MH1820 Introduction to Probability and Statistical Methods
Tutorial 9 (Week 10)

Problem 1 (Distribution of Sample Mean)

(a) Let $X_1, \ldots, X_{100} \sim N(0,1)$. Find $P(|\overline{X}| > 0.1)$ where $\overline{X} = \frac{1}{100}\sum_{i=1}^{100} X_i$.

(b) Let $X_1, \ldots, X_n$ be i.i.d $\sim$ Gamma(α, θ). Use MGFs to determine the distribution of $\overline{X} = \frac{1}{n}\sum_{i=1}^{n} X_i$.

(c) Let $X_1, \ldots, X_{100}$ be i.i.d $\sim$ Gamma$(10, 10)$. Use CLT to approximate $\mathbb{P}(95 \leq \overline{X} \leq 105)$.

(d) Let $X_1, \ldots, X_n$ be an i.i.d sample drawn from a population distribution D. Is there any relation between the sample mean and the population mean?

Problem 2 (Distribution of Sample Mean)

Let $X_1, \ldots, X_n$ be i.i.d $\sim$ N$(10, 100)$. How large does n need to be such that

$$P(|\overline{X} - 10| < 0.001) > 0.99 \ \ ?$$

Problem 3 (Sample Mean and Variances as Estimators)
The strength of chess players is measured by the so-called Elo rating. From a group of 1000 chess players, the Elo ratings of 10 players are sampled with the following results $x_1, \ldots, x_{10}$.

$$1628, 1534, 1630, 1634, 1789, 1443, 1584, 1353, 1559, 1430$$

We assume that the population distribution is $N(\mu, \sigma^2)$ with unknown μ and σ.

(a) Compute the sample mean $\overline{x} = \frac{1}{10}\sum_{i=1}^{10} x_i$ and and sample variance $s^2 = \frac{1}{9}\sum_{i=1}^{10}(x_i - \bar{x})^2$ based the the observations $x_1, \ldots, x_{10}$.

(b) The sample mean can be used to approximate μ and the sample variance to approximate σ^2. Based on this, find the approximate probability that a randomly chosen player from the group of 1000 players has an Elo rating higher than 1700.

Problem 4 (Function of Sample Mean as Estimator)
Let X_i be the time between the ith and $(i+1)$th eruption of a volcano and assume that $X_1, \ldots, X_n$ are i.i.d $\sim Exp(\theta)$, where θ is unknown. The following observations $x_1, \ldots, x_{10}$ (time in years) have been made.

$$1343.4, 1753.2, 1569.8, 645.4, 2617.0, 3897.0, 348.7, 3017.3, 2197.3, 245.1$$

Write $\overline{x} = \frac{1}{10}\sum_{i=1}^{10} x_i$.

Explain why $\overline{x}$ is a reasonable estimation of θ.

Problem 5 (Use of CLT for approximation)
Let $Y = X_1 + X_2 + \cdots + X_{15}$ be the sum of i.i.d random sample of size 15 from the distribution whose PDF is

$$f(x) = \frac{3}{2}x^2, \quad -1 < x < 1; \quad f(x) = 0 \ \text{ elsewhere.}$$

Use the central limit theorem to approximate the probability $\mathbb{P}(-0.3 \leq Y \leq 1.5)$.

Answer Keys.

Q1(a) 0.318 **Q1(b)** $Gamma\left(n\alpha, \frac{\theta}{n}\right)$ **Q1(c)** 0.8858 **2** 6.636×10^8 **3(a)** $\overline{x} = 1558.4,\ s^2 = 15789.6$ **3(b)** 0.13 **5** 0.2313