Combining Information by Using Likelihood with Application

Myung Soon Song*

Abstract
Numerical approximations are important research areas for dealing with complicated functional forms. Techniques for developing accurate and efficient calculation of combined likelihood functions in meta-analysis are studied. A multivariate numerical integration method for developing a better approximation of the likelihood for correlation matrices is studied. Analyses for inter-correlations among Cognitive Anxiety, Somatic Anxiety and Self Confidence from Competitive State Anxiety Inventory (CSAI-2) are explored. Evaluation and Visualization of the likelihood and the MLE is conducted. Comparison with two conventional methods (joint asymptotic weighted average method & marginal asymptotic weighted average method) is shown.

Key Words: likelihood, correlation matrix, combining information, MLE, meta-analysis

1. Introduction

How to combine information about population correlation matrices from many different independent studies is one of the hot questions in meta-analysis. Conventionally, asymptotic methods are used to tackle this problem. There are many ways to combine information from the studies. Using weighted averages are probably the most common choices. But these approaches are naïve and have some evident flaws. One of the most serious of these flaws is that we have to assume that the sample size from each study is sufficiently large to justify Central Limit approximations, an assumption which is violated in many situations.

Approximation methods are suggested for computing the combined likelihoods of correlation matrices from different studies in this paper.

As a good starting point for the general multivariate case, it is quite natural to deal with the trivariate case first. The author derived a very complicated infinite series form of the trivariate likelihood function in his Ph.D. dissertation (Song, 2010)

It turns out that representing the likelihood in terms of infinite series has serious drawbacks, both because of the mathematical complexity of the representation and also due to increasing difficulty in calculation of the likelihood as the dimensionality of the parameter space increases. In general, $\frac{p(p-1)}{2}$ combined infinite series are required for a $p \times p$ correlation matrix. Even in relatively low dimensions, computation can be difficult. For example, if you have 5 variables you need to use $\frac{5(5-1)}{2} = 10$ combined infinite series for the likelihood. Consequently, it was chosen to follow Fisher's (1962) approach and represents the likelihood function as a multiple integral to which numerical integration can be applied in this paper. It is proved that just $p-1$ multiple integrals need to be computed for the general p variable case. For example, $5-1=4$ multiple integrals are needed for the 5 dimensional case. When Fisher (1962) suggested a multiple integral representation of the likelihood for the

*State University of New York, Cortland, 22 Graham Ave, Cortland, NY 13045 myung-soon.song@cortland.edu

p-dimensional case, calculation was not easy, but we can now calculate a revised multiple integral form of the likelihood easily with advanced computing methods. (Song and Gleser, 2012)

As an application of inference, intercorrelation among cognitive anxiety, somatic anxiety, and self-confidence from the area of sports psychology.

2. Approach - Likelihood as a Form of Multiple Integral

It is shown in Song and Gleser (2012) that the likelihood function of the population correlation matrix P when in a sample of size N we observe the sample correlation matrix R is

$$L(P) = f(R|n,P) = C\int_0^1 \cdots \int_0^1 \frac{\left(\prod_{i=1}^{p} v_i\right)^{\frac{n}{2}-1} \cdot \prod_{k=2}^{p-1} b_k^{k-1}}{\left(\sum_{i=1}^{p}(P^{-1})_{ii}v_i + 2\sum_{1\le i<i\le p}(P^{-1})_{ij}R_{ij}\sqrt{v_i v_j}\right)^{\frac{np}{2}}} db_{p-1}\cdots db_1 \tag{1}$$

where

$$C = \frac{|R|^{\frac{n-p-1}{2}}\,\Gamma(\frac{np}{2})}{\pi^{\frac{p(p-1)}{4}}\,|P|^{\frac{n}{2}}\prod_{i=1}^{p}\Gamma\left[\frac{1}{2}(n+1-i)\right]}\ ,$$

$$v_i = (1-b_{i-1})\,b_i\cdots b_{p-1}\ ,\ \text{for } i=1,\cdots,p-1, \tag{2}$$

$$v_p = 1-b_{p-1},\ n = N-1,$$

and where P_{ij}, R_{ij} and $(P^{-1})_{ii}$ are the elements from the i^{th} row and j^{th} column in P, R and P^{-1}(the inverse matrix of P), respectively.

You can check Song(2010) to see how to derive (1).

Fisher (1962) uses a representation of the likelihood in terms of p multiple integrals in the p variable case. Its integration domain is $(0,\infty)$ for each coordinate. But this paper suggests a representation of the likelihood with $p-1$ multiple integrals in the p variable case with integration domain $(0,1)$ for each coordinate. This representation has obvious calculational advantage over Fisher's representation.

In the next section, three main numerical integration methods, which are Gauss-Legendre quadrature, adaptive integration and Monte Carlo method used for calculation, are introduced and compared.

3. Approach - Numerical Integration

The content of this section is from the author's previous publication.(Song and Gleser, 2012) If you know the content you can skip this section.

Numerical integration is the study of how the numerical value of an integral can be obtained by using approximate computational methods. It is sometimes called quadrature. Basically, all numerical integration methods are based on adding up the value of the integrand at a sequence of points in the range of integration.

We can formulize the situation in which we have main interest:

$$\int_{S^m}\cdots\int f(x_1,\cdots,x_m)\ dx_1\cdots dx_m \approx \sum_{i=1}^{M} W_i\ f(y_{i,1},\cdots,y_{i,m}) \tag{3}$$

where R^m is a m-dimensional Euclidean space, S^m is a specified region in R^m, f: $R^m \to R$ is a common function. The vector $(y_{i,1},\cdots,y_{i,m})$ is called the *point* of the formula. The W_i is called the *coefficient* of the formula. We say that formula (3) has *degree* r(or *degree of exactness* r) if it is exact for all polynomials in $x_1,\cdots,x_m$ of degree r and there is at least one polynomial of degree $r+1$ for which it is not exact.(Evans, 1993)

At this point, we need to mention a very important aspect of the numerical integration - the assessment of error in an approximation. The only absolutely certain method is to compare the approximation with the correct answer, which is not possible in practice. Various integration methods have their special characteristics with respect to error assessment. With iterative methods, we have the natural method of examining the approximations at successive stages and stopping the iteration when the changes become small for a number of iterations. Perhaps the best way to be confident that we have accurately approximated a particular integral is to use very different methods and see if the results agree.(Evans and Swartz, 1995) In the following sections, three main methods will be used, which are Gauss-Legendre quadrature, adaptive integration and Monte Carlo method.

Among many rules in quadrature, the most commonly used rule is the Gauss-Legendre rule with cubic interval $[0,1]\times\cdots\times[0,1]$. For the multi-dimensional case (3) can be changed into:

$$\int_0^1 dx_1\int_0^1 dx_2\cdots\int_0^1 f(x_1,\cdots,x_m)\ dx_m \approx \sum_{i_1=1}^{M}\cdots\sum_{i_m=1}^{M} W_{i_1}\cdots\ W_{i_m} f(y_{i,1},\cdots,y_{i,m}) \tag{4}$$

Even if an optimal extension giving a degree of exactness of $3M+1$ can be found for Gauss-Legendre quadrature, you still do not know the accuracy in terms of correct decimal places. To get a prescribed accuracy you need adaptive integration, which keeps reducing the step size until a specified error has been achieved.

Adaptive algorithms developed by Genz and Malik (1980) operate by repeated subdivision of the hyper-rectangular region into smaller hyper-rectangles. In each subregion, the integral is estimated using a rule of degree seven, and an error estimate is obtained by comparison with a rule of degree five which uses a subset of the same points. These subdivisions are designed to dynamically concentrate the computational work in the subregions where the integrand is most irregular, and thus adapt to the behavior of the integrand. But one of the disadvantages of adaptive algorithms is their slow speed, which can be considerably overcome by using Monte Carlo methods.

Monte Carlo (MC) methods can be loosely described as statistical simulation methods. We can refer to Robert and Casella (1999) for a comprehensive introduction. The classical MC method for approximating a multiple integral such as given in the left-hand side of (3), denoted by $I(f)$, is as follows. We choose M sets of points $\{y_{i,1}, \cdots, y_{i,m}\}$, $i = 1, \cdots, M$ at random, uniformly distributed in S^m. The integral is then estimated using $W_i = V/M$ in the right-hand side of (3),

$$I(f) \approx \widehat{I}(f) = \frac{V}{M} \sum_{i=1}^{M} f(y_{i,1}, \cdots, y_{i,m}) \tag{5}$$

where $V = I(1)$ is the m-dimensional volume of S^m. The basic MC method iteratively approximates a definite integral by uniformly sampling from the domain of integration,and averaging the function values at the samples. The integrand is treated as a random variable, and the sampling scheme yields a parameter estimate of the mean of the random variable. Since $\widehat{I}(f)$ in (5) estimates $I(f)$, the absolute error in this mean can be evaluated by considering the corresponding standard error of the mean,

$$\epsilon = \left| I(f) - \widehat{I}(f) \right| \approx \frac{\sigma}{M^{1/2}} \tag{6}$$

where $\sigma^2 = V \cdot I(f^2) - I^2(f)$. If $\{y_{i,1}, \cdots, y_{i,m}\}$, $i = 1, \cdots, M$ are regarded as independent random variables then $\widehat{I}(f)$ is a random variable with mean $I(f)$ and variance σ^2/M, which can also be estimated from the random sample through

$$\frac{V}{M^2} \sum_{i=1}^{M} \left\{ f(y_{i,1}, \cdots, y_{i,m}) - \widehat{I}(f) \right\}^2 . \tag{7}$$

The absolute error (6) has an average magnitude of $O(M^{-1/2})$.(Kuonen, 2003)

In the following sections, all of three methods are used to ensure the same results for calculation.

4. Data - Anxiety

4.1 Background

An inherent aspect of competitive athletics is the need for athletes to meet the demands of competition and to perform well under pressure. Depending on how the athlete perceives the demands of competition, he or she may interpret pressure situations in a variety of ways. For example, they may be perceived as a natural part of athletic competition, or they may invoke heightened levels of stress. When in stressful and anxiety-provoking circumstances, some athletes have been observed to experience deficits in performance, even to the point of "choking." Thus, the relationship between anxiety and athletic performance has received considerable attention from researchers in the field of sport psychology.(Craft et al., 2003)

The multidimensional approach to the study of sports anxiety (Martens et al., 1990) considers subcomponents of anxiety, specifically cognitive anxiety, somatic anxiety, and self-confidence. Cognitive anxiety is the mental component of anxiety and is caused by negative expectations about success or by negative self-evaluation. Somatic anxiety refers to the physiological and affective elements of the anxiety experience that develop directly from autonomic arousal. Self-confidence is the athlete's global perceptions of confidence.

In order to assess the multidimensional aspects of anxiety, Martens et al.(1990) develop the Competitive State Anxiety Inventory-2 (CSAI-2). This 27-item measure

Study ID	Study	Size
1	Caruso et al. (1990)	$n_1 = 24$
2	Edwards and Hardy(1996)	$n_2 = 45$
3	Maynard et al. (1995)	$n_3 = 24$

Table 1: Three Studies used for the Anxiety example

has three subscales: cognitive anxiety, somatic anxiety, and self-confidence. Athletes are asked to indicate "how you feel right now" for each item on a 4-point Likert scale ranging from "not at all" to "very much so." Examples of the cognitive anxiety items include "I am concerned about this competition." and "I am concerned about choking under pressure." These items differ from the somatic anxiety statements such as "I feel nervous." or "I feel tense in my stomach." The self-confidence subscale includes items such as "I feel at ease." and "I'm confident I can meet the challenge." Each of the three subscales has 9 items, which are summed to get a score representing the level of intensity the athlete is feeling for each component of anxiety, and for self-confidence about performing.(Craft et al., 2003)

In this section, I used the data from three studies which are Caruso et al.(1990), Edwards and Hardy(1996) and Maynard et al.(1995). These studies use the CSAI-2 to calculate correlations among cognitive anxiety, somatic anxiety, and self-confidence.

When I checked Craft et al.(2003), there are 29 studies which can be used for meta-analysis, but I decided to use just three studies partly because there is a limitation for access to the specific data and partly because my goal is not to conduct rigorous meta-analysis but to show that my likelihood approach is applicable to meta-analytic application.

The correlations matrices among the three variables in the i^{th} study for $i = 1, 2,$and 3 are represented as:

$$\begin{array}{c} \\ \text{CA} \\ \text{SA} \\ \text{SC} \end{array} \quad \begin{array}{c} \begin{array}{ccc} \text{CA} & \text{SA} & \text{SC} \end{array} \\ \begin{bmatrix} 1 & r_{(i)12} & r_{(i)13} \\ r_{(i)12} & 1 & r_{(i)23} \\ r_{(i)13} & r_{(i)23} & 1 \end{bmatrix} \end{array}$$

where CA: Cognitive Anxiety, SA: Somatic Anxiety, and SC: Self-Confidence

Also, these correlation matrices can be written as a vector $\mathbf{r}_{(i)}$, the relationships represented are

$$\mathbf{r}_{(i)} = \begin{bmatrix} r_{(i)12} \\ r_{(i)13} \\ r_{(i)23} \end{bmatrix} \quad \begin{array}{l} \text{CA-SA} \\ \text{CA-SC} \\ \text{SA-SC} \end{array} \tag{8}$$

Basic information about three studies are given in Table 1.
The $\mathbf{r}_{(i)}$ vectors for the three studies are:

$$\mathbf{r}_{(1)} = \begin{bmatrix} 0.42 \\ -0.42 \\ -0.48 \end{bmatrix}, \ \mathbf{r}_{(2)} = \begin{bmatrix} 0.47 \\ -0.37 \\ -0.50 \end{bmatrix}, \ \mathbf{r}_{(3)} = \begin{bmatrix} 0.67 \\ -0.36 \\ -0.72 \end{bmatrix}. \tag{9}$$

4.2 Test of Homogeneous Correlation Matrices

Summarizing the data from several studies by a single estimated correlation matrix is conceptually sensible only if the studies have the same population correlation matrix. To determine whether the data obtained from several studies are reasonably consistent with the hypothesis of a common correlation matrix, it is useful to have a formal hypothesis test.(Becker, 1992)

This test can be formalized as testing the hypotheses:

$$\begin{cases} H_0: & P_1 = \cdots = P_k \\ H_a: & \text{At least two } P_i\text{'s are different} \end{cases} \tag{10}$$

where P_i is a $p \times p$ population correlation matrix from the i^{th} study, $i = 1, \cdots, k$. This test uses the statistic

$$Q = -2\log\left[\frac{\sup_{H_0} L(P|R)}{\sup_{H_0 \cup H_a} L(P|R)}\right] \tag{11}$$

where $L(P|R)$ is the combined likelihood from the k studies.

Because Q has approximately a χ^2 distribution with $\frac{1}{2}p(p-1)(k-1)$ degrees of freedom if H_0 is true, a test of H_0 at the 100α % level of significance is given by rejecting H_0 if Q is greater than the $100(1-\alpha)$ percentile of the χ^2 distribution with $\frac{1}{2}p(p-1)(k-1)$ degrees of freedom.

In the Anxiety example, we have $Q = 6.38$ which is less than 7.23, the 30% critical value of the χ^2 distribution with $\frac{1}{2}3(3-1)(3-1) = 6$ degrees of freedom, so we conclude that the hypothesis of homogeneity of correlation matrices cannot be rejected even at the 30% level of significance.

4.3 Comparison of Estimators

Assuming that all of the four studies share a common population correlation matrix, i.e. $P_1 = \cdots = P_4$, pooling estimates from the studies to estimate the common correlation matrix is quite natural as a next step in the meta-analysis. Becker (1992) suggests two conventional methods to get a pooled estimate of a correlation matrix. One is to separately calculate a simple weighted average of corresponding sample correlations across studies for each population correlation whereas the other is a generalized least squares approach. In this paper these estimators will be referred to as the Marginal Asymptotic Weighted Average(MAWA) and the Joint Asymptotic Weighted Average(JAWA), respectively. In this section, the MLE is compared with these conventional pooled estimates.

If we have k studies in which each study uses the same p variables, the MAWA has the form:

$$\widetilde{\boldsymbol{\rho}}_M = \left(\widetilde{\rho}_{M12}\,,\ \widetilde{\rho}_{M13}\,, \cdots,\ \widetilde{\rho}_{M1p}\,,\ \widetilde{\rho}_{M23}\,, \cdots, \widetilde{\rho}_{M(1-p)p}\right)' \tag{12}$$

where

$$\widetilde{\rho}_{Mij} = \frac{\sum_{m=1}^{k} \frac{n_m}{\left(1 - r^2_{(m)ij}\right)^2} r_{(m)ij}}{\sum_{m=1}^{k} \frac{n_m}{\left(1 - r^2_{(m)ij}\right)^2}}, \qquad i < j \tag{13}$$

MAWA	JAWA	MLE
$\widetilde{\boldsymbol{\rho}}_M = \begin{pmatrix} 0.544\ (0.035) \\ -0.381\ (0.007) \\ -0.598\ (0.046) \end{pmatrix}$	$\widetilde{\boldsymbol{\rho}}_J = \begin{pmatrix} 0.605\ (0.111) \\ -0.378\ (0.004) \\ -0.653\ (0.115) \end{pmatrix}$	$\widehat{\boldsymbol{\rho}} = \begin{pmatrix} 0.511 \\ -0.374 \\ -0.555 \end{pmatrix}$

Table 2: MAWA, JAWA and MLE from the Anxiety data

and the JAWA has the form:

$$\widetilde{\boldsymbol{\rho}}_J = \left(\sum_{m=1}^{k} \widehat{\Gamma}_m^{-1}\right)^{-1} \cdot \sum_{m=1}^{k} \widehat{\Gamma}_m^{-1} r_{(m)} = \left(\widetilde{\rho}_{J12}\ ,\ \widetilde{\rho}_{J13}\ , \cdots ,\ \widetilde{\rho}_{J1p},\ \widetilde{\rho}_{J23}, \cdots ,\ \widetilde{\rho}_{J(1-p)p}\ \right)' \tag{14}$$

where

$$\widehat{\Gamma}_m = \begin{pmatrix} \widehat{Var}(r_{(m)12}) & \cdots & \widehat{Cov}\left(r_{(m)12}, r_{(m)(p-1)p}\right) \\ \cdots & \cdots & \cdots \\ \widehat{Cov}\left(r_{(m)12}, r_{(m)(p-1)p}\right) & \cdots & \widehat{Var}(r_{(m)(p-1)p}) \end{pmatrix} : \tag{15}$$

and where

$$\begin{cases} \widehat{Var}(r_{(m)ij}) = \dfrac{\left(1 - r_{(m)ij}^2\right)^2}{n_m} \\ \\ \widehat{Cov}\left(r_{(m)ij}, r_{(m)ik}\right) \\ = \dfrac{1}{n_m}\left\{\frac{1}{2} r_{(m)ij} r_{(m)ik}\left(r_{(m)ij}^2 + r_{(m)ik}^2 + r_{(m)jk}^2 - 1\right) + r_{(m)jk}\left(1 - r_{(m)ij}^2 - r_{(m)ik}^2\right)\right\} \end{cases}$$

$$\text{for } i < j, \qquad i < k, \qquad j < k, \text{ and } m = 1, 2, 3, \text{and } 4, \tag{16}$$

and $\widehat{\Gamma}_m$ is the associated asymptotic covariance matrix for each $\mathbf{r}_{(m)}$.(Olkin and Siotani, 1976)
Similarly, the MLE matrix can be written as a vector $\widehat{\boldsymbol{\rho}}$, where

$$\widehat{\boldsymbol{\rho}} = \left(\widehat{\rho}_{12}\ ,\ \widehat{\rho}_{13}\ , \cdots ,\ \widehat{\rho}_{1p}\ ,\ \widehat{\rho}_{23}\ , \cdots ,\ \widehat{\rho}_{(1-p)p}\ \right)'. \tag{17}$$

Because all of the three studies appear to share a common population correlation matrix, i.e. $P_1 = P_2 = P_3$, it is reasonable to pool estimates from the studies to estimate the common correlation matrix.

Table 2 shows the MAWA, the JAWA, and the MLE from (9). For the MAWA and the JAWA estimators, each value inside parentheses represents the corresponding relative difference from the MLE. The components of the MAWA are not so different from their counterparts in the MLE and their relative differences are relatively small, but the JAWA is quite different from the MLE. ($\mathbf{r}_{J12} = 0.605$ and $\mathbf{r}_{J23} = -0.653$ have the relative differences of 0.111 and 0.115, respectively)

Case	Population Correlations	Sample Sizes
1	$\rho_{12} = \rho_{13} = \rho_{23} = 0.7$	$n_1 = n_2 = n_3 = n_4 = n_5 = 10$
2	$\rho_{12} = \rho_{13} = \rho_{23} = 0.7$	$n_1 = n_2 = n_3 = n_4 = n_5 = 100$
3	$\rho_{12} = 0.7,\ \rho_{13} = 0.6,\ \rho_{23} = 0.4$	$n_1 = n_2 = 10,\ n_3 = n_4 = 15,\ n_5 = 100$
4	$\rho_{12} = 0.7,\ \rho_{13} = 0.6,\ \rho_{23} = 0.4$	$n_1 = n_2 = n_3 = 10,\ n_4 = n_5 = 15$

Table 3: Four Cases Used for Simulation

4.4 Visualization of Likelihoods

We can see level plot, contour plot and likelihood plot of ρ_{12} and ρ_{13} when $\rho_{23} = -0.755$, $\rho_{13} = -0.555$ and $\rho_{23} = -0.355$, respectively, from Fig 1, Fig 2 and Fig 3, . Those ρ_{23}'s are picked because the MLE of r_{23}'s is -0.555 from Table 2 and we want to check the behavior of the "slices" of likelihood with respect to ρ_{12} and ρ_{13} when ρ_{23} changes around the MLE.

In Fig 2, the global maximum of the likelihood is achieved when we take $\rho_{12} = 0.511$, $\rho_{13} = -0.374$ with the fixed value of $\rho_{23} = -0.555$, but the local maximums are achieved when we take $\rho_{12} = 0.586$, $\rho_{13} = -0.5$ with the fixed value of $\rho_{23} = -0.755$ and when we take $\rho_{12} = 0.471$, $\rho_{13} = -0.286$ with the fixed value of $\rho_{23} = -0.355$ in Fig 1 and Fig 3, respectively. The global maximum from Fig 2 is 1184.4 which is far greater than the local maximum 4.7 from Fig 1 and the local maximum 72.2 from Fig 3.

Fig 4 is a comprehensive summary of three different cases above, which practically combine contours and corresponding likelihoods.

5. Simulation

In this section, a simple simulation is conducted to compare the accuracies of the MAWA, the JAWA and the MLE in terms of Mean Squared Error (MSE hereafter).

The MSE of $\widetilde{P}$ of with respect to P is defined as:

$$MSE(\widetilde{P}) = E\left\{tr\left[(\widetilde{P} - P)'(\widetilde{P} - P)\right]\right\} = E\left[\sum_{1 \le i < i \le p} (\widetilde{\rho}_{ij} - \rho_{ij})^2\right] \tag{18}$$

where $P = \left[\rho_{ij}\right]$ is the $p \times p$ population matrix and $\widetilde{P} = \left[\widetilde{\rho}_{ij}\right]$ is the estimator matrix of P. ($\widetilde{P}$ can be the MAWA or the JAWA or the MLE in this section.)

In the simulation, four different cases are assumed. The basic information about the cases is given in Table 3.

Each case consists of five studies with the corresponding hypothesized three-dimensional population correlation matrix and the sample sizes. For example, case 3 has the population correlations $\rho_{12} = 0.7$, $\rho_{13} = 0.6$ and $\rho_{23} = 0.4$ and the five studies have the sample sizes $10, 10, 15, 15$ and 100, respectively.

In each case, the MAWA, the JAWA and the MLE is calculated 100 times and the corresponding MSEs are also calculated.

Table 4 shows the MSEs of the MAWA, the JAWA and the MLE for each case. The MLE shows consistently the best performance whereas the JAWA shows the worst performance. Case 2 shows that the three estimators work well when the sample sizes are large enough. Case 4 shows that all of the estimators have big

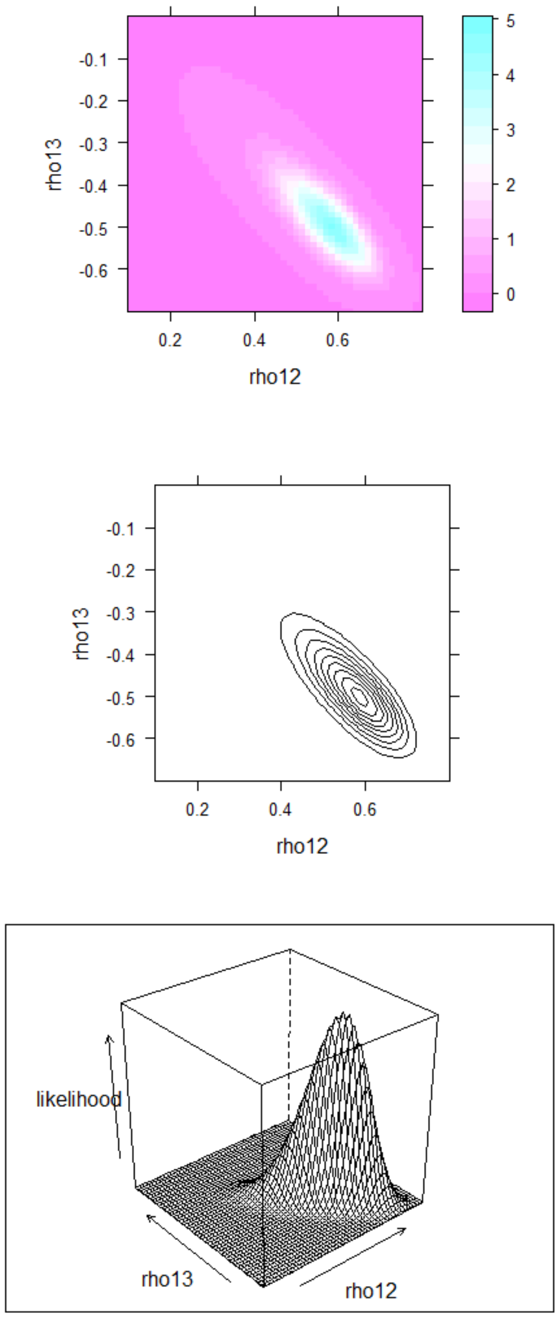

Figure 1: Level plot, contour plot and conditional likelihood I

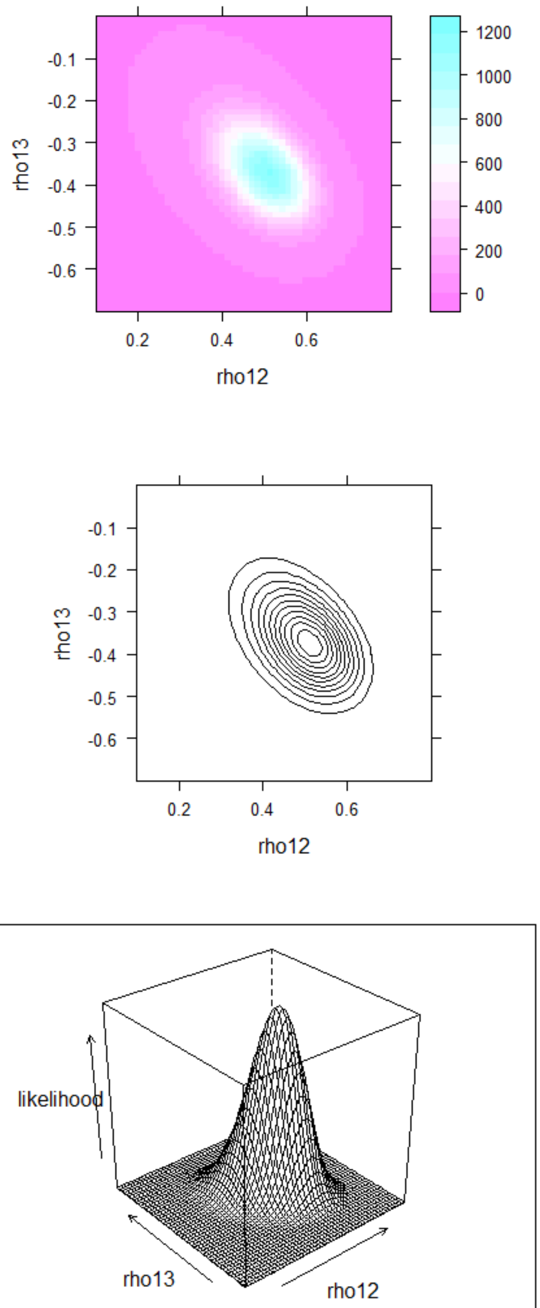

Figure 2: Level plot, contour plot and conditional likelihood II

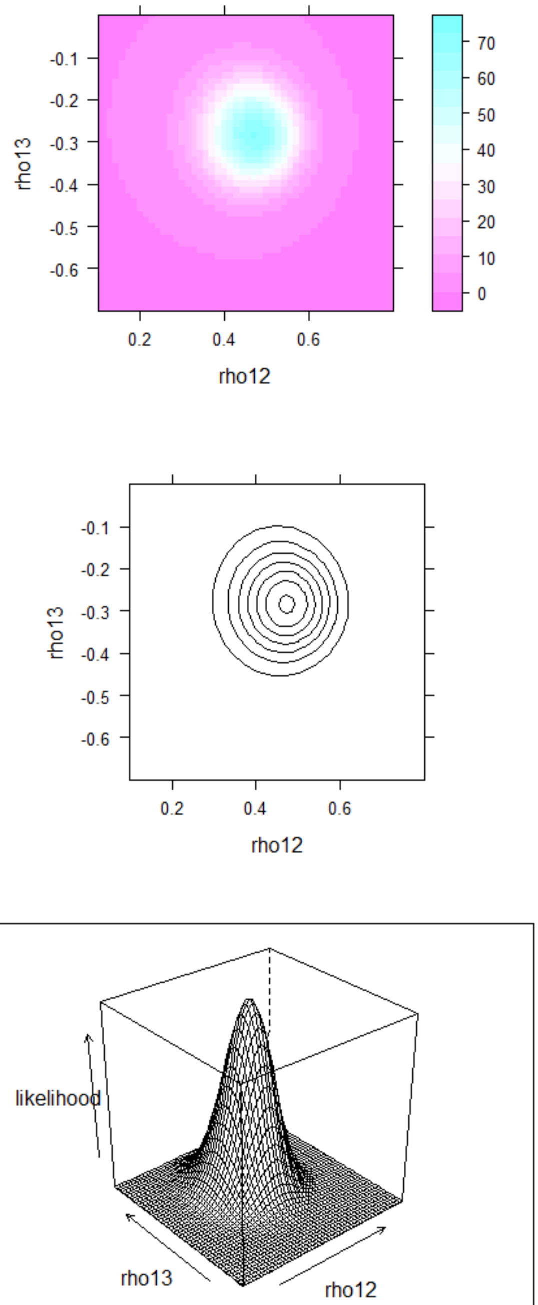

Figure 3: Level plot, contour plot and conditional likelihood III

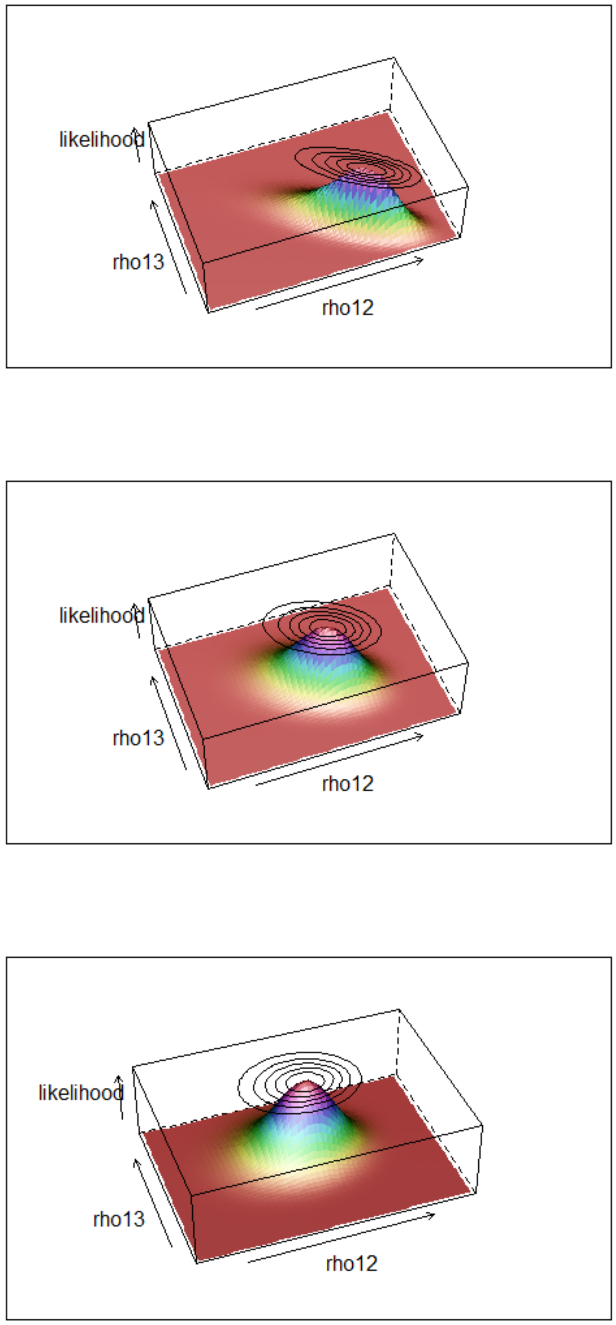

Figure 4: Three-dimensional view of Correlations and their corresponding conditional likelihoods and contours

Case	MAWA	JAWA	MLE
1	0.0487	0.0836	0.0194
2	0.0021	0.0025	0.0018
3	0.0236	0.0364	0.0111
4	0.0579	0.1026	0.0269

Table 4: The MSE's from the MAWA, the JAWA and the MLE

errors when sample sizes are small, even though the MLE is still the best among the three estimators. Interestingly enough, the MAWA is better than the JAWA for every case.

6. Discussion

In meta-analysis, asymptotic normal approximation approaches have been used to combine information about population correlation matrices from many different independent studies. But these approaches are questionable when we do not have large enough sample sizes. To overcome this problem, the likelihood approach using numerical integration to calculate the likelihood is used in this paper.

But the following are two limitations of the methods:

1) In relatively low dimensional spaces, say those for 3×3 or 4×4 correlation matrices, numerical integration works very well. But as dimension increases, the calculation becomes more difficult and cumbersome.

2) Because it is really difficult to get the first and second derivatives in a form useful for Newton's method of root- finding, numerical methods using grid points to find the MLE had to be used. Such an approach is computer intensive, especially when p is large.

Despite these limitations, the approach used in this paper is quite promising. The main ideas can be summarized as: 1) Derivation of the likelihood function for the population correlation matrix as a certain integral; 2) Calculation of the likelihood with numerical integration based on wise choices of grid points to balance precision of resolution and computational effort; and 3) Use of graphs to visualize many aspects of the likelihood function. These three steps above can be applied in many situations provided that we have way to calculate the likelihood at individual particular points.

7. Future Work

Sometimes it is impossible to get information about r_{12}, r_{13} and r_{23} from every study in the three dimentinal case, then how to combine the information from all the studies is a natural question for the next step. For example, suppose there are 10 studies having all of r_{12}, r_{13} and r_{23}, 5 studies having r_{12} and 5 studies having r_{13}, and it is desired to combine the information from these 20 studies. Under the assumption of the same population correlation matrices over all the studies, this goal can be achieved easily by multiplying the corresponding trivariate and bivariate likelihoods. Comparison between this approach and the large-sample approach described by Becker (1992) will be of interest.

Also, it is quite important to develop computing methods for the likelihood integral in the multivariate case. Among the three main methods used in multivariate case, the MC method is the fastest. But MC becomes slower as the dimension increases. Other methods need to developed to speed up getting the results

Beside the ideas and directions described above, many other good ideas will arise in the future because the topic in this paper is just at the starting point of research.

REFERENCES

Becker, B. J. (1992), "Using Results From Replicated Studies to Estimate Linear Models", *Journal of Educational and Behavioral Statistics*, 17(4), 341-362

Caruso, C. M., Dzewaltowski, D. A., Gill, D. L., and McElroy, M. A. (1990), "Psychological and physiological changes in competitive state anxiety during noncompetition and competitive success and failure", *Journal of Sport and Exercise Psychology*, 12, 6-20.

Craft, L. L., Magyar, T. M., Becker, B. J., and Feltz, D. L. (2003), "The relationship between the Competitive State Anxiety Inventory-2 and sport performance: A meta-analysis", *Journal of Sport and Exercise Psychology*, 35, 44-65.

Edwards, T. and Hardy, L. (1996), "The interactive effects of intensity and direction of cognitive and somatic anxiety and self-confidence upon performance", *Journal of Sport and Exercise Psychology*, 18, 296-312.

Evans, G. (1993), *Practical Numerical Integration*, New York: Wiley.

Eva95 : Evans, M., and Swartz, T. (1995), "Methods for approximating integrals in statistics with special emphasis on Bayesian integration problems", *Statistical Science*, 10(4), 254-272

Fisher, R. A. (1962), "The simultaneous distribution of correlation coefficients", *Sankhya: The Indian Journal of Statistics*, A, 1-8.

Genz, A., and Malik, A. A. (1980). "An adaptive algorithm for numerical integration over an N-dimensional rectangular region", *Journal of Computational and Applied Mathematics*, 6, 295-302.

Kuonen, D. (2003), "Numerical integration in S-PLUS or R: A survey", *Journal of Statistical Software*, 8, 1-14.

Martens, R., Vealey, R. S., and Burton, D. (1990), *Competitive anxiety in sport*, Human Kinetics.

May95 : Maynard, I. W., Smith, M. J., and Warwick-Evans, L. (1995), "The effects of a cognitive intervention strategy on competitive state anxiety and performance in semiprofessional soccer players", *Journal of Sport and Exercise Psychology*, 17. 428-446.

Olkin, I., and Siotani, M. (1976), "Asymptotic distribution of functions of a correlation matrix", *In S.Ideka (Ed.), Essays in Probability and Statistics*, 235-251. Tokyo: Sinko Tsusho.

Press, W.H., Teukolsky, S.A., Vetterling, W.T., and Flannery, B.P. (2001). *Numerical Recipes: The Art of Scientific Computing*, Cambridge: Cambridge University Press.

Robert, C. P., and Casella, G. (1999), *Monte Carlo Statistical Methods*, New York: Springer.

Song, M. S. (2010), “Unconventional Approach with the Likelihood of Correlation Matrices,” *Unpublished Ph.D. dissertation*, Department of Statistics, University of Pittsburgh.

Song, M. S., and Gleser, L. J. (2012), "A Numerical Likelihood-Based Approach to Combining Correlation Matrices", *Communications in Statistics-Simulation and Computation*, 41, 1679-1692.

Tanner, M. A. (1996), *Tools for statistical inference : methods for the exploration of posterior distributions and likelihood functions (3rd ed.)*, New York: Springer.