BAYESIAN INFERENCE FOR CRONBACH'S ALPHA

A thesis submitted in fulfillment of the requirements for the degree of

Doctor of Philosophy

in

Mathematical Statistics

Department of Statistics

Rhodes University

by

Sharkay Ruwade Izally

Student number: G10I3478

ORCID iD: 0000-0002-7686-138X

December 2024

Supervisor: Prof. L. Raubenheimer

Co-Supervisor: Prof. A. J. van der Merwe

Declaration

I, the undersigned, declare that the work contained in this thesis is my own work, except for references specifically indicated in the text, and that I have not previously submitted it elsewhere for degree purposes.

S. R. Izally

Date

Acknowledgements

I would like to acknowledge and express my gratitude to my supervisor Prof. Raubenheimer. This thesis would not have been possible without her meaningful advice, patience and motivation. I cannot thank her enough for all her efforts, emotional support and guidance. I am grateful to Prof. Raubenheimer for sharing her excellent knowledge with me and without her ideas this thesis would not have been possible. I would like to thank my co-supervisor Prof. Abrie van der Merwe for all the hours we spent on the phone discussing ideas and going through proofs with me. All my visits to Prof Abrie significantly improved my research. I want to thank Prof. Piet Groenewald for his valuable input on all the computational work. I also want to thank my colleagues, Rene, Jeremy, Amos, Sisipho, Amy and Rakesh, who have been very supportive and encouraging throughout this journey. The amount of support that I have received from of each of my colleagues was just so wonderful to experience. I thank my parents, Norman and Lydia for all the words of encouragement and prayers, my siblings, Heinrich and Sharday for always showing interest and my niece Shakira. I would like to thank my aunt, Teresa for all the special prayers.

I would like to express my gratitude to the Rhodes University UCDP and UCDP-Mass funding that assisted me in attending conferences and all my co-supervisor visits.

Lastly, I would like to thank God for guiding me and pulling me through to complete my thesis. Without God, none of my achievements would be possible.

Research Outputs

Academic Articles

1. Izally, S. R., Van der Merwe, A. J. and Raubenheimer, L. (2024). A comparison of objective priors for Cronbach's coefficient alpha using a balanced random effects model, *Communications in Statistics – Theory and Methods*, DOI: 10.1080/03610926.2024.2315300.

2. Van der Merwe, A. J., Izally, S. R. and Raubenheimer, L. (2024). Bayesian process control using Cronbach's alpha, revised paper submitted to *The South African Statistical Journal*.

3. Van der Merwe, A. J., Izally. S. R. and Raubenheimer, L. (2024). Bayesian estimation of Cronbach's alpha in the case of more than one balanced random effects experiment, revised paper submitted to *Communications in Statistics – Theory and Methods*.

4. Izally, S. R., Van der Merwe, A. J. and Raubenheimer, L. (2024). A comparison of objective priors for Cronbach's coefficient alpha using a three-component hierarchical model, submitted to *Communications in Statistics – Theory and Methods*.

5. Izally, S. R., Van der Merwe, A. J. and Raubenheimer, L. (2024). Fiducial and Bayesian estimation of Cronbach's alpha, submitted to Afrika Mathematika.

Conference Presentations

1. Izally, S. R., Van der Merwe, A. J., and Raubenheimer, L. (2023). Objective priors for Cronbach's coefficient alpha using a balanced random effects model, *The South African Statistical Association (SASA) Conference* hosted by the University of KwaZulu-Natal, Durban, South Africa, November 2023.

2. Izally, S. R., Van der Merwe, A. J., and Raubenheimer, L. (2023). Loss functions for Cronbach's coefficient alpha using a balanced random effects model, *The Bayesian Young Statisticians Meeting*, Virtual Conference, November 2023.

3. Van der Merwe, A. J., Izally, S. R., and Raubenheimer, L. (2024). Bayesian estimation of Cronbach's alpha in the case of more than one balanced random effects model, *The International*

Society for Bayesian Analysis (ISBA) Conference, The University of Venice, Venice, Italy, July 2024.

4. Izally, S. R., Van der Merwe, A. J., and Raubenheimer, L. (2024). Fiducial and Bayesian estimation of Cronbach's alpha, *The South African Statistical Association (SASA) Conference* hosted by the University of the Western Cape, Stellenbosch, South Africa, November 2024.

5. Izally, S. R., Van der Merwe, A. J. and Raubenheimer, L. (2025). A comparison of objective priors for Cronbach's coefficient alpha using a three-component hierarchical model, abstract submitted for presentation at the Objective Bayes Methodology Conference, Athens, Greece, June 2025.

Abstract

Cronbach's alpha is used as a measure of reliability in fields like education, psychology and sociology. The reason for the popularity of Cronbach's alpha is that it is computationally simple. Only the sample size and the variance components are needed and it can be computed for continuous as well as binary data. Cronbach's alpha has been studied extensively using maximum likelihood estimation. Since Cronbach's alpha is a function of the variance components, this often results in negative estimates of the variance components when the maximum likelihood method is considered as a method of estimation. In the field of Bayesian statistics, the parameters are random variables, and this can alleviate some of the problems of estimating negative variance estimates that often occur when the frequentist approach is used. The Bayesian approach also incorporates loss functions that considers the symmetry of the distribution of the parameters being estimated and adds some flexibility in obtaining better estimates of the unknown parameters. The Bayesian approach often results in better coverage probabilities than the frequentist approach especially for smaller sample sizes and it is therefore important to consider a Bayesian analysis in the estimation of Cronbach's alpha. The reference and probability matching priors for Cronbach's alpha will be derived using a one-way random effects model. The performance of these two priors will be compared to that of the well-known Jeffreys prior and a divergence prior. A simulation study will be considered to compare the performance of the priors, where the coverage rates, average interval lengths and standard deviations of the interval lengths will be computed. A second simulation study will be considered where the mean relative error will be compared for the various priors using the squared error, the absolute error and the linear in exponential (LINEX) loss functions. An illustrative example will also be considered. The combined Bayesian estimation of more than one Cronbach's alpha will also be considered for m experiments with equal α but possibly different variance components. It will be shown that the reference and the probability-matching priors are the same. The Bayesian theory and results will be applied to two examples. The intervals for the combined model are however much shorter than those of the individual models. Also, the point estimates of the combined model are more accurate than those of the individual models. It is further concluded that the posterior distribution of α for the combined model becomes more important as the number of samples and models increase. The reference and probability matching priors for Cronbach's alpha will be derived using a three-component hierarchical model. The performance of these two priors will be compared to that of the well-known Jeffreys prior and a divergence prior. A simulation study will be considered to compare the performance of the priors, where the coverage rates, average interval lengths and standard deviations of the interval lengths will be computed. Two illustrative examples will also be considered. Statistical control limits will be obtained for Cronbach's alpha in the case of a balanced one-way random effects model. This will be achieved by deriving the predictive distribution of a future Cronbach's alpha. The unconditional posterior predictive distribution will be determined using Monte Carlo simulation and the Rao-Blackwell procedure. The predictive distribution will be used to obtain control limits and to determine the run-length and average run-length. Cronbach's alpha will be estimated for a general covariance matrix using a Bayesian approach and comparing these results to the asymptotic frequentist interval valid under a general covariance matrix framework. Most of the results used in the literature require the compound symmetry assumption for analyses of Cronbach's alpha. Fiducial and posterior distributions will be derived for Cronbach's alpha in the case of the bivariate normal distribution. Various objective priors will be considered for the variance components and the correlation coefficient. One of the priors considered corresponds to the fiducial distribution. The performance of these priors will be compared to an asymptotic frequentist interval often used in the literature. A simulation study will be considered to compare the performance of the priors and the asymptotic interval, where the coverage rates and average interval lengths will be computed.

Keywords: Absolute error loss, Bayesian process control, Compound symmetry, Coverage, Cronbach's alpha, Divergence prior, Fiducial distribution, Hierarchical model, Jeffreys prior, LINEX loss, One-way random effects model, Predictive distribution, Probability matching prior, Reference prior, Run-length, Squared error loss.

Contents

List of Figures

List of Tables

List of Abbreviations

ADF	Asymptotically Distribution-Free
ANOVA	Analysis of Variance
HPD	Highest Posterior Density
LINEX	Linear in Exponential
MCMC	Markov Chain Monte Carlo
MLE	Maximum Likelihood Estimator
MRE	Mean Relative Error
MSE	Mean Squared Error
PMP	Probability Matching Prior

Mathematical Notation

log	Natural Logarithm
E	Expected Value of a Random Variable
Var	Variance of a Random Variable
Θ	Parameter Space
Σ	Variance-Covariance Matrix
$\ell(\theta)$	Likelihood Function of θ
$L(\theta)$	Log of the Likelihood Function of θ
$I_{(a,b)}(\theta)$	Indicator Function for θ Over the Interval (a,b)
$f(x\|\theta)$	Probability Density Function of x given θ
$\pi(\theta)$	Prior Distribution of θ
$\pi(\theta\|data)$	Posterior Distribution of θ
$F(\theta)$	Fisher Information Matrix of θ
$F^{-1}(\theta)$	Inverse of the Fisher Information Matrix of θ
$trace(\Sigma)$	Trace Operator of a Square Matrix
$\mathbf{X}'$	Transpose Operator of a Vector or a Matrix $\mathbf{X}$
R^k	The Euclidean space

Chapter 1

Introduction

Cronbach's alpha is a measure used to assess the reliability of a set of test items. Reliability refers to the consistency or stability of a measurement instrument. In other words, if you were to administer the same test multiple times, the results should be similar and stable. A higher Cronbach's alpha suggests that the items are highly consistent with one another, meaning they reliably measure the same concept and a lower Cronbach's alpha suggests that the items are not consistent and may not all be measuring the same underlying concept. Cronbach's alpha is well-known and by far the most frequently used and reported index of reliability. According to Cortina (1993), Cronbach's (1951) article has been cited approximately sixty times per year in the time period 1966 – 1990 and in a total of two hundred and seventy eight different journals. Recently it was cited 69 919 times on Google Scholar (Accessed on February 26, 2025). Cronbach's alpha is computed by correlating the score for each scale item with the total score for each observation and then comparing that to the variance of all individual item scores. The resulting coefficient of reliability typically ranges from zero to one although there are cases where Cronbach's alpha can be negative. In practice, it is difficult to compare values of alpha across studies and computing confidence or credibility intervals will thus be considered. Cronbach's alpha will be estimated using Bayesian inference. It is well-known that the posterior is proportional to the prior times the likelihood when using Bayesian inference. The choice of the prior distribution is critical in a Bayesian analysis, although this can also be the most difficult part. If the prior is chosen such that it influences the posterior minimally and the posterior is dominated by the data, then this prior is known as an objective prior. Objective priors are also known as non-informative, vague or flat priors. When an objective prior is considered then the results often mirror frequentist results. In order to compare frequentist results to Bayesian results, it makes sense to use objective priors to do so. The main focus of this research will be on estimating Cronbach's alpha using objective priors. The frequentist results often rely on large sample theory and the Bayesian approach can play a crucial role in the estimation of Cronbach's alpha for smaller sample sizes as well as for larger sample sizes.

1.1 Objectives

The main objectives and aims can be summarized as follows:

- to provide an overview of some objective priors and Cronbach's coefficient alpha;
- to derive the Fisher information matrix involving Cronbach's alpha for the following cases: the one-way analysis of variance model, an extension of the one-way model with possibly different variance components and a balanced two-factor nested random effects (hierarchical) model;
- to derive the probability matching prior and the reference prior using the Fisher information matrix under each of the above-mentioned cases;
- to compare the performance of the probability matching prior and the reference prior to other objective priors and to frequentist methods;
- to show the properness of the posteriors resulting from the different objective priors;
- to derive statistical process control limits for Cronbach's coefficient alpha in the case of the balanced one-way random effects model by deriving a predictive distribution of a future Cronbach's alpha coefficient using Bayesian inference and using the predictive distribution to determine the distribution of the run-length and the average run-length;
- to compare the performance of the Bayesian intervals with that of various frequentist intervals;
- to see how these Bayesian methods can be implemented in practice, and actually improve research in other fields.

1.2 Outline

In **Chapter 2** a literature review is provided giving a background on Cronbach's alpha and how it is estimated using frequentist methods and Bayesian methods in the literature. The chapter also summarizes some of the methods and notations used to derive objective priors in this thesis.

In **Chapter 3** a number of objective priors for Cronbach's alpha will be derived using a balanced one-way analysis of variance model. The following objective priors will be investigated: the well-known Jeffreys prior (from Jeffreys, 1939), a reference prior (see Berger & Bernardo, 1992 and Berger et al., 2009), the probability matching prior using the method proposed by Datta & Ghosh (1995) and a divergence prior developed by Ghosh et al. (2011). A reference prior for grouping order $\left\{\alpha, \theta, \sigma_1^2\right\}$, which means that α is the most important parameter and σ_1^2 is the least important parameter, will be derived using the algorithm in Berger & Bernardo (1992). The marginal posterior for α will be derived using the different priors and a simulation study will be conducted where the frequentist coverage rates, the interval lengths and the standard deviation of the interval lengths will be computed to compare the different priors. A second simulation study will be conducted to evaluate the performance of the Bayes estimators using the well-known squared error loss function, the absolute error loss function and the LINEX loss function, which was first introduced by Varian (1975).

In **Chapter 4** the combined Bayesian estimation of α for m experiments with equal α but possibly different variance components will be derived. The priors that will be considered are the reference prior by Berger & Bernardo (1992) and the probability matching prior using the method by Datta & Ghosh (1995). Since our model is a one-way balanced random effects model, the assumption of equicorrelated normal data is satisfied. Our approach is therefore different from the method of Van Zyl (2001) who had to make the assumption of equicorrelation and Van Zyl (2001) only considered a uniform prior. The work is therefore an extension of the work done in Chapter 3 and also an extension of the work done by Van Zyl (2001) .

In **Chapter 5** several objective priors for Cronbach's alpha will be derived under the balanced two-factor nested random effects (hierarchical) model of Van der Merwe & Hugo (2007). The marginal posterior for α will be derived using the different priors and a simulation study will be conducted where the frequentist coverage rates, the interval lengths and the standard deviation of the interval lengths will be computed to compare those of the different priors. This chapter is an extension of the work done by Izally et al. (2024).

In **Chapter 6** statistical process control limits will be obtained for Cronbach's coefficient alpha in the case of the balanced one-way random effects model. This will be achieved by deriving a predictive distribution of a future Cronbach's alpha coefficient using the Bayesian methods developed. The prior that will be considered in the Bayesian analysis is the Jeffreys independence prior from Box & Tiao (1973). The predictive distribution will also be used to determine the distribution of the run-length and the average run-length.

In **Chapter 7** the focus is on the estimation of Cronbach's alpha for a general covariance matrix using a Bayesian approach and comparing these results to the asymptotic frequentist interval valid under a general covariance matrix framework. In this chapter fiducial and posterior distributions will be derived for Cronbach's alpha in the case of the bivariate normal distribution. Various objective priors will be considered for the variance components and the correlation coefficient. It will be shown that the right-Haar prior corresponds to the fiducial distribution. The performance of these priors will be compared to an asymptotic frequentist interval derived by Van Zyl et al. (2000), which is often used in the literature. A simulation study will be considered to compare the performance of the priors and the asymptotic interval, where the coverage rates and average interval lengths will be computed.

In **Chapter 8** the distributions for Cronbach's coefficient alpha and the intra-class correlation will be derived in the case of single coefficients as well as for the case when there are two coefficients, using a one-way random effects model. Since there is a mathematical relationship between Cronbach's alpha and the intra-class correlation coefficient, the work that will be done in this chapter is an extension of the research done by Koning & Franses (2003) on Cronbach's coefficient alpha and the research done by Chung & Dey (1998) on the intra-class correlation coefficient.

In **Chapter 9** the conclusions are summarized for each chapter and possible future research in this area is proposed.

Appendix A contains the derivation of the Fisher information matrix for Cronbach's alpha using the one-way random effects model. It also contains the proofs for the reference prior for the group ordering: $\{\alpha, \sigma_1^2, \theta\}$,$\{\sigma_1^2, \theta, \alpha\}$, $\{\sigma_1^2, \alpha, \theta\}$, $\{\theta, \sigma_1^2, \alpha\}$ and $\{\theta, \alpha, \sigma_1^2\}$. The additional results for the simulation studies are also provided. The MATLAB code for the simulation studies and the example are provided in Appendix A.

Appendix B contains the MATLAB code for two examples and the MATHEMATICA code of all the figures for the combined Bayesian estimates for Cronbach's alpha.

Appendix C contains the derivation of the Fisher information matrix for Cronbach's alpha using a three-component hierarchical model. It also contains the proofs for the reference prior for the group orderings $\left\{\alpha, \sigma_1^2, \theta, \sigma_2^2\right\}$ and $\left\{\alpha, \theta, \sigma_1^2, \sigma_2^2\right\}$. Additional results and MATLAB code for the simulation study are provided as well as the MATLAB code for the two examples considered.

Appendix D contains the MATLAB code for the simulations done for Bayesian process control of Cronbach's alpha.

Appendix E contains the derivation of the fiducial distributions and the posterior distributions for $(\sigma_1, \sigma_2, \rho)$. MATLAB code for the simulation study is also provided.

Appendix F contains the derivation of the distributions of Cronbach's alpha and the intra-class correlation coefficient. The proof of the mean and variance of the statistic, $\frac{1-\alpha_1}{1-\alpha_2}$ is also provided. MATLAB code for all the figures and results is also provided.

Chapter 2

Literature Review

2.1 Cronbach's Alpha

Cronbach's coefficient alpha was introduced in Cronbach's (1951) article based on the work of Guttman (1945). It is important to ensure that instruments or tools used to collect data (such as surveys, tests, or questionnaires) produce consistent and dependable results. When conducting research, you are often measuring a specific concept or construct, such as intelligence, motivation, or satisfaction. However, these constructs can't directly be observed, so researchers create measurement tools, for example, a set of survey questions, to assess them. To determine the accuracy of measurement it makes sense to take two measurements and compare them. Cronbach (1951) mentions that it can be time consuming and costly to do this in practice. One of the methods used to alleviate this problem was to rescore a test, half the items at a time, to get two estimates. This is the split-half approach. The Spearman-Brown formula is then applied to get a coefficient similar to the correlation between the two forms. Cronbach (1951) states that this procedure was used for forty years. Kuder & Richardson (1937) criticized the split-half approach because it did not give a single coefficient for a test, instead, it gave different coefficients depending on which items are grouped when the test is split into two. In an attempt to improve on this, Kuder & Richardson (1937) developed the following formula

$$\alpha_{KR20} = \frac{n}{n-1}\left(1 - \frac{\sum_{i=1}^{n} p_i q_i}{\sigma_t^2}\right), \tag{2.1}$$

where i represents an item, p_i is the proportion receiving a score of 1, q_i is the proportion receiving a zero score on the item and σ_t^2 is the variance of the total test scores. Equation 2.1 is used to estimate the reliability of tests with binary items. Cronbach (1951) then generalized this formula by writing it as

$$\alpha = \frac{n}{n-1}\left(1 - \frac{\sum_{i=1}^{n} V_i}{V_t}\right), \tag{2.2}$$

where V_t is the variance of the test scores and V_i is the variance of item scores after weighting. The item variance,V_i, represents how much individual items vary in relation to the overall test. Since the items are binary, the variance V_i depends on the proportion of correct answers, p_i, and incorrect answers, q_i, for each item. The weighting occurs since each item's variance is not treated the same as the others which results in some items contributing more to the overall test score variance than others. Equation 2.2 reduces to Equation 2.1 when all items are scored one or zero. Van Zyl (2001) defines Cronbach's alpha using the normal distribution below. Suppose that $\mathbf{X}' = (X_1, ..., X_k)$ and $E(\mathbf{X}) = \mathbf{0}$, where $\mathbf{X_i}$ refers to an item in the test, with a variance-covariance matrix given by Σ. Then the covariance matrix, with the assumption of equicorrelation, is given by

$$\Sigma = \sigma^2 \begin{bmatrix} 1 & \rho & \rho & \dots & \rho \\ \rho & 1 & \rho & \dots & \rho \\ \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \rho & \rho & \rho & \cdots & 1 \end{bmatrix}_{k \times k} .$$

When the covariance matrix has equal variances and equicorrelation, this property is referred to a compound symmetry. Cronbach's alpha is therefore given by

$$\begin{aligned} \alpha &= \frac{k}{k-1}\left(1 - \frac{\sum_{i=1}^{k} Var(X_i)}{\text{Total Variance}}\right) \\ &= \frac{k}{k-1}\left(1 - \frac{\sum_{i=1}^{k} Var(X_i)}{\mathbf{1}'\Sigma\mathbf{1}}\right) \\ &= \frac{k}{k-1}\left(1 - \frac{trace(\Sigma)}{\mathbf{1}'\Sigma\mathbf{1}}\right) \\ &= \frac{k\rho}{1+(k-1)\rho}, \end{aligned} \tag{2.3}$$

where $\mathbf{1}$ is a $k \times 1$ column vector of ones. From Equation 2.3 it is clear that

$$1 - \alpha = \frac{1-\rho}{1+(k-1)\rho}$$

and

$$\rho = \frac{\alpha}{k - (k-1)\,\alpha}. \tag{2.4}$$

Equation 2.3 is commonly known as the Spearman-Brown formula. This formula was published independently by Spearman (1910) and Brown (1910). This mathematical relationship only exists under compound symmetry. Li & Woodruff (2002) mention that Cronbach's alpha equals reliability rather than just being a lower bound for reliability under compound symmetry. Equation 2.4 shows how the correlation coefficient relates to Cronbach's alpha.

2.2 Frequentist Methods

The method of maximum likelihood is a well-known tool for parameter estimation given the sample data. Suppose that random variables $X_1, \ldots, X_n$ have a joint density $f(x_1, x_2, \ldots, x_n | \boldsymbol{\theta})$, where the vector $\boldsymbol{\theta} = (\theta_1, \theta_2, \ldots \theta_k)$, and θ_i for $i = 1, \ldots k$ are the unknown parameters. Assume the X_i are independent and identically distributed random variables. Given the observed values $X_i = x_i$, for $i = 1, \ldots, n$, the likelihood function of $\boldsymbol{\theta}$ is given by

$$\ell(\boldsymbol{\theta}|x_1, x_2, \ldots, x_n) = \prod_{i=1}^{n} f(x_i|\boldsymbol{\theta}). \tag{2.5}$$

The maximum likelihood estimator (MLE) of $\boldsymbol{\theta}$ is the value of $\boldsymbol{\theta}$ that maximizes the likelihood or makes the observed data most probable. It is usually easier to maximize the natural logarithm of the likelihood function and this is given by

$$L = \log \ell(\boldsymbol{\theta}|x_1, x_2, \ldots, x_n) = \sum_{i=1}^{n} \log f(x_i|\boldsymbol{\theta}). \tag{2.6}$$

To find MLE's the partial derivatives of the log likelihood function should be taken with respect to the unknown parameter, equate them to zero and solve for the unknown parameters. This method works well for large samples of data. The maximum likelihood estimator for Cronbach's alpha, derived in Van Zyl et al. (2000), is given by

$$\hat{\alpha}_{MLE} = \frac{k}{k-1}\left(1 - \frac{trace(S)}{\mathbf{1}'S\mathbf{1}}\right),$$

where S is the unbiased estimator of Σ given by $S = \frac{1}{n-1}A$, where A has a Wishart distribution with parameters Σ and $n-1$ degrees of freedom, that is $A \sim W_k(n-1, \Sigma)$. A is a sample covariance matrix that is based on n independent and identically distributed samples from a multivariate normal distribution with covariance matrix, Σ. See Van Zyl et al. (2000) for more details.

2.3 Bayesian Methods

Bayesian statistics provide mathematical tools to update prior beliefs based on new observed data. In Bayesian statistics, the unknown parameter θ is treated as a random variable which follows some distribution. This distribution is known as the prior distribution, $\pi(\theta)$. The prior distribution represents what is known about the parameter before observing the data. Suppose that n continuous random variables, $\mathbf{X} = (X_1, X_2, ..., X_n)$, form a random sample from the distribution $\pi(\mathbf{x}|\theta)$, it then follows that their joint distribution will be given by

$$\pi(\mathbf{x} \mid \theta) = \pi(x_1, x_2, \ldots, x_n \mid \theta) = \pi(x_1 \mid \theta) \ldots \pi(x_n \mid \theta).$$

Then, using Bayes theorem

$$\begin{aligned} \pi(\theta|\mathbf{x}) &= \frac{\pi(\mathbf{x}|\theta)\pi(\theta)}{\pi(\mathbf{x})} \\ &= \frac{\pi(\mathbf{x} \mid \theta)\pi(\theta)}{\int \pi(\mathbf{x} \mid \theta)\pi(\theta)\, d\theta} \\ &= \frac{\ell(\theta|\mathbf{x})\pi(\theta)}{\int \ell(\theta|\mathbf{x})\pi(\theta)\, d\theta}, \end{aligned}$$

where $\ell(\theta|\mathbf{x})$ is the likelihood function as defined in Equation 2.5. The left-hand side of the equation, $\pi(\theta|\mathbf{x})$, is known as the posterior distribution and represents what is known about the parameter after the data had been observed. Since the denominator is not a function of θ it is conventional to write

$$\underbrace{\pi(\theta|\mathbf{x})}_{\text{Posterior}} \propto \underbrace{\ell(\theta|\mathbf{x})}_{\text{Likelihood}} \times \underbrace{\pi(\theta)}_{\text{Prior}}. \tag{2.7}$$

There can be more than one unknown parameter in which case θ will be denoted as a vector $\boldsymbol{\theta}$. Equation 2.7 shows that a Bayesian analysis involves two components, the likelihood which represents the data and a prior distribution. In the frequentist paradigm, the analysis is only based on the likelihood and the unknown parameter is considered as an unknown constant, which is in contrast to a Bayesian analysis where the unknown parameter represents a random variable. The difference is fundamentally in where the uncertainty lies. In Bayesian statistics, the uncertainty is over θ. In frequentist, it is over other data sets that might be have been observed. Bayesian inference therefore changes the interpretation of a confidence interval. When using Bayesian inference it is called a credibility interval. A credibility interval for θ of level $1-\beta$ can be interpreted as θ is in the interval with a probability of $1-\beta$.

When Bayesian models are specified, a decision must be made on which priors to use for the unknown parameters. As mentioned by Robert (2001), the most critical and most criticized point of Bayesian analysis deals with the choice of the prior distribution. Goligher & Harhay (2024) mention that the most difficult part of executing a Bayesian analysis is selecting priors. Goligher & Harhay (2024) also mention that prior selection is so important because it is the only place that a Bayesian analysis can go wrong. Objective Bayes is a framework that can be used to derive priors for the unknown parameters and will be used in this thesis. When the prior has minimal effect such that the posterior is dominated by the data then the prior used is known as an objective prior. The result of using an objective prior is a posterior distribution that often mirrors frequentist results, the only difference is in the interpretation of the results. Objective priors are used when little or no prior information is available. Objective priors are often improper which means the prior does not integrate to one or is not a proper density. This is usually not a problem as long as the posterior distribution results in a proper density. Daniels (1999) mention that the choice of prior distributions for variance components can be difficult. It is also mentioned how important it is for these priors to result in proper posterior distributions and how important it is for the posterior to be proper in order to use Markov Chain Monte Carlo methods such as Gibbs sampling. We will investigate a number of objective priors in this thesis and confirm in each case whether they are proper or not. This thesis will also focus on simulating from the posterior distribution without using MCMC methods, but instead using Monte Carlo simulation methods. This will alleviate any computational challenges that usually occur in the literature when using MCMC to simulate variance components.

2.3.1 The Jeffreys Prior

Jeffreys (1939) proposed this prior as a solution to the problem that the uniform prior does not yield an analysis invariant to the choice of parameterization. It is invariant to reparameterization, meaning that if the Jeffreys prior in one parameterization is transformed to a different parameterization, then the transformed prior will be the Jeffreys prior in the new parameterization. The Jeffreys prior is proportional to the square root of the determinant of the Fisher information matrix and will be denoted by π_J. Berger & Bernardo (1992) mention that Jeffreys prior is remarkably successful in one-dimensional problems and that Jeffreys himself noticed difficulties with the method in a multi-parameter setting.

2.3.2 The Divergence Prior

Ghosh et al. (2011) set out to find an objective prior which approximately maximizes the distance between the prior and the posterior. They maximized the distance to ensure that the prior distribution would have as little effect on the posterior as possible. The general expected divergence between the prior and the posterior after observing $\mathbf{x} = (x_1, \ldots, x_n)$ is given by,

$$R^{\beta}(\pi) \propto \frac{1 - \int \left[\int \pi^{\beta}(\theta)\, \pi^{1-\beta}(\theta|\mathbf{x})\, d\theta\right] m(\mathbf{x})\, \mu(d\mathbf{x})}{\beta(1-\beta)}, \tag{2.8}$$

where $m(\mathbf{x})$ is the marginal probability density function of $\mathbf{x}$ and $\mu(d\mathbf{x}) = (\mu(dx_1), \ldots, \mu(dx_n))$. Equation 2.8 is the Rényi divergence between two probability distributions, introduced by Rényi (1961). Mohammed-Djafari (2015) mention that the Rényi divergence is a generalization of the Kullback-Leibler divergence. Ghosh et al. (2011) noted that maximizing all general divergence criterion produced the Jeffreys rule prior, except for one special case - the chi-squared divergence. All other criteria only required a first order approximation to produce the prior, whereas the chi-squared divergence required a second-order approximation. The chi-squared distance is found by substituting $\beta = -1$ into Equation 2.8. The divergence prior for a single parameter θ, which will be denoted by π_D, was derived to be

$$\pi_D \propto \exp\left[\int \frac{2g_3(\theta) - F'(\theta)}{4F(\theta)} d\theta\right],$$

where $F(\theta) = -E\left[\frac{\partial^2 \ln \ell(\theta|\mathbf{x})}{\partial \theta^2}\right]$ is the Fisher Information matrix and $g_3(\theta) = -E\left[\frac{\partial^3 \ln \ell(\theta|\mathbf{x})}{\partial \theta^3}\right]$, where $\ell(\theta|\mathbf{x})$ is the likelihood function. After checking the priors for the exponential family of distributions, which included the binomial and Poisson distributions, these priors were compared to that of Jeffreys prior. This leads to the conclusion that the divergence prior is proportional to the positive fourth root of the determinant of the Fisher Information matrix. The Fisher information, $F(\theta)$, for the Poisson distribution and the binomial distribution is easy to compute. The term $g_3(\theta)$ simplifies to the first derivative of the Fisher information for any of the distributions that belong to the one-parameter exponential family of distributions. For this reason, the divergence prior results in being proportional to the positive fourth root of the determinant of the Fisher Information matrix. The method used to derive the divergence prior does not incorporate the importance of the parameters so similarly to the Jeffreys rule prior, the divergence prior might not perform optimally when there is more than one unknown parameter.

2.3.3 The Reference Prior

Bernardo (1979) investigated the reference prior approach to develop objective priors. For a multi-dimensional $\boldsymbol{\theta} = (\theta_1, \theta_2, ..., \theta_k)$, Bernardo (1979) divided $\boldsymbol{\theta}$ into parameters of interest and nuisance parameters. This enabled the development of an objective prior. Similar to the Jeffreys prior, this approach was based on the Fisher information matrix. Berger & Bernardo (1992) mention that this approach resulted in the Jeffreys prior in the one-dimensional case. The method to determine the reference prior by Bernardo (1979) is to first find the conditional reference prior for the nuisance parameters given the parameters of interest and then to determine the reference prior for the parameters of interest by integrating out the nuisance parameters. Berger & Bernado (1992) mention a limitation of the method by Bernardo (1979) which is that the conditional reference prior found in the first step is often improper. Berger & Bernardo (1992) showed that the improper conditional reference priors must be normalized. They also recommended that all parameters in the model must be grouped and ordered in terms of importance so that the reference prior is determined through a series of one-dimensional conditional steps. The following method from Berger & Bernardo (1992) describes the algorithm and the notation needed to derive the reference prior when each parameter is in its group and ordered in terms of their importance: Consider $\boldsymbol{\theta}$ as an ordered parameterization, $(\theta_1, \theta_2, ..., \theta_m)$. By successive conditioning, a reference prior relative to this ordered parameterization is given by

$$\pi(\boldsymbol{\theta}) = \pi(\theta_m|\theta_1, ..., \theta_{m-1}) \dots \pi(\theta_2|\theta_1)\pi(\theta_1).$$

Consider a parametric statistical problem in which the random observation $\mathbf{X}$ has density $f(\mathbf{x}|\boldsymbol{\theta})$ where $\boldsymbol{\theta} \in \Theta \subset R^k$ is the unknown parameter and the Fisher information matrix

$$H(\boldsymbol{\theta}) = -E_{\mathbf{x}|\boldsymbol{\theta}} \left\{ \frac{\partial^2}{\partial \theta_i \partial \theta_j} logp(\mathbf{x}|\boldsymbol{\theta}) \right\}$$

has full rank k. Strictly this defines $H_{ij}(\theta)$, the $(i, j)^{th}$ element of the information matrix. Define $S(\boldsymbol{\theta}) = H^{-1}(\boldsymbol{\theta})$ and define the component vectors

$$\boldsymbol{\theta}^{[j]} = (\theta_1, ..., \theta_j), \quad \boldsymbol{\theta}_{[j]} = (\theta_{j+1}, ..., \theta_m)$$

and denote $S_j(\boldsymbol{\theta})$ as the corresponding upper left $N_j \times N_j$ corner of S, with $S_m \equiv S$ and $H_j \equiv S_j^{-1}$. We assume that the θ_i are separated into m groups of sizes $n_1, ..., n_m$ and that these groups are given by

$$\theta_{(1)} = (\theta_1, ..., \theta_{n_1}), \theta_{(2)} = (\theta_1, ..., \theta_{n_2}), ...,$$

$$\theta_{(i)} = (\theta_{N_{i-1}+1}, ..., \theta_{N_i}), ...\theta_{(m)} = (\theta_{N_{m-1}+1}, ..., \theta_k),$$

where $N_j = n_1 + ... + n_j$ for $j = 1, ..., m$. The matrix S is written as

$$S = \begin{bmatrix} A_{11} & A'_{21} & \cdots & A'_{m1} \\ A_{21} & A_{22} & \cdots & A'_{m2} \\ \vdots & & \ddots & \\ A_{m_1} & A_{m2} & \cdots & A_{mm} \end{bmatrix},$$

such that A_{ij} is $n_i \times n_j$. In order to use the algorithm by Berger & Bernardo (1992), it is important to establish how to determine h_j, which is essentially the function needed to derive the reference prior. In order to determine the functions h_j, the inverse of the Fisher information matrix is needed. The functions h_j are defined to be the lower $n_j \times n_j$ corner of H_j. Berger & Bernardo (1992) note that $h_1 = A_{11}^{-1}$ and that if S is a block diagonal matrix then $h_j = A_{jj}^{-1}$. In the case of S being a block diagonal matrix, this simplifies the algorithm to determine the reference prior substantially. In the case where S is not a block diagonal matrix then the functions h_j can still be determined but it is required to partition

H corresponding to the unknown parameters according to the following procedure described in Berger & Bernardo (1992): Define $B_j = \left(A_{j1}...A_{j(j-1)}\right)$ for $j = 2, ..., m$, then

$$h_j = \left(A_{jj} - B_j H_{j-1} B_j'\right)^{-1}.$$

Consider a nested sequence $\Theta^1 \subset \Theta^2 \subset ...$ of compact subsets of Θ such that

$$\bigcup_{l=1}^{\infty} \Theta^l = \Theta.$$

The reference prior for the group ordering $(\theta_1, \theta_2, ..., \theta_m)$ is given by

$$\pi(\boldsymbol{\theta}) = \lim_{l \to \infty} \frac{\pi^l(\boldsymbol{\theta})}{\pi^l(\boldsymbol{\theta}^*)}$$

for some $\boldsymbol{\theta}^* \in \Theta$ and where $\pi^l(\boldsymbol{\theta})$ is determined using an iteration process according to the following procedure:

1. For $j = m$ and $\theta_m \in \Theta_m^l$,

$$\begin{aligned} \pi_m^l \left(\theta_{[m-1]} | \theta^{[m-1]}\right) &= \pi_m^l \left(\theta_m | \theta_1, ..., \theta_{m-1}\right) \\ &= \frac{\left\{h_m(\boldsymbol{\theta})\right\}^{\frac{1}{2}}}{\int_{\Theta_m^l} \left\{h_m(\boldsymbol{\theta})\right\}^{\frac{1}{2}} d\theta_m}. \end{aligned}$$

2. For $j = m-1, m-2, ..., 2$ and $\theta_j \in \Theta_j^l$,

$$\pi_j^l \left(\boldsymbol{\theta}_{[j-1]} | \boldsymbol{\theta}^{[j-1]}\right) = \pi_{j+1}^l \left(\boldsymbol{\theta}_{[j]} | \boldsymbol{\theta}^{[j]}\right) \frac{\exp\left\{E_j^l \left[log \left\{h_j(\boldsymbol{\theta})\right\}^{\frac{1}{2}}\right]\right\}}{\int_{\Theta_j^l} \exp\left\{E_j^l \left[log \left\{h_j(\boldsymbol{\theta})\right\}^{\frac{1}{2}}\right]\right\} d\theta_j},$$

   where

$$E_j^l \left[log \left\{h_j(\boldsymbol{\theta})\right\}^{\frac{1}{2}}\right] = \int_{\Theta_{[j]}^l} log \left\{h_j(\boldsymbol{\theta})\right\}^{\frac{1}{2}} \pi_{j+1}^l \left(\boldsymbol{\theta}_{[j]} | \boldsymbol{\theta}^{[j]}\right) d\boldsymbol{\theta}_{[j]}.$$

3. For $j = 1$, $\boldsymbol{\theta}_{[0]} = \boldsymbol{\theta}$, with $\boldsymbol{\theta}^{[0]}$ vacuous and

$$\pi^l(\boldsymbol{\theta}) = \pi_1^l \left(\boldsymbol{\theta}_{[0]} | \boldsymbol{\theta}^{[0]}\right).$$

The notation used for the reference prior algorithm is complicated. The following explanation aims to simplify the notation. Consider the case where the number of groups to be considered is $m = 3$ and the group ordering is $(\theta_1, \theta_2, \theta_3)$ where this means that θ_1 is the most important parameter and θ_3 is the least important parameter. Therefore, each parameter is in its own group. The elements of the Fisher information matrix have to be put in the order $H(\theta_1, \theta_2, \theta_3)$ so that it corresponds to this selected group ordering. Now for $j = m = 3$ according to Step 1 of the procedure above:

$$\begin{aligned}\pi_3^l\left(\theta_{[2]}|\theta^{[2]}\right) &= \pi_3^l\left(\theta_3|\theta_1,\theta_2\right)\\ &= \frac{\{h_3(\boldsymbol{\theta})\}^{\frac{1}{2}}}{\int_{\Theta_3^l}\{h_3(\boldsymbol{\theta})\}^{\frac{1}{2}}\,d\theta_3}.\end{aligned}$$

For Step 2, consider $j = 2$ then we have

$$\begin{aligned}E_2^l\left[log\left\{h_2(\boldsymbol{\theta})\right\}^{\frac{1}{2}}\right] &= \int_{\Theta_{[2]}^l} log\left\{h_2(\boldsymbol{\theta})\right\}^{\frac{1}{2}}\pi_3^l\left(\boldsymbol{\theta}_{[2]}|\boldsymbol{\theta}^{[2]}\right)d\boldsymbol{\theta}_{[2]}\\ &= \int_{\Theta_3^l} log\left\{h_2(\boldsymbol{\theta})\right\}^{\frac{1}{2}}\pi_3^l\left(\theta_3|\theta_1,\theta_2\right)d\theta_3\end{aligned}$$

where this expectation is used to determine

$$\begin{aligned}\pi_2^l\left(\boldsymbol{\theta}_{[1]}|\boldsymbol{\theta}^{[1]}\right) &= \pi_3^l\left(\boldsymbol{\theta}_{[2]}|\boldsymbol{\theta}^{[2]}\right)\frac{\exp\left\{E_2^l\left[log\left\{h_2(\boldsymbol{\theta})\right\}^{\frac{1}{2}}\right]\right\}}{\int_{\Theta_2^l}\exp\left\{E_2^l\left[log\left\{h_2(\boldsymbol{\theta})\right\}^{\frac{1}{2}}\right]\right\}d\theta_2}\\ \pi_2^l\left(\theta_2,\theta_3|\theta_1\right) &= \pi_3^l\left(\theta_3|\theta_1,\theta_2\right)\frac{\exp\left\{E_2^l\left[log\left\{h_2(\boldsymbol{\theta})\right\}^{\frac{1}{2}}\right]\right\}}{\int_{\Theta_2^l}\exp\left\{E_2^l\left[log\left\{h_2(\boldsymbol{\theta})\right\}^{\frac{1}{2}}\right]\right\}d\theta_2}.\end{aligned}$$

For $j = 1$ we have

$$\begin{aligned}E_1^l\left[log\left\{h_1(\boldsymbol{\theta})\right\}^{\frac{1}{2}}\right] &= \int_{\Theta_{[1]}^l} log\left\{h_1(\boldsymbol{\theta})\right\}^{\frac{1}{2}}\pi_2^l\left(\boldsymbol{\theta}_{[1]}|\boldsymbol{\theta}^{[1]}\right)d\boldsymbol{\theta}_{[1]}\\ &= \int_{\Theta_2^l}\int_{\Theta_3^l} log\left\{h_1(\boldsymbol{\theta})\right\}^{\frac{1}{2}}\pi_2^l\left(\theta_2,\theta_3|\theta_1\right)d\theta_2 d\theta_3\end{aligned}$$

where this expectation is used to determine

$$\begin{aligned}\pi_1^l\left(\boldsymbol{\theta}_{[0]}|\boldsymbol{\theta}^{[0]}\right) &= \pi_2^l\left(\boldsymbol{\theta}_{[1]}|\boldsymbol{\theta}^{[1]}\right)\frac{\exp\left\{E_1^l\left[log\left\{h_1(\boldsymbol{\theta})\right\}^{\frac{1}{2}}\right]\right\}}{\int_{\Theta_1^l}\exp\left\{E_1^l\left[log\left\{h_1(\boldsymbol{\theta})\right\}^{\frac{1}{2}}\right]\right\}d\theta_1}\\ \pi_1^l\left(\theta_1,\theta_2,\theta_3\right) &= \pi_2^l\left(\theta_2,\theta_3|\theta_1\right)\frac{\exp\left\{E_1^l\left[log\left\{h_1(\boldsymbol{\theta})\right\}^{\frac{1}{2}}\right]\right\}}{\int_{\Theta_1^l}\exp\left\{E_1^l\left[log\left\{h_1(\boldsymbol{\theta})\right\}^{\frac{1}{2}}\right]\right\}d\theta_1}.\end{aligned}$$

The reference prior for the group ordering $(\theta_1, \theta_2, \theta_3)$ is given by

$$\pi(\theta_1, \theta_2, \theta_3) = \lim_{l \to \infty} \frac{\pi_1^l(\theta_1, \theta_2, \theta_3)}{\pi_1^l(\theta_1^*, \theta_2^*, \theta_3^*)}.$$

The iterative procedure used to derive the reference prior can get tedious especially when there are 3 or more unknown parameters. For the three-parameter case, it is clear from above that the expected values can involve double integration. For the four-parameter case, it can involve triple integration and so on. If the expected values do not have closed forms, then it might prove difficult to derive the reference prior. Berger & Bernado (1992) suggested to use the reference prior where each parameter is in its own group and they also suggest that the ordering of the parameters should be in terms of inferential importance. Berger & Bernado (1992) also suggested to choosing the compact Θ^l as a collection of nested rectangles in Θ. The reference prior will be denoted by π_R.

2.3.4 The Probability Matching Prior

A probability matching prior, as introduced by Datta & Ghosh (1995), is a prior distribution in objective Bayesian statistics designed so that the posterior probabilities of certain outcomes closely align with their actual long-term frequencies. The posterior intervals of level $1-\beta$ will therefore result in good frequentist confidence intervals at the same level. Datta & Ghosh (1995) derived the differential equation that a prior must satisfy if the posterior probability of a one-sided credibility interval for a parametric function and its frequentist probability agree up to $O\left(n^{-1}\right)$ where n is the sample size. The method by Datta & Ghosh (1995) is as follows. Suppose that $\boldsymbol{\theta} = \begin{bmatrix} \theta_1 & \theta_2 & \dots & \theta_k \end{bmatrix}'$ is a k-dimensional vector of parameters and the parameter of interest is $t(\boldsymbol{\theta})$. The prior density function $\pi(\boldsymbol{\theta})$ for $t(\boldsymbol{\theta})$ which matches the frequentist and posterior probability can be derived in the following way. Let $\nabla_t(\boldsymbol{\theta}) = \begin{bmatrix} \frac{\partial}{\partial \theta_1} t(\boldsymbol{\theta}) & \cdots & \frac{\partial}{\partial \theta_k} t(\boldsymbol{\theta}) \end{bmatrix}'$ and

$$\begin{aligned} \eta(\boldsymbol{\theta}) &= \frac{F^{-1}(\boldsymbol{\theta}) \nabla_t(\boldsymbol{\theta})}{\sqrt{\nabla_t'(\boldsymbol{\theta}) F^{-1}(\boldsymbol{\theta}) \nabla_t(\boldsymbol{\theta})}} \\ &= \begin{bmatrix} \eta_1(\boldsymbol{\theta}) & \cdots & \eta_k(\boldsymbol{\theta}) \end{bmatrix}', \end{aligned}$$

where $F(\boldsymbol{\theta})$ is the Fisher information matrix of and $F^{-1}(\boldsymbol{\theta})$ is its inverse. It is evident that $\eta'(\boldsymbol{\theta}) F(\boldsymbol{\theta}) \eta(\boldsymbol{\theta}) = \mathbf{1}$ for all $\boldsymbol{\theta}$. Datta & Ghosh (1995) proved that the agreement between the posterior probability and the frequentist probability holds if and only if

$$\sum_{i=1}^{k} \frac{\partial}{\partial \theta_i} \left\{ \eta_i(\boldsymbol{\theta}) \pi(\boldsymbol{\theta}) \right\} = 0. \tag{2.9}$$

Therefore $\pi(\boldsymbol{\theta})$ is the probability-matching prior for $t(\boldsymbol{\theta})$, the parameters of interest. Equation 2.9 is a differential equation that needs to be solved to obtain an expression for $\pi(\boldsymbol{\theta})$. This equation usually results in a first-order linear partial differential equation and the method used to solve this differential equation is the method of characteristics. See Dennemeyer (1968) for more details on how to solve partial differential equations. The probability matching prior will be denoted by π_{PMP}.

2.4 Maximum Likelihood Estimation (MLE) of Cronbach's alpha

Frequentist methods used to estimate Cronbach's alpha are well developed and have been considered by many authors. Kristof (1963) first introduced confidence intervals for Cronbach's alpha followed by Feldt (1965). These confidence intervals required the assumption of normality as well as compound symmetry. Based on the same assumptions used by Feldt (1965), Hakstian & Whalen (1976) used a normalizing transformation to derive an approximate confidence interval for Cronbach's alpha using the normal distribution. Feldt et al. (1987) summarized the important aspects of the sampling theory for Cronbach's alpha. Feldt et al. (1987) used numerical illustrations to explain the hypothesis tests and confidence intervals derived previously in the literature for a single coefficient as well as for several alpha coefficients. Hulin et al. (1983) and Kaplan & Saccuzzo (1993) give detailed discussions on how test-retest reliability is assessed, how Cronbach's alpha differs from other methods of estimating reliability and how to increase the reliability of a test. Cronbach's alpha is used as a measure of reliability in for example psychological research (Cortina, 1993). Cortina (1993) presented a discussion of the assumptions and the meaning of coefficient alpha. This discussion is followed by a demonstration of the effects of test length and dimensionality on alpha by determining the statistic for tests with varying number of items, dimensions and average item intercorrelations. As mentioned by Padilla & Zhang (2011) the reasons for Cronbach's alpha's popularity is that it is computationally simple and can be computed for continuous as well as binary data. Duhachek & Iacobucci (2004) mention that the assumption of compound symmetry limits the application of Cronbach's alpha in applied research. Van Zyl et al. (2000) derived the maximum likelihood estimator (MLE) for Cronbach's alpha and its corresponding confidence interval. They also made the assumption of normality but the confidence interval holds even when the covariance matrix is not compound symmetric. The confidence interval is based on asymptotic theory. Van Zyl et al. (2000) also derived a pivotal quantity for Cronbach's alpha which can be used to construct an exact interval for alpha but this pivotal quantity only holds under compound symmetry. Koning & Franses (2003) considered various confidence intervals for Cronbach's alpha. They considered the asymptotic interval derived by Van Zyl et al. (2000) and also the exact interval for Cronbach's alpha using the pivotal quantity derived in Van Zyl et al. (2000). Koning & Franses (2003) mention that the asymptotic interval may yield anti-conservative intervals because the variance of the asymptotic result is unstable since it depends on alpha. It is also mentioned in Koning & Franses (2003) that anti-conservative intervals have nominal coverage less than the required coverage. Duhachek & Iacobucci (2004) conducted a simulation study comparing the performance of the confidence interval for Cronbach's alpha by Feldt (1965) with the confidence interval derived by Van Zyl et al. (2000) under a non-parallel measurement model. The results of their simulation study showed that the MLE outperformed across all their simulation conditions. In applications of psychology, most items are dichotomous or likert-type. For these item types the normality assumption may be unrealistic. For this reason, Yuan et al. (2003) and Maydeu-Olivares et al. (2007) looked at asymptotically distribution-free (ADF) confidence intervals for Cronbach's alpha. Yuan et al. (2003) compared the ADF, MLE and bootstrap confidence intervals for Cronbach's alpha using the Hopkins Symptom Checklist. The study showed that the ADF confidence intervals are between the MLE and the bootstrap methods in terms of their accuracy. Raykov (1998) proposed a non-parametric bootstrapping procedure to approximate the sampling distribution of Cronbach's alpha. The advantage of the bootstrapping procedure is that it does not rely on strict distributional assumptions. Cui & Li (2012) evaluated and compared various parametric and non-parametric methods for constructing confidence intervals for Cronbach's alpha. The coverage and width of different confidence intervals were compared across simulation conditions.

2.5 Bayesian Estimation of Cronbach's alpha

There has also been some work done using Bayesian methods on estimating Cronbach's alpha in the literature. Viana (1995) derived theory for the combined Bayesian estimation of a correlation where separate estimates are available. Viana (1995) derived the posterior density for the intraclass correlation coefficient using a uniform prior. This work proved to be useful when working with Cronbach's alpha since the intra-class correlation is related to Cronbach's alpha via the Spearman-Brown equation. Van Zyl (2001) considered the combined Bayesian estimation of Cronbach's alpha where seperate estimates are available by using the posterior distribution for ρ derived in Viana (1995) and the Spearman-Brown equation to derive a posterior distribution for α using a uniform prior. Van der Merwe & Hugo (2007) considered a balanced two-factor nested random effects model and determined Bayesian tolerance intervals using simulation methods. Although they did not consider Cronbach's alpha in their study, their work involves the study of variance components and Cronbach's alpha is a function of variance components. Padilla & Zhang (2011) set out to develop a Bayesian internal consistency estimate and to evaluate its performances through a simulation study. They considered a conjugate prior to derive the posterior distribution for α. Li & Woodruff (2002) also used a conjugate prior to derive Cronbach's alpha but their approach was based on the normal distribution theory for random and mixed effects analysis of variance (ANOVA). Li & Woodruff (2002) also considered a Markov Chain Monte Carlo (MCMC) method for Bayesian inference for Cronbach's alpha and compared it to the maximum likelihood estimate for α by selecting non-informative priors for the variance components. Payandeh Najafabadi & Najafabadi (2016) estimated Cronbach's alpha using non-informative priors.

They considered a gamma-type prior and estimated Cronbach's alpha using various loss functions. They considered the squared error loss function and the linear exponential (LINEX) loss function and found that under LINEX loss, the Bayes estimator does not overestimate reliability. From the literature, it appears that there is a lack of formal approaches used to derive objective priors to estimate Cronbach's alpha. However, Chung & Dey (1998) considered a Bayesian approach to estimating the intra-class correlation using various objective priors. The priors that were considered by Chung & Dey (1998) were the reference prior and the probability matching prior. Izally et al. (2024) estimated Cronbach's alpha using various objective priors to estimate Cronbach's alpha for the one-way random effects model. They considered the well-known Jeffreys rule prior, a divergence prior, the probability matching prior and various group orderings of the one-at-a-time reference prior. Since the intra-class correlation can be written in terms of Cronbach's alpha using the Spearman-Brown equation, this work was an extension of Chung & Dey (1998) with a focus on deriving objective priors for Cronbach's alpha. Izally et al. (2024) also considered various loss functions to estimate Cronbach's alpha. The loss functions that were considered were the well-known squared error loss, the absolute error loss and the LINEX loss. Similar to the conclusion of Payandeh Najafabadi & Najafabadi (2016), it was found that the LINEX loss provided the best estimates for Cronbach's alpha.

Chapter 3

Cronbach's Alpha for the One-Way Random Effects Model

3.1 Introduction

In this chapter a number of non-informative priors for Cronbach's alpha will be derived. The following non-informative priors will be investigated: the well-known Jeffreys prior (from Jeffreys, 1939), a reference prior (see Berger & Bernardo, 1992 and Berger et al., 2009), the probability matching prior using the method proposed by Datta & Ghosh (1995) and a divergence prior developed by Ghosh et al. (2011). A reference prior for grouping order $\left\{\alpha,\theta,\sigma_1^2\right\}$ which means that α is the most important parameter and σ_1^2 is the least important parameter will be derived using the algorithm in Berger & Bernardo (1992). The marginal posterior for α will be derived using the different priors and a simulation study will be conducted where the frequentist coverage rates, the interval lengths and the standard deviation of the interval lengths will be computed to compare the different priors. A second simulation study will be conducted to evaluate the performance of the Bayes estimators using the well-known squared error loss function, the absolute error loss function and the LINEX loss function which was first introduced by Varian (1975).

Consider a balanced variance components model

$$Y_{ij} = \theta + r_i + \varepsilon_{ij} \quad \text{for } i = 1,...I \text{ and } j = 1,...J, \tag{3.1}$$

where the mean effect θ is unknown. The r_i and ε_{ij} are independent normal variables with zero means and variances σ_2^2 and σ_1^2 respectively. Let $\mathbf{Y} = [Y_{i1}\, Y_{i2} \ldots Y_{iJ}]^{'}$. It can be shown that $Var\left(\mathbf{Y}|\theta,\sigma_1^2,\sigma_2^2\right) = \sigma_1^2\mathbf{I} + \mathbf{1}\mathbf{1}'\sigma_2^2$ where $\mathbf{I}$ is the $J \times J$ identity matrix and $\mathbf{1} = [1\,1\ldots 1]^{'}$ is a $J \times 1$ column vector of ones. Therefore the covariance matrix is given by

$$\Sigma = \begin{bmatrix} \sigma_1^2+\sigma_2^2 & \sigma_2^2 & \sigma_2^2 & \dots & \sigma_2^2 \\ \sigma_2^2 & \sigma_1^2+\sigma_2^2 & \sigma_2^2 & \dots & \sigma_2^2 \\ \sigma_2^2 & \dots & \sigma_1^2+\sigma_2^2 & \cdots & \sigma_2^2 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \sigma_2^2 & \sigma_2^2 & \sigma_2^2 & \cdots & \sigma_1^2+\sigma_2^2 \end{bmatrix}_{J\times J}.$$

Cronbach's alpha is defined by

$$\begin{aligned}\alpha =& \frac{J}{(J-1)}\left\{1-\frac{trace\,(\Sigma)}{\mathbf{1_J'}\Sigma\mathbf{1_J}}\right\}\\ =& \frac{J}{(J-1)}\left\{1-\frac{J\left(\sigma_1^2+\sigma_2^2\right)}{J\left(\sigma_1^2+\sigma_2^2\right)+J\,(J-1)\,\sigma_2^2}\right\}\\ =& \frac{J\sigma_2^2}{\sigma_1^2+J\sigma_2^2}=1-\frac{\sigma_1^2}{\sigma_1^2+J\sigma_2^2}. \end{aligned} \tag{3.2}$$

Cronbach's alpha can also be written in terms of the intra-class correlation coefficient. Let

$$\begin{aligned}\rho =& \frac{Cov\left(Y_{ij},Y_{ij'}\right)}{\sqrt{Var\left(Y_{ij}\right)Var\left(Y_{ij'}\right)}} \qquad \text{for } j\neq j'\\ =& \frac{\sigma_2^2}{\sigma_1^2+\sigma_2^2},\end{aligned}$$

then Cronbach's alpha is given by $\alpha = \frac{J\rho}{1+(J-1)\rho}$ using the Spearman-Brown formula as defined in Li & Woodruff (2002). The covariance matrix can therefore be expressed as

$$\Sigma = \phi^2 \begin{bmatrix} 1 & \rho & \rho & \dots & \rho \\ \rho & 1 & \rho & \dots & \rho \\ \vdots & \vdots & \ddots & & \vdots \\ \rho & \rho & \rho & \cdots & 1 \end{bmatrix},$$

where $\phi^2 = \left(\sigma_1^2+\sigma_2^2\right)$. The covariance matrix is therefore compound symmetric. The final form of Cronbach's alpha is true only when the covariance matrix is compound symmetric. In Palmer & Broemeling (1990), they considered the one-way random effects model for the balanced and unbalanced case. They wrote the intra-class correlation coefficient for the unbalanced case, in terms of the variance components. The relationship between the intra-class correlation coefficient and Cronbach's alpha using the Spearman-Brown formula could be used to derive an expression for Cronbach's alpha for the unbalanced one-way random effects model.

3.2 The Prior and Posterior Distibutions

In this section the prior and posterior distributions will be derived for Cronbach's alpha using the balanced one-way random effects model. The objective priors considered in this section are all derived using the Fisher information matrix and therefore it is important to establish the likelihood function for the one-way random effects model. The likelihood function for the one-way random effects model has been covered extensively by Box & Tiao (1973). They wrote the likelihood function based on sufficient statistics. Since $\bar{Y}_{i.}$ is well known to follow a normal distribution with mean θ and variance $\frac{\sigma_1^2+J\sigma_2^2}{J}$ and the $\bar{Y}_{i.}$ are distributed independently of $Y_{ij}-\bar{Y}_{i.}$. This implies that

$$\sum_{i=1}^{I}\sum_{j=1}^{J}\frac{\left(Y_{ij}-\bar{Y}_{i.}\right)^2}{\sigma_1^2}=\frac{v_1m_1}{\sigma_1^2}\sim\chi^2_{v_1=I(J-1)}.$$

Therefore v_1m_1 is a sufficient statistic for σ_1^2. The likelihood can be obtained using the group means $\bar{Y}_{i.}$ and the residuals $Y_{ij}-\bar{Y}_{i.}$ and can be written as

$$\begin{aligned}\ell(\theta,\sigma_1^2,\sigma_2^2|data)&\propto\left[\prod_{i=1}^{I}\left(\sigma_1^2+J\sigma_2^2\right)^{-1/2}\exp\left\{-\frac{J\left(\bar{Y}_{i.}-\theta\right)^2}{2\left(\sigma_1^2+J\sigma_2^2\right)}\right\}\right]\left(\sigma_1^2\right)^{-v_1/2}\exp\left\{-\frac{v_1m_1}{2}\left(\frac{1}{\sigma_1^2}\right)\right\}\\&=\left(\sigma_1^2+J\sigma_2^2\right)^{-I/2}\exp\left\{-\frac{1}{2}\left[\frac{J\sum_{i=1}^{I}\left(\bar{Y}_{i.}-\theta\right)^2}{\left(\sigma_1^2+J\sigma_2^2\right)}+\frac{v_1m_1}{\sigma_1^2}\right]\right\}\left(\sigma_1^2\right)^{-v_1/2}.\end{aligned}$$

Now since $\sum_{i=1}^{I}\left(\bar{Y}_{i.}-\theta\right)^2=I\left(\bar{Y}_{..}-\theta\right)^2+\sum_i\left(\bar{Y}_{i.}-\bar{Y}_{..}\right)^2$, the likelihood is given by

$$\ell(\theta,\sigma_1^2,\sigma_2^2|data)\propto\left(\sigma_1^2+J\sigma_2^2\right)^{-I/2}\left(\sigma_1^2\right)^{-v_1/2}\exp\left\{-\frac{1}{2}\left[\frac{IJ\left(\bar{Y}_{..}-\theta\right)^2}{\left(\sigma_1^2+J\sigma_2^2\right)}+\frac{J\sum_i\left(\bar{Y}_{i.}-\bar{Y}_{..}\right)^2}{\left(\sigma_1^2+J\sigma_2^2\right)}+\frac{v_1m_1}{\sigma_1^2}\right]\right\}$$

where $\bar{Y}_{i.}=\frac{1}{J}\sum_{j=1}^{J}Y_{ij}$ and $\bar{Y}_{..}=\frac{1}{IJ}\sum_{i=1}^{I}\sum_{j=1}^{J}Y_{ij}$. See Box & Tiao (1973) for more details on the likelihood of the one-way random effects model. From Equation 3.2, Cronbach's alpha can be written as

$$1-\alpha=\frac{\sigma_1^2}{\sigma_1^2+J\sigma_2^2},$$

therefore the likelihood function can be written in terms of Cronbach's alpha. The likelihood function for $(\theta,\sigma_1^2,\alpha)$ is given by

$$\ell(\theta,\sigma_1^2,\alpha|data)\propto(\sigma_1^2)^{-v_1/2}\left(\frac{\sigma_1^2}{1-\alpha}\right)^{-\left(\frac{v_2+1}{2}\right)}\exp\left\{-\frac{1}{2\sigma_1^2}\left[IJ\left(\bar{Y}_{..}-\theta\right)^2(1-\alpha)+v_2m_2(1-\alpha)+v_1m_1\right]\right\}\tag{3.3}$$

where $v_2m_2=J\sum_{i=1}^{I}(\bar{Y}_{i.}-\bar{Y}_{..})^2$, $v_1m_1=\sum_{i=1}^{I}\sum_{j=1}^{J}(Y_{ij}-\bar{Y}_{i.})^2$, $v_1=I(J-1)$ and $v_2=I-1$. v_1m_1 is the within group sums of squares and v_2m_2 is the between groups sums of squares. Also $m_1=$

$\frac{\sum_{i=1}^{I}\sum_{j=1}^{J}(Y_{ij}-\bar{Y}_{i.})^2}{v_1}$ and $m_2 = \frac{J\sum_{i=1}^{I}(\bar{Y}_{i.}-\bar{Y}_{..})^2}{v_2}$, the within group mean square error and the between group mean square error respectively. The Fisher information matrix is given by

$$F(\theta,\sigma_1^2,\alpha) = \begin{bmatrix} \frac{IJ(1-\alpha)}{\sigma_1^2} & 0 & 0 \\ 0 & \frac{(v_1+v_2+1)}{2(\sigma_1^2)^2} & \frac{(v_2+1)}{2(1-\alpha)\sigma_1^2} \\ 0 & \frac{(v_2+1)}{2(1-\alpha)\sigma_1^2} & \frac{(v_2+1)}{2(1-\alpha)^2} \end{bmatrix}. \tag{3.4}$$

The derivation of the Fisher information matrix can be found in Appendix A.1.

The well-known Jeffreys prior is proportional to the square root of the determinant of the Fisher information matrix, and is therefore

$$\pi_J(\theta,\sigma_1^2,\alpha) \propto (1-\alpha)^{-1/2}(\sigma_1^2)^{-3/2}. \tag{3.5}$$

The divergence prior is proportional to the positive fourth root of the determinant of the Fisher information matrix, and is therefore

$$\pi_D(\theta,\sigma_1^2,\alpha) \propto (1-\alpha)^{-1/4}(\sigma_1^2)^{-3/4}. \tag{3.6}$$

The probability matching prior will be derived in the following theorem.

Theorem 3.1. *The probability matching prior for Cronbach's alpha is given by*

$$\pi_{PMP}\left(\theta,\sigma_1^2,\alpha\right) \quad = \quad G\left(log\left(\sigma_1^2\right) - \frac{1}{J}log\left(1-\alpha\right)\right)\left(1-\alpha\right)^{-1}\left(\sigma_1^2\right)^{-1}$$

where G is an arbitrary continuously differentiable function.

Proof. The inverse of the Fisher information matrix is given by

$$F^{-1}(\boldsymbol{\theta}) = F^{-1}\left(\theta,\sigma_1^2,\alpha\right) = \begin{bmatrix} \frac{\sigma_1^2}{IJ(1-\alpha)} & 0 & 0 \\ 0 & \frac{2(\sigma_1^2)^2}{I(J-1)} & \frac{-2(1-\alpha)\left(\sigma_1^2\right)}{I(J-1)} \\ 0 & \frac{-2(1-\alpha)\left(\sigma_1^2\right)}{I(J-1)} & \frac{2J(1-\alpha)^2}{I(J-1)} \end{bmatrix}.$$

We are interested in a probability matching prior for $t(\boldsymbol{\theta}) = \alpha$. Now

$$\begin{aligned} \nabla_t'(\boldsymbol{\theta}) &= \begin{bmatrix} \frac{\partial t(\boldsymbol{\theta})}{\partial \theta} & \frac{\partial t(\boldsymbol{\theta})}{\partial \sigma_1^2} & \frac{\partial t(\boldsymbol{\theta})}{\partial \alpha} \end{bmatrix} \\ &= \begin{bmatrix} 0 & 0 & 1 \end{bmatrix}. \end{aligned}$$

Also,

$$\nabla_t'(\boldsymbol{\theta})F^{-1}(\boldsymbol{\theta}) = \left[\begin{array}{ccc} 0 & \frac{-2(1-\alpha)\left(\sigma_1^2\right)}{I(J-1)} & \frac{2J(1-\alpha)^2}{I(J-1)} \end{array}\right]$$

and

$$\nabla_t'(\boldsymbol{\theta})F^{-1}(\boldsymbol{\theta})\nabla_t(\boldsymbol{\theta}) = \frac{2J(1-\alpha)^2}{I(J-1)}.$$

Define

$$\begin{aligned}\eta'(\boldsymbol{\theta}) &= \frac{\nabla_t'(\boldsymbol{\theta})F^{-1}(\boldsymbol{\theta})}{\sqrt{\nabla_t'(\boldsymbol{\theta})F^{-1}(\boldsymbol{\theta})\nabla_t(\boldsymbol{\theta})}} \\ &= \left[\begin{array}{ccc} \eta_1(\boldsymbol{\theta}) & \eta_2(\boldsymbol{\theta}) & \eta_3(\boldsymbol{\theta}) \end{array}\right] \\ &= \left[\begin{array}{ccc} 0 & \frac{-\sqrt{2}\left(\sigma_1^2\right)}{\sqrt{IJ(J-1)}} & \frac{\sqrt{2J}(1-\alpha)}{\sqrt{I(J-1)}} \end{array}\right].\end{aligned}$$

The prior $\pi(\boldsymbol{\theta})$ is a probability matching prior if and only if the differential equation

$$\sum_{i=1}^{3}\frac{\partial}{\partial\theta_i}\left\{\eta_i(\boldsymbol{\theta})\pi(\boldsymbol{\theta})\right\} = 0$$

$$\frac{\partial}{\partial\alpha}\left\{\frac{\sqrt{2J}(1-\alpha)}{\sqrt{I(J-1)}}\pi\right\} - \frac{\partial}{\partial\sigma_1^2}\left\{\frac{\sqrt{2}\left(\sigma_1^2\right)}{\sqrt{IJ(J-1)}}\pi\right\} = 0 \tag{3.7}$$

is satisfied. To find a probability matching prior satisfying the differential equation, let

$$\pi^*\left(\theta,\sigma_1^2,\alpha\right) = (1-\alpha)\left(\sigma_1^2\right)\pi\left(\theta,\sigma_1^2,\alpha\right).$$

Then Equation 3.7 can be simplified to

$$\frac{\partial}{\partial\alpha}\left\{\frac{\sqrt{2J}}{\sqrt{I(J-1)}}\frac{\pi^*}{\sigma_1^2}\right\} - \frac{\partial}{\partial\sigma_1^2}\left\{\frac{\sqrt{2}}{\sqrt{IJ(J-1)}}\frac{\pi^*}{(1-\alpha)}\right\} = 0. \tag{3.8}$$

Now $\pi^* = 1$ is a solution to Equation 3.8, corresponding to

$$\pi\left(\theta,\sigma_1^2,\alpha\right) = (1-\alpha)^{-1}\left(\sigma_1^2\right)^{-1}.$$

Using the product rule and implicit differentiation it can be shown that Equation 3.8 is a first-order homogeneous and linear partial differential equation given by

$$\frac{\sqrt{2J}}{\sqrt{I(J-1)}}\frac{\pi^*_\alpha}{\sigma_1^2}-\frac{\sqrt{2}}{\sqrt{IJ(J-1)}}\frac{\pi^*_{\sigma_1^2}}{(1-\alpha)}=0 \tag{3.9}$$

where π^*_α and $\pi^*_{\sigma_1^2}$ denote the first derivative of π^* with respect to α and σ_1^2 respectively. The method by Dennemeyer (1968) will be used to solve the partial differential equation. The idea is to use a transformation to reduce the partial differential equation to an ordinary differential equation which is easier to solve. Let $\frac{d\sigma_1^2}{d\alpha}=\frac{-\sigma_1^2}{(1-\alpha)J}$ then re-arranging and integrating both sides we have

$$\begin{aligned}\int\left(\frac{1}{\sigma_1^2}\right)d\sigma_1^2 &= -\int\frac{1}{(1-\alpha)J}d\alpha\\ log\left(\sigma_1^2\right) &= \frac{1}{J}log\left(1-\alpha\right)+c_1 \qquad \text{where } c_1 \text{ is some constant}\\ c_1 &= log\left(\sigma_1^2\right)-\frac{1}{J}log\left(1-\alpha\right).\end{aligned}$$

Now let $\xi\left(\alpha,\sigma_1^2\right)=log\left(\sigma_1^2\right)-\frac{1}{J}log\left(1-\alpha\right)$ and $\eta\left(\alpha,\sigma_1^2\right)=\sigma_1^2$. Therefore

$$\pi^*_\alpha=\pi^*_\xi\xi_\alpha+\pi^*_\eta\eta_\alpha=\frac{1}{J(1-\alpha)}\pi^*_\xi$$

and

$$\pi^*_{\sigma_1^2}=\pi^*_\xi\xi_{\sigma_1^2}+\pi^*_\eta\eta_{\sigma_1^2}=\frac{1}{\sigma_1^2}\pi^*_\xi+\pi^*_\eta.$$

Substituting back into Equation 3.9 results in

$$\frac{-\sqrt{2}}{\sqrt{IJ(J-1)}}\frac{\pi^*_\eta}{(1-\alpha)}=0.$$

The solution is therefore $\pi^*=G(\xi)=G\left(log\left(\sigma_1^2\right)-\frac{1}{J}log\left(1-\alpha\right)\right)$ where G is an arbitrary continuously differentiable function. Finally since

$$\begin{aligned}\pi^*\left(\theta,\sigma_1^2,\alpha\right) &= (1-\alpha)\left(\sigma_1^2\right)\pi\left(\theta,\sigma_1^2,\alpha\right)\\ G\left(log\left(\sigma_1^2\right)-\frac{1}{J}log\left(1-\alpha\right)\right) &= (1-\alpha)\left(\sigma_1^2\right)\pi\left(\theta,\sigma_1^2,\alpha\right)\\ \pi_{PMP}\left(\theta,\sigma_1^2,\alpha\right) &= G\left(log\left(\sigma_1^2\right)-\frac{1}{J}log\left(1-\alpha\right)\right)(1-\alpha)^{-1}\left(\sigma_1^2\right)^{-1}.\end{aligned}$$

□

If we consider the function $G\left(log\left(\sigma_1^2\right)-\frac{1}{J}log\left(1-\alpha\right)\right)=1$ which is a constant function then we have that the probability matching prior is

$$\pi_{PMP}\left(\theta,\sigma_1^2,\alpha\right) \propto (1-\alpha)^{-1}\left(\sigma_1^2\right)^{-1}. \tag{3.10}$$

The following theorem shows that the reference prior for the group ordering $\left\{\alpha,\theta,\sigma_1^2\right\}$ is the same as the probability matching prior in Equation 3.10.

Theorem 3.2. *The reference prior for the group ordering* $\left\{\alpha,\theta,\sigma_1^2\right\}$ *is given by*

$$\pi_R\left(\alpha,\theta,\sigma_1^2\right) \propto (1-\alpha)^{-1}\left(\sigma_1^2\right)^{-1}.$$

Proof. We are interested in the reference prior for the group ordering $\left\{\alpha,\theta,\sigma_1^2\right\}$ which means that α is the most important parameter and σ_1^2 is the least important parameter. In order to derive the reference prior, the inverse of the Fisher information matrix is needed. Let $S(\boldsymbol{\theta}) = H^{-1}(\boldsymbol{\theta})$ where H is the Fisher information matrix. Now

$$S\left(\alpha,\theta,\sigma_1^2\right) = \begin{bmatrix} A_{11} & A_{21}' & \cdots & A_{m1}' \\ A_{21} & A_{22} & \cdots & A_{m2}' \\ \vdots & & \ddots & \\ A_{m_1} & A_{m2} & \cdots & A_{mm} \end{bmatrix} = \begin{bmatrix} \frac{2J(1-\alpha)^2}{I(J-1)} & 0 & \frac{-2(1-\alpha)\left(\sigma_1^2\right)}{I(J-1)} \\ 0 & \frac{\sigma_1^2}{IJ(1-\alpha)} & 0 \\ \frac{-2(1-\alpha)\left(\sigma_1^2\right)}{I(J-1)} & 0 & \frac{2(\sigma_1^2)^2}{I(J-1)} \end{bmatrix}.$$

Define the truncated ranges for the 3 parameters as $\alpha \in [a_l,b_l]$, $\theta \in [c_l,d_l]$ and $\sigma_1^2 \in [e_l,f_l]$ where $c_l \to -\infty$, $b_l \to 1$, $d_l,f_l \to \infty$ and $a_l,e_l \to 0$. Now

$$h_1 \equiv H_1 \equiv A_{11}^{-1} = \frac{I(J-1)}{2J(1-\alpha)^2},$$

and

$$h_2 = \left(A_{22} - B_2 H_1 B_2'\right)^{-1}.$$

Now S_1 is the upper left $N_1 \times N_1$ corner of S. We have $m=3$ groups with $n_1=1$, $n_2=1$ and $n_3=3$ which implies that S_1 is the upper left 1×1 corner of S, that is

$$S_1 = \frac{2J(1-\alpha)^2}{I(J-1)}.$$

Therefore

$$H_1 = S_1^{-1} = \frac{I(J-1)}{2J(1-\alpha)^2},$$

and $B_2 = A_{21} = 0$. Hence

$$\begin{aligned} h_2 &= \left[A_{22} - (0) H_1 (0)\right]^{-1}. \\ &= A_{22}^{-1} \\ &= \frac{IJ(1-\alpha)}{\sigma_1^2} \end{aligned}$$

Also

$$h_3 = \left(A_{33} - B_3 H_2 B_3'\right)^{-1},$$

where S_2 is the upper left $N_2 \times N_2$ corner of S. Now $N_2 = n_1 + n_2 = 2$. Therefore S_2 is the upper left 2×2 corner of S, that is

$$S_2 = \begin{bmatrix} \frac{2J(1-\alpha)^2}{I(J-1)} & 0 \\ 0 & \frac{\sigma_1^2}{IJ(1-\alpha)} \end{bmatrix}.$$

Now

$$H_2 = S_2^{-1} = \begin{bmatrix} \frac{I(J-1)}{2J(1-\alpha)^2} & 0 \\ 0 & \frac{IJ(1-\alpha)}{\sigma_1^2} \end{bmatrix},$$

and

$$B_3 = \begin{bmatrix} A_{31} & A_{32} \end{bmatrix} = \begin{bmatrix} \frac{-2(1-\alpha)\left(\sigma_1^2\right)}{I(J-1)} & 0 \end{bmatrix}.$$

Therefore

$$B_3 H_2 B_3' = \frac{2\left(\sigma_1^2\right)^2}{IJ(J-1)}.$$

So

$$\begin{aligned} h_3 &= \left(\frac{2(\sigma_1^2)^2}{I(J-1)} - \frac{2\left(\sigma_1^2\right)^2}{IJ(J-1)}\right)^{-1} \\ &= \frac{IJ}{2\left(\sigma_1^2\right)^2}. \end{aligned}$$

The three functions

$$h_1 = \frac{I(J-1)}{2J(1-\alpha)^2}, \qquad h_2 = \frac{IJ(1-\alpha)}{\sigma_1^2}, \qquad h_3 = \frac{IJ}{2\left(\sigma_1^2\right)^2}$$

are the functions needed to calculate the prior. During the iterations, first the truncated conditional function of σ_1^2 given α and θ can be computed as

$$\begin{aligned}
\pi_3^l\left(\sigma_1^2|\alpha,\theta\right) =& \frac{\{h_3\left(\boldsymbol{\theta}\right)\}^{\frac{1}{2}}}{\int_{\theta_3^l}\{h_3\left(\boldsymbol{\theta}\right)\}^{\frac{1}{2}}d\theta_3} \\
\propto& \frac{\sigma_1^{-2}}{\int_{e_l}^{f_l}\frac{1}{\sigma_1^2}d\sigma_1^2} \\
=& \frac{\sigma_1^{-2}}{log\left(f_l e_l^{-1}\right)} \qquad \text{for } e_l \leq \sigma_1^2 \leq f_l.
\end{aligned}$$

Now

$$\begin{aligned}
& E\left[log\{h_2\left(\boldsymbol{\theta}\right)\}^{\frac{1}{2}}\right] \\
\propto& \int_{e_l}^{f_l} log\left\{\left[(1-\alpha)\left(\sigma_1^2\right)^{-1}\right]^{\frac{1}{2}}\right\}\sigma_1^{-2}d\sigma_1^2 \\
=& -\frac{1}{2}log\left(f_l e_l^{-1}\right)\left[-log\left(1-\alpha\right)+\frac{1}{2}\frac{\left[log^2\left(f_l\right)-log^2\left(e_l\right)\right]}{log\left(f_l e_l^{-1}\right)}\right] \\
\propto& -log\left(1-\alpha\right)+\frac{1}{2}\frac{\left[log^2\left(f_l\right)-log^2\left(e_l\right)\right]}{log\left(f_l e_l^{-1}\right)} \\
=& -log\left(1-\alpha\right)+K
\end{aligned}$$

where $K=\frac{1}{2}\frac{\left[log^2(f_l)-log^2(e_l)\right]}{log\left(f_l e_l^{-1}\right)}$ is denoted as a constant which only relates to the ranges of the parameters. Therefore

$$\begin{aligned}
\pi_2^l\left(\theta,\sigma_1^2|\alpha\right) =& \frac{\pi_3^l\left(\sigma_1^2|\alpha,\theta\right)exp\left\{E\left[log\{h_2\left(\boldsymbol{\theta}\right)\}^{\frac{1}{2}}\right]\right\}}{\int_{c_l}^{d_l}exp\left\{E\left[log\{h_2\left(\boldsymbol{\theta}\right)\}^{\frac{1}{2}}\right]\right\}d\theta} \\
=& \frac{\pi_3^l\left(\sigma_1^2|\alpha,\theta\right)exp\{-log\left(1-\alpha\right)+K\}}{\int_{c_l}^{d_l}exp\{-log\left(1-\alpha\right)+K\}d\theta} \\
=& \frac{\pi_3^l\left(\sigma_1^2|\alpha,\theta\right)exp\{-log\left(1-\alpha\right)+K\}}{exp\{-log\left(1-\alpha\right)+K\}\left(d_l-c_l\right)} \\
=& \frac{\sigma_1^{-2}}{log\left(f_l e_l^{-1}\right)\left(d_l-c_l\right)}=\sigma_1^{-2}K' \qquad \text{for } e_l \leq \sigma_1^2 \leq f_l \text{ and } c_l \leq \theta \leq d_l
\end{aligned}$$

where $K' = \frac{1}{log\left(f_l e_l^{-1}\right)(d_l - c_l)}$. We now need the function, h_1, to determine $E\left[log\left\{h_1\left(\boldsymbol{\theta}\right)\right\}^{\frac{1}{2}}\right]$

$$\begin{aligned}
&E\left[log\left\{h_1\left(\boldsymbol{\theta}\right)\right\}^{\frac{1}{2}}\right] \\
\propto &\int_{e_l}^{f_l}\int_{c_l}^{d_l} log\left(1-\alpha\right)^{-1}\sigma_1^{-2}K' d\theta d\sigma_1^2 \\
= &\int_{e_l}^{f_l} log\left(1-\alpha\right)^{-1}\sigma_1^{-2}K'\left(d_l - c_l\right) \\
= &K' log\left(1-\alpha\right)^{-1}\left(d_l - c_l\right) log\left(f_l e_l^{-1}\right) \\
= &\frac{\left(d_l - c_l\right) log\left(f_l e_l^{-1}\right)}{log\left(f_l e_l^{-1}\right)\left(d_l - c_l\right)} \times log\left(1-\alpha\right)^{-1} \\
= &log\left(1-\alpha\right)^{-1}.
\end{aligned}$$

Also $\int_{a_l}^{b_l} exp\left\{log\left(1-\alpha\right)^{-1}\right\} d\alpha = log\left\{\frac{(1-a_l)}{(1-b_l)}\right\}$.Therefore

$$\begin{aligned}
\pi_1^l\left(\alpha,\theta,\sigma_1^2\right) = &\frac{\pi_2^l\left(\theta,\sigma_1^2|\alpha\right) exp\left\{\left[log\left(1-\alpha\right)^{-1}\right]\right\}}{log\left\{\frac{(1-a_l)}{(1-b_l)}\right\}} \\
= &\frac{\sigma_1^{-2}\left(1-\alpha\right)^{-1}}{log\left(f_l e_l^{-1}\right)\left(d_l - c_l\right) log\left\{\frac{(1-a_l)}{(1-b_l)}\right\}}.
\end{aligned}$$

Finally

$$\pi_R\left(\alpha,\theta,\sigma_1^2\right) \propto \lim_{l\to\infty}\frac{\pi_1^l\left(\alpha,\theta,\sigma_1^2\right)}{\pi_1^l\left(\alpha_0,\theta_0,\sigma_{10}^2\right)} \propto \sigma_1^{-2}\left(1-\alpha\right)^{-1},$$

where α_0, θ_0 and σ_{10}^2 are the three inner points in the ranges of the parameters. □

The reference prior for the group orderings $\{\alpha,\sigma_1^2,\theta\}$,$\{\sigma_1^2,\theta,\alpha\}$, $\{\sigma_1^2,\alpha,\theta\}$, $\{\theta,\sigma_1^2,\alpha\}$ and $\{\theta,\alpha,\sigma_1^2\}$ all resulted in the same reference prior obtained in Theorem 3.2. The proofs for the reference prior considering the above mentioned group orderings can be found from Appendix A.2 to Appendix A.6.

All the prior distributions considered are proportional to the negative powers of $(1-\alpha)$ and σ_1^2. Therefore a general form can be written as

$$\pi(\theta,\sigma_1^2,\alpha) \propto (1-\alpha)^{-a}(\sigma_1^2)^{-b} \tag{3.11}$$

for various non-negative a and b. Specifically for $a = b = 1$ we have the probability matching prior, which is also the reference prior. For $a = \frac{1}{2}$ and $b = \frac{3}{2}$ we have the Jeffreys prior and for $a = \frac{1}{4}$ and

$b = \frac{3}{4}$ then we have the divergence prior.

The posterior for $(\theta, \sigma_1^2, \alpha)$ using the prior in Equation 3.11 is given by

$$\pi(\theta,\sigma_1^2,\alpha|data) \propto (\sigma_1^2)^{-\frac{1}{2}(v_1+v_2+1+2b)}(1-\alpha)^{\frac{1}{2}(v_2+1-2a)}\exp\left\{\frac{-\left[IJ\left(\bar{Y}_{..}-\theta\right)^2(1-\alpha)+v_2m_2(1-\alpha)+v_1m_1\right]}{2\sigma_1^2}\right\}. \tag{3.12}$$

The parameter of interest is Cronbach's alpha which is the main interest of this chapter. After integrating out θ and σ_1^2, the resulting marginal posterior distribution of α is

$$\pi(\alpha|data) \propto (1-\alpha)^{\frac{1}{2}(v_2-2a)}\left[v_1m_1+v_2m_2(1-\alpha)\right]^{-\frac{1}{2}(v_1+v_2+2b-2)} \qquad \text{for } 0<\alpha\le 1. \tag{3.13}$$

Notice that in Equation 3.13, the marginal posterior distribution of α depends only on v_1m_1 and v_2m_2. Therefore it is only necessary to simulate v_1m_1 and v_2m_2 for inferences on α. In Bayesian statistics, when using objective priors, they are often improper. This is not a problem provided that the resulting posterior is proper. The following theorem proves that the marginal posterior of α is proper.

Theorem 3.3. *The marginal posterior distribution of α for model 3.1 using the prior in Equation 3.11 is proper and is given by*

$$\pi(\alpha|data) = \frac{\left(\frac{v_2m_2}{v_1m_1}\right)^{\frac{v_2}{2}-a+1}}{B_{t^*}\left(\frac{v_1}{2}+a+b-2,\frac{v_2}{2}-a+1\right)}(1-\alpha)^{\frac{1}{2}(v_2-2a)}\left[1+\frac{v_2m_2}{v_1m_1}(1-\alpha)\right]^{-\frac{1}{2}(v_1+v_2+2b-2)} \quad \textit{for } 0<\alpha\le 1, \tag{3.14}$$

where $B_x(a,b)=\int_x^1 t^{a-1}(1-t)^{b-1}dt$ is the upper incomplete beta function and $t^* = \left(1+\frac{v_2m_2}{v_1m_1}\right)^{-1}$.

Proof. For the posterior to be proper, the following should hold true:

$$\int_0^1 K(1-\alpha)^{\frac{1}{2}(v_2-2a)}\left[v_1m_1+v_2m_2(1-\alpha)\right]^{-\frac{1}{2}(v_1+v_2+2b-2)}d\alpha = 1$$

where K is the normalizing constant. Let $t = \left[1+\frac{v_2m_2}{v_1m_1}(1-\alpha)\right]^{-1}$ then by the substitution rule

$$\begin{aligned}
\int_{t^*}^1 K\frac{t^{\frac{v_1}{2}+a+b-2-1}(1-t)^{\frac{v_2}{2}-a+1-1}}{(v_1m_1)^{\frac{v_1}{2}+a+b-2}(v_2m_2)^{\frac{v_2}{2}-a+1}}dt &= 1\\
K\frac{B_{t^*}\left(\frac{v_1}{2}+a+b-2,\frac{v_2}{2}-a+1\right)}{(v_1m_1)^{\frac{v_1}{2}+a+b-2}(v_2m_2)^{\frac{v_2}{2}-a+1}} &= 1\\
K = \frac{(v_1m_1)^{\frac{v_1}{2}+a+b-2}(v_2m_2)^{\frac{v_2}{2}-a+1}}{B_{t^*}\left(\frac{v_1}{2}+a+b-2,\frac{v_2}{2}-a+1\right)}.&
\end{aligned}$$

Therefore the marginal posterior for α is proper and the posterior density is given by Equation 3.14. □

It can be shown that the mean of the marginal posterior of α is

$$E(\alpha|data) = 1 - \left(\frac{v_1 m_1}{v_2 m_2}\right)\left(\frac{B_{t^*}\left(\frac{v_1}{2}+a+b-3, \frac{v_2}{2}+2-a\right)}{B_{t^*}\left(\frac{v_1}{2}+a+b-2, \frac{v_2}{2}-a+1\right)}\right), \tag{3.15}$$

and the variance of the marginal posterior of α is given by

$$Var(\alpha|data) = \left(\frac{v_1 m_1}{v_2 m_2}\right)^2\left(\frac{B_{t^*}\left(\frac{v_1}{2}+a+b-4, \frac{v_2}{2}+3-a\right)}{B_{t^*}\left(\frac{v_1}{2}+a+b-2, \frac{v_2}{2}-a+1\right)}\right) + 2E(\alpha|data) - 1 - \left[E(\alpha|data)\right]^2, \tag{3.16}$$

where $B_x(a,b)$ and t^* is defined in Equation 3.14. The following theorem will be useful for simulating from the marginal posterior for α.

Theorem 3.4. *For the transformation* $Z = \frac{m_2}{m_1}(1-\alpha)$ *the posterior distribution is a truncated F distribution with* $df_1 = v_2 + 2 - 2a$ *numerator and* $df_2 = v_1 + 2b + 2a - 4$ *denominator degrees of freedom over the interval* $0 < Z < \frac{m_2}{m_1}$.

Proof. Let $Z = \frac{m_2}{m_1}(1-\alpha) = g(\alpha)$ then $g^{-1}(Z) = 1 - \frac{m_1}{m_2}Z$ and $\left|\frac{dg^{-1}(Z)}{dZ}\right| = \frac{m_1}{m_2}$. By the method of transformation we have

$$\begin{aligned}
\pi(Z|data) =& \frac{\left(\frac{v_2 m_2}{v_1 m_1}\right)^{\frac{v_2}{2}-a+1}}{B_{t^*}\left(\frac{v_1}{2}+a+b-2, \frac{v_2}{2}-a+1\right)}\left(\frac{m_1}{m_2}Z\right)^{\frac{1}{2}(v_2-2a)}\left[(1+\frac{v_2}{v_1}Z)\right]^{-\frac{1}{2}(v_1+v_2+2b-2)}\frac{m_1}{m_2}I_{(0,\frac{m_2}{m_1})}(Z) \\
=& \frac{(v_1)^{-\left(\frac{v_2}{2}-a+1\right)}(v_2)^{\frac{v_2}{2}-a+1}}{B_{t^*}\left(\frac{v_1}{2}+a+b-2, \frac{v_2}{2}-a+1\right)}Z^{\frac{v_2}{2}-a}\left[(1+\frac{v_2}{v_1}Z)\right]^{-\frac{1}{2}(v_1+v_2+2b-2)} I_{(0,\frac{m_2}{m_1})}(Z) \\
=& \frac{(v_1)^{\frac{1}{2}(v_1+2a+2b-4)}(v_2)^{\frac{1}{2}(v_2-2a+2)}Z^{\frac{v_2+2-2a}{2}-1}}{B_{t^*}\left(\frac{v_1+2b+2a-4}{2}, \frac{v_2+2-2a}{2}\right)\left[(v_1+v_2 Z)\right]^{\frac{1}{2}(v_2+v_1+2b-2)}} I_{(0,\frac{m_2}{m_1})}(Z),
\end{aligned}$$

where $I_{(0,\frac{m_2}{m_1})}(Z) = \begin{cases} 1, & 0 < Z < \frac{m_2}{m_1} \\ 0, & \text{otherwise} \end{cases}$ is the indicator function. The posterior density for Z is a truncated F distribution with $df_1 = v_2 + 2 - 2a$ numerator and $df_2 = v_1 + 2b + 2a - 4$ denominator degrees of freedom over the interval $0 < Z < \frac{m_2}{m_1}$. □

3.3 Simulation Studies

3.3.1 Simulation Study I

A simulation study is done and coverage probabilities are obtained for Cronbach's alpha using the random effects model. The Jeffreys prior, divergence prior and the probability matching prior, which is also the reference prior, will be used. The values considered in the simulation are considered using the following pairs: $I = 6$ and $J = 5$, $I = 20$ and $J = 5$ and $I = 6$ and $J = 20$. The rationale for the values chosen for the simulation is as follows: $I = 6$ and $J = 20$ is an example of an experiment with a small number of groups but with a large number of observations within each group. $I = 20$ and $J = 5$ on the other hand is an example of an experiment with a large number of groups but with a small number of observations within each group. $I = 6$ and $J = 5$ means that the number of groups are small and the number of observations in each group are also small. The variance values considered are $\sigma_1^2 = 1, 4, 6, 10, 15, 25$ and $\sigma_2^2 = 1, 2, 4, 6, 8, 10, 15, 20, 25$. The average length and standard deviation of the intervals are also given. If two priors have the same coverage probabilities then the one with the shortest interval length and smallest standard deviation of the interval length is preferable. The number of simulations is 10000. The average length $(\bar{I})$ and standard deviation (SD_I) of the intervals are calculated by using the following formulas, where I_i is the interval length and num is the number of intervals:

$$\bar{I} = \frac{1}{num}\sum_{i=1}^{num} I_i$$

and

$$SD_I = \sqrt{\frac{1}{num-1}\sum_{i=1}^{num}\left(I_i - \bar{I}\right)^2}.$$

The method used for simulating the coverage probabilities is as follows:

1. Since $\frac{v_1 m_1}{\sigma_1^2} \sim \chi^2_{v_1}$ we have that $v_1 m_1 = \sigma_1^2 \chi^2_{v_1}$. Similarly $\frac{v_2 m_2}{\sigma_1^2 + J\sigma_2^2} \sim \chi^2_{v_2}$, it follows that $v_2 m_2 = \chi^2_{v_2}\left(\sigma_1^2 + J\sigma_2^2\right)$.

2. Simulated $\chi^2_{v_1}$ and $\chi^2_{v_2}$ values are used to determine $v_1 m_1$ and $v_2 m_2$ which is substituted into the posterior for α in Equation 3.14. If the probability matching prior is used, $a = 1$ and $b = 1$, and for the Jeffreys prior, $a = \frac{1}{2}$, $b = \frac{3}{2}$. In the case of the divergence prior, $a = \frac{1}{4}$ and $b = \frac{3}{4}$.

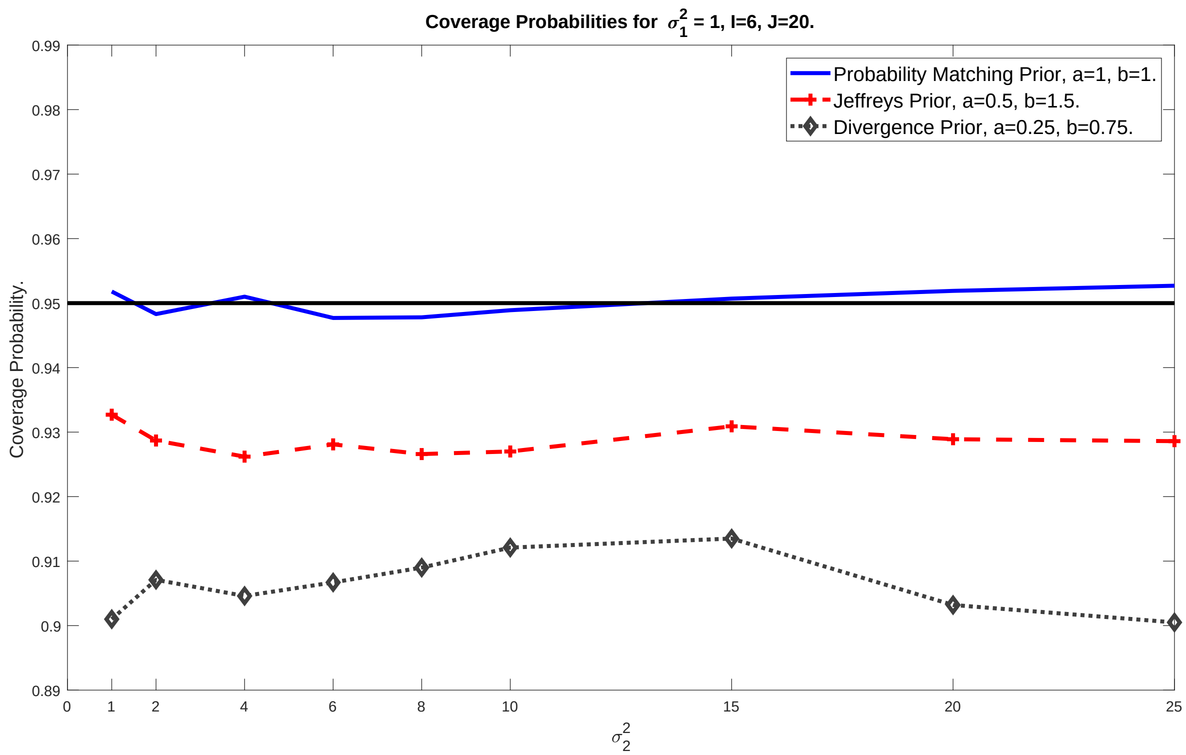
3. Determine for each simulated posterior a 95% credibility interval from the quantiles of the simulated posterior values.

4. All possible values that α can be, have been considered. When $J = 5$, $\sigma_1^2 = 25$ and $\sigma_2^2 = 1$, the smallest value that α can take on is 0.1667 and when $J = 20$, $\sigma_1^2 = 1$ and $\sigma_2^2 = 25$,the largest value that α can take on is 0.9980. On each iteration, the simulation counts whether these α values are in the posterior credibility interval or not. The proportion of times that α is contained within these intervals is the coverage probability.

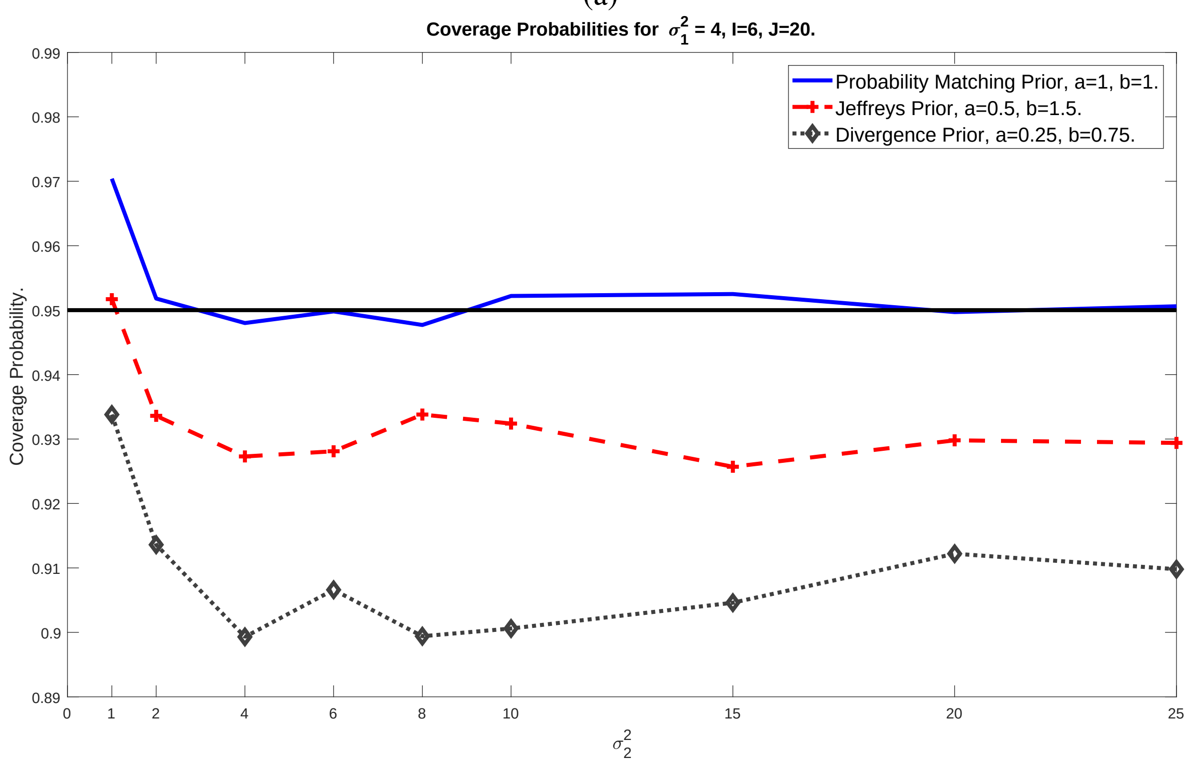
Figures 3.1 to 3.3 summarize the coverage rates obtained using each prior for Cronbach's alpha when $I = 6$, $J = 20$ and $\sigma_1^2 = 1, 4, 6, 10, 15, 25$ over all values of σ_2^2 chosen for this simulation study. Figures 3.4 to 3.6 summarize the coverage rates obtained using each prior for Cronbach's alpha when $I = 20$, $J = 5$ and $\sigma_1^2 = 1, 4, 6, 10, 15, 25$ over all values of σ_2^2 chosen for this simulation study. Figures 3.7 to 3.9 summarize the coverage rates obtained using each prior for Cronbach's alpha when $I = 6$, $J = 5$

and $\sigma_1^2 = 1, 4, 6, 10, 15, 25$ over all values of σ_2^2 chosen for this simulation study. Additional results for the simulation studies are given in Figures A.1 to A.7 in Appendix A.7.

From Figures 3.1 to 3.3 it is clear that the probability matching prior generally does better than the Jeffreys and divergence prior in estimating the coverage probabilities. The Jeffreys and divergence priors tend to underestimate the coverage probabilities over σ_2^2 for various σ_1^2 values. The average coverage probabilities of the probability matching prior, Jeffreys and divergence priors are 0.9558,0.9386 and 0.9194, respectively, while the average interval lengths are 0.2607, 0.2746 and 0.2836. The average standard deviations of the interval lengths are 0.0042, 0.0045 and 0.0046. The probability matching prior has an average coverage rate closest to 0.95 as well as the shortest average interval length and the smallest standard deviation of the interval length compared to the Jeffreys prior and the divergence prior. From Figures 3.4 to 3.6 it is clear that the probability matching prior has some over coverage over the smaller σ_2^2 values but generally outperforms the Jeffreys and divergence prior in estimating the coverage probabilities. The Jeffreys and divergence priors overestimate the coverage for the lower σ_2^2 values over the various σ_1^2 values. The Jeffreys and divergence priors tend to underestimate the coverage probabilities over σ_2^2 in the interval 4 to 25 for the various σ_1^2 values. The average coverage probabilities of the probability matching prior, Jeffreys and divergence priors are 0.9552,0.9522 and 0.9478, respectively, while the average interval lengths are 0.3044, 0.3081 and 0.3146. The average standard deviations of the interval lengths are 0.0052, 0.0052 and 0.0054. The average coverage probabilities for the probability matching prior and the Jeffreys prior are quite similar and the average standard deviation of the interval length is the same. Looking at the average interval length, it is clear that the probability matching prior performed slightly better but the Jeffreys prior also did well in comparison. The performance of the divergence prior was the worst in terms of its average coverage, average standard deviation and average interval length. Figures 3.7 to 3.9 show that there is not much difference between the three priors. The average coverage probabilities of the probability matching prior, Jeffreys and divergence priors are 0.9550,0.9568 and 0.9445, respectively, while the average interval lengths are 0.5261, 0.5259 and 0.5497. The average standard deviations of the interval lengths are 0.0094, 0.0095 and 0.0100.For the special case where $\sigma_1^2 = \sigma_2^2$, this corresponds to the intra-class correlation coefficient of $\rho = \frac{1}{2}$ and Cronbach's alpha is $\alpha = \frac{J}{J+1}$. It is interesting to see that for $\sigma_1^2 = \sigma_2^2$ the coverage rates for the probability matching prior were quite close to the nominal level for various chosen values for the variances except for $\sigma_1^2 = \sigma_2^2 = 1$ and $\sigma_1^2 = \sigma_2^2 = 4$ where there was some over coverage. It therefore seems that in general the probability matching prior is the best of the three priors.

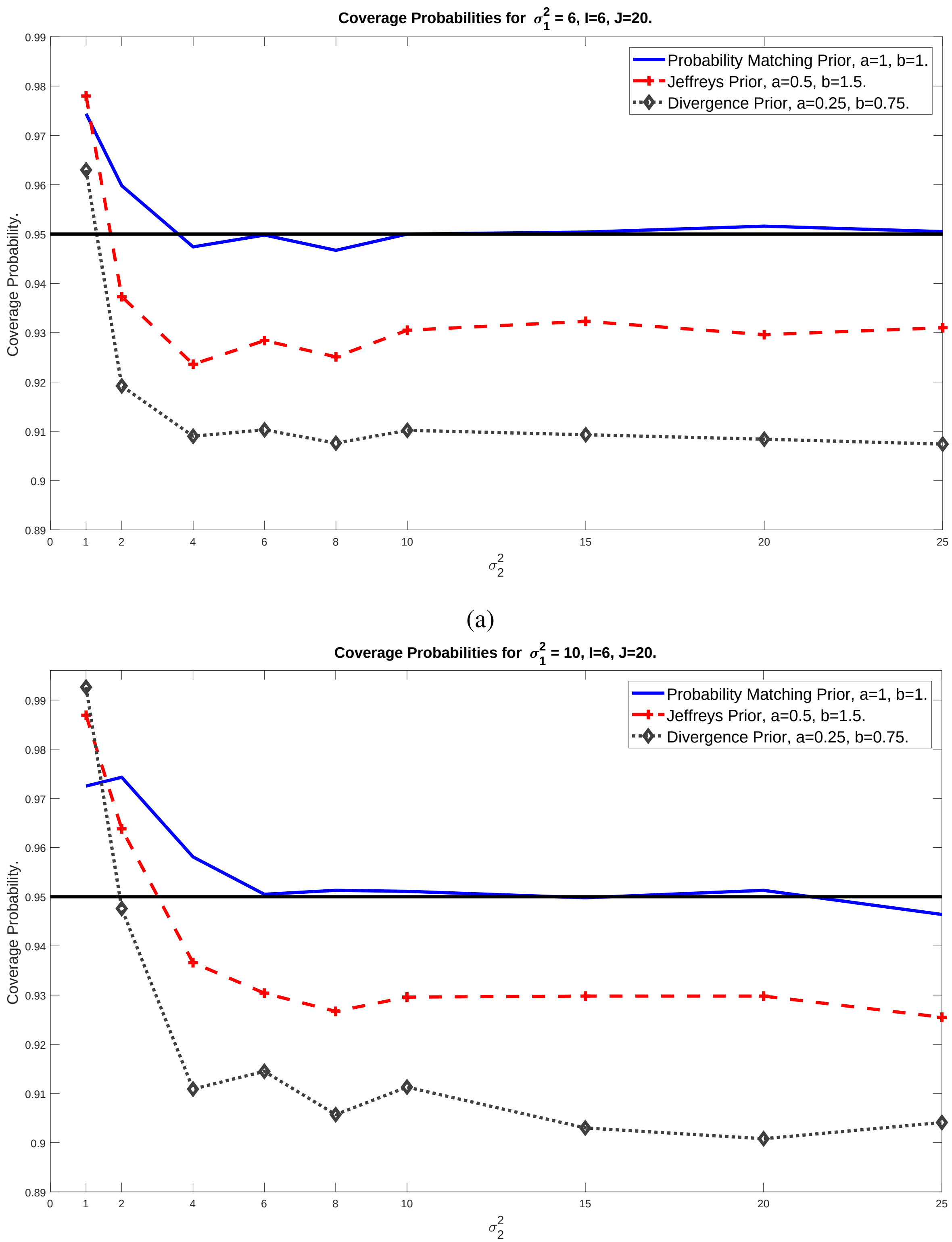


(a)



(b)

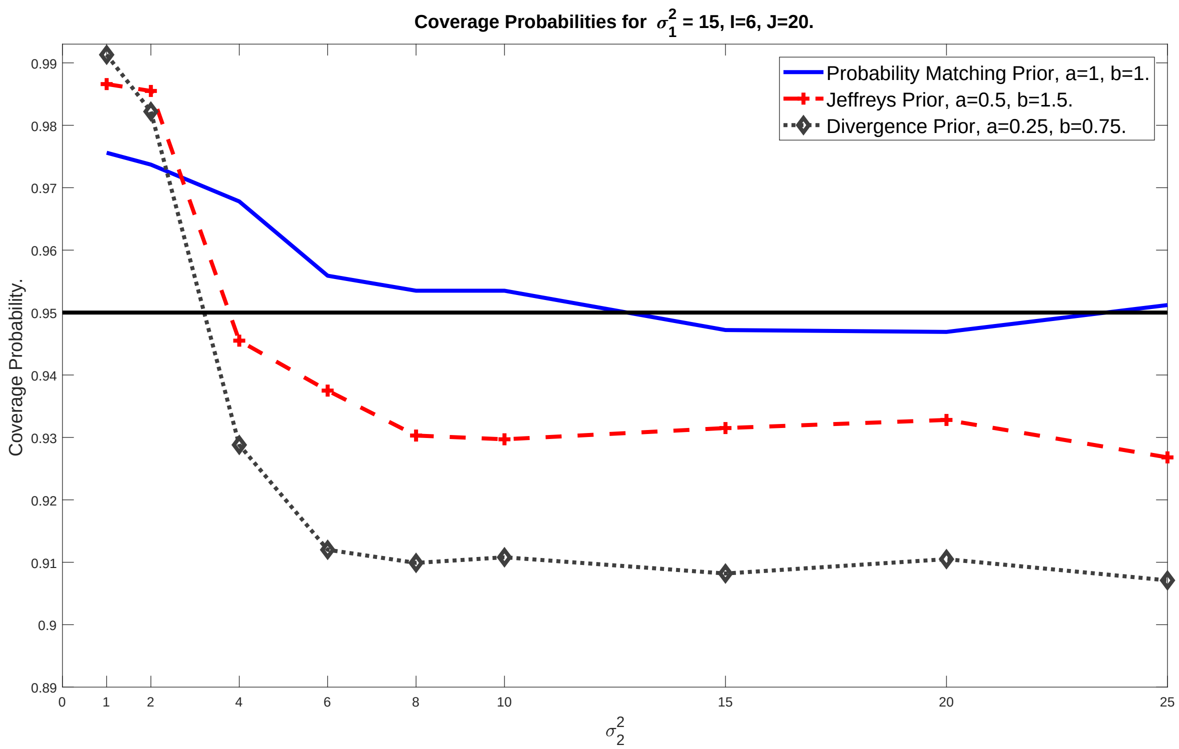
Figure 3.1: Coverage for (a) $\sigma_1^2 = 1$ and (b) $\sigma_1^2 = 4$.



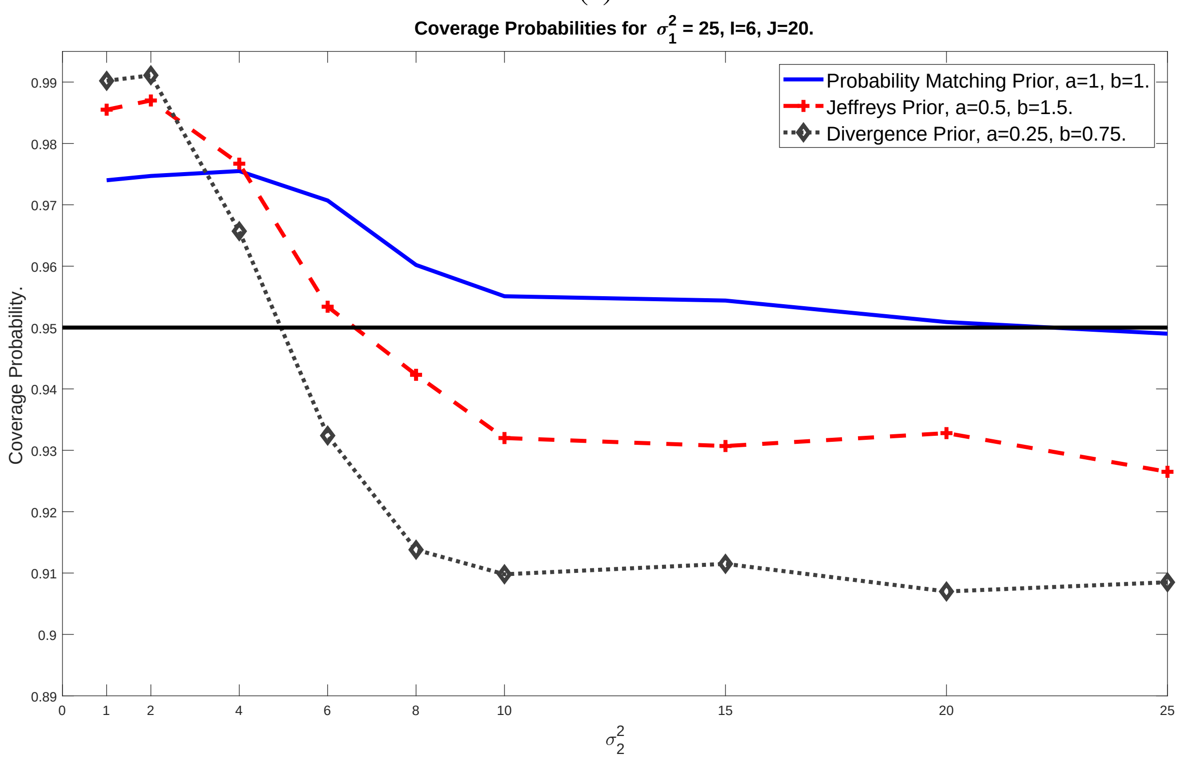
(a)

(b)

Figure 3.2: Coverage for (a) $\sigma_1^2 = 6$ and (b) $\sigma_1^2 = 10$.

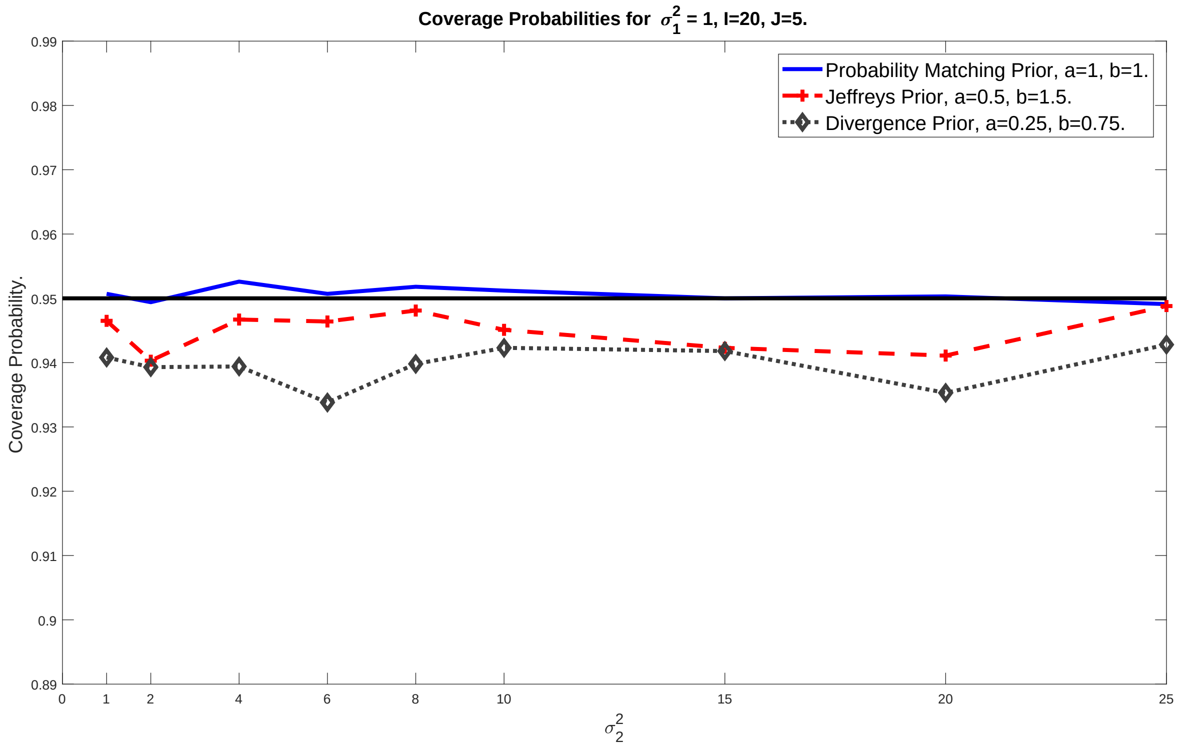


(a)

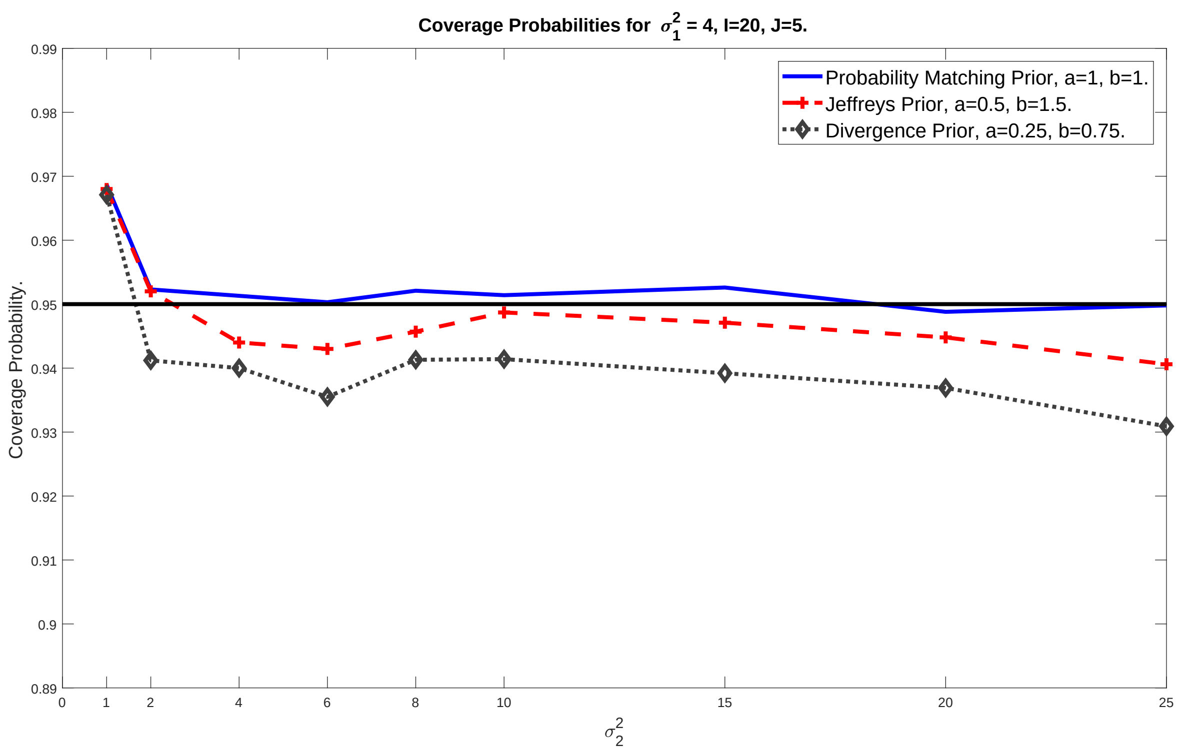


(b)

Figure 3.3: Coverage for (a) $\sigma_1^2 = 15$ and (b) $\sigma_1^2 = 25$.

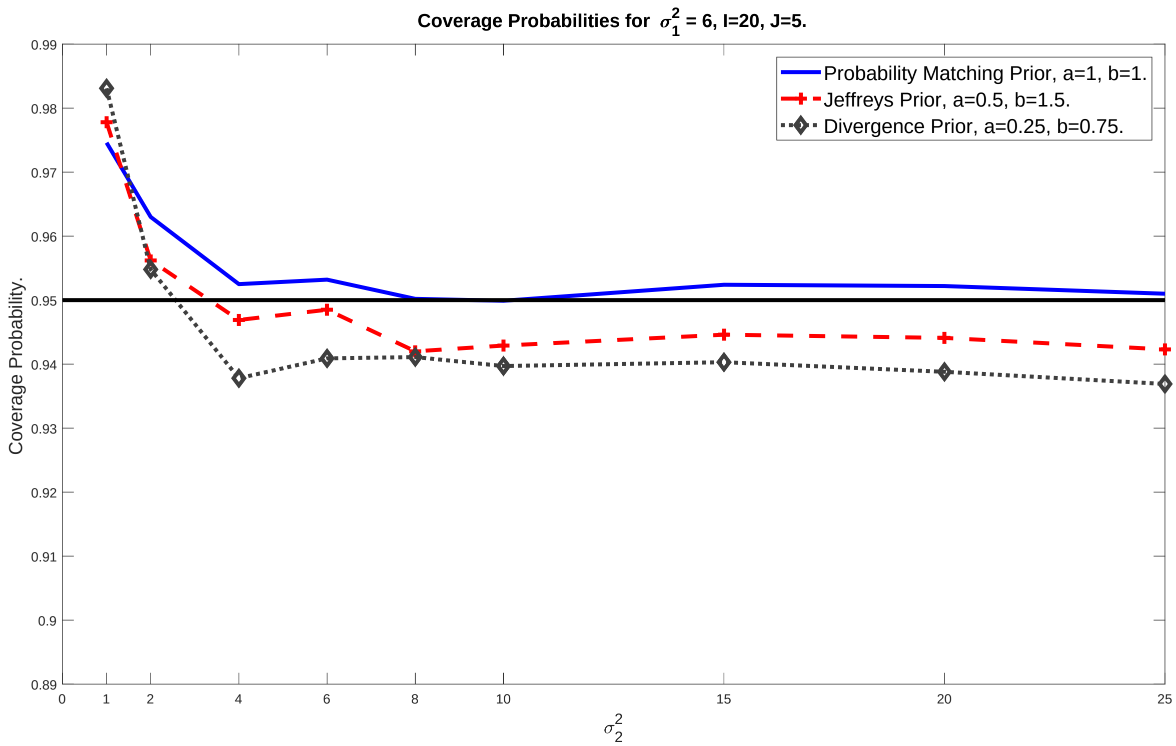


(a)

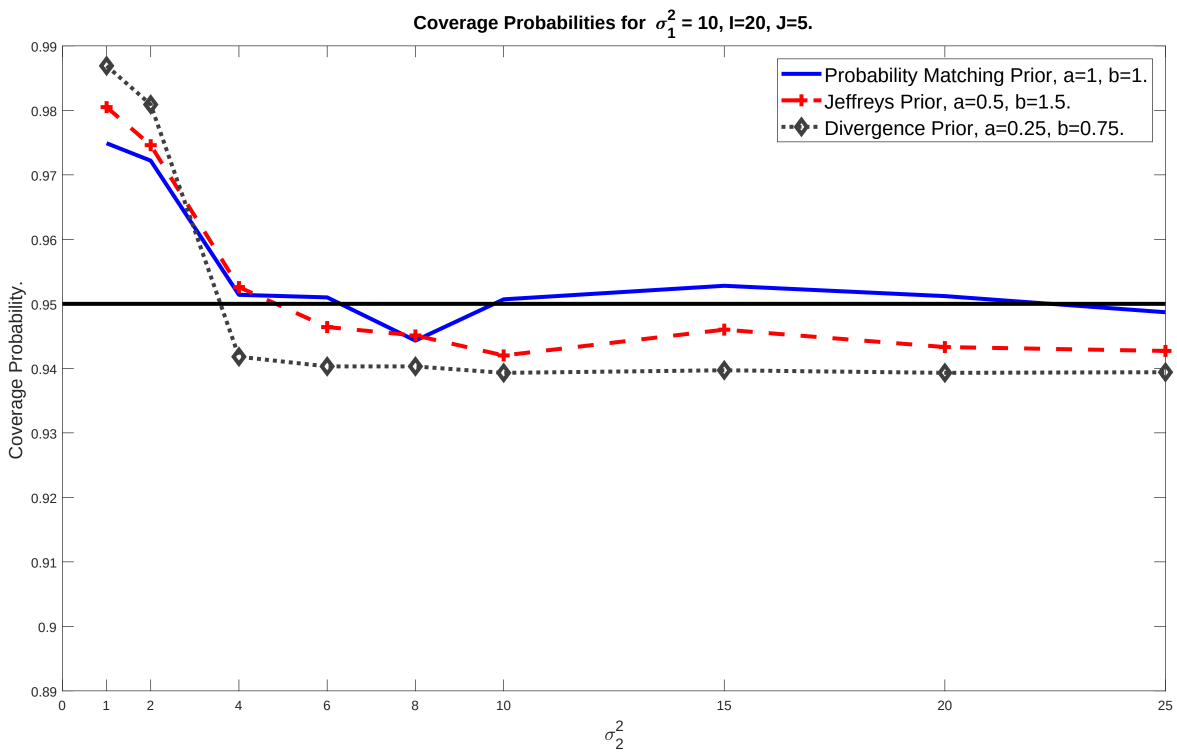


(b)

Figure 3.4: Coverage for (a) $\sigma_1^2 = 1$ and (b) $\sigma_1^2 = 4$.



(a)



(b)

Figure 3.5: Coverage for (a) $\sigma_1^2 = 6$ and (b) $\sigma_1^2 = 10$.

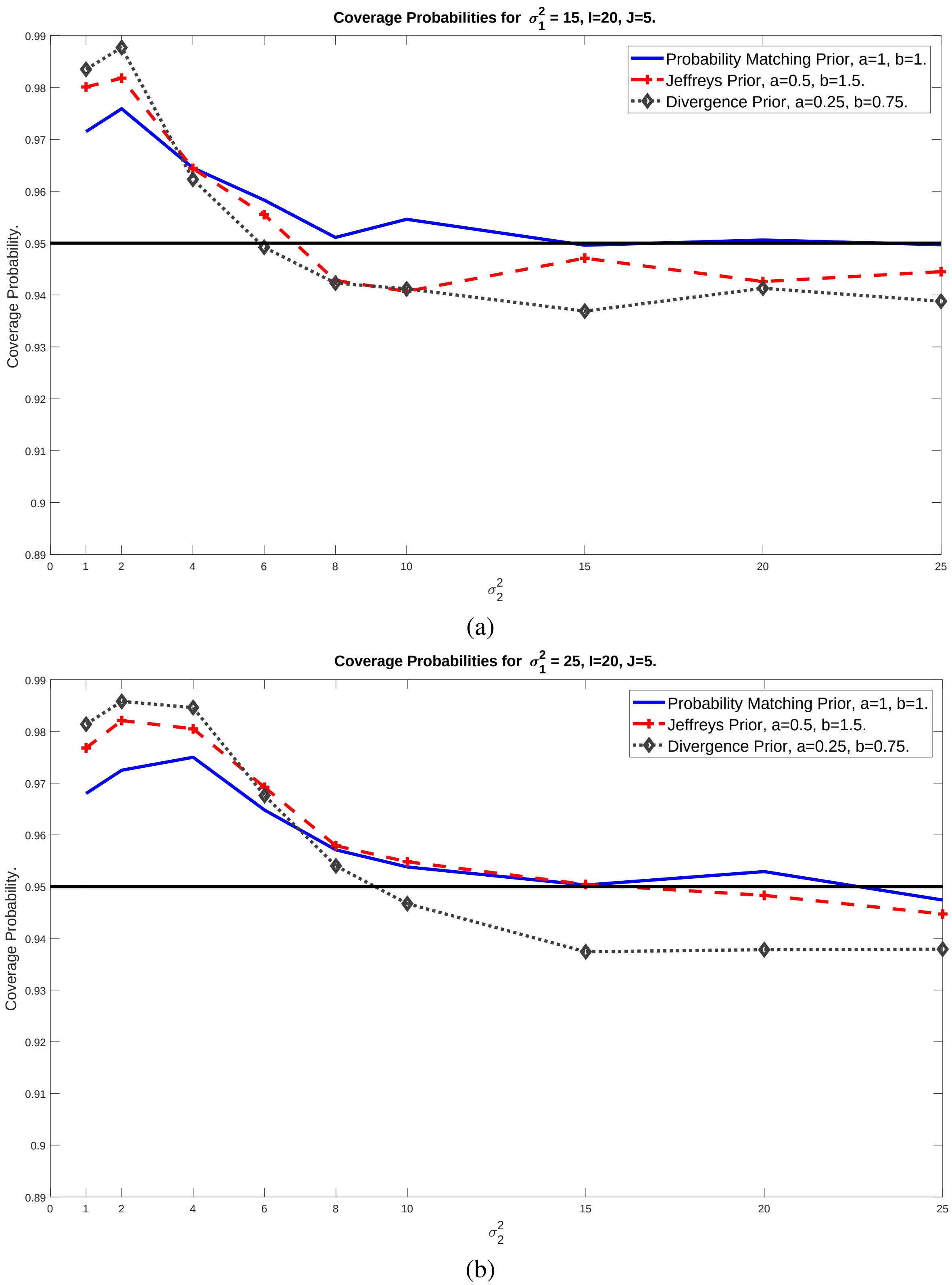
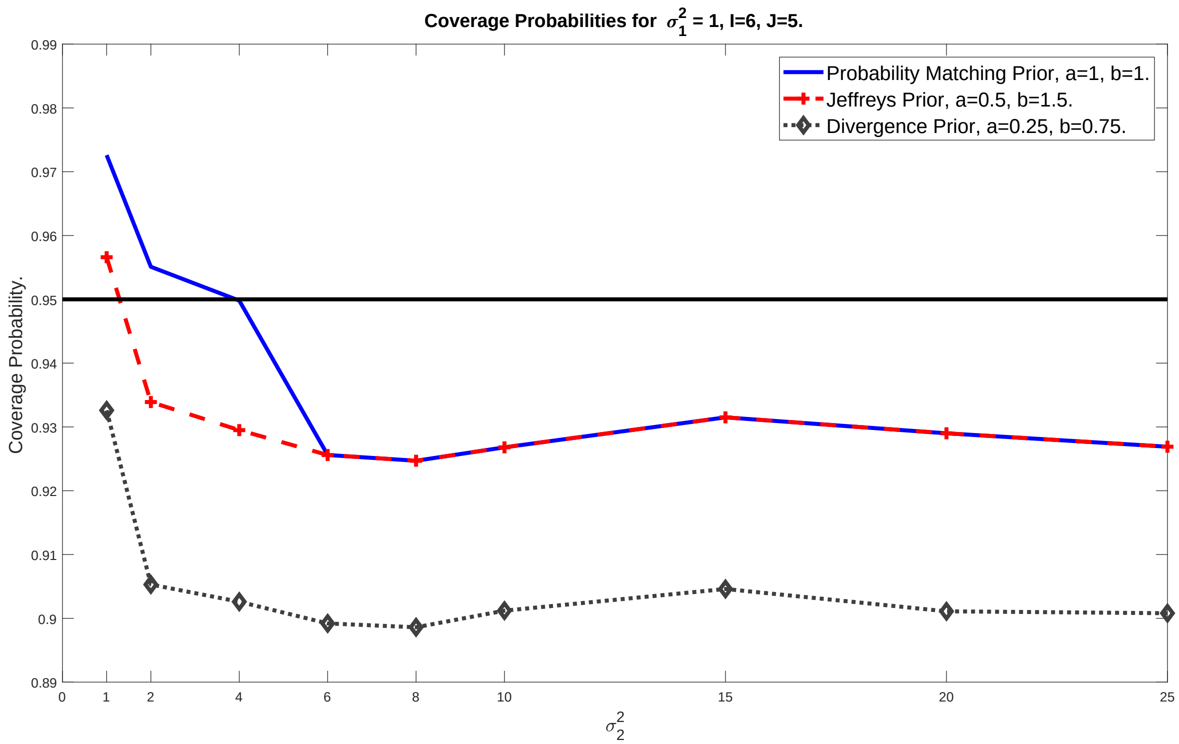
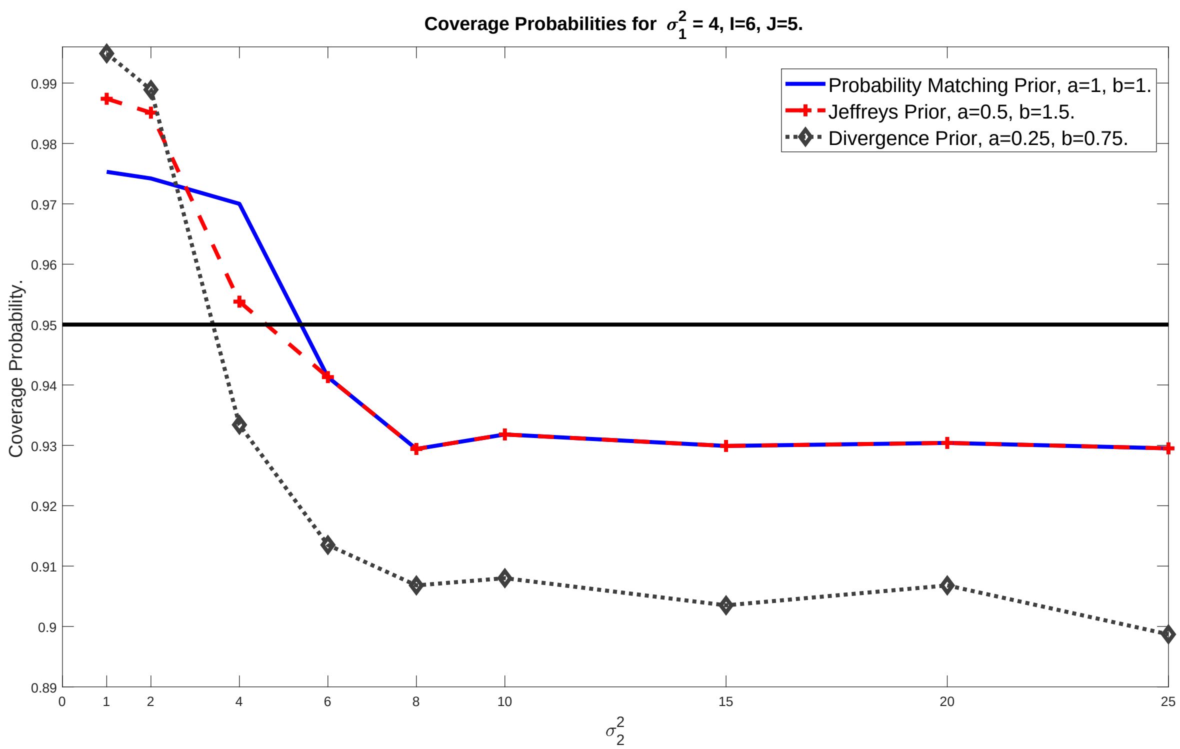


Figure 3.6: Coverage for (a) $\sigma_1^2 = 15$ and (b) $\sigma_1^2 = 25$.

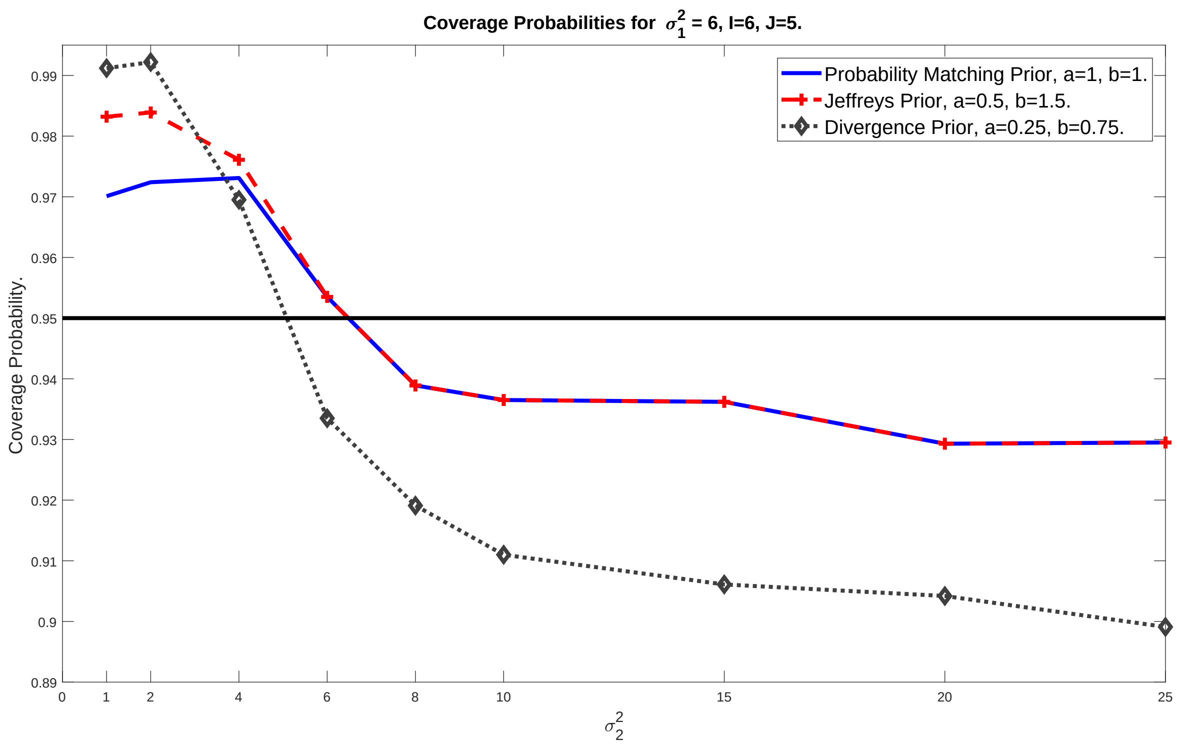


(a)

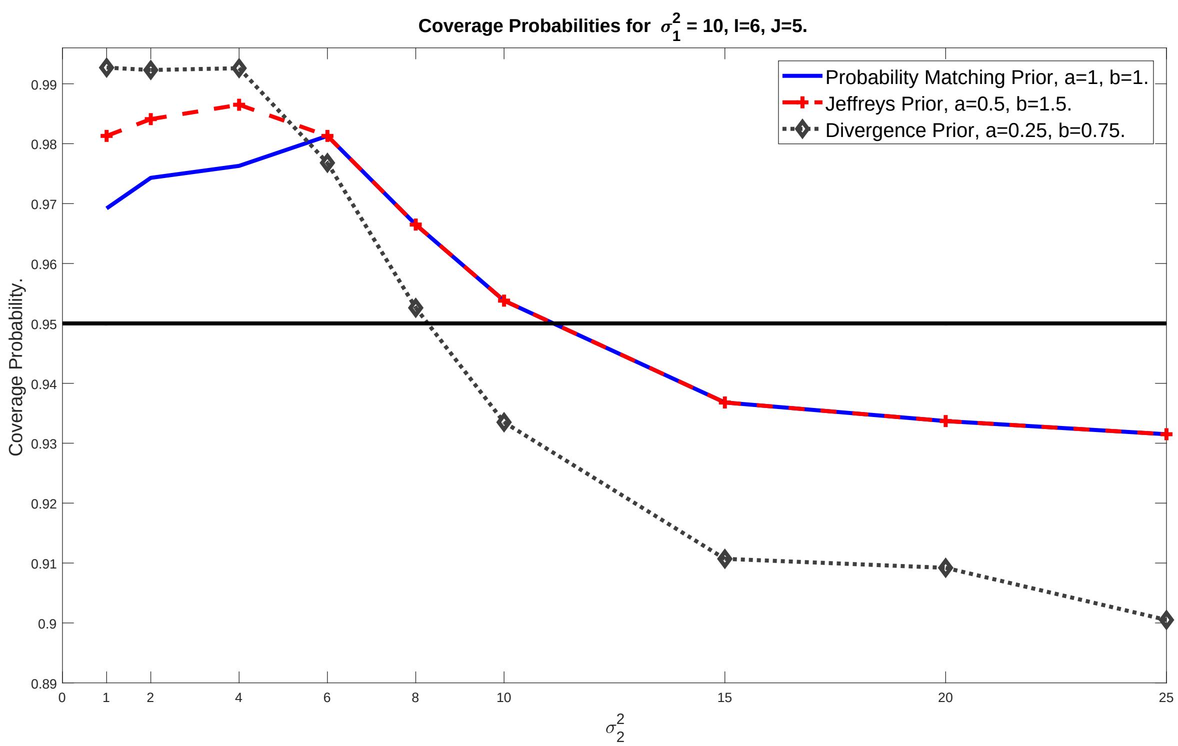


(b)

Figure 3.7: Coverage for (a) $\sigma_1^2 = 1$ and (b) $\sigma_1^2 = 4$.

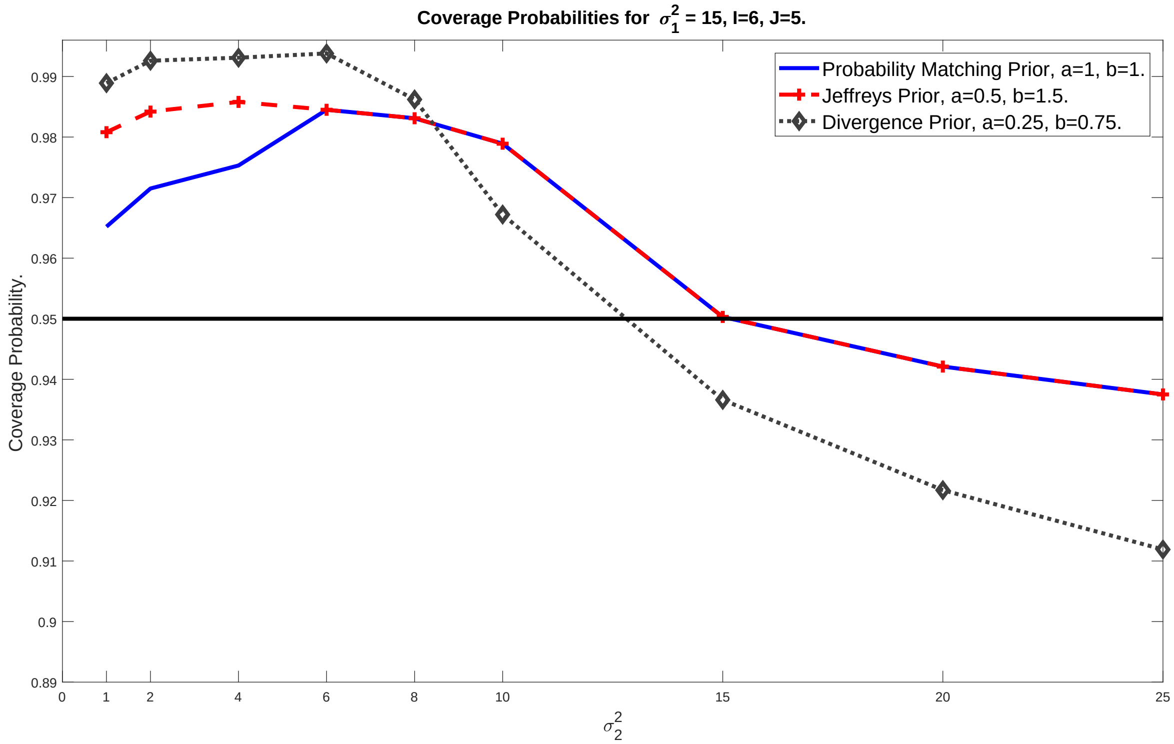


(a)

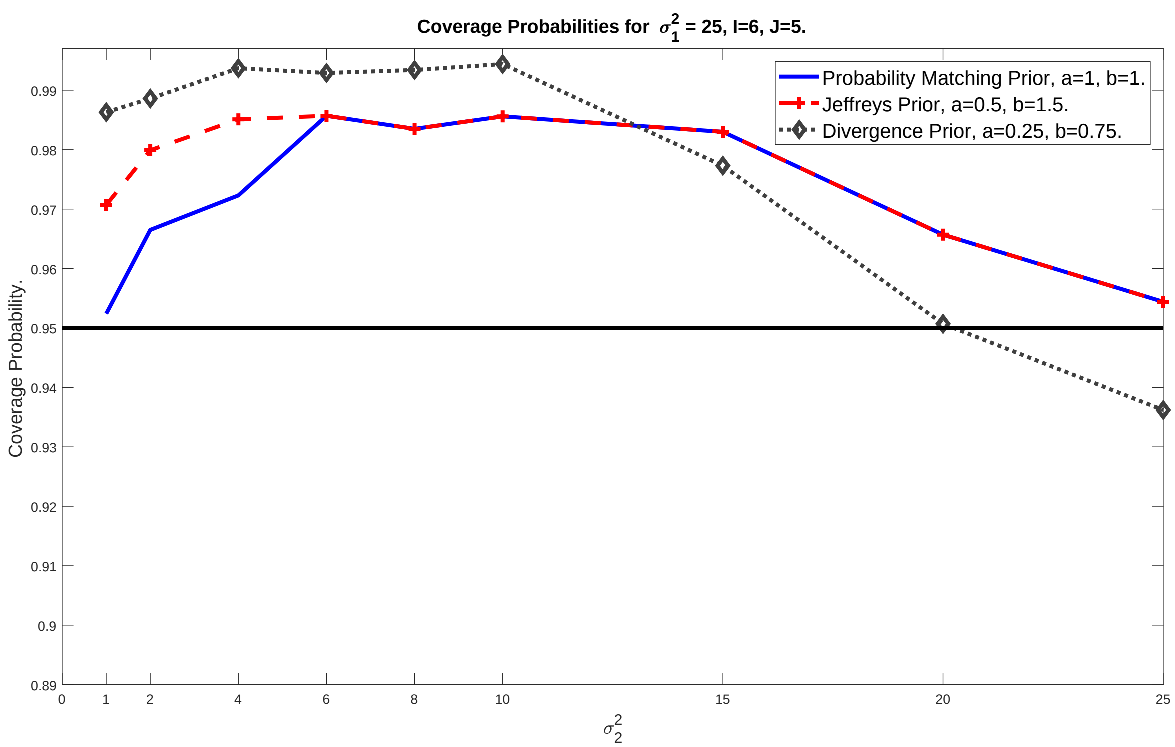


(b)

Figure 3.8: Coverage for (a) $\sigma_1^2 = 6$ and (b) $\sigma_1^2 = 10$.



(a)



(b)

Figure 3.9: Coverage for (a) $\sigma_1^2 = 15$ and (b) $\sigma_1^2 = 25$.

3.3.2 Simulation Study II

In this section, a simulation study will be conducted using various loss functions to compare the performance of the probability matching, Jeffreys and divergence prior. The three loss functions that will be considered is the well-known squared error loss, the absolute error loss and the LINEX loss function first introduced by Varian (1975). For the squared error loss, it is well known that the Bayes estimator is the mean of the posterior and under absolute error loss, the Bayes estimator is the median of the posterior. The LINEX loss function is given by $\ell(\hat{\alpha}-\alpha) = exp\{c[\hat{\alpha}-\alpha]\} - c(\hat{\alpha}-\alpha) - 1$ for $c \neq 0$. The parameter c serves to determine the shape of the loss function. The sign of c represents the direction and its magnitude represents the degree of symmetry. It can be shown that the Bayes estimator of α under LINEX loss is $\hat{\alpha}_{LIN} = -\frac{1}{c} ln[E_{\alpha}(exp\{-c\alpha\})]$ provided that $E_{\alpha}(exp\{-c\alpha\})$ exists and is finite. In this simulation study we have selected the values of α for $\sigma_1^2 = 6$ and $\sigma_2^2 = 2, 3, 6$ and 8 for $I = 5, J = 6$, $I = 5, J = 20$ and $I = 6, J = 6$. The different values considered for c in the LINEX loss function are $c = -9, -5, 0.5, 1$. Additional results for this simulation study are given in Tables A.1 to A.3 in Appendix A.8. The number of replications used was $R = 10000$. The simulation procedure is as follows:

1. Since $\frac{v_1 m_1}{\sigma_1^2} \sim \chi^2_{v_1}$ we have that $v_1 m_1 = \sigma_1^2 \chi^2_{v_1}$. Similarly $\frac{v_2 m_2}{\sigma_1^2 + J\sigma_2^2} \sim \chi^2_{v_2}$, it follows that $v_2 m_2 = \chi^2_{v_2}\left(\sigma_1^2 + J\sigma_2^2\right)$.

2. Simulated $\chi^2_{v_1}$ and $\chi^2_{v_2}$ values determine $v_1 m_1$ and $v_2 m_2$ and these values are used to simulate from the posterior using the fact that $Z = \frac{m_2}{m_1}(1-\alpha)$ follows a truncated F distribution with $df_1 = v_2 + 2 - 2a$ numerator and $df_2 = v_1 + 2b + 2a - 4$ denominator degrees of freedom over the interval $0 < Z < \frac{m_2}{m_1}$, where $m_1 = \frac{v_1 m_1}{v_1}$ and $m_2 = \frac{v_2 m_2}{v_2}$ which are obtained using the simulated values in step 1. If the probability matching prior is used, $a = 1$ and $b = 1$, and for the Jeffreys prior, $a = \frac{1}{2}$, $b = \frac{3}{2}$. In the case of the divergence prior, $a = \frac{1}{4}$ and $b = \frac{3}{4}$. Therefore, the posterior for Cronbach's α is $\alpha = 1 - \frac{Z m_1}{m_2}$.

3. Determine the Bayes estimates for α under each loss function, repeat this $R = 10000$ times.

4. After estimation of the parameters, the Mean Relative Error (MRE), and the estimates were determined to compare the methods of estimation. The Mean Relative Error (MRE) is defined as

$$MRE(\alpha) = \frac{\sum_{i=1}^{R} \frac{\hat{\alpha}_i}{\alpha}}{R}.$$

The results of the simulation study are summarized and tabulated in Tables 3.1, 3.2 and 3.3 for the MRE and the estimates of the three estimators for the chosen I, J, α and c values. The best method is the method that gives the value of MRE closest to one.

Table 3.1: MRE and Estimates Using Different Loss Functions and Values for α When $I = 5$ and $J = 6$.

$\alpha = 0.6667$						
	MATCHING		JEFFREYS		DIVERGENCE	
	MRE	EST	MRE	EST	MRE	EST
Squared Error	0.9561	0.6374	0.9399	0.6266	0.9228	0.6152
Absolute Error	**0.9994**	**0.6662**	0.9760	0.6506	0.9566	0.6378
Linex Loss c: -9	1.1671	0.7780	1.1385	0.7590	1.1208	0.7472
-5	1.0971	0.7314	1.0699	0.7133	1.0515	0.7010
0.5	0.9385	0.6256	0.9241	0.6161	0.9074	0.6049
1	0.9203	0.6135	0.9079	0.6053	0.8916	0.5944
$\alpha = 0.75$						
	MATCHING		JEFFREYS		DIVERGENCE	
	MRE	EST	MRE	EST	MRE	EST
Squared Error	0.9215	0.6911	0.9052	0.6789	0.8989	0.6742
Absolute Error	0.9638	0.7228	0.9408	0.7056	0.9333	0.7000
Linex Loss c: -9	1.0759	0.8070	1.0524	0.7893	1.0431	0.7823
-5	1.0252	0.7689	**1.0021**	**0.7516**	0.9933	0.7449
0.5	0.9081	0.6811	0.8932	0.6699	0.8872	0.6654
1	0.8942	0.6706	0.8807	0.6605	0.8752	0.6564
$\alpha = 0.8571$						
	MATCHING		JEFFREYS		DIVERGENCE	
	MRE	EST	MRE	EST	MRE	EST
Squared Error	0.9155	0.7847	0.9082	0.7785	0.9041	0.7749
Absolute Error	0.9501	0.8144	0.9381	0.8041	0.9330	0.7997
Linex Loss c: -9	**1.0009**	**0.8579**	0.9884	0.8472	0.9830	0.8426
-5	0.9729	0.8339	0.9611	0.8238	0.9559	0.8194
0.5	0.9079	0.7782	0.9014	0.7726	0.8975	0.7692
1	0.8999	0.7713	0.8942	0.7665	0.8905	0.7633
$\alpha = 0.8889$						
	MATCHING		JEFFREYS		DIVERGENCE	
	MRE	EST	MRE	EST	MRE	EST
Squared Error	0.9221	0.8196	0.9141	0.8126	0.9152	0.8135
Absolute Error	0.9524	0.8466	0.9404	0.8359	0.9405	0.8360
Linex Loss c: -9	**0.9873**	**0.8776**	0.9763	0.8678	0.9746	0.8663
-5	0.9658	0.8584	0.9551	0.8490	0.9542	0.8481
0.5	0.9163	0.8144	0.9088	0.8079	0.9102	0.8091
1	0.9101	0.8090	0.9033	0.8029	0.9050	0.8044

From Table 3.1 we see that the MRE values are closer to one under the absolute error loss for the probability matching prior for $\alpha = 0.6667$. The LINEX loss function seems to be doing well when n c is negative for the remainder of the cases in terms of estimating α and having the MRE closer to one. Except for the case where $\alpha = 0.75$, the probability matching prior provided the best estimates of α under each of the various loss functions considered in this simulation study. The divergence prior mostly did not do well in estimating α compared to the probability matching prior and Jeffreys prior.

Table 3.2: MRE and Estimates Using Different Loss Functions and Values for α When $I = 6$ and $J = 6$.

$\alpha = 0.6667$						
	MATCHING		JEFFREYS		DIVERGENCE	
	MRE	EST	MRE	EST	MRE	EST
Squared Error	0.9447	0.6298	0.9282	0.6188	0.9212	0.6141
Absolute Error	**0.9839**	**0.6559**	0.9623	0.6416	0.9540	0.6360
Linex Loss c: -9	1.1412	0.7608	1.1156	0.7438	1.1051	0.7367
-5	1.0740	0.7160	1.0496	0.6997	1.0396	0.6931
0.5	0.9288	0.6192	0.9136	0.6091	0.9070	0.6047
1	0.9124	0.6083	0.8986	0.5991	0.8925	0.5950
$\alpha = 0.75$						
	MATCHING		JEFFREYS		DIVERGENCE	
	MRE	EST	MRE	EST	MRE	EST
Squared Error	0.9176	0.6882	0.9074	0.6806	0.9028	0.6771
Absolute Error	0.9555	0.7166	0.9410	0.7057	0.9353	0.7015
Linex Loss c: -9	1.0587	0.7941	1.0413	0.7810	1.0340	0.7755
-5	1.0109	0.7582	**0.9946**	**0.7460**	0.9879	0.7409
0.5	0.9058	0.6793	0.8966	0.6725	0.8923	0.6693
1	0.8935	0.6702	0.8855	0.6641	0.8815	0.6612
$\alpha = 0.8571$						
	MATCHING		JEFFREYS		DIVERGENCE	
	MRE	EST	MRE	EST	MRE	EST
Squared Error	0.9227	0.7909	0.9190	0.7877	0.9169	0.7859
Absolute Error	0.9529	0.8168	0.9459	0.8108	0.9431	0.8083
Linex Loss c: -9	**0.9957**	**0.8534**	0.9872	0.8462	0.9836	0.8431
-5	0.9709	0.8322	0.9635	0.8259	0.9603	0.8231
0.5	0.9164	0.7855	0.9133	0.7829	0.9114	0.7812
1	0.9098	0.7799	0.9074	0.7778	0.9057	0.7763
$\alpha = 0.8889$						
	MATCHING		JEFFREYS		DIVERGENCE	
	MRE	EST	MRE	EST	MRE	EST
Squared Error	0.9330	0.8293	0.9307	0.8273	0.9292	0.8260
Absolute Error	0.9589	0.8523	0.9538	0.8478	0.9516	0.8459
Linex Loss c: -9	**0.9861**	**0.8765**	0.9801	0.8712	0.9774	0.8688
-5	0.9680	0.8604	0.9628	0.8558	0.9605	0.8538
0.5	0.9284	0.8253	0.9266	0.8237	0.9253	0.8224
1	0.9236	0.8210	0.9223	0.8198	0.9211	0.8187

From Table 3.2 we see the results are quite similar to the results from Table 3.1 where the Bayes estimates under Absolute error loss provide the MRE value closer to one for $\alpha = 0.6667$ and for the rest of the cases the LINEX loss function provides the MRE values closer to one. Except for the case where $\alpha = 0.75$, the probability matching prior provided the best estimates of α under each of the various loss functions considered in this simulation study. Compared to the probability matching prior and Jeffreys prior, the divergence prior did not produce good estimates of α.

Table 3.3: MRE and Estimates Using Different Loss Functions and Values for α When $I = 5$ and $J = 20$.

$\alpha = 0.8696$						
	MATCHING		JEFFREYS		DIVERGENCE	
	MRE	EST	MRE	EST	MRE	EST
Squared Error	0.9220	0.8018	0.9136	0.7945	0.9117	0.7928
Absolute Error	0.9505	0.8265	0.9371	0.8149	0.9333	0.8116
Linex Loss c: -9	**0.9944**	**0.8647**	0.9817	0.8536	0.9770	0.8495
-5	0.9701	0.8435	0.9579	0.8329	0.9538	0.8294
0,5	0.9158	0.7964	0.9081	0.7897	0.9066	0.7883
1	0.9093	0.7907	0.9024	0.7847	0.9012	0.7837
$\alpha = 0.9091$						
	MATCHING		JEFFREYS		DIVERGENCE	
	MRE	EST	MRE	EST	MRE	EST
Squared Error	0.9311	0.8464	0.9294	0.8449	0.9260	0.8418
Absolute Error	0.9538	0.8671	0.9480	0.8618	0.9434	0.8576
Linex Loss c: -9	**0.9800**	**0.8909**	0.9735	0.8850	0.9691	0.8810
-5	0.9633	0.8758	0.9579	0.8708	0.9537	0.8670
0,5	0.9269	0.8426	0.9258	0.8417	0.9226	0.8387
1	0.9226	0.8387	0.9221	0.8383	0.9191	0.8356
$\alpha = 0.9524$						
	MATCHING		JEFFREYS		DIVERGENCE	
	MRE	EST	MRE	EST	MRE	EST
Squared Error	0.9583	0.9127	0.9575	0.9120	0.9558	0.9103
Absolute Error	0.9723	0.9260	0.9690	0.9229	0.9666	0.9206
Linex Loss c: -9	**0.9791**	**0.9325**	0.9762	0.9297	0.9740	0.9276
-5	0.9718	0.9255	0.9694	0.9233	0.9674	0.9213
0,5	0.9566	0.9111	0.9561	0.9106	0.9544	0.9090
1	0.9548	0.9094	0.9546	0.9091	0.9530	0.9076
$\alpha = 0.9639$						
	MATCHING		JEFFREYS		DIVERGENCE	
	MRE	EST	MRE	EST	MRE	EST
Squared Error	0.9662	0.9312	0.9664	0.9315	0.9665	0.9316
Absolute Error	0.9774	0.9421	0.9755	0.9402	0.9750	0.9397
Linex Loss c: -9	**0.9807**	**0.9453**	0.9791	0.9437	0.9783	0.9430
-5	0.9756	0.9403	0.9744	0.9392	0.9740	0.9388
0,5	0.9650	0.9301	0.9654	0.9305	0.9656	0.9307
1	0.9637	0.9289	0.9644	0.9296	0.9647	0.9298

From Table 3.3, the LINEX loss function provides the best Bayes estimates for $c = -9$ using the probability matching prior. For $c = -9$, the Jeffreys and divergence prior provide MRE values close to one and also performs well in estimating α in comparison to the corresponding Bayes estimates under the Squared error and Absolute error loss functions. The divergence prior did not estimate α as well as the Jeffreys prior.

3.4 Example

The following example is from Box & Tiao (1973) which has to do with dyestuff data. The purpose of the experiment was to learn to what extent batch-to-batch variation in a certain raw material was responsible for variation in the final product yield. Five samples from each of six randomly chosen batches of raw material were taken and a single laboratory determination of product yield was made for each of the resulting 30 samples. In this example, $I = 6$ refers to the number of batches and $J = 5$ denotes the number of observations contained within each batch. The data is given in Table 3.4.

Table 3.4: Dyestuff Data.

Batch	1	2	3	4	5	6
Individual Observations	1545	1540	1595	1445	1595	1520
	1440	1555	1550	1440	1630	1455
	1440	1490	1605	1595	1515	1450
	1520	1560	1510	1465	1635	1480
	1580	1495	1560	1545	1625	1445

From the data, we obtained $v_1 m_1 = \sum_{i=1}^{I} \sum_{j=1}^{J} (Y_{ij} - \bar{Y}_{i.})^2 = 58\,830$ and $v_2 m_2 = J \sum_{i=1}^{I} (\bar{Y}_{i.} - \bar{Y}_{..})^2 = 56\,358$. From the results of the simulation studies involving the coverage probability and the Bayes estimators, it is suggested that the probability matching prior be used to analyze this data. The posterior distribution for α using the probability matching prior is given in Figure 3.10. The Bayes estimates under the various loss functions are given in Table 3.5 where the subscripts for $\hat{\alpha}$ in the table, namely, SE, ABS and LIN denote the squared error loss, absolute error loss and LINEX loss, respectively.

Table 3.5: Bayes Estimates of Cronbach's Alpha for the Dyestuff data.

$\hat{\alpha}_{SE}$	$\hat{\alpha}_{ABS}$	$\hat{\alpha}_{LIN,c_1=-9}$	$\hat{\alpha}_{LIN,c_1=-5}$	$\hat{\alpha}_{LIN,c_1=0.5}$	$\hat{\alpha}_{LIN,c_1=1}$
0.7670	0.8053	0.8377	0.8146	0.7603	0.7529

A 95% equal tailed credibility interval is given by $[0.3139; 0.9654]$, means that the probability is 0.95 that Cronbach's alpha will be included in the interval. A 95% highest posterior density (HPD) interval for α is given by $[0.4339; 0.9863]$. The MATLAB code used for this example can be found in Appendix A.11.

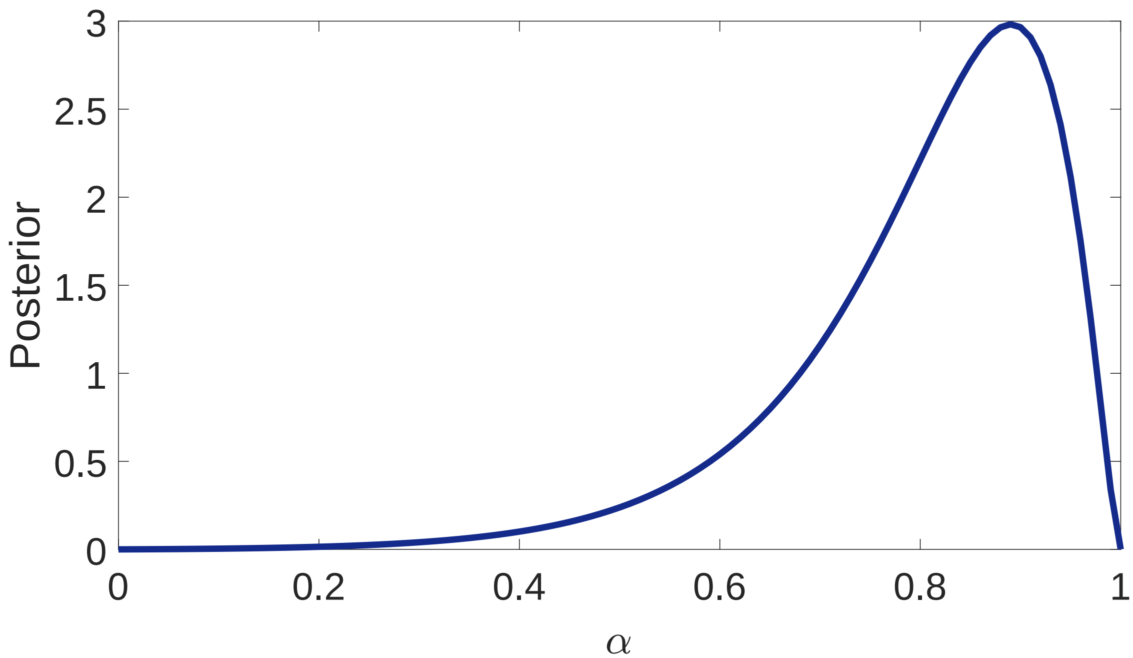


Figure 3.10: Posterior Distribution of α Using the Probability Matching Prior, $a = 1$ and $b = 1$, for the Dyestuff Data.

3.5 Conclusion

In this chapter, a number of objective priors for Cronbach's alpha have been derived. The reference prior for the group ordering $\{\alpha, \theta, \sigma_1^2\}$ was shown to be the same as the probability matching prior when Cronbach's alpha is the parameter of interest. For the group orderings $\{\alpha, \sigma_1^2, \theta\}$, $\{\theta, \sigma_1^2, \alpha\}$, $\{\theta, \alpha, \sigma_1^2\}$, $\{\sigma_1^2, \theta, \alpha\}$ and $\{\sigma_1^2, \alpha, \theta\}$, the result for the reference prior remained the same as the result for the group ordering $\{\alpha, \theta, \sigma_1^2\}$. The probability matching prior, Jeffreys prior and divergence prior were compared using a simulation study involving their coverage probabilities. The simulation study showed that the probability matching prior outperformed the Jeffreys prior and the divergence prior in terms of the coverage rates obtained. The divergence prior performed the worst in terms of its coverage rates which were either larger than the nominal rate or smaller than the nominal rate of 0.95. The second simulation study showed that the LINEX Loss function mostly performed the best with its c parameter equal to -9. It also showed that for $c = -9$ and if the probability matching prior is considered then the best estimates for Cronbach's alpha were obtained and the Mean Relative Error values were close to one. The Jeffreys prior did quite well in some cases but the divergence prior performed the worst in both of the simulation studies. It is recommended that the probability matching prior be used for the Bayesian analysis of Cronbach's alpha and to also use the LINEX loss function with $c = -9$ to obtain the best estimate of Cronbach's alpha.

Chapter 4

Combined Bayesian Estimates for Cronbach's Alpha

4.1 Introduction

In Chapter 3, a balanced random effects model given by

$$Y_{ij} = \theta + r_i + \varepsilon_{ij} \quad \text{for } i = 1, \ldots I \text{ and } j = 1, \ldots J, \tag{4.1}$$

was considered and objective priors for a single Cronbach's coefficient alpha were derived. Van Zyl (2001) considered p normally distributed random variables with equal variances and zero means. The combined Bayesian estimation of more than one Cronbach's alpha with possibly different variances was derived by Van Zyl (2001) using a uniform prior. In this chapter the combined Bayesian estimation of Cronbach's alpha is considered. The priors that will be considered are the reference prior by Berger & Bernardo (1992) and the probability matching prior using the method by Datta & Ghosh (1995). Since our model is the one-way balanced random effects model, the assumption of equicorrelated normal data is satisfied. Our approach is therefore different from the method of Van Zyl (2001) who had to make the assumption of equicorrelation and Van Zyl (2001) only considered a uniform prior. The work is therefore an extension of the work done in Chapter 3 and also an extension of the work done by Van Zyl (2001) .

4.2 The Likelihood and the Fisher Information Matrix

From Chapter 3, it was shown that the likelihood function, in terms of Cronbach's alpha is given by

$$\ell(\alpha,\theta,\sigma_1^2|\text{data}) \propto \left(\sigma_1^2\right)^{-\frac{1}{2}(I+v_1)}(1-\alpha)^{\frac{1}{2}I}\times$$
$$\exp\left\{-\frac{1}{2\sigma_1^2}\left[IJ\left(\bar{Y}_{..}-\theta\right)^2(1-\alpha)+v_2m_2(1-\alpha)+v_1m_1\right]\right\}, \quad (4.2)$$

where $\bar{Y}_{i.}=\frac{1}{J}\sum_{j=1}^{J}Y_{ij}$, $\bar{Y}_{..}=\frac{1}{IJ}\sum_{i=1}^{I}\sum_{j=1}^{J}Y_{ij}$, $v_1m_1=\sum_{i=1}^{I}\sum_{j=1}^{J}(Y_{ij}-\bar{Y}_{i.})^2$, $v_2m_2=J\sum_{i=1}^{I}(\bar{Y}_{i.}-\bar{Y}_{..})^2$, $v_1=I(J-1)$ and $v_2=I-1$. v_1m_1 is the within group sums of squares and v_2m_2 is the between groups sums of squares.

In the case of m experiments with equal α but possibly different variance components, if we denote $\boldsymbol{\theta}=\left(\alpha,\theta^{(1)},...,\theta^{(m)},\sigma_1^{2(1)},...,\sigma_1^{2(m)}\right)$, the likelihood defined in Equation 4.2 can be extended to

$$\ell(\boldsymbol{\theta}|data)\propto(1-\alpha)^{\frac{mI}{2}}\left\{\prod_{l=1}^{m}\left(\sigma_1^{2(l)}\right)^{-\frac{1}{2}(I+v_1)}\right\}\exp\left\{-\frac{1}{2}\sum_{l=1}^{m}\frac{1}{\sigma_1^{2(l)}}\left[v_2m_2^{(l)}(1-\alpha)+v_1m_1^{(l)}\right]\right\} \quad (4.3)$$
$$\times\exp\left\{-\frac{1}{2}\sum_{l=1}^{m}\frac{1}{\sigma_1^{2(l)}}\left[IJ\left(\bar{Y}_{..}^{(l)}-\theta^{(l)}\right)^2(1-\alpha)\right]\right\},$$

where $\bar{Y}_{i.}^{(l)}=\frac{1}{J}\sum_{j=1}^{J}Y_{ij}^{(l)}$, $\bar{Y}_{..}^{(l)}=\frac{1}{IJ}\sum_{i=1}^{I}\sum_{j=1}^{J}Y_{ij}^{(l)}$, $v_1m_1^{(l)}=\sum_{i=1}^{I}\sum_{j=1}^{J}(Y_{ij}^{(l)}-\bar{Y}_{i.}^{(l)})^2$ and $v_2m_2^{(l)}=J\sum_{i=1}^{I}(\bar{Y}_{i.}^{(l)}-\bar{Y}_{..}^{(l)})^2$.

When deriving objective priors, the Fisher information is needed. Therefore the Fisher information matrix will be derived in the following theorem.

Theorem 4.1. *The Fisher information matrix for the ordering* $\{\alpha,\theta^{(1)},...,\theta^{(m)},\sigma_1^{2(1)},...,\sigma_1^{2(m)}\}$ *is given by the following* $(2m+1)\times(2m+1)$ *matrix*

$$F\left(\alpha,\theta^{(1)},...,\theta^{(m)},\sigma_1^{2(1)},...,\sigma_1^{2(m)}\right)=\begin{bmatrix}\tilde{F}_{11} & \tilde{F}_{12}\\ \tilde{F}_{21} & \tilde{F}_{22}\end{bmatrix},$$

where $\tilde{F}_{11}=\frac{mI}{2(1-\alpha)^2}$. *The row vector*

$$\tilde{F}_{12}=\left[0\cdots0\,\frac{1}{2(1-\alpha)^2\sigma_1^{2(1)}}\cdots\frac{1}{2(1-\alpha)^2\sigma_1^{2(m)}}\right]$$

consists of $2m$ *elements and* $\tilde{F}_{12} = \tilde{F}_{21}'$. $\tilde{F}_{22} = \begin{bmatrix} D_{11} & \mathbf{0} \\ \mathbf{0} & D_{22} \end{bmatrix}$ *where*

$$D_{11} = IJ(1-\alpha)\begin{bmatrix} \frac{1}{\sigma_1^{2(1)}} & 0 & \cdots & 0 & \cdots & 0 \\ 0 & \frac{1}{\sigma_1^{2(2)}} & \cdots & 0 & \cdots & 0 \\ 0 & 0 & \frac{1}{\sigma_1^{2(3)}} & 0 & \cdots & 0 \\ \vdots & \vdots & & \ddots & & \vdots \\ 0 & 0 & \cdots & 0 & \cdots & \frac{1}{\sigma_1^{2(m)}} \end{bmatrix}$$

and

$$D_{22} = \frac{IJ}{2}\begin{bmatrix} \frac{1}{\left(\sigma_1^{2(1)}\right)^2} & 0 & \cdots & 0 & \cdots & 0 \\ 0 & \frac{1}{\left(\sigma_1^{2(2)}\right)^2} & \cdots & 0 & \cdots & 0 \\ 0 & 0 & \frac{1}{\left(\sigma_1^{2(3)}\right)^2} & 0 & \cdots & 0 \\ \vdots & \vdots & & \ddots & & \vdots \\ 0 & 0 & \cdots & 0 & \cdots & \frac{1}{\left(\sigma_1^{2(m)}\right)^2} \end{bmatrix}.$$

D_{11} *and* D_{22} *are both* $m \times m$ *diagonal matrices.* $\mathbf{0}$ *is a* $m \times m$ *matrix of zeros.*

Proof. As in the case of the Jeffreys prior, the derivations of the reference and probability matching priors are based on the Fisher information matrix. Differentiation of the log likelihood function twice with respect to the unknown parameters and taking expected values give the Fisher information matrix. From Equation 4.3 the log likelihood function is obtained as

$$\begin{aligned} L &= log\ell\left(\alpha, \theta^{(1)}, ..., \theta^{(m)}, \sigma_1^{2(1)}, ..., \sigma_1^{2(m)}\right) \\ &\propto \frac{mI}{2}log(1-\alpha) - \frac{1}{2}(I+v_1)\sum_{l=1}^{m}\left(\sigma_1^{2(l)}\right) - \frac{1}{2}\sum_{l=1}^{m}\frac{1}{\sigma_1^{2(l)}}\left[IJ\left(\bar{Y}_{..}^{(l)} - \theta^{(l)}\right)^2(1-\alpha) + v_2 m_2^{(l)}(1-\alpha) + v_1 m_1^{(l)}\right] \end{aligned}$$

$$\frac{\partial L}{\partial \sigma_1^{2(l)}} = -\frac{(I+v_1)}{2(\sigma_1^2)^{(l)}} + \frac{1}{2(\sigma_1^{2(l)})^2}\left[IJ\left(\bar{Y}_{..}^{(l)} - \theta^{(l)}\right)^2(1-\alpha) + v_2 m_2^{(l)}(1-\alpha) + v_1 m_1^{(l)}\right]$$

$$\frac{\partial L}{\partial \alpha} = \frac{-mI}{2(1-\alpha)} + \sum_{l=1}^{m}\left(\frac{IJ\left(\bar{Y}_{..}^{(l)} - \theta^{(l)}\right)^2 + v_2 m_2^{(l)}}{2\sigma_1^{2(l)}}\right)$$

$$\frac{\partial L}{\partial \theta^{(l)}} = \frac{1}{\sigma_1^{2(l)}}\left[IJ\left(\bar{Y}_{..}^{(l)} - \theta^{(l)}\right)(1-\alpha)\right]$$

$$\frac{\partial^2 L}{(\partial \alpha)^2} = \frac{-mI}{2(1-\alpha)^2}$$

$$\begin{aligned}
\frac{\partial^2 L}{\left(\partial\sigma_1^{2(l)}\right)^2} &= \frac{(I+v_1)}{2(\sigma_1^{2(l)})^2} - \frac{1}{(\sigma_1^{2(l)})^3}\left[IJ\left(\bar{Y}_{..}^{(l)}-\theta^{(l)}\right)^2(1-\alpha)+v_2 m_2^{(l)}(1-\alpha)+v_1 m_1^{(l)}\right] \\
\frac{\partial^2 L}{(\partial\theta^{(l)})^2} &= -\frac{IJ(1-\alpha)}{\sigma_1^{2(l)}} \\
\frac{\partial^2 L}{\partial\sigma_1^2\partial\theta^{(l)}} &= -\frac{IJ\left(\bar{Y}_{..}^{(l)}-\theta^{(l)}\right)(1-\alpha)}{(\sigma_1^{2(l)})^2} = \frac{\partial^2 L}{\partial\theta^{(l)}\partial\sigma_1^{2(l)}} \\
\frac{\partial^2 L}{\partial\sigma_1^{2(l)}\partial\alpha} &= -\frac{1}{2(\sigma_1^{2(l)})^2}\left[IJ\left(\bar{Y}_{..}^{(l)}-\theta^{(l)}\right)^2+v_2 m_2^{(l)}\right] = \frac{\partial^2 L}{\partial\alpha\partial\sigma_1^{2(l)}} \\
\frac{\partial^2 L}{\partial\alpha\partial\theta^{(l)}} &= \frac{-IJ\left(\bar{Y}_{..}^{(l)}-\theta^{(l)}\right)}{\sigma_1^{2(l)}} = \frac{\partial^2 L}{\partial\theta^{(l)}\partial\alpha}.
\end{aligned}$$

From the above it follows that

$$\begin{aligned}
& -E\left(\frac{\partial^2\ell}{(\partial\alpha)(\partial\sigma_1^{2(l)})}\right) \\
&= \frac{1}{2(\sigma_1^{2(l)})^2}\left[E\left(IJ\left(\bar{Y}_{..}^{(l)}-\theta^{(l)}\right)^2\right)+v_2 E\left(m_2^{(l)}\right)\right] \\
&= \frac{1}{2(\sigma_1^{2(l)})^2}\left[\frac{\sigma_1^{2(l)}}{1-\alpha}+v_2\frac{\sigma_1^{2(l)}}{1-\alpha}\right] = \frac{\sigma_1^{2(l)}}{2(\sigma_1^{2(l)})^2(1-\alpha)}[1+v_2] \\
&= \frac{I}{2(1-\alpha)\sigma_1^{2(l)}},
\end{aligned}$$

$$\begin{aligned}
& -E\left(\frac{\partial^2\ell}{(\partial\sigma_1^{2(l)})^2}\right) \\
&= -\frac{(I+v_1)}{2(\sigma_1^{2(l)})^2}+\frac{1}{(\sigma_1^{2(l)})^3}\left[E\left(IJ\left(\bar{Y}_{..}^{(l)}-\theta^{(l)}\right)^2\right)(1-\alpha)+E\left(v_2 m_2^{(l)}\right)(1-\alpha)+E\left(v_1 m_1^{(l)}\right)\right] \\
&= -\frac{(I+v_1)}{2(\sigma_1^{2(l)})^2}+\frac{1}{(\sigma_1^{2(l)})^3}\left[\frac{\sigma_1^{2(l)}}{1-\alpha}(1-\alpha)+v_2\frac{\sigma_1^{2(l)}}{1-\alpha}(1-\alpha)+v_1\sigma_1^{2(l)}\right] \\
&= \frac{1}{(\sigma_1^{2(l)})^2}\left[1+v_1+v_2-\frac{1}{2}(I+v_1)\right] \\
&= \frac{1}{(\sigma_1^{2(l)})^2}\left[1+IJ-1-\frac{1}{2}(IJ)\right] \quad \text{since } v_1 = IJ-I \text{ and } v_1+v_2 = IJ-1 \\
&= \frac{IJ}{2(\sigma_1^{2(l)})^2},
\end{aligned}$$

where $l = 1, 2, ..., m$. Finally
$-E\left(\frac{\partial^2 \ell}{\left(\partial\theta^{(l)}\right)^2}\right) = \frac{IJ(1-\alpha)}{\sigma_1^{2(l)}}$, $-E\left(\frac{\partial^2 \ell}{(\partial\alpha)^2}\right) = \frac{mI}{2(1-\alpha)^2}$, $-E\left(\frac{\partial^2 \ell}{\theta^{(l)}\theta^{(j)}}\right) = -E\left(\frac{\partial^2 \ell}{\sigma_1^{2(l)}\sigma_2^{2(j)}}\right) = 0$ for $l \neq j$.
The Fisher information matrix follows from these equations. □

4.3 Prior and Posterior Distributions

4.3.1 Probability Matching Priors

The probability matching prior where α is the parameter of interest is derived in the following theorem.

Theorem 4.2. *By making use of the properties of partitioned matrices, the probability matching prior for α is obtained as*

$$\pi_{PMP}\left(\alpha, \theta^{(1)}, ..., \theta^{(m)}, \sigma_1^{2(1)}, ..., \sigma_1^{2(m)}\right) \propto (1-\alpha)^{-1} \prod_{l=1}^{m} \sigma_1^{-2(l)}.$$

Proof. The inverse of the Fisher information matrix is $F^{-1}\left(\alpha, \theta^{(1)}, ..., \theta^{(m)}, \sigma_1^{2(1)}, ..., \sigma_1^{2(m)}\right)$. Denote it by

$$F^{-1}(\boldsymbol{\theta}) = \begin{bmatrix} F^{11} & F^{12} & \cdots & F^{1,2m+1} \\ F^{21} & F^{22} & \cdots & F^{2,2m+1} \\ \vdots & \vdots & & \vdots \\ F^{2m+1,1} & F^{2m+1,2} & \cdots & F^{2m+1,2m+1} \end{bmatrix}.$$

The parameter of interest is $t(\boldsymbol{\theta}) = \alpha$ which means that

$$\begin{aligned} \nabla_t(\boldsymbol{\theta}) &= \begin{bmatrix} \frac{\partial}{\partial\alpha} t(\boldsymbol{\theta}) & \frac{\partial}{\partial\theta^{(1)}} t(\boldsymbol{\theta}) & \cdots & \frac{\partial}{\partial\sigma_1^{2(m)}} t(\boldsymbol{\theta}) \end{bmatrix}' \\ &= \begin{bmatrix} 1 & 0 & \cdots & 0 \end{bmatrix}'. \end{aligned}$$

To get the inverse of the Fisher information matrix, we partition $F(\boldsymbol{\theta})$ as follows: $F(\boldsymbol{\theta}) = \begin{bmatrix} \tilde{F}_{11} & \tilde{F}_{12} \\ \tilde{F}_{21} & \tilde{F}_{22} \end{bmatrix}$ where the partition is defined in Theorem 4.1. The following result from Sahai & Ojeda (2004) states that the inverse of the partitioned Fisher information matrix $F(\boldsymbol{\theta})$ is given by

$$F^{-1}(\boldsymbol{\theta}) = \begin{bmatrix} \tilde{F}^{11} & \tilde{F}^{12} \\ \tilde{F}^{21} & \tilde{F}^{22} \end{bmatrix},$$

where $\tilde{F}^{11} = \left(\tilde{F}_{11} - \tilde{F}_{12}\tilde{F}_{22}^{-1}\tilde{F}_{21}\right)^{-1}$, $\tilde{F}^{12} = -\left(\tilde{F}_{11} - \tilde{F}_{12}\tilde{F}_{22}^{-1}\tilde{F}_{21}\right)^{-1}\tilde{F}_{12}\tilde{F}_{22}^{-1}$, $\tilde{F}^{21} = -\tilde{F}_{22}^{-1}\tilde{F}_{21}\left(\tilde{F}_{11} - \tilde{F}_{12}\tilde{F}_{22}^{-1}\tilde{F}_{21}\right)^{-1}$ and

$$\tilde{F}^{22} = \tilde{F}_{22}^{-1} + \tilde{F}_{22}^{-1}\tilde{F}_{21}(\tilde{F}_{11} - \tilde{F}_{12}\tilde{F}_{22}^{-1}\tilde{F}_{21})^{-1}\tilde{F}_{12}\tilde{F}_{22}^{-1}.$$

Now $\tilde{F}_{11} = \frac{mI}{2(1-\alpha)^2}$ and

$$\begin{aligned}
&\tilde{F}_{12}\tilde{F}_{22}^{-1}\tilde{F}_{21} \\
&= \left[0\cdots 0\,\frac{\sigma_1^{2(1)}}{J(1-\alpha)}\cdots\frac{\sigma_1^{2(m)}}{J(1-\alpha)}\right]\begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \\ \frac{I}{2\sigma_1^{2(1)}(1-\alpha)} \\ \vdots \\ \frac{I}{2\sigma_1^{2(m)}(1-\alpha)} \end{bmatrix} \\
&= \frac{mI}{J(1-\alpha)^2}.
\end{aligned}$$

Therefore

$$\begin{aligned}
\tilde{F}^{11} &= \left(\tilde{F}_{11}-\tilde{F}_{12}\tilde{F}_{22}^{-1}\tilde{F}_{21}\right)^{-1} \\
&= \left(\frac{mI}{2(1-\alpha)^2}-\frac{mI}{J(1-\alpha)^2}\right)^{-1} \\
&= \left(\frac{mI}{(1-\alpha)^2}\left[\frac{J-2}{2J}\right]\right)^{-1} \\
&= \frac{2J(1-\alpha)^2}{mI(J-2)}.
\end{aligned}$$

Now

$$\begin{aligned}
\tilde{F}^{12} &= -\left(\tilde{F}_{11}-\tilde{F}_{12}\tilde{F}_{22}^{-1}\tilde{F}_{21}\right)^{-1}\tilde{F}_{12}\tilde{F}_{22}^{-1} \\
&= -\frac{2J(1-\alpha)^2}{mI(J-2)}\times\left[0\cdots 0\,\frac{\sigma_1^{2(1)}}{J(1-\alpha)}\cdots\frac{\sigma_1^{2(m)}}{J(1-\alpha)}\right] \\
&= \left[0\cdots 0\,\frac{-2(1-\alpha)\sigma_1^{2(1)}}{mI(J-2)}\cdots\frac{-2(1-\alpha)\sigma_1^{2(m)}}{mI(J-2)}\right].
\end{aligned}$$

Now

$$\nabla_t'(\boldsymbol{\theta})F^{-1}(\boldsymbol{\theta}) = \left[\begin{array}{cccc} F^{11} & F^{12} & \cdots & F^{1,2m+1}\end{array}\right] = \left[\tilde{F}^{11}\tilde{F}^{12}\right],$$

and $\nabla_t'(\boldsymbol{\theta})F^{-1}(\boldsymbol{\theta})\nabla_t(\boldsymbol{\theta}) = F^{11} = \tilde{F}^{11}$. Therefore there is no need to determine the complete inverse of the Fisher information matrix to be able to derive the probability matching prior in this case.

Further $\sqrt{\nabla_t'(\boldsymbol{\theta})F^{-1}(\boldsymbol{\theta})\nabla_t(\boldsymbol{\theta})} = \sqrt{\tilde{F}^{11}} = (1-\alpha)\sqrt{\frac{2J}{mI(J-2)}}$ and

$$\begin{aligned}
\eta'(\boldsymbol{\theta}) &= \frac{\nabla_t'(\boldsymbol{\theta})F^{-1}(\boldsymbol{\theta})}{\sqrt{\nabla_t'(\boldsymbol{\theta})F^{-1}(\boldsymbol{\theta})\nabla_t(\boldsymbol{\theta})}} \\
&= \begin{bmatrix} \eta_1(\boldsymbol{\theta}) & \eta_2(\boldsymbol{\theta}) & \cdots & \eta_{m+1}(\boldsymbol{\theta})\,\eta_{m+2}(\boldsymbol{\theta})\cdots & \eta_{2m+1}(\boldsymbol{\theta}) \end{bmatrix} \\
&= \begin{bmatrix} \sqrt{F^{11}} & 0 & \cdots & 0 & \frac{F^{1,m+2}}{\sqrt{F^{11}}} & \cdots & \frac{F^{1,2m+1}}{\sqrt{F^{11}}} \end{bmatrix} \\
&= \frac{1}{\sqrt{\tilde{F}^{11}}}\left[\tilde{F}^{11}\tilde{F}^{12}\right] \\
&= \left[\sqrt{\tilde{F}^{11}}\,\frac{\tilde{F}^{12}}{\sqrt{\tilde{F}^{11}}}\right].
\end{aligned}$$

Now

$$\begin{aligned}
\left[\eta_2(\boldsymbol{\theta})\cdots\eta_{2m+1}(\boldsymbol{\theta})\right] &= \frac{\tilde{F}^{12}}{\sqrt{\tilde{F}^{11}}} \\
&= \frac{1}{(1-\alpha)\sqrt{\frac{2J}{mI(J-2)}}} \times \left[0\cdots 0\,\frac{-2(1-\alpha)\sigma_1^{2(1)}}{mI(J-2)}\cdots\frac{-2(1-\alpha)\sigma_1^{2(m)}}{mI(J-2)}\right] \\
&= \left[0\cdots 0\,\frac{-\sqrt{2}\sigma_1^{2(1)}}{\sqrt{mI(J-2)J}}\cdots\frac{-\sqrt{2}\sigma_1^{2(m)}}{\sqrt{mI(J-2)J}}\right].
\end{aligned}$$

Therefore

$$\eta'(\boldsymbol{\theta}) = \left[(1-\alpha)\sqrt{\frac{2J}{mI(J-2)}}\,0\cdots 0\,\frac{-\sqrt{2}\sigma_1^{2(1)}}{\sqrt{mI(J-2)J}}\cdots\frac{-\sqrt{2}\sigma_1^{2(m)}}{\sqrt{mI(J-2)J}}\right].$$

By using the prior $\pi(\boldsymbol{\theta}) \propto (1-\alpha)^{-1}\prod_{l=1}^{m}\sigma_1^{-2(l)}$ it follows that

$$\sum_{i=1}^{2m+1}\frac{\partial}{\partial\theta_i}\left\{\eta_i(\boldsymbol{\theta})\pi(\boldsymbol{\theta})\right\} = 0.$$

Therefore

$$\pi_{PMP}\left(\alpha,\theta^{(1)},...,\theta^{(m)},\sigma_1^{2(1)},...,\sigma_1^{2(m)}\right) \propto (1-\alpha)^{-1}\prod_{l=1}^{m}\sigma_1^{-2(l)}$$

is a probability matching prior for the parameter α. □

4.3.2 Reference Priors

Theorem 4.3. *The reference prior for the ordering* $\{\alpha, \theta^{(1)}, ..., \theta^{(m)}, \sigma_1^{2(1)}, ..., \sigma_1^{2(m)}\}$ *is*

$$\pi_R\left(\alpha, \theta^{(1)}, ..., \theta^{(m)}, \sigma_1^{2(1)}, ..., \sigma_1^{2(m)}\right) \propto (1-\alpha)^{-1} \prod_{l=1}^{m} \sigma_1^{-2(l)}.$$

Proof. To calculate the reference prior for the ordering $\{\alpha, \theta^{(1)}, ..., \theta^{(m)}, \sigma_1^{2(1)}, ..., \sigma_1^{2(m)}\}$,

$$h_1 = \tilde{F}_{11} - \tilde{F}_{12}\tilde{F}_{22}^{-1}\tilde{F}_{21}$$

must first be calculated. As shown $\tilde{F}_{11} = \frac{I}{(1-\alpha)^2}$,

$$\tilde{F}_{12} = \left[0 \cdots 0 \, \frac{1}{2(1-\alpha)^2 \sigma_1^{2(1)}} \cdots \frac{1}{2(1-\alpha)^2 \sigma_1^{2(m)}}\right],$$

and $\tilde{F}_{22} = \begin{bmatrix} D_{11} & \mathbf{0} \\ \mathbf{0} & D_{22} \end{bmatrix}$. Therefore $h_1 = \frac{1}{(1-\alpha)^2}\left(I - \frac{m}{2}\right)$ which means that $\pi(\alpha) \propto h_1^{\frac{1}{2}} = \frac{1}{1-\alpha}$. Since the Fisher information matrix does not contain $\theta^{(1)}, ..., \theta^{(m)}$,

$$\pi\left(\theta^{(1)}, ..., \theta^{(m)}\right) = \prod_{l=1}^{m} p\left(\theta^{(l)}\right),$$

where

$$\pi\left(\theta^{(l)}\right) \propto \text{constant}, \qquad l = 1, 2, ..., m.$$

To derive the reference prior for $\sigma_1^{2(l)}$, the matrix $H = F_{11}^* - F_{12}^* F_{22}^{*-1} F_{21}^*$ must first be obtained. The submatrix F_{11}^* is a $(m+2)\times(m+2)$ matrix that consists of the first $(m+2)$ rows and the first $(m+2)$ columns of the Fisher information matrix. Therefore

$$F_{11}^* = \begin{bmatrix} \frac{I}{(1-\alpha)^2} & 0 & 0 & \cdots & 0 & \frac{I}{2\sigma_1^{2(1)}(1-\alpha)} \\ 0 & \frac{IJ(1-\alpha)}{\sigma_1^{2(1)}} & 0 & \cdots & 0 & 0 \\ 0 & 0 & \frac{IJ(1-\alpha)}{\sigma_1^{2(2)}} & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & \frac{IJ(1-\alpha)}{\sigma_1^{2(m)}} & 0 \\ \frac{I}{2\sigma_1^{2(1)}(1-\alpha)} & 0 & 0 & \cdots & 0 & \frac{IJ}{2\left(\sigma_1^{2(1)}\right)^2} \end{bmatrix}.$$

Further $F_{12}^{*}=\begin{bmatrix}\frac{I}{2(1-\alpha)\sigma_1^{2(2)}} & \cdots & \frac{I}{2(1-\alpha)\sigma_1^{2(m)}}\\ 0 & \cdots & 0\\ \vdots & & \vdots\\ 0 & \cdots & 0\\ 0 & \cdots & 0\end{bmatrix}$. F_{12}^{*} is a $(m+2)\times(m-1)$ matrix that consists of the first $(m+2)$ rows and last $(m-1)$ columns of the Fisher information matrix. Also $F_{21}^{*}=F_{12}^{*'}$ and $F_{22}^{*}=\begin{bmatrix}\frac{IJ}{2\left(\sigma_1^{2(2)}\right)^2} & \cdots & 0\\ \vdots & \ddots & \vdots\\ 0 & & \frac{IJ}{2\left(\sigma_1^{2(m)}\right)^2}\end{bmatrix}$ is an $(m-1)\times(m-1)$ diagonal matrix consisting of the last $(m-1)$ rows and last $(m-1)$ columns of the Fisher information matrix. From this it follows that $\begin{bmatrix}H_{11} & H_{12} & \cdots & H_{1,m+2}\\ H_{21} & H_{22} & \cdots & H_{2,m+2}\\ \vdots & \vdots & & \vdots\\ H_{m+2,1} & H_{m+2,2} & \cdots & H_{m+2,m+2}\end{bmatrix}=F_{11}^{*}-F_{12}^{*}F_{22}^{*-1}F_{21}^{*}$ where $h_{m+2}=H_{m+2,m+2}=\frac{IJ}{2\left(\sigma_1^{2(1)}\right)^2}-0.$ Since $\pi\left(\sigma_1^2\right)\propto h_{m+2}^{\frac{1}{2}}$ it follows that the reference prior for $\sigma_1^{2(1)}$ is $\pi\left(\sigma_1^{2(1)}\right)\propto\frac{1}{\sigma_1^{2(1)}}$.Similar for the other variance components. Therefore

$$\pi_R\left(\alpha,\theta^{(1)},...,\theta^{(m)},\sigma_1^{2(1)},...,\sigma_1^{2(m)}\right)\propto(1-\alpha)^{-1}\prod_{l=1}^{m}\sigma_1^{-2(l)}$$

is a reference prior. □

It is therefore clear that the probability matching prior and reference prior are the same.

4.4 The Posterior Distribution of α

The parameter of interest is Cronbach's alpha, therefore the marginal posterior distribution for α will be derived in the following theorem.

Theorem 4.4. *The posterior distribution of α is given by*

$$\pi(\alpha|data)\propto(1-\alpha)^{\frac{m}{2}(I-1)-1}\prod_{l=1}^{m}\left\{\left(v_2m_2^{(l)}(1-\alpha)+v_1m_1^{(l)}\right)^{-\frac{1}{2}(IJ-1)}\right\}. \tag{4.4}$$

Proof. By multiplying the likelihood defined in Equation 4.3 by the reference (probability matching prior), the joint posterior

$$\pi\left(\boldsymbol{\theta}|data\right)\propto\left(1-\alpha\right)^{\frac{mI}{2}-1}\left\{\prod_{l=1}^{m}\left(\sigma_1^{2(l)}\right)^{-\frac{1}{2}(IJ+2)}\right\}\exp\left\{-\frac{1}{2}\sum_{l=1}^{m}\frac{1}{\sigma_1^{2(l)}}\left[v_2m_2^{(l)}\left(1-\alpha\right)+v_1m_1^{(l)}\right]\right\}$$
$$\times\exp\left\{-\frac{1}{2}\sum_{l=1}^{m}\frac{1}{\sigma_1^{2(l)}}\left[IJ\left(\bar{Y}_{..}^{(l)}-\theta^{(l)}\right)^2\left(1-\alpha\right)\right]\right\}$$

is obtained. Consider

$$\prod_{l=1}^{m}\left\{\int_{-\infty}^{\infty}\exp\left\{-\frac{1}{2\sigma_1^{2(l)}}\left[IJ\left(\bar{Y}_{..}^{(l)}-\theta^{(l)}\right)^2\left(1-\alpha\right)\right]\right\}d\theta^{(l)}\right\}$$
$$=\left\{\frac{\sqrt{2\pi}\sqrt{\sigma_1^{2(l)}}}{\sqrt{IJ\left(1-\alpha\right)}}\right\}.$$

From this it follows that

$$\pi\left(\boldsymbol{\theta}|data\right)\propto\left(1-\alpha\right)^{\frac{mI}{2}-1}\left\{\prod_{l=1}^{m}\left(\sigma_1^{2(l)}\right)^{-\frac{1}{2}(IJ+1)}\right\}\exp\left\{-\frac{1}{2}\sum_{l=1}^{m}\frac{1}{\sigma_1^{2(l)}}\left[v_2m_2^{(l)}\left(1-\alpha\right)+v_1m_1^{(l)}\right]\right\}.$$

Further

$$\prod_{l=1}^{m}\left\{\int_{0}^{\infty}\left(\sigma_1^{2(l)}\right)^{-\frac{1}{2}(IJ+1)}\exp\left\{-\frac{1}{2}\sum_{l=1}^{m}\frac{1}{\sigma_1^{2(l)}}\left[v_2m_2^{(l)}\left(1-\alpha\right)+v_1m_1^{(l)}\right]\right\}d\sigma_1^{2(l)}\right\}$$
$$=\prod_{l=1}^{m}\left\{\left(\frac{2}{v_2m_2^{(l)}\left(1-\alpha\right)+v_1m_1^{(l)}}\right)^{\frac{1}{2}(IJ-1)}\Gamma\left(\frac{IJ-1}{2}\right)\right\},$$

and

$$\pi\left(\alpha|data\right)\propto\left(1-\alpha\right)^{\frac{m}{2}(I-1)-1}\prod_{l=1}^{m}\left\{\left(v_2m_2^{(l)}\left(1-\alpha\right)+v_1m_1^{(l)}\right)^{-\frac{1}{2}(IJ-1)}\right\}.$$

□

4.5 Examples

4.5.1 Example 1

Two samples are obtained from random-effects models. The sample values for the first sample are $v_1m_1^{(1)}=414.0005$ and $v_2m_2^{(1)}=297.2577$ and for the second sample $v_1m_1^{(2)}=328.4378$ and

$v_2 m_2^{(2)} = 126.7261$. Also $I = 20$ and $J = 6$ for each sample. The variance components of the first random-effects model are $\sigma_1^{2(1)} = 4$ and $\sigma_2^{2(1)} = 1$ and for the second model $\sigma_1^{2(2)} = 3$ and $\sigma_2^{2(2)} = 0.75$. Although the corresponding variance components of the two models differ, they have the same Cronbach's alpha coefficient namely, $\alpha = 0.6$. By using Theorem 4.4 the posterior distributions of α are obtained and are displayed in Figure 4.1 and Table 4.1.

Table 4.1: Central Values, Variances and Credibility Intervals for the Posterior Distributions of α.

	Mean	Median	Mode	Variance	90% CI	95% CI
Combined	0.6469	0.6550	0.673	0.0080	(0.503-0.788)	(0.467-0.809)
1^{st} Sample	0.7128	0.7250	0.752	0.0101	(0.556-0.858)	(0.506-0.883)
2^{nd} Sample	0.4773	0.4940	0.538	0.0286	(0.197-0.749)	(0.123-0.772)

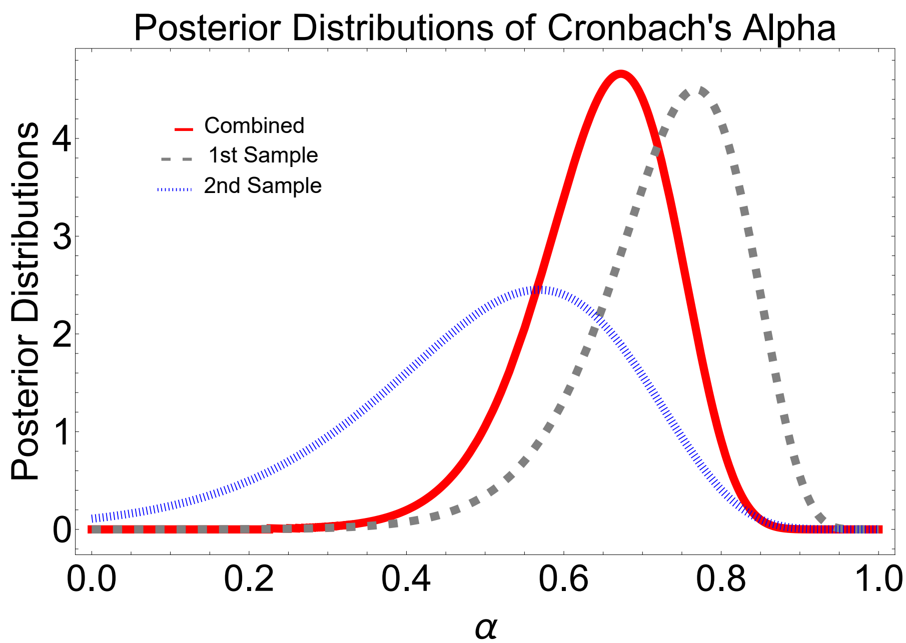


Figure 4.1: Posterior Distributions of Cronbach's Alpha.

Inspection of Figure 4.1 and Table 4.1 shows that for the combined sample the central values are all near 0.6. Also, the variance is less and the credibility intervals are shorter than those for the individual samples.

By making use of sufficient statistics $v_1 m_1^{(1)} \sim \sigma_1^{2(1)} \chi_{v_1}^2$ and $v_2 m_2^{(1)} \sim \sigma_2^{2(1)} \chi_{v_2}^2$ and also that $v_1 m_1^{(2)} \sim \sigma_1^{2(2)} \chi_{v_1}^2$ and $v_2 m_2^{(2)} \sim \sigma_2^{2(2)} \chi_{v_2}^2$ a 100000 first sample and a 100000 second sample observations are simulated. For the posterior mean the following are calculated

$$MSE_1 = Mean\left[(Posterior\,Mean - 0.6)^2\right]$$

and P_1 = Percentage MSE_1 that is the smallest. Also for the posterior mode

$$MSE_2 = Mean\left[(Posterior\,Mode - 0.6)^2\right]$$

and P_2 = Percentage MSE_2 that is the smallest. The results are given in Table 4.2.

Table 4.2: A Comparison of the Results for the Combined, First and Second Samples.

	MSE_1	P_1	MSE_2	P_2	95% Coverage	Mean Length
Combined	0.0119	37.29%	0.0109	30.06%	94.20%	0.4022
1^{st} Sample	0.0220	31.49%	0.0251	35.06%	95.09%	0.5431
2^{nd} Sample	0.0222	31.22%	0.0250	34.88%	94.96%	0.5426

The credibility intervals are H.P.D intervals. The frequentist coverage of the intervals are for all practical purposes 95%. The mean interval length for the combined sample is shorter than that for the individual samples. MSE_1 and MSE_2 for the combined sample are also much smaller than the corresponding values for samples 1 and 2.

4.5.2 Example 2

In this example the two-sample case is extended to three samples. Consider $\sigma_1^{2(1)} = 4$, $\sigma_2^{2(1)} = 1$,$\sigma_1^{2(2)} = 3$, $\sigma_2^{2(2)} = 0.75$, $\sigma_1^{2(3)} = 5$ and $\sigma_2^{2(3)} = 1.25$. The sample values are

Sample 1: $v_1 m_1^{(1)} = 349.9693$ and $v_2 m_2^{(1)} = 151.5623$

Sample 2: $v_1 m_1^{(2)} = 274.5413$ and $v_2 m_2^{(2)} = 100.2209$

Sample 3: $v_1 m_1^{(3)} = 386.3027$ and $v_2 m_2^{(3)} = 220.3655$.

The posterior density functions of α are illustrated in Figure 4.2

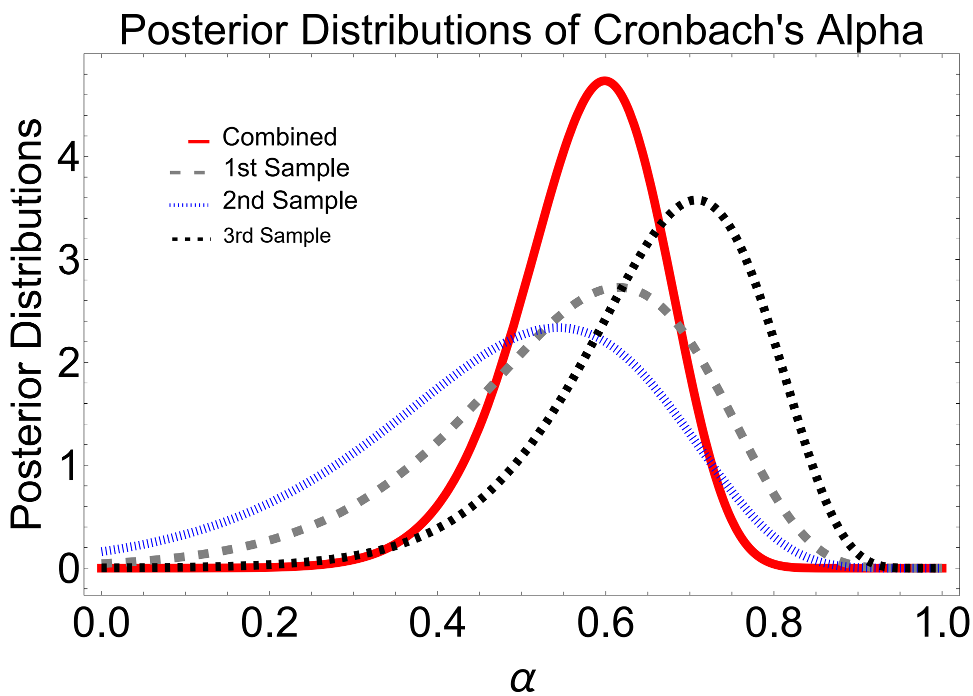


Figure 4.2: Posterior Distributions of Cronbach's Alpha.

and in Table 4.3 important statistics are given.

Table 4.3: Central Values, Variances and Credibility Intervals for the Posterior Distributions of α in the Case of Three Samples.

	Mean	Median	Mode	Variance	90% CI	95% CI
Combined	0.5783	0.5840	0.5990	0.0075	(0.425-0.708)	(0.389-0.728)
1^{st} Sample	0.5285	0.5460	0.5880	0.0251	(0.252-0.765)	(0.182-0.803)
2^{nd} Sample	0.4525	0.4670	0.5120	0.0299	(0.168-0.737)	(0.067-0.766)
3^{rd} Sample	0.6388	0.6540	0.6880	0.0159	(0.423-0.922)	(0.361-0.843)

As in the previous example the sufficient statistics are used to simulate a 100 000 observations for each of the three samples. The results are given in Table 4.4.

Table 4.4: Comparative Results for the Combined, First, Second and Third Samples.

	MSE_1	P_1	MSE_2	P_2	95% Coverage	Mean Length
Combined	0.0079	34.87%	0.0071	37.31%	94.75%	0.3302
1^{st} Sample	0.0221	21.70%	0.0253	21.19%	95.03%	0.5431
2^{nd} Sample	0.0220	21.67%	0.0251	20.89%	95.07%	0.5435
3^{rd} Sample	0.0222	21.76%	0.0254	20.61%	95.07%	0.5430

Inspection of Table 4.4 and also from the results of the previous example it can be concluded that the posterior distribution of α for the combined sample becomes more important as the number of samples increase since the mean square error is the smallest when the combined model is considered.

4.6 Conclusion

In this chapter the combined Bayesian estimate of α for m experiments with equal α but possibly different variance components are derived. Since the model considered is the one-way balanced random effects model, the assumption of equicorrelated normal data is satisfied. Our Bayesian results are therefore different from those of Van Zyl (2001) who had to make the assumption of equicorrelation. Reference and probability matching priors are derived in Section 4.3. They lead to procedures with properties that frequentists can relate to while still retaining Bayesian validity. The fact that the resulting Bayesian posterior intervals of level $1-\beta$ are also good frequentist intervals of the same level is a very desirable situation. It is also shown that the reference and probability matching priors for α are the same. The Bayesian theory and results derived in Sections 4.3 and 4.4 are applied to two examples in Section 4.5. The frequentist coverage of the credibility intervals are for all practical purposes 95%. The intervals for the combined sample are however much shorter than those of the individual samples. Also, the point estimates of the combined sample are more accurate. It is further concluded that the posterior distribution of α for the combined sample becomes more important as the number of samples increase.

Chapter 5

Cronbach's Alpha for a Three-Component Model

5.1 Introduction

This chapter is an extension of the work done by Izally et al. (2024) and the work done in Chapter 3. The work done in this chapter will focus on deriving objective priors for Cronbach's alpha where the balanced two-factor nested random effects model by Van der Merwe & Hugo (2007) will be used. In Izally et al. (2024), the one-way random effects model defined in Equation 3.1 was considered, where it was shown that the covariance matrix is given by

$$\Sigma = \begin{bmatrix} \sigma_1^2+\sigma_2^2 & \sigma_2^2 & \sigma_2^2 & \dots & \sigma_2^2 \\ \sigma_2^2 & \sigma_1^2+\sigma_2^2 & \sigma_2^2 & \dots & \sigma_2^2 \\ \sigma_2^2 & \dots & \sigma_1^2+\sigma_2^2 & \cdots & \sigma_2^2 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \sigma_2^2 & \sigma_2^2 & \sigma_2^2 & \cdots & \sigma_1^2+\sigma_2^2 \end{bmatrix}_{J\times J}$$

and is also compound symmetric. By definition, Cronbach's alpha is given by

$$\begin{aligned} \alpha =& \frac{J}{(J-1)}\left\{1-\frac{trace\,(\Sigma)}{\mathbf{1_J'}\Sigma\mathbf{1_J}}\right\} \\ =& 1-\frac{\sigma_1^2}{\sigma_1^2+J\sigma_2^2} \\ =& 1-\frac{E\,(\text{Within-Mean Squared Error})}{E\,(\text{Between-Mean Squared Error})}, \end{aligned} \tag{5.1}$$

where $\mathbf{I}$ is the $J\times J$ identity matrix and $\mathbf{1} = [1\,1 \ldots 1]'$ is a $J\times 1$ column vector of ones. The definition of Cronbach's alpha in Equation 5.1 will be used in this chapter to extend the work done by Izally et al.

(2024). Consider a three-component hierarchical variance components model

$$Y_{ijk} = \theta + r_i + c_{ij} + \varepsilon_{ijk} \quad \text{for } i = 1,...I \text{ and } j = 1,...J,\ k = 1,...K, \tag{5.2}$$

where the mean effect θ is unknown. The r_i, c_{ij} and ε_{ijk} are independent normal variables with zero means and unknown variances σ_3^2, σ_2^2 and σ_1^2 respectively. Let

$$\mathbf{Y} = \left[Y_{i11}\, Y_{i12} \ldots Y_{i1K} Y_{i21}\, Y_{i22} \ldots Y_{i2K} Y_{iJ1}\, Y_{iJ2} \ldots Y_{iJK}\right]'.$$

It can be shown that

$$Var\left(\mathbf{Y}|\theta, \sigma_1^2, \sigma_2^2, \sigma_3^2\right) = \tilde{\mathbf{1}}\tilde{\mathbf{1}}'\sigma_3^2 + \sigma_2^2 \begin{bmatrix} \mathbf{11}' & 0 & 0 \\ 0 & \mathbf{11}' & 0 \\ 0 & 0 & \mathbf{11}' \end{bmatrix} + \sigma_1^2 \mathbf{I},$$

where $\mathbf{I}$ is the $JK \times JK$ identity matrix, $\tilde{\mathbf{1}} = [1\,1 \ldots 1]'$ is a $JK \times 1$ column vector of ones and $\mathbf{1} = [1\,1 \ldots 1]'$ is a $k \times 1$ column vector of ones. If we let $J = 3$ and $K = 2$ then the covariance matrix is given by

$$\Sigma = \begin{bmatrix} \sigma_1^2+\sigma_2^2+\sigma_3^2 & \sigma_2^2+\sigma_3^2 & \sigma_3^2 & \sigma_3^2 & \sigma_3^2 & \sigma_3^2 \\ \sigma_2^2+\sigma_3^2 & \sigma_1^2+\sigma_2^2+\sigma_3^2 & \sigma_3^2 & \sigma_3^2 & \sigma_3^2 & \sigma_3^2 \\ \sigma_3^2 & \sigma_3^2 & \sigma_1^2+\sigma_2^2+\sigma_3^2 & \sigma_2^2+\sigma_3^2 & \sigma_3^2 & \sigma_3^2 \\ \sigma_3^2 & \sigma_3^2 & \sigma_2^2+\sigma_3^2 & \sigma_1^2+\sigma_2^2+\sigma_3^2 & \sigma_3^2 & \sigma_3^2 \\ \sigma_3^2 & \sigma_3^2 & \sigma_3^2 & \sigma_3^2 & \sigma_1^2+\sigma_2^2+\sigma_3^2 & \sigma_2^2+\sigma_3^2 \\ \sigma_3^2 & \sigma_3^2 & \sigma_3^2 & \sigma_3^2 & \sigma_2^2+\sigma_3^2 & \sigma_1^2+\sigma_2^2+\sigma_3^2 \end{bmatrix}_{6\times 6}.$$

Cronbach's alpha is defined by

$$\begin{aligned} \alpha^\star =& \frac{JK}{(JK-1)}\left\{1 - \frac{trace\,(\Sigma)}{\mathbf{1}_{\mathbf{IJ}}'\Sigma\mathbf{1}_{\mathbf{IJ}}}\right\} \\ =& \frac{JK}{(JK-1)}\left\{1 - \frac{JK\left(\sigma_1^2+\sigma_2^2+\sigma_3^2\right)}{JK\sigma_1^2 + JK^2\sigma_2^2 + J^2K^2\sigma_3^2}\right\} \\ =& \frac{JK}{(JK-1)}\left\{\frac{(K-1)\,\sigma_2^2 + (JK-1)\,\sigma_3^2}{\sigma_1^2 + K\sigma_2^2 + JK\sigma_3^2}\right\}. \end{aligned}$$

The covariance matrix is not compound symmetric and therefore the expression for Cronbach's alpha is not easy to work with. Cronbach's alpha will therefore be redefined such that the likelihood function for the hierarchical model can be expressed in terms of Cronbach's alpha to make it easier to derive objective priors for the hierarchical model. Define Cronbach's alpha using Equation 5.1 by

$$\begin{aligned}\alpha =& 1-\frac{E\left(\text{Within-Mean Squared Error}\right)}{E\left(\text{Between-Mean Squared Error}\right)}\\ =& 1-\frac{\sigma_1^2}{\sigma_1^2+K\sigma_2^2+JK\sigma_3^2}\\ =& \frac{K\sigma_2^2+JK\sigma_3^2}{\sigma_1^2+K\sigma_2^2+JK\sigma_3^2}.\end{aligned} \tag{5.3}$$

In this chapter a number of objective priors for Cronbach's alpha in the case of a balanced three-component hierarchical model will be derived. The following objective priors will be investigated: the well-known Jeffreys prior (from Jeffreys, 1939), a reference prior (see Berger & Bernardo, 1992 and Berger et al., 2009), the probability matching prior using the method proposed by Datta & Ghosh (1995) and a divergence prior developed by Ghosh et al. (2011). A reference prior for the group orderings $\left\{\alpha,\sigma_1^2,\sigma_2^2,\theta\right\}$, $\left\{\alpha,\sigma_1^2,\theta,\sigma_2^2,\right\}$, $\left\{\alpha,\theta,\sigma_1^2,\sigma_2^2\right\}$ where α is considered the most important parameter in each case, will be derived using the algorithm in Berger & Bernardo (1992). The marginal posterior for α will be derived using the different priors and a simulation study will be conducted where the frequentist coverage rates, the average interval lengths and the standard deviation of the interval lengths will be computed to compare the different priors. In Section 5.2 the priors and posteriors are derived, a simulation study is considered in Section 5.3 and two examples are considered in Section 5.4. Concluding remarks are given in Section 5.5 and all the proofs are provided in Appendix C.

5.2 Priors and Posteriors

In this section, the likelihood, priors and posteriors are derived. By using the results given in Box & Tiao (1973), the integrated likelihood can be written as

$$\begin{aligned}\ell(\theta,\sigma_1^2,\sigma_2^2,\sigma_3^2|\text{data}) \propto& \left(\sigma_1^2\right)^{-\frac{v_1}{2}}\left(\sigma_1^2+K\sigma_2^2\right)^{-\frac{v_2}{2}}\left(\sigma_1^2+K\sigma_2^2+JK\sigma_3^2\right)^{-\frac{1}{2}(v_3+1)}\times\\ &\exp\left\{-\frac{1}{2}\left[\frac{IJK\left(\bar{Y}_{...}-\theta\right)^2}{\sigma_1^2+K\sigma_2^2+JK\sigma_3^2}+\frac{v_3m_3}{\sigma_1^2+K\sigma_2^2+JK\sigma_3^2}+\frac{v_2m_2}{\sigma_1^2+K\sigma_2^2}+\frac{v_1m_1}{\sigma_1^2}\right]\right\},\end{aligned} \tag{5.4}$$

where $\bar{Y}_{ij.}=\frac{1}{K}\sum_{k=1}^{K}Y_{ijk}$, $\bar{Y}_{i..}=\frac{1}{JK}\sum_{j=1}^{J}\sum_{k=1}^{K}Y_{ijk}$, $\bar{Y}_{...}=\frac{1}{IJK}\sum_{i=1}^{I}\sum_{j=1}^{J}\sum_{k=1}^{K}Y_{ijk}$, $v_1m_1=\sum_{i=1}^{I}\sum_{j=1}^{J}\sum_{k=1}^{K}(Y_{ijk}-\bar{Y}_{ij.})^2$, $v_2m_2=K\sum_{i=1}^{I}\sum_{j=1}^{J}(\bar{Y}_{ij.}-\bar{Y}_{i..})^2$, $v_3m_3=KJ\sum_{i=1}^{I}(\bar{Y}_{i..}-\bar{Y}_{...})^2$, $v_1=IJ(K-1)$, $v_2=I(J-1)$ and $v_3=I-1$. By integrating the multivariate normal density from Equation 3.1 with respect to the random effects, r_i $(i=1,2,...,I)$ and $c_{ij}(j=1,2,...,J)$, the integrated likelihood can be obtained. Using Equation 5.3 we have

$$1-\alpha = \frac{\sigma_1^2}{\sigma_1^2 + K\sigma_2^2 + JK\sigma_3^2},$$

therefore the likelihood function can be written in terms of Cronbach's alpha. The likelihood can be shown to be

$$\ell(\theta,\sigma_1^2,\sigma_2^2,\alpha|\text{data}) \propto \left(\sigma_1^2\right)^{-\frac{1}{2}(v_1+v_3+1)}\left(\sigma_1^2+K\sigma_2^2\right)^{-\frac{v_2}{2}}(1-\alpha)^{\frac{1}{2}(v_3+1)} \times \exp\left\{-\frac{1}{2}\left[\frac{IJK\left(\bar{Y}_{\dots}-\theta\right)^2(1-\alpha)}{\sigma_1^2} + \frac{v_3 m_3 (1-\alpha)}{\sigma_1^2} + \frac{v_2 m_2}{\sigma_1^2+K\sigma_2^2} + \frac{v_1 m_1}{\sigma_1^2}\right]\right\}. \tag{5.5}$$

Using the likelihood function, it can be shown that the Fisher information matrix is given by

$$F(\theta,\sigma_1^2,\sigma_2^2,\alpha) = \begin{bmatrix} \frac{IJK(1-\alpha)}{\sigma_1^2} & 0 & 0 & 0 \\ 0 & \frac{(v_1+v_3+1)}{2\left(\sigma_1^2\right)^2} + \frac{v_2}{2\left(\sigma_1^2+K\sigma_2^2\right)^2} & \frac{v_2 K}{2\left(\sigma_1^2+K\sigma_2^2\right)^2} & \frac{(v_3+1)}{2(1-\alpha)\sigma_1^2} \\ 0 & \frac{v_2 K}{2\left(\sigma_1^2+K\sigma_2^2\right)^2} & \frac{v_2 K^2}{2\left(\sigma_1^2+K\sigma_2^2\right)^2} & 0 \\ 0 & \frac{(v_3+1)}{2(1-\alpha)\sigma_1^2} & 0 & \frac{(v_3+1)}{2(1-\alpha)^2} \end{bmatrix}. \tag{5.6}$$

The Jeffreys prior is given by

$$\pi_J(\theta,\sigma_1^2,\sigma_2^2,\alpha) \propto (\sigma_1^2)^{-3/2}(1-\alpha)^{-1/2}\left(K\sigma_2^2+\sigma_1^2\right)^{-\frac{1}{2}}. \tag{5.7}$$

The divergence prior is given by

$$\pi_D(\theta,\sigma_1^2,\sigma_2^2,\alpha) \propto (\sigma_1^2)^{-3/4}(1-\alpha)^{-1/4}\left(K\sigma_2^2+\sigma_1^2\right)^{-\frac{1}{4}}. \tag{5.8}$$

The reference prior for the group ordering $\left\{\alpha,\sigma_1^2,\sigma_2^2,\theta\right\}$ is derived in the following theorem.

Theorem 5.1. *The reference prior for the group ordering* $\left\{\alpha,\sigma_1^2,\sigma_2^2,\theta\right\}$ *is given by*

$$\pi_R\left(\alpha,\sigma_1^2,\sigma_2^2,\theta\right) \propto \left(\sigma_1^2\right)^{-1}(1-\alpha)^{-1}\left(K\sigma_2^2+\sigma_1^2\right)^{-1}.$$

Proof. We are interested in the reference prior for the group ordering $\left\{\alpha,\sigma_1^2,\sigma_2^2,\theta\right\}$ which means that α is the most important parameter and θ is the least important parameter. In order to derive the reference prior, the inverse of the Fisher information matrix is needed. Let $S(\boldsymbol{\theta}) = H^{-1}(\boldsymbol{\theta})$ where H

is the Fisher information matrix. Now

$$S\left(\alpha,\sigma_1^2,\sigma_2^2,\theta\right)=\begin{bmatrix}A_{11} & A_{12} & A_{13} & A_{14}\\ A_{21} & A_{22} & A_{23} & A_{24}\\ A_{31} & A_{32} & A_{33} & A_{34}\\ A_{41} & A_{42} & A_{43} & A_{44}\end{bmatrix}=\begin{bmatrix}\frac{2(1-\alpha)^2(v_1+v_3+1)}{v_1(v_3+1)} & \frac{-2\left(\sigma_1^2\right)(1-\alpha)}{v_1} & \frac{2\left(\sigma_1^2\right)(1-\alpha)}{v_1K} & 0\\ \frac{-2\left(\sigma_1^2\right)(1-\alpha)}{v_1} & \frac{2\left(\sigma_1^2\right)^2}{v_1} & \frac{-2\left(\sigma_1^2\right)^2}{v_1K} & 0\\ \frac{2\left(\sigma_1^2\right)(1-\alpha)}{v_1K} & \frac{-2\left(\sigma_1^2\right)^2}{v_1K} & \frac{2}{K^2}\left[\frac{\left(\sigma_1^2\right)^2}{v_1}+\frac{\left(K\sigma_2^2+\sigma_1^2\right)^2}{v_2}\right] & 0\\ 0 & 0 & 0 & \frac{\sigma_1^2}{IJK(1-\alpha)}\end{bmatrix}.$$

Define the truncated ranges for the 4 parameters as $\alpha\in[a_l,b_l]$, $\theta\in[c_l,d_l]$, $\sigma_1^2\in[e_l,f_l]$ and $\sigma_2^2\in[g_l,h_l]$, where $c_l\to-\infty$, $b_l\to 1$, $d_l,f_l,h_l\to\infty$ and $a_1,e_l,g_l\to 0$. Now

$$h_1\equiv H_1\equiv A_{11}^{-1}=\frac{I(J-1)}{2J(1-\alpha)^2},$$

and

$$h_2=\left(A_{22}-B_2H_1B_2'\right)^{-1}.$$

Now S_1 is the upper left $N_1\times N_1$ corner of S. We have $m=4$ groups with $n_1=1$, $n_2=1$, $n_3=1$ and $n_4=1$. This implies that S_1 is the upper left 1×1 corner of S, that is

$$S_1=\frac{2(1-\alpha)^2(v_1+v_3+1)}{v_1(v_3+1)}.$$

Therefore

$$H_1=S_1^{-1}=\frac{v_1(v_3+1)}{2(1-\alpha)^2(v_1+v_3+1)}$$

and

$$B_2=A_{21}=\frac{-2\left(\sigma_1^2\right)(1-\alpha)}{v_1}.$$

Hence

$$\begin{aligned}h_2&=\left[\frac{2\left(\sigma_1^2\right)^2}{v_1}-\left(\frac{-2\left(\sigma_1^2\right)(1-\alpha)}{v_1}\right)\frac{I(J-1)}{2J(1-\alpha)^2}\left(\frac{-2\left(\sigma_1^2\right)(1-\alpha)}{v_1}\right)\right]^{-1}\\&=\frac{(v_3+v_1+1)}{2\left(\sigma_1^2\right)^2}.\end{aligned}$$

Also

$$h_3=\left(A_{33}-B_3H_2B_3'\right)^{-1}$$

where S_2 is the upper left $N_2\times N_2$ corner of S. Now $N_2=n_1+n_2=2$. Therefore S_2 is the upper left

2×2 corner of S, that is

$$S_2 = \begin{bmatrix} \frac{2(1-\alpha)^2(v_1+v_3+1)}{v_1(v_3+1)} & \frac{-2\left(\sigma_1^2\right)(1-\alpha)}{v_1} \\ \frac{-2\left(\sigma_1^2\right)(1-\alpha)}{v_1} & \frac{2\left(\sigma_1^2\right)^2}{v_1} \end{bmatrix}.$$

Now

$$H_2 = S_2^{-1} = \begin{bmatrix} \frac{(v_3+1)}{2(1-\alpha)^2} & \frac{(v_3+1)}{2\left(\sigma_1^2\right)(1-\alpha)} \\ \frac{(v_3+1)}{2\left(\sigma_1^2\right)(1-\alpha)} & \frac{(v_1+v_3+1)}{2\left(\sigma_1^2\right)^2} \end{bmatrix}$$

and

$$B_3 = \begin{bmatrix} A_{31} & A_{32} \end{bmatrix} = \begin{bmatrix} \frac{2\left(\sigma_1^2\right)(1-\alpha)}{v_1 K} & \frac{-2\left(\sigma_1^2\right)^2}{v_1 K} \end{bmatrix}.$$

Therefore

$$B_3 H_2 B_3' = \frac{2\left(\sigma_1^2\right)^2}{v_1 K^2}.$$

So

$$\begin{aligned} h_3 &= \left(\frac{2}{K^2} \left[\frac{\left(\sigma_1^2\right)^2}{v_1} + \frac{\left(K\sigma_2^2 + \sigma_1^2\right)^2}{v_2} \right] - \frac{2\left(\sigma_1^2\right)^2}{v_1 K^2} \right)^{-1} \\ &= \frac{v_2 K^2}{2\left(K\sigma_2^2 + \sigma_1^2\right)^2}. \end{aligned}$$

Now

$$h_4 = \left(A_{44} - B_4 H_3 B_4'\right)^{-1}$$

where S_3 is the upper left 3×3 corner of S, that is

$$S_3 = \begin{bmatrix} \frac{2(1-\alpha)^2(v_1+v_3+1)}{v_1(v_3+1)} & \frac{-2\left(\sigma_1^2\right)(1-\alpha)}{v_1} & \frac{2\left(\sigma_1^2\right)(1-\alpha)}{v_1 K} \\ \frac{-2\left(\sigma_1^2\right)(1-\alpha)}{v_1} & \frac{2\left(\sigma_1^2\right)^2}{v_1} & \frac{-2\left(\sigma_1^2\right)^2}{v_1 K} \\ \frac{2\left(\sigma_1^2\right)(1-\alpha)}{v_1 K} & \frac{-2\left(\sigma_1^2\right)^2}{v_1 K} & \frac{2}{K^2}\left[\frac{\left(\sigma_1^2\right)^2}{v_1} + \frac{\left(K\sigma_2^2+\sigma_1^2\right)^2}{v_2} \right] \end{bmatrix}.$$

Now

$$\begin{aligned} H_3 &= S_3^{-1} \\ &= \begin{bmatrix} \frac{(v_3+1)}{2(1-\alpha)^2} & \frac{(v_3+1)}{2\sigma_1^2(1-\alpha)} & 0 \\ \frac{(v_3+1)}{2\sigma_1^2(1-\alpha)} & \frac{v_2}{2\left(K\sigma_2^2+\sigma_1^2\right)^2} + \frac{(v_1+v_3+1)}{2\left(\sigma_1^2\right)^2} & \frac{K v_2}{2\left(K\sigma_2^2+\sigma_1^2\right)^2} \\ 0 & \frac{K v_2}{2\left(K\sigma_2^2+\sigma_1^2\right)^2} & \frac{K^2 v_2}{2\left(K\sigma_2^2+\sigma_1^2\right)^2} \end{bmatrix}. \end{aligned}$$

Now $B_4 = \left[\begin{array}{ccc} A_{41} & A_{42} & A_{43} \end{array} \right] = \left[\begin{array}{ccc} 0 & 0 & 0 \end{array} \right]$. Therefore $B_4 H_3 B_4' = 0$. This implies that

$$h_4 = A_{44}^{-1} = \frac{IJK(1-\alpha)}{\sigma_1^2}.$$

The four functions

$$h_1 = \frac{I(J-1)}{2J(1-\alpha)^2}, \qquad h_2 = \frac{(v_3+v_1+1)}{2\left(\sigma_1^2\right)^2}, \qquad h_3 = \frac{v_2 K^2}{2\left(K\sigma_2^2+\sigma_1^2\right)^2}, \qquad h_4 = \frac{IJK(1-\alpha)}{\sigma_1^2}$$

are the functions needed to derive the prior. During the iterations, first the truncated conditional function of θ given α, σ_1^2 and σ_2^2 can be computed as

$$\begin{aligned}
\pi_4^l\left(\theta|\alpha,\sigma_1^2,\sigma_2^2\right) =& \frac{\{h_4(\boldsymbol{\theta})\}^{\frac{1}{2}}}{\int_{\theta_4^l}\{h_4(\boldsymbol{\theta})\}^{\frac{1}{2}}d\theta_4} \\
\propto& \frac{(1-\alpha)^{\frac{1}{2}}\left(\sigma_1^2\right)^{-\frac{1}{2}}}{\int_{c_l}^{d_l}(1-\alpha)^{\frac{1}{2}}\left(\sigma_1^2\right)^{-\frac{1}{2}}d\theta} \\
=& \frac{1}{(d_l-c_l)} \qquad \text{for } c_l \le \theta \le d_l.
\end{aligned}$$

Now

$$\begin{aligned}
& E\left[log\{h_3(\boldsymbol{\theta})\}^{\frac{1}{2}}\right] \\
\propto& \int_{c_l}^{d_l} log\left\{\left(K\sigma_2^2+\sigma_1^2\right)^{-1}\right\}\frac{1}{(d_l-c_l)}d\theta \\
=& log\left\{\left(K\sigma_2^2+\sigma_1^2\right)^{-1}\right\}.
\end{aligned}$$

Therefore,

$$\begin{aligned}
\pi_3^l\left(\sigma_2^2,\theta|\alpha,\sigma_1^2\right) =& \frac{\pi_4^l\left(\theta|\alpha,\sigma_1^2,\sigma_2^2\right)exp\left\{E\left[log\{h_3(\boldsymbol{\theta})\}^{\frac{1}{2}}\right]\right\}}{\int_{g_l}^{h_l}exp\left\{E\left[log\{h_3(\boldsymbol{\theta})\}^{\frac{1}{2}}\right]\right\}d\sigma_2^2} \\
=& \frac{\frac{1}{(d_l-c_l)}exp\left\{log\left\{\left(K\sigma_2^2+\sigma_1^2\right)^{-1}\right\}\right\}}{\int_{g_l}^{h_l}exp\left\{log\left\{\left(K\sigma_2^2+\sigma_1^2\right)^{-1}\right\}\right\}d\sigma_2^2} \\
=& \frac{K}{(d_l-c_l)}\frac{\left(K\sigma_2^2+\sigma_1^2\right)^{-1}}{log\left(\frac{Kh_l+\sigma_1^2}{Kg_l+\sigma_1^2}\right)} \qquad \text{for } g_l \le \sigma_2^2 \le h_l \text{ and } c_l \le \theta \le d_l.
\end{aligned}$$

The function, h_2, is now needed to determine $E\left[log\left\{h_2\left(\boldsymbol{\theta}\right)\right\}^{\frac{1}{2}}\right]$

$$\begin{aligned}
&E\left[log\left\{h_2\left(\boldsymbol{\theta}\right)\right\}^{\frac{1}{2}}\right]\\
&\propto\int_{c_l}^{d_l}\int_{g_l}^{h_l}-log\left(\sigma_1^2\right)\frac{K}{(d_l-c_l)}\frac{\left(K\sigma_2^2+\sigma_1^2\right)^{-1}}{log\left(\frac{Kh_l+\sigma_1^2}{Kg_l+\sigma_1^2}\right)}d\sigma_2^2d\theta\\
&=\int_{c_l}^{d_l}-log\left(\sigma_1^2\right)\frac{KK^{-1}}{(d_l-c_l)}\frac{log\left(\frac{Kh_l+\sigma_1^2}{Kg_l+\sigma_1^2}\right)}{log\left(\frac{Kh_l+\sigma_1^2}{Kg_l+\sigma_1^2}\right)}d\theta\\
&=log\left(\sigma_1^2\right)^{-1}.
\end{aligned}$$

Therefore

$$\begin{aligned}
\pi_2^l\left(\sigma_1^2,\sigma_2^2,\theta|\alpha\right)=&\frac{\pi_3^l\left(\sigma_2^2,\theta|\alpha,\sigma_1^2\right)exp\left\{E\left[log\left\{h_2\left(\boldsymbol{\theta}\right)\right\}^{\frac{1}{2}}\right]\right\}}{\int_{e_l}^{f_l}exp\left\{E\left[log\left\{h_2\left(\boldsymbol{\theta}\right)\right\}^{\frac{1}{2}}\right]\right\}d\sigma_1^2}\\
=&\frac{K}{(d_l-c_l)}\frac{\left(K\sigma_2^2+\sigma_1^2\right)^{-1}\left(\sigma_1^2\right)^{-1}}{log\left(\frac{Kh_l+\sigma_1^2}{Kg_l+\sigma_1^2}\right)\int_{e_l}^{f_l}\left(\sigma_1^2\right)^{-1}d\sigma_1^2}\\
=&\frac{K}{(d_l-c_l)}\frac{\left(K\sigma_2^2+\sigma_1^2\right)^{-1}\left(\sigma_1^2\right)^{-1}}{log\left(\frac{Kh_l+\sigma_1^2}{Kg_l+\sigma_1^2}\right)log\left(\frac{f_l}{e_l}\right)}
\end{aligned}$$

The function, h_1, is now needed to determine $E\left[log\left\{h_1\left(\boldsymbol{\theta}\right)\right\}^{\frac{1}{2}}\right]$ in order to proceed with the iterative process to determine $\pi_1^l\left(\alpha,\sigma_1^2,\sigma_2^2,\theta\right)$. We have

$$\begin{aligned}
&E\left[log\left\{h_1\left(\boldsymbol{\theta}\right)\right\}^{\frac{1}{2}}\right]\\
&\propto\int_{c_l}^{d_l}\int_{g_l}^{h_l}\int_{e_l}^{f_l}log\left(1-\alpha\right)^{-1}\frac{K}{(d_l-c_l)}\frac{\left(K\sigma_2^2+\sigma_1^2\right)^{-1}\left(\sigma_1^2\right)^{-1}}{log\left(\frac{Kh_l+\sigma_1^2}{Kg_l+\sigma_1^2}\right)log\left(\frac{f_l}{e_l}\right)}d\sigma_1^2d\sigma_2^2d\theta\\
&=\int_{e_l}^{f_l}\int_{g_l}^{h_l}(d_l-c_l)\,log\left(1-\alpha\right)^{-1}\frac{K}{(d_l-c_l)}\frac{\left(K\sigma_2^2+\sigma_1^2\right)^{-1}\left(\sigma_1^2\right)^{-1}}{log\left(\frac{Kh_l+\sigma_1^2}{Kg_l+\sigma_1^2}\right)log\left(\frac{f_l}{e_l}\right)}d\sigma_2^2d\sigma_1^2\\
&=log\left(1-\alpha\right)^{-1}.
\end{aligned}$$

Therefore,

$$\begin{aligned}
\pi_1^l\left(\alpha,\sigma_1^2,\sigma_2^2,\theta\right) =& \frac{\pi_2^l\left(\sigma_1^2,\sigma_2^2,\theta|\alpha\right) exp\left\{E\left[log\left\{h_1\left(\boldsymbol{\theta}\right)\right\}^{\frac{1}{2}}\right]\right\}}{\int_{a_l}^{b_l} exp\left\{E\left[log\left\{h_1\left(\boldsymbol{\theta}\right)\right\}^{\frac{1}{2}}\right]\right\} d\alpha} \\
=& \frac{K}{(d_l-c_l)} \frac{\left(K\sigma_2^2+\sigma_1^2\right)^{-1}\left(\sigma_1^2\right)^{-1}(1-\alpha)^{-1}}{log\left(\frac{Kh_l+\sigma_1^2}{Kg_l+\sigma_1^2}\right) log\left(\frac{f_l}{e_l}\right)\int_{a_l}^{b_l}(1-\alpha)^{-1} d\alpha} \\
=& \frac{K}{(d_l-c_l)} \frac{\left(K\sigma_2^2+\sigma_1^2\right)^{-1}\left(\sigma_1^2\right)^{-1}(1-\alpha)^{-1}}{log\left(\frac{Kh_l+\sigma_1^2}{Kg_l+\sigma_1^2}\right) log\left(\frac{f_l}{e_l}\right) log\left(\frac{1-a_l}{1-b_l}\right)}.
\end{aligned}$$

Finally

$$\pi_R\left(\alpha,\theta,\sigma_1^2,\sigma_2^2\right) \propto \lim_{l\to\infty}\frac{\pi_1^l\left(\alpha,\sigma_1^2,\sigma_2^2,\theta\right)}{\pi_1^l\left(\alpha_0,\sigma_{10}^2,\sigma_{20}^2,\theta_0\right)} \propto \left(K\sigma_2^2+\sigma_1^2\right)^{-1}\left(\sigma_1^2\right)^{-1}(1-\alpha)^{-1},$$

where $\alpha_0,\theta_0,\sigma_{10}^2$ and σ_{20}^2 are the four inner points in the ranges of the parameters. □

The reference prior for the group orderings $\left\{\alpha,\sigma_1^2,\theta,\sigma_2^2\right\}$, $\left\{\alpha,\theta,\sigma_1^2,\sigma_2^2\right\}$, where α is the most important parameter in each case resulted in the same reference prior derived in Theorem 5.1. The proof of the reference prior for these group orderings can be found in Appendices C.2 and C.3 respectively. The following group orderings where α is the most important parameter have also been considered: $\left\{\alpha,\theta,\sigma_2^2,\sigma_1^2\right\}$, $\left\{\alpha,\sigma_2^2,\sigma_1^2,\theta\right\}$ and $\left\{\alpha,\sigma_2^2,\theta,\sigma_1^2\right\}$. The reference prior could not be derived for these group orderings, mainly because the integrals for the $E\left[log\left\{h_j\left(\boldsymbol{\theta}\right)\right\}^{\frac{1}{2}}\right]$ described by the algorithm in Berger & Bernardo (1992) could not be expressed in a closed-form solution. There are other group orderings where α is not the most important parameter that may be considered to derive the reference prior, but the focus in this chapter is on the group ordering where Cronbach's alpha is the most important parameter.

The probability matching prior will be derived in the following theorem.

Theorem 5.2. *The probability matching prior for Cronbach's alpha is given by*

$$\pi_{PMP}\left(\theta,\sigma_1^2,\sigma_2^2,\alpha\right) = G\left(K\sigma_2^2+\sigma_1^2, \frac{\sigma_1^2}{(1-\alpha)^{\frac{(v_3+1)}{(v_1+v_3+1)}}}\right)\left(\sigma_1^2\right)^{-1}(1-\alpha)^{-1} \tag{5.9}$$

where G is an arbitrary continuously differentiable function.

Proof. The inverse of the Fisher information matrix is given by

$$F^{-1}(\boldsymbol{\theta}) = F^{-1}\left(\theta, \sigma_1^2, \sigma_2^2, \alpha\right) = \begin{bmatrix} \frac{\sigma_1^2}{IJK(1-\alpha)} & 0 & 0 & 0 \\ 0 & \frac{2(\sigma_1^2)^2}{v_1} & \frac{-2(\sigma_1^2)^2}{v_1 K} & \frac{-2(1-\alpha)\left(\sigma_1^2\right)}{v_1} \\ 0 & \frac{-2(\sigma_1^2)^2}{v_1 K} & \frac{2}{K^2}\left[\frac{\left(\sigma_1^2\right)^2}{v_1} + \frac{\left(K\sigma_2^2+\sigma_1^2\right)^2}{v_2}\right] & \frac{2(1-\alpha)\left(\sigma_1^2\right)}{v_1 K} \\ 0 & \frac{-2(1-\alpha)\left(\sigma_1^2\right)}{v_1} & \frac{2(1-\alpha)\left(\sigma_1^2\right)}{v_1 K} & \frac{2(1-\alpha)^2(v_1+v_3+1)}{v_1(v_3+1)} \end{bmatrix}.$$

We are interested in a probability matching prior for $t(\boldsymbol{\theta}) = \alpha$. Now

$$\begin{aligned} \nabla_t'(\boldsymbol{\theta}) &= \begin{bmatrix} \frac{\partial t(\boldsymbol{\theta})}{\partial \theta} & \frac{\partial t(\boldsymbol{\theta})}{\partial \sigma_1^2} & \frac{\partial t(\boldsymbol{\theta})}{\partial \sigma_2^2} & \frac{\partial t(\boldsymbol{\theta})}{\partial \alpha} \end{bmatrix} \\ &= \begin{bmatrix} 0 & 0 & 0 & 1 \end{bmatrix}. \end{aligned}$$

Also,

$$\nabla_t'(\boldsymbol{\theta}) F^{-1}(\boldsymbol{\theta}) = \begin{bmatrix} 0 & \frac{-2(1-\alpha)\left(\sigma_1^2\right)}{v_1} & \frac{2(1-\alpha)\left(\sigma_1^2\right)}{v_1 K} & \frac{2(1-\alpha)^2(v_1+v_3+1)}{v_1(v_3+1)} \end{bmatrix}$$

and

$$\nabla_t'(\boldsymbol{\theta}) F^{-1}(\boldsymbol{\theta}) \nabla_t(\boldsymbol{\theta}) = \frac{2(1-\alpha)^2(v_1+v_3+1)}{v_1(v_3+1)}.$$

Define

$$\begin{aligned} \eta'(\boldsymbol{\theta}) &= \frac{\nabla_t'(\boldsymbol{\theta}) F^{-1}(\boldsymbol{\theta})}{\sqrt{\nabla_t'(\boldsymbol{\theta}) F^{-1}(\boldsymbol{\theta}) \nabla_t(\boldsymbol{\theta})}} \\ &= \begin{bmatrix} \eta_1(\boldsymbol{\theta}) & \eta_2(\boldsymbol{\theta}) & \eta_3(\boldsymbol{\theta}) & \eta_4(\boldsymbol{\theta}) \end{bmatrix} \\ &= \begin{bmatrix} 0 & \frac{-2\left(\sigma_1^2\right)\sqrt{v_1(v_3+1)}}{v_1\sqrt{2(v_1+v_3+1)}} & \frac{2\left(\sigma_1^2\right)\sqrt{v_1(v_3+1)}}{v_1 K\sqrt{2(v_1+v_3+1)}} & \frac{2(1-\alpha)(v_1+v_3+1)\sqrt{v_1(v_3+1)}}{v_1(v_3+1)\sqrt{2(v_1+v_3+1)}} \end{bmatrix}. \end{aligned}$$

The prior $\pi(\boldsymbol{\theta})$ is a probability matching prior if and only if the differential equation

$$\sum_{i=1}^{4} \frac{\partial}{\partial \theta_i}\left\{\eta_i(\boldsymbol{\theta})\pi(\boldsymbol{\theta})\right\} = 0$$

$$\frac{\partial}{\partial \sigma_1^2}\left\{\frac{-2(\sigma_1^2)\sqrt{v_1(v_3+1)}}{v_1\sqrt{2(v_1+v_3+1)}}\pi\right\} + \frac{\partial}{\partial \sigma_2^2}\left\{\frac{2(\sigma_1^2)\sqrt{v_1(v_3+1)}}{v_1 K\sqrt{2(v_1+v_3+1)}}\pi\right\} + \frac{\partial}{\partial \alpha}\left\{\frac{2(1-\alpha)(v_1+v_3+1)\sqrt{v_1(v_3+1)}}{v_1(v_3+1)\sqrt{2(v_1+v_3+1)}}\pi\right\} = 0 \quad (5.10)$$

is satisfied. To find a probability matching prior satisfying the differential equation, let

$$\pi^*\left(\theta,\sigma_1^2,\sigma_2^2,\alpha\right)=(1-\alpha)\left(\sigma_1^2\right)\pi\left(\theta,\sigma_1^2,\sigma_2^2,\alpha\right). \tag{5.11}$$

Then Equation 5.10 can be simplified to

$$\frac{\partial}{\partial\sigma_1^2}\left\{\frac{-2\sqrt{v_1(v_3+1)}}{v_1\sqrt{2(v_1+v_3+1)}}\frac{\pi^*}{(1-\alpha)}\right\}+\frac{\partial}{\partial\sigma_2^2}\left\{\frac{2\sqrt{v_1(v_3+1)}}{v_1K\sqrt{2(v_1+v_3+1)}}\frac{\pi^*}{(1-\alpha)}\right\}+\frac{\partial}{\partial\alpha}\left\{\frac{2(v_1+v_3+1)\sqrt{v_1(v_3+1)}}{v_1(v_3+1)\sqrt{2(v_1+v_3+1)}}\frac{\pi^*}{\sigma_1^2}\right\}=0. \tag{5.12}$$

Now $\pi^*=1$ is a solution to Equation 5.12, corresponding to

$$\pi\left(\theta,\sigma_1^2,\sigma_2^2,\alpha\right)=(1-\alpha)^{-1}\left(\sigma_1^2\right)^{-1}.$$

Using the product rule and implicit differentiation it can be shown that Equation 5.12 is a first-order homogeneous and linear partial differential equation given by

$$\frac{-2\sqrt{v_1(v_3+1)}\pi^*_{\sigma_1^2}}{v_1\sqrt{2(v_1+v_3+1)}(1-\alpha)}+\frac{2\sqrt{v_1(v_3+1)}\pi^*_{\sigma_2^2}}{v_1K\sqrt{2(v_1+v_3+1)}(1-\alpha)}+\frac{2(v_1+v_3+1)\sqrt{v_1(v_3+1)}\pi^*_{\alpha}}{v_1(v_3+1)\sqrt{2(v_1+v_3+1)}\sigma_1^2}=0 \tag{5.13}$$

where $\pi^*_{\sigma_1^2}$, $\pi^*_{\sigma_2^2}$ and π^*_{α} denotes the first derivative of π^* with respect to σ_1^2, σ_2^2 and α respectively. The method of characteristics used in Dennemeyer (1968) will be used to solve the partial differential equation. Let

$$\begin{aligned}\frac{-v_1\sqrt{2(v_1+v_3+1)}(1-\alpha)}{2\sqrt{v_1(v_3+1)}}d\sigma_1^2 &= \frac{v_1K\sqrt{2(v_1+v_3+1)}(1-\alpha)}{2\sqrt{v_1(v_3+1)}}d\sigma_2^2\\ -d\sigma_1^2 &= Kd\sigma_2^2.\end{aligned}$$

This is a separable differential equation. Integrating both sides we have

$$\begin{aligned}\int -d\sigma_1^2 &= \int Kd\sigma_2^2\\ -\sigma_1^2+c_1^* &= K\sigma_2^2\\ c_1^* &= K\sigma_2^2+\sigma_1^2.\end{aligned}$$

Now let

$$\begin{aligned}\frac{-v_1\sqrt{2(v_1+v_3+1)}(1-\alpha)}{2\sqrt{v_1(v_3+1)}}d\sigma_1^2 &= \frac{v_1(v_3+1)\sqrt{2(v_1+v_3+1)}(\sigma_1^2)}{2(v_1+v_3+1)\sqrt{v_1(v_3+1)}}d\alpha\\ -(1-\alpha)d\sigma_1^2 &= \frac{(v_3+1)\sigma_1^2}{(v_1+v_3+1)}d\alpha.\end{aligned}$$

This is a separable differential equation therefore moving all terms involving σ_1^2 to the left hand side of the equation and all the terms involving α on the right hand side of the equation and integrating both sides we have

$$\begin{aligned}
\int \frac{d\sigma_1^2}{\sigma_1^2} &= \frac{-(v_3+1)}{(v_1+v_3+1)} \int \frac{1}{(1-\alpha)} d\alpha \\
log\left(\sigma_1^2\right) &= \frac{(v_3+1)}{(v_1+v_3+1)} log\left(1-\alpha\right) + c_2 \\
c_2 &= log\left(\sigma_1^2\right) - log\left(1-\alpha\right)^{\frac{(v_3+1)}{(v_1+v_3+1)}} \\
&= log\left(\frac{\sigma_1^2}{(1-\alpha)^{\frac{(v_3+1)}{(v_1+v_3+1)}}}\right) \\
c_2^* = \exp\{c_2\} &= \frac{\sigma_1^2}{(1-\alpha)^{\frac{(v_3+1)}{(v_1+v_3+1)}}}.
\end{aligned}$$

The solution is therefore $\pi^* = G\left(c_1^*, c_2^*\right) = G\left(K\sigma_2^2 + \sigma_1^2, \frac{\sigma_1^2}{(1-\alpha)^{\frac{(v_3+1)}{(v_1+v_3+1)}}}\right)$ where G is an arbitrary continuously differentiable function. Finally, since

$$\begin{aligned}
\pi^*\left(\theta, \sigma_1^2, \sigma_2^2, \alpha\right) &= (1-\alpha)\left(\sigma_1^2\right)\pi\left(\theta, \sigma_1^2, \sigma_2^2, \alpha\right) \\
G\left(K\sigma_2^2 + \sigma_1^2, \frac{\sigma_1^2}{(1-\alpha)^{\frac{(v_3+1)}{(v_1+v_3+1)}}}\right) &= (1-\alpha)\left(\sigma_1^2\right)\pi\left(\theta, \sigma_1^2, \sigma_2^2, \alpha\right) \\
\pi_{PMP}\left(\theta, \sigma_1^2, \sigma_2^2, \alpha\right) &= G\left(K\sigma_2^2 + \sigma_1^2, \frac{\sigma_1^2}{(1-\alpha)^{\frac{(v_3+1)}{(v_1+v_3+1)}}}\right)(1-\alpha)^{-1}\left(\sigma_1^2\right)^{-1}.
\end{aligned}$$

If we consider the function $G\left(K\sigma_2^2 + \sigma_1^2, \frac{\sigma_1^2}{(1-\alpha)^{\frac{(v_3+1)}{(v_1+v_3+1)}}}\right) = 1$ which is a constant function then we have that the probability matching prior is

$$\pi_{PMP}\left(\theta, \sigma_1^2, \sigma_2^2, \alpha\right) \propto (1-\alpha)^{-1}\left(\sigma_1^2\right)^{-1}. \tag{5.14}$$

□

The following theorem proves that the reference prior is also a probability matching prior.

Theorem 5.3. *The reference prior is also a probability matching prior.*

Proof. To show that the reference prior is also a probability matching prior, the following differential equation needs to be satisfied:

$$\sum_{i=1}^{4} \frac{\partial}{\partial \theta_i} \{\eta_i(\boldsymbol{\theta}) \pi(\boldsymbol{\theta})\} = 0.$$

If the reference prior given by $\pi_R(\alpha, \theta, \sigma_1^2, \sigma_2^2) \propto (K\sigma_2^2 + \sigma_1^2)^{-1} (\sigma_1^2)^{-1} (1-\alpha)^{-1}$ is substituted into this equation and the equation is satisfied then the reference prior is also a probability matching prior.

Now,

$$\frac{\partial}{\partial \sigma_1^2} \left\{ \frac{-2(\sigma_1^2)\sqrt{v_1(v_3+1)}}{v_1\sqrt{2(v_1+v_3+1)}} (K\sigma_2^2 + \sigma_1^2)^{-1} (\sigma_1^2)^{-1} (1-\alpha)^{-1} \right\}$$
$$= \frac{2\sqrt{v_1(v_3+1)}}{v_1\sqrt{2(v_1+v_3+1)}} (K\sigma_2^2 + \sigma_1^2)^{-2} (1-\alpha)^{-1}$$

$$\frac{\partial}{\partial \sigma_2^2} \left\{ \frac{2(\sigma_1^2)\sqrt{v_1(v_3+1)}}{v_1 K\sqrt{2(v_1+v_3+1)}} (K\sigma_2^2 + \sigma_1^2)^{-1} (\sigma_1^2)^{-1} (1-\alpha)^{-1} \right\}$$
$$= \frac{-2\sqrt{v_1(v_3+1)}}{v_1\sqrt{2(v_1+v_3+1)}} (K\sigma_2^2 + \sigma_1^2)^{-2} (1-\alpha)^{-1}$$

and

$$\frac{\partial}{\partial \alpha} \left\{ \frac{2(1-\alpha)(v_1+v_3+1)\sqrt{v_1(v_3+1)}}{v_1(v_3+1)\sqrt{2(v_1+v_3+1)}} (K\sigma_2^2 + \sigma_1^2)^{-1} (\sigma_1^2)^{-1} (1-\alpha)^{-1} \right\} = 0.$$

Therefore, the differential equation is satisfied and the reference prior is also a probability matching prior. □

All the prior distributions considered are proportional to the negative powers of σ_1^2, $(1-\alpha)$ and $(K\sigma_2^2 + \sigma_1^2)$. Therefore a general form can be written as

$$\pi(\theta, \sigma_1^2, \sigma_2^2, \alpha) \propto (\sigma_1^2)^{-a}(1-\alpha)^{-b} (K\sigma_2^2 + \sigma_1^2)^{-c} \tag{5.15}$$

for various non-negative a, b and c. Specifically for $a = b = 1$ and $c = 0$ we have the probability matching prior. For $a = b = c = 1$ we have the reference prior. For $a = \frac{3}{2}$, $b = \frac{1}{2}$ and $c = \frac{1}{2}$ we have the Jeffreys prior and for $a = \frac{3}{4}$, $b = \frac{1}{4}$ and $c = \frac{1}{4}$ then we have the divergence prior.

The posterior for $(\theta,\sigma_1^2,\sigma_2^2,\alpha)$ using the prior in Equation 5.15 is given by

$$\pi(\theta,\sigma_1^2,\sigma_2^2,\alpha|data) \propto (\sigma_1^2)^{-\frac{1}{2}(v_1+v_3+1+2a)}(1-\alpha)^{\frac{1}{2}(v_3+1)-b}\exp\left\{\frac{-\left[IJK\left(\bar{Y}_{...}-\theta\right)^2(1-\alpha)+v_3m_3(1-\alpha)\right]}{2\sigma_1^2}\right\}$$
$$\times\left(K\sigma_2^2+\sigma_1^2\right)^{-c-\frac{v_2}{2}}\exp\left\{-\frac{1}{2}\left[\frac{v_2m_2}{\left(K\sigma_2^2+\sigma_1^2\right)}+\frac{v_1m_1}{\sigma_1^2}\right]\right\}. \quad (5.16)$$

The marginal posterior for α is derived in the following theorem.

Theorem 5.4. *The marginal posterior distribution of Cronbach's α is given by*

$$\begin{aligned}\pi(\alpha|data) &\propto (1-\alpha)^{\frac{v_3}{2}-b}\left[v_1m_1+v_3m_3(1-\alpha)\right]^{-\frac{1}{2}(v_1+v_3+2a-2)} \quad (5.17)\\ &\times P\left(F_{v_1+v_3+2a-2,v_2+2c-2}>\frac{(v_2+2c-2)\left[v_3m_3(1-\alpha)+v_1m_1\right]}{v_2m_2(v_1+v_3+2a-2)}\right) \qquad \text{for } 0<\alpha\le 1+\frac{v_1m_1}{v_3m_3},\end{aligned}$$

where $P\left(F_{v_1+v_3+2a-2,v_2+2c-2}>\frac{(v_2+2c-2)[v_3m_3(1-\alpha)+v_1m_1]}{v_2m_2(v_1+v_3+2a-2)}\right)$ is a probability under the F distribution with numerator degrees of freedom v_1+v_3+2a-2 and denominator degrees of freedom given by v_2+2c-2.

Proof. To determine the marginal posterior for α, the parameters θ, σ_1^2 and σ_2^2 need to be integrated out of the joint posterior. Integrating out the parameter θ results in

$$\begin{aligned}\pi\left(\sigma_1^2,\sigma_2^2,\alpha|data\right) &\propto \int_{-\infty}^{\infty}(\sigma_1^2)^{-\frac{1}{2}(v_1+v_3+1+2a)}(1-\alpha)^{\frac{1}{2}(v_3+1)-b}\exp\left\{-\frac{1}{2}\left[\frac{IJK\left(\bar{Y}_{...}-\theta\right)^2(1-\alpha)}{\sigma_1^2}\right]\right\}\\ &\times\left(K\sigma_2^2+\sigma_1^2\right)^{-c-\frac{v_2}{2}}\exp\left\{-\frac{1}{2}\left[\frac{v_2m_2}{\left(K\sigma_2^2+\sigma_1^2\right)}+\frac{v_1m_1+v_3m_3(1-\alpha)}{\sigma_1^2}\right]\right\}d\theta\\ &\propto (\sigma_1^2)^{-\frac{1}{2}(v_1+v_3+2a)}(1-\alpha)^{\frac{v_3}{2}-b}\left(K\sigma_2^2+\sigma_1^2\right)^{-c-\frac{v_2}{2}}\exp\left\{\frac{-v_2m_2}{2\left(K\sigma_2^2+\sigma_1^2\right)}\right\}\\ &\times\exp\left\{-\frac{1}{2}\left[\frac{v_1m_1+v_3m_3(1-\alpha)}{\sigma_1^2}\right]\right\}.\end{aligned}$$

To determine the marginal posterior for α the remaining parameters σ_1^2 and σ_2^2 need to be integrated out. This is not an easy integral to evaluate. The transformation method using the Jacobian will be applied to determine the constant that makes the joint posterior of σ_1^2 and σ_2^2 integrate to one and knowing this constant will assist in determining the marginal posterior for α. Consider the joint posterior

$$\pi\left(\sigma_1^2,\sigma_2^2|data\right)=z(\sigma_1^2)^{-\frac{1}{2}(v_1+v_3+2a)}\left(K\sigma_2^2+\sigma_1^2\right)^{-c-\frac{v_2}{2}}\exp\left\{-\frac{1}{2}\left[\frac{v_1m_1+v_3m_3(1-\alpha)}{\sigma_1^2}+\frac{v_2m_2}{\left(K\sigma_2^2+\sigma_1^2\right)}\right]\right\}$$

where z is the normalising constant. Let $z^* = 1 + \frac{K\sigma_2^2}{\sigma_1^2}$ and $\sigma_1^2 = \sigma_1^2$. The Jacobian method gives rise to

$$\pi\left(\sigma_1^2, z^*|data\right) = zK^{-1}(\sigma_1^2)^{-\left(\frac{v_1+v_2+v_3}{2}+a+c-1\right)}(z^*)^{-c-\frac{v_2}{2}}\exp\left\{-\frac{1}{\sigma_1^2}\left[\frac{v_1m_1+v_3m_3(1-\alpha)}{2}+\frac{v_2m_2}{2z^*}\right]\right\}.$$

The idea is to integrate $\pi\left(\sigma_1^2, z^*|data\right)$ with respect to σ_1^2 and then integrate $\pi(z^*|data)$ over the entire domain where z^* is defined then z can be determined. Now

$$\begin{aligned}\pi(z^*|data) &= \int_0^\infty zK^{-1}(\sigma_1^2)^{-\left(\frac{v_1+v_2+v_3}{2}+a+c-1\right)}(z^*)^{-c-\frac{v_2}{2}}\exp\left\{-\frac{1}{\sigma_1^2}\left[\frac{v_1m_1+v_3m_3(1-\alpha)}{2}+\frac{v_2m_2}{2z^*}\right]\right\}d\sigma_1^2\\ &= \frac{zK^{-1}(z^*)^{-c-\frac{v_2}{2}}\Gamma\left(\frac{v_1+v_2+v_3+2a+2c-4}{2}\right)}{\left(2^{-1}\left[v_1m_1+v_3m_3(1-\alpha)+v_2m_2(z^*)^{-1}\right]\right)^{\frac{v_1+v_2+v_3+2a+2c-4}{2}}}\\ &= \frac{zK^{-1}2^{\frac{v_1+v_2+v_3+2a+2c-4}{2}}\Gamma\left(\frac{v_1+v_2+v_3+2a+2c-4}{2}\right)(z^*)^{\frac{v_1}{2}+\frac{v_3}{2}+a-2}}{\left(v_2m_2+z^*\left[v_1m_1+v_3m_3(1-\alpha)\right]\right)^{\frac{v_1+v_2+v_3+2a+2c-4}{2}}}\\ &= \frac{zK^{-1}2^{\frac{v_1+v_2+v_3+2a+2c-4}{2}}\Gamma\left(\frac{v_1+v_2+v_3+2a+2c-4}{2}\right)(z^*)^{\frac{v_1}{2}+\frac{v_3}{2}+a-2}(v_2m_2)^{-\frac{1}{2}\left(\frac{v_1+v_2+v_3+2a+2c-4}{2}\right)}}{\left(1+z^*\frac{[v_1m_1+v_3m_3(1-\alpha)]}{v_2m_2}\right)^{\frac{v_1+v_2+v_3+2a+2c-4}{2}}}.\end{aligned}$$

Also,

$$\int_1^\infty \pi(z^*|data)\,dz^* = 1$$

$$\int_1^\infty \frac{zK^{-1}2^{\frac{v_1+v_2+v_3+2a+2c-4}{2}}\Gamma\left(\frac{v_1+v_2+v_3+2a+2c-4}{2}\right)(z^*)^{\frac{v_1}{2}+\frac{v_3}{2}+a-2}(v_2m_2)^{-\frac{1}{2}\left(\frac{v_1+v_2+v_3+2a+2c-4}{2}\right)}dz^*}{\left(1+z^*\frac{[v_1m_1+v_3m_3(1-\alpha)]}{v_2m_2}\right)^{\frac{v_1+v_2+v_3+2a+2c-4}{2}}} = 1$$

$$\int_1^\infty \frac{(z^*)^{\frac{v_1}{2}+\frac{v_3}{2}+a-2}dz^*}{\left(1+z^*\frac{[v_1m_1+v_3m_3(1-\alpha)]}{v_2m_2}\right)^{\frac{v_1+v_2+v_3+2a+2c-4}{2}}} = \frac{z^{-1}K(v_2m_2)^{\frac{1}{2}(v_1+v_2+v_3+2a+2c-4)}}{2^{\frac{v_1+v_2+v_3+2a+2c-4}{2}}\Gamma\left(\frac{v_1+v_2+v_3+2a+2c-4}{2}\right)}$$

Consider the integral $\int_1^\infty \frac{(z^*)^{\frac{v_1}{2}+\frac{v_3}{2}+a-2}dz^*}{\left(1+z^*\frac{[v_1m_1+v_3m_3(1-\alpha)]}{v_2m_2}\right)^{\frac{v_1+v_2+v_3+2a+2c-4}{2}}}$. Let $w = \left(\frac{v_2+2c-2}{v_1+v_3+2a-2}\right)z^*\frac{[v_1m_1+v_3m_3(1-\alpha)]}{v_2m_2}$ then using the substitution rule of integration, it can be shown that

$$P\left(F_{v_1+v_3+2a-2,v_2+2c-2} > \frac{(v_2+2c-2)\left[v_3m_3(1-\alpha)+v_1m_1\right]}{v_2m_2(v_1+v_3+2a-2)}\right) = \frac{z^{-1}K(v_2m_2)^{\frac{1}{2}(v_2+2c-2)}\left[v_1m_1+v_3m_3(1-\alpha)\right]^{\frac{v_1+v_3+2a-2}{2}}}{2^{\frac{v_1+v_2+v_3+2a+2c-4}{2}}\Gamma\left(\frac{v_1+v_3+2a-2}{2}\right)\Gamma\left(\frac{v_2+2c-2}{2}\right)}.$$

Therefore,

$$z = \frac{K\left(v_2m_2\right)^{\frac{1}{2}\left(v_2+2c-2\right)}\left[v_1m_1+v_3m_3\left(1-\alpha\right)\right]^{\frac{v_1+v_3+2a-2}{2}}}{2^{\frac{v_1+v_2+v_3+2a+2c-4}{2}}\Gamma\left(\frac{v_1+v_3+2a-2}{2}\right)\Gamma\left(\frac{v_2+2c-2}{2}\right)P\left(F_{v_1+v_3+2a-2,v_2+2c-2} > \frac{(v_2+2c-2)[v_3m_3(1-\alpha)+v_1m_1]}{v_2m_2(v_1+v_3+2a-2)}\right)}.$$

Since the double integral of $\pi\left(\sigma_1^2,\sigma_2^2|data\right)$ is one, the double integral of the kernel is the reciprocal of z. Therefore

$$\begin{aligned}
\pi\left(\alpha|data\right) &\propto (1-\alpha)^{\frac{v_3}{2}-b}\int_0^\infty\int_0^\infty\left(\sigma_1^2\right)^{-\frac{1}{2}\left(v_1+v_3+2a\right)}\left(K\sigma_2^2+\sigma_1^2\right)^{-c-\frac{v_2}{2}}\exp\left\{\frac{-v_2m_2}{2\left(K\sigma_2^2+\sigma_1^2\right)}\right\} \\
&\times \exp\left\{-\frac{1}{2}\left[\frac{v_1m_1+v_3m_3\left(1-\alpha\right)}{\sigma_1^2}\right]\right\}d\sigma_1^2 d\sigma_2^2 \\
&\propto \frac{(1-\alpha)^{\frac{v_3}{2}-b}\Gamma\left(\frac{v_1+v_3+2a-2}{2}\right)\Gamma\left(\frac{v_2+2c-2}{2}\right)P\left(F_{v_1+v_3+2a-2,v_2+2c-2} > \frac{(v_2+2c-2)[v_3m_3(1-\alpha)+v_1m_1]}{v_2m_2(v_1+v_3+2a-2)}\right)}{K\left(v_2m_2\right)^{\frac{1}{2}\left(v_2+2c-2\right)}\left[v_1m_1+v_3m_3\left(1-\alpha\right)\right]^{\frac{v_1+v_3+2a-2}{2}}} \\
&\propto (1-\alpha)^{\frac{v_3}{2}-b}\left[v_1m_1+v_3m_3\left(1-\alpha\right)\right]^{-\frac{1}{2}\left(v_1+v_3+2a-2\right)} \\
&\times P\left(F_{v_1+v_3+2a-2,v_2+2c-2} > \frac{\left(v_2+2c-2\right)\left[v_3m_3\left(1-\alpha\right)+v_1m_1\right]}{v_2m_2\left(v_1+v_3+2a-2\right)}\right).
\end{aligned}$$

An interesting thing to note is that in Box & Tiao (1973) the following result is given: For $p_1>0$, $p_2>0$, $a_1>0$, $a_2>0$ and $c>0$,

$$\int_0^\infty\int_0^\infty x^{-(p_1+1)}(x+cy)^{-(p_2+1)}\exp\left[-\left(\frac{a_1}{x}+\frac{a_2}{x+cy}\right)\right]dxdy = \frac{\Gamma(p_1)\Gamma(p_2)}{ca_1^{p_1}a_2^{p_2}}I_{\frac{a_2}{a_1+a_2}}(p_2,p_1),$$

where $I_x(p,q)$ is the incomplete beta function. If this result is compared to the double integral above and the incomplete beta integral is expressed in terms of a probability under the F distribution then the result from Box & Tiao (1973) gives the same marginal posterior for α as the one derived in this chapter. □

Notice that in Equation 5.17, the marginal posterior distribution of α depends only on v_1m_1, v_2m_2 and v_3m_3. Therefore, it is only necessary to simulate v_1m_1, v_2m_2 and v_3m_3 for inferences on α. The marginal posterior for α is also not a closed-form expression due to the α parameter appearing in the lower limit of an F integral so it is not easy to determine the normalising constant by hand but it can be simulated or determined using numerical integration. It is also evident from Equation 5.17 that the boundary for α is $0<\alpha\leq 1+\frac{v_1m_1}{v_3m_3}$. Ideally α should be defined between zero and one since it is a reliability coefficient. It is not the case in general for the 3-component model since the variance covariance matrix is not compound symmetric which is a common assumption used for analysis involving Cronbach's alpha. However, when $v_3m_3>v_1m_1$ then Cronbach's alpha will be defined approximately between zero and one. For datasets where $v_3m_3 > v_1m_1$ for the 3-component model, Cronbach's alpha is an approximate reliability coefficient. For the case where $v_3m_3 > v_1m_1$, the marginal posterior for α is given by

$$\pi(\alpha|data) \propto (1-\alpha)^{\frac{v_3}{2}-b}\left[v_1m_1+v_3m_3(1-\alpha)\right]^{-\frac{1}{2}(v_1+v_3+2a-2)} \qquad \text{for } 0<\alpha\leq 1. \tag{5.18}$$

In the case of $v_3m_3 > v_1m_1$, the marginal posterior for α is analytically tractable, making it easy to simulate from it. It can be shown that the mode for this posterior density is given by

$$\alpha_{mode} = 1-\frac{v_1m_1(v_3-2b)}{v_3m_3(v_1+2a-2+2b)}.$$

The mean, variance and median do not have closed forms for this posterior density but they can easily be determined using numerical integration or they can be determined by simulation. The following theorem will be useful for simulating the mean, variance and median from the marginal posterior for α.

Theorem 5.5. *For the transformation $Z=\frac{m_3}{m_1}(1-\alpha)$ the posterior distribution is a truncated F distribution with $df1 = v_3-2b+2$ numerator and $df2 = v_1+2a+2b-4$ denominator degrees of freedom over the interval $0<Z<\frac{m_3}{m_1}$.*

Proof. Let $Z=\frac{m_3}{m_1}(1-\alpha)=g(\alpha)$ then $g^{-1}(Z)=1-\frac{m_1}{m_3}Z$ and $\left|\frac{dg^{-1}(Z)}{dZ}\right|=\frac{m_1}{m_3}$. By the method of transformation we have

$$\begin{aligned}\pi(Z|data) &\propto \left(\frac{m_1}{m_3}Z\right)^{\frac{v_3}{2}-b}\left[\left(v_1m_1+v_3m_3\left(\frac{m_1}{m_3}Z\right)\right)\right]^{-\frac{1}{2}(v_1+v_3+2a-2)}\frac{m_1}{m_3}I_{(0,\frac{m_3}{m_1})}(Z)\\ &\propto Z^{\frac{v_3}{2}-b}\left[(v_1+v_3Z)\right]^{-\frac{1}{2}(v_1+v_3+2a-2)}I_{(0,\frac{m_3}{m_1})}(Z),\end{aligned}$$

where $I_{(0,\frac{m_3}{m_1})}(Z)=\begin{cases}1, & 0<Z<\frac{m_3}{m_1}\\ 0, & \text{otherwise}\end{cases}$ is the indicator function. The posterior density for Z is a truncated F distribution with $df1 = v_3-2b+2$ numerator and $df2 = v_1+2a+2b-4$ denominator degrees of freedom over the interval $0<Z<\frac{m_3}{m_1}$. □

5.3 Simulation Study

A simulation study is done and coverage probabilities are obtained for Cronbach's alpha using the random effects model. The Jeffreys prior, divergence prior, probability matching prior and the reference prior will be considered in the simulation study. The values in the simulation are considered using the following pairs: $(I = 3 \& J = 4)$; $(I = 3 \& J = 8)$; $(I = 10 \& J = 4)$ and $(I = 10 \& J = 8)$ with $K = 2$ for each of these (I,J) pairs and then we considered $K = 3$ for the same (I,J) pairs that were chosen. The variance component values chosen for the simulation study are: $\sigma_1^2 = 1,2,5,7,8,9,10,11,12$, $\sigma_2^2 = 0.2,0.6,0.8,0.9,1,3,4,11,14$ and $\sigma_3^2 = 0.2,2,14,0.4,4$ and 8. The average length and standard deviation of the intervals are also given. If two priors have the same coverage probabilities then the one with the shortest interval length and smallest standard deviation of the interval length is preferable. The number of simulations is 10 000. The average length and standard deviation of the intervals are calculated by using the following formulas, where I_i is the interval length and *num* is the number of intervals:

$$\bar{I} = \frac{1}{num}\sum_{i=1}^{num} I_i$$

and

$$\text{SD}_I = \sqrt{\frac{1}{num-1}\sum_{i=1}^{num}\left(I_i - \bar{I}\right)^2}.$$

The method used for simulating the coverage probabilities is as follows:

1. Since $\frac{v_1 m_1}{\sigma_1^2} \sim \chi^2_{v_1}$ we have that $v_1 m_1 = \sigma_1^2 \chi^2_{v_1}$. Similarly $\frac{v_2 m_2}{\sigma_1^2 + K\sigma_2^2} \sim \chi^2_{v_2}$, it follows that $v_2 m_2 = \chi^2_{v_2}\left(\sigma_1^2 + J\sigma_2^2\right)$ and $\frac{v_3 m_3}{\sigma_1^2 + K\sigma_2^2 + JK\sigma_3^2} \sim \chi^2_{v_3}$ therefore $v_3 m_3 = \chi^2_{v_3}\left(\sigma_1^2 + K\sigma_2^2 + JK\sigma_3^2\right)$.

2. Iterate the process 10 000 times by simulating $\chi^2_{v_1}$, $\chi^2_{v_2}$ and $\chi^2_{v_3}$ values, determining $v_1 m_1$, $v_2 m_2$ and $v_3 m_3$ for each simulated value and substitute these values into the posterior for α in Equation 3.13. When the probability matching prior is used, $a = b = 1$ and $c = 0$, when the reference prior is used then $a = b = c = 1$. When the Jeffreys prior is used, $a = \frac{3}{2}$, $b = \frac{1}{2}$, $c = \frac{1}{2}$ and when the divergence prior is used then $a = \frac{3}{4}$,$b = \frac{1}{4}$ and $c = \frac{1}{4}$.

3. Determine for each simulated posterior a 95% equal tailed credibility interval from the quantiles of the simulated posterior values.

4. The proportion of times that α is contained within these intervals is the coverage probability. The values of alpha considered in the simulation are values of alpha between 0 and 1.

Tables 5.1 and 5.2 summarize the average coverage rates, average interval lengths and the average standard deviation of the interval lengths for the pairs: $(I = 3 \& J = 4)$; $(I = 3 \& J = 8)$; $(I = 10 \& J = 4)$ and $(I = 10 \& J = 8)$ with $K = 2$ for each of these (I,J) pairs. Tables 5.3 and 5.4 gives a summary using the same values in Tables 5.1 and 5.2 but for $K = 3$.

Table 5.1: Average Coverage, Interval Lengths and Standard Deviation of Interval Lengths for Each Prior.

$K = 2$						
	$I = 3$ and $J = 4$			$I = 3$ and $J = 8$		
Prior	Coverage	Length	SD	Coverage	Length	SD
Jeffreys	**0.9339**	0.7287	0.0434	0.8999	0.6407	0.0380
Divergence	0.8918	0.7508	0.0445	0.8508	0.6541	0.0387
Reference	0.9736	**0.7170**	**0.0428**	**0.9599**	**0.6267**	**0.0373**
PMP	0.9758	0.7258	0.0434	0.9602	0.6300	0.0375

For $I = 3$ and $J = 4$, the Jeffreys prior has an average coverage rate closest to 0.95 while the reference and probability matching prior has some over coverage although its values are not far from 0.95 in comparison to the Jeffreys prior. The reference prior has the smallest average interval length and standard deviation of the interval length. The reference prior has an average coverage closest to 0.95 as well as the shortest average interval length and the smallest standard deviation of the interval length compared to the rest of the priors. The probability matching prior has an average coverage rate, average interval length and average standard deviation of the interval length slightly larger than that of the reference prior but the values are quite close, so their performance is similar. The divergence prior has the lowest average coverage rate and largest interval length as well as the standard deviation of the interval length and therefore performed the worst. For $I = 3$ and $J = 8$, the reference prior has an average coverage rate closest to 0.95 as well as the shortest average interval length and the smallest standard deviation of the interval length compared to the rest of the priors. The probability matching prior has an average coverage rate, average interval length and average standard deviation of the interval length slightly larger than that of the reference prior but the values are quite close, so their performance is similar.

Table 5.2: Average Coverage, Interval Lengths and Standard Deviation of Interval Lengths for Each Prior.

$K = 2$						
	$I = 10$ and $J = 4$			$I = 10$ and $J = 8$		
Prior	Coverage	Length	SD	Coverage	Length	SD
Jeffreys	0.9422	0.4542	0.0180	0.9354	0.3588	0.0139
Divergence	0.9281	0.4664	0.0185	0.9207	0.3607	0.0141
Reference	**0.9537**	**0.4475**	**0.0178**	**0.9505**	**0.3532**	**0.0137**
PMP	0.9543	0.4499	0.0179	0.9506	0.3539	0.0138

For $I = 10$ and $J = 4$, the Jeffreys prior also has an average coverage quite close to 0.95, therefore the performance of the reference, probability matching and Jeffreys prior are similar. The divergence prior has the lowest average coverage rate and largest interval length as well as the standard deviation of the interval length and therefore performed the worst. For $I = 10$ and $J = 8$, the reference and probability matching prior have an average coverage closest to 0.95 as well as the shortest average interval length and the smallest standard deviation of the interval length compared to the rest of the priors. The Jeffreys prior also has an average coverage rate quite close to 0.95, therefore the performance of the reference, probability matching and Jeffreys prior are similar. The divergence prior has the lowest average coverage rate and largest interval length as well as the standard deviation of the interval length and therefore performed the worst.

Table 5.3: Average Coverage, Interval Lengths and Standard Deviation of Interval Lengths for Each Prior.

$K = 3$						
	$I = 3$ and $J = 4$			$I = 3$ and $J = 8$		
Prior	Coverage	Length	SD	Coverage	Length	SD
Jeffreys	0.9063	0.6691	0.0403	0.8752	0.5770	0.0346
Divergence	0.8553	0.6835	0.0409	0.8195	0.5868	0.0351
Reference	**0.9661**	**0.6554**	**0.0396**	**0.9502**	**0.5622**	**0.0338**
PMP	0.9667	0.6595	0.0399	**0.9502**	0.5634	0.0340

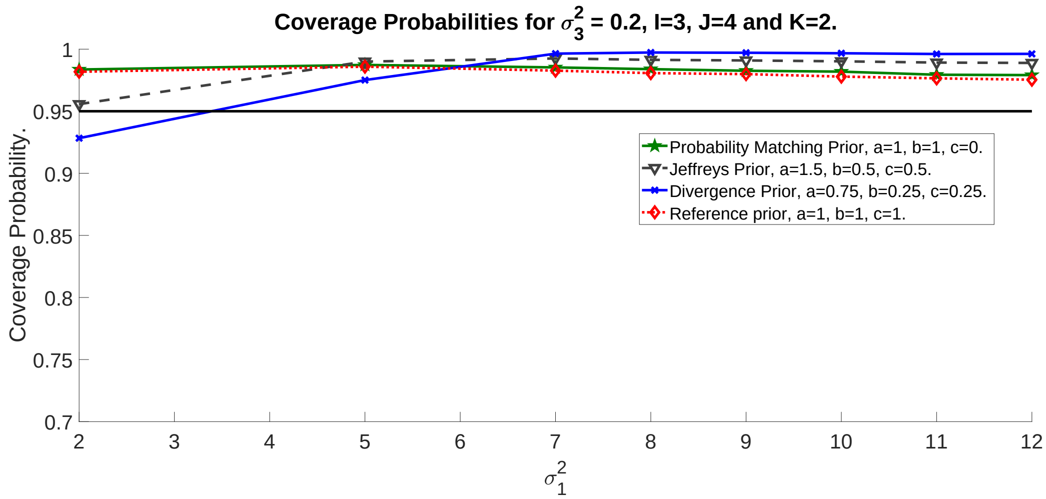
For $I = 3$ and $J = 4$, the reference prior has an average coverage closest to 0.95 as well as the shortest average interval length and the smallest standard deviation of the interval length compared to the rest of the priors. The probability matching prior has an average coverage rate, average interval length and average standard deviation of the interval length slightly larger than that of the reference prior but the values are quite close, so their performance is similar. The divergence prior has the lowest average coverage rate and largest interval length as well as the standard deviation of the interval length and therefore performed the worst. For $I = 3$ and $J = 8$, the reference and probability matching prior have an average coverage rate closest to 0.95 as well as the shortest average interval length and the smallest standard deviation of the interval length compared to the rest of the priors. The divergence prior has the lowest average coverage rate and largest interval length as well as the standard deviation of the interval length and therefore performed the worst.

Table 5.4: Average Coverage, Interval Lengths and Standard Deviation of Interval Lengths for Each Prior.

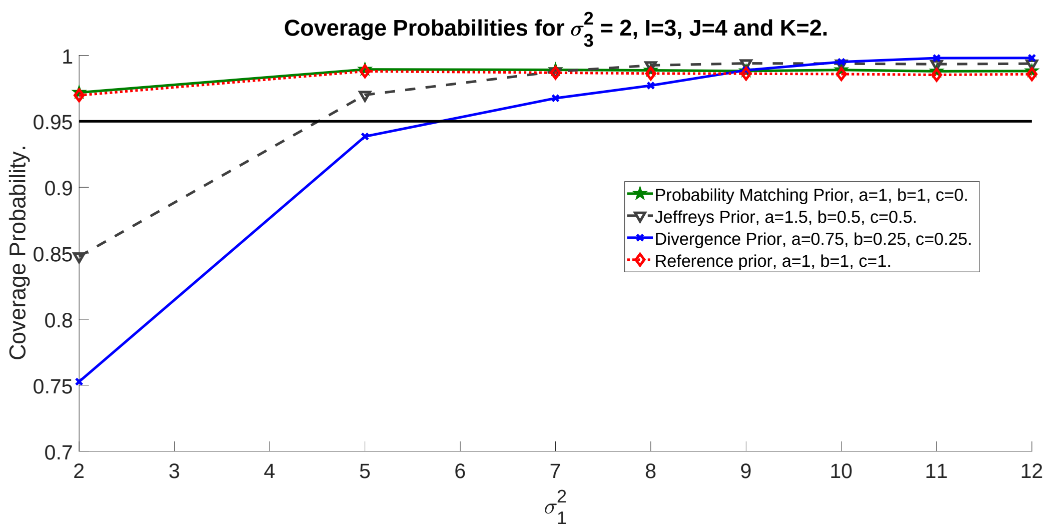
$K = 3$						
	$I = 10$ and $J = 4$			$I = 10$ and $J = 8$		
Prior	Coverage	Length	SD	Coverage	Length	SD
Jeffreys	0.9343	0.3787	0.0152	0.9298	0.2971	0.0117
Divergence	0.9188	0.3859	0.0155	0.9145	0.3013	0.0118
Reference	**0.9504**	**0.3725**	**0.0150**	**0.9481**	**0.2917**	**0.0115**
PMP	0.9507	0.3734	**0.0150**	0.9479	0.2921	**0.0115**

For $I = 10$ and $J = 4$, the reference prior has an average coverage closest to 0.95 as well as the shortest average interval length and the smallest standard deviation of the interval length compared to the rest of the priors. The probability matching prior has an average coverage rate, average interval length and average standard deviation of the interval length slightly larger than that of the reference prior but the values are quite close, so their performance is similar. The divergence prior has the lowest average coverage rate and largest interval length as well as the standard deviation of the interval length and therefore performed the worst. For $I = 10$ and $J = 8$, the reference prior has an average coverage rate closest to 0.95 as well as the shortest average interval length and the smallest standard deviation of the interval length compared to the rest of the priors. The probability matching prior has an average coverage rate, average interval length and average standard deviation of the interval length slightly larger than that of the reference prior but the values are quite close, so their performance is similar. The divergence prior has the lowest average coverage rate and largest average interval length as well as the standard deviation of the interval length and therefore performed the worst.

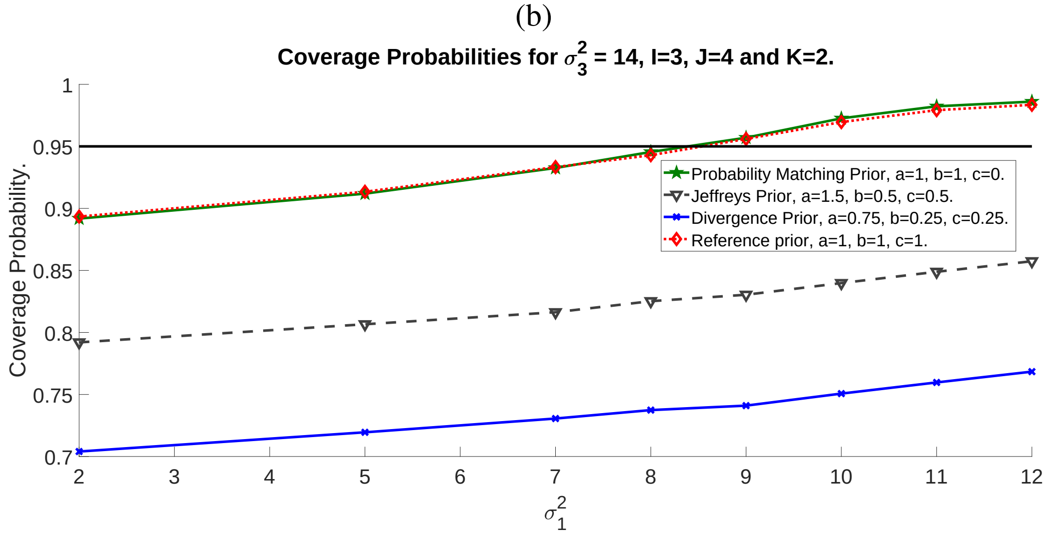
For $K = 2$, Figures 5.1 to 5.2 summarize the coverage rates obtained using each prior for Cronbach's alpha for the pair $(I = 3 \& J = 4)$. Figures 5.3 to 5.4 summarize the coverage rates obtained using each prior for Cronbach's alpha for the pair $(I = 10 \& J = 4)$. For $K = 3$, Figures 5.5 to 5.6 summarize the coverage rates obtained using each prior for Cronbach's alpha for the pair $(I = 3 \& J = 4)$. Figures 5.7 to 5.8 summarize the coverage rates obtained using each prior for Cronbach's alpha for the pair $(I = 10 \& J = 4)$. From Figures 5.1 and 5.2 we see that the reference prior and the probability matching prior perform similarly. Although there is some under-coverage and over-coverage from all four priors, the reference and probability matching prior generally does better than the Jeffreys and divergence prior in estimating the coverage probabilities. The divergence prior consistently performs the worst in terms of coverage rates. From Figures 5.3 and 5.4 we see that for $\sigma_3^2 = 0.2$ and 0.4, the performance of all four priors are similar. For the remainder of σ_3^2 values, the probability matching prior and the reference prior obtains the best coverage rates which are closest to 0.95. The divergence prior performs the worst and has the lowest coverage rates for $\sigma_3^2 = 2, 14, 4$ and 8. From Figures 5.5 and 5.6 we see that for $\sigma_3^2 = 0.2$ and 0.4, except for when $\sigma_1^2 = 2$, the divergence prior performs similarly to the other priors, in fact it slightly outperforms them for $\sigma_1^2 = 5, 7, 8$ and 9. For $\sigma_3^2 = 14, 4$ and 8, the probability matching prior and the reference prior obtains the best coverage rates which are closest to 0.95 although there are cases of under coverage and over coverage. From Figures 5.7 and 5.8 we see that for $\sigma_3^2 = 0.2$ and 0.4, the performance of all four priors are similar. For the remainder of σ_3^2 values, the probability matching prior and the reference prior obtains the best coverage rates which are closest to 0.95. The divergence prior performs the worst and has the lowest coverage rates for $\sigma_3^2 = 2, 14, 4$ and 8.



(a)

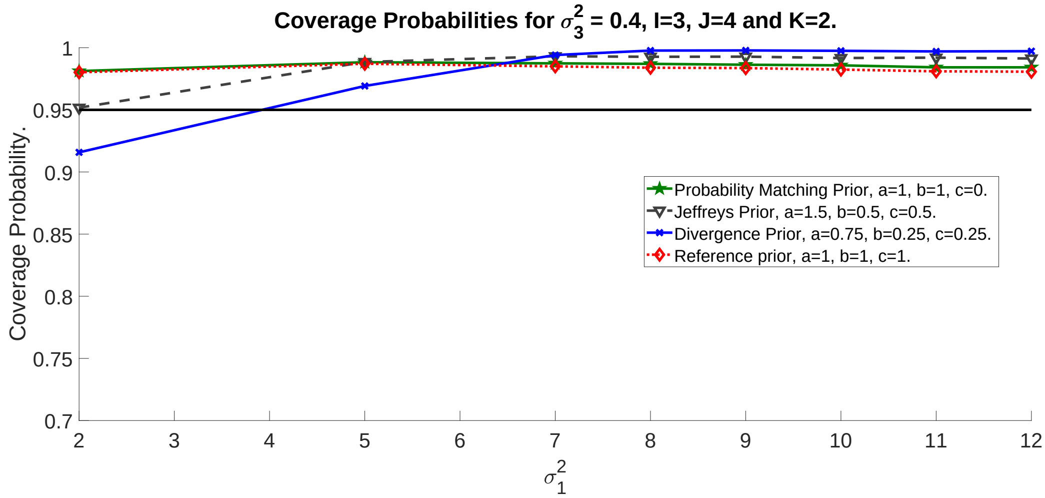


(b)

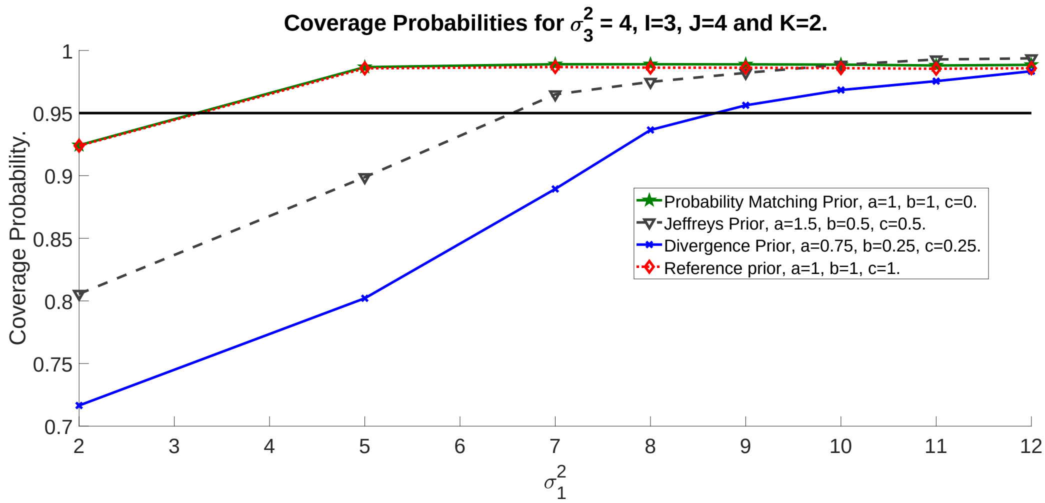


(c)

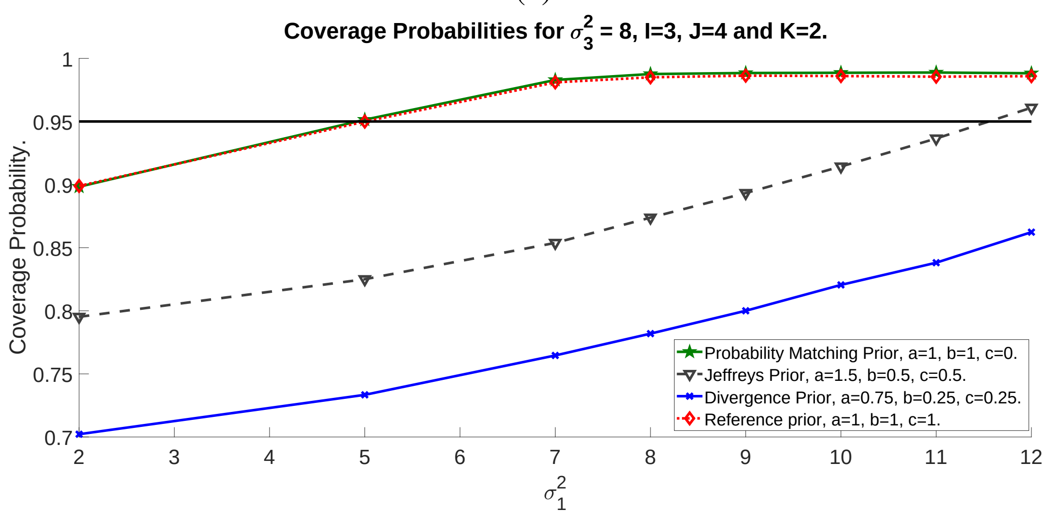
Figure 5.1: Coverage for (a) $\sigma_3^2 = 0.2$ (b) $\sigma_3^2 = 2$, and (c) $\sigma_3^2 = 14$.



(a)

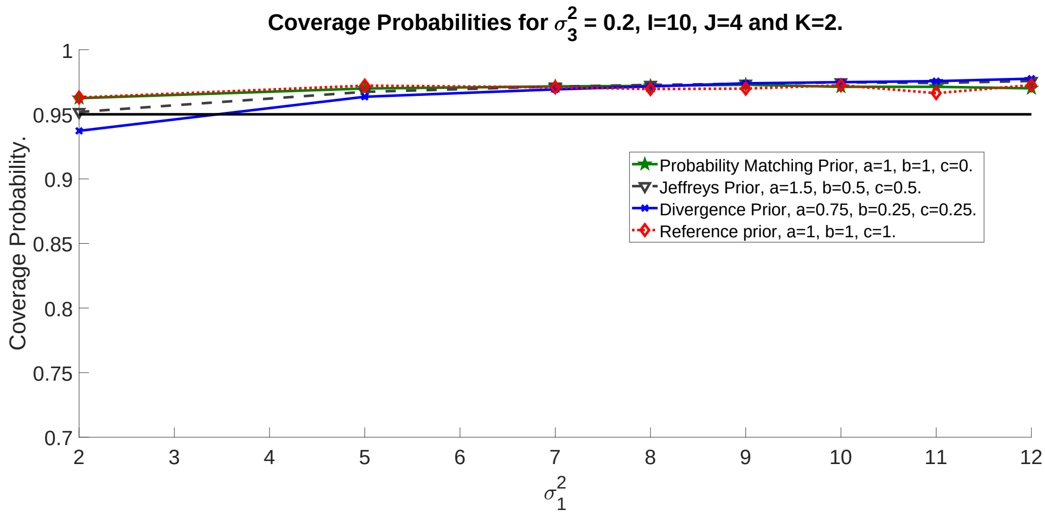


(b)

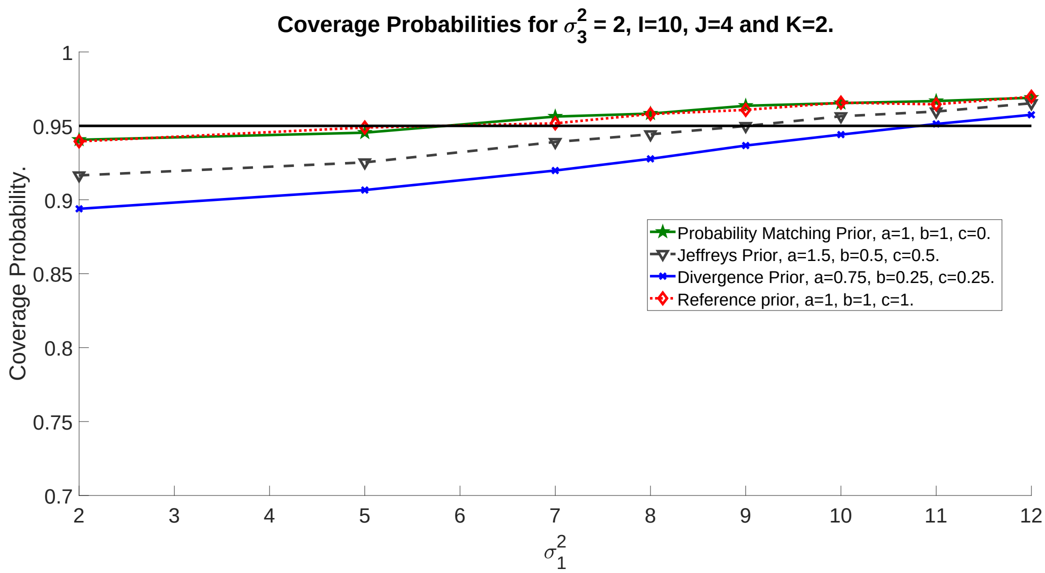


(c)

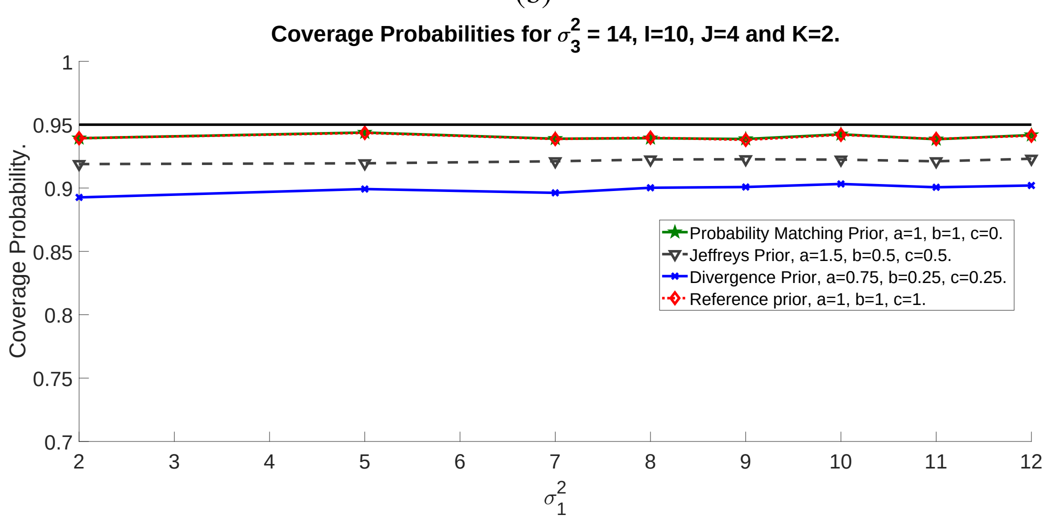
Figure 5.2: Coverage for (a) $\sigma_3^2 = 0.4$ (b) $\sigma_3^2 = 4$, and (c) $\sigma_3^2 = 8$.



(a)

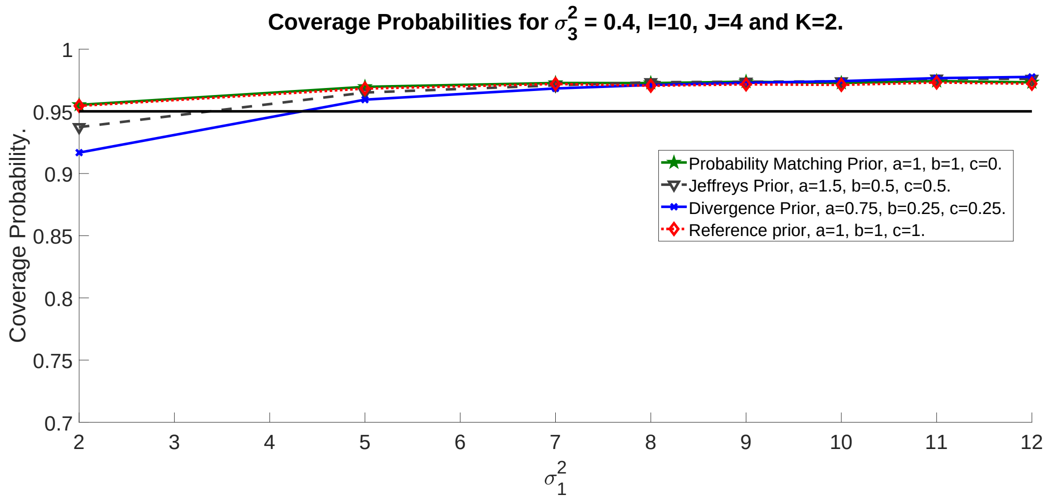


(b)

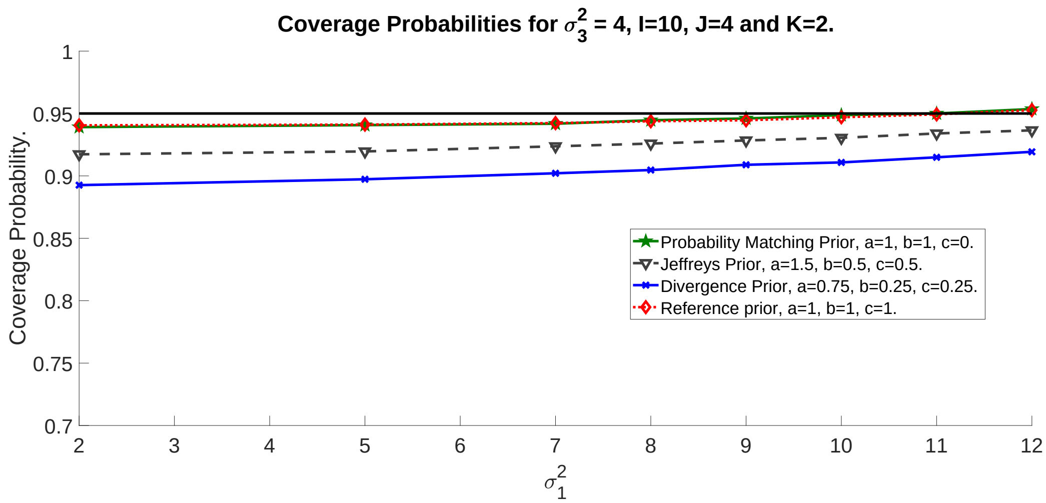


(c)

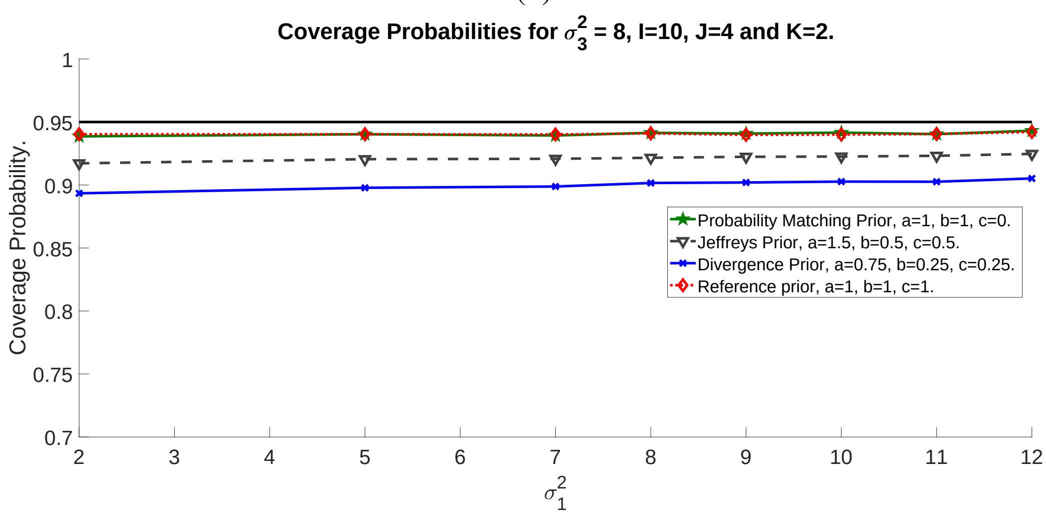
Figure 5.3: Coverage for (a) $\sigma_3^2 = 0.2$ (b) $\sigma_3^2 = 2$, and (c) $\sigma_3^2 = 14$.



(a)

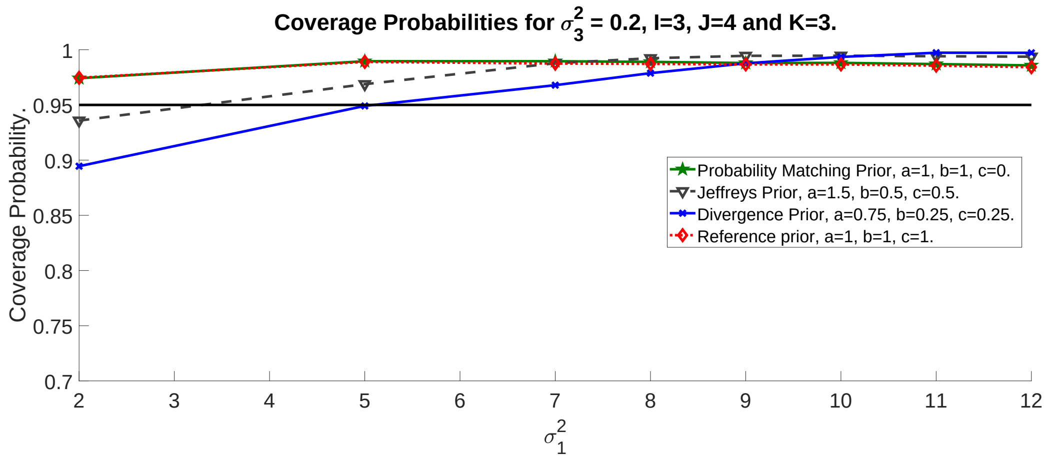


(b)

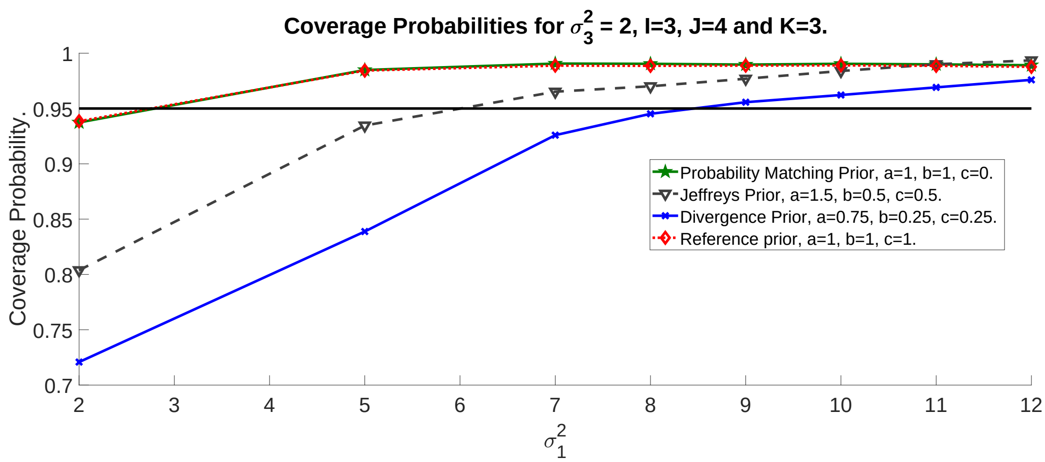


(c)

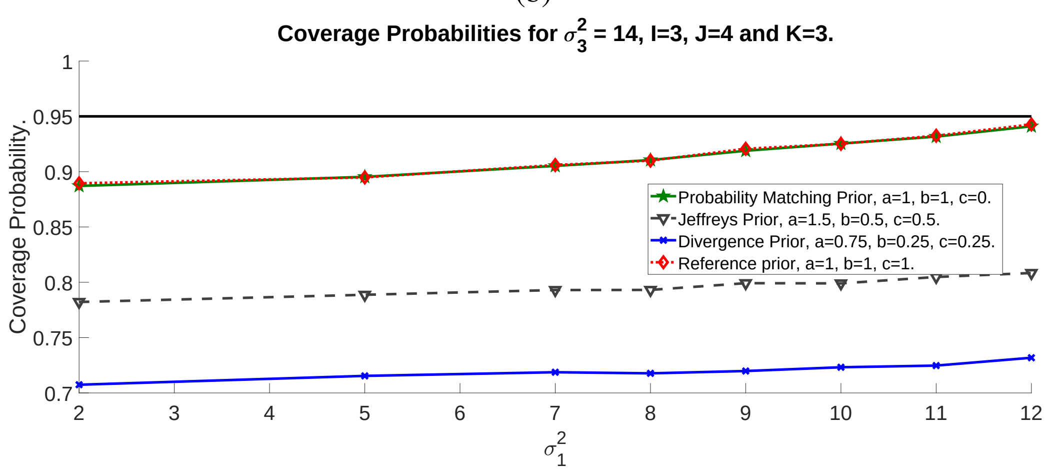
Figure 5.4: Coverage for (a) $\sigma_3^2 = 0.4$ (b) $\sigma_3^2 = 4$, and (c) $\sigma_3^2 = 8$.



(a)

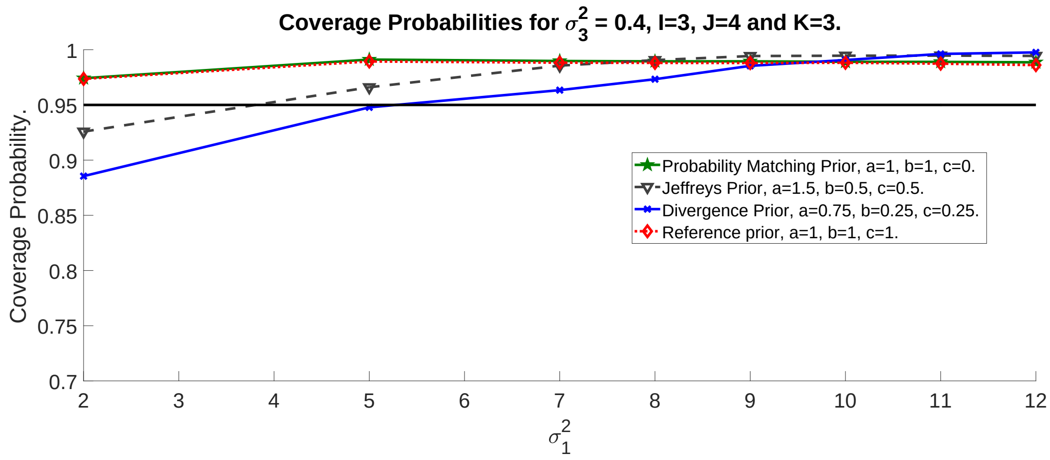


(b)

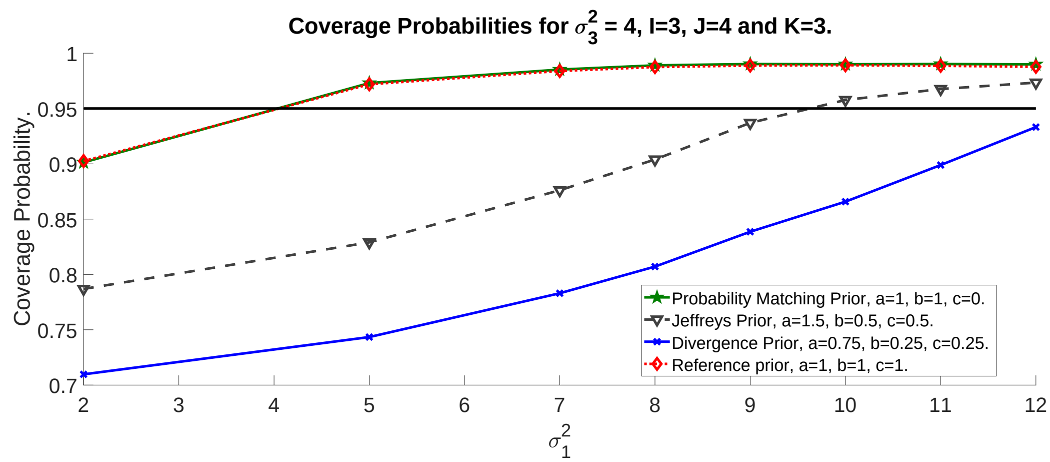


(c)

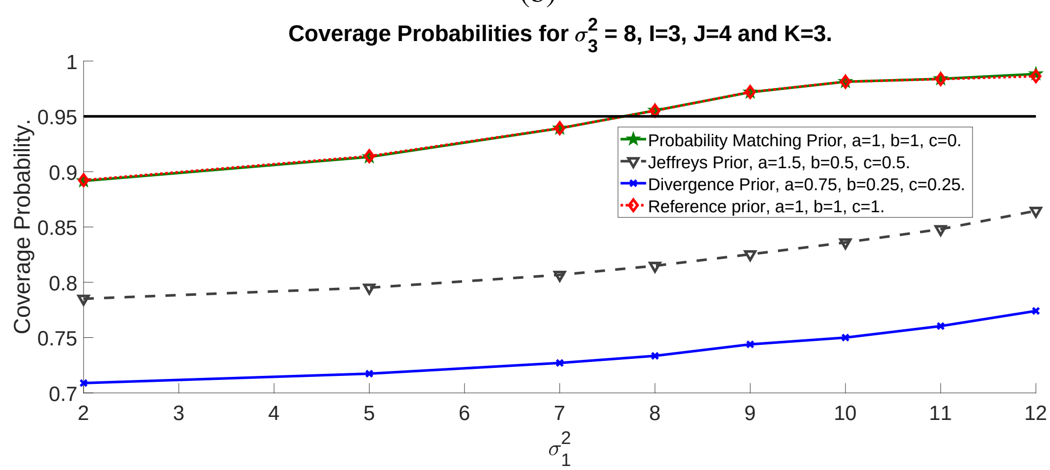
Figure 5.5: Coverage for (a) $\sigma_3^2 = 0.2$ (b) $\sigma_3^2 = 2$, and (c) $\sigma_3^2 = 14$.



(a)

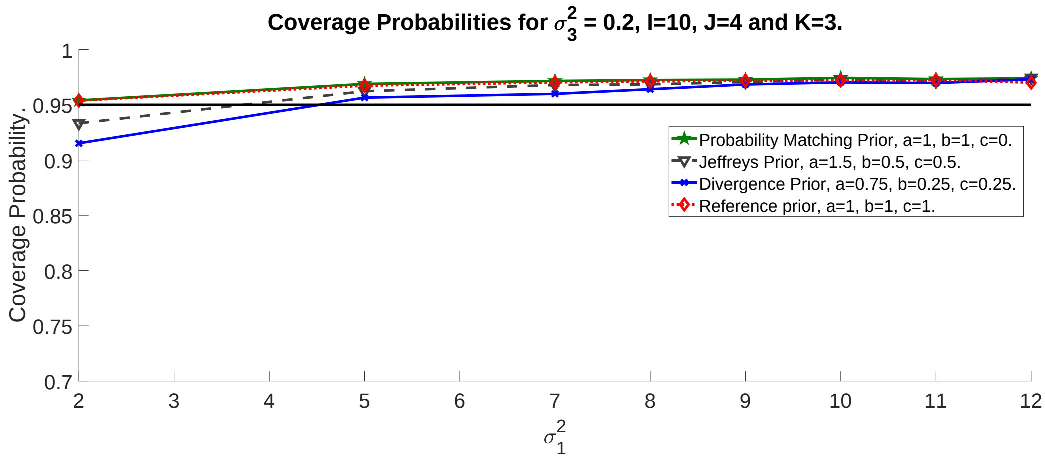


(b)

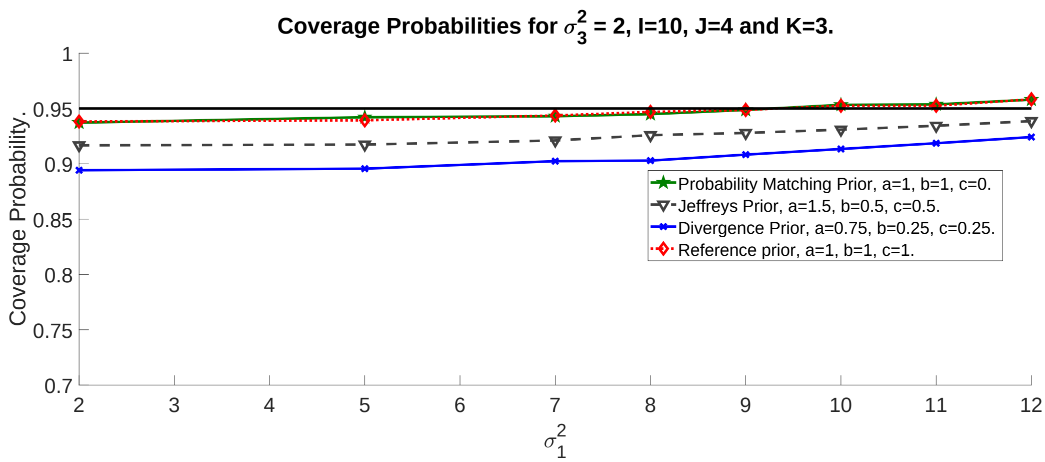


(c)

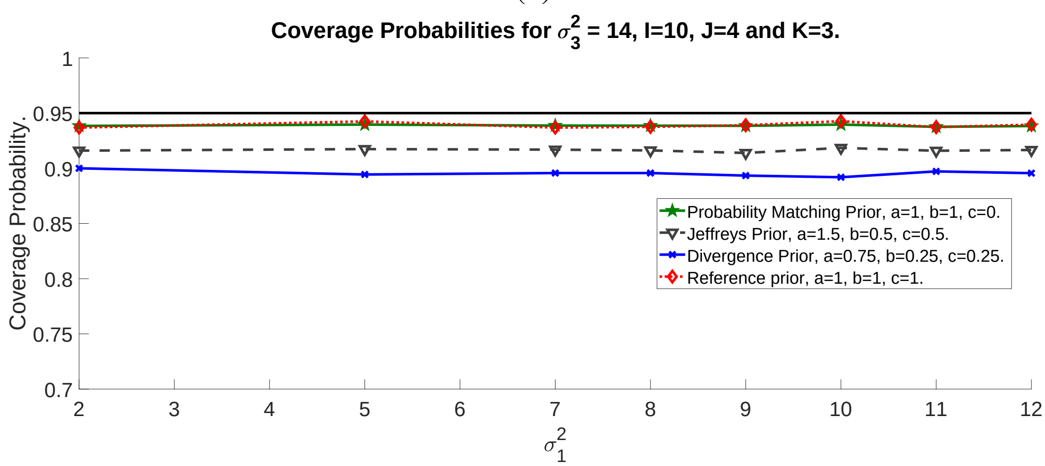
Figure 5.6: Coverage for (a) $\sigma_3^2 = 0.4$ (b) $\sigma_3^2 = 4$, and (c) $\sigma_3^2 = 8$.



(a)

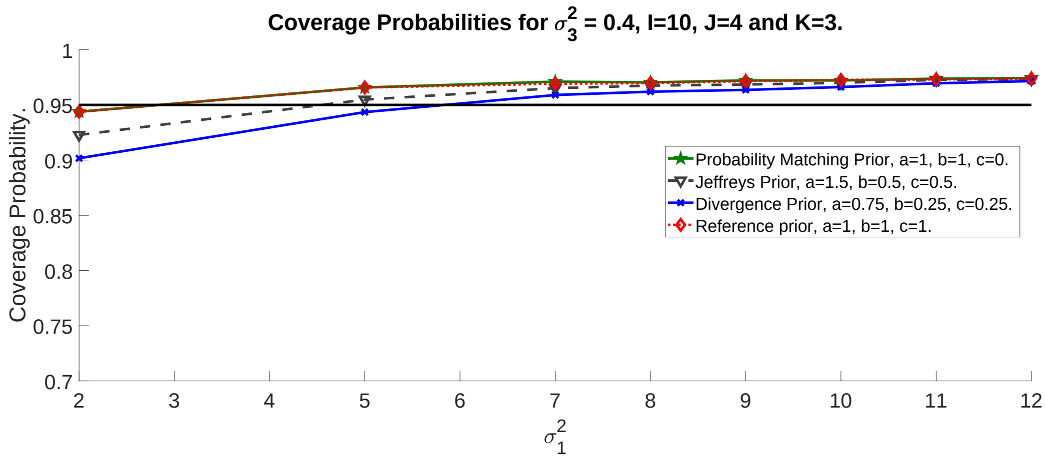


(b)

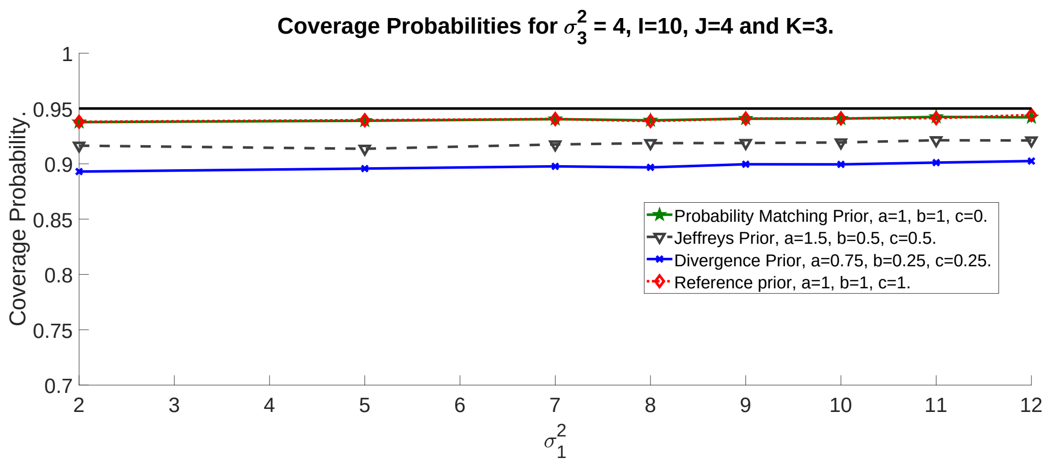


(c)

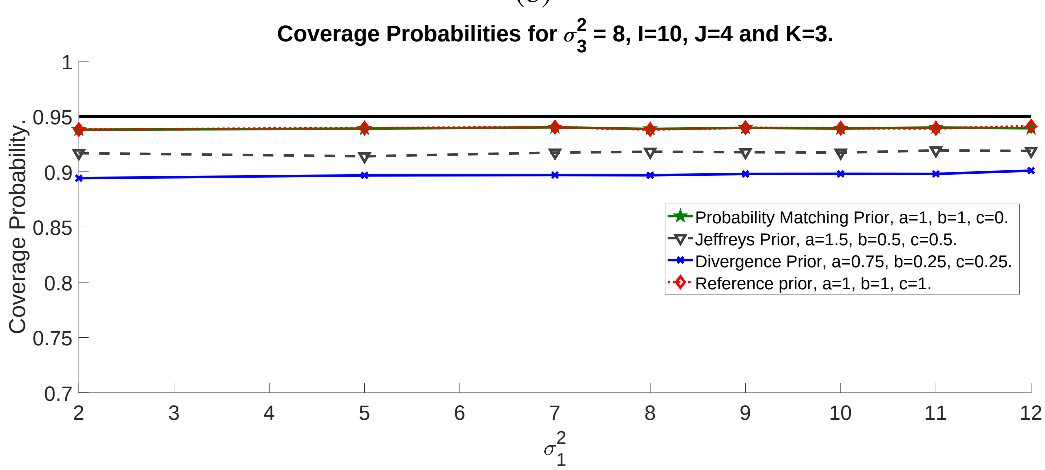
Figure 5.7: Coverage for (a) $\sigma_3^2 = 0.2$ (b) $\sigma_3^2 = 2$, and (c) $\sigma_3^2 = 14$.



(a)



(b)



(c)

Figure 5.8: Coverage for (a) $\sigma_3^2 = 0.4$ (b) $\sigma_3^2 = 4$, and (c) $\sigma_3^2 = 8$.

5.4 Examples

5.4.1 Example 1

The following is a simulated example used in Box & Tiao (1973). The data were generated using the 3-component hierarchical model for $\theta = 0$, $\sigma_1^2 = 1$, $\sigma_2^2 = 4$, $\sigma_3^2 = 2.25$, $I = 10$ and $J = K = 2$. The data are given in Table 5.5.

Table 5.5: Simulated Data from Box & Tiao (1973).

	$J=1$		$J=2$	
I	$K=1$	$K=2$	$K=1$	$K=2$
1	2.004	2.713	0.603	0.252
2	4.342	4.229	3.344	3.057
3	0.869	-2.621	-3.896	-3.696
4	3.531	4.185	1.722	0.380
5	2.579	4.271	-2.101	0.651
6	-1.404	-1.003	-0.775	-2.202
7	-1.676	-0.208	-9.139	-8.653
8	1.670	2.426	1.834	1.200
9	2.141	3.527	0.462	0.665
10	-1.229	-0.596	4.471	1.606

From the data, we obtained $v_1m_1 = \sum_{i=1}^{I}\sum_{j=1}^{J}\sum_{k=1}^{K}(Y_{ijk} - \bar{Y}_{ij.})^2 = 20.8696$ and $v_2m_2 = K\sum_{i=1}^{I}\sum_{j=1}^{J}(\bar{Y}_{ij.} - \bar{Y}_{i..})^2 = 122.9290$ and $v_3m_3 = KJ\sum_{i=1}^{I}(\bar{Y}_{i..} - \bar{Y}_{...})^2 = 240.2827$. Since $v_3m_3 > v_1m_1$, the marginal posterior for α in Equation 5.18 will be used to obtain results for Cronbach's alpha. From the results of the simulation studies involving the coverage probabilities, it is recommended that the reference prior and probability matching prior be used to analyse this data. The exact and approximate posterior density functions for α using the probability matching prior and reference prior are given in Figures 5.9 and 5.10 respectively. From Figures 5.9 and 5.10, it is clear that the exact posterior which was plotted using Equation 3.13 and the approximate posterior which was plotted using Equation 5.18 are exactly the same. This is because $v_3m_3 > v_1m_1$ for this data and the $P\left(F_{v_1+v_3+2a-2,v_2+2c-2} > \frac{(v_2+2c-2)[v_3m_3(1-\alpha)+v_1m_1]}{v_2m_2(v_1+v_3+2a-2)}\right)$ term in Equation 3.13 tends to one in this case which results in the exact posterior and approximate posterior being equal. It is also evident that the marginal posterior for α using the probability matching prior is the same as the marginal posterior using the reference prior. This is because the constant c does not appear in the marginal posterior for α in Equation 5.18. Therefore since the values of a and b are both equals to one for these two priors, the marginal posteriors using these priors are the same and any statistics obtained from the marginal posterior for α will be the same for both priors. Important statistics obtained from the posterior distribution of α are given in Table 5.6.

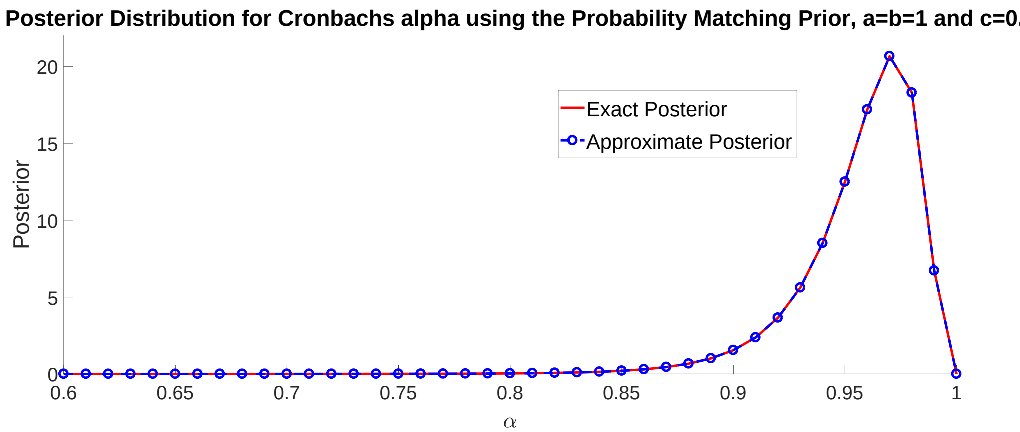


Figure 5.9: Posterior for α Using the Probability Matching Prior.

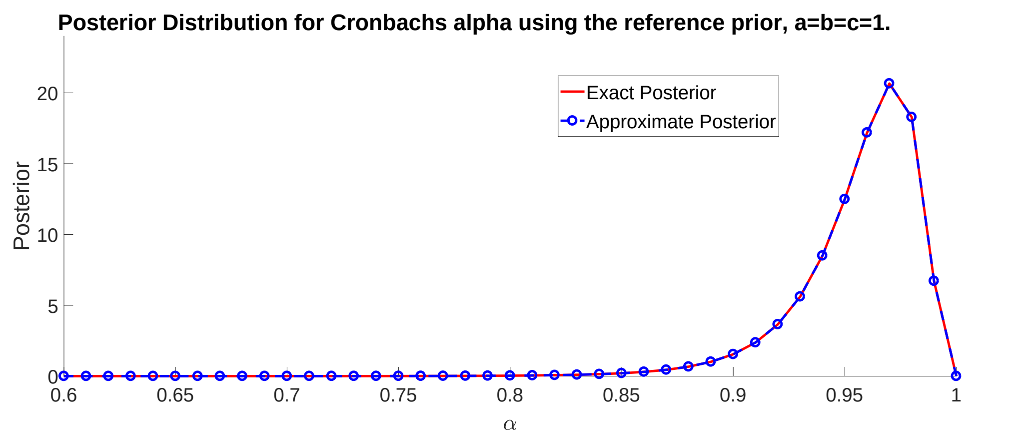


Figure 5.10: Posterior for α Using the Reference Prior.

Table 5.6: Central Values, Variances and Credibility Intervals for the Posterior Distributions of α for the Paste Data Using the Reference and Probability Matching Prior.

Mean	Median	Mode	Variance	90% CI	95% CI
0.9566	0.9625	0.9724	0.0007	(0.9065-0.9867)	(0.8891-0.9893)

5.4.2 Example 2

The following data are from Davies & Goldsmith (1972). They reported data from an experiment designed to measure batch-to-batch and cask-to-cask variation in the manufacture of a chemical paste product. The paste is shipped in batches and delivered in casks and it is suspected that there are variations in the mean strength of the paste between delivery batches. There are ten randomly selected delivery batches, from each batch, three casks are selected at random and two analyses are performed on the contents of each selected cask. The data comprising the percentage paste strength in the analysed sample are given in the following table:

Table 5.7: Paste Data from Davies & Goldsmith (1972).

	$J=1$		$J=2$		$J=3$	
I	$K=1$	$K=2$	$K=1$	$K=2$	$K=1$	$K=2$
1	62.8	62.6	60.1	62.3	62.7	63.1
2	60.0	61.4	57.5	56.9	61.1	58.9
3	58.7	57.5	63.9	63.1	65.4	63.7
4	57.1	56.4	56.9	58.6	64.7	64.5
5	55.1	55.1	54.7	54.2	58.8	57.5
6	63.4	64.9	59.3	58.1	60.5	60.0
7	62.5	62.6	61.0	58.7	56.9	57.7
8	59.2	59.4	65.2	66.0	64.8	64.1
9	54.8	54.8	64.0	64.0	57.7	56.8
10	58.3	59.3	59.2	59.2	58.9	56.6

From the data, we obtained $v_1m_1 = \sum_{i=1}^{I}\sum_{j=1}^{J}\sum_{k=1}^{K}(Y_{ijk} - \bar{Y}_{ij.})^2 = 20.3400$ and $v_2m_2 = K\sum_{i=1}^{I}\sum_{j=1}^{J}(\bar{Y}_{ij.} - \bar{Y}_{i..})^2 = 350.9067$ and $v_3m_3 = KJ\sum_{i=1}^{I}(\bar{Y}_{i..} - \bar{Y}_{...})^2 = 247.4027$. Since $v_3m_3 > v_1m_1$, the marginal posterior for α in Equation 5.18 will be used to obtain results for Cronbach's alpha. From the results of the simulation studies involving the coverage probabilities, it is recommended that the reference prior and probability matching prior be used to analyse this data. The exact and approximate posterior density functions of α using the probability matching prior and reference prior are given in Figures 5.11 and 5.12, respectively.

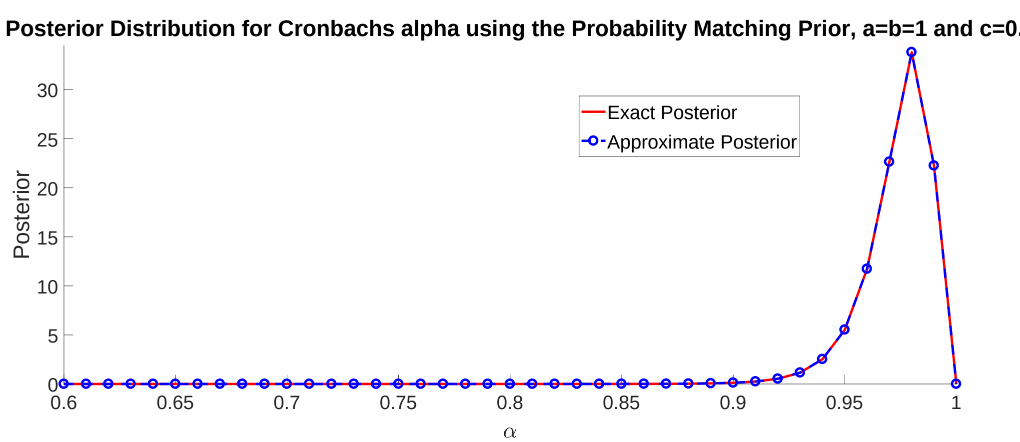


Figure 5.11: Posterior for α Using the Probability Matching Prior.

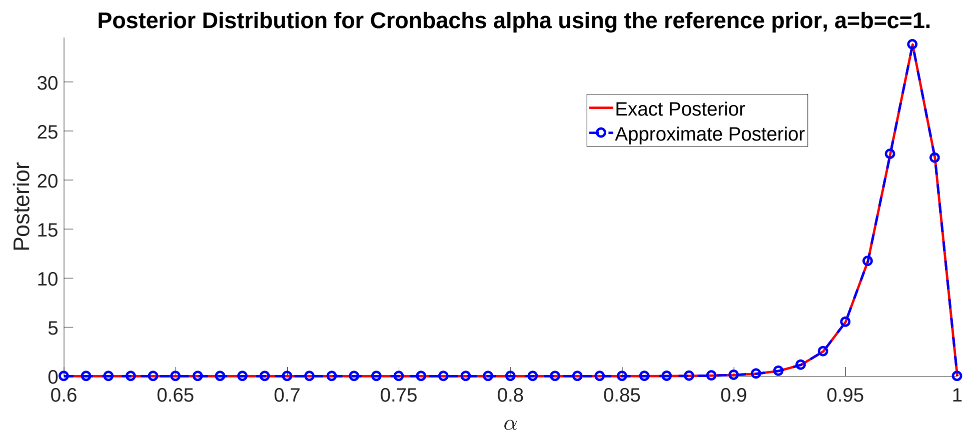


Figure 5.12: Posterior for α Using the Reference Prior.

From Figures 5.11 and 5.12, it is clear that the exact posterior which was plotted using Equation 3.13 and the approximate posterior which was plotted using Equation 5.18 are the same and the marginal posterior for α using the probability matching prior is the same as the marginal posterior using the reference prior. Important statistics obtained from the posterior distribution of α are given in Table 5.8.

Table 5.8: Central Values, Variances and Credibility Intervals for the Posterior Distributions of α for the Paste Data Using the Reference and Probability Matching Prior.

Mean	Median	Mode	Variance	90% CI	95% CI
0.9736	0.9766	0.9820	0.0002	(0.9455-0.9914)	(0.9365-0.9931)

The high reliability values implies that there is good consistency/reliability from batch to batch and cask to cask.

5.5 Conclusion

In this chapter, a number of objective priors for Cronbach's alpha in the case of a balanced three-component hierarchical model, have been derived. The reference prior for the following group orderings: $\{\alpha, \sigma_1^2, \sigma_2^2, \theta\}$, $\{\alpha, \sigma_2^2, \sigma_1^2, \theta\}$, $\{\alpha, \theta, \sigma_1^2, \sigma_2^2\}$, $\{\alpha, \theta, \sigma_2^2, \sigma_1^2\}$, were considered. It was shown that the reference prior for the group orderings $\{\alpha, \sigma_1^2, \sigma_2^2, \theta\}$, $\{\alpha, \sigma_1^2, \theta, \sigma_2^2\}$,$\{\alpha, \theta, \sigma_1^2, \sigma_2^2\}$, were equivalent to the probability matching prior when Cronbach's alpha is the parameter of interest. The probability matching prior that we have used when $c = 0$ when α is the parameter of interest is different from the reference prior that was obtained in this chapter. We compared the probability matching prior, Jeffreys prior and divergence prior using a simulation study involving their coverage probabilities. The simulation study showed that the reference prior and the probability matching prior outperformed the Jeffreys prior and the divergence prior in terms of the coverage rates obtained. The coverage rates obtained for the reference prior were the same as those for the probability matching prior. The mathematics showed that they both satisfy the condition to be a probability matching prior therefore it is expected that the coverage rates obtained using these priors should be similar. The divergence prior performed the worst in terms of its coverage rates which were either larger than the nominal rate or smaller than the nominal rate of 0.95. It is recommended that the reference prior and the probability matching prior should be used for the Bayesian analysis of Cronbach's alpha for the three-component hierarchical model.

Chapter 6

Bayesian Process Control for Cronbach's Alpha

6.1 Introduction

In this chapter statistical process control limits will be obtained for Cronbach's coefficient alpha in the case of the balanced one-way random effects model. This will be achieved by deriving a predictive distribution of a future Cronbach's alpha coefficient using Bayesian statistics. The prior that will be considered in the Bayesian analysis is the Jeffreys independence prior from Box & Tiao (1973). The predictive distribution will also be used to determine the distribution of the run-length and the average run-length.

While statistical process control (SPC) techniques such as control charts have not been widely used in applications of Cronbach's alpha (α), they could be a useful method to monitor if the reliability of a test remains stable across different samples or time points. If there is variation in alpha across different samples or time points SPC could be helpful to identify whether the variability is due to natural causes or special causes.

Since Bayesian statistical process control (SPC) techniques applied to Cronbach's alpha are even less common than classical (traditional) SPC methods and since Cronbach's coefficient alpha is one of the most important statistics in the research field of reliability, Bayesian statistical process control procedures will be developed on alpha.

In Section 6.2 an introduction to Bayesian statistical process control is given and in Section 6.3, the model, the choice of prior and the posterior distribution of α are derived. In Section 6.4 an example is given and it is shown that by using numerical integration or the Rao-Blackwell simulation procedure that the predictive density function of a future (unseen) Cronbach's alpha can be obtained. In Section 6.5 the run-length and average run-length are discussed. The predictive density function of a future run-length is the average of a large number of geometric distributions each with its own parameter value. In Section 6.6 the posterior and predictive densities for a larger example are calculated and in

Section 6.7 the "In Control" and "Out of Control" situations are considered. In Section 6.8 an example of measured values of "Bore diameter" is analysed. Section 6.9 is the conclusion section.

6.2 Statistical Process Control

Statistical process control (SPC) refers to statistical procedures and problem solving methods used to control and monitor the quality of the output of a production process. The aim of SPC is to detect and to eliminate uncontrolled variation in the process. For more information see for example Balakrishnan et al. (2006), Montgomery (2005) and Human (2009). Statistical process control is usually implemented in two phases. Control limits are established during Phase I, and in Phase II, the process is monitored to detect any breaches of these limits. It is indicated in Van Zyl & Van der Merwe (2019) that the conventional approach to statistical process control is frequentist in nature, where unknown parameters are estimated using maximum likelihood estimates in a Phase I study, and these estimates are then used for modelling distributions in Phase II. The Bayesian framework, postulates however, that assuming the true underlying parameter values can be derived in such a way, is unsatisfactory. Instead by using a prior distribution, the posterior distributions of the unknown parameters are derived. These posterior distributions illustrate the uncertainties in the parameter values and these uncertainties are incorporated through predictive distributions.

A number of Bayesian process monitoring schemes are based on predictive distributions. See for example Menzefricke (2002, 2007, 2010a, 2010b) and Van Zyl & Van der Merwe (2019). Also, in Tsiamyrtzis & Hawkins (2006) the following is mentioned "A particularly interesting feature in the Bayesian paradigm is forecasting. Namely one can use the available data to derive the predictive distribution of the next (unseen) observation". In this chapter it will be the next (unseen) Cronbach's alpha $(\hat{\alpha}_f)$. Given a stable Phase I process the predictive distribution $\left(f\left(\hat{\alpha}_f|\text{data}\right)\right)$ will be used to calculate central values, variances, prediction intervals, control limits, the run-length and the average run-length of a future Cronbach's alpha. Bayarri & García-Donato (2005) highlighted the following reasons for recommending a Bayesian analysis for control charts: (i) Bayesian methods allow naturally for prediction and control charts rely on future observations. (ii) Objective Bayesian procedures are possible without introduction of any other information other than the model. (iii) The numerical difficulties of a Bayesian procedure are easily handled via Monte Carlo simulation. By using Monte Carlo simulation and the Rao-Blackwell procedure or numerical integration the unconditional predictive distribution of a future Cronbach's coefficient, $\hat{\alpha}_f$, will be obtained. The advantage of a Bayesian approach to process monitoring arise from the sequential nature of Bayes theorem. As pointed out by Tsiamyrtzis & Hawkins (2006), Shiau & Feltz (2006), Alt (2006), Tagaras & Nenes (2006) and Graves (2006), a Bayesian approach allows a more flexible framework in particular with respect to the usual assumption made in SPC charts about known parameters. The above mentioned authors considered both univariate and multivariate process monitoring techniques. Applications and development of full

Bayesian approaches and empirical Bayes methods are discussed by them.

The researchers Peterson (2006) and Moreno (2006) focussed on Bayesian methods for process optimization. According to them the predictive approach to response surface optimization represented a major advance in response surface method techniques as it incorporates the uncertainty of the parameter estimates in the optimization process. They also mentioned that this has no frequentist counterpart.

6.3 The Balanced Random Effects Model

6.3.1 The Model

The model that will be used is the balanced one-way random effects (variance components) model, previously considered in Chapters 3 and 4, and is given by

$$Y_{ij} = \theta + r_i + \varepsilon_{ij} \quad \text{for } i = 1, \ldots I \text{ and } j = 1, \ldots J, \tag{6.1}$$

where Y_{ij} is the value corresponding to the j^{th} observation made at the i^{th} group (sample). θ is a constant referred to as the overall mean and it is unknown. The r_i and ε_{ij} are independent normal variables with zero means and variances, which will be denoted in this chapter by σ_r^2 and σ_ε^2 respectively. Let $\mathbf{Y} = [Y_{i1}, Y_{i2}, \ldots Y_{iJ}]'$. It was shown in Chapter 3 that $Var\left(\mathbf{Y}|\theta, \sigma_\varepsilon^2, \sigma_r^2\right) = \sigma_\varepsilon^2 \tilde{\mathbf{I}} + \mathbf{1}\mathbf{1}'\sigma_r^2 = \Sigma$ where $\tilde{\mathbf{I}}$ is the $J \times J$ identity matrix and $\mathbf{1} = [1\,1 \ldots 1]'$ is a $J \times 1$ column vector of ones. The model in Equation 6.1 is called the balanced one-way random effects model because the number of observations, J in each sample are the same.

$$\Sigma = \begin{bmatrix} \sigma_\varepsilon^2 + \sigma_r^2 & \sigma_r^2 & \sigma_r^2 & \ldots & \sigma_r^2 \\ \sigma_r^2 & \sigma_\varepsilon^2 + \sigma_r^2 & \sigma_r^2 & \ldots & \sigma_r^2 \\ \sigma_r^2 & \ldots & \sigma_\varepsilon^2 + \sigma_r^2 & \cdots & \sigma_r^2 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \sigma_r^2 & \sigma_r^2 & \sigma_r^2 & \cdots & \sigma_\varepsilon^2 + \sigma_r^2 \end{bmatrix}_{J \times J}.$$

The covariance matrix is called compound symmetric since all the variances along the diagonal are the same, namely $\sigma_\varepsilon^2 + \sigma_r^2$ and all the covariances σ_r^2 are equal. A general definition for α is

$$\alpha = \frac{J}{(J-1)} \left\{ 1 - \frac{trace\,(\Sigma)}{\mathbf{1}'\Sigma\mathbf{1}} \right\}$$

where $trace\,(\Sigma)$ is the sum of the diagonal elements (variances) of the covariance matrix. For the random effects model

$$\alpha = \frac{J}{(J-1)}\left\{1-\frac{J\left(\sigma_\varepsilon^2+\sigma_r^2\right)}{J\left(\sigma_\varepsilon^2+\sigma_r^2\right)+J(J-1)\sigma_r^2}\right\} = \frac{J\sigma_r^2}{\sigma_\varepsilon^2+J\sigma_r^2} = 1-\frac{\sigma_\varepsilon^2}{\sigma_\varepsilon^2+J\sigma_r^2}. \tag{6.2}$$

6.3.2 The Prior and Posterior Distribution of α

In Box & Tiao (1973) it is shown that the likelihood function for the model defined in Equation 6.1 is

$$\ell(\theta,\sigma_\varepsilon^2,\sigma_r^2|data) \propto \left(\sigma_\varepsilon^2\right)^{-v_1/2}\left(\sigma_\varepsilon^2+J\sigma_r^2\right)^{-(v_2+1)/2}\exp\left\{-\frac{1}{2}\left[\frac{IJ\left(\bar{Y}_{..}-\theta\right)^2}{\left(\sigma_\varepsilon^2+J\sigma_r^2\right)}+\frac{v_2m_2}{\left(\sigma_\varepsilon^2+J\sigma_r^2\right)}+\frac{v_1m_1}{\sigma_\varepsilon^2}\right]\right\}$$

where $v_1 = I(J-1)$, $v_2 = I-1$, $\bar{Y}_{i.} = \frac{1}{J}\sum_{j=1}^J Y_{ij}$, $\bar{Y}_{..} = \frac{1}{IJ}\sum_{i=1}^I\sum_{j=1}^J Y_{ij}$, $v_1m_1 = \sum_{i=1}^I\sum_{j=1}^J(Y_{ij}-\bar{Y}_{i.})^2$ and $v_2m_2 = J\sum_{i=1}^I(\bar{Y}_{i.}-\bar{Y}_{..})^2$. v_1m_1 is the within group sums of squares and v_2m_2 is the between groups sums of squares. Since $E(m_1) = \sigma_\varepsilon^2$ and $E(m_2) = \sigma_\varepsilon^2+J\sigma_r^2$ an estimate of α is

$$\hat{\alpha} = 1-\frac{m_1}{m_2}.$$

The prior that will be used in the analysis, is the Jeffreys independence prior

$$\pi\left(\sigma_\varepsilon^2,\sigma_r^2\right) \propto \left(\sigma_\varepsilon^2\right)^{-1}\left(\sigma_\varepsilon^2+J\sigma_r^2\right)^{-1}.$$

It can be shown that it is also a reference and probability matching prior. By multiplying the likelihood with the prior and integrating with respect to θ the posterior density function of the variance components is given by

$$\pi(\sigma_\varepsilon^2,\sigma_r^2|data) \propto \left(\sigma_\varepsilon^2\right)^{-(v_1+2)/2}\left(\sigma_\varepsilon^2+J\sigma_r^2\right)^{-(v_2+2)/2}\exp\left\{-\frac{1}{2}\left[\frac{v_1m_1}{\sigma_\varepsilon^2}+\frac{v_2m_2}{\sigma_\varepsilon^2+J\sigma_r^2}\right]\right\} \quad \sigma_\varepsilon^2>0,\sigma_r^2>0. \tag{6.3}$$

From Equation 6.3 it is clear that $\pi(\sigma_\varepsilon^2,\sigma_\varepsilon^2+J\sigma_r^2|data) = \pi(\sigma_\varepsilon^2|data)\pi(\sigma_\varepsilon^2+J\sigma_r^2|data)$. Since the posterior distributions of σ_ε^2 and $\sigma_\varepsilon^2+J\sigma_r^2$ are independent inverse-gamma distributions, it follows that

$$\sigma_\varepsilon^2 \sim \frac{v_1m_1}{\chi_{v_1}^2} \quad \text{and} \quad \sigma_\varepsilon^2+J\sigma_r^2 \sim \frac{v_2m_2}{\chi_{v_2}^2}.$$

The posterior density function of $\alpha = 1-\frac{\sigma_\varepsilon^2}{\sigma_\varepsilon^2+J\sigma_r^2}$ can therefore easily be obtained and is given in Theorem 6.1.

Theorem 6.1. *The posterior density function of* α *is given by*

$$\pi(\alpha|\hat{\alpha}) = K_1 \left(\frac{v_2}{v_1}\right)^{\frac{1}{2}v_2} \left(\frac{1}{1-\hat{\alpha}}\right)^{\frac{1}{2}v_2} (1-\alpha)^{\frac{1}{2}v_2-1} \times \left[1+\frac{v_2}{v_1}\left(\frac{1-\alpha}{1-\hat{\alpha}}\right)\right]^{-\frac{1}{2}(v_1+v_2)}$$

where $K_1 = \frac{\Gamma\left(\frac{v_1+v_2}{2}\right)}{\Gamma\left(\frac{v_1}{2}\right)\Gamma\left(\frac{v_2}{2}\right)}$ *and* $\hat{\alpha} = 1 - \frac{m_1}{m_2}$. *Since*

$$\alpha|\hat{\alpha} \sim 1-(1-\hat{\alpha})F_{v_2,v_1} \tag{6.4}$$

$$E(\alpha|\hat{\alpha}) = 1-(1-\hat{\alpha})\left(\frac{v_1}{v_1-2}\right)$$

and $Var(\alpha|\hat{\alpha}) = (1-\hat{\alpha})^2 \frac{2v_1^2(v_2+v_1-2)}{v_2(v_1-2)^2(v_1-4)}$.

Proof. The proof is given in Appendix D.1. □

6.4 Example and Simulation Procedure of the Predictive Density Function

6.4.1 Example: The Dyestuff Data

Consider the following example from Box & Tiao (1973) based on the dyestuff data. The purpose of the experiment was to learn to what extent batch to batch variation in a certain raw material was responsible for variation in the final product yield. Five samples from each of six randomly chosen batches of raw material were taken and a single laboratory determination of product yield was made for each of the resulting 30 samples. The data is from Davies & Goldsmith (1972), where they reported the data from an experiment designed to investigate the batch to batch variation in the quality of an intermediate product (H-acid) on the yield of a dyestuff (Naphthlence Black 1213) made from it. Six samples of the H-acid representing different batches of works manufacture were selected and five preparations of the dyestuff were made in the laboratory from each sample. The equivalent yields of each preparation as grams of standard color was determined by dye-trial and the data are given in Table 6.1 (Sahai & Ojeda, 2004). This is a repetitive process. In this example $I = 6$ refers to the number of batches and $J = 5$ denotes the number of observations contained within each batch.

Table 6.1: Dyestuff Data.

Batch					
1	**2**	**3**	**4**	**5**	**6**
1545	1540	1595	1445	1595	1520
1440	1555	1550	1440	1630	1455
1440	1490	1605	1595	1515	1450
1520	1560	1510	1465	1635	1480
1580	1495	1560	1545	1625	1445

From the data it follows that $v_1 m_1 = \sum_{i=1}^{I} \sum_{j=1}^{J} (y_{ij} - \bar{y}_{i.})^2 = 58\,830$ and $v_2 m_2 = J \sum_{i=1}^{I} (\bar{y}_{i.} - \bar{y}_{..})^2 = 56\,358$ and $\hat{\alpha} = 1 - \frac{m_1}{m_2} = 0.7825268$. The posterior distribution of α which is derived in Theorem 6.1 is illustrated for the Dyestuff data in Figure 6.1.

Table 6.2: Summary Statistics and Credibility Intervals of α.

Mean (α)	Median (α)	Mode (α)	Var (α)
0.7666	0.8051	0.879	0.0266
90% HPD Interval		95% HPD Interval	
(0.5187; 0.9705)		(0.3990; 0.9746)	

The values for the Mean $(\alpha) = 0.7666$ and Var $(\alpha) = 0.0266$ correspond well with the theoretical values

$$E(\alpha|\hat{\alpha}) = 1 - (1 - \hat{\alpha}) \left(\frac{v_1}{v_1 - 2} \right) = 0.7627564$$

and

$$Var(\alpha|\hat{\alpha}) = (1 - \hat{\alpha})^2 \frac{2v_1^2 (v_2 + v_1 - 2)}{v_2 (v_1 - 2)^2 (v_1 - 4)} = 0.0303936.$$

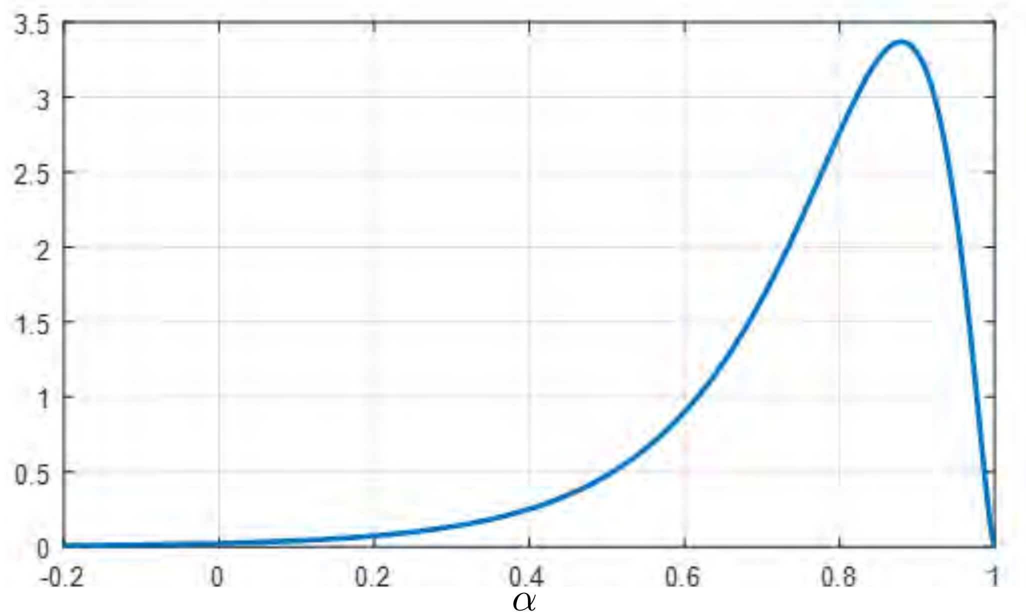


Figure 6.1: Posterior Density Function $\pi(\alpha|\hat{\alpha})$ in the Case of the Dyestuff Data.

6.4.2 The Predictive Density Function of a Future Cronbach's α

Consider a future (unseen) experiment

$$\tilde{Y}_{ij} = \theta + r_i + \varepsilon_{ij} \quad \text{for } i = 1, \ldots \tilde{I} \text{ and } j = 1, \ldots J$$

where $\tilde{Y}_{ij}$ is the value of the j^{th} observation in the i^{th} group (sample). As before the r_i and ε_{ij} are independent normal variables with zero means and variances σ_r^2 and σ_ε^2 respectively. Define

$$\tilde{v}_1\tilde{m}_1 = \sum_{i=1}^{\tilde{I}} \sum_{j=1}^{J} (\tilde{Y}_{ij} - \bar{\tilde{Y}}_{i.})^2$$

and

$$\tilde{v}_2\tilde{m}_2 = J \sum_{i=1}^{\tilde{I}} (\bar{\tilde{Y}}_{i.} - \bar{\tilde{Y}}_{..})^2$$

where $\tilde{v}_1 = \tilde{I}(J-1)$, $\tilde{v}_2 = \tilde{I} - 1$, $\bar{\tilde{Y}}_{i.} = \frac{1}{J}\sum_{j=1}^{J} \tilde{Y}_{ij}$, $\bar{\tilde{Y}}_{..} = \frac{1}{\tilde{I}J}\sum_{i=1}^{\tilde{I}}\sum_{j=1}^{J} \tilde{Y}_{ij}$. $\tilde{v}_1\tilde{m}_1$ is the within group sums of squares and $\tilde{v}_2\tilde{m}_2$ is the between groups sums of squares of the new (unseen) experiment (data set). It is well known from classical (traditional) statistics that $E(\widetilde{m}_1) = \sigma_\varepsilon^2$ and $E(\widetilde{m}_2) = \sigma_\varepsilon^2 + J\sigma_r^2$. A future (unseen) Cronbach's alpha is therefore defined as

$$\hat{\alpha}_f = 1 - \frac{\widetilde{m}_1}{\widetilde{m}_2}.$$

For given σ_ε^2 and $\sigma_\varepsilon^2 + J\sigma_r^2$, it follows that $\frac{\tilde{v}_1 \widetilde{m}_1}{\sigma_\varepsilon^2} \sim \chi^2_{\tilde{v}_1}$ and $\frac{\tilde{v}_2 \widetilde{m}_2}{\sigma_\varepsilon^2 + J\sigma_r^2} \sim \chi^2_{\tilde{v}_2}$. Therefore $\widetilde{m}_1 \sim \sigma_\varepsilon^2 \frac{\chi^2_{\tilde{v}_1}}{\tilde{v}_1}$, $\widetilde{m}_2 \sim \left(\sigma_\varepsilon^2 + J\sigma_r^2\right) \frac{\chi^2_{\tilde{v}_2}}{\tilde{v}_2}$ and

$$\frac{\widetilde{m}_1}{\widetilde{m}_2} \sim \frac{\sigma_\varepsilon^2}{\sigma_\varepsilon^2 + J\sigma_r^2} F_{\tilde{v}_1,\tilde{v}_2}.$$

Also $\hat{\alpha}_f = 1 - \frac{\widetilde{m}_1}{\widetilde{m}_2} \sim 1 - \frac{\sigma_\varepsilon^2}{\sigma_\varepsilon^2 + J\sigma_r^2} F_{\tilde{v}_1,\tilde{v}_2}.$ and therefore

$$\hat{\alpha}_f | \alpha \sim 1 - (1-\alpha) F_{\tilde{v}_1,\tilde{v}_2} \tag{6.5}$$

where $F_{\tilde{v}_1,\tilde{v}_2}$ is a F distribution with $\tilde{v}_1$ and $\tilde{v}_2$ degrees of freedom. The following theorem can now be proved.

Theorem 6.2. *The predictive density of a future Cronbach's alpha* $(\hat{\alpha}_f)$, *for given* α *is*

$$f\left(\hat{\alpha}_f|\alpha\right) = K_2 \left(\frac{\tilde{v}_1}{\tilde{v}_2}\right)^{\frac{\tilde{v}_1}{2}} \left(\frac{1}{1-\alpha}\right)^{\frac{\tilde{v}_1}{2}} \left(1-\hat{\alpha}_f\right)^{\frac{1}{2}\tilde{v}_1 - 1} \left[1 + \frac{\tilde{v}_1}{\tilde{v}_2}\left(\frac{1-\hat{\alpha}_f}{1-\alpha}\right)\right]^{-\frac{1}{2}(\tilde{v}_1+\tilde{v}_2)} \tag{6.6}$$

where $K_2 = \frac{\Gamma\left(\frac{\tilde{v}_1+\tilde{v}_2}{2}\right)}{\Gamma\left(\frac{\tilde{v}_1}{2}\right)\Gamma\left(\frac{\tilde{v}_2}{2}\right)}$. *Also*

$$E\left(\hat{\alpha}_f|\alpha\right) = 1 - (1-\alpha)\left(\frac{\tilde{v}_2}{\tilde{v}_2 - 2}\right)$$

and $Var\left(\hat{\alpha}_f|\alpha\right) = (1-\alpha)^2 \frac{2(\tilde{v}_2)^2(\tilde{v}_2+\tilde{v}_1-2)}{\tilde{v}_1(\tilde{v}_2-2)^2(\tilde{v}_2-4)}$.

Proof. The proof is given in Appendix D.2. □

6.4.3 Calculation of the Unconditional Predictive Density Function of a Future Cronbach's Alpha ($\hat{\alpha}_f$)

The predictive density of a future Cronbach's alpha is

$$f\left(\hat{\alpha}_f|\hat{\alpha}\right) = \int f\left(\hat{\alpha}_f|\alpha\right)\pi\left(\alpha|\hat{\alpha}\right)d\alpha \tag{6.7}$$

where $\pi(\alpha|\hat{\alpha})$ is the posterior density function of α which is derived in Theorem 6.1 and $f\left(\hat{\alpha}_f|\alpha\right)$ is the conditional predictive density function derived in Theorem 6.2. The integral in Equation 6.7 is difficult to solve analytically but can be obtained by either numerical integration or the following simulation procedure:

1. Calculate $\hat{\alpha} = 1 - \frac{m_1}{m_2}$.

2. Simulate α from its posterior distribution and substitute it in $f\left(\hat{\alpha}_f|\alpha\right)$ and draw the density function. In the proof of Theorem 6.1 it is shown that

$$\alpha|\hat{\alpha} \sim 1-(1-\hat{\alpha})F_{v_2,v_1}. \tag{6.8}$$

   It is therefore easy to simulate α from its posterior distribution.

3. Iterate step two a 100 000 times and determine the average of the 100 000 conditional predictive density functions to obtain $f\left(\hat{\alpha}_f|\hat{\alpha}\right)$ the unconditional predictive density function. This method is called the Rao-Blackwell procedure. $\hat{\alpha}$ represents the data.

4. Determine the mean, median, mode and variance of $f\left(\hat{\alpha}_f|\hat{\alpha}\right)$ as well as the 90^{th} and 95^{th} prediction intervals. For the Dyestuff data it will first be assumed that $\tilde{v}_1 = v_1 = 24$ and $\tilde{v}_2 = v_2 = 5$. In other words future experiments will have the same number of samples and observations per sample as the original data set.

In Figure 6.2 $f\left(\hat{\alpha}_f|\hat{\alpha}\right)$ is illustrated.

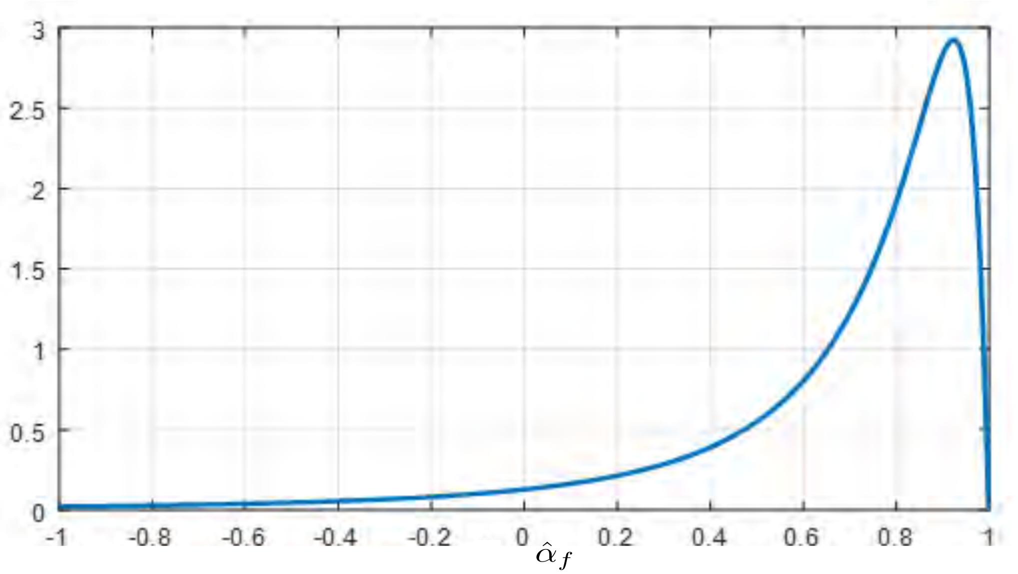


Figure 6.2: Predictive Density Function $f\left(\hat{\alpha}_f|\hat{\alpha}\right)$ for the Dyestuff Data.

Table 6.3: Summary Statistics and Credibility Intervals of $\hat{\alpha}_f$ for the Dyestuff Data, $\beta = 0.1$.

Mean $\left(\hat{\alpha}_f\right)$	Median $\left(\hat{\alpha}_f\right)$	Mode $\left(\hat{\alpha}_f\right)$	Var $\left(\hat{\alpha}_f\right)$
0.6038	0.780	0.922	0.6023
90% Equal tail Interval		95% HPD Interval	
$(-0.275; 0.9638)$		$(-0.332; 0.989)$	

For the predictive density function $f\left(\hat{\alpha}_f|\hat{\alpha}\right)$ the exact mean and variance can be derived analytically. The following theorem can now be stated.

Theorem 6.3. *The mean and variance of* $f\left(\hat{\alpha}_f|\hat{\alpha}\right)$, *the unconditional predictive density function of* $\hat{\alpha}_f$, *are given by*

$$E\left(\hat{\alpha}_f|\hat{\alpha}\right) = 1-(1-\hat{\alpha})\left(\frac{v_1}{v_1-2}\right)\left(\frac{\tilde{v}_2}{\tilde{v}_2-2}\right) \tag{6.9}$$

and

$$Var\left(\hat{\alpha}_f|\hat{\alpha}\right) = (1-\hat{\alpha})^2\left\{Var\left(F_{v_2,v_1}\right)+\left(\frac{v_1}{v_1-2}\right)^2\right\}Var\left(F_{\tilde{v}_1,\tilde{v}_2}\right)$$
$$+(1-\hat{\alpha})^2\left(\frac{\tilde{v}_2}{\tilde{v}_2-2}\right)^2 Var\left(F_{v_2,v_1}\right) \tag{6.10}$$

where $Var\left(F_{a,b}\right) = \frac{2b^2(a+b-2)}{a(b-2)^2(b-4)}$.

Proof. The proof is given in Appendix D.3. □

For the Dyestuff data $E\left(\hat{\alpha}_f|\hat{\alpha}\right) = 0.6045941$ and $Var\left(\hat{\alpha}_f|\hat{\alpha}\right) = 0.6261647$. The numerical values of the mean and variance given in Figure 6.2 are for all practical purposes the same as the exact values.

6.5 Distribution of the Run-length and the Average Run-length in the Case of the 90% Prediction Interval for $\hat{\alpha}_f$

Assuming that the process remains stable, the predictive distribution of $\hat{\alpha}_f$ can be used to derive the distribution of the run-length and the average run-length. Montgomery (1996) defines the average run length as the average number of points that must be plotted before a point indicates an out-of-control condition. The run-length (r) is defined as the number of future $\hat{\alpha}_f$ values until the process goes out of control (until the control charts signals for the first time). Note that r does not include that $\hat{\alpha}_f$ value when the control chart signals. The process goes out of control if $\hat{\alpha}_f$ of a future experiment is very small or very large.

The resulting region of size β for the determination of the run-length is given by

$$\beta = \int_{R(\beta)} f\left(\hat{\alpha}_f\,|\hat{\alpha}\right) d\hat{\alpha}_f, \tag{6.11}$$

where $f\left(\hat{\alpha}_f\,|\hat{\alpha}\right)$ is defined in Equation 6.7 and $R(\beta)$ is the size of the rejection region.

In the case of the 90% prediction interval $\beta = 0.10$ and $R(\beta)$ presents those values of $\hat{\alpha}_f$ that are smaller than -0.275 and larger than 0.9628. Given α and a stable Phase I process, the distribution of the run-length is geometrical with parameter

$$\Psi(\alpha) = \int_{R(\beta)} f\left(\hat{\alpha}_f\,|\alpha\right) d\hat{\alpha}_f, \tag{6.12}$$

where $f\left(\hat{\alpha}_f\,|\alpha\right)$ is the distribution of a future $\hat{\alpha}_f$ value given that α is known. See Theorem 6.2, Equation 6.6. The values of α is however unknown and its uncertainty is described by the posterior distribution given in Theorem 6.1. By simulating α from the posterior distribution and substituting it in Equation 6.12, $\Psi(\alpha)$ can be calculated. This must be done for each future experiment. Therefore by simulating a large number of α values, a large number of $\Psi(\alpha)$ values can be obtained. Also a large number of geometric distributions, i.e. a large number of run-length distributions, each with a different set of parameter values $\left\{\Psi\left(\alpha^{(1)}\right),\Psi\left(\alpha^{(2)}\right),\ldots,\Psi\left(\alpha^{(m)}\right)\right\}$ will be obtained. Since the run-length r for given α, is geometrical distributed with mean

$$E\left(r|\alpha\right)=\frac{1-\Psi\left(\alpha\right)}{\Psi\left(\alpha\right)}$$

and variance

$$Var\left(r|\alpha\right)=\frac{1-\Psi\left(\alpha\right)}{\Psi^{2}\left(\alpha\right)}.$$

The unconditional mean is

$$E\left(r\,|\hat{\alpha}\,\right)=E\left\{E\left(r|\alpha\right)\right\}$$

and the unconditional variance

$$Var\left(r\,|\hat{\alpha}\,\right)=E\left\{Var\left(r|\alpha\right)\right\}+Var\left\{E\left(r|\alpha\right)\right\}$$

can be calculated. The expectation and variance are taken with respect to the posterior distribution $\pi\left(\alpha\,|\hat{\alpha}\,\right)$.

In Figure 6.3, the simulated distribution of the geometric parameter, $\Psi(\alpha)$, using $f\left(\hat{\alpha}_f\,|\alpha\right)$ is given and the summary statistics and credibility intervals are given in Table 6.4.

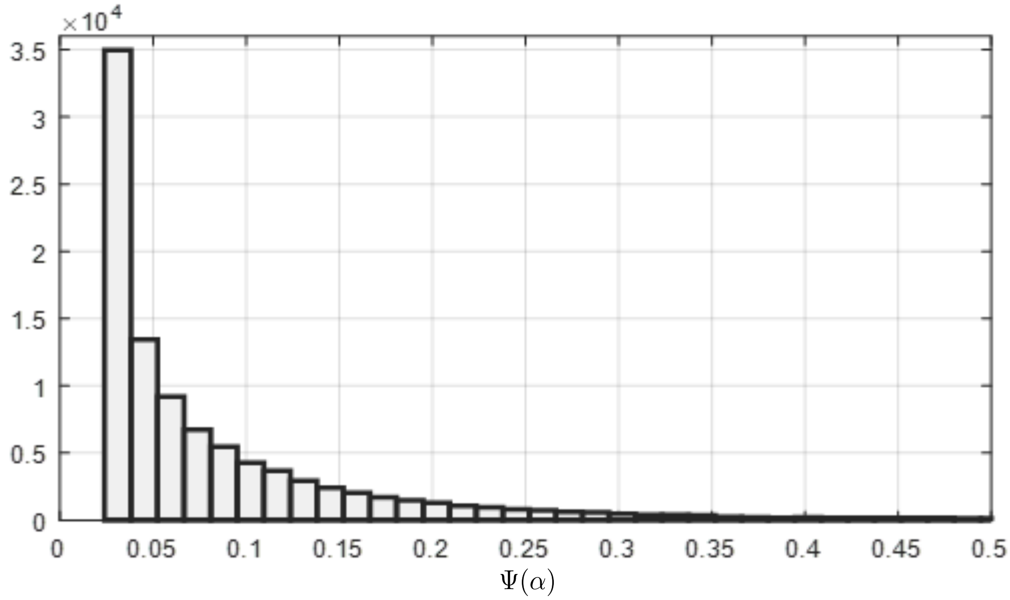


Figure 6.3: Histogram of $\Psi(\alpha)$.

Table 6.4: Summary Statistics and Credibility Intervals of $\Psi(\alpha)$.

Mean $(\Psi(\alpha))$	Median $(\Psi(\alpha))$	Mode $(\Psi(\alpha))$	Var $(\Psi(\alpha))$
0.0989	0.0545	0.0228	0.0149
90% Equal Tailed Interval		95% HPD Interval	
(0.0240; 0.4816)		(0.0239; 0.3185)	

As it should be, the mean of 0.0989 is in the line with $\beta = 0.1$. The long tail of the above distribution is an indication of the uncertainty in the parameter $\Psi(\alpha)$. In Figure 6.4 the simulated distribution of the average run-length ($E(r|\alpha)$) is illustrated. Summary statistics and credibility intervals are given in Table 6.5. In Figure 6.5 the distribution of $Var(r|\alpha)$ is given and in Figure 6.6 the predictive density function of a future run-length which is the average of a large number of geometric distributions each with its own parameter value, $\Psi(\alpha)$, is given. Summary statistics and credibility intervals are illustrated in Table 6.6.

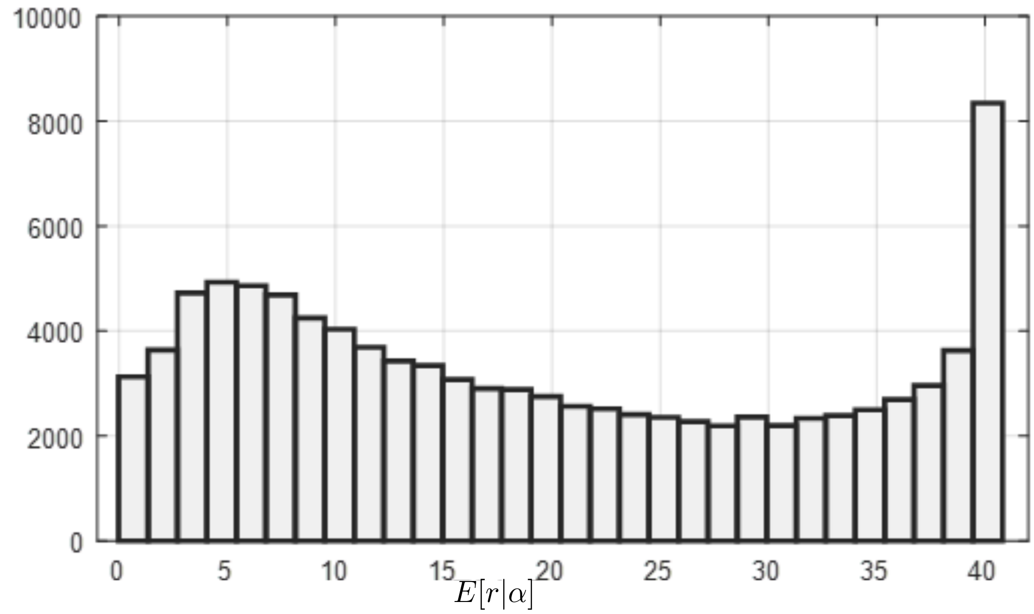


Figure 6.4: Distribution of the Expected Run-Length $E[r|\alpha]$ in the Case of the 90% Prediction Interval.

Table 6.5: Summary Statistics and Credibility Intervals of $E[r|\alpha]$.

Mean $(E[r\|\alpha])$	Median $(E[r\|\alpha])$	Mode $(E[r\|\alpha])$	Var $(E[r\|\alpha])$
19.4147	17.3609	40	164.2296
95% HPD Interval			
(2.1222; 40.8112)			

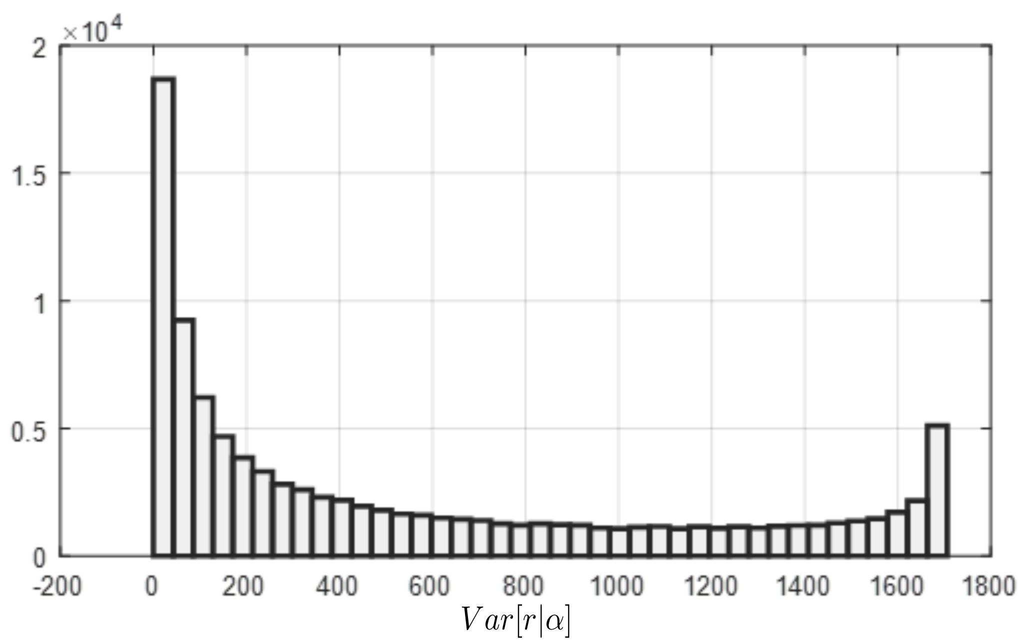


Figure 6.5: Histogram of $Var(r|\alpha)$.

$$E\{Var(r|\alpha)\} = 560.5730$$

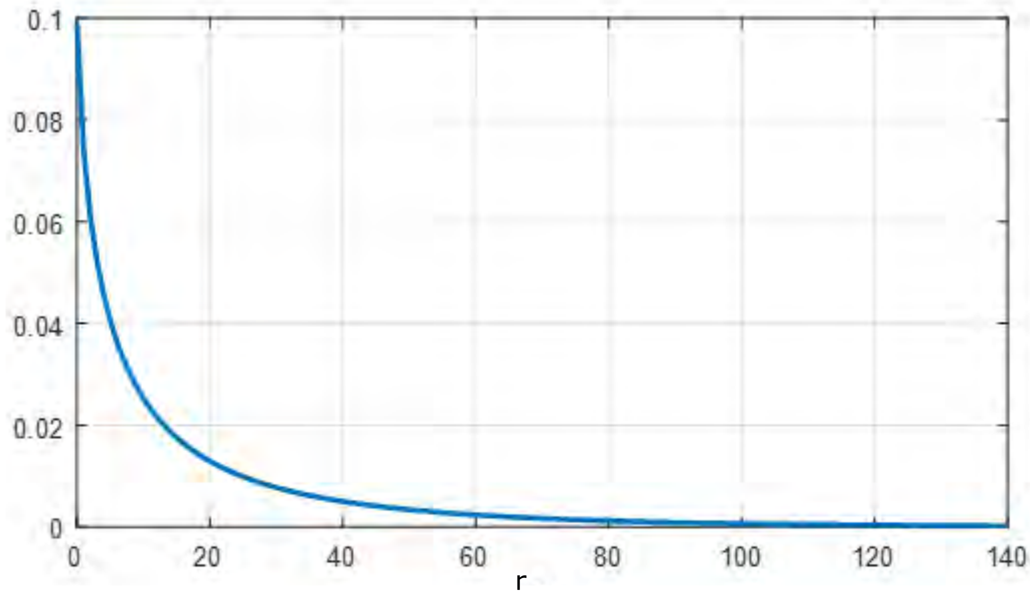


Figure 6.6: The Predictive Density Function $f(r|\hat{\alpha})$ in the Case of the 90% Prediction Interval ($\beta = 0.1$).

Table 6.6: Summary Statistics and Credibility Intervals of r.

Mean $(r\|\hat{\alpha})$	Median $(r\|\hat{\alpha})$	Var $(r\|\hat{\alpha})$
19.4113	9.01	723.41
90% HPD Interval		95% HPD Interval
$(0;50.33)$		$(0;72.27)$

- A comparison of Figures 6.4 and 6.6 show that $Mean(E[r|\alpha]) = 19.4147$ and $Mean(r|\hat{\alpha}) = 19.4113$. The means are for all practical purposes the same. Theoretically this should have been the case. Also $Var(r|\hat{\alpha}) = E\{Var(r|\alpha)\} + Var\{E(r|\alpha)\} = 560.5730 + 164.22 = 724.8026$. From Figure 6.6 it is clear that $Var(r|\hat{\alpha}) = 723.41$. The two methods of calculating $Var(r|\hat{\alpha})$ give for all practical purposes the same answer.

- A mean run-length of 19.41 is larger than $\frac{1-0.1}{0.1} = 9$ that would have been expected if $\beta = 0.1$. The reason for this larger average run-length is the small parameter values of some of the geometric distributions.

- The median run-length of 9.01 however corresponds well with the theoretical value of 9.

- A mean run-length of 19.4 is an indication that the charting statistic ($\hat{\alpha}_f$) will signal on average every 19^{th} or 20^{th} experiment even if the Phase I process is stable (in control). It can however take much longer. According to the 95% HPD Interval it can take as long as 72 experiments. It is however not impossible that $\hat{\alpha}_f$ will signal as early as the first experiment.

6.6 A Larger Experiment

In the example of the Dyestuff data, it was mentioned that it is a repetitive process. So let us assume that after a certain time period the number of samples obtained were hundred and twenty and the number of observations per sample is five which means that $I = 120$, $J = 5$, $v_1 = I(J-1) = 480$, $v_2 = I - 1 = 119$ and $\hat{\alpha} = 0.7825$. In Figure 6.7 the posterior distribution of α is illustrated.

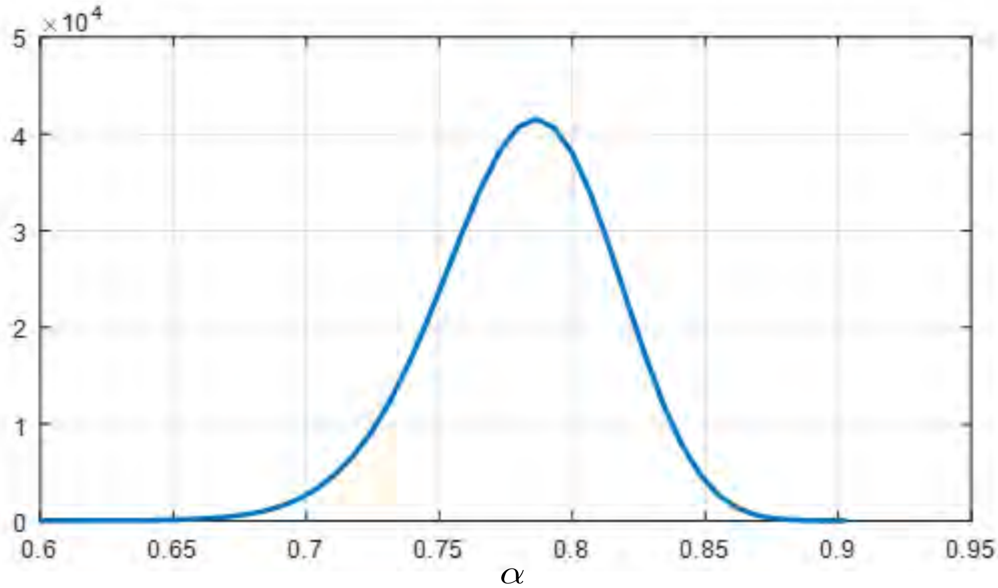


Figure 6.7: Posterior Density Function $\pi(\alpha|\hat{\alpha})$ with $v_1 = 480$, $v_2 = 119$ and $\hat{\alpha} = 0.7825$.

Table 6.7: Summary Statistics and Credibility Intervals of α.

Mean (α)	Median (α)	Mode (α)	Var (α)
0.7816	0.7834	0.786	0.0010
90% Equal Tail Interval		95% Equal Tail Interval	
(0.7264;0.8306)		(0.7143;0.8386)	

Using the formulas derived in Theorem 6.1, the exact mean and variance can be calculated $E(\alpha|\hat{\alpha}) = 0.7815899$ and $Var(\alpha|\hat{\alpha}) = 0.0010055$. The theoretical values are for all practical purposes the same as the numerical values, $Mean(\alpha) = 0.7816$ and $Var(\alpha) = 0.0010$. The predictive density function of $\hat{\alpha}_f$ for a future (unseen) experiment consisting of ninety samples and five observations per sample ($\tilde{I} = 90, J = 5, \tilde{v}_1 = \tilde{I}(J-1) = 360, \tilde{v}_2 = \tilde{I} - 1 = 89$ and $\hat{\alpha} = 0.7825$) is displayed in Figure 6.8.

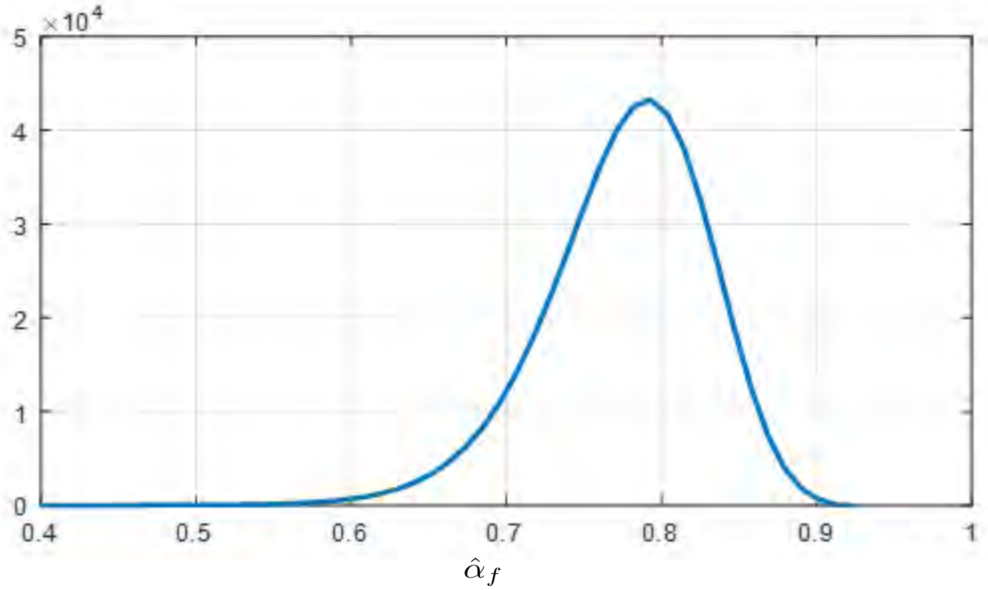


Figure 6.8: Predictive Density Function $f(\hat{\alpha}_f|\hat{\alpha})$ with $\tilde{v}_1 = 360$, $\tilde{v}_2 = 89$ and $\hat{\alpha} = 0.7825$.

Table 6.8: Summary Statistics and Credibility Intervals of $\hat{\alpha}_f$ for the Larger Data, $\beta = 0.1$.

Mean $(\hat{\alpha}_f)$	Median $(\hat{\alpha}_f)$	Mode $(\hat{\alpha}_f)$	Var $(\hat{\alpha}_f)$
0.7765	0.7822	0.79	0.0025
90% Equal tail Interval		95% HPD Interval	
(0.6854; 0.8486)		(0.6763; 0.8677)	

The formulas for the exact mean and variance are derived in Theorem 6.3, which is given by $E(\hat{\alpha}_f|\hat{\alpha}) = 0.776569$ and $Var(\hat{\alpha}_f|\hat{\alpha}) = 0.0025414$. The theoretical values are similar to those given in Figure 6.8, which are equal to $Mean(\hat{\alpha}_f) = 0.7765$ and $Var(\hat{\alpha}_f) = 0.0025$. The posterior and predictive density functions illustrated in Figures 6.7 and 6.8 are much more symmetrical than those shown in Figures 6.1 and 6.2. The reason for this is the larger sample sizes. Researchers are usually interested in a run-length of about 370. The reason for this originated from the fact that if a random variable $Z \sim N(0,1)$ then $P(-3 < Z < 3) = 0.0027$. In other words if $\beta = 0.0027$ the expected run-length $E(r|\alpha) = \frac{1-0.0027}{0.0027} = 369.37$. In practice however a $\beta = 0.0027$ will give a much larger average run-length than 370. The reason for this is the variation in α which is illustrated by the posterior distribution in Figure 6.7. The run-length will also be large if the parameter values of the geometric distributions are small. An advantage of the Bayesian procedure however is that β can be adjusted so that the average or median run-length takes on a value near 370. In our case a median run-length of 354 will be used. In Table 6.9 the average and median run-lengths for $\hat{\alpha}_f$ are given for the different values of β. The theoretical run-length is $\frac{1-\beta}{\beta}$.

Table 6.9: Run-Lengths (r) for $\hat{\alpha}_f$ in the Case of Different β Values. $v_1 = 480$, $v_2 = 119$, $\tilde{v}_1 = 360$, $\tilde{v}_2 = 89$ and $\hat{\alpha} = 0.7825$.

β	$E[r\|\hat{\alpha}]$	$Median[r\|\hat{\alpha}]$	Theoretical Run-Length
0.0050	1754.9	562	199.0
0.0060	1356.5	456	165.7
0.0070	1027.7	354	141.9
0.0080	843.7	305	124.0
0.0090	891.2	250	110.1
0.0100	577.8	215	99.0
0.0150	305.3	120	65.7
0.0200	194.2	80	49.0
0.0250	137.9	58	39.0
0.0300	103.3	44	32.3
0.0350	80.9	35	27.6
0.0400	66.7	30	24.0
0.0450	56.7	25	21.2
0.0500	47.5	21	19.0
0.0550	41.2	19	17.2
0.0600	36.4	17	15.7
0.0650	31.8	14	14.4
0.0700	28.5	13	13.3
0.0750	26.0	12	12.3
0.0800	23.6	11	11.5
0.0850	21.4	10	10.8
0.0900	19.8	9	10.0
0.0950	18.3	8	9.5
0.1000	16.6	7	9.0

6.7 The Out of Control Situation

Assume in a future experiment with $\tilde{I} = 90$ and $J = 5$ the true parameter value of α has changed from 0.78 to 0.74 then the average run-length might change dramatically. If the process is in control and if the desired median run-length is 354, then $\beta = 0.007$ and $R(\beta)$ presents those values of $\hat{\alpha}_f$ that are smaller than $A = 0.6003$ and larger than $B = 0.88$. For the out of control situation the parameter of the geometric distribution is

$$\Psi(\alpha) = \int_{R(\beta)} f(\hat{\alpha}_f|\alpha)\, d\hat{\alpha}_f = 0.0077$$

which means that for $\alpha = 0.74$ the average run-length is now

$$E(r|\alpha) = \frac{1-\Psi(\alpha)}{\Psi(\alpha)} = \frac{1-0.0077}{0.0077} = 128.87$$

and the variance of the run-length is

$$Var(r|\alpha) = \frac{1-\Psi(\alpha)}{\Psi^2(\alpha)} = 16736.38.$$

A change in run-length from 354 to 129 is quite large and is a clear indication that the process is out of control. In Figure 6.9 the "in control" and "out of control" situations are illustrated.

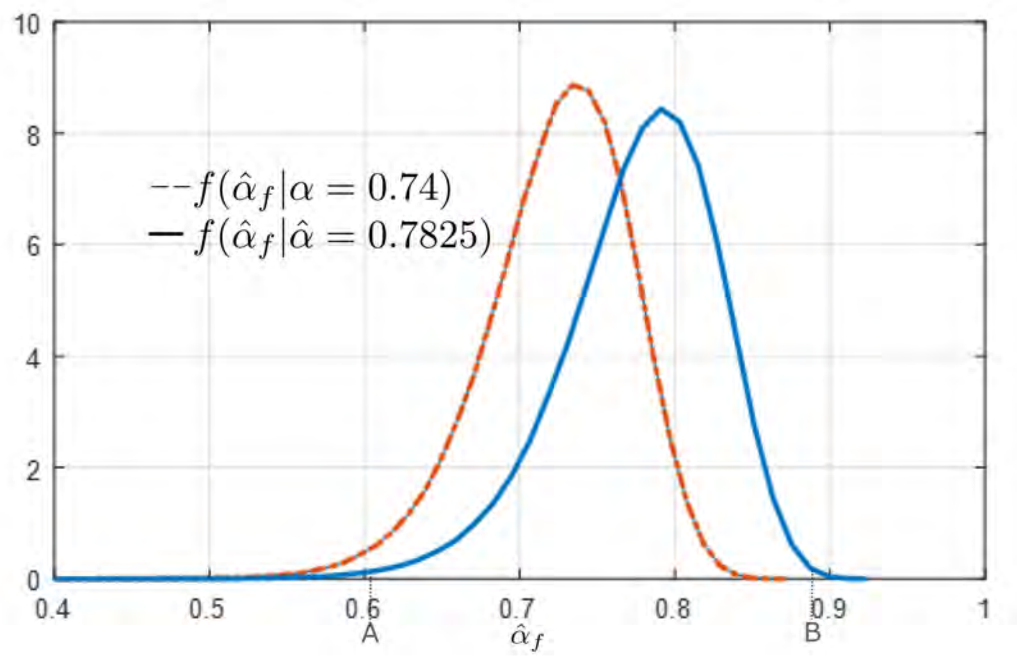


Figure 6.9: Graphs of the In Control and Out of Control Situations.

6.8 Box Diameter Data

6.8.1 Example 6.8

The data in this example are the data provided by Wooluru et al. (2014) from the manufacturing industry. The critical quality characteristic is the "Bore diameter" on the driver gear. The number of batches (samples) $I = 20$ and the number of observations per sample $J = 5$. The sample Cronbach's alpha, $\hat{\alpha} = 0.4952$. In Figure 6.10 the posterior density function of α is given and in Figure 6.11 the predictive density function of $\hat{\alpha}_f$ is given. The number of batches in the future data set are also twenty. Therefore $\tilde{I} = 20$ and $J = 5$.

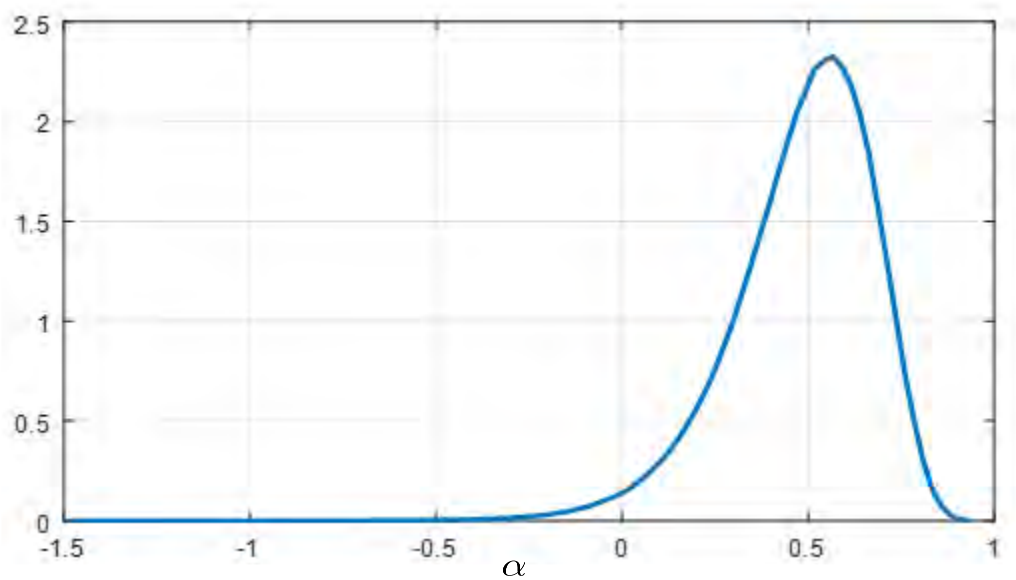


Figure 6.10: Posterior Distribution $f(\alpha|\hat{\alpha})$, $v_1 = 80$ and $v_2 = 19$.

Table 6.10: Summary Statistics and Credibility Intervals of α.

Mean (α)	Median (α)	Mode (α)	Var (α)
0.4822	0.5089	0.56	0.0359
90% Equal Tail Interval		95% Equal Tail Interval	
$(0.1330; 0.7414)$		$(0.0400; 0.7741)$	

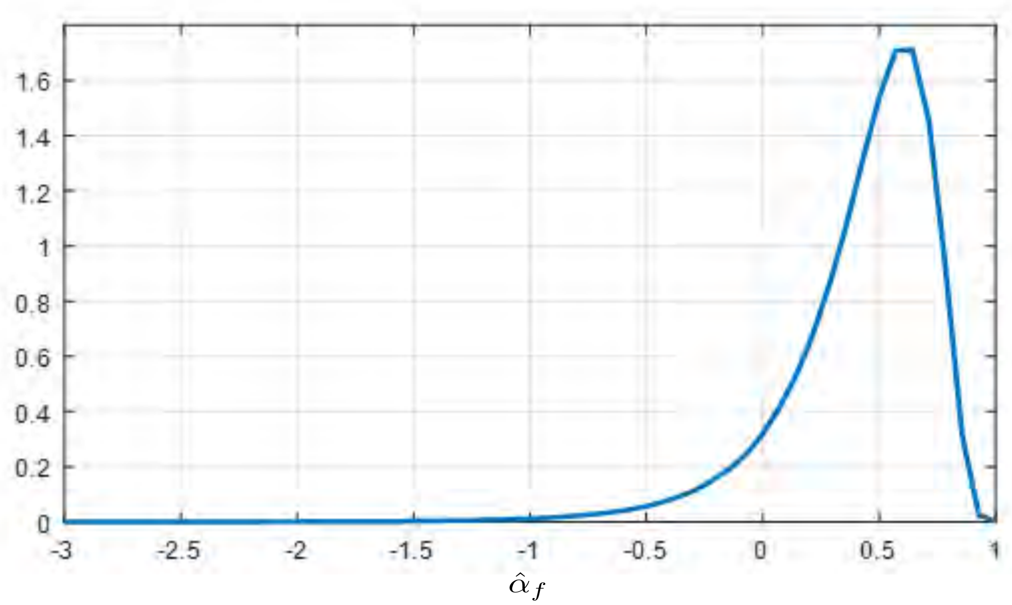


Figure 6.11: Predictive Density $f(\hat{\alpha}_f|\hat{\alpha})$, $\hat{v}_1 = 80$ and $\hat{v}_2 = 19$.

Table 6.11: Summary Statistics and Credibility Intervals of $\hat{\alpha}_f$ for the Bore Diameter Data, $\beta = 0.1$.

Mean $(\hat{\alpha}_f)$	Median $(\hat{\alpha}_f)$	Mode $(\hat{\alpha}_f)$	Var $(\hat{\alpha}_f)$
0.4213	0.4951	0.61	0.1064
90% Equal tail Interval		95% HPD Interval	
$(-0.1901; 0.7858)$		$(-0.2091; 0.8819)$	

In Table 6.12 the average and median run-lengths for $\hat{\alpha}_f$ are given for the different values of β.

Table 6.12: Run-Lengths (r) for $\hat{\alpha}_f$ in the Case of Different β Values. $v_1 = 80$, $v_2 = 19$, $\tilde{v}_1 = 80$, $\tilde{v}_2 = 19$.

β	$E[r\|\hat{\alpha}]$	$Median[r\|\hat{\alpha}]$
0.0050	3030.3	820
0.0060	2355.4	643
0.0070	1887.2	528
0.0080	1493.5	431
0.0090	1214.6	353
0.0100	1023.9	304
0.0110	870.8	266
0.0120	733.8	231
0.0130	638.7	203
0.0140	564.6	185
0.0150	504.9	165
0.0160	453.7	151
0.0170	406.5	138
0.0180	371.7	127
0.0190	336.5	115
0.0200	305.9	106

From the table it can be seen that for an average run-length in the vicinity of 370 $\beta = 0.0180$ should be used instead of $\beta = 0.0027$.

6.9 Conclusion

In this chapter statistical process control limits have been obtained for Cronbach's alpha coefficient in the case of the balanced one-way random effects model. This has been achieved by deriving the predictive distribution of a future (unseen) Cronbach's alpha coefficient. For given variance components, it was shown that the predictive density function of $f(\hat{\alpha}_f|\alpha)$ can be derived analytically. The Jeffreys independence prior was used to derive the posterior distribution of α. The unconditional posterior predictive density function $f(\hat{\alpha}_f|\hat{\alpha})$ can be obtained by Monte Carlo simulation or numerical integration. The predictive density function $f(\hat{\alpha}_f|\hat{\alpha})$ as well as the conditional predictive density functions $f(\hat{\alpha}_f|\alpha)$ can be used to determine the run-length and the average run-length. The distribution of the run-length $f(r|\hat{\alpha})$ is the average of a large number of geometric distributions each with its own parameter value. Three examples are considered. The first example has to do with Dyestuff data and it is from Box & Tiao (1973). In the second example it is assumed that the number of Dyestuff samples have increased from six to hundred and twenty and the number of samples in a future (unseen) data set is ninety. The third example is from Wooluru et al. (2014) and are measured values of "Box diameter" on the driver gear. From the results it can be seen that the average and median run-lengths are usually larger than the theoretical values. An advantage of the Bayesian procedure however is that control limits can be adjusted in such a way that the average or median run-length has a specific value.

Chapter 7

Fiducial and Bayesian Estimation for Cronbach's Alpha

7.1 Introduction

Although Cronbach's alpha is mostly used for assessing internal consistency for specific types of data, it can also be applied to more general covariance matrices. For example, covariance matrices that occur in structural equation models and in multilevel or nested data sets.

Without the assumption of the usual parallel condition, Van Zyl et al. (2000) derived an asymptotic normal distribution for $\hat{\alpha}$ the sample Cronbach's alpha for the case of multivariate normal data with a general covariance matrix. That is, if $n \to \infty$, then $\sqrt{n}\,(\alpha - \hat{\alpha})$ is normally distributed with mean zero and variance

$$\Omega = \left[\frac{2p^2}{(p-1)\,(\mathbf{1}'\Sigma\mathbf{1})^3}\right]\left\{(\mathbf{1}'\Sigma\mathbf{1})\left(trace\Sigma^2 + trace^2\Sigma\right) - 2\,(trace\Sigma)\left(\mathbf{1}'\Sigma^2\mathbf{1}\right)\right\}, \tag{7.1}$$

where n is the sample size, α is the population Cronbach's alpha, Σ is a $p \times p$ covariance matrix of the data and $\mathbf{1}$ is a $p \times 1$ vector of $1's$. Cronbach's alpha is defined as

$$\alpha = \frac{p}{p-1}\left\{1 - \frac{trace\Sigma}{\mathbf{1}'\Sigma\mathbf{1}}\right\}$$

where $trace$ is the trace operator. The sample Cronbach's alpha, $\hat{\alpha}$ is therefore normally distributed with mean α and variance $\frac{1}{n}\Omega$ and the $100\,(1-\beta)\,\%$ asymptotic confidence interval for α can therefore be obtained by

$$\hat{\alpha} \pm z_{1-\beta/2}\sqrt{\frac{1}{n}\Omega}. \tag{7.2}$$

Since Σ is unknown Van Zyl et al. (2000) used the estimator $\hat{\Sigma}$ in Equation 7.1 for the calculation of their confidence intervals. $\hat{\Sigma}$ is the unbiased estimator of Σ and

$$\hat{\alpha} = \frac{p}{p-1}\left\{1 - \frac{trace\hat{\Sigma}}{\mathbf{1}'\hat{\Sigma}\mathbf{1}}\right\}.$$

For the compound symmetric case, Van Zyl et al. (2000) shows that Equation 7.1 simplifies to

$$\Omega = \frac{2(p-1)(1-\alpha)^2}{p}. \tag{7.3}$$

Therefore, as $n \to \infty$, $\sqrt{n}\,(\alpha - \hat{\alpha})$ is normally distributed with mean zero and variance $\frac{2(p-1)(1-\alpha)^2}{p}$. Since α is unknown the variance is estimated by

$$\hat{\Omega} = \frac{2(p-1)(1-\hat{\alpha})^2}{p}.$$

Koning & Franses (2003) mention that it is well known that using the estimated asymptotic variance may well yield anti conservative intervals which are intervals that obtain coverage probabilities below the nominal rate and can therefore lead to a false impression that α is estimated with sufficient accuracy. For this reason, Koning & Franses (2003) proposed to use an exact interval for Cronbach's alpha under compound symmetry by using a result derived from Van Zyl et al. (2000) which states that

$$\frac{1-\hat{\alpha}}{1-\alpha} \sim F_{n(p-1),n}$$

which is used as a pivotal quantity to derive an exact $100(1-\beta)\%$ confidence interval for α given by

$$\left[1 - \frac{(1-\hat{\alpha})}{F_L}; 1 - \frac{(1-\hat{\alpha})}{F_U}\right] \tag{7.4}$$

where F_L and F_U are the lower and upper values of the interval from the F distribution with numerator degrees of freedom $n(p-1)$ and n denominator degrees of freedom. Koning & Franses (2003) showed that this interval obtained coverage probabilities closer to the nominal rate and performed better than the asymptotic confidence interval. However, this exact interval only holds under compound symmetry and it is not known if there exists an exact confidence interval for a general covariance matrix. As mentioned before, Van Zyl et al. (2000) only derived the asymptotic confidence interval for a general covariance matrix, the exact interval only works under compound symmetry. A number of authors used the asymptotic confidence interval, see Koning & Franses (2003), Duhachek & Iacobucci (2004), Cui & Li (2012) and Tsagris et al. (2013) for more details. Duhachek & Iacobucci (2004) studied the standard error of Cronbach's alpha and showed using simulation studies that the alpha standard error (ASE), which is the standard deviation of the variance derived by Van Zyl et al. (2000), is the most well-behaved error statistic as determined by both accuracy and precision. Duhachek & Iacobucci

(2004) concluded that the ASE's and asymptotic intervals function in a predictably robust manner even when covariance heterogeneity is severe.

Padilla & Zhang (2011) also considered a multivariate normal distribution with mean $\boldsymbol{\mu}\,(p\times 1)$ and covariance $\Sigma\,(p\times p)$. By choosing a conjugate prior for both the covariance matrix and the mean vector they derive the posterior distributions of $\boldsymbol{\mu}$ and Σ. By simulating values from the posterior distribution of Σ an estimation of the posterior distribution of α can be obtained. In terms of coverage, the Bayesian confidence intervals performed better than the frequentist intervals.

Payandeh Najafabadi & Najafabadi (2016) considered the estimation of α under a Bayesian framework. Simulation studies suggested that the Bayes estimator under the LINEX loss function reduces the bias of the ordinary maximum likelihood estimator.

In the next sections, fiducial and posterior distributions will be derived for α in the case of the bivariate normal distribution. Various priors will be considered for the variance components σ_1^2, σ_2^2 and the correlation coefficient ρ. These prior distributions are derived from Berger & Sun (2008). Since most of the posterior distributions are obtained in closed form, Monte Carlo simulation can be used to compute them. Credibility intervals for Cronbach's alpha using the objective priors derived by Berger & Sun (2008) and an asymptotic interval by Van Zyl et al. (2000), will be compared with respect to coverage probabilities and interval lengths. Most of the results that are used in the literature requires the compound symmetry assumption for analyses of Cronbach's alpha. This chapter focuses on estimation of Cronbach's alpha for a general covariance matrix using a Bayesian approach and comparing these results to the asymptotic frequentist interval valid under a general covariance matrix framework.

7.2 Fiducial Distributions

7.2.1 The Bivariate Normal Distribution

The bivariate normal distribution is denoted by $\mathbf{X}\sim N(\boldsymbol{\mu},\Sigma)$ where $\mathbf{X}=[X_1,X_2]'$ has mean parameters $\boldsymbol{\mu}=[\mu_1,\mu_2]'$ and covariance matrix

$$\Sigma=\begin{bmatrix}\sigma_1^2 & \sigma_{12}\\ \sigma_{21} & \sigma_2^2\end{bmatrix}=\begin{bmatrix}\sigma_1^2 & \rho\sigma_1\sigma_2\\ \rho\sigma_2\sigma_1 & \sigma_2^2\end{bmatrix},$$

where $\rho=\frac{\sigma_{12}}{\sigma_1\sigma_2}$ is the population correlation coefficient. For the bivariate normal distribution, Cronbach's alpha is given by

$$\alpha=2\left\{1-\frac{\sigma_1^2+\sigma_2^2}{\sigma_1^2+\sigma_2^2+2\rho\sigma_1\sigma_2}\right\}.$$

Since Cronbach's alpha is a function of the variance components and the population correlation coefficient, it is easy to simulate α if it is known how to simulate σ_1, σ_2 and ρ. The following sections will discuss closed-form results that can be used to simulate these parameters using Monte Carlo simulation.

7.2.2 Sufficient Statistics

For a random sample of $n \geq 3$ observations drawn from a bivariate normal distribution, the sufficient statistics are $\bar{\mathbf{x}} = [\bar{x}_1\ \bar{x}_2]'$ and $\bar{x}_i = \frac{1}{n}\sum_{k=1}^{n} x_{ik}$. Also $s_{ij} = \sum_{k=1}^{n}(x_{ik} - \bar{x}_i)\left(x_{jk} - \bar{x}_i\right)$ where $i, j = 1, 2$ and

$$S = \begin{bmatrix} s_{11} & s_{12} \\ s_{21} & s_{22} \end{bmatrix} = \begin{bmatrix} s_{11} & r\sqrt{s_{11}s_{22}} \\ r\sqrt{s_{22}s_{11}} & s_{22} \end{bmatrix},$$

where $r = \frac{s_{12}}{\sqrt{s_{11}s_{22}}}$ is the sample correlation coefficient between x_1 and x_2. The distribution of the elements of S will be used to derive the fiducial distributions of σ_1, σ_2 and ρ. The unbiased estimate of Σ is given by

$$\begin{aligned}\hat{\Sigma} &= \frac{1}{n-1}S \\ &= \frac{1}{n-1}\begin{bmatrix} s_{11} & s_{12} \\ s_{21} & s_{22} \end{bmatrix}\end{aligned}$$

which will be used in the calculation of the confidence interval proposed by Van Zyl et al. (2000).

7.2.3 Distribution of S

We are interested in the distribution of $f(s_{11}, s_{22}, s_{21}) = f(s_{11})\, f(s_{22}|s_{11})\, f(s_{12}|s_{11}, s_{22})$. By making use of Anderson (1958) the following can be obtained

$$s_{11} \sim \sigma_1^2 \chi^2_{n-1}, \tag{7.5}$$

$$s_{22}|s_{11} \sim \sigma_2^2\left(1-\rho^2\right)\chi^2_{n-2} \tag{7.6}$$

and

$$s_{12}|s_{11}, s_{22} \sim N\left(\frac{s_{11}\rho\sigma_2}{\sigma_1}, s_{11}\sigma_2^2\left(1-\rho^2\right)\right). \tag{7.7}$$

From Equations 7.5, 7.6 and 7.7 it follows that the fiducial distributions of $(\sigma_1, \sigma_2, \rho)$ given the data are

$$\sigma_1 = \sqrt{\frac{s_{11}}{\chi^2_{n-1}}}, \tag{7.8}$$

$$\sigma_2 = \sqrt{s_{22}\left(1-r^2\right)}\sqrt{\frac{1}{\chi^2_{n-2}} + \frac{1}{\chi^2_{n-1}}\left(\frac{Z}{\sqrt{\chi^2_{n-2}}} - \frac{r}{\sqrt{1-r^2}}\right)^2}, \tag{7.9}$$

$\rho = \psi(Y)$ where

$$Y = \frac{-Z}{\sqrt{\chi^2_{n-1}}} + \frac{\sqrt{\chi^2_{n-2}}}{\sqrt{\chi^2_{n-1}}}\frac{r}{\sqrt{1-r^2}}, \tag{7.10}$$

$\psi(X) = \frac{X}{\sqrt{1+X^2}}$ and $Z \sim N(0,1)$ is the standard normal distribution, χ^2_{n-1} and χ^2_{n-2} are chi-squared distributions with the indicated degrees of freedom. For further details see Berger & Sun (2008). The proofs of Equations 7.8, 7.9 and 7.10 are given in Appendix E.1.

7.3 Prior and Posterior Distributions

In this section, various prior distributions will be discussed for the variance components σ_1^2, σ_2^2 and the correlation coefficient ρ. As mentioned, these prior distributions are given in Berger & Sun (2008). They considered the prior

$$\pi_{ab}(\mu_1,\mu_2,\sigma_1,\sigma_2,\rho) \propto \frac{1}{\sigma_1^{3-a}\sigma_2^{2-b}(1-\rho^2)^{2-b/2}} \tag{7.11}$$

which is a subclass of the generalized Wishart priors from Brown et al. (1994). Special cases of this class of priors are the Jeffreys rule prior $\pi_J = \pi_{10}$, the independence Jeffreys prior $\pi_{IJ} = \pi_{21}$, the right-Haar prior $\pi_H = \pi_{12}$ and π_{RO} which has $a = b = 1$. The independence Jeffreys prior follows from using a constant prior for the means and the Jeffreys prior for the covariance matrix. It is proved by Berger & Sun (2006) that if the prior π_{ab} is applied the constructive posterior of (σ_1,σ_2,ρ) given the data can be expressed as

$$\sigma_1 = \sqrt{\frac{s_{11}}{\chi^2_{n-a}}}, \tag{7.12}$$

$$\sigma_2 = \sqrt{s_{22}\left(1-r^2\right)}\sqrt{\frac{1}{\chi^2_{n-b}} + \frac{1}{\chi^2_{n-a}}\left(\frac{Z}{\sqrt{\chi^2_{n-b}}} - \frac{r}{\sqrt{1-r^2}}\right)^2}, \tag{7.13}$$

$\rho = \psi(Y)$ where

$$Y = \frac{-Z}{\sqrt{\chi^2_{n-a}}} + \frac{\sqrt{\chi^2_{n-b}}}{\sqrt{\chi^2_{n-a}}}\frac{r}{\sqrt{1-r^2}} \tag{7.14}$$

and $\psi(X) = \frac{X}{\sqrt{1+X^2}}$, where as mentioned $Z \sim N(0,1)$ is the standard normal distribution, χ^2_{n-a} and χ^2_{n-b} are chi-squared distributions with the indicated degrees of freedom. The proofs of Equations

7.12, 7.13 and 7.14 are briefly explained in Appendix F.2. The most interesting prior of the form π_{ab} (besides the Jeffreys and independence Jeffreys priors) is the right-Haar prior $\pi_H = \pi_{12}$. Equations 7.8, 7.9 and 7.10 follow by substituting $a = 1$ and $b = 2$ in Equations 7.12, 7.13 and 7.14. It is shown by Berger & Sun (2008) that the Jeffreys rule prior $\pi_J = \pi_{10}$ is very often optimal while the independence Jeffreys prior $\pi_{IJ} = \pi_{21}$ is virtually never optimal. This is in contradiction to the common perception that the independence Jeffreys prior is better than the Jeffreys rule prior. It is also interesting to note that the Jeffreys rule posterior yields credible sets for σ_1, σ_2 that have exact frequentist coverage of $1-\beta$, while the right-Haar posterior distribution ($\pi_H = \pi_{12}$) on the other hand yields credible sets for σ_1 and ρ with the correct coverage. If this is the case, the fiducial/objective posterior distribution will be called exact frequentist matching. Berger & Sun (2008) also mentioned that it was not known that the right-Haar prior for ρ yields a posterior distribution that is exact frequentist matching (proved in Theorem 2 on page 980). "Indeed standard statistical software utilizes various approximations to arrive at frequentist confidence sets for ρ, missing the fact that a simple exact confidence set exists even for $n = 3$".

In Section 7.4 simulation studies will be performed for the priors π_J, π_{IJ}, π_H, π_{RO}, $\pi_{R\rho} \propto \frac{1}{\sigma_1\sigma_2(1-\rho^2)}$, $\pi_{R\sigma} \propto \frac{\sqrt{1+\rho^2}}{\sigma_1\sigma_2(1-\rho^2)}$ and $\pi_{MS} \propto \frac{1}{\sigma_1\sigma_2\sqrt{1-\rho^2}}$.

Coverage rate probabilities as well as interval lengths will be compared for the different posterior distributions and also the asymptotic normal distribution of $\hat{\alpha}$ derived by Van Zyl et al. (2000) and given in Equations 7.1 and 7.2.

The prior $\pi_{R\rho}$ was developed in Lindley (1965), using certain notions of transformations to constant information and studied extensively in Bayarri (1981), where it was shown to be a one-at-a-time reference prior. The second prior $\pi_{R\sigma}$ is derived in Berger & Sun (2008) where it was shown to be a reference prior for the parameter ordering $\{\sigma_1, \sigma_2, \rho\}$. The power of π_{MS} is smaller than that of $\pi_{R\rho}$ and $\pi_{R\sigma}$.

For the priors $\pi_{R\rho}$, $\pi_{R\sigma}$ and π_{MS} independent samples from their marginal posterior distributions $\pi(\sigma_1, \sigma_2, \rho|data)$ can be simulated by the following acceptance-rejection algorithm:

1. Generate $(\sigma_1, \sigma_2, \rho)$ from the independence Jeffreys posterior $\pi_{IJ}(\sigma_1, \sigma_2, \rho|data)$ and independently sample $u \sim uniform(0,1)$.

2. Suppose
$$M \equiv sup_{(\sigma_1,\sigma_2,\rho)} \frac{\pi(\sigma_1, \sigma_2, \rho)}{\pi_{IJ}(\sigma_1, \sigma_2, \rho)} < \infty.$$
If $u \leq \pi(\sigma_1, \sigma_2, \rho)/M\pi_{IJ}(\sigma_1, \sigma_2, \rho)$ accept $(\sigma_1, \sigma_2, \rho)$; If not, return to (1).

For each of the prior distributions listed in Table 7.1, the ratio π/π_{IJ} is listed in the table, along with the upper bound M, the rejection step and acceptance probability for $\rho = 0.8$, 0.95 and 0.99.

Table 7.1: Ratio π/π_{IJ}, Upper Bound M, Rejection Step and Acceptance Probability.

Priors	Ratio	Bound M	Rejection Step	Acceptance Probability		
				$\rho = 0.80$	$\rho = 0.95$	$\rho = 0.99$
$\pi_{R\rho}$	$\sqrt{1-\rho^2}$	1	$u \leq \sqrt{1-\rho^2}$	0.6000	0.3122	0.1410
$\pi_{R\sigma}$	$\sqrt{1-\rho^4}$	1	$u \leq \sqrt{1-\rho^4}$	0.7684	0.4307	0.1985
π_{MS}	$1-\rho^2$	1	$u \leq 1-\rho^2$	0.3600	0.0975	0.0199

Berger & Sun (2008) mentions that the algorithm is efficient for sampling these posteriors. For $\rho \approx 0$ the acceptance probability is close to one. They mention that for large $|\rho|$, the acceptance probabilities are reasonable for $\pi_{R\rho}$ and $\pi_{R\sigma}$ as evidenced by Table 7.1. For π_{MS} it is clear from Table 7.1 that the algorithm is less efficient for large $|\rho|$. Berger & Sun (2008) states that even though these acceptance rates are so low they may be fine in practice, given the simplicity of the algorithm. For further details see Berger & Sun (2008).

7.4 Simulation Study

A simulation study is done and coverage probabilities are obtained for Cronbach's alpha using the bivariate normal distribution. The values considered in the simulation are for sample sizes $n = 5, 10, 30$ and 50. These sample sizes represent small, medium and relatively large sample sizes. Suppose the exact covariance matrix Σ and sample size is given. The following represents the given covariance matrices for the simulation study.

$$\Sigma_1 = \begin{bmatrix} 1 & 1 \\ 1 & 4 \end{bmatrix} \qquad \Sigma_2 = \begin{bmatrix} 4 & 2 \\ 2 & 4 \end{bmatrix} \qquad \Sigma_3 = \begin{bmatrix} 5 & 6 \\ 6 & 8 \end{bmatrix}.$$

The parameter values considered for the bivariate normal distribution are $\mu_1 = 10$, $\mu_2 = 12$ for all three covariance matrices considered above. For Σ_1, the variance components are $\sigma_1 = 1$, $\sigma_2 = 2$ and $\rho = 0.5$ which corresponds to $\alpha = 0.5714$. For Σ_2, the variance components are $\sigma_1 = 2$, $\sigma_2 = 2$ and $\rho = 0.5$ which corresponds to $\alpha = 0.6667$. For Σ_3, the variance components are $\sigma_1 = \sqrt{5}$, $\sigma_2 = \sqrt{8}$ and $\rho = 0.9487$ which corresponds to $\alpha = 0.96$. Covariance matrices Σ_1 and Σ_3 have unequal variance components but equal covariances. Covariance matrix Σ_2 represents a compound symmetric matrix. The method used for simulating the coverage probabilities is as follows:

1. Simulate samples of size n from the bivariate normal distribution.

2. Iterate the process 10000 times by simulating Z values and χ^2_{n-a}, χ^2_{n-b} to determine the posterior distributions for σ_1, σ_2 and ρ. If the Jeffreys rule prior is used, $a = 1$ and $b = 0$, and for the independence Jeffreys prior, $a = 2$, $b = 1$. In the case of the right-Haar prior, $a = 1$, $b = 2$ and

π_{RO}, $a = b = 1$. For the priors $\pi_{R\rho}$, $\pi_{R\sigma}$ and π_{MS}, the simulation is done using the rejection sampling technique explained in Section 7.3.

3. Determine for each simulated posterior a 95% equal-tailed credibility interval.

4. The proportion of times that α is contained within these intervals is the coverage probability.

Tables 7.2, 7.4 and 7.6 summarize the coverage probabilities for covariance matrices Σ_1, Σ_2 and Σ_3 using $n = 5, 10, 30$ and 50. These tables also contain the coverage probabilities using the asymptotic interval of Van Zyl et al. (2000). Tables 7.3, 7.5 and 7.7 contain the average interval lengths for each prior and the frequentist asymptotic interval.

Table 7.2: Coverage Probabilities (in percentages) of Cronbach's Alpha for Σ_1 Using Different Priors and Sample Sizes.

Sample size	Priors							Asymptotic Interval
	π_J	π_{IJ}	π_H	π_{RO}	$\pi_{R\rho}$	$\pi_{R\sigma}$	π_{MS}	
$n = 5$	85.91	90.22	90.19	88.65	93.08	92.47	**93.33**	78.56
$n = 10$	92.09	93.54	93.29	92.99	**94.24**	94.01	94.07	87.59
$n = 30$	93.70	94.70	94.52	94.52	94.66	95.18	**94.96**	92.61
$n = 50$	94.65	94.54	94.60	**94.79**	94.65	94.66	94.61	93.27

Table 7.3: Average Interval Lengths of Cronbach's Alpha for Σ_1 Using Different Priors and Sample sizes.

Sample size	Priors							Asymptotic Interval
	π_J	π_{IJ}	π_H	π_{RO}	$\pi_{R\rho}$	$\pi_{R\sigma}$	π_{MS}	
$n = 5$	2.9140	3.8260	**2.5860**	2.8064	3.0701	3.2124	2.6203	1.6880
$n = 10$	**1.1730**	1.2778	1.2101	1.1837	1.2466	1.2581	1.2473	0.9027
$n = 30$	**0.5201**	0.5329	0.5324	0.5216	0.5336	0.5302	0.5380	0.4777
$n = 50$	**0.3833**	0.3878	0.3888	0.3863	0.3908	0.3880	0.3936	0.3654

For $n = 5$, π_{MS} has coverage closest to the nominal level of 0.95. The priors $\pi_{R\rho}$ and $\pi_{R\sigma}$ also performed well. The asymptotic interval performed the worst with coverage probabilities far below 0.95. For $n = 10$, the one at a time reference prior, $\pi_{R\rho}$ obtained a coverage rate closest to 0.95 and therefore performed the best. All three priors, $\pi_{R\rho}$,$\pi_{R\sigma}$ and π_{MS} performed similarly. For $n = 30$, the coverage probabilities improved for all the priors considered and the asymptotic interval. The prior, π_{MS} obtained a coverage rate closest to 0.95 and therefore performed the best, followed by $\pi_{R\sigma}$ and Jeffreys independence prior, π_{IJ}. The asymptotic interval improved but is still outperformed by the various priors considered. For $n = 50$, the coverage probabilities also improved for all the priors and the asymptotic interval. The prior, π_{RO} obtained a coverage rate closest to 0.95 and therefore performed the best. The asymptotic interval obtained the shortest interval lengths for all the sample sizes considered in the simulation study. The right-Haar prior obtained the shortest interval for $n = 5$

among all the priors considered. The Jeffreys prior obtained the shortest interval lengths for $n = 10, 30$ and 50 for all the priors considered.

Table 7.4: Coverage Probabilities (in percentages) of Cronbach's Alpha for Σ_2 Using Different Priors and Sample Sizes.

Sample size	Priors							Asymptotic Interval
	π_J	π_{IJ}	π_H	π_{RO}	$\pi_{R\rho}$	$\pi_{R\sigma}$	π_{MS}	
$n = 5$	85.97	90.37	92.15	89.53	92.34	92.82	**92.90**	78.71
$n = 10$	91.97	93.56	93.66	93.28	94.63	**94.66**	94.35	85.92
$n = 30$	93.89	94.51	**94.88**	94.70	94.71	94.59	**94.88**	91.53
$n = 50$	94.51	94.55	94.56	94.85	**95.06**	95.25	94.53	93.09

Table 7.5: Average Interval Lengths of Cronbach's Alpha for Σ_2 Using Different Priors and Sample Sizes.

Sample size	Priors							Asymptotic Interval
	π_J	π_{IJ}	π_H	π_{RO}	$\pi_{R\rho}$	$\pi_{R\sigma}$	π_{MS}	
$n = 5$	3.5897	4.8759	3.6806	3.8193	3.8534	4.3495	**3.1652**	1.9913
$n = 10$	**1.3626**	1.4882	1.4648	1.4039	1.4680	1.4607	1.4520	1.0009
$n = 30$	**0.5605**	0.5750	0.5790	0.5638	0.5783	0.5713	0.5853	0.49979
$n = 50$	**0.4048**	0.4091	0.4156	0.4104	0.4142	0.4110	0.4199	0.3820

For $n = 5$, π_{MS} has coverage closest to the nominal level of 0.95. The priors $\pi_{R\rho}$ and $\pi_{R\sigma}$ also performed well. The asymptotic interval performed the worst with coverage probabilities far below 0.95. For $n = 10$, $\pi_{R\sigma}$ obtained a coverage rate closest to 0.95 and therefore performed the best. All three priors, $\pi_{R\rho}$,$\pi_{R\sigma}$ and π_{MS} performed similarly. For $n = 30$, the coverage probabilities improved for all the priors considered and the asymptotic interval. The prior, π_{MS} and the right-Haar prior obtained a coverage rate closest to 0.95 and therefore performed the best, followed by $\pi_{R\rho}$ and π_{RO}. The Jeffreys rule prior performed the worst in terms of the coverage rate obtained for each prior. The asymptotic interval improved but is still outperformed by the various priors considered. For $n = 50$, the coverage probabilities also improved for all the priors and the asymptotic interval. The prior, $\pi_{R\rho}$ obtained a coverage rate closest to 0.95 and therefore performed the best. The asymptotic interval obtained the shortest interval lengths for all the sample sizes considered in the simulation study. The prior π_{MS} obtained the shortest interval for $n = 5$ among all the priors considered. The Jeffreys prior obtained the shortest interval lengths for $n = 10, 30$ and 50 from all the priors considered.

Table 7.6: Coverage Probabilities (in percentages) of Cronbach's Alpha for Σ_3 Using Different Priors and Sample Sizes.

Sample size	Priors							Asymptotic Interval
	π_J	π_{IJ}	π_H	π_{RO}	$\pi_{R\rho}$	$\pi_{R\sigma}$	π_{MS}	
$n = 5$	86.57	90.23	88.73	87.78	**91.15**	90.82	86.08	78.17
$n = 10$	92.16	**93.87**	91.87	92.48	93.03	93.44	91.54	86.89
$n = 30$	93.58	94.45	94.20	94.09	**94.74**	94.32	93.75	92.26
$n = 50$	94.33	**94.86**	94.60	94.66	94.74	94.67	93.91	94.05

Table 7.7: Average Interval Lengths of Cronbach's Alpha for Σ_3 Using Different Priors and Sample Sizes.

Sample size	Priors							Asymptotic Interval
	π_J	π_{IJ}	π_H	π_{RO}	$\pi_{R\rho}$	$\pi_{R\sigma}$	π_{MS}	
$n = 5$	**0.3008**	0.3909	0.5680	0.4111	0.5453	0.5058	0.7275	0.1698
$n = 10$	**0.1175**	0.1275	0.1623	0.1373	0.1579	0.1490	0.1891	0.0894
$n = 30$	**0.0519**	0.0523	0.0560	0.0534	0.0555	0.0554	0.0588	0.0477
$n = 50$	**0.0381**	0.0386	0.0399	0.0391	0.0398	0.0396	0.0411	0.0363

For $n = 5$, $\pi_{R\rho}$ has coverage closest to the nominal level of 0.95. The priors π_{IJ} and $\pi_{R\sigma}$ also performed well. The asymptotic interval performed the worst with coverage probabilities far below 0.95. For $n = 10$, the Jeffreys independence prior obtained a coverage rate closest to 0.95 and therefore performed the best. The prior $\pi_{R\sigma}$ obtained the second-best coverage rate and the one at a time reference prior, $\pi_{R\rho}$ got the third best coverage rate. For $n = 30$, the coverage probabilities improved for all the priors considered and asymptotic interval. The prior, $\pi_{R\rho}$ obtained a coverage rate closest to 0.95 and therefore performed the best, followed by π_{IJ} and $\pi_{R\sigma}$. The Jeffreys rule prior performed the worst in terms of the coverage rate obtained for each prior. The asymptotic interval improved but is still outperformed by the various priors considered. For $n = 50$, the coverage probabilities also improved for all the priors and the maximum likelihood estimate. The prior, π_{IJ} obtained a coverage rate closest to 0.95 and therefore performed the best. The asymptotic interval obtained the shortest interval lengths for all the sample sizes considered in the simulation study. The Jeffreys prior obtained the shortest interval lengths for all the sample sizes considered in the simulation study.

7.5 Conclusion

In this chapter a number of objective priors were considered in the Bayesian analysis of Cronbach's alpha for a bivariate normal distribution under a general covariance matrix. The asymptotic interval of Cronbach's alpha was also considered for a general covariance matrix. A simulation study was conducted for various sample sizes where the coverage probabilities and average interval lengths were computed for the posterior distributions considering all of the objective priors as well as the asymptotic interval. For the covariance matrices considered in the simulation study, it was evident that the objective priors outperformed the asymptotic interval where their coverage probabilities were all closer to the nominal rate of 0.95. The asymptotic interval underestimated the coverage probabilities for all the sample sizes considered in the simulation study. The performance of the asymptotic interval was especially bad for small sample sizes. This is not surprising since the interval was derived by Van Zyl et al. (2000) for large sample sizes using asymptotic results. The average interval lengths for the asymptotic interval were the shortest for all covariance matrices considered but with a shortcoming of not getting coverage probabilities closer to the nominal rate. The Jeffreys rule prior also obtained the shortest average interval lengths between all the priors considered but also fell short with the coverage probabilities being below the nominal level. In practice, for small sample sizes, it is not recommended to use the asymptotic interval for alpha. For small sample sizes, it is recommended to use either the one-at-a-time reference prior, $\pi_{R\rho}$, or the reference prior with group ordering $\{\sigma_1, \sigma_2, \rho\}$, $\pi_{R\sigma}$, or the prior π_{MS}. For larger sample sizes, all the objective priors considered performed similarly. It is recommended to also use $\pi_{R\rho}$, $\pi_{R\sigma}$ or π_{MS} for larger sample sizes although any of the priors considered may also be considered. Even though the asymptotic interval improved for larger sample sizes, the Bayesian approach still consistently outperformed the asymptotic frequentist interval.

Chapter 8

Estimating Cronbach's α and ρ under Compound Symmetry

8.1 Introduction

In this chapter, distributions for Cronbach's coefficient alpha and the intra-class correlation will be derived in the case of single coefficients as well as for the case when there are two coefficients, using a one-way random effects model. Since there is a mathematical relationship between Cronbach's alpha and the intra-class correlation coefficient, the work that will be done in this chapter is an extension of the research done by Koning & Franses (2003) on Cronbach's coefficient alpha and the research done by Chung & Dey (1998) on the intra-class correlation coefficient. This work is an extension of the work done in Chapter 3 since it involves inference on two alpha coefficients. Recall from Chapter 3 that the variance-covariance matrix for the one-way random effects model is given by

$$\Sigma = \begin{bmatrix} \sigma_1^2+\sigma_2^2 & \sigma_2^2 & \sigma_2^2 & \dots & \sigma_2^2 \\ \sigma_2^2 & \sigma_1^2+\sigma_2^2 & \sigma_2^2 & \dots & \sigma_2^2 \\ \sigma_2^2 & \dots & \sigma_1^2+\sigma_2^2 & \cdots & \sigma_2^2 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \sigma_2^2 & \sigma_2^2 & \sigma_2^2 & \cdots & \sigma_1^2+\sigma_2^2 \end{bmatrix}_{J\times J} . \tag{8.1}$$

It was also shown in Chapter 3 that for the one-way random effects model that Cronbach's alpha and the intra-class correlation is given by

$$\alpha = \frac{J\sigma_2^2}{\sigma_1^2+J\sigma_2^2}$$

and

$$\rho = \frac{\sigma_2^2}{\sigma_1^2+\sigma_2^2},$$

respectively. Using the well-known Spearman-Brown formula, Cronbach's alpha can be written in terms of the intra-class correlation coefficient given by $\alpha = \frac{J\rho}{1+(J-1)\rho}$, or equivalently, the intra-class correlation coefficient can be written in terms of Cronbach's alpha using the equation $\rho = \frac{\alpha}{J-\alpha(J-1)}$. These results will prove to be very useful in this chapter in deriving distributions for Cronbach's alpha and the intra-class correlation coefficient using a pivotal quantity $F = \frac{1-\hat{\alpha}}{1-\alpha}$. The main reason why this method works so well is because of the compound symmetrical structure of the covariance matrix of a one-way random effects model since the Spearman-Brown formula only works under compound symmetry.

8.2 The Distributions of Cronbach's Coefficient Alpha in the Case of the Balanced Random Effects Model

8.2.1 Inference for a Single Cronbach's Coefficient Alpha

It is well known from classical statistics that if σ_1^2 and σ_2^2 are known for the random effects model then $\frac{v_1 m_1}{\sigma_1^2} \sim \chi_{v_1}^2$ and $\frac{v_2 m_2}{\sigma_1^2 + J\sigma_2^2} \sim \chi_{v_2}^2$ where $v_1 m_1$ and $v_2 m_2$ are defined in Section 4.2. Also $v_1 = I(J-1)$ and $v_2 = I - 1$. From this it follows that $m_1 \sim \sigma_1^2 \frac{\chi_{v_1}^2}{v_1}$ and $m_2 \sim \left(\sigma_1^2 + J\sigma_2^2\right) \frac{\chi_{v_2}^2}{v_2}$. Since

$$\begin{aligned}\hat{\alpha} &= 1 - \frac{m_1}{m_2} = 1 - \frac{\sigma_1^2}{\sigma_1^2 + J\sigma_2^2} \frac{\chi_{v_1}^2 / v_1}{\chi_{v_2}^2 / v_2} \\ &= 1 - (1-\alpha) F_{v_1, v_2},\end{aligned}$$

it follows that

$$F_{v_1, v_2} = \frac{1-\hat{\alpha}}{1-\alpha} \tag{8.2}$$

has an $F-$ distribution with $I(J-1)$ and $(I-1)$ degrees of freedom. See also Van Zyl et al. (2000), Kristof (1963), Feldt (1965), Koning & Franses (2003), Duhachek & Iacobucci (2004) and Cui & Li (2012). Equation 8.2 is therefore a pivotal quantity for α and can be used to obtain accurate and precise confidence interval estimates for α as follows:

$$P\left(F_L < \frac{1-\hat{\alpha}}{1-\alpha} < F_U\right) = 1 - \beta$$

and therefore

$$P\left(\frac{1}{F_U} < \frac{1-\alpha}{1-\hat{\alpha}} < \frac{1}{F_L}\right) = 1 - \beta$$

which means that

$$P\left(1-\frac{(1-\hat{\alpha})}{F_L}<\alpha<1-\frac{(1-\hat{\alpha})}{F_U}\right)=1-\beta.$$

From Equation 8.2 it follows that

$$\begin{aligned}\alpha &= \frac{\hat{\alpha}-1}{F}+1 \\ &= (\hat{\alpha}-1)\tilde{F}+1 \qquad (8.3)\end{aligned}$$

where $\tilde{F} \sim F_{(I-1),I(J-1)}$. Equation 8.3 is the posterior distribution of α and it illustrates the uncertainty about the true value of the parameter α. Another method to prove that Equation 8.3 is in fact a posterior distribution is the following: It was proved in Equations 6.3 and 6.4 that the posterior distribution $\pi(\sigma_1^2, \sigma_1^2+J\sigma_2^2|data) = \pi(\sigma_1^2|data)\pi(\sigma_1^2+J\sigma_2^2|data)$. If the Jeffreys independence prior $\pi\left(\sigma_1^2,\sigma_2^2\right) \propto \left(\sigma_1^2\right)^{-1}\left(\sigma_1^2+J\sigma_2^2\right)^{-1}$ is used then the posterior distributions of σ_1^2 and $\sigma_1^2+J\sigma_2^2$ are inverse gamma distributions and it follows that

$$\sigma_1^2 \sim \frac{v_1 m_1}{\chi^2_{v_1}} \quad \text{and} \quad \sigma_1^2+J\sigma_2^2 \sim \frac{v_2 m_2}{\chi^2_{v_2}}.$$

The posterior distribution of α can therefore be easily derived. Since

$$\alpha = 1-\frac{\frac{v_1 m_1}{\chi^2_{v_1}}}{\frac{v_2 m_2}{\chi^2_{v_2}}} = 1-\frac{m_1}{m_2}\frac{\chi^2_{v_2}/v_2}{\chi^2_{v_1}/v_1},$$

it follows that

$$\alpha|\hat{\alpha} = 1-(1-\hat{\alpha})F_{v_2,v_1},$$

which is identical to Equation 8.3 because

$$\hat{\alpha} = 1-\frac{m_1}{m_2}, \qquad v_2 = I-1 \text{ and } v_1 = I(J-1).$$

The following theorem can now be proved.

Theorem 8.1. *The posterior distribution of* $\alpha = (\hat{\alpha}-1)F_{(I-1),I(J-1)}+1$ *is given by*

$$\pi(\alpha|\hat{\alpha}) = K\left(\frac{1}{1-\hat{\alpha}}\right)\left(\frac{1-\alpha}{1-\hat{\alpha}}\right)^{\frac{(I-3)}{2}}\left[1+\frac{(I-1)(1-\alpha)}{I(J-1)(1-\hat{\alpha})}\right]^{-\frac{1}{2}(IJ-1)}$$

where $K = \dfrac{\Gamma\left(\frac{(IJ-1)}{2}\right)}{\Gamma\left(\frac{(I-1)}{2}\right)\Gamma\left[\frac{I(J-1)}{2}\right]}(I-1)^{\frac{(I-1)}{2}}\left[I(J-1)\right]^{-\frac{(I-1)}{2}}.$

Proof. The proof is given in Appendix F.1.

Corollary 8.2. *The mean, variance and mode of* α *is given by*

$$E(\alpha|\hat{\alpha}) = (\hat{\alpha}-1)\left[\frac{I(J-1)}{I(J-1)-2}\right]+1 \tag{8.4}$$

$$Var(\alpha|\hat{\alpha}) = (\hat{\alpha}-1)^2 \frac{2I^2(J-1)^2(IJ-3)}{(I-1)[I(J-1)-2]^2[I(J-1)-4]}, \tag{8.5}$$

and

$$Mode(\alpha|\hat{\alpha}) = 1-\frac{I(I-3)(J-1)(1-\hat{\alpha})}{[(IJ-1)(I-1)-(I-3)(I-1)]}. \tag{8.6}$$

□

8.2.2 Example

Koning & Franses (2003) considered the following two studies. Study A consisted of 100 subjects $(I_1 = 100)$ and the number of items per subject were 4 $(J_1 = 4)$. The estimated Cronbach's alpha, $\hat{\alpha} = 0.7$. Study B also used 100 subjects $(I_2 = 100)$ but the number of items per subject were 6 $(J_2 = 6)$ and the estimated Cronbach's alpha, $\hat{\alpha}_2 = 0.8$. The probability density functions $\pi(\alpha_1|\hat{\alpha}_1)$ and $\pi(\alpha_2|\hat{\alpha}_2)$ are displayed in Figure 8.1. Using Equations 8.4 and 8.5 it follows that $E(\alpha_1|\hat{\alpha}_1) = 0.6980$ and $Var(\alpha_1|\hat{\alpha}_1) = 0.0025$ for study A. Also, the $E(\alpha_2|\hat{\alpha}_2) = 0.7992$ and $Var(\alpha_2|\hat{\alpha}_2) = 0.00098$ for study B. It is clear from Figure 8.1 that the numerical values are for all practical purposes the same as the theoretical values.

Table 8.1: Summary Statistics for α_1 and α_2.

	Mean	Median	Mode	Variance	Approximate 95% Interval	Exact 95% Interval
α_1	0.6980	0.7008	0.7080	0.0025	(0.591; 0.785)	(0.5910; 0.7857)
α_2	0.7992	0.8005	0.805	0.00098	(0.732, 0.855)	(0.7325, 0.8551)

Since the sample sizes are large, the mean, median and mode for the first sample are for all practical purposes 0.7. This is also the case for the second sample where the central values are 0.8. From the table it is also clear that the 95% confidence intervals for α calculated by using numerical integration and those obtained from the pivotal quantity are the same. These intervals are accurate and precise, that is, they have the correct coverage properties.

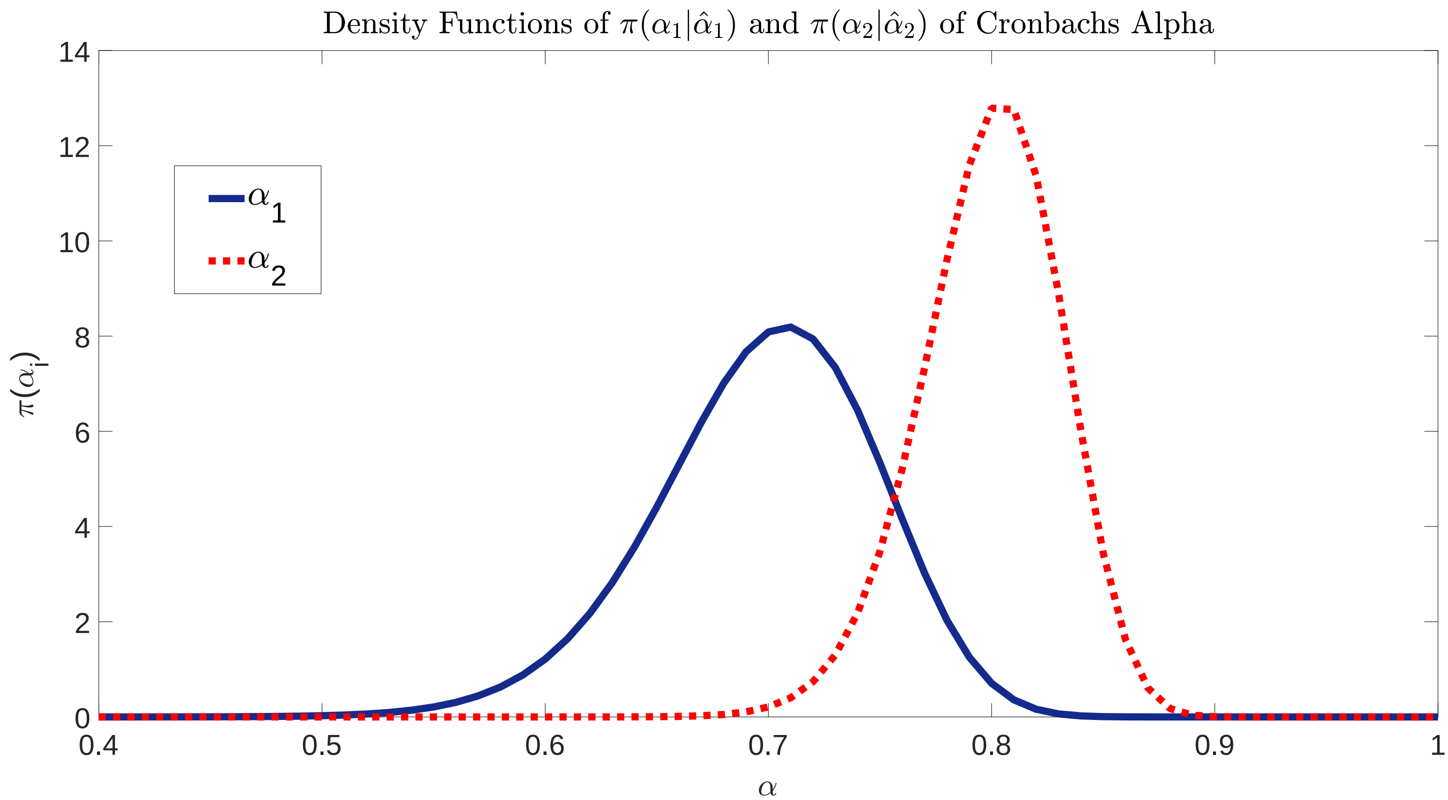


Figure 8.1: Posterior Density Functions $\pi(\alpha_1|\hat{\alpha}_1)$ and $\pi(\alpha_2|\hat{\alpha}_2)$ of Cronbach's Alpha.

8.2.3 Inference for Two Cronbach's Alpha Coefficients

Koning & Franses (2003) were also interested in testing $H_0: \alpha_1 = \alpha_2$ vs $H_a: \alpha_1 \neq \alpha_2$ at $\beta = 0.05$, that is, whether there is a significant difference between α_1 and α_2. The testing procedure can be done in the following way:

Since $\frac{\hat{\alpha}-1}{F}+1$ where $F \sim F_{I(J-1),(I-1)}$, α_1 and α_2 can easily be simulated.

- $\alpha_1 = \frac{0.7-1}{F_{300,99}} + 1$ and $\alpha_2 = \frac{0.8-1}{F_{500,99}} + 1$.
- Simulate α_1 and α_2 a large number of times- say a 100 000 times.
- At each iteration, calculate $\delta_\alpha = \frac{\alpha_2}{\alpha_1}$ and $\gamma_\alpha = \alpha_2 - \alpha_1$.

The summary statistics for δ_α using the data in the Example from Section 8.2.2 is given in Table 8.2.

Table 8.2: Summary Statistics for δ_α.

$Mean(\delta_\alpha)$	$Median(\delta_\alpha)$	$Mode(\delta_\alpha)$	$Var(\delta_\alpha)$	95% Interval	$P(\delta_\alpha > 1)$
1.1516	1.1418	1.125	0.0096	(0.9878; 1.3732)	0.9635

Since one is included in the 95% confidence interval, $H_0: \alpha_1 = \alpha_2$ will not be rejected at the $\beta = 0.05$ level. If we were interested in the alternative hypothesis $H_a: \alpha_1 < \alpha_2$ then H_0 would have been rejected. As mentioned $\gamma_\alpha = \alpha_2 - \alpha_1$ can also be used to test H_0. Now, $E(\gamma_\alpha|\hat{\alpha}_1, \hat{\alpha}_2) = E(\alpha_2|\hat{\alpha}_2) - E(\alpha_1|\hat{\alpha}_1) = 0.1012$ and $Var(\gamma_\alpha|\hat{\alpha}_1, \hat{\alpha}_2) = Var(\alpha_2|\hat{\alpha}_2) + Var(\alpha_1|\hat{\alpha}_1) = 0.0035$. The summary statistics for γ_α are given in Table 8.3.

Table 8.3: Summary Statistics for γ_α.

$Mean(\gamma_\alpha)$	$Median(\gamma_\alpha)$	$Mode(\gamma_\alpha)$	$Var(\gamma_\alpha)$	95% Interval	$P(\gamma_\alpha > 0)$
0.1015	0.0996	0.096	0.0035	$(-0.0093; 0.2230)$	0.9636

In Figure 8.2 the distribution of δ_α is illustrated.

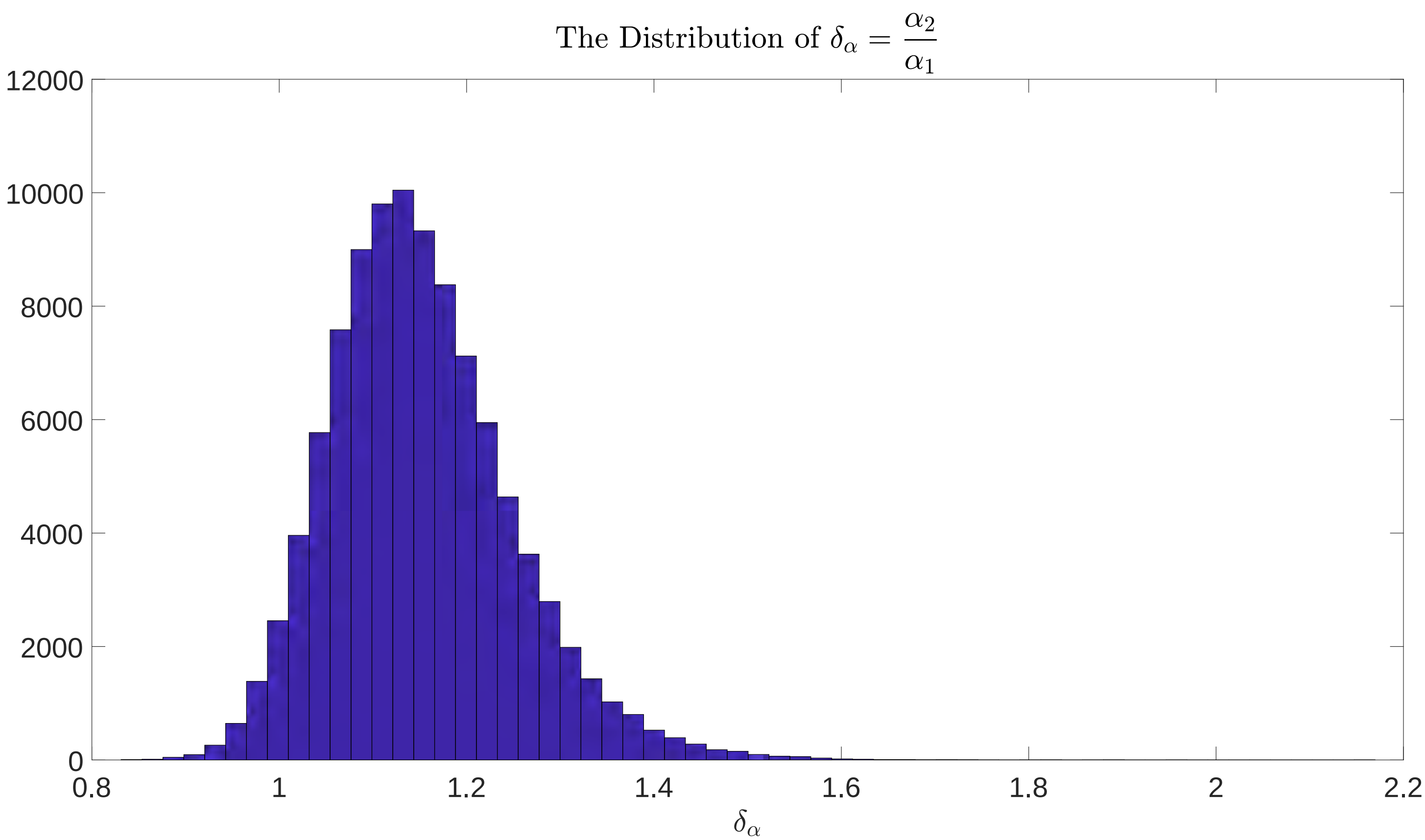


Figure 8.2: The Posterior Distribution of $\delta_\alpha = \frac{\alpha_2}{\alpha_1}$.

Since zero is included in the 95% confidence interval $H_0: \alpha_1 = \alpha_2$ will not be rejected. It is also clear from Figures 8.2 and 8.3 that $P(\delta_\alpha > 1) = P(\gamma_\alpha > 0) = 0.964$.

In Figure 8.3 the posterior distribution of γ_α is illustrated.

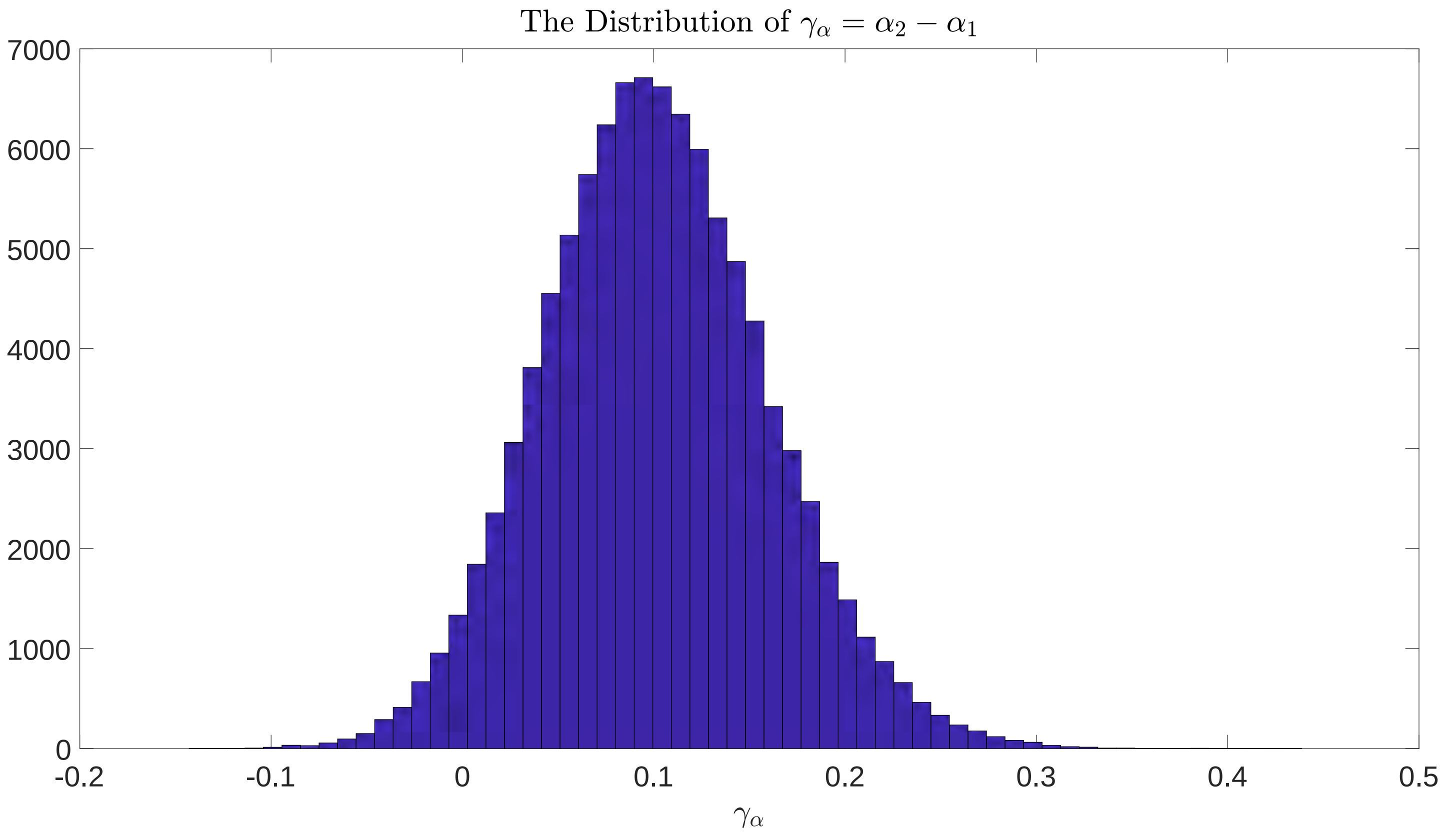


Figure 8.3: The Posterior Distribution of $\gamma = \alpha_2 - \alpha_1$.

Another statistic that can be used for testing $H_0: \alpha_1 = \alpha_2$ vs $H_a: \alpha_1 \neq \alpha_2$ is Ω where

$$\Omega = \frac{1-\alpha_1}{1-\alpha_2} = \frac{(1-\hat{\alpha}_1)\, F_{I_2(J_2-1),(I_2-1)}}{(1-\hat{\alpha}_2)\, F_{I_1(J_1-1),(I_1-1)}}. \tag{8.7}$$

The following theorem will first be proved.

Theorem 8.3. *The mean and variance of Ω are given by*

$$E\left(\Omega|\hat{\alpha}_1,\hat{\alpha}_2\right) = \frac{(1-\hat{\alpha}_1)}{(1-\hat{\alpha}_2)}\left(\frac{I_2-1}{I_2-3}\right)\left(\frac{I_1\,(J_1-1)}{I_1\,(J_1-1)-2}\right) \tag{8.8}$$

and

$$Var\left(\Omega|\hat{\alpha}_1,\hat{\alpha}_2\right) = Var\left\{E\left(\Omega|F_{I_2(J_2-1),(I_2-1)}\right)\right\} + E\left\{Var\left(\Omega|F_{I_2(J_2-1),(I_2-1)}\right)\right\}, \tag{8.9}$$

where

$$Var\left\{E\left(\Omega|F_{I_2(J_2-1),(I_2-1)}\right)\right\} = \frac{(1-\hat{\alpha}_1)^2}{(1-\hat{\alpha}_2)^2}\left(\frac{I_1\,(J_1-1)}{I_1\,(J_1-1)-2}\right)^2 \frac{2\,(I_2-1)^2\left\{I_2\,(J_2-1)+(I_2-3)\right\}}{I_2\,(J_2-1)\,(I_2-3)^2\,(I_2-5)},$$

and

$$E\left\{Var\left(\Omega|F_{I_2(J_2-1),(I_2-1)}\right)\right\} = \frac{(1-\hat{\alpha}_1)^2}{(1-\hat{\alpha}_2)^2}\left\{\frac{2\{I_1(J_1-1)\}^2\{(I_1J_1-3)\}}{(I_1-1)\{I_1(J_1-1)-2\}^2\{I_1(J_1-1)-4\}}\right\}$$
$$\times\left\{\frac{2(I_2-1)^2\{I_2(J_2-1)+I_2-3\}}{I_2(J_2-1)(I_2-3)^2(I_2-5)}+\left(\frac{I_2-1}{I_2-3}\right)^2\right\}.$$

Proof. The proof is given in Appendix F.3. □

If $I_1 = 100$, $J_1 = 4$, $\hat{\alpha}_1 = 0.7$, $I_2 = 100$, $J_2 = 6$ and $\hat{\alpha}_2 = 0.8$ are substituted into Equations 8.8 and 8.9, it follows that $E(\Omega|\hat{\alpha}_1,\hat{\alpha}_2) = 1.5413$ and $Var(\Omega|\hat{\alpha}_1,\hat{\alpha}_2) = 0.1260$. The summary statistics for Ω is given in Table 8.4.

Table 8.4: Summary Statistics for Ω.

$Mean(\Omega)$	$Median(\Omega)$	$Mode(\Omega)$	$Var(\Omega)$	95% Interval	$P(\Omega > 1)$
1.5407	1.5013	1.45	0.1257	(0.9616; 2.3443)	0.9632

It can be seen that the theoretical values and numerical values for the means and variances are for all practical purposes the same. Since one is included in the 95% interval H_0 will not be rejected. It is also clear that

$$P(\delta > 1) = P(\gamma > 0) = P(\Omega > 1) = 0.96.$$

Koning & Franses (2003) claimed that the 95% confidence interval for $log(\Omega)$ does not contain the value zero, indicating that there is a significant difference between α_1 and α_2. According to our simulation study the 95% confidence interval is $(-0.0381; 0.8504)$ which contains zero. From this it follows that $e^{-0.0381} = 0.9626$ and $e^{0.8504} = 2.3406$ which corresponds to the 95% confidence interval for Ω. It seems that Koning & Franses (2003) made a mistake with respect to the number of degrees of freedom in the $F-$ distribution. For example they used $n_i = 100$, n_i is in fact, $I_i - 1$ which is 99. In the next section inferences about the intra-class correlation coefficient ρ will be considered. In Figure 8.4 the distribution of Ω is illustrated.

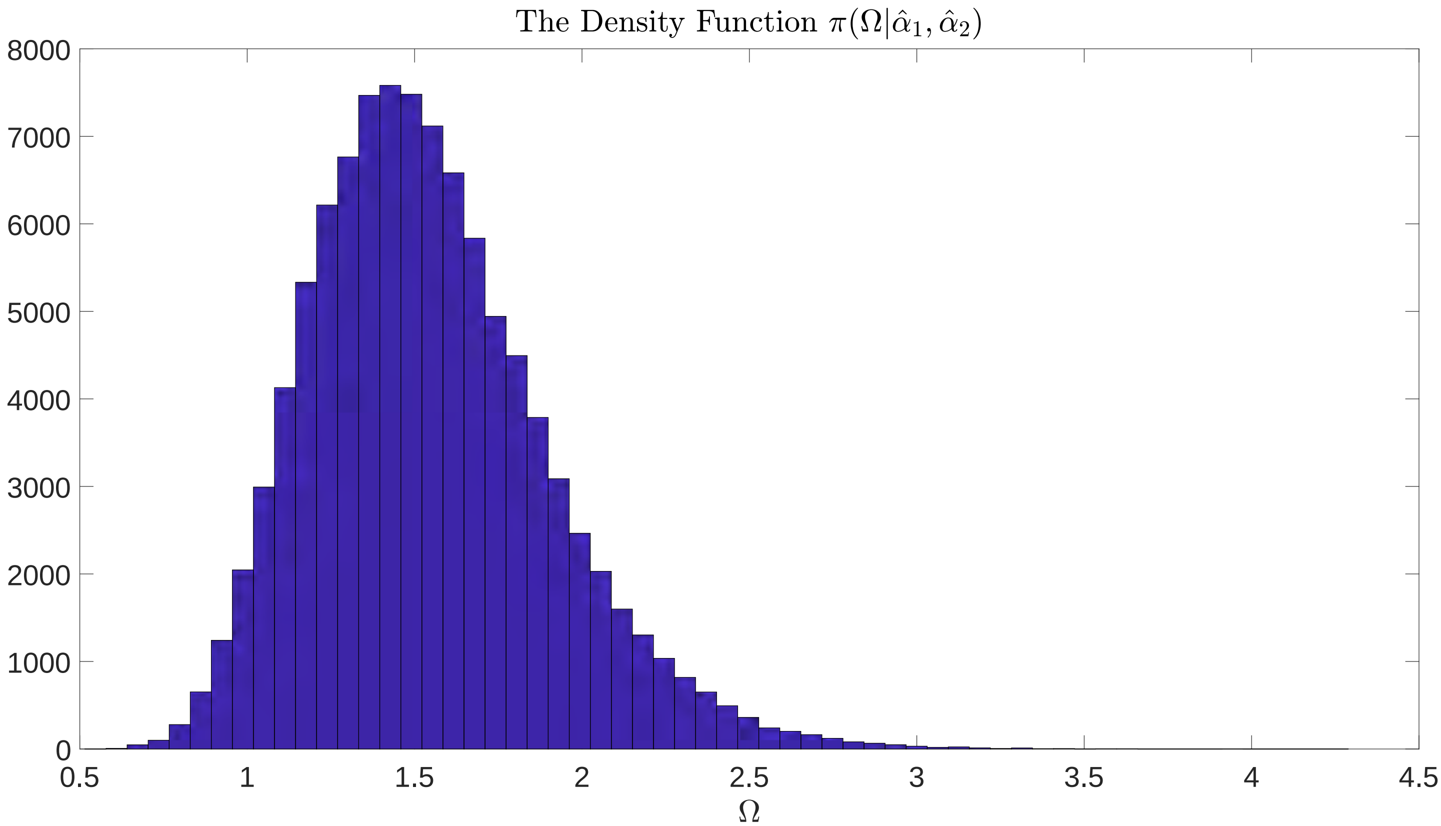


Figure 8.4: The Density Function $\pi(\Omega|\hat{\alpha}_1, \hat{\alpha}_2)$.

8.3 The Posterior Distributions of the Intra-Class Correlation Coefficient in the Case of the Balanced Random Effects Model

8.3.1 Inference for a Single Intra-Class Correlation Coefficient

Confidence interval estimates for ρ can be derived using the pivotal quantity in Equation 8.2 and the confidence interval for Cronbach's alpha that was derived in Section 8.2.

$$P\left(1-\frac{(1-\hat{\alpha})}{F_L} < \alpha < 1-\frac{(1-\hat{\alpha})}{F_U}\right) = 1-\beta.$$

Since $\alpha = \frac{J\rho}{1+(J-1)\rho}$ it also follows that

$$P\left(1-\frac{(1-\hat{\alpha})}{F_L} < \frac{J\rho}{1+(J-1)\rho} < 1-\frac{(1-\hat{\alpha})}{F_U}\right) = 1-\beta.$$

And therefore

$$P\left(\left(\frac{JF_L}{F_L-(1-\hat{\alpha})}-(J-1)\right)^{-1}<\rho<\left(\frac{JF_U}{F_U-(1-\hat{\alpha})}-(J-1)\right)\right)^{-1}=1-\beta$$

which is an exact confidence interval for ρ. F_U and F_L are the $100(\beta/2)$ and $100(1-\beta/2)$ percentile points of the central $F-$ distribution with $I(J-1)$ and $(I-1)$ degrees of freedom. The distribution of ρ, the intra-class correlation coefficient is given in Theorem 8.4.

Theorem 8.4. *The posterior distribution of* $\rho=\frac{\alpha}{J-\alpha(J-1)}$ *is given by*

$$\begin{aligned}\pi(\rho|\hat{\alpha}) &= K\left[\frac{1-\rho}{1+\rho(J-1)}\right]^{\frac{1}{2}(I-3)}\left[1+\frac{(I-1)}{I(J-1)}\frac{1}{(1-\hat{\alpha})}\frac{(1-\rho)}{[1+\rho(J-1)]}\right]^{-\frac{1}{2}(IJ-1)}\\ &\times \frac{J}{[1+\rho(J-1)]^2}\left(\frac{1}{1-\hat{\alpha}}\right)^{\frac{1}{2}(I-1)}.\end{aligned}$$

Proof. The proof is given in Appendix F.2. □

By using the data given in Example 8.2.2, the posterior density functions $\pi(\rho_1|\hat{\alpha}_1)$ and $\pi(\rho_2|\hat{\alpha}_2)$ can be calculated and are illustrated in Figure 8.5.

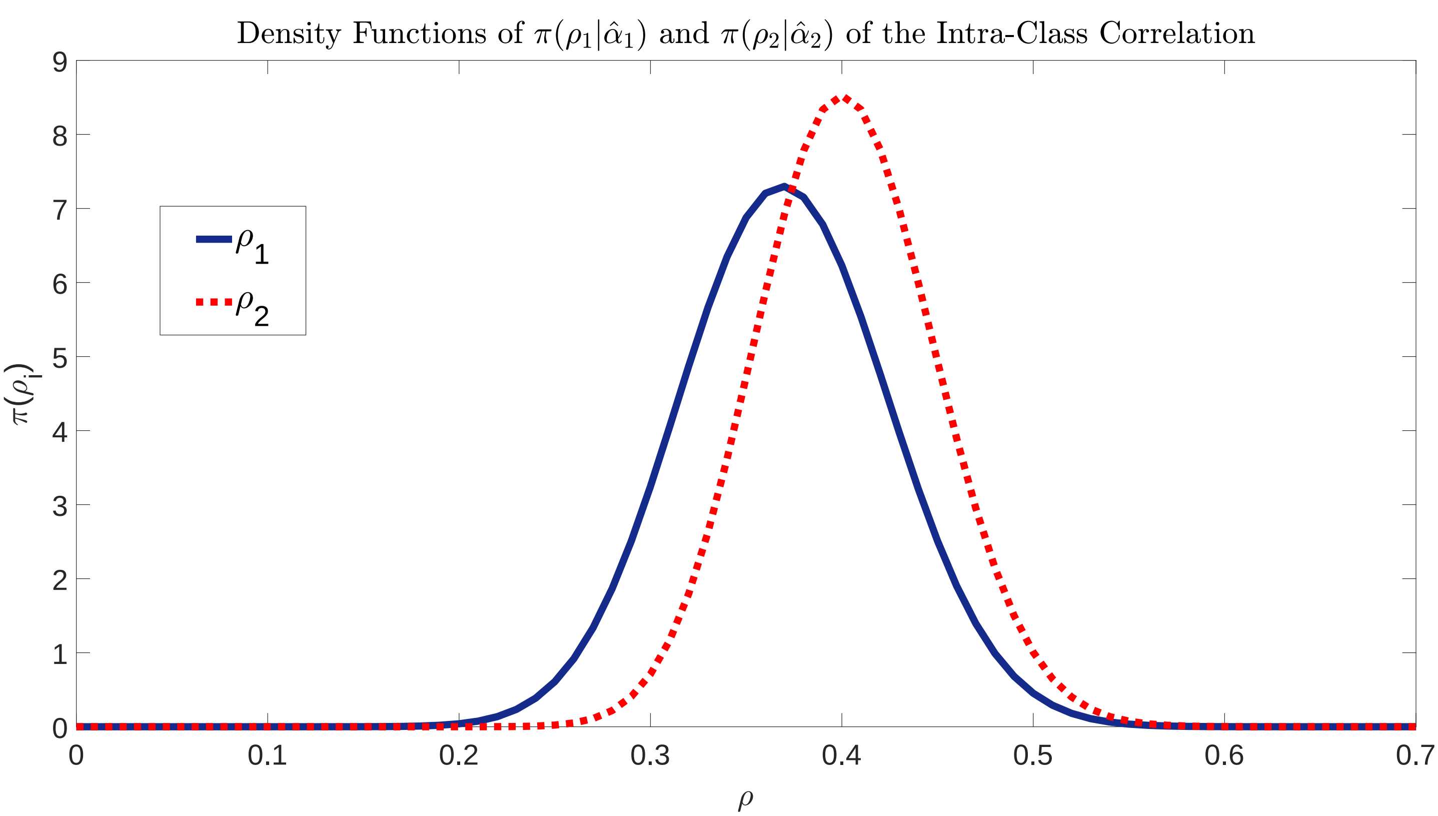


Figure 8.5: Posterior Density Functions $\pi(\rho_1|\hat{\alpha}_1)$ and $\pi(\rho_2|\hat{\alpha}_2)$ of the Intra-Class Correlation Coefficient.

Table 8.5: Summary Statistics for ρ_1 and ρ_2.

	Mean	Median	Mode	Variance	Approximate 95% Interval	Exact 95% Interval
ρ_1	0.3704	0.3695	0.369	0.0030	(0.269; 0.478)	(0.2717; 0.4738)
ρ_2	0.4024	0.4011	0.400	0.0022	(0.313, 0.495)	(0.3134, 0.4959)

From Table 8.5 it can be seen that the 95% confidence intervals for ρ, calculated using numerical integration and those obtained from the pivotal quantity are for all practical purposes the same.

8.3.2 Inference for Two Intra-Class Correlation Coefficients

If we are interested in testing $H_0: \rho_1 = \rho_2$ vs $H_a: \rho_1 \neq \rho_2$ then we can proceed as follows:

- As indicated before α_1 and α_2 can easily be simulated using the equation $\alpha = \frac{\hat{\alpha}-1}{F} + 1$.
- Since $\rho = \frac{\alpha}{J-\alpha(J-1)}$, ρ_1 and ρ_2 can easily be obtained.
- Simulate ρ_1 and ρ_2 a large number of times.
- At each iteration, calculate $\delta_\rho = \frac{\rho_2}{\rho_1}$ and $\gamma_\rho = \rho_2 - \rho_1$.

In Figure 8.6 the posterior distribution of δ_ρ is illustrated.

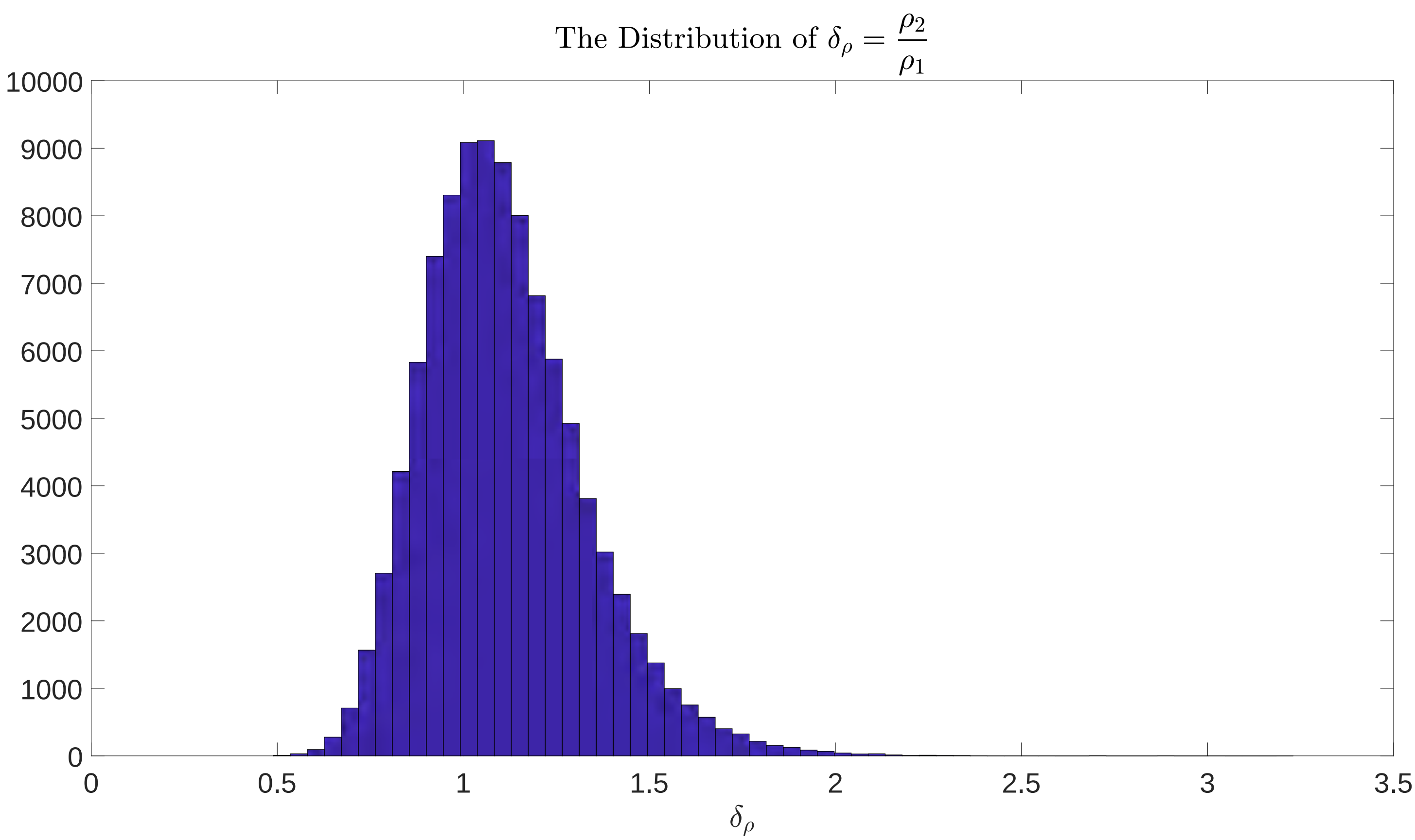


Figure 8.6: The Posterior Distribution of $\delta_\rho = \frac{\rho_2}{\rho_1}$.

The summary statistics for δ_ρ are given in Table 8.6.

Table 8.6: Summary Statistics for δ_ρ.

$Mean\left(\delta_\rho\right)$	$Median\left(\delta_\rho\right)$	$Mode\left(\delta_\rho\right)$	$Var\left(\delta_\rho\right)$	95% Interval	$P\left(\delta_\rho > 1\right)$
1.1117	1.0872	1.071	0.0473	(0.7600; 1.6091)	0.6724

Since one is included in the 95% confidence interval $H_0 : \rho_1 = \rho_2$ will not be rejected at the $\beta = 0.05$ level. As mentioned previously, $\gamma_\rho = \rho_2 - \rho_1$ can also be used for testing H_0. In Figure 8.7 the posterior distribution of γ_ρ is given.

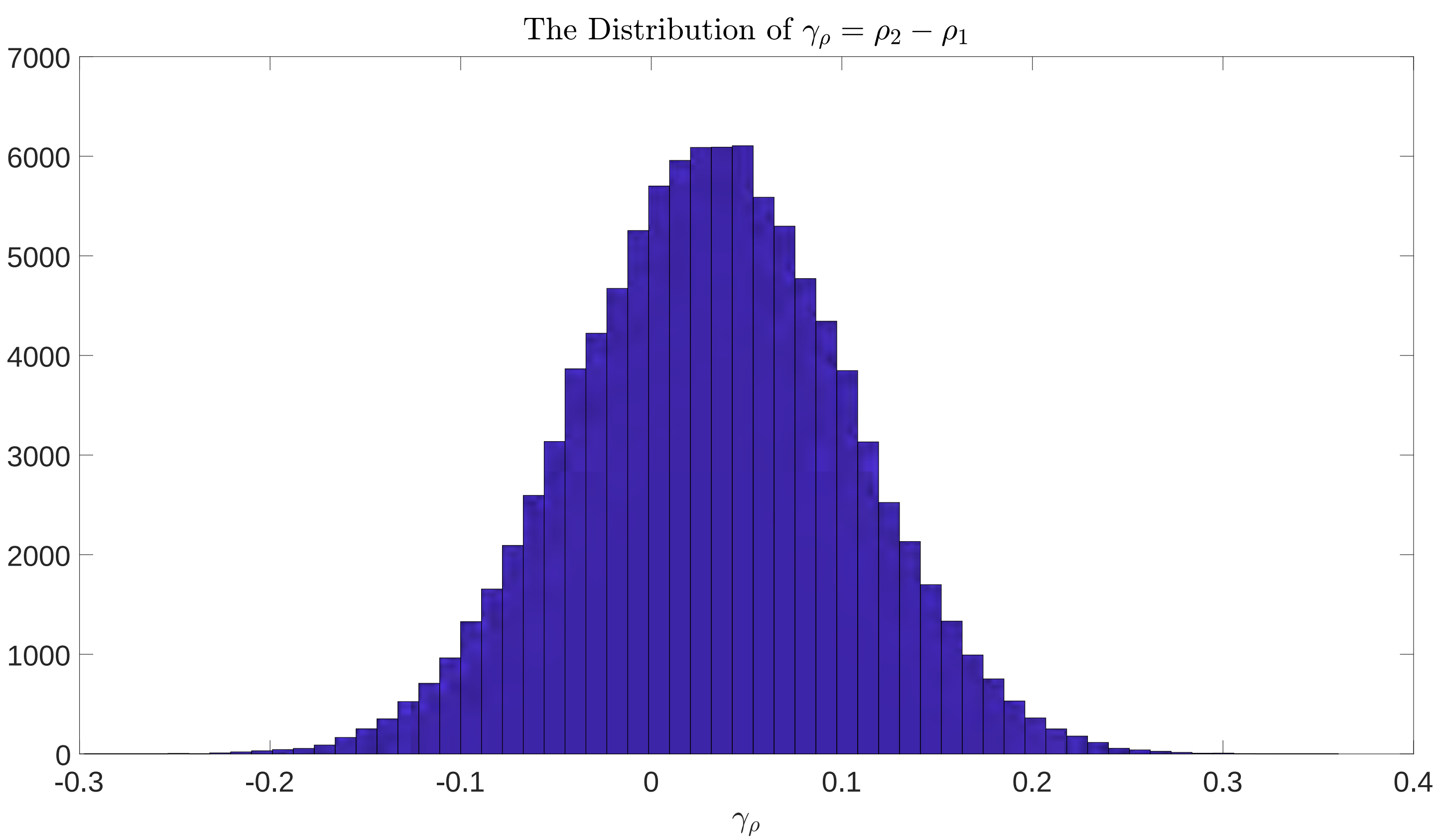


Figure 8.7: The Posterior Distribution of $\gamma_\rho = \rho_2 - \rho_1$.

The summary statistics for γ_ρ are given in Table 8.7.

Table 8.7: Summary Statistics for γ_ρ.

$Mean\left(\gamma_\rho\right)$	$Median\left(\gamma_\rho\right)$	$Mode\left(\gamma_\rho\right)$	$Var\left(\gamma_\rho\right)$	95% Interval	$P\left(\gamma_\rho > 0\right)$
0.0321	0.0322	0.033	0.0051	(−0.1077; 0.1727)	0.6724

Since zero is included in the 95% confidence interval, $H_0 : \rho_1 = \rho_2$ will not be rejected. It is also clear from Figures 8.6 and 8.7 that $P\left(\delta_\rho > 1\right) = P\left(\gamma_\rho > 0\right) = 0.6724$.

8.4 Conclusion

In this chapter a number of confidence intervals were derived for Cronbach's alpha and the intra-class correlation coefficient under compound symmetry. Since the one-way random effects model has a covariance matrix that is compound symmetric, it was used in the analysis of Cronbach's alpha and the intra-class correlation coefficient. In practice it is difficult to compare values of Cronbach's alpha across studies since it depends on the number of items, therefore computing confidence intervals for alpha is essential since it accounts for the differences across the number of items. Confidence intervals for Cronbach's alpha were derived and simulated to test significant differences of alpha across studies in this chapter. Various distributions of statistics were used to compare the significant differences of alpha, the ratio of two Cronbach's alpha values, $\delta_\alpha = \frac{\alpha_2}{\alpha_1}$, the difference between two Cronbach's alpha values, $\gamma_\alpha = \alpha_2 - \alpha_1$ and the distribution of $\Omega = \frac{1-\alpha_1}{1-\alpha_2}$, was also considered. The distributions of the intra-class correlation coefficient was also considered for the ratio of correlation coefficients from two studies, $\delta_\rho = \frac{\rho_2}{\rho_1}$, the difference between two correlation coefficients, $\gamma_\rho = \rho_2 - \rho_1$ and their confidence intervals were derived and simulated. These distributions were derived from the posterior distributions of Cronbach's alpha and the intra-class correlation coefficient.

Chapter 9

Conclusion and, Possible Future Research

This chapter gives a summary of the conclusions of the chapters in the thesis and possible future research.

9.1 Conclusion

In Chapter 3, a number of objective priors for Cronbach's alpha have been derived. The reference prior for the group ordering $\left\{\alpha, \theta, \sigma_1^2\right\}$ was shown to be the same as the probability matching prior when Cronbach's alpha is the parameter of interest. For the group orderings $\{\alpha, \sigma_1^2, \theta\}$, $\{\theta, \sigma_1^2, \alpha\}$, $\{\theta, \alpha, \sigma_1^2\}$, $\{\sigma_1^2, \theta, \alpha\}$ and $\{\sigma_1^2, \alpha, \theta\}$, the result for the reference prior remained the same as the result for the group ordering $\left\{\alpha, \theta, \sigma_1^2\right\}$. The probability matching prior, Jeffreys prior and divergence prior were compared using a simulation study involving their coverage probabilities. The simulation study showed that the probability matching prior outperformed the Jeffreys prior and the divergence prior in terms of the coverage rates obtained. The divergence prior performed the worst in terms of its coverage rates which were either larger than the nominal rate or smaller than the nominal rate of 0.95. The second simulation study showed that the LINEX Loss function mostly performed the best with its c parameter equal to -9. It also showed that for $c = -9$ and if the probability matching prior is considered then the best estimates for Cronbach's alpha were obtained and the Mean Relative Error values were close to one. The Jeffreys prior did quite well in some cases but the divergence prior performed the worst in both of the simulation studies. It is recommended that the probability matching prior be used for the Bayesian analysis of Cronbach's alpha and to also use the LINEX loss function with $c = -9$ to obtain the best estimate of Cronbach's alpha.

In Chapter 4, the combined Bayesian estimate of α for m experiments with equal α but possibly different variance components was derived. Since the model considered is the one-way balanced random effects model, the assumption of equicorrelated normal data is satisfied. Our Bayesian results are therefore different from those of Van Zyl (2001) who had to make the assumption of equicorrelation. Reference and probability matching priors were derived in Section 4.3. They lead to procedures with properties that frequentists can relate to while still retaining Bayesian validity. The fact that the resulting Bayesian posterior intervals of level $1-\beta$ are also good frequentist intervals of the same level is a very desirable situation. It is also shown that the reference and probability matching priors for α are the same. The Bayesian theory and results derived in Sections 4.3 and 4.4 were applied to two examples in Section 4.5. The frequentist coverage of the credibility intervals are for all practical purposes 95%. The intervals for the combined sample are however much shorter than those of the individual samples. Also, the point estimates of the combined sample are more accurate. It is further concluded that the posterior distribution of α for the combined sample becomes more important as the number of samples increase.

In Chapter 5, a number of objective priors for Cronbach's alpha in the case of a balanced three-component hierarchical model, have been derived. The reference prior for the following group orderings: $\left\{\alpha,\sigma_1^2,\sigma_2^2,\theta\right\}$, $\left\{\alpha,\sigma_2^2,\sigma_1^2,\theta\right\}$, $\left\{\alpha,\theta,\sigma_1^2,\sigma_2^2\right\}$, $\left\{\alpha,\theta,\sigma_2^2,\sigma_1^2\right\}$, were considered. It was shown that the reference prior for the group orderings $\left\{\alpha,\sigma_1^2,\sigma_2^2,\theta\right\}$, $\left\{\alpha,\sigma_1^2,\theta,\sigma_2^2\right\}$,$\left\{\alpha,\theta,\sigma_1^2,\sigma_2^2\right\}$, were equivalent to the probability matching prior when Cronbach's alpha is the parameter of interest. The probability matching prior that we have used when $c=0$ when α is the parameter of interest is different from the reference prior that was obtained in this chapter. We compared the probability matching prior, Jeffreys prior and divergence prior using a simulation study involving their coverage probabilities. The simulation study showed that the reference prior and the probability matching prior outperformed the Jeffreys prior and the divergence prior in terms of the coverage rates obtained. The coverage rates obtained for the reference prior were the same as those for the probability matching prior. The mathematics showed that they both satisfy the condition to be a probability matching prior therefore it is expected that the coverage rates obtained using these priors should be similar. The divergence prior performed the worst in terms of its coverage rates which were either larger than the nominal rate or smaller than the nominal rate of 0.95. It is recommended that the reference prior and the probability matching prior should be used for the Bayesian analysis of Cronbach's alpha for the three-component hierarchical model.

In Chapter 6, statistical process control limits were derived for Cronbach's alpha coefficient in the case of the balanced one-way random effects model. This was achieved by deriving the predictive distribution of a future (unseen) Cronbach's alpha coefficient. For given variance components, it was shown that the predictive density function of $f\left(\hat{\alpha}_f|\alpha\right)$ can be derived analytically. The Jeffreys independence prior was used to derive the posterior distribution of α. The unconditional posterior predictive density function $f\left(\hat{\alpha}_f|\hat{\alpha}\right)$ can be obtained by Monte Carlo simulation or numerical integration. The predictive density function $f\left(\hat{\alpha}_f|\hat{\alpha}\right)$ as well as the conditional predictive density functions $f\left(\hat{\alpha}_f|\alpha\right)$ can be used to determine the run-length and the average run-length. The distribution of the run-length $f\left(r|\hat{\alpha}\right)$ is the average of a large number of geometric distributions each with its own parameter value. Three examples were considered. The first example had to do with Dyestuff data and it is from Box & Tiao (1973). In the second example it was assumed that the number of Dyestuff samples have increased from six to hundred and twenty and the number of samples in a future (unseen) data set is ninety. The third example was from Wooluru et al. (2014) and were measured values of "Box diameter" on the driver gear. From the results it was observed that the average and median run-lengths are usually larger than the theoretical values. An advantage of the Bayesian procedure however is that control limits can be adjusted in such a way that the average or median run-length has a specific value.

In Chapter 7, a number of objective priors were considered in the Bayesian analysis of Cronbach's alpha for a bivariate normal distribution under a general covariance matrix. The asymptotic interval of Cronbach's alpha was also considered for a general covariance matrix. A simulation study was conducted for various sample sizes where the coverage probabilities and average interval lengths were computed for the posterior distributions considering all of the objective priors as well as the asymptotic interval. For the covariance matrices considered in the simulation study, it was evident that the objective priors outperformed the asymptotic interval where their coverage probabilities were all closer to the nominal rate of 0.95. The asymptotic interval underestimated the coverage probabilities for all the sample sizes considered in the simulation study. The performance of the asymptotic interval was especially bad for small sample sizes. This is not surprising since the interval was derived by Van Zyl et al. (2000) for large sample sizes using asymptotic results. The average interval lengths for the asymptotic interval were the shortest for all covariance matrices considered but with a shortcoming of not getting coverage probabilities closer to the nominal rate. The Jeffreys rule prior also obtained the shortest average interval lengths between all the priors considered but also fell short with the coverage probabilities being below the nominal level. In practice, for small sample sizes, it is not recommended to use the asymptotic interval for alpha. For small sample sizes, it is recommended to use either the one-at-a-time reference prior, $\pi_{R\rho}$, or the reference prior with group ordering $\{\sigma_1, \sigma_2, \rho\}$, $\pi_{R\sigma}$, or the prior π_{MS}. For larger sample sizes, all the objective priors considered performed similarly. It is recommended to also use $\pi_{R\rho}$, $\pi_{R\sigma}$ or π_{MS} for larger sample sizes although any of the priors considered may also be considered. Even though the asymptotic interval improved for larger sample sizes, the Bayesian approach still consistently outperformed the asymptotic frequentist interval.

In Chapter 8, a number of confidence intervals were derived for Cronbach's alpha and the intra-class correlation coefficient under compound symmetry. Since the one-way random effects model has a covariance matrix that is compound symmetric, it was the model considered in the analysis of Cronbach's alpha and the intra-class correlation coefficient. In practice it is difficult to compare values of Cronbach's alpha across studies since it depends on the number of items, therefore computing confidence intervals for alpha is essential since it accounts for the differences across the number of items. Confidence intervals for Cronbach's alpha were derived and simulated to test significant differences of alpha across studies in this chapter. Various distributions of statistics were used to compare the significant differences of alpha, the ratio of two Cronbach's alpha values, $\delta_\alpha = \frac{\alpha_2}{\alpha_1}$, the difference between two Cronbach's alpha values, $\gamma_\alpha = \alpha_2 - \alpha_1$ and the distribution of $\Omega = \frac{1-\alpha_1}{1-\alpha_2}$, was also considered. The distributions of the intra-class correlation coefficient was also considered for the ratio of correlation coefficients from two studies, $\delta_\rho = \frac{\rho_2}{\rho_1}$, the difference between two correlation coefficients, $\gamma_\rho = \rho_2 - \rho_1$ and their confidence intervals were derived and simulated. These distributions were derived using a pivotal quantity derived by Van Zyl et al. (2000) as a fiducial or objective Bayesian distribution for Cronbach's alpha and the intra-class correlation coefficient.

9.2 Possible Future Research

- For parameters defined on (0, 1), a common approach is to put a normal prior on the logit, for example, $log(\alpha) - log(1-\alpha)$. Taking such a prior with, say, 0 mean and large variance, would be another possible non-informative prior. The idea would be to extend the work done in this thesis using the vague prior.

- Possible further studies would be to investigate an unbalanced one-way random effects model and estimate Cronbach's alpha using various objective priors. Chaloner (1987) estimated variance components in the one-way random model with unequal sample sizes. Chaloner (1987) considered various objective priors in their study to estimate the variance components. They also did a simulation study for the ratio of the variance components. The idea is to extend the work done by Chaloner (1987) by deriving the reference and probability matching prior for Cronbach's alpha using the unbalanced one-way random effects model. Chaloner (1987) used the mode as the Bayes estimator, we will investigate the mode as well as other loss functions such as the absolute error, squared error and the LINEX loss function to estimate Cronbach's alpha.

- It is also possible to extend the work using Bayes factors. Kang & Lee (2004) developed a Bayesian test procedure for the intra-class correlation coefficient using the balanced and unbalanced one-way random effects model using a reference prior. They used the fractional Bayes factor by O'Hagan (1995) since the prior used was improper which can lead to the Bayes factor containing unspecified constants. They showed that the calculation of the fractional Bayes factor required only one-dimensional integration. Since Cronbach's alpha is related to the intra-class correlation coefficient using the Spearman-Brown equation, we can extend the work done by Kang & Lee (2004) by developing a Bayesian test procedure for Cronbach's alpha using objective priors. To gain further insights beyond estimation, the comparison of different priors in conjunction with fixed likelihoods for one-way or two-way ANOVA models using Bayes Factor could be considered. This would add a complementary perspective to the existing analysis.

- Development of general loss functions by exploring more general loss functions, such as the Kullback-Leibler divergence and invariant loss functions. These could provide additional insights into the robustness and applicability of the proposed Bayesian methods.

- Hierarchical and Empirical Bayesian Approaches: The inclusion of results on hierarchical and empirical Bayesian methods for estimating Cronbach's alpha.

References

Alt, F. B. (2006). A Bayesian approach to monitoring the mean of a multivariate normal process. In B. M. Colosimo & E. del Castillo (Eds.), *Bayesian process monitoring, control and optimization* (pp. 151–178). Chapman and Hall/CRC.

Anderson, T. W. (1958). *An Introduction to Multivariate Analysis*. John Wiley and Sons, New York, 1st edition.

Balakrishnan, N., Kotz, S., Read, C. B., Vidakovic, B., & Johnson, N. L. (2006). *Encyclopedia of Statistical Sciences, Volume 1*. John Wiley & Sons, 2nd edition.

Bayarri, M. J. (1981). Bayesian inference on the correlation coefficient of a bivariate normal population. *Statistical and Operational Research Papers*, 32(1), 18–31.

Bayarri, M. J. & García-Donato, G. (2005). A Bayesian sequential look at *u*-control charts. *Technometrics*, 47(2), 142 – 151.

Berger, J. O. & Bernado, J. M. (1992). Ordered group reference priors with application to the multinomial problem. *Biometrika*, 79(1), 25–37.

Berger, J. O. & Bernardo, J. M. (1992). On the development of reference priors. In *Bayesian Statistics 4*, Bernardo, J. M., Berger, J. O., Dawid, A. P. & Smith, A. F. M. (eds). Oxford University Press, Oxford, 35 - 60.

Berger, J. O., Bernardo, J. M., & Sun, D. (2009). The formal definition of reference priors. *The Annals of Statistics*, 37(2), 905 – 938.

Berger, J. O. & Sun, D. (2006). Objective priors for a bivariate normal model with multivariate generalizations. *Duke University ISDS Technical Report 2007*, 6.

Berger, J. O. & Sun, D. (2008). Objective priors for the bivariate normal model. *The Annals of Statistics*, 36(2), 963–982.

Bernardo, J. M. (1979). Reference posterior distributions for Bayesian inference. *Journal of the Royal Statistical Society, Series B*, 41(2), 113 – 147.

Box, G. & Tiao, G. (1973). *Bayesian Inference in Statistical Analysis*. Addison-Wesley Publishing Company, 1st edition.

Brown, P. J., Le, N. D., & Zidek, J. V. (1994). Inference for a covariance matrix. *Aspects of uncertainty: a tribute to DV Lindley*.

Brown, W. (1910). Some experimental results in the correlation of mental abilities. *British Journal of Psychology*, 3(3), 296–322.

Chaloner, K. (1987). A Bayesian approach to the estimation of variance components for the unbalanced one-way random model. *Technometrics*, 29(3), 323–337.

Chung, Y. & Dey, D. K. (1998). Bayesian approach to estimation of intraclass correlation using reference prior. *Communications in Statistics - Theory and Methods*, 27(9), 2241 – 2255.

Cortina, J. M. (1993). What is coefficient alpha? An examination of theory and applications. *Journal of Applied Psychology*, 78(1), 98 – 104.

Cronbach, L. J. (1951). Coefficient alpha and the internal structure of tests. *Psychometrika*, 16(3), 297 – 334.

Cui, Y. & Li, J. (2012). Evaluating the performance of different procedures for constructing confidence intervals for coefficient alpha: A simulation study. *British Journal of Mathematical and Statistical Psychology*, 65(1), 467–498.

Daniels, M. J. (1999). A prior for the variance in hierarchical models. *The Canadian Journal of Statistics*, 27(3), 567–578.

Datta, G. S. & Ghosh, J. K. (1995). On priors providing frequentist validity for Bayesian inference. *Biometrika*, 82(1), 37 – 45.

Davies, O. & Goldsmith, P. (1972). *Statistical Methods in Research and Production*. Longman Publishing Group, 4th edition.

Dennemeyer, R. (1968). *Introduction to Partial Differential Equations and Boundary Value Problems*. McGraw-Hill Book Company, 1st edition.

Duhachek, A. & Iacobucci, D. (2004). Alpha's Standard Error (ASA): An accurate and precice confidence interval estimate. *Journal of Applied Psychology*, 89(5), 792–808.

Feldt, L. (1965). The approximate sampling distribution of the Kuder-Richardson reliability coefficient twenty. *Psychometrika*, 30(3), 357–370.

Feldt, L., Woodruff, D., & Salih, F. (1987). Statistical inference for coefficient alpha. *Applied Psychological Measurement*, 11(1), 93–103.

Ghosh, M., Mergel, V., & Lui, R. (2011). A general divergence criterion for prior selection. *Annals of the Institute of Statistical Mathematics*, 63(1), 43–58.

Goligher, E. C. & Harhay, M. O. (2024). What is the point of Bayesian analysis? *American Journal of Respiratory and Critical Care Medicine*, 209(5), 485–487.

Graves, S. (2006). Bayes rule of information and monitoring in manufacturing integrated circuits. In B. M. Colosimo & E. del Castillo (Eds.), *Bayesian process monitoring, control and optimization* (pp. 199–224). Chapman and Hall/CRC.

Guttman, L. (1945). A basis for analyzing test-retest reliability. *Psychometrika*, 10(4), 255 – 282.

Hakstian, A. R. & Whalen, T. E. (1976). A *k*-sample significance test for independent alpha coefficients. *Psychometrika*, 41(2), 219–231.

Hulin, C. L., Drasgow, F., & Parsons, C. K. (1983). *Item Response Theory: Application to Psychological Measurement*. Dow Jones-Irwin, 1st edition.

Human, S. W. (2009). *Univariate parametric and nonparametric statistical quality control techniques with estimated process parameters*. phdthesis, University of Pretoria.

Izally, S. R., Van der Merwe, A. J., & Raubenheimer, L. (2024). A comparison of objective priors for Cronbach's coefficient alpha using a balanced random effects model. *Communications in Statistics - Theory and Methods*, doi:10.1080/03610926.2024.2315300.

Jeffreys, H. (1939). *Theory of Probability*. Oxford University Press, Oxford, 1st edition.

Kang, S.-G. & Lee, H.-C. (2004). Bayesian test for the intraclass correlation coefficient in the one-way random effect model. *Journal of the Korean Data and Information Science Society*, 15(3), 645–654.

Kaplan, R. M. & Saccuzzo, D. R. (1993). *Psychological Testing: Principles, Applications, and Issues*. Thomson Brooks/Cole Publishing Co., 3rd edition.

Koning, A. & Franses, P. (2003). *Confidence intervals for comparing Cronbach's alpha values*. Technical report, Erasmus Research Institute of Management.

Kristof, W. (1963). The statistical theory of stepped-up reliability coefficients when a test has been divided into several equivalent parts. *Psychometrika*, 28(3), 221–238.

Kuder, G. F. & Richardson, M. W. (1937). The theory of the estimation of test reliability. *Psychometrika*, 2(3), 151–160.

Li, J. C. & Woodruff, D. J. (2002). *Bayesian statistical inference for coefficient alpha*. Technical report, ACT Research Report Series.

Lindley, D. V. (1965). *Introduction to Probability and Statistics from a Bayesian Viewpoint*. Cambridge University Press, 1st edition.

Maydeu-Olivares, A., Coffman, D. L., & Hartmann, W. M. (2007). Asymptotically distribution-free (ADF) interval estimation of coefficient alpha. *Psychological Methods*, 12(2), 157–176.

Menzefricke, U. (2002). On the evaluation of control chart limits based on predictive distributions. *Communications in Statistics - Theory and Methods*, 31(8), 1423 – 1440.

Menzefricke, U. (2007). Control charts for the generalized variance based on its predictive distribution. *Communications in Statistics - Theory and Methods*, 36(5), 1031 – 1038.

Menzefricke, U. (2010a). Control charts for the variance and the coefficient of variation based on their predictive distribution. *Communications in Statistics - Theory and Methods*, 39(16), 2930 – 2941.

Menzefricke, U. (2010b). Multivariate exponentially weighted moving average charts for a mean based on its predictive distribution. *Communications in Statistics - Theory and Methods*, 39(16), 2942 – 2960.

Mohammed-Djafari, A. (2015). Entropy, information theory, information geometry and Bayesian inference in data, signal and image processing and inverse problems. *Entropy*, 17(6), 3989–4027.

Montgomery, D. C. (1996). *Introduction to Statistical Quality Control*. John Wiley & Sons, New York, 3rd edition.

Montgomery, D. C. (2005). *Introduction to Statistical Quality Control*. John Wiley & Sons, 5th edition.

Moreno, C. W. (2006). An application of Bayesian statistics to sequential empirical optimization. In B. M. Colosimo & E. del Castillo (Eds.), *Bayesian process monitoring, control and optimization* (pp. 303–322). Chapman and Hall/CRC.

O'Hagan, A. (1995). Fractional Bayes factors for model comparison. *Journal of the Royal Statistical Society: Series B (Methodological)*, 57(1), 99–118.

Padilla, M. & Zhang, G. (2011). Estimating internal consistency using Bayesian methods. *Journal of Modern Applied Statistical Methods*, 10(1), 277–286.

Palmer, J. L. & Broemeling, L. D. (1990). A comparison of Bayes and maximum likelihood estimation of the intraclass correlation coeffcient. *Communications in Statistics- Theory and Methods*, 19(3), 953–975.

Payandeh Najafabadi, A. T. & Najafabadi, M. O. (2016). On the Bayesian estimation for Cronbach's alpha. *Journal of Applied Statistics*, 43(13), 2416–2441.

Peterson, J. J. (2006). A review of Bayesian reliability approaches to multiple response surface optimization. In B. M. Colosimo & E. del Castillo (Eds.), *Bayesian process monitoring, control and optimization* (pp. 281–302). Chapman and Hall/CRC.

Raykov, T. (1998). A method for obtaining standard errors and confidence intervals of composite reliability for congeneric items. *Applied Psychological Measurement*, 22(4), 369–374.

Rényi, A. (1961). On measures of entropy and information. In *Proceedings of the fourth Berkeley symposium on mathematical statistics and probability, volume 1: contributions to the theory of statistics*, volume 4 (pp. 547–562).: University of California Press.

Robert, C. P. (2001). *The Bayesian Choice*. Springer, New York, 2nd edition.

Sahai, H. & Ojeda, M. M. (2004). *Analysis of Variance for Random Models*. Springer, 1st edition.

Shiau, J.-J. H. & Feltz, C. J. (2006). Empirical Bayes process monitoring techniques. In B. M. Colosimo & E. del Castillo (Eds.), *Bayesian process monitoring, control and optimization* (pp. 121–150). Chapman and Hall/CRC.

Spearman, C. (1910). Correlation calculated from faulty data. *British Journal of Psychology*, 3(3), 271–295.

Tagaras, G. & Nenes, G. (2006). Two-sided Bayesian X control charts for short production runs. In B. M. Colosimo & E. del Castillo (Eds.), *Bayesian process monitoring, control and optimization* (pp. 179–198). Chapman and Hall/CRC.

Tsagris, M., Frangos, C. C., & Frangos, C. C. (2013). Confidence intervals for Cronbach's reliability coefficient. In *Proceedings of the 1st International Conference on Computer Supported Education* (pp. 14–16).

Tsiamyrtzis, P. & Hawkins, D. M. (2006). A Bayesian approach to statistical process control. In B. M. Colosimo & E. del Castillo (Eds.), *Bayesian process monitoring, control and optimization* (pp. 87 – 108). Chapman & Hall/CRC.

Van der Merwe, A. & Hugo, J. (2007). Bayesian tolerance intervals for the balanced two-factor nested random effects model. *Test*, 16(1), 598–612.

Van Zyl, J., Neudecker, H., & Nel, D. (2000). On the distribution of the maximum likelihood estimator of Cronbach's alpha. *Psychometrika*, 65(3), 271–280.

Van Zyl, J. M. (2001). Combined Bayesian estimates of coefficient alpha. *South African Statistical Journal*, 35(2), 225 – 233.

Van Zyl, R. & Van der Merwe, A. J. (2019). Bayesian process monitoring schemes for the two-parameter exponential distribution. *Communications in Statistics - Theory and Methods*, 48(7), 1766 – 1797.

Varian, H. R. (1975). *Studies in Bayesian econometrics and statistics: In honour of Leonard J. Savage*, volume 86. North-Holland Publishing Company.

Viana, M. A. (1995). Combined Bayesian estimates for the equicorrelation coefficient. *Canadian Journal of Statistics*, 23(1), 43–53.

Wooluru, Y., Swamy, D., & Nagesh, P. (2014). The process capability analysis -A tool for process performance measures and metrics-A case study. *International Journal for Quality Research*, 8(3), 399–416.

Yuan, K.-H., Guarnaccia, C. A., & Hayslip Jr, B. (2003). A study of the distribution of sample coefficient alpha with the Hopkins symptom checklist: Bootstrap versus asymptotics. *Educational and Psychological Measurement*, 63(1), 5–23.

Appendix A: Cronbach's Alpha for the One-Way Random Effects Model

A.1 Fisher Information Matrix for Cronbach's Alpha using the One-Way Random effects Model

The likelihood function is given by

$$\ell(\theta,\sigma_1^2,\alpha|data) \propto (\sigma_1^2)^{-\frac{1}{2}(v_1+v_2+1)}(1-\alpha)^{\left(\frac{v_2+1}{2}\right)}\exp\left\{-\frac{1}{2\sigma_1^2}\left[IJ\left(\bar{Y}_{..}-\theta\right)^2(1-\alpha)+v_2m_2\left(1-\alpha\right)+v_1m_1\right]\right\}.$$

To obtain the Fisher information matrix we take the logarithm of the likelihood function and differentiate twice with respect to the unknown parameters.

$$\begin{aligned}
L &= log\left[\ell(\theta,\sigma_1^2,\alpha|data)\right] \\
&\propto -\frac{1}{2}(v_1+v_2+1)log(\sigma_1^2)+\frac{(v_2+1)}{2}log(1-\alpha)-\frac{1}{2\sigma_1^2}\left[IJ\left(\bar{Y}_{..}-\theta\right)^2(1-\alpha)+v_2m_2\left(1-\alpha\right)+v_1m_1\right] \\
\frac{\partial L}{\partial\sigma_1^2} &= -\frac{(v_1+v_2+1)}{2(\sigma_1^2)}+\frac{1}{2(\sigma_1^2)^2}\left[IJ\left(\bar{Y}_{..}-\theta\right)^2(1-\alpha)+v_2m_2\left(1-\alpha\right)+v_1m_1\right] \\
\frac{\partial L}{\partial\alpha} &= \frac{-(v_2+1)}{2(1-\alpha)}+\frac{1}{2\sigma_1^2}\left[IJ\left(\bar{Y}_{..}-\theta\right)^2+v_2m_2\right] \\
\frac{\partial L}{\partial\theta} &= \frac{1}{\sigma_1^2}\left[IJ\left(\bar{Y}_{..}-\theta\right)(1-\alpha)\right] \\
\frac{\partial^2 L}{\left(\partial\sigma_1^2\right)^2} &= \frac{(v_1+v_2+1)}{2(\sigma_1^2)^2}-\frac{1}{(\sigma_1^2)^3}\left[IJ\left(\bar{Y}_{..}-\theta\right)^2(1-\alpha)+v_2m_2\left(1-\alpha\right)+v_1m_1\right] \\
\frac{\partial^2 L}{(\partial\alpha)^2} &= \frac{-(v_2+1)}{2(1-\alpha)^2} \\
\frac{\partial^2 L}{(\partial\theta)^2} &= -\frac{IJ\left(1-\alpha\right)}{\sigma_1^2};\frac{\partial^2 L}{\partial\theta\partial\sigma_1^2}=-\frac{(1-\alpha)\left(v_1+v_2+1\right)\left(\bar{Y}_{..}-\theta\right)}{(\sigma_1^2)^2}=\frac{\partial^2 L}{\partial\sigma_1^2\partial\theta} \\
\frac{\partial^2 L}{\partial\sigma_1^2\partial\alpha} &= -\frac{1}{2(\sigma_1^2)^2}\left[IJ\left(\bar{Y}_{..}-\theta\right)^2+v_2m_2\right]=\frac{\partial^2 L}{\partial\alpha\partial\sigma_1^2}
\end{aligned}$$

$$\frac{\partial^2 L}{\partial\alpha\partial\theta} = \frac{-(v_1+v_2+1)(\bar{Y}_{..}-\theta)}{\sigma_1^2} = \frac{\partial^2 L}{\partial\theta\partial\alpha}$$

Therefore,

$$\begin{aligned}
&-E\left[\frac{\partial^2 L}{\left(\partial\sigma_1^2\right)^2}\right]\\
=&-E\left[\frac{(v_1+v_2+1)}{2(\sigma_1^2)^2}-\frac{1}{(\sigma_1^2)^3}\left[IJ\left(\bar{Y}_{..}-\theta\right)^2(1-\alpha)+v_2m_2(1-\alpha)+v_1m_1\right]\right]\\
=&\frac{-(v_1+v_2+1)}{2(\sigma_1^2)^2}+\frac{1}{(\sigma_1^2)^3}\left[E\left(IJ\left(\bar{Y}_{..}-\theta\right)^2\right)(1-\alpha)+v_2E(m_2)(1-\alpha)+v_1E(m_1)\right]\\
=&-\frac{(v_1+v_2+1)}{2(\sigma_1^2)^2}+\frac{1}{(\sigma_1^2)^3}\left[\frac{\sigma_1^2}{(1-\alpha)}(1-\alpha)+v_2\left(\frac{\sigma_1^2}{(1-\alpha)}\right)(1-\alpha)+v_1\sigma_1^2\right]\\
=&-\frac{(v_1+v_2+1)}{2(\sigma_1^2)^2}+\frac{(v_1+v_2+1)}{(\sigma_1^2)^2}=\frac{(v_1+v_2+1)}{(\sigma_1^2)^2}\left[1-\frac{1}{2}\right]=\frac{(v_1+v_2+1)}{2(\sigma_1^2)^2}.
\end{aligned}$$

Also, $-E\left[\frac{\partial^2 L}{(\partial\alpha)^2}\right]=-E\left[\frac{-(v_2+1)}{2(1-\alpha)^2}\right]=\frac{(v_2+1)}{2(1-\alpha)^2}$ and $-E\left[\frac{\partial^2 L}{(\partial\theta)^2}\right]=-E\left[-\frac{IJ(1-\alpha)}{\sigma_1^2}\right]=\frac{IJ(1-\alpha)}{\sigma_1^2}$. The following expected values are all zero: $-E\left[\frac{\partial^2 L}{\partial\theta\partial\sigma_1^2}\right]=-E\left[\frac{\partial^2 L}{\partial\sigma_1^2\partial\theta}\right]=0$ and $-E\left[\frac{\partial^2 L}{\partial\alpha\partial\theta}\right]=-E\left[\frac{\partial^2 L}{\partial\theta\partial\alpha}\right]=0$. Finally

$$\begin{aligned}
&-E\left[\frac{\partial^2 L}{\partial\sigma_1^2\partial\alpha}\right]\\
=&-E\left[-\frac{1}{2(\sigma_1^2)^2}\left[IJ\left(\bar{Y}_{..}-\theta\right)^2+v_2m_2\right]\right]\\
=&\frac{1}{2(\sigma_1^2)^2}\left[E\left(IJ\left(\bar{Y}_{..}-\theta\right)^2\right)+v_2E\left(m_2\right)\right]\\
=&\frac{1}{2(\sigma_1^2)^2}\left[\frac{\sigma_1^2}{(1-\alpha)}+\frac{v_2\sigma_1^2}{(1-\alpha)}\right]\\
=&\frac{(1+v_2)}{2\sigma_1^2(1-\alpha)}.
\end{aligned}$$

The Fisher information matrix is therefore given by

$$F(\theta,\sigma_1^2,\alpha)=\begin{bmatrix}\frac{IJ(1-\alpha)}{\sigma_1^2} & 0 & 0\\ 0 & \frac{(v_1+v_2+1)}{2(\sigma_1^2)^2} & \frac{(v_2+1)}{2(1-\alpha)\sigma_1^2}\\ 0 & \frac{(v_2+1)}{2(1-\alpha)\sigma_1^2} & \frac{(v_2+1)}{2(1-\alpha)^2}\end{bmatrix}.$$

A.2 Proof of the Reference Prior for the Group Ordering $\left\{\alpha, \sigma_1^2, \theta\right\}$

We are interested in the reference prior for the group ordering $\left\{\alpha, \sigma_1^2, \theta\right\}$ which means that α is the most important parameter and θ is the least important parameter. In order to derive the reference prior, the inverse of the Fisher information matrix is needed. Let $S(\boldsymbol{\theta}) = H^{-1}(\boldsymbol{\theta})$ where H is the Fisher information matrix. Now

$$S\left(\alpha, \sigma_1^2, \theta\right) = \begin{bmatrix} A_{11} & A'_{21} & \cdots & A'_{m1} \\ A_{21} & A_{22} & \cdots & A'_{m2} \\ \vdots & & \ddots & \\ A_{m_1} & A_{m2} & \cdots & A_{mm} \end{bmatrix} = \begin{bmatrix} \frac{2J(1-\alpha)^2}{I(J-1)} & \frac{-2(1-\alpha)\left(\sigma_1^2\right)}{I(J-1)} & 0 \\ \frac{-2(1-\alpha)\left(\sigma_1^2\right)}{I(J-1)} & \frac{2(\sigma_1^2)^2}{I(J-1)} & 0 \\ 0 & 0 & \frac{\sigma_1^2}{IJ(1-\alpha)} \end{bmatrix}$$

Define the truncated ranges for the 3 parameters as $\alpha \in [a_l, b_l]$, $\theta \in [c_l, d_l]$ and $\sigma_1^2 \in [e_l, f_l]$ where $c_l \to -\infty$, $b_l \to 1$, $d_l, f_l \to \infty$ and $a_l, e_l \to 0$. Now $h_1 \equiv H_1 \equiv A_{11}^{-1} = \frac{I(J-1)}{2J(1-\alpha)^2}$, $h_2 = \left(A_{22} - B_2 H_1 B'_2\right)^{-1}$. S_1 is the upper left $N_1 \times N_1$ corner of S. We have $m = 3$ groups with $n_1 = 1$, $n_2 = 1$ and $n_3 = 3$ $\therefore S_1$ is the upper left 1×1 corner of S, that is $S_1 = \frac{2J(1-\alpha)^2}{I(J-1)}$. Therefore $H_1 = S_1^{-1} = \frac{I(J-1)}{2J(1-\alpha)^2}$ and $B_2 = A_{21} = \frac{-2(1-\alpha)\left(\sigma_1^2\right)}{I(J-1)}$. Hence,

$$\begin{aligned} h_2 &= \left[A_{22} - B_2 H_1 B_2^{'}\right]^{-1} \\ &= \left[\frac{2(\sigma_1^2)^2}{I(J-1)} - \frac{2(\sigma_1^2)^2}{IJ(J-1)}\right]^{-1} \\ &= \left[\frac{2(\sigma_1^2)^2}{I(J-1)}\left(1 - \frac{1}{j}\right)\right]^{-1} \\ &= \frac{IJ}{2(\sigma_1^2)^2}. \end{aligned}$$

Also $h_3 = \left(A_{33} - B_3 H_2 B'_3\right)^{-1}$ where S_2 is the upper left $N_2 \times N_2$ corner of S. Now $N_2 = n_1 + n_2 = 2$. Therefore S_2 is the upper left 2×2 corner of S, that is $S_2 = \begin{bmatrix} \frac{2J(1-\alpha)^2}{I(J-1)} & \frac{-2(1-\alpha)\left(\sigma_1^2\right)}{I(J-1)} \\ \frac{-2(1-\alpha)\left(\sigma_1^2\right)}{I(J-1)} & \frac{\sigma_1^2}{IJ(1-\alpha)} \end{bmatrix}$. Now, $H_2 = S_2^{-1} = \begin{bmatrix} \frac{I}{2(1-\alpha)^2} & \frac{I}{2(1-\alpha)\sigma_1^2} \\ \frac{I}{2(1-\alpha)\sigma_1^2} & \frac{IJ}{2.(\sigma_1^2)^2} \end{bmatrix}$ and $B_3 = \begin{bmatrix} A_{31} & A_{32} \end{bmatrix} = \begin{bmatrix} 0 & 0 \end{bmatrix}$. Therefore, $B_3 H_2 B'_3 = 0$. Since the values of A_{31} and A_{32} are both zero, it was not really necessary to compute H_2 but we include it for completeness. Therefore, $h_3 = A_{33}^{-1} = \frac{IJ(1-\alpha)}{\sigma_1^2}$. The three functions

$$h_1 = \frac{I(J-1)}{2J(1-\alpha)^2}, \qquad h_2 = \frac{IJ}{2\left(\sigma_1^2\right)^2}, \qquad h_3 = \frac{IJ(1-\alpha)}{\sigma_1^2}$$

are the functions needed to calculate the prior. During the iterations, first the truncated conditional function of σ_1^2 given α and θ can be computed as

$$\begin{aligned}\pi_3^l\left(\theta|\alpha,\sigma_1^2\right) =& \frac{\{h_3(\boldsymbol{\theta})\}^{\frac{1}{2}}}{\int_{\theta_3^l}\{h_3(\boldsymbol{\theta})\}^{\frac{1}{2}}d\theta_3}\\ \propto& \frac{(1-\alpha)^{\frac{1}{2}}(\sigma_1^2)^{-\frac{1}{2}}}{\int_{c_l}^{d_l}(1-\alpha)^{\frac{1}{2}}(\sigma_1^2)^{-\frac{1}{2}}d\theta}\\ =& \frac{1}{d_l-c_l} \qquad \text{for } c_l \leq \theta \leq d_l.\end{aligned}$$

Now

$$\begin{aligned}&E\left[log\{h_2(\boldsymbol{\theta})\}^{\frac{1}{2}}\right]\\ \propto& \int_{c_l}^{d_l} log(\sigma_1^2)^{-1}\frac{1}{d_l-c_l}d\theta\\ =& log(\sigma_1^2)^{-1}.\end{aligned}$$

Therefore,

$$\begin{aligned}\pi_2^l\left(\sigma_1^2,\theta|\alpha\right) =& \frac{\pi_3^l\left(\theta|\alpha,\sigma_1^2\right)exp\left\{E\left[log\{h_2(\boldsymbol{\theta})\}^{\frac{1}{2}}\right]\right\}}{\int_{e_l}^{f_l}exp\left\{E\left[log\{h_2(\boldsymbol{\theta})\}^{\frac{1}{2}}\right]\right\}d\sigma_1^2}\\ =& \frac{\pi_3^l\left(\sigma_1^2|\alpha,\theta\right)exp\left\{log(\sigma_1^2)^{-1}\right\}}{\int_{e_l}^{f_l}exp\left\{log(\sigma_1^2)^{-1}\right\}d\sigma_1^2}\\ =& \frac{(d_l-c_l)^{-1}(\sigma_1^2)^{-1}}{log\left(f_l e_l^{-1}\right)} \qquad \text{for } e_l \leq \sigma_1^2 \leq f_l \text{ and } c_l \leq \theta \leq d_l.\end{aligned}$$

We now need the function, h_1, to determine $E\left[log\{h_1(\boldsymbol{\theta})\}^{\frac{1}{2}}\right]$

$$\begin{aligned}&E\left[log\{h_1(\boldsymbol{\theta})\}^{\frac{1}{2}}\right]\\ \propto& \int_{e_l}^{f_l}\int_{c_l}^{d_l} log(1-\alpha)^{-1}\frac{(d_l-c_l)^{-1}(\sigma_1^2)^{-1}}{log\left(f_l e_l^{-1}\right)}d\theta d\sigma_1^2\\ =& \int_{e_l}^{f_l}\frac{log(1-\alpha)^{-1}(\sigma_1^2)^{-1}}{log\left(f_l e_l^{-1}\right)}d\sigma_1^2\\ =& log(1-\alpha)^{-1}.\end{aligned}$$

Also $\int_{a_l}^{b_l} exp\left\{log\left(1-\alpha\right)^{-1}\right\} d\alpha = log\left\{\frac{(1-a_l)}{(1-b_l)}\right\}$.Therefore,

$$\begin{aligned}\pi_1^l\left(\alpha,\sigma_1^2,\theta\right) =& \frac{\pi_2^l\left(\sigma_1^2,\theta|\alpha\right)\left(1-\alpha\right)^{-1}}{\int_{a_l}^{b_l}\left(1-\alpha\right)^{-1}d\alpha}\\ =& \frac{(d_l-c_l)^{-1}(\sigma_1^2)^{-1}\left(1-\alpha\right)^{-1}}{log\left(f_l e_l^{-1}\right)log\left\{\frac{(1-a_l)}{(1-b_l)}\right\}}.\end{aligned}$$

Finally

$$\pi_R\left(\alpha,\sigma_1^2,\theta\right) \propto \lim_{l\to\infty}\frac{\pi_1^l\left(\alpha,\theta,\sigma_1^2\right)}{\pi_1^l\left(\alpha_0,\theta_0,\sigma_{10}^2\right)} \propto \sigma_1^{-2}\left(1-\alpha\right)^{-1},$$

where α_0, θ_0 and σ_{10}^2 are the three inner points in the ranges of the parameters.

A.3 Proof of the Reference Prior for the Group Ordering $\left\{\sigma_1^2, \theta, \alpha\right\}$

We are interested in the reference prior for the group ordering $\left\{\sigma_1^2, \theta, \alpha\right\}$ which means that σ_1^2 is the most important parameter and α is the least important parameter. In order to derive the reference prior, the inverse of the Fisher information matrix is needed. Let $S(\boldsymbol{\theta}) = H^{-1}(\boldsymbol{\theta})$ where H is the Fisher information matrix. Now

$$S\left(\sigma_1^2, \theta, \alpha\right) = \begin{bmatrix} A_{11} & A'_{21} & \cdots & A'_{m1} \\ A_{21} & A_{22} & \cdots & A'_{m2} \\ \vdots & & \ddots & \\ A_{m_1} & A_{m2} & \cdots & A_{mm} \end{bmatrix} = \begin{bmatrix} \frac{2(\sigma_1^2)^2}{I(J-1)} & 0 & \frac{-2(1-\alpha)\left(\sigma_1^2\right)}{I(J-1)} \\ 0 & \frac{\sigma_1^2}{IJ(1-\alpha)} & 0 \\ \frac{-2(1-\alpha)\left(\sigma_1^2\right)}{I(J-1)} & 0 & \frac{2J(1-\alpha)^2}{I(J-1)} \end{bmatrix}.$$

Define the truncated ranges for the 3 parameters as $\alpha \in [a_l, b_l]$, $\theta \in [c_l, d_l]$ and $\sigma_1^2 \in [e_l, f_l]$ where $c_l \to -\infty$, $b_l \to 1$, $d_l, f_l \to \infty$ and $a_1, e_l \to 0$. Now $h_1 \equiv H_1 \equiv A_{11}^{-1} = \frac{I(J-1)}{2(\sigma_1^2)^2}$, $h_2 = \left(A_{22} - B_2 H_1 B'_2\right)^{-1}$. S_1 is the upper left $N_1 \times N_1$ corner of S. We have $m = 3$ groups with $n_1 = 1$, $n_2 = 1$ and $n_3 = 3$ $\therefore S_1$ is the upper left 1×1 corner of S, that is $S_1 = \frac{2(\sigma_1^2)^2}{I(J-1)}$. Therefore $H_1 = S_1^{-1} = \frac{I(J-1)}{2(\sigma_1^2)^2}$ and $B_2 = A_{21} = 0$. Hence $h_2 = [A_{22} - (0) H_1 (0)]^{-1} = A_{22}^{-1} = \frac{IJ(1-\alpha)}{\sigma_1^2}$. Also $h_3 = \left(A_{33} - B_3 H_2 B'_3\right)^{-1}$ where S_2 is the upper left $N_2 \times N_2$ corner of S. Now $N_2 = n_1 + n_2 = 2$. Therefore S_2 is the upper left 2×2 corner of S, that is $S_2 = \begin{bmatrix} \frac{2(\sigma_1^2)^2}{I(J-1)} & 0 \\ 0 & \frac{\sigma_1^2}{IJ(1-\alpha)} \end{bmatrix}$. Now $H_2 = S_2^{-1} = \begin{bmatrix} \frac{I(J-1)}{2(\sigma_1^2)^2} & 0 \\ 0 & \frac{IJ(1-\alpha)}{\sigma_1^2} \end{bmatrix}$ and $B_3 = \begin{bmatrix} A_{31} & A_{32} \end{bmatrix} = \begin{bmatrix} \frac{-2(1-\alpha)\left(\sigma_1^2\right)}{I(J-1)} & 0 \end{bmatrix}$. Therefore, $B_3 H_2 B'_3 = \frac{2(1-\alpha)^2}{I(J-1)}$. So $h_3 = \left(\frac{2J(1-\alpha)^2}{I(J-1)} - \frac{2(1-\alpha)^2}{I(J-1)}\right)^{-1} = \frac{I}{2(1-\alpha)^2}$. The three functions

$$h_1 = \frac{I(J-1)}{2(\sigma_1^2)^2}, \qquad h_2 = \frac{IJ(1-\alpha)}{\sigma_1^2}, \qquad h_3 = \frac{I}{2(1-\alpha)^2}$$

are the functions needed to calculate the prior. During the iterations, first the truncated conditional function of α given σ_1^2 and θ can be computed as

$$\begin{aligned} \pi_3^l\left(\alpha|\sigma_1^2, \theta\right) &= \frac{\{h_3(\boldsymbol{\theta})\}^{\frac{1}{2}}}{\int_{\theta_3^l} \{h_3(\boldsymbol{\theta})\}^{\frac{1}{2}} d\theta_3} \\ &\propto \frac{(1-\alpha)^{-1}}{\int_{a_l}^{b_l} (1-\alpha)^{-1} d\alpha} \\ &= \frac{(1-\alpha)^{-1}}{log\left(\frac{1-a_l}{1-b_l}\right)} \qquad \text{for } a_l \leq \alpha \leq b_l. \end{aligned}$$

Now,

$$
\begin{aligned}
&E\left[log\left\{h_2\left(\boldsymbol{\theta}\right)\right\}^{\frac{1}{2}}\right]\\
&\propto \int_{a_l}^{b_l} \frac{1}{2}\frac{log\left(1-\alpha\right)}{\sigma_1} \times \frac{\left(1-\alpha\right)^{-1}}{log\left(\frac{1-a_l}{1-b_l}\right)} d\alpha\\
&= \frac{1}{2\sigma_1 log\left(\frac{1-a_l}{1-b_l}\right)} \int_{a_l}^{b_l} \frac{log\left(1-\alpha\right)}{\left(1-\alpha\right)} d\alpha\\
&= \frac{1}{2\sigma_1 log\left(\frac{1-a_l}{1-b_l}\right)} \left(\frac{1}{2}\left[log^2\left(1-a_l\right) - log^2\left(1-b_l\right)\right]\right)\\
&= \frac{K}{\sigma_1},
\end{aligned}
$$

where $K = \frac{1}{4}\frac{\left[log^2(1-a_l)-log^2(1-b_l)\right]}{log\left(\frac{1-a_l}{1-b_l}\right)}$ is denoted as a constant which only relates to the ranges of the parameters. Therefore,

$$
\begin{aligned}
\pi_2^l\left(\theta, \alpha | \sigma_1^2\right) &= \frac{\pi_3^l\left(\alpha | \sigma_1^2, \theta\right) exp\left\{E\left[log\left\{h_2\left(\boldsymbol{\theta}\right)\right\}^{\frac{1}{2}}\right]\right\}}{\int_{c_l}^{d_l} exp\left\{E\left[log\left\{h_2\left(\boldsymbol{\theta}\right)\right\}^{\frac{1}{2}}\right]\right\} d\theta}\\
&= \frac{\pi_3^l\left(\alpha | \sigma_1^2, \theta\right) exp\left\{\frac{K}{\sigma_1}\right\}}{\int_{c_l}^{d_l} exp\left\{\frac{K}{\sigma_1}\right\} d\theta}\\
&= \frac{\left(1-\alpha\right)^{-1}}{log\left(\frac{1-a_l}{1-b_l}\right)} \times \frac{exp\left\{\frac{K}{\sigma_1}\right\}}{exp\left\{\frac{K}{\sigma_1}\right\}\left(d_l - c_l\right)}\\
&= \frac{\left(1-\alpha\right)^{-1}}{log\left(\frac{1-a_l}{1-b_l}\right)\left(d_l - c_l\right)} \qquad \text{for } a_l \leq \alpha \leq b_l \text{ and } c_l \leq \theta \leq d_l.
\end{aligned}
$$

We now need the function, h_1, to determine $E\left[log\left\{h_1(\boldsymbol{\theta})\right\}^{\frac{1}{2}}\right]$

$$\begin{aligned}
&E\left[log\left\{h_1(\boldsymbol{\theta})\right\}^{\frac{1}{2}}\right] \\
&\propto \int_{a_l}^{b_l}\int_{c_l}^{d_l} log\left(\sigma_1^2\right)^{-1} \frac{(1-\alpha)^{-1}}{log\left(\frac{1-a_l}{1-b_l}\right)(d_l-c_l)} d\theta d\alpha \\
&= \frac{log\left(\sigma_1^2\right)^{-1}}{log\left(\frac{1-a_l}{1-b_l}\right)} \int_{a_l}^{b_l} (1-\alpha)^{-1} d\alpha \\
&= \frac{log\left(\sigma_1^2\right)^{-1}}{log\left(\frac{1-a_l}{1-b_l}\right)} \times log\left(\frac{1-a_l}{1-b_l}\right) \\
&= log\left(\sigma_1^2\right)^{-1}.
\end{aligned}$$

Therefore,

$$\begin{aligned}
\pi_1^l\left(\sigma_1^2,\theta,\alpha\right) &= \frac{\pi_2^l\left(\theta,\alpha|\sigma_1^2\right) exp\left\{log\left(\sigma_1^2\right)^{-1}\right\}}{\int_{e_l}^{f_l} exp\left\{log\left(\sigma_1^2\right)^{-1}\right\} d\sigma_1^2} \\
&= \frac{(1-\alpha)^{-1}}{log\left(\frac{1-a_l}{1-b_l}\right)(d_l-c_l)} \frac{\left(\sigma_1^2\right)^{-1}}{\int_{e_l}^{f_l}\left(\sigma_1^2\right)^{-1} d\sigma_1^2} \\
&= \frac{(1-\alpha)^{-1}\left(\sigma_1^2\right)^{-1}}{log\left(\frac{1-a_l}{1-b_l}\right)(d_l-c_l)\, log(f_l e_l^{-1})}
\end{aligned}$$

Finally,

$$\pi_R\left(\alpha,\theta,\sigma_1^2\right) \propto \lim_{l\to\infty} \frac{\pi_1^l\left(\sigma_1^2,\theta,\alpha\right)}{\pi_1^l\left(\sigma_{10}^2,\theta_0,\alpha_0\right)} \propto \sigma_1^{-2}(1-\alpha)^{-1}$$

where α_0, θ_0 and σ_{10}^2 are the three inner points in the ranges of the parameters.

A.4 Proof of the Reference Prior for the Group Ordering $\left\{\sigma_1^2, \alpha, \theta\right\}$

We are interested in the reference prior for the group ordering $\left\{\sigma_1^2, \theta, \alpha\right\}$ which means that σ_1^2 is the most important parameter and θ is the least important parameter. In order to derive the reference prior, the inverse of the Fisher information matrix is needed. Let $S(\boldsymbol{\theta}) = H^{-1}(\boldsymbol{\theta})$ where H is the Fisher information matrix. Now,

$$S\left(\sigma_1^2, \alpha, \theta\right) = \begin{bmatrix} A_{11} & A'_{21} & \cdots & A'_{m1} \\ A_{21} & A_{22} & \cdots & A'_{m2} \\ \vdots & & \ddots & \\ A_{m_1} & A_{m2} & \cdots & A_{mm} \end{bmatrix} = \begin{bmatrix} \frac{2(\sigma_1^2)^2}{I(J-1)} & \frac{-2(1-\alpha)\left(\sigma_1^2\right)}{I(J-1)} & 0 \\ \frac{-2(1-\alpha)\left(\sigma_1^2\right)}{I(J-1)} & \frac{2J(1-\alpha)^2}{I(J-1)} & 0 \\ 0 & 0 & \frac{\sigma_1^2}{IJ(1-\alpha)} \end{bmatrix}.$$

Define the truncated ranges for the 3 parameters as $\alpha \in [a_l, b_l]$, $\theta \in [c_l, d_l]$ and $\sigma_1^2 \in [e_l, f_l]$ where $c_l \to -\infty$, $b_l \to 1$, $d_l, f_l \to \infty$ and $a_1, e_l \to 0$. Now $h_1 \equiv H_1 \equiv A_{11}^{-1} = \frac{I(J-1)}{2(\sigma_1^2)^2}$, $h_2 = \left(A_{22} - B_2 H_1 B_2'\right)^{-1}$. S_1 is the upper left $N_1 \times N_1$ corner of S. We have $m = 3$ groups with $n_1 = 1$, $n_2 = 1$ and $n_3 = 3$ $\therefore S_1$ is the upper left 1×1 corner of S, that is $S_1 = \frac{2(\sigma_1^2)^2}{I(J-1)}$. Therefore $H_1 = S_1^{-1} = \frac{I(J-1)}{2(\sigma_1^2)^2}$ and $B_2 = A_{21} = \frac{-2(1-\alpha)\left(\sigma_1^2\right)}{I(J-1)}$. Hence,

$$\begin{aligned} h_2 &= \left[A_{22} - \left(\frac{-2(1-\alpha)\left(\sigma_1^2\right)}{I(J-1)}\right) H_1 \left(\frac{-2(1-\alpha)\left(\sigma_1^2\right)}{I(J-1)}\right)\right]^{-1} \\ &= \left[\frac{2J(1-\alpha)^2}{I(J-1)} - \frac{2(1-\alpha)^2}{I(J-1)}\right]^{-1} \\ &= \frac{I}{2(1-\alpha)^2}. \end{aligned}$$

Also, $h_3 = \left(A_{33} - B_3 H_2 B_3'\right)^{-1}$ where S_2 is the upper left $N_2 \times N_2$ corner of S. Now $N_2 = n_1 + n_2 = 2$. Therefore S_2 is the upper left 2×2 corner of S, that is $S_2 = \begin{bmatrix} \frac{2(\sigma_1^2)^2}{I(J-1)} & \frac{-2(1-\alpha)\left(\sigma_1^2\right)}{I(J-1)} \\ \frac{-2(1-\alpha)\left(\sigma_1^2\right)}{I(J-1)} & \frac{2J(1-\alpha)^2}{I(J-1)} \end{bmatrix}$. Now, $H_2 = S_2^{-1} = \begin{bmatrix} \frac{IJ}{2(\sigma_1^2)^2} & \frac{I}{2\left(\sigma_1^2\right)(1-\alpha)} \\ \frac{I}{2\left(\sigma_1^2\right)(1-\alpha)} & \frac{I}{2(1-\alpha)^2} \end{bmatrix}$ and $B_3 = \begin{bmatrix} A_{31} & A_{32} \end{bmatrix} = \begin{bmatrix} 0 & 0 \end{bmatrix}$. Therefore, $B_3 H_2 B_3' = 0$. So $h_3 = A_{33}^{-1} = \frac{IJ(1-\alpha)}{\sigma_1^2}$. The three functions

$$h_1 = \frac{I(J-1)}{2(\sigma_1^2)^2}, \qquad h_2 = \frac{I}{2(1-\alpha)^2}, \qquad h_3 = \frac{IJ(1-\alpha)}{\sigma_1^2}$$

are the functions needed to calculate the prior. During the iterations, first the truncated conditional function of θ given σ_1^2 and α can be computed as

$$\begin{aligned}
\pi_3^l\left(\theta|\sigma_1^2,\alpha\right) =& \frac{\{h_3(\boldsymbol{\theta})\}^{\frac{1}{2}}}{\int_{\theta_3^l}\{h_3(\boldsymbol{\theta})\}^{\frac{1}{2}}d\theta_3} \\
\propto& \frac{(1-\alpha)^{\frac{1}{2}}\sigma_1^{-1}}{\int_{c_l}^{d_l}(1-\alpha)^{\frac{1}{2}}\sigma_1^{-1}d\theta} \\
=& \frac{1}{(d_l-c_l)} \qquad \text{for } c_l \le \theta \le d_l.
\end{aligned}$$

Now,

$$\begin{aligned}
& E\left[log\{h_2(\boldsymbol{\theta})\}^{\frac{1}{2}}\right] \\
\propto& \int_{c_l}^{d_l} -log(1-\alpha)\frac{1}{(d_l-c_l)}d\theta \\
=& log(1-\alpha)^{-1}
\end{aligned}$$

Therefore,

$$\begin{aligned}
\pi_2^l\left(\alpha,\theta|\sigma_1^2\right) =& \frac{\pi_3^l\left(\theta|\sigma_1^2,\alpha\right)exp\left\{E\left[log\{h_2(\boldsymbol{\theta})\}^{\frac{1}{2}}\right]\right\}}{\int_{a_l}^{b_l}exp\left\{E\left[log\{h_2(\boldsymbol{\theta})\}^{\frac{1}{2}}\right]\right\}d\alpha} \\
=& \frac{\pi_3^l\left(\theta|\sigma_1^2,\alpha\right)(1-\alpha)^{-1}}{\int_{a_l}^{b_l}(1-\alpha)^{-1}d\alpha} \\
=& \frac{(1-\alpha)^{-1}}{(d_l-c_l)log\left(\frac{1-a_l}{1-b_l}\right)} \\
=& \frac{K}{(1-\alpha)} \qquad \text{for } a_l \le \alpha \le b_l \text{ and } c_l \le \theta \le d_l,
\end{aligned}$$

where $K = \frac{1}{(d_l-c_l)log\left(\frac{1-a_l}{1-b_l}\right)}$ is denoted as a constant which only relates to the ranges of the parameters. We now need the function, h_1, to determine $E\left[log\{h_1(\boldsymbol{\theta})\}^{\frac{1}{2}}\right]$

$$
\begin{aligned}
&E\left[log\left\{h_1\left(\boldsymbol{\theta}\right)\right\}^{\frac{1}{2}}\right]\\
\propto&\int_{c_l}^{d_l}\int_{a_l}^{b_l} log\left(\sigma_1^2\right)^{-1}\frac{K}{(1-\alpha)}d\alpha d\theta\\
=&\int_{c_l}^{d_l} Klog\left(\sigma_1^2\right)^{-1}\left(\int_{a_l}^{b_l}(1-\alpha)^{-1}d\alpha\right)d\theta\\
=&\int_{c_l}^{d_l} Klog\left(\sigma_1^2\right)^{-1}log\left(\frac{1-a_l}{1-b_l}\right)d\theta\\
=&Klog\left(\sigma_1^2\right)^{-1}log\left(\frac{1-a_l}{1-b_l}\right)(d_l-c_l)\\
=&\frac{1}{(d_l-c_l)\,log\left(\frac{1-a_l}{1-b_l}\right)}log\left(\sigma_1^2\right)^{-1}log\left(\frac{1-a_l}{1-b_l}\right)(d_l-c_l)\\
=&log\left(\sigma_1^2\right)^{-1}.
\end{aligned}
$$

Therefore,

$$
\begin{aligned}
\pi_1^l\left(\sigma_1^2,\theta,\alpha\right)=&\frac{\pi_2^l\left(\alpha,\theta|\sigma_1^2\right)exp\left\{log\left(\sigma_1^2\right)^{-1}\right\}}{\int_{e_l}^{f_l}exp\left\{log\left(\sigma_1^2\right)^{-1}\right\}d\sigma_1^2}\\
=&\frac{K}{(1-\alpha)}\frac{\left(\sigma_1^2\right)^{-1}}{\int_{e_l}^{f_l}\left(\sigma_1^2\right)^{-1}d\sigma_1^2}\\
=&\frac{K(1-\alpha)^{-1}\left(\sigma_1^2\right)^{-1}}{log(f_le_l^{-1})}
\end{aligned}
$$

Finally,

$$
\pi_R\left(\alpha,\theta,\sigma_1^2\right)\propto\lim_{l\to\infty}\frac{\pi_1^l\left(\sigma_1^2,\theta,\alpha\right)}{\pi_1^l\left(\sigma_{10}^2,\theta_0,\alpha_0\right)}\propto\sigma_1^{-2}(1-\alpha)^{-1}
$$

where α_0,θ_0 and σ_{10}^2 are the three inner points in the ranges of the parameters.

A.5 Proof of the Reference Prior for the Group Ordering $\left\{\theta, \sigma_1^2, \alpha,\right\}$

We are interested in the reference prior for the group ordering $\left\{\theta, \sigma_1^2, \alpha\right\}$ which means that θ is the most important parameter and α is the least important parameter. In order to derive the reference prior, the inverse of the Fisher information matrix is needed. Let $S(\boldsymbol{\theta}) = H^{-1}(\boldsymbol{\theta})$ where H is the Fisher information matrix. Now

$$S\left(\theta, \sigma_1^2, \alpha\right) = \begin{bmatrix} A_{11} & A'_{21} & \cdots & A'_{m1} \\ A_{21} & A_{22} & \cdots & A'_{m2} \\ \vdots & & \ddots & \\ A_{m_1} & A_{m2} & \cdots & A_{mm} \end{bmatrix} = \begin{bmatrix} \frac{\sigma_1^2}{IJ(1-\alpha)} & 0 & 0 \\ 0 & \frac{2(\sigma_1^2)^2}{I(J-1)} & \frac{-2(1-\alpha)\left(\sigma_1^2\right)}{I(J-1)} \\ 0 & \frac{-2(1-\alpha)\left(\sigma_1^2\right)}{I(J-1)} & \frac{2J(1-\alpha)^2}{I(J-1)} \end{bmatrix}.$$

Define the truncated ranges for the 3 parameters as $\alpha \in [a_l, b_l]$, $\theta \in [c_l, d_l]$ and $\sigma_1^2 \in [e_l, f_l]$ where $c_l \to -\infty$, $b_l \to 1$, $d_l, f_l \to \infty$ and $a_1, e_l \to 0$. Now $h_1 \equiv H_1 \equiv A_{11}^{-1} = \frac{IJ(1-\alpha)}{\sigma_1^2}$, $h_2 = \left(A_{22} - B_2 H_1 B_2'\right)^{-1}$. S_1 is the upper left $N_1 \times N_1$ corner of S. We have $m = 3$ groups with $n_1 = 1$, $n_2 = 1$ and $n_3 = 3 \therefore S_1$ is the upper left 1×1 corner of S, that is $S_1 = \frac{\sigma_1^2}{IJ(1-\alpha)}$. Therefore $H_1 = S_1^{-1} = \frac{IJ(1-\alpha)}{\sigma_1^2}$ and $B_2 = A_{21} = 0$. Hence $h_2 = A_{22}^{-1} = \frac{I(J-1)}{2(\sigma_1^2)^2}$ Also $h_3 = \left(A_{33} - B_3 H_2 B_3'\right)^{-1}$ where S_2 is the upper left $N_2 \times N_2$ corner of S. Now $N_2 = n_1 + n_2 = 2$. Therefore S_2 is the upper left 2×2 corner of S, that is $S_2 = \begin{bmatrix} \frac{\sigma_1^2}{IJ(1-\alpha)} & 0 \\ 0 & \frac{2(\sigma_1^2)^2}{I(J-1)} \end{bmatrix}$. Now $H_2 = S_2^{-1} = \begin{bmatrix} \frac{IJ(1-\alpha)}{\sigma_1^2} & 0 \\ 0 & \frac{I(J-1)}{2(\sigma_1^2)^2} \end{bmatrix}$ and $B_3 = \begin{bmatrix} A_{31} & A_{32} \end{bmatrix} = \begin{bmatrix} 0 & \frac{-2(1-\alpha)\left(\sigma_1^2\right)}{I(J-1)} \end{bmatrix}$. Therefore, $B_3 H_2 B_3' = \frac{2(1-\alpha)^2}{I(J-1)}$. So

$$\begin{aligned} h_3 &= \left[A_{33} - B_3 H_2 B_3'\right]^{-1} \\ &= \left[\frac{2J(1-\alpha)^2}{I(J-1)} - \frac{2(1-\alpha)^2}{I(J-1)}\right]^{-1} \\ &= \frac{I}{2(1-\alpha)^2}. \end{aligned}$$

The three functions

$$h_1 = \frac{IJ(1-\alpha)}{\sigma_1^2}, \qquad h_2 = \frac{I(J-1)}{2(\sigma_1^2)^2}, \qquad h_3 = \frac{I}{2(1-\alpha)^2}$$

are the functions needed to calculate the prior. During the iterations, first the truncated conditional function of α given θ and σ_1^2 can be computed as

$$\begin{aligned}\pi_3^l\left(\alpha|\theta,\sigma_1^2\right) &= \frac{\{h_3(\boldsymbol{\theta})\}^{\frac{1}{2}}}{\int_{\theta_3^l}\{h_3(\boldsymbol{\theta})\}^{\frac{1}{2}}\,d\theta_3} \\ &\propto \frac{(1-\alpha)^{-1}}{\int_{a_l}^{b_l}(1-\alpha)^{-1}\,d\alpha} \\ &= \frac{(1-\alpha)^{-1}}{log\left(\frac{1-a_l}{1-b_l}\right)} \qquad \text{for } a_l \le \alpha \le b_l.\end{aligned}$$

Now,

$$\begin{aligned}& E\left[log\{h_2(\boldsymbol{\theta})\}^{\frac{1}{2}}\right] \\ \propto & \int_{a_l}^{b_l} log\left(\sigma_1^2\right)^{-1}\frac{(1-\alpha)^{-1}}{log\left(\frac{1-a_l}{1-b_l}\right)}d\alpha \\ = & \frac{log\left(\sigma_1^2\right)^{-1}}{log\left(\frac{1-a_l}{1-b_l}\right)}log\left(\frac{1-a_l}{1-b_l}\right) \\ = & log\left(\sigma_1^2\right)^{-1}.\end{aligned}$$

Therefore,

$$\begin{aligned}\pi_2^l\left(\sigma_1^2,\alpha|\theta\right) &= \frac{\pi_3^l\left(\alpha|\theta,\sigma_1^2\right)exp\left\{E\left[log\{h_2(\boldsymbol{\theta})\}^{\frac{1}{2}}\right]\right\}}{\int_{e_l}^{f_l}exp\left\{E\left[log\{h_2(\boldsymbol{\theta})\}^{\frac{1}{2}}\right]\right\}d\sigma_1^2} \\ &= \frac{\pi_3^l\left(\theta|\sigma_1^2,\alpha\right)\left(\sigma_1^2\right)^{-1}}{\int_{e_l}^{f_l}\left(\sigma_1^2\right)^{-1}d\sigma_1^2} \\ &= \frac{(1-\alpha)^{-1}}{log\left(\frac{1-a_l}{1-b_l}\right)}\frac{\left(\sigma_1^2\right)^{-1}}{log\left(f_l e_l^{-1}\right)} \qquad \text{for } a_l \le \alpha \le b_l \text{ and } e_l \le \sigma_1^2 \le f_l.\end{aligned}$$

We now need the function, h_1, to determine $E\left[log\{h_1(\boldsymbol{\theta})\}^{\frac{1}{2}}\right]$

$$
\begin{aligned}
&E\left[log\left\{h_1(\boldsymbol{\theta})\right\}^{\frac{1}{2}}\right] \\
\propto&\int_{a_l}^{b_l}\int_{e_l}^{f_l} log\left[\frac{(1-\alpha)^{\frac{1}{2}}}{\left(\sigma_1^2\right)^{\frac{1}{2}}}\right]\frac{(1-\alpha)^{-1}}{log\left(\frac{1-a_l}{1-b_l}\right)}\frac{\left(\sigma_1^2\right)^{-1}}{log\left(f_l e_l^{-1}\right)}d\sigma_1^2 d\alpha \\
=&\frac{1}{2log\left(\frac{1-a_l}{1-b_l}\right)log\left(f_l e_l^{-1}\right)}\int_{a_l}^{b_l}(1-\alpha)^{-1}\int_{e_l}^{f_l}\left[log(1-\alpha)-log\left(\sigma_1^2\right)\right]\left(\sigma_1^2\right)^{-1}d\sigma_1^2 d\alpha \\
=&\frac{1}{2log\left(\frac{1-a_l}{1-b_l}\right)log\left(f_l e_l^{-1}\right)}\int_{a_l}^{b_l}(1-\alpha)^{-1}\left[\int_{e_l}^{f_l}log(1-\alpha)\left(\sigma_1^2\right)^{-1}d\sigma_1^2-\int_{e_l}^{f_l}log\left(\sigma_1^2\right)\left(\sigma_1^2\right)^{-1}d\sigma_1^2\right]d\alpha \\
=&K^{**}\int_{a_l}^{b_l}(1-\alpha)^{-1}\left[log(1-\alpha)\,log\left(f_l e_l^{-1}\right)-\frac{1}{2}\left[log^2(f_l)-log^2(e_l)\right]\right]d\alpha \\
=&K^{**}log\left(f_l e_l^{-1}\right)\int_{a_l}^{b_l}(1-\alpha)^{-1}log(1-\alpha)\,d\alpha-\frac{K^{**}\left[log^2(f_l)-log^2(e_l)\right]}{2}\int_{a_l}^{b_l}(1-\alpha)^{-1}d\alpha \\
=&\frac{K^{**}log\left(f_l e_l^{-1}\right)}{2}\left[log^2(1-a_l)-log^2(1-b_l)\right]-\frac{K^{**}\left[log^2(f_l)-log^2(e_l)\right]}{2}log\left(\frac{1-a_l}{1-b_l}\right) \\
=&K'
\end{aligned}
$$

where $K^{**}=\frac{1}{2log\left(\frac{1-a_l}{1-b_l}\right)log\left(f_l e_l^{-1}\right)}$ and K' are just constants which only relate to the ranges of the parameters. Therefore,

$$
\begin{aligned}
\pi_1^l\left(\theta,\sigma_1^2,\alpha\right)=&\frac{\pi_2^l\left(\sigma_1^2,\alpha|\theta\right)exp\left\{log\left(\sigma_1^2\right)^{-1}\right\}}{\int_{e_l}^{f_l}exp\left\{log\left(\sigma_1^2\right)^{-1}\right\}d\sigma_1^2} \\
=&\frac{(1-\alpha)^{-1}}{log\left(\frac{1-a_l}{1-b_l}\right)}\frac{\left(\sigma_1^2\right)^{-1}}{log\left(f_l e_l^{-1}\right)}\frac{exp(K')}{\int_{c_l}^{d_l}exp(K')\,d\theta} \\
=&\frac{(1-\alpha)^{-1}}{log\left(\frac{1-a_l}{1-b_l}\right)}\frac{\left(\sigma_1^2\right)^{-1}}{log\left(f_l e_l^{-1}\right)(d_l-c_l)}.
\end{aligned}
$$

Finally,

$$
\pi_R\left(\alpha,\theta,\sigma_1^2\right)\propto\lim_{l\to\infty}\frac{\pi_1^l\left(\sigma_1^2,\theta,\alpha\right)}{\pi_1^l\left(\sigma_{10}^2,\theta_0,\alpha_0\right)}\propto\sigma_1^{-2}(1-\alpha)^{-1}
$$

where α_0,θ_0 and σ_{10}^2 are the three inner points in the ranges of the parameters.

A.6 Proof of the Reference Prior for the Group Ordering $\left\{\theta, \alpha, \sigma_1^2\right\}$

We are interested in the reference prior for the group ordering $\left\{\theta, \alpha, \sigma_1^2\right\}$ which means that θ is the most important parameter and σ_1^2 is the least important parameter. In order to derive the reference prior, the inverse of the Fisher information matrix is needed. Let $S(\boldsymbol{\theta}) = H^{-1}(\boldsymbol{\theta})$ where H is the Fisher information matrix. Now

$$S\left(\theta, \sigma_1^2, \alpha\right) = \begin{bmatrix} A_{11} & A'_{21} & \cdots & A'_{m1} \\ A_{21} & A_{22} & \cdots & A'_{m2} \\ \vdots & & \ddots & \\ A_{m_1} & A_{m2} & \cdots & A_{mm} \end{bmatrix} = \begin{bmatrix} \frac{\sigma_1^2}{IJ(1-\alpha)} & 0 & 0 \\ 0 & \frac{2J(1-\alpha)^2}{I(J-1)} & \frac{-2(1-\alpha)\left(\sigma_1^2\right)}{I(J-1)} \\ 0 & \frac{-2(1-\alpha)\left(\sigma_1^2\right)}{I(J-1)} & \frac{2(\sigma_1^2)^2}{I(J-1)} \end{bmatrix}.$$

Define the truncated ranges for the 3 parameters as $\alpha \in [a_l, b_l]$, $\theta \in [c_l, d_l]$ and $\sigma_1^2 \in [e_l, f_l]$ where $c_l \to -\infty$, $b_l \to 1$, $d_l, f_l \to \infty$ and $a_l, e_l \to 0$. Now $h_1 \equiv H_1 \equiv A_{11}^{-1} = \frac{IJ(1-\alpha)}{\sigma_1^2}$, $h_2 = \left(A_{22} - B_2 H_1 B'_2\right)^{-1}$. S_1 is the upper left $N_1 \times N_1$ corner of S. We have $m = 3$ groups with $n_1 = 1$, $n_2 = 1$ and $n_3 = 3 \therefore S_1$ is the upper left 1×1 corner of S, that is $S_1 = \frac{\sigma_1^2}{IJ(1-\alpha)}$. Therefore $H_1 = S_1^{-1} = \frac{IJ(1-\alpha)}{\sigma_1^2}$ and $B_2 = A_{21} = 0$. Hence $h_2 = A_{22}^{-1} = \frac{I(J-1)}{2J(1-\alpha)^2}$ Also $h_3 = \left(A_{33} - B_3 H_2 B'_3\right)^{-1}$ where S_2 is the upper left $N_2 \times N_2$ corner of S. Now $N_2 = n_1 + n_2 = 2$. Therefore S_2 is the upper left 2×2 corner of S, that is $S_2 = \begin{bmatrix} \frac{\sigma_1^2}{IJ(1-\alpha)} & 0 \\ 0 & \frac{2J(1-\alpha)^2}{I(J-1)} \end{bmatrix}$.

Now $H_2 = S_2^{-1} = \begin{bmatrix} \frac{IJ(1-\alpha)}{\sigma_1^2} & 0 \\ 0 & \frac{I(J-1)}{2J(1-\alpha)^2} \end{bmatrix}$ and $B_3 = \begin{bmatrix} A_{31} & A_{32} \end{bmatrix} = \begin{bmatrix} 0 & \frac{-2(1-\alpha)\left(\sigma_1^2\right)}{I(J-1)} \end{bmatrix}$. Therefore, $B_3 H_2 B'_3 = \frac{2\left(\sigma_1^2\right)^2}{IJ(J-1)}$. So,

$$\begin{aligned} h_3 &= \left[A_{33} - B_3 H_2 B'_3\right]^{-1} \\ &= \left[\frac{2(\sigma_1^2)^2}{I(J-1)} - \frac{2\left(\sigma_1^2\right)^2}{IJ(J-1)}\right]^{-1} \\ &= \frac{IJ}{2\left(\sigma_1^2\right)^2}. \end{aligned}$$

The three functions

$$h_1 = \frac{IJ(1-\alpha)}{\sigma_1^2}, \qquad h_2 = \frac{I(J-1)}{2J(1-\alpha)^2}, \qquad h_3 = \frac{IJ}{2\left(\sigma_1^2\right)^2}$$

are the functions needed to calculate the prior. During the iterations, first the truncated conditional function of α given θ and σ_1^2 can be computed as

$$\begin{aligned}
\pi_3^l\left(\alpha|\theta,\sigma_1^2\right) =& \frac{\{h_3(\boldsymbol{\theta})\}^{\frac{1}{2}}}{\int_{\theta_3^l}\{h_3(\boldsymbol{\theta})\}^{\frac{1}{2}}\,d\theta_3} \\
\propto& \frac{(1-\alpha)^{-1}}{\int_{a_l}^{b_l}(1-\alpha)^{-1}\,d\alpha} \\
=& \frac{(1-\alpha)^{-1}}{log\left(\frac{1-a_l}{1-b_l}\right)} \qquad \text{for } a_l \le \alpha \le b_l.
\end{aligned}$$

Now,

$$\begin{aligned}
&E\left[log\{h_2(\boldsymbol{\theta})\}^{\frac{1}{2}}\right] \\
\propto& \int_{a_l}^{b_l} log\left(\sigma_1^2\right)^{-1}\frac{(1-\alpha)^{-1}}{log\left(\frac{1-a_l}{1-b_l}\right)}d\alpha \\
=& \frac{log\left(\sigma_1^2\right)^{-1}}{log\left(\frac{1-a_l}{1-b_l}\right)}log\left(\frac{1-a_l}{1-b_l}\right) \\
=& log\left(\sigma_1^2\right)^{-1}.
\end{aligned}$$

Therefore,

$$\begin{aligned}
\pi_2^l\left(\sigma_1^2,\alpha|\theta\right) =& \frac{\pi_3^l\left(\alpha|\theta,\sigma_1^2\right)exp\left\{E\left[log\{h_2(\boldsymbol{\theta})\}^{\frac{1}{2}}\right]\right\}}{\int_{e_l}^{f_l}exp\left\{E\left[log\{h_2(\boldsymbol{\theta})\}^{\frac{1}{2}}\right]\right\}d\sigma_1^2} \\
=& \frac{\pi_3^l\left(\theta|\sigma_1^2,\alpha\right)\left(\sigma_1^2\right)^{-1}}{\int_{e_l}^{f_l}\left(\sigma_1^2\right)^{-1}d\sigma_1^2} \\
=& \frac{(1-\alpha)^{-1}}{log\left(\frac{1-a_l}{1-b_l}\right)}\frac{\left(\sigma_1^2\right)^{-1}}{log\left(f_l e_l^{-1}\right)} \qquad \text{for } a_l \le \alpha \le b_l \text{ and } e_l \le \sigma_1^2 \le f_l.
\end{aligned}$$

We now need the function, h_1, to determine $E\left[log\{h_1(\boldsymbol{\theta})\}^{\frac{1}{2}}\right]$

$$
\begin{aligned}
& E\left[log\left\{h_1(\boldsymbol{\theta})\right\}^{\frac{1}{2}}\right] \\
\propto & \int_{a_l}^{b_l}\int_{e_l}^{f_l} log\left[\frac{(1-\alpha)^{\frac{1}{2}}}{\left(\sigma_1^2\right)^{\frac{1}{2}}}\right]\frac{(1-\alpha)^{-1}}{log\left(\frac{1-a_l}{1-b_l}\right)}\frac{\left(\sigma_1^2\right)^{-1}}{log\left(f_l e_l^{-1}\right)}d\sigma_1^2 d\alpha \\
= & \frac{1}{2log\left(\frac{1-a_l}{1-b_l}\right)log\left(f_l e_l^{-1}\right)}\int_{a_l}^{b_l}(1-\alpha)^{-1}\int_{e_l}^{f_l}\left[log(1-\alpha)-log\left(\sigma_1^2\right)\right]\left(\sigma_1^2\right)^{-1}d\sigma_1^2 d\alpha \\
= & \frac{1}{2log\left(\frac{1-a_l}{1-b_l}\right)log\left(f_l e_l^{-1}\right)}\int_{a_l}^{b_l}(1-\alpha)^{-1}\left[\int_{e_l}^{f_l} log(1-\alpha)\left(\sigma_1^2\right)^{-1}d\sigma_1^2-\int_{e_l}^{f_l} log\left(\sigma_1^2\right)\left(\sigma_1^2\right)^{-1}d\sigma_1^2\right]d\alpha \\
= & K^{**}\int_{a_l}^{b_l}(1-\alpha)^{-1}\left[log(1-\alpha)log\left(f_l e_l^{-1}\right)-\frac{1}{2}\left[log^2(f_l)-log^2(e_l)\right]\right]d\alpha \\
= & K^{**}log\left(f_l e_l^{-1}\right)\int_{a_l}^{b_l}(1-\alpha)^{-1}log(1-\alpha)d\alpha-\frac{K^{**}\left[log^2(f_l)-log^2(e_l)\right]}{2}\int_{a_l}^{b_l}(1-\alpha)^{-1}d\alpha \\
= & \frac{K^{**}log\left(f_l e_l^{-1}\right)}{2}\left[log^2(1-a_l)-log^2(1-b_l)\right]-\frac{K^{**}\left[log^2(f_l)-log^2(e_l)\right]}{2}log\left(\frac{1-a_l}{1-b_l}\right) \\
= & K'
\end{aligned}
$$

where $K^{**}=\frac{1}{2log\left(\frac{1-a_l}{1-b_l}\right)log\left(f_l e_l^{-1}\right)}$ and K' are just constants which only relate to the ranges of the parameters. Therefore,

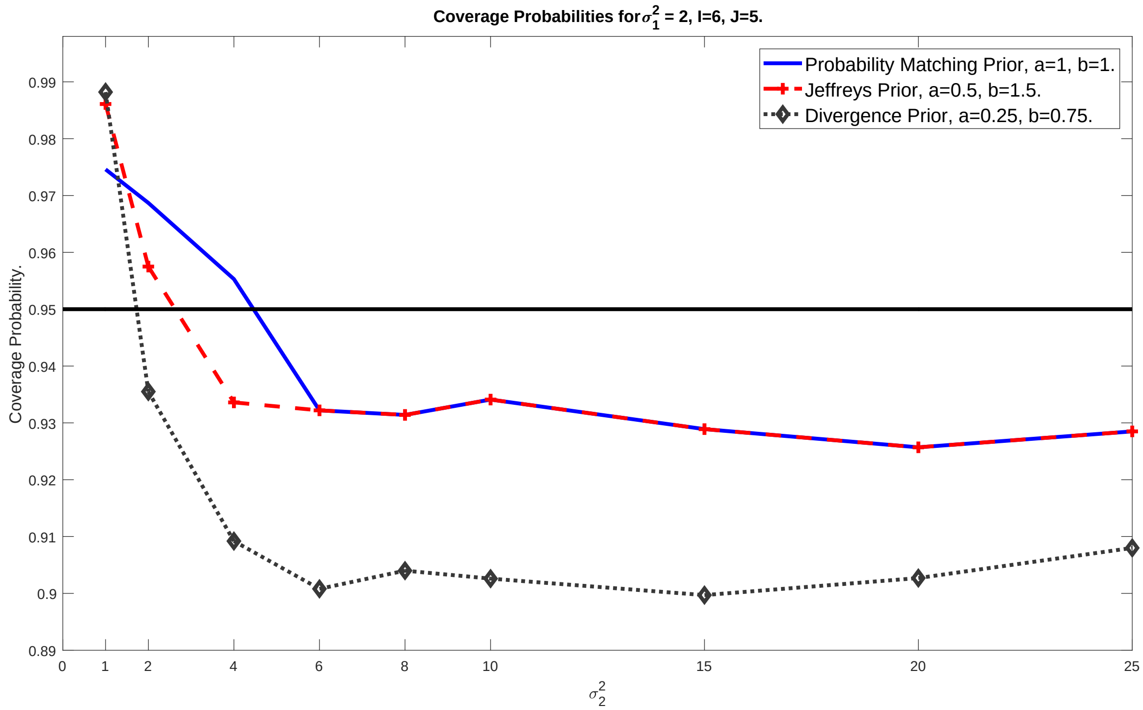
$$
\begin{aligned}
\pi_1^l\left(\theta,\sigma_1^2,\alpha\right)= & \frac{\pi_2^l\left(\sigma_1^2,\alpha|\theta\right)exp\left\{log\left(\sigma_1^2\right)^{-1}\right\}}{\int_{e_l}^{f_l} exp\left\{log\left(\sigma_1^2\right)^{-1}\right\}d\sigma_1^2} \\
= & \frac{(1-\alpha)^{-1}}{log\left(\frac{1-a_l}{1-b_l}\right)}\frac{\left(\sigma_1^2\right)^{-1}}{log\left(f_l e_l^{-1}\right)}\frac{exp(K')}{\int_{c_l}^{d_l} exp(K')d\theta} \\
= & \frac{(1-\alpha)^{-1}}{log\left(\frac{1-a_l}{1-b_l}\right)}\frac{\left(\sigma_1^2\right)^{-1}}{log\left(f_l e_l^{-1}\right)(d_l-c_l)}.
\end{aligned}
$$

Finally,

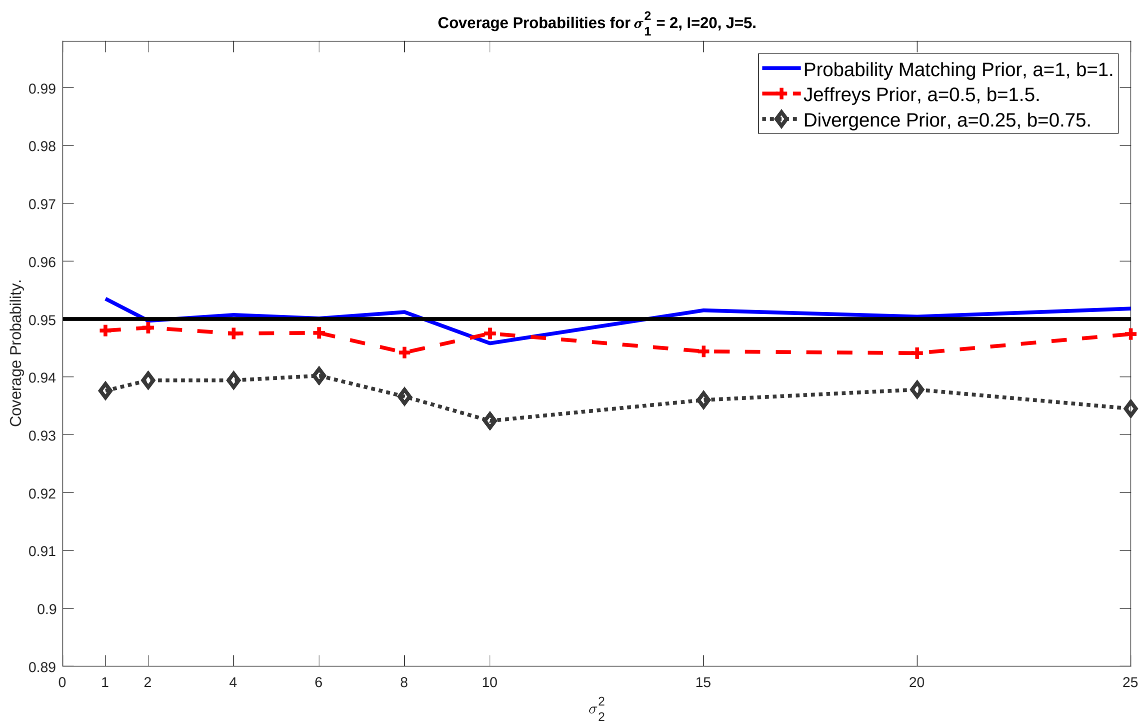
$$
\pi_R\left(\alpha,\theta,\sigma_1^2\right)\propto\lim_{l\to\infty}\frac{\pi_1^l\left(\sigma_1^2,\theta,\alpha\right)}{\pi_1^l\left(\sigma_{10}^2,\theta_0,\alpha_0\right)}\propto\sigma_1^{-2}(1-\alpha)^{-1}
$$

where α_0,θ_0 and σ_{10}^2 are the three inner points in the ranges of the parameters.

A.7 Additional Results for Simulation Study 1

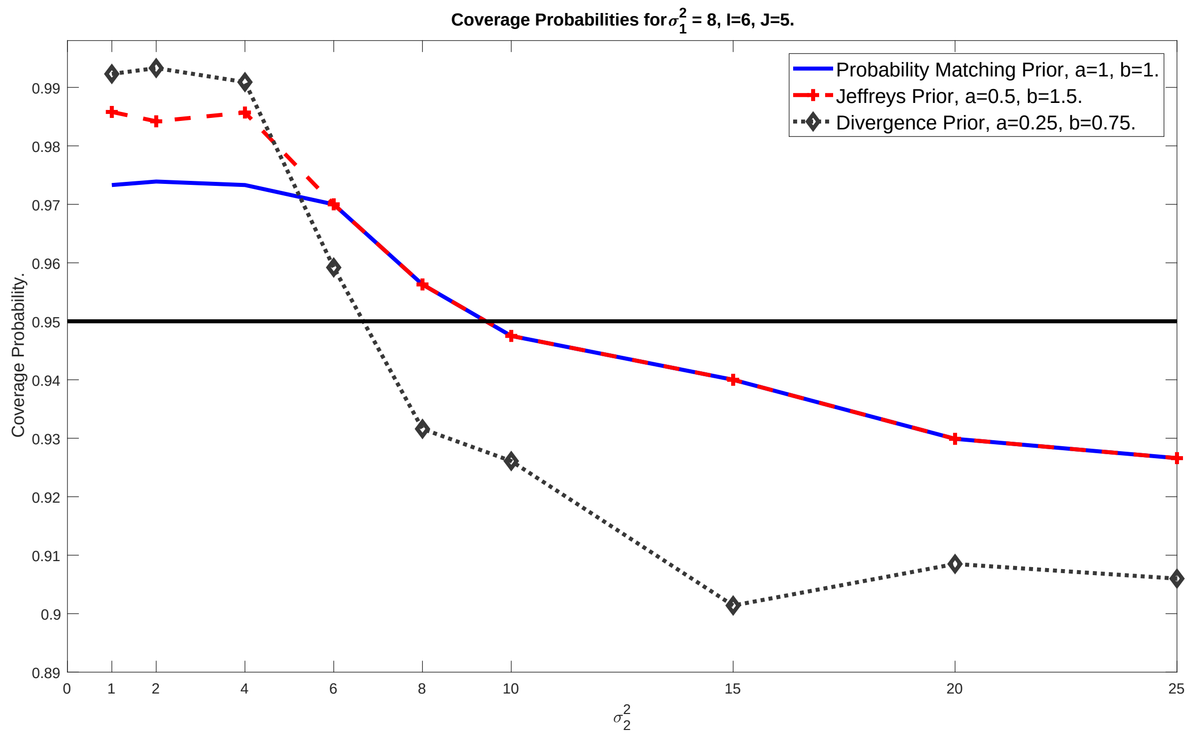


(a)

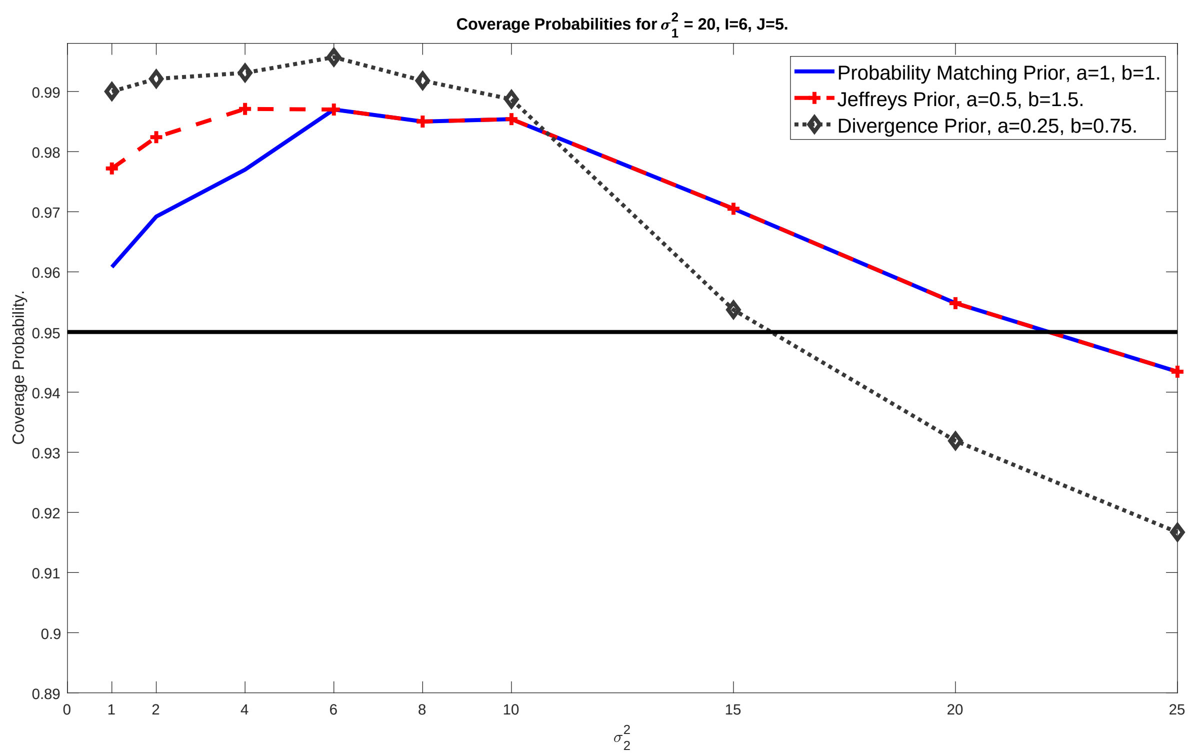


(b)

Figure A.1: Coverage for $\sigma_1^2 = 2$ When (a) $I = 6$, $J = 5$ and (b) $I = 20, J = 5$.

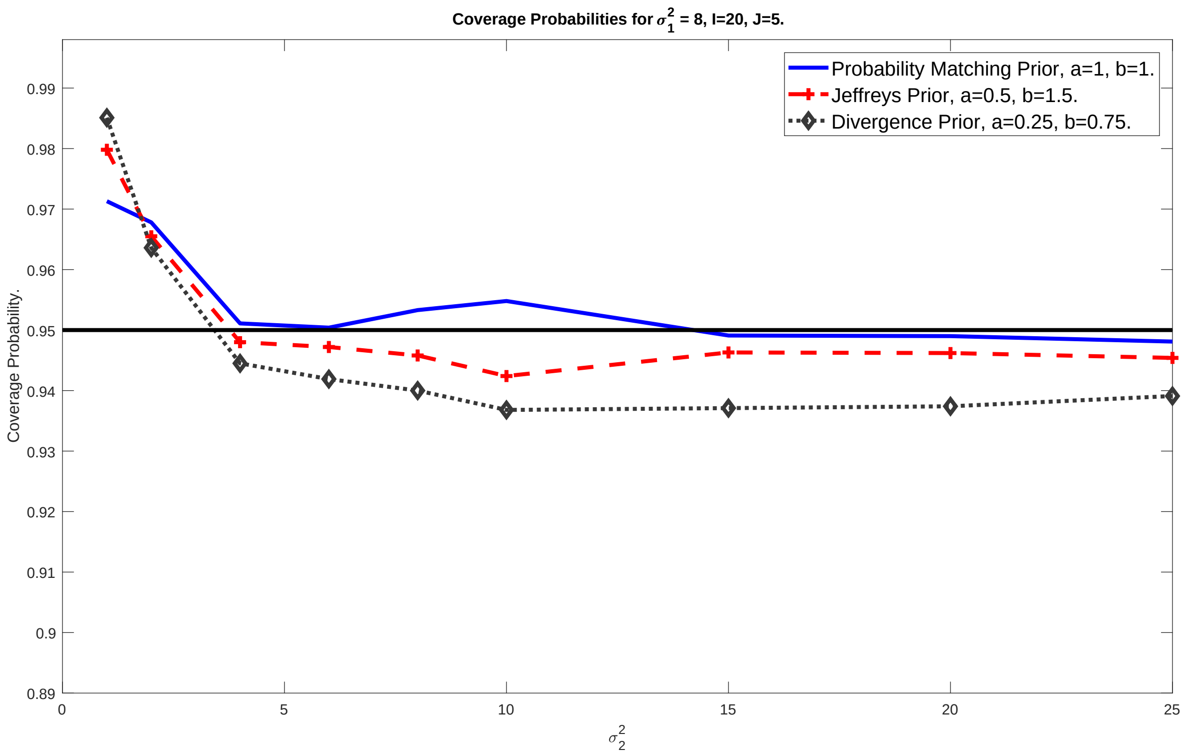


(a)

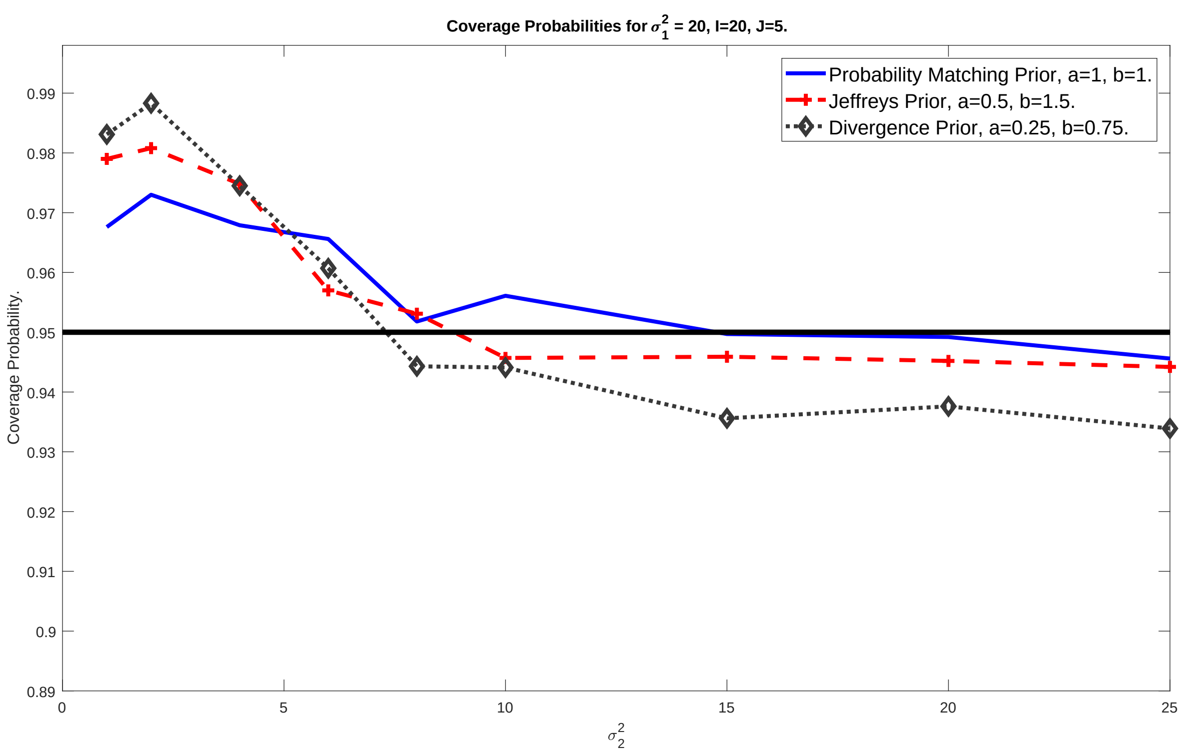


(b)

Figure A.2: Coverage for $I = 6$, $J = 5$, (a) $\sigma_1^2 = 8$ and (b) $\sigma_1^2 = 20$.

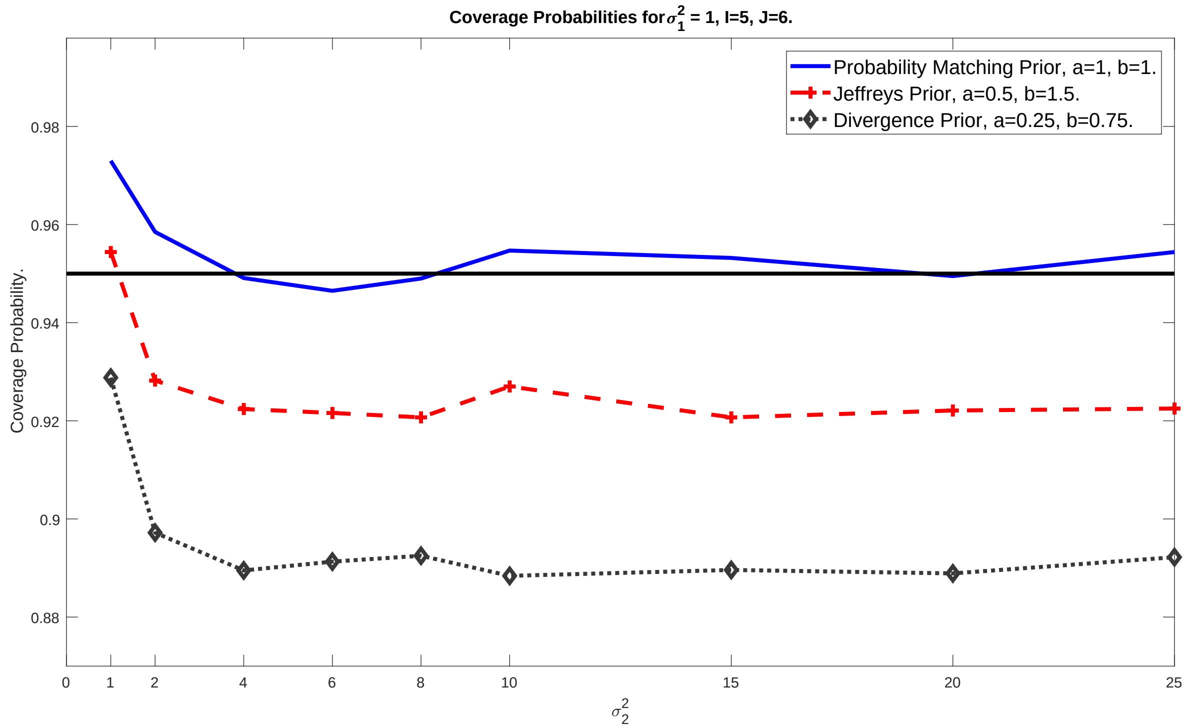


(a)

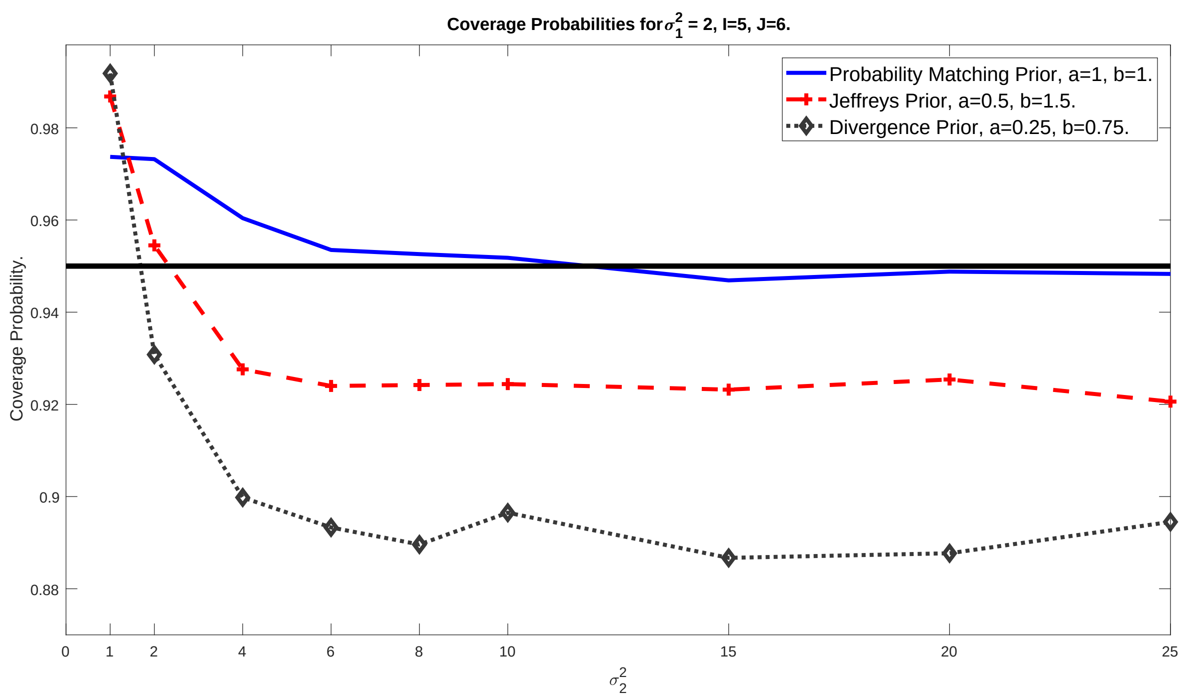


(b)

Figure A.3: Coverage for $I = 20$, $J = 5$, (a) $\sigma_1^2 = 8$ and (b) $\sigma_1^2 = 20$.

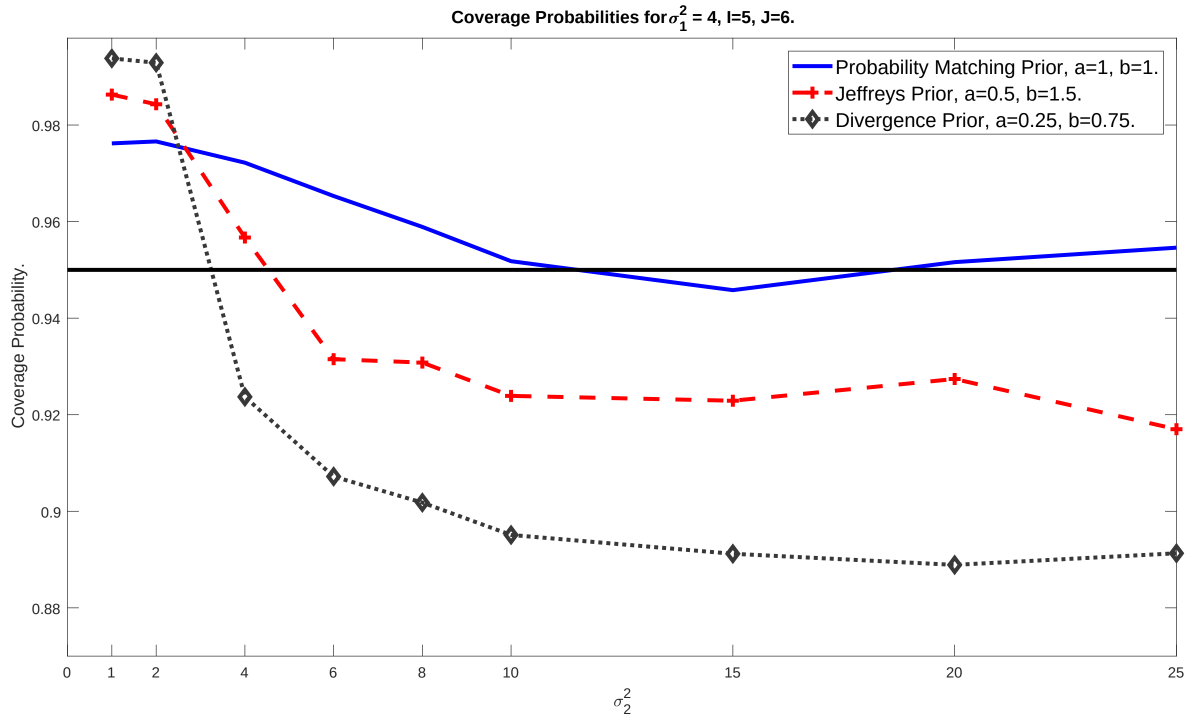


(a)

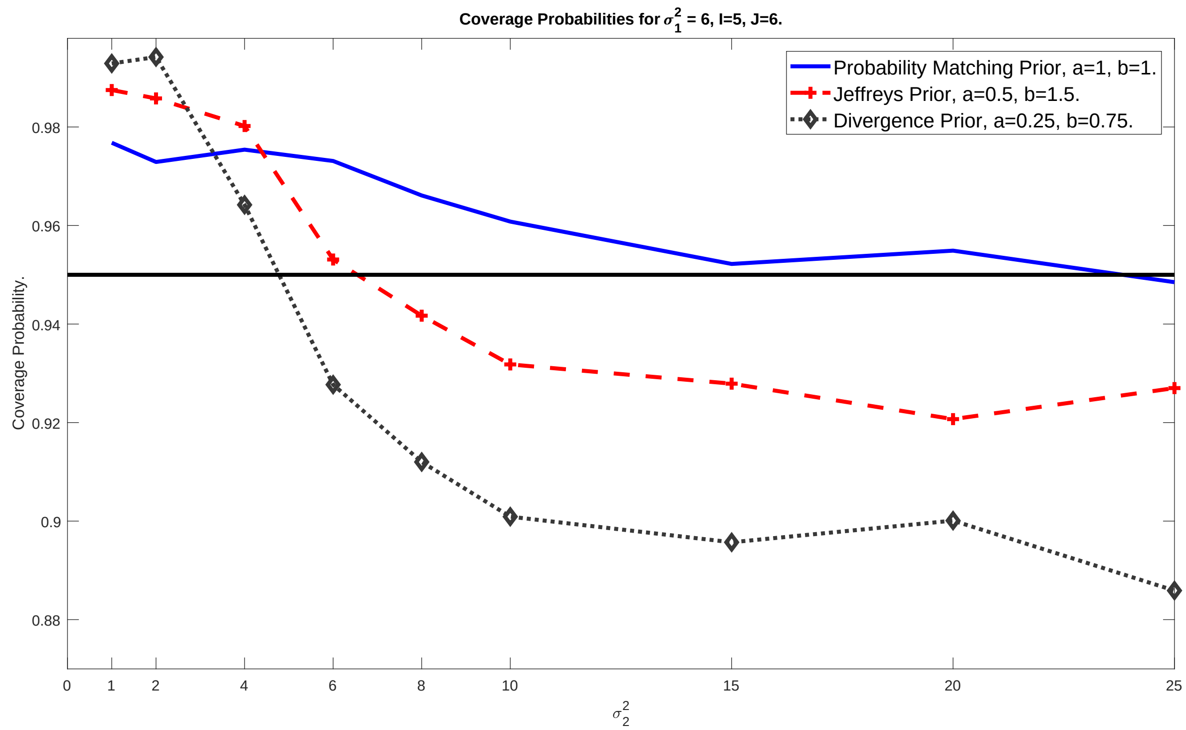


(b)

Figure A.4: Coverage for $I = 5$, $J = 6$, (a) $\sigma_1^2 = 1$ and (b) $\sigma_1^2 = 2$.

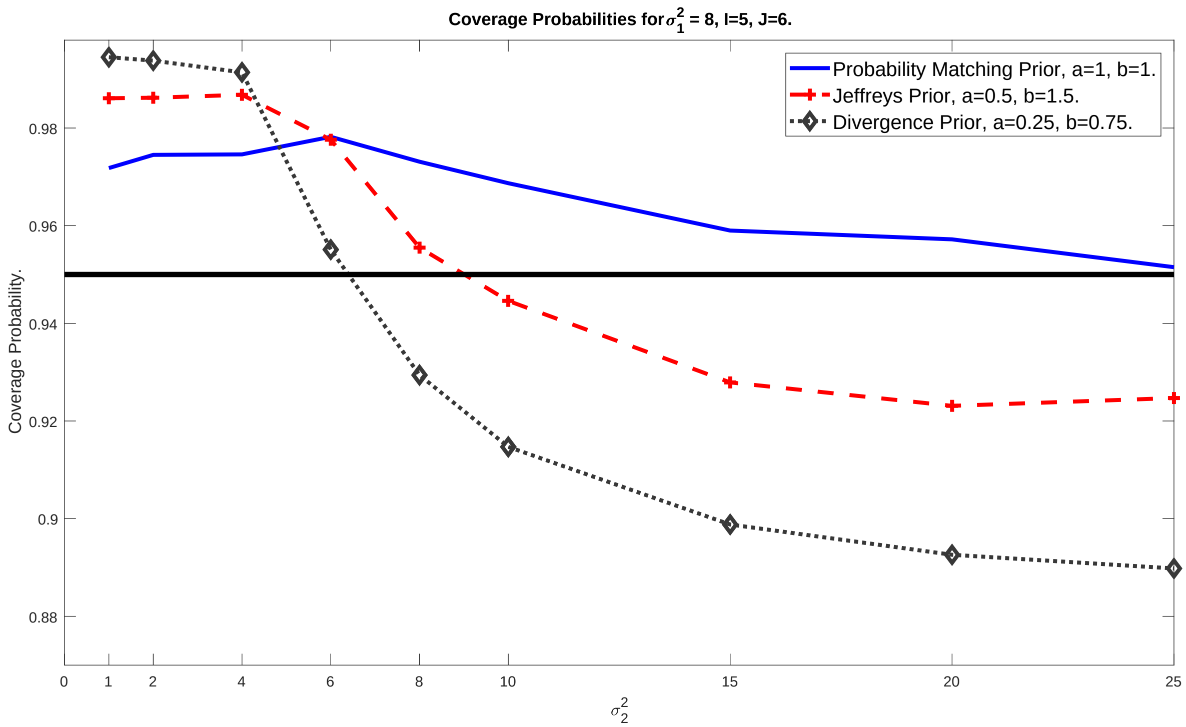


(a)

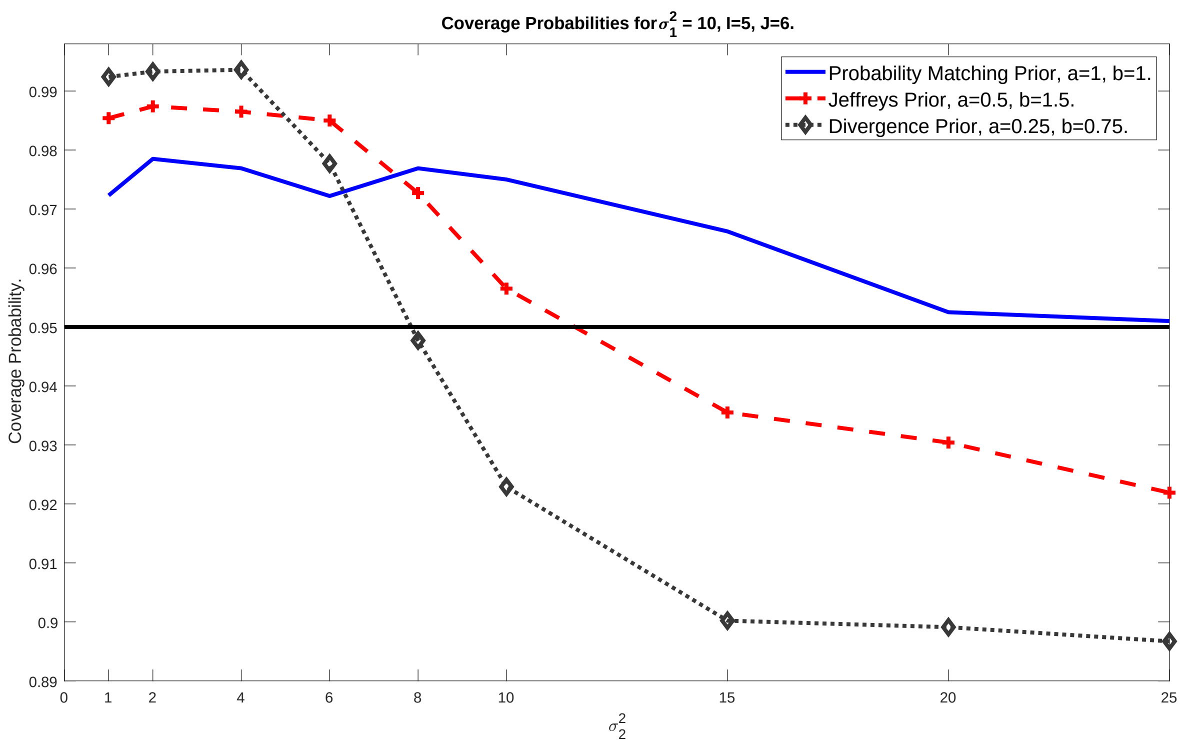


(b)

Figure A.5: Coverage for $I = 5$, $J = 6$, (a) $\sigma_1^2 = 4$ and (b) $\sigma_1^2 = 6$.

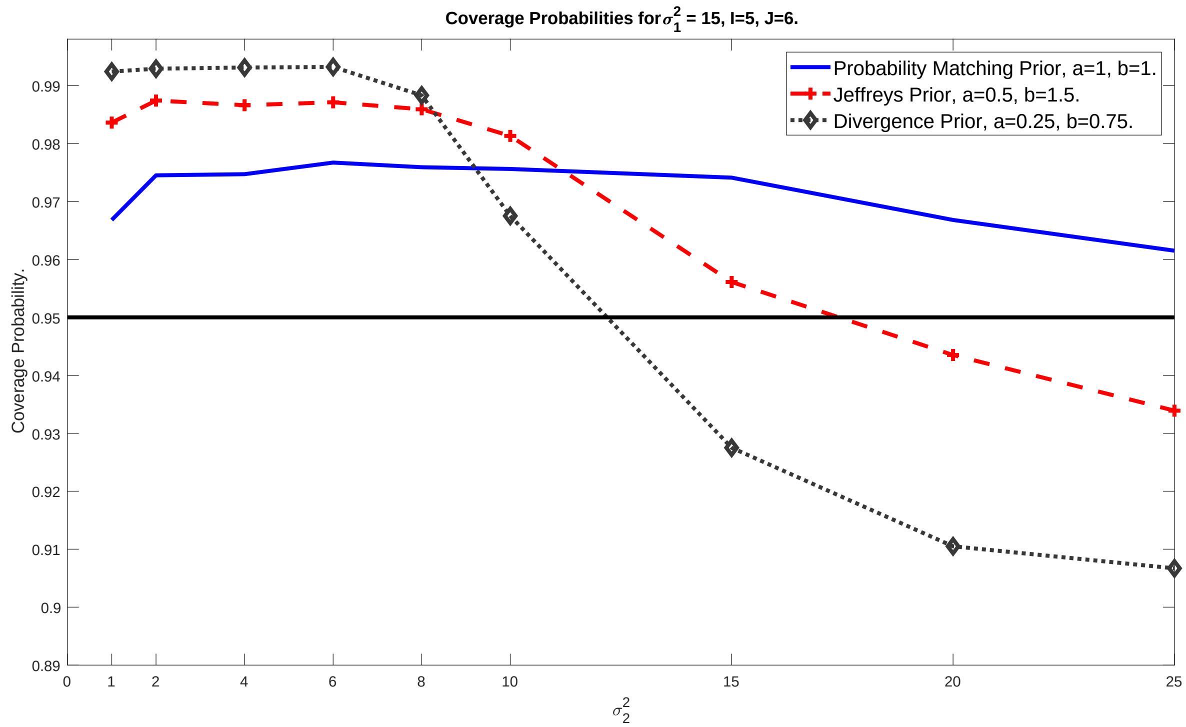


(a)

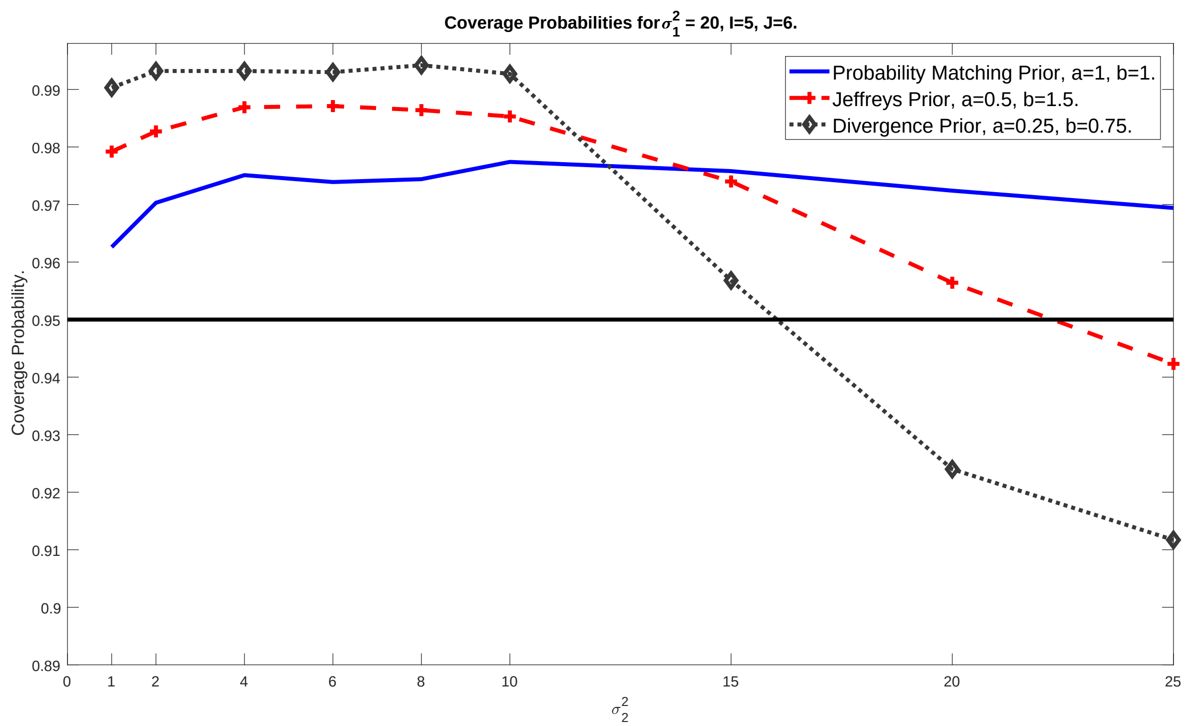


(b)

Figure A.6: Coverage for $I = 5$, $J = 6$, (a) $\sigma_1^2 = 8$ and (b) $\sigma_1^2 = 10$.



(a)



(b)

Figure A.7: Coverage for $I = 5$, $J = 6$, (a) $\sigma_1^2 = 8$ and (b) $\sigma_1^2 = 10$.

A.8 Additional Results for Simulation Study 2

Table A.1: MRE and Estimates Using Different Loss Functions and Values for α When $I = 6$ and $J = 6$ When $\sigma_1^2 = 6$ and $\sigma_2^2 = 3, 10, 15$ and 20.

$\alpha = 0.75$						
	MATCHING		JEFFREYS		DIVERGENCE	
	MRE	EST	MRE	EST	MRE	EST
Squared Error	0.9152	0.6864	0.9049	0.6786	0.9028	0.6771
Absolute Error	0.9529	0.7147	0.9382	0.7037	0.9353	0.7015
Linex Loss c: -9	1.0574	0.7931	1.0398	0.7798	1.0340	0.7755
-5	1.0092	0.7569	**0.9927**	**0.7445**	0.9879	0.7409
0.5	0.9033	0.6775	0.8940	0.6705	0.8923	0.6693
1	0.8910	0.6683	0.8828	0.6621	0.8815	0.6612
$\alpha = 0.9091$						
	MATCHING		JEFFREYS		DIVERGENCE	
	MRE	EST	MRE	EST	MRE	EST
Squared Error	0.9400	0.8546	0.9385	0.8532	0.9373	0.8521
Absolute Error	0.9626	0.8751	0.9585	0.8714	0.9568	0.8698
Linex Loss c: -9	**0.9815**	**0.8923**	0.9768	0.8880	0.9747	0.8861
-5	0.9672	0.8793	0.9633	0.8757	0.9615	0.8741
0.5	0.9365	0.8514	0.9353	0.8503	0.9342	0.8493
1	0.9328	0.8480	0.9320	0.8473	0.9310	0.8464
$\alpha = 0.9375$						
	MATCHING		JEFFREYS		DIVERGENCE	
	MRE	EST	MRE	EST	MRE	EST
Squared Error	0.9554	0.8957	0.9547	0.8950	0.9540	0.8944
Absolute Error	0.9722	0.9115	0.9696	0.9090	0.9685	0.9079
Linex Loss c: -9	**0.9801**	**0.9189**	0.9773	0.9163	0.9760	0.9150
-5	0.9715	0.9107	0.9692	0.9087	0.9681	0.9076
0,5	0.9533	0.8938	0.9529	0.8933	0.9522	0.8927
1	0.9512	0.8917	0.9509	0.8915	0.9503	0.8909
$\alpha = 0.9524$						
	MATCHING		JEFFREYS		DIVERGENCE	
	MRE	EST	MRE	EST	MRE	EST
Squared Error	0.9653	0.9193	0.9644	0.9185	0.9639	0.9180
Absolute Error	0.9785	0.9319	0.9762	0.9297	0.9754	0.9289
Linex Loss c: -9	**0.9817**	**0.9350**	0.9796	0.9329	0.9786	0.9320
-5	0.9759	0.9294	0.9741	0.9277	0.9733	0.9269
0,5	0.9639	0.9180	0.9632	0.9173	0.9627	0.9169
1	0.9625	0.9166	0.9619	0.9161	0.9615	0.9157

Table A.2: MRE and Estimates Using Different Loss Functions and Values for α When $I = 3$ and $J = 30$ When $\sigma_1^2 = 6$ and $\sigma_2^2 = 2, 3, 6$ and 8.

$\alpha = 0.9091$						
	MATCHING		JEFFREYS		DIVERGENCE	
	MRE	EST	MRE	EST	MRE	EST
Squared Error	0.9036	0.8215	0.8848	0.8044	0.8792	0.7993
Absolute Error	0.9435	0.8577	0.9115	0.8286	0.9018	0.8198
Linex Loss c: -9	**0.9881**	**0.8983**	0.9656	0.8778	0.9572	0.8702
-5	0.9629	0.8753	0.9393	0.8539	0.9310	0.8464
0.5	0.8952	0.8139	0.8778	0.7980	0.8728	0.7934
1	0.8864	0.8058	0.8705	0.7914	0.8661	0.7874
$\alpha = 0.9375$						
	MATCHING		JEFFREYS		DIVERGENCE	
	MRE	EST	MRE	EST	MRE	EST
Squared Error	0.9134	0.8563	0.9003	0.8441	0.8959	0.8399
Absolute Error	0.9466	0.8874	0.9228	0.8651	0.9152	0.8580
Linex Loss c: -9	**0.9762**	**0.9152**	0.9596	0.8996	0.9531	0.8936
-5	0.9573	0.8975	0.9402	0.8815	0.9338	0.8755
0.5	0.9072	0.8505	0.8952	0.8392	0.8911	0.8354
1	0.9006	0.8444	0.8898	0.8342	0.8862	0.8308
$\alpha = 0.9677$						
	MATCHING		JEFFREYS		DIVERGENCE	
	MRE	EST	MRE	EST	MRE	EST
Squared Error	0.9361	0.9059	0.9292	0.8993	0.9269	0.8970
Absolute Error	0.9589	0.9280	0.9450	0.9145	0.9405	0.9102
Linex Loss c: -9	**0.9723**	**0.9409**	0.9628	0.9317	0.9591	0.9282
-5	0.9612	0.9302	0.9517	0.9210	0.9482	0.9176
0.5	0.9325	0.9024	0.9263	0.8965	0.9243	0.8944
1	0.9288	0.8988	0.9233	0.8935	0.9215	0.8918
$\alpha = 0.9756$						
	MATCHING		JEFFREYS		DIVERGENCE	
	MRE	EST	MRE	EST	MRE	EST
Squared Error	0.9453	0.9223	0.9402	0.9172	0.9384	0.9155
Absolute Error	0.9646	0.9410	0.9534	0.9302	0.9499	0.9267
Linex Loss c: -9	**0.9737**	**0.9500**	0.9663	0.9427	0.9634	0.9399
-5	0.9650	0.9415	0.9576	0.9343	0.9549	0.9316
0.5	0.9426	0.9196	0.9379	0.9151	0.9364	0.9135
1	0.9397	0.9167	0.9356	0.9128	0.9342	0.9114

Table A.3: MRE and Estimates Using Different Loss Functions and Values for α When $I = 10$ and $J = 50$ When $\sigma_1^2 = 6$ and $\sigma_2^2 = 2, 3, 6$ and 8.

$\alpha = 0.9434$						
	MATCHING		JEFFREYS		DIVERGENCE	
	MRE	EST	MRE	EST	MRE	EST
Squared Error	0.9824	0.9268	0.9826	0.9270	0.9824	0.9268
Absolute Error	0.9882	0.9323	0.9878	0.9319	0.9874	0.9315
Linex Loss c: -9	**0.9892**	**0.9332**	0.9888	0.9328	0.9884	0.9324
-5	0.9865	0.9306	0.9863	0.9304	0.9859	0.9301
0.5	0.9820	0.9264	0.9822	0.9266	0.9820	0.9264
1	0.9815	0.9260	0.9818	0.9262	0.9816	0.9261
$\alpha = 0.9615$						
	MATCHING		JEFFREYS		DIVERGENCE	
	MRE	EST	MRE	EST	MRE	EST
Squared Error	0.9888	0.9508	0.9882	0.9502	0.9882	0.9502
Absolute Error	**0.9927**	**0.9545**	0.9918	0.9536	0.9916	0.9535
Linex Loss c: -9	0.9920	0.9538	0.9912	0.9531	0.9911	0.9530
-5	0.9907	0.9525	0.9900	0.9519	0.9899	0.9518
0.5	0.9886	0.9506	0.9880	0.9500	0.9880	0.9500
1	0.9884	0.9504	0.9878	0.9498	0.9879	0.9499
$\alpha = 0.9804$						
	MATCHING		JEFFREYS		DIVERGENCE	
	MRE	EST	MRE	EST	MRE	EST
Squared Error	0.9941	0.9746	0.9941	0.9746	0.9941	0.9746
Absolute Error	**0.9961**	**0.9765**	0.9959	0.9763	0.9958	0.9763
Linex Loss c: -9	0.9950	0.9755	0.9949	0.9754	0.9949	0.9754
-5	0.9946	0.9751	0.9946	0.9751	0.9945	0.9750
0.5	0.9940	0.9746	0.9941	0.9746	0.9941	0.9746
1	0.9940	0.9745	0.9940	0.9745	0.9940	0.9745
$\alpha = 0.9852$						
	MATCHING		JEFFREYS		DIVERGENCE	
	MRE	EST	MRE	EST	MRE	EST
Squared Error	0.9956	0.9809	0.9956	0.9809	0.9956	0.9809
Absolute Error	**0.9970**	**0.9823**	0.9969	0.9822	0.9969	0.9821
Linex Loss c: -9	0.9961	0.9814	0.9960	0.9813	0.9961	0.9813
-5	0.9959	0.9811	0.9958	0.9811	0.9959	0.9811
0.5	0.9955	0.9808	0.9955	0.9808	0.9956	0.9809
1	0.9955	0.9808	0.9955	0.9808	0.9956	0.9808

A.9 MATLAB Code for Simulation Study 1

```
clear
tic
I=5; J=6; % change I and J accordingly
v1=I*(J-1);
v2=I-1;
Count=[];
Length=[];
SI=[];
for s1=[1 2 4 6 8 10 15 20 25];
s2=25;
alf=J*s2/(s1+J*s2); %Cronbach's Alpha
N=10000;
ro=0:0.0001:0.9999;
count=0;
L=0;
for i=1:N
v1m1=s1*chi2rnd(v1);
v2m2=(s1+J*s2)*chi2rnd(v2);
%log of posterior kernel using PMP
lnf=((v2-2)/2)*log(1-ro)-((v1+v2)/2)*log(v1m1+v2m2*(1-ro));
%log of posterior kernel using Jeffreys prior
%lnf=((v2-1)/2)*log(1-ro)-((v1+v2+1)/2)*log(v1m1+v2m2*(1-ro));
%log of posterior kernel using divergence prior
%lnf=((v2-0.5)/2)*log(1-ro)-((v1+v2-0.5)/2)*log(v1m1+v2m2*(1-ro));
f=exp(lnf); %Kernel of Posterior for Cronbach's alpha
K=inv(0.0001*sum(f)); %normalizing constant
pro=K*f; % Normalized Posterior for Cronbach's alpha
C=0.0001*cumsum(pro);
lo95=ro(min(find(C>=0.025)));
up95=ro(min(find(C>=0.975)));
l=up95-lo95;
L=L+l;
if lo95<=alf & up95>=alf
count=count+1;
end
```

```
end
count;
leng=L/N;
Count=[Count;count];
Length=[Length;leng];
si=sqrt(1/(N-1)*(sum(Length.^2)-(sum(Length))^2/N));
SI=[SI;si];
end
toc
coverage=Count./N
coverage_2=coverage' Length' SI'
% Use this function to ouput results in the form of a vector in Excel
xlswrite('Probability Matching',[coverage_2 Length' SI'])
```

A.10 MATLAB Code for Simulation Study 2

```
%Matlab code for Cronbach alpha Loss functions
clc
clear
tic
R=10000;
I=5;
J=20;
v1=I*(J-1);
v2=I-1;
%a=1; b=1; %Probability Matching Prior
%a=1/2; b=3/2; % Jeffreys Prior
a=1/4; b=3/4; % Divergence Prior
df1=v2+2-2*a; % Numerator degrees of freedom for truncated F posterior of Z
df2=v1+2*b+2*a-4; % Denominator degrees of freedom for truncated F posterior of Z
s1=6;
s2=8;
alpha=J*s2/(s1+J*s2);
alpha;
n=10000;
%These are the values for the Linex Loss function
c1=-9;
```

```
c2=-5;
c3=0.5;
c4=1;
ALPHA_HAT1=[];
ALPHA_HAT2=[];
ALPHA_HAT_LIN1=[];
ALPHA_HAT_LIN2=[];
ALPHA_HAT_LIN3=[];
ALPHA_HAT_LIN4=[];
Z=[];
for i=1:R
v1_star=s1*chi2rnd(v1); %v1m1
v2_star=(s1+J*s2)*chi2rnd(v2); %v2m2
m1=v1_star/v1;
m2=v2_star/v2;
m=m2/m1;
u = unifrnd(0,1,1,n); % Generate n uniform random numbers on (0,1)
% Simulating from truncated F posterior distribution
z = finv(fcdf(0,df1,df2)+u*(fcdf(m,df1,df2)-fcdf(0,df1,df2)),df1,df2);
alpha_post=1-(z./m); % Simulated Cronbach alpha posterior values
% Squared error loss
alpha_hat1=mean(alpha_post);
% Absolute error loss
alpha_hat2=median(alpha_post);
%Linex Loss
%Linex Loss with c1=-9
alpha_hat_LIN1=(-1/c1)*log(mean(exp(-c1*alpha_post)));
%Linex Loss with c2=-5
alpha_hat_LIN2=(-1/c2)*log(mean(exp(-c2*alpha_post)));
%Linex Loss with c3=0.5
alpha_hat_LIN3=(-1/c3)*log(mean(exp(-c3*alpha_post)));
%Linex Loss with c4=1
alpha_hat_LIN4=(-1/c4)*log(mean(exp(-c4*alpha_post)));
Z=[Z z];
%Squared error loss
ALPHA_HAT1=[ALPHA_HAT1 alpha_hat1 ];
%Absolute error loss
```

```
ALPHA_HAT2=[ALPHA_HAT2 alpha_hat2 ];
%Linex loss
ALPHA_HAT_LIN1=[ALPHA_HAT_LIN1 alpha_hat_LIN1];
ALPHA_HAT_LIN2=[ALPHA_HAT_LIN2 alpha_hat_LIN2];
ALPHA_HAT_LIN3=[ALPHA_HAT_LIN3 alpha_hat_LIN3];
ALPHA_HAT_LIN4=[ALPHA_HAT_LIN4 alpha_hat_LIN4];
end
MRE1=(sum((ALPHA_HAT1./alpha)))./R;
MRE2=(sum((ALPHA_HAT2./alpha)))./R;
MRE3=(sum((ALPHA_HAT_LIN1./alpha)))./R;
MRE4=(sum((ALPHA_HAT_LIN2./alpha)))./R;
MRE5=(sum((ALPHA_HAT_LIN3./alpha)))./R;
MRE6=(sum((ALPHA_HAT_LIN4./alpha)))./R;
EST1=mean(ALPHA_HAT1);
EST2=mean(ALPHA_HAT2);
EST3=mean(ALPHA_HAT_LIN1);
EST4=mean(ALPHA_HAT_LIN2);
EST5=mean(ALPHA_HAT_LIN3);
EST6=mean(ALPHA_HAT_LIN4);
Results_MRE=[MRE1 MRE2 MRE3 MRE4 MRE5 MRE6]'
Results_EST=[EST1 EST2 EST3 EST4 EST5 EST6]'
% Use this function to output results in Excel
xlswrite('Output',[Results_MRE Results_EST])
toc
```

A.11 MATLAB Code for the Example

```
% Bayes estimation of Cronbach's alpha for the Dyestuff data
clc
clear
I=6; J=5;
% Data
y=[1545 1540 1595 1445 1595 1520; 1440 1555 1550 1440 1630 1455; 1440 1490 1605
1595 1515 1450; 1520 1560 1510 1465 1635 1480; 1580 1495 1560 1545 1625 1445];
y_new_i_dot=sum(y(:,1:I))/J;  %y_i_dot_bar
ydot_dot=mean(mean(y));       %y_dot_dot_bar
v_1_star=sum(sum((y-y_new_i_dot).^2)); Within  group sum of squares
```

```
v_2_star=J*sum((y_new_i_dot-ydot_dot).^2); Between group sum of squares
v1=I*(J-1);
v2=I-1;
m1=v_1_star/v1;  %Within group mean square error
m2=v_2_star/v2;  %Between group mean square error
m=m2/m1;
% Function to determine normalizing constant
fun1= @(alpha) %Posterior using the PMP:a=1,b=1 (not normalized)
(1-alpha).^(0.5*(v2-2)).*(v_1_star+v_2_star.*(1-alpha)).^(-0.5*(v1+v2));
q1 = integral(fun1,0,1);
% Therefore normalizing constant K=1/q1
K=1/q1;
%Function to determine mean of posterior of alpha
fun2= @(alpha)(1/q1)*alpha.*(1-alpha).^(0.5*(v2-2)).
*(v_1_star+v_2_star.*(1-alpha)).^(-0.5*(v1+v2));
% Bayes estimate under squared error loss
alpha_hat_SE=integral(fun2,0,1)
%Bayes estimate for alpha under Absolute error loss
alpha_hat_ABS=1-finv(0.5,v2,v1)/m   % Theorem 4
%Parameters for the Linex loss function
c1=-9;
c2=-5;
c3=0.5;
c4=1;
% Function to determine E(exp(-c1*alpha)) for the Bayes est under Linex loss
fun3=@(alpha) (1/q1)*exp(-c1*alpha).*(1-alpha).^(0.5*(v2-2)).
*(v_1_star+v_2_star.*(1-alpha)).^(-0.5*(v1+v2));
E_LIN_1=integral(fun3,0,1);
alpha_hat_LIN_c1=(-1/c1)*log(E_LIN_1)
% Function to determine E(exp(-c2*alpha)) for the Bayes est under Linex loss
fun4=@(alpha)(1/q1)*exp(-c2*alpha).*(1-alpha).^(0.5*(v2-2)).
*(v_1_star+v_2_star.*(1-alpha)).^(-0.5*(v1+v2));
E_LIN_2=integral(fun4,0,1);
alpha_hat_LIN_c2=(-1/c2)*log(E_LIN_2)
% Function to determine E(exp(-c3*alpha)) for the Bayes est under Linex loss
fun5=@(alpha) (1/q1)*exp(-c3*alpha).*(1-alpha).^(0.5*(v2-2)).
*(v_1_star+v_2_star.*(1-alpha)).^(-0.5*(v1+v2));
```

```
E_LIN_3=integral(fun5,0,1);
alpha_hat_LIN_c3=(-1/c3)*log(E_LIN_3)
% Function to determine E(exp(-c4*alpha)) for the Bayes est under Linex  loss
fun6=@(alpha) (1/q1)*exp(-c4*alpha).*(1-alpha).^(0.5*(v2-2)).
*(v_1_star+v_2_star.*(1-alpha)).^(-0.5*(v1+v2));
E_LIN_4=integral(fun6,0,1);
alpha_hat_LIN_c4=(-1/c4)*log(E_LIN_4)
% Credibility Interval for alpha (Equal tailed).
%I used Z=m(1-alpha) as the % pivotal quantity for the CI from Theorem 4
CI_95=[1-finv(0.975,v2,v1)/m 1-finv(0.025,v2,v1)/m]
%Plot of the  posterior for alpha
x=[0:0.01:1];
y_1=(1/q1)*x.*(1-x).^(0.5*(v2-2)).*(v_1_star+v_2_star.*(1-x)).^(-0.5*(v1+v2));
plot(x,y_1)
title('Posterior Distribution of \alpha')
xlabel('\alpha')
ylabel('p(\alpha)')
```

Appendix B: Combined Bayesian Estimates for Cronbach's Alpha

B.1 MATLAB Code for Example 1

```
%COMPARE TWO DISTRIBUTIONS OF THE SAME ALPHA
% Example 1
clear
tic
s1=4; s2=1;
s11=3; s22=0.75;
I=20; J=6;
v1=I*(J-1); v2=I-1;
D=[]; D1=[]; D2=[]; Q=[]; Mo=[]; Mo1=[]; Mo2=[]; U=[]; R=[];
count=0; Len=[]; count1=0; Len1=[]; count2=0; Len2=[];
for i=1:10000
v1m1=s1*chi2rnd(v1);
v2m2=(s1+J*s2)*chi2rnd(v2);
v1m11=s11*chi2rnd(v1);
v2m22=(s11+J*s22)*chi2rnd(v2);
alf=0:0.001:1;
A=(v2m2*(1-alf)+v1m1).^(-(v1+v2)/2);
B=(v2m22*(1-alf)+v1m11).^(-(v1+v2)/2);
pa=(10^300)*((1-alf).^(1.5*v2-1)).*A.*B;
K=1/sum(pa)/0.001;
pa1=((1-alf).^((v2-1)/2)).*A;
pa2=((1-alf).^((v2-1)/2)).*B;
K1=1/sum(pa1)/0.001;
K2=1/sum(pa2)/0.001;
figure(1)
```

```
plot(alf,K*pa,alf,K1*pa1,alf,K2*pa2)
grid
Fa=pa*K*0.001;
Fa1=pa1*K1*0.001;
Fa2=pa2*K2*0.001;
mFa=max(Fa); mFa1=max(Fa1); mFa2=max(Fa2);
mo=alf(find(Fa==mFa(1)));
mo1=alf(find(Fa1==mFa1(1)));
mo2=alf(find(Fa2==mFa2(1)));
r=(mo-0.6)^2;
r1=(mo1-0.6)^2;
r2=(mo2-0.6)^2;
Mo=[Mo;r]; Mo1=[Mo1;r1]; Mo2=[Mo2;r2];
rr=[r r1 r2];
R=[R;rr];
mr=min(rr);
u=find(rr==mr(1));
U=[U;u(1)];
me=alf*Fa';
me1=alf*Fa1';
me2=alf*Fa2';
d=(me-0.6)^2;
d1=(me1-0.6)^2;
d2=(me2-0.6)^2;
dd=[d;d1;d2];
q=find(dd==min(dd));
Q=[Q;q];
D=[D;d]; D1=[D1;d1]; D2=[D2;d2];
ca=cumsum(Fa)/sum(Fa);
ca1=cumsum(Fa1)/sum(Fa1);
ca2=cumsum(Fa2)/sum(Fa2);
L=[];
for j=0.001:0.001:0.025
a=max(find(ca<=0.025+j));
b=max(find(ca<=0.975+j));
l=alf(b)-alf(a);
L=[L;l];
```

```
end
M=find(L==min(L));
j=0.001:0.001:0.025;
Nlo=max(find(ca<=0.025+j(M(1))));
Nup=max(find(ca<=0.975+j(M(1))));
lo=alf(Nlo);
up=alf(Nup);
len=up-lo;
Len=[Len;len];
if lo<0.6 & up>0.6
count=count+1;
end
L1=[];
for j=0.001:0.001:0.025
a1=max(find(ca1<=0.025+j));
b1=max(find(ca1<=0.975+j));
l1=alf(b1)-alf(a1);
L1=[L1;l1];
end
M1=find(L1==min(L1));
j=0.001:0.001:0.025;
Nlo1=max(find(ca1<=0.025+j(M1(1))));
Nup1=max(find(ca1<=0.975+j(M1(1))));
lo1=alf(Nlo1);
up1=alf(Nup1);
len1=up1-lo1;
Len1=[Len1;len1];
if lo1<0.6 & up1>0.6
count1=count1+1;
end
L2=[];
for j=0.001:0.001:0.025
a2=max(find(ca2<=0.025+j));
b2=max(find(ca2<=0.975+j));
l2=alf(b2)-alf(a2);
L2=[L2;l2];
end
```

```
M2=find(L2==min(L2));
j=0.001:0.001:0.025;
Nlo2=max(find(ca2<=0.025+j(M2(1))));
Nup2=max(find(ca2<=0.975+j(M2(1))));
lo2=alf(Nlo2);
up2=alf(Nup2);
len2=up2-lo2;
Len2=[Len2;len2];
if lo2<0.6 & up2>0.6
count2=count2+1;
end
end
mme=length(find(Q==1));
mme1=length(find(Q==2));
mme2=length(find(Q==3));
mmo=length(find(U==1));
mmo1=length(find(U==2));
mmo2=length(find(U==3));
Mme=[mme mme1 mme2 ]
Mmo=[mmo mmo1 mmo2 ]
toc
```

B.2 MATLAB Code for Example 2

```
% COMPARE Three DISTRIBUTIONS OF THE SAME ALPHA
% EXAMPLE 2
clear
tic
s1=4; s2=1;
s11=3; s22=0.75;
s111=5; s222=1.25;
I=20; J=6;
v1=I*(J-1); v2=I-1;
D=[]; D1=[]; D2=[]; Q=[]; Mo=[]; Mo1=[]; Mo2=[]; U=[]; R=[];
Mo3=[]; D3=[];
count=0; Len=[]; count1=0; Len1=[]; count2=0; count3=0; Len2=[]; Len3=[];
for i=1:10000
```

```
v1m1=s1*chi2rnd(v1);
v2m2=(s1+J*s2)*chi2rnd(v2);
v1m11=s11*chi2rnd(v1);
v2m22=(s11+J*s22)*chi2rnd(v2);
v1m111=s111*chi2rnd(v1);
v2m222=(s111+J*s222)*chi2rnd(v2);
alf=0:0.001:1;
A=(v2m2*(1-alf)+v1m1).^(-(v1+v2)/2);
B=(v2m22*(1-alf)+v1m11).^(-(v1+v2)/2);
C=(v2m222*(1-alf)+v1m111).^(-(v1+v2)/2);
pa=(10^300)*((1-alf).^(1.5*v2-1)).*A.*B.*C;
K=1/sum(pa)/0.001;
pa1=(10^300)*((1-alf).^((v2-1)/2)).*A;
pa2=(10^300)*((1-alf).^((v2-1)/2)).*B;
pa3=(10^300)*((1-alf).^((v2-1)/2)).*C;
K1=1/sum(pa1)/0.001;
K2=1/sum(pa2)/0.001;
K3=1/sum(pa3)/0.001;
figure(1)
plot(alf,K*pa,alf,K1*pa1,alf,K2*pa2,alf,K3*pa3)
grid
stop
Fa=pa*K*0.001;
Fa1=pa1*K1*0.001;
Fa2=pa2*K2*0.001;
Fa3=pa3*K3*0.001;
mFa=max(Fa); mFa1=max(Fa1); mFa2=max(Fa2); mFa3=max(Fa3);
mo=alf(find(Fa==mFa(1)));
mo1=alf(find(Fa1==mFa1(1)));
mo2=alf(find(Fa2==mFa2(1)));
mo3=alf(find(Fa3==mFa3(1)));
r=(mo-0.6)^2;
r1=(mo1-0.6)^2;
r2=(mo2-0.6)^2;
r3=(mo3-0.6)^2;
Mo=[Mo;r]; Mo1=[Mo1;r1]; Mo2=[Mo2;r2]; Mo3=[Mo3;r3];
rr=[r r1 r2 r3];
```

```
R=[R;rr];
mr=min(rr);
u=find(rr==mr(1));
U=[U;u(1)];
me=alf*Fa';
me1=alf*Fa1';
me2=alf*Fa2';
me3=alf*Fa3';
d=(me-0.6)^2;
d1=(me1-0.6)^2;
d2=(me2-0.6)^2;
d3=(me3-0.6)^2;
dd=[d;d1;d2;d3];
q=find(dd==min(dd));
Q=[Q;q];
D=[D;d]; D1=[D1;d1]; D2=[D2;d2]; D3=[D3;d3];
ca=cumsum(Fa)/sum(Fa);
ca1=cumsum(Fa1)/sum(Fa1);
ca2=cumsum(Fa2)/sum(Fa2);
ca3=cumsum(Fa3)/sum(Fa3);
L=[];
for j=0.001:0.001:0.025
a=max(find(ca<=0.025+j));
b=max(find(ca<=0.975+j));
l=alf(b)-alf(a);
L=[L;l];
end
M=find(L==min(L));
j=0.001:0.001:0.025;
Nlo=max(find(ca<=0.025+j(M(1))));
Nup=max(find(ca<=0.975+j(M(1))));
lo=alf(Nlo);
up=alf(Nup);
len=up-lo;
Len=[Len;len];
if lo<0.6 & up>0.6
count=count+1;
```

```
end
L1=[];
for j=0.001:0.001:0.025
a1=max(find(ca1<=0.025+j));
b1=max(find(ca1<=0.975+j));
l1=alf(b1)-alf(a1);
L1=[L1;l1];
end
M1=find(L1==min(L1));
j=0.001:0.001:0.025;
Nlo1=max(find(ca1<=0.025+j(M1(1))));
Nup1=max(find(ca1<=0.975+j(M1(1))));
lo1=alf(Nlo1);
up1=alf(Nup1);
len1=up1-lo1;
Len1=[Len1;len1];
if lo1<0.6 & up1>0.6
count1=count1+1;
end
L2=[];
for j=0.001:0.001:0.025
a2=max(find(ca2<=0.025+j));
b2=max(find(ca2<=0.975+j));
l2=alf(b2)-alf(a2);
L2=[L2;l2];
end
M2=find(L2==min(L2));
j=0.001:0.001:0.025;
Nlo2=max(find(ca2<=0.025+j(M2(1))));
Nup2=max(find(ca2<=0.975+j(M2(1))));
lo2=alf(Nlo2);
up2=alf(Nup2);
len2=up2-lo2;
Len2=[Len2;len2];
if lo2<0.6 & up2>0.6
count2=count2+1;
end
```

```
L3=[];
for j=0.001:0.001:0.025
a3=max(find(ca3<=0.025+j));
b3=max(find(ca3<=0.975+j));
l3=alf(b3)-alf(a3);
L3=[L3;l3];
end
M3=find(L3==min(L3));
j=0.001:0.001:0.025;
Nlo3=max(find(ca3<=0.025+j(M3(1))));
Nup3=max(find(ca3<=0.975+j(M3(1))));
lo3=alf(Nlo3);
up3=alf(Nup3);
len3=up3-lo3;
Len3=[Len3;len3];
if lo2<0.6 & up2>0.6
count3=count3+1;
end
end
mme=length(find(Q==1));
mme1=length(find(Q==2));
mme2=length(find(Q==3));
mme3=length(find(Q==4));
mmo=length(find(U==1));
mmo1=length(find(U==2));
mmo2=length(find(U==3));
mmo3=length(find(U==4));
Mme=[mme mme1 mme2 mme3]
Mmo=[mmo mmo1 mmo2 mmo3]
toc
```

B.3 MATHEMATICA Code for Figure 4.1

```
i = 20
j = 6
m = 2
v1 = i*(j - 1)
```

```
v2 = i - 1
sig1 = 4
sig2 = 1
sig11 = 3
sig22 = 0.75
v1m11 = 414.0005
v2m21 = 297.2571
v1m12 = 328.4378
v2m22 = 126.7261
f[x_] := (1 - x)^((m/2 (i - 1)) -
1) {v2m21 (1 - x) + v1m11}^(-0.5 (i*j - 1))*{v2m22 (1 - x) +
v1m12}^(-0.5 (i*j - 1));
c = NIntegrate[f[x], {x, 0, 1}, WorkingPrecision -> 40,
PrecisionGoal -> 6, MaxRecursion -> 20]
g[x_] := 1/
c (1 - x)^((m/2 (i - 1)) -
1) {v2m21 (1 - x) + v1m11}^(-0.5 (i*j - 1))*{v2m22 (1 - x) +
v1m12}^(-0.5 (i*j - 1));
NIntegrate[g[x], {x, 0, 1}]
h[x_] :=
1/c x (1 - x)^((m/2 (i - 1)) -
1) {v2m21 (1 - x) + v1m11}^(-0.5 (i*j - 1))*{v2m22 (1 - x) +
v1m12}^(-0.5 (i*j - 1));
meanpost = NIntegrate[h[x], {x, 0, 1}]
f1[t_] := (1 - t)^(
1/2*(v2 - 2)) {v2m21 (1 - t) + v1m11}^(-0.5 (v1 + v2));
k1 = Integrate[f1[t], {t, 0, 1}]
g2[x_] := 1/k1 (1 - x)^(
1/2*(v2 - 2)) {v2m21 (1 - x) + v1m11}^(-0.5 (v1 + v2))
Integrate[g2[x], {x, 0, 1}]
f2[t_] := (1 - t)^(
1/2*(v2 - 2)) {v2m22 (1 - t) + v1m12}^(-0.5 (v1 + v2));
k2 = Integrate[f2[t], {t, 0, 1}]
g3[x_] := 1/k2 (1 - x)^(
1/2*(v2 - 2)) {v2m22 (1 - x) + v1m12}^(-0.5 (v1 + v2))
Integrate[g3[x], {x, 0, 1}]
p1 = Plot[g[x], {x, 0, 1},
```

```
PlotLabel ->
Style["Posterior Distributions of Cronbach's Alpha",
FontSize -> 40], PlotStyle -> {Red, Thickness[0.01]},
Frame -> True, FrameStyle -> Directive[Black, 40],
FrameLabel -> {\[Alpha], "Posterior Distributions"},
PlotLegends -> Placed[LineLegend[{R = "Combined"}], {0.2, 0.85}]];
p2 = Plot[g2[x], {x, 0, 1},
PlotStyle -> { Gray, Thickness[0.01], Dashing[Large]},
Frame -> True, FrameLabel -> {\[Alpha]},
PlotLegends -> Placed[LineLegend[{R = "1st Sample"}], {0.2, 0.79}]];
p3 = Plot[g3[x], {x, 0, 1},
PlotStyle -> {Blue, Thickness[0.01], Dashing[Tiny]}, Frame -> True,
FrameLabel -> {\[Alpha]},
PlotLegends -> Placed[LineLegend[{R = "2nd Sample"}], {0.2, 0.73}]];
Show[p1, p2, p3]
```

B.4 MATHEMATICA Code for Figure 4.2

```
i = 20
j = 6
m = 3
v1 = i*(j - 1)
v2 = i - 1
sig1 = 4
sig2 = 1
sig11 = 3
sig22 = 0.75
sig111 = 5
sig222 = 1.25
v1m11 = 349.9693
v2m21 = 151.5623
v1m12 = 274.5413
v2m22 = 100.2209
v1m13 = 386.3027
v2m23 = 220.3655
f[x_] := (1 - x)^((m/2 (i - 1)) -
1) {v2m21 (1 - x) + v1m11}^(-0.5 (i*j - 1))*{v2m22 (1 - x) +
```

```
v1m12}^(-0.5 (i*j - 1))*{v2m23 (1 - x) +
v1m13}^(-0.5 (i*j - 1));
c = NIntegrate[f[x], {x, 0, 1}, WorkingPrecision -> 40,
PrecisionGoal -> 6, MaxRecursion -> 20]
g[x_] := 1/
c (1 - x)^((m/2 (i - 1)) -
1) {v2m21 (1 - x) + v1m11}^(-0.5 (i*j - 1))*{v2m22 (1 - x) +
v1m12}^(-0.5 (i*j - 1))*{v2m23 (1 - x) +
v1m13}^(-0.5 (i*j - 1));
NIntegrate[g[x], {x, 0, 1}]
h[x_] :=
1/c x (1 - x)^((m/2 (i - 1)) -
1) {v2m21 (1 - x) + v1m11}^(-0.5 (i*j - 1))*{v2m22 (1 - x) +
v1m12}^(-0.5 (i*j - 1))*{v2m23 (1 - x) +
v1m13}^(-0.5 (i*j - 1));
meanpost = NIntegrate[h[x], {x, 0, 1}]
f1[t_] := (1 - t)^(
1/2*(v2 - 2)) {v2m21 (1 - t) + v1m11}^(-0.5 (v1 + v2));
k1 = Integrate[f1[t], {t, 0, 1}]
g2[x_] := 1/k1 (1 - x)^(
1/2*(v2 - 2)) {v2m21 (1 - x) + v1m11}^(-0.5 (v1 + v2))
Integrate[g2[x], {x, 0, 1}]
f2[t_] := (1 - t)^(
1/2*(v2 - 2)) {v2m22 (1 - t) + v1m12}^(-0.5 (v1 + v2));
k2 = Integrate[f2[t], {t, 0, 1}]
g3[x_] := 1/k2 (1 - x)^(
1/2*(v2 - 2)) {v2m22 (1 - x) + v1m12}^(-0.5 (v1 + v2))
Integrate[g3[x], {x, 0, 1}]
f3[t_] := (1 - t)^(
1/2*(v2 - 2)) {v2m23 (1 - t) + v1m13}^(-0.5 (v1 + v2));
k3 = Integrate[f3[t], {t, 0, 1}]
g4[x_] := 1/k3 (1 - x)^(
1/2*(v2 - 2)) {v2m23 (1 - x) + v1m13}^(-0.5 (v1 + v2))
Integrate[g3[x], {x, 0, 1}]
p1 = Plot[g[x], {x, 0, 1},
PlotLabel ->
Style["Posterior Distributions of Cronbach's Alpha",
```

```
FontSize -> 40], PlotStyle -> {Red, Thickness[0.01]},
Frame -> True, FrameStyle -> Directive[Black, 40],
FrameLabel -> {\[Alpha], "Posterior Distributions"},
PlotLegends -> Placed[LineLegend[{R = "Combined"}], {0.2, 0.85}]];
p2 = Plot[g2[x], {x, 0, 1},
PlotStyle -> {Gray, Thickness[0.01], Dashing[Large]},
Frame -> True, FrameLabel -> {\[Alpha]},
PlotLegends -> Placed[LineLegend[{R = "1st Sample"}], {0.2, 0.79}]];
p3 = Plot[g3[x], {x, 0, 1},
PlotStyle -> {Blue, Thickness[0.01], Dashing[Tiny]}, Frame -> True,
FrameLabel -> {\[Alpha]},
PlotLegends -> Placed[LineLegend[{R = "2nd Sample"}], {0.2, 0.73}]];
p4 = Plot[g4[x], {x, 0, 1},
PlotStyle -> {Black, Thickness[0.01], Dashing[Medium]},
Frame -> True, FrameLabel -> {\[Alpha]},
PlotLegends -> Placed[LineLegend[{R = "3rd Sample"}], {0.2, 0.66}]];
Show[p1, p2, p3, p4]
```

Appendix C: Cronbach's Alpha for the Three-Component Model

C.1 Fisher Information Matrix for Cronbach's Alpha the Three Component Model

The likelihood function is given by

$$\ell(\theta, \sigma_1^2, \sigma_2^2, \alpha|\text{data}) \propto \left(\sigma_1^2\right)^{-\frac{1}{2}(v_1+v_3+1)} \left(\sigma_1^2 + K\sigma_2^2\right)^{-\frac{v_2}{2}} (1-\alpha)^{\frac{1}{2}(v_3+1)} \times$$
$$\exp\left\{-\frac{1}{2}\left[\frac{IJK\left(\bar{Y}_{...}-\theta\right)^2(1-\alpha)}{\sigma_1^2} + \frac{v_3 m_3 (1-\alpha)}{\sigma_1^2} + \frac{v_2 m_2}{\sigma_1^2 + K\sigma_2^2} + \frac{v_1 m_1}{\sigma_1^2}\right]\right\}.$$

To obtain the Fisher information matrix we take the logarithm of the likelihood function and differentiate twice with respect to the unknown parameters.

$$\begin{aligned}
L &= log\left[\ell(\theta, \sigma_1^2, \alpha|data)\right] \\
&\propto -\frac{1}{2}(v_1+v_3+1)\, log(\sigma_1^2) - \frac{v_2}{2} log\left(\sigma_1^2 + K\sigma_2^2\right) + \frac{(v_3+1)}{2} log(1-\alpha) - \frac{1}{2\sigma_1^2}\left[IJK\left(\bar{Y}_{...}-\theta\right)^2(1-\alpha)\right] \\
&- \frac{1}{2\sigma_1^2}\left[v_3 m_3 (1-\alpha) + v_1 m_1\right] - \frac{v_2 m_2}{2\left(\sigma_1^2 + K\sigma_2^2\right)} \\
\frac{\partial L}{\partial \sigma_1^2} &= -\frac{(v_1+v_3+1)}{2(\sigma_1^2)} - \frac{v_2}{2\left(\sigma_1^2 + K\sigma_2^2\right)} + \frac{1}{2(\sigma_1^2)^2}\left[IJK\left(\bar{Y}_{...}-\theta\right)^2(1-\alpha) + v_3 m_3 (1-\alpha) + v_1 m_1\right] + \frac{v_2 m_2}{2\left(\sigma_1^2 + K\sigma_2^2\right)^2} \\
\frac{\partial L}{\partial \alpha} &= \frac{-(v_3+1)}{2(1-\alpha)} + \frac{1}{2\sigma_1^2}\left[IJK\left(\bar{Y}_{...}-\theta\right)^2 + v_3 m_3\right] \\
\frac{\partial L}{\partial \theta} &= \frac{1}{\sigma_1^2}\left[IJK\left(\bar{Y}_{...}-\theta\right)(1-\alpha)\right] \\
\frac{\partial L}{\partial \sigma_2^2} &= \frac{-v_2 K}{2\left(\sigma_1^2 + K\sigma_2^2\right)} + \frac{K\left(v_2 m_2\right)}{2\left(\sigma_1^2 + K\sigma_2^2\right)^2}
\end{aligned}$$

$$\begin{aligned}
\frac{\partial^2 L}{\left(\partial \sigma_1^2\right)^2} &= \frac{(v_1+v_3+1)}{2(\sigma_1^2)^2} + \frac{v_2}{2\left(\sigma_1^2+K\sigma_2^2\right)^2} - \frac{1}{(\sigma_1^2)^3}\left[IJK\,(\bar{Y}_{...}-\theta)^2(1-\alpha)+v_3m_3(1-\alpha)+v_1m_1\right] - \frac{v_2m_2}{\left(\sigma_1^2+K\sigma_2^2\right)^3} \\
\frac{\partial^2 L}{(\partial \alpha)^2} &= \frac{-(v_3+1)}{2(1-\alpha)^2} \\
\frac{\partial^2 L}{(\partial \theta)^2} &= -\frac{IJK\,(1-\alpha)}{\sigma_1^2} \\
\frac{\partial^2 L}{\left(\partial \sigma_2^2\right)^2} &= \frac{v_2K^2}{2\left(\sigma_1^2+K\sigma_2^2\right)^2} - \frac{K^2\,(v_2m_2)}{\left(\sigma_1^2+K\sigma_2^2\right)^3} \\
\frac{\partial^2 L}{\partial \sigma_1^2 \partial \theta} &= -\frac{IJK\,(\bar{Y}_{...}-\theta)\,(1-\alpha)}{(\sigma_1^2)^2} = \frac{\partial^2 L}{\partial \theta \partial \sigma_1^2} \\
\frac{\partial^2 L}{\partial \sigma_1^2 \partial \alpha} &= -\frac{1}{2(\sigma_1^2)^2}\left[IJK\,(\bar{Y}_{...}-\theta)^2+v_3m_3\right] = \frac{\partial^2 L}{\partial \alpha \partial \sigma_1^2} \\
\frac{\partial^2 L}{\partial \sigma_1^2 \partial \sigma_2^2} &= \frac{v_2K}{2\left(\sigma_1^2+K\sigma_2^2\right)^2} - \frac{(v_2m_2)\,K}{\left(\sigma_1^2+K\sigma_2^2\right)^3} = \frac{\partial^2 L}{\partial \sigma_2^2 \partial \sigma_1^2} \\
\frac{\partial^2 L}{\partial \sigma_2^2 \partial \theta} &= 0 = \frac{\partial^2 L}{\partial \theta \partial \sigma_2^2} \\
\frac{\partial^2 L}{\partial \sigma_2^2 \partial \alpha} &= 0 = \frac{\partial^2 L}{\partial \alpha \partial \sigma_2^2} \\
\frac{\partial^2 L}{\partial \alpha \partial \theta} &= \frac{-IJK\,(\bar{Y}_{..}-\theta)}{\sigma_1^2} = \frac{\partial^2 L}{\partial \theta \partial \alpha}
\end{aligned}$$

Therefore,

$$\begin{aligned}
& -E\left[\frac{\partial^2 L}{\left(\partial \sigma_1^2\right)^2}\right] \\
&= -E\left[\frac{(v_1+v_3+1)}{2(\sigma_1^2)^2} + \frac{v_2}{2\left(\sigma_1^2+K\sigma_2^2\right)^2} - \frac{\left[IJK\,(\bar{Y}_{...}-\theta)^2(1-\alpha)+v_3m_3(1-\alpha)+v_1m_1\right]}{(\sigma_1^2)^3} - \frac{v_2m_2}{\left(\sigma_1^2+K\sigma_2^2\right)^3}\right] \\
&= \frac{-(v_1+v_3+1)}{2(\sigma_1^2)^2} - \frac{v_2}{2\left(\sigma_1^2+K\sigma_2^2\right)^2} + \frac{1}{(\sigma_1^2)^3}\left[E\left(IJK\,(\bar{Y}_{...}-\theta)^2\right)(1-\alpha)+v_3E\,(m_3)\,(1-\alpha)+v_1E\,(m_1)\right] \\
&+ \frac{v_2E\,(m_2)}{\left(\sigma_1^2+K\sigma_2^2\right)^3} \\
&= \frac{-(v_1+v_3+1)}{2(\sigma_1^2)^2} - \frac{v_2}{2\left(\sigma_1^2+K\sigma_2^2\right)^2} + \frac{1}{(\sigma_1^2)^3}\left[\frac{\sigma_1^2}{(1-\alpha)}(1-\alpha)+v_3\left(\frac{\sigma_1^2}{(1-\alpha)}\right)(1-\alpha)+v_1\sigma_1^2\right] + \frac{v_2\left(\sigma_1^2+K\sigma_2^2\right)}{\left(\sigma_1^2+K\sigma_2^2\right)^3} \\
&= \frac{-(v_1+v_3+1)}{2(\sigma_1^2)^2} + \frac{v_2}{2\left(\sigma_1^2+K\sigma_2^2\right)^2} + \frac{1}{(\sigma_1^2)^2}\left[1+v_3+v_1\right] = \frac{(v_1+v_3+1)}{2(\sigma_1^2)^2} + \frac{v_2}{2\left(\sigma_1^2+K\sigma_2^2\right)^2}.
\end{aligned}$$

$$
\begin{aligned}
&-E\left[\frac{\partial^2 L}{\left(\partial\sigma_2^2\right)^2}\right]\\
=&-E\left[\frac{v_2K^2}{2\left(\sigma_1^2+K\sigma_2^2\right)^2}-\frac{K^2\left(v_2m_2\right)}{\left(\sigma_1^2+K\sigma_2^2\right)^3}\right]\\
=&\frac{-v_2K^2}{2\left(\sigma_1^2+K\sigma_2^2\right)^2}+\frac{v_2K^2E\left(m_2\right)}{\left(\sigma_1^2+K\sigma_2^2\right)^3}\\
=&\frac{v_2K^2}{2\left(\sigma_1^2+K\sigma_2^2\right)^2}.
\end{aligned}
$$

Also,

$$
-E\left[\frac{\partial^2 L}{(\partial\alpha)^2}\right]=-E\left[\frac{-(v_3+1)}{2(1-\alpha)^2}\right]=\frac{(v_3+1)}{2(1-\alpha)^2}
$$

and

$$
-E\left[\frac{\partial^2 L}{(\partial\theta)^2}\right]=-E\left[-\frac{IJK\left(1-\alpha\right)}{\sigma_1^2}\right]=\frac{IJK\left(1-\alpha\right)}{\sigma_1^2}.
$$

$$
\begin{aligned}
&-E\left[\frac{\partial^2 L}{\partial\sigma_1^2\partial\alpha}\right]\\
=&-E\left[-\frac{1}{2(\sigma_1^2)^2}\left[IJK\left(\bar{Y}_{...}-\theta\right)^2+v_3m_3\right]\right]\\
=&\frac{1}{2(\sigma_1^2)^2}\left[E\left(IJK\left(\bar{Y}_{...}-\theta\right)^2\right)+v_3E\left(m_3\right)\right]\\
=&\frac{1}{2(\sigma_1^2)^2}\left[\frac{\sigma_1^2}{(1-\alpha)}+\frac{v_3\sigma_1^2}{(1-\alpha)}\right]\\
=&\frac{(1+v_3)}{2\sigma_1^2\left(1-\alpha\right)}.
\end{aligned}
$$

$$
\begin{aligned}
&-E\left[\frac{\partial^2 L}{\partial\sigma_1^2\partial\sigma_2^2}\right]\\
=&-E\left[\frac{v_2K}{2\left(\sigma_1^2+K\sigma_2^2\right)^2}-\frac{\left(v_2m_2\right)K}{\left(\sigma_1^2+K\sigma_2^2\right)^3}\right]\\
=&\frac{-v_2K}{2\left(\sigma_1^2+K\sigma_2^2\right)^2}+\frac{v_2KE\left(m_2\right)}{\left(\sigma_1^2+K\sigma_2^2\right)^3}\\
=&\frac{v_2K}{2\left(\sigma_1^2+K\sigma_2^2\right)^2}.
\end{aligned}
$$

The following expected values are all zero: $-E\left[\frac{\partial^2 L}{\partial\theta\partial\sigma_1^2}\right] = -E\left[\frac{\partial^2 L}{\partial\sigma_1^2\partial\theta}\right] = 0$, $-E\left[\frac{\partial^2 L}{\partial\theta\partial\sigma_2^2}\right] = -E\left[\frac{\partial^2 L}{\partial\sigma_2^2\partial\theta}\right] = 0$, $-E\left[\frac{\partial^2 L}{\partial\alpha\partial\theta}\right] = -E\left[\frac{\partial^2 L}{\partial\theta\partial\alpha}\right] = 0$ and $-E\left[\frac{\partial^2 L}{\partial\alpha\partial\sigma_2^2}\right] = -E\left[\frac{\partial^2 L}{\partial\sigma_2^2\partial\alpha}\right] = 0$.

The Fisher information matrix is therefore given by

$$F(\theta,\sigma_1^2,\sigma_2^2,\alpha) = \begin{bmatrix} \frac{IJK(1-\alpha)}{\sigma_1^2} & 0 & 0 & 0 \\ 0 & \frac{(v_1+v_3+1)}{2\left(\sigma_1^2\right)^2} + \frac{v_2}{2\left(\sigma_1^2+K\sigma_2^2\right)^2} & \frac{v_2 K}{2\left(\sigma_1^2+K\sigma_2^2\right)^2} & \frac{(v_3+1)}{2(1-\alpha)\sigma_1^2} \\ 0 & \frac{v_2 K}{2\left(\sigma_1^2+K\sigma_2^2\right)^2} & \frac{v_2 K^2}{2\left(\sigma_1^2+K\sigma_2^2\right)^2} & 0 \\ 0 & \frac{(v_3+1)}{2(1-\alpha)\sigma_1^2} & 0 & \frac{(v_3+1)}{2(1-\alpha)^2} \end{bmatrix}.$$

C.2 Proof of the Reference Prior for the Group Ordering $\left\{\alpha,\sigma_1^2,\theta,\sigma_2^2\right\}$

We are interested in the reference prior for the group ordering $\left\{\alpha,\sigma_1^2,\theta,\sigma_2^2\right\}$ which means that α is the most important parameter and σ_2^2 is the least important parameter. In order to derive the reference prior, the inverse of the Fisher information matrix is needed. Since the h functions needed to derive the reference prior has already been determined in Theorem 5.1, it is easy to see what they should be for the group ordering $\left\{\alpha,\sigma_1^2,\theta,\sigma_2^2\right\}$. Define the truncated ranges for the 4 parameters as $\alpha \in [a_l,b_l]$, $\theta \in [c_l,d_l]$, $\sigma_1^2 \in [e_l,f_l]$ and $\sigma_2^2 \in [g_l,h_l]$ where $c_l \to -\infty$, $b_l \to 1$, $d_l,f_l,h_l \to \infty$ and $a_1,e_l,g_l \to 0$. The four functions

$$h_1 \propto \frac{1}{(1-\alpha)^2}, \qquad h_2 \propto \frac{1}{\left(\sigma_1^2\right)^2}, \qquad h_3 \propto \frac{(1-\alpha)}{\sigma_1^2}, \qquad h_4 \propto \frac{1}{\left(K\sigma_2^2+\sigma_1^2\right)^2}$$

are the functions needed to derive the prior. During the iterations, first the truncated conditional function of σ_2^2 given α, σ_1^2 and θ can be computed as

$$\begin{aligned}
\pi_4^l\left(\sigma_2^2|\alpha,\sigma_1^2,\theta\right) &= \frac{\{h_4(\boldsymbol{\theta})\}^{\frac{1}{2}}}{\int_{\theta_4^l}\{h_4(\boldsymbol{\theta})\}^{\frac{1}{2}}\,d\theta_4} \\
&\propto \frac{\left(K\sigma_2^2+\sigma_1^2\right)^{-1}}{\int_{g_l}^{h_l}\left(K\sigma_2^2+\sigma_1^2\right)^{-1}d\sigma_2^2} \\
&= \frac{\left(K\sigma_2^2+\sigma_1^2\right)^{-1}}{K^{-1}log\left(\frac{Kh_l+\sigma_1^2}{Kg_l+\sigma_1^2}\right)} \qquad \text{for } g_l \le \sigma_2^2 \le h_l.
\end{aligned}$$

Now,

$$\begin{aligned}
&E\left[log\{h_3(\boldsymbol{\theta})\}^{\frac{1}{2}}\right] \\
&\propto \int_{g_l}^{h_l} log\left\{\frac{(1-\alpha)}{\sigma_1^2}\right\}^{\frac{1}{2}} \frac{\left(K\sigma_2^2+\sigma_1^2\right)^{-1}}{K^{-1}log\left(\frac{Kh_l+\sigma_1^2}{Kg_l+\sigma_1^2}\right)} d\sigma_2^2 \\
&= log\left\{\frac{(1-\alpha)}{\sigma_1^2}\right\}^{\frac{1}{2}}.
\end{aligned}$$

Therefore,

$$\pi_3^l\left(\theta,\sigma_2^2|\alpha,\sigma_1^2\right) = \frac{\pi_4^l\left(\sigma_2^2|\alpha,\sigma_1^2,\theta\right) exp\left\{E\left[log\left\{h_3\left(\boldsymbol{\theta}\right)\right\}^{\frac{1}{2}}\right]\right\}}{\int_{c_l}^{d_l} exp\left\{E\left[log\left\{h_3\left(\boldsymbol{\theta}\right)\right\}^{\frac{1}{2}}\right]\right\}d\theta}$$
$$= \frac{\frac{\left(K\sigma_2^2+\sigma_1^2\right)^{-1}}{K^{-1}log\left(\frac{Kh_l+\sigma_1^2}{Kg_l+\sigma_1^2}\right)} exp\left\{log\left\{\frac{(1-\alpha)}{\sigma_1^2}\right\}^{\frac{1}{2}}\right\}}{\int_{c_l}^{d_l} exp\left\{log\left\{\frac{(1-\alpha)}{\sigma_1^2}\right\}^{\frac{1}{2}}\right\}d\theta}$$
$$= \frac{\left(K\sigma_2^2+\sigma_1^2\right)^{-1}}{K^{-1}log\left(\frac{Kh_l+\sigma_1^2}{Kg_l+\sigma_1^2}\right)} \times \left(d_l-c_l\right)^{-1} \qquad \text{for } g_l \le \sigma_2^2 \le h_l \text{ and } c_l \le \theta \le d_l.$$

We now need the function, h_2, to determine $E\left[log\left\{h_2\left(\boldsymbol{\theta}\right)\right\}^{\frac{1}{2}}\right]$

$$E\left[log\left\{h_2\left(\boldsymbol{\theta}\right)\right\}^{\frac{1}{2}}\right]$$
$$\propto \int_{g_l}^{h_l}\int_{c_l}^{d_l} -log\left(\sigma_1^2\right)\frac{\left(K\sigma_2^2+\sigma_1^2\right)^{-1}}{K^{-1}log\left(\frac{Kh_l+\sigma_1^2}{Kg_l+\sigma_1^2}\right)} \times \left(d_l-c_l\right)^{-1} d\theta d\sigma_2^2$$
$$= \int_{g_l}^{h_l} -log\left(\sigma_1^2\right)\frac{\left(K\sigma_2^2+\sigma_1^2\right)^{-1}}{K^{-1}log\left(\frac{Kh_l+\sigma_1^2}{Kg_l+\sigma_1^2}\right)} d\sigma_2^2$$
$$= log\left(\sigma_1^2\right)^{-1}.$$

Therefore,

$$\pi_2^l\left(\sigma_1^2,\theta,\sigma_2^2|\alpha\right) = \frac{\pi_3^l\left(\theta,\sigma_2^2|\alpha,\sigma_1^2\right) exp\left\{E\left[log\left\{h_2\left(\boldsymbol{\theta}\right)\right\}^{\frac{1}{2}}\right]\right\}}{\int_{e_l}^{f_l} exp\left\{E\left[log\left\{h_2\left(\boldsymbol{\theta}\right)\right\}^{\frac{1}{2}}\right]\right\}d\sigma_1^2}$$
$$= \frac{\left(K\sigma_2^2+\sigma_1^2\right)^{-1}\left(d_l-c_l\right)^{-1}\left(\sigma_1^2\right)^{-1}}{K^{-1}log\left(\frac{Kh_l+\sigma_1^2}{Kg_l+\sigma_1^2}\right)} \times \frac{1}{\int_{e_l}^{f_l}\left(\sigma_1^2\right)^{-1} d\sigma_1^2}$$
$$= \frac{\left(K\sigma_2^2+\sigma_1^2\right)^{-1}\left(d_l-c_l\right)^{-1}\left(\sigma_1^2\right)^{-1}}{K^{-1}log\left(\frac{Kh_l+\sigma_1^2}{Kg_l+\sigma_1^2}\right) log\left(f_l e_l^{-1}\right)}.$$

We now need the function, h_1, to determine $E\left[log\left\{h_1\left(\boldsymbol{\theta}\right)\right\}^{\frac{1}{2}}\right]$ in order to proceed with the iterative process to determine $\pi_1^l\left(\alpha,\sigma_1^2,\theta,\sigma_2^2\right)$. We have

$$\begin{aligned}
&E\left[log\left\{h_1\left(\boldsymbol{\theta}\right)\right\}^{\frac{1}{2}}\right]\\
\propto&\int_{e_l}^{f_l}\int_{g_l}^{h_l}\int_{c_l}^{d_l} log\left(1-\alpha\right)^{-1}\frac{\left(K\sigma_2^2+\sigma_1^2\right)^{-1}\left(d_l-c_l\right)^{-1}\left(\sigma_1^2\right)^{-1}}{K^{-1}log\left(\frac{Kh_l+\sigma_1^2}{Kg_l+\sigma_1^2}\right)log\left(f_le_l^{-1}\right)}d\theta d\sigma_2^2 d\sigma_1^2\\
=&\int_{e_l}^{f_l}\int_{g_l}^{h_l} log\left(1-\alpha\right)^{-1}\frac{\left(K\sigma_2^2+\sigma_1^2\right)^{-1}\left(\sigma_1^2\right)^{-1}}{K^{-1}log\left(\frac{Kh_l+\sigma_1^2}{Kg_l+\sigma_1^2}\right)log\left(f_le_l^{-1}\right)}d\sigma_2^2 d\sigma_1^2\\
=&log\left(1-\alpha\right)^{-1}
\end{aligned}$$

Therefore,

$$\begin{aligned}
\pi_1^l\left(\alpha,\sigma_1^2,\theta,\sigma_2^2\right)=&\frac{\pi_2^l\left(\sigma_1^2,\theta,\sigma_2^2|\alpha\right)exp\left\{E\left[log\left\{h_1\left(\boldsymbol{\theta}\right)\right\}^{\frac{1}{2}}\right]\right\}}{\int_{a_l}^{b_l}exp\left\{E\left[log\left\{h_1\left(\boldsymbol{\theta}\right)\right\}^{\frac{1}{2}}\right]\right\}d\alpha}\\
=&\frac{\left(K\sigma_2^2+\sigma_1^2\right)^{-1}\left(d_l-c_l\right)^{-1}\left(\sigma_1^2\right)^{-1}\left(1-\alpha\right)^{-1}}{K^{-1}log\left(\frac{Kh_l+\sigma_1^2}{Kg_l+\sigma_1^2}\right)log\left(f_le_l^{-1}\right)}\frac{1}{\int_{a_l}^{b_l}\left(1-\alpha\right)^{-1}d\alpha}\\
=&\frac{\left(K\sigma_2^2+\sigma_1^2\right)^{-1}\left(d_l-c_l\right)^{-1}\left(\sigma_1^2\right)^{-1}\left(1-\alpha\right)^{-1}}{K^{-1}log\left(\frac{Kh_l+\sigma_1^2}{Kg_l+\sigma_1^2}\right)log\left(f_le_l^{-1}\right)}\times\frac{1}{log\left(\frac{1-a_l}{1-b_l}\right)}.
\end{aligned}$$

Finally,

$$\pi_R\left(\alpha,\sigma_1^2,\theta,\sigma_2^2\right)\propto\lim_{l\to\infty}\frac{\pi_1^l\left(\alpha,\sigma_1^2,\theta,\sigma_2^2\right)}{\pi_1^l\left(\alpha_0,\sigma_{10}^2,\theta_0,\sigma_{20}^2\right)}\propto\left(K\sigma_2^2+\sigma_1^2\right)^{-1}\left(\sigma_1^2\right)^{-1}\left(1-\alpha\right)^{-1}$$

where $\alpha_0,\theta_0,\sigma_{10}^2$ and σ_{20}^2 are the four inner points in the ranges of the parameters.

C.3 Proof of the Reference Prior for the Group Ordering $\left\{\alpha,\theta,\sigma_1^2,\sigma_2^2\right\}$

We are interested in the reference prior for the group ordering $\left\{\alpha,\theta,\sigma_1^2,\sigma_2^2\right\}$ which means that α is the most important parameter and σ_2^2 is the least important parameter. In order to derive the reference prior, the inverse of the Fisher information matrix is needed. Since the h functions needed to derive the reference prior has already been determined in Theorem 5.1, it is easy to see what they should be for the group ordering $\left\{\alpha,\sigma_1^2,\theta,\sigma_2^2\right\}$. Define the truncated ranges for the 4 parameters as $\alpha\in[a_l,b_l]$, $\theta\in[c_l,d_l]$, $\sigma_1^2\in[e_l,f_l]$ and $\sigma_2^2\in[g_l,h_l]$ where $c_l\to-\infty$, $b_l\to 1$, $d_l,f_l,h_l\to\infty$ and $a_1,e_l,g_l\to 0$. The four functions

$$h_1\propto\frac{1}{(1-\alpha)^2},\qquad h_2\propto\frac{(1-\alpha)}{\sigma_1^2},\qquad h_3\propto\frac{1}{\left(\sigma_1^2\right)^2},\qquad h_4\propto\frac{1}{\left(K\sigma_2^2+\sigma_1^2\right)^2}$$

are the functions needed to derive the prior. During the iterations, first the truncated conditional function of σ_2^2 given α, θ and σ_1^2 can be computed as

$$\begin{aligned}\pi_4^l\left(\sigma_2^2|\alpha,\theta,\sigma_1^2\right)&=\frac{\left\{h_4\left(\boldsymbol{\theta}\right)\right\}^{\frac{1}{2}}}{\int_{\Theta_4^l}\left\{h_4\left(\boldsymbol{\theta}\right)\right\}^{\frac{1}{2}}d\theta_4}\\&\propto\frac{\left(K\sigma_2^2+\sigma_1^2\right)^{-1}}{\int_{g_l}^{h_l}\left(K\sigma_2^2+\sigma_1^2\right)^{-1}d\sigma_2^2}\\&=\frac{\left(K\sigma_2^2+\sigma_1^2\right)^{-1}}{K^{-1}log\left(\frac{Kh_l+\sigma_1^2}{Kg_l+\sigma_1^2}\right)}\qquad\text{for } g_l\le\sigma_2^2\le h_l.\end{aligned}$$

Now,

$$\begin{aligned}&E\left[log\left\{h_3\left(\boldsymbol{\theta}\right)\right\}^{\frac{1}{2}}\right]\\&\propto\int_{g_l}^{h_l}log\left(\sigma_1^2\right)^{-1}\frac{\left(K\sigma_2^2+\sigma_1^2\right)^{-1}}{K^{-1}log\left(\frac{Kh_l+\sigma_1^2}{Kg_l+\sigma_1^2}\right)}d\sigma_2^2\\&=log\left(\sigma_1^2\right)^{-1}\end{aligned}$$

Therefore,

$$
\begin{aligned}
\pi_3^l\left(\sigma_1^2,\sigma_2^2|\alpha,\theta\right) =& \frac{\pi_4^l\left(\sigma_2^2|\alpha,\theta,\sigma_1^2\right) exp\left\{E\left[log\left\{h_3\left(\boldsymbol{\theta}\right)\right\}^{\frac{1}{2}}\right]\right\}}{\int_{c_l}^{d_l} exp\left\{E\left[log\left\{h_3\left(\boldsymbol{\theta}\right)\right\}^{\frac{1}{2}}\right]\right\} d\theta} \\
=& \frac{\left(K\sigma_2^2+\sigma_1^2\right)^{-1}\left(\sigma_1^2\right)^{-1}}{K^{-1}log\left(\frac{Kh_l+\sigma_1^2}{Kg_l+\sigma_1^2}\right)} \times \frac{1}{\int_{e_l}^{f_l}\left(\sigma_1^2\right)^{-1} d\sigma_1^2} \\
=& \frac{\left(K\sigma_2^2+\sigma_1^2\right)^{-1}\left(\sigma_1^2\right)^{-1}}{K^{-1}log\left(\frac{Kh_l+\sigma_1^2}{Kg_l+\sigma_1^2}\right)} \times \left[log\left(f_l e_l^{-1}\right)\right]^{-1} \qquad \text{for } e_l \le \sigma_1^2 \le f_l \text{ and } g_l \le \sigma_2^2 \le h_l.
\end{aligned}
$$

We now need the function, h_2, to determine $E\left[log\left\{h_2\left(\boldsymbol{\theta}\right)\right\}^{\frac{1}{2}}\right]$

$$
\begin{aligned}
& E\left[log\left\{h_2\left(\boldsymbol{\theta}\right)\right\}^{\frac{1}{2}}\right] \\
\propto & \int_{e_l}^{f_l}\int_{g_l}^{h_l} log\left\{\frac{(1-\alpha)}{\sigma_1^2}\right\}^{\frac{1}{2}} \frac{\left(K\sigma_2^2+\sigma_1^2\right)^{-1}\left(\sigma_1^2\right)^{-1}}{K^{-1}log\left(\frac{Kh_l+\sigma_1^2}{Kg_l+\sigma_1^2}\right)} \times \left[log\left(f_l e_l^{-1}\right)\right]^{-1} d\sigma_2^2 d\sigma_1^2 \\
=& \int_{e_l}^{f_l} \frac{1}{2}\left[log\left(1-\alpha\right) - log\left(\sigma_1^2\right)\right]\left(\sigma_1^2\right)^{-1}\left[log\left(f_l e_l^{-1}\right)\right]^{-1} d\sigma_1^2 \\
=& \frac{log\left(1-\alpha\right)}{2} - \frac{1}{4log\left(f_l e_l^{-1}\right)}\left[log^2\left(f_l\right) - log^2\left(e_l\right)\right] \\
=& log(1-\alpha)^{\frac{1}{2}} + K^*
\end{aligned}
$$

where $K^* = -\frac{1}{4log\left(f_l e_l^{-1}\right)}\left[log^2\left(f_l\right) - log^2\left(e_l\right)\right]$ is denoted as a constant which only relates to the ranges of the parameters. Therefore

$$
\begin{aligned}
\pi_2^l\left(\theta,\sigma_1^2,\sigma_2^2|\alpha\right) =& \frac{\pi_3^l\left(\sigma_1^2,\sigma_2^2|\alpha,\theta\right) exp\left\{E\left[log\left\{h_2\left(\boldsymbol{\theta}\right)\right\}^{\frac{1}{2}}\right]\right\}}{\int_{c_l}^{d_l} exp\left\{E\left[log\left\{h_2\left(\boldsymbol{\theta}\right)\right\}^{\frac{1}{2}}\right]\right\} d\theta} \\
=& \frac{\left(K\sigma_2^2+\sigma_1^2\right)^{-1}\left(\sigma_1^2\right)^{-1}}{K^{-1}log\left(\frac{Kh_l+\sigma_1^2}{Kg_l+\sigma_1^2}\right)} \times \left[log\left(f_l e_l^{-1}\right)\right]^{-1} \times \frac{(1-\alpha)^{\frac{1}{2}}\exp\left\{K^*\right\}}{\int_{c_l}^{d_l}(1-\alpha)^{\frac{1}{2}}\exp\left\{K^*\right\} d\theta} \\
=& \frac{\left(K\sigma_2^2+\sigma_1^2\right)^{-1}\left(d_l-c_l\right)^{-1}\left(\sigma_1^2\right)^{-1}}{K^{-1}log\left(\frac{Kh_l+\sigma_1^2}{Kg_l+\sigma_1^2}\right) log\left(f_l e_l^{-1}\right)}
\end{aligned}
$$

We now need the function, h_1, to determine $E\left[log\left\{h_1\left(\boldsymbol{\theta}\right)\right\}^{\frac{1}{2}}\right]$ in order to proceed with the iterative process to determine $\pi_1^l\left(\alpha,\theta,\sigma_1^2,\sigma_2^2\right)$. We have

$$
\begin{aligned}
&E\left[log\left\{h_1\left(\boldsymbol{\theta}\right)\right\}^{\frac{1}{2}}\right]\\
\propto&\int_{e_l}^{f_l}\int_{g_l}^{h_l}\int_{c_l}^{d_l} log\left(1-\alpha\right)^{-1}\frac{\left(K\sigma_2^2+\sigma_1^2\right)^{-1}\left(d_l-c_l\right)^{-1}\left(\sigma_1^2\right)^{-1}}{K^{-1}log\left(\frac{Kh_l+\sigma_1^2}{Kg_l+\sigma_1^2}\right)log\left(f_le_l^{-1}\right)}d\theta d\sigma_2^2 d\sigma_1^2\\
=&\int_{e_l}^{f_l}\int_{g_l}^{h_l} log\left(1-\alpha\right)^{-1}\frac{\left(K\sigma_2^2+\sigma_1^2\right)^{-1}\left(\sigma_1^2\right)^{-1}}{K^{-1}log\left(\frac{Kh_l+\sigma_1^2}{Kg_l+\sigma_1^2}\right)log\left(f_le_l^{-1}\right)}d\sigma_2^2 d\sigma_1^2\\
=&log\left(1-\alpha\right)^{-1}.
\end{aligned}
$$

Therefore,

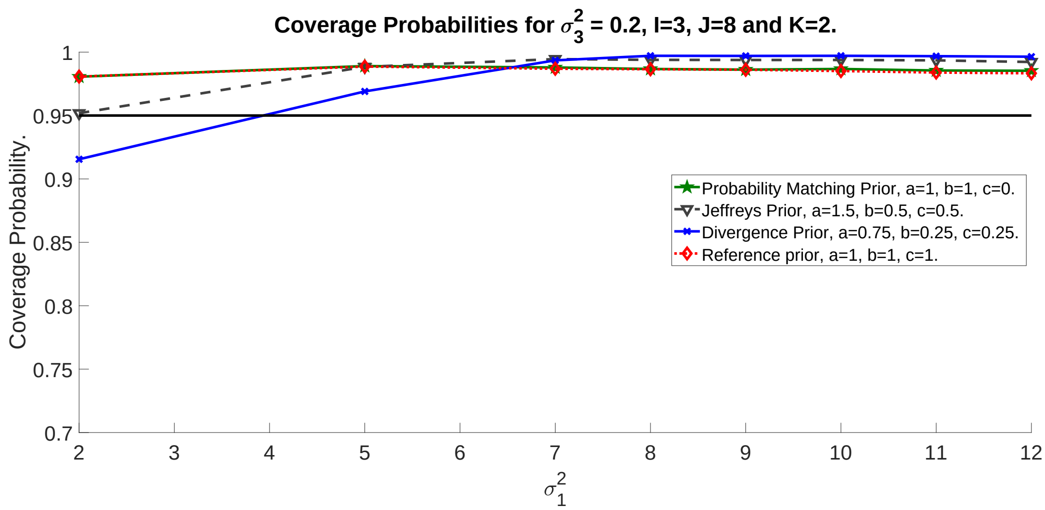
$$
\begin{aligned}
\pi_1^l\left(\alpha,\sigma_1^2,\theta,\sigma_2^2\right)=&\frac{\pi_2^l\left(\sigma_1^2,\theta,\sigma_2^2|\alpha\right)exp\left\{E\left[log\left\{h_1\left(\boldsymbol{\theta}\right)\right\}^{\frac{1}{2}}\right]\right\}}{\int_{a_l}^{b_l}exp\left\{E\left[log\left\{h_1\left(\boldsymbol{\theta}\right)\right\}^{\frac{1}{2}}\right]\right\}d\alpha}\\
=&\frac{\left(K\sigma_2^2+\sigma_1^2\right)^{-1}\left(d_l-c_l\right)^{-1}\left(\sigma_1^2\right)^{-1}\left(1-\alpha\right)^{-1}}{K^{-1}log\left(\frac{Kh_l+\sigma_1^2}{Kg_l+\sigma_1^2}\right)log\left(f_le_l^{-1}\right)}\frac{1}{\int_{a_l}^{b_l}\left(1-\alpha\right)^{-1}d\alpha}\\
=&\frac{\left(K\sigma_2^2+\sigma_1^2\right)^{-1}\left(d_l-c_l\right)^{-1}\left(\sigma_1^2\right)^{-1}\left(1-\alpha\right)^{-1}}{K^{-1}log\left(\frac{Kh_l+\sigma_1^2}{Kg_l+\sigma_1^2}\right)log\left(f_le_l^{-1}\right)}\times\frac{1}{log\left(\frac{1-a_l}{1-b_l}\right)}.
\end{aligned}
$$

Finally,

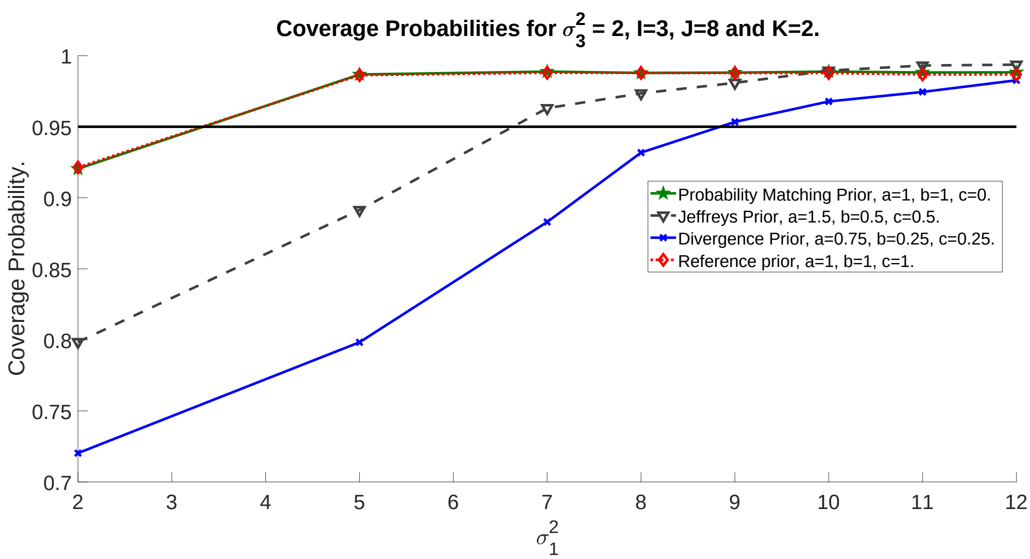
$$
\pi_R\left(\alpha,\theta,\sigma_1^2,\sigma_2^2\right)\propto\lim_{l\to\infty}\frac{\pi_1^l\left(\alpha,\theta,\sigma_1^2,\sigma_2^2\right)}{\pi_1^l\left(\alpha_0,\theta_0,\sigma_{10}^2,\sigma_{20}^2\right)}\propto\left(K\sigma_2^2+\sigma_1^2\right)^{-1}\left(\sigma_1^2\right)^{-1}\left(1-\alpha\right)^{-1}
$$

where $\alpha_0,\theta_0,\sigma_{10}^2$ and σ_{20}^2 are the four inner points in the ranges of the parameters.

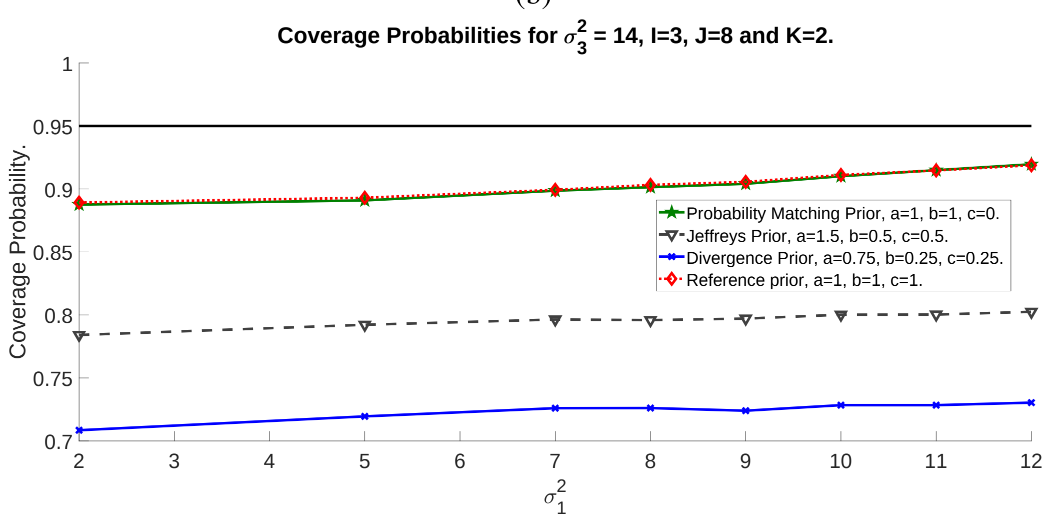
C.4 Additional Simulation Results



(a)



(b)



(c)

Figure A.8: Coverage for (a) $\sigma_3^2 = 0.2$ (b) $\sigma_3^2 = 2$, and (c) $\sigma_3^2 = 14$.

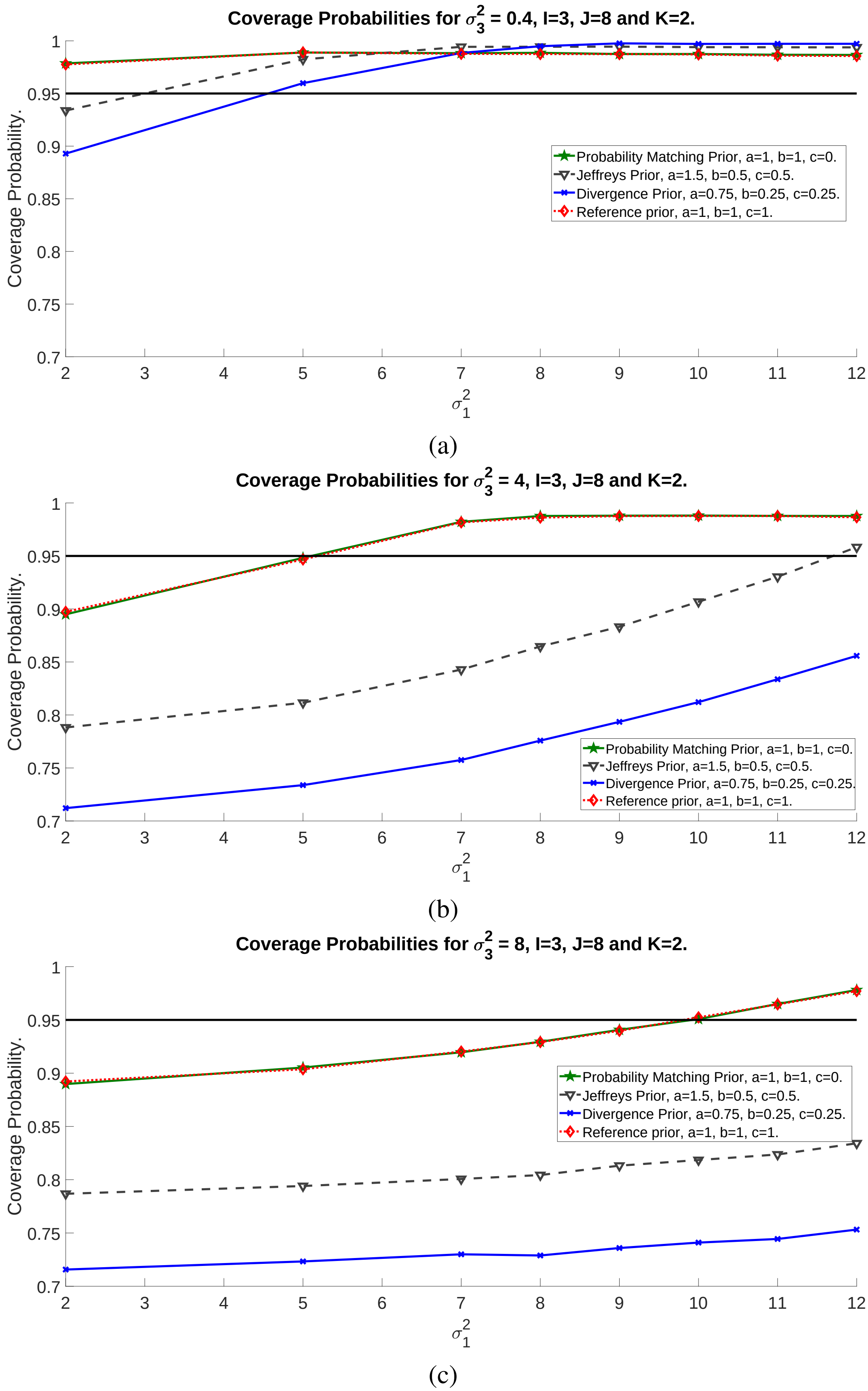
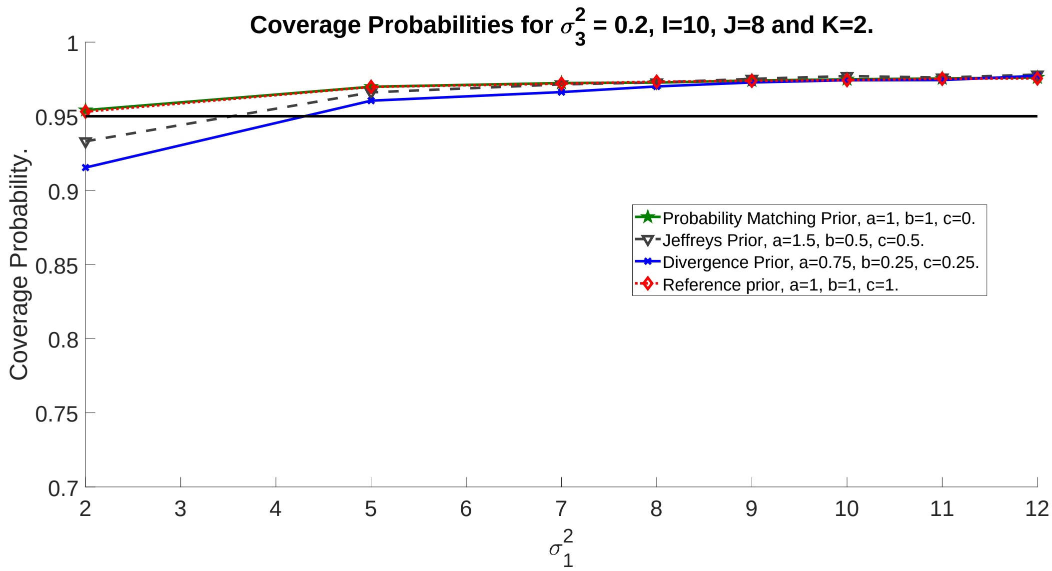
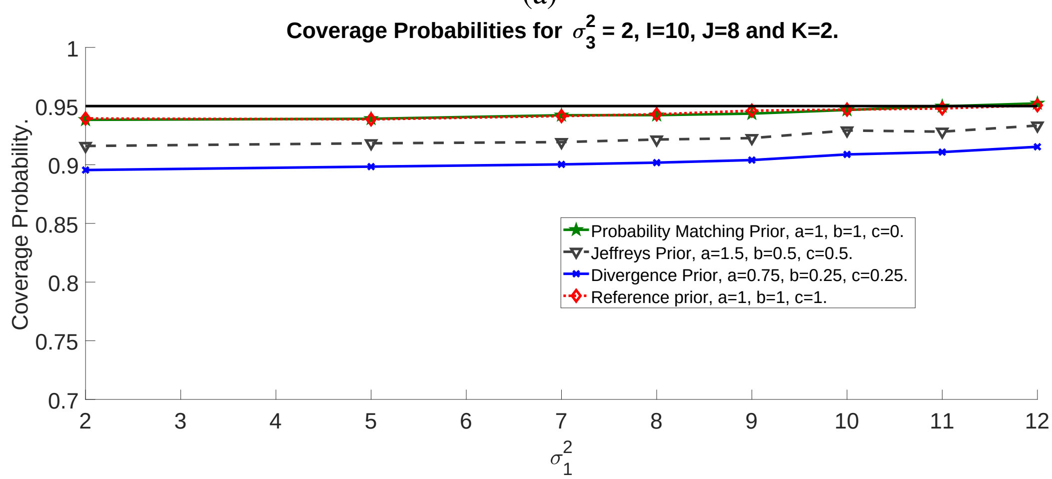


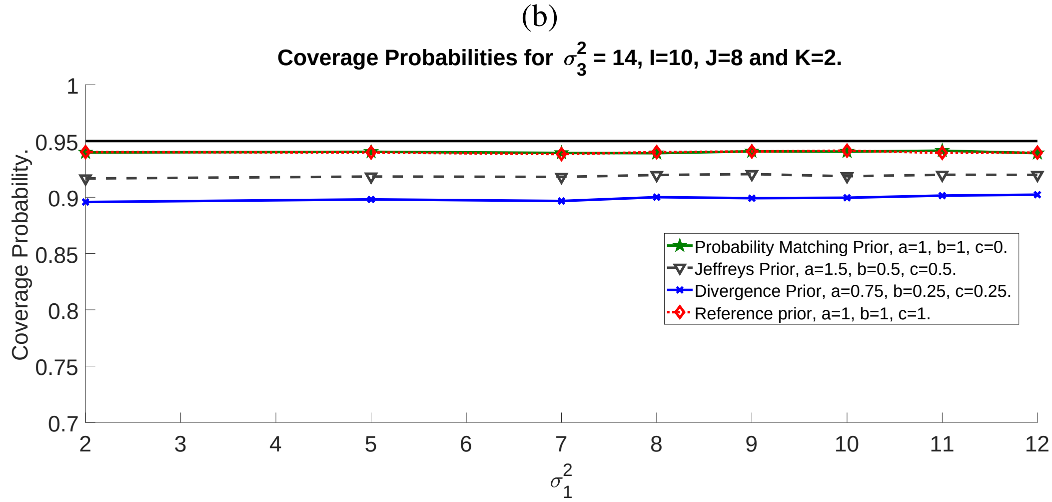
Figure A.9: Coverage for (a) $\sigma_3^2 = 0.4$ (b) $\sigma_3^2 = 4$, and (c) $\sigma_3^2 = 8$.



(a)

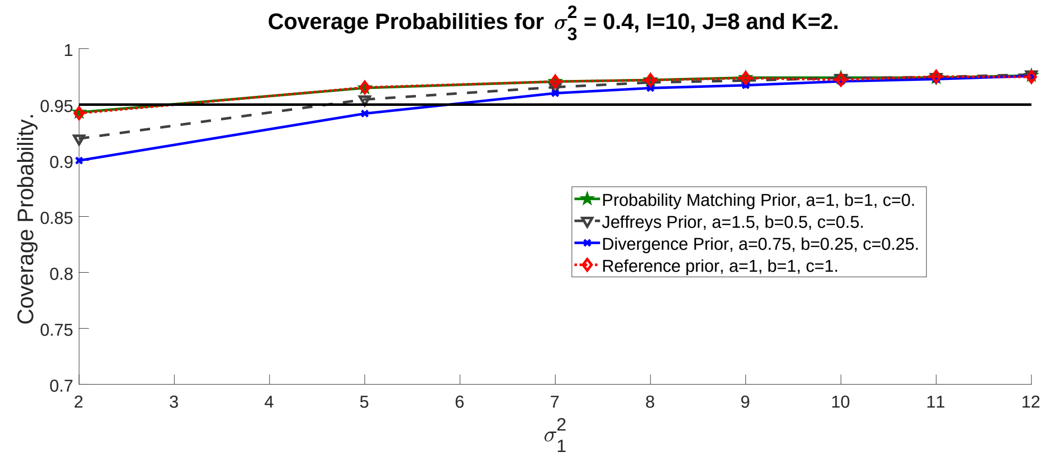


(b)

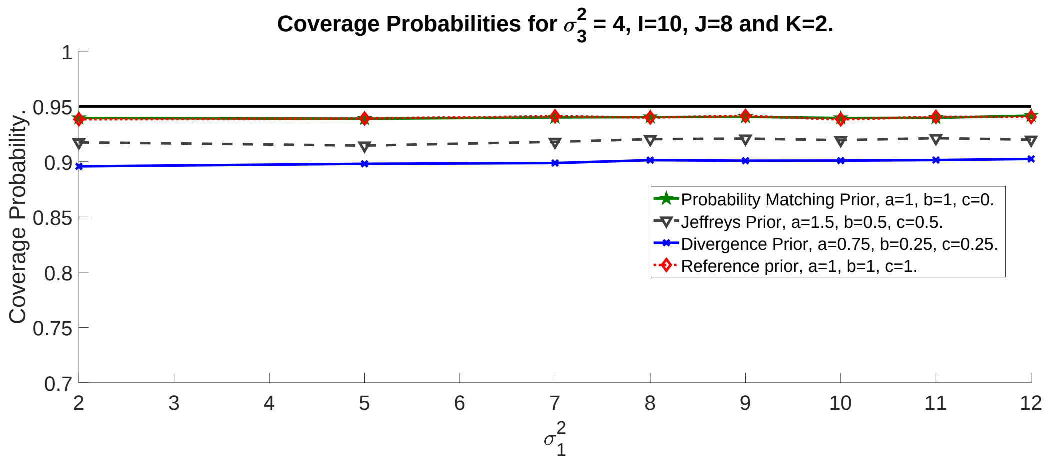


(c)

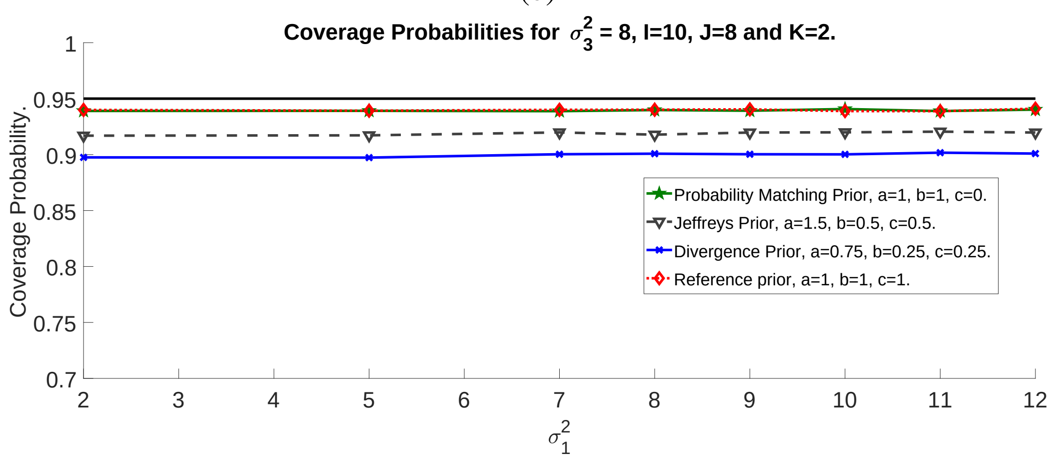
Figure A.10: Coverage for (a) $\sigma_3^2 = 0.2$ (b) $\sigma_3^2 = 2$, and (c) $\sigma_3^2 = 14$.



(a)

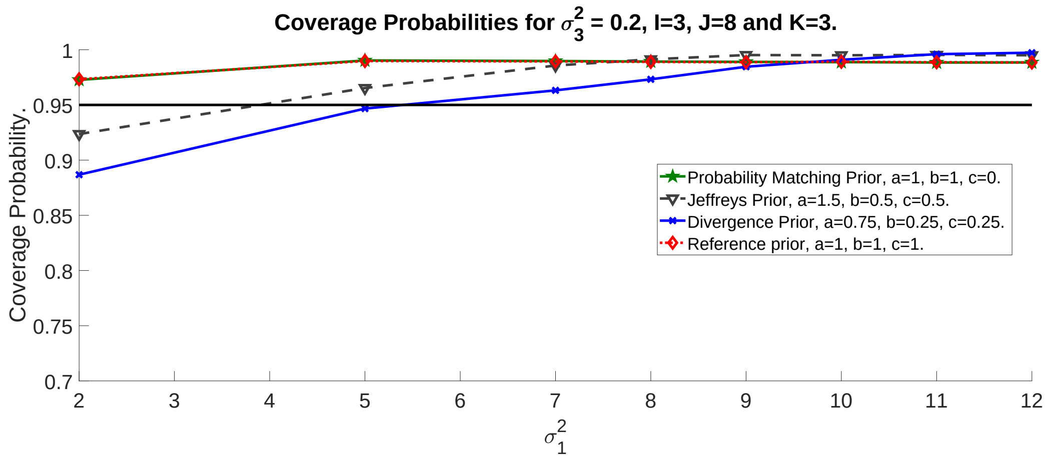


(b)

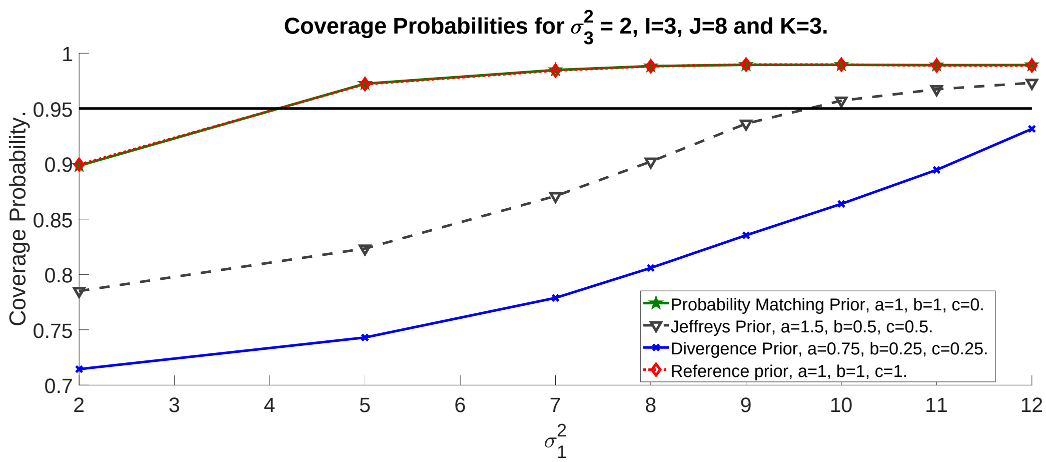


(c)

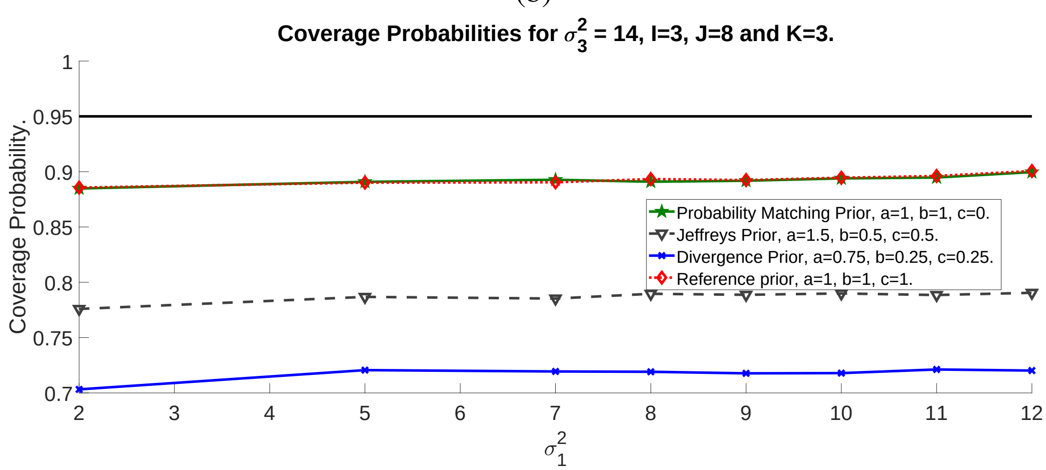
Figure A.11: Coverage for (a) $\sigma_3^2 = 0.4$ (b) $\sigma_3^2 = 4$, and (c) $\sigma_3^2 = 8$.



(a)

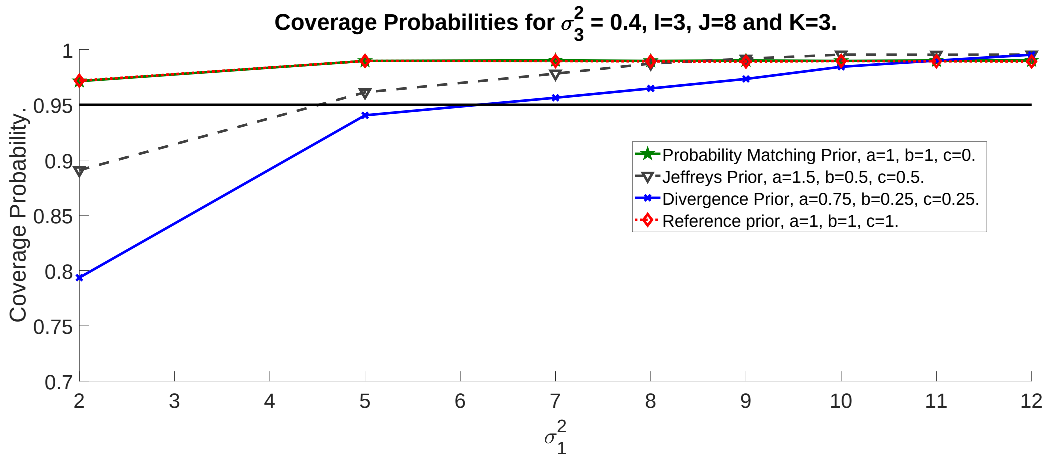


(b)

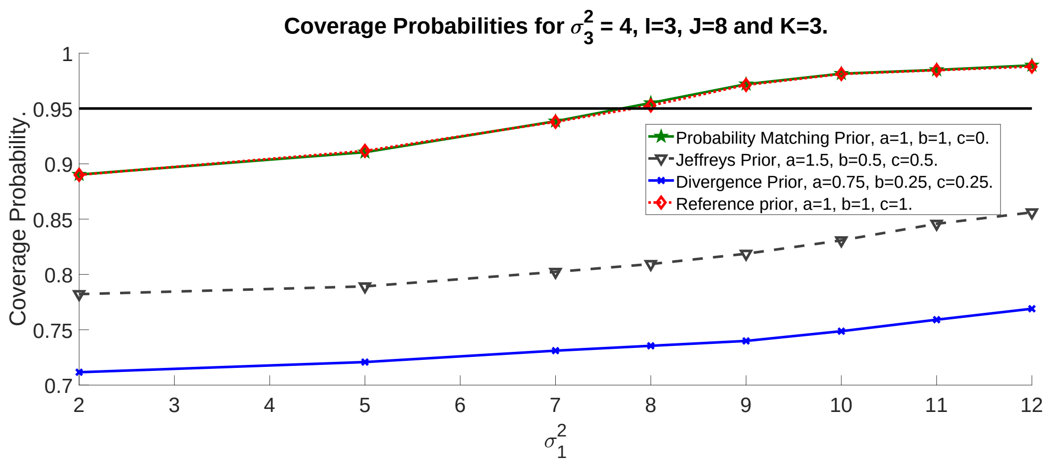


(c)

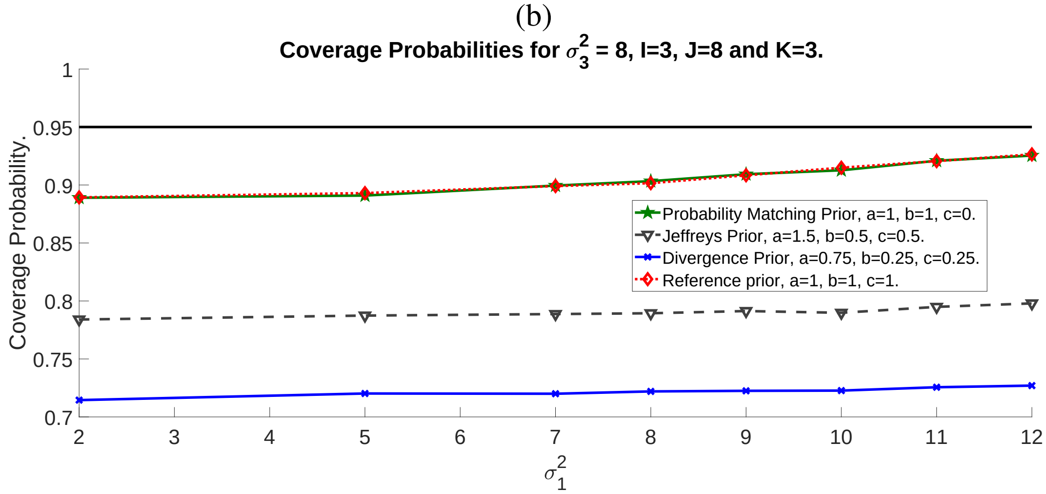
Figure A.12: Coverage for (a) $\sigma_3^2 = 0.2$, (b) $\sigma_3^2 = 2$, and (c) $\sigma_3^2 = 14$.



(a)

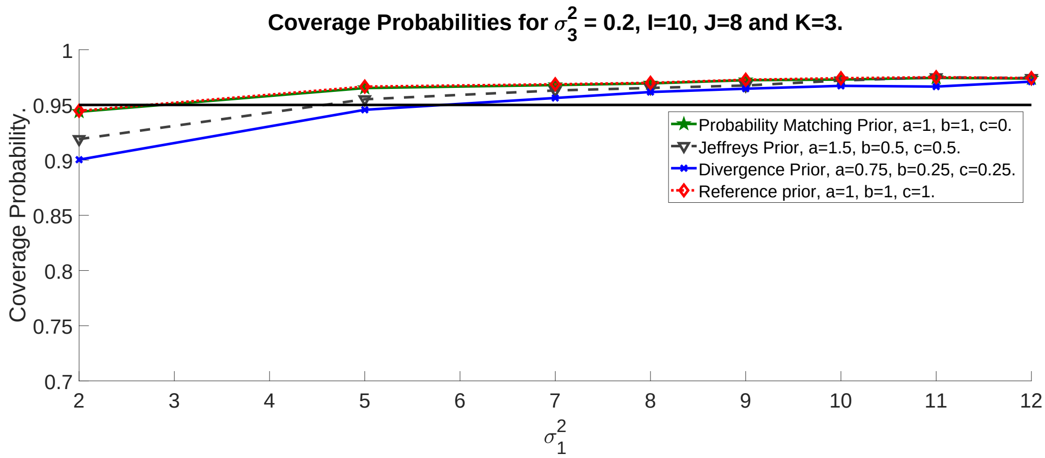


(b)

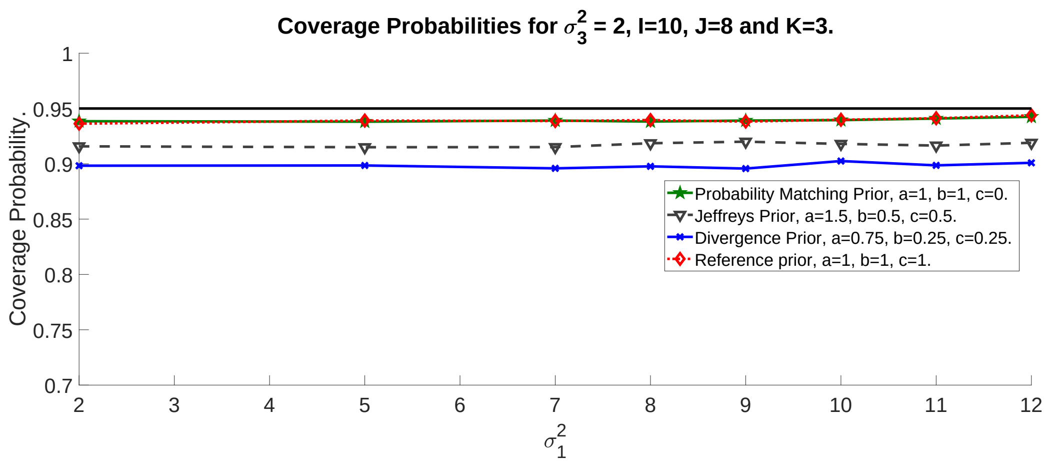


(c)

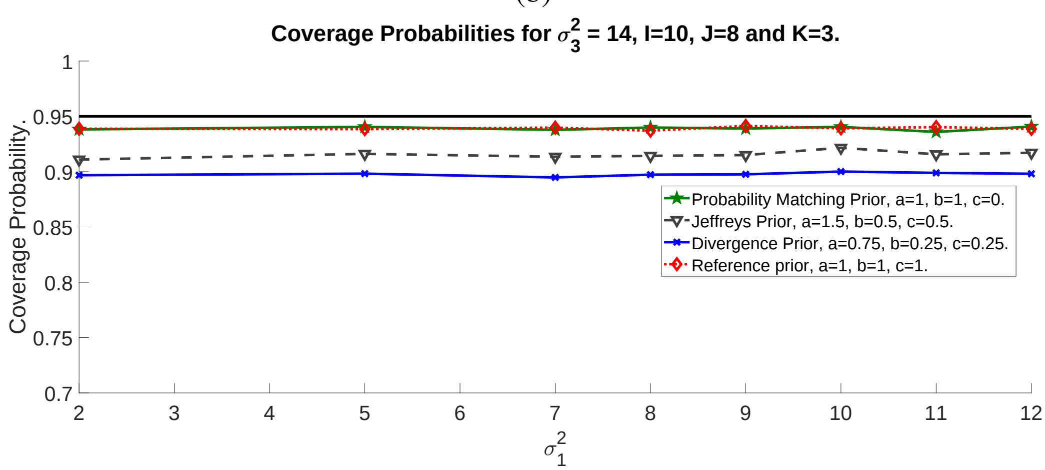
Figure A.13: Coverage for (a) $\sigma_3^2 = 0.4$ (b) $\sigma_3^2 = 4$, and (c) $\sigma_3^2 = 8$.



(a)

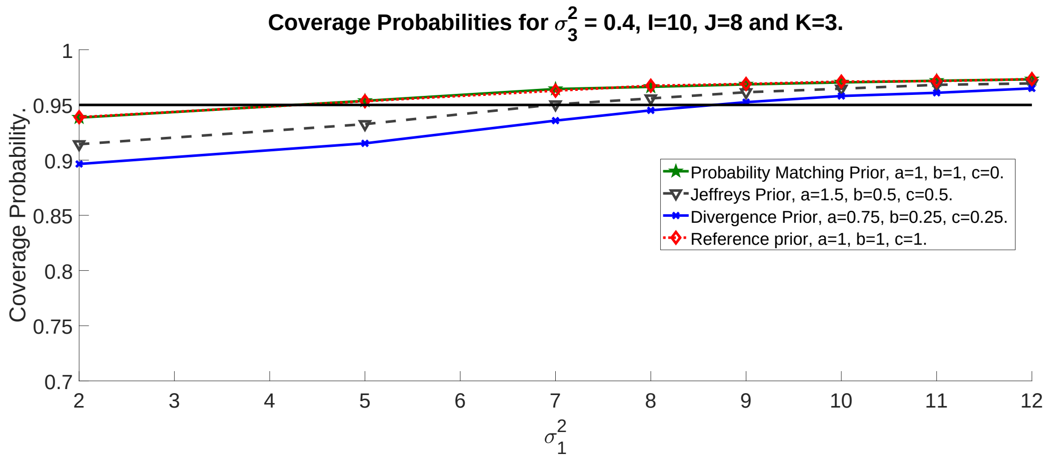


(b)

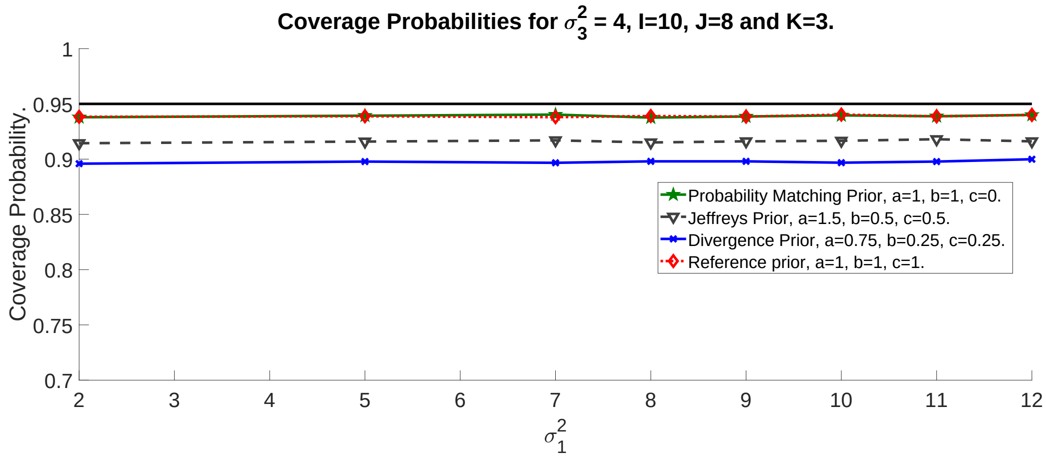


(c)

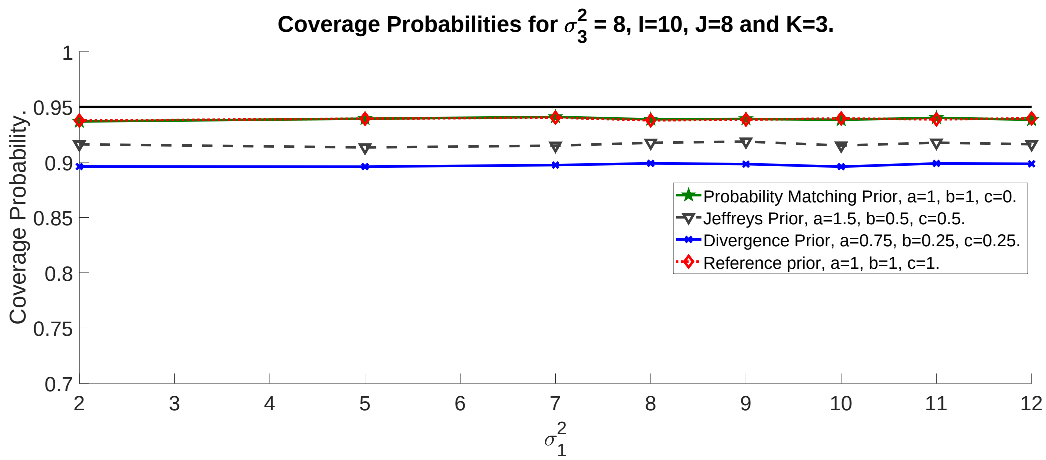
Figure A.14: Coverage for (a) $\sigma_3^2 = 0.2$ (b) $\sigma_3^2 = 2$, and (c) $\sigma_3^2 = 14$.



(a)

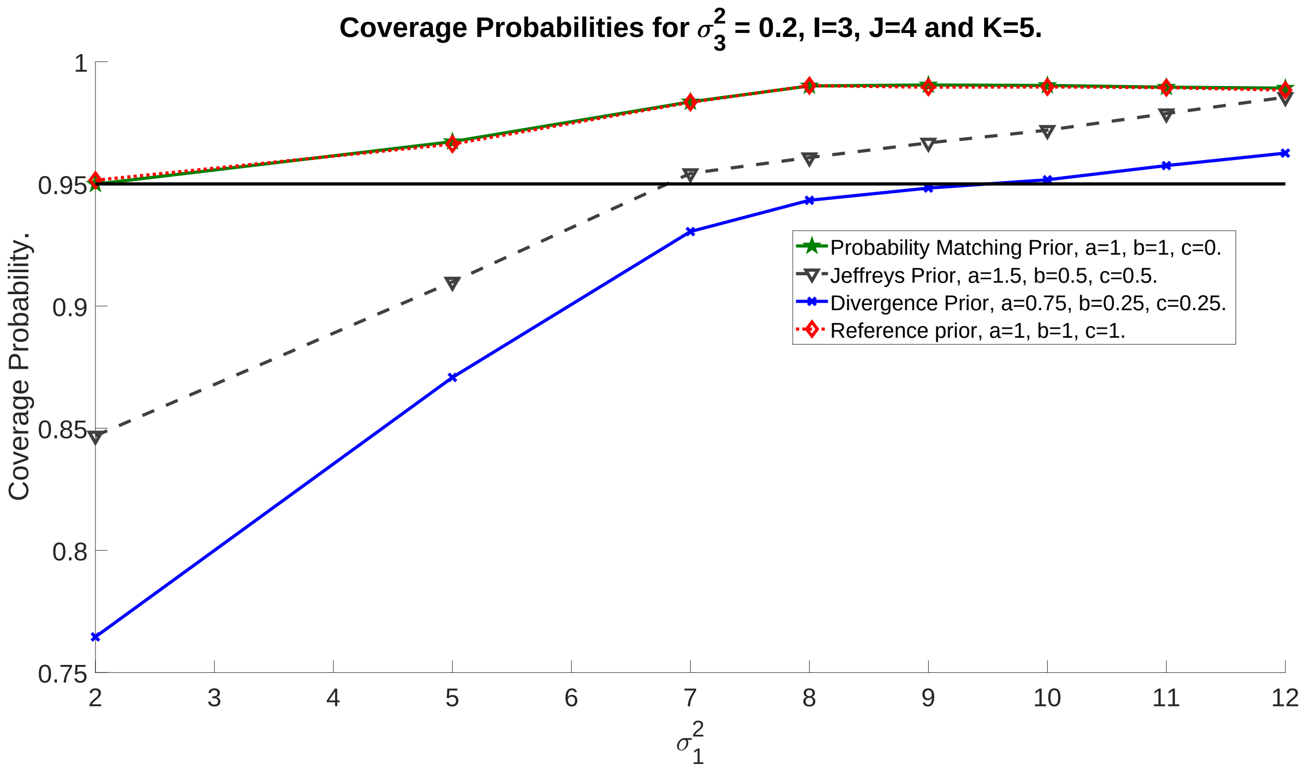


(b)

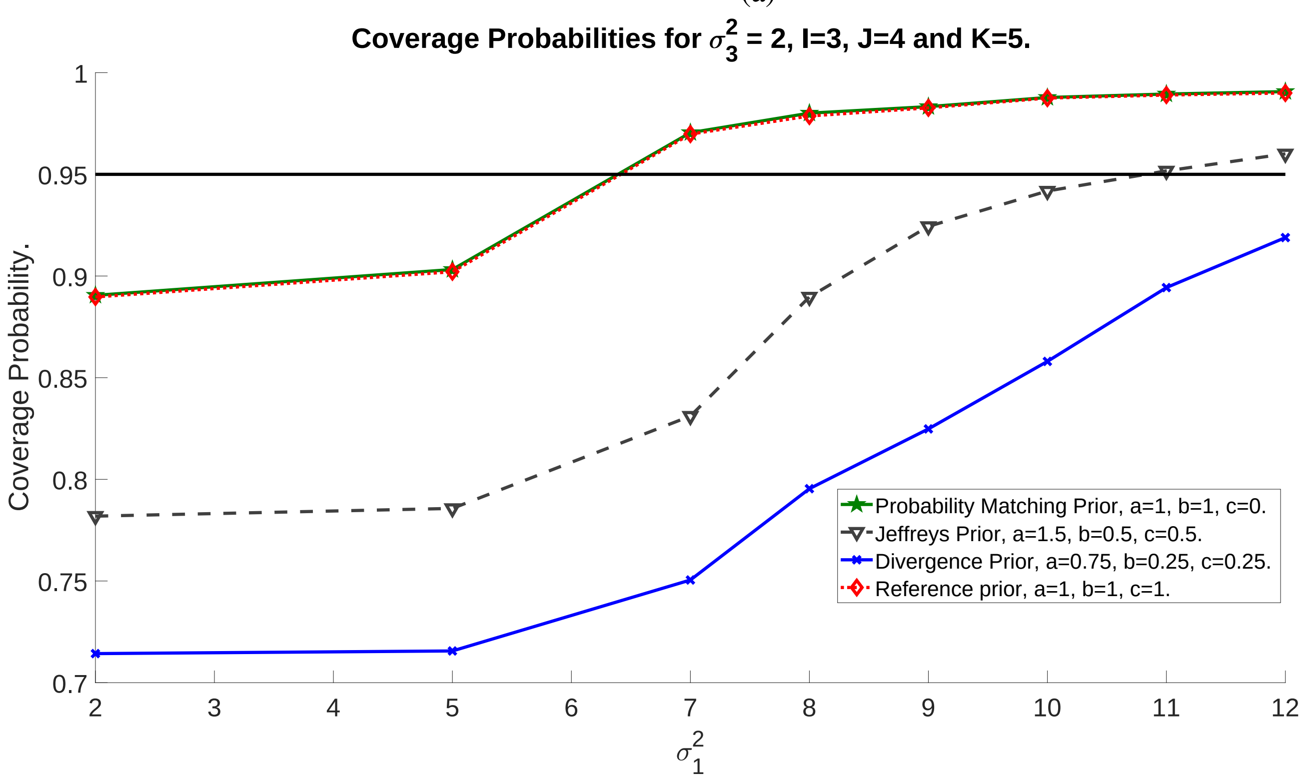


(c)

Figure A.15: Coverage for (a) $\sigma_3^2 = 0.4$ (b) $\sigma_3^2 = 4$, and (c) $\sigma_3^2 = 8$.

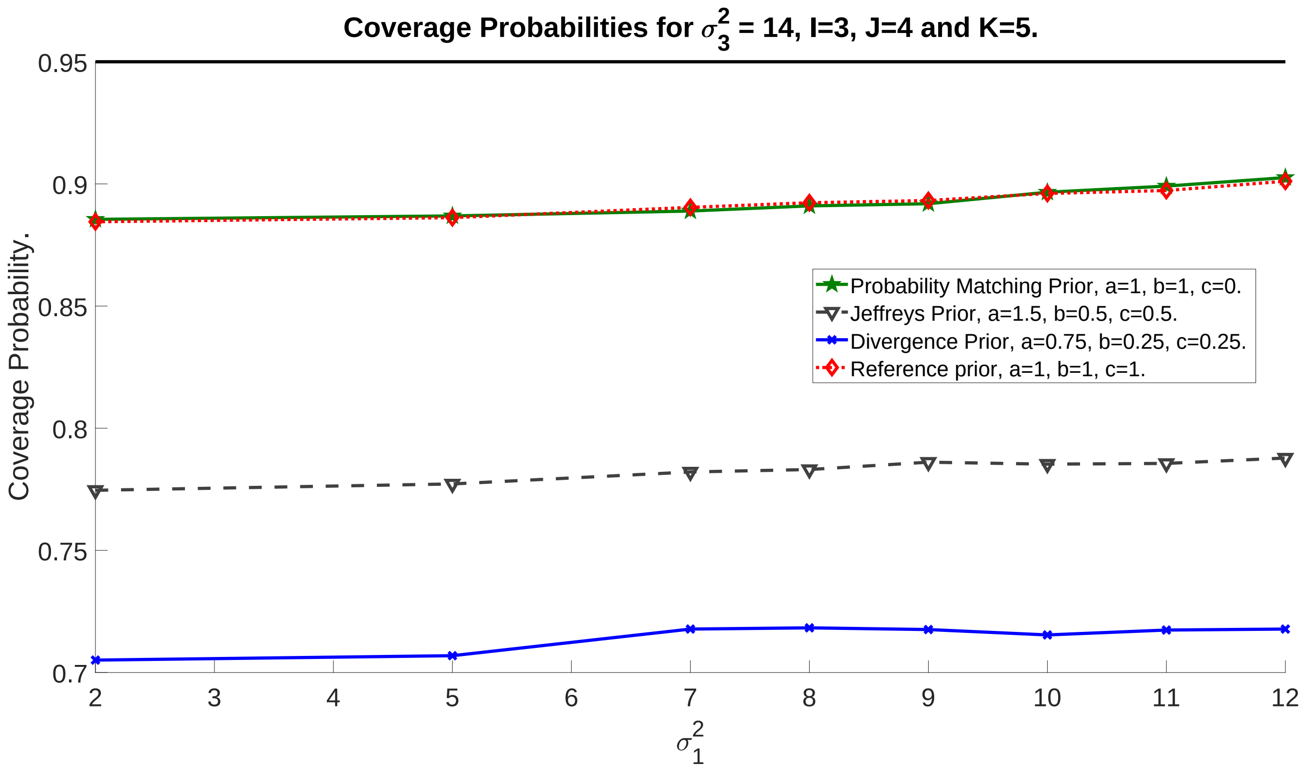


(a)

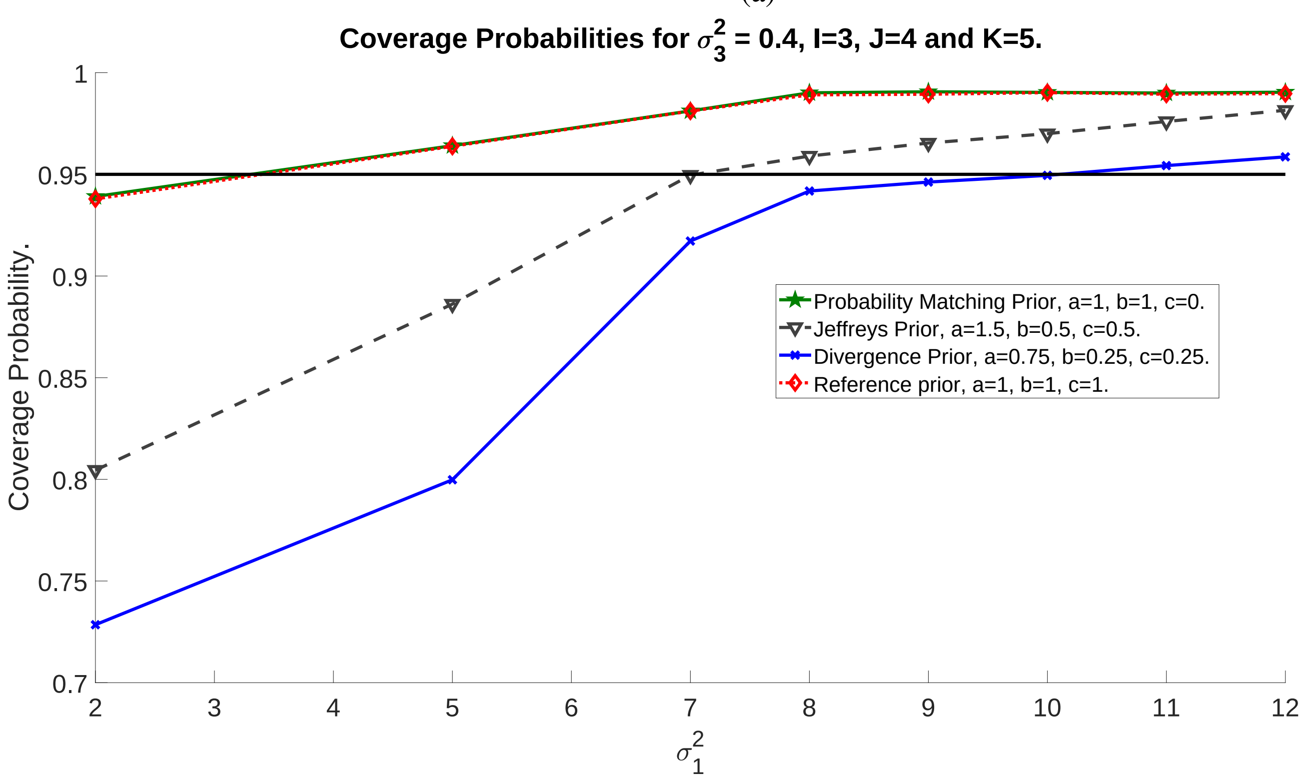


(b)

Figure A.16: Coverage for $I = 3$, $J = 4$, $K = 5$ When (a) $\sigma_3^2 = 0.2$ and (b) $\sigma_3^2 = 2$.

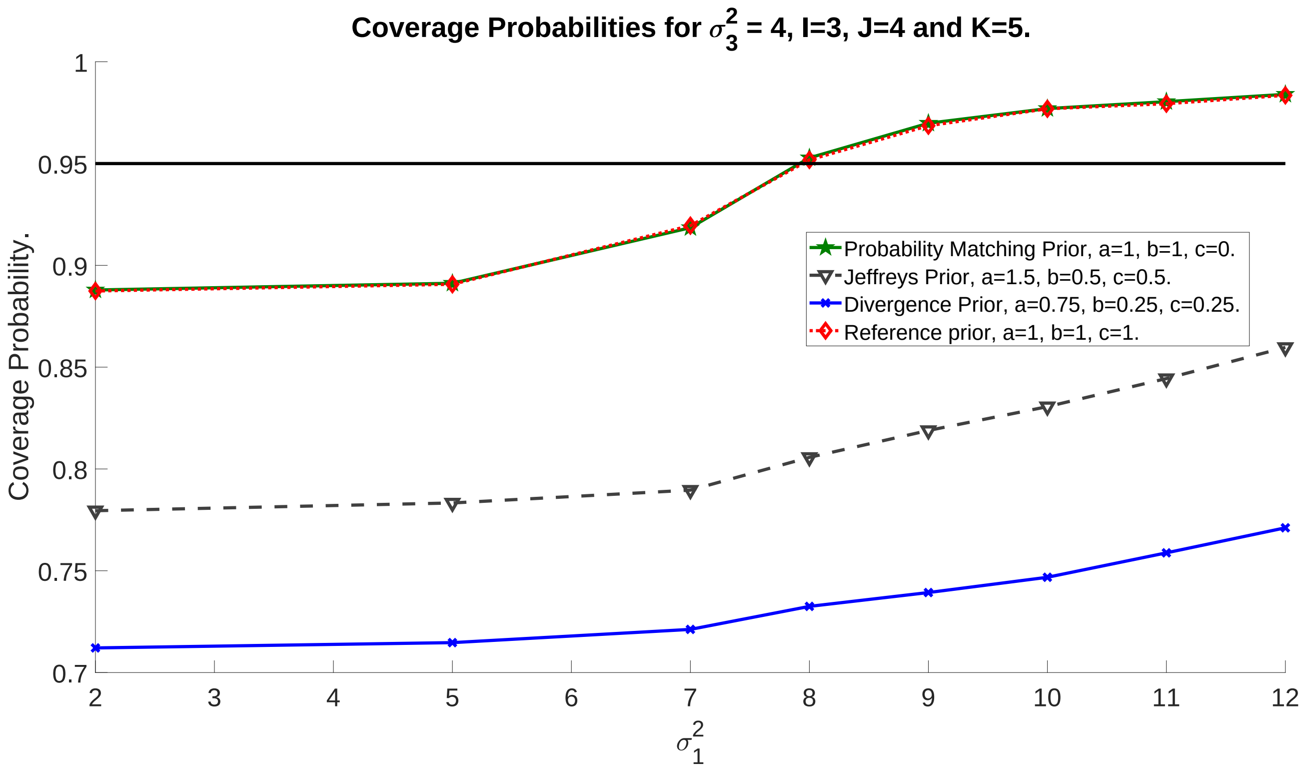


(a)

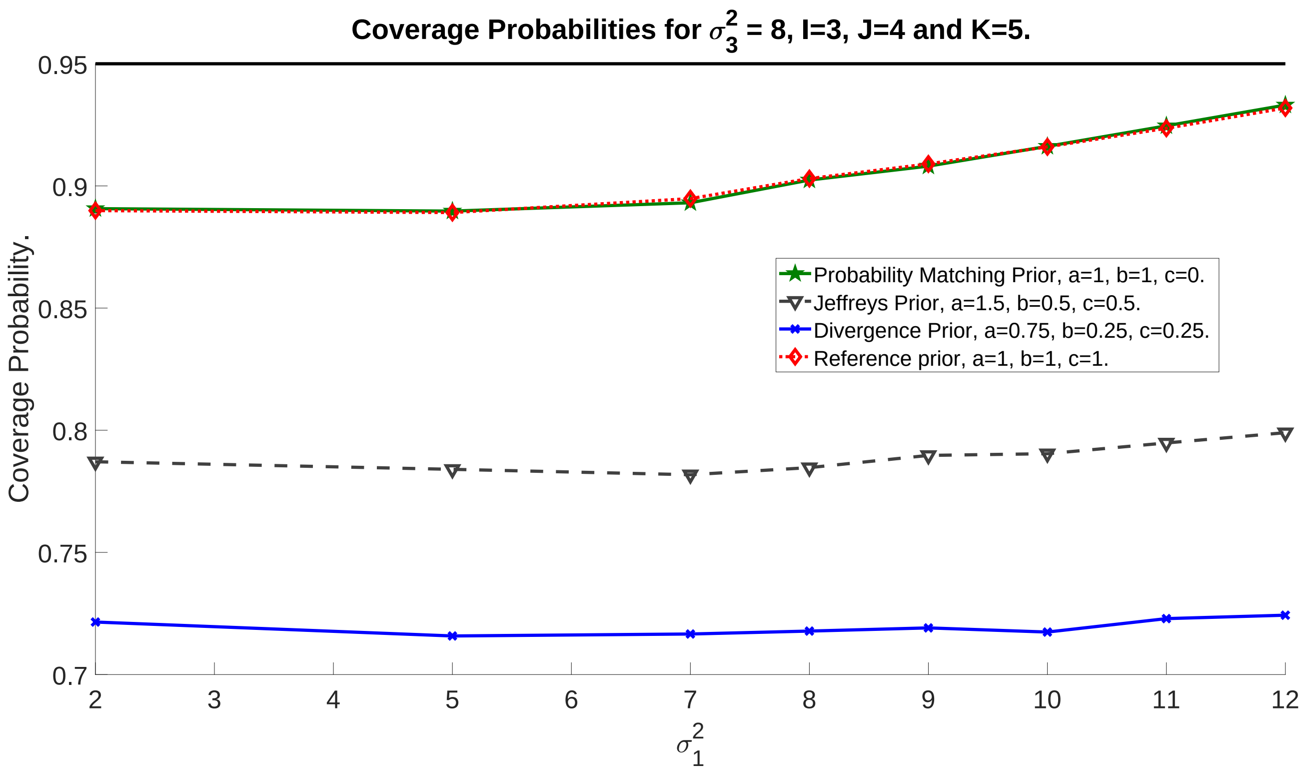


(b)

Figure A.17: Coverage for $I = 3$, $J = 4$, $K = 5$ When (a) $\sigma_3^2 = 14$ and (b) $\sigma_3^2 = 0.4$.

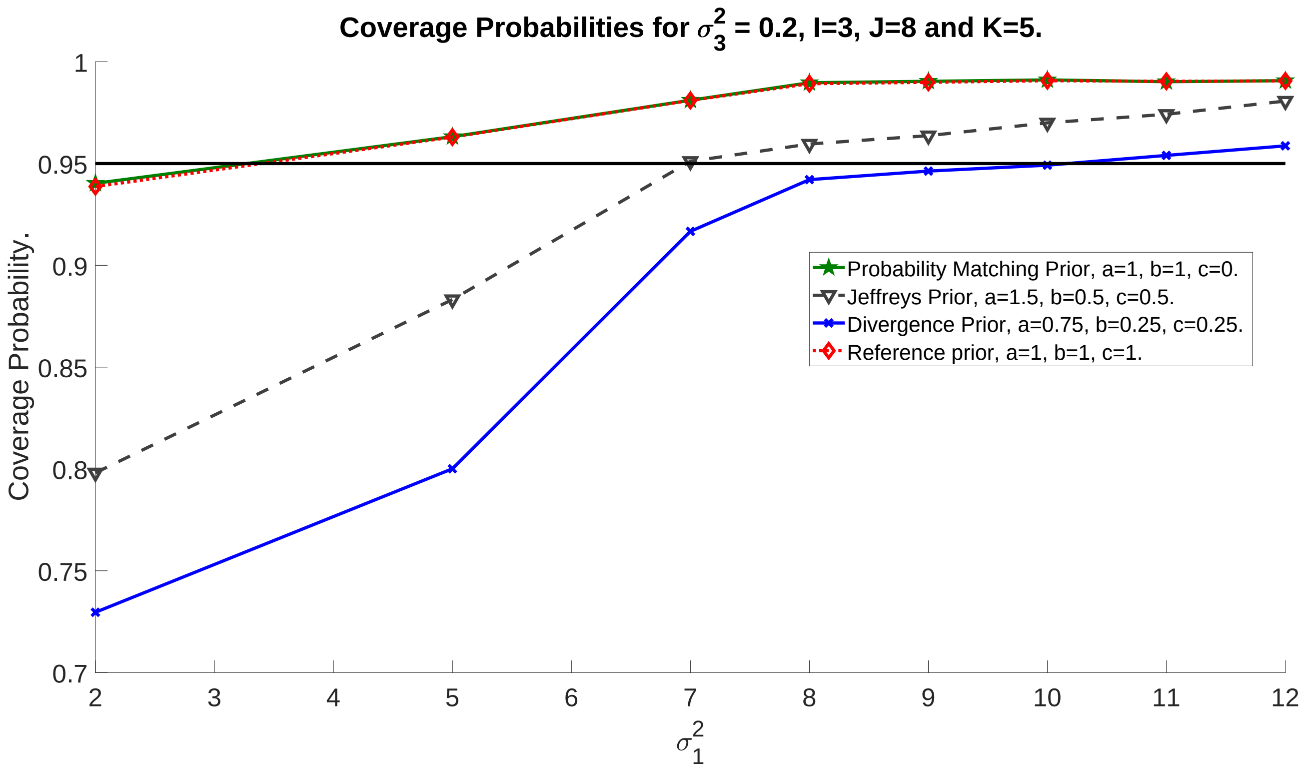


(a)

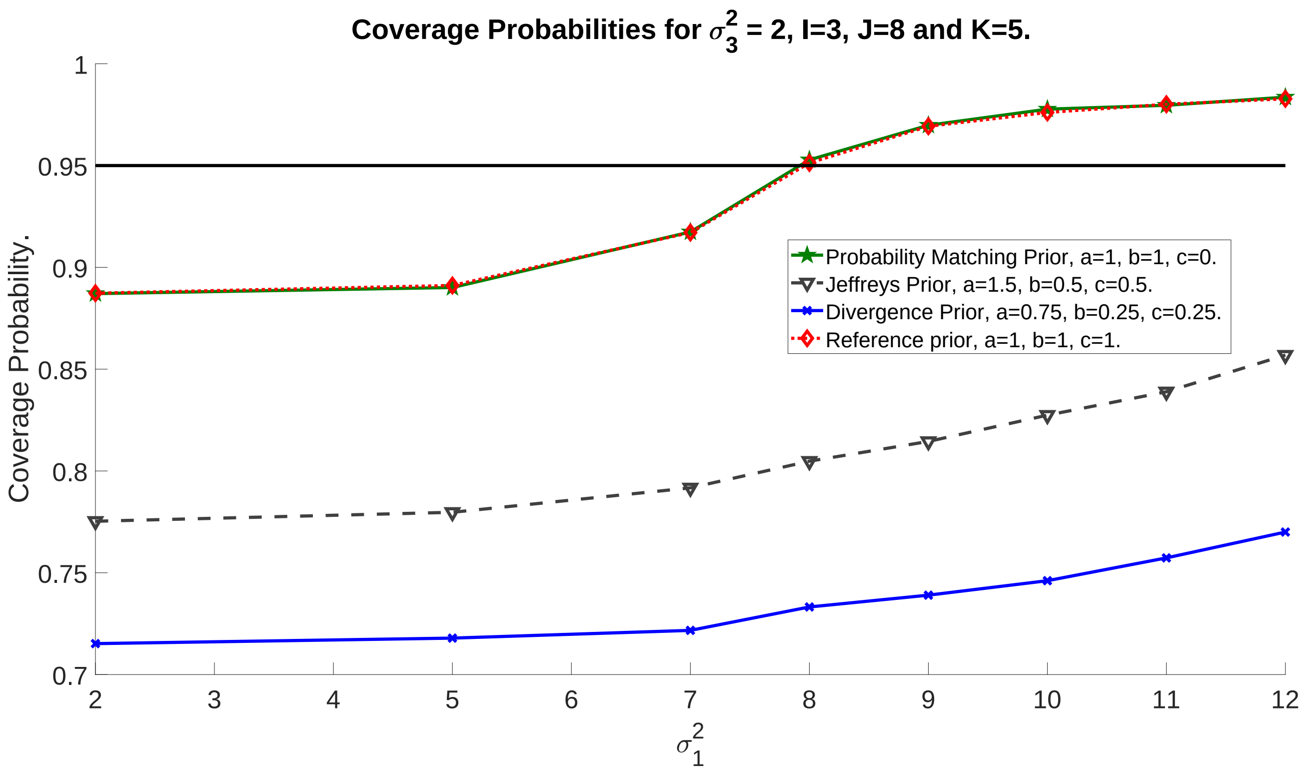


(b)

Figure A.18: Coverage for $I = 3$, $J = 4$, $K = 5$ When (a) $\sigma_3^2 = 4$ and (b) $\sigma_3^2 = 8$.

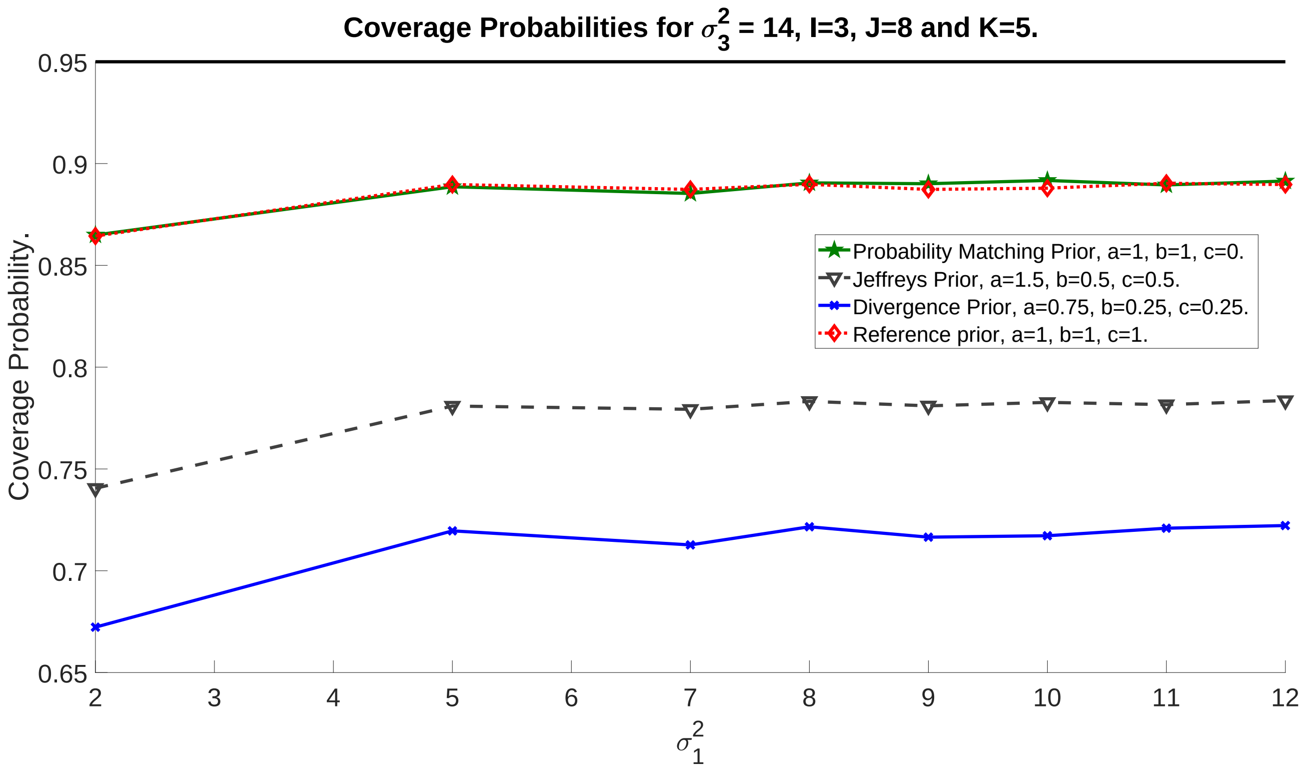


(a)

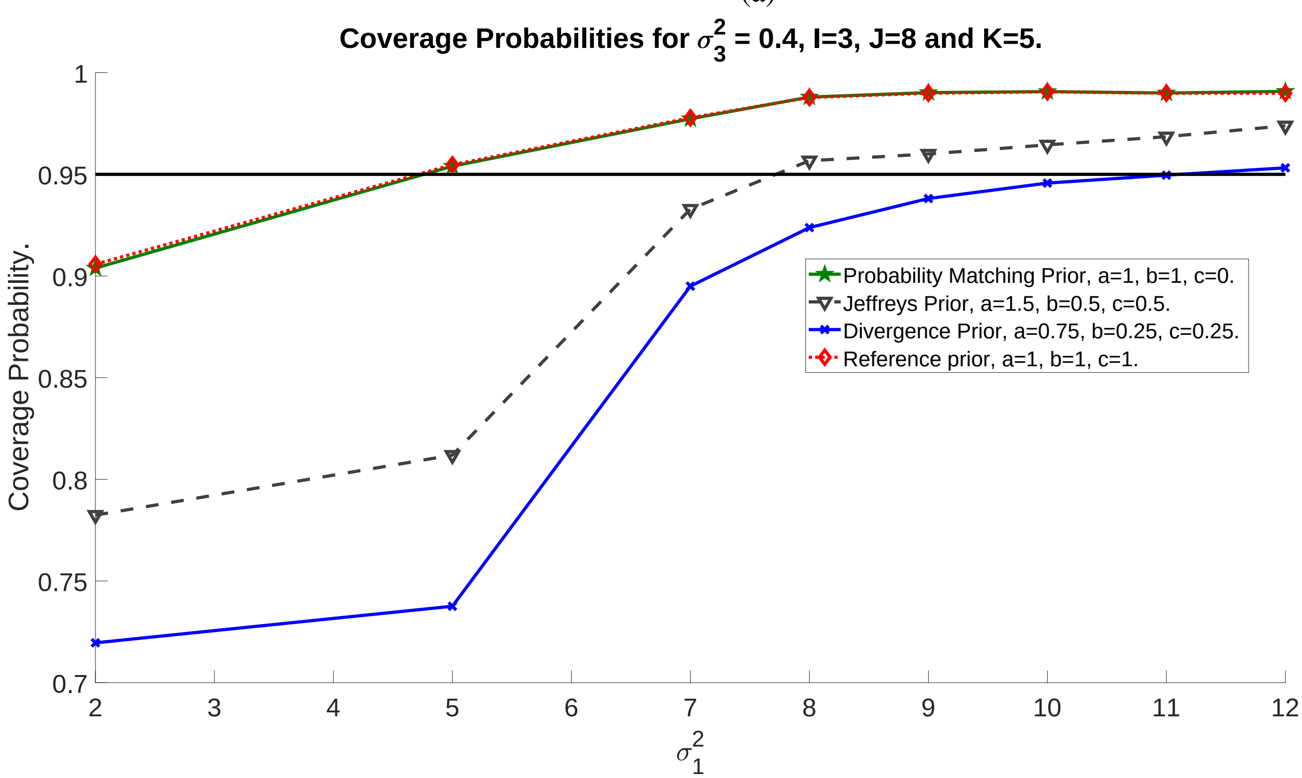


(b)

Figure A.19: Coverage for $I = 3$, $J = 8$, $K = 5$ When (a) $\sigma_3^2 = 0.2$ and (b) $\sigma_3^2 = 2$.

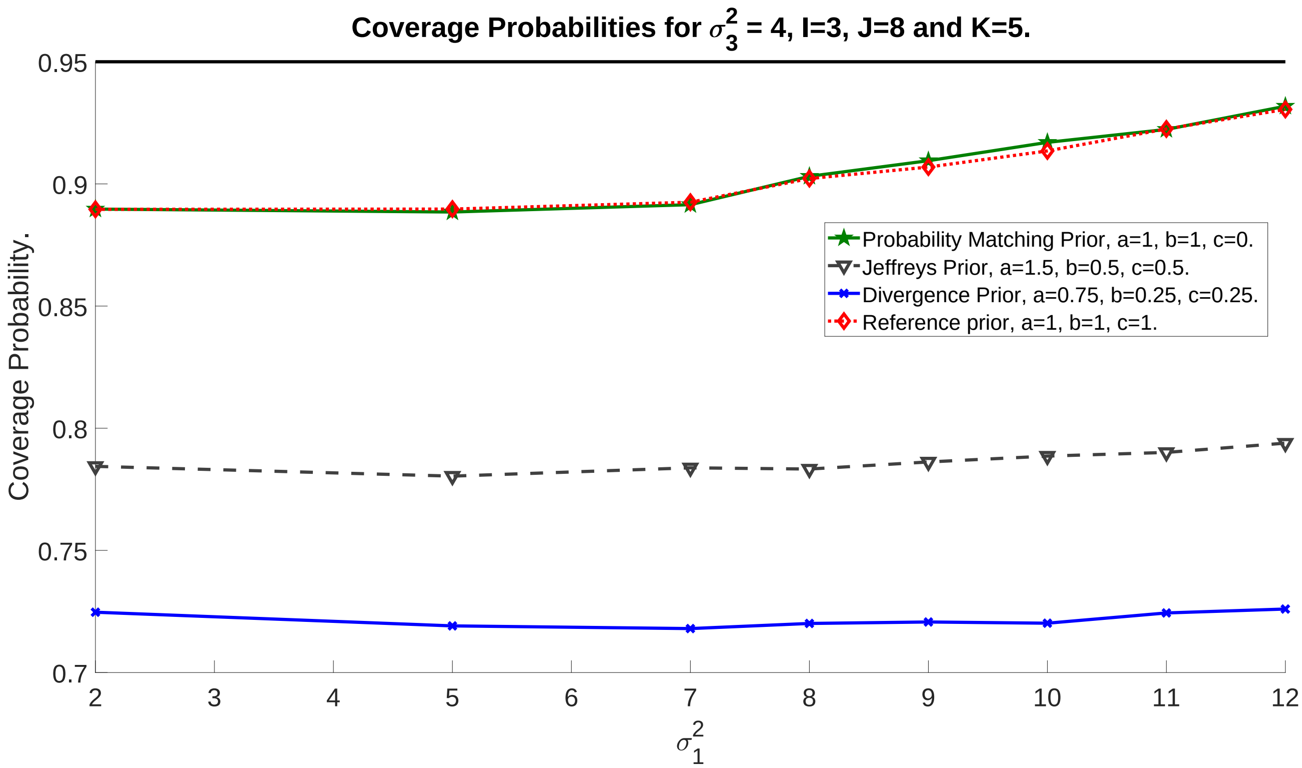


(a)

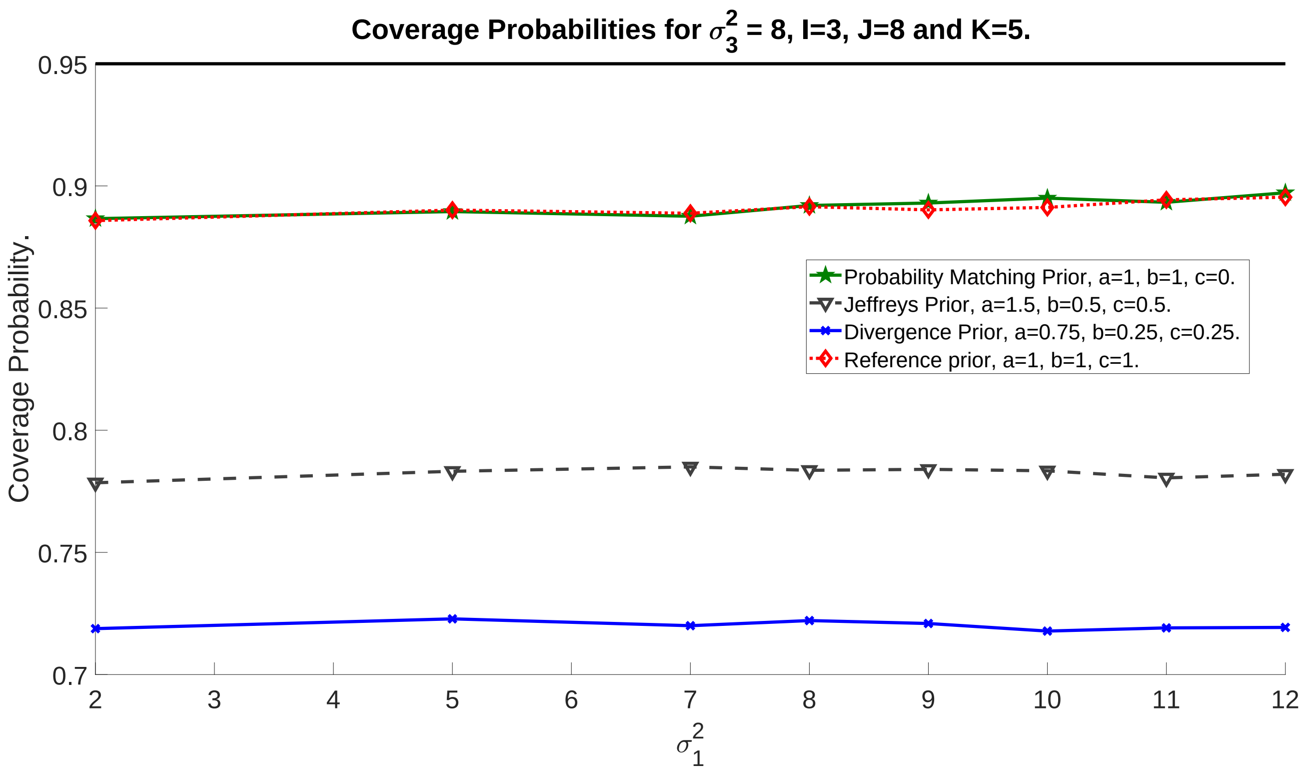


(b)

Figure A.20: Coverage for $I = 3$, $J = 8$, $K = 5$ When (a) $\sigma_3^2 = 14$ and (b) $\sigma_3^2 = 0.4$.



(a)



(b)

Figure A.21: Coverage for $I = 3$, $J = 8$, $K = 5$ When (a) $\sigma_3^2 = 4$ and (b) $\sigma_3^2 = 8$.

C.5 MATLAB Code for the Simulation Study

```
clc
clear
tic
I=3;
J=8;
K=3; % Change I, J and K accordingly
v1=I*J*(K-1);
v2=I*(J-1);
v3=I-1;
a=1;
b=1;
c=0; %a=b=1 and c=0 is the PMP
df1=v2+2*c-2; % Numerator df of posterior
df2=v1+v3+2*a-2;% Denominator df of posterior
alf=[];
Count=[];
Length=[];
SI=[];
for j=1:9
s1=[1 2 5 7 8 9 10 11 12 ] ;
s2=[0.2 0.6 0.8 0.9 1 3 4 11 14];
s3=0.2;
for k=1:9
alf(k,j)=1-(s1(k)/(s1(k)+K*s2(j)+J*K*s3)); %Cronbach's Alpha
N=10000;
ro=0:0.001:1;
count=0;
L=0;
for i=1:N
v1m1=s1(k)*chi2rnd(v1);
v2m2=(s1(k)+K*s2(j))*chi2rnd(v2);
v3m3=(s1(k)+K*s2(j)+J*K*s3)*chi2rnd(v3);
%log of posterior kernel using Jeffreys prior
lnf=((v3+1-2*b)/2)*log(1-ro)-((v1+v3+2*a-2)/2)*log(v1m1+v3m3*(1-ro))
+log(fcdf(v2m2*(v1+v3+2*a-2).*(v2+2*c-2).^-1.*(v1m1+v3m3.*(1-ro)).^-1,df1,df2));
```

```
f=exp(lnf); %Kernel of Posterior for Cronbach's alpha
K_star=inv(0.0001*sum(f)); %normalizing constant
pro=K_star*f; % Normalized Posterior for Cronbach's alpha
C=0.0001*cumsum(pro);
lo95=ro(min(find(C>=0.025)));
up95=ro(min(find(C>=0.975)));
l=up95-lo95;
L=L+l;
if lo95<=alf(k,j) & up95>=alf(k,j)
count=count+1;
end
end
count;
leng=L./N;
Count=[Count;count];
Length=[Length;leng];
si=sqrt(1/(N-1)*(sum(Length.^2)-(sum(Length))^2/N));
SI=[SI;si];
end
end
toc
coverage=Count./N;
coverage_2=coverage'
Length_2=Length';
SI_2=SI';
MATRIX=[coverage_2(1:9); Length_2(1:9); SI_2(1:9); coverage_2(10:18)
; Length_2(10:18); SI_2(10:18); coverage_2(19:27)
; Length_2(19:27); SI_2(19:27); coverage_2(28:36)
; Length_2(28:36); SI_2(28:36); coverage_2(37:45)
; Length_2(37:45); SI_2(37:45); coverage_2(46:54)
; Length_2(46:54); SI_2(46:54); coverage_2(55:63)
; Length_2(55:63); SI_2(55:63); coverage_2(64:72)
; Length_2(64:72); SI_2(64:72); coverage_2(73:81)
; Length_2(73:81); SI_2(73:81)];
% Use this function to ouput results in the form of a matrix in Excel
xlswrite('PMP1sig3=0.2.xlsx',MATRIX)
```

C.6 MATLAB Code for Examples 1 and 2

```
%  Examples 1 and 2 for  the  3 Component model
clear
clc
I=10;
J=2;
K=2;
% Data for Example 1
y=[2.004 2.713 0.603 0.252; 4.342 4.229 3.344 3.057; 0.869 -2.621 -3.896 -3.696;
3.531 4.185 1.722 0.380; 2.579 4.271 -2.101 0.651; -1.404 -1.003 -0.775 -2.202;
-1.676 -0.208 -9.139 -8.653; 1.670 2.426 1.834 1.2; 2.141 3.527 0.462 0.665;
-1.229 -0.596 4.471 1.606];
% Data for Example 2
% y=[62.8 62.6 60.1 62.3 62.7 63.1; 60.0 61.4 57.5 56.9 61.1 58.9; 58.7 57.5
% 63.9 63.1 65.4 63.7; 57.1 56.4 56.9 58.6 64.7 64.5; 55.1 55.1 54.7 54.2 58.8
% 57.5; 63.4 64.9 59.3 58.1 60.5 60.0; 62.5 62.6 61.0 58.7 56.9 57.7; 59.2 59.4
% 65.2 66.0 64.8 64.1; 54.8 54.8 64 64 57.7 56.8; 58.3 59.3 59.2 59.2 58.9 56.6]
y_i_j_dotbar=[(y(:,1)+y(:,2))/K (y(:,3)+y(:,4))/K] ;
y_i_dotdotbar=(y_i_j_dotbar(:,1)+y_i_j_dotbar(:,2))/J ;
ydotdotdotbar=(1/I)*sum(y_i_dotdotbar) ;
y_ij_dot_square=sum([y(:,1).^2+y(:,2).^2 y(:,3).^2+y(:,4).^2]) ;
v1m1=sum(y_ij_dot_square)-K*sum(sum(y_i_j_dotbar.^2)) ;
v2m2=K*sum(sum(y_i_j_dotbar.^2))-J*K*sum(y_i_dotdotbar.^2) ;
v3m3=K*J*sum(y_i_dotdotbar.^2)-I*J*K*ydotdotdotbar.^2 ;
% Posterior using the Probability Matching prior
v1=I*J*(K-1);
v2=I*(J-1);
v3=I-1;
m1=v1m1/v1;
m2=v2m2/v2;
m3=v3m3/v3;
% a=b=1 and c=0 The probability matching prior
% a=b=c=1 The reference prior
a=1;
b=1;
c=0;
```

```
df2=v2+2*c-2; % Denominator df of posterior
df1=v1+v3+2*a-2; % Numerator df of posterior
fun1= @(alpha) (1-alpha).^(0.5*(v3-2*b)).*(v1m1+v3m3.*(1-alpha)).^(-0.5*(v1+v3+2*a-2))
     .*(1-fcdf(((v2+2*c-2).*(v1m1+v3m3.*(1-alpha))/(v2m2.*(v1+v3+2*a-2))),df1,df2));
q1=integral(fun1,0,1)
K1=1/q1
%Check posterior integrates to one
posterior=@(alpha) K1.*(1-alpha).^(0.5*(v3-2*b))
        .*(v1m1+v3m3.*(1-alpha)).^(-0.5*(v1+v3+2*a-2))
        .*(1-fcdf(((v2+2*c-2).*(v1m1+v3m3.*(1-alpha))/(v2m2.*(v1+v3+2*a-2))),df1,df2));
Area=integral(posterior,0,1)
%Plotting Posterior
figure
hold on
x=[0:0.01:1];
y_1=K1.*(1-x).^(0.5*(v3-2*b)).*(v1m1+v3m3.*(1-x)).^(-0.5*(v1+v3+2*a-2))
    .*(1-fcdf(((v2+2*c-2).*(v1m1+v3m3.*(1-x))/(v2m2.*(v1+v3+2*a-2))),df1,df2));
y_2=K1.*(1-x).^(0.5*(v3-2*b)).*(v1m1+v3m3.*(1-x)).^(-0.5*(v1+v3+2*a-2));
plot(x,y_1,'r','LineWidth',3,'MarkerSize',10)
plot(x,y_2,'b--o','LineWidth',3,'MarkerSize',10)
title('Posterior Distribution for Cronbachs alpha using
      the Probability Matching Prior, a=b=1 and c=0.','FontSize',26)
xlabel('\alpha','FontSize',26)
ylabel('Posterior','FontSize',26)
set(gca,'fontsize',24) % Setting the fontsize of the values on the axes
legend({'Exact Posterior', 'Approximate Posterior'},'FontSize',20)
%Mean
fun2= @(alpha) alpha.*K1.*(1-alpha).^(0.5*(v3-2*b))
     .*(v1m1+v3m3.*(1-alpha)).^(-0.5*(v1+v3+2*a-2))
     .*(1-fcdf(((v2+2*c-2).*(v1m1+v3m3.*(1-alpha))/(v2m2.*(v1+v3+2*a-2))),df1,df2));
MEAN_1=integral(fun2,0,1)
fun3= @(alpha) (1-alpha).^(0.5*(v3-2*b)).*(v1m1+v3m3.*(1-alpha)).^(-0.5*(v1+v3+2*a-2));
q1=integral(fun3,0,1);
K2=1/q1;
fun4= @(alpha) K2.*alpha.*(1-alpha).^(0.5*(v3-2*b))
     .*(v1m1+v3m3.*(1-alpha)).^(-0.5*(v1+v3+2*a-2));
% First Moment
MEAN_2=integral(fun4,0,1)
```

```
fun5= @(alpha) K2.*alpha.^(2).*(1-alpha).^(0.5*(v3-2*b))
    .*(v1m1+v3m3.*(1-alpha)).^(-0.5*(v1+v3+2*a-2));
% Second Moment
E_X_2=integral(fun5,0,1)
Variance_post=E_X_2-(MEAN_1)^2
m=m3/m1;
MEAN_3=mean(1-frnd(v3,v1,1,100000)/m) % Simulated Mean
MEDIAN=1-finv(0.5,v3,v1)/m MODE=1-v1m1*(v3-2*b)/(v3m3*(v1+2*a-2+2*b))
CI_90=[1-finv(0.95,v3,v1)/m 1-finv(0.05,v3,v1)/m]
CI_95=[1-finv(0.975,v3,v1)/m 1-finv(0.025,v3,v1)/m]
```

Appendix D: Bayesian Process Control for Cronbach's Alpha

D.1 Proof of Theorem 6.1

From the posterior distribution it follows that

$$\frac{\sigma_\varepsilon^2}{\sigma_\varepsilon^2 + J\sigma_r^2} \sim \frac{\frac{v_1 m_1}{\chi^2_{v_1}}}{\frac{v_2 m_2}{\chi^2_{v_2}}} = \frac{m_1}{m_2}\frac{\frac{\chi^2_{v_2}}{v_2}}{\frac{\chi^2_{v_1}}{v_1}}.$$

Therefore

$$\frac{\sigma_\varepsilon^2}{\sigma_\varepsilon^2 + J\sigma_r^2} \sim (1-\hat{\alpha})\,F_{v_2,v_1}$$

and

$$1-\alpha \sim (1-\hat{\alpha})\,F_{v_2,v_1}.$$

Therefore

$$\alpha|\hat{\alpha} \sim 1-(1-\hat{\alpha})\,F_{v_2,v_1}.$$

Let $f = F_{v_2,v_1}$, then

$$g(f) = K_1\left(\frac{v_2}{v_1}\right)^{\frac{1}{2}v_2} f^{\frac{1}{2}v_2-1}\left(1+\frac{v_2}{v_1}f\right)^{-\frac{1}{2}(v_1+v_2)}.$$

Since $\alpha = 1-(1-\hat{\alpha})\,f$, it follows that

$$f = \frac{1-\alpha}{1-\hat{\alpha}} \text{ and } \quad \left|\frac{df}{d\alpha}\right| = \frac{1}{1-\hat{\alpha}}.$$

Therefore

$$\pi(\alpha|\hat{\alpha}) = K_1 \left(\frac{v_2}{v_1}\right)^{\frac{1}{2}v_2} \left(\frac{1}{1-\hat{\alpha}}\right)^{\frac{1}{2}v_2} (1-\alpha)^{\frac{1}{2}v_2-1} \times \left[1+\frac{v_2}{v_1}\left(\frac{1-\alpha}{1-\hat{\alpha}}\right)\right]^{-\frac{1}{2}(v_1+v_2)}$$

where $K_1 = \frac{\Gamma\left(\frac{v_1+v_2}{2}\right)}{\Gamma\left(\frac{v_1}{2}\right)\Gamma\left(\frac{v_2}{2}\right)}$. Since

$$\begin{aligned} E(\alpha|\hat{\alpha}) &= 1-(1-\hat{\alpha})E(F_{v_2,v_1}) \\ &= 1-(1-\hat{\alpha})\left(\frac{v_1}{v_1-2}\right) \end{aligned}$$

and

$$\begin{aligned} Var(\alpha|\hat{\alpha}) &= (1-\hat{\alpha})^2 Var(F_{v_2,v_1}) \\ &= (1-\hat{\alpha})^2 \frac{2v_1^2(v_2+v_1-2)}{v_2(v_1-2)^2(v_1-4)}. \end{aligned}$$

D.2 Proof of Theorem 6.2

Let $\tilde{f} = F_{\tilde{v}_1,\tilde{v}_2}$. Therefore

$$f(\tilde{f}) = K_2\left(\frac{\tilde{v}_1}{\tilde{v}_2}\right)^{\frac{1}{2}\tilde{v}_1} \tilde{f}^{\frac{1}{2}\tilde{v}_1-1}\left(1+\frac{\tilde{v}_1}{\tilde{v}_2}\tilde{f}\right)^{-\frac{1}{2}(\tilde{v}_1+\tilde{v}_2)}.$$

Since $\hat{\alpha}_f = 1-(1-\alpha)\tilde{f}$, it follows that

$$\tilde{f} = \frac{1-\hat{\alpha}_f}{1-\alpha} \text{ and } \left|\frac{d\tilde{f}}{d\hat{\alpha}_f}\right| = \frac{1}{1-\alpha}.$$

Therefore

$$\begin{aligned} f(\hat{\alpha}_f|\alpha) &= K_2\left(\frac{\tilde{v}_1}{\tilde{v}_2}\right)^{\frac{1}{2}\tilde{v}_1}\left(\frac{1-\hat{\alpha}_f}{1-\alpha}\right)^{\frac{1}{2}\tilde{v}_1-1}\left[1+\frac{\tilde{v}_1}{\tilde{v}_2}\left(\frac{1-\hat{\alpha}_f}{1-\alpha}\right)\right]^{-\frac{1}{2}(\tilde{v}_1+\tilde{v}_2)}\left(\frac{1}{1-\alpha}\right) \\ &= K_2\left(\frac{\tilde{v}_1}{\tilde{v}_2}\right)^{\frac{1}{2}\tilde{v}_1}\left(\frac{1}{1-\alpha}\right)^{\frac{1}{2}\tilde{v}_1}(1-\hat{\alpha}_f)^{\frac{1}{2}\tilde{v}_1-1}\left[1+\frac{\tilde{v}_1}{\tilde{v}_2}\left(\frac{1-\hat{\alpha}_f}{1-\alpha}\right)\right]^{-\frac{1}{2}(\tilde{v}_1+\tilde{v}_2)} \end{aligned}$$

where $K_2 = \frac{\Gamma\left(\frac{\tilde{v}_1+\tilde{v}_2}{2}\right)}{\Gamma\left(\frac{\tilde{v}_1}{2}\right)\Gamma\left(\frac{\tilde{v}_2}{2}\right)}$. Also

$$\begin{aligned} E\left(\hat{\alpha}_f|\alpha\right) &= 1-(1-\alpha)\,E\left(F_{\tilde{v}_1,\tilde{v}_2}\right) \\ &= 1-(1-\alpha)\left(\frac{\tilde{v}_2}{\tilde{v}_2-2}\right) \end{aligned}$$

and

$$\begin{aligned} Var\left(\hat{\alpha}_f|\hat{\alpha}\right) &= (1-\alpha)^2\,Var\left(F_{\tilde{v}_1,\tilde{v}_2}\right) \\ &= (1-\alpha)^2\,\frac{2\left(\tilde{v}_2\right)^2\left(\tilde{v}_2+\tilde{v}_1-2\right)}{\tilde{v}_1\left(\tilde{v}_2-2\right)^2\left(\tilde{v}_2-4\right)}. \end{aligned}$$

D.3 Proof of Theorem 6.3

$$\hat{\alpha}_f|\alpha = 1-(1-\alpha)\,F_{\tilde{v}_1,\tilde{v}_2} \tag{D.1}$$

and

$$\alpha|\hat{\alpha} = 1-(1-\hat{\alpha})\,F_{v_2,v_1} \tag{D.2}$$

From Equation D.1 it follows that

$$\begin{aligned} E\left(\hat{\alpha}_f|\alpha\right) &= 1-(1-\alpha)\,E\left(F_{\tilde{v}_1,\tilde{v}_2}\right) \\ &= 1-(1-\alpha)\left(\frac{\tilde{v}_2}{\tilde{v}_2-2}\right) \end{aligned}$$

and from Equation D.2 it follows that

$$(1-\alpha) \sim (1-\hat{\alpha})\,F_{v_2,v_1}.$$

Therefore

$$\begin{aligned} E\left(\hat{\alpha}_f|\hat{\alpha}\right) &= E\left\{E\left(\hat{\alpha}_f|\alpha\right)\right\} \\ &= 1-(1-\hat{\alpha})\left(\frac{v_1}{v_1-2}\right)\left(\frac{\tilde{v}_2}{\tilde{v}_2-2}\right). \end{aligned}$$

Also for the variance

$$Var\left(\hat{\alpha}_f|\hat{\alpha}\right) = E\left\{Var\left(\hat{\alpha}_f|\alpha\right)\right\} + Var\left\{E\left(\hat{\alpha}_f|\alpha\right)\right\}.$$

Now

$$Var\left(\hat{\alpha}_f|\alpha\right) = (1-\alpha)^2 Var\left(F_{\tilde{v}_1,\tilde{v}_2}\right)$$

and

$$E\left(\hat{\alpha}_f|\alpha\right) = 1-(1-\alpha)\left(\frac{\tilde{v}_2}{\tilde{v}_2-2}\right).$$

Further

$$\begin{aligned}E\left\{Var\left(\hat{\alpha}_f|\alpha\right)\right\} =& E\left(1-\alpha\right)^2 Var\left(F_{\tilde{v}_1,\tilde{v}_2}\right)\\ =& \left\{Var\left(1-\alpha\right)+\left[E\left(1-\alpha\right)\right]^2\right\} Var\left(F_{\tilde{v}_1,\tilde{v}_2}\right).\end{aligned}$$

Since

$$(1-\alpha)\sim(1-\hat{\alpha})\,F_{v_2,v_1}$$

it follows that

$$Var\left(1-\alpha\right) = (1-\hat{\alpha})^2 Var\left(F_{v_2,v_1}\right)$$

and

$$\left[E\left(1-\alpha\right)\right]^2 = (1-\hat{\alpha})^2\left(\frac{v_1}{v_1-2}\right)^2.$$

Therefore

$$\begin{aligned}&E\left\{Var\left(\hat{\alpha}_f|\alpha\right)\right\}\\ &=(1-\hat{\alpha})^2\left\{Var\left(F_{v_2,v_1}\right)+\left(\frac{v_1}{v_1-2}\right)^2\right\} Var\left(F_{\tilde{v}_1,\tilde{v}_2}\right)\end{aligned}$$

and

$$\begin{aligned}Var\left\{E\left(\hat{\alpha}_f|\alpha\right)\right\} =& Var\left\{1-(1-\alpha)\left(\frac{\tilde{v}_2}{\tilde{v}_2-2}\right)\right\}\\ =& \left(\frac{\tilde{v}_2}{\tilde{v}_2-2}\right)^2(1-\hat{\alpha})^2 Var\left(F_{v_2,v_1}\right).\end{aligned}$$

From this it follows that

$$\begin{aligned}Var\left(\hat{\alpha}_f|\hat{\alpha}\right) =& (1-\hat{\alpha})^2\left\{Var\left(F_{v_2,v_1}\right)+\left(\frac{v_1}{v_1-2}\right)^2\right\} Var\left(F_{\tilde{v}_1,\tilde{v}_2}\right)\\ &+(1-\hat{\alpha})^2\left(\frac{\tilde{v}_2}{\tilde{v}_2-2}\right)^2 Var\left(F_{v_2,v_1}\right)\end{aligned}$$

D.4 MATLAB Code

```
%POSTERIOR DISTRIBUTION OF CRONBACH ALPHA
clear
m1=2451.25; m2=11271.5;
v1=24; v2=5;
K=5.664409273746257e+71;
%K=gamma((v1+v2)/2)/gamma(v1/2)/gamma(v2/2)*((v2*m2/v1/m1)^(v2/2))
ro=-0.2:0.0001:1;
pro=K*((1-ro).^((v2-2)/2)).*((1+v1*m1+v2*m2*(1-ro)).^(-(v1+v2)/2));
T=sum(pro)*0.0001;
C=0.0001*cumsum(pro);
lo95=ro(max(find(C<=0.025)))
up95=ro(max(find(C<=0.975)))
lo90=ro(max(find(C<=0.05)))
up90=ro(max(find(C<=0.95)))
a=pro(4339); b=pro(9863);
plot(ro,pro)
grid
%PREDICTIVE DISTRIBUTION FOR CRONBACH, GIVEN v1, v2, m1, m2
v1m1=58830; v2m2=56357;
v1=24; v2=5;
m1=v1m1/v1; m2=v2m2/v2;
N=1000;
alf=1-m1*frnd(v2,v1,N,1)/m2;
K=gamma((v1+v2)/2)/gamma(v1/2)/gamma(v2/2);
K2=K*((v1/v2)^(v1/2));
faf=[];
for af=-10:0.001:1;
fa1=((1-alf).^(-v1/2))*((1-af)^(v1/2-1));
fa2=(1+v1*(1-af)/v2./(1-alf)).^(-(v1+v2)/2);
fa=K2*fa1.*fa2;
faf=[faf fa];
end
Fa=sum(faf)/N;
af=-10:0.001:1;
plot(af,Fa)
```

```
grid
A=sum(Fa)*0.001
%RUNLENGTHS FOR DYESTUFF DATA
v1m1=58830; v2m2=56357;
v1=24; v2=5;
m1=v1m1/v1; m2=v2m2/v2;
lo=-0.048; up=0.963;
K=gamma((v1+v2)/2)/gamma(v1/2)/gamma(v2/2);
K2=K*((v1/v2)^(v1/2));
N=100000; Phi=[]; ERa=[]; VRa=[]; Fr=0;
for i=1:N
alf=1-m1*frnd(v2,v1)/m2;
faf=[];
for af=-1:0.001:1;
fa1=((1-alf).^(-v1/2))*((1-af)^(v1/2-1));
fa2=(1+v1*(1-af)/v2./(1-alf)).^(-(v1+v2)/2);
fa=K2*fa1.*fa2;
faf=[faf fa];
end
%Area of Predictive density < point953=0.05 & Area > point1964=0.05
phi=(sum(faf(1:953))+sum(faf(1964:2001)))*0.001;
Phi=[Phi;phi];
Era=(1-phi)/phi;
Vra=(1-phi)/(phi^2);
ERa=[ERa;Era];
VRa=[VRa;Vra];
r=0:400;
fr=((1-phi).^r)*phi;
Fr=Fr+fr;
end
figure(1)
plot(r,Fr/N)
grid
figure(2)
hist(Phi,30)
grid
figure(3)
```

```
hist(ERa,30)
grid
figure(4)
hist(VRa,30)
grid
%RUNLENGTHS
clear
check_predict
v1=80; v2=19;
v11=80; v22=19;
ahat=0.4952;
Me=[];
for j=1:6
    j
    lo=Lim(j,2); up=Lim(j,3);
P=[]; ERa=[]; VRa=[]; Fr=0;
for i=1:10000
a=1-(1-ahat)*frnd(v2,v1);
N=50000;
af=1-(1-a)*frnd(v11,v22,N,1);
L1=length(find(af<=lo));
L2=length(find(af>=up));
p=(L1+L2)/N;
P=[P;p];
Era=(1-p)/p;
Vra=(1-p)/(p^2);
ERa=[ERa;Era];
VRa=[VRa;Vra];
r=0:40000;
fr=((1-p).^r)*p;
Fr=Fr+fr;
end
cr=cumsum(Fr)/sum(Fr);
med=r(max(find(cr<=0.5)));
me=r*Fr'/10000;
Me=[Me;[Lim(j,1) me med]];
end
```

```
%POSTERIOR AND PREDICTIVE DISTRIBUTIONS
clear
v1=80; v2=19;
v11=80; v22=19;
N=1000000;
ahat=0.4952;
a=1-(1-ahat)*frnd(v2,v1,N,1);
af=1-(1-a).*frnd(v11,v22,N,1);
me=mean(af);
va=var(af);
saf=sort(af);
Lim=[];
for i=0.005:0.001:0.01
    lo=saf((round(N*i/2)));
    up=saf(round(N*(1-i/2)));
    Lim=[Lim;[i lo up]];
end
```

Appendix E: Fiducial and Bayesian Estimation for Cronbach's Alpha

E.1 Proof of Fiducial Distributions for $(\sigma_1, \sigma_2, \rho)$.

We are interested in the distribution of $f(s_{11}, s_{22}, s_{21}) = f(s_{11})\, f(s_{22}|s_{11})\, f(s_{12}|s_{11}, s_{22})$. According to Anderson (1958)

$$s_{11} \sim \sigma_1^2 \chi^2_{n-1} \text{ and } \quad \frac{s_{11}}{\sigma_1^2} \sim \chi^2_{n-1}.$$

Therefore

$$\sigma_1 = \sqrt{\frac{s_{11}}{\chi^2_{n-1}}}. \tag{E.1}$$

Using the notation $s_{22}|s_{11} = s_{22.1}$ (See (Anderson, 1958)) it follows that

$$\begin{aligned} s_{22.1} &= s_{22} - s_{21} s_{11}^{-1} s_{12} \\ &= s_{22}(1 - r^2) \end{aligned}$$

and

$$\begin{aligned} s_{22.1} &\sim \sigma_{22.1}\, \chi^2_{n-2} \\ &\sim \left(\sigma_2^2 - \sigma_{21}\sigma_1^{-2}\sigma_{12}\right) \chi^2_{n-2}. \end{aligned}$$

Therefore

$$s_{22.1} \sim \sigma_2^2 (1 - \rho^2)\, \chi^2_{n-2}$$

and

$$\frac{s_{22}(1 - r^2)}{\sigma_2^2 (1 - \rho^2)} \sim \chi^2_{n-2} \tag{E.2}$$

Accoring to Lemma 4.2 in Anderson (1958), s_{12} is independently distrubuted of s_{22} but depends on s_{11}. It can be shown that for given s_{11}

$$\frac{s_{12}}{s_{11}} \sim N\left[\frac{\rho\sigma_2}{\sigma_1}, \frac{\sigma_2^2(1-\rho^2)}{s_{11}}\right]$$

and

$$\frac{s_{12}}{\sqrt{s_{11}}} \sim N\left[\frac{\sqrt{s_{11}}\rho\sigma_2}{\sigma_1}, \sigma_2^2(1-\rho^2)\right]$$

which can be written as

$$\frac{s_{12}}{\sqrt{s_{11}}} = \sigma_2\sqrt{(1-\rho^2)}Z + \frac{\sqrt{s_{11}}\rho\sigma_2}{\sigma_1}$$

where $Z \sim N(0,1)$.

According to Equation E.2

$$\frac{r}{\sqrt{1-r^2}} = \frac{s_{12}/\sqrt{s_{11}}}{\sqrt{s_{22}(1-r^2)}} = \frac{\sigma_2\sqrt{1-\rho^2}Z + \frac{\sqrt{s_{11}}\rho\sigma_2}{\sigma_1}}{\sigma_2\sqrt{1-\rho^2}\sqrt{\chi_{n-2}^2}}$$

and according to Equation E.1

$$s_{11} \sim \sigma_1^2\chi_{n-1}^2.$$

Therefore,

$$\begin{aligned}\frac{r}{\sqrt{1-r^2}} &= \frac{\sigma_2\sqrt{(1-\rho^2)}Z + \rho\sigma_2\sqrt{\chi_{n-1}^2}}{\sigma_2\sqrt{1-\rho^2}\sqrt{\chi_{n-2}^2}} \\ &= \frac{Z}{\sqrt{\chi_{n-2}^2}} + \frac{\rho}{\sqrt{1-\rho^2}}\frac{\sqrt{\chi_{n-1}^2}}{\sqrt{\chi_{n-2}^2}}. \qquad \text{(E.3)}\end{aligned}$$

See Berger & Sun (2006) page 980. From Equation E.3 it follows that $\rho = \psi(Y)$ where

$$Y = \frac{-Z}{\sqrt{\chi_{n-1}^2}} + \frac{\sqrt{\chi_{n-2}^2}}{\sqrt{\chi_{n-1}^2}}\frac{r}{\sqrt{1-r^2}} \qquad \text{(E.4)}$$

and $\psi(X) = \frac{X}{\sqrt{1+X^2}}$. Since $\frac{s_{22}(1-r^2)}{\sigma_2^2(1-\rho^2)} \sim \chi_{n-2}^2$ it follows that

$$\sigma_2^2 = \frac{s_{22}(1-r^2)}{\chi_{n-2}^2}\frac{1}{(1-\rho^2)}.$$

Also $\rho = \psi(Y)$ and $1-\rho^2 = 1-\psi^2(Y)$. Therefore

$$\frac{1}{1-\rho^2} = \frac{1}{1-\psi^2(Y)}$$

and

$$\frac{1}{1-\rho^2} = \frac{1}{1-\frac{Y^2}{1+Y^2}} = 1+Y^2.$$

From this it follows that

$$\begin{aligned}\sigma_2^2 &= \frac{s_{22}(1-r^2)}{\chi^2_{n-2}}(1+Y^2)\\ &= s_{22}(1-r^2)\left\{\frac{1}{\chi^2_{n-2}}+\frac{Y^2}{\chi^2_{n-2}}\right\}\end{aligned} \quad \text{(E.5)}$$

If $Y = \frac{-Z}{\sqrt{\chi^2_{n-1}}} + \frac{\sqrt{\chi^2_{n-2}}}{\sqrt{\chi^2_{n-1}}}\frac{r}{\sqrt{1-r^2}}$ is substituted into Equation E.5, it follows that

$$\begin{aligned}\sigma_2^2 &= s_{22}\left(1-r^2\right)\left\{\frac{1}{\chi^2_{n-2}}+\frac{1}{\chi^2_{n-2}}\left(\frac{Z}{\sqrt{\chi^2_{n-1}}}-\frac{\sqrt{\chi^2_{n-2}}}{\sqrt{\chi^2_{n-1}}}\frac{r}{\sqrt{1-r^2}}\right)^2\right\}\\ &= s_{22}\left(1-r^2\right)\left\{\frac{1}{\chi^2_{n-2}}+\frac{1}{\chi^2_{n-1}}\left(\frac{Z}{\sqrt{\chi^2_{n-2}}}-\frac{\sqrt{\chi^2_{n-2}}}{\sqrt{\chi^2_{n-2}}}\frac{r}{\sqrt{1-r^2}}\right)^2\right\}\\ &= s_{22}\left(1-r^2\right)\left\{\frac{1}{\chi^2_{n-2}}+\frac{1}{\chi^2_{n-1}}\left(\frac{Z}{\sqrt{\chi^2_{n-2}}}-\frac{r}{\sqrt{1-r^2}}\right)^2\right\}\end{aligned}$$

and

$$\sigma_2 = \sqrt{s_{22}\left(1-r^2\right)}\sqrt{\frac{1}{\chi^2_{n-2}}+\frac{1}{\chi^2_{n-1}}\left(\frac{Z}{\sqrt{\chi^2_{n-2}}}-\frac{r}{\sqrt{1-r^2}}\right)^2}. \quad \text{(E.6)}$$

Equations E.1, E.6 and E.4 correspond to Equations 7.8, 7.9 and 7.10.

E.2 Proof of Posterior distributions for $(\sigma_1, \sigma_2, \rho)$.

It is easier to use the parameters (η_1, η_2, η_3) instead of σ_1, σ_2 and ρ where

$$\eta_1 = \frac{1}{\sigma_1}, \qquad \eta_2 = \frac{1}{\sigma_2\sqrt{1-\rho^2}} \text{ and } \qquad \eta_3 = \frac{-\rho}{\sigma_1\sqrt{1-\rho^2}}. \quad \text{(E.7)}$$

From Equation E.7 it follows that

$$\sigma_1 = \frac{1}{\eta_1}, \qquad \sigma_2 = \frac{\sqrt{\eta_1^2+\eta_3^2}}{\eta_1\eta_2} \text{ and } \qquad \rho = \frac{-\eta_3}{\sqrt{\eta_1^2+\eta_3^2}}$$

and Σ^{-1} can be written as

$$\Sigma^{-1} = \begin{bmatrix} \eta_1 & \eta_3 \\ 0 & \eta_2 \end{bmatrix} \begin{bmatrix} \eta_1 & 0 \\ \eta_3 & \eta_2 \end{bmatrix}.$$

Since the Jacobian between the two parameterizations are

$$J = \left| \frac{\partial(\mu_1,\mu_2,\sigma_1,\sigma_2,\rho)}{\partial(\mu_1,\mu_2,\eta_1,\eta_2,\eta_3)} \right| = \frac{1}{\eta_1\eta_2^2(\eta_1+\eta_3^2)},$$

the prior

$$\pi_{ab}(\mu_1,\mu_2,\eta_1,\eta_2,\eta_3) = \frac{1}{\eta_1^a\eta_2^b}. \tag{E.8}$$

Under the prior defined in Equation E.8, the marginal posterior density function of (η_1,η_2,η_3) is

$$\pi(\eta_1,\eta_2,\eta_3|data) \propto \eta_1^{n-a-1}\eta_2^{n-b-1}\exp\left\{-\frac{1}{2}\left[\eta_1^2 s_{11} + \eta_2^2 s_{22}(1-r^2) + s_{11}\left(\eta_3 + \eta_2 r\sqrt{\frac{s_{22}}{s_{11}}}\right)^2\right]\right\}. \tag{E.9}$$

From Equation E.9 it follows that

- The marginal posterior distribution of η_3 given η_1,η_2 is $N\left(-\eta_2 r\sqrt{\frac{s_{22}}{s_{11}}}, \frac{1}{s_{11}}\right)$.
- The marginal posterior distributions of η_1 and η_2 are independent,

$$\eta_1^2|data \sim Gamma\left(\frac{1}{2}(n-a), \frac{s_{11}}{2}\right)$$

and

$$\eta_2^2|data \sim Gamma\left(\frac{1}{2}(n-b), \frac{s_{22}(1-r^2)}{2}\right).$$

The constructive posterior distributions of η_1, η_2 and η_3 given the data can therefore be written as

$$\eta_1 = \sqrt{\frac{\chi^2_{n-a}}{s_{11}}}, \tag{E.10}$$

$$\eta_2 = \sqrt{\frac{\chi^2_{n-b}}{s_{22}(1-r^2)}} \tag{E.11}$$

and

$$\eta_3 = \frac{z_3}{\sqrt{s_{11}}} - \eta_2 r \sqrt{\frac{s_{22}}{s_{11}}}$$
$$= \frac{z_3}{\sqrt{s_{11}}} - \sqrt{\frac{\chi^2_{n-b}}{s_{22}}} \frac{r}{\sqrt{1-r^2}}. \tag{E.12}$$

Since $\eta_1 = \frac{1}{\sigma_1}$, $\eta_2 = \frac{1}{\sigma_2\sqrt{1-\rho^2}}$ and $\eta_3 = \frac{-\rho}{\sigma_1\sqrt{1-\rho^2}}$, Equations 7.12, 7.13 and 7.14 follow.

E.3 MATLAB Code

```
%CRONBACH ALPHA
clc
clear
tic
mu=[10;12];
sig=[1 1;1 4];
n=5; % Change the sample size n accordingly
p=2;
a=1; b=0; % Jeffreys Prior
%a=1; b=2; % Right-Haar Prior
%a=2; b=1; % Independence Jeffreys Prior
%a=1; b=1; % pi_Ro
tcron=0; MC=[];MM=[]; Ii=[];
for j=1:10000
X=mvnrnd(mu,sig,n);
mx=mean(X)';
S=(n-1)*cov(X);
r=S(1,2)/sqrt(S(1,1)*S(2,2));
N=10000;
RHO=sig(1,2)/sqrt((sig(1,1)*sig(2,2)));
CRON=2*(1-(sig(1,1)+sig(2,2))/(sig(1,1)+sig(2,2)+2*RHO*sqrt(sig(1,1))*sqrt(sig(2,2))));
Cron=[];
for i=1:N
chi1=chi2rnd(n-a);
chi2=chi2rnd(n-b);
Z3=normrnd(0,1);
s1=sqrt(S(1,1)/chi1);
```

```
A=1/chi2+(1/chi1)*((Z3/sqrt(chi2)-r/sqrt(1-r^2))^2);
s2=sqrt(S(2,2)*(1-r^2)*A);
Y=-Z3/sqrt(chi1)+sqrt(chi2/chi1)*r/sqrt(1-r^2);
ro=Y/sqrt(1+Y^2); z=rand(1,1);
Use one of these if statements to simulate using the following priors:
%if z<sqrt(1-ro^2);  Pi_R_rho
%end
%if z<sqrt(1-ro^4);  Pi_R_sigma
%end
%if z<(1-ro^2);      Pi_MS
%end
cron=(p/(p-1))*(1-(s1^2+s2^2)/(s1^2+s2^2+2*ro*s1*s2));
Cron=[Cron;cron];
end
mc=mean(Cron);
MC=[MC;mc];
scron=sort(Cron);
M=length(scron);
lower=scron((0.025*M));
upper=scron((0.975*M));
leng=upper-lower;
Ii=[Ii leng];
MM=[MM;M];
if scron(round(0.025*M))<CRON & scron(round(0.975*M))>CRON
tcron=tcron+1;
end
end
toc
Coverage=tcron/M
Interval_Length=mean(Ii)
```

Appendix F: Estimating Cronbach's α and ρ under Compound Symmetry

F.1 Proof of Theorem 8.1.

We are interested in the density function of $(\hat{\alpha}-1)\,F_{(I-1),I(J-1)}+1$. Let $f \sim F_{(I-1),I(J-1)}$, then

$$\pi(f) = K f^{\frac{1}{2}(I-1)-1}\left[1+\frac{(I-1)}{I(J-1)}f\right]^{-\frac{1}{2}(IJ-1)}$$

and $K = \frac{\Gamma\left(\frac{(IJ-1)}{2}\right)}{\Gamma\left(\frac{(I-1)}{2}\right)\Gamma\left[\frac{I(J-1)}{2}\right]}(I-1)^{\frac{(I-1)}{2}}\left[I(J-1)\right]^{-\frac{(I-1)}{2}}$. To obtain the density function for $y=(\hat{\alpha}-1)\,f$, we must first get the Jacobian. Now, $f=\frac{y}{(\hat{\alpha}-1)}$ and $\left|\frac{df}{dy}\right| = \frac{1}{1-\hat{\alpha}}$. Therefore

$$\pi(y|\hat{\alpha}) = K\left(\frac{y}{(\hat{\alpha}-1)}\right)^{\frac{1}{2}(I-1)-1}\left[1+\frac{(I-1)}{I(J-1)}\left(\frac{y}{(\hat{\alpha}-1)}\right)\right]^{-\frac{1}{2}(IJ-1)}\frac{1}{1-\hat{\alpha}}.$$

Let $\alpha = y+1$, then $y=\alpha-1$ and $\left|\frac{dy}{d\alpha}\right| = 1$. Therefore,

$$\pi(\alpha|\hat{\alpha}) = K\left(\frac{1}{1-\hat{\alpha}}\right)\left(\frac{1-\alpha}{1-\hat{\alpha}}\right)^{\frac{(I-3)}{2}}\left[1+\frac{(I-1)(1-\alpha)}{I(J-1)(1-\hat{\alpha})}\right]^{-\frac{1}{2}(IJ-1)}.$$

F.2 Proof of Theorem 8.3.

Now $\alpha = \frac{J\rho}{1+(J-1)\rho}$, $1-\alpha = \frac{1-\rho}{1+\rho(J-1)}$ and $\rho = \frac{\alpha}{J-\rho(J-1)}$. The Jacobian is

$$\frac{d\alpha}{d\rho} = \frac{J\left[1+\rho(J-1)\right]-J(J-1)\rho}{\left[1+(J-1)\rho\right]^2} = \frac{J}{\left[1+(J-1)\rho\right]^2}.$$

Therefore,

$$\begin{aligned}\pi(\rho|\hat{\alpha}) &= K\left[\frac{1-\rho}{1+\rho(J-1)}\right]^{\frac{1}{2}(I-3)}\left[1+\frac{(I-1)}{I(J-1)}\frac{1}{(1-\hat{\alpha})}\frac{(1-\rho)}{[1+\rho(J-1)]}\right]^{-\frac{1}{2}(IJ-1)} \\ &\times \frac{J}{[1+\rho(J-1)]^2}\left(\frac{1}{1-\hat{\alpha}}\right)^{\frac{1}{2}(I-1)}.\end{aligned}$$

F.3 Proof of Theorem 8.4.

$$\Omega=\frac{1-\alpha_1}{1-\alpha_2}=\frac{(1-\hat{\alpha}_1)F_{I_2(J_2-1),(I_2-1)}}{(1-\hat{\alpha}_2)F_{I_1(J_1-1),(I_1-1)}}. \tag{F.1}$$

$$\begin{aligned}E(\Omega|\hat{\alpha}_1,\hat{\alpha}_2) &= \frac{(1-\hat{\alpha}_1)}{(1-\hat{\alpha}_2)}E\left\{\frac{F_{I_2(J_2-1),(I_2-1)}}{F_{I_1(J_1-1),(I_1-1)}}\right\} \\ &= \frac{(1-\hat{\alpha}_1)}{(1-\hat{\alpha}_2)}E\left\{F_{I_2(J_2-1),(I_2-1)}\right\}E\left\{F_{(I_1-1),I_1(J_1-1)}\right\} \\ &= \frac{(1-\hat{\alpha}_1)}{(1-\hat{\alpha}_2)}\left(\frac{I_2-1}{I_2-3}\right)\left(\frac{I_1(J_1-1)}{I_1(J_1-1)-2}\right)\end{aligned}$$

and

$$Var(\Omega|\hat{\alpha}_1,\hat{\alpha}_2)=Var\left\{E\left(\Omega|F_{I_2(J_2-1),(I_2-1)}\right)\right\}+E\left\{Var\left(\Omega|F_{I_2(J_2-1),(I_2-1)}\right)\right\}$$

where

$$E\left(\Omega|F_{I_2(J_2-1),(I_2-1)}\right)=\frac{(1-\hat{\alpha}_1)}{(1-\hat{\alpha}_2)}F_{I_2(J_2-1),(I_2-1)}\left(\frac{I_1(J_1-1)}{I_1(J_1-1)-2}\right)$$

and

$$\begin{aligned}Var\left\{E\left(\Omega|F_{I_2(J_2-1),(I_2-1)}\right)\right\} &= \frac{(1-\hat{\alpha}_1)^2}{(1-\hat{\alpha}_2)^2}\left\{\frac{I_1(J_1-1)}{I_1(J_1-1)-2}\right\}^2 Var\left[F_{I_2(J_2-1),(I_2-1)}\right] \\ &= \frac{(1-\hat{\alpha}_1)^2}{(1-\hat{\alpha}_2)^2}\left\{\frac{I_1(J_1-1)}{I_1(J_1-1)-2}\right\}^2\frac{2(I_2-1)^2\{I_2(J_2-1)+(I_2-3)\}}{I_2(J_2-1)(I_2-3)^2(I_2-5)}.\end{aligned}$$

Also,

$$\begin{aligned}&Var\left(\Omega|F_{I_2(J_2-1),(I_2-1)}\right) \\ =&\frac{(1-\hat{\alpha}_1)^2}{(1-\hat{\alpha}_2)^2}\left[F_{I_2(J_2-1),(I_2-1)}\right]^2 Var\left(F_{(I_1-1),I_1(J_1-1)}\right)\end{aligned}$$

and

$$E\left\{Var\left(\Omega|F_{I_2(J_2-1),(I_2-1)}\right)\right\} = \frac{(1-\hat{\alpha}_1)^2}{(1-\hat{\alpha}_2)^2}\left\{\frac{2\left\{I_1\left(J_1-1\right)\right\}^2\left\{(I_1J_1-3)\right\}}{(I_1-1)\left\{I_1\left(J_1-1\right)-2\right\}^2\left\{I_1\left(J_1-1\right)-4\right\}}\right\} \times\left\{\frac{2\left(I_2-1\right)^2\left\{I_2\left(J_2-1\right)+I_2-3\right\}}{I_2\left(J_2-1\right)\left(I_2-3\right)^2\left(I_2-5\right)}+\left(\frac{I_2-1}{I_2-3}\right)^2\right\}$$

F.4 MATLAB Code

```
% Simulate from two populations Cronbach's alpha
clc
clear
I_1=100;
J_1=4;
I_2=100;
J_2=6;
v_11=(I_1)*(J_1-1);
v_21=I_1-1;
v_12=(I_2)*(J_2-1);
v_22=I_2-1;
alpha_hat1=0.7;
alpha_hat2=0.8;
% Results for the Differences in Cronbach's alpha, Delta
n=100000; % Number of iterations
% Generate 100 0000 F values
F_1=frnd(v_11,v_21,1,n);
F_2=frnd(v_12,v_22,1,n);
alpha_1=1+(alpha_hat1-1)./F_1;
alpha_2=1+(alpha_hat2-1)./F_2;
delta=alpha_2./alpha_1;
MEAN=mean(delta)
temp=sort(delta);
MEDIAN=temp(0.5*n)
VARIANCE=mean(delta.^2)-MEAN^2
CI_95=[temp(0.025*n) temp(0.975*n)]
P_1=sum(delta>1)/n
% Results for the Differences in Cronbach's alpha, Gamma
gamma=alpha_2-alpha_1;
```

```
MEAN_2=mean(gamma)
temp_2=sort(gamma);
MEDIAN_2=temp_2(0.5*n)
VARIANCE_2=mean(gamma.^2)-MEAN_2^2
CI_95=[temp_2(0.025*n) temp_2(0.975*n)]
P_2=sum(gamma>0)/n
% Results for Omega
omega=(1-alpha_hat1).*F_2./((1-alpha_hat2).*F_1);
MEAN_3=mean(omega)
temp_3=sort(omega);
MEDIAN_3=temp_3(0.5*n)
VARIANCE_3=mean(omega.^2)-MEAN_3^2
CI_95=[temp_3(0.025*n) temp_3(0.975*n)]
P_3=sum(omega>1)/n
figure
hist(delta,60)
title('The Distribution of $$\delta_\alpha=\frac{\alpha_2}{\alpha_1}$$',
'Interpreter','latex','FontSize',26)
xlabel('$\delta_\alpha$','Interpreter','latex','FontSize',26)
set(gca,'fontsize',24)
figure
hist(gamma,60)
title('The Distribution of $$\gamma_\alpha=\alpha_2-\alpha_1$$',
'Interpreter','latex','FontSize',26)
xlabel('$\gamma_\alpha$','Interpreter','latex','FontSize',26)
set(gca,'fontsize',24)
figure
hist(omega,60)
title('The Density Function $$\pi(\Omega| \hat{\alpha}_1,\hat{\alpha}_2)$$',
'Interpreter','latex','FontSize',26)
xlabel('$\Omega$','Interpreter','latex','FontSize',26)
set(gca,'fontsize',24)
% Intra-Class Correlation Coefficient
% 1st Population
rho_1=alpha_1./(J_1-alpha_1.*(J_1-1));
MEAN_4=mean(rho_1)
temp_4=sort(rho_1);
```

```
MEDIAN_4=temp_4(0.5*n)
VARIANCE_4=mean(rho_1.^2)-MEAN_4^2
CI_95=[temp_4(0.025*n) temp_4(0.975*n)]
% 2nd Population
rho_2=alpha_2./(J_2-alpha_2.*(J_2-1));
MEAN_5=mean(rho_2)
temp_5=sort(rho_2);
MEDIAN_5=temp_5(0.5*n)
VARIANCE_5=mean(rho_2.^2)-MEAN_5^2
CI_95=[temp_5(0.025*n) temp_5(0.975*n)]
% Ratio of Correlations
delta_r=rho_2./rho_1;
MEAN_6=mean(delta_r)
temp_6=sort(delta_r);
MEDIAN_6=temp_6(0.5*n)
VARIANCE_6=mean(delta_r.^2)-MEAN_6^2
CI_95=[temp_6(0.025*n) temp_6(0.975*n)]
P_1=sum(delta_r>1)/n
% Differences of Correlations
gamma_r=rho_2-rho_1;
MEAN_7=mean(gamma_r)
temp_7=sort(gamma_r);
MEDIAN_7=temp_7(0.5*n)
VARIANCE_7=mean(gamma_r.^2)-MEAN_7^2
CI_95=[temp_7(0.025*n) temp_7(0.975*n)]
P_2=sum(gamma_r>0)/n
figure
hist(delta_r,60)
title('The Distribution of $$\delta_\rho=\frac{\rho_2}{\rho_1}$$','Interpreter'
,'latex','FontSize',26)
xlabel('$\delta_\rho$','Interpreter','latex','FontSize',26)
set(gca,'fontsize',24)
figure
hist(gamma_r,60)
title('The Distribution of $$\gamma_\rho=\rho_2-\rho_1$$','Interpreter'
,'latex','FontSize',26)
xlabel('$\gamma_\rho$','Interpreter','latex','FontSize',26)
```

```
set(gca,'fontsize',24)
% Posterior for Cronbach's alpha-Compound Symmetry
% Exact Inference for Cronbach's alpha using closed
% form results and Plots of Posteriors for
% Cronbach's alpha-Compound Symmetry
clc
clear
I_1=100;
I_2=100;
J_1=4;
J_2=6;
v_11=(I_1)*(J_1-1);
v_21=I_1-1;
v_12=(I_2)*(J_2-1);
v_22=I_2-1;
alpha_hat_1=0.7;
alpha_hat_2=0.8;
% Posterior for Alpha_1
% Function to determine normalizing constant
fun1= @(alpha) (1/(1-alpha_hat_1)).*((1-alpha)./(1-alpha_hat_1)).^(0.5*(I_1-3))
.*(1+(I_1-1).*(1-alpha)./(I_1.*(J_1-1).*(1-alpha_hat_1))).^(-0.5*(I_1*J_1-1));
q1 = integral(fun1,0,1);
% Therefore normalizing constant K=1/q1
K=1/q1;
%Function to determine mean of posterior of alpha
fun2= @(alpha) (1/q1)*alpha.*(1/(1-alpha_hat_1)).*((1-alpha)./(1-alpha_hat_1))
.^(0.5*(I_1-3)).*(1+(I_1-1).*(1-alpha)./(I_1.*(J_1-1)
.*(1-alpha_hat_1))).^(-0.5*(I_1*J_1-1));
% Bayes estimate under squared error loss
alpha_hat_SE=integral(fun2,0,1)
% Posterior for Alpha_2
fun3= @(alpha) (1/(1-alpha_hat_2)).*((1-alpha)./(1-alpha_hat_2)).^(0.5*(I_2-3))
.*(1+(I_2-1).*(1-alpha)./(I_2.*(J_2-1).*(1-alpha_hat_2))).^(-0.5*(I_2*J_2-1));
q2 = integral(fun3,0,1);
%Function to determine mean of posterior of alpha
fun3= @(alpha) (1/q2)*alpha.*(1/(1-alpha_hat_2)).*((1-alpha)./(1-alpha_hat_2))
.^(0.5*(I_2-3)).*(1+(I_2-1).*(1-alpha)./(I_2.*(J_2-1).*(1-alpha_hat_2)))
```

```
.^(-0.5*(I_2*J_2-1));
alpha_hat_SE_2=integral(fun3,0,1)

%Plot of the  posterior for alpha
x=[0:0.01:1];
y_1=(1/q1).*(1/(1-alpha_hat_1)).*((1-x)./(1-alpha_hat_1)).^(0.5*(I_1-3))
.*(1+(I_1-1).*(1-x)./(I_1.*(J_1-1).*(1-alpha_hat_1))).^(-0.5*(I_1*J_1-1));
y_2=(1/q2).*(1/(1-alpha_hat_2)).*((1-x)./(1-alpha_hat_2)).^(0.5*(I_2-3))
.*(1+(I_2-1).*(1-x)./(I_2.*(J_2-1).*(1-alpha_hat_2))).^(-0.5*(I_2*J_2-1));
plot(x,y_1,x,y_2,':','LineWidth',3)
xlim([0.4 1])
%ylim([0 10])
title('Density Functions of $$\pi(\alpha_1|\hat{\alpha}_1)$$ and
$$\pi(\alpha_2|\hat{\alpha}_2)$$ of Cronbachs Alpha','Interpreter',
'latex','FontSize',26)
xlabel('\alpha','FontSize',26)
ylabel('\pi(\alpha)','FontSize',30)
set(gca,'fontsize',24) % Setting the fontsize of the values on the axes
% Mode of alpha_1
Mode_alpha_1=1-(I_1*(I_1-3)*(J_1-1)*(1-alpha_hat_1))/((I_1*J_1-1)
*(I_1-1)-(I_1-3)*(I_1-1))
% Mode of alpha_2
Mode_alpha_2=1-(I_2*(I_2-3)*(J_2-1)*(1-alpha_hat_2))/((I_2*J_2-1)
*(I_2-1)-(I_2-3)*(I_2-1))
% Variance of alpha_1
Variance_alpha_1=(alpha_hat_1-1)^(2)*2*I_1^2*(J_1-1)^2*(I_1*J_1-3)
/((I_1-1)*(I_1*(J_1-1)-2)^2*(I_1*(J_1-1)-4))
% Variance of alpha_2
Variance_alpha_2=(alpha_hat_2-1)^(2)*2*I_2^2*(J_2-1)^2*(I_2*J_2-3)
/((I_2-1)*(I_2*(J_2-1)-2)^2*(I_2*(J_2-1)-4))
% 95% Interval of alpha_1
CI_95_alpha_1=[1-(1-alpha_hat_1)/finv(0.025,v_11,v_21)
1-(1-alpha_hat_1)/finv(0.975,v_11,v_21)]
% 95% Interval of alpha_2
CI_95_alpha_2=[1-(1-alpha_hat_2)/finv(0.025,v_12,v_22)
1-(1-alpha_hat_2)/finv(0.975,v_12,v_22)]
% Posterior for Intra-Class Correlation-Compound Symmetry
```

```
%Exact Inference for the Intra-class correlation coefficient
%using some of the closed form results and numerical integration
clc
clear
I_1=100;
I_2=100
J_1=4;
J_2=6
v_11=(I_1)*(J_1-1);
v_21=I_1-1;
v_12=(I_2)*(J_2-1);
v_22=I_2-1;
alpha_hat_1=0.7
alpha_hat_2=0.8
fun4= @(rho) ((1-rho)./(1+rho.*(J_1-1))).^(0.5.*(I_1-3)).*(1+(I_1-1).*(1-rho)
./(I_1.*((J_1-1)).*(1-alpha_hat_1).*(1+rho.*(J_1-1)))).^(-0.5.*(I_1.*J_1-1))
.*J_1./(1+rho*(J_1-1)).^2.*(1-alpha_hat_1).^(-0.5*(I_1-1));
q4 = integral(fun4,0,1);
% Posterior for rho_1
% Therefore normalizing constant K=1/q4
K_2=1/q4;
% Function to determine the mean of posterior for rho_1
fun5= @(rho) K_2.*rho.*((1-rho)./(1+rho.*(J_1-1))).^(0.5.*(I_1-3)).*(1+(I_1-1)
.*(1-rho)./(I_1.*((J_1-1)).*(1-alpha_hat_1).*(1+rho.*(J_1-1))))
.^(-0.5.*(I_1.*J_1-1)).*J_1./(1+rho*(J_1-1)).^2.*(1-alpha_hat_1).^(-0.5*(I_1-1));
MEAN_RHO_1=integral(fun5,0,1)
% Posterior for rho_2
fun6= @(rho) ((1-rho)./(1+rho.*(J_2-1))).^(0.5.*(I_2-3)).*(1+(I_2-1)
.*(1-rho)./(I_2.*((J_2-1)).*(1-alpha_hat_2).*(1+rho.*(J_2-1))))
.^(-0.5.*(I_2.*J_2-1)).*J_2./(1+rho*(J_2-1)).^2.*(1-alpha_hat_2).^(-0.5*(I_2-1));
q5 = integral(fun6,0,1);
% Therefore normalizing constant K=1/q5
K_3=1/q5;
fun7= @(rho) K_3.*rho.*((1-rho)./(1+rho.*(J_2-1))).^(0.5.*(I_2-3)).*(1+(I_2-1)
.*(1-rho)./(I_2.*((J_2-1)).*(1-alpha_hat_2).*(1+rho.*(J_2-1))))
.^(-0.5.*(I_2.*J_2-1)).*J_2./(1+rho*(J_2-1)).^2.*(1-alpha_hat_2)
.^(-0.5*(I_2-1));
```

```
MEAN_RHO_2=integral(fun7,0,1)
%Plot of the  posterior for rho
x=[0:0.01:1];
y_3=K_2.*((1-x)./(1+x.*(J_1-1))).^(0.5.*(I_1-3)).*(1+(1-x)./((J_1-1)
.*(1-alpha_hat_1).*(1+x.*(J_1-1)))).^(-0.5.*(I_1-1).*J_1).*J_1
./(1+x*(J_1-1)).^2.*(1-alpha_hat_1).^(-0.5*(I_1-1));
y_4=K_3.*((1-x)./(1+x.*(J_2-1))).^(0.5.*(I_2-3)).*(1+(1-x)./((J_2-1)
.*(1-alpha_hat_2).*(1+x.*(J_2-1)))).^(-0.5.*(I_2-1).*J_2).*J_2
./(1+x*(J_2-1)).^2.*(1-alpha_hat_2).^(-0.5*(I_2-1));
plot(x,y_3,x,y_4,':','LineWidth',3)
xlim([0 0.7])
%ylim([0 10])
title('Density Functions of $$\pi(\rho_1|\hat{\alpha}_1)$$
and $$\pi(\rho_2|\hat{\alpha}_2)$$ of the Intra-Class
Correlation','Interpreter','latex','FontSize',26)
xlabel('\rho','FontSize',26)
ylabel('\pi(\rho)','FontSize',30)
set(gca,'fontsize',24) % Setting the fontsize of the
values on the axes
%Variance of rho_1
fun8= @(rho) K_2.*rho.^2.*((1-rho)./(1+rho.*(J_1-1))).^(0.5.*(I_1-3))
.*(1+(I_1-1).*(1-rho)./(I_1.*((J_1-1)).*(1-alpha_hat_1).*(1+rho.*(J_1-1))))
.^(-0.5.*(I_1.*J_1-1)).*J_1./(1+rho*(J_1-1)).^2.*(1-alpha_hat_1).^(-0.5*(I_1-1));
E_RHO__SQ_1=integral(fun8,0,1)
V_RHO_1=E_RHO__SQ_1-(MEAN_RHO_1)^2
%Variance of rho_2
fun9= @(rho) K_3.*rho.^2.*((1-rho)./(1+rho.*(J_2-1))).^(0.5.*(I_2-3)).*(1+(I_2-1)
.*(1-rho)./(I_2.*((J_2-1)).*(1-alpha_hat_2).*(1+rho.*(J_2-1)))).^(-0.5
.*(I_2.*J_2-1))
.*J_2./(1+rho*(J_2-1)).^2.*(1-alpha_hat_2).^(-0.5*(I_2-1));
E_RHO__SQ_2=integral(fun9,0,1)
V_RHO_2=E_RHO__SQ_2-(MEAN_RHO_2)^2
% 95% Interval of rho_1
CI_95_rho_1=[((J_1*finv(0.025,v_11,v_12)/(finv(0.025,v_11,v_12)-(1-alpha_hat_1)))
-(J_1-1))^(-1)
((J_1*finv(0.975,v_11,v_12)/(finv(0.975,v_11,v_12)-(1-alpha_hat_1)))-(J_1-1))^(-1)]
%Median of rho_1
```

```
median_rho_1=((J_1*finv(0.5,v_11,v_12)/(finv(0.5,v_11,v_12)-(1-alpha_hat_1)))
-(J_1-1))^(-1)
% 95% Interval of rho_2
CI_95_rho_2=[((J_2*finv(0.025,v_21,v_22)/(finv(0.025,v_21,v_22)-(1-alpha_hat_2)))
-(J_2-1))^(-1)
((J_2*finv(0.975,v_21,v_22)/(finv(0.975,v_21,v_22)-(1-alpha_hat_2)))-(J_2-1))^(-1)]
%Median of rho_2
median_rho_2=((J_2*finv(0.5,v_21,v_22)/(finv(0.5,v_21,v_22)-(1-alpha_hat_2)))
-(J_2-1))^(-1)
```