STAT 471: Homework 4

Due: November 1, 2025 at 11:59pm

1 Instructions

Please make sure to submit your solutions to the following questions in an .rmd file or preferably a knitted html file.

2 Question 1 (25 points)

Create three simulations of the Exponential distribution with lambda being set to 0.8. You can use rexp() and set the rate parameter to be 0.8. Let the first simulation be a random sample of size 50, the second simulation of size 1000, and the third simulation of size 10000. Plot the histograms of your samples.

(a) Does the Central Limit Theorem apply here? How do you know?
(b) Does the Law of Large Numbers apply here? Verify this by checking the sample mean of each of your simulations and comparing with the true lambda (0.8).

3 Question 2 (25 points)

Suppose that the random variable $X_1 \sim Gamma(2,3)$ and $X_2 \sim Gamma(4,5)$. Generate 1000 random samples from the mixture distribution F, defined as $F = 0.7X_1 + 0.3X_2$. This means that you will mix 70% of samples from X_1 and 30% of samples from X_2. Plot the histogram of your samples.

4 Question 3 (25 points)

Use the continuous inverse CDF method to generate a simulation of 500 random samples from the given pdf, $f_X(x) = 3x^2$ which has support $x \in (0,1)$, similar to the example done in class for the Exponential distribution. Plot the histogram of your samples. *Hint: Find the CDF of the probability density function first, and then write your inverse CDF code.*

5 Question 4 (25 points)

Brownian Motion is often used in financial applications to simulate stock market fluctuations. You will model and plot Brownian motion by constructing your own Brownian motion function.

Step 1. Name your function simBM. Your function should take in the arguments n (sample size) and T (time steps).
Step 2. Create an object within the function called times which stores the sequence of times from 0 to T using the seq() function. The length parameter should be specified to be n+1.
Step 3. Let z be the variable that stores your random samples from the standard normal distribution of size n.
Step 4. Set w to be from 0 to n using the rep() function, and set s to be the square root of the lags in between your time steps, this is called "drift". This means that you take the square root of the difference between your times using diff().

Step 5. Write a for loop to create the Brownian Motion stochastic process. The scheme is given below. return the w's and the t's (the sampled values and time steps).

$w_i = w_{i-1} + s_i * z_i$

Step 6. Set a random seed and use the below code to plot three instances of Brownian motion of size 200 using your simBM function.

```
n = 200
x1 = simBM(n, 1)
x2 = simBM(n, 1)
x3 = simBM(n, 1)
r = range(c(x1$w, x2$w, x3$w))
plot(x1$w, type="l", main="", xlab="t", ylab="W", ylim=r)
lines(x2$w, lty=2)
lines(x3$w, lty=3)
```