Computational Statistics Homework 3

Liam Dillingham

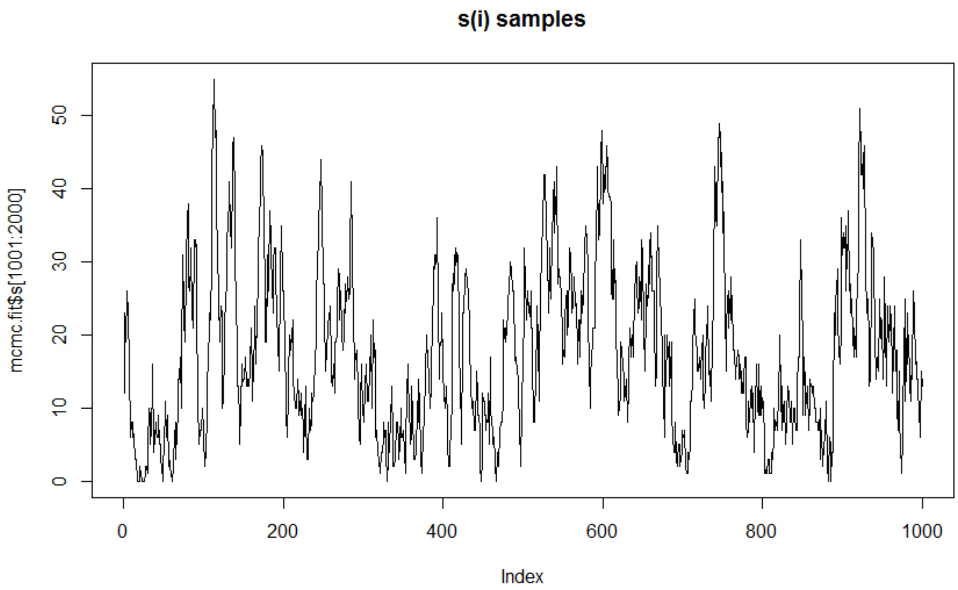
March 11, 2019

1 Question 1

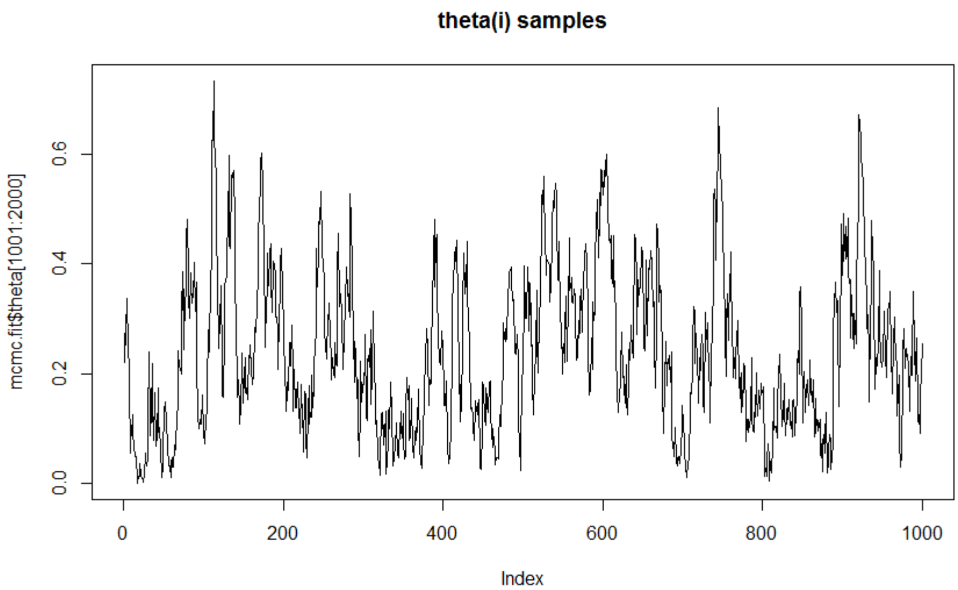
Implement the Gibbs sampler from the assignment sheet for generating bivariate samples from the join density (s, θ). ___________

```
n <- 74
s_init <- 16
theta_init <- s_init/n
burn = 1000 # burn-in
nmc = 1000 # number of Monte Carlo iterations

gibbs <- function(s_init, theta_init, burn = 1000, nmc = 1000,
        alpha_0 = 2.0, beta_0 = 6.4){

  theta <- rep(0, nmc+burn)
  s <- rep(0, nmc+burn)

  s[1] = s_init; theta[1] = theta_init;

  for (i in 2:(burn+nmc)) {

    s[i] <- rbinom(1, n, theta[i-1])
    theta[i] <- rbeta(1, alpha_0 + s[i], beta_0 + n-s[i])
  }
  return(list(s=s,theta=theta))
}

mcmc.fit <- gibbs(s_init, theta_init)

plot(mcmc.fit$theta[1001:2000], type = "l", main="theta(i) samples")
plot(mcmc.fit$s[1001:2000], type = "l", main="s(i) samples")
hist(mcmc.fit$s[1001:2000], breaks = 30, main="s(i) samples histogram", freq = F)
acf(mcmc.fit$theta[1001:2000])
mean(mcmc.fit$theta[1001:2000])
median(mcmc.fit$theta[1001:2000])
```

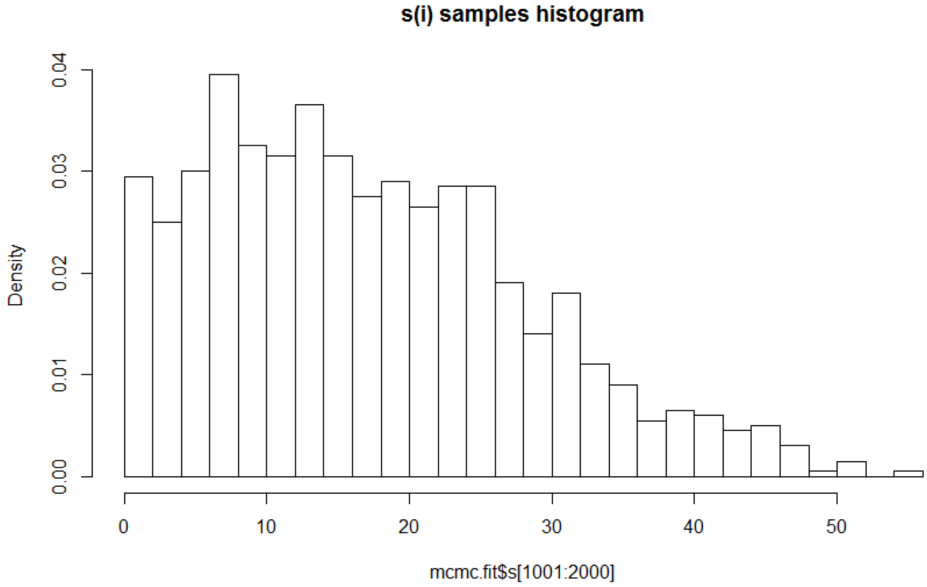
1.1 Draw the trace plot for the $s^{(i)}$ samples



1.2 Draw the traceplot for the $\theta^{(i)}$ samples



1.3 Draw the histogram for the $s^{(i)}$ samples



1.4 Estimate the posterior median of θ based on the samples drawn. Is it close to the maximum likelihood estimate of s/n?

From the problem description, it states that the maximum likelihood estimate of theta, $\hat{\theta}_{mle}$, is equal to $s/n = 74/16 = 0.2162162$. By calculating the post burn-in simulation of our Gibbs sampler, we achieve a median value of 0.2198654

1.5 How sensitive is the posterior median to the choice of initial values?

It doesn't seem to be terribly sensitive, since it is still able to converge to the median estimate s/n. the actual value of the median shifts by 0.1 0.2 when n is changed by a few integers

2 Question 2

Implement the Gibbs sampler from before but treat n as an unknown paramter as well with a Poisson prior on n, i.e. $\pi(n) = \text{Poisson}(\lambda)$. Assume $(\lambda = 64)$. ___________

```
lambda = 64
n_init = 74
s_init <- 16
theta_init <- s_init/n_init
burn = 1000 # burn-in
nmc = 1000 # number of Monte Carlo iterations

gibbs <- function(s_init, theta_init, n_init, burn = 1000, nmc = 1000,
        alpha_0 = 2.0, beta_0 = 6.4){

  n = rep(0, nmc+burn)
  theta <- rep(0, nmc+burn)
  s <- rep(0, nmc+burn)

  s[1] = s_init; theta[1] = theta_init; n[1] = n_init;

  for (i in 2:(burn+nmc)) {
    n[i] = rpois(1, lambda * (1-theta[i-1]))
    s[i] <- rbinom(1, n[i], theta[i-1])
    theta[i] <- rbeta(1, alpha_0 + s[i], beta_0 + n[i]-s[i])
  }
  return(list(s=s,theta=theta,n=n))
}

mcmc.fit <- gibbs(s_init, theta_init, n_init)

plot(mcmc.fit$theta[1001:2000], type = "l")
plot(mcmc.fit$s[1001:2000], type = "l")
plot(mcmc.fit$n[1001:2000], type = "l")
acf(mcmc.fit$theta[1001:2000])
mean(mcmc.fit$theta[1001:2000])
median(mcmc.fit$theta[1001:2000])
```

2.1 Derive the three full conditional densities needed for the Gibbs sampler. Show full calculations

Note:

$$f(n,\theta,s) = (\frac{\lambda^n e^{-\lambda}}{n!})(\frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)}\theta^{\alpha-1}(1-\theta)^{\beta-1})(\binom{n}{s}\theta^s(1-\theta)^{n-s})$$

Due to a property of the Gibbs sampler, when a certain prior is given, we can elminate terms that don't include the posterior variable we are solving for (proportionality property).

$$f(n|\theta,s) \propto (e^{-\lambda}\frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)}\theta^{\alpha-1}(1-\theta)^{\beta-1}\theta^s(1-\theta)^{-s})\binom{n}{s}\frac{\lambda^n(1-\theta)^n}{n!}$$

$$f(n|\theta, s) \propto \binom{n}{s} \frac{\lambda^n (1-\theta)^n}{n!}$$
$$f(n|\theta, s) \propto \frac{n!}{s!(n-s)!} \frac{\lambda^n (1-\theta)^n}{n!}$$
$$f(n|\theta, s) \propto \frac{\lambda^n (1-\theta)^n}{(n-s)!}$$

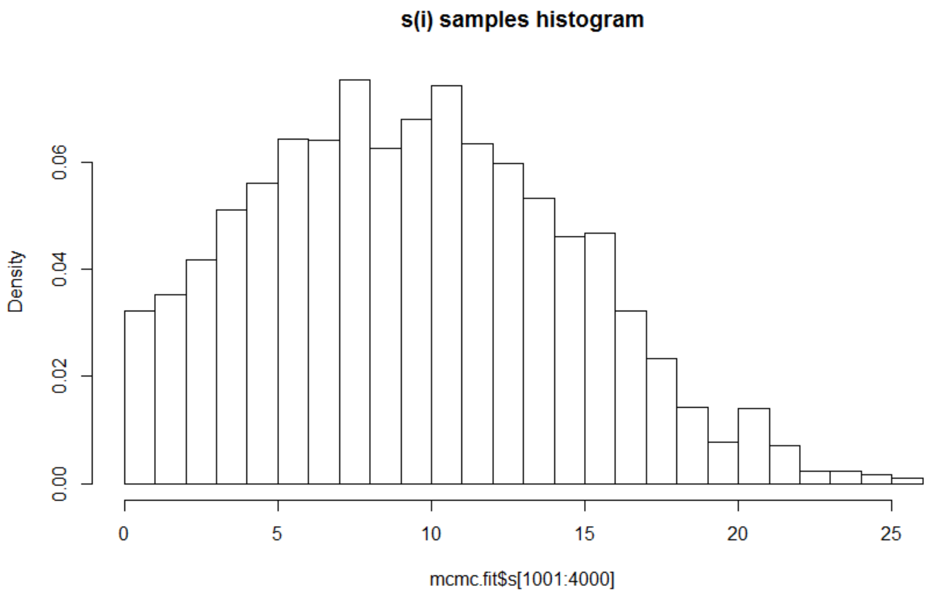
Next,

$$f(\theta|n, s) \propto (\binom{n}{s} \frac{\lambda^n e^{-\lambda}}{n!} \frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)})\theta^{\alpha-1}(1-\theta)^{\beta-1}\theta^s(1-\theta)^{n-s}$$
$$f(\theta|n, s) \propto \theta^{\alpha-1}(1-\theta)^{\beta-1}\theta^s(1-\theta)^{n-s}$$
$$f(\theta|n, s) \propto \theta^{\alpha+s-1}(1-\theta)^{\beta+n-s-1}$$

And finally,

$$f(s|\theta, n) \propto (\frac{\lambda^n e^{-\lambda}}{n!} \frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)}\theta^{\alpha-1}(1-\theta)^{\beta-1})\binom{n}{s}\theta^s(1-\theta)^{n-s}$$
$$f(s|\theta, n) \propto \binom{n}{s}\theta^s(1-\theta)^{n-s}$$

2.2 Draw the histogram of $s^{(i)}$ samples and compare your new posterior median with the one you obtained in problem 1



2.3 Do you get similar convergence results as before?

I had to run it for a few thousand more iterations that the first gibbs sampler, but the results tend to stick in the 0.2 to 0.23 range, so yes.