DISCUSSION WORKSHEET 04

PSTAT 120B: Mathematical Statistics, I
Summer Session A, 2024 with Instructor: Ethan P. Marzban

Conceptual Review

(a) What are the main goals of **inferential statistics**?
(b) What is a **population**? What is a **sample**? What is the difference between a sample and an **observed instance** (aka **realization**) of a sample?
(c) What is a **statistic**? When does a statistic become an **estimator**?
(d) What is the notion of a **sampling distribution**? How does that tie into the notions of **bias** and **unbiasedness**?
(e) What is the **mean square error** of an estimator?

Problem 1: Exercise 8.12 from the Textbook

The reading on a voltage meter connected to a test circuit is uniformly distributed over the interval $(\theta\ ,\ \theta+1)$, where θ is the true but unknown voltage of the circuit. Suppose that $Y_1, Y_2, \cdots Y_n$ denote a random sample of such readings.

(a) Show that $\overline{Y}_n$ is a biased estimator of θ and compute the bias.

> **Solution:** We have seen (several times) that the expected value of the sample mean is the population mean. Hence,
>
> $$\mathbb{E}[\overline{Y}_n] = \mathbb{E}[Y_1] = \frac{\theta + \theta + 1}{2} = \frac{2\theta + 1}{2} = \theta + \frac{1}{2}$$
>
> Since this is (in general) not equal to 0, we have shown that $\overline{Y}_n$ is a biased estimator of θ. Additionally, by the definition of bias,
>
> $$\text{Bias}(\overline{Y}_n\ ,\ \theta) := \mathbb{E}[\overline{Y}_n] - \theta = \theta + \frac{1}{2} - \theta = \boxed{\frac{1}{2}}$$

(b) Find a function of $\overline{Y}_n$ that is an unbiased estimator of θ.

> **Solution:** Since $\overline{Y}_n$ (on average) overestimates θ by a constant factor of $1/2$, it seems plausible to construct an unbiased estimator for θ by taking $\overline{Y}_n$ and subtracting $1/2$. That is, define the following estimator for θ:
>
> $$\boxed{\widehat{\theta}_n := \overline{Y}_n - \frac{1}{2}}$$
>
> Then,
>
> $$\mathbb{E}[\widehat{\theta}_n] = \mathbb{E}\left[\overline{Y}_n - \frac{1}{2}\right] = \mathbb{E}[\overline{Y}_n] - \frac{1}{2} = \frac{\theta\ 1}{+\ 2} - \frac{1}{2} = \theta$$

thereby showing that $\widehat{\theta}_n := \overline{Y}_n - \frac{1}{2}$ is an unbiased estimator of θ.

(c) Find MSE$(\overline{Y}_n\ ,\ \theta)$.

Solution: By the Bias-Variance Decomposition,

$$\text{MSE}(\overline{Y}_n\ ,\ \theta) = \left[\text{Bias}(\overline{Y}_n\ ,\ \theta)\right]^2 + \text{Var}\left(\overline{Y}_n\right)$$

We computed Bias$(\overline{Y}_n\ ,\ \theta)$ in part (a); additionally, we know that

$$\text{Var}(\overline{Y}_n) = \frac{\text{Var}(Y_1)}{n} = \frac{1}{12n}(\theta + 1 - \theta)^2 = \frac{1}{12n}$$

Therefore, putting everything together,

$$\begin{aligned} \text{MSE}(\overline{Y}_n\ ,\ \theta) &= \left[\text{Bias}(\overline{Y}_n\ ,\ \theta)\right]^2 + \text{Var}\left(\overline{Y}_n\right) \\ &= \left(\frac{1}{2}\right)^2 + \left(\frac{1}{12n}\right) = \boxed{\frac{3n+1}{12n}} \end{aligned}$$

©Ethan P. Marzban, 2024