Published in final edited form as:

Struct Equ Modeling. 2012 October 1; 19(4): . doi:10.1080/10705511.2012.713267.

Bayesian Evaluation of inequality-constrained Hypotheses in SEM Models using Mplus

Rens van de Schoot,

Department of Methods and Statistics, Utrecht University, The Netherlands Optentia research program, faculty of Humanities, North-West University, South Africa

Herbert Hoijtink,

Department of Methods and Statistics, Utrecht University, The Netherlands

Michael N. Hallquist, and

Department of Psychiatry, University of Pittsburgh, USA

Paul A. Boelen

Department of Child and Health Psychology, Utrecht University, The Netherlands

Abstract

Researchers in the behavioural and social sciences often have expectations that can be expressed in the form of inequality constraints among the parameters of a structural equation model resulting in an informative hypothesis. The question they would like an answer to is “Is the Hypothesis Correct” or “Is the hypothesis incorrect?”. We demonstrate a Bayesian approach to compare an inequality-constrained hypothesis with its complement in an SEM framework. The method is introduced and its utility is illustrated by means of an example. Furthermore, the influence of the specification of the prior distribution is examined. Finally, it is shown how the approach proposed can be implemented using Mplus.

Keywords

structural equation modeling; Bayes factor; Informative Hypothesis; order restricted inference; Mplus

Many researchers in the behavioural and social sciences have expectations that can be expressed in the form of inequality constraints among the parameters of a Structural Equation Model (SEM). Based on previous research, literature reviews, and the current academic debate, they formulate order restrictions among, for example, regression coefficients. For example, in health psychology research, an important distinction is the distinctiveness of symptoms of complicated grief (CG) and uncomplicated grief (UG) (Boelen & van den Bout, 2008). As we will discuss in more details later, CG symptoms should be expected to be more strongly associated with symptoms of depression and anxiety than responses of UG. The question they would like an answer to is “Is the Hypothesis Correct” or “Is the hypothesis incorrect?”. The evaluation of inequality-constrained hypotheses has been studied within the SEM framework (Gonzalez & Griffin, 2001; Stoel, Galindo-Garre, Dolan, & Van den Wittenboer, 2006; Van de Schoot, Hoijtink, & Dekovi, 2010; Van de Schoot & Strohmeier, 2011) as well as outside the SEM framework (Barlow, Bartholomew, Bremner, & Brunk, 1972; Hoijtink, Klugkist, & Boelen, 2008; Klugkist,

Laudy, & Hoijtink, 2005; Robertson, Wright, & Dykstra, 1988; Silvapulle & Sen, 2004; Van de Schoot, Hoijtink, et al., 2011; Van de Schoot, Mulder, et al., 2011). The procedures described in these studies tested an inequality-constrained hypothesis against the classical null hypothesis (i.e., “Nothing is going on”), against an unconstrained hypothesis (i.e., “Something is going on, but not the null hypothesis”), or against other inequality-constrained hypotheses (i.e., CG symptoms are expected to be more strongly associated with symptoms of depression than responses of UG, but the opposite holds for anxiety). For an overview of literature see Van de Schoot, Romeijn, and Hoijtink (2011). A comparison of an inequality-constrained hypothesis with its complement received hardly any attention in literature, but only this latter comparison would provide an answer to the question of interest: “Is the Hypothesis Correct” or “Is the hypothesis incorrect?”. One exception is the study of van Rossum, Van de Schoot, and Hoijtink (2011) (see also Hoijtink, 2012), in the context of analysis of variance. In the current paper we propose a method for the evaluation of an inequality constrained-hypothesis with respect to its complement using Bayesian statistics available in the software Mplus (Muthén & Muthén, 2007).

In what follows we first introduce Bayesian SEM, followed by the formulation of inequality-constrained hypotheses in Bayesian SEM models, and the use of Bayes factors to evaluate such hypotheses. Then we discuss specification of the prior distribution for the evaluation of inequality-constrained hypotheses and finally we apply the proposed methodology to a real-life data set from clinical and health psychology.

Bayesian SEM Models

For an introduction to Bayesian estimation see Lynch (2007), for more technical details see Gelman, Carlin, Stern, and Rubin (2004), for Bayesian model selection see Hoijtink et al. (2008) and Hoijtink (2012), for Bayesian structural equation modelling see Lee (2004). There is a variety of software packages available for Bayesian analyses: R packages (e.g., mcmc), WinBUGS, AMOS, OpenBUGS, MlwiN, Mplus v6.x. For the current paper we use Mplus (1998–2010) because of its popularity among applied researchers. Bayesian estimation first became available in Mplus in version 6, which was released in the summer of 2010. For a technical implementation of Bayesian statistics in Mplus, see Asparouhov and Muthén (2010b), Asparouhov and Muthén (2010a) and the website of Mplus: www.statmodel.com.

Let $\mathbf{y}_i = (y_1, \dots, y_p)$ be a $p \times 1$ vector of observed variables for person i , and let $\boldsymbol{\omega}_i = (\omega_1, \dots, \omega_q)$ be a $q \times 1$ vector of latent variables for person i . The measurement model is then given by

$$\mathbf{y}_i = \boldsymbol{\Lambda} \boldsymbol{\omega}_i + \boldsymbol{\varepsilon}_i, \quad (1)$$

where $\boldsymbol{\Lambda}$ is a $p \times q$ matrix of factor loadings, and $\boldsymbol{\varepsilon}$ is a $p \times 1$ vector of measurement errors with $\boldsymbol{\varepsilon} \sim N(0, \boldsymbol{\Psi}_{\boldsymbol{\varepsilon}})$, and $\boldsymbol{\Psi}_{\boldsymbol{\varepsilon}}$ being a diagonal matrix.

For the structural model, let $\boldsymbol{\omega}_i = \{\boldsymbol{\eta}_i, \boldsymbol{\xi}_i\}$ with $\boldsymbol{\eta} = (\eta_1, \dots, \eta_{q_1})$ being a $q_1 \times 1$ vector of outcomes variables and $\boldsymbol{\xi} = (\xi_1, \dots, \xi_{q_2})$ being a $q_2 (= p - q_1) \times 1$ vector of explanatory latent variables, then

$$\boldsymbol{\eta}_i = \boldsymbol{\Pi} \boldsymbol{\eta}_i + \boldsymbol{\Gamma} \boldsymbol{\xi}_i + \boldsymbol{\delta}_i, \quad (2)$$

where Π and Γ are matrices of unknown regression coefficients. Furthermore, $\delta \sim N(0, \Psi_\delta)$ and $\xi \sim N(0, \Phi)$. It is assumed that ε and δ are independent. Equations (1) and (2) describe the general SEM framework.

Let $\theta = \{\theta^1, \theta^2\}$ which consist of unknown parameters with $\theta^1 = \text{vec}\{\Pi, \Gamma\}$ and $\theta^2 = \{\Lambda, \Psi_\varepsilon, \Psi_\delta, \Phi\}$, and let Y be the observed data set with sample size n . In a Bayesian SEM model θ is considered to be random. The behavior of θ given Y in such a Bayesian model can be described by $g(\theta|Y) \propto h(\theta) \times f(Y|\theta)$, where $f(Y|\theta)$ is the likelihood function, $h(\theta)$ is the prior distribution, and $g(\theta|Y)$ is the posterior distribution.

Using the posterior distribution the mean, mode or median of the desired parameters can be estimated. Bayesian estimation uses Markov Chain Monte Carlo (MCMC) algorithms to obtain a sample from $g(\theta|Y)$. The idea behind MCMC is that the conditional distribution of one set of parameters given other sets can be used to make random draws of parameter values, ultimately resulting in an approximation of the joint distribution of all the parameters. The default MCMC sequence used in Mplus is a Gibbs sampler where t iterations ($t = 1, \dots, T$) are used to obtain new values for θ in each step drawing from a conditional posterior parameter distribution. The Gibbs sampler has three steps:

- Step 1: Sample η_i and ξ_i given $Y, \Pi, \Gamma, \Lambda, \Phi, \Psi_\varepsilon, \Psi_\delta$
- Step 2: Sample $\Pi, \Gamma, \Lambda, \Phi$ given $Y, \eta_i, \xi_i, \Psi_\varepsilon, \Psi_\delta$
- Step 3: Sample Ψ_ε and Ψ_δ given $Y, \eta_i, \xi_i, \Pi, \Gamma, \Lambda, \Phi$

These three conditional distributions have been derived in Lee (2004) and for the implementation in Mplus see Asparouhov and Muthén (2010a). In Mplus it is easy to sample from these conditional distributions. Typically, several MCMC chains are used (default in Mplus is two chains) and some part of each chain is discarded as a burn-in phase (in Mplus by default the first half). Mplus automatically monitors convergence using the Gelman-Rubin convergence criterion which considers within and between chain variability of the parameter estimates (Gelman et al., 2004; Asparouhov & Muthén, 2010b). In Mplus it is possible to first estimate the SEM model using maximum likelihood estimation and use these parameter estimates as starting values in the Bayesian analysis.

The default prior specification in Mplus is

$$h(\theta^1) = \prod_{k=1}^K h(\theta_k^1), \quad (3)$$

where $h(\theta_k^1) = \mathcal{N}(\theta_0, \tau_0^2)$ with $\theta_0 = 0$ and $\tau_0^2 = \infty$. In Mplus this prior distribution is used for means and intercepts of observed and latent continuous variables, thresholds of observed categorical dependent variables, factor loadings, and regression coefficients. For the prior distribution of θ^2 Mplus uses non-informative inverse Gamma distributions for variances and residual variances of observed and latent parameters, an inverse Wishart distribution if more than one latent variable is estimated and a Dirichlet distribution for categorical latent variable parameters. Below it will be shown that the default prior distribution used by Mplus is suited for the evaluation of inequality-constrained hypotheses by means of the Bayes factor (which will be elaborated in the next section). This is because the default prior distribution leads to an adequacy quantification of the complexity of inequality-constrained hypotheses (elaborated in the section after the next).

inequality-constrained hypotheses

In SEM, researchers have typically expectations about θ^1 and not about θ^2 . Hence, we will focus on inequality constraints imposed on the Π or Γ matrix. If M is the number of inequality constraints imposed on θ^1 , and K the number of parameters involved, then let $\mathbf{A}$ be an $M \times K$ matrix of full rank with known constants. Now, an inequality-constrained hypothesis can be specified as follows:

$$H_i: \mathbf{A}\theta^1 > \mathbf{d}, \quad (4)$$

where $\mathbf{d}$ is an $m \times 1$ vector of known constants.

Lets consider a hypothetical example and say, based on some theory, we expect the following ordering of four regression coefficients $H_i: \theta_1^1 > \theta_2^1 > \theta_3^1 > \theta_4^1$, then

$$\mathbf{A} = \begin{bmatrix} 1 & -1 & 0 & 0 & 0 \\ 0 & 1 & -1 & 0 & 0 \\ 0 & 0 & 1 & -1 & 0 \\ 0 & 0 & 0 & 1 & -1 \end{bmatrix}, \theta^1 = \begin{bmatrix} \theta_1^1 \\ \theta_2^1 \\ \theta_3^1 \\ \theta_4^1 \end{bmatrix} \text{ and } \mathbf{d} = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \end{bmatrix}, \quad (5)$$

which renders

$$H_i = \begin{bmatrix} \theta_1^1 - \theta_2^1 > 0 \\ \theta_2^1 - \theta_3^1 > 0 \\ \theta_3^1 - \theta_4^1 > 0 \end{bmatrix}, \quad (6)$$

which is equivalent to $H_i: \theta_1^1 > \theta_2^1 > \theta_3^1 > \theta_4^1$.

Another example is $H_i: \theta_2^1 - \theta_1^1 > \theta_3^1 - \theta_2^1 > 0.10$, where the difference between the first two regression coefficients is larger than the difference between the last two regression coefficients. Moreover, the difference between the coefficients is expected to be larger than 0.10. This hypothesis can be constructed using:

$$\mathbf{A} = \begin{bmatrix} -1 & 2 & -1 \\ 0 & -1 & 1 \end{bmatrix}, \theta^1 = \begin{bmatrix} \theta_1^1 \\ \theta_2^1 \\ \theta_3^1 \end{bmatrix} \text{ and } \mathbf{d} = \begin{bmatrix} 0 \\ 0.10 \end{bmatrix}. \quad (7)$$

Note that, the requirement that $\mathbf{A}$ is of full rank requirement excludes about equality constraints like $|\theta_1^1 - \theta_2^1| < d$ which could be constructed using

$$\mathbf{A} = \begin{bmatrix} 1 & -1 \\ -1 & 1 \end{bmatrix} \quad (8)$$

and

$$\mathbf{d} = [-d, -d]. \quad (9)$$

Bayes Factors for the evaluation of H_i

Typically applied researchers test the hypothesis

$$H_0: \theta^1 = 0, \quad (10)$$

or,

$$H_0: \theta_1^1 = \theta_2^1 = \dots = \theta_K^1, \quad (11)$$

for K parameters against the alternative hypothesis

$$H_1: \text{not } H_0. \quad (12)$$

We argue, however, that researchers are not really interested in such a test (Hojtink et al., 2008; Klugkist et al., 2005; Van de Schoot, Hoijtink, et al., 2011; Van de Schoot, Romeijn, & Hoijtink, 2011; Wagenmakers, 2007). Rather, they are interested in the evaluation of their expectations directly. These expectations are about the expected ordering of, for example, regression coefficients and are based on previous research or a literature study. Therefore we propose to evaluate the inequality-constrained hypothesis presented in (4) instead of the null hypotheses in (10) or (11).

The evaluation of H_i versus either H_0 or H_1 using p-values has been described by, for example Silvapulle and Sen (2004); Van de Schoot, Hoijtink, et al. (2011); Van de Schoot, Mulder, et al. (2011). However, H_i is contained in H_1 which might be considered logically inconsistent. Moreover, we argue that many applied researchers are interested in the question “Is the Hypothesis Correct” or “Is the hypothesis incorrect?”. Therefore we will explore the comparison of H_i versus its complement,

$$H_{-i}: \text{not } H_i. \quad (13)$$

As was shown by, Mulder et al. (2009) and Mulder, Hoijtink, and Klugkist (2010) the Bayes factor of H_i versus an unconstrained hypothesis, H_u , can be written as

$$BF_{iu} = \frac{f_i}{c_i}, \quad (14)$$

where f_i can be interpreted as model fit, or the proportion of the posterior distribution of H_u in agreement with the inequality-constrained hypothesis H_i . Furthermore, c_i can be interpreted as the complexity of H_i or as the proportion of the prior distribution of H_u in agreement with the constraints of H_i . Using $f_{-i} = 1 - f_i$ and $c_{-i} = 1 - c_i$,

$$BF_{-iu} = \frac{f_{-i}}{c_{-i}} = \frac{1 - f_i}{1 - c_i}. \quad (15)$$

Combining Equations (14) and (15) leads to

$$BF_{i-i} = \frac{BF_{iu}}{BF_{-iu}} = \frac{f_i/c_i}{(1-f_i)/(1-c_i)}. \quad (16)$$

The resulting Bayes factor can be interpreted as a relative measure of support for the research question “Is the Hypothesis Correct” or “Is the hypothesis incorrect?”. If $BF_{i-i} > 1$, H_i is more supported by the data than H_{-i} . If $BF_{i-i} \approx 1$, none of the two hypotheses is preferred by the data. For $BF_{i-i} < 1$, H_{-i} is more supported by the data than H_i . See Kass and Raftery (1995) for more information about the Bayes factor and its interpretation and for the interpretation in the context of inequality-constrained hypotheses Van de Schoot, Mulder, et al. (2011).

Complexity

In this section we will focus on the prior distribution, $h(\boldsymbol{\theta}^1)$ and we will show that the default prior specification, i.e. (3), that is used in Mplus is suitable for the computation of the Bayes factor (16). This is because this default specification results in an adequate quantification of the complexity of inequality-constrained hypotheses.

Hoijtink (2012) defines the complexity of inequality-constrained hypotheses using an argument based on equivalent sets of hypotheses. Consider, for example, the hypothesis $H_{iz}:\theta_1^1 > \theta_2^1 > \theta_3^1$. There are $z = 1, \dots, 3!$ permutations of the θ^1 's rendering hypotheses that have an equivalent structure. A logical implication is to consider these hypotheses to be of the same complexity. This leads to the following definition:

Definition 1

An *Equivalent Set* consist of equivalent hypotheses $H_{i1}, \dots, H_{iz}$ for which $H_{i1} \cup \dots \cup H_{iz}$ encompasses 100% of the parameter space.

With $3!=6$ equivalent hypotheses, 16.6% of the parameter space is in agreement with each hypothesis, with Z equivalent hypotheses the proportion of the parameter space in agreement with each is $1/Z$. This leads to the following definition of complexity:

Definition 2

The *Complexity of an inequality-constrained Hypothesis* is the proportion of the parameter space in agreement with H_{iz} .

Let $\mathbf{A}_m = \{A_{m1}, \dots, A_{mK}\}$ denote the m -th row from $\mathbf{A}$. The following restrictions on $\mathbf{A}$ render an hypothesis that is a member of an equivalent set:

1. Each $A_{mk} \in \{-1, 0, 1\}$.
2. For $m = 1, \dots, M$, $\sum_k A_{mk} = 0$.
3. $\mathbf{A}_1$ can be divided into D subsets of the same size, such that for $m = 2, \dots, M$, $\mathbf{A}_m$ is a permutation of these subsets.

The number of elements in an equivalent set can be determined as follows:

1. Obtain all $D!$ permutations of the D subsets in $\boldsymbol{\theta}^1$, that is divide $\theta_1^1, \dots, \theta_K^1$ in the same subsets as $A_{11}, \dots, A_{1K}$.
2. Denote the number of permutations for which $\mathbf{A}\boldsymbol{\theta}^1 > 0$ is in agreement with H_i by B .
3. Then $Z = D!/B$.

The following theorem shows that prior distributions of the form Equation (16) render a complexity measure in agreement with Definitions 1 and 2 for hypotheses that belong to an equivalent set:

Theorem 1

(for the proof see Appendix A). Let

$$h(\theta^1, \theta^2) = h(\theta^1)h(\theta^2). \quad (17)$$

For prior distributions with the property that $h(\theta_1^1, \dots, \theta_K^1) = h(\theta_{(1)}^1, \dots, \theta_{(K)}^1)$ for all permutations (1), ..., (K) of 1, ..., K, it holds that $c_i = 1/Z$ for each hypothesis in an equivalent set.

Note that the default prior used by Mplus, see (3), is in agreement with the requirements specified in Theorem 1.

Not all inequality-constrained hypotheses belong to an equivalent set. Consider, for example,

$$H_{i1}: \begin{matrix} \theta_1^1 - \theta_2^1 > \theta_3^1 - \theta_4^1 \\ \theta_1^1 > \theta_2^1 \\ \theta_3^1 > \theta_4^1. \end{matrix} \quad (18)$$

Each element of H_{i1} can be permuted in two ways, however, although all $2 \times 2 \times 2$ permutations cover 100% of the parameter space, not each permutation is equally complex. For example, 0% of the parameter space is in agreement with

$$H_{i2}: \begin{matrix} \theta_1^1 - \theta_2^1 > \theta_3^1 - \theta_4^1 \\ \theta_2^1 > \theta_1^1 \\ \theta_3^1 > \theta_4^1 \end{matrix}. \quad (19)$$

Consider also $H_{i1}: \theta_1^1 > 0, \theta_2^1 > 0$. If the ordinate (0,0) is treated as the natural midpoint of the parameter space. The permutation argument renders four hypotheses (another is $H_{i2}: \theta_1^1 < 0, \theta_2^1 < 0$) that are of equal complexity. Note that, Jeffreys (1961) (Chapters 5 and 6) and Berger and Mortera (2006) in the univariate counterpart of H_{i1} use 0 as the mean of the prior distribution and thus as the natural midpoint of the parameter space.

It can be proven that for hypotheses of the form (4), and thus also for hypotheses belonging to equivalent sets, that the complexity of an hypothesis is independent of the mean and variance of the prior distribution in Mplus.

Theorem 2

(for the proof see Appendix B) If $h(\theta_k^1) = \mathcal{N}(\theta_0, \tau_0^2)$ for $k = 1, \dots, K$ and $\mathbf{A}$ is of full rank, c_i is independent of θ_0 for $\tau_0^2 \rightarrow \infty$.

An implication of Theorem 2 for hypotheses of the form $H_i: \mathbf{A}\boldsymbol{\theta}^1 > \mathbf{d}$ can be illustrated using the hypothesis $H_i: \theta_1^1 > \theta_2^1 + 1 > \theta_3^1 + 4$. Using the prior distribution from Theorem 2, it can be seen that $c_i = 1/6$, that is, the complexity of the hypothesis $H_i: \theta_1^* > \theta_2^* > \theta_3^*$ where

$\theta_1^* = \theta_1$, $\theta_2^* = \theta_2 + 1$ and $\theta_3^* = \theta_3 + 4$. Similarly, the complexity of $H_i : \theta^l > 2$ is $1/2$, that is the complexity of the hypothesis $H_i : \theta^* > 0$, where $\theta^* = \theta - 2$. Stated otherwise, application of Theorem 2 to hypotheses of the form (4) implies that the parameter space is centered around d instead of 0.

As was illustrated above, the complexity c_i of an informative hypothesis is independent of the prior distribution for the nuisance parameters. Furthermore, since the prior distribution for the nuisance parameter is specified to be vague, the posterior distribution and consequently the fit f_i of an informative hypothesis are virtually independent of the prior distribution for the nuisance parameters.

The overall conclusion is that the Bayes factor for the comparison of an informative hypothesis with its complement is independent of the prior distribution, that is, independent of subjective decisions that have to be made by a researcher, and thus, in this sense, objective. Moreover, the default prior specification in Mplus are suitable for computation of the Bayes factor in Equation (16).

Estimation of c_i and f_i

According to Definition 2, complexity can be computed by the proportion of parameter space in agreement with the inequality-constrained hypothesis. For simple hypotheses, one can use $Z = D!/B$. However, if more complex hypotheses are used c_i can be obtained using sampling from the prior distribution $h(\theta^l)$. In this situation, c_i can be obtained using two steps for $t = 1, \dots, T$:

1. Sample θ_{tk}^1 from $\mathcal{N}(\theta_k^1 | \theta_0, \tau_0)$ for $k = 1, \dots, K$ and $t = 1, \dots, T$;
2. Estimate c_i by the proportion of θ_t^1 for $t = 1, \dots, T$ in agreement with H_i .

The fit of the model, f_i , can be estimated using a sample from the posterior distribution, $g(\theta^l / Y)$:

1. Sample θ_{tk} from the posterior distribution for $k = 1, \dots, K$;
2. Estimate f_i by the proportion of θ_t for $t = 1, \dots, T$ in agreement with H_i .

Samples from $g(\theta^l / Y)$ can be obtained using Mplus using the option ANALYSIS: ESTIMATOR = BAYES in combination with the option SAVEDATA: BPARAMETERS ARE. These statements result in a file where all the parameters of the statistical model are saved for each iteration of the Gibbs sampler after omission of the burn-in phase. The second step of the computation of f_i is to count the number of iterations where the inequality constraints are satisfied. One could do this by hand, but there is also an R package called MplusAutomation (Hallquist, 2003) that provides functions to compute these calculations based on Mplus output. For simple order restrictions c_i can be computed by hand using Theorem 1, or it can be sampled using the procedure described above. Unfortunately, it is not yet possible to sample from the prior distribution in Mplus. The final step is to compute the BF_{i-i} using Equation (16), and this can easily be done when the estimates for c_i and f_i are computed.

Precision of c_i and f_i

In this section we will evaluate the precision of the Bayes factor when evaluating inequality constraint hypotheses. As was shown in the previous section three steps have to be executed in order to compute the required BF:

- Obtain a sample of size T from the unconstrained prior and posterior distribution.

- Use this sample to determine c_i and f_i for the informative hypothesis under consideration, that is, determine the proportion of the sample from the unconstrained prior and posterior distribution in agreement with the hypotheses under consideration.
- Compute the required Bayes factors.

This procedure will work fine if the number of parameters K and the number of constraints M are relatively small. For larger values of K and M the size of T , needed to obtain accurate estimates of c_i and f_i , becomes so large that the procedure may no longer be of practical value. The question how large K and M may be such that simply sampling from the prior and posterior of the unconstrained model will render accurate estimates of c_i and f_i will now be considered.

The larger M the smaller c_i and the larger the sample that is needed in order to be able to accurately estimate c_i . Therefore the elaboration below is based on a complete ordering of the K parameters, that is $M = K - 1$ with $H_i: \theta_1^1 > \dots > \theta_K^1$. If T is large enough to accurately estimate c_i for this hypothesis, it will also be large enough to estimate c_i for informative hypotheses based on fewer inequality constraints. In Table 1 the sizes T of the sample from the prior distribution needed in order to be able to accurately estimate c_i for $K = 2, \dots, 8$ are displayed. Note that, by accurately it is meant that the difference between the true c_i and the lower bound and upper bound of a 95% central credibility interval for the estimate of c_i is less than 10%. Assuming that f_i is exactly known, this implies that $BF_{iu} = f_i/c_i$ is never off by more than 10%.

The following can be concluded from Table 1. The complexity c_i can accurately be evaluated for a complete ordering of 6 parameters using a sample of $T = 360,000$ from the prior distribution. The implication is that c_i can almost exactly be computed for any inequality-constrained hypothesis based on $K = 6$ parameters using a sample of $T = 1,000,000$ from the prior distribution. Accurate computation of complexity for a complete ordering of 7 parameters can be achieved using a sample of, 5,000,000. For more than 7 parameters accurate computation of c_i for a complete ordering may not always be possible. Note that not always a complete ordering of parameters is of interest. Consider, for example, $\theta_1^1 > \theta_2^1 > \theta_3^1 > \theta_4^1, \theta_5^1 > \theta_6^1 > \theta_7^1 > \theta_8^1$, the combination of two independent sets of constraints. The complexity of this hypothesis is $1/24 \times 1/24 = .0017$. According to Table 1 a complexity of about .0017 can accurately be estimated using $T = 360,000$ and virtually exactly using $T = 1,000,000$. As exemplified in the previous bullet, Table 1 does not only apply to hypotheses in which the parameters at hand are completely ordered. As long as there is an idea of the size of the complexity to be estimated, Table 1 gives an indication of the sample sizes needed in order to be able to obtain an accurate estimate.

Now the T rendering accurate estimates of c_i has been discussed, the implications for f_i can be elaborated. If the sample from the prior is of size T , use also a sample from the posterior of size T . If the fit of a hypothesis is really bad, irrespective of the size of T , the number of parameter vectors sampled from the posterior distribution that are in agreement with H_i will be very small. This is not a problem if one only wants to compare the hypothesis at hand to its complement. The conclusion will simply be that H_i is not supported by the data.

Real-life Example

To illustrate the Bayesian procedure of evaluating an inequality-constrained hypothesis versus its complement, we re-evaluate the data of Boelen and van den Bout (2008) (see also, Boelen, Van de Schoot, de Keijser, & van den Bout, 2010). The authors collected data to examine the distinctiveness of symptoms of complicated grief (CG) and uncomplicated grief

(UG). Symptoms of CG represent distressing and disabling responses to bereavement, including persistent yearning, feeling that life is empty, and numbness. Symptoms of UG represent relatively benign grief reactions, including missing the lost person, crying, and having emotional recollections related to the death. To examine the distinctiveness of CG and UG, Boelen and van den Bout (2008) used data from 130 mourners who all received mental help after their loss. They were recruited via grief counselors, therapists, and other caretakers.

Symptoms of CG were assessed using the Dutch version of the Inventory of Complicated Grief (ICG Boelen, van den Bout, de Keijser, & Hoijtink, 2003). This is a 30-item questionnaire tapping potentially problematic grief reactions. Respondents rate the presence of symptoms in the preceding month on 5-point scales ranging from "never" to "all the time". Items representing UG were taken from the "Present Feelings" subscale of the Texas Revised Inventory of Grief. Respondents rate the presence of grief reactions on 5-point scales ranging from "completely true" to "completely false". For the purpose of the present illustrative analyses we selected only three items of both questionnaires with the highest explained variance in the original factor analysis. Symptoms of depression were assessed using the Beck Depression Inventory. Anxiety was assessed with the state version of the State-Trait Anxiety Inventory.

There is evidence that symptoms of CG are phenomenologically and qualitatively distinguishable from reactions representing UG with symptoms of CG but not UG being associated with concomitant mental health problems (Boelen et al., 2003; Dillen, Fontaine, & Verhofstadt-Denve, 2008). This is relevant, as it suggests that CG represents a clinical condition worthy of being included in psychiatric classification systems, such as the Diagnostic and Statistical Manual of mental disorders. Notably, CG is not yet included in these systems.

We re-analysed these data, testing the prediction that, if CG symptoms are indeed more debilitating than symptoms of UG as prior research has shown, CG symptoms should be expected to be more strongly associated with symptoms of depression and anxiety than responses of UG. See Figure 1 for a graphical representation of the model where the indicators of the latent constructs are defined as categorical variables. For the Mplus syntax see Appendix C. The inequality-constrained hypothesis can be written as

$$\mathbf{A} = \begin{bmatrix} 1 & 0 & -1 & 0 \\ 0 & 1 & 0 & -1 \end{bmatrix}, \mu = \begin{bmatrix} \theta_1^1 \\ \theta_2^1 \\ \theta_3^1 \\ \theta_4^1 \end{bmatrix} \text{ and } \mathbf{0} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}, \quad (20)$$

which renders:

$$H_i = \begin{bmatrix} \theta_1^1 - \theta_3^1 > 0 \\ \theta_2^1 - \theta_4^1 > 0 \end{bmatrix}, \quad (21)$$

which is equivalent to $H_i: \theta_1^1 > \theta_3^1$ and $\theta_2^1 > \theta_4^1$.

Mplus v6.11 was used to estimate the model shown in Figure 1, using four chains, maximum likelihood estimates as starting values and a Gelman-Rubin convergence criterion of .01 (see Appendix C). The default prior distributions of Mplus were used. The posterior results are shown in Table 2 (note that for reason of space we did not include all thresholds) and the posterior distributions of $\theta_1^1 \dots \theta_4^1$ are shown in Figure 2. For description purposes it

can be interesting to inspect the credibility intervals. In the table it can be seen that the 95% credibility interval of the four parameters of interest does contain zero. Moreover, it can be seen that the correlations between the independent and between the dependent variables are substantial. It is easy to show that $c_i = .25$ because H_i consist of four equivalent hypotheses:

1. $H_1: \theta_1^1 > \theta_3^1$ and $\theta_2^1 > \theta_4^1$;
2. $H_2: \theta_1^1 < \theta_3^1$ and $\theta_2^1 > \theta_4^1$;
3. $H_3: \theta_1^1 > \theta_3^1$ and $\theta_2^1 < \theta_4^1$;
4. $H_4: \theta_1^1 < \theta_3^1$ and $\theta_2^1 < \theta_4^1$

Furthermore, for f_i the proportion of iterations of the Gibbs sampler where the parameter estimates are in line with the inequality constraints can be computed using MplusAutomation. For our model, in total 82,800 iterations were performed and 57,934 of these iterations met the constraints, which renders an estimate for f_i of .699. Using these results

$$BF_{i-i} = \frac{f_i/c_i}{(1-f_i)/(1-c_i)} = \frac{.699/.25}{(1-.699)/(1-.25)} = 6.97. \quad (22)$$

In conclusion, H_i is 6.97 times more likely than H_{-i} . This finding indicates that the evidence for complicated grief being a stronger predictor of depressive and anxious symptoms than uncomplicated is approximately 7 times greater than the converse (i.e., uncomplicated grief is a stronger correlate of negative emotional symptoms).

Conclusion

Traditional hypotheses tests are not equipped to deal with informative hypotheses formulated in terms of inequality constraints among the parameters of a structural equation model. Also, the statistical tools that are available to evaluate an inequality-constrained hypothesis test it against the classical null hypothesis, against an unconstrained hypothesis, or “Is the hypothesis incorrect?”, but tools to evaluate this hypothesis in SEM were not yet available. In the current paper we presented a solution for this problem using Mplus.

Bayesian estimation was first made available in Mplus in 2010 which makes it possible for applied researchers to switch to Bayesian statistics easily. Also, the Bayesian toolkit in Mplus allows for the computation of a Bayes factor for the comparison of an informative hypothesis against its complement. The prior specifications are an important aspect of this comparison, but the defaults in Mplus are specified in such a way that the prior distributions are specified correctly. A next step would be to include the computation of the Bayes factor in the syntax Mplus so that researchers do not have to switch to other software to compute the Bayes factor. On the other hand, the computation is very simple and the R package MplusAutomation can compute the necessary estimates.

Acknowledgments

The first author was supported by a grant from the Netherlands organization for scientific research: NWO-VENI-451-11-008. The second author was supported by a grant from the Netherlands organization for scientific research: NWO-VICI-453-05-002.

References

- Asparouhov, T.; Muthen, B. Bayesian analysis of latent variable models using mplus. 2010a. Manuscript submitted for publication, available via www.statmodel.com
- Asparouhov, T.; Muthen, B. Bayesian analysis using mplus: Technical implementation. 2010b. Manuscript submitted for publication, available via www.statmodel.com
- Barlow, RE.; Bartholomew, DJ.; Bremner, HM.; Brunk, HD. Statistical inference under order restrictions. New York: Wiley; 1972.
- Berger J, Mortera J. Bayes factors for non-nested hypothesis testing. Journal of the American Statistical Association. 2006; 94:542–554.
- Boelen PA, Van de Schoot R, de Keijser J, van den Bout J. Prolonged grief disorder, depression, and posttraumatic stress-disorder are distinguishable syndromes. Journal of Affective Disorders. 2010; 125:374–378. [PubMed: 20189657]
- Boelen PA, van den Bout J. Complicated grief and uncomplicated grief are distinguishable constructs. Psychiatry Research. 2008; 157:311–314. [PubMed: 17916387]
- Boelen PA, van den Bout J, de Keijser J, Hooijink H. Reliability and validity of the dutch version of the inventory of traumatic grief. Death Studies. 2003; 27:227–247. [PubMed: 12703504]
- Dillen L, Fontaine J, Verhofstadt-Denve L. Are normal and complicated grief different constructs? a confirmatory factor analytic test. Clinical Psychology and Psychotherapy. 2008; 15:386–395. [PubMed: 19115457]
- Gelman, A.; Carlin, JB.; Stern, HS.; Rubin, DB. Bayesian data analysis. 2. London: Chapman&Hall/CRC; 2004.
- Gonzalez R, Griffin D. Testing parameters in structural equation modelling: Every “One” matters. Psychological Methods. 2001; 6:258–269. [PubMed: 11570231]
- Hallquist, MN. Mplusautomation: Automating mplus model estimation and interpretation. an R package. (version 0.4–2). 2003. [computer software]. Available from <http://cran.r-project.org/web/packages/MplusAutomation/>
- Hooijink, H. Informative hypotheses: Theory and practice for behavioral and social scientists. London: Chapman and Hall/CRC; 2012.
- Hooijink, H.; Klugkist, I.; Boelen, PA. Bayesian evaluation of informative hypotheses. New-York: Springer; 2008.
- Jeffreys, H. Theory of probability. Oxford: Oxford University Press; 1961.
- Kass RE, Raftery AE. Bayes factors. Journal of the American Statistical Association. 1995; 90:773–795.
- Klugkist I, Laudy O, Hooijink H. Inequality constrained analysis of variance: A bayesian approach. Psychological Methods. 2005; 10:477–493. [PubMed: 16393001]
- Lee, S-Y. Structural equation modeling: A bayesian approach. West Sussex: Wiley; 2004.
- Lynch, S. Introduction to applied bayesian statistics and estimation for social scientists. New-York: Springer; 2007.
- Mulder J, Hooijink H, Klugkist I. Equality and inequality constrained multivariate linear models: Objective model selection using constrained posterior priors. Journal of Statistical Planning and Inference. 2010; 140:887–906.
- Mulder J, Klugkist I, Van de Schoot R, Meeus W, Selfhout M, Hooijink H. Bayesian model selection of informative hypotheses for repeated measurements. Journal of Mathematical Psychology. 2009; 53:530–546.
- Muthén, LK.; Muthén, BO. Mplus: Statistical analysis with latent variables: User’s guide. Los Angeles, CA: Muthén & Muth; 2007.
- Robertson, T.; Wright, FT.; Dykstra, RL. Order restricted statistical inference. New York: Wiley; 1988.
- Silvapulle, MJ.; Sen, PK. Constrained statistical inference: Order, inequality, and shape constraints. London: John Wiley Sons; 2004.

- Stoel RD, Galindo-Garre F, Dolan C, Van den Wittenboer G. On the likelihood ratio test in structural equation modeling when parameters are subject to boundary constraints. *Psychological Methods*. 2006; 4:439–455. [PubMed: 17154756]
- Van de Schoot R, Hoijtink H, Deković M. Testing inequality constrained hypotheses in sem models. *Structural Equation Modeling*. 2010; 17:443–463.
- Van de Schoot R, Hoijtink H, Mulder J, Van Aken MAG, Orobio de Castro B, Meeus W, et al. Evaluating expectations about negative emotional states of aggressive boys using bayesian model selection. *Developmental Psychology*. 2011; 47:203–212. [PubMed: 21244159]
- Van de Schoot R, Mulder J, Hoijtink H, Van Aken M, Dubas J, Orobio de Castro B, et al. Psychological functioning, personality and support from family: An introduction bayesian model selection. *European Journal of Developmental Psychology*. 2011
- Van de Schoot R, Romeijn J-W, Hoijtink H. Moving beyond traditional null hypothesis testing: Evaluating expectations directly. *Frontiers in Quantitative Psychology and Measurement*. 2011; 2:24.
- Van de Schoot R, Strohmeier D. Testing informative hypotheses in sem increases power: An illustration contrasting classical hypothesis testing with a parametric bootstrap approach. *International Journal of Behavioral Development*. 2011; 35:180–190.
- van Rossum M, Van de Schoot R, Hoijtink H. ‘is the hypothesis correct’ or ‘is it not’? bayesian evaluation of one informative hypothesis in anova. *Methodology*. 2011
- Wagenmakers EJ. A practical solution to the pervasive problems of p values. *Psychonomic Bulletin & Review*. 2007; 14:779–804. [PubMed: 18087943]

Appendix A. Proof of Theorem 1

Proof of Theorem 1

If A_1 is divided in D subsets of the same size, $\theta^1 = [\theta_1, \dots, \theta_D]$, where each subset of θ contains the same number of elements, there are $D!$ permutations of $\theta^1, \dots, \theta^{D!}$.

If the number of equivalent hypotheses is Z , $D!/Z = B$ of these permutations are in agreement with each equivalent hypothesis, that is, $[\theta^1, \dots, \theta^{D!}] = [\dots, \theta^1, \dots, \theta^B, \dots]$, leading to $H_{iz} : H_{iz1} \cup \dots \cup H_{izB}$.

Because each $\theta \in H_{iz}$ can be permuted to a $\theta' \in H_{iz'}$ for which it holds that $h(\theta) = h(\theta')$ it holds that

$$c_{iz} = \int_{\theta \in H_{iz}} h(\theta) d\theta = \int_{\theta' \in H_{iz'}} h(\theta') d\theta' = c_{iz'}. \quad (23)$$

Since $\sum_z c_{iz} = 1$ this implies that $c_{iz} = 1/Z$ for $z = 1, \dots, Z$.

Appendix B. Proof of Theorem 2

Proof of Theorem 2

Let the rows of A be denoted by $A_1, \dots, A_M$, $A_m = [A_{m1}, \dots, A_{mK}]$ and $d = [d_1, \dots, d_M]$. Then

$$c_i = P(A_1\theta - d_1 > 0, \dots, A_M\theta - d_M > 0) \left[\begin{array}{c} A_1\theta - d_1 \\ \dots \\ A_M\theta - d_M \end{array} \right] \sim \mathcal{N}(V\theta_0 - d, W\tau_0^2), \quad (24)$$

where

$$\mathbf{V} = \begin{bmatrix} \sum_k A_{1k} \\ \dots \\ \sum_k A_{Mk} \end{bmatrix}, \quad (25)$$

and

$$\mathbf{W} = \begin{bmatrix} \sum_k A_{1k}^2 & \dots & \sum_k A_{Mk} A_{1k} \\ \dots & \dots & \dots \\ \sum_k A_{1k} A_{Mk} & \dots & \sum_k A_{Mk}^2 \end{bmatrix}. \quad (26)$$

Since

$\lim_{\tau_0^2 \rightarrow \infty} P(\mathbf{Z} > \mathbf{0} | \mathbf{Z} \sim \mathcal{N}(\mathbf{V}\theta_0 - \mathbf{d}, \mathbf{W}\tau_0^2)) = P(\mathbf{Z} > \mathbf{0} | \mathbf{Z} \sim \mathcal{N}(\mathbf{0}, \mathbf{W}\tau_0^2)) = P(\mathbf{Z} > \mathbf{0} | \mathbf{Z} \sim \mathcal{N}(\mathbf{0}, \mathbf{W}))$
, where $\mathbf{Z}$ is a vector of length M containing $[\mathbf{A}_1\theta - d_1, \dots, \mathbf{A}_M\theta - d_M]$, the $\lim_{\tau_0^2 \rightarrow \infty} c_i$ reduces to

$$c_i = P(\mathbf{A}_1\theta - d_1 > 0, \dots, \mathbf{A}_M\theta - d_M > 0 | [\begin{matrix} \mathbf{A}_1\theta - d_1 \\ \dots \\ \mathbf{A}_M\theta - d_M \end{matrix}] \sim \mathcal{N}(\mathbf{0}, \mathbf{W})), \quad (27)$$

that is, independent of θ_0 .

Appendix C. Mplus syntax

```
DATA: FILE = data.dat;
VARIABLE:
NAMES ARE TRIG4 TRIG5 TRIG9 ITG4 ITG5 ITG10 STAITOT BDIM;
CATEGORICAL ARE TRIG4 TRIG5 TRIG9 ITG4 ITG5 ITG10;
MISSING ARE ALL ( -999);
ANALYSIS:
ESTIMATOR = bayes; !here you select the Bayesian estimator
CHAINS = 4; !change the number of chains
PROCESSOR= 4 (starts); !increase computation speed
BCONVERGENCE = .01; !select a smaller convergence criterium
STVALUES = ml; !use maximum likelihood estimates as starting values
MODEL:
STAITOT BDIM on CG UCG;
UCG by TRIG4 TRIG5 TRIG9;
CG BY ITG4 ITG5 ITG10;
SAVEDATA: BPARAMETERS ARE c:/Bresults2.dat; !save the parameter estimates
for each iteration after burn-in
```

Figure 1.
Path model between symptoms of complicated grief (CG), uncomplicated grief (UG), anxiety and depression.

Figure 2.

Posterior distributions with posterior mean, mode, median and 95% C.I. for $\theta_1^1 \dots \theta_4^1$.

Table 1

K versus *T*

<i>K</i>	2	3	4	5	6	7	8
True c_i	1/2 = .5	1/6 = .166	1/24 = .042	1/120 = .008	1/720 = .0014	1/5040 = .0002	1/40320 = .000025
<i>T</i>	1000	3000	9600	120,000	360,000	2,520,000	20,160,000
LB-95%	.47	.154	.038	.0078	.00127	.000182	.000023
UB-95%	.53	.180	.046	.0089	.00152	.000217	.000027

Table 2

Posterior parameter estimates with in boldface the parameters of interest. Note that we left out the information about the thresholds for reasons of space.

	Est.	SD	95% C.I.
Anxiety ON			
CG (θ_1^1)	0.560	0.343	-0.099–1.280
UCG (θ_3^1)	0.083	0.356	-0.677–0.751
Depr. ON			
CG (θ_2^1)	0.595	0.334	-0.046–1.310
UCG (θ_4^1)	0.086	0.354	-0.683–0.754
CG WITH			
UCG	0.821	0.069	0.654–0.919
Anxiety WITH			
Depr.	0.700	0.081	0.514–0.804
Intercepts			
Anxiety	3.472	0.237	3.011–3.943
Depr.	3.535	0.242	3.071–4.020
Res. Var.			
Anxiety	0.572	0.104	0.349–0.754
Depr.	0.522	0.099	0.312–0.703
Factor loadings			
λ_1	0.669	0.064	0.533–0.780
λ_2	0.830	0.066	0.679–0.939
λ_3	0.598	0.081	0.421–0.739
λ_4	0.569	0.079	0.416–0.725
λ_5	0.550	0.095	0.345–0.716
λ_6	0.679	0.099	0.456–0.840