STA321

STATISTICAL INFERENCE III

Criteria of Good Estimators

Series 1 outline

- **Part 1.1: Introduction to Estimators**
  - Definition of an estimator
  - Role of estimators in statistical inference
- **Part 1.2: Consistency of Estimators**
  - Definition of consistency
  - Conditions for consistency
  - Examples of consistent and inconsistent estimators
- **Part 1.3: Unbiasedness and Efficiency**
  - Definition of unbiasedness
  - Examples of unbiased and biased estimators
  - Efficiency and comparison of estimators
- **Part 1.4: Minimum Variance and Sufficiency**
  - Minimum variance unbiased estimators (MVUE)
  - Concept of sufficiency
  - Factorisation theorem and examples
- **Part 1.5: Summary and Applications**
  - Recap of properties of good estimators
  - Practical applications in statistical inference

Introduction

In statistical inference, the quality of an estimator determines the reliability of conclusions drawn from data. When researchers or analysts attempt to estimate unknown parameters, they rely on certain desirable properties to judge whether their chosen estimator is suitable. These properties ensure that the estimator not only provides accurate results but also performs well across repeated samples. Understanding these criteria is fundamental, as they form the backbone of sound statistical practice and connect directly to later topics such as estimation methods, confidence intervals, and hypothesis testing.

The aim of this Series is to equip you with the ability to identify and explain the essential properties that make an estimator "good", and to evaluate estimators against these criteria in practical situations.

We shall commence this Series by introducing the concept of estimators and their role in statistical inference. Next, we will examine the property of consistency, followed by a detailed discussion of unbiasedness and efficiency. Subsequently, we will explore minimum variance and sufficiency, including the factorisation theorem. Finally, we will wrap up by summarising the properties of good estimators and considering their applications in practice.

Series Learning Outcomes

Upon completing this Series, you should be able to:

- **SLO 1.1** define an estimator and explain its role in statistical inference,
- **SLO 1.2** describe the property of consistency and evaluate whether an estimator is consistent,
- **SLO 1.3** define unbiasedness and efficiency, and compare estimators using these properties,
- **SLO 1.4** explain the concepts of minimum variance and sufficiency, and apply the factorisation theorem to identify sufficient statistics, and
- **SLO 1.5** summarise the properties of good estimators and illustrate their applications in statistical inference

1.1 Introduction to Estimators

An **estimator** is a rule, formula, or statistic used to approximate an unknown population parameter based on sample data. For example, the sample mean is often used as an estimator of the population mean. Estimators are central to statistical inference because they allow us to make conclusions about populations using limited information from samples.

Fill in the blank: An estimator is a __________________ used to approximate an unknown population parameter.

1.1.1 Definition of an estimator

An estimator is formally defined as a function of observed sample data that provides an estimate of a population parameter. If is the parameter of interest, then denotes the estimator. Estimators can be expressed as mathematical formulas, such as the sample mean $\bar{X}=1ni=1nXi$.

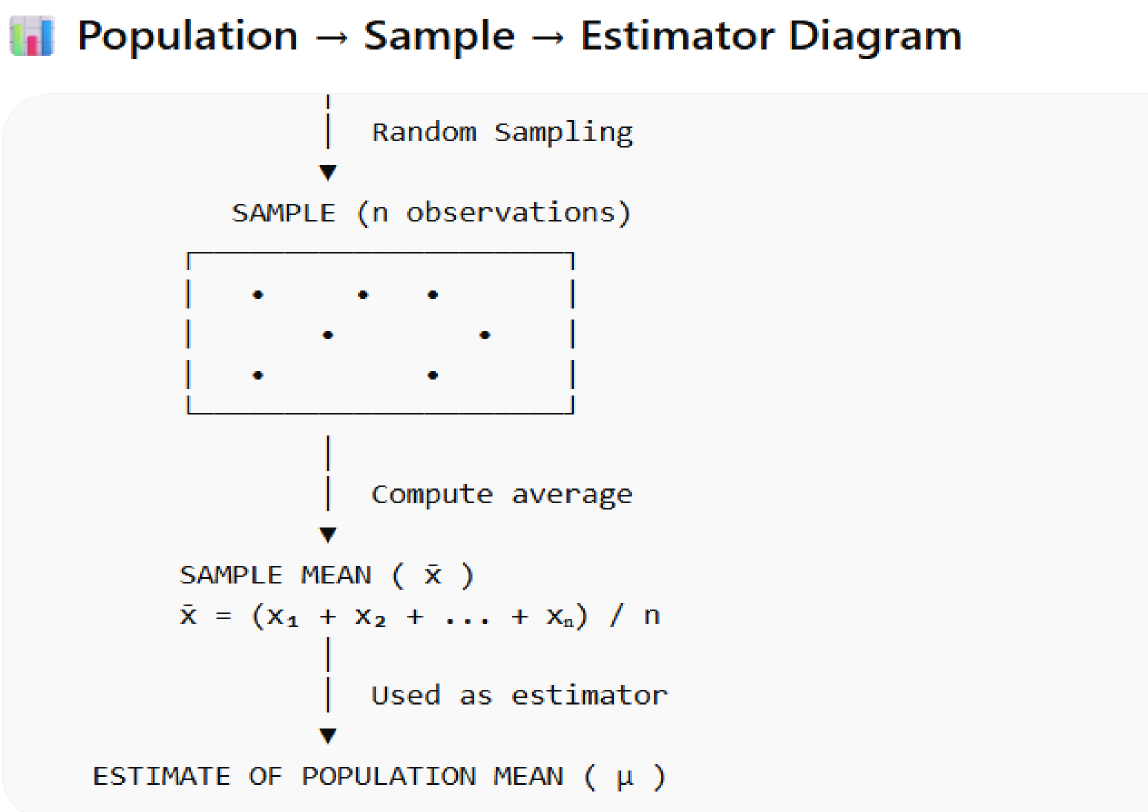

Illustration of how sample data is used to construct an estimator of a population parameter.

1.1.2 Role of estimators in statistical inference

Estimators provide the foundation for statistical inference. They are used to make decisions, test hypotheses, and construct confidence intervals. A good estimator should capture the true parameter as closely as possible, while also being reliable across repeated samples.

Fill in the blank: Estimators are essential in statistical inference because they allow us to make ___________________ about populations using sample data.

1.1.3 Examples of common estimators

- **Sample mean ($\bar{X}$)**: Estimates the population mean .
- **Sample variance (**S2**)**: Estimates the population variance 2.
- **Sample proportion (**p**)**: Estimates the population proportion p.

These examples highlight how estimators are practical tools for summarising data and making inferences.

Examples of common estimators and their corresponding parameters.

Estimators vs Population Parameters

Concept	Sample Statistic (Estimator)	Formula (Sample)	Population Parameter	Formula (Population)
Mean	Sample Mean ($\bar{x}$)	$\bar{x} = (x_1 + x_2 + \ldots + x_n) / n$	Population Mean (μ)	$\mu = (\Sigma x) / N$
Variance	Sample Variance (s^2)	$s^2 = \Sigma(x_i - \bar{x})^2 / (n - 1)$	Population Variance (σ^2)	$\sigma^2 = \Sigma(x_i - \mu)^2 / N$
Proportion	Sample Proportion ($\hat{p}$)	$\hat{p} = x / n$	Population Proportion (p)	$p = X / N$

Pilot Questions 1.1

1. An estimator is defined as: A. A function of sample data used to approximate a population parameter B. A population parameter itself C. A random experiment D. None of the above
2. The symbol is commonly used to denote: A. An estimator of a parameter B. The true parameter C. A random variable D. None of the above
3. The sample mean is an estimator of: A. The population mean B. The population variance C. The population proportion D. None of the above
4. The sample variance is an estimator of: A. The population variance B. The population mean C. The population proportion D. None of the above

5. Estimators are essential in statistical inference because they allow us to make: A. Conclusions about populations using sample data B. Predictions without data C. Decisions without parameters D. None of the above

1.2 Consistency of Estimators

An important property of a good estimator is **consistency**. An estimator is said to be consistent if, as the sample size increases, it converges in probability to the true value of the parameter being estimated. In other words, with larger samples, the estimator becomes increasingly reliable and accurate.

Fill in the blank: An estimator is consistent if it converges in ___________________ to the true parameter as the sample size increases.

1.2.1 Definition of consistency

Formally, an estimator nof a parameter is consistent if:

$$\lim_{n\to} P(|n-\theta|<\epsilon)=1 \text{ for all } \epsilon>0$$

This definition means that the probability of the estimator being close to the true parameter approaches one as the sample size grows.

📊 Consistency of an Estimator (Convergence to True Parameter μ)

```
Density
  ^
  |      Small Sample (n small)
  |         /\
  |        /  \
  |       /    \        Medium Sample
  |      /      \         /\
  |     /        \       /  \
  |    /          \     /    \        Large Sample
  |   /            \   /      \         /\
  |  /              \ /        \       /  \
  |/                 X----------\-----/----\------> θ
                   True Parameter (μ)
```

Illustration of consistency showing how the distribution of an estimator becomes concentrated around the true parameter as sample size increases.

1.2.2 Conditions for consistency

Consistency often requires two conditions:

- The estimator must be **unbiased** or asymptotically unbiased.
- The variance of the estimator must shrink to zero as the sample size increases.

For example, the sample mean is a consistent estimator of the population mean because its variance decreases as the sample size grows.

Fill in the blank: For an estimator to be consistent, its variance must ___________________ as the sample size increases.

1.2.3 Examples of consistent and inconsistent estimators

- **Consistent estimator:** The sample mean $\bar{X}$ is consistent for the population mean .
- **Inconsistent estimator:** The sample maximum is not a consistent estimator of the population mean, since it does not converge to as sample size increases.

Examples of consistent and inconsistent estimators.

Estimator	Type	What It Estimates	Consistency	Explanation
Sample Mean ($\bar{x}$)	Estimator	Population Mean (μ)	☑ Consistent	As sample size (**n**) increases, $\bar{x}$ gets closer to μ. Its variance shrinks (σ^2/n), so estimates concentrate around the true mean.
Sample Maximum (max)	Estimator	Population Mean (μ) *(incorrect target)*	✗ Inconsistent	The maximum value does not stabilize around μ. As **n** grows, it tends to increase (especially in unbounded distributions), moving away from μ instead of converging to it.

Pilot Questions 1.2

1. An estimator is consistent if: A. It converges in probability to the true parameter as sample size increases B. It always equals the true parameter C. It has zero variance in small samples D. None of the above
2. The formal definition of consistency requires: A. Probability of estimator being close to the true parameter approaches one as sample size grows B. Estimator always equals the parameter in finite samples C. Estimator has maximum variance D. None of the above
3. For an estimator to be consistent, its variance must: A. Shrink to zero as sample size increases B. Remain constant C. Increase with sample size D. None of the above
4. The sample mean is: A. A consistent estimator of the population mean B. An inconsistent estimator of the population mean C. Always biased D. None of the above
5. The sample maximum is: A. An inconsistent estimator of the population mean B. A consistent estimator of the population mean C. Always unbiased D. None of the above

1.3 Unbiasedness and Efficiency

Two further properties that help us judge the quality of an estimator are **unbiasedness** and **efficiency**. These properties ensure that an estimator not only targets the correct parameter but also does so with minimal error compared to other possible estimators.

Fill in the blank: An estimator is unbiased if its expected value is equal to the ________________ it is estimating.

1.3.1 Definition of unbiasedness

An estimator of a parameter is said to be **unbiased** if:

$$E[]=\theta$$

This means that, on average, the estimator hits the true parameter value. For example, the sample mean $\bar{X}$ is an unbiased estimator of the population mean .

📊 Unbiased Estimator (Sample Mean x̄)

```
|             *   *
|           *       *
|         *    |      *
|       *      |        *
|     *        |          *
|   *          |            *
|*             |             *
+------------|-----------------------> x̄
             μ
       (Population Mean)

 Individual sample means:
 x̄₁   x̄₂   x̄₃   x̄₄   x̄₅   ...
  •     •     •     •     •
   \    |    / \    /
    \   |   /   \  /
     ----μ---------------
```

showing unbiasedness, where the average of repeated estimates equals the true parameter.

1.3.2 Examples of unbiased and biased estimators

- **Unbiased estimator:** The sample mean $\bar{X}$ for .
- **Biased estimator:** The sample variance calculated as $\frac{1}{n}\sum(X_i-\bar{X})^2$ is biased, because its expected value is slightly less than the true variance 2. The corrected version using $\frac{1}{n-1}$ is unbiased.

Fill in the blank: The corrected sample variance using $\frac{1}{n-1}$ instead of $\frac{1}{n}$ is an ___________________ estimator of the population variance.

1.3.3 Definition of efficiency

An estimator is said to be **efficient** if it has the smallest variance among all unbiased estimators of the same parameter. Efficiency ensures that the estimator is not only correct on average but also precise.

For example, among unbiased estimators of the mean, the sample mean is the most efficient because it has the lowest variance.

Pilot Questions 1.3

1. An estimator is unbiased if: A. Its expected value equals the parameter being estimated B. Its variance is zero C. It always equals the parameter in every sample D. None of the above
2. The sample mean is: A. An unbiased estimator of the population mean B. A biased estimator of the population mean C. Always inconsistent D. None of the above
3. The sample variance using 1nis: A. Biased B. Unbiased C. Efficient D. None of the above
4. An efficient estimator is one that: A. Has the smallest variance among unbiased estimators B. Always equals the parameter C. Is biased but precise D. None of the above
5. Among unbiased estimators of the mean, the most efficient is: A. The sample mean B. The sample maximum C. The sample median D. None of the above

1.4 Minimum Variance and Sufficiency

Two additional properties that distinguish good estimators are **minimum variance** and **sufficiency**. These properties ensure that estimators are not only accurate but also optimal in terms of precision and information content.

Fill in the blank: An estimator is said to have minimum variance if it has the ________________ variance among all unbiased estimators of the same parameter.

1.4.1 Minimum variance unbiased estimators (MVUE)

A **minimum variance unbiased estimator (MVUE)** is an estimator that is both unbiased and has the lowest possible variance among all unbiased estimators of a parameter. This makes it the most reliable choice when estimating parameters.

For example, the sample mean is the MVUE of the population mean under normal distribution assumptions.

📊 Unbiased Estimators vs MVUE

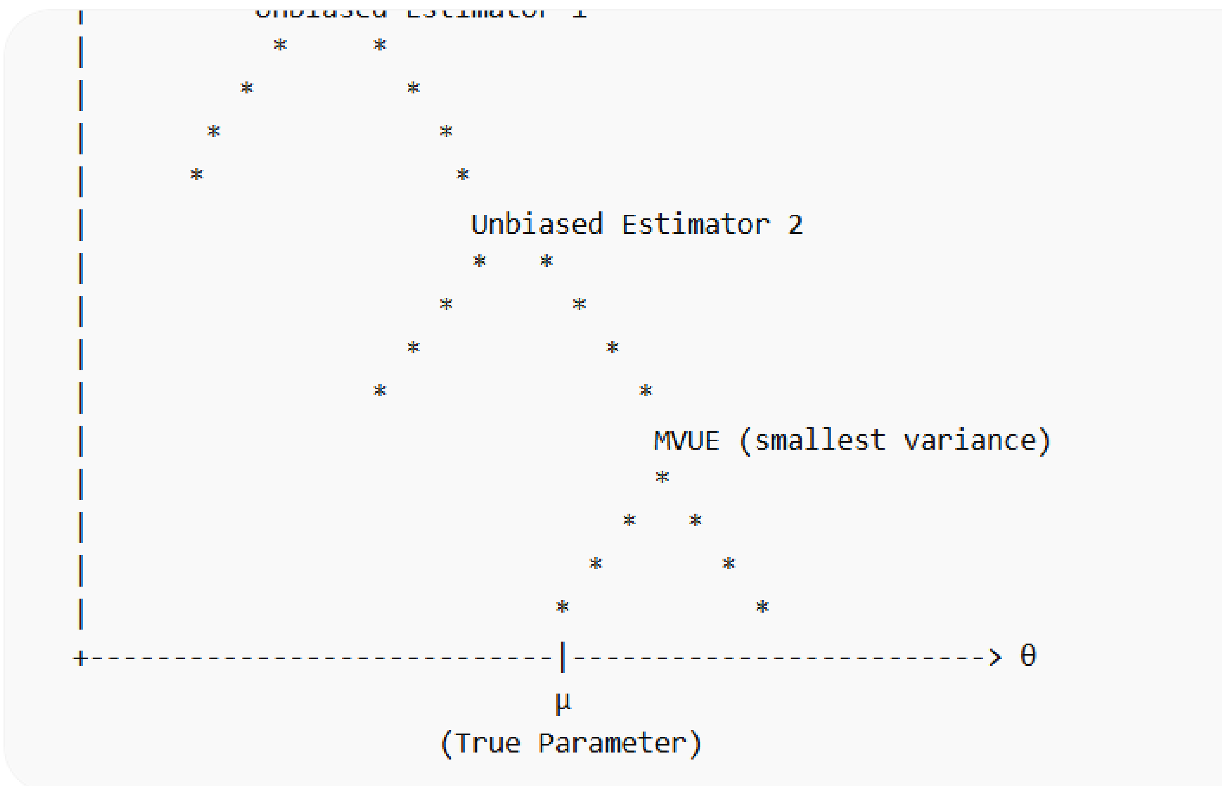

Illustration showing how the MVUE has the narrowest spread around the true parameter compared to other unbiased estimators.

1.4.2 Concept of sufficiency

An estimator is **sufficient** if it captures all the information in the sample about the parameter being estimated. Sufficiency ensures that no other statistic can provide additional information about the parameter beyond what the sufficient statistic already contains.

Formally, a statistic T(X)is sufficient for parameter if the conditional distribution of the sample given T(X)does not depend on .

Fill in the blank: A statistic is sufficient if it contains all the __________________ in the sample about the parameter.

1.4.3 Factorisation theorem

The **factorisation theorem** provides a practical way to identify sufficient statistics. It states that a statistic T(X)is sufficient for parameter if the joint probability density (or mass) function can be factorised into:

$$f(x1,x2,...,xn|\theta)=g(T(x),\theta)\cdot h(x)$$

where gdepends on the statistic and the parameter, and hdepends only on the data.

For example, in a normal distribution with mean and known variance, the sample mean is a sufficient statistic for .

Pilot Questions 1.4

1. An MVUE is defined as: A. An unbiased estimator with the smallest variance B. A biased estimator with the largest variance C. An estimator with no variance D. None of the above
2. The sample mean is the MVUE of: A. The population mean under normal distribution assumptions B. The population variance C. The population proportion D. None of the above
3. A statistic is sufficient if: A. It contains all the information in the sample about the parameter B. It ignores information in the sample C. It has maximum variance D. None of the above
4. The factorisation theorem helps to: A. Identify sufficient statistics B. Calculate sample variance C. Test unbiasedness D. None of the above
5. In a normal distribution with known variance, the sufficient statistic for the mean is: A. The sample mean B. The sample maximum C. The sample median D. None of the above

1.5 Summary and Applications

Having explored the properties of good estimators, it is useful to consolidate what we have learned and consider how these properties are applied in practice. A **good estimator** is one that is consistent, unbiased, efficient, has minimum variance, and is sufficient. These criteria ensure that the estimator not only targets the true parameter but also does so reliably and optimally.

Fill in the blank: A good estimator should be consistent, unbiased, efficient, have minimum variance, and be ___________________.

1.5.1 Recap of properties of good estimators

- **Consistency:** The estimator converges to the true parameter as sample size increases.

- **Unbiasedness:** The expected value of the estimator equals the parameter.
- **Efficiency:** Among unbiased estimators, it has the smallest variance.
- **Minimum variance:** It achieves the lowest possible spread around the parameter.
- **Sufficiency:** It captures all information in the sample about the parameter.

Checklist of properties of good estimators.

Five Properties of Good Estimators

Property	Definition / Key Idea
1. Unbiasedness	The estimator's expected value equals the true parameter. $E(\hat{\theta}) = \theta$. It does not systematically over- or underestimate.
2. Consistency	As sample size $n \to \infty$, the estimator converges in probability to the true parameter. Estimates become increasingly accurate.
3. Efficiency	Among unbiased estimators, it has the **smallest variance**. Provides the most precise estimates.
4. Sufficiency	The estimator captures **all information about the parameter** in the sample. No other statistic provides extra information.
5. Robustness	The estimator is **not overly sensitive to outliers or violations of assumptions**. Remains reliable under small deviations from model conditions.

1.5.2 Applications in statistical inference

These properties are not abstract ideas but practical tools. For example:

- In **survey sampling**, unbiasedness ensures that estimates of population proportions are correct on average.
- In **quality control**, efficiency ensures that estimators of defect rates are precise, reducing uncertainty in decision-making.
- In **econometrics**, sufficiency allows analysts to use statistics that fully capture information about parameters, improving model reliability.

Fill in the blank: In econometrics, sufficiency ensures that statistics capture all the __________________ about parameters in the sample.

Illustration of applications of estimator properties in different fields such as survey sampling, quality control, and econometrics.

📊 Applications of Estimator Properties

```
                         |
       -------------------------------------------------
       |                       |                       |
       ▼                       ▼                       ▼

+----------------+      +----------------+      +----------------+
| Survey Sampling|      | Quality Control|      | Econometrics   |
|----------------|      |----------------|      |----------------|
| - Estimate      |      | - Monitor       |      | - Regression    |
|   population    |      |   process       |      |   coefficients |
|   mean/proportion|     |   parameters    |      | - Forecasting   |
| - Use unbiased  |      | - Use efficient|      | - Policy effect|
|   and efficient |      |   estimators    |      |   estimation    |
|   sample mean   |      | - Detect shifts|      | - Hypothesis    |
| - Ensure        |      |   quickly       |      |   testing       |
|   consistency   |      | - Robust to     |      | - Standard      |
|                 |      |   small errors  |      |   errors in X   |
+----------------+      +----------------+      +----------------+
```

Pilot Questions 1.5

1. A good estimator should have which of the following properties? A. Consistency, unbiasedness, efficiency, minimum variance, and sufficiency B. Bias, inconsistency, and maximum variance C. Randomness only D. None of the above

2. Efficiency ensures that an estimator: A. Has the smallest variance among unbiased estimators B. Always equals the parameter C. Ignores sample information D. None of the above
3. Sufficiency means that a statistic: A. Contains all the information in the sample about the parameter B. Provides no information about the parameter C. Has maximum variance D. None of the above
4. In survey sampling, unbiasedness ensures that: A. Estimates of population proportions are correct on average B. Variance is always zero C. The sample maximum is used D. None of the above
5. In econometrics, sufficiency improves: A. Model reliability B. Randomness of data C. Bias in estimators D. None of the above

Series 1 Summary

In this Series, we examined the **criteria of good estimators**. We began by defining estimators and explaining their role in statistical inference. We then explored the property of **consistency**, which ensures that estimators converge to the true parameter as sample size increases. Following that, we studied **unbiasedness** and **efficiency**, highlighting how unbiased estimators target the correct parameter on average and how efficiency ensures minimal variance among unbiased estimators. We continued with **minimum variance** and **sufficiency**, introducing the concept of MVUE and the factorisation theorem as tools for identifying sufficient statistics. Finally, we summarised these properties and discussed their practical applications in fields such as survey sampling, quality control, and econometrics.

Glossary of Terms

- **Estimator:** A function of sample data used to approximate a population parameter.
- **Consistency:** Property of an estimator that converges to the true parameter as sample size increases.
- **Unbiasedness:** When the expected value of an estimator equals the parameter being estimated.
- **Efficiency:** Property of an estimator that has the smallest variance among unbiased estimators.
- **Minimum variance unbiased estimator (MVUE):** An estimator that is unbiased and has the lowest variance.

- **Sufficiency:** A statistic that contains all information in the sample about the parameter.
- **Factorisation theorem:** A method for determining whether a statistic is sufficient for a parameter.

Concept Check Questions 1

Objective Questions

1. Define an estimator and give one example. (Linked to SLO 1.1)
2. State the formal definition of consistency. (Linked to SLO 1.2)
3. What condition must the variance of an estimator satisfy for consistency? (Linked to SLO 1.2)
4. Write the definition of unbiasedness. (Linked to SLO 1.3)
5. What is the MVUE? (Linked to SLO 1.4)

Short-Answer Questions

6. Explain why the corrected sample variance using 1n-1is unbiased. (Linked to SLO 1.3)

7. Describe the concept of sufficiency and give one example. (Linked to SLO 1.4)

8. Discuss how efficiency improves the reliability of estimators. (Linked to SLO 1.3)

Long-Answer Question

9. Compare and contrast consistency, unbiasedness, efficiency, minimum variance, and sufficiency, explaining why each property is important in statistical inference. (Linked to SLOs 1.2–1.4)

Answers to Pilot Questions 1

Part 1.1: 1-A, 2-A, 3-A, 4-A, 5-A **Part 1.2:** 1-A, 2-A, 3-A, 4-A, 5-A **Part 1.3:** 1-A, 2-A, 3-A, 4-A, 5-A **Part 1.4:** 1-A, 2-A, 3-A, 4-A, 5-A **Part 1.5:** 1-A, 2-A, 3-A, 4-A, 5-A

References and Attributions 1

- **Figure 1.1:** Illustration of estimator construction. AI prompt: "Generate a diagram showing a population, a sample drawn from it, and the sample mean used as an estimator of the population mean."

- **Figure 1.2:** Illustration of consistency. AI prompt: "Generate a diagram showing distributions of an estimator with small, medium, and large sample sizes converging to the true parameter."
- **Figure 1.3:** Illustration of unbiasedness. AI prompt: "Generate a diagram showing repeated sample means centred around the population mean to illustrate unbiasedness."
- **Table 1.1:** Examples of common estimators. AI prompt: "Generate a table showing sample mean, sample variance, and sample proportion with their corresponding population parameters."
- **Table 1.2:** Consistent vs inconsistent estimators. AI prompt: "Generate a table listing sample mean as consistent and sample maximum as inconsistent, with brief explanations."
- **Table 1.3:** Comparison of unbiasedness and efficiency. AI prompt: "Generate a table comparing unbiasedness and efficiency, with examples such as sample mean and sample variance."
- **Figure 1.4:** Illustration of MVUE. AI prompt: "Generate a diagram comparing distributions of different unbiased estimators, highlighting the MVUE with the smallest variance."
- **Box 1.1:** Factorisation theorem summary. AI prompt: "Generate a box summarising the factorisation theorem with formula and example for the normal distribution."
- **Box 1.2:** Checklist of properties of good estimators. AI prompt: "Generate a box summarising the five properties of good estimators with brief definitions."
- **Figure 1.5:** Applications of estimator properties. AI prompt: "Generate a diagram showing applications of estimator properties in survey sampling, quality control, and econometrics."

2 Methods of Estimation – Maximum Likelihood, Least Squares, and Method of Moments

Series outline

- **Part 2.1: Introduction to Methods of Estimation**

- Definition and role of estimation methods
- Importance in statistical inference

- **Part 2.2: Maximum Likelihood Estimation (MLE)**
  - Concept and definition
  - Properties of MLE
  - Examples and applications
- **Part 2.3: Least Squares Estimation (LSE)**
  - Concept and definition
  - Properties of LSE
  - Examples and applications
- **Part 2.4: Method of Moments (MoM)**
  - Concept and definition
  - Properties of MoM
  - Examples and applications
- **Part 2.5: Comparative Discussion and Applications**
  - Comparison of MLE, LSE, and MoM
  - Practical applications in different fields

Introduction

In statistical inference, choosing the right method of estimation is crucial for obtaining reliable results. Different methods have been developed to estimate parameters, each with its own strengths and limitations. These methods provide the tools that statisticians and researchers use to bridge the gap between theory and practice.

The aim of this Series is to introduce you to the three major methods of estimation—maximum likelihood, least squares, and method of moments—and to equip you with the ability to apply and evaluate them in practical situations.

We shall commence this Series by introducing the general concept of estimation methods and their role in statistical inference. Next, we will explore maximum likelihood estimation, followed by least squares estimation. Subsequently, we will study the method of moments. Finally, we will conclude with a comparative discussion of these methods and their applications in practice.

Series Learning Outcomes

Upon completing this Series, you should be able to:

- **SLO 2.1** define estimation methods and explain their role in statistical inference,
- **SLO 2.2** describe the concept of maximum likelihood estimation and apply it to parameter estimation problems,
- **SLO 2.3** explain least squares estimation and demonstrate its use in regression analysis,
- **SLO 2.4** define the method of moments and apply it to derive estimators, and
- **SLO 2.5** compare maximum likelihood, least squares, and method of moments, and evaluate their applications in statistical practice.

2.1 Introduction to Methods of Estimation

In statistical inference, **methods of estimation** are systematic approaches used to derive values of unknown parameters from sample data. These methods provide the foundation for making reliable conclusions about populations. Without them, statistical analysis would lack structure and precision.

Fill in the blank: Methods of estimation are systematic ___________________ used to derive values of unknown parameters from sample data.

2.1.1 Definition and role of estimation methods

An estimation method is a procedure that specifies how to construct an estimator from observed data. For example, the sample mean is derived using a straightforward formula, but more complex parameters often require specialised methods. These methods ensure that estimators are not chosen arbitrarily but follow logical and mathematical principles.

Illustration of the role of estimation methods in linking sample data to population parameters.

📊 From Sample Data to Parameter Estimates

```
        |
        ▼
+--------------------+
| Estimation Methods |
|--------------------|
| - Method of Moments|
| - Maximum Likelihood|
| - Least Squares     |
| - Bayesian Estimation|
+--------------------+
        |
        |
        ▼
PARAMETER ESTIMATES
(θ̂₁, θ̂₂, ...)
- Sample mean x̄ → μ
- Sample variance s² → σ²
- Sample proportion p̂ → p
```

2.1.2 Importance in statistical inference

Estimation methods are important because they determine the quality of the estimators produced. A poor method may lead to biased or inconsistent estimators, while a good method ensures accuracy and reliability. The choice of method depends on the nature of the data, the parameter being estimated, and the desired properties of the estimator.

Fill in the blank: The choice of estimation method depends on the nature of the ________________, the parameter, and the desired properties of the estimator.

2.1.3 Overview of major methods

The three major methods of estimation that we will study in this Series are:

- **Maximum likelihood estimation (MLE):** Based on maximising the likelihood function.

- **Least squares estimation (LSE):** Based on minimising the sum of squared errors.
- **Method of moments (MoM):** Based on equating sample moments to population moments.

These methods are widely used in practice and form the backbone of modern statistical inference.

Overview of major estimation methods.

📦 Overview of Estimation Methods

Method	Definition / Key Idea	Typical Use / Example
Maximum Likelihood (MLE)	Finds the parameter values that **maximize the likelihood** of observing the given sample.	Estimating μ and σ^2 in normal distribution; logistic regression coefficients.
Least Squares (LSE)	Minimizes the **sum of squared differences** between observed and predicted values.	Linear regression: estimating slope and intercept.
Method of Moments (MoM)	Sets **sample moments equal to population moments** to solve for parameter estimates.	Estimating λ in Poisson distribution or μ, σ^2 in normal distribution.

Pilot Questions 2.1

1. Methods of estimation are defined as: A. Systematic procedures used to derive parameter values from sample data B. Random guesses of parameters C. Arbitrary formulas without logic D. None of the above

2. The role of estimation methods is to: A. Provide structured procedures for constructing estimators B. Eliminate the need for data C. Always guarantee zero variance D. None of the above
3. The choice of estimation method depends on: A. The nature of the data, the parameter, and desired properties of the estimator B. Random preference of the researcher C. The largest possible variance D. None of the above
4. Maximum likelihood estimation is based on: A. Maximising the likelihood function B. Minimising squared errors C. Equating sample moments to population moments D. None of the above
5. Least squares estimation is based on: A. Minimising the sum of squared errors B. Maximising the likelihood function C. Equating sample moments to population moments D. None of the above

2.2 Maximum Likelihood Estimation (MLE)

The **maximum likelihood estimation (MLE)** method is one of the most widely used approaches in statistics. It is based on the principle of choosing parameter values that maximise the likelihood of observing the given sample data. This method is powerful because it often produces estimators with desirable properties such as consistency and efficiency.

Fill in the blank: Maximum likelihood estimation chooses parameter values that ___________________ the likelihood of observing the sample data.

2.2.1 Concept and definition

The **likelihood function** is defined as the joint probability of the observed sample, expressed as a function of the unknown parameter. MLE involves finding the parameter value that maximises this function.

Formally, if $X_1, X_2, \ldots, X_n$are independent observations with probability density function $f(x|\theta)$, then the likelihood function is:

$$L(\theta)=i=1nf(X_i|\theta)$$

The MLE, denoted MLE, is the value of that maximises $L(\theta)$.

Likelihood Function and MLE

```
Likelihood L(θ)
  ^
  |           *
  |          * *
  |         *   *
  |        *     *
  |       *       *
  |      *         *
  |     *           *
  |    *             *
  |   *               *
  |  *                 *
  | *                   *
  |*                     *
  +----------------------------->
           θ
          θ̂_MLE
```

- **Curve**: Shows the likelihood $L(\theta)$ as a function of the parameter θ.
- **θ̂_MLE**: The **parameter value that maximizes the likelihood**; the peak of the curve.

Illustration of a likelihood function curve showing the maximum point as the MLE.

2.2.2 Properties of MLE

MLE has several important properties:

- **Consistency:** The estimator converges to the true parameter as sample size increases.
- **Asymptotic normality:** For large samples, the distribution of the MLE approaches a normal distribution.
- **Efficiency:** MLE often achieves the lowest possible variance among consistent estimators in large samples.

Fill in the blank: For large samples, the distribution of the MLE approaches a _______________________ distribution.

2.2.3 Examples and applications

- **Example 1:** Estimating the mean of a normal distribution with known variance. The MLE is the sample mean.
- **Example 2:** Estimating the probability of success in a Bernoulli distribution. The MLE is the sample proportion of successes.
- **Application:** MLE is widely used in econometrics, biostatistics, and machine learning because of its flexibility and strong theoretical foundation.

Examples of MLE for different distributions. **(MLE)**

Distribution	**Parameter**	**Likelihood Function L(θ)**	**MLE (θ̂) Formula**	**Notes**
Normal (N(μ, σ²))	μ (mean), σ² known	L(μ)=i=1n12π2exp[-()222]	MLE=X‾=1ni=1 nxi	Sample mean maximizes the likelihood of observing the data
Bernoulli (p)	p (success probability)	L(p)=i=1npxi(1-p)1-xi	pMLE=1ni=1nxi	Sample proportion of successes is the MLE

Pilot Questions 2.2

1. Maximum likelihood estimation is based on: A. Maximising the likelihood function B. Minimising squared errors C. Equating sample moments to population moments D. None of the above
2. The likelihood function is defined as: A. The joint probability of the observed sample expressed as a function of the parameter B. The variance of the sample C. The mean of the sample D. None of the above

3. The MLE of the mean of a normal distribution with known variance is: A. The sample mean B. The sample variance C. The sample maximum D. None of the above
4. For large samples, the distribution of the MLE approaches: A. A normal distribution B. A uniform distribution C. A binomial distribution D. None of the above
5. The MLE of the probability of success in a Bernoulli distribution is: A. The sample proportion of successes B. The sample mean of failures C. The sample variance D. None of the above

2.3 Least Squares Estimation (LSE)

The **least squares estimation (LSE)** method is another widely used approach in statistics, particularly in regression analysis. It is based on the principle of minimising the sum of squared differences between observed values and the values predicted by a model. This method is intuitive and computationally straightforward, making it one of the most popular techniques in applied statistics.

Fill in the blank: Least squares estimation is based on minimising the sum of __________________ between observed and predicted values.

2.3.1 Concept and definition

Suppose we have data points yiand we want to fit a line y=0+1x. The least squares method chooses 0and 1such that the sum of squared errors is minimised:

S(0,1)=i=1n(yi-(0+1xi))2

The values of 0and 1that minimise this function are the least squares estimators.

📊 Least Squares Regression Line

```
Y
^
|        *
|   *         *
|        *    *
|  *       *
|        *
|      *
|  *
|             *
|------------------------------> X
       |       |       |
       -----------------
         Fitted Line: Ŷ = b0 + b1 X
```

Illustration of least squares regression line fitted to data points.

2.3.2 Properties of LSE

Least squares estimators have several important properties:

- **Unbiasedness:** Under certain conditions, the least squares estimators are unbiased.
- **Efficiency:** They are efficient within the class of linear unbiased estimators, as shown by the Gauss–Markov theorem.
- **Simplicity:** The method is easy to compute and interpret, which explains its widespread use.

Fill in the blank: According to the Gauss–Markov theorem, least squares estimators are the most __________________ linear unbiased estimators.

2.3.3 Examples and applications

- **Example 1:** Fitting a straight line to data in simple linear regression.

- **Example 2:** Estimating coefficients in multiple regression models.
- **Application:** LSE is widely used in economics, engineering, and social sciences to model relationships between variables.

Examples of least squares estimation in practice.

📊 Examples of Least Squares Estimation (Regression)

Regression Type	Model Equation	LSE Formula for Coefficients	Notes
Simple Regression	Y=0+1X+ϵ	1=∑(Xi-Xˉ)(Yi-Yˉ)∑(Xi-Xˉ)20=Yˉ-1Xˉ	Minimizes ∑(Yi-Yi)2. Used when there is **one predictor** X.
Multiple Regression	Y=0+1X1+⋯+kXk+ϵ	=(X′X)-1X′Y	Minimizes ∑(Yi-Yi)2. Used when there are **multiple predictors** X1,...,Xk.

Pilot Questions 2.3

1. Least squares estimation is based on: A. Minimising the sum of squared errors B. Maximising the likelihood function C. Equating sample moments to population moments D. None of the above
2. In simple linear regression, least squares estimation finds values of: A. 0and 1that minimise squared errors B. The sample mean and variance C. The maximum likelihood function D. None of the above
3. The Gauss–Markov theorem states that least squares estimators are: A. The most efficient linear unbiased estimators B. Always biased C. Inconsistent D. None of the above
4. Least squares estimation is widely used in: A. Regression analysis B. Probability theory only C. Random sampling only D. None of the above

5. An application of least squares estimation is: A. Fitting regression models in economics and engineering B. Estimating probabilities in Bernoulli trials C. Calculating maximum likelihood functions D. None of the above

2.4 Method of Moments (MoM)

The **method of moments (MoM)** is a classical approach to parameter estimation. It is based on the idea of equating sample moments to their corresponding population moments. This method is straightforward and often provides quick estimators, especially when maximum likelihood or least squares methods are difficult to apply.

Fill in the blank: The method of moments is based on equating __________________ moments to population moments.

2.4.1 Concept and definition

A **moment** is a measure of the shape of a distribution. The kthpopulation moment is defined as:

$$k=E[Xk]$$

The method of moments estimator is obtained by setting the sample moment equal to the population moment and solving for the parameter. For example, the first sample moment (the sample mean) is equated to the first population moment (the expected value).

📊 Method of Moments: Sample Mean → Population Mean

```
Population Parameter μ
        |
        |
        ▼
   True Population Mean μ
   ----------------------
   |                     |
   |   X₁  X₂  ...  Xₙ   |  ← Sample drawn from population
   ----------------------
        |
        | Compute sample mean
        ▼
   Sample Mean x̄ = (1/n) ∑ Xᵢ
        |
        | Equate to population mean
        ▼
   Method of Moments Estimate:
        μ̂_MoM = x̄
```

Illustration of the method of moments showing how sample moments are equated to population moments.

2.4.2 Properties of MoM

- **Simplicity:** The method is easy to apply and requires minimal computation.
- **Generality:** It can be applied to a wide range of distributions.
- **Limitations:** MoM estimators are not always efficient or unbiased, but they often provide reasonable starting values for more advanced methods.

Fill in the blank: A limitation of the method of moments is that its estimators are not always __________________ or efficient.

2.4.3 Examples and applications

- **Example 1:** For a Poisson distribution with mean , the first sample moment (sample mean) is equated to , giving the estimator $=\bar{X}$.

- **Example 2:** For a normal distribution with mean and variance 2, the first two sample moments are equated to the population mean and variance, yielding estimators for and 2.
- **Application:** MoM is often used in actuarial science and applied probability, where quick and simple estimators are needed.

Examples of method of moments estimators for Poisson and normal distributions.

📊 Method of Moments Estimators

Distribution	Parameter(s)	Population Moment(s)	Sample Moment(s)	MoM Estimator (θ̂)
Poisson (λ)	λ (mean)	E[X]=λ	X¯=1n∑Xi	MoM=X¯
Normal (μ, σ²)	μ (mean), σ² (variance)	E[X]=μ,Var(X)=2	X¯=1n∑Xi,S2=1n∑(Xi-X¯)2	MoM=X¯,MoM2=S2

Pilot Questions 2.4

1. The method of moments is based on: A. Equating sample moments to population moments B. Maximising the likelihood function C. Minimising squared errors D. None of the above
2. The first population moment is: A. The expected value of the random variable B. The variance of the random variable C. The skewness of the distribution D. None of the above
3. For a Poisson distribution, the method of moments estimator of is: A. The sample mean B. The sample variance C. The sample maximum D. None of the above
4. A limitation of the method of moments is: A. Its estimators are not always unbiased or efficient B. It cannot be applied to any distribution C. It requires complex computation D. None of the above
5. The method of moments is often used in: A. Actuarial science and applied probability B. Regression analysis only C. Machine learning exclusively D. None of the above

2.5 Comparative Discussion and Applications

Having studied maximum likelihood, least squares, and method of moments individually, it is important to compare them and understand their practical applications. Each method has strengths and limitations, and the choice of method often depends on the context of the problem, the nature of the data, and the properties desired in the estimator.

Fill in the blank: The choice of estimation method depends on the context of the problem, the nature of the data, and the ___________________ desired in the estimator.

2.5.1 Comparison of MLE, LSE, and MoM

- **Maximum likelihood estimation (MLE):** Provides estimators with strong theoretical properties such as consistency, asymptotic normality, and efficiency. However, it can be computationally intensive.
- **Least squares estimation (LSE):** Simple and intuitive, particularly useful in regression analysis. It is efficient within the class of linear unbiased estimators but may not be optimal in non-linear contexts.
- **Method of moments (MoM):** Easy to apply and general, but its estimators are not always unbiased or efficient. Often used as a quick alternative or starting point for more advanced methods.

Comparison of MLE, LSE, and MoM.

📊 **Comparison of Estimation Methods**

Method	Properties	Strengths	Limitations
Maximum Likelihood (MLE)	- Often **consistent, asymptotically unbiased, and efficient** - Uses **full likelihood**	- Produces **efficient estimators** for large samples - Applicable to many parametric models - Can handle complex models	- Can be **computationally intensive** - May not exist or be unique for small samples - Sensitive to model assumptions
Least Squares (LSE)	- **Unbiased** for linear regression parameters (if Gauss-Markov assumptions hold)	- Simple, intuitive, widely used in regression - Minimizes sum	- Assumes **linearity and homoscedasticity** - Sensitive to outliers - Not directly

	- Often **efficient** when errors are normal	of squared errors - Works well for prediction	applicable to non-linear models
Method of Moments (MoM)	- **Unbiased** under certain conditions - Simple to compute - Relies on matching sample and population moments	- Easy to use when **likelihood is complicated or unknown** - Intuitive and flexible	- Can be **less efficient** than MLE - May produce estimates outside parameter space - Accuracy depends on chosen moments

2.5.2 Applications in practice

- **MLE:** Widely used in econometrics, biostatistics, and machine learning due to its flexibility and efficiency in large samples.
- **LSE:** Commonly applied in regression analysis across economics, engineering, and social sciences.
- **MoM:** Frequently used in actuarial science and applied probability for quick and simple estimations.

Fill in the blank: Least squares estimation is most commonly applied in ____________________ analysis.

Pilot Questions 2.5

1. Maximum likelihood estimation is particularly valued for: A. Consistency, asymptotic normality, and efficiency B. Simplicity only C. Always producing unbiased estimators in small samples D. None of the above
2. Least squares estimation is most commonly applied in: A. Regression analysis B. Probability theory only C. Sampling design exclusively D. None of the above
3. The method of moments is often used because: A. It is simple and general B. It always produces efficient estimators C. It requires maximum computation D. None of the above
4. A limitation of MLE is: A. It can be computationally intensive B. It is never consistent C. It ignores sample information D. None of the above

5. Which method is frequently used in actuarial science? A. Method of moments B. Maximum likelihood estimation C. Least squares estimation D. None of the above

Series 2 Summary

In this Series, we explored the **methods of estimation** that statisticians use to derive parameter values from sample data. We began with an introduction to estimation methods and their role in statistical inference. We then studied **maximum likelihood estimation (MLE)**, which is based on maximising the likelihood function and is valued for its consistency, asymptotic normality, and efficiency. Next, we examined **least squares estimation (LSE)**, which minimises the sum of squared errors and is widely applied in regression analysis. Following that, we discussed the **method of moments (MoM)**, which equates sample moments to population moments and provides simple though sometimes less efficient estimators. Finally, we compared these three methods and highlighted their practical applications across fields such as econometrics, engineering, actuarial science, and machine learning.

Glossary of Terms

- **Estimation method:** A systematic procedure for deriving parameter values from sample data.
- **Maximum likelihood estimation (MLE):** A method that chooses parameter values by maximising the likelihood of observing the sample data.
- **Likelihood function:** The joint probability of observed data expressed as a function of the parameter.
- **Least squares estimation (LSE):** A method that minimises the sum of squared differences between observed and predicted values.
- **Gauss–Markov theorem:** A theorem stating that least squares estimators are the most efficient linear unbiased estimators.
- **Method of moments (MoM):** A method that equates sample moments to population moments to estimate parameters.
- **Moment:** A measure of the shape of a distribution, such as mean or variance.

Concept Check Questions 2

Objective Questions

1. Define an estimation method and explain its role in statistical inference. (Linked to SLO 2.1)
2. State the formula for the likelihood function. (Linked to SLO 2.2)
3. What does the Gauss–Markov theorem say about least squares estimators? (Linked to SLO 2.3)
4. Write the definition of the method of moments. (Linked to SLO 2.4)
5. Compare MLE, LSE, and MoM in terms of strengths and limitations. (Linked to SLO 2.5)

Short-Answer Questions

6. Explain why MLE is considered efficient in large samples. (Linked to SLO 2.2)

7. Describe how least squares estimation is applied in regression analysis. (Linked to SLO 2.3)

8. Give one example of a method of moments estimator. (Linked to SLO 2.4)

Long-Answer Question

9. Discuss the comparative advantages and disadvantages of maximum likelihood, least squares, and method of moments, and explain how the choice of method depends on the context of the problem. (Linked to SLO 2.5)

Answers to Pilot Questions 2

Part 2.1: 1-A, 2-A, 3-A, 4-A, 5-A **Part 2.2:** 1-A, 2-A, 3-A, 4-A, 5-A **Part 2.3:** 1-A, 2-A, 3-A, 4-A, 5-A **Part 2.4:** 1-A, 2-A, 3-A, 4-A, 5-A **Part 2.5:** 1-A, 2-A, 3-A, 4-A, 5-A

References and Attributions 2

- **Figure 2.1:** Role of estimation methods. AI prompt: "Generate a diagram showing sample data being processed through estimation methods to produce parameter estimates."
- **Figure 2.2:** Likelihood function curve. AI prompt: "Generate a diagram showing a likelihood function curve with the maximum point labelled as the MLE."

- **Table 2.1:** Examples of MLE. AI prompt: "Generate a table showing examples of MLE for normal mean and Bernoulli probability with formulas."
- **Figure 2.3:** Least squares regression line. AI prompt: "Generate a diagram showing a scatter plot of data points with a regression line fitted using least squares."
- **Table 2.2:** Examples of least squares estimation. AI prompt: "Generate a table showing examples of least squares estimation in simple and multiple regression with formulas."
- **Figure 2.4:** Method of moments illustration. AI prompt: "Generate a diagram showing sample mean equated to population mean to illustrate the method of moments."
- **Table 2.3:** Examples of MoM estimators. AI prompt: "Generate a table showing method of moments estimators for Poisson mean and normal mean/variance."
- **Table 2.4:** Comparison of estimation methods. AI prompt: "Generate a table comparing maximum likelihood, least squares, and method of moments in terms of properties, strengths, and limitations."
- **Figure 2.5:** Applications of estimation methods. AI prompt: "Generate a diagram showing applications of MLE in econometrics, LSE in regression, and MoM in actuarial science."
- **Box 2.1: Overview of major estimation methods. AI prompt: "Generate a box summarising maximum likelihood, least squares, and method of moments with brief definitions."**

3 Confidence Intervals for Statistical Inference

Series outline

- **Part 3.1: Introduction to Confidence Intervals**
    - Concept and definition
    - Role in statistical inference
- **Part 3.2: Confidence Intervals for the Mean**

- Known variance case
- Unknown variance case
- Examples and applications
- **Part 3.3: Confidence Intervals for Proportions**
  - Derivation and interpretation
  - Practical examples
- **Part 3.4: Confidence Intervals for Variance**
  - Using chi-square distribution
  - Applications in practice
- **Part 3.5: Comparative Discussion and Applications**
  - Comparing confidence intervals across parameters
  - Practical uses in research and decision-making

Introduction

In statistical inference, point estimates provide single values for unknown parameters, but they do not convey the uncertainty inherent in sampling. To address this, statisticians use **confidence intervals**, which provide a range of plausible values for the parameter along with a specified level of confidence. Confidence intervals are essential because they combine estimation with a measure of reliability, making them more informative than point estimates alone.

The aim of this Series is to introduce you to the concept of confidence intervals, explain how they are constructed for different parameters, and show how they are applied in practice to support sound statistical conclusions.

We shall commence this Series by defining confidence intervals and explaining their role in statistical inference. Next, we will study confidence intervals for the mean, followed by confidence intervals for proportions. Subsequently, we will explore confidence intervals for variance. Finally, we will conclude with a comparative discussion of confidence intervals across parameters and their practical applications in research and decision-making.

Series Learning Outcomes

Upon completing this Series, you should be able to:

- **SLO 3.1** define confidence intervals and explain their role in statistical inference,
- **SLO 3.2** construct and interpret confidence intervals for the mean under both known and unknown variance cases,
- **SLO 3.3** derive and interpret confidence intervals for proportions,
- **SLO 3.4** calculate confidence intervals for variance using the chi-square distribution, and
- **SLO 3.5** compare confidence intervals across parameters and evaluate their applications in research and practice.

3.1 Introduction to Confidence Intervals

When we estimate a population parameter using sample data, a single value called a **point estimate** is often calculated. While useful, point estimates do not reflect the uncertainty that arises from sampling variation. To address this, statisticians use **confidence intervals**, which provide a range of plausible values for the parameter along with a specified confidence level (such as 95%). Confidence intervals therefore combine estimation with reliability, making them more informative than point estimates alone.

Fill in the blank: A confidence interval provides a ________________ of plausible values for a parameter along with a confidence level.

3.1.1 Concept and definition

A confidence interval for a parameter is an interval estimate constructed from sample data, expressed as:

Margin of Error

The confidence level (e.g., 95%) indicates the proportion of times that intervals constructed in this way would contain the true parameter if repeated samples were taken.

Illustration of confidence intervals showing repeated samples and intervals, most of which contain the true parameter.

3.1.2 Role in statistical inference

Confidence intervals play a crucial role in statistical inference because they:

- Provide a measure of uncertainty around point estimates.

- Allow researchers to assess the precision of their estimates.
- Support decision-making by indicating the range within which the true parameter is likely to lie.

For example, if a 95% confidence interval for the mean income is [₦50,000, ₦55,000], we can be reasonably confident that the true mean lies within this range.

Fill in the blank: Confidence intervals support decision-making by indicating the __________________ within which the true parameter is likely to lie.

3.1.3 Confidence level interpretation

It is important to interpret confidence levels correctly. A 95% confidence interval does not mean there is a 95% probability that the true parameter lies within the interval for a single sample. Instead, it means that if we were to take many samples and construct intervals in the same way, about 95% of those intervals would contain the true parameter.

Key points about interpreting confidence levels.

Pilot Questions 3.1

1. A confidence interval provides: A. A range of plausible values for a parameter along with a confidence level B. A single point estimate only C. The variance of the sample D. None of the above
2. The formula for a confidence interval is generally expressed as: A. Point estimate ± Margin of Error B. Variance × Mean C. Likelihood function × Parameter D. None of the above
3. A 95% confidence interval means: A. About 95% of intervals constructed from repeated samples will contain the true parameter B. The true parameter has a 95% probability of lying in the interval for a single sample C. The interval is always correct D. None of the above
4. Confidence intervals are important because they: A. Provide a measure of uncertainty around point estimates B. Eliminate the need for sampling C. Always guarantee unbiasedness D. None of the above
5. If a 95% confidence interval for the mean income is [₦50,000, ₦55,000], we can say: A. We are reasonably confident the true mean lies within this range B. The true mean is exactly ₦52,500 C. The variance is zero D. None of the above

3.2 Confidence Intervals for the Mean

Confidence intervals for the mean are among the most commonly used tools in statistics. They provide a range of values within which the true population mean is likely to lie, based on sample data. The construction of these intervals depends on whether the population variance is known or unknown.

Fill in the blank: Confidence intervals for the mean provide a __________________ of values within which the true population mean is likely to lie.

3.2.1 Known variance case

When the population variance 2is known, the confidence interval for the mean is given by:

$$\bar{X} Z/2n$$

Here, $\bar{X}$is the sample mean, $Z/2$is the critical value from the standard normal distribution, is the population standard deviation, and $n$is the sample size.

Illustration of confidence interval for the mean with known variance.

3.2.2 Unknown variance case

When the population variance is unknown, we use the sample standard deviation sand the **t-distribution** instead of the normal distribution. The confidence interval becomes:

$$\bar{X} t/2, n\text{-}1sn$$

Here, $t/2, n\text{-}1$is the critical value from the t-distribution with $n\text{-}1$degrees of freedom. This adjustment accounts for the additional uncertainty introduced by estimating the variance from the sample.

Fill in the blank: When the population variance is unknown, the confidence interval for the mean uses the __________________ distribution.

Comparison of confidence intervals for the mean with known and unknown variance.

3.2.3 Examples and applications

- **Example 1:** A sample of 100 students has a mean test score of 70 with a known population variance of 25. A 95% confidence interval for the mean can be constructed using the Z-distribution.
- **Example 2:** A sample of 20 households has a mean income of ₦200,000 with a sample standard deviation of ₦15,000. Since the population variance is unknown, the t-distribution is used to construct the confidence interval.
- **Application:** Confidence intervals for the mean are widely used in education, economics, and health sciences to estimate average outcomes with a measure of reliability.

Fill in the blank: Confidence intervals for the mean are widely used in fields such as __________________, economics, and health sciences.

Pilot Questions 3.2

1. When the population variance is known, the confidence interval for the mean uses: A. The Z-distribution B. The t-distribution C. The chi-square distribution D. None of the above
2. The formula for the confidence interval with known variance is: A. X¯Z/2nB. X¯t/2,n-1snC. X¯/22s2nD. None of the above
3. When the population variance is unknown, the confidence interval for the mean uses: A. The t-distribution B. The Z-distribution C. The chi-square distribution D. None of the above
4. The adjustment for unknown variance accounts for: A. Additional uncertainty introduced by estimating variance from the sample B. Elimination of sampling error C. Exact knowledge of the population parameter D. None of the above
5. Confidence intervals for the mean are widely applied in: A. Education, economics, and health sciences B. Astronomy only C. Linguistics exclusively D. None of the above

3.3 Confidence Intervals for Proportions

Confidence intervals for proportions are used when we want to estimate the proportion of a population that has a certain characteristic. For example, the proportion of voters who support a candidate, or the proportion of defective items in a batch. These intervals provide a range of plausible values for the population proportion based on sample data.

Fill in the blank: Confidence intervals for proportions provide a __________________ of plausible values for the population proportion.

3.3.1 Derivation and formula

If pis the sample proportion and nis the sample size, the confidence interval for the population proportion pis:

$$pZ/2p(1\text{-}p)n$$

Here, $Z/2$is the critical value from the standard normal distribution. This formula assumes a sufficiently large sample size so that the normal approximation is valid.

Illustration of confidence interval for a proportion.

3.3.2 Interpretation

The confidence interval indicates the range within which the true population proportion is likely to lie, given the sample data. For example, if a survey finds that 60% of respondents support a policy, with a 95% confidence interval of [0.55, 0.65], we can be reasonably confident that the true proportion of supporters in the population lies between 55% and 65%.

Fill in the blank: A 95% confidence interval for a proportion means we are reasonably confident the true proportion lies within the __________________.

3.3.3 Examples and applications

- **Example 1:** In a sample of 200 voters, 120 support a candidate. The sample proportion is $p=0.6$. A 95% confidence interval can be constructed using the formula above.
- **Example 2:** In a batch of 500 items, 25 are defective. The sample proportion of defective items is $p=0.05$. A confidence interval provides a range for the true defect rate in the population.
- **Application:** Confidence intervals for proportions are widely used in public opinion polling, quality control, and epidemiology.

Examples of confidence intervals for proportions in voting and quality control.

Pilot Questions 3.3

1. The formula for a confidence interval for a proportion is: A. pZ/2p(1-p)nB. X̄t/2,n-1snC. p/22s2nD. None of the above
2. A confidence interval for a proportion provides: A. A range of plausible values for the population proportion B. The exact proportion in the population C. The variance of the sample D. None of the above
3. A 95% confidence interval for a proportion means: A. We are reasonably confident the true proportion lies within the interval B. The true proportion has a 95% probability of being exactly the sample proportion C. The interval is always correct D. None of the above
4. Confidence intervals for proportions are widely used in: A. Public opinion polling, quality control, and epidemiology B. Astronomy only C. Linguistics exclusively D. None of the above
5. In a sample of 200 voters, 120 support a candidate. The sample proportion is: A. 0.6 B. 0.55 C. 0.65 D. None of the above

3.4 Confidence Intervals for Variance

Confidence intervals can also be constructed for the **variance** of a population. Unlike intervals for the mean or proportion, these rely on the **chi-square distribution**, since the sample variance follows a chi-square distribution under normality assumptions.

Fill in the blank: Confidence intervals for variance are constructed using the ___________________ distribution.

3.4.1 Formula for confidence interval of variance

If s2is the sample variance from a sample of size n, then the confidence interval for the population variance 2is:

$$(n-1)s2/2, n-122(n-1)s21-\alpha/2, n-12$$

Here, /2, n-12and 1-α/2, n-12are critical values from the chi-square distribution with n-1degrees of freedom.

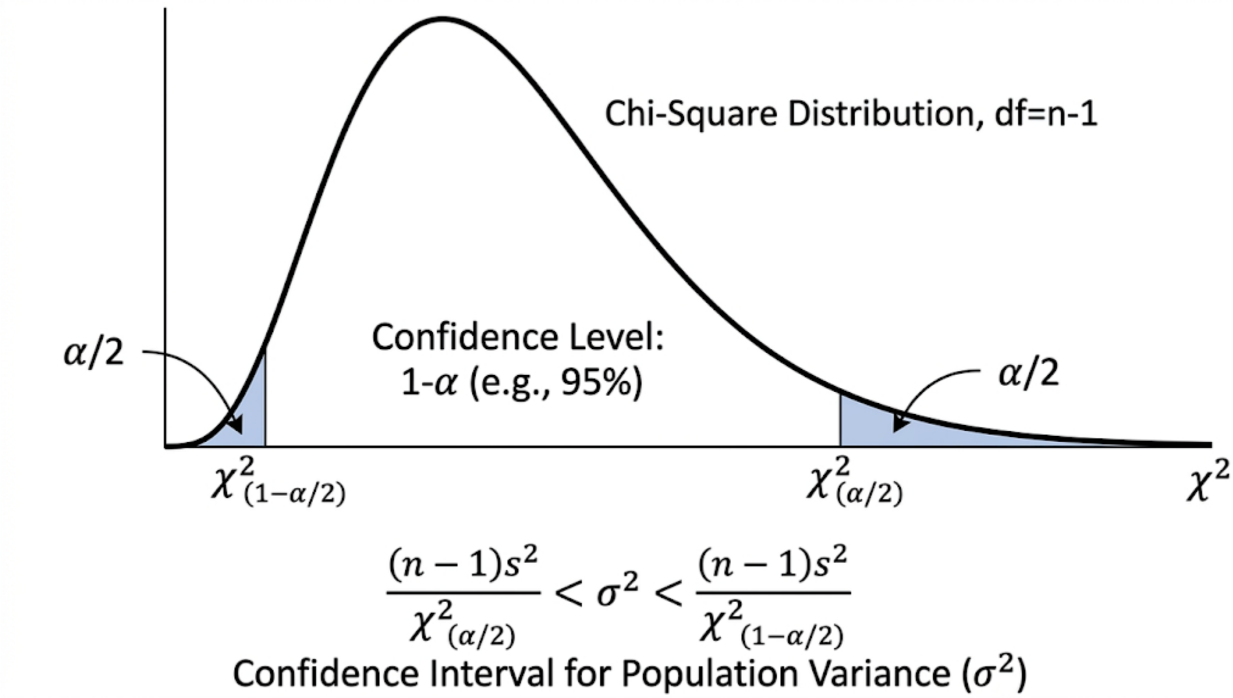

Illustration of confidence interval for variance using chi-square distribution.

3.4.2 Interpretation

This interval provides a range of plausible values for the population variance. For example, if the sample variance of test scores is 25 from a sample of 30 students, a 95% confidence interval can be constructed using chi-square critical values.

Fill in the blank: A confidence interval for variance provides a ____________________ of plausible values for the population variance.

3.4.3 Applications

- **Quality control:** Estimating variability in production processes.
- **Education:** Measuring variability in student performance.
- **Health sciences:** Assessing variability in patient responses to treatment.

Applications of confidence intervals for variance in different fields.

Pilot Questions 3.4

1. Confidence intervals for variance are constructed using: A. The chi-square distribution B. The Z-distribution C. The t-distribution D. None of the above

2. The formula for the confidence interval of variance involves: A. Sample variance and chi-square critical values B. Sample mean and Z critical values C. Sample proportion and normal approximation D. None of the above
3. A confidence interval for variance provides: A. A range of plausible values for the population variance B. The exact variance of the population C. The mean of the sample D. None of the above
4. Confidence intervals for variance are useful in: A. Quality control, education, and health sciences B. Astronomy only C. Linguistics exclusively D. None of the above
5. If the sample variance is 25 from a sample of 30 students, a 95% confidence interval for variance can be constructed using: A. Chi-square critical values B. Z critical values C. t critical values D. None of the above

3.5 Comparative Discussion and Applications

Now that we have studied confidence intervals for the mean, proportions, and variance, it is useful to compare them and highlight their practical applications. Confidence intervals differ in their construction depending on the parameter being estimated, but they all serve the same purpose: to provide a range of plausible values with a specified level of confidence.

Fill in the blank: Confidence intervals differ in their construction depending on the parameter being estimated, but they all provide a __________________ of plausible values with a confidence level.

3.5.1 Comparison across parameters

- **Mean:** Uses the Z-distribution when variance is known, and the t-distribution when variance is unknown.
- **Proportion:** Uses the normal approximation for large samples.
- **Variance:** Uses the chi-square distribution under normality assumptions.

Comparison of confidence intervals for mean, proportion, and variance.

3.5.2 Applications in practice

- **Research:** Confidence intervals are used to report estimates with uncertainty, making results more credible.

- **Policy-making:** Intervals help decision-makers understand the reliability of survey results.
- **Quality control:** Confidence intervals for variance and proportions are applied to monitor production processes.
- **Health sciences:** Confidence intervals for means and proportions are used to evaluate treatment effects and prevalence rates.

Fill in the blank: Confidence intervals are widely applied in research, policy-making, quality control, and __________________ sciences.

Illustration of applications of confidence intervals in research, policy-making, quality control, and health sciences.

Pilot Questions 3.5

1. Confidence intervals for the mean use: A. Z-distribution if variance is known, t-distribution if variance is unknown B. Chi-square distribution always C. Normal approximation only D. None of the above
2. Confidence intervals for proportions typically use: A. Normal approximation for large samples B. Chi-square distribution C. t-distribution D. None of the above
3. Confidence intervals for variance are based on: A. Chi-square distribution B. Z-distribution C. Normal approximation D. None of the above
4. Confidence intervals are applied in: A. Research, policy-making, quality control, and health sciences B. Astronomy only C. Linguistics exclusively D. None of the above
5. The common purpose of all confidence intervals is to: A. Provide a range of plausible values for a parameter with a confidence level B. Give the exact parameter value C. Eliminate uncertainty completely D. None of the above

Series 3 Summary

In this Series, we studied **confidence intervals** as a way to express uncertainty around point estimates. We began with the definition and role of confidence intervals, highlighting how they provide ranges of plausible values with a specified confidence level. We then examined confidence intervals for the **mean**, distinguishing between known and unknown variance cases. Next, we explored confidence intervals for **proportions**, which are widely used in surveys and quality control. We continued with confidence intervals for **variance**, constructed using the chi-square distribution. Finally, we compared confidence intervals across parameters and discussed their applications in research, policy-making, quality control, and health sciences.

Glossary of Terms

- **Confidence interval (CI):** A range of values that likely contains the true parameter, with a specified confidence level.
- **Confidence level:** The proportion of intervals that would contain the true parameter if repeated samples were taken (e.g., 95%).
- **Margin of error:** The amount added and subtracted from the point estimate to form the confidence interval.
- **Z-distribution:** Standard normal distribution used when variance is known.
- **t-distribution:** Distribution used when variance is unknown, accounting for extra uncertainty.
- **Chi-square distribution:** Distribution used to construct confidence intervals for variance.
- **Sample proportion (p):** The fraction of sample observations with a given characteristic.

Concept Check Questions 3

Objective Questions

1. Define a confidence interval and explain its role in statistical inference. (Linked to SLO 3.1)
2. Write the formula for a confidence interval for the mean with known variance. (Linked to SLO 3.2)
3. What distribution is used when variance is unknown in constructing confidence intervals for the mean? (Linked to SLO 3.2)
4. State the formula for a confidence interval for a proportion. (Linked to SLO 3.3)
5. Write the formula for a confidence interval for variance. (Linked to SLO 3.4)

Short-Answer Questions

6. Explain the correct interpretation of a 95% confidence interval. (Linked to SLO 3.1)

7. Give one example of a confidence interval for a proportion in practice. (Linked to SLO 3.3)

8. Describe how confidence intervals for variance are applied in quality control. (Linked to SLO 3.4)

Long-Answer Question

9. Compare confidence intervals for the mean, proportion, and variance, explaining the distributions used and their practical applications. (Linked to SLOs 3.2–3.4)

Answers to Pilot Questions 3

Part 3.1: 1-A, 2-A, 3-A, 4-A, 5-A **Part 3.2:** 1-A, 2-A, 3-A, 4-A, 5-A **Part 3.3:** 1-A, 2-A, 3-A, 4-A, 5-A **Part 3.4:** 1-A, 2-A, 3-A, 4-A, 5-A **Part 3.5:** 1-A, 2-A, 3-A, 4-A, 5-A

References and Attributions 3

- **Figure 3.1:** Confidence intervals from repeated samples. AI prompt: "Generate a diagram showing multiple confidence intervals from repeated samples, with most intervals containing the true parameter."
- **Box 3.1:** Confidence level interpretation. AI prompt: "Generate a box summarising correct interpretation of confidence levels with examples."
- **Figure 3.2:** Confidence interval for mean with known variance. AI prompt: "Generate a diagram showing a normal distribution curve with confidence interval bounds around the sample mean when variance is known."
- **Table 3.1:** Comparison of CI formulas for mean. AI prompt: "Generate a table comparing formulas for confidence intervals for the mean with known variance (Z-distribution) and unknown variance (t-distribution)."
- **Figure 3.3:** Confidence interval for proportion. AI prompt: "Generate a diagram showing a sample proportion with confidence interval bounds using the normal approximation."
- **Table 3.2:** Examples of CI for proportions. AI prompt: "Generate a table showing examples of confidence intervals for proportions in voter surveys and defect rates."

- **Figure 3.4:** Confidence interval for variance. AI prompt: "Generate a diagram showing chi-square distribution curve with confidence interval bounds for variance."
- **Table 3.3:** Applications of CI for variance. AI prompt: "Generate a table showing applications of confidence intervals for variance in quality control, education, and health sciences."
- **Table 3.4:** Comparison of CI across parameters. AI prompt: "Generate a table comparing confidence intervals for mean, proportion, and variance with formulas and distributions used."
- **Figure 3.5:** Applications of confidence intervals. AI prompt: "Generate a diagram showing applications of confidence intervals in research, policy-making, quality control, and health sciences."

4 Hypothesis Testing – Simple and Composite Hypotheses

Series Outline

- **Part 4.1: Introduction to Hypothesis Testing**
  - Concept and role in statistical inference
  - Null and alternative hypotheses
- **Part 4.2: Simple Hypotheses**
  - Definition and examples
  - Testing procedures
- **Part 4.3: Composite Hypotheses**
  - Definition and examples
  - Testing procedures
- **Part 4.4: Errors in Hypothesis Testing**
  - Type I and Type II errors
  - Power of a test
- **Part 4.5: Comparative Discussion and Applications**

- Comparing simple and composite hypotheses
- Practical applications in research and decision-making

Introduction

Hypothesis testing is a fundamental aspect of statistical inference. It provides a structured way to make decisions about population parameters based on sample data. A **hypothesis** is a statement about a parameter, and hypothesis testing involves evaluating whether the data provide sufficient evidence to reject or accept that statement.

In this Series, we focus on **simple and composite hypotheses**. A **simple hypothesis** specifies the parameter completely (e.g., $\mu=50$), while a **composite hypothesis** specifies the parameter only partially (e.g., $\mu>50$). Understanding the distinction between these two types of hypotheses is crucial for applying appropriate testing procedures.

We shall commence this Series by introducing the general concept of hypothesis testing and the role of null and alternative hypotheses. Next, we will study simple hypotheses, followed by composite hypotheses. Subsequently, we will examine errors in hypothesis testing and the concept of test power. Finally, we will conclude with a comparative discussion of simple and composite hypotheses and their applications in practice.

Series Learning Outcomes

Upon completing this Series, you should be able to:

- **SLO 4.1** define hypothesis testing and explain its role in statistical inference,
- **SLO 4.2** describe simple hypotheses and apply testing procedures,
- **SLO 4.3** explain composite hypotheses and apply testing procedures,
- **SLO 4.4** distinguish between Type I and Type II errors and explain the concept of test power, and
- **SLO 4.5** compare simple and composite hypotheses and evaluate their applications in research and practice.

4.1 Introduction to Hypothesis Testing

Hypothesis testing is a cornerstone of statistical inference. It provides a formal framework for making decisions about population parameters based on sample data. A **hypothesis** is essentially a claim or statement about a parameter, and hypothesis testing helps us determine whether the evidence from the sample supports or contradicts that claim.

Fill in the blank: Hypothesis testing provides a __________________ framework for making decisions about population parameters based on sample data.

4.1.1 Concept and definition

A hypothesis test involves two competing statements:

- **Null hypothesis** (H0): The default assumption or claim, often representing no effect or no difference.
- **Alternative hypothesis** (H1): The competing claim, representing an effect, difference, or change.

The goal of hypothesis testing is to use sample data to decide whether to reject H0in favor of H1.

Illustration of hypothesis testing showing null and alternative hypotheses with decision regions.

4.1.2 Role in statistical inference

Hypothesis testing plays a critical role in:

- **Scientific research:** Testing theories and claims.
- **Policy-making:** Evaluating evidence before decisions.
- **Business and industry:** Assessing product quality or market trends.

It provides a structured way to move from data to conclusions, reducing subjectivity in decision-making.

Fill in the blank: Hypothesis testing reduces __________________ in decision-making by providing a structured framework.

4.1.3 Steps in hypothesis testing

1. **State the hypotheses:** Define H0and H1.
2. **Choose a significance level ():** Commonly 0.05 or 0.01.
3. **Select the test statistic:** Depends on the parameter and distribution.
4. **Determine the rejection region:** Based on critical values.

5. **Compute the test statistic from sample data.**
6. **Make a decision:** Reject or fail to reject H0.

Steps in hypothesis testing.

Pilot Questions 4.1

1. Hypothesis testing provides: A. A formal framework for making decisions about population parameters B. Random guesses about data C. Arbitrary conclusions without evidence D. None of the above
2. The null hypothesis (H0) usually represents: A. No effect or no difference B. The presence of an effect C. Always the true parameter D. None of the above
3. The alternative hypothesis (H1) represents: A. An effect, difference, or change B. No effect or no difference C. Always the false parameter D. None of the above
4. A common significance level used in hypothesis testing is: A. 0.05 or 0.01 B. 0.5 or 1.0 C. 5 or 10 D. None of the above
5. The final step in hypothesis testing is to: A. Make a decision to reject or fail to reject H0B. Always accept H1C. Ignore the data D. None of the above

4.2 Simple Hypotheses

A **simple hypothesis** is one that completely specifies the distribution of the population. In other words, it states the parameter(s) of the distribution exactly, leaving no ambiguity. For example, saying that the population mean is exactly 50 (μ=50) is a simple hypothesis because it fully defines the parameter.

Fill in the blank: A simple hypothesis completely __________________ the distribution of the population by specifying the parameter(s) exactly.

4.2.1 Definition and examples

- **Definition:** A hypothesis that specifies the population distribution fully by fixing the parameter(s).
- **Example 1:** H0:μ=100(the mean is exactly 100).
- **Example 2:** H0:p=0.3(the proportion is exactly 0.3).
- **Example 3:** H0:2=25(the variance is exactly 25).

Illustration of a simple hypothesis showing a fixed mean value in a normal distribution.

4.2.2 Testing procedures

Testing a simple hypothesis involves:

1. Stating H0and H1.
2. Choosing a test statistic appropriate for the parameter.
3. Determining the rejection region based on the chosen significance level.
4. Computing the test statistic from sample data.
5. Making a decision to reject or fail to reject H0.

Fill in the blank: Testing a simple hypothesis involves computing a ___________________ from sample data and comparing it to the rejection region.

4.2.3 Applications

- **Manufacturing:** Testing whether the defect rate is exactly a specified value.
- **Education:** Testing whether the average test score is exactly a benchmark.
- **Medicine:** Testing whether the mean response time to a drug is exactly a target value.

Applications of simple hypotheses in different fields.

Pilot Questions 4.2

1. A simple hypothesis is one that: A. Completely specifies the distribution of the population B. Partially specifies the distribution C. Provides no information about the parameter D. None of the above
2. An example of a simple hypothesis is: A. H0:μ=50B. H0:μ>50C. H0:μ50D. None of the above
3. Testing a simple hypothesis involves: A. Computing a test statistic and comparing it to the rejection region B. Guessing the parameter value C. Ignoring the sample data D. None of the above
4. Simple hypotheses are applied in: A. Manufacturing, education, and medicine B. Astronomy only C. Linguistics exclusively D. None of the above

5. A hypothesis that states $p=0.3$is: A. A simple hypothesis B. A composite hypothesis C. Not a hypothesis D. None of the above

4.3 Composite Hypotheses

A **composite hypothesis** is one that does not completely specify the distribution of the population. Instead of fixing the parameter(s) to a single value, it allows them to vary within a range or set. For example, stating that the population mean is greater than 50 ($\mu>50$) is a composite hypothesis because it does not specify the exact value of , only a condition.

Fill in the blank: A composite hypothesis does not completely specify the distribution but allows the parameter(s) to __________________ within a range or set.

4.3.1 Definition and examples

- **Definition:** A hypothesis that specifies the population distribution only partially, leaving the parameter(s) free to vary within certain bounds.
- **Example 1:** $H0:\mu\geq100$(the mean is at least 100).
- **Example 2:** $H0:p0.3$(the proportion is not equal to 0.3).
- **Example 3:** $H0:2>25$(the variance is greater than 25).

Illustration of a composite hypothesis showing a range of possible mean values in a normal distribution.

4.3.2 Testing procedures

Testing a composite hypothesis involves:

1. Stating H0and H1.
2. Choosing the appropriate test statistic (Z, t, chi-square, etc.).
3. Defining the rejection region based on inequalities or ranges.
4. Computing the test statistic from sample data.
5. Making a decision to reject or fail to reject H0.

Fill in the blank: Testing a composite hypothesis often requires defining rejection regions based on __________________ or ranges.

4.3.3 Applications

- **Economics:** Testing whether average income is greater than a benchmark.
- **Medicine:** Testing whether a drug's effect is different from zero.
- **Engineering:** Testing whether the variance in production exceeds a tolerance level.

Applications of composite hypotheses in economics, medicine, and engineering.

Pilot Questions 4.3

1. A composite hypothesis is one that: A. Does not completely specify the distribution of the population B. Completely specifies the distribution C. Provides no information about the parameter D. None of the above
2. An example of a composite hypothesis is: A. $H0:\mu>50$B. $H0:\mu=50$C. $H0:p=0.3$D. None of the above
3. Testing a composite hypothesis involves: A. Defining rejection regions based on inequalities or ranges B. Fixing the parameter to a single value C. Ignoring the sample data D. None of the above
4. Composite hypotheses are applied in: A. Economics, medicine, and engineering B. Astronomy only C. Linguistics exclusively D. None of the above
5. A hypothesis that states p0.3is: A. A composite hypothesis B. A simple hypothesis C. Not a hypothesis D. None of the above

4.4 Errors in Hypothesis Testing

When conducting hypothesis tests, decisions are made based on sample data. Because samples are subject to variability, errors can occur. Two types of errors are particularly important: **Type I** and **Type II** errors. Understanding these errors, along with the concept of **power**, is essential for interpreting hypothesis tests correctly.

Fill in the blank: Hypothesis testing errors occur because decisions are made based on ___________________ data, which is subject to variability.

4.4.1 Type I error

- **Definition:** Rejecting the null hypothesis (H0) when it is actually true.

- **Probability of Type I error:** Denoted by , the significance level of the test.
- **Example:** Concluding that a new drug is effective when in fact it is not.

Fill in the blank: A Type I error occurs when we ___________________ the null hypothesis even though it is true.

4.4.2 Type II error

- **Definition:** Failing to reject the null hypothesis when the alternative hypothesis is true.
- **Probability of Type II error:** Denoted by .
- **Example:** Concluding that a new drug is not effective when in fact it is.

Fill in the blank: A Type II error occurs when we ___________________ to reject the null hypothesis even though the alternative is true.

4.4.3 Power of a test

- **Definition:** The probability of correctly rejecting the null hypothesis when the alternative hypothesis is true.
- **Formula:** Power = $1-\beta$.
- **Importance:** A powerful test is more likely to detect true effects.

Illustration of Type I and Type II errors with rejection regions and power.

4.4.4 Applications

- **Medicine:** Balancing risks of false positives (Type I) and false negatives (Type II) in drug trials.
- **Engineering:** Ensuring product quality by minimizing both types of errors.
- **Social sciences:** Designing surveys and experiments with adequate power to detect meaningful effects.

Applications of Type I and Type II errors in different fields

📊 Applications of Type I and Type II Errors

Domain	Type I Error (False Positive)	Type II Error (False Negative)
Medicine	Diagnosing a patient with a disease when they are healthy. Leads to unnecessary treatment, anxiety, and costs.	Failing to diagnose a patient who actually has the disease. Leads to delayed treatment, worsening health, or death.
Engineering	Concluding a system component is faulty when it is actually fine. Causes unnecessary repairs, downtime, and costs.	Concluding a component is safe when it is actually faulty. Leads to system failure, accidents, or safety hazards.
Social Sciences	Concluding a social program or policy is effective when it is not. Wastes resources and misguides policy decisions.	Concluding a program or policy is ineffective when it actually works. Prevents adoption of beneficial interventions.

.

Pilot Questions 4.4

1. A Type I error occurs when: A. We reject H0even though it is true B. We fail to reject H0even though it is false C. We always accept H1D. None of the above
2. The probability of a Type I error is denoted by: A. B. C. Power D. None of the above
3. A Type II error occurs when: A. We fail to reject H0even though the alternative is true B. We reject H0even though it is true C. We always accept H0D. None of the above
4. The power of a test is defined as: A. 1-βB. C. D. None of the above
5. Type I and Type II errors are important in: A. Medicine, engineering, and social sciences B. Astronomy only C. Linguistics exclusively D. None of the above

4.5 Comparative Discussion and Applications

We have now examined **simple hypotheses** (which fully specify the parameter) and **composite hypotheses** (which allow parameters to vary within a range). To conclude this Series, let's compare them and highlight their practical applications.

Fill in the blank: Simple hypotheses specify parameters exactly, while composite hypotheses allow parameters to __________________ within a range or set.

4.5.1 Comparison of simple and composite hypotheses

- **Simple hypotheses:**
  - Fully specify the distribution.
  - Easier to test because the parameter is fixed.
  - Example: $H_0: \mu = 50$.
- **Composite hypotheses:**
  - Partially specify the distribution.
  - More complex to test because parameters vary.
  - Example: $H_0: \mu > 50$.

📊 Comparison of Simple vs Composite Hypotheses

Aspect	Simple Hypothesis	Composite Hypothesis
Definition	Specifies the distribution completely by fixing all parameters.	Does not specify all parameters; allows a range of possible values.
Formulation	States that a parameter equals a specific value: $H_0: \theta = 0$.	States that a parameter belongs to a set of values: $H_0: \theta \in 0$.
Example (Medicine)	$H_0: \mu = 120$mmHg (mean blood pressure is exactly 120).	$H_0: \mu \leq 120$mmHg (mean blood pressure is less than or equal to 120).
Example (Engineering)	$H_0: p = 0.02$(failure probability is exactly 2%).	$H_0: p \leq 0.02$(failure probability is at most 2%).
Example (Social Sci.)	$H_0: \mu = 50$(average test score is exactly 50).	$H_0: \mu \geq 50$(average test score is at least 50).
Testing Procedure	Uses **exact tests** (e.g., Z-test, t-test) since	Requires **generalized tests** (e.g., likelihood ratio test, chi-square test) to handle parameter ranges.

	distribution is fully specified.	
Complexity	Easier to test; results are more straightforward.	More flexible but harder to test; often requires approximations or asymptotic methods.

Comparison of simple and composite hypotheses.

4.5.2 Applications in practice

- **Simple hypotheses:**
  - Used in controlled experiments where exact parameter values are tested.
  - Example: Testing whether a machine produces items with exactly 2 cm diameter.
- **Composite hypotheses:**
  - Used in real-world scenarios where parameters are not fixed but conditions are tested.
  - Example: Testing whether average income is greater than a benchmark.

Fill in the blank: Composite hypotheses are often used in __________________ scenarios where parameters are not fixed but conditions are tested.

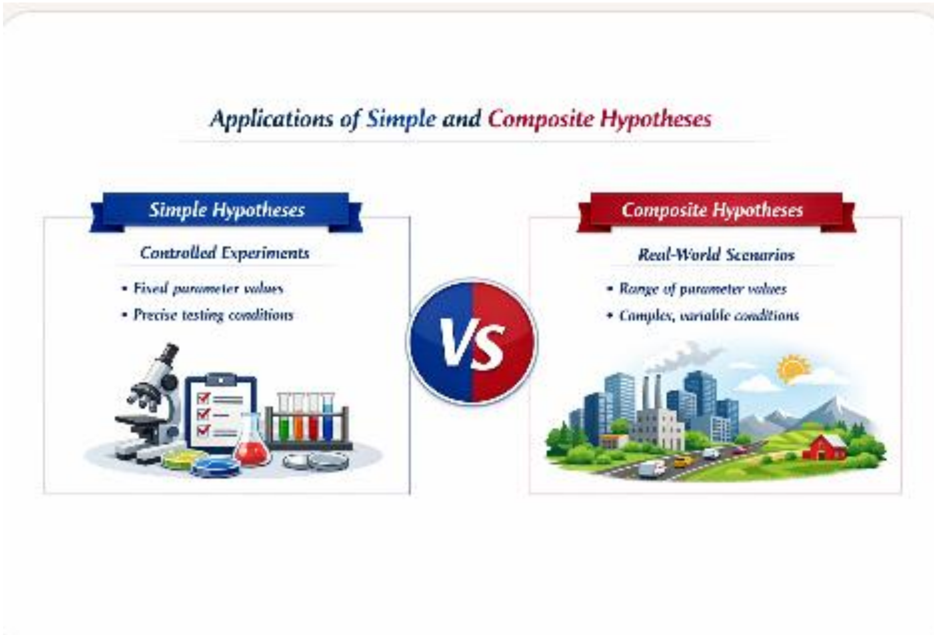

Illustration of applications of simple and composite hypotheses in controlled experiments and real-world scenarios.

Pilot Questions 4.5

1. A simple hypothesis specifies: A. Parameters exactly B. Parameters within a range C. No parameters D. None of the above
2. A composite hypothesis allows: A. Parameters to vary within a range B. Parameters to be fixed exactly C. No testing procedure D. None of the above
3. An example of a simple hypothesis is: A. $H0:\mu=50$B. $H0:\mu>50$C. $H0:p0.3$D. None of the above
4. An example of a composite hypothesis is: A. $H0:\mu>50$B. $H0:\mu=50$C. $H0:p=0.3$D. None of the above
5. Simple hypotheses are often used in: A. Controlled experiments B. Real-world scenarios C. Surveys only D. None of the above

Series 4 Summary

In this Series, we explored **hypothesis testing** with a focus on **simple and composite hypotheses**. We began with an introduction to hypothesis testing, defining null and alternative hypotheses and outlining the steps involved. We then studied **simple hypotheses**, which fully specify the distribution by fixing parameters exactly. Next, we examined **composite hypotheses**, which specify parameters only partially, allowing them to vary within a range. We continued with a discussion of **errors in hypothesis testing**, including Type I and Type II errors, and introduced the concept of **power**. Finally, we compared simple and composite hypotheses and highlighted their applications in controlled experiments and real-world scenarios.

Glossary of Terms

- **Hypothesis testing:** A formal framework for making decisions about population parameters using sample data.
- **Null hypothesis (H0):** The default assumption, often representing no effect or no difference.
- **Alternative hypothesis (H1):** The competing claim, representing an effect or difference.
- **Simple hypothesis:** A hypothesis that fully specifies the distribution by fixing parameters exactly.
- **Composite hypothesis:** A hypothesis that partially specifies the distribution, allowing parameters to vary within a range.
- **Type I error:** Rejecting H0when it is true.
- **Type II error:** Failing to reject H0when H1is true.
- **Power of a test:** The probability of correctly rejecting H0when H1is true, equal to 1-β.

Concept Check Questions 4

Objective Questions

1. Define hypothesis testing and explain its role in statistical inference. (Linked to SLO 4.1)
2. What is the difference between a simple and a composite hypothesis? (Linked to SLOs 4.2 & 4.3)
3. Define Type I and Type II errors. (Linked to SLO 4.4)
4. What is the formula for the power of a test? (Linked to SLO 4.4)
5. Give one example each of a simple and a composite hypothesis. (Linked to SLOs 4.2 & 4.3)

Short-Answer Questions

6. Explain why hypothesis testing reduces subjectivity in decision-making. (Linked to SLO 4.1)

7. Describe the steps involved in testing a simple hypothesis. (Linked to SLO 4.2)

8. Give one application of composite hypotheses in economics. (Linked to SLO 4.3)

Long-Answer Question

9. Compare simple and composite hypotheses, discuss the types of errors in hypothesis testing, and explain the importance of test power. (Linked to SLOs 4.2–4.4)

Answers to Pilot Questions 4

Part 4.1: 1-A, 2-A, 3-A, 4-A, 5-A **Part 4.2:** 1-A, 2-A, 3-A, 4-A, 5-A **Part 4.3:** 1-A, 2-A, 3-A, 4-A, 5-A **Part 4.4:** 1-A, 2-A, 3-A, 4-A, 5-A **Part 4.5:** 1-A, 2-A, 3-A, 4-A, 5-A

References and Attributions 4

- **Figure 4.1:** Null vs alternative hypotheses. AI prompt: "Generate a diagram showing null hypothesis vs alternative hypothesis with rejection region."
- **Box 4.1:** Steps in hypothesis testing. AI prompt: "Generate a box summarising the six steps in hypothesis testing."
- **Figure 4.2:** Simple hypothesis illustration. AI prompt: "Generate a diagram showing a normal distribution curve with mean fixed at a specific value to illustrate a simple hypothesis."
- **Table 4.1:** Applications of simple hypotheses. AI prompt: "Generate a table showing applications of simple hypotheses in manufacturing, education, and medicine."
- **Figure 4.3:** Composite hypothesis illustration. AI prompt: "Generate a diagram showing a normal distribution curve with mean allowed to vary above a threshold to illustrate a composite hypothesis."
- **Table 4.2:** Applications of composite hypotheses. AI prompt: "Generate a table showing applications of composite hypotheses in economics, medicine, and engineering."

- **Figure 4.4:** Type I and Type II errors. AI prompt: "Generate a diagram showing Type I and Type II errors with rejection regions and power in hypothesis testing."
- **Table 4.3:** Applications of Type I and Type II errors. AI prompt: "Generate a table showing applications of Type I and Type II errors in medicine, engineering, and social sciences."
- **Table 4.4:** Comparison of simple and composite hypotheses. AI prompt: "Generate a table comparing simple and composite hypotheses with definitions, examples, and testing procedures."
- **Figure 4.5:** Applications of simple and composite hypotheses. AI prompt: "Generate a diagram showing applications of simple hypotheses in controlled experiments and composite hypotheses in real-world scenarios."

5 Likelihood Ratio Test and Inferences about Means and Variance

Series Outline

- **Part 5.1: Introduction to the Likelihood Ratio Test (LRT)**
  - Concept and definition
  - Role in hypothesis testing
- **Part 5.2: Properties of the Likelihood Ratio Test**
  - General procedure
  - Asymptotic distribution
  - Examples
- **Part 5.3: Inferences about Means using LRT**
  - Testing hypotheses about population mean
  - Applications in practice
- **Part 5.4: Inferences about Variance using LRT**
  - Testing hypotheses about population variance
  - Applications in practice

- **Part 5.5: Comparative Discussion and Applications**
    - Comparing LRT with other tests
    - Practical applications in research and industry

Introduction

The **Likelihood Ratio Test (LRT)** is a powerful method in hypothesis testing. It compares the likelihoods of two competing hypotheses — the null and the alternative — and uses their ratio as the basis for decision-making. The LRT is widely regarded for its generality and optimality properties, especially in large samples.

In this Series, we focus on the LRT and its applications in making inferences about **means** and **variance**. We begin with an introduction to the concept and role of the LRT in hypothesis testing. Next, we study its properties and general procedure. We then apply the LRT to hypotheses about the mean, followed by hypotheses about the variance. Finally, we conclude with a comparative discussion of the LRT and its applications in practice.

Series Learning Outcomes

Upon completing this Series, you should be able to:

- **SLO 5.1** define the likelihood ratio test and explain its role in hypothesis testing,
- **SLO 5.2** describe the properties and general procedure of the LRT,
- **SLO 5.3** apply the LRT to make inferences about population means,
- **SLO 5.4** apply the LRT to make inferences about population variance, and
- **SLO 5.5** compare the LRT with other tests and evaluate its applications in research and practice.

5.1 Introduction to the Likelihood Ratio Test (LRT)

The **Likelihood Ratio Test (LRT)** is a fundamental method in hypothesis testing. It compares how well two competing hypotheses explain the observed data by examining the ratio of their likelihoods. The LRT is widely used because of its generality, optimality properties, and applicability across many statistical models.

Fill in the blank: The likelihood ratio test compares how well two competing __________________ explain the observed data.

5.1.1 Concept and definition

- **Likelihood function:** Measures the probability of observing the given data under specific parameter values.
- **Likelihood ratio:** The ratio of the maximum likelihood under the null hypothesis to the maximum likelihood under the alternative hypothesis.
- **Decision rule:** If the ratio is too small (indicating the null hypothesis explains the data poorly compared to the alternative), we reject the null hypothesis.

Mathematically, the likelihood ratio statistic is:

$$=\sup \theta\in 0 L(\theta) \sup \theta\in L(\theta)$$

where:

- $L(\theta)$is the likelihood function,
- $0$is the parameter space under the null hypothesis,
- is the full parameter space.

5.1.2 Role in hypothesis testing

The LRT plays a crucial role in hypothesis testing because:

- It provides a general framework applicable to many models.
- It is often the most powerful test in large samples (asymptotically optimal).
- It allows comparison of nested models (null vs alternative).

Fill in the blank: The likelihood ratio test is often the most ___________________ test in large samples.

5.1.3 Steps in conducting an LRT

1. **Specify the null and alternative hypotheses.**
2. **Compute the likelihood function** under both hypotheses.
3. **Calculate the likelihood ratio statistic** .
4. **Determine the distribution of the test statistic** (often chi-square asymptotically).
5. **Compare with critical values** and make a decision.

Five Steps in Conducting a Likelihood Ratio Test

Step	Description	Purpose / Formula		
1. Define Hypotheses	Specify the **null hypothesis (H_0)** and **alternative hypothesis (H_1)**.	Example: H0:μ=0vs. H1:μ0.		
2. Calculate Likelihoods	Compute the **likelihood function** under each hypothesis using sample data.	\(L(H_0) = f(x	\theta_0) \), \(L(H_1) = f(x	\hat{\theta}) \).
3. Compute the Likelihood Ratio	Form the ratio of the two likelihoods.	λ=L(H0)L(H1).		
4. Determine the Test Statistic	Transform the ratio into a test statistic for easier comparison.	-2ln⁡(λ)—follows a **χ^2 distribution** under H0.		
5. Make a Decision	Compare the test statistic to the **critical value** or use the **p-value**.	If -2ln⁡(λ)>α,df2, **reject H_0.**		

Steps in conducting a likelihood ratio test.

Pilot Questions 5.1

1. The likelihood ratio test compares: A. How well two competing hypotheses explain the observed data B. The mean and variance directly C. Random guesses about data D. None of the above

2. The likelihood ratio statistic is defined as: A. =sup⁡θ∈0L(θ)sup⁡θ∈L(θ)B. X‾Z/2nC. p±Z/2p(1-p)nD. None of the above
3. The LRT is often the most powerful test in: A. Large samples B. Small samples only C. Random guesses D. None of the above
4. The likelihood function measures: A. The probability of observing the data under specific parameter values B. The variance of the sample C. The mean of the population D. None of the above
5. One key application of the LRT is: A. Comparing nested models B. Estimating variance directly C. Ignoring sample data D. None of the above

5.2 Properties of the Likelihood Ratio Test

The **Likelihood Ratio Test (LRT)** is valued for its generality and optimality. It is applicable across a wide range of statistical models and has strong theoretical properties that make it a preferred choice in many inference problems.

Fill in the blank: The likelihood ratio test is valued for its __________________ and optimality.

5.2.1 General procedure

The LRT works by comparing the maximum likelihood under the null hypothesis with the maximum likelihood under the alternative hypothesis. The ratio forms the basis of the test statistic.

Steps:

1. Compute the likelihood under the null hypothesis (L0).
2. Compute the likelihood under the alternative hypothesis (L1).
3. Form the ratio =L0L1.
4. Use the distribution of -2ln⁡to make a decision.

5.2.2 Asymptotic distribution

- For large samples, the statistic -2ln⁡follows a **chi-square distribution** with degrees of freedom equal to the difference in the number of parameters estimated under the null and alternative hypotheses.
- This property makes the LRT widely applicable, even when exact distributions are complicated.

Fill in the blank: For large samples, the statistic -2ln follows a __________________ distribution.

5.2.3 Examples

- **Mean testing:** Comparing whether the mean is equal to a specified value versus not equal.
- **Variance testing:** Comparing whether the variance equals a specified value versus not equal.
- **Model comparison:** Testing nested models in regression or likelihood-based frameworks.

Pilot Questions 5.2

1. The likelihood ratio test compares: A. Maximum likelihood under the null and alternative hypotheses B. Sample mean and variance directly C. Random guesses about data D. None of the above
2. The likelihood ratio statistic is: A. =L0L1B. X‾Z/2nC. p±Z/2p(1-p)nD. None of the above
3. For large samples, the distribution of -2ln is: A. Chi-square distribution B. Normal distribution C. t-distribution D. None of the above
4. The degrees of freedom in the chi-square distribution for LRT are equal to: A. The difference in the number of parameters estimated under null and alternative hypotheses B. The sample size directly C. Always 1 D. None of the above
5. The LRT can be applied to: A. Mean testing, variance testing, and model comparison B. Only variance testing C. Only regression models D. None of the above

5.3 Inferences about Means using LRT

The **Likelihood Ratio Test (LRT)** can be applied to test hypotheses about the population mean. This is particularly useful when comparing whether the mean equals a specified value or differs from it, under the assumption of normality.

Fill in the blank: The likelihood ratio test can be applied to test hypotheses about the __________________ mean.

5.3.1 Hypotheses about the mean

- **Null hypothesis (H0):** The mean equals a specified value, e.g., $\mu=0$.
- **Alternative hypothesis (H1):** The mean differs from the specified value, e.g., 0.

The LRT compares the likelihood of the data under H0with the likelihood under H1.

5.3.2 LRT procedure for the mean

1. **Specify hypotheses:** H0:$\mu=0$, H1:$\mu 0$.
2. **Compute likelihood under H0:** Using the fixed mean 0.
3. **Compute likelihood under H1:** Using the maximum likelihood estimate (MLE) of , which is the sample mean $\bar{X}$.
4. **Form the likelihood ratio statistic:**

$$=L(0)L(\bar{X})$$

5. **Decision rule:** Reject H0if -2ln exceeds the chi-square critical value (with 1 degree of freedom).

Fill in the blank: The maximum likelihood estimate of the mean is the __________________ mean.

5.3.3 Example

- Suppose we have a sample of exam scores with mean $\bar{X}=70$.
- We test H0:$\mu=65$versus H1:$\mu 65$.
- The LRT compares the likelihood of the data under $\mu=65$with the likelihood under $\mu=70$.
- If the ratio is too small, we reject H0.
-

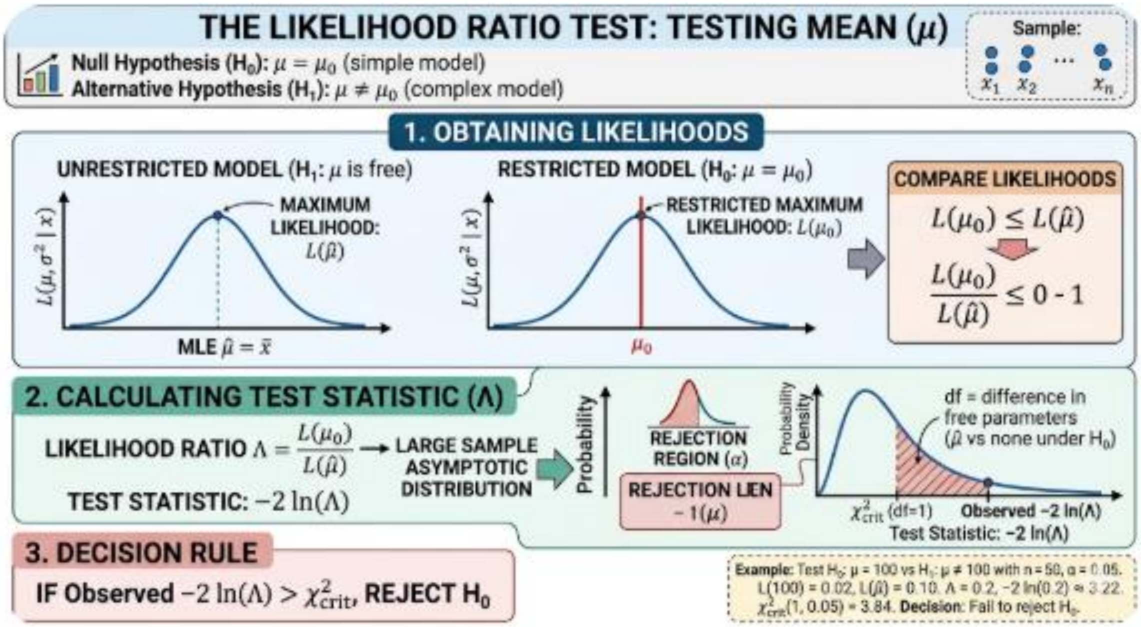

Illustration of LRT applied to testing the mean.

5.3.4 Applications

- **Education:** Testing whether average test scores meet a benchmark.
- **Economics:** Testing whether average income equals a policy target.
- **Medicine:** Testing whether average recovery time equals a standard.

Applications of LRT for means in education, economics, and medicine.

Field		**Application Scenario**	**Null vs Alternative Hypothesis (Means)**	**How LRT is Used**	**Interpretation of Result**
Education		Comparing average test scores between two teaching methods	H_0: $\mu_1 = \mu_2$ (no difference in mean scores) H_1: $\mu_1 \neq \mu_2$	LRT compares likelihood of observed scores under	Rejecting H_0 suggests one teaching method significantly affects

				equal means vs different means	performance
Education		Evaluating intervention programs (e.g., tutoring impact)	$H_0: \mu = \mu_0$ (no improvement) $H_1: \mu > \mu_0$	Uses sample mean vs benchmark to compute likelihood ratio	Significant result indicates intervention improves outcomes
Economics		Testing average income differences between regions	$H_0: \mu_1 = \mu_2$ $H_1: \mu_1 \neq \mu_2$	Compares likelihood under pooled vs separate mean models	Rejecting H_0 implies regional income disparity
Economics		Assessing policy impact on average consumption	H_0: $\mu\_before = \mu\_after$ H_1: $\mu\_before \neq \mu\_after$	LRT evaluates whether policy shifts mean consumption	Significant result suggests policy effect
Economics		Comparing mean returns of two investments	$H_0: \mu_1 = \mu_2$ $H_1: \mu_1 > \mu_2$	Likelihoods computed under constrained vs unconstrained means	Rejecting H_0 supports better-performing investment
Medicine		Comparing mean blood pressure	H_0: $\mu\_treat = \mu\_control$ H_1: μ_treat	LRT evaluates if treatment	Significant result indicates treatment

		between treatment and control groups	≠ μ_control	shifts mean outcome	effectiveness
Medicine		Testing drug effect on mean recovery time	$H_0: \mu = \mu_0$ (standard recovery time) $H_1: \mu < \mu_0$	Likelihood ratio compares observed mean vs standard	Rejecting H_0 shows faster recovery due to drug
Medicine		Evaluating biomarker levels across patient groups	$H_0: \mu_1 = \mu_2 = \mu_3$ H_1: at least one mean differs	LRT extends to multiple groups (ANOVA-like setting)	Significant result suggests variation across condition

Pilot Questions 5.3

1. The LRT can be applied to test hypotheses about: A. The population mean B. The population variance only C. Random guesses D. None of the above
2. The null hypothesis for testing the mean is usually: A. μ=0B. μ>0C. μ<0D. None of the above
3. The maximum likelihood estimate of the mean is: A. The sample mean B. The population mean directly C. The variance D. None of the above
4. The LRT statistic for the mean is: A. =L(0)L(X¯)B. X¯Z/2nC. p±Z/2p(1-p)nD. None of the above
5. Applications of LRT for means include: A. Education, economics, and medicine B. Astronomy only C. Linguistics exclusively D. None of the above

5.4 Inferences about Variance using LRT

The **Likelihood Ratio Test (LRT)** can also be applied to test hypotheses about the population variance. This is important in contexts where variability itself is of interest, such as quality control, risk assessment, or medical studies.

Fill in the blank: The likelihood ratio test can be applied to test hypotheses about the population __________________.

5.4.1 Hypotheses about variance

- **Null hypothesis (**H0**):** The variance equals a specified value, e.g., 2=02.
- **Alternative hypothesis (**H1**):** The variance differs from the specified value, e.g., 202.

5.4.2 LRT procedure for variance

1. **Specify hypotheses:** H0:2=02, H1:202.
2. **Compute likelihood under** H0**:** Using the fixed variance 02.
3. **Compute likelihood under** H1**:** Using the maximum likelihood estimate (MLE) of variance, which is the sample variance s2.
4. **Form the likelihood ratio statistic:**

=L(02)L(s2)

5. **Decision rule:** Reject H0if -2ln exceeds the chi-square critical value (with 1 degree of freedom).

Fill in the blank: The maximum likelihood estimate of variance is the __________________ variance.

5.4.3 Example

- Suppose we have a sample of product weights with variance s2=4.
- We test H0:2=5versus H1:25.
- The LRT compares the likelihood of the data under 2=5with the likelihood under 2=4.
- If the ratio is too small, we reject H0.

5.4.4 Applications

- **Manufacturing:** Testing whether variability in product dimensions meets tolerance standards.

- **Finance:** Testing whether variance in returns equals a benchmark risk level.
- **Medicine:** Testing whether variability in patient responses matches expected levels.

Pilot Questions 5.4

1. The LRT can be applied to test hypotheses about: A. The population variance B. The population mean only C. Random guesses D. None of the above
2. The null hypothesis for testing variance is usually: A. 2=02B. 2>02C. 2<02D. None of the above
3. The maximum likelihood estimate of variance is: A. The sample variance B. The population variance directly C. The mean D. None of the above
4. The LRT statistic for variance is: A. =L(02)L(s2)B. X‾Z/2nC. p±Z/2p(1-p)nD. None of the above
5. Applications of LRT for variance include: A. Manufacturing, finance, and medicine B. Astronomy only C. Linguistics exclusively D. None of the above

5.5 Comparative Discussion and Applications

We have now applied the **Likelihood Ratio Test (LRT)** to hypotheses about both the **mean** and the **variance**. To conclude this Series, let's compare the LRT with other testing methods and highlight its practical applications.

Fill in the blank: The likelihood ratio test is often compared with other tests because of its ___________________ and generality.

5.5.1 Comparison with other tests

- **Z-test / t-test:**
    - Designed specifically for testing means.
    - Easier to compute but less general than LRT.
- **Chi-square test:**
    - Designed for variance and categorical data.
    - LRT generalizes this approach by comparing likelihoods.

- **LRT:**
  - Applicable to both mean and variance.
  - Provides a unified framework for hypothesis testing.
  - Asymptotically optimal in large samples.

Comparison of LRT with Z-test, t-test, and chi-square test.

Criterion	Likelihood Ratio Test (LRT)	Z-test	t-test	Chi-square test
Scope	Very general; applies to a wide class of statistical models (nested models, parametric families)	Narrow; tests hypotheses about means or proportions with known variance	Moderate; tests means when variance is unknown (small samples)	Tests relationships between categorical variables or goodness-of-fit
Underlying Idea	Compares likelihoods under null and alternative hypotheses	Uses standard normal distribution to test standardized statistic	Uses Student's t distribution for small samples	Compares observed vs expected frequencies
Typical Application	Model comparison, parameter restrictions, regression, ANOVA-	Large-sample mean or proportion testing	Small-sample mean comparisons (one-sample, two-	Independence tests, goodness-of-fit, contingency tables

	type problems		sample, paired)	
Type of Data	**Continuous or discrete (very flexible)**	**Continuous or binary (proportions)**	**Continuous**	**Categorical (counts/frequencies)**
Assumptions	**Depends on model (often assumes distribution like normal, exponential, etc.)**	**Normality or large sample (CLT), known variance**	**Normality, unknown variance**	**Expected frequencies sufficiently large, independence**
Test Statistic Distribution	**Asymptotically follows Chi-square distribution**	**Standard normal (Z)**	**t distribution**	**Chi-square distribution**
Optimality	**Asymptotically optimal (Neyman–Pearson framework); efficient for large samples**	**Optimal for large samples when variance known**	**Best unbiased test for small samples under normality**	**Good for categorical data; not optimal for continuous mean testing**
Flexibility	**Very high; extends to complex models**	**Low**	**Moderate**	**Low (specific to categorical data)**

	(e.g., regression, GLMs)			
Relationship to LRT	**— (baseline method)**	**Special case of LRT under normality and known variance**	**Special case of LRT under normality with unknown variance**	**Often derived from or related to LRT for multinomial models**
When Preferred	**When comparing nested models or working in a general modeling framework**	**Large samples with known variance or proportions**	**Small samples with unknown variance**	**Categorical data analysis**

5.5.2 Applications in practice

- **Economics:** Testing whether average income or variance in income meets policy targets.
- **Medicine:** Testing whether mean recovery time or variability in patient responses matches expectations.
- **Engineering:** Testing whether mean product dimensions and variance in quality meet standards.
- **Social sciences:** Testing hypotheses about mean survey responses and variability across groups.

Fill in the blank: The likelihood ratio test is widely applied in economics, medicine, engineering, and ___________________ sciences.

Pilot Questions 5.5

1. The LRT is compared with other tests because of its: A. Optimality and generality B. Simplicity only C. Randomness D. None of the above
2. The Z-test and t-test are designed specifically for: A. Testing means B. Testing variance C. Random guesses D. None of the above
3. The chi-square test is designed for: A. Variance and categorical data B. Means only C. Random guesses D. None of the above
4. The LRT provides: A. A unified framework for hypothesis testing B. Only variance testing C. No practical applications D. None of the above
5. Applications of LRT include: A. Economics, medicine, engineering, and social sciences B. Astronomy only C. Linguistics exclusively D. None of the above

Series 5 Summary

In this Series, we studied the **Likelihood Ratio Test (LRT)** and its applications in making inferences about **means** and **variance**. We began with an introduction to the LRT, defining the likelihood function and the likelihood ratio statistic. We then examined its properties, including the general procedure and asymptotic chi-square distribution. Next, we applied the LRT to hypotheses about the mean, followed by hypotheses about the variance. Finally, we compared the LRT with other tests (Z-test, t-test, chi-square) and discussed its wide applications in economics, medicine, engineering, and social sciences.

Glossary of Terms

- **Likelihood function:** A function that measures the probability of observing the data under specific parameter values.
- **Likelihood ratio statistic ():** Ratio of maximum likelihood under the null hypothesis to maximum likelihood under the alternative hypothesis.
- **Likelihood Ratio Test (LRT):** A hypothesis test based on comparing likelihoods under null and alternative hypotheses.
- **Asymptotic distribution:** For large samples, -2ln follows a chi-square distribution.
- **Maximum likelihood estimate (MLE):** The parameter value that maximizes the likelihood function.

- **Inference about mean:** Using LRT to test whether the population mean equals a specified value.
- **Inference about variance:** Using LRT to test whether the population variance equals a specified value.

Concept Check Questions 5

Objective Questions

1. Define the likelihood ratio test and explain its role in hypothesis testing. (Linked to SLO 5.1)
2. State the formula for the likelihood ratio statistic. (Linked to SLO 5.1)
3. What is the asymptotic distribution of -2ln[illegible]? (Linked to SLO 5.2)
4. What is the MLE of the mean? (Linked to SLO 5.3)
5. What is the MLE of the variance? (Linked to SLO 5.4)

Short-Answer Questions

6. Explain why the LRT is considered asymptotically optimal. (Linked to SLO 5.2)

7. Give one example of applying LRT to test a mean in practice. (Linked to SLO 5.3)

8. Give one example of applying LRT to test a variance in practice. (Linked to SLO 5.4)

Long-Answer Question

9. Compare the LRT with Z-test, t-test, and chi-square test, and explain its applications in economics, medicine, engineering, and social sciences. (Linked to SLO 5.5)

Answers to Pilot Questions 5

Part 5.1: 1-A, 2-A, 3-A, 4-A, 5-A **Part 5.2:** 1-A, 2-A, 3-A, 4-A, 5-A **Part 5.3:** 1-A, 2-A, 3-A, 4-A, 5-A **Part 5.4:** 1-A, 2-A, 3-A, 4-A, 5-A **Part 5.5:** 1-A, 2-A, 3-A, 4-A, 5-A

References and Attributions 5

- **Box 5.1:** Steps in conducting LRT. AI prompt: “Generate a box summarising the five steps in conducting a likelihood ratio test.”

- **Table 5.1:** Examples of LRT applications. AI prompt: "Generate a table showing examples of likelihood ratio test applications in mean testing, variance testing, and model comparison."
- **Figure 5.1:** LRT applied to mean. AI prompt: "Generate a diagram showing likelihood ratio test applied to testing whether the mean equals a specified value."
- **Table 5.2:** Applications of LRT for means. AI prompt: "Generate a table showing applications of likelihood ratio test for means in education, economics, and medicine."
- **Figure 5.2:** LRT applied to variance. AI prompt: "Generate a diagram showing likelihood ratio test applied to testing whether variance equals a specified value."
- **Table 5.3:** Applications of LRT for variance. AI prompt: "Generate a table showing applications of likelihood ratio test for variance in manufacturing, finance, and medicine."
- **Table 5.4:** Comparison of LRT with other tests. AI prompt: "Generate a table comparing likelihood ratio test with Z-test, t-test, and chi-square test in terms of scope, application, and optimality."
- **Figure 5.3:** Applications of LRT. AI prompt: "Generate a diagram showing applications of likelihood ratio test in economics, medicine, engineering, and social sciences."