Michael Lin
STA 360
Lab 3

Part 2:
1. Here is the density:
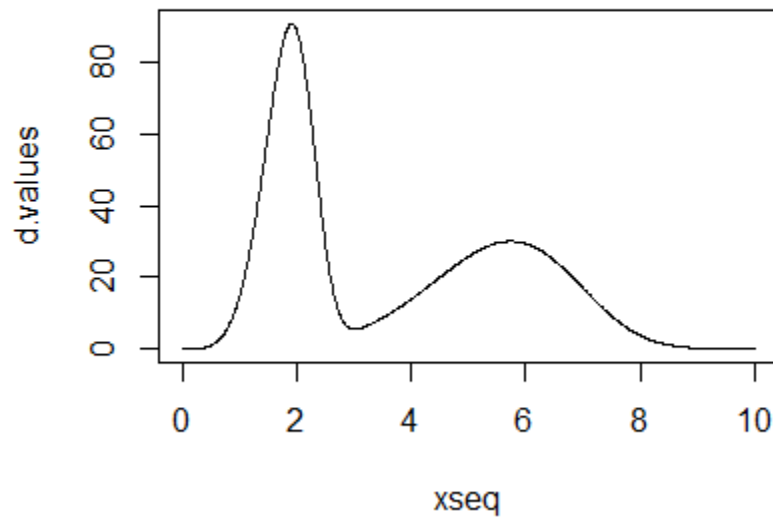

2. Acceptance ratio is 0.2001.

Histogram of samples.uniform
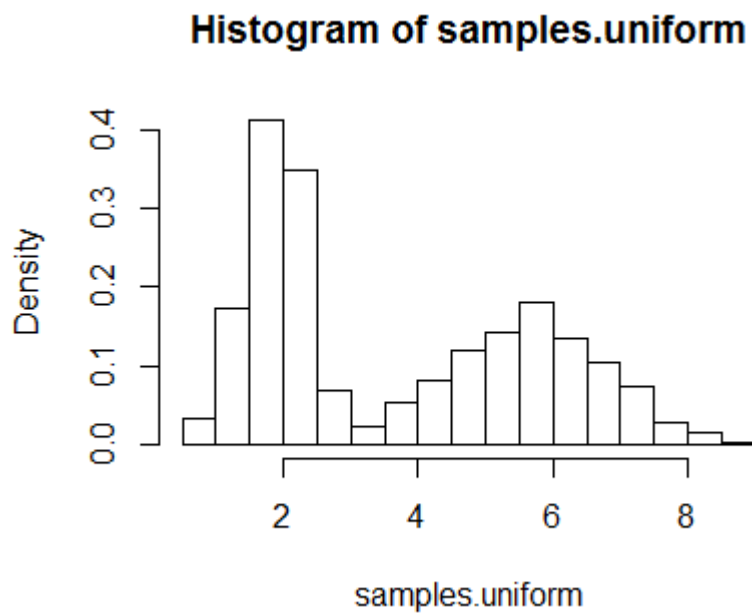
The method produces a histogram that has a distribution approximately equal to the target distribution. This is because the method takes a sample from the envelope distribution: if this sample is also in the target distribution (i.e. less than function value of target distribution), then this point is accepted. To ensure that we gather distribution information for the entire support of the target distribution, the envelope distribution has the same support. In the code, the ratio is compared with a Uniform(0,1) which takes into account that the entire support has been sampled. The acceptance ratio is an empirical value proportional to the ratio between the area under the target distribution and the area under the scaled envelope distribution.

3. A good envelope distribution would roughly follow the shape of the target distribution, has function value greater than the target distribution everywhere, and has most if not all the density within the range of the target distribution's support. In addition, the envelope distribution should be a common distribution (i.e. easy to sample from). The one that I found through trial and error is a gamma distribution, particularly a scaled version (by 500) of the Gamma(3, 0.8) distribution. This is a good distribution because it "envelopes" the target distribution almost completely, and 98.6% of the density is within the support of the target distribution.

4. Acceptance ratio: 0.3815

Histogram of samples.gamma
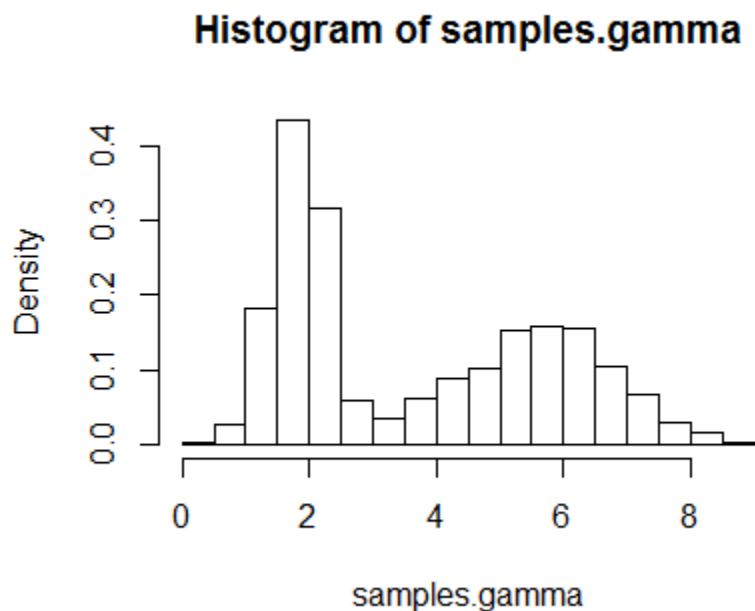
5. It would not work well with high-dimensional multivariate distributions because M (the scaling factor) is generally large, which means more samples will be rejected and it would take a lot longer to construct an adequate histogram approximation for the target distribution.

Part 3:

1. The cdf is:

$$P(x) = \begin{cases} 0, & x \in (-\infty, a] \\ \dfrac{\Phi(x) - \Phi(a)}{\Phi(b) - \Phi(a)}, & x \in (a, b) \\ 1, & x \in [b, \infty) \end{cases}$$

Where $\Phi(\cdot)$ is the normal cdf. Since the pdf is 0 for all x < a, and all x > b, and assume a < b, we know that cdf is 0 for all x<a and 1 for all x>b. For a<x<b, since the pdf is proportional to the normal distribution in this interval, we can find the cdf as if finding the cdf of a normal within this interval, but scaled by a factor to account for the fact that pdf of p(x) from a to b integrates to 1 while pdf of normal from a to b integrates to something less than 1.

It turns out that the distribution of p(x) is called a **truncated normal distribution**.

2. I would sample from distribution using rejection sampling with Normal(m,v) as the envelope distribution (scaled by a factor) so that it covers all of the target distribution. Then, just use the same method as Part 2 to construct a histogram approximation of the target distribution.

3. Sample using inverse transform sampling. First take a sample from a uniform distribution, call this t. Since we have the quantile-function F^{-1} for the cdf of p(x), we can find a random variable x such that

$$x = F^{-1}(t)$$

We then simply take x as the random sample from p(x). This procedure would be repeated numerous times to get many x values, which we can then use to construct a histogram approximation of the distribution p(x).

A. Uniform envelope distribution, R code:

Plot the density ####
xseq <- seq(from=0,to=10,by=0.01)

d.values <- xseq^4*(15*exp(-(xseq/2)^5)+5/81*exp(-(xseq/6)^5))

unif.thresh = 980

plot(xseq,d.values,type="l", ylim=c(0,110))
lines(xseq, unif.thresh*dunif(xseq, min = 0, max = 10))

Uniform envelope function ####
density.function <- function(x){
  return(x^4*(15*exp(-(x/2)^5)+5/81*exp(-(x/6)^5)))
}

samples.uniform <- NULL
N=10000
for (i in 1:N) {
  proposal <- runif(1, min = 0, max = 10) # Here we get a proposal value
  density.ratio <- density.function(proposal)/(unif.thresh*dunif(proposal, min = 0, max = 10)) # We calculate the ratio of the densities
  if (runif(1) < density.ratio) samples.uniform <- c(samples.uniform,proposal) # If a random uniform is lower than our ratio, we accept our sample, otherwise we reject. Then we repeat this process
}

hist(samples.uniform,freq=FALSE)
print(paste("Acceptance Ratio: ",length(samples.uniform)/N))

B. Gamma (improved) envelope distribution, R code.

Plot the density ####
```
xseq <- seq(from=0,to=10,by=0.01)

d.values <- xseq^4*(15*exp(-(xseq/2)^5)+5/81*exp(-(xseq/6)^5))

a = 3; b = 0.8;
thresh = 500

plot(xseq,d.values,type="l", ylim=c(0,100))
lines(xseq, thresh*dgamma(xseq, shape = a, rate = b), col = "red")
```

Uniform envelope function ####
```
density.function <- function(x){
  return(x^4*(15*exp(-(x/2)^5)+5/81*exp(-(x/6)^5)))
}


samples.gamma <- NULL
N=10000
for ( i in 1:N ) {
  proposal <- rgamma(1, shape = a, rate = b) # Here we get a proposal value
  density.ratio <- density.function(proposal)/(thresh*dgamma(proposal, a, b))  # We calculate the ratio of the densities
  if ( runif(1) < density.ratio ) samples.gamma <- c(samples.gamma,proposal) # If a random uniform is lower than our ratio, we accept our sample, otherwise we reject. Then we repeat this process
}

hist(samples.gamma,freq=FALSE)
print(paste("Acceptance Ratio: ",length(samples.gamma)/N))
```