Bayesian data analysis – reading instructions

Aki Vehtari

Chapter 1 – outline

Outline of the chapter 1

- 1.1-1.3 important terms, especially 1.3 for the notation
- 1.4 an example related to the first excerise, and another practical example
- 1.5 foundations
- 1.6 good example related to visualisation exercise
- 1.7 example which can be skipped
- 1.8 background material, good to read before doing the first exercises
- 1.9 background material, good to read before doing the second exercises
- 1.10 a point of view for using Bayesian inference

Chapter 1 – most important terms

Find all the terms and symbols listed below. Note that some of the terms are now only briefly introduced and will be covered later in more detail. After the course you should be able to explain all these terms (these are part of the final exam). When reading the chapter, write down questions related to things unclear for you or things you think might be unclear for others. After reading the chapter, read the file comments1.pdf

- full probability model
- posterior distribution
- potentially observable quantity
- quantities that are not directly observable
- exchangeability
- independently and identically distributed
- $\theta, y, \tilde{y}, x, X, p(\cdot|\cdot), p(\cdot), \Pr(\cdot), \sim, H$
- sd, E, var
- Bayes rule
- prior distribution
- sampling distribution, data distribution
- joint probability distribution
- posterior density
- probability
- density
- distribution
- $p(y|\theta)$ as a function of y or θ
- likelihood
- posterior predictive distribution

- probability as measure of uncertainty
- subjectivity and objectivity
- transformation of variables
- simulation
- inverse cumulative distribution function

Model and likelihood

Term $p(y|\theta, M)$ has two different names depending on the situation. Due to the short notation used, there is possibility of confusion.

1) Term $p(y|\theta, M)$ is called a *model* (sometimes more specifically *observation model* or *statistical model*) when it is used to describe uncertainty about y given θ and M. Longer notation $p_y(y|\theta, M)$ shows explicitly that it is a function of y.

2) In Bayes rule, the term $p(y|\theta, M)$ is called *likelihood function*. Posterior distribution describes the probability (or probability density) for different values of θ given a fixed y, and thus when the posterior is computed the terms on the right hand side (in Bayes rule) are also evaluated as a function of θ given fixed y. Longer notation $p_\theta(y|\theta, M)$ shows explicitly that it is a function of θ. Term has it's own name (likelihood) to make the differene to the model. The likelihood function is unnormalized probability distribution describing uncertainty related to θ (and that's why Bayes rule has the normalization term to get the posterior distribution).

Two types of uncertainty

Epistemic and aletory uncertainty are reviewed nicely in the article: Tony O'Hagan, "Dicing with unknown" Significance 1(3):132-133, 2004. Link to the article is in the course web page.

Transformation of variables

- Logic, Probability, and Bayesian Inference by Michael Betancourt `https://github.com/betanalpha/stan_intro/blob/master/stan_intro.pdf`