Name: ____________________________ PUID ____________________________

STAT 350 Worksheet #12

Point Estimators: Unbiased Estimation and Estimator Properties

In previous lessons, we studied *sampling distributions* and how they describe the variability of statistics computed from samples. We focused on scenarios involving Normal distributions or approximate Normality via the Central Limit Theorem (CLT). Now, we turn our attention specifically to **estimating unknown parameters** using sample data.

When analyzing data, we often want to estimate unknown numerical characteristics (parameters) of a population. Examples of parameters include:

- Population mean $(\boldsymbol{\mu})$
- Population variance $(\boldsymbol{\sigma}^2)$
- Probability of success $(\boldsymbol{p})$
- Rate parameter of a Poisson or Exponential distribution $(\boldsymbol{\lambda})$

A **point estimator** is a rule or formula that uses the sample data to produce a single "best guess" for the unknown parameter. Formally:

- A **parameter** $(\boldsymbol{\theta})$ is a fixed, unknown number describing a population characteristic.
- A **point estimator** $(\widehat{\boldsymbol{\theta}})$ is a statistic computed from a sample, intended to approximate $\boldsymbol{\theta}$.

For example, if we wish to estimate **the population mean** $(\boldsymbol{\mu})$**,** a natural estimator is that of the **sample mean** $(\overline{\boldsymbol{X}})$. Similarly, to estimate **a probability of success** $(\boldsymbol{p})$**,** the **sample proportion of successes** $(\widehat{\boldsymbol{p}})$ can be used as an estimator.

One desirable property of an estimator is that it produces estimates which are correct *on average*. This leads to the concept of **unbiasedness**:

- The **bias** of an **estimator** $\widehat{\boldsymbol{\theta}}$ is measured as:
$$\textbf{bias}(\widehat{\boldsymbol{\theta}}) = \boldsymbol{E}[\widehat{\boldsymbol{\theta}}] - \boldsymbol{\theta}$$
- If an estimator satisfies:
$$\boldsymbol{E}[\widehat{\boldsymbol{\theta}}] = \boldsymbol{\theta}$$
it is called an **unbiased estimator** of $\boldsymbol{\theta}$. In other words, the **estimator** hits the **true parameter** on average over many repeated samples.
- If an estimator is not **unbiased** it is said to be **biased**.

An **unbiased estimator** does not consistently underestimate or overestimate the **parameter** it targets. Instead, the estimation errors "balance out" across many samples.

1. Consider a random variable $\boldsymbol{X} \sim \textbf{Exponential}(\boldsymbol{\lambda})$, where $\boldsymbol{\lambda} > \mathbf{0}$ is the rate parameter. Recall that the population mean $\boldsymbol{\mu} = \boldsymbol{E}[\boldsymbol{X}] = \frac{\mathbf{1}}{\boldsymbol{\lambda}}$.
Suppose we take $\boldsymbol{n}$ **independent random samples** from this distribution, i.e., $\boldsymbol{X}_1, \boldsymbol{X}_2, \dots, \boldsymbol{X}_n \overset{\text{i.i.d.}}{\sim} \textbf{Exponential}(\boldsymbol{\lambda})$.
We wish to estimate the parameter $\boldsymbol{\mu}$ (the mean) and the parameter $\boldsymbol{\lambda}$ (the rate) using these samples.

a) Show mathematically that the sample mean $\overline{\boldsymbol{X}}$ is an unbiased estimator for $\boldsymbol{\mu} = \mathbf{1}/\boldsymbol{\lambda}$. In other words, show that the $\textbf{bias}(\overline{\boldsymbol{X}}) = \mathbf{0}$**.**

b) Show by simulation that the sample mean $\overline{\boldsymbol{X}}$ is an unbiased estimator for $\boldsymbol{\mu} = \mathbf{1}/\boldsymbol{\lambda}$ when $\boldsymbol{\lambda} = \mathbf{10}$. In other words, option 1500 random samples of size $\boldsymbol{n} = \mathbf{25}$ from $\textbf{Exponential}(\boldsymbol{\lambda} = \mathbf{10})$ and compute 1500 sample means. Compute and report the following summary statistics of the sample means:

- **i.** The average of the sample means
- **ii.** The standard deviation of the sample means
- **iii.** The proportion of sample means that exceed the true mean $\boldsymbol{\mu} = \mathbf{1}/\mathbf{10}$.
- **iv.** The proportion of sample means that are at most $\boldsymbol{\mu} = \mathbf{1}/\mathbf{10}$**.**

c) Using your summary statistics do you think that the sample means were ***close*** to the expected value $\mu = 1/10$ and also does the sample means appear to be unbiased?

d) Next, suppose instead we want to estimate the rate parameter $\boldsymbol{\lambda}$. Consider the estimator $\hat{\lambda} = \frac{1}{\bar{X}}$ using your simulation data from part **b)** compute $\hat{\lambda} = \frac{1}{\bar{X}}$. Based on your results, is $\hat{\lambda}$ likely unbiased as an estimator of λ? If you think the estimator is biased, approximate that bias using your simulated data. Repeat the bias approximation for different values of $\boldsymbol{n}$. Do you observe that the bias remains roughly constant, or does it change with $\boldsymbol{n}$?

2. Suppose we have $\boldsymbol{n}$ **i.i.d.** random variables from a Uniform distribution, i.e., $X_1, X_2, \dots, X_n \overset{\text{i.i.d.}}{\sim} \textbf{Uniform}(\mathbf{0}, \boldsymbol{\theta})$, where $\boldsymbol{\theta} > \mathbf{0}$ is an unknown parameter representing the largest value that the random variables can possibly take on. Suppose we are able to obtain samples from this population and wish to estimate $\boldsymbol{\theta}$. What estimator would you choose? A natural guess might be the sample maximum:

$$\mathcal{M} = \max\{X_1, X_2, \dots, X_n\}.$$

In this exercise, you will show that this estimator is biased but can be corrected to be an unbiased estimator.

a) Show that $\mathcal{M}$ is an unbiased estimator for $\boldsymbol{\theta}$. We will work through deriving the distribution function for $\mathcal{M}$.

- **Step 1:** Determine the cumulative distribution function for $\mathcal{M}$. Since we know that each X_i are independent and follow the same **Uniform distribution** we can easily derive the **cumulative distribution function**:

$$F_{\mathcal{M}}(x) = P(\mathcal{M} \leq x)$$

Think carefully: the event $\{\mathcal{M} \leq x\}$ means that the maximum of the sample is at most x. In other words, if the largest is at most x, then all of the observations must be at most x:

$$P(\mathcal{M} \leq x) = P(\{X_1 \leq x\} \cap \{X_2 \leq x\} \cap \dots \cap \{X_n \leq x\}).$$

Using the independence of the random variables and the known CDF of the $\textbf{Uniform}(\mathbf{0}, \boldsymbol{\theta})$ distribution determine and clearly write out the full piecewise function for **CDF** of $\mathcal{M}$.

- **Step 2**: After you have obtained the CDF for $\mathcal{M}$, determine the corresponding probability density function $f_{\mathcal{M}}(x)$. The probability density function will simply be a piece-wise function where it is 0 everywhere except within the interval $(\mathbf{0}, \boldsymbol{\theta})$. To determine the functional form for the PDF in this region $(\mathbf{0}, \boldsymbol{\theta})$, we simply need to realize that the PDF can be obtained by taking the derivative of the CDF with respect to x.

$$f_{\mathcal{M}}(x) = \frac{d}{dx} F_{\mathcal{M}}(x).$$

Obtain the probability density function $f_{\mathcal{M}}(x)$ for the random variable $\mathcal{M}$ by taking the derivative of the CDF with respect for the region $(\mathbf{0}, \boldsymbol{\theta})$. Be sure to write out the full piece wise function.

- **Step 3:** Using the probability density function, calculate the **expected value** of the estimator $\mathcal{M}$.

- **Step 4**: Using your result from **Step 3**, Determine the bias of your estimator $\mathcal{M}$:

$$\textbf{bias}(\mathcal{M}) =$$

b) Since the estimator $\mathcal{M}$ is biased, propose a simple modification to $\mathcal{M}$ that results in an unbiased estimator, and denote this estimator as $\hat{\theta}$.

$$\hat{\boldsymbol{\theta}} =$$

Minimum Variance Unbiased Estimators

However, there could be multiple unbiased estimators for the same parameter. Among all unbiased estimators, we typically prefer the one with the smallest variability (variance).
Such an estimator is called the **Minimum Variance Unbiased Estimator (MVUE)**. The MVUE is the "best" unbiased estimator in the sense that it provides estimates consistently closest to the true parameter, on average.
Formally (but intuitively), an estimator $\hat{\theta}$ is **MVUE** if:

- It is **unbiased**, meaning $E[\hat{\theta}] = \theta$
- It has the **smallest variance** among all possible unbiased estimators for θ.

Although deriving an MVUE theoretically involves advanced methods beyond this course, we can illustrate the concept clearly through simulation.

3. Suppose you collect data from a population that follows a Normal distribution with mean μ and variance σ^2. Specifically, consider the population distribution:

$$X \sim N(\mu = 50, \sigma^2 = 25),$$

Two natural and commonly used estimators for the population mean μ are:

- Estimator A: the sample mean $\bar{X}$.
- Estimator B: the sample median $\tilde{X}$.

Both estimators intuitively seem plausible, and indeed both are unbiased for the mean

a) Generate 2000 independent samples, each of size $n = 15$, from the distribution $N(\mu = 50, \sigma^2 = 25)$. Compute the sample mean and median for each of the 2000 independent samples. Compute and report the average and standard deviation of each of 2000 estimates. Were they both close to the true mean $\mu = 50$ on average? Which estimator had less variability?

b) In practice we will never have access to 2000 samples instead we will only have a single sample of size n. If you have a good reason to assume that the population you sampled from is Normal or nearly Normal, which estimator would you prefer for the central tendencies. Use your exploration in a) to justify your answer.

Note: While unbiased estimators are generally desirable, sometimes an estimator with slight bias but much smaller variability can provide better estimates overall. In practice, small bias may be an acceptable tradeoff for reduced uncertainty.