

Muhammad Rauf Ahmad

**Analysis of High Dimensional
Repeated Measures Designs:
The One- and Two-Sample Test Statistics**



Cuvillier Verlag Göttingen

Analysis of High Dimensional Repeated Measures Designs: The One- and Two-Sample Test Statistics

Dissertation
zur Erlangung des Doktorgrades
der Mathematisch-Naturwissenschaftlichen Fakultäten
der Georg-August-Universität zu Göttingen

vorgelegt von
Muhammad Rauf Ahmad
aus
Faisalabad, Pakistan

GÖTTINGEN
(2008)

Bibliografische Information der Deutschen Nationalbibliothek

Die Deutsche Nationalbibliothek verzeichnet diese Publikation in der Deutschen Nationalbibliografie; detaillierte bibliografische Daten sind im Internet über <http://dnb.ddb.de> abrufbar.

1. Aufl. - Göttingen : Cuvillier, 2008

Zugl.: Göttingen, Univ., Diss., 2008

978-3-86727-734-1

D7

Referent: Prof. Dr. Edgar Brunner

Koreferent: Prof. Dr. Manfred Denker

Tag der Mündlichen Prüfung: 07.07.2008

© CUVILLIER VERLAG, Göttingen 2008

Nonnenstieg 8, 37075 Göttingen

Telefon: 0551-54724-0

Telefax: 0551-54724-21

www.cuvillier.de

Alle Rechte vorbehalten. Ohne ausdrückliche Genehmigung des Verlages ist es nicht gestattet, das Buch oder Teile daraus auf fotomechanischem Weg (Fotokopie, Mikrokopie) zu vervielfältigen.

1. Auflage, 2008

Gedruckt auf säurefreiem Papier

978-3-86727-734-1

Acknowledgement

First of all, I am indebted to my supervisor, PROF. DR. EDGAR BRUNNER, for his thoughtful supervision throughout the conduct of this research project. I gained a lot from his insight into the field, astute judgement and, particularly, his fastidious approach to address the problems. I also owe a lot to my co-supervisor, PROF. DR. MANFRED DENKER, for his continuous help and encouragement during my Ph.D. studies. His sagacious guidance has always been a big push to lead the work on this project further.

My heartiest thanks are also due to many other members of the center of statistics, both in the faculty and among the fellow students, from whom, I learned, from an iota to a lot. Among the faculty, special mention is to Prof. Dr. W. Zucchini, Prof. Dr. Axel Munk and Dr. Gudrun Freitag. Among the fellow members, Rada Matic deserves special mention. As I learned the rigorous mathematical theory of linear models from Prof. Brunner, from Rada I learned the tips and tricks of how to use this theory to solve cumbersome problems of linear models.

Arne Schillert is the name I should mention with special emphasis. The two years time we shared in the same office was full of his unforgettable conviviality during which I also benefitted from his R expertise. Among other fellows, I must name a few with whom I enjoyed a close, intimate collaboration. They include: Frank Konietzschke, Antonia Zapf, Yesilda Balavarca, J. P. Lozano, Benjamin Baker, Melanie Sohns and Rico Ihle.

Chapter 2 of my dissertation is a revision of the work Carola Werner did for her Diploma thesis. Although, the results are thoroughly revised, with much simpler and more elegant proofs, it was Carola who did the ab initio work on this one-sample normal case. I am grateful to her for her sincere help and, particularly, for several useful discussions I had with her.

I am also thankful to the members of the Department of Medical Statistics, particularly to the Secretary, Karola Riemenschneider. It was not without her untiring efforts and timely co-operation that my scholarship extensions with the DAAD and leave extensions back home always worked smoothly without any bureaucratic hassle.

Thanks are also due to the Higher Education Commission (HEC), Pakistan, for their financial support to my studies, in administrative collaboration with Deutscher Akademischer Austausch Dienst (DAAD), Germany.

My family deserves special thanks for their affection, passion, and most of all, their unrelenting patience. I am extremely indebted to their sincere prayers and best wishes which were always a source of resuscitation during the last four stressful years.

(Muhammad Rauf Ahmad)

To
Shah Jee

Scaffolding

Reacting to criticism concerning the lack of motivation in his writings, Carl Friedrich Gauss remarked that *the architects of great cathedrals do not obscure the beauty of their work by leaving the scaffolding in place after the construction has been completed.*

adapted from: Meyer, C. D. (2001). *Matrix Analysis and Applied Linear Algebra*. SIAM, PA.

Abstract

All models are wrong; only some are useful. (G. E. P. Box)

In this project, we have analyzed some useful models, based on an approximation introduced by G. E. P. Box; hence, the next few chapters map an odyssey wherein Box and his adage go hand in hand. In a nutshell, one- and two-sample test statistics are developed for the analysis of repeated measures designs when the dimension, d , can be large compared to the sample size, n ($d > n$).

The statistics do not depend on any specific structure of the covariance matrix and can be used in a variety of situations: they are valid for testing any general linear hypothesis, are equally applicable to the design set up of profile analysis and to the usual multivariate structure, are invariant to an orthogonal linear transformation, and are also valid when the data are not high dimensional.

The test statistics, a modification of the ANOVA-type statistic (Brunner, 2001), are based on Box's approximation (Box, 1954a), and follow a χ^2_f -distribution. The estimators, the building blocks of the test statistics, are composed of quadratic and symmetric bilinear forms, and are proved to be unbiased, L_2 -consistent and uniformly bounded in dimension, d . This last property of estimators helps us in the asymptotic derivations in that we need not let both n and d approach infinity. We let $n \rightarrow \infty$, while keep d fixed, such that the approximation of the distribution of the test statistic to the χ^2 distribution remains accurate when $d > n$, or even $d \gg n$.

The performance of the statistics is evaluated through simulations and it is shown that, for n as small as 10 or 20, the approximation is quite accurate, whatever be d . The statistic is also applied to a number of real data sets for numerical illustrations.

Contents

Acknowledgement	i
Abstract	vii
1 Introduction and Motivation	1
1.1 Introduction	1
1.2 Classical Approaches of Analysis	2
1.2.1 Univariate Analysis	3
1.2.2 Multivariate Analysis	4
1.3 The Case of High Dimensional Data	5
1.4 Motivating Examples	6
1.5 Review of Literature	8
2 The One-sample Statistic	11
2.1 The ANOVA-type Statistic	11
2.2 The Estimators	18
2.3 The Approximating Distribution	20
2.4 Simulation Results	22
2.4.1 Moments of Estimators	22
2.4.2 Level	23
2.4.3 Power	23
2.5 Applications	24
2.5.1 Analysis of Sleep Lab Data	24
2.5.2 Analysis of Amylase Data	26
3 The Two-sample Statistic	29
3.1 The Statistical Model	29
3.1.1 Model and Assumptions	29

3.1.2	Formulation of Hypotheses	30
3.1.3	Formulation of the Test Statistic	31
3.2	The Interaction Effect	31
3.2.1	The Covariance Matrix	32
3.2.2	Defining The Statistic	33
3.2.3	Case I: $n_1 \neq n_2$, $\Sigma_1 = \Sigma_2$	36
3.2.3.1	The Estimators	36
3.2.3.2	The Approximating Distribution	39
3.2.4	Case II: $n_1 = n_2$, $\Sigma_1 \neq \Sigma_2$	40
3.2.4.1	The Estimators	40
3.2.4.2	The Approximating Distribution	42
3.2.5	Simulation Results	44
3.2.5.1	Level	44
3.2.5.2	Power	46
3.3	The Time Effect	47
3.3.1	The Covariance Matrix	47
3.3.2	Defining The Statistic	47
3.3.3	Case I: $n_1 \neq n_2$, $\Sigma_1 = \Sigma_2$	49
3.3.4	Case II: $n_1 = n_2$, $\Sigma_1 \neq \Sigma_2$	49
3.3.5	Simulation Results	50
3.4	The Group Effect	52
3.5	Applications	55
3.5.1	Analysis of Complete Sleep Lab Data	55
3.5.2	Analysis of Wistar Rats Data	56
4	The Multivariate Case	59
5	Summary and Outlook	61
5.1	Summary and Conclusions	61
5.2	Future Outlook	64
A	Quadratic and Bilinear Forms	67
A.1	Moments of Quadratic Forms	68
A.2	Moments of Bilinear Forms	72
B	Some Miscellaneous Results	77

C	Data Sets	81
C.1	α -Amylase in Saliva	81
C.2	Body Weight of Wistar Rats	82
C.3	Sleep Disorder	83

List of Tables

2.1	THE GENERAL DATA TABLE: ONE-SAMPLE CASE	12
2.2	A_n WITH TRUE COVARIANCE STRUCTURES	15
2.3	PROPERTIES OF TRACES WITH COVARIANCE ESTIMATOR, $\hat{\Sigma}_n$	15
2.4	COMPARISON OF A_n , W_n AND H_n	16
2.5	PROPERTIES OF ESTIMATORS	22
2.6	ESTIMATED QUANTILES OF $\tilde{A}_n$: <i>Normal Distribution</i>	23
2.7	ANALYSIS OF SLEEP LAB DATA	25
2.8	ANALYSIS OF AMYLASE DATA	27
3.1	THE GENERAL DATA TABLE: TWO-SAMPLE CASE	29
3.2	ESTIMATED QUANTILES OF $\tilde{A}_N^{AB}$: <i>Normal Distribution</i>	44
3.3	ESTIMATED QUANTILES OF $\tilde{A}_N^B$: <i>Normal Distribution</i>	51
3.4	ANALYSIS OF COMPLETE SLEEP LAB DATA	55
3.5	ANALYSIS OF RATS DATA	57
5.1	SUMMARY ANOVA TABLE: ONE-SAMPLE CASE	61
5.2	SUMMARY ANOVA TABLE: TWO-SAMPLE CASE	62
C.1	MEASUREMENTS ON α -AMYLASE IN SALIVA	81
C.2	BODY WEIGHTS OF WISTAR RATS	82
C.3	PROSTAGLANDIN-D SYNTHASE ON SLEEP PROBANDS	83

List of Figures

1.1	TIME PROFILES FOR α -AMYLASE STUDY	6
1.2	TIME PROFILES FOR BODY WEIGHT STUDY	7
1.3	TIME PROFILES FOR SLEEP DISORDER STUDY	8
2.1	POWER CURVES FOR $\tilde{A}_n$: <i>Normal Distribution</i>	24
2.2	BAR PLOT FOR SLEEP LAB DATA	25
2.3	DAY $\times$ TIME INTERACTION PLOT FOR SLEEP LAB DATA . .	26
2.4	BAR PLOT FOR AMYLASE DATA	26
2.5	DAY $\times$ TIME INTERACTION PLOT FOR AMYLASE DATA . . .	27
3.1	POWER CURVES FOR $\tilde{A}_N^{AB}$: <i>Normal Distribution</i>	45
3.2	POWER CURVES FOR $\tilde{A}_N^B$: <i>Normal Distribution</i>	51
3.3	BAR PLOT FOR COMPLETE SLEEP LAB DATA	55
3.4	DAY $\times$ TIME INTERACTION PLOT FOR COMPLETE SLEEP LAB DATA	56
3.5	BAR PLOT FOR WISTAR RATS DATA	57
3.6	GROUP $\times$ TIME INTERACTION PLOT FOR RATS DATA	57

Chapter 1

Introduction and Motivation

1.1 Introduction

Repeated Measures Designs (RMD) are one of the most frequently studied and applied designs in a variety of applied fields. A design in which the same experimental unit is *repeatedly* observed under multiple treatments is called *repeated measures design*. This is but a broad concept and in practice a repeated measures design is laid out in a variety of ways, from a very simple set up of one-way repeated measures design to a very complex framework of longitudinal data or some other mixed model set up. There may be more than one treatments(factors) in the design and the repeated measurements may be observed on one or more of the factors. The treatments may be conditions, occasions, time points and the units may be individuals, animals, cell cultures etc.. Due to its wide application in the realm of applied research, a repeated measures design is conceived of, and planned in a number of ways. It is basically a set up that can be used to plan any standard experiment. Mostly, any design with single observation can also be planned with repeated observations (Crowder and Hand, 1990, Ch. 1). For example, a paired design is a simple formulation of a repeated measures design where same unit is observed twice. A cross-over design is another example of a repeated measures design (Hinkelmann and Kempthorne, 2005, Ch. 19). Similarly, a factorial experiment and a covariance experiment may be laid out in a repeated measures set up (Stuart et al, 1999, Ch. 31). A split-plot experiment is a typical example of a repeated measures design wherein the main-plot factor contains a set of repeated measurements taken on the sub-plot factor (Hinkelmann and

Kempthorne, 1994, Ch. 13). A comprehensive introduction to the analysis and different plans of RMD is Crowder and Hand (1990) and applications to several real-life, mostly medical, experiments is given by Hand and Taylor (1987).

The most attractive feature of a repeated measures design is the potential advantages it offers. For example,

- (1) Maximum error control (Winer et al, 1991, Ch. 4)
- (2) Economy of subjects
- (3) Study of patterned behavior of individuals over different treatments, conditions or time points, and
- (4) Data more reliable than in a cross sectional study.

For details, see Davis (2002, Ch. 1). Although the advantages of an RMD outweigh its disadvantages, there are some issues to be seriously taken care of before planning an RMD. Since the data constitute repeated observations under essentially the same conditions, hence correlated observations, the independence assumption is no longer viable. At times, the structure of correlation matrix can be extremely complicated and the issue needs a careful treatment and a good amount of knowledge. Second, the data may not be complete due to one reason or another. An incomplete or unbalanced repeated measures design is much more cumbersome to deal with than a design which does not have repeated observations.

1.2 Classical Approaches of Analysis

Suppose $\mathbf{X}_k = (X_{k1}, \dots, X_{kd})' \stackrel{i.i.d.}{\sim} \mathcal{N}_d(\boldsymbol{\mu}, \boldsymbol{\Sigma})$, $\boldsymbol{\Sigma} > 0$, denotes a vector of d repeated measurements on k th individual, where $k = 1, \dots, n$. A unique feature of a repeated measures design is that the data obtained from such a design can be analyzed in a variety of ways. Some comprehensive references for the classical analysis of repeated measures designs are Crowder and Hand (1990), Lindsey (1999), Stuart et al (1999, Ch. 31), Davis (2002) and Maxwell and Delaney (2004). A special, detailed account of longitudinal data analysis can be found in Diggle et al (2002) whereas more complicated repeated measures models are discussed in Vonesh and Chinchilli (1996) and Davidian and Giltinan (1995). Broadly, the analysis of a repeated measures

design can be carried out using both univariate and multivariate techniques. The two approaches basically differ on the assumption about the underlying covariance matrix, Σ : univariate approach requires Σ to be compound symmetric (Definition B.6) while multivariate approach only requires that $\Sigma > 0$. This project basically addresses the problem of analyzing repeated measures data when $d > n$, i.e., when the data are *high dimensional*. Both the univariate and multivariate approaches have limitations in dealing with the high dimensional data. Following is a brief review of these two approaches, augmented with their shortcomings when the data are high dimensional.

1.2.1 Univariate Analysis

Let $\mathbf{X}_k$ be as defined above. If Σ is compound symmetry (see Definition B.6), i.e. if all variances are equal and all covariances are equal, we call the model to be *univariate repeated measures model*. Interestingly, the advantage the assumption of compound symmetry offers and the problem it generates move quite confluent in the univariate analysis: if the assumption is tenable, then univariate analysis is more powerful than the multivariate approach (Stuart et al, 1999, Ch. 31), but in practice it is very rare that this assumption is met. Note that, in general, the univariate approach can be validly applied to analyze high dimensional data ($d > n$), but it is this restrictive assumption, put together with the sensitivity of univariate approach to the normality assumption (Stuart et al, 1999, p 719), that hinders a wide spectrum of application of this approach.

Huynh and Feldt (1970) showed that compound symmetry assumption is sufficient but not necessary for the valid univariate analysis. The necessary condition is that Σ is circular. A matrix $\Sigma = (\sigma_{ij})_{i,j=1}^d$ is *circular* if, $\forall i \neq j$, $\sigma_i^2 + \sigma_j^2 - 2\sigma_{ij} = 2\lambda$, where λ is any constant. The assumption of circularity is, in practice, more frequently studied in the form of *sphericity*. Any circular matrix can be transformed to a spherical matrix using the transformation $\mathbf{C}\Sigma\mathbf{C}' = \Sigma_y$ where $\mathbf{C}_{(d-1) \times d}$ is a matrix of constants with rank $d - 1$ and $\mathbf{y} = \mathbf{C}\mathbf{x}$ refers to the new transformed variables. Now, if Σ is circular, then $\Sigma_y = \lambda\mathbf{I}$, where any matrix of the form $\lambda\mathbf{I}$ is called spherical matrix (Winer et al, 1991; Kirk, 1982). Since Σ is mostly unknown, we need to know if the sphericity condition is satisfied. Mauchley (1940) derived a test of sphericity for the one sample case which was extended to the multi-sample case by Mendoza (1980). The inevitability of sphericity condition in the univariate analysis has been extensively studied. For some very useful references, see

Collier et al (1967), Mendoza et al (1976), Huynh and Mandeville (1979), Keselman et al (1980), Keselman (1998) and Keselman et al (2000). Several studies have shown that, by all practical means, the sphericity assumption is met in practice only in the form of compound symmetry assumption (Maxwell and Delaney, 2004, p 541).

In general, when the sphericity assumption is not satisfied, the probability of type I error associated with the F test seriously inflates. Box (1954a,b) introduced a correction for this bias, computed as

$$\epsilon = \frac{[\text{tr}(\mathbf{\Sigma}_y)]^2}{(d-1)\text{tr}(\mathbf{\Sigma}_y^2)} = \frac{(\sum_i \lambda_i)^2}{(d-1)(\sum_i \lambda_i^2)} \quad (1.1)$$

where $\mathbf{\Sigma}_y$ is as defined above and λ_i s are the eigenvalues of $\mathbf{\Sigma}_y$. Further, $\frac{1}{d-1} \leq \epsilon \leq 1$. Some estimators of ϵ are suggested by Geisser and Greenhouse (1958), Greenhouse and Geisser (1959), and Huynh and Feldt (1976). Muller and Barton (1989) study the power of the F test using these approximations for $n > d$. For more details and applications of these estimators, see Kirk (1982, p 259) and Winer et al (1991, pp 253-54).

The restrictive sphericity assumption of univariate approach is replaced with reasonable flexibility on $\mathbf{\Sigma}$ by the recently developed linear mixed model approach. Since the subjects in a repeated measures setting are usually random, the repeated measures model can also be considered as a mixed model with repeated measures factor as fixed. The mixed model approach accommodates several structures of $\mathbf{\Sigma}$, including compound symmetry. But when $d > n$, only a limited number of these structures can be entertained since, otherwise, $\mathbf{\Sigma}$ cannot be estimated (Davis, 2002, Ch. 6).

1.2.2 Multivariate Analysis

The main advantage of multivariate analysis is that we need not specify any structure of $\mathbf{\Sigma}$. The only assumption we need is that $\mathbf{\Sigma} > 0$. Therefore, for one sample analysis, we assume $\mathbf{X}_k \sim \mathcal{N}(\boldsymbol{\mu}, \mathbf{\Sigma})$, $\mathbf{\Sigma} > 0$, where $\mathbf{X}_k = (X_{k1}, \dots, X_{kd})'$, $k = 1, \dots, n$, are independent, identically distributed random vectors, each of d repeated measures on the same individual. Usually, we are interested to test the hypothesis $H_0 : \mathbf{T}\boldsymbol{\mu} = \mathbf{0}$ where $\mathbf{T}_{r \times d}$ is a matrix of contrasts with rank $r < d$. The maximum likelihood estimators of $\boldsymbol{\mu}$ and $\mathbf{\Sigma}$ are $\bar{\mathbf{X}}$ and $\frac{n-1}{n}\hat{\mathbf{\Sigma}}$, respectively, where, $\bar{\mathbf{X}} = \frac{1}{n} \sum_{k=1}^n \mathbf{X}_k$ and $\hat{\mathbf{\Sigma}} = \frac{1}{n-1} \sum_{k=1}^n (\mathbf{X}_k - \bar{\mathbf{X}})(\mathbf{X}_k - \bar{\mathbf{X}})'$. The hypothesis, H_0 , can be tested using

Hotelling's T^2 statistic, usually expressed as F statistic as

$$F = \frac{n-d+1}{(n-1)(d-1)} T^2 \quad (1.2)$$

where $T^2 = n(\mathbf{T}\bar{\mathbf{X}})'(\mathbf{T}\hat{\Sigma}\mathbf{T}')^{-1}(\mathbf{T}\bar{\mathbf{X}})$. Further, $(n-1)\mathbf{C}\hat{\Sigma}\mathbf{C}' \sim W_{d-1}(n-1, \mathbf{C}\Sigma\mathbf{C}')$, where W denotes the Wishart distribution (Davis, 2002, Ch. 3). This one sample problem can be extended to $a \geq 2$ populations. For details, see Stuart et al (1999, Ch. 31). Note that, the formulation of the hypothesis matrix $\mathbf{T}$ covers a wide variety of general linear hypotheses. One such formulation is to test profile hypothesis. There is a vast literature on the multivariate analysis, including profile analysis. For more details, see Johnson and Wichern (2002), Rencher (1995) and Davis (2002). Timm (2002) has discussed the profile analysis in detail under the set up of general linear model based on Gauss-Markov theorem whereas Srivastava (2002, Chs. 6, 7) gives extensions of simple MANOVA to other designs including factorial experiments. Davis (2002, Ch. 4) compares different multivariate techniques regarding level control and power.

Clearly, for the one sample case above, $\hat{\Sigma} > 0 \Rightarrow n-1 > d$. This indicates that the T^2 approach totally collapses when the data are high dimensional. There are other multivariate approaches, for example Wald-type statistic, which do not totally break down in this case but they suffer from other shortcomings. A comparison, based on simulation results, is given in Chapter 2 to substantiate the idea that it is only the ANOVA-type statistic (see Section 2.1) that can be repaired to be used for the analysis of high dimensional data.

1.3 The Case of High Dimensional Data

Both the univariate and multivariate approaches have problems in dealing with the data from a repeated measures design when $d > n$. The univariate approach, as mentioned above, strongly depends on the assumption of sphericity which is rarely satisfied in practice (Stuart et al, 1999, p 719-720). The multivariate approach, although does not need specification of the covariance structure, is totally unworkable when $d > n$, since the estimated covariance estimator is no longer non-singular (Stuart et al, 1999, p 716 and passim).

The present project addresses the problem of high dimensional data for

the analysis of repeated measures designs. A modification to the ANOVA-type statistic (Brunner, 2001), based on Box's approximation (Box, 1954a), is proposed. The statistics for one- and two-sample designs are derived. The one sample statistic is derived in Chapter 2 and is extended for the two sample case in Chapter 3. The statistics, in both cases, are evaluated using simulation studies, supplemented with numerical illustrations. The data used for these illustrations are provided in Appendix C. Appendix A gives the general results, mainly regarding quadratic and bilinear forms, which are frequently used in the derivations of the statistics, whereas some other useful results are collected in Appendix B.

1.4 Motivating Examples

Following are a few motivating examples depicting the need for some durable and practical solution to deal with the high dimensional data. The data and the experimental description for the first two examples are taken from Brunner, Domhof and Langer (2002). The first example is a typical one where the data are not high dimensional but is included to show that the test statistics derived in this project are also applicable when the data are not high dimensional. The experiment for the third example is reported in Jordan et al (2004). The data sets are provided in Appendix C.

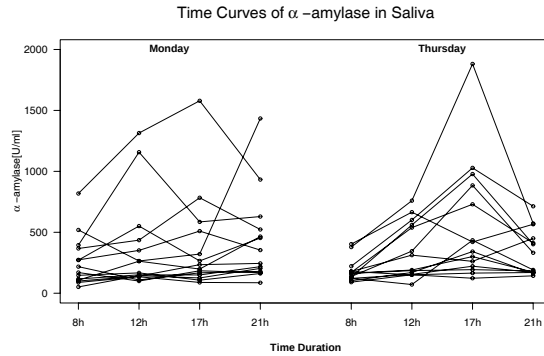


Figure 1.1: TIME PROFILES FOR α -AMYLASE STUDY

Example 1.1 (α -Amylase in Saliva) The objective of this experiment is to study the effect of time on the α -amylase activity in saliva. The α -

amylase level in saliva is repeatedly measured for each of 14 individuals in the experiment. The repeated measurements on each subject are taken at four time points per day, 8 a.m., 12 a.m., 5 p.m. and 9 p.m., and the experiment is repeated on two days, Monday and Thursday. Fig. 1.1 shows the profiles of the subjects on all four time points for both days. The data for this experiment is given in Section C.1 and the analysis is carried out in Section 2.5.2. $\triangle$

Example 1.2 (Body Weight of Wistar Rats) The body weight of male

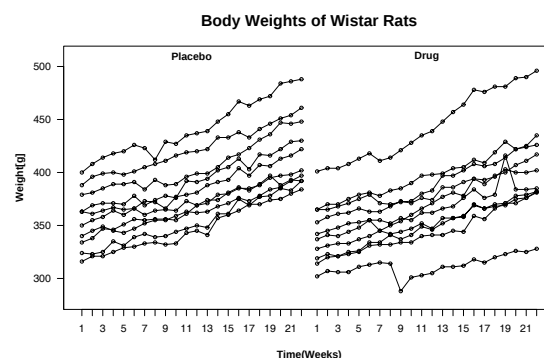


Figure 1.2: TIME PROFILES FOR BODY WEIGHT STUDY

Wistar rats was observed over a period of 22 weeks to assess the toxicity of a drug. A group of ten animals was given a placebo, while a second group of ten was given a high dose of the drug. The time curves of the body weights for both groups are shown in Fig. 1.3. The main question to be addressed is whether the body-weights of the two test groups differ in their evolution over time. The data for this experiment is given in Section C.2 and the analysis is carried out in Section 3.5.2. $\triangle$

Example 1.3 (Sleep Disorder) The aim of this experiment was to investigate the activity of Prostaglandin-D-synthase (β -trace) in relation to human sleep. The variable of interest, serum concentration, was measured on each of 10 men and 10 women at 18 time points. Measurements were taken every 4 hours for three consecutive nights classified as Normal Sleep(NS), Total Sleep Deprivation(TSD) and Recovery Night(RN).

The data are given in Section C.3 and is shown in Fig. 1.4. Considering

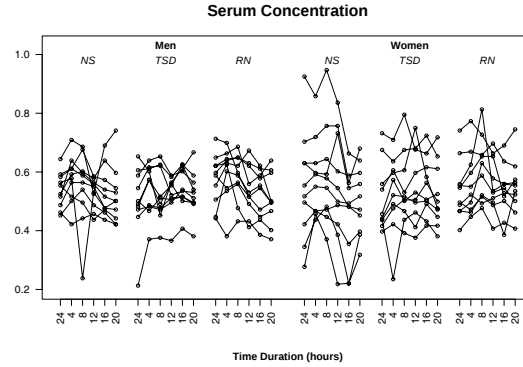


Figure 1.3: TIME PROFILES FOR SLEEP DISORDER STUDY

only women group (one sample), the data are first analyzed in Section 2.5.1. Then the complete data are analyzed in Section 3.5.1. $\triangle$

1.5 Review of Literature

Analysis of high dimensional data, be it for repeated measures designs or otherwise, has long been an uncharted territory of research. It is relatively recently that a systematic, target driven research has emerged significantly. Recent advances in the analysis of proteomics and microarray data have galvanized the need for techniques to analyze such data. Apart from a very few examples, most of the research work has surfaced in the scientific body of literature in the past decade or so. Numerous researchers have addressed the problem and have meaningfully contributed to the development of the theory. But this recent development typically addresses the problem of multivariate set up of the high dimensional data. Since, the mainstream effort in this dissertation concerns the analysis of high dimensional repeated measures data, therefore, these developments are not of direct interest for us. But, is briefly discussed in Chapter 4, the statistics, developed in this project, can be appropriately modified to test the multivariate hypothesis. It can also be verified that our statistic, in this particular case, outperforms the available competitors, in several aspects. Since we shall not deal with this multivariate case in detail, therefore, the relevant references are only listed in the following without giving any further details.

A one sample statistic for high dimensional multivariate data is discussed in Läuter et al (1998) and Läuter (2004), based on the idea of spherical distributions (Tai-Kai and Yao-Ting, 1990). Both one and two sample statistics are discussed in Srivastava (2007) and Srivastava and Du (2008) whereas only two sample statistics are considered by Depmster (1958) and Bai and Saranadasa (1996). The multivariate general linear model for high dimensional data is considered in Srivastava and Fujikoshi (2006) and a multivariate extension for one-way MANOVA is given by Schott (2007).

It is stated above that the classical multivariate approach is unworkable for high dimensional data since the estimated covariance matrix is singular. An alternative approach to deal with high dimensional data is, therefore, to find a reasonable, well-conditioned estimator of the large dimensional covariance matrix. Recently, there have been some attempts to address this question. For example, see Ledoit and Wolf (2002), Ledoit and Wolf (2004), Srivastava (2005), Birke and Dette (2005), and Schäfer and Strimmer (2005).

Attempts have also been made to study the classical ANOVA F test when the number of treatments is large; two recent references include Bathke (2002) and Akritas and Arnold (2000). Similarly, non-parametric analysis of high dimensional data are discussed by Bathke and Harrar (2006a,b) and Harrar and Bathke (2006). A comprehensive review of a variety of methods for the analysis of multivariate data, high dimensional or otherwise, is given in Bathke et al (2008).

Chapter 2

The One-sample Statistic

In this chapter, the modified ANOVA-type statistic for the one sample problem will be derived.

2.1 The ANOVA-type Statistic

Let $\mathbf{X}_k = (X_{k1}, \dots, X_{kd})'$ be a vector of d repeated measurements taken on k th subject/unit where $k = 1, \dots, n$. The data from such an experiment may be arranged as shown in Table 2.1. We assume that

$$\mathbf{X}_k \stackrel{i.i.d.}{\sim} \mathcal{N}(\boldsymbol{\mu}, \boldsymbol{\Sigma}) \quad (2.1)$$

where $\boldsymbol{\mu} = (\mu_1, \dots, \mu_d)'$ is the population mean vector and

$$\boldsymbol{\Sigma}_d = \begin{pmatrix} \sigma_1^2 & \sigma_{12} & \dots & \sigma_{1d} \\ \sigma_{21} & \sigma_2^2 & \dots & \sigma_{2d} \\ \vdots & \vdots & \ddots & \vdots \\ \sigma_{d1} & \sigma_{d2} & \dots & \sigma_d^2 \end{pmatrix}$$

is the covariance matrix. For the derivation of the proposed test statistic, no specific form of $\boldsymbol{\Sigma}$ is assumed. However, the results simplify immediately in case some structure is imposed on $\boldsymbol{\Sigma}$, e.g. *compound symmetric* structure (Definition B.6).

For model $\mathbf{X}_k \sim \mathcal{N}(\boldsymbol{\mu}, \boldsymbol{\Sigma})$, $\boldsymbol{\Sigma} > 0$, the null hypothesis to be tested is $H_0 : \mathbf{H}\boldsymbol{\mu} = \mathbf{0}$ where $\mathbf{H}$ is the hypothesis matrix. The matrix $\mathbf{H}$ can be formulated in different settings depending on the objectives of the experiment.

Table 2.1: THE GENERAL DATA TABLE: ONE-SAMPLE CASE

	1	...	s	...	d
1	X_{11}	...	X_{1s}	...	X_{1d}
$\vdots$	$\vdots$	$\ddots$		$\ddots$	$\vdots$
k	X_{k1}	...	X_{ks}	...	X_{kd}
$\vdots$	$\vdots$	$\ddots$		$\ddots$	$\vdots$
n	X_{n1}	...	X_{ns}	...	X_{nd}
Means	$\bar{X}_{.1}$	...	$\bar{X}_{.s}$	...	$\bar{X}_{.d}$

Specifically, we may set $\mathbf{H} = \mathbf{I}$ to test the usual multivariate hypothesis $H_0 : \boldsymbol{\mu} = \mathbf{0}$. For a simple, unstructured repeated measures design, we have $\mathbf{H} = \mathbf{P}_d$, where $\mathbf{P}_d = \mathbf{I}_d - \frac{1}{d}\mathbf{J}_d$, being the centering matrix, is symmetric and idempotent, hence projection matrix, $\mathbf{J}$ is the matrix of 1s and $\mathbf{I}$ is the identity matrix.

To give a unique representation to the null hypothesis, we can write $H_0 : \mathbf{T}\boldsymbol{\mu} = \mathbf{0}$ where $\mathbf{T} = \mathbf{H}'(\mathbf{H}\mathbf{H}')^{-}\mathbf{H}$ is the general hypothesis matrix with $(\mathbf{H}\mathbf{H}')^{-}$ denoting a g -inverse of $\mathbf{H}\mathbf{H}'$. We note that $\mathbf{H}\boldsymbol{\mu} = \mathbf{0} \Leftrightarrow \mathbf{T}\boldsymbol{\mu} = \mathbf{0}$. The matrix $\mathbf{T}$ can be formulated so as to represent any general linear hypothesis, including, for example, any factorial structure of repeated measures. For example, for the sleep lab experiment, discussed in Section 2.5.1, we have $\mathbf{T} = \mathbf{P}_3 \otimes \frac{1}{6}\mathbf{J}_6$ and $\mathbf{T} = \frac{1}{3}\mathbf{J}_3 \otimes \mathbf{P}_6$, for the main effects and $\mathbf{P}_3 \otimes \mathbf{P}_6$ for the interaction effect.

As discussed in Section 1.2, when $d < n$, such data can be analyzed using multivariate approach without assuming any particular structure of $\boldsymbol{\Sigma}$ (Stuart et al, 1999, Ch. 31). But when $d > n$, the case of high dimensional data, the multivariate approach, for example Hotelling's T^2 , totally collapses since the estimated covariance matrix is singular (Rencher, 1995, p 135). For $n > d$, there are several competing statistics to be used to test H_0 , including Wald-type statistic, Hotelling's T^2 statistic and ANOVA-type statistic which are discussed in the following.

The Wald-type statistic (WTS) is defined as (Rao, 1973; Timm, 2002)

$$W_n = n\bar{\mathbf{X}}'\mathbf{H}'\left(\mathbf{H}\hat{\boldsymbol{\Sigma}}\mathbf{H}'\right)^{-}\mathbf{H}\bar{\mathbf{X}} \quad (2.2)$$

where $\bar{\mathbf{X}}$ and $\hat{\Sigma}$ are the sample estimators of $\boldsymbol{\mu}$ and $\Sigma > 0$, respectively. Clearly, we can replace the g -inverse, $(\cdot)^-$, with the unique Moore-Penrose inverse, $(\cdot)^+$. Then, $W_n \sim \chi_r^2$, as $n \rightarrow \infty$, where $r = \rho(\mathbf{H})$ and $\rho(\cdot)$ denotes the rank of a matrix (see Campbell and Meyer, 1979, Ch. 10).

W_n can be used to define the Hotelling's T^2 statistic, usually transformed to an F statistic, as (Davis, 2002)

$$H_n = \frac{n-r}{(n-1)r} W_n \quad (2.3)$$

where r and W_n are as defined above. Under H_0 , $H_n \sim F_{(r, n-r)}$.

The ANOVA-type statistic (ATS), as considered by Box (1954a), Brunner, Dette and Munk (1997), Brunner (2001), Brunner, Domhof and Langer (2002), and Tian and Wilcox (2007), is based on the traces of products of matrices involving the sample covariance matrix and is defined as

$$A_n = \frac{n\bar{\mathbf{X}}'\mathbf{T}\bar{\mathbf{X}}}{\text{tr}(\mathbf{T}\hat{\Sigma})} \quad (2.4)$$

where $\text{tr}(\cdot)$ denotes the trace. Further, $A_n \sim \chi_{\hat{f}}^2/\hat{f}$ where

$$\hat{f} = \frac{[\text{tr}(\mathbf{T}\hat{\Sigma})]^2}{\text{tr}(\mathbf{T}\hat{\Sigma})^2}$$

is the degrees of freedom estimated from the sample.

The rationale behind the ANOVA-type statistic stems from two main shortcomings of the Wald-type statistic. W_n needs very large sample size to yield a good approximation to the chi-square distribution and the approximation is still very bad if at least some of the diagonal entries in the covariance matrix, $\hat{\Sigma}$, are zero or near zero. The solution approached by A_n , in this situation, is to take $\hat{\Sigma}$ out of W_n and write $\mathbf{H}'(\mathbf{H}\mathbf{H}')^{-1}\mathbf{H} = \mathbf{T}$ where $\mathbf{T}$ is called the general hypothesis matrix. We know that $\mathbf{H}\boldsymbol{\mu} = \mathbf{0} \Leftrightarrow \mathbf{T}\boldsymbol{\mu} = \mathbf{0}$. Then, W_n reduces to

$$Q_n = n\bar{\mathbf{X}}'\mathbf{T}\bar{\mathbf{X}}. \quad (2.5)$$

But it can be easily demonstrated that Q_n is not asymptotically distributed as a χ^2 random variable. This problem, however, can be approached using Box's approximation (Box, 1954a). First, we need the following well-known representation theorem for the quadratic form from the theory of linear models (Box, 1954a; Mathai and Provost, 1992, Ch. 3).

Theorem 2.1 *Let $\mathbf{X} \sim \mathcal{N}_d(\mathbf{0}, \Sigma)$ and let $\mathbf{T}$ be any symmetric matrix. Then*

$$Q = \mathbf{X}'\mathbf{T}\mathbf{X} \sim \sum_{i=1}^d \lambda_i C_i$$

where λ_i are the eigenvalues of $\mathbf{T}\Sigma$, $C_i \sim \chi_1^2$ and C_i are independent. $\square$

The identity Q in Theorem 2.1 represents a quadratic form as a weighted sum of independent single-degree-of-freedom chi-square random variables. The idea is to replace the distribution of Q in Theorem 2.1 with that of $g\chi_f^2$ distribution, where g and f are chosen such that the first two moments of Q and $g\chi_f^2$ are same. Such an approximation is referred to as Box's approximation although the idea was first proposed by Patnaik (1949); see Box (1954a) and Mathai and Provost (1992, Ch. 4). Based on this approximation, we get (Box, 1954a; Brunner, 2001; Mathai and Provost, 1992)

$$gf = E(g\chi_f^2) = E\left(\sum_i \lambda_i C_i\right) = \sum_i \lambda_i = \text{tr}(\mathbf{T}\Sigma) \quad (2.6)$$

$$2g^2f = \text{Var}(g\chi_f^2) = \text{Var}\left(\sum_i \lambda_i C_i\right) = 2 \sum_i \lambda_i^2 = 2\text{tr}(\mathbf{T}\Sigma)^2 \quad (2.7)$$

and solving these equations, we have

$$f = \frac{[\text{tr}(\mathbf{T}\Sigma)]^2}{\text{tr}(\mathbf{T}\Sigma)^2} \quad \text{and} \quad g = \frac{\text{tr}(\mathbf{T}\Sigma)^2}{\text{tr}(\mathbf{T}\Sigma)}.$$

The distribution of the quadratic form $\mathbf{X}'\mathbf{T}\mathbf{X}$ is then approximated as

$$\frac{\mathbf{X}'\mathbf{T}\mathbf{X}}{\text{tr}(\mathbf{T}\Sigma)} \sim \frac{\chi_f^2}{f}. \quad (2.8)$$

First let us evaluate the performance of A_n when Σ is known. Table 2.2 reports the simulation results for the test sizes for A_n for compound symmetric(CS), autoregressive(AR(1)) and unstructured(UN) covariance patterns (see Definitions B.6, B.7 and B.8). These are the results of 10,000 simulation runs for $n = 10$, $d \in \{5, 10, 50, 100, 200\}$ while assuming $\sigma^2 = 2, \kappa = 1$ for CS, $\sigma^2 = 1, \rho = 0.6$ for AR(1) and $\sigma_{ij} = 1(1)d$ ($i = j$), $\rho_{ij} = (i-1)/d$ ($i > j$), for UN.

Table 2.2: A_n WITH TRUE COVARIANCE STRUCTURES

d	Covariance Structure/ $1 - \alpha$								
	CS			AR(0.6)			UN		
	0.90	0.95	0.99	0.90	0.95	0.99	0.90	0.95	0.99
5	0.8952	0.9498	0.9888	0.8852	0.9404	0.9848	0.9215	0.9596	0.9907
10	0.9006	0.9480	0.9890	0.8966	0.9432	0.9856	0.9091	0.9536	0.9910
50	0.9016	0.9526	0.9918	0.9016	0.9448	0.9848	0.9003	0.9499	0.9895
100	0.8994	0.9454	0.9880	0.8978	0.9458	0.9874	0.8993	0.9486	0.9897
200	0.8918	0.9432	0.9884	0.8968	0.9428	0.9860	0.8990	0.9500	0.9897

From Table 2.2 we observe that the ANOVA-type statistic maintains the pre-assigned level quite accurately when the population covariance is known, even for n as small as 10. This performance is the same for all three types of covariances and does not seem to depend on the dimension. The evidence we get from these results is that, given the covariance matrix, the Box approximation is accurate whatever be the dimension, d .

But, in practice, of course, we need to estimate f and g ; i.e. we need to estimate Σ . Using the classical unbiased estimator,

$$\hat{\Sigma}_n = \frac{1}{n-1} \sum_{k=1}^n (\mathbf{X}_k - \bar{\mathbf{X}})(\mathbf{X}_k - \bar{\mathbf{X}})',$$

the estimator of the functional $\text{tr}(\mathbf{T}\Sigma)$ is unbiased but the estimators of the functionals $[\text{tr}(\mathbf{T}\Sigma)]^2$ and $\text{tr}(\mathbf{T}\Sigma)^2$ are biased. Table 2.3 reports some results to show this behavior. From this table, we also observe that the bias of

Table 2.3: PROPERTIES OF TRACES WITH COVARIANCE ESTIMATOR, $\hat{\Sigma}_n$

n	d	$\frac{E[\text{tr}(\mathbf{T}\hat{\Sigma}_n)]}{\text{tr}(\mathbf{T}\Sigma)}$	$\frac{\text{Var}[\text{tr}(\mathbf{T}\hat{\Sigma}_n)]}{[\text{tr}(\mathbf{T}\Sigma)]^2}$	$\frac{E[\text{tr}(\mathbf{T}\hat{\Sigma})]^2}{[\text{tr}(\mathbf{T}\Sigma)]^2}$	$\frac{\text{Var}[\text{tr}(\mathbf{T}\hat{\Sigma})]^2}{[\text{tr}(\mathbf{T}\Sigma)]^4}$	$\frac{E[\text{tr}(\mathbf{T}\hat{\Sigma})^2]}{\text{tr}(\mathbf{T}\Sigma)^2}$	$\frac{\text{Var}[\text{tr}(\mathbf{T}\hat{\Sigma})^2]}{[\text{tr}(\mathbf{T}\Sigma)^2]^2}$	$\frac{E(\hat{f})}{f}$	$\frac{\text{Var}(\hat{f})}{f^2}$
10	20	1.0008	0.0117	1.0134	0.0480	3.2320	0.5483	0.3151	0.0005
	50	1.0002	0.0046	1.0050	0.0184	6.5574	0.8308	0.1534	0.0000
	100	1.0002	0.0023	1.0026	0.0091	12.1155	1.3866	0.0828	0.0000
	200	0.9999	0.0011	1.0010	0.0044	23.2152	2.4519	0.0431	0.0000
	300	0.9997	0.0007	1.0002	0.0029	34.3139	3.5270	0.0291	0.0000
50	100	1.0003	0.0004	1.0009	0.0017	3.0428	0.0172	0.3290	0.0000
	200	1.0002	0.0002	1.0006	0.0008	5.0840	0.0229	0.1968	0.0000
	300	1.0001	0.0001	1.0003	0.0005	7.1234	0.0290	0.1404	0.0000

$\text{tr}(\mathbf{T}\Sigma)^2$ increases with increasing d , and this bad performance remains so even if n is increased five times. Note that, since $\text{tr}(\mathbf{T}\Sigma)^2$ is the denominator of f , therefore, f is also badly affected due to the biased estimation of trace. Although the variance of $\hat{f}$ reduces with increasing d , but this is an illusive

Table 2.4: COMPARISON OF A_n , W_n AND H_n

$(d = 5)$		Simulated Quantiles				
Statistic	Quantile	$n = 10$	$n = 15$	$n = 20$	$n = 30$	$n = 50$
A_n	0.90	0.9042	0.9043	0.9035	0.9031	0.9004
	0.95	0.9536	0.9519	0.9515	0.9526	0.9534
	0.99	0.9913	0.9904	0.9889	0.9928	0.9921
W_n	0.90	0.6315	0.7377	0.7898	0.8302	0.8592
	0.95	0.7062	0.8161	0.8573	0.8947	0.9179
	0.99	0.8189	0.9074	0.9390	0.9618	0.9761
H_n	0.90	0.8963	0.9008	0.9020	0.9008	0.8995
	0.95	0.9491	0.9506	0.9513	0.9503	0.9847
	0.99	0.9906	0.9906	0.9905	0.9895	0.9911

$(n = 10)$		Simulated Quantiles				
Statistic	Quantile	$d = 5$	$d = 10$	$d = 20$	$d = 30$	$d = 50$
A_n	0.90	0.9014	0.9368	0.9719	0.9849	0.9966
	0.95	0.9532	0.9774	0.9938	0.9979	0.9997
	0.99	0.9902	0.9980	1.0000	1.0000	1.0000
W_n	0.90	0.6234	0.0435	0.9780	1.0000	1.0000
	0.95	0.7005	0.0546	0.9836	1.0000	1.0000
	0.99	0.8130	0.0840	0.9918	1.0000	1.0000
H_n	0.90	0.8949	0.8982	-	-	-
	0.95	0.9440	0.9484	-	-	-
	0.99	0.9881	0.9899	-	-	-

behavior since the ratio $\frac{E(\hat{f})}{f}$ also vanishes with increasing d while we expect $\frac{E(\hat{f})}{f} = 1$ whatever be d .

Now, Table 2.4 gives the simulation results for the type I error rates of three test statistics, W_n , H_n and A_n , as discussed above. For the upper part of the table, $d = 5$ is fixed and $n \in \{10, 15, 20, 30, 50\}$ while for the lower part, $n = 10$ is fixed and $d \in \{5, 10, 20, 30, 50\}$.

Note that both A_n and H_n keep the nominal level when d is fixed and n increases, i.e. when $n > d$, but the Wald-type statistic is liberal. It tends to be less and less liberal as n increases but still at $n = 50$ the gap is large especially when compared to the other two statistics. On the other hand, when n is fixed and d increases, Hotelling's T^2 totally collapses for $d > n$ and W_n ranges from being very liberal to very conservative, while A_n is also conservative when d increases. Note once again that, both A_n and H_n keep the nominal level when $n \geq d$ but the W_n fails even in this range. One can use similar results to show that f , which is proportional to the Box's ϵ (Equation

1.1), is biased and the bias increases with increasing d . The results of Tables 2.3 and 2.4 hint at a possibility of improving A_n while Hotelling's T^2 and W_n are totally unworkable for the case of high dimensionality.

To improve the performance of A_n , we need to define the component estimators of the statistic such that they are consistent and the variances of estimators are uniformly bounded with respect to the dimension, d . In other words, the estimators must be consistent and dimensionally stable. To this end, we need Definition 2.2 which extends the concept of consistency of a sequence of estimators to an array of estimators which depend, simultaneously, on n and d . For the general theory of consistency of an array of estimators, see Lehmann (1999, Ch. 2) and Serfling (1980, Ch. 1).

Definition 2.2 An array of estimators $\hat{\theta}_{n,d}$ of functional θ_d is consistent if $\hat{\theta}_{n,d} - \theta_d \xrightarrow{P} 0$ for fixed d . „

A straightforward calculation, based on Chebychev's inequality (Theorem B.4), shows that consistency follows if $\forall 1 \leq d < \infty$,

$\lim_{n \rightarrow \infty} \{E(\hat{\theta}_{n,d})\} = \theta_d$ and $\lim_{n \rightarrow \infty} \{\text{Var}(\hat{\theta}_{n,d})\} = 0$. In the sequel, we shall show that for the estimators we define for the modified A_n , $E(\hat{\theta}_{n,d}) = \theta_d$, for all n , and $\lim_{n \rightarrow \infty} \left\{ \frac{\text{Var}(\hat{\theta}_{n,d})}{\theta_d^2} \right\} = 0$, where these variances are uniformly bounded with respect to d , so that the quality of the approximation depends only on n . Finally, to settle the issue of dimensional stability, we shall ensure that the estimators defined for the modified ANOVA-type statistic follow the criteria given below.

Definition 2.3 An array of estimators $\hat{\theta}_{n,d}$ of functional θ_d is dimensionally stable if, $\forall d \geq 1, n \geq 1$,

1. $E(\hat{\theta}_{n,d}) = \theta_d$,
2. $\text{Var}\left(\frac{\hat{\theta}_{n,d}}{\theta_d}\right) \leq B(n) < \infty$,

where $B(n) \rightarrow 0$, as $n \rightarrow \infty$ and is uniformly bounded with respect to d . „

Based on these results, we shall, in the next section, define the new estimators such that they are unbiased, consistent and are uniformly bounded with respect to d .

2.2 The Estimators

We have $\mathbf{X}_k \sim \mathcal{N}(\boldsymbol{\mu}, \boldsymbol{\Sigma})$, $\boldsymbol{\Sigma} > 0$, $k = 1, \dots, n$. For $H_0 : \mathbf{T}\boldsymbol{\mu} = \mathbf{0}$, let $\mathbf{T}\mathbf{X}_k = \mathbf{Y}_k$. Under H_0 , $E(\mathbf{Y}_k) = \mathbf{0}$, $\text{Var}(\mathbf{Y}_k) = \text{Var}(\mathbf{T}\mathbf{X}_k) = \mathbf{T}\boldsymbol{\Sigma}\mathbf{T} = \mathbf{S}$. Then, we define the covariance estimator as

$$\widehat{\mathbf{S}}_n = \frac{1}{n} \sum_{k=1}^n \mathbf{S}_k = \frac{1}{n} \sum_{k=1}^n \mathbf{Y}_k \mathbf{Y}_k'. \quad (2.9)$$

It is straightforward to verify that $\widehat{\mathbf{S}}_n$, under H_0 , is an unbiased estimator of $\mathbf{S}$. Let $A_k = \mathbf{Y}_k' \mathbf{Y}_k$ and $A_{kl} = \mathbf{Y}_k' \mathbf{Y}_l$, $k \neq l$, be, respectively, a quadratic and a symmetric bilinear form. Let B_0 , B_1 and B_2 be the following estimators of $\text{tr}(\mathbf{T}\boldsymbol{\Sigma})$, $[\text{tr}(\mathbf{T}\boldsymbol{\Sigma})]^2$ and $\text{tr}(\mathbf{T}\boldsymbol{\Sigma})^2$, respectively, where,

$$\left. \begin{aligned} B_0 &= \frac{1}{n} \sum_{k=1}^n A_k, \\ B_1 &= \frac{1}{n(n-1)} \sum_{\substack{k=1 \\ k \neq l}}^n \sum_{l=1}^n A_k A_l \\ B_2 &= \frac{1}{n(n-1)} \sum_{\substack{k=1 \\ k \neq l}}^n \sum_{l=1}^n A_{kl}^2. \end{aligned} \right\} \quad (2.10)$$

We prove the following theorem on the properties of these estimators.

Theorem 2.4 *The estimators B_0 , B_1 and B_2 , as defined in Equations (2.10), are unbiased and consistent estimators of $\text{tr}(\mathbf{T}\boldsymbol{\Sigma})$, $[\text{tr}(\mathbf{T}\boldsymbol{\Sigma})]^2$ and $\text{tr}(\mathbf{T}\boldsymbol{\Sigma})^2$, respectively. Further, the variances of the ratios of these estimators to the traces that they estimate are of order $O(1/n)$ and are uniformly bounded in d . $\square$*

Proof The unbiasedness is quite straightforward using Lemmas A.11 and A.19 and using the independence since $k \neq l$. Then we begin with $\text{Var}(B_0)$.

$$\text{Var}(B_0) = \frac{1}{n^2} \sum_{k=1}^n \text{Var}(A_k) = \frac{2}{n} \text{tr}(\mathbf{T}\boldsymbol{\Sigma})^2,$$

from Equations (A.5). Now, with $k \neq l$, $r \neq s$, we have

$$\begin{aligned}
 \text{Var}(B_1) &= \frac{1}{n^2(n-1)^2} \left(\underbrace{\sum_{k=1}^n \sum_{l=1}^n}_{k \neq l} \text{Var}(A_k A_l) + \underbrace{\sum_{k=1}^n \sum_{l=1}^n \sum_{r=1}^n \sum_{s=1}^n}_{\substack{k \neq l \quad r \neq s \\ (k, r) \neq (l, s)}} \text{Cov}(A_k A_l, A_r A_s) \right) \\
 &= \frac{1}{n^2(n-1)^2} \left[2n(n-1) \left(4[\text{tr}(\mathbf{T}\Sigma)^2]^2 + 4\text{tr}(\mathbf{T}\Sigma)^2[\text{tr}(\mathbf{T}\Sigma)]^2 \right) \right. \\
 &\quad \left. + 4n(n-1)(n-2) \cdot 2\text{tr}(\mathbf{T}\Sigma)^2[\text{tr}(\mathbf{T}\Sigma)]^2 \right] \\
 &= \frac{8}{n(n-1)} \left[[\text{tr}(\mathbf{T}\Sigma)^2]^2 + (n-1)\text{tr}(\mathbf{T}\Sigma)^2[\text{tr}(\mathbf{T}\Sigma)]^2 \right]
 \end{aligned}$$

using again Equations (A.5). Similarly, from Equations (A.19), we get

$$\begin{aligned}
 \text{Var}(B_2) &= \frac{1}{n^2(n-1)^2} \left(\underbrace{\sum_{k=1}^n \sum_{l=1}^n}_{k \neq l} \text{Var}(A_{kl}^2) + \underbrace{\sum_{k=1}^n \sum_{l=1}^n \sum_{r=1}^n \sum_{s=1}^n}_{\substack{k \neq l \quad r \neq s \\ (k, r) \neq (l, s)}} \text{Cov}(A_{kl}^2, A_{rs}^2) \right) \\
 &= \frac{1}{n^2(n-1)^2} \left[2n(n-1) \left(6\text{tr}(\mathbf{T}\Sigma)^4 + 2[\text{tr}(\mathbf{T}\Sigma)^2]^2 \right) \right. \\
 &\quad \left. + 4n(n-1)(n-2) (2\text{tr}(\mathbf{T}\Sigma)^4) \right] \\
 &= \frac{4}{n(n-1)} \left[(2n-1)\text{tr}(\mathbf{T}\Sigma)^4 + [\text{tr}(\mathbf{T}\Sigma)^2]^2 \right].
 \end{aligned}$$

Clearly, all three variances vanish when $n \rightarrow \infty$, proving the estimators to be consistent.

Further, we observe that

$$\begin{aligned} \text{Var} \left(\frac{B_0}{\text{tr}(\mathbf{T}\Sigma)} \right) &= \frac{2}{n} \frac{\text{tr}(\mathbf{T}\Sigma)^2}{[\text{tr}(\mathbf{T}\Sigma)]^2} \leq \frac{2}{n} \\ \text{Var} \left(\frac{B_1}{[\text{tr}(\mathbf{T}\Sigma)]^2} \right) &= \frac{8}{n(n-1)} \left(\frac{[\text{tr}(\mathbf{T}\Sigma)^2]^2}{[\text{tr}(\mathbf{T}\Sigma)]^4} + (n-1) \frac{\text{tr}(\mathbf{T}\Sigma)^2}{[\text{tr}(\mathbf{T}\Sigma)]^2} \right) \leq \frac{8}{n-1} \\ \text{Var} \left(\frac{B_2}{\text{tr}(\mathbf{T}\Sigma)^2} \right) &= \frac{4}{n(n-1)} \left(\frac{(2n-1)\text{tr}(\mathbf{T}\Sigma)^4}{[\text{tr}(\mathbf{T}\Sigma)^2]^2} + 1 \right) \leq \frac{8}{n-1}. \end{aligned}$$

where the inequalities come immediately from Lemma B.11. ■

2.3 The Approximating Distribution

Now, we re-write Q_n in Equation (2.5) as

$$Q_n = n\bar{\mathbf{X}}' \mathbf{T} \bar{\mathbf{X}} = \frac{1}{n} \sum_{k=1}^n \sum_{l=1}^n A_{kl} \quad (2.11)$$

where A_{kl} is the bilinear form as defined right after Equations (2.10). Then the new test statistic is defined as

$$\tilde{A}_n = \frac{Q_n}{B_0} \quad (2.12)$$

with degrees of freedom $\tilde{f} = \frac{B_1}{B_2}$. Using the moments of quadratic and bilinear forms (Sections A.1 and A.2), we get the following results.

Theorem 2.5 *For Q_n and B_0 , as defined above, we have*

$$E(Q_n) = \text{tr}(\mathbf{T}\Sigma) \quad (2.13)$$

$$\text{Var}(Q_n) = 2\text{tr}(\mathbf{T}\Sigma)^2 \quad (2.14)$$

$$\text{Cov}(Q_n, B_0) = \frac{2}{n} \text{tr}(\mathbf{T}\Sigma)^2. \quad (2.15)$$

□

Proof Here, Q_n is the quadratic form of the test statistic and from Equation (2.11) we have $Q_n = \frac{1}{n} \sum_{k=1}^n \sum_{l=1}^n A_{kl} = n\bar{\mathbf{X}}' \mathbf{T} \bar{\mathbf{X}}$. Since, $\text{Var}(\bar{\mathbf{X}}) = \frac{1}{n} \Sigma$, the first two identities come directly from the moments of a quadratic form

(Equations (A.5)). Hence $E(Q_n) = \text{tr}(\mathbf{T}\Sigma)$ and $\text{Var}(Q_n) = 2\text{tr}(\mathbf{T}\Sigma)^2$. For the covariance of Q_n and B_0 , we write Q_n as

$$Q_n = \frac{1}{n} \left(\sum_{k=1}^n A_k + \underbrace{\sum_{k=1}^n \sum_{\substack{l=1 \\ k \neq l}}^n A_{kl}} \right) \quad (2.16)$$

where A_k is the quadratic form when $k = l$ and A_{kl} is the bilinear form when $k \neq l$. Since $B_0 = \frac{1}{n} \sum_{k=1}^n A_k$ where A_k is again the quadratic form, therefore, we have

$$\begin{aligned} \text{Cov}(Q_n, B_0) &= \frac{1}{n^2} \left[\text{Cov} \left(\sum_{k=1}^n \sum_{l=1}^n A_{kl}, \sum_{m=1}^n A_m \right) + \text{Cov} \left(\sum_{k=1}^n A_k, \sum_{m=1}^n A_m \right) \right] \\ &= \frac{1}{n^2} \text{Cov} \left(\sum_{k=1}^n A_k, \sum_{m=1}^n A_m \right) = \frac{1}{n^2} \text{Var} \left(\sum_{k=1}^n A_k \right) \\ &= \text{Var}(B_0) = \frac{2}{n} \text{tr}(\mathbf{T}\Sigma)^2 \end{aligned}$$

from Equations (A.22). This completes the proof. ■

The statistic $\tilde{A}_n = Q_n/B_0$ is the ratio of two random variables. Substituting the results from Theorems 2.4 and 2.5 into the approximation formulas, Equations B.1 and B.2, we get the following moments of $\tilde{A}_n$. Note that, in the following theorem, the order of the remainder for the mean of the statistic is at least $O(n^{-2})$ and for the variance is at least $o(n^{-2})$ (see Stuart and Ord, 1994, p 350). We compute the moments ignoring these remainders for approximation purposes (see Casella and Berger, 2002, p 245).

Theorem 2.6 *The first two moments of $\tilde{A}_n$ are computed as*

$$\begin{aligned} E(\tilde{A}_n) &\approx 1 \\ \text{Var}(\tilde{A}_n) &\approx \frac{2}{f} \left(1 - \frac{1}{n} \right) \end{aligned}$$

where $f = \frac{[\text{tr}(\mathbf{T}\Sigma)]^2}{\text{tr}(\mathbf{T}\Sigma)^2}$. □

Note that the first two moments of the χ_f^2/f -distribution are, respectively, 1 and $2/f$. Comparing with the corresponding moments of $\tilde{A}_n$, we observe that the sampling distribution of $\tilde{A}_n$ is very closely approximated by the χ_f^2/f -distribution, especially when $n \rightarrow \infty$. Moreover, the moments of $\tilde{A}_n$ are uniformly bounded with respect to d .

For convenience, we can write

$$\tilde{F}_n = \tilde{A}_n \cdot \frac{B_1}{B_2}$$

such that $E(\tilde{F}_n) \approx f$ and $\text{Var}(\tilde{F}_n) \approx 2f$ where the moments, now, correspond to the first two moments of the χ_f^2 -distribution; hence $\tilde{F}_n \approx \chi_f^2$, as $n \rightarrow \infty$.

2.4 Simulation Results

2.4.1 Moments of Estimators

Table 2.5: PROPERTIES OF ESTIMATORS

n	d	B_0		B_1		B_2		$\tilde{f}$	
		$\frac{E(B_0)}{\text{tr}(\mathbf{T}\mathbf{\Sigma})}$	$\frac{\text{Var}(B_0)}{[\text{tr}(\mathbf{T}\mathbf{\Sigma})]^2}$	$\frac{E(B_1)}{[\text{tr}(\mathbf{T}\mathbf{\Sigma})]^2}$	$\frac{\text{Var}(B_1)}{[\text{tr}(\mathbf{T}\mathbf{\Sigma})]^4}$	$\frac{E(B_2)}{\text{tr}(\mathbf{T}\mathbf{\Sigma})^2}$	$\frac{\text{Var}(B_2)}{[\text{tr}(\mathbf{T}\mathbf{\Sigma})^2]^2}$	$\frac{E(f)}{f}$	$\frac{\text{Var}(f)}{f^2}$
10	20	1.0004	0.0107	1.0011	0.0429	0.9994	0.0889	1.1628	0.0611
	50	0.9997	0.0041	0.9994	0.0165	0.9973	0.0598	1.1639	0.0643
	100	0.9995	0.0020	0.9990	0.0080	0.9979	0.0543	1.1652	0.0672
	300	1.0000	0.0007	1.0001	0.0026	0.9999	0.0474	1.1631	0.0659
	500	1.0001	0.0004	1.0003	0.0016	1.0008	0.0464	1.1617	0.0641
	1000	1.0000	0.0002	1.0001	0.0008	0.9999	0.0465	1.1645	0.0683
20	30	0.9995	0.0033	0.9990	0.0133	0.9990	0.0244	1.0639	0.0120
	50	0.9997	0.0021	0.9994	0.0082	0.9983	0.0186	1.0649	0.0121
	100	1.0000	0.0010	1.0000	0.0041	1.0000	0.0145	1.0638	0.0119
	300	1.0000	0.0003	1.0005	0.0013	0.9991	0.0121	1.0655	0.0124
	500	1.0000	0.0002	0.9999	0.0008	0.9977	0.0113	1.0662	0.0122
	1000	0.9998	0.0001	0.9996	0.0004	1.0007	0.0111	1.0627	0.0121
50	100	0.9999	0.0004	0.9999	0.0016	1.0005	0.0033	1.0215	0.0018
	300	0.9999	0.0001	0.9999	0.0005	1.0000	0.0021	1.0223	0.0017
	500	0.9998	0.0001	1.0000	0.0003	1.0000	0.0019	1.0221	0.0017
	1000	1.0000	0.0000	0.9999	0.0002	1.0001	0.0018	1.0219	0.0017

First, the moments of the estimators, B_0, B_1 and B_2 , and of degrees of freedom, $\tilde{f}$, are studied. Table 2.5 reports the results where $n \in \{10, 20, 50\}$, $d \in \{20, 50, 100, 300, 500, 1000\}$ and $\mathbf{\Sigma}$ is assumed to be compound symmetric. But, as is clear from the theoretical results, the behavior of the estimators is generally true for any covariance structure.

Three points are important to note in Table 2.5. First, the estimators do not need even moderately large sample size to be unbiased and consistent; even $n = 10$ is sufficient. Second, the expectation of the degrees of freedom vary very closely around the target value, 1, and gets closer and closer for increasing sample size. The variance already vanishes even for n as small as 10 or 20. Third, the relationships remain stable around the target values irrespective of the number of dimensions. Note that even for d as large as 1000, the stability of the moments is quite evident.

Table 2.6: ESTIMATED QUANTILES OF $\tilde{A}_n$: *Normal Distribution*

n	d	CS			AR(0.6)			UN		
		0.90	0.95	0.99	0.90	0.95	0.99	0.90	0.95	0.99
10	20	0.9008	0.9542	0.9944	0.8961	0.9480	0.9924	0.9017	0.9544	0.9925
	50	0.9072	0.9491	0.9938	0.8965	0.9485	0.9900	0.8967	0.9480	0.9911
	100	0.9084	0.9535	0.9949	0.8980	0.9482	0.9902	0.9012	0.9496	0.9915
	200	0.9089	0.9540	0.9935	0.8950	0.9472	0.9902	0.8973	0.9485	0.9913
20	50	0.9006	0.9476	0.9913	0.8971	0.9442	0.9873	0.9050	0.9520	0.9911
	100	0.9015	0.9522	0.9901	0.9025	0.9510	0.9892	0.8983	0.9491	0.9888
	200	0.9038	0.9524	0.9921	0.8983	0.9469	0.9884	0.8989	0.9524	0.9911

2.4.2 Level

Table 2.6 reports the estimated quantiles $(1 - \alpha)$ for the modified ANOVA-type statistic for the same three covariance structures discussed in the context of Table 2.2. These are the results of 10,000 simulation runs with $n \in \{10, 20\}$ and $d \in \{20, 50, 100, 200\}$ where $\alpha \in \{0.01, 0.05, 0.10\}$. The test statistic very accurately estimates the nominal quantiles even for n as small as 10 and d as large as 200. The accuracy is not affected by increasing the dimension and by changing the covariance structure.

2.4.3 Power

Power of the statistic is also examined for $n = 10$, $d \in \{20, 50, 100\}$ and the same covariance structures, CS, AR(0.6) and UN. The corresponding power curves are shown in Fig. 2.1. For these power computations, we define the alternative as

$$\mu_Y = \delta_i \mathbf{d}^{-1} \quad (2.17)$$

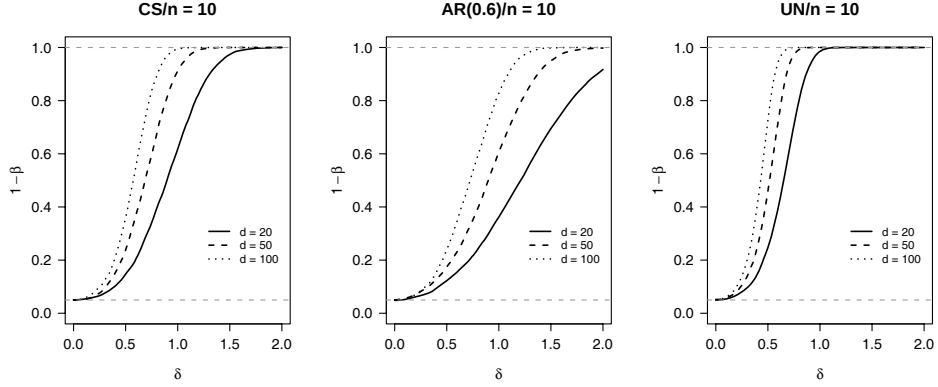


Figure 2.1: POWER CURVES FOR $\tilde{A}_n$: *Normal Distribution*

where $\boldsymbol{\mu}_Y = E(\mathbf{Y}_k) = E(\mathbf{T}\mathbf{X}_k)$, $\mathbf{d}^{-1} = (\frac{1}{d}, \frac{2}{d}, \dots, \frac{d}{d})'$ and δ_i is the i th element of the vector of constants, $\boldsymbol{\delta} = 0(0.02)2$.

We observe that the power of ATS is not only very high, even for $n = 10$, but also increases for increasing d , for all covariance structures. Note that, when $n > d$, the statistic can also be compared with the Hotelling's T^2 statistic. Although the results are not reported here, but it has been examined that none of these statistics is uniformly better than the other and one of them outperforms the other depending on the covariance structure and the type of alternative. Further, both level and power computations for ATS were also carried out for other covariance structures, for example, Heterogeneous AR(1) structure or AR(1) structure with low correlation, and the performance of the statistic was found to be similar to the one reported above for CS, AR(0.6) and UN structures. This fact is also clear from the analytical results where we do not assume any specific covariance structure and derive the results under very general conditions.

2.5 Applications

2.5.1 Analysis of Sleep Lab Data

The test statistic is used to analyze the sleep lab example introduced in Section 1.4 (Example 1.3). The complete data, produced in Section C.3, consists of two independent groups, of men and women. Here only the data

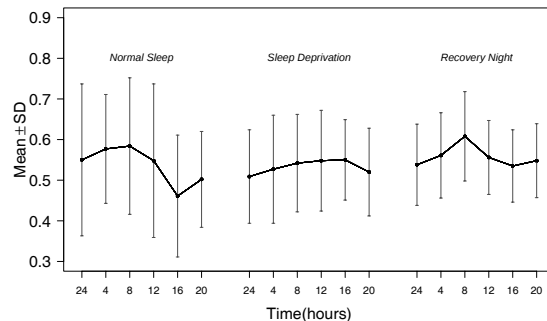


Figure 2.2: BAR PLOT FOR SLEEP LAB DATA

on the group of women is analyzed. The analysis of complete data is given in Section 6.4.2. The variable of interest is serum concentration measured on each of $n = 10$ women at $d = 18$ time points. Fig 2.2 depicts the mean profiles of the subjects for three consecutive nights.

Table 2.7: ANALYSIS OF SLEEP LAB DATA

Effect	$\mathbf{T}$ -matrix	$\tilde{F}_n$	$\tilde{f}$	p -value
Day	$\mathbf{P}_3 \otimes \frac{1}{6}\mathbf{J}_6$	0.25	1.55	0.7200
Time	$\frac{1}{3}\mathbf{J}_3 \otimes \mathbf{P}_6$	4.04	3.86	0.0032
Day $\times$ Time	$\mathbf{P}_3 \otimes \mathbf{P}_6$	2.66	9.24	0.0040
Time(NS-TSD)	$\mathbf{P}_2 \otimes \mathbf{P}_6$	3.39	4.67	0.0057
Time(NS-RN)	$\mathbf{P}_2 \otimes \mathbf{P}_6$	2.51	5.00	0.0279
Time(TSD-RN)	$\mathbf{P}_2 \otimes \mathbf{P}_6$	1.54	7.60	0.1404
Time (NS)	$\mathbf{P}_6$	4.54	2.98	0.0035
Time (TSD)	$\mathbf{P}_6$	1.08	8.46	0.3709
Time (RN)	$\mathbf{P}_6$	2.80	4.74	0.0175

Table 2.7 reports the results where the $\mathbf{T}$ -matrix represents the hypothesis matrix for the corresponding effect being tested. For each effect, $\tilde{F}_n$ follows a $\chi^2_{\tilde{f}}$ -distribution with $\tilde{f}$ degrees of freedom as reported in the penultimate column of the table. We observe highly significant Time main effect and Day $\times$ Time interaction effect. The significant interaction effect, shown in Fig. 2.3, is further explored, in the middle panel of the table, to study the effect for each pair of days. Note that, in this case, $n = 10$ and $d = 12$.

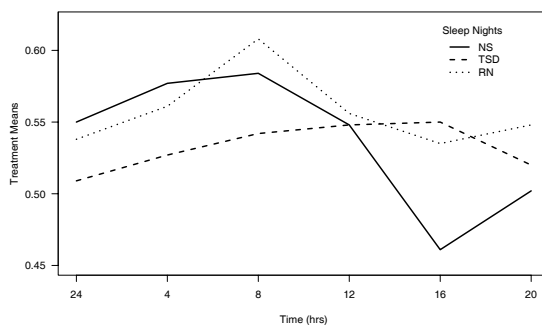


Figure 2.3: DAY×TIME INTERACTION PLOT FOR SLEEP LAB DATA

We observe that the interaction is significant for the pairs wherein NS is common, i.e. NS-TSD and NS-RN. The last panel of the table explores the same interaction further by evaluating the simple effects of time for each day level. Clearly, as expected, the contribution to the interaction from the simple effects of NS and RN is highly significant.

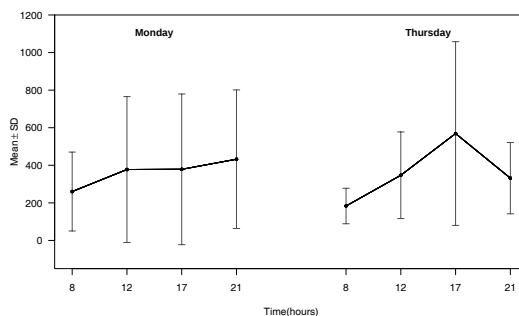


Figure 2.4: BAR PLOT FOR AMYLASE DATA

2.5.2 Analysis of Amylase Data

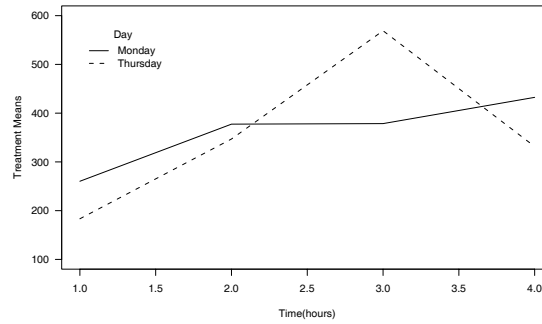
As stated in Section 1.4, the Amylase data (Example 1.1) is not high dimensional but is included to show that the modified ATS is also applicable when $d < n$. Table 2.8 reports the results for this two-factor factorial experiment

wherein each of 14 probands is repeatedly observed at 8 time points, 4 on each of two days (Monday and Thursday). The average treatment combination effects, shown in Fig. 2.4, already clue to a significant within-subject variation.

Table 2.8: ANALYSIS OF AMYLASE DATA

Effect	T -matrix	$\tilde{F}_n$	$\tilde{f}$	p -value
Day	$\mathbf{P}_2 \otimes \frac{1}{4}\mathbf{J}_4$	0.02	1.08	0.9119
Time	$\frac{1}{2}\mathbf{J}_2 \otimes \mathbf{P}_4$	4.09	2.92	0.0070
Day $\times$ Time	$\mathbf{P}_2 \otimes \mathbf{P}_4$	3.43	2.52	0.0226
Time(Monday)	$\mathbf{P}_4$	1.65	4.88	0.1450
Time(Thursday)	$\mathbf{P}_4$	5.39	1.36	0.0119

We observe, from Table 2.8, that the significant interaction effect masks the significant time effect. This significant interaction effect (Fig. 2.5) is further explored for simple effects for each day, presented in the lower panel of Table 2.8. We observe a significant effect of Thursday on the α -amylase in saliva, a fact also evident in the bar and interaction graphs.

Figure 2.5: DAY $\times$ TIME INTERACTION PLOT FOR AMYLASE DATA

Chapter 3

The Two-sample Statistic

Here we begin with the two sample extension to the modified ANOVA-type statistic developed for the one sample case in Chapter 2.

3.1 The Statistical Model

3.1.1 Model and Assumptions

Suppose $\mathbf{X}_{1k} = (X_{1k1}, \dots, X_{1kd})'$ and $\mathbf{X}_{2l} = (X_{2l1}, \dots, X_{2ld})'$ be the vectors, each of d repeated observations measured on k th subject in sample 1 and l th subject in sample 2, respectively, where the samples are drawn independently from two populations. The observations from such a design can be arranged as shown in Table 3.1.

Table 3.1: THE GENERAL DATA TABLE: TWO-SAMPLE CASE

	Sample 1						Sample 2				
	1	...	s	...	d		1	...	s	...	d
1	X_{111}	...	X_{11s}	...	X_{11d}	1	X_{211}	...	X_{21s}	...	X_{21d}
$\vdots$	$\vdots$	$\ddots$		$\ddots$	$\vdots$	$\vdots$	$\vdots$	$\ddots$		$\ddots$	$\vdots$
k	X_{1k1}	...	X_{1ks}	...	X_{1kd}	l	X_{2l1}	...	X_{2ls}	...	X_{2ld}
$\vdots$	$\vdots$	$\ddots$		$\ddots$	$\vdots$	$\vdots$	$\vdots$	$\ddots$		$\ddots$	$\vdots$
n_1	X_{1n_11}	...	X_{1n_1s}	...	X_{1n_1d}	n_2	X_{2n_21}	...	X_{2n_2s}	...	X_{2n_2d}
Means	$\bar{X}_{1.1}$	...	$\bar{X}_{1.s}$	...	$\bar{X}_{1.d}$		$\bar{X}_{2.1}$	...	$\bar{X}_{2.s}$	...	$\bar{X}_{2.d}$

Let $E(\mathbf{X}_{1k}) = \boldsymbol{\mu}_1$, $E(\mathbf{X}_{2l}) = \boldsymbol{\mu}_2$, $\text{Cov}(\mathbf{X}_{1k}) = \boldsymbol{\Sigma}_1$ and $\text{Cov}(\mathbf{X}_{2l}) = \boldsymbol{\Sigma}_2$. Then $\mathbf{X} = (\mathbf{X}'_{1k}, \mathbf{X}'_{2l})'$ is the vector of all observations from k th subject in sample 1 and l th subject in sample 2, $k = 1, \dots, n_1$, $l = 1, \dots, n_2$, with $E(\mathbf{X}) = \boldsymbol{\mu} = (\boldsymbol{\mu}'_1, \boldsymbol{\mu}'_2)'$ and $\text{Cov}(\mathbf{X}) = \boldsymbol{\Sigma} = \boldsymbol{\Sigma}_1 \oplus \boldsymbol{\Sigma}_2$. The corresponding sample estimators are $\bar{\mathbf{X}} = (\bar{\mathbf{X}}'_1, \bar{\mathbf{X}}'_2)'_{2d \times 1}$ and $\hat{\boldsymbol{\Sigma}} = \hat{\boldsymbol{\Sigma}}_1 \oplus \hat{\boldsymbol{\Sigma}}_2$, respectively, where $\bar{\mathbf{X}}_i = (\bar{X}_{i.1}, \dots, \bar{X}_{i.d})'_{d \times 1}$ is the vector of means of i th sample, $i = 1, 2$, as shown at the bottom of Table 3.1.

Like in one sample case, we do not assume any structure of $\boldsymbol{\Sigma}$ to keep the results valid in general, whereas for simulation studies, compound symmetric, autoregressive and unstructured covariance patterns (Definitions B.6, B.7 and B.8) will be used to typically show the behavior of the test statistic.

3.1.2 Formulation of Hypotheses

Let $\mathbf{1}_d$ be a vector of 1s and $\mathbf{P}_d = \mathbf{I}_d - \frac{1}{d}\mathbf{J}_d$ be the centering matrix, where $\mathbf{I}_d$ is the identity matrix and $\mathbf{J}_d = \mathbf{1}_d\mathbf{1}'_d$. For the data set up given above, we are interested to test the following three hypotheses.

$$H_0^{AB} : (\mathbf{P}_2 \otimes \mathbf{P}_d) \boldsymbol{\mu} = \mathbf{0} \quad (\text{Interaction Hypothesis}) \quad (3.1)$$

$$H_0^A : (\mathbf{P}_2 \otimes \frac{1}{d}\mathbf{1}'_d) \boldsymbol{\mu} = \mathbf{0} \quad (\text{Group Hypothesis}) \quad (3.2)$$

$$H_0^B : (\frac{1}{2}\mathbf{1}'_2 \otimes \mathbf{P}_d) \boldsymbol{\mu} = \mathbf{0} \quad (\text{Time Hypothesis}) \quad (3.3)$$

where $\otimes$ denotes the Kronecker product. Replacing 2 with a in Equations (3.1)-(3.3) will give us the corresponding hypotheses for the general case with a populations. We use $\mathbf{T} = \mathbf{H}'(\mathbf{H}\mathbf{H}')^{-}\mathbf{H}$ to uniquely represent each of the three hypotheses where $\mathbf{T}$ is a projection matrix and $\mathbf{H}\boldsymbol{\mu} = \mathbf{0} \iff \mathbf{T}\boldsymbol{\mu} = \mathbf{0}$. The $\mathbf{T}$ matrices for Equations (3.1)-(3.3) are, respectively,

$$\left. \begin{aligned} \mathbf{T}^{AB} &= (\mathbf{P}_2 \otimes \mathbf{P}_d) \\ \mathbf{T}^A &= (\mathbf{P}_2 \otimes \frac{1}{d}\mathbf{J}_d) \\ \mathbf{T}^B &= (\frac{1}{2}\mathbf{J}_2 \otimes \mathbf{P}_d) \end{aligned} \right\} \quad (3.4)$$

These $\mathbf{T}$ matrices will be used to define the estimators and the test statistic in the sequel that follows. For convenience, it is sufficient to consider even the simpler versions of the $\mathbf{T}$ matrices. For instance, for the interaction effect, $H_0^{AB} : (\mathbf{P}_2 \otimes \mathbf{P}_d) \boldsymbol{\mu} \iff H_0^{AB} : \frac{1}{2}(\mathbf{1} - \mathbf{1}) \otimes \mathbf{P}_d$ so that $\mathbf{P}_d$ can replace $\mathbf{T}^{AB}$ as the generating matrix without any loss of generality, where $\mathbf{P}_d$ is a projection matrix. Similarly, $\mathbf{P}_d$ is also the generating matrix for H_0^B .

Remark 3.1 It is clear from Equations (3.1)-(3.3) that the high dimensionality influences only H_0^{AB} and H_0^B whereas H_0^A is a univariate hypothesis and does not depend on d . H_0^A , therefore, is not of our main interest. But to keep the analysis complete and the notations coherent, the analysis of group effect will be tangentially touched in Section 3.4. $\triangle$

3.1.3 Formulation of the Test Statistic

As is clear from Equations (3.1)-(3.3), the formulation of the test statistic for the interaction effect involves differences while that for the time effect involves sums of the vectors $\mathbf{X}_{1k}$ and $\mathbf{X}_{2l}$. Therefore, here we give only a general formulation of the test statistic which follows exactly the same lines as used in Section 2.1 for the one sample case.

Let $\mathbf{Z}_{kl}$ denote the vector of sums or differences of the vectors $\mathbf{T}\mathbf{X}_{1k}$ and $\mathbf{T}\mathbf{X}_{2l}$ with $\text{Cov}(\mathbf{Z}_{kl}) = \mathbf{T}(\boldsymbol{\Sigma}_1 + \boldsymbol{\Sigma}_2)\mathbf{T}$. Note that $\text{Cov}(\mathbf{Z}_{kl})$ remains same whether $\mathbf{Z}_{kl}$ is composed of sums or differences. This fact will greatly help us minimizing the computational burden since, as will also be discussed later on, the final results for the interaction and time effects are same. Therefore, the detailed derivations will be presented only for the interaction effect where for the time effect the results follow the same pattern.

Using $\mathbf{Z}_{kl}$, the two-sample ANOVA-type statistic is given as

$$\tilde{A}_N = \frac{\mathbf{Z}'_{kl}\mathbf{T}\mathbf{Z}_{kl}}{\text{tr}[\mathbf{T}(\boldsymbol{\Sigma}_1 + \boldsymbol{\Sigma}_2)\mathbf{T}]} \quad (3.5)$$

where $\tilde{A}_N \sim \frac{\chi_f^2}{f}$ with

$$f = \frac{[\text{tr}[\mathbf{T}(\boldsymbol{\Sigma}_1 + \boldsymbol{\Sigma}_2)\mathbf{T}]]^2}{\text{tr}[\mathbf{T}(\boldsymbol{\Sigma}_1 + \boldsymbol{\Sigma}_2)\mathbf{T}]^2}$$

and $N = n_1 + n_2$.

3.2 The Interaction Effect

First, the test statistic for the interaction effect will be derived. From Equation (3.1), $H_0^{AB} : (\mathbf{P}_2 \otimes \mathbf{P}_d)\boldsymbol{\mu} = \mathbf{0}$, which can also be written as $H_0^{AB} : \mathbf{P}_d(\boldsymbol{\mu}_1 - \boldsymbol{\mu}_2) = \mathbf{0}$. As already mentioned, the generating matrix for the interaction effect is $\mathbf{T} = \mathbf{P}_d$, although for convenience, the general form of the hypothesis matrix, $\mathbf{T}$, will be used.

3.2.1 The Covariance Matrix

Let $\mathbf{X}_{1k}$ and $\mathbf{X}_{2l}$ be as defined in Section 3.1. Define the differences $\mathbf{X}_{1k} - \mathbf{X}_{2l}$, $\forall k = 1 \dots, n_1, l = 1, \dots, n_2$, where $E(\mathbf{X}_{1k} - \mathbf{X}_{2l}) = \boldsymbol{\mu}_1 - \boldsymbol{\mu}_2$ and $\text{Cov}(\mathbf{X}_{1k} - \mathbf{X}_{2l}) = \boldsymbol{\Sigma}_1 + \boldsymbol{\Sigma}_2$. Let $\mathbf{T}(\mathbf{X}_{1k} - \mathbf{X}_{2l}) = \mathbf{Y}_{1k} - \mathbf{Y}_{2l}$, with $E(\mathbf{Y}_{1k} - \mathbf{Y}_{2l}) = \mathbf{0}$ and $\text{Cov}(\mathbf{Y}_{1k} - \mathbf{Y}_{2l}) = \mathbf{T}(\boldsymbol{\Sigma}_1 + \boldsymbol{\Sigma}_2)\mathbf{T}$, under H_0^{AB} , where $\mathbf{T}(\boldsymbol{\Sigma}_1 + \boldsymbol{\Sigma}_2)\mathbf{T}$ is estimated as

$$\hat{\boldsymbol{\Sigma}}_N = \frac{1}{n_1 n_2} \sum_{k=1}^{n_1} \sum_{l=1}^{n_2} (\mathbf{Y}_{1k} - \mathbf{Y}_{2l})(\mathbf{Y}_{1k} - \mathbf{Y}_{2l})'. \quad (3.6)$$

As will be seen later, the identities involved in the statistic are conveniently tractable when expressed in terms of matrices. To begin with, note that $\mathbf{X}_{1k} \sim \mathcal{N}_d(\boldsymbol{\mu}_1, \boldsymbol{\Sigma}_1)$, $\mathbf{X}_{2l} \sim \mathcal{N}_d(\boldsymbol{\mu}_2, \boldsymbol{\Sigma}_2)$. Let $\mathbf{T}\mathbf{X}_{1k} = \mathbf{Y}_{1k}$, $\mathbf{T}\mathbf{X}_{2l} = \mathbf{Y}_{2l}$ so that $\mathbf{Y}_{1k} \sim \mathcal{N}_d(\mathbf{T}\boldsymbol{\mu}_1, \mathbf{T}\boldsymbol{\Sigma}_1\mathbf{T})$, $\mathbf{Y}_{2l} \sim \mathcal{N}_d(\mathbf{T}\boldsymbol{\mu}_2, \mathbf{T}\boldsymbol{\Sigma}_2\mathbf{T})$. Define $\mathbf{Y} = (\mathbf{Y}'_1, \mathbf{Y}'_2)'$, where $\mathbf{Y}_i = (\mathbf{Y}'_{i1}, \dots, \mathbf{Y}'_{in_i})$, $i = 1, 2$. Denote $E(\mathbf{Y}) = \boldsymbol{\xi}$, $\text{Cov}(\mathbf{Y}) = \boldsymbol{\Xi}$, with

$$\boldsymbol{\xi} = (\mathbf{1}'_{n_1} \otimes \boldsymbol{\mu}_1 \quad \mathbf{1}'_{n_2} \otimes \boldsymbol{\mu}_2)' \quad (3.7)$$

$$\boldsymbol{\Xi} = (\mathbf{I}_{n_1} \otimes \mathbf{T}\boldsymbol{\Sigma}_1\mathbf{T}) \oplus (\mathbf{I}_{n_2} \otimes \mathbf{T}\boldsymbol{\Sigma}_2\mathbf{T}). \quad (3.8)$$

Further, define $\mathbf{M} = (\mathbf{I}_{n_1} \otimes \mathbf{1}_{n_2} \mid -\mathbf{1}_{n_1} \otimes \mathbf{I}_{n_2}) \otimes \mathbf{I}_d$ such that $\mathbf{Z} = \mathbf{M}\mathbf{Y}$ denotes the vector of differences, $\mathbf{Z}_{kl} = \mathbf{Y}_{1k} - \mathbf{Y}_{2l}$, $\forall k, l$, with $E(\mathbf{Z}_{kl}) = \mathbf{0}$ and $\text{Cov}(\mathbf{Z}_{kl}) = \mathbf{T}(\boldsymbol{\Sigma}_1 + \boldsymbol{\Sigma}_2)\mathbf{T}$. Then

$$\hat{\boldsymbol{\Sigma}}_N = \frac{1}{n_1 n_2} \sum_{k=1}^{n_1} \sum_{l=1}^{n_2} \mathbf{Z}_{kl} \mathbf{Z}'_{kl}. \quad (3.9)$$

Under H_0^{AB} , $\hat{\boldsymbol{\Sigma}}_N$ is an unbiased estimator of $\mathbf{T}(\boldsymbol{\Sigma}_1 + \boldsymbol{\Sigma}_2)\mathbf{T}$, as proved in the following theorem.

Theorem 3.2 *The covariance estimator, $\hat{\boldsymbol{\Sigma}}_N$, defined in Equation (5.4), is an unbiased estimator of $\mathbf{T}(\boldsymbol{\Sigma}_1 + \boldsymbol{\Sigma}_2)\mathbf{T}$. $\square$*

Proof Clearly, $\hat{\boldsymbol{\Sigma}}_N$ is just the average of $n_1 n_2$ matrices of the form $\mathbf{Z}_{kl} \mathbf{Z}'_{kl}$, each of order $d \times d$, and we need the expectation of this sum. Since

$$E(\mathbf{Z}_{kl} \mathbf{Z}'_{kl}) = E(\mathbf{Y}_{1k} - \mathbf{Y}_{2l})(\mathbf{Y}_{1k} - \mathbf{Y}_{2l})' = \mathbf{T}E[(\mathbf{X}_{1k} - \mathbf{X}_{2l})(\mathbf{X}_{1k} - \mathbf{X}_{2l})']\mathbf{T},$$

therefore, by the independence of $\mathbf{X}_{1k}$ and $\mathbf{X}_{2l}$, we get, after simplification,

$$\begin{aligned} E(\mathbf{X}_{1k} - \mathbf{X}_{2l})(\mathbf{X}_{1k} - \mathbf{X}_{2l})' &= E[\{(\mathbf{X}_{1k} - \boldsymbol{\mu}_1) - (\mathbf{X}_{2l} - \boldsymbol{\mu}_2) + (\boldsymbol{\mu}_1 - \boldsymbol{\mu}_2)\} \\ &\quad \{(\mathbf{X}_{1k} - \boldsymbol{\mu}_1) - (\mathbf{X}_{2l} - \boldsymbol{\mu}_2) + (\boldsymbol{\mu}_1 - \boldsymbol{\mu}_2)\}'] \\ &= \boldsymbol{\Sigma}_1 + \boldsymbol{\Sigma}_2 + (\boldsymbol{\mu}_1 - \boldsymbol{\mu}_2)(\boldsymbol{\mu}_1 - \boldsymbol{\mu}_2)' \end{aligned}$$

where, $E[(\mathbf{X}_{1k} - \boldsymbol{\mu}_1)(\mathbf{X}_{1k} - \boldsymbol{\mu}_1)'] = \boldsymbol{\Sigma}_1$ and $E[(\mathbf{X}_{2l} - \boldsymbol{\mu}_2)(\mathbf{X}_{2l} - \boldsymbol{\mu}_2)'] = \boldsymbol{\Sigma}_2$. Under $H_0 : \mathbf{T}(\boldsymbol{\mu}_1 - \boldsymbol{\mu}_2) = \mathbf{0}$,

$$E(\mathbf{Y}_{1k} - \mathbf{Y}_{2l})(\mathbf{Y}_{1k} - \mathbf{Y}_{2l})' = \mathbf{T}(\boldsymbol{\Sigma}_1 + \boldsymbol{\Sigma}_2)\mathbf{T}.$$

and the result follows immediately. $\blacksquare$

3.2.2 Defining The Statistic

Following the general formulation of the ANOVA-type statistic, explained in Section 3.1.3, we proceed as following. Define $\bar{\mathbf{Y}}_1, -\bar{\mathbf{Y}}_2 = \mathbf{T}(\bar{\mathbf{X}}_1, -\bar{\mathbf{X}}_2)$ such that, under H_0^{AB} , $E(\bar{\mathbf{Y}}_1, -\bar{\mathbf{Y}}_2) = \mathbf{0}$ and $\text{Var}(\bar{\mathbf{Y}}_1, -\bar{\mathbf{Y}}_2) = \mathbf{T} \left(\frac{1}{n_1} \boldsymbol{\Sigma}_1 + \frac{1}{n_2} \boldsymbol{\Sigma}_2 \right) \mathbf{T}$. Then, with $N = n_1 + n_2$, Equation (4.5) gives

$$\tilde{A}_N^{AB} = \frac{\frac{2n_1n_2}{N}(\bar{\mathbf{Y}}_1, -\bar{\mathbf{Y}}_2)'(\bar{\mathbf{Y}}_1, -\bar{\mathbf{Y}}_2)}{\text{tr}[\mathbf{T}(\boldsymbol{\Sigma}_1 + \boldsymbol{\Sigma}_2)\mathbf{T}]} \quad (3.10)$$

where $\tilde{A}_N^{AB} \sim \chi_{f^{AB}}^2 / f^{AB}$ with

$$f^{AB} = \frac{[\text{tr}\{\mathbf{T}(\boldsymbol{\Sigma}_1 + \boldsymbol{\Sigma}_2)\mathbf{T}\}]^2}{\text{tr}[\mathbf{T}(\boldsymbol{\Sigma}_1 + \boldsymbol{\Sigma}_2)\mathbf{T}]^2}. \quad (3.11)$$

Denote $\tilde{B}_0^{AB}$, $\tilde{B}_1^{AB}$ and $\tilde{B}_2^{AB}$ as the estimators of $\text{tr}[\mathbf{T}(\boldsymbol{\Sigma}_1 + \boldsymbol{\Sigma}_2)\mathbf{T}]$, $[\text{tr}\{\mathbf{T}(\boldsymbol{\Sigma}_1 + \boldsymbol{\Sigma}_2)\mathbf{T}\}]^2$ and $\text{tr}[\mathbf{T}(\boldsymbol{\Sigma}_1 + \boldsymbol{\Sigma}_2)\mathbf{T}]^2$, respectively, where

$$\tilde{B}_0^{AB} = \frac{1}{n_1n_2} \sum_{k=1}^{n_1} \sum_{l=1}^{n_2} A_{kl}^{AB} \quad (3.12)$$

$$\tilde{B}_1^{AB} = \frac{1}{n_1n_2(n_1-1)(n_2-1)} \underbrace{\sum_{k=1}^{n_1} \sum_{l=1}^{n_2} \sum_{r=1}^{n_1} \sum_{s=1}^{n_2} A_{kl}^{AB} A_{rs}^{AB}}_{k \neq r, l \neq s} \quad (3.13)$$

$$\tilde{B}_2^{AB} = \frac{1}{n_1n_2(n_1-1)(n_2-1)} \underbrace{\sum_{k=1}^{n_1} \sum_{l=1}^{n_2} \sum_{r=1}^{n_1} \sum_{s=1}^{n_2} A_{klrs}^{AB^2}}_{k \neq r, l \neq s} \quad (3.14)$$

with $A_{kl}^{AB} = (\mathbf{Y}_{1k} - \mathbf{Y}_{2l})'(\mathbf{Y}_{1k} - \mathbf{Y}_{2l})$ and $A_{rs}^{AB} = (\mathbf{Y}_{1r} - \mathbf{Y}_{2s})'(\mathbf{Y}_{1r} - \mathbf{Y}_{2s})$ as the quadratic forms, $A_{klrs}^{AB} = (\mathbf{Y}_{1k} - \mathbf{Y}_{2l})'(\mathbf{Y}_{1r} - \mathbf{Y}_{2s})$ as the symmetric

bilinear form (see Sections A.1 and A.2).

Note that n_1, n_2 and Σ_1, Σ_2 , are implicitly coalesced with each other in $\text{Var}(\bar{\mathbf{Y}}_1 - \bar{\mathbf{Y}}_2)$ which seems to be computationally tedious to deal with. For convenience, we consider the following two cases.

Case I: $n_1 \neq n_2, \Sigma_1 = \Sigma_2$

Case II: $n_1 = n_2, \Sigma_1 \neq \Sigma_2$

The estimators, their properties and the asymptotic distribution of the statistic will be discussed according to this bifurcation. But to avoid many lengthy computations, we first give some general results in the following theorem, the special cases of which will then be extracted for cases I and II.

Denote $Q_N^{AB} = (\bar{\mathbf{Y}}_1 - \bar{\mathbf{Y}}_2)'(\bar{\mathbf{Y}}_1 - \bar{\mathbf{Y}}_2)$ in Equation (5.5). Then

$$Q_N^{AB} = \frac{2}{Nn_1n_2} \mathbf{Y}' \mathbf{M}' (\mathbf{J}_{n_1} \otimes \mathbf{J}_{n_2} \otimes \mathbf{I}_d) \mathbf{M} \mathbf{Y} \quad (3.15)$$

$$\tilde{B}_0^{AB} = \frac{1}{n_1n_2} \mathbf{Y}' \mathbf{M}' \mathbf{M} \mathbf{Y} \quad (3.16)$$

where $\mathbf{M}$ and $\mathbf{Y}$ are as defined in Section 3.2.1 and $N = n_1 + n_2$.

Theorem 3.3 Let Q_N^{AB} , $\tilde{B}_0^{AB}$, $\tilde{B}_1^{AB}$ and $\tilde{B}_2^{AB}$ be as defined above. Then,

$$E(\tilde{B}_0^{AB}) = \text{tr}[\mathbf{T}(\Sigma_1 + \Sigma_2)\mathbf{T}] \quad (3.17)$$

$$E(\tilde{B}_1^{AB}) = [\text{tr}\{\mathbf{T}(\Sigma_1 + \Sigma_2)\mathbf{T}\}]^2 \quad (3.18)$$

$$E(\tilde{B}_2^{AB}) = \text{tr}[\mathbf{T}(\Sigma_1 + \Sigma_2)\mathbf{T}]^2 \quad (3.19)$$

$$E(Q_N^{AB}) = \frac{2}{N} [n_2 \text{tr}(\mathbf{T}\Sigma_1\mathbf{T}) + n_1 \text{tr}(\mathbf{T}\Sigma_2\mathbf{T})] \quad (3.20)$$

$$\text{Var}(\tilde{B}_0^{AB}) = \frac{2}{n_1n_2} [n_2 \text{tr}(\mathbf{T}\Sigma_1\mathbf{T})^2 + 2\text{tr}(\mathbf{T}\Sigma_1\mathbf{T}\Sigma_2\mathbf{T}) + n_1 \text{tr}(\mathbf{T}\Sigma_2\mathbf{T})^2] \quad (3.21)$$

$$\text{Var}(Q_N^{AB}) = \frac{8}{N^2} \text{tr}[n_2(\mathbf{T}\Sigma_1\mathbf{T}) + n_1(\mathbf{T}\Sigma_2\mathbf{T})]^2 \quad (3.22)$$

$$\text{Cov}(Q_N^{AB}, \tilde{B}_0^{AB}) = \frac{4}{Nn_1n_2} \text{tr}[n_2(\mathbf{T}\Sigma_1\mathbf{T}) + n_1(\mathbf{T}\Sigma_2\mathbf{T})]^2. \quad (3.23)$$

□

Proof Let $\mathbf{M}'\mathbf{M} = \mathbf{U}$ in Equation (3.16). Then $\tilde{B}_0^{AB} = \frac{1}{n_1n_2} \mathbf{Y}'\mathbf{U}\mathbf{Y}$ where

$$\mathbf{U} = \begin{pmatrix} n_2 \mathbf{I}_{n_1} & -\mathbf{J}_{n_1 \times n_2} \\ -\mathbf{J}_{n_2 \times n_1} & n_1 \mathbf{I}_{n_2} \end{pmatrix} \otimes \mathbf{I}_d. \quad (3.24)$$

Now, with $\Xi = \text{Cov}(\mathbf{Y})$,

$$\mathbf{U}\Xi = \begin{pmatrix} n_2 \mathbf{I}_{n_1} \otimes \mathbf{T}\Sigma_1 \mathbf{T} & -\mathbf{J}_{n_1 \times n_2} \otimes \mathbf{T}\Sigma_2 \mathbf{T} \\ -\mathbf{J}_{n_2 \times n_1} \otimes \mathbf{T}\Sigma_1 \mathbf{T} & n_1 \mathbf{I}_{n_2} \otimes \mathbf{T}\Sigma_2 \mathbf{T} \end{pmatrix}$$

$$\text{and } (\mathbf{U}\Xi)^2 = \begin{pmatrix} \mathbf{K}_1 & \mathbf{K}_2 \\ \mathbf{K}_3 & \mathbf{K}_4 \end{pmatrix}$$

$$\begin{aligned} \text{where } \mathbf{K}_1 &= n_2^2 \mathbf{I}_{n_1} \otimes (\mathbf{T}\Sigma_1 \mathbf{T})^2 + n_2 \mathbf{J}_{n_1} \otimes (\mathbf{T}\Sigma_2 \mathbf{T}\Sigma_1 \mathbf{T}) \\ \mathbf{K}_2 &= -n_2 \mathbf{J}_{n_1 \times n_2} \otimes (\mathbf{T}\Sigma_1 \mathbf{T}\Sigma_2 \mathbf{T}) - n_1 \mathbf{J}_{n_1 \times n_2} \otimes (\mathbf{T}\Sigma_2 \mathbf{T})^2 \\ \mathbf{K}_3 &= -n_2 \mathbf{J}_{n_2 \times n_1} \otimes (\mathbf{T}\Sigma_1 \mathbf{T})^2 - n_1 \mathbf{J}_{n_2 \times n_1} \otimes (\mathbf{T}\Sigma_2 \mathbf{T}\Sigma_1 \mathbf{T}) \\ \mathbf{K}_4 &= n_1 \mathbf{J}_{n_2} \otimes (\mathbf{T}\Sigma_1 \mathbf{T}\Sigma_2 \mathbf{T}) + n_1^2 \mathbf{I}_{n_2} \otimes (\mathbf{T}\Sigma_2 \mathbf{T})^2. \end{aligned}$$

Then, using the properties of Kronecker product and of trace,

$$\text{tr}(\mathbf{U}\Xi) = n_1 n_2 \text{tr}[\mathbf{T}(\Sigma_1 + \Sigma_2)\mathbf{T}] \quad (3.25)$$

$$\text{tr}(\mathbf{U}\Xi)^2 = n_1 n_2 [n_2 \text{tr}(\mathbf{T}\Sigma_1 \mathbf{T})^2 + 2\text{tr}(\mathbf{T}\Sigma_1 \mathbf{T}\Sigma_2 \mathbf{T}) + n_1 \text{tr}(\mathbf{T}\Sigma_2 \mathbf{T})^2]. \quad (3.26)$$

Mean and variance of $\mathbf{Y}'\mathbf{U}\mathbf{Y}$ are obtained from Equations (A.8) and (A.9). Substitution of these two moments yields Equations (3.17) and (3.21).

Similarly, let $\mathbf{M}'(\mathbf{J}_{n_1} \otimes \mathbf{J}_{n_2} \otimes \mathbf{I}_d)\mathbf{M} = \mathbf{W}$ in Equation (3.15), so that

$$\mathbf{W} = \begin{pmatrix} n_2^2 \mathbf{J}_{n_1} & -n_1 n_2 \mathbf{J}_{n_1 \times n_2} \\ -n_1 n_2 \mathbf{J}_{n_2 \times n_1} & n_1^2 \mathbf{J}_{n_2} \end{pmatrix} \otimes \mathbf{I}_d \quad (3.27)$$

and $Q_N^{AB} = \frac{2}{N n_1 n_2} \mathbf{Y}'\mathbf{W}\mathbf{Y}$. With $\Xi = \text{Cov}(\mathbf{Y})$, we have

$$\mathbf{W}\Xi = \begin{pmatrix} n_2^2 \mathbf{J}_{n_1} \otimes \mathbf{T}\Sigma_1 \mathbf{T} & -n_1 n_2 \mathbf{J}_{n_1 \times n_2} \otimes \mathbf{T}\Sigma_2 \mathbf{T} \\ -n_1 n_2 \mathbf{J}_{n_2 \times n_1} \otimes \mathbf{T}\Sigma_1 \mathbf{T} & n_1^2 \mathbf{J}_{n_2} \otimes \mathbf{T}\Sigma_2 \mathbf{T} \end{pmatrix}$$

$$(\mathbf{W}\Xi)^2 = \begin{pmatrix} \mathbf{L}_1 & \mathbf{L}_2 \\ \mathbf{L}_3 & \mathbf{L}_4 \end{pmatrix}$$

$$\begin{aligned} \text{where } \mathbf{L}_1 &= n_1 n_2^4 \mathbf{J}_{n_1} \otimes (\mathbf{T}\Sigma_1 \mathbf{T})^2 + n_1^2 n_2^3 \mathbf{J}_{n_1} \otimes (\mathbf{T}\Sigma_2 \mathbf{T}\Sigma_1 \mathbf{T}) \\ \mathbf{L}_2 &= -n_1^2 n_2^3 \mathbf{J}_{n_1 \times n_2} \otimes (\mathbf{T}\Sigma_1 \mathbf{T}\Sigma_2 \mathbf{T}) - n_1^3 n_2^2 \mathbf{J}_{n_1 \times n_2} \otimes (\mathbf{T}\Sigma_2 \mathbf{T})^2 \\ \mathbf{L}_3 &= -n_1^2 n_2^3 \mathbf{J}_{n_2 \times n_1} \otimes (\mathbf{T}\Sigma_1 \mathbf{T})^2 - n_1^3 n_2^2 \mathbf{J}_{n_2 \times n_1} \otimes (\mathbf{T}\Sigma_2 \mathbf{T}\Sigma_1 \mathbf{T}) \\ \mathbf{L}_4 &= n_1^3 n_2^2 \mathbf{J}_{n_2} \otimes (\mathbf{T}\Sigma_1 \mathbf{T}\Sigma_2 \mathbf{T}) + n_1^4 n_2 \mathbf{J}_{n_2} \otimes (\mathbf{T}\Sigma_2 \mathbf{T})^2. \end{aligned}$$

so that

$$\text{tr}(\mathbf{W}\mathbf{\Xi}) = n_1 n_2 [n_2 \text{tr}(\mathbf{T}\mathbf{\Sigma}_1 \mathbf{T}) + n_1 \text{tr}(\mathbf{T}\mathbf{\Sigma}_2 \mathbf{T})] \quad (3.28)$$

$$\text{tr}(\mathbf{W}\mathbf{\Xi})^2 = n_1^2 n_2^2 \text{tr} [n_2 (\mathbf{T}\mathbf{\Sigma}_1 \mathbf{T}) + n_1 (\mathbf{T}\mathbf{\Sigma}_2 \mathbf{T})]^2. \quad (3.29)$$

Substituting these in Equations (A.8) and (A.9) give Equations (3.20) and (3.22).

Equations (3.18) and (3.19) also come immediately from Equations (A.8) and (A.19) and by noting the independence of the quadratic and bilinear forms involved since $k \neq r, l \neq s$.

Finally, for Equation (3.23), we need $\text{tr}(\mathbf{U}\mathbf{\Xi}\mathbf{W}\mathbf{\Xi})$, where it can be easily verified that $\text{tr}(\mathbf{U}\mathbf{\Xi}\mathbf{W}\mathbf{\Xi}) = \frac{1}{n_1 n_2} \text{tr}(\mathbf{W}\mathbf{\Xi})^2$. Hence,

$$\text{tr}(\mathbf{U}\mathbf{\Xi}\mathbf{W}\mathbf{\Xi}) = n_1 n_2 \text{tr} [n_2 (\mathbf{T}\mathbf{\Sigma}_1 \mathbf{T}) + n_1 (\mathbf{T}\mathbf{\Sigma}_2 \mathbf{T})]^2. \quad (3.30)$$

Substitution of this result gives Equation (3.23) which completes the proof of the theorem. $\blacksquare$

Now, we partition the general result into cases I and II, as described above, and derive the test statistic for each case separately.

3.2.3 Case I: $n_1 \neq n_2, \mathbf{\Sigma}_1 = \mathbf{\Sigma}_2$

Here we assume that the covariance matrices are equal while the sample sizes may be unequal.

3.2.3.1 The Estimators

First, the properties of the estimators are evaluated in the following theorem.

Theorem 3.4 *For $\mathbf{\Sigma}_1 = \mathbf{\Sigma}_2 = \mathbf{\Sigma}$, the estimators, $\tilde{B}_0^{AB}, \tilde{B}_1^{AB}$ and $\tilde{B}_2^{AB}$, as defined in Equations (3.12)-(3.14), are unbiased and consistent estimators of $2\text{tr}(\mathbf{T}\mathbf{\Sigma}\mathbf{T})$, $4[\text{tr}(\mathbf{T}\mathbf{\Sigma}\mathbf{T})]^2$ and $4\text{tr}(\mathbf{T}\mathbf{\Sigma}\mathbf{T})^2$, respectively. Further, the variances of the ratios of the estimators to the traces that they estimate are uniformly bounded with respect to d .* $\square$

Proof Substituting $\mathbf{\Sigma}_1 = \mathbf{\Sigma}_2 = \mathbf{\Sigma}$ in Equations (3.25) and (3.26), we have

$$\begin{aligned} \text{tr}(\mathbf{U}\mathbf{\Xi}) &= 2n_1 n_2 \text{tr}(\mathbf{T}\mathbf{\Sigma}\mathbf{T}) \\ \text{tr}(\mathbf{U}\mathbf{\Xi})^2 &= n_1 n_2 (N+2) \text{tr}(\mathbf{T}\mathbf{\Sigma}\mathbf{T})^2 \end{aligned}$$

where $N = n_1 + n_2$. Then from Equations (3.17) and (3.21) we get

$$\begin{aligned} E(\tilde{B}_0^{AB}) &= 2\text{tr}(\mathbf{T}\Sigma\mathbf{T}) \\ \text{Var}(\tilde{B}_0^{AB}) &= \frac{2(N+2)}{n_1n_2}\text{tr}(\mathbf{T}\Sigma\mathbf{T})^2 \end{aligned}$$

which prove, respectively, the unbiasedness and consistency of $\tilde{B}_0^{AB}$. Further, the dimensional stability is proved by Lemma B.11, since

$$\begin{aligned} \text{Var}\left(\frac{\tilde{B}_0^{AB}}{2\text{tr}(\mathbf{T}\Sigma\mathbf{T})}\right) &= \frac{N+2}{2n_1n_2} \frac{\text{tr}(\mathbf{T}\Sigma\mathbf{T})^2}{[\text{tr}(\mathbf{T}\Sigma\mathbf{T})]^2} \\ &= O\left(\frac{1}{n_1} + \frac{1}{n_2}\right). \end{aligned}$$

The unbiasedness of $\tilde{B}_1^{AB}$ and $\tilde{B}_2^{AB}$ follows directly from Equations (3.18) and (3.19) by taking $\Sigma_1 = \Sigma_2$. For consistency of $\tilde{B}_1^{AB}$, we write

$$\begin{aligned} \sum_{k=1}^{n_1} \sum_{l=1}^{n_2} \sum_{r=1}^{n_1} \sum_{s=1}^{n_2} A_{kl}^{AB} A_{rs}^{AB} &= \sum_{k=1}^{n_1} \sum_{l=1}^{n_2} \sum_{r=1}^{n_1} \sum_{s=1}^{n_2} A_{kl}^{AB} A_{rs}^{AB} - \underbrace{\sum_{k=1}^{n_1} \sum_{l=1}^{n_2} \sum_{s=1}^{n_2} A_{kl}^{AB} A_{ks}^{AB}}_{k=r} \\ &\quad - \underbrace{\sum_{k=1}^{n_1} \sum_{l=1}^{n_2} \sum_{r=1}^{n_1} A_{kl}^{AB} A_{rl}^{AB}}_{l=s} + \underbrace{\sum_{k=1}^{n_1} \sum_{l=1}^{n_2} A_{kl}^{AB} A_{kl}^{AB}}_{k=r, l=s} \\ &\approx \left(\sum_{k=1}^{n_1} \sum_{l=1}^{n_2} A_{kl}^{AB} \right)^2, \end{aligned}$$

where $\approx$ means 'approximately equal to'. Then

$$\frac{1}{n_1^2 n_2^2} \underbrace{\sum_{k=1}^{n_1} \sum_{l=1}^{n_2} \sum_{r=1}^{n_1} \sum_{s=1}^{n_2} A_{kl}^{AB} A_{rs}^{AB}}_{k \neq r, l \neq s} \approx \left(\frac{1}{n_1 n_2} \sum_{k=1}^{n_1} \sum_{l=1}^{n_2} A_{kl}^{AB} \right)^2 = \left(\tilde{B}_0^{AB} \right)^2,$$

using Equation (3.12). Since $\tilde{B}_0^{AB}$ is consistent, hence the identity on the left

hand side of the above expression is also consistent, and since

$$\tilde{B}_1^{AB} = \frac{n_1 n_2}{(n_1 - 1)(n_2 - 1)} \left[\frac{1}{n_1^2 n_2^2} \underbrace{\sum_{k=1}^{n_1} \sum_{l=1}^{n_2} \sum_{r=1}^{n_1} \sum_{s=1}^{n_2}}_{k \neq r, l \neq s} A_{kl}^{AB} A_{rs}^{AB} \right],$$

therefore, $\tilde{B}_1^{AB}$ is also consistent. Now, by Theorem B.5,

$$\begin{aligned} \text{Var}(\tilde{B}_1^{AB}) &= \frac{1}{n_1^2 n_2^2 (n_1 - 1)^2 (n_2 - 1)^2} \\ &\quad \sum_{\substack{k=1 \\ k \neq r, l \neq s}}^{n_1} \sum_{l=1}^{n_2} \sum_{r=1}^{n_1} \sum_{s=1}^{n_2} \sum_{\substack{k'=1 \\ k' \neq r', l' \neq s'}}^{n_1} \sum_{l'=1}^{n_2} \sum_{r'=1}^{n_1} \sum_{s'=1}^{n_2} \text{Cov}(A_{kl}^{AB} A_{rs}^{AB}, A_{k'l'}^{AB} A_{r's'}^{AB}) \\ &\ll 4 \left([\text{tr}(\mathbf{T}\Sigma\mathbf{T})^2]^2 + \text{tr}(\mathbf{T}\Sigma\mathbf{T})^2 [\text{tr}(\mathbf{T}\Sigma\mathbf{T})]^2 \right) (1/n_1 + 1/n_2), \end{aligned}$$

so that

$$\begin{aligned} \text{Var}\left(\frac{\tilde{B}_1^{AB}}{4 [\text{tr}(\mathbf{T}\Sigma\mathbf{T})]^2}\right) &\ll \frac{1}{4} \left(\frac{[\text{tr}(\mathbf{T}\Sigma\mathbf{T})^2]^2}{[\text{tr}(\mathbf{T}\Sigma\mathbf{T})]^4} + \frac{\text{tr}(\mathbf{T}\Sigma\mathbf{T})^2}{[\text{tr}(\mathbf{T}\Sigma\mathbf{T})]^2} \right) (N/n_1 n_2) \\ &\ll \frac{N}{n_1 n_2} \end{aligned}$$

by Lemma B.11, which proves the dimensional stability of $\tilde{B}_1^{AB}$, as $N/n_1 n_2$ is independent of d . Now, for consistency of $\tilde{B}_2^{AB}$, we write

$$\begin{aligned} \text{Var}(\tilde{B}_2^{AB}) &= \frac{1}{n_1^2 n_2^2 (n_1 - 1)^2 (n_2 - 1)^2} \\ &\quad \sum_{\substack{k=1 \\ k \neq r, l \neq s}}^{n_1} \sum_{l=1}^{n_2} \sum_{r=1}^{n_1} \sum_{s=1}^{n_2} \sum_{\substack{k'=1 \\ k' \neq r', l' \neq s'}}^{n_1} \sum_{l'=1}^{n_2} \sum_{r'=1}^{n_1} \sum_{s'=1}^{n_2} \text{Cov}(A_{klrs}^{AB^2}, A_{k'l'r's'}^{AB^2}). \quad (3.31) \end{aligned}$$

There are four possible pairs, (k, k') , (l, l') , (r, r') and (s, s') , with 8 positions, to be permuted to cover all the covariance terms, including variances. When all pairs are unequal, i.e., $k \neq k', l \neq l', r \neq r', s \neq s'$, then the covariance is zero, by independence, and the number of such covariances is of order $O(n_1^4 n_2^4)$, same as that of the denominator, on the right side of Equation (5.26) above. For all other cases, we have a non-zero covariance.

Now, suppose $\text{Cov}(A_{klrs}^{AB^2}, A_{k'l'r's'}^{AB^2}) \neq 0$, so that $\exists$ at least one pair where

indices are equal. The number of terms of all such covariances is of order no more than $O(n_1^4 n_2^3)$ or $O(n_1^3 n_2^4)$, so that the order of the expression on the right side of Equation (5.26) is $O(1/n_1)$ if $n_2 \geq n_1$ or $O(1/n_2)$ if $n_1 \geq n_2$, which proves the consistency of $\tilde{B}_2^{AB}$.

Finally, Theorem B.5 gives the dimensional stability, since

$$\begin{aligned} \text{Var}(\tilde{B}_2^{AB}) &= \frac{1}{n_1^2 n_2^2 (n_1 - 1)^2 (n_2 - 1)^2} \\ &\quad \sum_{\substack{k=1 \\ k \neq r}}^{n_1} \sum_{\substack{l=1 \\ l \neq s}}^{n_2} \sum_{r=1}^{n_1} \sum_{s=1}^{n_2} \sum_{\substack{k'=1 \\ k' \neq r'}}^{n_1} \sum_{\substack{l'=1 \\ l' \neq s'}}^{n_2} \sum_{r'=1}^{n_1} \sum_{s'=1}^{n_2} \text{Cov}(A_{klrs}^{AB^2}, A_{k'l'r's'}^{AB^2}) \\ &\ll 2 \left(3\text{tr}(\mathbf{T}\Sigma\mathbf{T})^4 + [\text{tr}(\mathbf{T}\Sigma\mathbf{T})^2]^2 \right) O(1/n_1 + 1/n_2) \end{aligned}$$

so that

$$\begin{aligned} \text{Var}\left(\frac{\tilde{B}_2^{AB}}{4\text{tr}(\mathbf{T}\Sigma\mathbf{T})^2}\right) &\ll \frac{1}{8} \left(\frac{3\text{tr}(\mathbf{T}\Sigma\mathbf{T})^4}{[\text{tr}(\mathbf{T}\Sigma\mathbf{T})^2]^2} + 1 \right) (N/n_1 n_2) \\ &\ll \frac{N}{n_1 n_2}. \end{aligned}$$

This completes the proof of the theorem. ■

3.2.3.2 The Approximating Distribution

Following lemma is needed to compute the moments of the statistic.

Lemma 3.5 Let Q_N^{AB} and $\tilde{B}_0^{AB}$ be as defined in Equations (3.15) and (3.16), respectively. Then, with $\Sigma_1 = \Sigma_2$, we have

$$\text{E}(Q_N^{AB}) = 2\text{tr}(\mathbf{T}\Sigma\mathbf{T}) \quad (3.32)$$

$$\text{Var}(Q_N^{AB}) = 8\text{tr}(\mathbf{T}\Sigma\mathbf{T})^2 \quad (3.33)$$

$$\text{Cov}(Q_N^{AB}, \tilde{B}_0^{AB}) = \frac{4N}{n_1 n_2} \text{tr}(\mathbf{T}\Sigma\mathbf{T})^2. \quad (3.34)$$

□

Proof For $\Sigma_1 = \Sigma_2$, Equations (3.28), (3.29) and (5.25) reduce to

$$\begin{aligned} \text{tr}(\mathbf{W}\Xi) &= n_1 n_2 N \text{tr}(\mathbf{T}(\Sigma\mathbf{T})) \\ \text{tr}(\mathbf{W}\Xi)^2 &= n_1^2 n_2^2 N^2 \text{tr}(\mathbf{T}(\Sigma\mathbf{T})^2) \\ \text{tr}(\mathbf{U}\Xi\mathbf{W}\Xi) &= n_1 n_2 N^2 \text{tr}(\mathbf{T}(\Sigma\mathbf{T})^2) \end{aligned}$$

where $N = n_1 + n_2$. Further, substitution of $\Sigma_1 = \Sigma_2$ in Equations (3.20), (3.22) and (3.23) gives the required result. ■

Now, define the statistic as $\tilde{A}_N = Q^{AB}/\tilde{B}_0^{AB}$, where $\tilde{A}_N \sim \chi_{f^{AB}}^2/f^{AB}$. Equations (B.1) and (B.2) give the following moments of the statistic, where the order of the remainder of the mean of the statistic is at least $O(1/n_1 + 1/n_2)$ and that for the variance is at least $o(1/n_1 + 1/n_2)$ (Stuart and Ord, 1994, p 350). We compute the moments of the statistic ignoring these remainders for approximation purposes (Casella and Berger, 2002, p 245).

Theorem 3.6 *For $\Sigma_1 = \Sigma_2$, the first two moments of $\tilde{A}_N^{AB}$ are given as following, where $N = n_1 + n_2$.*

$$\begin{aligned} E(\tilde{A}_N^{AB}) &\approx 1 \\ \text{Var}(\tilde{A}_N^{AB}) &\approx \frac{2}{f^{AB}} \left(1 - \frac{3N-2}{4n_1n_2}\right). \end{aligned} \quad \square$$

Hence, we conclude that $\tilde{F}_N^{AB} = \tilde{f}^{AB} \tilde{A}_N^{AB} \sim \chi_{f^{AB}}^2$, asymptotically.

3.2.4 Case II: $n_1 = n_2$, $\Sigma_1 \neq \Sigma_2$

Now we assume that the sample sizes are equal while the covariance matrices may be unequal.

3.2.4.1 The Estimators

Following theorem gives the properties of the estimators.

Theorem 3.7 *For $n_1 = n_2 = n$, the estimators, $\tilde{B}_0^{AB}$, $\tilde{B}_1^{AB}$ and $\tilde{B}_2^{AB}$, as defined in Equations (3.12)-(3.14), are unbiased and consistent estimators of $\text{tr}[\mathbf{T}(\Sigma_1 + \Sigma_2)\mathbf{T}]$, $[\text{tr}\{\mathbf{T}(\Sigma_1 + \Sigma_2)\mathbf{T}\}]^2$ and $\text{tr}[\mathbf{T}(\Sigma_1 + \Sigma_2)\mathbf{T}]^2$, respectively. Further, the variances of the ratios of the estimators to the traces that they estimate are uniformly bounded with respect to d .* □

Proof For $n_1 = n_2$, Equation (3.16) reduces to

$$\tilde{B}_0^{AB} = \frac{1}{n^2} \mathbf{Y}' \mathbf{M}' \mathbf{M} \mathbf{Y}. \quad (3.35)$$

Similarly, from Equations (3.25) and (3.26), we have

$$\begin{aligned}\text{tr}(\mathbf{U}\Xi) &= n^2 \text{tr}[\mathbf{T}(\Sigma_1 + \Sigma_2)\mathbf{T}] \\ \text{tr}(\mathbf{U}\Xi)^2 &= n^3 [\text{tr}(\mathbf{T}\Sigma_1\mathbf{T})^2 + \frac{2}{n} \text{tr}(\mathbf{T}\Sigma_1\mathbf{T}\Sigma_2\mathbf{T}) + \text{tr}(\mathbf{T}\Sigma_2\mathbf{T})^2]\end{aligned}$$

and Equations (3.17) and (3.21) give

$$\begin{aligned}\mathbb{E}(\tilde{B}_0^{AB}) &= \text{tr}[\mathbf{T}(\Sigma_1 + \Sigma_2)\mathbf{T}] \\ \text{Var}(\tilde{B}_0^{AB}) &= \frac{2}{n} [\text{tr}(\mathbf{T}\Sigma_1\mathbf{T})^2 + \frac{2}{n} \text{tr}(\mathbf{T}\Sigma_1\mathbf{T}\Sigma_2\mathbf{T}) + \text{tr}(\mathbf{T}\Sigma_2\mathbf{T})^2]\end{aligned}$$

which prove the unbiasedness and consistency of $\tilde{B}_0^{AB}$, respectively. Further

$$\begin{aligned}\text{Var}\left(\frac{\tilde{B}_0^{AB}}{\text{tr}[\mathbf{T}(\Sigma_1 + \Sigma_2)\mathbf{T}]}\right) &= \frac{2}{n} \left(\frac{[\text{tr}(\mathbf{T}\Sigma_1\mathbf{T})^2 + \frac{2}{n} \text{tr}(\mathbf{T}\Sigma_1\mathbf{T}\Sigma_2\mathbf{T}) + \text{tr}(\mathbf{T}\Sigma_2\mathbf{T})^2]}{[\text{tr}[\mathbf{T}(\Sigma_1 + \Sigma_2)\mathbf{T}]]^2} \right) \\ &\leq \frac{2}{n} \left(2 + \frac{2}{n} \right) = O\left(\frac{1}{n}\right)\end{aligned}$$

by Lemma B.11, which proves the dimensional stability of $\tilde{B}_0^{AB}$.

The unbiasedness of $\tilde{B}_1^{AB}$ and $\tilde{B}_2^{AB}$ follows directly from Equations (3.18) and (3.19) by taking $n_1 = n_2$. For consistency of $\tilde{B}_1^{AB}$, we proceed as in Case I, so that

$$\frac{1}{n^4} \underbrace{\sum_{k=1}^n \sum_{l=1}^n \sum_{r=1}^n \sum_{s=1}^n}_{k \neq r, l \neq s} A_{kl}^{AB} A_{rs}^{AB} \approx \left(\tilde{B}_0^{AB}\right)^2,$$

using Equation (3.12), where $\approx$ means 'approximately equal to'. It proves that $\tilde{B}_1^{AB}$ is also consistent, since

$$\tilde{B}_1^{AB} = \frac{n^4}{n^2(n-1)^2} \left[\frac{1}{n^4} \sum_{k=1}^n \sum_{l=1}^n \sum_{r=1}^n \sum_{s=1}^n \underbrace{A_{kl}^{AB} A_{rs}^{AB}}_{k \neq r, l \neq s} \right].$$

For uniform boundedness in d , we write

$$\text{Var}\left(\tilde{B}_1^{AB}\right) = \frac{1}{n^4(n-1)^4} \sum_{k=1}^n \sum_{l=1}^n \sum_{r=1}^n \sum_{s=1}^n \underbrace{\sum_{k'=1}^n \sum_{l'=1}^n \sum_{r'=1}^n \sum_{s'=1}^n}_{k' \neq r', l' \neq s'} \text{Cov}\left(A_{kl}^{AB} A_{rs}^{AB}, A_{k'l'}^{AB} A_{r's'}^{AB}\right),$$

and by Cauchy-Schwarz inequality (Theorem B.5), we have

$$\begin{aligned} \text{Cov} (A_{kl}^{AB} A_{rs}^{AB}, A_{k'l'}^{AB} A_{r's'}^{AB}) &\leq \sqrt{\text{Var} (A_{kl}^{AB} A_{rs}^{AB}) \text{Var} (A_{k'l'}^{AB} A_{r's'}^{AB})} \\ &\leq 4 [\text{tr} \{ \mathbf{T} (\boldsymbol{\Sigma}_1 + \boldsymbol{\Sigma}_2) \mathbf{T} \}^2]^2 \\ &\quad + 4 \text{tr} \{ \mathbf{T} (\boldsymbol{\Sigma}_1 + \boldsymbol{\Sigma}_2) \mathbf{T} \}^2 [\text{tr} \{ \mathbf{T} (\boldsymbol{\Sigma}_1 + \boldsymbol{\Sigma}_2) \mathbf{T} \}]^2 \end{aligned}$$

using Equations (A.14). Hence,

$$\begin{aligned} \text{Var} \left(\frac{\tilde{B}_1^{AB}}{[\text{tr} \{ \mathbf{T} (\boldsymbol{\Sigma}_1 + \boldsymbol{\Sigma}_2) \mathbf{T} \}]^2} \right) &\ll 4 \left(\frac{[\text{tr} \{ \mathbf{T} (\boldsymbol{\Sigma}_1 + \boldsymbol{\Sigma}_2) \mathbf{T} \}^2]^2}{[\text{tr} \{ \mathbf{T} (\boldsymbol{\Sigma}_1 + \boldsymbol{\Sigma}_2) \mathbf{T} \}]^4} \right. \\ &\quad \left. + \frac{\text{tr} \{ \mathbf{T} (\boldsymbol{\Sigma}_1 + \boldsymbol{\Sigma}_2) \mathbf{T} \}^2}{[\text{tr} \{ \mathbf{T} (\boldsymbol{\Sigma}_1 + \boldsymbol{\Sigma}_2) \mathbf{T} \}]^2} \right) (1/n) \\ &\ll 1/n \end{aligned}$$

by Lemma B.11, where the left-over expression on the right side of the last inequality is independent of d . The consistency of $\tilde{B}_2^{AB}$ can be proved using the same arguments as in Case I, where, for $n_1 = n_2$, the order of convergence is $O(1/n^5)$. For dimensional stability, write

$$\begin{aligned} \text{Cov} (A_{klrs}^{AB^2}, A_{k'l'r's'}^{AB^2}) &\leq \sqrt{\text{Var} (A_{klrs}^{AB^2}) \text{Var} (A_{k'l'r's'}^{AB^2})} \\ &\leq 6 \text{tr} \{ \mathbf{T} (\boldsymbol{\Sigma}_1 + \boldsymbol{\Sigma}_2) \mathbf{T} \}^4 + 2 [\text{tr} \{ \mathbf{T} (\boldsymbol{\Sigma}_1 + \boldsymbol{\Sigma}_2) \mathbf{T} \}^2]^2 \end{aligned}$$

by Theorem B.5, so that,

$$\begin{aligned} \text{Var} \left(\frac{\tilde{B}_2^{AB}}{\text{tr} \{ \mathbf{T} (\boldsymbol{\Sigma}_1 + \boldsymbol{\Sigma}_2) \mathbf{T} \}^2} \right) &\ll \left(\frac{6 \text{tr} \{ \mathbf{T} (\boldsymbol{\Sigma}_1 + \boldsymbol{\Sigma}_2) \mathbf{T} \}^4}{[\text{tr} \{ \mathbf{T} (\boldsymbol{\Sigma}_1 + \boldsymbol{\Sigma}_2) \mathbf{T} \}^2]^2} + 2 \right) (1/n) \\ &\ll 1/n \end{aligned}$$

which completes the proof. ■

3.2.4.2 The Approximating Distribution

We need the following lemma for the computation of the moments of the test statistic, $\tilde{A}_N^{AB}$.

Lemma 3.8 For Q_N^{AB} and $\tilde{B}_0^{AB}$, as defined in Equations (3.15) and (3.16), respectively, we have, with $n_1 = n_2$,

$$E(Q_N^{AB}) = \text{tr}[\mathbf{T}(\boldsymbol{\Sigma}_1 + \boldsymbol{\Sigma}_2)\mathbf{T}] \quad (3.36)$$

$$\text{Var}(Q_N^{AB}) = 2\text{tr}[\mathbf{T}(\boldsymbol{\Sigma}_1 + \boldsymbol{\Sigma}_2)\mathbf{T}]^2 \quad (3.37)$$

$$\text{Cov}(Q_N^{AB}, \tilde{B}_0^{AB}) = \frac{2}{n}\text{tr}[\mathbf{T}(\boldsymbol{\Sigma}_1 + \boldsymbol{\Sigma}_2)\mathbf{T}]^2. \quad (3.38)$$

□

Proof For $n_1 = n_2$, we have $Q_N^{AB} = \frac{1}{n^3}\mathbf{Y}'\mathbf{M}'(\mathbf{J}_{n^2} \otimes \mathbf{I}_d)\mathbf{M}\mathbf{Y}$ and from Equations (3.28), (3.29) and (5.25), we get

$$\begin{aligned} \text{tr}(\mathbf{W}\boldsymbol{\Xi}) &= n^3\text{tr}[\mathbf{T}(\boldsymbol{\Sigma}_1 + \boldsymbol{\Sigma}_2)\mathbf{T}] \\ \text{tr}(\mathbf{W}\boldsymbol{\Xi})^2 &= n^6\text{tr}[\mathbf{T}(\boldsymbol{\Sigma}_1 + \boldsymbol{\Sigma}_2)\mathbf{T}]^2 \\ \text{tr}(\mathbf{U}\boldsymbol{\Xi}\mathbf{W}\boldsymbol{\Xi}) &= n^4\text{tr}[\mathbf{T}(\boldsymbol{\Sigma}_1 + \boldsymbol{\Sigma}_2)\mathbf{T}]^2. \end{aligned}$$

Finally, Equations (3.20), (3.22) and (3.23) give the required result. ■

Now, it is just a matter of substituting the relevant identities into Equations (B.1) and (B.2) and we get the following moments of $\tilde{A}_N^{AB}$. Note that, in Theorem 3.9, the order of the remainder for the mean of the statistic is at least $O(1/n)$ and for the variance of the statistic is at least $o(1/n)$ (Stuart and Ord, 1994, p 350). We compute the moments ignoring these remainders for approximation purposes (Casella and Berger, 2002, p 245).

Theorem 3.9 For $n_1 = n_2$, the first two moments of the test statistic, $\tilde{A}_N^{AB}$, are given as following.

$$\begin{aligned} E(\tilde{A}_N^{AB}) &\approx 1 \\ \text{Var}(\tilde{A}_N^{AB}) &\approx \frac{2}{\tilde{f}^{AB}} \left(1 + \frac{\frac{1}{n} [\text{tr}(\mathbf{T}\boldsymbol{\Sigma}_1\mathbf{T})^2 + \frac{2}{n}\text{tr}(\mathbf{T}\boldsymbol{\Sigma}_1\mathbf{T}\boldsymbol{\Sigma}_2\mathbf{T}) + \text{tr}(\mathbf{T}\boldsymbol{\Sigma}_2\mathbf{T})^2]}{\text{tr}[\mathbf{T}(\boldsymbol{\Sigma}_1 + \boldsymbol{\Sigma}_2)\mathbf{T}]^2} - \frac{2}{n} \right). \end{aligned}$$

□

The moments of $\tilde{A}_N^{AB}$ closely approximate the moments of χ_f^2/f -distribution, which are 1 and $2/f$, particularly when $n \rightarrow \infty$. Like in one sample case, we can write $\tilde{F}_N^{AB} = \tilde{f}^{AB}\tilde{A}_N^{AB}$, so that $\tilde{F}_N^{AB} \sim \chi_{f^{AB}}^2$.

The middle term in $\text{Var}(\tilde{A}_N^{AB})$, consisting of traces, is exactly 1 for $n = 1$ and gets decreasing for increasing n . We need to estimate these traces. Since,

$$\text{tr}[\mathbf{T}(\boldsymbol{\Sigma}_1 + \boldsymbol{\Sigma}_2)\mathbf{T}]^2 = \text{tr}(\mathbf{T}\boldsymbol{\Sigma}_1\mathbf{T})^2 + 2\text{tr}(\mathbf{T}\boldsymbol{\Sigma}_1\mathbf{T}\boldsymbol{\Sigma}_2\mathbf{T}) + \text{tr}(\mathbf{T}\boldsymbol{\Sigma}_2\mathbf{T})^2,$$

therefore,

$$\text{Var}(\tilde{A}_N^{AB}) = \frac{2}{f^{AB}} \left(1 - \frac{1}{n}\right) \left(1 - \frac{2\eta}{n}\right)$$

where $\eta = \frac{\text{tr}(\mathbf{T}\boldsymbol{\Sigma}_1\mathbf{T}\boldsymbol{\Sigma}_2\mathbf{T})}{\text{tr}[\mathbf{T}(\boldsymbol{\Sigma}_1 + \boldsymbol{\Sigma}_2)\mathbf{T}]^2}$. Note that $\eta = \frac{1}{4}$ when $\boldsymbol{\Sigma}_1 = \boldsymbol{\Sigma}_2$. Otherwise, we expect η vanishing since the numerator is a part of the denominator. We plug in the classical estimator, $\hat{\boldsymbol{\Sigma}}_i = \sum_{k=1}^n (\mathbf{X}_{ik} - \bar{\mathbf{X}}_i)(\mathbf{X}_{ik} - \bar{\mathbf{X}}_i)' / (n_i - 1)$, $i = 1, 2$, so that $\hat{\eta} = \frac{\text{tr}(\mathbf{T}\hat{\boldsymbol{\Sigma}}_1\mathbf{T}\hat{\boldsymbol{\Sigma}}_2\mathbf{T})}{\text{tr}[\mathbf{T}(\hat{\boldsymbol{\Sigma}}_1 + \hat{\boldsymbol{\Sigma}}_2)\mathbf{T}]^2}$ replaces η in $\text{Var}(\tilde{F}^{AB})$. It will be shown in the simulation study in the next section that using $\hat{\boldsymbol{\Sigma}}_i$, although leaves $\hat{\eta}$ biased, does not jeopardize the approximation which is still accurate for moderate n and any d .

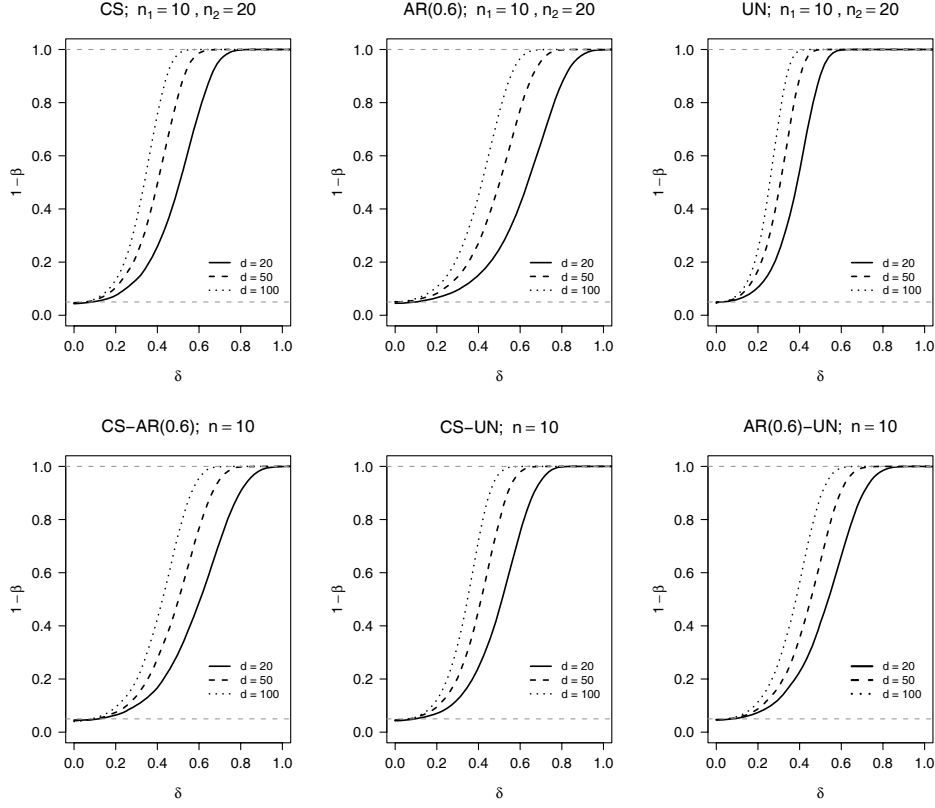
3.2.5 Simulation Results

3.2.5.1 Level

Table 3.2: ESTIMATED QUANTILES OF $\tilde{A}_N^{AB}$: *Normal Distribution*

n_1	n_2	d	0.90	0.95	0.99	0.90	0.95	0.99	0.90	0.95	0.99
$(\boldsymbol{\Sigma}_1 = \boldsymbol{\Sigma}_2)$			CS			AR(0.6)			UN		
10	20	20	0.9045	0.9526	0.9923	0.9005	0.9504	0.9918	0.9006	0.9543	0.9925
		50	0.9042	0.9526	0.9913	0.8974	0.9445	0.9888	0.9004	0.9507	0.9923
		100	0.9029	0.9631	0.9938	0.9029	0.9517	0.9909	0.9085	0.9581	0.9937
		200	0.8992	0.9503	0.9918	0.9046	0.9531	0.9912	0.9069	0.9511	0.9913
20	30	30	0.9041	0.9502	0.9911	0.9084	0.9515	0.9868	0.9035	0.9502	0.9914
		50	0.8981	0.9484	0.9912	0.9027	0.9494	0.9883	0.9058	0.9530	0.9917
		100	0.9007	0.9509	0.9906	0.9022	0.9491	0.9892	0.9026	0.9482	0.9904
		200	0.8986	0.9513	0.9900	0.9022	0.9489	0.9882	0.9016	0.9473	0.9914
$(n_1 = n_2)$			CS-AR(0.6)			CS-UN			AR(0.6)-UN		
10	10	20	0.9069	0.9509	0.9923	0.9069	0.9495	0.9877	0.8996	0.9501	0.9928
		50	0.9033	0.9432	0.9848	0.8999	0.9505	0.9941	0.9004	0.9504	0.9928
		100	0.9010	0.9423	0.9817	0.9003	0.9508	0.9912	0.9026	0.9536	0.9927
		200	0.9007	0.9486	0.9896	0.8925	0.9501	0.9909	0.9067	0.9576	0.9933
20	20	30	0.9104	0.9490	0.9857	0.9058	0.9503	0.9902	0.9063	0.9521	0.9893
		50	0.9044	0.9490	0.9914	0.9050	0.9498	0.9908	0.9079	0.9535	0.9898
		100	0.9066	0.9505	0.9890	0.9082	0.9501	0.9892	0.9020	0.9513	0.9916
		200	0.9065	0.9510	0.9891	0.9001	0.9504	0.9907	0.9006	0.9465	0.9893

Table 3.2 reports the results for the estimated quantiles for the ANOVA-type statistic for the interaction effect. The upper panel of the table refers

Figure 3.1: POWER CURVES FOR $\tilde{A}_N^{AB}$: *Normal Distribution*

to Case I ($\Sigma_1 = \Sigma_2$) whereas the lower panel refers to Case II ($n_1 = n_2$). For Case II, we evaluate the statistic for all pairs of the three covariance structures, CS, AR(0.6) and UN (Definitions B.6, B.7 and B.8). These results are for 10,000 simulation runs with $n \in \{10, 20\}$ and $d \in \{20, 30, 50, 100, 200\}$.

We observe a very accurate approximation of the ATS for all three covariance structures. Further, the stability of the estimated quantiles is well maintained with increasing dimension, d . Particularly noticeable are the results for Case II, the so-called Behrens-Fisher case, where the accuracy is evident even for $n = 10$, for all pairs of covariance matrices. Note that, like for one sample case, we also investigated the behavior of two-sample statistic for other covariance structures and found similar results. Likewise, for $\Sigma_1 \neq \Sigma_2$, we also got similar results for other pairs, for example AR(0.2)-AR(0.6).

3.2.5.2 Power

The power curves for $\tilde{A}_N^{AB}$ are shown in Fig. 5.1 with upper and lower panels referring, respectively, to Cases I and II, as in Table 3.2. We observe a similar performance as for the one sample case, particularly, that the power increases for increasing d . The most important feature of Fig. 5.1 is that the power curves for the two cases are similar, further substantiating the fact that the statistic does not depend on any specific covariance structure. Note that, for these power curves, the alternative is defined as in Equation (2.17) except that now $\delta = 0(0.02)1$.

A point worth mentioning is the behavior of the statistic under the two cases, viz. $\Sigma_1 = \Sigma_2$ and $n_1 = n_2$. Although, theoretically, we have to partition the general case into these two cases, but, as the simulation results for the level and power indicate, the statistic can, by all practical means, be used without distinguishing it into the two sub-cases. This feature is further substantiated by the numerical illustrations, since in all applications, we observed virtually the same result in both cases and, therefore, only the results for Case II ($\Sigma_1 \neq \Sigma_2, n_1 = n_2$) are presented (see Section 3.5).

3.3 The Time Effect

The hypothesis for the main effect of time is $H_0^B : (\frac{1}{2}\mathbf{J}_2 \otimes \mathbf{P}_d) \boldsymbol{\mu} = \mathbf{0}$ which can also be written as $H_0^B : \mathbf{P}_d(\boldsymbol{\mu}_1 + \boldsymbol{\mu}_2) = \mathbf{0}$. The statistical model, data set up and assumptions are as defined in Section 3.1. As for the interaction effect, the generating matrix for the time effect is $\mathbf{T} = \mathbf{P}_d$, although for convenience, the general form of the hypothesis matrix, $\mathbf{T}$, will be used.

As mentioned in Section 3.1.3, the final results for all expressions involved in the computations of the time effect are the same as those for the interaction effect, since the covariance matrix for the differences $\mathbf{Y}_{1k} - \mathbf{Y}_{2l}$ and the sums $\mathbf{Y}_{1k} + \mathbf{Y}_{2l}$ is same (see Sections A.1 and A.2). Therefore, in this section, only the main results will be summarized, the proofs of which are similar to those given in Section 3.2.

3.3.1 The Covariance Matrix

Let $\mathbf{X}_{1k}$ and $\mathbf{X}_{2l}$, be as defined in Section 3.1. Define the sums $\mathbf{X}_{1k} + \mathbf{X}_{2l}$, $\forall k, l$, where $E(\mathbf{X}_{1k} + \mathbf{X}_{2l}) = \boldsymbol{\mu}_1 + \boldsymbol{\mu}_2$ and $\text{Var}(\mathbf{X}_{1k} + \mathbf{X}_{2l}) = \boldsymbol{\Sigma}_1 + \boldsymbol{\Sigma}_2$. Let $\mathbf{T}(\mathbf{X}_{1k} + \mathbf{X}_{2l}) = \mathbf{Y}_{1k} + \mathbf{Y}_{2l}$, with $E(\mathbf{Y}_{1k} + \mathbf{Y}_{2l}) = \mathbf{0}$ and $\text{Cov}(\mathbf{Y}_{1k} + \mathbf{Y}_{2l}) = \mathbf{T}(\boldsymbol{\Sigma}_1 + \boldsymbol{\Sigma}_2)\mathbf{T}$, under H_0^B , where $\hat{\boldsymbol{\Sigma}}_N$ can now be defined by replacing $\mathbf{Y}_{1k} - \mathbf{Y}_{2l}$ with $\mathbf{Y}_{1k} + \mathbf{Y}_{2l}$ in Equation (5.1). In matrix form, with $\mathbf{Z}$ as a vector of sums $\mathbf{Z}_{kl} = \mathbf{Y}_{1k} + \mathbf{Y}_{2l}$,

$$\hat{\boldsymbol{\Sigma}}_N = \frac{1}{n_1 n_2} \sum_{k=1}^{n_1} \sum_{l=1}^{n_2} \mathbf{Z}_{kl} \mathbf{Z}_{kl}' \quad (3.39)$$

We have the following theorem.

Theorem 3.10 *Under H_0^B , $\hat{\boldsymbol{\Sigma}}_N$, as defined in Equation (5.34), is an unbiased estimator of $\mathbf{T}(\boldsymbol{\Sigma}_1 + \boldsymbol{\Sigma}_2)\mathbf{T}$. $\square$*

3.3.2 Defining The Statistic

Define $\bar{\mathbf{Y}}_1 + \bar{\mathbf{Y}}_2 = \mathbf{T}(\bar{\mathbf{X}}_1 + \bar{\mathbf{X}}_2)$ such that, under H_0^B , $E(\bar{\mathbf{Y}}_1 + \bar{\mathbf{Y}}_2) = \mathbf{0}$ and $\text{Cov}(\bar{\mathbf{Y}}_1 + \bar{\mathbf{Y}}_2) = \mathbf{T} \left(\frac{1}{n_1} \boldsymbol{\Sigma}_1 + \frac{1}{n_2} \boldsymbol{\Sigma}_2 \right) \mathbf{T}$. Then, from Equation (4.5),

$$\tilde{A}_N^B = \frac{\frac{2n_1 n_2}{N} (\bar{\mathbf{Y}}_1 + \bar{\mathbf{Y}}_2)' (\bar{\mathbf{Y}}_1 + \bar{\mathbf{Y}}_2)}{\text{tr} [\mathbf{T} (\boldsymbol{\Sigma}_1 + \boldsymbol{\Sigma}_2) \mathbf{T}]} \quad (3.40)$$

where $\tilde{A}_N^B \sim \chi_{f^B}^2 / f^B$ with

$$f^B = \frac{[\text{tr} \{ \mathbf{T} (\boldsymbol{\Sigma}_1 + \boldsymbol{\Sigma}_2) \mathbf{T} \}]^2}{\text{tr} [\mathbf{T} (\boldsymbol{\Sigma}_1 + \boldsymbol{\Sigma}_2) \mathbf{T}]^2}. \quad (3.41)$$

The estimators, $\tilde{B}_0^B$, $\tilde{B}_1^B$ and $\tilde{B}_2^B$, from Equations (3.12)-(3.14), now take the following form.

$$\tilde{B}_0^B = \frac{1}{n_1 n_2} \sum_{k=1}^{n_1} \sum_{l=1}^{n_2} A_{kl}^B \quad (3.42)$$

$$\tilde{B}_1^B = \frac{1}{n_1 n_2 (n_1 - 1)(n_2 - 1)} \underbrace{\sum_{k=1}^{n_1} \sum_{l=1}^{n_2} \sum_{r=1}^{n_1} \sum_{s=1}^{n_2} A_{kl}^B A_{rs}^B}_{k \neq r, l \neq s} \quad (3.43)$$

$$\tilde{B}_2^B = \frac{1}{n_1 n_2 (n_1 - 1)(n_2 - 1)} \underbrace{\sum_{k=1}^{n_1} \sum_{l=1}^{n_2} \sum_{r=1}^{n_1} \sum_{s=1}^{n_2} A_{klrs}^{B^2}}_{k \neq r, l \neq s} \quad (3.44)$$

where $A_{kl}^B = (\mathbf{Y}_{1k} + \mathbf{Y}_{2l})'(\mathbf{Y}_{1k} + \mathbf{Y}_{2l})$ and $A_{rs}^B = (\mathbf{Y}_{1r} + \mathbf{Y}_{2s})'(\mathbf{Y}_{1r} + \mathbf{Y}_{2s})$ are the quadratic forms and $A_{klrs}^B = (\mathbf{Y}_{1k} + \mathbf{Y}_{2l})'(\mathbf{Y}_{1r} + \mathbf{Y}_{2s})$ is the symmetric bilinear form, as defined in Sections A.1 and A.2.

To express Q_N^B and $\tilde{B}_0^B$ in matrix form, let the random vectors $\mathbf{Y}$, $\boldsymbol{\xi}$ and $\boldsymbol{\Xi}$ be as defined in Section 3.3.1 where, now, $\mathbf{M} = (\mathbf{I}_{n_1} \otimes \mathbf{1}_{n_2} \mid \mathbf{1}_{n_1} \otimes \mathbf{I}_{n_2}) \otimes \mathbf{I}_d$ such that $\mathbf{Z} = \mathbf{M}\mathbf{Y}$ denotes the vector of sums, $\mathbf{Z}_{kl} = \mathbf{Y}_{1k} + \mathbf{Y}_{2l}$, $k = 1, \dots, n_1$ and $l = 1, \dots, n_2$, with $E(\mathbf{Z}_{kl}) = \mathbf{0}$ and $\text{Cov}(\mathbf{Z}_{kl}) = \mathbf{T}(\boldsymbol{\Sigma}_1 + \boldsymbol{\Sigma}_2)\mathbf{T}$. Then

$$Q_N^B = \frac{2}{N n_1 n_2} \mathbf{Y}' \mathbf{M}' (\mathbf{J}_{n_1} \otimes \mathbf{J}_{n_2} \otimes \mathbf{I}_d) \mathbf{M} \mathbf{Y} \quad (3.45)$$

$$\tilde{B}_0^B = \frac{1}{n_1 n_2} \mathbf{Y}' \mathbf{M}' \mathbf{M} \mathbf{Y} \quad (3.46)$$

where $N = n_1 + n_2$.

As for the interaction effect (Section 3.2), we shall derive the test statistic for the two cases, Case I: $\boldsymbol{\Sigma}_1 = \boldsymbol{\Sigma}_2$ and Case II: $n_1 = n_2$, separately. The general results for the moments, derived in Theorem 3.3, remain exactly the same.

3.3.3 Case I: $n_1 \neq n_2$, $\Sigma_1 = \Sigma_2$

Following lemma summarizes the properties of estimators.

Theorem 3.11 *For $\Sigma_1 = \Sigma_2 = \Sigma$, the estimators, $\tilde{B}_0^B, \tilde{B}_1^B$ and $\tilde{B}_2^B$, as defined in Equations (3.42)-(3.44), are the unbiased and consistent estimators of $2\text{tr}(\mathbf{T}\Sigma\mathbf{T})$, $4[\text{tr}(\mathbf{T}\Sigma\mathbf{T})]^2$ and $4\text{tr}(\mathbf{T}\Sigma\mathbf{T})^2$, respectively. Further, the variances of the ratios of the estimators to the traces that they estimate are uniformly bounded with respect to d .* $\square$

Following lemma summarizes the results needed for the moments of $\tilde{A}_N^B$.

Lemma 3.12 Let Q_N^B and $\tilde{B}_0^B$ be as defined in Equations (3.45) and (3.46), respectively. Then, for $\Sigma_1 = \Sigma_2$, we have the following where $N = n_1 + n_2$.

$$E(Q_N^B) = 2\text{tr}(\mathbf{T}\Sigma\mathbf{T}) \quad (3.47)$$

$$\text{Var}(Q_N^B) = 8\text{tr}(\mathbf{T}\Sigma\mathbf{T})^2 \quad (3.48)$$

$$\text{Cov}(Q_N^B, \tilde{B}_0^B) = \frac{4N}{n_1 n_2} \text{tr}(\mathbf{T}\Sigma\mathbf{T})^2. \quad (3.49)$$

$\square$

Substitution of these results into Equations (B.1) and (B.2) gives the following.

Theorem 3.13 *For $\Sigma_1 = \Sigma_2$, the first two moments of $\tilde{A}_N^B$ are given as following where $N = n_1 + n_2$.*

$$E(\tilde{A}_N^B) \approx 1$$

$$\text{Var}(\tilde{A}_N^B) \approx \frac{2}{f^B} \left(1 - \frac{3N-2}{4n_1 n_2} \right) \quad \square$$

3.3.4 Case II: $n_1 = n_2$, $\Sigma_1 \neq \Sigma_2$

The proof of the following theorem is similar to that of Theorem 3.7.

Theorem 3.14 *For $n_1 = n_2$, the estimators, $\tilde{B}_0^B, \tilde{B}_1^B$ and $\tilde{B}_2^B$, as defined in Equations (3.42)-(3.44), are unbiased and consistent estimators of $\text{tr}[\mathbf{T}(\Sigma_1 + \Sigma_2)\mathbf{T}]$, $[\text{tr}\{\mathbf{T}(\Sigma_1 + \Sigma_2)\mathbf{T}\}]^2$ and $\text{tr}[\mathbf{T}(\Sigma_1 + \Sigma_2)\mathbf{T}]^2$, respectively. Further, the variances of the ratios of estimators to the traces that they estimate are uniformly bounded with respect to d .* $\square$

Further, we need the following lemma to compute the moments of the test statistic.

Lemma 3.15 Let Q_N^B and $\tilde{B}_0^B$ be as defined in Equations (3.45) and (3.46), respectively. Then, with $n_1 = n_2$, we have

$$E(Q_N^B) = \text{tr}[\mathbf{T}(\boldsymbol{\Sigma}_1 + \boldsymbol{\Sigma}_2)\mathbf{T}] \quad (3.50)$$

$$\text{Var}(Q_N^B) = 2\text{tr}[\mathbf{T}(\boldsymbol{\Sigma}_1 + \boldsymbol{\Sigma}_2)\mathbf{T}]^2 \quad (3.51)$$

$$\text{Cov}(Q_N^B, \tilde{B}_0^B) = \frac{2}{n}\text{tr}[\mathbf{T}(\boldsymbol{\Sigma}_1 + \boldsymbol{\Sigma}_2)\mathbf{T}]^2. \quad (3.52)$$

□

Finally, Equations (B.1) and (B.2) give the following moments of $\tilde{A}_N^B$.

Theorem 3.16 For $n_1 = n_2$, the first two moments of the test statistic, $\tilde{A}_N^B$, are computed as following.

$$E(\tilde{A}_N^B) \approx 1$$

$$\text{Var}(\tilde{A}_N^B) \approx \frac{2}{f^B} \left(1 + \frac{\frac{1}{n} [\text{tr}(\mathbf{T}\boldsymbol{\Sigma}_1\mathbf{T})^2 + \frac{2}{n}\text{tr}(\mathbf{T}\boldsymbol{\Sigma}_1\mathbf{T}\boldsymbol{\Sigma}_2\mathbf{T}) + \text{tr}(\mathbf{T}\boldsymbol{\Sigma}_2\mathbf{T})^2]}{\text{tr}[\mathbf{T}(\boldsymbol{\Sigma}_1 + \boldsymbol{\Sigma}_2)\mathbf{T}]^2} - \frac{2}{n} \right)$$

□

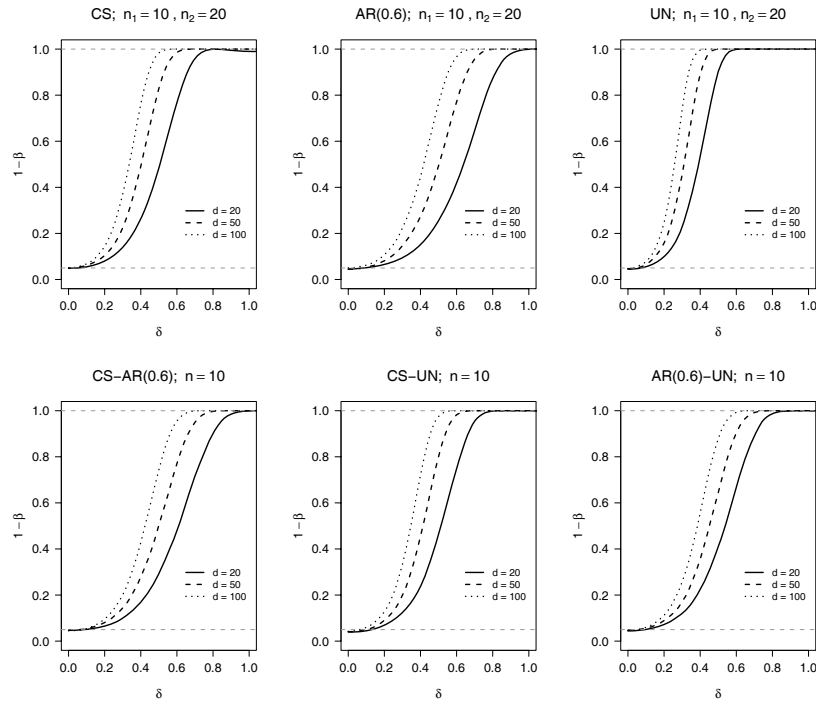
Since $\text{Var}(\tilde{A}_N^B)$ and $\text{Var}(\tilde{A}_N^{AB})$ are same, we expect the same accuracy of the statistic by estimating η using the plug in estimator, as discussed after Theorem 3.9; see also the simulation results in the next section.

3.3.5 Simulation Results

Table 3.3 reports the estimated quantiles for the modified ANOVA-type statistic for the time effect, $\tilde{A}_N^B$, whereas the power curves are shown in Fig. 5.2. These results are based on 10,000 simulation runs using exactly the same parameter settings as explained for the interaction effect in Section 3.2.5, including the form of alternative. We observe precisely the same behavior of the statistic as witnessed for the interaction effect, for both level and power, as expected, based on the same theoretical moments of the statistic for the two cases.

Table 3.3: ESTIMATED QUANTILES OF $\tilde{A}_N^B$: *Normal Distribution*

n_1	n_2	d	0.90	0.95	0.99	0.90	0.95	0.99	0.90	0.95	0.99
$(\Sigma_1 = \Sigma_2)$			CS			AR(0.6)			UN		
10	20	20	0.9007	0.9518	0.9941	0.9033	0.9520	0.9903	0.9032	0.9528	0.9917
		50	0.9054	0.9538	0.9928	0.9045	0.9526	0.9917	0.8987	0.9550	0.9927
		100	0.8992	0.9485	0.9913	0.9028	0.9517	0.9906	0.8980	0.9496	0.9914
		200	0.9022	0.9525	0.9918	0.9000	0.9502	0.9899	0.8993	0.9517	0.9930
20	30	30	0.8978	0.9503	0.9907	0.9014	0.9515	0.9881	0.9015	0.9485	0.9895
		50	0.9010	0.9545	0.9920	0.9098	0.9541	0.9905	0.8959	0.9456	0.9899
		100	0.9092	0.9553	0.9914	0.9043	0.9508	0.9884	0.8986	0.9482	0.9889
		200	0.9018	0.9519	0.9903	0.8977	0.9501	0.9884	0.9007	0.9496	0.9901
$(n_1 = n_2)$			CS-AR(0.6)			CS-UN			AR(0.6)-UN		
10	10	20	0.9065	0.9553	0.9952	0.9091	0.9592	0.9930	0.9069	0.9561	0.9934
		50	0.9059	0.9575	0.9938	0.9101	0.9576	0.9939	0.9069	0.9577	0.9943
		100	0.9059	0.9564	0.9946	0.8998	0.9511	0.9932	0.9038	0.9545	0.9938
		200	0.9066	0.9572	0.9940	0.9148	0.9594	0.9933	0.9100	0.9573	0.9936
20	20	30	0.9035	0.9524	0.9928	0.9042	0.9505	0.9906	0.9046	0.9535	0.9904
		50	0.9039	0.9574	0.9927	0.9046	0.9475	0.9889	0.9032	0.9515	0.9907
		100	0.9054	0.9542	0.9919	0.9084	0.9549	0.9901	0.9041	0.9500	0.9911
		200	0.9046	0.9508	0.9917	0.9018	0.9498	0.9895	0.8960	0.9515	0.9898

Figure 3.2: POWER CURVES FOR $\tilde{A}_N^B$: *Normal Distribution*

3.4 The Group Effect

As mentioned in Section 3.1.2 (see Remark 3.1), the group effect does not involve the dimensionality of data since H_0^A is a univariate hypothesis of the means of two independently sampled normal populations. Hence, H_0^A can be tested using the usual t -test for two independent samples. But, to keep the notations coherent for all model components and present the complete analysis in a unified form, the analysis of group effect is very briefly sketched in the following. For the classical analysis of the group effect, see Morrison (1967, Ch. 4), Rencher (1995, Ch. 5), Johnson and Wichern (2002, Ch. 6) and Timm (2002, Ch. 3).

The hypothesis for the group effect is $H_0^A : (\mathbf{P}_2 \otimes \frac{1}{d}\mathbf{J}_d') \boldsymbol{\mu} = \mathbf{0}$ which can also be written as $H_0^A : \frac{1}{d}\mathbf{1}'_d(\boldsymbol{\mu}_1 - \boldsymbol{\mu}_2) = \mathbf{0}$ or simply as $\bar{\mu}_{1.} = \bar{\mu}_{2.}$, where $\boldsymbol{\mu}_i = (\mu_{i1}, \dots, \mu_{id})'$, $i = 1, 2$, is the vector of means for the i th population and $\bar{\mu}_{i.}$ is the overall mean of the i th population. The $\mathbf{T}$ matrix for H_0^A is $\mathbf{T}^A = \frac{1}{d}\mathbf{J}_d$, where $\mathbf{T}^A$ is a projection matrix. Then, with $\bar{\mathbf{Y}}_1 - \bar{\mathbf{Y}}_2 = \mathbf{T}(\bar{\mathbf{X}}_1 - \bar{\mathbf{X}}_2)$, $\mathbf{T} = \frac{1}{d}\mathbf{J}_d$, $Q_N^A = \frac{2n_1n_2}{N}(\bar{\mathbf{Y}}_1 - \bar{\mathbf{Y}}_2)'(\bar{\mathbf{Y}}_1 - \bar{\mathbf{Y}}_2)$ is the quadratic form of the test statistic. Note that the formulation of Q_N^A , and similarly all other expressions, is the same as those for the interaction effect except the different $\mathbf{T}$ matrices for the two effects. Since, the derivations for the interaction effect (Section 3.2) are carried out using the general notation $\mathbf{T}$, hence, all the computations stand valid for the group effect as well. Therefore, no explicit derivations will be carried out for the group effect. But, to show how the quantities reduce to the scalar forms, we only summarize the main results for the sake of record.

Let $\mathbf{1}'\boldsymbol{\Sigma}_i\mathbf{1} = \sum_{s=1}^d \sigma_{is} = \sigma_i^2$, $i = 1, 2$, and $\mathbf{1}'(\mathbf{X}_{1k} - \mathbf{X}_{2l}) = X_{1k.} - X_{2l.}$. Then, the estimators, $\tilde{B}_0^A$, $\tilde{B}_1^A$ and $\tilde{B}_2^A$, are defined as

$$\tilde{B}_0^A = \frac{1}{n_1n_2} \sum_{k=1}^{n_1} \sum_{l=1}^{n_2} A_{kl}^A \quad (3.53)$$

$$\tilde{B}_1^A = \frac{1}{n_1n_2(n_1-1)(n_2-1)} \underbrace{\sum_{k=1}^{n_1} \sum_{l=1}^{n_2} \sum_{r=1}^{n_1} \sum_{s=1}^{n_2} A_{kl}^A A_{rs}^A}_{k \neq r, l \neq s} \quad (3.54)$$

$$\tilde{B}_2^A = \frac{1}{n_1n_2(n_1-1)(n_2-1)} \underbrace{\sum_{k=1}^{n_1} \sum_{l=1}^{n_2} \sum_{r=1}^{n_1} \sum_{s=1}^{n_2} A_{klrs}^{A^2}}_{k \neq r, l \neq s} \quad (3.55)$$

where $A_{kl}^A = \frac{1}{d}(X_{1k.} - X_{2l.})^2$, $A_{rs}^A = \frac{1}{d}(X_{1r.} - X_{2s.})^2$ are the quadratic forms and $A_{klrs}^A = \frac{1}{d}(X_{1k.} - X_{2l.})(X_{1r.} - X_{2s.})$ is the bilinear form.

Note that $\tilde{B}_1^A = \tilde{B}_2^A$, so that $\tilde{f}^A = 1$. Then, the following theorem is similar to Theorem 3.3.

Theorem 3.17 *Let Q_N^A , $\tilde{B}_0^A$, $\tilde{B}_1^A$ and $\tilde{B}_2^A$ be as defined above. Then, we have the following, where $N = n_1 + n_2$.*

$$E(\tilde{B}_0^A) = \frac{1}{d}(\sigma_1^2 + \sigma_2^2) \quad (3.56)$$

$$E(\tilde{B}_1^A) = \frac{1}{d^2}(\sigma_1^2 + \sigma_2^2)^2 \quad (3.57)$$

$$E(\tilde{B}_2^A) = \frac{1}{d^2}(\sigma_1^2 + \sigma_2^2)^2 \quad (3.58)$$

$$E(Q_N^A) = \frac{2}{Nd}(n_2\sigma_1^2 + n_1\sigma_2^2) \quad (3.59)$$

$$\text{Var}(\tilde{B}_0^A) = \frac{2}{n_1n_2d^2}(n_2\sigma_1^4 + 2\sigma_1^2\sigma_2^2 + n_1\sigma_2^4) \quad (3.60)$$

$$\text{Var}(Q_N^A) = \frac{8}{N^2d^2}(n_2\sigma_1^2 + n_1\sigma_2^2)^2 \quad (3.61)$$

$$\text{Cov}(Q_N^A, \tilde{B}_0^A) = \frac{4}{Nn_1n_2d^2}(n_2\sigma_1^2 + n_1\sigma_2^2)^2. \quad (3.62)$$

□

The results of Theorem 3.17 can be partitioned for the cases, when $\sigma_1^2 = \sigma_2^2 = \sigma^2$ and when $n_1 = n_2$. We summarize the properties of the estimators, for the two cases, in the following theorems.

Theorem 3.18 (Case I: $\sigma_1^2 = \sigma_2^2$) *The estimators, $\tilde{B}_0^A$, $\tilde{B}_1^A$ and $\tilde{B}_2^A$, as defined in Equations (3.53)-(3.55) above, are unbiased and consistent estimators of $2\sigma^2/d$, $4\sigma^4/d^2$ and $4\sigma^4/d^2$, respectively.* □

Theorem 3.19 (Case II: $n_1 = n_2$) *The estimators, $\tilde{B}_0^A$, $\tilde{B}_1^A$ and $\tilde{B}_2^A$, as defined in Equations (3.53)-(3.55) above, are unbiased and consistent estimators of $(\sigma_1^2 + \sigma_2^2)/d$, $(\sigma_1^2 + \sigma_2^2)^2/d^2$ and $(\sigma_1^2 + \sigma_2^2)^2/d^2$, respectively.* □

Remark 3.20 From Theorems 3.19 and 3.18, we deduce two interesting features of the estimators for the group effect. First, both $\tilde{B}_1^A$ and $\tilde{B}_2^A$ estimate the same quantity, since $[\text{tr}[\mathbf{T}(\boldsymbol{\Sigma}_1 + \boldsymbol{\Sigma}_2)\mathbf{T}]]^2 = \text{tr}[\mathbf{T}(\boldsymbol{\Sigma}_1 + \boldsymbol{\Sigma}_2)\mathbf{T}]^2$ when $\mathbf{T} = \mathbf{J}_d/d$. Further, the result of these traces is the square of the result of the third trace, $\text{tr}[\mathbf{T}(\boldsymbol{\Sigma}_1 + \boldsymbol{\Sigma}_2)\mathbf{T}]$, estimated by $\tilde{B}_0^A$. △

Now, let $\mathbf{1}'(\bar{\mathbf{X}}_{1..} - \bar{\mathbf{X}}_{2..}) = \bar{X}_{1..} - \bar{X}_{2..}$, so that $Q_N^A = \frac{1}{d}(\bar{X}_{1..} - \bar{X}_{2..})^2$. Then, the test statistic for the group effect is defined as $\tilde{A}_N^A = \frac{Q_N^A}{\tilde{B}_0^A}$, where, under H_0^A , $\tilde{A}_N^A \sim \chi_1^2$ since $f^A = 1$. Note that, with $\hat{\sigma}^2$ as the estimator of common σ^2 for Case I, we have

$$\tilde{A}_N^A = \frac{(\bar{X}_{1..} - \bar{X}_{2..})^2}{\hat{\sigma}^2 \left(\frac{1}{n_1} + \frac{1}{n_2} \right)} = \left(\frac{\bar{X}_{1..} - \bar{X}_{2..}}{\sqrt{\hat{\sigma}^2 \left(\frac{1}{n_1} + \frac{1}{n_2} \right)}} \right)^2 = t^2$$

where $t \sim t_{n_1+n_2-2}$ is the usual t statistic. We state the following theorem for the moments of the test statistic for both cases.

Theorem 3.21 *With $N = n_1 + n_2$, the first two moments of the test statistic, $\tilde{A}_N^A = \frac{Q_N^A}{\tilde{B}_0^A}$, are given as following.*

$$E(\tilde{A}_N^A) \approx 1 \quad (\text{Case I/II}) \quad (3.63)$$

$$\text{Var}(\tilde{A}_N^A) \approx 2 \left(1 - \frac{3N-2}{2n_1n_2} \right) \quad (\text{Case I}) \quad (3.64)$$

$$\approx 2 \left(1 + \frac{\frac{1}{n}(\sigma_1^4 + \frac{2}{n}\sigma_1^2\sigma_2^2 + \sigma_2^4)}{(\sigma_1^2 + \sigma_2^2)^2} - \frac{2}{n} \right) \quad (\text{Case II}) \quad (3.65)$$

where $\tilde{A}_N^A \sim \chi_1^2$. □

Since, for the group effect, $\tilde{f}^A = 1$, therefore, $\tilde{F}_N^A = \tilde{A}_N^A$.

Note that, like interaction and time effects, $\text{Var}(\tilde{A}_N^A)$ involves σ_i 's, $i = 1, 2$, for Case II. Simplifying it as we did for interaction and time effects, we have

$$\text{Var}(\tilde{A}_N^A) = 2 \left(1 - \frac{1}{n} \right) \left(1 - \frac{2\eta}{n} \right)$$

where $\eta = \frac{\sigma_1^2\sigma_2^2}{(\sigma_1^2 + \sigma_2^2)^2}$. Clearly, it is trivial for the group effect to replace σ_i^2 with the usual estimator $\hat{\sigma}_i^2$, $i = 1, 2$, to estimate η .

3.5 Applications

3.5.1 Analysis of Complete Sleep Lab Data

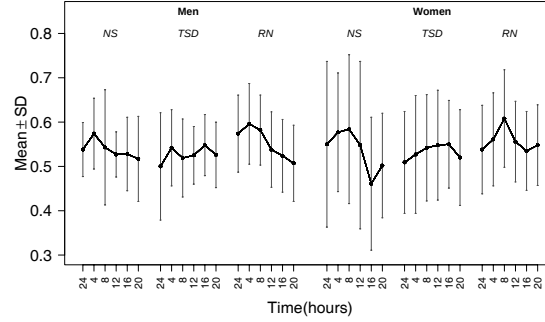


Figure 3.3: BAR PLOT FOR COMPLETE SLEEP LAB DATA

Table 3.4: ANALYSIS OF COMPLETE SLEEP LAB DATA

Effect	T-matrix	$\tilde{F}_N$	$\tilde{f}$	p-value
Gender	$\mathbf{P}_2 \otimes \frac{1}{3}\mathbf{J}_3 \otimes \frac{1}{6}\mathbf{J}_6$	0.01	1.24	0.9561
Day	$\frac{1}{2}\mathbf{J}_2 \otimes \mathbf{P}_3 \otimes \frac{1}{6}\mathbf{J}_6$	0.66	2.59	0.5529
Time	$\frac{1}{2}\mathbf{J}_2 \otimes \frac{1}{3}\mathbf{J}_3 \otimes \mathbf{P}_6$	4.75	3.53	0.0014
Gender×Day	$\mathbf{P}_2 \otimes \mathbf{P}_3 \otimes \frac{1}{6}\mathbf{J}_6$	0.01	2.39	0.9938
Gender×Time	$\mathbf{P}_2 \otimes \frac{1}{3}\mathbf{J}_3 \otimes \mathbf{P}_6$	2.00	7.02	0.0515
Day×Time	$\frac{1}{2}\mathbf{J}_2 \otimes \mathbf{P}_3 \otimes \mathbf{P}_6$	2.82	9.49	0.0021
Gender×Day×Time	$\mathbf{P}_2 \otimes \mathbf{P}_3 \otimes \mathbf{P}_6$	1.48	14.52	0.1067
Time(NS-TSD)	$\mathbf{P}_2 \otimes \mathbf{P}_6$	3.73	4.38	0.0036
Time(NS-RN)	$\mathbf{P}_2 \otimes \mathbf{P}_6$	0.73	5.98	0.6212
Time(TSD-RN)	$\mathbf{P}_2 \otimes \mathbf{P}_6$	3.14	5.44	0.0061
Time (NS)	$\mathbf{P}_6$	4.01	4.27	0.0023
Time (TSD)	$\mathbf{P}_6$	1.84	6.78	0.0783
Time (RN)	$\mathbf{P}_6$	4.33	4.32	0.0012

The analysis of women group of sleep lab example (Example 1.3) was carried out in Section 2.5.1. Here we apply the two sample modified ANOVA-type statistic to analyze the whole data given in Section C.3.

We have analyzed the data for both cases I and II, as discussed above, and found the results virtually same. Therefore, the results in Tables 3.4 and 3.5 refer to Case II ($\Sigma_1 \neq \Sigma_2$). From Fig. 5.3, a statistically discernable contribution of NS can be expected. We observe a significant Day $\times$ Time interaction effect, shown in Fig. 5.4. This interaction is partitioned in the middle panel of Table 3.4 and further, for simple effects, in the bottom panel of the table. The final conclusion is the same as drawn for the women data in Section 2.5.1, that RN and NS are the main contributors to the significant interaction effect. This fact is also clear, to some extent, from the non-significant group effect.

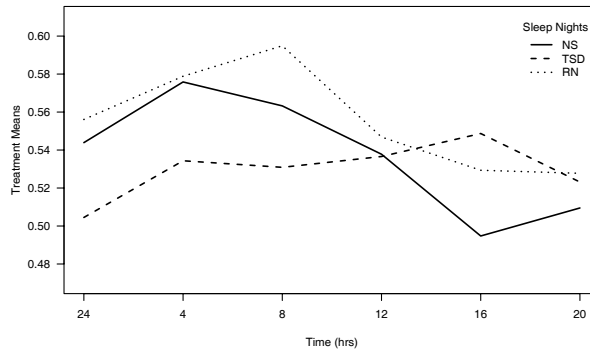


Figure 3.4: DAY $\times$ TIME INTERACTION PLOT FOR COMPLETE SLEEP LAB DATA

3.5.2 Analysis of Wistar Rats Data

The analysis of the Wistar rats experiment (Example 1.2) is reported in Table 3.5. We have $d = 22$ and $n_i = 10$, $i = 1, 2$. We observe a highly significant time effect but a non-significant interaction. The continuously increasing time profiles of rats in Fig. 5.5 clue to the significant time effect whereas the parallel lines for the drug and placebo groups in Fig. 5.6 clearly support the conclusion that there is no interaction between the treatment and the time.

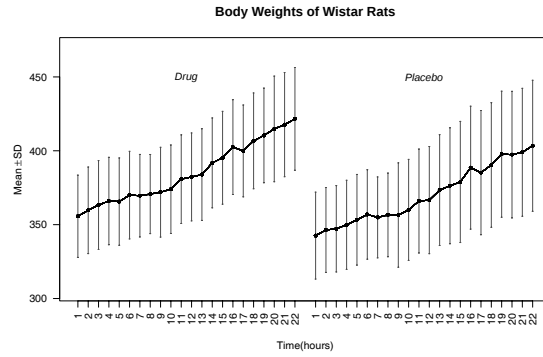


Figure 3.5: BAR PLOT FOR WISTAR RATS DATA

Table 3.5: ANALYSIS OF RATS DATA

Effect	T-matrix	$\tilde{F}_N$	$\tilde{f}$	p-value
Group	$\mathbf{P}_2 \otimes \mathbf{P}_{22}$	1.00	1.17	0.3293
Time	$\frac{1}{2}\mathbf{J}_2 \otimes \mathbf{P}_{22}$	9.56	1.21	0.0010
Group×Time	$\mathbf{P}_2 \otimes \frac{1}{22}\mathbf{J}_{22}$	0.51	3.48	0.7015

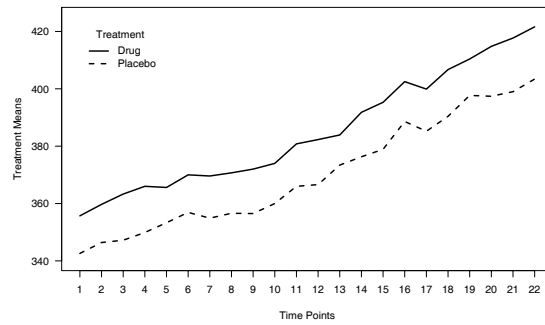


Figure 3.6: GROUP×TIME INTERACTION PLOT FOR RATS DATA

Chapter 4

The Multivariate Case

As mentioned in Section 2.1, the projection matrix $\mathbf{T}$ in $H_0 : \mathbf{T}\boldsymbol{\mu} = \mathbf{0}$ can be formulated so that H_0 is any general linear hypothesis. Perhaps the most commonly tested multivariate one-sample hypothesis is $H_0 : \boldsymbol{\mu} = \mathbf{0}$. For $n > d$, the most commonly used test statistic for this hypothesis is Hotelling's T^2 , as discussed in Section 1.2.2, where it is also stated that T^2 statistic is not applicable when $d > n$ (Stuart et al, 1999, Ch. 31).

The hypothesis $H_0 : \mathbf{T}\boldsymbol{\mu} = \mathbf{0}$ reduces to $H_0 : \boldsymbol{\mu} = \mathbf{0}$ when $\mathbf{T} = \mathbf{I}$, and $\mathbf{I}$ is a projection matrix. Similarly, in the two sample case, the hypotheses can be reduced to the multivariate case by replacing $\mathbf{T}$ with $\mathbf{I}$. For example, the interaction hypothesis then becomes $H_0^{AB} : \boldsymbol{\mu}_1 = \boldsymbol{\mu}_2$. The Hotelling's T^2 statistic can not be used to test this hypothesis when $n_1 + n_2 - 2 < d$. It is quite straightforward to prove that the modified ANOVA-type statistic, for both one and two sample cases, is also valid to test the multivariate hypotheses when the matrix $\mathbf{T}$ is replaced with $\mathbf{I}$ in all the expressions involved. The data structure remains the same as introduced in Section 2.1 except that the d observations measured on each of n individuals, in the multivariate case, can be any set of measurements taken on the same individual, for example, observations taken under different conditions, or observations taken on different characteristics (variables). It is also easy to see that the properties of the estimators and the statistics remain intact for the multivariate case. Due to its triviality, this case will not be dealt with in detail here.

Some recent references which also address the testing of multivariate hypotheses for high dimensional data are reported in Section 1.5. A point that deserves explicit treatment in the high dimensional multivariate case is invariance. The modified ANOVA-type statistic, both for one and two sample

cases, is not intrinsically invariant under affine transformation (Definition B.9). It is, however, invariant to an orthogonal linear transformation (Definition B.10). Here, *not intrinsically invariant under affine transformation* means that the statistic, although not invariant by construction, remains so when used for a repeated measures set up. It is because in a repeated measures set up, the observations are measured on a commensurate scale. It further implies that the statistic is also invariant when each component undergoes the same linear transformation, even if the design set up is not of repeated measurements.

It must be noted that, none of the statistics developed for the high dimensional data so far (see references in Section 1.5), is affine invariant. The reason, all these statistics primarily share for not being affine invariant, is that all are approximate statistics. Lehmann (1959, p 318) states that for high dimensional data, only an exact α level test can be affine invariant.

Chapter 5

Summary and Outlook

5.1 Summary and Conclusions

Table 5.1: SUMMARY ANOVA TABLE: ONE-SAMPLE CASE

Effect	T-matrix	Identities	$\tilde{F}_n \sim \chi_f^2$	$\tilde{f}$
Time	$\mathbf{P}_d$	$Q_n = n\bar{\mathbf{X}}'\mathbf{T}\bar{\mathbf{X}} = n\bar{\mathbf{Y}}'\bar{\mathbf{Y}}$ $B_0 = \frac{1}{n} \sum_{k=1}^n A_k$ $B_1 = \frac{1}{n(n-1)} \sum_{k=1}^n \sum_{\substack{l=1 \\ k \neq l}}^n A_k A_l$ $B_2 = \frac{1}{n(n-1)} \sum_{k=1}^n \sum_{\substack{l=1 \\ k \neq l}}^n A_{kl}^2$ $A_k = \mathbf{Y}_k' \mathbf{Y}_k, A_l = \mathbf{Y}_l' \mathbf{Y}_l, A_{kl} = \mathbf{Y}_k' \mathbf{Y}_l$	$\frac{Q_n}{B_0} \cdot \frac{B_1}{B_2}$	$\frac{B_1}{B_2}$

First, in Chapter 2, a one sample test statistic is derived for the analysis of repeated measures data when the dimension, d , can (but not must) be large compared to the sample size, n , and the underlying distribution is assumed multivariate normal. This case is then extended to the two sample case in Chapter 3. A small digression to the multivariate case is briefly sketched in Chapter 4.

The statistic, the modified ANOVA-type statistic (Brunner, 2001), is

Table 5.2: SUMMARY ANOVA TABLE: TWO-SAMPLE CASE

Effect	T-matrix	Identities	$\tilde{F}_n \sim \chi_f^2$	$\tilde{f}$
Interaction	$\mathbf{P}_2 \otimes \mathbf{P}_d$	$Q_N^{AB} = \frac{2n_1n_2}{N}(\bar{\mathbf{Y}}_{1.} - \bar{\mathbf{Y}}_{2.})'(\bar{\mathbf{Y}}_{1.} - \bar{\mathbf{Y}}_{2.})$ $\tilde{B}_0^{AB} = \frac{1}{n_1n_2} \sum_{k=1}^{n_1} \sum_{l=1}^{n_2} A_{kl}^{AB}$ $\tilde{B}_1^{AB} = \frac{1}{n_1n_2(n_1-1)(n_2-1)} \sum_{k=1}^{n_1} \sum_{l=1}^{n_2} \sum_{r=1}^{n_1} \sum_{s=1}^{n_2} A_{kl}^{AB} A_{rs}^{AB}$ $\tilde{B}_2^{AB} = \frac{1}{n_1n_2(n_1-1)(n_2-1)} \sum_{k=1}^{n_1} \sum_{l=1}^{n_2} \sum_{r=1}^{n_1} \sum_{s=1}^{n_2} A_{klrs}^{AB^2}$ $A_{kl}^{AB} = (\mathbf{Y}_{1k} - \mathbf{Y}_{2l})'(\mathbf{Y}_{1k} - \mathbf{Y}_{2l}), A_{rs}^{AB} = (\mathbf{Y}_{1r} - \mathbf{Y}_{2s})'(\mathbf{Y}_{1r} - \mathbf{Y}_{2s}), A_{klrs}^{AB} = (\mathbf{Y}_{1k} - \mathbf{Y}_{2l})'(\mathbf{Y}_{1r} - \mathbf{Y}_{2s})$	$\frac{Q_N^{AB}}{\tilde{B}_0^{AB}} \cdot \frac{\tilde{B}_1^{AB}}{\tilde{B}_2^{AB}}$	$\frac{\tilde{B}_1^{AB}}{\tilde{B}_2^{AB}}$
Time	$\frac{1}{2}\mathbf{J}_2 \otimes \mathbf{P}_d$	$Q_N^B = \frac{2n_1n_2}{N}(\bar{\mathbf{Y}}_{1.} + \bar{\mathbf{Y}}_{2.})'(\bar{\mathbf{Y}}_{1.} + \bar{\mathbf{Y}}_{2.})$ $\tilde{B}_0^B = \frac{1}{n_1n_2} \sum_{k=1}^{n_1} \sum_{l=1}^{n_2} A_{kl}^B$ $\tilde{B}_1^B = \frac{1}{n_1n_2(n_1-1)(n_2-1)} \sum_{k=1}^{n_1} \sum_{l=1}^{n_2} \sum_{r=1}^{n_1} \sum_{s=1}^{n_2} A_{kl}^B A_{rs}^B$ $\tilde{B}_2^B = \frac{1}{n_1n_2(n_1-1)(n_2-1)} \sum_{k=1}^{n_1} \sum_{l=1}^{n_2} \sum_{r=1}^{n_1} \sum_{s=1}^{n_2} A_{klrs}^{B^2}$ $A_{kl}^B = (\mathbf{Y}_{1k} + \mathbf{Y}_{2l})'(\mathbf{Y}_{1k} + \mathbf{Y}_{2l}), A_{rs}^B = (\mathbf{Y}_{1r} + \mathbf{Y}_{2s})'(\mathbf{Y}_{1r} + \mathbf{Y}_{2s}), A_{klrs}^B = (\mathbf{Y}_{1k} + \mathbf{Y}_{2l})'(\mathbf{Y}_{1r} + \mathbf{Y}_{2s})$	$\frac{Q_N^B}{\tilde{B}_0^B} \cdot \frac{\tilde{B}_1^B}{\tilde{B}_2^B}$	$\frac{\tilde{B}_1^B}{\tilde{B}_2^B}$
Group	$\mathbf{P}_2 \otimes \frac{1}{d}\mathbf{J}_d$	$Q_N^A = \frac{2n_1n_2}{N}(\bar{\mathbf{Y}}_{1.} - \bar{\mathbf{Y}}_{2.})'(\bar{\mathbf{Y}}_{1.} - \bar{\mathbf{Y}}_{2.})$ $\tilde{B}_0^A = \frac{1}{n_1n_2} \sum_{k=1}^{n_1} \sum_{l=1}^{n_2} A_{kl}^A$ $\tilde{B}_1^A = \frac{1}{n_1n_2(n_1-1)(n_2-1)} \sum_{k=1}^{n_1} \sum_{l=1}^{n_2} \sum_{r=1}^{n_1} \sum_{s=1}^{n_2} A_{kl}^A A_{rs}^A$ $\tilde{B}_2^A = \frac{1}{n_1n_2(n_1-1)(n_2-1)} \sum_{k=1}^{n_1} \sum_{l=1}^{n_2} \sum_{r=1}^{n_1} \sum_{s=1}^{n_2} A_{klrs}^{A^2}$ $A_{kl}^A = \frac{1}{d}(X_{1k.} - X_{2l.})^2, A_{rs}^A = \frac{1}{d}(X_{1r.} - X_{2s.})^2, A_{klrs}^A = \frac{1}{d}(X_{1k.} - X_{2l.})(X_{1r.} - X_{2s.}), N = n_1 + n_2$	$\frac{Q_N^A}{\tilde{B}_0^A}$	$\frac{\tilde{B}_1^A}{\tilde{B}_2^A} = 1$

based on Box's approximation and follows a χ_f^2 distribution, asymptotically. The formulations of the statistic are summarized in Tables 5.1 and 5.2, respectively, for one and two sample cases. Following are the main characteristics, shared by both one and two sample test statistics.

C-1 The statistics follow a χ_f^2 distribution, asymptotically.

- C-2** The statistics are composed of the unbiased, $\mathcal{L}_2$ -consistent and uniformly bounded estimators. Although, unbiasedness and consistency are two of the most desirable properties of an estimator, in the present context, it is the uniform boundedness of the estimators with respect to the dimension, d , which plays the most crucial role in determining the accuracy of the approximating distribution of the statistics to the χ^2 distribution. This property ensures that the asymptotic approximation of the test statistics is achieved by letting only $n \rightarrow \infty$ where d is kept finite and fixed, such that, the accuracy is not jeopardized by $d > n$, or even by $d \gg n$.
- C-3** The statistics can be used to analyze any general linear hypothesis, which further implies that they can be used to analyze unstructured, or factorially structured repeated measures data.
- C-4** The statistics can also be used when the data are not high dimensional, i.e. when $d < n$. In this case, the statistics can be compared with the most frequently used classical multivariate techniques, e.g., Hotelling's T^2 statistic. There is no uniformly better statistic among these competitors and the performance strongly depends on the underlying covariance matrix and the type of alternatives.
- C-5** The statistics are invariant to an orthogonal transformation. As far as the affine invariance (Definition B.9) is concerned, it is a windfall for the statistics that the problem of invariance does not potentially emerge. It is simply because the measurements in a repeated measures set up are made on a commensurate scale. It implies that, intrinsically, the statistics also remain invariant when each component undergoes the same transformation, even if the data set up is not of repeated measures.
- C-6** The statistics do not depend on any specific covariance structure and are derived under fairly general conditions. This advantage gives the statistic a unique flavor and a potential appeal of application. As discussed in Chapter 1, repeated measures data can be analyzed using both univariate and multivariate techniques. A univariate analysis reduces the computational burden by using the traces of the concerned matrices but is limited by the almost inevitable sphericity assumption on the covariance matrix. A multivariate analysis, on the other hand,

enjoys the flexibility of not depending on a particular covariance structure, but totally collapses when the data are high dimensional. The modified ANOVA-type statistic efficiently collects the positives of the two approaches and shuns the negatives: it is computationally based on the traces of the functions of the underlying covariance matrix but does not call for imposing any particular structure on this covariance matrix.

C-7 The statistics are subjected to the evaluation based on simulation studies which indicate that the approximation is accurate enough for moderate sample size, usually $n = 10$, disregarding how large d is.

5.2 Future Outlook

Given the immense amount of applications of the repeated measures theory and the rapidly emerging challenge of dealing with huge amounts of data in a variety of applied research problems, be they in the biological sciences, behavioral sciences, or even in engineering and astronomy, it is absolutely fair to say that much more needs to be done than what is presented in this manuscript. As mentioned in Section 1.5, the analysis of high dimensional data has attracted a reasonable amount of researchers in the last couple of decades and they have approached the problem from a variety of perspectives. Although not limited to, but the analysis of proteomics and microarray data has significantly galvanized the need to enhance the efficiency of statistical methods so that they can handle the large amounts of data. Add to it the recent revolution of modern statistical softwares and highly efficient algorithmic approach to tune the traditional statistical instruments so they can play better when called for in the ever growing realm of target-driven, application-oriented research environment. The following list of potential extensions of the present work is, therefore, by no means exhaustive.

1. Extension to the general case of $a > 2$ populations.
2. Developing a sound methodology to deal with the multiple comparisons problem in high dimensional data.
3. Constructing an affine invariant statistic for the analysis of high dimensional data.

4. As an alternative, but currently much sought after, approach is to develop, as Ledoit and Wolf (2004) put it, a well-conditioned estimator of the high dimensional covariance matrix that can be used as a plug-in estimator for the classical test statistics.
5. A completely non-parametric test statistic is needed to deal with those types of data sets and alternatives which are beyond the grasp of subtle parametric statistical theory.

Appendix A

Quadratic and Bilinear Forms

This chapter contains the main results on quadratic and bilinear forms and their moments which are frequently used in the derivations of the test statistics. There is an extensive literature on quadratic and bilinear forms containing most of the fundamental results we need; see, for example, Searle (1971), Mathai and Provost (1992), Mathai et al (1995). The derivations of most of the results for the present work are based on the transformation of the form $\mathbf{Y} = \mathbf{TX}$. We, therefore, state the following result from the multivariate theory for reference, the proof of which is quite trivial (see Rencher, 1995, Ch. 3).

Theorem A.1 *Let $\mathbf{X}$ be a vector of n observations with $E(\mathbf{X}) = \boldsymbol{\mu}$ and $Cov(\mathbf{X}) = \boldsymbol{\Sigma}$. Define the transformation $\mathbf{Y} = \mathbf{AX}$, where $\mathbf{A}_{k \times n}$ be any matrix of constants. Then, $E(\mathbf{Y}) = \mathbf{A}\boldsymbol{\mu}$ and $Cov(\mathbf{Y}) = \mathbf{A}\boldsymbol{\Sigma}\mathbf{A}'$. $\square$*

Corollary A.2 In our case, the matrix $\mathbf{T}$ in the transformation $\mathbf{Y} = \mathbf{TX}$ is symmetric (and idempotent). Hence, $Cov(\mathbf{Y}) = \mathbf{T}\boldsymbol{\Sigma}\mathbf{T}$. $\square$

Theorem A.1 is valid without any distributional assumption for $\mathbf{X}$. Clearly, from the properties of normal distribution, we have the following special case (assuming again symmetric $\mathbf{T}$, although the result holds in general). For details, see Anderson (1984, Ch. 2).

Lemma A.3 If $\mathbf{X} \sim \mathcal{N}(\boldsymbol{\mu}, \boldsymbol{\Sigma})$, then $\mathbf{Y} = \mathbf{TX} \sim \mathcal{N}(\mathbf{T}\boldsymbol{\mu}, \mathbf{T}\boldsymbol{\Sigma}\mathbf{T})$. $\square$

A.1 Moments of Quadratic Forms

We begin with the following general representation of a quadratic form. For details, see Mathai and Provost (1992, Ch. 3).

Lemma A.4 (Representation of a Quadratic Form) Let $\mathbf{X}$ be a random vector with $E(\mathbf{X}) = \boldsymbol{\mu}$ and $\text{Cov}(\mathbf{X}) = \boldsymbol{\Sigma} > 0$ and $\mathbf{A} = \mathbf{A}'$ be any matrix. Then the quadratic form $\mathbf{X}'\mathbf{A}\mathbf{X}$ can be represented as

$$Q(\mathbf{X}) = \mathbf{X}'\mathbf{A}\mathbf{X} = \sum_{j=1}^d \lambda_j (U_j + b_j)^2 \quad \boldsymbol{\mu} \neq \mathbf{0} \quad (\text{A.1})$$

$$= \sum_{j=1}^d \lambda_j U_j^2 \quad \boldsymbol{\mu} = \mathbf{0} \quad (\text{A.2})$$

where $\lambda_1, \dots, \lambda_d$ are the eigenvalues of $\boldsymbol{\Sigma}^{\frac{1}{2}}\mathbf{A}\boldsymbol{\Sigma}^{\frac{1}{2}}$, $E(\mathbf{U}) = \mathbf{0}$, $\text{Cov}(\mathbf{U}) = \mathbf{I}_d$ with $\mathbf{U} = (U_1, \dots, U_d)'$, $\mathbf{b} = (b_1, \dots, b_d)' = \left(\mathbf{P}'\boldsymbol{\Sigma}^{-\frac{1}{2}}\boldsymbol{\mu}\right)'$ and $\mathbf{P}$ be any orthogonal matrix such that $\mathbf{P}\mathbf{P}' = \mathbf{I}$. $\square$

Remark A.5 This representation does not assume normal, or any, distribution for $\mathbf{X}$ except that the covariance matrix is positive definite. Hence, it is the general representation and is valid for any underlying non-singular distribution. Several variants of this representation, including those for the singular case, are given in Mathai and Provost (1992, Ch. 3). $\triangle$

Remark A.6 When $U_j \sim N(0, 1)$, $\forall j$, then $U_j^2 \sim \chi_1^2$ and Lemma A.4 can be replaced with Theorem 2.1. $\triangle$

Though, frequently we shall need the first two moments of the quadratic and bilinear forms, in some cases the higher moments will also be required. Following theorem gives the general moment generating function of a quadratic form under the assumption of normal distribution.

Theorem A.7 (Moment Generating Function) Let $\mathbf{X} \sim \mathcal{N}_d(\boldsymbol{\mu}, \boldsymbol{\Sigma})$, $\boldsymbol{\Sigma} > 0$ and $\mathbf{A} = \mathbf{A}'$ be any matrix of constants. Then the mgf of the quadratic form $\mathbf{X}'\mathbf{A}\mathbf{X}$ is given as

$$\begin{aligned} M_{\mathbf{X}'\mathbf{A}\mathbf{X}}(t) &= |\mathbf{I} - 2t\mathbf{A}\boldsymbol{\Sigma}|^{-\frac{1}{2}} e^{-\frac{1}{2}\boldsymbol{\mu}'[\mathbf{I} - (\mathbf{I} - 2t\mathbf{A}\boldsymbol{\Sigma})^{-1}]\boldsymbol{\Sigma}^{-1}\boldsymbol{\mu}} \\ &= |\mathbf{I} - 2t\mathbf{A}\boldsymbol{\Sigma}|^{-\frac{1}{2}} \quad \boldsymbol{\mu} = \mathbf{0}. \end{aligned} \quad \square$$

Proof See Searle (1971, p 55). ■

Remark A.8 In Theorem A.7, we assume the covariance matrix to be non-singular, which suffices our needs. For the more general case, with singular covariance matrix, see Mathai and Provost (1992, Ch. 3). △

As is obvious from Theorem A.7, it is not easy to deal with the mgf of a quadratic form. Although, we can write $|\mathbf{I} - 2t\mathbf{A}\mathbf{\Sigma}|^{-\frac{1}{2}} = \prod_{j=1}^d (1 - 2t\lambda_j)^{-\frac{1}{2}}$, but even then it is much easier to deal with the cumulant generating function which is the log of the mgf. The moment-cumulant relationships (Stuart and Ord, 1994, Ch. 3) can then be used to compute the moments. The general form of the cgf of a quadratic form is given in the following theorem.

Theorem A.9 (Cumulant Generating Function) *Let $\mathbf{X} \sim \mathcal{N}(\boldsymbol{\mu}, \mathbf{\Sigma})$, $\mathbf{\Sigma} > 0$, and $\mathbf{A} = \mathbf{A}'$ be any matrix of constants. Then the r th cumulant of the quadratic form $\mathbf{X}'\mathbf{A}\mathbf{X}$ is given as*

$$\kappa_r(\mathbf{X}'\mathbf{A}\mathbf{X}) = 2^{r-1}(r-1)! [\text{tr}(\mathbf{A}\mathbf{\Sigma})^r + r\boldsymbol{\mu}'\mathbf{A}(\mathbf{\Sigma}\mathbf{A})^{r-1}\boldsymbol{\mu}], \quad r = 1, 2, \dots \quad (\text{A.3})$$

where tr denotes the trace. □

Proof See Searle (1971, Ch. 2) or Mathai et al (1995, Ch. 3). ■

Clearly, when $\boldsymbol{\mu} = \mathbf{0}$, Equation (A.3) reduces to a the following simpler form.

Corollary A.10 When $\boldsymbol{\mu} = \mathbf{0}$, the r th cumulant of $\mathbf{X}'\mathbf{A}\mathbf{X}$ reduces to

$$\kappa_r(\mathbf{X}'\mathbf{A}\mathbf{X}) = 2^{r-1}(r-1)!\text{tr}(\mathbf{A}\mathbf{\Sigma})^r, \quad r = 1, 2, \dots \quad (\text{A.4}) \quad \square$$

Since the computations of the test statistic are based on the moments of quadratic forms for the transformed variables, therefore, we use Corollary A.10 to deduce the first four moments. For the one sample case, $\mathbf{X}_k \sim \mathcal{N}(\boldsymbol{\mu}, \mathbf{\Sigma})$, such that $\mathbf{Y}_k = \mathbf{T}\mathbf{X}_k \sim \mathcal{N}(\mathbf{0}, \mathbf{T}\mathbf{\Sigma}\mathbf{T})$, $k = 1, \dots, n$, and the quadratic form is defined as $A_k = \mathbf{X}_k'\mathbf{T}\mathbf{X}_k = \mathbf{Y}_k'\mathbf{Y}_k$, since $\mathbf{T}$ is idempotent. Following lemma summarizes the final results for the one sample case.

Lemma A.11 Under $H_0 : \mathbf{T}\boldsymbol{\mu} = \mathbf{0}$, the first four moments of the quadratic form $\mathbf{X}_k'\mathbf{T}\mathbf{X}_k$ are

$$\left. \begin{aligned} \mu'_1 &= \text{tr}(\mathbf{T}\mathbf{\Sigma}) \\ \mu'_2 &= 2\text{tr}(\mathbf{T}\mathbf{\Sigma})^2 + [\text{tr}(\mathbf{T}\mathbf{\Sigma})]^2 \\ \mu'_3 &= [\text{tr}(\mathbf{T}\mathbf{\Sigma})]^3 + 6\text{tr}(\mathbf{T}\mathbf{\Sigma})^2\text{tr}(\mathbf{T}\mathbf{\Sigma}) + 8\text{tr}(\mathbf{T}\mathbf{\Sigma})^3 \\ \mu'_4 &= 48\text{tr}(\mathbf{T}\mathbf{\Sigma})^4 + 32\text{tr}(\mathbf{T}\mathbf{\Sigma})^3\text{tr}(\mathbf{T}\mathbf{\Sigma}) + 12\text{tr}(\mathbf{T}\mathbf{\Sigma})^2[\text{tr}(\mathbf{T}\mathbf{\Sigma})]^2 \\ &\quad + 12[\text{tr}(\mathbf{T}\mathbf{\Sigma})^2]^2 + [\text{tr}(\mathbf{T}\mathbf{\Sigma})]^4 \end{aligned} \right\} \quad (\text{A.5})$$

where $\mu'_i = E(\mathbf{X}'_k \mathbf{T} \mathbf{X}_k)^i$, $i = 1, \dots, 4$. $\square$

The mean and variance of a quadratic form, being most frequently needed identities, are stated, again under $H_0 : \mathbf{T}\boldsymbol{\mu} = \mathbf{0}$, in the following.

$$E(\mathbf{X}'_k \mathbf{T} \mathbf{X}_k) = \text{tr}(\mathbf{T}\boldsymbol{\Sigma}) \quad (\text{A.6})$$

$$\text{Var}(\mathbf{X}'_k \mathbf{T} \mathbf{X}_k) = 2\text{tr}(\mathbf{T}\boldsymbol{\Sigma})^2. \quad (\text{A.7})$$

Similarly, we can proceed for the two sample case. We have $\mathbf{X}_{1k} \stackrel{i.i.d.}{\sim} \mathcal{N}(\boldsymbol{\mu}_1, \boldsymbol{\Sigma}_1)$ and $\mathbf{X}_{2l} \stackrel{i.i.d.}{\sim} \mathcal{N}(\boldsymbol{\mu}_2, \boldsymbol{\Sigma}_2)$. Since, under $H_0^{AB} : \mathbf{T}(\boldsymbol{\mu}_1 - \boldsymbol{\mu}_2) = \mathbf{0}$,

$$\mathbf{T}(\mathbf{X}_{1k} - \mathbf{X}_{2l}) = (\mathbf{Y}_{1k} - \mathbf{Y}_{2l}) \sim \mathcal{N}(\mathbf{0}, \mathbf{T}(\boldsymbol{\Sigma}_1 + \boldsymbol{\Sigma}_2)\mathbf{T})$$

and under $H_0^B : \mathbf{T}(\boldsymbol{\mu}_1 + \boldsymbol{\mu}_2) = \mathbf{0}$,

$$\mathbf{T}(\mathbf{X}_{1k} + \mathbf{X}_{2l}) = (\mathbf{Y}_{1k} + \mathbf{Y}_{2l}) \sim \mathcal{N}(\mathbf{0}, \mathbf{T}(\boldsymbol{\Sigma}_1 + \boldsymbol{\Sigma}_2)\mathbf{T}),$$

therefore, the final results for both interaction and time effects are same. The moments of quadratic forms, $A_{kl}^{AB} = (\mathbf{Y}_{1k} - \mathbf{Y}_{2l})'(\mathbf{Y}_{1k} - \mathbf{Y}_{2l})$ and $A_{kl}^B = (\mathbf{Y}_{1k} + \mathbf{Y}_{2l})'(\mathbf{Y}_{1k} + \mathbf{Y}_{2l})$, come directly from Lemma A.11. The mean and variance of the quadratic form for the two sample case are collected in the following, where $A_{kl} = A_{kl}^{AB}$ or $A_{kl} = A_{kl}^B$.

$$E(A_{kl}) = \text{tr}[\mathbf{T}(\boldsymbol{\Sigma}_1 + \boldsymbol{\Sigma}_2)\mathbf{T}] \quad (\text{A.8})$$

$$\text{Var}(A_{kl}) = 2\text{tr}[\mathbf{T}(\boldsymbol{\Sigma}_1 + \boldsymbol{\Sigma}_2)\mathbf{T}]^2. \quad (\text{A.9})$$

Due to the independence of the quadratic forms involved in most of the expressions required for the derivations of the statistics, the proofs regarding joint moments are quite simple. For more general and detailed treatment on the joint moments, see Mathai and Provost (1992). In the following theorem, we compute some useful joint moments of the quadratic form, for the one sample case.

Theorem A.12 For $A_j = \mathbf{Y}'_j \mathbf{Y}_j$, $j = k, l, r, s$, as the quadratic forms for

the one sample case, we have the following, under $H_0 : \mathbf{T}\boldsymbol{\mu} = \mathbf{0}$.

$$\text{Cov}(A_k, A_l) = \begin{cases} 0 & k \neq l \\ 2\text{tr}(\mathbf{T}\boldsymbol{\Sigma}\mathbf{T})^2 & k = l \end{cases} \quad (\text{A.10})$$

$$\text{Cov}(A_k^2, A_l^2) = \begin{cases} 0 & k \neq l \\ 48\text{tr}(\mathbf{T}\boldsymbol{\Sigma}\mathbf{T})^2 + 12[\text{tr}(\mathbf{T}\boldsymbol{\Sigma}\mathbf{T})^2]^2 & k = l \end{cases} \quad (\text{A.11})$$

$$\text{Cov}(A_k A_l, A_r A_s) = \begin{cases} 0 & k \neq r, l \neq s \\ 2\text{tr}(\mathbf{T}\boldsymbol{\Sigma}\mathbf{T})^2 [\text{tr}(\mathbf{T}\boldsymbol{\Sigma}\mathbf{T})]^2 & k \neq r, l = s \text{ or } k = r, l \neq s \\ 4[\text{tr}(\mathbf{T}\boldsymbol{\Sigma}\mathbf{T})^2]^2 + 4\text{tr}(\mathbf{T}\boldsymbol{\Sigma}\mathbf{T})^2 [\text{tr}(\mathbf{T}\boldsymbol{\Sigma}\mathbf{T})]^2 & k = r, l = s \end{cases} \quad (\text{A.12})$$

where, for the last covariance, it is assumed that $k \neq l, r \neq s$. $\square$

Proof The proofs of (A.11) and (A.12) are trivial. When $k \neq l$, covariances vanish because of independence. When $k = l$, covariances reduce to variances and the results come from Lemma A.11. For (A.13), the zero covariance is again clear because of independence. Now, let $k = l, r \neq s$. Then

$$\begin{aligned} \text{Cov}(A_k A_l, A_k A_s) &= \text{E}(A_k A_l A_k A_s) - \text{E}(A_k A_l) \text{E}(A_k A_s) \\ &= \text{E}(A_k^2) [\text{E}(A_k)]^2 - [\text{E}(A_k)]^4 = 2\text{tr}(\mathbf{T}\boldsymbol{\Sigma}\mathbf{T})^2 [\text{tr}(\mathbf{T}\boldsymbol{\Sigma}\mathbf{T})]^2 \end{aligned}$$

using Equations (A.6).

When $k = r, l = s$, $\text{Cov}(A_k A_l, A_k A_s) = \text{Var}(A_k A_l)$, where

$$\begin{aligned} \text{Var}(A_k A_l) &= \text{E}(A_k A_l)^2 - [\text{E}(A_k A_l)]^2 = [\text{E}(A_k^2)]^2 - [\text{E}(A_k)]^4 \\ &= 4[\text{tr}(\mathbf{T}\boldsymbol{\Sigma}\mathbf{T})^2]^2 + 4\text{tr}(\mathbf{T}\boldsymbol{\Sigma}\mathbf{T})^2 [\text{tr}(\mathbf{T}\boldsymbol{\Sigma}\mathbf{T})]^2 \end{aligned}$$

since $k \neq l$. Hence the result. $\blacksquare$

For the two sample case, we need only one result, namely $\text{Var}(A_{kl} A_{rs})$, which comes directly from the simple moments (Lemma A.11). Hence, we have the following result, using the independence of the quadratic forms, since $k \neq r, l \neq s$.

$$\text{Var}(A_{kl} A_{rs}) = 4 \left[\text{tr} \{ \mathbf{T} (\boldsymbol{\Sigma}_1 + \boldsymbol{\Sigma}_2) \mathbf{T} \}^2 \right]^2 + 4 \text{tr} \{ \mathbf{T} (\boldsymbol{\Sigma}_1 + \boldsymbol{\Sigma}_2) \mathbf{T} \}^2 [\text{tr} \{ \mathbf{T} (\boldsymbol{\Sigma}_1 + \boldsymbol{\Sigma}_2) \mathbf{T} \}]^2 \quad (\text{A.13})$$

A.2 Moments of Bilinear Forms

Properties of bilinear forms can be conveniently studied using the properties of quadratic forms since a bilinear form can be expressed as a quadratic form. First, the following definition differentiates between a quadratic and a bilinear form. For a comprehensive treatment of bilinear forms, see Mathai et al (1995). A few results can also be found in Mathai (1992) and Searle (1971, Ch. 2).

Definition A.13 Suppose $\mathbf{X}_k$ and $\mathbf{X}_l$ be two random vectors and $\mathbf{A}$ be any symmetric matrix. Then

$$\mathbf{X}_k' \mathbf{A} \mathbf{X}_l = \begin{cases} \text{Quadratic form} & \text{if } k = l \\ \text{Bilinear form} & \text{if } k \neq l. \end{cases}$$

We present two representations of a bilinear form. The first form, based on the eigenvalues of $\mathbf{T}\Sigma$, is an extension of Lemma A.4.

Lemma A.14 (Representation of a Bilinear Form-I) Let $\mathbf{X}_k$ and $\mathbf{X}_l$ be independent and identically distributed random vectors with $E(\mathbf{X}_k) = E(\mathbf{X}_l) = \boldsymbol{\mu}$ and $\text{Cov}(\mathbf{X}_k) = \text{Cov}(\mathbf{X}_l) = \Sigma > 0$. Let $\mathbf{T}$ be any symmetric matrix. Then, the bilinear form $\mathbf{X}_k' \mathbf{T} \mathbf{X}_l$, $k \neq l$, can be represented as

$$\mathbf{X}_k' \mathbf{T} \mathbf{X}_l = \sum_{j=1}^d \lambda_j (U_j + b_j)(W_j + b_j) \quad \boldsymbol{\mu} \neq \mathbf{0} \quad (\text{A.14})$$

$$= \sum_{j=1}^d \lambda_j U_j W_j \quad \boldsymbol{\mu} = \mathbf{0} \quad (\text{A.15})$$

where $\lambda_1, \dots, \lambda_d$ are the eigenvalues of $\mathbf{T}\Sigma$, and $\mathbf{U} = (U_1, \dots, U_d)'$, $\mathbf{V} = (V_1, \dots, V_d)'$ are such that $E(\mathbf{U}) = \mathbf{0} = E(\mathbf{V})$, $\text{Var}(\mathbf{U}) = \mathbf{I} = \text{Var}(\mathbf{V})$ and $\text{Cov}(\mathbf{U}, \mathbf{V}) = \mathbf{0}$. Further $\mathbf{b} = (b_1, \dots, b_d) = \mathbf{P}\Sigma^{\frac{1}{2}}\boldsymbol{\mu}$ with $\mathbf{P}\mathbf{P}' = \mathbf{I}$. $\square$

Next we represent the bilinear form as a quadratic form. This representation is extremely helpful when it is relatively easier to compute the moments of a quadratic form so that they can be converted into those of a bilinear form.

Lemma A.15 (Representation of a Bilinear Form-II) Let $\mathbf{X}_k$ and $\mathbf{X}_l$ be two random vectors with means $\boldsymbol{\mu}_i$ and variances Σ_i , $i = 1, 2$, respectively.

Let $\mathbf{T}$ be any matrix of constants. Then the bilinear form $\mathbf{X}'_k \mathbf{T} \mathbf{X}_l$, $k \neq l$, can be represented as

$$\mathbf{X}'_k \mathbf{T} \mathbf{X}_l = \frac{1}{2} (\mathbf{X}_k \quad \mathbf{X}_l) \begin{pmatrix} \mathbf{0} & \mathbf{T} \\ \mathbf{T} & \mathbf{0} \end{pmatrix} \begin{pmatrix} \mathbf{X}_k \\ \mathbf{X}_l \end{pmatrix} = \frac{1}{2} \mathbf{X}'_m \mathbf{T}^* \mathbf{X}_m \quad (\text{A.16})$$

where $\mathbf{X}_m = \begin{pmatrix} \mathbf{X}_k \\ \mathbf{X}_l \end{pmatrix}$, $\mathbf{T}^* = \begin{pmatrix} \mathbf{0} & \mathbf{T} \\ \mathbf{T} & \mathbf{0} \end{pmatrix}$ and $\mathbf{T}$, $\mathbf{T}^*$ are symmetric matrices.

The mean and variance of the transformed vector $\mathbf{X}_m$ are, respectively,

$$\boldsymbol{\mu}^* = \begin{pmatrix} \boldsymbol{\mu}_1 \\ \boldsymbol{\mu}_2 \end{pmatrix}, \quad \boldsymbol{\Sigma}^* = \begin{pmatrix} \boldsymbol{\Sigma}_1 & \mathbf{0} \\ \mathbf{0} & \boldsymbol{\Sigma}_2 \end{pmatrix}. \quad \square$$

Remark A.16 This representation of a bilinear form is not based on any distributional assumption. For more on this representation, see Mathai et al (1995, Ch. 2). The same representation under the assumption of normality is given by Searle (1971, Ch. 2). Further, the general version of this representation also involves a covariance between $\mathbf{X}_k$ and $\mathbf{X}_l$, say $\boldsymbol{\Sigma}_{12}$ (Searle, 1971, Ch. 2). $\triangle$

All the results of quadratic form given above can now be written for the bilinear form using Equation (A.17). The cumulant generating function of a bilinear form is given in the following lemma using Theorem A.9.

Lemma A.17 (Cumulant Generating Function) For the bilinear form $\mathbf{X}'_k \mathbf{T} \mathbf{X}_l$, as defined above, the r th cumulant is computed as

$$\kappa_r(\mathbf{X}'_k \mathbf{T} \mathbf{X}_l) = \frac{1}{2} (r-1)! [\text{tr}(\mathbf{T} \boldsymbol{\Sigma})^r + r \boldsymbol{\mu}' \mathbf{T} (\boldsymbol{\Sigma} \mathbf{T})^{r-1} \boldsymbol{\mu}], \quad r = 1, 2, \dots \quad (\text{A.17})$$

where $\mathbf{T}$ is the projection matrix, as defined above. $\square$

Corollary A.18 For $\boldsymbol{\mu} = \mathbf{0}$, the cumulant generating function reduces to the following simple form.

$$\kappa_r(\mathbf{X}'_k \mathbf{T} \mathbf{X}_l) = \frac{1}{2} (r-1)! \text{tr}(\mathbf{T} \boldsymbol{\Sigma})^r, \quad r = 1, 2, \dots \quad (\text{A.18}) \quad \square$$

Now, using Corollary A.18, the first four moments of the bilinear form for the one sample case, i.e., $A_{kl} = \mathbf{X}'_k \mathbf{T} \mathbf{X}_l = \mathbf{Y}'_k \mathbf{Y}_l$, $k \neq l$, are computed as following.

Lemma A.19 Under $H_0 : \mathbf{T}\boldsymbol{\mu} = \mathbf{0}$, the first four moments of the bilinear form $A_{kl} = \mathbf{Y}'_k \mathbf{Y}_l$ are

$$\left. \begin{aligned} \mu'_1 &= 0 \\ \mu'_2 &= \text{tr}(\mathbf{T}\boldsymbol{\Sigma})^2 = \text{Var}(A_{kl}) \\ \mu'_3 &= 0 \\ \mu'_4 &= 6\text{tr}(\mathbf{T}\boldsymbol{\Sigma})^4 + 3[\text{tr}(\mathbf{T}\boldsymbol{\Sigma})^2]^2. \end{aligned} \right\} \quad (\text{A.19}) \quad \square$$

It is trivial to extend Lemma A.19 for the corresponding moments of the two sample bilinear forms, i.e., $A_{klrs}^{AB} = (\mathbf{X}_{1k} - \mathbf{X}_{2l})'\mathbf{T}(\mathbf{X}_{1r} - \mathbf{X}_{2s})$ and $A_{klrs}^B = (\mathbf{X}_{1k} + \mathbf{X}_{2l})'\mathbf{T}(\mathbf{X}_{1r} + \mathbf{X}_{2s})$. Clearly, the moments will again be same for both interaction and time effects, as those of the quadratic forms.

The required results for the joint moments of the bilinear forms, for the one sample case, are computed in the following theorem.

Theorem A.20 Suppose A_{kl} , $k \neq l$, be the bilinear form for the one sample case as defined above. Then, under $H_0 : \mathbf{T}\boldsymbol{\mu} = \mathbf{0}$, we have the following identities.

$$\text{Cov}(A_{kl}, A_{rs}) = \begin{cases} 0 & k \neq r, l \neq s; k = r, l \neq s; k \neq r, l = s \\ \text{tr}(\mathbf{T}\boldsymbol{\Sigma}\mathbf{T})^2 & k = r, l = s \end{cases} \quad (\text{A.20})$$

$$\text{Cov}(A_{kl}^2, A_{rs}^2) = \begin{cases} 0 & k \neq r, l \neq s \\ 2\text{tr}(\mathbf{T}\boldsymbol{\Sigma}\mathbf{T})^4 & k = r, l \neq s; k \neq r, l = s \\ 6\text{tr}(\mathbf{T}\boldsymbol{\Sigma}\mathbf{T})^4 + 2[\text{tr}(\mathbf{T}\boldsymbol{\Sigma}\mathbf{T})^2]^2 & k = r, l = s \end{cases} \quad (\text{A.21})$$

$$\text{Cov}(A_{kl}, A_m) = 0, \quad k = m \neq l; l = m \neq k \quad (\text{A.22})$$

□

Proof For the first two covariances, the case when $k \neq r, l \neq s$ is trivial due to the independence of the bilinear forms involved. Now, let $k = r, l \neq s$ for the first covariance. Then, with $E(A_{kl}) = 0$, we have

$$\begin{aligned} \text{Cov}(A_{kl}, A_{ks}) &= E(A_{kl}A_{ks}) = E(\mathbf{Y}'_k \mathbf{Y}_l \mathbf{Y}'_k \mathbf{Y}_s) \\ &= E(\mathbf{Y}'_k \mathbf{Y}_l \mathbf{Y}'_s \mathbf{Y}_k) = E[\text{tr}(\mathbf{Y}_k \mathbf{Y}'_k \mathbf{Y}_l \mathbf{Y}'_s)] = 0 \end{aligned}$$

since $k \neq l \neq s$ and $E(\mathbf{Y}_k) = \mathbf{0}$, under H_0 . We get the same result when $k \neq r, l = s$. Finally, when $k = r, l = s$, $\text{Cov}(A_{kl}, A_{rs}) = \text{Var}(A_{kl})$ and the

result immediately follows from Lemma A.19.

The last part of the second covariance is also clear using the simple moments, since $\text{Cov}(A_{kl}^2, A_{rs}^2) = \text{Var}(A_{kl}^2)$. For the middle part, we use Lemma A.14. We know that, under $H_0 : \mathbf{T}\boldsymbol{\mu} = \mathbf{0}$, $E(\mathbf{Y}_k) = \mathbf{0} = E(\mathbf{Y}_l)$, $\text{Var}(\mathbf{Y}_k) = \mathbf{T}\boldsymbol{\Sigma}\mathbf{T} = \text{Var}(\mathbf{Y}_l)$ and $\text{Cov}(\mathbf{Y}_k, \mathbf{Y}_l) = \mathbf{0}$, $k \neq l$. Then we can write $\mathbf{X}'_k \mathbf{T} \mathbf{X}_l = \mathbf{Y}'_k \mathbf{Y}_l = \sum_{j=1}^d \lambda_j U_j V_j$ where the λ_j are the eigenvalues of $\mathbf{T}\boldsymbol{\Sigma}$, and U_j and V_j are as defined in Lemma A.14. Further, $E(U_j^2) = 1$, and, with $U_j \sim N(0, 1)$, $E(U_j^4) = 3$. Then, for $k = r, l \neq s$, we have

$$\text{Cov}(A_{kl}^2, A_{ks}^2) = E(A_{kl}^2 A_{ks}^2) - [E(A_{kl}^2)]^2,$$

where $E(A_{kl}^2) = \text{tr}(\mathbf{T}\boldsymbol{\Sigma})^2$, from Theorem A.19, and

$$\begin{aligned} E(A_{kl}^2 A_{ks}^2) &= E \left[\left(\sum_l \lambda_l U_l V_l \right)^2 \left(\sum_m \lambda_m U_m W_m \right)^2 \right] \\ &= E \left[\sum_l \lambda_l^4 U_l^4 V_l^2 W_l^2 + \underbrace{\sum_l \sum_{m \neq l} \lambda_l^2 \lambda_m^2 U_l^2 V_l^2 U_m^2 W_m^2}_{l \neq m} \right] \\ &= 3 \sum_l \lambda_l^4 + \underbrace{\sum_l \sum_{m \neq l} \lambda_l^2 \lambda_m^2}_{l \neq m} \\ &= 2 \sum_l \lambda_l^4 + \sum_l \sum_m \lambda_l^2 \lambda_m^2 = 2 \text{tr}(\mathbf{T}\boldsymbol{\Sigma})^4 + [\text{tr}(\mathbf{T}\boldsymbol{\Sigma})^2]^2, \end{aligned}$$

so that $\text{Cov}(A_{kl}^2, A_{ks}^2) = 2 \text{tr}(\mathbf{T}\boldsymbol{\Sigma})^4$.

The last identity is the covariance of a bilinear form and a quadratic form. Let, for example, $k = m \neq l$. Then,

$$\begin{aligned} \text{Cov}(A_{kl}, A_k) &= E(A_{kl} A_k) = E(\mathbf{Y}'_k \mathbf{Y}_l \mathbf{Y}'_k \mathbf{Y}_k) \\ &= E(\mathbf{Y}'_l \mathbf{Y}_k \mathbf{Y}'_k \mathbf{Y}_k) = E(\mathbf{Y}'_l) E(\mathbf{Y}_k \mathbf{Y}'_k \mathbf{Y}_k) = 0 \end{aligned}$$

since $E(A_{kl}) = 0$ and $E(\mathbf{Y}'_l) = 0$.

This completes the proof of the theorem. ■

Finally, for the two sample case, we need only one result, namely $\text{Var}(A_{klrs}^2)$. Using again the simple moments from Lemma A.19, we have

$$\text{Var}(A_{klrs}^2) = 6\text{tr}\{\mathbf{T}(\boldsymbol{\Sigma}_1 + \boldsymbol{\Sigma}_2)\mathbf{T}\}^4 + 2\left[\text{tr}\{\mathbf{T}(\boldsymbol{\Sigma}_1 + \boldsymbol{\Sigma}_2)\mathbf{T}\}^2\right]^2, \quad (\text{A.23})$$

where $k \neq r, l \neq s$.

Appendix B

Some Miscellaneous Results

As mentioned in the introduction, the distribution of the modified ANOVA-type statistics for both one- and two-sample cases is approximated with the corresponding χ_f^2/f -distribution. This approximation is approached by comparing the first two moments of the sampling distributions of the derived test statistics with those of the χ_f^2/f distribution (Lindsay and Basak, 2000). Clearly, the first two moments of χ_f^2/f distribution are 1 and $\frac{2}{f}$. The moments of the sampling distributions are derived using the approximation formulas for the mean and variance of the function (here, ratio) of two random dependent variables. The approximation formulas are based on the bivariate Taylor series expansion (Finney et al, 2001, pp 965-967). In the following, we state these formulas for the vector (U, V) , which can be used for both one and two sample cases, using, for example $U = Q_n$ and $V = B_0$. For details, see Stuart and Ord (1994, Ch. 10), Casella and Berger (2002, Ch. 5) and Rohatgi (1984, Ch. 5).

Theorem B.1 (Approximation Formulas) *Let U and V be the random variables with their means $E(U)$, $E(V)$, variances $Var(U)$, $Var(V)$, and the covariance $Cov(U, V)$. Then, the first two moments of the ratio, $\frac{U}{V}$, are computed as following.*

$$E\left(\frac{U}{V}\right) \approx \frac{E(U)}{E(V)} \quad (\text{B.1})$$

$$Var\left(\frac{U}{V}\right) \approx \frac{[E(U)]^2}{[E(V)]^2} \left(\frac{Var(U)}{[E(U)]^2} + \frac{Var(V)}{[E(V)]^2} - 2 \frac{Cov(U, V)}{[E(U)][E(V)]} \right) \quad (\text{B.2})$$

□

Remark B.2 Note that the expression within parentheses of Equation (B.2) is the sum of the two coefficients of variation minus twice the joint coefficient of variation of the two variables. $\triangle$

Remark B.3 The only condition for the use of Equations (B.1) and (B.2) is that the variances and covariances of the random variables are of order $O(n^{-r})$, $r > 0$. Then, a remainder of order $O(n^{-r})$ for Equation (B.1) and of order $o(n^{-r})$ for Equation (B.2) can be ignored, for large sample approximation (Stuart and Ord, 1994, p 350); see also Casella and Berger (2002, p 245). Stuart and Ord (1994) further point out that, in practice, usually, $r = 1$. We have proved in the derivations of the modified ANOVA-type statistic, that, for all cases, $\text{Var}(Q_n)$, $\text{Var}(B_0)$ and $\text{Cov}(Q_n, B_0)$ tend to zero, asymptotically. $\triangle$

For details and some interesting examples for the following, and other, inequalities, see Casella and Berger (2002, Ch. 5).

Theorem B.4 (Chebychev's Inequality) *Let X be a random variable and let $g(x)$ be a nonnegative function. Then,*

$$P[g(X) \geq r] \leq \frac{E[g(X)]}{r} \quad (\text{B.3})$$

where $r > 0$. $\square$

Theorem B.5 (Cauchy-Schwarz's Inequality) *For any two random variables X and Y ,*

$$|E(XY)| \leq E|XY| \leq (E|X|^2)^{1/2} (E|Y|^2)^{1/2}. \quad (\text{B.4})$$

If μ_X , μ_Y , σ_X^2 and σ_Y^2 , are the means and variances of X and Y , respectively, then using the above inequality, we have

$$|E(X - \mu_X)(Y - \mu_Y)| \leq \{E(X - \mu_X)^2\}^{1/2} \{E(Y - \mu_Y)^2\}^{1/2} \quad (\text{B.5})$$

Squaring on both sides we get $[Cov(X, Y)]^2 \leq \sigma_X^2 \sigma_Y^2$. This result is usually known as Covariance Inequality. $\square$

In the following, we define the three covariance structures typically used for the simulations to assess the small sample behavior of the test statistics in both one and two sample cases.

Definition B.6 (Compound Symmetry) Let $\mathbf{X} = (X_1, \dots, X_d)'$ be a vector of random variables. Then the covariance structure is called compound symmetric if $\text{Var}(X_k) = \sigma^2$ and $\text{Cov}(X_k, X_l) = \kappa$, $k \neq l$, $\forall k, l = 1, \dots, d$, where $\sigma^2 > 0$ and $\kappa \in \mathcal{R}^d$.

In matrix form, we can write $\Sigma = \sigma^2 \mathbf{I} + \kappa(\mathbf{J} - \mathbf{I})$, where $\mathbf{I}$ is the identity matrix and $\mathbf{J}$ is the matrix of 1s. ||

Definition B.7 (Autoregressive) Let $\mathbf{X} = (X_1, \dots, X_d)'$ be a vector of random variables. Then the covariance structure is called autoregressive if $\text{Cov}(X_k, X_l) = \sigma^2 \rho^{|k-l|}$, $\forall k, l = 1, \dots, d$, where $\sigma^2 > 0$ and $-1 \leq \rho \leq 1$. ||

Definition B.8 (Unstructured) The unstructured covariance matrix is defined as $\Sigma = (\sigma_{ij})_{i,j=1}^d$ where $\sigma_{ij} = 1(1)d$, $i = j$, and $\rho_{ij} = \frac{i-1}{d}$, $i > j$. ||

Definition B.9 (Affine Transformation) Let $\mathbf{A}$ be a $d \times d$ matrix and $\mathbf{c}$ be a $d \times 1$ vector. A transformation $\mathbf{g}$ is an affine transformation of $\mathcal{R}^d$ if $\mathbf{g}(\mathbf{x}) = \mathbf{A}\mathbf{x} + \mathbf{c}$. If $\mathbf{c} = \mathbf{0}$, then $\mathbf{g}$ is the linear transformation. (Bickel and Doksum, 2001, p 487) ||

Definition B.10 (Orthogonal Transformation) Let $\mathbf{g}(\mathbf{x})$ be a linear transformation, as defined in Definition B.9 and let $\mathbf{A}$ be a non-singular matrix. Any such linear transformation is orthogonal if it preserves lengths, distances and angles, and satisfies the condition that $\mathbf{A}\mathbf{A}' = \mathbf{I} = \mathbf{A}'\mathbf{A}$. The matrix $\mathbf{A}$ is called the orthogonal matrix. (Timm, 2002, p 62) ||

For the proof of the following lemma, see Zhang (1999, p 166).

Lemma B.11 (A Property of the Trace of PSD Matrices) Let $\mathbf{A}$ and $\mathbf{B}$ be positive semi-definite matrices of the same order. Then,

$$\text{tr}(\mathbf{A}\mathbf{B}) \leq \text{tr}(\mathbf{A})\text{tr}(\mathbf{B}). \quad (\text{B.6}) \quad \square$$

Remark B.12 Lemma B.11 has been extensively used to prove the uniform boundedness of the estimators. Note that, with $\mathbf{B} = \mathbf{A}$, Lemma B.11 immediately gives $\text{tr}(\mathbf{A}^2) \leq [\text{tr}(\mathbf{A})]^2$, which, replacing $\mathbf{A}$ with $\mathbf{A}^2$, further gives $\text{tr}(\mathbf{A}^4) \leq [\text{tr}(\mathbf{A}^2)]^2$ and so on. Actually, the general inequality can be expressed as $0 \leq \text{tr}(\mathbf{A}\mathbf{B})^n \leq \text{tr}(\mathbf{A})^n \text{tr}(\mathbf{B})^n$, for any integer n . Now, since $\text{tr}(\mathbf{A})^s = \sum_j \lambda_j^s$, $s = 1, 2, \dots$, where $\lambda_j \geq 0$ are the eigenvalues of $\mathbf{A}$, therefore, the inequalities can also be expressed in terms of eigenvalues. △

Lemma B.13 Let $\mathbf{X}$ and $\mathbf{Y}$ be any two real vectors and $\mathbf{A} > 0$. Then

$$(\mathbf{X}'\mathbf{A}\mathbf{Y})^2 \leq (\mathbf{X}'\mathbf{A}\mathbf{X})(\mathbf{Y}'\mathbf{A}\mathbf{Y}). \quad (\text{B.7})$$

(Ravishanker and Dey, 2002, p 31)

□

Appendix C

Data Sets

C.1 α -Amylase in Saliva

Table C.1: MEASUREMENTS ON α -AMYLASE IN SALIVA

Proband	Monday				Thursday			
	8h	12h	17h	21h	8h	12h	17h	21h
1	146.8	167.0	107.2	161.8	90.8	151.6	123.0	142.8
2	818.2	1314.2	1578.8	932.5	378.8	759.5	1881.2	572.6
3	394.4	1157.4	585.2	629.2	171.0	538.4	729.8	412.1
4	100.2	140.4	234.4	244.8	121.6	154.6	221.8	170.6
5	169.8	99.9	184.2	168.8	103.0	170.0	342.0	162.2
6	107.2	262.8	198.4	465.1	178.8	312.6	261.6	450.5
7	272.0	551.2	265.2	453.2	133.4	560.4	977.9	402.0
8	51.8	144.4	125.4	203.8	122.2	71.4	434.9	191.2
9	273.6	351.6	510.0	354.0	403.0	665.4	420.4	566.0
10	367.2	435.6	783.3	523.1	221.8	601.2	1028.5	713.4
11	519.2	264.6	321.4	1433.8	137.2	345.6	884.9	331.8
12	88.6	135.0	88.6	86.2	164.2	190.4	301.0	173.2
13	218.0	109.2	167.6	179.4	162.8	185.6	193.6	183.2
14	117.2	151.0	150.0	218.0	178.2	151.0	165.2	170.0

C.2 Body Weight of Wistar Rats

Table C.2: BODY WEIGHTS OF WISTAR RATS

Rat	Week																					
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22
<i>(Drug)</i>																						
1	342	345	348	352	353	355	345	350	354	360	366	371	377	381	378	394	393	396	403	400	400	402
2	337	341	340	344	348	355	355	352	357	355	362	362	366	368	376	384	376	379	416	384	384	385
3	314	320	321	323	325	331	332	332	334	334	340	341	341	345	344	359	356	366	371	371	376	382
4	365	370	370	375	379	381	378	383	385	390	397	398	399	404	405	412	409	419	429	422	424	426
5	319	323	321	325	326	334	334	341	337	341	349	346	352	357	359	369	366	368	369	375	376	381
6	401	404	404	408	413	418	411	414	421	428	435	439	448	457	464	478	476	481	481	489	490	496
7	365	365	368	371	375	379	371	370	372	373	380	383	397	397	402	408	406	408	414	422	425	435
8	328	331	333	333	337	340	345	342	344	347	352	347	357	357	358	370	366	370	371	378	379	382
9	353	358	361	362	366	363	363	368	373	371	376	374	386	386	391	394	389	397	400	407	411	417
10	302	307	306	306	311	313	315	314	288	301	303	305	311	311	312	318	315	320	323	326	325	328
<i>(Placebo)</i>																						
11	363	369	371	371	370	378	369	374	378	376	392	391	394	402	405	413	403	417	416	422	429	430
12	388	396	399	400	398	401	405	408	411	416	419	420	422	433	433	438	433	441	446	451	454	461
13	334	338	347	346	351	356	355	356	356	355	361	369	374	374	381	386	383	389	397	388	393	392
14	363	361	364	367	365	366	373	372	366	377	379	381	388	391	393	404	397	407	406	413	416	422
15	379	381	385	389	389	391	384	393	388	389	396	399	399	405	414	417	423	431	436	447	446	448
16	400	408	414	418	420	426	423	412	429	427	435	437	439	448	455	467	463	469	472	484	486	488
17	324	323	325	335	331	339	342	339	340	344	347	350	348	361	361	375	369	377	378	385	383	392
18	340	345	349	345	343	347	352	355	359	363	363	362	363	368	371	376	373	378	384	386	392	397
19	350	355	358	364	360	366	360	364	365	364	373	369	371	379	380	385	385	388	395	397	398	402
20	316	321	321	325	329	330	333	334	332	333	343	345	341	357	360	364	370	370	374	375	380	384

C.3 Sleep Disorder

Table C.3: PROSTAGLANDIN-D SYNTHASE ON SLEEP PROBANDS

Proband	Sleep Type/Measurement Time																	
	Normal Sleep						Total Sleep deprivation											
	24h	4h	8h	12h	16h	20h	24h	4h	8h	12h	16h	20h	24h	4h	8h	12h	16h	20h
(Men)																		
1	0.516	0.639	0.597	0.564	0.540	0.530	0.491	0.573	0.516	0.554	0.625	0.540	0.554	0.625	0.477	0.412	0.437	0.402
2	0.525	0.592	0.602	0.578	0.639	0.597	0.606	0.611	0.625	0.583	0.625	0.588	0.621	0.644	0.649	0.630	0.611	0.606
3	0.583	0.611	0.676	0.585	0.573	0.545	0.545	0.616	0.621	0.521	0.535	0.530	0.621	0.630	0.621	0.673	0.621	0.501
4	0.554	0.516	0.496	0.437	0.472	0.442	0.501	0.467	0.511	0.506	0.516	0.496	0.588	0.545	0.564	0.516	0.545	0.496
5	0.564	0.578	0.238	0.530	0.477	0.472	0.482	0.573	0.491	0.559	0.501	0.496	0.442	0.386	0.432	0.432	0.386	0.371
6	0.462	0.422	0.442	0.457	0.437	0.422	0.653	0.602	0.452	0.564	0.611	0.667	0.447	0.602	0.588	0.530	0.554	0.496
7	0.644	0.709	0.686	0.525	0.690	0.741	0.588	0.639	0.653	0.588	0.602	0.564	0.649	0.663	0.686	0.559	0.588	0.639
8	0.592	0.611	0.588	0.559	0.516	0.501	0.447	0.487	0.477	0.511	0.516	0.491	0.597	0.639	0.649	0.621	0.578	0.597
9	0.452	0.501	0.540	0.487	0.462	0.422	0.472	0.482	0.472	0.496	0.540	0.511	0.506	0.535	0.559	0.491	0.447	0.467
10	0.487	0.564	0.564	0.550	0.477	0.501	0.213	0.371	0.376	0.366	0.407	0.381	0.713	0.699	0.597	0.516	0.472	0.496
(Women)																		
1	0.277	0.457	0.371	0.218	0.221	0.318	0.437	0.235	0.437	0.462	0.432	0.381	0.402	0.447	0.477	0.407	0.427	0.407
2	0.554	0.592	0.578	0.545	0.482	0.472	0.457	0.573	0.501	0.530	0.506	0.525	0.559	0.625	0.813	0.535	0.559	0.564
3	0.422	0.467	0.447	0.422	0.355	0.397	0.397	0.422	0.392	0.376	0.417	0.417	0.467	0.496	0.630	0.578	0.559	0.554
4	0.516	0.639	0.597	0.564	0.540	0.530	0.491	0.573	0.516	0.554	0.625	0.540	0.554	0.625	0.477	0.412	0.437	0.402
5	0.630	0.630	0.644	0.602	0.588	0.597	0.540	0.606	0.530	0.597	0.616	0.611	0.554	0.550	0.588	0.530	0.545	0.501
6	0.525	0.592	0.602	0.578	0.639	0.597	0.606	0.611	0.625	0.583	0.625	0.588	0.621	0.644	0.649	0.630	0.611	0.606
7	0.583	0.611	0.676	0.585	0.573	0.545	0.545	0.616	0.621	0.521	0.535	0.530	0.621	0.630	0.621	0.673	0.621	0.501
8	0.630	0.597	0.592	0.732	0.545	0.559	0.559	0.397	0.606	0.750	0.583	0.476	0.546	0.597	0.653	0.653	0.517	0.574
9	0.703	0.719	0.757	0.757	0.563	0.680	0.675	0.636	0.675	0.680	0.724	0.653	0.664	0.669	0.658	0.697	0.636	0.619
10	0.517	0.550	0.546	0.499	0.493	0.517	0.416	0.476	0.505	0.505	0.563	0.499	0.488	0.523	0.493	0.511	0.528	0.523

Bibliography

- Ahmad, M. Rauf, C. Werner and E. Brunner (2008). Analysis of high dimensional repeated measures designs: The one-sample case. *Computational Statistics & Data Analysis*. Accepted.
- Akritis, M. and S. Arnold (2000). Asymptotics for analysis of variance when the number of levels is large. *Journal of the American Statistical Association*, **95**(449), 212-226.
- Anderson, T. W. (1984). *An introduction to multivariate statistical analysis*. 2nd Edition. Wiley, New York.
- Arnold, Steven F. (1981). *The theory of linear models and multivariate analysis*. Wiley, New York.
- Bai, Z. and H. Saranadasa (1996). Effect of high dimension: By an example of a two sample problem. *Statistica Sinica*, **6**, 311-329.
- Bathke, Arne (2002). ANOVA for a large number of treatments. *Mathematical Methods of Statistics*, **11**(1), 118-132.
- Bathke, Arne and Solomon W. Harrar (2006a). Testing treatment effects in multivariate factorial designs with many factor levels. *Proceedings of the 2006 Joint Statistical Meeting, Alexandria, VA, American Statistical Association*, 1664-1667.
- Bathke, Arne and Solomon W. Harrar (2006b). Nonparametric methods in multivariate factorial designs for large number of factor levels. *Journal of Statistical Planning and Inference*. **138**(3), 588-610.
- Bathke, Arne, Solomon W. Harrar and M. Rauf Ahmad (2008). Some contributions to the analysis of multivariate data. *Biometrical Journal*. Revision Submitted.

- Bickel, P. J. and K. A. Doksum (2001). *Mathematical statistics: Basic ideas and selected topics*. Vol. 1, 2nd Edition, Prentice Hall, New Jersey.
- Birke, M. and H. Dette (2005). A note on testing the covariance matrix for large dimension. *Statistics and Probability Letters*, **74**, 281-289.
- Box, G. E. P. (1954a). Some theorems on quadratic forms applied in the study of analysis of variance problems, I: Effect of inequality of variance in the one-way classification. *Annals of Mathematical Statistics*, **25**(2), 290-302.
- Box, G. E. P. (1954b). Some theorems on quadratic forms applied in the study of analysis of variance problems, II: Effect of inequality of variance and of correlation between errors in the two-way classification. *Annals of Mathematical Statistics*, **25**(3), 484-498.
- Brunner, E., H. Dette and A. Munk (1997). Box-type approximations in nonparametric factorial designs. *Journal of the American Statistical Association*, **92**(440), 1494-1502.
- Brunner, E., U. Munzel and M. Puri (1999). Rank score tests in factorial designs with repeated measures. *Journal of Multivariate Analysis*, **70**, 286-317.
- Brunner, E. (2001). Asymptotic and approximate analysis of repeated measures designs under heteroscedasticity. In: *Mathematical Statistics with Applications in Biometry*, Eds.: Kunert, J. and G. Trenkler. Josef Eul Verlag, Lohmar.
- Brunner, E., S. Domhof and Langer, F. (2002). *Nonparametric analysis of longitudinal data in factorial experiments*. Wiley, New York
- Campbell, S. L. and Meyer, C. D. (1979). *Generalized inverses of linear transformations*. Dover Publications, New York
- Casella, G. and R. L. Berger (2002). *Statistical inference*. 2nd Edition, Duxbury Press, Pacific Grove, CA.
- Collier, R. O., F. B. Baker, G. K. Mandeville and T. F. Hayes (1967). Estimates of test size for several test procedures based on conventional variance ratios in the repeated measures design. *Psychometrika*, **32**(3), 339-353.

- Crowder, M. J. and D. J. Hand (1990). *Analysis of repeated measures*. Chapman & Hall, London.
- Davidian, M. and D. M. Giltinan (1995). *Nonlinear models for repeated measures data*. Chapman & Hall, London.
- Davis, C. S. (2002). *Statistical methods for the analysis of repeated measurements*. Springer, New York.
- Depmster, A. P. (1958). A high dimensional two sample significance test. *Annals of Mathematical Statistics*, **29**(4), 995-1010.
- Diggle, P. J., P. Heagerty, K-Y. Liang and S. Zeger (2002). *Analysis of longitudinal data*. Oxford University Press, Oxford.
- Finney, Ross L., Maurice D. Weir and Frank R. Giordano (2001). *Thomas' calculus: Early transcendentals*. 10th Edition. Addison Wesley, Boston.
- Fujikoshi, Y., T. Himeno and H. Wakaki (2004). Asymptotic results of a high dimensional MANOVA test and power comparison when the dimension is large compared to the sample size. *Journal of Japan Statistics Society*, **1**, 19-26.
- Geissser, S. and W. Greenhouse (1958). An extension of Box's results on the use of F distribution in multivariate analysis. *Annals of Mathenatical Statistics*, **29**(3), 885-891.
- Graybill, F. (1976). *Theory and application of linear models*. Duxbury Press, MA.
- Greenhouse, W. and S. Geissser (1959). On methods in the analysis of profile data. *Psychometrika*, **24**(2), 95-112.
- Hand, D. J. and C. C. Taylor (1987). *Multivariate analysis of variance and repeated measures*. Chapman and Hall, London.
- Harrar, Solomon W. and A. Bathke (2006). Nonparametric methods for unbalanced multivariate data and many factor levels. *Journal of Multivariate Analysis*. Submitted.
- Hinkelmann, K. and O. kempthorne (1994). *Design and analysis of experiments*. Vol. I: Introduction to experimental design. Wiley, New York.

- Hinkelmann, K. and O. Kempthorne (2005). *Design and analysis of experiments*. Vol. II: Advanced experimental design. Wiley, New York.
- Huynh, H. (19778). Some approximate tests for repeated measures designs. *Psychometrika*, **43**(2), 161-175.
- Huynh, H. and L. S. Feldt (1970). Conditions under which mean square ratios in repeated measurements designs have exact F -distributions. *Journal of the American Statistical Association*, **65**(332), 1582-1589.
- Huynh, H. and L. S. Feldt (1976). Estimation of the Box correction for the degrees of freedom from sample data in randomized block and split-plot designs. *Journal of Educational Statistics*, **1**(1), 69-82.
- Huynh, H. and G. K. Mandeville (1979). Validity conditions in repeated measures designs. *Psychological Bulletin*, **86**(5), 964-973.
- Johnson, R. A. and D. W. Wichern (2002). *Applied multivariate statistical analysis*. 5th Edition, Prentice Hall, New Jersey.
- Jordan, W., H. Tumani, S. Cohrs, et al (2004). Prostaglandin D synthase (β -trace) in healthy human sleep. *Sleep*, **27**(5), 867-874.
- Keselman, H. J. (1998). Testing treatment effects in repeated measures designs: An update for psychophysiological researchers. *Psychophysiology*, **35**, 470-478.
- Keselman, H. J., J. Algina, R. R. Wilcox and R. K. Kowalchuk (2000). Testing repeated measures hypotheses when covariance matrices are heterogeneous: Revisiting the robustness of the Welch-James test again. *Educational and Psychological Measurement*, **60**(6), 925-938.
- Keselman, H. J., J. C. Rogan, J. L. Mendoza and L. J. Breen (1980). Testing the validity conditions of repeated measures F tests. *Psychological Bulletin*, **87**(3), 479-481.
- Kirk, Roger E. (1982). *Experimental design*. 2nd Edition. Brooks/Cole Pub. Co., CA.
- Läuter, Jürgen (2004). Two new multivariate tests, in particular for a high dimension. *Acta et Commentationes Universitatis Tartuensis de Mathematica*, **8**, 179-186.

- Läuter, J., E. Glimm and S. Kropf (1998). Multivariate tests based on left-spherically distributed linear scores. *The Annals of Statistics*, **26**, 1972-1988. Corrections: **27**, 1441.
- Ledoit, O. and M. Wolf (2002). Some hypothesis tests for the covariance matrix when the dimension is large compared to the sample size. *The Annals of Statistics*, **30**(4), 1081-1102.
- Ledoit, O. and M. Wolf (2004). A well-conditioned estimator for large-dimensional covariance matrix. *Journal of Multivariate Analysis*, **88**, 365-411.
- Lehmann, E. L. (1959). *Testing statistical hypothesis*. Wiley, New York.
- Lehmann, E. L. (1999). *Elements of large-sample theory*. Springer, New York.
- Lindsay, B. G. and P. Basak (2000). Moments determine the tail of a distribution (but not much else). *The American Statistician*. **54**(4), 248-251.
- Lindsey, J. K. (1999). *Models for repeated measurements*. Oxford University Press, New York.
- Mathai, A. M. (1992). On bilinear forms in normal variables. *Annals of the Institute of Statistical Mathematics*, **44**(4), 769-779.
- Mathai, A. M. and S. B. Provost (1992). *Quadratic forms in random variables*. Marcel Dekker Inc., New York.
- Mathai, A. M., S. B. Provost and T. Hayakawa (1995). *Bilinear forms and zonal polynomials*. Lecture Notes in Statistics, No. 102. Springer, New York.
- Mauchley, J. W. (1940). Significance test for sphericity of a normal n -variate distribution. *Annals of Mathematical Statistics*, **11**, 204-209.
- Maxwell, Scott E. and Harold D. Delaney (2004). *Designing experiments and analyzing data: A model comparison perspective*. 2nd Edition, Lawrence Earlbaum Associates, New Jersey.
- Mendoza, Jorg L. (1980). A significance test for multisample sphericity. *Psychometrika*, **45**(4), 495-498.

- Mendoza, J. L., L. E. Toothaker and B. R. Crain (1976). Necessary and sufficient conditions for F ratios in the $L \times J \times K$ factorial design with two repeated factors. *Journal of the American Statistical Association*, **71**(356), 992-993.
- Morrison, Donald F. (1967). *Multivariate statistical methods*. McGraw-Hill Book Co., New York.
- Muirhead, Robb J. (1982). *Aspects of multivariate statistical theory*. Wiley, New York.
- Muller, K. E. and C. N. Barton (1989). Approximate power for repeated measures ANOVA lacking sphericity. *Journal of the American Statistical Association*, **84**(406), 549-555. Corections: (1991), **86**(413), 255-256.
- Patnaik, P. B. (1949). The non-central χ^2 -and F-distributions and their applications. *Biometrika*, **36**, 202-232.
- Rao, C. R. (1973). *Linear statistical inference and its applications*. 2nd Edition, Wiley, New York
- Ravishanker, Nalini and Dipak K. Dey (2002). *A first course in linear model theory*. Chapman & Hall, Boca Raton.
- Rencher, Alvin, C. (1995). *Methods of multivariate analysis*. Wiley, New York.
- Rohatgi, Vijay K. (1984). *Statistical inference*. Wiley, New York.
- Schäfer, J. and K. Strimmer (2005). A shrinkage approach to large-scale covariance matrix estimation and implications for functional genomics. *Statistical Applications in Genetics and Molecular Biology*, **4**(1), 1-32.
- Schott, James R. (2007). Some high dimensional tests for a one-way MANOVA. *Journal of Multivariate Analysis*, **98**, 1825-1839.
- Searle, S. R. (1971). *Linear models*. Wiley, New York.
- Serfling, Robert J. (1980). *Approximate theorems of mathematical statistics*. Wiley, New York
- Srivastava, M. S. (2002). *Methods of multivariate statistics*. Wiley, New York.

- Srivastava, M. S. (2005). Some tests concerning the covariance matrix in high dimensional data. *Journal of Japan Statistical Association*, **35**(2), 251-272.
- Srivastava, M. S. (2007). Multivariate theory for analyzing high dimensional data. *Journal of Japan Statistical Association*, **37**(1), 53-86.
- Srivastava, M. S. and M. Du (2007). A test for the mean vector with fewer observations than the dimension. *Journal of Multivariate Analysis*, **99**, 386-402.
- Srivastava, M. S. and Y. Fujikoshi (2006). Multivariate analysis of variance with fewer observations than the dimension. *Journal of Multivariate Analysis*, **97**, 1927-1940.
- Stuart, A. and J. K. Ord (1994). *Kendall's advanced theory of statistics*, Vol. 1: Distribution theory. 6th Edition. Arnold pubs., London.
- Stuart, A., J. K. Ord and S. Arnold (1999). *Kendall's advanced theory of statistics*, Vol. 2A: Classical inference and the linear model. 6th Edition. Arnold pubs., London.
- Tai-Kai, Fang and Yao-Ting, Zhang (1990). *Generalized multivariate analysis*. Springer, New York.
- Tian, T. and R. Wilcox (2007). A comparison of two rank tests for repeated measures designs. *Journal of Modern Applied Statistical Methods*, **6**(1), 331-335.
- Timm, Neil H. (2002). *Applied multivariate analysis*. Springer, New York.
- Vonesh, E. F. and V. M. Chinchilli (1996). *Linear and nonlinear models for the analysis of repeated measurements*. Marcel Dekker, New York.
- Winer, B. J., D. R. Brown and K. M. Michels (1991). *Statistical principles in experimental design*. 3rd Edition. McGraw-Hill, New York.
- Zhang, F. (1999). *Matrix theory. Basic results and techniques*. Springer, New York.

Index

- affine invariance, 60
- affine transformation, 79
- ANOVA-type statistic, 11, 13
 - general formulation, 14
 - modified one-sample, 20
 - modified two-sample
 - group effect, 54
 - interaction effect, 33
 - time effect, 47
- approximation formulas, 77
- bilinear form
 - as quadratic form, 72
 - cgf, 73
 - general representation, 72
- estimators
 - consistency, 17
 - one-sample, 18
 - two-sample
 - group effect, 52
 - interaction effect, 33
 - time effect, 48
- high dimensional data
 - classical methods of analysis, 16
 - problems of analysis, 5
- Hotelling's T^2 statistic, 5, 13
- inequality
 - Cauchy-Schwarz, 78
 - Chebychev, 78
 - covariance, 78
 - for psd matrices, 79
- matrix
 - autoregressive, 79
 - centering, 30
 - circular, 3
 - compound symmetric, 79
 - general linear hypothesis, 12, 30
 - spherical, 3
 - unstructured, 79
- orthogonal transformation, 79
- quadratic form
 - cgf, 69
 - general representation, 68
 - mgf, 68
 - normal representation, 14
- repeated measures design, 1
 - classical analysis, 2
- Wald-type statistic, 12

Curriculum Vitae

08.03.1973	Born in Faisalabad, Pakistan
1985-1987	Secondary School Certificate Examination (Matriculation) Government Comprehensive High School, Faisalabad
1987-1989	Higher Secondary School Certificate Examination (Intermediate) Government College of Science, Faisalabad
1989-1991	B.Sc. (Statistics, Mathematics): Govt College of Science, Faisalabad Final Exam: Punjab University, Lahore <i>Distinctions:</i> <i>Certificate of Merit; Roll of Honour</i>
1992-1994	M.Sc. (Statistics): Department of Mathematics and Statistics, Faculty of Basic Sciences, University of Agriculture, Faisalabad <i>Distinctions:</i> <i>Merit Scholarship; Medal of Distinction</i>
Since 25.05.1995	Lecturer (Statistics) Government Gordon College, Rawalpindi
2002-2004	M.Phil. (Statistics): Department of Mathematics and Statistics, Faculty of Basic Sciences, University of Agriculture, Faisalabad <i>Distinctions:</i> <i>Merit Scholarship</i>
10.2004 - 07.2008	Ph.D. (Statistics): Department of Medical Statistics (Center of Statistics), Gerog-August University, Göttingen, Germany Supervisor: Prof. Dr. Edgar Brunner Co-supervisor: Prof. Dr. Manfred Denker <i>HEC(Pakistan)/DAAD(Germany) Scholarship</i>

