CONTRIBUTIONS TO ESTIMATION OF MEASURES FOR ASSESSING RATER RELIABILITY

A Dissertation
Submitted to
the Temple University Graduate Board

---

in Partial Fulfillment
of the Requirements for the Degree of
DOCTOR OF PHILOSOPHY

---

by
Luqiang Wang
April, 2009

ABSTRACT

CONTRIBUTIONS TO ESTIMATION OF MEASURES FOR ASSESSING RATER RELIABILITY

Luqiang Wang
DOCTOR OF PHILOSOPHY

Temple University, April, 2009

Professor Burt Holland, Chair

Reliability measures have been well studied over many years, beginning with an entire chapter devoted to intraclass correlation in the first edition of Fisher (1925). Such measures have been thoroughly studied for two factor models. This dissertation, motivated by a medical research problem, extends point and confidence interval estimation of both intraclass correlation coefficient and interater reliability coefficient to models containing three crossed random factors — subjects, raters and occasions. The intraclass correlation coefficient is used when decision is made on an absolute basis with rater's scores, while the interater reliability coefficient is defined for decisions made on a relative basis. The estimation is conducted using both ANOVA and MCMC methods. The results from the two methods are compared. The MCMC method is preferred for analyses of small data sets when ICC values are high. Besides, the bias of estimator of intraclass correlation coefficient in one-way random effects model is evaluated.

ACKNOWLEDGEMENTS

I would like to thank the members of my dissertation committee. Their guidance and assistance ensured the quality of this dissertation and have been vital to my training to be a PhD. Without their help, this dissertation would not have been possible. I can't express my gratitude enough to them.

I would also like to take this opportunity to thank my wife. Her support is indispensable.

Further thanks should also go to my sons — James and Michael. They are my source of joy and drive. Whatever anxiety and tiredness dissolve when I see their smiles and funny behaviors. They will be happy seeing their names appearing here when they can read it.

Finally I would like to thank Dr. Mason, my former professor at the University of Delaware, and Dr. Stanley, my supervisor of a internship at DuPont. Their encouragement sent me onto the PhD journey and went a long way. "You will never regret having a PhD degree"— I could't agree more on this now.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

CHAPTER 1

INTRODUCTION

1.1 Introduction

Reliability measures have been well studied over many years, beginning with an entire chapter devoted to intraclass correlation in the first edition of the classic text by Fisher (1925). Class relates to one of the factors or categories in the study, and intraclass refers to comparing the measurements within classes. Shrout and Fleiss (1979) provided a comprehensive overview of estimation of the intraclass correlation covering scenarios involving two random or mixed factors with or without interaction. These authors referred to the two factors as targets (objects of measurement, subjects or patients) and judges (raters or observers). More recent surveys of this topic were produced by McGraw and Wong (1996), and Dunn (2004). Their presentations, still restricted to two factors or just one factor, focused on construction of frequentist confidence intervals and hypothesis tests.

Measures of reliability are used in much of science, particularly the social science areas of psychology and sociology and the natural science areas including medicine and biology. The present research is motivated by a problem that arose in medical research and will be described in more detail in Chapter 5. It was desired to evaluate two distinct types of consistency among the raters in an experiment designed to assess each subject on more than one occasions.

This multi-occasion scenario required an extension to models involving three crossed factors: subjects, raters, and occasions. Potential interaction among pairs of these factors is allowed. In order to extend results to well beyond this experiment's subjects, raters and occasions, the experiment was designed so that the further assumption that all factors are random is justifiable. This means that the subjects, raters and occasions had no qualities that would appreciably differ in a replication of the experiment in another setting.

Intraclass correlation coefficient (ICC) assesses the consistency of the raters across their multiple measurements on different occasions on each subject, it is in essence an index of dependability. Interrater reliability coefficient (IRC) is in essence a generalization coefficient. ICC is used when decision is made on an absolute basis with rater's scores, while IRC is defined for decisions made on a relative basis (Shavelson and Webb, 1991).

In Chapter 2, I shall discuss ICCs for one-way and two-way random effects models. Included is work on approximating bias for ICC estimation in the one-way random effects model. This is groundwork for novel work on three-way random effects models, incorporating all two factor interactions, discussed in Chapter 3. Chapter 4 examines a reduced three-way model that assumes one of the two-way interactions is negligible. In Chapter 5, I offer data analysis and simulation to demonstrate the quality of the new three-way estimation procedures. Comparison is made between confidence intervals from ANOVA-based procedures and confidence intervals using the bootstrap resampling method. In recent years, Markov chain Monte Carlo (MCMC) methodology has been used in various applications in different disciplines. Chapter 6 examines MCMC estimation of the ICC. Appendices contain the SAS code for the work in Chapters 5–6.

1.2 Intraclass Correlation Coefficient

The intraclass correlation coefficient is a reliability index. Such an index is important in many fields, as in behavioral and medical science. For example, when reading radiology films, different physicians or chiropractors may report different findings; indeed the same individual may report different findings on two reading occasions of the same radiology film. Our task is to assess the reliability of such readings. Intraclass correlation coefficients are estimators of homogeneity, not only for pairs of measurements but for larger sets as well.

Definition 1.1 Intraclass correlation coefficient *is the correlation between one measurement (either a single rating or a mean of several ratings) on a target and another measurement obtained on the same target (Shrout and Fleiss, 1979).*

The ICC is typically a ratio of the variance of interest to the sum of the variance of interest plus error and other components of variance (Bartko, 1966; Ebel, 1951; Haggard, 1958). For example, if a model contains two independent random factors (one-way random model), one is from targets and the other is from the residual term, with σ_t^2 and σ_e^2 the corresponding components of variance respectively, then the ICC is defined as $\rho_I = \sigma_t^2/(\sigma_t^2 + \sigma_e^2)$. This correlation appears mainly in the psychometric and population genetics literature. It is defined in Snedecor and Cochran, 8th Edition, (1989). Calling ρ_I a correlation dates back to Fisher (1925).

The variance components are estimated from the mean squares corresponding to the variance sources. Inferences about the ICC can be based on the analysis of variance summarizing the mean square results. An estimator of the ICC and an approximate confidence interval on the ICC can be presented as functions of the mean squares.

There are numerous versions of ICC that can give quite different results when applied to the same data. The form depends on the experimental design, the conceptual intent of the study and the assumptions of the model. For a one-way random effects model (containing only one random factor apart from the residual term) and a two-way random effects model (containing two random factors apart from the residual term), researchers have derived formulas for the various ICCs and presented discussions of the forms in various situations.

This dissertation considers the bias of the ICC in one-way models, and how to estimate an ICC and a confidence interval in a three-way random effects model (containing three crossed random factors apart from the residual term).

CHAPTER 2

ICCs FOR ONE-WAY AND TWO-WAY RANDOM MODELS

2.1 Introduction

First I review what other authors have done to estimate ICCs and construct a confidence interval around it for one-way and two-way models. As I mentioned in Chapter 1, Fleiss and Shrout (1978) gave a brief derivation of estimation of the ICCs and analogous confidence intervals in two-way random models. McGraw and Wong (1996) summarized various forms of the ICCs and presented procedures available for calculating confidence intervals and conducting tests on ICCs developed using data from one-way and two-way random and mixed-effect analysis of variance models. In order to get insight on the three-way crossed and random model, I next review in detail point estimation and confidence intervals for ICCs in one-way and two-way random models.

Since a one-way random model has a simple structure, I have the advantage of knowing the degrees of freedom relating to each variance component when the design is balanced. Then I can estimate the bias of the estimator of ICC.

In this chapter, the delta method was used to accomplish this task.

In a reliability study when each of a random sample of n subjects is rated independently by k raters, three designs can arise for the study:

1. Each subject is rated by a different set of k raters, randomly selected from a larger population of raters.

2. A random sample of k raters is selected from a larger population, and each rater rates each subject, that is, each rater rates n subject.

3. A random sample of k raters is selected from a larger population, and each rater rates each subject on m different occasions that is, each rater rates n subjects m times.

Various models result from the above cases: one-way random model corresponds to Case 1, two-way random model corresponds to Case 2 and three-way random model corresponds to Case 3. Henceforth in this dissertation, *subjects*, *targets* and *patients* will be used interchangeably and *raters*, *observers* and *judges* will be used interchangeably.

2.2 ICC for One-way Random Model

I begin with a one-way random model.

Notation 2.1 *A one-way random model may be written*

$$x_{ij} = \mu + r_i + w_{ij} \qquad i = 1, \cdots, n \quad j = 1, \cdots, k \tag{2.1}$$

where μ *is the population mean;* r_i *represents subject effects and are assumed to be i.i.d. Gaussian random variables with mean* 0 *and variance* σ_r^2*; and* w_{ij} *represents the residual terms and are assumed to be i.i.d. Gaussian random variables with mean* 0 *and variance* σ_w^2*. Moreover,* r_i *and* w_{ij} *are mutually independent.*

Table 2.1: ANOVA table for one-way model

Source of variation	df	MS	EMS
Between subjects	$n-1$	$\mathsf{MS_R}$	$k\sigma_r^2+\sigma_w^2$
Within subjects	$n(k-1)$	$\mathsf{MS_W}$	σ_w^2

Note that in the above notation, the normality assumption is required for confidence interval estimation. Since the inferences about the ICC must be based on the two mean squares from the analysis of variance summarizing the results, between subjects and within subjects (residual), I present the analysis of variance table based on this model in Table 2.1.

Because

$$\begin{aligned}\mathrm{Cov}(x_{ij},x_{ij'}) &= \sigma_r^2\\ \mathrm{Var}(x_{ij})=\mathrm{Var}(x_{ij'}) &= \sigma_r^2+\sigma_w^2\end{aligned}$$

according Definition 1.1, the ICC for the one-way random model is defined as

$$\begin{aligned}\rho_I &= \frac{\mathrm{Cov}(x_{ij},x_{ij'})}{\sqrt{\mathrm{Var}(x_{ij})}\sqrt{\mathrm{Var}(x_{ij'})}}\\ &= \frac{\sigma_r^2}{\sigma_r^2+\sigma_w^2} \qquad (2.2)\end{aligned}$$

An unbiased estimator of σ_w^2 is $\mathrm{MS_W}$ and an unbiased estimator of σ_r^2 is $(\mathrm{MS_R}-\mathrm{MS_W})/\mathrm{k}$. Placing these two terms into Equation (2.2), Fisher (1925) defined an estimator of ρ_I as

$$r_I=\frac{\mathrm{MS_R}-\mathrm{MS_W}}{\mathrm{MS_R}+(\mathrm{k}-1)\mathrm{MS_W}} \qquad (2.3)$$

which is the ratio of an unbiased estimator of the numerator of ρ_I to an unbiased estimator of the denominator of ρ_I. This estimator is a consistent but biased estimator of ρ_I (Olkin and Pratt, 1958). It is possible that r_I is as small as $-1/(n-1)$. Actually, $\rho_I\in(-\frac{1}{n-1},1)$, so r_I is admissible for ρ_I.

From the Equations (2.2) and (2.3) above, I derive a $100(1-\alpha)\%$ confidence interval for ρ_I. First by simple algebraic manipulation I get

$$\begin{aligned} 1-\rho_I &= \frac{\sigma_w^2}{\sigma_r^2+\sigma_w^2} \\ \frac{\rho_I}{1-\rho_I} &= \frac{\sigma_r^2}{\sigma_w^2} \end{aligned} \tag{2.4}$$

And since MS_R, MS_W are independent and $E(\mathrm{MS}_\mathsf{R}) = \mathrm{k}\,\sigma_\mathrm{r}^2 + \sigma_\mathrm{w}^2$, $E(\mathrm{MS}_\mathsf{W}) = \sigma_\mathrm{w}^2$. So

$$\frac{\mathrm{MS}_\mathsf{R}}{\mathrm{MS}_\mathsf{W}} \frac{\sigma_w^2}{k\,\sigma_r^2+\sigma_w^2}$$

has an F distribution with degrees of freedom $(n-1)$ and $n(k-1)$. So does

$$\begin{aligned} \frac{\mathrm{MS}_\mathsf{R}}{\mathrm{MS}_\mathsf{W}} \frac{\sigma_w^2}{k\,\sigma_r^2+\sigma_w^2} &= \frac{\mathrm{MS}_\mathsf{R}}{\mathrm{MS}_\mathsf{W}} \left(k\,\frac{\sigma_r^2}{\sigma_w^2}+1\right)^{-1} \\ &= \frac{\mathrm{MS}_\mathsf{R}}{\mathrm{MS}_\mathsf{W}} \left(k\,\frac{\rho_I}{1-\rho_I}+1\right)^{-1} \end{aligned}$$

Therefore,

$$\begin{aligned} Pr&\left\{F_{\frac{\alpha}{2}}(n-1,n(k-1)) \le \frac{\mathrm{MS}_\mathsf{R}}{\mathrm{MS}_\mathsf{W}} \left(k\,\frac{\rho_I}{1-\rho_I}+1\right)^{-1} \le F_{1-\frac{\alpha}{2}}(n-1,n(k-1))\right\} \\ &= 1-\alpha, \end{aligned}$$

where $F_{1-\frac{\alpha}{2}}(n-1,n(k-1))$ is the $(1-\alpha/2)\times 100$th percentile of an F distribution with degrees of freedom $(n-1)$ and $n(k-1)$.

From the above formula, I get the lower and upper limits of a confidence interval for ρ_I. Let $\mathrm{MS}_\mathsf{R}/\mathrm{MS}_\mathsf{W} = F$,

$$\begin{aligned} F\left(k\,\frac{\rho_I}{1-\rho_I}+1\right)^{-1} &\le F_{1-\frac{\alpha}{2}}(n-1,n(k-1)) \\ k\,\frac{\rho_I}{1-\rho_I}+1 &\ge F/F_{1-\frac{\alpha}{2}}(n-1,n(k-1)) \\ \frac{\rho_I}{1-\rho_I} &\ge \left[F/F_{1-\frac{\alpha}{2}}(n-1,n(k-1))-1\right]/k \\ \rho_I &\ge \frac{\left[F/F_{1-\frac{\alpha}{2}}(n-1,n(k-1))-1\right]/k}{1+\left[F/F_{1-\frac{\alpha}{2}}(n-1,n(k-1))-1\right]/k} \\ &= \frac{F/F_{1-\frac{\alpha}{2}}(n-1,n(k-1))-1}{F/F_{1-\frac{\alpha}{2}}(n-1,n(k-1))+(k-1)} \end{aligned}$$

Similarly,

$$\begin{aligned}
F\left(k\,\frac{\rho_I}{1-\rho_I}+1\right)^{-1} &\geq F_{\frac{\alpha}{2}}(n-1,n(k-1))\\
k\,\frac{\rho_I}{1-\rho_I}+1 &\leq F/F_{\frac{\alpha}{2}}(n-1,n(k-1))\\
\frac{\rho_I}{1-\rho_I} &\leq \left[F/F_{\frac{\alpha}{2}}(n-1,n(k-1))-1\right]/k\\
\rho_I &\leq \frac{\left[F/F_{\frac{\alpha}{2}}(n-1,n(k-1))-1\right]/k}{1+\left[F/F_{\frac{\alpha}{2}}(n-1,n(k-1))-1\right]/k}\\
&= \frac{F/F_{\frac{\alpha}{2}}(n-1,n(k-1))-1}{F/F_{\frac{\alpha}{2}}(n-1,n(k-1))-1+(k-1)}\\
&= \frac{F\cdot F_{1-\frac{\alpha}{2}}(n(k-1),n-1)-1}{F\cdot F_{1-\frac{\alpha}{2}}(n(k-1),n-1)+(k-1)}
\end{aligned}$$

So a $100(1-\alpha)\%$ confidence interval for ρ_I based on the ANOVA table from a one-way random model is (Fisher, 1925)

$$\left[\frac{F/F_{1-\frac{\alpha}{2}}(n-1,n(k-1))-1}{F/F_{1-\frac{\alpha}{2}}(n-1,n(k-1))+(k-1)},\frac{F\cdot F_{1-\frac{\alpha}{2}}(n(k-1),n-1)-1}{F\cdot F_{1-\frac{\alpha}{2}}(n(k-1),n-1)+(k-1)}\right]$$

2.3 Bias of Estimator of ICC for One-way Random Model

The estimator r_I is a biased estimator. Due to the simplicity of one-way random models, I can estimate its bias.

As defined in Equation (2.3), an estimator of ρ_I is

$$r_I = \frac{\mathrm{MS}_{\mathsf{R}} - \mathrm{MS}_{\mathsf{W}}}{\mathrm{MS}_{\mathsf{R}} + (\mathrm{k}-1)\mathrm{MS}_{\mathsf{W}}}$$

and from this the following equations hold

$$\begin{aligned}
\frac{1}{r_I} &= \frac{\mathrm{MS}_{\mathsf{R}} + (\mathrm{k}-1)\mathrm{MS}_{\mathsf{W}}}{\mathrm{MS}_{\mathsf{R}} - \mathrm{MS}_{\mathsf{W}}} \\
&= 1 + \frac{k \cdot \mathrm{MS}_{\mathsf{W}}}{\mathrm{MS}_{\mathsf{R}} - \mathrm{MS}_{\mathsf{W}}} \\
\frac{1}{r_I} - 1 &= \frac{k \cdot \mathrm{MS}_{\mathsf{W}}}{\mathrm{MS}_{\mathsf{R}} - \mathrm{MS}_{\mathsf{W}}} \\
\frac{r_I}{1-r_I} &= \frac{\mathrm{MS}_{\mathsf{R}} - \mathrm{MS}_{\mathsf{W}}}{k \cdot \mathrm{MS}_{\mathsf{W}}} \\
\frac{k r_I}{1-r_I} + 1 &= \frac{\mathrm{MS}_{\mathsf{R}}}{\mathrm{MS}_{\mathsf{W}}}
\end{aligned}$$

Since $F = \dfrac{\mathrm{MS}_{\mathsf{R}}/(\mathrm{k}\sigma_{\mathrm{r}}^2 + \sigma_{\mathrm{w}}^2)}{\mathrm{MS}_{\mathsf{W}}/(\sigma_{\mathrm{w}}^2)} \sim F_{\nu_1,\nu_2}$, with $\nu_1 = n-1$ and $\nu_2 = n(k-1)$, I calculate the expectation and variance of a function of r_I, $W = 1/(1-r_I)$.

$$\begin{aligned}
E\left(\frac{\mathrm{MS}_{\mathsf{R}}/(\mathrm{k}\sigma_{\mathrm{r}}^2 + \sigma_{\mathrm{w}}^2)}{\mathrm{MS}_{\mathsf{W}}/(\sigma_{\mathrm{w}}^2)}\right) &= \frac{\nu_2}{\nu_2 - 2}, \quad \nu_2 > 2 \\
E\left(\frac{[1+(k-1)r_I]/(k\sigma_r^2 + \sigma_w^2)}{(1-r_I)/\sigma_w^2}\right) &= \frac{\nu_2}{\nu_2 - 2} \\
E\left(\frac{1+(k-1)r_I}{1-r_I}\right) &= \frac{\nu_2}{\nu_2 - 2}\frac{k\sigma_r^2 + \sigma_w^2}{\sigma_w^2} \\
E\left(1 + \frac{k}{1-r_I} - k\right) &= \frac{\nu_2}{\nu_2 - 2}\frac{k\sigma_r^2 + \sigma_w^2}{\sigma_w^2} \\
E\left(\frac{1}{1-r_I}\right) &= \frac{1}{k}\frac{\nu_2}{\nu_2 - 2}\frac{k\sigma_r^2 + \sigma_w^2}{\sigma_w^2} + \frac{k-1}{k}
\end{aligned}$$

Replacing Equation (2.4) in the above equation gives

$$E(W) = \frac{1}{k}\left(\frac{\nu_2}{\nu_2 - 2}\right)\left(\frac{1+(k-1)\rho_I}{1-\rho_I}\right) + \frac{k-1}{k}$$

Next I derive $\mathrm{Var}(W)$. Since

$$W = \frac{1}{1-r_I} = \frac{1}{1 - \frac{\mathrm{MS}_{\mathsf{R}} - \mathrm{MS}_{\mathsf{W}}}{\mathrm{MS}_{\mathsf{R}} + (\mathrm{k}-1)\mathrm{MS}_{\mathsf{W}}}} = \frac{\mathrm{MS}_{\mathsf{R}} + (\mathrm{k}-1)\mathrm{MS}_{\mathsf{W}}}{k \cdot \mathrm{MS}_{\mathsf{W}}} = \frac{\mathrm{MS}_{\mathsf{R}}}{k \cdot \mathrm{MS}_{\mathsf{W}}} + \frac{k-1}{k}$$

I have $\mathrm{Var}(W) = \mathrm{Var}\left(\mathrm{MS}_{\mathsf{R}}/(k \cdot \mathrm{MS}_{\mathsf{W}})\right)$.

Remark 2.1 *If* $X \sim F_{\nu_1,\nu_2}$, *then* $\mathrm{Var}(X) = \left[2\nu_2^2(\nu_1+\nu_2-2)\right]/\left[\nu_1(\nu_2-2)^2(\nu_2-4)\right]$, *with* $\nu_2 > 4$.

By the above Remark,

$$\begin{aligned} \mathrm{Var}\left(\frac{\mathrm{MS_R}/(\mathrm{k}\sigma_\mathrm{r}^2+\sigma_\mathrm{w}^2)}{\mathrm{MS_W}/(\sigma_\mathrm{w}^2)}\right) &= \frac{2\nu_2^2(\nu_1+\nu_2-2)}{\nu_1(\nu_2-2)^2(\nu_2-4)} \\ \mathrm{Var}\left(\frac{\mathrm{MS_R}}{\mathrm{MS_W}}\right) &= \left(\frac{k\sigma_r^2+\sigma_w^2}{\sigma_w^2}\right)^2 \frac{2\nu_2^2(\nu_1+\nu_2-2)}{\nu_1(\nu_2-2)^2(\nu_2-4)} \end{aligned}$$

So

$$\begin{aligned} \mathrm{Var}(W) &= \mathrm{Var}\left(\frac{\mathrm{MS_R}}{k\cdot\mathrm{MS_W}}\right) \\ &= \left(\frac{k\sigma_r^2+\sigma_w^2}{k\sigma_w^2}\right)^2 \frac{2\nu_2^2(\nu_1+\nu_2-2)}{\nu_1(\nu_2-2)^2(\nu_2-4)} \\ &= \left(\frac{1+(k-1)\rho_I}{k(1-\rho_I)}\right)^2 \frac{2\nu_2^2(\nu_1+\nu_2-2)}{\nu_1(\nu_2-2)^2(\nu_2-4)} \end{aligned}$$

Remark 2.2 **Delta method:** *let* X *be a random variable with* $E(X) = \mu$ *and* $g(X)$ *be a function of* X. *Then*

$$\begin{aligned} E(g(X)) &\approx g(\mu) + \frac{1}{2}\frac{\partial^2 g(X)}{\partial X^2}\mathrm{Var}(X) \\ \mathrm{Var}(g(X)) &\approx \left(\frac{\partial g(X)}{\partial X}\right)^2 \mathrm{Var}(X) \end{aligned}$$

Since $W = 1/(1-r_I)$, then $r_I = g(W) = 1 - 1/W$, and

$$\begin{aligned} \frac{\partial r_I}{\partial W} &= 1/W^2 \\ \frac{\partial^2 r_I}{\partial W^2} &= -2/W^3 \end{aligned}$$

Therefore, using the delta method,

$$\begin{aligned} E(r_I) &= E(g(W)) \\ &\approx \left(1-\frac{1}{E(W)}\right) - \frac{1}{2}\left(\frac{2}{E(W)^3}\right)\mathrm{Var}(W) \\ &= 1 - \frac{k}{(\frac{\nu_2}{\nu_2-2})(\frac{1+(k-1)\rho_I}{1-\rho_I}) + k - 1} \\ &\quad - \frac{k^3}{\left[(\frac{\nu_2}{\nu_2-2})(\frac{1+(k-1)\rho_I}{1-\rho_I}) + k - 1\right]^3}\left(\frac{1+(k-1)\rho_I}{k(1-\rho_I)}\right)^2 \frac{2\nu_2^2(\nu_1+\nu_2-2)}{\nu_1(\nu_2-2)^2(\nu_2-4)} \end{aligned}$$

Further

$$
\begin{aligned}
\mathrm{Var}(r_I) =& \mathrm{Var}(g(W)) \\
\approx& \left(\frac{\partial r_I}{\partial W}\right)^2 \mathrm{Var}(W) \\
=& \frac{1}{(E(W))^4} \mathrm{Var}(W) \\
=& \frac{k^4}{\left[(\frac{\nu_2}{\nu_2-2})(\frac{1+(k-1)\rho_I}{1-\rho_I}) + k - 1\right]^4} \left(\frac{1+(k-1)\rho_I}{k(1-\rho_I)}\right)^2 \frac{2\nu_2^2(\nu_1+\nu_2-2)}{\nu_1(\nu_2-2)^2(\nu_2-4)}
\end{aligned}
$$

Table 2.2 and Table 2.3 list the approximate bias of the estimator of ICC for some specific parameters. Bias is defined as Bias $= E(r_I) - \rho_I$ and RE (relative error) is defined as RE $= (\mathrm{Bias}/\rho_I) \times 100\%$. When $k = 3$ and n (500–2000) is large, the RE is very small, with the largest value 0.2%. When n (30–60) is relatively small, which is more common in reality, the REs increase, but are still acceptable. These results provide confidence on the accuracy of the ANOVA-based point estimator of this ICC.

Table 2.2: Bias of ICC when $k = 3$ and n is large

	ρ_I	n	k	$E(r_I)$	Bias	RE(%)
1	0.1	500	3	0.0999	−0.0001	0.1
2	0.1	1000	3	0.0999	−0.0001	0.1
3	0.1	1500	3	0.1000	0.0000	< 0.001
4	0.1	2000	3	0.1000	0.0000	< 0.001
5	0.2	500	3	0.1997	−0.0003	0.2
6	0.2	1000	3	0.1998	−0.0002	0.1
7	0.2	1500	3	0.1999	−0.0001	0.1
8	0.2	2000	3	0.1999	−0.0001	< 0.001
9	0.3	500	3	0.2995	−0.0005	0.2
10	0.3	1000	3	0.2998	−0.0002	0.1
11	0.3	1500	3	0.2999	−0.0001	< 0.001
12	0.3	2000	3	0.2999	−0.0001	< 0.001
13	0.4	500	3	0.3994	−0.0006	0.1
14	0.4	1000	3	0.3997	−0.0003	0.1
15	0.4	1500	3	0.3998	−0.0002	< 0.001
16	0.4	2000	3	0.3999	−0.0001	< 0.001
17	0.5	500	3	0.4993	−0.0007	0.1
18	0.5	1000	3	0.4997	−0.0003	0.1
19	0.5	1500	3	0.4998	−0.0002	< 0.001
20	0.5	2000	3	0.4998	−0.0002	< 0.001
21	0.6	500	3	0.5993	−0.0007	0.1
22	0.6	1000	3	0.5996	−0.0004	0.1
23	0.6	1500	3	0.5998	−0.0002	< 0.001
24	0.6	2000	3	0.5998	−0.0002	< 0.001
25	0.7	500	3	0.6993	−0.0007	0.1
26	0.7	1000	3	0.6997	−0.0003	< 0.001
27	0.7	1500	3	0.6998	−0.0002	< 0.001
28	0.7	2000	3	0.6998	−0.0002	< 0.001
29	0.8	500	3	0.7994	−0.0006	0.1
30	0.8	1000	3	0.7997	−0.0003	< 0.001
31	0.8	1500	3	0.7998	−0.0002	< 0.001
32	0.8	2000	3	0.7999	−0.0001	< 0.001
33	0.9	500	3	0.8997	−0.0003	< 0.001
34	0.9	1000	3	0.8998	−0.0002	< 0.001
35	0.9	1500	3	0.8999	−0.0001	< 0.001
36	0.9	2000	3	0.8999	−0.0001	< 0.001

Table 2.3: Bias of ICC when $k = 3$ and n is small

	ρ_I	n	k	$E(r_I)$	Bias	RE(%)
1	0.1	30	3	0.0964	−0.0036	3.6
2	0.1	40	3	0.0975	−0.0025	2.5
3	0.1	50	3	0.0981	−0.0019	1.9
4	0.1	60	3	0.0985	−0.0015	1.5
5	0.2	30	3	0.1936	−0.0064	3.2
6	0.2	40	3	0.1955	−0.0045	2.3
7	0.2	50	3	0.1965	−0.0035	1.7
8	0.2	60	3	0.1972	−0.0028	1.4
9	0.3	30	3	0.2910	−0.0090	3.0
10	0.3	40	3	0.2936	−0.0064	2.1
11	0.3	50	3	0.2950	−0.0050	1.7
12	0.3	60	3	0.2959	−0.0041	1.4
13	0.4	30	3	0.3889	−0.0111	2.8
14	0.4	40	3	0.3920	−0.0080	2.0
15	0.4	50	3	0.3937	−0.0063	1.6
16	0.4	60	3	0.3948	−0.0052	1.3
17	0.5	30	3	0.4874	−0.0126	2.5
18	0.5	40	3	0.4909	−0.0091	1.8
19	0.5	50	3	0.4928	−0.0072	1.4
20	0.5	60	3	0.4941	−0.0059	1.2
21	0.6	30	3	0.5870	−0.0130	2.2
22	0.6	40	3	0.5905	−0.0095	1.6
23	0.6	50	3	0.5925	−0.0075	1.2
24	0.6	60	3	0.5938	−0.0062	1.0
25	0.7	30	3	0.6878	−0.0122	1.7
26	0.7	40	3	0.6910	−0.0090	1.3
27	0.7	50	3	0.6929	−0.0071	1.0
28	0.7	60	3	0.6941	−0.0059	0.8
29	0.8	30	3	0.7900	−0.0100	1.2
30	0.8	40	3	0.7927	−0.0073	0.9
31	0.8	50	3	0.7942	−0.0058	0.7
32	0.8	60	3	0.7952	−0.0048	0.6
33	0.9	30	3	0.8940	−0.0060	0.7
34	0.9	40	3	0.8956	−0.0044	0.5
35	0.9	50	3	0.8965	−0.0035	0.4
36	0.9	60	3	0.8971	−0.0029	0.3

Table 2.4: ANOVA table for two-way model

Source of variation	df	MS	EMS
Between raters	$k-1$	MS_C	$n\sigma_c^2 + \sigma_{rc}^2 + \sigma_e^2$
Between subjects	$n-1$	MS_R	$k\sigma_r^2 + \sigma_{rc}^2 + \sigma_e^2$
Residual	$(n-1)(k-1)$	MS_E	$\sigma_{rc}^2 + \sigma_e^2$

2.4 ICC for Two-way Random Model

Next I turn to the ICC for two-way random models.

Notation 2.2 *A two-way random model may be written*

$$x_{ij} = \mu + r_i + c_j + (rc)_{ij} + e_{ij} \qquad i = 1, \cdots, n \quad j = 1, \cdots, k \tag{2.5}$$

where μ is the population mean; r_i represents the subject effects and are assumed to be i.i.d. Gaussian random variables with mean 0 and variance σ_r^2; c_j represents rater effects and are assumed to be i.i.d. Gaussian random variables with mean 0 and variance σ_c^2; $(rc)_{ij}$ represents interaction effects and are assumed to be i.i.d. Gaussian random variables with mean 0 and variance σ_{rc}^2; and e_{ij} represents the residual terms and are assumed to be i.i.d. Gaussian random variables with mean 0 and variance σ_e^2. Moreover, all the random effects are assumed mutually independent.

As for one-way random model, the inferences about the ICC are based on the mean squares from the analysis of variance summarizing the results, which is presented in Table 2.4:

According to Definition 1.1, the ICC for the two-way random model is defined as

$$\rho_I = \frac{\sigma_r^2}{\sigma_r^2 + \sigma_c^2 + \sigma_{rc}^2 + \sigma_e^2} \tag{2.6}$$

Shrout and Fleiss (1979) defined an estimator of ρ_I as

$$r_I = \frac{\text{MS}_\mathsf{R} - \text{MS}_\mathsf{E}}{\text{MS}_\mathsf{R} + (\text{k}-1)\text{MS}_\mathsf{E} + \text{k}(\text{MS}_\mathsf{C} - \text{MS}_\mathsf{E})/\text{n}} \tag{2.7}$$

which is the ratio of an unbiased estimator of the numerator of ρ_I to an unbiased estimator of the denominator of ρ_I, again consistent but biased.

Next I derive a $100(1-\alpha)\%$ confidence interval for ρ_I. First from Definition (2.6), I get

$$1-\rho_I = \frac{\sigma_c^2+(\sigma_{rc}^2+\sigma_e^2)}{\sigma_r^2+\sigma_c^2+(\sigma_{rc}^2+\sigma_e^2)}$$
$$\frac{\rho_I}{1-\rho_I} = \frac{\sigma_r^2}{\sigma_c^2+(\sigma_{rc}^2+\sigma_e^2)} \tag{2.8}$$

Substituting σ_r^2 from Equation (2.8) in the expectation of MS_{R} in Table 2.4, I find

$$\begin{aligned}
E(\mathrm{MS}_{\mathsf{R}}) &= k\sigma_r^2+\sigma_{rc}^2+\sigma_e^2\\
&= k\frac{\rho_I}{1-\rho_I}\left(\sigma_c^2+(\sigma_{rc}^2+\sigma_e^2)\right)+(\sigma_{rc}^2+\sigma_e^2)\\
&= k\frac{\rho_I}{1-\rho_I}\sigma_c^2+\left(k\frac{\rho_I}{1-\rho_I}+1\right)(\sigma_{rc}^2+\sigma_e^2)\\
&= k\frac{\rho_I}{1-\rho_I}\sigma_c^2+\frac{1+(k-1)\rho_I}{1-\rho_I}(\sigma_{rc}^2+\sigma_e^2)\\
&= \frac{1}{1-\rho_I}\left\{k\rho_I\sigma_c^2+[1+(k-1)\rho_I](\sigma_{rc}^2+\sigma_e^2)\right\}
\end{aligned}$$

Further, since $\sigma_c^2 = E(\mathrm{MS}_{\mathsf{C}}-\mathrm{MS}_{\mathsf{E}})/\mathrm{n}$ and $\sigma_{rc}^2+\sigma_e^2 = E(\mathrm{MS}_{\mathsf{E}})$, replacing σ_c^2 by $(\mathrm{MS}_{\mathsf{C}}-\mathrm{MS}_{\mathsf{E}})/\mathrm{n}$ and $\sigma_{rc}^2+\sigma_e^2$ by MS_{E} in the above equation, I construct a variable V so that $E(V)=E(\mathrm{MS}_{\mathsf{R}})$.

$$\begin{aligned}
V &= \frac{1}{1-\rho_I}\left\{k\rho_I\cdot(\mathrm{MS}_{\mathsf{C}}-\mathrm{MS}_{\mathsf{E}})/\mathrm{n}+[1+(\mathrm{k}-1)\rho_\mathrm{I}]\cdot(\mathrm{MS}_{\mathsf{E}})\right\}\\
&= \frac{1}{n(1-\rho_I)}\left\{k\rho_I\cdot(\mathrm{MS}_{\mathsf{C}})+[\mathrm{n}\,(1+(\mathrm{k}-1)\rho_\mathrm{I})-\mathrm{k}\rho_\mathrm{I}]\cdot(\mathrm{MS}_{\mathsf{E}})\right\}\\
&= \frac{k\rho_I}{n(1-\rho_I)}\mathrm{MS}_{\mathsf{C}}+\left[1+\frac{\mathrm{k}\rho_\mathrm{I}}{1-\rho_\mathrm{I}}-\frac{\mathrm{k}\rho_\mathrm{I}}{\mathrm{n}(1-\rho_\mathrm{I})}\right]\mathrm{MS}_{\mathsf{E}}
\end{aligned}$$

Remark 2.3 Satterthwaite's Approximation*(Satterthwaite, 1946): If* $\mathrm{MS}_1, \mathrm{MS}_2, \cdots, \mathrm{MS}_\mathrm{k}$ *are independent mean squares with* $r_1, r_2, \cdots, r_k$ *degrees of freedom and*

$$\hat{V}_s = a_1(\mathrm{MS}_1)+\mathrm{a}_2(\mathrm{MS}_2)+\cdots+\mathrm{a}_\mathrm{k}(\mathrm{MS}_\mathrm{k})$$

the number of degrees of freedom of the approximating chi-square is found to be given by

$$r_s = \frac{[a_1E(\mathrm{MS}_1)+\mathrm{a}_2\mathrm{E}(\mathrm{MS}_2)+\cdots+\mathrm{a}_\mathrm{k}\mathrm{E}(\mathrm{MS}_\mathrm{k})]^2}{[a_1E(\mathrm{MS}_1)]^2/r_1+[a_2E(\mathrm{MS}_2)]^2/r_2+\cdots+[a_kE(\mathrm{MS}_\mathrm{k})]^2/r_k}$$

In practice the expected values of the independent mean squares will not be known. The observed values will usually be substituted in, giving, as an estimate of r_s

$$\hat{r_s} = \frac{[a_1(\mathrm{MS_1}) + \mathrm{a_2(MS_2)} + \cdots + \mathrm{a_k(MS_k)}]^2}{[a_1(\mathrm{MS_1})]^2/r_1 + [a_2(\mathrm{MS_2})]^2/r_2 + \cdots + [a_k(\mathrm{MS_k})]^2/r_k}$$

The variable V is a linear combination of two independent mean squares:$\mathrm{MS_C}$ and $\mathrm{MS_E}$. According to Satterthwaite's approximation, V is approximately distributed as $c\chi^2_\nu/\nu$ with $c = E(V)$ and ν the degrees of freedom. Let $F_c = \mathrm{MS_C}/\mathrm{MS_E}$,

$$\nu = \frac{(a\mathrm{MS_C} + \mathrm{bMS_E})^2}{\frac{(a\mathrm{MS_C})^2}{k-1} + \frac{(b\mathrm{MS_E})^2}{(n-1)(k-1)}} \qquad (\text{where} \quad a = \frac{k\rho_I}{n(1-\rho_I)}, \quad b = 1 + \frac{k\rho_I}{1-\rho_I} - \frac{k\rho_I}{n(1-\rho_I)})$$

$$= \frac{\left\{\frac{k\rho_I}{n(1-\rho_I)}\mathrm{MS_C} + \left[1 + \frac{\mathrm{k}\rho_\mathrm{I}}{1-\rho_\mathrm{I}} - \frac{\mathrm{k}\rho_\mathrm{I}}{\mathrm{n}(1-\rho_\mathrm{I})}\right]\mathrm{MS_E}\right\}^2}{(\frac{k\rho_I}{n(1-\rho_I)}\mathrm{MS_C})^2/(\mathrm{k}-1) + \left[(1 + \frac{\mathrm{k}\rho_\mathrm{I}}{1-\rho_\mathrm{I}} - \frac{\mathrm{k}\rho_\mathrm{I}}{\mathrm{n}(1-\rho_\mathrm{I})})\mathrm{MS_E}\right]^2/[(\mathrm{n}-1)(\mathrm{k}-1)]}$$

$$= \frac{(k-1)(n-1)\left\{k\rho_I F_c + n[1+(k-1)\rho_I] - k\rho_I\right\}^2}{(n-1)k^2\rho_I^2F_c^2 + \left\{n[1+(k-1)\rho_I] - k\rho_I\right\}^2}$$

To estimate ν, use r_I to replace ρ_I.

V and $\mathrm{MS_R}$ are each scaled χ^2 variates with the same expected values and they are independent, so

$$F = \mathrm{MS_R}/\mathrm{V}$$
$$= \mathrm{MS_R}/(\mathrm{aMS_C} + \mathrm{bMS_E})$$

has an approximate F distribution with degrees of freedom $(n-1)$ and ν. Then

$$Pr\left\{F_{\frac{\alpha}{2}}(n-1,\nu) \leq \mathrm{MS_R}/(\mathrm{aMS_C} + \mathrm{bMS_E}) \leq \mathrm{F}_{1-\frac{\alpha}{2}}(\mathrm{n}-1,\nu)\right\} = 1-\alpha,$$

where $F_{1-\frac{\alpha}{2}}(n-1,\nu)$ is the $(1-\frac{\alpha}{2}) \times 100$th percentile of an F distribution with degrees of freedom $(n-1)$ and ν.

From the above formula, I get the lower and upper limits for a confidence interval on ρ_I. Let $\rho_I^* = \rho_I/(1-\rho_I)$ and $F_{1-\frac{\alpha}{2}}(n-1,\nu) = F_*$,

$$
\begin{aligned}
F_* &\geq \frac{\mathrm{MS_R}}{a\mathrm{MS_C} + \mathrm{bMS_E}} \\
F_* &\geq \frac{\mathrm{MS_R}}{(k/n)\rho_I^*\mathrm{MS_C} + [1 + \mathrm{k}\rho_\mathrm{I}^* - (\mathrm{k/n})\rho_\mathrm{I}^*]\mathrm{MS_E}} \\
\mathrm{MS_R} &\leq F_* \left\{(k/n)\rho_I^*\mathrm{MS_C} + [1 + \mathrm{k}\rho_\mathrm{I}^* - (\mathrm{k/n})\rho_\mathrm{I}^*]\mathrm{MS_E}\right\} \\
\frac{\mathrm{MS_R}}{F_*} &\leq \rho_I^*\{(k/n)\mathrm{MS_C} + [\mathrm{k} - (\mathrm{k/n})]\mathrm{MS_E}\} + \mathrm{MS_E} \\
\rho_I^* &\geq \frac{\mathrm{MS_R/F_*} - \mathrm{MS_E}}{(k/n)\mathrm{MS_C} + [\mathrm{k} - (\mathrm{k/n})]\mathrm{MS_E}} \\
\rho_I &\geq \left\{\frac{\mathrm{MS_R/F_*} - \mathrm{MS_E}}{(k/n)\mathrm{MS_C} + [\mathrm{k} - (\mathrm{k/n})]\mathrm{MS_E}}\right\} \Big/ \left\{1 + \frac{\mathrm{MS_R/F_*} - \mathrm{MS_E}}{(k/n)\mathrm{MS_C} + [\mathrm{k} - (\mathrm{k/n})]\mathrm{MS_E}}\right\} \\
\rho_I &\geq \frac{n(\mathrm{MS_R} - \mathrm{F_*MS_E})}{F_*[k\mathrm{MS_C} + (\mathrm{nk - k - n})\mathrm{MS_E}] + \mathrm{nMS_R}}
\end{aligned}
$$

The lower limit for the approximate $100(1-\alpha)\%$ confidence interval on ρ_I is complete.

Similarly, I derive the upper limit. Let $F_{1-\frac{\alpha}{2}}(\nu, n-1) = F^*$,

$$
\begin{aligned}
\rho_I &\leq \frac{n[\mathrm{MS_R} - \mathrm{F}_{\frac{\alpha}{2}}(\mathrm{n}-1,\nu)\mathrm{MS_E}]}{F_{\frac{\alpha}{2}}(n-1,\nu)[k\mathrm{MS_C} + (\mathrm{nk - k - n})\mathrm{MS_E}] + \mathrm{nMS_R}} \\
&= \frac{n[\mathrm{MS_R} - \mathrm{MS_E/F^*}]}{[k\mathrm{MS_C} + (\mathrm{nk - k - n})\mathrm{MS_E}]/\mathrm{F^*} + \mathrm{nMS_R}} \\
&= \frac{n(F^*\mathrm{MS_R} - \mathrm{MS_E})}{k\mathrm{MS_C} + (\mathrm{nk - k - n})\mathrm{MS_E} + \mathrm{nF^*MS_R}}.
\end{aligned}
$$

Since V is approximately distributed as $c\chi^2_\nu/\nu$, the upper and lower limits are approximate. So an approximate $100(1-\alpha)\%$ confidence interval for ρ_I for a two-way random model is (Fleiss and Shrout, 1978)

$$
\left[\frac{n(\mathrm{MS_R} - \mathrm{F_*MS_E})}{F_*[k\mathrm{MS_C} + (\mathrm{nk - k - n})\mathrm{MS_E}] + \mathrm{nMS_R}}, \frac{n(F^*\mathrm{MS_R} - \mathrm{MS_E})}{k\mathrm{MS_C} + (\mathrm{nk - k - n})\mathrm{MS_E} + \mathrm{nF^*MS_R}}\right]
$$

CHAPTER 3

RELIABILITY PARAMETERS FOR FULL THREE-WAY RANDOM EFFECTS MODEL

3.1 Introduction

In this case, in addition to subjects and raters, there is a third random factor, occasions, that needs to be considered. To accommodate the effect of occasions, a three-way random effects model is proposed.

Notation 3.1 *A three-way random effects model may be written*

$$x_{ijk} = \mu + p_i + r_j + o_k + (pr)_{ij} + (po)_{ik} + (ro)_{jk} + e_{ijk} \qquad (3.1)$$
$$i = 1, \cdots, n_p \quad j = 1, \cdots, n_r \quad k = 1, \cdots, n_o$$

where μ is the population mean; p_i represents subject effects and are assumed to be i.i.d. Gaussian random variables with mean 0 and variance σ_p^2; r_j represents rater effects and are assumed to be i.i.d. Gaussian random variables with mean 0 and variance σ_r^2; o_k represents occasion effects and are assumed to be i.i.d. Gaussian random variables with mean 0 and variance σ_o^2; $(pr)_{ij}$, $(po)_{ik}$, and $(ro)_{jk}$ represent the interaction effects and are assumed to be i.i.d.

Gaussian random variables with mean 0 *and variance* σ^2_{pr}*,* σ^2_{po} *and* σ^2_{ro}*; and* e_{ijk} *represents the residual terms and are assumed to be i.i.d. Gaussian random variables with mean* 0 *and variance* σ^2_e*. Moreover, all the random effects are mutually independent.*

In the preceding, the normality assumption is needed only for construction of confidence intervals. Since usually there is only one observation for each subject-rater-occasion combination, the three-way interaction is not estimable. Therefore I call the above model the full three-way random effects model and throughout the dissertation I refer to any model lacking one or more of the two-way interactions as a reduced three-way random effects model.

To draw inferences about the ICC, ANOVA decomposition results like that for one-way and two-way random effect model are needed. Each mean square corresponding to a random effect is a quadratic form based on the random variables. The expected values of the mean squares can be calculated from the quadratic forms. In this dissertation, the initial intent of developing the quadratic forms was to deduce the exact distribution of the ICC, which is a ratio of quadratic forms. But the complexity of the variance-covariance matrix of the data vector makes this difficult. Then I reverted to an ANOVA-based estimator. The expected values of the mean squares are still required. I present the results in matrix notation. Comparable results for simpler models appear in Searle, Casella and McCulloch (1992).

3.2 Quadratic Forms and Expected Values of Mean Squares for Full Model

3.2.1 Quadratic Form of Mean Squares

Let the data be presented as

$$\mathbf{x} = \{x_{ijk}\} = \begin{bmatrix} x_{111} \\ x_{112} \\ \vdots \\ x_{11n_o} \\ x_{121} \\ x_{122} \\ \vdots \\ x_{12n_o} \\ \vdots \\ x_{1n_r1} \\ x_{1n_r2} \\ \vdots \\ x_{1n_r\,n_o} \\ \vdots \\ x_{n_p\,n_r\,n_o} \end{bmatrix}.$$

Let $N = n_p \times n_r \times n_o$, then the grand mean and mean vectors of main categories are

$$\begin{aligned}
\bar{x}_{...} &= \sum_{i=1}^{n_p}\sum_{j=1}^{n_r}\sum_{k=1}^{n_o} x_{ijk}/N \\
&= \frac{1}{n_p\, n_r\, n_o}\mathbf{1}'_N\mathbf{x} \\
\bar{\mathbf{x}}_{i..} &= \frac{1}{n_r\, n_o}(\mathbf{I}_{n_p} \otimes \mathbf{1}'_{n_r\, n_o})\mathbf{x} \\
\bar{\mathbf{x}}_{.j.} &= \frac{1}{n_p\, n_o}(\mathbf{1}'_{n_p} \otimes \mathbf{I}_{n_r}\mathbf{1}'_{n_o})\mathbf{x} \\
\bar{\mathbf{x}}_{..k} &= \frac{1}{n_p\, n_r}(\mathbf{1}'_{n_p\, n_r} \otimes \mathbf{I}_{n_o})\mathbf{x}
\end{aligned}$$

The total sum of squares is

$$\begin{aligned}
\mathrm{SS}_T &= \sum_{i=1}^{n_p}\sum_{j=1}^{n_r}\sum_{k=1}^{n_o}(x_{ijk} - \bar{x}_{...})^2 \\
&= \mathbf{x}'(\mathbf{I}_N - \bar{\mathbf{J}}_N)\mathbf{x}
\end{aligned}$$

where $\bar{\mathbf{J}}_N$ is a $N \times N$ matrix with all entries to be $1/N$.

The sum of squares due to subjects is

$$\begin{aligned}
\mathrm{SS}_p &= n_r\, n_o \sum_{i=1}^{n_p}(\bar{x}_{i..} - \bar{x}_{...})^2 \\
&= n_r\, n_o(\bar{\mathbf{x}}'_{i..}\bar{\mathbf{x}}_{i..} - n_p\bar{x}^2_{...}) \\
&= n_r\, n_o\mathbf{x}'\left[\frac{1}{n_r^2\, n_o^2}(\mathbf{I}_{n_p} \otimes \mathbf{1}_{n_r\, n_o})(\mathbf{I}_{n_p} \otimes \mathbf{1}'_{n_r\, n_o}) - \frac{n_p}{N^2}\mathbf{1}_N\mathbf{1}'_N\right]\mathbf{x} \\
&= \mathbf{x}'\left[\mathbf{I}_{n_p} \otimes \bar{\mathbf{J}}_{n_r\, n_o} - \bar{\mathbf{J}}_N\right]\mathbf{x}
\end{aligned}$$

The sum of squares due to raters is

$$
\begin{aligned}
\mathrm{SS}_r &= n_p\, n_o \sum_{j=1}^{n_r} (\bar{x}_{.j.} - \bar{x}_{...})^2 \\
&= n_p\, n_o (\bar{\mathbf{x}}'_{.j.}\bar{\mathbf{x}}_{.j.} - n_r \bar{x}^2_{...}) \\
&= \mathbf{x}' \left[\Big((\mathbf{1}_{n_p} \otimes \mathbf{I}_{n_r})(\mathbf{1}'_{n_p} \otimes \mathbf{I}_{n_r}) \Big) \otimes \mathbf{J}_{n_o}/(n_p\, n_o) - \frac{1}{N}\mathbf{J}_N \right] \mathbf{x} \\
&= \mathbf{x}' \left[\bar{\mathbf{J}}_{n_p} \otimes \mathbf{I}_{n_r} \otimes \bar{\mathbf{J}}_{n_o} - \bar{\mathbf{J}}_N \right] \mathbf{x}
\end{aligned}
$$

And the sum of squares due to occasions is

$$
\begin{aligned}
\mathrm{SS}_o &= n_p\, n_r \sum_{k=1}^{n_o} (\bar{x}_{..k} - \bar{x}_{...})^2 \\
&= n_p\, n_r (\bar{\mathbf{x}}'_{..k}\bar{\mathbf{x}}_{..k} - n_o \bar{x}^2_{...}) \\
&= n_p\, n_r \mathbf{x}' \left[\frac{1}{n_p^2\, n_r^2} (\mathbf{1}_{n_p\, n_r} \otimes \mathbf{I}_{n_o})(\mathbf{1}'_{n_p\, n_r} \otimes \mathbf{I}_{n_o}) - \frac{n_o}{N^2} \mathbf{1}_N \mathbf{1}'_N \right] \mathbf{x} \\
&= \mathbf{x}' \left[\bar{\mathbf{J}}_{n_p\, n_r} \otimes \mathbf{I}_{n_o} - \bar{\mathbf{J}}_N \right] \mathbf{x}
\end{aligned}
$$

The mean vectors of category combinations are

$$
\begin{aligned}
\bar{\mathbf{x}}_{ij.} &= \frac{1}{n_o} (\mathbf{I}_{n_p\, n_r} \otimes \mathbf{1}'_{n_o}) \mathbf{x} \\
\bar{\mathbf{x}}_{i.k} &= \frac{1}{n_r} (\mathbf{I}_{n_p} \otimes \mathbf{1}'_{n_r} \otimes \mathbf{I}_{n_o}) \mathbf{x} \\
\bar{\mathbf{x}}_{.jk} &= \frac{1}{n_p} (\mathbf{1}'_{n_p} \otimes \mathbf{I}_{n_r\, n_o}) \mathbf{x}
\end{aligned}
$$

The sum of squares due to interaction between subjects and raters is

$$
\begin{aligned}
\mathrm{SS}_{pr} &= n_o \sum_{i=1}^{n_p} \sum_{j=1}^{n_r} (\bar{x}_{ij.} - \bar{x}_{i..} - \bar{x}_{.j.} + \bar{x}_{...})^2 \\
&= n_o \bar{\mathbf{x}}'_{ij.}\bar{\mathbf{x}}_{ij.} - N\bar{x}^2_{...} - \mathrm{SS}_p - \mathrm{SS}_r \\
&= \frac{1}{n_o} \mathbf{x}' (\mathbf{I}_{n_p\, n_r} \otimes \mathbf{1}_{n_o})(\mathbf{I}_{n_p\, n_r} \otimes \mathbf{1}'_{n_o}) \mathbf{x} - \frac{1}{N} \mathbf{x}' \mathbf{J}_N \mathbf{x} - \mathbf{x}' (\mathbf{I}_{n_p} \otimes \bar{\mathbf{J}}_{n_r\, n_o} - \bar{\mathbf{J}}_N) \mathbf{x} \\
&\quad - \mathbf{x}' (\bar{\mathbf{J}}_{n_p} \otimes \mathbf{I}_{n_r} \otimes \bar{\mathbf{J}}_{n_o} - \bar{\mathbf{J}}_N) \mathbf{x} \\
&= \mathbf{x}' [\mathbf{I}_{n_p\, n_r} \otimes \bar{\mathbf{J}}_{n_o} - \mathbf{I}_{n_p} \otimes \bar{\mathbf{J}}_{n_r\, n_o} - \bar{\mathbf{J}}_{n_p} \otimes \mathbf{I}_{n_r} \otimes \bar{\mathbf{J}}_{n_o} + \bar{\mathbf{J}}_N] \mathbf{x}
\end{aligned}
$$

The sum of squares due to interaction between subjects and occasions is

$$\begin{aligned}
\mathrm{SS}_{po} &= n_r \sum_{i=1}^{n_p} \sum_{k=1}^{n_o} (\bar{x}_{i.k} - \bar{x}_{i..} - \bar{x}_{..k} + \bar{x}_{...})^2 \\
&= n_r \bar{\mathbf{x}}'_{i.k} \bar{\mathbf{x}}_{i.k} - N\bar{x}^2_{...} - \mathrm{SS}_p - \mathrm{SS}_o \\
&= \frac{1}{n_r} \mathbf{x}' \; \{\mathbf{I}_{n_p} \otimes [(\mathbf{1}_{n_r} \otimes \mathbf{I}_{n_o})(\mathbf{1}'_{n_r} \otimes \mathbf{I}_{n_o})]\} \; \mathbf{x} - \mathbf{x}'\bar{\mathbf{J}}_N \mathbf{x} - \mathrm{SS}_p - \mathrm{SS}_o \\
&= \mathbf{x}'(\mathbf{I}_{n_p} \otimes \bar{\mathbf{J}}_{n_r} \otimes \mathbf{I}_{n_o} - \mathbf{I}_{n_p} \otimes \bar{\mathbf{J}}_{n_r n_o} - \bar{\mathbf{J}}_{n_p n_r} \otimes \mathbf{I}_{n_o} + \bar{\mathbf{J}}_N)\mathbf{x}
\end{aligned}$$

The sum of squares due to interaction between raters and occasions is

$$\begin{aligned}
\mathrm{SS}_{ro} &= n_p \sum_{j=1}^{n_r} \sum_{k=1}^{n_o} (\bar{x}_{.jk} - \bar{x}_{.j.} - \bar{x}_{..k} + \bar{x}_{...})^2 \\
&= n_p \bar{\mathbf{x}}'_{.jk} \bar{\mathbf{x}}_{.jk} - N\bar{x}^2_{...} - \mathrm{SS}_r - \mathrm{SS}_o \\
&= \frac{1}{n_p} \mathbf{x}'(\mathbf{1}_{n_p} \otimes \mathbf{I}_{n_r n_o})(\mathbf{1}'_{n_p} \otimes \mathbf{I}_{n_r n_o})\mathbf{x} - \mathbf{x}'\bar{\mathbf{J}}_N \mathbf{x} - \mathrm{SS}_r - \mathrm{SS}_o \\
&= \mathbf{x}'(\bar{\mathbf{J}}_{n_p} \otimes \mathbf{I}_{n_r n_o} - \bar{\mathbf{J}}_{n_p} \otimes \mathbf{I}_{n_r} \otimes \bar{\mathbf{J}}_{n_o} - \bar{\mathbf{J}}_{n_p n_r} \otimes \mathbf{I}_{n_o} + \bar{\mathbf{J}}_N)\mathbf{x}
\end{aligned}$$

The sum of squares due to residual is total sum of squares then subtracting the sum of squares due to all other effects. So it is

$$\begin{aligned}
\mathrm{SS}_E &= \mathrm{SS}_T - \mathrm{SS}_p - \mathrm{SS}_r - \mathrm{SS}_o - \mathrm{SS}_{pr} - \mathrm{SS}_{po} - \mathrm{SS}_{ro} \\
&= \mathbf{x}'[\mathbf{I}_N - \bar{\mathbf{J}}_N - \mathbf{I}_{n_p} \otimes \bar{\mathbf{J}}_{n_r n_o} - \bar{\mathbf{J}}_{n_p} \otimes \mathbf{I}_{n_r} \otimes \bar{\mathbf{J}}_{n_o} - \bar{\mathbf{J}}_{n_p n_r} \otimes \mathbf{I}_{n_o} \\
&\qquad - \mathbf{I}_{n_p n_r} \otimes \bar{\mathbf{J}}_{n_o} + \mathbf{I}_{n_p} \otimes \bar{\mathbf{J}}_{n_r n_o} + \bar{\mathbf{J}}_{n_p} \otimes \mathbf{I}_{n_r} \otimes \bar{\mathbf{J}}_{n_o} \\
&\qquad - \mathbf{I}_{n_p} \otimes \bar{\mathbf{J}}_{n_r} \otimes \mathbf{I}_{n_o} + \mathbf{I}_{n_p} \otimes \bar{\mathbf{J}}_{n_r n_o} + \bar{\mathbf{J}}_{n_p n_r} \otimes \mathbf{I}_{n_o} \\
&\qquad - \bar{\mathbf{J}}_{n_p} \otimes \mathbf{I}_{n_r n_o} + \bar{\mathbf{J}}_{n_p} \otimes \mathbf{I}_{n_r} \otimes \bar{\mathbf{J}}_{n_o} + \bar{\mathbf{J}}_{n_p n_r} \otimes \mathbf{I}_{n_o}]\mathbf{x}
\end{aligned}$$

3.2.2 Expectation of Mean Squares

The expected values of mean squares of three-way random effects model can be derived from the quadratic form of the sum of squares using the following theorem.

Theorem 3.1 *For* $\mathbf{X} \sim (\boldsymbol{\mu}, \mathbf{V})$, *meaning that* $E(\mathbf{X}) = \boldsymbol{\mu}$ *and* $\text{Var}(\mathbf{X}) = \mathbf{V}$,

$$E(\mathbf{X}'\mathbf{A}\mathbf{X}) = \text{tr}(\mathbf{A}\mathbf{V}) + \boldsymbol{\mu}'\mathbf{A}\boldsymbol{\mu}$$

First the variance-covariance matrix $\boldsymbol{\Sigma}$ of the data vector needs to be derived. Under the full three-way random effects model, the variance of X_{ijk} is

$$\sigma_X^2 = \text{Var}(X_{ijk}) = \sigma_p^2 + \sigma_r^2 + \sigma_o^2 + \sigma_{pr}^2 + \sigma_{po}^2 + \sigma_{ro}^2 + \sigma_e^2$$

$\boldsymbol{\Sigma}$ can be partitioned into within subjects variance-covariance matrices $\boldsymbol{\Sigma}_p$ and between subjects variance-covariance matrices $\boldsymbol{\Sigma}_{in}$. To simplify the calculation, $\boldsymbol{\Sigma}_p$ and $\boldsymbol{\Sigma}_{in}$ can be further partitioned into smaller matrices, such as within subjects and raters variance-covariance matrix $\mathbf{V}_d$. Let $\sigma_{p:r}^2 = \sigma_p^2 + \sigma_r^2 + \sigma_{pr}^2$, $\sigma_{p:o}^2 = \sigma_p^2 + \sigma_o^2 + \sigma_{po}^2$ and $\sigma_{r:o}^2 = \sigma_r^2 + \sigma_o^2 + \sigma_{ro}^2$, $\mathbf{V}_d$, $\mathbf{V}_{p:o}$, $\mathbf{V}_{r:o}$ and $\mathbf{V}_o$ be $n_o \times n_o$ symmetric matrices,

$$\mathbf{V}_d = \begin{pmatrix} \sigma_X^2 & \sigma_{p:r}^2 & \cdots & \sigma_{p:r}^2 \\ \sigma_{p:r}^2 & \sigma_X^2 & \cdots & \sigma_{p:r}^2 \\ \vdots & \vdots & \ddots & \vdots \\ \sigma_{p:r}^2 & \sigma_{p:r}^2 & \cdots & \sigma_X^2 \end{pmatrix} \qquad \mathbf{V}_{p:o} = \begin{pmatrix} \sigma_{p:o}^2 & \sigma_p^2 & \cdots & \sigma_p^2 \\ \sigma_p^2 & \sigma_{p:o}^2 & \cdots & \sigma_p^2 \\ \vdots & \vdots & \ddots & \vdots \\ \sigma_p^2 & \sigma_p^2 & \cdots & \sigma_{p:o}^2 \end{pmatrix}$$

$$\mathbf{V}_{r:o} = \begin{pmatrix} \sigma_{r:o}^2 & \sigma_r^2 & \cdots & \sigma_r^2 \\ \sigma_r^2 & \sigma_{r:o}^2 & \cdots & \sigma_r^2 \\ \vdots & \vdots & \ddots & \vdots \\ \sigma_r^2 & \sigma_r^2 & \cdots & \sigma_{r:o}^2 \end{pmatrix} \qquad \mathbf{V}_o = \begin{pmatrix} \sigma_o^2 & 0 & \cdots & 0 \\ 0 & \sigma_o^2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \sigma_o^2 \end{pmatrix}$$

Then

$$\begin{aligned} \boldsymbol{\Sigma}_p &= \mathbf{I}_{n_r} \otimes \mathbf{V}_d + (\mathbf{J}_{n_r} - \mathbf{I}_{n_r}) \otimes \mathbf{V}_{p:o} \\ \boldsymbol{\Sigma}_{in} &= \mathbf{I}_{n_r} \otimes \mathbf{V}_{r:o} + (\mathbf{J}_{n_r} - \mathbf{I}_{n_r}) \otimes \mathbf{V}_o \\ \boldsymbol{\Sigma} &= \mathbf{I}_{n_p} \otimes \boldsymbol{\Sigma}_p + (\mathbf{J}_{n_p} - \mathbf{I}_{n_p}) \otimes \boldsymbol{\Sigma}_{in} \\ &= \mathbf{I}_{n_p} \otimes \Big[\mathbf{I}_{n_r} \otimes \mathbf{V}_d + (\mathbf{J}_{n_r} - \mathbf{I}_{n_r}) \otimes \mathbf{V}_{p:o}\Big] \\ &\quad + (\mathbf{J}_{n_p} - \mathbf{I}_{n_p}) \otimes \Big[\mathbf{I}_{n_r} \otimes \mathbf{V}_{r:o} + (\mathbf{J}_{n_r} - \mathbf{I}_{n_r}) \otimes \mathbf{V}_o\Big] \end{aligned}$$

Now the expectations of mean squares can be calculated. From $\mathrm{SS_p} = \mathbf{x}' \left[\mathbf{I}_{\mathrm{n_p}} \otimes \bar{\mathbf{J}}_{\mathrm{n_r\,n_o}} - \bar{\mathbf{J}}_{\mathrm{N}}\right] \mathbf{x}$, I have $E(\mathrm{SS_p}) = \mathrm{tr}\left((\mathbf{I}_{\mathrm{n_p}} \otimes \bar{\mathbf{J}}_{\mathrm{n_r\,n_o}} - \bar{\mathbf{J}}_{\mathrm{N}})\boldsymbol{\Sigma}\right)$. Hence,

$$
\begin{aligned}
&\mathrm{tr}\left\{(\mathbf{I}_{n_p}\otimes\bar{\mathbf{J}}_{n_r\,n_o})\boldsymbol{\Sigma}\right\}\\
&= \mathrm{tr}\Big\{(\mathbf{I}_{n_p}\otimes\bar{\mathbf{J}}_{n_r\,n_o})\Big[\mathbf{I}_{n_p}\otimes[\mathbf{I}_{n_r} \otimes \mathbf{V}_d + (\mathbf{J}_{n_r}-\mathbf{I}_{n_r})\otimes\mathbf{V}_{p:o}]+\\
&\qquad (\mathbf{J}_{n_p}-\mathbf{I}_{n_p})\otimes[\mathbf{I}_{n_r}\otimes\mathbf{V}_{r:o} + (\mathbf{J}_{n_r}-\mathbf{I}_{n_r})\otimes\mathbf{V}_o]\Big]\Big\}\\
&= n_p\left\{\sigma_X^2 + (n_o-1)\sigma_{p:r}^2 + (n_r-1)\sigma_{p:o}^2 + (n_r-1)(n_o-1)\sigma_p^2\right\},\\
&\mathrm{tr}\left\{\bar{\mathbf{J}}_N\boldsymbol{\Sigma}\right\}\\
&= \mathrm{tr}\Big\{\bar{\mathbf{J}}_N\Big[\mathbf{I}_{n_p}\otimes[\mathbf{I}_{n_r} \otimes \mathbf{V}_d + (\mathbf{J}_{n_r}-\mathbf{I}_{n_r})\otimes\mathbf{V}_{p:o}]\\
&\qquad + (\mathbf{J}_{n_p}-\mathbf{I}_{n_p})\otimes[\mathbf{I}_{n_r}\otimes\mathbf{V}_{r:o} + (\mathbf{J}_{n_r}-\mathbf{I}_{n_r})\otimes\mathbf{V}_o]\Big]\Big\}\\
&= \sigma_X^2 + (n_o-1)\sigma_{p:r}^2 + (n_r-1)\sigma_{p:o}^2 + (n_r-1)(n_o-1)\sigma_p^2 + (n_p-1)\sigma_{r:o}^2\\
&\quad + (n_p-1)(n_o-1)\sigma_r^2 + (n_p-1)(n_r-1)\sigma_o^2,
\end{aligned}
$$

$$
\begin{aligned}
E(\mathrm{SS_p}) &= \mathrm{tr}\left\{(\mathbf{I}_{n_p}\otimes\bar{\mathbf{J}}_{n_r\,n_o})\boldsymbol{\Sigma}\right\} - \mathrm{tr}\left\{\bar{\mathbf{J}}_N\boldsymbol{\Sigma}\right\}\\
&= (n_p-1)\sigma_X^2 + (n_p-1)(n_o-1)\sigma_{p:r}^2 + (n_p-1)(n_r-1)\sigma_{p:o}^2\\
&\quad + (n_p-1)(n_r-1)(n_o-1)\sigma_p^2\\
&\quad - (n_p-1)[\sigma_{r:o}^2 + (n_o-1)\sigma_r^2 + (n_r-1)\sigma_o^2],\\
E(\mathrm{MS_p}) &= E(\mathrm{SS_p})/(\mathrm{n_p}-1)\\
&= \sigma_e^2 + n_r\,n_o\sigma_p^2 + n_o\sigma_{pr}^2 + n_r\sigma_{po}^2.
\end{aligned}
$$

Similar to the calculation of $E(\mathrm{MS_p})$, the expected values of other mean squares can be calculated. From $\mathrm{SS_r} = \mathbf{x}' \left[\bar{\mathbf{J}}_{\mathrm{n_p}} \otimes \mathbf{I}_{\mathrm{n_r}} \otimes \bar{\mathbf{J}}_{\mathrm{n_o}} - \bar{\mathbf{J}}_{\mathrm{N}}\right] \mathbf{x}$,

$$\begin{aligned}
E(\mathrm{SS_r}) &= \mathrm{tr}\left\{(\bar{\mathbf{J}}_{n_p} \otimes \mathbf{I}_{n_r} \otimes \bar{\mathbf{J}}_{n_o})\boldsymbol{\Sigma}\right\} - \mathrm{tr}\left\{\bar{\mathbf{J}}_N \boldsymbol{\Sigma}\right\} \\
&= (n_r-1)\sigma^2_X + (n_r-1)(n_o-1)\sigma^2_{p:r} + (n_p-1)(n_r-1)\sigma^2_{r:o} \\
&\quad + (n_p-1)(n_r-1)(n_o-1)\sigma^2_r - (n_r-1)\sigma^2_{p:o} \\
&\quad - (n_r-1)(n_o-1)\sigma^2_p - (n_p-1)(n_r-1)\sigma^2_o, \\
E(\mathrm{MS_r}) &= E(\mathrm{SS_r})/(\mathrm{n_r}-1) \\
&= \sigma^2_e + n_p n_o \sigma^2_r + n_o \sigma^2_{pr} + n_p \sigma^2_{ro}.
\end{aligned}$$

From $\mathrm{SS_o} = \mathbf{x}'\left[\bar{\mathbf{J}}_{\mathrm{n_p\,n_r}} \otimes \mathbf{I}_{\mathrm{n_o}} - \bar{\mathbf{J}}_{\mathrm{N}}\right]\mathbf{x}$,

$$\begin{aligned}
E(\mathrm{SS_o}) &= \mathrm{tr}\left\{(\bar{\mathbf{J}}_{n_p\,n_r} \otimes \mathbf{I}_{n_o})\boldsymbol{\Sigma}\right\} - \mathrm{tr}\left\{\bar{\mathbf{J}}_{\mathbf{N}} \boldsymbol{\Sigma}\right\} \\
&= (n_o-1)\sigma^2_X - (n_o-1)\sigma^2_{p:r} + (n_p-1)(n_o-1)\sigma^2_{r:o} \\
&\quad - (n_p-1)(n_o-1)\sigma^2_r + (n_r-1)(n_o-1)\sigma^2_{p:o} \\
&\quad - (n_r-1)(n_o-1)\sigma^2_p + (n_p-1)(n_r-1)(n_o-1)\sigma^2_o, \\
E(\mathrm{MS_o}) &= E(\mathrm{SS_o})/(\mathrm{n_o}-1) \\
&= \sigma^2_e + n_p n_r \sigma^2_o + n_r \sigma^2_{po} + n_p \sigma^2_{ro}.
\end{aligned}$$

From $\mathrm{SS_{pr}} = \mathbf{x}'\left[\mathbf{I}_{\mathrm{n_p\,n_r}} \otimes \bar{\mathbf{J}}_{\mathrm{n_o}} - \mathbf{I}_{\mathrm{n_p}} \otimes \bar{\mathbf{J}}_{\mathrm{n_r\,n_o}} - \bar{\mathbf{J}}_{\mathrm{n_p}} \otimes \mathbf{I}_{\mathrm{n_r}} \otimes \bar{\mathbf{J}}_{\mathrm{n_o}} + \bar{\mathbf{J}}_{\mathrm{N}}\right]\mathbf{x}$,

$$\begin{aligned}
E(\mathrm{SS_{pr}}) &= \mathrm{tr}\left\{(\mathbf{I}_{n_p\,n_r} \otimes \bar{\mathbf{J}}_{n_o} - \mathbf{I}_{n_p} \otimes \bar{\mathbf{J}}_{n_r\,n_o} - \bar{\mathbf{J}}_{n_p} \otimes \mathbf{I}_{n_r} \otimes \bar{\mathbf{J}}_{n_o} + \bar{\mathbf{J}}_N)\boldsymbol{\Sigma}\right\} \\
&= (n_p-1)(n_r-1)\sigma^2_X + (n_p-1)(n_r-1)(n_o-1)\sigma^2_{p:r} \\
&\quad - (n_p-1)(n_r-1)\sigma^2_{p:o} - (n_p-1)(n_r-1)\sigma^2_{r:o} \\
&\quad - (n_p-1)(n_r-1)(n_o-1)\sigma^2_p - (n_p-1)(n_r-1)(n_o-1)\sigma^2_r \\
&\quad + (n_p-1)(n_r-1)\sigma^2_o, \\
E(\mathrm{MS_{pr}}) &= E(\mathrm{SS_{pr}})/\left[(\mathrm{n_p}-1)(\mathrm{n_r}-1)\right] \\
&= \sigma^2_e + n_o \sigma^2_{pr}.
\end{aligned}$$

From $\mathrm{SS_{po}} = \mathbf{x}'\big[\mathbf{I}_{\mathrm{n_p}} \otimes \bar{\mathbf{J}}_{\mathrm{n_r}} \otimes \mathbf{I}_{\mathrm{n_o}} - \mathbf{I}_{\mathrm{n_p}} \otimes \bar{\mathbf{J}}_{\mathrm{n_r\,n_o}} - \bar{\mathbf{J}}_{\mathrm{n_p\,n_r}} \otimes \mathbf{I}_{\mathrm{n_r}} + \bar{\mathbf{J}}_{\mathrm{N}}\big]\mathbf{x}$,

$$\begin{aligned} E(\mathrm{SS_{po}}) = \mathrm{tr}\big\{&(\mathbf{I}_{n_p} \otimes \bar{\mathbf{J}}_{n_r} \otimes \mathbf{I}_{n_o} - \mathbf{I}_{n_p} \otimes \bar{\mathbf{J}}_{n_r\,n_o} - \bar{\mathbf{J}}_{n_p\,n_r} \otimes \mathbf{I}_{n_r} + \bar{\mathbf{J}}_N)\mathbf{\Sigma}\big\} \\ &+ (n_p-1)(n_r-1)(n_o-1)\sigma^2_{p:o} \\ &- (n_p-1)(n_o-1)\sigma^2_{r:o} - (n_p-1)(n_r-1)(n_o-1)\sigma^2_p \\ &+ (n_p-1)(n_o-1)\sigma^2_r - (n_p-1)(n_r-1)n_o-1)\sigma^2_o, \\ E(\mathrm{MS_{po}}) &= \mathrm{SS_{po}}/\left[(\mathrm{n_p}-1)(\mathrm{n_o}-1)\right] \\ &= \sigma^2_e + n_r\sigma^2_{po}. \end{aligned}$$

From $\mathrm{SS_{ro}} = \mathbf{x}'\big[\bar{\mathbf{J}}_{\mathrm{n_p}} \otimes \mathbf{I}_{\mathrm{n_r\,n_o}} - \bar{\mathbf{J}}_{\mathrm{n_p}} \otimes \mathbf{I}_{\mathrm{n_r}} \otimes \bar{\mathbf{J}}_{\mathrm{n_o}} - \bar{\mathbf{J}}_{\mathrm{n_p\,n_r}} \otimes \mathbf{I}_{\mathrm{n_r}} + \bar{\mathbf{J}}_{\mathrm{N}}\big]\mathbf{x}$,

$$\begin{aligned} E(\mathrm{SS_{ro}}) &= \mathrm{tr}\big\{(\bar{\mathbf{J}}_{n_p} \otimes \mathbf{I}_{n_r\,n_o} - \bar{\mathbf{J}}_{n_p} \otimes \mathbf{I}_{n_r} \otimes \bar{\mathbf{J}}_{n_o} - \bar{\mathbf{J}}_{n_p\,n_r} \otimes \mathbf{I}_{n_r} + \bar{\mathbf{J}}_N)\mathbf{\Sigma}\big\} \\ &= (n_r-1)(n_o-1)\sigma^2_X - (n_r-1)(n_o-1)\sigma^2_{p:r} - (n_r-1)(n_o-1)\sigma^2_{p:o} \\ &\quad + (n_p-1)(n_r-1)(n_o-1)\sigma^2_{r:o} + (n_r-1)(n_o-1)\sigma^2_p \\ &\quad - (n_p-1)(n_r-1)(n_o-1)\sigma^2_r - (n_p-1)(n_r-1)(n_o-1)\sigma^2_o, \\ E(\mathrm{MS_{ro}}) &= E(\mathrm{SS_{ro}})/\left[(\mathrm{n_r}-1)(\mathrm{n_o}-1)\right] \\ &= \sigma^2_e + n_p\sigma^2_{ro}. \end{aligned}$$

From $\mathrm{SS_T} = \mathbf{x}'\big[\mathbf{I}_{\mathrm{N}} - \bar{\mathbf{J}}_{\mathrm{N}}\big]\mathbf{x}$,

$$\begin{aligned} E(\mathrm{SS_T}) = \mathrm{tr}\big\{&(\mathbf{I}_N - \bar{\mathbf{J}}_N)\mathbf{\Sigma}\big\} \\ = (N&-1)\sigma^2_e + [N - n_r - n_o + 1 - (n_r-1)(n_o-1)]\sigma^2_p \\ &+ [N - n_p - n_o + 1 - (n_p-1)(n_o-1)]\sigma^2_r \\ &+ [N - n_r - n_p + 1 - (n_p-1)(n_r-1)]\sigma^2_o \\ &+ [N - n_o]\sigma^2_{pr} + [N - n_r]\sigma^2_{po} + [N - n_p]\sigma^2_{ro}. \end{aligned}$$

Then

$$
\begin{aligned}
E(\mathrm{SS_e}) &= E(\mathrm{SS_T} - \mathrm{SS_p} - \mathrm{SS_r} - \mathrm{SS_o} - \mathrm{SS_{pr}} - \mathrm{SS_{po}} - \mathrm{SS_{ro}}) \\
&= (n_p - 1)(n_r - 1)(n_o - 1)\sigma_e^2 \\
&\quad + [N - n_r - n_o + 1 - (n_r - 1)(n_o - 1) - (n_p - 1)n_r n_o]\sigma_p^2 \\
&\quad + [N - n_p - n_o + 1 - (n_p - 1)(n_o - 1) - (n_r - 1)n_p n_o]\sigma_r^2 \\
&\quad + [N - n_r - n_p + 1 - (n_p - 1)(n_r - 1) - (n_o - 1)n_p n_r]\sigma_o^2 \\
&\quad + [N - n_o - (n_p - 1)n_o - (n_r - 1)n_o - (n_p - 1)(n_r - 1)n_o]\sigma_{pr}^2 \\
&\quad + [N - n_r - (n_p - 1)n_r - (n_o - 1)n_r - (n_p - 1)(n_o - 1)n_r]\sigma_{po}^2 \\
&\quad + [N - n_p - (n_r - 1)n_p - (n_o - 1)n_p - (n_r - 1)(n_o - 1)n_p]\sigma_{ro}^2 \\
&= (n_p - 1)(n_r - 1)(n_o - 1)\sigma_e^2, \\
E(\mathrm{MS_e}) &= \sigma_e^2.
\end{aligned}
$$

3.3 Intraclass Correlation Coefficient for Full Model

Table 3.1 is the ANOVA table for the full three-way random effects model. From the results in the previous section, the expected values of mean squares and the unbiased estimators of the variance components can be found. Based on the unbiased estimators of the variance components, the two rater reliability indices—intraclass correlation coefficient and interrater reliability coefficient—are estimable. Table 3.2 lists unbiased estimators of the variance components.

Because

$$
\begin{aligned}
\mathrm{Cov}(x_{ijk}, x_{ij'k'}) &= \sigma_p^2 \\
\mathrm{Var}(x_{ijk}) = \mathrm{Var}(x_{ij'k'}) &= \sigma_p^2 + \sigma_r^2 + \sigma_o^2 + \sigma_{pr}^2 + \sigma_{po}^2 + \sigma_{ro}^2 + \sigma_e^2
\end{aligned}
$$

according to Definition 1.1, ICC for the three-way random effects model is defined as

$$
\rho_I = \frac{\sigma_p^2}{\sigma_p^2 + \sigma_r^2 + \sigma_o^2 + \sigma_{pr}^2 + \sigma_{po}^2 + \sigma_{ro}^2 + \sigma_e^2}. \tag{3.2}
$$

Table 3.1: ANOVA table for full three-way model

Source of variation	df	MS	EMS
Subjects	$n_p - 1$	$\mathrm{MS_p}$	$\sigma_e^2 + n_r\sigma_{po}^2 + n_o\sigma_{pr}^2 + n_r n_o\sigma_p^2$
Raters	$n_r - 1$	$\mathrm{MS_r}$	$\sigma_e^2 + n_p\sigma_{ro}^2 + n_o\sigma_{pr}^2 + n_p n_o\sigma_r^2$
Occasions	$n_o - 1$	$\mathrm{MS_o}$	$\sigma_e^2 + n_p\sigma_{ro}^2 + n_r\sigma_{po}^2 + n_p n_r\sigma_o^2$
Subjects×Raters	$(n_p - 1)(n_r - 1)$	$\mathrm{MS_{pr}}$	$\sigma_e^2 + n_o\sigma_{pr}^2$
Subjects×Occasions	$(n_p - 1)(n_o - 1)$	$\mathrm{MS_{po}}$	$\sigma_e^2 + n_r\sigma_{po}^2$
Raters×Occasions	$(n_r - 1)(n_o - 1)$	$\mathrm{MS_{ro}}$	$\sigma_e^2 + n_p\sigma_{ro}^2$
Residual	$(n_p - 1)(n_r - 1)(n_o - 1)$	$\mathrm{MS_e}$	σ_e^2

Table 3.2: Unbiased estimators of variance components for full model

Variance Component	Unbiased Estimator
σ_p^2	$\frac{1}{n_r n_o}(\mathrm{MS_p} + \mathrm{MS_e} - \mathrm{MS_{pr}} - \mathrm{MS_{po}})$
σ_r^2	$\frac{1}{n_p n_o}(\mathrm{MS_r} + \mathrm{MS_e} - \mathrm{MS_{pr}} - \mathrm{MS_{ro}})$
σ_o^2	$\frac{1}{n_p n_r}(\mathrm{MS_o} + \mathrm{MS_e} - \mathrm{MS_{ro}} - \mathrm{MS_{po}})$
σ_{pr}^2	$\frac{1}{n_o}(\mathrm{MS_{pr}} - \mathrm{MS_e})$
σ_{po}^2	$\frac{1}{n_r}(\mathrm{MS_{po}} - \mathrm{MS_e})$
σ_{ro}^2	$\frac{1}{n_p}(\mathrm{MS_{ro}} - \mathrm{MS_e})$
σ_e^2	$\mathrm{MS_e}$

An estimator of ρ_I is

$$
\begin{aligned}
r_I = &\left\{\frac{1}{n_r n_o}\mathrm{MS_p} + \frac{1}{\mathrm{n_r n_o}}(\mathrm{MS_e} - \mathrm{MS_{pr}} - \mathrm{MS_{po}})\right\} \Big/ \left\{\frac{1}{\mathrm{n_r n_o}}\mathrm{MS_p}\right. \\
&+ \frac{1}{n_p n_o}\mathrm{MS_r} + \frac{1}{\mathrm{n_p n_r}}\mathrm{MS_o} + (\frac{1}{\mathrm{n_o}} - \frac{1}{\mathrm{n_r n_o}} - \frac{1}{\mathrm{n_p n_o}})\mathrm{MS_{pr}} \\
&+ (\frac{1}{n_r} - \frac{1}{n_r n_o} - \frac{1}{n_p n_r})\mathrm{MS_{po}} + (\frac{1}{\mathrm{n_p}} - \frac{1}{\mathrm{n_p n_o}} - \frac{1}{\mathrm{n_p n_r}})\mathrm{MS_{ro}} \\
&\left. + (1 + \frac{1}{n_r n_o} + \frac{1}{n_p n_o} + \frac{1}{n_p n_r} - \frac{1}{n_p} - \frac{1}{n_r} - \frac{1}{n_o})\mathrm{MS_e}\right\}
\end{aligned}
\quad (3.3)
$$

which is the ratio of an unbiased estimator of the numerator of ρ_I to an unbiased estimator of the denominator of ρ_I. This estimator is consistent but biased. The proof of consistency parallels that for the one-way model.

Next I derive a $100(1-\alpha)\%$ confidence interval for ρ_I. First from Definition (3.2), I have

$$
\begin{aligned}
1 - \rho_I &= \frac{\sigma_r^2 + \sigma_o^2 + \sigma_{pr}^2 + \sigma_{po}^2 + \sigma_{ro}^2 + \sigma_e^2}{\sigma_p^2 + \sigma_r^2 + \sigma_o^2 + \sigma_{pr}^2 + \sigma_{po}^2 + \sigma_{ro}^2 + \sigma_e^2} \\
\frac{\rho_I}{1 - \rho_I} &= \frac{\sigma_p^2}{\sigma_r^2 + \sigma_o^2 + \sigma_{pr}^2 + \sigma_{po}^2 + \sigma_{ro}^2 + \sigma_e^2} \\
\sigma_p^2 &= \frac{\rho_I}{1 - \rho_I}(\sigma_r^2 + \sigma_o^2 + \sigma_{pr}^2 + \sigma_{po}^2 + \sigma_{ro}^2 + \sigma_e^2). \quad (3.4)
\end{aligned}
$$

Substituting σ_p^2 from Equation (3.4) in the expectation of $\mathrm{MS_p}$ in Table 3.1 gives

$$
\begin{aligned}
E(\mathrm{MS_p}) &= n_r n_o \sigma_p^2 + n_o \sigma_{pr}^2 + n_r \sigma_{po}^2 + \sigma_e^2 \\
&= n_r n_o \frac{\rho_I}{1 - \rho_I}\left(\sigma_r^2 + \sigma_o^2 + \sigma_{pr}^2 + \sigma_{po}^2 + \sigma_{ro}^2 + \sigma_e^2\right) + n_o \sigma_{pr}^2 + n_r \sigma_{po}^2 + \sigma_e^2 \\
&= n_r n_o \frac{\rho_I}{1 - \rho_I}\sigma_r^2 + n_r n_o \frac{\rho_I}{1 - \rho_I}\sigma_o^2 + \left(n_r n_o \frac{\rho_I}{1 - \rho_I} + n_o\right)\sigma_{pr}^2 \\
&\quad + \left(n_r n_o \frac{\rho_I}{1 - \rho_I} + n_r\right)\sigma_{po}^2 + n_r n_o \frac{\rho_I}{1 - \rho_I}\sigma_{ro}^2 + \left(1 + n_r n_o \frac{\rho_I}{1 - \rho_I}\right)\sigma_e^2.
\end{aligned}
$$

Now substituting the variance components with their unbiased estimators in the above equation, I define a variable V so that $E(V) = E(\mathrm{MS_p})$.

$$
\begin{aligned}
V = & \frac{n_r n_o}{n_p n_o}\frac{\rho_I}{1-\rho_I}(\mathrm{MS_r} + \mathrm{MS_e} - \mathrm{MS_{pr}} - \mathrm{MS_{ro}}) \\
& + \frac{n_r n_o}{n_p n_r}\frac{\rho_I}{1-\rho_I}(\mathrm{MS_o} + \mathrm{MS_e} - \mathrm{MS_{ro}} - \mathrm{MS_{po}}) \\
& + (1 + n_r\frac{\rho_I}{1-\rho_I})(\mathrm{MS_{pr}} - \mathrm{MS_e}) + (1 + \mathrm{n_o}\frac{\rho_\mathrm{I}}{1-\rho_\mathrm{I}})(\mathrm{MS_{po}} - \mathrm{MS_e}) \\
& + \frac{n_r n_o}{n_p}\frac{\rho_I}{1-\rho_I}(\mathrm{MS_{ro}} - \mathrm{MS_e}) + (1 + \mathrm{n_r n_o}\frac{\rho_\mathrm{I}}{1-\rho_\mathrm{I}})\mathrm{MS_e} \\
= & \frac{n_r}{n_p}\frac{\rho_I}{1-\rho_I}\mathrm{MS_r} + \frac{\mathrm{n_o}}{\mathrm{n_p}}\frac{\rho_\mathrm{I}}{1-\rho_\mathrm{I}}\mathrm{MS_o} + (1 + \mathrm{n_r}\frac{\rho_\mathrm{I}}{1-\rho_\mathrm{I}} - \frac{\mathrm{n_r}}{\mathrm{n_p}}\frac{\rho_\mathrm{I}}{1-\rho_\mathrm{I}})\mathrm{MS_{pr}} \\
& + (1 + n_o\frac{\rho_I}{1-\rho_I} - \frac{n_o}{n_p}\frac{\rho_I}{1-\rho_I})\mathrm{MS_{po}} + \frac{\rho_\mathrm{I}}{1-\rho_\mathrm{I}}(\frac{\mathrm{n_r n_o}}{\mathrm{n_p}} - \frac{\mathrm{n_r}}{\mathrm{n_p}} - \frac{\mathrm{n_o}}{\mathrm{n_p}})\mathrm{MS_{ro}} \\
& + (\frac{n_r}{n_p}\frac{\rho_I}{1-\rho_I} + \frac{n_o}{n_p}\frac{\rho_I}{1-\rho_I} - 1 - n_r\frac{\rho_I}{1-\rho_I} - 1 - n_o\frac{\rho_I}{1-\rho_I} - \frac{n_r n_o}{n_p}\frac{\rho_I}{1-\rho_I} + 1 \\
& \quad + n_r n_o\frac{\rho_I}{1-\rho_I})\mathrm{MS_e} \\
= & \frac{\rho_I}{1-\rho_I}\left[\frac{n_r}{n_p}\mathrm{MS_r} + \frac{\mathrm{n_o}}{\mathrm{n_p}}\mathrm{MS_o} + \left(\mathrm{n_r} - \frac{\mathrm{n_r}}{\mathrm{n_p}}\right)\mathrm{MS_{pr}} + \left(\mathrm{n_o} - \frac{\mathrm{n_o}}{\mathrm{n_p}}\right)\mathrm{MS_{po}} \right. \\
& \left. + \left(\frac{n_r n_o - n_r - n_o}{n_p}\right)\mathrm{MS_{ro}} + \left(\frac{\mathrm{n_r + n_o - n_r n_p - n_o n_p - n_r n_o + n_p n_r n_o}}{\mathrm{n_p}}\right)\mathrm{MS_e}\right] \\
& + \mathrm{MS_{pr}} + \mathrm{MS_{po}} - \mathrm{MS_e}.
\end{aligned}
$$

V is a linear combination of independent mean squares: $\mathrm{MS_r}$, $\mathrm{MS_o}$, $\mathrm{MS_{pr}}$, $\mathrm{MS_{po}}$, $\mathrm{MS_{ro}}$ and $\mathrm{MS_e}$. V is independent of $\mathrm{MS_p}$. According to Satterthwaite's approximation, V is approximately distributed as $d\chi^2_\nu/\nu$ with $d = E(V)$ and degrees of freedom

$$
\nu = \frac{(e\mathrm{MS_r} + \mathrm{fMS_o} + \mathrm{gMS_{pr}} + \mathrm{hMS_{po}} + \mathrm{iMS_{ro}} + \mathrm{jMS_e})^2}{\frac{(e\mathrm{MS_r})^2}{n_r-1} + \frac{(f\mathrm{MS_o})^2}{n_o-1} + \frac{(g\mathrm{MS_{pr}})^2}{(n_p-1)(n_r-1)} + \frac{(h\mathrm{MS_{po}})^2}{(n_p-1)(n_o-1)} + \frac{(i\mathrm{MS_{ro}})^2}{(n_r-1)(n_o-1)} + \frac{(j\mathrm{MS_e})^2}{(n_p-1)(n_r-1)(n_o-1)}}
$$

where

$$\begin{aligned}
e &= \frac{n_r}{n_p}\frac{\rho_I}{1-\rho_I}, \\
f &= \frac{n_o}{n_p}\frac{\rho_I}{1-\rho_I}, \\
g &= 1 + n_r\frac{\rho_I}{1-\rho_I} - \frac{n_r}{n_p}\frac{\rho_I}{1-\rho_I}, \\
h &= 1 + n_o\frac{\rho_I}{1-\rho_I} - \frac{n_o}{n_p}\frac{\rho_I}{1-\rho_I}, \\
i &= \frac{n_r n_o - n_r - n_o}{n_p}\frac{\rho_I}{1-\rho_I}, \\
j &= \left(\frac{n_r}{n_p} + \frac{n_o}{n_p} - n_r - n_o - \frac{n_r n_o}{n_p} + n_r n_o\right)\frac{\rho_I}{1-\rho_I} - 1.
\end{aligned}$$

To estimate ν, use r_I to replace ρ_I.

Then

$$\begin{aligned}
F &= \mathrm{MS_p/V} \\
&= \mathrm{MS_p/(eMS_r + fMS_o + gMS_{pr} + hMS_{po} + iMS_{ro} + jMS_e)}
\end{aligned}$$

has an approximate F distribution with degrees of freedom $(n_p - 1)$ and ν. Therefore

$$Pr\{F_{\frac{\alpha}{2}}(n_p - 1, \nu) \leq \mathrm{MS_p/V} \leq \mathrm{F}_{1-\frac{\alpha}{2}}(\mathrm{n_p} - 1, \nu)\} = 1 - \alpha,$$

where $F_{1-\frac{\alpha}{2}}(n_p - 1, \nu)$ is the $(1 - \alpha/2) \times 100$th percentile of an F distribution with degrees of freedom $(n_p - 1)$ and ν.

From the above formula, the lower and upper limits of a confidence interval for ρ_I are estimable. Let $\rho_I^* = \rho_I/(1 - \rho_I)$, $F_* = F_{1-\frac{\alpha}{2}}(n_p - 1, \nu)$ and $q = (n_r/n_p)\mathrm{MS_r} + (\mathrm{n_o/n_p})\mathrm{MS_o} + (\mathrm{n_r - n_r/n_p})\mathrm{MS_{pr}} + (\mathrm{n_o - n_o/n_p})\mathrm{MS_{po}} + [(\mathrm{n_r n_o - n_r - n_o})/\mathrm{n_p}]\mathrm{MS_{ro}} + [(\mathrm{n_r + n_o - n_p n_r - n_p n_o - n_r n_o + n_p n_r n_o})/\mathrm{n_p}]\mathrm{MS_e}$,

$$\begin{aligned}
\frac{\mathrm{MS_p}}{V} &\le F_{1-\frac{\alpha}{2}}(n_p-1,\nu), \\
\frac{\mathrm{MS_p}}{F_*} &\le \rho_I^* q + \mathrm{MS_{pr}} + \mathrm{MS_{po}} - \mathrm{MS_e}, \\
\rho_I^* q &\ge \frac{\mathrm{MS_p}}{F_*} - \mathrm{MS_{pr}} - \mathrm{MS_{po}} + \mathrm{MS_e}, \\
\rho_I^* &\ge \frac{\mathrm{MS_p/F_*} - \mathrm{MS_{pr}} - \mathrm{MS_{po}} + \mathrm{MS_e}}{q}, \\
\rho_I &\ge \frac{(\mathrm{MS_p/F_*} - \mathrm{MS_{pr}} - \mathrm{MS_{po}} + \mathrm{MS_e})/\mathrm{q}}{1 + (\mathrm{MS_p/F_*} - \mathrm{MS_{pr}} - \mathrm{MS_{po}} + \mathrm{MS_e})/\mathrm{q}} \\
&= \frac{\mathrm{MS_p/F_*} - \mathrm{MS_{pr}} - \mathrm{MS_{po}} + \mathrm{MS_e}}{q + (\mathrm{MS_p/F_*} - \mathrm{MS_{pr}} - \mathrm{MS_{po}} + \mathrm{MS_e})} \\
&= \frac{\mathrm{MS_p} - \mathrm{F_*}(\mathrm{MS_{pr}} + \mathrm{MS_{po}} - \mathrm{MS_e})}{F_*(q - \mathrm{MS_{pr}} - \mathrm{MS_{po}} + \mathrm{MS_e}) + \mathrm{MS_p}}.
\end{aligned}$$

Similarly, the upper limit can be derived. Let $F_{1-\frac{\alpha}{2}}(\nu, n_p-1) = F^*$,

$$\begin{aligned}
\frac{\mathrm{MS_p}}{V} &\ge F_{\frac{\alpha}{2}}(n_p-1,\nu), \\
\rho_I &\le \frac{(\mathrm{MS_p/F_{\frac{\alpha}{2}}(n_p-1,\nu)} - \mathrm{MS_{pr}} - \mathrm{MS_{po}} + \mathrm{MS_e})/\mathrm{q}}{1 + (\mathrm{MS_p/F_{\frac{\alpha}{2}}(n_p-1,\nu)} - \mathrm{MS_{pr}} - \mathrm{MS_{po}} + \mathrm{MS_e})/\mathrm{q}} \\
&= \frac{F^*\mathrm{MS_p} - \mathrm{MS_{pr}} - \mathrm{MS_{po}} + \mathrm{MS_e}}{q + (F^*\mathrm{MS_p} - \mathrm{MS_{pr}} - \mathrm{MS_{po}} + \mathrm{MS_e})}.
\end{aligned}$$

So an approximate $100(1-\alpha)\%$ confidence interval for ρ_I is

$$\left[\frac{\mathrm{MS_p} - \hat{\mathrm{F}}_*(\mathrm{MS_{pr}} + \mathrm{MS_{po}} - \mathrm{MS_e})}{\hat{F}_*(q - \mathrm{MS_{pr}} - \mathrm{MS_{po}} + \mathrm{MS_e}) + \mathrm{MS_p}}, \frac{\hat{F}^*\mathrm{MS_p} - \mathrm{MS_{pr}} - \mathrm{MS_{po}} + \mathrm{MS_e}}{q + (\hat{F}^*\mathrm{MS_p} - \mathrm{MS_{pr}} - \mathrm{MS_{po}} + \mathrm{MS_e})}\right] \tag{3.5}$$

where $\hat{F}_* = F_{1-\frac{\alpha}{2}}(n_p-1,\hat{\nu})$ and $\hat{F}^* = F_{1-\frac{\alpha}{2}}(\hat{\nu}, n_p-1)$.

3.4 Interrater Reliability for Full Model

The *interrater reliability coefficient* for three-way random effects model is defined by

$$\rho_e = \frac{\sigma_p^2}{\sigma_p^2 + \sigma_{pr}^2 + \sigma_{ro}^2 + \sigma_e^2} \tag{3.6}$$

comparable to definition of ρ_I in Equation (3.2). The variance component σ_{po}^2 is omitted from the denominator of ρ_I because an observer examines the same subject on all occasions and this should have no impact on the estimate of interrater reliability. The variance component σ_r^2 is omitted because *a priori* not all raters are alike and this component is, therefore, permitted to be nonzero. The variance component σ_o^2 is omitted because conditions may vary from occasion to occasion which may uniformly affect subjects' outcomes.

An estimator of ρ_e is

$$\begin{aligned} r_e = &\left\{\frac{1}{n_r n_o}\mathrm{MS_p} + \frac{1}{\mathrm{n_r n_o}}(\mathrm{MS_e} - \mathrm{MS_{pr}} - \mathrm{MS_{po}})\right\}\Big/\left\{\frac{1}{\mathrm{n_r n_o}}\mathrm{MS_p}\right. \\ &+ \left(\frac{1}{n_o} - \frac{1}{n_r n_o}\right)\mathrm{MS_{pr}} - \frac{1}{\mathrm{n_r n_o}}\mathrm{MS_{po}} + \frac{1}{\mathrm{n_p}}\mathrm{MS_{ro}} \\ &\left. + \left(1 + \frac{1}{n_r n_o} - \frac{1}{n_p} - \frac{1}{n_o}\right)\mathrm{MS_e}\right\} \end{aligned} \tag{3.7}$$

which is the ratio of an unbiased estimator of the numerator of ρ_e to an unbiased estimator of the denominator of ρ_e. This estimator is consistent but biased. The proof of consistency parallels that for the one-way model.

Follow the same theme of the previous section, I derive a $100(1-\alpha)\%$ confidence interval for ρ_e. First, from definition (3.6), I get

$$\begin{aligned} 1 - \rho_e &= \frac{\sigma_{pr}^2 + \sigma_{ro}^2 + \sigma_e^2}{\sigma_p^2 + \sigma_{pr}^2 + \sigma_{ro}^2 + \sigma_e^2} \\ \frac{\rho_e}{1-\rho_e} &= \frac{\sigma_p^2}{\sigma_{pr}^2 + \sigma_{ro}^2 + \sigma_e^2} \\ \sigma_p^2 &= \frac{\rho_e}{1-\rho_e}(\sigma_{pr}^2 + \sigma_{ro}^2 + \sigma_e^2) \\ &= \rho_e^*(\sigma_{pr}^2 + \sigma_{ro}^2 + \sigma_e^2). \end{aligned} \tag{3.8}$$

where $\rho_e^* = \rho_e/(1-\rho_e)$. Substituting σ_p^2 from Equation (3.8) in the expectation of $\mathrm{MS_p}$ in Table 3.1 gives

$$\begin{aligned} E(\mathrm{MS_p}) &= n_r n_o \sigma_p^2 + n_o \sigma_{pr}^2 + n_r \sigma_{po}^2 + \sigma_e^2 \\ &= n_r n_o \rho_e^* (\sigma_{pr}^2 + \sigma_{ro}^2 + \sigma_e^2) + n_o \sigma_{pr}^2 + n_r \sigma_{po}^2 + \sigma_e^2 \\ &= (n_r n_o \rho_e^* + n_o)\sigma_{pr}^2 + n_r \sigma_{po}^2 + n_r n_o \rho_e^* \sigma_{ro}^2 + (1 + n_r n_o \rho_e^*)\sigma_e^2. \end{aligned}$$

Next, substituting the variance components with their unbiased estimators in the above equation, I define a variable U so that $E(U) = E(\mathrm{MS_p})$.

$$\begin{aligned} U &= (1 + n_r \rho_e^*)(\mathrm{MS_{pr}} - \mathrm{MS_e}) + (\mathrm{MS_{po}} - \mathrm{MS_e}) \\ &\quad + \frac{n_r n_o}{n_p} \rho_e^* (\mathrm{MS_{ro}} - \mathrm{MS_e}) + (1 + \mathrm{n_r n_o} \rho_\mathrm{e}^*)\mathrm{MS_e} \\ &= (1 + n_r \rho_e^*)\mathrm{MS_{pr}} + \mathrm{MS_{po}} + \frac{\mathrm{n_r n_o}}{\mathrm{n_p}} \rho_\mathrm{e}^* \mathrm{MS_{ro}} \\ &\quad + \left(-1 - n_r \rho_e^* - \frac{n_r n_o}{n_p} \rho_e^* + n_r n_o \rho_e^*\right) \mathrm{MS_e} \\ &= \rho_e^* \left[n_r \mathrm{MS_{pr}} + \frac{\mathrm{n_r n_o}}{\mathrm{n_p}} \mathrm{MS_{ro}} + \frac{\mathrm{n_p n_r n_o - n_p n_r - n_r n_o}}{\mathrm{n_p}} \mathrm{MS_e} \right] \\ &\quad + \mathrm{MS_{pr}} + \mathrm{MS_{po}} - \mathrm{MS_e}. \end{aligned}$$

U is a linear combination of independent mean squares: $\mathrm{MS_{pr}}$, $\mathrm{MS_{po}}$, $\mathrm{MS_{ro}}$ and $\mathrm{MS_e}$. U is independent of $\mathrm{MS_p}$. According to Satterthwaite's approximation, U is approximately distributed as $b\chi_\nu^2/\xi$ with $b = E(U)$ and degrees of freedom

$$\xi = \frac{(l_1 \mathrm{MS_{pr}} + \mathrm{l_2 MS_{po}} + \mathrm{l_3 MS_{ro}} + \mathrm{l_4 MS_e})^2}{\frac{(l_1 \mathrm{MS_{pr}})^2}{(n_p-1)(n_r-1)} + \frac{(l_2 \mathrm{MS_{po}})^2}{(n_p-1)(n_o-1)} + \frac{(l_3 \mathrm{MS_{ro}})^2}{(n_r-1)(n_o-1)} + \frac{(l_4 \mathrm{MS_e})^2}{(n_p-1)(n_r-1)(n_o-1)}}$$

where

$$\begin{aligned} l_1 &= 1 + n_r \rho_e^*, \\ l_2 &= 1, \\ l_3 &= \frac{n_r n_o}{n_p} \rho_e^*, \\ l_4 &= \left(-n_r - \frac{n_r n_o}{n_p} + n_r n_o\right) \rho_e^* - 1. \end{aligned}$$

To estimate ξ, use r_e to replace ρ_e.

Then

$$F = \mathrm{MS_p/U}$$
$$= \mathrm{MS_p/(l_1MS_{pr} + l_2MS_{po} + l_3MS_{ro} + l_4MS_e)}$$

has an approximate F distribution with degrees of freedom $(n_p - 1)$ and ξ. Therefore

$$Pr\left\{F_{\frac{\alpha}{2}}(n_p - 1, \xi) \leq \mathrm{MS_p/U} \leq \mathrm{F_{1-\frac{\alpha}{2}}(n_p - 1, \xi)}\right\} = 1 - \alpha,$$

where $F_{1-\frac{\alpha}{2}}(n_p - 1, \xi)$ is the $(1 - \alpha/2) \times 100$th percentile of an F distribution with degrees of freedom $(n_p - 1)$ and ξ.

From the above formula, the lower and upper limits of a confidence interval for ρ_e are estimable. Let $F_{*\xi} = F_{1-\frac{\alpha}{2}}(n_p - 1, \xi)$ and $q_* = n_r\mathrm{MS_{pr}} + \mathrm{[(n_rn_o)/n_p]MS_{ro} + [(n_pn_rn_o - n_pn_r - n_rn_o)/n_p)]MS_e}$,

$$\frac{\mathrm{MS_p}}{U} \leq F_{1-\frac{\alpha}{2}}(n_p - 1, \xi),$$
$$\frac{\mathrm{MS_p}}{F_{*\xi}} \leq \rho_e^* q_* + \mathrm{MS_{pr}} + \mathrm{MS_{po}} - \mathrm{MS_e},$$
$$\rho_e^* q_* \geq \frac{\mathrm{MS_p}}{F_{*\xi}} - \mathrm{MS_{pr}} - \mathrm{MS_{po}} + \mathrm{MS_e},$$
$$\rho_e^* \geq \frac{\mathrm{MS_p/F_{*\xi} - MS_{pr} - MS_{po} + MS_e}}{q_*},$$
$$\rho_e \geq \frac{\mathrm{(MS_p/F_{*\xi} - MS_{pr} - MS_{po} + MS_e)/q_*}}{\mathrm{1 + (MS_p/F_{*\xi} - MS_{pr} - MS_{po} + MS_e)/q_*}}$$
$$= \frac{\mathrm{MS_p/F_{*\xi} - MS_{pr} - MS_{po} + MS_e}}{q_* + \mathrm{(MS_p/F_{*\xi} - MS_{pr} - MS_{po} + MS_e)}}$$
$$= \frac{\mathrm{MS_p - F_{*\xi}(MS_{pr} + MS_{po} - MS_e)}}{F_{*\xi}(q_* - \mathrm{MS_{pr} - MS_{po} + MS_e}) + \mathrm{MS_p}}.$$

Similarly, the upper limit can be derived. Let $F^*_\xi = F_{1-\frac{\alpha}{2}}(\xi, n_p - 1)$,

$$\begin{aligned}
\frac{\mathrm{MS_p}}{U} &\geq F_{\frac{\alpha}{2}}(n_p - 1, \xi), \\
\rho_e &\leq \frac{(\mathrm{MS_p}/\mathrm{F}_{\frac{\alpha}{2}}(\mathrm{n_p} - 1, \xi) - \mathrm{MS_{pr}} - \mathrm{MS_{po}} + \mathrm{MS_e})/\mathrm{q_*}}{1 + (\mathrm{MS_p}/\mathrm{F}_{\frac{\alpha}{2}}(\mathrm{n_p} - 1, \xi) - \mathrm{MS_{pr}} - \mathrm{MS_{po}} + \mathrm{MS_e})/\mathrm{q_*}} \\
&= \frac{F^*_\xi \mathrm{MS_p} - \mathrm{MS_{pr}} - \mathrm{MS_{po}} + \mathrm{MS_e}}{q_* + (F^*_\xi \mathrm{MS_p} - \mathrm{MS_{pr}} - \mathrm{MS_{po}} + \mathrm{MS_e})}.
\end{aligned}$$

So an approximate $100(1-\alpha)\%$ confidence interval for ρ_e is

$$\left[\frac{\mathrm{MS_p} - \hat{\mathrm{F}}_{*\xi}(\mathrm{MS_{pr}} + \mathrm{MS_{po}} - \mathrm{MS_e})}{\hat{F}_{*\xi}(q_* - \mathrm{MS_{pr}} - \mathrm{MS_{po}} + \mathrm{MS_e}) + \mathrm{MS_p}}, \frac{\hat{F}^*_\xi \mathrm{MS_p} - \mathrm{MS_{pr}} - \mathrm{MS_{po}} + \mathrm{MS_e}}{q_* + (\hat{F}^*_\xi \mathrm{MS_p} - \mathrm{MS_{pr}} - \mathrm{MS_{po}} + \mathrm{MS_e})}\right] \tag{3.9}$$

where $\hat{F}_{*\xi} = F_{1-\frac{\alpha}{2}}(n_p - 1, \hat{\xi})$ and $\hat{F}^*_\xi = F_{1-\frac{\alpha}{2}}(\hat{\xi}, n_p - 1)$. This confidence interval has a form similar to the one for ρ_I, but with different parameters, such as q_* and $\hat{\xi}$.

CHAPTER 4

RELIABILITY PARAMETERS FOR REDUCED THREE-WAY RANDOM EFFECTS MODEL

4.1 Introduction

When the performance of subjects (e.g., x-ray films) is not affected by occasions, it is reasonable to ignore the interaction effect between them in the full three-way random effects model. These kinds of models are called reduced three-way random effects models. In this section, I shall use a model without $subjects \times occasions$ interaction as an example to derive the estimator of reliability parameters and the confidence interval for these parameters. For other reduced models, derivations would be similar.

Notation 4.1 *A reduced three-way random effects model may be written*

$$x_{ijk} = \mu + p_i + r_j + o_k + (pr)_{ij} + (ro)_{jk} + e_{ijk} \tag{4.1}$$

$$i = 1, \cdots, n_p \quad j = 1, \cdots, n_r \quad k = 1, \cdots, n_o$$

where all the parameters have the same meaning as that for full three-way random effects model, except $(po)_{ik}$ is omitted from the full model.

Again, I shall start with the ANOVA.

4.2 Expected Values of Mean Squares for Reduced Model

In Section 3.2, I derived the quadratic forms and expected values of the mean squares for full random three-way random effects model. Under the reduced three-way random effects model, the mean squares remain the same. The variance-covariance matrix of $\mathbf{x}$ can be obtained from the full model by setting $\sigma^2_{po} = 0$. The degrees of freedom of residual term in the model increases by the amount of $(n_p - 1)(n_o - 1)$ to $(n_p - 1)n_r(n_o - 1)$. So

$$\begin{aligned}
E(\mathrm{SS_e}) &= E(\mathrm{SS_T} - \mathrm{SS_p} - \mathrm{SS_r} - \mathrm{SS_o} - \mathrm{SS_{pr}} - \mathrm{SS_{po}} - \mathrm{SS_{ro}}) \\
&= (n_p - 1)n_r(n_o - 1)\sigma^2_e \\
&\quad + [N - n_r - n_o + 1 - (n_r - 1)(n_o - 1) - (n_p - 1)n_r n_o]\sigma^2_p \\
&\quad + [N - n_p - n_o + 1 - (n_p - 1)(n_o - 1) - (n_r - 1)n_p n_o]\sigma^2_r \\
&\quad + [N - n_r - n_p + 1 - (n_p - 1)(n_r - 1) - (n_o - 1)n_p n_r]\sigma^2_o \\
&\quad + [N - n_o - (n_p - 1)n_o - (n_r - 1)n_o - (n_p - 1)(n_r - 1)n_o]\sigma^2_{pr} \\
&\quad + [N - n_p - (n_r - 1)n_p - (n_o - 1)n_p - (n_r - 1)(n_o - 1)n_p]\sigma^2_{ro} \\
&= (n_p - 1)n_r(n_o - 1)\sigma^2_e, \\
E(\mathrm{MS_e}) &= \left[(n_p - 1)n_r(n_o - 1)\sigma^2_e\right] / \left[(n_p - 1)n_r(n_o - 1)\right] = \sigma^2_e.
\end{aligned}$$

Table 4.1: ANOVA table for reduced three-way model

Source of variation	df	MS	EMS
Subjects	$n_p - 1$	$\mathrm{MS_p}$	$\sigma_e^2 + n_o\sigma_{pr}^2 + n_r n_o \sigma_p^2$
Raters	$n_r - 1$	$\mathrm{MS_r}$	$\sigma_e^2 + n_p\sigma_{ro}^2 + n_o\sigma_{pr}^2 + n_p n_o \sigma_r^2$
Occasions	$n_o - 1$	$\mathrm{MS_o}$	$\sigma_e^2 + n_p\sigma_{ro}^2 + n_p n_r \sigma_o^2$
Subjects×Raters	$(n_p - 1)(n_r - 1)$	$\mathrm{MS_{pr}}$	$\sigma_e^2 + n_o\sigma_{pr}^2$
Raters×Occasions	$(n_r - 1)(n_o - 1)$	$\mathrm{MS_{ro}}$	$\sigma_e^2 + n_p\sigma_{ro}^2$
Residual	$(n_p - 1)n_r(n_o - 1)$	$\mathrm{MS_e}$	σ_e^2

So the expectations of mean squares become

$$
\begin{aligned}
E(\mathrm{MS_p}) &= \sigma_e^2 + n_r n_o \sigma_p^2 + n_o \sigma_{pr}^2 \\
E(\mathrm{MS_r}) &= \sigma_e^2 + n_p n_o \sigma_r^2 + n_o \sigma_{pr}^2 + n_p \sigma_{ro}^2 \\
E(\mathrm{MS_o}) &= \sigma_e^2 + n_p n_r \sigma_o^2 + n_p \sigma_{ro}^2 \\
E(\mathrm{MS_{pr}}) &= \sigma_e^2 + n_o \sigma_{pr}^2 \\
E(\mathrm{MS_{ro}}) &= \sigma_e^2 + n_p \sigma_{ro}^2 \\
E(\mathrm{MS_e}) &= \sigma_e^2.
\end{aligned}
$$

4.3 Intraclass Correlation Coefficient for Reduced Model

Based on the results from the previous section, the ANOVA table for the reduced three-way random effects model is summarized as Table 4.1.

Table 4.2 lists the unbiased estimators of the variance components.

The ICC for a reduced three-way random effects model is defined as

$$\rho_I = \frac{\sigma_p^2}{\sigma_p^2 + \sigma_r^2 + \sigma_o^2 + \sigma_{pr}^2 + \sigma_{ro}^2 + \sigma_e^2}. \tag{4.2}$$

Table 4.2: Unbiased estimators of variance components for reduced model

Variance Component	Unbiased Estimator
σ_p^2	$\frac{1}{n_r n_o}(\mathrm{MS_p} - \mathrm{MS_{pr}})$
σ_r^2	$\frac{1}{n_p n_o}(\mathrm{MS_r} + \mathrm{MS_e} - \mathrm{MS_{pr}} - \mathrm{MS_{ro}})$
σ_o^2	$\frac{1}{n_p n_r}(\mathrm{MS_o} - \mathrm{MS_{ro}})$
σ_{pr}^2	$\frac{1}{n_o}(\mathrm{MS_{pr}} - \mathrm{MS_e})$
σ_{ro}^2	$\frac{1}{n_p}(\mathrm{MS_{ro}} - \mathrm{MS_e})$
σ_e^2	$\mathrm{MS_e}$

An estimator of ρ_I is

$$\begin{aligned} r_I = \{&\frac{1}{n_r n_o}(MS_p - MS_{pr})\}/\{\frac{1}{n_r n_o}MS_p + \frac{1}{n_p n_o}MS_r + \frac{1}{n_p n_r}MS_o \\ &+ (\frac{1}{n_o} - \frac{1}{n_r n_o} - \frac{1}{n_p n_o})MS_{pr} + (\frac{1}{n_p} - \frac{1}{n_p n_o} - \frac{1}{n_p n_r})MS_{ro} \\ &+ (1 + \frac{1}{n_p n_o} - \frac{1}{n_p} - \frac{1}{n_o})MS_e\} \end{aligned} \tag{4.3}$$

which is the ratio of an unbiased estimator of the numerator of ρ_I to an unbiased estimator of the denominator of ρ_I. This estimator is consistent but biased. The proof of consistency parallels that for the one-way model.

As I did for ρ_I in full three-way random effects model, a $100(1-\alpha)\%$ confidence interval for ρ_I in the reduced three-way random effects model can be derived. First from Definition (4.2), I have

$$\begin{aligned} 1 - \rho_I &= \frac{\sigma_r^2 + \sigma_o^2 + \sigma_{pr}^2 + \sigma_{ro}^2 + \sigma_e^2}{\sigma_p^2 + \sigma_r^2 + \sigma_o^2 + \sigma_{pr}^2 + \sigma_{ro}^2 + \sigma_e^2} \\ \frac{\rho_I}{1-\rho_I} &= \frac{\sigma_p^2}{\sigma_r^2 + \sigma_o^2 + \sigma_{pr}^2 + \sigma_{ro}^2 + \sigma_e^2} \\ \sigma_p^2 &= \frac{\rho_I}{1-\rho_I}(\sigma_r^2 + \sigma_o^2 + \sigma_{pr}^2 + \sigma_{ro}^2 + \sigma_e^2) \\ &= \rho_I^*(\sigma_r^2 + \sigma_o^2 + \sigma_{pr}^2 + \sigma_{ro}^2 + \sigma_e^2) \end{aligned} \tag{4.4}$$

where $\rho_I^* = \rho_I/(1-\rho_I)$. Substituting σ_p^2 from Equation (4.4) in the expectation of $\mathrm{MS_p}$ in Table 4.1 gives

$$\begin{aligned}E(\mathrm{MS_p}) &= n_r n_o \sigma_p^2 + n_o \sigma_{pr}^2 + \sigma_e^2 \\ &= n_r n_o \rho_I^* (\sigma_r^2 + \sigma_o^2 + \sigma_{pr}^2 + \sigma_{ro}^2 + \sigma_e^2) + n_o \sigma_{pr}^2 + \sigma_e^2 \\ &= n_r n_o \rho_I^* \sigma_r^2 + n_r n_o \rho_I^* \sigma_o^2 + (n_r n_o \rho_I^* + n_o)\sigma_{pr}^2 \\ &\quad + n_r n_o \rho_I^* \sigma_{ro}^2 + (1 + n_r n_o \rho_I^*)\sigma_e^2.\end{aligned}$$

Next substituting the variance components with their unbiased estimators in the above equation, I define a variable V so that $E(V) = E(\mathrm{MS_p})$.

$$\begin{aligned}V &= \frac{n_r n_o}{n_p n_o}\rho_I^*(\mathrm{MS_r} + \mathrm{MS_e} - \mathrm{MS_{pr}} - \mathrm{MS_{ro}}) + \frac{\mathrm{n_r n_o}}{\mathrm{n_p n_r}}\rho_\mathrm{I}^*(\mathrm{MS_o} - \mathrm{MS_{ro}}) \\ &\quad + (1 + n_r \rho_I^*)(\mathrm{MS_{pr}} - \mathrm{MS_e}) + \frac{\mathrm{n_r n_o}}{\mathrm{n_p}}\rho_\mathrm{I}^*(\mathrm{MS_{ro}} - \mathrm{MS_e}) \\ &\quad + (1 + n_r n_o \rho_I^*)\mathrm{MS_e} \\ &= \frac{n_r}{n_p}\rho_I^* \mathrm{MS_r} + \frac{\mathrm{n_o}}{\mathrm{n_p}}\rho_\mathrm{I}^* \mathrm{MS_o} + \left(1 + \mathrm{n_r}\rho_\mathrm{I}^* - \frac{\mathrm{n_r}}{\mathrm{n_p}}\rho_\mathrm{I}^*\right)\mathrm{MS_{pr}} \\ &\quad + \left(\frac{n_r n_o}{n_p}\rho_I^* - \frac{n_r}{n_p}\rho_I^* - \frac{n_o}{n_p}\rho_I^*\right)\mathrm{MS_{ro}} \\ &\quad + \left(\frac{n_r}{n_p}\rho_I^* - 1 - n_r \rho_I^* - \frac{n_r n_o}{n_p}\rho_I^* + 1 + n_r n_o \rho_I^*\right)\mathrm{MS_e} \\ &= \rho_I^*\left[\frac{n_r}{n_p}\mathrm{MS_r} + \frac{\mathrm{n_o}}{\mathrm{n_p}}\mathrm{MS_o} + \left(\mathrm{n_r} - \frac{\mathrm{n_r}}{\mathrm{n_p}}\right)\mathrm{MS_{pr}} + \left(\frac{\mathrm{n_r n_o - n_r - n_o}}{\mathrm{n_p}}\right)\mathrm{MS_{ro}} \right. \\ &\quad \left. + \left(\frac{n_r - n_p n_r - n_r n_o + n_p n_r n_o}{n_p}\right)\mathrm{MS_e}\right] + \mathrm{MS_{pr}}.\end{aligned}$$

V is a linear combination of independent mean squares: $\mathrm{MS_r}$, $\mathrm{MS_o}$, $\mathrm{MS_{pr}}$, $\mathrm{MS_{ro}}$ and $\mathrm{MS_e}$. V is independent of $\mathrm{MS_p}$. According to Satterthwaite's approximation, V is approximately distributed as $d\chi_\nu^2/\nu$ with $d = E(V)$ and degrees of freedom ν. Note that degrees of freedom of $\mathrm{MS_e}$ are now $(n_p - 1)n_r(n_o - 1)$. Hence,

$$\nu = \frac{(m_1 \mathrm{MS_r} + \mathrm{m_2 MS_o} + \mathrm{m_3 MS_{pr}} + \mathrm{m_4 MS_{ro}} + \mathrm{m_5 MS_e})^2}{\frac{(m_1 \mathrm{MS_r})^2}{n_r - 1} + \frac{(m_2 \mathrm{MS_o})^2}{n_o - 1} + \frac{(m_3 \mathrm{MS_{pr}})^2}{(n_p - 1)(n_r - 1)} + \frac{(m_4 \mathrm{MS_{ro}})^2}{(n_r - 1)(n_o - 1)} + \frac{(m_5 \mathrm{MS_e})^2}{(n_p - 1)n_r(n_o - 1)}}$$

where

$$m_1 = \frac{n_r}{n_p}\rho_I^*,$$
$$m_2 = \frac{n_o}{n_p}\rho_I^*,$$
$$m_3 = 1 + n_r\rho_I^* - \frac{n_r}{n_p}\rho_I^*,$$
$$m_4 = \frac{n_r n_o - n_r - n_o}{n_p}\rho_I^*,$$
$$m_5 = \left(\frac{n_r}{n_p} - n_r - \frac{n_r n_o}{n_p} + n_r n_o\right)\rho_I^*.$$

To estimate ν, use r_I to replace ρ_I.

Then

$$F = \mathrm{MS_p/V}$$
$$= \mathrm{MS_p/(m_1MS_r + m_2MS_o + m_3MS_{pr} + m_4MS_{ro} + m_5MS_e)}$$

has an approximate F distribution with degrees of freedom $(n_p - 1)$ and ν. Therefore

$$Pr\left\{F_{\frac{\alpha}{2}}(n_p - 1, \nu) \le \mathrm{MS_p/V} \le \mathrm{F}_{1-\frac{\alpha}{2}}(\mathrm{n_p} - 1, \nu)\right\} = 1 - \alpha,$$

where $F_{1-\frac{\alpha}{2}}(n_p - 1, \nu)$ is the $(1 - \alpha/2) \times 100$th percentile of an F distribution with degrees of freedom $(n_p - 1)$ and ν.

From the above formula, the lower and upper limits of a confidence interval for ρ_I are estimable. Let $F_* = F_{1-\frac{\alpha}{2}}(n_p - 1, \nu)$ and $q = (n_r/n_p)\mathrm{MS_r} + (\mathrm{n_o/n_p})\mathrm{MS_o} + (\mathrm{n_r} - \mathrm{n_r/n_p})\mathrm{MS_{pr}} + [(\mathrm{n_rn_o} - \mathrm{n_r} - \mathrm{n_o})/\mathrm{n_p}]\mathrm{MS_{ro}} + [(\mathrm{n_r} - \mathrm{n_pn_r} - \mathrm{n_rn_o} + \mathrm{n_pn_rn_o})/\mathrm{n_p}]\mathrm{MS_e}$,

$$
\begin{aligned}
\frac{\mathrm{MS_p}}{V} &\le F_{1-\frac{\alpha}{2}}(n_p-1,\nu),\\
\frac{\mathrm{MS_p}}{F_*} &\le \rho_I^* q + \mathrm{MS_{pr}},\\
\rho_I^* q &\ge \frac{\mathrm{MS_p}}{F_*} - \mathrm{MS_{pr}},\\
\rho_I^* &\ge \frac{\mathrm{MS_p/F_* - MS_{pr}}}{q},\\
\rho_I &\ge \frac{\mathrm{(MS_p/F_* - MS_{pr})/q}}{\mathrm{1 + (MS_p/F_* - MS_{pr})/q}}\\
&= \frac{\mathrm{MS_p/F_* - MS_{pr}}}{q + \mathrm{(MS_p/F_* - MS_{pr})}}\\
&= \frac{\mathrm{MS_p - F_*(MS_{pr})}}{F_*(q - \mathrm{MS_{pr}}) + \mathrm{MS_p}}.
\end{aligned}
$$

Similarly, the upper limit can be derived. Let $F^* = F_{1-\frac{\alpha}{2}}(\nu, n_p - 1)$,

$$
\begin{aligned}
\frac{\mathrm{MS_p}}{V} &\ge F_{\frac{\alpha}{2}}(n_p-1,\nu),\\
\rho_I &\le \frac{\left(\mathrm{MS_p/F_{\frac{\alpha}{2}}(n_p-1,\nu) - MS_{pr}}\right)/q}{1 + \left(\mathrm{MS_p/F_{\frac{\alpha}{2}}(n_p,\nu)_* - MS_{pr}}\right)/q}\\
&= \frac{F^*\mathrm{MS_p - MS_{pr}}}{q + (F^*\mathrm{MS_p - MS_{pr}})}.
\end{aligned}
$$

So an approximate $100(1-\alpha)\%$ confidence interval for ρ_I in the reduced three-way random effects effects model is

$$
\left[\frac{\mathrm{MS_p - \hat{F}_*(MS_{pr})}}{\hat{F}_*(q - \mathrm{MS_{pr}}) + \mathrm{MS_p}}, \frac{\hat{F}^*\mathrm{MS_p - MS_{pr}}}{q + (\hat{F}^*\mathrm{MS_p - MS_{pr}})}\right] \tag{4.5}
$$

where $\hat{F}_* = F_{1-\frac{\alpha}{2}}(n_p - 1, \hat{\nu})$ and $\hat{F}^* = F_{1-\frac{\alpha}{2}}(\hat{\nu}, n_p - 1)$.

4.4 Interrater Reliability Coefficient for Reduced Model

The *interrater reliability coefficient* for reduced three-way random effects model is defined by

$$\rho_e = \frac{\sigma_p^2}{\sigma_p^2 + \sigma_{pr}^2 + \sigma_{ro}^2 + \sigma_e^2}. \tag{4.6}$$

It has the same form as the IRC for the full three-way random effects model, Equation (3.6).

An estimator of ρ_e is

$$r_e = \left\{\frac{1}{n_r n_o}(\mathrm{MS_p} - \mathrm{MS_{pr}})\right\} \Big/ \left\{\frac{1}{\mathrm{n_r n_o}}\mathrm{MS_p} \right. \tag{4.7}$$
$$\left. + (\frac{1}{n_o} - \frac{1}{n_r n_o})\mathrm{MS_{pr}} + \frac{1}{\mathrm{n_p}}\mathrm{MS_{ro}} + (1 - \frac{1}{\mathrm{n_p}} - \frac{1}{\mathrm{n_o}})\mathrm{MS_e}\right\}$$

which is the ratio of an unbiased estimator of the numerator of ρ_e to an unbiased estimator of the denominator of ρ_e. This estimator is consistent but biased. The proof of consistency parallels that for the one-way model.

Following the same theme of previous section, a $100(1-\alpha)\%$ confidence interval for ρ_e in the reduced three-way random effects model can be derived. First from the definition in equation (4.6), I have

$$1 - \rho_e = \frac{\sigma_{pr}^2 + \sigma_{ro}^2 + \sigma_e^2}{\sigma_p^2 + \sigma_{pr}^2 + \sigma_{ro}^2 + \sigma_e^2},$$
$$\frac{\rho_e}{1-\rho_e} = \frac{\sigma_p^2}{\sigma_{pr}^2 + \sigma_{ro}^2 + \sigma_e^2}$$
$$\sigma_p^2 = \frac{\rho_e}{1-\rho_e}(\sigma_{pr}^2 + \sigma_{ro}^2 + \sigma_e^2)$$
$$= \rho_e^*(\sigma_{pr}^2 + \sigma_{ro}^2 + \sigma_e^2) \tag{4.8}$$

where $\rho_e^* = \rho_e/(1-\rho_e)$. Substituting σ_p^2 from Equation (4.8) in the expectation of $\mathrm{MS_p}$ in Table 4.1, I get

$$E(\mathrm{MS_p}) = n_r n_o \sigma_p^2 + n_o \sigma_{pr}^2 + \sigma_e^2$$
$$= n_r n_o \rho_e^*(\sigma_{pr}^2 + \sigma_{ro}^2 + \sigma_e^2) + n_o \sigma_{pr}^2 + \sigma_e^2$$
$$= (n_r n_o \rho_e^* + n_o)\sigma_{pr}^2 + n_r n_o \rho_e^* \sigma_{ro}^2 + (1 + n_r n_o \rho_e^*)\sigma_e^2.$$

Next, substituting the variance components with their unbiased estimators in the above equation, I define a variable U so that $E(U) = E(\mathrm{MS_p})$.

$$\begin{aligned}
U &= (1 + n_r\rho_e^*)(\mathrm{MS_{pr}} - \mathrm{MS_e}) + \frac{\mathrm{n_r n_o}}{\mathrm{n_p}}\rho_\mathrm{e}^*(\mathrm{MS_{ro}} - \mathrm{MS_e}) \\
&\quad + (1 + n_r n_o \rho_e^*)\mathrm{MS_e} \\
&= (1 + n_r\rho_e^*)\mathrm{MS_{pr}} + \frac{\mathrm{n_r n_o}}{\mathrm{n_p}}\rho_\mathrm{e}^*\mathrm{MS_{ro}} + \left(-\mathrm{n_r}\rho_\mathrm{e}^* - \frac{\mathrm{n_r n_o}}{\mathrm{n_p}}\rho_\mathrm{e}^* + \mathrm{n_r n_o}\rho_\mathrm{e}^*\right)\mathrm{MS_e} \\
&= \rho_e^*\left[n_r\mathrm{MS_{pr}} + \frac{\mathrm{n_r n_o}}{\mathrm{n_p}}\mathrm{MS_{ro}} + \frac{\mathrm{n_p n_r n_o} - \mathrm{n_p n_r} - \mathrm{n_r n_o}}{\mathrm{n_p}}\mathrm{MS_e}\right] + \mathrm{MS_{pr}}.
\end{aligned}$$

U is a linear combination of independent mean squares: $\mathrm{MS_{pr}}$, $\mathrm{MS_{ro}}$ and $\mathrm{MS_e}$. U is independent of $\mathrm{MS_p}$. According to Satterthwaite's approximation, U is approximately distributed as $b\chi^2_\xi/\xi$ with $b = E(U)$ and degrees of freedom

$$\xi = \frac{(n_1\mathrm{MS_{pr}} + \mathrm{n_2 MS_{ro}} + \mathrm{n_3 MS_e})^2}{\frac{(n_1\mathrm{MS_{pr}})^2}{(n_p-1)(n_r-1)} + \frac{(n_2\mathrm{MS_{ro}})^2}{(n_r-1)(n_o-1)} + \frac{(n_3\mathrm{MS_e})^2}{(n_p-1)n_r(n_o-1)}}$$

where

$$\begin{aligned}
n_1 &= 1 + n_r\rho_e^*, \\
n_2 &= \frac{n_r n_o}{n_p}\rho_e^*, \\
n_3 &= \left(-n_r - \frac{n_r n_o}{n_p} + n_r n_o\right)\rho_e^*.
\end{aligned}$$

To estimate ξ, use r_e to replace ρ_e.

Then

$$\begin{aligned}
F &= \mathrm{MS_p/U} \\
&= \mathrm{MS_p}/\left(\mathrm{n_1 MS_{pr}} + \mathrm{n_2 MS_{ro}} + \mathrm{n_3 MS_e}\right)
\end{aligned}$$

has an approximate F distribution with degrees of freedom $(n_p - 1)$ and ξ. Therefore

$$Pr\left\{F_{\frac{\alpha}{2}}(n_p - 1, \xi) \leq \mathrm{MS_p/U} \leq \mathrm{F}_{1-\frac{\alpha}{2}}(\mathrm{n_p} - 1, \xi)\right\} = 1 - \alpha,$$

where $F_{1-\frac{\alpha}{2}}(n_p - 1, \xi)$ is the $(1 - \alpha/2) \times 100$th percentile of an F distribution with degrees of freedom $(n_p - 1)$ and ξ.

From the above formula, the lower and upper limits of a confidence interval for ρ_e are estimable. Let $F_{*\xi} = F_{1-\frac{\alpha}{2}}(n_p - 1, \xi)$ and $q_* = n_r\mathrm{MS_{pr}} + [(\mathrm{n_r n_o})/\mathrm{n_p}]\mathrm{MS_{ro}} + [(\mathrm{n_p n_r n_o} - \mathrm{n_p n_r} - \mathrm{n_r n_o})/\mathrm{n_p})]\mathrm{MS_e}$,

$$
\begin{aligned}
\frac{\mathrm{MS_p}}{U} &\leq F_{1-\frac{\alpha}{2}}(n_p - 1, \xi), \\
\frac{\mathrm{MS_p}}{F_{*\xi}} &\leq \rho_e^* q_* + \mathrm{MS_{pr}}, \\
\rho_e^* q_* &\geq \frac{\mathrm{MS_p}}{F_{*\xi}} - \mathrm{MS_{pr}}, \\
\rho_e^* &\geq \frac{\mathrm{MS_p/F_{*\xi} - MS_{pr}}}{q_*}, \\
\rho_e &\geq \frac{(\mathrm{MS_p/F_{*\xi} - MS_{pr})/q_*}}{1 + (\mathrm{MS_p/F_{*\xi} - MS_{pr})/q_*}} \\
&= \frac{\mathrm{MS_p/F_{*\xi} - MS_{pr}}}{q_* + (\mathrm{MS_p/F_{*\xi} - MS_{pr}})} \\
&= \frac{\mathrm{MS_p - F_{*\xi}(MS_{pr})}}{F_{*\xi}(q_* - \mathrm{MS_{pr}}) + \mathrm{MS_p}}.
\end{aligned}
$$

Similarly, the upper limit can be derived. Let $F_\xi^* = F_{1-\frac{\alpha}{2}}(\xi, n_p - 1)$,

$$
\begin{aligned}
\frac{\mathrm{MS_p}}{U} &\geq F_{\frac{\alpha}{2}}(n_p - 1, \xi), \\
\rho_e &\leq \frac{\left(\mathrm{MS_p/F_{\frac{\alpha}{2}}(n_p - 1, \xi) - MS_{pr}}\right)/q_*}{1 + \left(\mathrm{MS_p/F_{\frac{\alpha}{2}}(n_p - 1, \xi) - MS_{pr}}\right)/q_*}, \\
&= \frac{F_\xi^*\mathrm{MS_p} - \mathrm{MS_{pr}}}{q_* + (F_\xi^*\mathrm{MS_p} - \mathrm{MS_{pr}})}.
\end{aligned}
$$

So an approximate $100(1-\alpha)\%$ confidence interval for ρ_e in the reduced three-way random effects effects model is

$$
\left[\frac{\mathrm{MS_p - \hat{F}_{*\xi}(MS_{pr})}}{\hat{F}_{*\xi}(q_* - \mathrm{MS_{pr}}) + \mathrm{MS_p}}, \frac{\hat{F}_\xi^*\mathrm{MS_p} - \mathrm{MS_{pr}}}{q_*(\hat{F}_\xi^*\mathrm{MS_p} - \mathrm{MS_{pr}})}\right] \tag{4.9}
$$

where $\hat{F}_{*\xi} = F_{1-\frac{\alpha}{2}}(n_p - 1, \hat{\xi})$ and $\hat{F}_\xi^* = F_{1-\frac{\alpha}{2}}(\hat{\xi}, n_p - 1)$.

In next chapter, I shall use these methods to analyze a real data set of small size and examine the estimators through simulations.

CHAPTER 5

DATA ANALYSIS AND SIMULATION

5.1 Introduction

In this Chapter, I include data analysis and simulation. In the second section, I apply the methods developed in Chapter 3 to analyze a real data set. The confidence intervals are also calculated by bootstrapping methods. Then the results from ANOVA-based formulas and bootstrapping are compared. In the third section, I do simulation to assess the interval coverage probabilities. I also study the robustness of the confidence intervals to alternative distributional assumptions.

5.2 Data Analysis

5.2.1 Data Source

The data come from an orthopedic study, Harrison et al. (2000), designed to assess the reliability of readings of x-ray films of each of 30 patients by 3 different examiners on 2 different occasions. On two successive Saturdays, each of the 3 examiners evaluated an x-ray of the thoracic spine from the same 30 subjects. On the second Saturday, the ordering of the x-rays was re-randomized for the examinations. Using the same x-ray films throughout, each examiner assessed 16 measurements. Each measurement is an angle between tangents to consecutive thoracic vertebrae. For example, `RRAVT2T3` is the relative rotation angle between segments in the thoracic spine called `T2` and `T3`. These are standard skeletal measurements.

5.2.2 Analysis Procedure

For this situation, the x-ray films do not change from occasion to occasion. That is, $\sigma^2_{po} = 0$. The ICC in the full three-way random effects model has complicated point and confidence interval estimates. I assume full model for the data and do the estimation. The calculations were coded in SAS 9.0 software. The SAS program is given in Appendix A. For other reliability parameters, minor modifications to the program are needed for estimation.

The data were assessed for normality. There were 30 data points for each combination of raters and occasions. The normality of the data within each of the six rater-occasion combinations for one measurement was tested using the Shapiro-Wilk test. Altogether there were 96 ($= 3 \times 2 \times 16$) such tests. Only 3 out of the 96 sets of data failed this test at the 0.05 level. Table 5.1 lists the failures and corresponding p-values. None failed Shapiro-Wilk at the 0.01 level. So it is reasonable to assume that these data are at least approximately Gaussian. Multivariate normality was also tested on the 96 variables. But due to the high correlations among them, which in turn caused a singularity problem, the multivariate test could not be conducted. The fact that the number of variables, 96, is greater than the sample size, is the likely culprit.

The point estimates and confidence intervals of ρ_I were calculated using equations (3.3) and (3.5). The bootstrap resampling method was then used to validate the confidence interval. Efron (1987) recommended 1000 iterations to get good estimates for the endpoints of bootstrap confidence intervals. Here 10,000 samples of the same size of original data were drawn from the original

Table 5.1: Test of normality

Num	Measure	Rater-occasion	Test	Stat	p-value
1	RRAVT2T3	B1	Shapiro-Wilk	0.9297	0.0483
2	RRAVT3T4	B2	Shapiro-Wilk	0.9209	0.0283
3	RRT11T12	C2	Shapiro-Wilk	0.9250	0.0363

data by film to be conservative. For each bootstrap sample, the point estimate of ρ_I was calculated. Then Efron percentile method and bias-corrected percentile method (Davison and Hinkley, 1997) were employed to get the bootstrap confidence intervals. Because the natural upper limit of ICC is 1, sometimes, only the the lower confidence limit of ρ_I is of interest. Both one-sided and two-sided confidence intervals were calculated.

Table 5.2 lists two-sided confidence intervals for ρ_I calculated from the developed formulas and bootstrapping methods, where the column headings `ANOVA`, `Efron` and `Bias-corrected` denote ANOVA method, Efron percentile method, and the bias-corrected percentile method, respective.

Table 5.3 lists one-sided confidence intervals for ρ_I, with columns carrying the same meaning as that of two-sided confidence intervals.

5.2.3 Conclusion

Comparing the different methods for confidence intervals, I generally conclude that all confidence interval methods are close, especially between those from the developed formula and the Efron percentile method — the difference is within two places after the decimal. Among the three methods, confidence intervals using the developed formulas have the smallest width for most of the measurements. Therefore the formula derived for confidence intervals on ρ_I is justified from the bootstrapping validation.

Table 5.2: Two-sided confidence interval of ρ_I

Measure	ICC	Lower limit			Upper limit			Length of CI		
		ANOVA	Efron	Bias-corrected	ANOVA	Efron	Bias-corrected	ANOVA	Efron	Bias-corrected
T1T12VER	1.00	1.00	1.00	1.00	1.00	1.00	1.00	0.00	0.00	0.00
T2T11VER	1.00	1.00	1.00	1.00	1.00	1.00	1.00	0.00	0.00	0.00
RRAVT1T2	0.91	0.86	0.84	0.81	0.95	0.96	0.95	0.09	0.12	0.14
RRAVT2T3	0.94	0.90	0.90	0.88	0.97	0.96	0.96	0.07	0.06	0.08
RRAVT3T4	0.88	0.81	0.79	0.77	0.94	0.93	0.93	0.12	0.14	0.16
RRAVT4T5	0.93	0.88	0.83	0.81	0.96	0.97	0.96	0.08	0.13	0.15
RRAVT5T6	0.90	0.83	0.81	0.78	0.94	0.95	0.94	0.11	0.14	0.17
RRAVT6T7	0.76	0.63	0.61	0.58	0.86	0.85	0.84	0.23	0.24	0.26
RRAVT7T8	0.80	0.70	0.71	0.68	0.89	0.88	0.87	0.19	0.17	0.19
RRAVT8T9	0.79	0.68	0.68	0.65	0.88	0.87	0.86	0.20	0.19	0.21
RRAT9T10	0.81	0.71	0.72	0.68	0.89	0.88	0.87	0.18	0.16	0.18
RRT10T11	0.83	0.73	0.69	0.68	0.90	0.91	0.91	0.18	0.22	0.23
RRT11T12	0.80	0.70	0.66	0.62	0.89	0.89	0.88	0.19	0.23	0.26
ARAT1T12	0.99	0.98	0.98	0.97	0.99	0.99	0.99	0.02	0.02	0.02
ARAT2T11	0.98	0.97	0.97	0.96	0.99	0.99	0.99	0.02	0.03	0.03
ARAT3T10	0.98	0.96	0.96	0.95	0.99	0.99	0.99	0.03	0.03	0.03

Table 5.3: One-sided confidence interval of ρ_I

Measure	ICC	Lower limit			Length of CI		
		ANOVA	Efron	Bias-corrected	ANOVA	Efron	Bias-corrected
T1T12VER	1.00	1.00	1.00	1.00	0.00	0.00	0.00
T2T11VER	1.00	1.00	1.00	1.00	0.00	0.00	0.00
RRAVT1T2	0.91	0.87	0.86	0.83	0.13	0.14	0.17
RRAVT2T3	0.94	0.91	0.91	0.89	0.09	0.09	0.11
RRAVT3T4	0.88	0.83	0.81	0.79	0.17	0.19	0.21
RRAVT4T5	0.93	0.89	0.86	0.84	0.11	0.14	0.16
RRAVT5T6	0.90	0.84	0.83	0.80	0.16	0.17	0.20
RRAVT6T7	0.76	0.66	0.64	0.61	0.34	0.36	0.39
RRAVT7T8	0.80	0.72	0.73	0.70	0.28	0.27	0.30
RRAVT8T9	0.79	0.70	0.70	0.68	0.30	0.30	0.32
RRAT9T10	0.81	0.73	0.74	0.71	0.27	0.26	0.29
RRT10T11	0.83	0.75	0.72	0.70	0.25	0.28	0.30
RRT11T12	0.80	0.72	0.69	0.65	0.28	0.31	0.35
ARAT1T12	0.99	0.98	0.98	0.97	0.02	0.02	0.03
ARAT2T11	0.98	0.97	0.97	0.96	0.03	0.03	0.04
ARAT3T10	0.98	0.96	0.96	0.96	0.04	0.04	0.04

5.3 Validation by Simulation

5.3.1 Validation from Gaussian Data

In this section, I simulate data to validate the developed formula (3.5) for a confidence interval on ρ_I. For each random effect in the model, I specify a Gaussian distribution with mean 0 and variance of one, except for that due to the subjects. For each sample of simulated data, I estimate the confidence interval of ρ_I with formula (3.5). The coverage probability is the percentage of samples with the estimated confidence interval containing ρ_I. Finally, I compare the calculated coverage probability to the nominal coverage probability through Wald and score statistical hypothesis tests.

Tables 5.4, 5.5 and 5.6 list the coverage probabilities for several settings of the simulated data. To keep the same number of samples with bootstrap, the number of simulated samples N is set to be 10,000. In these tables, ρ_I is the true value of ICC, the p-value is for the testing whether the coverage probability is 95% against the alternative that it is not by either the Wald test or score test.

Tables 5.4, 5.5 and 5.6 illustrate statistically significant undercoverage, especially Table 5.6. Undercoverage is most likely the result of small values of

Table 5.4: Coverage probability when $n_p = 30$ and p_i is distributed as Gaussian

ρ_I	n_p	n_r	n_o	Coverage (%)	p-value	
					Wald	Score
0.00	30	3	2	98.6	0.000	0.000
0.14	30	3	2	94.9	0.682	0.680
0.40	30	3	2	93.8	0.000	0.000
0.60	30	3	2	93.2	0.000	0.000
0.73	30	3	2	92.8	0.000	0.000
0.81	30	3	2	92.8	0.000	0.000
0.86	30	3	2	92.6	0.000	0.000
0.89	30	3	2	92.6	0.000	0.000
0.91	30	3	2	92.5	0.000	0.000
0.93	30	3	2	92.5	0.000	0.000
0.00	30	4	3	98.2	0.000	0.000
0.14	30	4	3	94.9	0.649	0.646
0.40	30	4	3	94.6	0.101	0.090
0.60	30	4	3	94.5	0.031	0.025
0.73	30	4	3	94.3	0.003	0.002
0.81	30	4	3	94.3	0.003	0.001
0.86	30	4	3	94.3	0.003	0.002
0.89	30	4	3	94.2	0.000	0.000
0.91	30	4	3	94.2	0.000	0.000
0.93	30	4	3	94.1	0.000	0.000

p-value is for the Wald or score test: $coverage = 95\%$ vs. $coverage \neq 95\%$.

n_r and n_o. In each of Tables 5.4 and 5.5, there are many cases for which the null hypothesis of 95% coverage probability cannot be rejected at level 0.05. From a practical perspective, the results in Table 5.4 are respectable and as expected. There are some trends worth of noting in these results: general speaking, when n_r and n_o are more comparable to n_p, the coverage probability will be closer to 95% and when ρ_I increases, coverage probability decreases given n_p, n_r and n_o are held fixed.

Table 5.5: Coverage probability when $n_p = 60$ and p_i is distributed as Gaussian

ρ_I	n_p	n_r	n_o	Coverage (%)	p-value	
					Wald	Score
0.00	60	3	2	98.5	0.000	0.000
0.14	60	3	2	94.3	0.003	0.002
0.40	60	3	2	92.4	0.000	0.000
0.60	60	3	2	91.3	0.000	0.000
0.73	60	3	2	90.6	0.000	0.000
0.81	60	3	2	90.4	0.000	0.000
0.86	60	3	2	90.2	0.000	0.000
0.89	60	3	2	90.1	0.000	0.000
0.91	60	3	2	90.1	0.000	0.000
0.93	60	3	2	90.1	0.000	0.000
0.00	60	4	3	97.8	0.000	0.000
0.14	60	4	3	95.2	0.485	0.491
0.40	60	4	3	94.4	0.015	0.010
0.60	60	4	3	94.3	0.004	0.002
0.73	60	4	3	94.1	0.000	0.000
0.81	60	4	3	93.9	0.000	0.000
0.86	60	4	3	93.9	0.000	0.000
0.89	60	4	3	93.9	0.000	0.000
0.91	60	4	3	93.9	0.000	0.000
0.93	60	4	3	93.8	0.000	0.000

p-value is for the Wald or score test: $coverage = 95\%$ vs. $coverage \neq 95\%$.

Table 5.6: Coverage probability when $n_p = 100$ and p_i is distributed as Gaussian

ρ_I	n_p	n_r	n_o	Coverage (%)	p-value	
					Wald	Score
0.00	100	3	2	98.7	0.000	0.000
0.14	100	3	2	93.3	0.000	0.000
0.40	100	3	2	90.0	0.000	0.000
0.60	100	3	2	88.6	0.000	0.000
0.73	100	3	2	88.1	0.000	0.000
0.81	100	3	2	87.9	0.000	0.000
0.86	100	3	2	87.7	0.000	0.000
0.89	100	3	2	87.6	0.000	0.000
0.91	100	3	2	87.5	0.000	0.000
0.93	100	3	2	87.5	0.000	0.000
0.00	100	4	3	98.2	0.000	0.000
0.14	100	4	3	94.5	0.039	0.031
0.40	100	4	3	92.9	0.000	0.000
0.60	100	4	3	92.1	0.000	0.000
0.73	100	4	3	91.9	0.000	0.000
0.81	100	4	3	91.8	0.000	0.000
0.86	100	4	3	91.7	0.000	0.000
0.89	100	4	3	91.7	0.000	0.000
0.91	100	4	3	91.7	0.000	0.000
0.93	100	4	3	91.7	0.000	0.000

p-value is for the Wald or score test: $coverage = 95\%$ vs. $coverage \neq 95\%$.

5.3.2 Validation Study with Non-Gaussian Data

In previous section, all the random effects are simulated to be Gaussian. In this section, I change p_i to non-Gaussian, distributed as uniform or double exponential, and repeat what I did in the previous section. p_i represent the subject effects with specific mean and variance. Then the calculated coverage probability is compared to the nominal coverage probability.

Table 5.7 – 5.10 list the calculated coverage probabilities in the settings that p_i is distributed as unform and double exponential while variances of all the variance components are set to 1 except σ_p^2. Again to keep the same number of samples with bootstrap, the number of simulated samples is set to

Table 5.7: Coverage probability when $n_p = 30$ and p_i is distributed as uniform

ρ_I	n_p	n_r	n_o	Coverage (%)	p-value	
					Wald	Score
0.05	30	3	2	98.4	0.000	0.000
0.18	30	3	2	95.7	0.001	0.002
0.33	30	3	2	96.0	0.000	0.000
0.47	30	3	2	95.6	0.008	0.012
0.58	30	3	2	95.1	0.515	0.521
0.67	30	3	2	94.9	0.682	0.680
0.73	30	3	2	94.9	0.556	0.551
0.78	30	3	2	94.7	0.244	0.233
0.82	30	3	2	94.6	0.077	0.066
0.85	30	3	2	94.4	0.015	0.010
0.87	30	3	2	94.3	0.004	0.002
0.89	30	3	2	94.2	0.001	0.000
0.90	30	3	2	94.1	0.000	0.000
0.92	30	3	2	94.1	0.000	0.000
0.93	30	3	2	94.1	0.000	0.000

p-value is for the Wald or score test: $coverage = 95\%$ vs. $coverage \neq 95\%$.

be 10000. From the results, the coverage probability varies among different settings, with one value as high as 98.4% and one value as low as 62.2%. The trend that when ρ_I increases, coverage probability decreases given n_p, n_r and n_o are fixed still holds regardless of the distribution of p_i. Although overcoverages occur, undercoverage is more commonplace, especially when p_i is distributed as double exponential and ρ_I is greater than about 0.20. So the estimation of the confidence interval for ρ_I is not robust, and it is desirable to check the normality of the data before using the formulas for confidence intervals.

Table 5.8: Coverage probability when $n_p = 100$ and p_i is distributed as uniform

ρ_I	n_p	n_r	n_o	Coverage (%)	p-value	
					Wald	Score
0.05	100	3	2	95.7	0.000	0.001
0.18	100	3	2	94.1	0.000	0.000
0.33	100	3	2	91.7	0.000	0.000
0.47	100	3	2	90.4	0.000	0.000
0.58	100	3	2	89.6	0.000	0.000
0.67	100	3	2	89.0	0.000	0.000
0.73	100	3	2	88.8	0.000	0.000
0.78	100	3	2	88.6	0.000	0.000
0.82	100	3	2	88.3	0.000	0.000
0.85	100	3	2	88.2	0.000	0.000
0.87	100	3	2	88.2	0.000	0.000
0.89	100	3	2	88.1	0.000	0.000
0.90	100	3	2	88.1	0.000	0.000
0.92	100	3	2	88.0	0.000	0.000
0.93	100	3	2	88.0	0.000	0.000

p-value is for the Wald or score test: $coverage = 95\%$ vs. $coverage \neq 95\%$.

Table 5.9: Coverage probability when $n_p = 30$ and p_i is distributed as double exponential

ρ_I	n_p	n_r	n_o	Coverage (%)	p-value	
					Wald	Score
0.08	30	3	2	96.6	0.000	0.000
0.25	30	3	2	91.0	0.000	0.000
0.46	30	3	2	89.6	0.000	0.000
0.62	30	3	2	88.9	0.000	0.000
0.75	30	3	2	87.7	0.000	0.000
0.81	30	3	2	86.2	0.000	0.000
0.85	30	3	2	84.5	0.000	0.000
0.88	30	3	2	82.4	0.000	0.000
0.91	30	3	2	80.1	0.000	0.000
0.92	30	3	2	77.5	0.000	0.000
0.94	30	3	2	74.7	0.000	0.000
0.95	30	3	2	72.6	0.000	0.000
0.96	30	3	2	69.8	0.000	0.000
0.96	30	3	2	66.0	0.000	0.000
0.97	30	3	2	62.2	0.000	0.000

p-value is for the Wald or score test: $coverage = 95\%$ vs. $coverage \neq 95\%$.

Table 5.10: Coverage probability when $n_p = 100$ and p_i is distributed as double exponential

ρ_I	n_p	n_r	n_o	Coverage (%)	p-value	
					Wald	Score
0.08	100	3	2	92.5	0.000	0.000
0.25	100	3	2	91.3	0.000	0.000
0.46	100	3	2	89.5	0.000	0.000
0.62	100	3	2	89.0	0.000	0.000
0.75	100	3	2	88.5	0.000	0.000
0.81	100	3	2	88.3	0.000	0.000
0.85	100	3	2	88.1	0.000	0.000
0.88	100	3	2	87.6	0.000	0.000
0.91	100	3	2	87.2	0.000	0.000
0.92	100	3	2	86.1	0.000	0.000
0.94	100	3	2	84.9	0.000	0.000
0.95	100	3	2	83.4	0.000	0.000
0.96	100	3	2	81.4	0.000	0.000
0.96	100	3	2	78.4	0.000	0.000
0.97	100	3	2	75.5	0.000	0.000

p-value is for the Wald or score test: $coverage = 95\%$ vs. $coverage \neq 95\%$.

CHAPTER 6

MCMC STUDY OF THE RELIABILITY PARAMETERS

6.1 Introduction

Markov chain Monte Carlo (MCMC) employs Markov chains for the purpose of simulating posterior distributions. In recent years, this methodology has come to provide a unifying framework within which many complex problems can be analyzed using generic software. In my study, the confidence intervals for estimating the reliability parameters are asymptotically accurate. For large data sets they work fairly well. For small data sets, the accuracy of the estimators and confidence intervals is questionable; this is evident from the simulation results in Chapter 5. When n_r and n_o are relatively small, some calculated coverage probabilities are far lower than the nominal coverage of 95%. An alternative method for estimating the parameters is therefore required in this case. In this chapter, I shall explore the use of MCMC techniques for estimating the reliability parameters.

Among all the MCMC algorithms, Gibbs sampling is very widely used. Under the MCMC framework, all quantities which are not observed are treated as random.

The full three-way random effects model in Chapter 3 can be written in the Bayesian framework as

$$x_{ijk} = \mu + p_i + r_j + o_k + (pr)_{ij} + (po)_{ik} + (ro)_{jk} + e_{ijk}$$
$$i = 1, \cdots, n_p \quad j = 1, \cdots, n_r \quad k = 1, \cdots, n_o$$

where

$$(p_i|\sigma_p) \sim N(0, \sigma_p^2),$$
$$(r_j|\sigma_r) \sim N(0, \sigma_r^2),$$
$$(o_k|\sigma_o) \sim N(0, \sigma_o^2),$$
$$((pr)_{ij}|\sigma_{pr}) \sim N(0, \sigma_{pr}^2),$$
$$((po)_{ik}|\sigma_{po}) \sim N(0, \sigma_{po}^2),$$
$$((ro)_{ik}|\sigma_{ro}) \sim N(0, \sigma_{ro}^2),$$
$$(e_{ijk}|\sigma_e) \sim N(0, \sigma_e^2),$$

and

$$(x_{ijk}|\mu, p_i, r_j, o_k, (pr)_{ij}, (po)_{ik}, (ro)_{jk}, \sigma_e) \sim$$
$$N(\mu + p_i + r_j + o_k + (pr)_{ij} + (po)_{ik} + (ro)_{jk}, \sigma_e^2).$$

For the full three-way random effects model, under the aforementioned framework, the random effects, p_i, r_j, o_k, $(pr)_{ij}$, $(po)_{ik}$ and $(ro)_{jk}$ and the parameters, μ, σ_p^2, σ_r^2, σ_o^2, σ_{pr}^2, σ_{po}^2, σ_{ro}^2 and σ_e^2 are considered random.

It is assumed that X_{ijk} are independent random variables conditional on the aforementioned quantities, μ, p_i, r_j, o_k, $(pr)_{ij}$, $(po)_{ik}$, $(ro)_{jk}$ and σ_e.

A directed acyclic graph (DAG) can be used to represent these assumptions. This is depicted in Figure 6.1. Each random quantity in the model appears as a node in the graph. This graph is acyclic because it has no cycles; it is directed because nodes are linked by arrows (Spiegelhalter et al., 1995).

Notation 6.1 *Let v be a node in the graph, and V be the set of all nodes. A* ***parent*** *of v is defined to be any node with an arrow emanating from it pointing*

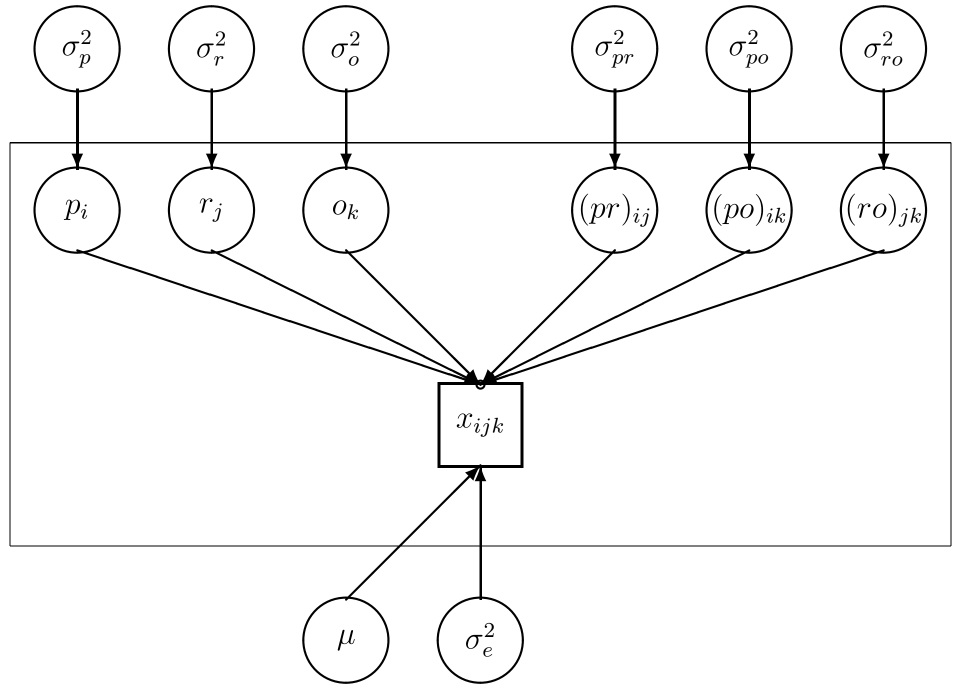

Figure 6.1: Directed acyclic graph representing the structure of the full random model

*to v and a **descendent** is any node on a directed path starting from v. The nodes without parents are **founders** in the structure.*

For example, in the following DAG for the full three-way model, σ_p^2 is the parent of p_i, p_i is one parent of x_{ijk}. The parameters σ_p^2, σ_r^2, σ_o^2, σ_{pr}^2, σ_{po}^2, σ_{ro}^2, σ_e^2 and μ are the founders in the structure.

The joint distribution of all random variables in the model is the product of the conditional distribution of each random variable given its parents. Let v denote a generic random variable in the structure. In this case, the joint distribution is

$$P(V) = \prod_{v \in V} p(v|\text{parents}[v]).$$

I chose conjugate prior distributions for the founders in the model. For this reason, the posterior distributions have the same form as their prior counterparts. The prior distributions are:

$$\mu \sim N(0, \eta),$$
$$\tau, \tau_p, \tau_r, \tau_o, \tau_{pr}, \tau_{po}, \tau_{ro} \sim \text{Ga}(\alpha, \beta)$$

where $\tau = 1/\sigma_e^2$, $\tau_p = 1/\sigma_p^2$, $\tau_r = 1/\sigma_r^2$, $\tau_o = 1/\sigma_o^2$, $\tau_{pr} = 1/\sigma_{pr}^2$, $\tau_{po} = 1/\sigma_{po}^2$, $\tau_{ro} = 1/\sigma_{ro}^2$. η, α and β are prior parameters.

It follows that the joint distribution is

$$\prod_i \prod_j \prod_k P(x_{ijk}|\mu, p_i, r_j, o_k, pr_{ij}, po_{ik}, ro_{jk}, \tau) \cdot \prod_i P(p_i|\tau_p) \cdot \prod_j P(r_j|\tau_r) \cdot$$
$$\prod_k P(o_k|\tau_o) \cdot \prod_i \prod_j P((pr)_{ij}|\tau_{pr}) \cdot \prod_i \prod_k P((po)_{ik}|\tau_{po}) \cdot \prod_j \prod_k P((ro)_{jk}|\tau_{ro}) \cdot$$
$$P(\mu) \cdot P(\tau) \cdot P(\tau_p) \cdot P(\tau_r) \cdot P(\tau_o) \cdot P(\tau_{pr}) \cdot P(\tau_{po}) \cdot P(\tau_{ro}).$$

6.2 Full Conditional Distribution of Random Variables

The Gibbs sampler is a special case of Metropolis–Hastings algorithm whereby all proposed moves are accepted. It provides samples from the full conditional posterior distributions of all random variables in the model.

Notation 6.2 *Let $\boldsymbol{X}$ be a vector of k components. The full conditional posterior distribution is the distribution of the i^{th} component of $\boldsymbol{X}$ conditioned on all the remaining components.*

In the structure of a DAG, the full conditional posterior distribution for a node is the distribution of that node given current or known values for all the other nodes in the graph. For any node v, let V_{-v} denote the full set of nodes with v removed. The full conditional posterior distribution $P(v|V_{-v})$ takes the form

$$\begin{aligned} P(v|V_{-v}) &\propto P(v, V_{-v}) \\ &\propto \text{terms in } P(V) \text{ containing } v \\ &= P(v|\text{parents}[v]) \times \prod_{w \in \text{children}[v]} P(w|\text{parents}[w]). \end{aligned}$$

In what follows I derive the full conditional posterior distribution for each random (non-observed) variable in the full three-way random effects model. Let "$\cdot$" denote all data nodes, random effect nodes and parameter nodes except v, (i.e. V_{-v}), $\mu_{ijk} = \mu + p_i + r_j + o_k + (pr)_{ij} + (po)_{ik} + (ro)_{jk}$, $\mu_{-p_i} = \mu + r_j + o_k + (pr)_{ij} + (po)_{ik} + (ro)_{jk}$, $\mu_{-r_j} = \mu + p_i + o_k + (pr)_{ij} + (po)_{ik} + (ro)_{jk}$ and etc. In the derivations, the "$\propto$" proportion symbol is used instead of "$=$" to avoid the need for keeping track of constants. Then the full conditional posterior distribution of p_i is

$$\begin{aligned} P(p_i|\cdot) &\propto \left\{ \prod_{j=1}^{n_r} \prod_{k=1}^{n_o} \exp\left(-\frac{1}{2}\tau (x_{ijk} - \mu_{ijk})^2 \right) \right\} \cdot \exp\left(-\frac{1}{2}\tau_p p_i^2 \right) \\ &= \exp\left\{ -\frac{1}{2}\tau \Big[\sum_j \sum_k \left(x_{ijk} - \mu_{-p_i} - p_i \right)^2 \Big] - \frac{1}{2}\tau_p p_i^2 \right\} \\ &= \exp\Bigg\{ -\frac{\tau}{2} \Big[\sum_j \sum_k \left(x_{ijk} - \mu_{-p_i} \right)^2 \\ &\qquad - 2p_i \sum_j \sum_k \left(x_{ijk} - \mu_{-p_i} \right) + n_r n_o p_i^2 \Big] - \frac{1}{2}\tau_p p_i^2 \Bigg\} \\ &\propto \exp\left\{ -\frac{1}{2} \Big[(n_r n_o \tau + \tau_p) p_i^2 - 2\tau p_i \sum_j \sum_k \left(x_{ijk} - \mu_{-p_i} \right) \Big] \right\} \\ &\propto \exp\left\{ -\frac{1}{2}(n_r n_o \tau + \tau_p) \left[p_i - \frac{\tau \sum_j \sum_k \left(x_{ijk} - \mu_{-p_i} \right)}{n_r n_o \tau + \tau_p} \right]^2 \right\}. \end{aligned}$$

So $p_i|\cdot$ has a Gaussian distribution and

$$(p_i|\cdot) \sim N\left(\frac{\tau \sum_j \sum_k \left(x_{ijk} - \mu_{p_i} \right)}{n_r n_o \tau + \tau_p}, \frac{1}{n_r n_o \tau + \tau_p} \right).$$

The derivation of the full conditional posterior distributions for the other random effects is analogous. The full conditional posterior distribution of r_j is

$$\begin{aligned}P(r_j|\cdot) &\propto \left\{\prod_{i=1}^{n_p}\prod_{k=1}^{n_o} \exp\left(-\frac{1}{2}\tau(x_{ijk}-\mu_{ijk})^2\right)\right\}\cdot \exp\left(-\frac{1}{2}\tau_r r_j^2\right)\\ &\propto \exp\left\{-\frac{1}{2}(n_p n_o \tau+\tau_r)\left[r_j - \frac{\tau\sum_i\sum_k\left(x_{ijk}-\mu_{-r_j}\right)}{n_p n_o\tau+\tau_r}\right]^2\right\}.\end{aligned}$$

So $r_j|\cdot$ has a Gaussian distribution and

$$(r_j|\cdot) \sim N\left(\frac{\tau\sum_i\sum_k\left(x_{ijk}-\mu_{-r_j}\right)}{n_p n_o\tau+\tau_r}, \frac{1}{n_p n_o\tau+\tau_r}\right).$$

The full conditional posterior distribution of o_k is

$$\begin{aligned}P(o_k|\cdot) &\propto \left\{\prod_{i=1}^{n_p}\prod_{j=1}^{n_r} \exp\left(-\frac{1}{2}\tau(x_{ijk}-\mu_{ijk})^2\right)\right\}\cdot \exp\left(-\frac{1}{2}\tau_o o_k^2\right)\\ &\propto \exp\left\{-\frac{1}{2}(n_p n_r \tau+\tau_o)\left[o_k - \frac{\tau\sum_i\sum_j\left(x_{ijk}-\mu_{-o_k}\right)}{n_p n_r\tau+\tau_o}\right]^2\right\}.\end{aligned}$$

So $o_k|\cdot$ has a Gaussian distribution and

$$(o_k|\cdot) \sim N\left(\frac{\tau\sum_i\sum_j\left(x_{ijk}-\mu_{-o_k}\right)}{n_p n_r\tau+\tau_o}, \frac{1}{n_p n_r\tau+\tau_o}\right).$$

The full conditional posterior distribution of $(pr)_{ij}$ is

$$\begin{aligned}P\left((pr)_{ij}|\cdot\right) &\propto \left\{\prod_{k=1}^{n_o} \exp\left(-\frac{1}{2}\tau(x_{ijk}-\mu_{ijk})^2\right)\right\}\cdot \exp\left(-\frac{1}{2}\tau_{pr}(pr)_{ij}^2\right)\\ &\propto \exp\left\{-\frac{1}{2}(n_o \tau+\tau_{pr})\left[(pr)_{ij} - \frac{\tau\sum_k\left(x_{ijk}-\mu_{-(pr)_{ij}}\right)}{n_o\tau+\tau_{pr}}\right]^2\right\}.\end{aligned}$$

So $(pr)_{ij}|\cdot$ has a Gaussian distribution and

$$((pr)_{ij}|\cdot) \sim N\left(\frac{\tau\sum_k\left(x_{ijk}-\mu_{-(pr)_{ij}}\right)}{n_o\tau+\tau_{pr}}, \frac{1}{n_o\tau+\tau_{pr}}\right).$$

The full conditional posterior distribution of $(po)_{ik}$ is

$$\begin{aligned} P\left((po)_{ik}|\cdot\right) &\propto \left\{\prod_{j=1}^{n_r}\exp\left(-\frac{1}{2}\tau(x_{ijk}-\mu_{ijk})^2\right)\right\}\cdot\exp\left(-\frac{1}{2}\tau_{po}(po)_{ik}^2\right) \\ &\propto \exp\left\{-\frac{1}{2}(n_r\tau+\tau_{po})\left[(po)_{ik}-\frac{\tau\sum_j\left(x_{ijk}-\mu_{-(po)_{ik}}\right)}{n_r\tau+\tau_{po}}\right]^2\right\}. \end{aligned}$$

So $(po)_{ik}|\cdot$ has a Gaussian distribution and

$$((po)_{ik}|\cdot) \sim N\left(\frac{\tau\sum_j\left(x_{ijk}-\mu_{-(po)_{ik}}\right)}{n_r\tau+\tau_{po}}, \frac{1}{n_r\tau+\tau_{po}}\right).$$

The full conditional posterior distribution of $(ro)_{jk}$ is

$$\begin{aligned} P\left((ro)_{jk}|\cdot\right) &\propto \left\{\prod_{i=1}^{n_p}\exp\left(-\frac{1}{2}\tau(x_{ijk}-\mu_{ijk})^2\right)\right\}\cdot\exp\left(-\frac{1}{2}\tau_{ro}(ro)_{jk}^2\right) \\ &\propto \exp\left\{-\frac{1}{2}(n_p\tau+\tau_{ro})\left[(ro)_{jk}-\frac{\tau\sum_i\left(x_{ijk}-\mu_{-(ro)_{jk}}\right)}{n_p\tau+\tau_{ro}}\right]^2\right\}. \end{aligned}$$

So $(ro)_{jk}|\cdot$ has a Gaussian distribution and

$$((ro)_{jk}|\cdot) \sim N\left(\frac{\tau\sum_i\left(x_{ijk}-\mu_{-(ro)_{jk}}\right)}{n_p\tau+\tau_{ro}}, \frac{1}{n_p\tau+\tau_{ro}}\right).$$

The full conditional posterior distribution of μ is

$$P(\mu|\cdot) \propto \left\{\prod_{i=1}^{n_p}\prod_{j=1}^{n_r}\prod_{k=1}^{n_o} \exp\left(-\frac{1}{2}\tau(x_{ijk}-\mu_{ijk})^2\right)\right\} \cdot \exp\left(-\frac{1}{2\eta}\mu^2\right)$$

$$\propto \exp\Bigg\{ -\frac{1}{2}\left(n_p n_r n_o \tau + \frac{1}{\eta}\right)$$

$$\left[\mu - \frac{\tau \sum_i \sum_j \sum_k \left(x_{ijk}-\mu_{-\mu}\right)}{n_p n_r n_o \tau + \frac{1}{\eta}}\right]^2\Bigg\}.$$

Thus $\mu|\cdot$ has a Gaussian distribution and

$$(\mu|\cdot) \sim N\left(\frac{\tau \sum_i \sum_j \sum_k \left(x_{ijk}-\mu_{-\mu}\right)}{n_p n_r n_o \tau + \frac{1}{\eta}}, \frac{1}{n_p n_r n_o \tau + \frac{1}{\eta}}\right).$$

The full conditional posterior distributions of all random effects and μ are Gaussian, as a result of the choice of conjugate prior distributions.

In what follows, I derive the full conditional posterior distributions of variance components. The full conditional posterior distribution of τ_p is

$$P(\tau_p|\cdot) \propto \tau_p^{\alpha-1}\exp(-\tau_p/\beta)\prod_{i=1}^{n_p}\left[\tau_p^{1/2}\exp(-\tau_p p_i^2/2)\right]$$

$$= \tau_p^{\alpha+n_p/2-1}\exp\left(-\tau_p/\beta - \tau_p\sum_{i=1}^{n_p} p_i^2/2\right)$$

$$= \tau_p^{\alpha+n_p/2-1}\exp\left[-\tau_p(1/\beta + \sum_{i=1}^{n_p} p_i^2/2)\right]$$

so

$$(\tau_p|\cdot) \sim \mathrm{Ga}\left(\alpha + \mathrm{n_p}/2, 1/(1/\beta + \sum_{\mathrm{i}=1}^{\mathrm{n_p}} \mathrm{p_i^2}/2)\right).$$

By analogy with the above, the full conditional posterior distribution of other variance components are

$$P(\tau_r|\cdot) \propto \tau_r^{\alpha-1}\exp(-\tau_r/\beta)\prod_{j=1}^{n_r}\left[\tau_r^{1/2}\exp(-\tau_r r_j^2/2)\right]$$
$$= \tau_r^{\alpha+n_r/2-1}\exp\left[-\tau_r(1/\beta+\sum_{j=1}^{n_r} r_j^2/2)\right],$$

$$P(\tau_o|\cdot) \propto \tau_o^{\alpha-1}\exp(-\tau_o/\beta)\prod_{k=1}^{n_o}\left[\tau_o^{1/2}\exp(-\tau_o r_j^2/2)\right]$$
$$= \tau_o^{\alpha+n_o/2-1}\exp\left[-\tau_o(1/\beta+\sum_{k=1}^{n_o} r_j^2/2)\right],$$

$$P(\tau_{pr}|\cdot) \propto \tau_{pr}^{\alpha-1}\exp(-\tau_{pr}/\beta)\prod_{i=1}^{n_p}\prod_{j=1}^{n_r}\left[\tau_{pr}^{1/2}\exp\big(-\tau_{pr}(pr)_{ij}^2/2\big)\right]$$
$$= \tau_{pr}^{\alpha+n_p n_r/2-1}\exp\left[-\tau_{pr}(1/\beta+\sum_{i=1}^{n_p}\sum_{j=1}^{n_r}(pr)_{ij}^2/2)\right],$$

$$P(\tau_{po}|\cdot) \propto \tau_{po}^{\alpha-1}\exp(-\tau_{po}/\beta)\prod_{i=1}^{n_p}\prod_{k=1}^{n_o}\left[\tau_{po}^{1/2}\exp\big(-\tau_{po}(po)_{ik}^2/2\big)\right]$$
$$= \tau_{po}^{\alpha+n_p n_o/2-1}\exp\left[-\tau_{po}(1/\beta+\sum_{i=1}^{n_p}\sum_{k=1}^{n_o}(po)_{ik}^2/2)\right],$$

$$P(\tau_{ro}|\cdot) \propto \tau_{ro}^{\alpha-1}\exp(-\tau_{ro}/\beta)\prod_{j=1}^{n_r}\prod_{k=1}^{n_o}\left[\tau_{ro}^{1/2}\exp\big(-\tau_{ro}(ro)_{jk}^2/2\big)\right]$$
$$= \tau_{ro}^{\alpha+n_r n_o/2-1}\exp\left[-\tau_{ro}(1/\beta+\sum_{j=1}^{n_r}\sum_{k=1}^{n_o}(ro)_{jk}^2/2)\right],$$

$$P(\tau|\cdot) \propto \tau^{\alpha-1}\exp(-\tau/\beta)\prod_{i=1}^{n_p}\prod_{j=1}^{n_r}\prod_{k=1}^{n_o}\left[\tau^{1/2}\exp\big(-\tau(x_{ijk}-\mu_{ijk})^2/2\big)\right]$$
$$= \tau^{\alpha+n_p n_r n_o/2-1}\exp\left[-\tau\big(1/\beta+\sum_{i=1}^{n_p}\sum_{j=1}^{n_r}\sum_{k=1}^{n_o}(x_{ijk}-\mu_{ijk})^2/2\big)\right].$$

Therefore they all have Gamma distributions and

$$(\tau_r|\cdot) \sim \text{Ga}\left(\alpha + \text{n}_\text{r}/2, 1/(1/\beta + \sum_{\text{j}=1}^{\text{n}_\text{r}} \text{r}_\text{j}^2/2)\right),$$

$$(\tau_o|\cdot) \sim \text{Ga}\left(\alpha + \text{n}_\text{o}/2, 1/(1/\beta + \sum_{\text{k}=1}^{\text{n}_\text{o}} \text{o}_\text{k}^2/2)\right),$$

$$(\tau_{pr}|\cdot) \sim \text{Ga}\left(\alpha + \text{n}_\text{p}\text{n}_\text{r}/2, 1/(1/\beta + \sum_{\text{i}=1}^{\text{n}_\text{p}}\sum_{\text{j}=1}^{\text{n}_\text{r}} (\text{pr})_\text{ij}^2/2)\right),$$

$$(\tau_{po}|\cdot) \sim \text{Ga}\left(\alpha + \text{n}_\text{p}\text{n}_\text{o}/2, 1/(1/\beta + \sum_{\text{i}=1}^{\text{n}_\text{p}}\sum_{\text{k}=1}^{\text{n}_\text{o}} (\text{po})_\text{ik}^2/2)\right),$$

$$(\tau_{ro}|\cdot) \sim \text{Ga}\left(\alpha + \text{n}_\text{r}\text{n}_\text{o}/2, 1/(1/\beta + \sum_{\text{j}=1}^{\text{n}_\text{r}}\sum_{\text{k}=1}^{\text{n}_\text{o}} (\text{ro})_\text{jk}^2/2)\right),$$

$$(\tau|\cdot) \sim \text{Ga}\left(\alpha + \text{n}_\text{p}\text{n}_\text{r}\text{n}_\text{o}/2, 1/(1/\beta + \sum_{\text{i}=1}^{\text{n}_\text{p}}\sum_{\text{j}=1}^{\text{n}_\text{r}}\sum_{\text{k}=1}^{\text{n}_\text{o}} (\text{x}_\text{ijk} - \mu_\text{ijk})^2/2)\right).$$

6.3 Point and Confidence Interval Estimation for ICC

I simulate the variance components using the full conditional posterior distributions calculated in the previous section. After each iteration, ICC is calculated by

$$\rho_I = \frac{\sigma_p^2}{\sigma_p^2 + \sigma_r^2 + \sigma_o^2 + \sigma_{pr}^2 + \sigma_{po}^2 + \sigma_{ro}^2 + \sigma_e^2}$$

which results in the simulated draws from the posterior distribution for the ICC.

Suppose a Markov chain consists of a sequence $\{X_t, t = 1, \ldots, n\}$. After a sufficiently long burn-in of m iterations, $\{X_t, t = m+1, \ldots, n\}$ remains. Discarding the burn-in samples, the mean and variance of X are estimated by

$$\bar{X} = \frac{1}{n-m} \sum_{t=m+1}^{n} X_t \tag{6.1}$$

and

$$S^2 = \frac{1}{n-m-1} \sum_{t=m+1}^{n} (X_t - \bar{X})^2.$$

A $100(1-2p)\%$ credible interval $[c_p, c_{1-p}]$ for X can be estimated by setting c_p equal to the p^{th} quantile of $\{X_t, t = m+1, \ldots, n\}$, and c_{1-p} equal to the $(1-p)^{th}$ quantile.

The most obvious informal method for determining run length n is to run several chains in parallel, with different starting values, and compare the estimates from (6.1). If they do not agree adequately, n must be increased (Gilks et al., 1995). Geyer (1992) points out that, since m is likely to be less than 1% of the total length of a run, calculating the proper burn-in length to obtain adequate precision is unnecessary for the estimator in (6.1). Geyer suggests using a burn-in m of length between 1% and 2% of the run length n. This makes sense as long as extreme starting values are avoided. In order to avoid risk, I used a burn-in m which is at least 40% of the run length n.

6.4 Results of Analyzing Film Data using MCMC Method

The data from the orthopedic study designed to assess the reliability of readings of x-ray films are used here again. There were x-ray films from 30 patients examined by 3 different examiners on 2 different occasions. The numbers of raters and occasions are rather small. The point estimate and credible interval of ICC for each of the 16 measurements are calculated using MCMC. The results are compared with ANOVA based estimates. The SAS program for this is given in Appendix B.

To determine the run length n for each measurement, two chains are run in parallel with over-dispersed starting values. For most of the measurements, 5000 is large enough for the value of n to get chains producing close estimates. The lower and upper limits of the 95% credible interval are chosen to be the 0.025 and 0.975 quantiles respectively.

Table 6.1 lists the calculated point estimate and credible (respectively, confidence) interval of ICC using the MCMC (respectively, ANOVA) based method. For all of the measurements, the point estimates and the upper limits of the intervals are very close, although the estimates from MCMC are usually a little bit smaller than their counterparts from ANOVA. The lower limits are not that close. In some cases, this difference is substantial (e.g., `RRT11T12`). The big discrepancy between lower limits from the MCMC and ANOVA methods for some measurements cast some doubts about convergence. To clarify this problem, two longer chains with $n = 10000$ and $m = 8500$ were run for `RRT11T12`. Now I got ICC (0.74), lower limit (0.43) and upper limit (0.87). While the estimates of ICC and upper limit stay stable, the lower limit is closer to its counterpart from ANOVA, although still smaller.

In addition to the length of the Markov chains, the scheme of generating random numbers in SAS may also have some effect on the results. I chose `RRAVT6T7` to explore this issue. The point and credible interval estimates were calculated from each of the 5 chains using 5 different apart seeds, as well as that from the combined chains. The results are listed in Table 6.2. From the results I can see that estimates of ICC and upper limit are stable as expected, but the lower limit varies slightly with the seed. It is a good strategy to try multiple chains and avoid the chains that produce extreme results.

The smaller lower limits in turn cause more conservative credible intervals. Recall that the justification for turning to the MCMC method was the low coverage probability of the ANOVA confidence interval. The coverage probability of each of the MCMC credible intervals, which are wider than ANOVA confidence intervals, should be closer to nominal. Next I will show this with simulation results.

Table 6.1: Point and 95% confidence interval estimates for ρ_I for 16 orthopedic measurements, comparing ANOVA and MCMC (with the mean of the Markov chain as the estimate of ICC and $[c_{0.025}, c_{0.975}]$ as CI) methods

Measure	n	m	ICC		Lower limit		Upper limit		Length of CI	
			MCMC	ANOVA	MCMC	ANOVA	MCMC	ANOVA	MCMC	ANOVA
T1T12VER	5000	3500	0.99	1.00	0.95	1.00	1.00	1.00	0.05	0.00
T2T11VER	5000	3500	0.99	1.00	0.94	1.00	1.00	1.00	0.06	0.00
RRAVT1T2	5000	3500	0.88	0.91	0.66	0.86	0.95	0.95	0.28	0.09
RRAVT2T3	5000	3500	0.91	0.94	0.75	0.90	0.96	0.97	0.21	0.07
RRAVT3T4	5000	3500	0.85	0.88	0.61	0.81	0.93	0.94	0.32	0.12
RRAVT4T5	5000	3500	0.90	0.93	0.75	0.88	0.95	0.96	0.20	0.08
RRAVT5T6	5000	3500	0.86	0.90	0.64	0.83	0.93	0.94	0.29	0.11
RRAVT6T7	5000	3500	0.71	0.76	0.44	0.63	0.84	0.86	0.41	0.23
RRAVT7T8	5000	3500	0.75	0.80	0.46	0.70	0.87	0.89	0.41	0.19
RRAVT8T9	5000	3500	0.72	0.79	0.42	0.68	0.85	0.88	0.43	0.20
RRAT9T10	5000	3500	0.76	0.81	0.51	0.71	0.87	0.89	0.36	0.18
RRT10T11	5000	3500	0.78	0.83	0.51	0.73	0.89	0.90	0.38	0.18
RRT11T12	5000	3500	0.73	0.80	0.25	0.70	0.87	0.89	0.62	0.19
ARAT1T12	5000	3500	0.98	0.99	0.93	0.98	0.99	0.99	0.06	0.02
ARAT2T11	5000	3500	0.97	0.98	0.91	0.97	0.99	0.99	0.07	0.02
ARAT3T10	5000	3500	0.97	0.98	0.92	0.96	0.99	0.99	0.07	0.03

Table 6.2: Point and 95% confidence interval estimates for ρ_I for `RRAVT6T7` from MCMC method with $n = 5000$ and $m = 3500$

Seed	ICC	Lower limit	Upper limit	Length of CI
1	0.72	0.43	0.85	0.42
1234	0.72	0.48	0.85	0.37
12300530	0.72	0.50	0.85	0.35
20081017	0.70	0.30	0.85	0.54
12300530567890	0.70	0.30	0.85	0.54
combined	0.71	0.41	0.85	0.44

6.5 Coverage Probability Validation through Simulation

The data were simulated in the same way as in Chapter 5. The true ICC values range from 0.1 to 0.9. For each value of ICC, 1000 samples with 30 subjects, 3 raters and 2 occasions were generated and then each sample was analyzed using the MCMC and ANOVA methods. The proportion of samples which gave a CI covering the true ICC value (ρ_I) is the coverage probability. Each sample was analyzed using two Markov chains of length 2000 (5000 for ICC 0.1, 0.2 and 0.3 , for small ICC values, the Markov chain converges more slowly than for large ICC values). From Table 6.3, I can see that when ICC reaches 0.8 the coverage probabilities from the MCMC method are equal to 95% while the coverage probabilities from the ANOVA method are lower than 95%. In practice, we hope ICC exceeds 0.85. From the coverage probability point of view, the MCMC method is superior to ANOVA.

I also studied the location of the MCMC estimator of ICC from the simulation results. The center of the histogram of the ICC estimates from the 1000 sample usually shifts to the left of the true value of ICC. So most of time the ICC estimate is smaller than the true ICC. To choose the median instead of the mean of the Markov chain as the estimate of ICC will alleviate this bias.

Table 6.3: Coverage probability of the 95% credible interval calculated from MCMC and ANOVA methods

ρ_I	Coverage(%)		p-value	
	MCMC	ANOVA	MCMC	ANOVA
0.1	93.5	95.7	0.054	0.275
0.2	91.5	94.3	< 0.001	0.340
0.3	92.6	94.1	0.004	0.227
0.4	90.0	93.7	< 0.001	0.090
0.5	91.3	93.2	< 0.001	0.023
0.6	92.3	92.9	0.001	0.010
0.7	92.8	92.7	0.007	0.005
0.8	93.5	92.2	0.054	0.001
0.9	93.6	91.9	0.070	< 0.001

p-value is for the Wald test: $coverage = 95\%$ vs. $coverage \neq 95\%$.

6.6 Discussion and Conclusion

Unlike the ANOVA method, the MCMC method is more subjective, such as the choice the length of the Markov chain and the scheme generating random numbers. When the estimates from the MCMC method are doubtful, different options such as increasing the length of the Markov chain and running multiple chains should be tried to see if the problem is resolved. From the data analysis results it is clear that usually the credible interval from the MCMC method is twice as wide as the confidence interval from the ANOVA method. When the estimate of ICC is high, the length of the CI from the MCMC method is longer than that from the ANOVA method but still acceptable compared to the value of ICC. Because the CI from the ANOVA method has an undercoverage problem when the ICC values are high, I recommend the MCMC method for analyses of small data sets.

REFERENCES

[1] Bartko, J.J. (1966). *The Intraclass Correlation Coefficient as a Measure of Reliability.* Psychological Reports **19**, no. 1: 3–11.

[2] Davison, A.C., and Hinkley, D.V. (1997). *Bootstrap Methods and their Application.* Cambridge: Cambridge University Press.

[3] Dunn, G. (2004). *Statistical Evaluation of Measurement Errors* (2nd ed.). London: Arnold.

[4] Ebel, R.L. (1951). *Estimation of the Reliability of Ratings.* Psychometrika **16**, no. 4: 407–424.

[5] Efron, B. (1987). *Better Bootstrap Confidence Intervals (with discussion).* Journal of the American Statistical Association **82**, no. 397: 171–200.

[6] Fisher, R.A. (1925). *Statistical Methods for Research Workers.* Edinburgh: Oliver and Boyd.

[7] Fleiss, J.L., and Shrout, P.E. (1978). *Approximate Interval Estimation for a Certain Intraclass Correlation Coefficient.* Psychometrika **43**, no. 2: 259–262.

[8] Geyer, C.J. (1992). *Practical Markov chain Monte Carlo.* Statist. Sci. **7**, no. 4: 473–483.

[9] Gilks, W.R., Richardson, S., and Spiegelhalter, D.J. (1995). *Introducing Markov chain Monte Carlo.* In W.R. Gilks, S. Richardson and D.J. Spiegelhalter (Eds.), *Markov Chain Monte Carlo in Practice* (pp. 1–19). London: Chapman & Hall.

[10] Haggard, E.A. (1958). *Intraclass Correlation and the Analysis of Variance.* New York: Dryden Press.

[11] Harrison, D.E., Cailliet, R., Harrison, D.D., Janik, T.J., and Holland, B. (2001). *Reliability of Centroid, Cobb and Harrison Posterior Tangent Methods: Which to Choose for Analysis of Thoracic Kyphosis.* Spine **26**, no. 11: 227–234.

[12] Johnson, N.L., Kotz, S., and Balakrishnan, N. (1994). *Continuous Univariate Distributions.* New York: Wiley & Sons.

[13] Johnson, V.E. (1996). *On Bayesian Analysis of Multirater Ordinal Data: an Application to Automated Essay Grading.* Journal of the American Statistical Association **91**, no. 433: 42–51.

[14] Mandrekar, J., Sargent, D.J., Novotny, P.J., and Sloan, J.A. (1999). *A General Gibbs Sampling Algorithm for Analyzing Linear Models Using the SAS System.* Statistics, Data Analysis, and Modeling at SAS Users Group International Conference, Paper 266.

[15] McCulloch, C.E., and Searle, S.R. (2001). *Generalized, Linear, and Mixed Models.* New York: John Wiley & Sons.

[16] McGraw, K.O., and Wong, S.P. (1996). *Forming Inferences about Some Intraclass Correlation Coefficients.* Psychological Methods **1**, no. 1: 30–46.

[17] Olkin, I., and Pratt, J.W. (1958). *Unbiased Estimation of Certain*

Correlation Coefficients. Annal of Mathematical Statistics **29**, no. 1: 201–211.

[18] Satterthwaite, F.E. (1946). *An Approximate Distribution of Estimates of Variance Components.* Biometrics Bulletin **2**, no. 6: 110–114.

[19] Searle, S.R., Casella, G., and McCulloch, C.E. (1992). *Variance Components.* New York: John Wiley & Sons.

[20] Shavelson, R.J., and Webb, N.M. (1991). *Generalizability theory : a primer.* Newbury Park, Calif.: Sage Publications.

[21] Shoukri, M.M. (2004). *Measures of Interobserver Agreement.* Boca Raton: Chapman & Hall/CRC.

[22] Shrout, P.E., and Fleiss, J.L. (1979). *Intraclass Correlations: Uses in Assessing Rater Reliability.* Psychological Bulletin **86**, no. 2: 420–428.

[23] Snedecor, G.W., and Cochran, W.G. (1989). *Statistical Methods* (8th ed.). Ames: Iowa State University Press.

[24] Spiegelhalter, D.J., Best, N.G., Gilks, W.R., and Inskip, H. (1995). *Hepatitis B: a Case Study in MCMC Methods.* In W.R. Gilks, S. Richardson and D.J. Spiegelhalter (Eds.), *Markov Chain Monte Carlo in Practice* (pp. 20–42). London: Chapman & Hall.

APPENDIX A

SAS Code Estimating ICC by ANOVA and Bootstrap Methods

```
%Macro analyze(data=,response=,out=);
DATA &data;
    SET &data;
    WHERE MEASURE="&response";
RUN;

PROC SORT DATA=&data;
    BY &byobs FILM;
RUN;

/*Calculate ICC from data */
PROC TRANSPOSE DATA=&data OUT=TWO;
    ID MEASURE;
    BY &byobs FILM;
RUN;

DATA TWO;SET TWO;
    RATER=SUBSTR(_NAME_,1,1);
    OCCAS=SUBSTR(_NAME_,2,1);
RUN;

PROC SORT DATA=TWO;
    BY &byobs _NAME_ FILM;
RUN;
```

```
/* Analysis */
PROC GLM DATA=TWO OUTSTAT=STAT NOPRINT;
    CLASS FILM RATER OCCAS;
    MODEL &response=FILM RATER OCCAS FILM*RATER RATER*OCCAS FILM*OCCAS/SS3;
    BY &by;
RUN;

DATA MS;SET STAT;
    MS=SS/DF;
    KEEP &BY _NAME_ _SOURCE_ MS;
RUN;

DATA DF;SET STAT;
    KEEP &BY _NAME_ _SOURCE_ DF;
RUN;

PROC TRANSPOSE DATA=MS OUT=MS1;
    BY  &BY _NAME_;
    ID _SOURCE_;
    VAR MS;
RUN;

PROC TRANSPOSE DATA=DF OUT=DF1 PREFIX=DF;
    BY &BY _NAME_;
    ID _SOURCE_;
    VAR DF;
RUN;

DATA THREE;
    SET MS1;
    SET DF1;
    BY &BY;
    rename ERROR= MS_e; rename FILM=MS_p; rename RATER=MS_r; rename OCCAS=MS_o;
    rename FILM_RATER=MS_pr; rename RATER_OCCAS=MS_ro; rename FILM_OCCAS=MS_po;

    rename DFERROR=DF_e; rename DFFILM=DF_p; rename DFRATER=DF_r; rename DFOCCAS=DF_o;
    rename DFFILM_RATER=DF_pr; rename DFRATER_OCCAS=DF_ro; rename DFFILM_OCCAS=DF_po;

    n_p=DFFILM+1; n_r=DFRATER+1; n_o=DFOCCAS+1;
RUN;

/* Calculation estimate and CI of ICC*/
DATA &out;SET THREE;
    numer1=(MS_p+MS_e-MS_pr-MS_po)/(n_r*n_o);
    denom1=MS_p/(n_r*n_o)+MS_r/(n_p*n_o)+MS_o/(n_p*n_r)+MS_pr*(1/n_o-1/(n_r*n_o)-1/(n_p*n_o))
```

```
       +MS_po*(1/n_r-1/(n_r*n_o)-1/(n_p*n_r))+MS_ro*(1/n_p-1/(n_p*n_o)-1/(n_p*n_r))
       +MS_e*(1+1/(n_r*n_o)+1/(n_p*n_o)+1/(n_p*n_r)-1/n_p-1/n_r-1/n_o);
    ICC=numer1/denom1;         /* estimate of ICC*/
    Keep &BY ICC;
RUN;
%MEND analyze;

%Macro boot(data=, response=,  samples=,   random=);/* uniform bootstrap resampling */

options nonotes;

%let alpha=0.05;
%local by byobs;

DATA data_&response;
    SET &data;
    WHERE MEASURE="&response";
RUN;

%global _nobs;
Data _null_;
      call symput('_nobs',trim(left(put(_nobs,12.))));
      if 0 then set data_&response nobs=_nobs;
      stop;
Run;

/*** Generate the bootstrap data   ***/

Data BOOTDATA;/*view=BOOTDATA*/
     * drop  _i;
      do _sample_=1 to &samples;
         do _i=1 to &_nobs;
               _j=ceil(ranuni(&random)*&_nobs);    /* choose a cell */
               set data_&response point=_j;output;
         end;
      end;
      stop;
Run;


/*** Calculate point and confidence interval estimates for ICC from original data ***/
%Let by=;
%Let byobs=;
%analyze(data=data_&response,response=&response, out=_ACTUALL_);

DATA CI_theo;SET THREE;
```

```
    numer1=(MS_p+MS_e-MS_pr-MS_po)/(n_r*n_o);
    denom1=MS_p/(n_r*n_o)+MS_r/(n_p*n_o)+MS_o/(n_p*n_r)+MS_pr*(1/n_o-1/(n_r*n_o)-1/(n_p*n_o))
       +MS_po*(1/n_r-1/(n_r*n_o)-1/(n_p*n_r))+MS_ro*(1/n_p-1/(n_p*n_o)-1/(n_p*n_r))
       +MS_e*(1+1/(n_r*n_o)+1/(n_p*n_o)+1/(n_p*n_r)-1/n_p-1/n_r-1/n_o);
    ICC=numer1/denom1;         /* estimate of ICC*/

    ICCR=ICC/(1-ICC);
    e=(n_r/n_p)*ICCR;
    f=(n_o/n_p)*ICCR;
    g=1+n_r*ICCR-(n_r/n_p)*ICCR;
    h=1+n_o*ICCR-(n_o/n_p)*ICCR;
    i=(n_r*n_o-n_r-n_o)/n_p*ICCR;
    j=(n_r/n_p+n_o/n_p-n_r-n_o-n_r*n_o/n_p+n_r*n_o)*ICCR-1;

    numer2=(e*MS_r+f*MS_o+g*MS_pr+h*MS_po+i*MS_ro+j*MS_e)**2;
    denom2=(e*MS_r)**2/DF_r+(f*MS_o)**2/DF_o+(g*MS_pr)**2/DF_pr
         +(h*MS_po)**2/DF_po+(i*MS_ro)**2/DF_ro+(j*MS_e)**2/DF_e;

    *v=ROUND(numer2/denom2,1);  /* denominator df */
    v=numer2/denom2;
    q=(n_r/n_p)*MS_r+(n_o/n_p)*MS_o+(n_r-n_r/n_p)*MS_pr+(n_o-n_o/n_p)*MS_po
     +(n_r*n_o-n_r-n_o)/n_p*MS_ro+(n_r+n_o-n_r*n_p-n_o*n_p-n_r*n_o+n_p*n_r*n_o)/n_p*MS_e;

    E025 = FINV(0.025,DF_p,v);
    E975 = FINV(0.975,DF_p,v);
    E950 = FINV(0.95,DF_p,v);   /* For one-side lower limit */

  /* two-sided CI limits */
    L_limit = (MS_p-E975*(MS_pr+MS_po-MS_e))/(E975*(q-MS_pr-MS_po+MS_e)+MS_p);
    U_limit = (MS_p-E025*(MS_pr+MS_po-MS_e))/(E025*(q-MS_pr-MS_po+MS_e)+MS_p);
  /* one-sided lower limit */
    Lrho   =(MS_p-E950*(MS_pr+MS_po-MS_e))/(E950*(q-MS_pr-MS_po+MS_e)+MS_p);

    Keep _Name_  ICC L_limit U_limit Lrho;
RUN;

DATA _NULL_;
    SET _ACTUALL_;
    CALL symput('estimate',ICC);
RUN;
/*%Put what in estiamte: &estimate;*/


/*** Calculate ICC from each bootstrapped sample ***/
%Let by=_sample_;
%Let byobs=_sample_ _i;
```

```
%analyze(data=BOOTDATA,response=&response, out=BOOTDIST);

PROC SORT data=BOOTDIST; By ICC; RUN;

DATA BOOTDIST;
    SET BOOTDIST;
    e_b=ICC-&estimate;

    If ICC > &estimate Then indicator=1;
    Else indicator=0;
RUN;

/*** Calculate Lower and Upper limit of confidence interval ***/
/* 1-The first percentile method (Efron)  p41 eq3.3 */
/* 2-The Second percentile method (Hall)  p42 eq3.4 */

PROC UNIVARIATE DATA=BOOTDIST NOPRINT;
     VAR ICC e_b;
     OUTPUT OUT=CI1_2
            N=n pctlpre=P_ICC P_E_b pctlpts=2.5,97.5, 5;
RUN;

DATA CI1_2;
    SET CI1_2;
    llmt2=&estimate-P_E_b97_5;
    ulmt2=&estimate-P_E_b2_5;
    Lrho2=&estimate-P_E_b5;

    rename P_ICC2_5=llmt1 P_ICC97_5=ulmt1  P_ICC5=Lrho1 ;

    Keep P_ICC2_5 P_ICC97_5 llmt2 ulmt2 P_ICC5 Lrho2;
RUN;

/* 3-Biascorrected percentile CI (Efron, 1981a) p47 eq3.7 */
PROC UNIVARIATE DATA=BOOTDIST NOPRINT;
    Var indicator;
    Output out=z_0 mean=proportion;
RUN;

%Global Fhi_l Fhi_u rho_l;
DATA z_0;
    SET z_0;
    pro=1-proportion;
    z_0=probit(pro);
    z_alpha_half=probit(1-&alpha/2);
```

```
    z_alpha=probit(1-&alpha);

    Fhi_l=PROBNORM(2*z_0-z_alpha_half)*100;
    Fhi_u=PROBNORM(2*z_0+z_alpha_half)*100;

    rho_l=PROBNORM(2*z_0-z_alpha)*100;

    IF Fhi_l < 1.0E-6 Then Fhi_l=0.00001;
    IF Fhi_u < 1.0E-6 Then Fhi_u=0.00001;

    IF rho_l < 1.0E-6 Then rho_l=0.00001;

    CALL symput('Fhi_l', Fhi_l);
    CALL symput('Fhi_u', Fhi_u);
    CALL symput('rho_l', rho_l);
RUN;

PROC UNIVARIATE DATA=BOOTDIST NOPRINT;
     VAR ICC;
     OUTPUT OUT=CI3
            pctlpre=P_Cor pctlpts=&Fhi_l,&Fhi_u,&rho_l;
RUN;

DATA CI3_name;
    Length partL1 $2.0; Length partL2 $2.0;Length partU1 $3.0; Length partU2 $2.0;
    Length SpartU1 $3.0; Length SpartU2 $2.0;

    partL1=scan("&Fhi_l",1); partL2=substr(scan("&Fhi_l",2),1,2);
    lname=trim(left("P_Cor"||trim(left(partL1))||"_"||trim(left(partL2))));

    partU1=trim(left(scan("&Fhi_u",1))); partU2=substr(scan("&Fhi_u",2),1,2);

    leng=length(partU1);
    If length(partU1) LE 2 Then
    Uname=trim(left("P_Cor"||trim(left(partU1))||"_"||trim(left(partU2))));

    Else Uname=trim(left("P_Cor"||partU1));

/*For one-side lower limit*/
    SpartU1=trim(left(scan("&rho_l",1))); SpartU2=substr(scan("&rho_l",2),1,2);

    leng=length(SpartU1);
    If length(SpartU1) LE 2 Then
    Uname_1s=trim(left("P_Cor"||trim(left(SpartU1))||"_"||trim(left(SpartU2))));

    Else Uname_1s=trim(left("P_Cor"||SpartU1));
```

```
    CALL symput('llmt3', lname);
    CALL symput('ulmt3', uname);
    CALL symput('Lrho_3', uname_1s);
RUN;

DATA CI3;
    SET CI3;
    rename &llmt3=llmt3;
    rename &ulmt3=ulmt3;
    rename &Lrho_3=Lrho3;
RUN;

DATA CI_&response;
    SET CI_theo;
    SET CI1_2;
    SET CI3;

    Format ICC L_limit U_limit length llmt1 ulmt1 length1 llmt2 ulmt2 length2
           llmt3 ulmt3 length3 Lrho Lrho1 Lrho2 Lrho3 lth lth1 lth2 lth3 6.4;

    ICC=round(ICC,0.0001);
    L_limit=round(L_limit,0.0001);
    U_limit=round(U_limit,0.0001);
    length=U_limit-L_limit;

    llmt1=round(llmt1,0.0001);
    ulmt1=round(ulmt1,0.0001);
    length1=ulmt1-llmt1;

    llmt2=round(llmt2,0.0001);
    ulmt2=round(ulmt2,0.0001);
    length2=ulmt2-llmt2;

    llmt3=round(llmt3,0.0001);
    ulmt3=round(ulmt3,0.0001);
    length3=ulmt3-llmt3;

    Lrho=round(Lrho,0.0001);
    lth=1-Lrho;
    Lrho1=round(Lrho1,0.0001);
    lth1=1-Lrho1;
    Lrho2=round(Lrho2,0.0001);
    lth2=1-Lrho2;
    Lrho3=round(Lrho3,0.0001);
    lth3=1-Lrho3;
```

```
RUN;

Proc APPEND BASE=CI Data=CI_&response FORCE APPENDVER=V6; RUN;

options notes;
%Mend boot;



DM editor 'clear log' editor;
DM editor 'clear output' editor;
/*****************************************************************************************/
/* NOTE:To run the program, location need to be changed to the path where the original    */
/*      dataset, the macro BOOT, and the results will be stored.                          */
/*****************************************************************************************/

%Let Location=C:\Documents and Settings\wang\My Documents\paper\Bootstrap\;
*%include "&location.BootstrapAllFull.sas";

/* Preparation of data */
DATA ONE;
    INFILE "&location.posterio.dat" PAD MISSOVER;
    INPUT FILM  A1 A2 B1 B2 C1 C2;
        IF   1<=_N_<= 30 THEN MEASURE='T1T12VER';
    ELSE IF  31<=_N_<= 60 THEN MEASURE='T2T11VER';
    ELSE IF  61<=_N_<= 90 THEN MEASURE='RRAVT1T2';
    ELSE IF  91<=_N_<=120 THEN MEASURE='RRAVT2T3';
    ELSE IF 121<=_N_<=150 THEN MEASURE='RRAVT3T4';
    ELSE IF 151<=_N_<=180 THEN MEASURE='RRAVT4T5';
    ELSE IF 181<=_N_<=210 THEN MEASURE='RRAVT5T6';
    ELSE IF 211<=_N_<=240 THEN MEASURE='RRAVT6T7';
    ELSE IF 241<=_N_<=270 THEN MEASURE='RRAVT7T8';
    ELSE IF 271<=_N_<=300 THEN MEASURE='RRAVT8T9';
    ELSE IF 301<=_N_<=330 THEN MEASURE='RRAT9T10';
    ELSE IF 331<=_N_<=360 THEN MEASURE='RRT10T11';
    ELSE IF 361<=_N_<=390 THEN MEASURE='RRT11T12';
    ELSE IF 391<=_N_<=420 THEN MEASURE='ARAT1T12';
    ELSE IF 421<=_N_<=450 THEN MEASURE='ARAT2T11';
    ELSE IF 451<=_N_<=480 THEN MEASURE='ARAT3T10';
RUN;

PROC SORT DATA=ONE;
    BY FILM;
RUN;

%boot(data=one, response=T1T12VER, samples=10000, random=123);
```

```
%boot(data=one, response=T2T11VER, samples=10000, random=123);
%boot(data=one, response=RRAVT1T2, samples=10000, random=123);
%boot(data=one, response=RRAVT2T3, samples=10000, random=123);
%boot(data=one, response=RRAVT3T4, samples=10000, random=123);
%boot(data=one, response=RRAVT4T5, samples=10000, random=123);
%boot(data=one, response=RRAVT5T6, samples=10000, random=123);
%boot(data=one, response=RRAVT6T7, samples=10000, random=123);
%boot(data=one, response=RRAVT7T8, samples=10000, random=123);
%boot(data=one, response=RRAVT8T9, samples=10000, random=123);
%boot(data=one, response=RRAT9T10, samples=10000, random=123);
%boot(data=one, response=RRT10T11, samples=10000, random=123);
%boot(data=one, response=RRT11T12, samples=10000, random=123);
%boot(data=one, response=ARAT1T12, samples=10000, random=123);
%boot(data=one, response=ARAT2T11, samples=10000, random=123);
%boot(data=one, response=ARAT3T10, samples=10000, random=123);


ODS  HTMLCSS FILE="&location.Case30_2sFull.xls";
PROC PRINT DATA=CI;
    Var _NAME_ ICC L_limit U_limit length llmt1 ulmt1 length1 llmt2 ulmt2 length2
        llmt3   ulmt3 length3;
RUN;
ODS HTMLCSS CLOSE;

ODS  HTMLCSS FILE="&location.Case30_1sFull.xls";
PROC PRINT DATA=CI;
    Var _NAME_ ICC Lrho lth Lrho1 lth1 Lrho2 lth2 Lrho3 lth3;
RUN; ODS HTMLCSS CLOSE;
```

APPENDIX B

SAS Code Applying MCMC

```
/*************************************************************************************/
/*   Gibbs Sampling                                                                  */
/* This macro is designed to estimate ICC and CI for full 3-way random model with up to  */
/* 6 raters(coded as A,B,C, ---, F) and 4 occasions (coded as 1,2,3)                 */
/* Argument discription: n_p, n_r, n_o are numbers of patient(film), raters and occasions*/
/*                      alpha, beta and eita are prior parameters;                   */
/*                      nrep is the chain length and burnin is burnin period, burnin<nrep*/
/*                      measurement is the response variable;                        */
/*                      seed is for generating random number;                        */
/*                      startValue is a binary number, if =0 then the starting values of */
/*                        all random effects are set to be 1; if=1, deviation components */
/*                        that comprise the Sum of Squares will be used.             */
/* Usage: change Macro variable "location" to where original data set is stored.  The    */
/*        analysis results will also be stored there.                                */
/*************************************************************************************/
DM Editor 'clear log' editor; DM Editor 'clear output' editor;

options ls=125 ps=60 nodate nonumber;
%Let location=C:\Documents and Settings\wang\My Documents\Dissertation\MCMC\revise\;
LIBNAME sto "&location";

*options nomprint;
%Macro gibbs(n_p=, n_r=, n_o=, alpha=, beta=, eita=, nrep=, burnin=, measurement=, seed=,
             startValue=);

/**********************************************/
/* Read the original data                     */
/**********************************************/
DATA ONE;
INFILE "&location.posterio.dat" PAD MISSOVER;
INPUT FILM  A1 A2 B1 B2 C1 C2;
```

```
      IF   1<=_N_<= 30 THEN MEASURE='T1T12VER';
 ELSE IF  31<=_N_<= 60 THEN MEASURE='T2T11VER';
 ELSE IF  61<=_N_<= 90 THEN MEASURE='RRAVT1T2';
 ELSE IF  91<=_N_<=120 THEN MEASURE='RRAVT2T3';
 ELSE IF 121<=_N_<=150 THEN MEASURE='RRAVT3T4';
 ELSE IF 151<=_N_<=180 THEN MEASURE='RRAVT4T5';
 ELSE IF 181<=_N_<=210 THEN MEASURE='RRAVT5T6';
 ELSE IF 211<=_N_<=240 THEN MEASURE='RRAVT6T7';
 ELSE IF 241<=_N_<=270 THEN MEASURE='RRAVT7T8';
 ELSE IF 271<=_N_<=300 THEN MEASURE='RRAVT8T9';
 ELSE IF 301<=_N_<=330 THEN MEASURE='RRAT9T10';
 ELSE IF 331<=_N_<=360 THEN MEASURE='RRT10T11';
 ELSE IF 361<=_N_<=390 THEN MEASURE='RRT11T12';
 ELSE IF 391<=_N_<=420 THEN MEASURE='ARAT1T12';
 ELSE IF 421<=_N_<=450 THEN MEASURE='ARAT2T11';
 ELSE IF 451<=_N_<=480 THEN MEASURE='ARAT3T10';
RUN;

PROC SORT DATA=ONE; BY FILM; RUN;

Data ONE ; Set ONE; where Measure="&measurement"; Run;

PROC TRANSPOSE DATA=ONE OUT=ONE1_2; ID MEASURE; BY FILM; RUN;

DATA TWO (DROP=_NAME_ &measurement film);SET ONE1_2;
  patient=film;

  IF      UPCASE(SUBSTR(_NAME_,1,1))="A" then rater=1;
  Else if UPCASE(SUBSTR(_NAME_,1,1))="B" then rater=2;
  Else if UPCASE(SUBSTR(_NAME_,1,1))="C" then rater=3;
  Else if UPCASE(SUBSTR(_NAME_,1,1))="D" then rater=4;
  Else if UPCASE(SUBSTR(_NAME_,1,1))="E" then rater=5;
  Else if UPCASE(SUBSTR(_NAME_,1,1))="F" then rater=6;

  IF      SUBSTR(_NAME_,2,1)="1" then ocass=1;
  Else if SUBSTR(_NAME_,2,1)="2" then ocass=2;
  Else if SUBSTR(_NAME_,2,1)="3" then ocass=3;
  Else if SUBSTR(_NAME_,2,1)="4" then ocass=4;

  x=&measurement;
  resid=.; resid_wp=.; resid_wr=.; resid_wo=.;resid_wpr=.; resid_wpo=.;
  resid_wro=.; resid_wmu=.;
RUN;

/*********************************************************************************/
/* Evaluate the initial values of the random effects of mu, p, r, o, pr, po and ro */
```

```
/*******************************************************************************/
proc sql;
    create table mu as
    select mean(x) as mu
    from Two;
/*random effects of P*/
    create table p as
    select unique patient, mean(x)-mu as p
    from two, mu
    group by patient;
/*random effects of R*/
    create table r as
    select unique rater, mean(x)-mu as r
    from two, mu
    group by rater;
/*random effects of O*/
    create table o as
    select unique ocass, mean(x)-mu as o
    from two, mu
    group by ocass;
/*random effects of PR -- interaction between Patient and Rater*/
    create table pr as
    select unique two.patient, two.rater, mean(x)-p-r-mu as pr
    from two, mu, p, r
    where two.patient=p.patient and two.rater=r.rater
    group by two.patient, two.rater;
/*random effects of PO -- interaction between Patient and Occasion*/
    create table po as
    select unique two.patient, two.ocass, mean(x)-p-o-mu as po
    from two, mu, p, o
    where two.patient=p.patient and two.ocass=o.ocass
    group by two.patient, two.ocass;
/*random effects of RO -- interaction between Rater and Occasion*/
    create table ro as
    select unique two.rater, two.ocass, mean(x)-r-o-mu as ro
    from two, mu, r, o
    where two.rater=r.rater and two.ocass=o.ocass
    group by two.rater, two.ocass;
quit;

Options nonotes;

PROC IML;
/*************   Starting values of random terms        **********************************/
/* If startValue is 0, then the starting values of all random effects are set to be 1;    */
/* Else, they are set to be deviation components that comprise the their Sum of Squares.  */
```

```
/*******************************************************************************************/
 %If %bquote(&startValue)=  OR &startValue=0 %Then %Do;
    mu=0;
    pa=repeat(1,&n_p,1);    ra=repeat(1,&n_r,1);    oc=repeat(1,&n_o,1);
    pr=repeat(1,&n_p,&n_r); po=repeat(1,&n_p,&n_o); ro=repeat(1,&n_r,&n_o);
 %End;
 %Else %Do;
    Use mu; Read all Var{mu} into mu; Close mu; Use P ; Read all Var{p} into PA;  Close P;
    Use R ; Read all Var{r} into RA;  Close R;  Use O ; Read all Var{o} into OC;  Close O;
    Use PR ; Read all Var{pr} into pr; pr=shape(pr,&n_p,&n_r); Close PR;
    Use PO ; Read all Var{po} into po; po=shape(po,&n_p,&n_o); Close PO;
    Use RO ; Read all Var{ro} into ro; ro=shape(ro,&n_r,&n_o); Close RO;
 %End;
    tau_p=1; tau_r=1; tau_o=1; tau=1; tau_pr=1; tau_po=1; tau_ro=1;/**/
    n_p=&n_p;   n_r=&n_r;  n_o=&n_o;

/*********************************  start Gibbs sampling  ******************************/
Do rep=1  to &nrep;

    var_p =1/tau_p;     var_r =1/tau_r;     var_o =1/tau_o;     var   =1/tau;
    var_pr=1/tau_pr;    var_po=1/tau_po;    var_ro=1/tau_ro;
    rho=var_p/(var_p + var_r + var_o + var_pr + var_po + var_ro + var);

    Use TWO; Read all into Rdata; Close two;   /* Read in original data from data set TWO */
    mattrib Rdata colname=
         ({'subj' 'rate' 'occa' 'x' 'x_mu' 'x_p' 'x_r' 'x_o' 'x_pr' 'x_po' 'x_ro' 'resid'});

/***************************** 1. Update Variance Components ***************************/
/****** tau ---  alpha=&alpha+n_p*n_r*n_o/2;  beta =1/(1/&beta+ssq(Rdata[,'resid'])/2);  */
    Do i=1  to n_p;
         Do j=1  to n_r;
            Do k=1 to n_o;
                Rdata[loc(Rdata[ ,'subj']=i & Rdata[ ,'rate']=j & Rdata[ ,'occa']=k),'resid']
                =Rdata[loc(Rdata[ ,'subj']=i & Rdata[ ,'rate']=j & Rdata[ ,'occa']=k) ,'x']
                    -mu-pa[i]-ra[j]-oc[k]-pr[i,j]-po[i,k]-ro[j,k];
            End;
         End;
    End;
    tau=(1/(1/&beta+ssq(Rdata[,'resid'])/2))*RANGAM(&SEED,&alpha+n_p*n_r*n_o/2);

/****** tau_p --- alpha=&alpha+n_p/2;  beta =1/(1/&beta +ssq(pa)/2);   */
    tau_p=(1/(1/&beta +ssq(pa)/2))*RANGAM(&SEED,&alpha+n_p/2);

/****** tau_r --- alpha=&alpha+n_r/2;  beta =1/(1/&beta +ssq(ra)/2);   */
    tau_r=(1/(1/&beta +ssq(ra)/2))*RANGAM(&SEED,&alpha+n_r/2);
```

```
/****** tau_o --- alpha=&alpha+n_o/2;  beta =1/(1/&beta +ssq(oc)/2);   */
    tau_o=(1/(1/&beta +ssq(oc)/2))*RANGAM(&SEED,&alpha+n_o/2);

/****** tau_pr -- alpha=&alpha+n_p*n_r/2; beta =1/(1/&beta +ssq(pr)/2);*/
    tau_pr=(1/(1/&beta +ssq(pr)/2))*RANGAM(&SEED,&alpha+n_p*n_r/2);

/****** tau_po -- alpha=&alpha+n_p*n_o/2; beta =1/(1/&beta +ssq(po)/2);*/
    tau_po=(1/(1/&beta +ssq(po)/2))*RANGAM(&SEED,&alpha+n_p*n_o/2);

/****** tau_ro -- alpha=&alpha+n_r*n_o/2; beta =1/(1/&beta +ssq(ro)/2);*/
    tau_ro=(1/(1/&beta +ssq(ro)/2))*RANGAM(&SEED,&alpha+n_r*n_o/2);

/************************* 2. Update Random Effects  ***********************************/
/****** Pa_i ---  Mean_p= tau*sum(Rdata[loc(Rdata[ ,'subj']=i), 'x_p'])/(n_r*n_o*tau+tau_p);
                  v_p   = 1/(n_r*n_o*tau+tau_p);    ***/
    Do i=1 to n_p;
        Do j=1 to n_r;
            Do k=1 to n_o;
                Rdata[loc(Rdata[ ,'subj']=i & Rdata[ ,'rate']=j & Rdata[ ,'occa']=k) ,'x_p']
                =Rdata[loc(Rdata[ ,'subj']=i & Rdata[ ,'rate']=j & Rdata[ ,'occa']=k) ,'x']
                    -mu     -ra[j]-oc[k]-pr[i,j]-po[i,k]-ro[j,k];
            End;
        End;
        pa[i]=tau*sum(Rdata[loc(Rdata[ ,'subj']=i), 'x_p'])/(n_r*n_o*tau+tau_p)
                +sqrt(1/(n_r*n_o*tau+tau_p))*NORMAL(&seed);
    End;

/****** Ra_j ---  Mean_r= tau*sum(Rdata[loc(Rdata[ ,'rate']=j), 'x_r'])/(n_p*n_o*tau+tau_r);
                  v_r   = 1/(n_p*n_o*tau+tau_r);   ***/
    Do j=1 to n_r;
        Do i=1 to n_p;
            Do k=1 to n_o;
                Rdata[loc(Rdata[ ,'subj']=i & Rdata[ ,'rate']=j & Rdata[ ,'occa']=k) ,'x_r']
                =Rdata[loc(Rdata[ ,'subj']=i & Rdata[ ,'rate']=j & Rdata[ ,'occa']=k) ,'x']
                    -mu-pa[i]     -oc[k]-pr[i,j]-po[i,k]-ro[j,k];
            End;
        End;
        ra[j]=tau*sum(Rdata[loc(Rdata[ ,'rate']=j), 'x_r'])/(n_p*n_o*tau+tau_r)
                +sqrt(1/(n_p*n_o*tau+tau_r))*NORMAL(&seed);
    End;

/****** Oc_k --- Mean_o= tau*sum(Rdata[loc(Rdata[ ,'occa']=k), 'x_o'])/(n_p*n_r*tau+tau_o);
                 v_o   = 1/(n_p*n_r*tau+tau_o);  ***/
    Do k=1 to n_o;
        Do i=1 to n_p;
            Do j=1 to n_r;
```

```
                Rdata[loc(Rdata[ ,'subj']=i & Rdata[ ,'rate']=j & Rdata[ ,'occa']=k) ,'x_o']
                =Rdata[loc(Rdata[ ,'subj']=i & Rdata[ ,'rate']=j & Rdata[ ,'occa']=k) ,'x']
                    -mu-pa[i]-ra[j]     -pr[i,j]-po[i,k]-ro[j,k];
            End;
        End;
        oc[k]=tau*sum(Rdata[loc(Rdata[ ,'occa']=k), 'x_o'])/(n_p*n_r*tau+tau_o)
                +sqrt(1/(n_p*n_r*tau+tau_o))*NORMAL(&seed);
    End;

/****** pr_ij ******/
    Do i=1 to n_p;
        Do j=1 to n_r;
            Do k=1 to n_o;
                Rdata[loc(Rdata[ ,'subj']=i & Rdata[ ,'rate']=j & Rdata[ ,'occa']=k) ,'x_pr']
                =Rdata[loc(Rdata[ ,'subj']=i & Rdata[ ,'rate']=j & Rdata[ ,'occa']=k) ,'x']
                    -mu-pa[i]-ra[j]-oc[k]       -po[i,k]-ro[j,k];
            End;
            pr[i,j]=tau*sum(Rdata[loc(Rdata[ ,'subj']=i & Rdata[ ,'rate']=j),'x_pr'])
                    /(n_o*tau+tau_pr) +sqrt(1/(n_o*tau+tau_pr))*NORMAL(&seed);
        End;
    End;

/****** po_ik ********/
    Do i=1 to n_p;
        Do k=1 to n_o;
            Do j=1 to n_r;
                Rdata[loc(Rdata[ ,'subj']=i & Rdata[ ,'rate']=j & Rdata[ ,'occa']=k) ,'x_po']
                =Rdata[loc(Rdata[ ,'subj']=i & Rdata[ ,'rate']=j & Rdata[ ,'occa']=k) ,'x']
                    -mu-pa[i]-ra[j]-oc[k]-pr[i,j]      -ro[j,k];
            End;
            po[i,k]=tau*sum(Rdata[loc(Rdata[ ,'subj']=i & Rdata[ ,'occa']=k),'x_po'])
                     /(n_r*tau+tau_po)+sqrt(1/(n_r*tau+tau_po))*NORMAL(&seed);
        End;
    End;

/****** ro_jk **********/
    Do j=1 to n_r;
        Do k=1 to n_o;
            Do i=1 to n_p;
                Rdata[loc(Rdata[ ,'subj']=i & Rdata[ ,'rate']=j & Rdata[ ,'occa']=k) ,'x_ro']
                =Rdata[loc(Rdata[ ,'subj']=i & Rdata[ ,'rate']=j & Rdata[ ,'occa']=k) ,'x']
                    -mu-pa[i]-ra[j]-oc[k]-pr[i,j]-po[i,k]     ;
            End;
            ro[j,k]=tau*sum(Rdata[loc(Rdata[ ,'rate']=j & Rdata[ ,'occa']=k),'x_ro'])
                     /(n_p*tau+tau_ro)+sqrt(1/(n_p*tau+tau_ro))*NORMAL(&seed);
        End;
```

```
    End;

/****** mu   --- mean_mu=tau*sum(Rdata[ ,'x_mu'])/(1/&eita+tau*n_p*n_r*n_o);
                 v_mu   =1/(1/&eita+tau*n_p*n_r*n_o);***/
    Do i=1  to n_p;
         Do j=1  to n_r;
            Do k=1 to n_o;
                Rdata[loc(Rdata[ ,'subj']=i & Rdata[ ,'rate']=j & Rdata[ ,'occa']=k) ,'x_mu']
                =Rdata[loc(Rdata[ ,'subj']=i & Rdata[ ,'rate']=j & Rdata[ ,'occa']=k) ,'x']
                       -pa[i]-ra[j]-oc[k]-pr[i,j]-po[i,k]-ro[j,k];
            End;
         End;
    End;
    mu=tau*sum(Rdata[ ,'x_mu'])/(1/&eita+tau*n_p*n_r*n_o)
        +sqrt(1/(1/&eita+tau*n_p*n_r*n_o))*NORMAL(&seed);
/********************** End of one iteration of updating parameters  *******************/

    mattrib outMatrix colname={'var_p' 'var_r' 'var_o' 'var_pr' 'var_po'
                              'var_ro' 'var' 'rho'};
    outMatrix= outMatrix // (var_p || var_r || var_o || var_pr || var_po ||
                              var_ro || var || rho);

    *print mu PA RA OC PR PO RO;
    *print tau tau_p tau_r tau_o tau_pr tau_po tau_ro;
End;

create result from outMatrix [colname={'var_p' 'var_r' 'var_o' 'var_pr' 'var_po' 'var_ro'
                                      'var' 'rho'}];
   append from outMatrix;
close result;

quit; /* End of IML proc */

/* Export the results of Markov Chain for further use*/
Data sto.&measurement._I&nrep._sd&seed;
    Set result;
Run;

Data trunResult;
    Set result;
    If _N_ > &burnin and  _N_ <= &nrep;
Run;

Proc univariate data=trunResult noprint;
    Var rho;
    Output out=estimate median=ICC_md mean=ICC pctlpre=CI pctlpts=2.5,97.5;;
```

```
Run;

Data estimate;
    Set estimate;
    measure=SUBSTR("&measurement", 1,8);
    n=&nrep;
    m=&burnin;
    ICC_estimate=round(ICC,0.01);
    lowerCI=round(CI2_5, 0.01);
    UpperCI=round(CI97_5,0.01);
    CI_length=round(CI97_5-CI2_5, 0.01);
    Drop ICC  CI2_5  CI97_5;
run;
/*Proc Print data=estimate; Run;*/

Proc APPEND BASE=estimate_all Data=estimate FORCE APPENDVER=V6; RUN;

/* Export the estimation results */
Data sto.CI_estimate;   Set estimate_all;  Run;
Proc Print data=estimate_all; Run;
Quit;Run;
%Mend;

%gibbs(n_p=30, n_r=3, n_o=2, alpha=0.01, beta=100, eita=100, nrep=10000, burnin=6000,
       measurement=RRT11T12, seed=12300530,startValue=1);
%gibbs(n_p=30, n_r=3, n_o=2, alpha=0.01, beta=100, eita=100, nrep=10000, burnin=6000,
       measurement=RRT11T12, seed=123,startValue=1);
```