Homework 8

Bingying Liu

April 15, 2020

Question 1: Biased coin problem:

Suppose we have univariate data $y_1, \dots, y_n|\theta \sim Bernoulli(\theta)$ and wish to test: $H_0 : \theta = 0.5$ vs. $H_1 : \theta \neq 0.5$.

Part (a): Formulate this hypothesis testing problem in a Bayesian way. Specify all the necessary steps and come up with your own priors where necessary.

Steps:
1. Put a prior on actual hypotheses, that is on $\pi(H_0) = Pr(H_0)$ and $\pi(H_1) = Pr(H_1)$. In this case, since we have no prior information, we set $Pr(H_0) = Pr(H_1) = 0.5$.
2. Put a prior on the parameter in each model. In this case, we set an uninformative prior $\pi(\theta) = Beta(1,1) = 1$.
3. Compute marginal posterior probabilities for each hypothesis: $Pr(H_0|Y)$ and $Pr(H_1|Y)$.
4. Conclude based on the magnitude of $Pr(H_1|Y) relative to $Pr(H_0|Y)$.

Part (b): Derive and simplify the marginal likelihoods $L[Y|H_0]$ and $L[Y|H_1]$.

$$
\begin{aligned}
L[Y|H_0] &= \int_{\theta=0.5} p(Y,\theta|H_0)d\theta \\
&= \int_{\theta=0.5} L(Y|H_0,\theta)\pi(\theta|H_0)d\theta \\
&= 0.5^{\sum y_i} 0.5^{n-\sum y_i} \\
&= 0.5^n
\end{aligned}
$$

$$
\begin{aligned}
L[Y|H_1] &= \int_{\theta=0}^{1} p(Y,\theta|H_1)d\theta \\
&= \int_{\theta=0}^{1} L(Y|H_1,\theta)\pi(\theta|H_1)d\theta \\
&= \int_{\theta=0}^{1} \theta^{\sum y_i}(1-\theta)^{n-\sum y_i} Beta(1,1)d\theta \\
&= B(\sum y_i + 1, n - \sum y_i + 1)
\end{aligned}
$$

Part (c): Derive the Bayes factor in favor of H_1. Also, derive the posterior probability of H_1 being true. Simplify both as much as possible.

$$
BF_{10} = \frac{L[Y|H_1]}{L[Y|H_0]} = \frac{B(\sum y_i + 1, n - \sum y_i + 1)}{0.5^n}
$$

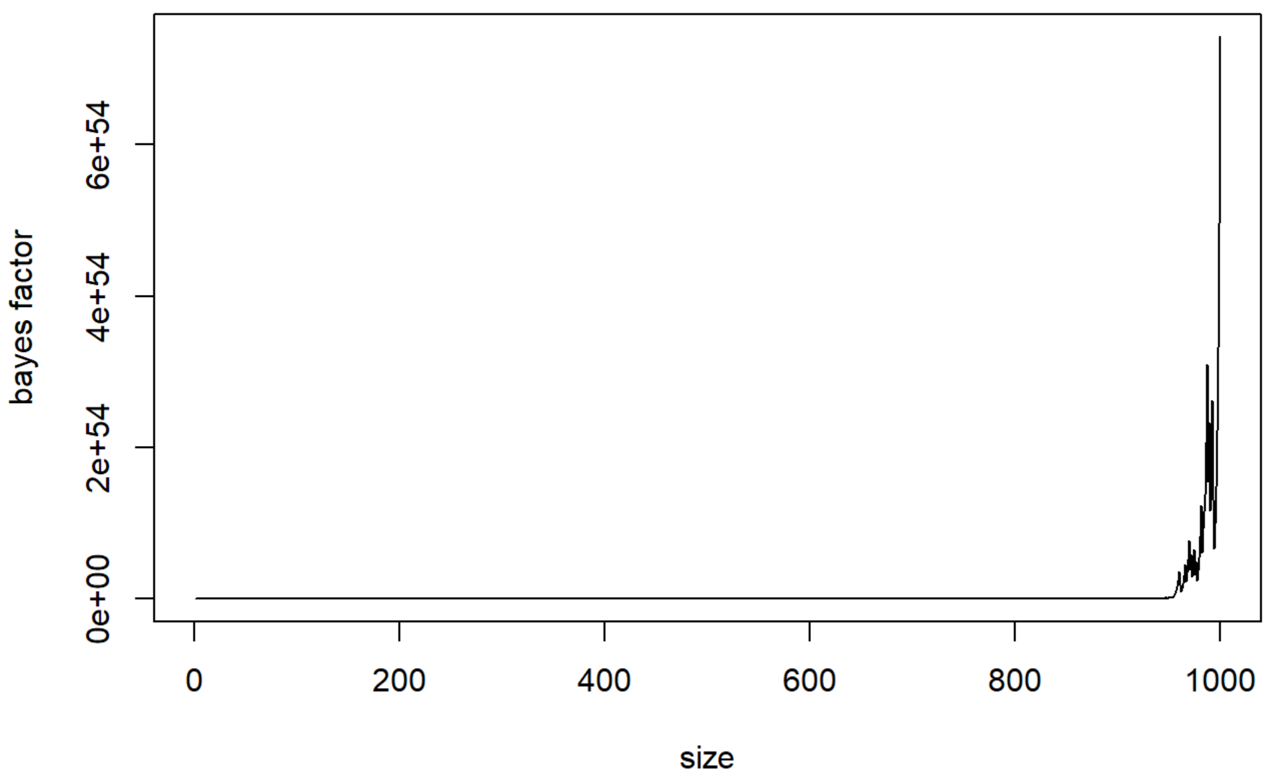
$$
Pr[H_1|Y] = \frac{1}{BF_{01}+1} = \frac{1}{\frac{1}{BF_{10}}+1} = \frac{1}{\frac{0.5^n}{B(\sum y_i+1,n-\sum y_i+1)}+1} = \frac{B(\sum y_i + 1, n - \sum y_i + 1)}{B(\sum y_i + 1, n - \sum y_i + 1) + 0.5^n}
$$

Part (d): Study the asymptotic behavior of the Bayes factor in favor of H_1. For $\theta \in \{0.25, 0.46, 0.5, 0.54\}$, make a plot of the Bayes factor (y-axis) against sample size (x-axis). You should have four plots. Comment (in detail) on the implications of the true value of θ on the behavior of the Bayes factor in favor of H_1, as a function of sample size. When answering this question, remind yourself of what Bayes factors actually mean and represent! One line answers will not be sufficient here; explain clearly and in detail what you think the plots mean or represent.

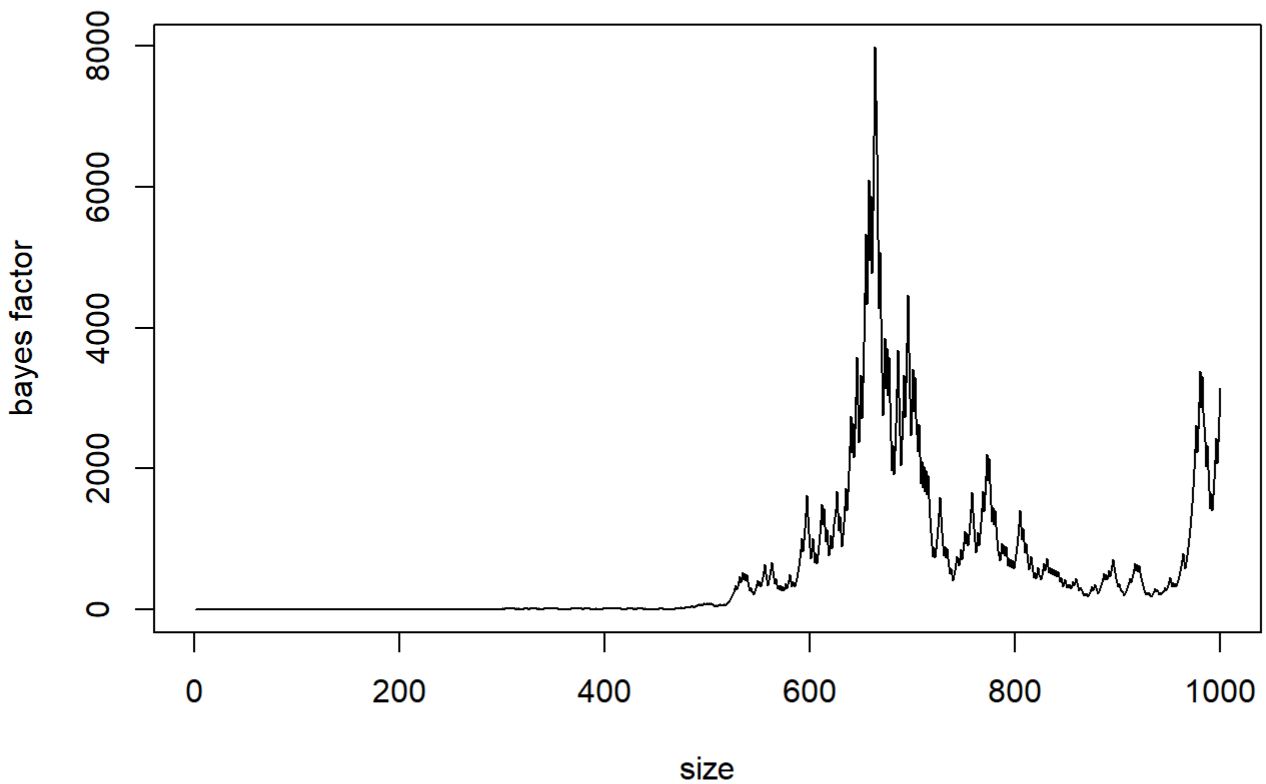
```
theta_list = c(0.25,0.46,0.5, 0.54)
a = b = 1


for (theta in theta_list){
  n = 1000
  y = c()
  bf10_vec = c()
  for(i in 1:n){
    y_new = rbinom(1,1,theta)
    y = c(y, y_new)
    Y = sum(y)
    bf10 = beta(a+Y, b+i-Y)/(0.5^i)
    bf10_vec = c(bf10_vec, bf10)
  }
  plot(x=1:n, y=bf10_vec, main=paste("theta = ",theta), ylab = "bayes factor", xlab = "size",typ
e="l")
}
```

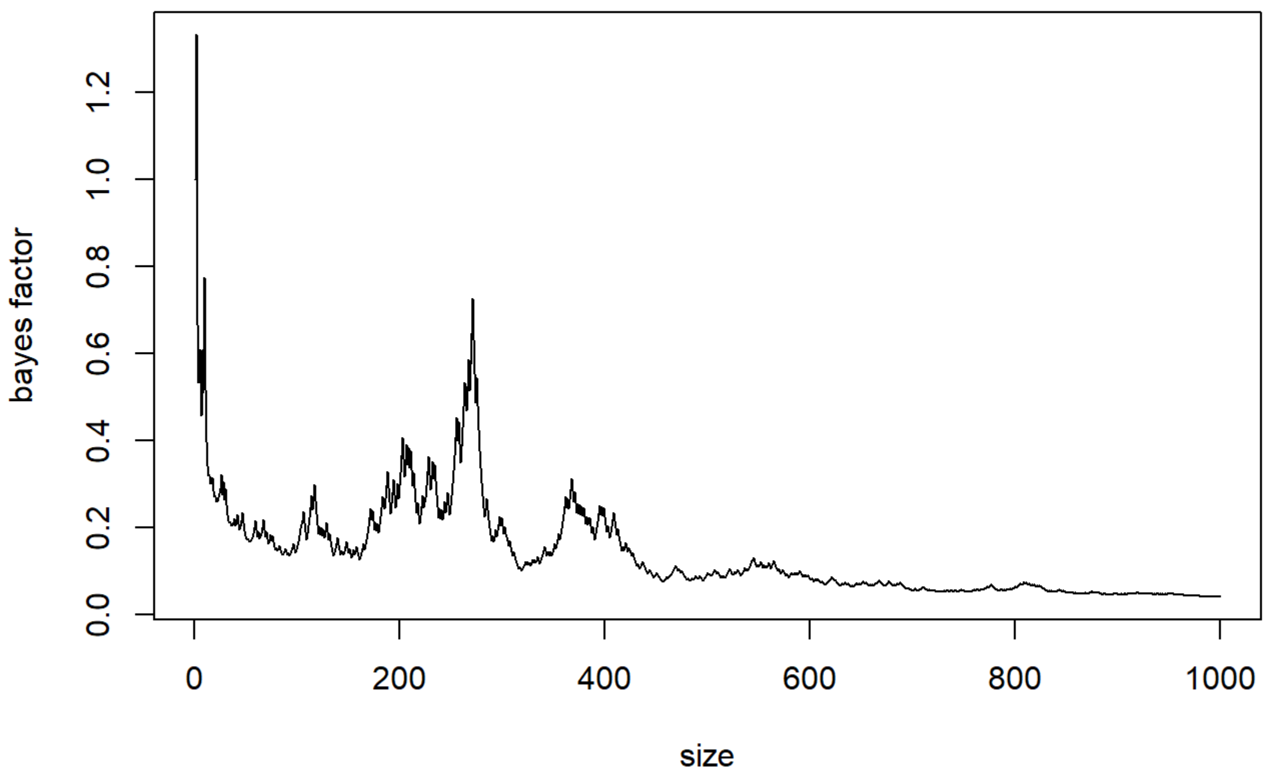
theta = 0.25



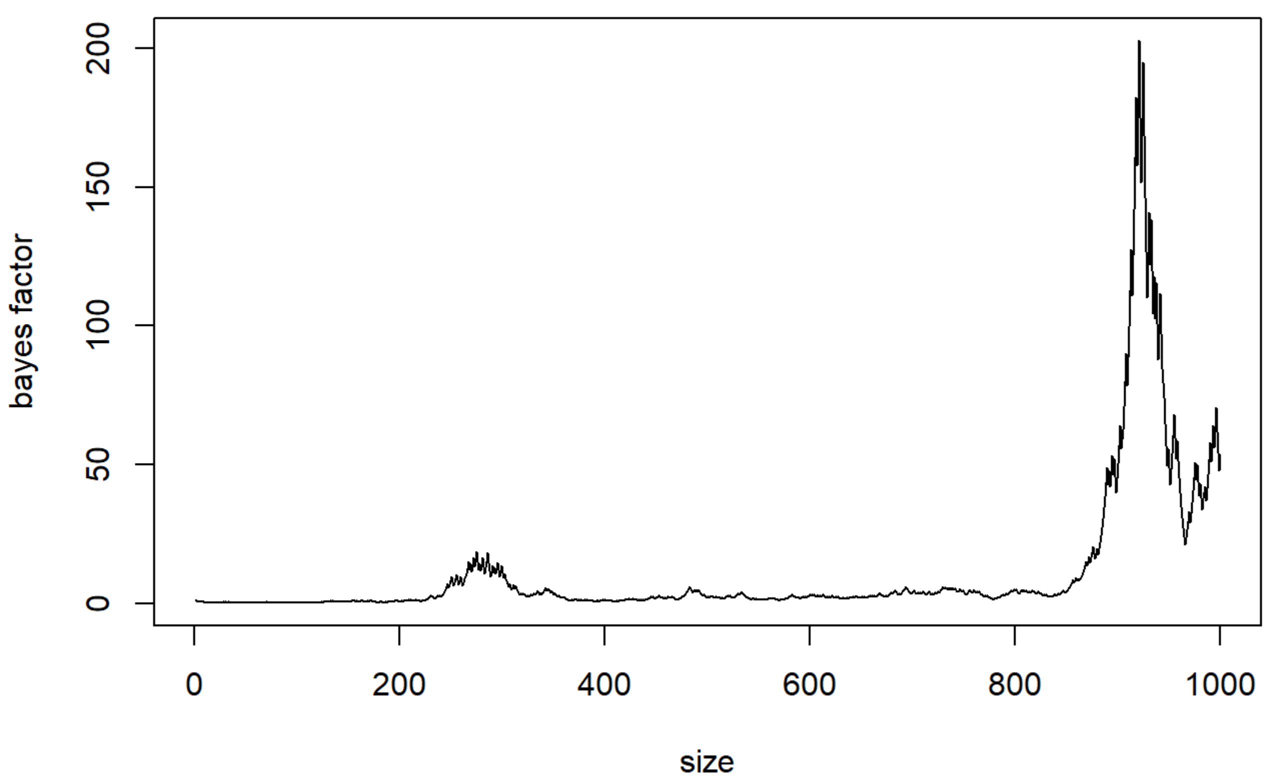
theta = 0.46



theta = 0.5



theta = 0.54



In the first plot when true $\theta = 0.25$, as sample size becomes larger, especially towards 800 - 1000, we can see that the bayes factor BF_{10} increases significantly. This shows that when sample size is large, $Pr[H_0|Y]$ becomes small enough that the null hypothesis should be rejected.

Whereas in the third plot when true $\theta = 0.5$, as sample size becomes larger, the bayes factor BF_{10} stabilizes around 0, which shows decisive evidence that $Pr[H_0|Y]$ is close enough to 1 and we should not reject the null.

For second and fourth plots, since the true θs are close to but not equal to 0.5, the plots for bayes factors don't have a fixed pattern (however, the range of bayes factor is narrower than that of plot 1), suggesting that data can either support H_0 or H_1.

Question 2: Metropolis-Hastings

Part (a): Full Conditionals:

$$g_{\theta_1}[\theta_1^*|\theta_1^{(s)}, \theta_2^{(s)}] = p(\theta_1^*|y_1, \dots y_n, \theta_2^{(s)});$$
$$g_{\theta_2}[\theta_2^*|\theta_1^{(s)}, \theta_2^{(s)}] = p(\theta_2^*|y_1, \dots y_n, \theta_2^{(s)}).$$

$$\begin{aligned} r &= \frac{p(\theta_1^*, \theta_2^{(s)}|y_{1:n})}{p(\theta_1^{(s)}, \theta_2^{(s)}|y_{1:n})} \frac{g_{\theta_1}[\theta_1^s|\theta_1^{(s)}, \theta_2^{(s)}]}{g_{\theta_1}[\theta_1^*|\theta_1^{(s)}, \theta_2^{(s)}]} \\ &= \frac{p(\theta_1^*|y_{1:n}, \theta_2^{(s)})p(\theta_2^*|y_{1:n})}{p(\theta_1^{(s)}|y_{1:n}, \theta_2^{(s)})p(\theta_2^*|y_{1:n})} \frac{p(\theta_1^{(s)}|y_{1:n}, \theta_2^{(s)})}{p(\theta_1^*|y_{1:n}, \theta_2^{(s)})} \\ &= 1 \end{aligned}$$

Similarly, acceptance ratio for $r = \frac{p(\theta_1^{(s)}, \theta_2^*|y_{1:n})}{p(\theta_1^{(s)}, \theta_2^{(s)}|y_{1:n})} \frac{g_{\theta_2}[\theta_2^s|\theta_1^{(s)}, \theta_2^{(s)}]}{g_{\theta_2}[\theta_2^*|\theta_1^{(s)}, \theta_2^{(s)}]} = 1$

Part (b): Priors:

$$g_{\theta_1}[\theta_1^*|\theta_1^{(s)}, \theta_2^{(s)}] = \pi_1(\theta_1^*);$$
$$g_{\theta_2}[\theta_2^*|\theta_1^{(s)}, \theta_2^{(s)}] = \pi_2(\theta_2^*).$$

$$\begin{aligned} r &= \frac{p(\theta_1^*, \theta_2^{(s)}|y_{1:n})}{p(\theta_1^{(s)}, \theta_2^{(s)}|y_{1:n})} \frac{g_{\theta_1}[\theta_1^s|\theta_1^{(s)}, \theta_2^{(s)}]}{g_{\theta_1}[\theta_1^*|\theta_1^{(s)}, \theta_2^{(s)}]} \\ &= \frac{p(y_{1:n}|\theta_1^*, \theta_2^{(s)})\pi_1(\theta_1^*)\pi_2(\theta_2^{(s)})}{p(y_{1:n}|\theta_1^{(s)}, \theta_2^{(s)})\pi_1(\theta_1^{(s)})\pi_2(\theta_2^{(s)})} \frac{\pi_1(\theta_1^{(s)})}{\pi_1(\theta_1^*)} \\ &= \frac{p(y_{1:n}|\theta_1^*, \theta_2^{(s)})}{p(y_{1:n}|\theta_1^{(s)}, \theta_2^{(s)})} \end{aligned}$$

Similarly, acceptance ratio for $r = \frac{p(\theta_1^{(s)}, \theta_2^*|y_{1:n})}{p(\theta_1^{(s)}, \theta_2^{(s)}|y_{1:n})} \frac{g_{\theta_2}[\theta_2^s|\theta_1^{(s)}, \theta_2^{(s)}]}{g_{\theta_2}[\theta_2^*|\theta_1^{(s)}, \theta_2^{(s)}]} = \frac{p(y_{1:n}|\theta_1^{(s)}, \theta_2^*)}{p(y_{1:n}|\theta_1^{(s)}, \theta_2^{(s)})}$

Part (c): Random Walk:

$$g_{\theta_1}[\theta_1^*|\theta_1^{(s)}, \theta_2^{(s)}] = N(\theta_1^{(s)}, \delta^2);$$
$$g_{\theta_2}[\theta_2^*|\theta_1^{(s)}, \theta_2^{(s)}] = N(\theta_2^{(s)}, \delta^2).$$

$$
\begin{aligned}
r &= \frac{p(\theta_1^*, \theta_2^{(s)}|y_{1:n})}{p(\theta_1^{(s)}, \theta_2^{(s)}|y_{1:n})} \times \frac{g_{\theta_1}[\theta_1^s|\theta_1^{(s)}, \theta_2^{(s)}]}{g_{\theta_1}[\theta_1^*|\theta_1^{(s)}, \theta_2^{(s)}]} \\
&= \frac{p(y_{1:n}|\theta_1^*, \theta_2^{(s)})\pi_1(\theta_1^*)\pi_2(\theta_2^{(s)})}{p(y_{1:n}|\theta_1^{(s)}, \theta_2^{(s)})\pi_1(\theta_1^{(s)})\pi_2(\theta_2^{(s)})} \times \frac{\frac{1}{\sqrt{2\pi\sigma^2}}\exp\left(-\frac{1}{2\delta^2}(\theta_1^{(s)} - \theta_1^\star)^2\right)}{\frac{1}{\sqrt{2\pi\sigma^2}}\exp\left(-\frac{1}{2\delta^2}(\theta_1^\star - \theta_1^{(s)})^2\right)} \\
&= \frac{p(y_{1:n}|\theta_1^*, \theta_2^{(s)})\pi_1(\theta_1^*)}{p(y_{1:n}|\theta_1^{(s)}, \theta_2^{(s)})\pi_1(\theta_1^{(s)})}
\end{aligned}
$$

Similarly, acceptance ratio for $r = \frac{p(\theta_1^{(s)}, \theta_2^*|y_{1:n})}{p(\theta_1^{(s)}, \theta_2^{(s)}|y_{1:n})} \frac{g_{\theta_2}[\theta_2^s|\theta_1^{(s)}, \theta_2^{(s)}]}{g_{\theta_2}[\theta_2^*|\theta_1^{(s)}, \theta_2^{(s)}]} = \frac{p(y_{1:n}|\theta_1^{(s)}, \theta_2^*)\pi_2(\theta_2^*)}{p(y_{1:n}|\theta_1^{(s)}, \theta_2^{(s)})\pi_2(\theta_2^{(s)})}$