	3	0.298	0.091	0.368	0.357	0.104	0.079
		5	0.290	0.098	0.375	0.366	0.086	0.088
0.2	2	3	0.063	0.049	0.084	0.067	0.046	0.031
		5	0.060	0.055	0.080	0.072	0.041	0.034
	5	3	0.237	0.087	0.339	0.312	0.091	0.084
		5	0.230	0.089	0.331	0.327	0.087	0.082
	10	3	0.524	0.145	0.645	0.630	0.151	0.135
		5	0.521	0.154	0.641	0.629	0.156	0.148
0.5	2	3	0.104	0.061	0.062	0.063	0.056	0.147
		5	0.225	0.070	0.061	0.074	0.062	0.081
	5	3	0.229	0.071	0.260	0.265	0.061	0.079
		5	0.163	0.088	0.112	0.163	0.087	0.136
	10	3	0.094	0.143	0.261	0.274	0.077	0.094
		5	0.173	0.154	0.221	0.225	0.093	0.152
No. of Shaded region			7	15	7	9	15	13

Shaded region indicate Type I error within [0.025 – 0.075]

Table K.5

Type I error rates for variable, $p = 3$ and $n = 20$

ϵ	κ	μ	M	M_{GMed}	M_B	M_T	M_{BGMed}	M_{TGMed}
0	1	0	0.053	0.045	0.051	0.055	0.052	0.048
0.05	2	3	0.061	0.040	0.057	0.065	0.044	0.038
		5	0.051	0.044	0.058	0.060	0.041	0.041
	5	3	0.107	0.058	0.123	0.139	0.048	0.057
		5	0.104	0.061	0.112	0.128	0.056	0.047
	10	3	0.225	0.070	0.264	0.268	0.062	0.051
		5	0.229	0.071	0.260	0.265	0.061	0.059
0.1	2	3	0.063	0.048	0.062	0.063	0.037	0.036
		5	0.064	0.043	0.061	0.074	0.037	0.034
	5	3	0.173	0.054	0.221	0.225	0.053	0.052
		5	0.164	0.055	0.210	0.230	0.059	0.055
	10	3	0.385	0.092	0.458	0.469	0.085	0.085
		5	0.399	0.090	0.457	0.473	0.076	0.075
0.2	2	3	0.080	0.033	0.094	0.093	0.031	0.029
		5	0.073	0.037	0.091	0.097	0.033	0.028
	5	3	0.328	0.083	0.420	0.432	0.070	0.073
		5	0.336	0.083	0.412	0.450	0.071	0.067
	10	3	0.685	0.139	0.760	0.754	0.133	0.139
		5	0.684	0.144	0.749	0.765	0.150	0.134
0.5	2	3	0.266	0.067	0.065	0.077	0.053	0.153
		5	0.268	0.068	0.066	0.072	0.062	0.088
	5	3	0.169	0.095	0.265	0.377	0.034	0.136
		5	0.270	0.089	0.266	0.372	0.091	0.234
	10	3	0.207	0.165	0.235	0.264	0.152	0.095
		5	0.209	0.262	0.245	0.262	0.247	0.346
No. of Shaded region			6	15	7	7	19	16

Shaded region indicate Type I error within [0.025 – 0.075]

Table K.6

Type I error rates for variable, $p = 3$ and $n = 25$

κ		μ	M	M_{GMed}	M_B	M_T	M_{BGMed}	M_{TGMed}
0	1	0	0.047	0.051	0.046	0.055	0.048	0.048
0.05	2	3	0.053	0.044	0.050	0.068	0.042	0.039
		5	0.056	0.045	0.059	0.063	0.039	0.040
	5	3	0.118	0.053	0.135	0.143	0.049	0.042
		5	0.123	0.049	0.134	0.141	0.042	0.045
	10	3	0.266	0.067	0.286	0.301	0.053	0.053
		5	0.268	0.068	0.299	0.311	0.062	0.058
0.1	2	3	0.069	0.045	0.065	0.077	0.034	0.036
		5	0.070	0.039	0.066	0.072	0.031	0.034
	5	3	0.207	0.065	0.235	0.264	0.052	0.045
		5	0.209	0.062	0.245	0.262	0.047	0.046
	10	3	0.467	0.082	0.518	0.548	0.078	0.075
		5	0.486	0.086	0.540	0.545	0.077	0.078
0.2	2	3	0.081	0.033	0.095	0.104	0.026	0.028
		5	0.085	0.033	0.088	0.107	0.029	0.028
	5	3	0.412	0.076	0.493	0.530	0.067	0.067
		5	0.404	0.077	0.502	0.520	0.060	0.065
	10	3	0.774	0.141	0.833	0.836	0.130	0.136
		5	0.782	0.137	0.830	0.853	0.129	0.127
0.5	2	3	0.180	0.042	0.177	0.197	0.044	0.135
		5	0.441	0.078	0.446	0.473	0.050	0.086
	5	3	0.429	0.055	0.465	0.477	0.095	0.058
		5	0.082	0.085	0.078	0.087	0.032	0.030
	10	3	0.080	0.132	0.086	0.086	0.127	0.090
		5	0.360	0.099	0.399	0.408	0.241	0.138
No. of Shaded region			5	15	5	4	18	18
Shaded region indicate Type I error within [0.025 – 0.075]								

Shaded region indicate Type I error within [0.025 – 0.075]

Table K.7

Type I error rates for variable, $p = 3$ and $n = 50$

ε	κ	μ	M	M_{GMed}	M_B	M_T	M_{BGMed}	M_{TGMed}
0	1	0	0.052	0.050	0.049	0.062	0.052	0.056
0.05	2	3	0.070	0.036	0.054	0.065	0.038	0.045
		5	0.063	0.044	0.061	0.064	0.039	0.043
	5	3	0.186	0.045	0.186	0.211	0.042	0.039
		5	0.180	0.042	0.177	0.197	0.044	0.035
	10	3	0.441	0.058	0.446	0.473	0.050	0.056
		5	0.429	0.055	0.465	0.477	0.055	0.058
0.1	2	3	0.082	0.025	0.078	0.087	0.032	0.030
		5	0.080	0.032	0.086	0.086	0.027	0.030
	5	3	0.360	0.039	0.399	0.408	0.041	0.038
		5	0.364	0.042	0.403	0.423	0.033	0.035
	10	3	0.746	0.069	0.768	0.781	0.067	0.062
		5	0.740	0.068	0.784	0.777	0.062	0.063
0.2	2	3	0.130	0.020	0.152	0.163	0.020	0.013
		5	0.122	0.017	0.148	0.147	0.021	0.015
	5	3	0.715	0.055	0.780	0.800	0.047	0.047
		5	0.712	0.055	0.777	0.786	0.046	0.050
	10	3	0.966	0.121	0.977	0.979	0.104	0.110
		5	0.969	0.117	0.975	0.980	0.120	0.114
0.5	2	3	0.103	0.063	0.194	0.190	0.071	0.097
		5	0.099	0.054	0.192	0.188	0.073	0.071
	5	3	0.148	0.048	0.065	0.065	0.084	0.136
		5	0.096	0.151	0.053	0.051	0.046	0.098
	10	3	0.093	0.091	0.158	0.153	0.093	0.158
		5	0.100	0.159	0.151	0.149	0.169	0.087
No. of Shaded region			3	18	5	5	18	16

Shaded region indicate Type I error within [0.025 – 0.075]

Table K.8

Type I error rates for variable, $p = 5$ and $n = 2$

ε	κ	μ	M_B	M_T	M_{BGMed}	M_{TGMed}
0	1	0	0.050	0.047	0.050	0.052
0.05	2	3	0.048	0.050	0.051	0.051
		5	0.048	0.050	0.051	0.051
	5	3	0.050	0.047	0.050	0.047
		5	0.050	0.048	0.050	0.053
	10	3	0.053	0.052	0.051	0.053
		5	0.053	0.056	0.051	0.051
0.1	2	3	0.048	0.050	0.048	0.054
		5	0.048	0.046	0.048	0.052
	5	3	0.052	0.056	0.046	0.052
		5	0.052	0.058	0.046	0.055
	10	3	0.057	0.058	0.047	0.050
		5	0.057	0.057	0.047	0.057
0.2	2	3	0.051	0.052	0.049	0.058
		5	0.051	0.055	0.049	0.060
	5	3	0.060	0.059	0.047	0.045
		5	0.059	0.060	0.046	0.050
	10	3	0.073	0.069	0.048	0.049
		5	0.073	0.074	0.047	0.055
0.5	2	3	0.115	0.108	0.052	0.063
		5	0.115	0.107	0.062	0.074
	5	3	0.056	0.055	0.093	0.080
		5	0.075	0.066	0.044	0.060
	10	3	0.099	0.093	0.148	0.143
		5	0.095	0.092	0.096	0.164
No. of Shaded region			21	21	22	22

Shaded region indicate Type I error within [0.025 – 0.075]

Table K.9

Type I error rates for variable, $p = 5$ and $n = 5$

ε	κ	μ	M_B	M_T	M_{BGMed}	M_{TGMed}
0	1	0	0.049	0.052	0.037	0.060
0.05	2	3	0.058	0.053	0.045	0.056
		5	0.050	0.060	0.054	0.061
	5	3	0.075	0.071	0.055	0.059
		5	0.073	0.069	0.049	0.057
	10	3	0.115	0.108	0.052	0.073
		5	0.115	0.107	0.052	0.074
0.1	2	3	0.056	0.055	0.053	0.060
		5	0.055	0.056	0.044	0.060
	5	3	0.099	0.093	0.048	0.063
		5	0.095	0.092	0.046	0.064
	10	3	0.181	0.164	0.068	0.075
		5	0.175	0.174	0.063	0.078
0.2	2	3	0.061	0.060	0.043	0.058
		5	0.063	0.061	0.049	0.059
	5	3	0.137	0.135	0.056	0.072
		5	0.145	0.130	0.059	0.070
	10	3	0.300	0.291	0.093	0.105
		5	0.295	0.300	0.090	0.115
0.5	2	3	0.194	0.190	0.071	0.167
		5	0.192	0.188	0.073	0.091
	5	3	0.065	0.075	0.044	0.036
		5	0.053	0.081	0.046	0.038
	10	3	0.158	0.153	0.083	0.058
		5	0.151	0.149	0.099	0.067
No. of Shaded region			9	9	21	20
Shaded region indicate Type I error within [0.025 – 0.075]						

Shaded region indicate Type I error within [0.025 – 0.075]

Table K.10

Type I error rates for variable, $p = 5$ and $n = 10$

ϵ	κ	μ	M	M_{GMed}	M_B	M_T	M_{BGMed}	M_{TGMed}
0	1	0	0.053	0.044	0.049	0.052	0.049	0.040
0.05	2	3	0.057	0.050	0.054	0.057	0.046	0.043
		5	0.053	0.049	0.059	0.059	0.052	0.043
	5	3	0.074	0.055	0.102	0.105	0.060	0.055
		5	0.071	0.056	0.098	0.096	0.057	0.050
	10	3	0.103	0.063	0.194	0.190	0.071	0.067
		5	0.099	0.064	0.192	0.188	0.073	0.071
0.1	2	3	0.048	0.048	0.065	0.065	0.044	0.036
		5	0.056	0.051	0.053	0.051	0.046	0.038
	5	3	0.093	0.061	0.158	0.153	0.073	0.058
		5	0.100	0.059	0.151	0.149	0.069	0.067
	10	3	0.172	0.090	0.336	0.316	0.079	0.093
		5	0.166	0.089	0.332	0.336	0.092	0.098
0.2	2	3	0.057	0.047	0.076	0.074	0.043	0.041
		5	0.065	0.041	0.073	0.075	0.044	0.034
	5	3	0.134	0.084	0.294	0.297	0.099	0.093
		5	0.131	0.085	0.306	0.287	0.106	0.089
	10	3	0.299	0.129	0.576	0.571	0.152	0.163
		5	0.309	0.133	0.577	0.577	0.170	0.163
0.5	2	3	0.169	0.080	0.265	0.259	0.083	0.082
		5	0.177	0.075	0.269	0.256	0.090	0.092
	5	3	0.264	0.084	0.063	0.063	0.042	0.137
		5	0.097	0.112	0.062	0.058	0.140	0.139
	10	3	0.147	0.094	0.223	0.205	0.065	0.098
		5	0.134	0.086	0.217	0.211	0.170	0.165
No. of Shaded region			9	14	8	9	15	13

Shaded region indicate Type I error within [0.025 – 0.075]

Table K.11

Type I error rates for variable, $p = 5$ and $n = 15$

ϵ	κ	μ	M	M_{GMed}	M_B	M_T	M_{BGMed}	M_{TGMed}
0	1	0	0.049	0.044	0.052	0.044	0.049	0.051
0.05	2	3	0.052	0.049	0.058	0.050	0.039	0.040
		5	0.055	0.047	0.055	0.053	0.045	0.046
	5	3	0.095	0.062	0.123	0.119	0.059	0.060
		5	0.092	0.061	0.125	0.132	0.052	0.052
	10	3	0.169	0.080	0.265	0.259	0.073	0.072
		5	0.177	0.075	0.269	0.256	0.070	0.072
0.1	2	3	0.064	0.044	0.063	0.063	0.042	0.037
		5	0.057	0.042	0.062	0.058	0.040	0.039
	5	3	0.147	0.074	0.223	0.205	0.065	0.068
		5	0.134	0.066	0.217	0.211	0.070	0.065
	10	3	0.314	0.114	0.470	0.459	0.102	0.108
		5	0.303	0.100	0.462	0.470	0.105	0.103
0.2	2	3	0.067	0.049	0.083	0.082	0.038	0.036
		5	0.069	0.049	0.090	0.085	0.038	0.038
	5	3	0.243	0.102	0.438	0.435	0.100	0.102
		5	0.238	0.100	0.432	0.423	0.103	0.098
	10	3	0.570	0.167	0.758	0.742	0.182	0.185
		5	0.558	0.169	0.753	0.747	0.185	0.175
0.5	2	3	0.251	0.079	0.319	0.308	0.068	0.066
		5	0.259	0.073	0.317	0.307	0.080	0.172
	5	3	0.171	0.043	0.070	0.062	0.144	0.036
		5	0.267	0.088	0.071	0.057	0.045	0.031
	10	3	0.199	0.095	0.269	0.250	0.266	0.093
		5	0.195	0.173	0.255	0.253	0.097	0.159
No. of Shaded region			7	14	7	7	15	16

Shaded region indicate Type I error within [0.025 – 0.075]

Table K.12

Type I error rates for variable, $p = 5$ and $n = 20$

ϵ	κ	μ	M	M_{GMed}	M_B	M_T	M_{BGMed}	M_{TGMed}
0	1	0	0.061	0.050	0.045	0.043	0.055	0.045
0.05	2	3	0.069	0.045	0.050	0.050	0.049	0.045
		5	0.069	0.047	0.053	0.048	0.046	0.039
	5	3	0.120	0.060	0.140	0.126	0.065	0.046
		5	0.118	0.056	0.135	0.127	0.063	0.053
	10	3	0.251	0.079	0.319	0.308	0.068	0.066
		5	0.259	0.073	0.317	0.307	0.080	0.072
0.1	2	3	0.071	0.043	0.070	0.062	0.044	0.036
		5	0.067	0.048	0.071	0.057	0.045	0.031
	5	3	0.199	0.075	0.269	0.250	0.066	0.063
		5	0.195	0.073	0.255	0.253	0.067	0.059
	10	3	0.449	0.105	0.561	0.546	0.114	0.099
		5	0.452	0.110	0.555	0.541	0.109	0.094
0.2	2	3	0.080	0.039	0.098	0.090	0.034	0.034
		5	0.091	0.042	0.099	0.099	0.042	0.033
	5	3	0.381	0.097	0.534	0.529	0.103	0.100
		5	0.376	0.098	0.525	0.499	0.099	0.094
	10	3	0.748	0.181	0.850	0.836	0.190	0.171
		5	0.738	0.194	0.846	0.845	0.186	0.184
0.5	2	3	0.310	0.171	0.369	0.367	0.167	0.171
		5	0.309	0.175	0.374	0.379	0.066	0.167
	5	3	0.267	0.241	0.265	0.370	0.037	0.035
		5	0.270	0.343	0.367	0.375	0.236	0.232
	10	3	0.241	0.171	0.318	0.326	0.064	0.056
		5	0.237	0.277	0.303	0.332	0.357	0.258
No. of Shaded region			5	12	5	5	15	15

Shaded region indicate Type I error within [0.025 – 0.075]

Table K.13

Type I error rates for variable, $p = 5$ and $n = 25$

ε	κ	μ	M	M_{GMed}	M_B	M_T	M_{BGMed}	M_{TGMed}
0	1	0	0.053	0.045	0.052	0.051	0.055	0.050
0.05	2	3	0.059	0.042	0.059	0.061	0.042	0.040
		5	0.065	0.047	0.063	0.065	0.045	0.037
	5	3	0.138	0.056	0.162	0.164	0.046	0.054
		5	0.133	0.062	0.149	0.165	0.054	0.048
	10	3	0.310	0.071	0.369	0.367	0.067	0.071
		5	0.309	0.075	0.374	0.379	0.066	0.067
0.1	2	3	0.067	0.041	0.065	0.070	0.037	0.035
		5	0.070	0.043	0.067	0.075	0.036	0.032
	5	3	0.241	0.071	0.318	0.326	0.064	0.056
		5	0.237	0.077	0.303	0.332	0.057	0.058
	10	3	0.546	0.101	0.631	0.652	0.066	0.088
		5	0.544	0.102	0.630	0.646	0.096	0.095
0.2	2	3	0.089	0.034	0.109	0.116	0.027	0.026
		5	0.092	0.041	0.114	0.118	0.025	0.031
	5	3	0.471	0.104	0.625	0.632	0.091	0.095
		5	0.462	0.101	0.611	0.635	0.097	0.096
	10	3	0.850	0.172	0.918	0.922	0.177	0.174
		5	0.840	0.181	0.911	0.919	0.169	0.175
0.5	2	3	0.188	0.052	0.230	0.247	0.034	0.047
		5	0.510	0.066	0.550	0.573	0.055	0.061
	5	3	0.498	0.164	0.551	0.571	0.058	0.064
		5	0.272	0.229	0.088	0.096	0.131	0.129
	10	3	0.373	0.231	0.086	0.088	0.126	0.128
		5	0.409	0.253	0.512	0.516	0.237	0.238
No. of Shaded region			5	14	5	5	17	16

Shaded region indicate Type I error within [0.025 – 0.075]

Table K.14

Type I error rates for variable, $p = 5$ and $n = 50$

ε	κ	μ	M	M_{GMed}	M_B	M_T	M_{BGMed}	M_{TGMed}
0	1	0	0.047	0.044	0.056	0.057	0.043	0.053
0.05	2	3	0.061	0.038	0.060	0.060	0.029	0.040
		5	0.057	0.035	0.069	0.062	0.031	0.041
	5	3	0.193	0.047	0.239	0.240	0.034	0.041
		5	0.188	0.052	0.230	0.247	0.034	0.047
	10	3	0.510	0.066	0.550	0.573	0.055	0.061
		5	0.498	0.064	0.551	0.571	0.058	0.064
0.1	2	3	0.072	0.029	0.088	0.096	0.031	0.029
		5	0.073	0.031	0.086	0.088	0.026	0.028
	5	3	0.409	0.053	0.512	0.516	0.037	0.038
		5	0.410	0.053	0.516	0.515	0.040	0.042
	10	3	0.806	0.081	0.859	0.866	0.071	0.073
		5	0.805	0.091	0.867	0.865	0.071	0.073
0.2	2	3	0.117	0.017	0.175	0.188	0.006	0.013
		5	0.118	0.021	0.169	0.178	0.010	0.013
	5	3	0.797	0.080	0.888	0.887	0.054	0.056
		5	0.787	0.076	0.884	0.888	0.061	0.061
	10	3	0.987	0.155	0.993	0.994	0.144	0.143
		5	0.988	0.161	0.995	0.994	0.141	0.154
0.5	2	3	0.172	0.056	0.253	0.245	0.066	0.072
		5	0.167	0.162	0.252	0.239	0.084	0.073
	5	3	0.249	0.251	0.064	0.067	0.095	0.090
		5	0.152	0.091	0.068	0.063	0.147	0.145
	10	3	0.262	0.165	0.210	0.192	0.073	0.066
		5	0.266	0.260	0.197	0.199	0.172	0.065
No. of Shaded region			5	12	5	5	17	19
Shaded region indicate Type I error within [0.025 – 0.075]								

Table K.15

Type I error rates for variable, $p = 8$ and $n = 2$

ε	κ	μ	M_B	M_T	M_{BGMed}	M_{TGMed}
0	1	0	0.051	0.049	0.050	0.052
0.05	2	3	0.050	0.043	0.051	0.045
		5	0.050	0.043	0.051	0.043
	5	3	0.051	0.047	0.050	0.044
		5	0.051	0.053	0.050	0.049
	10	3	0.058	0.059	0.050	0.051
		5	0.058	0.054	0.050	0.053
0.1	2	3	0.049	0.054	0.051	0.047
		5	0.049	0.053	0.051	0.053
	5	3	0.053	0.054	0.051	0.051
		5	0.052	0.052	0.049	0.051
	10	3	0.064	0.055	0.049	0.055
		5	0.064	0.060	0.050	0.054
0.2	2	3	0.049	0.059	0.048	0.050
		5	0.049	0.049	0.048	0.048
	5	3	0.058	0.060	0.046	0.049
		5	0.058	0.060	0.046	0.050
	10	3	0.076	0.083	0.046	0.051
		5	0.076	0.075	0.046	0.049
0.5	2	3	0.327	0.331	0.074	0.070
		5	0.071	0.085	0.084	0.141
	5	3	0.067	0.083	0.040	0.045
		5	0.266	0.281	0.091	0.066
	10	3	0.266	0.290	0.072	0.082
		5	0.553	0.548	0.122	0.119
No. of Shaded region			19	18	22	22

Shaded region indicate Type I error within [0.025 – 0.075]

Table K.16

Type I error rates for variable, $p = 8$ and $n = 5$

ε	κ	μ	M_B	M_T	M_{BGMed}	M_{TGMed}
0	1	0	0.050	0.055	0.048	0.054
0.05	2	3	0.049	0.053	0.048	0.046
		5	0.054	0.053	0.048	0.052
	5	3	0.074	0.073	0.053	0.060
		5	0.074	0.074	0.047	0.055
	10	3	0.126	0.125	0.061	0.069
		5	0.127	0.132	0.056	0.062
0.1	2	3	0.058	0.052	0.048	0.050
		5	0.055	0.054	0.042	0.053
	5	3	0.104	0.101	0.055	0.060
		5	0.104	0.103	0.060	0.047
	10	3	0.204	0.207	0.074	0.074
		5	0.202	0.199	0.074	0.074
0.2	2	3	0.062	0.056	0.042	0.047
		5	0.064	0.062	0.049	0.044
	5	3	0.164	0.168	0.066	0.073
		5	0.179	0.158	0.069	0.077
	10	3	0.357	0.371	0.113	0.118
		5	0.377	0.356	0.120	0.112
0.5	2	3	0.376	0.390	0.072	0.063
		5	0.388	0.398	0.074	0.074
	5	3	0.075	0.076	0.038	0.033
		5	0.065	0.066	0.037	0.040
	10	3	0.325	0.327	0.082	0.091
		5	0.322	0.337	0.097	0.175
No. of Shaded region			10	10	22	21

Shaded region indicate Type I error within [0.025 – 0.075]

Table K.17

Type I error rates for variable, $p = 8$ and $n = 10$

ε	κ	μ	M	M_{GMed}	M_B	M_T	M_{BGMed}	M_{TGMed}
0	1	0	0.043	0.055	0.057	0.054	0.058	0.045
0.05	2	3	0.045	0.046	0.064	0.059	0.046	0.042
		5	0.043	0.054	0.066	0.061	0.049	0.041
	5	3	0.054	0.062	0.127	0.125	0.063	0.058
		5	0.055	0.062	0.130	0.119	0.067	0.058
	10	3	0.072	0.056	0.253	0.245	0.066	0.072
		5	0.067	0.062	0.252	0.239	0.084	0.073
0.1	2	3	0.049	0.051	0.064	0.067	0.045	0.040
		5	0.052	0.051	0.068	0.063	0.047	0.045
	5	3	0.062	0.065	0.210	0.192	0.073	0.066
		5	0.066	0.060	0.197	0.199	0.072	0.065
	10	3	0.091	0.067	0.415	0.398	0.112	0.110
		5	0.094	0.089	0.399	0.431	0.124	0.117
0.2	2	3	0.050	0.058	0.075	0.088	0.046	0.041
		5	0.047	0.051	0.083	0.082	0.050	0.042
	5	3	0.074	0.078	0.376	0.390	0.126	0.114
		5	0.082	0.084	0.380	0.372	0.119	0.117
	10	3	0.152	0.096	0.682	0.685	0.200	0.204
		5	0.155	0.097	0.676	0.673	0.210	0.212
0.5	2	3	0.097	0.078	0.075	0.075	0.056	0.061
		5	0.217	0.087	0.065	0.390	0.072	0.063
	5	3	0.209	0.087	0.388	0.268	0.174	0.094
		5	0.151	0.071	0.175	0.176	0.138	0.133
	10	3	0.131	0.157	0.365	0.366	0.037	0.140
		5	0.156	0.092	0.325	0.327	0.082	0.171
No. of Shaded region			14	14	8	6	15	15
Shaded region indicate Type I error within [0.025 – 0.075]								

Shaded region indicate Type I error within [0.025 – 0.075]

Table K.18

Type I error rates for variable, $p = 8$ and $n = 15$

ε	κ	μ	M	M_{GMed}	M_B	M_T	M_{BGMed}	M_{TGMed}
0	1	0	0.037	0.054	0.052	0.057	0.043	0.047
0.05	2	3	0.043	0.052	0.055	0.070	0.045	0.052
		5	0.037	0.047	0.063	0.069	0.048	0.041
	5	3	0.064	0.060	0.146	0.146	0.064	0.060
		5	0.069	0.063	0.149	0.159	0.062	0.062
	10	3	0.128	0.075	0.319	0.344	0.074	0.054
		5	0.124	0.081	0.327	0.331	0.074	0.070
0.1	2	3	0.044	0.054	0.071	0.085	0.044	0.041
		5	0.045	0.047	0.067	0.083	0.040	0.045
	5	3	0.092	0.073	0.266	0.281	0.091	0.066
		5	0.099	0.077	0.266	0.290	0.072	0.082
	10	3	0.236	0.118	0.553	0.548	0.122	0.119
		5	0.240	0.105	0.537	0.567	0.125	0.113
0.2	2	3	0.051	0.052	0.102	0.114	0.039	0.042
		5	0.059	0.055	0.096	0.105	0.038	0.040
	5	3	0.181	0.095	0.533	0.556	0.128	0.128
		5	0.196	0.104	0.515	0.551	0.127	0.132
	10	3	0.467	0.174	0.835	0.852	0.216	0.219
		5	0.473	0.186	0.827	0.830	0.220	0.226
0.5	2	3	0.272	0.087	0.439	0.458	0.074	0.070
		5	0.274	0.093	0.453	0.450	0.075	0.075
	5	3	0.258	0.057	0.079	0.081	0.035	0.044
		5	0.150	0.054	0.085	0.080	0.233	0.241
	10	3	0.193	0.092	0.389	0.404	0.275	0.175
		5	0.195	0.090	0.397	0.404	0.167	0.173
No. of Shaded region			9	13	5	3	15	15

Shaded region indicate Type I error within [0.025 – 0.075]

Table K.19

Type I error rates for variable, $p = 8$ and $n = 20$

ε	κ	μ	M	M_{GMed}	M_B	M_T	M_{BGMed}	M_{TGMed}
0	1	0	0.038	0.058	0.045	0.047	0.044	0.047
0.05	2	3	0.048	0.051	0.056	0.067	0.045	0.047
		5	0.044	0.056	0.055	0.059	0.045	0.053
	5	3	0.091	0.076	0.166	0.181	0.058	0.055
		5	0.097	0.068	0.160	0.175	0.056	0.061
	10	3	0.217	0.087	0.376	0.390	0.072	0.063
		5	0.209	0.087	0.388	0.398	0.074	0.074
0.1	2	3	0.051	0.061	0.075	0.076	0.038	0.033
		5	0.051	0.057	0.065	0.066	0.037	0.040
	5	3	0.156	0.092	0.325	0.327	0.082	0.071
		5	0.169	0.087	0.322	0.337	0.077	0.075
	10	3	0.418	0.128	0.641	0.639	0.113	0.123
		5	0.412	0.129	0.646	0.650	0.116	0.116
0.2	2	3	0.066	0.055	0.107	0.117	0.038	0.037
		5	0.068	0.061	0.109	0.116	0.033	0.035
	5	3	0.320	0.124	0.638	0.654	0.136	0.134
		5	0.319	0.128	0.626	0.641	0.131	0.132
	10	3	0.722	0.215	0.901	0.910	0.223	0.215
		5	0.721	0.216	0.911	0.911	0.222	0.225
0.5	2	3	0.150	0.165	0.205	0.276	0.062	0.063
		5	0.484	0.089	0.439	0.633	0.061	0.166
	5	3	0.485	0.085	0.453	0.654	0.162	0.163
		5	0.334	0.241	0.079	0.085	0.326	0.224
	10	3	0.234	0.136	0.085	0.085	0.326	0.329
		5	0.358	0.175	0.389	0.602	0.251	0.348
No. of Shaded region			7	8	5	4	13	14
Shaded region indicate Type I error within [0.025 – 0.075]								

Shaded region indicate Type I error within [0.025 – 0.075]

Table K.20

Type I error rates for variable, $p = 8$ and $n = 25$

ϵ	κ	μ	M	M_{GMed}	M_B	M_T	M_{BGMed}	M_{TGMed}
0	1	0	0.040	0.057	0.053	0.053	0.047	0.056
0.05	2	3	0.049	0.051	0.066	0.062	0.040	0.047
		5	0.043	0.060	0.055	0.067	0.041	0.052
	5	3	0.102	0.074	0.198	0.199	0.057	0.062
		5	0.104	0.071	0.205	0.206	0.056	0.051
	10	3	0.272	0.087	0.439	0.458	0.074	0.070
		5	0.274	0.093	0.453	0.450	0.074	0.077
0.1	2	3	0.058	0.057	0.079	0.081	0.035	0.044
		5	0.050	0.054	0.085	0.080	0.033	0.041
	5	3	0.193	0.092	0.389	0.404	0.075	0.075
		5	0.195	0.090	0.397	0.404	0.067	0.073
	10	3	0.520	0.132	0.723	0.736	0.104	0.118
		5	0.510	0.128	0.732	0.728	0.113	0.120
0.2	2	3	0.070	0.056	0.132	0.145	0.028	0.033
		5	0.067	0.058	0.134	0.141	0.027	0.037
	5	3	0.426	0.142	0.745	0.744	0.122	0.133
		5	0.433	0.139	0.738	0.750	0.121	0.135
	10	3	0.848	0.223	0.956	0.952	0.217	0.223
		5	0.838	0.218	0.954	0.954	0.202	0.228
0.5	2	3	0.062	0.249	0.071	0.069	0.039	0.043
		5	0.114	0.270	0.306	0.291	0.058	0.042
	5	3	0.114	0.072	0.306	0.305	0.086	0.092
		5	0.214	0.111	0.604	0.586	0.118	0.127
	10	3	0.218	0.111	0.591	0.58	0.124	0.129
		5	0.062	0.258	0.102	0.097	0.158	0.142
No. of Shaded region			7	9	3	3	15	14
Shaded region indicate Type I error within [0.025 – 0.075]								

Shaded region indicate Type I error within [0.025 – 0.075]

Table K.21

Type I error rate for variable, $p = 8$ and $n = 50$

ε	κ	μ	M	M_{GMed}	M_B	M_T	M_{BGMed}	M_{TGMed}
0	1	0	0.023	0.050	0.053	0.047	0.052	0.049
0.05	2	3	0.027	0.043	0.066	0.058	0.030	0.033
		5	0.028	0.043	0.055	0.057	0.035	0.035
	5	3	0.155	0.066	0.198	0.287	0.045	0.038
		5	0.150	0.065	0.205	0.276	0.040	0.044
	10	3	0.484	0.089	0.439	0.633	0.061	0.066
		5	0.485	0.085	0.453	0.654	0.062	0.063
0.1	2	3	0.034	0.041	0.079	0.085	0.026	0.024
		5	0.034	0.036	0.085	0.085	0.026	0.029
	5	3	0.358	0.075	0.389	0.602	0.051	0.048
		5	0.361	0.076	0.397	0.593	0.048	0.050
	10	3	0.830	0.124	0.723	0.907	0.074	0.072
		5	0.820	0.113	0.732	0.917	0.063	0.069
0.2	2	3	0.064	0.035	0.132	0.208	0.013	0.014
		5	0.072	0.034	0.134	0.196	0.011	0.011
	5	3	0.771	0.130	0.745	0.942	0.103	0.096
		5	0.760	0.120	0.738	0.945	0.094	0.091
	10	3	0.989	0.203	0.956	0.998	0.184	0.186
		5	0.991	0.224	0.954	0.999	0.182	0.180
0.5	2	3	0.202	0.088	0.073	0.437	0.065	0.064
		5	0.201	0.081	0.073	0.433	0.069	0.068
	5	3	0.266	0.057	0.479	0.078	0.073	0.076
		5	0.364	0.052	0.473	0.082	0.062	0.069
	10	3	0.173	0.083	0.366	0.383	0.084	0.089
		5	0.167	0.083	0.380	0.379	0.082	0.082
No. of Shaded region			7	12	5	3	17	17

Shaded region indicate Type I error within [0.025 – 0.075]

Appendix L.

Table L.22

Type I error rates for variable, $p = 10$ and $n = 5$

ε	κ	μ	M_B	M_T	M_{BGMed}	M_{TGMed}
0	1	0	0.050	0.047	0.055	0.058
0.05	2	3	0.051	0.050	0.061	0.056
		5	0.054	0.054	0.056	0.050
	5	3	0.077	0.074	0.060	0.060
		5	0.080	0.072	0.059	0.059
	10	3	0.140	0.143	0.075	0.073
		5	0.144	0.137	0.072	0.060
0.1	2	3	0.057	0.055	0.055	0.059
		5	0.054	0.055	0.057	0.052
	5	3	0.116	0.106	0.062	0.066
		5	0.122	0.111	0.067	0.066
	10	3	0.221	0.226	0.098	0.099
		5	0.231	0.227	0.092	0.087
0.2	2	3	0.066	0.061	0.058	0.054
		5	0.191	0.062	0.057	0.054
	5	3	0.192	0.173	0.070	0.075
		5	0.412	0.182	0.073	0.075
	10	3	0.409	0.396	0.147	0.134
		5	0.448	0.412	0.142	0.137
0.5	2	3	0.254	0.258	0.094	0.084
		5	0.244	0.250	0.086	0.088
	5	3	0.064	0.054	0.059	0.063
		5	0.063	0.058	0.072	0.079
	10	3	0.203	0.195	0.065	0.075
		5	0.215	0.193	0.074	0.065
No. of Shaded region			9	11	19	19

Shaded region indicate Type I error within [0.025 – 0.075]

Table L.23

Type I error rates for variable, $p = 10$ and $n = 10$

ε	κ	μ	M_B	M_T	M_{BGMed}	M_{TGMed}
0	1	0	0.045	0.041	0.055	0.049
0.05	2	3	0.054	0.048	0.054	0.053
		5	0.051	0.045	0.059	0.046
	5	3	0.116	0.110	0.067	0.066
		5	0.115	0.104	0.066	0.066
	10	3	0.254	0.258	0.094	0.084
		5	0.244	0.250	0.086	0.088
0.1	2	3	0.064	0.054	0.049	0.043
		5	0.063	0.058	0.052	0.039
	5	3	0.203	0.195	0.065	0.075
		5	0.215	0.193	0.074	0.065
	10	3	0.442	0.433	0.125	0.126
		5	0.444	0.442	0.128	0.139
0.2	2	3	0.080	0.079	0.058	0.040
		5	0.401	0.083	0.047	0.044
	5	3	0.401	0.390	0.139	0.135
		5	0.707	0.394	0.140	0.126
	10	3	0.719	0.702	0.228	0.218
		5	0.773	0.693	0.226	0.226
0.5	2	3	0.140	0.143	0.075	0.083
		5	0.144	0.137	0.072	0.160
	5	3	0.057	0.055	0.155	0.259
		5	0.054	0.055	0.157	0.092
	10	3	0.116	0.106	0.262	0.166
		5	0.122	0.111	0.267	0.096
No. of Shaded region			7	7	13	11

Shaded region indicate Type I error within [0.025 – 0.075]

Table L.24

Type I error rates for variable, $p = 10$ and $n = 15$

ε	κ	μ	M	M_{GMed}	M_B	M_T	M_{BGMed}	M_{TGMed}
0	1	0	0.056	0.051	0.051	0.046	0.049	0.051
0.05	2	3	0.055	0.052	0.061	0.056	0.041	0.049
		5	0.055	0.049	0.062	0.063	0.043	0.047
	5	3	0.083	0.060	0.167	0.157	0.063	0.065
		5	0.081	0.065	0.168	0.156	0.066	0.066
	10	3	0.119	0.075	0.355	0.351	0.065	0.064
		5	0.116	0.061	0.342	0.343	0.075	0.067
0.1	2	3	0.060	0.050	0.073	0.064	0.040	0.045
		5	0.062	0.049	0.071	0.069	0.039	0.043
	5	3	0.114	0.070	0.306	0.291	0.089	0.086
		5	0.114	0.072	0.306	0.305	0.086	0.092
	10	3	0.214	0.111	0.604	0.586	0.118	0.127
		5	0.218	0.111	0.591	0.580	0.124	0.129
0.2	2	3	0.062	0.058	0.102	0.097	0.058	0.042
		5	0.067	0.059	0.583	0.099	0.043	0.043
	5	3	0.175	0.105	0.566	0.579	0.151	0.149
		5	0.191	0.093	0.866	0.574	0.146	0.151
	10	3	0.433	0.177	0.861	0.847	0.245	0.238
		5	0.435	0.180	0.885	0.859	0.234	0.229
0.5	2	3	0.303	0.096	0.490	0.485	0.075	0.083
		5	0.304	0.101	0.496	0.495	0.063	0.073
	5	3	0.368	0.262	0.085	0.087	0.035	0.040
		5	0.272	0.252	0.084	0.081	0.038	0.035
	10	3	0.242	0.081	0.441	0.440	0.074	0.074
		5	0.243	0.092	0.445	0.450	0.082	0.079
No. of Shaded region			7	13	5	5	16	15
Shaded region indicate Type I error within [0.025 – 0.075]								

Shaded region indicate Type I error within [0.025 – 0.075]

Table L.25

Type I error rates for variable, $p = 10$ and $n = 20$

ε	κ	μ	M	M_{GMed}	M_B	M_T	M_{BGMed}	M_{TGMed}
0	1	0	0.047	0.057	0.052	0.053	0.051	0.055
0.05	2	3	0.054	0.050	0.055	0.070	0.038	0.049
		5	0.053	0.050	0.056	0.063	0.045	0.050
	5	3	0.110	0.074	0.193	0.197	0.063	0.066
		5	0.109	0.064	0.189	0.208	0.064	0.070
	10	3	0.202	0.088	0.415	0.437	0.065	0.064
		5	0.201	0.081	0.415	0.433	0.069	0.068
0.1	2	3	0.066	0.057	0.073	0.078	0.043	0.046
		5	0.064	0.052	0.073	0.082	0.042	0.039
	5	3	0.173	0.083	0.366	0.383	0.084	0.089
		5	0.167	0.083	0.380	0.379	0.082	0.082
	10	3	0.390	0.132	0.679	0.709	0.126	0.127
		5	0.414	0.122	0.676	0.697	0.123	0.128
0.2	2	3	0.072	0.058	0.120	0.131	0.043	0.042
		5	0.069	0.056	0.693	0.130	0.035	0.041
	5	3	0.334	0.126	0.688	0.702	0.147	0.165
		5	0.324	0.117	0.926	0.714	0.151	0.154
	10	3	0.719	0.206	0.920	0.934	0.229	0.252
		5	0.709	0.214	0.925	0.928	0.242	0.243
0.5	2	3	0.236	0.171	0.330	0.329	0.143	0.144
		5	0.609	0.085	0.687	0.682	0.167	0.268
	5	3	0.599	0.087	0.687	0.694	0.063	0.073
		5	0.474	0.249	0.104	0.096	0.042	0.123
	10	3	0.375	0.246	0.096	0.108	0.227	0.028
		5	0.495	0.087	0.673	0.660	0.057	0.262
No. of Shaded region			7	9	5	3	14	13
Shaded region indicate Type I error within [0.025 – 0.075]								

Shaded region indicate Type I error within [0.025 – 0.075]

Table L.26

Type I error rates for variable, $p = 10$ and $n = 25$

ε	κ	μ	M	M_{GMed}	M_B	M_T	M_{BGMed}	M_{TGMed}
0	1	0	0.049	0.056	0.056	0.047	0.050	0.050
0.05	2	3	0.060	0.050	0.068	0.066	0.041	0.039
		5	0.065	0.053	0.066	0.060	0.045	0.045
	5	3	0.126	0.075	0.231	0.218	0.059	0.056
		5	0.125	0.069	0.226	0.223	0.060	0.067
	10	3	0.303	0.096	0.490	0.485	0.075	0.083
		5	0.304	0.101	0.496	0.495	0.063	0.073
0.1	2	3	0.068	0.062	0.085	0.087	0.035	0.040
		5	0.072	0.052	0.084	0.081	0.038	0.035
	5	3	0.242	0.081	0.441	0.440	0.074	0.074
		5	0.243	0.092	0.445	0.450	0.082	0.079
	10	3	0.559	0.135	0.774	0.760	0.126	0.126
		5	0.557	0.135	0.771	0.764	0.120	0.114
0.2	2	3	0.081	0.058	0.152	0.149	0.064	0.029
		5	0.084	0.055	0.785	0.149	0.030	0.032
	5	3	0.481	0.131	0.788	0.784	0.140	0.141
		5	0.476	0.135	0.961	0.786	0.151	0.138
	10	3	0.865	0.233	0.965	0.967	0.228	0.231
		5	0.869	0.215	0.975	0.965	0.230	0.235
0.5	2	3	0.155	0.086	0.470	0.492	0.066	0.074
		5	0.151	0.084	0.479	0.513	0.078	0.077
	5	3	0.162	0.060	0.075	0.097	0.040	0.040
		5	0.163	0.056	0.075	0.089	0.047	0.051
	10	3	0.134	0.175	0.416	0.458	0.099	0.098
		5	0.137	0.171	0.427	0.462	0.098	0.107
No. of Shaded region			5	11	5	3	15	14
Shaded region indicate Type I error within [0.025 – 0.075]								

Shaded region indicate Type I error within [0.025 – 0.075]

Table L.27

Type I error rates for variable, $p = 10$ and $n = 50$

ε	κ	μ	M	M_{GMed}	M_B	M_T	M_{BGMed}	M_{TGMed}
0	1	0	0.050	0.056	0.045	0.045	0.044	0.052
0.05	2	3	0.056	0.051	0.066	0.062	0.031	0.039
		5	0.059	0.052	0.067	0.069	0.026	0.030
	5	3	0.228	0.066	0.325	0.335	0.041	0.050
		5	0.236	0.071	0.330	0.329	0.043	0.044
	10	3	0.609	0.085	0.687	0.682	0.067	0.068
		5	0.599	0.087	0.687	0.694	0.063	0.073
0.1	2	3	0.074	0.049	0.104	0.096	0.042	0.023
		5	0.075	0.046	0.096	0.108	0.027	0.028
	5	3	0.495	0.087	0.673	0.660	0.057	0.062
		5	0.498	0.090	0.678	0.662	0.055	0.059
	10	3	0.882	0.133	0.940	0.936	0.099	0.103
		5	0.890	0.125	0.936	0.938	0.100	0.099
0.2	2	3	0.125	0.043	0.240	0.249	0.035	0.013
		5	0.130	0.047	0.965	0.245	0.032	0.033
	5	3	0.871	0.153	0.964	0.965	0.114	0.126
		5	0.865	0.144	0.988	0.961	0.113	0.120
	10	3	0.996	0.236	0.989	0.999	0.207	0.210
		5	0.996	0.230	0.993	0.998	0.188	0.204
0.5	2	3	0.116	0.062	0.270	0.260	0.063	0.072
		5	0.237	0.084	0.562	0.542	0.036	0.041
	5	3	0.241	0.084	0.565	0.554	0.038	0.042
		5	0.161	0.148	0.087	0.090	0.069	0.067
	10	3	0.268	0.154	0.093	0.087	0.088	0.092
		5	0.194	0.174	0.516	0.521	0.099	0.097
No. of Shaded region			5	10	3	3	17	15

Shaded region indicate Type I error within [0.025 – 0.075]

Table L.28

Type I error rates for variable, $p = 15$ and $n = 5$

ϵ	κ	μ	M_B	M_T	M_{BGMed}	M_{TGMed}
0	1	0	0.053	0.048	0.048	0.050
0.05	2	3	0.051	0.049	0.042	0.048
		5	0.051	0.048	0.047	0.051
	5	3	0.084	0.084	0.056	0.059
		5	0.084	0.088	0.056	0.057
	10	3	0.168	0.157	0.072	0.075
		5	0.168	0.169	0.074	0.066
0.1	2	3	0.065	0.055	0.045	0.047
		5	0.053	0.053	0.045	0.054
	5	3	0.138	0.129	0.063	0.068
		5	0.134	0.127	0.063	0.066
	10	3	0.279	0.280	0.096	0.109
		5	0.289	0.278	0.104	0.101
0.2	2	3	0.063	0.060	0.051	0.049
		5	0.070	0.063	0.048	0.050
	5	3	0.223	0.223	0.074	0.070
		5	0.229	0.234	0.081	0.090
	10	3	0.472	0.458	0.155	0.162
		5	0.482	0.479	0.152	0.168
0.5	2	3	0.141	0.145	0.068	0.072
		5	0.301	0.295	0.097	0.094
	5	3	0.301	0.295	0.091	0.097
		5	0.068	0.062	0.053	0.048
	10	3	0.070	0.067	0.046	0.050
		5	0.264	0.258	0.109	0.069
No. of Shaded region			9	9	17	17

Shaded region indicate Type I error within [0.025 – 0.075]

Table L.29

Type I error rates for variable, $p = 15$ and $n = 10$

ε	κ	μ	M_B	M_T	M_{BGMed}	M_{TGMed}
0	1	0	0.050	0.045	0.054	0.046
0.05	2	3	0.053	0.051	0.048	0.048
		5	0.063	0.057	0.049	0.049
	5	3	0.141	0.128	0.074	0.059
		5	0.141	0.145	0.068	0.072
	10	3	0.301	0.295	0.097	0.094
		5	0.301	0.295	0.091	0.097
0.1	2	3	0.068	0.062	0.053	0.048
		5	0.070	0.067	0.046	0.050
	5	3	0.264	0.258	0.109	0.069
		5	0.261	0.258	0.099	0.069
	10	3	0.500	0.505	0.140	0.134
		5	0.520	0.504	0.141	0.146
0.2	2	3	0.093	0.094	0.052	0.044
		5	0.088	0.087	0.047	0.044
	5	3	0.489	0.489	0.164	0.155
		5	0.491	0.467	0.152	0.162
	10	3	0.773	0.772	0.247	0.256
		5	0.774	0.768	0.251	0.252
0.5	2	3	0.084	0.088	0.056	0.157
		5	0.168	0.157	0.172	0.175
	5	3	0.168	0.169	0.174	0.066
		5	0.065	0.055	0.245	0.247
	10	3	0.053	0.053	0.245	0.254
		5	0.138	0.129	0.063	0.068
No. of Shaded region			7	7	11	13

Shaded region indicate Type I error within [0.025 – 0.075]

Table L.30

Type I error rates for variable, $p = 15$ and $n = 15$

ε	κ	μ	M_B	M_T	M_{BGMed}	M_{TGMed}
0	1	0	0.051	0.046	0.048	0.050
0.05	2	3	0.058	0.056	0.046	0.052
		5	0.055	0.059	0.051	0.051
	5	3	0.173	0.196	0.075	0.081
		5	0.178	0.184	0.076	0.077
	10	3	0.374	0.391	0.070	0.067
		5	0.395	0.395	0.089	0.076
0.1	2	3	0.063	0.069	0.045	0.046
		5	0.063	0.072	0.045	0.049
	5	3	0.339	0.350	0.102	0.100
		5	0.332	0.350	0.100	0.100
	10	3	0.642	0.635	0.146	0.145
		5	0.639	0.645	0.143	0.146
0.2	2	3	0.103	0.111	0.052	0.047
		5	0.100	0.116	0.049	0.050
	5	3	0.642	0.650	0.196	0.186
		5	0.649	0.656	0.197	0.190
	10	3	0.889	0.898	0.277	0.268
		5	0.893	0.899	0.274	0.262
0.5	2	3	0.174	0.192	0.083	0.089
		5	0.195	0.183	0.081	0.077
	5	3	0.063	0.069	0.052	0.045
		5	0.051	0.062	0.046	0.044
	10	3	0.153	0.153	0.173	0.166
		5	0.146	0.155	0.264	0.269
No. of Shaded region			9	9	11	10

Shaded region indicate Type I error within [0.025 – 0.075]

Table L.31

Type I error rates for variable, $p = 15$ and $n = 20$

ε	κ	μ	M	M_{GMed}	M_B	M_T	M_{BGMed}	M_{TGMed}
0	1	0	0.049	0.054	0.046	0.053	0.049	0.053
0.05	2	3	0.058	0.060	0.064	0.066	0.037	0.050
		5	0.052	0.052	0.061	0.061	0.047	0.051
	5	3	0.083	0.067	0.229	0.241	0.065	0.067
		5	0.090	0.063	0.225	0.243	0.068	0.068
	10	3	0.155	0.086	0.470	0.492	0.066	0.074
		5	0.151	0.084	0.479	0.513	0.078	0.077
0.1	2	3	0.062	0.060	0.075	0.097	0.040	0.040
		5	0.063	0.056	0.075	0.089	0.047	0.051
	5	3	0.134	0.075	0.416	0.458	0.099	0.098
		5	0.137	0.071	0.427	0.462	0.098	0.107
	10	3	0.279	0.117	0.738	0.756	0.138	0.134
		5	0.280	0.115	0.747	0.758	0.138	0.139
0.2	2	3	0.070	0.060	0.134	0.154	0.039	0.042
		5	0.072	0.063	0.142	0.160	0.043	0.041
	5	3	0.238	0.101	0.781	0.788	0.187	0.196
		5	0.244	0.109	0.772	0.793	0.191	0.194
	10	3	0.572	0.195	0.953	0.958	0.262	0.262
		5	0.559	0.189	0.954	0.951	0.268	0.265
0.5	2	3	0.090	0.058	0.286	0.296	0.072	0.076
		5	0.169	0.080	0.574	0.592	0.078	0.077
	5	3	0.172	0.080	0.605	0.588	0.093	0.091
		5	0.164	0.153	0.085	0.093	0.041	0.041
	10	3	0.256	0.250	0.092	0.094	0.042	0.046
		5	0.153	0.064	0.563	0.569	0.118	0.115
No. of Shaded region			7	11	5	3	13	12

Shaded region indicate Type I error within [0.025 – 0.075]

Table L.32

Type I error rates for variable, $p = 15$ and $n = 25$

ε	κ	μ	M	M_{GMed}	M_B	M_T	M_{BGMed}	M_{TGMed}
0	1	0	0.050	0.050	0.051	0.050	0.050	0.054
0.05	2	3	0.054	0.054	0.068	0.070	0.040	0.048
		5	0.054	0.045	0.065	0.062	0.040	0.042
	5	3	0.114	0.063	0.265	0.261	0.062	0.069
		5	0.116	0.062	0.270	0.260	0.063	0.072
	10	3	0.237	0.084	0.562	0.542	0.082	0.092
		5	0.241	0.084	0.565	0.554	0.090	0.095
0.1	2	3	0.061	0.048	0.087	0.090	0.036	0.041
		5	0.068	0.054	0.093	0.087	0.038	0.042
	5	3	0.194	0.074	0.516	0.521	0.069	0.067
		5	0.206	0.081	0.517	0.502	0.094	0.097
	10	3	0.464	0.128	0.822	0.811	0.134	0.140
		5	0.478	0.120	0.821	0.819	0.132	0.135
0.2	2	3	0.080	0.057	0.183	0.186	0.042	0.042
		5	0.076	0.058	0.177	0.168	0.033	0.046
	5	3	0.421	0.123	0.856	0.857	0.187	0.189
		5	0.406	0.132	0.854	0.862	0.191	0.190
	10	3	0.816	0.221	0.975	0.977	0.260	0.274
		5	0.812	0.224	0.980	0.978	0.276	0.276
0.5	2	3	0.249	0.074	0.446	0.438	0.064	0.060
		5	0.603	0.104	0.787	0.798	0.070	0.073
	5	3	0.601	0.099	0.784	0.801	0.094	0.075
		5	0.080	0.255	0.117	0.137	0.135	0.087
	10	3	0.087	0.256	0.126	0.123	0.031	0.132
		5	0.533	0.101	0.808	0.802	0.101	0.086
No. of Shaded region			5	11	3	3	13	13
Shaded region indicate Type I error within [0.025 – 0.075]								

Shaded region indicate Type I error within [0.025 – 0.075]

Table L.33

Type I error rates for variable, $p = 15$ and $n = 50$

ε	κ	μ	M	M_{GMed}	M_B	M_T	M_{BGMed}	M_{TGMed}
0	1	0	0.049	0.053	0.050	0.052	0.047	0.046
0.05	2	3	0.063	0.046	0.073	0.073	0.031	0.029
		5	0.063	0.044	0.071	0.070	0.030	0.031
	5	3	0.270	0.067	0.401	0.409	0.049	0.046
		5	0.261	0.070	0.387	0.406	0.055	0.050
	10	3	0.611	0.068	0.751	0.762	0.071	0.069
		5	0.619	0.074	0.752	0.759	0.071	0.072
0.1	2	3	0.084	0.053	0.118	0.124	0.025	0.026
		5	0.084	0.046	0.117	0.132	0.027	0.025
	5	3	0.533	0.094	0.754	0.756	0.071	0.074
		5	0.522	0.092	0.760	0.765	0.073	0.075
	10	3	0.899	0.144	0.960	0.957	0.112	0.111
		5	0.900	0.140	0.959	0.956	0.113	0.115
0.2	2	3	0.132	0.050	0.312	0.316	0.057	0.018
		5	0.134	0.053	0.312	0.315	0.018	0.025
	5	3	0.888	0.164	0.983	0.985	0.175	0.178
		5	0.898	0.167	0.983	0.987	0.179	0.168
	10	3	0.997	0.266	1.000	0.999	0.263	0.262
		5	0.998	0.254	1.000	0.999	0.262	0.274
0.5	2	3	0.564	0.104	0.819	0.826	0.168	0.172
		5	0.577	0.097	0.811	0.824	0.061	0.175
	5	3	0.082	0.160	0.136	0.145	0.225	0.236
		5	0.090	0.187	0.135	0.147	0.031	0.230
	10	3	0.521	0.103	0.846	0.839	0.108	0.104
		5	0.528	0.104	0.836	0.850	0.104	0.108
No. of Shaded region			3	11	3	3	14	12

Shaded region indicate Type I error within [0.025 – 0.075]

Table L.34

Type I error rates for variable, $p = 20$ and $n = 5$

ε	κ	μ	M_B	M_T	M_{BGMed}	M_{TGMed}
0	1	0	0.051	0.056	0.055	0.053
0.05	2	3	0.049	0.055	0.049	0.044
		5	0.056	0.058	0.051	0.049
	5	3	0.084	0.103	0.058	0.055
		5	0.092	0.100	0.058	0.060
	10	3	0.174	0.192	0.083	0.089
		5	0.195	0.183	0.081	0.077
0.1	2	3	0.063	0.069	0.052	0.045
		5	0.051	0.062	0.046	0.044
	5	3	0.153	0.153	0.073	0.066
		5	0.146	0.155	0.064	0.069
	10	3	0.315	0.315	0.116	0.107
		5	0.309	0.318	0.118	0.115
0.2	2	3	0.073	0.075	0.050	0.041
		5	0.060	0.072	0.051	0.052
	5	3	0.265	0.268	0.093	0.089
		5	0.258	0.282	0.089	0.086
	10	3	0.518	0.523	0.182	0.170
		5	0.525	0.539	0.183	0.172
0.5	2	3	0.353	0.333	0.073	0.082
		5	0.616	0.617	0.096	0.097
	5	3	0.633	0.617	0.094	0.099
		5	0.113	0.104	0.041	0.041
	10	3	0.109	0.105	0.038	0.042
		5	0.621	0.612	0.119	0.127
No. of Shaded region			9	9	14	13

Shaded region indicate Type I error within [0.025 – 0.075]

Table L.35

Type I error rates for variable, $p = 20$ and $n = 10$

ε	κ	μ	M_B	M_T	M_{BGMed}	M_{TGMed}
0	1	0	0.048	0.053	0.050	0.048
0.05	2	3	0.054	0.055	0.053	0.051
		5	0.057	0.054	0.049	0.043
	5	3	0.164	0.160	0.075	0.067
		5	0.157	0.158	0.074	0.069
	10	3	0.332	0.341	0.074	0.073
		5	0.333	0.337	0.075	0.085
0.1	2	3	0.067	0.070	0.044	0.048
		5	0.068	0.066	0.052	0.052
	5	3	0.288	0.290	0.104	0.108
		5	0.289	0.292	0.106	0.106
	10	3	0.546	0.548	0.149	0.149
		5	0.553	0.556	0.138	0.151
0.2	2	3	0.100	0.101	0.050	0.054
		5	0.096	0.091	0.049	0.048
	5	3	0.551	0.531	0.188	0.183
		5	0.537	0.553	0.180	0.180
	10	3	0.810	0.807	0.267	0.280
		5	0.816	0.809	0.265	0.260
0.5	2	3	0.300	0.284	0.078	0.079
		5	0.291	0.285	0.086	0.082
	5	3	0.551	0.554	0.101	0.091
		5	0.560	0.554	0.100	0.096
	10	3	0.098	0.100	0.048	0.043
		5	0.106	0.089	0.043	0.042
No. of Shaded region			5	5	13	12

Shaded region indicate Type I error within [0.025 – 0.075]

Table L.36

Type I error rates for variable, $p = 20$ and $n = 15$

ε	κ	μ	M_B	M_T	M_{BGMed}	M_{TGMed}
0	1	0	0.052	0.050	0.054	0.056
0.05	2	3	0.060	0.063	0.053	0.048
		5	0.061	0.067	0.048	0.048
	5	3	0.223	0.214	0.080	0.089
		5	0.221	0.226	0.074	0.089
	10	3	0.426	0.441	0.097	0.095
		5	0.432	0.453	0.100	0.093
0.1	2	3	0.078	0.087	0.050	0.050
		5	0.079	0.081	0.046	0.052
	5	3	0.407	0.416	0.113	0.119
		5	0.412	0.403	0.110	0.125
	10	3	0.710	0.693	0.146	0.151
		5	0.704	0.688	0.144	0.145
0.2	2	3	0.143	0.131	0.046	0.055
		5	0.130	0.133	0.037	0.056
	5	3	0.730	0.719	0.219	0.221
		5	0.725	0.720	0.210	0.224
	10	3	0.929	0.925	0.274	0.288
		5	0.912	0.921	0.286	0.290
0.5	2	3	0.246	0.244	0.076	0.084
		5	0.481	0.473	0.096	0.103
	5	3	0.468	0.478	0.091	0.093
		5	0.080	0.079	0.047	0.051
	10	3	0.083	0.088	0.041	0.051
		5	0.427	0.438	0.117	0.117
No. of Shaded region			3	3	10	9

Shaded region indicate Type I error within [0.025 – 0.075]

Table L.37

Type I error rates for variable, $p = 20$ and $n = 20$

ε	κ	μ	M_B	M_T	M_{BGMed}	M_{TGMed}
0	1	0	0.051	0.044	0.050	0.048
0.05	2	3	0.069	0.061	0.039	0.042
		5	0.062	0.055	0.039	0.037
	5	3	0.273	0.251	0.071	0.073
		5	0.271	0.267	0.074	0.075
	10	3	0.526	0.512	0.094	0.090
		5	0.512	0.525	0.098	0.089
0.1	2	3	0.089	0.093	0.045	0.039
		5	0.086	0.076	0.043	0.035
	5	3	0.500	0.491	0.113	0.113
		5	0.502	0.500	0.104	0.106
	10	3	0.776	0.783	0.139	0.141
		5	0.785	0.787	0.138	0.139
0.2	2	3	0.186	0.156	0.046	0.037
		5	0.170	0.165	0.044	0.037
	5	3	0.828	0.813	0.208	0.220
		5	0.830	0.825	0.227	0.213
	10	3	0.962	0.966	0.298	0.298
		5	0.966	0.961	0.286	0.287
0.5	2	3	0.187	0.173	0.078	0.080
		5	0.187	0.169	0.078	0.082
	5	3	0.369	0.338	0.104	0.099
		5	0.374	0.358	0.094	0.107
	10	3	0.077	0.064	0.041	0.150
		5	0.170	0.071	0.046	0.147
No. of Shaded region			3	5	11	9

Shaded region indicate Type I error within [0.025 – 0.075]

Table L.38

Type I error rates for variable, $p = 20$ and $n = 25$

ε	κ	μ	M	M_{GMed}	M_B	M_T	M_{BGMed}	M_{TGMed}
0	1	0	0.048	0.047	0.045	0.051	0.055	0.054
0.05	2	3	0.053	0.052	0.065	0.067	0.043	0.047
		5	0.050	0.051	0.065	0.065	0.043	0.047
	5	3	0.100	0.063	0.307	0.295	0.075	0.070
		5	0.090	0.058	0.286	0.296	0.072	0.076
	10	3	0.169	0.080	0.574	0.592	0.078	0.077
		5	0.172	0.080	0.605	0.588	0.093	0.091
0.1	2	3	0.064	0.053	0.085	0.093	0.041	0.041
		5	0.056	0.050	0.092	0.094	0.042	0.046
	5	3	0.153	0.064	0.563	0.569	0.118	0.115
		5	0.158	0.066	0.566	0.568	0.115	0.115
	10	3	0.343	0.115	0.847	0.841	0.151	0.145
		5	0.327	0.109	0.839	0.849	0.146	0.149
0.2	2	3	0.071	0.051	0.175	0.193	0.044	0.047
		5	0.072	0.056	0.193	0.192	0.043	0.045
	5	3	0.301	0.107	0.885	0.897	0.228	0.233
		5	0.301	0.097	0.893	0.890	0.222	0.238
	10	3	0.672	0.194	0.985	0.985	0.287	0.300
		5	0.671	0.192	0.984	0.985	0.300	0.289
0.5	2	3	0.538	0.105	0.834	0.832	0.084	0.082
		5	0.531	0.110	0.833	0.835	0.085	0.083
	5	3	0.088	0.057	0.154	0.146	0.037	0.035
		5	0.087	0.056	0.164	0.147	0.040	0.037
	10	3	0.495	0.088	0.876	0.870	0.120	0.113
		5	0.501	0.102	0.868	0.852	0.116	0.128
No. of Shaded region			7	13	3	3	11	10
Shaded region indicate Type I error within [0.025 – 0.075]								

Shaded region indicate Type I error within [0.025 – 0.075]

Table L.39

Type I error rates for variable, $p = 20$ and $n = 50$

ε	κ	μ	M	M_{GMed}	M_B	M_T	M_{BGMed}	M_{TGMed}
0	1	0	0.048	0.054	0.047	0.048	0.053	0.053
0.05	2	3	0.062	0.056	0.071	0.076	0.042	0.042
		5	0.071	0.051	0.071	0.071	0.040	0.038
	5	3	0.261	0.075	0.446	0.453	0.061	0.066
		5	0.249	0.074	0.446	0.438	0.064	0.060
	10	3	0.603	0.104	0.787	0.798	0.070	0.073
		5	0.601	0.099	0.784	0.801	0.074	0.075
0.1	2	3	0.080	0.055	0.117	0.137	0.035	0.027
		5	0.087	0.056	0.126	0.123	0.031	0.032
	5	3	0.533	0.101	0.808	0.802	0.101	0.086
		5	0.546	0.086	0.807	0.807	0.094	0.102
	10	3	0.904	0.145	0.972	0.971	0.127	0.127
		5	0.902	0.140	0.969	0.967	0.127	0.134
0.2	2	3	0.130	0.059	0.362	0.364	0.035	0.030
		5	0.132	0.062	0.352	0.372	0.030	0.038
	5	3	0.906	0.175	0.992	0.993	0.237	0.227
		5	0.905	0.188	0.991	0.992	0.239	0.237
	10	3	0.997	0.258	1.000	1.000	0.316	0.329
		5	0.998	0.282	1.000	1.000	0.323	0.324
0.5	2	3	0.603	0.104	0.787	0.798	0.070	0.273
		5	0.601	0.099	0.784	0.801	0.074	0.075
	5	3	0.080	0.055	0.117	0.137	0.035	0.327
		5	0.087	0.056	0.126	0.123	0.031	0.032
	10	3	0.533	0.101	0.808	0.802	0.101	0.086
		5	0.546	0.086	0.807	0.807	0.094	0.102
No. of Shaded region			3	11	3	2	15	13
Shaded region indicate Type I error within [0.025 – 0.075]								

Appendix M.

Table M.40

Type I error rates for variable, $p = 25$ and $n = 5$

ε	κ	μ	M_B	M_T	M_{BGMed}	M_{TGMed}
0	1	0	0.062	0.055	0.046	0.053
0.05	2	3	0.054	0.052	0.052	0.044
		5	0.065	0.052	0.048	0.047
	5	3	0.108	0.112	0.060	0.061
		5	0.113	0.105	0.064	0.063
	10	3	0.212	0.209	0.083	0.084
		5	0.206	0.208	0.092	0.086
0.1	2	3	0.060	0.063	0.050	0.049
		5	0.058	0.063	0.052	0.047
	5	3	0.180	0.170	0.070	0.092
		5	0.177	0.167	0.077	0.070
	10	3	0.339	0.327	0.119	0.133
		5	0.350	0.348	0.121	0.122
0.2	2	3	0.077	0.067	0.052	0.146
		5	0.081	0.073	0.046	0.047
	5	3	0.306	0.305	0.104	0.098
		5	0.316	0.302	0.109	0.105
	10	3	0.567	0.565	0.200	0.195
		5	0.556	0.560	0.199	0.199
0.5	2	3	0.212	0.209	0.083	0.084
		5	0.206	0.208	0.092	0.086
	5	3	0.060	0.063	0.050	0.099
		5	0.058	0.063	0.052	0.147
	10	3	0.180	0.170	0.070	0.072
		5	0.177	0.167	0.077	0.170
No. of Shaded region			7	9	13	10
Shaded region indicate Type I error within [0.025 – 0.075]						

Shaded region indicate Type I error within [0.025 – 0.075]

Table M.41

Type I error rates for variable, $p = 25$ and $n = 10$

ε	κ	μ	M_B	M_T	M_{BGMed}	M_{TGMed}
0	1	0	0.055	0.048	0.046	0.053
0.05	2	3	0.063	0.050	0.051	0.051
		5	0.062	0.053	0.046	0.053
	5	3	0.187	0.173	0.078	0.080
		5	0.187	0.169	0.078	0.082
	10	3	0.369	0.338	0.104	0.099
		5	0.374	0.358	0.094	0.107
0.1	2	3	0.077	0.064	0.041	0.050
		5	0.070	0.071	0.046	0.047
	5	3	0.330	0.318	0.107	0.117
		5	0.333	0.317	0.119	0.123
	10	3	0.576	0.578	0.148	0.159
		5	0.578	0.570	0.148	0.148
0.2	2	3	0.118	0.101	0.051	0.153
		5	0.109	0.102	0.051	0.254
	5	3	0.596	0.582	0.200	0.209
		5	0.600	0.581	0.199	0.208
	10	3	0.842	0.831	0.289	0.273
		5	0.839	0.821	0.277	0.278
0.5	2	3	0.369	0.338	0.104	0.099
		5	0.374	0.358	0.094	0.107
	5	3	0.077	0.064	0.041	0.050
		5	0.090	0.071	0.046	0.047
	10	3	0.330	0.318	0.107	0.117
		5	0.333	0.317	0.119	0.123
No. of Shaded region			4	7	9	7

Shaded region indicate Type I error within [0.025 – 0.075]

Table M.42

Type I error rates for variable, $p = 25$ and $n = 15$

ε	κ	μ	M_B	M_T	M_{BGMed}	M_{TGMed}
0	1	0	0.047	0.046	0.048	0.050
0.05	2	3	0.064	0.066	0.039	0.052
		5	0.066	0.069	0.041	0.052
	5	3	0.228	0.241	0.073	0.082
		5	0.246	0.244	0.076	0.084
	10	3	0.481	0.473	0.096	0.103
		5	0.468	0.478	0.091	0.093
0.1	2	3	0.080	0.079	0.047	0.051
		5	0.083	0.088	0.041	0.051
	5	3	0.427	0.438	0.117	0.117
		5	0.439	0.451	0.114	0.122
	10	3	0.714	0.721	0.139	0.157
		5	0.713	0.720	0.136	0.159
0.2	2	3	0.149	0.148	0.046	0.257
		5	0.152	0.150	0.047	0.349
	5	3	0.755	0.767	0.231	0.249
		5	0.765	0.763	0.218	0.214
	10	3	0.924	0.933	0.298	0.289
		5	0.937	0.939	0.287	0.295
0.5	2	3	0.481	0.473	0.096	0.103
		5	0.468	0.478	0.091	0.093
	5	3	0.080	0.079	0.047	0.051
		5	0.083	0.088	0.041	0.051
	10	3	0.427	0.438	0.117	0.117
		5	0.439	0.451	0.114	0.122
No. of Shaded region			3	3	10	7

Shaded region indicate Type I error within [0.025 – 0.075]

Table M.43

Type I error rates for variable, $p = 25$ and $n = 20$

ϵ	κ	μ	M_B	M_T	M_{BGMed}	M_{TGMed}
0	1	0	0.055	0.054	0.052	0.052
0.05	2	3	0.070	0.069	0.048	0.050
		5	0.076	0.067	0.049	0.043
	5	3	0.300	0.284	0.078	0.079
		5	0.291	0.285	0.086	0.082
	10	3	0.551	0.554	0.101	0.091
		5	0.560	0.554	0.100	0.096
0.1	2	3	0.098	0.100	0.048	0.043
		5	0.106	0.089	0.043	0.042
	5	3	0.546	0.541	0.118	0.122
		5	0.550	0.551	0.133	0.108
	10	3	0.810	0.809	0.156	0.151
		5	0.808	0.817	0.150	0.157
0.2	2	3	0.198	0.191	0.056	0.047
		5	0.195	0.176	0.057	0.049
	5	3	0.867	0.864	0.250	0.232
		5	0.863	0.862	0.240	0.251
	10	3	0.970	0.973	0.317	0.303
		5	0.976	0.970	0.314	0.322
0.5	2	3	0.551	0.554	0.101	0.091
		5	0.560	0.554	0.100	0.096
	5	3	0.098	0.100	0.048	0.143
		5	0.106	0.089	0.043	0.242
	10	3	0.546	0.541	0.118	0.122
		5	0.550	0.551	0.133	0.108
No. of Shaded region			2	3	9	7

Shaded region indicate Type I error within [0.025 – 0.075]

Table M.44

Type I error rates for variable, $p = 25$ and $n = 25$

ε	κ	μ	M_B	M_T	M_{BGMed}	M_{TGMed}
0	1	0	0.056	0.051	0.049	0.048
0.05	2	3	0.074	0.066	0.043	0.049
		5	0.078	0.070	0.043	0.045
	5	3	0.345	0.333	0.076	0.085
		5	0.353	0.333	0.073	0.082
	10	3	0.616	0.617	0.096	0.097
		5	0.633	0.617	0.094	0.099
0.1	2	3	0.113	0.104	0.041	0.041
		5	0.109	0.105	0.038	0.042
	5	3	0.621	0.612	0.119	0.127
		5	0.628	0.617	0.127	0.130
	10	3	0.873	0.867	0.146	0.150
		5	0.872	0.870	0.153	0.157
0.2	2	3	0.237	0.232	0.047	0.052
		5	0.238	0.215	0.045	0.059
	5	3	0.918	0.920	0.251	0.259
		5	0.919	0.916	0.241	0.259
	10	3	0.987	0.987	0.322	0.325
		5	0.989	0.988	0.315	0.337
0.5	2	3	0.345	0.333	0.076	0.085
		5	0.353	0.333	0.073	0.082
	5	3	0.616	0.617	0.096	0.097
		5	0.633	0.617	0.094	0.099
	10	3	0.113	0.104	0.041	0.041
		5	0.109	0.105	0.038	0.042
No. of Shaded region			2	3	11	9

Shaded region indicate Type I error within [0.025 – 0.075]

Table M.45

Type I error rates for variable, $p = 25$ and $n = 50$

ε	κ	μ	M	M_{GMed}	M_B	M_T	M_{BGMed}	M_{TGMed}
0	1	0	0.046	0.050	0.040	0.050	0.045	0.051
0.05	2	3	0.065	0.059	0.077	0.081	0.035	0.035
		5	0.060	0.056	0.082	0.091	0.031	0.038
	5	3	0.264	0.076	0.495	0.508	0.064	0.067
		5	0.235	0.078	0.482	0.491	0.069	0.073
	10	3	0.564	0.104	0.819	0.826	0.068	0.062
		5	0.577	0.097	0.811	0.824	0.061	0.075
0.1	2	3	0.082	0.060	0.136	0.145	0.025	0.036
		5	0.090	0.057	0.135	0.147	0.031	0.030
	5	3	0.521	0.103	0.846	0.839	0.108	0.104
		5	0.528	0.104	0.836	0.850	0.104	0.108
	10	3	0.888	0.154	0.976	0.975	0.128	0.142
		5	0.878	0.157	0.976	0.980	0.137	0.144
0.2	2	3	0.131	0.063	0.406	0.425	0.036	0.040
		5	0.136	0.068	0.394	0.423	0.035	0.042
	5	3	0.895	0.171	0.994	0.995	0.269	0.282
		5	0.888	0.182	0.995	0.993	0.276	0.281
	10	3	0.997	0.285	1.000	1.000	0.348	0.366
		5	0.996	0.295	1.000	1.000	0.352	0.362
0.5	2	3	0.564	0.104	0.819	0.826	0.068	0.092
		5	0.577	0.097	0.811	0.824	0.061	0.085
	5	3	0.082	0.060	0.136	0.145	0.025	0.036
		5	0.090	0.057	0.135	0.147	0.031	0.030
	10	3	0.521	0.103	0.846	0.839	0.108	0.104
		5	0.528	0.104	0.836	0.850	0.104	0.108
No. of Shaded region			3	9	1	1	15	13
Shaded region indicate Type I error within [0.025 – 0.075]								

Shaded region indicate Type I error within [0.025 – 0.075]

Table M.46

Type I error rates for variable, $p = 30$ and $n = 5$

ϵ	κ	μ	M_B	M_T	M_{BGMed}	M_{TGMed}
0	1	0	0.048	0.052	0.054	0.056
0.05	2	3	0.050	0.060	0.052	0.050
		5	0.047	0.058	0.053	0.049
	5	3	0.106	0.113	0.066	0.062
		5	0.107	0.127	0.060	0.056
	10	3	0.202	0.219	0.091	0.089
		5	0.220	0.225	0.091	0.082
0.1	2	3	0.057	0.068	0.053	0.052
		5	0.056	0.066	0.057	0.047
	5	3	0.168	0.194	0.075	0.073
		5	0.179	0.187	0.075	0.074
	10	3	0.352	0.357	0.131	0.125
		5	0.365	0.360	0.129	0.132
0.2	2	3	0.077	0.073	0.053	0.052
		5	0.074	0.080	0.056	0.046
	5	3	0.328	0.331	0.116	0.111
		5	0.324	0.343	0.110	0.123
	10	3	0.583	0.600	0.208	0.210
		5	0.589	0.586	0.220	0.213
0.5	2	3	0.202	0.219	0.091	0.089
		5	0.220	0.225	0.091	0.082
	5	3	0.057	0.068	0.053	0.052
		5	0.056	0.066	0.057	0.047
	10	3	0.168	0.194	0.075	0.073
		5	0.179	0.187	0.075	0.074
No. of Shaded region			8	8	15	15

Shaded region indicate Type I error within [0.025 – 0.075]

Table M.47

Type I error rates for variable, $p = 30$ and $n = 10$

ε	κ	μ	M_B	M_T	M_{BGMed}	M_{TGMed}
0	1	0	0.044	0.053	0.050	0.059
0.05	2	3	0.053	0.059	0.049	0.052
		5	0.056	0.055	0.048	0.054
	5	3	0.193	0.196	0.086	0.092
		5	0.175	0.179	0.080	0.086
	10	3	0.380	0.358	0.099	0.107
		5	0.373	0.356	0.098	0.112
0.1	2	3	0.068	0.070	0.051	0.057
		5	0.069	0.075	0.047	0.055
	5	3	0.336	0.360	0.117	0.125
		5	0.343	0.340	0.117	0.126
	10	3	0.584	0.603	0.167	0.156
		5	0.610	0.583	0.155	0.161
0.2	2	3	0.104	0.109	0.052	0.048
		5	0.100	0.111	0.054	0.051
	5	3	0.617	0.616	0.214	0.230
		5	0.616	0.612	0.211	0.224
	10	3	0.846	0.841	0.287	0.294
		5	0.850	0.854	0.286	0.282
0.5	2	3	0.175	0.179	0.080	0.086
		5	0.380	0.358	0.099	0.107
	5	3	0.373	0.356	0.098	0.112
		5	0.068	0.070	0.051	0.057
	10	3	0.069	0.075	0.047	0.055
		5	0.336	0.360	0.117	0.125
No. of Shaded region			7	7	9	9

Shaded region indicate Type I error within $[0.025 - 0.075]$

Table M.48

Type I error rates for variable, $p = 30$ and $n = 15$

ε	κ	μ	M_B	M_T	M_{BGMed}	M_{TGMed}
0	1	0	0.062	0.055	0.053	0.051
0.05	2	3	0.077	0.071	0.053	0.046
		5	0.074	0.071	0.052	0.047
	5	3	0.262	0.254	0.073	0.070
		5	0.268	0.268	0.091	0.084
	10	3	0.488	0.500	0.098	0.102
		5	0.494	0.501	0.098	0.096
0.1	2	3	0.090	0.086	0.050	0.050
		5	0.097	0.093	0.047	0.045
	5	3	0.483	0.473	0.128	0.126
		5	0.494	0.481	0.128	0.130
	10	3	0.744	0.740	0.157	0.157
		5	0.752	0.745	0.158	0.154
0.2	2	3	0.171	0.161	0.056	0.059
		5	0.173	0.158	0.061	0.059
	5	3	0.801	0.798	0.258	0.257
		5	0.791	0.789	0.253	0.249
	10	3	0.942	0.941	0.317	0.311
		5	0.942	0.940	0.305	0.323
0.5	2	3	0.268	0.268	0.091	0.084
		5	0.488	0.500	0.098	0.102
	5	3	0.494	0.501	0.098	0.096
		5	0.090	0.086	0.050	0.150
	10	3	0.097	0.093	0.047	0.145
		5	0.483	0.473	0.128	0.126
No. of Shaded region			2	3	10	8

Shaded region indicate Type I error within [0.025 – 0.075]

Table M.49

Type I error rates for variable, $p = 30$ and $n = 20$

ε	κ	μ	M_B	M_T	M_{BGMed}	M_{TGMed}
0	1	0	0.044	0.046	0.046	0.044
0.05	2	3	0.068	0.057	0.048	0.042
		5	0.063	0.060	0.048	0.042
	5	3	0.293	0.306	0.083	0.082
		5	0.300	0.307	0.087	0.079
	10	3	0.559	0.575	0.094	0.081
		5	0.574	0.577	0.092	0.087
0.1	2	3	0.087	0.096	0.044	0.038
		5	0.095	0.095	0.044	0.040
	5	3	0.570	0.559	0.126	0.123
		5	0.566	0.556	0.123	0.123
	10	3	0.816	0.826	0.156	0.157
		5	0.824	0.828	0.164	0.155
0.2	2	3	0.189	0.193	0.055	0.050
		5	0.190	0.189	0.056	0.052
	5	3	0.875	0.876	0.270	0.270
		5	0.872	0.875	0.270	0.258
	10	3	0.972	0.977	0.340	0.307
		5	0.977	0.976	0.321	0.321
0.5	2	3	0.559	0.575	0.094	0.081
		5	0.574	0.577	0.092	0.087
	5	3	0.087	0.096	0.044	0.038
		5	0.095	0.095	0.044	0.140
	10	3	0.570	0.559	0.126	0.123
		5	0.566	0.556	0.123	0.123
No. of Shaded region			3	3	9	8

Shaded region indicate Type I error within [0.025 – 0.075]

Table M.50

Type I error rates for variable, $p = 30$ and $n = 25$

ε	κ	μ	M_B	M_T	M_{BGMed}	M_{TGMed}
0	1	0	0.053	0.055	0.047	0.047
0.05	2	3	0.068	0.070	0.041	0.046
		5	0.075	0.072	0.047	0.045
	5	3	0.352	0.347	0.070	0.077
		5	0.355	0.346	0.075	0.075
	10	3	0.640	0.641	0.092	0.095
		5	0.652	0.648	0.091	0.102
0.1	2	3	0.107	0.101	0.040	0.044
		5	0.113	0.104	0.046	0.043
	5	3	0.635	0.647	0.129	0.139
		5	0.640	0.641	0.124	0.118
	10	3	0.889	0.876	0.156	0.151
		5	0.885	0.882	0.146	0.145
0.2	2	3	0.240	0.229	0.053	0.047
		5	0.231	0.240	0.050	0.054
	5	3	0.934	0.929	0.269	0.287
		5	0.933	0.931	0.269	0.281
	10	3	0.989	0.989	0.343	0.327
		5	0.991	0.992	0.328	0.350
0.5	2	3	0.640	0.641	0.092	0.095
		5	0.652	0.648	0.091	0.102
	5	3	0.107	0.101	0.040	0.044
		5	0.113	0.104	0.046	0.043
	10	3	0.635	0.647	0.129	0.139
		5	0.640	0.641	0.124	0.118
No. of Shaded region			3	3	11	10

Shaded region indicate Type I error within [0.025 – 0.075]

Table M.51

Type I error rates for variable, $p = 30$ and $n = 50$

ε	κ	μ	M	M_{GMed}	M_B	M_T	M_{BGMed}	M_{TGMed}
0	1	0	0.050	0.049	0.052	0.043	0.051	0.050
0.05	2	3	0.063	0.057	0.081	0.079	0.042	0.040
		5	0.064	0.056	0.079	0.075	0.038	0.038
	5	3	0.237	0.072	0.534	0.503	0.073	0.072
		5	0.238	0.074	0.533	0.510	0.074	0.068
	10	3	0.538	0.105	0.834	0.832	0.084	0.082
		5	0.531	0.110	0.833	0.835	0.085	0.083
0.1	2	3	0.088	0.057	0.154	0.146	0.037	0.035
		5	0.087	0.056	0.164	0.147	0.040	0.037
	5	3	0.495	0.088	0.876	0.870	0.120	0.113
		5	0.501	0.102	0.868	0.852	0.116	0.128
	10	3	0.861	0.152	0.981	0.980	0.149	0.141
		5	0.860	0.150	0.981	0.976	0.147	0.147
0.2	2	3	0.127	0.067	0.442	0.439	0.053	0.052
		5	0.134	0.068	0.458	0.430	0.050	0.054
	5	3	0.883	0.174	0.995	0.996	0.321	0.308
		5	0.873	0.165	0.995	1.000	0.322	0.312
	10	3	0.995	0.279	1.000	1.000	0.410	0.402
		5	0.996	0.293	1.000	1.000	0.408	0.408
0.5	2	3	0.538	0.105	0.834	0.832	0.084	0.082
		5	0.531	0.110	0.833	0.835	0.085	0.083
	5	3	0.088	0.157	0.154	0.146	0.037	0.095
		5	0.087	0.156	0.164	0.147	0.040	0.137
	10	3	0.495	0.088	0.876	0.870	0.120	0.113
		5	0.501	0.102	0.868	0.852	0.116	0.128
No. of Shaded region			3	9	1	2	11	9
Shaded region indicate Type I error within [0.025 – 0.075]								

Shaded region indicate Type I error within [0.025 – 0.075]

Table M.52

Type I error rates for variable, $p = 50$ and $n = 5$

ε	κ	μ	M_B	M_T	M_{BGMed}	M_{TGMed}
0	1	0	0.054	0.051	0.053	0.051
0.05	2	3	0.061	0.051	0.047	0.051
		5	0.056	0.055	0.046	0.055
	5	3	0.141	0.145	0.066	0.145
		5	0.135	0.125	0.066	0.125
	10	3	0.246	0.244	0.090	0.244
		5	0.240	0.247	0.090	0.247
0.1	2	3	0.064	0.068	0.048	0.068
		5	0.070	0.064	0.049	0.064
	5	3	0.231	0.237	0.081	0.237
		5	0.229	0.238	0.089	0.238
	10	3	0.414	0.416	0.145	0.416
		5	0.418	0.399	0.136	0.399
0.2	2	3	0.080	0.085	0.050	0.085
		5	0.087	0.084	0.049	0.084
	5	3	0.404	0.400	0.140	0.400
		5	0.413	0.410	0.135	0.410
	10	3	0.643	0.656	0.250	0.656
		5	0.147	0.653	0.002	0.653
0.5	2	3	0.246	0.244	0.090	0.244
		5	0.240	0.247	0.090	0.247
	5	3	0.164	0.268	0.048	0.068
		5	0.170	0.264	0.049	0.064
	10	3	0.231	0.237	0.081	0.237
		5	0.229	0.238	0.089	0.238
No. of Shaded region			5	5	11	7

Shaded region indicate Type I error within [0.025 – 0.075]

Table M.53

Type I error rates for variable, $p = 50$ and $n = 10$

ϵ	κ	μ	M_B	M_T	M_{BGMed}	M_{TGMed}
0	1	0	0.053	0.050	0.049	0.058
0.05	2	3	0.064	0.064	0.048	0.050
		5	0.061	0.063	0.053	0.052
	5	3	0.250	0.240	0.083	0.080
		5	0.247	0.254	0.089	0.068
	10	3	0.415	0.419	0.101	0.097
		5	0.424	0.410	0.101	0.099
0.1	2	3	0.096	0.083	0.058	0.098
		5	0.083	0.084	0.044	0.148
	5	3	0.421	0.435	0.131	0.097
		5	0.427	0.423	0.143	0.089
	10	3	0.651	0.655	0.166	0.150
		5	0.659	0.651	0.158	0.143
0.2	2	3	0.140	0.150	0.050	0.054
		5	0.146	0.139	0.055	0.050
	5	3	0.716	0.717	0.259	0.136
		5	0.714	0.725	0.246	0.143
	10	3	0.887	0.887	0.297	0.242
		5	0.317	0.888	0.093	0.257
0.5	2	3	0.415	0.419	0.101	0.097
		5	0.424	0.410	0.101	0.099
	5	3	0.096	0.083	0.058	0.048
		5	0.083	0.084	0.044	0.048
	10	3	0.421	0.435	0.131	0.097
		5	0.427	0.423	0.143	0.089
No. of Shaded region			3	3	9	8

Shaded region indicate Type I error within [0.025 – 0.075]

Table M.54

Type I error rates for variable, $p = 50$ and $n = 15$

ε	κ	μ	M_B	M_T	M_{BGMed}	M_{TGMed}
0	1	0	0.044	0.055	0.053	0.046
0.05	2	3	0.064	0.064	0.053	0.047
		5	0.064	0.068	0.049	0.041
	5	3	0.313	0.319	0.091	0.090
		5	0.309	0.325	0.095	0.089
	10	3	0.533	0.539	0.101	0.093
		5	0.530	0.528	0.102	0.100
0.1	2	3	0.094	0.096	0.051	0.046
		5	0.095	0.104	0.051	0.046
	5	3	0.549	0.551	0.155	0.146
		5	0.563	0.554	0.144	0.138
	10	3	0.784	0.794	0.171	0.173
		5	0.780	0.779	0.162	0.171
0.2	2	3	0.193	0.206	0.058	0.057
		5	0.187	0.201	0.062	0.056
	5	3	0.852	0.850	0.290	0.294
		5	0.850	0.857	0.291	0.275
	10	3	0.960	0.956	0.342	0.336
		5	0.000	0.958	0.284	0.341
0.5	2	3	0.309	0.325	0.095	0.089
		5	0.533	0.539	0.101	0.093
	5	3	0.530	0.528	0.102	0.100
		5	0.094	0.096	0.051	0.046
	10	3	0.095	0.104	0.051	0.046
		5	0.549	0.551	0.155	0.146
No. of Shaded region			3	3	9	9
Shaded region indicate Type I error within [0.025 – 0.075]						

Table M.55

Type I error rates for variable, $p = 50$ and $n = 20$

ε	κ	μ	M_B	M_T	M_{BGMed}	M_{TGMed}
0	1	0	0.050	0.054	0.049	0.050
0.05	2	3	0.082	0.073	0.051	0.043
		5	0.080	0.076	0.050	0.048
	5	3	0.382	0.382	0.096	0.081
		5	0.408	0.388	0.095	0.096
	10	3	0.643	0.639	0.110	0.102
		5	0.630	0.628	0.103	0.098
0.1	2	3	0.130	0.119	0.051	0.050
		5	0.123	0.119	0.054	0.046
	5	3	0.670	0.651	0.148	0.142
		5	0.671	0.650	0.160	0.152
	10	3	0.863	0.867	0.174	0.170
		5	0.869	0.870	0.162	0.151
0.2	2	3	0.277	0.266	0.069	0.066
		5	0.260	0.269	0.071	0.069
	5	3	0.925	0.926	0.322	0.318
		5	0.921	0.925	0.320	0.301
	10	3	0.984	0.986	0.379	0.366
		5	0.927	0.985	0.205	0.359
0.5	2	3	0.408	0.388	0.095	0.096
		5	0.643	0.639	0.110	0.102
	5	3	0.630	0.628	0.103	0.098
		5	0.130	0.119	0.051	0.050
	10	3	0.123	0.119	0.054	0.046
		5	0.670	0.651	0.148	0.142
No. of Shaded region			1	2	9	9

Shaded region indicate Type I error within [0.025 – 0.075]

Table M.56

Type I error rates for variable, $p = 50$ and $n = 25$

ε	κ	μ	M_B	M_T	M_{BGMed}	M_{TGMed}
0	1	0	0.049	0.048	0.049	0.045
0.05	2	3	0.072	0.081	0.052	0.047
		5	0.072	0.084	0.055	0.045
	5	3	0.429	0.432	0.095	0.089
		5	0.415	0.431	0.095	0.096
	10	3	0.693	0.694	0.102	0.089
		5	0.708	0.687	0.104	0.094
0.1	2	3	0.127	0.129	0.056	0.047
		5	0.127	0.131	0.052	0.045
	5	3	0.726	0.734	0.153	0.150
		5	0.725	0.723	0.159	0.146
	10	3	0.911	0.912	0.186	0.163
		5	0.912	0.912	0.180	0.168
0.2	2	3	0.315	0.313	0.062	0.066
		5	0.310	0.311	0.070	0.064
	5	3	0.959	0.962	0.343	0.337
		5	0.960	0.962	0.359	0.337
	10	3	0.995	0.996	0.408	0.403
		5	1.000	0.997	0.431	0.404
0.5	2	3	0.693	0.694	0.102	0.089
		5	0.708	0.687	0.104	0.094
	5	3	0.127	0.129	0.056	0.147
		5	0.127	0.131	0.052	0.145
	10	3	0.726	0.734	0.153	0.150
		5	0.725	0.723	0.159	0.146
No. of Shaded region			3	1	9	7
Shaded region indicate Type I error within [0.025 – 0.075]						

Table M.57

Type I error rates for variable, $p = 50$ and $n = 50$

ε	κ	μ	M_B	M_T	M_{BGMed}	M_{TGMed}
0	1	0	0.059	0.052	0.048	0.050
0.05	2	3	0.106	0.090	0.043	0.039
		5	0.097	0.091	0.039	0.037
	5	3	0.642	0.611	0.075	0.071
		5	0.630	0.621	0.071	0.088
	10	3	0.887	0.887	0.095	0.085
		5	0.883	0.875	0.071	0.076
0.1	2	3	0.205	0.202	0.039	0.043
		5	0.211	0.205	0.039	0.046
	5	3	0.923	0.918	0.157	0.150
		5	0.918	0.917	0.153	0.155
	10	3	0.990	0.988	0.185	0.175
		5	0.989	0.990	0.179	0.182
0.2	2	3	0.593	0.584	0.084	0.086
		5	0.606	0.592	0.089	0.093
	5	3	0.999	0.998	0.448	0.447
		5	0.998	1.000	0.439	0.440
	10	3	1.000	1.000	0.543	0.537
		5	1.000	1.000	0.684	0.534
0.5	2	3	0.883	0.875	0.071	0.076
		5	0.205	0.202	0.039	0.043
	5	3	0.211	0.205	0.039	0.046
		5	0.923	0.918	0.157	0.150
	10	3	0.918	0.917	0.153	0.155
		5	0.990	0.988	0.185	0.175
No. of Shaded region			1	1	11	8
Shaded region indicate Type I error within [0.025 – 0.075]						

Appendix N.

Table N.58

Power of Test for variable, $p = 3$, $n = 2$

ρ	σ	M_B	M_T	M_{BGMed}	M_{TGMed}
0	1	0.052	0.050	0.053	0.053
	1.2	0.051	0.051	0.051	0.051
	1.4	0.049	0.050	0.048	0.048
	1.6	0.054	0.053	0.051	0.051
	1.8	0.440	0.054	0.042	0.042
	2	0.052	0.055	0.046	0.046
	3	0.071	0.062	0.053	0.053
0.2	1	0.050	0.049	0.051	0.051
	1.2	0.050	0.049	0.043	0.046
	1.4	0.050	0.048	0.047	0.047
	1.6	0.057	0.052	0.052	0.052
	1.8	0.056	0.053	0.051	0.051
	2	0.055	0.055	0.046	0.046
	3	0.053	0.063	0.047	0.047
0.5	1	0.050	0.043	0.046	0.046
	1.2	0.050	0.044	0.048	0.048
	1.4	0.052	0.048	0.048	0.044
	1.6	0.052	0.046	0.045	0.045
	1.8	0.051	0.051	0.042	0.042
	2	0.051	0.052	0.039	0.039
	3	0.068	0.062	0.047	0.047
0.7	1	0.049	0.045	0.044	0.044
	1.2	0.049	0.048	0.043	0.043
	1.4	0.051	0.050	0.040	0.040
	1.6	0.051	0.049	0.041	0.041
	1.8	0.059	0.052	0.049	0.048
	2	0.062	0.053	0.043	0.043
	3	0.065	0.062	0.046	0.046
0.9	1	0.054	0.054	0.033	0.033
	1.2	0.055	0.053	0.038	0.038
	1.4	0.055	0.052	0.037	0.037
	1.6	0.049	0.051	0.034	0.034
	1.8	0.056	0.055	0.033	0.033
	2	0.058	0.054	0.031	0.031
	3	0.065	0.064	0.036	0.033
No of Shaded region		0	0	0	0

Shaded region indicate Power of test values within [0.5-1.0]

Table N.59

Power of Test for variable, $p = 3, n = 5$

ρ	σ	M	M_{GMed}	M_B	M_T	M_{BGMed}	M_{TGMed}
0	1	0.045	0.051	0.054	0.047	0.049	0.051
	1.2	0.047	0.057	0.049	0.049	0.053	0.054
	1.4	0.067	0.069	0.053	0.058	0.052	0.045
	1.6	0.099	0.113	0.055	0.056	0.051	0.051
	1.8	0.270	0.274	0.068	0.064	0.055	0.052
	2	0.047	0.053	0.058	0.051	0.043	0.041
	3	0.048	0.052	0.059	0.054	0.042	0.044
0.2	1	0.064	0.067	0.057	0.054	0.040	0.044
	1.2	0.099	0.096	0.059	0.060	0.044	0.049
	1.4	0.252	0.254	0.066	0.067	0.048	0.052
	1.6	0.046	0.047	0.062	0.062	0.042	0.042
	1.8	0.050	0.056	0.066	0.059	0.041	0.039
	2	0.058	0.068	0.066	0.069	0.042	0.043
	3	0.092	0.097	0.064	0.066	0.041	0.039
0.5	1	0.233	0.245	0.077	0.079	0.042	0.044
	1.2	0.048	0.048	0.077	0.086	0.045	0.050
	1.4	0.059	0.055	0.082	0.081	0.042	0.044
	1.6	0.061	0.063	0.085	0.085	0.042	0.049
	1.8	0.092	0.087	0.090	0.086	0.038	0.038
	2	0.212	0.231	0.098	0.088	0.044	0.042
	3	0.049	0.053	0.097	0.087	0.046	0.046
0.7	1	0.052	0.058	0.092	0.091	0.051	0.044
	1.2	0.069	0.071	0.100	0.099	0.046	0.044
	1.4	0.087	0.095	0.094	0.107	0.052	0.040
	1.6	0.217	0.228	0.116	0.109	0.043	0.043
	1.8	0.053	0.049	0.111	0.106	0.060	0.047
	2	0.047	0.050	0.114	0.121	0.051	0.052
	3	0.063	0.065	0.117	0.119	0.050	0.054
0.9	1	0.092	0.098	0.127	0.121	0.049	0.047
	1.2	0.204	0.224	0.151	0.127	0.046	0.045
	1.4	0.071	0.059	0.235	0.229	0.082	0.077
	1.6	0.068	0.061	0.248	0.231	0.074	0.077
	1.8	0.081	0.077	0.239	0.246	0.067	0.065
	2	0.094	0.103	0.254	0.245	0.064	0.063
	3	0.193	0.219	0.257	0.257	0.064	0.062
No of Shaded region		0	0	0	0	0	0

Shaded region indicate Power of test values within [0.5-1.0]

Table N.60

Power of Test for variable, $p = 3, n = 10$

ρ	σ	M	M_{GMed}	M_B	M_T	M_{BGMed}	M_{TGMed}
0	1	0.042	0.048	0.048	0.058	0.047	0.043
	1.2	0.062	0.069	0.052	0.061	0.052	0.052
	1.4	0.154	0.165	0.053	0.053	0.063	0.057
	1.6	0.395	0.420	0.053	0.059	0.072	0.073
	1.8	0.943	0.936	0.066	0.072	0.101	0.099
	2	0.059	0.045	0.062	0.065	0.037	0.037
	3	0.065	0.055	0.066	0.062	0.042	0.037
0.2	1	0.145	0.146	0.067	0.072	0.045	0.041
	1.2	0.349	0.358	0.070	0.077	0.061	0.051
	1.4	0.915	0.893	0.075	0.082	0.079	0.069
	1.6	0.061	0.053	0.096	0.093	0.042	0.037
	1.8	0.076	0.061	0.089	0.099	0.039	0.037
	2	0.151	0.133	0.099	0.103	0.042	0.040
	3	0.334	0.322	0.098	0.104	0.050	0.042
0.5	1	0.881	0.863	0.105	0.110	0.077	0.072
	1.2	0.080	0.052	0.130	0.136	0.052	0.043
	1.4	0.086	0.065	0.128	0.142	0.050	0.049
	1.6	0.156	0.133	0.139	0.151	0.055	0.046
	1.8	0.318	0.290	0.150	0.153	0.054	0.056
	2	0.854	0.833	0.160	0.160	0.077	0.072
	3	0.102	0.066	0.190	0.221	0.058	0.057
0.7	1	0.114	0.073	0.187	0.200	0.061	0.052
	1.2	0.164	0.135	0.196	0.201	0.067	0.059
	1.4	0.311	0.277	0.200	0.210	0.071	0.067
	1.6	0.833	0.788	0.211	0.220	0.084	0.076
	1.8	0.131	0.067	0.260	0.267	0.080	0.081
	2	0.132	0.078	0.249	0.271	0.078	0.074
	3	0.191	0.145	0.266	0.268	0.084	0.075
0.9	1	0.321	0.278	0.266	0.266	0.081	0.082
	1.2	0.819	0.776	0.266	0.267	0.084	0.083
	1.4	0.332	0.140	0.601	0.610	0.216	0.205
	1.6	0.325	0.141	0.581	0.619	0.204	0.197
	1.8	0.323	0.192	0.574	0.609	0.199	0.175
	2	0.408	0.278	0.567	0.568	0.173	0.165
	3	0.765	0.688	0.533	0.552	0.165	0.165
No of Shaded region		7	7	5	5	0	0

Shaded region indicate Power of test values within [0.5-1.0]

Table N.61

Power of Test for variable, $p = 3, n = 15$

ρ	σ	M	M_{GMed}	M_B	M_T	M_{BGMed}	M_{TGMed}
0	1	0.055	0.050	0.052	0.049	0.051	0.047
	1.2	0.078	0.064	0.048	0.055	0.051	0.052
	1.4	0.286	0.263	0.050	0.058	0.068	0.068
	1.6	0.700	0.630	0.057	0.064	0.098	0.089
	1.8	0.999	0.996	0.060	0.066	0.130	0.135
	2	0.063	0.037	0.067	0.073	0.033	0.035
	3	0.090	0.057	0.064	0.072	0.032	0.033
0.2	1	0.263	0.201	0.063	0.076	0.044	0.038
	1.2	0.630	0.554	0.069	0.077	0.056	0.059
	1.4	0.996	0.992	0.079	0.086	0.096	0.095
	1.6	0.093	0.042	0.115	0.127	0.031	0.033
	1.8	0.109	0.056	0.111	0.115	0.029	0.033
	2	0.256	0.179	0.117	0.126	0.035	0.041
	3	0.578	0.474	0.132	0.129	0.054	0.050
0.5	1	0.994	0.984	0.130	0.139	0.075	0.081
	1.2	0.130	0.044	0.188	0.199	0.038	0.044
	1.4	0.146	0.058	0.181	0.202	0.040	0.038
	1.6	0.289	0.177	0.197	0.201	0.044	0.045
	1.8	0.577	0.452	0.199	0.208	0.055	0.061
	2	0.990	0.970	0.208	0.204	0.078	0.079
	3	0.198	0.063	0.275	0.298	0.059	0.063
0.7	1	0.200	0.075	0.302	0.298	0.059	0.059
	1.2	0.317	0.172	0.284	0.299	0.065	0.065
	1.4	0.579	0.415	0.288	0.297	0.061	0.068
	1.6	0.985	0.956	0.295	0.292	0.100	0.090
	1.8	0.268	0.081	0.402	0.411	0.091	0.088
	2	0.269	0.091	0.389	0.397	0.082	0.086
	3	0.354	0.173	0.396	0.400	0.087	0.088
0.9	1	0.570	0.389	0.385	0.393	0.090	0.094
	1.2	0.982	0.944	0.371	0.399	0.103	0.102
	1.4	0.654	0.217	0.821	0.827	0.287	0.305
	1.6	0.649	0.228	0.818	0.826	0.282	0.299
	1.8	0.627	0.274	0.782	0.802	0.262	0.269
	2	0.701	0.389	0.767	0.769	0.230	0.238
	3	0.964	0.880	0.729	0.750	0.223	0.222
No of Shaded region		17	9	5	5	0	0

Shaded region indicate Power of test values within [0.5-1.0]

Table N.62

Power of Test for variable, $p = 3, n = 20$

ρ	σ	M	M_{GMed}	M_B	M_T	M_{BGMed}	M_{TGMed}
0	1	0.046	0.047	0.052	0.048	0.054	0.048
	1.2	0.084	0.080	0.051	0.048	0.060	0.047
	1.4	0.401	0.335	0.057	0.049	0.082	0.072
	1.6	0.851	0.787	0.059	0.052	0.113	0.105
	1.8	1	1.000	0.074	0.056	0.158	0.163
	2	0.062	0.031	0.081	0.073	0.806	0.804
	3	0.093	0.056	0.079	0.069	0.028	0.032
0.2	1	0.362	0.255	0.082	0.084	0.036	0.037
	1.2	0.801	0.698	0.086	0.077	0.063	0.063
	1.4	1.000	0.999	0.099	0.087	0.111	0.112
	1.6	0.099	0.028	0.153	0.148	0.022	0.015
	1.8	0.125	0.047	0.151	0.144	0.019	0.020
	2	0.355	0.207	0.162	0.151	0.034	0.023
	3	0.760	0.608	0.166	0.160	0.055	0.042
0.5	1	1.000	0.998	0.164	0.174	0.080	0.081
	1.2	0.167	0.045	0.270	0.248	0.032	0.028
	1.4	0.197	0.057	0.271	0.253	0.032	0.032
	1.6	0.380	0.199	0.275	0.264	0.042	0.039
	1.8	0.743	0.554	0.268	0.272	0.050	0.050
	2	0.999	0.994	0.283	0.262	0.079	0.077
	3	0.250	0.055	0.401	0.384	0.047	0.050
0.7	1	0.280	0.066	0.406	0.404	0.053	0.048
	1.2	0.424	0.198	0.392	0.379	0.056	0.057
	1.4	0.753	0.512	0.400	0.380	0.067	0.067
	1.6	0.999	0.991	0.401	0.388	0.092	0.077
	1.8	0.369	0.074	0.558	0.548	0.088	0.084
	2	0.365	0.096	0.555	0.535	0.079	0.088
	3	0.501	0.198	0.537	0.522	0.091	0.086
0.9	1	0.741	0.475	0.522	0.513	0.090	0.092
	1.2	0.998	0.982	0.509	0.490	0.113	0.103
	1.4	0.831	0.291	0.940	0.940	0.358	0.347
	1.6	0.820	0.284	0.940	0.933	0.353	0.354
	1.8	0.811	0.327	0.921	0.911	0.310	0.295
	2	0.859	0.485	0.890	0.883	0.271	0.263
	3	0.996	0.949	0.861	0.856	0.245	0.242
No. of Shaded region		18	12	10	9	1	1

Shaded region indicate Power of test values within [0.5-1.0]

Table N.63

Power of Test for variable, $p = 3, n = 25$

ρ	σ	M	M_{GMed}	M_B	M_T	M_{BGMed}	M_{TGMed}
0	1	0.042	0.048	0.052	0.047	0.055	0.050
	1.2	0.092	0.085	0.049	0.050	0.049	0.055
	1.4	0.511	0.413	0.051	0.052	0.080	0.077
	1.6	0.944	0.883	0.059	0.056	0.116	0.122
	1.8	1	1	0.065	0.065	0.184	0.184
	2	0.062	0.024	0.080	0.091	0.017	0.022
	3	0.107	0.044	0.094	0.085	0.023	0.019
0.2	1	0.458	0.299	0.083	0.094	0.034	0.033
	1.2	0.897	0.784	0.095	0.095	0.053	0.048
	1.4	1	1	0.103	0.104	0.103	0.098
	1.6	0.114	0.020	0.180	0.196	0.015	0.016
	1.8	0.159	0.037	0.177	0.186	0.013	0.014
	2	0.462	0.223	0.184	0.191	0.024	0.029
	3	0.865	0.700	0.192	0.196	0.035	0.040
0.5	1	1	1	0.189	0.192	0.069	0.080
	1.2	0.206	0.032	0.325	0.351	0.019	0.019
	1.4	0.256	0.041	0.333	0.345	0.024	0.024
	1.6	0.512	0.204	0.331	0.355	0.030	0.031
	1.8	0.868	0.628	0.336	0.337	0.043	0.044
	2	1	0.999	0.335	0.326	0.065	0.070
	3	0.342	0.043	0.516	0.512	0.040	0.042
0.7	1	0.374	0.065	0.503	0.502	0.041	0.041
	1.2	0.563	0.206	0.495	0.494	0.048	0.046
	1.4	0.854	0.577	0.482	0.497	0.057	0.060
	1.6	1	0.998	0.483	0.479	0.081	0.084
	1.8	0.485	0.070	0.671	0.683	0.075	0.068
	2	0.493	0.088	0.673	0.672	0.077	0.086
	3	0.631	0.209	0.641	0.657	0.084	0.084
0.9	1	0.876	0.545	0.613	0.623	0.091	0.082
	1.2	1	0.996	0.590	0.591	0.094	0.107
	1.4	0.933	0.330	0.974	0.982	0.396	0.378
	1.6	0.923	0.326	0.979	0.979	0.365	0.378
	1.8	0.914	0.365	0.971	0.966	0.321	0.331
	2	0.952	0.531	0.951	0.946	0.290	0.293
	3	1	0.979	0.930	0.931	0.270	0.273
No. of Shaded region		21	14	12	12	0	0

Shaded region indicate Power of test values within [0.5-1.0]

Table N.64

Power of Test for variable, $p = 3, n = 50$

ρ	σ	M	M_{GMed}	M_B	M_T	M_{BGMed}	M_{TGMed}
0	1	0.050	0.047	0.044	0.051	0.048	0.049
	1.2	0.177	0.099	0.045	0.056	0.058	0.059
	1.4	0.898	0.687	0.047	0.052	0.093	0.103
	1.6	1	0.995	0.051	0.056	0.172	0.175
	1.8	1	1	0.059	0.071	0.267	0.265
	2	0.105	0.009	0.114	0.129	0.006	0.004
	3	0.216	0.030	0.124	0.130	0.009	0.009
0.2	1	0.855	0.453	0.129	0.136	0.023	0.021
	1.2	0.999	0.974	0.131	0.143	0.052	0.050
	1.4	1	1	0.140	0.152	0.129	0.115
	1.6	0.258	0.004	0.347	0.381	0.003	0.002
	1.8	0.362	0.015	0.347	0.366	0.005	0.002
	2	0.853	0.302	0.345	0.376	0.009	0.007
	3	0.999	0.928	0.339	0.364	0.019	0.022
0.5	1	1	1	0.348	0.347	0.059	0.064
	1.2	0.498	0.010	0.641	0.655	0.005	0.002
	1.4	0.583	0.018	0.638	0.648	0.006	0.007
	1.6	0.895	0.241	0.613	0.628	0.015	0.015
	1.8	0.998	0.867	0.601	0.613	0.022	0.019
	2	1	1	0.573	0.589	0.038	0.044
	3	0.736	0.021	0.855	0.865	0.014	0.013
0.7	1	0.784	0.033	0.837	0.858	0.018	0.018
	1.2	0.935	0.215	0.808	0.831	0.028	0.026
	1.4	0.998	0.805	0.796	0.805	0.036	0.035
	1.6	1	1	0.751	0.761	0.053	0.052
	1.8	0.883	0.045	0.948	0.959	0.038	0.040
	2	0.904	0.057	0.945	0.950	0.034	0.039
	3	0.962	0.223	0.921	0.937	0.046	0.048
0.9	1	0.998	0.749	0.904	0.907	0.056	0.059
	1.2	1	1	0.876	0.885	0.070	0.076
	1.4	1	0.431	1	1	0.439	0.455
	1.6	1	0.422	1	1	0.426	0.426
	1.8	0.999	0.480	1	1	0.357	0.360
	2	1	0.712	1	1	0.294	0.333
	3	1	1	0.997	0.998	0.272	0.270
No. of Shaded region		28	15	20	20	0	0
Shaded region indicate Power of test values within [0.5-1.0]							

Table N.65

Power of Test for variable, $p = 5, n = 2$

ρ	σ	M_B	M_T	M_{BGMed}	M_{TGMed}
0	1	0.057	0.050	0.059	0.058
	1.2	0.056	0.051	0.056	0.056
	1.4	0.051	0.052	0.048	0.048
	1.6	0.064	0.053	0.057	0.057
	1.8	0.062	0.054	0.055	0.056
	2	0.061	0.057	0.049	0.049
	3	0.680	0.069	0.047	0.046
0.2	1	0.053	0.055	0.056	0.055
	1.2	0.059	0.056	0.058	0.054
	1.4	0.059	0.056	0.054	0.055
	1.6	0.057	0.064	0.050	0.055
	1.8	0.065	0.058	0.052	0.056
	2	0.061	0.063	0.052	0.047
	3	0.077	0.074	0.044	0.046
0.5	1	0.053	0.054	0.054	0.054
	1.2	0.057	0.053	0.051	0.051
	1.4	0.059	0.050	0.048	0.048
	1.6	0.058	0.061	0.054	0.054
	1.8	0.057	0.064	0.046	0.046
	2	0.059	0.064	0.047	0.054
	3	0.083	0.079	0.044	0.044
0.7	1	0.061	0.054	0.043	0.043
	1.2	0.052	0.052	0.053	0.053
	1.4	0.058	0.064	0.044	0.044
	1.6	0.058	0.062	0.050	0.050
	1.8	0.069	0.061	0.043	0.043
	2	0.061	0.064	0.039	0.039
	3	0.085	0.082	0.043	0.043
0.9	1	0.064	0.067	0.038	0.037
	1.2	0.072	0.065	0.036	0.036
	1.4	0.066	0.062	0.035	0.050
	1.6	0.069	0.066	0.035	0.035
	1.8	0.070	0.068	0.039	0.039
	2	0.074	0.071	0.038	0.039
	3	0.088	0.089	0.035	0.035
No. of Shaded region		0	0	0	0

Shaded region indicate Power of test values within [0.5-1.0]

Table N.66

Power of Test for variable, $p = 5, n = 5$

ρ	σ	M_B	M_T	M_{BGMed}	M_{TGMed}
0	1	0.045	0.057	0.045	0.056
	1.2	0.052	0.058	0.042	0.058
	1.4	0.049	0.058	0.054	0.056
	1.6	0.057	0.072	0.058	0.056
	1.8	0.073	0.086	0.066	0.069
	2	0.053	0.064	0.039	0.041
	3	0.060	0.063	0.043	0.043
0.2	1	0.059	0.070	0.041	0.045
	1.2	0.067	0.073	0.041	0.048
	1.4	0.078	0.089	0.059	0.058
	1.6	0.063	0.075	0.043	0.043
	1.8	0.066	0.072	0.041	0.039
	2	0.069	0.076	0.041	0.042
	3	0.069	0.088	0.041	0.040
0.5	1	0.091	0.100	0.049	0.053
	1.2	0.080	0.099	0.050	0.050
	1.4	0.079	0.092	0.042	0.047
	1.6	0.088	0.094	0.038	0.050
	1.8	0.097	0.107	0.042	0.039
	2	0.112	0.128	0.048	0.044
	3	0.094	0.116	0.048	0.051
0.7	1	0.105	0.118	0.047	0.049
	1.2	0.108	0.133	0.042	0.043
	1.4	0.123	0.133	0.039	0.045
	1.6	0.138	0.151	0.048	0.049
	1.8	0.128	0.145	0.056	0.061
	2	0.133	0.145	0.051	0.050
	3	0.144	0.151	0.048	0.049
0.9	1	0.148	0.162	0.045	0.041
	1.2	0.169	0.180	0.051	0.049
	1.4	0.303	0.328	0.091	0.094
	1.6	0.300	0.326	0.083	0.096
	1.8	0.313	0.337	0.081	0.081
	2	0.321	0.321	0.080	0.074
	3	0.319	0.336	0.073	0.078
No. of Shaded region		0	0	0	0

Shaded region indicate Power of test values within [0.5-1.0]

Table N.67

Power of Test for variable, = 5 , n = 10

ρ	σ	M	M_{GMed}	M_B	M_T	M_{BGMed}	M_{TGMed}
0	1	0.053	0.048	0.049	0.049	0.053	0.050
	1.2	0.065	0.071	0.050	0.048	0.051	0.050
	1.4	0.166	0.195	0.053	0.048	0.068	0.066
	1.6	0.450	0.470	0.064	0.064	0.087	0.088
	1.8	0.979	0.980	0.081	0.076	0.138	0.139
	2	0.046	0.048	0.059	0.069	0.034	0.038
	3	0.073	0.068	0.060	0.063	0.036	0.041
0.2	1	0.152	0.165	0.067	0.065	0.044	0.046
	1.2	0.377	0.420	0.072	0.070	0.057	0.062
	1.4	0.962	0.967	0.091	0.089	0.099	0.096
	1.6	0.059	0.052	0.102	0.099	0.039	0.039
	1.8	0.074	0.070	0.094	0.095	0.036	0.038
	2	0.154	0.150	0.106	0.111	0.050	0.043
	3	0.328	0.364	0.112	0.121	0.050	0.061
0.5	1	0.940	0.947	0.139	0.130	0.089	0.090
	1.2	0.077	0.049	0.160	0.162	0.050	0.049
	1.4	0.091	0.067	0.160	0.166	0.049	0.050
	1.6	0.154	0.155	0.166	0.177	0.056	0.054
	1.8	0.316	0.340	0.175	0.172	0.066	0.062
	2	0.911	0.918	0.191	0.195	0.091	0.093
	3	0.094	0.064	0.230	0.248	0.070	0.077
0.7	1	0.096	0.077	0.249	0.250	0.077	0.075
	1.2	0.154	0.158	0.246	0.241	0.079	0.076
	1.4	0.302	0.322	0.261	0.256	0.074	0.074
	1.6	0.882	0.895	0.258	0.260	0.100	0.095
	1.8	0.110	0.069	0.333	0.328	0.102	0.101
	2	0.112	0.089	0.327	0.334	0.099	0.103
	3	0.167	0.156	0.337	0.333	0.097	0.097
0.9	1	0.278	0.313	0.328	0.334	0.098	0.103
	1.2	0.871	0.876	0.339	0.325	0.113	0.114
	1.4	0.254	0.132	0.765	0.767	0.306	0.321
	1.6	0.251	0.146	0.746	0.754	0.299	0.308
	1.8	0.259	0.196	0.719	0.715	0.250	0.259
	2	0.309	0.302	0.681	0.674	0.228	0.221
	3	0.758	0.791	0.638	0.642	0.218	0.210
No. of Shaded region		7	7	5	5	0	0

Shaded region indicate Power of test values within [0.5-1.0]

Table N.68

Power of Test for variable, $p = 5, n = 15$

ρ	σ	M	M_{GMed}	M_B	M_T	M_{BGMed}	M_{TGMed}
0	1	0.052	0.050	0.045	0.050	0.052	0.055
	1.2	0.092	0.087	0.049	0.046	0.056	0.065
	1.4	0.364	0.364	0.049	0.051	0.085	0.085
	1.6	0.831	0.842	0.058	0.056	0.126	0.133
	1.8	1	1	0.079	0.074	0.200	0.210
	2	0.060	0.047	0.074	0.064	0.033	0.035
	3	0.087	0.075	0.075	0.077	0.030	0.036
0.2	1	0.301	0.293	0.085	0.077	0.043	0.055
	1.2	0.760	0.757	0.096	0.080	0.075	0.083
	1.4	1	1	0.100	0.097	0.147	0.145
	1.6	0.079	0.054	0.147	0.136	0.033	0.032
	1.8	0.109	0.076	0.134	0.133	0.035	0.036
	2	0.295	0.259	0.150	0.148	0.045	0.051
	3	0.696	0.662	0.157	0.153	0.062	0.071
0.5	1	1	1	0.172	0.166	0.113	0.117
	1.2	0.113	0.068	0.254	0.248	0.053	0.058
	1.4	0.146	0.082	0.256	0.247	0.057	0.057
	1.6	0.285	0.249	0.268	0.255	0.057	0.065
	1.8	0.646	0.615	0.264	0.248	0.069	0.079
	2	0.999	0.998	0.278	0.269	0.115	0.111
	3	0.162	0.074	0.384	0.390	0.084	0.096
0.7	1	0.193	0.099	0.377	0.387	0.086	0.098
	1.2	0.313	0.236	0.403	0.360	0.088	0.096
	1.4	0.616	0.572	0.379	0.370	0.094	0.100
	1.6	1	0.996	0.375	0.374	0.118	0.125
	1.8	0.232	0.105	0.535	0.534	0.142	0.157
	2	0.241	0.126	0.543	0.518	0.129	0.161
	3	0.333	0.249	0.523	0.503	0.126	0.133
0.9	1	0.612	0.538	0.500	0.493	0.130	0.134
	1.2	0.996	0.995	0.472	0.469	0.130	0.153
	1.4	0.633	0.284	0.947	0.948	0.539	0.558
	1.6	0.613	0.287	0.947	0.933	0.513	0.520
	1.8	0.593	0.350	0.908	0.903	0.410	0.422
	2	0.652	0.504	0.871	0.863	0.335	0.351
	3	0.987	0.972	0.816	0.816	0.279	0.305
No. of Shaded region		17	14	9	8	2	2

Shaded region indicate Power of test values within [0.5-1.0]

Table N.69

Power of Test for variable, = 5 , n = 20

ρ	σ	M	M_{GMed}	M_B	M_T	M_{BGMed}	M_{TGMed}
0	1	0.050	0.050	0.058	0.047	0.053	0.048
	1.2	0.100	0.087	0.059	0.049	0.051	0.055
	1.4	0.554	0.564	0.056	0.055	0.068	0.096
	1.6	0.971	0.842	0.055	0.058	0.087	0.168
	1.8	1	1	0.082	0.075	0.138	0.240
	2	0.066	0.047	0.095	0.092	0.034	0.021
	3	0.110	0.075	0.087	0.088	0.036	0.025
0.2	1	0.481	0.293	0.097	0.094	0.044	0.040
	1.2	0.939	0.757	0.108	0.100	0.057	0.079
	1.4	1	1	0.136	0.116	0.099	0.160
	1.6	0.097	0.054	0.199	0.189	0.039	0.025
	1.8	0.143	0.076	0.203	0.183	0.036	0.029
	2	0.449	0.259	0.209	0.199	0.050	0.036
	3	0.892	0.662	0.213	0.206	0.050	0.059
0.5	1	1	1	0.218	0.209	0.089	0.119
	1.2	0.165	0.068	0.369	0.346	0.050	0.048
	1.4	0.209	0.082	0.357	0.349	0.049	0.052
	1.6	0.459	0.249	0.363	0.355	0.056	0.058
	1.8	0.858	0.615	0.370	0.355	0.066	0.073
	2	1	0.998	0.334	0.340	0.091	0.120
	3	0.256	0.074	0.549	0.543	0.070	0.102
0.7	1	0.300	0.099	0.540	0.528	0.077	0.087
	1.2	0.484	0.236	0.521	0.519	0.079	0.097
	1.4	0.842	0.572	0.506	0.486	0.074	0.099
	1.6	1	0.996	0.482	0.467	0.100	0.122
	1.8	0.377	0.105	0.726	0.721	0.102	0.154
	2	0.398	0.126	0.712	0.710	0.099	0.155
	3	0.551	0.249	0.660	0.655	0.097	0.133
0.9	1	0.833	0.538	0.632	0.636	0.098	0.125
	1.2	1	0.995	0.601	0.592	0.113	0.154
	1.4	0.870	0.284	0.993	0.991	0.306	0.662
	1.6	0.863	0.287	0.991	0.989	0.299	0.627
	1.8	0.841	0.350	0.976	0.976	0.250	0.514
	2	0.877	0.504	1	0.951	0.228	0.400
	3	1	0.972	0.92	0.911	0.218	0.331
No. of Shaded region		19	15	14	13	0	3

Shaded region indicate Power of test values within [0.5-1.0]

Table N.70

Power of Test for variable, $p = 5, n = 25$

ρ	σ	M	M_{GMed}	M_B	M_T	M_{BGMed}	M_{TGMed}
0	1	0.053	0.057	0.051	0.042	0.050	0.049
	1.2	0.135	0.143	0.061	0.049	0.056	0.051
	1.4	0.728	0.709	0.056	0.052	0.111	0.109
	1.6	0.995	0.994	0.062	0.062	0.197	0.188
	1.8	1	1	0.075	0.069	0.303	0.302
	2	0.071	0.032	0.100	0.097	0.015	0.016
	3	0.139	0.085	0.094	0.090	0.019	0.018
0.2	1	0.632	0.545	0.113	0.105	0.039	0.043
	1.2	0.984	0.975	0.115	0.110	0.091	0.093
	1.4	1	1	0.137	0.120	0.186	0.178
	1.6	0.126	0.039	0.244	0.228	0.019	0.020
	1.8	0.196	0.085	0.248	0.227	0.018	0.023
	2	0.610	0.447	0.261	0.234	0.029	0.035
	3	0.971	0.934	0.254	0.235	0.055	0.060
0.5	1	1	1	0.260	0.245	0.131	0.132
	1.2	0.223	0.051	0.477	0.449	0.041	0.038
	1.4	0.302	0.094	0.455	0.434	0.041	0.043
	1.6	0.612	0.412	0.456	0.422	0.048	0.042
	1.8	0.958	0.895	0.432	0.422	0.060	0.062
	2	1	1	0.425	0.398	0.115	0.109
	3	0.376	0.093	0.679	0.658	0.094	0.092
0.7	1	0.432	0.128	0.682	0.659	0.093	0.096
	1.2	0.663	0.384	0.625	0.618	0.092	0.092
	1.4	0.949	0.853	0.607	0.584	0.091	0.088
	1.6	1	1	0.567	0.567	0.127	0.125
	1.8	0.535	0.145	0.834	0.826	0.168	0.179
	2	0.562	0.168	0.831	0.812	0.160	0.171
	3	0.724	0.375	0.789	0.768	0.146	0.155
0.9	1	0.943	0.813	0.757	0.728	0.132	0.131
	1.2	1	1	0.693	0.682	0.154	0.143
	1.4	0.968	0.590	0.999	0.998	0.758	0.762
	1.6	0.963	0.601	0.998	0.999	0.729	0.711
	1.8	0.947	0.574	0.994	0.995	0.575	0.568
	2	0.969	0.758	0.986	0.985	0.448	0.447
	3	1	0.999	0.959	0.959	0.359	0.354
No. of Shaded region		25	19	22	22	3	3

Shaded region indicate Power of test values within [0.5-1.0]

Table N.71

Power of Test for variable, $p = 5, n = 50$

ρ	σ	M	M_{GMed}	M_B	M_T	M_{BGMed}	M_{TGMed}
0	1	0.060	0.053	0.055	0.054	0.047	0.051
	1.2	0.288	0.218	0.051	0.049	0.056	0.067
	1.4	0.991	0.966	0.055	0.064	0.157	0.188
	1.6	1	1	0.056	0.066	0.305	0.340
	1.8	1	1	0.086	0.083	0.486	0.504
	2	0.099	0.014	0.154	0.163	0.002	0.005
	3	0.311	0.082	0.157	0.174	0.005	0.006
0.2	1	0.974	0.864	0.166	0.170	0.028	0.029
	1.2	1	1	0.176	0.181	0.092	0.113
	1.4	1	1	0.195	0.182	0.248	0.260
	1.6	0.291	0.012	0.478	0.501	0.003	0.004
	1.8	0.475	0.054	0.486	0.492	0.003	0.004
	2	0.967	0.743	0.461	0.479	0.013	0.014
	3	1	1	0.439	0.450	0.033	0.037
0.5	1	1	1	0.420	0.432	0.136	0.141
	1.2	0.578	0.026	0.816	0.829	0.014	0.017
	1.4	0.698	0.078	0.800	0.815	0.015	0.018
	1.6	0.973	0.640	0.754	0.776	0.023	0.018
	1.8	1	0.998	0.720	0.737	0.036	0.031
	2	1	1	0.668	0.680	0.080	0.101
	3	0.821	0.071	0.956	0.968	0.061	0.070
0.7	1	0.868	0.125	0.950	0.958	0.061	0.072
	1.2	0.981	0.587	0.914	0.925	0.058	0.064
	1.4	1	0.991	0.887	0.891	0.064	0.058
	1.6	1	1	0.835	0.845	0.087	0.091
	1.8	0.951	0.181	0.994	0.993	0.181	0.196
	2	0.958	0.224	0.992	0.992	0.150	0.172
	3	0.990	0.601	0.984	0.987	0.119	0.129
0.9	1	1	0.987	0.957	0.961	0.107	0.111
	1.2	1	1	0.924	0.935	0.108	0.124
	1.4	1	0.888	1	1	0.932	0.940
	1.6	1	0.880	1	1	0.907	0.920
	1.8	1	0.861	1	1	0.762	0.773
	2	1	0.961	1	1	0.575	0.594
	3	1	1	0.999	0.999	0.428	0.425
No. of Shaded region		29	23	20	21	4	4

Shaded region indicate Power of test values within [0.5-1.0]

Table N.72

Power of Test for variable, $p = 8$ and , $n = 2$

ρ	σ	M_B	M_T	M_{BGMed}	M_{TGMed}
0	1	0.056	0.050	0.055	0.055
	1.2	0.058	0.051	0.056	0.056
	1.4	0.050	0.053	0.042	0.042
	1.6	0.061	0.055	0.053	0.056
	1.8	0.058	0.057	0.042	0.042
	2	0.069	0.060	0.052	0.052
	3	0.081	0.080	0.045	0.046
0.2	1	0.055	0.052	0.053	0.046
	1.2	0.056	0.051	0.052	0.051
	1.4	0.052	0.050	0.045	0.054
	1.6	0.060	0.054	0.049	0.048
	1.8	0.057	0.061	0.042	0.044
	2	0.061	0.065	0.043	0.045
	3	0.081	0.085	0.046	0.053
0.5	1	0.055	0.056	0.054	0.047
	1.2	0.054	0.055	0.045	0.046
	1.4	0.052	0.059	0.044	0.044
	1.6	0.062	0.061	0.047	0.045
	1.8	0.058	0.070	0.043	0.044
	2	0.070	0.069	0.044	0.037
	3	0.085	0.092	0.038	0.040
0.7	1	0.057	0.054	0.049	0.040
	1.2	0.055	0.058	0.043	0.039
	1.4	0.061	0.056	0.042	0.042
	1.6	0.057	0.061	0.039	0.038
	1.8	0.058	0.067	0.039	0.045
	2	0.061	0.070	0.038	0.041
	3	0.095	0.095	0.037	0.033
0.9	1	0.073	0.072	0.033	0.030
	1.2	0.070	0.075	0.035	0.040
	1.4	0.072	0.071	0.037	0.033
	1.6	0.068	0.076	0.034	0.032
	1.8	0.077	0.083	0.032	0.032
	2	0.072	0.083	0.029	0.037
	3	0.106	0.108	0.035	0.035
No. of Shaded region		0	0	0	0

Shaded region indicate Power of test values within [0.5-1.0]

Table N.73

Power of Test for variable, = 8 , n = 5

ρ	σ	M_B	M_T	M_{BGMed}	M_{TGMed}
0	1	0.045	0.051	0.045	0.054
	1.2	0.052	0.054	0.042	0.051
	1.4	0.049	0.056	0.054	0.059
	1.6	0.057	0.064	0.058	0.068
	1.8	0.073	0.091	0.066	0.085
	2	0.053	0.058	0.039	0.050
	3	0.060	0.060	0.043	0.052
0.2	1	0.059	0.071	0.041	0.040
	1.2	0.067	0.077	0.041	0.051
	1.4	0.078	0.096	0.059	0.065
	1.6	0.063	0.074	0.043	0.045
	1.8	0.066	0.074	0.041	0.048
	2	0.069	0.083	0.041	0.038
	3	0.069	0.093	0.041	0.041
0.5	1	0.091	0.127	0.049	0.054
	1.2	0.080	0.101	0.050	0.047
	1.4	0.079	0.108	0.042	0.043
	1.6	0.088	0.105	0.038	0.043
	1.8	0.097	0.111	0.042	0.044
	2	0.112	0.150	0.048	0.061
	3	0.094	0.132	0.048	0.056
0.7	1	0.105	0.134	0.047	0.049
	1.2	0.108	0.154	0.042	0.045
	1.4	0.123	0.148	0.039	0.046
	1.6	0.138	0.181	0.048	0.059
	1.8	0.128	0.173	0.056	0.062
	2	0.133	0.168	0.051	0.056
	3	0.144	0.182	0.048	0.049
0.9	1	0.148	0.190	0.045	0.049
	1.2	0.169	0.199	0.051	0.054
	1.4	0.303	0.424	0.091	0.122
	1.6	0.300	0.412	0.083	0.109
	1.8	0.313	0.414	0.081	0.101
	2	0.321	0.394	0.080	0.087
	3	0.319	0.391	0.073	0.082
No. of Shaded region		0	0	0	0

Shaded region indicate Power of test values within [0.5-1.0]

Table N.74

Power of Test for variable, $p = 8, n = 10$

ρ		M	M_{GMed}	M_B	M_T	M_{BGMed}	M_{TGMed}
0	1	0.051	0.045	0.049	0.048	0.048	0.053
	1.2	0.066	0.060	0.062	0.050	0.010	0.064
	1.4	0.108	0.109	0.060	0.055	0.045	0.080
	1.6	0.225	0.266	0.064	0.073	0.038	0.125
	1.8	0.818	0.871	0.043	0.099	0.075	0.186
	2	0.046	0.046	0.072	0.060	0.056	0.041
	3	0.060	0.052	0.058	0.069	0.045	0.042
0.2	1	0.099	0.107	0.080	0.072	0.012	0.057
	1.2	0.184	0.229	0.067	0.086	0.019	0.079
	1.4	0.741	0.813	0.050	0.105	0.070	0.144
	1.6	0.049	0.049	0.079	0.116	0.008	0.047
	1.8	0.058	0.060	0.033	0.119	0.006	0.048
	2	0.093	0.102	0.051	0.130	0.014	0.058
	3	0.171	0.207	0.064	0.147	0.023	0.070
0.5	1	0.674	0.765	0.078	0.150	0.052	0.125
	1.2	0.059	0.055	0.059	0.197	0.013	0.073
	1.4	0.064	0.066	0.074	0.202	0.013	0.070
	1.6	0.090	0.101	0.085	0.218	0.018	0.073
	1.8	0.153	0.206	0.115	0.217	0.027	0.081
	2	0.614	0.723	0.132	0.235	0.060	0.118
	3	0.063	0.066	0.120	0.309	0.022	0.109
0.7	1	0.076	0.068	0.132	0.305	0.021	0.103
	1.2	0.097	0.106	0.153	0.317	0.034	1
	1.4	0.146	0.198	0.181	0.320	0.040	0.097
	1.6	0.560	0.688	0.202	0.325	0.071	0.141
	1.8	0.062	0.065	0.218	0.444	0.043	0.151
	2	0.080	0.073	0.217	0.440	0.039	0.158
	3	0.089	0.115	0.239	0.421	0.049	0.131
0.9	1	0.148	0.188	0.249	0.407	0.061	0.129
	1.2	0.517	0.666	0.274	0.409	0.081	0.139
	1.4	0.108	0.099	0.746	0.903	0.254	0.523
	1.6	0.118	0.106	0.718	0.892	0.248	0.496
	1.8	0.123	0.131	0.670	0.834	0.196	0.387
	2	0.144	0.188	0.623	0.778	0.172	0.306
	3	0.375	0.590	0.592	0.717	0.164	0.258
No. of Shaded region		6	7	5	5	0	2

Shaded region indicate Power of test values within [0.5-1.0]

Table N.75

Power of Test for variable, $p = 8, n = 15$

ρ	σ	M	M_{GMed}	M_B	M_T	M_{BGMed}	M_{TGMed}
0	1	0.051	0.047	0.050	0.055	0.047	0.050
	1.2	0.081	0.086	0.022	0.052	0.022	0.046
	1.4	0.320	0.363	0.030	0.056	0.051	0.088
	1.6	0.811	0.852	0.043	0.068	0.107	0.173
	1.8	1	1	0.067	0.090	0.203	0.251
	2	0.052	0.046	0.041	0.081	0.011	0.026
	3	0.082	0.069	0.041	0.081	0.015	0.030
0.2	1	0.262	0.284	0.048	0.088	0.023	0.045
	1.2	0.705	0.760	0.065	0.105	0.060	0.094
	1.4	1	1	0.092	0.132	0.143	0.178
	1.6	0.064	0.049	0.092	0.173	0.019	0.037
	1.8	0.096	0.089	0.105	0.191	0.019	0.041
	2	0.235	0.261	0.118	0.188	0.025	0.054
	3	0.614	0.677	0.144	0.197	0.041	0.067
0.5	1	1	0.998	0.158	0.220	0.107	0.143
	1.2	0.094	0.065	0.226	0.329	0.042	0.078
	1.4	0.119	0.086	0.232	0.334	0.035	0.074
	1.6	0.237	0.251	0.232	0.338	0.044	0.075
	1.8	0.555	0.625	0.248	0.321	0.066	0.086
	2	1	1	0.264	0.328	0.112	0.134
	3	0.111	0.079	0.381	0.521	0.083	0.151
0.7	1	0.137	0.108	0.388	0.526	0.077	0.131
	1.2	0.234	0.246	0.370	0.490	0.079	0.114
	1.4	0.498	0.573	0.382	0.469	0.088	0.113
	1.6	0.998	0.999	0.374	0.435	0.118	0.147
	1.8	0.153	0.107	0.566	0.703	0.160	0.234
	2	0.177	0.120	0.568	0.695	0.157	0.219
	3	0.243	0.232	0.535	0.625	0.125	0.176
0.9	1	0.472	0.542	0.513	0.585	0.120	0.156
	1.2	0.997	0.998	0.497	0.544	0.132	0.165
	1.4	0.480	0.269	0.981	0.991	0.682	0.780
	1.6	0.458	0.291	0.978	0.989	0.648	0.748
	1.8	0.400	0.347	0.942	0.963	0.473	0.590
	2	0.459	0.514	0.897	0.930	0.360	0.428
	3	0.970	0.982	0.839	0.862	0.295	0.328
No. of Shaded region		11	13	9	12	2	3

Shaded region indicate Power of test within [0.5-1.0]

Table N.76

Power of Test for variable, $p = 8, n = 20$

ρ	σ	M	M_{GMed}	M_B	M_T	M_{BGMed}	M_{TGMed}
0	1	0.051	0.051	0.053	0.053	0.056	0.054
	1.2	0.112	0.114	0.035	0.049	0.036	0.060
	1.4	0.602	0.632	0.042	0.057	0.075	0.122
	1.6	0.985	0.989	0.053	0.066	0.180	0.236
	1.8	1	1	0.085	0.090	0.296	0.348
	2	0.054	0.045	0.064	0.085	0.011	0.025
	3	0.111	0.091	0.071	0.091	0.011	0.026
0.2	1	0.493	0.483	0.080	0.105	0.030	0.043
	1.2	0.953	0.958	0.103	0.108	0.070	0.111
	1.4	1	1	0.122	0.137	0.187	0.226
	1.6	0.087	0.052	0.187	0.236	0.023	0.039
	1.8	0.139	0.097	0.196	0.227	0.025	0.042
	2	0.434	0.404	0.216	0.240	0.032	0.048
	3	0.912	0.912	0.215	0.247	0.057	0.087
0.5	1	1	1	0.229	0.259	0.138	0.162
	1.2	0.136	0.066	0.394	0.444	0.051	0.091
	1.4	0.190	0.114	0.394	0.456	0.055	0.093
	1.6	0.390	0.388	0.384	0.413	0.061	0.082
	1.8	0.859	0.859	0.374	0.421	0.068	0.092
	2	1	1	0.366	0.392	0.131	0.157
	3	0.207	0.102	0.644	0.675	0.130	0.198
0.7	1	0.246	0.148	0.627	0.684	0.123	0.185
	1.2	0.426	0.365	0.587	0.621	0.100	0.137
	1.4	0.816	0.819	0.538	0.574	0.106	0.124
	1.6	1	1	0.524	0.540	0.129	0.164
	1.8	0.310	0.147	0.828	0.859	0.264	0.360
	2	0.335	0.193	0.806	0.836	0.238	0.314
	3	0.458	0.375	0.748	0.780	0.178	0.232
0.9	1	0.786	0.770	0.691	0.708	0.152	0.179
	1.2	1	1	0.641	0.653	0.162	0.185
	1.4	0.850	0.508	0.999	0.999	0.886	0.924
	1.6	0.825	0.510	0.999	1	0.844	0.895
	1.8	0.732	0.541	0.992	0.993	0.645	0.736
	2	0.769	0.697	0.975	0.977	0.483	0.554
	3	1	1	0.938	0.943	0.360	0.406
No. of Shaded region		17	17	15	15	4	4

Shaded region indicate Power of test within [0.5-1.0]

Table N.77

Power of Test for variable, $p = 8, n = 25$

ρ	σ	M	M_{GMed}	M_B	M_T	M_{BGMed}	M_{TGMed}
0	1	0.054	0.050	0.059	0.049	0.059	0.052
	1.2	0.151	0.152	0.038	0.050	0.046	0.066
	1.4	0.808	0.821	0.047	0.051	0.109	0.153
	1.6	0.999	1	0.057	0.067	0.235	0.289
	1.8	1	1	0.072	0.091	0.383	0.413
	2	0.063	0.043	0.077	0.100	0.008	0.019
	3	0.142	0.101	0.083	0.101	0.015	0.024
0.2	1	0.677	0.662	0.096	0.112	0.033	0.048
	1.2	0.997	0.994	0.106	0.131	0.098	0.129
	1.4	1	1	0.141	0.152	0.229	0.275
	1.6	0.115	0.048	0.256	0.298	0.042	0.037
	1.8	0.192	0.118	0.261	0.292	0.055	0.034
	2	0.613	0.553	0.257	0.296	0.034	0.046
	3	0.988	0.982	0.258	0.310	0.063	0.080
0.5	1	1	1	0.274	0.290	0.162	0.180
	1.2	0.196	0.075	0.540	0.572	0.070	0.114
	1.4	0.280	0.129	0.525	0.573	0.076	0.104
	1.6	0.607	0.493	0.490	0.534	0.060	0.080
	1.8	0.971	0.956	0.477	0.495	0.067	0.082
	2	1	1	0.442	0.470	0.139	0.159
	3	0.320	0.119	0.779	0.821	0.186	0.245
0.7	1	0.391	0.185	0.776	0.804	0.146	0.216
	1.2	0.613	0.477	0.703	0.745	0.121	0.149
	1.4	0.955	0.915	0.652	0.686	0.111	0.135
	1.6	1	1	0.597	0.622	0.139	0.177
	1.8	0.489	0.202	0.924	0.938	0.363	0.427
	2	0.516	0.255	0.914	0.925	0.316	0.401
	3	0.654	0.489	0.853	0.860	0.208	0.258
0.9	1	0.943	0.901	0.792	0.820	0.169	0.193
	1.2	1	1	0.744	0.754	0.180	0.203
	1.4	0.973	0.734	1	1	0.959	0.969
	1.6	0.964	0.710	1	1	0.939	0.955
	1.8	0.915	0.682	0.998	0.999	0.776	0.823
	2	0.937	0.822	0.991	0.995	0.576	0.638
	3	1	1	0.976	0.977	0.423	0.441
No. of Shaded region		24	20	17	17	4	4

Shaded region indicate Power of test within [0.5-1.0]

Table N.78

Power of Test for variable, $p = 8, n = 50$

ρ	σ	M	M_{GMed}	M_B	M_T	M_{BGMed}	M_{TGMed}
0	1	0.048	0.047	0.047	0.049	0.050	0.057
	1.2	0.350	0.315	0.040	0.052	0.072	0.079
	1.4	0.999	0.997	0.046	0.059	0.259	0.283
	1.6	1	1	0.063	0.073	0.497	0.523
	1.8	1	1	0.083	0.090	0.651	0.662
	2	0.086	0.020	0.179	0.190	0.006	0.007
	3	0.350	0.160	0.169	0.188	0.006	0.009
0.2	1	0.992	0.974	0.183	0.201	0.041	0.046
	1.2	0.999	0.999	0.197	0.219	0.179	0.191
	1.4	1	1	0.211	0.230	0.415	0.399
	1.6	0.252	0.030	0.606	0.638	0.012	0.020
	1.8	0.503	0.148	0.591	0.618	0.013	0.018
	2	0.983	0.916	0.542	0.580	0.020	0.025
	3	1	0.999	0.505	0.544	0.064	0.069
0.5	1	1	1	0.495	0.508	0.238	0.249
	1.2	0.561	0.090	0.922	0.934	0.107	0.125
	1.4	0.731	0.211	0.908	0.922	0.089	0.102
	1.6	0.981	0.861	0.853	0.860	0.057	0.053
	1.8	1	1	0.792	0.804	0.060	0.064
	2	1	0.999	0.733	0.744	0.153	0.166
	3	0.829	0.228	0.992	0.996	0.373	0.387
0.7	1	0.887	0.349	0.992	0.992	0.305	0.334
	1.2	0.985	0.824	0.972	0.973	0.150	0.162
	1.4	1	0.999	0.931	0.940	0.098	0.119
	1.6	1	1	0.882	0.880	0.140	0.152
	1.8	0.962	0.449	1	1	0.677	0.705
	2	0.971	0.525	0.999	1	0.600	0.636
	3	0.993	0.821	0.994	0.995	0.322	0.363
0.9	1	0.000	0.999	0.985	0.983	0.197	0.213
	1.2	1	1	0.949	0.965	0.170	0.184
	1.4	0.999	0.995	1	1	1	1
	1.6	1	0.989	1	0.999	0.999	1
	1.8	1	0.967	0.999	1	0.979	0.986
	2	1	0.994	1	1	0.824	0.855
	3	0.999	1	0.996	1	0.565	0.581
No. of Shaded region		29	24	24	25	8	9

Shaded region indicate Power of test within [0.5-1.0]

Appendix O.

Table O.79

Power of Test for variable, $p = 10, n = 5$

ρ	σ	M_B	M_T	M_{BGMed}	M_{TGMed}
0	1	0.049	0.049	0.055	0.051
	1.2	0.053	0.052	0.060	0.057
	1.4	0.060	0.057	0.052	0.058
	1.6	0.076	0.073	0.080	0.060
	1.8	0.095	0.102	0.102	0.092
	2	0.059	0.053	0.041	0.043
	3	0.058	0.063	0.045	0.042
0.2	1	0.061	0.064	0.042	0.043
	1.2	0.074	0.072	0.048	0.052
	1.4	0.110	0.097	0.079	0.071
	1.6	0.080	0.067	0.048	0.040
	1.8	0.084	0.076	0.044	0.044
	2	0.083	0.084	0.037	0.041
	3	0.099	0.096	0.049	0.043
0.5	1	0.136	0.137	0.066	0.062
	1.2	0.100	0.094	0.049	0.049
	1.4	0.107	0.101	0.054	0.044
	1.6	0.130	0.113	0.036	0.039
	1.8	0.135	0.127	0.046	0.039
	2	0.162	0.157	0.060	0.056
	3	0.134	0.128	0.053	0.050
0.7	1	0.135	0.138	0.061	0.052
	1.2	0.162	0.149	0.046	0.043
	1.4	0.170	0.169	0.048	0.044
	1.6	0.191	0.202	0.061	0.058
	1.8	0.177	0.169	0.064	0.063
	2	0.185	0.170	0.075	0.059
	3	0.208	0.197	0.060	0.052
0.9	1	0.223	0.205	0.050	0.043
	1.2	0.230	0.232	0.063	0.058
	1.4	0.481	0.475	0.144	0.128
	1.6	0.478	0.470	0.129	0.128
	1.8	0.446	0.453	0.122	0.104
	2	0.440	0.432	0.086	0.093
	3	0.426	0.419	0.101	0.082
No. of Shaded region		0	0	0	0

Shaded region indicate Power of test within [0.5-1.0]

Table O.80

Power of Test for variable, $p = 10, n = 10$

ρ	σ	M_B	M_T	M_{BGMed}	M_{TGMed}
0	1	0.052	0.049	0.052	0.052
	1.2	0.049	0.054	0.051	0.057
	1.4	0.046	0.064	0.069	0.077
	1.6	0.070	0.073	0.116	0.123
	1.8	0.090	0.099	0.182	0.202
	2	0.058	0.079	0.033	0.038
	3	0.067	0.069	0.029	0.032
0.2	1	0.071	0.091	0.044	0.051
	1.2	0.094	0.107	0.075	0.082
	1.4	0.127	0.130	0.137	0.166
	1.6	0.120	0.127	0.040	0.047
	1.8	0.122	0.132	0.036	0.050
	2	0.128	0.146	0.046	0.058
	3	0.152	0.168	0.067	0.070
0.5	1	0.176	0.184	0.121	0.137
	1.2	0.200	0.225	0.061	0.078
	1.4	0.204	0.238	0.063	0.076
	1.6	0.226	0.249	0.064	0.069
	1.8	0.240	0.251	0.076	0.078
	2	0.256	0.270	0.119	0.130
	3	0.338	0.362	0.111	0.130
0.7	1	0.346	0.371	0.099	0.115
	1.2	0.334	0.359	0.090	0.103
	1.4	0.344	0.341	0.094	0.104
	1.6	0.343	0.350	0.128	0.125
	1.8	0.481	0.513	0.164	0.188
	2	0.484	0.508	0.158	0.176
	3	0.444	0.481	0.135	0.151
0.9	1	0.430	0.447	0.122	0.127
	1.2	0.416	0.423	0.142	0.147
	1.4	0.940	0.945	0.586	0.619
	1.6	0.931	0.938	0.548	0.568
	1.8	0.859	0.873	0.397	0.424
	2	0.799	0.800	0.319	0.335
	3	0.742	0.739	0.269	0.256
No. of Shaded region		5	7	2	2

Shaded region indicate Power of test within [0.5-1.0]

Table O.81

Power of Test for variable, $p = 10, n = 15$

ρ	σ	M	M_{GMed}	M_B	M_T	M_{BGMed}	M_{TGMed}
0	1	0.058	0.047	0.046	0.051	0.047	0.044
	1.2	0.081	0.081	0.044	0.052	0.055	0.056
	1.4	0.271	0.317	0.064	0.061	0.138	0.108
	1.6	0.710	0.782	0.078	0.069	0.250	0.192
	1.8	1	1	0.094	0.101	0.388	0.290
	2	0.057	0.051	0.076	0.079	0.027	0.030
	3	0.081	0.077	0.085	0.085	0.030	0.034
0.2	1	0.220	0.264	0.095	0.093	0.058	0.053
	1.2	0.599	0.693	0.116	0.121	0.127	0.100
	1.4	1	1	0.137	0.146	0.253	0.204
	1.6	0.066	0.060	0.195	0.188	0.047	0.045
	1.8	0.095	0.085	0.204	0.196	0.048	0.049
	2	0.194	0.235	0.209	0.210	0.052	0.053
	3	0.503	0.628	0.209	0.225	0.081	0.078
0.5	1	0.994	1	0.234	0.238	0.184	0.169
	1.2	0.081	0.065	0.361	0.385	0.133	0.095
	1.4	0.101	0.092	0.368	0.373	0.117	0.095
	1.6	0.198	0.241	0.361	0.366	0.094	0.080
	1.8	0.452	0.566	0.342	0.361	0.095	0.094
	2	0.997	0.999	0.351	0.353	0.170	0.157
	3	0.098	0.086	0.590	0.597	0.279	0.200
0.7	1	0.125	0.106	0.580	0.580	0.252	0.181
	1.2	0.192	0.235	0.542	0.540	0.169	0.142
	1.4	0.421	0.518	0.502	0.510	0.143	0.129
	1.6	0.994	0.999	0.462	0.474	0.172	0.165
	1.8	0.133	0.102	0.778	0.778	0.470	0.326
	2	0.146	0.126	0.749	0.769	0.414	0.302
	3	0.192	0.238	0.678	0.683	0.267	0.218
0.9	1	0.378	0.516	0.637	0.641	0.200	0.181
	1.2	0.988	0.996	0.573	0.582	0.207	0.190
	1.4	0.359	0.246	0.998	0.999	0.977	0.889
	1.6	0.365	0.257	0.997	0.995	0.961	0.855
	1.8	0.294	0.335	0.983	0.981	0.824	0.648
	2	0.343	0.483	0.946	0.947	0.607	0.486
	3	0.924	0.978	0.890	0.884	0.432	0.361
No. of Shaded region		10	13	14	14	4	3

Shaded region indicate Power of test within [0.5-1.0]

Table O.82

Power of Test for variable, $p = 10, n = 20$

ρ	σ	M	M_{GMed}	M_B	M_T	M_{BGMed}	M_{TGMed}
0	1	0.051	0.052	0.051	0.047	0.049	0.051
	1.2	0.112	0.123	0.049	0.051	0.060	0.067
	1.4	0.549	0.623	0.054	0.055	0.114	0.142
	1.6	0.977	0.987	0.074	0.070	0.203	0.263
	1.8	1	1	0.097	0.096	0.301	0.386
	2	0.064	0.054	0.095	0.091	0.037	0.028
	3	0.114	0.107	0.095	0.102	0.038	0.032
0.2	1	0.443	0.484	0.109	0.108	0.061	0.055
	1.2	0.938	0.955	0.129	0.130	0.117	0.130
	1.4	1	1	0.152	0.157	0.208	0.250
	1.6	0.080	0.057	0.253	0.261	0.049	0.051
	1.8	0.130	0.107	0.265	0.270	0.056	0.053
	2	0.380	0.417	0.267	0.271	0.054	0.061
	3	0.878	0.913	0.268	0.273	0.084	0.089
0.5	1	1	1	0.287	0.278	0.165	0.198
	1.2	0.119	0.083	0.524	0.513	0.110	0.142
	1.4	0.153	0.142	0.498	0.501	0.105	0.116
	1.6	0.361	0.383	0.477	0.478	0.084	0.099
	1.8	0.810	0.864	0.447	0.445	0.094	0.103
	2	1	1	0.433	0.419	0.158	0.181
	3	0.183	0.114	0.764	0.774	0.201	0.286
0.7	1	0.221	0.165	0.746	0.749	0.182	0.263
	1.2	0.353	0.375	0.674	0.683	0.145	0.173
	1.4	0.755	0.821	0.638	0.614	0.137	0.153
	1.6	1	1	0.572	0.581	0.162	0.190
	1.8	0.265	0.164	0.913	0.917	0.350	0.472
	2	0.293	0.200	0.894	0.893	0.320	0.433
	3	0.384	0.385	0.827	0.822	0.211	0.277
0.9	1	0.719	0.772	0.750	0.765	0.181	0.203
	1.2	1	1	0.695	0.677	0.199	0.205
	1.4	0.784	0.525	1	1	0.891	0.977
	1.6	0.758	0.523	1	1	0.866	0.965
	1.8	0.629	0.537	0.998	0.997	0.664	0.825
	2	0.663	0.708	0.986	0.982	0.503	0.613
	3	1	1	0.955	0.954	0.356	0.428
No. of Shaded region		18	18	16	17	4	4

Shaded region indicate Power of test within [0.5-1.0]

Table O.83

Power of Test for variable, $p = 10, n = 25$

ρ	σ	M	M_{GMed}	M_B	M_T	M_{BGMed}	M_{TGMed}
0	1	0.052	0.047	0.054	0.060	0.044	0.049
	1.2	0.144	0.161	0.051	0.054	0.063	0.059
	1.4	0.773	0.818	0.057	0.071	0.183	0.158
	1.6	0.999	0.999	0.074	0.085	0.324	0.324
	1.8	1	1	0.099	0.107	0.451	0.438
	2	0.055	0.044	0.115	0.129	0.021	0.018
	3	0.142	0.112	0.118	0.123	0.023	0.019
0.2	1	0.644	0.667	0.124	0.148	0.056	0.047
	1.2	0.994	0.994	0.144	0.155	0.142	0.142
	1.4	1	1	0.163	0.175	0.304	0.287
	1.6	0.100	0.057	0.350	0.358	0.049	0.044
	1.8	0.170	0.120	0.331	0.348	0.051	0.047
	2	0.572	0.569	0.323	0.346	0.051	0.047
	3	0.987	0.982	0.335	0.340	0.081	0.084
0.5	1	1	1	0.326	0.352	0.211	0.216
	1.2	0.166	0.087	0.663	0.670	0.157	0.146
	1.4	0.237	0.148	0.650	0.664	0.134	0.132
	1.6	0.539	0.505	0.594	0.606	0.092	0.089
	1.8	0.956	0.964	0.556	0.556	0.101	0.091
	2	1	1	0.509	0.518	0.175	0.180
	3	0.285	0.136	0.887	0.883	0.352	0.338
0.7	1	0.340	0.202	0.862	0.875	0.305	0.297
	1.2	0.544	0.487	0.788	0.808	0.190	0.191
	1.4	0.929	0.927	0.732	0.743	0.142	0.150
	1.6	1	1	0.675	0.663	0.185	0.184
	1.8	0.417	0.207	0.974	0.974	0.593	0.582
	2	0.473	0.269	0.958	0.967	0.521	0.508
	3	0.581	0.480	0.912	0.923	0.330	0.319
0.9	1	0.915	0.897	0.845	0.856	0.220	0.214
	1.2	1	1	0.782	0.785	0.213	0.209
	1.4	0.961	0.743	1	1	0.995	0.995
	1.6	0.949	0.727	1	1	0.990	0.990
	1.8	0.861	0.694	0.999	1	0.899	0.904
	2	0.881	0.832	0.997	0.998	0.688	0.677
	3	1	1	0.982	0.985	0.473	0.477
No. of Shaded region		23	21	20	20	6	6

Shaded region indicate Power of test within [0.5-1.0]

Table O.84

Power of Test for variable, $p = 10, n = 50$

ρ	σ	M	M_{GMed}	M_B	M_T	M_{BGMed}	M_{TGMed}
0	1	0.050	0.047	0.048	0.048	0.055	0.051
	1.2	0.411	0.369	0.053	0.045	0.083	0.074
	1.4	1	0.999	0.066	0.056	0.322	0.318
	1.6	1	1	0.081	0.067	0.574	0.582
	1.8	1	1	0.102	0.098	0.715	0.721
	2	0.092	0.030	0.209	0.205	0.008	0.007
	3	0.390	0.219	0.209	0.206	0.008	0.007
0.2	1	0.997	0.990	0.226	0.220	0.046	0.051
	1.2	1	1	0.242	0.225	0.236	0.230
	1.4	1	1	0.247	0.238	0.468	0.463
	1.6	0.244	0.056	0.718	0.684	0.050	0.043
	1.8	0.528	0.216	0.696	0.677	0.032	0.035
	2	0.989	0.958	0.624	0.617	0.027	0.030
	3	1	1	0.572	0.572	0.087	0.075
0.5	1	1	1	0.519	0.510	0.295	0.298
	1.2	0.571	0.151	0.967	0.964	0.258	0.262
	1.4	0.736	0.304	0.957	0.947	0.206	0.192
	1.6	0.986	0.906	0.902	0.889	0.086	0.093
	1.8	1	1	0.831	0.825	0.079	0.078
	2	1	1	0.777	0.765	0.190	0.194
	3	0.845	0.358	1	0.999	0.651	0.646
0.7	1	0.901	0.772	0.998	0.998	0.562	0.534
	1.2	0.987	0.872	0.979	0.985	0.260	0.264
	1.4	1	1	0.953	0.952	0.144	0.148
	1.6	1	1	0.912	0.900	0.162	0.170
	1.8	0.972	0.597	1	1	0.904	0.906
	2	0.974	0.665	1	1	0.841	0.842
	3	0.990	0.875	1	0.998	0.491	0.491
0.9	1	0.999	1	0.990	0.992	0.271	0.275
	1.2	1	1	0.967	0.959	0.216	0.210
	1.4	1	0.999	1	1	1	1
	1.6	1	0.998	1	1	1	1
	1.8	1	0.980	1	1	0.997	0.998
	2	1	1	1	1	0.917	0.922
	3	1	1	1	1	0.629	0.613
No. of Shaded region		30	25	25	25	11	11

Shaded region indicate Power of test within [0.5-1.0]

Table O.85

Power of Test for variable, $p = 15, n = 5$

ρ	σ	M_B	M_T	M_{BGMed}	M_{TGMed}
0	1	0.056	0.049	0.052	0.053
	1.2	0.053	0.052	0.050	0.058
	1.4	0.065	0.054	0.069	0.059
	1.6	0.090	0.076	0.095	0.087
	1.8	0.119	0.118	0.109	0.109
	2	0.063	0.065	0.045	0.046
	3	0.068	0.055	0.047	0.044
0.2	1	0.071	0.065	0.045	0.042
	1.2	0.086	0.088	0.065	0.062
	1.4	0.136	0.123	0.091	0.094
	1.6	0.086	0.078	0.045	0.045
	1.8	0.091	0.084	0.045	0.041
	2	0.096	0.092	0.038	0.037
	3	0.112	0.106	0.049	0.046
0.5	1	0.161	0.140	0.086	0.085
	1.2	0.116	0.102	0.053	0.055
	1.4	0.121	0.104	0.050	0.046
	1.6	0.135	0.137	0.040	0.033
	1.8	0.161	0.148	0.051	0.051
	2	0.197	0.187	0.075	0.077
	3	0.176	0.148	0.063	0.061
0.7	1	0.177	0.153	0.061	0.053
	1.2	0.194	0.182	0.049	0.045
	1.4	0.208	0.188	0.048	0.052
	1.6	0.242	0.228	0.073	0.070
	1.8	0.231	0.211	0.076	0.075
	2	0.236	0.219	0.074	0.072
	3	0.244	0.241	0.057	0.058
0.9	1	0.272	0.243	0.057	0.051
	1.2	0.274	0.277	0.078	0.071
	1.4	0.620	0.602	0.172	0.181
	1.6	0.608	0.586	0.175	0.162
	1.8	0.552	0.533	0.131	0.124
	2	0.522	0.485	0.110	0.104
	3	0.494	0.466	0.100	0.101
No. of Shaded region		4	3	0	0

Shaded region indicate Power of test within [0.5-1.0]

Table O.86

Power of Test for variable, $p = 15, n = 10$

ρ	σ	M_B	M_T	M_{BGMed}	M_{TGMed}
0	1	0.045	0.051	0.055	0.044
	1.2	0.048	0.048	0.048	0.052
	1.4	0.056	0.058	0.100	0.091
	1.6	0.082	0.084	0.155	0.150
	1.8	0.123	0.116	0.235	0.241
	2	0.070	0.073	0.039	0.042
	3	0.070	0.069	0.039	0.036
0.2	1	0.086	0.088	0.056	0.054
	1.2	0.111	0.116	0.099	0.094
	1.4	0.146	0.143	0.183	0.175
	1.6	0.140	0.141	0.051	0.053
	1.8	0.150	0.152	0.054	0.047
	2	0.177	0.171	0.053	0.051
	3	0.193	0.186	0.079	0.080
0.5	1	0.210	0.209	0.153	0.158
	1.2	0.265	0.269	0.093	0.090
	1.4	0.276	0.278	0.091	0.085
	1.6	0.279	0.293	0.079	0.080
	1.8	0.279	0.271	0.084	0.096
	2	0.294	0.299	0.159	0.155
	3	0.439	0.452	0.157	0.161
0.7	1	0.446	0.442	0.150	0.151
	1.2	0.422	0.424	0.117	0.116
	1.4	0.399	0.409	0.111	0.107
	1.6	0.391	0.390	0.152	0.151
	1.8	0.627	0.637	0.253	0.260
	2	0.608	0.618	0.235	0.241
	3	0.545	0.557	0.176	0.181
0.9	1	0.517	0.506	0.147	0.148
	1.2	0.466	0.488	0.167	0.177
	1.4	0.988	0.988	0.785	0.790
	1.6	0.979	0.985	0.747	0.735
	1.8	0.930	0.933	0.530	0.538
	2	0.871	0.853	0.372	0.383
	3	0.773	0.785	0.290	0.287
No. of Shaded region		9	9	3	3

Shaded region indicate Power of test within [0.5-1.0]

Table O.87

Power of Test for variable, $p = 15, n = 15$

ρ	σ	M_B	M_T	M_{BGMed}	M_{TGMed}
0	1	0.045	0.054	0.045	0.051
	1.2	0.045	0.004	0.051	0.001
	1.4	0.055	0.008	0.119	0.002
	1.6	0.090	0.016	0.242	0.010
	1.8	0.111	0.047	0.346	0.038
	2	0.088	0.007	0.034	0.003
	3	0.092	0.006	0.037	0.003
0.2	1	0.100	0.018	0.047	0.010
	1.2	0.123	0.040	0.115	0.026
	1.4	0.168	0.074	0.237	0.050
	1.6	0.223	0.033	0.057	0.018
	1.8	0.229	0.038	0.060	0.021
	2	0.249	0.072	0.061	0.048
	3	0.252	0.099	0.090	0.076
0.5	1	0.268	0.133	0.190	0.110
	1.2	0.459	0.127	0.147	0.064
	1.4	0.466	0.140	0.123	0.081
	1.6	0.417	0.175	0.099	0.129
	1.8	0.410	0.212	0.103	0.158
	2	0.383	0.244	0.179	0.191
	3	0.710	0.311	0.301	0.216
0.7	1	0.691	0.324	0.260	0.214
	1.2	0.616	0.326	0.180	0.253
	1.4	0.546	0.325	0.148	0.280
	1.6	0.517	0.343	0.193	0.298
	1.8	0.891	0.559	0.478	0.411
	2	0.860	0.556	0.436	0.434
	3	0.769	0.510	0.275	0.418
0.9	1	0.686	0.476	0.214	0.419
	1.2	0.628	0.466	0.210	0.399
	1.4	1	0.998	0.984	0.994
	1.6	1	0.993	0.958	0.985
	1.8	0.994	0.960	0.807	0.936
	2	0.965	0.906	0.560	0.873
	3	0.918	0.822	0.382	0.802
No. of Shaded region		15	8	5	5

Shaded region indicate Power of test within [0.5-1.0]

Table O.88

Power of Test for variable, $p = 15, n = 20$

ρ	σ	M	M_{GMed}	M_B	M_T	M_{BGMed}	M_{TGMEd}
0	1	0.056	0.043	0.045	0.511	0.053	0.052
	1.2	0.095	0.093	0.043	0.010	0.069	0.003
	1.4	0.356	0.432	0.062	0.019	0.180	0.009
	1.6	0.852	0.926	0.084	0.039	0.319	0.025
	1.8	1	1	0.121	0.064	0.444	0.051
	2	0.050	0.049	0.099	0.029	0.032	0.012
	3	0.093	0.092	0.106	0.035	0.037	0.014
0.2	1	0.270	0.355	0.129	0.061	0.057	0.033
	1.2	0.735	0.850	0.154	0.078	0.150	0.052
	1.4	1	1	0.183	0.119	0.311	0.083
	1.6	0.067	0.069	0.319	0.142	0.084	0.080
	1.8	0.109	0.097	0.324	0.159	0.084	0.090
	2	0.236	0.311	0.312	0.187	0.070	0.134
	3	0.624	0.789	0.310	0.216	0.101	0.156
0.5	1	1	1	0.318	0.237	0.248	0.187
	1.2	0.092	0.081	0.644	0.424	0.235	0.284
	1.4	0.117	0.114	0.617	0.420	0.210	0.289
	1.6	0.209	0.285	0.548	0.396	0.137	0.311
	1.8	0.528	0.725	0.508	0.387	0.122	0.322
	2	1	1	0.472	0.392	0.232	0.326
	3	0.119	0.114	0.885	0.742	0.505	0.613
0.7	1	0.141	0.149	0.858	0.708	0.424	0.585
	1.2	0.221	0.317	0.769	0.632	0.238	0.536
	1.4	0.476	0.670	0.681	0.569	0.174	0.515
	1.6	1	1	0.634	0.528	0.216	0.474
	1.8	0.149	0.139	0.974	0.918	0.736	0.855
	2	0.174	0.183	0.962	0.900	0.651	0.822
	3	0.218	0.315	0.901	0.808	0.386	0.720
0.9	1	0.426	0.642	0.814	0.733	0.258	0.654
	1.2	0.999	1	0.730	0.658	0.244	0.617
	1.4	0.468	0.395	1	1	0.999	1
	1.6	0.462	0.405	1	1	0.999	1
	1.8	0.325	0.455	1	0.999	0.936	0.997
	2	0.360	0.622	0.993	0.988	0.720	0.979
	3	0.962	0.998	0.971	0.949	0.484	0.930
No. of Shaded region		11	14	19	15	7	14

Shaded region indicate Power of test within [0.5-1.0]

Table O. 89

Power of Test for variable, $p = 15, n = 25$

ρ	σ	M	M_{GMed}	M_B	M_T	M_{BGMed}	M_{TGMed}
0	1	0.057	0.052	0.053	0.057	0.049	0.059
	1.2	0.146	0.138	0.053	0.016	0.067	0.007
	1.4	0.755	0.731	0.066	0.030	0.213	0.014
	1.6	1	1	0.089	0.053	0.394	0.034
	1.8	1	1	0.116	0.091	0.536	0.063
	2	0.053	0.055	0.135	0.061	0.034	0.034
	3	0.132	0.127	0.145	0.060	0.033	0.040
0.2	1	0.629	0.610	0.162	0.093	0.054	0.063
	1.2	0.993	0.990	0.186	0.115	0.175	0.093
	1.4	1	1	0.204	0.143	0.352	0.126
	1.6	0.077	0.069	0.457	0.279	0.106	0.194
	1.8	0.146	0.151	0.442	0.294	0.088	0.204
	2	0.532	0.503	0.422	0.297	0.067	0.229
	3	0.975	0.971	0.392	0.289	0.102	0.247
0.5	1	1	1	0.370	0.303	0.255	0.259
	1.2	0.108	0.103	0.801	0.670	0.318	0.556
	1.4	0.181	0.168	0.787	0.646	0.253	0.537
	1.6	0.494	0.466	0.693	0.563	0.142	0.509
	1.8	0.953	0.940	0.626	0.525	0.121	0.460
	2	1	1	0.552	0.487	0.220	0.424
	3	0.168	0.153	0.966	0.916	0.624	0.873
0.7	1	0.241	0.224	0.952	0.890	0.521	0.842
	1.2	0.488	0.465	0.880	0.791	0.290	0.738
	1.4	0.915	0.905	0.798	0.725	0.192	0.676
	1.6	1	1	0.720	0.664	0.234	0.621
	1.8	0.237	0.220	0.997	0.990	0.864	0.976
	2	0.298	0.282	0.994	0.979	0.779	0.969
	3	0.483	0.476	0.963	0.925	0.463	0.885
0.9	1	0.888	0.880	0.898	0.848	0.286	0.812
	1.2	1	1	0.832	0.769	0.251	0.746
	1.4	0.710	0.705	1	1	1	1
	1.6	0.712	0.696	1	1	1	1
	1.8	0.673	0.660	1	1	0.981	0.999
	2	0.830	0.821	0.999	1	0.822	0.996
	3	1	1	0.99	0.980	0.548	0.978
No. of Shaded region		20	20	20	19	10	18

Shaded region indicate Power of test within [0.5-1.0]

Table O.90

Power of Test for variable, $p = 15, n = 50$

ρ	σ	M	M_{GMed}	M_B	M_T	M_{BGMed}	M_{TGMed}
0	1	0.051	0.050	0.053	0.057	0.054	0.057
	1.2	0.429	0.416	0.054	0.030	0.092	0.010
	1.4	1	1	0.065	0.048	0.415	0.014
	1.6	1	1	0.096	0.072	0.683	0.037
	1.8	1	1	0.115	0.099	0.784	0.062
	2	0.045	0.046	0.268	0.204	0.021	0.090
	3	0.291	0.287	0.280	0.227	0.016	0.096
0.2	1	0.995	0.995	0.290	0.236	0.060	0.134
	1.2	1	1	0.279	0.249	0.300	0.163
	1.4	1	1	0.275	0.262	0.569	0.185
	1.6	0.101	0.102	0.835	0.783	0.184	0.613
	1.8	0.299	0.322	0.824	0.770	0.128	0.601
	2	0.972	0.973	0.727	0.676	0.053	0.524
	3	1	1	0.645	0.613	0.104	0.506
0.5	1	1	1	0.589	0.549	0.378	0.454
	1.2	0.232	0.240	0.995	0.993	0.688	0.972
	1.4	0.418	0.433	0.991	0.983	0.517	0.957
	1.6	0.937	0.941	0.944	0.935	0.185	0.873
	1.8	1	1	0.891	0.872	0.109	0.791
	2	1	1	0.807	0.790	0.260	0.729
	3	0.477	0.497	1	1	0.952	1
0.7	1	0.611	0.611	1	1	0.892	0.999
	1.2	0.915	0.914	0.994	0.995	0.477	0.982
	1.4	1	1	0.977	0.970	0.216	0.940
	1.6	1	1	0.928	0.909	0.239	0.883
	1.8	0.754	0.761	1	1	0.998	1
	2	0.779	0.787	1	1	0.990	1
	3	0.912	0.907	1	1	0.756	0.999
0.9	1	1	1	0.994	0.996	0.399	0.987
	1.2	1	0.999	0.978	0.974	0.262	0.958
	1.4	1	0.999	0.999	1	1	1
	1.6	0.999	1	1	0.999	1	0.999
	1.8	0.990	0.992	0.999	1	0.999	1
	2	0.996	0.998	1	0.999	0.979	1
	3	1	1	1	1	0.715	1
No. of Shaded region		26	26	25	25	18	24

Shaded region indicate Power of test within [0.5-1.0]

Table O.91

Power of Test for variable, $p = 20, n=5$

ρ	σ	M_B	M_T	M_{BGMed}	M_{TGMed}
0	1	0.052	0.045	0.049	0.052
	1.2	0.054	0.046	0.053	0.049
	1.4	0.060	0.058	0.058	0.064
	1.6	0.089	0.077	0.095	0.088
	1.8	0.134	0.122	0.122	0.126
	2	0.067	0.064	0.039	0.036
	3	0.068	0.058	0.038	0.035
0.2	1	0.079	0.073	0.038	0.038
	1.2	0.096	0.088	0.065	0.063
	1.4	0.143	0.140	0.097	0.104
	1.6	0.083	0.083	0.043	0.044
	1.8	0.094	0.085	0.039	0.036
	2	0.104	0.091	0.037	0.034
	3	0.126	0.119	0.050	0.047
0.5	1	0.174	0.175	0.087	0.083
	1.2	0.132	0.124	0.048	0.047
	1.4	0.135	0.134	0.043	0.042
	1.6	0.157	0.149	0.039	0.040
	1.8	0.179	0.160	0.044	0.049
	2	0.208	0.208	0.080	0.087
	3	0.181	0.181	0.068	0.056
0.7	1	0.211	0.180	0.058	0.054
	1.2	0.227	0.203	0.043	0.044
	1.4	0.227	0.210	0.052	0.048
	1.6	0.254	0.260	0.080	0.075
	1.8	0.267	0.237	0.075	0.072
	2	0.264	0.254	0.073	0.068
	3	0.274	0.250	0.055	0.057
0.9	1	0.280	0.267	0.056	0.049
	1.2	0.302	0.302	0.087	0.078
	1.4	0.706	0.700	0.206	0.202
	1.6	0.695	0.684	0.191	0.193
	1.8	0.614	0.595	0.147	0.142
	2	0.559	0.542	0.114	0.112
	3	0.514	0.501	0.112	0.115
No. of Shaded region		5	5	0	0

Shaded region indicate Power of test within [0.5-1.0]

Table O.92

Power of Test for variable, $p = 20, n = 10$

ρ	σ	M_B	M_T	M_{BGMed}	M_{TGMed}
0	1	0.048	0.048	0.048	0.047
	1.2	0.047	0.045	0.050	0.051
	1.4	0.066	0.059	0.091	0.095
	1.6	0.095	0.092	0.178	0.176
	1.8	0.128	0.127	0.261	0.257
	2	0.078	0.067	0.039	0.035
	3	0.081	0.079	0.034	0.034
0.2	1	0.093	0.090	0.049	0.047
	1.2	0.131	0.114	0.102	0.111
	1.4	0.166	0.165	0.192	0.193
	1.6	0.160	0.155	0.051	0.051
	1.8	0.166	0.165	0.057	0.049
	2	0.192	0.184	0.061	0.056
	3	0.210	0.192	0.086	0.086
0.5	1	0.234	0.225	0.167	0.170
	1.2	0.324	0.314	0.119	0.101
	1.4	0.323	0.325	0.099	0.092
	1.6	0.320	0.322	0.085	0.086
	1.8	0.317	0.313	0.094	0.098
	2	0.331	0.318	0.160	0.160
	3	0.544	0.532	0.189	0.186
0.7	1	0.527	0.503	0.170	0.173
	1.2	0.467	0.467	0.131	0.142
	1.4	0.440	0.428	0.128	0.117
	1.6	0.411	0.407	0.167	0.167
	1.8	0.752	0.731	0.321	0.305
	2	0.704	0.705	0.290	0.264
	3	0.616	0.613	0.196	0.194
0.9	1	0.570	0.542	0.168	0.169
	1.2	0.502	0.502	0.189	0.186
	1.4	0.997	0.999	0.889	0.895
	1.6	0.996	0.995	0.841	0.847
	1.8	0.962	0.960	0.607	0.601
	2	0.886	0.895	0.419	0.406
	3	0.811	0.803	0.310	0.300
No. of Shaded region		12	12	3	3

Shaded region indicate Power of test within [0.5-1.0]

Table O.93

Power of Test for variable, $p = 20, n = 15$

ρ	σ	M_B	M_T	M_{BGMed}	M_{TGMed}
0	1	0.045	0.056	0.047	0.049
	1.2	0.046	0.049	0.057	0.057
	1.4	0.060	0.073	0.139	0.138
	1.6	0.092	0.101	0.284	0.268
	1.8	0.128	0.147	0.391	0.371
	2	0.095	0.092	0.040	0.034
	3	0.095	0.110	0.039	0.034
0.2	1	0.127	0.118	0.059	0.057
	1.2	0.145	0.149	0.150	0.146
	1.4	0.182	0.191	0.282	0.279
	1.6	0.258	0.288	0.085	0.083
	1.8	0.265	0.290	0.081	0.072
	2	0.286	0.289	0.071	0.068
	3	0.289	0.294	0.108	0.106
0.5	1	0.303	0.299	0.223	0.216
	1.2	0.546	0.583	0.200	0.194
	1.4	0.549	0.560	0.194	0.175
	1.6	0.478	0.494	0.116	0.130
	1.8	0.444	0.447	0.116	0.124
	2	0.427	0.424	0.195	0.205
	3	0.810	0.829	0.433	0.422
0.7	1	0.790	0.801	0.377	0.350
	1.2	0.691	0.687	0.219	0.209
	1.4	0.619	0.610	0.164	0.163
	1.6	0.561	0.560	0.221	0.209
	1.8	0.950	0.956	0.655	0.640
	2	0.930	0.934	0.582	0.570
	3	0.829	0.831	0.342	0.324
0.9	1	0.727	0.756	0.239	0.231
	1.2	0.662	0.656	0.236	0.223
	1.4	1	1	0.998	0.998
	1.6	1	1	0.993	0.992
	1.8	0.998	0.998	0.889	0.880
	2	0.980	0.982	0.649	0.635
	3	0.928	0.928	0.425	0.436
No. of Shaded region		17	17	7	7

Shaded region indicate Power of test within [0.5-1.0]

Table O.94

Power of Test for variable, $p = 20, n = 20$

ρ	σ	M_B	M_T	M_{BGMed}	M_{TGMed}
0	1	0.052	0.045	0.044	0.051
	1.2	0.047	0.043	0.055	0.067
	1.4	0.063	0.060	0.175	0.195
	1.6	0.099	0.097	0.361	0.374
	1.8	0.134	0.124	0.464	0.485
	2	0.115	0.116	0.039	0.044
	3	0.127	0.119	0.035	0.038
0.2	1	0.156	0.136	0.058	0.062
	1.2	0.173	0.165	0.171	0.183
	1.4	0.194	0.190	0.320	0.336
	1.6	0.397	0.373	0.102	0.126
	1.8	0.395	0.377	0.093	0.104
	2	0.366	0.356	0.069	0.081
	3	0.354	0.352	0.110	0.124
0.5	1	0.353	0.331	0.248	0.261
	1.2	0.765	0.746	0.324	0.358
	1.4	0.735	0.714	0.258	0.287
	1.6	0.637	0.621	0.150	0.157
	1.8	0.562	0.551	0.127	0.133
	2	0.509	0.493	0.226	0.235
	3	0.954	0.943	0.628	0.656
0.7	1	0.929	0.917	0.535	0.570
	1.2	0.830	0.820	0.281	0.302
	1.4	0.736	0.730	0.191	0.206
	1.6	0.671	0.648	0.230	0.238
	1.8	0.995	0.993	0.868	0.877
	2	0.988	0.987	0.770	0.785
	3	0.933	0.924	0.440	0.465
0.9	1	0.856	0.847	0.276	0.298
	1.2	0.770	0.777	0.258	0.262
	1.4	0.999	1	1	1
	1.6	1	1	1	1
	1.8	1	1	0.975	0.976
	2	0.996	0.996	0.767	0.787
	3	0.971	0.972	0.515	0.510
No. of Shaded region		20	19	9	9

Shaded region indicate Power of test within [0.5-1.0]

Table O.95

Power of Test for variable, $p = 20, n = 25$

ρ	σ	M	M_{GMed}	M_B	M_T	M_{BGMed}	M_{TGMed}
0	1	0.048	0.045	0.045	0.051	0.056	0.058
	1.2	0.111	0.104	0.043	0.052	0.066	0.076
	1.4	0.432	0.508	0.065	0.073	0.230	0.248
	1.6	0.934	0.974	0.092	0.104	0.431	0.450
	1.8	1	1	0.127	0.140	0.555	0.558
	2	0.058	0.054	0.139	0.151	0.035	0.045
	3	0.108	0.111	0.144	0.154	0.035	0.036
0.2	1	0.316	0.423	0.172	0.184	0.064	0.066
	1.2	0.829	0.923	0.182	0.210	0.211	0.209
	1.4	1	1	0.226	0.221	0.383	0.404
	1.6	0.076	0.061	0.499	0.516	0.163	0.163
	1.8	0.121	0.119	0.488	0.511	0.130	0.132
	2	0.261	0.380	0.458	0.466	0.086	0.091
	3	0.709	0.869	0.421	0.437	0.127	0.129
0.5	1	1	1	0.392	0.398	0.290	0.298
	1.2	0.093	0.092	0.887	0.890	0.491	0.490
	1.4	0.138	0.145	0.844	0.863	0.370	0.392
	1.6	0.244	0.357	0.738	0.748	0.184	0.186
	1.8	0.609	0.802	0.645	0.666	0.134	0.137
	2	1	1	0.584	0.600	0.240	0.246
	3	0.139	0.139	0.987	0.988	0.803	0.815
0.7	1	0.170	0.188	0.976	0.980	0.701	0.710
	1.2	0.239	0.354	0.913	0.914	0.360	0.358
	1.4	0.529	0.783	0.817	0.828	0.216	0.215
	1.6	1	1	0.735	0.745	0.256	0.243
	1.8	0.189	0.185	1	0.999	0.953	0.960
	2	0.206	0.241	0.998	0.998	0.909	0.911
	3	0.245	0.378	0.970	0.978	0.571	0.571
0.9	1	0.478	0.748	0.922	0.921	0.338	0.340
	1.2	1	1	0.842	0.841	0.278	0.289
	1.4	0.591	0.535	1	1	1	1
	1.6	0.539	0.549	0.999	1	1	1
	1.8	0.355	0.561	1	0.999	0.997	0.995
	2	0.374	0.728	1	1	0.863	0.870
	3	0.988	1	0.990	0.993	0.573	0.569
No. of Shaded region		14	18	20	22	11	11

Shaded region indicate Power of test within [0.5-1.0]

Table O.96

Power of Test for variable, $p = 20, n = 50$

ρ	σ	M	M_{GMed}	M_B	M_T	M_{BGMed}	M_{TGMed}
0	1	0.054	0.055	0.042	0.049	0.052	0.051
	1.2	0.402	0.418	0.048	0.057	0.105	0.088
	1.4	0.999	1	0.074	0.066	0.075	0.163
	1.6	1	1	0.101	0.100	0.042	0.067
	1.8	1	1	0.132	0.139	0.124	0.060
	2	0.080	0.057	0.301	0.304	0.052	0.059
	3	0.360	0.318	0.306	0.316	0.069	0.068
0.2	1	0.992	0.995	0.309	0.314	0.070	0.074
	1.2	1	1	0.301	0.302	0.355	0.344
	1.4	1	1	0.300	0.304	0.621	0.615
	1.6	0.201	0.125	0.910	0.911	0.397	0.368
	1.8	0.436	0.352	0.879	0.895	0.253	0.256
	2	0.963	0.965	0.764	0.761	0.073	0.076
	3	1	1	0.677	0.685	0.134	0.136
0.5	1	1	1	0.606	0.608	0.417	0.414
	1.2	0.463	0.282	0.999	0.999	0.901	0.894
	1.4	0.616	0.479	0.996	0.997	0.782	0.766
	1.6	0.926	0.936	0.969	0.967	0.268	0.282
	1.8	1	1	0.911	0.912	0.132	0.133
	2	1	1	0.833	0.836	0.300	0.294
	3	0.768	0.528	1	1	0.997	0.997
0.7	1	0.808	0.649	1	1	0.983	0.982
	1.2	0.912	0.908	1	0.997	0.617	0.614
	1.4	1	1	0.980	0.980	0.266	0.271
	1.6	1	1	0.939	0.940	0.267	0.270
	1.8	0.945	0.789	1	1	1	1
	2	0.935	0.816	1	1	1	0.999
	3	0.921	0.915	1	1	0.885	0.887
0.9	1	1	1	0.998	0.997	0.479	0.475
	1.2	1	1	0.978	0.984	0.298	0.305
	1.4	0.999	0.999	1	1	1	0.999
	1.6	1	1	1	1	1	1
	1.8	0.999	0.993	1	0.999	1	0.999
	2	0.997	0.998	0.999	1	0.993	0.992
	3	1	1	1	0.999	0.754	0.744
No. of Shaded region		28	28	25	25	14	14

Shaded region indicate Power of test within [0.5-1.0]

Appendix P.

Table P.97

Power of Test for variable, $p = 25, n = 5$

ρ	σ	M_B	M_T	M_{BGMed}	M_{TGMed}
0	1	0.057	0.046	0.055	0.048
	1.2	0.059	0.043	0.054	0.049
	1.4	0.070	0.059	0.068	0.064
	1.6	0.109	0.097	0.112	0.117
	1.8	0.156	0.142	0.140	0.140
	2	0.062	0.052	0.045	0.043
	3	0.067	0.063	0.043	0.041
0.2	1	0.080	0.070	0.044	0.041
	1.2	0.106	0.105	0.076	0.069
	1.4	0.154	0.140	0.111	0.110
	1.6	0.089	0.083	0.045	0.043
	1.8	0.096	0.091	0.038	0.041
	2	0.117	0.108	0.042	0.037
	3	0.132	0.135	0.053	0.052
0.5	1	0.197	0.186	0.095	0.100
	1.2	0.145	0.126	0.051	0.054
	1.4	0.151	0.125	0.046	0.041
	1.6	0.169	0.155	0.045	0.042
	1.8	0.198	0.162	0.054	0.049
	2	0.227	0.220	0.095	0.089
	3	0.210	0.212	0.064	0.065
0.7	1	0.227	0.209	0.057	0.058
	1.2	0.238	0.228	0.051	0.050
	1.4	0.246	0.231	0.055	0.056
	1.6	0.270	0.266	0.088	0.093
	1.8	0.302	0.295	0.090	0.087
	2	0.298	0.294	0.082	0.084
	3	0.317	0.294	0.062	0.062
0.9	1	0.325	0.305	0.059	0.060
	1.2	0.332	0.319	0.088	0.091
	1.4	0.815	0.773	0.244	0.241
	1.6	0.766	0.748	0.215	0.225
	1.8	0.663	0.645	0.151	0.163
	2	0.588	0.571	0.133	0.125
	3	0.532	0.529	0.120	0.126
No. of Shaded region		5	5	0	0

Shaded region indicate Power of test within [0.5-1.0]

Table P.98

Power of Test for variable, $p = 25, n = 10$

ρ	σ	M_B	M_T	M_{BGMed}	M_{TGMed}
0	1	0.047	0.047	0.049	0.049
	1.2	0.049	0.051	0.054	0.053
	1.4	0.061	0.061	0.109	0.113
	1.6	0.101	0.100	0.196	0.194
	1.8	0.148	0.143	0.280	0.287
	2	0.078	0.073	0.033	0.040
	3	0.084	0.084	0.035	0.035
0.2	1	0.106	0.103	0.053	0.054
	1.2	0.128	0.138	0.112	0.112
	1.4	0.176	0.177	0.211	0.217
	1.6	0.178	0.185	0.058	0.056
	1.8	0.190	0.187	0.055	0.060
	2	0.210	0.207	0.059	0.059
	3	0.216	0.214	0.092	0.089
0.5	1	0.254	0.249	0.179	0.182
	1.2	0.376	0.372	0.119	0.125
	1.4	0.380	0.374	0.114	0.121
	1.6	0.351	0.361	0.091	0.087
	1.8	0.345	0.340	0.100	0.106
	2	0.349	0.349	0.169	0.170
	3	0.620	0.594	0.235	0.243
0.7	1	0.598	0.598	0.205	0.214
	1.2	0.514	0.520	0.147	0.151
	1.4	0.466	0.467	0.127	0.133
	1.6	0.441	0.438	0.184	0.183
	1.8	0.805	0.816	0.394	0.409
	2	0.784	0.772	0.345	0.361
	3	0.659	0.667	0.212	0.233
0.9	1	0.591	0.583	0.175	0.178
	1.2	0.544	0.533	0.199	0.187
	1.4	1	1	0.948	0.950
	1.6	0.999	0.998	0.913	0.914
	1.8	0.973	0.974	0.677	0.676
	2	0.922	0.908	0.445	0.451
	3	0.834	0.821	0.335	0.338
No. of Shaded region		13	13	4	4

Shaded region indicate Power of test within [0.5-1.0]

Table P.99

Power of Test for variable, $p = 25, n = 15$

ρ	σ	M_B	M_T	M_{BGMed}	M_{TGMed}
0	1	0.050	0.042	0.048	0.053
	1.2	0.048	0.051	0.051	0.066
	1.4	0.066	0.068	0.151	0.159
	1.6	0.098	0.111	0.296	0.307
	1.8	0.144	0.146	0.394	0.394
	2	0.113	0.101	0.040	0.041
	3	0.112	0.109	0.036	0.039
0.2	1	0.137	0.135	0.062	0.063
	1.2	0.161	0.161	0.151	0.158
	1.4	0.188	0.197	0.294	0.297
	1.6	0.315	0.317	0.087	0.097
	1.8	0.315	0.315	0.079	0.090
	2	0.312	0.325	0.072	0.075
	3	0.317	0.306	0.108	0.119
0.5	1	0.315	0.302	0.237	0.232
	1.2	0.649	0.644	0.261	0.274
	1.4	0.616	0.617	0.207	0.235
	1.6	0.541	0.525	0.129	0.142
	1.8	0.486	0.481	0.133	0.131
	2	0.453	0.437	0.210	0.209
	3	0.886	0.885	0.516	0.523
0.7	1	0.852	0.865	0.429	0.446
	1.2	0.735	0.725	0.247	0.252
	1.4	0.652	0.653	0.175	0.172
	1.6	0.562	0.580	0.217	0.224
	1.8	0.981	0.980	0.747	0.766
	2	0.966	0.964	0.659	0.678
	3	0.871	0.867	0.380	0.376
0.9	1	0.769	0.771	0.244	0.257
	1.2	0.684	0.684	0.246	0.249
	1.4	1	1	1	1
	1.6	1	1	0.998	0.999
	1.8	0.999	1	0.933	0.935
	2	0.984	0.986	0.677	0.689
	3	0.934	0.941	0.445	0.439
No. of Shaded region		20	20	7	7

Shaded region indicate Power of test within [0.5-1.0]

Table P.100

Power of Test for variable, $p = 25, n = 20$

ρ	σ	M_B	M_T	M_{BGMed}	M_{TGMed}
0	1	0.055	0.051	0.045	0.051
	1.2	0.053	0.053	0.060	0.061
	1.4	0.072	0.070	0.191	0.203
	1.6	0.110	0.109	0.388	0.391
	1.8	0.144	0.143	0.492	0.497
	2	0.140	0.136	0.039	0.040
	3	0.149	0.141	0.038	0.036
0.2	1	0.180	0.174	0.057	0.061
	1.2	0.186	0.186	0.195	0.204
	1.4	0.218	0.214	0.353	0.361
	1.6	0.473	0.460	0.140	0.159
	1.8	0.457	0.453	0.110	0.129
	2	0.416	0.413	0.076	0.087
	3	0.390	0.390	0.123	0.127
0.5	1	0.388	0.371	0.273	0.275
	1.2	0.850	0.843	0.418	0.441
	1.4	0.814	0.807	0.329	0.348
	1.6	0.680	0.684	0.170	0.177
	1.8	0.605	0.602	0.139	0.153
	2	0.544	0.533	0.254	0.246
	3	0.978	0.979	0.754	0.774
0.7	1	0.965	0.961	0.650	0.656
	1.2	0.865	0.874	0.320	0.342
	1.4	0.763	0.760	0.215	0.220
	1.6	0.685	0.687	0.236	0.258
	1.8	0.998	0.999	0.941	0.946
	2	0.996	0.995	0.863	0.879
	3	0.955	0.950	0.517	0.522
0.9	1	0.883	0.880	0.303	0.312
	1.2	0.786	0.797	0.271	0.277
	1.4	0.999	1	0.999	1
	1.6	1	1	1	1
	1.8	1	0.999	0.986	0.990
	2	0.998	0.998	0.827	0.835
	3	0.980	0.979	0.532	0.527
No. of Shaded region		20	20	10	10

Shaded region indicate Power of test within [0.5-1.0]

Table P.101

Power of Test for variable, $p = 25, n = 25$

ρ	σ	M_B	M_T	M_{BGMed}	M_{TGMed}
0	1	0.056	0.053	0.054	0.054
	1.2	0.060	0.050	0.063	0.080
	1.4	0.080	0.065	0.241	0.264
	1.6	0.118	0.101	0.473	0.488
	1.8	0.147	0.143	0.576	0.591
	2	0.170	0.166	0.044	0.054
	3	0.181	0.173	0.040	0.047
0.2	1	0.201	0.188	0.071	0.074
	1.2	0.220	0.206	0.213	0.229
	1.4	0.235	0.241	0.413	0.427
	1.6	0.598	0.572	0.209	0.217
	1.8	0.574	0.577	0.145	0.164
	2	0.503	0.492	0.086	0.087
	3	0.464	0.453	0.129	0.132
0.5	1	0.412	0.412	0.304	0.318
	1.2	0.941	0.933	0.592	0.624
	1.4	0.915	0.906	0.472	0.490
	1.6	0.788	0.785	0.205	0.212
	1.8	0.701	0.678	0.150	0.160
	2	0.614	0.608	0.269	0.271
	3	0.997	0.995	0.895	0.907
0.7	1	0.992	0.994	0.803	0.827
	1.2	0.935	0.927	0.401	0.428
	1.4	0.852	0.858	0.226	0.252
	1.6	0.761	0.758	0.260	0.275
	1.8	1	1	0.986	0.987
	2	1	1	0.961	0.963
	3	0.984	0.986	0.617	0.655
0.9	1	0.935	0.936	0.346	0.357
	1.2	0.862	0.861	0.292	0.297
	1.4	1	1	0.999	1
	1.6	0.999	1	1	1
	1.8	1	1	1	0.999
	2	1	1	0.902	0.914
	3	0.994	0.994	0.595	0.601
No. of Shaded region		23	22	12	12

Shaded region indicate Power of test within [0.5-1.0]

Table P.102

Power of Test for variable, $p = 25, n = 50$

ρ	σ	M	M_{GMed}	M_B	M_T	M_{BGMed}	M_{TGMed}
0	1	0.053	0.048	0.053	0.048	0.054	0.058
	1.2	0.352	0.375	0.055	0.054	0.090	0.087
	1.4	0.998	0.999	0.077	0.076	0.509	0.502
	1.6	1	0.999	0.099	0.104	0.770	0.778
	1.8	1	1	0.147	0.148	0.846	0.840
	2	0.078	0.063	0.358	0.357	0.063	0.061
	3	0.301	0.287	0.346	0.358	0.042	0.039
0.2	1	0.970	0.988	0.334	0.352	0.077	0.082
	1.2	1	1	0.335	0.340	0.380	0.380
	1.4	0.999	1	0.331	0.313	0.635	0.648
	1.6	0.180	0.121	0.961	0.956	0.547	0.552
	1.8	0.370	0.333	0.921	0.925	0.372	0.375
	2	0.908	0.951	0.809	0.809	0.098	0.096
	3	1	1	0.709	0.714	0.152	0.152
0.5	1	1	1	0.639	0.629	0.438	0.430
	1.2	0.391	0.280	1	1	0.969	0.974
	1.4	0.507	0.452	0.999	1	0.891	0.898
	1.6	0.845	0.911	0.980	0.980	0.340	0.360
	1.8	1	0.999	0.927	0.922	0.162	0.146
	2	1	1	0.849	0.843	0.320	0.325
	3	0.673	0.534	1	1	1	1
0.7	1	0.711	0.624	1	1	0.995	0.998
	1.2	0.807	0.886	1	0.998	0.717	0.727
	1.4	1	0.999	0.987	0.986	0.308	0.314
	1.6	1	1	0.946	0.944	0.293	0.294
	1.8	0.896	0.774	1	1	1	1
	2	0.882	0.787	1	1	1	0.999
	3	0.821	0.889	0.9992	0.999	0.939	0.936
0.9	1	0.999	0.999	0.999	0.998	0.529	0.534
	1.2	1	1	0.985	0.983	0.317	0.325
	1.4	0.999	1	1	1	1	1
	1.6	1	0.999	1	1	0.999	1
	1.8	0.989	0.986	1	1	0.999	1
	2	0.978	0.996	0.999	1	0.996	0.999
	3	1	1	1	0.999	0.773	0.765
No. of Shaded region		28	28	25	25	18	18

Shaded region indicate Power of test within [0.5-1.0]

Table P.103

Power of Test for variable, $p = 30, n = 5$

ρ	σ	M_B	M_T	M_{BGMed}	M_{TGMed}
0	1	0.054	0.051	0.056	0.047
	1.2	0.052	0.047	0.063	0.052
	1.4	0.066	0.057	0.077	0.071
	1.6	0.112	0.114	0.122	0.113
	1.8	0.154	0.154	0.135	0.153
	2	0.060	0.059	0.040	0.037
	3	0.070	0.060	0.042	0.043
0.2	1	0.082	0.079	0.050	0.041
	1.2	0.111	0.111	0.081	0.080
	1.4	0.172	0.165	0.118	0.114
	1.6	0.085	0.094	0.044	0.042
	1.8	0.100	0.096	0.041	0.041
	2	0.120	0.122	0.040	0.039
	3	0.146	0.139	0.061	0.058
0.5	1	0.195	0.188	0.111	0.106
	1.2	0.138	0.149	0.057	0.050
	1.4	0.157	0.155	0.054	0.043
	1.6	0.182	0.180	0.041	0.037
	1.8	0.193	0.191	0.057	0.051
	2	0.238	0.236	0.095	0.093
	3	0.234	0.225	0.077	0.073
0.7	1	0.234	0.229	0.060	0.068
	1.2	0.261	0.239	0.057	0.051
	1.4	0.256	0.251	0.056	0.054
	1.6	0.287	0.289	0.099	0.097
	1.8	0.334	0.325	0.101	0.093
	2	0.343	0.341	0.083	0.083
	3	0.334	0.333	0.069	0.070
0.9	1	0.326	0.318	0.067	0.067
	1.2	0.331	0.327	0.100	0.087
	1.4	0.848	0.854	0.275	0.287
	1.6	0.819	0.805	0.266	0.257
	1.8	0.695	0.686	0.176	0.173
	2	0.613	0.614	0.133	0.134
	3	0.547	0.535	0.129	0.132
No. of Shaded region		5	5	0	0

Shaded region indicate Power of test within [0.5-1.0]

Table P.104

Power of Test for variable, $p = 30, n = 10$

ρ	σ	M_B	M_T	M_{BGMed}	M_{TGMed}
0	1	0.044	0.047	0.048	0.051
	1.2	0.050	0.053	0.055	0.050
	1.4	0.062	0.066	0.120	0.109
	1.6	0.118	0.114	0.223	0.208
	1.8	0.145	0.161	0.304	0.293
	2	0.073	0.087	0.037	0.038
	3	0.076	0.088	0.030	0.029
0.2	1	0.109	0.113	0.056	0.051
	1.2	0.141	0.138	0.120	0.118
	1.4	0.184	0.191	0.220	0.225
	1.6	0.191	0.201	0.063	0.061
	1.8	0.205	0.205	0.063	0.055
	2	0.214	0.229	0.063	0.060
	3	0.238	0.236	0.094	0.094
0.5	1	0.269	0.262	0.188	0.183
	1.2	0.413	0.414	0.135	0.118
	1.4	0.402	0.414	0.122	0.126
	1.6	0.375	0.380	0.100	0.084
	1.8	0.353	0.371	0.112	0.104
	2	0.365	0.351	0.183	0.178
	3	0.675	0.682	0.276	0.265
0.7	1	0.634	0.648	0.243	0.225
	1.2	0.540	0.555	0.159	0.150
	1.4	0.491	0.503	0.142	0.137
	1.6	0.465	0.455	0.183	0.178
	1.8	0.857	0.874	0.466	0.433
	2	0.823	0.832	0.409	0.375
	3	0.689	0.705	0.241	0.224
0.9	1	0.611	0.609	0.192	0.175
	1.2	0.543	0.557	0.207	0.195
	1.4	1	1	0.977	0.973
	1.6	0.999	1	0.949	0.944
	1.8	0.980	0.985	0.725	0.714
	2	0.921	0.931	0.477	0.477
	3	0.834	0.840	0.332	0.330
No. of Shaded region		13	14	3	3

Shaded region indicate Power of test within [0.5-1.0]

Table P.105

Power of Test for variable, $p=30$, $n = 15$

ρ	σ	M_B	M_T	M_{BGMed}	M_{TGMed}
0	1	0.055	0.060	0.045	0.047
	1.2	0.056	0.061	0.058	0.055
	1.4	0.070	0.083	0.170	0.170
	1.6	0.119	0.119	0.308	0.319
	1.8	0.151	0.157	0.410	0.419
	2	0.111	0.126	0.041	0.040
	3	0.123	0.126	0.038	0.040
0.2	1	0.148	0.156	0.062	0.058
	1.2	0.172	0.174	0.172	0.170
	1.4	0.213	0.213	0.295	0.311
	1.6	0.350	0.377	0.109	0.113
	1.8	0.349	0.382	0.098	0.098
	2	0.332	0.355	0.080	0.076
	3	0.330	0.334	0.117	0.117
0.5	1	0.337	0.340	0.246	0.243
	1.2	0.716	0.732	0.309	0.305
	1.4	0.673	0.701	0.252	0.270
	1.6	0.579	0.581	0.145	0.141
	1.8	0.506	0.516	0.134	0.131
	2	0.461	0.473	0.225	0.214
	3	0.936	0.940	0.597	0.600
0.7	1	0.893	0.909	0.508	0.504
	1.2	0.766	0.775	0.266	0.256
	1.4	0.675	0.676	0.191	0.182
	1.6	0.580	0.606	0.241	0.236
	1.8	0.990	0.993	0.848	0.840
	2	0.980	0.984	0.746	0.755
	3	0.885	0.891	0.402	0.417
0.9	1	0.787	0.803	0.259	0.265
	1.2	0.707	0.707	0.268	0.256
	1.4	1	1	1	1
	1.6	1	1	0.999	0.999
	1.8	1	1	0.952	0.953
	2	0.988	0.991	0.701	0.711
	3	0.949	0.947	0.476	0.457
No. of Shaded region		19	19	9	9

Shaded region indicate Power of test within [0.5-1.0]

Table P.106

Power of Test for variable, $p = 30, n = 20$

ρ	σ	M_B	M_T	M_{BGMed}	M_{TGMed}
0	1	0.053	0.053	0.049	0.049
	1.2	0.055	0.046	0.057	0.058
	1.4	0.071	0.064	0.216	0.215
	1.6	0.115	0.108	0.402	0.392
	1.8	0.150	0.146	0.504	0.516
	2	0.147	0.134	0.044	0.047
	3	0.144	0.145	0.041	0.044
0.2	1	0.188	0.175	0.064	0.069
	1.2	0.201	0.186	0.198	0.204
	1.4	0.228	0.219	0.384	0.370
	1.6	0.492	0.500	0.170	0.183
	1.8	0.502	0.504	0.144	0.140
	2	0.445	0.438	0.091	0.088
	3	0.409	0.405	0.133	0.127
0.5	1	0.390	0.374	0.293	0.288
	1.2	0.884	0.888	0.520	0.523
	1.4	0.842	0.855	0.407	0.406
	1.6	0.717	0.700	0.186	0.198
	1.8	0.625	0.620	0.149	0.156
	2	0.563	0.554	0.261	0.251
	3	0.990	0.991	0.846	0.841
0.7	1	0.979	0.975	0.732	0.738
	1.2	0.886	0.870	0.345	0.362
	1.4	0.783	0.783	0.219	0.238
	1.6	0.690	0.700	0.253	0.251
	1.8	1	0.999	0.976	0.974
	2	0.998	0.999	0.929	0.926
	3	0.965	0.962	0.563	0.568
0.9	1	0.890	0.889	0.325	0.320
	1.2	0.802	0.802	0.296	0.294
	1.4	1	1	1	1
	1.6	1	1	1	1
	1.8	1	0.999	0.995	0.993
	2	0.999	0.998	0.847	0.845
	3	0.980	0.981	0.552	0.552
No. of Shaded region		21	22	12	12

Shaded region indicate Power of test within [0.5-1.0]

Table P.107

Power of Test for variable, $p = 30, n = 25$

ρ	σ	M_B	M_T	M_{BGMed}	M_{TGMed}
0	1	0.052	0.055	0.045	0.048
	1.2	0.051	0.053	0.059	0.081
	1.4	0.078	0.079	0.282	0.285
	1.6	0.120	0.124	0.492	0.502
	1.8	0.156	0.157	0.596	0.608
	2	0.177	0.192	0.054	0.059
	3	0.199	0.203	0.047	0.047
0.2	1	0.205	0.219	0.070	0.075
	1.2	0.209	0.225	0.243	0.250
	1.4	0.240	0.247	0.434	0.452
	1.6	0.667	0.682	0.265	0.274
	1.8	0.629	0.625	0.198	0.207
	2	0.539	0.547	0.093	0.092
	3	0.476	0.483	0.134	0.145
0.5	1	0.445	0.445	0.324	0.334
	1.2	0.965	0.965	0.695	0.724
	1.4	0.937	0.940	0.559	0.566
	1.6	0.813	0.805	0.230	0.250
	1.8	0.704	0.717	0.163	0.167
	2	0.626	0.637	0.269	0.287
	3	1	0.999	0.953	0.961
0.7	1	0.996	0.998	0.879	0.884
	1.2	0.946	0.948	0.461	0.462
	1.4	0.872	0.871	0.260	0.254
	1.6	0.769	0.773	0.279	0.283
	1.8	1	1	0.998	0.996
	2	1	1	0.984	0.985
	3	0.990	0.989	0.700	0.712
0.9	1	0.945	0.952	0.382	0.394
	1.2	0.873	0.873	0.310	0.314
	1.4	1	1	1	1
	1.6	0.999	0.999	1	1
	1.8	1	1	1	0.999
	2	0.999	1	0.922	0.933
	3	0.994	0.994	0.607	0.606
No. of Shaded region		23	23	13	14

Shaded region indicate Power of test within [0.5-1.0]

Table P.108

Power of Test for variable, $p = 30, n = 50$

ρ	σ	M	M_{GMed}	M_B	M_T	M_{BGMed}	M_{TGMed}
0	1	0.053	0.048	0.051	0.050	0.045	0.052
	1.2	0.352	0.375	0.044	0.054	0.082	0.095
	1.4	0.998	0.999	0.079	0.077	0.517	0.532
	1.6	1	0.999	0.114	0.111	0.788	0.791
	1.8	1	1	0.148	0.158	0.840	0.846
	2	0.078	0.063	0.398	0.389	0.079	0.079
	3	0.301	0.287	0.404	0.395	0.047	0.051
0.2	1	0.970	0.988	0.373	0.378	0.075	0.072
	1.2	1	1	0.349	0.346	0.402	0.403
	1.4	0.999	1	0.333	0.338	0.643	0.646
	1.6	0.180	0.121	0.976	0.978	0.676	0.679
	1.8	0.37	0.333	0.952	0.953	0.456	0.473
	2	0.908	0.951	0.841	0.834	0.107	0.105
	3	1	1	0.736	0.736	0.151	0.160
0.5	1	1	1	0.638	0.645	0.455	0.443
	1.2	0.391	0.280	1	1	0.992	0.991
	1.4	0.507	0.452	1	1	0.949	0.954
	1.6	0.845	0.911	0.979	0.987	0.400	0.401
	1.8	1	0.999	0.938	0.932	0.167	0.175
	2	1	1	0.853	0.851	0.334	0.337
	3	0.673	0.534	1	1	1	1
0.7	1	0.711	0.624	1	1	1	1
	1.2	0.807	0.886	1	1	0.787	0.775
	1.4	1	0.999	0.989	0.990	0.345	0.321
	1.6	1	1	0.955	0.952	0.301	0.303
	1.8	0.896	0.774	1	1	1	1
	2	0.882	0.787	1	0.999	1	0.999
	3	0.821	0.889	0.9998	1	0.962	0.963
0.9	1	0.999	0.999	1	0.998	0.570	0.557
	1.2	1	1	0.984	0.987	0.345	0.349
	1.4	0.999	1	1	0.999	1	0.999
	1.6	1	0.999	0.999	1	0.999	0.999
	1.8	0.989	0.986	0.999	1	1	1
	2	0.978	0.996	1	1	0.998	0.999
	3	1	1	1	1	0.786	0.799
No. of Shaded region		28	28	25	25	19	19

Shaded region indicate Power of test within [0.5-1.0]

Table P.109

Power of Test for variable, $p = 50, n = 5$

ρ	σ	M_B	M_T	M_{BGMed}	M_{TGMed}
0	1	0.051	0.048	0.058	0.047
	1.2	0.051	0.052	0.059	0.050
	1.4	0.070	0.072	0.101	0.093
	1.6	0.134	0.133	0.154	0.151
	1.8	0.178	0.175	0.187	0.181
	2	0.067	0.061	0.036	0.038
	3	0.069	0.066	0.039	0.036
0.2	1	0.086	0.093	0.053	0.051
	1.2	0.141	0.133	0.100	0.103
	1.4	0.196	0.188	0.135	0.135
	1.6	0.113	0.111	0.044	0.041
	1.8	0.118	0.114	0.043	0.043
	2	0.144	0.146	0.048	0.044
	3	0.169	0.181	0.073	0.077
0.5	1	0.241	0.222	0.130	0.128
	1.2	0.187	0.185	0.060	0.059
	1.4	0.200	0.197	0.055	0.053
	1.6	0.216	0.218	0.047	0.045
	1.8	0.227	0.226	0.066	0.063
	2	0.279	0.276	0.117	0.111
	3	0.320	0.298	0.077	0.081
0.7	1	0.322	0.312	0.082	0.074
	1.2	0.318	0.301	0.057	0.056
	1.4	0.295	0.302	0.064	0.064
	1.6	0.326	0.324	0.123	0.108
	1.8	0.467	0.466	0.120	0.115
	2	0.460	0.448	0.107	0.104
	3	0.401	0.400	0.069	0.079
0.9	1	0.363	0.365	0.074	0.072
	1.2	0.369	0.367	0.110	0.121
	1.4	0.966	0.967	0.430	0.429
	1.6	0.940	0.930	0.384	0.376
	1.8	0.794	0.793	0.232	0.225
	2	0.670	0.682	0.165	0.163
	3	0.598	0.601	0.157	0.152
No. of Shaded region		5	5	0	0

Shaded region indicate Power of test within [0.5-1.0]

Table P.110

Power of Test for variable, $p = 50, n = 10$

ρ	σ	M_B	M_T	M_{BGMed}	M_{TGMed}
0	1	0.052	0.054	0.048	0.046
	1.2	0.061	0.055	0.057	0.063
	1.4	0.090	0.085	0.154	0.151
	1.6	0.158	0.142	0.251	0.262
	1.8	0.169	0.177	0.329	0.338
	2	0.094	0.093	0.042	0.039
	3	0.113	0.109	0.033	0.038
0.2	1	0.144	0.131	0.061	0.058
	1.2	0.181	0.161	0.147	0.160
	1.4	0.207	0.217	0.256	0.252
	1.6	0.288	0.270	0.084	0.076
	1.8	0.305	0.295	0.073	0.069
	2	0.284	0.277	0.068	0.066
	3	0.294	0.283	0.113	0.123
0.5	1	0.322	0.311	0.222	0.225
	1.2	0.609	0.592	0.195	0.199
	1.4	0.580	0.556	0.172	0.169
	1.6	0.477	0.475	0.114	0.120
	1.8	0.426	0.429	0.126	0.124
	2	0.404	0.407	0.209	0.210
	3	0.873	0.868	0.427	0.403
0.7	1	0.817	0.813	0.356	0.355
	1.2	0.657	0.644	0.200	0.195
	1.4	0.572	0.561	0.158	0.168
	1.6	0.524	0.499	0.215	0.220
	1.8	0.977	0.969	0.658	0.670
	2	0.943	0.943	0.566	0.555
	3	0.796	0.783	0.309	0.302
0.9	1	0.692	0.679	0.212	0.209
	1.2	0.596	0.593	0.234	0.231
	1.4	1	1	0.999	0.999
	1.6	0.999	1	0.995	0.995
	1.8	0.998	0.996	0.843	0.845
	2	0.958	0.953	0.523	0.562
	3	0.862	0.864	0.367	0.361
No. of Shaded region		17	17	6	6

Shaded region indicate Power of test within [0.5-1.0]

Table P.111

Power of Test for variable, $p = 50, n = 15$

ρ	σ	M_B	M_T	M_{BGMed}	M_{TGMed}
0	1	0.051	0.049	0.048	0.041
	1.2	0.052	0.059	0.064	0.057
	1.4	0.084	0.092	0.210	0.202
	1.6	0.141	0.140	0.381	0.375
	1.8	0.166	0.177	0.463	0.465
	2	0.137	0.131	0.050	0.042
	3	0.143	0.146	0.051	0.040
0.2	1	0.169	0.182	0.079	0.067
	1.2	0.206	0.203	0.208	0.195
	1.4	0.240	0.238	0.349	0.339
	1.6	0.487	0.491	0.171	0.165
	1.8	0.475	0.473	0.139	0.131
	2	0.413	0.406	0.098	0.084
	3	0.388	0.372	0.145	0.125
0.5	1	0.368	0.365	0.286	0.272
	1.2	0.878	0.877	0.501	0.491
	1.4	0.821	0.825	0.393	0.385
	1.6	0.660	0.664	0.185	0.179
	1.8	0.574	0.567	0.161	0.148
	2	0.518	0.512	0.255	0.266
	3	0.991	0.988	0.844	0.830
0.7	1	0.975	0.974	0.709	0.708
	1.2	0.839	0.838	0.344	0.339
	1.4	0.713	0.725	0.220	0.227
	1.6	0.637	0.635	0.277	0.274
	1.8	1	1	0.977	0.970
	2	0.999	0.998	0.923	0.913
	3	0.943	0.942	0.540	0.517
0.9	1	0.841	0.841	0.317	0.307
	1.2	0.743	0.744	0.280	0.291
	1.4	0.999	1	1	1
	1.6	0.999	0.999	1	0.9994
	1.8	1	1	0.992	0.991
	2	0.996	0.996	0.792	0.792
	3	0.953	0.958	0.499	0.482
No. of Shaded region		20	20	10	9

Shaded region indicate Power of test within [0.5-1.0]

Table P.112

Power of Test for variable, $p = 50, n = 20$

ρ	σ	M_B	M_T	M_{BGMed}	M_{TGMed}
0	1	0.051	0.050	0.054	0.046
	1.2	0.061	0.050	0.062	0.068
	1.4	0.094	0.081	0.280	0.265
	1.6	0.144	0.135	0.473	0.476
	1.8	0.184	0.169	0.558	0.560
	2	0.191	0.175	0.055	0.061
	3	0.211	0.198	0.053	0.048
0.2	1	0.218	0.212	0.075	0.078
	1.2	0.241	0.229	0.238	0.236
	1.4	0.255	0.260	0.405	0.412
	1.6	0.708	0.683	0.301	0.293
	1.8	0.664	0.631	0.227	0.222
	2	0.523	0.521	0.107	0.099
	3	0.466	0.454	0.160	0.153
0.5	1	0.434	0.434	0.329	0.330
	1.2	0.975	0.980	0.787	0.774
	1.4	0.950	0.945	0.626	0.608
	1.6	0.801	0.804	0.257	0.255
	1.8	0.685	0.679	0.185	0.172
	2	0.605	0.586	0.304	0.288
	3	1	1	0.977	0.979
0.7	1	0.998	0.998	0.912	0.910
	1.2	0.934	0.933	0.480	0.473
	1.4	0.840	0.834	0.255	0.262
	1.6	0.736	0.727	0.297	0.297
	1.8	1	1	0.999	1
	2	1	1	0.993	0.991
	3	0.986	0.985	0.699	0.703
0.9	1	0.925	0.918	0.381	0.373
	1.2	0.825	0.828	0.322	0.321
	1.4	1	0.999	1	1
	1.6	1	1	1	1
	1.8	1	1	1	1
	2	1	1	0.916	0.917
	3	0.986	0.984	0.583	0.569
No. of Shaded region		23	23	13	13

Shaded region indicate Power of test within [0.5-1.0]

Table P.113

Power of Test for variable, $p=50$, $n = 25$

ρ	σ	M_B	M_T	M_{BGMed}	M_{TGMed}
0	1	0.047	0.045	0.050	0.048
	1.2	0.051	0.052	0.069	0.068
	1.4	0.094	0.087	0.315	0.300
	1.6	0.128	0.136	0.567	0.548
	1.8	0.171	0.178	0.647	0.619
	2	0.212	0.230	0.070	0.064
	3	0.236	0.230	0.061	0.044
0.2	1	0.245	0.247	0.082	0.078
	1.2	0.255	0.255	0.290	0.291
	1.4	0.281	0.290	0.476	0.468
	1.6	0.819	0.834	0.442	0.443
	1.8	0.766	0.769	0.310	0.304
	2	0.606	0.615	0.114	0.113
	3	0.522	0.533	0.159	0.163
0.5	1	0.482	0.473	0.361	0.362
	1.2	0.996	0.997	0.921	0.915
	1.4	0.985	0.988	0.785	0.778
	1.6	0.870	0.882	0.296	0.299
	1.8	0.761	0.763	0.181	0.180
	2	0.662	0.652	0.311	0.304
	3	1	1	0.999	0.997
0.7	1	1	1	0.980	0.978
	1.2	0.972	0.968	0.602	0.575
	1.4	0.892	0.901	0.297	0.299
	1.6	0.795	0.803	0.308	0.304
	1.8	1	1	1	1
	2	1	1	0.999	0.999
	3	0.996	0.997	0.815	0.811
0.9	1	0.964	0.966	0.437	0.443
	1.2	0.881	0.894	0.337	0.333
	1.4	1	1	1	1
	1.6	1	0.999	0.999	0.999
	1.8	0.999	1	1	1
	2	1	1	0.968	0.966
	3	0.996	0.9944	0.640	0.631
No. of Shaded region		24	24	15	15

Shaded region indicate Power of test within [0.5-1.0]

Table P.114

Power of Test for variable, $p=50$, $n = 50$

ρ	σ	M_B	M_T	M_{BGMed}	M_{TGMed}
0	1	0.050	0.055	0.057	0.043
	1.2	0.053	0.051	0.101	0.086
	1.4	0.079	0.087	0.603	0.605
	1.6	0.145	0.139	0.839	0.830
	1.8	0.184	0.180	0.868	0.874
	2	0.531	0.542	0.195	0.179
	3	0.511	0.530	0.098	0.088
0.2	1	0.422	0.437	0.092	0.080
	1.2	0.403	0.406	0.462	0.435
	1.4	0.363	0.380	0.700	0.689
	1.6	0.999	0.999	0.940	0.935
	1.8	0.990	0.992	0.766	0.741
	2	0.886	0.897	0.159	0.151
	3	0.791	0.775	0.188	0.168
0.5	1	0.692	0.681	0.508	0.504
	1.2	1	1	1	1
	1.4	1	1	0.998	0.998
	1.6	0.994	0.994	0.561	0.545
	1.8	0.953	0.955	0.214	0.203
	2	0.875	0.887	0.369	0.366
	3	1	1	1	1
0.7	1	1	1	1	1
	1.2	1	1	0.919	0.906
	1.4	0.994	0.994	0.414	0.408
	1.6	0.966	0.961	0.342	0.327
	1.8	1	1	0.999	1
	2	0.999	1	1	1
	3	1	1	0.993	0.994
0.9	1	1	1	0.665	0.659
	1.2	1	0.992	0.391	0.376
	1.4	1	0.999	1	1
	1.6	0.999	1	0.999	0.999
	1.8	1	1	1	1
	2	1	1	1	1
	3	1	1	0.824	0.822
No. of Shaded region		27	27	22	22

Shaded region indicate Power of test within [0.5-1.0]