Statistical tests

By Saurav Poudel

Why do we need Statistical tests?

Remember Hypothesis testing?

Steps for Hypothesis testing

- ***Step 1 : State the hypothesis. Both Null and Alternate. And they should be mutually exclusive.***
- ***Step 2 : Identify the test statistic.***
- ***Step 3 : Calculate the p-value. P-value is the probability that the test statistic at least as significant as the one observed would be obtained assuming the null hypothesis were true.***
- ***Step 4 : Compare your p-value with your alpha value. Alpha depends on our confidence level.***

Let's think of some examples where we need Statistical tests first!

Statistical tests

Only the ones we are covering!

- T-Test
- ANOVA Test (Extra - Out of Syllabus)

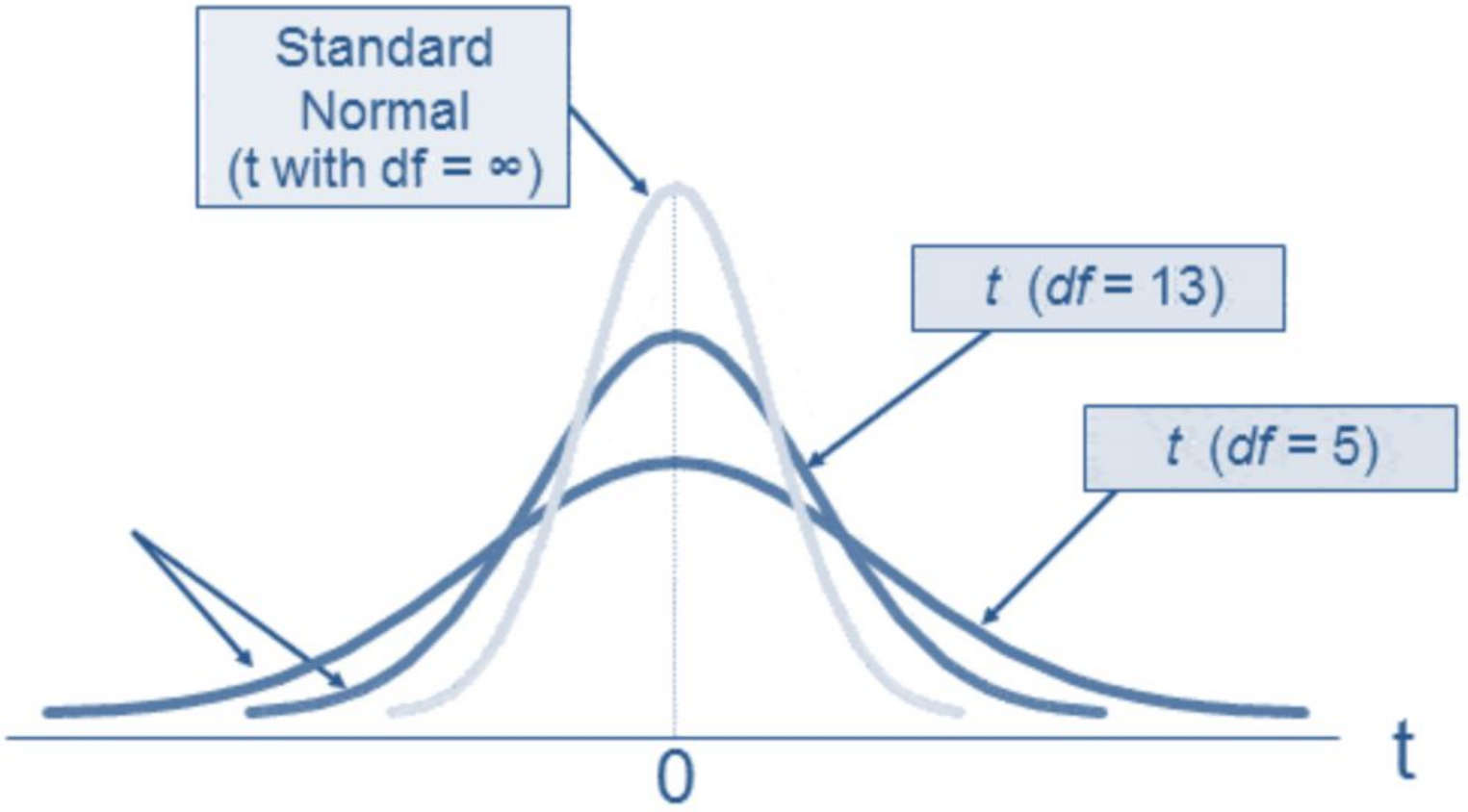
T - Test & T - Distribution

Less Data -> Sample Means Distribution Not Normal

Less Data -> More Variance -> More SD

Normal Distribution -> Critical Value 1.96

More SD -> Greater Critical Value than 1.96


Standard
Normal
(t with df = ∞)
t (df = 13)
t (df = 5)
0
t

Degrees of freedom	Significance level					
	20% (0.20)	10% (0.10)	5% (0.05)	2% (0.02)	1% (0.01)	0.1% (0.001)
1	3.078	6.314	12.706	31.821	63.657	636.619
2	1.886	2.920	4.303	6.965	9.925	31.598
3	1.638	2.353	3.182	4.541	5.841	12.941
4	1.533	2.132	2.776	3.747	4.604	8.610
5	1.476	2.015	2.571	3.365	4.032	6.859
40	1.303	1.684	2.021	2.423	2.704	3.551
60	1.296	1.671	2.000	2.390	2.660	3.460
120	1.289	1.658	1.980	2.158	2.617	3.373
∝	1.282	1.645	1.960	2.326	2.576	3.291

Today, we take 'formulae and memos to memorize' route!

Things to remember!

P-Value less than 0.05 or not!

- For every test, we get a test score.
- For every score, there is a corresponding P-Value.
- For only P-Value less than 0.05 (for 95 % Confidence Level, which is the most common one) we reject Null.

T-Test is a procedure used for comparing Sample Means to see if there is sufficient evidence to infer that the means of the corresponding population distributions also differ. The important things are;

1. Two (t-test always compare two different means)

2. Some variable of interest

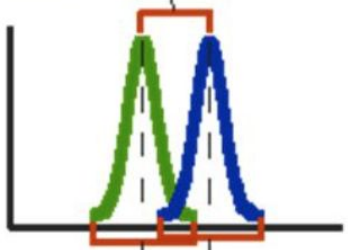
Company A		Company B
10,000		25,000
10,000		25,000
10,000		30,000
20,000		30,000
20,000		30,000
15,000		30,000
15,000		40,000
10,000		40,000
2,00,000		40,000
2,00,000		50,000

What does a t-test measure?

$$\frac{\text{signal}}{\text{noise}} = \frac{\text{difference between group means}}{\text{variability of groups}}$$

$$= \frac{\bar{X}_T - \bar{X}_C}{SE(\bar{X}_T - \bar{X}_C)}$$

$$= \text{t-value}$$



Note: T = treatment group and C = control group. (The above depicts a comparison in experimental research; in most discussions these will just be shown as groups 1 and 2, indicating different groups.)

$$t = \frac{\overline{x_1} - \overline{x_2}}{\sqrt{\frac{S_1^{\,2}}{N_1} + \frac{S_2^{\,2}}{N_2}}}$$

Where

$\mathbf{x_1}$ is the mean of first data set

$\mathbf{x_2}$ is the mean of first data set

$\mathbf{S_1^2}$ is the standard deviation of first data set

$\mathbf{S_2^2}$ is the standard deviation of first data set

$\mathbf{N_1}$ is the number of elements in the first data set

$\mathbf{N_2}$ is the number of elements in the first data set

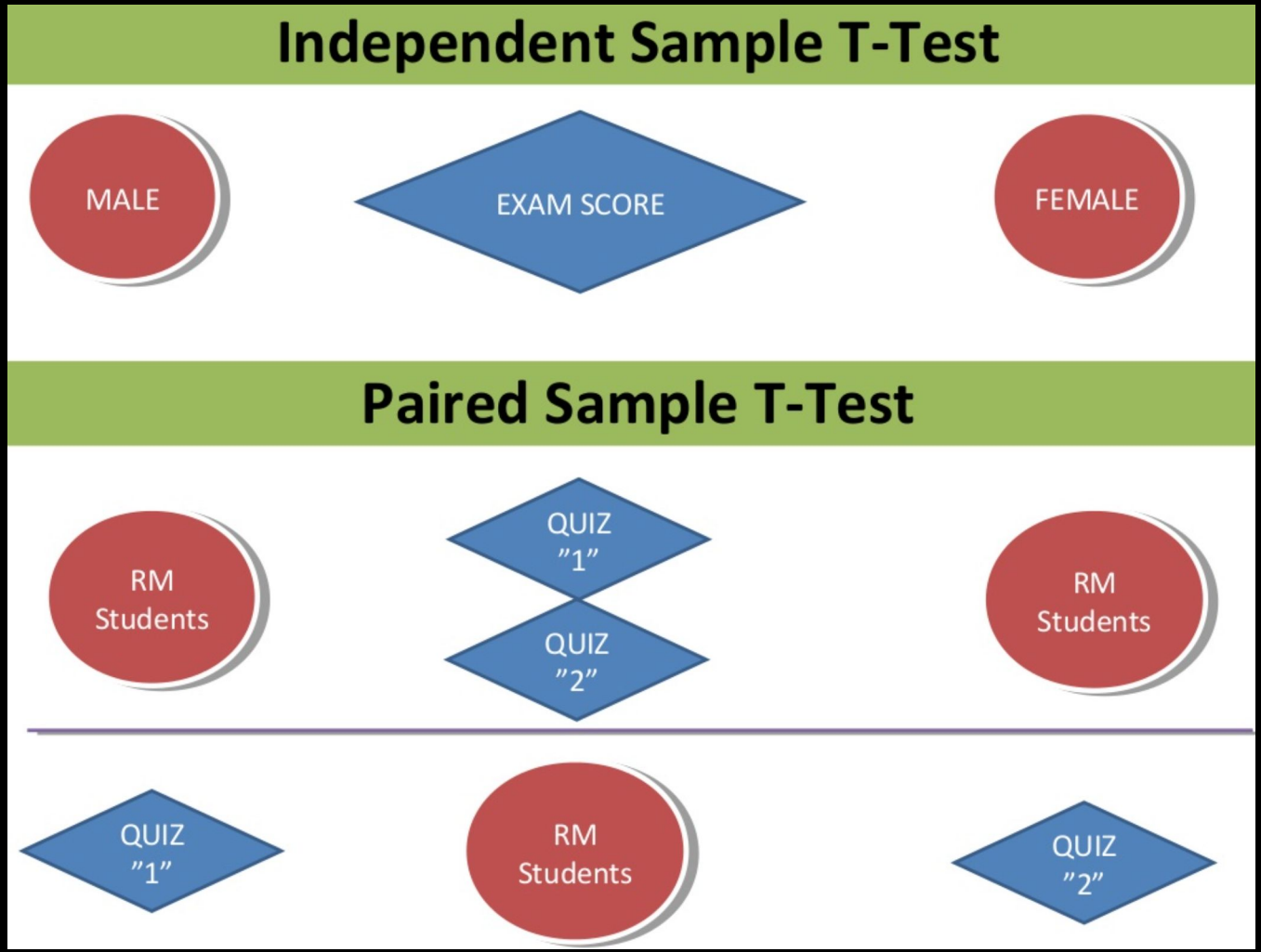
So far we have learned the following things about a t-test;

1. The t-test produces a single value, *t*, which grows larger as the difference between the means of two samples grows larger;
2. *t* does not cover a fixed range such as 0 to 1 like probabilities do;
3. You can convert a t-value into a probability, called a p-value;
4. The p-value is always between 0 and 1 and it tells you the probability of the difference in your data being due to sampling error;
5. The p-value should be lower than a chosen significance level (0.05 for example) before you can reject your null hypothesis.

Types of T-Test

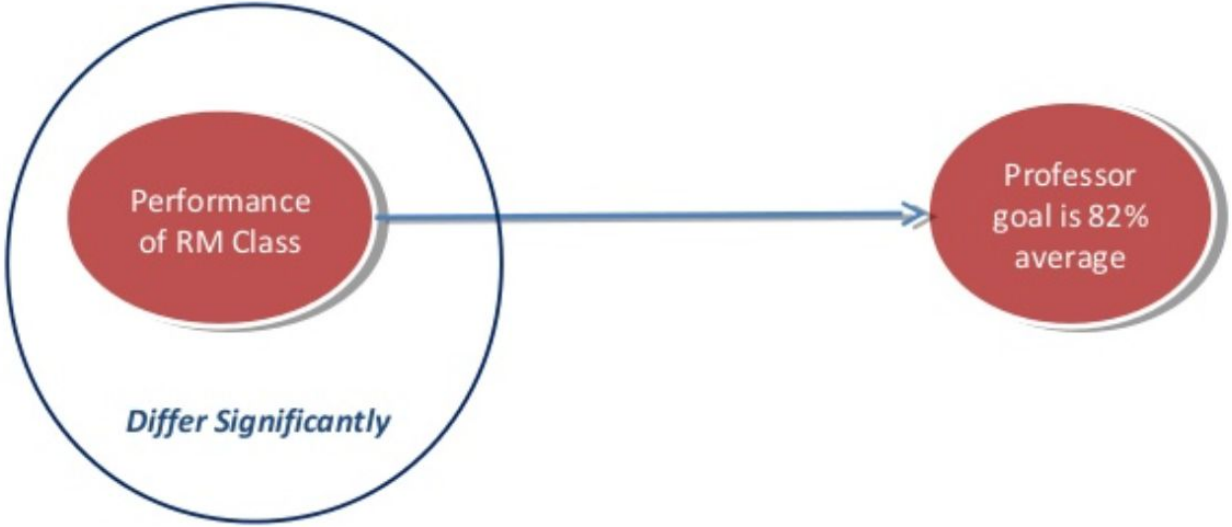
- Independent Sample T-Test
- Paired Sample T-Test
- One Sample T-Test

Independent Sample T-Test	Paired Sample T-Test
• The two samples share some variable of interest in common, but there is no overlap between membership of two groups. • Compare the running speeds of horses and zebra would be an independent design as there is no sensible way to pair off each horse with each zebra.	• Usually based on the group of individuals who experience both conditions of the variable of interest. • Also called a Repeated Measure Design or a Paired Design. • Compare the running speed of horses for a week of eating one type of feed with the ***same*** horses for a week on a different type of feed would be a paired design as you can pair off measurements from the same horse


Independent Sample T-Test
MALE
EXAM SCORE
FEMALE
Paired Sample T-Test
RM Students
QUIZ "1"
QUIZ "2"
RM Students
QUIZ "1"
RM Students
QUIZ "2"

One Sample T-Test

It is designed to test whether the mean of a distribution differs significantly from some present value e.g.



Test of Significance

Test of Significance can be one-tailed or two tailed test;

Two tailed test examines whether the mean of one distribution differs significantly from the mean of other distribution. (Regardless of the direction +ve or –ve)

The one tailed test measures only whether the second distribution differs in a particular direction from the first.

One Tailed OR Two Tailed

If you have stated your experimental hypothesis with care, it will tell you which type of effect you are looking for.

For example, the hypothesis that "*Coffee improves memory*" is _____tailed test.

The hypothesis, "*Men weigh a different amount from women*" suggests a _____tailed test.

So remember, don't be vague with your hypothesis if you are looking for a specific effect! Be careful with the null hypothesis too - avoid "*A does not effect B*" if you really mean "*A does not improve B*".

Full Name of T - Test:

Student's T - Test

Why is it called Student's T - Test?

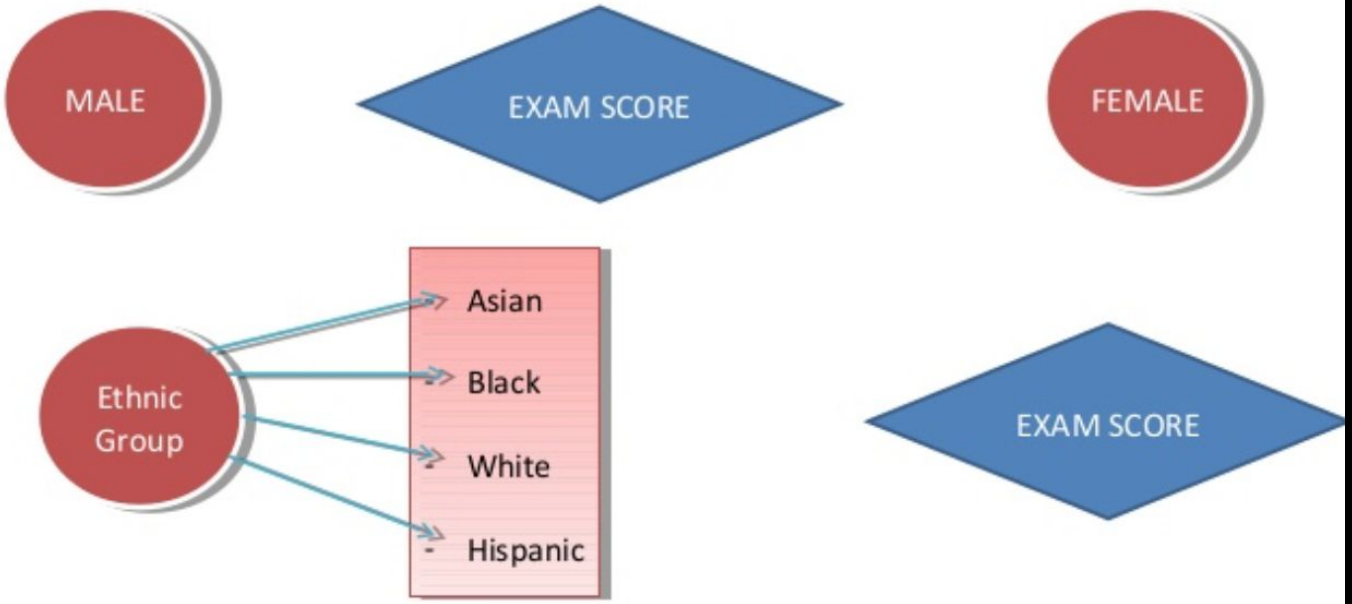
This is for you to figure it out!

ANOVA

- Analysis of variance is a procedure used for comparing sample means to see if there is sufficient evidence to infer that the means of the corresponding population distributions also differ.

- Where t-test compare only two distributions, analysis of variance is able to compare many.

ANOVA



One-way ANOVA will generate a significance value indicating whether there are significant differences within the comparisons being made. This significance value does not indicate where the difference is or what the differences are; but a 'Test' can identify which groups differ significantly from each other.

The difference between t-tests and ANOVAs is that t-tests are used to analyze data from independent variables with just two levels

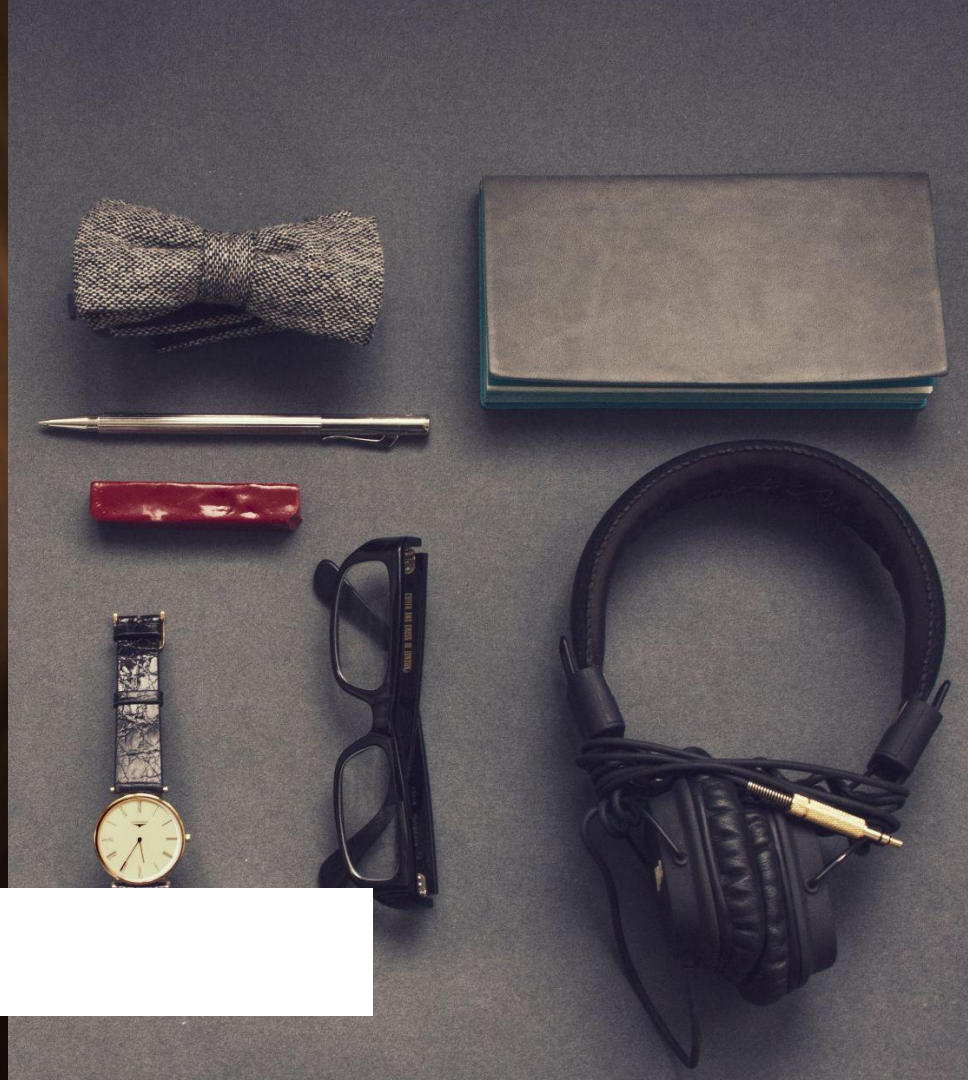
For example – studying driving speed across **two levels** of athlete *(football and basketball players).*

Whereas ANOVAs are used to analyze data from independent variables with three or more levels

For example – studying driving speed across **three levels** of athlete *(football, basketball, soccer players).*

Can you think of more applications of Statistical Tests?

Causality



Thank You!