List of Key References in Missing Data Analysis

Provided by:
Todd D. Little, Ph.D.

Medicine: Mind the Gap Seminar
Optimizing Inferences Using Principled Missing Data Treatments
June 29, 2016

Allison, P. D. (2002). *Missing data*. Thousand Oaks, CA: Sage Publications.

Anderson, T. W. (1957). Maximum likelihood estimates for a multivariate normal distribution when some observations are missing. *Journal of the American Statistical Association, 52*(278), 200–203. doi:10.1080/01621459.1957.10501379

Arbuckle, J. L. (1996). Full information estimation in the presence of incomplete data. In G. A. Marcoulides & R. E. Schumacker (Eds.), *Advanced structural equation modeling* (pp. 243–277). Mahwah, NJ: Lawrence Erlbaum Associates, Inc.

Carpenter, J. R., & Kenward, M. G. (2012). Multiple imputation and its applications. John Wiley & Sons.

Collins, L. M., Schafer, J. L., & Kam, C. M. (2001). A comparison of inclusive and restrictive strategies in modern missing data procedures. *Psychological Methods, 6*, 330−351. doi: 10.1037/1082-989X.6.4.330

Dempster, A. P., Laird, N. M., & Rubin, D. B. (1977). Maximum likelihood from incomplete data via the EM algorithm. *Journal of the Royal Statistical Society, Series B, 39*, 1–38.

Enders, C. K. (2010). *Applied missing data analysis*. New York, NY: Guilford.

Enders, C. K., & Bandalos, D. L. (2001). The relative performance of full information maximum likelihood estimation for missing data in structural equation models. *Structural Equation Modeling, 8*, 430–457. doi: 10.1207/S15328007SEM0803_5

Graham, J. W. (2003). Adding missing-data-relevant variables to FIML-based structural equation models. *Structural Equation Modeling, 10*, 80–100. doi: 10.1207/S15328007SEM1001_4

Graham, J. W. (2012). *Missing data: Analysis and design*. New York, NY: Springer. doi:10.1007/978-1-46144018-5_1

Graham, J. W., Olchowski, A. E., & Gilreath, T. D. (2007). How many imputations are really needed? Some practical clarifications of multiple imputation theory. *Prevention Science, 8*, 206−213. doi: 10.1007/s11121-007-0070-9

Graham, J. W., Taylor, B. J., & Cumsille, P. E. (2001). Planned missing data designs in the analysis of change. In L. M. Collins & A. G. Sayer (Eds.), *New methods for the analysis of change* (pp. 335–353). Washington, DC: American Psychological Association. doi: 10.1037/10409-011

Graham, J. W., Taylor, B. J., Olchowski, A. E., & Cumsille, P. E. (2006). Planned missing data designs in psychological research. *Psychological Methods*, 11, 323−343. doi: 10.1037/1082-989X.11.4.323

Honaker, J., & King, G. (2010). What to do about missing values in time-series cross-section data. *American Journal of Political Science, 54*(2), 561–581. doi: 10.1111/j.1540-5907.2010.00447.x

Little, R. J. A., & Rubin, D. B. (2002). *Statistical analysis with missing data*. Hoboken, NJ: John Wiley & Sons.

Orchard, T., & Woodbury, M. A. (1972). *A missing information principle: Theory and applications.* Paper presented at the sixth Berkeley symposium on mathematical statistics and probability, University of California, Berkeley, CA.

Raghunathan, T. E., Lepkowski, J. M., Van Hoewyk, J., & Solenberger, P. (2001). A multivariate technique for multiply imputing missing values using a sequence of regression models. *Survey Methodology, 27*(1), 85–96.

Rubin, D. B. (1976). Inference and missing data. *Biometrika, 63*, 581–592. doi: 10.1093/biomet/63.3.581

Rubin, D. B. (1978). Multiple imputations in sample surveys—A phenomenological bayesian approach to nonresponse. *Proceedings of the Survey Research Methods Section of the American Statistical Association,* 30–34.

Rubin, D. B. (1987). *Multiple imputation for nonresponse in surveys.* New York, NY: Wiley. doi: 10.1002/9780470316696

Rubin, D. B. (1996). Multiple imputation after 18+ years. *Journal of the American Statistical Association, 91*, 473−489. doi: 10.1080/01621459.1996.10476908

Savalei, V., & Rhemtulla, M. (2012). Teacher's corner: On obtaining estimates of the fraction of missing information from FIML. *Structural Equation Modeling, 19*, 477–494. doi: 10.1080/10705511.2012.687669

Schafer, J. L. (1997). *Analysis of incomplete multivariate data*. New York, NY: Chapman Hall.

Tanner, M. A., & Wong, W. H. (1987). The calculation of posterior distributions by data augmentation. *Journal of the American Statistical Association, 82,* 528–540. doi: 10.1080/01621459.1987.10478458

Van Buuren, S. (2012). *Flexible imputation of missing data*. Boca Raton, FL: CRC Press. doi: 10.1201/b11826

Van Buuren, S., Brand, J. P. L., Groothuis-Oudshoorn, C. G. M., & Rubin, D. B. (2006). Fully conditional specification in multivarite imputation. *Journal of Statistical Computation and Simulation, 76*(12), 1049–1064.

Von Hippel, P. T. (2009). How to impute interactions, squares, and other transformed variables. *Sociological Methodology, 39,* 265–291. doi: 10.1111/j.1467-9531.2009.01215.x

Self Citations:

Garnier-Villarreal, M., Rhemtulla, M., & Little, T. D. (2014). Two-method planned missing designs for longitudinal research. *International Journal of Behavioral Development, 38,* 411-422. (doi: 10.1177/0165025414542711).^

Howard, W. J., Rhemtulla, M., & Little, T. D. (2015). Using principal component analysis (PCA) to obtain auxiliary variables for missing data estimation in large data sets. *Multivariate Behavioral Research, 50,* 285-299. (doi 10.1080/00273171.2014.999267).

Jia, F., Moore, E. W. G., Kinai, R., Crowe, K. S., Schoemann, A. M., & Little, T. D. (2014). Planned missing data design on small sample size: How small is too small? *International Journal of Behavioral Development, 38,* 1-18.

Jorgensen, T. D., Rhemtulla, M., Schoemann, A., McPherson, B., Wu, W., & Little, T. D. (2014). Optimal assignment methods in three-form planned missing data designs for longitudinal panel studies. *International Journal of Behavioral Development, 38,* 397-410.

Lang, K. M., & Little, T. D. (2014). The supermatrix technique: A simple framework for hypothesis testing with missing data. *International Journal of Behavioral Development, 38,* 461–470.

Lang, K. M., & Little, T. D. (in press). Principled missing data treatments. *Prevention Science.*

Little, T. D., Jorgensen, T. D., Lang, K. M., & Moore, E. W. G. (2014). On the joys of missing data. *Journal of Pediatric Psychology, 39,* 151-162.

Little, T. D., Lang, K. M., Wu, w., & Rhemtulla, M. (in press). Missing data. In D. Cicchetti (Ed.), *Developmental Psychopathology* (3rd Ed., pp 000-000). New York, NY: Wiley.

Little, T. D., & Rhemtulla, M. (2013). Planned missing data designs for developmental researchers. *Child Development Perspectives, 7,* 199-204. (doi:10.1111/cdep.12043).

Rhemtulla, M., Jia, F., Wu, W., & Little, T. D. (2014). Planned missing designs to optimize the efficiency of latent growth parameter estimates. *International Journal of Behavioral Development, 38,* 423-434.

Rhemtulla, M., & Little, T. D. (2012). Planned missing data designs for research in cognitive development. *Journal of Cognition and Development, 13,* 425-438.

Rhemtulla, M., & Little, T. D. (2014). Planned missing data designs for longitudinal organizational research. In M. Hassett & E. Paavilainen-Mäntymäki (Eds.), *Handbook of longitudinal research methods in studying organisations* (pp. 56-77). Cheltenham Glos, UK: Edward Elgar Publishing.

Rhemtulla, M., Savalei, V., & Little, T. D. (2014). On the asymptotic relative efficiency of planned missingness designs. *Psychometrika, 1-28*. (doi:10.1007/s11336-014-9422-0).^

Wu, W., Jia, F., Rhemtulla, M., & Little, T.D. (2015). Search for efficient complete and planned missing data designs for analysis of change. *Behavior Research Methods, 47.*

Hogue, C. M., Pornprasertmanit, S., Fry, M. D., Rhemtulla, M., & Little, T. D. (2013). Planned missing data designs for spline growth models in salivary cortisol research. *Measurement in Physical Education and Exercise Science, 17,* 310-325.