Metropolis-Hastings

Pohsiang Peng, Jinjie Zhang

Problem 1

We will use Metropolis-Hastings to draw samples from beta dsitribution with parameters $a = 6$ and $b = 4$. The proposal distribution has the from

$$\phi_{\text{prop}} \,|\, \phi_{\text{old}} \sim \text{Beta}(c\,\phi_{\text{old}}, c\,(1-\phi_{\text{old}}))$$

First of all, we need to decide an initial value. In this project, we use a random number from uniform distribution. Given the initial value x_0, we can draw a random sample from proposal distribution. In particular

$$x_1 \sim \text{Beta}(c\,x_0, c\,(1-x_0))$$

Then we accept this sample with probability

$$A = \min\left(1, \frac{\pi(x_1)\,p(x_0|x_1)}{\pi(x_0)\,p(x_1|x_0)}\right)$$

where $\pi(x)$ is the pdf for Beta$(6, 4)$ evaluated at x, and $\pi(x_i|x_j)$ is the pdf for Beta$(c(x_j), c(1-x_j))$ evaluated at x_i. Note that if the proposal distribution is symmetric, the probability of acceptance can be reduced as

$$A = \min\left(1, \frac{\pi(x_1)}{\pi(x_0)}\right)$$

Once we have x_1, we can draw x_2 by exactly the same procedure. Repeat this many times, we get the random samples from Beta$(6, 4)$.

Code

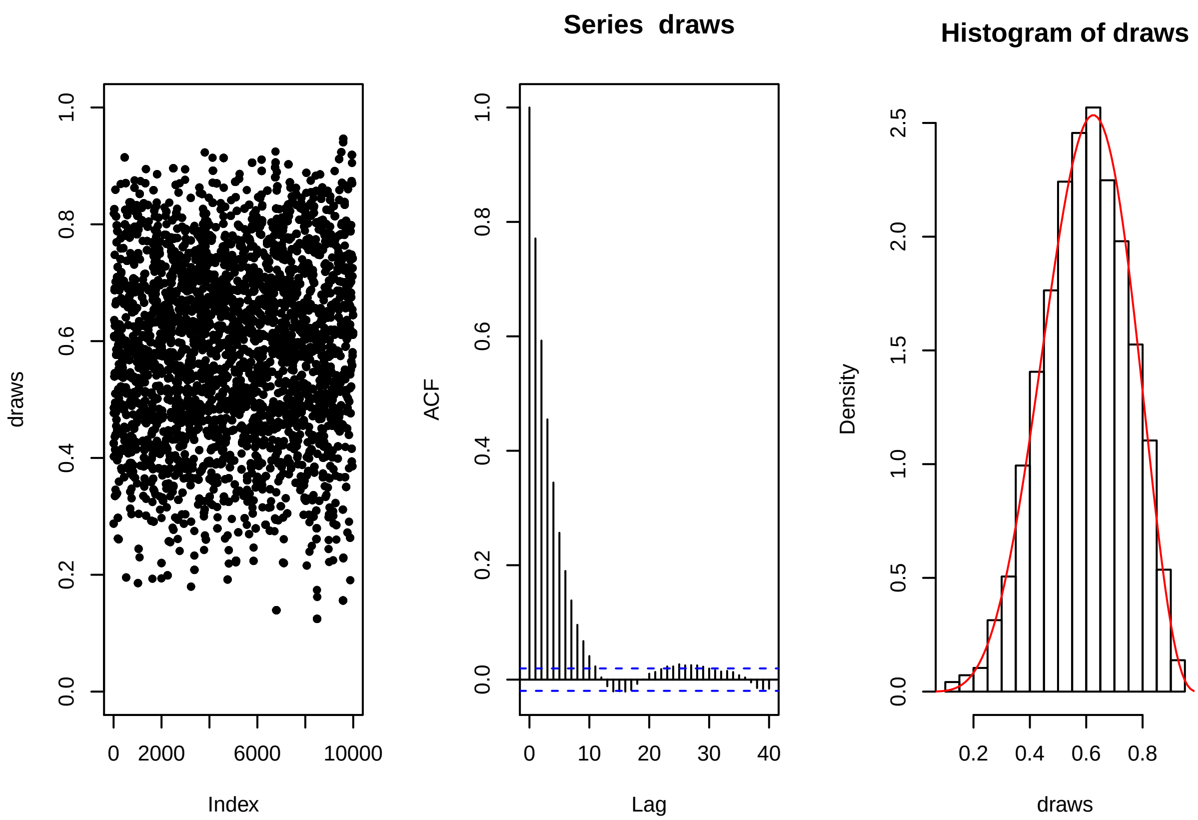
```r
# This function returns p(x/y)
prop_pdf = function(x, y, c) dbeta(x, c*y, c*(1-y))
```

```r
# @iter: iteration numbers
# @initial: initial value x0
# @c: the parameter for proposal distribution
# @alpha, beta: the parameter for beta distriubtion, which we would like to draw random samples
# @seed: random seed
# output: random samples from Beta(alpha,beta)
MH_Beta = function(iter, initial, c, alpha, beta, seed = 123) {
  set.seed(seed)

  draws = array(dim=iter)
  draws[1] = initial

  for(i in 2:iter) {
    current = draws[i-1]
    prop = rbeta(1, c*current, c*(1-current))

    A = dbeta(prop,alpha,beta)/dbeta(current,alpha,beta) *
        prop_pdf(current,prop,c)/prop_pdf(prop,current,c)
```

```
    if(runif(1) < A)
      draws[i] = prop
    else
      draws[i] = current
  }
  return(draws)
}
```

Problem 2

We first set $c = 1$ and run for 10,000 total iteration.



We can see that the histogram is in agreement with pdf for target distribution.

We then consider Kolmogorov-Smirnov statistic. Note that the Kolmogorov-Smirnov statistic for a given cumulative distribution function $F(x)$ is

$$D_n = \sup_x |F_n(x) - F(x)|$$

Thus, to compute D_n, we first compute $C_1 = \max_i |F_n(x_i) - F(x_i)|$ and $C_2 = \max_i |F_n(x_i^-) - F(x_i)|$. Then Kolmogorov-Smirnov statistic can be calculated as

$$D_n = \max(C_1, C_2)$$

Code

```
ks = function(x) {
  y = sort(x)

  cdf = pbeta(y,6,4) # F(x)
  ECDF = ecdf(y)
  F1 = ECDF(y) # C1
  F2 = ECDF(y-.Machine$double.eps) # C2

  return(max(max(abs(cdf-F1)),max(abs(cdf-F2)))) #max(C1,C2)
}
ks(draws)
```

```
## [1] 0.02026074
```

Problem 3

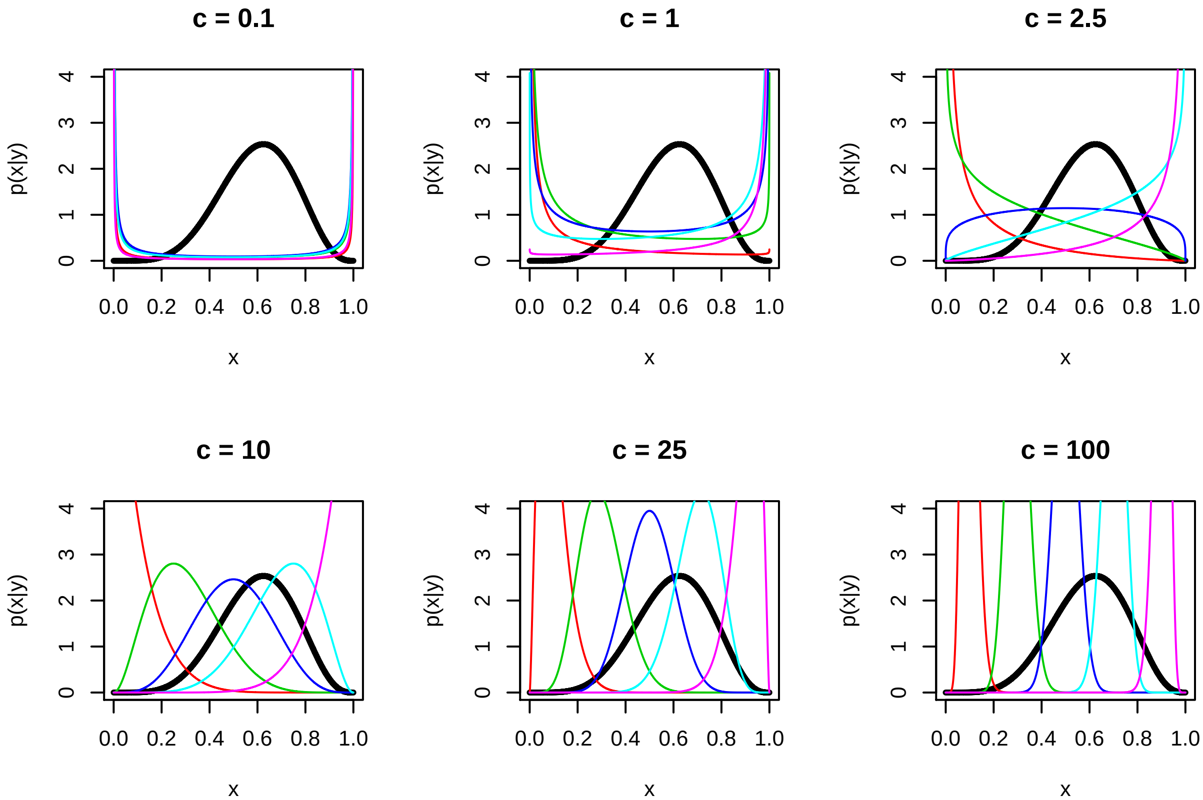
Now we will try different c to see its effect. First note that

$$\begin{aligned}
\mathrm{E}[\phi_{\text{prop}} \,|\, \phi_{\text{old}}] &= \frac{c\,\phi_{\text{old}}}{c\,\phi_{\text{old}} + c\,(1-\phi_{\text{old}})} = \phi_{\text{old}} \\
\mathrm{Var}(\phi_{\text{prop}} \,|\, \phi_{\text{old}}) &= \frac{c\,\phi_{\text{old}}\,c(1-\phi_{\text{old}})}{(c\,\phi_{\text{old}} + c\,(1-\phi_{\text{old}}))^2 + (c\,\phi_{\text{old}} + c\,(1-\phi_{\text{old}}) + 1)} = \frac{\phi_{\text{old}}(1-\phi_{\text{old}})}{c+1}
\end{aligned}$$

Thus, if c is too big, then $\mathrm{Var}(\phi_{\text{prop}} \,|\, \phi_{\text{old}})$ will be too small, or ϕ_{prop} will be close to ϕ_{old}. In this case, this sampler will be very ineffective.

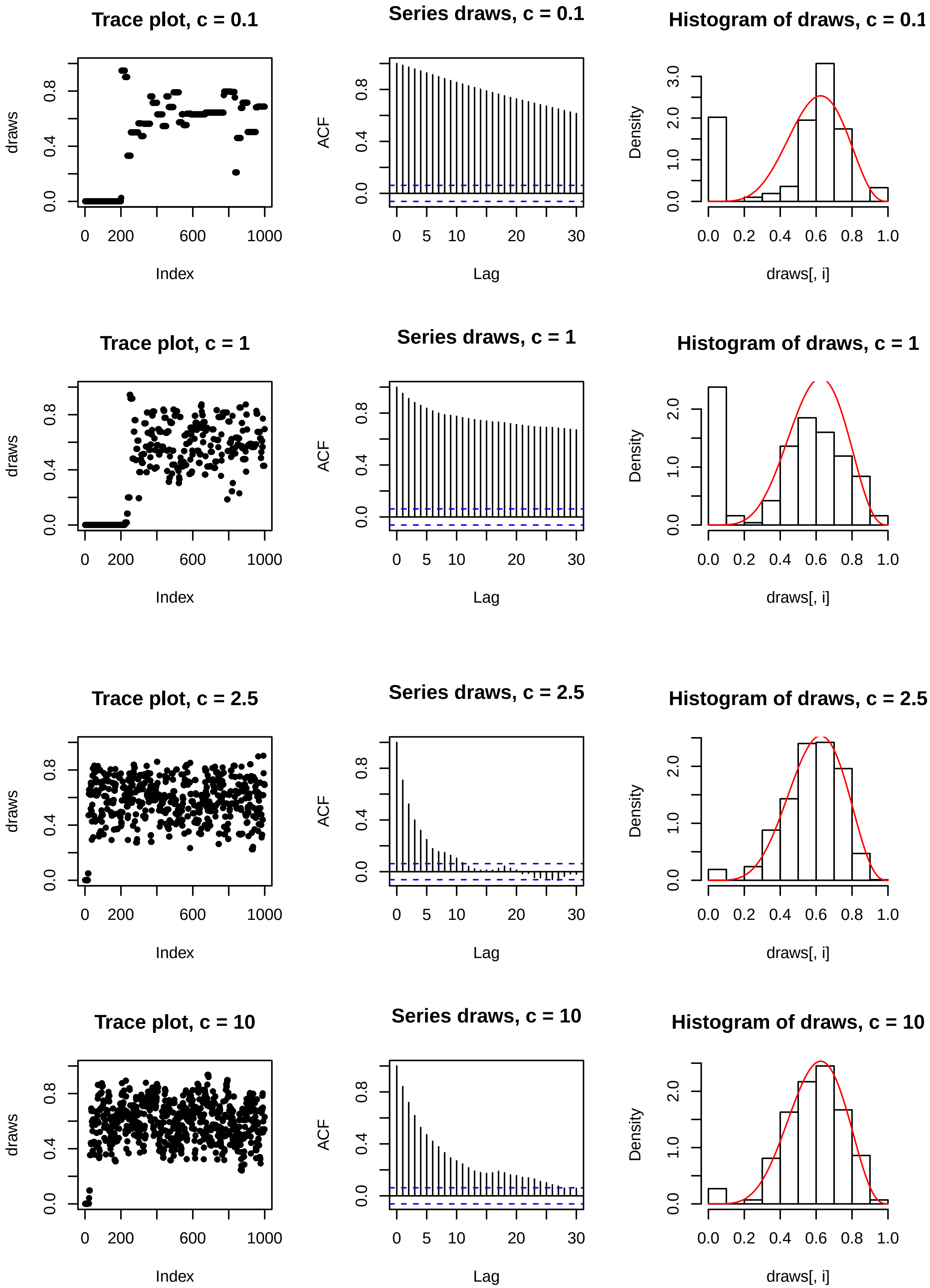
On the other hand, if c is too small, then $\mathrm{Var}(\phi_{\text{prop}} \,|\, \phi_{\text{old}})$ will be too big, or the probability that ϕ_{prop} close to 0 or 1 will be too high. In this case, the acceptance rate will be very low because our target distribution has less probability near 0 or 1.

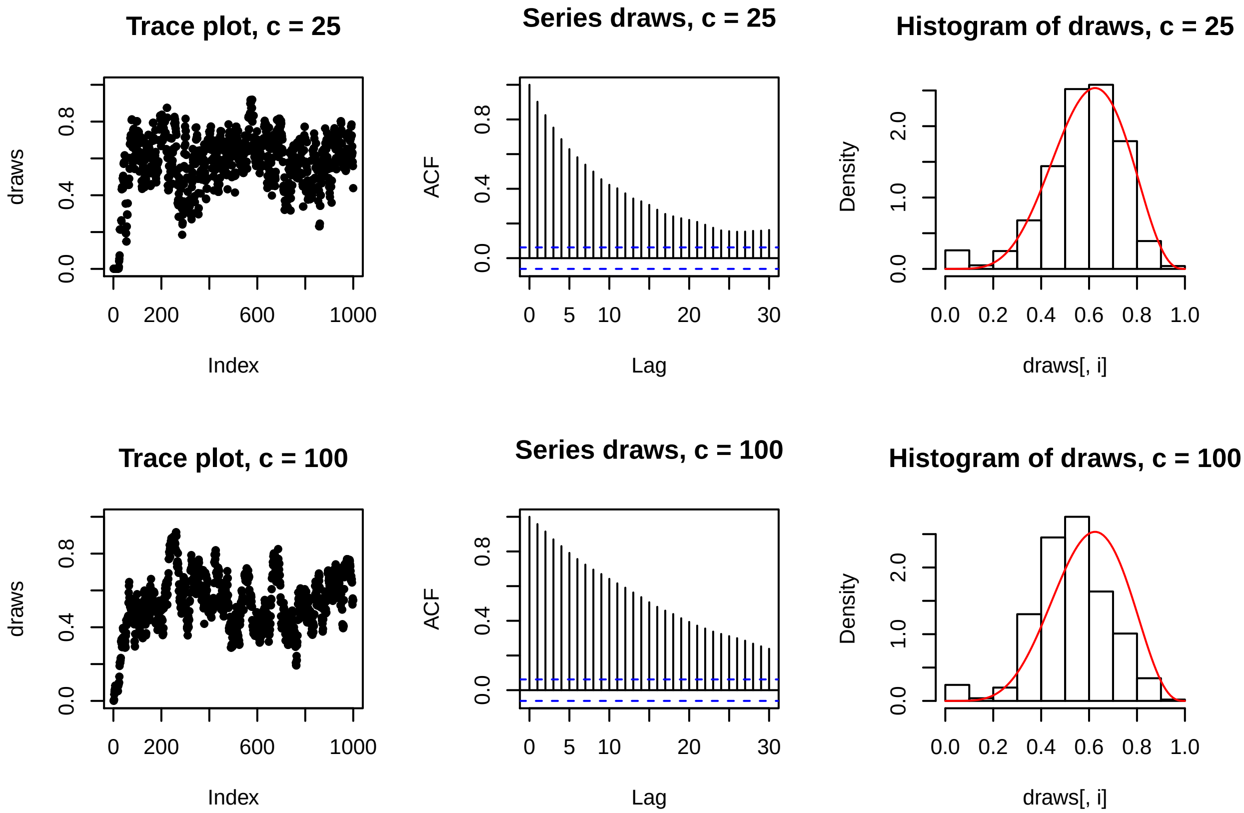
We can plot the pdf of proposal distribution $p(x|y)$ for different c and different y.



According to the above figure, we can see that $c = 2.5$ and $c = 10$ are reasonable choices. To see this, we plot some diagnostic plots for different c (see next page). Note that we only run for 1,000 iteration because we want to compare the effectiveness of these samplers. We also set the initial value to be close to 0, which is the worst case for this target distribution.

From the histogram, we can see that $c = 0.1$ and $c = 1$ are very ineffective. Furthermore, for $c = 25$ and $c = 100$, the draw x_i is too close to its previous draw x_{i-1}. To choose between $c = 2.5$ and $c = 10$, we can use ACF plot. Since the effect of previous draws last longer for $c = 10$, which is undesirable, we will use $c = 2.5$ in this case.


Trace plot, c = 0.1
Series draws, c = 0.1
Histogram of draws, c = 0.1
Trace plot, c = 1
Series draws, c = 1
Histogram of draws, c = 1
Trace plot, c = 2.5
Series draws, c = 2.5
Histogram of draws, c = 2.5
Trace plot, c = 10
Series draws, c = 10
Histogram of draws, c = 10
draws
Index
ACF
Lag
Density
draws[, i]


Trace plot, c = 25
Series draws, c = 25
Histogram of draws, c = 25
draws
0.0
0.4
0.8
0
200
600
1000
Index
ACF
Lag
0
5
10
20
30
Density
0.0
1.0
2.0
0.2
0.6
1.0
draws[, i]
Trace plot, c = 100
Series draws, c = 100
Histogram of draws, c = 100