STAT40380/40390 - Bayesian Analysis

Computer Lab 2 – Fitting models in WinBUGS with different prior distributions

1 Revision: running WinBUGS

Remember the steps to run a WinBUGS file:

1. Click Model > Specification
2. Highlight the word `model` and click 'check model'
3. Highlight the word `list` in the data section and click 'load data'
4. Click 'compile'
5. Highlight the word `list` in the initial values section and click 'load inits'
6. Click Inferences > Samples
7. Type the name of the parameters you are interested in and click 'set' after each one.
8. Choose the number of updates (the default 1000 is fine for simpler models) and Click Model > Update
9. Click 'Update'

When WinBUGS runs, it produces *samples* from the posterior distribution. We will cover more about how these samples are created in lectures, but for now we can use them to explore the posterior distributions of various models and data sets we might like to fit.

Task:

- Open the `dating.odc` example and run it for 10,000 iterations. Using history, density and stats, look at the posterior sample properties

2 Files with more than one observation

It is relatively simple to create WinBUGS code to run models with multiple observations. We simply need to include a loop. In the `chest.odc` example (taken from the chest measurements example in Lecture 2), we specify the likelihood using a `for` statement:

```
for(i in 1:N) {
  x[i] ~ dnorm(theta,phi.inv)
}
```

This looping statement says that, for each different observation (up to observation `N`), each data point comes from a normal distribution with mean `theta` and precision `phi.inv`. To complete the syntax we need to give a value for `N` and a list of the data (comma-seperated and surrounded by `c(...)`). These parts can be found at the end of the file.

Task:

- Run the `chest.odc` example for 10,000 iterations, and confirm the posterior distribution of θ as approximately $N(\bar{x}, \phi/n)$. (note: $\bar{x}$ here is 39.8)

3 Reference and flat prior distributions

WinBUGS allows for lots of different probability distributions, including the normal, uniform, exponential, gamma, Poisson and Binomial. If, however, we want to specify an improper prior distribution, two of the most useful options are:

- `dflat()` for a flat prior distribution, for example on a mean parameter
- `dgamma(0.01,0.01)` ie a gamma distribution with shape and rate as very small values. Recall that the gamma distribution has pdf: $p(x|\alpha, \beta) \propto x^{\alpha-1}e^{-\beta x}$, thus

$$\lim_{\alpha,\beta\downarrow 0} p(x|\alpha, \beta) \propto x^{-1}$$

  Hence this distribution can be used as the Jeffreys' prior on a variance or precision parameter.

> Task:
>
> - Change the prior distribution for the dating and chest examples to use a flat prior distribution on the mean. What is the effect on the posterior distributions?

4 Multi-parameter models

WinBUGS can easily handle models with many parameters. In such circumstances, the only change we need to make is to ensure we specify prior distributions and starting values for each unknown parameter. When running the model, be sure to type the parameters into the 'samples' dialog box individually and click 'set' after each. WinBUGS will now produce joint samples of all the parameters and will summarise them individually (ie provide their marginal posterior distributions) through the sample monitor tool.

> Task:
>
> - Open the file `rats.odc` and check that the formulation of the problem corresponds to slides 8 and 9 of lecture 6. Run the model and show that the 95% posterior credible intervals for the mean and variance are approximately (18,24) and (20,74) respectively.

5 Comparing different prior distributions

A useful trick for comparing models with differing prior distributions is shown in the file `Chest2.odc`. Here, we compare 2 different prior distributions (an informative $N(38, 9)$ prior and a flat prior) on the parameter of interest θ. The script works by looping through the two models (with index `j`) and creating a separate data set for each. This new data set (called `x2`) contains the data in a matrix format such that each row of the matrix comes from a normal distribution with a different mean (`theta`). There are now two prior distributions, one for each element of the vector `theta`.

After running the model, a neat way of comparing the two parameters is to choose the Inference > Compare dialog box, and enter `theta` in the 'node' box. Clicking 'box plot' or 'caterpillar' will produce a boxplot or credible interval comparison for the values of `theta` under the two different prior distributions.

> Task:
>
> - Run the file `Chest2.odc` for 10,000 iterations. Use the 'compare' dialog box to show that the flat prior distribution and the informative prior distribution yield very similar posterior values of θ.

6 Homework exercises

Produce a short report (in either word, open office or pdf format) which answers the following questions and includes all used WinBUGS code and appropriate plots:

1. Write a WinBUGS file to fit the Poisson-Gamma misprints example of Lecture 4. Show that the results obtained are identical to those found in class

2. Create another script to enable comparison for 3 different prior distributions:

   (a) The $Ga(9, 6)$ prior used originally

   (b) A $U(0, 10)$ prior distribution

   (c) The reference prior for the Poisson distribution (hint: the reference prior for the Poisson distribution with parameter λ is $p(\lambda) \propto \sqrt{1/\lambda}$. Use an appropriate gamma distribution to approximate this prior)

   Use the compare dialog box to show the different posterior distributions under each of these assumptions. Which of the prior distributions is most informative? Why might two of these distributions be similar?

3. Some more misprints are observed on the following pages as 2,1,6, and 0. Use your posterior distribution created in (1) as a prior distribution for this new data. Show that the new posterior distribution is the same as that if the data had all been observed simultaneously.