

SUPPORTING BAYESIAN MODELING WITH VISUALIZATIONS

Mine Dogucu and Alicia A. Johnson

Department of Statistical Science, University College London

Department of Mathematics, Statistics, and Computer Science, Macalester College

m.dogucu@ucl.ac.uk

With computational advances, Bayesian modeling is becoming more accessible. But because Bayesian thinking often differs from learners' previous statistics training, it can be challenging for novice Bayesian learners to conceptualize and interpret the three major components of a Bayesian analysis: the prior, likelihood, and posterior. To this end, we developed an R package, bayesrules, which provides tools for exploring common introductory Bayesian models: beta-binomial, gamma-Poisson, and normal-normal. Specifically, within these model settings, the bayesrules functions provide an active learning opportunity to interact with the three Bayesian model components, as well as the effects of different model settings on the model results. We present here the package's visualization functions and how they can be utilized in a statistics classroom.

INTRODUCTION

The typical statistics curriculum exposes students to statistical modeling through the frequentist paradigm. In this setting, the models are driven by one component—a likelihood function. Thus, when students are eventually exposed to Bayesian modeling, it can be difficult to grasp the *three* required components as well as the synergies among them: the prior, the likelihood, and the posterior. Visualizations can help bring these concepts to life and deepen learners' understanding of the concepts (Cole & Davidson, 1989; Micallef et al., 2012; Mosca et al., 2021). For example, tree diagrams (Binder et al., 2018), contingency tables (Cole & Davidson, 1989), and mosaic plots (White & Warr, 2021) assist the teaching of conditional probability and Bayes' Rule, two foundational Bayesian concepts. Yet one limitation of these existing visualization tools is that they are designed around Bayesian models for simple *events*, whereas modern Bayesian courses often expand into Bayesian models for *random variables* (Dogucu & Hu, 2022), i.e., Bayesian regression. These generalized model settings require a unique suite of visualizations.

As the scientific community draws attention to mis- and over-use of p -values, Bayesian methods have been suggested as an alternative to frequentist methods for both answering scientific questions (Wasserstein & Lazar, 2016) as well as teaching inferential thinking in science classes (Rosenberg et al., 2022) and statistics classes (Dogucu & Hu, 2022). To this end, there exist several examples of teaching Bayesian statistics at the undergraduate level (Dogucu & Hu, 2022; Hoegh, 2020; Hu, 2020; Johnson et al., 2020; Witmer, 2017).

As is true for teaching statistics in general, computational tools are recommended for enhancing student learning in Bayesian statistics (Albert & Hu, 2020). To this end, to support students' exploration of Bayesian modeling for random variables, we developed the R package *bayesrules* (Dogucu et al., 2021) to supplement the *Bayes Rules!* book (Johnson et al., 2022). The package and book are products of our own, as well as our coauthors, undergraduate Bayesian courses and yet are also appropriate for comparably trained practitioners. Though we will focus here on the package's visualization functions, the package supports the book and Bayesian classrooms in three other key ways: (a) it contains datasets for Bayesian modeling applications; (b) it provides functions for evaluating Bayesian models; and (c) it includes functions for numerical summaries of basic Bayesian models.

SAMPLE VISUALIZATIONS

In this section, we illustrate the *bayesrules* package's visualization functions for exploring and interacting with the foundational beta-binomial model. The package also includes exploratory tools for the gamma-Poisson and the normal-normal models. Because their motivation and syntax are similar, we omit a discussion of these tools here. The syntax across the various model functions is outlined in Table 1, and interested readers can learn more about the gamma-Poisson and the normal-normal models in Chapter 5 of the *Bayes Rules!* book.

In a beta-binomial model, the random variable of interest is some proportion, π . For example, π might be the proportion of voters that support a given candidate or the proportion of incoming text

messages that are spam. The initial step in this beta-binomial analysis is to identify and tune a beta prior model, i.e., a model that reflects our prior knowledge of π before observing any data.

For $\pi \sim \text{Beta}(\alpha, \beta)$, the continuous probability density function of beta can be specified by:

$$f(\pi) = \frac{\Gamma(\alpha + \beta)}{\Gamma(\alpha)\Gamma(\beta)} \pi^{\alpha-1} (1 - \pi)^{\beta-1} \quad \text{for } \pi \in [0,1]$$

where $\Gamma(z) = \int_0^\infty x^{z-1} e^{-x} dx$ and $\Gamma(z + 1) = z \Gamma(z)$.

The `plot_beta()` visualization function helps familiarize students with the beta model. For instance, the following code visualizes the probability density function (pdf), expected value (mean), and mode of the beta model with shape and rate parameters of 1 and 13, respectively (Figure 1): `plot_beta(alpha = 1, beta = 13, mean = TRUE, mode = TRUE)`.

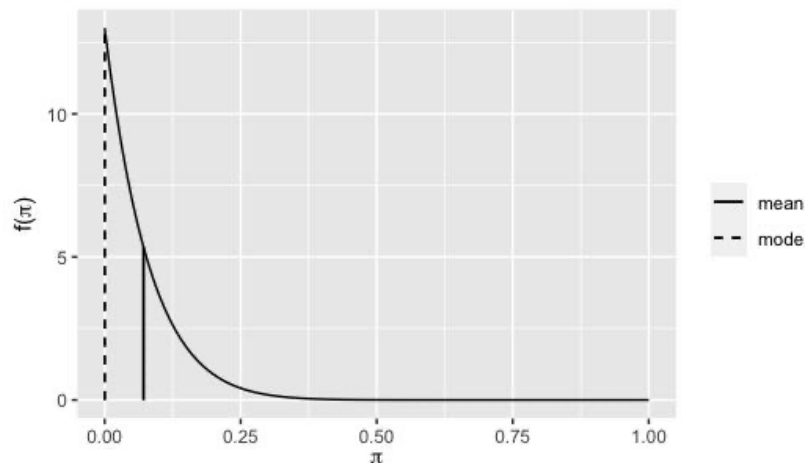


Figure 1. A Beta (1, 13) prior model plotted with its mean and mode

If π were to represent the proportion of voters that support a political candidate, then this particular prior model reflects an understanding that there is little voter support for the candidate. Specifically, even though π could be anywhere between 0 and 1, there is a strong prior understanding that π is likely near 0 and almost certainly below 0.25, with an expected value around 0.07 and a mode at 0. Yet by changing the alpha and beta arguments, students can explore the impact of these parameters and how to *tune* them to accurately reflect their prior knowledge of π . With or without visualizations, students are often challenged by the fact that the appropriate prior model is not typically defined by the context—they must tune their prior and there is not necessarily a singular “correct” choice. Thus, though the `plot_beta()` function can facilitate exploration, instructors still need to support their students in switching to Bayesian thinking and deepening their understanding of the prior model.

Following the tuning of a prior model, the second step in a beta-binomial analysis is to incorporate the observed data through the binomial likelihood function. To quantify this with numbers, reconsider the voter support example. Assume that in a survey of 50 voters, 25 have indicated their support for the political candidate. The posterior model of π then reflects some balance or consideration of the information in the prior model and data through the likelihood function. Students can explore these three components, and their synergies, using the `plot_beta_binomial()` function. For instance, the code below visualizes a Beta (1, 13) prior model, a binomial likelihood function reflecting 25 successes out of 50 trials, and the resulting beta posterior model (Figure 2): `plot_beta_binomial(alpha = 1, beta = 13, y = 25, n = 50, prior = TRUE, likelihood = TRUE, posterior = TRUE)`

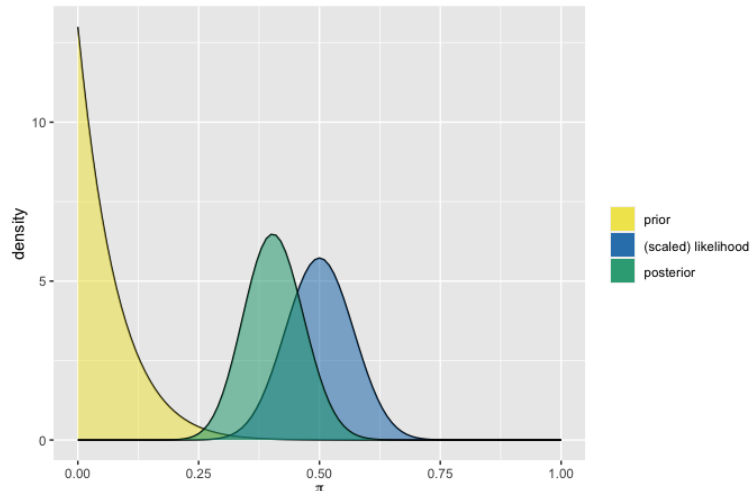


Figure 2. Prior, likelihood, and the posterior

By observing all three components of a Bayesian analysis side-by-side, students can more easily explore the fundamental balance that our posterior understanding of π strikes between our prior knowledge and the data. Specifically, through repeated usage of this function, learners will arrive at the bigger picture shown in Figure 3. This figure displays three different prior scenarios, from highly informative to flat priors, and three different likelihood scenarios where the maximum likelihood estimate is the same but the sample size is different. Through the exploration of these scenarios, students will directly observe that the weight of the prior is heavier when the prior is highly informative. Similarly, the data, as represented through the likelihood function, has a higher weight when the sample size is larger (i.e., when we have more evidence).

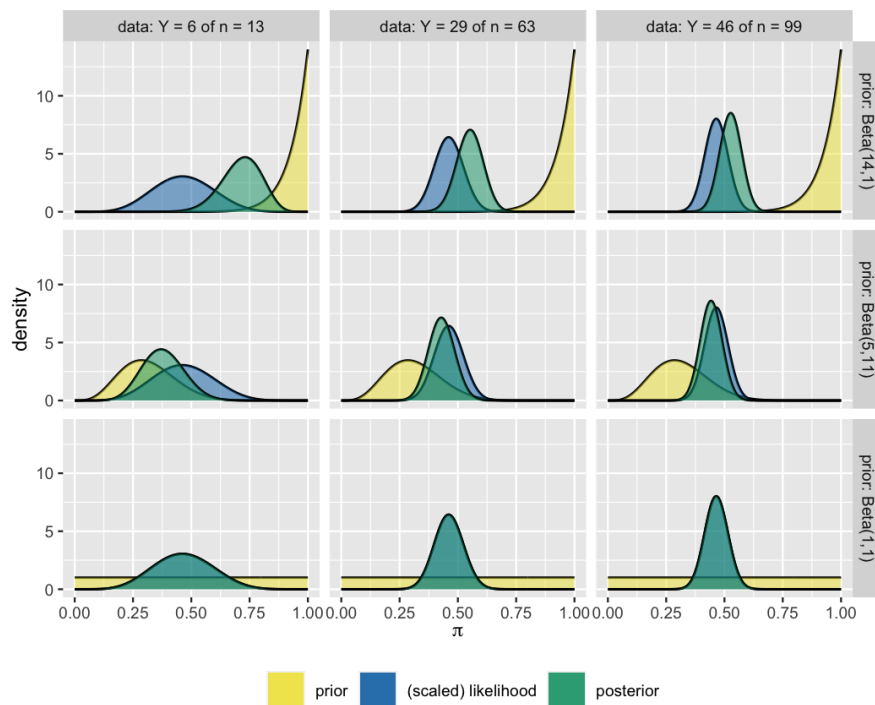


Figure 3. Posterior models constructed under different combinations of prior models and observed data

Table 1. A summary of the visualization functions within the *bayesrules* package

Model Name	Prior Model	Full Bayesian Model
Beta-Binomial	<code>plot_beta</code> (<code>alpha = 1, beta = 13, mean = TRUE, mode = TRUE</code>)	<code>plot_beta_binomial</code> (<code>alpha = 1, beta = 13, y = 25, n = 50, prior = TRUE, likelihood = TRUE, posterior = TRUE</code>)
Gamma-Poisson	<code>plot_gamma</code> (<code>shape = 2, rate = 11, mean = TRUE, mode = TRUE</code>)	<code>plot_gamma_poisson</code> (<code>shape = 100, rate = 20, sum_y = 39, n = 6, prior = TRUE, likelihood = TRUE, posterior = TRUE</code>)
Normal-Normal	<code>plot_normal</code> (<code>mean = 3.5, sd = 0.5</code>)	<code>plot_normal_normal</code> (<code>mean = 0, sd = 3, sigma = 4, y_bar = 5, n = 3, prior = TRUE, likelihood = TRUE, posterior = TRUE</code>)

CONCLUSION

One of the six recommendations in the Guidelines for Assessment and Instruction in Statistics Education (GAISE) college report is to foster active learning (GAISE College Report ASA Revision Committee, 2016). The functions we have presented here provide an opportunity for students to actively engage with the key Bayesian modeling components and concepts. In general, regardless of whether a course is Bayesian or not, we recommend the inclusion of R tools that allow students to explore statistical modeling outcomes under different scenarios. For instructors interested in developing their own course-specific R functions, we highly recommend a package infrastructure. We have relied heavily on Wickham's (2015) R packages book in writing and maintaining our package and recommend it to others.

REFERENCES

- Albert, J., & Hu, J. (2020). Bayesian computing in the undergraduate statistics curriculum. *Journal of Statistics Education*, 28(3), 236–247. <https://doi.org/10.1080/10691898.2020.1847008>
- Binder, K., Krauss, S., Bruckmaier, G., & Marienhagen, J. (2018). Visualizing the Bayesian 2-test case: The effect of tree diagrams on medical decision making. *PLOS One*, 13(3), Article e0195029. <https://doi.org/10.1371/journal.pone.0195029>
- Cole, W., & Davidson, J. (1989). Graphic representation can lead to fast and accurate Bayesian reasoning. In L. C. Kingsland (Ed.), *Proceedings of the thirteenth Annual Symposium on Computer Applications in Medical Care* (pp. 227–231). Institute of Electrical and Electronic Engineers Computer Society Press.
- Dogucu, M., & Hu, J. (2022). The current state of undergraduate Bayesian education and recommendations for the future. *The American Statistician*. Advance online publication. <https://doi.org/10.1080/00031305.2022.2089232>
- Dogucu, M., Johnson, A., & Ott, M. (2021). Datasets and supplemental functions from Bayes rules! book: bayesrules R package (Version 0.0.2) [Computer software]. <https://cran.r-project.org/web/packages/bayesrules/index.html>
- GAISE College Report ASA Revision Committee. (2016). *Guidelines for assessment and instruction in statistics education (GAISE) college report 2016*. American Statistical Association. https://www.amstat.org/docs/default-source/amstat-documents/gaiecollege_full.pdf
- Hoegh, A. (2020). Why Bayesian ideas should be introduced in the statistics curricula and how to do so. *Journal of Statistics Education*, 28(3), 222–228. <https://doi.org/10.1080/10691898.2020.1841591>
- Hu, J. (2020). A Bayesian statistics course for undergraduates: Bayesian thinking, computing, and research. *Journal of Statistics Education*, 28(3), 229–235. <https://doi.org/10.1080/10691898.2020.1817815>
- Johnson, A. A., Ott, M. Q., & Dogucu, M. (2022). *Bayes rules! An introduction to applied Bayesian modeling*. CRC Press. <https://doi.org/10.1201/9780429288340>

- Johnson, A., Rundel, C., Hu, J., Ross, K., & Rossman, A. (2020). Teaching an undergraduate course in Bayesian statistics: A panel discussion. *Journal of Statistics Education*, 28(3), 251–261. <https://doi.org/10.1080/10691898.2020.1845499>
- Micallef, L., Dragicevic, P., & Fekete, J. D. (2012). Assessing the effect of visualizations on Bayesian reasoning through crowdsourcing. *IEEE Transactions on Visualization and Computer Graphics*, 18(12), 2536–2545. <https://doi.org/10.1109/TVCG.2012.199>
- Mosca, A., Ottley, A., & Chang, R. (2021, May). Does interaction improve Bayesian reasoning with visualization? In *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems (CHI '21)* (Article 609). Association for Computing Machinery. <https://doi.org/10.1145/3411764.3445176>
- Rosenberg, J. M., Kubsch, M., Wagenmakers, E. J., & Dogucu, M. (2022). Making sense of uncertainty in the science classroom. *Science & Education*. Advance online publication. <https://doi.org/10.1007/s11191-022-00341-3>
- Wasserstein, R. L., & Lazar, N. A. (2016). The ASA statement on p-values: Context, process, and purpose. *The American Statistician*, 70(2), 129–133. <https://doi.org/10.1080/00031305.2016.1154108>
- White, E. D., & Warr, R. L. (2021). Teaching Bayes' rule using mosaic plots. *arXiv:2112.00162*. <https://doi.org/10.48550/arXiv.2112.00162>
- Wickham, H. (2015). *R packages: Organize, test, document, and share your code*. O'Reilly Media.
- Witmer, J. (2017). Bayes and MCMC for undergraduates. *The American Statistician*, 71(3), 259–264. <https://doi.org/10.1080/00031305.2017.1305289>