Homework #5 STA 360

Jerry Xin STA 360: Homework 5

Due Friday September 24th, 5 PM EDT

```
library(plyr)
library(ggplot2)
library(dplyr)
library(xtable)
library(reshape)
library(tidyverse)
```

1. (15 points, 5 points each) Hoff, 3.12 (Jeffery's Prior).

(a)

See attached Document.

(b)

See attached Document.

(c)

See attached Document.

2. Lab component (25 points total) Please refer to lab 4 and complete tasks 4—5.

(a) Task 4 (Finish For Homework)

Plot the densities of $f(x)$ and the Unif(0,1) on the same plot.

Let's first create a sequence of points from 0 to 1, so that we can have a grid of points for plotting both of the proposed functions.

```
# grid of points
x <- seq(0, 1, 10^-2)
```

```
fx <- function(x) sin(pi * x)^2
plot(fx, xlim = c(0,1), ylim = c(0,1.5), ylab = "f(x)", lwd = 2)
curve(dunif, add = TRUE, col = "blue", lwd = 2)
curve(dbeta(x,2,2), add = TRUE, col = "red", lwd = 2)
legend("bottom", legend = c(expression(paste("sin(",pi,"x)"^"2")),"Unif(0,1)",
"Beta(2,2)"), col = c("black", "blue", "red"), lty = c(1,1,1), bty = "n", cex = 1.1, lwd = 2)
```

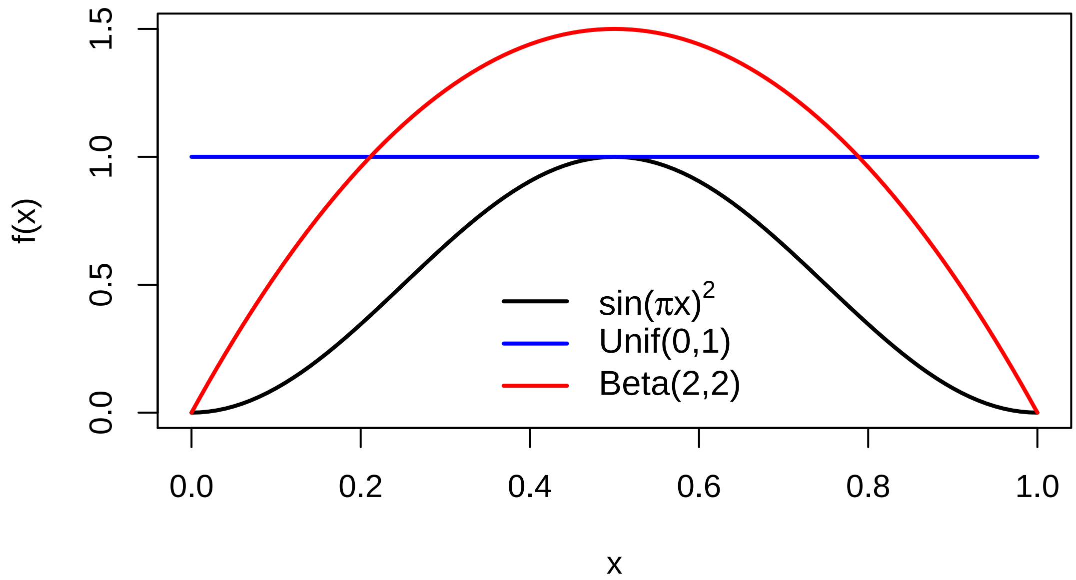

Figure 1: Comparision of the target function and the Unif(0,1) and the Beta(2,2) densities on the same plot.

According to the rejection sampling approach sample from f(x) using the Unif(0,1) pdf as an enveloping function. In order to do this, we write a general rejection sampling function that also allows us to plot the historams for any simulation size. Finally, our function also allows us to look at task 4 quite easily.

```
set.seed(1)
sim_fun <- function(f, envelope = "unif", par1 = 0, par2 = 1, n = 10^2,
                    plot = TRUE){

  r_envelope <- match.fun(paste0("r", envelope))
  d_envelope <- match.fun(paste0("d", envelope))
  proposal <- r_envelope(n, par1, par2)
  density_ratio <- f(proposal) / d_envelope(proposal, par1, par2)
  samples <- proposal[runif(n) < density_ratio]
  acceptance_ratio <- length(samples) / n
  if (plot) {
    hist(samples, probability = TRUE,
         main = paste0("Histogram of ",
                       n, " samples from ",
                       envelope, "(", par1, ",", par2,
                       ").\n Acceptance ratio: ",
                       round(acceptance_ratio,2)),
                       cex.main = 0.75)
  }
  list(x = samples, acceptance_ratio = acceptance_ratio)
}
```

```
set.seed(1)
par(mfrow = c(2,2), mar = rep(4, 4))
unif_1 <- sim_fun(fx, envelope = "unif", par1 = 0, par2 = 1, n = 10^2)
unif_2 <- sim_fun(fx, envelope = "unif", par1 = 0, par2 = 1, n = 10^5)
# ATTN: You will need to add in the
# Beta(2,2) densities on your own to finish task 4.
```

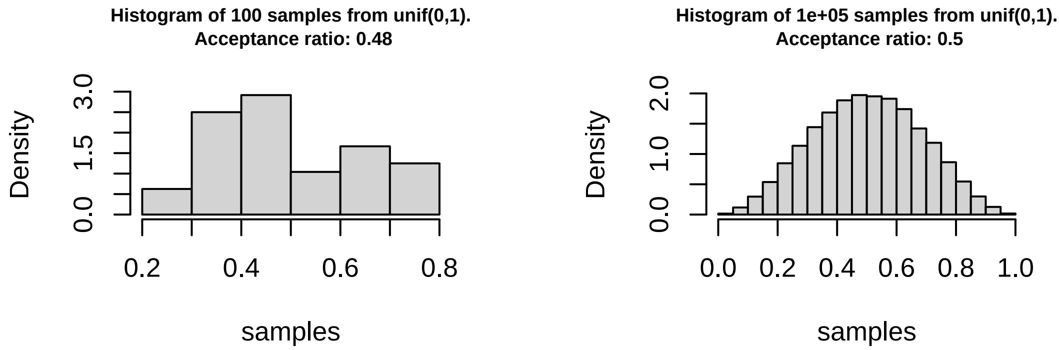

Figure 2: Comparision of the output of the rejection sampling for 100 versus 100,000 simulations with the uniform distribution as the envelope function.

```r
par(mfrow = c(1,1))
```

```r
set.seed(1)
par(mfrow = c(2,2), mar = rep(4, 4))
beta_1 <- sim_fun(fx, envelope = "beta", par1 = 2, par2 = 2, n = 10^2)
beta_2 <- sim_fun(fx, envelope = "beta", par1 = 2, par2 = 2, n = 10^5)
# ATTN: You will need to add in the
# Beta(2,2) densities on your own to finish task 4.
```

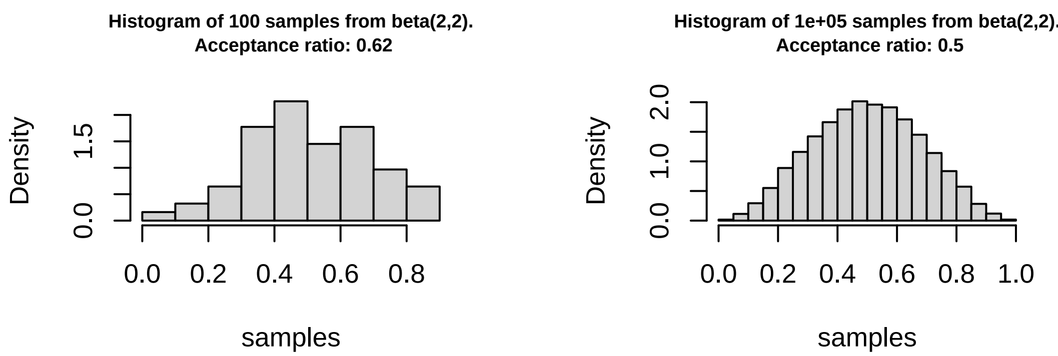

Figure 3: Comparision of the output of the rejection sampling for 100 versus 100,000 simulations with the beta distribution as the envelope function.

Results: The acceptance ratio for sampling 100 samples from a beta(2,2) function is 0.62. The acceptance ratio for sampling 10ˆ5 samples from a beta(2,2) function is 0.5. The acceptance ratio for sampling 100 samples from a uniform(0,1) function is 0.48. The acceptance ratio for sampling 10ˆ5 samples from a uniform(0,1) function is 0.5.

The histogram of sampling 100 samples from a beta(2,2) seems very roughly in the shape of a normal distribution. It seems that sampling more, in this case sampling 10ˆ5 compared to 100, will cause the histogram of samples from a beta(2,2) to roughly approach a normal distribution.

The histograms of Beta(2,2) and unif(0,1) are approximately equal at 10ˆ5 samples. Both seem to approach normal distributions that are symmetric and unimodal. There are no outliers for beta(2,2) or the unif(0,1) graphs for both 100 and 10ˆ5 samples.

The histogram of sampling 100 samples from a uniform(0,1) seems very roughly skewed right, or a bimodal distribution. It seems that sampling more, in this case sampling 10ˆ5 compared to 100, will cause the histogram of samples from a uniform(0,1) to roughly approach a normal distribution.

(b) Task 5 (Finish For Homework)

Comparing the acceptance ratios of 100 samples of the beta(2,2) and uniform(0,1) initially might indicate that the beta(2,2) function is better, since the acceptance ratio of 0.62 is greater than 0.48, which means that overall more samples got accepted. However, when expanding to 10^5 samples and looking at the bigger picture in the long run, the acceptance ratios of the beta(2,2) an the uniform(0,1) function become the same at 0.5 each, which suggests that the beta(2,2) function is no better than a uniform(0,1) function in rejection sampling.

This is the case because the area between the unif(0,1) and the sine function is approximately equal to the area between the beta(2,2) and the sine function. This is the acceptance ratios are similar(both are approx. 0.5), and it is hard to recommend which one is better.

Overall, there seems to be no statisitically signifigant difference between the beta(2,2) and the uniform(0,1) as an enveloping functions, so I would not be able to recommend one over the other since they are both roughly the same when looking at the acceptance ratio over a large quantity such as 10^5.(both have acceptance ratios of approximately 0.5)

If I were trying to find an enveloping function that had an high acceptance ratio, I would personally try the Gaussian Distribution(Normal Distribution). The shape of

$$f(x) = sin^2(\pi x)$$

is approximately a bellcurve shape, and it has a maximum as well as inflection points which indicate concavity. This is really similar to the Gaussian distribution, so it seems like it would have a high acceptance ratio(smaller gap between enveloping function and the target function), so I would try that.