Stat 570 HW5

Dongyang Wang

2022-10-29

Q1

a

$P(Y=1|X=1) = \frac{P(X=1,Y=1)}{P(X=1)} = \frac{P(X=1|Y=1)P(Y=1)}{P(X=1)} = p_1\frac{P(Y=1)}{P(X=1)}$

$P(Y=0|X=1) = \frac{P(X=1,Y=0)}{P(X=1)} = \frac{P(X=1|Y=0)P(Y=0)}{P(X=1)} = p_2\frac{P(Y=0)}{P(X=1)}$

$P(Y=1|X=0) = \frac{P(X=0,Y=1)}{P(X=0)} = \frac{P(X=0|Y=1)P(Y=1)}{P(X=0)} = (1-p_1)\frac{P(Y=1)}{P(X=0)}$

$P(Y=0|X=0) = \frac{P(X=0,Y=0)}{P(X=0)} = \frac{P(X=0|Y=0)P(Y=0)}{P(X=0)} = (1-p_2)\frac{P(Y=0)}{P(X=0)}$

$\theta = \frac{P(Y=1|X=1)P(Y=0|X=0)}{P(Y=0|X=1)P(Y=1|X=0)} = \frac{p_1/(1-p_1)}{p_2/(1-p_2)}$.

b

Multiplying the pdf for p_1, p_2 and taking log, we obtain

$l(x_1, x_2|p_1, p_2) = \log\binom{n_1}{x_1} + x_1\log(p_1) + (n_1-x_1)\log(1-p_1) + \log\binom{n_2}{x_2} + x_2\log(p_2) + (n_2-x_2)\log(1-p_2)$.

Taking derivative wrt p_1, p_2, we have

$\frac{dl}{dp_1} = \frac{x_1}{p_1} + \frac{x_1-n_1}{1-p_1}$ $\frac{dl}{dp_2} = \frac{x_2}{p_2} + \frac{x_2-n_2}{1-p_2}$

Solving, we have $\hat{p_1} = \frac{96}{200}$ and $\hat{p_2} = \frac{109}{775}$.

Thus, $\hat{\theta} = \frac{\hat{p_1}/(1-\hat{p_1})}{\hat{p_2}/(1-\hat{p_2})} = 5.640085$ by invariance of MLE.

```
## [1] 5.640085
```

It's a little difficult to compute but the logic is using the Fisher Information Matrix to calculate the variance with Delta Method and therefore construct a confidence interval based on the standard deviation.

c

$\pi(p_1, p_2|x_1, x_2) \propto \pi(p_1|x_1)\pi(p_2|x_2) \propto p(x_1|p_1)\pi(p)p(x_2|p_2)\pi(p) \propto p_1^{x_1}(1-p_1)^{n_1-x_1}p_1^{a-1}(1-p_1)^{b-1}p_2^{x_2}(1-p_2)^{n_2-x_2}p_2^{a-1}(1-p_2)^{b-1} \propto Beta(a+x_1, b+n_1+x_1) * Beta(a+x_2, b+n_2+x_2)$.

d

From Beta distribution we know, $E(p_1|x_1) = \frac{a+x_1}{a+b+n_1}$ $Mode(p_1|x_1) = \frac{a+x_1-1}{a+b+n_1-2}$ $SD(p_1|x_1) = \sqrt{\frac{(a+x_1)(b+n_1-x_1)}{(a+b+n_1)^2(a+b+n_1+1)}}$

$E(p_2|x_2) = \frac{a+x_2}{a+b+n_2}$ $Mode(p_2|x_2) = \frac{a+x_2-1}{a+b+n_2-2}$ $SD(p_2|x_2) = \sqrt{\frac{(a+x_2)(b+n_2-x_2)}{(a+b+n_2)^2(a+b+n_2+1)}}$

```
##    Expectation      Mode         SD Quantile 5% Quantile 95%
## p1   0.4801980 0.4800000 0.03506559   0.4226119    0.5380077
```

```
## p2   0.1415701 0.1406452 0.01249823   0.1215485    0.1626412
```

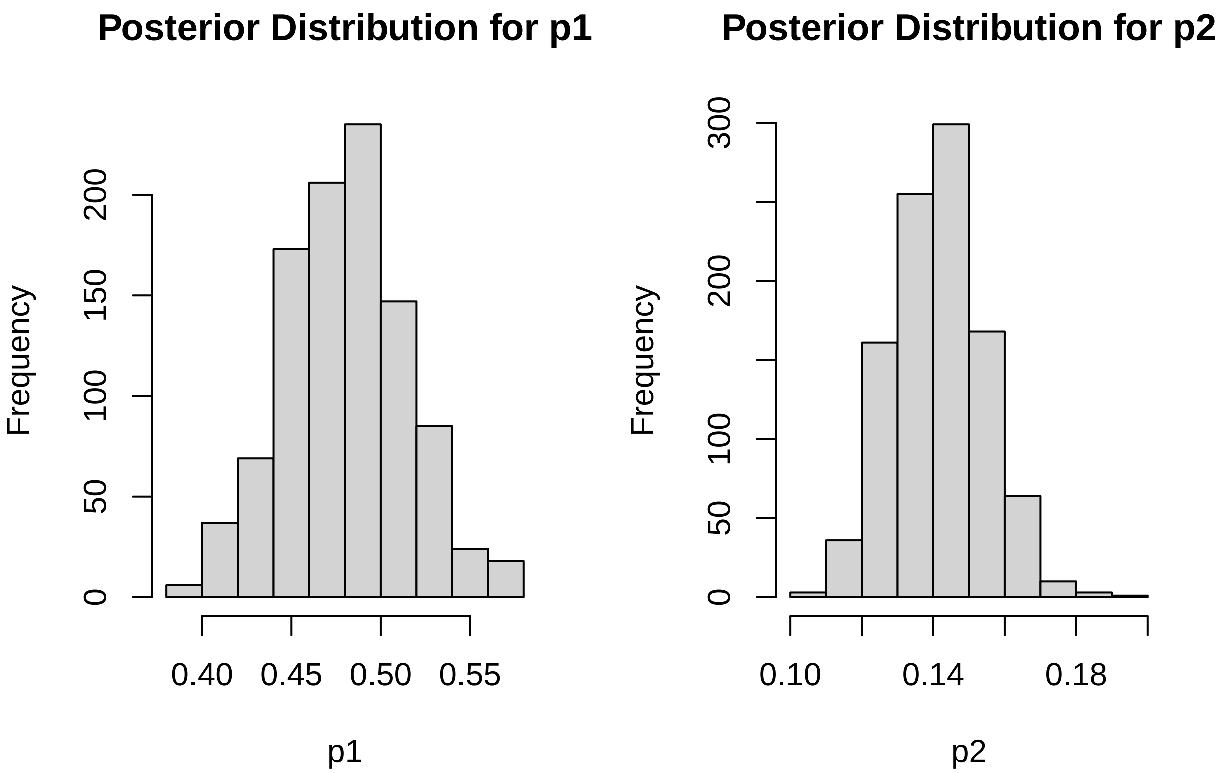
e

The interval for implied prior distribution for theta is as follows.

```
##           5%         95%
##   0.01435962 64.16813032
```

f

The histograms are as follows, as well as the credible intervals for p1, p2.

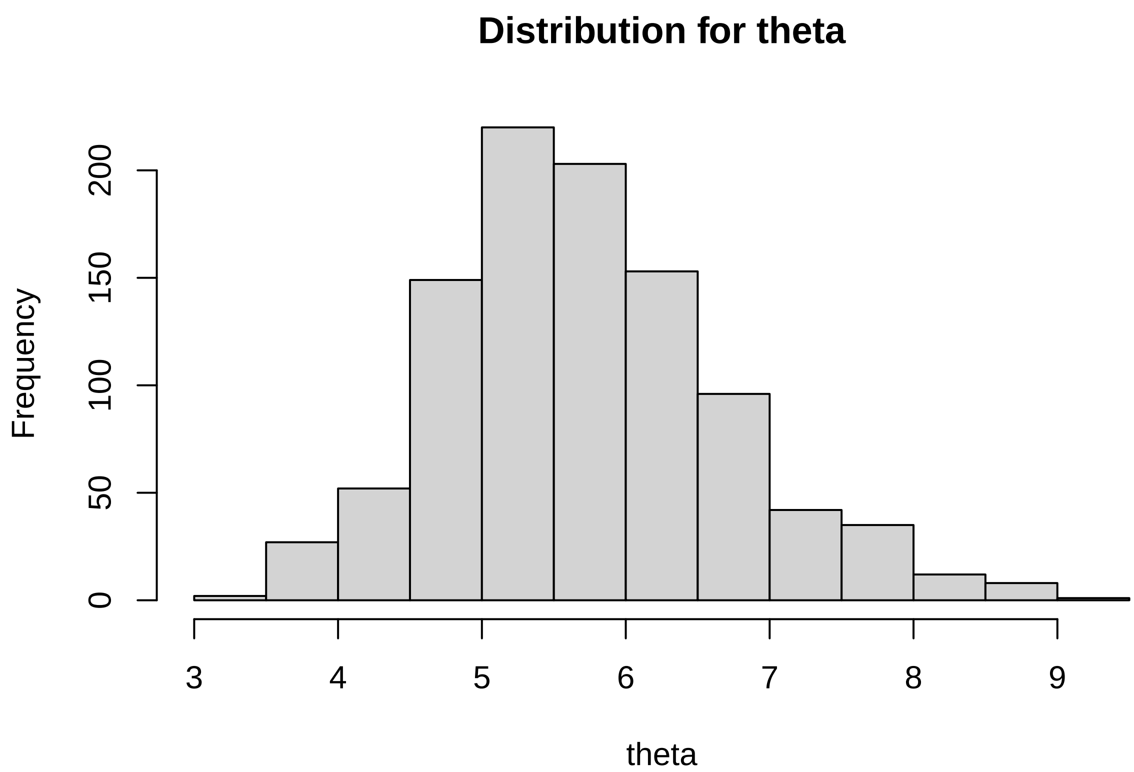



```
##         5%       95%
## 0.4219537 0.5379629
```

```
##         5%       95%
## 0.1209931 0.1624018
```

g

The interval and median for theta are below respectively.




```
##       5%      95%
## 4.223760 7.559624
```

```
##      50%
## 5.607552
```

Compared with the likelihood results, they seem close but the intervals are a bit wider.

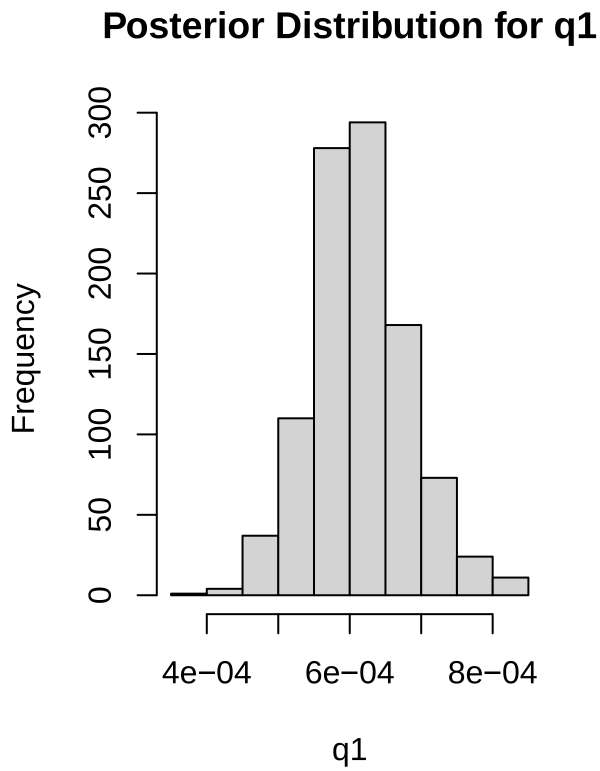
h

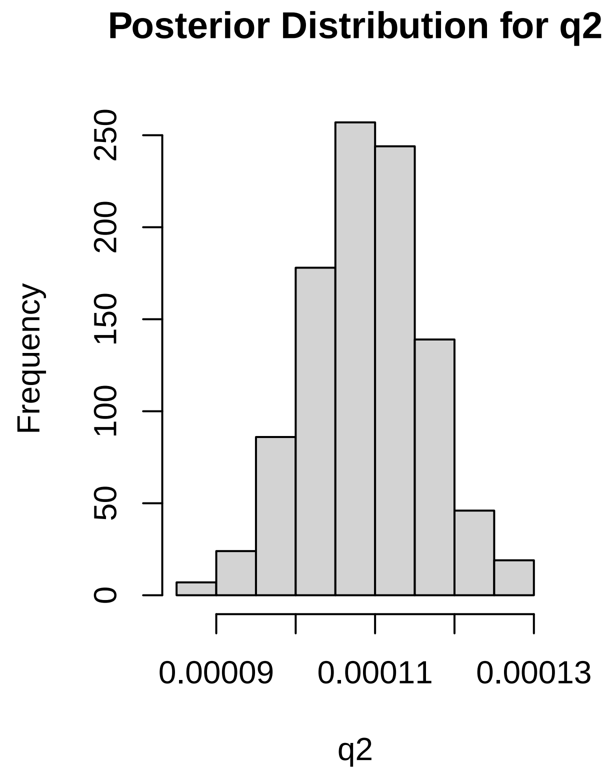
We know $P(Y=1) = 18/100000$. So,

$q_1 = \frac{P(X=1|Y=1)P(Y=1)}{P(X=1|Y=1)P(Y=1)+P(X=1|Y=0)P(Y=0)} = \frac{p_1\frac{18}{100000}}{p_1\frac{18}{100000}+p_2(1-\frac{18}{100000})}$

and $q_2 = \frac{P(X=0|Y=1)P(Y=1)}{P(X=0|Y=1)P(Y=1)+P(X=0|Y=0)P(Y=0)} = \frac{(1-p_1)\frac{18}{100000}}{(1-p_1)\frac{18}{100000}+(1-p_2)(1-\frac{18}{100000})}$

To calculate as in above sections, the histograms and the interval for q1, q2 are as follows respectively.






```
##           5%          95%
## 0.0005055006 0.0007416874
```

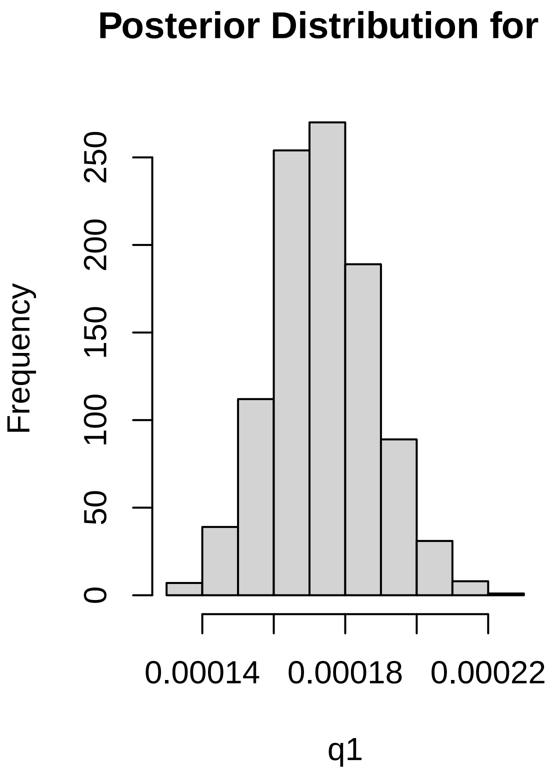
```
##           5%          95%
## 9.656161e-05 1.216228e-04
```

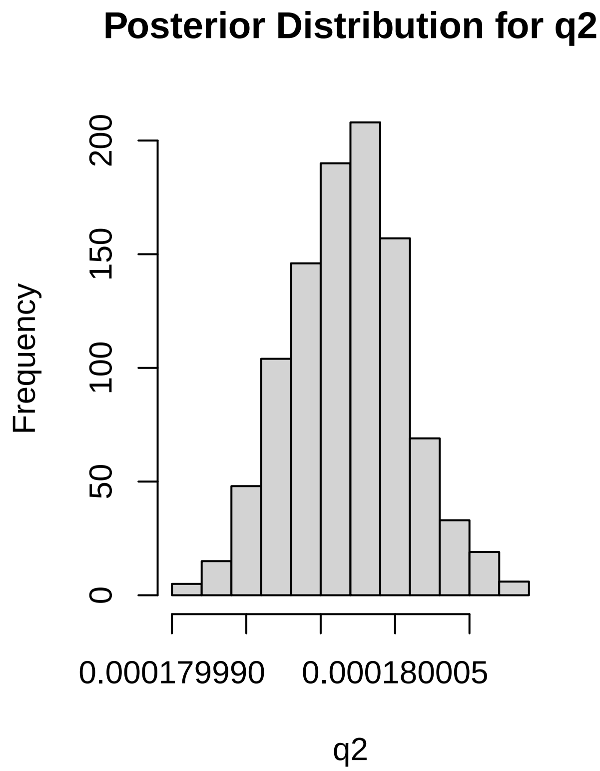
i

a,b are calculated as follows.

```
## [1]     217.6115 1211867.6981
```

The histograms and the interval for q1, q2 are as follows respectively.






```
##           5%          95%
## 0.0001505564 0.0001976295
```

```
##           5%          95%
## 0.0001799954 0.0001800085
```

Q2

a

$p(\hat{\theta}|\theta) \propto p(\theta)p(\theta|\hat{\theta}) \propto \exp(-\frac{1}{2W}\theta^2)\exp(-\frac{1}{2V}(\hat{\theta}-\theta)^2) \propto \exp(-\frac{1}{2W}\theta^2 - \frac{1}{2V}(\hat{\theta}^2+\theta^2-2\hat{\theta}\theta)) \propto \exp(-\frac{1}{rV}\theta^2 + \frac{2}{V}\theta\hat{\theta}) \propto \exp(-\frac{1}{2rV}(\theta - r\hat{\theta})^2)$ and this is proportional to $N(r\hat{\theta}, rV)$.

b

$$BF = \frac{\frac{1}{\sqrt{2\pi V}}\exp(\frac{\hat{\theta}^2}{2V})}{\frac{1}{2\pi\sqrt{VW}}\int \exp(-\frac{(\hat{\theta}-\theta)^2}{2V} - \frac{\theta^2}{2W})d\theta} = \frac{\frac{1}{\sqrt{2\pi V}}\exp(\frac{\hat{\theta}^2}{2V})}{\frac{1}{2\pi\sqrt{VW}}\int \exp(-\frac{(\hat{\theta}-\theta)^2}{2V} - \frac{\theta^2}{2W})d\theta}$$

After simplification, $BF = \frac{\frac{V+W}{V}}{\exp(\frac{-\hat{\theta}^2 r}{2V})} = \frac{1}{\sqrt{1-r}}\exp(\frac{-Z^2 r}{2})$.

c

$P(M_1|\hat{\theta}) = \frac{P(\hat{\theta}|M_1)P(M_1)}{P(\hat{\theta}|M_1)P(M_1)+P(\hat{\theta}|M_0)P(M_0)} = \frac{\pi_! P(\hat{\theta}|M_0)/BF}{\pi_! P(\hat{\theta}|M_0)/BF + \hat{\theta}|M_0)(1-\pi_1)} = \frac{\pi_1}{\pi_1 + BF(1-\pi_1)}$.

d

$p(\theta|\hat{\theta_1}, \hat{\theta_2}) \propto p(\hat{\theta_1}, \hat{\theta_2}|\theta)\pi(\theta) \propto \exp(-(\frac{(\hat{\theta_1}-\theta)^2}{2V_1} + \frac{(\hat{\theta_2}-\theta)^2}{2V_2} + \frac{\theta^2}{2W}))$

After simplification, we assign $a = V_1^{-1}(V_1^{-1} + V_2^{-2} + W^{-1})^{-1}$ and $b = V_2^{-1}(V_1^{-1} + V_2^{-2} + W^{-1})^{-1}$ and $c = (V_1^{-1} + V_2^{-2} + W^{-1})^{-1}$,

$p(\theta|\hat{\theta_1},\hat{\theta_2}) \propto \exp(-\frac{(\theta-\frac{V_1^{-1}\hat{\theta_1}+V_2^{-1}\hat{\theta_2}}{(V_1^{-1}+V_2^{-2}+W^{-1})^{-1}})^2}{2(V_1^{-1}+V_2^{-2}+W^{-1})^{-1}}) \sim N(a\hat{\theta_1}+b\hat{\theta_2},c)$.

e

$BF = \frac{p(\hat{\theta_1},\hat{\theta_2}|\theta=0}{\int_{\theta\neq 0}p(\hat{\theta_1},\hat{\theta_2}|\theta)\pi(\theta)d\theta} = \frac{2\pi\sqrt{V_1V_2W}}{2\pi\sqrt{V_1V_2}}\frac{\exp(-(\frac{(\hat{\theta_1})^2}{2V_1}+\frac{(\hat{\theta_2})^2}{2V_2}))}{\int_{\theta\neq 0}\exp(-(\frac{(\hat{\theta_1}-\theta)^2}{2V_1}+\frac{(\hat{\theta_2}-\theta)^2}{2V_2}+\frac{\theta^2}{2W}))d\theta}$

Based on results from d and simplifications,

$BF = \frac{W}{c}\exp(-\frac{(a\hat{\theta_1}+b\hat{\theta_2})^2}{2c})$ where $a = V_1^{-1}(V_1^{-1}+V_2^{-2}+W^{-1})^{-1}$ and $b = V_2^{-1}(V_1^{-1}+V_2^{-2}+W^{-1})^{-1}$ and $c = (V_1^{-1}+V_2^{-2}+W^{-1})^{-1}$, as defined in d.

f

Based on the CI, W will take the value

```
## [1] 0.04279675
```

V1, V2 are respectively

```
## [1] 0.0018017814 0.0009502544
```

The median based on first dataset is the normal mean, and ci will be calculated as follows

```
## [1] 0.3890843
```

and the results for

```
##                 median  CI Lower CI Higher
## One dataset  0.2293606 0.1478631 0.3108581
## Two datasets 0.1715401 0.1230048 0.2200755
```

g

The Bayes factors for 1 set and two sets are as follows

```
## [1] 1.229929e-06
```

```
## [1] 3.176441e-10
```

h

The probabilities are as follows.

```
## [1] 0.9938892 0.9999984
```

Q3

a

The result is as follows:

```
## Loading required package: Matrix
```

```
## Loading required package: foreach
```

```
## Loading required package: parallel
```

```
## Loading required package: sp
```

```
## This is INLA_22.05.07 built 2022-05-07 09:58:26 UTC.
##  - See www.r-inla.org/contact-us for how to get help.
##  - To enable PARDISO sparse library; see inla.pardiso()
```

```
##                     mean         sd  0.025quant    0.5quant 0.975quant mode
## (Intercept)  0.08659774 0.06464031 -0.04043483  0.08677722 0.21326047   NA
## newx        -0.01784798 0.01429274 -0.04596179 -0.01780025 0.01013564   NA
##                       kld
## (Intercept) 4.224437e-06
## newx        4.671814e-13
```

b

For some reason the code wouldn't run. See appendix for details.

Appendix

Q1

```
p1 = 96/200
p2 = 109/775
theta = (p1/(1-p1))/(p2/(1-p2))
theta

a = 1
b = 1
x1 = 96
n1 = 200

x2 = 109
n2 = 775

e1 <- (a + x1)/(a + b + n1)
m1 <- (a + x1 -1)/(a + b + n1 -2)
sd1 <- sqrt( ((a + x1) * (b+n1 -x1))/((a+b+n1)^2 * (a+b+n1+1)) )

e2 <- (a + x2)/(a + b + n2)
m2 <- (a + x2 -1)/(a + b + n2 -2)
sd2 <- sqrt( ((a + x2) * (b+n2 -x2))/((a+b+n2)^2 * (a+b+n2+1)) )

q1 <- c(qbeta(0.05, a + x1 , b + n1 -x1), qbeta(0.95, a + x1 , b + n1 -x1))
q2 <- c(qbeta(0.05, a + x2 , b + n2 -x2), qbeta(0.95, a + x2 , b + n2 -x2))

mat_d <- matrix(c(e1, m1, sd1, q1, e2, m2, sd2, q2), byrow = TRUE, nrow = 2)
colnames(mat_d) <- c("Expectation", "Mode", "SD", "Quantile 5%", "Quantile 95%")
rownames(mat_d) <- c("p1", "p2")
mat_d

nsample <- 1e4

prior1 <- rbeta(nsample, a , b)
prior2 <- rbeta(nsample, a , b)

theta_e <- (prior1/(1-prior1))/(prior2/(1-prior2))

ci_theta_e <- quantile(theta_e, c(0.05, 0.95))
ci_theta_e
```

```
t = 1e3
post1 <- rbeta(t, a + x1 , b + n1 -x1)
post2 <- rbeta(t, a + x2 , b + n2 -x2)

par(mfrow = c(1, 2))
hist(post1, xlab = "p1", main = "Posterior Distribution for p1")
hist(post2, xlab = "p2", main = "Posterior Distribution for p2")

post_q1 <- quantile(post1, c(0.05, 0.95))
post_q2 <- quantile(post2, c(0.05, 0.95))

post_q1
post_q2

theta_g <- (post1/(1-post1))/(post2/(1-post2))
hist(theta_g, xlab = "theta", main = "Distribution for theta")

theta_g_quantile <- quantile(theta_g, c(0.05, 0.95))
theta_g_quantile

theta_g_median <- quantile(theta_g, c(0.5))
theta_g_median

py <- 18/1e5
q_1 <- post1 * py/ (post1*py + (post2 * (1 - py)))
q_2 <- (1 - post1) * py/ ((1 - post1)*py + (1-post2) * (1 - py) )

par(mfrow = c(1, 2))
hist(q_1, xlab = "q1", main = "Posterior Distribution for q1")
hist(q_2, xlab = "q2", main = "Posterior Distribution for q2")

quantile_q1 <- quantile(q_1, c(0.05, 0.95))
quantile_q2 <- quantile(q_2, c(0.05, 0.95))

quantile_q1
quantile_q2

get_ab <- function(res){
  sum((pbeta(16/100000, shape1 =  res[1], shape2 = res[2]) - 0.05)^2
      + (pbeta(20/100000, shape1 =  res[1], shape2 = res[2]) - 0.95)^2)
}

ab <- optim(par = c(1,1), fn = get_ab)$par
a <- ab[1]
b <- ab[2]
c(a,b)

t = 1e3
post1 <- rbeta(t, a + x1 , b + n1 -x1)
post2 <- rbeta(t, a + x2 , b + n2 -x2)

py <- 18/1e5
q_1 <- post1 * py/ (post1*py + (post2 * (1 - py)))
```

```r
q_2 <- (1 - post1) * py/ ((1 - post1)*py + (1-post2) * (1 - py) )

par(mfrow = c(1, 2))
hist(q_1, xlab = "q1", main = "Posterior Distribution for q1")
hist(q_2, xlab = "q2", main = "Posterior Distribution for q2")

quantile_q1 <- quantile(q_1, c(0.05, 0.95))
quantile_q2 <- quantile(q_2, c(0.05, 0.95))

quantile_q1
quantile_q2
```

Q2

```r
(log(3/2)/qnorm(0.975))^2
((log(c(1.37, 1.23)) - log(c(1.16, 1.09)))/(qnorm(0.975) * 2))^2
0.04279675/(0.0018017814+0.04279675) *log(3/2)
theta = c(log(1.27), log(1.15))

V <- ((log(c(1.37, 1.23)) - log(c(1.16, 1.09)))/(qnorm(0.975) * 2))^2
r <- W/(W+V[1])
v <- 1/(sum(1/V) + 1/W)

rboth <- 1/V*v

res1 <- qnorm(c(0.5, 0.025,0.975),mean=r*theta[1],sd=sqrt(r*V[1]))
res2 <- qnorm(c(0.5,0.025,0.975),mean=sum(rboth*theta),sd=sqrt(v))
dat_f <- matrix(c(res1, res2), byrow = TRUE, nrow = 2)
colnames(dat_f) <- c("median", "CI Lower", "CI Higher")
rownames(dat_f) <- c("One dataset", "Two datasets")
dat_f

z <- theta[1]/sqrt(V[1])
bf1 <- 1/sqrt(1-r)*exp(-z^2/2*r)
bf2 <- sqrt(W/v)*exp(-(sum(rboth*theta))^2/(2*v))
bf1
bf2

pi1 <- 1/5000
p1 <- pi1/(pi1 + bf1*(1-pi1))
p2 <- pi1/(pi1 + bf2*(1-pi1))

c(p1, p2)
```

Q3

```r
library(data.table)
lung <- as.data.frame(fread('http://faculty.washington.edu/jonno/book/MNlung.txt'))
radon <- as.data.frame(fread('http://faculty.washington.edu/jonno/book/MNradon.txt'))
#install.packages("INLA",repos=c(getOption("repos"),INLA="https://inla.r-inla-download.org/R/stable"),
#inla.upgrade()
library(INLA)
```

```
# Code from Jon's website
Obs <- apply(cbind(lung[,3], lung[,5]), 1, sum)
Exp <- apply(cbind(lung[,4], lung[,6]), 1, sum)
rad.avg <- rep(0, length(lung))
for(i in 1:length(lung)) {
        rad.avg[i] <- mean(radon[radon$county==i,2])
}
x <- rad.avg
rad.avg[26]<-0
rad.avg[63]<-0
x[26] <- NA
x[63] <- NA
newy <- Obs[is.na(x)==F]
newx <- x[is.na(x)==F]
newE <- Exp[is.na(x)==F]

df <- as.data.frame(cbind(newy,newx,newE))
inla_model <- inla(newy ~ newx, data = df, family = "poisson", E=newE)
inla_model$summary.fixed
plot(inla_model)

# Code from Jon's website
LogNormalPriorCh <- function(theta1,theta2,prob1,prob2){
  zq1 <- qnorm(prob1)
  zq2 <- qnorm(prob2)
  mu <- log(theta1)*zq2/(zq2-zq1) - log(theta2)*zq1/(zq2-zq1)
  sigma <- (log(theta1)-log(theta2))/(zq1-zq2)
  list(mu=mu,sigma=sigma)
}

lungradprior0 <- LogNormalPriorCh(1.5,0.67,.975,.025)
lungradprior1 <- LogNormalPriorCh(1.2,0.8,.975,.025)

newx1 <- newx-mean(newx)
new_model <- inla(newy ~ newx1, data = df, family = "poisson", E=newE,
              control.fixed=list(mean.intercept=lungradprior0$mu,
                                 prec.intercept=1/(lungradprior0$sigma^2),
                                 mean=c(lungradprior1$mu),
                                 prec=c(1/(lungradprior1$sigma^2)) ) )

new_model$summary.fixed
plot(new_model)
```