

- Albert, J. H., & Chib, S. (1993). Bayesian analysis of binary and polychotomous response data. *Journal of the American Statistical Association*, 88(422), 669-679.
- Attias, H. (1999). A variational bayesian framework for graphical models. *Advances in neural information processing systems*, 12.
- Belin, T. R., & Rubin, D. B. (1995). The analysis of repeated-measures data on schizophrenic reaction times using mixture models. *Statistics in Medicine*, 14(8), 747-768.
- Berry, D. A. (2006). Bayesian clinical trials. *Nature Reviews Drug Discovery*, 5(1), 27-36.
- Best, N., Richardson, S., & Thomson, A. (2005). A comparison of Bayesian spatial models for disease mapping. *Statistical methods in medical research*, 14(1), 35-59.
- Bishop, C. M. (1999). Bayesian PCA. *Advances in neural information processing systems*, 382-388. <https://proceedings.neurips.cc/paper/1998/file/c88d8d0a6097754525e02c2246d8d27f-Paper.pdf>
- Blei, D. M., Ng, A. Y., & Jordan, M. I. (2003). Latent dirichlet allocation. *Journal of machine Learning research*, 3(Jan), 993-1022.
- Box, G. E. P., & Tiao, G. C. (1962). A further look at robustness via Bayes's theorem. *Biometrika*, 49(3), 419-432.
- Box, G. E., & Cox, D. R. (1964). An analysis of transformations. *Journal of the Royal Statistical Society: Series B*, 26(2), 211-243.
- Box, G.E.P. and Tiao, G.C. (1968). A Bayesian approach to some outlier problems. *Biometrika*, 55(1), 119-129.
- Carlin, B. P., Chaloner, K. M., Louis, T. A., & Rhame, F. S. (1995). Elicitation, monitoring, and analysis for an AIDS clinical trial. In *Case Studies in Bayesian Statistics, Volume II* (pp. 48-89). Springer, New York, NY.
- Carpenter, B., Gelman, A., Hoffman, M.D., Lee, D., Goodrich, B., Betancourt, M., Brubaker, M., Guo, J., Li, P. and Riddell, A. (2017). Stan: A probabilistic programming language. *Journal of Statistical Software*, 76(1), 1-32.
- Carvalho, C. M., Polson, N. G., & Scott, J. G. (2009, April). Handling sparsity via the horseshoe. In *Artificial Intelligence and Statistics* (pp. 73-80). PMLR. Available from: <http://proceedings.mlr.press/v5/carvalho09a/carvalho09a.pdf>
- Carvalho, C. M., Polson, N. G., & Scott, J. G. (2010). The horseshoe estimator for sparse signals. *Biometrika*, 97(2), 465-480.
- Chaloner, K., & Brant, R. (1988). A Bayesian approach to outlier detection and residual analysis. *Biometrika*, 75(4), 651-659.
- Chib, S. (1995). Marginal likelihood from the Gibbs output. *Journal of the American Statistical Association*, 90(432), 1313-1321.

Chaloner, K., Church, T., Louis, T. A., & Matts, J. P. (1993). Graphical elicitation of a prior distribution for a clinical trial. *Journal of the Royal Statistical Society: Series D (The Statistician)*, 42(4), 341-353.

Chib, S., & Carlin, B. P. (1999). On MCMC sampling in hierarchical longitudinal models. *Statistics and Computing*, 9(1), 17-26.

Comulada, W. S., & Weiss, R. E. (2007). On models for binomial data with random numbers of trials. *Biometrics*, 63(2), 610-617.

Frangakis, C. E., & Rubin, D. B. (2002). Principal stratification in causal inference. *Biometrics*, 58(1), 21-29.

Gelfand, A. E., Hills, S. E., Racine-Poon, A., & Smith, A. F. (1990). Illustration of Bayesian inference in normal data models using Gibbs sampling. *Journal of the American Statistical Association*, 85(412), 972-985.

Gelman, A. (2006). Prior distributions for variance parameters in hierarchical models (comment on article by Browne and Draper). *Bayesian Analysis*, 1(3), 515-534.

Gelman, A., Hwang, J., & Vehtari, A. (2014). Understanding predictive information criteria for Bayesian models. *Statistics and Computing*, 24(6), 997-1016.

Gelman, A., Jakulin, A., Pittau, M. G., & Su, Y. S. (2008). A weakly informative default prior distribution for logistic and other regression models. *The Annals of Applied Statistics*, 2(4), 1360-1383.

Gelman, A., & Meng, X. L. (1998). Simulating normalizing constants: From importance sampling to bridge sampling to path sampling. *Statistical Science*, 163-185.

Gelman, A., & Stern, H. (2006). The difference between “significant” and “not significant” is not itself statistically significant. *The American Statistician*, 60(4), 328-331.

Gelman, A., Meng, X. L., & Stern, H. (1996). Posterior predictive assessment of model fitness via realized discrepancies. *Statistica Sinica*, 733-760.

Ghosh, S. K., Mukhopadhyay, P., & Lu, J. C. J. (2006). Bayesian analysis of zero-inflated regression models. *Journal of Statistical Planning and Inference*, 136(4), 1360-1375.

Hoeting, J. A., Madigan, D., Raftery, A. E., & Volinsky, C. T. (1999). Bayesian model averaging: a tutorial. *Statistical Science*, 382-401.

Hoff, P. (2007). Modeling homophily and stochastic equivalence in symmetric relational data. *Advances in neural information processing systems*, 20.

Hoff, P. D., Raftery, A. E., & Handcock, M. S. (2002). Latent space approaches to social network analysis. *Journal of the American Statistical Association*, 97(460), 1090-1098.

Hoffman, M. D., & Gelman, A. (2014). The No-U-Turn sampler: adaptively setting path lengths in Hamiltonian Monte Carlo. *Journal of Machine Learning Research*, 15(1), 1593-1623.

Hyvärinen, A., & Dayan, P. (2005). Estimation of non-normalized statistical models by score matching. *Journal of Machine Learning Research*, 6(4).

- Kass, R. E., & Raftery, A. E. (1995). Bayes factors. *Journal of the American Statistical Association*, 90(430), 773-795.
- Laud, P. W., & Ibrahim, J. G. (1995). Predictive model selection. *Journal of the Royal Statistical Society: Series B (Methodological)*, 57(1), 247-262.
- Liu, Q., & Wang, D. (2016). Stein variational gradient descent: A general purpose bayesian inference algorithm. *Advances in neural information processing systems*, 29.
- Madigan, D., & Raftery, A. E. (1994). Model selection and accounting for model uncertainty in graphical models using Occam's window. *Journal of the American Statistical Association*, 89(428), 1535-1546.
- McCandless, L. C., Gustafson, P., & Austin, P. C. (2009). Bayesian propensity score analysis for observational data. *Statistics in Medicine*, 28(1), 4-112.
- Meng, X. L. (1994). Posterior predictive p -values. *The Annals of Statistics*, 22(3), 1142-1160.
- Minka, T. P. (2001, August). Expectation propagation for approximate Bayesian inference. In *Proceedings of the Seventeenth conference on Uncertainty in artificial intelligence* (pp. 362-369).
- Murray, I., Adams, R., & MacKay, D. (2010, March). Elliptical slice sampling. In *Proceedings of the thirteenth international conference on artificial intelligence and statistics* (pp. 541-548). *JMLR Workshop and Conference Proceedings*.
- Neelon, B. H., O'Malley, A. J., & Normand, S. L. T. (2010). A Bayesian model for repeated measures zero-inflated count data with application to outpatient psychiatric service use. *Statistical Modelling*, 10(4), 421-439.
- Oh, M. S., & Raftery, A. E. (2001). Bayesian multidimensional scaling and choice of dimension. *Journal of the American Statistical Association*, 96(455), 1031-1044.
- Rosenbaum, P. R., & Rubin, D. B. (1983). The central role of the propensity score in observational studies for causal effects. *Biometrika*, 70(1), 41-55.
- Rubin, D. B. (1976). Inference and missing data. *Biometrika*, 63(3), 581-592.
- Rubin, D. B. (1984). Bayesianly justifiable and relevant frequency calculations for the applied statistician. *The Annals of Statistics*, 1151-1172.
- Spiegelhalter, D. J., Best, N. G., Carlin, B. P., & Van Der Linde, A. (2002). Bayesian measures of model complexity and fit. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 64(4), 583-639.
- Spiegelhalter, D. J., Best, N. G., Carlin, B. P., & Van der Linde, A. (2014). The deviance information criterion: 12 years on. *Journal of the Royal Statistical Society: Series B: Statistical Methodology*, 485-493.
- Spiegelhalter, D. J., Freedman, L. S., & Parmar, M. K. (1994). Bayesian approaches to randomized trials. *Journal of the Royal Statistical Society: Series A (Statistics in Society)*, 157(3), 357-387.
- Vehtari, A., Gelman, A., & Gabry, J. (2017). Practical Bayesian model evaluation using leave-one-out cross-validation and WAIC. *Statistics and Computing*, 27(5), 1413-1432.

Verdinelli, I., & Wasserman, L. (1991). Bayesian analysis of outlier problems using the Gibbs sampler. *Statistics and Computing*, 1(2), 105-117.

Weiss, R. (1994). Pediatric pain, predictive inference, and sensitivity analysis. *Evaluation Review*, 18(6), 651-677.

Weiss, R. (1996). An approach to Bayesian sensitivity analysis. *Journal of the Royal Statistical Society: Series B (Methodological)*, 58(4), 739-750.

Weiss, R. (1997). Bayesian sample size calculations for hypothesis testing. *Journal of the Royal Statistical Society: Series D (The Statistician)*, 46(2), 185-191.

Weiss, R. E., & Marshall, C. R. (1999). The uncertainty in the true end point of a fossil's stratigraphic range when stratigraphic sections are sampled discretely. *Mathematical Geology*, 31(4), 435-453.

Weiss, R. E., Jia, J., & Suchard, M. A. (2011). A Bayesian model for the common effects of multiple predictors on mixed outcomes. *Interface Focus*, 1(6), 886-894.

Williams, C., & Rasmussen, C. (1995). Gaussian processes for regression. *Advances in neural information processing systems*, 8.