

STAT 4224 HW #3

Carlyle Morgan

10/19/2021

```
set.seed(1)
knitr::opts_chunk$set(echo = TRUE)
```

1.

a.

```
a_data<-c(12,9,12,14,13,13,15,8,15,6)
b_data<-c(11,11,10,9,9,8,7,10,6,8,8,9,7)

a_sim<-rgamma(5000,120+sum(a_data),10+length(a_data))
b_sim<-rgamma(5000,12+sum(b_data),1+length(b_data))

mean(a_sim-b_sim)
```

```
## [1] 2.922571
```

```
mean(a_sim>b_sim)
```

```
## [1] 0.9962
```

```
quantile(a_sim-b_sim,c(0.025,0.975))
```

```
##      2.5%      97.5%
```

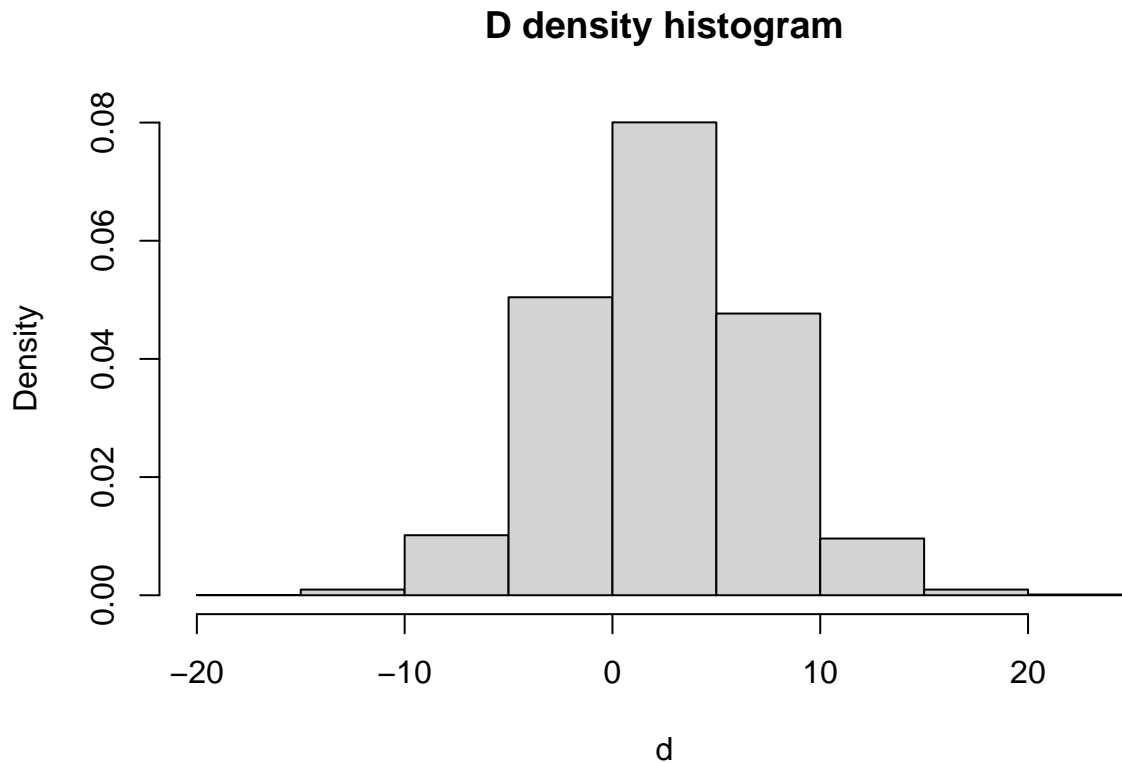
```
## 0.7736724 5.1374764
```

It is invariably clear that the underlying rate parameter is bigger for A than B, and we can expect this difference is ~2.9. We are 99.6% sure that the underlying rate parameter is bigger for A than B. We can be 95% confident that the difference between the parameters for A and B is between 0.7736 and 5.137.

b.

```
yA_mc<-rpois(5000,a_sim)
yB_mc<-rpois(5000,b_sim)
d<-yA_mc-yB_mc

hist(d, freq=F,main="D density histogram")
```



The plot looks has an almost normal shape, but its weird shape is likely because it is the result of subtraction of two negative binomial distributions. It looks like the distribution of d is centered a little to the right of zero, but that is just because of the scale. This confirms our thinking that the expected number of tumors is probably greater for the A population than the B one.

2.

a.

```

y1 <- scan("http://www2.stat.duke.edu/~pdh10/FCBS/Exercises/school1.dat")
y2 <- scan("http://www2.stat.duke.edu/~pdh10/FCBS/Exercises/school2.dat")
y3 <- scan("http://www2.stat.duke.edu/~pdh10/FCBS/Exercises/school3.dat")

mu.0<-5
sigma2.0<-4
kappa.0<-1
nu.0<-2

n1 <- length(y1); ybar1 <- mean(y1); s2.1 <- var(y1);

nu.n1 <-nu.0 + n1; kappa.n1 <- kappa.0 + n1;

mu.n1 <- (kappa.0*mu.0 + n1*ybar1) / kappa.n1

sigma2.n1 <- (1/nu.n1) * (nu.0*sigma2.0 + (n1-1)*s2.1 +

```

```

kappa.0*n1*(ybar1-mu.0)^2 / kappa.n1)

n2 <- length(y2); ybar2 <- mean(y2); s2.2 <- var(y2);

nu.n2 <- nu.0 + n2; kappa.n2 <- kappa.0 + n2;

mu.n2 <- (kappa.0*mu.0 + n2*ybar2) / kappa.n2

sigma2.n2 <- (1/nu.n2) * (nu.0*sigma2.0 + (n2-1)*s2.2 +
  kappa.0*n2*(ybar2-mu.0)^2 / kappa.n2)

n3 <- length(y3); ybar3 <- mean(y3); s2.3 <- var(y3);

nu.n3 <- nu.0 + n3; kappa.n3 <- kappa.0 + n3;

mu.n3 <- (kappa.0*mu.0 + n3*ybar3) / kappa.n3

sigma2.n3 <- (1/nu.n3) * (nu.0*sigma2.0 + (n3-1)*s2.3 +
  kappa.0*n3*(ybar3-mu.0)^2 / kappa.n3)

gs <- 800

mu <- seq(5, 13.5, length=gs)

I.sig2 <- exp(seq(log(.01), log(.18), length=gs))

sigma2 <- exp(seq(log(3), log(45), length=gs))

log.post <- matrix(NA, gs, gs);

for(i in 1:gs){ for(j in 1:gs){ log.post[i,j] <-
  dnorm(mu[i], mu.n1, 1/sqrt(I.sig2[j]*kappa.n1), log=T) +
  dchisq(nu.n1*sigma2.n1*I.sig2[j], df=nu.n1, log=T) }}

maxie <- max(log.post); log.post <- log.post - maxie; rm(maxie);

post.P <- exp(log.post); rm(log.post);

log.post <- matrix(NA, gs, gs);

for(i in 1:gs){ for(j in 1:gs){ log.post[i,j] <-
  dnorm(mu[i], mu.n1, sqrt(sigma2[j]/kappa.n1), log=T) +
  dchisq(nu.n1*sigma2.n1/sigma2[j], df=nu.n1, log=T) -
  2*log(sigma2[j]) }}

maxie <- max(log.post); log.post <- log.post - maxie; rm(maxie);

post.V <- exp(log.post); rm(log.post);

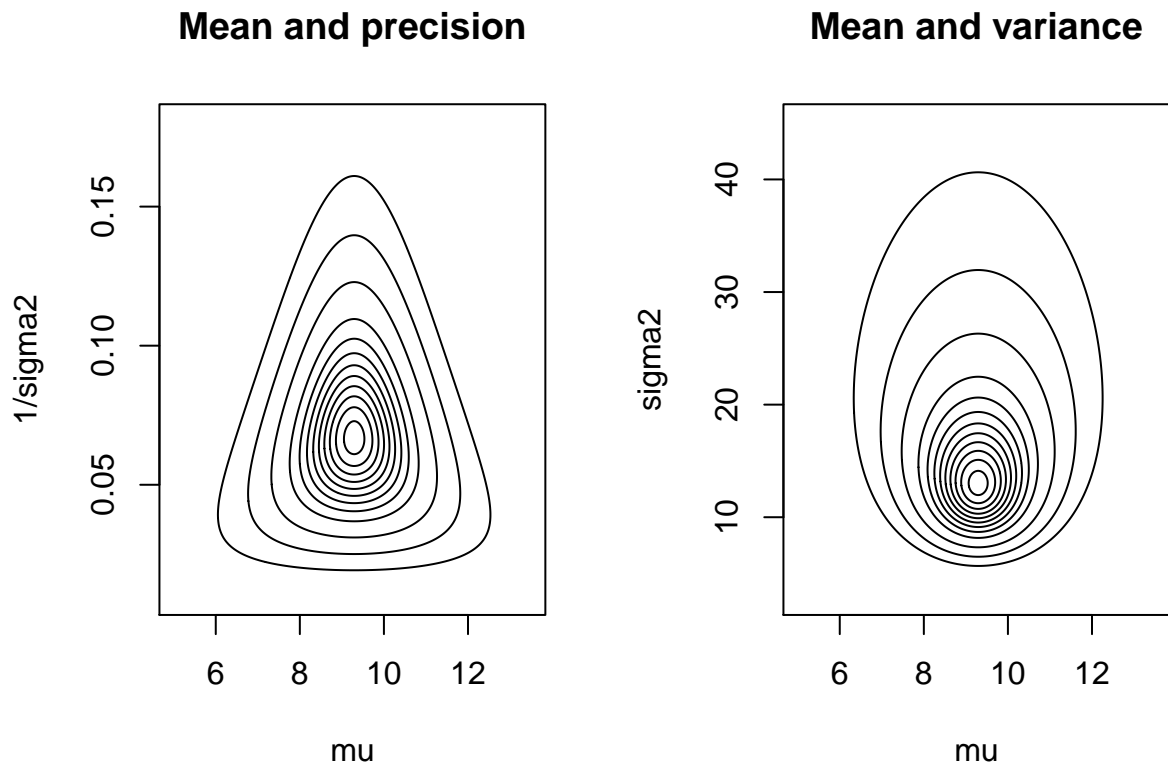
contours <- c(.001, .01, seq(.05, .95, .10))

```

```
op <- par(mfrow=c(1,2))

contour(mu, I.sig2, post.P, levels=contours, drawlabels=F,
        xlab="mu", ylab="1/sigma2", main="Mean and precision")

contour(mu, sigma2, post.V, levels=contours, drawlabels=F,
        xlab="mu", ylab="sigma2", main="Mean and variance")
```



```
par(op)
```

b.

```
S <- 10000
sigma2.sim1 <- (nu.n1 * sigma2.n1 / rchisq(S, df=(nu.n1)))
mu.sim1 <- rnorm(S, mu.n1, sqrt(sigma2.sim1/kappa.n1))

quantile(sigma2.sim1,c(0.025,0.975))
```

```
##      2.5%      97.5%
## 8.962084 26.514058
```

```
quantile(mu.sim1,c(0.025,0.975))
```

```
##      2.5%      97.5%
## 7.764778 10.836542
```

A 95% CI for σ_1^2 is [8.962, 26.51]. A 95% CI for μ_1 is [7.764, 10.837].

c.

```
sigma2.sim2 <- (nu.n2 * sigma2.n2 / rchisq(S, df=nu.n2))
mu.sim2 <- rnorm(S, mu.n2, sqrt(sigma2.sim2/kappa.n2))
sigma2.sim3 <- (nu.n3 * sigma2.n3 / rchisq(S, df=nu.n3))
mu.sim3 <- rnorm(S, mu.n3, sqrt(sigma2.sim3/kappa.n3))

mean(mu.sim1>pmax(mu.sim2,mu.sim3))
```

```
## [1] 0.8933
```

The probability that $\mu_1 > \max(\mu_2, \mu_3)$ is 0.89, that the mean hours studied for students at school 1 is greater than both school 2 and school 3.

d.

```
y1tilde.sim<-rnorm (10000 , mu.sim1 , sqrt(sigma2.sim1))
y2tilde.sim<-rnorm (10000 , mu.sim2 , sqrt(sigma2.sim2))
y3tilde.sim<-rnorm (10000 , mu.sim1 , sqrt(sigma2.sim3))

mean(y1tilde.sim>pmax(y2tilde.sim,y3tilde.sim))
```

```
## [1] 0.3962
```

The probability that $\tilde{y}_1 > \max(\tilde{y}_2, \tilde{y}_3)$ is ~0.4. This is significantly less than the value calculated in part c because the posterior predictive distribution has a higher variance than the distribution of the means. More extreme values of μ_2 and μ_3 are less common such that μ_2 or μ_3 being greater than μ_1 is less likely than in the PPD case.

3.

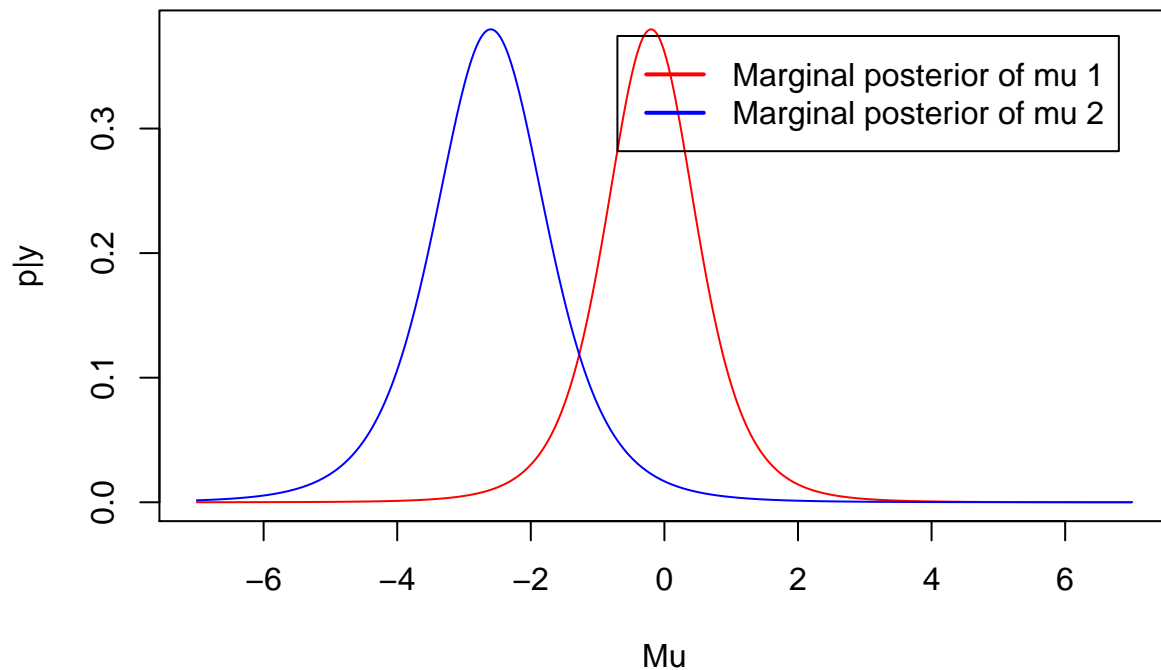
```
y1 <- c(+2.0, -3.1, -1.0, +0.2, +0.3, +0.4)
y2 <- c(-3.5, -1.6, -4.6, -0.9, -5.1, +0.1)
ybar1 <- mean(y1); s1 <- sd(y1);
ybar2 <- mean(y2); s2 <- sd(y2);
```

```
marginalposterior1<-function(mu){
  dt((mu-ybar1)/(s1/sqrt(6)),5)
}
```

```
marginalposterior2<-function(mu){
  dt((mu-ybar2)/(s2/sqrt(6)),5)
}
```

```
mus<-seq(-7,7,0.01)
```

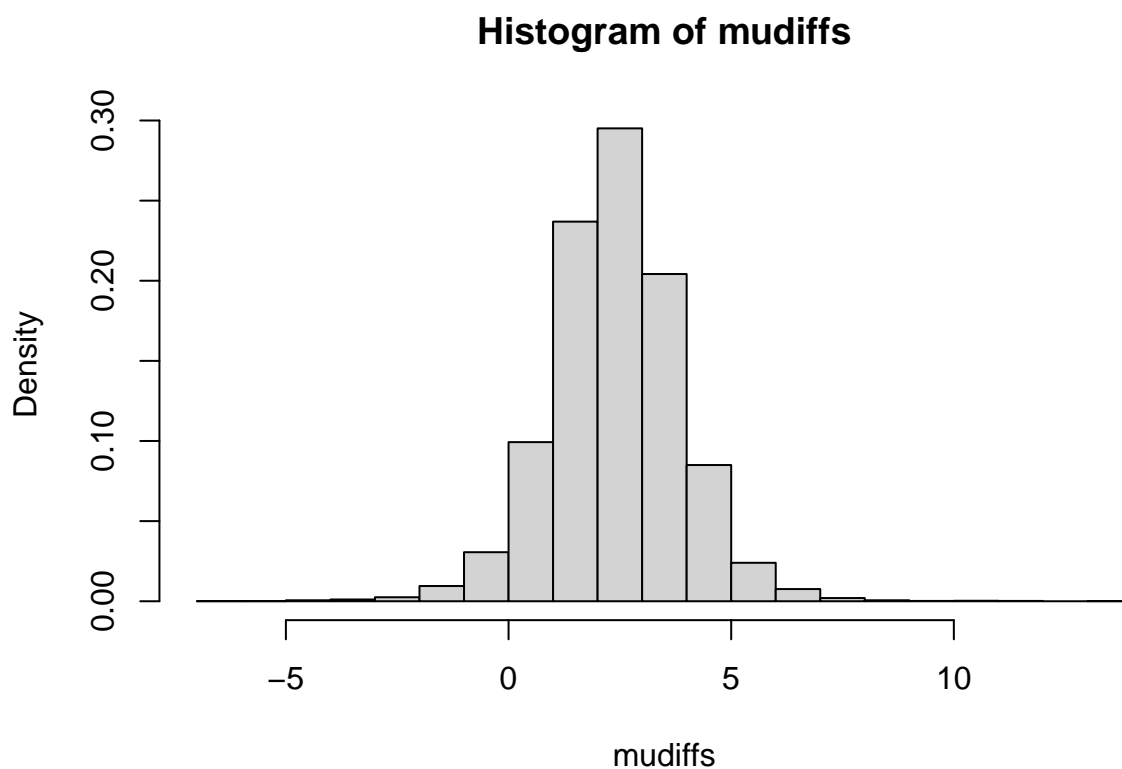
```
plot(mus, marginalposterior1(mus), type = "l", col = "red", xlab = "Mu", ylab = "ply")
lines(mus, marginalposterior2(mus), col = "blue")
legend("topright", inset=.05, lwd=2, col=c("red","blue"),legend=c("Marginal posterior of mu 1", "Marginal posterior of mu 2"))
```



It appears that the majority of samples from the distribution of the treatment group will be less than the majority of the samples from the distribution of the control group, as $\mu_1 > \mu_2$ in most instances. Thus, it appears that the weight loss treatment is effective.

b.

```
mu1.sim<-rt(10000,5)*s1/sqrt(6) + ybar1
mu2.sim<-rt(10000,5)*s2/sqrt(6) + ybar2
mudiffs<-mu1.sim-mu2.sim
hist(mudiffs, freq = F)
```



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

```
##      2.5%      97.5%
## -0.4672251  5.2625065
```

A 95% CI for $\mu_1 - \mu_2$ is $[-0.39, 5.23]$.

c.

```
SSY1<-sum((y1-ybar1)^2)
SSY2<-sum((y2-ybar2)^2)

sigma21.sim<-1/rgamma(10000,5/2,SSY1/2)
sigma22.sim<-1/rgamma(10000,5/2,SSY2/2)

y1tilde.sim<-rnorm(10000,mu1.sim,sigma21.sim)
y2tilde.sim<-rnorm(10000,mu2.sim,sigma22.sim)

mean(y1tilde.sim>y2tilde.sim)
```

```
## [1] 0.6313
```

The probability that $\tilde{y}_1 > \tilde{y}_2$ is ~ 0.63 .

4.

a.

```
library(boot)
prob4 <- read.csv("prob4.csv")
colnames(prob4)<-c("County", "Approve", "Disapprove")
prob4$Total<-prob4$Approve+prob4$Disapprove

y<-prob4$Approve
n<-prob4$Total

mu <- mean(y/n)
psi <- mu * (1-mu) / var(y/n) - 1
logit.mu <- seq ( -1 , 1 , .20)
log.psi <- seq ( 0 , 6 , .20)

log.post.fun <- function(alpha, beta, y, n)
{
  J <- length(y)
  lpost <- -2.5 * log(alpha + beta)
  lpost <- lpost - J * lbeta(alpha, beta)
  lpost <- lpost + sum( lbeta(alpha+y, beta+n-y) )
  lpost
}

I <- length(logit.mu); J <- length(log.psi);

log.post <- matrix(NA, I, J);

for(i in 1:I)
{
  for(j in 1:J)
  {
    mu <- inv.logit(logit.mu[i]); psi <- exp(log.psi[j]);
    alpha <- mu*psi; beta <- (1-mu)*psi;
    log.post[i,j] <- log(alpha) + log(beta) + log.post.fun(alpha,beta,y,n)
  }}

rm(mu, psi, alpha, beta, i, j, I, J)

maxie <- max(log.post);

log.post <- log.post - maxie; rm(maxie);

post <- exp(log.post); rm(log.post);

contours <- c(.001, .01, seq(.05, .95, .10))

delta <- (logit.mu[2] - logit.mu[1]) / 2
epsilon <- (log.psi[2] - log.psi[1]) / 2

post.lm <- apply(post, 1, sum)
```



```

S <- 10000;

logit.mu.sim <- rep(NA, S); log.psi.sim <- rep(NA, S);

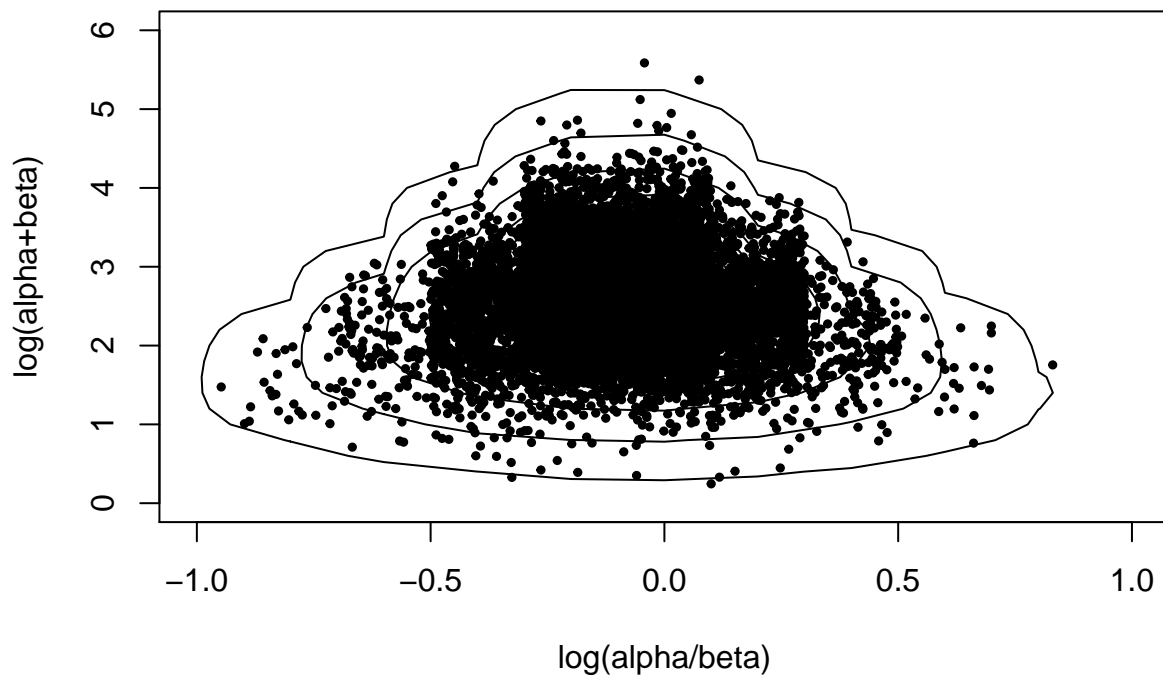
I <- length(logit.mu); J <- length(log.psi);

for(s in 1:S)
{
  i <- sample(I, 1, prob=post.lm)
  j <- sample(J, 1, prob=post[i,])
  logit.mu.sim[s] <- logit.mu[i] + runif(1, -delta, delta)
  log.psi.sim[s] <- log.psi[j] + runif(1, -epsilon, epsilon)
}

contour(logit.mu, log.psi, post, levels=contours, drawlabels=F,
        xlab="log(alpha/beta)", ylab="log(alpha+beta)")

points(logit.mu.sim, log.psi.sim, cex=.5, pch=19)

```



```

J <- length(y)

mu.sim <- inv.logit(logit.mu.sim); psi.sim <- exp(log.psi.sim);

alpha.sim <- mu.sim * psi.sim; beta.sim <- (1-mu.sim) * psi.sim;

theta.sim <- rbeta(J*S, shape1=outer(y, alpha.sim, "+"),

```

```

shape2=outer(n-y, beta.sim, "+")

theta.sim <- matrix(theta.sim, J, S)

meds <- apply(theta.sim, 1, median)

ints <- apply(theta.sim, 1, quantile, prob=c(.05, .95))

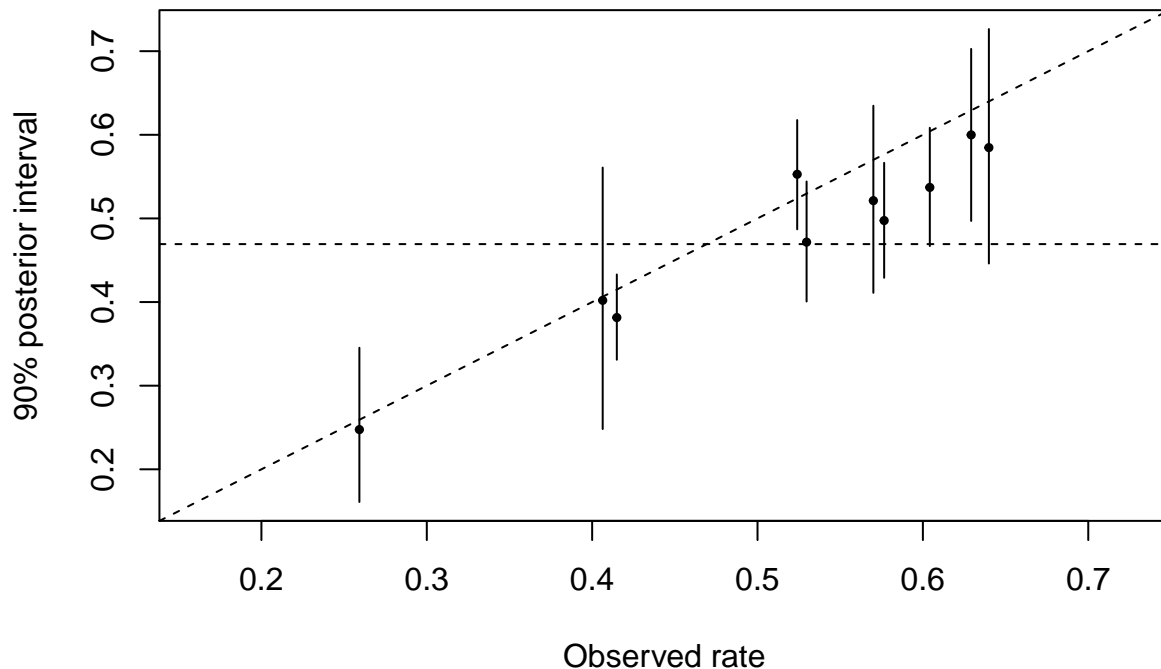
obs <- jitter(y/n, factor=50)

plot(obs, meds, xlim=range(ints), ylim=range(ints), pch=19, cex=.50,
      xlab="Observed rate", ylab="90% posterior interval")

abline(0, 1, lty=2); abline(h=sum(y)/sum(n), lty=2);

for(j in 1:J){ lines(rep(obs[j],2), ints[,j]) }

```



The majority of the estimates we produce using the hierarchical model lie between the diagonal line (unpooled results) and the horizontal line (pooled results). This illustrates how the hierarchical model combines information from both strategies.

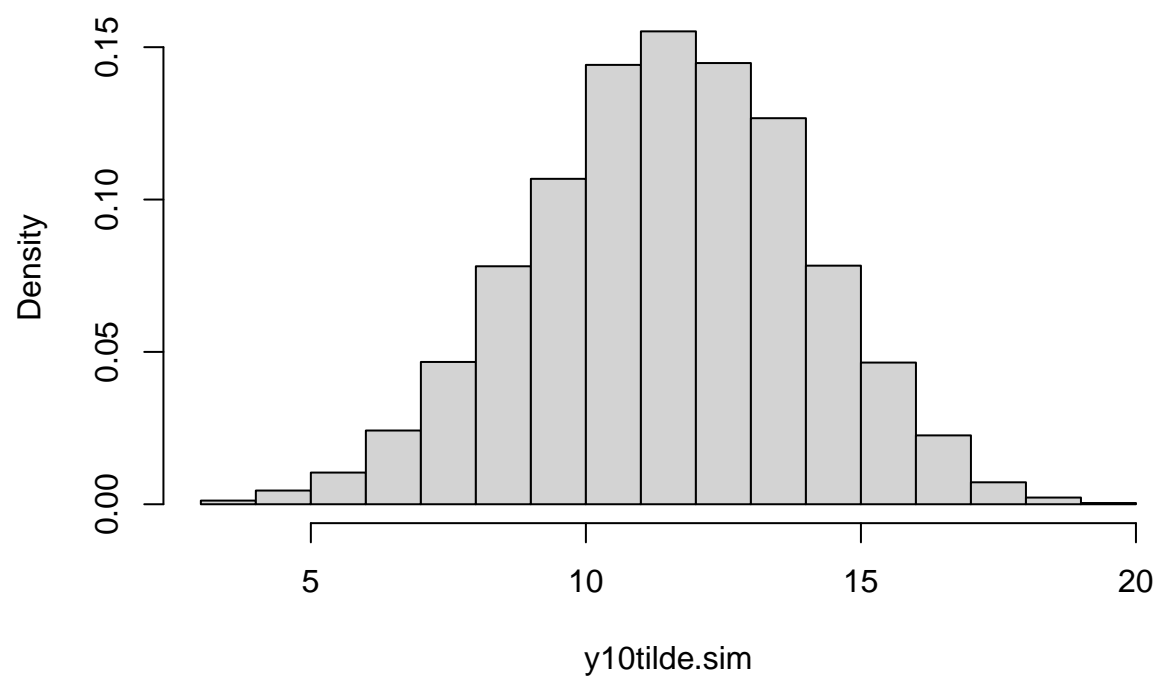
b.

```

y10tilde.sim <- rbinom(10000, 20, theta.sim[10,])
hist(y10tilde.sim, freq = F)

```

Histogram of y10tilde.sim



```
quantile(y10tilde.sim, c(0.05,0.95))
```

```
## 5% 95%
```

```
## 8 16
```

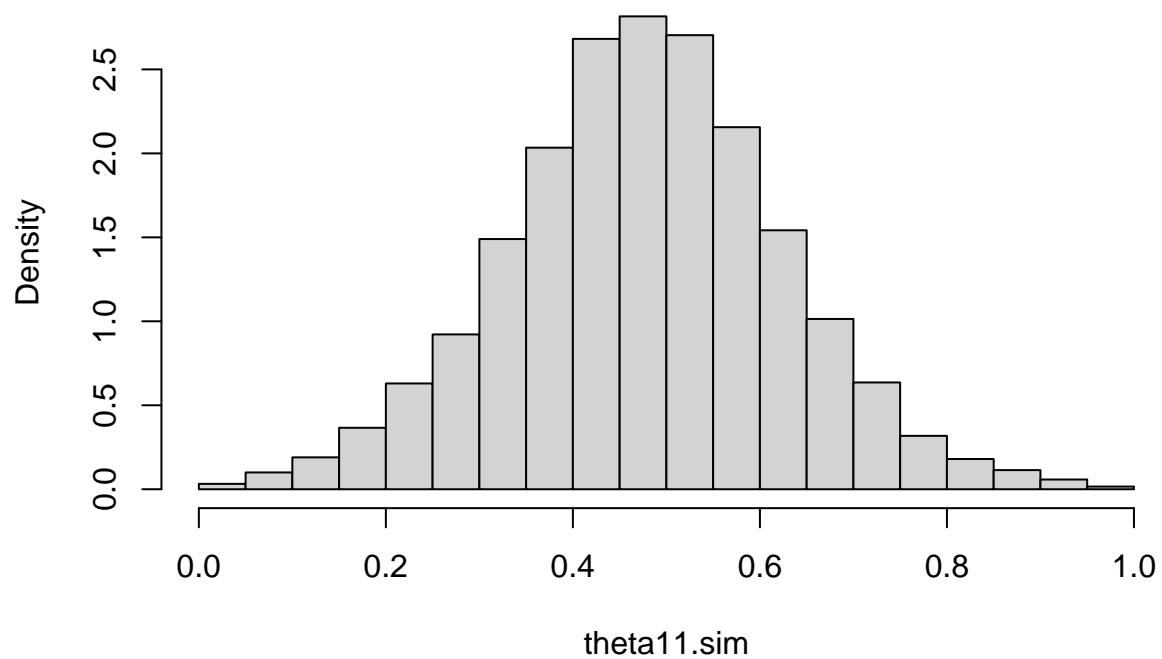
A 90% CI for \tilde{y}_{11} is [8,16].

c.

```
theta11.sim<-rbeta(10000,alpha.sim,beta.sim)
```

```
hist(theta11.sim, freq = F)
```

Histogram of theta11.sim



```
quantile(theta11.sim, c(0.05,0.95))
```

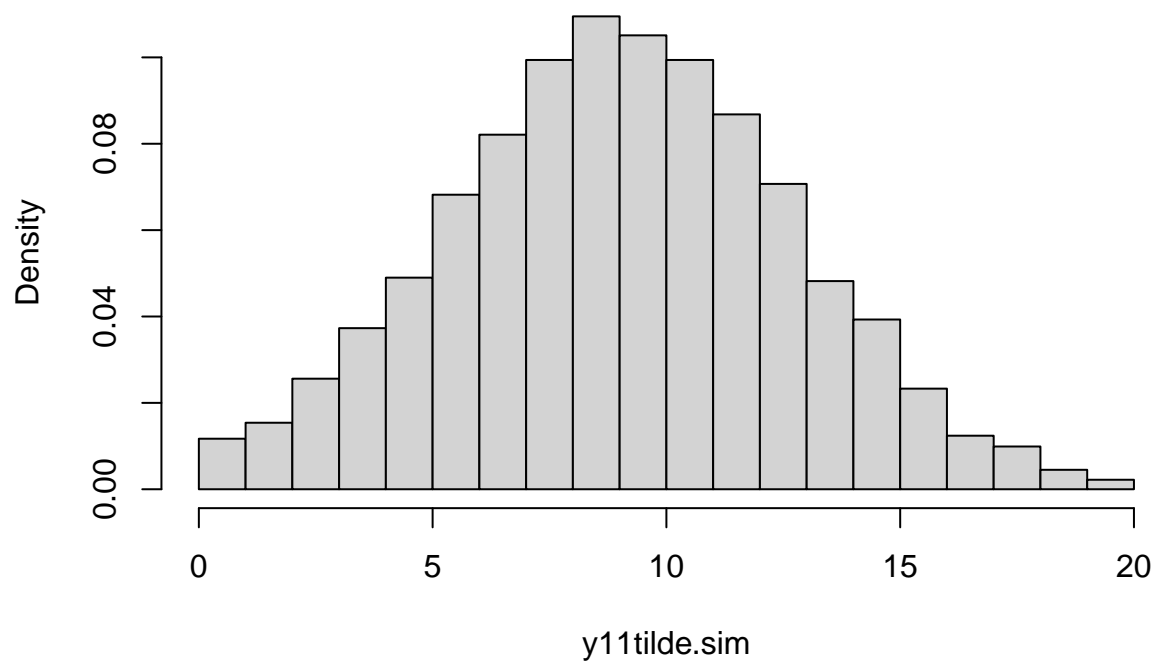
```
##          5%          95%  
## 0.2278693 0.7253743
```

A 90% CI for θ_{11} is [0.23,0.73].

d.

```
y11tilde.sim<-rbinom(10000,20,theta11.sim)  
hist(y11tilde.sim, freq = F)
```

Histogram of y11tilde.sim



```
quantile(y11tilde.sim, c(0.05,0.95))
```

```
## 5% 95%
```

```
## 3 16
```

A 90% CI for \tilde{y}_{11} is [4,16].