# Book outline

Bayesian analysis is the most intuitive theory of statistics, with an applicability that is wider than currently covered by ‘traditional’ methods. It is however easy to get lost in the relatively advanced maths on which the subject is based. The maths poses such a sufficient challenge to those wanting to learn Bayesian methodology that many students give up after attempting to get to grips with the theory. This is especially common for students of the social sciences; almost everyone has a general idea of the concepts inherent in Bayesian statistics, but very few actually practice it.  
  
Current literature aimed at students generally falls into one of three categories: most often the texts are very mathematical (examples being, “Bayesian Data Analysis” by Gelman, “Bayesian Analysis for the Social Sciences” by Jackman, and Koop’s “Bayesian Econometric Methods”), with limited practical examples, limiting their appeal to students; alternatively the texts are overly basic (examples include, “Bayes’ rule: A Tutorial Introduction to Bayesian Analysis” by Stone), or fail to provide a sufficiently broad introduction to the field (for example, “Doing Bayesian Data Analysis” by Kruschke).  
  
It is my view that a book can be written which places an emphasis on the intuition not the maths behind Bayesian theory, with a focus on practical examples, yet without compromising on the breadth of material covered. With this in mind, the book would follow the following principles:  
  
1. The student should be able to come away from reading a chapter, and feel that they would be able to apply the theory learned to real world data; the central aim of the book is to provide a practical guide to modern Bayesian statistical techniques.  As such, the book will provide data-based examples using the open source BUGS software.  
  
2.    The level of mathematics will be kept as simple as possible. There will of course be occasions when it is necessary to introduce mathematical concepts which are *prima facie* difficult to understand. In these circumstances, a very low level of prior knowledge will be assumed, and an emphasis will be placed instead on understanding the intuition behind the equations.  
  
3.    Whenever new theory is introduced, it will be supplemented with one or more examples.  
  
4.    New concepts will be explained through the use of videos. Bayesian theory is full of intuition, even though it is sometimes difficult to convey this through traditional static media. Videos will not merely be a recapitulation of the theory and examples explained in the text. Wherever possible the material in the videos will be unique, and complementary to the text.  
  
5.    After reading the book the student should be able to read the majority of modern research papers in the social sciences which make use of Bayesian statistics, and be able to understand the theory contained within them. They should also have an idea as to how to go about replicating the research practically.  
  
6.    The chapters will be as self-contained as possible. This principle is guided by the fact that students will often use the text for reference with specific practical goals in mind. Of course, the book will be most effective when read in order, but attempts will be made to limit the reliance of each chapter on those preceding it.  
  
7.    Problem sets will be focused on analysis of real data. The data will be introduced in the text, and by video.  
  
The text produced would be an ‘essential textbook’, which when followed will provide a comprehensive self-contained introduction to the subject. Whilst the time taken to cover the material contained within the book would be institution-dependent, the course covered would span over one or two semesters. The book would be usable as a source of information for undergraduates, graduates and professionals in the field. It would also have appeal as a ‘supplementary book’ since it would be full of examples, and due to the self-contained nature of chapters. I would anticipate that the length of the book would be of order 300 pages. It is my belief that whilst more pages provide more room for explaining theory, students are perhaps discouraged by lengthy texts.

# Book contents

This section is a fairly brief description of how I would perhaps order the book. It is split into sections which have overarching aims, which are accomplished bit-wise in each of their chapters.

1. The purpose of the book, and how best to use it
   1. The goal of this book: after reading it, the student should be able to understand the majority of the Bayesian statistical methods that are used in modern applied social sciences research papers. Furthermore they should have an idea about how to practically go about recapitulating their results.
   2. How to find the videos associated with the relevant material.
   3. How to obtain BUGs/WinBugs. Note this will be a *very* comprehensive guide to getting the software to working starting from how to install the software (across Windows, Apple, and Linux.)
   4. How to use the problem sets. Where to find the videos associated with each of these datasets.

**Part I: Understanding the Bayesian formula**

This section is devoted to developing an understanding of the central formula in Bayesian statistics. The first chapter will explain the purpose of Bayesian statistics, and highlight its differences with classical statistics; culminating with an introduction of the Bayesian formula. The second chapter (of this section) will explain the first part of the Bayesian formula: the likelihood. The third chapter explains the second part of the Bayesian formula: the prior distribution. The final chapter of this section will introduce the student to the last component necessary to build the posterior distribution: the denominator of the Bayesian formula.

1. The subjective and the seemingly objective: An introduction to classical and Bayesian statistics
   1. The goal of this chapter: Introduce the purpose of statistics in general; classical statistics will be compared and contrasted with Bayesian statistics; the tangible (non-academic) benefits of Bayesian methods will be highlighted; finally, the Bayesian formula will be introduced, defining each of its elements.
   2. The purpose of statistics.
   3. The world according to classical statistics.
   4. The Bayesian central dogma: The Bayesian formula (along with a short biography of Bayes – the tragedy of posthumous success).
   5. An introduction to the Bayesian inference process: Choose a model for data (specify a likelihood); What do you know about the situation? (specify a prior). These two choices together result in ‘updated’ knowledge about the situation (the posterior distribution.)
   6. The intuition behind the Bayesian formula.
   7. How do classical and Bayesian theories address the purpose of statistics?
   8. What is a probability? The flexibility of the Bayesian notion opposed to the classical (frequentist) view.
   9. Explicit vs implicit subjectivity: the danger of the word ‘objective’. This would be a critique of the notion of objectivity commonly thought to hold in frequentist statistics.
   10. What are the tangible (non-academic) benefits of Bayesian theory vs classical statistics? It yields the best predictions, allows easier interpretation of results, and is more intuitive.
   11. Why don’t more people use Bayesian statistics? Existing literature is very heavy with the advanced mathematics behind Bayesian theory. For many students, continuing with using classical statistics appears to be the path of least resistance.
   12. Chapter conclusion: where we are in understanding the Bayesian formula.
   13. An introduction to the material of the next few chapters: introducing the elements of the Bayesian formula.
   14. Problem set introduction.
2. Choosing an appropriate model for the data: specifying a likelihood
   1. The goal of this chapter: introduce the concept of likelihood; explain how to choose a likelihood; explain the idea behind maximum likelihood estimation.
   2. What is a likelihood?
   3. Why likelihood is not a probability.
   4. An aside to a classical method: maximum likelihood estimation
      1. Maximising the ‘probability’ of obtaining the sample from a population.
      2. Example: what’s the probability someone has blood type B? Writing down likelihood.
      3. Why do we maximise log likelihood?
      4. Example continued: Estimating the probability that an individual has blood type B.
      5. How to estimate uncertainty of maximum likelihood estimates? The Cramer-Rao Lower Bound and its intuition.
      6. Example continued: Estimating the imprecision of estimated probability that an individual has blood type B.
   5. How to choose a likelihood appropriate to the situation.
   6. The subjectivity of model choice.
   7. Chapter summary. An update of where we are in part I: understanding the Bayesian formula.
   8. Problem set introduction.
3. The prior: representing your pre-investigation knowledge of the phenomena in question
   1. The goal of this chapter: introduction to the concept of a prior; an introduction as to how to specify a prior; a short section on objective Bayesian priors.
   2. Introduction to priors. A way of representing current knowledge of the situation.
   3. Example: probability an individual has blood type B. What values would we consider plausible beforehand? How we can express our uncertainty in a prior?
   4. How to specify a prior. A guide as to how to write down a prior in a given situation. This will only be a short section, as it is better explained later on after more theory has been covered.
   5. Objective Bayesian priors. What if we have no idea about the value of a model’s parameters? This concept will be only explained in brief, as the purpose isn’t to go into depth about how to carry out Objective Bayesian inference. It is more to put the student at ease about the supposed subjectivity of Bayesian methods.
   6. ‘Good’ model robustness. The predictions and inferences from a good model should be robust to changes in the prior. The prior becomes less important the more data is available.
   7. Chapter summary. An update of where we are in part I: understanding the Bayesian formula.
   8. Problem set introduction.
4. The denominator
   1. The goal of this chapter.
   2. What does the denominator mean? This is a short introduction to the meaning of the denominator as the probability of the data (given a choice of model). The denominator as a weighted average. The denominator as a nuisance normalising constant.
   3. The difficulty with the denominator. Why is the denominator often hard to evaluate?
   4. Example: the probability an individual has blood type B. How to evaluate the probability of obtaining the data.
   5. Example made more complex. The addition of more parameters to the previously simple probability example to highlight the increase in complexity of evaluating the denominator, especially in the presence of a high degree of uncertainty surrounding the priors.
   6. The difficult made easy. Modern Bayesian (computational) methods ignore denominator, and can still make exact simulations from the posterior distribution. This will not be a full introduction to, for example, Metropolis-Hastings, but is intended to convey the message, ‘not all hope is lost’, to the student.
   7. Chapter summary. Description of the fact that we have now covered all the parts of the Bayesian formula, and are now ready to start applying it.

**Part II: A practical guide to doing (and understanding) Bayesian analysis**

This part of the book will introduce the student to the practicalities of doing Bayesian statistics. The first part of the section will focus on analytical Bayesian data analysis: conjugate priors, calculation of posteriors, and prediction of dependent variables. The reason behind this first half is partly to gain experience with using Bayesian theory on datasets, and also to allow the student to be able to read and understand a significant proportion of the texts and papers written using this type of statistical analysis. The second part of this section will focus on real life applications of Bayesian analysis; straying away from the comfort and analytical (over-) simplicity of the distributions and assumptions used in the first half of this chapter, by introducing the student to computational Bayesian analysis (algorithms used to simulate draws from the posterior distribution) using BUGs. I believe that explanations of the algorithms on which modern Bayesian analysis are built is often where existing texts falter. I also believe that, fortunately, these algorithms are intuitive, and hence can be explained in a manner which can be understood by anyone.

1. An introduction to well known (and often used) probability distributions.
2. Conjugate priors and their uses in Bayesian data analysis.
3. How to forecast in Bayesian statistics.
4. An introduction to BUGs.
5. Leaving conjugate priors behind: the motivation for simulating draws from the posterior distribution.
6. Computational Bayes introduction part 1: Grid approximations.
7. Computational Bayes introduction part 2: the Metropolis-Hastings algorithm.
8. Computational Bayes introduction part 3: the Gibbs sampler.

**Part III: Regression analysis and hierarchical models**

The first part of this section will introduce the reader to some of the quirks and benefits of regression analysis with Bayesian statistics. The benefits in part come as a result of the use of hierarchical models within a regression framework; leading onto the second half of the section, which focuses on hierarchical models. This section will focus on simple (cross sectional) linear regression models, allowing the student to gain practice in Bayesian regression, before moving onto more complex models in the next part.

**Part IV: Hypothesis testing and evaluating model fit**

This section will aim to explain state-of-the art hypothesis testing in Bayesian statistics. It will focus first of all on the analogues of hypothesis tests in frequentist statistics, and will then move onto novel techniques and concepts. In this part I will give ample comparison of model building in classical vs Bayesian statistics; extolling the benefits of the latter.

**Part V: Regression analysis beyond simple cross sectional data**

This section will build on the previous section by introducing the student to multilevel models on different types of data and situations: generalised linear modelling, time series data and panel data. After this part, the student should be comfortable with reading the majority of texts and research papers which use Bayesian analysis. Instead of focussing on the (somewhat involved) theory behind these applications of Bayesian theory, each of the chapters will be built around case studies.

**Part VI: Objective Bayesian data analysis**

Considerable attention in modern literature is around attempts to make Bayesian analysis, ‘as objective as possible’. I anticipate that this will be an area of increasing importance in coming years. As such, I would like to devote a part of this book on the principles of so-called *Objective Bayesian data analysis.* Individual chapters will introduce the reader to the Jeffrey’s prior, reference priors and Zellner’s G-priors. This section will necessarily be a little more focussed on the philosophy and theory behind Bayesian analysis, but will continue to be grounded in data-based examples.

**Part VII: Advanced simulation methods**

This last section of the book will introduce the reader to the most modern methods for simulating draws from the posterior distribution. The individual chapters will focus on particular techniques for speeding up simulations of multilevel, multivariate models. The BUGs code (along with any other supplementary scripts in R) will be available, and all theory will be introduced through means of examples which use real world data. It is my belief that this section of existing textbooks is often given less emphasis than its importance (especially in modern applications of Bayesian theory) merits. It is my hope that by devoting a significant part of the book, suitably brought to life through examples, that this text can avoid falling into the same trap.