Problem Set 2

Abdulah Abrahim

ECON 6100, Applied Bayesian Methods
March 16th, 2025

Problem 1: Use the probability integral transformation method to simulate from the distribution

$$f(x) = \begin{cases} \frac{2}{a^2}x, & \textbf{if } 0 \le x \le a \\ 0, & \textbf{otherwise} \end{cases} \tag{1}$$

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

To simulate from the given distribution using the probability integral transformation method, follow these steps:

1. **Find the cumulative distribution function (CDF):** The CDF is obtained by integrating the PDF:

$$F(x) = \int_0^x \frac{2}{a^2} t \, dt = \frac{x^2}{a^2}, \quad \text{for } 0 \le x \le a.$$

2. **Set the CDF equal to a uniform random variable U:** Let $U \sim$ Uniform$(0, 1)$. Then:

$$F(X) = U \implies \frac{X^2}{a^2} = U.$$

3. **Solve for X:**
$$X = a\sqrt{U}.$$

4. **Simulate samples for various sample sizes:** - Choose $a = 2$. - Generate U from Uniform$(0, 1)$. - Compute $X = 2\sqrt{U}$.

5. **Compare the empirical CDF to the true CDF:** - For sample sizes $n = 100, 1000, 10000$, compute the empirical CDF and compare it to the true CDF $F(x) = \frac{x^2}{4}$. - As n increases, the empirical CDF should converge to the true CDF.

See python code file for histogram visualization of comparison to the true PDF.

Problem 2: Generate samples from the distribution

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

using the finite mixture approach.

The given density is a finite mixture of two exponential distributions:

$$f(x) = \frac{2}{3} \cdot 2e^{-2x} + \frac{1}{3} \cdot 3e^{-3x}.$$

Here, the mixture weights are $w_1 = \frac{2}{3}$ and $w_2 = \frac{1}{3}$, and the component distributions are $\text{Exp}(2)$ and $\text{Exp}(3)$, respectively.

1. **Simulate from the mixture:** - Generate $U \sim \text{Uniform}(0,1)$. - If $U \le \frac{2}{3}$, draw X from $\text{Exp}(2)$. - Otherwise, draw X from $\text{Exp}(3)$.

2. **Verify the simulation:** - The theoretical mean and variance of the mixture distribution are:

$$\text{Mean} = \frac{2}{3} \cdot \frac{1}{2} + \frac{1}{3} \cdot \frac{1}{3} = \frac{1}{3} + \frac{1}{9} = \frac{4}{9} \approx 0.4444,$$

$$\text{Variance} = \frac{2}{3} \cdot \frac{1}{4} + \frac{1}{3} \cdot \frac{1}{9} = \frac{1}{6} + \frac{1}{27} = \frac{11}{54} \approx 0.2037.$$

- Compute the empirical mean and variance of the simulated data and compare them to the theoretical values.

See python code file for histogram visualization of comparison to the true PDF.. For example, with $n = 1000$, the empirical mean should be close to 0.4444, and the empirical variance should be close to 0.2037.

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

To draw samples from Beta$(3,3)$ using the accept-reject algorithm:

1. **Choose a proposal distribution:** Use Uniform$(0,1)$ as the proposal distribution $g(x)$.

2. **Find the maximum ratio M:** The density of Beta$(3,3)$ is:

$$f(x) = \frac{x^2(1-x)^2}{B(3,3)}, \quad \text{where } B(3,3) = \frac{\Gamma(3)\Gamma(3)}{\Gamma(6)} = \frac{4!}{5!} = \frac{1}{30}.$$

The maximum of $f(x)$ occurs at $x = 0.5$, so:

$$M = \frac{f(0.5)}{g(0.5)} = \frac{(0.5)^2(0.5)^2/(1/30)}{1} = \frac{1/16}{1/30} = \frac{30}{16} = 1.875.$$

3. **Implement the accept-reject algorithm:** - Generate $U \sim \text{Uniform}(0,1)$ and $Y \sim \text{Uniform}(0,1)$. - Accept Y if $U \leq \frac{f(Y)}{M \cdot g(Y)} = \frac{f(Y)}{1.875}$.

4. **Compute the sample mean and variance:** - The theoretical mean and variance of Beta$(3,3)$ are:

$$\text{Mean} = \frac{3}{3+3} = 0.5,$$

$$\text{Variance} = \frac{3 \cdot 3}{(3+3)^2(3+3+1)} = \frac{9}{6^2 \cdot 7} = \frac{1}{28} \approx 0.0357.$$

- Compute the empirical mean and variance of the simulated data and compare them to the theoretical values.

See python code file for histogram visualization of comparison to the true PDF. The histogram should closely match the theoretical density, confirming the correctness of the accept-reject algorithm. For example, with $n = 500$, the empirical mean should be close to 0.5, and the empirical variance should be close to 0.0357.