BSEM 2.0

Bengt Muthén & Tihomir Asparouhov

Mplus
www.statmodel.com

Presentation at the Mplus Users' Meeting
Utrecht, January 13, 2016

Overview

How to make the case for Bayes:

- Non-informative priors
  - Regression analysis with missing on x's
  - Mediation analysis with missing on mediator and x's
- Informative priors
  - BSEM
  - Time-series factor analysis

1. Bayes' Advantage Over ML: Non-Informative Priors

Using Bayes with non-informative priors as a computational device to obtain results that are essentially the same as ML if ML could have been used:

The example of missing data on covariates

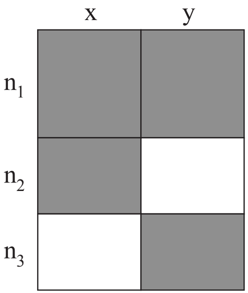
- Regression analysis
- Mediation analysis

Regressing y On x: Bringing x's Into The Model

ML estimation maximizes the log likelihood for the bivariate distribution of y and x expressed as,

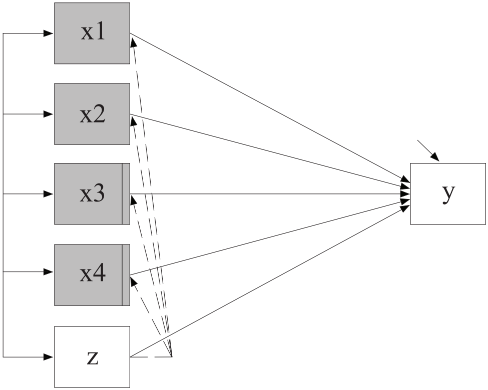
$$logL = \sum_i log[y_i, x_i] = \sum_{i=1}^{n_1} log[y_i \mid x_i] + \sum_{i=1}^{n_1+n_2} log[x_i] + \sum_{i=n_2+1}^{n_2+n_3} log[y_i]. \quad (1)$$

Figure : Missing data patterns. White areas represent missing data



Example: Monte Carlo Simulation Study

- Linear regression with 40% missing on x_1 - x_4; no missing on y
- x_3 and x_4 s are binary split 86/16
- MAR holds as a function of the covariate z with no missing
- $n = 200$
- Comparison of Bayes and ML



Bayes Treating Binary X's As Binary

```
DATA:           FILE = MARn200replist.dat;
                TYPE = MONTECARLO;
VARIABLE:       NAMES = y x1-x4 z;
                USEVARIABLES = y x1-z;
                CATEGORICAL = x3-x4;
DEFINE:         IF(z gt .25)THEN x1=_MISSING;
                IF(z gt .25)THEN x2=_MISSING;
                IF(-z gt .25)THEN x3=_MISSING;
                IF(-z gt .25)THEN x4=_MISSING;
ANALYSIS:       ESTIMATOR = BAYES;
                PROCESSORS = 2;
                BITERATIONS = (10000);
                MEDIATOR = OBSERVED;
MODEL:          y ON x1-z*.5;
                y*1;
                x1-z WITH x1-z;
```

ML Versus Bayes Treating Binary X's As Binary

- Attempting to estimate the same model using ML leads to much heavier computations due to the need for numerical integration over several dimensions
- Already in this simple model ML requires three dimensions of integration, two for the x_3, x_4 covariates and one for a factor capturing the association between x_3 and x_4.
- Bayes uses a multivariate probit model that generates correlated latent response variables underlying the binary x's - no need for numerical integration

Monte Carlo Simulation Results (500 replications)

	Population	Average	Std. Dev.	S.E. Average	M.S.E.	95% Cover	% Sig Coeff
			MLR with binary x's treated as normal				
x1	0.500	0.5087	0.1474	0.1358	0.0218	0.928	0.912
x2	0.500	0.5016	0.1453	0.1380	0.0211	0.932	0.910
x3	0.500	0.4684	0.4193	0.3440	0.1764	0.904	0.366
x4	0.500	0.4970	0.4255	0.3508	0.1807	0.888	0.380
			Bayes with binary x's treated as normal				
x1	0.500	0.5024	0.1351	0.1366	0.0182	0.968	0.926
x2	0.500	0.4966	0.1327	0.1380	0.0176	0.960	0.912
x3	0.500	0.4874	0.3518	0.3516	0.1237	0.958	0.300
x4	0.500	0.5066	0.3549	0.3519	0.1257	0.962	0.318
			Bayes with binary x's treated as binary				
x1	0.500	0.5106	0.1227	0.1214	0.0151	0.946	0.984
x2	0.500	0.5069	0.1212	0.1221	0.0147	0.948	0.972
x3	0.500	0.4765	0.3363	0.3345	0.1134	0.960	0.300
x4	0.500	0.5016	0.3459	0.3354	0.1194	0.950	0.332

Binary y With Missing Data On Covariates

- Treating all covariates as normal, ML needs 4 dimensions of integration for the 4 covariates with missing data
- Treating all covariates as normal, Bayes takes 30% of the ML computational time
- Treating x_3, x_4 as binary ML needs 5 dimensions of integration
- With a categorical y and many covariates with missing data that are brought into the model Bayes is the only practical alternative

Example: Mediation Analysis With Missing Data On The Mediator And The Covariates

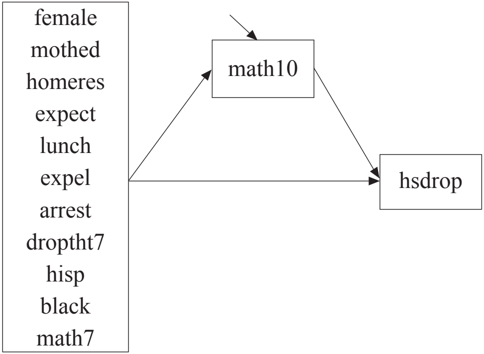


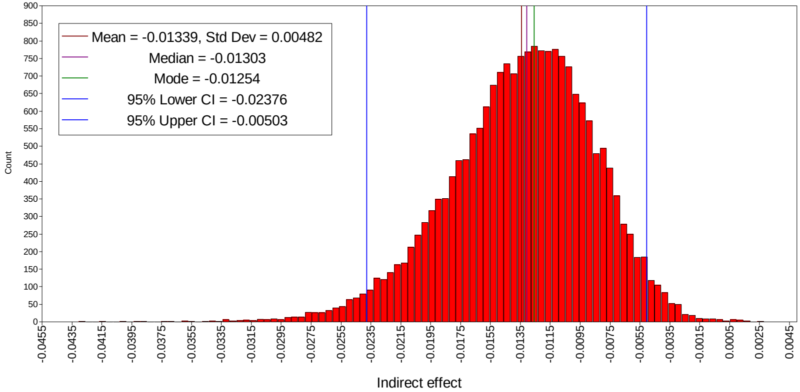
Figure : Mediation model for a binary outcome of dropping out of high school (n=2898)

```
                     CATEGORICAL = hsdrop;
ANALYSIS:            ESTIMATOR = BAYES;
                     PROCESSORS = 2;
                     BITERATIONS = (20000);
MODEL:               hsdrop ON math10 female-math7;
                     math10 ON female-math7;
MODEL INDIRECT:
                     hsdrop IND math10 math7(61.01 50.88);
OUTPUT:              SAMPSTAT PATTERNS TECH1 TECH8 CINTERVAL;
PLOT:                TYPE = PLOT3;
```

Indirect and direct effects computed in probability scale using counterfactually-based causal effects:

Muthén, B. & Asparouhov, T. (2015). Causal effects in mediation modeling: An introduction with applications to latent variables. Structural Equation Modeling: A Multidisciplinary Journal.

Bayesian Posterior Distribution Of Indirect Effect For High School Dropout



Missing On The Mediator: ML Versus Bayes

ML estimates are almost identical to Bayes, but:

- ML needs Monte Carlo integration with 250 points because the mediator is a partially latent variable due to missing data
- ML needs bootstrapping (1,000 draws) to capture CIs for the non-normal indirect effect
- ML takes 21 minutes

- Bayes takes 21 seconds
- Bayes posterior distribution for the indirect effect is based on 20,000 draws as compared to 1,000 bootstraps for ML

Missing On The Mediator And The Covariates Treating All Covariates As Normal: ML Versus Bayes

- ML requires integration over 10 dimensions
- ML needs 2,500 Monte Carlo integration points for sufficient precision
- ML takes 6 hours with 1,000 bootstraps

- Bayes takes less than a minute
- Bayes posterior based on 20,000 draws as compared to 1,000 bootstraps for ML

Missing On The Mediator And The Covariates
Treating Binary Covariates As Binary: ML Versus Bayes

6 covariates are binary.

- ML requires 10 + 15 = 35 dimensions of integration: intractable
- Bayes takes 3 minutes for 20,000 draws

Speed Of Bayes In Mplus

Wang & Preacher (2014). Moderated mediation analysis using Bayesian methods. Structural Equation Modeling.

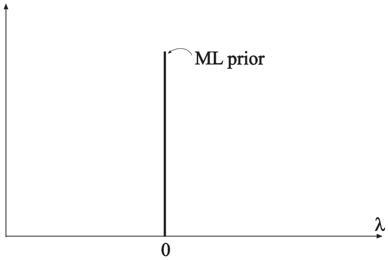
- Comparison of ML (with bootstrap) and Bayes: Similar statistical performance
- Comparison of Bayes using BUGS versus Mplus: Mplus is 15 times faster

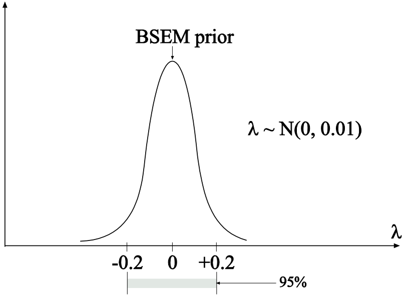
2. Bayes' Advantage Over ML: Informative Priors

- Frequentists often object to Bayes using informative priors
- But they already do use such priors in many cases in unrealistic ways
- Bayes can let informative priors reflect prior studies
- Bayes can let informative priors identify models that are unidentified by ML which is useful for model modification
- Example: CFA

ML Versus BESEM: CFA Cross-Loadings

- ML uses a very strict zero-mean, zero-variance prior
- BSEM uses a zero-mean, small-variance prior for the parameter:





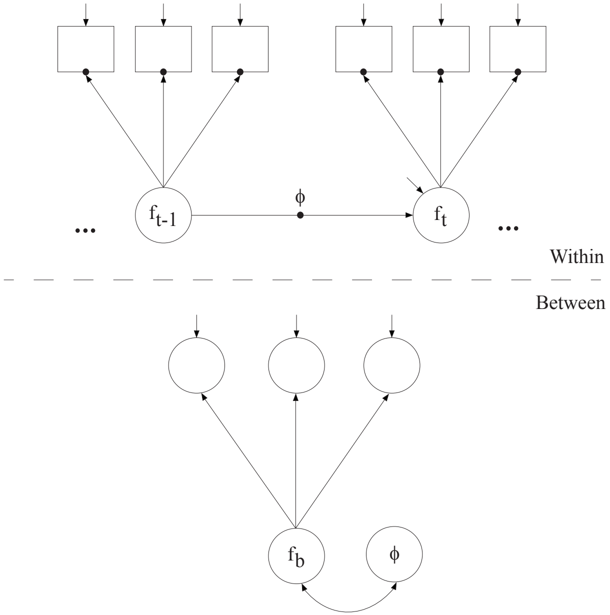
- EFA <BSEM <CFA

The Several Uses Of BSEM

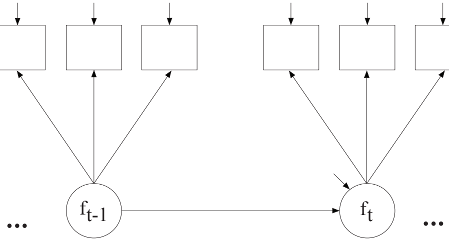
Non-identified models in ML made identified in Bayes using zero-mean, small-variance priors. Produces a Bayes version of "modification indices".

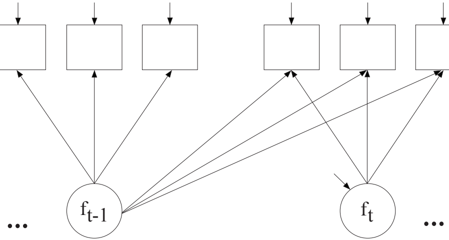
- Single-group analysis (2012 Muthén-Asparouhov article in Psychological Methods):
  - Cross-loadings in CFA
  - Direct effects in MIMIC
  - Residual covariances in CFA (2015 Asparouhov-Muthén-Morin article in Journal of Management)
- Multiple-group analysis:
  - Configural and scalar analysis with cross-loadings and/or residual covariances
  - Approximate measurement invariance (Web Note 17)
  - BSEM-based alignment optimization (Web Note 18):
    - Residual covariances
    - Approximate measurement invariance

Multilevel Time-Series Factor Analysis

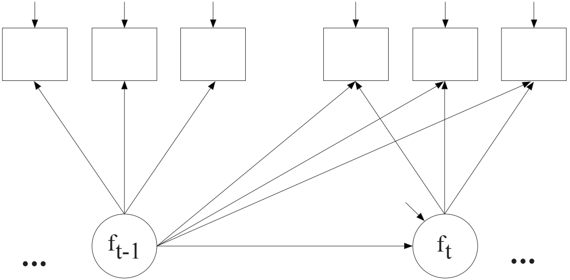


Time-Series Factor Analysis: DAFS and WNFS Models





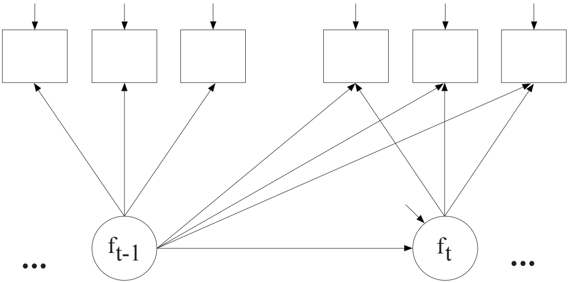
Time-Series Factor Analysis: Combined DAFS And WNFS Model



Example: Affective Instability In Ecological Momentary Assessment

- Jahng S., Wood, P. K.,& Trull, T. J., (2008). Analysis of Affective Instability in Ecological Momentary Assessment: Indices Using Successive Difference and Group Comparison via Multilevel Modeling. Psychological Methods, 13, 354-375
- An example of the growing amount of EMA data
- 84 outpatient subjects: 46 meeting borderline personality disorder (BPD) and 38 meeting MDD or DYS
- Each individual is measured several times a day for 4 weeks for total of about 100 assessments
- A mood factor for each individual is measured with 21 self-rated continuous items
- The research question is if the BPD group demonstrates more temporal negative mood instability than the MDD/DYS group

BSEM For The Combined DAFS And WNFS Model



Input for BSEM Of The Combined DAFS-WNFS Model

```
                USEVARIABLES = jittery-scornful group;
                BETWEEN = group;
                CLUSTER = id;
DEFINE:         group = group-1;
ANALYSIS:       TYPE = TWOLEVEL;
                ESTIMATOR = BAYES;
                PROCESSORS = 2; THIN = 5; BITERATIONS = (2000);
MODEL:          %WITHIN%
                f BY jittery-scornful*(& 1);
                f@1;
                f ON f&1;
                jittery-scornful ON f&1 (p1-p21);
                %BETWEEN%
                fb BY jittery-scornful*;
                fb ON group; fb@1;
MODEL
PRIORS:         p1-p21~N(0,0.01);
```

BSEM Direct Effects From f_{t-1} To y_t

Items with the largest direct effects:

- Upset
- Distressed
- Angry
- Irritable

Effects are negative, indicating that these items have lower auto-correlation than the rest. The factor auto-correlation therefore goes up. These direct effects can be freed, but...

Might these items measure a separate factor?

3-Factor EFA/CFA DAFS

Although time-series ESEM is needed, crude EFA suggests 3 factors:

- Angry: Upset, Distressed, Angry, Irritable
- Sad: Downhearted, Sad, Blue, Lonely
- Afraid: Afraid, Frightened, Scared

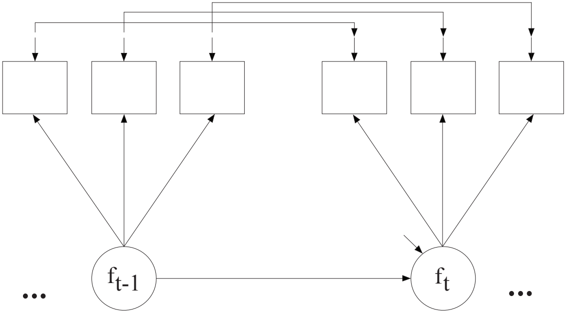
3-factor EFA/CFA DAFS factor autocorrelation (single-factor auto-corr = 0.596):

0.536 (Angry), 0.578 (Sad), 0.623 (Afraid).

- to which you could add random effects for the factor auto-correlations to see if they have different variability across subjects.

BSEM can be used again to search for direct effects from f_{t-1} to y_t.

Extended DAFS Model: Direct Effects From y_{t-1} To y_t



Factor Auto-Correlations With And Without Direct Effects From y_{t-1} To y_t

All but one of the 21 direct effects are significant and positive. Direct effects vary in size.

Table : Factor auto-correlations

	Angry	Sad	Afraid
Without direct effects	0.536	0.587	0.623
With direct effects	0.479	0.543	0.597

How to Learn More About Bayesian Analysis In Mplus: www.statmodel.com

- Topic 9 handout and video from the 6/1/11 Mplus session at Johns Hopkins
- Part 1 - Part 3 handouts and video from the August 2012 Mplus Version 7 training session at Utrecht University