

# Homework 4

*Amil Khan, Calvin*

***Due on November 11, 2018 at 11:59 pm***

**Note:** If you are working with a partner, please submit only one homework per group with both names and whether you are taking the course for graduate credit or not. Submit your Rmarkdown (.Rmd) and the compiled pdf on Gauchospace.

## Problem 1

Importance sampling of the Galenshore distribution. A *Galenshore*( $\alpha, \theta$ ) distribution has density:

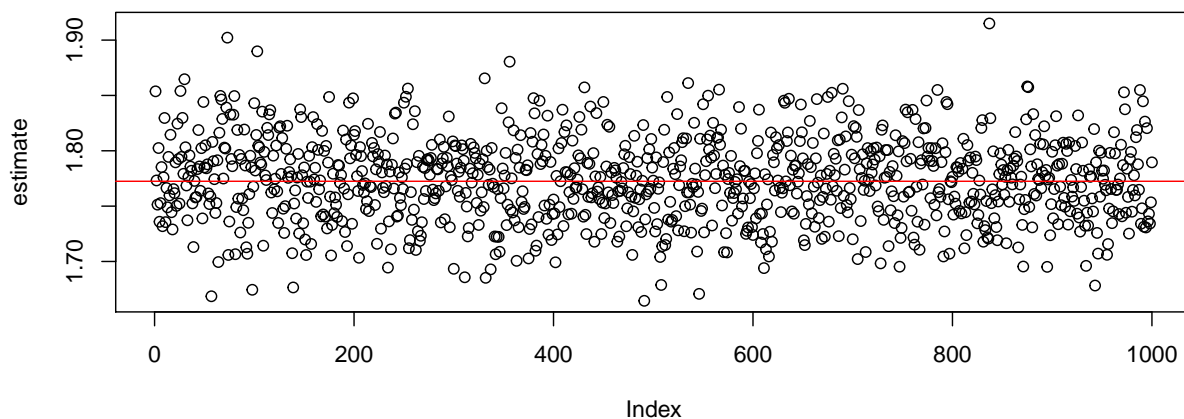
$$p(y) = \frac{2}{\Gamma(\alpha)} \theta^{2\alpha} y^{2\alpha-1} e^{-\theta^2 y^2}$$

for  $y > 0$ ,  $\theta > 0$  and  $\alpha > 0$ .

### Part a.)

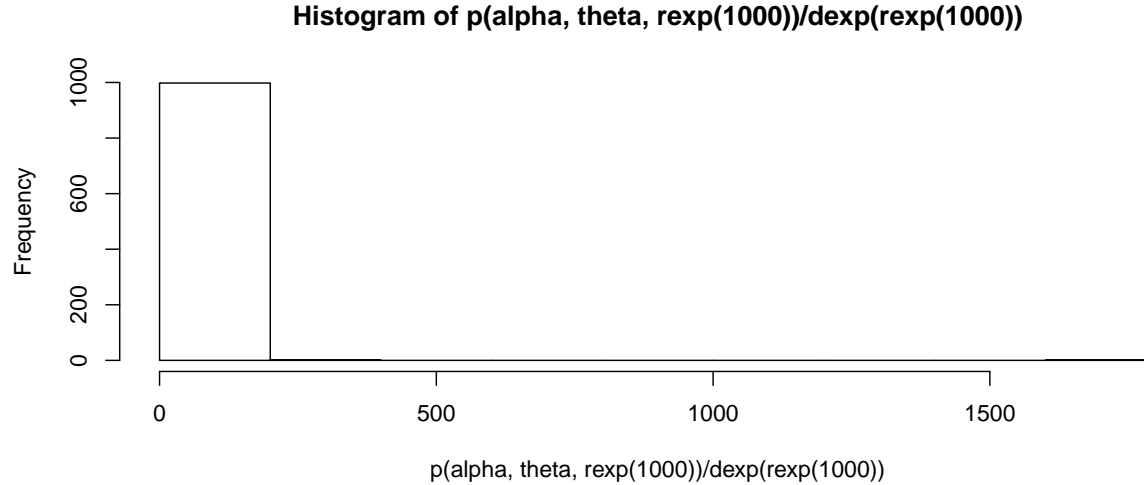
Assume  $\alpha = 1$  and  $\theta = 0.5$ . For  $k = 10, 100$  and  $1000$ , draw samples of size  $k$  from an  $\text{Exponential}(1)$  distribution. Use these samples as important sampling proposals to estimate the mean of the Galenshore. Compare your estimates for each  $k$  to the true mean:

$$E[Y|\alpha, \theta] = \frac{\Gamma(\alpha + 1/2)}{\theta \Gamma(\alpha)}$$



### Part b.)

Plot a histogram of the log importance weights for  $k = 1000$ .



**Part c.)**

Now draw samples of size  $k \in \{10, 100, 1000\}$  from  $Z \sim N(1.5, 1)$ . For each  $k$ , compute and print the estimated mean. Plot the histogram of log importance weights for  $k = 1000$  (exclude the points for which  $Z < 0$  when plotting this histogram). Based on the two histograms and the comparison between estimates and true mean, which proposal do you think is better?

```
## [1] -7.704892
```

```
## [1] 1.175353
```

## Problem 2

Rejection Sampling for the Semi-Circle Distribution. Let

$$p(x|R) = \frac{2}{\pi R^2} \sqrt{R^2 - x^2} \mathbf{1}_{\{x \in [-R, R]\}}$$

**Part a.)**

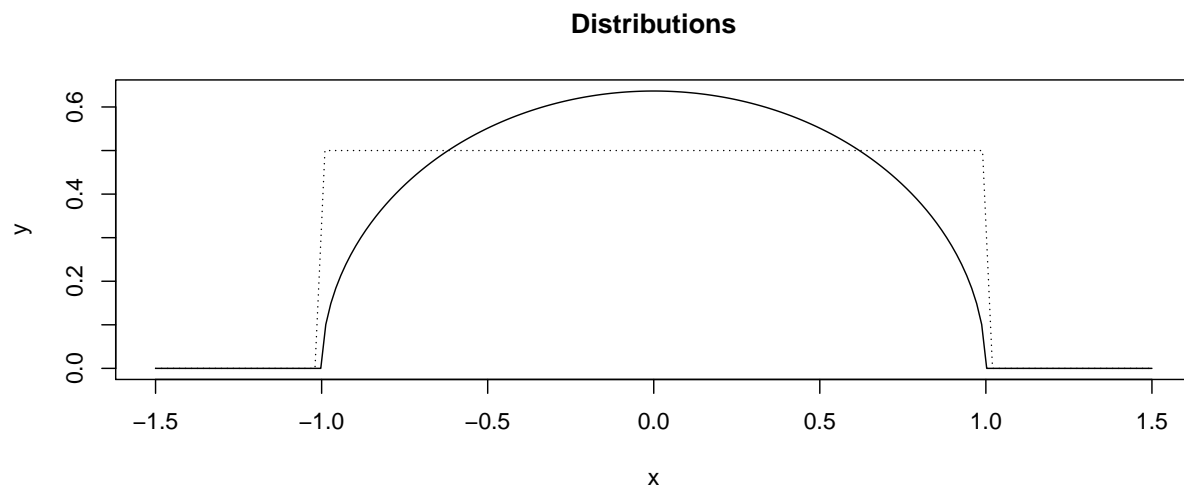
Identify a proposal distribution,  $g(x|R)$  for rejection sampling. Show that there exists some  $M$  such that  $Mg(x|R) \geq p(x|R) \forall x \in [-R, R]$ . What is  $M$ ?

For rejection sampling, let  $g(x|R) = \frac{1}{2R} \mathbf{1}_{[-R, R]}$ .

We see that, if  $M = \frac{4}{\pi}$ , then  $p(x|R) = \frac{2}{\pi R^2} \sqrt{R^2 - x^2} \mathbf{1}_{[-R, R]} \leq \frac{2}{\pi R} \mathbf{1}_{[-R, R]} = Mg(x|R)$ .

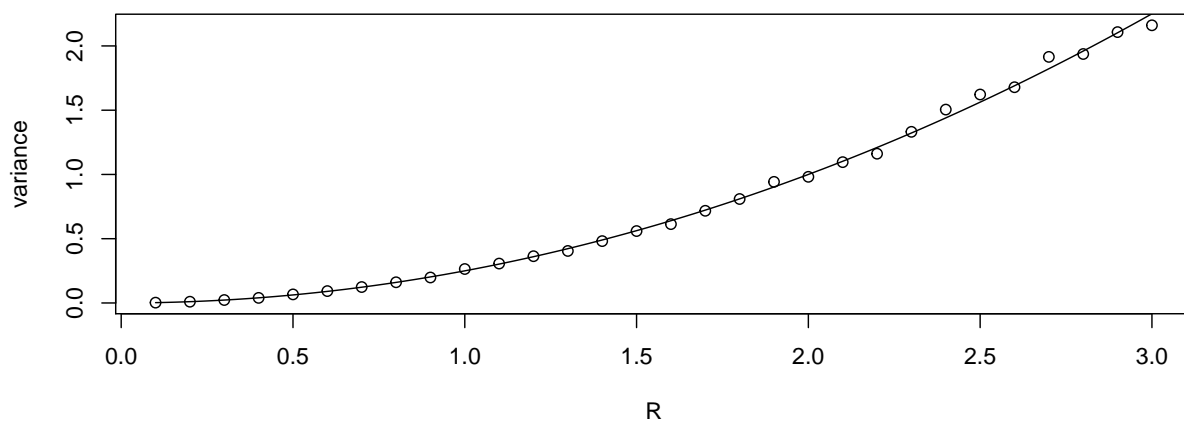
**Part b.)**

Plot the semi-circle distribution for  $R = 1$  and your proposal distribution on the same plot using the `curve` function. Note that because the density must be normalized to integrate to 1, the semi-circle distribution is really a “semi-ellipse”.



**Part c.)**

For  $R = 0.1, 0.2, \dots, 3.0$ , use rejection sampling to sample from  $p(x|R)$ . Plot the sample variance for each value of  $R$ , and show graphically that the variance for the semi-circle distribution is  $\frac{R^2}{4}$ .



## Problem 3

Rejection sampling and interval estimation.

### Part a.)

Use rejection sampling to sample from the following density:

$$p(x) = |\sin(x)| \times \mathbf{1}[0, 2\pi]$$

Generate at least 1000 true samples from the above density. Compute and report the Monte Carlo estimate of the upper and lower bound for the 50% quantile interval using the `quantile` function on your samples. Compare this to the 50% HPD region calculated on the samples. What are the bounds on the HPD region? Report the length of the quantile interval and the total length of the HPD region. What explains the difference? Hint: to compute the HPD use the `hdi` function from the `HDInterval` package. As the first argument pass in `density(samples)`, where `samples` is the name of your vector of true samples from the density. Set the `allowSplit` argument to true and use the `credMass` argument to set the total probability mass in the HPD region to 50%.

```
##      lower      upper
## 0.01987962 0.24999762
## attr("credMass")
## [1] 0.95
```

### Part b.)

Plot  $p(x)$  using the `curve` function. Add lines corresponding to the intervals / probability regions computed in the previous part to your plot using them `segments` function. To ensure that the lines don't overlap visually, for the HPD region set `y0` and `y1` to 0 and for the quantile interval set `y0` and `y1` to 0.01. Make the segments for HPD region colored red and the segment for quantile interval colored blue.