(Lecture No:-12) (27/9/23)

$$\mu_1 - \mu_2 = 0$$

$$\Rightarrow \mu_1 = \mu_2$$

$$\mu_1 = \mu_2 = \mu_3 = \cdots = \mu_K$$

$$\mu_1 = \mu_2 = \mu_3 = \mu_4$$

$${}^4C_2 = 6 \text{ Possibilities}$$

(ANOVA)

It is called the Analysis of Variance.
It is used to test the equality of more than two populations means $(\mu_1 = \mu_2 = \dots = \mu_k)$
The six hypotheses testing steps are as follows:-

① $H_0: \mu_1 = \mu_2 = \cdots = \mu_K$
(All the population means are equal)

$H_1: \mu_1 \neq \mu_2 \neq \cdots \neq \mu_K$
(At least '2' population means are not equal)

② Level of significance $= \alpha$

③ Test statistic

$$F = \frac{MSTR}{MSE}$$

MSTR $\longrightarrow$ Mean Squares of Treatments

MSE $\longrightarrow$ Mean Squared Errors.

Error or Residual.

4 (Computation)

(ANOVA Table)

$\downarrow$

Source of Variation	d.f.	Sum of Squares	Mean Squares	F-ratio
Treatments	$K-1$	$SSTR$	$MSTR = \frac{SSTR}{K-1}$	$F = \frac{MSTR}{MSE}$ ( )
Error	$n-K$	SSE	$MSE = \frac{SSE}{n-K}$	—
Total	$n-1$	SST	—	—

SSTR (Sum of Squares of Treatment)

$$= \sum_{j=1}^{K} n_j (\bar{x}_j - \bar{x})^2 .$$

SSE (Sum of Squares of Errors)

$$= \sum_{j=1}^{K} (n_j - 1) S_j^2 .$$

SST (Sum of Squares of Total) $= \sum_{j=1}^{K} (x_j - \bar{x})^2$

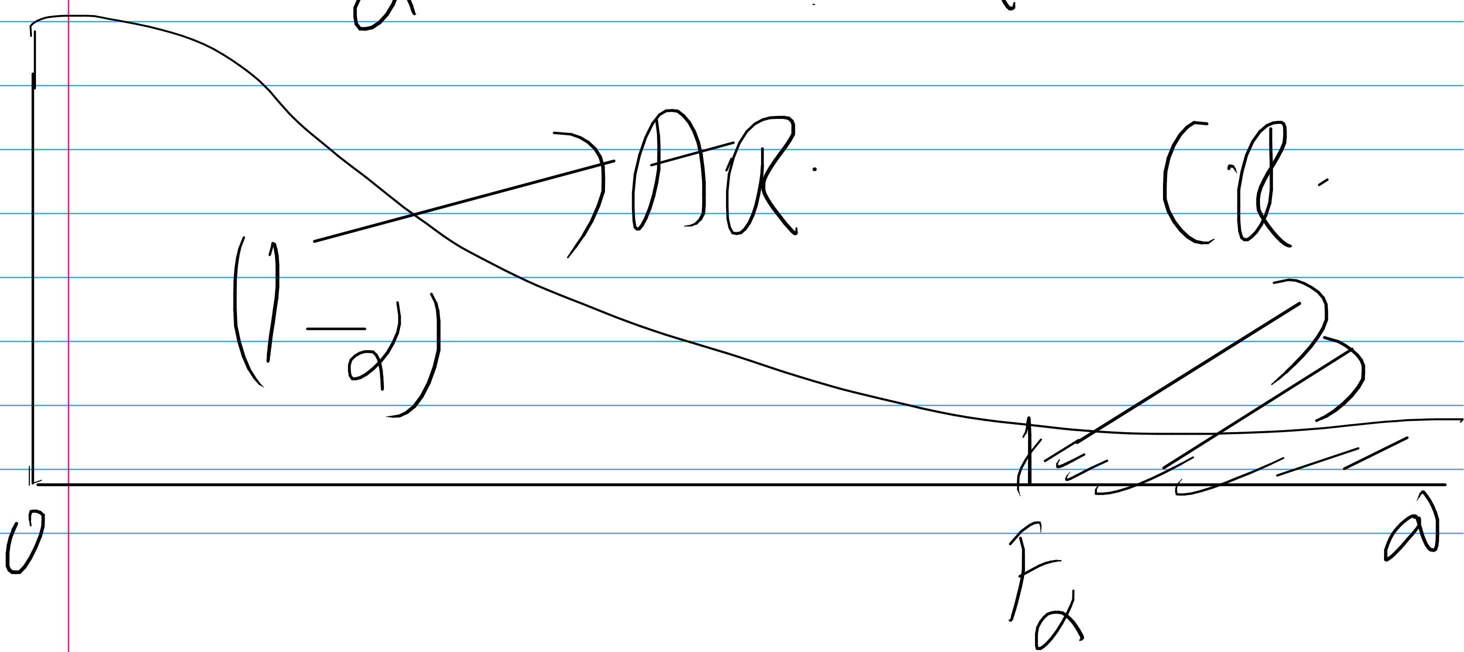
$SST = SSTR + SSE.$

5 (Critical Region)

$$F_{cal} \geq F_\alpha (v_1, v_2).$$

Where, $v_1 = K - 1.$

$v_2 = n - K.$



6 Conclusion Reject H_0 if the calculated value lies in the C.R. otherwise do not reject H_0.

Weiss, P# 726

Example 11.2

① $H_0: \mu_1 = \mu_2 = \mu_3 = \mu_4$

$H_1: \mu_1 \neq \mu_2 \neq \mu_3 \neq \mu_4$

② Level of significane

$\alpha = 0.05$.

③ Test statistic

$$F = \frac{MSTR}{MSE}.$$

④ Computation

(X_1)	(X_2)	(X_3)	(X_4)
N·E	M·W	S	W
13	15	5	8
8	10	11	10
11	16	9	6
12	11	5	5
11	13		7
	10		
$n_1 = 5$	$n_2 = 6$	$n_3 = 4$	$n_4 = 5$
$\bar{X}_1 = 11$	$\bar{X}_2 = 12.5$	$\bar{X}_3 = 7.5$	$\bar{X}_4 = 7.2$

$K = 4$

$$n = n_1 + n_2 + n_3 + n_4$$

$$n = 5 + 6 + 4 + 5 = 20$$

$\bar{x}$ (Grand mean)

$$= \frac{n_1 \bar{x}_1 + n_2 \bar{x}_2 + n_3 \bar{x}_3 + n_4 \bar{x}_4}{n_1 + n_2 + n_3 + n_4}$$

$$= \frac{(5)(11) + (6)(12.5) + 4(7.5) + 5(7.2)}{5 + 6 + 4 + 5}$$

$$\bar{x} = 9.80 \checkmark$$

$$SSTR = \underline{\sum n_j (\bar{x}_j - \bar{x})^2}$$

$$= n_1(\bar{x}_1 - \bar{x})^2 + n_2(\bar{x}_2 - \bar{x})^2$$

$$+ n_3(\bar{x}_3 - \bar{x})^2 + n_4(\bar{x}_4 - \bar{x})^2$$

$$= 5(11 - 9.8)^2 + 6(12.5 - 9.8)^2$$

$$+ 4(7.5 - 9.8)^2 + 5(7.2 - 9.8)^2$$

$$\underline{SSTR = 105.90}$$

$$SSE = \sum_{j=1}^{4}(n_j - 1)S_j^2$$

$$= (n_1 - 1)S_1^2 + (n_2 - 1)S_2^2 + (n_3 - 1)S_3^2$$

$$+ (n_4 - 1)S_4^2.$$

$$S_1^2 = \frac{1}{n_1 - 1}\left(\sum x_1^2 - \frac{(\sum x_1)^2}{n_1}\right).$$

$$S_2^2 = \frac{1}{n_2 - 1}\left(\sum x_2^2 - \frac{(\sum x_2)^2}{n_2}\right).$$

$$S_3^2 = \frac{1}{n_3 - 1}\left(\sum x_3^2 - \frac{(\sum x_3)^2}{n_3}\right).$$

$$S_4^2 = \frac{1}{n_4 - 1}\left(\sum x_4^2 - \frac{(\sum x_4)^2}{n_4}\right).$$

$S_1^2 = \underline{3.5}$ ✓, $S_2^2 = 6.7$ ✓

$S_3^2 = 9.6$ ✓, $S_4^2 = 3.7$ ✓

$$SSE = (n_1-1)S_1^2 + (n_2-1)S_2^2 + (n_3-1)S_3^2 + (n_4-1)S_4^2$$

$$= (6-1)(3.5) + (5-1)(6.7) + (4-1)(9) + (5-1)(3.7)$$

$$SSE = 89.30$$

$$MSTR = \frac{SSTR}{K-1} = \frac{105.90}{4-1}$$

$$MSTR = 35.3$$

$$MSE = \frac{SSE}{n-K} = \frac{89.3}{20-4}$$

$$MSE = 5.581$$

ANOVA Table

S.O.V.	d.f.	SS	MS	F-ratio
Treatments	$K-1 = 4-1 = 3$	105.90	35.3	$F = \frac{MSTR}{MSE} = 35.3 / 5.581$ $F = 6.32$
Error	$n-K = 20-4 = 16$	89.3	5.581	
Total	$n-1 = 20-1$ 19	195.2	–	–

$F_{cal} = \underline{6.32}$

5 $\underline{\underline{(\text{Critical Region})}}$

$F_{cal} \geq F_{\alpha}(V_1, V_2)$

$V_1 = K - 1 = 4 - 1 = 3$

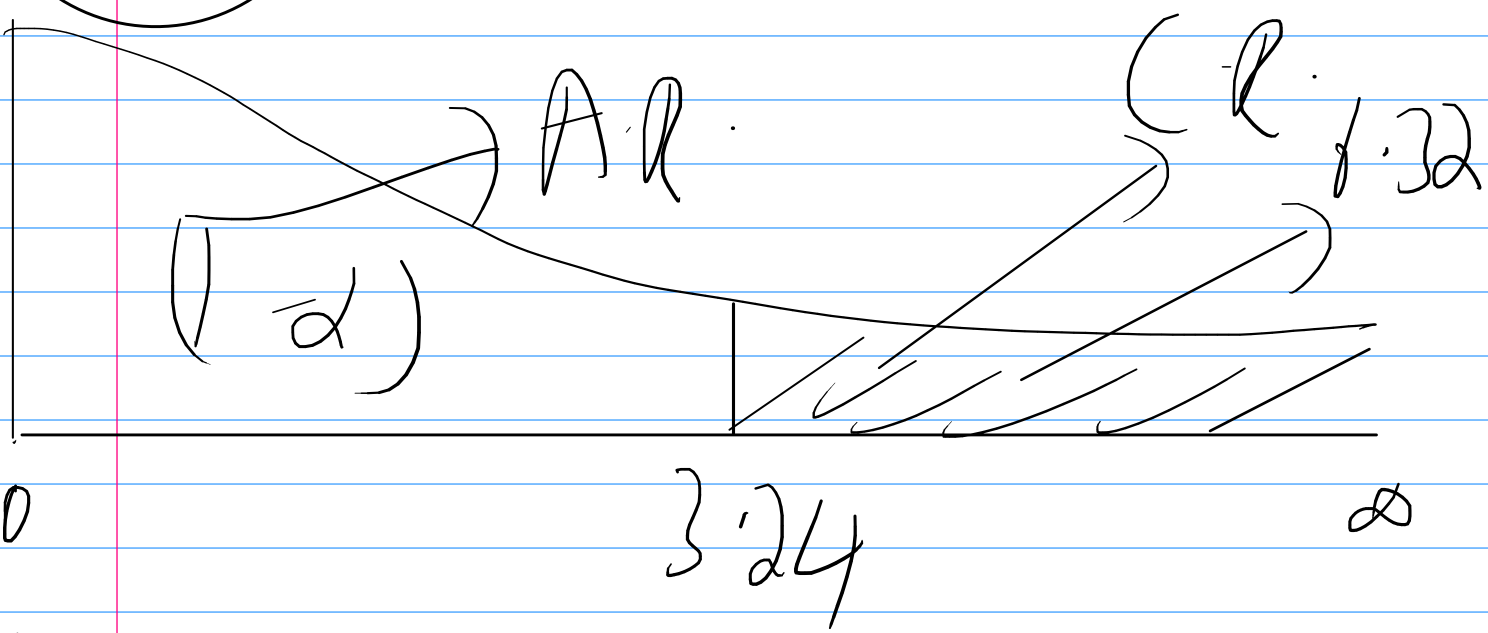
$V_2 = n - K = 20 - 4 = 16$

$F_{cal} \geq F_{0.05}(3, 16)$

$6.32 \geq F_{0.05}(3, 16)$

v_1 v_2

6.327 3.24



7 Conclusion Reject H_0.

Therefore, the average energy consumption in 4 regions is not equal.

ANOVA Assumptions

1. We assume that the populations variances are equal.
2. Treatments are applied randomly and Independently to the subject.
3. Errors are

normally distributed

Anderson, P# 614

Exercises 9, 10 and

12 P45