UNBALANCED COMLETELY RANDOMIZED DESIGNS
UNBALANCED CRD

Recall that an experimental design is called unbalanced if the sample sizes for the treatment combinations are **not all equal**.

Reasons why you may need to be able to work with unbalanced designs:

1. Balanced designs produce unbalanced data when **something goes wrong**. (e.g., plants die, machinery breaks down, shipments of raw materials don't come in on time, subjects get sick, etc.)
2. **Some treatments** may be more **expensive** or more **difficult** to run than others.
3. Some treatment combinations may be of **particular interest**, so the experimenter chooses to sample more heavily from them.

Example:

The following data represent the shoot dry weight (g) of wheat plant in a glass house experiment after spraying with four types of insecticide to control a certain pest. Complete analysis of variance table and decide if the insecticides used significantly differ or not?

	A	B	C	D
r1	2	1.7	2	2.1
r2	2.2	1.9	2.4	2.2
r3	1.8	1.5	2.7	2.2
r4	2.3		2.5	1.9
r5	1.7		2.4	
∑ti	∑ A= 10	∑ B= 5.1	∑ C= 12	∑ D= 8.4

Step 1 Write the hypotheses to be tested.

Ho: µ1 = µ2 =………… = µk or There is no significant difference between the treatments

Step 2. Statistical analysis:

Df t= t-1 = 4-1 =3

Df E = ∑ri – t = 5 +3 +5+ 4= 17 – 4= 13

∑ri T = ∑ri – 1= 17-1= 16

$$(C.F) = \frac{GT2}{\sum ri} = \frac{(35.5)2}{5+3+5+4} = \mathbf{74.132}$$

$\boldsymbol{Total\ Sum\ of\ Squares\ (TSS) = \sum Yij2 - C.F = (2)2 + (2.2)2 + (1.8)2 + \cdots . + (1.9)2 - 74.132 = 75.77 - 74.132 = 1.638}$

$$Treatmnent\ Sum\ of\ squares\ (t\ SS) = \frac{\sum ti\ 2}{r} - C.F = \frac{\sum A\ 2}{rA} + \frac{\sum B\ 2}{rB} + \frac{\sum C\ 2}{rC} + \frac{\sum D\ 2}{rD} - C.F$$

$$= \left(\frac{(10)2}{5} + \frac{(5.1)2}{3} + \frac{(12)2}{5} + \frac{(8.4)2}{4}\right) - 74.132 = 0.978$$

$$Error\ sum\ of\ squares\ (E\ SS) = TSS - t\ SS = 1.638 - 0.978 = \mathbf{0.66}$$

$$Mean\ Square\ of\ Treatment\ (MS\ t) = \frac{t\ SS}{Df\ t} = \frac{0.978}{3} = \mathbf{0.326}$$

$$Mean\ square\ of\ Eeeor\ (MS\ E) = \frac{ESS}{DfE} = \frac{0.66}{13} = \mathbf{0.051}$$

1. Calculated – F:

$$Calculated\ F = \frac{MS\ t}{MS\ E} = \frac{0.326}{0.051} = 6.932$$

S. O. V.	Df	SS	MS	Calculate F	Tabulated F
Treatment	3	0.978	0.326	6.932 **	**$F_{0.05;\ 3,\ 13}$ = 3.41**
Error	13	0.66	0.051		
Total	16	1.638			**$F_{0.01;\ 3,\ 13}$ = 5.74**

- Since F-calculated (6.932) > F tabulated (3.41) under level of significant (0.05) we reject null hypothesis; Ho: µ1 = µ2 =………… = µk.

- Since F-calculated (6.932) > F tabulated (5.74) under level of significant (0.01) we reject null hypothesis; Ho: µ1 = µ2 =………… = µk.

Post-ANOVA Comparison of Means

The analysis of variance method is a useful and powerful tool to compare several treatments means. In comparing k treatments, the null hypothesis tested is that the true means are all equal (HO: µ1 = µ2 = … = µ k). If a significant F test is found, one accepts the alternative hypothesis which merely states that they are not all equal; (HA: µ1 ≠ µ2 ≠ … ≠ µ k). Further comparisons to determine which treatments are different can be carried out by using so-called multiple comparison procedures.

1. Fisher's Least Significant Difference (LSD)

Fisher (1935) described a procedure for pairwise comparisons called the least significant difference (LSD) test. This test is to be used only if the hypothesis that all means are equal is rejected by the overall F test. If the overall test is significant, a procedure analogous to ordinary Student's t test is used to test any pair of means. If the overall F ratio is not significant, no further tests are performed. When it is used, the two treatments will be declared different if the absolute difference between two samples means.

When the number of treatment replications is equal the following equation used:

$$LSD = t\ (\alpha\ and\ df\ E) * \sqrt{\frac{2MSE}{r}}$$

When the number of treatment replications is un- equal (un balanced designs) the following equation used:

$$LSD = t\ (\alpha and\ df\ E) * \sqrt{MSE\ (\frac{1}{r1} + \frac{1}{r2} + \frac{1}{r3} + \frac{1}{rn})}$$

Construct comparison table as follow:

- Rank the means from largest to smallest,
- Create table rows beginning with the largest mean and going through the next-to-smallest mean,
- Create table columns starting with the smallest mean and going through the next-to-largest mean,
- Compute the absolute difference between each row and column intersection/mean.
- Two means are significantly different if there difference equal and greater than the LSD value for e.g. $\bar{A} - \bar{B} \geq$ **LSD**

2. Dennett's Test

In statistics, Dunnett's test is a multiple comparison procedure developed by Canadian statistician Charles Dunnett to compare each of a number of treatments with a single control. Multiple comparisons to a control are also referred to as many-to-one comparisons.

Dennett's test was devised for the situation when:

1. one condition (e.g., the control condition) is to be compared against all other conditions (e.g., the treatment conditions)
2. No other pairwise comparisons are required.

$$Calculated\ Dunett\ value = Dunett's\ critical\ value\ (\alpha, dft\ and\ dfE) \sqrt{\frac{2MSE}{r}}$$

$$Calculated\ Dunett\ value = D\ critical\ value\ (\alpha, dft\ and\ dfE) * \sqrt{MSE\ (\frac{1}{r1} + \frac{1}{r2} + \frac{1}{r3} + \frac{1}{rn})}$$