Spring 4-20-2011

Modified Profile Likelihood Approach for Certain Intraclass Correlation Coefficient

Huayu Liu

MODIFIED PROFILE LIKELIHOOD APPROACH FOR CERTAIN INTRACLASS CORRELATION COEFFICIENTS

by

HUAYU LIU

Under the Direction of Dr. YUANHUI XIAO

ABSTRACT

In this paper we consider the problem of constructing confidence intervals and lower bounds for the intraclass correlation coefficient in an interrater reliability study where the raters are randomly selected from a population of raters.The likelihood function of the interrater reliability is derived and simplified, and the profile likelihood based approach is readily available for computing the confidence intervals of the interrater reliability. Unfortunately, the confidence intervals computed by using the profile likelihood function are in general too narrow to have the desired coverage probabilities. From the point view of practice, a conservative approach, if is at least as precise as any existing method, is preferred since it gives the correct results with a probability higher than claimed. Under this rationale, we propose the so-called modified profile likelihoodapproach in this paper. Simulation study shows that, the proposed method in general has better performance than currently used methods.

INDEX WORDS: Interrater reliability, Profile likelihood, Confidence intervals, Generalized P-values, Modified large-sample approach

MODIFIED PROFILE LIKELIHOOD APPROACH FOR CERTAIN INTRACLASS CORRELATION COEFFICIENTS

by

HUAYU LIU

A Dissertation Submitted in Partial Fulfillment of the Requirements for the Degree of

Master of Science

in the College of Arts and Sciences

Georgia State University

2011

MODIFIED PROFILE LIKELIHOOD APPROACH FOR CERTAIN INTRACLASS CORRELATION COEFFICIENTSK

by

HUAYU LIU

Committee Chair:	Dr. Yuanhui Xiao
Committee:	Dr. Yichuan Zhao
	Dr. Xu Zhang

Electronic Version Approved:

Office of Graduate Studies

College of Arts and Sciences

Georgia State University

May 2011

ACKNOWLEDGEMENTS

Foremost, I would like to give most sincere thanks to my advisor Dr. Yuanhui Xiao for his constant support, abundant help and invaluable advice for my research and study in Georgia State University. With patience and motivation, he has given me timely guidance in the critical stages of my research and especially towards the completion of this thesis.

Secondly, I am thankful for the fact that Dr. Zhao and Dr. Zhang would like to be my thesis committee members. I appreciate their time to read my thesis and give me salient advice.

Last but most importantly, I would like to thank my husband Shenjia Zhang for his caring love and selfless support, which makes my life, study and research in Georgia State university an enjoyable experience.

TABLE OF CONTENTS

LIST OF TABLES

Chapter 1

INTRODUCTION

The intraclass correlation coefficient,which defined by Harggard [?] as "the measure of relative homogeneity of the score within classes in relation to the total variance", has been widely adapted in behavior and biomedical science. Its first application dates back to Pearson in his study of measuring family resemblance of height of brothers. There are basically two major objects with which Intraclass correlation coefficient is used: to measure the sameness for unit in the same group, or reliability test, where measurement error is introduced by rates.

In this paper we consider the interrater reliability study in which each of R raters measures each of S subjects. Suppose each of the random sample of R raters in the reliability study rates each of the corresponding random sample of S subjects independently, the rating score Y_{ij} of the ith rater on the jth subject may be represented as

$$Y_{ij} = \mu + r_i + s_j + e_{ij}, \quad (i = 1, 2, \ldots, R;\ j = 1, 2, \ldots, S), \tag{1.1}$$

where μ is the overall population mean of the measurements, r_i reflects the effect of the ith rater, s_j characterizes the effect of the jth subject and e_{ij} is the measurement error associated with this rating. The random variables r_i, s_j and e_{ij} are assumed to be mutually independent and normally distributed with zero mean 0 and variances σ_r^2, σ_s^2 and σ_e^2, respectively. The variance σ^2 of Y_{ij} is

$$\sigma^2 \equiv \text{var}(Y_{ij}) = \sigma_r^2 + \sigma_s^2 + \sigma_e^2 \tag{1.2}$$

and the covariance between two measurements taken by the ith and i'th raters on the same subject j is

$$\mathrm{cov}(Y_{ij}, Y_{i'j}) = \sigma_s^2. \tag{1.3}$$

It follows that the appropriate intraclass correlation coefficient to measure interrater reliability is

$$\rho = \frac{\sigma_s^2}{\sigma^2} \equiv \frac{\sigma_s^2}{\sigma_s^2 + \sigma_r^2 + \sigma_e^2}. \tag{1.4}$$

The value of ρ represents the proportion of total variability on observed scores accounted for by the subject-to-subject variability in the true but unobservable scores. To emphasize this fact, it is also denoted by ρ_s. Throughout this paper, ρ and ρ_s are used interchangeably.

(Similarly, the quantity ρ_r given by the following equation

$$\rho_r = \frac{\sigma_r^2}{\sigma^2} \equiv \frac{\sigma_r^2}{\sigma_s^2 + \sigma_r^2 + \sigma_e^2} \tag{1.5}$$

is the proportion of total variability on observed scores accounted for by the rater-to-rater variability in the true score. The ratio of rater-to-error variability is defined as

$$\delta = \frac{\sigma_r^2}{\sigma_e^2} \equiv \frac{\rho_r}{1 - \rho_s - \rho_r}, \tag{1.6}$$

which is also an important parameter of the model (1.1). Obviously, $\rho_s > 0,\ \rho_r > 0,\ \rho_s + \rho_r < 1$.)

As the assessment of reliability of measurement is of great importance in medical study, where measurement error may have serious unwanted consequence, considerable amount of research has been done for ρ. See [7], [1], [14], [2] and [10], etc. Table 1.1 presents the analysis of variance layout corresponding to the two-way random effects model (1.1). From this table, the unbiased estimators of the three components of variances σ_s^2, σ_r^2, σ_e^2 are $s_s^2 = (\mathrm{SMS} - \mathrm{EMS})/R$, $s_r^2 = (\mathrm{RMS} - \mathrm{EMS})/S$ and s_e^2, respectively. The interrater reliability

Table 1.1. Analysis of variance for the results of an interrater reliability study

Source of variation	df	Sum of squares	MS	E(MS)
Subjects	$S-1$	$R\sum(\bar{y}_{\cdot j}-\bar{y}_{\cdot\cdot})^2$	SMS	$\theta_S=\sigma_e^2+R\sigma_s^2$
Raters	$R-1$	$S\sum(\bar{y}_{i\cdot}-\bar{y}_{\cdot\cdot})^2$	RMS	$\theta_R=\sigma_e^2+S\sigma_r^2$
Error	$(R-1)(S-1)$	By subtraction	EMS	$\theta_E=\sigma_e^2$
Total	$RS-1$	$\sum\sum(Y_{ij}-\bar{y}_{\cdot\cdot})^2$		

is then estimated by

$$\hat{\rho}=\frac{s_s^2}{s_s^2+s_r^2+s_e^2}=\frac{S\times(SMS-EMS)}{S\times SMS+R\times RMS+(RS-R-S)\times EMS} \tag{1.7}$$

This approach was proposed by Rajaratnam [7] and Rajaratnam[1]. The method is in general biased and the resulting estimates may be negative. Fortunately, the bias decreases as both R and S increase.

Fleiss [4] developed an approach for interval estimation of ρ based on Satterthwaite's[8] two-moment approximation. This method has been widely used, but it understates the coverage probabilities substantially in certain cases. The problem has been noticed by Zou *et al.*[14], so they proposed a three-moment approximation and a four-moment approximation by using the Pearson system of distribution under the rationale that a better approximation may be achieved by using higher moments. In general, their higher-moment approaches produce confidence intervals that are more conservative and satisfactory than those produced by two-moment approaches.

However, Cappelleri [2] found that, in certain situations, (for example, when there are only three raters, or the ratio of rater-to-error variability is relatively high.), the higher-moment approaches tend to understate the coverage probabilities. Therefore, they proposed a modified large-sample approach (MLS), which produces either correct or conservative confidence intervals with more precise (narrower) widths than those generated by the higher-moment approaches.

By using generalized variables (Section 2), Tian *et al.*[10] proposed an approach for estimating the two-sided confidence intervals or one-side confidence lower bounds of ρ. The generalized variable method (GV) in general has a behavior similar to that of the MLS.

In certain situations, (e.g., when the ratio of rater-to-error variability is about 0.5), the coverage probabilities of one-sided confidence intervals produced by the GV method (i.e., the method using generalized variables) are marginally more conservative than that produced by MLS. Based on simulation studies, the GV method has the best overall performance among existing methods.

It seems that no existing method for estimating the confidence intervals of ρ is based on likelihood approach. However, through a series of algebraic operations, we find found that the likelihood function of (1.1) has an explicit expression which is extremely simple (Appendix), so we will develop methods for constructing confidence intervals and computing the lower bounds for ρ based on profile likelihood approach. In general, the confidence intervals of ρ produced by the traditional profile-likelihood approach are too narrow to have the desired coverage probabilities, so a remedy is necessary, resulting in the modified profile likelihood (MPL) approach proposed in this paper. The detail is described in Section 3. The proposed approach is assessed by a Monte-Carlo simulation study, where its performance is compared with that of the GV approach in terms of coverage probabilities and average lengths of the respective confidence intervals. The GV method is chosen for comparison purpose not only because it is one of the most competent methods, but also easy to implement. The results of the simulation study are presented and analyzed in Section 4. The performance of the modified profile likelihood approach will be furtherly discussed in the last section. Before we introduce the proposed approach, we will give a short description of the generalized variable method.

Chapter 2

THE GENERAL VARIABLE APPROACH

Suppose that $\boldsymbol{X} = (X_1, X_2, \ldots, X_n)$ constitute a random sample from a distribution which depends on the parameters $\beta = (\theta, \boldsymbol{\nu})$, where θ is of interest and $\boldsymbol{\nu}$ is a vector of nuisance parameters. Let $\boldsymbol{x}$ is be an observed value of $\boldsymbol{X}$, a generalized variable $R(\boldsymbol{X}; \boldsymbol{x}, \theta, \boldsymbol{\nu})$ for interval estimation of θ has the following two properties:

1. $R(\boldsymbol{X}; \boldsymbol{x}, \theta, \boldsymbol{\nu})$ has a distribution free of unknown parameters;

2. $R(\boldsymbol{x}; \boldsymbol{x}, \theta, \boldsymbol{\nu}) = \theta$.

See Weerahandi [12]. To construct confidence intervals for the intraclass correlation coefficient ρ, [10] proposed the following generalized variable

$$
\begin{aligned}
R_\rho &= \frac{\mathrm{sms}\frac{\theta_S}{\mathrm{SMS}} - \mathrm{ems}\frac{\theta_E}{\mathrm{EMS}}}{\mathrm{sms}\frac{\theta_S}{\mathrm{SMS}} + (R/S)\mathrm{rms}\frac{\theta_R}{\mathrm{RMS}} + [R-1-(R/S)]\mathrm{ems}\frac{\theta_E}{\mathrm{EMS}}} \\
&= \frac{\mathrm{sms}\frac{S-1}{Q_S} - \mathrm{ems}\frac{(R-1)(S-1)}{Q_E}}{\mathrm{sms}\frac{S-1}{Q_S} + (R/S)\mathrm{rms}\frac{R-1}{Q_R} + [R-1-(R/S)]\mathrm{ems}\frac{(R-1)(S-1)}{Q_E}},
\end{aligned}
\tag{2.1}
$$

where $Q_S = (S-1)\mathrm{SMS}/\theta_S$, $Q_R = (R-1)\mathrm{RMS}/\theta_R$, and $Q_E = (R-1)(S-1)\mathrm{EMS}/\theta_E$. (Note that, here $\boldsymbol{X} = (\mathrm{SMS}, \mathrm{RMS}, \mathrm{EMS})$ and $\boldsymbol{x} = (\mathrm{sms}, \mathrm{rms}, \mathrm{ems})$.) The random variables Q_S, Q_R and Q_E are independent and distributed as χ^2_{S-1}, χ^2_{R-1}, and $\chi^2_{(R-1)(S-1)}$, respectively.

To construct a confidence interval for ρ, first we simulate a large number of triples (Q_S, Q_R, Q_E) of random numbers from the chi-square distributions with degrees of freedoms $S-1$, $S-1$ and $(R-1)(S-1)$, respectively. Then the same number of values of R_ρ are computed via (2.1). The sample quantiles (denoted by $R_{\rho,\gamma}$, where $0 < \gamma < 1$) of those values of R_ρ are used to construct confidence intervals for ρ: the two-sided $100(1-\alpha)\%$

confidence interval is computed as $(R_{\rho,\alpha/2},\ R_{\rho,1-\alpha/2})$ and the one-sided $100(1-\alpha)\%$ lower bound is estimated as $R_{\rho,\alpha}$.

Chapter 3

LIKELIHOOD BASED INTERVAL ESTIMATION FOR ρ

It follows from (28), (29), (30), (31) and (56) in the Appendix that the negative twice logarithm of the likelihood function l of ρ $(= \rho_s)$ and ρ_r, is given by

$$
\begin{aligned}
-2l = c' &+ \ln[1 + (R-1)\rho_s + (S-1)\rho_r] && (3.1)\\
&+ (S-1)\ln[1 + (R-1)\rho_s - \rho_r] && (3.2)\\
&+ (R-1)\ln[1 - \rho_s + (S-1)\rho_r] && (3.3)\\
&+ (R-1)(S-1)\ln[1 - \rho_s - \rho_r] && (3.4)\\
&+ RS\ln\left[\frac{(S-1)\mathrm{SMS}}{1-\rho_r+(R-1)\rho_s} + \frac{(R-1)\mathrm{RMS}}{1+(S-1)\rho_r-\rho_s} + \frac{(R-1)(S-1)\mathrm{EMS}}{1-\rho_s-\rho_r}\right], && (3.5)
\end{aligned}
$$

where c' is a constant free from the data and the parameters. Interestingly, the log-likelihood function is a "function" of the ANOVA layout in Table 1.1. The profile log likelihood function for ρ $(\equiv \rho_s)$ is

$$l^{\dagger}(\rho; \boldsymbol{y}) = \max\{l(\rho, \rho_r; \boldsymbol{y}) : \ 0 < \rho_r < 1 - \rho\}. \tag{3.6}$$

Suppose $l^{\dagger}(\rho; \boldsymbol{y})$ achieves its maximum at $\hat{\rho}$, then $\hat{\rho}$ is also the maximum likelihood estimate (MLE) of ρ. An approximate two-side $100(1-\alpha)\%$ $(0 < \alpha < 1)$ confidence interval for ρ is obtained as

$$\left\{\rho : \ 2l^{\dagger}(\hat{\rho}; \boldsymbol{y}) - 2l^{\dagger}(\rho; \boldsymbol{y}) \le (1+\kappa)\chi^2_{1,1-\alpha}\right\}, \tag{3.7}$$

and an approximate one-side $100(1-\alpha)\%$ lower bound is computed as the smaller root of the following equation for ρ,

$$2l^{\dagger}(\hat{\rho}; \boldsymbol{y}) - 2l^{\dagger}(\rho; \boldsymbol{y}) = (1+\kappa)\chi^2_{1,1-\alpha}, \tag{3.8}$$

where κ is an appropriately chosen constant in (3.7) and (3.8), respectively. The confidence bounds in (3.7) and (3.8) are actually computed by a one-dimensional root-finding routine nesting a one-dimensional optimization method.

According to the traditional likelihood theory, the constant κ is set to be zero, but in our case the coverage probabilities will be in general understated, as shown in the simulation study in Section 4. Ideally, we wish to choose the value of κ that would result in correct confidence intervals. This value of κ is termed the "correct value" of κ and denoted by κ_{corr}. Accordingly to the likelihood theory, the correct value κ_{corr} approaches to zero as both the numbers of raters and subjects increase.

The correct value κ_{corr} depends on ρ_s, ρ_r, R (the number of raters) and S (the number of subjects). a trackable expression is hardly available for it, thus it is infeasible to use κ_{corr} in practice. If we use a value of κ that is larger than κ_{corr}, then the resulting confidence intervals will be conservative in the sense that the coverage probability is higher than $1 - \alpha$. The cost of achieving higher coverage probabilities is the loss of accuracy since the confidence intervals would become wider. Suppose $(\rho_L,\ \rho_U)$ and $(\delta_L,\ \delta_U)$ are known ranges of ρ and δ, respectively, then for fixed values of R and S, letting κ equal to the value of

$$\kappa_m = \max\{\kappa_{corr} \equiv \kappa_{corr}(\rho, \delta) :\ \rho_L \le \rho \le \rho_U;\ \delta_L \le \delta \le \delta_U\}, \tag{3.9}$$

would minimize the loss of accuracy. If we use κ_m or an estimate of κ_m in (3.7) or (3.8), the resulting method will be called as modified profile likelihood (MPL) approach. Suppose we know the modified profile likelihood would produce confidence intervals which are in average shorter than those by a widely used method, then we may put this approach in our tool-box for practical use since it is not only more precise than the widely used approach, but also captures the true parameter value with a probability higher than claimed. That the modified profile likelihood approach is more accurate than existing methods, which is illustrated by our simulation study in Section 4.

In real life, we can hardly expect a value of ρ that is lower than 0.6, so we can set the lower bound of ρ to be $\rho_L = 0.6$. As for the upper bound ρ_U of ρ, we can safely set $\rho_U = 0.98$ in most cases. Several authors used the values $0.5,\ 1.0,\ 4.0$ of δ in their simulation study. Conservatively, in most cases we may set $\delta_L = 0.5$ and $\delta_U = 16$.

Simulation study (not presented, but the reader is referred to Table 4.2 and Table 4.5) shows that, roughly κ_{corr} is an increasing function of δ for fixed value of ρ but a decreasing function of ρ for fixed value of δ. Thus, an estimate of κ_m can be obtained by using a grid search as follows. First we choose two positive values d and r, and define the grid points (ρ_i, δ_j) as

$$\begin{aligned} &\rho_i = \rho_L + (i-1)d,\ i = 1, 2, \ldots, n_1\ (\le \frac{\rho_U - \rho_L}{d} + 1), \\ &\delta_j = r^j,\ k_1 \le j \le k_2\ (\text{ so } \delta_L = r^{k_1},\ \text{ and } \delta_U = r^{k_2}). \end{aligned} \tag{3.10}$$

Then a reasonable estimate for κ_m is

$$\hat{\kappa}_m = \max\{\kappa_{corr}(\rho_i, \delta_j) :\ 1 \le i \le n_1;\ k_1 \le j \le k_2\}. \tag{3.11}$$

Usually we use $d = 0.1$ or $d = 0.05$ and $r = 2,\ k_1 = -1,\ k_2 = 4$ for computing $\hat{\kappa}_m$.

For known parameter values of ρ and δ, the optimal value κ_{corr} can be obtained through a grid search incorporated with Monte-Carlo simulation. Suppose we want to search for κ_{corr} in the interval $[\kappa_L, \kappa_U]$ of κ, we can simulate the coverage probability at the grid points $\kappa_i = \kappa_L + (i-1) * d_\kappa,\ i = 1, 2, \ldots; i \le (\kappa_U - \kappa_L)/d_\kappa$, where $d_\kappa > 0$ is the increment, and select the smallest κ_i for which the simulated coverage probability is not less than the expected coverage probability. Here are the steps:

1. Choose κ_L, κ_U and d_κ, and compute the value for each κ_i.

2. Generate a sample according to (1.1) by using given parameter values.

3. Compute the upper confidence bound $\hat{\rho}_{U,0}$ of (3.7) and/or the lower confidence bound $\hat{\rho}_{L,0}$ of (3.7) or (3.8) for a_1.

4. For $i = 2, \ldots,$ search for the upper confidence bound $\hat{\rho}_{U,i}$ of (3.7) in the interval $[\hat{\rho}_{U,i-1}, 1]$ for κ_i, and/or the lower confidence bound $\hat{\rho}_{L,i}$ of (3.7) or (3.8) in the interval $[0, \hat{\rho}_{L,i-1}]$ for κ_i.

5. For each κ_i, check if the corresponding confidence interval includes the true value of ρ.

6. Repeat the steps 2 - 5 m times, where m is a large number (say 10, 000 or 20,000). For each κ_i, record the number of confidence intervals that include the true value of ρ.

7. Finally, compute the proportion of confidence intervals that conclude the true value of ρ for each κ_i (which is the simulated coverage probability), and select the smallest κ_i for which the simulated coverage probability is at least $1 - \alpha$ as an approximation to κ_{corr}.

With probability one, the value of κ thus obtained is not less than the actual value of κ_{corr}. Usually, we may set $\kappa_L = -0.20$, $\kappa_U = 0.8$ and $d_\kappa = 0.05$.

Table 3.1 presents the estimated values of κ_m for various numbers of raters and subjects by using the above methods with the following settings: $m = 20,000$; $\rho_L = 0.6$, $\rho_U = 0.9$, $d = 0.1$; $r = 2$, $k_1 = -1$ (so $\delta_L = 0.5$), $k_2 = 4$ (so δ_U is equal to the values in the table); $d_\kappa = 0.01$; $\alpha = 0.10$, $\kappa_L = 0.0$, $\kappa_U = 0.80$ for two-sided confidence intervals and $\alpha = 0.05$, $\kappa_L = -.20$, $\kappa_U = 1.20$ for one-sided confidence lower bounds.

Table 3.1. Estimates of κ_m for two-sided 90% confidence intervals and one-sided 95% confidence lower bounds of ρ (based on 20,000 simulated samples)

		Upper bound δ_U of $\delta = \sigma_r^2/\sigma_e^2$							
		1	4	8	16	1	4	8	16
R	S	Two-sided ($\alpha = 0.10$)				One-sided ($\alpha = 0.05$)			
3	10	0.04	0.24	0.31	0.32	0.05	0.51	0.64	0.72
3	25	0.13	0.44	0.50	0.52	0.40	0.90	1.00	1.03
3	50	0.35	0.60	0.62	0.67	0.75	1.12	1.16	1.20
5	10	0.12	0.13	0.13	0.13	-0.08	0.19	0.29	0.33
5	25	0.06	0.17	0.23	0.23	0.26	0.53	0.57	0.59
5	50	0.14	0.29	0.32	0.33	0.47	0.75	0.77	0.77

Chapter 4

A SIMULATION STUDY

The main purpose of this Monte-Carlo simulation study is to compare the performance of the modified profile likelihood approach (MPL) with that of the GV approach described in [10], so the parameter settings in [10] were used here. That is, the number of raters were $R = 3$ and 5; the number of subjects were $S = 10$, 25 and 50; the value of ratio of ratio-to-error variability σ_r^2/σ_e^2 were 0.5, 1.0, 4.0; the value of ρ were 0.6, 0.75, 0.90. For each parameter setting, 20,000 random samples were generated. For the GV approach, 10,000 values of R_ρ's were created for each of the 20,000 random samples.

Tables 4.1 - 4.3 present the results about two-sided confidence intervals for ρ produced by the GV approach and the PL approach with different values of κ. The results for the PL method with $\kappa = 0$ are presented in Table 4.1, where we can see that, in all cases the average length of the confidence intervals produced by the PL approach is significantly shorter than that by the GV method. Unfortunately, the PL approach consistently understates the coverage probabilities. This phenomenon is obvious especially when the number of raters is only three, or the number of subjects is large ($S = 50$), or the value of δ is high.

After a careful examination of the results in Table 4.1 we find that, the confidence intervals produced by the PL approach with $\kappa = 0$ is "unnecessarily" narrow. For example, when $R = 3,\ S = 50, \delta = 4.0$ and $\rho = 0.6$, (This is the worst case for the PL method in the sense that the coverage probability is extremely low, only 79.6%), the average length of the confidence intervals from the PL approach is 0.420, while that from the GV method is 0.604. This fact gives us the confidence to increase the value of κ for achieving the desired coverage probabilities with comparable precision. Table 4.2 has the results for the PL approach with estimated correct values of κ, whose superiority to the GV method is clearly illustrated. The

Table 4.1. Empirical probabilities (×1000) and average lengths of approximate 90% confidence intervals for ρ (based on 20,000 simulations). For the PL method, $\kappa = 0$.

		GV		PL		GV		PL	
$\frac{\sigma_r^2}{\sigma_e^2}$	ρ	CP	AL	CP	AL	CP	AL	CP	AL
		$R = 3,\ S = 10$				$R = 5,\ S = 10$			
0.5	0.60	914	0.606	902	0.498	915	0.483	891	0.428
0.5	0.75	912	0.564	902	0.419	912	0.415	889	0.361
0.5	0.90	915	0.406	895	0.263	910	0.250	881	0.213
1.0	0.60	914	0.619	898	0.497	910	0.499	895	0.438
1.0	0.75	906	0.591	898	0.430	909	0.437	887	0.376
1.0	0.90	910	0.449	892	0.286	910	0.274	886	0.231
4.0	0.60	899	0.653	863	0.516	901	0.536	887	0.471
4.0	0.75	905	0.635	871	0.459	901	0.484	887	0.412
4.0	0.90	898	0.505	878	0.322	901	0.321	885	0.266
		$R = 3,\ S = 25$				$R = 5,\ S = 25$			
0.5	0.60	909	0.504	902	0.362	908	0.369	900	0.309
0.5	0.75	915	0.481	901	0.301	911	0.317	900	0.255
0.5	0.90	912	0.346	903	0.183	906	0.184	901	0.143
1.0	0.60	904	0.545	878	0.388	908	0.408	891	0.339
1.0	0.75	906	0.535	885	0.335	905	0.362	891	0.286
1.0	0.90	902	0.404	891	0.212	905	0.220	894	0.167
4.0	0.60	898	0.616	819	0.449	898	0.481	868	0.402
4.0	0.75	897	0.607	834	0.389	900	0.437	877	0.344
4.0	0.90	895	0.482	864	0.259	897	0.283	880	0.210
		$R = 3,\ S = 50$				$R = 5,\ S = 50$			
0.5	0.60	911	0.460	882	0.298	904	0.319	889	0.252
0.5	0.75	907	0.446	885	0.251	904	0.276	892	0.208
0.5	0.90	905	0.324	897	0.150	901	0.159	898	0.114
1.0	0.60	902	0.513	838	0.340	901	0.370	875	0.294
1.0	0.75	902	0.506	850	0.291	902	0.332	883	0.247
1.0	0.90	903	0.390	871	0.181	903	0.201	882	0.140
4.0	0.60	901	0.604	796	0.420	898	0.462	849	0.372
4.0	0.75	898	0.598	812	0.362	899	0.420	857	0.314
4.0	0.90	899	0.477	845	0.232	896	0.269	875	0.186

GV: Generalized variable method
PL: profile-likelihood method
CP: coverage probability
AL: average length of the confidence intervals

estimates of κ_{corr} shown in the table were searched by using the increment $d_\kappa = 0.01$ and $m = 20,000$ simulated random samples.

If we have no prior information of δ, then we can use the modified profile likelihood (MPL) approach with the estimated values of κ_m corresponding to $\delta_U = 16$ in Table 3.1 The resulting confidence intervals are conservative, but their average lengths are still significantly shorter than that by the GV method, as shown in Table 4.3.

The results for one-sided lower bounds are similar. See Tables 4.4 - 4.6. Since ρ is bounded by one from above, a one-sided lower bound and one constitute a confidence interval for ρ. Thus, the average length of such confidence intervals, (which is equal to one minus the average of the one-sided lower bounds), is a good measure to assess the accuracy of a method that produces one-sided lower bounds. The average lengths (AL) in Tables 4.4 - 4.6 are computed as one minus the averages of one-sided lower bounds.

Table 4.2. Empirical probabilities ($\times 1000$) and average lengths of approximate 90% confidence intervals for ρ (based on 20,000 simulations) as well as the estimates of κ. For the PL method, $\kappa = \hat{\kappa}_{opt}$.

		GV		PL			GV		PL		
$\frac{\sigma_r^2}{\sigma_e^2}$	ρ	CP	AL	CP	AL	$\hat{\kappa}_{opt}$	CP	AL	CP	AL	$\hat{\kappa}_{opt}$
		$R = 3,\ S = 10$					$R = 5,\ S = 10$				
0.5	0.60	914	0.606	904	0.498	0.00	915	0.483	900	0.439	0.05
0.5	0.75	912	0.564	900	0.421	0.01	912	0.415	901	0.379	0.10
0.5	0.90	915	0.406	900	0.270	0.03	910	0.250	901	0.228	0.11
1.0	0.60	914	0.619	901	0.504	0.03	910	0.499	901	0.450	0.06
1.0	0.75	906	0.591	900	0.433	0.01	909	0.437	901	0.385	0.04
1.0	0.90	910	0.449	900	0.294	0.04	910	0.274	902	0.251	0.12
4.0	0.60	899	0.653	901	0.570	0.24	901	0.536	900	0.489	0.09
4.0	0.75	905	0.635	901	0.501	0.17	901	0.484	903	0.433	0.10
4.0	0.90	898	0.505	900	0.356	0.15	901	0.321	901	0.290	0.13
		$R = 3,\ S = 25$					$R = 5,\ S = 25$				
0.5	0.60	909	0.504	901	0.364	0.01	908	0.369	903	0.310	0.00
0.5	0.75	915	0.481	903	0.303	0.00	911	0.317	901	0.254	0.00
0.5	0.90	912	0.346	906	0.183	0.00	906	0.184	900	0.142	0.00
1.0	0.60	904	0.545	901	0.417	0.13	908	0.408	900	0.349	0.05
1.0	0.75	906	0.535	901	0.359	0.12	905	0.362	900	0.295	0.05
1.0	0.90	902	0.404	901	0.227	0.08	905	0.220	901	0.171	0.03
4.0	0.60	898	0.616	901	0.547	0.44	898	0.481	900	0.436	0.17
4.0	0.75	897	0.607	901	0.480	0.38	900	0.437	901	0.368	0.12
4.0	0.90	895	0.482	901	0.305	0.22	897	0.283	901	0.223	0.09
		$R = 3,\ S = 50$					$R = 5,\ S = 50$				
0.5	0.60	911	0.460	900	0.315	0.09	904	0.319	902	0.256	0.03
0.5	0.75	907	0.446	900	0.262	0.07	904	0.276	900	0.209	0.02
0.5	0.90	905	0.324	900	0.152	0.02	901	0.159	901	0.115	0.01
1.0	0.60	902	0.513	900	0.408	0.35	901	0.370	901	0.316	0.14
1.0	0.75	902	0.506	901	0.351	0.29	902	0.332	900	0.266	0.13
1.0	0.90	903	0.390	900	0.211	0.18	903	0.201	901	0.152	0.10
4.0	0.60	901	0.604	901	0.546	0.60	898	0.462	901	0.427	0.29
4.0	0.75	898	0.598	901	0.481	0.53	899	0.420	901	0.358	0.24
4.0	0.90	899	0.477	901	0.297	0.33	896	0.269	901	0.203	0.13

GV: Generalized variable method
PL: profile-likelihood method
CP: coverage probability
AL: average length of the confidence intervals

Table 4.3. Empirical probabilities (×1000) and average lengths of approximate 90% confidence intervals for ρ (based on 20,000 simulations). For the MPL method, κ_m equals to the estimates corresponding to $\delta_U = 16$ in Table 3.1.

		GV		MPL		GV		MPL	
$\frac{\sigma_r^2}{\sigma_e^2}$	ρ	CP	AL	CP	AL	CP	AL	CP	AL
		$R = 3,\ S = 10,\ \kappa_m = 0.32$				$R = 5,\ S = 10,\ \kappa_m = 0.13$			
0.5	0.60	914	0.606	945	0.569	915	0.483	912	0.454
0.5	0.75	912	0.564	943	0.489	912	0.415	908	0.384
0.5	0.90	915	0.406	939	0.327	910	0.250	901	0.232
1.0	0.60	914	0.619	941	0.570	910	0.499	918	0.464
1.0	0.75	906	0.591	941	0.502	909	0.437	908	0.401
1.0	0.90	910	0.449	939	0.352	910	0.274	906	0.252
4.0	0.60	899	0.653	914	0.589	901	0.536	906	0.498
4.0	0.75	905	0.635	921	0.536	901	0.484	907	0.439
4.0	0.90	898	0.505	925	0.396	901	0.321	904	0.288
		$R = 3,\ S = 25,\ \kappa_m = 0.52$				$R = 5,\ S = 25,\ \kappa_m = 0.23$			
0.5	0.60	909	0.504	961	0.460	908	0.369	928	0.345
0.5	0.75	915	0.481	962	0.402	911	0.317	933	0.289
0.5	0.90	912	0.346	963	0.269	906	0.184	930	0.165
1.0	0.60	904	0.545	945	0.494	908	0.408	924	0.379
1.0	0.75	906	0.535	950	0.443	905	0.362	929	0.324
1.0	0.90	902	0.404	953	0.309	905	0.220	926	0.194
4.0	0.60	898	0.616	908	0.562	898	0.481	909	0.449
4.0	0.75	897	0.607	919	0.511	900	0.437	911	0.388
4.0	0.90	895	0.482	934	0.370	897	0.283	918	0.242
		$R = 3,\ S = 50,\ \kappa_m = 0.67$				$R = 5,\ S = 50,\ \kappa_m = 0.33$			
0.5	0.60	911	0.460	963	0.416	904	0.319	940	0.298
0.5	0.75	907	0.446	964	0.369	904	0.276	939	0.250
0.5	0.90	905	0.324	966	0.248	901	0.159	942	0.142
1.0	0.60	902	0.513	936	0.466	901	0.370	919	0.346
1.0	0.75	902	0.506	942	0.424	902	0.332	928	0.297
1.0	0.90	903	0.390	950	0.296	903	0.201	934	0.176
4.0	0.60	901	0.604	908	0.559	898	0.462	907	0.435
4.0	0.75	898	0.598	921	0.509	899	0.420	916	0.375
4.0	0.90	899	0.477	937	0.364	896	0.269	927	0.230

GV: Generalized variable method
PL: profile-likelihood method
CP: coverage probability
AL: average length of the confidence intervals

Table 4.4. Empirical probabilities (×1000) and "average lengths" of approximate 95% confidence lower bounds for ρ (based on 20,000 simulations). For the PL method, $\kappa = 0$.

		GV		PL		GV		PL	
$\frac{\sigma_r^2}{\sigma_e^2}$	ρ	CP	AL	CP	AL	CP	AL	CP	AL
		$R = 3,\ S = 10$				$R = 5,\ S = 10$			
0.5	0.60	987	0.818	959	0.707	976	0.703	960	0.510
0.5	0.75	985	0.689	937	0.607	974	0.544	937	0.435
0.5	0.90	984	0.456	908	0.567	976	0.297	912	0.405
1.0	0.60	980	0.824	968	0.662	971	0.713	956	0.331
1.0	0.75	980	0.712	951	0.576	971	0.565	924	0.263
1.0	0.90	980	0.494	929	0.541	972	0.322	891	0.236
4.0	0.60	961	0.831	964	0.544	960	0.730	958	0.282
4.0	0.75	961	0.735	942	0.450	956	0.595	939	0.224
4.0	0.90	959	0.541	916	0.410	957	0.362	924	0.203
		$R = 3,\ S = 25$				$R = 5,\ S = 25$			
0.5	0.60	984	0.771	966	0.503	976	0.643	892	0.672
0.5	0.75	982	0.639	952	0.415	976	0.480	838	0.620
0.5	0.90	983	0.408	934	0.381	975	0.244	807	0.599
1.0	0.60	975	0.793	967	0.314	966	0.670	933	0.663
1.0	0.75	972	0.676	951	0.240	967	0.517	897	0.612
1.0	0.90	970	0.460	926	0.210	966	0.281	873	0.588
4.0	0.60	954	0.820	968	0.270	955	0.711	905	0.550
4.0	0.75	952	0.722	958	0.205	954	0.570	855	0.489
4.0	0.90	953	0.525	940	0.180	952	0.331	828	0.465
		$R = 3,\ S = 50$				$R = 5,\ S = 50$			
0.5	0.60	979	0.752	947	0.693	973	0.621	938	0.528
0.5	0.75	977	0.622	900	0.611	970	0.455	908	0.465
0.5	0.90	976	0.393	858	0.581	973	0.228	886	0.441
1.0	0.60	967	0.785	955	0.661	962	0.656	920	0.359
1.0	0.75	964	0.669	930	0.590	962	0.501	888	0.298
1.0	0.90	961	0.445	900	0.561	958	0.265	861	0.269
4.0	0.60	948	0.814	951	0.546	951	0.702	943	0.311
4.0	0.75	950	0.721	910	0.466	954	0.561	922	0.256
4.0	0.90	954	0.523	865	0.433	952	0.322	905	0.233

GV: Generalized variable method
PL: profile-likelihood method
CP: coverage probability
AL: One minus the average of lower bounds

Table 4.5. Empirical probabilities (×1000) and average lengths of approximate 95% confidence lower bounds for ρ (based on 20,000 simulations) as well as the estimates of κ_{corr}. For the PL method, $\kappa = \hat{\kappa}_{opt}$.

		GV		PL			GV		PL		
$\frac{\sigma_r^2}{\sigma_e^2}$	ρ	CP	AL	CP	AL	$\hat{\kappa}_{opt}$	CP	AL	CP	AL	$\hat{\kappa}_{opt}$
		$R=3,\ S=10$					$R=5,\ S=10$				
0.5	0.60	987	0.818	950	0.690	-0.12	976	0.703	951	0.641	-0.20
0.5	0.75	985	0.689	950	0.519	-0.14	974	0.544	952	0.476	-0.20
0.5	0.90	984	0.456	951	0.282	-0.20	976	0.297	951	0.476	-0.20
1.0	0.60	980	0.824	950	0.703	0.05	971	0.713	950	0.651	-0.08
1.0	0.75	980	0.712	951	0.547	0.00	971	0.565	950	0.493	-0.13
1.0	0.90	980	0.494	951	0.325	-0.40	972	0.322	951	0.267	-0.14
4.0	0.60	961	0.831	951	0.758	0.51	960	0.730	951	0.692	0.19
4.0	0.75	961	0.735	951	0.638	0.43	956	0.595	950	0.547	0.15
4.0	0.90	959	0.541	951	0.437	0.36	957	0.362	950	0.324	0.03
		$R=3,\ S=25$					$R=5,\ S=25$				
0.5	0.60	984	0.771	950	0.621	0.10	976	0.643	951	0.580	0.02
0.5	0.75	982	0.639	950	0.459	0.07	976	0.480	950	0.412	-0.01
0.5	0.90	983	0.408	950	0.243	0.02	975	0.244	951	0.200	-0.06
1.0	0.60	975	0.793	950	0.675	0.40	966	0.670	950	0.620	0.26
1.0	0.75	972	0.676	950	0.529	0.35	967	0.517	951	0.463	0.23
1.0	0.90	970	0.460	950	0.306	0.26	966	0.281	950	0.238	0.14
4.0	0.60	954	0.820	950	0.766	0.90	955	0.711	951	0.684	0.53
4.0	0.75	952	0.722	950	0.641	0.77	954	0.570	950	0.531	0.44
4.0	0.90	953	0.525	951	0.419	0.61	952	0.331	951	0.298	0.33
		$R=3,\ S=50$					$R=5,\ S=50$				
0.5	0.60	979	0.752	950	0.617	0.38	973	0.621	950	0.565	0.24
0.5	0.75	977	0.622	950	0.458	0.32	970	0.455	950	0.399	0.20
0.5	0.90	976	0.393	951	0.239	0.23	973	0.228	950	0.188	0.10
1.0	0.60	967	0.785	950	0.755	1.07	962	0.656	950	0.615	0.47
1.0	0.75	964	0.669	950	0.556	0.71	962	0.501	951	0.458	0.43
1.0	0.90	961	0.445	950	0.317	0.53	958	0.265	950	0.229	0.30
4.0	0.60	948	0.814	950	0.776	1.12	951	0.702	950	0.687	0.75
4.0	0.75	950	0.721	950	0.649	0.96	954	0.561	950	0.529	0.62
4.0	0.90	954	0.523	950	0.421	0.79	952	0.322	950	0.289	0.45

GV: Generalized variable method
PL: profile-likelihood method
CP: coverage probability
AL: One minus the average of lower bounds

Table 4.6. Empirical probabilities (×1000) and average lengths of approximate 90% confidence lower bounds for ρ (based on 20,000 simulations). For the MPL method, κ_m equals to the values corresponding to $\delta_U = 16$ in Table 3.1.

		GV		MPL		GV		MPL	
$\frac{\sigma_r^2}{\sigma_e^2}$	ρ	CP	AL	CP	AL	CP	AL	CP	AL
		$R=3,\ S=10,\ \kappa_m=0.72$				$R=5,\ S=10,\ \kappa_m=0.23$			
0.5	0.60	987	0.818	991	0.804	976	0.703	983	0.700
0.5	0.75	985	0.689	992	0.664	974	0.544	983	0.543
0.5	0.90	984	0.456	992	0.440	976	0.297	983	0.304
1.0	0.60	980	0.824	985	0.793	971	0.713	975	0.698
1.0	0.75	980	0.712	988	0.673	971	0.565	977	0.554
1.0	0.90	980	0.494	987	0.469	972	0.322	979	0.326
4.0	0.60	961	0.831	964	0.787	960	0.730	958	0.707
4.0	0.75	961	0.735	966	0.689	956	0.595	960	0.576
4.0	0.90	959	0.541	972	0.512	957	0.362	965	0.360
		$R=3,\ S=25,\ \kappa_m=1.03$				$R=5,\ S=25,\ \kappa_m=0.59$			
0.5	0.60	984	0.771	991	0.739	976	0.643	981	0.632
0.5	0.75	982	0.639	991	0.610	976	0.480	981	0.476
0.5	0.90	983	0.408	992	0.403	975	0.244	984	0.254
1.0	0.60	975	0.793	980	0.756	966	0.670	970	0.657
1.0	0.75	972	0.676	985	0.643	967	0.517	976	0.510
1.0	0.90	970	0.460	986	0.452	966	0.281	975	0.285
4.0	0.60	954	0.820	957	0.781	955	0.711	955	0.688
4.0	0.75	952	0.722	963	0.684	954	0.570	960	0.552
4.0	0.90	953	0.525	974	0.504	952	0.331	965	0.334
		$R=3,\ S=50,\ \kappa_m=1.20$				$R=5,\ S=50,\ \kappa_m=0.77$			
0.5	0.60	979	0.752	988	0.723	973	0.621	976	0.614
0.5	0.75	977	0.622	988	0.596	970	0.455	980	0.457
0.5	0.90	976	0.393	991	0.391	973	0.228	983	0.240
1.0	0.60	967	0.785	972	0.754	962	0.656	967	0.647
1.0	0.75	964	0.669	975	0.638	962	0.501	972	0.498
1.0	0.90	961	0.445	981	0.439	958	0.265	974	0.275
4.0	0.60	948	0.814	956	0.788	951	0.702	953	0.691
4.0	0.75	950	0.721	963	0.684	954	0.561	961	0.551
4.0	0.90	954	0.523	972	0.501	952	0.322	969	0.329

GV: Generalized variable method
PL: profile-likelihood method
CP: coverage probability
AL: One minus the average of lower bounds

Chapter 5

EXAMPLES

Example 1. Fleiss *et al.*[5] conducted a reliability test for four raters ($R = 4$), each of whom evaluated the teeth of ten patients ($S = 10$) independently and recorded the number of decayed, missing and filled surfaces of patients' permanent teeth (DMFS score). We estimate ρ to be 0.8987, with the estimate of rater-to-ratio error variability $\delta = 1.26$ by profile-likelihood(PL) method . By PL method, the two-sided 90% confidence interval is $(0.7120, 0.9598)$, and the lower bound of 95% confidence is 0.7120. The modified profile-likelihood method (MPL) gives a two-sided 90% confidence interval of $(0.6808, 0.9631)$ with $\kappa = 0.18$ and the lower bound for 95% one-sided confidence interval of 0.6290 with $\kappa = 0.47$. The data were also analyzed by [2] using the large sample approach and by [10] using the generalized variable(GV) method. The estimate of ρ given by GV method is 0.9037, 90% two-sided confidence interval is (0.6295,0.9614) and lower bound of one-sided 95%confidence interval is 0.6201.

Example 2. Streiner *et al.*[9] have presented a example of rates' effect on patients' sadness score where each of the three observers($R = 3$) gave a ten points scale to measure a patient's sadness ($S = 10$). The estimate of ρ is 0.7304 by MPL method, and 90% two-sided confidence interval is $(0.3062, 0.9023)$, with value of κ set as 0.29. MPL method also estimates the 95% one-sided lower bound to be 0.2221 with $\kappa = 0.74$. GV method gives an estimate of ρ of 0.7141, 90% two-sided confidence interval of $(0.1481, 0.8772)$ and lower bound of 95% one-sided confidence interval of 0.1499.

Example 3. Cuttie *et al.*[3] tested measurement reliability between two raters($R = 2$) for measuring hips and knees static flexion angles for both legs on nine ($S = 18$) healthy children, when their bodies were oriented in order to define Outwalk anatomical coordinate systems. For value h (hip static flexion) , MPL method yields a 90 % confidence interval for ρ of $(0.4653, 0.9444)$, while GV gives a 90% confidence interval of $(0.1587, 0.9221)$. Lower bound for 95% test is 0.3839 by MPL method and 0.1474 by GV method. For value k (knee static flexion), MPL method estimates two-sided 90% confidence interval and one-sided 95% lower bound to be $(0.2686, 0.8889)$ and 0.2251, respectively, while GV gives the estimates of $(0.0536, 0.8429)$ and 0.0495, respectively. The estimates of rater-to-ratio error variability are less than 0.15 in both cases.

Example 4. Yi *et al.*[13] *et al.* used measurements of systolic blood pressure by observer J on eighty-five patients ($I = 3, S = 85$) to show agreement across tools. Estimate of ρ is 0.9611 by MPL and 0.9615 by GV method. Length of 90% confidence interval is 0.0368 by MPL and 0.0352 by GV approach. 95% confidence lower bound by GV method is 0.9352 while that given by MPL approach is 0.9039. The estimated value for δ is less than 0.2.

Example 5. In a study conducted by Vaid *et al.* [11], four different methods to segment magnetic resonance images (MRI) of brain were utilized to measure the volumes of four patients' tumors.($I = 4, S = 4$). The estimated value of δ is 2.18. ρ is estimated to be 0.7281 by MPL method. The 90% confidence interval is determined to be $(0.2672, 0.9359)$ when $\kappa = 0.45$. Point estimate of ρ by GV method is 0.7644 and 90% two-sided confidence interval is $(0.2672, 0.9539)$. For the lower bound of 95% one-sided confidence interval , MPL gives a value of 0.3085 with $\kappa = 0.21$ while GV gives a value of 0.2898.

Chapter 6

DISCUSSIONS AND FUTURE WORKS

In this paper, it is of interest to study the interrater reliability coefficient ρ from a two-way random effects analysis of variance model in which every rater scores each subject. The likelihood function is simplified, the profile-likelihood approach is derived, studied and applied to obtain one-sided lower bounds and two-sided bounds for ρ. The profile-likelihood (PL) approach is more accurate than existing methods in the sense that it always produces confidence intervals of shorter average lengths. However, the PL approach with $\kappa = 0$ understates the coverage probabilities in almost all cases. Fortunately, simulation study shows that, we can increase the value of κ appropriately so that the PL approach is still more precise than existing methods, but does not understate the coverage probabilities. A good choice for κ is κ_m defined in (3.9), resulting in the modified profile likelihood (MPL) approach. For the parameter settings in Section 4, the MPL approach is always more accurate than the GV approach. Estimates of κ_m can can be obtained through Monte-Carlo simulation. Table 3.1 gives the estimates of κ_m for several cases.

The MPL approach is in general conservative in the sense that the coverage probability are in general higher than the expected coverage probability $1-\alpha$. In other words, the actual probability that the MPL approach produces correct results is higher than claimed. Thus, from the point view of application, this feature of the MPL approach is not evil and should be appreciated since it is also more precise than existing methods.

How conservative the MPL approach can be depends on the value of

$$\kappa_m - \kappa_{corr} (\equiv \kappa_m - \kappa_{corr}(\rho_0,\ \delta_0)), \tag{6.1}$$

where ρ_0 and δ_0 are the true parameter values of ρ and δ. A larger value of $\kappa_m - \kappa_{corr}$ indicates that the MPL approach is more conservative. For example, for the case $R = 3,\ S = 50$ and $\rho_0 = 0.90,\ \delta_0 = 0.5$, the optimal value of κ is $\kappa_{corr} = 0.02$ for producing two-sided 90% confidence intervals (Table 4.2). If we have no knowledge about δ, we may set $\kappa_m = 0.67$ (the number corresponding to $\delta_U = 16$ in Table 3.1), and the MPL approach becomes substantially conservative with the coverage probability 99.6% (Table 4.3). The MPL approach is not always so conservative. For example, when $R = 3,\ 5$ and $S = 10$, the values of κ_{corr} are less variable (Table 4.2), so the value of $\kappa_m - \kappa_{corr}$ is not expected to be high and the MPL approach is less conservative than the GV method for certain parameter settings (Table 4.3). Simulation study shows that, for fixed number R of raters, the value of $\kappa_m - \kappa_{corr}$ becomes larger as the number S of subjects increases. When there are only a few raters but many subjects, the MPL approach may be very conservative for certain parameter settings. On the other hand, the value of $\kappa_m - \kappa_{corr}$ decreases as R increases (see Table 3.1) for fixed number of subjects. In general, if the ratio R/S is not too low, the value of $\kappa_m - \kappa_{corr}$ is near to zero and the actual coverage probability of the MPLapproach is close to the expected one.

If a less conservative MPL approach is more desirable, the search for an estimate of κ_m can be made on a narrower range of δ. For example, with the advance of technology, measuring instruments may have excellent intrarater or test-rerest reliability, so the rater component of variability in the ratings is large relative to the random error component, that is, the value of δ is high. Therefore, we should search for an estimate of κ_m in the range of high values of δ. As a result, the MPL approach is less conservative, but also more precise.

In conclusion, the modified profile likelihood approach proposed here is recommended as having the best overall performance among the methods mentioned in this paper.

REFERENCES

[1] Bartko, J.J(1966). The intraclass correlation coefficient as a measure of reliability,*Psychological Reports*,**19**,3-11.

[2] Cappelleri, J.C, and Ting, N.(2003). A modified large sample approach to approximate interval estimation for a particular intraclass correlation coefficient. *Statistics in Medicine*, **22**, 1861-1877.

[3] Cutti, A.G.*et al.* (2010). 'Outwalk' : a protocal for clinical gait analysis based on inertial and magnetic sensors. *Medical & Biological Engineering & Computing*, **48**,17-25.

[4] Fleiss, J.L. and Shrout, P.E.(1978). Approximate interval estimation for a certain intraclass correlation coefficient. *Psychometrika*, **43**, 259-262.

[5] Fleiss J.L. (1986)*The Design and Analysis of Clinical Experiment*. Wiley: New York.

[6] Nelder, J.A., Mead, R.(1965). A simplex method for function minimization, *Computer Journal*, **7**,308-313.

[7] Rajaratnam, N.(1960). Reliability forumlas for independent decision data when reliability data matched.*Psychometrika*,**25**,262-271.

[8] Satterthwaite, F.E.(1946). An approximate distribution of estimates of variance components, *Biometrics* ,**2**, 110-114.

[9] Streiner, D.L., Norman, G.R.(1995) *Health Measurement Scales: A Practical Guide to Their Development and Use*. Second edition. Oxford University Press, NY.

[10] Tian, L., and Cappelleri, J.C.(2003). A new approach for interval estimation and hypothesis testing of a certain intraclass correlation coefficient: the generalized variable method. *Statistics in Medicine*, **23**, 2125-2135

[11] Vaidyanathan, M.*et al.* (1995). Comparison of supervised MRI segementation methods for tumor volume determination during therapy. *Magnetic Resonance Imaging*, **5**,719-728.

[12] Weerahandi, S.(1993). Generalized confidence intervals. *Journal of American Statistical Association* **88**, 899-905.

[13] Yi, Q., Wang, P., and He,Y. (2008). Reliability analysis for continuous measurements: Equivalence test for agreement.*Statist. Med.*,**27** 2816-2825.

[14] Zou, H., and McDermott, M.P.(1999). Higher-moments approaches to approximate interval estimation for a certain intraclass correlation coefficient. *Statistics in Medicine*, **18**, 2051-2061.

q

APPENDICES

Derivation of of Likelihood Function

In this section we will perform a series of algebraic operations for simplifying the expression of the likelihood function. First we introduce several notations. For an integer n, let $\mathbf{1}_n$ denote the n-dimensional one-vector whose components are one

$$\mathbf{1}_n = \begin{bmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{bmatrix}, \tag{2}$$

$\boldsymbol{I}_n$ the $n \times n$ identity matrix, and $\boldsymbol{J}_n$ the $n \times n$ one-matrix

$$\boldsymbol{J}_n = \begin{bmatrix} 1 & 1 & \cdots & 1 \\ 1 & 1 & \cdots & 1 \\ \cdots & \cdots & \cdots & \cdots \\ 1 & 1 & \cdots & 1 \end{bmatrix} \equiv \mathbf{1}_n \mathbf{1}_n^T. \tag{3}$$

For simplicity, the subscript n is suppressed in case of no confusion. Let y_j be the data on the jth subject:

$$y_j = \begin{bmatrix} y_{1j} \\ y_{2j} \\ \vdots \\ y_{Rj} \end{bmatrix}, \tag{4}$$

and y is the RS-dimensional vector of all data

$$\boldsymbol{y} = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_S \end{bmatrix}. \tag{5}$$

It follows that the covariance matrix of y is the $RS \times RS$ matrix $\sigma^2 \boldsymbol{V}$, where

$$\boldsymbol{V} = \begin{bmatrix} (1-\rho_s)\boldsymbol{I}_R + \rho_s \boldsymbol{J}_R & \rho_r \boldsymbol{I}_R & \cdots & \rho_r \boldsymbol{I}_R \\ \rho_r \boldsymbol{I}_R & (1-\rho_s)\boldsymbol{I}_R + \rho_s \boldsymbol{J}_R & \cdots & \rho_r \boldsymbol{I}_R \\ \cdots & \cdots & \cdots & \cdots \\ \rho_r \boldsymbol{I}_R & \rho_r \boldsymbol{I}_R & \cdots & (1-\rho_s)\boldsymbol{I}_R + \rho_s \boldsymbol{J}_R \end{bmatrix}. \tag{6}$$

Thus the log-likelihood function is given by

$$\begin{aligned} l &= -\frac{1}{2}\left\{RS\ln(2\pi) + \ln|\sigma^2\boldsymbol{V}| + (\boldsymbol{y}-\mu\boldsymbol{1})^T(\sigma^2\boldsymbol{V})^{-1}(\boldsymbol{y}-\mu\boldsymbol{1})\right\} \\ &= -\frac{1}{2}\left\{RS\ln(2\pi) + RS\ln\sigma^2 + \ln|\boldsymbol{V}| + \frac{(\boldsymbol{y}-\mu\boldsymbol{1})^T\boldsymbol{V}^{-1}(\boldsymbol{y}-\mu\boldsymbol{1})}{\sigma^2}\right\} \end{aligned} \tag{7}$$

and

$$-2l = RS\ln(2\pi) + RS\ln\sigma^2 + \ln|\boldsymbol{V}| + \frac{(\boldsymbol{y}-\mu\boldsymbol{1})^T\boldsymbol{V}^{-1}(\boldsymbol{y}-\mu\boldsymbol{1})}{\sigma^2}. \tag{8}$$

Setting to zero the partial derivative of $-2l$ with respect to μ

$$\frac{\partial(-2l)}{\partial\mu} = -\frac{2}{\sigma^2}(\boldsymbol{1}^T\boldsymbol{V}^{-1}y - \mu\boldsymbol{1}^T\boldsymbol{V}^{-1}\boldsymbol{1}). \tag{9}$$

gives the maximum likelihood estimator (MLE) $\hat{\mu}$ of μ

$$\hat{\mu} = \frac{\boldsymbol{1}^T\boldsymbol{V}^{-1}\boldsymbol{y}}{\boldsymbol{1}^T\boldsymbol{V}^{-1}\boldsymbol{1}} = \frac{\boldsymbol{1}^T\boldsymbol{y}}{\boldsymbol{1}^T\boldsymbol{1}} = \bar{y}_{\cdot\cdot}. \tag{10}$$

since the vector **1** is an eigenvalue of the matrix $\boldsymbol{V}^{-1}$ (and $\boldsymbol{V}$). Similarly, equating with zero the partial derivative of $-2l$ with respect to σ^2

$$\frac{\partial(-2l)}{\partial\sigma^2} = \frac{RS}{\sigma^2} - \frac{(\mathbf{1}^T\boldsymbol{V}^{-1}\boldsymbol{y} - \mu\mathbf{1}^T\boldsymbol{V}^{-1}\mathbf{1})}{(\sigma^2)^2} \tag{11}$$

yields the MLE $\hat{\sigma^2}$ of σ^2

$$\hat{\sigma^2} = \frac{(\boldsymbol{y} - \mu\mathbf{1})^T\boldsymbol{V}^{-1}(\boldsymbol{y} - \mu\mathbf{1})}{RS}. \tag{12}$$

Replace μ by its maximum likelihood estimate $\bar{y}_{..}$, we shall have

$$\hat{\sigma^2} = \frac{(\boldsymbol{y} - \bar{y}_{..}\mathbf{1})^T\boldsymbol{V}^{-1}(\boldsymbol{y} - \bar{y}_{..}\mathbf{1})}{RS}. \tag{13}$$

Let

$$\Delta \overset{def}{=} (\boldsymbol{y} - \bar{y}_{..}\mathbf{1})^T\boldsymbol{V}^{-1}(\boldsymbol{y} - \bar{y}_{..}\mathbf{1}), \tag{14}$$

then

$$\hat{\sigma^2} = \frac{\Delta}{RS}, \tag{15}$$

and

$$l = -\frac{1}{2}\left\{RS\ln(2\pi) + RS\ln\sigma^2 + \ln|\boldsymbol{V}| + \frac{\Delta}{\sigma^2}\right\}, \tag{16}$$

$$-2l = RS\ln(2\pi) + RS\ln\sigma^2 + \ln|\boldsymbol{V}| + \frac{\Delta}{\sigma^2}. \tag{17}$$

The nuisance parameter μ is not involved in (14), (15), (16) and (17).

To evaluate the determinant $|\boldsymbol{V}|$ of the matrix $\boldsymbol{V}$ and simplify the quadratic form Δ in (14), we need to find out the eigenvalues and eigenvectors of the matrix $\boldsymbol{V}$.

For an integer n, let $h_1^{(n)},\ h_2^{(n)}, \ldots,\ h_n^{(n)}$ be the $n \times 1$ vectors given by

$$h_1^{(n)} = d_1^{(n)} \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \\ \vdots \\ 1 \end{bmatrix},\ h_2^{(n)} = d_2 \begin{bmatrix} 1 \\ -1 \\ 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix},\ h_3^{(n)} = d_3 \begin{bmatrix} 1 \\ 1 \\ -2 \\ 0 \\ \vdots \\ 0 \end{bmatrix},\ \ldots,\ h_n^{(n)} = d_n \begin{bmatrix} 1 \\ 1 \\ 1 \\ \vdots \\ 1 \\ -(n-1) \end{bmatrix}, \tag{18}$$

where

$$d_1^{(n)} = \frac{1}{\sqrt{n}}, \quad d_i = \frac{1}{\sqrt{(i-1)i}} \quad \text{for } i = 2,\ \ldots, n. \tag{19}$$

The superscript $^{(n)}$ is suppressed if no confusion arises from this omission. Then the vectors $\{h_i\}_{i=1}^{n}$ are eigenvectors of the matrix $\boldsymbol{J}_n$. Indeed,

$$\boldsymbol{J}_n h_1 = n h_1, \quad \boldsymbol{J}_n h_2 = 0 \ \text{ for all } i = 2,\ \ldots, n. \tag{20}$$

The matrix

$$\boldsymbol{H}_n = \begin{bmatrix} h_1 & h_2 & \ldots & h_n \end{bmatrix} \tag{21}$$

is the $n \times n$ orthogonal matrix due to Helmert. Now for $i = 1,\ 2,\ \ldots, R$, let

$$q_{i1} = d_1^{(S)} \begin{bmatrix} h_i^{(R)} \\ h_i^{(R)} \\ h_i^{(R)} \\ h_i^{(R)} \\ \vdots \\ h_i^{(R)} \end{bmatrix},\ q_{i2} = d_2 \begin{bmatrix} h_i^{(R)} \\ -h_i^{(R)} \\ 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix},\ q_{i3} = d_3 \begin{bmatrix} h_i^{(R)} \\ h_i^{(R)} \\ -2h_i^{(R)} \\ 0 \\ \vdots \\ 0 \end{bmatrix},\ \ldots,\ q_{iS} = d_S \begin{bmatrix} h_i^{(R)} \\ h_i^{(R)} \\ h_i^{(R)} \\ \vdots \\ h_i^{(R)} \\ -(S-1)h_i^{(R)} \end{bmatrix}, \tag{22}$$

then the matrix $\boldsymbol{Q}$ given by

$$\boldsymbol{Q} = \begin{bmatrix} q_{11} & \dots & q_{1S}; & q_{21} & \dots & q_{2S}; & \dots & ; & q_{R1} & \dots & q_{RS} \end{bmatrix} \tag{23}$$

is orthogonal. Furthermore,

$$\boldsymbol{V} q_{11} = [1 - \rho_s - \rho_r + R\rho_s + S\rho_r] q_{11} \tag{24}$$

$$\boldsymbol{V} q_{1j} = [1 - \rho_s - \rho_r + R\rho_s] q_{1j} \quad \text{for } j = 2,\ 3,\ \dots, S \tag{25}$$

$$\boldsymbol{V} q_{i1} = [1 - \rho_s - \rho_r + S\rho_r] q_{i1} \quad \text{for } i = 2,\ 3,\ \dots, R \tag{26}$$

$$\boldsymbol{V} q_{ij} = [1 - \rho_s - \rho_r] q_{ij} \quad \text{for } i = 2,\ 3,\ \dots, R \text{ and } j = 2,\ 3,\ \dots,\ S, \tag{27}$$

by (20), so q_{ij}'s are the eigenvectors of $\boldsymbol{V}$ and

$$\lambda_1 \stackrel{def}{=} 1 - \rho_s - \rho_r + R\rho_s + S\rho_r, \tag{28}$$

$$\lambda_2 \stackrel{def}{=} 1 - \rho_s - \rho_r + R\rho_s, \tag{29}$$

$$\lambda_3 \stackrel{def}{=} 1 - \rho_s - \rho_r + S\rho_r, \tag{30}$$

$$\lambda_4 \stackrel{def}{=} 1 - \rho_s - \rho_r, \tag{31}$$

are the eigenvalues of the matrix $\boldsymbol{V}$.

Define two diagonal $S \times S$ matrices $\boldsymbol{\Lambda}_1$ and $\boldsymbol{\Lambda}_2$ by

$$\boldsymbol{\Lambda}_1 = \begin{bmatrix} \lambda_1 & 0 & 0 & \dots & 0 \\ 0 & \lambda_2 & 0 & \dots & 0 \\ 0 & 0 & \lambda_2 & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & \dots & \lambda_2 \end{bmatrix} \text{ and } \boldsymbol{\Lambda}_2 = \begin{bmatrix} \lambda_3 & 0 & 0 & \dots & 0 \\ 0 & \lambda_4 & 0 & \dots & 0 \\ 0 & 0 & \lambda_4 & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & \dots & \lambda_4 \end{bmatrix}, \tag{32}$$

and denote by $\mathbf{\Lambda}$ the following $RS \times RS$ diagonal matrix

$$\mathbf{\Lambda} = \begin{bmatrix} \mathbf{\Lambda}_1 & 0 & 0 & \dots & 0 \\ 0 & \mathbf{\Lambda}_2 & 0 & \dots & 0 \\ 0 & 0 & \mathbf{\Lambda}_2 & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & \dots & \mathbf{\Lambda}_2 \end{bmatrix}. \tag{33}$$

It follows that

$$\boldsymbol{Q}^T \boldsymbol{V} \boldsymbol{Q} = \mathbf{\Lambda} \text{ and } \boldsymbol{V}^{-1} = \boldsymbol{Q} \mathbf{\Lambda}^{-1} \boldsymbol{Q}^T. \tag{34}$$

Since $\boldsymbol{Q}$ is orthogonal, the determinant of the matrix $\boldsymbol{V}$ is

$$|\boldsymbol{V}| = |\mathbf{\Lambda}| = |\mathbf{\Lambda}_1| \times |\mathbf{\Lambda}_2|^{R-1} = \lambda_1 \times \lambda_2^{S-1} \times \lambda_3^{R-1} \times \lambda_4^{(R-1)(S-1)}. \tag{35}$$

It follows that from (34) that

$$\Delta = (\boldsymbol{y} - \bar{y}_{..}\mathbf{1})^T \boldsymbol{Q} \mathbf{\Lambda}^{-1} \boldsymbol{Q}^T (\boldsymbol{y} - \bar{y}_{..}\mathbf{1}) = [\boldsymbol{Q}^T (\boldsymbol{y} - \bar{y}_{..}\mathbf{1})]^T \mathbf{\Lambda}^{-1} [\boldsymbol{Q}^T (\boldsymbol{y} - \bar{y}_{..}\mathbf{1})]. \tag{36}$$

Define $\boldsymbol{z} \stackrel{def}{=} \boldsymbol{Q}^T (\boldsymbol{y} - \bar{y}_{..}\mathbf{1})$, then

$$\begin{aligned} \Delta &= \frac{z_{(1)}^2}{\lambda_1} + \frac{\sum_{j=2}^S z_{(j)}^2}{\lambda_2} + \frac{\sum_{i=2}^R z_{(S(i-1)+1)}^2}{\lambda_3} + \frac{\sum_{i=2}^R \sum_{j=2}^S z_{(S(i-1)+j)}^2}{\lambda_4} \\ &= \frac{a}{\lambda_1} + \frac{b}{\lambda_2} + \frac{c}{\lambda_3} + \frac{d}{\lambda_4}, \end{aligned} \tag{37}$$

where $z_{(k)}$ denotes the kth component of $\boldsymbol{z}$ and

$$a = z_{(1)}^2, \quad b = \sum_{j=2}^S z_{(j)}^2, \quad c = \sum_{i=2}^R z_{(S(i-1)+1)}^2, \quad d = \sum_{i=2}^R \sum_{j=2}^S z_{(S(i-1)+j)}^2. \tag{38}$$

To simplify the expressions of a, b, c and d, let $x_j = y_j - \bar{y}_{\cdot\cdot}\mathbf{1}_R$ for all $j = 1,\ 2,\ \ldots, S$, $x = y - \bar{y}_{\cdot\cdot} \equiv (x_1^T, x_2^T, \ldots, x_S^T)^T$, and

$$\begin{aligned} w_1 &= x_1 + x_2 + \ldots + x_S \equiv d_1^{(S)}[h_1^{(S)}]^T x \\ w_2 &= x_1 - x_2 \equiv d_2^{(S)}[h_2^{(S)}]^T x, \\ w_3 &= x_1 + x_2 - 2y_3 \equiv d_3^{(S)}[h_2^{(S)}]^T x, \\ &\cdots\cdots\cdots\cdots \\ w_S &= x_1 + \ldots + x_{S-1} - (S-1)x_S \equiv d_S^{(S)}[h_S^{(S)}]^T x, \end{aligned} \tag{39}$$

where $h^T x$ is understood as

$$h^T x = \sum_{j=1}^{S} h_{(j)} x_j. \tag{40}$$

By (22) and (39),

$$w_1 = \begin{bmatrix} x_{1\cdot} \\ x_{2\cdot} \\ \vdots \\ x_{R\cdot} \end{bmatrix}, \tag{41}$$

and

$$z_{(S(i-1)+j)} = q_{ij}^T x = d_j^{(S)} h_i^T w_j. \tag{42}$$

It follows that

$$h_1^T w_1 = \frac{1}{R} \times 1^T w_1 = \frac{1}{S} \times x_{\cdot\cdot} = 0, \tag{43}$$

which implies that

$$a = z_{(1)}^2 = (d_1^{(S)} h_1^T w_1)^2 = 0. \tag{44}$$

By (38), (42), (43) and the fact that H_S is an orthogonal matrix,

$$\begin{aligned} b &= \sum_{j=1}^{S} (d_j^{(S)} h_1^T w_j)^2 = \sum_{j=1}^{S} (h_1^T x_j)^2 \\ &= \sum_{j=1}^{S} \left(\frac{x_{\cdot j}}{\sqrt{R}} \right)^2 = R \sum_{j=1}^{S} \bar{x}_{\cdot j}^2 \\ &= R \sum_{j=1}^{S} (\bar{y}_{\cdot j} - \bar{y}_{\cdot\cdot})^2 = \text{SSBS}. \end{aligned} \tag{45}$$

By (38), (42), (43) and the fact that H_R is an orthogonal matrix,

$$\begin{aligned} c &= \sum_{i=1}^{R} (d_1^{(S)} h_i^T w_1)^2 = [d_1^{(S)}]^2 \sum_{i=1}^{R} (h_i^T w_1)^2 \\ &= \frac{1}{S} \times (H_R w_1)^T H_R w_1 = \frac{1}{S} \times w_1^T w_1 \\ &= \frac{1}{S} \sum_{i=1}^{R} x_{i\cdot}^2 = S \sum_{i=1}^{R} \bar{x}_{i\cdot}^2 \\ &= S \sum_{i=1}^{R} (\bar{y}_{i\cdot} - \bar{y}_{\cdot\cdot})^2 = \text{SSBR}. \end{aligned} \tag{46}$$

By the proof of (45) and (46),

$$\text{SSBS} = \sum_{j=1}^{S} (d_j^{(S)} h_1^T w_j)^2, \tag{47}$$

$$\text{SSBR} = \sum_{i=1}^{R} (d_1^{(S)} h_i^T w_1)^2. \tag{48}$$

Using the fact that H_S is orthogonal again yields that

$$\sum_{j=1}^{S} (d_j^{(S)} h_i^T w_j)^2 = \sum_{j=1}^{S} (h_i^T x_j)^2 \quad \text{for all } i = 1, 2, \ldots, R. \tag{49}$$

Thus, by (38), (42), (43), (47), (48) and (49),

$$\begin{aligned} d &= \sum_{i=2}^{R}\sum_{j=2}^{S}(d_j^{(S)}h_i^T w_j)^2 \\ &= \sum_{i=1}^{R}\sum_{j=1}^{S}(d_j^{(S)}h_i^T w_j)^2 - \sum_{j=1}^{S}(d_j^{(S)}h_1^T w_j)^2 - \sum_{i=1}^{R}(d_1^{(S)}h_i^T w_1)^2 \\ &= \sum_{i=1}^{R}\sum_{j=1}^{S}(h_i^T x_j)^2 - \text{SSBS} - \text{SSBR} \\ &= \sum_{j=1}^{S}(H_R x_j)^T(H_R x_j) - \text{SSBS} - \text{SSBR} \\ &= \sum_{j=1}^{S} x_j^T x_j - \text{SSBS} - \text{SSBR} \\ &= \text{TOT} - \text{SSBS} - \text{SSBR} = \text{SSE}. \end{aligned} \tag{50}$$

Hence,

$$\Delta = \frac{\text{SSBS}}{\lambda_2} + \frac{\text{SSBR}}{\lambda_3} + \frac{\text{SSE}}{\lambda_4}, \tag{51}$$

and

$$\hat{\sigma^2} = \frac{\Delta}{RS} = \frac{1}{RS}\left(\frac{\text{SSBS}}{\lambda_2} + \frac{\text{SSBR}}{\lambda_3} + \frac{\text{SSE}}{\lambda_4}\right). \tag{52}$$

It follows that the log-likelihood function is

$$\begin{aligned} l = &-\frac{1}{2}\left[RS\ln(2\pi) + RS\ln\sigma^2\right. \\ &+ \ln\lambda_1 + (S-1)\ln\lambda_2 + (R-1)\ln\lambda_3 + (R-1)(S-1)\ln\lambda_4 \\ &\left.+\frac{1}{\sigma^2}\left(\frac{\text{SSBS}}{\lambda_2} + \frac{\text{SSBR}}{\lambda_3} + \frac{\text{SSE}}{\lambda_4}\right)\right], \end{aligned} \tag{53}$$

which implies that SSBS, SSBR and SSE are mutually independent. Furthermore, for true parameter values, the distributions of

$$\frac{\text{SSBS}}{\lambda_2\sigma^2}, \quad \frac{\text{SSBR}}{\lambda_3\sigma^2} \quad \text{and} \quad \frac{\text{SSE}}{\lambda_4\sigma^2} \tag{54}$$

are Chi-square distributions with degrees of freedom $S-1$, $R-1$ and $(R-1)(S-1)$, respectively. It is easy to verify that

$$\theta_S = \lambda_2\sigma^2, \quad \theta_R = \lambda_3\sigma^2, \quad \sigma_e^2 = \lambda_4\sigma^2. \tag{55}$$

Replacing σ^2 in (53) by its MLE in (52) yields the following log-likelihood function of ρ_s and ρ_r:

$$\begin{aligned} l &= -\frac{1}{2}\left[c_0 + D(\lambda_1, \lambda_2, \lambda_3, \lambda_4) + RS\ln\Delta\right] \\ &= -\frac{1}{2}\left[c_0 + D(\lambda_1, \lambda_2, \lambda_3, \lambda_4) + RS\ln(\frac{\text{SSBS}}{\lambda_2} + \frac{\text{SSBR}}{\lambda_3} + \frac{\text{SSE}}{\lambda_4})\right], \end{aligned} \tag{56}$$

where

$$c_0 = (1 + \ln 2\pi - \ln RS)RS, \tag{57}$$

$$D(\lambda_1, \lambda_2, \lambda_3, \lambda_4) = \ln\lambda_1 + (S-1)\ln\lambda_2 + (R-1)\ln\lambda_3 + (R-1)(S-1)\ln\lambda_4. \tag{58}$$

The constant c_0 is free of parameters, and $D = D(\lambda_1, \lambda_2, \lambda_3, \lambda_4)$ is the determinant of the matrix $\boldsymbol{V}$.

A notable fact is that, the log-likelihood function in (56) depends on only on the two parameters: ρ_s and and ρ_r.