Assignment on F-Test

① F-Test for variances among populations

Hyd Branch: 156 278 134 202 236 198 187 199 143 165 223

Mumbai Branch: 345 322 309 367 388 312 355 363 381

Apply F-distribution,

①	②	③	④	⑤	⑥	⑦	⑧	⑨	⑩	⑪
156	278	134	202	236	198	187	199	143	165	223

Mean = 2121/11 = 192.81

Mean-x

$(192.81-156)^2 + (192.81-278)^2 + (192.81-134)^2 + (192.81-202)^2 +$
$(192.81-236)^2 + (192.81-198)^2 + (192.81-187)^2 + (192.81-199)^2$
$+ (192.81-143)^2 + (192.81-165)^2 + (192.81-223)^2$

$= (36.81)^2 + (-85.19)^2 + (58.81)^2 + (-9.19)^2 + (-43.19)^2$
$+ (-5.19)^2 + (5.81)^2 + (-6.19)^2 + (49.81)^2 + (27.81)^2$
$+ (-30.19)^2$

$= \underset{1}{1354.97} + \underset{2}{7257.33} + \underset{3}{3458.61} + \underset{4}{84.45} + \underset{9}{2481.03} + \underset{10}{773.39} + \underset{5}{1865.37}$
$+ \underset{6}{26.93} + \underset{7}{33.75} + \underset{11}{911.43} + \underset{8}{38.31}$

$(\text{Mean}-x)^2 = 18285.57$

$\sigma^2 = 18285.57/11 = \cancel{1662.32}$

$= 18285.57/10 = 1828.557$

Mumbai Branch (2nd dataset)

$\text{Mean} = \frac{345+322+309+367+388+312+355+363+}{9}$

Mean = ~~349.11~~ 3152/9 = 350.22

$(350.22-345)^2 + (350.22-332)^2 + (350.22-309)^2 +$
$(350.22-367)^2 + (350.22-388)^2 + (350.22-312)^2 +$
$(350.22-355)^2 + (350.22-363)^2 + (350.22-381)^2$

$= (5.22)^2 + (18.22)^2 + (41.22)^2 + (-16.78)^2 + (-37.78)^2$
$+ (38.22)^2 + (-4.78)^2 + (-12.78)^2 + (-30.78)^2$

$= 27.24 + 331.96 + 1699.08 + 281.56 + 1427.32$
$+ 1460.76 + 22.84 + 163.32 + 947.40$

$= 6361.48/n-1 = 6361.48/8 = 795.185$

F-statistic = Var of 1st Dataset / var of 2nd Dataset

$= 1828.557/795.185$

$= 2.299$

$\simeq 2.30$

Null hypothesis: $\sigma^2_{hyd} = \sigma^2_{mumb}$ (same variances)

Alternate hypothesis: $\sigma^2_{hyd} \neq \sigma^2_{mumbai}$

df = 10,8

F(10,8) at $\alpha = 0.05$

$\alpha = 0.025$ in Right Tail

F-critical value is 4.2951

Obtained F-statistic $2.30 <$ F-critical value

So, we cannot reject Null Hypothesis;

Accept that variances are same for both offices Hyd & Mumbai.

(8) Line 1 210 215 205 180 175 190

Line 2 180 160 195 190 170 155

Line 3 145 170 165 160 155 175

$n = 18$, $K = 3$ (no of groups)

$df1 = K-1$, $df_2 = n-k = 18-3$

$= 3-1 = 2$ $df_2 = 15$

Line 1 → $\bar{X}_1 = 210+215+205+180+175+190/6 = 1175/6 = 195.83$

Line 2 → $\bar{X}_2 = 180+160+195+190+170+155/6 = 1050/6 = 175$

Line 3 → $\bar{X}_3 = 145+170+165+160+155+175/6 = 970/6 = 161.66$

$\bar{X} = (210+215+205+180+175+190)+(180+160+195+190+170+155)+$

$(145+170+165+160+155+175)/18$

$= 3195/18 = 177.5$

Null hypothesis: $\mu_1 = \mu_2 = \mu_3$ at $\alpha = 0.05$

Alternate Hypothesis: $\mu_1 \neq \mu_2 \neq \mu_3$

$$\text{F-statistic} = \frac{\text{Variance between Groups}}{\text{Variance within Groups}}$$

Sum of Squares Between Groups $= 6(177.5-195.83)^2 +$

$6(177.5-175)^2 + 6(177.5-161.66)^2 / 3-1$

$= 6(-18.33)^2 + 6(2.5)^2 + 6(15.84)^2 / 2$

$= 6(335.98) + 6(6.25) + 6(250.90) / 2$

$= 2015.88 + 37.5 + 1505.4 / 2$

$= 3558.78/2 = 1779.39$

Sum of Squares within groups

$(195.83-210)^2 + (195.83-215)^2 + (195.83-205)^2 +$
$(195.83-180)^2 + (195.83-175)^2 + (195.83-190)^2 +$
$(175-180)^2 + (175-160)^2 + (175-195)^2 + (175-190)^2 +$
$(175-170)^2 + (175-155)^2 +$
$(\overset{161.66}{\cancel{175}}-145)^2 + (\overset{161.66}{\cancel{175}}-170)^2 + (\overset{161.66}{\cancel{175}}-165)^2 + (\overset{161.66}{\cancel{175}}-160)^2$
$+ (\overset{161.66}{\cancel{175}}-155)^2 + (\overset{161.66}{\cancel{175}}-175)^2 / 18-3$

$= 1370.78 + 1300 + 583.3 / 15$

$= 3254.08/15$

$= 216.93$

$$F = \frac{1779.39}{216.93} = 8.202$$

f-value at $f(2,15)$ $\alpha = 0.05$ from table is 3.6823

Obtained F-value $8.202 > 3.6823$, we reject Null Hypothesis i.e all three lines of production have different variance of products in respective production lines.

(3) Given,

Pop	Population SD	Sample SD
Women	30	35
Men	50	45

7 women from women population, 12 men from men population

$$f\text{-statistic} = \frac{S_1^2 / \sigma_1^2}{S_2^2 / \sigma_2^2}$$

S1 is sample S.D of women, σ_1 is population SD of women
S2 is sample SD of men, σ_2^2 is population SD of men

$$= \frac{(35)^2 / (30)^2}{(45)^2 / (50)^2}$$

$$= \frac{1225}{900} \times \frac{2500}{2025}$$

$$f\text{-statistic} = \frac{30625}{18225} = \boxed{1.6803}$$

(4) $df_1 = 7 - 1 = 6$ To find, cumulative Probability associated
$df_2 = 12 - 1 = 11$ with f-statistic, we need numerator
f at (6, 11) and denominator degrees of freedom

From F-distribution calculator (Online)

$v_1 = 6$, $v_2 = 11$, f-value = 1.68

$P(F \le f)$ i.e $P(F \le 1.68)$ is $\boxed{0.78}$

If men is taken as first pop in numerator & women pop in denominator, we can compute f-statistic as follows.

$$f = \frac{s_1^2 / \sigma_1^2}{s_2^2 / \sigma_2^2}$$

$$= \frac{(45)^2 / (50)^2}{(35)^2 / (30)^2}$$

$v_1 = 12 - 1 = 11$

$v_2 = 7 - 1 = 6$

$$= \frac{2025}{2500} \times \frac{900}{1225}$$

$$= \frac{18225}{30625} = 0.59$$

$P(F \le 0.59)$ at $f(6, 11)$

Numerator degree of freedom

Denominator degree of freedom

From F-distribution calculator, we have

$P(F \le 0.59) = \boxed{0.21}$

(5) Given

	N	Mean	Std. Dev
midsize	31	25.8	2.56
SUV's	31	22.68	3.67
Pickup Trucks	14	21.29	2.76

Null Hypothesis: $\mu_{mid} = \mu_{suv} = \mu_{pickup}$ at $\alpha = 0.01$

Avg Highway Gas mileage same for all vehicles

Alternate Hypothesis: $\mu_{mid} \neq \mu_{suv} \neq \mu_{pickup}$

Overall mean $= (31 \times 25.8) + (31 \times 22.68) + (14 \times 21.29) / 76$

$= 799.8 + 703.08 + 298.06 / 76$

$\bar{x} = 1800.94 / 76 = 23.69$

Variance between Groups $= 31(23.69 - 25.8)^2 + 31(23.69 - 22.68)^2$

$+ 14(23.69 - 21.29)^2 / \cancel{76}\ 3 - 1$

$= 31(-2.11)^2 + 31(1.01)^2 + 14(2.4)^2 / \cancel{7}3 - 1$

$= 31(4.45) + 31(1.02) + 14(5.76) / \cancel{7}3 - 1$

$= 137.95 + 31.62 + 80.64 / \cancel{7}3 - 1$

$= 2\cancel{8}0.\cancel{8}1 / \cancel{7}\cancel{3} = 3.\cancel{4}27$

$= 250.21 / 2 = 125.105$

Sum of Squares within Groups

midsize Stddev $= 2.56$

$\sigma^2 = 6.55$

$\sigma^2 = \frac{\Sigma (x - \bar{x})^2}{n-1}$

$6.55 = \frac{\text{unknown}}{30}$

unknown1 = 6.55×30 = 196.5

SUV's StdDev = 3.67

$13.468 = \frac{\text{unknown 2}}{30}$

unknown 2 = 404.067

Pickup Trucks StdDev = 2.76

$7.617 = \frac{\text{unknown 3}}{13}$

unknown 3 = 99.021

Variance within Group = Sum of squares of individual observation from Mean for all 3 groups / 76-3

= 196.5 + 404.067 + 99.021/73

= 699.588/73

= 9.583

Computing F-statistic = $\frac{\text{Variance B/n Groups}}{\text{Variance within Groups}}$

$= \frac{125.105}{9.583}$

= 13.054

F at (2,73) with $\alpha = 0.01$ level of significance f-value is 3.12;

Obtained f-statistic 13.054 is greater than f-critical value and so we reject Null hypothesis i/e occurance at 2,73 df < 0.05; the f-value should be 3.12, but 13.05 is by chance. we conclude that Avg gas Mileage is diff for all vehicles.