

(1) Use the probability integral transformation method to simulate from the distribution

$$f(x) = \begin{cases} \frac{2}{a^2}x, & \text{if } 0 \leq x \leq a \\ 0, & \text{otherwise} \end{cases} \quad (1)$$

where  $a > 0$ . Set a value for  $a$ , simulate various sample sizes, and compare results to the true distribution.

Solution

1. In order to solve this problem the following needs to be done. First we must take the given PDF equation and transform it into a CDF equation using integration. We will integrate the function from 0 to  $x$  as all values outside of this range are 0. Then, once we define the CDF we will need to find the inverse function. Once the inverse function is defined we will take a random sample from a uniform distribution. Using these values we will plug them into the inverse function we defined.

The CDF can be found by integrating the PDF from 0 to  $x$

$$F(x) = \int_0^x \frac{2}{a^2}t \, dt \quad (2)$$

We can factor out the constant:

$$F(x) = \frac{2}{a^2} \int_0^x t \, dt \quad (3)$$

Now we take the antiderivative of  $t$ :

$$\int t \, dt = \frac{t^2}{2} \quad (4)$$

This gives:

$$F(x) = \frac{2}{a^2} \left[ \frac{t^2}{2} \right]_0^x \quad (5)$$

Plugging in upper and lower limits:

$$F(x) = \frac{2}{a^2} \left( \frac{x^2}{2} - \frac{0^2}{2} \right) \quad (6)$$

We are then left with

$$F(x) = \frac{2}{a^2} * \frac{x^2}{2} \quad (7)$$

Which can be simplified as

$$F(x) = \frac{x^2}{a^2} \quad (8)$$

As we will be using a uniform distribution we will define the following

$$u = \frac{x^2}{a^2} \quad (9)$$

We will then need to solve for x in terms of u. Starting by multiplying each side by  $a^2$

$$ua^2 = x^2 \quad (10)$$

We can then take the root of this entire expression and rearrange to put x first

$$x = a\sqrt{u} \quad (11)$$

Using this inverse function we can sample a distribution by first taking a sample from a uniform distribution and then transforming it with the inverse equation

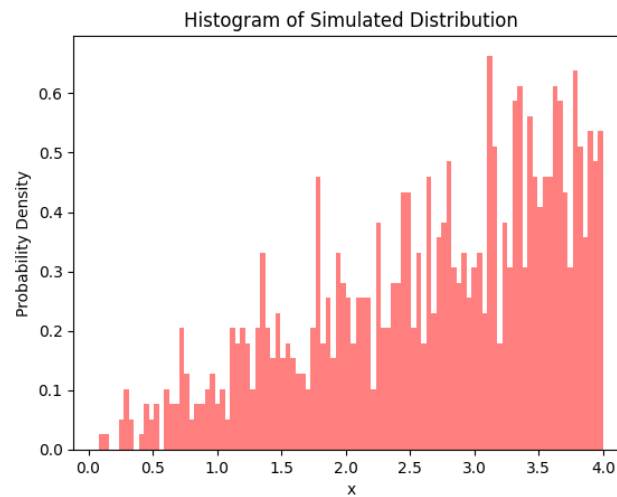


Figure 1: Histogram of simulated distribution using inverse CDF method. SS=1,000

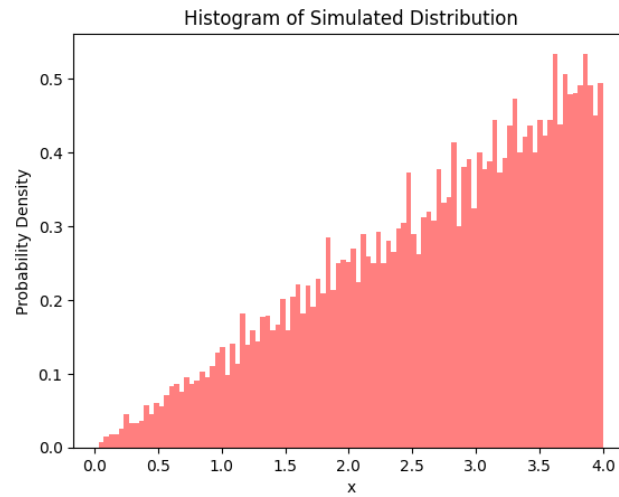


Figure 2: Histogram of simulated distribution using inverse CDF method. SS=10,000

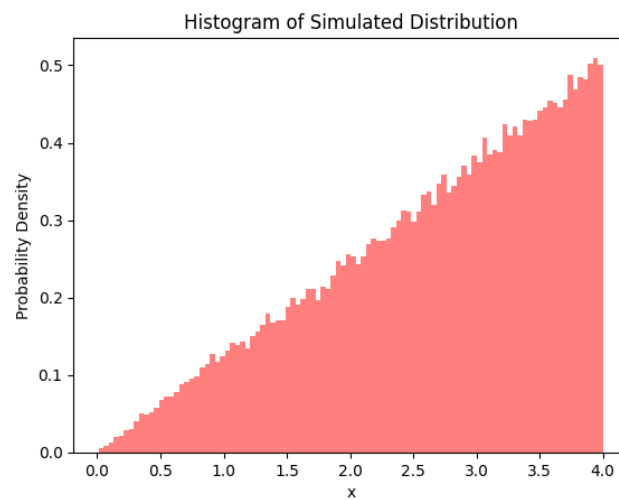


Figure 3: Histogram of simulated distribution using inverse CDF method. SS=100,000

As the number of simulated observations increases, the histogram of the simulated data approaches the shape of the true probability density function, which is a linear function over the interval  $[0, a]$

(2) Generate samples from the distribution

$$f(x) = \frac{2}{3}e^{-2x} + 2e^{-3x} \quad (12)$$

using the finite mixture approach.

1. In order to solve this problem, we first need to determine the weight of each term. The weights are initially  $\frac{2}{3}$  and 2, which will be normalized to 0.25 and 0.75, respectively. Next, we can observe that, rather than the distribution being comprised of two normal distributions as in the in-class example, this problem involves two exponential distributions. Therefore, we will need to adjust the code to allow us to sample from exponential distributions instead. Specifically, for each observation, we will randomly select which exponential distribution to sample from based on the normalized weights. After determining which component to use, we will sample from the corresponding exponential distribution, which have rate parameters  $\lambda = 2$  and  $\lambda = 3$ , corresponding to scales of  $\frac{1}{2}$  and  $\frac{1}{3}$ , respectively. Finally, we will generate and plot histograms of the simulated data for various sample sizes to compare the results.

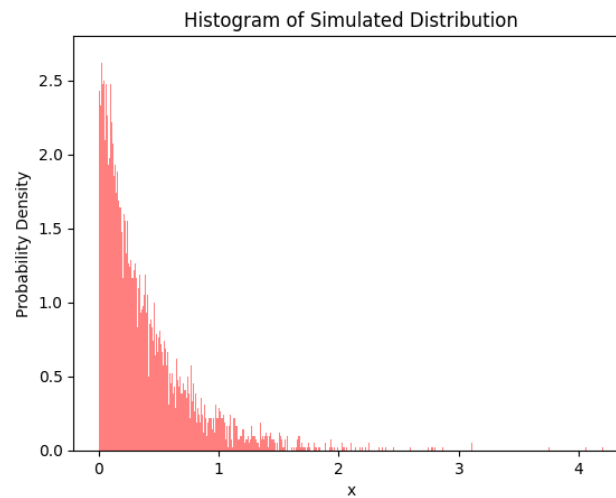


Figure 4: Histogram of simulated distribution. SS=10,000

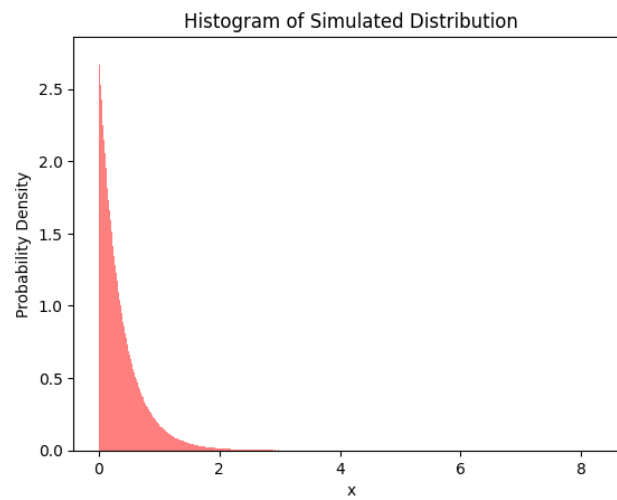


Figure 5: Histogram of simulated distribution. SS=1,000,000

As the sample size increases, the histogram closely approximates the shape of the true mixture distribution.

(3) Draw 500 observations from  $\text{Beta}(3,3)$  using the accept-reject algorithm. Compute the mean and variance of the sample and compare them to the true values.

1. In order to solve this problem, we will use the accept-reject algorithm to generate 500 observations from the  $\text{Beta}(3,3)$  distribution. First, we define the target distribution, which is the  $\text{Beta}(3,3)$  probability density function. Next, we choose a simple proposal distribution that is easy to sample from; in this case, we use the  $\text{Uniform}(0,1)$  distribution. We then determine the appropriate scaling constant  $c$  by evaluating the ratio of the target density to the proposal density at the mode of the target distribution. After setting up the accept-reject sampling function, we repeatedly generate candidate samples from the proposal distribution and accept them based on the acceptance criterion involving the target and proposal densities. Once 500 observations are drawn, we compute the sample mean and variance. Finally, we compare the sample statistics to the true mean and variance of the  $\text{Beta}(3,3)$  distribution, which are 0.5 and 0.035714, respectively.

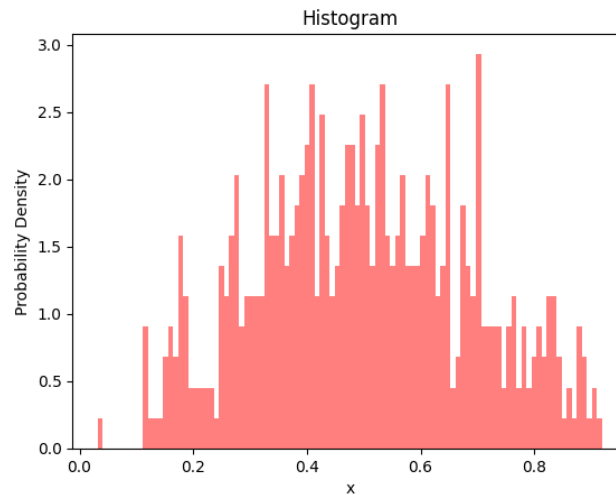


Figure 6: Histogram of simulated distribution. 500 observations

Sample mean: 0.4996254654115793

Sample variance: 0.035416819475811136

The sample mean is 0.4996 and the sample variance is 0.0354. Both the sample mean and variance closely approximate the true mean (0.5) and true variance (0.035714), differing only by small amounts in the thousandths place.