Statistics

① ANOVA (F-Test) → 1 hour }
② EDA → { Solve Some Examples. }

ANOVA : { Analysis of Variance }

ANOVA Is a statistical method used to compare the means of 2 or more group

ANOVA :

① Factors (variables) — Medicine
② Levels

Anxiety reducing

{ Dosage }

Gender }

M F

factor : Dosage
Levels : 0mg, 50mg, 100mg

0mg	50mg	100mg
9	7	4
8	6	3
7	6	2
8	7	3
8	8	

Types of ANOVA

One Way ANOVA : One factor with atleast 2 levels, levels are independent.

② Repeated Measures ANOVA — One factor with atleast 2 levels, but levels are { dependent }

Factor — Running kms

Levels

KARTIK

	Day 1	Day 2	Day 3	Day 4
1	6	10	2	10 km
2	–	–	–	–

99% Study hours of KARTIK

④ Factorial ANOVA

1% Day 1 ⟷ Day 2 ⟷ Day 3

10 hrs — 5 hrs — 2 hrs

GYM

Day 1 ⟷ Day 2 ⟷ Day 3

2 hrs (5 Eggs) — 1.5 hr (2 Eggs) — 1 hr

⑧ Factorial Anova : Two or more factor (each of which with atleast 2 levels), levels can be either independent, dependent or both (mixed)

Eg

factor (Men/Women); factor (Day 1/Day 2/Day 3)

	Day 1	Day 2	Day 3
Men	9	7	4
	8	6	3
	7	5	2
Women	8	7	3
	8	8	4
	9	7	3

One way ANOVA (F-test) $\Rightarrow$ Inferential Stats

$\Downarrow$

Comparing means of 2 or more groups

A) Researchers want to test a new anxiety medication. They split participants into 3 conditions (0mg, 50mg, 100mg), then ask them to rate their anxiety level on scale of 1-10. Are there any difference between the 3 conditions using $\alpha = 0.05$?

0mg	50mg	100mg
9	7	4
8	6	3
7	6	2
8	7	3
8	8	4
9	7	3
8	6	2

① $H_0 = \mu_{0mg} = \mu_{50mg} = \mu_{100mg}$

$H_1 =$ not all μ's are equal

② State α and C.I

$\alpha = 0.05$ C.I = 95%

③ Calculate the Degree of freedom

Statistics

$N = 21$ $n = 7$

$a = 3 \rightarrow$ {No. of levels}

$\rightarrow df_{Between} = a-1 = 3-1 = 2$

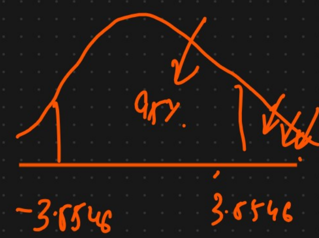
$\rightarrow df_{within} = N-a = 21-3 = 18$

$\rightarrow df_{Total} = N-1 = 21-1 = 20$

④ State Decision Rule

$df_{Between} = a-1 = 3-1 = 2$ $\{(2,18)\}$

$df_{within} = N-a = 21-3 = 18$



If F test is greater than 3.5546, Reject the Null Hypothesis

If F test is less than -3.5546 " " " "

⑤ Calculate F Test Statistics

$$F_{Test} = \frac{MS_{between}}{MS_{within}} = \frac{49.34}{0.57} =$$

	SS	df	MS	F Test
Between	98.67 →	2	49.34	86.56
Within	10.29	18	0.57	
Total	108.96	20		

$N=21 \quad n=7$

$$SS_{between} = \left[\frac{\sum(\sum a_i)^2}{n}\right] - \frac{T^2}{N}$$

$$T^2 = [57+47+21]^2 = (125)^2$$

$$\sum(\sum a_i)^2 = (9+8+7+8+8+9+8)^2 + (7+6+6+7+8+7+6)^2 + (4+3+2+3+4+3+2)^2$$

$$= 57^2 + 47^2 + 21^2$$

$$SS_{Between} = \frac{57^2+47^2+21^2}{7} - \frac{125^2}{21} = 98.67$$

② $SS_{within} = \sum y^2 - \frac{\sum(\sum a_i)^2}{n}$

P 0.75

$\alpha = 0.05$

$= \sum y^2 - \left[\frac{57^2 + 47^2 + 21^2}{7}\right] = 10.29$

$\sum y^2 \rightarrow 853$

$\sum y^2 = 9^2 + 8^2 + 7^2 + 8^2 + 8^2 + 9^2 + \ldots + 2^2 = 853$

$0.75 > 0.05$

$\Downarrow$

Accept

Final Conclusion

A 86.56 > 35.46 So we reject the Null hypothesis?

$H_0 = \mu =$ Some value

$H_1 : \mu \neq$ Some value $\rightarrow$ 95% C.I

Virginica

Petal width

→ $H_0 = \mu_{virgn} = \mu_{setosa} = \mu_{\ldots}$

$H_1 = \neq$ pvalue. $\neq$ ⇒ Reject the Null Hypothesis

$0.0228 < 0.05$

$1 - 0.025 = 0.975$

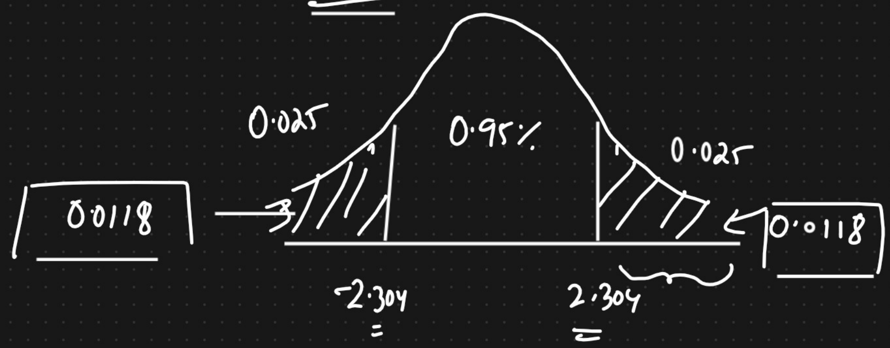
$\alpha =$

Z test

$$\begin{array}{r} 0.0118 \\ 0.0118 \\ \hline 0.0228 \end{array}$$

Z test statistics

$$Z = \frac{\bar{x} - \mu}{\sigma / \sqrt{n}}$$



$Z = 2.304$

$2.304 > 1.96$ Reject