State of the Art of Missing Data in Meta-analysis

4/29/2020

Aims to answer 3 questions:

1. What are people talking about theory of missing data in meta-regression?
2. What are they doing in practice? Why?
3. What should they be doing?

## 1. Methods for Handling missing data in research synthesis

This section will attempt to answer question 1. But will be mostly focused on missing predictor variables.

- 1.1 "Ad Hoc" methods:  
 1.1.1 Complete Case Analysis or listwise deletion  
 1.1.2 Available Case Analysis  
 1.1.3 Single-Value Imputation Methods  
 3.1 Complete Case Mean  
 3.2 With Conditional means (regression model)  
- 1.2 "Model Based" Methods:  
 1.2.1 Maximun Likelihood Methods Using EM Algorithm  
 1.2.2 Multiple Imputation for Multivariate Normal Data

**NOTE FROM JAKE:** You might mention something like hot-deck imputation, last-observation-carried-forward, and random imputation (replacing a cell from another row chosen at random, which can happen with hot-deck) in your “ad-hoc” methods.

**NOTE FROM JAKE:** There are some great statistical papers from the 1990s by Joe Ibrahim (see our Zotero repo) on EM for missing predictors in GLM, including methods that incorporate MNAR. It would be *tough*, but a worthwhile contribution to derive and implement those for meta-analysis.

**NOTE FROM JAKE:** It’s important to point out “missing data in meta-analysis” is inherently ambiguous. You could be missing data from individual participants within studies. You could be missing entire studies (that are in the filedrawer). You could be missing effects from studies that you don’t even know are missing (i.e., some effects from a study are in the file drawer). You could also be missing information about effects or covariates, which is what I think you’re getting at. Not a huge thing to fix, but seems like it would help situate what you’re talking about.

**NOTE FROM JAKE:** It strikes me there could be a neat paper to write about missingness within studies versus missingness between studies. Like, if there is attrition within studies, how does that affect a meta-analysis? Conversely, if there is little attrition within studies, but we’re missing effects (e.g., from publication selection), how much can that screw up a meta-analysis?

## 2. Current Practices in meta-analysis

This section will try to summarized current practices for handling missing data in meta-anylisis.(What methods are prefered and why?)

- 2.1 Commonly used methods:   
 "45% of studies report missing data problems. When missing data were found, the majority (86%) of synteses reported using ad hoc approach to handle missingess. Only 3% reported using multiple imputation."[2] (Tipton et al., 2019).  
 - 2.2 Fundamental problems of ad hoc approaches  
 - 2.3 Advantages of using Model based methods   
 - 2.4 Remaining work needed   
 "Remains a need for translational work on missing data methods, including compelling cases studies, examples of software implementation, and simulation studies on the comparative performance of current practies versus methods such as multiple imputation."[3] (Tipton et al., 2019)

**NOTE FROM JAKE:** Per point 2.1, are they talking about missingness within studies (i.e., we’re missing individual participant data in a given study) or missingness from the meta-analytic dataset itself? That would be an important thing to know.

**NOTE FROM JAKE:** A key thing to keep in mind is that missing data corrections are all about assumptions. There’s the standard assumptions about why data are missing, which we refer to as MNAR/MAR/MCAR. But, there’s also assumptions about what we *would have* observed. That is, model-based methods make assumptions about *why* data are missing and *what we would likely have seen* if there was no data missing. Neither of those are particularly testable in an explicit sense. However, we can think about ways to examine if they are feasible (e.g., from a Bayesian perspective we could look at posterior predictive distributions).

## References

[1] Pigott, T. (2019). Missing data in Meta-analysis. In Cooper H., Hedges L., & Valentine J. (Eds.) The Handbook of Research Synthesis and Meta-Analysis  (pp. 367-382). NEW YORK: Russell Sage Foundation. Retrieved April 29, 2020, from www.jstor.org/stable/10.7758/9781610448864.20

[2] Tipton E, Pustejovsky JE, Ahmadi H. A history of meta‐regression: Technical, conceptual, and practical developments between 1974 and 2018. Res Syn Meth. 2019;10:161–179. <https://doi.org/10.1002/jrsm.1338>

[3] Tipton E, Pustejovsky JE, Ahmadi H. Current practices in meta‐ regression in psychology, education, and medicine. Res Syn Meth. 2019;10:180‐194. <https://doi.org/> 10.1002/jrsm.1339