# STAT 431 Homework 4

Robert Garrett

10/19/2022

## Problem 1

```
n1 <- 24;   ybar1 <- 7.8730;  s1 <- 0.05353
n2 <- 123;  ybar2 <- 7.9725;  s2 <- 0.01409
Nsim <- 100000
sigma1.2s <- 1 / rgamma(Nsim, (n1-1)/2, (n1-1)*s1^2/2)
sigma2.2s <- 1 / rgamma(Nsim, (n2-1)/2, (n2-1)*s2^2/2)
mu1s <- rnorm(Nsim, ybar1, sqrt(sigma1.2s/n1))
mu2s <- rnorm(Nsim, ybar2, sqrt(sigma2.2s/n2))
```

### Part a

```
sigma2_ratio = sigma1.2s/sigma2.2s
mean(sigma2_ratio)
```

Posterior mean is: 15.829

### Part b

```
sd(sigma2_ratio)/sqrt(Nsim)
```

Monte Carlo error is: 0.01765

### Part c

```
sd(sigma2_ratio)
```

Posterior Standard Deviation is: 5.583

### Part d

```
quantile(sigma2_ratio,c(0.025,0.975))
```

95% Equal-tailed Credible interval is: (8.125, 29.506)

### Part e

```
lower = 1/qf(0.025,n1-1,n2-1,lower.tail=FALSE)*(s1/s2)^2
upper = qf(0.025,n2-1,n1-1,lower.tail=FALSE)*(s1/s2)^2
c(lower,upper)
```

95% Frequentist confidence interval is: (8.144, 29.449), slightly narrower than the previous credible interval.

## Problem 2

### Part a

$$
\begin{aligned}
D(y=12|\theta) &= -2ln(f(y=12|\theta)) \\
&= -2ln\left(\binom{70}{12}\theta^{12}(1-\theta)^{70-12}\right) \\
&= -2\left(ln\binom{70}{12} + 12ln(\theta) + 58ln(1-\theta)\right)
\end{aligned}
$$

### Part b

$$
\begin{aligned}
E(D(y|\theta)|y) &= \int_0^1 D(y|\theta)p(\theta|y)d\theta \\
&= \int_0^1 D(y|\theta)p(\theta|y)d\theta \\
&= \int_0^1 D(y|\theta)\left(\frac{f(y|\theta)\pi(\theta)}{m(y)}\right)d\theta \\
&= \int_0^1 D(y|\theta)\left(\frac{f(y|\theta)\pi(\theta)}{\int_0^1 f(y|\theta)\pi(\theta)d\theta}\right)d\theta \\
&= \frac{\int_0^1 D(y|\theta)f(y|\theta)\pi(\theta)d\theta}{\int_0^1 f(y|\theta)\pi(\theta)d\theta}
\end{aligned}
$$

Plugging in $y = 12$ along with the Jeffrey's prior for $\theta$, $Beta(1/2, 1/2)$, we get the following integrals:

$$
E(D(y|\theta)|y=12) = \frac{\int_0^1 -2\left(ln\binom{70}{12} + 12ln(\theta) + 58ln(1-\theta)\right)\left(\binom{70}{12}\theta^{12}(1-\theta)^{58}\right)\left(\frac{\Gamma(1)}{\Gamma(1/2)\Gamma(1/2)}\theta^{-1/2}(1-\theta)^{-1/2}\right)d\theta}{\int_0^1 \left(\binom{70}{12}\theta^{12}(1-\theta)^{58}\right)\left(\frac{\Gamma(1)}{\Gamma(1/2)\Gamma(1/2)}\theta^{-1/2}(1-\theta)^{-1/2}\right)d\theta}
$$

### Part c

```
(top = integrate(function(t){
  -2*log(dbinom(12,70,t))*dbinom(12,70,t)*dbeta(t,1/2,1/2)
},0,1))
```

```
## 0.06122251 with absolute error < 6.4e-07
```

```
(bottom = integrate(function(t){
  dbinom(12,70,t)*dbeta(t,1/2,1/2)
},0,1))
```

```
## 0.01191481 with absolute error < 9.6e-09
```

```
top$value/bottom$value
```

```
## [1] 5.138356
```

So, the posterior expected deviance is: 5.138

## Problem 3

### Part a

$$\begin{aligned} p(\theta|y,n) &\propto f(y|\theta,n)\pi(\theta)\pi(n) \\ &\propto \binom{n}{y}\theta^y(1-\theta)^{n-y} \\ &\propto \theta^{(y+1)-1}(1-\theta)^{(n-y+1)-1} \end{aligned}$$

From the kernel, we see $\theta|Y=y, N=n \sim Beta(y+1, n-y+1)$

### Part b

Note that $\binom{r+y}{y} = \binom{r+y}{r}$. $y$ is given, so use a substitution to find the conditional distribution of $R$:

$$\begin{aligned} p(n|\theta,y) &\propto \binom{n}{y}\theta^y(1-\theta)^{n-y}0.9^n \\ p(r|\theta,y) &\propto \binom{r+y}{y}\theta^y(1-\theta)^r 0.9^{r+y} \\ &\propto \binom{r+y}{r}(1-\theta)^r 0.9^r \\ &= \binom{r+y}{r}(0.9-0.9\theta)^r \\ &= \binom{r+(y+1)-1}{r}(1-(0.1+0.9\theta))^r \end{aligned}$$

So, the conditional distribution is $R|\theta, Y=y \sim NegBinomial(0.1+0.9\theta, y+1)$ in the ordering used in BSM or $NegBinomial(y+1, 0.1+0.9\theta)$ in the ordering used in R.

### Part c

```r
y = 20
n_iterations = 1000000
R = numeric(n_iterations)
theta = numeric(n_iterations)
R[1] = 10
theta[1] = rbeta(1,y+1,R[1]+1)
for(i in 2:n_iterations){
  R[i] = rnbinom(1,y+1,0.1+0.9*theta[i-1])
  theta[i] = rbeta(1,y+1,R[i]+1)
}
mean(R+y<=2*y)
```

```
## [1] 0.938537
```

So, the posterior probability that at least half of the birds in the forest were sighted is about 0.939.