ISE 315: Engineering Statistics

Lecture 2: Review of Estimation

Instructor: Mansur M. Arief, PhD
Industrial and Systems Engineering, KFUPM

Office: 22-219 — Email: mansur.arief@kfupm.edu.sa

Based on Montgomery & Runger, Applied Statistics and Probability for Engineers, 6th Ed.

Lecture 2

Review of Estimation

Lecture 2 Outline: (Chapter 7.1-7.4)

- Point estimation (7.1)
- Sampling distributions (7.2)
- General concepts of point estimation (unbiased, standard error) (7.3) and Central Limit Theorem (7.2)
- (Next week) Methods of point estimation (method of moments, maximum likelihood) (7.4)

Chapter 7 Concept Review

Probability (ISE 205): Known $\rightarrow$ Observed

Population
(Parameters: μ, σ, p)

Sample
(Statistics: $\bar{x}, s, \hat{p}$)

Statistical Inference (ISE 315): Observed $\rightarrow$ Unknown

Given a random sample $X_1, X_2, \cdots, X_n$

Parameter Estimation
"What is μ?"

Hypothesis Testing
"Is $\mu = 50$?"

Point estimate: $\hat{\mu} = \bar{x}$
Interval estimate: $\bar{x} \pm t \cdot \frac{s}{\sqrt{n}}$

$H_0 : \mu = 50$ vs $H_1 : \mu \neq 50$
Decision: Reject or Fail to Reject

Key Terminologies

1. **Random Sample (of size n):**
   - $X_1, X_2, \ldots, X_n$ are **independent**
   - Each X_i has the **same distribution**
   - Observed values: $x_1, x_2, \ldots, x_n$
2. **Statistic:**
   - Any function of sample observations
   - E.g: $\bar{X}$ ("X bar"), S^2, mode, etc.
   - A statistic is a **random variable**
3. **Point Estimator:**
   - A statistic $\hat{\theta}$ ("theta hat") used to estimate an unknown parameter θ ("theta")

     $$\hat{\theta} = f(x_1, x_2, \ldots, x_n)$$

   - $\hat{\mu} = \bar{X}$ (sample average) estimates μ ("mu")
   - $\hat{\sigma}^2 = S^2$ (sample variance) estimates σ^2 ("sigma squared")

Key Terminologies

4. Sampling Distribution:

- Probability distribution of a statistic
  - Not about the population distribution of X_i's
  - But the sampling distribution (of the statistic), e.g $\bar{X}$
- Why do statistics vary?
  - Because different set of samples yield different statistics
- What affects the properties of the sampling distribution?
  - Depends on the statistic formula, sample size n, etc.

Different samples yield different estimates—this variability is what we study!

Key Terminologies

5. Properties of Estimators

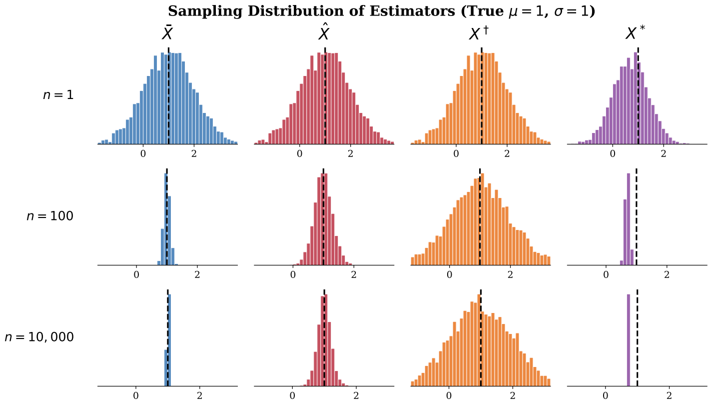
- **Biased/unbiased:** Is the mean **equal** to the true parameter?
- **Low/high bias:** How much is it **off** from the true parameter?
- **Low/high variance:** How much does it **vary** from sample to sample?
- **Rate of shrinkage:** How **quickly** does it converge as n increases?
- **Low/high Mean Squared Error (MSE):** Combines **both** variance & bias

$$MSE = \mathbb{E}[(\hat{\theta} - \theta)^2] = \mathrm{Var}(\hat{\theta}) + \mathrm{Bias}^2$$

Ideal estimator: Low MSE (low variance and low bias)

Some estimators are better than others

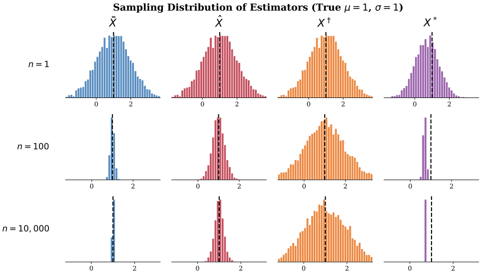
- Consider estimators $\bar{X}$, $\hat{X}$, $X^\dagger$, X^* for population mean μ:



Which one would you choose and why?

Some statistics are better than others

- What you have seen is the sampling distribution of
  - Sample mean: $\bar{X} = \frac{1}{n}\sum_{i=1}^{n} X_i$
  - Half-range: $\hat{X} = \frac{\max(X_i)+\min(X_i)}{2}$
  - First sample: $X^{\dagger} = X_1$
  - Shrinkage: $X^{*} = 0.7\bar{X}$
- Which one(s) are unbiased?
- Which one(s) has smallest variance?
- Which one has lowest MSE for $n = 1$? For larger n?

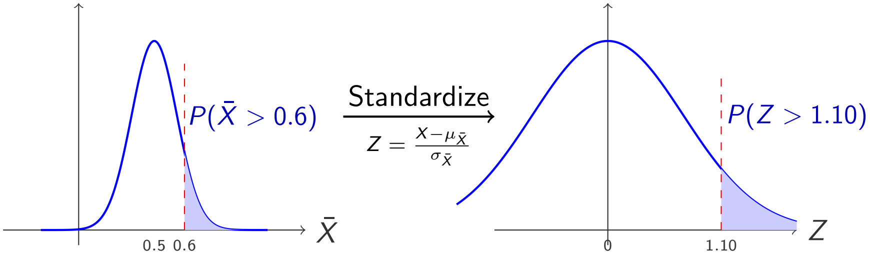


Notes about Estimator Properties

- About **sampling distribution**, NOT **population distribution**!
- The sampling dist **varies** with samples, but the population dist is **fixed**
- Example: Population is $X_i \sim Uniform(0,1)$
  - Population mean $\mu = \frac{0+1}{2} = 0.5$
  - Population variance $\sigma^2 = \frac{(1-0)^2}{12} = 1/12$
  - From ISE 205 (mean and var of common distributions)
- Sampling distribution of the sample average $\bar{X}$ for $n = 10$ is
  - Mean $\mu_{\bar{X}} = \mu = 0.5$ (regardless of n – **unbiased**)
  - Variance $Var(\bar{X}) = \frac{\sigma^2}{n} = \frac{1/12}{10} = 1/120$ (**shrinks** as n increases)
  - Standard deviation $\sigma_{\bar{X}} = \sqrt{Var(\bar{X})} = \sqrt{1/120} \approx 0.091$ (**also shrinks** in n)
  - By CLT, $\bar{X}$ is **loosely approximately** $N(\mu_{\bar{X}}, \sigma_{\bar{X}}) = N(0.5, 0.091)$
  - Use Normal distribution properties for analysis (e.g., finding $P(\bar{X} > 0.6)$)

Review: Computing Probabilities for Normal Distribution

Problem: Find $P(\bar{X} > 0.6)$ where $\bar{X} \sim N(0.5, 0.091)$



Method 1: Direct
$\bar{X} \sim N(0.5, 0.091)$

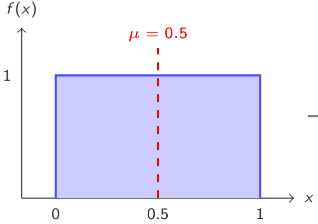
Method 2: Standardized
$Z \sim N(0, 1)$

Threshold for Z is $\frac{0.6-0.5}{0.091} = \frac{0.1}{0.091} \approx 1.10 \Rightarrow P(\bar{X} > 0.6) = P(Z > 1.10) \approx 0.136$

Population vs Sampling Distribution

Population Distribution

$X_i \sim \text{Uniform}(0, 1)$



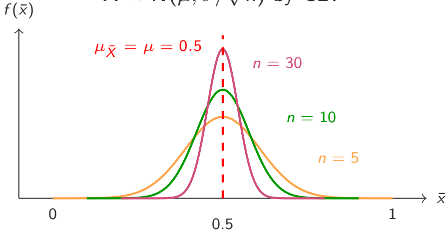
$\mu = 0.5, \quad \sigma^2 = 1/12$

Fixed shape (doesn't change)

Take many samples of size n

Sampling Distribution of $\bar{X}$

$\bar{X} \approx N(\mu, \sigma/\sqrt{n})$ by CLT



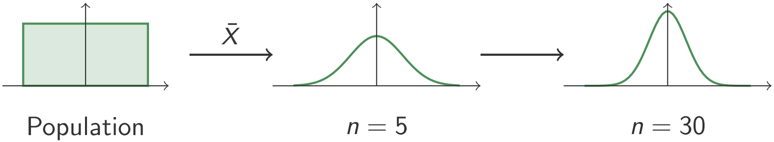
$\mu_{\bar{X}} = \mu, \quad \sigma^2_{\bar{X}} = \sigma^2/n$

Shape depends on n (gets narrower as $n \uparrow$)

Both have different distributions: **same mean**, but **different shape** and **variance**

Central Limit Theorem (CLT)

Visual illustration



As $n \uparrow$, distribution of $\bar{X}$ becomes Normal and more concentrated around μ.

Central Limit Theorem (CLT)

If $X_1, X_2, \ldots, X_n$ is a random sample of size n taken from a population with mean μ and variance σ^2, and if $\bar{X}$ is the sample mean

$$\bar{X} = \frac{1}{n}\sum_{i=1}^{n} X_i,$$

then the limiting form of the sampling distribution $\bar{X}$ is $N(\mu, \sigma/\sqrt{n})$.

Key insight: Regardless of the population distribution, $\bar{X}$ is approximately Normal for large n!

CLT for Z-statistic

Using the properties of the normal distribution, we can standardize $\bar{X}$ and calculate the Z-statistic:

$$Z = \frac{\bar{X} - \mu}{\sigma/\sqrt{n}}.$$

For large n, the Z-statistic has a sampling distribution of the Standard Normal $N(0, 1)$.

Key insight: Regardless of the population distribution, Z is approximately standard Normal for large n!

Summary #1

- The random variable $X_1, X_2, \ldots, X_n$ is a **random sample** of size n if:
  - The X_i's are **independent** random variables
  - Every X_i has the **same probability distribution**
- A statistic (also called point estimator) is a **function** of the realizations of the random variable $X_1, X_2, \ldots, X_n$
- A **point estimate** of some population parameter θ is a single numerical value $\hat{\theta}$ of a statistic $\hat{\Theta}$
- The probability distribution of a statistic is called the **sampling distribution**
- The **Central Limit Theorem (CLT)** is a theory that the sampling distribution of $\bar{X} \approx$ Normal for large n

Questions you might be asked about sampling distribution

- Given a population distribution, what is the sampling distribution of the sample average $\bar{X}$ with sample size n?
- Given a population distribution, what is the probability that the sample average $\bar{X}$ is above/below a certain value?
- Given a population distribution, what is the minimum sample size n required to achieve a certain standard error for the sample average $\bar{X}$?

Questions you might be asked about sampling distribution

How to answer these questions?

- Find μ and σ^2 of the **population distribution**
- Use CLT to find the sampling distribution of the statistic
    - For $\bar{X}$, it will be Normal
    - Mean: $\mu_{\bar{X}} = \mu$ (unbiased)
    - Variance: $\sigma^2_{\bar{X}} = \frac{\sigma^2}{n}$ (shrinkage)
    - Standard error: $\sigma_{\bar{X}} = \frac{\sigma}{\sqrt{n}}$
    - Sometimes, it is useful to find the standardized statistic $Z = \frac{\bar{X}-\mu}{\sigma_{\bar{X}}}$
    - Use table/calculator to find the probability using standard Normal $N(0, 1)$

Questions you might be asked about sampling distribution

What variants we might see?

- The population mean and variance might not be provided
- There are multiple populations and random samples involved
- The statistic is not the sample mean $\bar{X}$ but something else
  - Shifted mean $(\bar{X} - c)$
  - Scaled shifted mean $(a(\bar{X} - c))$—Z-statistic is an example of this
  - Sample proportion $\left(\hat{p} = \frac{X_1 + X_2 + \cdots + X_n}{n}\right)$
    (also mean but the random variables are binaries (0 or 1))
  - Difference in means $(\bar{X}_1 - \bar{X}_2)$ or proportions $(\hat{p}_1 - \hat{p}_2)$

Same approach:
population μ and σ^2 $\Rightarrow$ sampling dist $\mu_{\hat{\theta}}, \sigma^2_{\hat{\theta}}$ $\Rightarrow$ apply $N(\mu_{\hat{\theta}}, \sigma^2_{\hat{\theta}})$ $\Rightarrow$ $Z \sim N(0, 1)$

What parameters we care about in this class?

- Mean μ (population average)
- Variance σ^2 (spread of the data)
- Proportion p (fraction of success in a population)
- Difference in means $\mu_1 - \mu_2$ (effect size)
- Difference in proportions $p_1 - p_2$ (difference in success rates)

Can you give examples why we care about these parameters?
What estimators would you use for each parameter?

What parameters we care about in this class?

- Mean μ (population average) (sample mean $\bar{X}$)
- Variance σ^2 (spread of the data) (sample variance S^2)
- Proportion p (fraction of success in a population) (sample proportion $\hat{p}$)
- Difference in means $\mu_1 - \mu_2$ (effect size) ($\bar{X}_1 - \bar{X}_2$)
- Difference in proportions $p_1 - p_2$ (difference in success rates) ($\hat{p}_1 - \hat{p}_2$)

Summary #2

- **General recipe** for sampling distribution problems:
  1. Find population parameters: μ and σ^2
  2. Apply CLT to find sampling distribution: $\mu_{\hat{\theta}}$ and $\sigma^2_{\hat{\theta}}$
  3. Use Normal approximation: $\hat{\theta} \sim N(\mu_{\hat{\theta}}, \sigma_{\hat{\theta}})$
  4. (Optional) Standardize to $Z \sim N(0, 1)$ for easier calculation
- **Key parameters**:
  - Mean μ (estimate: $\bar{X}$), Var σ^2 (estimate: S^2), Proportion p (estimate: $\hat{p}$)
  - Differences: $\mu_1 - \mu_2$ (estimate: $\bar{X}_1 - \bar{X}_2$), $p_1 - p_2$ (estimate: $\hat{p}_1 - \hat{p}_2$)
- **Remember**: Sampling distribution $\neq$ Population distribution!
  - Same mean ($\mu_{\bar{X}} = \mu$), but different variance ($\sigma^2_{\bar{X}} = \sigma^2/n$)

Class Announcement Week 1

- Syllabus updated, check Blackboard for the latest
- Homework 1 will be posted at 5pm today, due next Wednesday before class (in Gradescope)
- Office hour: Tuesday 9-10 AM (22-219 or Zoom, link in Blackboard)

Have a great first week!