Exercises 2

Question 1, finding a prior: If you went to a biology department in America, what proportion of the staff would be male? Give a best estimate, and some indication of the range (you can take this as a 95% confidence interval). Find a Beta distribution that summarises your belief.

Question 2: Go to the web page of the biology department of an American university, and find a list of staff. Count how many there are, and how many are male. How does this compare with your estimate from Question 1?

Question 3: Find the posterior distribution for the proportion of males. In particular, calculate:
a. The prior and posterior mean
b. The prior and posterior median
c. The prior and posterior standard deviation
d. The prior and posterior probability that the proportion of males is larger than 0.5
e. The prior and posterior predicted number of males in a department of 20 people. What are the mean and standard deviation?

NOTE: This can all be done easily in R, for example, and some nice plots can be produced (e.g. curve(dbeta(x, A, B)) will draw a curve of the density of a beta distribution with parameters A and B). If you are not familiar with R, or any other software available that can do this analysis, then you can use OpenBUGS. To do this:

1. Start OpenBUGS.
2. Click in File -> New
3. Copy the following code into the window:

```
model {  p ~ dbeta(100,31)    }
```

4. Replace 100 and 31 by the values you need.
5. Click on Model -> Specification
6. Click on Check Model (with the window with the code in it raised). A message should appear in the bottom left hand corner saying "model is syntactically correct"
7. Click on Compile. You should get the message "model compiled"
8. Click on Gen Inits
9. Click on Inference -> Samples. Type p in the node box, and click Set.
10. Click on Model -> Update. Change the 1000 to 10000, then click Update
11. When BUGS has finished updating, type p in the node box in the Sample Monitor Tool, and click density. This will produce a plot of the density.
12. For the statistics, click stats in the Sample Monitor Tool.
13. To calculate the probability that p>0.5, copy the following code into a new BUGS window:

```
model {
     p ~ dbeta(100,31)
     Ind <- step(p-0.5)
}
```

Check, compile etc. as before, but now put both p and Ind in the Sample Monitor Tool (type one in, press Set, then type the other in and press Set again). The mean of Ind will be the probability that p>0.5.