Monte Carlo Inference - Lab 2
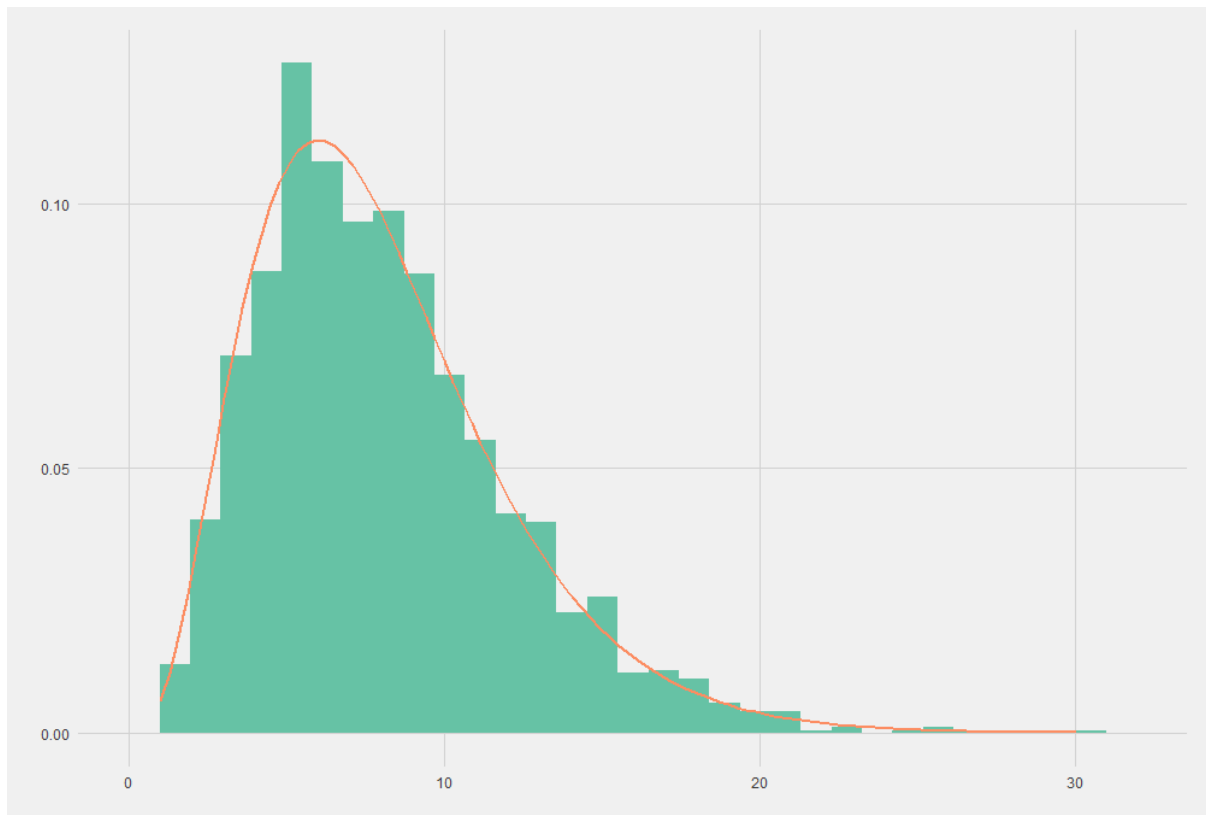
Daniel Jordan 13496112

This lab details methods of simulating distributions using the method of inversion and rejection sampling.

Inversion Method

The first histogram is an approximation of the gamma distribution by distribution taking 2000 samples random uniform samples and inputting them to the inverted gamma distribution.


The orange line is the plot of a gamma distribution with a shape of 4 and rate of 0.5. As we can see, the sampled distribution is well apporximated using the inversion method.

The Kolmogorov-Smirnov test outputs:

```
D = 0.021824, p-value = 0.2966
alternative hypothesis: two-sided
```

Rejection Sampling

Below is the plot of a sample gamma distribution with shape = 4.5, rate = 3. We use rejection sampling with a known distribution - gamma with shape = 4.5, rate = 3.
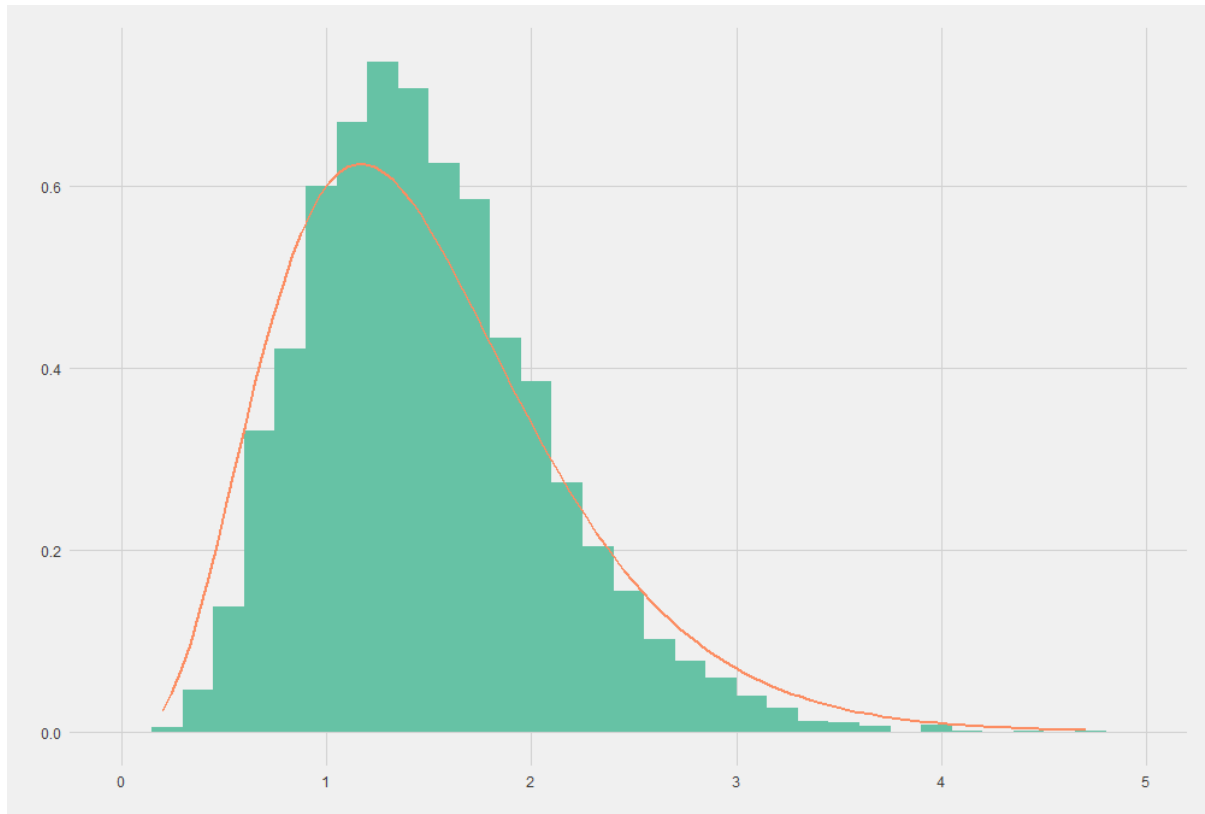
The sample approximates the target distribution well. The Kolmogorov-Smirnov test outputs:

```
D = 0.064943, p-value < 2.2e-16
alternative hypothesis: two-sided
```

By maximising the ratio of the target distribution to the known distribution, I found M = 1.93986

The proportion of accepted samples and rejected samples should converge to 1/M, the proportion of the target distribution and the known distribution.

Issues & Questions

I struggled with this lab, and am quite confused as a result. I'm quite certain I understand rejection sampling however I know my answers are incorrect and it's killing me.

Why did we pick Gamma(4, 2) as the envelope function when sampling from Gamma(4.5, 3)?
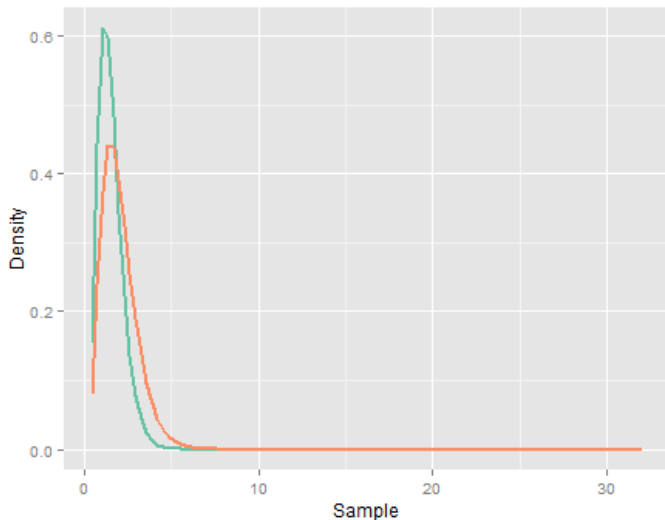
I plotted these two distributions against each other and it shows that the envelope function does not encapsulate the target distribution.

The green line is Gamma(4, 2) and the orange line is Gamma(4.5, 3). It's clear they cross, so it cannot be a valid envelope function.

We use Gamma(4, 2) because it is the closest distribution with integer parameters, but I don't understand why this is a suitable known distribution to sample from.
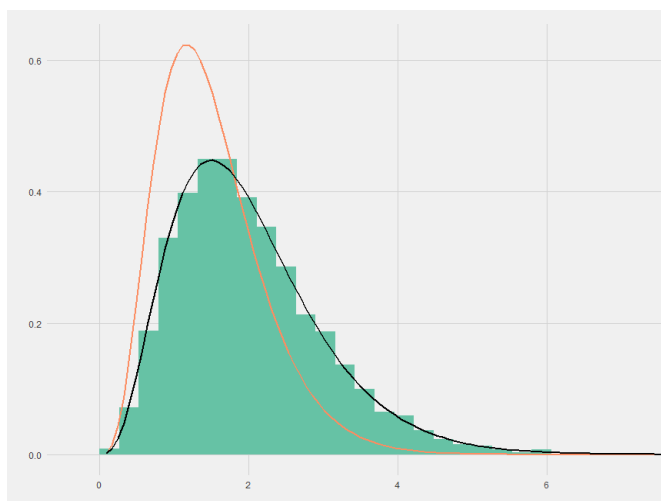
So I decided to go back, rewrite the rejection sampling function and plot all the samples from the known distribution, regardless of whether they were accepted as valid or not. That is the second plot on the page. It matches the distribution really well which makes sense.

Then I filtered for only accepted samples from the known distribution and plotted them. These are the samples that we are hoping will approximate to our target distribution, however the plot looks incorrect.
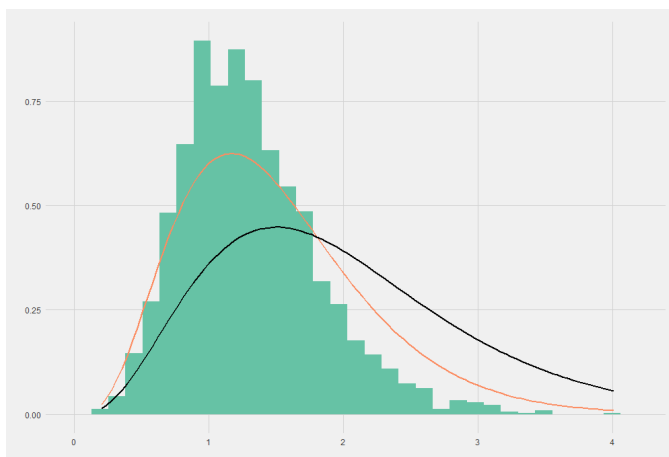
I'm sure I have some fundamental misunderstanding here, or I've made some basic mistake but I can't seem to work it out.