Maths & Stats Pre-Sessional

Estimation and Hypothesis Testing

Lecturer: Claudio Vallar
School of Economics and Finance

Estimation and Hypothesis Testing

In this session:

- We will review some basic statistical concepts in inferential statistics
- How can we infer from a random sample drawn from a population the population parameters?

For more extensive reading, refer to Chapter 6, 7 and 9 of Newbold, P., Carlson, W., and Thorne, B. (2010). Statistics for Business and Economics, Pearson, 7^{th} Edition

Maths & Stats Pre-Sessional

Z-Scores

Lecturer: Claudio Vallar
School of Economics and Finance

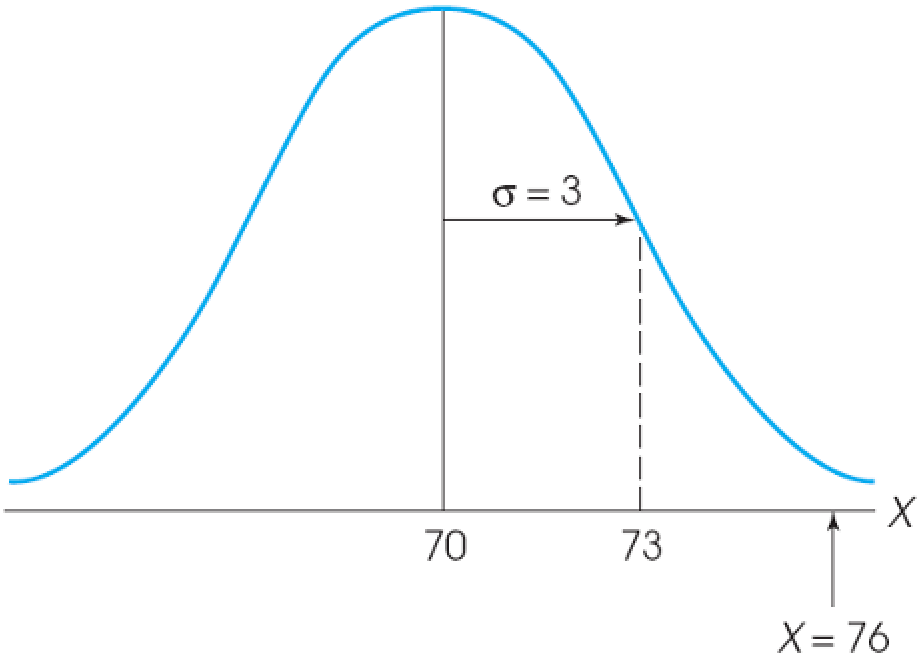
Overview of Z-Scores

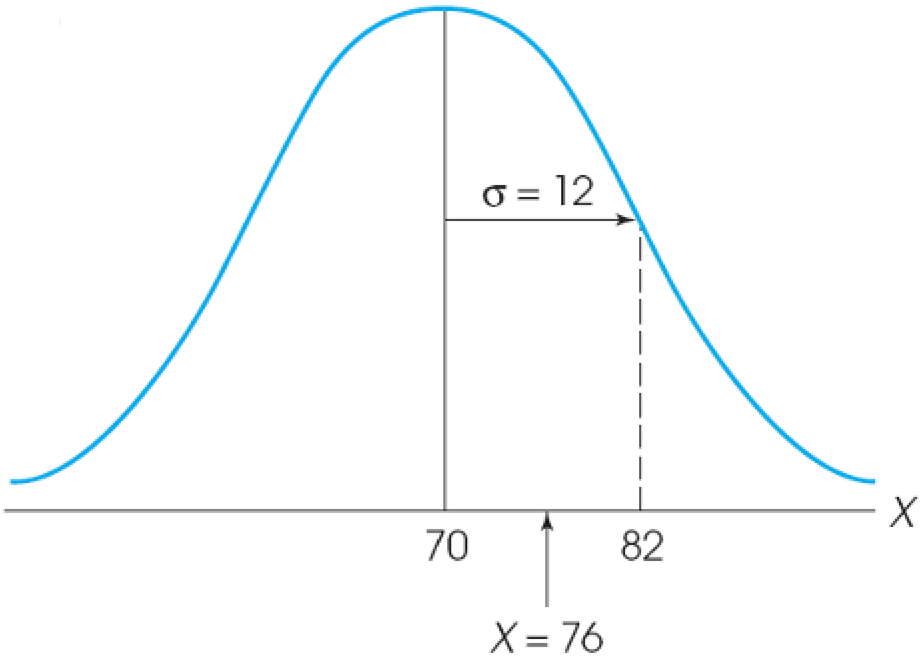
- A Student earned a mark of 76 on an exam
- How does a mark of 76 compare to other students?
  - 76 the lowest mark in the class?
  - Anyone earn a mark higher than 76?

Z - Score

Z-Score -> standardized value that specifies the exact location of an X value within a distribution by describing its distance from the mean in terms of standard deviation units.

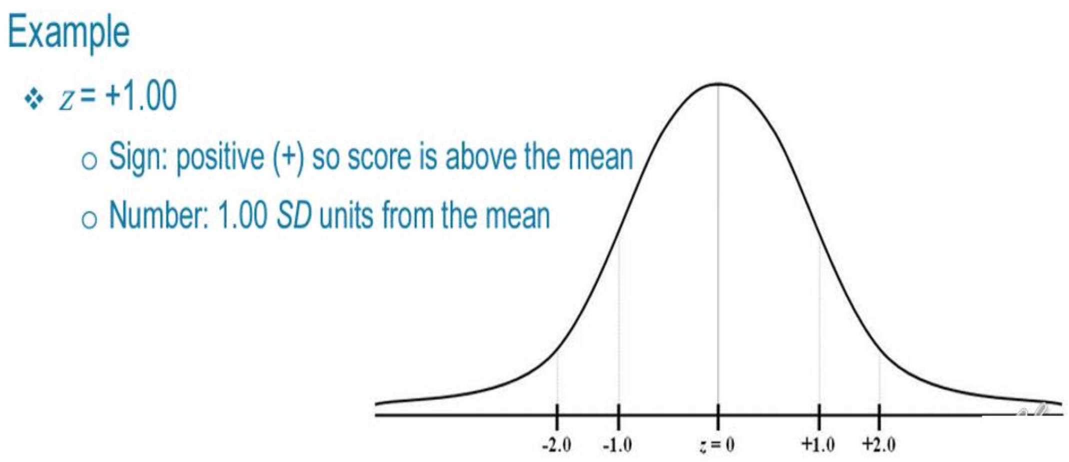
Overview of Z-Scores

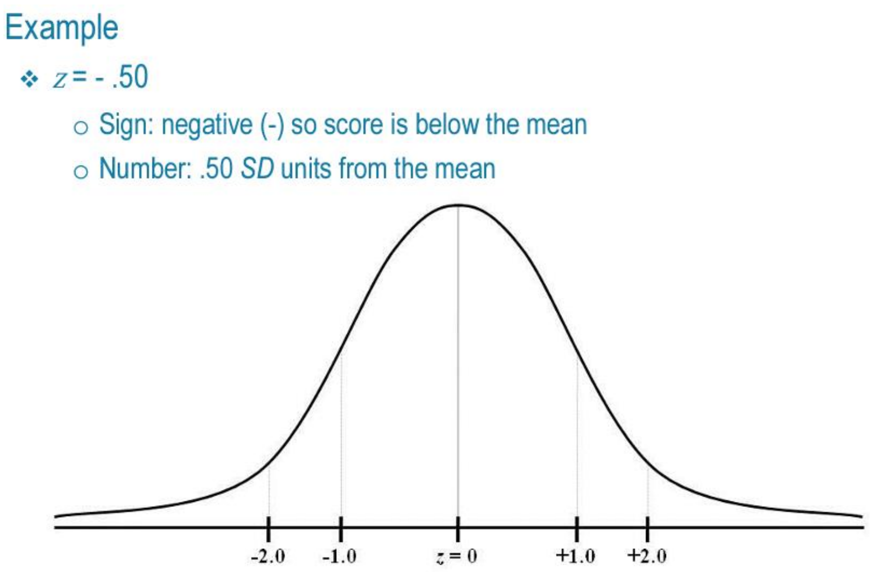




Overview of Z-Scores

- z-Scores describe the exact location of a score within a distribution
  - Sign: Whether score is above (+) or below (-) the mean
  - Number: Distance between score and mean in standard deviation units
  - Standard Deviation Unit: Standardized value(i.e., 1 SD unit = value of 1 SD before standardization)





Overview of Z-Scores

- How to transform a X value to z-Score:

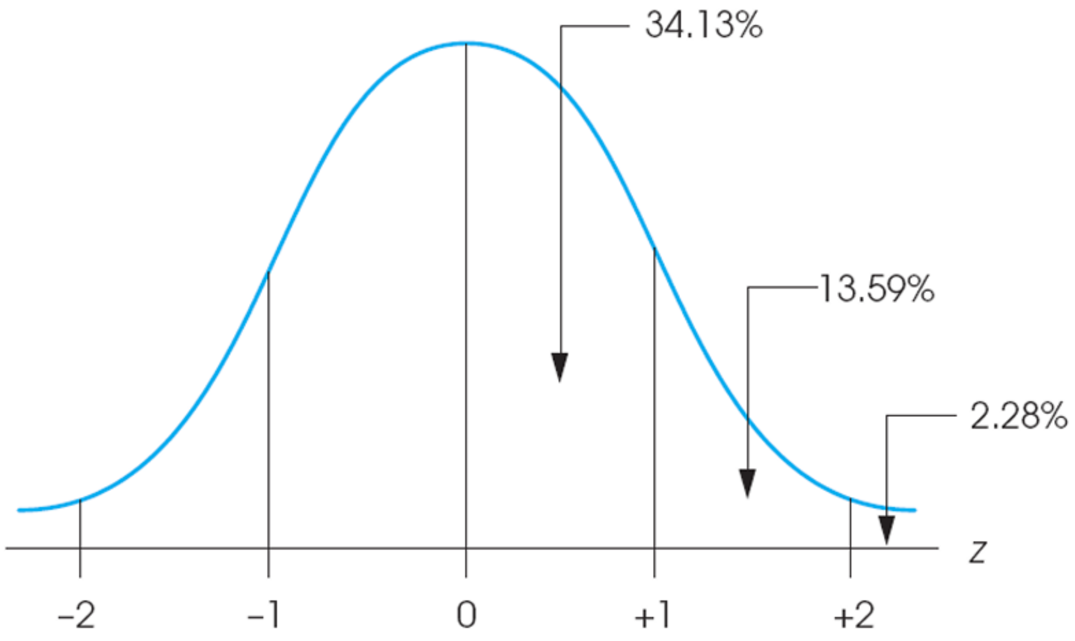
$$z = \left(\frac{X - \mu}{\sigma} \right)$$

- They will produce standardized distributions: distribution composed of scores that have been transformed to create predetermined values for μ and σ; distributions used to make dissimilar distributions comparable.

- Characteristics
  - Same shape as original distribution
  - Mean will always equal zero (0)
  - Standard deviation will always equal one (1)

Overview of Z-Scores

- Advantages:
  - Possible to compare scores or individuals from different distributions
  - Results more generalizable
  - z-Score distributions have equal means (0) and standard deviations (1)



Probability and Normal Distribution

- Example
- p (X > 80) = ?
- Translate into a proportion question: Out of all possible marks, what proportion consists of values greater than 80”?
- The set of “all possible marks” is the population distribution
- We are interested in all the marks greater than 80”, so we shade in the area of the graph to the right of where 80” falls on the distribution

Probability and Normal Distribution

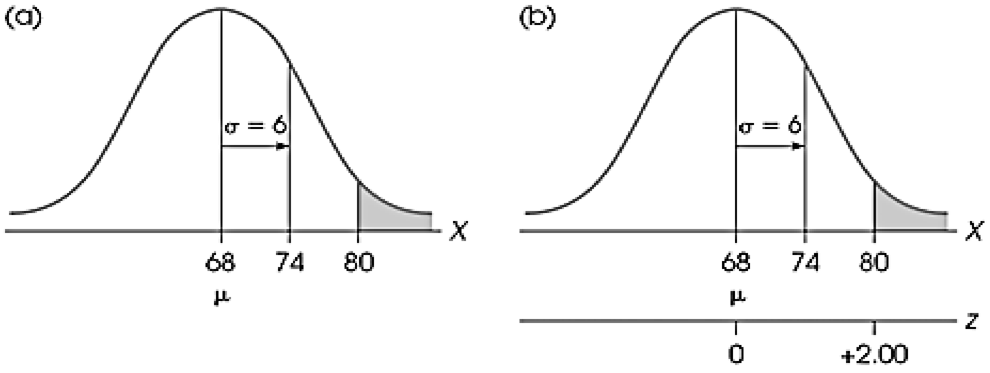
Example (continued)

- Transform $X = 80$ to a z-score

$$z = (X - \mu) / \sigma = (80 - 68) / 6 = 12 / 6 = 2.00$$

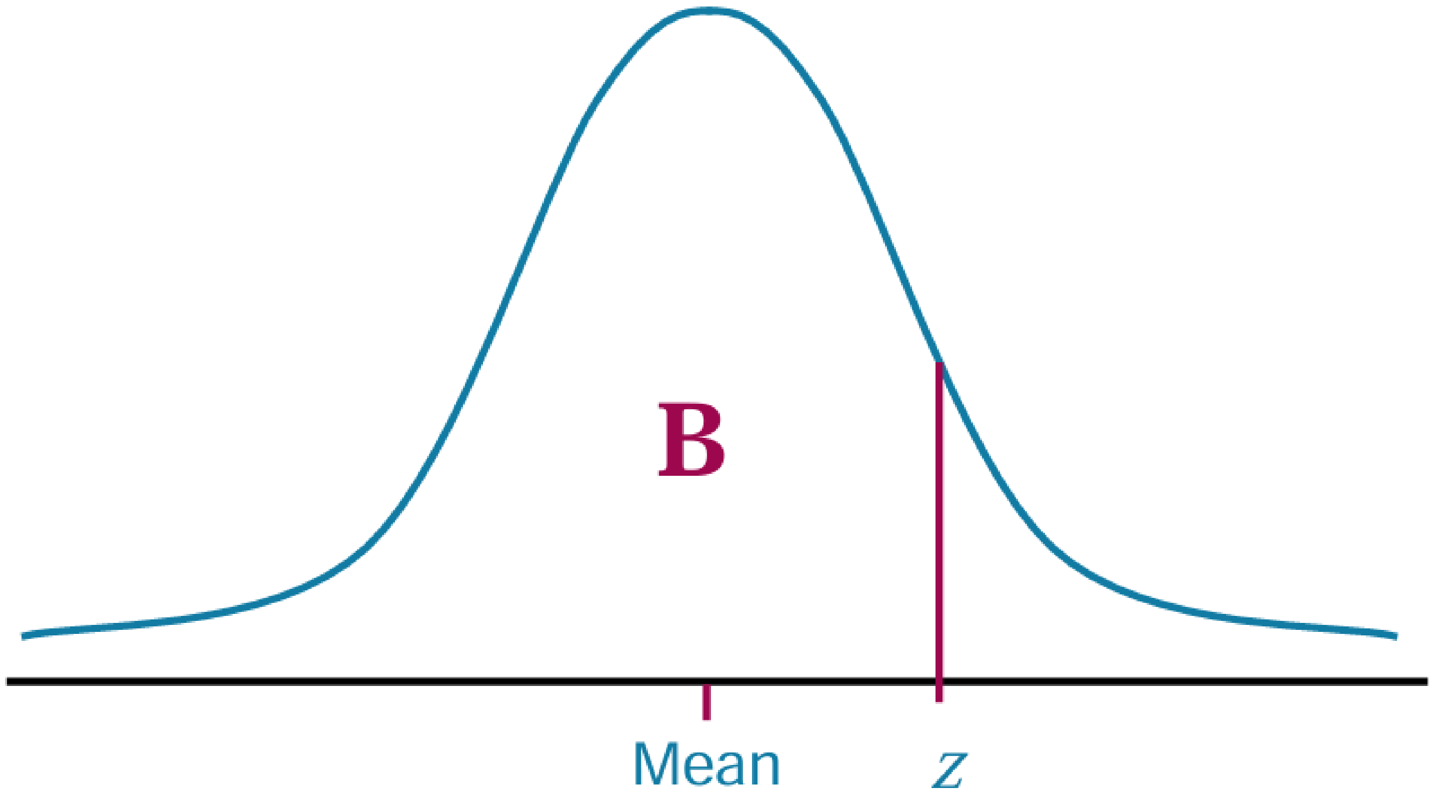
- Express the proportion we are trying to find in terms of the z-score: $p(z > 2.00) = ?$

- By Figure 6.4, $p(X > 80) = p(z > +2.00) = 2.28\%$

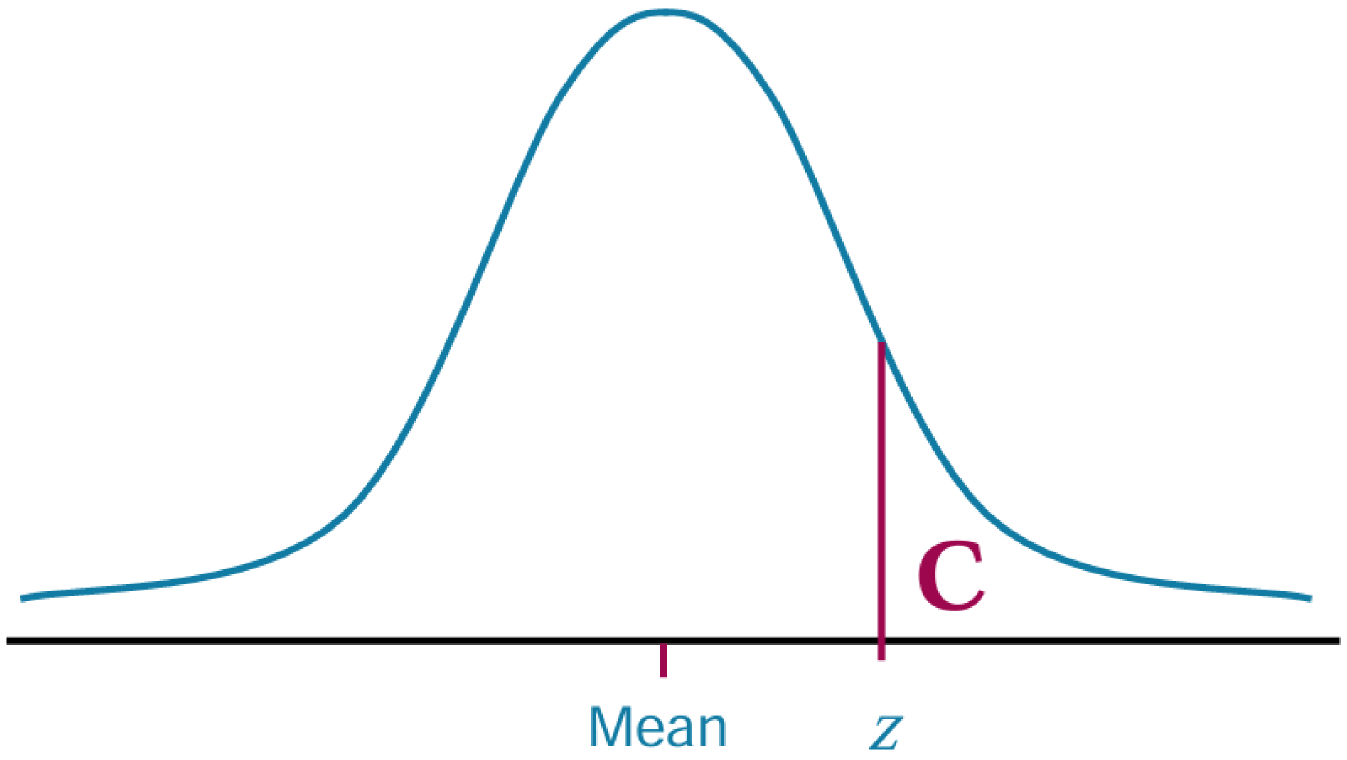


Probability and Normal Distribution

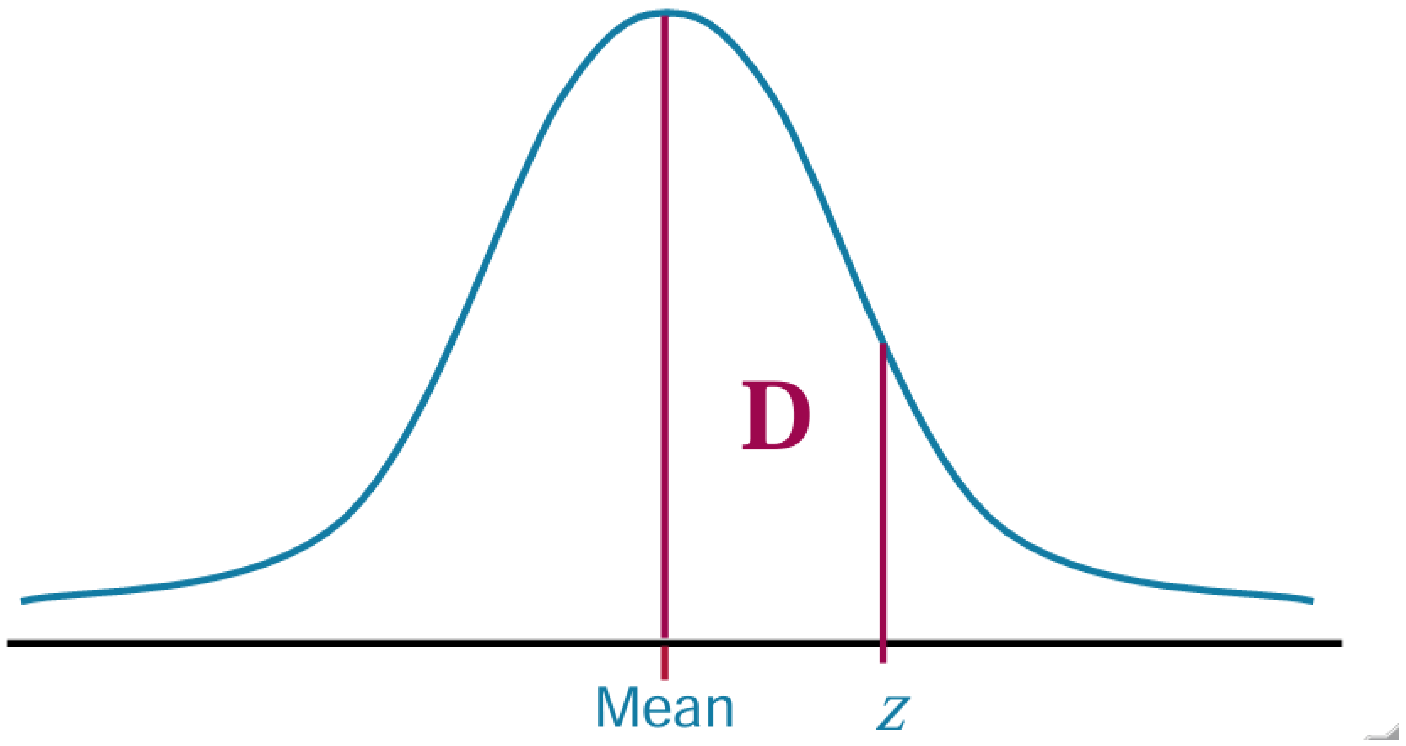
- How to calculate Probabilities?



Probability and Normal Distribution



Probability and Normal Distribution



Overview of Z-Scores

- Example:
- Assume a normal distribution with $\mu = 58$ and $\sigma = 10$ for average speed of cars on a section of interstate highway.
- What proportion of cars traveled between 55 and 65 miles per hour?

$$p(55 < X < 65) = ?$$

- What proportion of cars traveled between 65 and 75 miles per hour?

$$p(65 < X < 75) = ?$$

Step 1: Convert X values to z-Scores

Step 2: Use Unit Normal Table to convert z-scores to corresponding proportions

Maths & Stats Pre-Sessional

Statistical Inference

Lecturer: Claudio Vallar
School of Economics and Finance

Statistical Inference

Statistical inference draws conclusions for the population value of the parameter of interest by using information from a sample.

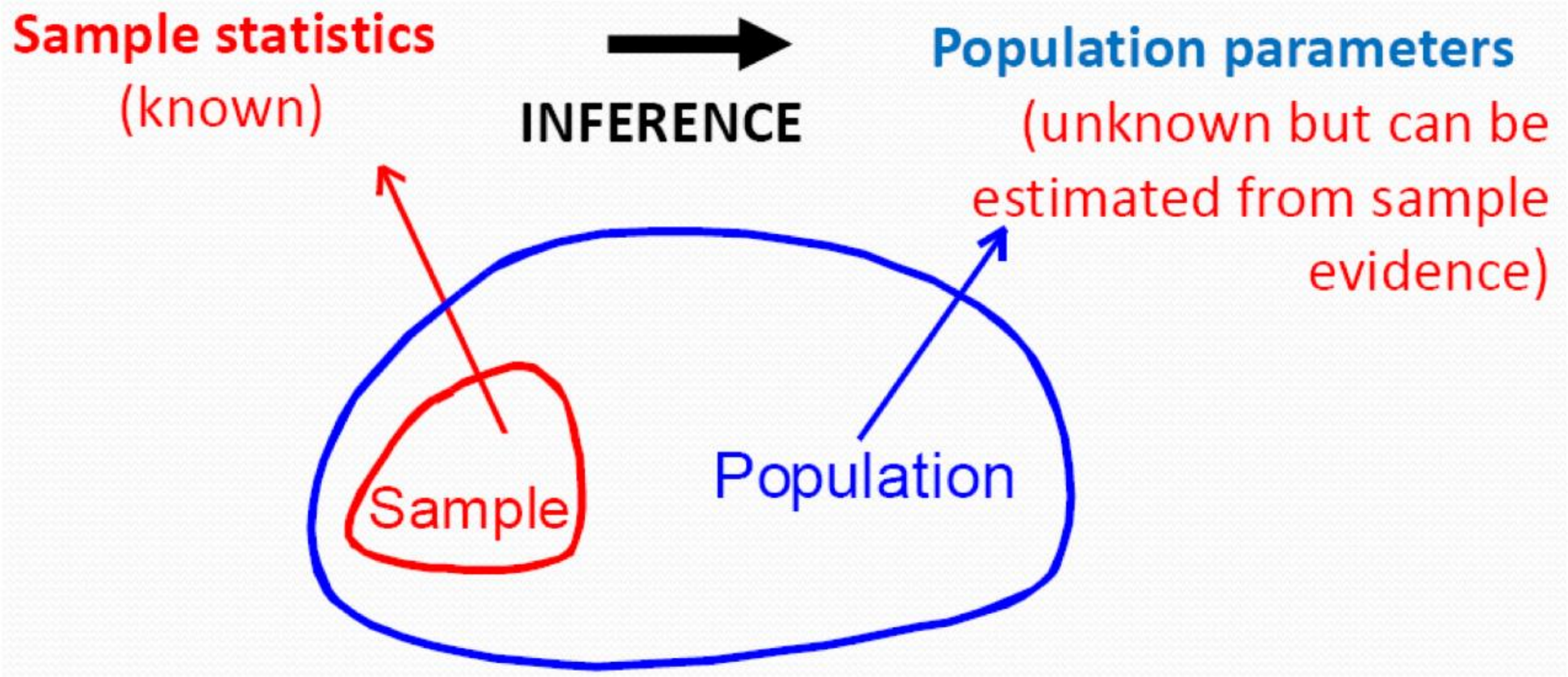
We use statistical inference because we are interested in:

- population **moments** of the distribution: e.g., the population mean and the population variance
- obtaining from our sample a single value, i.e., a **point estimate** for the parameter of interest, or rather a range, i.e., an interval estimate.
- **testing hypotheses** about the population under investigation.

  E.g., $consumption_i = \alpha + \beta income_i + \varepsilon_i$

Statistical Inference

Make inference about the population by examining sample results.



Estimation

To implement a model, we need to know its parameters. However, **parameters are unknown**. We need to estimate them by using a sample

- **Sample**: A collection $(x_{1,} x_{2,} \dots, x_n)$ of observations of the variable X.
- **Estimator $\widehat{\boldsymbol{\theta}}$**: A function of the sample values:

$$\hat{\theta} = f(x_{1,} x_{2,} \dots, x_n)$$

  Note that the estimator $\hat{\theta}$ is a random variable that depends on the sample information and its value provides approximations of this unknown parameter.

  Its distribution is called sampling distribution

- **Estimate:** The particular numerical value taken by the estimator.

Estimation: Example

Consider a population parameter such as the population mean μ.

- An **estimator** of a population parameter is a function of the sample information that produces a single number called a point estimate. For example, the sample mean $\bar{X}$ is an estimator of the population mean.

- The value that $\bar{X}$ assumes for a given set of data is called the **point estimate**, $\bar{x}$.

Maths & Stats Pre-Sessional

Estimators and their Properties

Lecturer: Claudio Vallar
School of Economics and Finance

Estimators Properties

Finite sample properties: they hold for a sample of any size

- **Unbiasedness**
- **Efficiency**

Asymptotic properties: they hold when the sample size grows without bound

- **Consistency**

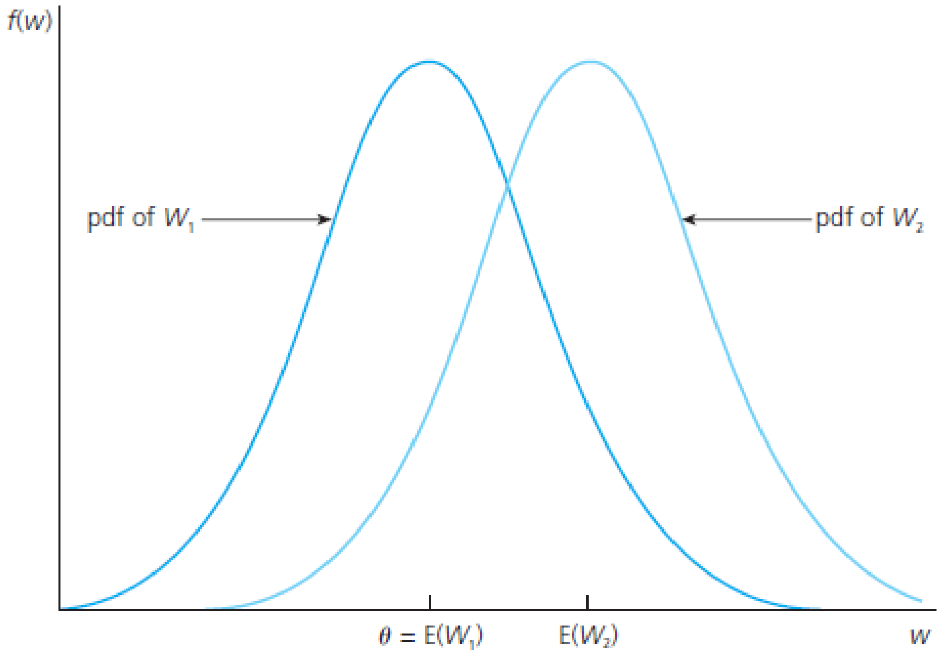
Estimator - Unbiasedness

An estimator $\hat{\theta}$ for the parameter θ is **unbiased** if:

$$E(\hat{\theta}) = \theta$$

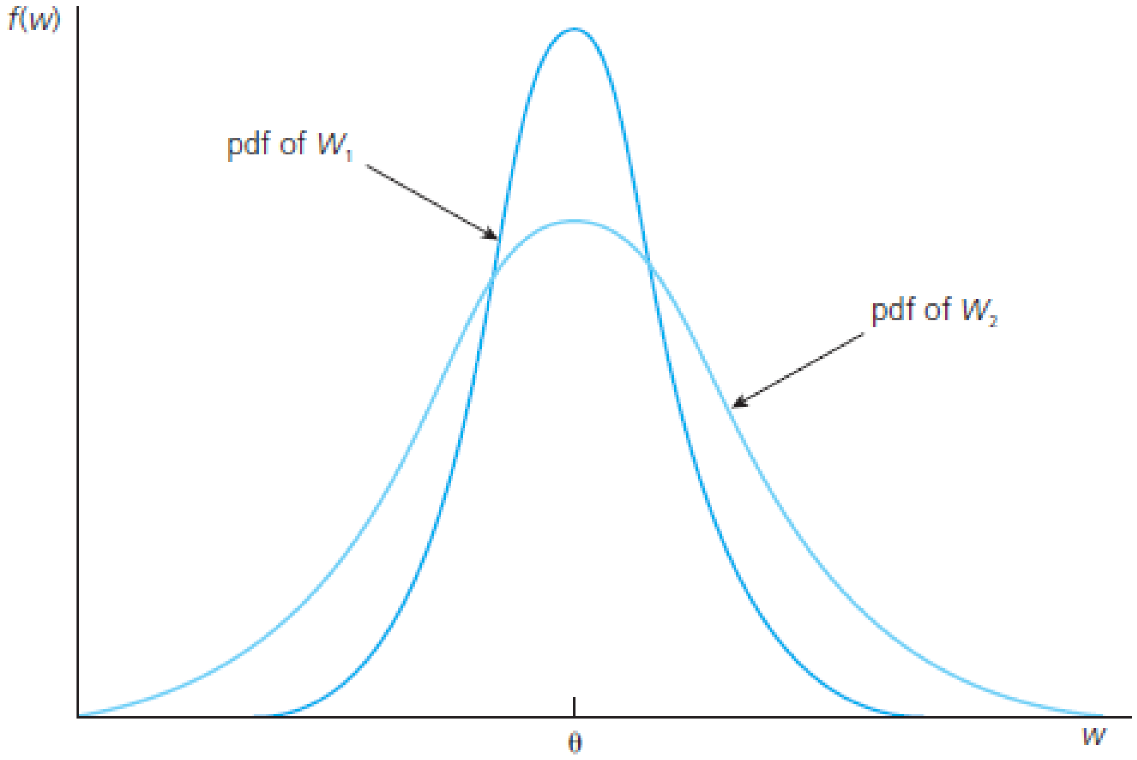
- If an estimator is unbiased, then its probability distribution has an expected value equal to the parameter it is supposed to estimate.
- If we drew infinitely many samples and computed an estimate for each sample, the average of all these estimates would give the true value of the parameter.
- If the estimator is biased, then:

$$bias(\hat{\theta}) = E(\hat{\theta}) - \theta$$



Estimator - Efficiency

- Suppose $\hat{\theta}_1$ and $\hat{\theta}_2$ are two unbiased estimators for θ. Then, $\hat{\theta}_1$ is **efficient** relative to $\hat{\theta}_2$ if Var $(\hat{\theta}_1) \leq$ Var $(\hat{\theta}_2)$ for all θ, with strict inequality for at least one value of θ.
- When comparing two unbiased estimators, we should prefer the one with lower variance (i.e. the efficient one).



Estimator - Consistency

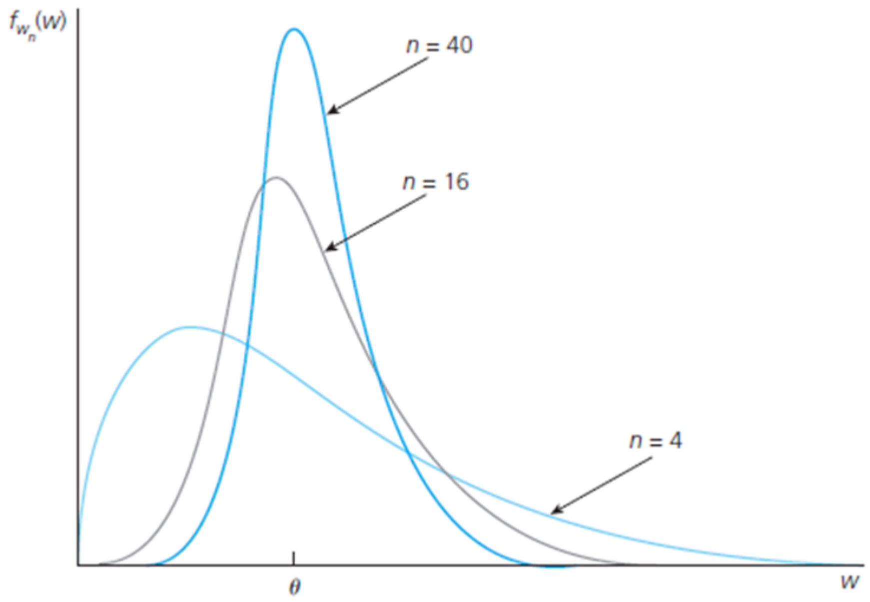
Let be $\hat{\theta}_n$ an estimator for θ based on a sample $x_{1,}\, x_{2,}\, \dots, x_n$ of size n. Then $\hat{\theta}_n$ is a **consistent** estimator for θ if:

$$\forall\epsilon > 0: \qquad Pr\left(\left|\hat{\theta}_n - \theta\right| > \epsilon\right) \to 0 \qquad as\ n \to \infty$$

- The probability that the estimator is close to the true value of the parameter increases to 1 as the sample size gets larger.
- If $\hat{\theta}_n$ is consistent, θ is the probability limit of $\hat{\theta}_n$: $\operatorname*{plim}_{n\to\infty} \hat{\theta} = \theta$

Estimator - Consistency

- Consistency means that, as the sample size increases, the distribution of the estimator becomes more and more concentrated about θ
- Unbiasedness does not necessarily implies consistency (and vice versa). An unbiased estimator is consistent if its variance shrinks to zero as n increases.
- Consistency is typically a minimal requirement of an estimator used in econometrics



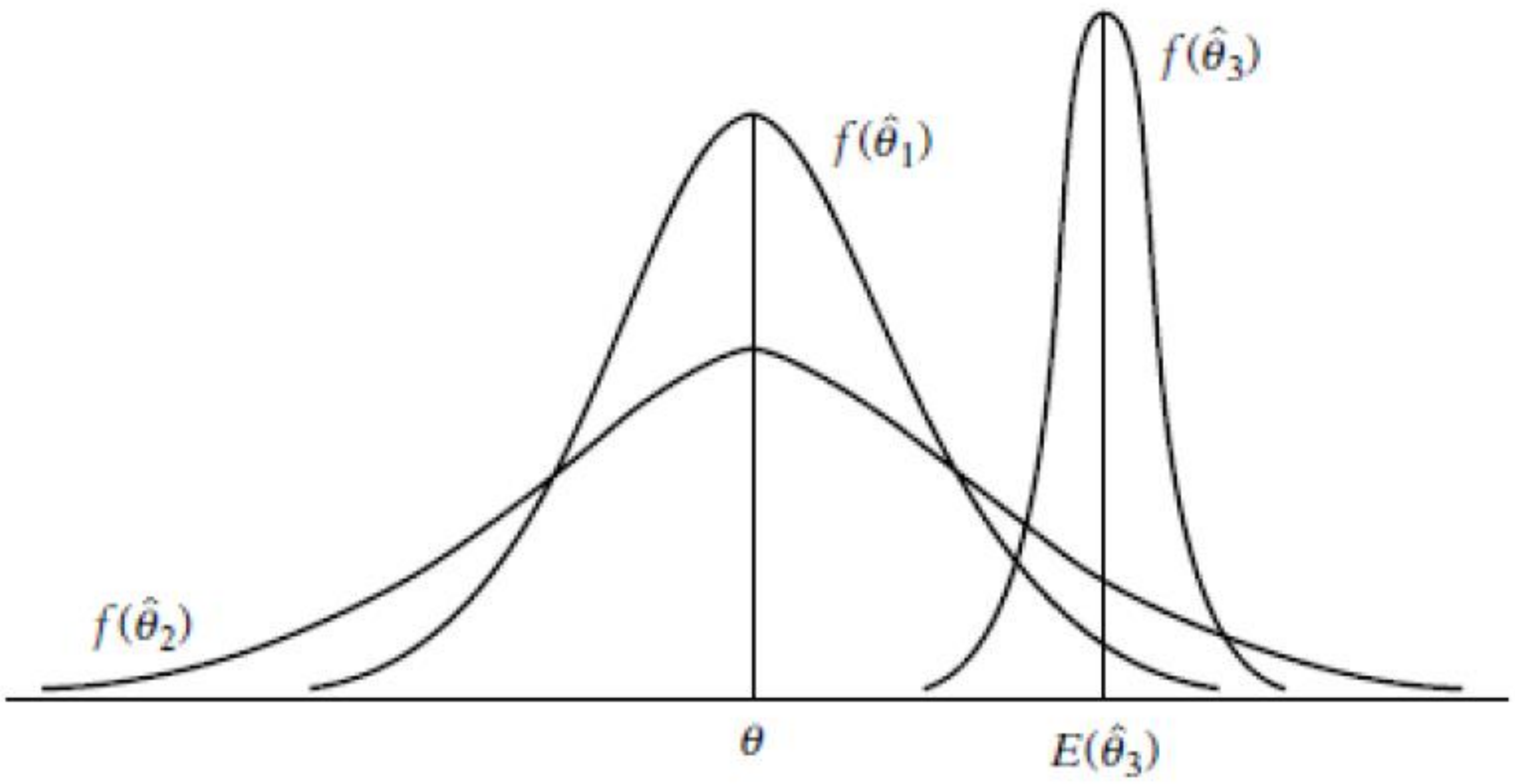
BLUE Estimator

Linear estimator: An estimator $\hat{\theta}$ is said to be linear estimator of θ if it is a linear function of the sample observations.

Best Linear Unbiased Estimator (BLUE): An estimator $\hat{\theta}$ is said to be BLUE if it is:

- Linear,
- Unbiased
- Has the smallest variance in the class of all linear and unbiased estimators of θ.

Estimator



Maths & Stats Pre-Sessional

Estimator: Sample Mean

Lecturer: Claudio Vallar
School of Economics and Finance

Sample Mean

A natural estimator of the population mean μ_Y is the mean of the random sample, that is, the sample mean:

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

- The estimator $\bar{Y}$ is a random variable itself, as it depends on which elements of the population were drawn randomly.
- To determine the properties of the estimator, we need to determine the mean of this random variable $E(\bar{Y})$, and its variance $var(\bar{Y})$.

Sample Mean (optional)

The mean of $\overline{Y}$ is:

$$\mathbb{E}\left(\overline{Y}\right) = \mathbb{E}\left(\frac{1}{n}\sum_{i=1}^{n} Y_i\right) = \frac{1}{n}\sum_{i=1}^{n}\mathbb{E}\left(Y_i\right) = \frac{1}{n}\sum_{i=1}^{n}\mu_Y = \mu_Y$$

The variance of $\overline{Y}$ is:

$$\begin{aligned}\mathrm{var}\left(\overline{Y}\right) = \mathrm{var}\left(\frac{1}{n}\sum_{i=1}^{n} Y_i\right) &= \frac{1}{n^2}\sum_{i=1}^{n}\mathrm{var}\left(Y_i\right) + \frac{1}{n^2}\sum_{i=1}^{n}\sum_{j=1, j\neq i}^{n}\mathrm{cov}(Y_i, Y_j)\\ &= \frac{1}{n^2}\sum_{i=1}^{n}\sigma_Y^2 + 0 = \frac{1}{n^2} n\sigma_Y^2 = \frac{\sigma_Y^2}{n}\end{aligned}$$

Sample Mean

The mean and the variance of the sample mean are:

- $E(\bar{Y}) = \mu_Y$
- $var(\bar{Y}) = \frac{\sigma_Y^2}{n}$

Hence:

- $\bar{Y}$ is an unbiased estimator of μ_Y .
- $var(\bar{Y})$ shrinks as n increases: $var(\bar{Y}_n) \to 0 \quad as\ n \to \infty$
- This implies that $\bar{Y}_n$ is a consistent estimator of μ_Y

Maths & Stats Pre-Sessional

Law of Large Numbers and Central Limit Theorem (optional)

Lecturer: Claudio Vallar
School of Economics and Finance

Law of Large Numbers (LLN)

The consistency of the sample mean is known as the **Law of Large Numbers (LLN)**:

Let $Y_1, Y_2, ..., Y_n$ be independent and identically distributed random variables with mean $E(Y_i) = \mu_Y$ then:

$$p\lim(\overline{Y}_n) = \mu_Y$$

- that is, the sample average $\overline{Y}_n$ **converges in probability** to μ_Y as the sample size n grows indefinitely.

A further result about the sample mean $\overline{Y}_n$ regards its asymptotic distribution...

Central Limit Theorem (CLT)

Central Limit Theorem. Let $X_1, X_2, \ldots, X_n$ be a set of n independent random variables having identical distributions with mean μ, variance σ^2, and $\bar{X}$ is the mean of these random variables. As n becomes large, the **central limit theorem** states that the distribution of

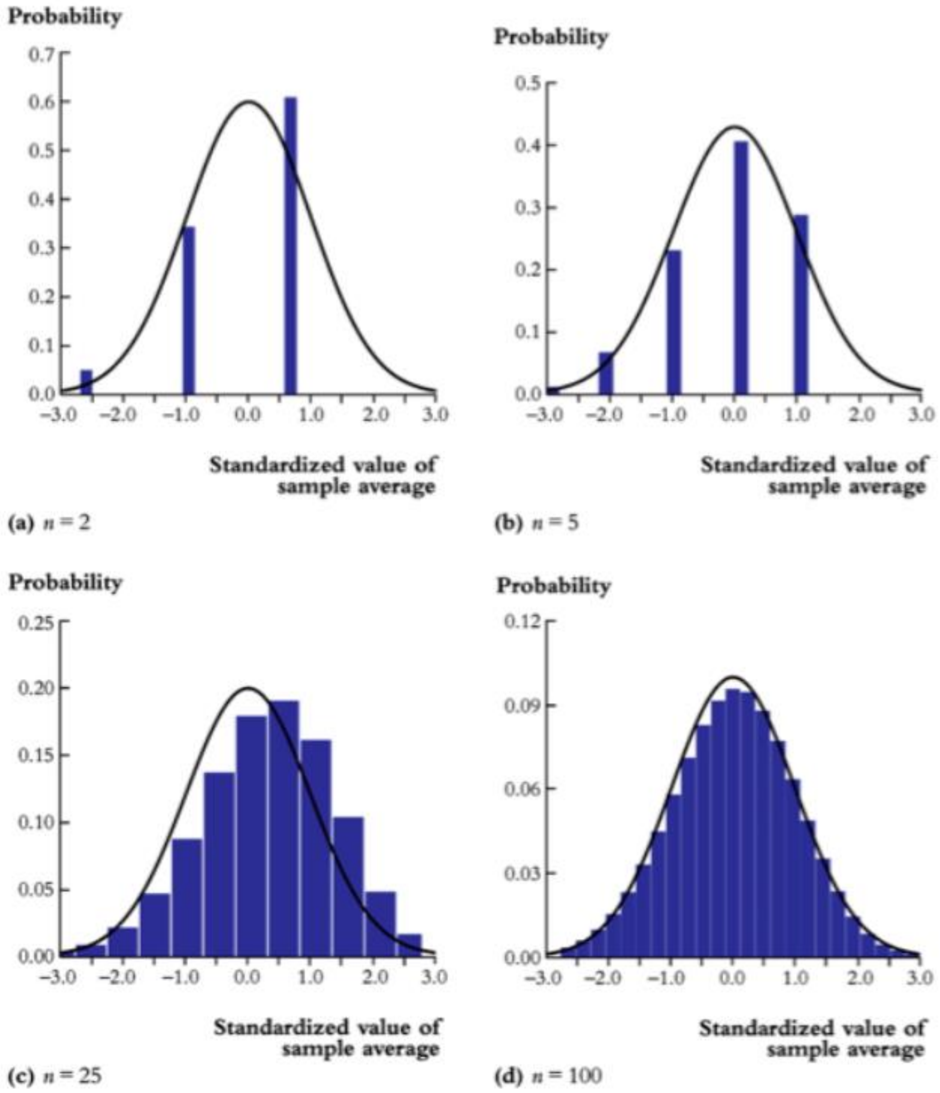
$$Z = \frac{\bar{X} - \mu_{\bar{X}}}{\sigma_{\bar{X}}}$$

Approaches the standard normal distribution.

We can say that :

- Z_n has an asymptotic standard normal distribution
- Z_n converges in distribution to a standard normal distribution

Central Limit Theorem (CLT)



Maths & Stats Pre-Sessional

Hypothesis Testing

Lecturer: Claudio Vallar
School of Economics and Finance

Introduction

In statistical inference, we commonly want to:

- Learn the value of parameter $\Rightarrow$ Use an estimator.
- Test if parameter's value equals specific value (e.g., from theory or intuition) $\Rightarrow$ Use hypothesis testing.

We may be interesting in answering some questions like "Does a job training programme effectively increase average worker productivity?"

A method for answering such questions, using a sample of data, is known as hypothesis testing.

Introduction

Steps to perform hypothesis testing:

- State the null hypothesis and the alternative hypothesis.
- Select the test statistic and determine its distribution.
- Select significance level.
- Perform test statistic using data in your sample.
- Reach a decision about the null hypothesis you stated.

Null Hypothesis / Alternative Hypothesis

- The true population value of the parameter is unknown.

- Hypothesis – A statement about the value of an unknown population parameters.

- The two complementary hypotheses in a hypothesis testing are called the null hypothesis H_0 and the alternative hypothesis H_1 :

  - **Null hypothesis $\boldsymbol{H_0}$** : the hypothesis to be tested. Formation of H_0 is navigated by empirical evidence, intuition or financial/economic theory.
  - **Alternative hypothesis $\boldsymbol{H_1}$** contains all the other possible outcomes.

Significance Level

- We choose the significance level α. This along with the distribution of the test determines the critical value.
- Significance level α usually is equal to 0.01, 0.05 or 0.10.

- The decision rule depends on the way H_1 is formulated.
  - $H_0\colon \mu = \mu_0\,; H_1\colon \mu > \mu_0$

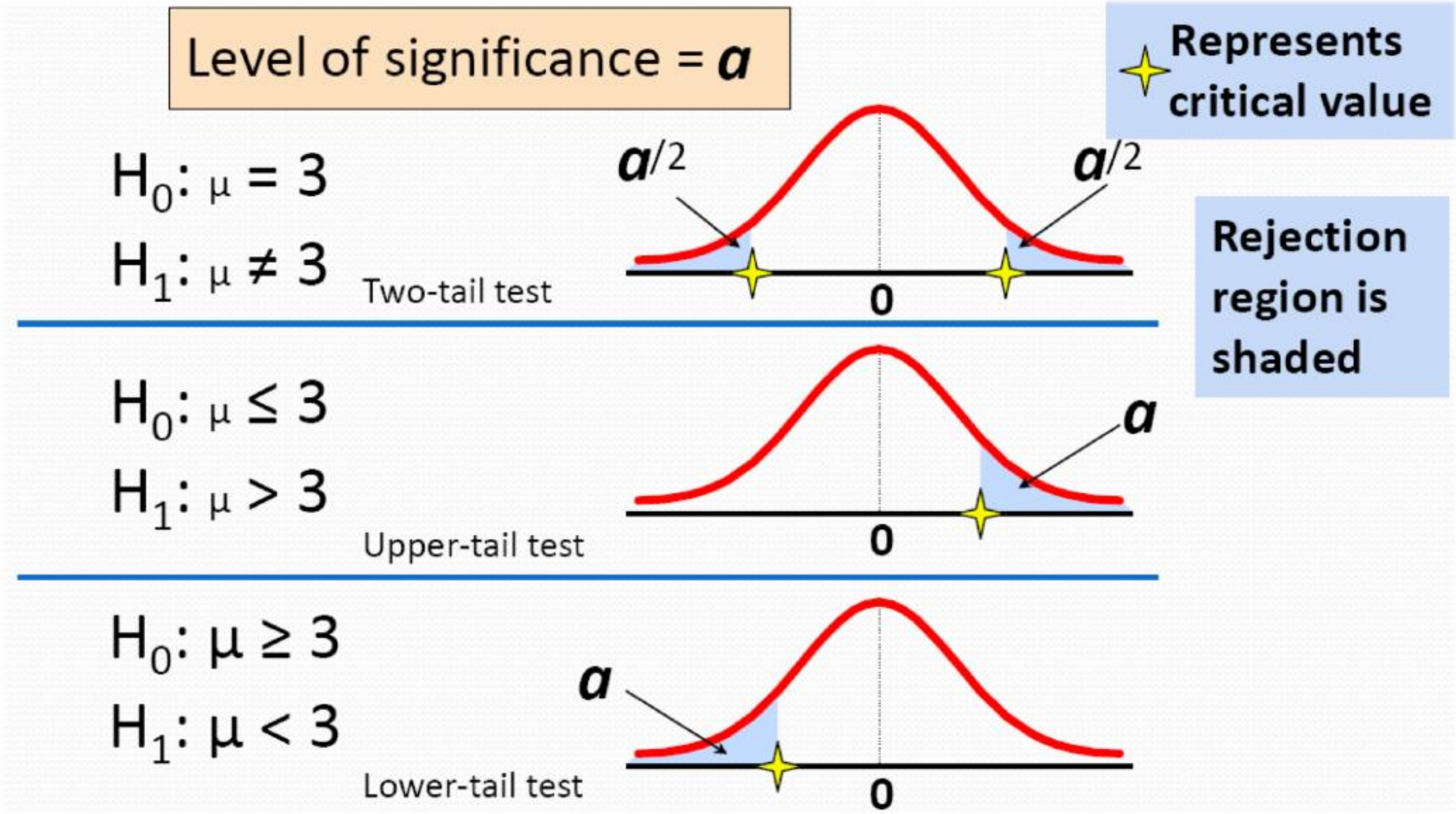
    Reject H_0 if test statistic value > critical value corresponding to α.
  - $H_0\colon \mu = \mu_0\,; H_1\colon \mu < \mu_0$

    Reject H0 if test statistic value < - critical value corresponding to a.
  - $H_0\colon \mu = \mu_0\,; H_1\colon \mu \neq \mu_0$

    Reject H0 if test statistic value < -critical value corresponding to a/2 or test statistic value > critical value corresponding to a/2.

Rejection Area

The level of significance and the rejection region



Hypothesis Testing

Consider the random sample $x_{1,} x_{2,} \ldots, x_n$ drawn from a population $X \sim N(\mu, \sigma^2)$).

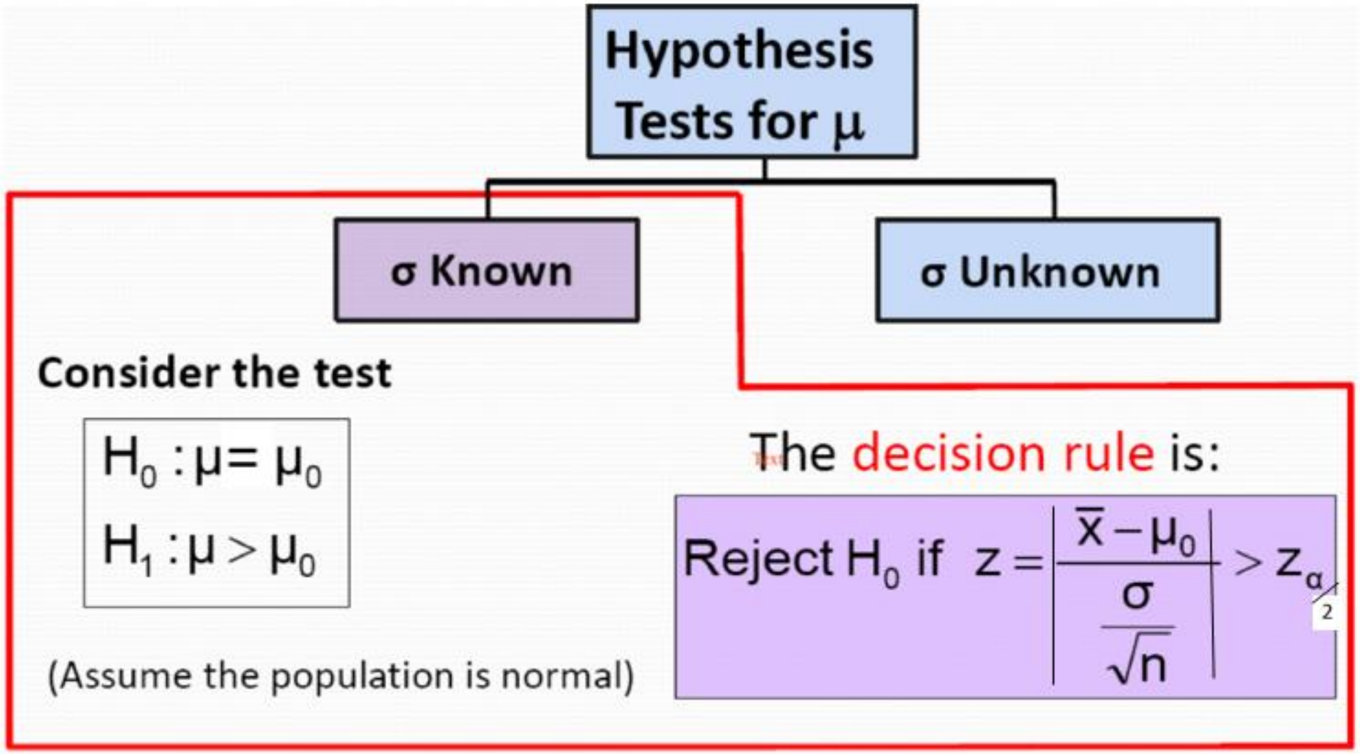
We want to test hypothesis on the mean, e.g., $\mu = \mu_0$

We need to distinguish two cases:

- population variance σ^2 is known;
- population variance σ^2 is unknown.

Hypothesis Testing

Test for the mean of a normal population when population variance σ^2 is **known.**

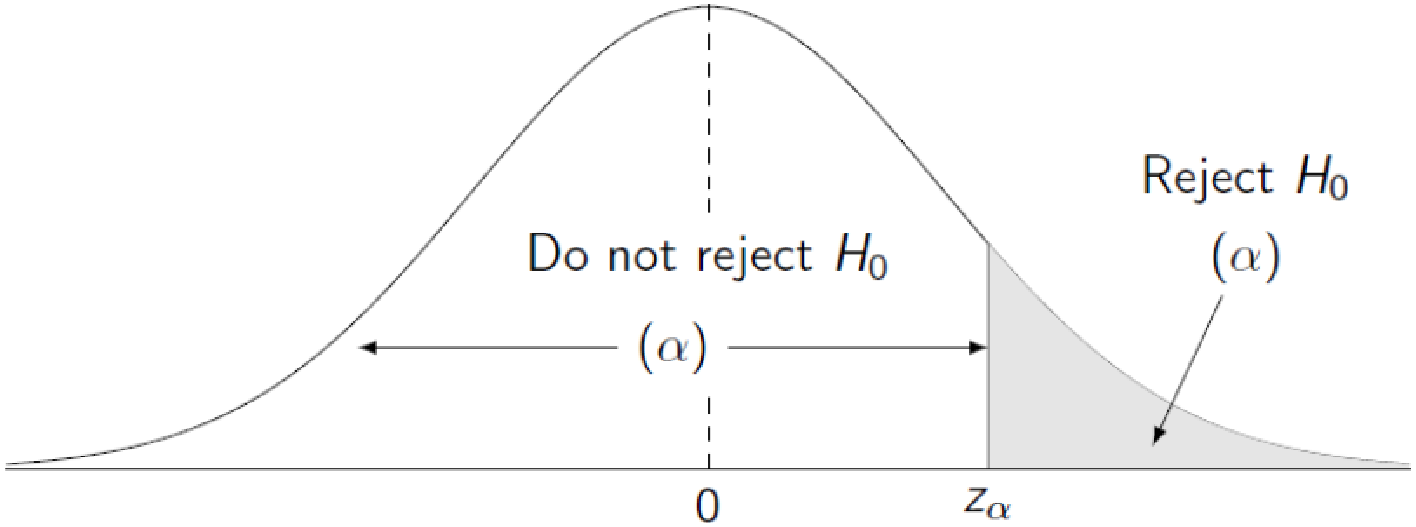


Note that the sample mean $\bar{X} \sim N(\mu, \sigma^2/N)$. Hence, its standardized version has a **standard normal distribution**.

Hypothesis Testing

$$H_0 : \mu = \mu_0$$
$$H_a : \mu > \mu_0$$

$$\text{Test statistic} = \frac{\bar{X} - \mu_0}{\sigma/\sqrt{n}} \overset{H_0}{\sim} N(0,1)$$



Note that: $P(Z > z_\alpha) = \alpha$

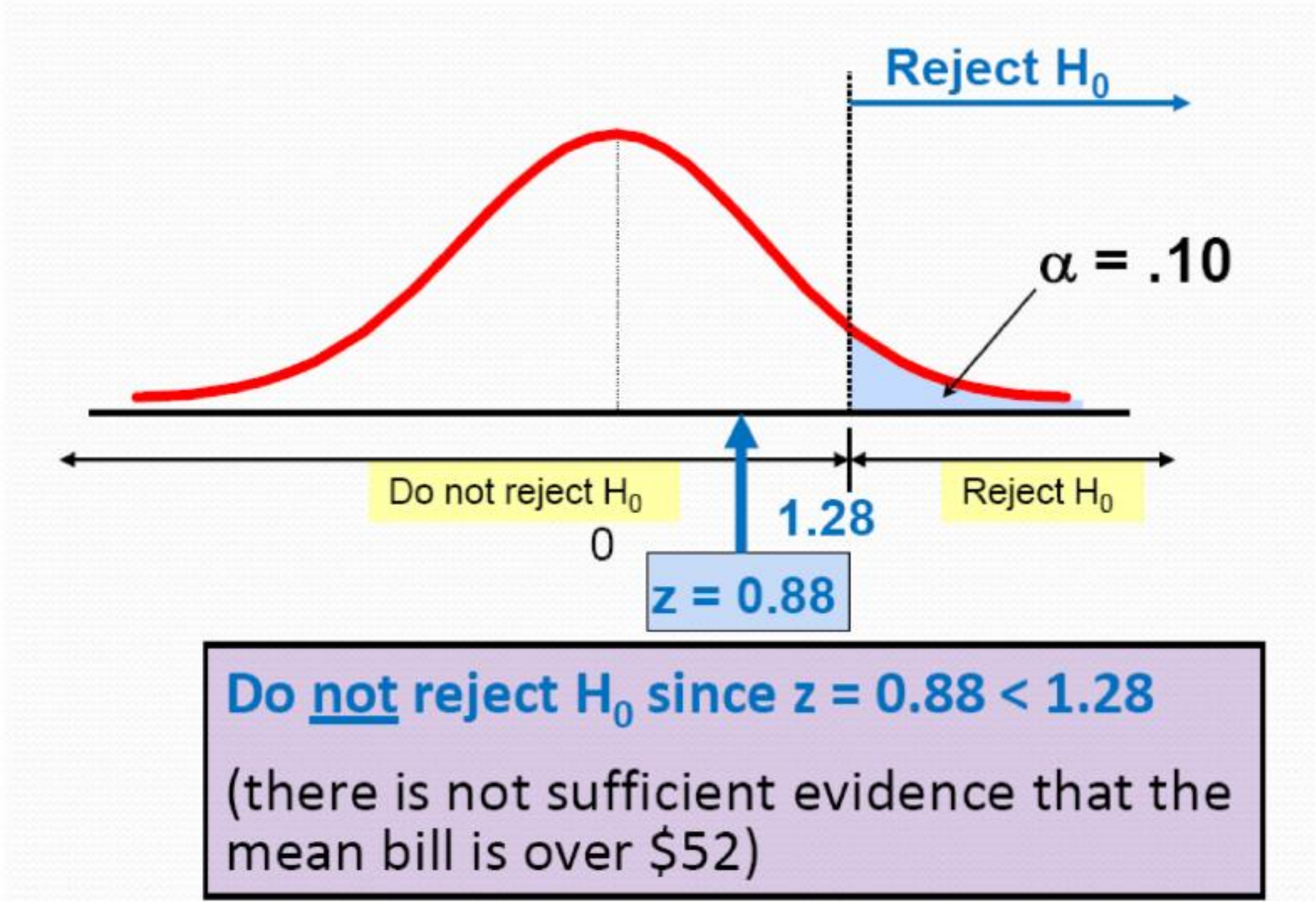
Hypothesis Testing – Example 1

<u>Example</u>: Tom is evaluating his energy bill. He believes that the energy bill is normally distributed with variance 100. He looks at the bills in the past 64 months, which averages at £53.1 per month. He would switch to a new energy provider if it does not cost him more than £52 per month. So he tests the hypothesis that the mean energy bill is at most £52 at the $\alpha = 0.10$ level, in which case he would stay with the current provider.

- The null hypothesis is $H_0\colon\ \mu = 52$
- The alternative hypothesis is $H_1\colon\ \mu > 52$
- The test statistic is

$$z = \frac{\bar{\mu} - \mu_0}{\sigma/\sqrt{n}} = \frac{53.1 - 52}{10/\sqrt{64}} = 0.88$$

Hypothesis Testing – Example 1



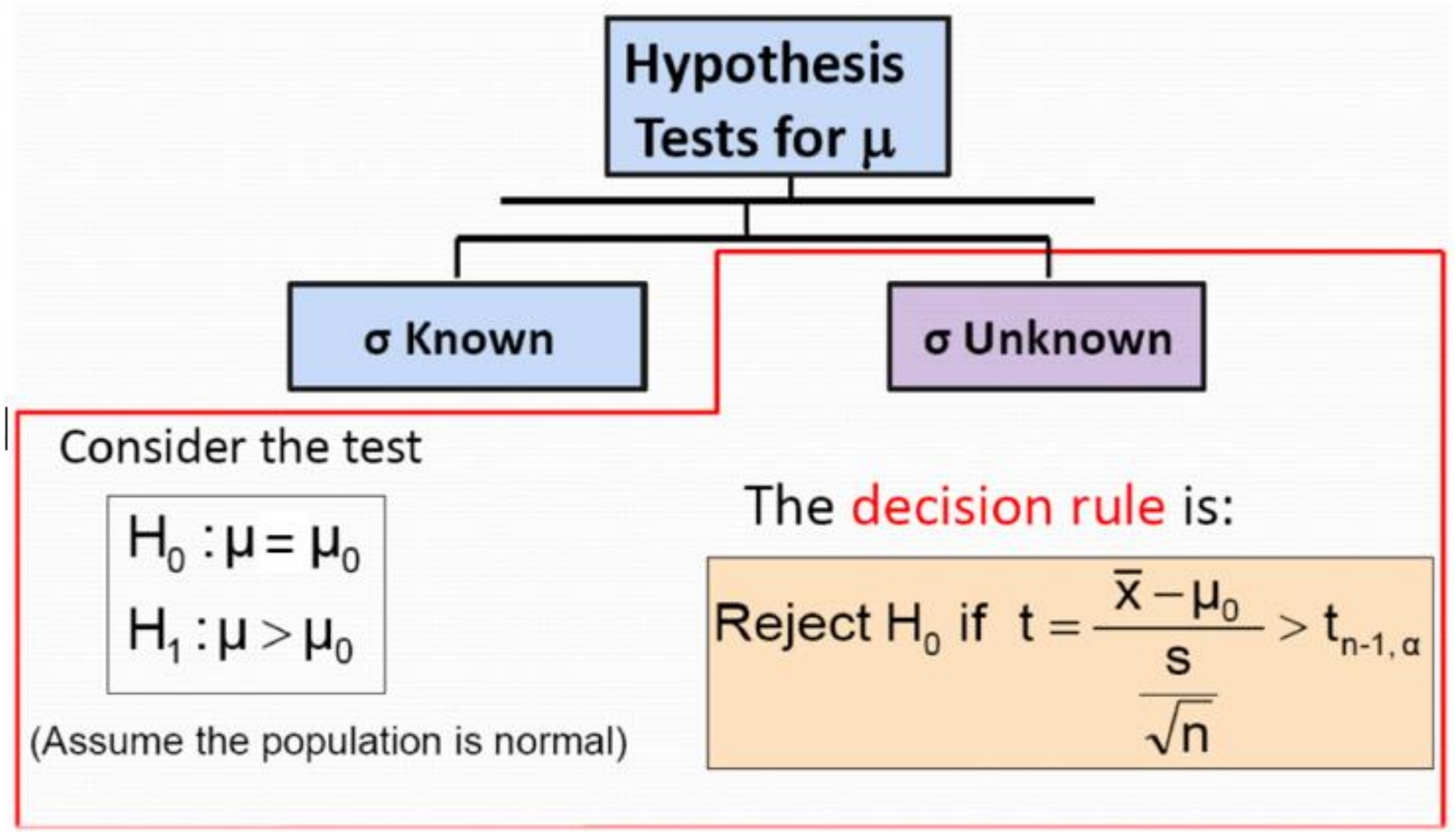
So Tom will change his current energy provider.

Confidence Interval Estimators: Definitions

- A **confidence interval estimator** for a population parameter is a rule for determining (based on sample information) an interval that is likely to include the parameter.
- The corresponding estimate is called a **confidence interval estimate**.

Hypothesis testing

Test for the mean of a normal population when population variance σ^2 is <u>**unknown.**</u>



We perform a **t-test**