CHAPTER 15 – Section 3 (and some concepts of Ch. 14)

Randomized Block Design

I. Introduction

An experiment is called a **randomized block design** when comparing t treatments in b blocks. The blocks consist of t homogeneous experimental units. Treatments are randomly assigned to experimental units within a block, with each treatment appearing exactly once in every block.

<u>Advantages</u>

1. The design is useful for comparing t treatment means in the presence of a single extraneous source of variability.
2. The statistical analysis is simple.
3. The design is easy to construct.

<u>Disadvantages</u>

1. Because the experimental units within a block must be homogeneous, the design is best suited for a relatively small number of treatments.
2. This design controls for only one extraneous source of variability (due to blocks). Additional extraneous sources of variability tend to increase the error term, making it more difficulty to detect treatment differences.
3. The effect of each treatment on the response must be approximately the same from block to block.

Randomized Block Design Layout

	Block			
Treatment	1	2	...	b
1	y_{11}	y_{12}	...	y_{1b}
2	y_{21}	y_{22}	...	y_{2b}
...	...	...	...	...
t	y_{t1}	y_{t2}	...	y_{tb}

Example: Ten plots of land are available to determine if a new genetically-engineered variety of corn will increase the yield. The researcher plans to plant five plots with the genetically-engineered variety (the treated group) and five plots with the same variety that has not been altered (the control group). Two plots of land are available for each of five regions of the state. We might expect that the yield would be different for the various regions of the state. In order to reduce the influence of the region-to-region variability on the treatment comparison, it seems logical to pair the plots by region. We would then plant the control group variety in one plot and the genetically-engineered variety in the other plot. Below are the results of the experiment.

	Region				
Treatment	Clemson	Aiken	Rock Hill	Florence	Charleston
Control	64.3	101.4	70.6	84.7	62.6
Genetically-Engineered	70.0	105.6	70.3	90.0	65.8

<u>The ANOVA Model for a Randomized Block Design</u>

Observations for a Randomized Block Design can be expressed as the sum of three terms.

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

where

- μ an overall mean that is an unknown constant
- α: an effect due to treatment i; α_i is an unknown constant
- β_j: an effect due to block j; β_j is an unknown constant
- ε_{ij}: a random error associated with the response on treatment i, block j.

Assumptions about the ε_{ij}'s

1. The ε_{ij}'s are independent.
2. The ε_{ij}'s have a mean of 0 and a variance of σ_ε^2.
3. The ε_{ij}'s are normally distributed.

II. Hypothesis Tests for the Randomized Block Design

1. Hypotheses (Treatments)

H_0: $\alpha_1 = \alpha_2 = ... = \alpha_t = 0$
H_a: Not all of the above equal to zero.

Hypotheses (Blocks)

H_0: $\beta_1 = \beta_2 = ... = \beta_b = 0$
H_a: Not all of the above equal to zero.

2. ANOVA Calculations

Analysis of Variance (ANOVA) for a Randomized Block Design

Source Due to	**Sum of Squares (SS)**	**df**	**Mean Square(MS)**	**F**
Treatments	SST	t-1	MST = SST/(t-1)	MST/MSE
Blocks	SSB	b-1	MSB = SSB/(b-1)	MSB/MSE
Error	SSE	(b-1)(t-1)	MSE = SSE/(b-1)(t-1)	
Totals	TSS	bt-1		

3. Rejection Region (Treatments)

It is an F-distribution with df_1=t-1 and df_2=(b-1)(t-1).

Rejection Region (Blocks)

It is an F-distribution with df_1=b-1 and df_2=(b-1)(t-1).

4. Notation

y_{ij}: observation for treatment i in block j.

t: number of treatments.

b: number of blocks.

$y_{i.}$: total for observations receiving treatment i.

$y_{.j}$: total for observations in block j.

$y_{..}$: total of all sample observations.

$\bar{y}_{i.}$: sample mean for treatment i ($y_{i.}/b$).

$\bar{y}_{.j}$: sample mean for block j ($y_{.j}/t$).

$\bar{y}_{..}$: overall sample mean ($y_{..}/bt$).

$$TSS = \sum_{i,j} y_{ij}^2 - \frac{y_{..}^2}{bt}$$

$$SST = \sum_{i} \frac{y_{i.}^2}{b} - \frac{y_{..}^2}{bt}$$

$$SSB = \sum_{j} \frac{y_{.j}^2}{t} - \frac{y_{..}^2}{bt}$$

$$SSE = TSS - SST - SSB$$

Calculations for the Genetically-Engineered Corn Example

	Region					
Treatment	Clemson	Aiken	Rock Hill	Florence	Charleston	Sum
Control	64.3	101.4	70.6	84.7	62.6	383.6
Genetically-Engineered	70.0	105.6	70.3	90.0	65.8	401.7
Sum	134.3	207.0	140.9	174.7	128.4	785.3

$$TSS = 2247.141$$

$$SST = \frac{383.6^2 + 401.7^2}{5} - \frac{616696.09}{10} = 32.761$$

$$SSB = \frac{134.3^2 + 207.0^2 + 140.9^2 + 174.7^2 + 128.4^2}{2} - \frac{616696.09}{10} = 2202.866$$

$$SSE = 2247.141 - 32.761 - 2202.866 = 11.514$$

Analysis of Variance (ANOVA)
for a Randomized Block Design

Source Due to	**Sum of Squares (SS)**	**df**	**Mean Square(MS)**	**F**
Treatments	32.761	1	32.761	11.381
Blocks	2202.866	4	550.716	191.320
Error	11.514	4	2.878	
Totals	2247.141	9		

III. Multiple Comparison Procedures of Treatment Means for Randomized Complete Block Designs

The Linear Contrast, Fisher's LSD, and Tukey's W procedures follow the methods discussed for the Completely Randomized Design except for the following differences: b replaces n for sample size per treatment, MSE replaces S^2_W in the standard deviation formula, and (b-1)(t-1) replaces n_T-t as error degrees of freedom for looking up the t-value or q-value from a table. Note that in this example no mean comparison procedure is necessary.

IV. Relative Efficiency

To determine if using a randomized block design has increased the precision in comparing treatment means, we compare the standard errors between the randomized block design and the completely randomized design.

The standard error of a mean for the completely randomized design is given by the value MSE_{CR}/r, where r is the number of replications. An estimate of the MSE_{CR} may be obtained from the ANOVA table of the randomized block design and is given by

$$MSE_{CR} = \frac{(b-1)MSB + b(t-1)MSE_{RB}}{(bt-1)}$$

The standard error of a mean for the randomized block design is given by the value MSE_{RB}/b.

If the two designs have the same precision then

$$\frac{MSE_{CR}}{r} = \frac{MSE_{RB}}{b} \quad or \quad \frac{MSE_{CR}}{MSE_{RB}} = \frac{r}{b} = 1$$

The ratio MSE_{CR}/MSE_{RB} is called the **relative efficiency**. If the relative efficiency is greater than one, then the randomized block design has increased the precision in comparing treatment means. The relative efficiency also tells us approximately how many times more observations it would take for the completely randomized design to have the same precision as the randomized block design.

The relative efficiency for the genetically-engineered corn experiment is

$$\frac{MSE_{CR}}{MSE_{RB}} = \frac{(b-1)MSB + b(t-1)MSE_{RB}}{(bt-1)MSE_{RB}}$$

$$= \frac{(5-1)(550.716) + 5(2-1)(2.878)}{(5*2-1)2.878}$$

$$= 85.6$$