Homework 9

Ben Buzzee

November 22, 2020

1a

Sample from a dgamma(3,5) prior

```
my.data <- list()
my.fname <- "model_a.txt"
my.jags.model <- jags.model(
  file = my.fname, data = my.data,
  n.chains = 1, n.adapt = 1000, quiet = FALSE )
```

```
## Compiling model graph
##    Resolving undeclared variables
##    Allocating nodes
## Graph information:
##    Observed stochastic nodes: 0
##    Unobserved stochastic nodes: 1
##    Total graph size: 3
##
## Initializing model
```

```
my.variables <- c("theta")
my.coda.samples <- coda.samples(
  my.jags.model,my.variables,10000,thin=1)
#plot(my.coda.samples)
```

After sampling from the prior using the code above, we see that the estimated mean and standard deviation for theta are .598 and .346 respectively. We also know alpha = 3 and beta = 5. So it is clear the mean is $\frac{\alpha}{\beta}$ and variance is $\frac{\alpha}{\beta^2}$

1b

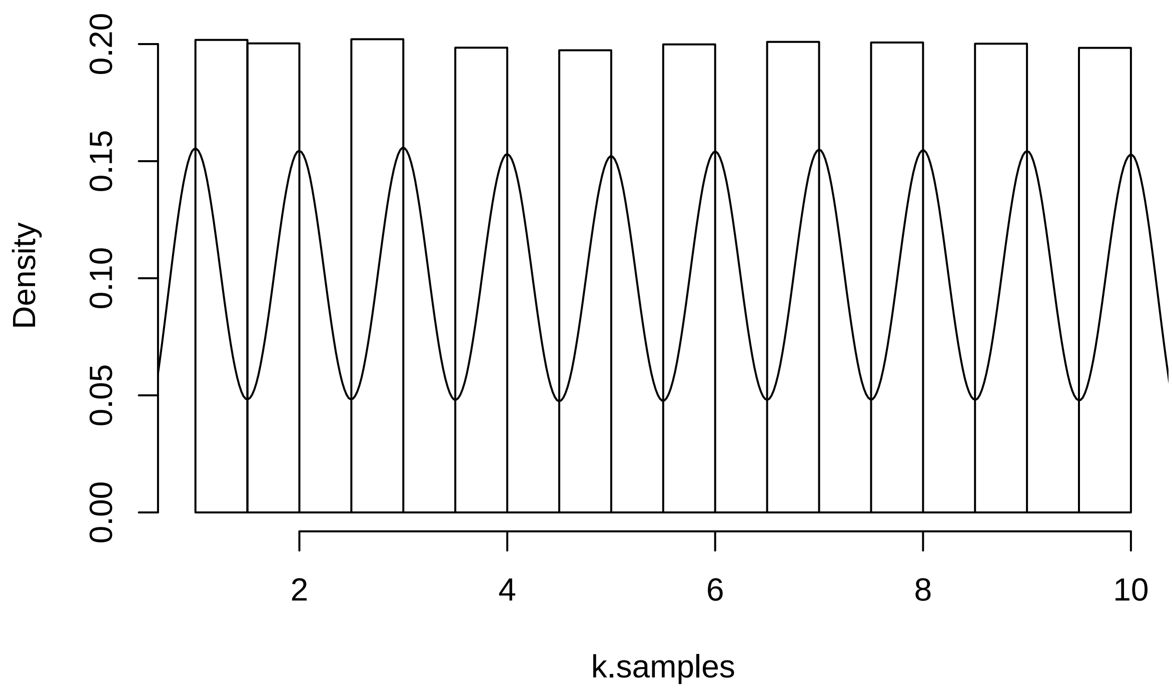
1/theta

Here we again know alpha = 3 and beta = 5. After sampling we find that we have a mean for theta of 2.5 and standard deviation of 2.6 (variance = 6.76). This is roughly what we would expect if the mean were $\frac{\beta}{\alpha-1}$ and the variance were $\frac{\beta^2}{(\alpha-1)^2(\alpha-1)}$.

1c

dcat[]

```r
hist(k.samples, prob=TRUE, main="")
lines(density(k.samples))
```



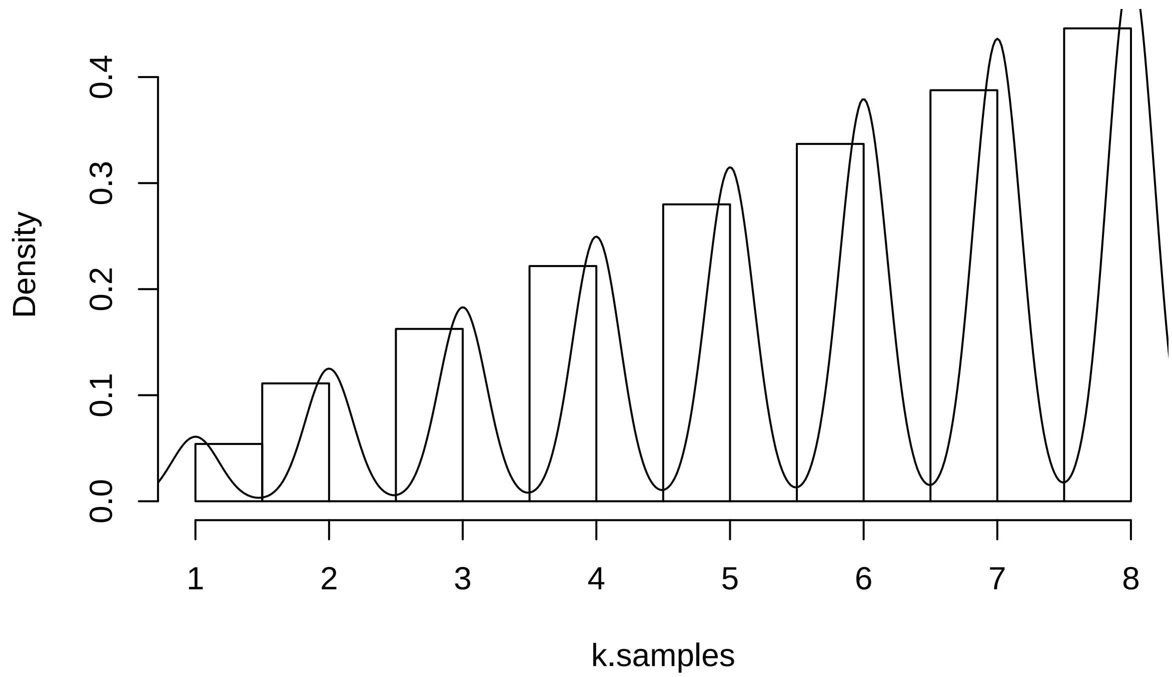

```r
table(k.samples)
```

```
## k.samples
##     1     2     3     4     5     6     7     8     9    10
## 10089 10015 10104  9923  9868  9993 10047 10034 10008  9919
```

K appears to follow a multinomial distribution with 10 categories, each category having a .1 probability of "success." The values it can take on are 1-10. Out of 100,000 samples, approximately 10,000 occurrences of each level is simulated.

1c (iii)

```r
hist(k.samples, prob=TRUE, main="")
lines(density(k.samples))
```




```r
table(k.samples)
```

```
## k.samples
##     1     2     3     4     5     6     7     8
##  2702  5557  8126 11090 13998 16849 19381 22297
```

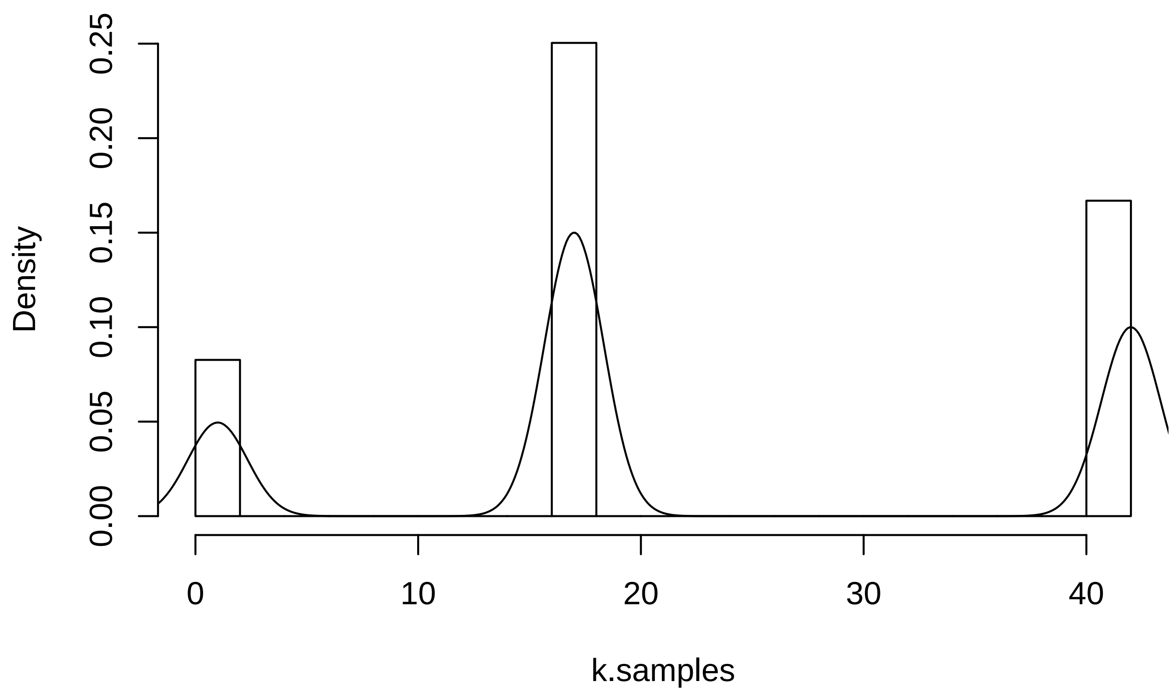
Here the distribution is still multinomial, except with 8 categories instead of 10. The values k can take on are 1-8. The probability of an outcome occurring in each category is 1/36, 2/36 … 8/36 for each of the respective categories.

1d

```r
unique( k.samples )
```

```
## [1]  1 42 17
```

```r
hist(k.samples, prob=TRUE, main="")
lines(density(k.samples))
```



Here we are again dealing with a multinomial distribution. K can take on the values 1, 17, and 42. The probability of each value occuring is 1/6, 3/6 and 2/6 respectively.

2a

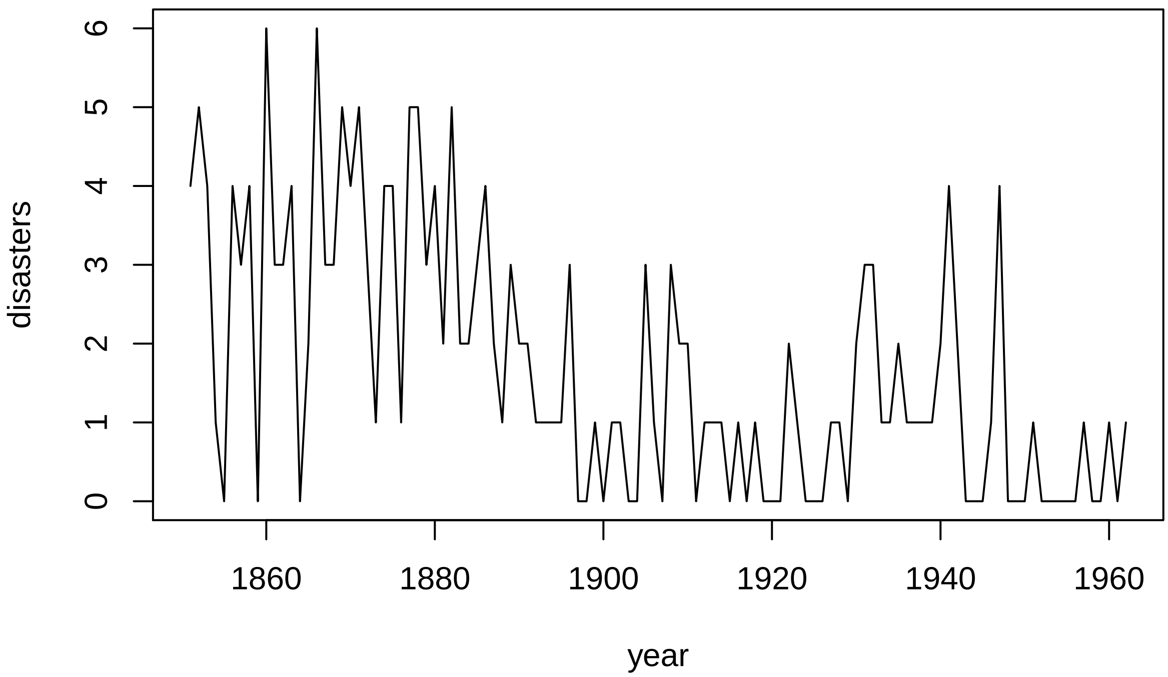
```r
my.data <- read.table("data.txt",header=TRUE)
names(my.data)
```

```
## [1] "year"      "disasters"
```

```r
disasters <- my.data$disasters
year <- my.data$year
N <- length(disasters); N
```

```
## [1] 112
```

```r
plot(year,disasters, type='l')
```



I would say there does appear to be a change in the rate of disasters. From 1851 to the 1890s there appears to be roughly 2-3 disasters a year on average. Then after 1900 it seems that average declines to about 1 disaster a year, with many years having zero disasters.

2b

For k=40, it is implied a change in the rate of disaster occurred starting in year 41. Since our time series starts at 1851, this would indicate a change in the disaster rate in 1892.

2c

The posterior distribution would be Poisson($\lambda_1 + 1$) for i = 1, 2, ... 40 and Poisson($\lambda_2 + 1$) for i = 41, 42, ... 112.

2d

idx[i] <- 1+step(i-k-0.5)

whenever the step function is negative, it will return a zero, resulting in idx[i] being equal to one. For k = 40, it will be zero until i = 41. Then from 41 forward it will start to return a 1, resulting in idx[i] = 2. It should always result in either a zero or a one.

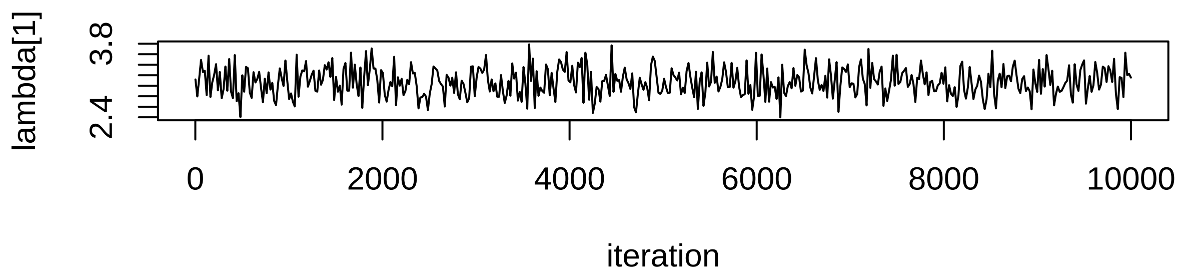
2e

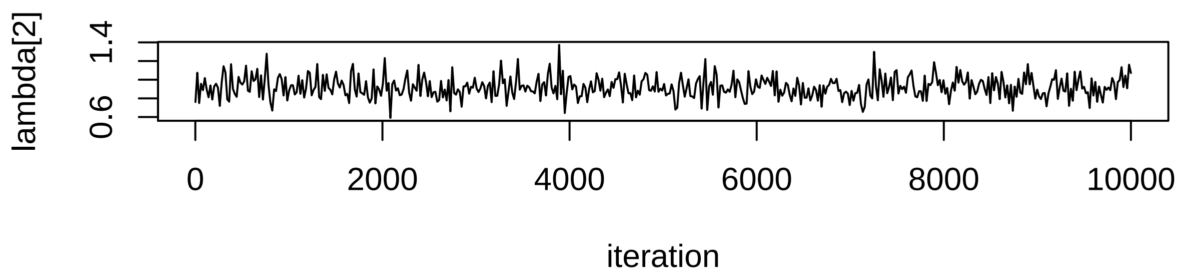
Fit the model

2f

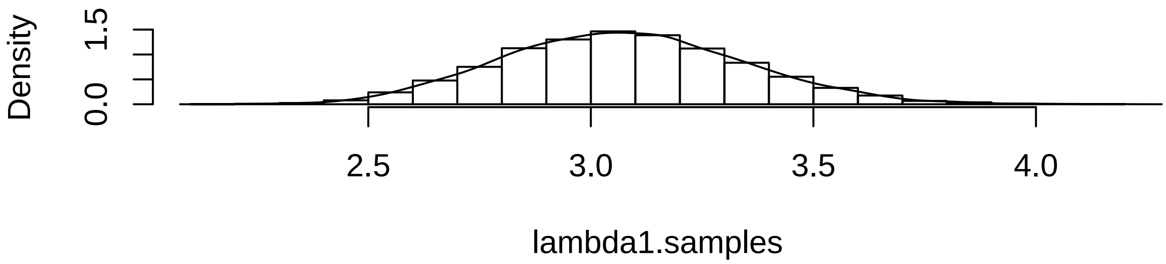
Results

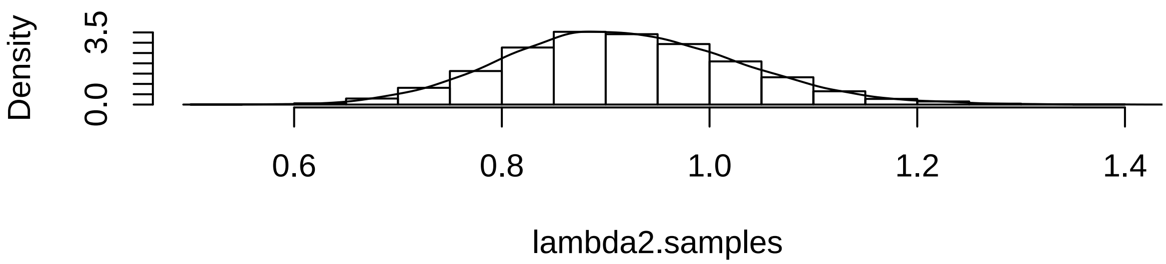
```
par(mfrow = c(2,1))
plot( which.ones, some.lam1.samples, type='l',
      xlab="iteration", ylab="lambda[1]")
plot( which.ones, some.lam2.samples, type='l',
      xlab="iteration", ylab="lambda[2]")
```






```
hist(lambda1.samples, prob=TRUE, main="")
lines(density(lambda1.samples))
hist(lambda2.samples, prob=TRUE, main="")
lines(density(lambda2.samples))
```





The effective sample size for each parameter is approximately 30,000. The rejection rate is zero, which implies a gibbs sampler is being used with a 100% success rate.