Bayesian data analysis – reading instructions 3

Aki Vehtari

Chapter 3

Outline of the chapter 3

- 3.1 Marginalisation
- 3.2 Normal distribution with a noninformative prior (very important)
- 3.3 Normal distribution with a conjugate prior (very important)
- 3.4 Multinomial model (can be skipped)
- 3.5 Multivariate normal with known variance (needed later)
- 3.6 Multivariate normal with unknown variance (glance through)
- 3.7 Bioassay example (very important, related to one of the exercises)
- 3.8 Summary (summary)

Normal model is used a lot as a building block of the models in the later chapters, so it is important to learn it now. Bioassay example is good example used to illustrate many important concepts and it is used in several exercises over the course.

R and Python demos at https://avehtari.github.io/BDA_course_Aalto/demos.html

- demo3_1: visualise joint density and marginal densities of posterior of normal distribution with unknown mean and variance
- demo3_2: visualise factored sampling and corresponding marginal and conditional density
- demo3_3: visualise marginal distribution of mu as a mixture of normals
- demo3_4: visualise sampling from the posterior predictive distribution
- demo3_5: visualise Newcomb's data
- demo3_6: visualise posterior in bioassay example

Find all the terms and symbols listed below. When reading the chapter, write down questions related to things unclear for you or things you think might be unclear for others. See also the additional comments below.

- marginal distribution/density
- conditional distribution/density
- joint distribution/density
- nuisance parameter
- mixture
- normal distribution with a noninformative prior
- normal distribution with a conjugate prior
- sample variance
- sufficient statistics
- μ,σ^2,$\bar{y}$,s^2
- a simple normal integral

- Inv-χ^2
- factored density
- t_{n-1}
- degrees of freedom
- posterior predictive distribution
- to draw
- N-Inv-χ^2
- variance matrix Σ
- nonconjugate model
- generalized linear model
- exchangeable
- binomial model
- logistic transformation
- density at a grid

Conjugate prior for Gaussian distribution

BDA p. 67 (BDA3) mentions that the conjugate prior for Gaussian distribution has to have a product form $p(\sigma^2)p(\mu|\sigma^2)$. The book refers to (3.2) and the following discussion. As additional hint is useful to think the relation of terms $(n-1)s^2$ and $n(\bar{y}-\mu)^2$ in 3.2 to equations 3.3 and 3.4.

Trace of square matrix

Trace of square matrix, trace, $\mathrm{tr}\,A$, $\mathrm{trace}(A)$, $\mathrm{tr}(A)$, is the sum of diagonal elements. To derive equation 3.11 the following property has been used $\mathrm{tr}(ABC) = \mathrm{tr}(CAB) = \mathrm{tr}(BCA)$.

History and naming of distributions

See *Earliest Known Uses of Some of the Words of Mathematics* http://jeff560.tripod.com/mathword.html.

The number of required Monte Carlo draws

This will discussed in chapter 10. Meanwhile, e.g., 1000 draws is sufficient.

Bioassay

Bioassay example is is an example of very common statistical inference task typical, for example, medicine, pharmacology, health care, cognitive science, genomics, industrial processes etc.

The example is from Racine et al (1986) (see ref in the end of the BDA3). Swiss company makes classification of chemicals to different toxicity categories defined by authorities (like EU). Toxicity classification is based on lethal dose 50% (LD50) which tells what amount of chemical kills 50% of the subjects. Smaller the LD50 more lethal the chemical is. The original paper mentions "1983 Swiss poison Regulation" which defines following categories for chemicals orally given to rats (mg/ml)

Class	LD50
1	<5
2	5-50
3	50-500
4	500-2000
5	2000-5000

To reduce the number of rats needed in the experiments, the company started to use Bayesian methods. The paper mentions that in those days use of just 20 rats to define the classification was very little. Book gives LD50 in log(g/ml). When the result from demo3_6 is transformed to scale mg/ml, we see that the mean LD50 is about 900 and $p(500 < \text{LD50} < 2000) \approx 0.99$. Thus, the tested chemical can be classified as category 4 toxic.

Note that the chemical testing is moving away from using rats and other animals to using, for example, human cells grown in chips, tissue models and human blood cells. The human-cell based approaches are also more accurate to predict the effect for humans.

logit transformation can be justified information theoretically when binomial likelihood is used.

Example codes in demo3_6 can be helpful in exercises related to Bioassay example.

Bayesian vs. frequentist statements in two group comparisons

Frank Harrell's recommendations how to state results in two group comparisons are excellent http://www.fharrell.com/2017/10/bayesian-vs-frequentist-statements.html.