

TEST	T test	Z test	ANOVA
GROUP	2 Group	2 Group	> 2 Group
N (total rows)	$N \leq 30$	$N > 30$	

Q) What is T test? (Explain the same for Z test).

Ans - It works good with normal distribution.

- T test assess whether the mean of two groups are statistically different from each other.

$$T \text{ test} = \frac{\text{Difference between mean of two groups}}{\text{Standard error of difference between mean}} = \frac{M_2 - M_1}{SE \text{ of different mean}}$$

A Big t-value = Different group      A small t-value = similar groups

- Each t-value has a p-value. The p-value tells us the likelihood that there is a real-difference.
  - P value will tell us if two group are real different or just by a fluke. P value is the probability that the pattern of data in the sample could be produced by a random state. if CI is at 95% level.
- $p = 0.10$ , there is 10% chances there is no real difference (rejected)  
 $p = 0.01$ , there is 1% chances there is no real difference (accepted)  
 $p = 0.02$ , there is 2% chances there is no real difference (accept)

Q) What is the Null hypothesis and alternate hypothesis in T test?

T test, Null  $\rightarrow$  There is no difference between means of two group.

Alternate  $\rightarrow$  There is a difference between means of two group.

Null hypothesis means two group are equal by mean.

So, if p value < alpha level or p value < 0.05, we can reject the null hypothesis that there is no difference between means (Null).

Q) Suppose there are two groups. Mean(Group<sub>1</sub>) = 40 and Mean(Group<sub>2</sub>) = 60. So difference of mean is 20. So why not directly say mean of two Group (Group<sub>1</sub> and Group<sub>2</sub>) are different, rather than use T test?

$\rightarrow$  Yes difference of mean is 20, but we can't be sure if its reliable difference. Suppose we toss a coin 100 times and found head = 52 and tail = 48. So this doesn't suggest, we will get heads more in future. This is only a chance.

So we use Inferential statistics rather than Descriptive statistics.  
 (Doesn't happen by a chance)      (Can be happen by a chance)



DESCRIPTIVE STATISTICS	INFERENCEAL STATISTICS
<p>1) Gives information about raw data which describe data in some manner.</p> <p>2) It is used to describe a situation. Eg mean of the salary of population.</p> <p>3) It explain already known data and limited to a sample/ population having small size.</p> <p>4) Example <math>\rightarrow</math> Mean, median, range, mode, SD</p>	<p>1) Makes inference about population using data drawn from population.</p> <p>2) It is used to explain the chance of occurrence of the event. (p values).</p> <p>3) It attempts to reach the conclusion about the population.</p> <p>4) Example - T test, z test ANOVA etc.</p>

**Inferential statistics** allow you to make predictions by taking a small sample instead of working on whole population.

Q) **What type of t-test should I use?**

Ans  $\rightarrow$  i) One sample, two sample, pair t test.

a) **Paired t-test**  $\rightarrow$  If the groups comes from single population  
Eg measuring before & after an experiment treatment.

b) **Two sample t-test**  $\rightarrow$  If the groups comes from two different population  
(Independent t test) Eg two different species/people from two separate cities.

c) **One-sample t test**  $\rightarrow$  If there is one group compared against standard value. Eg comparing acidity of a liquid to neutral pH 7.

ii) One tailed or two tailed t test.

a) **Two tailed t test**  $\rightarrow$  If we want to know/only care whether two population are different from one another.

b) **One tailed t test**  $\rightarrow$  If we want to know whether one population mean is greater than or less than the other.

Q) **Give an example of t-test?**

Ans - Suppose Fixed deposit tenure (1 year/2y/5y) differ by Gender (male/female)

- Our observation comes from two different population (male, female), use two sample t-test.

- we don't care direction of difference, only whether there is difference. so we choose two tailed t-test. NULL - No difference, both are same. Alternate - Both are different.

So based on P value we decide to reject/accept Null hypothesis.