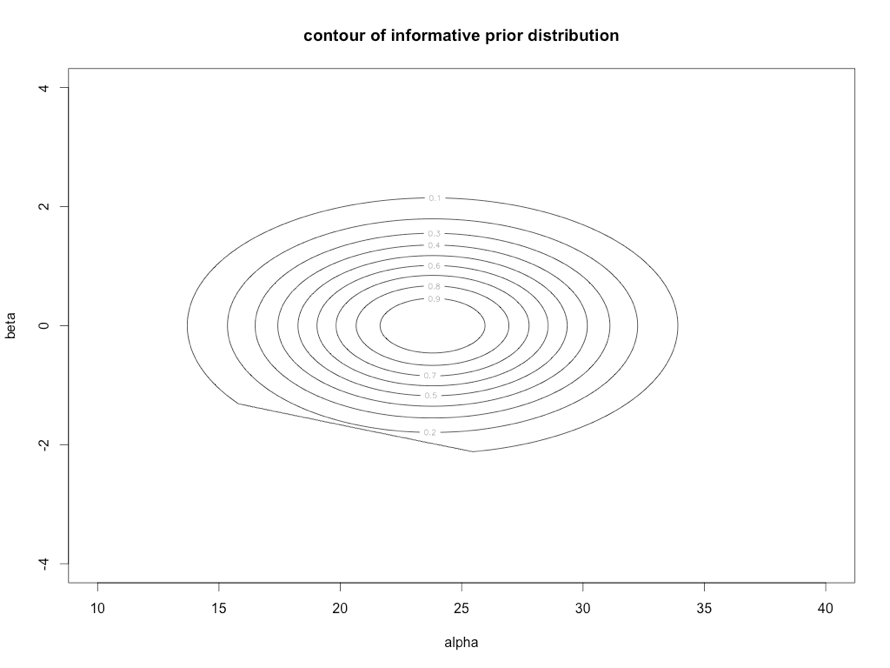
STAT W4640 | HW3 | Haoyang Chen | hc2812

1.

(a)

(b)



y <- c(24,25,31,31,22,21,26,20,16,22)

alpha <- seq(mean(y) - 4 \* sd(y), mean(y) + 4\*sd(y), length.out = 1000)

beta <- seq(-4, 4, length.out = 1000)

prior<-matrix(rep(0,1000000), 1000, 1000)

for (i in 1:1000){

for (j in 1:1000){

if (alpha[i] + 12 \* beta[j] > 0){

prior[i, j] <- exp(-(alpha[i] - mean(y)) ^ 2 \* 0.5 / var(y) - beta[j] ^ 2 / 2)

}

}

}

contour(alpha, beta, prior, xlim = c(10, 40), ylim = c(-4, 4), xlab = "alpha", ylab = "beta", main = "contour of informative prior distribution")

(c)

(d)

(e)

lm(formula = y ~ t)

Residuals:

Min 1Q Median 3Q Max

-4.576 -2.320 -1.761 3.273 5.818

Coefficients:

Estimate Std. Error t value Pr(>|t|)

(Intercept) 28.8667 2.7494 10.499 5.89e-06 \*\*\*

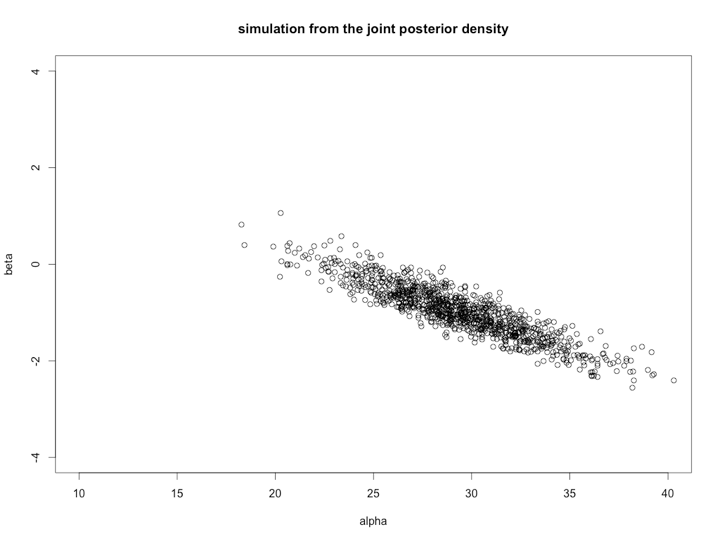
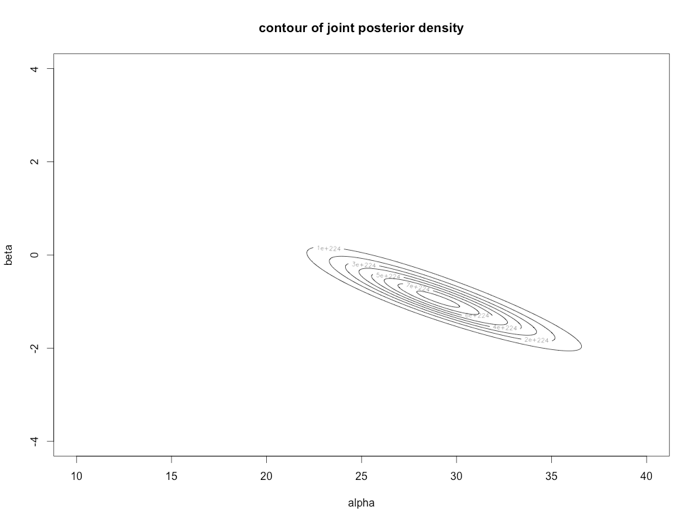
t -0.9212 0.4431 -2.079 0.0712 .

t <- seq(1:10)

fit <- lm(y ~ t)

summary(fit)

(f)



posterior <- matrix(rep(0, 1000000), 1000, 1000)

for (i in 1:1000){

for (j in 1:1000){

if (alpha[i] + 12 \* beta[j] > 0){

# simulate posterior

posterior[i, j] <- prod(exp(-(alpha[i] + beta[j] \* t)) \* (alpha[i] + beta[j] \* t) ^ y)

}

}

}

contour(alpha, beta, posterior, xlim=c(10,40), ylim=c(-4,4), xlab="alpha", ylab="beta", main="contour of joint posterior density")

#take 1000 draws from the joint posterior density

post\_alpha <- apply(posterior, 1, sum)

alpha\_sim<-rep(0, 1000)

beta\_sim<-rep(0, 1000)

for (k in 1:1000){

i <- sample(1000, 1, prob = post\_alpha)

j <- sample(1000, 1, prob=posterior[i,])

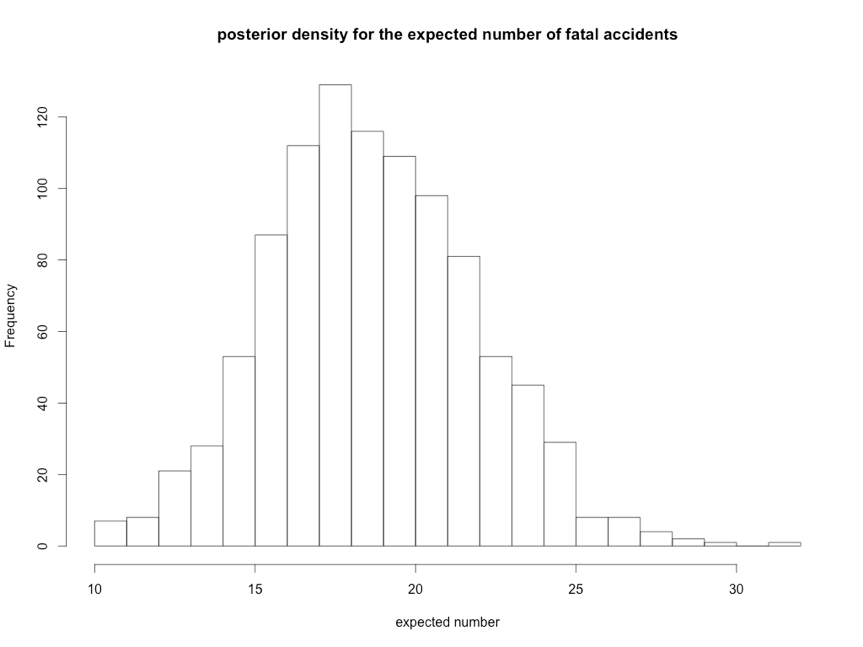
alpha\_sim[k] <- alpha[i]

beta\_sim[k] <- beta[j]

}

plot(alpha\_sim, beta\_sim, xlim=c(10,40), ylim=c(-4,4), xlab="alpha", ylab="beta", main="simulation from the joint posterior density")

(g)



hist(alpha\_sim + 11 \* beta\_sim, breaks = 20, xlab="expected number", main="posterior density for the expected number of fatal accidents")

(h)

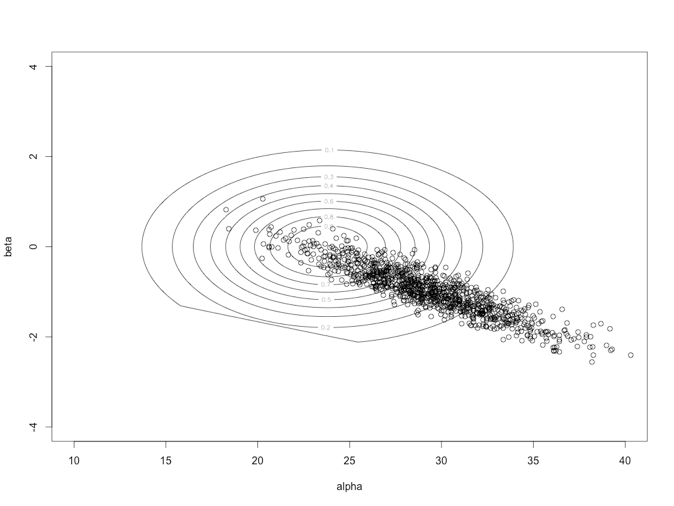
2.5% 97.5%

9.975 31.000

sim <- rpois(1000, alpha\_sim + 11 \* beta\_sim)

quantile(sim, c(0.025,0.975))

(i)



contour(alpha, beta, prior, xlim=c(10,40), ylim=c(-4,4),xlab="alpha",ylab="beta")

points(alpha\_sim, beta\_sim)

2.

(a)

> sum(0.01 \*p \* p\_theta)

[1] 0.5003242

theta <- -seq(-10, 10, 0.01)

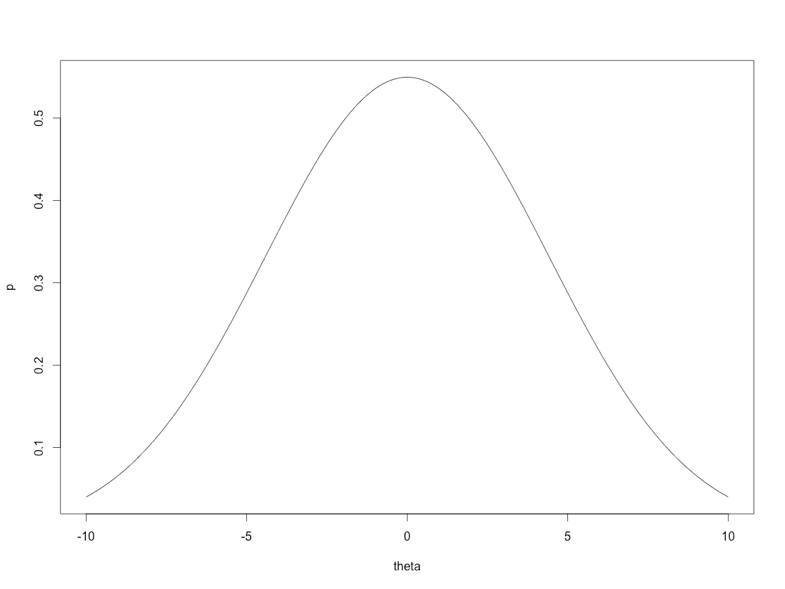
p <- pnorm(theta/4 + 0.675 \* sqrt(5/4)) - pnorm(theta / 4 - 0.675 \* sqrt(5/4))

p\_theta <- 1 / sqrt(8 \* 3.14159265) \* exp(-theta ^ 2 /8)

sum(0.01 \*p \* p\_theta)

(b)

(c)



plot(theta, p, type = 'l')

3.

4.

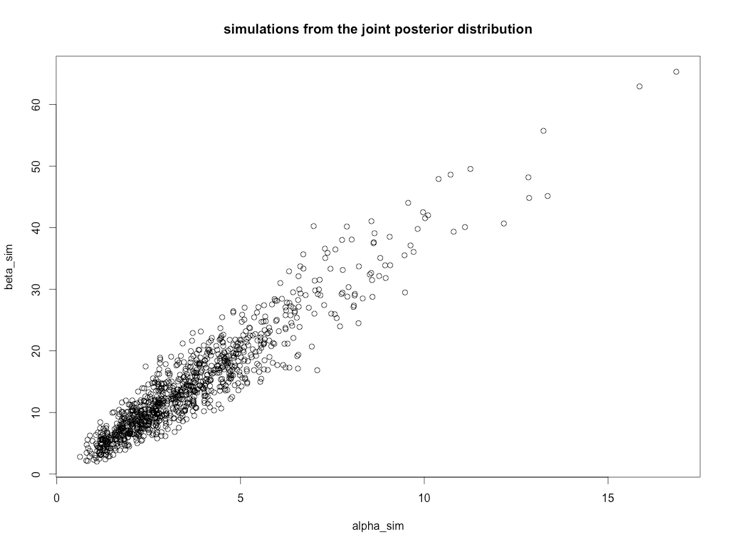
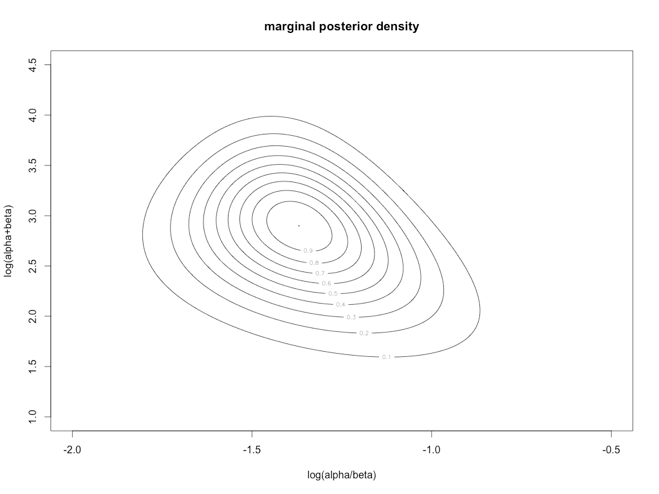
(a)

(b)

5.

(a)

(b)



y <- c(16,9,10,13,19,20,18,17,35,55)

n <- c(58,90,48,57,103,57,86,112,273,64) + y

r <- y/n

alpha <- (mean(r) ^ 2 \* (1 - mean(r))) / var(r) - mean(r)

beta <- (mean(r) \* (1 - mean(r)) ^ 2) / var(r) + mean(r) - 1

a <- seq(-2.0, -0.5, length.out=1000)

b <- seq(1.0, 4.5, length.out=1000)

prior <- matrix(0, 1000, 1000)

posterior <- matrix(0, 1000, 1000)

for (i in 1:1000){

for (j in 1:1000){

alpha\_sim <- exp(a[i] + b[j]) / (exp(a[i]) + 1)

beta\_sim <- exp(b[j]) / (exp(a[i]) + 1)

prior[i, j] <- (alpha\_sim + beta\_sim) ^ (-2.5) \* alpha\_sim \* beta\_sim

posterior[i, j] <- prior[i, j] \* prod(beta(alpha\_sim + y, beta\_sim + n - y) / beta(alpha\_sim, beta\_sim))

}

}

contour(a, b, posterior/max(posterior), xlab="log(alpha/beta)", ylab="log(alpha+beta)", main="marginal posterior density")

post\_alpha <- apply(posterior, 1, sum)

alpha\_sim = rep(0,1000)

beta\_sim = rep(0,1000)

for(i in 1:1000){

loc <- sample(1000, 1, prob=post\_alpha)

a\_sim <- a[loc]

post\_beta <- posterior[loc,]

b\_sim <- sample(b, 1, prob=post\_beta)

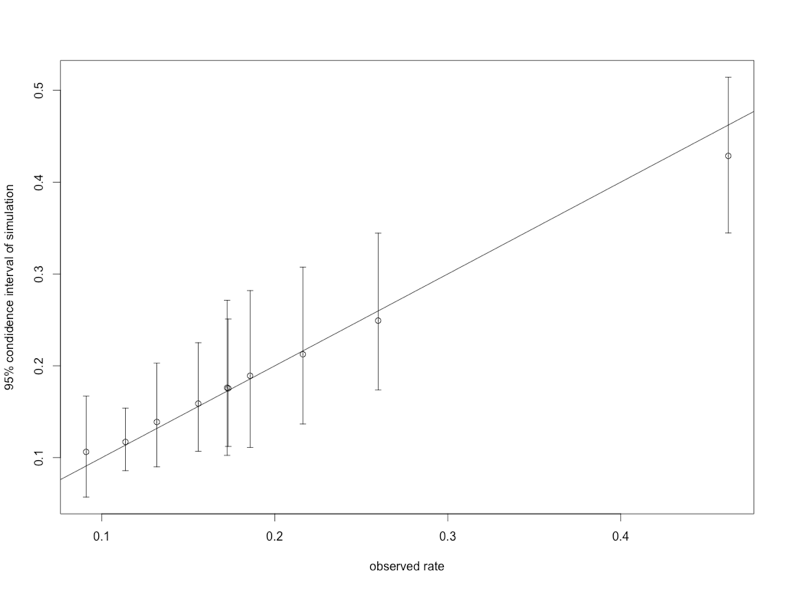
alpha\_sim[i] <- exp(a\_sim + b\_sim) / (exp(a\_sim) + 1)

beta\_sim[i] <- exp(b\_sim) / (exp(a\_sim) + 1)

}

plot(alpha\_sim, beta\_sim, main="simulations from the joint posterior distribution")

(c)



library(plotrix)

theta <- matrix(rep(0, 10 \* 1000), 10, 1000)

for(i in 1:10)

for(j in 1:1000){

{

theta[i,j]=rbeta(1, alpha\_sim[j] + y[i], beta\_sim[j] + n[i] - y[i])

}

}

mean <- rowMeans(theta)

median <- apply(theta, 1, median)

lower <- vector()

upper <- vector()

for(i in 1:10){

lower[i] <- quantile(theta[i,],0.025)

upper[i] <- quantile(theta[i,],0.975)

}

theta\_raw <- y / n

plotCI(theta\_raw, median, ui=upper, li=lower, xlab="observed rate", ylab="95% condidence interval of simulation")

abline(a=0,b=1)

(d)

> mean(lower)

[1] 0.1320606

> mean(upper)

[1] 0.2718957

(e)

> CI

[1] 11 28

y\_new<-seq(0, 100, 0.1)

theta\_avg<-colMeans(theta)

post\_pred<-rep(0, 101)

for (i in 1:1000){

post.pred<-post\_pred + dbinom(y\_new, 100, theta\_avg[i]) / 1000

}

prob <- 0

CI <- c(0, 0)

for (i in y\_new){

prob <- prob + post\_pred[i + 1]

if(prob <= 0.025) CI[1] <- i

if(prob >= 0.975) {

CI[2] <- i

break

}

}

CI

(f)

8.

(a)

(b)

(c)