Resources to Learn Bayesian Stats

Ilan Reinstein

2020-06-09

In this post, I summarize a series of resources to get started with Bayesian Statistics. I compiled these references myself based on my own experience and opinion as to what a good introduction and next steps are in this process. This is not an academic curriculum or anything tremendously rigorous, but it is a comprehensive list that will surely get you embarked on the journey to revisiting/starting your statistics. Many of the references below were recommened to me in several workshops I’ve attended and I wanted to share with those like me that want to be better at statistics and Machine Learning (ML).

The first resource I can think of out there for beginners interested in Bayesian statistics and modeling is Richard McElreath’s [Statistical Rethinking](https://xcelab.net/rm/statistical-rethinking/). Here you’ll learn everything from applying Bayes’ rule in simple problems to complex multilevel/hierarchical models. Since this is not only a book but also a full course, you should definitely follow along your reading with the [video lectures on YouTube](https://www.youtube.com/playlist?list=PLDcUM9US4XdNM4Edgs7weiyIguLSToZRI). I would encourage you to work through the code and exercises at the end of each chapter, it is the best way to start getting hands on expertise in this topic. Another great thing is that the book’s R package [rethinking](https://github.com/rmcelreath/rethinking) and all the example code has been translated into multiple programming languages, so you are free to choose the one of your preference to go through the reading.

A powerful reference for those that feel like brushing up your stats with real-world examples, Andrew Gelman and Jennifer Hill’s [Data Analysis Using Regression and Multilevel/Hierarchical Models](http://www.stat.columbia.edu/~gelman/arm/) is a fantastic book for those with more interest on Applied Statistics on a social science setting. An updated version of this book is called [Regression and Other Stories](https://avehtari.github.io/ROS-Examples/) (Andrew Gelman, Jennifer Hill, Aki Vehtari) and it will be released later this year. If you’d like to work through another more advanced course on Bayesian Statistics, I suggest you visit [Aki Vehtari’s teaching page](https://github.com/avehtari/BDA_course_Aalto). Although more challengin than McElreath’s class it is worth checking it out. Again, the course material is available in R and Python.

Now, you’ve refreshed your basic stats and gone through most of McElreath’s book, but you are also looking into ML. You may wonder, where should I go? What to do next? Challenge yourself to read Christopher Bishop’s [Pattern Recognition and Machine Learning](https://www.amazon.com/Pattern-Recognition-Learning-Information-Statistics/dp/0387310738). Bishop’s book presents all the well known ML concepts from regression and classification to neural networks through a concealed Bayesian lense. It is fair to say that I’ve deeply strengthen my knowledge in both ML and Bayesian statistics just by working through the first chapters. Although I highly recommend it, keep in mind this is an advanced book and it takes some time to get your head around the concepts.

At the same level as Bishop’s book, you can also find a rigorous and detailed explanation of Bayesian statistics and modeling on [David MacKay’s Information Theory, Inference, and Learning Algorithms](http://www.inference.org.uk/mackay/itila/book.html). I was fortunate enough to use this book in college, and it still proves to be essential whenever I want to revisit some probability concepts, Bayesian statistics and ML. The book’s webpage also provides software resources and examples for you to experiment and play around as you navigate the text.

Whenever you feel comfortable enough with the basics to go deeper into the math and the theoretical concepts behind probabilistic modeling, probabilistic computation and applied Bayesian Statistics I strongly encourage you to read *all* of [Michael Betancourt’s case studies](https://betanalpha.github.io/writing/) in due order. This will be challenging to grasp at first, but as you go back over the content and dedicate the time to understand these concepts, the reward is huge.

It is not very often that you hear about the success stories of Bayesian methods in industry, specially nowadays against the huge amount of attention on ML and AI. For those that like podcasts to listen to stories about practitioners and people applying Bayesian statistics to their fields I strongly advise that you check out Alexandre Andorra’s podcast Learning Bayesian Statistics. Alex does an amazing job at interviewing professionals from a wide variety of fields, with incredibly diverse backgrounds and experiences, hence introducing you to the vast universe of applications of Bayesian statistics.

One final thing that it is a hard requirement and a common thread between the references and resources listed above is Markov Chain Monte Carlo (MCMC). In order to fully explore probability spaces and distributions you need efficient and reliable computational methods like MCMC and its advanced variants. Today, many programming languages are capable of implementing such advanced estimation algorithm but the most popular are 1) [Stan](https://mc-stan.org), which is built on C++ and has multiple interfaces to R ([rstanarm](http://mc-stan.org/rstanarm/), [brms](https://paul-buerkner.github.io/brms/)), Python ([PyStan](https://pystan.readthedocs.io/en/latest/)), Julia and others; and 2) [PyMC3](https://docs.pymc.io). If you are interested in learning the basics you may visit their webpages to see examples with code.

Some personal remarks on my own journey through Bayesian Statistics. This topic is not easy, you should invest some time to see some progress. The references above are a great starting point and I’ve noticed an important step forward in my learning path by trying to apply Bayesian methods to my own work. For me, its fundamentals are more intuitive and transparent, and overall simpler to grasp (at least conceptually), but considerably harder to apply.