MULTIVARIATE CONCORDANCE CORRELATION COEFFICIENT

A Dissertation in
Statistics
by
Sasiprapa Hiriote

© 2009 Sasiprapa Hiriote

Submitted in Partial Fulfillment
of the Requirements
for the Degree of

Doctor of Philosophy

August 2009
The dissertation of Sasiprapa Hiriote was reviewed and approved* by the following:

Vernon M. Chinchilli  
Distinguished Professor of Public Health Sciences  
and Professor of Statistics  
Dissertation Adviser  
Chair of Committee

Donald St.P. Richards  
Professor of Statistics

Bing Li  
Professor of Statistics

Tonya S. King  
Associate Professor of Public Health Sciences

Peter C. M. Molenaar  
Professor of Human Development

Bruce G. Lindsay  
Willaman Professor of Statistics  
Head of the Department of Statistics

*Signatures on file in the Graduate School.
Abstract

In many clinical studies, the Lin's concordance correlation coefficient (CCC) is a common tool to assess the agreement of a continuous response measured by two raters or methods. However, the need of measures of agreement may arise for more complex situations, such as when the responses are measured on more than one occasion by each rater or method. In this work, we propose a new CCC in the presence of repeated measurements, called the multivariate concordance correlation coefficient. We constructed the multivariate CCC based on a matrix that possesses the properties needed to characterize the level of agreement between two $p \times 1$ vectors of random variables. For ease of interpretation, we transformed this matrix to a scalar whose value is scaled to range between -1 and 1 by using three distinct functions, namely trace, highest eigenvalue, and determinant. It can be shown that the multivariate CCC reduces to Lin's CCC when $p = 1$. For inference, we proposed an asymptotically unbiased estimator based on U-statistics and derived its asymptotic distribution for each form of the function. The proposed estimators are proven to be asymptotically normal and their performances are evaluated via simulation studies. To obtain a confidence interval or a test statistic, we considered a sample moment estimator of the asymptotic variance and the Z-transformation to improve the normal approximation and bound the confidence limits. The simulation studies confirmed that overall in terms of accuracy, precision, and the coverage probabilities, the estimator of the multivariate CCC based on the determinant function works relatively well in general cases even with small samples. However, for a skewed underlying distribution with moderate or weaker correlation between the two variables, the trace multivariate CCC is slightly more robust. Finally, We used real data from an Asthma Clinical Research Network (ACRN) study and the Penn State Young Women’s Health Study for demonstration.
Table of Contents

List of Tables .................................................. v
List of Figures ................................................... vi
Acknowledgments .................................................. x

Chapter 1. Introduction ............................................. 1

Chapter 2. Literature Review ...................................... 8
  2.1 Lin’s Concordance Correlation Coefficient ................. 8
  2.2 CCC in the presence of repeated measures ................. 12
  2.3 The Review of U-Statistics ................................. 19
    2.3.1 U-statistics ........................................ 19
    2.3.2 The variance of a U-statistic ...................... 21
    2.3.3 The covariance of two U-statistics ............... 24
    2.3.4 The asymptotic distribution of U-statistics .... 26

Chapter 3. Multivariate Concordance Correlation Coefficient .... 28
  3.1 Multivariate Concordance Correlation Coefficient ...... 28
  3.2 Inference ................................................. 38

Chapter 4. Monte Carlo Simulation ............................... 47
  4.1 Multivariate normal distribution ......................... 47
  4.2 Multivariate Student’s t-distribution ................... 53
  4.3 Multivariate lognormal distribution .................... 55
  4.4 Summary .................................................. 59

Chapter 5. Examples ............................................. 121
  5.1 Blood draws data ........................................ 121
  5.2 Body fat data ........................................... 123

Chapter 6. Future Studies ....................................... 129

Bibliography ..................................................... 131
## List of Tables

<table>
<thead>
<tr>
<th></th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>4.1</td>
<td>Mean of estimates of $\hat{\rho}_g$, relative bias, the estimated asymptotic variances $\hat{\sigma}_Z^2$, and the empirical variance of $\hat{Z}$ based on simulation of 1000 runs from the multivariate normal distribution: three repeated measurements.</td>
<td>73</td>
</tr>
<tr>
<td>4.2</td>
<td>Mean of estimates of $\hat{\rho}_g$, relative bias, the estimated asymptotic variances $\hat{\sigma}_Z^2$, and the empirical variance of $\hat{Z}$ based on simulation of 1000 runs from the multivariate normal distribution: five repeated measurements.</td>
<td>76</td>
</tr>
<tr>
<td>4.3</td>
<td>Empirical coverage probabilities for the 95% confidence intervals of $\rho_g$ based on simulation of 1000 runs generated from the multivariate normal distribution.</td>
<td>79</td>
</tr>
<tr>
<td>4.4</td>
<td>Mean of estimates of $\hat{\rho}_g$, relative bias, the estimated asymptotic variances $\hat{\sigma}_Z^2$, and the empirical variance of $\hat{Z}$ based on simulation of 1000 runs from multivariate Student’s t-distribution: three repeated measurements.</td>
<td>93</td>
</tr>
<tr>
<td>4.5</td>
<td>Mean of estimates of $\hat{\rho}_g$, relative bias, the estimated asymptotic variances $\hat{\sigma}_Z^2$, and the empirical variance of $\hat{Z}$ based on simulation of 1000 runs from multivariate Student’s t-distribution: five repeated measurements.</td>
<td>96</td>
</tr>
<tr>
<td>4.6</td>
<td>Empirical coverage probabilities for the 95% confidence intervals of $\rho_g$ based on simulation of 1000 runs generated from multivariate Student’s t distribution.</td>
<td>99</td>
</tr>
<tr>
<td>4.7</td>
<td>Mean of estimates of $\hat{\rho}_g$, relative bias, the estimated asymptotic variances $\hat{\sigma}_Z^2$, and the empirical variance of $\hat{Z}$ based on simulation of 1000 runs from the multivariate lognormal distribution: three repeated measurements.</td>
<td>113</td>
</tr>
<tr>
<td>4.8</td>
<td>Mean of estimates of $\hat{\rho}_g$, relative bias, the estimated asymptotic variances $\hat{\sigma}_Z^2$, and the empirical variance of $\hat{Z}$ based on simulation of 1000 runs from the multivariate lognormal distribution: five repeated measurements.</td>
<td>116</td>
</tr>
<tr>
<td>4.9</td>
<td>Empirical coverage probabilities for the 95% confidence intervals of $\rho_g$ based on simulation of 1000 runs generated from the multivariate lognormal distribution.</td>
<td>119</td>
</tr>
</tbody>
</table>
List of Figures

1.1 Cases when Pearson correlation coefficient, paired t-test, and least square test fail to detect poor agreement. ........................................ 6
1.2 Cases when paired t-test and least square test reject high reproducibility. .................................................. 7

4.1 Q-Q plot of the estimates of $\hat{\rho}_g$ and $\hat{Z}$ for sample size 20 based on simulation of 1000 runs from the multivariate normal distribution with three repeated measure, high within-unit correlation($\rho = 0.8$) ,and strong between-unit correlation(Case 1). ............................... 61
4.2 Q-Q plot of the estimates of $\hat{\rho}_g$ and $\hat{Z}$ for sample size 40 based on simulation of 1000 runs from the multivariate normal distribution with three repeated measure, high within-unit correlation($\rho = 0.8$) ,and strong between-unit correlation(Case 1). ............................... 62
4.3 Q-Q plot of the estimates of $\hat{\rho}_g$ and $\hat{Z}$ for sample size 80 based on simulation of 1000 runs from the multivariate normal distribution with three repeated measure, high within-unit correlation($\rho = 0.8$) ,and strong between-unit correlation(Case 1). ............................... 63
4.4 Q-Q plot of the estimates of $\hat{\rho}_g$ and $\hat{Z}$ for sample size 160 based on simulation of 1000 runs from the multivariate normal distribution with three repeated measure, high within-unit correlation($\rho = 0.8$) ,and strong between-unit correlation(Case 1). ............................... 64
4.5 Q-Q plot of the estimates of $\hat{\rho}_g$ and $\hat{Z}$ for sample size 20 based on simulation of 1000 runs from the multivariate normal distribution with three repeated measure, high within-unit correlation($\rho = 0.8$) ,and moderate between-unit correlation(Case 2). ........................................... 65
4.6 Q-Q plot of the estimates of $\hat{\rho}_g$ and $\hat{Z}$ for sample size 40 based on simulation of 1000 runs from the multivariate normal distribution with three repeated measure, high within-unit correlation($\rho = 0.8$) ,and moderate between-unit correlation(Case 2). ........................................... 66
4.7 Q-Q plot of the estimates of $\hat{\rho}_g$ and $\hat{Z}$ for sample size 80 based on simulation of 1000 runs from the multivariate normal distribution with three repeated measure, high within-unit correlation($\rho = 0.8$) ,and moderate between-unit correlation(Case 2). ........................................... 67
4.8 Q-Q plot of the estimates of $\hat{\rho}_g$ and $\hat{Z}$ for sample size 160 based on simulation of 1000 runs from the multivariate normal distribution with three repeated measure, high within-unit correlation($\rho = 0.8$) ,and moderate between-unit correlation(Case 2). ........................................... 68
4.9 Q-Q plot of the estimates of $\hat{\rho}_g$ and $\hat{Z}$ for sample size 20 based on simulation of 1000 runs from the multivariate normal distribution with three repeated measure, high within-unit correlation($\rho = 0.8$) ,and weak between-unit correlation(Case 3). ........................................... 69
4.10 Q-Q plot of the estimates of $\hat{\rho}_g$ and $\hat{Z}$ for sample size 40 based on simulation of 1000 runs from the multivariate normal distribution with three repeated measure, high within-unit correlation ($\rho = 0.8$) and weak between-unit correlation (Case 3). .................................................. 70
4.11 Q-Q plot of the estimates of $\hat{\rho}_g$ and $\hat{Z}$ for sample size 80 based on simulation of 1000 runs from the multivariate normal distribution with three repeated measure, high within-unit correlation ($\rho = 0.8$) and weak between-unit correlation (Case 3). .................................................. 71
4.12 Q-Q plot of the estimates of $\hat{\rho}_g$ and $\hat{Z}$ for sample size 160 based on simulation of 1000 runs from the multivariate normal distribution with three repeated measure, high within-unit correlation ($\rho = 0.8$) and weak between-unit correlation (Case 3). .................................................. 72
4.13 Q-Q plot of the estimates of $\hat{\rho}_g$ and $\hat{Z}$ for sample size 20 based on simulation of 1000 runs from multivariate Student’s t-distribution with three repeated measure, high within-unit correlation ($\rho = 0.8$) and strong between-unit correlation (Case 1). .................................................. 81
4.14 Q-Q plot of the estimates of $\hat{\rho}_g$ and $\hat{Z}$ for sample size 40 based on simulation of 1000 runs from multivariate Student’s t-distribution with three repeated measure, high within-unit correlation ($\rho = 0.8$) and strong between-unit correlation (Case 1). .................................................. 82
4.15 Q-Q plot of the estimates of $\hat{\rho}_g$ and $\hat{Z}$ for sample size 80 based on simulation of 1000 runs from multivariate Student’s t-distribution with three repeated measure, high within-unit correlation ($\rho = 0.8$) and strong between-unit correlation (Case 1). .................................................. 83
4.16 Q-Q plot of the estimates of $\hat{\rho}_g$ and $\hat{Z}$ for sample size 160 based on simulation of 1000 runs from multivariate Student’s t-distribution with three repeated measure, high within-unit correlation ($\rho = 0.8$) and strong between-unit correlation (Case 1). .................................................. 84
4.17 Q-Q plot of the estimates of $\hat{\rho}_g$ and $\hat{Z}$ for sample size 20 based on simulation of 1000 runs from multivariate Student’s t-distribution with three repeated measure, high within-unit correlation ($\rho = 0.8$) and moderate between-unit correlation (Case 2). .................................................. 85
4.18 Q-Q plot of the estimates of $\hat{\rho}_g$ and $\hat{Z}$ for sample size 40 based on simulation of 1000 runs from multivariate Student’s t-distribution with three repeated measure, high within-unit correlation ($\rho = 0.8$) and moderate between-unit correlation (Case 2). .................................................. 86
4.19 Q-Q plot of the estimates of $\hat{\rho}_g$ and $\hat{Z}$ for sample size 80 based on simulation of 1000 runs from multivariate Student’s t-distribution with three repeated measure, high within-unit correlation ($\rho = 0.8$) and moderate between-unit correlation (Case 2). .................................................. 87
4.20 Q-Q plot of the estimates of $\hat{\rho}_g$ and $\hat{Z}$ for sample size 160 based on simulation of 1000 runs from multivariate Student’s t-distribution with three repeated measure, high within-unit correlation ($\rho = 0.8$) and moderate between-unit correlation (Case 2). .................................................. 88
4.21 Q-Q plot of the estimates of $\hat{\rho}_g$ and $\hat{Z}$ for sample size 20 based on simulation of 1000 runs from multivariate Student’s $t$-distribution with three repeated measure, high within-unit correlation ($\rho = 0.8$) and weaker between-unit correlation (Case 3). ..................................................... 89

4.22 Q-Q plot of the estimates of $\hat{\rho}_g$ and $\hat{Z}$ for sample size 40 based on simulation of 1000 runs from multivariate Student’s $t$-distribution with three repeated measure, high within-unit correlation ($\rho = 0.8$) and weaker between-unit correlation (Case 3). ..................................................... 90

4.23 Q-Q plot of the estimates of $\hat{\rho}_g$ and $\hat{Z}$ for sample size 80 based on simulation of 1000 runs from multivariate Student’s $t$-distribution with three repeated measure, high within-unit correlation ($\rho = 0.8$) and weaker between-unit correlation (Case 3). ..................................................... 91

4.24 Q-Q plot of the estimates of $\hat{\rho}_g$ and $\hat{Z}$ for sample size 160 based on simulation of 1000 runs from multivariate Student’s $t$-distribution with three repeated measure, high within-unit correlation ($\rho = 0.8$) and weaker between-unit correlation (Case 3). ..................................................... 92

4.25 Q-Q plot of the estimates of $\hat{\rho}_g$ and $\hat{Z}$ for sample size 20 based on simulation of 1000 runs from the multivariate lognormal distribution with three repeated measure, high within-unit correlation ($\rho = 0.8$) and strong between-unit correlation (Case 1). ..................................................... 101

4.26 Q-Q plot of the estimates of $\hat{\rho}_g$ and $\hat{Z}$ for sample size 40 based on simulation of 1000 runs from the multivariate lognormal distribution with three repeated measure, high within-unit correlation ($\rho = 0.8$) and strong between-unit correlation (Case 1). ..................................................... 102

4.27 Q-Q plot of the estimates of $\hat{\rho}_g$ and $\hat{Z}$ for sample size 80 based on simulation of 1000 runs from the multivariate lognormal distribution with three repeated measure, high within-unit correlation ($\rho = 0.8$) and strong between-unit correlation (Case 1). ..................................................... 103

4.28 Q-Q plot of the estimates of $\hat{\rho}_g$ and $\hat{Z}$ for sample size 160 based on simulation of 1000 runs from the multivariate lognormal distribution with three repeated measure, high within-unit correlation ($\rho = 0.8$) and strong between-unit correlation (Case 1). ..................................................... 104

4.29 Q-Q plot of the estimates of $\hat{\rho}_g$ and $\hat{Z}$ for sample size 20 based on simulation of 1000 runs from the multivariate lognormal distribution with three repeated measure, high within-unit correlation ($\rho = 0.8$) and moderate between-unit correlation (Case 2). ..................................................... 105

4.30 Q-Q plot of the estimates of $\hat{\rho}_g$ and $\hat{Z}$ for sample size 40 based on simulation of 1000 runs from the multivariate lognormal distribution with three repeated measure, high within-unit correlation ($\rho = 0.8$) and moderate between-unit correlation (Case 2). ..................................................... 106

4.31 Q-Q plot of the estimates of $\hat{\rho}_g$ and $\hat{Z}$ for sample size 80 based on simulation of 1000 runs from the multivariate lognormal distribution with three repeated measure, high within-unit correlation ($\rho = 0.8$) and moderate between-unit correlation (Case 2). ..................................................... 107
4.32 Q-Q plot of the estimates of $\hat{\rho}_g$ and $\hat{Z}$ for sample size 160 based on simulation of 1000 runs from the multivariate lognormal distribution with three repeated measure, high within-unit correlation ($\rho = 0.8$), and moderate between-unit correlation (Case 2). ................................. 108

4.33 Q-Q plot of the estimates of $\hat{\rho}_g$ and $\hat{Z}$ for sample size 20 based on simulation of 1000 runs from the multivariate lognormal distribution with three repeated measure, high within-unit correlation ($\rho = 0.8$), and weaker between-unit correlation (Case 3). ................................. 109

4.34 Q-Q plot of the estimates of $\hat{\rho}_g$ and $\hat{Z}$ for sample size 40 based on simulation of 1000 runs from the multivariate lognormal distribution with three repeated measure, high within-unit correlation ($\rho = 0.8$), and weaker between-unit correlation (Case 3). ................................. 110

4.35 Q-Q plot of the estimates of $\hat{\rho}_g$ and $\hat{Z}$ for sample size 80 based on simulation of 1000 runs from the multivariate lognormal distribution with three repeated measure, high within-unit correlation ($\rho = 0.8$), and weaker between-unit correlation (Case 3). ................................. 111

4.36 Q-Q plot of the estimates of $\hat{\rho}_g$ and $\hat{Z}$ for sample size 160 based on simulation of 1000 runs from the multivariate lognormal distribution with three repeated measure, high within-unit correlation ($\rho = 0.8$), and weaker between-unit correlation (Case 3). ................................. 112

5.1 Summary statistics of blood draw data for each visit with Pearson Correlation Coefficients. ......................................................... 125

5.2 Scatter plots of blood draw data for each visit with Lin’s CCCs. ...... 126

5.3 Summary statistics of body fat data for each visit with Pearson Correlation Coefficients. ......................................................... 127

5.4 Scatter plots of body fat data for each visit with Lin’s CCCs. ...... 128
I would like to deeply thank my dissertation advisor Professor Vernon M. Chinchilli for his great help throughout the years that I have been working on my research. Without his guidance, support, and encouragement, this dissertation could not have been completed. I would also like to thank my other committee members, Professor Donald St.P. Richards, Professor Tonya S. King, Professor Bing Li, and Professor Peter C. M. Molenaar for their valuable questions and comments. I want to give a special thank to Professor Donald Richards who devoted his time to discuss with me about some theoretical and simulation problems. I also want to give a special thank to Professor Tonya King for providing me all the dataset that I used for the demonstration of this work and helping me solve some programming problems.

In addition, I would like to thank all of those professors, staff, and colleagues in the Department of Statistics at Penn State for being so supportive during the years that I have been here. I would like to specially thank the Department for the financial support. I would also like to thank Professor Michele L. Shaffer and Professor Vernon M. Chinchilli for giving me the research assistantship in the Public Health Science Department at the Hershey Medical Center.

Last but not least, I would like to thank my father Prasan Hiriote, my mother Aapa Hiriote, and my three sisters for their unconditional love and care.
Chapter 1

Introduction

A measure of agreement between two responses, whose observations are made by two raters or methods, is often needed in many research studies. For example, in the area of medical diagnostic testing, the main research interest is to compare the results of the new technique with those of the gold standard practice. Moreover, in many scientific studies, the reproducibility of measurements from trial to trial is one of the research objectives. When the responses are continuous, the concordance correlation coefficient (CCC), introduced by Lin (1989), is one of the most popular measures of agreement. The CCC evaluates the agreement between two readings from the same sample by measuring how far each paired data point deviates from the 45° line through the origin, called the concordance line. Unlike the traditional approaches, e.g. the Pearson correlation coefficient, the paired t-test, and the least square test, which sometimes fail to detect departure from the concordance line (See Figure 1.1) or falsely reject strong agreement (See Figure 1.2), the CCC can fully assess the desired reproducibility characteristics. The CCC contains not only the measurements of precision, which is the deviation of observations from the fitted line, but also the measurements of accuracy, which is the deviation of the fitted line from the concordance line.

To characterize the index of agreement between two random variables, X and Y, Lin (1989) considered the expected value of the squared difference, $E(X - Y)^2$, and
constructed the CCC as

$$\rho_c = 1 - \frac{\mathbb{E}[(X - Y)^2]}{\mathbb{E}[(X - Y)^2 | X, Y \text{ are independent}]} = \frac{2\text{Cov}(X, Y)}{\text{Var}(X) + \text{Var}(Y) + [\mathbb{E}(X) - \mathbb{E}(Y)]^2}$$

(1.1)

The value of $\rho_c$ ranges between -1 and 1 with the equality to 1 for perfect positive agreement, 0 for no agreement, and -1 for perfect negative agreement. Based on a bivariate random sample $(X_1, Y_1), \cdots, (X_n, Y_n)$, $\rho_c$ is estimated by replacing each of the parameters in (1.1) with its sample moment. The sample CCC, $\hat{\rho}_c$, has been shown by Lin (1989) to be consistent and have an asymptotic normal distribution with mean $\rho_c$ and variance $(1 - \rho_c^2)^2 \sigma^2_\hat{Z}$ for bivariate normal samples, where $\hat{Z}$ is the sample version of the $Z$-transformation of $\rho_c$, $Z = \tanh^{-1} \rho_c$ and $\sigma^2_\hat{Z}$ is the asymptotic variance of $\hat{Z}$. He first showed the asymptotic normality of $\hat{Z}$ and then used the theorem on functions of asymptotically normal statistics to obtain the asymptotic distribution of $\hat{\rho}_c$. The sample size formula for testing the assay’s reproducibility has also been derived by Lin (1992) so that the smaller the sample size, the less likely the chance of accepting the assay.

To deal with other types of problems involving agreement, there has been several articles in the literatures proposing extended versions of Lin’s CCC. In addition to the squared function of distance used in its original formula, King and Chinchilli (2001a) considered alternative functions, e.g. the absolute distance function and Huber’s function, to construct more robust versions of the CCC. Other than this extension, the generalized CCC, introduced by King and Chinchilli (2001b) can be applied to the cases with more than two responses measured in either a continuous or categorical scale.
In many fields of science, especially medical sciences, the need of the CCC to assess the agreement between two responses often arises when each response is a vector containing repeated measurements. For example, in a longitudinal asthma clinical trial, one of the research goals was to study the amount of agreement between plasma cortisol AUC (area under the curve) measured every hour and every two hours at several visits. For paired or unpaired repeated measurements study design, Chinchilli et al. (1996) developed a weighted CCC based on a random coefficient model which allows the within-subject variances to change across subjects. For each subject, the CCC was constructed as an average of q CCC’s of the least squares random vectors, whose variance-covariance matrices were of dimension $q \times q$. Then, the global CCC was defined as a weighted average of the coefficients using a weight function based on the amount of variation within each subject. The estimator of the weighted CCC was obtained by substituting each parameter in the CCC and weight function for each subject with its unbiased estimator.

King et al. (2007) proposed another version of CCC in the presence of repeated measurements. They characterized the amount of agreement between two $p \times 1$ random vectors, $X$ and $Y$, by $\text{E} \left[ (X - Y)'D(X - Y) \right]$, where $D$ is a $p \times p$ non-negative definite matrix of weights among the different repeated measurements. Then, the repeated measure CCC was defined as

$$
\rho_{c, rm} = 1 - \frac{\text{E} \left[ (X - Y)'D(X - Y) \right]}{\text{E} \left[ (X - Y)'D(X - Y) \big| X, Y \text{ are independent} \right]}
= \frac{\text{trace} \left( D\text{Cov}(X, Y) + D\text{Cov}(Y, X) \right)}{\text{trace} \left( D\text{Var}(X) + D\text{Var}(Y) \right) + (\text{E}(X) - \text{E}(Y))' D (\text{E}(X) - \text{E}(Y))}
$$
This repeated measures CCC reduces to Lin’s CCC when \( p = 1 \). The authors obtained the estimator of the repeated measures CCC by replacing each parameter in \( \rho_{c,rm} \) with its sample counterparts, denoted by \( \hat{\rho}_{c,rm} \). Moreover, they expressed this \( \hat{\rho}_{c,rm} \) as a ratio of functions of U-statistics so that it could be shown that the \( \hat{\rho}_{c,rm} \) has an asymptotically normal distribution with mean \( \rho_{c,rm} \) and a variance whose consistent estimator could be derived by using the delta method.

In this work, we are also interested in a measure of overall agreement between two vectors of repeated measurements. However, we noticed that to use the subject-specific repeated measure CCC proposed by Chinchilli et al. (1996), an appropriate model is needed to be fitted to obtain accurate variance component estimates. In addition, for some models, a highly complex within-unit variance-covariance structure has to be assumed in order for the variance components to be estimable. Moreover, to use the weighted repeated measure CCC introduced by King et al. (2007), a suitable weight matrix is needed to be presumed based on the study design and the reliability of the two vectors, in which the latter depends on the source or the amount of available data.

Therefore, we introduce a new repeated measure CCC, called multivariate CCC, which can be easily used without any complicated restriction or assumption. We first construct a matrix denoted by \( \mathbf{M}_{\rho_c} \) which possesses the characteristics needed to measure the amount of agreement between two \( p \times 1 \) vectors of random variables and reduces to Lin’s CCC when \( p=1 \). Then, to ease the problem of interpretation, we transform the matrix \( \mathbf{M}_{\rho_c} \) to a scalar whose value is scaled to range between -1 and 1. We consider three forms of the multivariate CCC based on three different functions denoted by \( \rho_{gi}, i = 1, 2, 3 \). These \( \rho_{gi} \) can be used to measure the level of agreement between the two vectors as \( \mathbf{M}_{\rho_c} \).
but it is much easier to interpret, especially when \( p \) is large. To estimate each of \( M_{\rho c} \) and \( \rho_{gi}, i = 1, 2, 3 \), we consider an estimator based on U-statistics. For inference, we derive the asymptotic distributions of the estimators of \( \rho_{gi}, i = 1, 2, 3 \).

In Chapter 2, we provide more details of Lin’s CCC, which is the basis of our motivation. The literature about the repeated measure CCC is reviewed as well. Furthermore, we review U-statistics that are used to construct our statistics for inference about the proposed multivariate CCC. In Chapter 3, the three versions of multivariate CCC are introduced and shown to have the desired characteristics. Also, the estimators based on U-statistics and the asymptotic distributions of the multivariate CCC are established. In Chapter 4, Monte Carlo simulation is performed to assess and compare the properties of the three estimators of the multivariate CCC based on finite samples. Finally, in Chapter 5, some real examples are used to demonstrate the application of the multivariate CCC, and a number of potential future studies are discussed in Chapter 6.
Fig. 1.1: Cases when Pearson correlation coefficient, paired t-test, and least square test fail to detect poor agreement.
**Fig. 1.2:** Cases when paired t-test and least square test reject high reproducibility.
Chapter 2

Literature Review

In this Chapter, the original version of Lin’s CCC, and the modified versions of Lin’s CCC in the presence of repeated measurements, is reviewed. In addition, $U$-statistics, which are the significant tool used to derive the asymptotic distribution of the statistics proposed in this work, is summarized.

2.1 Lin’s Concordance Correlation Coefficient

Assume that $(X_i, Y_i), i = 1, \cdots, n$ are random sample pairs from a bivariate distribution with mean $(\mu_X, \mu_Y)$ and covariance matrix

$$
\begin{pmatrix}
\sigma_X^2 & \sigma_{XY} \\
\sigma_{XY} & \sigma_Y^2
\end{pmatrix}
$$

Lin (1989) defined the CCC for $X$ and $Y$ as

$$
\rho_c = 1 - \frac{E[(X - Y)^2]}{E[(X - Y)^2 | X, Y \text{ are independent}]}
= \frac{2\sigma_{XY}}{\sigma_X^2 + \sigma_Y^2 + (\mu_X - \mu_Y)^2} = \rho_C b
$$
where

$$\rho = \text{Pearson Correlation Coefficient}$$

$$C_b = \frac{2}{v + \frac{1}{v} + u^2}$$

$$v = \frac{\sigma_X}{\sigma_Y} = \text{scale shift}$$

$$u = \frac{(\mu_X - \mu_Y)}{\sqrt{\sigma_X \sigma_Y}} = \text{location shift relative to scale shift}$$

Whereas the Pearson correlation coefficient measures how far the observations deviate from the fitted line, a bias correction factor, $C_b$, measures how far the fitted line deviates from the concordance line. Note that $0 < C_b \leq 1$ and the equality occurs when there is no deviation from the concordance line. By definition, the values of $\rho_c$ range between -1 and 1; $\rho_c = \pm 1$ iff $\rho = \pm 1$, $\sigma_1 = \sigma_2$, and $\mu_1 = \mu_2$, or equivalently $\rho_c = \pm 1$ iff each observation falls on the concordance line. Note that $\rho_c = 0$ iff $\rho = 0$. That is, $\rho_c$ measures how far the observations deviate from the concordance line. The further the observations from the concordance line, the greater $\rho_c$ differs from ±1.

Lin (1989) used the sample counterparts to estimate $\rho_c$ as follows,

$$\hat{\rho}_c = \frac{2S_{XY}}{S_X^2 + S_Y^2 + (X - Y)^2}$$
where

\[ \bar{X} = \frac{1}{n} \sum_{i=1}^{n} X_i, \quad \bar{Y} = \frac{1}{n} \sum_{i=1}^{n} Y_i \]

\[ S_X^2 = \frac{1}{n} \sum_{i=1}^{n} (X_i - \bar{X})^2, \quad S_Y^2 = \frac{1}{n} \sum_{i=1}^{n} (Y_i - \bar{Y})^2 \]

\[ S_{XY} = \frac{1}{n} \sum_{i=1}^{n} (X_i - \bar{X})(Y_i - \bar{Y}) \]

Assuming that the n sample pairs are from a bivariate normal distribution, \( \hat{\rho}_c \) is shown to be a consistent estimator of \( \rho_c \) and has an asymptotic normal distribution with mean \( \rho_c \) and variance \( (1 - \rho_c^2)^2 \sigma_Z^2 \) by first proving that the Z-transformation,

\[ \hat{Z} = \tanh^{-1} \hat{\rho}_c = \frac{1}{2} \ln \frac{1 + \hat{\rho}_c}{1 - \hat{\rho}_c} \]

is asymptotically normal with mean \( Z = \frac{1}{2} \ln \frac{1+\rho_c}{1-\rho_c} \) and variance

\[ \sigma_Z^2 = \frac{1}{n} \left[ \frac{(1 - \rho^2) \hat{\rho}_c^2}{(1 - \rho_c^2) \rho^2} + \frac{2\rho_c^3 (1 - \rho_c) u^2}{\rho (1 - \rho_c^2)^2} - \frac{\rho_c^4 u^4}{2\rho^2 (1 - \rho_c^2)^2} \right]. \]

Lin (1989) examined the convergence to asymptotic normality of \( \hat{\rho}_c \) and \( \hat{Z} \) based on paired samples from some bivariate normal distributions by performing a Monte Carlo simulation for five different combinations of \( \rho_c \), location and scale parameters and three sample sizes (n = 10, 20, and 50). Using a Kolmogorov goodness-of-fit test, he found that the closer \( \rho_c \) is to zero and/or the larger the n is, the closer the distribution of \( \hat{\rho}_c \) is to normality. By the results of this simulation, the distribution of \( \hat{Z} \) approached normality faster than that of \( \hat{\rho}_c \) for all 15 cases. Also, the simulation showed that the
asymptotic normality of $\hat{\rho}_c$ and $\hat{Z}$ was robust even when the samples were drawn from uniform and Poisson distributions.

In order to demonstrate the application of the proposed CCC, Lin (1989) studied two interesting questions. The first question was about the reproducibility of a new machine for measuring total bilirubin in blood compared to that of the gold-standard one. The second example was the comparison between two assays used in screening the toxicity of biomaterials for medical devices in terms of the reproducibility of the measurements from trial to trial.

In the study of the first question, ten blood samples were taken and measured for bilirubin level by three operators (well-trained, with minimal and no training), using the new and standard machine. The results showed the high reproducibility by the well-trained personnel ($\hat{\rho}_c = 0.995$), medium reproducibility by the personnel with minimal training ($\hat{\rho}_c = 0.838$), and low reproducibility by the personnel with no training ($\hat{\rho}_c = 0.624$). In addition, the estimates of precision ($r$), location shift ($\hat{u}$), and scale shift ($\hat{v}$) were shown for each operator. Whereas the well-trained personnel performed with very high precision ($r = 0.999$), minimum location shift ($\hat{u} = -.041$) and scale shift ($\hat{v} = 0.916$), the personnel with minimal training performed with less precision ($r = 0.936$), higher location shift ($\hat{u} = 0.768$) and scale shift ($\hat{v} = 1.878$). In the second example, two assays (ATP-76 and CLA-74) were used to assess the toxicity of ten materials for two independent trials. The analysis showed that ATP-76 ($\hat{\rho}_c = 0.969$) performed much better than CLA-74 ($\hat{\rho}_c = 0.283$) in terms of the reproducibility between trials.
2.2 CCC in the presence of repeated measures

Chinchilli et al. (1996) introduced a weighted concordance correlation coefficient for two scenarios of repeated measurement designs in which the observations of X and Y are paired or unpaired over time. For the unpaired case, the data of X and Y from the $i^{th}$ experimental unit are denoted by $X_i = [X_{i1}, \cdots, X_{im_i}]'$ and $Y_i = [Y_{i1}, \cdots, Y_{im_i}]'$. They considered a version of a random-coefficient generalized multivariate analysis model for $X_i$ and $Y_i$ with some assumptions about the independence among random effects and the first two moments of those random effects. Then the least squares vectors on X and Y for the $i^{th}$ unit, where $E(X_i) = A_X \beta_X$ and $E(Y_i) = A_Y \beta_Y$, are constructed as follows:

$$X_i^* = (A'_{X_i} A_{X_i})^{-1} A'_{X_i} X_i \quad \text{and} \quad Y_i^* = (A'_{Y_i} A_{Y_i})^{-1} A'_{Y_i} Y_i$$

The CCC for the $i^{th}$ unit, $i = 1, \cdots, n$ is defined by

$$\rho_{c,i} = \frac{1}{q} \sum_{j=1}^{q} \text{CCC}(X_{ij}^*, Y_{ij}^*)$$

where

$$\begin{bmatrix}
\mu_{X_i^*} \\
\mu_{Y_i^*}
\end{bmatrix} = E \begin{bmatrix}
X_i^* \\
Y_i^*
\end{bmatrix} = \begin{bmatrix}
\beta_X \\
\beta_Y
\end{bmatrix}$$
and

\[
\begin{bmatrix}
\Delta_{XX_i} & \Delta_{XY_i} \\
\Delta_{YX_i} & \Delta_{YY_i}
\end{bmatrix}
= \text{var}
\begin{bmatrix}
X_i^* \\
Y_i^*
\end{bmatrix}
\]

is obtained by the above transformations and the assumptions of the models and \((\Gamma)_{jj}\) denotes the \(j^{th}\) diagonal element of a square matrix \(\Gamma\). From these CCC, the weighted CCC is defined by

\[
\rho_c = \left( \sum_{i=1}^{n} w_i \right)^{-1} \left( \sum_{i=1}^{n} w_i \rho_{c,i} \right)
\]

where \(w_i, i = 1, \cdots, n\) are non-negative weights. The authors used the weights

\[
w_i = \frac{1}{(1 + \sigma_{XX_i})(1 + \sigma_{YY_i})}.
\]

This gives higher weights to smaller variations within each variable for each unit. For the paired scenario, \(m_{X_i} = m_{Y_i} = m_i\). In this case, the variance-covariance matrix of the error terms is different from the unpaired case because of the dependence between the observations on \(X\) and \(Y\) within the same unit. As a result, the variance-covariance matrix of \((X_i^*, Y_i^*)'\), where \(X_i^*\) and \(Y_i^*\) are defined as in the former case, is different. The CCC for each unit, \(\rho_{c,i}\), and the weighted CCC, \(\rho_c\), are defined as in the unpaired case, but the weight function, \(w_i\), is modified to account for the covariance between \(X\) and \(Y\) within unit,

\[
w_i = \frac{1}{(1 + \sigma_{XX_i})(1 + \sigma_{YY_i}) - \sigma_{XY_i}^2}.
\]
For both cases, the authors proposed estimators for all parameters and substituted them into the weighted CCC to obtain its estimator. In addition, a bootstrapping algorithm was used to calculate a $100(1 - \alpha)\%$ confidence interval for $\rho_c$.

Chinchilli et al. (1996) illustrated the application of the weighted CCC with three examples. In the first example, two different assays were used to evaluate serum cholesterol on ten occasions for each of 100 subjects. This is an unpaired example with strong agreement between the two assays ($\hat{\rho}_c = 0.994$). The second example was an ancillary study in which two methods of collecting nutrition information, namely a food frequency questionnaire on one occasion and a 24-hour dietary call on six occasions, were compared based on average daily protein intake evaluated from 96 subjects. The agreement between the two methods was weak ($\hat{\rho}_c = 0.35$). In this case, $m_{X_i} = q = 1, i = 1, \cdots, n$ so that $\sigma_{XX_i}$ is not estimable. Thus, the authors recommended assuming that $\sigma_{XX_1} = \cdots = \sigma_{XX_n} = \sigma_{XX}$ and adding a restriction to the variance components of the $(m_{X_i} + m_Y) \times 1$ vector of random effects for identifiability among them. The last example was also an ancillary study of which the objective was to assess the level of agreement between the skinfold caliper and dual energy x-ray absorptiometry (DEXA) measurements of percentage body fat. The data were collected from 90 adolescent girls, whose first visit occurred at age 12 and other 7 visits occurred every 6 months. The overall agreement between the skinfold caliper and DEXA measurements of percentage body fat was not as high as expected ($\hat{\rho}_c = 0.42$).

King et al. (2007) proposed another version of a repeated measures CCC, which evaluates the amount of agreement between two responses measured by two raters or methods in the presence of the repeated measures, denoted by $X$ and $Y$. The elements
of the vector \([X, Y]\) are assumed to be drawn from a multivariate normal distribution with \(2p \times 1\) mean vector \([\mu_X, \mu_Y]\), and \(2p \times 2p\) covariance matrix \(\Sigma\), consisting of four \(p \times p\) matrices: \(\Sigma_{XX}, \Sigma_{XY}, \Sigma_{YX},\) and \(\Sigma_{YY}\). The degree of agreement is characterized by

\[
E \left[ (X - Y)'D(X - Y) \right] = \text{trace} \left\{ D \left( \Sigma_{XX} + \Sigma_{YY} - \Sigma_{XY} - \Sigma_{YX} \right) \right\}
\]

\[+ (\mu_X - \mu_Y)'D(\mu_X - \mu_Y)\]

where \(D\) is a \(p \times p\) non-negative definite matrix of weights among the \(p\) repeated measurements.

King et al. (2007) defined the repeated measures CCC by

\[
\rho_{c,rm} = 1 - \frac{E \left[ (X - Y)'D(X - Y) \right]}{E \left[ (X - Y)'D(X - Y) \right]} = \frac{\text{trace} \left( D\Sigma_{XY} + D\Sigma_{YX} \right)}{\text{trace} \left( D\Sigma_{XX} + D\Sigma_{YY} \right) + (\mu_X - \mu_Y)'D(\mu_X - \mu_Y)}
\]

This \(\rho_{c,rm}\) is the generalization of Lin’s CCC and reduces to Lin’s CCC when \(p=1\).

The authors considered four options for the \(D\) matrix with different emphasis patterns of the within-visit(diagonal elements) and between-visit(off-diagonal elements) agreements as follows:

1. \(D = I_{p \times p}\)
2. \( D = (d_{jk}) \)

where

\[
d_{jk} = \begin{cases} 
  p & , j = k \text{ of greatest interest} \\
  p-1 & \\
  \\
  1 & , j = k \text{ of least interest} \\
  \\
  0 & \text{ when } j \neq k 
\end{cases}
\]

3. \( D = (d_{jk}) \) where \( d_{jk} = d^{j-k} \) \( 0 < d < 1 \)

4. \( D = (d_{jk}) \) where \( d_{jk} = 1 \)

where \( j = 1, \cdots , p \) and \( k = 1, \cdots , p \)

This \( \rho_{c,rm} \) possesses the desired characteristics of the CCC, namely \(-1 \leq \rho_{c,rm} \leq 1\), where

\[
\rho_{c,rm} = \begin{cases} 
  1 & \text{if X and Y are in perfect positive agreement} \\
  -1 & \text{if X and Y are in perfect negative agreement} \\
  0 & \text{if there is no agreement between X and Y} 
\end{cases}
\]

The estimator of the \( \rho_{c,rm} \), denoted by \( \hat{\rho}_{c,rm} \), was obtained by replacing \( \Sigma_{XY} \), \( \Sigma_{YX} \), \( \Sigma_{XX} \), \( \Sigma_{YY} \), \( \mu_X \) and \( \mu_Y \) with their sample moments. The estimator, \( \hat{\rho}_{c,rm} \) also could be
written in terms of U-statistics as follows

\[
\hat{\rho}_{c,rm} = 1 - \frac{nU}{U + (n-1)V} = \frac{(n-1)(V-U)}{U + (n-1)V}
\]

where

\[
U = \frac{1}{n} \sum_i (X_i - Y_i)'D(X_i - Y_i)
\]

\[
V = \frac{1}{n(n-1)} \sum_{i \neq j} (X_i - Y_j)'D(X_i - Y_j)
\]

\[
X_i = \left(X_{i1}, \ldots, X_{ip}\right)
\]

\[
Y_i = \left(Y_{i1}, \ldots, Y_{ip}\right)
\]

By using the delta method with the function \(g(W) = g((U,V)) = (n-1)(V-U)/(U+(n-1)V)\), the asymptotic distribution of \(\hat{\rho}_{c,rm}\) is derived as a normal distribution with mean \(\rho_{c,rm}\) and a variance consistently estimated with

\[
V(\rho_{c,rm}) \doteq d \Sigma d'
\]

where

\[
d = \left(\frac{\partial g}{\partial U} = \frac{-n(n-1)V}{\{U + (n-1)V\}^2}, \quad \frac{\partial g}{\partial V} = \frac{n(n-1)U}{\{U + (n-1)V\}^2}\right)
\]

\[
\Sigma = V(W), \text{ the } 2 \times 2 \text{ variance-covariance matrix of } W, \text{ consistently estimated by }
\]

\[
\Sigma \doteq C \left[\frac{1}{n^2}(\phi_i - W)(\phi_i - W)\right]'C
\]
where

\[
C = \begin{pmatrix} 1 & 0 \\ 0 & 2 \end{pmatrix}, \quad \phi_i = \begin{pmatrix} \phi_{1i} \\ \phi_{2i} \end{pmatrix}, \quad \phi_{1i} = \frac{1}{n-1} \sum_{i \neq j} \phi_{1ij}, \quad \phi_{2i} = \frac{1}{n-1} \sum_{i \neq j} \phi_{2ij},
\]

\[
\phi_{1ij} = \frac{1}{2} (X_i - Y_i)' D (X_i - Y_i) + \frac{1}{2} (X_j - Y_j)' D (X_j - Y_j)
\]

\[
\phi_{2ij} = \frac{1}{2} (X_i - Y_j)' D (X_i - Y_j) + \frac{1}{2} (X_j - Y_i)' D (X_j - Y_i)
\]

Note that \( \phi_{1ij} \) and \( \phi_{2ij} \) are the kernels of the U-statistics

\[
U = \frac{1}{n} \sum_{i} \left\{ \sum_{i \neq j} \frac{\phi_{1ij}}{n-1} \right\}
\]

and

\[
V = \frac{1}{n} \sum_{i} \left\{ \sum_{i \neq j} \frac{\phi_{2ij}}{n-1} \right\}
\]

Moreover, King et al. (2007) considered the Z-transformation to improve the asymptotic normality of \( \hat{\rho}_{c,rm} \), i.e.,

\[
\hat{Z} = \tanh^{-1} \hat{\rho}_{c,rm} = \frac{1}{2} \ln \frac{1 + \hat{\rho}_{c,rm}}{1 - \hat{\rho}_{c,rm}}
\]
By the delta method, \( \hat{Z} \) has an asymptotically normal distribution with mean \( Z = \frac{1}{2} \ln \frac{1 + \rho_{c,rm}}{1 - \rho_{c,rm}} \) and variance \( V(\hat{\rho}_{c,rm})/(1 - \rho_{c,rm}^2)^2 \). Using the transformation from the confidence limits for \( Z \), the confidence limits for \( \rho_{c,rm} \) are

\[
\left( \frac{\exp(2\hat{Z}_L) - 1}{\exp(2\hat{Z}_L) + 1}, \frac{\exp(2\hat{Z}_U) - 1}{\exp(2\hat{Z}_U) + 1} \right)
\]

2.3 The Review of U-Statistics

The review of U-statistics in this section is taken from Hoeffding (1948).

2.3.1 U-statistics

Let \( X_1, X_2, \ldots, X_n \) be \( n \) independent random vectors, \( X_\nu = (X^{(1)}_\nu, \ldots, X^{(r)}_\nu) \), and \( \Phi(x_1, \ldots, x_m) \) a function of \( m(\leq n) \) vectors \( x_\nu = (x^{(1)}_\nu, \ldots, x^{(r)}_\nu) \). Consider the function of the form

\[
U = U(X_1, X_2, \ldots, X_n) = \frac{1}{n(n-1) \cdots (n-m+1)} \sum \Phi(X_{\alpha_1}, \ldots, X_{\alpha_m})
\]  

(2.1)

where \( \sum \) represents summation over all permutations \( (\alpha_1, \ldots, \alpha_m) \) of \( m \) integers such that

\[
1 \leq \alpha_i \leq n, \quad \alpha_i \neq \alpha_j \quad \text{if} \quad i \neq j, \quad (i, j = 1, \ldots, m)
\]

Any statistic of the form (2.1) is called a U-statistic with a kernel function \( \Phi(X_{\alpha_1}, \ldots, X_{\alpha_m}) \). If \( X_1, X_2, \ldots, X_n \) have the same distribution function \( F(x) \), \( U \)
is an unbiased estimator of

$$\theta(F) = \int \cdots \int \Phi(x_1, \ldots, x_m) \, dF(x_1) \cdots dF(x_m). \quad (2.2)$$

$\theta(F)$ is called a regular functional of the distribution function $F(x)$ and $\Phi(x_1, \ldots, x_m)$ is referred to as a kernel of $\theta(F)$.

For any regular functional $\theta(F)$, there exists a kernel $\Phi_0(x_1, \ldots, x_m)$ symmetric in $(x_1, \ldots, x_m)$, that is, if $\Phi(x_1, \ldots, x_m)$ is a kernel of $\theta(F)$,

$$\Phi_0(x_1, \ldots, x_m) = \frac{1}{m!} \sum \Phi(x_{\alpha_1}, \ldots, x_{\alpha_m}), \quad (2.3)$$

where the sum is extended over all permutations $(\alpha_1, \ldots, \alpha_m)$ of $(1, \ldots, m)$.

By (2.1) and (2.3), we may write a U-statistic in the form

$$U(X_1, X_2, \ldots, X_n) = \left( \begin{array}{c} n \\ m \end{array} \right)^{-1} \sum \Phi_0(X_{\alpha_1}, \ldots, X_{\alpha_m}) \quad (2.4)$$

where the kernel $\Phi_0$ is symmetric in its $m$ vector arguments and the sum $\sum$ is taken over all subscripts $\alpha$ such that

$$1 \leq \alpha_1 < \alpha_2 < \cdots < \alpha_m \leq n.$$
2.3.2 The variance of a U-statistic

Let \((X_1, X_2, \cdots, X_n)\) be \(n\) independent random vectors with the same distribution function \(F(x) = F(x^{(1)}, \cdots, x^{(r)})\), and let

\[
U = U(X_1, X_2, \cdots, X_n) = \left( \begin{array}{c} n \\ m \end{array} \right)^{-1} \sum \Phi(X_{\alpha_1}, \cdots, X_{\alpha_m}) \tag{2.5}
\]

where \(\Phi(x_1, \cdots, x_m)\) is symmetric in its argument and \(\sum\) is the same as the summation in (2.4). Assume that the function \(\Phi\) does not involve \(n\).

Consider \(\theta = \theta(F)\) defined by (2.2). We have

\[
E\{U\} = E\{\Phi(X_1, \cdots, X_m)\} = \theta
\]

Let

\[
\Phi_c(x_1, \cdots, x_c) = E\{\Phi(x_1, \cdots, x_c, X_{c+1}, \cdots, X_m)\}, \quad (c = 1, \cdots, m), \tag{2.6}
\]

where \(x_1, \cdots, x_c\) are arbitrary fixed vectors and the expected value is taken with respect to the random vectors \(X_{c+1}, \cdots, X_m\). Then

\[
\Phi_{c-1}(x_1, \cdots, x_{c-1}) = E\{\Phi(x_1, \cdots, x_{c-1}, X_c)\}, \tag{2.7}
\]
and

$$E \{ \Phi_c (X_1, \cdots, X_c) \} = \theta, \quad (c = 1, \cdots, m),$$  \hspace{1cm} (2.8)

Define

$$\Psi (x_1, \cdots, x_m) = \Phi (x_1, \cdots, x_m) - \theta, \quad (2.9)$$

$$\Psi_c (x_1, \cdots, x_m) = \Phi_c (x_1, \cdots, x_m) - \theta, \quad (c = 1, \cdots, m). \quad (2.10)$$

Then

$$\Psi_{c-1} (x_1, \cdots, x_{c-1}) = E \{ \Psi_c (x_1, \cdots, x_{c-1}, X_c) \}, \quad (2.11)$$

$$E \{ \Psi_c (X_1, \cdots, X_c) \} = E \{ \Psi (X_1, \cdots, X_m) \} = 0, \quad (c = 1, \cdots, m) \quad (2.12)$$

Suppose that the variance of \( \Psi_c (X_1, \cdots, X_c) \) exists, and let

$$\zeta_0 = 0, \quad \zeta_c = E \left\{ \Psi_c^2 (X_1, \cdots, X_c) \right\}, \quad (c = 1, \cdots, m) \quad (2.13)$$

Then

$$\zeta_c = E \left\{ \Phi_c^2 (X_1, \cdots, X_c) \right\} - \theta^2. \quad (2.14)$$

If \((\alpha_1, \cdots, \alpha_m)\) and \((\beta_1, \cdots, \beta_m)\) are two sets of \(m\) different integers, \(1 \leq \alpha_i, \beta_i \leq n\), and \(c\) is the number of integers common to the two sets, we have, by the symmetry
of $\Psi$,
\[
E\left\{\Psi\left(X_{\alpha_1}, \ldots, X_{\alpha_m}\right) \Psi\left(X_{\beta_1}, \ldots, X_{\beta_m}\right)\right\} = \zeta_c.
\] (2.15)

If the variance of $U$ exists, it is equal to
\[
\sigma^2(U) = \left(\frac{n}{m}\right)^{-2} E\left\{\sum \Psi\left(X_{\alpha_1}, \ldots, X_{\alpha_m}\right)^2\right\}
\]
\[
= \left(\frac{n}{m}\right)^{-2} \sum_{c=0}^{m} \sum \{\Psi\left(X_{\alpha_1}, \ldots, X_{\alpha_m}\right) \Psi\left(X_{\beta_1}, \ldots, X_{\beta_m}\right)\},
\]
where the summation $\sum$ is taken over all subscripts such that
\[
1 \leq \alpha_1 < \alpha_2 < \cdots < \alpha_m \leq n, \quad 1 \leq \beta_1 < \beta_2 < \cdots < \beta_m \leq n,
\]
and exactly $c$ equations
\[
\alpha_i = \beta_j
\]
are satisfied. By (2.15), each term in $\sum$ is equal to $\zeta_c$. The number of terms in $\sum$ is
\[
\left(\begin{array}{c} m \\ c \end{array}\right) \left(\begin{array}{c} n-m \\ m-c \end{array}\right) \left(\begin{array}{c} n \\ m \end{array}\right),
\]
and hence, since $\zeta_0 = 0$,
\[
\sigma^2(U) = \left(\frac{n}{m}\right)^{-1} \sum_{c=1}^{m} \left(\begin{array}{c} m \\ c \end{array}\right) \left(\begin{array}{c} n-m \\ m-c \end{array}\right) \zeta_c.
\] (2.16)
Consider the variance $\sigma^2(U_n)$ of a U-statistic $U_n = \mathbf{U} (X_1, X_2, \cdots, X_n)$, where $(X_1, X_2, \cdots, X_n)$ are independent and identically distributed. The quantities $\zeta_1, \cdots, \zeta_m$ are defined by (2.13). It was proven by Hoeffding (1948) that

$$\sigma^2(U_m) = \zeta_m, \quad \text{(2.17)}$$

$$\lim_{n \to \infty} n\sigma^2(U_n) = m^2\zeta_1. \quad \text{(2.18)}$$

### 2.3.3 The covariance of two U-statistics

Let

$$U^{(\gamma)} = \begin{pmatrix} n \\ m(\gamma) \end{pmatrix}^{-1} \sum' \Phi^{(\gamma)}(X_{\alpha_1}, \cdots, X_{\alpha_m(\gamma)}), \quad (\gamma = 1, \cdots, g),$$

be a set of $g$ U-statistics, each of which is a function of the same $n$ independent, identically distributed random vectors $X_1, X_2, \cdots, X_n$ with the function $\Phi^{(\gamma)}$ assuming to be
symmetric in its $m(\gamma)$ arguments ($\gamma = 1, \cdots, g$), and let

$$E\left\{ \Phi^{(\gamma)} \left( X_{\alpha_1}, \cdots, X_{\alpha_{m(\gamma)}} \right) \right\} = \theta^{(\gamma)}, \quad (\gamma = 1, \cdots, g);$$

$$\Psi^{(\gamma)} \left( x_{\alpha_1}, \cdots, x_{\alpha_{m(\gamma)}} \right) = \Phi^{(\gamma)} \left( x_{\alpha_1}, \cdots, x_{\alpha_{m(\gamma)}} \right) - \theta^{(\gamma)}, \quad (\gamma = 1, \cdots, g);$$

(2.19)

$$\Psi^{(\gamma)} \left( x_1, \cdots, x_c \right) = E\left\{ \Psi^{(\gamma)} \left( x_1, \cdots, x_c, X_{c+1}, \cdots, X_{m(\gamma)} \right) \right\};$$

(2.20)

$(c = 1, \cdots, m(\gamma); \gamma = 1, \cdots, g);$

$$\zeta^{(\gamma, \delta)} = E\left\{ \Psi^{(\gamma)} \left( X_1, \cdots, X_c \right) \Psi^{(\delta)} \left( X_1, \cdots, X_c \right) \right\}, \quad (\gamma, \delta = 1, \cdots, g).$$

(2.21)

If $\gamma = \delta$, we have

$$\zeta^{(\gamma)} = \zeta^{(\gamma, \gamma)} = E\left\{ \Psi^{(\gamma)} \left( X_1, \cdots, X_c \right) \right\}^2$$

(2.22)

Let

$$\sigma \left( U^{(\gamma)}, U^{(\delta)} \right) = E\left\{ \left( U^{(\gamma)} - \theta^{(\gamma)} \right) \left( U^{(\delta)} - \theta^{(\delta)} \right) \right\}$$

be the covariance of $U^{(\gamma)}$ and $U^{(\delta)}$.

Similar to the variance, we have, if $m(\gamma) \leq m(\delta),

$$\sigma \left( U^{(\gamma)}, U^{(\delta)} \right) = \left( \begin{array}{c} n \\ m(\gamma) \end{array} \right)^{-1} \sum_{c=1}^{m(\gamma)} \left( \begin{array}{c} m(\delta) \\ c \end{array} \right) \left( \begin{array}{c} n - m(\delta) \\ m(\gamma) - c \end{array} \right) \zeta^{(\gamma, \delta)}.$$

(2.23)

For $\gamma = \delta$, (2.23) is the variance of $U^{(\gamma)}$. 
From (2.18) and (2.23), we have

\[
\lim_{n \to \infty} n \sigma^2 \left( U^{(\gamma)} \right) = m^2(\gamma) \zeta_1^{(\gamma)},
\]
\[
\lim_{n \to \infty} n \sigma \left( U^{(\gamma)}, U^{(\delta)} \right) = m(\gamma) m(\delta) \zeta_1^{(\gamma, \delta)}.
\]

### 2.3.4 The asymptotic distribution of U-statistics

The following theorem was proven by Hoeffding (1948) and will be used in Chapter 3.

**Theorem 2.1.** Let \( \mathbf{X}_1, \ldots, \mathbf{X}_n \) be \( n \) independent, identically distributed random vectors,

\[
\mathbf{X}_\alpha = \left( X^{(1)}_{\alpha}, \ldots, X^{(r)}_{\alpha} \right), \quad (\alpha = 1, \ldots, n).
\]

Let

\[
\Phi^{(\gamma)} \left( x_1, \ldots, x_{m(\gamma)} \right), \quad (\gamma = 1, \ldots, g).
\]

be \( g \) real-valued functions not involving \( n \), \( \Phi^{(\gamma)} \) being symmetric in its \( m(\gamma) (\leq n) \) vector arguments \( \mathbf{x}_\alpha = \left( x^{(1)}_{\alpha}, \ldots, x^{(r)}_{\alpha} \right) \), \((\alpha = 1, \ldots, m(\gamma) ; \gamma = 1, \ldots, g \)). Define

\[
U^{(\gamma)} = \left( \begin{array}{c} n \\ m(\gamma) \end{array} \right)^{-1} \sum_{i} \Phi^{(\gamma)} \left( \mathbf{X}_{\alpha_1}, \ldots, \mathbf{X}_{\alpha_{m(\gamma)}} \right), \quad (\gamma = 1, \ldots, g), \quad (2.24)
\]
where the summation is over all subscripts such that $1 \leq \alpha_1 < \alpha_2 < \cdots < \alpha_m \leq n$.

Then, if the expected values

$$
\theta^{(\gamma)} = E \left\{ \Phi^{(\gamma)} \left( X_1, \ldots, X_{m(\gamma)} \right) \right\}, \quad (\gamma = 1, \cdots, g), \quad (2.25)
$$

and

$$
E \left\{ \Phi^{(\gamma)} \left( X_1, \ldots, X_{m(\gamma)} \right) \right\}^2, \quad (\gamma = 1, \cdots, g), \quad (2.26)
$$

exist, the joint distribution function of

$$
\sqrt{n} \left( U^{(1)} - \theta^{(1)} \right), \ldots, \sqrt{n} \left( U^{(g)} - \theta^{(g)} \right)
$$

tends, as $n \to \infty$, to the $g$-variate normal distribution function with zero means and covariance matrix $\left( m(\gamma)m(\delta) \zeta^{(\gamma,\delta)} \right)$, where $\zeta^{(\gamma,\delta)}$ is defined by $(2.21)$. The limiting distribution is non-singular if the determinant $|\zeta^{(\gamma,\delta)}|$ is positive.
Chapter 3

Multivariate Concordance Correlation Coefficient

3.1 Multivariate Concordance Correlation Coefficient

Let \((X, Y)\) be a \(2p \times 1\) random vector from a \(2p\)-variate distribution with a finite \(2p \times 1\) mean vector \((\mu_X, \mu_Y)\) and a positive definite \(2p \times 2p\) covariance matrix

\[
\begin{pmatrix}
\Sigma_{XX} & \Sigma_{XY} \\
\Sigma_{YX} & \Sigma_{YY}
\end{pmatrix}
\]

where

\[
E(X) = \mu_X, \quad E(Y) = \mu_Y, \quad \text{Var}(X) = \Sigma_{XX}, \quad \text{Var}(Y) = \Sigma_{YY}, \quad \text{and} \quad \text{Cov}(X, Y) = \Sigma_{XY}.
\]

To characterize the level of agreement between the two \(p \times 1\) vectors \(X\) and \(Y\), let us consider the following \(p \times p\) matrices,

\[
V_D = E((X - Y)(X - Y)')
\]

\[
= \Sigma_{XX} + \Sigma_{YY} - \Sigma_{XY} - \Sigma_{YX} + (\mu_X - \mu_Y)(\mu_X - \mu_Y)'
\]  

(3.1)
and

\[ V_1 = E_{\text{indep}} \{ (X - Y)(X - Y)' \} \]
\[ = \Sigma_{XX} + \Sigma_{YY} + (\mu_X - \mu_Y)(\mu_X - \mu_Y)' . \] (3.2)

Then, we construct a matrix version of multivariate concordance correlation coefficient denoted by \( M_{\rho_c} \), as follows

\[ M_{\rho_c} = V_1^{-1/2} (V_1 - V_D) V_1^{-1/2} \]
\[ = I_{p \times p} - V_1^{-1/2} V_D V_1^{-1/2} \] (3.3)

where \( I_{p \times p} \) denotes the \( p \times p \) identity matrix, \( V_D \) is non-negative definite, \( V_I \) is positive definite, and \( V_I^{-1/2} \) denotes the symmetric square-root decomposition of the inverse of \( V_I \).

For ease of notation, we write \( A > 0_{p \times p} \) if a \( p \times p \), symmetric matrix \( A \) is positive definite and \( A \geq 0_{p \times p} \) if \( A \) is non-negative definite.

\( M_{\rho_c} \) has the following properties:

1. \(-I_{p \times p} \leq M_{\rho_c} \leq I_{p \times p} \).

2. \( M_{\rho_c} = 0_{p \times p} \) if and only if \( \Sigma_{XY} = 0_{p \times p} \).

3. \( M_{\rho_c} = I_{p \times p} \) if and only if \( X = Y \) with probability one.

4. \( M_{\rho_c} = -I_{p \times p} \) if and only if \( X = -Y \) with probability one and \( \mu_X = \mu_Y = 0_{p \times 1} \).
5. If $p = 1$, then $M_{\rho_c}$ reduces to Lin’s CCC.

6. If $\Sigma_{XX}$ and $\Sigma_{YY}$ are diagonal matrices and $\mu_X = \mu_Y$, then each of the diagonal elements of $M_{\rho_c}$ corresponds to Lin’s CCC.

**Proof**  
1. First we will show that $I_{p \times p} - M_{\rho_c} \geq 0_{p \times p}$. By (3.3), it suffices to show that $V^{-1/2} DV^{-1/2} \geq 0_{p \times p}$.

Note that $V_D \geq 0_{p \times p}$, i.e. $x'V_D x \geq 0_{p \times p}$ for any vector $x$.

For any vector $y$, let $x = V^{-1/2}_1 y$. Thus,

$$y'V^{-1/2}_1 V_D V^{-1/2}_1 y = x'V_D x \geq 0_{p \times p}.$$ 

Next, we will show that $M_{\rho_c} + I_{p \times p} \geq 0_{p \times p}$. Note that

$$M_{\rho_c} + I_{p \times p} = 2I_{p \times p} - V^{-1/2}_1 V_D V^{-1/2}_1$$

$$= V^{-1/2}_1 (2V_1 - V_D) V^{-1/2}_1 \geq 0_{p \times p} \quad \text{if} \quad 2V_1 - V_D \geq 0_{p \times p}.$$
So it suffices to show that $2V_I - V_D \geq 0_{p \times p}$. But

$$2V_I - V_D = \Sigma_{XX} + \Sigma_{YY} + \Sigma_{XY} + \Sigma_{YX} + (\mu_X - \mu_Y)(\mu_X - \mu_Y)' \geq 0_{p \times p}$$

because $\begin{bmatrix} \Sigma_{XX} & \Sigma_{XY} \\ \Sigma_{YX} & \Sigma_{YY} \end{bmatrix}$ is non-negative definite, so that

$$\begin{bmatrix} I & I \\ \Sigma_{XX} & \Sigma_{XY} \\ \Sigma_{YX} & \Sigma_{YY} \end{bmatrix} \begin{bmatrix} I \\ I \end{bmatrix}$$

is non-negative definite, and $(\mu_X - \mu_Y)(\mu_X - \mu_Y)'$ is non-negative definite.

2. $V_I > 0_{p \times p}$, implying that $V_I^{-1/2} > 0_{p \times p}$, so we have that

$$M_{c} = V_I^{-1/2}(\Sigma_{XY} + \Sigma_{YX})V_I^{-1/2} = 0_{p \times p} \quad \text{if and only if} \quad \Sigma_{XY} = 0_{p \times p}.$$

3. By (3.3), we have

$$M_{c} = 0_{p \times p} \iff V_I^{-1/2}V_DV_I^{-1/2} = 0_{p \times p} \iff V_D = 0_{p \times p} \quad \text{(because} \quad V_I^{-1/2} > 0_{p \times p}) \iff X = Y \quad \text{with probability one.}$$
4. Note that

\[ M_{p_e} = -I_{p \times p} \iff I_{p \times p} + M_{p_e} = V_I^{-1/2} (2 V_I - V_D) V_I^{-1/2} = 0_{p \times p} \]

\[ \iff 2 V_I - V_D = \begin{bmatrix} I & I \\ \Sigma_{XX} & \Sigma_{XY} \\ \Sigma_{YX} & \Sigma_{YY} \end{bmatrix} \begin{bmatrix} I \\ I \end{bmatrix} \]

\[ \quad \quad + (\mu_X - \mu_Y)(\mu_X - \mu_Y)' = 0_{p \times p} \]

\[ \iff \Sigma_{XX} + \Sigma_{XY} = 0_{p \times p}, \Sigma_{YX} + \Sigma_{YY} = 0_{p \times p}, \text{ and } \]

\[ (\mu_X - \mu_Y)(\mu_X - \mu_Y)' = 0_{p \times p} \]

\[ \iff X = -Y \quad \text{with probability one} \quad \text{and} \quad \mu_X = \mu_Y = 0_{p \times 1}. \]

5. If \( p = 1 \) then \( V_I = E_{\text{indep}} (X - Y)^2 = \sigma_X^2 \sigma_Y^2 + 2 \sigma_{XY} (\mu_X - \mu_Y)^2 \) and \( V_D = E (X - Y)^2 = \sigma_X^2 \sigma_Y^2 - 2 \sigma_{XY} + (\mu_X - \mu_Y)^2 \). Thus,

\[ M_{p_e} = 1 - \frac{V_D}{V_I} = 1 - \frac{E (X - Y)^2}{E_{\text{indep}} (X - Y)^2} \]

\[ = \frac{2 \sigma_{XY}}{\sigma_X^2 + \sigma_Y^2 + (\mu_X - \mu_Y)^2} \]

which is Lin’s CCC.

6. If \( \Sigma_{XX} \) and \( \Sigma_{YY} \) are diagonal matrices and \( \mu_X = \mu_Y \), then \( V_I = \Sigma_{XX} + \Sigma_{YY} \) is a diagonal matrix with diagonal elements, \( \sigma_X^2 + \sigma_Y^2 \), \( i = 1, \ldots, p \) and hence \( V_I^{-1/2} \) is a diagonal matrix with diagonal elements, \( \sigma_{X_i}^2 + \sigma_{Y_i}^2 \), \( i = 1, \ldots, p \). Therefore
\[ M_{\rho_c} = I_{p \times p} - V_1^{-1/2} D V_1^{-1/2} \] has diagonal elements,

\[ 1 - \frac{\sigma^2_{X_i} + \sigma^2_{Y_i} - 2\sigma_{X_i Y_i}}{\sigma^2_{X_i} + \sigma^2_{Y_i}} = \frac{2\sigma_{X_i Y_i}}{\sigma^2_{X_i} + \sigma^2_{Y_i}}, \quad i = 1, \ldots, p \]

each of which corresponds to Lin’s CCC.

Based on these properties, we can use the multivariate CCC \( M_{\rho_c} \) to measure the amount of agreement between two vectors of random variables. The closer \( M_{\rho_c} \) is to the identity matrix, the higher the level of positive agreement between the two vectors. Conversely, the closer \( M_{\rho_c} \) is to the negative identity matrix, the higher the level of negative agreement between the two vectors. If \( M_{\rho_c} \) is equal to zero, it means that the two vectors are independent, in the other words, there is no agreement between the two vectors.

Note that \( M_{\rho_c} \) is a symmetric matrix whose eigenvalues are real and can be used to summarize the information in \( M_{\rho_c} \). As a result, we construct a scalar version of multivariate concordance correlation coefficient denoted by \( \rho_g \), by using a matrix function \( g \) to quantify the distance between \( M_{\rho_c} \) and identity matrix and then scale it to range between -1 and 1. This multivariate CCC \( \rho_g \) can be used to assess the amount of agreement between any two vectors as \( M_{\rho_c} \) but it is much easier to interpret and make inference, especially when \( p \) is large.

Let \( g \) be a matrix function satisfying the following properties:

1. if \( A, B \) are matrices such that \( A \geq B > 0 \), then \( g(A) \geq g(B) \), and

2. if \( b \) is a nonnegative constant, then \( g(bA) = bg(A) \).
Note that the matrix functions satisfying the first property are called \textit{the monotone matrix functions} and those matrix functions satisfying the second property are said to be \textit{positive homogeneous of degree 1} (Donoghue (1974)).

Define $\rho_g$ as

$$
\rho_g = 1 - \frac{g(I - M_{\rho_c})}{g(I)} = 1 - \frac{g(V_1^{-1/2}V_DV_1^{-1/2})}{g(I)}
$$

(3.4)

It can be proven that the following three functions of interest possess the above two properties.

$$
g_1(A) = \text{trace}(A)/p
$$

(3.5)

$$
g_2(A) = \text{the largest eigenvalue of } A
$$

$$
g_3(A) = |A|^{1/p}
$$

\textbf{Proof} We will first prove the property (1) for all three functions. Let $A, B$ be matrices such that $A \geq B > 0$. Then

$$
g_1(A) = \text{trace}(A)/p = \text{trace}(A - B)/p + \text{trace}(B)/p \geq \text{trace}(B)/p = g_1(B) \quad (\text{because } A - B \geq 0)
$$

$$
g_2(A) = \sup_{t \in \mathbb{R}^p, t' t = 1} \left\{ t' At \right\} = \sup_{t \in \mathbb{R}^p, t' t = 1} \left\{ t' (A - B)t + t' Bt \right\} \geq \sup_{t \in \mathbb{R}^p, t' t = 1} \left\{ t' Bt \right\} = g_2(B) \quad (\text{because } A - B \geq 0)
$$
\[ g_3(A) = |A|^{1/p} = |A - B + B|^{1/p} \]
\[ \geq (|A - B| + |B|)^{1/p} \quad \text{(by "Determinantal inequality", Abadir and Magnus (2005), p. 225)} \]
\[ \geq |A - B|^{1/p} + |B|^{1/p} \quad \text{(by "Triangle inequality")} \]
\[ \geq |B|^{1/p} = g_3(B). \]

Next, we will prove the second property for all three functions. Let \( b \) be a non-negative constant. Then

\[ g_1(bA) = \text{trace}(bA)/p = b \text{trace}(A)/p = bg_1(A) \]
\[ g_2(bA) = \sup_{t \in \mathbb{R}^p, t't = 1} \left\{ t'bAt \right\} = b \sup_{t \in \mathbb{R}^p, t't = 1} \left\{ t'At \right\} = bg_2(A) \]
\[ g_3(bA) = |bA|^{1/p} = b |A|^{1/p} = bg_3(A). \]
that there is no agreement between the two vectors, \( \rho_g = 1 \) if \( \mathbf{M}_{\rho_c} = \mathbf{I}_{p \times p} \), which is the case that there is a perfectly positive agreement between the two vectors, and \( \rho_g = -1 \) if \( \mathbf{M}_{\rho_c} = -\mathbf{I}_{p \times p} \), which is the case that there is a perfectly negative agreement between the two vectors. In addition, the closer \( \mathbf{M}_{\rho_c} \) is to \( \mathbf{I}_{p \times p} \), the closer \( \rho_g \) is to 1 and the closer \( \mathbf{M}_{\rho_c} \) is to \( -\mathbf{I}_{p \times p} \) the closer \( \rho_g \) is to -1.

To estimate \( \mathbf{M}_{\rho_c} \) and \( \rho_g \), we shall first estimate \( \mathbf{V}_D \) and \( \mathbf{V}_I \) with their unbiased estimators based on U-statistics.

Assume that \((\mathbf{X}_i, \mathbf{Y}_i), i = 1, \ldots, n\) are independent and identical-distributed random vectors from a \( 2p \)-variate distribution with finite fourth moments. Define

\[
\hat{\mathbf{V}}_D = \frac{1}{n} \sum_{i=1}^{n} (\mathbf{X}_i - \mathbf{Y}_i)(\mathbf{X}_i - \mathbf{Y}_i)'^{(3.8)}
\]

and

\[
\hat{\mathbf{V}}_I = \frac{1}{n(n-1)} \sum_{i \neq j} (\mathbf{X}_i - \mathbf{Y}_j)(\mathbf{X}_i - \mathbf{Y}_j)'^{(3.9)}
\]

Now, we construct the estimator of \( \mathbf{M}_{\rho_c} \) and \( \rho_g \) as

\[
\hat{\mathbf{M}}_{\rho_c} = \mathbf{I}_{p \times p} - \hat{\mathbf{V}}_I^{-1/2} \hat{\mathbf{V}}_D \hat{\mathbf{V}}_I^{-1/2}
\]

and

\[
\hat{\rho}_g = 1 - g \left( \hat{\mathbf{V}}_I^{-1/2} \hat{\mathbf{V}}_D \hat{\mathbf{V}}_I^{-1/2} \right)
\]
Let
\[
\Phi_{Dij}\{(X_i, Y_i), (X_j, Y_j)\} = \frac{1}{2} \{ \text{vec} \left[ (X_i - Y_i) (X_i - Y_i)' \right] + \text{vec} \left[ (X_j - Y_j) (X_j - Y_j)' \right] \}
\]
(3.12)

and
\[
\Phi_{ij}\{(X_i, Y_i), (X_j, Y_j)\} = \frac{1}{2} \{ \text{vec} \left[ (X_i - Y_j) (X_i - Y_j)' \right] + \text{vec} \left[ (X_j - Y_i) (X_j - Y_i)' \right] \}
\]
(3.13)

where the vec operator vectorizes a matrix by stacking its columns.

Note that
\[
\text{vec} \hat{V}_D = \frac{1}{n} \sum_{i=1}^{n} \left\{ \sum_{i \neq j}^{n} \Phi_{Dij} \right\} (n - 1)
\]
(3.14)

and
\[
\text{vec} \hat{V}_I = \frac{1}{n} \sum_{i=1}^{n} \left\{ \sum_{i \neq j}^{n} \Phi_{ij} \right\} (n - 1)
\]
(3.15)

That is, vec \( \hat{V}_D \) and vec \( \hat{V}_I \) are U-statistics with kernels \( \Phi_{Dij}\{(X_i, Y_i), (X_j, Y_j)\} \) and \( \Phi_{ij}\{(X_i, Y_i), (X_j, Y_j)\} \), respectively. Since \( \text{E} \left( \Phi_{Dij} \right) = \text{vec} \ V_D \) and \( \text{E} \left( \Phi_{ij} \right) = \text{vec} \ V_I \), vec \( \hat{V}_D \), vec \( \hat{V}_D \) and vec \( \hat{V}_I \) are unbiased estimators of vec \( V_D \) and vec \( V_I \), respectively.
3.2 Inference

To make inference about $\rho_g$, we will derive the asymptotic distribution of $\hat{\rho}_g$.

By (3.11), $\hat{\rho}_g$ is a function of $\hat{\mathbf{V}}^{-1/2}_I \hat{\mathbf{V}}_D \hat{\mathbf{V}}_I^{-1/2}$. Thus, we will first derive the limiting distribution of $\hat{\mathbf{V}}^{-1/2}_I \hat{\mathbf{V}}_D \hat{\mathbf{V}}_I^{-1/2}$.

**Theorem 3.1.** Assume that $(\mathbf{X}_1, \mathbf{Y}_1), \cdots, (\mathbf{X}_n, \mathbf{Y}_n)$ are independent and identically distributed random vectors from a $2p$-variate distribution with finite fourth moments.

Let $\text{vec} \hat{\mathbf{V}}_D$ and $\text{vec} \hat{\mathbf{V}}_I$ be defined as in (3.14) and (3.15), respectively. Then

$$\sqrt{n} \left\{ \text{vec} \left( \hat{\mathbf{V}}^{-1/2}_I \hat{\mathbf{V}}_D \hat{\mathbf{V}}_I^{-1/2} \right) - \text{vec} \left( \mathbf{V}_I^{-1/2} \mathbf{V}_D \mathbf{V}_I^{-1/2} \right) \right\} \xrightarrow{d} N_p(0, \Sigma^*) \quad (3.16)$$
where

\[ \Sigma^* = \Gamma \Sigma' \Gamma', \]

\[ \Gamma = \begin{bmatrix} \Gamma_1 & \Gamma_2 \end{bmatrix}, \]

\[ \Gamma_1 = V_I^{-1/2} \otimes V_I^{-1/2}, \]

\[ \Gamma_2 = -V_I^{-1} \otimes V_I^{-1}\left(V_I^{-1/2} \otimes I_p + I_p \otimes V_I^{-1/2}\right)^{-1}\left(V_D V_I^{-1/2} \otimes I_p + I_p \otimes V_D V_I^{-1/2}\right), \]

\[ \Sigma = \begin{pmatrix} \Sigma_{DD} & \Sigma_{DI} \\ \Sigma_{ID} & \Sigma_{II} \end{pmatrix}, \]

\[ \Sigma_{DD} = 4 \mathbb{E}\left\{ \Psi_{Di}(X_i, Y_i) \Psi_{Di}'(X_i, Y_i) \right\}, \]

\[ \Sigma_{DI} = 4 \mathbb{E}\left\{ \Psi_{Di}(X_i, Y_i) \Psi_{Ii}(X_i, Y_i) \right\}, \]

\[ \Sigma_{ID} = \Sigma_{DI}', \]

\[ \Sigma_{II} = 4 \mathbb{E}\left\{ \Psi_{Ii}(X_i, Y_i) \Psi_{Ii}'(X_i, Y_i) \right\}, \]

\[ \Psi_{Di}(x_i, y_i) = \mathbb{E}\left\{ \Phi_{Di}[(x_i, y_i), (X_j, Y_j)] - \text{vec} V_D \right\}, \]

\[ \Psi_{Ii}(x_i, y_i) = \mathbb{E}\left\{ \Phi_{Ii}[(x_i, y_i), (X_j, Y_j)] - \text{vec} V_I \right\}, \]

where \((x_i, y_i)\) is an arbitrary fixed vector and the expected value is taken with respect to the random vector \((X_j, Y_j), i < j\).

**Proof** Let

\[ \text{vec}\left(V_I^{-1/2} V_D V_I^{-1/2}\right) = h \left\{(\text{vec} V_D, \text{vec} V_I)\right\}' \] (3.17)
We will first derive the asymptotic distribution of \( \left( \text{vec } \hat{\mathbf{V}}_D, \text{vec } \hat{\mathbf{V}}_I \right)' \), and then use the theory on functions of asymptotically normal vectors (Serfling (1980), Theorem 3.3) to obtain the limiting distribution of \( \text{vec } \left( \hat{\mathbf{V}}_I^{-1/2} \hat{\mathbf{V}}_D \hat{\mathbf{V}}_I^{-1/2} \right) \).

It can be easily proven by using Theorem 2.1 that

\[
\sqrt{n} \left\{ \left( \text{vec } \hat{\mathbf{V}}_D, \text{vec } \hat{\mathbf{V}}_I \right)' - \left( \text{vec } \mathbf{V}_D, \text{vec } \mathbf{V}_I \right)' \right\} \xrightarrow{d} N_{2p^2}(0, \Sigma) \tag{3.18}
\]

where

\[
\Sigma = \begin{pmatrix}
\Sigma_{DD} & \Sigma_{DI} \\
\Sigma_{ID} & \Sigma_{II}
\end{pmatrix},
\]

and

\[
\Sigma_{DD} = 4 \mathbb{E} \left\{ \Psi_{Di} (\mathbf{X}_i, \mathbf{Y}_i) \Psi_{Di}' (\mathbf{X}_i, \mathbf{Y}_i) \right\},
\]

\[
\Sigma_{DI} = 4 \mathbb{E} \left\{ \Psi_{Di} (\mathbf{X}_i, \mathbf{Y}_i) \Psi_{Ii}' (\mathbf{X}_i, \mathbf{Y}_i) \right\},
\]

\[
\Sigma_{ID} = \Sigma_{DI}',
\]

\[
\Sigma_{II} = 4 \mathbb{E} \left\{ \Psi_{Ii} (\mathbf{X}_i, \mathbf{Y}_i) \Psi_{Ii}' (\mathbf{X}_i, \mathbf{Y}_i) \right\},
\]

where

\[
\Psi_{Di} (\mathbf{x}_i, \mathbf{y}_i) = \mathbb{E} \left\{ \Phi_{Dij} \left[ (\mathbf{x}_i, \mathbf{y}_i), (\mathbf{X}_j, \mathbf{Y}_j) \right] - \text{vec } \mathbf{V}_D \right\},
\]

\[
\Psi_{Ii} (\mathbf{x}_i, \mathbf{y}_i) = \mathbb{E} \left\{ \Phi_{Iij} \left[ (\mathbf{x}_i, \mathbf{y}_i), (\mathbf{X}_j, \mathbf{Y}_j) \right] - \text{vec } \mathbf{V}_I \right\},
\]
where \((x_i, y_i)\) is an arbitrary fixed vector and the expected value is taken with respect to the random vector \((X_j, X_j)\), \(i < j\).

Then, by the theory on functions of asymptotically normal vectors (Serfling (1980), Theorem 3.3), we have

\[
\sqrt{n} \left\{ \text{vec} \left( \hat{V}_I^{-1/2} \hat{V}_D \hat{V}_I^{-1/2} \right) - \text{vec} \left( V_I^{-1/2} V_D V_I^{-1/2} \right) \right\} \overset{d}{\to} N_p^2(0, \Gamma \Sigma \Gamma')
\]

where

\[
\Gamma = \begin{bmatrix}
\frac{\partial h}{\partial \text{vec} (V_D)} & \frac{\partial h}{\partial \text{vec} (V_I)} \\
\end{bmatrix}
\]

and

\[
h = \text{vec} \left( V_I^{-1/2} V_D V_I^{-1/2} \right)
\]
By the properties of matrix derivatives*, we have

\[
\frac{\partial h}{\partial \text{vec} (V_D)} = \frac{\partial V_1^{-1/2} V_D V_1^{-1/2}}{\partial V_D} = V_1^{-1/2} \otimes V_1^{-1/2}
\]

\[
\frac{\partial h}{\partial \text{vec} (V_I)} = \frac{\partial V_1^{-1/2} V_D V_I^{-1/2}}{\partial V_I}
\]

\[
= \frac{\partial V_1^{-1/2}}{\partial V_I} \left( V_D V_I^{-1/2} \otimes I_p \right) + \frac{\partial V_D V_I^{-1/2}}{\partial V_I} \left( I_p \otimes V_I^{-1/2} \right)
\]

\[
= \frac{\partial V_1^{-1/2}}{\partial V_I} \left( V_D V_I^{-1/2} \otimes I_p + I_p \otimes V_D V_I^{-1/2} \right)
\]

\[
= -V_1^{-1} \otimes V_1^{-1} \left( V_1^{-1/2} \otimes I_p + I_p \otimes V_1^{-1/2} \right)^{-1} \left( V_D V_I^{-1/2} \otimes I_p + I_p \otimes V_D V_I^{-1/2} \right).
\]

The last equation is true because we know that

\[
\frac{\partial V_1^{-1}}{\partial \text{vec} (V_I)} = -V_1^{-1} \otimes V_1^{-1}
\]

and

\[
\frac{\partial V_1^{-1}}{\partial \text{vec} (V_I)} = \frac{\partial V_1^{-1/2} V_1^{-1/2}}{\partial \text{vec} (V_I)}
\]

\[
= \frac{\partial V_1^{-1/2}}{\partial V_I} \left( V_1^{-1/2} \otimes I_p \right) + \frac{\partial V_1^{-1/2}}{\partial V_I} \left( I_p \otimes V_1^{-1/2} \right)
\]

\[
= \frac{\partial V_1^{-1/2}}{\partial V_I} \left( V_1^{-1/2} \otimes I_p + I_p \otimes V_1^{-1/2} \right).
\]

*See Kollo and Rosen (2005) p.148-149
Hence
\[
\frac{\partial V_{-1/2}}{\partial V} = -V_{-1} \otimes V_{-1} \left( V_{-1/2} \otimes I_p + I_p \otimes V_{-1/2} \right)^{-1}.
\]

Finally, we apply the theory on functions of asymptotically normal vectors (Serfling (1980), Theorem 3.3) to the result from the above theorem to obtain the asymptotic distribution of $\hat{\rho}_g$ as follows.

**Theorem 3.2.** Assume that $(X_1, Y_1), \cdots, (X_n, Y_n)$ are independent and identically distributed random vectors from a $2p$ - variate distribution with finite fourth moments.

Let $g_1, g_2,$ and $g_3$ and $\hat{\rho}_g$ be defined as in (3.5) and (3.11), respectively. Then

\[
\sqrt{n} \left( \hat{\rho}_{g_i} - \rho_{g_i} \right) \xrightarrow{d} N \left( 0, \frac{1}{p^2} \text{vec}' (I) \Sigma^* \text{vec}(I) \right)
\]

(3.19)

\[
\sqrt{n} \left( \hat{\rho}_{g_2} - \rho_{g_2} \right) \xrightarrow{d} N \left( 0, \text{vec}' (e_1 e_1') \Sigma^* \text{vec}(e_1 e_1') \right)
\]

(3.20)

\[
\sqrt{n} \left( \hat{\rho}_{g_3} - \rho_{g_3} \right) \xrightarrow{d} N \left( 0, \frac{1}{p^2} \left| V_{-1/2} V_D V_{-1/2} \right|^{2/3} \text{vec}' \left[ \left( V_{-1/2} V_D V_{-1/2} \right)^{-1} \right] \Sigma^* \text{vec} \left[ \left( V_{-1/2} V_D V_{-1/2} \right)^{-1} \right] \right)
\]

(3.21)

where $\Sigma^*$ is defined as in Theorem 3.1 and $e_1$ is the largest eigenvalue of $V_{-1/2} V_D V_{-1/2}$. 

**Proof** By applying the theory on functions of asymptotically normal vectors (Serfling (1980), Theorem 3.3) to (3.16), we have, for $i = 1, 2, 3$,

\[
\sqrt{n} \left( \hat{\rho}_{g_i} - \rho_{g_i} \right) \xrightarrow{d} N \left( 0, \frac{\partial g_i(X)}{d \text{vec}(X)} \right) \left. \times \right|_{X = V_{1/2} V_D V_{1/2}^{-1}} \Sigma^* \left. \frac{\partial g_i(X)}{d \text{vec}(X)} \right) \times = V_{1/2} V_D V_{1/2}^{-1}
\]

(3.22)
By the properties of matrix derivatives,

\[
\frac{dg_1(X)}{d \text{vec}(X)} = \frac{1}{p} \frac{d \text{trace}(X)}{d \text{vec}(X)} = \frac{1}{p} \text{vec}(I).
\]

Note that by the spectrum value decomposition for a non-negative definite matrix \(X\),

\[
X = EE', \quad E = \begin{bmatrix} e_1 & \cdots & e_p \end{bmatrix}_{p \times p}, \quad \Lambda = \begin{bmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \cdots & 0 \\ \vdots & \ddots & \ddots & \vdots \\ 0 & \cdots & 0 & \lambda_p \end{bmatrix}_{p \times p}
\]

so that

\[
g_2(X) = \lambda_1 = e_1'e_1
\]

Again by the properties of matrix derivatives,

\[
\frac{dg_2(X)}{d \text{vec}(X)} = \frac{d\lambda_1}{d \text{vec}(X)} = \text{vec}(e_1'e_1)
\]

and

\[
\frac{dg_3(X)}{d \text{vec}(X)} = \frac{d|X|^{1/p}}{d \text{vec}(X)} = \frac{1}{p} |X|^{1/p} \text{vec} \left( X^{-1} \right)'
\]

To obtain confidence intervals or test statistics for hypothesis testing about \(\rho_g\), we need to calculate the estimates of the parameters of the asymptotic variances in
(3.19), (3.20), or (3.21), depending on the function $g$ of interest. In addition to the estimates of $\vec{V}_D$ and $\vec{V}_I$, defined in (3.14) and (3.15), we need the estimate of the variance-covariance matrix $\Sigma$. According to Sen (1960), $\Sigma$ can be consistently estimated by

$$
\hat{\Sigma} = \frac{4}{n-1} \sum_{i=1}^{n-1} \left( \Phi_i^* - U \right) \left( \Phi_i^* - U \right)'
$$

(3.23)

where

$$
\Phi_i^* = (\Phi_{Di}, \Phi_{Ii})', \quad \Phi_{Di} = \sum_{i<j} \frac{\Phi_{Dij}}{(n-i)}, \quad \Phi_{Ii} = \sum_{i<j} \frac{\Phi_{Iij}}{(n-i)}
$$

and

$$
U = \left( \vec{V}_D, \vec{V}_I \right)'.
$$

As shown in the paper by Lin (1989), the normal approximation of Lin’s CCC can be improved by using the inverse hyperbolic tangent transformation or $Z$-transformation. Confirmed by Monte Carlo study in Lin’s paper, the $Z$-transformation can accelerate the convergence of normality of the sample CCC not only when the samples are from the normal distribution but also when the samples are from the short-tailed symmetric distribution like uniform and the long-tailed skewed-to-the-right distribution like Poisson. The $Z$-transformation was also shown to effectively improve the normality approximation of the sample repeated measure CCC for both normal and contaminated normal data in the paper by King et al. (2007).
To improve the asymptotic normality of our sample multivariate CCC \((\hat{\rho}_g)\), we will also consider using the following Z-transformation for inference about the multivariate CCC, \(\rho_g\).

Let \(\hat{Z}\) be

\[
\hat{Z} = \tan^{-1}(\hat{\rho}_g) = \frac{1}{2} \ln \frac{1 + \hat{\rho}_g}{1 - \hat{\rho}_g}.
\]

(3.24)

Then it follows from the theory on functions of asymptotically normal statistics (Serfling (1980), Theorem 3.1) that \(\hat{Z}\) is asymptotically normal with mean

\[
Z = \frac{1}{2} \ln \frac{1 + \rho_g}{1 - \rho_g}
\]

and variance

\[
\sigma^2_{\hat{Z}} = \frac{1}{n} \frac{\text{Var}(\hat{\rho}_g)}{(1 - \rho^2_g)^2}.
\]

By replacing the parameters in the variance of \(\hat{Z}\) with their estimates, we can obtain the confidence interval for \(Z\) denoted by \((\hat{Z}_U, \hat{Z}_L)\) and then by transformation we can obtain the confidence interval for \(\rho_g\) based on the Z-transformation as follows

\[
\left( \frac{\exp(2\hat{Z}_L) - 1}{\exp(2\hat{Z}_L) + 1}, \frac{\exp(2\hat{Z}_U) - 1}{\exp(2\hat{Z}_U) + 1} \right)
\]

As noted by Lin (1989), the confidence interval of the multivariate CCC based on the Z-transformation will be bounded in the open interval (-1,1) and more realistic asymmetric.
Chapter 4

Monte Carlo Simulation

To assess the finite sample properties of the sample multivariate CCC, \( \hat{\rho}_{gi} \) and the corresponding Z-transformation, \( \hat{Z}_i \), \( i = 1, 2, 3 \), as described in Section 3.2, we performed a Monte Carlo simulation for three cases with different combinations of location and scale shifts and levels of correlation between X and Y. In each case, we consider three and five repeated measurements per unit for three levels of within-unit correlation \( (\rho = 0, 0.4, 0.8) \) with sample sizes of \( n = 20, n = 40, n = 80, \) and \( n = 160 \). For each of the 72 situations, 1,000 runs were performed. In Section 4.1, the repeated measures paired samples were generated from the multivariate normal distribution. In Section 4.2 and 4.3, we generated the data from multivariate Student’s t-distribution and multivariate lognormal distribution, respectively. The scenarios considered here for this simulation study are similar to those considered by King et al. (2007).

4.1 Multivariate normal distribution

In this section, five repeated measures paired samples were generated from each of the following cases of the multivariate normal distribution using SAS/IML software.
**Case 1:** Means $\mu_X = (4, 6, 8, 10, 12)$ and $\mu_Y = (5, 7, 9, 11, 13)$ and covariance matrix $\Lambda_1 \otimes \Lambda_2$ where

$$
\Lambda_1 = \begin{pmatrix}
8 & 0.95 \times \sqrt{8} \times \sqrt{10} \\
0.95 \times \sqrt{8} \times \sqrt{10} & 10
\end{pmatrix}
$$

and $\Lambda_2$ is a $5 \times 5$ compound symmetric within-unit correlation structure with $\rho = 0.4$, assuming repeated measures have equal variance for both $X$ and $Y$. This case represents a slight difference in location and scale parameters, and strong positive correlation between $X$ and $Y$.

**Case 2:** Means $\mu_X = (4, 6, 8, 10, 12)$ and $\mu_Y = (6, 8, 10, 12, 14)$ and covariance matrix $\Lambda_1 \otimes \Lambda_2$ where

$$
\Lambda_1 = \begin{pmatrix}
8 & 0.80 \times \sqrt{8} \times \sqrt{12} \\
0.80 \times \sqrt{8} \times \sqrt{12} & 12
\end{pmatrix}
$$

and $\Lambda_2$ is a $5 \times 5$ compound symmetric within-unit correlation structure with $\rho = 0.4$, assuming repeated measures have equal variance for both $X$ and $Y$. This case represents a moderate difference in location and scale parameters, and moderate positive correlation between $X$ and $Y$. 
Case 3: Means $\mu_X = (4, 6, 8, 10, 12)$ and $\mu_Y = (7, 9, 11, 13, 15)$ and covariance matrix $\Lambda_1 \otimes \Lambda_2$ where

$$
\Lambda_1 = \begin{pmatrix}
8 & 0.5 \times \sqrt{8} \times \sqrt{15} \\
0.5 \times \sqrt{8} \times \sqrt{15} & 15
\end{pmatrix}
$$

and $\Lambda_2$ is a $5 \times 5$ compound symmetric within-unit correlation structure with $\rho = 0.4$, assuming repeated measures have equal variance for both X and Y. This case represents a large difference in location and scale parameters, and weaker positive correlation between X and Y.

Then, the three repeated measures paired samples were generated from the same three cases of the multivariate normal distribution using the first three observations in $\mu_X$ and $\mu_Y$ and a $3 \times 3$ compound symmetric within-unit correlation structure. In addition, all six situations were repeated with $\rho = 0$ and 0.8 instead of 0.4.

In each run, we calculated $\hat{\rho}_{gi}$, $\hat{Z}_i$, their estimated asymptotic variances, and the 95% confidence intervals of $\rho_{gi}$ and $Z_i$, $i = 1, 2, 3$ as described in Section 3.2. Based on 1,000 runs, for each scenario and each estimator, we evaluated the normality, accuracy, precision, and coverage probability of the confidence interval. To assess the normality, we examined Q-Q plots of each estimator $\hat{\rho}_{gi}$ and $\hat{Z}_i$. To evaluate accuracy, we calculated the mean of the estimates of $\hat{\rho}_{gi}$ and then computed its average relative bias by subtracting the mean from the true parameter value and then dividing the result with the true parameter value. To assess precision, we calculated the mean of the estimated asymptotic variances and the empirical variance for each estimator $\hat{Z}_i$. To determine coverage
probability of the confidence interval, we calculated the empirical coverage probabilities for the 95% confidence intervals of each multivariate CCC, $\rho_{gi}$.

The Q-Q plots of the estimates of $\hat{\rho}_{gi}$ and $\hat{Z}_i$, $i = 1, 2, 3$, based on simulation of 1000 runs from the multivariate normal distribution for the scenario with three repeated measures and high within-unit correlation, are shown in Figure 4.1 - 4.12. With strong or moderate correlation (Case 1 and 2), the distribution of $\hat{Z}_i$ is much closer to normality than that of $\hat{\rho}_{gi}$ for all $i = 1, 2, 3$ and for all sample sizes. With weaker correlation (Case 3), the distribution of $\hat{\rho}_{gi}$ is closer to normality for all $i = 1, 2, 3$ and all sample sizes, especially when sample sizes are large ($n \geq 80$), in which the distribution of $\hat{\rho}_{gi}$ and $\hat{Z}_i$ are very close to normality for all $i = 1, 2, 3$. However, in Case 3, when sample sizes are small ($n \leq 40$), the distribution of $\hat{\rho}_{g1}$ and $\hat{\rho}_{g2}$ are still improved by the Z-transformation. The Q-Q plots are similar for the scenarios with zero and moderate within-unit correlation so the results are not shown. In addition, the Q-Q plots for the case of five repeated measures are similar to those for the case of three repeated measures (results are not shown).

The mean of estimates of $\hat{\rho}_g$, relative bias, the estimated asymptotic variances $\hat{\sigma}_Z^2$, and the empirical variance of $\hat{Z}_i$, $i = 1, 2, 3$, based on simulation of 1000 runs from the multivariate normal distribution, are shown in Table 4.1 for three repeated measurements and in Table 4.2 for five repeated measurements. For all scenarios, the magnitude of the average relative bias of $\hat{\rho}_{gi}$ decreases as sample size increases for all $i = 1, 2, 3$. In all situations, the estimator of the multivariate CCC based on the determinant function $\left(\hat{\rho}_{g_3}\right)$ has the lowest magnitude of the average relative bias whereas the estimator of the multivariate CCC based on the largest eigenvalue function $\left(\hat{\rho}_{g_2}\right)$ has the highest
magnitude of the average relative bias. In most scenarios, $\hat{\rho}_{g3}$, the estimator of the multivariate CCC based on the determinant function, has much lower absolute value of the average relative bias than $\hat{\rho}_{g1}$, the estimator of the multivariate CCC based on the trace function. However, the magnitude of the average relative bias of $\hat{\rho}_{g1}$ is closer to that of $\hat{\rho}_{g3}$ when sample sizes are larger. With strong or moderate correlation (Case 1 and 2) for all sample sizes, the absolute value of the average relative bias of $\hat{\rho}_{g3}$ is less than 0.5% when $p = 3$ and less than 1% when $p = 5$. For weak correlation (Case 3), the absolute values of the average relative of $\hat{\rho}_{g3}$ is less than 5% for all sample sizes when $p = 3$ and not more than 6% for all sample sizes when $p = 5$. With strong correlation (Case 1), the magnitude of the average relative bias of $\hat{\rho}_{g1}$ is less than 5% for all sample sizes and closer to that of $\hat{\rho}_{g3}$ when sample sizes are at least 80. For moderate or weaker correlation (Case 2 and 3), the absolute value of the average relative bias of $\hat{\rho}_{g1}$ is less than or equal to 5.3% when $n \geq 40$ and closer to that of $\hat{\rho}_{g3}$ when $n \geq 160$. Overall, in terms of the magnitude of the average relative bias, $\hat{\rho}_{g3}$ performs relatively well in all situations; however when sample sizes are at least 160, $\hat{\rho}_{g1}$ performs as well as $\hat{\rho}_{g3}$. To evaluate precision, we compared the mean of the estimated asymptotic variances of $\hat{Z}_i$ with the empirical variance of $\hat{Z}_i$ for $i = 1, 2, 3$ in each situation. For all scenarios, we found that the average asymptotic variance estimates of $\hat{Z}_1$ and $\hat{Z}_3$ were very close to the empirical variances. The means of the estimated asymptotic variances of $\hat{Z}_2$ were quite close to the empirical variances for most situations except for $n = 20$ in which they tended to somewhat overestimate.
The empirical coverage probabilities for the 95% confidence intervals of $\rho_g$, based on simulation of 1000 runs generated from the multivariate normal distribution, are reported in Table 4.3. For all three multivariate CCC, the coverage probabilities for the 95% confidence intervals of $\rho_g$ based on the Z-transformation are closer to 0.95 as sample size increases in all scenarios. For most situations, in terms of the coverage probabilities, the estimator of the multivariate CCC based on the determinant function $\hat{\rho}_{g3}$ performs the best and the estimator of the multivariate CCC based on the trace function $\hat{\rho}_{g1}$ performs better than that of the multivariate CCC based on the largest eigenvalue function $\hat{\rho}_{g2}$. However, when sample sizes are larger the coverage probabilities for all estimators became closer to 0.95. The coverage probabilities for $\hat{\rho}_{g3}$ are very close to 0.95 when $n \geq 40$ for $p = 3$ and when $n \geq 80$ for $p = 5$ in all cases. With strong or moderate correlation (Case 1 and 2), the coverage probabilities for $\hat{\rho}_{g1}$ are close to 0.95 when $n \geq 80$ for $p = 3$ and when $n = 160$ for $p = 5$. With weak correlation (Case 3), the coverage probabilities for $\hat{\rho}_{g1}$ are close to 0.95 when $n \geq 40$ for $p = 3$ and when $n \geq 80$ for $p = 5$. The coverage probabilities for $\hat{\rho}_{g2}$ are close to 0.95 when $n = 160$ for Case 1 and Case 2. Overall, in terms of the coverage probabilities, when sample size is at least 160, $\hat{\rho}_{g1}$ performs as well as $\hat{\rho}_{g3}$ for all cases and all the estimators perform comparably well when the correlation between $X$ and $Y$ is strong or moderate.

In general, based on the accuracy, precision, and coverage probabilities, when the data are from the multivariate normal distribution, the estimator of the multivariate CCC based on the determinant function, $\hat{\rho}_{g3}$, performs notably well in most scenarios. However, when the sample size is at least 160, $\hat{\rho}_{g1}$, the estimator of the multivariate CCC based on the trace function, performs as well as $\hat{\rho}_{g3}$.
4.2 Multivariate Student’s t-distribution

In this section, three and five repeated measures paired samples were generated from the multivariate Student’s t-distribution with 10 degrees of freedom using the same scenarios as in the case of the multivariate normal distribution. As in Section 4.1, for each scenario and each estimator, we evaluated the normality, accuracy, precision, and coverage probability of the confidence interval based on simulation of 1,000 runs.

The Q-Q plots of the estimates of $\hat{\rho}_{gi}$ and $\hat{Z}_i$, $i = 1, 2, 3$, based on simulation of 1000 runs from multivariate Student’s t-distribution for the scenario with three repeated measures and high within-unit correlation, are shown in Figure 4.13 - 4.24. As in the normal case, with strong or moderate correlation (Case 1 and 2), the distribution of $\hat{Z}_i$ is much closer to normality than that of $\hat{\rho}_{gi}$ for all $i = 1, 2, 3$ and for all sample sizes. In Case 3, with weaker correlation, when sample sizes are large, the distribution of $\hat{\rho}_{gi}$ is very close to normality for all $i = 1, 2, 3$, but when sample sizes are small, only the distribution of $\hat{\rho}_{g1}$ and $\hat{\rho}_{g3}$ are improved by the Z-transformation. The results are similar for the scenario with zero or moderate within-unit correlation and with five repeated measures (results are not shown).

The mean of estimates of $\hat{\rho}_{gi}$, relative bias, the estimated asymptotic variances and the empirical variance of $\hat{Z}_i$, $i = 1, 2, 3$, based on simulation of 1000 runs from the multivariate Student’s t-distribution, are shown in Table 4.4 for three repeated measurements and in Table 4.5 for five repeated measurements. As when the underlying distribution is multivariate normal, the magnitude of the average relative bias of $\hat{\rho}_{gi}$ decreases as sample size increases for all $i = 1, 2, 3$ in all situations. For all scenarios,
\( \hat{\rho}_{g3} \), the estimator of the multivariate CCC based on the determinant function, performs the best, and \( \hat{\rho}_{g1} \), the estimator of the multivariate CCC based on the trace function, performs better than \( \hat{\rho}_{g2} \), the estimator of the multivariate CCC based on the largest eigenvalue function, in terms of the magnitude of the average relative bias. In most scenarios, the magnitude of the average relative bias of \( \hat{\rho}_{g3} \) are much lower than those of the others. However, the absolute value of the average relative bias of \( \hat{\rho}_{g1} \) is closer to that of \( \hat{\rho}_{g3} \) when sample size is larger. As in the normal case, with strong or moderate correlation (Case 1 and 2), the magnitude of the average relative bias of \( \hat{\rho}_{g3} \) is less than 0.5% for \( p = 3 \) and less than 1% for \( p = 5 \) for all sample sizes. For Case 3, with weak correlation, the absolute values of the average relative bias of \( \hat{\rho}_{g3} \) is less than 5% for all sample sizes when \( p = 3 \) but when \( p = 5 \) they are less than 5% for \( n \geq 40 \). With strong correlation (Case 1), the absolute value of the average relative bias of \( \hat{\rho}_{g1} \) is lower than 5% for all sample sizes and closer to that of \( \hat{\rho}_{g3} \) when \( n \geq 80 \). For Case 2 and 3, with moderate and weaker correlation, the magnitude of the average relative bias of \( \hat{\rho}_{g1} \) is at most 6.3% when \( n \geq 40 \) and closer to that of \( \hat{\rho}_{g3} \) when \( n \geq 160 \). As in the normal case, \( \hat{\rho}_{g3} \) performs much better than the others in terms of the magnitude of the average relative bias except when the sample size is 160 or more in which \( \hat{\rho}_{g1} \) is comparable to \( \hat{\rho}_{g3} \). To evaluate precision, we compared the mean of the estimated asymptotic variances of \( \hat{Z}_i \) with the empirical variance of \( \hat{Z}_i \) for \( i = 1, 2, 3 \) in each situation. Similar to the normal case, we found that the means of the estimated asymptotic variances of \( \hat{Z}_1 \) and \( \hat{Z}_3 \) were very close to the empirical variances in all situations. For most scenarios, the average asymptotic variance estimates of \( \hat{Z}_2 \) were adequately close to the empirical variances except for \( n = 20 \), in which they tended to slightly overestimate.
The empirical coverage probabilities for the 95% confidence intervals of $\rho_g$ based on simulation of 1000 runs generated from the multivariate Student’s t-distribution are shown in Table 4.6. As in the normal case, the coverage probabilities for the 95% confidence intervals of $\rho_{gi}$ based on the Z-transformation are closer to 0.95 as sample size increases in all scenarios for all $i = 1, 2, 3$. For most situations, in terms of the coverage probabilities, $\hat{\rho}_{g3}$ performs the best except for $n \geq 80$ when $p = 3$ and $n \geq 160$ when $p = 5$, in which $\hat{\rho}_{g1}$ performs as well as $\hat{\rho}_{g3}$ for all cases and for Case 3, respectively.

Overall, based on the accuracy, precision, and coverage probabilities, when the data are from the multivariate Student’s t-distribution, the estimator of the multivariate CCC based on the determinant function, $\hat{\rho}_{g3}$, performs relatively well in most scenarios. However, when the sample size is at least 160 and the correlation is weak, the estimator of the multivariate CCC based on the trace function, $\hat{\rho}_{g1}$, is comparable to $\hat{\rho}_{g3}$.

### 4.3 Multivariate lognormal distribution

In this section, five repeated measures paired samples were generated from each of the following cases of the multivariate normal distribution and then transformed to multivariate lognormal distribution using SAS/IML software.

**Case 1:** Means $\mu_X = (1.4, 1.75, 2.0, 2.2, 2.35)$ and $\mu_Y = (1.6, 1.9, 2.15, 2.33, 2.45)$ and covariance matrix $\Lambda_1 \otimes \Lambda_2$ where

$$
\Lambda_1 = \begin{pmatrix}
0.2 & 0.95 \times \sqrt{0.2} \times \sqrt{0.25} \\
0.95 \times \sqrt{0.2} \times \sqrt{0.25} & 0.25
\end{pmatrix}
$$
and Λ₂ is a 5 × 5 compound symmetric within-unit correlation structure with \( \rho = 0.4 \), assuming repeated measures have equal variance for both X and Y. This case represents a slight difference in location and scale parameters, and strong positive correlation between X and Y.

**Case 2:** Means \( \mu_X = (1.4, 1.75, 2.0, 2.3, 2.55) \) and \( \mu_Y = (1.75, 2.0, 2.3, 2.55, 2.8) \) and covariance matrix \( \Lambda_1 \otimes \Lambda_2 \) where

\[
\Lambda_1 = \begin{pmatrix}
0.2 & 0.2 \times \sqrt{0.2} \times \sqrt{0.35} \\
0.2 \times \sqrt{0.2} \times \sqrt{0.35} & 0.35
\end{pmatrix}
\]

and \( \Lambda_2 \) is a 5 × 5 compound symmetric within-unit correlation structure with \( \rho = 0.4 \), assuming repeated measures have equal variance for both X and Y. This case represents a moderate difference in location and scale parameters, and moderate positive correlation between X and Y.

**Case 3:** Means \( \mu_X = (1.4, 1.75, 2.0, 2.35, 2.6) \) and \( \mu_Y = (1.9, 2.2, 2.4, 2.7, 2.95) \) and covariance matrix \( \Lambda_1 \otimes \Lambda_2 \) where

\[
\Lambda_1 = \begin{pmatrix}
0.2 & 0.5 \times \sqrt{0.2} \times \sqrt{0.5} \\
0.5 \times \sqrt{0.2} \times \sqrt{0.5} & 0.5
\end{pmatrix}
\]

and \( \Lambda_2 \) is a 5 × 5 compound symmetric within-unit correlation structure with \( \rho = 0.4 \), assuming repeated measures have equal variance for both X and Y. This case represents a large difference in location and scale parameters, and weaker positive correlation between X and Y.
Then, the three repeated measures paired samples were generated from the same three cases of the multivariate lognormal distribution using the first three observations in $\mu_X$ and $\mu_Y$ and a $3 \times 3$ compound symmetric within-unit correlation structure. In addition, all six situations were repeated with $\rho = 0$ and 0.8 instead of 0.4.

As in Section 4.1 and 4.2, for each scenario and each estimator, we assessed the normality, accuracy, precision, and coverage probability of the confidence interval based on simulation of 1,000 runs.

The Q-Q plots of the estimates of $\hat{\rho}_{gi}$ and $\hat{Z}_i$, $i = 1, 2, 3$, based on simulation of 1000 runs from the multivariate lognormal distribution for the scenario with three repeated measures and high within-unit correlation, are shown in Figure 4.25 - 4.36. Unlike the former underlying distributions, only with strong correlation (Case 1), the distribution of $\hat{\rho}_{gi}$ is improved by the Z-transformation for all $i = 1, 2, 3$ and for all sample sizes. In Case 2 and 3, with moderate and weaker correlation, the distribution of $\hat{Z}_i$ is as close to normality as that of $\hat{\rho}_{gi}$ for all $i = 1, 2, 3$. Moreover, the distribution of $\hat{\rho}_{g1}$ and $\hat{\rho}_{g3}$ are closer to normality than that of $\hat{\rho}_{g2}$ for Case 2 and 3. The results are similar for the scenario with zero or moderate within-unit correlation and with five repeated measures (results are not shown).

The mean of estimates of $\hat{\rho}_{gi}$, relative bias, the estimated asymptotic variances and the empirical variance of $\hat{Z}_i$, $i = 1, 2, 3$, based on simulation of 1000 runs from the multivariate lognormal distribution, are reported in Table 4.7 for three repeated measurements and in Table 4.8 for five repeated measurements. As before, the magnitude of the average relative bias of $\hat{\rho}_{gi}$ decreases as the sample size increases for all $i = 1, 2, 3$ in all situations. With strong correlation (Case 1), $\hat{\rho}_{g3}$, the estimator of the multivariate
CCC based on the determinant function, has the least magnitude of the average relative bias and \( \hat{\rho}_{g1} \), the estimator of the multivariate CCC based on the trace function, has lower magnitude of the average relative bias than \( \hat{\rho}_{g2} \), the estimator of the multivariate CCC based on the largest eigenvalue function. However, for Case 1, the absolute value of the average relative bias of \( \hat{\rho}_{g1} \) is closer to that of \( \hat{\rho}_{g3} \) as \( n \) increases. In Case 1, the magnitude of the average relative bias of \( \hat{\rho}_{g3} \) is less than 0.3% when \( p = 3 \) and less than 1% when \( p = 5 \) for all sample sizes, whereas the absolute value of the average relative bias of \( \hat{\rho}_{g1} \) is less than 5% for all sample sizes and closer to that of \( \hat{\rho}_{g3} \) for \( n \geq 160 \). With moderate or weak correlation (Case 2 and 3), \( \hat{\rho}_{g1} \) become the best with the magnitude of the average relative bias less than 10% for all sample sizes; however, \( \hat{\rho}_{g3} \) is still better than \( \hat{\rho}_{g2} \). Again, to evaluate precision, we compared the mean of the estimated asymptotic variances of \( \hat{Z}_i \) with the empirical variance of \( \hat{Z}_i \) for \( i = 1, 2, 3 \) in each situation. As in the former underlying distributions, we found that the average asymptotic variance estimates of \( \hat{Z}_1 \) and \( \hat{Z}_3 \) were considerably close to the empirical variances for all scenarios. The means of the estimated asymptotic variances of \( \hat{Z}_2 \) were significantly close to the empirical variances for most situations except for \( n = 20 \) in which they tended to slightly overestimate.

The empirical coverage probabilities for the 95% confidence intervals of \( \rho_g \), based on simulation of 1000 runs generated from the multivariate lognormal distribution, are shown in Table 4.6. As for the other parent distributions, the coverage probabilities for the 95% confidence intervals of \( \rho_{gi} \) based on the Z-transformation are closer to 0.95 as the sample size increases in all scenarios for all \( i = 1, 2, 3 \). With strong correlation (Case 1), when the sample size is small, \( \hat{\rho}_{g3} \), the estimator of the multivariate CCC based on
the determinant function, tends to have the coverage probabilities closest to 0.95 and \( \hat{\rho}_{g1} \), the estimator of the multivariate CCC based on the trace function, tends to have the coverage probabilities closer to 0.95 than those of \( \hat{\rho}_{g2} \), the multivariate CCC based on the largest eigenvalue function. However, when the sample size is larger, all estimators have the coverage probabilities closer to each other and closer to 0.95. For Case 2 and 3, with moderate and weaker correlation, \( \hat{\rho}_{g1} \) performs the best and \( \hat{\rho}_{g3} \) performs better than \( \hat{\rho}_{g2} \) for all sample sizes except when \( n = 160 \) and \( p = 3 \), in which \( \hat{\rho}_{g3} \) performs as well as \( \hat{\rho}_{g1} \) in Case 3.

In general, based on the accuracy, precision, and coverage probabilities, when the data are from the multivariate lognormal distribution, the estimator of the multivariate CCC based on the determinant function \( \hat{\rho}_{g3} \) performs comparatively well for strong correlation (Case 1) whereas the estimator of the multivariate CCC based on the trace function \( \hat{\rho}_{g1} \) performs the best for moderate or weaker correlation. However, for Case 1, when the sample sizes are at least 160, \( \hat{\rho}_{g1} \) and \( \hat{\rho}_{g3} \) are of comparable performances.

4.4 Summary

To assess and compare the properties of the three estimators of the multivariate CCC in case of finite samples, the simulation study has been conducted using three and five repeated measurements per unit with four different sample sizes under three different underlying distributions, three different levels of correlation between \( X \) and \( Y \), and three different levels of within-unit correlation. It was found that the normality of \( \hat{\rho}_{g1} \) and \( \hat{\rho}_{g3} \), the estimators based on the trace and determinant functions, are improved by the Z-transformation for all sample sizes and all levels of correlation when the underlying
distribution is multivariate normal or multivariate Student’s t, but only for the case of strong correlation between $X$ and $Y$ when the data are from the multivariate lognormal distribution. The distribution of $\hat{\rho}_{g2}$, the estimator based on the largest eigenvalue function, is closer to normality by using the Z-transformation for only large sample sizes ($n \geq 80$) in the normal and Student’s t case and when $X$ and $Y$ have strong correlation in the lognormal case. Moreover, we found that overall in terms of the accuracy, precision, and coverage probabilities, the estimator $\hat{\rho}_{g3}$ performs relatively well for all scenarios under each underlying distribution except for moderate or weaker correlation in the lognormal case, in which $\hat{\rho}_{g1}$ works the best for all sample sizes. Also, when the sample sizes are as large as 160, $\hat{\rho}_{g1}$ is comparable to $\hat{\rho}_{g3}$ for all cases under the normal or Student’s t distribution and for strong correlation between $X$ and $Y$ under the lognormal distribution. This suggests that the proposed test statistics or the confidence intervals based on the Z-transformation for finite samples are more suitable for making inference about the multivariate CCC based on the trace or determinant function than the largest eigenvalue function in all scenarios. This also indicates that the distributions of the estimators of the multivariate CCC based on the trace and determinant functions converge to their asymptotic normality faster than that based on the largest eigenvalue function.
Fig. 4.1: Q-Q plot of the estimates of $\hat{\rho}_g$ and $\hat{Z}$ for sample size 20 based on simulation of 1000 runs from the multivariate normal distribution with three repeated measure, high within-unit correlation($\rho = 0.8$), and strong between-unit correlation (Case 1).
Fig. 4.2: Q-Q plot of the estimates of $\hat{\rho}_g$ and $\hat{Z}$ for sample size 40 based on simulation of 1000 runs from the multivariate normal distribution with three repeated measure, high within-unit correlation ($\rho = 0.8$), and strong between-unit correlation (Case 1).
Fig. 4.3: Q-Q plot of the estimates of $\hat{\rho}_g$ and $\hat{Z}$ for sample size 80 based on simulation of 1000 runs from the multivariate normal distribution with three repeated measure, high within-unit correlation ($\rho = 0.8$), and strong between-unit correlation (Case 1).
Fig. 4.4: Q-Q plot of the estimates of $\hat{\rho}_g$ and $\hat{Z}$ for sample size 160 based on simulation of 1000 runs from the multivariate normal distribution with three repeated measure, high within-unit correlation ($\rho = 0.8$), and strong between-unit correlation (Case 1).
Fig. 4.5: Q-Q plot of the estimates of $\hat{\rho}_g$ and $\hat{Z}$ for sample size 20 based on simulation of 1000 runs from the multivariate normal distribution with three repeated measure, high within-unit correlation ($\rho = 0.8$), and moderate between-unit correlation (Case 2).
Fig. 4.6: Q-Q plot of the estimates of $\hat{\rho}_g$ and $\hat{Z}$ for sample size 40 based on simulation of 1000 runs from the multivariate normal distribution with three repeated measure, high within-unit correlation ($\rho = 0.8$), and moderate between-unit correlation (Case 2).
Fig. 4.7: Q-Q plot of the estimates of $\hat{\rho}_g$ and $\hat{Z}$ for sample size 80 based on simulation of 1000 runs from the multivariate normal distribution with three repeated measure, high within-unit correlation ($\rho = 0.8$), and moderate between-unit correlation (Case 2).
Fig. 4.8: Q-Q plot of the estimates of $\hat{\rho}_g$ and $\hat{Z}$ for sample size 160 based on simulation of 1000 runs from the multivariate normal distribution with three repeated measure, high within-unit correlation ($\rho = 0.8$), and moderate between-unit correlation (Case 2).
Fig. 4.9: Q-Q plot of the estimates of $\hat{\rho}_g$ and $\hat{Z}$ for sample size 20 based on simulation of 1000 runs from the multivariate normal distribution with three repeated measure, high within-unit correlation ($\rho = 0.8$), and weak between-unit correlation (Case 3).
Fig. 4.10: Q-Q plot of the estimates of $\hat{\rho}_g$ and $\hat{Z}$ for sample size 40 based on simulation of 1000 runs from the multivariate normal distribution with three repeated measure, high within-unit correlation ($\rho = 0.8$), and weak between-unit correlation (Case 3).
Fig. 4.11: Q-Q plot of the estimates of $\hat{\rho}_g$ and $\hat{Z}$ for sample size 80 based on simulation of 1000 runs from the multivariate normal distribution with three repeated measure, high within-unit correlation ($\rho = 0.8$), and weak between-unit correlation (Case 3).
Fig. 4.12: Q-Q plot of the estimates of $\hat{\rho}_g$ and $\hat{Z}$ for sample size 160 based on simulation of 1000 runs from the multivariate normal distribution with three repeated measure, high within-unit correlation ($\rho = 0.8$), and weak between-unit correlation (Case 3).
Table 4.1: Mean of estimates of $\hat{\rho}_g$, relative bias, the estimated asymptotic variances $\hat{\sigma}_Z^2$, and the empirical variance of $\hat{Z}$ based on simulation of 1000 runs from the multivariate normal distribution: three repeated measurements.

<table>
<thead>
<tr>
<th>Case</th>
<th>Measure</th>
<th>$\rho_{g1}$</th>
<th>$\rho_{g2}$</th>
<th>$\rho_{g3}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>Rel. bias</td>
<td>0.020</td>
<td>0.009</td>
<td>0.005</td>
</tr>
<tr>
<td></td>
<td>$\hat{\sigma}_Z^2$</td>
<td>0.015</td>
<td>0.007</td>
<td>0.004</td>
</tr>
<tr>
<td></td>
<td>$\hat{\sigma}_Z^2$</td>
<td>0.017</td>
<td>0.007</td>
<td>0.004</td>
</tr>
<tr>
<td></td>
<td>$\hat{\sigma}_Z^2$</td>
<td>0.011</td>
<td>0.005</td>
<td>0.001</td>
</tr>
<tr>
<td>2</td>
<td>Rel. bias</td>
<td>0.019</td>
<td>0.010</td>
<td>0.005</td>
</tr>
<tr>
<td></td>
<td>$\hat{\sigma}_Z^2$</td>
<td>0.012</td>
<td>0.006</td>
<td>0.001</td>
</tr>
<tr>
<td></td>
<td>$\hat{\sigma}_Z^2$</td>
<td>0.014</td>
<td>0.007</td>
<td>0.003</td>
</tr>
<tr>
<td>3</td>
<td>Rel. bias</td>
<td>0.020</td>
<td>0.011</td>
<td>0.004</td>
</tr>
<tr>
<td></td>
<td>$\hat{\sigma}_Z^2$</td>
<td>0.010</td>
<td>0.005</td>
<td>0.001</td>
</tr>
<tr>
<td></td>
<td>$\hat{\sigma}_Z^2$</td>
<td>0.013</td>
<td>0.006</td>
<td>0.001</td>
</tr>
<tr>
<td>0.4</td>
<td>Rel. bias</td>
<td>0.018</td>
<td>0.008</td>
<td>0.004</td>
</tr>
<tr>
<td></td>
<td>$\hat{\sigma}_Z^2$</td>
<td>0.017</td>
<td>0.008</td>
<td>0.004</td>
</tr>
<tr>
<td></td>
<td>$\hat{\sigma}_Z^2$</td>
<td>0.019</td>
<td>0.008</td>
<td>0.004</td>
</tr>
</tbody>
</table>

Continued on Next Page.
<table>
<thead>
<tr>
<th>Within-unit correlation</th>
<th>Case</th>
<th>Measure</th>
<th>20</th>
<th>40</th>
<th>80</th>
<th>160</th>
<th>20</th>
<th>40</th>
<th>80</th>
<th>160</th>
<th>20</th>
<th>40</th>
<th>80</th>
<th>160</th>
<th>20</th>
<th>40</th>
<th>80</th>
<th>160</th>
<th>20</th>
<th>40</th>
<th>80</th>
<th>160</th>
<th>20</th>
<th>40</th>
<th>80</th>
<th>160</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.4</td>
<td>2</td>
<td>$\hat{\rho}_g$</td>
<td>0.687</td>
<td>0.704</td>
<td>0.711</td>
<td>0.715</td>
<td>0.719</td>
<td>0.484</td>
<td>0.538</td>
<td>0.563</td>
<td>0.576</td>
<td>0.588</td>
<td>0.731</td>
<td>0.732</td>
<td>0.732</td>
<td>0.732</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Rel. bias</td>
<td>0.044</td>
<td>0.020</td>
<td>0.011</td>
<td>0.005</td>
<td>0.177</td>
<td>0.085</td>
<td>0.042</td>
<td>0.020</td>
<td>0.001</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>$\hat{\sigma}_Z^2$</td>
<td>0.012</td>
<td>0.006</td>
<td>0.003</td>
<td>0.002</td>
<td>0.033</td>
<td>0.015</td>
<td>0.007</td>
<td>0.004</td>
<td>0.012</td>
<td>0.007</td>
<td>0.003</td>
<td>0.002</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Var($\hat{Z}$)</td>
<td>0.013</td>
<td>0.006</td>
<td>0.003</td>
<td>0.001</td>
<td>0.024</td>
<td>0.012</td>
<td>0.006</td>
<td>0.003</td>
<td>0.015</td>
<td>0.007</td>
<td>0.003</td>
<td>0.002</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3</td>
<td></td>
<td>$\hat{\rho}_g$</td>
<td>0.390</td>
<td>0.404</td>
<td>0.408</td>
<td>0.412</td>
<td>0.414</td>
<td>0.130</td>
<td>0.208</td>
<td>0.247</td>
<td>0.269</td>
<td>0.288</td>
<td>0.433</td>
<td>0.428</td>
<td>0.424</td>
<td>0.422</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Rel. bias</td>
<td>0.058</td>
<td>0.023</td>
<td>0.013</td>
<td>0.005</td>
<td>0.549</td>
<td>0.279</td>
<td>0.142</td>
<td>0.068</td>
<td>-0.032</td>
<td>-0.020</td>
<td>-0.009</td>
<td>-0.006</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>$\hat{\sigma}_Z^2$</td>
<td>0.011</td>
<td>0.006</td>
<td>0.003</td>
<td>0.001</td>
<td>0.033</td>
<td>0.013</td>
<td>0.006</td>
<td>0.003</td>
<td>0.012</td>
<td>0.007</td>
<td>0.003</td>
<td>0.002</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Var($\hat{Z}$)</td>
<td>0.012</td>
<td>0.005</td>
<td>0.003</td>
<td>0.001</td>
<td>0.018</td>
<td>0.008</td>
<td>0.004</td>
<td>0.002</td>
<td>0.014</td>
<td>0.006</td>
<td>0.003</td>
<td>0.002</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0.8</td>
<td>1</td>
<td>$\hat{\rho}_g$</td>
<td>0.909</td>
<td>0.918</td>
<td>0.922</td>
<td>0.924</td>
<td>0.925</td>
<td>0.837</td>
<td>0.865</td>
<td>0.877</td>
<td>0.882</td>
<td>0.887</td>
<td>0.927</td>
<td>0.928</td>
<td>0.929</td>
<td>0.929</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Rel. bias</td>
<td>0.017</td>
<td>0.007</td>
<td>0.004</td>
<td>0.002</td>
<td>0.056</td>
<td>0.025</td>
<td>0.012</td>
<td>0.006</td>
<td>0.003</td>
<td>0.001</td>
<td>0.001</td>
<td>0.000</td>
<td>0.000</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>$\hat{\sigma}_Z^2$</td>
<td>0.022</td>
<td>0.009</td>
<td>0.004</td>
<td>0.002</td>
<td>0.051</td>
<td>0.023</td>
<td>0.011</td>
<td>0.005</td>
<td>0.021</td>
<td>0.008</td>
<td>0.004</td>
<td>0.002</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Var($\hat{Z}$)</td>
<td>0.019</td>
<td>0.008</td>
<td>0.004</td>
<td>0.002</td>
<td>0.037</td>
<td>0.017</td>
<td>0.009</td>
<td>0.004</td>
<td>0.016</td>
<td>0.007</td>
<td>0.004</td>
<td>0.002</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td></td>
<td>$\hat{\rho}_g$</td>
<td>0.702</td>
<td>0.720</td>
<td>0.727</td>
<td>0.731</td>
<td>0.735</td>
<td>0.516</td>
<td>0.579</td>
<td>0.608</td>
<td>0.623</td>
<td>0.637</td>
<td>0.742</td>
<td>0.743</td>
<td>0.743</td>
<td>0.743</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Rel. bias</td>
<td>0.044</td>
<td>0.020</td>
<td>0.010</td>
<td>0.005</td>
<td>0.190</td>
<td>0.091</td>
<td>0.045</td>
<td>0.022</td>
<td>0.002</td>
<td>0.000</td>
<td>0.001</td>
<td>0.000</td>
<td>0.000</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>$\hat{\sigma}_Z^2$</td>
<td>0.015</td>
<td>0.007</td>
<td>0.003</td>
<td>0.002</td>
<td>0.041</td>
<td>0.019</td>
<td>0.009</td>
<td>0.004</td>
<td>0.016</td>
<td>0.007</td>
<td>0.004</td>
<td>0.002</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Var($\hat{Z}$)</td>
<td>0.014</td>
<td>0.007</td>
<td>0.003</td>
<td>0.002</td>
<td>0.027</td>
<td>0.013</td>
<td>0.007</td>
<td>0.004</td>
<td>0.016</td>
<td>0.007</td>
<td>0.003</td>
<td>0.002</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Continued on Next Page...
<table>
<thead>
<tr>
<th>Within-unit correlation</th>
<th>Case</th>
<th>Measure</th>
<th>$\rho_{g1}^a$</th>
<th>$\rho_{g2}^b$</th>
<th>$\rho_{g3}^c$</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>0.8</td>
<td>3</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>$\hat{\rho}_g$</td>
<td>0.401</td>
<td>0.416</td>
<td>0.421</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Rel. bias</td>
<td>0.062</td>
<td>0.025</td>
<td>0.014</td>
</tr>
<tr>
<td></td>
<td></td>
<td>$\hat{\sigma}_Z^2$</td>
<td>0.012</td>
<td>0.006</td>
<td>0.003</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Var($\hat{Z}$)</td>
<td>0.012</td>
<td>0.006</td>
<td>0.003</td>
</tr>
</tbody>
</table>

$a$ The multivariate CCC based on the trace function

$b$ The multivariate CCC based on the largest eigenvalue function

$c$ The multivariate CCC based on the determinant function

$d$ Relative bias $= \frac{\rho_g - \hat{\rho}_g}{\rho_g}$
Table 4.2: Mean of estimates of $\hat{\rho}_g$, relative bias, the estimated asymptotic variances $\hat{\sigma}_Z^2$, and the empirical variance of $\hat{Z}$ based on simulation of 1000 runs from the multivariate normal distribution: five repeated measurements.

<table>
<thead>
<tr>
<th>Within-unit correlation</th>
<th>Case</th>
<th>Measure</th>
<th>$\rho_g$</th>
<th>$\hat{\rho}_g$</th>
<th>Rel. bias</th>
<th>$\hat{\sigma}_Z^2$</th>
<th>Var($\hat{Z}$)</th>
<th>$\rho_g$</th>
<th>$\hat{\rho}_g$</th>
<th>Rel. bias</th>
<th>$\hat{\sigma}_Z^2$</th>
<th>Var($\hat{Z}$)</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>1</td>
<td></td>
<td>$\rho_g$</td>
<td>0.877</td>
<td>0.891</td>
<td>0.898</td>
<td>0.901</td>
<td>0.903</td>
<td>0.651</td>
<td>0.700</td>
<td>0.724</td>
<td>0.731</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Rel. bias</td>
<td></td>
<td>0.029</td>
<td>0.013</td>
<td>0.005</td>
<td>0.003</td>
<td>0.119</td>
<td>0.052</td>
<td>0.021</td>
<td>0.010</td>
<td>0.001</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>$\hat{\sigma}_Z^2$</td>
<td></td>
<td>0.009</td>
<td>0.005</td>
<td>0.002</td>
<td>0.001</td>
<td>0.029</td>
<td>0.014</td>
<td>0.007</td>
<td>0.003</td>
<td>0.007</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Var($\hat{Z}$)</td>
<td></td>
<td>0.010</td>
<td>0.005</td>
<td>0.002</td>
<td>0.001</td>
<td>0.026</td>
<td>0.013</td>
<td>0.006</td>
<td>0.003</td>
<td>0.009</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td></td>
<td>$\rho_g$</td>
<td>0.661</td>
<td>0.685</td>
<td>0.696</td>
<td>0.701</td>
<td>0.705</td>
<td>0.270</td>
<td>0.341</td>
<td>0.371</td>
<td>0.381</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Rel. bias</td>
<td></td>
<td>0.063</td>
<td>0.030</td>
<td>0.013</td>
<td>0.006</td>
<td>0.311</td>
<td>0.131</td>
<td>0.054</td>
<td>0.027</td>
<td>-0.007</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>$\hat{\sigma}_Z^2$</td>
<td></td>
<td>0.006</td>
<td>0.003</td>
<td>0.001</td>
<td>0.001</td>
<td>0.018</td>
<td>0.007</td>
<td>0.003</td>
<td>0.002</td>
<td>0.007</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Var($\hat{Z}$)</td>
<td></td>
<td>0.007</td>
<td>0.003</td>
<td>0.001</td>
<td>0.001</td>
<td>0.012</td>
<td>0.006</td>
<td>0.003</td>
<td>0.001</td>
<td>0.009</td>
</tr>
<tr>
<td></td>
<td>3</td>
<td></td>
<td>$\rho_g$</td>
<td>0.374</td>
<td>0.394</td>
<td>0.405</td>
<td>0.409</td>
<td>0.413</td>
<td>-0.041</td>
<td>0.080</td>
<td>0.132</td>
<td>0.148</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Rel. bias</td>
<td></td>
<td>0.094</td>
<td>0.047</td>
<td>0.021</td>
<td>0.010</td>
<td>1.257</td>
<td>0.506</td>
<td>0.182</td>
<td>0.082</td>
<td>-0.060</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>$\hat{\sigma}_Z^2$</td>
<td></td>
<td>0.006</td>
<td>0.003</td>
<td>0.002</td>
<td>0.001</td>
<td>0.036</td>
<td>0.009</td>
<td>0.002</td>
<td>0.001</td>
<td>0.007</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Var($\hat{Z}$)</td>
<td></td>
<td>0.007</td>
<td>0.003</td>
<td>0.002</td>
<td>0.001</td>
<td>0.018</td>
<td>0.005</td>
<td>0.002</td>
<td>0.001</td>
<td>0.009</td>
</tr>
<tr>
<td></td>
<td>0.4</td>
<td>1</td>
<td>$\rho_g$</td>
<td>0.902</td>
<td>0.915</td>
<td>0.922</td>
<td>0.924</td>
<td>0.926</td>
<td>0.761</td>
<td>0.815</td>
<td>0.838</td>
<td>0.846</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Rel. bias</td>
<td></td>
<td>0.026</td>
<td>0.011</td>
<td>0.005</td>
<td>0.002</td>
<td>0.107</td>
<td>0.045</td>
<td>0.017</td>
<td>0.008</td>
<td>0.002</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>$\hat{\sigma}_Z^2$</td>
<td></td>
<td>0.012</td>
<td>0.005</td>
<td>0.003</td>
<td>0.001</td>
<td>0.046</td>
<td>0.020</td>
<td>0.009</td>
<td>0.004</td>
<td>0.009</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Var($\hat{Z}$)</td>
<td></td>
<td>0.012</td>
<td>0.005</td>
<td>0.003</td>
<td>0.001</td>
<td>0.032</td>
<td>0.016</td>
<td>0.009</td>
<td>0.004</td>
<td>0.010</td>
</tr>
</tbody>
</table>

Continued on Next Page...
Table 4.2 – Continued

<table>
<thead>
<tr>
<th>Within-unit correlation</th>
<th>Case</th>
<th>Measure</th>
<th>$\hat{\rho}_g$</th>
<th>$\rho_{g1}^a$</th>
<th>$\rho_{g2}^b$</th>
<th>$\rho_{g3}^c$</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.4</td>
<td>2</td>
<td></td>
<td></td>
<td>0.691</td>
<td>0.717</td>
<td>0.730</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Rel. bias</td>
<td>0.066</td>
<td>0.031</td>
<td>0.014</td>
<td>0.007</td>
</tr>
<tr>
<td></td>
<td></td>
<td>$\hat{\sigma}_Z^2$</td>
<td>0.008</td>
<td>0.004</td>
<td>0.002</td>
<td>0.001</td>
</tr>
<tr>
<td></td>
<td></td>
<td>$\text{Var}(\hat{Z})$</td>
<td>0.009</td>
<td>0.004</td>
<td>0.002</td>
<td>0.001</td>
</tr>
<tr>
<td>0.8</td>
<td>1</td>
<td>$\hat{\rho}_g$</td>
<td>0.910</td>
<td>0.923</td>
<td>0.928</td>
<td>0.930</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Rel. bias</td>
<td>0.024</td>
<td>0.011</td>
<td>0.004</td>
<td>0.002</td>
</tr>
<tr>
<td></td>
<td></td>
<td>$\hat{\sigma}_Z^2$</td>
<td>0.020</td>
<td>0.007</td>
<td>0.003</td>
<td>0.001</td>
</tr>
<tr>
<td></td>
<td></td>
<td>$\text{Var}(\hat{Z})$</td>
<td>0.013</td>
<td>0.005</td>
<td>0.003</td>
<td>0.001</td>
</tr>
<tr>
<td>2</td>
<td></td>
<td>$\hat{\rho}_g$</td>
<td>0.703</td>
<td>0.730</td>
<td>0.743</td>
<td>0.754</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Rel. bias</td>
<td>0.067</td>
<td>0.032</td>
<td>0.014</td>
<td>0.007</td>
</tr>
<tr>
<td></td>
<td></td>
<td>$\hat{\sigma}_Z^2$</td>
<td>0.012</td>
<td>0.005</td>
<td>0.002</td>
<td>0.001</td>
</tr>
<tr>
<td></td>
<td></td>
<td>$\text{Var}(\hat{Z})$</td>
<td>0.010</td>
<td>0.004</td>
<td>0.002</td>
<td>0.001</td>
</tr>
</tbody>
</table>

Continued on Next Page...
<table>
<thead>
<tr>
<th>Within-unit correlation</th>
<th>Case</th>
<th>Measure</th>
<th>( \rho_g^a )</th>
<th>( \rho_g^b )</th>
<th>( \rho_g^c )</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>20</td>
<td>40</td>
<td>80</td>
<td>160</td>
<td>20</td>
</tr>
<tr>
<td>0.8</td>
<td>3</td>
<td>( \hat{\rho}_g )</td>
<td>0.398</td>
<td>0.422</td>
<td>0.435</td>
</tr>
<tr>
<td>Rel. bias</td>
<td>0.107</td>
<td>0.053</td>
<td>0.024</td>
<td>0.012</td>
<td>1.114</td>
</tr>
<tr>
<td>( \hat{\sigma}_Z^2 )</td>
<td>0.009</td>
<td>0.004</td>
<td>0.002</td>
<td>0.001</td>
<td>0.072</td>
</tr>
<tr>
<td>( \text{Var}(\hat{Z}) )</td>
<td>0.008</td>
<td>0.004</td>
<td>0.002</td>
<td>0.001</td>
<td>0.025</td>
</tr>
</tbody>
</table>

*a* The multivariate CCC based on the trace function  
*b* The multivariate CCC based on the largest eigenvalue function  
*c* The multivariate CCC based on the determinant function  
*d* Relative bias = \( \frac{\rho_g - \hat{\rho}_g}{\rho_g} \)
Table 4.3: Empirical coverage probabilities for the 95% confidence intervals of \( \rho_g \) based on simulation of 1000 runs generated from the multivariate normal distribution.

<table>
<thead>
<tr>
<th>Within-unit correlation</th>
<th>Case</th>
<th>n</th>
<th>Three repeated measure</th>
<th>Five repeated measure</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td>( \rho_{g1}^a )</td>
<td>( \rho_{g2}^b )</td>
</tr>
<tr>
<td>0</td>
<td>1</td>
<td>20</td>
<td>0.855</td>
<td>0.887</td>
</tr>
<tr>
<td></td>
<td></td>
<td>40</td>
<td>0.918</td>
<td>0.922</td>
</tr>
<tr>
<td></td>
<td></td>
<td>80</td>
<td>0.940</td>
<td>0.929</td>
</tr>
<tr>
<td></td>
<td></td>
<td>160</td>
<td>0.950</td>
<td>0.940</td>
</tr>
<tr>
<td>2</td>
<td>20</td>
<td></td>
<td>0.876</td>
<td>0.858</td>
</tr>
<tr>
<td></td>
<td></td>
<td>40</td>
<td>0.924</td>
<td>0.902</td>
</tr>
<tr>
<td></td>
<td></td>
<td>80</td>
<td>0.932</td>
<td>0.919</td>
</tr>
<tr>
<td></td>
<td></td>
<td>160</td>
<td>0.943</td>
<td>0.942</td>
</tr>
<tr>
<td>3</td>
<td>20</td>
<td></td>
<td>0.912</td>
<td>0.878</td>
</tr>
<tr>
<td></td>
<td></td>
<td>40</td>
<td>0.941</td>
<td>0.895</td>
</tr>
<tr>
<td></td>
<td></td>
<td>80</td>
<td>0.946</td>
<td>0.903</td>
</tr>
<tr>
<td></td>
<td></td>
<td>160</td>
<td>0.951</td>
<td>0.931</td>
</tr>
<tr>
<td>0.4</td>
<td>1</td>
<td>20</td>
<td>0.870</td>
<td>0.871</td>
</tr>
<tr>
<td></td>
<td></td>
<td>40</td>
<td>0.926</td>
<td>0.921</td>
</tr>
<tr>
<td></td>
<td></td>
<td>80</td>
<td>0.937</td>
<td>0.940</td>
</tr>
<tr>
<td></td>
<td></td>
<td>160</td>
<td>0.949</td>
<td>0.945</td>
</tr>
<tr>
<td>2</td>
<td>20</td>
<td></td>
<td>0.874</td>
<td>0.854</td>
</tr>
<tr>
<td></td>
<td></td>
<td>40</td>
<td>0.924</td>
<td>0.902</td>
</tr>
<tr>
<td></td>
<td></td>
<td>80</td>
<td>0.935</td>
<td>0.932</td>
</tr>
<tr>
<td></td>
<td></td>
<td>160</td>
<td>0.949</td>
<td>0.933</td>
</tr>
<tr>
<td>3</td>
<td>20</td>
<td></td>
<td>0.911</td>
<td>0.853</td>
</tr>
<tr>
<td></td>
<td></td>
<td>40</td>
<td>0.945</td>
<td>0.883</td>
</tr>
<tr>
<td></td>
<td></td>
<td>80</td>
<td>0.944</td>
<td>0.907</td>
</tr>
<tr>
<td></td>
<td></td>
<td>160</td>
<td>0.950</td>
<td>0.926</td>
</tr>
<tr>
<td>0.8</td>
<td>1</td>
<td>20</td>
<td>0.903</td>
<td>0.872</td>
</tr>
<tr>
<td></td>
<td></td>
<td>40</td>
<td>0.939</td>
<td>0.930</td>
</tr>
<tr>
<td></td>
<td></td>
<td>80</td>
<td>0.944</td>
<td>0.939</td>
</tr>
<tr>
<td></td>
<td></td>
<td>160</td>
<td>0.960</td>
<td>0.942</td>
</tr>
</tbody>
</table>

Continued on Next Page...
<table>
<thead>
<tr>
<th>Within-unit correlation</th>
<th>Case</th>
<th>n</th>
<th>Three repeated measure</th>
<th>Five repeated measure</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td>$\rho_{g1}$</td>
<td>$\rho_{g2}$</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>20</td>
<td>0.893</td>
<td>0.861</td>
</tr>
<tr>
<td></td>
<td></td>
<td>40</td>
<td>0.926</td>
<td>0.911</td>
</tr>
<tr>
<td></td>
<td></td>
<td>80</td>
<td>0.940</td>
<td>0.928</td>
</tr>
<tr>
<td></td>
<td></td>
<td>160</td>
<td>0.949</td>
<td>0.950</td>
</tr>
<tr>
<td></td>
<td>3</td>
<td>20</td>
<td>0.920</td>
<td>0.853</td>
</tr>
<tr>
<td></td>
<td></td>
<td>40</td>
<td>0.951</td>
<td>0.868</td>
</tr>
<tr>
<td></td>
<td></td>
<td>80</td>
<td>0.948</td>
<td>0.905</td>
</tr>
<tr>
<td></td>
<td></td>
<td>160</td>
<td>0.950</td>
<td>0.929</td>
</tr>
</tbody>
</table>

$^a$The multivariate CCC based on the trace function
$^b$The multivariate CCC based on the largest eigenvalue function
$^c$The multivariate CCC based on the determinant function
Fig. 4.13: Q-Q plot of the estimates of $\hat{\rho}_g$ and $\hat{Z}$ for sample size 20 based on simulation of 1000 runs from multivariate Student’s t-distribution with three repeated measure, high within-unit correlation ($\rho = 0.8$), and strong between-unit correlation (Case 1).
Fig. 4.14: Q-Q plot of the estimates of $\hat{\rho}_g$ and $\hat{Z}$ for sample size 40 based on simulation of 1000 runs from multivariate Student’s t-distribution with three repeated measure, high within-unit correlation($\rho = 0.8$), and strong between-unit correlation(Case 1).
Fig. 4.15: Q-Q plot of the estimates of $\hat{\rho}_g$ and $\hat{Z}$ for sample size 80 based on simulation of 1000 runs from multivariate Student’s t-distribution with three repeated measure, high within-unit correlation ($\rho = 0.8$), and strong between-unit correlation (Case 1).
Fig. 4.16: Q-Q plot of the estimates of $\hat{\rho}_g$ and $\hat{Z}$ for sample size 160 based on simulation of 1000 runs from multivariate Student’s t-distribution with three repeated measure, high within-unit correlation($\rho = 0.8$), and strong between-unit correlation(Case 1).
Fig. 4.17: Q-Q plot of the estimates of $\hat{\rho}_g$ and $\hat{Z}$ for sample size 20 based on simulation of 1000 runs from multivariate Student’s t-distribution with three repeated measure, high within-unit correlation ($\rho = 0.8$), and moderate between-unit correlation (Case 2).
Fig. 4.18: Q-Q plot of the estimates of $\hat{\rho}_g$ and $\hat{Z}$ for sample size 40 based on simulation of 1000 runs from multivariate Student’s t-distribution with three repeated measure, high within-unit correlation ($\rho = 0.8$), and moderate between-unit correlation (Case 2).
Fig. 4.19: Q-Q plot of the estimates of $\hat{\rho}_g$ and $\hat{Z}$ for sample size 80 based on simulation of 1000 runs from multivariate Student’s t-distribution with three repeated measure, high within-unit correlation ($\rho = 0.8$), and moderate between-unit correlation (Case 2).
Fig. 4.20: Q-Q plot of the estimates of $\hat{\rho}_g$ and $\hat{Z}$ for sample size 160 based on simulation of 1000 runs from multivariate Student’s t-distribution with three repeated measure, high within-unit correlation ($\rho = 0.8$), and moderate between-unit correlation (Case 2).
Fig. 4.21: Q-Q plot of the estimates of $\hat{\rho}_g$ and $\hat{Z}$ for sample size 20 based on simulation of 1000 runs from multivariate Student’s t-distribution with three repeated measure, high within-unit correlation ($\rho = 0.8$), and weaker between-unit correlation (Case 3).
Fig. 4.22: Q-Q plot of the estimates of $\hat{\rho}_g$ and $\hat{Z}$ for sample size 40 based on simulation of 1000 runs from multivariate Student’s t-distribution with three repeated measure, high within-unit correlation($\rho = 0.8$), and weaker between-unit correlation (Case 3).
Fig. 4.23: Q-Q plot of the estimates of $\hat{\rho}_g$ and $\hat{Z}$ for sample size 80 based on simulation of 1000 runs from multivariate Student’s t-distribution with three repeated measure, high within-unit correlation ($\rho = 0.8$), and weaker between-unit correlation (Case 3).
Fig. 4.24: Q-Q plot of the estimates of $\hat{\rho}_g$ and $\hat{Z}$ for sample size 160 based on simulation of 1000 runs from multivariate Student’s t-distribution with three repeated measure, high within-unit correlation ($\rho = 0.8$), and weaker between-unit correlation (Case 3).
Table 4.4: Mean of estimates of $\hat{\rho}_g$, relative bias, the estimated asymptotic variances $\hat{\sigma}_Z^2$, and the empirical variance of $\hat{Z}$ based on simulation of 1000 runs from multivariate Student’s t-distribution: three repeated measurements.

<table>
<thead>
<tr>
<th>Within-unit correlation</th>
<th>Case</th>
<th>Measure</th>
<th>$\rho_{g1}^a$</th>
<th>$\rho_{g2}^b$</th>
<th>$\rho_{g3}^c$</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td>20  40  80  160</td>
<td>20  40  80  160</td>
<td>20  40  80  160</td>
</tr>
<tr>
<td>0</td>
<td>1</td>
<td>$\hat{\rho}_g$</td>
<td>0.887 0.898 0.903 0.905</td>
<td>0.907 0.777 0.807 0.821</td>
<td>0.827 0.833 0.916 0.918</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Rel. bias</td>
<td>0.023 0.011 0.005 0.002</td>
<td>0.067 0.031 0.015 0.007</td>
<td>0.003 0.002 0.001 0.000</td>
</tr>
<tr>
<td></td>
<td></td>
<td>$\hat{\sigma}_Z^2$</td>
<td>0.017 0.008 0.004 0.002</td>
<td>0.038 0.019 0.010 0.005</td>
<td>0.013 0.007 0.004 0.002</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Var($\hat{Z}$)</td>
<td>0.020 0.009 0.004 0.002</td>
<td>0.037 0.018 0.009 0.005</td>
<td>0.018 0.009 0.004 0.002</td>
</tr>
<tr>
<td>2</td>
<td>1</td>
<td>$\hat{\rho}_g$</td>
<td>0.663 0.681 0.691 0.695</td>
<td>0.699 0.425 0.480 0.506</td>
<td>0.518 0.530 0.719 0.720</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Rel. bias</td>
<td>0.052 0.025 0.012 0.006</td>
<td>0.197 0.093 0.045 0.021</td>
<td>0.001 0.001 -0.001 0.000</td>
</tr>
<tr>
<td></td>
<td></td>
<td>$\hat{\sigma}_Z^2$</td>
<td>0.011 0.006 0.003 0.002</td>
<td>0.029 0.014 0.007 0.003</td>
<td>0.013 0.007 0.004 0.002</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Var($\hat{Z}$)</td>
<td>0.014 0.007 0.003 0.002</td>
<td>0.023 0.012 0.006 0.003</td>
<td>0.018 0.009 0.004 0.002</td>
</tr>
<tr>
<td>3</td>
<td>1</td>
<td>$\hat{\rho}_g$</td>
<td>0.371 0.385 0.394 0.396</td>
<td>0.399 0.090 0.167 0.209</td>
<td>0.229 0.246 0.423 0.415</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Rel. bias</td>
<td>0.071 0.036 0.014 0.007</td>
<td>0.632 0.321 0.148 0.068</td>
<td>-0.034 -0.015 -0.012 -0.005</td>
</tr>
<tr>
<td></td>
<td></td>
<td>$\hat{\sigma}_Z^2$</td>
<td>0.011 0.006 0.003 0.002</td>
<td>0.034 0.013 0.005 0.002</td>
<td>0.013 0.007 0.004 0.002</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Var($\hat{Z}$)</td>
<td>0.014 0.007 0.003 0.002</td>
<td>0.021 0.008 0.004 0.002</td>
<td>0.018 0.008 0.004 0.002</td>
</tr>
<tr>
<td>0.4</td>
<td>1</td>
<td>$\hat{\rho}_g$</td>
<td>0.904 0.914 0.918 0.920</td>
<td>0.922 0.824 0.854 0.867</td>
<td>0.873 0.879 0.925 0.926</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Rel. bias</td>
<td>0.020 0.009 0.004 0.002</td>
<td>0.062 0.028 0.014 0.007</td>
<td>0.003 0.001 0.001 0.000</td>
</tr>
<tr>
<td></td>
<td></td>
<td>$\hat{\sigma}_Z^2$</td>
<td>0.019 0.009 0.005 0.002</td>
<td>0.047 0.023 0.011 0.006</td>
<td>0.015 0.008 0.004 0.002</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Var($\hat{Z}$)</td>
<td>0.021 0.010 0.005 0.002</td>
<td>0.038 0.019 0.010 0.005</td>
<td>0.018 0.009 0.004 0.002</td>
</tr>
</tbody>
</table>

Continued on Next Page...
### Table 4.4 – Continued

<table>
<thead>
<tr>
<th>Within-unit correlation</th>
<th>Case</th>
<th>Measure</th>
<th>$\rho_g^a$</th>
<th>$\rho_g^b$</th>
<th>$\rho_g^c$</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>20 40 80 160</td>
<td>20 40 80 160</td>
<td>20 40 80 160</td>
<td>20 40 80 160</td>
</tr>
<tr>
<td>0.4</td>
<td>2</td>
<td>$\hat{\rho}_g$</td>
<td>0.691 0.711 0.720 0.724</td>
<td>0.488 0.555 0.588 0.604</td>
<td>0.619</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Rel. bias</td>
<td>0.052 0.025 0.012 0.006</td>
<td>0.212 0.103 0.050 0.024</td>
<td>0.002</td>
</tr>
<tr>
<td></td>
<td></td>
<td>$\hat{\sigma}_Z^2$</td>
<td>0.014 0.007 0.004 0.002</td>
<td>0.040 0.019 0.009 0.005</td>
<td>0.013</td>
</tr>
<tr>
<td></td>
<td></td>
<td>$Var(\hat{Z})$</td>
<td>0.016 0.008 0.004 0.002</td>
<td>0.029 0.014 0.008 0.004</td>
<td>0.018</td>
</tr>
<tr>
<td>3</td>
<td>3</td>
<td>$\hat{\rho}_g$</td>
<td>0.390 0.405 0.415 0.418</td>
<td>0.109 0.200 0.256 0.284</td>
<td>0.313</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Rel. bias</td>
<td>0.076 0.039 0.016 0.008</td>
<td>0.653 0.361 0.183 0.093</td>
<td>-0.031 -0.014 -0.011 -0.006</td>
</tr>
<tr>
<td></td>
<td></td>
<td>$\hat{\sigma}_Z^2$</td>
<td>0.012 0.006 0.003 0.002</td>
<td>0.044 0.019 0.009 0.004</td>
<td>0.013</td>
</tr>
<tr>
<td></td>
<td></td>
<td>$Var(\hat{Z})$</td>
<td>0.015 0.007 0.004 0.002</td>
<td>0.027 0.012 0.006 0.003</td>
<td>0.018</td>
</tr>
<tr>
<td>0.8</td>
<td>1</td>
<td>$\hat{\rho}_g$</td>
<td>0.912 0.921 0.925 0.927</td>
<td>0.929 0.842 0.872 0.885</td>
<td>0.892 0.898 0.929 0.931 0.931 0.931 0.932</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Rel. bias</td>
<td>0.018 0.008 0.004 0.002</td>
<td>0.063 0.029 0.014 0.007</td>
<td>0.003</td>
</tr>
<tr>
<td></td>
<td></td>
<td>$\hat{\sigma}_Z^2$</td>
<td>0.025 0.011 0.005 0.002</td>
<td>0.059 0.027 0.013 0.006</td>
<td>0.023</td>
</tr>
<tr>
<td></td>
<td></td>
<td>$Var(\hat{Z})$</td>
<td>0.022 0.010 0.005 0.002</td>
<td>0.039 0.019 0.010 0.006</td>
<td>0.019</td>
</tr>
<tr>
<td>2</td>
<td>2</td>
<td>$\hat{\rho}_g$</td>
<td>0.705 0.725 0.734 0.739</td>
<td>0.743 0.513 0.587 0.625</td>
<td>0.643 0.662 0.747 0.748 0.749 0.749 0.749</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Rel. bias</td>
<td>0.051 0.025 0.012 0.006</td>
<td>0.224 0.112 0.055 0.028</td>
<td>0.003</td>
</tr>
<tr>
<td></td>
<td></td>
<td>$\hat{\sigma}_Z^2$</td>
<td>0.018 0.008 0.004 0.002</td>
<td>0.051 0.024 0.011 0.005</td>
<td>0.018</td>
</tr>
<tr>
<td></td>
<td></td>
<td>$Var(\hat{Z})$</td>
<td>0.018 0.009 0.004 0.002</td>
<td>0.033 0.016 0.008 0.005</td>
<td>0.018</td>
</tr>
</tbody>
</table>

Continued on Next Page...
<table>
<thead>
<tr>
<th>Within-unit correlation</th>
<th>Case</th>
<th>Measure</th>
<th>( \rho_{g1}^a )</th>
<th>( \rho_{g1}^b )</th>
<th>( \rho_{g1}^c )</th>
<th>( \rho_{g2}^b )</th>
<th>( \rho_{g3}^a )</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>0.8</td>
<td>3</td>
<td>0.400 0.417 0.427 0.430 0.434 0.117 0.216 0.277 0.310 0.350 0.450 0.443 0.442 0.440 0.437</td>
<td>0.079 0.040 0.017 0.009 0.665 0.384 0.208 0.113 -0.030 -0.014 -0.010 -0.006</td>
<td>0.014 0.007 0.004 0.002 0.054 0.023 0.011 0.005 0.015 0.007 0.004 0.002</td>
<td>0.016 0.008 0.004 0.002 0.031 0.014 0.007 0.003 0.018 0.009 0.004 0.002</td>
<td></td>
</tr>
</tbody>
</table>

\( \hat{\rho}_g \) indicates the multivariate CCC based on the trace function.

\( \hat{\sigma}_Z^2 \) indicates the multivariate CCC based on the largest eigenvalue function.

\( \hat{\text{Var}}(\hat{Z}) \) indicates the multivariate CCC based on the determinant function.

Relative bias = \( \frac{\hat{\rho}_g - \hat{\rho}_g}{\rho_g} \).
Table 4.5: Mean of estimates of $\hat{\rho}_g$, relative bias, the estimated asymptotic variances $\hat{\sigma}_Z^2$, and the empirical variance of $\hat{Z}$ based on simulation of 1000 runs from multivariate Student’s t-distribution: five repeated measurements.

<table>
<thead>
<tr>
<th>Within-unit correlation</th>
<th>Case</th>
<th>Measure</th>
<th>$\rho_{g1}$</th>
<th>$\rho_{g2}$</th>
<th>$\rho_{g3}$</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td>20</td>
<td>40</td>
<td>80</td>
</tr>
<tr>
<td>0</td>
<td>1</td>
<td>$\hat{\rho}_g$</td>
<td>0.878</td>
<td>0.896</td>
<td>0.903</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Rel. bias</td>
<td>0.034</td>
<td>0.015</td>
<td>0.007</td>
</tr>
<tr>
<td></td>
<td></td>
<td>$\hat{\sigma}_Z^2$</td>
<td>0.010</td>
<td>0.005</td>
<td>0.003</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Var($\hat{Z}$)</td>
<td>0.013</td>
<td>0.006</td>
<td>0.003</td>
</tr>
<tr>
<td>2</td>
<td>1</td>
<td>$\hat{\rho}_g$</td>
<td>0.661</td>
<td>0.689</td>
<td>0.702</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Rel. bias</td>
<td>0.075</td>
<td>0.036</td>
<td>0.017</td>
</tr>
<tr>
<td></td>
<td></td>
<td>$\hat{\sigma}_Z^2$</td>
<td>0.007</td>
<td>0.003</td>
<td>0.002</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Var($\hat{Z}$)</td>
<td>0.009</td>
<td>0.004</td>
<td>0.002</td>
</tr>
<tr>
<td>3</td>
<td>1</td>
<td>$\hat{\rho}_g$</td>
<td>0.370</td>
<td>0.394</td>
<td>0.406</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Rel. bias</td>
<td>0.114</td>
<td>0.057</td>
<td>0.028</td>
</tr>
<tr>
<td></td>
<td></td>
<td>$\hat{\sigma}_Z^2$</td>
<td>0.006</td>
<td>0.003</td>
<td>0.002</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Var($\hat{Z}$)</td>
<td>0.008</td>
<td>0.004</td>
<td>0.002</td>
</tr>
<tr>
<td>0.4</td>
<td>1</td>
<td>$\hat{\rho}_g$</td>
<td>0.901</td>
<td>0.917</td>
<td>0.924</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Rel. bias</td>
<td>0.030</td>
<td>0.013</td>
<td>0.006</td>
</tr>
<tr>
<td></td>
<td></td>
<td>$\hat{\sigma}_Z^2$</td>
<td>0.014</td>
<td>0.006</td>
<td>0.003</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Var($\hat{Z}$)</td>
<td>0.016</td>
<td>0.006</td>
<td>0.003</td>
</tr>
</tbody>
</table>

Continued on Next Page...
Table 4.5 – Continued

<table>
<thead>
<tr>
<th>Within-unit correlation</th>
<th>Case</th>
<th>Measure</th>
<th>$\rho_g$</th>
<th>$\rho_{g1}^a$</th>
<th>$\rho_{g2}^b$</th>
<th>$\rho_{g3}^c$</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>20 40 80 160</td>
<td></td>
<td>20 40 80 160</td>
<td>20 40 80 160</td>
<td></td>
</tr>
<tr>
<td>0.4</td>
<td></td>
<td>ρg</td>
<td>0.688</td>
<td>0.719 0.734 0.740 0.747</td>
<td>0.343 0.479 0.545 0.572</td>
<td>0.599 0.761 0.758 0.757</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Rel. bias</td>
<td>0.078</td>
<td>0.037 0.018 0.010</td>
<td>0.429 0.200 0.092 0.045</td>
<td>-0.008 -0.004 -0.002 0.000</td>
</tr>
<tr>
<td></td>
<td></td>
<td>$\hat{\sigma}_Z^2$</td>
<td>0.009</td>
<td>0.004 0.002 0.001</td>
<td>0.046 0.020 0.009 0.004</td>
<td>0.008 0.004 0.002 0.001</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Var(\hat{Z})</td>
<td>0.011</td>
<td>0.005 0.003 0.001</td>
<td>0.029 0.012 0.007 0.004</td>
<td>0.011 0.005 0.003 0.001</td>
</tr>
<tr>
<td>3</td>
<td>3</td>
<td>ρg</td>
<td>0.386</td>
<td>0.413 0.427 0.433 0.440</td>
<td>-0.082 0.088 0.190 0.245</td>
<td>0.297 0.475 0.459 0.451 0.447 0.445</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Rel. bias</td>
<td>0.124</td>
<td>0.061 0.031 0.017</td>
<td>1.276 0.705 0.361 0.177</td>
<td>-0.068 -0.032 -0.015 -0.006</td>
</tr>
<tr>
<td></td>
<td></td>
<td>$\hat{\sigma}_Z^2$</td>
<td>0.007</td>
<td>0.004 0.002 0.001</td>
<td>0.063 0.023 0.010 0.004</td>
<td>0.007 0.004 0.002 0.001</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Var(\hat{Z})</td>
<td>0.009</td>
<td>0.004 0.002 0.001</td>
<td>0.033 0.010 0.005 0.003</td>
<td>0.010 0.005 0.003 0.001</td>
</tr>
<tr>
<td>0.8</td>
<td>1</td>
<td>ρg</td>
<td>0.908</td>
<td>0.923 0.929 0.932 0.935</td>
<td>0.778 0.844 0.872 0.885</td>
<td>0.897 0.936 0.936 0.937 0.937 0.937</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Rel. bias</td>
<td>0.028</td>
<td>0.012 0.006 0.003</td>
<td>0.132 0.058 0.027 0.013</td>
<td>0.001 0.000 0.000 0.000 0.000</td>
</tr>
<tr>
<td></td>
<td></td>
<td>$\hat{\sigma}_Z^2$</td>
<td>0.023</td>
<td>0.008 0.003 0.002</td>
<td>0.079 0.032 0.014 0.006</td>
<td>0.030 0.007 0.003 0.001</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Var(\hat{Z})</td>
<td>0.016</td>
<td>0.007 0.003 0.001</td>
<td>0.039 0.017 0.009 0.005</td>
<td>0.012 0.006 0.003 0.001</td>
</tr>
<tr>
<td>2</td>
<td>2</td>
<td>ρg</td>
<td>0.699</td>
<td>0.731 0.745 0.751 0.759</td>
<td>0.365 0.513 0.587 0.623</td>
<td>0.658 0.768 0.766 0.764 0.763 0.763</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Rel. bias</td>
<td>0.079</td>
<td>0.037 0.018 0.010</td>
<td>0.445 0.221 0.109 0.054</td>
<td>-0.007 -0.003 -0.002 0.000</td>
</tr>
<tr>
<td></td>
<td></td>
<td>$\hat{\sigma}_Z^2$</td>
<td>0.014</td>
<td>0.006 0.003 0.001</td>
<td>0.070 0.028 0.013 0.006</td>
<td>0.018 0.006 0.003 0.001</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Var(\hat{Z})</td>
<td>0.012</td>
<td>0.005 0.003 0.001</td>
<td>0.033 0.013 0.007 0.004</td>
<td>0.011 0.005 0.003 0.001</td>
</tr>
</tbody>
</table>

Continued on Next Page...
<table>
<thead>
<tr>
<th>Within-unit correlation</th>
<th>Case</th>
<th>Measure</th>
<th>$\rho_g^a$</th>
<th>$\rho_g^b$</th>
<th>$\rho_g^c$</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>0.8</td>
<td>3</td>
<td>$\hat{\rho}_g$</td>
<td>$\hat{\rho}_g$</td>
<td>$\hat{\rho}_g$</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>0.393</td>
<td>0.422</td>
<td>0.426</td>
</tr>
<tr>
<td>Rel. bias</td>
<td></td>
<td></td>
<td>0.128</td>
<td>0.063</td>
<td>0.032</td>
</tr>
<tr>
<td>$\hat{\sigma}_Z^2$</td>
<td></td>
<td></td>
<td>0.010</td>
<td>0.004</td>
<td>0.002</td>
</tr>
<tr>
<td>$Var(\hat{Z})$</td>
<td></td>
<td></td>
<td>0.009</td>
<td>0.004</td>
<td>0.002</td>
</tr>
</tbody>
</table>

*a* The multivariate CCC based on the trace function  
*b* The multivariate CCC based on the largest eigenvalue function  
*c* The multivariate CCC based on the determinant function  
*d* Relative bias = $\frac{\rho_g - \hat{\rho}_g}{\rho_g}$
Table 4.6: Empirical coverage probabilities for the 95% confidence intervals of $\rho_g$ based on simulation of 1000 runs generated from multivariate Student’s t distribution.

<table>
<thead>
<tr>
<th>Within-unit correlation</th>
<th>Case</th>
<th>n</th>
<th>Three repeated measure</th>
<th>Five repeated measure</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td>$\rho_{g1}$</td>
<td>$\rho_{g2}$</td>
</tr>
<tr>
<td>0</td>
<td>1</td>
<td>20</td>
<td>0.846</td>
<td>0.864</td>
</tr>
<tr>
<td></td>
<td></td>
<td>40</td>
<td>0.902</td>
<td>0.902</td>
</tr>
<tr>
<td></td>
<td></td>
<td>80</td>
<td>0.932</td>
<td>0.931</td>
</tr>
<tr>
<td></td>
<td></td>
<td>160</td>
<td>0.942</td>
<td>0.943</td>
</tr>
<tr>
<td>2</td>
<td>1</td>
<td>20</td>
<td>0.861</td>
<td>0.835</td>
</tr>
<tr>
<td></td>
<td></td>
<td>40</td>
<td>0.909</td>
<td>0.890</td>
</tr>
<tr>
<td></td>
<td></td>
<td>80</td>
<td>0.928</td>
<td>0.920</td>
</tr>
<tr>
<td></td>
<td></td>
<td>160</td>
<td>0.944</td>
<td>0.943</td>
</tr>
<tr>
<td>3</td>
<td>1</td>
<td>20</td>
<td>0.886</td>
<td>0.851</td>
</tr>
<tr>
<td></td>
<td></td>
<td>40</td>
<td>0.923</td>
<td>0.888</td>
</tr>
<tr>
<td></td>
<td></td>
<td>80</td>
<td>0.937</td>
<td>0.917</td>
</tr>
<tr>
<td></td>
<td></td>
<td>160</td>
<td>0.940</td>
<td>0.949</td>
</tr>
<tr>
<td>0.4</td>
<td>1</td>
<td>20</td>
<td>0.861</td>
<td>0.868</td>
</tr>
<tr>
<td></td>
<td></td>
<td>40</td>
<td>0.906</td>
<td>0.912</td>
</tr>
<tr>
<td></td>
<td></td>
<td>80</td>
<td>0.928</td>
<td>0.925</td>
</tr>
<tr>
<td></td>
<td></td>
<td>160</td>
<td>0.941</td>
<td>0.933</td>
</tr>
<tr>
<td>2</td>
<td>1</td>
<td>20</td>
<td>0.855</td>
<td>0.822</td>
</tr>
<tr>
<td></td>
<td></td>
<td>40</td>
<td>0.904</td>
<td>0.888</td>
</tr>
<tr>
<td></td>
<td></td>
<td>80</td>
<td>0.925</td>
<td>0.918</td>
</tr>
<tr>
<td></td>
<td></td>
<td>160</td>
<td>0.945</td>
<td>0.926</td>
</tr>
<tr>
<td>3</td>
<td>1</td>
<td>20</td>
<td>0.889</td>
<td>0.816</td>
</tr>
<tr>
<td></td>
<td></td>
<td>40</td>
<td>0.918</td>
<td>0.872</td>
</tr>
<tr>
<td></td>
<td></td>
<td>80</td>
<td>0.935</td>
<td>0.912</td>
</tr>
<tr>
<td></td>
<td></td>
<td>160</td>
<td>0.946</td>
<td>0.936</td>
</tr>
<tr>
<td>0.8</td>
<td>1</td>
<td>20</td>
<td>0.905</td>
<td>0.873</td>
</tr>
<tr>
<td></td>
<td></td>
<td>40</td>
<td>0.928</td>
<td>0.918</td>
</tr>
<tr>
<td></td>
<td></td>
<td>80</td>
<td>0.934</td>
<td>0.918</td>
</tr>
<tr>
<td></td>
<td></td>
<td>160</td>
<td>0.943</td>
<td>0.934</td>
</tr>
</tbody>
</table>

Continued on Next Page...
<table>
<thead>
<tr>
<th>Within-unit correlation</th>
<th>Case</th>
<th>n</th>
<th>Three repeated measure</th>
<th>Five repeated measure</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td>$\rho_g^1$</td>
<td>$\rho_g^2$</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>20</td>
<td>0.880</td>
<td>0.832</td>
</tr>
<tr>
<td></td>
<td></td>
<td>40</td>
<td>0.925</td>
<td>0.898</td>
</tr>
<tr>
<td></td>
<td></td>
<td>80</td>
<td>0.929</td>
<td>0.914</td>
</tr>
<tr>
<td></td>
<td></td>
<td>160</td>
<td>0.945</td>
<td>0.924</td>
</tr>
<tr>
<td></td>
<td>3</td>
<td>20</td>
<td>0.902</td>
<td>0.828</td>
</tr>
<tr>
<td></td>
<td></td>
<td>40</td>
<td>0.916</td>
<td>0.864</td>
</tr>
<tr>
<td></td>
<td></td>
<td>80</td>
<td>0.936</td>
<td>0.905</td>
</tr>
<tr>
<td></td>
<td></td>
<td>160</td>
<td>0.941</td>
<td>0.922</td>
</tr>
</tbody>
</table>

$^a$ The scalar version of CCC using the trace function

$^b$ The scalar version of CCC using the largest eigenvalue function

$^c$ The scalar version of CCC using the determinant function
Fig. 4.25: Q-Q plot of the estimates of $\hat{\rho}$ and $\hat{Z}$ for sample size 20 based on simulation of 1000 runs from the multivariate lognormal distribution with three repeated measures, high within-unit correlation ($\rho = 0.8$), and strong between-unit correlation (Case 1).
Fig. 4.26: Q-Q plot of the estimates of $\hat{\rho}_g$ and $\hat{Z}$ for sample size 40 based on simulation of 1000 runs from the multivariate lognormal distribution with three repeated measure, high within-unit correlation ($\rho = 0.8$), and strong between-unit correlation (Case 1).
Fig. 4.27: Q-Q plot of the estimates of $\hat{\rho}_g$ and $\hat{Z}$ for sample size 80 based on simulation of 1000 runs from the multivariate lognormal distribution with three repeated measure, high within-unit correlation ($\rho = 0.8$), and strong between-unit correlation (Case 1).
Fig. 4.28: Q-Q plot of the estimates of $\hat{\rho}_g$ and $\hat{Z}$ for sample size 160 based on simulation of 1000 runs from the multivariate lognormal distribution with three repeated measure, high within-unit correlation ($\rho = 0.8$), and strong between-unit correlation (Case 1).
Fig. 4.29: Q-Q plot of the estimates of $\hat{\rho}_g$ and $\hat{Z}$ for sample size 20 based on simulation of 1000 runs from the multivariate lognormal distribution with three repeated measure, high within-unit correlation ($\rho = 0.8$), and moderate between-unit correlation (Case 2).
Fig. 4.30: Q-Q plot of the estimates of $\hat{\rho}$ and $\hat{Z}$ for sample size 40 based on simulation of 1000 runs from the multivariate lognormal distribution with three repeated measure, high within-unit correlation ($\rho = 0.8$), and moderate between-unit correlation (Case 2).
Fig. 4.31: Q-Q plot of the estimates of $\hat{\rho}_g$ and $\hat{Z}$ for sample size 80 based on simulation of 1000 runs from the multivariate lognormal distribution with three repeated measure, high within-unit correlation ($\rho = 0.8$), and moderate between-unit correlation (Case 2).
Fig. 4.32: Q-Q plot of the estimates of $\hat{\rho}_g$ and $\hat{Z}$ for sample size 160 based on simulation of 1000 runs from the multivariate lognormal distribution with three repeated measure, high within-unit correlation ($\rho = 0.8$), and moderate between-unit correlation (Case 2).
Fig. 4.33: Q-Q plot of the estimates of \( \hat{\rho}_g \) and \( \hat{Z} \) for sample size 20 based on simulation of 1000 runs from the multivariate lognormal distribution with three repeated measure, high within-unit correlation(\( \rho = 0.8 \)), and weaker between-unit correlation(Case 3).
Fig. 4.34: Q-Q plot of the estimates of $\hat{\rho}_g$ and $\hat{Z}$ for sample size 40 based on simulation of 1000 runs from the multivariate lognormal distribution with three repeated measure, high within-unit correlation($\rho = 0.8$), and weaker between-unit correlation(Case 3).
Fig. 4.35: Q-Q plot of the estimates of $\hat{\rho}_g$ and $\hat{Z}$ for sample size 80 based on simulation of 1000 runs from the multivariate lognormal distribution with three repeated measure, high within-unit correlation ($\rho = 0.8$), and weaker between-unit correlation (Case 3).
Fig. 4.36: Q-Q plot of the estimates of $\hat{\rho}_g$ and $\hat{Z}$ for sample size 160 based on simulation of 1000 runs from the multivariate lognormal distribution with three repeated measures, high within-unit correlation ($\rho = 0.8$), and weaker between-unit correlation (Case 3).
Table 4.7: Mean of estimates of $\hat{\rho}_g$, relative bias, the estimated asymptotic variances $\hat{\sigma}_Z^2$, and the empirical variance of $\hat{Z}$ based on simulation of 1000 runs from the multivariate lognormal distribution: three repeated measurements.

<table>
<thead>
<tr>
<th>Within-unit correlation</th>
<th>Measure</th>
<th>$\rho_g^a$</th>
<th>Rel. bias</th>
<th>$\hat{\sigma}_Z^2$</th>
<th>Var($\hat{Z}$)</th>
<th>$\rho_g^b$</th>
<th>Rel. bias</th>
<th>$\hat{\sigma}_Z^2$</th>
<th>Var($\hat{Z}$)</th>
<th>$\rho_g^c$</th>
<th>Rel. bias</th>
<th>$\hat{\sigma}_Z^2$</th>
<th>Var($\hat{Z}$)</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>$\rho_g$</td>
<td>0.827</td>
<td>0.023</td>
<td>0.012</td>
<td>0.014</td>
<td>0.836</td>
<td>0.012</td>
<td>0.006</td>
<td>0.007</td>
<td>0.840</td>
<td>0.003</td>
<td>0.002</td>
<td>0.002</td>
</tr>
<tr>
<td></td>
<td></td>
<td>0.843</td>
<td>0.081</td>
<td>0.026</td>
<td>0.023</td>
<td>0.846</td>
<td>0.026</td>
<td>0.006</td>
<td>0.011</td>
<td>0.679</td>
<td>0.003</td>
<td>0.002</td>
<td>0.003</td>
</tr>
<tr>
<td></td>
<td></td>
<td>0.722</td>
<td>0.044</td>
<td>0.011</td>
<td>0.011</td>
<td>0.731</td>
<td>0.012</td>
<td>0.003</td>
<td>0.016</td>
<td>0.739</td>
<td>0.003</td>
<td>0.002</td>
<td>0.004</td>
</tr>
<tr>
<td></td>
<td></td>
<td>0.861</td>
<td>0.011</td>
<td>0.004</td>
<td>0.008</td>
<td>0.862</td>
<td>0.006</td>
<td>0.004</td>
<td>0.004</td>
<td>0.861</td>
<td>0.004</td>
<td>0.002</td>
<td>0.002</td>
</tr>
<tr>
<td></td>
<td></td>
<td>0.862</td>
<td>0.004</td>
<td>0.002</td>
<td>0.002</td>
<td>0.862</td>
<td>0.004</td>
<td>0.004</td>
<td>0.002</td>
<td>0.862</td>
<td>0.002</td>
<td>0.002</td>
<td>0.002</td>
</tr>
<tr>
<td></td>
<td></td>
<td>0.862</td>
<td>0.004</td>
<td>0.002</td>
<td>0.002</td>
<td>0.862</td>
<td>0.004</td>
<td>0.004</td>
<td>0.002</td>
<td>0.862</td>
<td>0.002</td>
<td>0.002</td>
<td>0.002</td>
</tr>
<tr>
<td>2</td>
<td>$\rho_g$</td>
<td>0.535</td>
<td>0.018</td>
<td>0.006</td>
<td>0.010</td>
<td>0.541</td>
<td>0.004</td>
<td>0.004</td>
<td>0.005</td>
<td>0.543</td>
<td>0.004</td>
<td>0.002</td>
<td>0.002</td>
</tr>
<tr>
<td></td>
<td></td>
<td>0.544</td>
<td>0.004</td>
<td>0.003</td>
<td>0.002</td>
<td>0.544</td>
<td>0.012</td>
<td>0.003</td>
<td>0.003</td>
<td>0.540</td>
<td>0.003</td>
<td>0.002</td>
<td>0.002</td>
</tr>
<tr>
<td></td>
<td></td>
<td>0.309</td>
<td>0.026</td>
<td>0.006</td>
<td>0.023</td>
<td>0.348</td>
<td>0.007</td>
<td>0.003</td>
<td>0.002</td>
<td>0.369</td>
<td>0.004</td>
<td>0.002</td>
<td>0.002</td>
</tr>
<tr>
<td></td>
<td></td>
<td>0.381</td>
<td>0.016</td>
<td>0.005</td>
<td>0.011</td>
<td>0.393</td>
<td>0.007</td>
<td>0.005</td>
<td>0.003</td>
<td>0.381</td>
<td>0.004</td>
<td>0.002</td>
<td>0.002</td>
</tr>
<tr>
<td></td>
<td></td>
<td>0.574</td>
<td>0.014</td>
<td>0.005</td>
<td>0.005</td>
<td>0.567</td>
<td>0.007</td>
<td>0.003</td>
<td>0.003</td>
<td>0.562</td>
<td>0.003</td>
<td>0.002</td>
<td>0.002</td>
</tr>
<tr>
<td></td>
<td></td>
<td>0.559</td>
<td>0.016</td>
<td>0.005</td>
<td>0.003</td>
<td>0.559</td>
<td>0.007</td>
<td>0.003</td>
<td>0.003</td>
<td>0.556</td>
<td>0.002</td>
<td>0.002</td>
<td>0.002</td>
</tr>
<tr>
<td>3</td>
<td>$\rho_g$</td>
<td>0.237</td>
<td>-0.063</td>
<td>0.005</td>
<td>0.007</td>
<td>0.234</td>
<td>-0.052</td>
<td>0.003</td>
<td>0.003</td>
<td>0.230</td>
<td>0.003</td>
<td>0.002</td>
<td>0.002</td>
</tr>
<tr>
<td></td>
<td></td>
<td>0.227</td>
<td>-0.052</td>
<td>0.003</td>
<td>0.002</td>
<td>0.223</td>
<td>-0.052</td>
<td>0.004</td>
<td>0.003</td>
<td>0.223</td>
<td>0.004</td>
<td>0.002</td>
<td>0.002</td>
</tr>
<tr>
<td></td>
<td></td>
<td>0.042</td>
<td>-0.052</td>
<td>0.004</td>
<td>0.002</td>
<td>0.089</td>
<td>-0.052</td>
<td>0.005</td>
<td>0.004</td>
<td>0.114</td>
<td>0.005</td>
<td>0.002</td>
<td>0.002</td>
</tr>
<tr>
<td></td>
<td></td>
<td>0.127</td>
<td>-0.052</td>
<td>0.006</td>
<td>0.004</td>
<td>0.127</td>
<td>-0.052</td>
<td>0.005</td>
<td>0.004</td>
<td>0.141</td>
<td>0.005</td>
<td>0.002</td>
<td>0.002</td>
</tr>
<tr>
<td></td>
<td></td>
<td>0.260</td>
<td>-0.052</td>
<td>0.006</td>
<td>0.004</td>
<td>0.260</td>
<td>-0.052</td>
<td>0.006</td>
<td>0.004</td>
<td>0.246</td>
<td>0.005</td>
<td>0.002</td>
<td>0.002</td>
</tr>
<tr>
<td></td>
<td></td>
<td>0.237</td>
<td>-0.052</td>
<td>0.006</td>
<td>0.004</td>
<td>0.237</td>
<td>-0.052</td>
<td>0.006</td>
<td>0.004</td>
<td>0.232</td>
<td>0.005</td>
<td>0.002</td>
<td>0.002</td>
</tr>
<tr>
<td></td>
<td></td>
<td>0.225</td>
<td>-0.052</td>
<td>0.006</td>
<td>0.004</td>
<td>0.225</td>
<td>-0.052</td>
<td>0.006</td>
<td>0.004</td>
<td>0.225</td>
<td>0.005</td>
<td>0.002</td>
<td>0.002</td>
</tr>
<tr>
<td>0.4</td>
<td>$\rho_g$</td>
<td>0.846</td>
<td>0.021</td>
<td>0.014</td>
<td>0.016</td>
<td>0.855</td>
<td>0.010</td>
<td>0.007</td>
<td>0.002</td>
<td>0.859</td>
<td>0.005</td>
<td>0.004</td>
<td>0.002</td>
</tr>
<tr>
<td></td>
<td></td>
<td>0.864</td>
<td>0.005</td>
<td>0.005</td>
<td>0.003</td>
<td>0.864</td>
<td>0.006</td>
<td>0.004</td>
<td>0.002</td>
<td>0.862</td>
<td>0.005</td>
<td>0.004</td>
<td>0.002</td>
</tr>
<tr>
<td></td>
<td></td>
<td>0.731</td>
<td>0.003</td>
<td>0.005</td>
<td>0.002</td>
<td>0.764</td>
<td>0.005</td>
<td>0.002</td>
<td>0.003</td>
<td>0.780</td>
<td>0.005</td>
<td>0.002</td>
<td>0.002</td>
</tr>
<tr>
<td></td>
<td></td>
<td>0.789</td>
<td>0.007</td>
<td>0.004</td>
<td>0.002</td>
<td>0.789</td>
<td>0.006</td>
<td>0.003</td>
<td>0.004</td>
<td>0.797</td>
<td>0.005</td>
<td>0.002</td>
<td>0.002</td>
</tr>
<tr>
<td></td>
<td></td>
<td>0.872</td>
<td>0.007</td>
<td>0.005</td>
<td>0.004</td>
<td>0.872</td>
<td>0.007</td>
<td>0.003</td>
<td>0.004</td>
<td>0.871</td>
<td>0.005</td>
<td>0.002</td>
<td>0.002</td>
</tr>
<tr>
<td></td>
<td></td>
<td>0.871</td>
<td>0.007</td>
<td>0.005</td>
<td>0.004</td>
<td>0.871</td>
<td>0.007</td>
<td>0.003</td>
<td>0.004</td>
<td>0.871</td>
<td>0.005</td>
<td>0.002</td>
<td>0.002</td>
</tr>
<tr>
<td></td>
<td></td>
<td>0.871</td>
<td>0.007</td>
<td>0.005</td>
<td>0.004</td>
<td>0.871</td>
<td>0.007</td>
<td>0.003</td>
<td>0.004</td>
<td>0.871</td>
<td>0.005</td>
<td>0.002</td>
<td>0.002</td>
</tr>
</tbody>
</table>

Continued on Next Page...
<table>
<thead>
<tr>
<th>Case</th>
<th>Measure</th>
<th>$\rho_a$</th>
<th>$\rho_b$</th>
<th>$\rho_c$</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0.234</td>
<td>0.012</td>
<td>0.011</td>
<td>0.010</td>
</tr>
<tr>
<td></td>
<td>0.235</td>
<td>0.010</td>
<td>0.009</td>
<td>0.008</td>
</tr>
<tr>
<td></td>
<td>0.236</td>
<td>0.008</td>
<td>0.007</td>
<td>0.006</td>
</tr>
<tr>
<td></td>
<td>0.237</td>
<td>0.006</td>
<td>0.005</td>
<td>0.004</td>
</tr>
</tbody>
</table>

Continued on Next Page...
Table 4.7 – Continued

<table>
<thead>
<tr>
<th>Within-unit correlation</th>
<th>Case</th>
<th>Measure</th>
<th>$\rho_g$</th>
<th>$\rho_g$</th>
<th>$\rho_g$</th>
<th>$\rho_g$</th>
<th>$\rho_g$</th>
<th>$\rho_g$</th>
<th>$\rho_g$</th>
<th>$\rho_g$</th>
<th>$\rho_g$</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>3</td>
<td></td>
<td>$\rho_g$</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>20</td>
<td>40</td>
<td>80</td>
<td>160</td>
<td>20</td>
<td>40</td>
<td>80</td>
<td>160</td>
<td>20</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>$\hat{\rho}_g$</td>
<td>0.244</td>
<td>0.240</td>
<td>0.234</td>
<td>0.230</td>
<td>0.222</td>
<td>0.026</td>
<td>0.092</td>
<td>0.129</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Rel. bias</td>
<td>-0.096</td>
<td>-0.078</td>
<td>-0.053</td>
<td>-0.033</td>
<td>0.862</td>
<td>0.512</td>
<td>0.316</td>
<td>0.196</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>$\hat{\sigma}^2_{\hat{Z}}$</td>
<td>0.008</td>
<td>0.004</td>
<td>0.002</td>
<td>0.001</td>
<td>0.027</td>
<td>0.009</td>
<td>0.004</td>
<td>0.002</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>$\text{Var}(\hat{Z})$</td>
<td>0.009</td>
<td>0.004</td>
<td>0.002</td>
<td>0.001</td>
<td>0.013</td>
<td>0.006</td>
<td>0.003</td>
<td>0.002</td>
</tr>
</tbody>
</table>

*The multivariate CCC based on the trace function

bThe multivariate CCC based on the largest eigenvalue function

cThe multivariate CCC based on the determinant function

dRelative bias = $\frac{\hat{\rho}_g - \rho_g}{\rho_g}$
Table 4.8: Mean of estimates of $\hat{\rho}_g$, relative bias, the estimated asymptotic variances $\hat{\sigma}_Z^2$, and the empirical variance of $\hat{Z}$ based on simulation of 1000 runs from the multivariate lognormal distribution: five repeated measurements.

<table>
<thead>
<tr>
<th>Within-unit correlation</th>
<th>Case</th>
<th>Measure</th>
<th>$\rho_{g1}$</th>
<th>$\rho_{g2}$</th>
<th>$\rho_{g3}$</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td>20 40 80 160</td>
<td>20 40 80 160</td>
<td>20 40 80 160</td>
</tr>
<tr>
<td>0</td>
<td>1</td>
<td>$\rho_g$</td>
<td>0.837 0.851 0.858 0.861 0.863 0.595</td>
<td>0.650 0.678 0.688 0.698 0.883 0.882</td>
<td>0.881 0.881 0.880</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Rel. bias</td>
<td>0.030 0.014 0.006 0.003</td>
<td>0.147 0.068 0.029 0.015</td>
<td>-0.004 -0.002 -0.002 -0.001</td>
</tr>
<tr>
<td></td>
<td></td>
<td>$\hat{\sigma}_Z^2$</td>
<td>0.008 0.004 0.002 0.001</td>
<td>0.027 0.012 0.006 0.003</td>
<td>0.007 0.004 0.002 0.001</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Var($\hat{Z}$)</td>
<td>0.009 0.005 0.003 0.001</td>
<td>0.021 0.011 0.006 0.003</td>
<td>0.010 0.005 0.003 0.002</td>
</tr>
<tr>
<td>2</td>
<td></td>
<td>$\rho_g$</td>
<td>0.550 0.563 0.568 0.569 0.202</td>
<td>0.276 0.311 0.323 0.336 0.610</td>
<td>0.598 0.592 0.586 0.581</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Rel. bias</td>
<td>0.033 0.011 0.001 0.001</td>
<td>0.398 0.177 0.074 0.039</td>
<td>-0.049 -0.028 -0.017 -0.008</td>
</tr>
<tr>
<td></td>
<td></td>
<td>$\hat{\sigma}_Z^2$</td>
<td>0.005 0.003 0.001 0.001</td>
<td>0.017 0.006 0.003 0.001</td>
<td>0.005 0.003 0.002 0.001</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Var($\hat{Z}$)</td>
<td>0.006 0.003 0.002 0.001</td>
<td>0.009 0.005 0.003 0.001</td>
<td>0.009 0.004 0.002 0.001</td>
</tr>
<tr>
<td>3</td>
<td></td>
<td>$\rho_g$</td>
<td>0.250 0.251 0.250 0.245 0.241</td>
<td>-0.071 0.030 0.082 0.102 0.118</td>
<td>0.293 0.271 0.260 0.252 0.244</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Rel. bias</td>
<td>-0.035 -0.040 -0.034 -0.017</td>
<td>1.601 0.749 0.308 0.138</td>
<td>-0.201 -0.112 -0.067 -0.033</td>
</tr>
<tr>
<td></td>
<td></td>
<td>$\hat{\sigma}_Z^2$</td>
<td>0.004 0.002 0.001 0.001</td>
<td>0.027 0.007 0.002 0.001</td>
<td>0.004 0.002 0.001 0.001</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Var($\hat{Z}$)</td>
<td>0.005 0.003 0.001 0.001</td>
<td>0.013 0.004 0.001 0.001</td>
<td>0.007 0.003 0.001 0.001</td>
</tr>
<tr>
<td>0.4</td>
<td>1</td>
<td>$\rho_g$</td>
<td>0.855 0.870 0.877 0.879 0.881</td>
<td>0.673 0.741 0.772 0.784 0.796</td>
<td>0.892 0.890 0.889 0.888 0.887</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Rel. bias</td>
<td>0.029 0.013 0.005 0.002</td>
<td>0.154 0.070 0.031 0.014</td>
<td>-0.006 -0.004 -0.003 -0.001</td>
</tr>
<tr>
<td></td>
<td></td>
<td>$\hat{\sigma}_Z^2$</td>
<td>0.011 0.005 0.003 0.001</td>
<td>0.043 0.018 0.008 0.004</td>
<td>0.010 0.004 0.003 0.001</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Var($\hat{Z}$)</td>
<td>0.011 0.006 0.003 0.002</td>
<td>0.026 0.013 0.007 0.004</td>
<td>0.012 0.006 0.003 0.002</td>
</tr>
</tbody>
</table>
Table 4.8 – Continued

<table>
<thead>
<tr>
<th>Within-unit correlation</th>
<th>Case</th>
<th>Measure</th>
<th>$\rho_g^a$</th>
<th>$\rho_g^b$</th>
<th>$\rho_g^c$</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td>20  40  80 160</td>
<td>20  40  80 160</td>
<td>20  40  80 160</td>
</tr>
<tr>
<td>0.4</td>
<td>2</td>
<td>$\hat{\rho}_g$</td>
<td>0.566</td>
<td>0.580</td>
<td>0.587</td>
</tr>
<tr>
<td>Rel. bias</td>
<td></td>
<td></td>
<td>0.033</td>
<td>0.008</td>
<td>-0.002</td>
</tr>
<tr>
<td>$\hat{\sigma}_Z^2$</td>
<td></td>
<td></td>
<td>0.006</td>
<td>0.003</td>
<td>0.002</td>
</tr>
<tr>
<td>Var($\hat{Z}$)</td>
<td></td>
<td></td>
<td>0.008</td>
<td>0.004</td>
<td>0.003</td>
</tr>
<tr>
<td>3</td>
<td></td>
<td>$\hat{\rho}_g$</td>
<td>0.257</td>
<td>0.257</td>
<td>0.254</td>
</tr>
<tr>
<td>Rel. bias</td>
<td></td>
<td></td>
<td>-0.058</td>
<td>-0.057</td>
<td>-0.047</td>
</tr>
<tr>
<td>$\hat{\sigma}_Z^2$</td>
<td></td>
<td></td>
<td>0.004</td>
<td>0.002</td>
<td>0.001</td>
</tr>
<tr>
<td>Var($\hat{Z}$)</td>
<td></td>
<td></td>
<td>0.006</td>
<td>0.003</td>
<td>0.002</td>
</tr>
<tr>
<td>0.8</td>
<td>1</td>
<td>$\hat{\rho}_g$</td>
<td>0.850</td>
<td>0.867</td>
<td>0.874</td>
</tr>
<tr>
<td>Rel. bias</td>
<td></td>
<td></td>
<td>0.033</td>
<td>0.014</td>
<td>0.006</td>
</tr>
<tr>
<td>$\hat{\sigma}_Z^2$</td>
<td></td>
<td></td>
<td>0.025</td>
<td>0.010</td>
<td>0.005</td>
</tr>
<tr>
<td>Var($\hat{Z}$)</td>
<td></td>
<td></td>
<td>0.012</td>
<td>0.006</td>
<td>0.003</td>
</tr>
<tr>
<td>2</td>
<td></td>
<td>$\hat{\rho}_g$</td>
<td>0.565</td>
<td>0.580</td>
<td>0.585</td>
</tr>
<tr>
<td>Rel. bias</td>
<td></td>
<td></td>
<td>0.030</td>
<td>0.004</td>
<td>-0.005</td>
</tr>
<tr>
<td>$\hat{\sigma}_Z^2$</td>
<td></td>
<td></td>
<td>0.014</td>
<td>0.006</td>
<td>0.003</td>
</tr>
<tr>
<td>Var($\hat{Z}$)</td>
<td></td>
<td></td>
<td>0.009</td>
<td>0.005</td>
<td>0.003</td>
</tr>
</tbody>
</table>

Continued on Next Page...
Table 4.8 – Continued

<table>
<thead>
<tr>
<th>Within-unit correlation</th>
<th>Case</th>
<th>Measure</th>
<th>$\rho g_1^a$</th>
<th>$\rho g_2^b$</th>
<th>$\rho g_3^c$</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>0.8</td>
<td>3</td>
<td>$\hat{\rho} g$</td>
<td>$\hat{\rho} g$</td>
<td>$\hat{\rho} g$</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>20</td>
<td>40</td>
<td>80</td>
</tr>
<tr>
<td>Rel. bias</td>
<td></td>
<td></td>
<td>-0.093</td>
<td>-0.083</td>
<td>-0.068</td>
</tr>
<tr>
<td>$\hat{\sigma}^2_{\hat{Z}}$</td>
<td>0.009</td>
<td>0.003</td>
<td>0.001</td>
<td>0.001</td>
<td>0.077</td>
</tr>
<tr>
<td>$\text{Var}(\hat{Z})$</td>
<td>0.006</td>
<td>0.003</td>
<td>0.002</td>
<td>0.001</td>
<td>0.022</td>
</tr>
</tbody>
</table>

$^a$The multivariate CCC based on the trace function
$^b$The multivariate CCC based on the largest eigenvalue function
$^c$The multivariate CCC based on the determinant function
$^d$Relative bias = \( \frac{\hat{\rho}_g - \rho_g}{\rho_g} \)
Table 4.9: Empirical coverage probabilities for the 95% confidence intervals of $\rho_g$ based on simulation of 1000 runs generated from the multivariate lognormal distribution.

<table>
<thead>
<tr>
<th>Within-unit correlation</th>
<th>Case</th>
<th>n</th>
<th>Three repeated measure</th>
<th>Five repeated measure</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td>$\rho_g^a$</td>
<td>$\rho_g^b$</td>
</tr>
<tr>
<td>0</td>
<td>1</td>
<td>20</td>
<td>0.859</td>
<td>0.862</td>
</tr>
<tr>
<td></td>
<td></td>
<td>40</td>
<td>0.909</td>
<td>0.892</td>
</tr>
<tr>
<td></td>
<td></td>
<td>80</td>
<td>0.918</td>
<td>0.915</td>
</tr>
<tr>
<td></td>
<td></td>
<td>160</td>
<td>0.937</td>
<td>0.932</td>
</tr>
<tr>
<td>2</td>
<td>20</td>
<td>0.868</td>
<td>0.831</td>
<td>0.854</td>
</tr>
<tr>
<td></td>
<td></td>
<td>40</td>
<td>0.920</td>
<td>0.875</td>
</tr>
<tr>
<td></td>
<td></td>
<td>80</td>
<td>0.918</td>
<td>0.895</td>
</tr>
<tr>
<td></td>
<td></td>
<td>160</td>
<td>0.932</td>
<td>0.925</td>
</tr>
<tr>
<td>3</td>
<td>20</td>
<td>0.862</td>
<td>0.819</td>
<td>0.838</td>
</tr>
<tr>
<td></td>
<td></td>
<td>40</td>
<td>0.906</td>
<td>0.843</td>
</tr>
<tr>
<td></td>
<td></td>
<td>80</td>
<td>0.908</td>
<td>0.870</td>
</tr>
<tr>
<td></td>
<td></td>
<td>160</td>
<td>0.929</td>
<td>0.911</td>
</tr>
<tr>
<td>0.4</td>
<td>1</td>
<td>20</td>
<td>0.852</td>
<td>0.836</td>
</tr>
<tr>
<td></td>
<td></td>
<td>40</td>
<td>0.904</td>
<td>0.895</td>
</tr>
<tr>
<td></td>
<td></td>
<td>80</td>
<td>0.926</td>
<td>0.906</td>
</tr>
<tr>
<td></td>
<td></td>
<td>160</td>
<td>0.934</td>
<td>0.915</td>
</tr>
<tr>
<td>2</td>
<td>20</td>
<td>0.873</td>
<td>0.787</td>
<td>0.845</td>
</tr>
<tr>
<td></td>
<td></td>
<td>40</td>
<td>0.903</td>
<td>0.843</td>
</tr>
<tr>
<td></td>
<td></td>
<td>80</td>
<td>0.923</td>
<td>0.873</td>
</tr>
<tr>
<td></td>
<td></td>
<td>160</td>
<td>0.933</td>
<td>0.910</td>
</tr>
<tr>
<td>3</td>
<td>20</td>
<td>0.868</td>
<td>0.760</td>
<td>0.850</td>
</tr>
<tr>
<td></td>
<td></td>
<td>40</td>
<td>0.884</td>
<td>0.772</td>
</tr>
<tr>
<td></td>
<td></td>
<td>80</td>
<td>0.908</td>
<td>0.827</td>
</tr>
<tr>
<td></td>
<td></td>
<td>160</td>
<td>0.916</td>
<td>0.873</td>
</tr>
<tr>
<td>0.8</td>
<td>1</td>
<td>20</td>
<td>0.899</td>
<td>0.843</td>
</tr>
<tr>
<td></td>
<td></td>
<td>40</td>
<td>0.917</td>
<td>0.890</td>
</tr>
<tr>
<td></td>
<td></td>
<td>80</td>
<td>0.925</td>
<td>0.913</td>
</tr>
<tr>
<td></td>
<td></td>
<td>160</td>
<td>0.946</td>
<td>0.935</td>
</tr>
</tbody>
</table>

Continued on Next Page…
Table 4.9 – Continued

<table>
<thead>
<tr>
<th>Within-unit correlation</th>
<th>Case</th>
<th>n</th>
<th>Three repeated measure</th>
<th>Five repeated measure</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td>ρg₁^a</td>
<td>ρg₂^b</td>
</tr>
<tr>
<td>2</td>
<td>20</td>
<td>0.892</td>
<td>0.797</td>
<td>0.873</td>
</tr>
<tr>
<td></td>
<td>40</td>
<td>0.904</td>
<td>0.840</td>
<td>0.885</td>
</tr>
<tr>
<td></td>
<td>80</td>
<td>0.909</td>
<td>0.873</td>
<td>0.905</td>
</tr>
<tr>
<td></td>
<td>160</td>
<td>0.933</td>
<td>0.893</td>
<td>0.927</td>
</tr>
<tr>
<td>3</td>
<td>20</td>
<td>0.896</td>
<td>0.778</td>
<td>0.863</td>
</tr>
<tr>
<td></td>
<td>40</td>
<td>0.886</td>
<td>0.757</td>
<td>0.880</td>
</tr>
<tr>
<td></td>
<td>80</td>
<td>0.904</td>
<td>0.786</td>
<td>0.899</td>
</tr>
<tr>
<td></td>
<td>160</td>
<td>0.918</td>
<td>0.838</td>
<td>0.918</td>
</tr>
</tbody>
</table>

^a The scalar version of CCC using the trace function
^b The scalar version of CCC using the largest eigenvalue function
^c The scalar version of CCC using the determinant function
Chapter 5

Examples

In this chapter, we demonstrate the use of the multivariate CCC for measuring an overall agreement between two vectors of repeated measures presented in Chapter 3 using some real examples.

5.1 Blood draws data

The data for this example are taken from an Asthma Clinical Research Network (ACRN) study reported by Martin et al. (2002). The main objective of this trial was to develop a reliable method to compare six different available inhaled corticosteroid (ICS) preparations in terms of systemic bioavailability as measured by effect on cortisol suppression. Three different outcomes were considered to evaluate this systematic effect, namely hourly plasma cortisol concentrations, 12- and 24-hour urine cortisol concentrations, and a morning blood osteocalcin. After one week of placebo run-in period, corticosteroid-naive asthma subjects enrolled at six ACRN centers were randomized to one of the six ICS and matched placebo groups. Following randomization, another placebo week was continued and then the subjects were admitted for an overnight testing at each of the next five weekly visits. During an overnight stay, an out-of-laboratory 12-hour urine collection was conducted between 8 A.M. and 8 P.M. and then in-laboratory urine cortisol collection and hourly blood sampling for cortisol was performed between
8 P.M. and 8 A.M.; blood for osteocalcin concentration was taken at 7 A.M. It was found that the area under the concentration-time curve (AUC) for hourly plasma cortisol measurements was the most reliable method to assess systematic effect. An additional interesting goal was to assess the agreement between the plasma cortisol AUC calculated from measurements taken every hour and measurements taken every two hours. This is very useful for future studies because the every two hour analysis requires less sleep interruption and lower budget. Since the plasma cortisol AUC were obtained at five visits for each subject, a repeated measures CCC is a proper coefficient for measuring the overall agreement between hourly and every two hours measurements based on five visits.

Figure 5.1 shows the summary statistics of the blood draw data for each visit including the Pearson Correlation Coefficients. The scatter plots of the plasma cortisol AUC calculated from the measurements taken every hour and every two hours at visit 3, 4, 5, 6, and 7 are shown in Figure 5.2. The Lin’s sample concordance correlation coefficients and the corresponding 95% confidence intervals for each visit are also included in the graph. The scatter plots and the Lin’s CCCs indicate strong agreement between the plasma AUCs based on the hourly and every other hour data for all five visits. For this data, the point estimates and the 95% confidence interval of the multivariate CCC based on the trace, highest eigenvalue, and determinant function are respectively calculated as follows: \( \hat{\rho}_{g_1} = 0.925 \) (95%CI = (0.909, 0.939)), \( \hat{\rho}_{g_2} = 0.859 \) (95%CI = (0.801, 0.901)), and \( \hat{\rho}_{g_3} = 0.941 \) (95%CI = (0.930, 0.951)). Since the estimate of multivariate CCC based on the determinant function works relatively well for most situations especially when the correlations between X and Y are high which corresponds to this example, we would
recommend to use it to assess an overall agreement for this case. Using the same data, the estimate of the repeated measures CCC proposed by King et al. (2007) is 0.958 with 95% confidence interval = (0.950, 0.965). This result, which is based on the weight matrix consisting of equal on-diagonal elements and zero off-diagonal elements, is close to the estimate of the multivariate CCC based on the determinant function. According to King et al. (2007), the point estimate of the weighted CCC created by Chinchilli et al. (1996) for these data is 0.971 and the corresponding 95% confidence interval is (0.962, 0.977). All of these results suggest high level of overall agreement between the two sets of plasma AUCs.

5.2 Body fat data

For this example, we use the data from the Penn State Young Women’s Health Study conducted by Lloyd et al. (1998). In this data, percentages of body fat are obtained from 82 white female subjects at age 12, 12.5, and 13 years based on whole-body composition measurements made by dual-energy x-ray absorbtiometer (DEXA) and skinfold caliper. The summary statistics of these data are shown in Figure 5.3. Figure 5.4 shows the scatter plots of the percentage of body fat along with the estimates of Lin’s CCC and 95% confidence intervals, indicating moderate agreement for all three visits. Following are the point estimates of the multivariate CCC based on the trace, highest eigenvalue, and determinant function along with the corresponding 95% confidence intervals: \( \hat{\rho}_{g1} = 0.476 \) (95%CI = (0.365, 0.574)), \( \hat{\rho}_{g2} = 0.217 \) (95%CI = (0.001, 0.414)), and \( \hat{\rho}_{g3} = 0.528 \) (95%CI = (0.438, 0.607)). For this example, we also recommend the multivariate CCC based on the determinant function as a measure of overall agreement. Based
on this data, the estimate of the repeated measures CCC using the approach suggested by King et al. (2007) is 0.568 and the 95% confidence interval is (0.406, 0.695). This point estimate is similar to that of our determinant multivariate CCC but the corresponding 95% confidence interval is much wider than our result. Reported by King et al. (2007), the weighted CCC estimate proposed by Chinchilli et al. (1996) for the same data is 0.597 with the corresponding 95% confidence interval = (0.514, 0.661). These results are similar in that they indicate moderate agreement between the percentage of body fat measured by the DEXA and skinfold caliper.
Fig. 5.1: Summary statistics of blood draw data for each visit with Pearson Correlation Coefficients.
Fig. 5.2: Scatter plots of blood draw data for each visit with Lin's CCCs.
Fig. 5.3: Summary statistics of body fat data for each visit with Pearson Correlation Coefficients.
Fig. 5.4: Scatter plots of body fat data for each visit with Lin’s CCCs.
Chapter 6

Future Studies

We have introduced an index of overall agreement between two responses in the presence of repeated measurements which is an extension of Lin’s concordance correlation coefficient. First, we developed a matrix that possesses the properties needed for assessing the amount of agreement between two vectors of random variables. For ease of interpretation we used a function to transform this matrix to a scalar whose value is scaled to range between -1 and 1. In this work, we have considered three distinct functions, namely trace, highest eigenvalue, and determinant. We called this repeated measures CCC “the multivariate concordance correlation coefficient”. This multivariate CCC not only has desirable characteristics but also is easy to use. For inference, we constructed an asymptotically unbiased estimator based on U-statistics and derived its asymptotic distribution for each form of the function. An estimator of its asymptotic variance has also been proposed for obtaining confidence intervals or testing hypotheses. Moreover, we used the Z-transformation to bound the confidence limits and improve the rate of convergence. The simulation results confirmed that overall in terms of accuracy, precision, and the coverage probabilities, the estimator of the multivariate CCC based on the determinant function works relatively well in general cases even with small samples. However, for a skewed underlying distribution with moderate or weaker correlation between the two variables, the trace multivariate CCC is slightly more robust.
To handle both symmetric and asymmetric underlying distribution, a weighted function based on skewness might be considered to form a mixture of multivariate CCCs based on both trace and determinant functions. Furthermore, when the data are obtained by stratified random sampling, where each sample comes from different sub-population, one may need a weighted average of the multivariate CCCs to evaluate overall agreement across strata. In addition, the multivariate CCC may be generalized to evaluate agreement among more than two vectors of variables. These topics for extensions will be explored in future work.
Bibliography


Vita

Sasiprapa Hiriote

Education

- Ph.D. in Statistics, The Pennsylvania State University, August 2009
- M.S. in Statistics, Chulalongkorn University, Thailand March 1997
- B.S. in Statistics, Silpakorn University, Thailand March 1994

Academic Experience at The Pennsylvania State University

- Fall 2004 - Spring 2009 Teaching Assistant at the Statistics Department
- Summer 2007 - Fall 2007 Research Assistant at the Public Health Science Department