Demonstration of an Online Educational Application on Bayesian Default Priors

Sonja D. Winter – University of California, Merced

swinter@ucmerced.edu

Sanne C. Smid - Utrecht University
s.c.smid@uu.nl

The Plan

1. Review: Dangers of the default
2. How to recognize a misbehaving prior?
3. Introducing our App
4. Using our app as a teacher
5. Next steps

Review: Dangers of the Defaults

1. Relatively larger impact of prior on posterior

2. Wide range of plausible parameter values

3. The idea that default priors are
non-informative and ‘let the data speak’

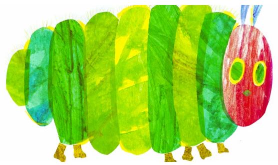
Review: Dangers of the Defaults

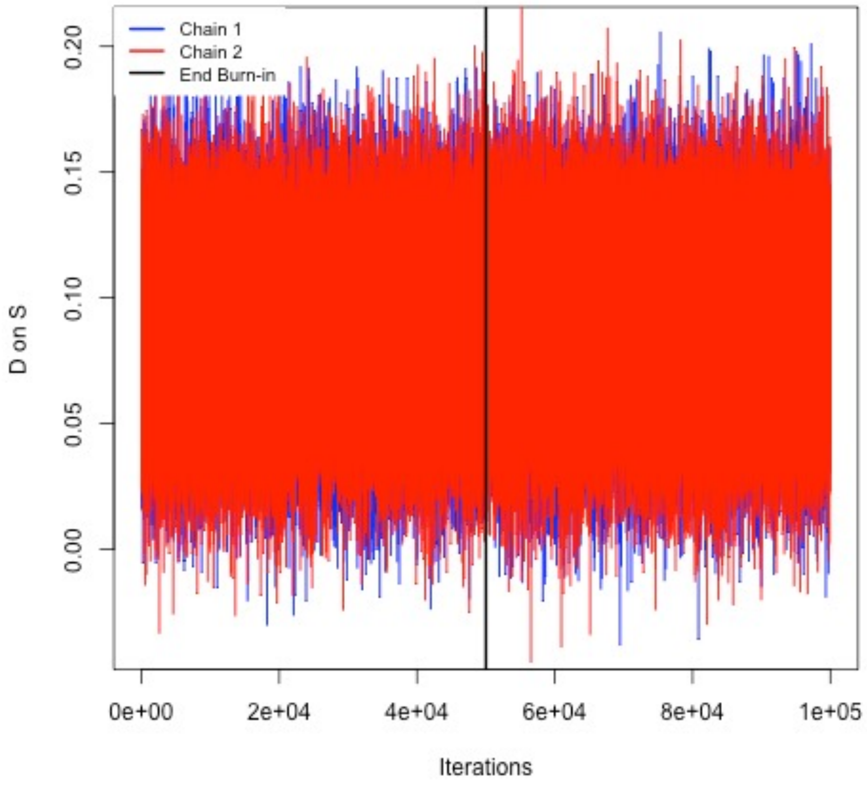
- Software packages make it very easy to naively switch to the Bayesian framework
- For example, M*plus*:

```
TITLE:
  Latent Growth Model, 4 Y, 1 Dist

  DATA:
  FILE = data_mplus.dat;

  VARIABLE:
  NAMES = t1-t4 D;

  ANALYSIS:
  ESTIMATOR = BAYES;

  MODEL:
  I BY t1-t4@1;
  S BY t1@0 t2@1 t3@2 t4@3;

  D ON I;
  D ON S;


  OUTPUT:
  tech1 tech8;

  PLOT:
  TYPE = PLOT2;
  SERIES = t1-t4(s);
```

How to recognize a misbehaving prior?

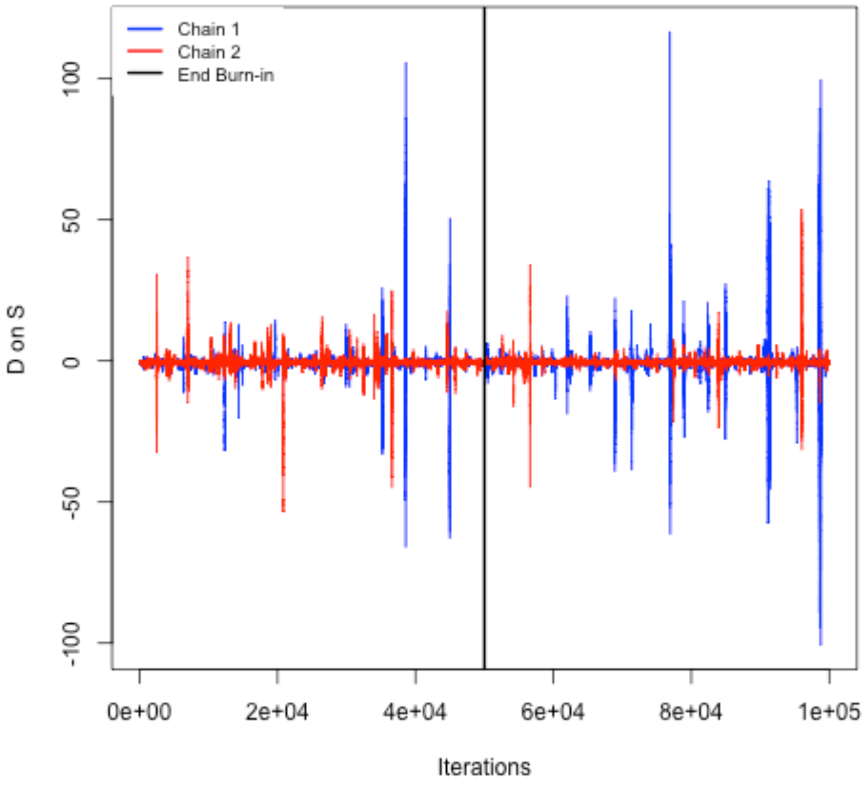
How to recognize a misbehaving prior?

- Characteristics of a trace plot when the prior behaves:
  1. Shape looks like the hungry caterpillar after it ate everything
  
  2. Range of values in posterior is relatively narrow
  3. All chains occupy the same parameter space (high degree of overlap)



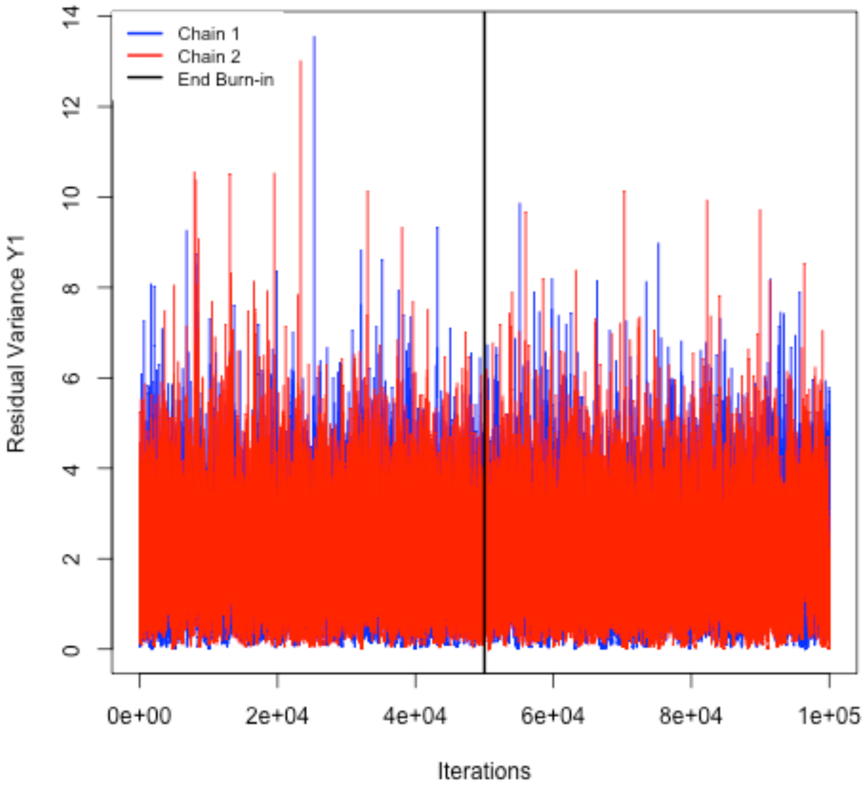
How to recognize a misbehaving prior?

- Characteristics of a trace plot when the prior misbehaves:
  1. Shape includes spikes
  2. Range of values in posterior is relatively wide
  3. Chains do not seem to have a high degree of overlap



How to recognize a misbehaving prior?

- Characteristics of a trace plot when the prior misbehaves:
  1. Shape includes spikes
  2. Range of values in posterior is relatively wide
  3. Chains do not seem to have a high degree of overlap



Introducing our App

- Why create an App?
  - To provide supporting illustration for Smid, Depaoli, & Van de Schoot (in press)
  - Create an educational tool that can be used by anyone to learn more about Bayesian statistics and priors
  - Create a worksheet that can be used by teachers who teach their students about responsible use of the Bayesian framework

  - We thought it might be fun to do?

Introducing our App

- Decisions we needed to make
  - Which defaults
  - What model
  - Which sample sizes
  - Which alternative priors

Introducing our App

- Which defaults do we use?
  - M*plus* default priors

- Why?
  - Popular software package for user friendly estimation through Bayesian framework
  - Makes it extremely easy to be a naïve user of Bayes

Introducing our App

- Challenges with these defaults

- Some M*plus* default priors are impossible to visualize
  - Default prior for the latent variable covariance matrix: IW(0, -p – 1)
  - Default prior for residual variances: IG(-1, 0)

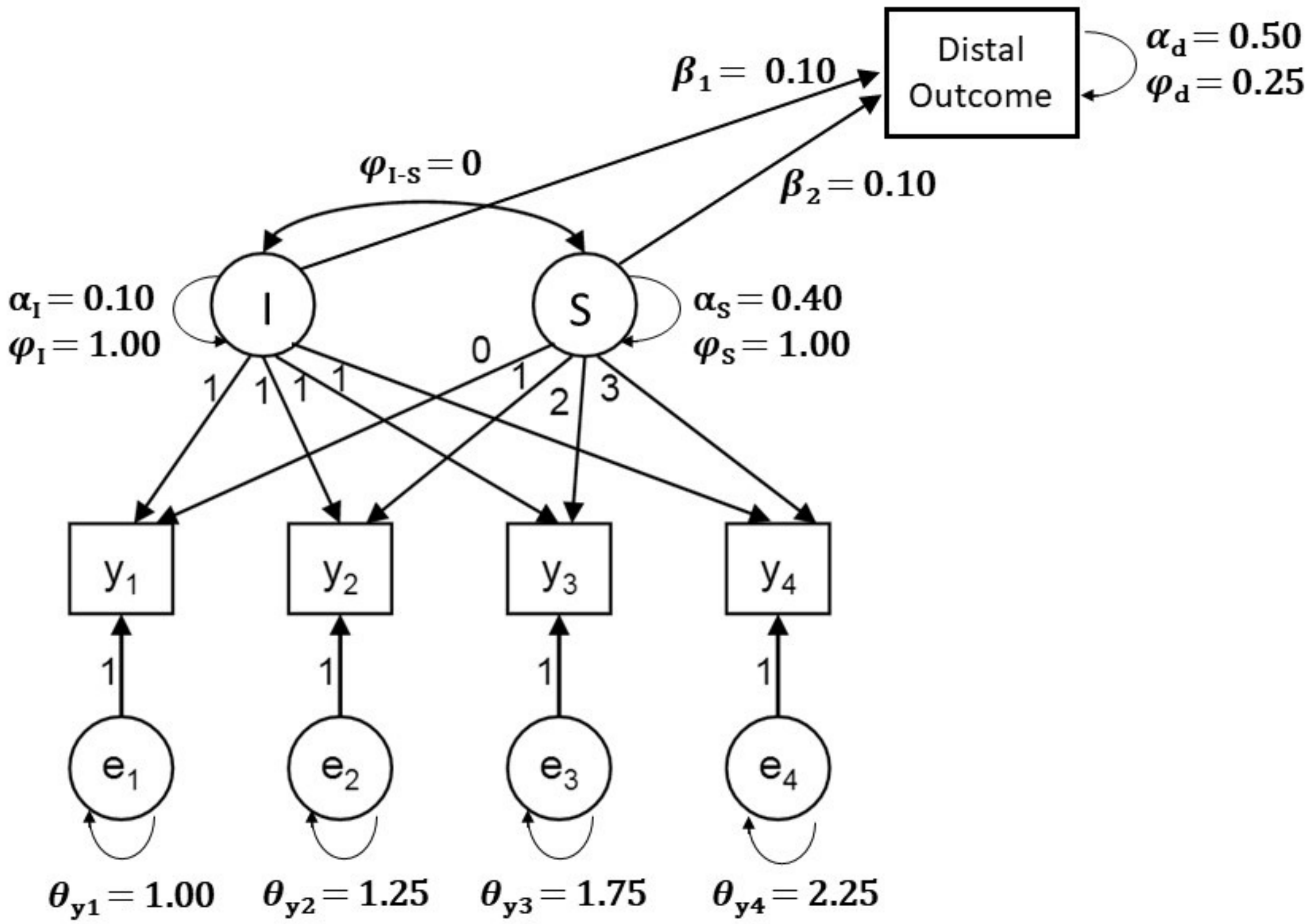
Introducing our App

- Challenges with these defaults

- We based our visualization on info from Asparouhov & Muthén (2010)
  - For IW: “This prior is essentially the uniform prior on $(-\infty,\infty)$ for all Σ parameters.” (p. 35)
  - For IG: “The default non-informative prior in Mplus is IG(−1, 0) which has constant density of 1 on the interval $(-\infty,\infty)$.” (p. 35)

Introducing our App

- The Model
  - Latent Growth Model with a distal outcome
  - Four time points
  - Intercept and linear slope



Introducing our App

- Which sample sizes?
  - N = 26 (very small)
  - N = 52 (small)
  - N = 325 (large)

- Why?
  - Taken from Smid, Depaoli, & Van de Schoot (in press) conditions
  - Based on formula $n = da$, where a is number of unknown parameters (13) and d is a constant > 2 that quantifies how large the sample should be (Lee & Song, 2004)

Introducing our App

- Which alternative priors?

- We added two additional prior specifications
    1. **Partial informative:** includes informative priors for the estimated means (I, S, D) and regression coefficients
    2. **Informative:** includes the above, in addition to an alternative prior specification for the latent variable covariance matrix using a so-called “Separation prior” (Liu, Zhang and Grimm (2016))

Introducing our App

You’ve waited long enough!

- https://utrecht-university.shinyapps.io/impact_of_prior_distributions/

- Or download and run locally: https://osf.io/m6byv/

~/Dropbox/UC Merced/Group Projects/Shiny - PriorEduc/Shiny App v2.0 - Shiny

http://127.0.0.1:7788 Open in Browser Publish

The Impact of Prior Distributions in a Bayesian Latent Growth Model

Utrecht University
UCMERCED
UNIVERSITY OF CALIFORNIA

DIY | Priors and settings | Resources

This Shiny App is created as an educational tool to show how varying prior distributions can affect parameter estimates in a Bayesian Latent Growth Model under varying sample sizes. We specified a Latent Growth Model with an intercept, linear slope, four time points and a distal outcome.
Below, you can play around with different prior specifications and sample sizes and explore their effect on the parameter estimates in the model. Note that all variations of the model were externally run using the software Mplus (Muthén & Muthén, 2017).

All code to reproduce this shiny app, generate the data and run the models in Mplus, can be found on the OSF: https://osf.io/m6byv/

By using this app you agree to be bound by the Terms of Usage.

Step 0. Start from Scratch

Reset Model

Step 1. Decide on Sample Size

- n = 26
- n = 52
- n = 325

Step 2. Choose your Priors

More info

- Mplus default priors
- Partial informative priors
- Informative priors

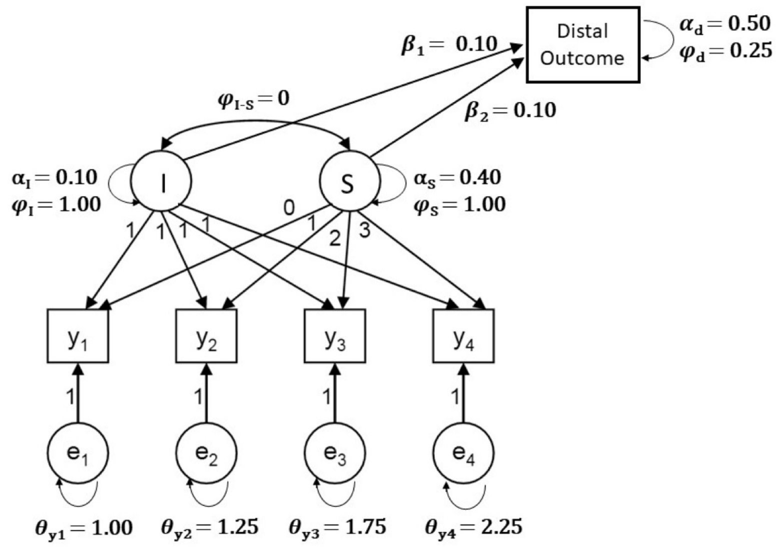
Step 3. Run the Model

Run Model

Step 4. Run the Model Again

Run Model Again

Convergence | Plots | Estimates | Interpretation Guide | The Model



Latent Growth Model with an Intercept, Linear Slope, four time points and a continuous distal outcome. A distal outcome variable is also known as a long-term variable. It refers to a wave of data collection that occurs long after the other waves of data collection in the Latent Growth Model. The model and population values used in the Shiny App are similar to the model and population values examined in Smid, Depaoli & van de Schoot (in press).

Using Our App as a Teacher

- Download the worksheet and answer key from the App

- Let students work in groups so they can explore more sample x prior combinations

- Students can exports all plots and tables they create and write up a report.

Using Our App as a Teach

- Download the worksheet and answer k
- Let students work in groups so they ca combinations
- Students can exports all plots and table

The Impact of Prior Distributions in a Bayesian Latent Growth Model
Worksheet

Question 1
Before looking at any of the results in the App, how do you expect that *sample size* affects the posterior estimates? Try to think about this in terms of convergence, accuracy, and precision.

Question 2
Before looking at any of the results in the App, how do you expect the *informativeness of the prior distributions* to affect the posterior estimates? Again, try to think about this in terms of convergence, accuracy, and precision.

Now, choose a sample size and a prior specification (click on “more info” to learn more about the specific priors used). Run the model and then run it again with twice as many iterations by clicking the buttons in the menu on the left.

Question 3
What sample size and prior specification did you select?

Question 4
Based on the PSR values, did the model converge?

Question 5
After running the model with twice as many iterations, move on to the “Plots” tab in the App. Do you think the trace plots of all individual parameter agree with the PSR values?

Question 6
Which (if any) parameters have trace plots that do not look like the desired caterpillar shape?

Question 7
Continue to the “Estimates” tab. Do any of the 95% HPD intervals reported in the table look unusually wide? What about the posterior standard deviation (SD)?

Next Steps

- This App is very much a work-in-progress
- We are not web developers!
- Let us know what you'd like to see!

- Limitations:
  - Results currently based on one pre-run (in *M*plus) replication of sample x prior combination
  - We only look at one software package's default priors
  - Only one type of model is included

Next Steps

- This App is very much a work-in-progress
- We are not web developers!
- Let us know what you'd like to see!

- Our wish list:
  - Look at impact of default priors in Blavaan
  - Run samplers ”live” in App
    - This opens up the possibility of users specifying their own priors and selecting more finetuned sample sizes
  - Include more visuals (e.g., autocorrelation plots, posterior histograms)

References

- Asparouhov, T., & Muthén, B. (2010). *Bayesian Analysis Using Mplus: Technical Implementation*. Retrieved from https://www.statmodel.com/download/Bayes3.pdf
- Lee, S. Y., & Song, X. Y. (2004). Evaluation of the Bayesian and maximum likelihood approaches in analyzing structural equation models with small sample sizes. *Multivariate Behavioral Research*, *39*(4), 653-686.
- Smid, S.C., Depaoli, S., & van de Schoot, R. (in press). Predicting a Distal Outcome Variable from a Latent Growth Model: ML versus Bayesian Estimation. *Structural Equation Modeling: A Multidisciplinary Journal.* https://doi.org/10.1080/10705511.2019.1604140