

Middle East Technical University
Department of Economics
Econ 206:
Statistics for Economists-II

Güneş A. Aşık

November 5, 2025

Recitation

Question 1: Sampling Distribution

Consider a population with the following information: Population size: $N = 4$, X : age of individuals; Population values: 18, 20, 22, 24 (years)

1. What is the distribution for the population?
2. Find the mean and variance for this population.
3. Assume that we select a sample from this population with $n = 2$. Construct a table for all possible samples of size $n = 2$ with replacement. What is the mean and variance for the sample?
4. Construct a table for all possible sample means based on $n = 2$ and plot a histogram.

Solution:

The population has a uniform distribution with mean, $\mu = 21$ and variance, σ^2 is 5. Sample mean $\bar{X} = 21$ and $Var(\bar{X}) = \sigma^2/n = 2.5$

Since we are sampling with replacement from a finite population, the number of possible samples is:

$$N^n = 4^2 = 16$$

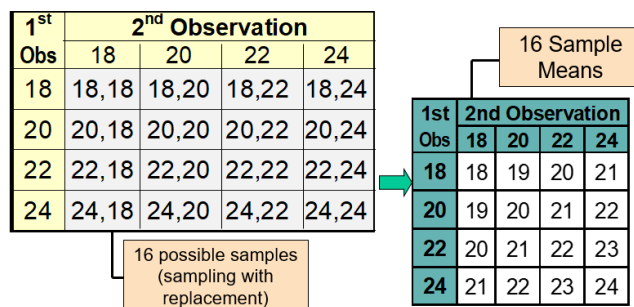


Figure 1: Sampling Distribution

Sampling Distribution of All Sample Means

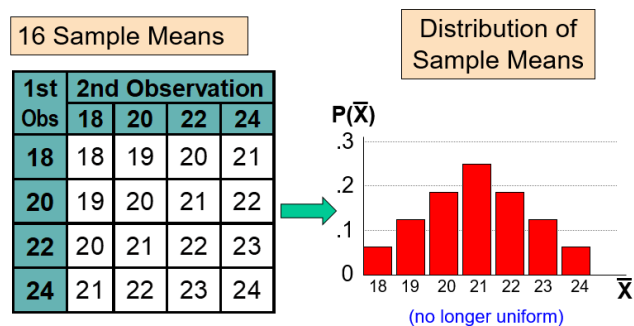


Figure 2: Distribution of Sample Means

Question 2

Let X_1 and X_2 be two random observations from a population with mean μ and variance σ^2 . Consider the following point estimators of μ :

$$\hat{\mu}_1 = \frac{1}{2}X_1 + \frac{1}{2}X_2, \quad \hat{\mu}_2 = \frac{1}{4}X_1 + \frac{3}{4}X_2, \quad \hat{\mu}_3 = \frac{1}{3}X_1 + \frac{2}{3}X_2$$

1. Show that each estimator is an unbiased estimator of μ .
2. Which one is the most efficient estimator?

Solution:

Part 1) - Unbiasedness

An estimator $\hat{\mu}$ is unbiased if $E[\hat{\mu}] = \mu$.

Since $E[X_1] = E[X_2] = \mu$, we have:

$$\begin{aligned}
E[\hat{\mu}_1] &= E\left[\frac{1}{2}X_1 + \frac{1}{2}X_2\right] = \frac{1}{2}\mu + \frac{1}{2}\mu = \mu \\
E[\hat{\mu}_2] &= E\left[\frac{1}{4}X_1 + \frac{3}{4}X_2\right] = \frac{1}{4}\mu + \frac{3}{4}\mu = \mu \\
E[\hat{\mu}_3] &= E\left[\frac{1}{3}X_1 + \frac{2}{3}X_2\right] = \frac{1}{3}\mu + \frac{2}{3}\mu = \mu
\end{aligned}$$

All three estimators are unbiased.

Part 2) - Most Efficient Estimator

The variance of each estimator (assuming X_1 and X_2 are independent):

$$\begin{aligned}
\text{Var}(\hat{\mu}_1) &= \text{Var}\left(\frac{1}{2}X_1 + \frac{1}{2}X_2\right) = \left(\frac{1}{2}\right)^2 \sigma^2 + \left(\frac{1}{2}\right)^2 \sigma^2 = \frac{1}{4}\sigma^2 + \frac{1}{4}\sigma^2 = \frac{1}{2}\sigma^2 \\
\text{Var}(\hat{\mu}_2) &= \text{Var}\left(\frac{1}{4}X_1 + \frac{3}{4}X_2\right) = \left(\frac{1}{4}\right)^2 \sigma^2 + \left(\frac{3}{4}\right)^2 \sigma^2 = \frac{1}{16}\sigma^2 + \frac{9}{16}\sigma^2 = \frac{10}{16}\sigma^2 = \frac{5}{8}\sigma^2 \\
\text{Var}(\hat{\mu}_3) &= \text{Var}\left(\frac{1}{3}X_1 + \frac{2}{3}X_2\right) = \left(\frac{1}{3}\right)^2 \sigma^2 + \left(\frac{2}{3}\right)^2 \sigma^2 = \frac{1}{9}\sigma^2 + \frac{4}{9}\sigma^2 = \frac{5}{9}\sigma^2
\end{aligned}$$

Comparing the variances:

$$\text{Var}(\hat{\mu}_1) = \frac{1}{2}\sigma^2 = 0.5\sigma^2, \quad \text{Var}(\hat{\mu}_2) = \frac{5}{8}\sigma^2 = 0.625\sigma^2, \quad \text{Var}(\hat{\mu}_3) = \frac{5}{9}\sigma^2 \approx 0.556\sigma^2$$

Since $\text{Var}(\hat{\mu}_1) < \text{Var}(\hat{\mu}_3) < \text{Var}(\hat{\mu}_2)$, $\hat{\mu}_1$ has the smallest variance and is therefore the most efficient estimator.

Question 3

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 $Y_1, Y_2, \dots, Y_n$ denote a random sample of such readings.

- Show that $\bar{Y}$ is a biased estimator of θ and compute the bias.
- Find a function of $\bar{Y}$ that is an unbiased estimator of θ .

Solution

Part (a)

Each $Y_i \sim \text{Uniform}(\theta, \theta + 1)$, so:

$$E[Y_i] = \theta + \frac{1}{2}, \quad \text{Var}(Y_i) = \frac{1}{12}$$

The expected value of the sample mean is:

$$E[\bar{Y}] = E\left[\frac{1}{n} \sum_{i=1}^n Y_i\right] = \frac{1}{n} \sum_{i=1}^n E[Y_i] = \frac{1}{n} \sum_{i=1}^n \left(\theta + \frac{1}{2}\right) = \theta + \frac{1}{2}$$

The bias of $\bar{Y}$ is:

$$\text{Bias}(\bar{Y}) = E[\bar{Y}] - \theta = \left(\theta + \frac{1}{2}\right) - \theta = \frac{1}{2}$$

Therefore, $\bar{Y}$ is a biased estimator of θ with bias $+\frac{1}{2}$.

Part (b)

Since $E[\bar{Y}] = \theta + \frac{1}{2}$, we can define:

$$\hat{\theta} = \bar{Y} - \frac{1}{2}$$

Verifying unbiasedness:

$$E[\hat{\theta}] = E\left[\bar{Y} - \frac{1}{2}\right] = E[\bar{Y}] - \frac{1}{2} = \left(\theta + \frac{1}{2}\right) - \frac{1}{2} = \theta$$

Therefore, $\bar{Y} - \frac{1}{2}$ is an unbiased estimator of θ .

Question 3

Consider a population that is uniformly distributed over an interval $[0, \beta]$ where β is an unknown positive number that we would like to estimate. We take a sample of size 3 from this population and use the sample maximum as an estimate for the population maximum.

1. What is the PDF of this estimator, i.e., what is the sampling distribution of the sample maximum?
2. What is the expected value of this estimator?
3. Is this estimator an unbiased estimator for β ?
4. Suggest one more estimator for β which is unbiased.

Solution:

1) $\hat{\beta}$ is defined as:

$$\hat{\beta} = \max\{x_1, x_2, x_3\}$$

Where x_1, x_2 and x_3 are the sample observations. Representing each observation x_i as a random variable X_i :

$$\hat{\beta} = \max\{X_1, X_2, X_3\}$$

CDF of $\hat{\beta}$, $F_{\hat{\beta}}(t)$ is defined as a function of t as follows:

$$F_{\hat{\beta}}(t) = P(\hat{\beta} \leq t) = P(\max\{X_1, X_2, X_3\} \leq t).$$

Assuming that X_1, X_2 and X_3 are independent random variables,

$$\begin{aligned} P(\max\{X_1, X_2, X_3\} \leq t) &= P(X_1 \leq t \text{ and } X_2 \leq t \text{ and } X_3 \leq t) \\ &= P(X_1 \leq t) \cdot P(X_2 \leq t) \cdot P(X_3 \leq t) \end{aligned}$$

can be written (note how we used the statistical independence granted by the i.i.d. assumption here). Since X_1, X_2 and X_3 come from a $\text{Uniform}(0, \beta)$ distribution:

$$P(X_1 \leq t) = \frac{t}{\beta}, \quad P(X_2 \leq t) = \frac{t}{\beta}, \quad P(X_3 \leq t) = \frac{t}{\beta}$$

Then, $F_{\hat{\beta}}(t) = \frac{t^3}{\beta^3}$ is obtained. This is just the CDF but we want to find the PDF because we want to estimate the expected value. Therefore we differentiate the CDF to find the PDF:

$$f_{\hat{\beta}}(t) = \frac{dF_{\hat{\beta}}(t)}{dt} = \frac{3t^2}{\beta^3}.$$

2) We will now find the expected value as:

$$\begin{aligned} E(\hat{\beta}) &= \int_{-\infty}^{\infty} t f_{\hat{\beta}}(t) dt = \int_0^{\beta} t \cdot \frac{3t^2}{\beta^3} dt \\ &= \frac{3}{\beta^3} \int_0^{\beta} t^3 dt = \frac{3}{\beta^3} \left[\frac{t^4}{4} \right]_0^{\beta} = \frac{3}{4} \beta. \end{aligned}$$

3) As $E(\hat{\beta}) = \frac{3}{4}\beta < \beta$, $\hat{\beta}$ is not an unbiased estimator of β .

4) $\tilde{\beta} = \frac{4}{3} \max\{x_1, x_2, x_3\}$ is an unbiased estimator of β . Check why/how.

Question 4: Convergence in Probability of the Minimum to θ

Let $X_1, \dots, X_n$ be i.i.d. random variables with common pdf:

$$f(x) = \begin{cases} e^{-(x-\theta)} & x > \theta, \quad -\infty < \theta < \infty \\ 0 & \text{elsewhere} \end{cases}$$

This is the shifted exponential distribution. Let $Y_n = \min\{X_1, \dots, X_n\}$. We will prove that $Y_n \xrightarrow{P} \theta$.

Finding the CDF of Y_n

First, find the CDF of an individual X_i . For $x > \theta$:

$$F_X(x) = P(X_i \leq x) = \int_{\theta}^x e^{-(t-\theta)} dt = [-e^{-(t-\theta)}]_{\theta}^x = 1 - e^{-(x-\theta)}$$

So the complete CDF is:

$$F_X(x) = \begin{cases} 0 & x \leq \theta \\ 1 - e^{-(x-\theta)} & x > \theta \end{cases}$$

Now, for $Y_n = \min\{X_1, \dots, X_n\}$, the CDF is:

$$F_{Y_n}(y) = P(Y_n \leq y) = 1 - P(Y_n > y) = 1 - P(\min\{X_1, \dots, X_n\} > y)$$

Since the minimum exceeds y if and only if all X_i exceed y :

$$F_{Y_n}(y) = 1 - [P(X_1 > y)]^n = 1 - [1 - F_X(y)]^n$$

For $y \leq \theta$, $F_X(y) = 0$, so:

$$F_{Y_n}(y) = 1 - [1 - 0]^n = 0$$

For $y > \theta$, $F_X(y) = 1 - e^{-(y-\theta)}$, so:

$$F_{Y_n}(y) = 1 - [e^{-(y-\theta)}]^n = 1 - e^{-n(y-\theta)}$$

Therefore, the complete CDF of Y_n is:

$$F_{Y_n}(y) = \begin{cases} 0 & y \leq \theta \\ 1 - e^{-n(y-\theta)} & y > \theta \end{cases}$$

Proving $Y_n \xrightarrow{P} \theta$

To show $Y_n \xrightarrow{P} \theta$, we need to prove that for any $\epsilon > 0$:

$$\lim_{n \rightarrow \infty} P(|Y_n - \theta| \geq \epsilon) = 0$$

Note that since $Y_n = \min\{X_1, \dots, X_n\} > \theta$ almost surely, we have $Y_n - \theta > 0$. Therefore:

$$P(|Y_n - \theta| \geq \epsilon) = P(Y_n - \theta \geq \epsilon) = P(Y_n \geq \theta + \epsilon)$$

Using the CDF of Y_n :

$$P(Y_n \geq \theta + \epsilon) = 1 - F_{Y_n}(\theta + \epsilon) = 1 - [1 - e^{-n(\theta + \epsilon - \theta)}] = e^{-n\epsilon}$$

Therefore:

$$P(|Y_n - \theta| \geq \epsilon) = e^{-n\epsilon}$$

Taking the limit as $n \rightarrow \infty$:

$$\lim_{n \rightarrow \infty} P(|Y_n - \theta| \geq \epsilon) = \lim_{n \rightarrow \infty} e^{-n\epsilon} = 0$$

Since for any $\epsilon > 0$:

$$\lim_{n \rightarrow \infty} P(|Y_n - \theta| \geq \epsilon) = 0$$

we have shown that:

$$\boxed{Y_n \xrightarrow{P} \theta}$$

This proves that the minimum of a sample from the shifted exponential distribution converges in probability to the shift parameter θ .

Question 4

Problem: Let $Y_1, Y_2, \dots, Y_n$ constitute a random sample from the probability density function given by

$$f(y | \theta) = \begin{cases} \frac{2}{\theta^2}(\theta - y), & 0 \leq y \leq \theta, \\ 0, & \text{elsewhere.} \end{cases}$$

Find an estimator for θ by using the method of moments.

Solution:

Find the first population moment $E[Y]$ using direct integration

$$\begin{aligned} E[Y] &= \int_0^\theta y \cdot f(y | \theta) dy = \int_0^\theta y \cdot \frac{2}{\theta^2}(\theta - y) dy \\ &= \frac{2}{\theta^2} \int_0^\theta (\theta y - y^2) dy \\ &= \frac{2}{\theta^2} \left[\frac{\theta y^2}{2} - \frac{y^3}{3} \right]_0^\theta \\ &= \frac{2}{\theta^2} \left(\frac{\theta \cdot \theta^2}{2} - \frac{\theta^3}{3} \right) = \frac{2}{\theta^2} \left(\frac{\theta^3}{2} - \frac{\theta^3}{3} \right) \\ &= \frac{2}{\theta^2} \cdot \frac{\theta^3}{6} = \frac{\theta}{3} \end{aligned}$$

Now apply the Method of Moments:

The first sample moment is the sample mean:

$$m_1 = \bar{Y} = \frac{1}{n} \sum_{i=1}^n Y_i$$

Equate the population moment to the sample moment:

$$E[Y] = m_1 \quad \Rightarrow \quad \frac{\theta}{3} = \bar{Y}$$

Solve for the estimator:

$$\theta = 3\bar{Y}$$

Therefore, the method of moments estimator for θ is:

$$\boxed{\hat{\theta}_{\text{MM}} = 3\bar{Y}}$$

Question 5

Let $Y_1, Y_2, \dots, Y_n$ denote independent and identically distributed random variables from a power family distribution with parameters $\alpha > 0$, with probability density function:

$$f(y | \alpha) = \begin{cases} \frac{\alpha y^{\alpha-1}}{3^\alpha}, & 0 \leq y \leq 3, \\ 0, & \text{elsewhere.} \end{cases}$$

Given that $E(Y_1) = \frac{3\alpha}{\alpha+1}$, derive the method-of-moments estimator for α and determine whether it is consistent.

Solution

The first population moment is:

$$E[Y] = \frac{3\alpha}{\alpha+1}$$

Equating the sample mean to the population moment:

$$\bar{Y} = \frac{3\alpha}{\alpha+1}$$

Solving for α :

$$\begin{aligned} \bar{Y}(\alpha+1) &= 3\alpha \\ \bar{Y}\alpha + \bar{Y} &= 3\alpha \\ \bar{Y} &= 3\alpha - \bar{Y}\alpha \\ \bar{Y} &= \alpha(3 - \bar{Y}) \\ \alpha &= \frac{\bar{Y}}{3 - \bar{Y}} \end{aligned}$$

Thus, the method-of-moments estimator is:

$$\hat{\alpha} = \frac{\bar{Y}}{3 - \bar{Y}}, \quad \text{where } \bar{Y} = \frac{1}{n} \sum_{i=1}^n Y_i$$

Consistency of $\hat{\alpha}$:

To prove consistency, we need to show that $\hat{\alpha} \xrightarrow{P} \alpha$ as $n \rightarrow \infty$.

By the Law of Large Numbers:

$$\bar{Y} \xrightarrow{P} E[Y] = \frac{3\alpha}{\alpha+1}$$

Define the function $g(m) = \frac{m}{3-m}$, which is continuous for $m \neq 3$. Since $E[Y] = \frac{3\alpha}{\alpha+1} < 3$ for all $\alpha > 0$, g is continuous at $E[Y]$.

By the Continuous Mapping Theorem:

$$\hat{\alpha} = g(\bar{Y}) \xrightarrow{P} g(E[Y]) = \frac{\frac{3\alpha}{\alpha+1}}{3 - \frac{3\alpha}{\alpha+1}}$$

Simplifying:

$$g(E[Y]) = \frac{\frac{3\alpha}{\alpha+1}}{\frac{3(\alpha+1)-3\alpha}{\alpha+1}} = \frac{\frac{3\alpha}{\alpha+1}}{\frac{3}{\alpha+1}} = \alpha$$

Therefore:

$$\hat{\alpha} \xrightarrow{P} \alpha$$

We conclude that $\hat{\alpha}$ is a consistent estimator for α .