Name: Solutions

Bayesian Statistics, 22S:138
Midterm 2, 2010

Eight college statistics majors are playing basketball together for the first time. Just for the fun of it, they decide to use Bayesian methods to estimate each individual student's success probability of making a basket from 10 feet away, as well as the overall average success probability of the group. They gather data in the following way:

One at a time, each student stands at a position 10 feet from the basket and keeps shooting until he finally makes a basket. Another student records x_i – how many failures shooter i had before successfully making his first basket.

The geometric probability mass function, which we met in midterm 1, is appropriate for modeling each student's number of failures before the first success. The students have never played together before, so they don't have any knowledge about who the good shooters and poor shooters might be. Thus, in their Bayesian model, they consider the success probabilities p_i, $i = 1, \ldots, 8$ to be random draws from a common Beta density. They complete their Bayesian model by specifying priors on the parameters of the Beta density.

OpenBUGS code and output for fitting the students' model to their data are attached. Note that the geometric distribution is a special case of the negative binomial distribution, namely negative binomial with the second parameter equal to 1. Thus,

```
x[i] ~ dnegbin(p[i], 1)
```

says that x_i is drawn from a geometric distribution with parameter p_i.

1. On the OpenBUGS code, indicate which line or lines represent the second stage of the model.

2. What quantity in the OpenBUGS model should be monitored to get samples from the posterior density of the overall average success probability? If it's already named in the model code, just write the name here. If not, in the OpenBUGS code itself, write the line(s) that should be added to define the quantity, and then write the name here.

   theta

3. If the students had used a Gamma(1,2) prior on the parameter `beta` instead of Gamma(1,1), would that have been likely to make any difference in the resulting posterior means of the individual p_is? Explain briefly.

   Yes. Changing prior in this way would have encouraged β to be smaller (mean $\frac{1}{2}$ instead of 1) which in turn would encourage $\frac{\alpha}{\alpha+\beta}$ to be larger, thus encouraging p_i's to be larger as well.

4. Three plots are included in the OpenBUGS output provided. Refer to them in answering these questions.

   (a) The autocorrelation between values of the `beta` parameter drawn 50 iterations apart is closest to (circle one):

      i. 1.0
      ii. 0.5
      iii. 0.0
      iv. -0.5
      v. -1.0
      vi. Plots give no information on this.

   (b) How many iterations would you discard as burn-in? Explain how you decided.

   At least 350, probably more. Red line in BGR plot stabilizes near 1.0 earlier than that, but blue and green lines must also come together and stay horizontal.

5. What is the 95% credible set for p_2? (Give numeric values from OpenBUGS output).

   (0.1049, 0.9783)

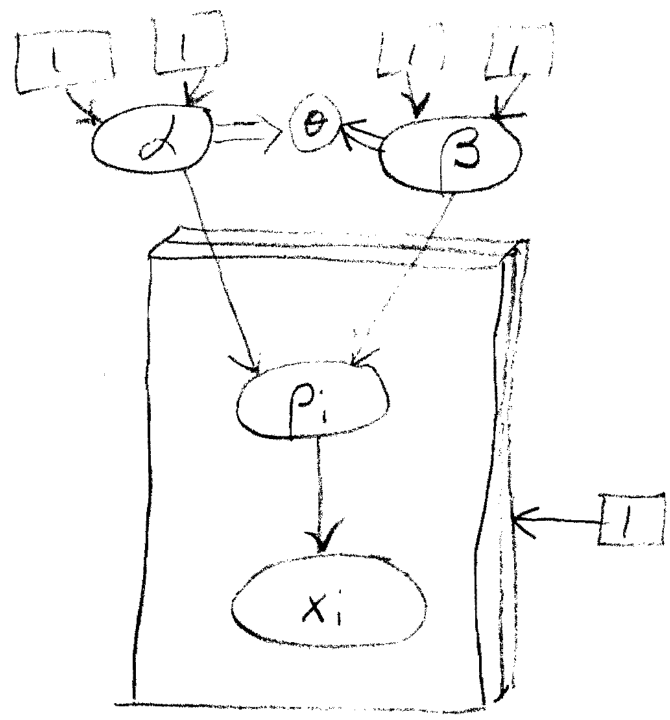
6. Student 2 and Student 8 have exactly the same data values: $x_2 = x_8 = 0$. But the estimated posterior means shown in the OpenBUGS output for p_2 and p_8 are not equal. Why might that be the case? OpenBUGS estimates are based on random sampling. The true posterior means of p_2 and p_8 have to be equal, but random sampling variability affects their estimates.

7. If the data for student number 2 was analyzed separately, the frequentist mle would be $\hat{p} = 1$. However, the Bayesian posterior mean for p_2 from this hierarchical model is only about 0.55. This is an example of a phenomenon in hierarchical models called (circle one):

   (a) exchangeability
   (b) invariance to transformations
   (c) shrinkage
   (d) multiple stages
   (e) none of the above

8. In specifying their model, the students treated all the p_is as random draws from the same Beta density. This shows that they considered the p_is to be (circle one):
   (a) exchangeable
   (b) invariant to transformations
   (c) marginally independent
   (d) nuisance parameters
   (e) none of the above

9. Draw a directed graph of the students' model.



```
model
{
  for( i in 1:N) {
        x[i] ~ dnegbin( p[i], 1 )
        p[i] ~ dbeta( alpha, beta )
   }

   alpha ~ dgamma( 1, 1)
    beta  ~ dgamma( 1, 1)
```
second stage

theta <- alpha/(alpha + beta)

```
}

# data
list( x = c( 4,0,1, 7, 3, 2, 8, 0), N = 8)

# inits

list( alpha = 1, beta = 1 )
list( alpha = 10, beta = 1)
list( alpha = 1, beta = 10 )
```