جامعة الملك سعود
King Saud University

College of Science
Department of Statistics & OR

STAT 109
Biostatistics

Chapter 7

Using Sample Statistics To Test Hypotheses About Population Parameters

October 2025

NOTE: This presentation is based on the presentation prepared thankfully by Professor Abdullah al-Shiha

7.1 Introduction

7.2 Hypothesis Testing: A Single Population Mean (μ)

7.3 Hypothesis Testing: The Difference Between Two Population Mean independent ($\mu_1 - \mu_2$)

7.4 Paired Comparisons ($\mu_D = \mu_1 - \mu_2$)

7.5 Hypothesis Testing: A Single Population Proportion (P)

7.6 Hypothesis Testing: The Difference Between Two Population Proportions ($P_1 - P_2$)

7.1 Introduction

Consider a population with some unknown parameter θ. We are interested in testing (confirming or denying) some conjectures about θ. For example, we might be interested in testing the conjecture that $\theta > \theta_o$, where θ_o is a given value.

- A hypothesis is a statement about one or more populations.
- A research hypothesis is the conjecture or supposition that motivates the research.
- A statistical hypothesis is a conjecture (or a statement) concerning the population which can be evaluated by appropriate statistical technique.
- For example, if θ is an unknown parameter of the population, we might be interested in testing the conjecture sating that $\theta \geq \theta_o$ against $\theta < \theta_o$ (for some specific value θ_o).

The hypothesis

- We usually test the null hypothesis (H_o) against the alternative (or the research) hypothesis (H_1 or H_A) by choosing one of the following situations:

(i)	$H_o: \theta = \theta_o$	against	$H_A: \theta \neq \theta_o$
(ii)	$H_o: \theta \geq \theta_o$	against	$H_A: \theta < \theta_o$
(iii)	$H_o: \theta \leq \theta_o$	against	$H_A: \theta > \theta_o$

- Equality sign must appear in the null hypothesis (H_o) .

- H_o is the null hypothesis, and H_A is the alternative hypothesis.
  (H_o and H_A are complement of each other)

- The null hypothesis (H_o) is also called "the hypothesis of no difference".

- The alternative hypothesis (H_A) is also called the research hypothesis.

There are 4 possible situations in testing a statistical hypothesis:

		Condition of Null Hypothesis Ho (Nature/reality)	
		Ho is true	Ho is false
Possible Action (Decision)	Accepting Ho	Correct Decision	Type II error (β)
	Rejecting Ho	Type I error (α)	Correct Decision

- There are two types of Errors:
  - Type I error = Rejecting Ho when Ho is true
  - P(Type I error) = P(Rejecting Ho | Ho is true) = α

  - Type II error = Accepting Ho when Ho is false
  - P(Type II error) = P(Accepting Ho | Ho is false) = β

- The level of significance of the test is the probability of rejecting true Ho:

  α = P(Rejecting Ho | Ho is true) = P(Type I error)

- There are 2 types of alternative hypothesis:
  - One-sided alternative hypothesis (H_A):
    - $H_0: \theta \geq \theta_0$ against $H_A: \theta < \theta_0$
    - $H_0: \theta \leq \theta_0$ against $H_A: \theta > \theta_0$
  - Two-sided alternative hypothesis (H_A):
    - $H_0: \theta = \theta_0$ against $H_A: \theta \neq \theta_0$

- We will use the terms "accepting" and "not rejecting" interchangeably. Also, we will use the terms "acceptance" and "nonrejection" interchangeably.

  We will use the terms "accept" and "fail to reject" interchangeably

The Procedure of Testing H_o (against H_A):

The test procedure for rejecting H_0 (accepting H_A) or accepting H_0 (rejecting H_A) involves the following steps:

1. **Determine the hypothesis :** Null hypothesis (H_0) and Alternative hypothesis (H_A) .

2. **Determining a test statistic (T.S.)**

We choose the appropriate test statistic <u>based on the point estimator</u> of the parameter.

The test statistic has the following form:

$$\textbf{Test statistic} = \frac{\textbf{Estimate} - \textbf{hypothezed parameter}}{\textbf{Standard error of the estimate}}$$

3. **Determining the level of significance (α):**

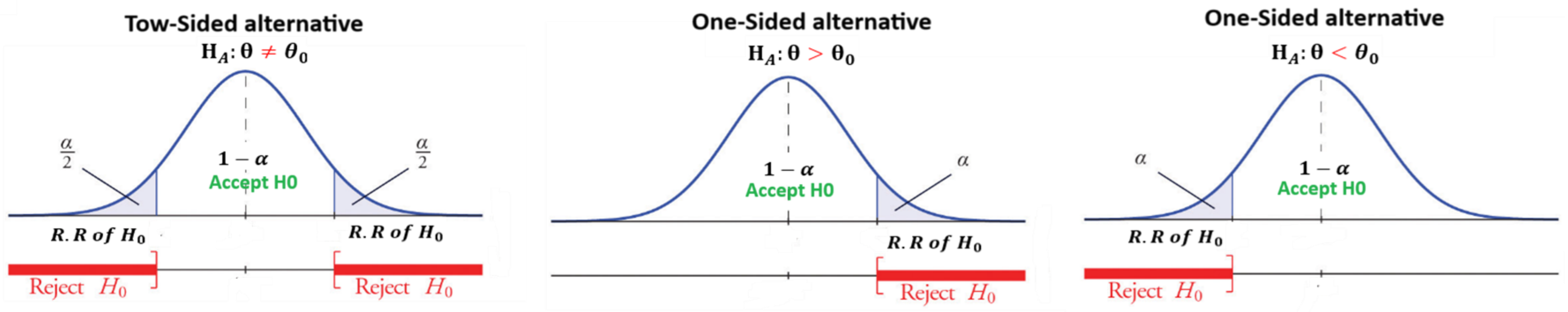
$\alpha = 0.01, 0.025, 0.05, 0.10$

4. **Determining** the **rejection region of H_0** (R.R.) and the **acceptance region of H_0** (A.R.).

The R.R. of H_0 depends on H_A and α

- H_A determines the direction of the R.R. of H_0 .
- α determines the size of the R.R. of H_0 . (α = the size of the R.R. of H_0 = shaded area)

The rejection region(R.R of H_0) depends on the sign of H_A and α :



5. **Decision:**

We reject H_0 (and accept H_A) if the value of the test statistic (T.S.) belongs to the R.R. of H_0 ,and vice versa.

Notes:

1. The rejection region of H_0 (R.R.) is sometimes called "the critical region".

2. The values which separate the rejection region (R.R.) and the acceptance region (A.R.) are called the critical values or "Reliability coefficient" .

7.2 Hypothesis Testing: A Single Population Mean (μ)

King Saud University

Suppose that $x_1, x_2, \dots, x_n$ is a random sample of size n from a distribution (or population) with mean μ and variance σ^2.
We need to test some hypotheses (make some statistical inference) about the mean (μ).

(1) First case:

Assumptions:

-The variance σ^2 is known + Normal distribution with any sample size .

or

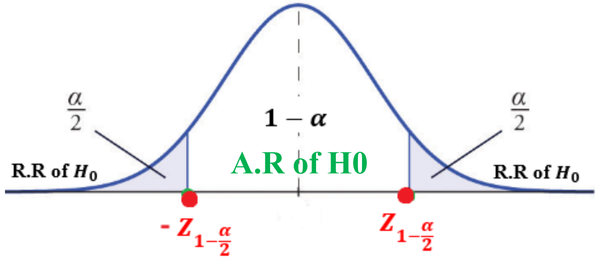
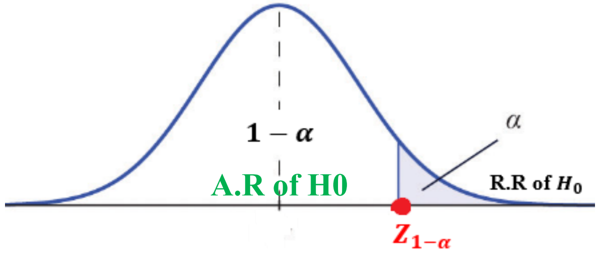
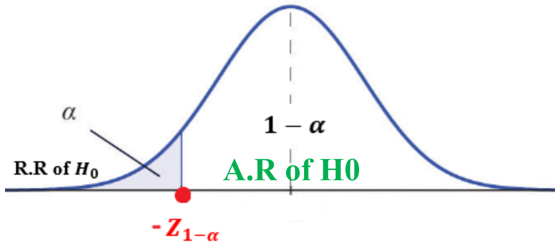
-The variance σ^2 is known +Non-normal distribution with large sample size ($n \geq 30$).

(2) Second case:

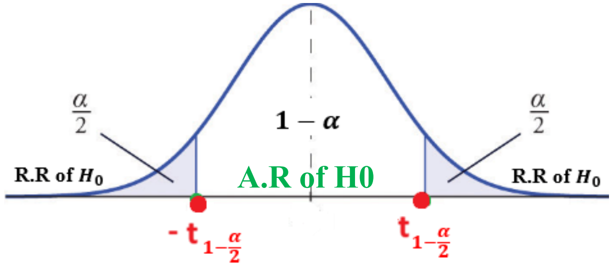
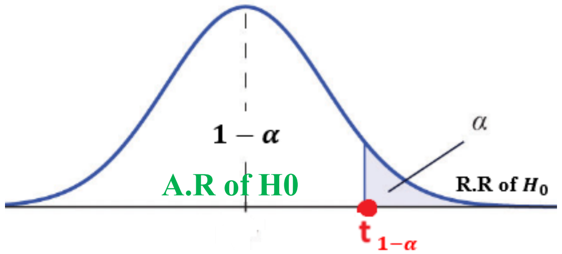
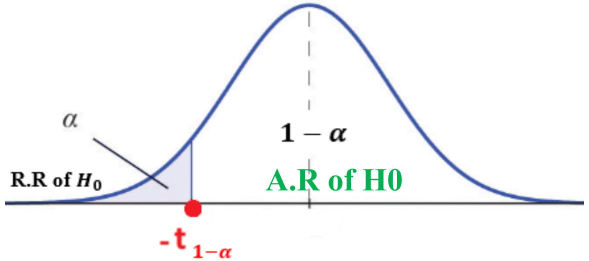
Assumptions:

-The variance σ^2 is unknown +Normal distribution + Small sample size ($n < 30$).

Test procedures:

Hypotheses	$H_0: \mu = \mu_0$ $H_A: \mu \neq \mu_0$	$H_0: \mu \leq \mu_0$ $H_A: \mu > \mu_0$	$H_0: \mu \geq \mu_0$ $H_A: \mu < \mu_0$
First case	σ **is known +Normal distribution or** σ **is known +Non-normal distribution (n large =** $n \geq 30$ **)**		
Test Statistic (T.S.)	$Z = \frac{\bar{X} - \mu_0}{\sigma / \sqrt{n}} \sim N(0,1)$		
R.R. & A.R. of H_0			
Critical value(s)	$Z_{1-\frac{\alpha}{2}}$ ***and*** $-Z_{1-\frac{\alpha}{2}}$	$Z_{1-\alpha}$	$-Z_{1-\alpha}$
Decision:	We reject H_0 (and accept H_A) at the significance level α if:		
	$Z_{T.S} > Z_{1-\frac{\alpha}{2}}$ or $Z_{T.S} < -Z_{1-\frac{\alpha}{2}}$	$Z_{T.S} > Z_{1-\alpha}$	$Z_{T.S} < -Z_{1-\alpha}$
	T.S. ∈ R.R. **Two - Sided Test**	**T.S. ∈ R.R.** **One - Sided Test**	**T.S. ∈ R.R.** **One - Sided Test**

Test procedures:

Hypotheses	$H_0 : \mu = \mu_0$ $H_A : \mu \neq \mu_0$	$H_0 : \mu \leq \mu_0$ $H_A : \mu > \mu_0$	$H_0 : \mu \geq \mu_0$ $H_A : \mu < \mu_0$
Second case	σ **is unknown, Normal distribution and n small (n <30)**		
Test Statistic (T.S.)	$T = \frac{\bar{X}-\mu_0}{S/\sqrt{n}} \sim t_{n-1} \quad df = n-1$		
R.R. & A.R. of H_0			
Critical value(s)	$t_{1-\frac{\alpha}{2}}$ ***and*** $-t_{1-\frac{\alpha}{2}}$	$t_{1-\alpha}$	$-t_{1-\alpha}$
Decision:	We reject H_0 (and accept H_A) at the significance level α if:		
	$T_{T.S} > t_{1-\frac{\alpha}{2}}$ or $T_{T.S} < -t_{1-\frac{\alpha}{2}}$	$T_{T.S} > t_{1-\alpha}$	$T_{T.S} < -t_{1-\alpha}$
	T.S. ∈ R.R. **Two - Sided Test**	**T.S. ∈ R.R.** **One - Sided Test**	**T.S. ∈ R.R.** **One - Sided Test**

Example : First case (Variance σ^2 known)

A random sample of 100 recorded deaths in the United States during the past year showed an average of 71.8 years. Assuming a population standard deviation of 8.9 year . Does this seem to indicate that the mean life span today is greater than 70 years? Use a 0.05 level of significance.

Solution:

$n = 100$ (large $n \geq 30$)

$\sigma = 8.9$ (σ known)

$\overline{X} = 71.8$

$\alpha = 0.05$ (level of significance)

μ = average (mean)life span

$\mu_0 = 70$

First case:

Assumptions:

- The variance σ^2 is known .
- Non-normal distribution with large sample size

Hypotheses:

$$H_0 : \mu \leq 70 \qquad (\mu_0 = 70)$$
$$H_A : \mu > 70 \quad \text{(research hypothesis)}$$

Test statistics (T.S.)

$$Z = \frac{\bar{X} - \mu_0}{\sigma / \sqrt{n}} = \frac{71.8 - 70}{8.9/\sqrt{100}} = 2.02$$

Rejection Region of Ho (R.R.): (critical region): $(1.645 , \infty)$

$\alpha = 0.05 \gg$ **Critical Value :** $Z_{1-\alpha} = Z_{1-0.05} = Z_{0.95} = 1.645$

Decision :

Reject $H_0 : \mu \leq 70$ *if* :

$$Z_{T.S} > Z_{1-\alpha}$$

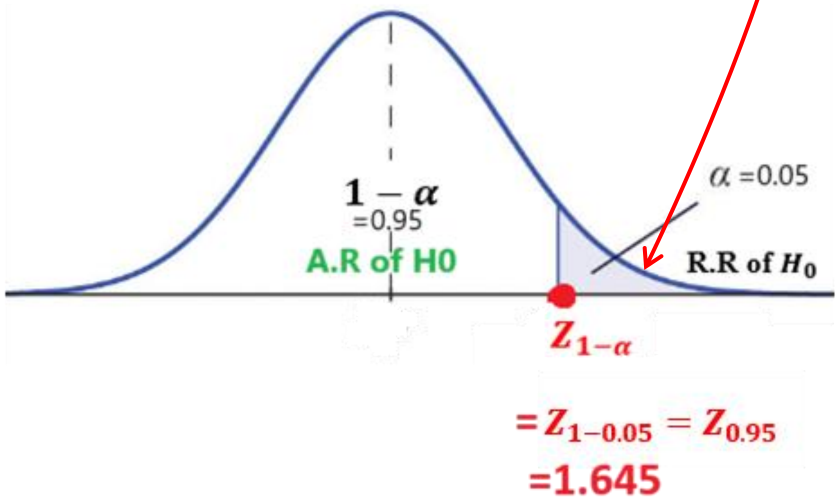
$2.02 > 1.645$ condition satisfied

Decision is Reject $H_0 : \mu \leq 70$ and Accept $H_A : \mu > 70$

Therefore, we conclude that the mean life span today is greater than 70 years.

Another way to take decision is by graph :

Determine test statistics value on the graph



Third way we can use to take decision by Using P- Value as a decision tool:

P-value : is the smallest value of α for which we can reject the null hypothesis H_0.

Calculate P-value :

- Calculating P-value depends on the alternative hypothesis H_A.
- Suppose that $Z_c = Z_{T.S} = \frac{\bar{X}-\mu_0}{\sigma/\sqrt{n}}$ is the computed value of the test Statistic.

-The following table illustrates how to compute P-value, and how to use P-value for testing the null hypothesis:

Alternative Hypothesis (H_A)	$H_A : \mu \neq \mu_0$	$H_A : \mu > \mu_0$	$H_A : \mu < \mu_0$
P-Value =	$2 \times P(Z > \mid Z_c \mid)$	$P(Z > Z_c)$	$P(Z < Z_c)$
Significance Level =	α		
Decision:	Reject H_0 if P-value $<$ α		

Example : For the previous example

Hypotheses:

$H_0 : \mu \leq 70 \qquad (\mu_0 = 70)$

$H_A : \mu > 70$ (research hypothesis)

Test statistics (T.S.)

$$Z = \frac{\bar{X} - \mu_0}{\sigma / \sqrt{n}} = \frac{71.8 - 70}{8.9/\sqrt{100}} = 2.02$$

To calculate P- value The alternative hypothesis was H_A: $\mu > 70$, $\alpha = 0.05$

P-value = $P(Z > Z_c) = P(Z > 2.02) = 1 - P(Z < 2.02) = 1 - 0.9783 = 0.0217$

Decision :

Reject H_0 if P-value $< \alpha = 0.05$

$0.0217 < \alpha = 0.05$ (condition satisfied)

Then , we Reject H_0 .

Example: Second case (Variance σ^2 unknown)

The manager of a private clinic claims that the mean time of the patient-doctor visit in his clinic is 8 minutes. Test the hypothesis that $\mu=8$ minutes against the alternative that $\mu\neq8$ minutes . If a random sample of 25 patient-doctor visits yielded a mean time of 7.8 minutes with a standard deviation of 0.5 minutes. It is assumed that the distribution of the time of this type of visits is normal. Use a 0.01 level of significance.

Solution:

Normal distribution ,

$\overline{X} = 7.8$,

$\alpha=0.01$ (level of significance)

$S = 0.5$ Sample standard deviation (σ is unknown)

$n = 25$ (n small $n < 30$)

μ= mean time of the visit

Second case:

Assumptions:

- The variance σ^2 is unknown .
- Normal distribution
- Any sample size .

Hypotheses:

$H_0 : \mu = 8 \quad (\mu_0 = 8)$

$H_A : \mu \neq 8$ (research hypothesis)

Test statistics (T.S.)

$$T = \frac{\bar{X} - \mu_0}{S/\sqrt{n}} = \frac{7.8 - 8}{0.5/\sqrt{25}} = -2$$

Rejection Region of Ho (R.R.): (critical region) :$(-\infty, -2.797) \cup (2.797, \infty)$

$\alpha = 0.01 \gg$ **Critical Value :** $\pm t_{1-\frac{\alpha}{2}} = \pm t_{1-\frac{0.01}{2}} = \pm t_{0.995} = \pm 2.797$

df =n-1 =25-1 =24

Decision :

Reject $H_0 : \mu = 8$ *if* : $T_{T.S} < -T_{1-\frac{\alpha}{2}}$ *or* $T_{T.S} > T_{1-\frac{\alpha}{2}}$

$-2 \nless -2.797$ (First condition not satisfied, then try the second condition)

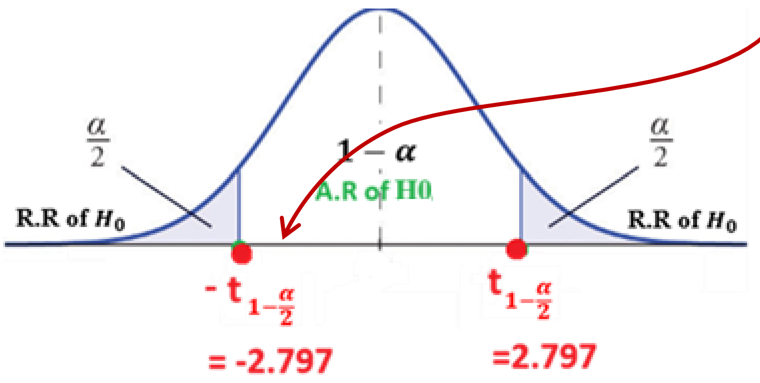
$-2 \ngtr 2.797$ (Second condition not satisfied)

Decision is Accepte $H_0 : \mu = 8$ and Reject $H_A : \mu \neq 8$

Therefore, we conclude that the claim is correct.

Another way to take decision is by graph :

Determine test statistics value on the graph



Special case :

For the case of non-normal population with unknown variance, and when the sample size is large ($n \geq 30$),

we may use the following test statistic:

$$Z = \frac{\bar{X} - \mu_0}{S/\sqrt{n}} \sim N(0,1)$$

That is, we replace the population standard deviation (σ) by the sample standard deviation (S), and we conduct the test as described for the first case.

7.3 Hypothesis Testing: The Difference Between Two Population Means ($\mu_1 - \mu_2$) : (Independent Populations)

Suppose that we have two (independent) populations:

- 1-st population with mean μ_1 and variance σ_1^2 .
- 2-nd population with mean μ_2 and variance σ_2^2 .
- We are interested in comparing μ_1 and μ_2 , or equivalently, making inferences about the difference between the means $(\mu_1 - \mu_2)$.
- We independently select a random sample of size n_1 from the 1-st population and another random sample of size n_2 from the 2-nd population:
- Let $\bar{X}_1$ and S_1^2 be the sample mean and the sample variance of the 1-st sample.
- Let $\bar{X}_2$ and S_2^2 be the sample mean and the sample variance of the 2-nd sample.
- The sampling distribution of $\bar{X}_1$ - $\bar{X}_2$ is used to make inferences about $\mu_1 - \mu_2$.

We wish to test some hypotheses comparing the population means.

Hypotheses:

We choose one of the following situations:

(i) $H_0: \mu_1 = \mu_2$ against $H_A: \mu_1 \neq \mu_2$
(ii) $H_0: \mu_1 \geq \mu_2$ against $H_A: \mu_1 < \mu_2$
(iii) $H_0: \mu_1 \leq \mu_2$ against $H_A: \mu_1 > \mu_2$

or equivalently,

(i) $H_0: \mu_1 - \mu_2 = 0 (\mu_0)$ against $H_A: \mu_1 - \mu_2 \neq 0 \; (\mu_0)$
(ii) $H_0: \mu_1 - \mu_2 \geq 0 (\mu_0)$ against $H_A: \mu_1 - \mu_2 < 0 \; (\mu_0)$
(iii) $H_0: \mu_1 - \mu_2 \leq 0 (\mu_0)$ against $H_A: \mu_1 - \mu_2 > 0 \; (\mu_0)$

Test Statistic (T.S):

(1) First Case:

Assumptions:

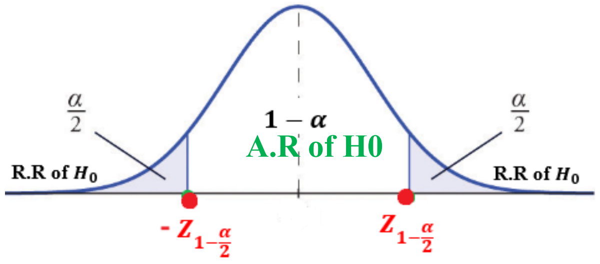
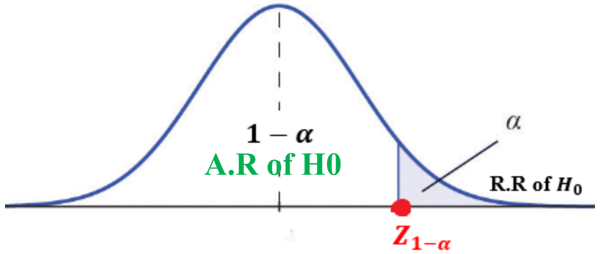
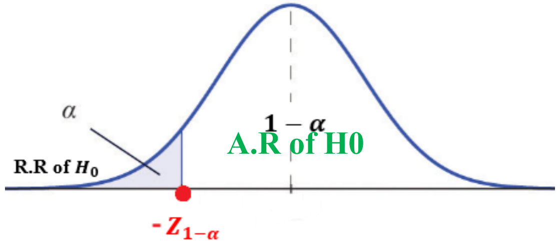
-Normal populations + σ_1^2 and σ_2^2 are known.

- Non-normal populations with large sample sizes$(n_1, n_2 \geq 30)$ + σ_1^2 and σ_2^2 are known

The test statistic is:

$$\mathbf{Z} = \frac{(\bar{X}_1 - \bar{X}_2) - \mu_0}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}} \sim N(0,1)$$

Test procedures:

Hypotheses	$H_0: \mu_1 - \mu_2 = \mu_0$ $H_A: \mu_1 - \mu_2 \neq \mu_0$	$H_0: \mu_1 - \mu_2 \leq \mu_0$ $H_A: \mu_1 - \mu_2 > \mu_0$	$H_0: \mu_1 - \mu_2 \geq \mu_0$ $H_A: \mu_1 - \mu_2 < \mu_0$
First case	**if σ_1^2 and σ_2^2 are known , Normal populations or** **σ_1^2 and σ_2^2 are known , non-normal populations with large sample sizes**		
Test Statistic (T.S.)	$Z = \frac{(\bar{X}_1 - \bar{X}_2) - \mu_0}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}} \sim N(0,1)$		
R.R. & A.R. of H_0			
Critical value(s)	$Z_{1-\frac{\alpha}{2}}$ *and* $-Z_{1-\frac{\alpha}{2}}$	$Z_{1-\alpha}$	$-Z_{1-\alpha}$
Decision:	We reject H_0 (and accept H_A) at the significance level α if:		
	$Z_{T.S} > Z_{1-\frac{\alpha}{2}}$ or $Z_{T.S} < -Z_{1-\frac{\alpha}{2}}$	$Z_{T.S} > Z_{1-\alpha}$	$Z_{T.S} < -Z_{1-\alpha}$
	T.S. ∈ R.R. **Two - Sided Test**	**T.S. ∈ R.R.** **One - Sided Test**	**T.S. ∈ R.R.** **One - Sided Test**

(2) Second Case:

Assumptions:

- Normal populations .
- σ_1^2 and σ_2^2 are unknown but equal $\sigma_1^2 = \sigma_2^2 = \sigma^2$,
- small sample size ($n_1, n_2 < 30$)

then the test statistic is:

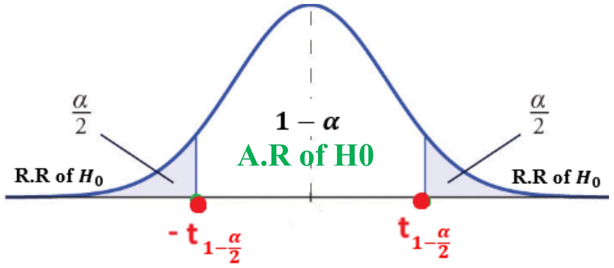
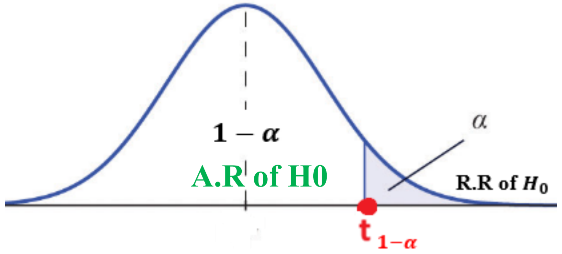
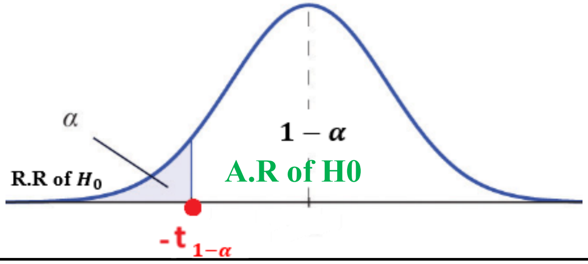
$$T = \frac{(\bar{X}_1 - \bar{X}_2) - \mu_0}{\sqrt{\frac{S_P^2}{n_1} + \frac{S_P^2}{n_2}}} \sim t(n_1 + n_2 - 2)$$

where the pooled variance estimate of σ^2 is

$$S_P^2 = \frac{(n_1 - 1)S_1^2 + (n_2 - 1)S_2^2}{n_1 + n_2 - 2}$$

and the degrees of freedom of is $df = \text{v} = n_1 + n_2 - 2.$

Test procedures:

Hypotheses	$H_0: \mu_1 - \mu_2 = \mu_0$ $H_A: \mu_1 - \mu_2 \neq \mu_0$	$H_0: \mu_1 - \mu_2 \leq \mu_0$ $H_A: \mu_1 - \mu_2 > \mu_0$	$H_0: \mu_1 - \mu_2 \geq \mu_0$ $H_A: \mu_1 - \mu_2 < \mu_0$
Second case	**if $\sigma_1^2 = \sigma_2^2 = \sigma^2$ is unknown + Normal populations + $n_1, n_2 < 30$**		
Test Statistic (T.S.)	$T = \frac{(\bar{X}-\bar{X}_2)-\mu_0}{\sqrt{\frac{S_P^2}{n_1}+\frac{S_P^2}{n_2}}} \sim t_{n_1+n_2-2}$, $S_P^2 = \frac{(n_1-1)S_1^2+(n_2-1)S_2^2}{n_1+n_2-2}$, $df = n_1 + n_2 - 2$		
R.R. & A.R. of H_0			
Critical value(s)	$t_{1-\frac{\alpha}{2}}$ ***and*** $-t_{1-\frac{\alpha}{2}}$	$t_{1-\alpha}$	$-t_{1-\alpha}$
Decision:	We reject H_0 (and accept H_A) at the significance level α if:		
	$T_{T.S} > t_{1-\frac{\alpha}{2}}$ or $T_{T.S} < -t_{1-\frac{\alpha}{2}}$	$T_{T.S} > t_{1-\alpha}$	$T_{T.S} < -t_{1-\alpha}$
	T.S. ∈ R.R. **Two - Sided Test**	**T.S. ∈ R.R.** **One - Sided Test**	**T.S. ∈ R.R.** **One - Sided Test**

Example: (σ_1^2 and σ_2^2 are known)

Researchers wish to know if the data they have collected provide sufficient evidence to indicate the difference in mean serum uric acid levels between individuals with Down's syndrome(1) and normal individuals(2). The data consist of serum uric acid on 12 individuals with Down's syndrome and 15 normal individuals.

The sample means are $\bar{X}_1 = 4.5$ mg/100 ml and $\bar{X}_2 = 3.4$ mg/100 ml. Assume the populations are normal with variances $\sigma_1^2 = 1$ and $\sigma_2^2 = 1.5$. Use significance level α=0.05.

Solution:

μ_1= mean serum uric acid levels for the individuals with Down's syndrome.

μ_2 = mean serum uric acid levels for the normal individuals.

$n_1 = 12$, $\bar{X}_1 = 4.5$, $\sigma_1^2 = 1$ (σ_1^2 known)

$n_2 = 15$, $\bar{X}_2 = 3.4$, $\sigma_2^2 = 1.5$ (σ_2^2 known)

Normal populations

First case:

Assumptions:

- The variance σ_1^2, σ_2^2, is known .
- Normal distribution .

Hypotheses:

$$H_0 : \mu_1 = \mu_2 \quad \text{OR} \quad (H_0: \mu_1 - \mu_2 = 0)$$
$$H_A : \mu_1 \neq \mu_2 \qquad (H_0: \mu_1 - \mu_2 \neq 0)$$

Test statistics (T.S.)

$$Z = \frac{(\bar{X}_1 - \bar{X}_2) - \mu_0}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}} = \frac{4.5 - 3.4}{\sqrt{\frac{1}{12} + \frac{1.5}{15}}} = 2.569$$

Rejection Region of Ho (R.R.): (critical region): $(-\infty, -1.96) \cup (1.96, \infty)$

$\alpha = 0.05 \gg$ **Critical Value :** $\pm Z_{1-\frac{\alpha}{2}} = \pm Z_{1-\frac{0.05}{2}} = \pm Z_{0.975} = \pm 1.96$

Decision :

Reject $H_0: \mu_1 = \mu_2$ *if* :

$$Z_{T.S} < -Z_{1-\frac{\alpha}{2}} \quad or \quad Z_{T.S} > Z_{1-\frac{\alpha}{2}}$$

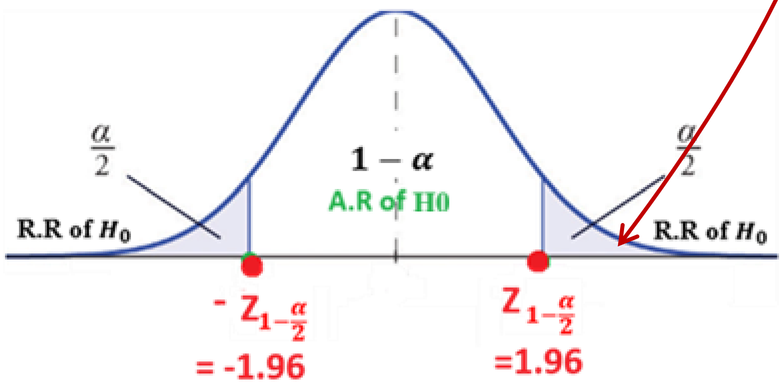
$2.569 \not< -1.96$ (First condition not satisfied, then try the second condition)

$2.569 > 1.96$ (Second condition satisfied)

Decision is Reject $H_0: \mu_1 = \mu_2$ and Accept $H_A: \mu_1 \neq \mu_2$
we conclude that the two populations means are not equal.

Another way to take decision is by graph :

Determine test statistics value on the graph



Decision by using P-value :

Calculate the P-value

P-value = 2x P(Z > $|Z_c|$) (where Z_c = Z(Test statistic)=2.569)

= 2x P(Z > |2.569|)

= 2x P(Z > 2.57)

= 2x [1-P(Z < 2.57)]

= 2x [1- 0.99492]

= 2x 0.00508

= 0.01016

P-value = 0.01016

Decision : **Reject H_0 if P-value < α =0.05**

Since P-value= 0.0106 < **α =0.05**

We reject H_0 and accept H_A

Example: ($\sigma_1^2 = \sigma_2^2 = \sigma^2$ is unknown)

An experiment was performed to compare the abrasive wear of two different materials used in making artificial teeth. 12 pieces of material(1) were tested by exposing each piece to a machine measuring wear. 10 pieces of material(2) were similarly tested. In each case, the depth of wear was observed. The samples of material(1) gave an average wear of 85 units with a sample standard deviation of 4, while the samples of materials(2) gave an average wear of 81 and a sample standard deviation of 5.

Can we conclude at the 0.05 level of significance that the mean abrasive wear of material(1) is greater than that of material(2)? **Assume normal populations with equal variances.**

Solution:

Normal populations

$\alpha=0.05$ (level of significance)

$\sigma_1^2 = \sigma_2^2 = \sigma^2$ **is unknown**

Material(1)	Material(2)
$n_1 = 12$	$n_2 = 10$
$\bar{X}_1 = 85$	$\bar{X}_2 = 81$
$S_1 = 4$	$S_2 = 5$

Second case:

Assumptions:

-The variance σ_1^2, σ_2^2, is Unknown and equal

- Normal distribution .

-$n_1, n_2 < 30$ small

Hypotheses:

$H_0: \mu_1 \leq \mu_2$ OR $(H_0: \mu_1 - \mu_2 \leq 0 \quad \mu_0 = 0)$

$H_A: \mu_1 > \mu_2$ $\quad (H_0: \mu_1 - \mu_2 > 0 \quad \mu_0 = 0)$

Test statistics (T.S.) $T = \frac{(\bar{X}_1 - \bar{X}_2) - \mu_0}{\sqrt{\frac{S_p^2}{n_1} + \frac{S_p^2}{n_2}}} = \frac{85-81}{\sqrt{\frac{20.05}{12} + \frac{20.05}{10}}} = 2.09$

pooled variance (S_p^2) : $S_P^2 = \frac{(n_1-1)\times S_1^2 + (n_2-1)\times S_2^2}{n_1+n_2-2} = \frac{(12-1)\times(4^2)+(10-1)\times(5^2)}{12+10-2} = 20.05$

Rejection Region of Ho (R.R.): (critical region : H_A one sided) $(1.725, \infty)$

$\alpha = 0.05 \gg$ **Critical Value :** $t_{1-\alpha} = t_{1-0.05} = t_{0.95} = 1.725$

df $= n_1 + n_2 - 2 = 12 + 10 - 2 = 20$

Decision :

Reject $H_0: \mu_1 \leq \mu_2$ *if* : $\quad T_{T.S} > t_{1-\alpha}$

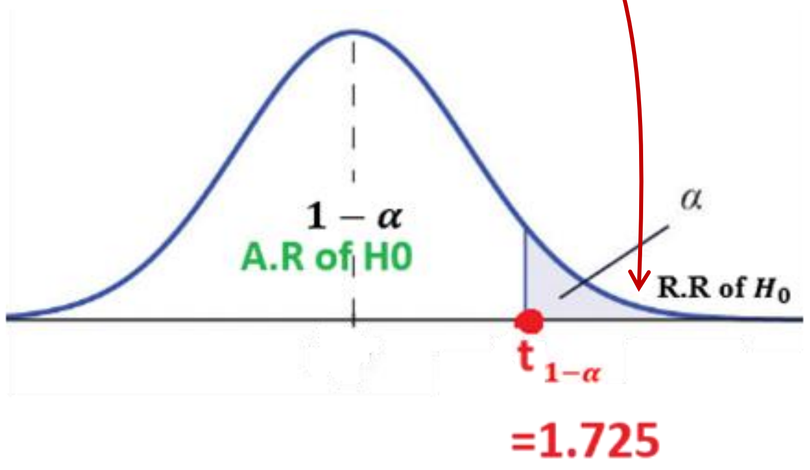
$2.09 > 1.725$ (condition satisfied)

Decision is Reject $H_0: \mu_1 \leq \mu_2$ and Accept $H_A: \mu_1 > \mu_2$

Therefore, we conclude that the mean abrasive wear of material (1) is greater than that of material (2).

Another way to take decision is by graph :

Determine test statistics value on the graph



7.4 Paired Comparisons:

- In this section, we are interested in comparing the means of two related (non-independent/dependent) normal populations.
- In other words, we wish to make statistical inference for the difference between the means of two related normal populations.
- Paired t-Test concerns about testing the equality of the means of two related normal populations.

Examples of related (dependent) populations are:

1. Height of the father and height of his son.
2. Mark of the student in MATH and his mark in STAT.
3. Pulse rate of the patient before and after the medical treatment.
4. Hemoglobin level of the patient before and after the medical treatment.

Test procedure

Let

X : the values of the first population .

Y : the values of the second population .

D = Values of X – Values of Y

Means

μ_1: Mean of the first population .

μ_2: Mean of the second population .

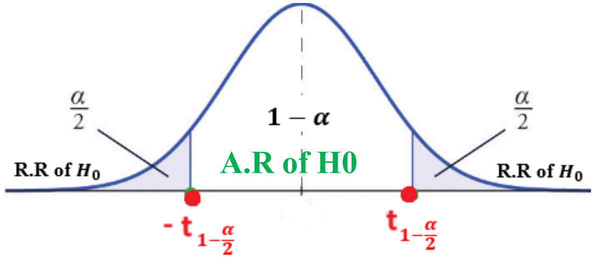
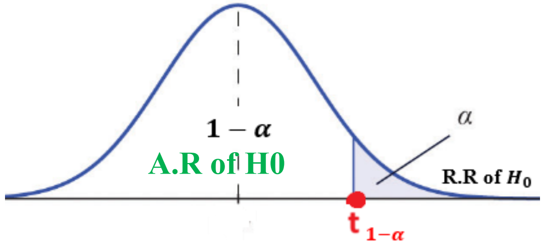
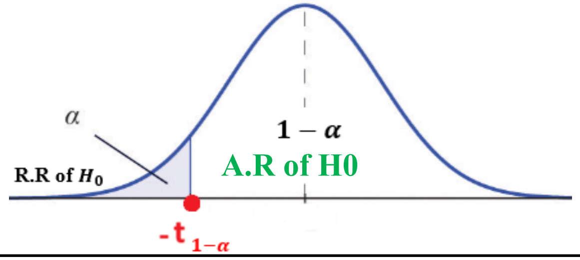
μ_D =Mean of X – Mean of Y

($\mu_D = \mu_1 - \mu_2$)

Confident interval about a difference of two population means ($\mu_D = \mu_1 - \mu_2$) : (Dependent , Related populations)

calculate the following quantities: Or (Can use the calculator)	**-The differences (D-observations):** $D_i = X_i - Y_i$ (i=1, 2, …, n) **-Sample mean of the D-observations:** $\bar{D} = \frac{\sum_{i=1}^{n} D_i}{n}$ **-Sample variance of the D-observations:** $S_D^2 = \frac{\sum_{i=1}^{n}(D_i - \bar{D})}{n-1}$ **-Sample standard deviation of the D-observations:** $S_D = \sqrt{S_D^2}$
Confident interval for $\mu_D = \mu_1 - \mu_2$	
A 100% confidence interval for μ_D	$\bar{D} \pm t_{1-\frac{\alpha}{2}} \frac{S_D}{\sqrt{n}}$, df =n-1

Testing hypothesis about a difference of two population means ($\mu_D = \mu_1 - \mu_2$) : (Dependent , Related populations)

Hypothesis	$H_o: \mu_1 - \mu_2 = \mu_0$ vs $H_A: \mu_1 - \mu_2 \neq \mu_0$ Or $H_o: \mu_D = \mu_0$ vs $H_A: \mu_D \neq \mu_0$	$H_o: \mu_1 - \mu_2 \leq \mu_0$ vs $H_A: \mu_1 - \mu_2 > \mu_0$ Or $H_o: \mu_D \leq \mu_0$ vs $H_A: \mu_D > \mu_0$	$H_o: \mu_1 - \mu_2 \geq \mu_0$ vs $H_A: \mu_1 - \mu_2 < \mu_0$ Or $H_o: \mu_D \geq \mu_0$ vs $H_A: \mu_D < \mu_0$
Test statistic(T.S)	$T = \frac{\bar{D} - (\mu_0)}{S/\sqrt{n}}$, df =n-1		
R.R. and A.R. of Ho			
Critical value (s)	$t_{(1-\frac{\alpha}{2})}$ and $-t_{(1-\frac{\alpha}{2})}$	$t_{1-\alpha}$	$-t_{1-\alpha}$
Decision :	Reject Ho (and accept HA) at the significance level α if:		
Reject Ho if the following condition satisfied.	$T_{T.S} > t_{1-\frac{\alpha}{2}}$ or $T_{T.S} < -t_{1-\frac{\alpha}{2}}$	$T_{T.S} > t_{1-\alpha}$	$T_{T.S} < -t_{1-\alpha}$
	T.S. ∈ R.R. Two-Sided Test	T.S. ∈ R.R. One-Sided Test	T.S. ∈ R.R. One-Sided Test

Example (effectiveness of a diet program)

Suppose that we are interested in study the effectiveness of a certain diet program on ten individual . Let the random variable X and Y given as following table :

Individual (i)	1	2	3	4	5	6	7	8	9	10
Weight before (Xi)	86.6	80.2	91.5	80.6	82.3	81.9	88.4	85.3	83.1	82.1
Weight after (Yi)	79.7	85.9	81.7	82.5	77.9	85.8	81.3	74.7	68.3	69.7

Assume that the populations are normal.

1. Find a 95% confidence interval for the difference between the mean of weights before the diet program (μ_1) and the mean of weights after the diet program (μ_2) [$\mu_D = \mu_1 - \mu_2$].

2. Does these data provide sufficient evidence to allow us to conclude that the diet program is effective ? Use α=0.05 .

Solution:

Let the random variables X and Y are as follows:

X = the weight of the individual **before** the diet program
Y= the weight of the same individual **after** the diet program.

We assume that the distributions of these random variables are normal with means μ_1 and μ_2, respectively .

Populations:

1-st population (X): weights **before** a diet program mean
2-nd population (Y): weights **after** the diet program mean

These two variables are related (dependent/non-independent) because they are measured on the same individual.

We select a random sample of n individuals. At the beginning of the study, we record the individuals' weights before the diet program (X) . At the end of the diet program, we record the individuals' weights after the program (Y). We end up with the following information and calculations:

Individual I	Weight before X_i	Weight after Y_i	Difference $D_i = X_i - Y_i$
1	X_1	Y_1	$D_1 = X_1 - Y_1$
2	X_2	Y_2	$D_2 = X_2 - Y_2$
. . .	. . .	. . .	. . .
n	X_n	Y_n	$D_n = X_n - Y_n$

Find the following measures by calculator or rules

- **The sample mean of the D-observations:** $\bar{D} = \frac{\sum_{i=1}^{n} D_i}{n} = \frac{54.5}{10} = 5.45$

-**Sample variance of the D-observations:** $S_D^2 = \frac{\sum_{i=1}^{n}(D_i - \bar{D})^2}{n-1} = \frac{(6.9-5.45)^2+\cdots+(12.4-5.45)^2}{10-1} = 50.328$

-**Sample standard deviation of the D-observations:** $S_D = \sqrt{S_D^2} = \sqrt{50.328} = 7.09$

First :

Calculate difference $D_i = X_i - Y_i$ (Last column)

Then, From calculator,

-Enter the data in the last column ($D_i = X_i - Y_i$)

- Find the sample mean of the D-observations ($\bar{D} = \frac{\sum_{i=1}^{n} D_i}{n} = \frac{54.5}{10} = 5.45$)

-Find the sample standard deviation of the D-observations ($S_D = 7.09$)

i	X_i	Y_i	$D_i = X_i - Y_i$
1	86.6	79.7	6.9
2	80.2	85.9	-5.7
3	91.5	81.7	9.8
4	80.6	82.5	-1.9
5	82.3	77.9	4.4
6	81.9	85.8	-3.9
7	88.4	81.3	7.1
8	85.3	74.7	10.6
9	83.1	68.3	14.8
10	82.1	69.7	12.4
Sum			$\sum_{i=1}^{10} D_i = 54.5$

1. We need to find a 95% confidence interval for $\mu_D = \mu_1 - \mu_2$

Reliability coefficient $t_{1-\frac{\alpha}{2}}$:

$$\alpha = 0.05\,, \quad t_{1-\frac{\alpha}{2}} = t_{1-\frac{0.05}{2}} = t_{0.975} = 2.262 \quad (df = n-1 = 10-1 = 9)$$

The 95% C.I for $\mu_D = \mu_1 - \mu_2$

$$\bar{D} \pm t_{1-\alpha/2} \frac{S_D}{\sqrt{n}}$$

$$5.45 \pm 2.262 \text{ x } \frac{7.09}{\sqrt{10}}$$

$$5.45 \pm 5.0715$$

$$(5.45 - 5.0715\,,\; 5.45 + 5.0715)$$

$$(0.38\;,\; 10.52)$$

$$0.38 < \mu_D < 10.52 \qquad >> \qquad \mu_D(0.38\,, 10.52\,)$$

Note :Since the confidence interval does not include zero, 0 $\notin$ **(0.38 , 10.52)** $(\mu_1 - \mu_2 \neq 0 \Leftrightarrow \mu_1 \neq \mu_2)$

Mean of weights **before** a diet program ≠ Mean of weights **after** a diet program mean

2. Does these data provide sufficient evidence to allow us to conclude that the diet program is effective? Use α=0.05 and assume that the populations are normal.

$\boldsymbol{\mu_1}$ =Mean of weights **before** a diet program mean
$\boldsymbol{\mu_2}$ =Mean of weights **after** the diet program mean

μ_D =Mean of X – Mean of Y ($\mu_D = \mu_1 - \mu_2$)

Hypotheses:

H_o: the diet program has no effect on weight
H_A: the diet program has an effect on weight

Equivalently,

$H_o: \mu_1 = \mu_2$ (no effect) VS $H_A: \mu_1 \neq \mu_2$ (There is effect)

$H_o: \mu_1 - \mu_2 = 0$ VS $H_A: \mu_1 - \mu_2 \neq 0$

$H_o: \mu_D = 0$ VS $H_A: \mu_D \neq 0$ (*where* : $\mu_D = \mu_1 - \mu_2$)

Hypotheses:

$H_0: \mu_1 = \mu_2$ **OR** $(H_0: \mu_1 - \mu_2 \neq 0)$ **OR** $H_0: \mu_D \neq 0$

$H_A: \mu_1 \neq \mu_2$ $\quad$ $(H_A: \mu_1 - \mu_2 \neq 0)$ $\quad$ $H_A: \mu_D \neq 0$

Test statistics (T.S.) $T = \frac{\bar{D} - \mu_0}{s_D/\sqrt{n}} = \frac{5.45 - 0}{7.09/\sqrt{10}} = 2.43$

Rejection Region of Ho (R.R.): (critical region : H_A Two sided)

$\alpha = 0.05$ ≫ **Critical Value :** $\pm t_{1-\frac{\alpha}{2}} = \pm t_{1-\frac{0.05}{2}} = \pm t_{0.975} = \pm 2.262$

df $= n - 1 = 10 - 1 = 9$

Decision :

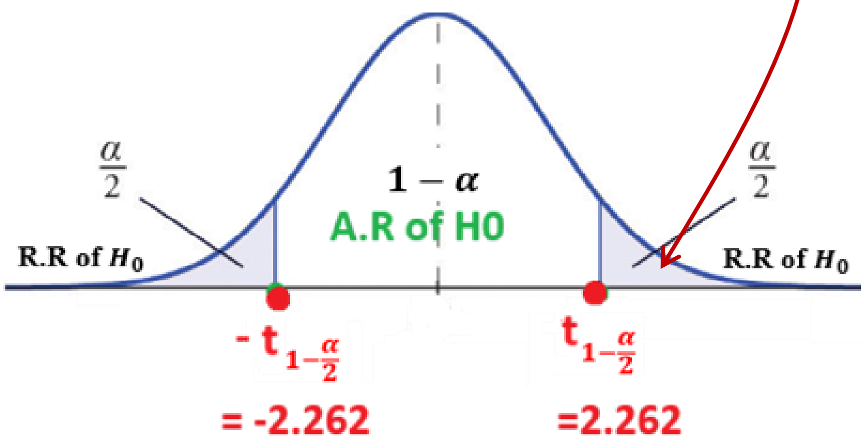
Reject $H_0: \mu_1 = \mu_2$ *if* : $T_{T.S} < -t_{1-\frac{\alpha}{2}}$ or $T_{T.S} > t_{1-\frac{\alpha}{2}}$

$2.43 \not< -2.262$ (First condition not satisfied, then try the second condition)

$2.43 > 2.262$ (Second condition satisfied)

Decision is Reject $H_0: \mu_1 = \mu_2$ (no effect) and Accept $H_A: \mu_1 \neq \mu_2$ (effect)
We conclude that there is effect of the diet program.

Another way to take decision is by graph :
Determine test statistics value on the graph



Note:

The **sample mean** of the weights **before** the program is ($\bar{X} = 84.2$)

The **sample mean** of the weights **after** the program is ($\bar{Y} = 78.75$)

i	X_i	Y_i
1	86.6	79.7
2	80.2	85.9
3	91.5	81.7
4	80.6	82.5
5	82.3	77.9
6	81.9	85.8
7	88.4	81.3
8	85.3	74.7
9	83.1	68.3
10	82.1	69.7
Sample Mean	$\bar{X}$ =84.2	$\bar{Y}$ =78.75

Since the diet program is effective and since

$$\bar{X} > \bar{Y}$$

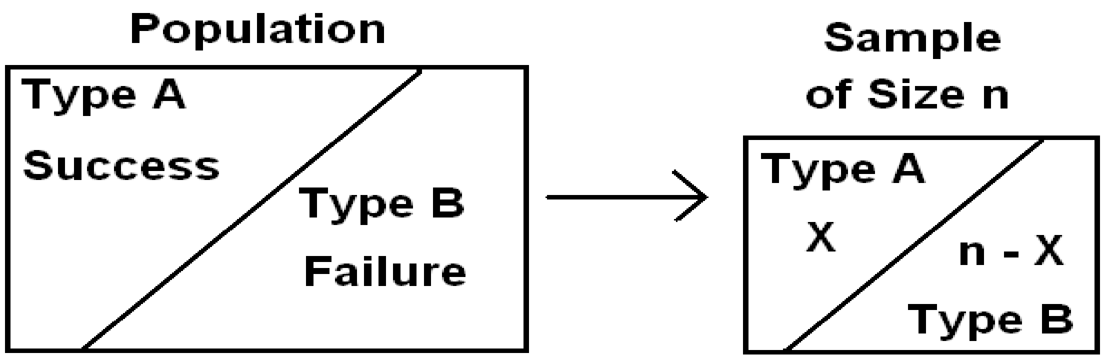
84.2 > 78.75

Weight before > weight after

we can conclude that the program is effective in reducing the weight (So, we have a good diet program).

7.5 Hypothesis Testing: A Single Population Proportion (p):

In this section, we are interested in testing some hypotheses about the population proportion (p).



Recall:

p = Population proportion of elements of Type A in the population.

$$P = \frac{\text{No. of elements of type A in the population}}{\text{Total number of elements in the population}} = \frac{A}{N}$$ (N= Population size)

$$\hat{P} = \frac{\text{No. of elements of type A in the sample}}{\text{Total number of elements in the sample}} = \frac{X}{n}$$ (n= Sample size)

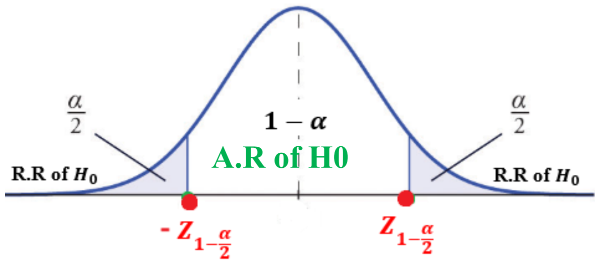
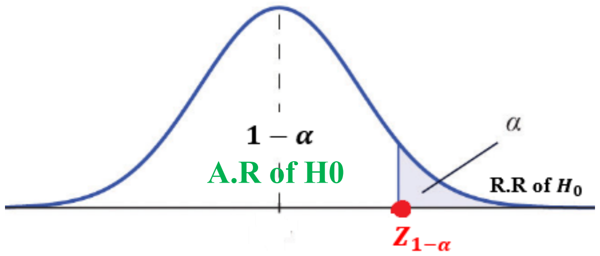
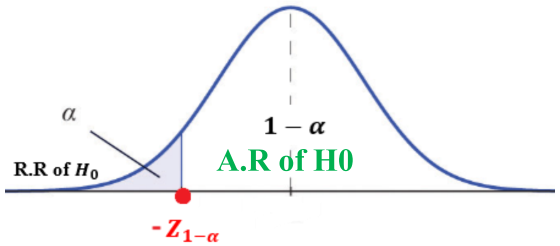
$\hat{P}$ is a "good" point estimate for p.

From Chapter 5:

For large n , ($n \geq 30$, $np > 5$, $nq > 5$)

$$Z = \frac{\hat{P} - P}{\sqrt{\frac{P(1-P)}{n}}} \sim N(0,1)$$

Test procedures:

Hypotheses	$H_0 : P = P_0$ $H_A : P \neq P_0$	$H_0 : P \leq P_0$ $H_A : P > P_0$	$H_0 : P \geq P_0$ $H_A : P < P_0$
Test Statistic (T.S.)	$\mathbf{Z} = \dfrac{\widehat{\mathrm{P}} - \mathrm{P_0}}{\sqrt{\dfrac{\mathrm{P_0 q_0}}{\mathrm{n}}}} \sim \mathbf{N(0, 1)}$ $\quad P_0$: be a given known value and $\mathbf{q_0} = 1 - P_0$ $\widehat{\mathrm{P}}$ = Sample propotion		
R.R. & A.R. of H_0			
Critical value(s)	$Z_{1-\frac{\alpha}{2}}$ and $-Z_{1-\frac{\alpha}{2}}$	$Z_{1-\alpha}$	$-Z_{1-\alpha}$
Decision:	We reject H_0 (and accept H_A) at the significance level α if:		
	$Z_{T.S} > Z_{1-\frac{\alpha}{2}}$ or $Z_{T.S} < -Z_{1-\frac{\alpha}{2}}$	$Z_{T.S} > Z_{1-\alpha}$	$Z_{T.S} < -Z_{1-\alpha}$
	T.S. ∈ R.R. **Two - Sided Test**	**T.S. ∈ R.R.** **One - Sided Test**	**T.S. ∈ R.R.** **One - Sided Test**

Example

A researcher was interested in the proportion of females in the population of all patients visiting a certain clinic. The researcher claims that 70% of all patients in this population are females. Would you agree with this claim if a random survey shows that 24 out of 45 patients are females? Use a 0.10 level of significance.

Solution:

p = Proportion of female in the population.

n=45 (large)

X= no. of female in the sample = 24

$\hat{P}$ = proportion of females in the sample

$$\hat{P} = \frac{X}{n} = \frac{24}{45} = 0.533$$

α=0.10 (level of significance)

Hypotheses:

$$H_0 : P = 0.7 \qquad (P_0 = 0.7)$$
$$H_A : P \neq 0.7$$

Test statistics (T.S.) $q_o = 1 - P_0 = 1 - 0.7 = 0.3$

$$Z = \frac{\hat{P} - P_0}{\sqrt{\frac{P_0 q_0}{n}}} = \frac{0.533 - 0.7}{\sqrt{\frac{0.7 \times 0.3}{45}}} = -2.44$$

Rejection Region of Ho (R.R.): (critical region) : $(-\infty, -1.645) \cup (1.645, \infty)$

$\alpha = 0.10 \gg$ **Critical Value :** $\pm Z_{1-\frac{\alpha}{2}} = \pm Z_{1-\frac{0.10}{2}} = \pm Z_{0.95} = \pm 1.645$

Decision :

Reject $H_0: P = 0.7$ *if* :

$$Z_{T.S} < -Z_{1-\frac{\alpha}{2}} \quad or \quad Z_{T.S} > Z_{1-\frac{\alpha}{2}}$$

$$-2.44 < -1.645$$

(First condition satisfied, not need to try the second condition)

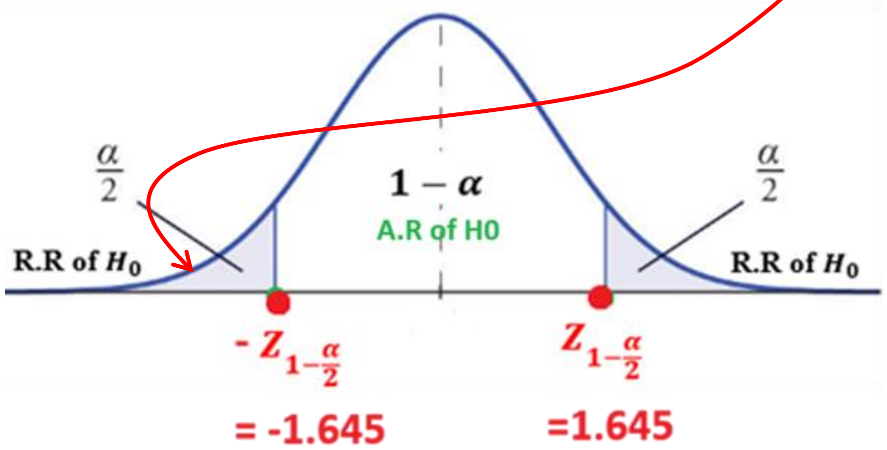
Decision is Reject $H_0: P = 0.7$ and Accept $H_A: P \neq 0.7$

Therefore, we do not agree with the claim stating that 70% of the patients in this population are females.

.

Another way to take decision is by graph :

Determine test statistics value on the graph



Example (Reading)

In a study on the fear of dental care in a certain city, a survey showed that 60 out of 200 adults said that they would hesitate to take a dental appointment due to fear. Test whether the proportion of adults in this city who hesitate to take dental appointment is less than 0.25. Use a level of significance of 0.025.

Solution:

p = Proportion of adults in the city who hesitate to take a dental appointment.

n= 200 (large)

X= no. of adults who hesitate in the sample = 60

$\hat{P}$ = proportion of adults who hesitate in the sample.

$$\hat{P} = \frac{X}{n} = \frac{60}{200} = 0.3$$

α=0.025 (level of significance)

Hypotheses:

$$H_0 : P \geq 0.25 \qquad (P_0 = 0.25)$$
$$H_A : P < 0.25$$

$$q_0 = 1 - P_0 = 1 - 0.25 = 0.75$$

Test statistics (T.S.)

$$Z = \frac{\hat{P} - P_0}{\sqrt{\frac{P_0 q_0}{n}}} = \frac{0.3 - 0.25}{\sqrt{\frac{0.25 \times 0.75}{200}}} = 1.633$$

Rejection Region of Ho (R.R.): (critical region): $(-\infty, -1.96)$

$\alpha = 0.025 \gg$ **Critical Value :** $-Z_{1-\alpha} = -Z_{1-0.025} = -Z_{0.975} = -1.96$

Decision :

Reject $H_0 : P = 0.7$ *if* :

$$Z_{T.S} < -Z_{1-\alpha}$$
$$1.633 \not< -1.96$$

condition not satisfied

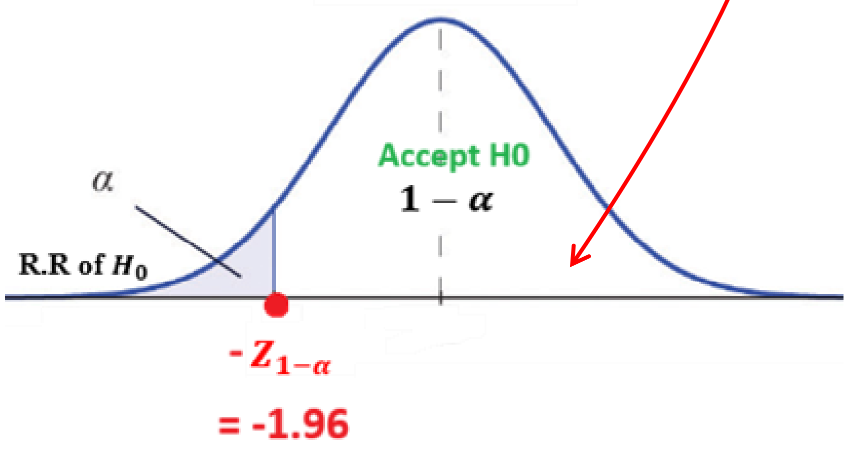
Decision is Accept $H_0 : P \geq 0.25$ and Reject $H_A : P < 0.25$

.

.

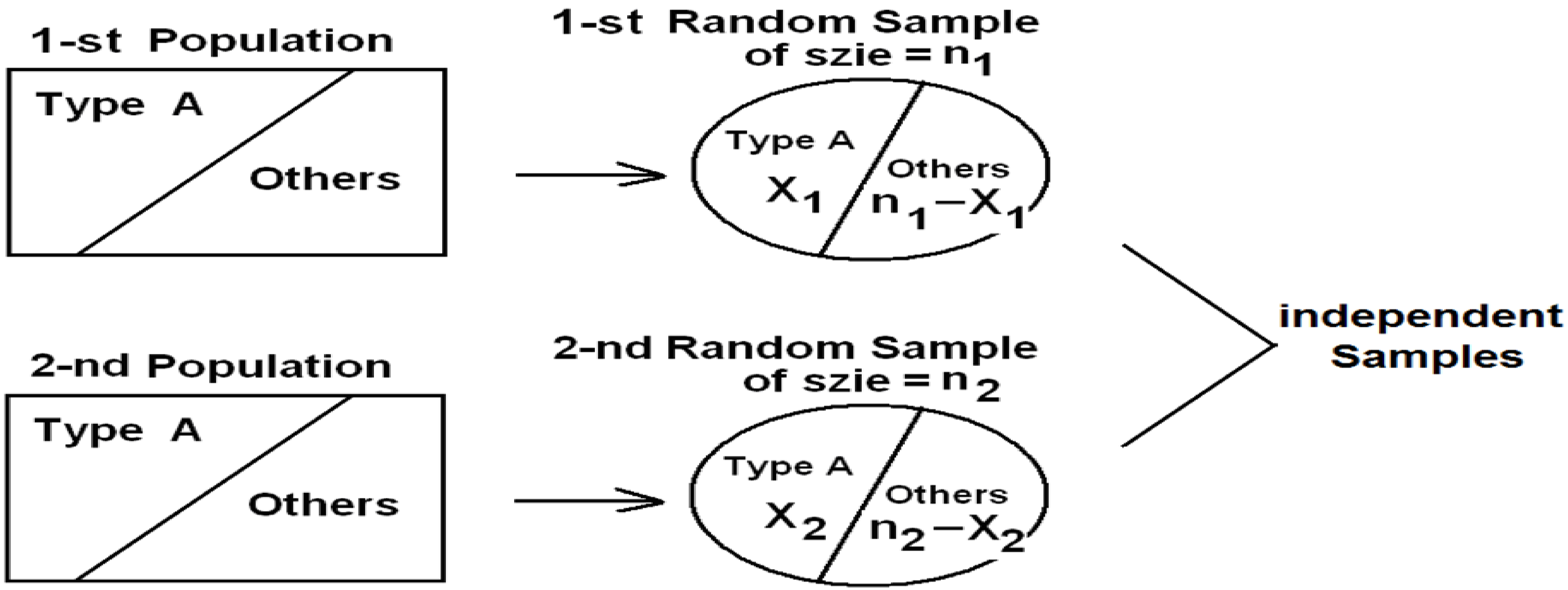
Another way to take decision is by graph :

Determine test statistics value on the graph



7.6 Hypothesis Testing:
The Difference Between Two Population Proportions ($P_1 - P_2$):

In this section, we are interested in testing some hypotheses about the difference between two population proportions ($P_1 - P_2$)



Suppose that we have two populations:

- P_1 = population proportion of the 1-st population.
- P_2 = population proportion of the 2-nd population.
- We are interested in comparing P_1 and P_2, or equivalently, making inferences about $P_1 - P_2$.
- We independently select a random sample of size n_1 from the 1-st population and another random sample of size n_2 from the 2-nd population:
- Let X_1 = no. of elements of type A in the 1-st sample.
- Let X_2= no. of elements of type A in the 2-nd sample.
- $\hat{P}_1 = \frac{X_1}{n_1}$ the sample proportion of the 1-st sample.
- $\hat{P}_2 = \frac{X_2}{n_2}$ the sample proportion of the 1-st sample.

Hypotheses:

We choose one of the following situations:

(i) $H_0: P_1 = P_2$ against $H_A: P_1 \neq P_2$

(ii) $H_0: P_1 \geq P_2$ against $H_A: P_1 < P_2$

(iii) $H_0: P_1 \leq P_2$ against $H_A: P_1 > P_2$

or equivalently,

(i) $H_0: P_1 - P_2 = 0$ against $H_A: P_1 - P_2 \neq 0$

(ii) $H_0: P_1 - P_2 \geq 0$ against $H_A: P_1 - P_2 < 0$

(iii) $H_0: P_1 - P_2 \leq 0$ against $H_A: P_1 - P_2 > 0$

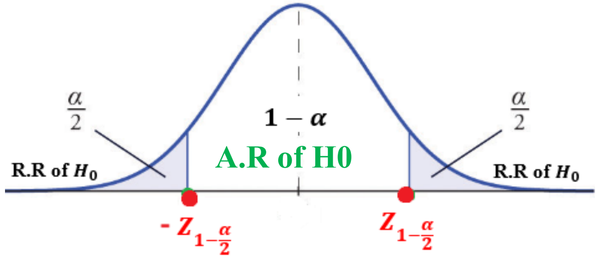
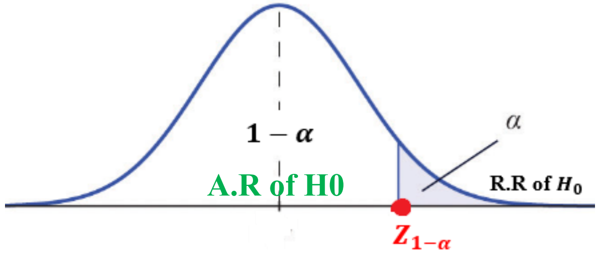
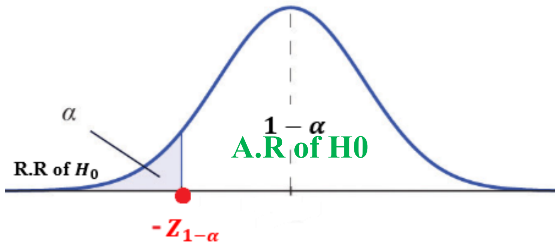
Note, under the assumption of the equality of the two population proportions ($\mathbf{H_0: P_1 = P_2 = P}$), the pooled estimate of the common proportion p (**pooled proportion**)is:

$$\bar{P} = \frac{X_1+X_2}{n_1+n_2} \qquad (\bar{q} = 1 - \bar{P})$$

The test statistic (T.S.) is

$$Z = \frac{\hat{P}_1 - \hat{P}_2}{\sqrt{\frac{\bar{P}\,\bar{q}}{n_1} + \frac{\bar{P}\,\bar{q}}{n_2}}} \sim N(0,1)$$

Test procedures:

Hypotheses	$H_0 : P_1 - P_2 = 0$ $H_A : P_1 - P_2 \neq 0$	$H_0 : P_1 - P_2 \leq 0$ $H_A : P_1 - P_2 > 0$	$H_0 : P_1 - P_2 \geq 0$ $H_A : P_1 - P_2 < 0$
Test Statistic (T.S.)	$Z = \frac{\widehat{P}_1 - \widehat{P}_2}{\sqrt{\frac{\bar{P}\bar{q}}{n_1} + \frac{\bar{P}\bar{q}}{n_2}}} \sim N(0,1)$	Pooled Proportion $\bar{P} = \frac{X_1 + X_2}{n_1 + n_2}$ and $\bar{q} = 1 - \bar{P}$	
R.R. & A.R. of H_0			
Critical value(s)	$Z_{1-\frac{\alpha}{2}}$ and $-Z_{1-\frac{\alpha}{2}}$	$Z_{1-\alpha}$	$-Z_{1-\alpha}$
Decision:	We reject H_0 (and accept H_A) at the significance level α if:		
	$Z_{T.S} > Z_{1-\frac{\alpha}{2}}$ or $Z_{T.S} < -Z_{1-\frac{\alpha}{2}}$	$Z_{T.S} > Z_{1-\alpha}$	$Z_{T.S} < -Z_{1-\alpha}$
	T.S. ∈ R.R. **Two - Sided Test**	**T.S. ∈ R.R.** **One - Sided Test**	**T.S. ∈ R.R.** **One - Sided Test**

Example:

In a study about the obesity (overweight), a researcher was interested in comparing the proportion of obesity between males and females. The researcher has obtained a random sample of 150 males and another independent random sample of 200 females. The following results were obtained from this study.

	n	Number of obese people (X)
Males	150	21
Females	200	48

Can we conclude from these data that there is a difference between the proportion of obese males and proportion of obese females? Use $\alpha = 0.05$, assume that the two population proportions are equal .

Solution :

P_1 = population proportion of obese males.

P_2 = population proportion of obese females .

$\hat{P}_1$ = sample proportion of obese males

$\hat{P}_2$ = sample proportion of obese females

$\alpha = 0.05$ (level of significance)

	Sample size (n)	X	Sample proportion $\hat{P}$
Male (1)	150 large	21	$\hat{P}_1 = \frac{X_1}{n_1} = \frac{21}{150} = 0.14$
Female (2)	200 large	48	$\hat{P}_2 = \frac{X_2}{n_2} = \frac{48}{200} = 0.24$

The pooled estimate of the common proportion *P* is:

$$\bar{P} = \frac{X_1+X_2}{n_1+n_2} = \frac{21+48}{150+200} = 0.197 \qquad (\bar{q} = 1 - \bar{P} = 1 - 0.197 = 0.803)$$

Hypotheses:

$H_0: P_1 = P_2$ against $H_A: P_1 \neq P_2$

OR

$H_0: P_1 - P_2 = 0$ against $H_A: P_1 - P_2 \neq 0$

Hypotheses:

$H_0 : P_1 = P_2$ **OR** $(H_0 : P_1 - P_2 = 0)$

$H_A : P_1 \neq P_2$ $(H_A : P_1 - P_2 \neq 0)$

Test statistics (T.S.)

$$Z = \frac{\hat{P}_1 - \hat{P}_2}{\sqrt{\frac{\bar{P}\bar{q}}{n_1} + \frac{\bar{P}\bar{q}}{n_2}}} = \frac{0.14 - 0.24}{\sqrt{\frac{0.196 \times 0.803}{150} + \frac{0.196 \times 0.803}{200}}} = -2.328$$

Rejection Region of Ho (R.R.): (critical region) $(-\infty, -1.96) \cup (1.96, \infty)$

$\alpha = 0.05$ ≫ **Critical Value :** $\pm Z_{1-\frac{\alpha}{2}} = \pm Z_{1-\frac{0.05}{2}} = \pm Z_{(0.975)} = \pm 1.96$

Decision :

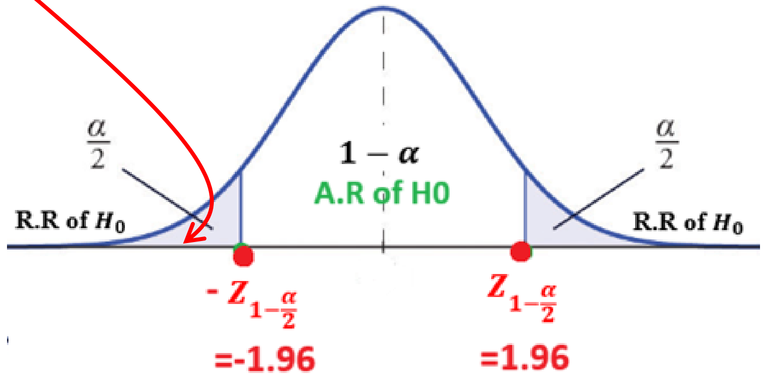
Reject $H_0 : P_1 = P_2$ *if* :

$$Z_{T.S} < -Z_{1-\frac{\alpha}{2}} \quad or \quad Z_{T.S} > Z_{1-\frac{\alpha}{2}}$$

$-2.328 < -1.96$ (First condition satisfied, then do not try the second condition)

Another way to take decision is by graph :

Determine test statistics value on the graph



Decision is Reject $H_0 : P_1 = P_2$ and Accept $H_A : P_1 \neq P_2$

Therefore, we conclude that there is a difference between the proportion of obese males and the proportion of obese females

Since $Z = -2.328 \in R.R.$, we reject $H_0: P_1 = P_0$ and accept $H_A: P_1 \neq P_2$ at $\alpha = 0.05$. Therefore,

we conclude that there **is a difference between the proportion of obese males and the proportion of obese females.** Additionally, since,

$$\widehat{P}_1 = 0.14 < \widehat{P}_2 = 0.24$$

Sample proportion of male < Sample proportion of female

We may conclude that the **proportion of obesity for females is <u>larger than</u> that for males**.

TABLE OF CONTENTS

CHAPTER 2: Basic Summary Statistics	
Measures of Central Tendency	
Mean	
-Population Mean	
-Sample Mean	
-Advantages and Disadvantages of the Mean	
Median	
-Advantages and Disadvantages of the Median	
Mode	
-Advantages and Disadvantages of the Mode	
Measures of Dispersion (Variation)	
Range	
Variance	
- Deviations of Sample Values from the Sample Mean	
-Population Variance	
-Sample Variance	
- Calculating Formula for the Sample Variance	
Standard Deviation	
Coefficient of Variation	

CHPTER 3: Basic Probability Concepts	
General Definitions and Concepts	
Probability	
An Experiment	
Sample Space	
Events	
Equally Likely Outcomes	
Probability of an Event	
Some Operations on Events	
Union of Two events	
Intersection of Two Events	
Complement of an Event	
General Probability Rules	
Applications	
Conditional Probability	
Independent Events	
Bayes' Theorem	

CHAPTER 4: Probability Distributions	
Introduction	
Probability Distributions of Discrete R.V.'s	
Graphical Presentation	
Population Mean of a Discrete Random Variable	
Cumulative Distributions	
Binomial Distribution	
Poisson Distribution	
Probability Distributions of Continuous R.V.	
Normal Distribution	
Standard Normal Distribution	
Calculating Probabilities of Standard Normal Distribution	
Calculating Probabilities of Normal Distribution	
t-distribution	

CHAPTER 5: Sampling Distribution	
Results about Sampling Distribution of the Sample Mean	
-Sampling Distribution of the single Sample Mean	
-Sampling Distribution of the Difference between Two Means	
Results about Sampling Distribution of the Sample Proportion	
-Sampling Distribution of Proportion	
-Sampling Distribution of the Difference between Two Proportions	

CHAPTER 6 & 7: Statistical Inferences (Estimation and Hypotheses Testing)	
CHAPTER 6 : Estimation	
Estimation of the Population Mean:	
- Point Estimation of the population mean	
- Interval Estimation of the population mean	
Estimation for the Population Proportion :	
- Point Estimate for the Population Proportion	
- Interval Estimation for the Population Proportion	
CHAPTER 7: Tests of Hypotheses	
Tests of Hypotheses	
Single Sample: Tests Concerning a Single Mean	
Two Samples: Tests on Two Means	
Two-sample: Paired t-test	
Single Sample: Tests on a Single Proportion:	
Two Samples: Tests on Two Proportions	