ANOVA + X^2 Test

Hypothesis Testing Framework

1) Setup the Null and Alternate Hypothesis

2) Choose the right test statistic

3) Left tailed vs Right tailed vs Two-Tailed

4) Compute P-value

5) If P- value is less than alpha , then reject the null hypothesis.

① Chi square
Test of Independance

② ANOVA ✓ Get Code
Assumption X ✓

Analysis of Variance

Amazon

observation

	F	M		
Offline	527	72	599	66%
Online	206	102	308	33%
	733	174	907	

Expected

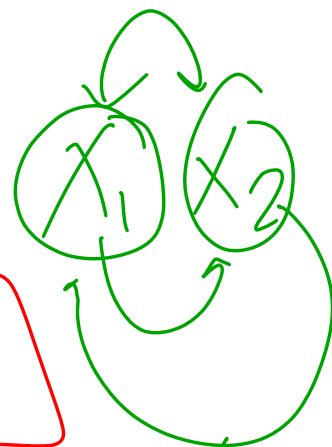|	F	M	
offline	484	115	599
Online	249	59	308
	733	174	907

66% of 733

H_0 : Gender doesn't affect (Independant)

H_a : Gender affect (dependant)

If Gender dosn't affect & we know 66% of Total people byzer offline shopping → how many females are expected to shop offlin.

Hands	Legs	Ears	Eyes	X_5	X_6
X_1	X_2	X_3	X_4		
2	2	2	2		
0	6	0	2		
0	4	2	2		

X_2

human
Cockroach
Lion

model

Linear depends

$$A_1 X_1 + A_2 X_2 + A_3 X_3$$

- - - -

$$x_1 \quad \overbrace{(x_1 \times 10)}^{x_2}$$

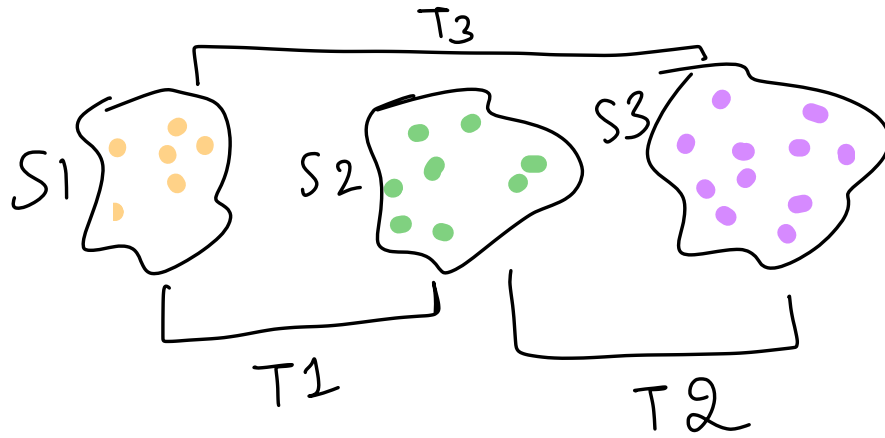| x_1	x_2
1	10
2	20
3	30
4	40
5	50
6	60

garbage

module
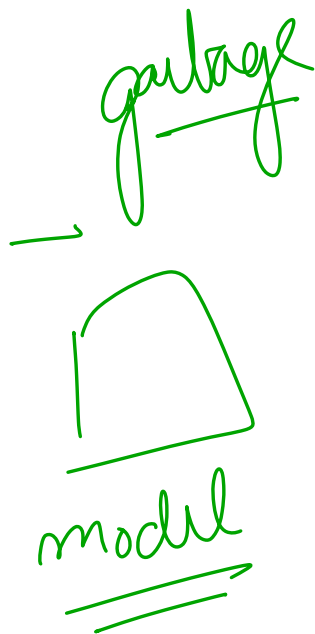
Garbage
In
Garbage
Out.

feature Selection

ANOVA
↳ Analysis of Variance

① Numerical v/s Categorical → t test
(2 Categories)

② Categorical v/s Categorical → X^2 test

③ Numerical v/s Numerical ↳ Covariance → Correlation
pearson
spearman

④ Numerical v/s Categorical → ANOVA
(> 2 Categories)

T3

S1 S2 S3

T1 T2

3 Sample

$3C_2 \quad \dfrac{3!}{1! \, 2!}$

$= \boxed{3}$

10 Sample

$10C_2 \quad \dfrac{10!}{8! \, 2!} = \boxed{45}$

① Too many T test (nC_2)

② Error compounding
Not all Ason all going to be perfect.

(I) ① American Basket ball players   80 inches   low (within)

Ha ② Indonesian College Student   64 inches

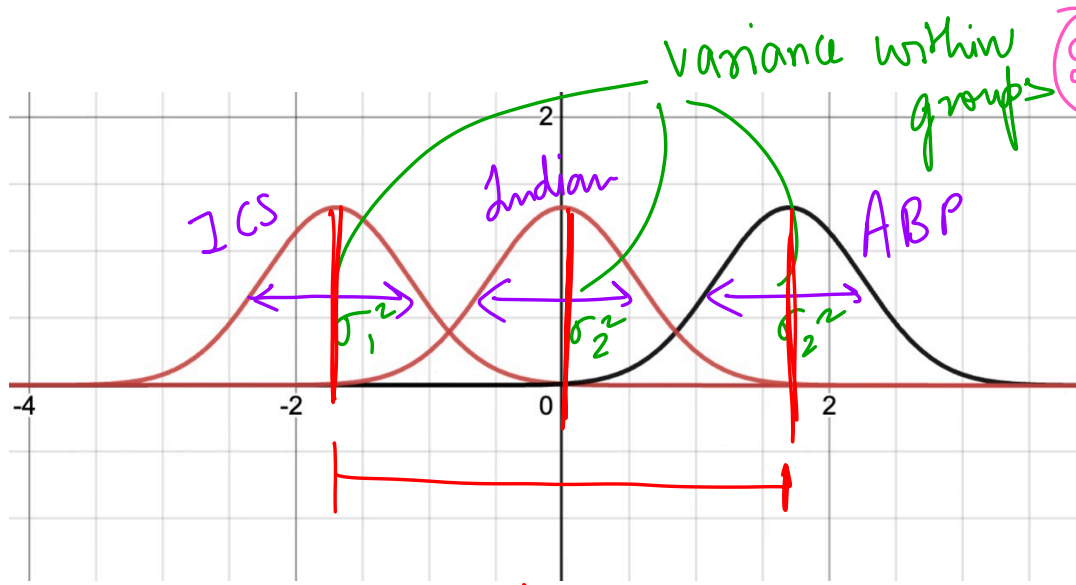③ Indian Cricket Team   70 inches
high (between)

① Variance within group

② Variance b/w groups.

(II) Sorting Alphabetically

① A - F } high (within)

② G - N } low (b/w)

③ O - Z }   Ho

$$F \, ratio = \frac{Variance \; b/w \, grps}{Variance \; within \; grp.}$$

$Ho \Rightarrow F \, ratio = \frac{low}{high}$ | $Ha \therefore F \, ratio = \frac{high}{low}$

variance within groups ② (circled)

ICS

Indian

ABP

σ_1^2 σ_2^2 σ_2^2

Var b/w grps. (4) (circled)

Setup 1

H_a

$$f \text{ ratio} = \frac{4}{2}$$
H_a

$H_0 = \quad \mu_1 = \mu_2 = \mu_3$

$H_a = \quad \mu_1 \neq \mu_2 \neq \mu_3$

Var. within grps. ⑤ setup 2
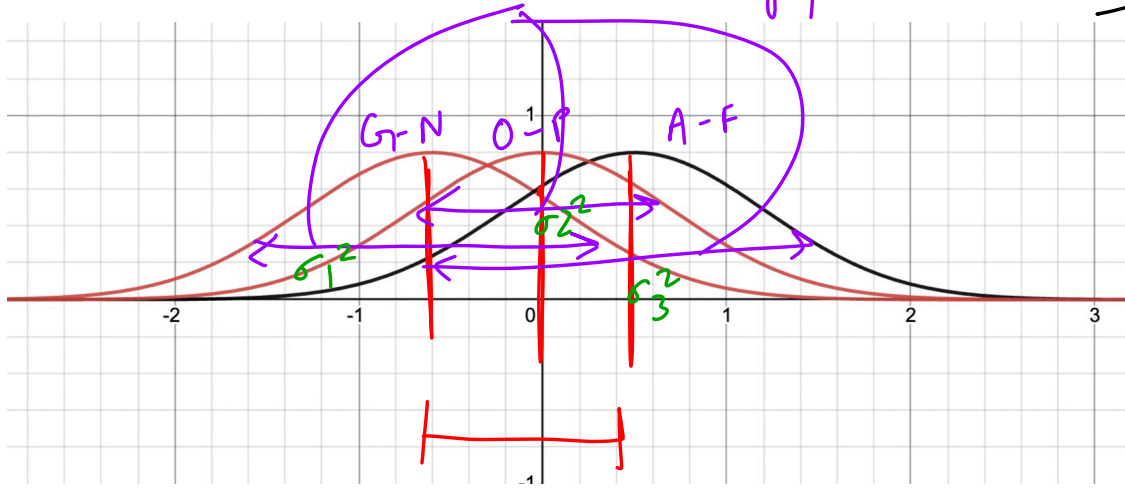
G-N   O-P   A-F

σ_1^2 σ_2^2 σ_3^2

Var b/w grps (3)

F ratio = $\frac{3}{5}$

H₀

F Distribution PDF

Legend:
- n=4, m=4
- n=10, m=4
- n=10, m=10
- n=4, m=10

Axes: Probability (y-axis), Random Variable (x-axis)
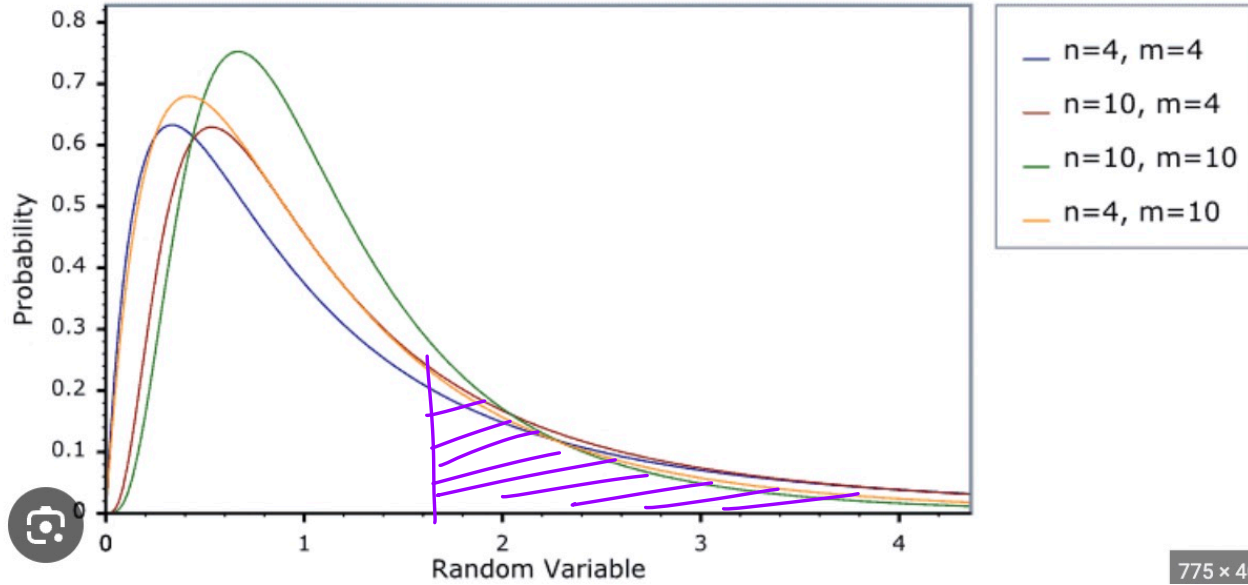
$\Rightarrow$ Right tailed test

Prob

Next Class

① Computation
② Automated
③ Business Problems

Fail to Reject Ho

Reject Ho

PDF

F-ratio