732A73 Bayesian Learning

Computer lab 1

Johannes Hedström & Mikael Montén

STIMA
Department of Computer and Information Science
Linköpings universitet

2024-04-04

Contents

1 1 Daniel Bernoulli

```
s <- 22 # successes
f <- 70 - 22 # failures
n <- 70 # trials

A0 <- 8
B0 <- 8
```

1.1 a

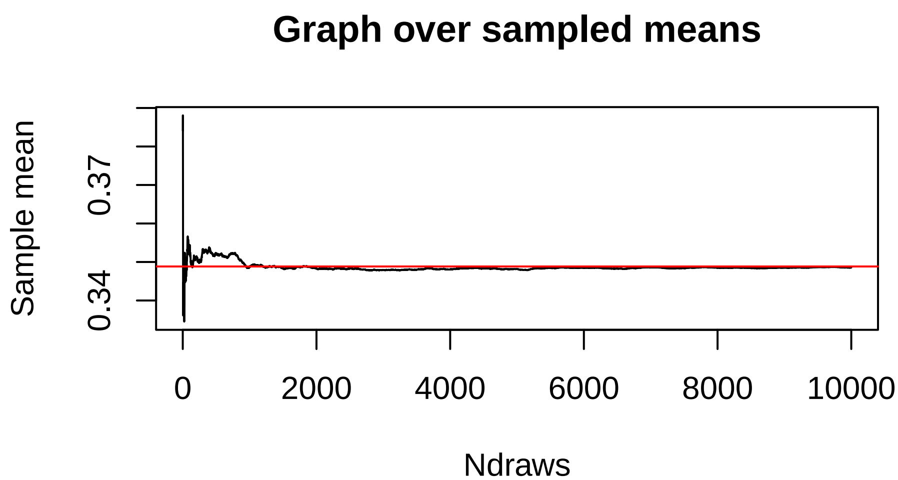
$$\theta|y \sim Beta(\alpha_0 + s, \beta_0 + f)$$

$$\text{True mean} = \frac{\alpha_0 + s}{\alpha_0 + \beta_0 + n}$$

$$\text{True sd} = \sqrt{\frac{(\alpha_0 + s) \cdot (\beta_0 + f)}{(\alpha_0 + \beta_0 + n)^2 (\alpha_0 + \beta_0 + n + 1)}}$$

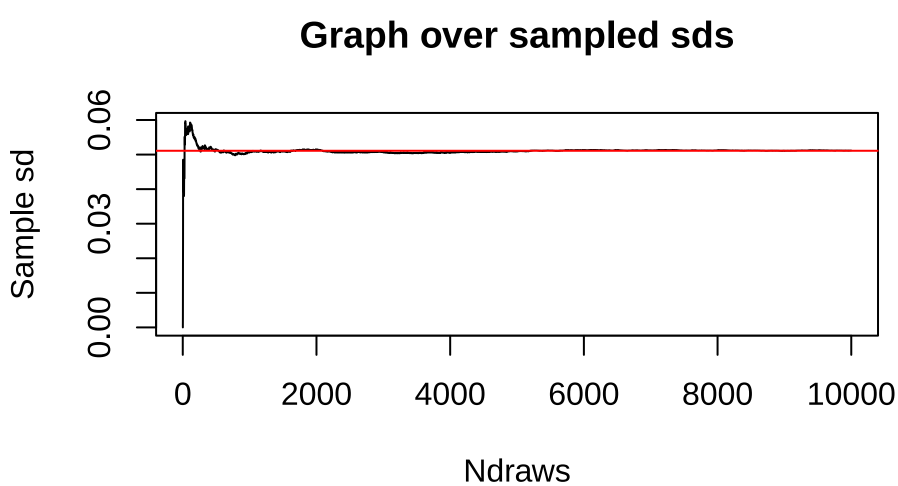
```
nDraws <- 10000 # nr of draws

true_mean <- (A0 +s) / (A0 + B0 + n) # calculating true mean and sd for posterior dist
true_sd  <- sqrt(((A0+s)*(B0+f))/((A0+B0+n)^2 * (A0+ B0+n+1)))

posterior_sample <- rbeta(nDraws,A0+s,B0+f) # draw values from posterior

# Calculate sample cumulative means and standard deviations to show convergence
sample_means <- cumsum(posterior_sample)/(1:nDraws)
sample_sds <- sqrt(cumsum((posterior_sample - sample_means)^2) / (1:nDraws))
```

```
plot(sample_means,type='line',xlab='Ndraws',ylab='Sample mean', main='Graph over sampled means')
abline(h=true_mean,col='red')
```



We can see that the sample mean of θ is getting closer to the true mean posterior mean of θ after about 1000 draws but doesn't converge until after 2000 draws.

```
plot(sample_sds,type='line',xlab='Ndraws',ylab='Sample sd', main='Graph over sampled sds')
abline(h=true_sd,col='red')
```



The sampled standard deviation gets closer to the true sd after about 1000 draws but doesn't fully converge to the true sd before 5000 draws.

1.2 b

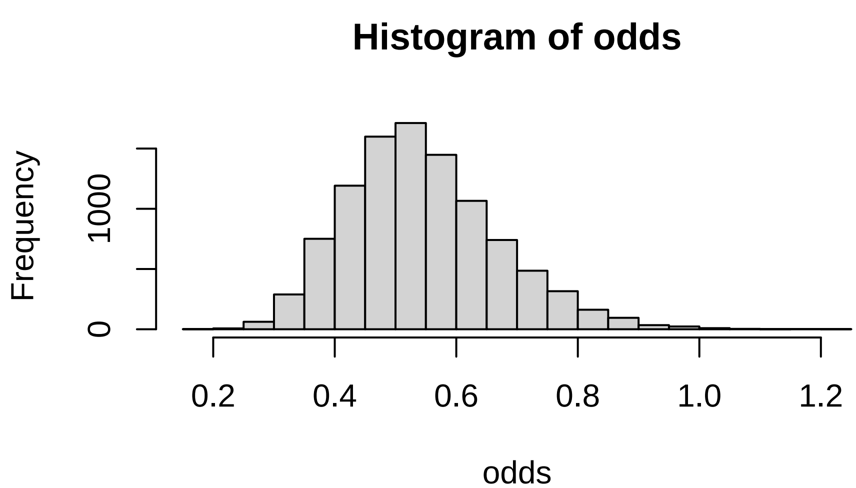
```r
# draws from posterior
post <- rbeta(nDraws, A0+s, B0+f)

# mean of samples over 0.3
prob <- mean(post>0.3)


prob_exact <- 1 - pbeta(0.3,A0+s, B0+f)

df <- data.frame('Posterior prob' =prob, 'Exact value from beta post' = prob_exact)
colnames(df) <- c('Posterior prob','Exact value from beta post')

knitr::kable(df)
```

Posterior prob	Exact value from beta post
0.8251	0.8285936

The approximate posterior probability of $\theta > 0.3$ given y is 0.8207 which is close to the exact value from the beta posterior 0.83(rounded).

1.3 c

```r
odds <- post / (1- post)

hist(odds)
```



1.4 kommentar ska in här

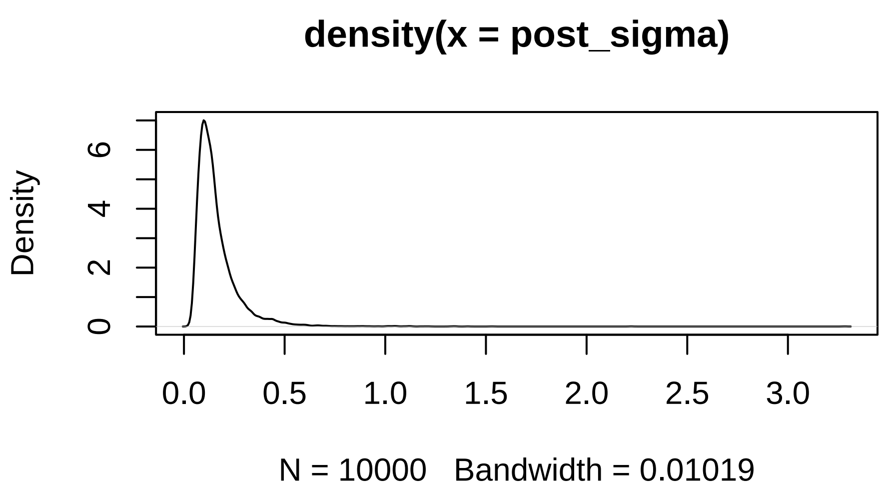
2 2 Log-normal distribution and the Gini coefficient.

```
income <- c(33,24,48,32,55,74,23,17)

tau2 <- function(y,n,mu){

  sum((log(y)-mu)^2)/n
}
```

2.1 a

```
library(invgamma)
n <-8
mu <- 3.6
tau <- tau2(income,n,mu)
post_sigma <- rinvchisq(10000,n,tau)
```

```
plot(density(post_sigma))
```

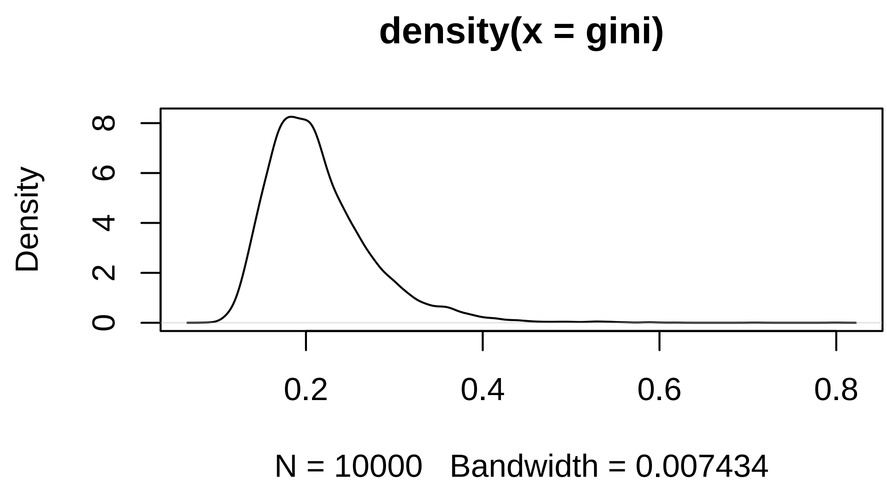


Most of our values for our σ^2 posterior are under 0.5.

2.2 b

```
gini <- 2*pnorm(mean=0,sd=1,sqrt(post_sigma)/sqrt(2)) -1
```

```
plot(density(gini))
```



2.3 c

```
eti <- quantile(gini,c(0.025,0.975))

print(eti)
```

```
##      2.5%     97.5%
## 0.1325730 0.3644386
```

The probability of the Gini coefficient being in this interval is 95 %.

2.4 d

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

3.1 a

3.2 b