

Bayesian Learning Lab1

Weng Hang Wong

4/8/2020

1. Bernoulli ... again.

Let $y_1, \dots, y_n | \theta \sim \text{Bern}(\theta)$ and assume that you have obtained a sample with $s = 5$ successes in $n = 20$ trials. Assume a $\text{Beta}(\alpha_0, \beta_0)$ prior for θ and let $\alpha_0 = \beta_0 = 2$

(a) Draw random numbers from the posterior $\theta | y \sim \text{Beta}(\alpha_0 + s, \beta_0 + f)$, $y = (y_1, \dots, y_n)$ and verify graphically that the posterior mean and standard deviation converges to the true values as the number of random draws grows large.

Verification

To verify graphically that the posterior mean and standard deviation, meaning that it is needed to use the mean and variance from the Beta Probability Density Function.

Mean of Beta PDF: $\frac{\alpha}{\alpha + \beta}$

Variance of Beta PDF: $\frac{\alpha\beta}{(\alpha + \beta)^2(\alpha + \beta + 1)}$

Bernouli Distruibution

The probability density function of Bernouli distribution is:

$$p(s) = \binom{n}{s} \theta^s (1 - \theta)^{n-s}$$

The likelihood function is:

$$p(y_1, \dots, y_n | \theta) = \binom{n}{s} \theta^s (1 - \theta)^{n-s}$$

The prior is given $\text{Beta}(\alpha_0, \beta_0)$:

$$p(\theta) = \frac{\Gamma(\alpha_0 + \beta_0)}{\Gamma(\alpha_0)\Gamma(\beta_0)} \theta^{\alpha_0-1} (1 - \theta)^{\beta_0-1}$$

We have:

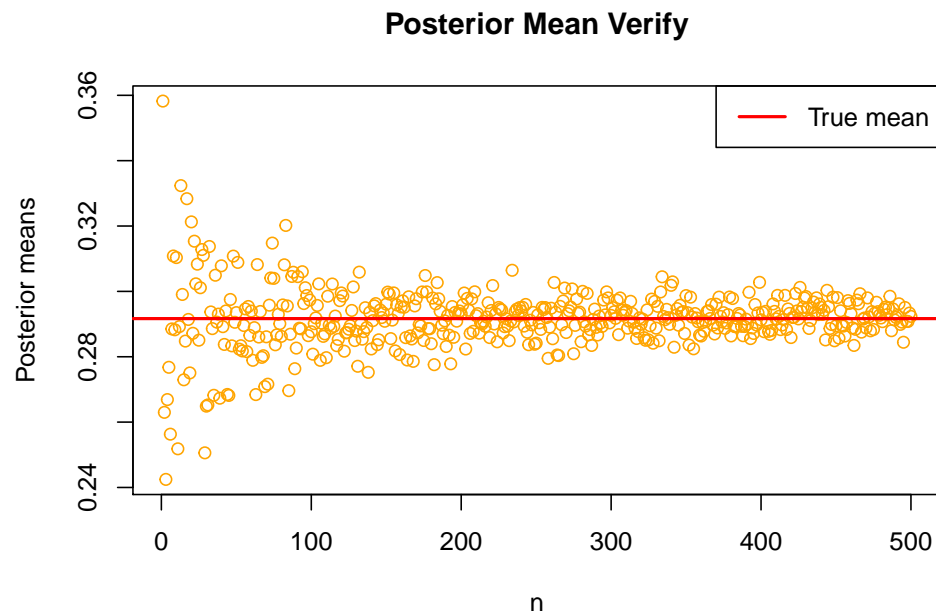
Posterior \propto likelihood \cdot prior

$$\theta | y \sim \text{Beta}(\alpha_0 + s, \beta_0 + f)$$

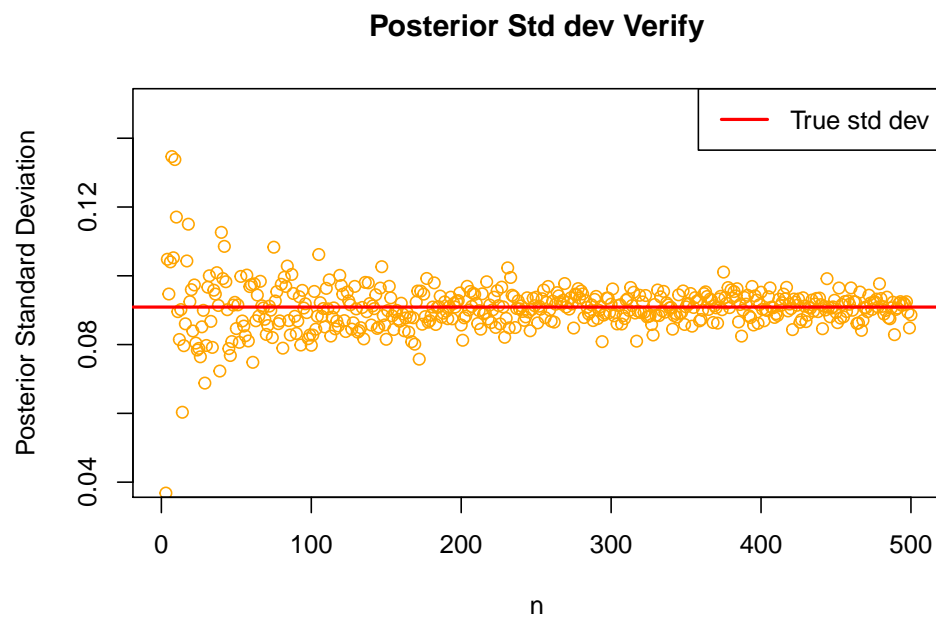
From the above given posterior, we know that $\alpha = \alpha_0 + s = 7$, $\beta = \beta_0 + f = 17$. By using the alpha and beta we have, it is possible to simulate the random samples from the `rbeta()` function. With the growing number of random draws (1:500), we can plot the samples and investigate how the mean and standard deviation distributed.

```
## [1] 0.09090593
```

```
## The true mean is: 0.2916667
```



The true Standard deviation is: 0.09090593



According to the two figures above, it is successfully verified the posterior mean and standard deviations are converged to the true value.

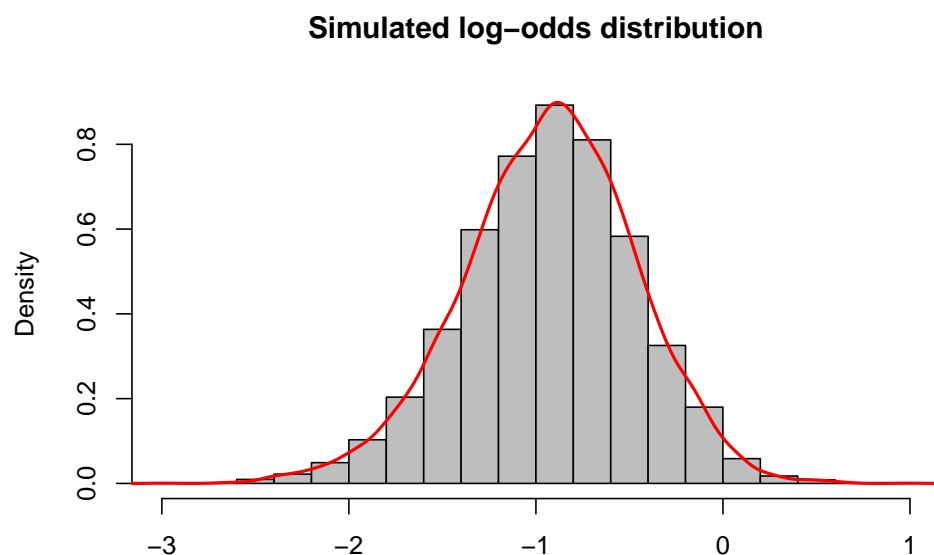
(b) Use simulation (nDraws = 10000) to compute the posterior probability $Pr(\theta > 0.3|y)$ and compare with the exact value [Hint: pbeta()].

From the simulation of nDraws = 10000, the posterior probability $Pr(\theta > 0.3|y)$ is 0.4392, and it is very close to the true probability computed from the pbeta() function, which is 0.4399472.

```
## The simulation posterior probability (> 0.3) is: 0.4392
```

```
## The true probability is : 0.4399472
```

(c) Compute the posterior distribution of the log-odds $\phi = \log \frac{\theta}{1-\theta}$ (nDraws = 10000). [Hint: hist() and density()] might come in handy]



2 Log-normal distribution and the Gini coefficient.

Assume that you have asked 10 randomly selected persons about their monthly income (in thousands Swedish Krona) and obtained the following ten observations: 44, 25, 45, 52, 30, 63, 19, 50, 34 and 67. A common model for non-negative continuous variables is the log-normal distribution. The log-normal distribution $\log(N(\mu, \sigma^2))$ has density function: *

$$p(y|\mu, \sigma^2) = \frac{1}{y * \sqrt{2\pi\sigma^2}} e^{-\frac{1}{2\sigma^2}(\log(y)-\mu)^2}$$

For $y > 0$, $\mu > 0$ and $\sigma^2 > 0$. The log-normal distribution is related to the normal distribution as follows: if $y \sim \log N(\mu, \sigma^2)$ then $\log(y) \sim N(\mu, \sigma^2)$. Let $y_1, \dots, y_n | \mu, \sigma^2 \stackrel{iid}{\sim} \log N(\mu, \sigma^2)$, where $\mu = 3.7$ is assumed to be known but σ^2 is unknown with non-informative prior $p(\sigma^2) \propto \frac{1}{\sigma^2}$. The posterior for the σ^2 is the $Inv - \chi^2(n, \tau^2)$ distribution, where

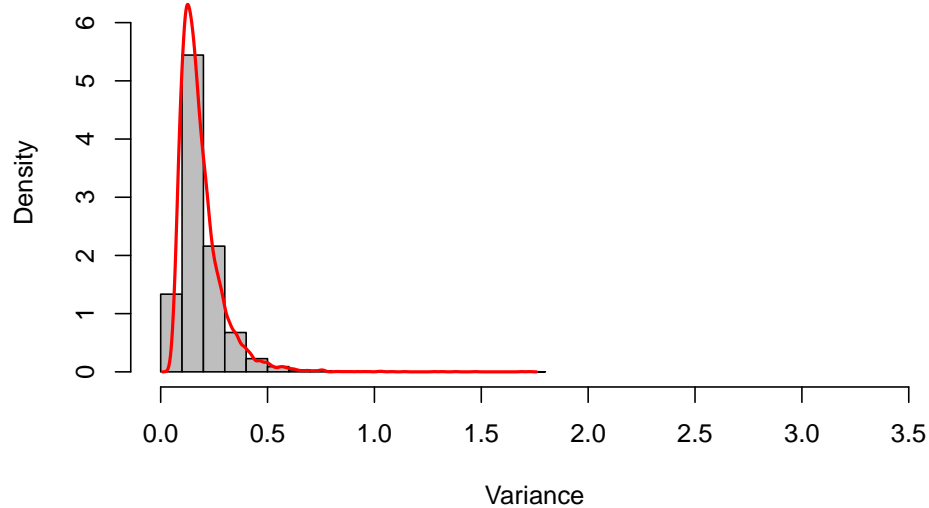
$$\tau^2 = \frac{\sum_{i=1}^n (\log(y_i) - \mu)^2}{n}$$

(a) Simulate 10,000 draws from the posterior of σ^2 (assuming $\mu = 3.7$) and compare with the theoretical $Inv - \chi^2(n, \tau^2)$ posterior distribution.

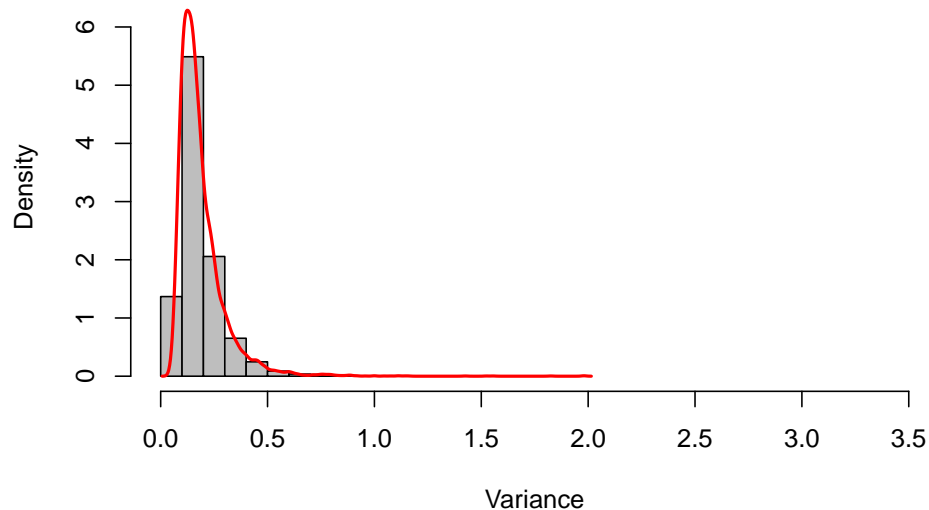
We can simulate 10,000 samples from the posterior of σ^2 using `rchisq()`. First, we draw $X \sim \chi^2(n - 1)$ then compute $\sigma^2 = \frac{\tau^2 n}{X}$. To compare with the theoretical $Inv - \chi^2(n, \tau^2)$, where the scaled parameter is τ^2 , we can use the `invrchisq()` draw 10,000 samples of the variance. (Source: Wikipedia https://en.wikipedia.org/wiki/Scaled_inverse_chi-squared_distribution)

Looking at the two histogram belows, there are no difference between them.

Simulated Inv-Chi-Square



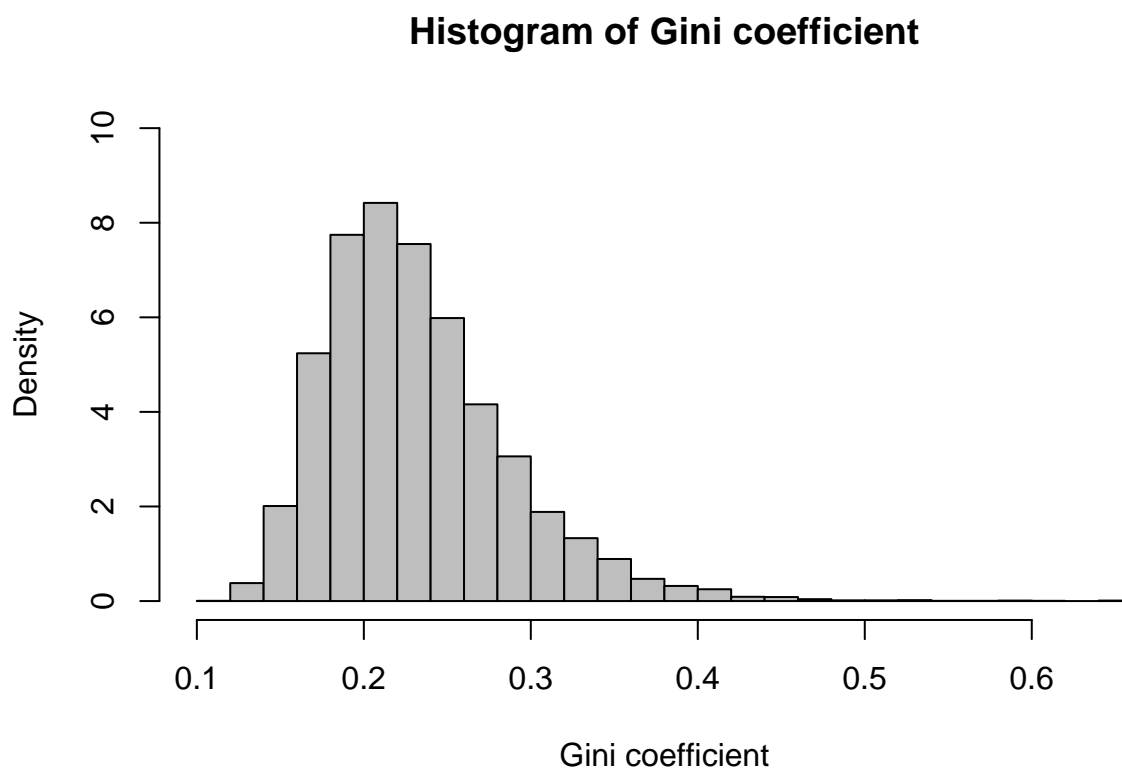
Theoretical Inv-Chi-Square distribution



(b) The most common measure of income inequality is the Gini coefficient, G , where $0 \leq G \leq 1$. $G = 0$ means a completely equal income distribution, whereas $G = 1$ means complete income inequality. See Wikipedia for more information. It can be shown that $G = 2\phi(\frac{\sigma}{\sqrt{2}} - 1)$ when incomes follow a $\log N(\mu, \sigma^2)$ distribution. $\phi(z)$ is the cumulative distribution function (CDF) for the standard normal distribution with mean zero and unit variance. Use the posterior draws in a) to compute the posterior distribution of the Gini coefficient G for the current data set.

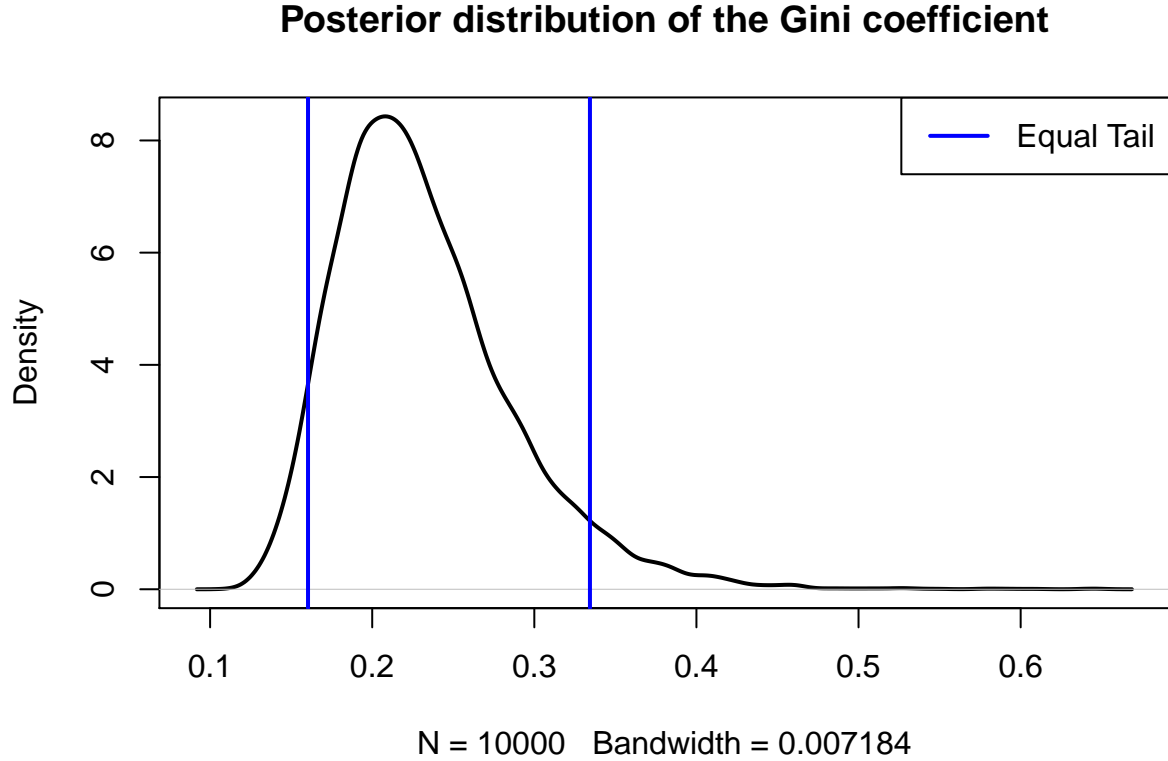
To compute the posterior distribution of the Gini coefficient, we use the posterior variance σ^2 we had from above and apply it to the cumulative distribution function for the standard normal distribution.

$$G = 2\Phi(\sigma/\sqrt{2}) - 1$$



From the above histogram, the Gini coefficient is mainly concentrated on about 0.2 and it's overall lower than 0.5, meaning that we can easily say the income distribution is quite equal.

(c) Use the posterior draws from b) to compute a 90% equal tail credible interval for G. A 90% equal tail interval (a, b) cuts off 5% percent of the posterior probability mass to the left of a, and 5% to the right of b. Also, do a kernel density estimate of the posterior of G using the density function in R with default settings, and use that kernel density estimate to compute a 90% Highest Posterior Density interval for G. Compare the two intervals.



3. Bayesian inference for the concentration parameter in the von Mises distribution.

This exercise is concerned with directional data. The point is to show you that the posterior distribution for somewhat weird models can be obtained by plotting it over a grid of values. The data points are observed wind directions at a given location on ten different days. The data are recorded in degrees:

(40, 303, 326, 285, 296, 314, 20, 308, 299, 296),

where North is located at zero degrees. To fit with Wikipedia's description of probability distributions for circular data we convert the data into radians $-\pi \leq y \leq \pi$. The 10 observations in radians are

(-2.44, 2.14, 2.54, 1.83, 2.02, 2.33, -2.79, 2.23, 2.07, 2.02).

Assume that these data points are independent observations following the von Mises distribution

$$p(y|\mu, \kappa) = \frac{\exp[\kappa \cdot \cos(y - \mu)]}{2\pi I_0(\kappa)}, \pi \leq y \leq \pi,$$

where $I_0(\kappa)$ is the modified Bessel function of the first kind of order zero. The parameter $\mu(\pi \leq y \leq \pi)$ is the mean direction and $\kappa > 0$ is called the concentration parameter. Large κ gives a small variance around μ , and vice versa. Assume that μ is known to be 2.39. Let $\kappa \sim \text{Exponential}(\lambda = 1)$ a priori, where λ is the rate parameter of the exponential distribution (so that the mean is $1/\lambda$).

(a). Plot the posterior distribution of κ for the wind direction data over a fine grid of κ values.

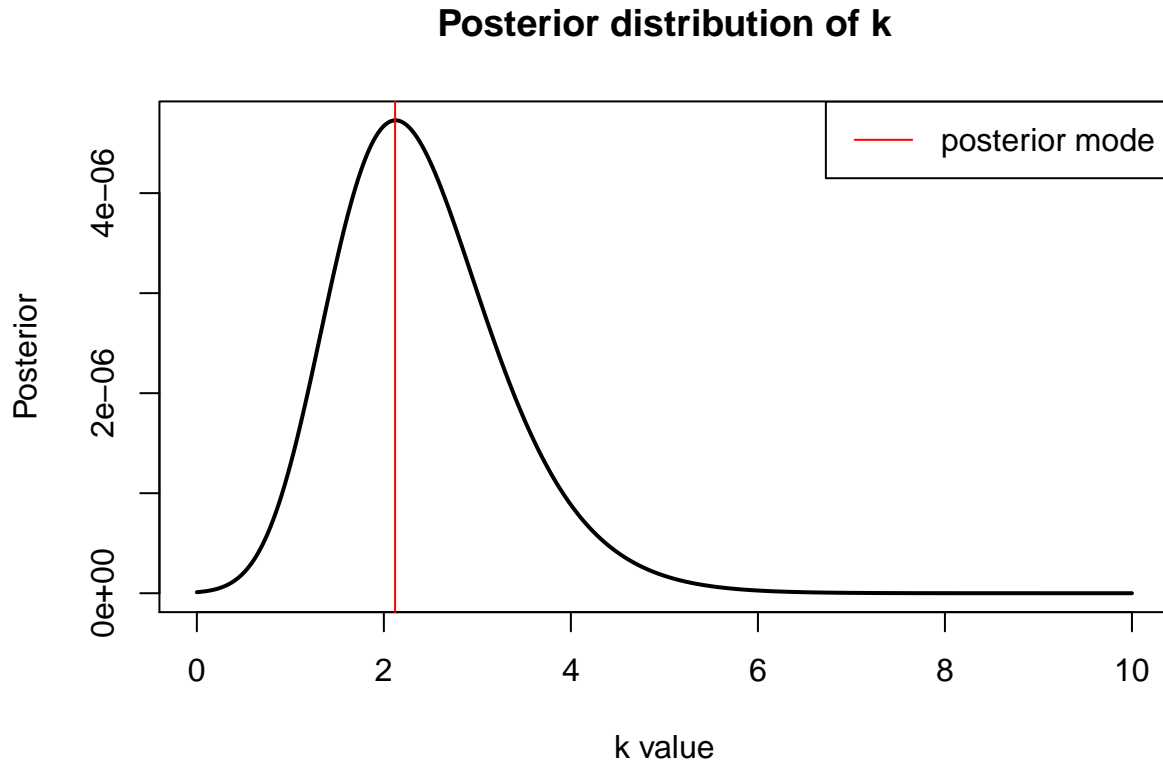
To find the Posterior distribution, we need to compute $\text{Posterior} \propto \text{Likelihood} \cdot \text{prior}$.

Posterior

$$p(\mu, \kappa | y) \propto \frac{\exp[\kappa \cdot (\sum_{i=1}^n \cos(y_i - \mu) - 1)]}{(2\pi I_0(\kappa))^n}$$

(b). Find the (approximate) posterior mode of κ from the information in a).

The posterior mode of k is: 2.12



Appendix

```

#1.a

## Draw random numbers from the posterior
set.seed(12345)
beta_sample = function(n,alpha, beta){
  sample = rbeta(n, alpha, beta)
  mean_sam = mean(sample)
  sd_sam = sd(sample)
  return(list("Mean"=mean_sam,"sd"= sd_sam))
}

alpha=7
beta=17

true_mean = alpha/(alpha+beta)
true_sd = sqrt((alpha*beta) / (((alpha+beta)^2)*(alpha+beta+1)))
true_sd

##verify of n =500
n=500
veri_means=c()
veri_sd = c()
for(i in 1:n){
  veri_means[i] = beta_sample(i, alpha, beta)$Mean
  veri_sd[i]=beta_sample(i, alpha, beta)$sd
}

cat("The true mean is:", true_mean)
plot( veri_means, main="Posterior Mean Verify", xlab="n", ylab="Posterior means",col="orange")
abline(h=true_mean, col="red", lwd=2)
legend("topright",legend="True mean", lty=1,col="red",lwd=2)

cat("The true Standard deviation is:", true_sd)
plot(veri_sd, main="Posterior Std dev Verify", xlab="n", ylab="Posterior Standard Deviation",ylim=c(0.0
abline(h=true_sd, col="red", lwd=2)
legend("topright",legend="True std dev", lty=1,col="red",lwd=2)

# 1.b
set.seed(12345)
nDraws=10000
beta_sample2 = rbeta(nDraws, alpha, beta)

sample_prob = sum(beta_sample2 > 0.3)/nDraws
real_prob = pbeta(q=0.3,alpha, beta, lower.tail = FALSE)

cat("The simulation posterior probability (> 0.3) is: ",sample_prob)
cat("The true probability is :", real_prob)

#1.c
log_odd = log(beta_sample2/(1-beta_sample2))

hist(log_odd, freq = F, main="Simulated log-odds distribution", xlab="",col="grey")

```



```

lines(density(log_odd), col="red", lwd=2)
#log_odd
#2.a
library(LaplacesDemon)
y = c(44, 25, 45, 52, 30, 63, 19, 50, 34 ,67)
n=10
mu = 3.7
nDraws =10000

tau_2 = sum((log(y)-mu)^2)/n

###simulation
#set.seed(12345)
var=c()
for(i in 1:nDraws){
  chi_sample = rchisq(n=1, df=n)
  var[i] = (tau_2*n)/chi_sample
}

library(LaplacesDemon)
###therotical inv-chi-square distuibtion (n,T^2)
thero_tau = rinvcchisq(nDraws, n, tau_2)

### plot the posterior var and the theoretical var
hist(var, freq = F,xlim=c(0,3.5),ylim=c(0,6.5), main = "Simulated Inv-Chi-Square", xlab="Variance",col="red", lwd=2)
lines(density(var), col="red", lwd=2)

hist(thero_tau,freq=F,, xlim=c(0,3.5),ylim=c(0,6.5), main="Theoretical Inv-Chi-Square distribution", xlab="Variance",col="red", lwd=2)
lines(density(thero_tau), lwd=2, col="red")

# 2.b

### gini coefficient
gini = 2*pnorm(sqrt(var/2))-1

hist(gini, freq=F, ylim=c(0,10),breaks=30, main="Histogram of Gini coefficient", xlab="Gini coefficient",col="red", lwd=2)

#2.c

### equal tail CI
CI=quantile(gini, c(0.05,0.95))
#0.1602085 0.3357992

### HPD
gini_x = density(gini)$x
gini_dens = density(gini)$y

plot(density(gini), lwd=2, main="Posterior distribution of the Gini coefficient")

```

```

abline(v=CI, col="blue",lwd=2)
legend("topright",legend="Equal Tail", lty=1,col="blue",lwd=2)

## 3.a
y = c(-2.44, 2.14, 2.54, 1.83, 2.02, 2.33, -2.79, 2.23, 2.07, 2.02)
mu = 2.39
n = length(y)

#kvalues
k = seq(0,10,0.01)

#posterior distri
pos = exp(k * (sum(cos(y-mu)) -1)) / (2*pi*besseli(k,0))^n

# 3.b
cat("The posterior mode of k is:", k[which.max(pos)])
# 2.12

plot(k,pos, type="l",main="Posterior distribution of k", xlab="k value",ylab="Posterior", lwd=2)
abline(v=k[which.max(pos)], col="red")
legend("topright",legend="posterior mode", lty=1,col="red")

```