



Why Bayesian Ideas Should Be Introduced in the Statistics Curricula and How to Do So

Andrew Hoegh

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.
10.1080/10691898.2020.1841591

Made available through Montana State University's [ScholarWorks](https://scholarworks.montana.edu)
scholarworks.montana.edu

Why Bayesian Ideas Should Be Introduced in the Statistics Curricula and How to Do So

Andrew Hoegh

Department of Mathematical Sciences, Montana State University, Bozeman, MT

ABSTRACT

While computing has become an important part of the statistics field, course offerings are still influenced by a legacy of mathematically centric thinking. Due to this legacy, Bayesian ideas are not required for undergraduate degrees and have largely been taught at the graduate level; however, with recent advances in software and emphasis on computational thinking, Bayesian ideas are more accessible. Statistics curricula need to continue to evolve and students at all levels should be taught Bayesian thinking. This article advocates for adding Bayesian ideas for three groups of students: intro-statistics students, undergraduate statistics majors, and graduate student scientists; and furthermore, provides guidance and materials for creating Bayesian-themed courses for these audiences. Supplementary files for this article are available online.

KEYWORDS

Bayesian statistics; MCMC;
Statistics curricula

1. Introduction

In recent years, computing has taken a more central role in statistics curricula, and furthermore, there has been robust growth in statistical computing and data science. However, despite these advances the lasting legacy of “mathematical thinking” (Brown and Kass 2009) still influences curricula. In 2010, Nolan and Temple Lang encouraged action to be “bold and design curricula from scratch that embrace new and innovative topics and paradigms for teaching.” While there are good examples of adopting innovative ideas and topics for undergraduate programs (Steel, Liermann, and Guttorp 2019; Çetinkaya-Rundel and Rundel 2018), particularly those that have orchestrated statistics-centric data science programs and courses (Baumer 2015; Hardin et al. 2015; Hicks and Irizarry 2018), on the whole, statistics curricula could still use a *bold* redesign.

Historically, the combination of mathematics-centric statistical curricula along with the difficulty of implementing Bayesian models resulted in Bayesian courses being relegated to the margins of curricula. Even today, when the computing power exists to fit sophisticated Bayesian models *and* the curricular balance between computing and mathematics has begun to equalize, the most common place that students encounter Bayesian statistics is as a graduate elective, often in the 2nd year of a graduate program and even then, it is generally not a required course. While some larger programs do offer an undergraduate Bayesian courses as an elective, many smaller programs still do not regularly offer undergraduate Bayesian courses. Based on our reading of the course catalogs, none of the 10 largest undergraduate statistics programs (UC-Berkeley, Purdue, Illinois, UC-Davis, UCLA, Minnesota, UC-Santa Barbara, Michigan, Carnegie Mellon, Brigham Young University) require

a Bayesian course to earn a B.S. degree. Some of them don't even have an undergraduate Bayesian class listed in their course catalog. A conservative estimate would put about 10 statistics courses in the undergraduate curriculum. Hence, many graduate students are taking about 15 statistics courses before being exposed to Bayesian ideas. Furthermore many undergraduate statistics students, as well as research scientists from other disciplines, receive no formal training in Bayesian statistics.

With the advent of Markov chain Monte Carlo (MCMC) methods and the accessibility of software for Bayesian computing there have been large shifts in scientific research; in particular, more-and-more scientific domains are commonly using Bayesian methods. For instance there are now highly cited Bayesian textbooks for “Social and Behavior Sciences” (Jackman 2009; Gill 2014), “Ecology” (McCarthy 2007; Hobbs and Hooten 2015), and “Econometrics” (Koop 2003), amongst others. In response a change needs to be made so that understanding and implementing Bayesian methods becomes an essential part of the statistics curriculum. Nolan and Temple Lang summarized Brown and Kass's paper, “What is Statistics,” by pointing out that “historically ‘mathematical thinking influenced both research and infrastructure’ in statistics departments and currently may not be serving the field well.” With this article we describe how introducing Bayesian ideas serves both the field of statistics and science in general.

While not suggesting Bayesian methods should be taught exclusively or in place of frequentist methods, in the spirit of Stangl (1998), rather than an either-or approach, students should be exposed to frequentist *and* Bayesian ideas. Including Bayesian ideas in the curricula presents several advantages:

1. Bayesian methodology is inherently very computational, or algorithmic, which may appeal to or attract students that

often end up in other computationally focused data analysis fields.

2. The Bayesian data analysis process is straightforward: specify prior, state sampling model, and collect posterior samples for inference. This is not intended to diminish the role of computation or model checking in that process, but this single approach can be applied to virtually any statistical problem.
3. Equipping students with a toolbox consisting of a broad range of methods is helpful. Many B.S. and M.S. level students are more generalists than specialists. As more students enter the statistics and data science profession without advanced degrees, training in Bayesian methods provides more statistical tools.
4. While Bayesian methods and philosophy are useful in their own right, the philosophical contrasts between Bayesian approaches and classical statistical methods are profound and enhance learning, regardless of which side of the classical-Bayesian spectrum students embrace. Bayesian methods are not a complete panacea to all the ills detailed in the ASA statement on p -values (Wasserstein and Lazar 2016), but even if students continue to primarily use frequentist methods the exposure to the Bayesian paradigm would encourage deeper thought about the challenges with Null Hypothesis Significance Testing (NHST) and associated p -values.
5. Bayesian methods provide a natural way to understand uncertainty in statistical learning frameworks. While we would advocate including Breiman's (2001) ideas on the two cultures of statistical modeling in statistics curricula, the advantage of Bayesian formulations of many statistical learning ideas, such as tree-based models, is that the Bayesian approach gives a familiar treatment to model uncertainty.
6. Bayesian methods are becoming more common in many scientific fields, as well as in the social sciences and even the humanities, formal training gives practitioners the ability to articulate priors, formulate models, and implement these methods.

The remainder of this article describes why Bayesian methods are beneficial for three specific audiences: intro-statistics students, undergraduate statistics majors, and graduate non-statistician scientists; and furthermore, we provide suggestions on how to include Bayesian ideas in courses for these three audiences.

2. Bayesian Intro Stats

2.1. Why Should Bayesian Thinking Be Included in Intro Stat Courses?

There have been dramatic changes in the teaching approaches of intro stat courses to include simulation-based approaches (Tintle et al. 2014; Rossman and Chance 2014; Lock et al. 2016) and rely more on computational thinking than mathematical thinking. Furthermore, Bayesian thinking arguably maps onto human thinking better than computation, mathematical, or statistical thinking. In this article, intro stat refers to a first course in statistics; at the author's institution, intro stat requires college algebra but not calculus. In line with the *Guidelines for Assessment and Instruction in Statistics Education* (GAISE)

reports (Franklin et al. 2007; Carver et al. 2016), rather than pen-and-paper calculations of standard errors and looking up z and t statistics in the back of a textbook, a simulation-based approach uses computation to create a simulation distribution. The simulation distribution is used for inference and encourages students to "think distributionally."

One major advantage of simulation-based methods is that the data analysis process can be taught a single time and then used repeatedly in different settings. Traditionally, an intro stat course would iteratively cover one-sample t -tests, two-sample t -tests, paired t -tests, etc. and then require students to memorize or derive the mathematical formulas for standard errors. However, simulation, which is reliant on computational thinking and implementation, can be used in all of these modeling scenarios rather than the mathematical foundations necessary for calculating z or t scores.

While most simulation-based courses do not explicitly cover Bayesian ideas, simulation-based approaches are conceptually very similar to Bayesian methods, especially Bayesian ideas seen through the lens of generative models and approximate Bayesian computation (Baath 2017). Bayesian methods make inferences with the posterior distribution in the same way that simulation distributions are used in the simulation-based framework; both teach inference using a distributional framework. Furthermore, similar to the simulation-based approaches, a Bayesian approach to data analysis can simplify the data analysis process by not requiring students to derive or memorize specific mathematical details for each setting. Witmer (2017) advocated for teaching the Bayesian approach early in statistics curricula, such as the second class, and states that "Bayes via Markov chain Monte Carlo (MCMC)" provides a single unified framework for answering a set of interesting questions.

The move from rote mathematical calculations in traditional intro stat classrooms to the simulation approaches supporting (or supported by) GAISE has been popular (delMas, Garfield, and Chance 1998; Jamie 2002; Garfield and Ben-Zvi 2007). This paradigm shift can be phrased as a move from mathematical thinking to computational thinking. In particular, computational thinking shifts the focus from memorization and mathematical derivations to distributional thinking, an idea with a heavy Bayesian influence. The general purpose approach of simulation distributions provides a way to address a range of interesting data-oriented problems. Furthermore, as an additional benefit, a computationally focused intro stat course may help draw a new population of students into the statistics field.

2.2. How to Include Bayesian Thinking in Intro Stat Courses

One major challenge with intro stats courses is the wide range of backgrounds and future aspirations of students. Intro stat generally serves as not only the first course for all students that eventually grow into Ph.D. statisticians, but also is the only statistics course that many students will take. Given that intro stat is both the pipeline for potential statistics majors and serves as the lasting legacy of statistics for many other students, the course should provide a complete statistical foundation, serve as a building block for future statistics courses, and also

attract students to the statistical discipline. A simulation-based approach provides an intuitive multi-purpose procedure for statistical analysis that would serve both students not taking another stats class and future stat majors. This approach might even encourage students to take more stats courses.

Despite the similarities between simulation approaches and Bayesian methods, there are challenges in directly presenting Bayesian ideas to an intro stat audience. Moore's (1997) article titled "Bayes for Beginners? Some Reasons to Hesitate" suggested that, in 1997, it was premature to teach Bayes for beginners. However, shortly after this article, Stangl (1998) argued that many of these challenges are in fact the very reason to include Bayesian ideas in courses. Now, over 20 years later, some of the arguments in Moore (1997), such as "Bayesian methods are relatively rarely used in practice" and "It is unclear what Bayesian methods we should teach," are dated.

The major challenge that we see for teaching Bayesian ideas at an intro stat audience is the probability concepts necessary for conditional probability and generative probability distributions. Note that probability theory, along with basic statistics; drills with z , t , χ^2 , and F tests; making plots by hand; and advanced programming, are on the GAISE list of topics that might be excluded from intro stat courses (Carver et al. 2016). Some might suggest that prior distributions would be an additional challenge. We'd initially recommend starting with uniform priors to avoid some of the issues with teaching objective and subjective probability. Nevertheless, we describe two ways to introduce Bayesian thinking for an intro stats audience.

While we have not yet had the opportunity to teach a course like this, our preferred option would be to teach an *honors* version, or a course targeting students interested in statistics or data science, that would explicitly incorporate Bayesian ideas within a computationally focused paradigm. This intro stats course would look similar to the course Witmer (2017) described and use Witmer's "Bayes via MCMC" approach. Coupled with a computational overview, in place of time traditionally spent on mathematical derivations, Witmer's "Bayes via MCMC" could be more generally "Statistics via MCMC" for intro classes.

With only a college algebra pre-req, a Bayesian course might require additional instruction on probability concepts, specifically conditional probability and probability distributions. However, the approach discussed in Baath (2017), using approximate Bayesian computation (ABC), and a generative model, would be more in line with GAISE recommendations and would limit the need for some of this material. Specifically, using the generative model is very similar to the sampling-based approach common in intro stat courses and would alleviate the need for a comprehensive overview of probability. Computing, and in this case ABC, would be presented with R Shiny applets (Chang et al. 2019) and R, primarily through instructor generated functions.

The best textbook that we have found for this course would be Albert and Rossman (2009); however, given the textbook is over ten years old, we would supplement the coding and graphics section and also introduce the ABC framework from Baath (2017). The course learning outcomes would be:

- Understand and appreciate how statistics affects your daily life and the fundamental role of statistics in all disciplines;

- Create graphical visualizations to explore relationships in data;
- Use approximate Bayesian computational methods to conduct and analyze statistical studies; and
- Understand how data can be collected, and how data collection dictates the choice of statistical method and appropriate statistical inference.

We envision the course being taught using weekly, or biweekly, modules with homework and in-class labs. The GAISE report makes the distinction between intro courses focusing on literacy and methods as the difference between *users* and *producers* of statistical analyses (Carver et al. 2016). This course would be focused on creating producers of information and place a premium on reporting and communicating. Thus, in lieu of exams, an ongoing course project would serve as the final means of assessment.

Using the topics in Albert and Rossman (2009) as a rough guide, we'd envision modules focusing on: data and variables, R overview and graphics with ggplot2, measures of center and spread, comparing distributions with ggplot2, random sampling and scope of inference, probability distributions/generative models, approximate Bayesian computation (ABC), ABC models for one parameter models, and ABC models for regression for two parameter models. While this list of topics is admittedly ambitious, there are a total of 9 weeks of modules. So a 14 or 15 week semester would have extra time for modules that require more than a single week, especially the last couple of modules where the transition from one parameter models to two parameter models can be a little difficult.

An alternative to the Bayesian-themed honors intro stat course, and perhaps a more realistic approach given the constraints on most of these courses, would be to maintain the computationally-focused simulation-based approaches, but use this as a building block for Bayesian ideas. Implicitly, emphasizing computational thinking and the simulation approach would plant the seed for future Bayesian courses, such as those described in Sections 3 and 4. Explicitly, lectures and small snippets could be used to directly link simulation approaches and Bayesian methods. In particular, Baath (2017) provided a set of three video lectures that use approximate Bayesian computation to explain Bayesian inference. These videos are designed to be accessible to the audience that we see in an intro stat course. Furthermore, due to the simulation-based approaches in intro stat, a rising generation of students well versed in computational thinking and simulation approaches will be more prepared for Bayesian ideas and computation after intro to stat.

3. Undergraduate Bayes

3.1. Why Should a Bayesian Course Be Included in the Undergraduate Statistics Curricula

With increasing enrollment in undergraduate statistics programs (American Statistical Association 2015) and more of those students bypassing graduate school and going directly into the workforce, Bayesian statistics should be included in undergraduate statistics curricula. Undergraduate statistics programs that emphasize "mathematical thinking" tend to focus

on preparing students for the mathematical rigor of graduate school. With the emphasis on “computational thinking,” the focus should also include providing students computational thinking skills for analyzing modern datasets with a broad range of tools, including Bayesian methods.

Statisticians, especially those without terminal degrees, tend to be generalists that work on a wide range of problems. The exposure to more, diverse statistical ideas enables better consulting, collaboration and statistical practice. Bayesian statistics is a growing part of statistical practice and science. Hence, undergraduates should have the opportunity to learn Bayesian ideas in the classroom to broaden their skill set, deepen their understanding, and ultimately enhance their hiring appeal.

With an emphasis on computational thinking and tools for analyzing data, predictive modeling and statistical learning ideas should be an important inclusion in modern statistics curricula. One hurdle for statisticians introducing statistical learning frameworks is that there is not a familiar process for handling variability. However, Bayesian predictive modeling approaches for statistical learning, such as tree-based models (Chipman, George, and McCulloch 1998, 2010), retain the familiar accounting for uncertainty through the posterior distribution. In other words, the same data analysis process can be used with the variance expressed through the posterior distribution or the posterior predictive distribution.

In addition to new tools and approaches, a course in Bayesian statistics provides a powerful contrast with the classical approach to statistics. As an instructor of Bayesian courses, the goal of teaching Bayesian courses is not to convince all students to see the world from a Bayesian viewpoint, but rather to help students think deeply about their approach to statistical modeling and by extension, science. While Bayesian methods are not a silver bullet for the p -value dilemma (Wasserstein and Lazar 2016), by providing an alternative to NHST they do provide an opportunity for student to think critically about their own philosophy and approach to statistical analysis.

3.2. How to Include a Bayesian Undergraduate Course in the Stats Curriculum

While Bayesian statistics courses at the undergraduate level are becoming more common, there are options for the best place to put this course in the curriculum. In addition to an intro to stat course infused with Bayesian thinking such as those described in Section 2, we detail two additional options for an undergraduate Bayesian course. A common offering is a senior-level Bayesian statistics course that requires a probability and/or math stat sequence. Traditionally, this course often serves as an elective and is not required for all students. The other option would be a course earlier in the curriculum that would not require probability as a prerequisite.

Teaching a senior-level Bayesian statistics course will guarantee that students have had exposure to a math stat sequence, a methods course, and have previous computing training. This background with mathematical, statistical, and computational expertise will allow advanced treatment of Bayesian modeling and computation. In particular, time can be devoted to the mechanics of MCMC and Bayesian model fitting using stan

and/or JAGS. As students have seen regression in previous methods courses, the transition from one-parameter models to two-parameter models will be fairly intuitive, and furthermore, instructors can focus on understanding and reporting differences between classical and Bayesian analyses. A suitable textbook for this audience might be Bolstad and Curran (2016). Another option is Reich and Ghosh (2019) which is intended for a one semester course for “advanced undergraduates or graduate students.”

One challenge in offering a Bayesian undergraduate course has traditionally been the requirement for a sequence of mathematically rigorous probability and inference courses that often fall in the junior or senior year. This limits the accessibility of the course, and other upper level electives, to senior-level students that have completed the math stat sequence. Furthermore, in smaller programs many electives are offered bi-annually; if the course required a math stat sequence offered during the junior year some students may be out of cycle and not able to take the course in their fourth year. Partially for this reason, Cobb (2015) recommended teaching a broader range of courses that do not require probability. The courses are not necessarily taught at a lower level, but should be data-focused courses and introduce research ideas. Using Cobb’s philosophy, we taught a Bayesian course for undergraduates using Kruschke’s (2014) text. For future versions of the course, we would also consider using Albert and Hu’s (2019) text.

The course that we taught was an elective course designed for undergraduate statistics majors. The course did not require probability, but did require Calculus II and at least one statistics course that included regression concepts. Related to that, Kruschke (2014)’s mathematical prerequisite is stated as

There is no avoiding mathematics when doing data analysis. On the other hand, this book is definitely not a mathematical statistics textbook, in that it does *not* emphasize theorem proving or formal analysis. But I *do* expect that you are coming to this book with a dim knowledge of basic calculus.

A course covering the basics of R is not required, but most students have experience using R in previous statistics courses. A course syllabus is provided in the supplementary materials. The learning outcomes for this course were

- Describe fundamental differences between Bayesian and classical inference;
- Select models and priors, write likelihoods, derive posterior distributions, and verify model and prior assumptions;
- Use computer code, including R, Stan, and JAGS, to sample from posterior distributions; and
- Make inferences from posterior distributions.

The course included a 3–4 week intro to probability, conditional probability, and Bayes’ rule. The course then moved on to discuss exact posterior inference with conjugate priors. MCMC techniques in JAGS were explained and then the course concluded by fitting multi-parameter models through linear and generalized linear models. The course assessments included a mix of quizzes, homework, exams, and projects. The focus of the project was conducting a Bayesian data analysis *and* clearly communicating and reporting the results. The use of project

rubrics and allowing project rewrites provided a way to refine how Bayesian results are reported. The supplemental materials contain a project description and the take home portion of the final exam.

The course primarily used JAGS (Plummer 2003) and stan (Carpenter et al. 2017) for computation, with an emphasis on making inferences from the posterior samples rather than deriving the details of the MCMC algorithms. In future iterations, we'd most likely do model fitting with the `rstanarm` package (Goodrich et al. 2020) as detailed in Gelman, Hill, and Vehtari (2020). This decision would place less emphasis on the MCMC details and the syntax of JAGS or stan and more emphasis on implementing and interpreting Bayesian analyses.

4. Bayes for Graduate Student Scientists

4.1. Why a Bayesian Course for Graduate Student Scientists

While a graduate-level Bayesian methods course is standard in many Statistics M.S. and Ph.D. programs, there generally are not Bayesian methods courses designed for non-statistician scientists; however, Bayesian methods are becoming more commonly used to address scientific research questions. In general, when discussing statistics curricula, little attention is given to courses designed for scientists. While there may be administrative challenges in developing and offering new methods courses designed for non-statistics students, many of these courses are currently taught by statistics faculty and could be adapted to include Bayesian ideas. Furthermore, at some institutions undergraduate statistics majors take these courses concurrently with graduate students from other disciplines.

Despite an increase in the use of Bayesian methods in many fields, including the natural sciences, social sciences, and even the humanities, there has been little change in the delivery of the methods course that serves many graduate student scientists. Often these courses are focused on traditional methods within the purview of the null-hypothesis significance testing framework. As many scientific journals move away from using p -values and statistical significance through the NHST framework, training in alternative approaches will be important. While Bayesian methods would be one alternative, implementation requires specifying a model, formalizing and defending priors, generating posterior samples, and performing model checks. Hence, formal training in Bayesian methods is important for valid scientific inferences.

Graduate students that take statistical methods courses can be extremely influential scientists later in their careers as they start research labs, teach courses, and advise students. While some methods courses are taught by statistics faculty, there are also situations where statistical methods courses are taught in other departments by nonstatistics faculty. Hence, formal training in Bayesian ideas and contrasts with the NHST framework would have a trickle-down effect too.

4.2. How to Teach a Bayesian Course for Graduate Student Scientists

We see two options for teaching Bayesian ideas to graduate student scientists. The first option would be to offer a

Bayesian-centric methods course designed for scientists. There are examples of how to teach the graduate methods sequence from a Bayesian perspective, most famously by Gelman (2008) and the associated textbook, *Bayesian Data Analysis* (BDA), (Gelman et al. 2013), but also Pullenayegum and Thabane (2009) detail the need and development of a Bayesian course for health-scientists. BDA mentions three intended roles for the textbook, one of which is to serve as “an introductory text on Bayesian inference starting from first principles” and another is to be “a handbook of Bayesian methods in applied statistics for general users of and researchers in applied statistics,” both of which are well suited for a methods course for graduate student scientists. The McElreath (2018) textbook is another good textbook intended for this type of audience that also provides a nice set of video lectures that go along with the textbook. Finally, *Regression and Other Stories* (ROS) (Gelman, Hill, and Vehtari 2020) could also be used in this setting. While this text is not overtly Bayesian, most of the methods are implemented from a Bayesian perspective.

Depending on the computing background of the participants and the textbook of choice, there are several ways that computing and MCMC could be conducted in this course. BDA is the most mathematical of the three options and contains fewer code examples. McElreath's text focuses on implementing MCMC in Stan, through R, and includes a comprehensive website with examples. One point from this textbook is that writing out the detailed model along with distributional assumptions requires deeper knowledge of the modeling framework. ROS would require the least amount of MCMC detail, as the `rstanarm` functions (Goodrich et al. 2020) work very similar to the base `lm` or `glm` functions.

We wouldn't envision a complete rethinking of the methods course. The content of the course, the methods themselves, wouldn't change. The course would still cover material from one-parameter models up through linear and generalized linear model frameworks. Rather, computing and software would be presented from a Bayesian perspective. Furthermore, the course would still include a data analysis project, as many methods courses require, but the analysis, results, and interpretation would all be Bayesian.

The second option would be for students to take either the undergraduate course, as described in Section 3, or in rare cases, the course traditionally offered for stats graduate students. We have had success implemented both of these approaches. At many institutions, this would be a more practical choice as developing a new course for out of department students could be logistically difficult. Having statistics students and graduate scientists take the same course provides a nice collaborative setting where, with intentional instruction, the strengths of each group enhance the other group's learning experience. In a similar setting, the textbook by Reich and Ghosh (2019) was developed for a course consisting of “undergraduate statistics majors, non-statistics graduate students from all over campus, and students in the Masters of Science in Statistics program” and would be another suitable textbook choice.

We have taught two courses for graduate student scientists. The first course was explicitly designed for undergraduate statistics students, and is detailed in Section 3, but there were also graduate students from other departments that took the course.

Ideally, moving forward, the course would include both undergraduate statistics students and graduate student scientists. The other course graduate student scientists have taken is one primarily offered for second-year graduate students in the statistics department.

We include the main details about a statistics graduate course here, but also include a syllabus and sample assessments in the supplementary materials. This course requires an inference course, or second semester of a math stat sequence, along with an advanced linear models course, which may limit the accessibility for most graduate student scientists. The course learning outcomes include the following:

- Demonstrate a basic understanding of the fundamental concepts underlying Bayesian inference;
- Demonstrate connections and make comparisons among frequentist, likelihood, and Bayesian methods, both from a practical and a philosophical perspective;
- Demonstrate ability to program methods for taking samples from posterior distributions, including rejection sampling, Metropolis-Hastings algorithm, and Gibbs sampling; and
- Demonstrate ability to use available and common software to carry out Bayesian data analysis.

A complete list of the topics covered is included in the syllabus in the supplementary materials, but the course started with basic Bayesian principles for settings with conjugate priors. The course next moves to discuss Monte Carlo techniques and Markov chain Monte Carlo techniques for linear and generalized linear models. Students employ a combination of their own Gibbs and Metropolis-Hastings samplers written in R along with both JAGS and Stan. The course also covers classical p -values and contrasts Bayesian estimation and testing with null-hypothesis significance testing frameworks from a classical perspective. The course assessments include a mix of quizzes, homework, exams, and a final project. The project focuses on the complete Bayesian data analysis cycle, including summarizing the results. The supplementary materials also contain a project description and the final exam, both the in-class and the take-home portions.

5. Discussion

There are many benefits for teaching Bayesian statistics across the curriculum. In this article we have detailed six benefits: (i) Bayesian ideas and computing may attract a set of students that would otherwise end up in other computationally focused data analysis fields; (ii) The Bayesian data analysis approach can simplify inference; (iii) Bayesian methods give students more tools for data analysis; (iv) Bayesian ideas and philosophy provide a contrast with classical approaches; (v) Bayesian methods provide a natural way to understand uncertainty in statistical learning frameworks; and (vi) Bayesian methods are becoming more common in many scientific fields, and formal training gives practitioners the ability to implement these methods. The list of benefits described here is not intended to be exhaustive. There are many other scenarios that occur in research, such as model selection problems and complicated latent variable models, where Bayesian methods are useful. Rather the intent

of the list is to capture benefits that broadly apply across the statistics curriculum.

The arguments for including Bayesian ideas across the curriculum are not exclusive to the categories we have described. For instance, additional data analysis tools would also benefit scientists, but we placed the benefits where they are *most* applicable. In reality, the benefits of Bayesian methods generally apply to all students regardless of where in the curriculum the ideas are experienced.

In addition to the promoting the benefits of Bayesian ideas, we also described how to implement Bayesian approaches for a set of courses. For each proposed Bayesian course, we have discussed the necessary background for students as well as learning outcomes, course topics, and appropriate textbooks. To help other instructors develop methods, we have included two resources in the supplementary materials. The first resource is a starter guide for teaching Bayesian statistics that contains a list of references for compiling materials for these courses. The second resource is material from our own courses. While these courses are not required at our institution, we strongly feel the undergrad version should be required. Consider flipping the script and only teaching frequentist ideas (NHST) in a single elective course; this sounds ridiculous, but is not all that different from the current treatment of Bayesian ideas.

Beyond including dedicated courses on Bayesian statistics, exposure early in the curriculum gives instructors the ability to include Bayesian approaches in other classes too. Hence, while we discuss specific courses, incorporating Bayesian ideas in other courses, such as a math stat/inference course, would be strongly encouraged as well. Thus, rather than teaching elective courses such as time series analysis, spatial statistics, and categorical data and then having students take a single Bayesian course, both Bayesian and classical ideas can be included in these courses.

Supplementary Materials

The supplementary files contain a compilation of materials that can be used to create Bayesian courses. Monika Hu also maintains a GitHub repo with undergraduate Bayesian education resources: <https://monika76five.github.io/Undergrad-Bayesian-Education-Resources/>.

References

- Albert, J. (2009), *Bayesian Computation with R*, New York: Springer.
- Albert, J., and Hu, J. (2019), *Probability and Bayesian Modeling*, Boca Raton, London, NY: CRC Press.
- Albert, J. and Rossman, A. (2009), *Workshop Statistics: Discovery with Data, a Bayesian Approach*. New York: Springer-Verlag.
- Allenby, G. M. and Rossi, P. E. (2008), "Teaching Bayesian Statistics to Marketing and Business Students," *The American Statistician*, 62, 195–198.
- American Statistical Association (2015), "A Peek Into the Largest, Fastest-growing Undergraduate Statistics Departments," *AMSTAT News*.
- Baath, R. (2017), *Video Introduction to Bayesian Data Analysis*. Publishable Stuff: Rasmus Baath's Research Blog.
- Baglin, J., and Da Costa, C. (2009), "Integrated Statistical Inference: The Amalgamation of Conventional and Bayesian Statistical Inference in Introductory Statistics Courses," in *Third Annual Applied Statistics Education and Research Collaboration Conference (ASEARC)*, pp. 1–4. Applied Statistics Education and Research Collaboration (ASEARC).
- Baumer, B. (2015), "A Data Science Course for Undergraduates: Thinking with Data," *The American Statistician*, 69, 334–342.

- Berry, D. A. (1997), "Teaching Elementary Bayesian Statistics With Real Applications in Science," *The American Statistician*, 51, 241–246.
- Bolstad, W. (2002), "Teaching Bayesian Statistics to Undergraduates: Who, What, Where, When, Why, and How," *Proceedings of the Sixth International Conference on Teaching of Statistics*, pp. 1–6. Citeseer. Available at https://iase-web.org/Conference_Proceedings.php?p=ICOTS_6_2002
- Bolstad, W. M., and Curran, J. M. (2016). *Introduction to Bayesian statistics*. John Wiley & Sons.
- Breiman, L. (2001), "Statistical Modeling: The Two Cultures (with Comments and a Rejoinder by the Author)," *Statistical Science*, 16, 199–231.
- Brown, E. N., and Kass, R. E. (2009), "What is Statistics?" *The American Statistician*, 63, 105–110.
- 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, 1–32.
- Carver, R., Everson, M., Gabrosek, J., Horton, N., Lock, R., Mocko, M., Rossman, A., Roswell, G. H., Velleman, P., Witmer, J., and Wood, B. (2016), Guidelines for Assessment and Instruction in Statistics Education (GAISE), College Report 2016.
- Çetinkaya-Rundel, M. and Rundel, C. (2018), "Infrastructure and Tools for Teaching Computing Throughout the Statistical Curriculum," *The American Statistician*, 72, 58–65.
- Chang, W., Cheng, J., Allaire, J., Xie, Y., and McPherson, J. (2019), *Shiny: Web Application Framework for R*. R package version 1.4.0.
- Chipman, H. A., George, E. I., and McCulloch, R. E. (1998), "Bayesian CART Model Search," *Journal of the American Statistical Association*, 93, 935–948.
- (2010), "BART: Bayesian Additive Regression Trees," *The Annals of Applied Statistics*, 4, 266–298.
- Christensen, R., Johnson, W., Branscum, A., and Hanson, T. E. (2011), *Bayesian Ideas and Data Analysis: An Introduction for Scientists and Statisticians*, Boca Raton, London, NY: CRC Press.
- Clyde, M., Cetinkaya-Rundel, M., Rundel, C., Banks, D., Chai, C., and Huang, L. (2020), *An Introduction to Bayesian Thinking. A Companion to the Statistics with R Course*. self published.
- Cobb, G. (2015), Mere Renovation is Too Little Too Late: We Need to Rethink Our Undergraduate Curriculum from the Ground Up," *The American Statistician*, 69, 266–282.
- delMas, R., Garfield, J., and Chance, B. (1998), "Assessing the Effects of a Computer Microworld on Statistical Reasoning," In *Proceedings of the Fifth International Conference on Teaching Statistics*, pp. 1083–1089. The Netherlands: International Statistical Institute Voorburg.
- Downey, A. (2013). *Think Bayes: Bayesian statistics in python*. "O'Reilly Media, Inc."
- Eadie, G., Huppenkothen, D., Springford, A., and McCormick, T. (2019), "Introducing Bayesian Analysis With m&m's®: An Active-learning Exercise for Undergraduates," *Journal of Statistics Education*, 27, 60–67.
- Franklin, C., Kader, G., Mewborn, D., Moreno, J., Peck, R., Perry, M., and Scheaffer, R. (2007), *Guidelines for Assessment and Instruction in Statistics Education (GAISE) Report* Alexandria: American Statistical Association.
- Garfield, J. and Ben-Zvi, D. (2007), "How Students Learn Statistics Revisited: A Current Review of Research on Teaching and Learning Statistics," *International Statistical Review*, 75, 372–396.
- Gelman, A. (2008), "Teaching Bayes to Graduate Students in Political Science, Sociology, Public Health, Education, Economics," *The American Statistician*, 62, 202–205.
- Gelman, A., Carlin, J. B., Stern, H. S., Dunson, D. B., Vehtari, A., and Rubin, D. B. (2013), *Bayesian Data Analysis*, Boca Raton, FL: Chapman and Hall/CRC.
- Gelman, A., Hill, J., and Vehtari, A. (2020), *Regression and Other Stories*, Cambridge CB2 8BS, UK: Cambridge University Press.
- Gill, J. (2014), *Bayesian Methods: A Social and Behavioral Sciences Approach* (Vol. 20), Boca Raton, FL: CRC Press.
- Goodrich, B., Gabry, J., Ali, I., and Brilleman, S. (2020), "rstanarm: Bayesian Applied Regression Modeling Via Stan," R package version 2.21.1.
- Hardin, J., Hoerl, R., Horton, N. J., Nolan, D., Baumer, B., Hall-Holt, O., Murrell, P., Peng, R., Roback, P., Lang, D. T., and Ward, M. D. (2015), "Data Science in Statistics Curricula: Preparing Students to 'Think With Data,'" *The American Statistician*, 69, 343–353.
- Hicks, S. C. and Irizarry, R. A. (2018), "A Guide to Teaching Data Science," *The American Statistician*, 72, 382–391.
- Hobbs, N. T. and Hooten, M. B. (2015), *Bayesian Models: A Statistical Primer for Ecologists*, Princeton and Oxford: Princeton University Press.
- Hoff, P. D. (2009), *A First Course in Bayesian Statistical Methods* (Vol. 580), Dordrecht, Heidelberg, London, NY: Springer.
- Hu, J. (2020), "A Bayesian Statistics Course for Undergraduates: Bayesian Thinking, Computing, and Research," *Journal of Statistics Education*, 1–18, DOI:10.1080/10691898.2020.1817815.
- Jackman, S. (2009), *Bayesian Analysis for the Social Sciences* (Vol. 846), West Sussex, UK: John Wiley & Sons.
- Jamie, D. M. (2002), "Using Computer Simulation Methods to Teach Statistics: A Review of the Literature," *Journal of Statistics Education*, 10, 1–20.
- Koop, G. M. (2003). *Bayesian Econometrics*, West Sussex, England: John Wiley & Sons Inc.
- Kruschke, J. (2014). *Doing Bayesian Data Analysis: A Tutorial With R, JAGS, and Stan*, Amsterdam, Boston, Heidelberg, London, New York, Oxford, Paris, San Diego, San Francisco, Singapore, Sydney, Tokyo: Academic Press.
- Lambert, B. (2018). *A Student's Guide to Bayesian Statistics*, Los Angeles, London, New Delhi, Singapore, Washington DC, Melbourne: Sage.
- Lock, R. H., Lock, P. F., Morgan, K. L., Lock, E. F., and Lock, D. F. (2016), *Statistics: Unlocking the Power of Data*, Hoboken, NJ: John Wiley & Sons.
- Marin, J.-M. and Robert, C. P. (2014), *Bayesian Essentials With R* (Vol.48), New York, Heidelberg, Dordrecht, London: Springer.
- McCarthy, M. A. (2007), *Bayesian Methods for Ecology*, New York, NY: Cambridge University Press.
- McElreath, R. (2018), *Statistical Rethinking: A Bayesian Course With Examples in R and Stan*, Boca Raton, FL: Chapman and Hall/CRC.
- Moore, D. S. (1997), "Bayes for Beginners? Some Reasons to Hesitate," *The American Statistician*, 51, 254–261.
- Plummer, M. (2003), "Jags: A Program for Analysis of Bayesian Graphical Models Using Gibbs Sampling," in *Proceedings of the 3rd International Workshop on Distributed Statistical Computing*, vol. 124, p. 10. Vienna, Austria.
- Pullenayegum, E. M., and Thabane, L. (2009), "Teaching Bayesian Statistics in a Health Research Methodology Program," *Journal of Statistics Education*, 17.
- Reich, B. J., and Ghosh, S. K. (2019), *Bayesian Statistical Methods*, Boca Raton, FL: CRC Press.
- Robert, C. (2007), *The Bayesian Choice: From Decision-theoretic Foundations to Computational Implementation*. Springer Science & Business Media.
- Rossman, A. J. and Chance, B. L. (2014), "Using Simulation-based Inference for Learning Introductory Statistics," *Wiley Interdisciplinary Reviews: Computational Statistics*, 6, 211–221.
- Stangl, D. (1998), "Classical and Bayesian Paradigms: Can We Teach Both? In *Proceedings of the Fifth International Conference on Teaching Statistics*, eds. L. Pereira-Mendoza, L. S. Kea, T. W. Kee, and W. Wong, International Statistics Institute, vol. 1, pp. 251–258. Citeseer. Available at https://iase-web.org/Conference_Proceedings.php?p=ICOTS_5_1998
- Steel, E. A., Liermann, M., and Guttrop, P. (2019), "Beyond Calculations: A Course in Statistical Thinking," *The American Statistician*, 73, 392–401.
- Tintle, N. L., Rogers, A., Chance, B., Cobb, G., Rossman, A., Roy, S., Swanson, T., and VanderStoep, J. (2014), *Quantitative Evidence for the Use of Simulation and Randomization in the Introductory Statistics Course*. In *Proceedings of the Ninth International Conference on Teaching Statistics, volume ICOTS-9*. Available at http://iase-web.org/icots/9/proceedings/pdfs/ICOTS9_8A3_TINTLE.pdf
- Utts, J., and Johnson, W. (2008), "The Evolution of Teaching Bayesian Statistics to Nonstatisticians: A Partisan View from the Trenches," *The American Statistician*, 62, 199–201.
- Wasserstein, R. L., and Lazar, N. A. (2016), "The ASA's Statement on p-values: Context, Process, and Purpose," *The American Statistician*, 70, 129–133.
- Witmer, J. (2017), "Bayes and MCMC for Undergraduates," *The American Statistician*, 71, 259–264.