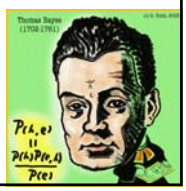
Mplus Users Group August 30, 2012

Some Examples of Bayesian Modeling in Mplus

Various examples coauthored by Joop Hox, Rens vd Schoot, Suzette Matthijssen, Seda Can



mcmc

Why Bayesian Statistics?

- Can do some things that cannot be done or are very difficult in classical statistics
- Valid in small samples
  - Maximum Likelihood is not
    - "Asymptotically we are all dead ..." (Novick)
- Always proper estimates
  - No negative variances

2

Bayesian Statistics Reaches Parts that other Statistics Do Not Reach…

- Complex models with complex constraints
  - E.g. include a model for the missingness
- Estimation including missing data
  - *Each missing data point is just another parameter to estimate…*
- Multiple Imputation (MI) of complex data
- Estimation of scores for latent variables
  - Which are *all* missing…

3

Simulating the Posterior Distribution

- Markov Chain Monte Carlo (MCMC)
- Given a draw from a specific probability distribution, MCMC produces a new pseudorandom draw from that distribution
  - …then Repeat, Repeat, Repeat…
  - Distributions typically multivariate
- Obtain point estimates and 95% credibility interval from posterior

4

MCMC Issues: Burn In

- Sequence of draws $Z^{(1)} \rightarrow Z^{(2)} \rightarrow,\ldots, \rightarrow Z^{(t)}$
  - From target distribution *f(Z)*
  - Even if $Z^{(1)}$ not from *f(Z)*, the distribution of $Z^{(t)}$ is f(Z), as $t \rightarrow \infty$
  - So, for arbitrary $Z^{(1)}$, if *t* is sufficiently large, $Z^{(t)}$ is from target distribution f(Z)
    - But having good starting values helps

$\rightarrow$ MCMC must run *t* iterations 'burn in' before we reach target distribution f(Z)

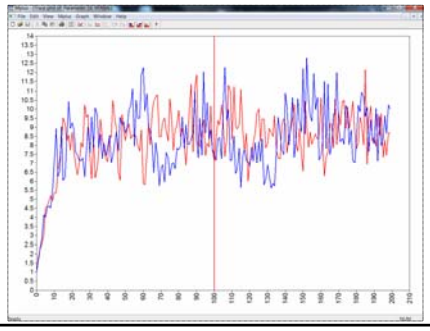
5

Assessing Convergence of MCMC Chain on Correct *Distribution*

- Burn-in: Mplus deletes first half of chain
- Run multiple chains (Mplus default 2)
  - PSR statistic compares variances within to variance between chains: must be close to 1
  - Mplus default BCONVERGENCE=0.05
- Graphical evaluation
  - Plots of chain(s), visual inspection

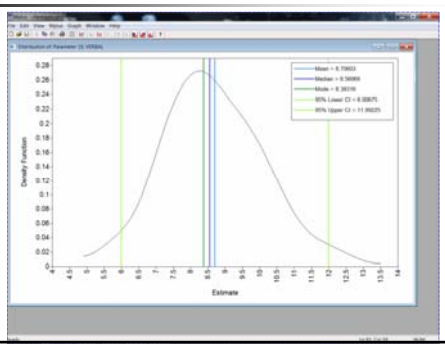
6

Trace Plot, Two Chains



7

Posterior Distribution, Kernel Plot



8

Example 1: Analyzing Comparative Surveys

- Current best analysis method:
  - multigroup SEM
- But multilevel SEM very attractive option
  - because country level variables can be included
  - Respondents nested within countries

9

Problem with Multilevel SEM: Too Few Countries

- Most comparative surveys involve a relatively low number of countries
  - SHARE ±13; GGP ±14; ESS ±30; TIMMS 36-48; PISA ±65
- How large a sample is needed for accurate estimation in SEM?
  - Boomsma (1983): 200 OK, at least 100
  - Hox, Maas Brinkhuis (2010): at least 100 groups

10

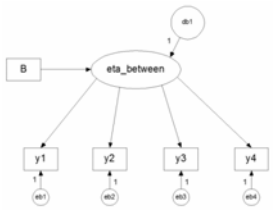
SEM with (too) Small Samples

- Typical problems with small samples
  - Nonconvergence
  - Inadmissible estimates
    - Large negative variances
  - Biased chi-squares and standard errors
  - Low power
  - Less robust against nonnormality

11

Meuleman & Billiet (2009) Monte Carlo Study

- Simulation based on typical ESS sample design, with # of countries (NC) 20-100
- Model 4 indicator CFA with extraneous observed predictor on latent factor
- ICC observed indicators 0.08
- Within country sample size ±1755



  - Conclusions (country level): serious nonconvergence only for NC=20; parameter bias small if NC≥40; CI coverage acceptable but not good if NC≥60; power low

12

Bayesian Replication of Meuleman & Billiet

- Simulation design follows M&B (2009) closely
- # of countries (NC) 10, 15, 20, 40
- Simulations in Mplus 6.1
  - Completely relying on Mplus 6.1 Bayes estimation defaults
- Priors N(0,10^{10}) for coefficients, IG (-1,0) for variances
- Run 2 chains
- Convergence when PSR close to 1 (.05)
- Point estimate is median of posterior
- 95% CI by percentile method

Results: Bias & Coverage

Table 1. Mean absolute bias for various country level sample sizes

	Number of countries					
	10	15	20	20	40	60
	Bayesian estimation			ML estimation (M&B, 2009)		
Parameter bias						
Within factor loadings	0.00	0.00	0.00	0.00	0.00	0.00
Within error variances	0.00	0.00	0.00	0.00	0.00	0.00
Within structural effect	0.00	0.00	0.00	0.00	0.00	0.00
Between factor loadings	0.50	0.04	0.03	-.07	-.03	-.02
Between error variances	0.59	0.33	0.24	-.10	-.05	-.04
Between structural effect	-.05	-.05	-.04	0.11	0.05	0.04
Coverage						
Within factor loadings	0.95	0.96	0.95	0.93	0.94	0.94
Within error variances	0.93	0.94	0.93	0.93	0.94	0.94
Within structural effect	0.95	0.95	0.96	0.93	0.94	0.94
Between factor loadings	0.96	0.96	0.94	0.84	0.89	0.91
Between error variances	0.95	0.95	0.94	0.81	0.88	0.90
Between structural effect	0.96	0.94	0.95	0.85	0.90	0.92

14

Results: Alpha Control & Power

Table 2. Statistical power for detecting the country level structural effect, for various effect sizes and country level sample sizes

	Number of countries					
	10	15	20	20	40	60
	Bayesian estimation			ML estimation (M&B, 2009)		
Effect size						
None (0.00)	0.03	0.05	0.05	0.16	0.10	0.08
Small (0.10)	0.04	0.06	0.06	0.18	0.15	0.16
Medium (0.25)	0.08	0.13	0.15	0.31	0.41	0.53
Large (0.50)	0.26	0.43	0.58	0.75	0.94	0.99
Very large (0.75)	0.67	0.89	0.97	1.00	1.00	1.00

15

Conclusions

- Bayesian estimation works well with far smaller country level sample sizes than maximum likelihood needs
- Bias is larger, but coverage is much better
- Comparative analysis with NC=20 OK
- With NC=10 problems occur
  - These are related to outliers in the posterior distributions (convergence problems), which PRS statistic does not pick up

16

Some Conclusions on Bayesian Estimation in Mplus

- Priors N(0,10^{10}) for coefficients, IG (-1,0) for variances
- Run 2 chains
- Convergence when PSR close to 1 (.05)
- Always use graphical tools to monitor burn-in, convergence & autocorrelations!

- Priors N(0,10^{10}) for coefficients, IG (-1,0) for variances
- Run ≥ 4 chains
- Convergence when PSR close to 1 (.01)
- Point estimate is median of posterior
- 95% CI by percentile method

Example 2: Small Experiments

- Same simulation setup
- Group sizes not 1750 but only 100

- Same results, accuracy and coverage comparable with accuracy and coverage given 1750 group sizes!
- Within group sizes matter very little when between groups effects are studies

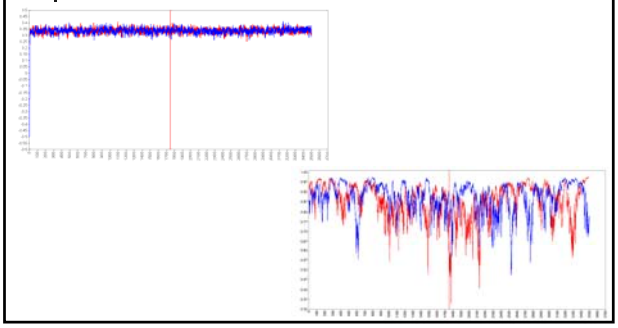
18

Example 3

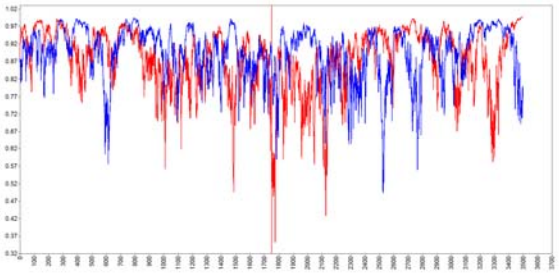
- Two-level data with a very high correlation between two factors at the between level
  - Two factors, each with 4 indicators
  - 5000 cases in 200 clusters
  - Same model within and between
  - High correlation FB1 and FB2
- ML estimation: $r_{FB1\text{-}FB2}$ = 1.11
  - Bayes estimation: $r_{FB1\text{-}FB2}$ = 0.90
- Problem solved *or is it?*

19

Example 3 trace plots $r_{FW1\text{-}FW2}$ and $r_{FB1\text{-}FB2}$

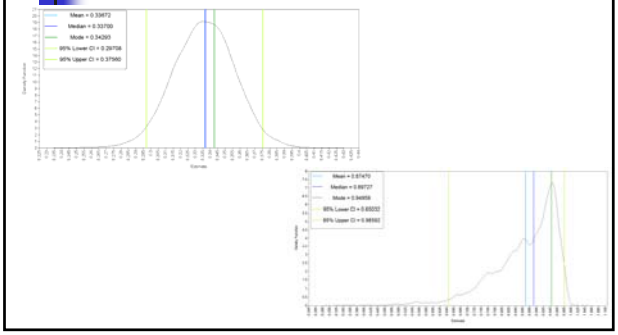


Example 3, trace plot $r_{FB1\text{-}FB2}$

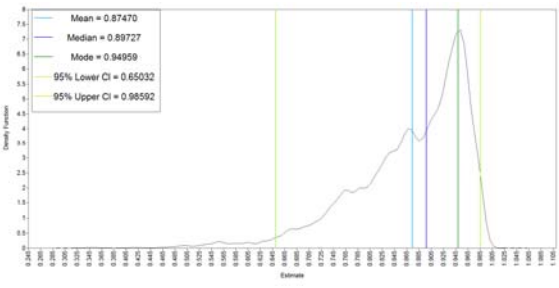


21

Example 3 posteriors $r_{FW1\text{-}FW2}$ and $r_{FB1\text{-}FB2}$

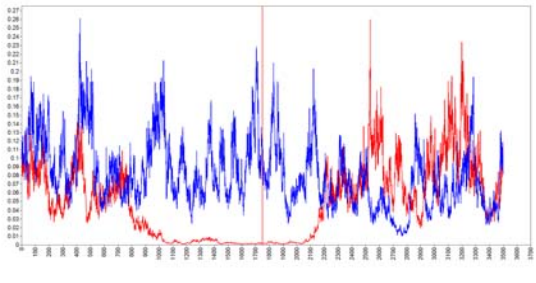


Example 3, trace plot $r_{FB1\text{-}FB2}$



23

Example 3, variance of FB2



24

Example 3, tweaking Mplus Bayes estimation settings

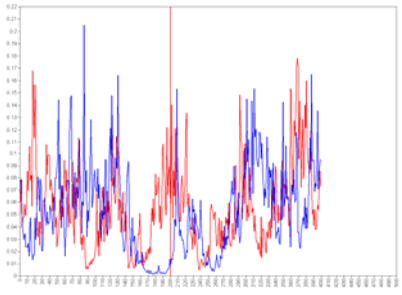
- More strict value for BCONVERGENCE
  - default 0.05→3600 iters, 0.001→45000 iters
    - convergence did not improve
- Set iterations 1000000 plus use THIN=20
  - remove autocorrelation

25

Example 3, variance of FB2

26

Example 3, variance of FB2 after thinning = 20



27

Example 3 *answers* and questions

- Answer #1: Bayes is statistics not magic, not all problems go quietly away
- Answer #2: vastly increasing iterations and thinning and # of chains did not help
- Answer #3: all between estimates are permissible values, wide 95% CI's

28

Example 3 answers and *questions*

- Question #1: accept model and estimates anyway?
- Question #2: could we ‘clip off’ (winsorize) the spikes in the posterior?
- Question #3: can we tweak the estimation procedure further (MH estimation?)
- Question #4: would informative priors help?

29