Bayesian Data Analysis
WinBUGS exercise 1

Consider early investigation of a new drug, with the aim to study its success rate.
Data. Suppose we treat $n = 20$ volunteers with the compound and observe $y = 15$ positive responses: $y \sim Bin(n, \theta)$.
Prior distribution. Experience with similar compounds has suggested that response rates between 0.2 and 0.6 could be feasible; this can be interpreted this as a distribution with mean = 0.4, standard deviation 0.1. A Beta distribution $B(9.2, 13.8)$ has these properties, so use this as the prior distribution: $\theta \sim B(9.2, 13.8)$.

1. Fit the model using WinBUGS.

   The WinBUGS code for fitting the model to the drug data can be found in files drug-model.odc. The data are in file drug-dat.odc and 2 sets of initial values are in files drug-in1.odc and drug-in2.odc respectively. Open these files and carry out a WinBUGS run for this model.

   i. Check and compile the model, following the instructions in section Running a model in WinBUGS of the hints handout. Note that you should specify the number of chains to be run before you compile the model. (You should run 2 chains using the 2 different sets of initial values).

   ii. Before you start updating, make sure you set sample monitors for the drug success rate θ, the odds *odds*, and the indicator $p.high$ of whether θ is greater than 0.5 (See section Monitoring parameter values of the hints handout).

   iii. Run 1000 iterations (updates), then look at history plots and autocorrelation plots of the sample traces. Do the simulations look like they have converged? If not, carry out some more updates and check again. (See section Checking convergence of the hints handout).

   iv. Once you are happy with convergence, carry out a further 10000 iterations (or more if you wish) to obtain samples from the posterior distribution of the model parameters.

   v. Produce summary statistics for all the variables you have monitored. Remember to discard any burn-in samples first. Check the Monte Carlo standard error of θ and *odds* to assess the accuracy of your estimates. By changing the values of beg and end (which define the start and end iteration of the subset of samples you wish to use for posterior summaries), examine how the MC error depends on the posterior sample size. (See also section Obtaining summary statistics of the posterior distribution of the hints handout).

[Please turn over]

2. Provide point and interval estimates of θ:

   the posterior mean $E(\theta \mid y) =$

   the 95% credible interval for θ is

   Produce kernel density plots of the posterior distribution for θ.

3. Find the posterior mean and 95% credible interval for posterior odds in favour of the drug $odds = \frac{\theta}{1-\theta}$:

   $E(odds \mid y) =$

   95% credible interval for odds is

4. Find the posterior probability $P(\theta \geq 0.5 \mid y) =$.

   What conclusion can you make about the success rate of the drug?

5. Edit the model code to specify a Uniform(0, 1) prior on the success rate θ, and re-run the analysis. (Note: the syntax for the uniform prior in WinBUGS is dunif(a, b) where a and b are the lower and upper bounds.) How is the posterior estimate of θ affected? How is the probability $P(\theta \geq 0.5 \mid y)$ is affected?

   $E(\theta \mid y) =$, the 95% credible interval for θ is

   $E(odds \mid y) =$, 95% credible interval for odds is

   $P(\theta \geq 0.5 \mid y) =$