BS1112 - Statistics for Economics

Lecture 7 - Hypothesis Testing

Dr. Robert Riegler

What is a Hypothesis?

- A hypothesis is a claim (assumption) about a population parameter (e.g. population mean)
- For example:
  - The mean monthly cell phone bill of this city is $\mu =$ $42
  - The mean wealth in the UK is $\mu =$ £180,000.
  - The mean wealth in the UK is $\mu =$ £250,000.
  - The mean mark in the statistics exam is $\mu = 50$

The Null Hypothesis, H_0

- States the assumption to be tested.
- Example: The average number of TV sets in U.S. homes is three ($H_0 : \mu = 3$)
- Is always about a population parameter, not about a sample statistic.

$$H_0 : \mu = 3 \quad \boxtimes \qquad H_0 : \bar{x} = 3 \quad \square$$

The Null Hypothesis, H_0, continued

- Begin with the assumption that the null hypothesis is true
- Similar to the notion of innocent until proven guilty
- Refers to the status quo
- Always contains '$=$', '$\leq$' or '$\geq$' sign
- May or may not be rejected

The Alternative Hypothesis, H_A

- Is the opposite of the null hypothesis
- E.g.: The average number of TV sets in U.S. homes not equal to 3 ($H_A : \mu \neq 3$)
- Challenges the status quo
- Never contains '$=$', '$\leq$' or '$\geq$' sign
- May or may not be accepted

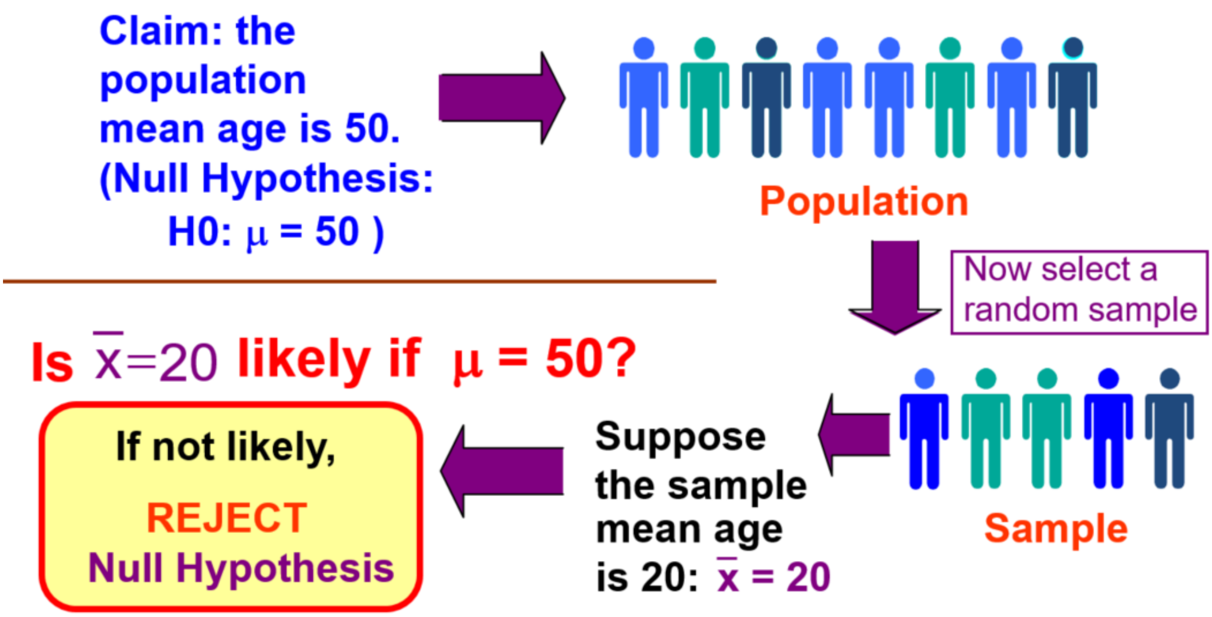
More on Setting up Hypotheses

- Sometimes researchers want to test statements including "greater, smaller, at least, etc."
- Example: The average number of TV sets in U.S. Homes is at least three.
- The hypotheses are:
  - $H_0 : \mu \geq 3$ and $H_A : \mu < 3$
- Note: Here H_A is generally the hypothesis that is believed (or needs to be supported) by the researcher

Student Task I

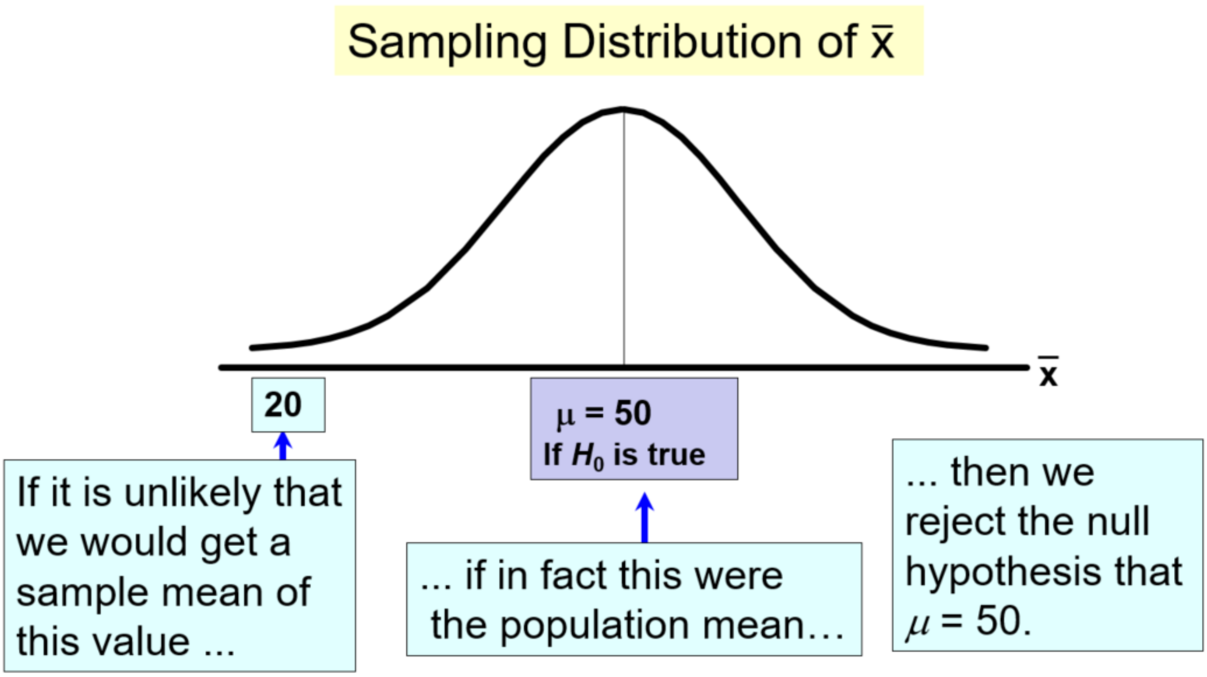
- Set up the H_0 and H_A:
- The average mark in the in-class test was 60.
-
- The average mark is greater than 55.
-

Hypothesis Testing Process



Slide Source : Business Statistics: A Decision Making Approach, 6e. 2005 Prentice-Hall, Inc.

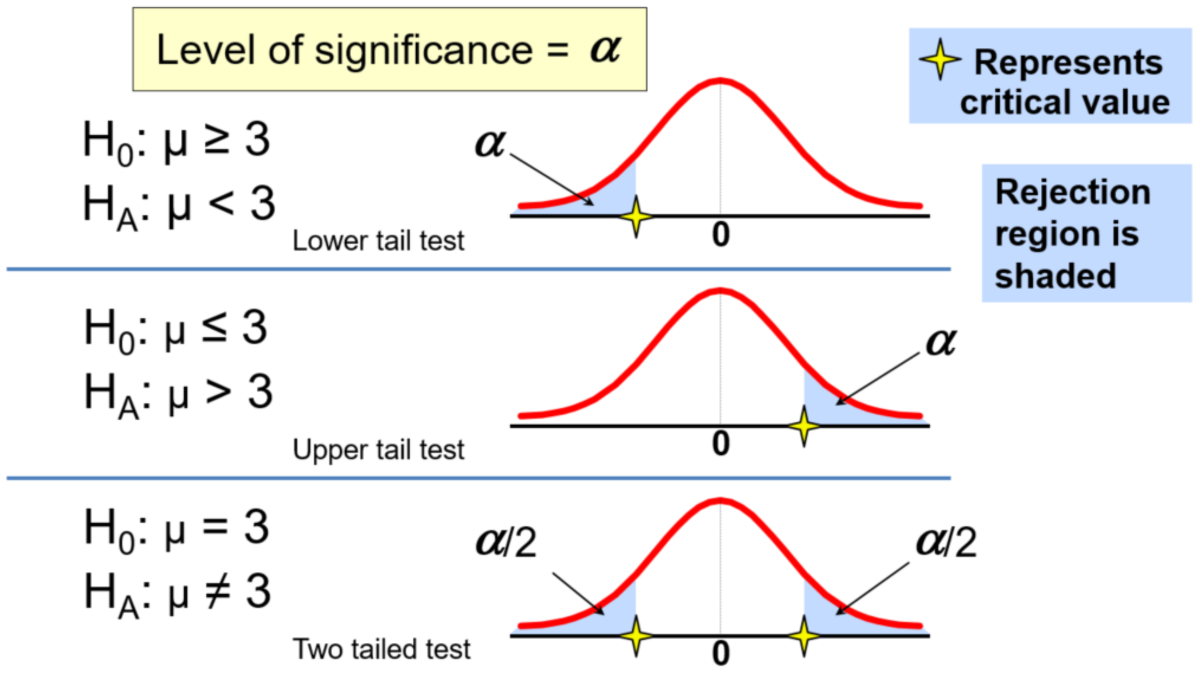
Reason for Rejection H_0



Level of Significance

- **Defines unlikely values of sample statistic if null hypothesis is true**
  - Defines rejection region of the sampling distribution
- Is designated by α, (level of significance)
  - Typical values are .01, .05, or .10
- Is selected by the researcher at the beginning
- Provides the critical value(s) of the test

Level of Significance and the Rejection Region



Errors in Making Decisions

- **Type I error**
  - Reject a true null hypothesis
  - Considered as a serious type of error

The probability of Type I Error is α

- Called level of significance of the test
- Set by researcher in advance

Errors in Making Decisions

- **Type II error**
  - Fail to reject a false null hypothesis

> The probability of Type II Error is β

Hypothesis Fundamentals Testing procedure Hypothesis Testing Comparing Means

Outcomes and Probabilities

Possible Hypothesis Test Outcomes

	State of Nature	
Decision	H_0 True	H_0 False
Do Not Reject **H_0**	**No error** **$(1 - \alpha)$**	**Type II Error** **(β)**
Reject **H_0**	**Type I Error** **(α)**	**No Error** **$(1 - \beta)$**

Key:
Outcome
(Probability)

Type I & II Error Relationship

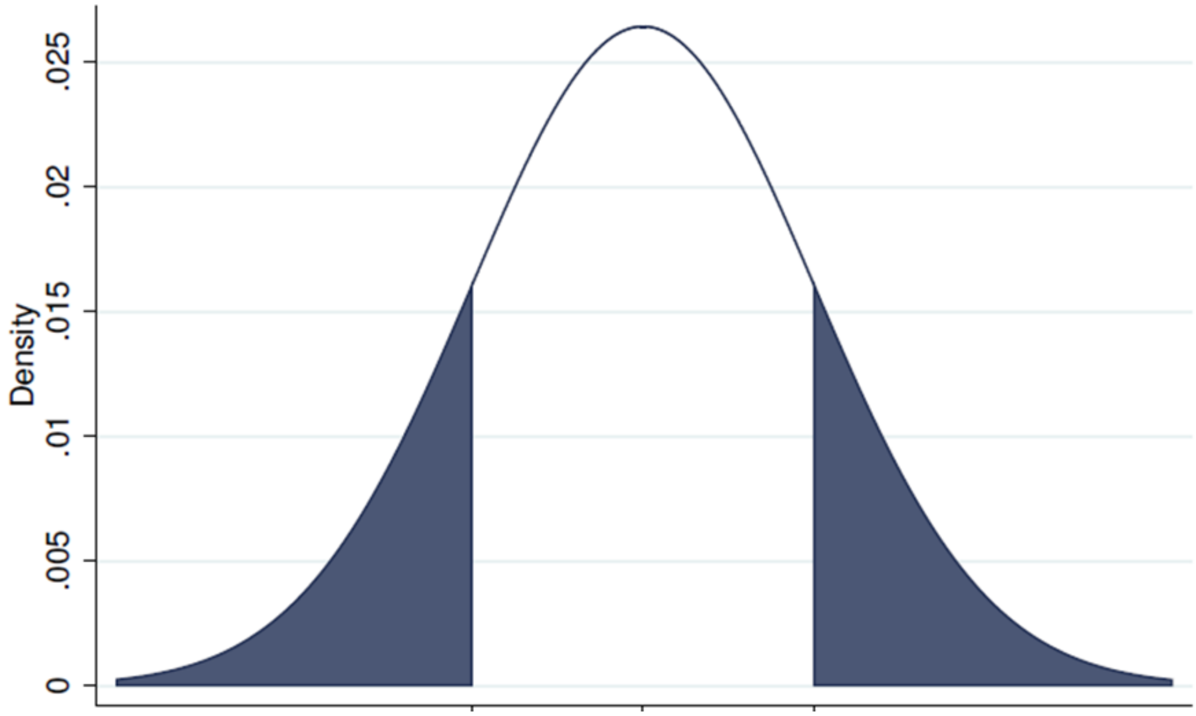
- Type I and Type II errors cannot happen at the same time
  - Type I error can only occur if H_0 is true
  - Type II error can only occur if H_0 is false

If Type I error probability (α) $\uparrow$, then Type II error probability (β) $\downarrow$.

Student Task II

- Referring to the diagram on the next slide:
  - Are the set significance levels (blue areas on both ends) rather high or low?
  - Add a sample mean which could suffer from a type I error.
  - If you increase the significance level, is a type I error more or less likely to appear?

Student Task 1 cont.

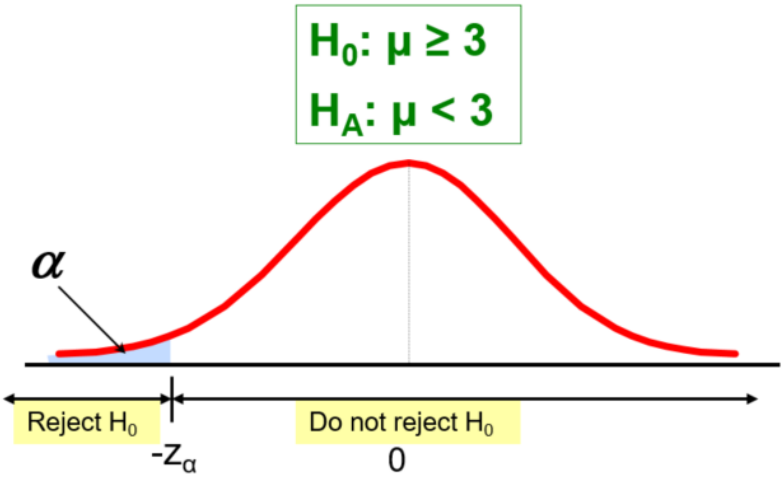


Critical Value Approach to Testing

- Convert sample statistic (e.g.: $\bar{x}$) to test statistic (z or t statistic)
- Determine the critical value(s) for a specified level of significance α from a table or computer
- If the test statistic falls in the rejection region, reject H_0; otherwise do not reject H_0

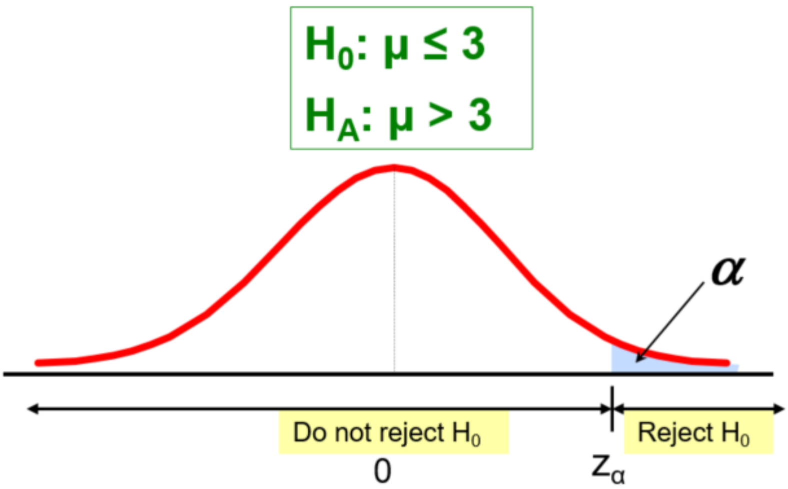
Lower Tail Test

The cut-off value $-z_\alpha$ is called critical value



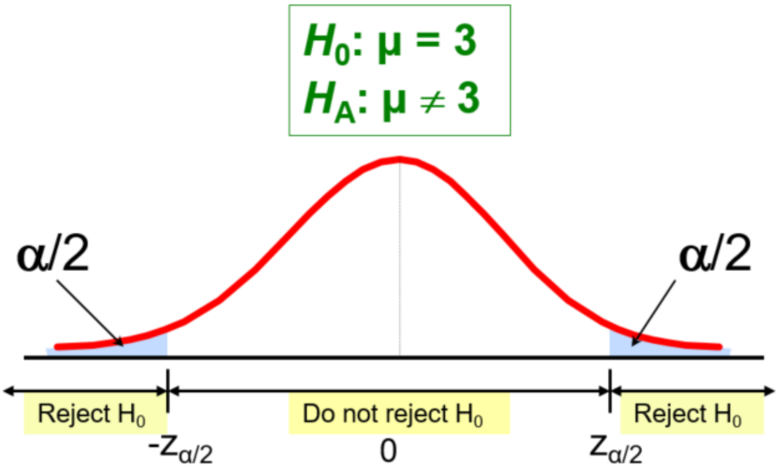
Upper Tail Test

The cut-off value z_α is called critical value



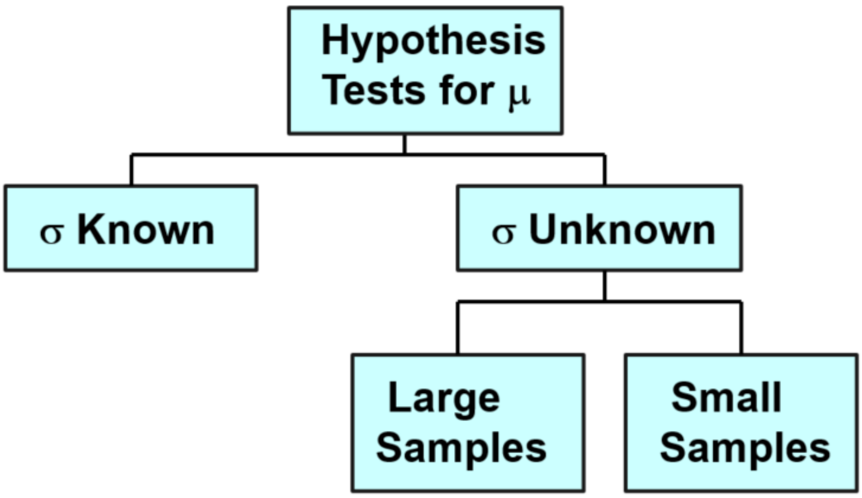
Two Tailed Test

There are two cutoff values (critical values): $\pm z_{\frac{\alpha}{2}}$

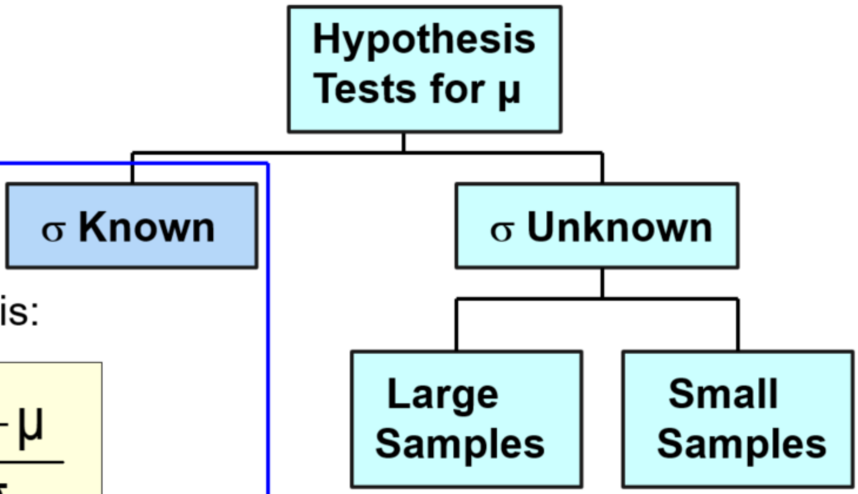


Critical Value Approach to Testing

- Convert sample statistic ($\bar{x}$) to a test statistic (z or t statistic)



Calculating the Test Statistic

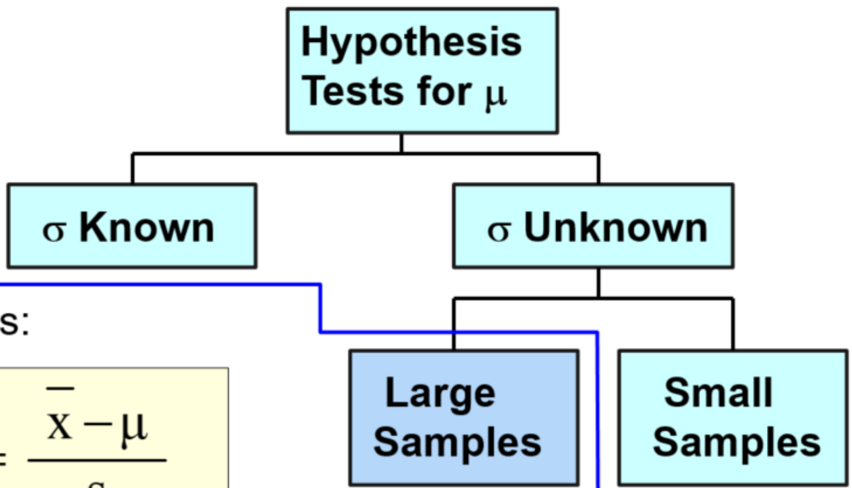


The test statistic is:

$$z = \frac{\bar{x} - \mu}{\frac{\sigma}{\sqrt{n}}}$$

Hypothesis Fundamentals Testing procedure Hypothesis Testing Comparing Means

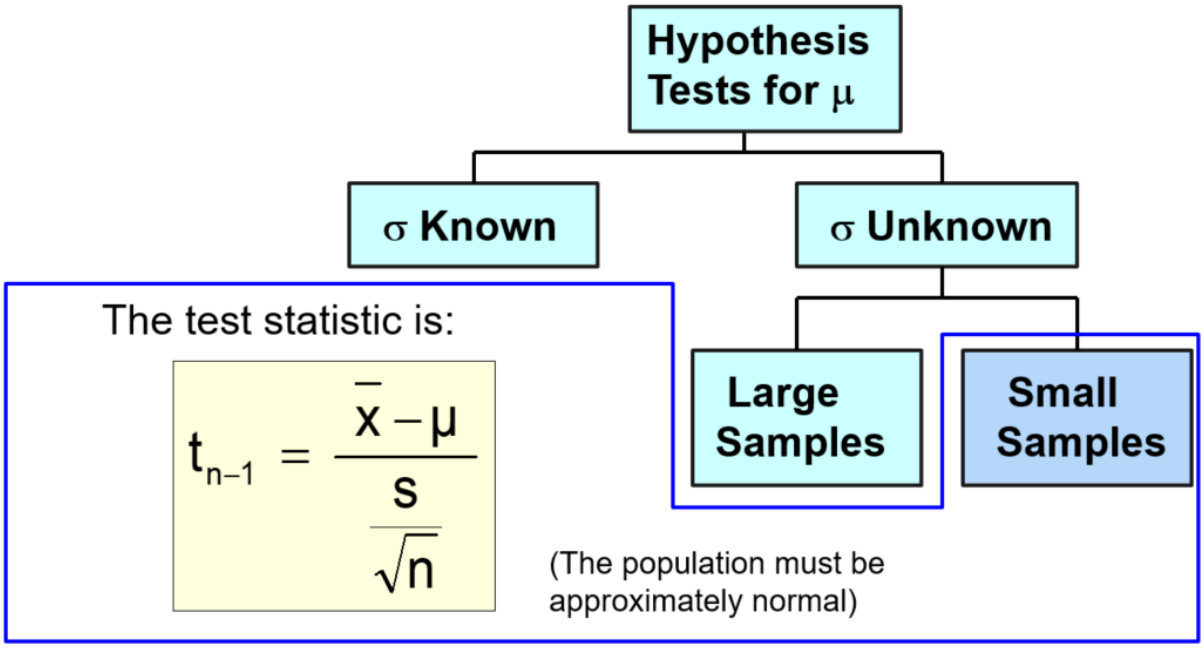
Calculating the Test Statistic



The test statistic is:

$$z = \frac{\bar{x} - \mu}{\frac{s}{\sqrt{n}}}$$

Calculating the Test Statistic



Review: Steps in Hypothesis Testing

1. Specify the population value of interest
2. Formulate the appropriate null and alternative hypotheses
3. Specify the desired level of significance
4. Determine the rejection region
5. Obtain sample evidence and compute the test statistic
6. Reach a decision and interpret the result

Hypothesis Testing Examples

σ is known, n is large

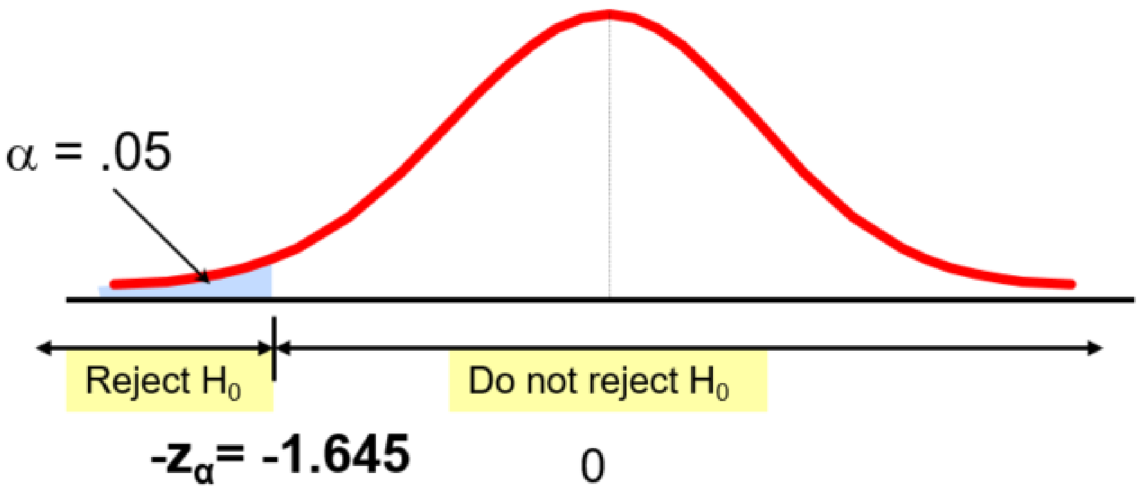
Hypothesis Testing Example

Test the claim that the true mean # of TV sets in US homes is at least 3. Assume $\sigma = 0.8$.

1. Specify the population value of interest
   - The mean number of TVs in US homes
2. Formulate the appropriate null and alternative hypotheses
   - $H_0 : \mu \geq 3$ $\quad H_A : \mu < 3$ (This is a lower tail test)
3. Specify the desired level of significance
   - Suppose that $\alpha = 0.05$ is chosen for this test

Hypothesis Testing Example cont.

4. Determine the rejection region



This is a one-tailed test with $\alpha = 0.05$. Since σ is known, the cut-off value is a z value:

> Reject H_0 if $z < z_\alpha = -1.645$; otherwise do not reject H_0

Hypothesis Testing Example cont.

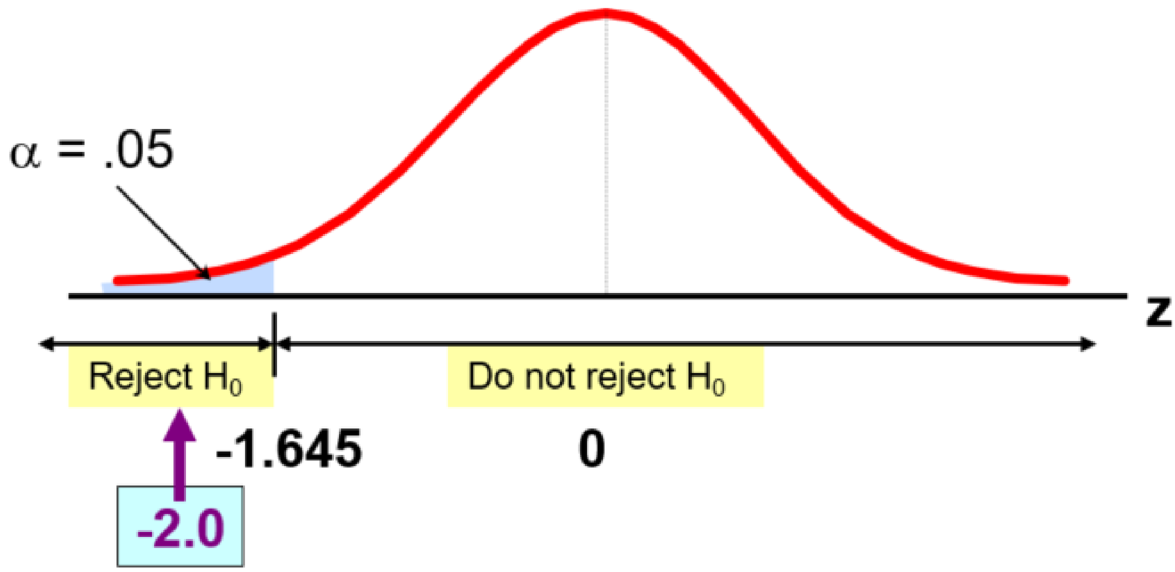
5. Obtain sample evidence and compute the test statistic

- Suppose a sample is taken with the following results:
- $n = 100$, $\bar{x} = 2.84$ ($\sigma = 0.8$ is assumed known)
- Then the test statistic is:

$$z = \frac{\bar{x} - \mu}{\frac{\sigma}{\sqrt{n}}} = \frac{2.84 - 3}{\frac{0.8}{\sqrt{100}}} = \frac{-0.16}{0.08} = -2.0$$

Hypothesis Testing Example cont.

6. Determine the rejection region



Since $z = -2.0 < -1.645$, we **reject the null hypothesis** that the mean number of TVs in US homes is at least 3

Example: Upper Tail z Test for Mean (σ known)

A phone industry manager thinks that customer monthly cell phone bill have increased, and now average over £52 per month. The company wishes to test this claim. Assume $\sigma = 10$ is known.

From the hypothesis test:

$H_0 : \mu \leq 52$ the average is not over £52 per month

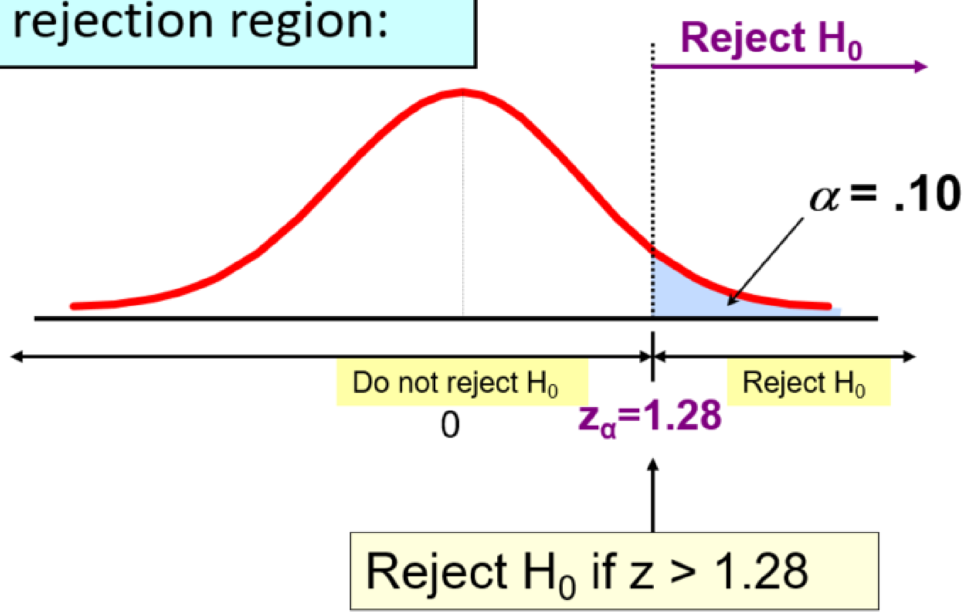
$H_A : \mu > 52$ the average is greater than £52 per month

(i.e., sufficient evidence exists to support the manager's claim)

Example: Find Rejection Region

- Suppose that $\alpha = 0.10$ is chosen for this test

Find the rejection region:



Example: Test Statistic

Obtain sample evidence and compute the test statistic

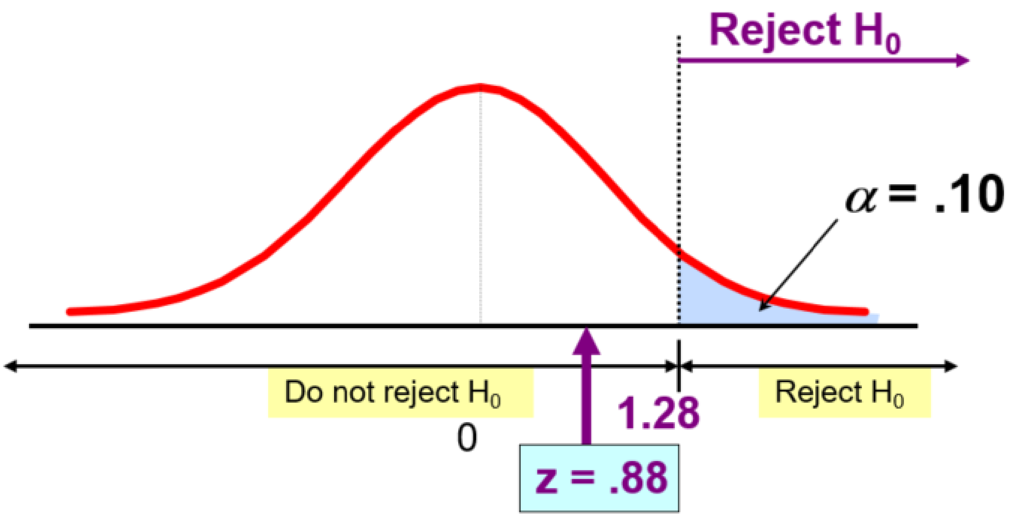
Suppose a sample is taken with the following results: $n = 64$, $\bar{x} = 53.1$ ($\sigma = 10$ was assumed known)

Then the test statistic is:

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

Example: Decision

Reach a decision and interpret the result:



Do not reject H_0 since $z = 0.88 \leq 1.28$. I.e.: there is not sufficient evidence that the mean bill is over $52.

Student Task III: Two tailed test

- It is claimed that an average child spends 15 hours per week watching television. A survey of $n = 100$ children finds an average of $\bar{x} = 14.5$ hours per week, with standard deviation $s = 8$ hours. Is the claim justified?
- The claim would be wrong if children spend either more or less than 15 hours watching TV. The rejection region is split across the two tails of the distribution. This is a two tailed test.
- Conduct a hypothesis test!

Hypothesis Testing Examples

σ is known, n is small ($n \leq 30$)

Example: Two-Tail Test (σ unknown)

The average cost of a hotel room in New York is said to be \$168 per night. A random sample of 25 hotels resulted in $\bar{x} = \$172.50$ and $s = \$15.40$. Test at the $\alpha = 0.05$ level. Assume the population distribution is normal.

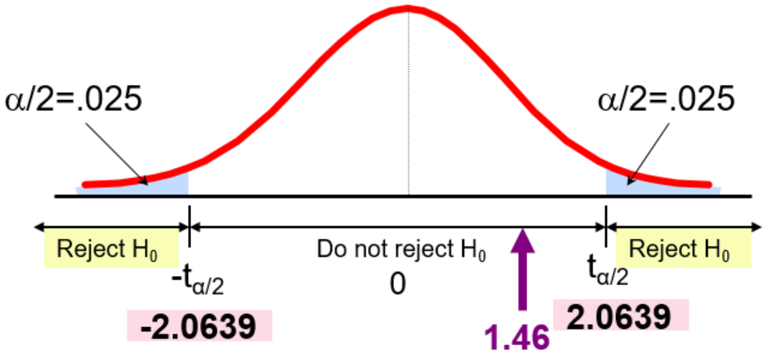
- $H_0 : \mu = 168$
- $H_A : \mu \neq 168$

Example Solution: Two-Tail Test

H_0: $\mu = 168$
H_A: $\mu \neq 168$

- $\alpha = 0.05$
- n = 25
- σ is unknown, so use a **t statistic**
- Critical Value:

$t_{24} = \pm\ 2.0639$



$$t_{n-1} = \frac{\bar{x} - \mu}{\frac{s}{\sqrt{n}}} = \frac{172.50 - 168}{\frac{15.40}{\sqrt{25}}} = 1.46$$

Do not reject H_0: not sufficient evidence that true mean cost is different than \$168

Student Task IV: Small sample testing

- A sample of 12 cars of a particular type has on average 35 mpg, with standard deviation 15.
- Test the manufacturer's claim of 40 mpg or more as the true average.

Comparison-of-means testing

- Is the average growth rate of the UK over the last 30 years the same as the growth rate in Ireland?
- Is the average wage of men the same as the average wage of women?
- We require *independent sample* comparison-of-means tests
- Does the Covid vaccine reduce the frequency of Covid symptoms?
- Does a university degree make students independent learners?
- We require *paired / matched* comparison-of-means tests

Comparison-of-means testing (indep. sample, *n* large)

- Assume σ_1^2 and σ_2^2 are unknown, but sample is large. Sample variances may be different
- To test whether two samples are drawn from populations with the same mean, we do a 2-tail test.
- Hypotheses:
  $H_0 : \mu_1 = \mu_2$ or $H_0 : \mu_1 - \mu_2 = 0$
  $H_1 : \mu_1 \neq \mu_2$ or $H_1 : \mu_1 - \mu_2 \neq 0$
- The test statistic is:

$$z = \frac{(\bar{x}_1 - \bar{x}_2) - (\mu_1 - \mu_2)}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}}$$

Example

- Assume weekly wages of men and women that work for a delivery company have been collected. Test if the averages wages of men and women are different.
- Sample:
  Male: $\bar{x}_1 = 420$, $s_1 = 25$, $n_1 = 30$
  Female: $\bar{x}_2 = 408$, $s_2 = 20$, $n_2 = 30$
- Test calculations:

$$z = \frac{(420 - 408) - 0}{\sqrt{\frac{25^2}{30} + \frac{20^2}{30}}} = 2.05$$

- Set significance level at 1%, therefore the critical value from the Z table is 2.575 (2-tailed test).
- We cannot reject the null as the z-score is smaller than z-critical. There is not a statistically significant difference between the 2 means.

Comparison-of-means testing (indep. sample, *n* small)

- σ_1^2 and σ_2^2 are unknown, and sample is small.
- Two conditions must hold:
  1. Population variances are normal
  2. t-distribution requires that $\sigma_1^2 = \sigma_2^2$
- The test statistic is:

$$t = \frac{(\bar{x}_1 - \bar{x}_2) - (\mu_1 - \mu_2)}{\sqrt{\frac{S^2}{n_1} + \frac{S^2}{n_2}}}$$

- where S^2 is the pooled variance

$$S^2 = \frac{(n_1 - 1)s_1^2 + (n_2 - 1)s_2^2}{n_1 + n_2 - 2}$$

Comparison-of-means testing (paired)

- Similar to one-sample hypothesis test!
- Calculate difference between the pre- and post intervention values for each sample unit
- E.g. difference in output of a worker before and after training session
- Calculate sample mean and standard deviation of differences
- If n is large, use z-statistic, if n is small, use t-statistic