Lab1 Report

Ludvig Noring, Michael Sörsäter

April 10, 2017

1. Bernoulli

(a)

Using the sample with 14 successes and 6 failures and a Beta prior with $\alpha = \beta = 2$ we get the posterior $\theta|y \sim Beta(16, 8)$ from which we draw random numbers. As the number of draws increases the sampled mean and standard deviation converges towards the theoretical mean and standard deviation for the posterior. This can be seen in Figure 1 and Figure 2.

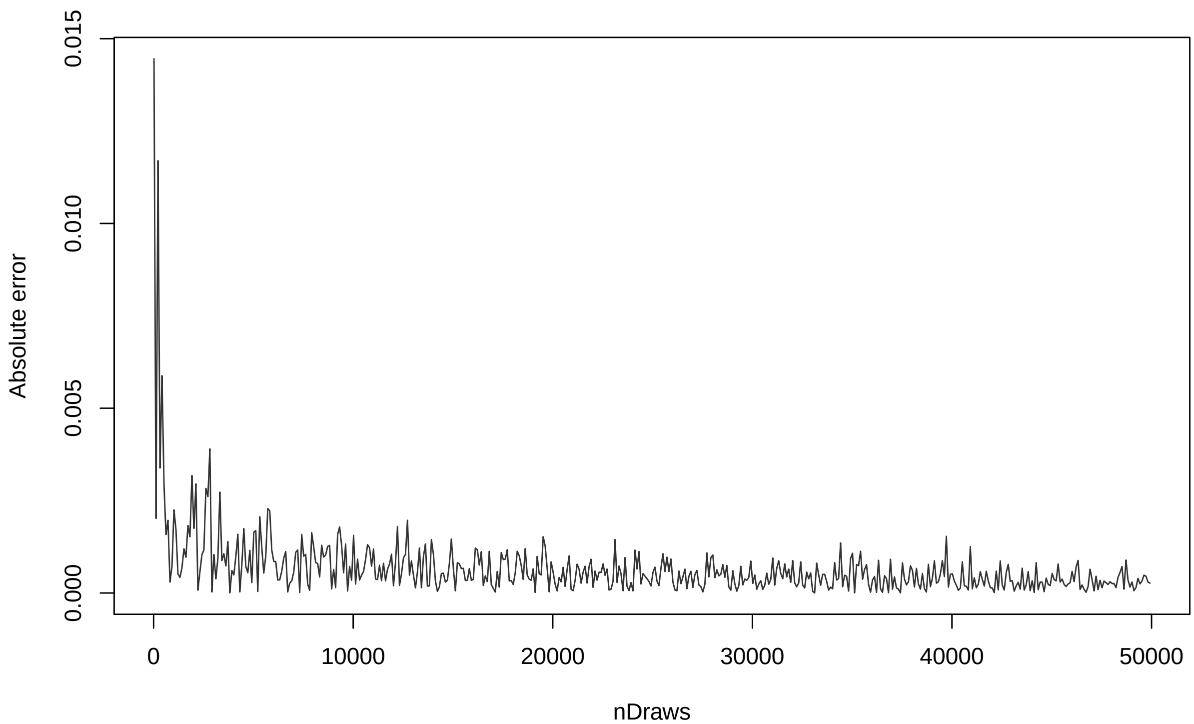


Figure 1: Absolute error of the sampled and theoretical mean

(b)

With 10 000 simulated draws we compute the posterior probability to $Pr(\theta < 0.4|y) = 0.35\%$. This can be compared with the theoretical value 0.397%.

With just so few as 10 000 draws we get a probability close to the exact value.

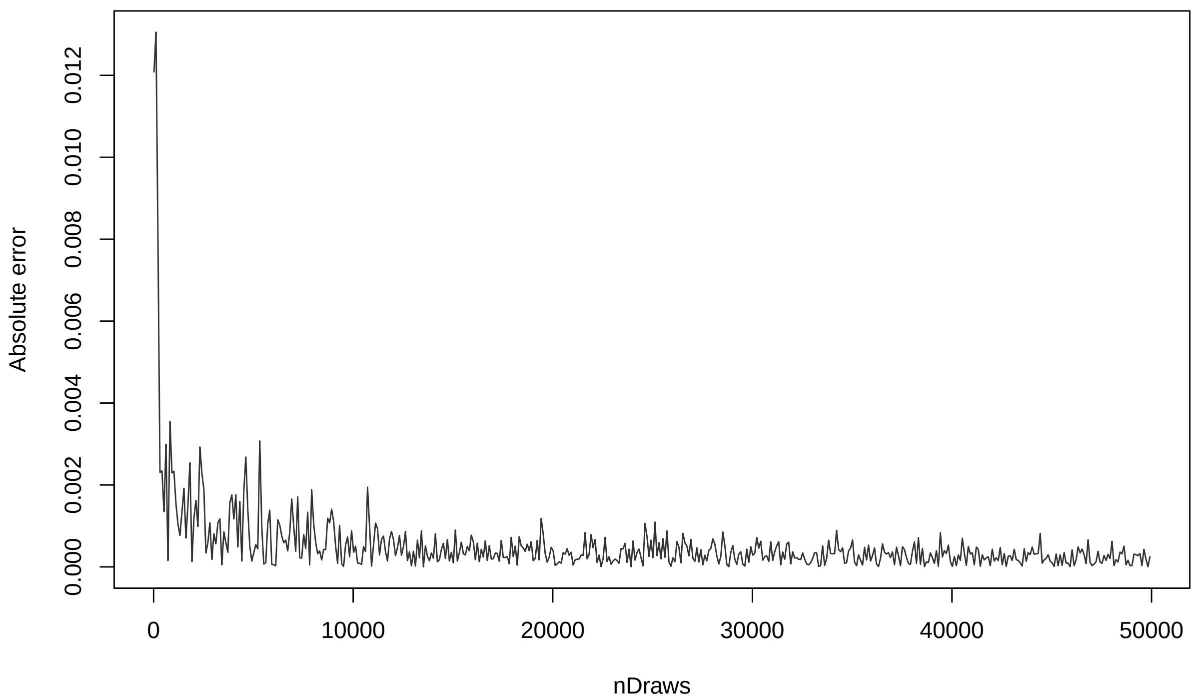


Figure 2: Absolute error of the sampled and theoretical standard deviation

(c)

It is easy to compute the log-odds of the posterior distribution. The plot is seen in Figure 3.

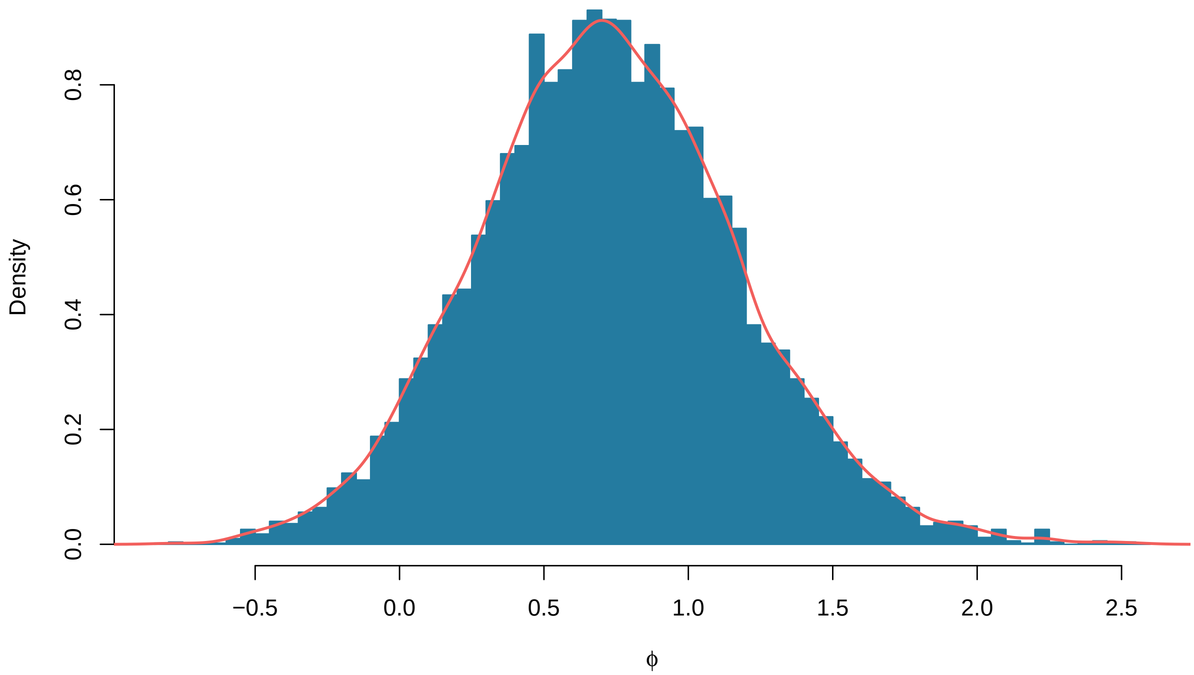


Figure 3: Log-odds

2. Log-normal distribution and the Gini coefficient

(a)

10 000 draws are simulated from the posterior $Inv - \chi^2(n, \tau^2)$. The result is plotted with the theoretical distribution and is seen in Figure 4. They are very similar.

(b)

We calculate the Gini coefficient for different values of σ and plot the result which can be seen in Figure 5.

(c)

TODO

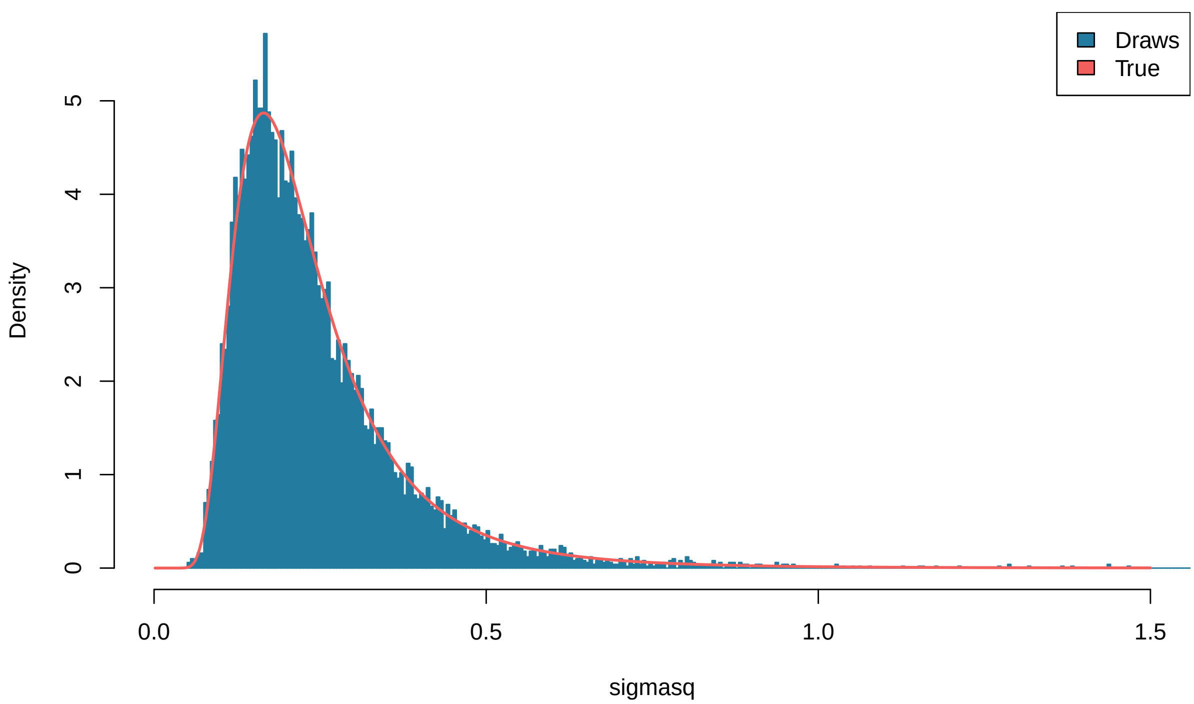


Figure 4: $Inv - \chi^2(n, \tau^2)$ Simulated and theoretical

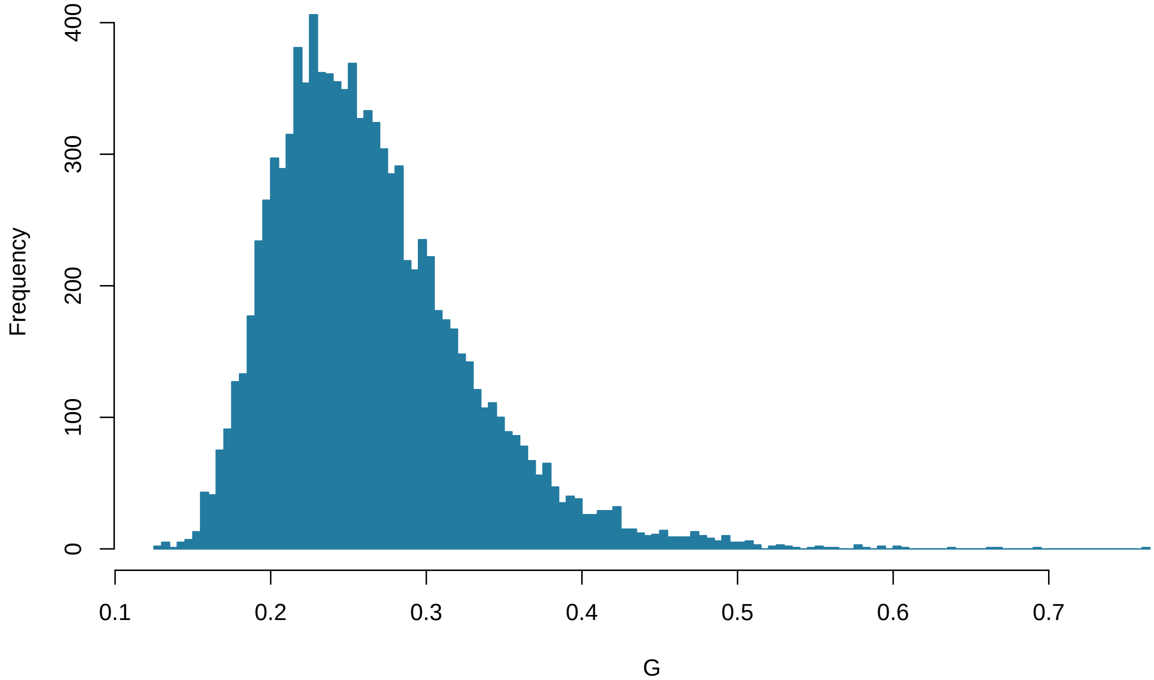
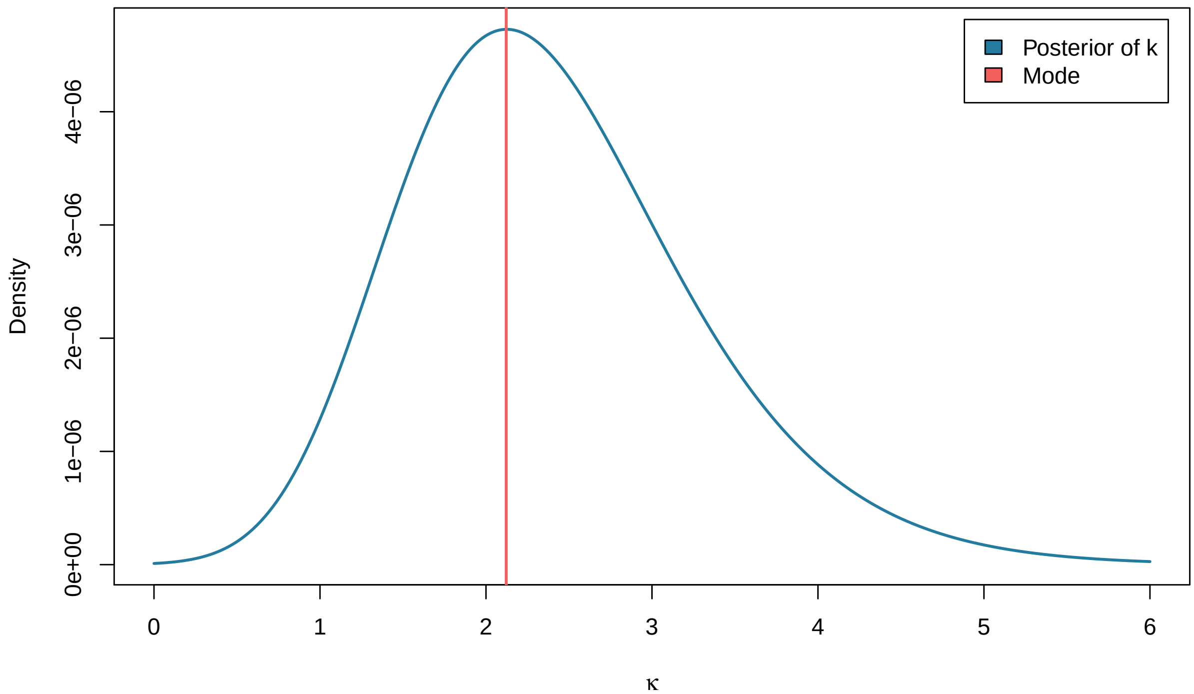


Figure 5: Gini coefficient

3. Von Mises distribution

The posterior distribution is calculated by multiplying the individual observations together with the prior for a fine grid of κ values. The mode is calculated for the posterior. The plot can be found in Figure 6.



Appendix A - Code for assignment 1

```
a = c(1,23)
```

Appendix B - Code for assignment 2

```r
set.seed(12345)
col1 = '#247ba0'
col2 = '#f25f5c'
col3 = '#333333'

a = 2
b = 2
s = 14
n = 20

xGrid <- seq(0.001, 0.999, 0.001)

## 1a
draw = function(nDraws){
  sample = rbeta(nDraws, a+s, b+(n-s))
  mean_sample = mean(sample)
  sd_sample = sd(sample)

  return (c(mean_sample, sd_sample))
}

alpha = a + s
beta = b + (n-s)

intervals = seq(20,50000,100)
data = sapply(intervals, draw)

true = dbeta(xGrid, a+s, b+(n-s))
mean_true = alpha / (alpha + beta)
sd_true = sqrt( (alpha*beta) / ((alpha+beta)^2*(alpha+beta+1)) )

# Mean
pdf("converge_mean.pdf")
  plot(intervals, abs(data[1,]-mean_true), type='l', col=col3, xlab = 'nDraws', ylab='Absolute error')
dev.off()

# SD
pdf("converge_sd.pdf")
  plot(intervals, abs(data[2,]-sd_true), type='l', col=col3, xlab = 'nDraws', ylab='Absolute error')
dev.off()

## 1b
nDraws = 10000
posterior = rbeta(nDraws, a+s, b+(n-s))
true = pbeta(0.4, a+s, b+(n-s))

message((sum(posterior <= 0.4) / length(posterior)) * 100)
message(round(true*100, 3))

## 1c
nDraws = 10000
```

```r
posterior = rbeta(nDraws, a+s, b+(n-s))
phi = log(posterior / (1 - posterior))

pdf("phi.pdf")
  hist(phi, 100, prob=TRUE, col=col1, border=col1, main='')
  lines(density(phi), lwd=2, col=col2)
dev.off()
```

Appendix C - Code for assignment 3

```
a = c(1,23)
print('hej')
```