Bayesian and Frequentist Multilevel Modeling

Andy Grogan-Kaylor

2020-01-14

Contents

1 *Introduction*

$$y_{ij} = \beta_0 + \beta_1 x_{1i} + u_{0j} + e_{ij}$$

All multilevel models account for group structure, in estimating the association of x and y.

Bayesian models may offer some advantages over frequentist models, but may be substantially slower to converge.

2 *Prior Information*

Bayesian models allow one to incorporate prior information about a parameter of interest.

Prior information may come from the prior research literature, e.g. from systematic reviews or meta-analyses, or expert opinion or clinical wisdom.

3 *Smaller Samples*

Bayesian multilevel models may be better with small samples, especially samples with small numbers of Level 2 units (Hox et al. 2012). It is not clear to what degree this improvement in performance is dependent upon the use of informative priors.

4 Full Distribution of Parameters

Bayesian models of all kinds provide full distributions of the parameters (e.g. β's and random effects (u's))–both singly and jointly–rather than only point estimates.

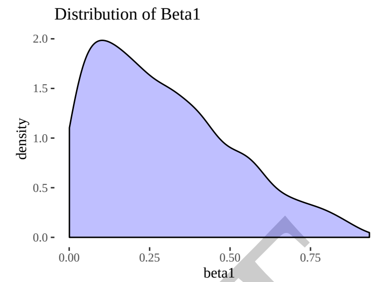

Figure 1: Distribution of a Single Parameter

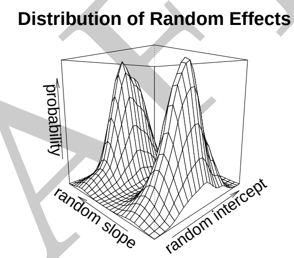

Figure 2: Joint Distribution of Parameters

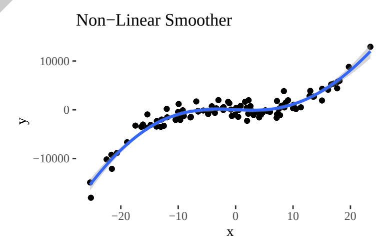

Figure 3: Non-Linear Terms

5 Accepting H_0 is Possible

Bayesian models allow one to both accept and reject H_0 (Kruschke and Liddell 2018). This may have consequences for affirming similarity, universality, or treatment invariance (Morey, Homer, and Proulx 2018).

Also, accepting certain null hypotheses may allow for the simplification of theory.

6 Distributional Models

Bayesian estimators allow one to directly model σ_{u_0}, the variance of the Level 2 units as a function of covariates (Burkner 2018).

7 Non-Linear Terms

Bayesian estimators allow for the incorporation of non-linear terms (Burkner 2018).[1]

[1] Such non-linear terms offer ways of non-parametrically fitting curvature. Do they represent over-fitting? Do they provide substantively interpretable results?

8 Maximal Models

Bayesian estimators allow for the inclusion of a large number of random slopes, e.g. $u_1, u_2, u_3, ..., etc.$ even when some of those estimated slopes are close to 0.

References

Burkner, Paul-Christian. 2018. "Advanced Bayesian Multilevel Modeling with the R Package brms." *The R Journal* 10 (1): 395–411.

Hox, Joop, Joop J. C.M. Hox, Rens van de Schoot, and Suzette Matthijsse. 2012. "How few countries will do? Comparative survey analysis from a Bayesian perspective." *Survey Research Methods*. https://doi.org/10.18148/srm/2012.v6i2.5033.

Kruschke, John K, and Torrin M Liddell. 2018. "The Bayesian New Statistics: Hypothesis testing, estimation, meta-analysis, and power analysis from a Bayesian perspective." *Psychonomic Bulletin & Review* 25 (1): 178–206. https://doi.org/10.3758/s13423-016-1221-4.

Morey, Richard D., Saskia Homer, and Travis Proulx. 2018. "Beyond Statistics: Accepting the Null Hypothesis in Mature Sciences." *Advances in Methods and Practices in Psychological Science*. https://doi.org/10.1177/2515245918776023.