Hypothesis Testing for One Sample

Learning Outcomes

Learning Outcomes

At the end of this lesson, the learner should be able to:

1. Explain and formulate the procedure for testing a hypothesis
2. Explain Type I and Type II errors with reference to the possible outcomes of a hypothesis test.
3. Evaluate the hypothesis of a population mean by using the z-test or t-test.
4. Perform inferential statistics on real-life business data using statistical hypothesis tests.

Introduction to Hypothesis testing

- You have been working several years, and have saved up to buy your first car. You came across an advertisement:

BMW 3 Series Automatic Transmission Diesel Variant on the Expressway is 22.69 km/litre

- Based on feedback from your friends who drive, this figure seems high. So, what do you make of this claim? How would you verify its validity?

- In this lesson, you will apply concepts of Hypothesis Testing to real life problems to decide whether to reject or not to reject a claim, such as the one above.

Hypothesis Testing

- Hypothesis testing begins with a hypothesis statement about a population parameter.

> **What is a Hypothesis?**
> A claim about a population parameter. It may or may not be true.

- Examples:
  - BMW 3 Series Automatic Transmission Diesel Variant on the Expressway is 22.69 km/litre
  - Mean cost to remodel a kitchen is $20,000

Hypothesis Testing

- The objective of hypothesis testing is to verify the validity of a claim about a population parameter.

What is Hypothesis Testing?
A procedure based on sample evidence and probability theory to determine whether the hypothesis is a reasonable statement.

STEP 1
State Null and Alternate Hypothesis

STEP 2
Identity type of test & level of significance

STEP 3
Compute test statistics

STEP 4
Formulate the decision rule

STEP 5
Summarise results

Step 1: State Null and Alternate hypothesis

- To test a population parameter we state a pair of hypothesis

Null hypothesis, H_o

- Statement about the value of a population parameter developed for the purpose of testing numerical evidence
- Contains a statement of equality such as $\leq$, $=$, $\geq$

Alternate hypothesis, H_a

- Statement that is accepted if the sample data provide sufficient evidence that the null hypothesis is false
- Complement of H_o
- Contains statement of strict inequality such as $<$, $\neq$, $>$

Step 1: State Null and Alternate hypothesis

- To state a Null and Alternate Hypothesis, translate the claim into a mathematical statement.
- Example 1:

Claim	: BMW 3 Series Automatic Transmission Diesel Variant on the Expressway is 22.69 km/litre
Null hypothesis, Ho	: Ho: $\mu = 22.69$ (Claim)
Alternate hypothesis, Ha	: Ha: $\mu \neq 22.69$

Step 1: State Null and Alternate hypothesis

- Besides the previous example, there are 2 other possible pairs of Null and Alternate Hypothesis, where k is the claim value:

Ho: $\mu \leq k$	Ho: $\mu \geq k$
Ha: $\mu > k$	Ha: $\mu < k$

- Example 2:

Claim	: Pizza Chain claims that its mean delivery time is less than 30 minutes.
Null hypothesis, Ho	: Ho: $\mu \geq 30$
Alternate hypothesis, Ha	: Ha: $\mu < 30$ (Claim)

Step 1: State Null and Alternate hypothesis

- Example 3:

Claim	:	CEO of Supercomputer Ltd claims that, with his company's computer system installed, the current average 22 customers served per hour at an ATM, will increase.
Null hypothesis, Ho	:	Ho: $\mu \leq 22$
Alternate hypothesis, Ha	:	Ha: $\mu > 22$ (Claim)

Step 2: Identify type of test & level of significance

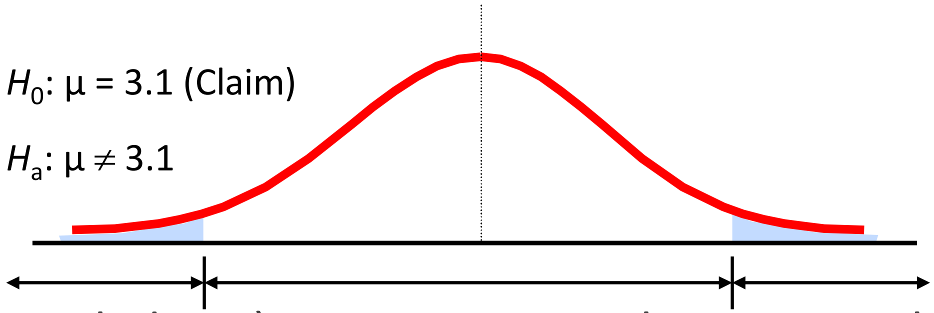
- Knowing the type of hypothesis tests helps us to decide the criteria for rejecting the null hypothesis.
- There are three types of hypothesis tests—two-tailed, a left-tailed, right-tailed test.

Types of Tests: Two tailed Test

- Example 4a. (Two tailed test): A nutritionist claims that the mean consumption of fish in Singapore per person is 3.1 kg per year.

H_0: $\mu = 3.1$ (Claim)

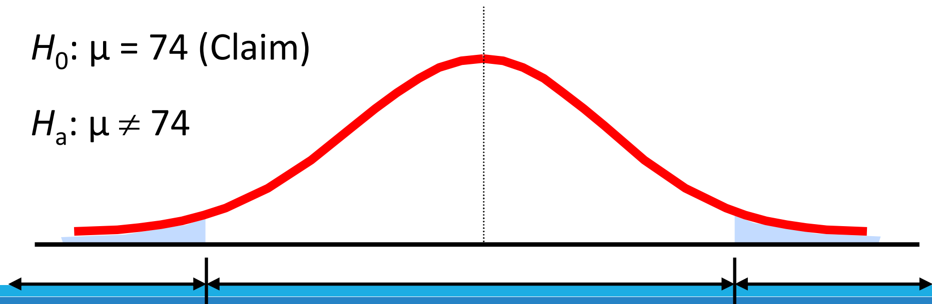
H_a: $\mu \neq 3.1$



- Example 4b. (Two tailed test): A consumer analyst reports that the mean life of a certain type of automobile battery is 74 months.

H_0: $\mu = 74$ (Claim)

H_a: $\mu \neq 74$

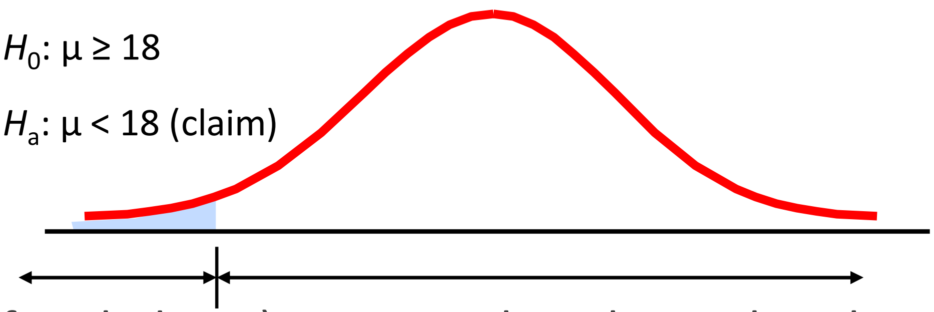


Types of Tests: Left tailed test

- Example 5a (Left tailed test): An engineer claims mean defects of compact disks can be reduced by using robots. The mean defects of compact disk is 18 per 1000

H_0: $\mu \geq 18$

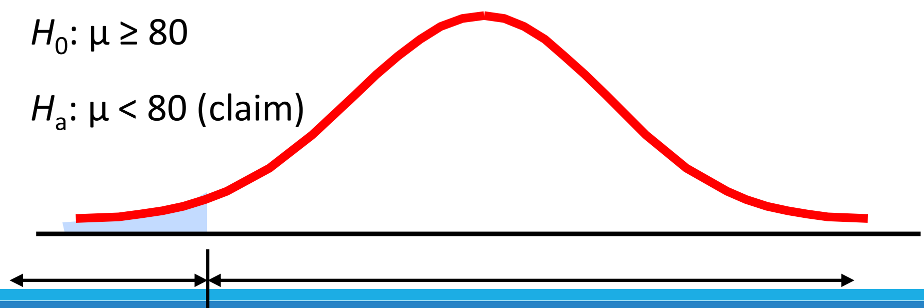
H_a: $\mu < 18$ (claim)



- Example 5b (Left tailed test): A researcher claims that the average cost of men's jeans is less than $80

H_0: $\mu \geq 80$

H_a: $\mu < 80$ (claim)

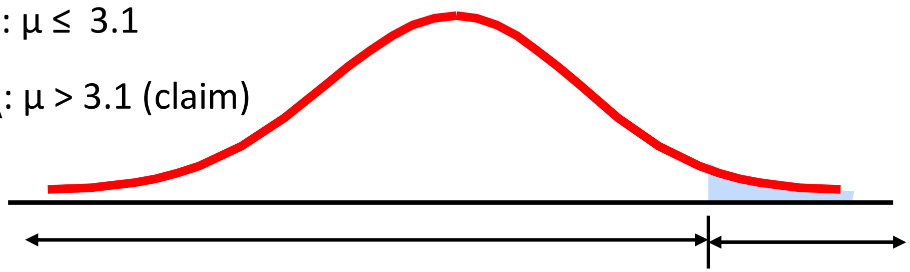


Types of Tests: Right tailed test

- Example 6a (Right tailed test): A researcher says if expectant mothers take vitamin pills, birth weight of babies will increase. Average birth weight of the babies population is 3.1 kg

$H_0: \mu \leq 3.1$

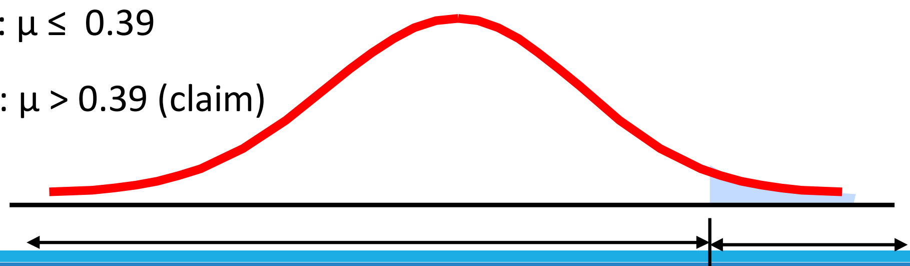
$H_A: \mu > 3.1$ (claim)



- Example 6b (Right tailed test): A radio station publicises that more than 39% of the population is listening to its station.

$H_0: \mu \leq 0.39$

$H_A: \mu > 0.39$ (claim)



Level of significance

- Since our decision is based on a sample rather than the entire population, there is always the possibility that we will make the wrong decision.
- The Probability of rejecting the null hypothesis when it is true is known as Level of Significance. It is denoted by α

Level of Significance, α

- Probability of rejecting the null hypothesis when it is true
- There is no one Level of Significance, is it depends on the risk level you can accept
- Traditionally,
  - $\alpha = 0.05$ for consumer research projects
  - $\alpha = 0.01$ quality assurance

Types of Errors

- There are 2 Types of errors in Hypothesis Testing: Type I and Type II Errors.

Type I Error, α : Null hypothesis is rejected when it is true.

Type II Error, β : Null hypothesis is not rejected when it is false.

Decision	H_o is TRUE	H_o is FALSE
Reject H_o	Type 1 Error	Correct Decision
DO NOT reject H_o	Correct Decision	Type II Error

Reject Ho when it is actually true

Did not reject Ho when it is actually false

Types of Errors

- Example 7: The allowable limit by Agri-Food Authority for salmonella contamination for chicken is 20%. A inspector reports that the chicken produced by a company exceeds this the limit. You perform a hypothesis test to determine whether the inspector's claim is true. When will a type I or II error occur? Which is more serious?

Claim	: Chicken produced exceeds allowable contamination limit of 20%
Null hypothesis, Ho	: Ho: $\mu \leq 0.2$
Alternate hypothesis, Ha	: Ha: $\mu > 0.2$ (Claim)

Type I Error: H_o is true, that is, contaminated chicken is less than or equal to 0.2 but we <u>reject</u> null hypothesis.

Type II Error: Ho is false that is, contaminated chicken is greater than 0.2 but we do <u>not reject</u> the null hypothesis. Type II error is more serious because we allow chicken that exceeded contamination limit to be sold to consumers, which could result in sickness and death.

Types of Errors

- Example 8: A company specialising in parachute assembly states that its main parachute failure rate is less than or equal to 1%. You perform a hypothesis test to determine if its claim is false. When will a type I or type II error occur? Which is more serious?

Claim	:	Failure rate of parachute not more than 1%
Null hypothesis, Ho	:	Ho: $\mu \leq 0.01$ (claim)
Alternate hypothesis, Ha	:	Ha: $\mu > 0.01$

Type I Error: H_o is true, that is, failure rate ≤ 0.01 but we <u>reject</u> null hypothesis.

Type II Error: H_o is false, that is, failure rate > 0.01 but we <u>do not reject</u> the null hypothesis
Type II error is more serious because defective parachutes to be used, which could result in death.

Step 3: Compute the test statistic

- A Test Statistic, z is a value, determined from sample information, used to determine whether to reject the null hypothesis.
- In this lesson, we will use z or t as the test statistic.

	Population Variance, σ^2	Sample size, n	Test Statistic
Testing a single sample value	known	–	$z = \frac{x-\mu}{\sigma}$
Testing a mean	known	any n	$z = \frac{\bar{x}-\mu}{\sigma/\sqrt{n}}$
Testing a mean	unknown	$n \geq 30$	$z = \frac{\bar{x}-\mu}{s/\sqrt{n}}$
Testing a mean	unknown	$n < 30$	$t = \frac{\bar{x}-\mu}{s/\sqrt{n}}$ with $df = n-1$

Step 4: Formulate the Decision rule

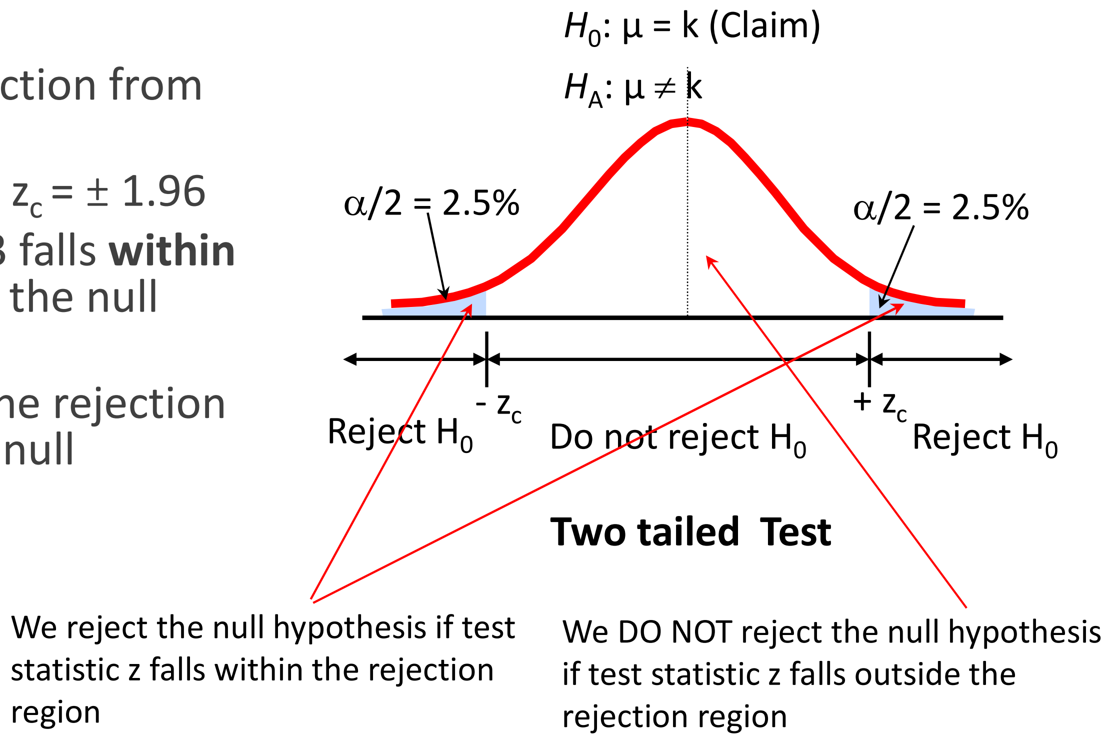
- The decision rule specifies the conditions under which the null hypothesis is rejected or not rejected.
- The decision is
  - ✓ **To reject** the null hypothesis if the test statistic falls **in** the rejection region
  - ✓ **Not to reject** the null hypothesis if the test statistic falls **outside** the rejection region
- What is the rejection region?

Rejection Region

A rejection region of the sampling distribution is the range of values for which the null hypothesis is not probable.

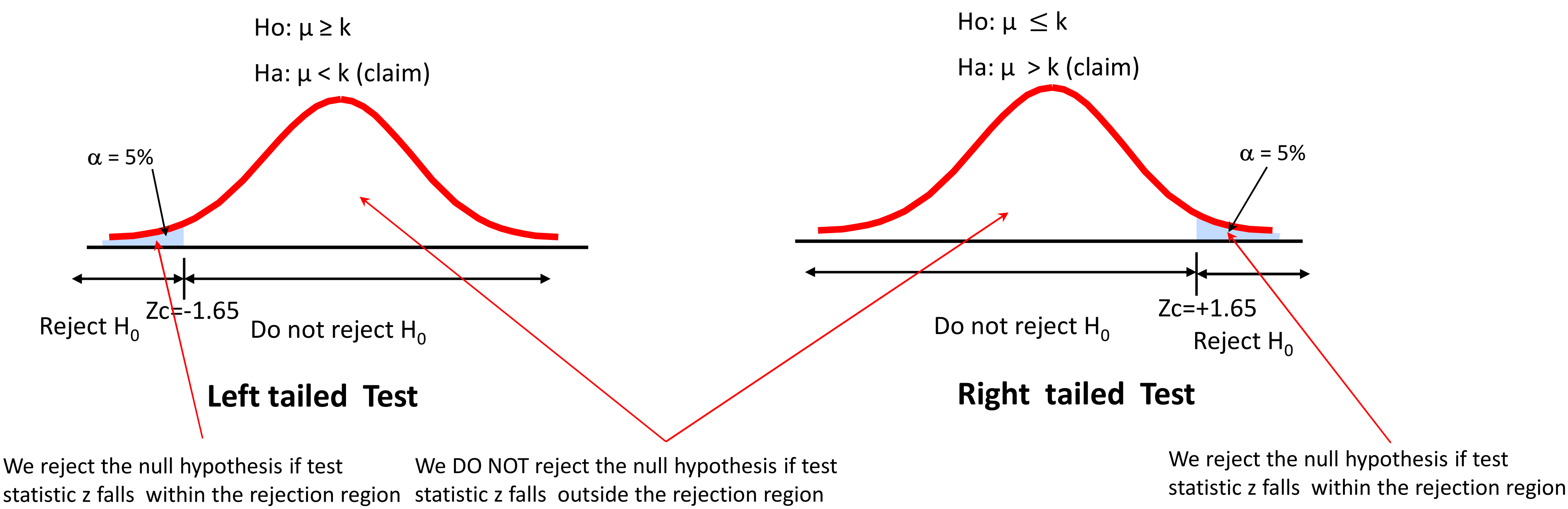
Step 4: Formulate the Decision rule

- Example 9a: For a level of significance, α = 0.05, the rejection region for Two Tailed test is illustrated below.
  - ✓ The Critical Value, z_c, separates the rejection from non-rejection region
  - ✓ Based on α = 0.05 for a Two Tailed test, $z_c = \pm 1.96$
  - ✓ If computed test statistic, z from Step 3 falls **within** the rejection region (in blue), we **reject** the null hypothesis
  - ✓ If computed test statistic falls **outside** the rejection region (in white), we **do not reject** the null hypothesis



Step 4: Formulate the Decision rule

- Example 9b: For a level of significance, α = 0.05, the rejection region for Left and Right Tailed tests is illustrated below.



Step 5: Summarise results

- In hypothesis testing, we are not trying to prove the hypothesis is true.
- If the decision is **not to** reject the null hypothesis, based on sample data, we can only conclude that the difference between the sample mean and the hypothesized population mean was not large enough to reject the null hypothesis.
- The following table will help you summarize your results.

Decision	Claim is H_o	Claim is H_a
Reject H_o	There is enough evidence to reject the claim	There is enough evidence to support the claim
DO NOT reject H_o	There is not enough evidence to reject the claim	There is not enough evidence to support the claim

Example 10 (known σ)

A recent newspaper report claimed that the average amount of money in the POS Bank saving accounts held by school children was $1,000 with a standard deviation of $50. A random sample of 50 accounts is taken and found to have a mean of $950. At level of significance = 0.01, test the claim made by the report.

1. State the Null and Alternate hypothesis

   H_o: μ = 1000 (Claim) H_a : $\mu \neq 1000$

Rejection Criteria:

Reject H_o if $z < -2.575$ or $z > 2.575$

2. Identify (a) Type of Test (b) Level of Significance

   (a) Two Tailed Test (b) $\alpha = 0.01$

   Since σ^2 is known, then use z-distribution table

   From z-distribution table, for $\alpha/2=0.005$, critical value, $z_c = \pm 2.575$

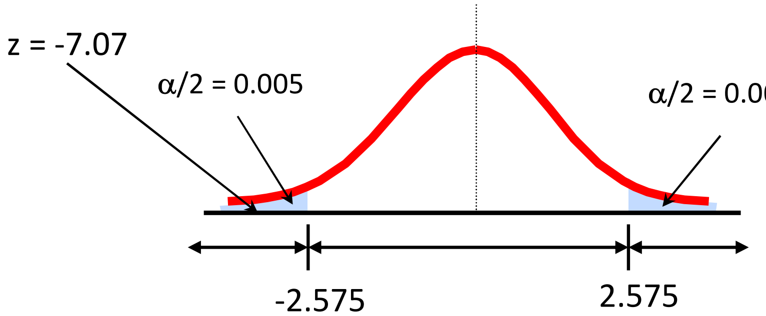
Example 10 (known σ)

3. Compute the test statistic

μ = 1000 n =50

$\bar{x}$ = 950 σ = 50

test statistic, $z = \frac{\bar{x}-\mu}{\frac{\sigma}{\sqrt{n}}} = \frac{950-100}{\frac{50}{\sqrt{50}}}$ = -7.07



4. Formulate the decision rule

Since the test statistic, z = -7.07 is in the rejection region, we reject null hypothesis

5. Summarise results

At α = 0.01, there is **enough evidence to reject** the claim that the average amount of money in the POS Bank saving accounts held by school children was $1,000

Example 11 (for $n \geq 30$ and unknown σ)

The Nanyang Polytechnic Student Union maintains that, on the average, students travel less than 25 minutes in order to reach the campus each day. The Student Affairs Office obtained a random sample of 36 one-way travel times from the students. The sample had a mean of 19.4 minutes and a standard deviation of 9.6 minutes. Does the Student Affairs Office have sufficient evidence to support the Student Union Claim? Use level of significance, $\alpha = 0.01$.

1. State the Null and Alternate hypothesis

   $H_o: \mu \geq 25$ $\qquad$ $H_a: \mu < 25$ (claim)

2. Identify (a) Type of Test (b) Level of Significance

   (a) Left Tailed Test (b) $\alpha = 0.01$

   From z-distribution table, for $\alpha = 0.01$, critical value, $z_c = -2.33$

Rejection Criteria:

Reject H_o if $z < -2.33$

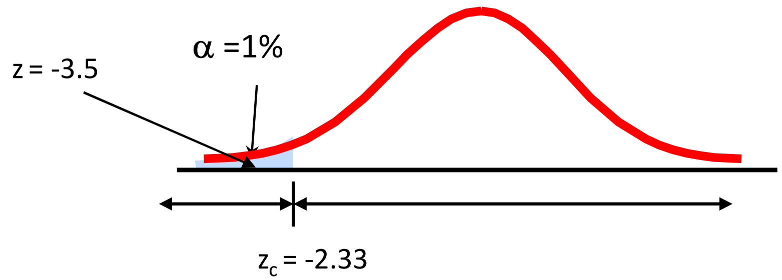
Example 11 (for $n \geq 30$ and unknown σ)

3. Compute the test statistic

   $\mu = 25$ $\quad$ $n = 36$

   $\bar{x} = 19.4$ $\quad$ $s = 9.6$

   test statistic, $z = \dfrac{\bar{x}-\mu}{\frac{s}{\sqrt{n}}} = \dfrac{19.4-25}{\frac{9.6}{\sqrt{36}}} = -3.5$



4. Formulate the decision rule

   Since the test statistic, z = -3.5 is in the rejection region, we reject null hypothesis

5. Summarise results

   At α=0.01, Student Affairs office has **enough evidence to support** the claim that students travel less than 25 minutes in order to reach the campus each day.

Example 12 (for $n < 30$ and unknown σ)

A used car dealer says that the mean price of a 6 year old Suzuki is at least $23,900. You suspect this claim is incorrect and find that a random sample of 14 similar vehicles has a mean price of $23,000 and a standard deviation of $1113. Is there enough evidence to reject the dealer's claim at $\alpha=0.05$? Assume the population is normally distributed.

1. State the Null and Alternate hypothesis

   $H_o: \mu \geq 23900$ (Claim) $H_a: \mu < 23900$

   Rejection Criteria:
   Reject H_o if $t < -1.771$

2. Identify (a) Type of Test (b) Level of Significance

   (a) Left Tailed Test (b) $\alpha=0.05$.

   From t-distribution table, for df = 14-1 =13, and $\alpha=0.05$, critical value, $t_c = -1.771$

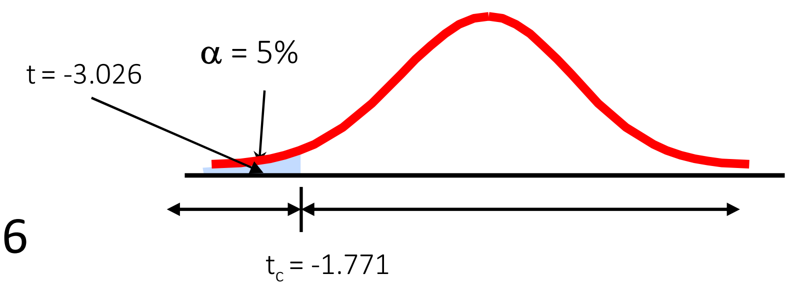
Example 12 (for n < 30 and unknown σ)

3. Compute the test statistic

   μ = 23900   n =14

   $\bar{x}$ = 23000   s = 1113

   test statistic, $t = \frac{\bar{x}-\mu}{\frac{s}{\sqrt{n}}} = \frac{23000-2390}{\frac{1113}{\sqrt{14}}}$ = -3.026



4. Formulate the decision rule

   Since the test statistic, t = -3.026 is in the rejection region, we reject null hypothesis

5. Summarise results

   At α=0.05, there is **enough evidence to reject** the claim that mean price of a 6 year old Suzuki is at least $23,900.

Example 13 (for $n < 30$ and unknown σ)

A personnel specialist of a major company conducted an aptitude test when recruiting a large number of employees. During the testing process, management asks how things are going and she replies " Fine, I think the average score will be 90." When the specialist reviews 20 of the tests results, she finds that the mean score is 84 and the standard deviation is 11. Test at 0.01 level of significance if the claim that the average score on the aptitude test is 90 can be supported?

1. State the Null and Alternate hypothesis

   $H_o: \mu = 90$ (Claim) $H_a: \mu \neq 90$

Rejection Criteria:
Reject H_o if $t < -2.861$ or $t > 2.861$

2. Identify (a) Type of Test (b) Level of Significance

   (a) Two Tailed Test (b) $\alpha=0.01$.
   From t-distribution table, for df = 20-1 =19, and $\alpha=0.01$, critical value, $t_c = \pm 2.861$

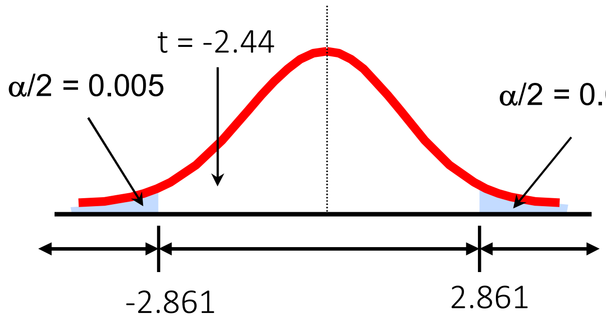
Example 13 (for n < 30 and unknown σ)

3. Compute the test statistic

   $\mu = 90 \qquad n = 20$

   $\bar{x} = 84 \qquad s = 11$

   test statistic, $t = \dfrac{\bar{x}-\mu}{\frac{s}{\sqrt{n}}} = \dfrac{84-90}{\frac{11}{\sqrt{20}}} = -2.44$



4. Formulate the decision rule

   Since the test statistic, t = -2.44 is outside the rejection region, we do not reject null hypothesis

5. Summarise results

   At α=0.01, there is **not enough evidence to reject** the claim that average score on the aptitude test is 90. Hence, the claim can be supported.