MATH 456 — Mathematical Modeling

Assignment #4 **Due Date:** February 28 2022, 11:59 PM

Question 1. Suppose we generate a random variable X in the following way. First we flip a fair coin. If the coin is heads, take X to have a Uniform$(0, 1)$ distribution. If the coin is tails, take X to have a Uniform$(3, 4)$ distribution.

(a) Find the mean of X.

(b) Find the standard deviation of X.

Question 2. Prove that

$$\mathbb{V}(Y) = \mathbb{E}\mathbb{V}(Y|X) + \mathbb{V}\mathbb{E}(Y|X)$$

Recall that $\mathbb{E}(Y|X)$ and $\mathbb{V}(Y|X)$ are the conditional mean and conditional variance of Y given X.

Question 3. [Coding] Let $X_1, \ldots, X_n$ be $N(0, 1)$ random variables and let $\overline{X}_n = n^{-1}\sum_{i=1}^n X_i$. Plot $\overline{X}_n$ versus n for $n = 1, \ldots, 10,000$. Repeat for $X_1, \ldots, X_n \sim$ Cauchy. Explain why there is such a difference.

Question 4. [Coding: Simulating the Stock Market] Let $Y_1, Y_2, \ldots$ be independent random variables such that $\mathbb{P}(Y_i = 1) = \mathbb{P}(Y_i = -1) = 1/2$. Let $X_n = \sum_{i=1}^n Y_i$. Think of $Y_i = 1$ as "the stock price increased by one dollar " $Y_i = -1$ as "the stock price decreased by one dollar" and X_n as the value of the stock on day n.

(a) Find $\mathbb{E}(X_n)$ and $\mathbb{V}(X_n)$.

(b) Simulate X_n and plot X_n versus n for $n = 1, 2, \ldots, 10,000$. Repeat the whole simulation several times. Notice two things. First, it's easy to "see" patterns in the sequence even though it is random. Second, you will find that the runs look very different even though they were generated the same way. How do the calculations in (a) explain the second observation?

Question 5. [Coding] This question is to help you understand the idea of **sampling distribution**. Let $X_1, \ldots, X_n$ be IID with mean μ and variance σ^2. Let $\overline{X}_n = n^{-1}\sum_{i=1}^n X_i$. Then $\overline{X}_n$ is a **statistic**, that is, a function of the data. Since $\overline{X}_n$ is a random variable, it has a distribution. This distribution is called the **sampling distribution of the statistic**. Recall from Theorem 4.16 that $\mathbb{E}(\overline{X}_n) = \mu$ and $\mathbb{V}(\overline{X}_n) = \sigma^2/n$. Don't confuse the distribution of the data f_X and the distribution of the statistic $f_{\overline{X}_n}$. To make this clear, let $X_1, \ldots, X_n \sim$ Uniform$(0, 1)$. Let f_X be the density of the Uniform$(0, 1)$. Plot f_X. Now let $\overline{X}_n = n^{-1}\sum_{i=1}^n X_i$. Find $\mathbb{E}(\overline{X}_n)$ and $\mathbb{V}(\overline{X}_n)$. Plot them as a function of n. Comment. Now simulate the distribution of $\overline{X}_n$ for $n = 1, 5, 25, 100$. Check the simulated values of $\mathbb{E}(\overline{X}_n)$ and $\mathbb{V}(\overline{X}_n)$ agree with your theoretical calculations. What do you notice about the sampling distribution of $\overline{X}_n$ as it increases?

Note: Document your results for coding questions using word, latex or markdown, and submit the documents along with the codes.