ANOVA

M T W T F S S
Page No.:
Date:
YOUVA

One Way ANOVA

① Define null & alternate hypothesis

data ⇒

A	B	C
3	1	8
6	8	6
3	9	10

(2) Perform ANOVA

NULL Hyp (H0): $\mu_A = \mu_B = \mu_C$ (All 3 pop mean are equal)

Alternative (H1): atleast one of them is signifi antly different

② Calculate grand mean of all groups combined and mean of individual groups.

Given $n = 9$ (total samples)
$k = 3$ (no of categories)

(i) grand mean $= \frac{3+6+3+1+8+9+8+6+10}{9}$

$\bar{\bar{x}} = \frac{54}{9} \Rightarrow \boxed{\bar{\bar{x}} = 6}$

Individual mean :

$\boxed{\bar{x}_A = 4}$, $\boxed{\bar{x}_B = 6}$, $\boxed{\bar{x}_C = 8}$

(3) Calculate "between group" and "within group" sum of square (SS)

Calculate the table.

i) SST → sum of square total

Calculating spread of each point from grand mean

Σ (each point - $\bar{x}$)^2

$= (6-3)^2 + (6-6)^2 + (6-3)^2 + (6-1)^2 + (6-6)^2 + (6-9)^2 + (6-8)^2 + (6-6)^2 + (6-10)^2$

SST = 76

(ii) SSW → sum of squares within

calculating variance of each group with its respective mean

A spread + B spread + C spread

$[(4-3)^2 + (4-6)^2 + (4-3)^2]$

$+$

$[(6-1)^2 + (6-8)^2 + (6-9)^2]$

$+$

$[(8-8)^2 + (8-6)^2 + (8-10)^2]$

$=$ SSW = 52

(iii) Calculating degrees of freedom for SSW

$= n - k$

$= 9 - 3$

$= 6$

(iv) Calculating degrees of freedom for SST

$= n - 1$

$= 9 - 1$

$= 8$

(v) SSB → sum of square between

Calculating spread of grand mean from each group mean and multiplied by ~~no~~ with their no of samples

$3 \times (\bar{x} - A_m)^2 + 3 \times (\bar{x} - B_m)^2 + 3 \times (\bar{x} - C_m)^2$

$= 3 \times (6 - 4)^2 + 3 \times (6 - 6)^2 + 3 \times (6 - 8)^2$

$\boxed{SSB = 24}$

(vi) Degree of freedom for SSB

$k - 1$

$= 3 - 1$

$= 2$

TABLE

quantity	value	df
SSB	24	2
SSW	52	6
SST	76	8

Trick to verify ⇒ SST = SSW + SSB
= 52 + 24
= 76

SST (Overall variance)
↙ ↘
internal variance groups (SSW) + SSB

(4) Calculate F-ratio (Test Statistic)

$$\frac{\text{MS between}}{\text{MS within}}$$

where MS between = $\frac{SSB}{\text{its df}}$ =

2

MS within = $\frac{SSW}{\text{its df}}$

$$F\ Ratio = \frac{\frac{SSB}{df_{ssb}}}{\frac{SSW}{df_{ssw}}}$$

& This will follow F-Distribution as you can see behind the scene

2 chi² distributions are getting divided by their df. that's how you get F-distribution

$$F = \frac{\frac{24}{2}}{\frac{52}{6}} = 1.4$$

F = 1.4

(5) Calculate P value

Use the python function

P value = 0.31

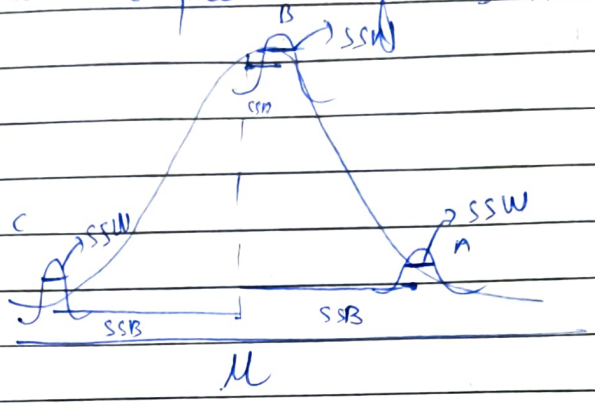
Since P value is more than α (0.05) Null Hypothesis can't be rejected.

∴ all three pop means are equal

GEOMETRIC INTUITION

Our Null hypo is $\mu_A = \mu_B = \mu_C$

→ indirectly we are saying that these all are part of same population



SSB → distance from grand mean
SSW → distance mean ~~between~~ within

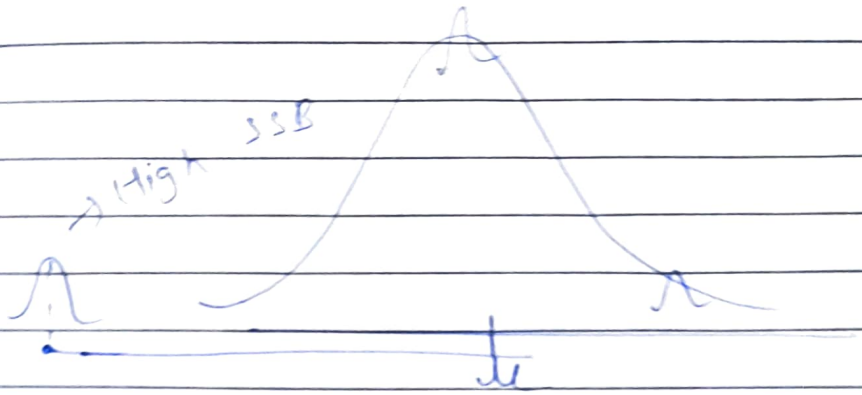
~~gg gg F-value~~

Q When will f-stat increase

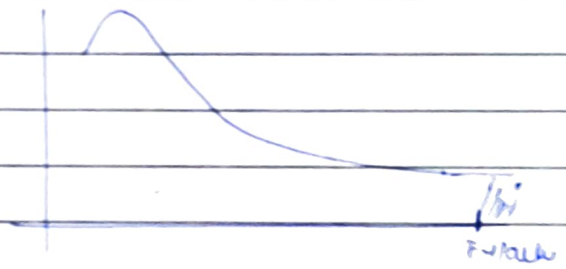
Ans → When SSW ↑ & SSB ↓

because F-stat = $\frac{\frac{SSB}{df_{SSB}}}{\frac{SSW}{df_{SSW}}}$

& if SS W increase this means distance from population mean is very high or even it may not be a part of population Ex.



∴ If high F-stat, that means in f dist, F. value will lie here

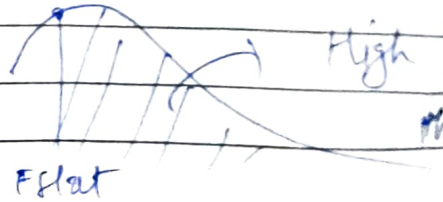


Here Pvalue is very small, hence the we can reject null hypothesis

& ~~if~~ Pv is means atleast one of the pop mean is not equal

Or we can say ~~att~~ A, B, C are not the part of same population.

if ~~P-value~~ F stat very small, the high P-value



High P-value, Hence that means we could reject null hyp