

順序カテゴリ項目の真値相関係数

True Scale Correlations for Ordinal Categorical Items

岡 本 安 晴

Yasuharu OKAMOTO

(日本女子大学大学院人間社会研究科 心理学専攻)

要 約

本研究において、順序カテゴリ項目尺度を用いて測られた概念の真値間の相関係数を推定するための階層的非線形モデルを提案する。連続項目に対しては、相関係数希薄化の修正式はよく知られており、観測スコアの相関係数が概念の真値の相関係数より小さいときに参照される。しかし、順序カテゴリ項目に対しては、この修正式は成り立たない。本研究において、順序カテゴリ項目に対して提案されたモデルの直接ベイズ法が用いられ、人工データの分析において成功したことが示された。本提案方法においては、平行項目の仮定は必要ない。

[Abstract]

This study proposes hierarchical nonlinear models to estimate the correlation coefficients for the true values of the concepts, each of which is measured using an ordinal categorical item scale. For continuous items, the correction formula for correlation attenuation is well known and refers to when the correlation coefficient of the observed scores is smaller than that of the true concept values. However, for ordinal categorical items, this correction formula does not hold. In this study, direct Bayesian methods were employed in the proposed models for ordinal categorical items and were successfully applied to hypothetical data sets. The proposed methods do not need a parallel item assumption.

1. Introduction

Indices for the relationship between variables are correlation coefficients, early contributors to which were F. Galton and K. Pearson (Stigler, 1986). This study presents a Bayesian method to directly estimate concept correlation coefficients, in which psychological scales are used to measure ordinal categorical items. However, the estimation of scale correlation coefficients for ordered categorical items has a number of difficulties, which are discussed below.

A correlation coefficient calculated using observed scale values underestimates the true value of the concept correlation coefficients; known as a correlation attenuation. For continuous scales, the correction formula is well known (Lord & Novick, 1968; McDonald, 1999; Raykov & Marcoulides, 2011; Spearman, 1904):

$$\rho_{T_X T_Y} = \rho_{XY} / \sqrt{\rho_X \rho_Y}, \quad (1)$$

where ρ_{XY} is a correlation coefficient for the X and Y scales; ρ_X and ρ_Y are the X and Y scale reliability coefficients for the scales and $\rho_{T_X T_Y}$ is the correlation coefficient for the true values of the X and Y scales. Equation 1 is based on classical error theory; which states that the observed scores for X_{ob} and Y_{ob} are represented as sums of the true values T_X and T_Y and errors E_X and E_Y (McDonald, 1999; Raykov & Marcoulides, 2011), from which the following equations were derived:

$$X_{ob} = T_X + E_X, \text{ and } Y_{ob} = T_Y + E_Y.$$

Consider the case where both reliability coefficients are of the same value and represented by coefficient alpha α , that is, $\rho_X = \rho_Y = \alpha$. The common value of the reliability coefficients, so in this case α , is equal to the attenuation value, as in the following equation:

$$\frac{\rho_{XY}}{\rho_{T_X T_Y}} = \sqrt{\rho_X \rho_Y} = \alpha. \quad (2)$$

Equation 1 is based on the assumption that scale item scores are represented by continuous variables. In psychology, scales are usually composed of ordinal categorical items; therefore, as the continuous variable values must be categorized to respond to the items, Equation 1 does not hold. Okamoto (2016b) reported on a simulation study, in which the number of items in each scale was five and the number of categories in each item was two. The results showed that the attenuation indicated by $\rho_{XY}/\rho_{T_X T_Y}$ was 0.804, while the reliability coefficient alpha was $\alpha = 0.925$. Reflecting on Equation 2, the difference between 0.804 and 0.925 cannot be ignored because the sample size was 1,000,000 and therefore large enough to obtain stable values. This discrepancy indicated that there was a need for an ordinal categorical item model.

The categorization process can be represented as a nonlinear Thurstonian model (Green & Yang, 2009; Okamoto, 2013; Yang & Green, 2011; see also Guilford, 1954; Torgerson, 1958). Based on a nonlinear model for ordinal categorical items, the relationship between the two scales can be represented as a hierarchical model (Kruschke, 2011) composed of two nonlinear models for categorical items. With this hierarchical model, it is possible to estimate true value correlation coefficients. Details of the model and the associated Bayesian estimation method are presented in the next section (for Bayesian methods, see, e.g. Gelman, Carlin, Stern, Dunson, Vehtari, & Rubin, 2014).

Lee (2007) presented a Bayesian approach to general structural equation modelling (SEM; for SEM, see, e.g. Bollen, 1989; Jöreskog, 1974), which included a model for the categorical variables; therefore, the model in this study is a special version of Lee (id.). however, since Lee's model was general, the associated algorithm was rather complicated and the identification conditions for the parameters, which are discussed in the next section, were not shown. The models and algorithms in this study are specified for true value scale correlation coefficients for ordinal categorical items, the identification conditions for which are explicitly presented in the next section. The algorithms for item response theory (IRT; for IRT, see, e.g. Baker & Kim, 2004), which are an extension of Patz and Junker (1999), are

simpler, and the associated programs, which are available as ‘Supporting Information’, are easier to use.

2. Models

Observed categorical responses X_{is} and Y_{it} on ordinal categorical items are assumed to be measured based on the latent continuous variables U_{is} and V_{it} , where i denotes person i , and s and t denote the items s and t on the X and Y scales. The classical error model for variables U_{is} and V_{it} is;

$$U_{is} = \mu_{X_s} + \lambda_{X_s} F_i + E_{iX_s} \quad (3)$$

$$V_{it} = \mu_{Y_t} + \lambda_{Y_t} G_i + E_{iY_t} \quad (4)$$

$$E_{iX_s} \sim N(0, \psi_{X_s}^2), \quad E_{iY_t} \sim N(0, \psi_{Y_t}^2).$$

F_i and G_i are the respective common factors (i.e. true values) for person i for the concepts measured by the X and Y scales. μ_{X_s} and μ_{Y_t} are the means, and λ_{X_s} and λ_{Y_t} are the factor loadings for the one factor models 3 and 4. E_{iX_s} and E_{iY_t} are the error terms with normal distributions and are independent of the other variables.

Nonlinear models are set for ordinal categorical responses X_{is} and Y_{it} , as follows:

$$X_{is} = h_s, \text{ when } C_{h_s-1} < U_{is} \leq C_{h_s} \quad (5)$$

$$Y_{it} = k_t, \text{ when } D_{k_t-1} < V_{it} \leq D_{k_t} \quad (6)$$

where C_{h_s} and D_{k_t} are the category boundaries for the latent variables U_{is} and V_{it} , which are given by models 3 and 4. The observed categorical responses X_{is} and Y_{it} are represented by the integer values h_s and k_t , where $1 \leq h_s \leq K_X$ and $1 \leq k_t \leq K_Y$. K_X and K_Y are the number of categories in the X and Y scales. It is therefore assumed that

$$C_0 = D_0 = -\infty, \text{ and } C_{K_X} = D_{K_Y} = +\infty.$$

To represent the relationship between the concepts measured by the X and Y scales, the following hierarchical model is set.

$$(F_i, G_i) \sim N(\mathbf{0}, \Sigma) \quad (7)$$

$$\Sigma = \begin{pmatrix} 1 & \rho \\ \rho & 1 \end{pmatrix}, \quad \rho \sim u(-1, 1)$$

where $N(\mathbf{0}, \Sigma)$ and $u(-1, 1)$ denote a normal distribution with a mean vector $\mathbf{0}$ and a covariance matrix Σ , and a uniform distribution on an interval $(-1, 1)$, respectively.

Nonlinear models 5 and 6 are connected by Model 7 as a hierarchical (Kruschke, 2011) or multilevel (McElreath, 2016) nonlinear model.

The model parameters are organized as follows:

$$\begin{aligned}\mathbf{C} &= (C_0, \dots, C_{K_X}), \quad \mathbf{D} = (D_0, \dots, D_{K_Y}), \\ \boldsymbol{\mu} &= (\mu_{X1}, \dots, \mu_{Xm_X}, \mu_{Y1}, \dots, \mu_{Ym_Y}), \quad \boldsymbol{\lambda} = (\lambda_{X1}, \dots, \lambda_{Xm_X}, \lambda_{Y1}, \dots, \lambda_{Ym_Y}), \\ \boldsymbol{\psi} &= (\psi_{X1}, \dots, \psi_{Xm_X}, \psi_{Y1}, \dots, \psi_{Ym_Y}), \\ \boldsymbol{\theta} &= (\rho, \mathbf{C}, \mathbf{D}, \boldsymbol{\mu}, \boldsymbol{\lambda}, \boldsymbol{\psi}),\end{aligned}$$

where m_X and m_Y are respectively the numbers of items on the X and Y scales.

The probability of the response $X_{is} = h_s$ is conditional on parameters F_i and $\boldsymbol{\theta}$ and is given by the following equation:

$$\begin{aligned}P(X_{is} = h_s | F_i, \boldsymbol{\theta}) &= P(C_{h_s-1} < U_{is} \leq C_{h_s} | F_i, \boldsymbol{\theta}) \\ &= P(C_{h_s-1} < \mu_{Xs} + \lambda_{Xs}F_i + E_{iX_s} \leq C_{h_s} | F_i, \boldsymbol{\theta}) \\ &= P(C_{h_s-1} - (\mu_{Xs} + \lambda_{Xs}F_i) < E_{iX_s} \leq C_{h_s} - (\mu_{Xs} + \lambda_{Xs}F_i) | F_i, \boldsymbol{\theta}) \\ &= \Phi\left(\frac{C_{h_s} - (\mu_{Xs} + \lambda_{Xs}F_i)}{\psi_{Xs}}\right) - \Phi\left(\frac{C_{h_s-1} - (\mu_{Xs} + \lambda_{Xs}F_i)}{\psi_{Xs}}\right),\end{aligned}\tag{8}$$

where $\Phi(z)$ is the cumulative function for the standard normal distribution with conventions

$$\Phi(-\infty) = 0 \text{ and } \Phi(+\infty) = 1.$$

Likewise, the probability of response $Y_{it} = k_t$ is conditional on parameters G_i and $\boldsymbol{\theta}$, as follows:

$$P(Y_{it} = k_t | G_i, \boldsymbol{\theta}) = \Phi\left(\frac{D_{k_t} - (\mu_{Yt} + \lambda_{Yt}G_i)}{\psi_{Yt}}\right) - \Phi\left(\frac{D_{k_t-1} - (\mu_{Yt} + \lambda_{Yt}G_i)}{\psi_{Yt}}\right)\tag{9}$$

The set of responses on the X and Y scales for person i are denoted by the following vectors:

$$\mathbf{X}_i = (X_{i1}, \dots, X_{im_X}), \quad \mathbf{Y}_i = (Y_{i1}, \dots, Y_{im_Y})$$

The following can then be set under the independence assumption

$$P(\mathbf{X}_i, \mathbf{Y}_i | F_i, G_i, \boldsymbol{\theta}) = \left\{ \prod_{s=1}^{m_X} P(X_{is} = h_s | F_i, \boldsymbol{\theta}) \right\} \left\{ \prod_{t=1}^{m_Y} P(Y_{it} = k_t | G_i, \boldsymbol{\theta}) \right\}$$

\mathbf{X}_i , \mathbf{Y}_i , F_i , and G_i are organized as follows:

$$\mathbf{X} = \begin{pmatrix} X_1 \\ \vdots \\ X_N \end{pmatrix}, \quad \mathbf{Y} = \begin{pmatrix} Y_1 \\ \vdots \\ Y_N \end{pmatrix}, \quad \mathbf{F} = \begin{pmatrix} F_1 \\ \vdots \\ F_N \end{pmatrix}, \quad \mathbf{G} = \begin{pmatrix} G_1 \\ \vdots \\ G_N \end{pmatrix},$$

The following is set under the independence assumption

$$P(\mathbf{X}, \mathbf{Y} | \mathbf{F}, \mathbf{G}, \boldsymbol{\theta}) = \prod_{i=1}^N P(X_i, Y_i | F_i, G_i, \boldsymbol{\theta})$$

To estimate these parameters, Bayesian methods are used. The joint probability (cf. Gelman, Carlin, Stern, Dunson, Vehtari, & Rubin, 2014) is set as follows:

$$P(\mathbf{X}, \mathbf{Y}, \mathbf{F}, \mathbf{G}, \boldsymbol{\theta}) = P_0(\mathbf{F}, \mathbf{G}, \boldsymbol{\theta}) P(\mathbf{X}, \mathbf{Y} | \mathbf{F}, \mathbf{G}, \boldsymbol{\theta}),$$

where $P_0(\mathbf{F}, \mathbf{G}, \boldsymbol{\theta})$ is a prior distribution. The form of P_0 is given in section 3. Algorithms.

The posterior probability is given by

$$P(\mathbf{F}, \mathbf{G}, \boldsymbol{\theta} | \mathbf{X}, \mathbf{Y}) = \frac{P(\mathbf{X}, \mathbf{Y}, \mathbf{F}, \mathbf{G}, \boldsymbol{\theta})}{P(\mathbf{X}, \mathbf{Y})} \propto P_0(\mathbf{F}, \mathbf{G}, \boldsymbol{\theta}) P(\mathbf{X}, \mathbf{Y} | \mathbf{F}, \mathbf{G}, \boldsymbol{\theta}), \quad (10)$$

as $P(\mathbf{X}, \mathbf{Y})$ is a constant under the condition that data \mathbf{X} and \mathbf{Y} are given and fixed. Equation 10 directly provides posterior distributions for \mathbf{F} , \mathbf{G} and $\boldsymbol{\theta}$ (Okamoto, 2013; Patz & Junker, 1999). Other methods, which employ structural equation modelling (SEM) approaches for categorical items, first estimate the correlation coefficients using polychoric correlations and then estimate the model parameters (Green & Yang, 2009).

2.1. Identification of parameter values

The basic model probabilities are given by Equations 8 and 9. The equation forms indicate that the origins and units are arbitrary because the arguments for $\boldsymbol{\Phi}$ consist of the ratios of the differences between the category boundaries and the latent values to the standard deviations. To simplify the algorithms, the origins and units are set according to the number of the categories as follows.

2.1.1. Items with more than two categories

In this case, origins and units are set so that

$$C_1 = -1 \text{ and } C_{K_X-1} = 1 \quad \text{when } K_X > 2 \quad (11)$$

$$D_1 = -1 \text{ and } D_{K_Y-1} = 1 \quad \text{when } K_Y > 2 \quad (12)$$

2.1.2. Binary items

For binary items, Equation 11 or 12 cannot be used, because $C_1 = C_{K_X-1}$ when $K_X = 2$, or $D_1 = D_{K_Y-1}$ when $K_Y = 2$.

Since origins are arbitrary with respect to Equations 8 and 9, we set

$$C_1 = 0 \text{ when } K_X = 2, \quad (13)$$

and

$$D_1 = 0 \text{ when } K_Y = 2. \quad (14)$$

In this case, basic probabilities are given using the following equations, which correspond to Equations 8 and 9:

$$\begin{aligned} P(X_{is} = 1|F_i, \boldsymbol{\theta}) &= \Phi\left(\frac{C_1 - (\mu_{Xs} + \lambda_{Xs}F_i)}{\psi_{Xs}}\right) - \Phi\left(\frac{C_0 - (\mu_{Xs} + \lambda_{Xs}F_i)}{\psi_{Xs}}\right) \\ &= \Phi\left(-\left(\frac{\mu_{Xs}}{\psi_{Xs}} + \frac{\lambda_{Xs}}{\psi_{Xs}}F_i\right)\right) - \Phi(-\infty) \\ &= \Phi\left(-\left(\frac{\mu_{Xs}}{\psi_{Xs}} + \frac{\lambda_{Xs}}{\psi_{Xs}}F_i\right)\right) \end{aligned} \quad (15)$$

or

$$\begin{aligned} P(Y_{it} = 1|G_i, \boldsymbol{\theta}) &= \Phi\left(\frac{D_1 - (\mu_{Yt} + \lambda_{Yt}G_i)}{\psi_{Yt}}\right) - \Phi\left(\frac{C_0 - (\mu_{Yt} + \lambda_{Yt}G_i)}{\psi_{Yt}}\right) \\ &= \Phi\left(-\left(\frac{\mu_{Yt}}{\psi_{Yt}} + \frac{\lambda_{Yt}}{\psi_{Yt}}G_i\right)\right) - \Phi(-\infty) \\ &= \Phi\left(-\left(\frac{\mu_{Yt}}{\psi_{Yt}} + \frac{\lambda_{Yt}}{\psi_{Yt}}G_i\right)\right) \end{aligned} \quad (16)$$

For Equations 15 and 16, units can be chosen so that $\psi_{Xs} = 1$ and $\psi_{Yt} = 1$ for the binary items, or the parameters can be replaced as follows:

$$P(X_{is} = 1|F_i, \boldsymbol{\theta}) = \Phi\left(-\left(\frac{\mu_{Xs}}{\psi_{Xs}} + \frac{\lambda_{Xs}}{\psi_{Xs}}F_i\right)\right) = \Phi(-(\mu_{BXs} + \lambda_{BXs}F_i)), \quad (17)$$

and

$$P(Y_{it} = 1|G_i, \boldsymbol{\theta}) = \Phi\left(-\left(\frac{\mu_{Yt}}{\psi_{Yt}} + \frac{\lambda_{Yt}}{\psi_{Yt}}G_i\right)\right) = \Phi(-(\mu_{BYt} + \lambda_{BYt}G_i)) \quad (18)$$

where

$$\mu_{BXs} = \frac{\mu_{Xs}}{\psi_{Xs}}, \lambda_{BXs} = \frac{\lambda_{Xs}}{\psi_{Xs}}, \mu_{BYt} = \frac{\mu_{Yt}}{\psi_{Yt}} \text{ and } \lambda_{BYt} = \frac{\lambda_{Yt}}{\psi_{Yt}}.$$

For the binary items, we have,

$$P(X_{is} = 2|F_i, \boldsymbol{\theta}) = 1 - P(X_{is} = 1|F_i, \boldsymbol{\theta})$$

and

$$P(Y_{it} = 2|G_i, \boldsymbol{\theta}) = 1 - P(Y_{it} = 1|G_i, \boldsymbol{\theta}).$$

3. Algorithms

In this study, the prior distributions are as follows:

$$P_0(\mathbf{F}, \mathbf{G}, \boldsymbol{\theta}) = P(\mathbf{F}, \mathbf{G}|\boldsymbol{\theta})P_0(\boldsymbol{\theta})$$

Prior distribution $P(\mathbf{F}, \mathbf{G}|\boldsymbol{\theta})$ is set to construct hierarchical Model 7, and is given as follows:

$$P(\mathbf{F}, \mathbf{G}|\boldsymbol{\theta}) = \prod_{i=1}^N P(F_i, G_i|\boldsymbol{\theta}),$$

where

$$P(F_i, G_i|\boldsymbol{\theta}) = \phi(F_i, G_i|\mathbf{0}, \rho) = \frac{1}{2\pi\sqrt{|\Sigma|}} \exp\left(-\frac{1}{2}\mathbf{f}_i'\Sigma^{-1}\mathbf{f}_i\right),$$

$$\mathbf{f}_i = \begin{pmatrix} F_i \\ G_i \end{pmatrix}, \quad \Sigma = \begin{pmatrix} 1 & \rho \\ \rho & 1 \end{pmatrix}.$$

The function $\phi(F_i, G_i|\mathbf{0}, \rho)$ denotes a density function for a two-dimensional normal distribution.

The prior distribution $P_0(\boldsymbol{\theta})$ is given by

$$P_0(\boldsymbol{\theta}) = P_0(\rho)P_0(\mathbf{C})P_0(\mathbf{D})P_0(\boldsymbol{\mu})P_0(\boldsymbol{\lambda})P_0(\boldsymbol{\psi}).$$

Bounded uniform distributions are employed for the prior distributions $P_0(\rho), \dots, P_0(\boldsymbol{\psi})$ on the right side. For nonlinear models similar to those in this study, Okamoto (2013) obtained successful results using bounded uniform distributions as the priors. Uniform distributions have been employed as priors by many authors (Alcalá-Quintana & García-Pérez, 2004; Carlin & Louis, 2009; Gelman, 2006; Kingdom & Prins, 2010).

With these prior distributions, the posterior distribution given by Equation 10 can be estimated using samples from the Markov chain Monte Carlo (MCMC) method. Specifically, the author's programs use an algorithm called Metropolis-within-Gibbs (Robert & Casella, 2010a, 2010b) or the component-wise version of the Metropolis–Hastings algorithm (Gamerman & Lopes, 2006). The algorithms for the programs in this study are extended versions of those used in Okamoto (2013), which treated the nonlinear model 5 or 6. The extensions are simple connections of the algorithms for the two nonlinear models 5 and 6 using hierarchical model 7. Details of the algorithms for the nonlinear models are explained in Okamoto (2013). The source code files for the programs in this study are available as ‘Supporting Information.’

4. Applications

Applications of the proposed models and methods for two types of hypothetical data sets are reported. One data set (Table 1) consisted of responses by 500 persons on two scales, each of which had five items with four categories. This was generated using models 5, 6, and 7 with $\mu_{Xs} = \mu_{Yt} = 0$, $\lambda_{Xs} = \lambda_{Yt} = 0.8$, $\psi_{Xs} = \psi_{Yt} = 0.6$, $C_1 = D_1 = -1$, $C_2 = D_2 = 0$, $C_3 = D_3 = 1$, and $\rho = 0.7$. The sample correlation coefficient was 0.722 for the true values F_{is} and G_{it} s generated to produce the data set.

Table 1

Hypothetical data set of 500 persons, who responded on two scales X and Y , each of which has five items with four categories. Data were generated using models 5, 6, and 7 with parameter values, $\mu_{Xs} = \mu_{Yt} = 0$, $\lambda_{Xs} = \lambda_{Yt} = 0.8$, $\psi_{Xs} = \psi_{Yt} = 0.6$, $C_1 = D_1 = -1$, $C_2 = D_2 = 0$, $C_3 = D_3 = 1$, and $\rho = 0.7$.

ID	X_{i1}	X_{i2}	X_{i3}	X_{i4}	X_{i5}	Y_{i1}	Y_{i2}	Y_{i3}	Y_{i4}	Y_{i5}
1	2	1	2	3	2	3	3	3	2	4
2	2	3	2	2	3	4	3	3	3	3
3	4	2	3	2	3	1	3	1	2	1
.										
.										
.										
498	3	3	3	2	3	2	2	1	2	3
499	2	3	2	3	2	2	3	3	4	3
500	2	3	2	2	1	3	4	2	4	4

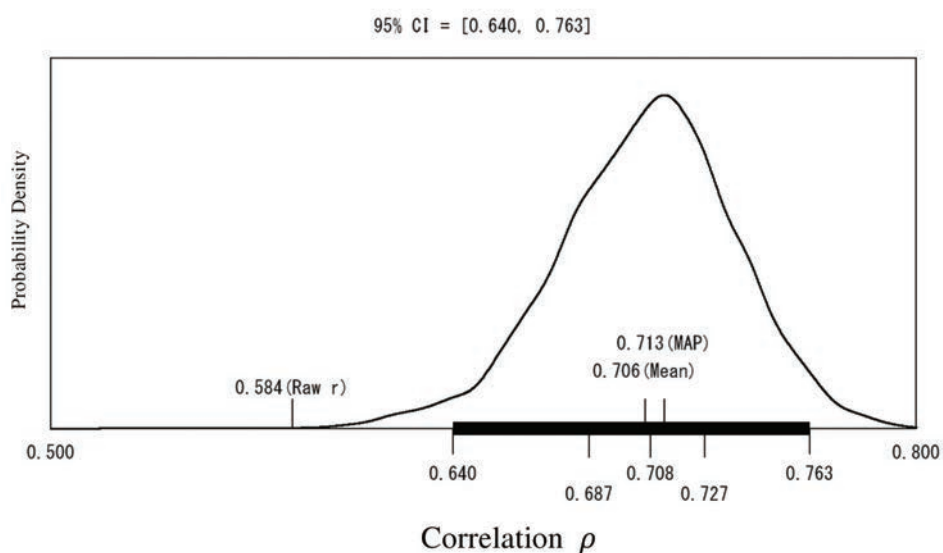


Figure 1. Posterior distribution for correlation coefficient ρ for the true values F_{is} and G_{is} for the concepts in the data set in Table 1. The maximum a posteriori (MAP) estimation is 0.713 and the mean for the posterior distribution of ρ is 0.706. The 95% confidence interval (CI) is [0.640, 0.763], which is shown as the wide line segment on the horizontal axis, on which the first quartile 0.687, the median 0.708, and the third quartile 0.727 are also shown. The sample correlation coefficient 0.584 for the observed X_{is} and Y_{is} scale values is also shown as 'Raw r' on the horizontal axis.

The posterior distribution of the correlation coefficient ρ for the true values F_{it} s and G_{it} s for the data set in Table 1 was estimated using 12,000 samples generated by the MCMC (see Figure 1). Figure 1 shows that the maximum a posteriori (MAP) estimation for ρ is 0.713 and the mean of the samples for ρ is 0.706. These values are close to the sample correlation coefficient 0.722 for the true values F_{it} s and G_{it} s used when generating the data set in Table 1. The 95% confidence interval (CI) is [0.640, 0.763], to the left of which the sample correlation coefficient $r_{XY} = 0.584$ is positioned for the X_i and Y_i sums (where $X_i = \sum_{s=1}^5 X_{is}$ and $Y_i = \sum_{t=1}^5 Y_{it}$). The correlation coefficient r_{XY} for the observed X and Y scores is significantly smaller than the correlation coefficient ρ for the true values.

The other data set (Table 2) was produced in the same way as the data set shown in Table 1, except that there were two item categories and $C_1 = D_1 = 0$. The sample correlation coefficient was 0.680 for the true values F_{it} s and G_{it} s generated to produce the data. The posterior distribution for the correlation coefficient ρ for the true values F_{it} s and G_{it} s for data set Table 2 was estimated using 12,000 samples generated using MCMC (see Figure 2). Figure 2 shows that the maximum a posteriori (MAP) estimation for ρ is 0.695 and the mean of the samples for ρ is 0.684, which are close to the sample correlation coefficient 0.680 for the true values F_{it} s and G_{it} s used when generating the data set shown in Table 2. The 95% confidence interval (CI) is [0.605, 0.755], to the left of which the sample correlation coefficient $r_{XY} = 0.527$ was positioned for the X_i and Y_i sums (where $X_i = \sum_{s=1}^5 X_{is}$ and $Y_i = \sum_{t=1}^5 Y_{it}$). The correlation coefficient r_{XY} for the observed X and Y scores is significantly smaller than the correlation coefficient ρ for the true values.

Table 2.

Hypothetical data set of 500 persons, who responded on two scales X and Y , each of which has five binary items. Data were generated using models 5, 6, and 7 with parameter values, $\mu_{Xs} = \mu_{Yt} = 0$, $\lambda_{Xs} = \lambda_{Yt} = 0.8$, $\psi_{Xs} = \psi_{Yt} = 0.6$, $C_1 = D_1 = 0$, and $\rho = 0.7$.

ID	X_{i1}	X_{i2}	X_{i3}	X_{i4}	X_{i5}	Y_{i1}	Y_{i2}	Y_{i3}	Y_{i4}	Y_{i5}
1	1	1	1	1	1	2	1	1	1	2
2	1	1	1	1	1	1	1	1	1	1
3	2	2	2	1	2	2	2	1	2	2
.										
.										
.										
498	2	1	1	1	2	2	2	2	2	2
499	1	1	1	1	1	1	1	1	1	1
500	1	2	2	2	2	1	1	1	1	1

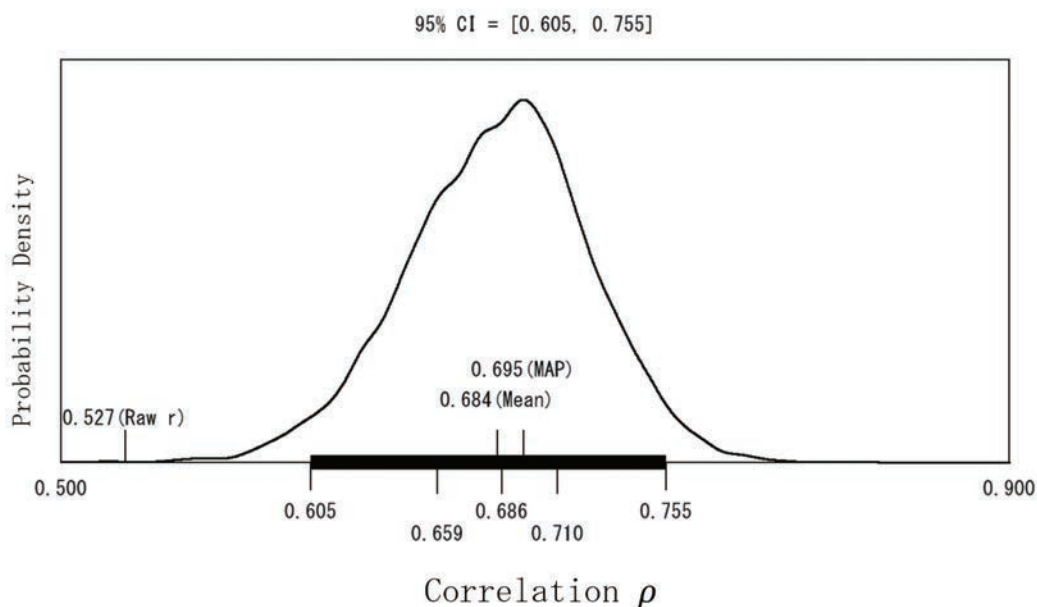


Figure 2. Posterior distribution for correlation coefficient ρ for the true values F_i s and G_i s for the concepts in the data set in Table 2. The maximum a posteriori (MAP) estimation is 0.695 and the mean for the posterior distribution of ρ is 0.684. The 95% confidence interval (CI) is [0.606, 0.755], which is shown as the wide line segment on the horizontal axis, on which the first quartile 0.659, the median 0.686, and the third quartile 0.710 are also shown. The sample correlation coefficient 0.527 for the observed X_i s and Y_i s scale values is also shown as ‘Raw r’ on the horizontal axis.

5. Discussion

The correlation coefficient of observed scores for the X and Y scales is smaller than that for the true values, which is known as attenuation. For continuous items, attenuation can be corrected using correction Formula 1.

$$\rho_{T_X T_Y} = \rho_{XY} / \sqrt{\rho_X \rho_Y} \quad (1)$$

However, correction Formula 1 does not hold for categorical items. This study demonstrated that the correlation coefficient for the true values of ordinal categorical items can be estimated using direct Bayesian methods for hierarchical nonlinear models. In general, to analyze relationships such as reliability coefficients, SEM estimation methods based on covariance matrices can also be employed (Bollen, 1989). To estimate a reliability coefficient for ordinal categorical items, SEM approaches estimate the covariances using polychoric correlations, which are correlation coefficients that are estimated for each pair of ordinal categorical items (Olsson, 1979). The model parameters can then be estimated using

weighted least squares estimation methods (Green & Yang, 2009; Yang & Green, 2011). These SEM methods, however, have some estimation problems, as noted by Yang and Green (2010), who elucidated the convergence problems in SEM analyses.

SEM estimation in the second step depends on the polychoric correlation estimations in the first step. In contrast, the direct Bayesian methods proposed in this study estimate the posterior parameter distributions that are directly dependent on or given by the data set (Equation 10). While these Bayesian methods may have posterior distribution convergence problems, these can be controlled using weakly informative priors (Gelman, Carlin, Stern, Dunson, Vehtari, & Rubin, 2014) or by regularizing the priors (McElreath, 2016). To estimate the posterior distributions for variances in the error terms in generalizability theory, Okamoto (2016a) successfully used inverse gamma distributions as weakly informative priors. Fortunately, the Bayesian methods proposed in this study do not need any weakly informative or regularizing priors.

For binary items, Okamoto (2013) assumed that the items were parallel to estimate the reliability of the omega coefficient. To estimate the correlation coefficient for the concepts, however, an assumption that the items are parallel is not needed, as shown in Equations 15 and 16 and 17 and 18. Using direct Bayesian methods, it is possible to estimate a correlation coefficient for the concepts measured by items that may not be parallel.

References

- Alcalá-Quintana, R., & García-Pérez, M. A. (2004). The role of parametric assumptions in adaptive Bayesian estimation. *Psychological Methods*, 9, 250-271.
- Baker, F. B., & Kim, S.-H. (2004). *Item response theory: Parameter estimation techniques, second edition, revised and expanded*. New York: Marcel Dekker, Inc.
- Bollen, K. A. (1989). *Structural equations with latent variables*. New York: John Wiley & Sons.
- Gamerman, D., & Lopes, H. F. (2006). *Markov chain Monte Carlo: Stochastic simulation for Bayesian inference, 2nd ed.* New York: Chapman & Hall/CRC.
- Carlin, B. P., & Louis, T. A. (2009). *Bayesian methods for data analysis, 3rd ed.* Boca Raton: Chapman & Hall/CRC.
- Gelman, A. (2006). Prior distributions for variance parameters in hierarchical models. *Bayesian Analysis*, 1, 515-533.
- Gelman, A., Carlin, J. B., Stern, H. S., Dunson, D. B., Vehtari, A., & Rubin, D. B. (2014). *Bayesian data analysis, third edition*. Boca Raton: CRC Press.
- Green, S. B., & Yang, Y. (2009). Reliability of summed item scores using structural equation modeling: An alternative to coefficient alpha. *Psychometrika*, 74, 155-167.
- Guilford, J.P. (1954). *Psychometric methods, second edition*. New York: McGraw-Hill Book Company, Inc.
- Jöreskog, K. G. (1974). Analyzing psychological data by structural analysis of covariance matrices. In D. H. Krantz, R. D. Luce, R. C. Atkinson, & P. Suppes (Eds.), *Contemporary developments in mathematical psychology II: Measurement, psychophysics, and neural information processing* (pp. 1-56). San Francisco: W. H. Freeman and company.

- Kingdom, F. A. A., & Prins, N. (2010). *Psychophysics: A practical introduction*. San Diego: Academic Press.
- Kruschke, J. K. (2011). *Doing Bayesian data analysis*. Burlington: Academic Press.
- Lee, S.-Y. (2007). *Structural equation modeling: A Bayesian approach*. West Sussex: John Wiley & Sons Ltd.
- Lord, F. M., & Novick, M. R. (1968). *Statistical theories of mental test scores*. Reading: Addison-Wesley Publishing Company.
- McDonald, R. P. (1999). *Test theory: A unified treatment*. Mahwah: Lawrence Erlbaum Associates, Publishers.
- McElreath, R. (2016). *Statistical rethinking: A Bayesian course with examples in R and Stan*. Boca Raton: CRC Press.
- Okamoto, Y. (2013). A direct Bayesian estimation of reliability. *Behaviormetrika*, 40, 149-168.
- Okamoto, Y. (2016a). Hierarchical modeling for Bayesian approach in generalizability theory. *Japan Women's University Journal: The Graduate School of Integrated Arts and Social Sciences*, 22, 153-169.
- Okamoto, Y. (2016b). On the correction for attenuation of correlations of categorical items. Proceedings of the 14th Annual Conference of The Japan Association for Research on Testing, 112-115. (In Japanese, translated by the author)
- Olsson, U. (1979). Maximum likelihood estimation of the polychoric correlation coefficient. *Psychometrika*, 44, 443-460.
- Patz, R. J., & Junker, B. W. (1999). A straightforward approach to Markov chain Monte Carlo methods for item response models. *Journal of Education and Behavioral Statistics*, 24, 146-178.
- Raykov, T., & Marcoulides, G. A. (2011). *Introduction to psychometric theory*. New York: Routledge.
- Robert, C. P., & Casella, G. (2010a). *Introducing Monte Carlo methods with R*. New York: Springer.
- Robert, C. P., & Casella, G. (2010b). *Monte Carlo statistical methods*. New York: Springer.
- Spearman, C. (1904). General intelligence: Objectively determined and measured. *The American Journal of Psychology*, 15, 201-292.
- Stigler, S. M. (1986). *The history of statistics: The measurement of uncertainty before 1900*. Cambridge: The Belknap Press of Harvard University Press.
- Torgerson, W. S. (1958). *Methods of scaling*. New York: John Wiley & Sons, Inc.
- Yang, Y. & Green, S. B. (2010). A note on structural equation modeling estimates of reliability. *Structural Equation Modeling*, 17, 66-81.
- Yang, Y., & Green, S. B. (2011). Coefficient alpha: A reliability coefficient for the 21st century? *Journal of Psychoeducational Assessment*, 29, 377-392.