ANOVA

a : [25 25 27 30 23 20]
b : [30 30 21 24 26 28]
c : [18 30 29 29 24 26]

$\mu_A = 25$

$\mu_B = 26.5$

$\mu_C = 26$

H_0 : No difference

$\mu_1 = \mu_2 = \mu_3$

H_a : At least 1 is different

$\mu_1 \neq \mu_2 \neq \mu_3$

Step 1 : Compute individual grps means.

Step 2 : Compute means of means of grps.

$$\bar{m} = \frac{\mu_A + \mu_B + \mu_C}{3} = \frac{25 + 26.5 + 26}{3} = 25.83$$

a : [25 25 27 30 23 20] $\mu_A = 25$

b : [30 30 21 24 26 28] $\mu_B = 26.5$

c : [18 30 29 29 24 26] $\mu_C = 24$

Step3: Compute variance b/w groups.

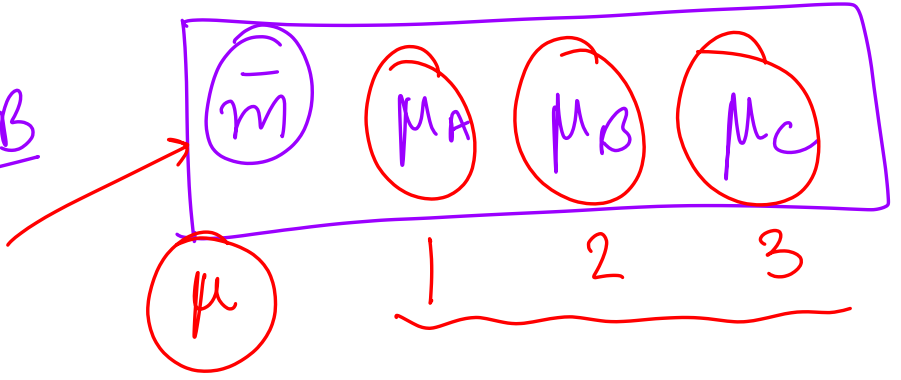
SSB. Sum of squares between $\sum n_i(\mu - \bar{m})^2$

$= 6(25 - 25.83)^2 + 6(26.5 - 25.83)^2 + 6(26 - 25.83)^2$

$= 7.0002$

$DOF_B = 2$

SSB



Step 4: MSB

Mean Sum of Squared between grps

$$MSB = \frac{SSB}{DOF_B} = 3.501$$

$$MSW = \frac{SSW}{DOF_W} = \frac{223.5}{15} = 14.9$$

Mean of Sum of Squared within groups.

a : [25 25 27 30 23 20] $\mu_A = 25$

b : [30 30 21 24 26 28] $\mu_B = 26.5$

c : [18 30 29 29 24 26] $\mu_C = 21$

Step 5: Compute variance within groups.

SSW: Sum of Squared within group.

$(25-25)^2 + (25-25)^2 + \dots + (20-25)^2 \}$ A

$(30-26.5)^2 + (30-26.5)^2 + \dots + (28-26.5)^2 \}$ B

$(18-26)^2 + (30-26)^2 + \dots + (26-26)^2 \}$ C

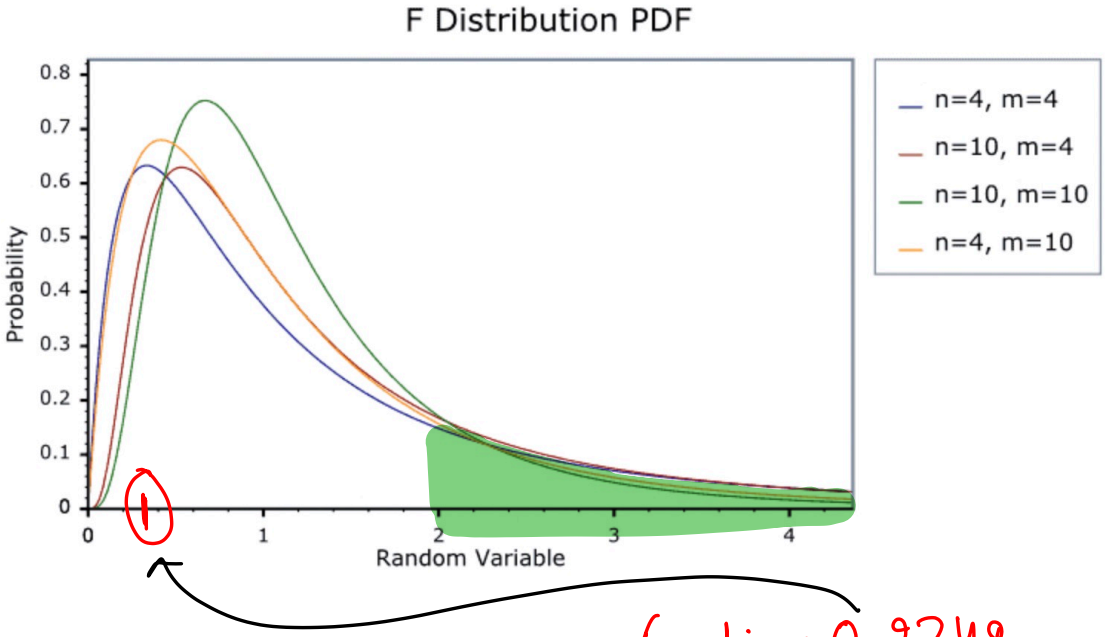
$(x_i - \mu_i)^2 \rightarrow$ SSW

223.5

$DOF_W = 15$

$$F_{ratio} = \frac{\text{Variance b/w groups}}{\text{Variance within groups}} = \frac{MSB}{MSW}$$

$$= \frac{3.501}{14.9} = 0.2348$$



Right tailed Test

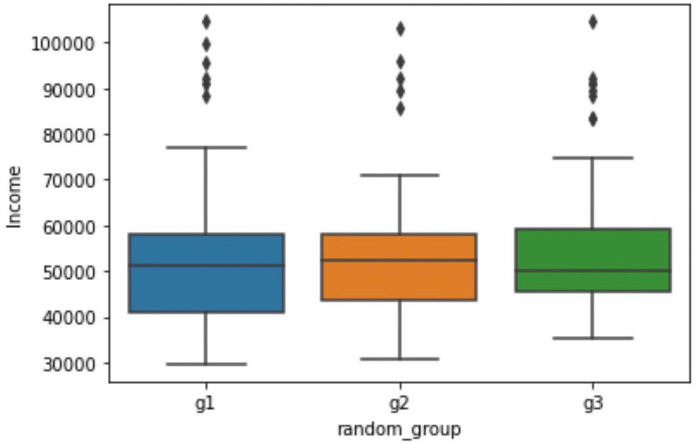
$f_{ratio} = 0.2348$

$$p\text{value} = 1 - f.cdf(0.2348, df_n = 2, df_d = 15)$$

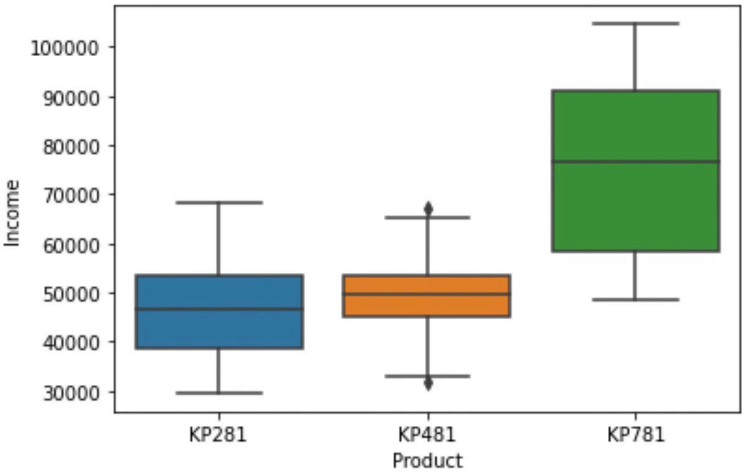
df_n : between

df_d : within

(1)



(2)



Similar / H_0 /

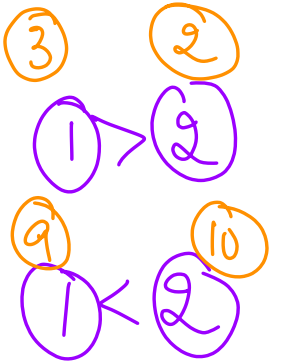
Different / H_a /

$H_0 =$ $\mu_1 = \mu_2 = \mu_3$ (Groups are similar)

$H_a =$ $\mu_1 \neq \mu_2 \neq \mu_3$ (" " not similar)

Variance Within groups.

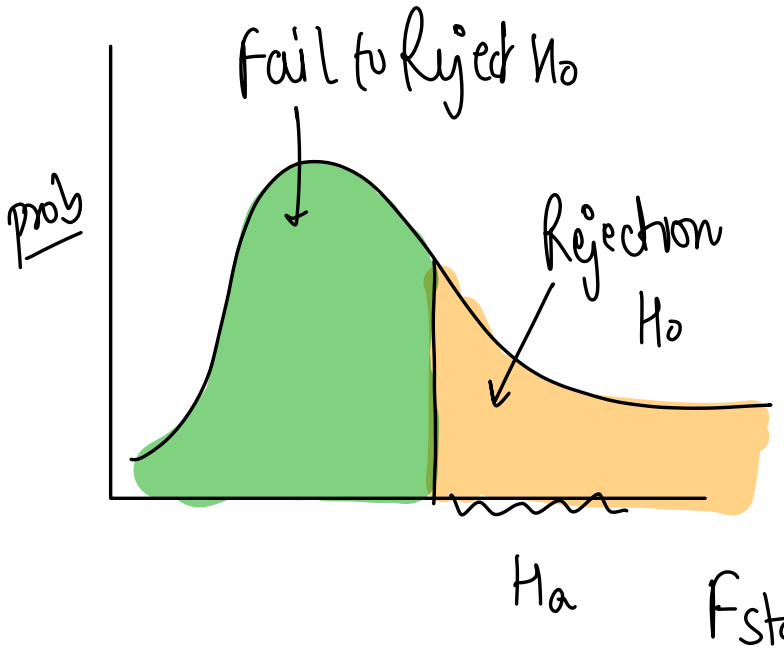
Variance b/w groups



F ratio : b/w / within

F ratio (2nd) H_a = $\frac{10}{2}$ = 5

F ratio 1st set up H_o = $\frac{9}{3}$ = 3



Right tailed test

ASSUMPTIONS

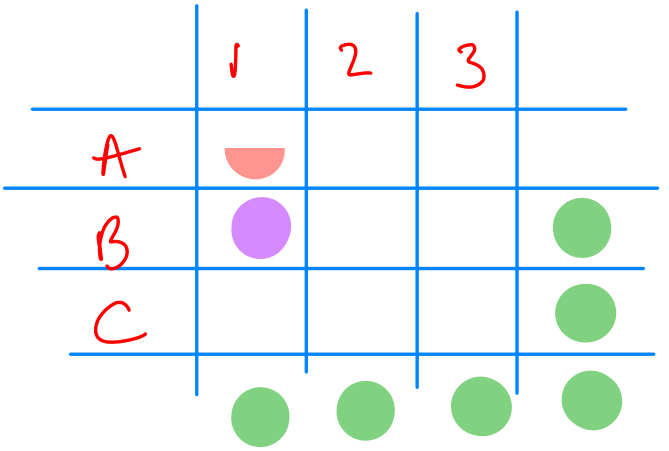
1. Data should be Gaussian
   - Visualise → Q Q Plot
   - Test → Wilkin Shapiro Test
2. Independant.
3. Equal Variances among diff group. → Levene Test

If these cond^{n} are not met / don't hold true

KRUSKAL'S TEST

Next Class

1. Gaussian → Test ← ECDFs
2. Gaussian → QQ Plot
3. Business Understanding → { ANOVA } → Feature Engineering



$(n-1)(m-1)$

$= (3-1)(3-1)$

$= 2 \times 2$

$\mu_0 = \mu_1 = \mu_2$

$\mu_a =$

ttest.sul →

$H_0 : \mu_1 = \mu_2$

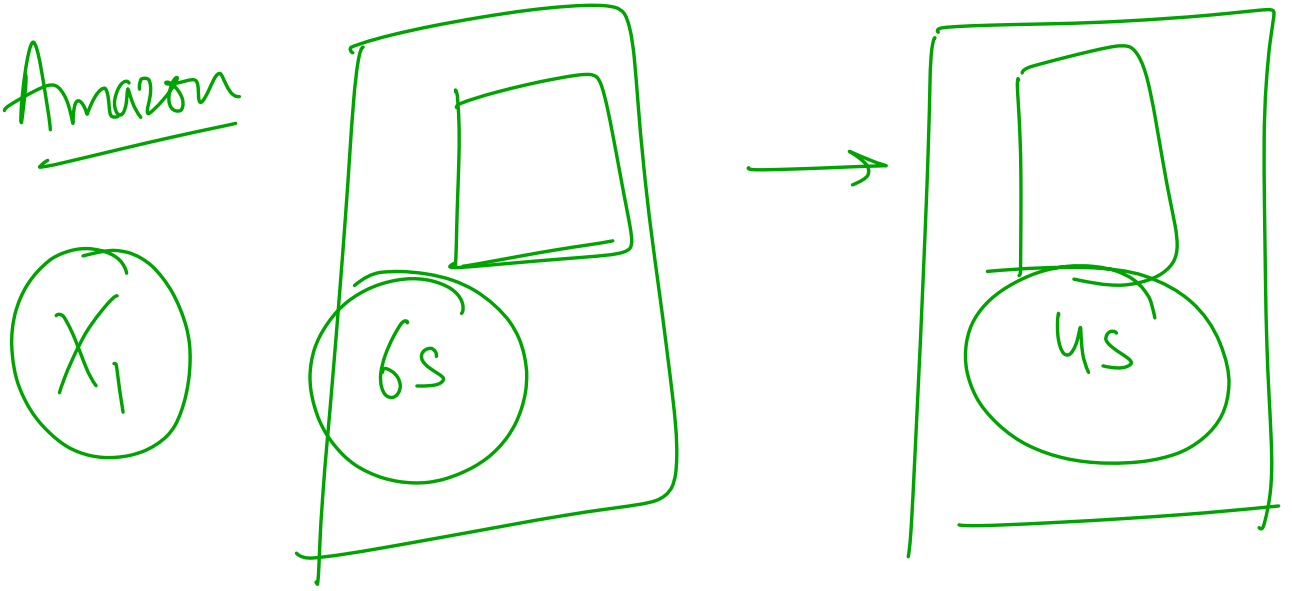
$H_0 :$

Session had no effect

Region had no effect

Region

Region


Amazon
X_1
6s
4s