

Lab 11: Comparing Binomial Proportions

Wayne Stewart

April 6, 2018

Contents

A good reference book	1
The story	1
Task 1	2
Task 2	2
Task 3	2
Task 4	2
Task 5	2
Task 6	3

A good reference book

The following problem is taken from page 105 Bayesian Ideas and Data Analysis: An Introduction for Scientists and Statisticians

The story

534 medflies that lived at least 34 days. These are cross classified by two factors:

- A) whether the fly produced at least 1,000 eggs in the first 30 days of life and
- B) whether the fly lived at least 44 days

```
mat = matrix(c(54,80,224,176), nr = 2,nc=2, byrow = TRUE)
colnames(mat)= c("Long-lived","short-lived")
rownames(mat)= c("High", "Low")
mat
```

```
##      Long-lived short-lived
## High         54         80
## Low          224        176
```

These are multinomial data but if we view the Longevity given reproductive factor we shall see these as two independent binomials

$$y_1|\theta \sim \text{Bin}(n_1, \theta) \perp y_2|\theta_2 \sim \text{Bin}(n_2, \theta_2)$$

The data can then be obtained from the above table:

$$n_1 = 134, n_2 = 400, y_1 = 54, y_2 = 224$$

Task 1

Find the conjugate solution – that is find the *analytical* posteriors for $\theta_1|y_1$ and $\theta_2|y_2$ using $\theta_i \sim \text{Beta}(a_i, b_i)$ where $i \in \{1, 2\}$

Task 2

Make plots of the two posteriors in R using

- $a_i = 1, b_i = 1$
- $a_i = 4, b_i = 1$
- $a_i = 1, b_i = 4$
- $a_i = 3, b_i = 3$

Task 3

Create a jags model for the above problem by correcting and/or completing the given code. Note that there are some useful logical nodes made. Comment on them.

```
model{
  for(i in 1:2){
    y[i] ~ dbin(theta[i], n[i])

    theta[i] ~ ...
    odds[i] <- theta[i]/(1-theta[i])
  }

  RD <- theta[2]-theta[1] # Risk difference
  RR <-theta[2]/theta[1] #Relative risk
  test<- step(RD) # What does this test?
}
```

Task 4

Now make the complete jags code. Please include some logical constant assignments so that you can assign new priors easily

Example:

```
a[i]<-1
b[i]<-1
```

Make sure you monitor `theta`, `RR`, `OR` and `test`.

Task 5

Please give point and interval estimates for the above nodes. Interpret them.

Task 6

Make a function `mybibin()` that will solve the same type of problem where the data inputs will be a vector of y's and a vector of n's. The priors will be beta's and will therefore need a vector of a's and b's. The command line output should be a list of point and interval estimates for a given α . The function should produce analytical posterior, prior and likelihood plots and MCMC posterior plots of the theta's and all other stochastic nodes. Run your function with the same data as given and $a = c(1, 2)$, $b = c(1, 2)$ and $\alpha = 0.2$

How will you change the model code? There are many ways to do this – one way is to break the model “text” into parts and rejoin them after changing values. You can use `paste()`