14/03

ANALYSIS OF VARIANCE (ANNOVA)

Annova is a statistical method used to compare the means of 2 or more groups

Components of Annova:-

① Factors (variable) → Independant / → dependent

② Levels

ex:- medicine (factors)

(dosage) {5 mg, 10 mg, 15 mg} → Levels

ex:→ Mode of Payment (factor)

G-Pay, Phone pe, Imps, NEFT } (Levels)

Assumptions of Annova:-

① **Normality of Sampling Distribution of Mean:-**

The distribution of sample mean is normally distributed.

② **Absence of Outliers:-**

Outlying score needs to be removed from the dataset

ex: using boxplot etc.

③ **Homogenity of Varience:-**

Population varience in different levels of each independent variable are equal.

$[\sigma_1^2 = \sigma_2^2 = \sigma_3^2]$

④ **Samples are independent and random:**

Types of Annova:-

① **One Way Annova** :- One factor with atleast 2 levels, these levels are independent

ex: Doctor want to test a new medication to decrease headache. They split the participants in 3 conditions/groups based on no. of dosage of medication given to them [10mg, 20mg, 30mg]. Doctor ask participants to rate the headache from [1-10]

here,

factor → medication

level → 10 mg, 20 mg, 30 mg (independent levels)

10 mg	20 mg	30 mg
5	7	2
3	4	6
⋮	⋮	⋮

② **Repeated Measures Annova** :- One factor annova with atleast 2 levels, levels are dependent.

ex: factor → running

levels → Day 1, Day 2, Day 3 (dependent levels)

Day 1	Day 2	Day 3
8	5	4
7	4	9
⋮	⋮	⋮

③ **Factorial Annova** :- 2 or more factors each of which with at least 2 levels, levels can be independent and dependent.

ex: factors: Running, Gender

levels: Day 1, Day 2, Day 3 | Male, Female

Day 1	Day 2	Day 3
8	5	4
9	4	3
2	4	2
7	8	9

Hypothesis Testing in Annova (Partitioning of Variance in Annova)

Step 1) Define Null Hypothesis (H_0): $\mu_1 = \mu_2 = \mu_3 \ldots\ldots$ (compaining mean of 2 or more groups)

Step 2) Alternate Hypothesis H_1: atleast one of the mean is not equal

★ To perform this we use F-Test:

$$\boxed{F = \frac{\text{Variance b/w samples}}{\text{Variance within Samples}}}$$

Q1) Explain Variance b/w & Variance within Sample?

Solⁿ →

x_1	x_2	x_3
1	6	5
2	7	6
4	3	3
5	2	2
3	1	4
$\bar{x}_1 = 3$	$\bar{x}_2 = 19/5$	$\bar{x}_3 = 4$

Variance b/w Sample:- is the difference b/w the variances of x_1, x_2, x_3 (i.e $\sigma_1, \sigma_2, \sigma_3$)

Varience within Sample: How much the variace of a sample differs from its mean.

One Way Annova

one factor with atleast 2 independent levels

Qy Doctors want to test a new medication which reduces headaches. They splits the participants into 3 condition [15mg, 30mg, 45mg]. Later on the doctor ask the patient to rate the headaches b/w [1-10]. Are there any differences b/w the 3 conditions using alpha = 0.05?

15mg	30mg	45mg
9	7	4
8	6	3
7	6	2
8	7	3
8	8	4
9	7	3
8	6	2

Soln → Step1) $H_0 = \mu_{15} = \mu_{30} = \mu_{45}$

H_1 = not all μ are equal

Step2) Significance value 0.05 ⇒ CI = 95%

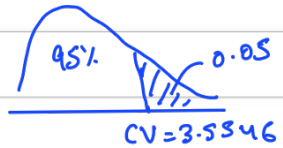
Step3) Calculate dof: total samples $N = 7 \times 3 = 21$

dof (b/w samples) = $a - 1 = 3 - 1 = \boxed{2}$ df1

dof (within sample) = $N - a = 21 - 3 = \boxed{18}$ df2

dof (total) = $N - 1 = 20$

Step4) Construct Decision Boundry (F test ⇒ Righ skewed)



if F. is > 3.5546 → reject H_0

15 mg	30 mg	45 mg
9	7	4
8	6	3
7	6	2
8	7	3
8	8	4
9	7	3
8	6	2

- $F = \dfrac{\text{Variance b/w Sample}}{\text{Variance within Sample}}$
- SS – Sum of Squares
- MS → mean square

	SS	dof	MS	F
b/w				
within				
Total				

① $$\boxed{SS(b/w) = \frac{\Sigma(\Sigma a_i)^2}{n} - \frac{T^2}{N}}$$

Σ 15 mg : $9+8+7+8+4+8 = 57$

Σ 30 mg : . . . . . $= 47$

Σ 45 mg : . . . . . $= 21$

$\Sigma(\Sigma a_i)^2 = (57)^2 + (47)^2 + (21)^2 =$

$T^2 = (57)^2 + (47)^2 + (21)^2 =$

$$\Rightarrow SS(b/w) = \frac{57^2 + (47)^2 + (21)^2}{7} - \frac{[(57)^2 + (47)^2 + (21)^2]}{21}$$

$= 98.67$

② $$\boxed{SS(\text{within}) = \Sigma Y^2 - \frac{\Sigma(\Sigma a_i)^2}{n}}$$

$\Sigma Y^2 =$ sum of sq of every value of all the samples

$= 853$

$$SS(\text{within}) = 853 - \frac{[(57)^2 + (47)^2 + (21)^2]}{7} = 10.29$$

	SS	dof	$MS = \frac{SS}{dof}$	F
b/w	98.67	3-1 = 2	49.34	
within	10.29	21-3=18	0.54	
Total	108.96	20		

Now,

$$\boxed{F\text{-test} = \frac{MS(b/w)}{MS(\text{within})}} = \frac{49.34}{0.54} = 86.56$$

Since, $F > 3.5546$ $\therefore$ we reject H_0