

## CHAPTERS 8, 9, and 15.2 (and some concepts of Ch. 14) Analysis of Variance and the Completely Randomized Design

### I. Introduction

An experiment is called a **completely randomized design** when a random sample of observations is taken for each of  $t$  different populations.

#### Advantage

1. The design is extremely easy to construct.
2. The design is easy to analyze even though the sample sizes might not be the same for each treatment.
3. The design can be used for any number of treatments.

#### Disadvantage

1. Although the completely randomized design can be used for any number of treatments, it is best suited for situations in which there are relatively few treatments.
2. The experimental units to which treatments are applied must be as homogeneous as possible. Any extraneous sources of variability will tend to inflate the error term, making it more difficult to detect differences among the treatment means.

#### Assumptions

1. The samples are independent random samples.
2. Each sample is selected from a normal population.
3. The mean and variance for population  $i$  are, respectively,  $\mu_i$  and  $\sigma^2$  ( $i=1,2,\dots,t$ ).

### Statistical Model for a Completely Randomized Design

The  $j$  th sample observation from population  $i$ , can be expressed as the sum of three terms.

$$y_{ij} = \mu + \alpha_i + \varepsilon_{ij}$$

The term  $\mu$  denotes an overall mean that is an unknown constant.

The term  $\alpha_i$  denotes an effect due to population  $i$ ;  $\alpha_i$  is an unknown constant.

The term  $\varepsilon_{ij}$  denotes a random error associated with the  $j$  th observation from population  $i$ . We assume that  $\varepsilon_{ij}$  is independent and normally distributed with a mean of 0 and a variance  $\sigma^2$ .

## II. Hypothesis Test on Means for special case when all sample sizes are equal to n.

Comments:

- a. If assumption 3 is valid, then an estimate of  $\sigma^2$  may be found by pooling the sample variances for each population. This estimate is:

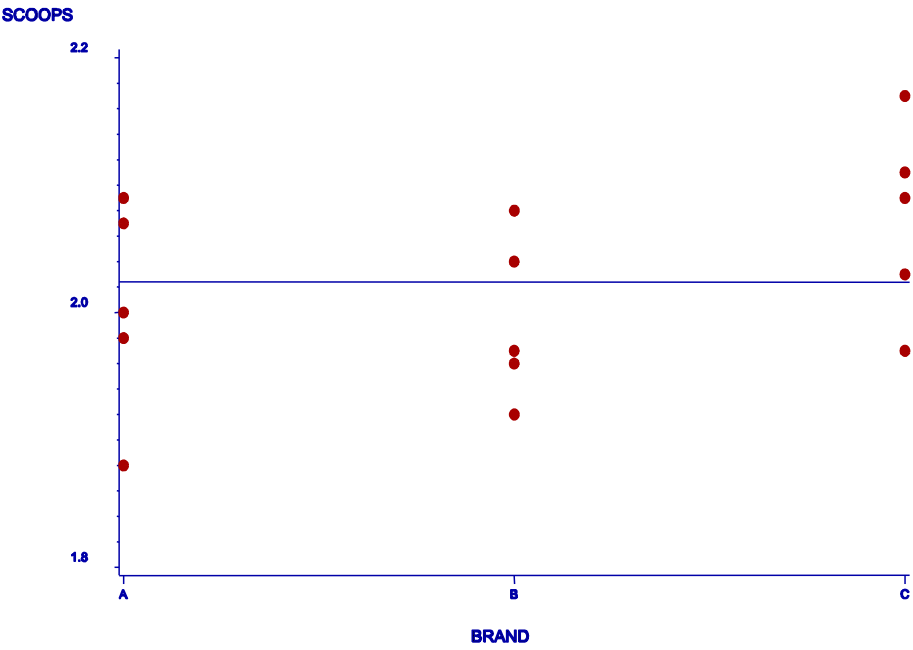
$$s_w^2 = \frac{(n_1 - 1)s_1^2 + (n_2 - 1)s_2^2 + \dots + (n_t - 1)s_t^2}{n_1 + n_2 + \dots + n_t - t}$$

The value  $s_w^2$  is called the **within-sample variance**.

- b. If all of the population means are equal to the same value  $\mu$ , then the sample means all come from the same sampling distribution. The variance for the sampling distribution is given by  $\sigma^2/n$ .
- c. If all of the population means are equal and the variance of the sample means is calculated, then another estimate for  $\sigma^2$  would be  $s_B^2 = n \times (\text{variance of the sample means})$ . The value  $s_B^2$  is called the **between-samples variance**.
- d. If all of the populations means are not equal, then  $s_B^2$  will, on average, be larger than  $\sigma^2$ .
- e. Comments a and c give two estimates of the same value of  $\sigma^2$  when all population means are equal. So the ratio  $F_{\text{obs}} = s_B^2 / s_w^2$  will have an F-distribution with  $df_1 = t - 1$  and  $df_2 = nt - t$ .
- f. If all of the populations means are not equal, then the  $F_{\text{obs}}$  value will be large, and so this is an upper tail test.

Example: Random samples from three brands of raisin bran were taken to determine if there is a difference in the mean number of scoops among the three brands. Below are the sample data. (Use a significance level of .05.)

Brand	Number of Scoops					Sample Mean	Sample Variance
A	2.09	2.07	1.98	1.88	2.00	2.004	.00693
B	2.08	1.92	2.04	1.96	1.97	1.994	.00418
C	2.17	2.09	2.11	1.97	2.03	2.074	.00588



## Hypothesis Test for the Raisin Bran Example

### Hypotheses

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

$H_a$ : Not all the means are equal.

### Within-Sample Variance:

$$s_w^2 = \frac{(5-1)(.00693) + (5-1)(.00418) + (5-1)(.00588)}{5+5+5-3} = .00566$$

Variance of Sample Means: 0.0019

Between-Samples Variance:  $s_B^2 = 5(.0019) = .00950$

Test Statistic:  $F_{\text{obs}} = .00950/.00566 = 1.68$

Rejection Region: Reject  $H_0$  if  $F_{\text{obs}} > F_{.05,2,12}=3.89$

Decision: Fail to Reject  $H_0$

Conclusion: Insufficient evidence, at a .05 significance level, to conclude that the population means are different.

### III. General Hypothesis Test for the Comparison of More than Two Population Means

#### 1. Hypotheses

$$H_0: \mu_1 = \mu_2 = \dots = \mu_t$$

$H_a$ : Not all of the above equal.

#### 2. ANOVA Calculations: **Analysis of Variance (ANOVA) for a Completely Randomized Design**

Source Due to	Sum of Squares (SS)	df	Mean Square (MS)	F
Between samples	SSB	t-1	$s_B^2 = SSB/(t-1)$	$F = s_B^2/s_W^2$
Within samples	SSW	$n_T - t$	$s_W^2 = SSW/(n_T - t)$	
Totals	TSS	$n_T - 1$		

#### 3. Rejection Region

It is an F-distribution with  $df_1 = t - 1$  and  $df_2 = n_T - t$ .

#### 4. Notation

$y_{ij}$ : The  $j$ th sample observation selected from population  $i$ . For example,  $y_{23}$  denotes the third sample observation drawn from population 2.

$n_i$ : The number of sample observations selected from population  $i$ .

$n_T$ : The total sample size:  $n_T = \sum n_i$ .

$y_{i.}$ : The sum (total) of the sample measurements obtained from population  $i$ .

$y_{..}$ : The sum (grand total) of all sample observations:  $y_{..} = \sum y_{i.}$ .

$\bar{y}_{i.}$ : The average of the  $n_i$  sample observations drawn from population  $i$ .

$\bar{y}_{..}$ : The average of all sample observations.

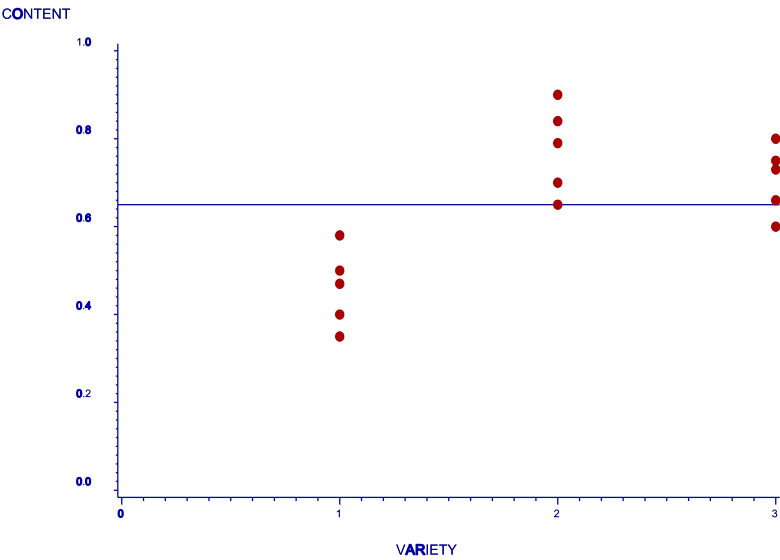
$$TSS = \sum_{ij} y_{ij}^2 - \frac{y_{..}^2}{n_T}$$

$$SSB = \sum_i \frac{y_{i.}^2}{n_i} - \frac{y_{..}^2}{n_T}$$

$$SSW = TSS - SSB$$

**Example:** A horticulturist was investigating the phosphorus content of tree leaves from three different varieties of apple trees (1, 2, and 3). Random samples of five leaves from each of the three varieties were analyzed for phosphorus content. The data are given in the table.

Variety	Phosphorus Content				
1	.35	.40	.58	.50	.47
2	.65	.70	.90	.84	.79
3	.60	.80	.75	.73	.66



Calculations for the Phosphorus Content Example

Variety	Phosphorus Content					y <sub>i</sub>
1	.35	.40	.58	.50	.47	2.30
2	.65	.70	.90	.84	.79	3.88
3	.60	.80	.75	.73	.66	3.54

$$TSS = .37444$$

$$SSB = \frac{2.30^2}{5} + \frac{3.88^2}{5} + \frac{3.54^2}{5} - \frac{(2.30 + 3.88 + 3.54)^2}{15} = .27664$$

$$SSW = .37444 - .27664 = .0978$$

**Analysis of Variance (ANOVA)  
for a Completely Randomized Design**

Source Due to	Sum of Squares (SS)	df	Mean Square (MS)	F
Between samples	.27664	2		
Within samples	.0978	12		
Totals	.37444	14		

#### IV. Checking the Equal Variance Assumption

Hypotheses

$$H_0: \sigma_1^2 = \sigma_2^2 = \dots = \sigma_t^2$$

$H_a$ : Not all the variances are equal.

Test Statistic

$$F_{\max} = \frac{s_{\max}^2}{s_{\min}^2}$$

Rejection Region

Table 12 for  $\alpha$ ,  $df_1=t$ , and  $df_2=n-1$ , where  $n$  is the number of observations in each sample.

Checking Equality of Variances for the Phosphorus Example

Hypotheses

$$H_0: \sigma_1^2 = \sigma_2^2 = \sigma_3^2$$

$H_a$ : Not all of the variances equal

Test Statistic

$$F_{\max} = .01033/.00617 = 1.674$$

Rejection Region

Reject  $H_0$  if  $F_{\max} > F_{\max, .05, 3, 4} = 15.5$ .

Decision

Fail to reject  $H_0$ .

Conclusion

There is insufficient evidence, at the .05 significance level, to conclude that the variances among the varieties are different.



## V. Multiple Comparison Procedures

### Fisher's Least Significant Difference (LSD) Procedure and Tukey's W Procedure

1. In some instances there are no clearly defined hypotheses of interest prior to the experiment or survey. In these cases the only hypotheses to evaluate may be  $H_0: \mu_i = \mu_j$  which can be rewritten as  $H_0: \mu_i - \mu_j = 0$ , where  $i$  and  $j$  indicate all pairs of treatments.

For the Phosphorous Content Example the hypotheses would be

$$H_0: \mu_1 - \mu_2 = 0$$

$$H_0: \mu_1 - \mu_3 = 0$$

$$H_0: \mu_2 - \mu_3 = 0$$

(Note 1: This set of hypotheses is said to be **all pairwise comparisons** of the means)

(Note 2: These hypotheses should be tested **only** if the ANOVA F-test for the hypothesis  $H_0: \mu_1 = \mu_2 = \dots = \mu_t$  leads to rejection of the hypothesis.)

2. One popular procedure for determining differences in means for all pairwise comparisons is called **Fisher's LSD**. The amount by which a pair of means must differ to be deemed significantly different is called **Fisher's LSD value** and the formula is

$$LSD = t_{\alpha/2, (n_T - t)} \sqrt{s_w^2 \left( \frac{1}{n_i} + \frac{1}{n_j} \right)}$$

(The standard deviation reduces to a simple formula and is exactly the same for all pairwise comparisons when sample sizes are equal for each treatment.)

3. All pairs of means are then compared. If  $|\bar{y}_i - \bar{y}_j| \geq LSD$ , we conclude that the populations means  $\mu_i$  and  $\mu_j$  are different
4. The probability for a Type I error for each of the individual hypothesis tests or comparisons is  $\alpha$  when the LSD value is used to make the decision. However, the **experiment wise error rate** across all hypothesis tests or pairwise comparisons is greater than  $\alpha$  when the LSD value is used. To control the experiment wise error rate, **Tukey's W value** can be used for decisions about the hypotheses. Tukey's W value is similar to Fisher's LSD value except that a q-value from Table 10 in the text is substituted for the t-value. The formula for Tukey's W value is:

$$W = \frac{q_\alpha(t, v)}{\sqrt{2}} * \sqrt{s_w^2 \left( \frac{1}{n_i} + \frac{1}{n_j} \right)}$$

where  $v = n_T - t$ .

For the Phosphorous Content Example, a table with both the Fisher's LSD values and the Tukey's W values would be

<b><u>Hypothesis</u></b>	<b><u>Point Estimate</u></b>	<b><u>LSD value</u></b>	<b><u>Dec</u></b>	<b><u>W value</u></b>	<b><u>Dec</u></b>
$H_0: \mu_1 - \mu_2 = 0$	-0.316	0.1244	R $H_0$	0.1521	R $H_0$
$H_0: \mu_1 - \mu_3 = 0$	-0.248	0.1244	R $H_0$	0.1521	R $H_0$
$H_0: \mu_2 - \mu_3 = 0$	0.068	0.1244	FTR $H_0$	0.1521	FTR $H_0$

(Note 1: The t-value is 2.179, the calculated standard error is 0.05709, and the q-value is 3.77 for this example)

(Note 2: The decisions do not change in this example whether LSD or Tukey's is the chosen technique.)