Module-4
Test of Significance

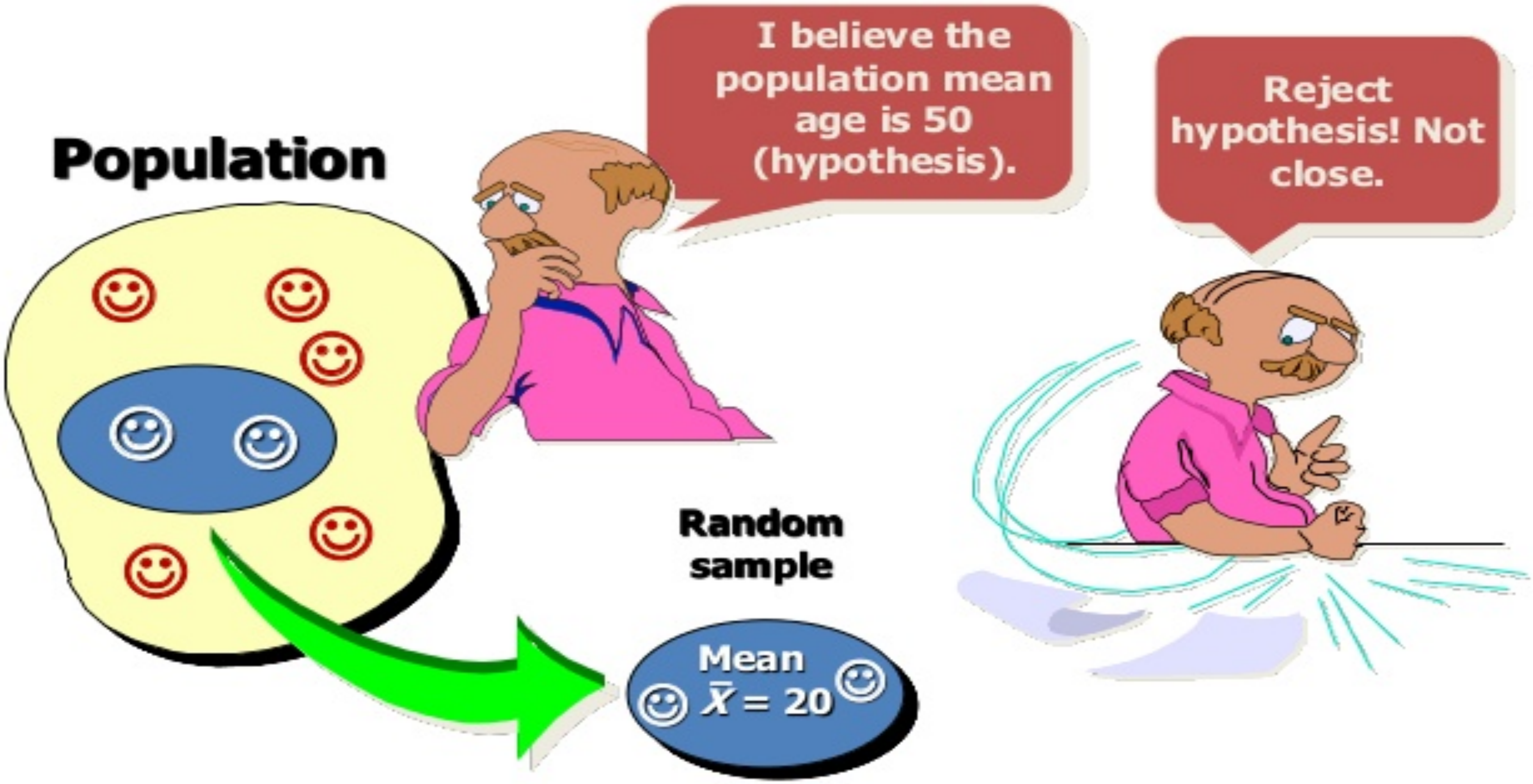
Outline

- ***Test of significance***
- ***Practical steps involved in the test of significance***
- ***Test of significance of large samples***
- ***Test of significance of small samples***

Test of significance (Test of Hypothesis)

- It is a process of making inferences about population on the basis of sample information
- It is a decision making tool

HYPOTHESIS TESTING



Null Hypothesis (H_0)

-The assumption you're beginning with

-The opposite of what you're testing

Alternative Hypothesis (H_1)

-The claim you're testing

example

- Null hypothesis- there has been no significant decrease in the consumption of tea after the increase in excise duty
- Alternative hypothesis- there has been significant decrease in the consumption of tea after the increase in excise duty

example

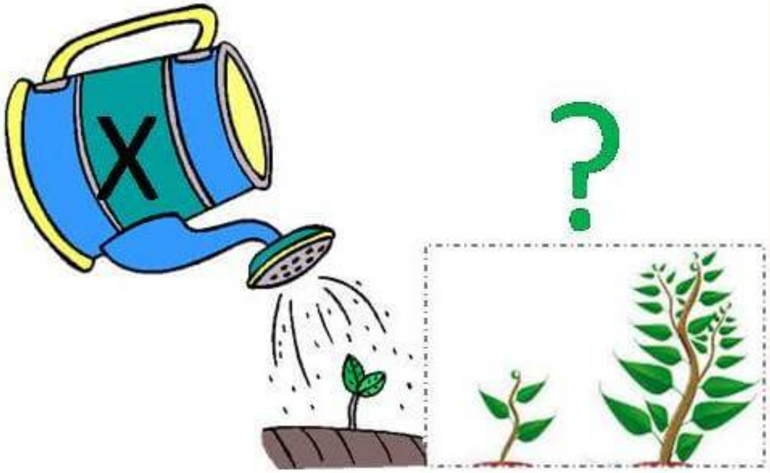
- Null hypothesis- the coin is unbiased.
- Alternative hypothesis- the coin is biased

Effect of Bio-fertilizer 'x' on Plant growth

www.majordifferences.com

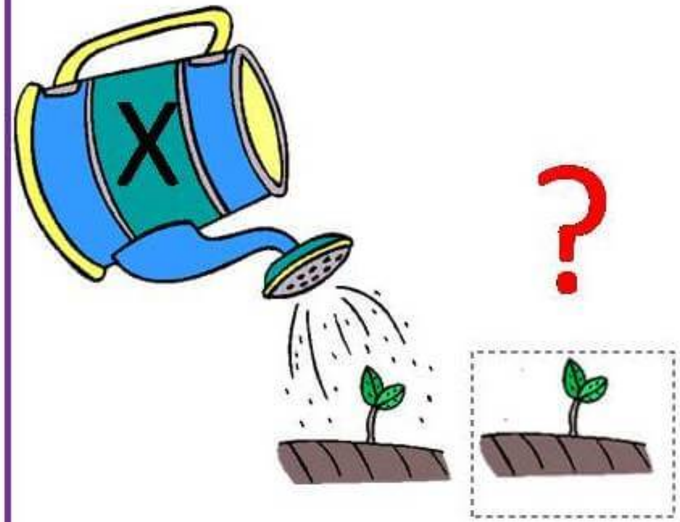
Alternative Hypothesis

H_1: Application of bio-fertilizer 'x' increase plant growth.



Null Hypothesis

H_0: Application of bio-fertilizer 'x' do not increase plant growth.



Type I and Type II Error

		Given the Null Hypothesis Is	
		True	False
Your Decision Based On a Random Sample	Reject	Type I Error	Correct Decision
	Do Not Reject	Correct Decision	Type II Error

Two Types of Errors in Decision Making

		The Truth (Based on Entire Population)	
		Nothing Is There (H_0 Is True)	Something Is There (H_0 Is False)
Your Conclusion (Based on Your Sample)	I Don't See Anything (Nonsignificant)	Right!	Wrong (Type II Error)
	I See Something (Significant)	Wrong (Type I Error)	Right!

Practical steps involved in the test of significance or testing hypothesis

- Step1- Specify the Null and Alternative Hypothesis
- Step 2: Specify the appropriate test statistic to be used
- Step 3 – Specify the level of significance such as 5% or 1%.
- Step4 – Compute the value of test statistic used in testing.
- Step 5 – Find the critical value of the test statistic used at the selected level of significance from the table of respective statistics distribution.
- Step 6: Interpret the result

SMALL SAMPLE AND LARGE SAMPLE

Large sampling

- $n>30$

Small sampling

- $n<= 30$

small sampling test

- t-test
- F-test

t- test

t-test is used when

- The sample size is 30 or less
- The variance of the population is unknown
- The sample is random sample
- The population is normal and selection of items is independent

Test for specified mean of a small sample	$t = \frac{(\overline{X} - \mu)\sqrt{n}}{S}$ $S = \sqrt{\frac{\sum (X - \overline{X})^2}{n-1}}$

$$\mu = population\,mean$$

$$\overline{X} = sample\,mean$$

$$s = s.d\ of\ sample$$

$$n = sample\,size$$

Test for difference between the means of two independent random samples	$t = \frac{\lvert \overline{X}_1 - \overline{X}_2 \rvert}{S} \times \sqrt{\frac{n_1 n_2}{n_1 + n_2}}$ $S = \sqrt{\frac{(n_1 - 1)S_1^2 + (n_2 - 1)S_2^2}{n_1 + n_2 - 2}}$

$\overline{X1} = mean\ of\ \ first\ sample$

$\overline{X2} = mean\ of\ \sec ond\ sample$

$n1 = samplesize\ \ of\ \ first\ sample$

$n2 = sample\ size\ of\ \sec ond\ sample$

$s = s.d\ of\ sample$

Test for difference between the means of two dependent samples	$$t = \frac{\bar{d}\sqrt{n}}{S}$$ $$S = \sqrt{\frac{\sum d^2 - n(\sum d)^2}{(n-1)}}$$ **d = difference between values of pair** **n= sample size** **S= standard deviation of sample**

Question

- A fertilizer mixing machine is set to give 4 kg of nitrate for every quintal bag of fertilizers. Five of 100 kg bags are examined. The percentages of nitrate are: 2, 6, 4, 3, 1. Is there reason to believe that the machine is defective?

solution

$$H_0 : \mu = 4$$

$$H_1 : \mu \neq 4$$

Calculate Mean and Standard Deviation

mean

X
2
6
4
3
1
Total=16

$$Mean$$

$$\overline{X} = \frac{\sum X}{n}$$

$$\overline{X} = \frac{16}{5} = 3.2$$

Standard deviation

X	$X-\overline{X}$	$(X-\overline{X})^2$
2	-1.2	1.44
6	2.8	7.84
4	0.8	0.64
3	-0.2	0.04
1	-2.2	4.84
Total= 16		Total= 14.8

$$S=\sqrt{\frac{(X-\overline{X})^2}{n-1}}$$

$$S=\sqrt{\frac{14.8}{5-1}}=\sqrt{\frac{14.8}{4}}=1.93$$

$$t = \frac{(\overline{X} - \mu)\sqrt{n}}{S}$$

$$t = \frac{(3.2 - 4)\sqrt{5}}{1.93} = -0.927$$

- Critical Value of t at 5% level of significance = 2.78

Interpretation

- Since the computed value of t is less than table value of t, we accept the null hypothesis and conclude that the population mean is 4

To test the significance of the difference between means of two independent samples

$$t = \frac{\left|\overline{X_1} - \overline{X_2}\right|}{S} \times \sqrt{\frac{n_1 n_2}{n_1 + n_2}}$$

where

$$\overline{X_1} = \frac{\sum X_1}{n_1}$$

$$\overline{X_2} = \frac{\sum X_2}{n_2}$$

$$S = \sqrt{\frac{\sum (X_1 - \overline{X_1})^2 + \sum (X_2 - \overline{X_2})^2}{n_1 + n_2 - 2}}$$

Question

- A group of 5 patients treated with medicine A weigh 42, 39, 48, 60, 41 kg. A second group of 5 patients treated with medicine B weigh 38, 42, 48, 67, 40 kg. Do the two medicines differ significantly with regard to their effect in increasing weight?

given : Value of t at 5% level = 2.31

Solution

X1	X2	$(X_1-\overline{X_1})$	$(X_1-\overline{X}_1)^2$	$(X_2-\overline{X_2})$	$(X_2-\overline{X}_2)^2$
42	38				
39	42				
48	48				
60	67				
41	40				
TOTAL=					

MEAN X1 =
MEAN X2 =
STANDARD DEVIATION S =
t =

To test the significance of the difference between the means of two dependent samples

$$t = \frac{\bar{d}\sqrt{n}}{S}$$

$$S = \sqrt{\frac{\sum d^2 - n(\bar{d})^2}{n-1}}$$

$$d = differenes\, between\, the\, values\, of\;\; pair$$

Question

- A certain medicine was given to each of the 5 patients. The results are given below:

	I	II	III	IV	V
Weight before medicine	42	39	48	60	41
Weight after medicine	38	42	48	67	40

- Given tabulated value of $t = 2.78$

Solution

Before X	After Y	d= Y-X	d^2
42	38		
39	42		
48	48		
60	67		
41	40		
n=5			

F- Test

- F-test is based on the ratio rather than difference between variances.

$$f = \frac{s_1^{\,2}}{s_2^{\,2}}$$

$$f = \frac{\sum (X_1 - \overline{X}_1)^2 / (n_1 - 1)}{\sum (X_2 - \overline{X}_2)^2 / (n_2 - 1)}$$

To compare the price of a certain commodity in two towns, ten shops were selected in each town. The following figures give the prices found:

Town A	**61**	**60**	**56**	**63**	**56**	**63**	**59**	**56**	**44**	**61**
Town B	**55**	**54**	**47**	**69**	**51**	**61**	**57**	**54**	**62**	**58**

Test whether the average price can be said to be same in the two towns.

Test of significance of large samples

- Large sample test (z test) = size of sample>30
- **Assumptions**
  - Sampling distribution of a statistic is approximately normal
  - Samples are approximately close to the population value
- Standard error is used in z test
- Standard error = $\dfrac{\sigma}{\sqrt{n}}$

$$\sigma = population\ S.D$$

$$n = sample\ size$$

Tabulated value of z

- At 5% level of significance = 1.96
- At 1% level of significance = 2.58

- A coin is tossed 400 times and it turns up head 216 times. Discuss whether the coin is unbiased.
- **Solution**
- Null hypothesis: the coin is unbiased
- Alternative hypothesis: the coin is biased
- Probability of turning up head in each toss (p) = ½
- Probability of turning up tail in each toss (q) = ½

$$actual\ head = 216$$

$$expected\ head(np) = 400 * \frac{1}{2} = 200$$

$$\sqrt{npq} = \sqrt{400 * \frac{1}{2} * \frac{1}{2}} = 10$$

$$z = \frac{difference}{SE} = \frac{216 - 200}{10} = 1.6$$

$$standard\ value = 1.96$$

Decision :the coin is unbiased

Question

- Intelligence test on two groups of boys and girls gave the following results:

	Mean	S.D	N
Girls	61	2	64
Boys	60	4	100

- Is there a significant difference in the mean scores obtained by boys and girls? Test at 5% level of significance

solution

$$Z = \frac{\overline{X_1} - \overline{X_2}}{S.E}$$

$$S.E = \sqrt{\frac{\sigma_1{}^2}{n_1} + \frac{\sigma_2{}^2}{n_2}}$$

$$S.E = 0.4717$$

$$Z = 2.12$$

At 5% level, the critical value of $Z = 1.96$

reject $H0$