Lecture 15 Worksheet
Chrysafis Vogiatzis

Every worksheet will work as follows.
1. You will be entered into a Zoom breakout session with other students in the class.
2. Read through the worksheet, discussing any questions with the other participants in your breakout session.
   - You can call me using the “Ask for help” button.
   - Keep in mind that I will be going through all rooms during the session so it might take me a while to get to you.
3. Answer each question (preferably in the order provided) to the best of your knowledge.
4. While collaboration between students in a breakout session is highly encouraged and expected, each student has to submit their own version.
5. You will have 24 hours (see Compass) to submit your work.

Worksheet 1: Biases

You are interested in estimating the (unknown) mean $\mu$ of a population $X$. You have been able to collect only a sample of $n = 2$ observations, so you are worried about your estimating the mean. You are already aware of one good estimator: take the average of the 2 elements and use that as a proxy of the unknown mean $\mu$.

However, a friend of yours tells you about this revolutionary technique they read about online! First, flip a fair coin. If it comes up Heads (with probability 50%) take the first element $X_1$ and report that the mean is actually $\frac{3X_1}{2}$. If the coin comes up Tails (with probability 50%) take the first two elements $X_1, X_2$ and report that the mean is $\frac{X_1 + 2X_2}{6}$.

Problem 1: $\hat{\Theta}_1 = \frac{X_1 + X_2}{2}$

Let $\hat{\Theta}_1$ be equal to $\frac{X_1 + X_2}{2}$ (the sample average). What is the estimator’s bias? 1

\[ \text{Answer to Problem 1.} \]

1 Recall that because $X_1, X_2$ have come from the population $X$ you know that
\[ E [X_1] = E [X_2] = E [X] = \mu \]
\[ \text{Var} [X_1] = \text{Var} [X_2] = \text{Var} [X] = \sigma^2. \]
You will not need the variance in this question, but you may need it later!
**Problem 2:** \( \hat{\Theta}_2 = \frac{3X_1}{2} \)

Let \( \hat{\Theta}_2 \) be equal to \( \frac{3X_1}{2} \) (the weird estimator your friend recommended if the coin comes up Heads). What is its bias?

**Answer to Problem 2.**

**Problem 3:** \( \hat{\Theta}_3 = \frac{X_1 + 2X_2}{6} \)

Let \( \hat{\Theta}_3 \) be equal to \( \frac{X_1 + 2X_2}{6} \) (the other weird estimator your friend recommended if the coin comes up Tails). What is its bias?

**Answer to Problem 3.**

**Problem 4:** *Bias is a weird thing*

Based on your answers in Problems 3 and 4, what is the bias of the technique your friend is recommending?

**Answer to Problem 4.**

\(^2\) Let us revisit what the law of total expectation states for two mutually exclusive events \( A, \overline{A} \):

\[
E[X] = E[X|A] \cdot P(A) + E[X|\overline{A}] \cdot P(\overline{A}).
\]
Worksheet 2: Weird point estimators

Assume that a population is distributed with pdf \( f(x) = c(1 + \theta x) \), 
\(-1 \leq x \leq 1\), where \( \theta \) is an unknown parameter, and \( c \) a constant.  

Problem 5: Back to basics

Let’s return to the basics for a second! What should \( c \) be equal to in order for \( f(x) \) to be a valid continuous pdf?

Answer to Problem 5.

Problem 6: Where did you come up with this?

Assume you obtain a sample of \( n \) observations. Consider the sample average \( \bar{X} = (X_1 + X_2 + \ldots + X_n) / n \). Show that \( \hat{\Theta} = 3\bar{X} \) is an unbiased estimator for \( \theta \).

Answer to Problem 6.
Problem 7: Variance and standard error

What is the standard error of the point estimator $\hat{\Theta} = 3X$?

To calculate this you will first need to calculate the expectation and the variance of population $X$. They could very well be a function of $\theta$ as you do not know what the parameter is equal to...

Answer to Problem 7.
Worksheet 3: Comparing point estimators

Assume we have collected a sample of \( n = 3 \) observations \( X_1, X_2, X_3 \) coming from a population \( X \) distributed with some pdf with unknown \( \mu \) and known \( \sigma^2 = 16 \). We have devised three point estimators for the unknown population mean:

- Get the average from the first two observations omitting the third, i.e.,
  \[ \hat{\Theta}_1 = \frac{X_1 + X_2}{2}. \]

- Add the “odd” observations once and the “even” observations doubled and divide everything by 4, i.e.,
  \[ \hat{\Theta}_2 = \frac{X_1 + 2X_2 + X_3}{4}. \]

- Once again omit the third observation and simply add the first two and divide by 4, i.e.,
  \[ \hat{\Theta}_3 = \frac{X_1 + X_2}{4}. \]

Problem 8: Comparison I

For one last time in this worksheet, calculate the bias and variance of each of the estimators.

Answer to Problem 8.
Problem 9: Comparison II

What is the MSE of each of the estimators? Which estimator is the best according to its MSE, if we have been told that \( \mu > 4 \)?

Answer to Problem 9.

Problem 10: Observation

In Problem 8, you must have gotten that the first two estimators are unbiased (i.e., zero bias). In general, assume you are collecting a sample of \( n \) observations \((X_1, X_2, \ldots, X_n)\) and are using \( \hat{\Theta} = a_1X_1 + a_2X_2 + \ldots + a_nX_n \) to estimate the unknown mean. What condition should \( a_1 + a_2 + \ldots + a_n \) satisfy in order for \( \hat{\Theta} \) to have bias equal to zero?

\[ E \left[ \sum_{i=1}^{n} a_iX_i \right] \]

Additionally, never forget that \( E[X_1] = E[X_2] = \ldots = E[X_n] = E[X] \) because all observations come from the same population \( X \)!