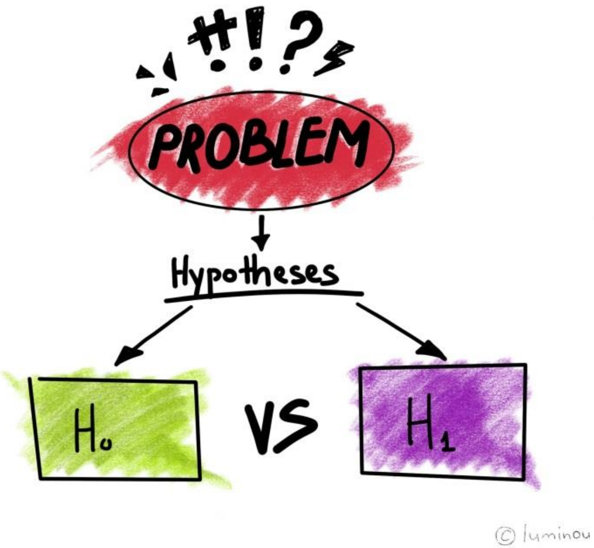
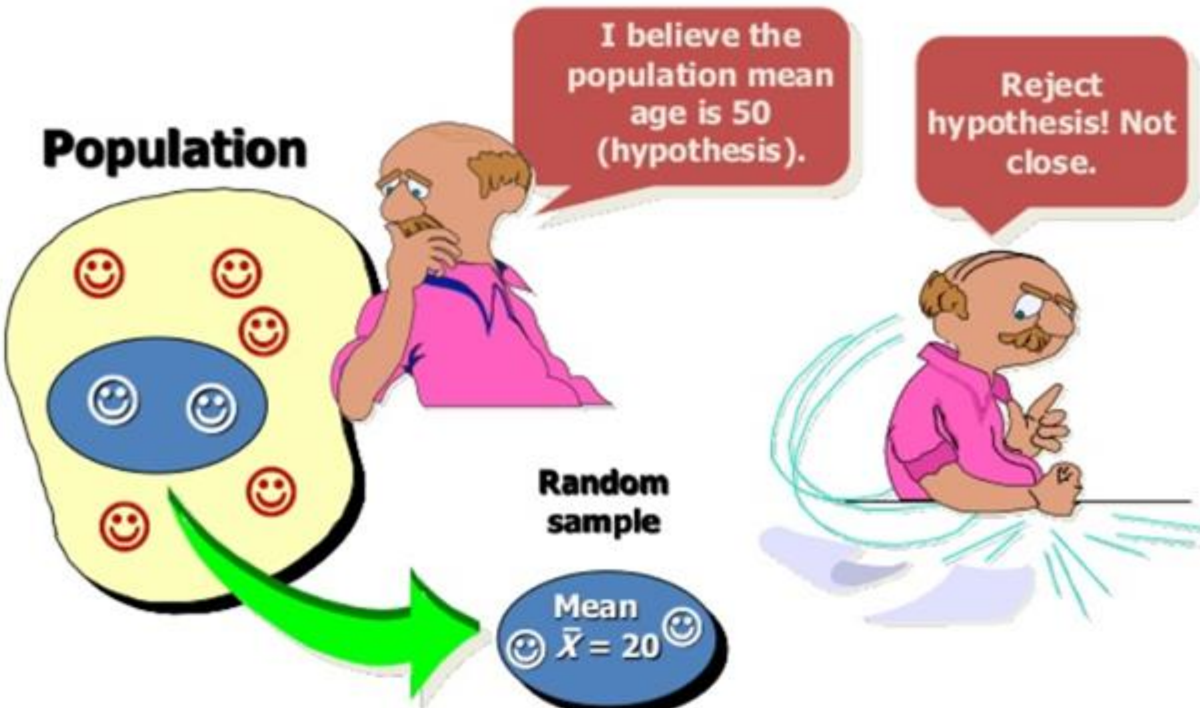
Hypothesis Testing





Hypothesis Testing

- The general goal of a hypothesis test is to rule out chance (sampling error) as a plausible explanation for the results from a research study.
- Hypothesis testing is a technique to help determining whether a hypothesis is true (e.g. treatment, procedure has an effect in a population), or simply if a relationship exists between two or more variables.

Research Hypothesis **vs** Null Hypothesis

Research hypothesis (H_1) is what the research believes to be a true reflection on the general population. In another word, a true explanation for a phenomena in the population. The researcher wants to prove that his sample statistics is different than the population parameters. Research hypothesis is also called alternative hypothesis.

Null hypothesis (H_0) is the opposite of H_1. The H_0 assumes no difference of test statistics and the population parameter. This means that the researcher hypothesis about a certain phenomena is not correct, and that there is no real difference between sample and population for a certain feature or difference is due another reason (that is not tested).

Null & Alternative Hypothesis

Null Hypothesis or the Hypothesis to be Tested (H_0)

- A claim that there is NO difference between the population parameter like mean (μ) and the hypothesized value. In the testing process the null hypothesis either is rejected or not rejected (we do not say accepted).

Alternative Hypothesis or the Research Hypothesis (H_A or H_I)

- A statement of what we will believe is true if our sample data cause us to reject the null hypothesis. Usually the alternative hypothesis is the research hypothesis (the conclusion that the researcher is seeking to reach).

Null Hypothesis as an Assumption to be Challenged

- We might begin with a belief or assumption that a statement about the value of a population parameter is true.
- We then using a hypothesis test to challenge the assumption and determine if there is statistical evidence to conclude that the assumption is incorrect.
- In these situations, it is helpful to develop the null hypothesis first.

Example

- A new drug is developed with the goal of lowering blood pressure more than the existing drug.
- **Null hypothesis**

➢ The new drug does not lower BP more than the existing drug.

- **Alternative hypothesis**

➢ The new drug lowers blood pressure more than the existing drug.

Null & Alternative Hypothesis

Gabapentin has no pharmacological effect

- The mean for population A is 20 (H_0: $\mu = 20$)
- The mean for population A is less than 20 (H_0: $\mu \leq 20$)
- The mean for population A is larger than 20 (H_0: $\mu \geq 20$)

Gabapentin has a pharmacological effect

- The mean for population A is not 20 (H_1: $\mu \neq 20$)
- The mean for population A is larger than 20 (H_1: $\mu > 20$)
- The mean for population A is less than 20 (H_1: $\mu < 20$)

Accepting or rejecting a hypothesis is not a proof of the hypothesis!
Null hypothesis can be true or false, we only can reject it or not to reject it

Null Hypothesis vs Alternative Hypothesis

Null hypothesis H_0	Alternative Hypothesis H_1
There is no relationship or difference	There is a relationship or difference
Refers to the population	Refers to the examined sample
Research aims to reject the null	Research aims to accept the alternative
Represent an original assumption	Prove statistically a systemic difference or relationship
Assumes a difference is due to chance	Assumes that difference is less likely to be due to chance.

Types of Statistical Errors in Hypotheses Testing

Because hypothesis tests are based on sample data, we must allow for the possibility errors.

Type I error

- A type I error is rejecting H_0 when it is true.
- The probability of making a Type I error when the null hypothesis is true as an equality is called the level of significance (α).

Type II error

- A type II error is accepting H_0 when it is false.
- It is difficult to control for the probability of making a type II error (β).
- Statisticians avoid the risk of making a type II error by using “do not reject H_0” and not “accept H_0”

Errors in hypothesis testing

TRUTH

DECISION

	Ho is true (A=B)	Ho is false (A≠B)
Reject Ho (A≠B)	Type I error "giving a treatment that does not work"	Correct
Do not reject Ho (A=B)	Correct	Type II error "not giving a treatment that works"

How to establish a good hypothesis?

- Clear and declarative statement. Not a question.
- Show a relationship between variables
- Reflect a body of literature or a theory
- Be direct, explicit, and to the point.
- Be testable and measurable.

Critical Value Approach to Hypotheses Testing

- **Step 1**: develop the null and alternative hypotheses.
- **Step 2**: specify the level of significance α. Common value of α are: α =1% or 0.01, 5% or 0.05, 10% or 0.1
- **Step 3**: collect the sample data and compute the value of the test statistics Z or t.
- **Step 4**: use the level of significance (α) to determine the critical value (tabulated value).
- **Step 5**: use the suitable rejection rule.
- **Step 6**: state the appropriate conclusions.

Step 1

- Set your hypothesis

  Set both H_0 & H_1.

Example

It is believed that a candy machine makes chocolate bars that are on average 5g. A worker claims that the machine after maintenance no longer makes 5g bars. Write Ho and Ha.

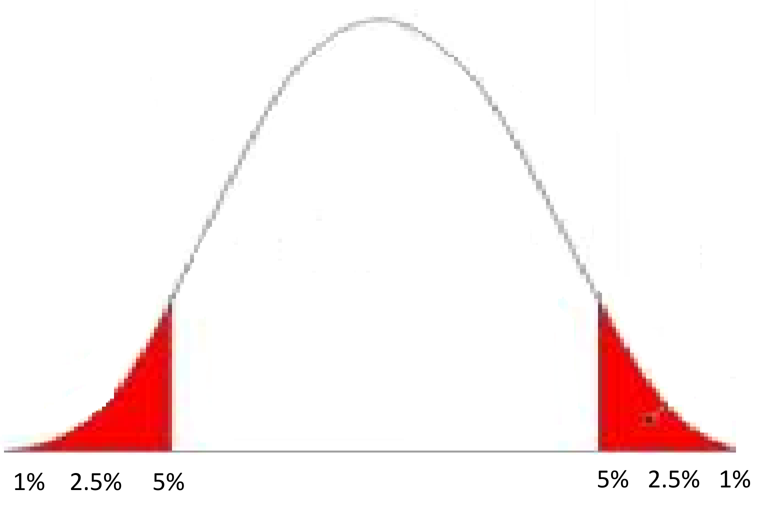
Answer:
H_0: $\mu = 5$
H_1: $\mu \neq 5$

A company has stated that their straw machine makes straws that are 4 mm diameter. A worker believes the machine no longer makes straws of this size and samples 100 straws to perform a hypothesis test with 99% confidence.

Answer:
H_0: $\mu = 4$
H_1: $\mu \neq 4$

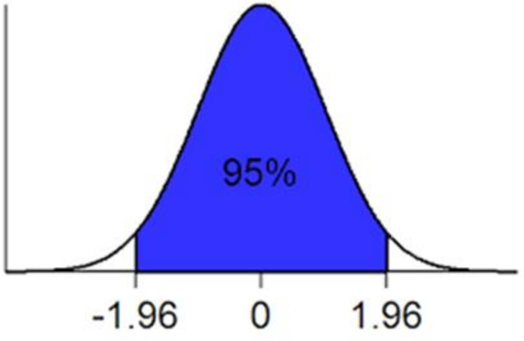
Step 2

- Set level of significance associated with the hypothesis.



Step 3

- Set the critical value needed to reject the null hypothesis (from Tables).
- You might need the table of Z-test, t-test, F-test.
- Based on the chosen table, look for the cut off value based on the level of significance you determined in step 2.

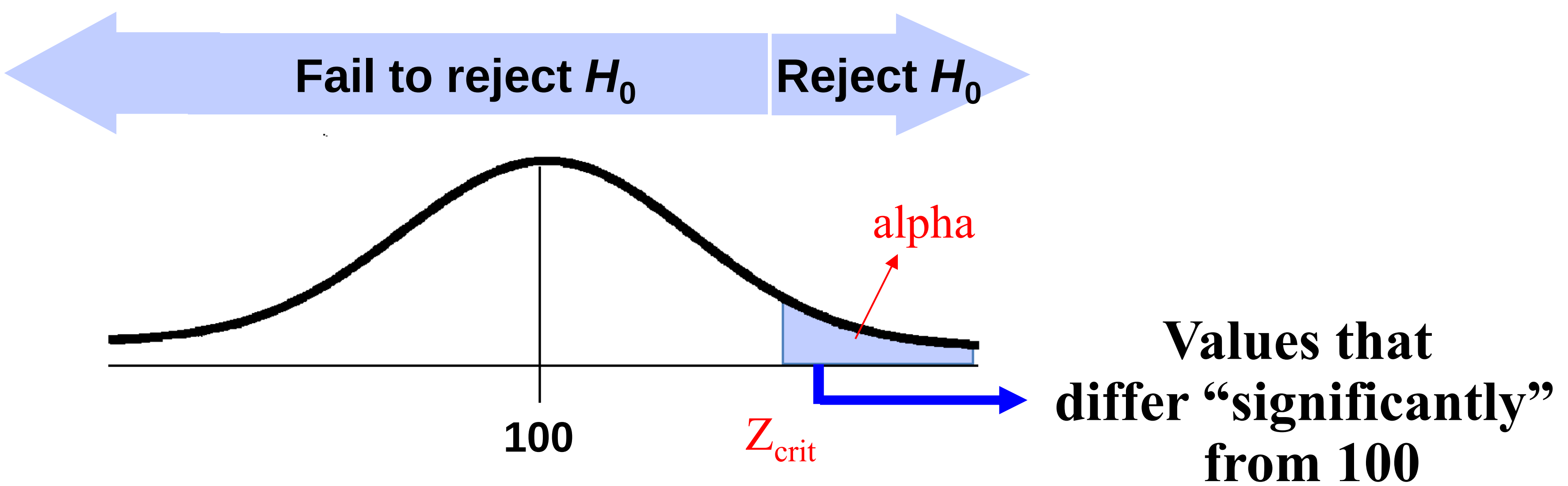


Right-tailed tests

H_0: $\mu = 100$

H_1: $\mu > 100$

Points Right

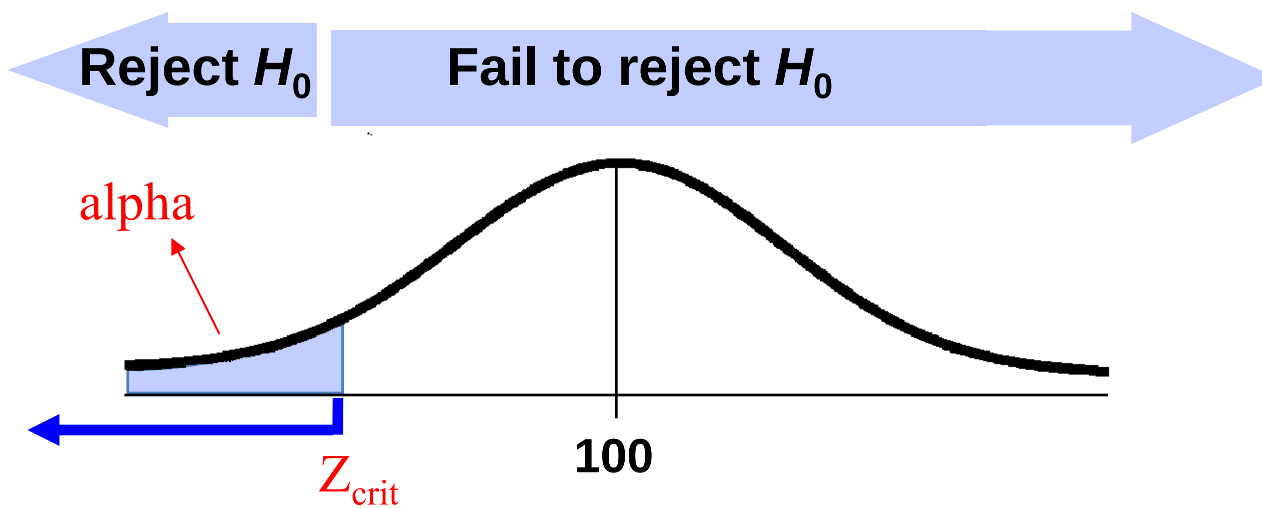


Left-tailed tests

$H_0: \mu = 100$

$H_1: \mu < 100$

Points Left



Values that differ "significantly" from 100

Two-tailed hypothesis testing

- H_A is that μ is *either* greater or less than μ_{H0}

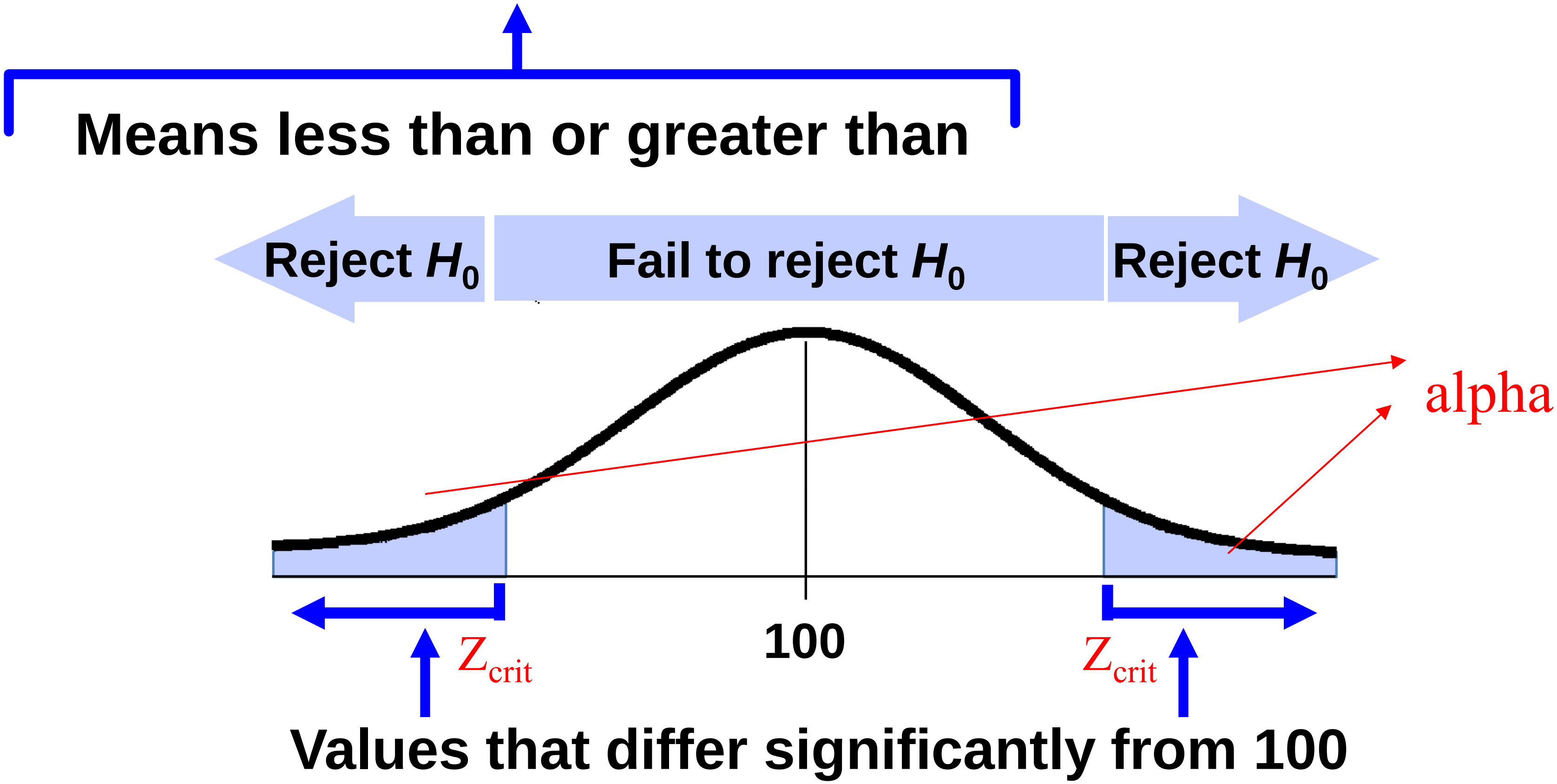
  H_A: $\mu \neq \mu_{H0}$

- α is divided equally between the two tails of the critical region

Two-tailed hypothesis testing

$H_0: \mu = 100$

$H_1: \mu \neq 100$



One tale critical values

alpha .05, $Z_{crit}=1.64$;

alpha .01, $Z_{crit}=2.33$

Two tale critical values

alpha .05, $Z_{crit}=1.96$;

alpha .01, $Z_{crit}=2.58$

Normal Distribution
(Population Standard Deviation σ is known or Standard Deviation σ is known but n is Large)

Normal Distribution

σ is known

[n < 30 (small) or n ≥ 30 (large)]

$$z = \frac{\bar{X} - \mu_0}{\sigma / \sqrt{n}}$$

Normal (σ is unknown) or unknown or non-normal distribution

but n ≥ 30 (large)

$$z = \frac{\bar{X} - \mu_0}{S / \sqrt{n}}$$

Normal Distribution
(Population Standard Deviation σ is Unknown and n is small)

Normal Distribution

σ is unknown

N < 30 (small)

$$t = \frac{\bar{X} - \mu_0}{S/\sqrt{n}}$$

Summary of Forms for Null and Alternative Hypotheses about a Population Mean

The equality part of the hypotheses always appears in the null hypothesis H_0. In general, a hypothesis test about the value of a population mean μ must take one of the following three forms (where μ_0 is the hypothesized value of the population mean).

a) One – tailed test

i. $H_0: \mu = \mu_0$ vs $H_1: \mu < \mu_0$ (less than or smaller than)

ii. Upper (right)-tailed test

$H_0: \mu = \mu_0$ vs $H_1: \mu > \mu_0$ (more than or greater than)

b) Two – tailed test

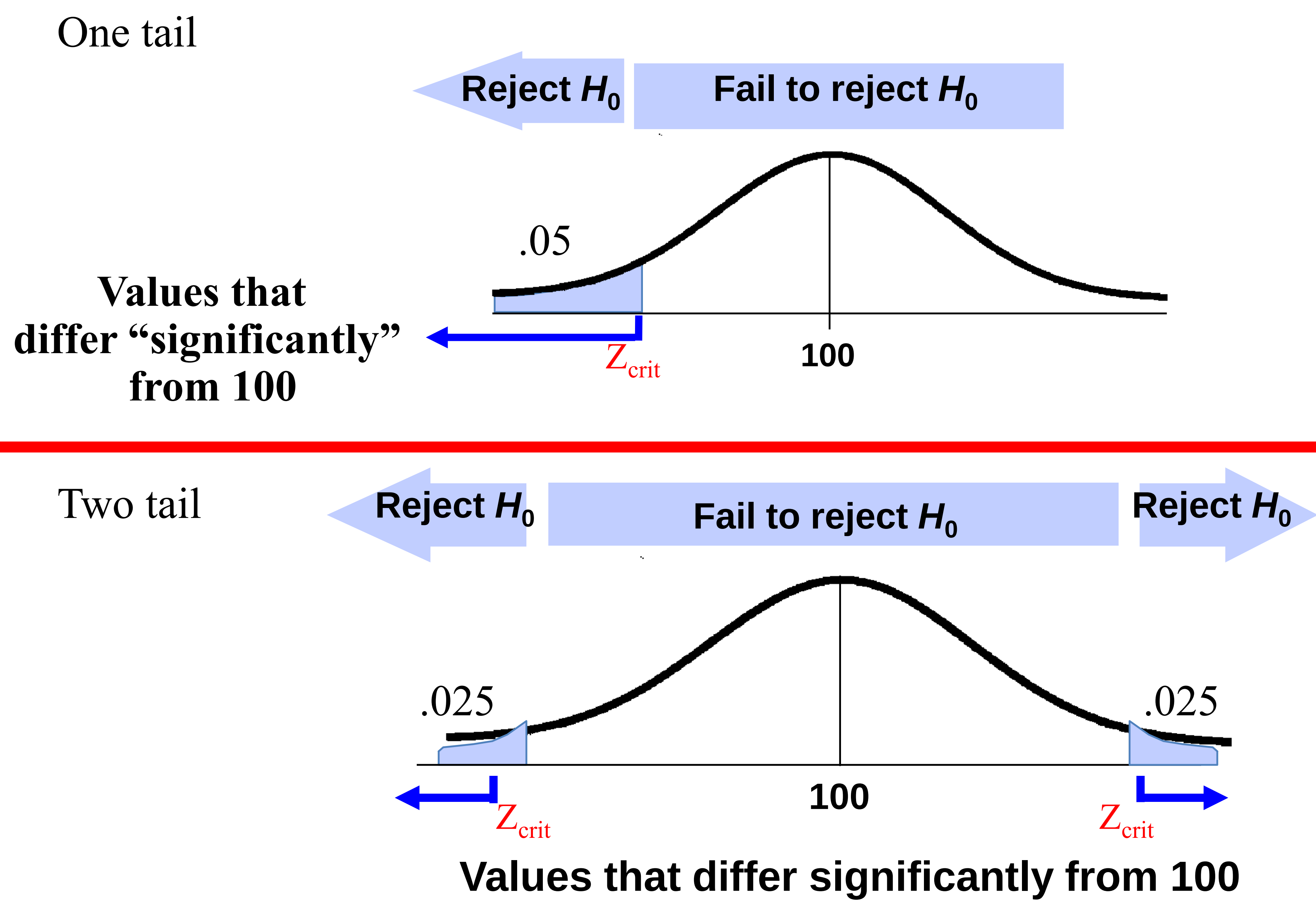
$H_0: \mu = \mu_0$ vs $H_1: \mu \neq \mu_0$ (does not equal or different from)

One- vs. Two-Tailed Tests

- In theory, you should use one-tailed when
  1. Change in opposite direction would be meaningless
  2. Change in opposite direction would be uninteresting
  3. No opposing theory predicts change in opposite direction

- By convention/default in the social sciences, two-tailed is standard

- Why? Because it is a more strict criterion. A more conservative test.


One tail
Reject H_0
Fail to reject H_0
.05
Values that differ "significantly" from 100
Z_{crit}
100
Two tail
Reject H_0
Fail to reject H_0
Reject H_0
.025
.025
100
Z_{crit}
Z_{crit}
Values that differ significantly from 100

Step 5

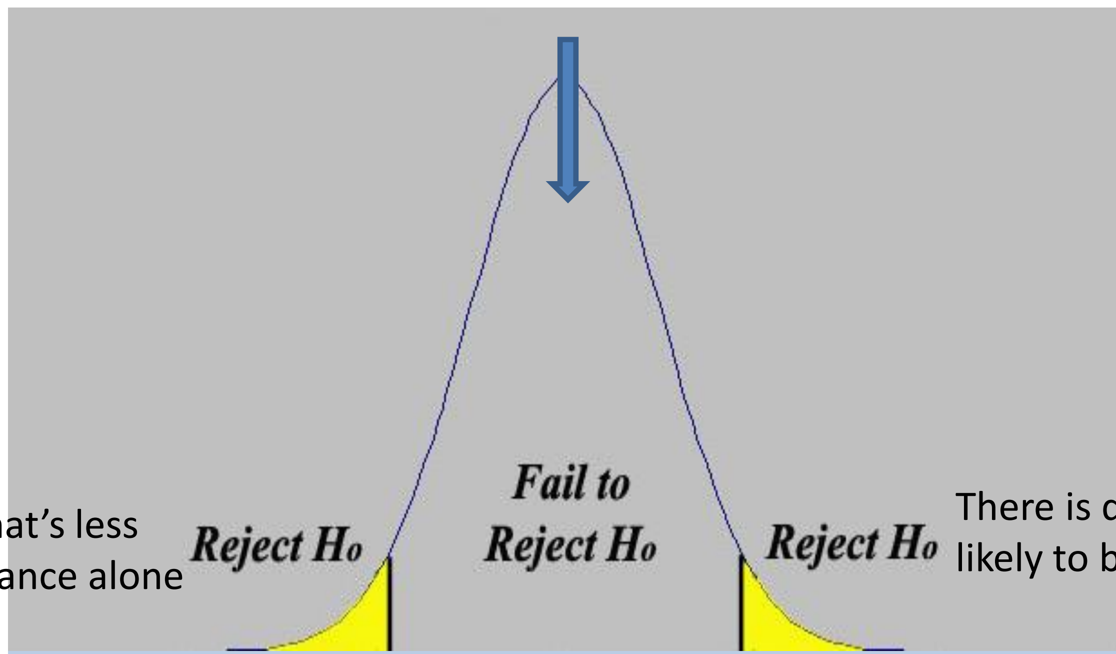
- Compare the test statistics value with the critical value of rejection.

  Does Z-value = or $\neq$ value from the table

Step 6

- Decide whether to reject the null hypothesis and confidence statement.

There is difference that's less likely to be due to chance alone

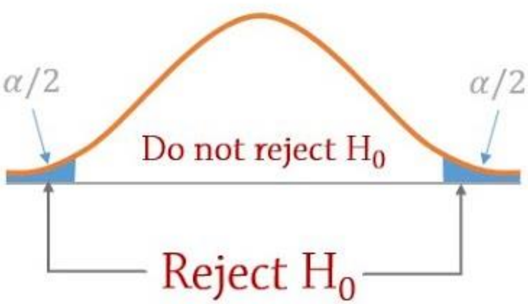


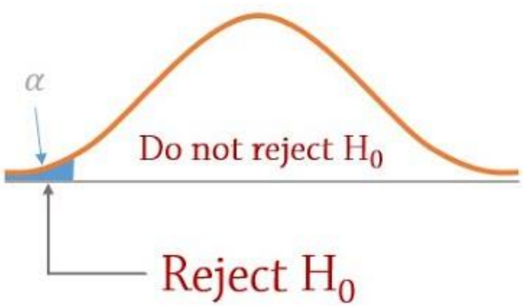
There is difference that's less likely to be due to chance alone

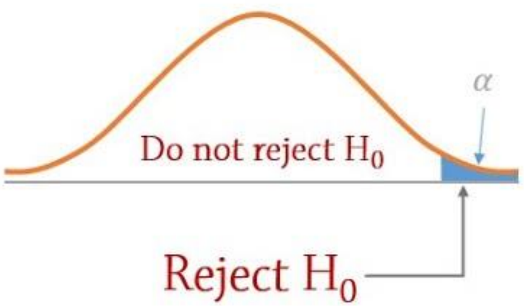
Hypothesis Testing

One-tailed

Two-tailed	Left-tailed	Right-tailed
$H_0: \mu = 23$	$H_0: \mu \geq 23$	$H_0: \mu \leq 23$
$H_1: \mu \neq 23$	$H_1: \mu < 23$	$H_1: \mu > 23$







Rejection Rule

Lower (left)-tailed test

Hypotheses

H_0: $\mu = \mu_0$

vs H_1: $\mu < \mu_0$ (less than or smaller than)

<u>Rejection Rule</u>

Reject H_0 at a level of significance α if:

i. $Z < -Z_{1-\alpha}$ (In case of using Z-distribution)

ii. $t < -t_{(\alpha,\ n-1)}$ (In case of using t-distribution)

Rejection Rule

Upper (right)-tailed test

Hypotheses

H_0: $\mu = \mu_0$

vs H_1: $\mu > \mu_0$ (more than or greater than)

Rejection Rule

Reject H_0 at a level of significance α if:

i. $Z > Z_{1-\alpha}$ (In case of using Z-distribution)

ii. $t > t_{(\alpha, n-1)}$ (In case of using t-distribution)

Rejection Rule

Two-tailed test

Hypotheses

H_0: $\mu = \mu_0$

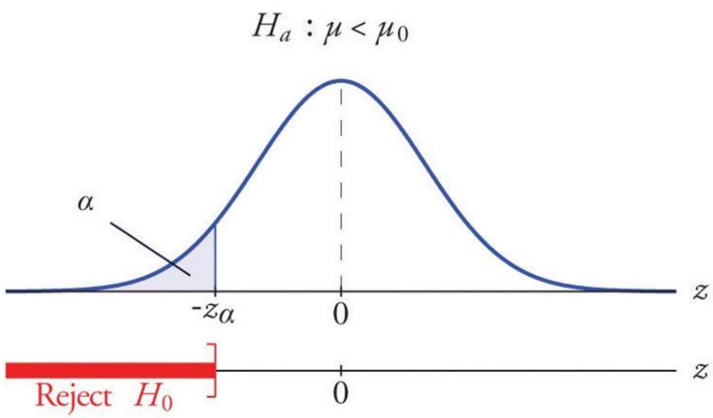
vs H_1: $\mu \neq \mu_0$ (does not equal or different from)

<u>Rejection Rule</u>

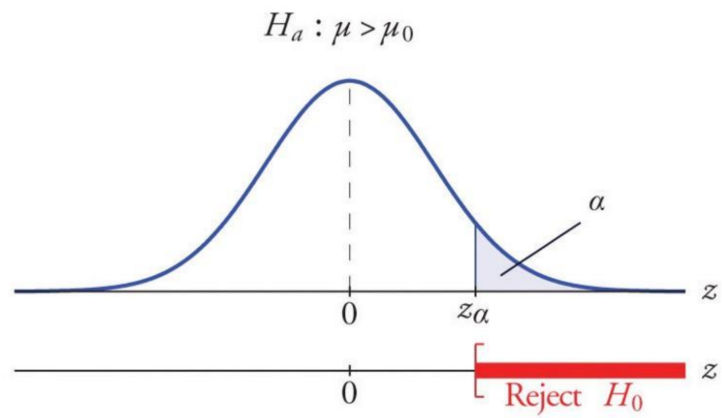
Reject H_0 at a level of significance α if:

i. $|Z| > Z_{1-\frac{\alpha}{2}}$ (In case of using Z-distribution)

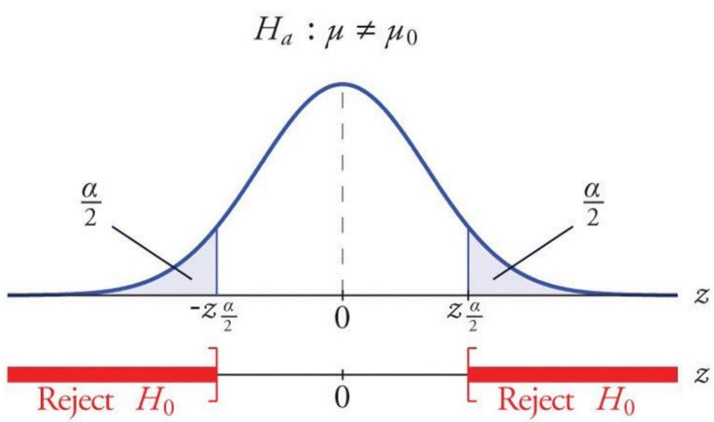
ii. $|t| > t_{(\frac{\alpha}{2},\, n-1)}$ (In case of using t-distribution)



$H_a : \mu < \mu_0$
α
$-z_\alpha$
0
z
Reject H_0
0
z

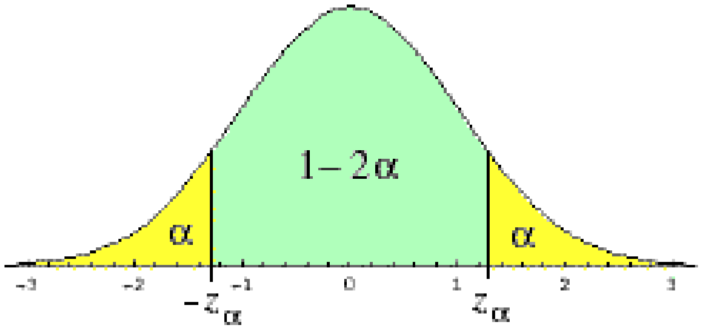


$H_a : \mu > \mu_0$
α
0
z_α
z
0
Reject H_0
z



$H_a : \mu \neq \mu_0$
$\frac{\alpha}{2}$
$\frac{\alpha}{2}$
$-z_{\frac{\alpha}{2}}$
0
$z_{\frac{\alpha}{2}}$
z
Reject H_0
0
Reject H_0
z

Common Critical Values



α = tail area	central area = 1 − 2α	z_α
0.10	0.80	$z_{.10} = \mathbf{1.28}$
0.05	0.90	$z_{.05} = \mathbf{1.645}$
0.025	0.95	$z_{.025} = \mathbf{1.96}$
0.01	0.98	$z_{.01} = \mathbf{2.33}$
0.005	0.99	$z_{.005} = \mathbf{2.58}$

α	1 − α	$z_{1-\alpha}$
0.10	0.90	$z_{0.90} = 1.28$
0.05	0.95	$z_{0.95} = 1.645$
0.01	0.99	$z_{0.99} = 2.33$

Example Weight

Salem believes that his “true weight” is 72kg with a standard deviation of 3kg.

Salem weighs himself once a week for four weeks. The average of these four measurements is 75.4kg.

Are the data consistent with Salem’s belief?

Example Weight

1. H_0: μ =72 H_1: μ >72   This is a one tail test

2. a=0.05

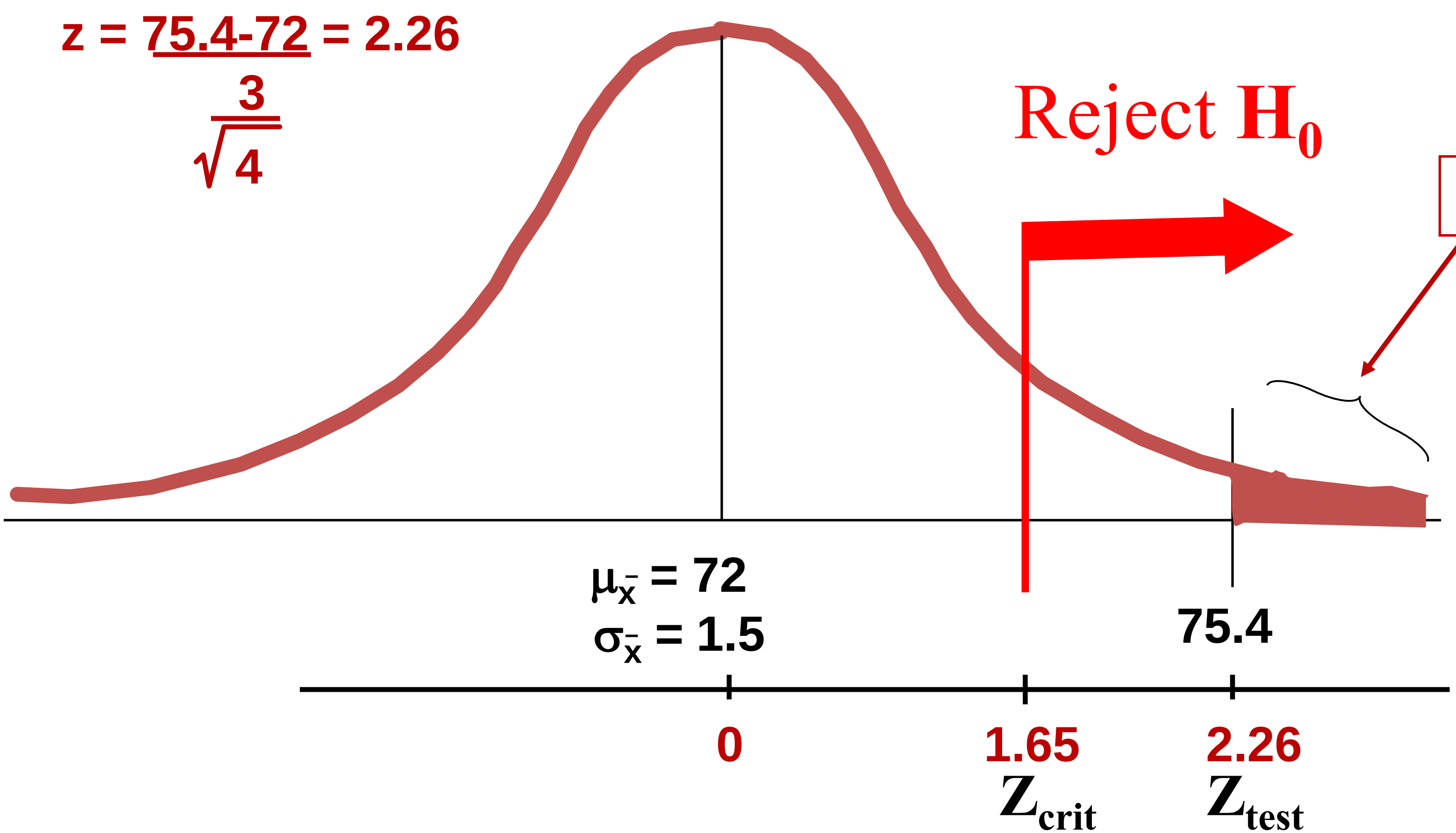
3. μ > 72 (one tail test)

4. Z_{crit} = ZReject H_0 if z $\geq$ 1.645

5. $$Z = \frac{\overline{X} - \mu_0}{s/\sqrt{n}} = \frac{75.4 - 72}{3/\sqrt{4}} = 2.26$$ $$P(Z > 2.26) = .012$$

6. Since 2.26 >1.645, we <u>Reject</u> H_0. There is a statistically significant evidence at a=0.05 to show that the mean weight measured is higher than his original belief about his weight. The chance that the measured weight and initial (belief) weight means are different due to chance only is less than 5%.

Example Weight illustrated



Example

Researcher are interested in the mean age of a certain population. They are wondering if the **mean age** is more than **25** years. Assuming that the **population** is **normally** distributed with **variance** equal to **20**. A **random sample** of **10** individuals drawn from the population of interest. From this sample, a **mean** of **27** is calculated. Construct the proper hypothesis, test your hypothesis, and then state the proper conclusion? Use **α = 0.05** to test the hypothesis?

<u>Solution</u>

We have:

1- Normal distribution.

2- The standard deviation σ is known.

⇒ Then we will use the standard normal distribution (Z).

$n = 10$, $\mu_0 = 25$, $\bar{X} = 27$, $\sigma = \sqrt{\sigma^2} = \sqrt{20} = 4.472$, $\alpha = 0.05$

Example

Hypotheses

$H_0: \mu = \mu_0$ → $H_0: \mu = 25$

vs $H_1: \mu > \mu_0$ → $H_1: \mu > 25$ (more than)

Rejection Rule

Reject H_0 at a level of significance $\alpha = 0.05$ if:

$Z > Z_{1-\alpha}$, $Z_{1-\alpha} = Z_{1-0.05} = Z_{0.95} = 1.645$

Test statistics (calculated value)

$$z = \frac{\bar{X} - \mu_0}{\sigma / \sqrt{n}} = \frac{27 - 25}{4.472 / \sqrt{10}} = 1.414$$

Continued

Decision

We get **$Z = 1.414$** < **$Z_{1-\alpha} = 1.645$**

Then the rejection is not satisfied and therefore the decision will be do not reject (accept) H_0 at $\alpha = 0.05$ and therefore we conclude that the mean age is 25 years, that is, $\mu = 25$ years

Example

The **mean** maximum oxygen uptake for a sample of **242** women was **32.3** with a **standard deviation** of **12.14**, we wish to know if, on the basis of the data, can we conclude that the **mean** score for a population of such women is **smaller** than **33**? Use **α = 0.01** to test the hypothesis?

<u>Solution</u>

We have:

1- Unknown distribution (population).

2- The standard deviation σ is unknown (S = 12.14).

3- The sample size (n) is large (n = 242 > 30).

⇒ Then we will use the standard normal distribution (Z).

$$n = 242 \ , \ \mu_0 = 33 \ , \ \bar{X} = 32.30 \ , \ S = 12.14 \ , \ \alpha = 0.01$$

Continued

Lower (Left) -Tailed Test

Hypotheses

$H_0 : \mu = 33$
vs $H_1 : \mu < 33$ (smaller than).

Rejection Rule

Reject H_0 at level of significance $\alpha = 0.01$ if :

$Z < -Z_{1-\alpha}$

Test Statistic (*Calculated Value*)

$$Z = \frac{\bar{X} - \mu_0}{S/\sqrt{n}}$$

$$\Rightarrow Z = \frac{32.30 - 33}{12.14/\sqrt{242}} = -0.897$$

Critical Value(*Tabulated Value*)

$Z_{1-\alpha} = Z_{1-0.01} = Z_{0.99} = 2.33$

$\Rightarrow -Z_{1-\alpha} = -Z_{1-0.01} = -Z_{0.99} = -2.33$

Decision

We get $Z = -0.897 > -Z_{1-\alpha} = -2.33$

Then the rejection rule is not satisfied and therefore the decision will be do not reject (accept) H_0 at $\alpha = 0.01$ and therefore we conclude that the mean score for a population of such women is 33, that is, $\mu = 33$.

Example

The body mass index (BMI) of a **group** of **14** healthy adult males has a **mean of 30.5** and a **standard deviation of 10.6392**, can we conclude that the **mean** BMI of the **population** is equal to **36** assuming that the population is normally distributed? Use **α = 0.1** to test the hypothesis?

Solution

We have:

1- Normal distribution (Normal population).

2- The standard deviation σ is unknown ($S = 10.6392$).

3- The sample size (n) is small ($n = 14 < 30$).

➡ Then we will use the t-distribution.

$n = 14$, $\mu_0 = 36$, $\bar{X} = 30.5$, $S = 10.6392$, $\alpha = 0.10$

<u>Two -Tailed Test</u>

<u>Hypotheses</u>

$H_0 : \mu = 36$

vs $H_1 : \mu \neq 36$ (does not equal).

<u>Rejection Rule</u>

Reject H_0 at level of significance $\alpha = 0.10$ if :

$$|t| > t_{(\frac{\alpha}{2}, n-1)}$$

Test Statistic (*Calculated Value*)

$$t = \frac{\bar{X} - \mu_0}{S/\sqrt{n}}$$

$$\Rightarrow \quad t = \frac{30.5 - 36}{10.6392/\sqrt{14}} = -1.934$$

Critical Value(*Tabulated Value*)

$$t_{(\frac{\alpha}{2}, n-1)} = t_{(\frac{0.10}{2}, 14-1)} = t_{(0.05, 13)} = 1.771$$

Decision

We get $|t| = |-1.934| = 1.934 > t_{(\frac{\alpha}{2}, n-1)} = t_{(0.05, 13)} = 1.771$

Then the rejection rule is satisfied and therefore the decision will be <u>reject</u> H_0 at $\alpha = 0.10$ and therefore we conclude that the mean BMI of the population is not equal to 36, that is, $\mu \neq 36$. In other words H_1 is accepted at $\alpha = 0.10$.