

Name: -----

**Bayesian Statistics, 22S:138**  
PRACTICE PROBLEMS for Midterm 2, 2005  
Midterm 2

Experiments are performed on laboratory animals to assess whether chemicals are harmful to unborn baby animals. In these experiments, pregnant females are fed (or otherwise exposed to) the chemical of interest and responses are recorded on each individual baby.

Weil (1970) reported an experiment in which 16 pregnant female rats were fed a diet treated with a chemical. The parameter of interest was the population proportion of baby rats that would survive 21 days after birth if their mother ingested this chemical during pregnancy. The group of baby rats born to the same mother at one time are called a "litter." The following notation will be used:

- L = 16    total number of litters
- $n_i$     total number in litter  $i, i = 1 \dots 16$
- $x_i$     number surviving in litter  $i, i = 1 \dots 16$
- N = 145    total number of baby rats in all litters
- X = 112    total number of survivors in all litters

The data for the individual litters were as follows (presented as  $x_i/n_i$ ):

12/12	11/11	10/10	9/9	10/11	9/10	9/10	8/9	8/9	4/5	7/9
4/7	5/10	3/6	3/10	0/7						

Two models for these data were fit using WinBUGS. The models are shown below, and the output is attached.

Model 1

```
model
{
  X ~ dbin(p,N);
  p ~ dbeta(1,1);
}
```

Model 2

```
model
{
  for (i in 1:L) {
    x[i] ~ dbin(p[i],n[i]);
    p[i] ~ dbeta(a,b);
  }
  theta <- a/(a + b);
  a ~ dgamma(1,.2);
  b ~ dgamma(1,.2);
}
```

1. Draw a directed graph of Model 2.

2. Which of these models is or are hierarchical?
- (a) Model 1
  - (b) Model 2
  - (c) Both
  - (d) Neither
3. In which of these models are the outcomes (survive or not survive) for all the individual baby rats in all litters considered exchangeable?
- (a) Model 1
  - (b) Model 2
  - (c) Both
  - (d) Neither
4. In which of these models are the outcomes for the individual baby rats in any single litter considered exchangeable?
- (a) Model 1
  - (b) Model 2
  - (c) Both
  - (d) Neither
5. Which model is more appropriate for this application and why?
6. From the attached WinBUGS output, what are the posterior mean and the 95% central posterior interval for the parameter of primary interest (the population proportion of baby rats that would survive 21 days after birth if their mother ingested the chemical during pregnancy)?
- (a) according to Model 1
  - (b) according to Model 2
7. Three samplers were run for Model 2 using different sets of initial values. WinBUGS history plots, autocorrelation plots, and BGR diagnostic plots are given for the parameters `a`, `b`, and `theta` and for one of the `p`'s. (The plots for the remaining `p`'s are similar.) In addition, the table corresponding to the BGR diagnostic plot for `b` is shown. Use the plots and table to answer the following questions.

- (a) For which model parameter was the autocorrelation the highest?
  - (b) How many burn-in iterations would you discard? (Circle the best answer.)
    - i. none
    - ii. about 250
    - iii. about 500
    - iv. about 750
    - v. 1000 or more
  - (c) Briefly justify your answer to the previous question, referring to specific WinBUGS output.
8. Suppose that we had data on only the first litter, in which all 12 of the baby rats survived. We wish to use it to infer about the population proportion  $p$ . We wish to use an improper prior on  $p$ , namely
- $$p \sim \text{Beta}(0, 0)$$
- (a) Is it possible to carry out a valid Bayesian analysis using this prior and the data from litter 1 alone? (Yes/no)
  - (b) If your answer to the previous question was "Yes," what is the resulting posterior distribution,  $p(p|\mathbf{x})$ ? If your answer to the previous question was "No," why not?