

ANOVA (Analysis of Variance) and its assumption

Definition \Rightarrow ANOVA is a statistical method used to compare the means of 2 or more groups

* Using variance, you want to determine, if the mean value of two or more than two group are same or different.

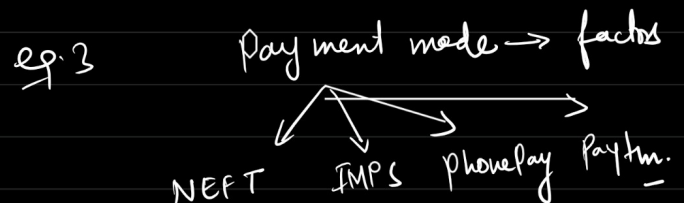
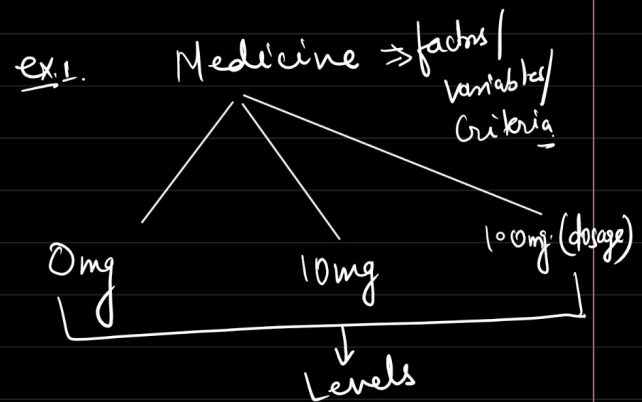
\rightarrow Generally ANOVA is used for more than 2 groups

\rightarrow Reason: for two groups you can use t-test.

ANOVA terms

* Factor / variables / criteria

* Levels



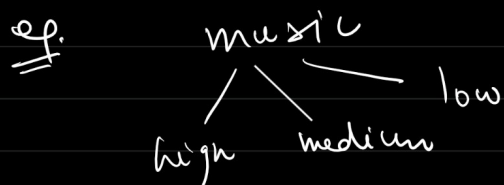
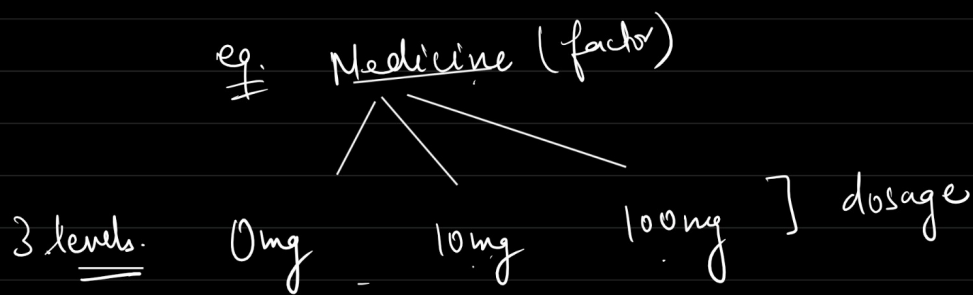
* Assumptions of ANOVA

- The population from which samples are drawn should be normally distributed.
- The sample should be independent of each other/random.
- Absence of outliers
- Homogeneity of variance:- Homogeneity means that the variance among the groups should be approximately equal.

	Pop.	Sample	S_1	S_2	S_3
std dev σ		S			
var σ^2		S^2			
			$S_1^2 = S_2^2 = S_3^2$		

Types of ANOVA (3 types)

① One way ANOVA : One factor with atleast two levels, and levels are independent.



eg. Stree level of employees

Dept A	Dept B	Dept C
8	1	6
5	2	5
6	3	2
2	4	9

② Repeated measure ANOVA : One factor with atleast two levels but level are dependent.

eg. No of hours studied.

Day 1 ← Day 2 ← Day 3 ← Day 4
10 8 4 10

eg. Gym

Day 1 ← Day 2 ← Day 3 ← Day 4
5 1 hr 0.5 1 hr

③ Factorial ANOVA :- Two or more factors each of which with atleast two levels. Levels can be either dependent, independent or both.

→ Also called as two way ANOVA → It deals with only two factors.

		Medicine (factor)			
		0mg	10mg	100mg	(dosage) → levels
Gender (2nd factor)	male	2	3	3	
		3	4	2	
		4	1	1	
		5	5	6	
	female	1	2	7	
		2	3	8	
		1	1	6	
		3	4	5	

Runnig shoes Brand

days
(2nd factor)

	brand A	brand B	brand C
Monday	10 12 13	100 23 24	101 48 95
wednesday	15 16 17	16 18 23	23 25 26
Saturday	19 20 21	25 26 27	28 27 17