# Design and Analysis of Experiments

## Lecture 5 - THE RANDOMIZED COMPLETE BLOCK DESIGN

Ahmed A.(Msc)

DIRE DAWA UNIVERSITY
COLLEGE OF NATURAL AND COMPUTATIONAL SCIENCES
DEPARTMENT OF SATATISTICS

October 29, 2024

# Outline

# The Randomized Complete Block Design

- In any experiment, variability arising from a nuisance factor can affect the results.
- Generally, a nuisance factor is defined as a design factor that likely affects the response, but whose effect is not of primary interest.
- Sometimes, a nuisance factor is unknown and uncontrolled, meaning we may not be aware of its existence, and it may change levels during the experiment.
- Randomization is the design technique used to protect against such "lurking" nuisance factors.
- In other cases, the nuisance factor is known but uncontrollable. Observing the value of the nuisance factor at each experiment run allows us to adjust for it using analysis of covariance

# The Randomized Complete Block Design

- When the source of variability is known and controllable, blocking can systematically remove its effect on statistical comparisons among treatments.
- Blocking is crucial in industrial experimentation and is the main topic of this lecture.

**Example:**

Reconsider the hardness testing experiment (lecture 2). Suppose we want to determine if four different tips yield different readings on a hardness testing machine. Such an experiment might be part of a quality control study.

# RCBD...

**Randomized Complete Block Design for the Hardness Testing Experiment**

| Test Coupon (Block) | | | |
|---|---|---|---|
| **1** | **2** | **3** | **4** |
| Tip 3 | Tip 3 | Tip 2 | Tip 1 |
| Tip 1 | Tip 4 | Tip 1 | Tip 4 |
| Tip 4 | Tip 2 | Tip 3 | Tip 2 |
| Tip 2 | Tip 1 | Tip 4 | Tip 3 |

**Figure:** Table 5.1

# The Machine Operation

The machine operates by pressing the tip into a metal test coupon, and from the depth of the resulting depression, the hardness of the coupon can be determined.

# The Experimental Design

- The experimenter has decided to obtain four observations on Rockwell C-scale hardness for each tip.
- There is only one factor—tip type—and a completely randomized single-factor design would consist of randomly assigning each one of the $4 \times 4 = 16$ runs to an **experimental unit**, that is, a metal coupon, and observing the hardness reading that results.
- 16 different metal test coupons would be required in this experiment, one for each run in the design.

# The Problem with Completely Randomized Design

- If the metal coupons differ slightly in their hardness, as might happen if they are taken from ingots that are produced in different heats, the experimental units (the coupons) will contribute to the variability observed in the hardness data.
- The experimental error will reflect *both* random error *and* variability between coupons.
- We would like to make the experimental error as small as possible by removing the variability between coupons from the experimental error.

# The Randomized Complete Block Design (RCBD)

- The design requires the experimenter to test each tip once on each of four coupons.
- The blocks, or coupons, form a more homogeneous experimental unit on which to compare the tips.
- This design strategy improves the accuracy of the comparisons among tips by eliminating the variability among the coupons.
- Within a block, the order in which the four tips are tested is randomly determined.
- The RCBD is a generalization of the paired $t$-test concept.

# Applications of RCBD

- Units of test equipment or machinery are often different in their operating characteristics and would be a typical blocking factor.
- Batches of raw material, people, and time are also common nuisance sources of variability in an experiment that can be systematically controlled through blocking.
- Blocking may also be useful in situations that do not necessarily involve nuisance factors, such as testing the effect of catalyst feed rate on polymer viscosity.
- Blocks can be used to test the **robustness** of a process variable to conditions that cannot be easily controlled.

# RCBD...

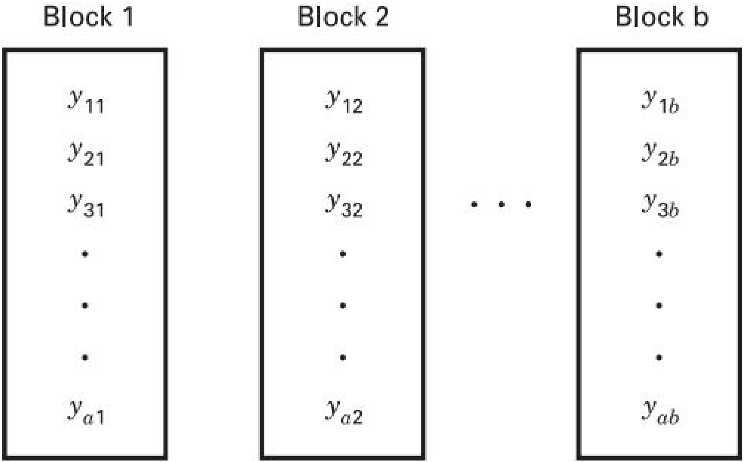


**Figure:** 5.1 The randomized complete block design

# The Randomized Complete Block Design

- $a$ treatments are compared in $b$ blocks
- One observation per treatment in each block
- Order of treatments within each block is randomly determined
- Blocks represent a **restriction on randomization**

# The Statistical Model

The traditional **effects model** for the RCBD is:

$$y_{ij} = \mu + \tau_i + \beta_j + \epsilon_{ij} \begin{cases} i = 1, 2, \ldots, a \\ j = 1, 2, \ldots, b \end{cases}$$

where:

- $\mu$ is the overall mean
- $\tau_i$ is the effect of the $i$th treatment
- $\beta_j$ is the effect of the $j$th block
- $\epsilon_{ij}$ is the random error term

We initially consider treatments and blocks as fixed factors.

# Overspecified Model

Just as in the single-factor experimental design model, the effects model for the RCBD is an overspecified model. Consequently, we usually think of the treatment and block effects as deviations from the overall mean, such that:

$$\sum_{i=1}^{a} \tau_i = 0$$

$$\sum_{j=1}^{b} \beta_j = 0$$

# Means Model

It is also possible to use a **means model** for the RCBD, where:

$$y_{ij} = \mu_{ij} + \epsilon_{ij} \begin{cases} i = 1, 2, \ldots, a \\ j = 1, 2, \ldots, b \end{cases}$$

However, the effects model will be used throughout this chapter.

# Treatment Means in the RCBD

The $i$th treatment mean in the RCBD is:

$$\mu_i = \frac{1}{b} \sum_{j=1}^{b} (\mu + \tau_i + \beta_j)$$

which simplifies to:

$$\mu_i = \mu + \tau_i$$

# Hypotheses in Terms of Treatment Effects

The hypotheses of interest for the RCBD are:

$$H_0 : \mu_1 = \mu_2 = \cdots = \mu_a$$
$$H_1 : \text{at least one } \mu_i \neq \mu_j$$

Since $\mu_i = \mu + \tau_i$, an equivalent way to write the hypotheses is in terms of the treatment effects:

$$H_0 : \tau_1 = \tau_2 = \cdots = \tau_a = 0$$
$$H_1 : \text{at least one } \tau_i \neq 0$$

# Random Blocks

If the blocks are considered random factors, the statistical model becomes:

$$y_{ij} = \mu + \tau_i + \beta_j + \epsilon_{ij}$$

where $\beta_j \sim \text{NID}(0, \sigma_\beta^2)$ and $\epsilon_{ij} \sim \text{NID}(0, \sigma^2)$.

- The F-test for treatments is $\frac{\text{MST}}{\text{MSB}}$.
- The expected mean squares are more complex in this case.

# Notation and Definitions

Let:

- $y_{i.}$ be the total of all observations taken under treatment $i$
- $y_{.j}$ be the total of all observations in block $j$
- $y_{..}$ be the grand total of all observations
- $N = ab$ be the total number of observations

Also, let:

$$\bar{y}_{i.} = \frac{y_{i.}}{b} \qquad \text{(average of treatment } i\text{)}$$

$$\bar{y}_{.j} = \frac{y_{.j}}{a} \qquad \text{(average of block } j\text{)}$$

$$\bar{y}_{..} = \frac{y_{..}}{N} \qquad \text{(grand average)}$$

# Notation and Definitions

The notation $y_{i.}$ represents the total of all observations taken under treatment $i$. This can be expressed as:

$$y_{i.} = \sum_{j=1}^{b} y_{ij}$$

$$y_{.j} = \sum_{i=1}^{a} y_{ij}$$

$$y_{..} = \sum_{i=1}^{a} \sum_{j=1}^{b} y_{ij}$$

# Total Corrected Sum of Squares

The total corrected sum of squares can be expressed as:

$$SST = \sum_{i=1}^{a}\sum_{j=1}^{b}(y_{ij} - \bar{y}_{..})^2 \tag{1}$$

Expanding this expression, we get:

$$SST = \sum_{i=1}^{a}\sum_{j=1}^{b}(y_{ij} - \bar{y}_{i.} - \bar{y}_{.j} + \bar{y}_{..})^2 \tag{2}$$

# ANOVA Equation for RCBD

Simple algebra shows that the three cross-product terms are zero, so we can write:

$$SST = SS_{\text{Treatments}} + SS_{\text{Blocks}} + SS_{\text{Error}} \tag{3}$$

Expressing the sums of squares symbolically:

$$SST = b\sum_{i=1}^{a}(\bar{y}_{i.} - \bar{y}_{..})^2 + a\sum_{j=1}^{b}(\bar{y}_{.j} - \bar{y}_{..})^2 + \sum_{i=1}^{a}\sum_{j=1}^{b}(y_{ij} - \bar{y}_{i.} - \bar{y}_{.j} + \bar{y}_{..})^2 \tag{4}$$

# Statistical Testing in RCBD

- The sums of squares for Treatments (SSTreatments), Blocks (SSBlocks), and Error (SSE) divided by their degrees of freedom are independently distributed as chi-square random variables.
- Each sum of squares divided by its degrees of freedom is a mean square.
- The expected values of the mean squares, assuming fixed treatments and blocks, are:

$$E(\mathrm{MSE}) = \boldsymbol{\sigma}^2$$

$$E(\mathrm{MSBlocks}) = \boldsymbol{\sigma}^2 + \frac{a}{b-1}\sum_{j=1}^{b} \tau_j^2$$

$$E(\mathrm{MSTreatments}) = \boldsymbol{\sigma}^2 + \frac{b}{a-1}\sum_{i=1}^{a} \beta_i^2$$

# Testing Equality of Treatment Means:

- To test if treatment means are equal, use the test statistic:

$$F_0 = \frac{\text{MSTreatments}}{\text{MSE}}$$

which follows an $F$ distribution with $(a-1)$ and $(a-1)(b-1)$ degrees of freedom if the null hypothesis is true.

- Reject $H_0$ if $F_0 > F_{\alpha,a-1,(a-1)(b-1)}$. Alternatively, a P-value approach can be used.

# Comparing Block Means:

- To test the hypothesis $H_0 : \tau_j = 0$, the statistic

$$F_0 = \frac{\text{MSBlocks}}{\text{MSE}}$$

can be compared to $F_{\alpha,b-1,(a-1)(b-1)}$.

- Note that randomization is applied only within blocks; thus, blocks impose a restriction on randomization.
- Box, Hunter, and Hunter (2005) support using the F-test based on randomization, while Anderson and McLean (1974) argue that this restriction makes $F_0$ less meaningful for block mean comparisons.

# Practical Considerations in Using F-Test

- Due to potential non-normality, using $F_0 = \text{MSBlocks}/\text{MSE}$ as an exact test for block means is generally not advisable.
- However, examining the ratio MSBlocks/MSE is reasonable to gauge the effect of blocking.
- A large ratio implies that blocking reduces noise and likely improves precision in treatment comparisons.
- An ANOVA table, as in Table 4.2, typically summarizes this procedure. Calculations are often done using statistical software.

**Sums of Squares Computation:**

- The sums of squares in the ANOVA can be derived from:

$$y_{ij} - y_{..} = (y_{i.} - y_{..}) + (y_{.j} - y_{..}) + (y_{ij} - y_{i.} - y_{.j} + y_{..})$$

# RCBD...

**TABLE 4.2**

**Analysis of Variance for a Randomized Complete Block Design**

| Source of Variation | Sum of Squares | Degrees of Freedom | Mean Square | $F_0$ |
|---|---|---|---|---|
| Treatments | $SS_{\text{Treatments}}$ | $a-1$ | $\frac{SS_{\text{Treatments}}}{a-1}$ | $\frac{MS_{\text{Treatments}}}{MS_E}$ |
| Blocks | $SS_{\text{Blocks}}$ | $b-1$ | $\frac{SS_{\text{Blocks}}}{b-1}$ | |
| Error | $SS_E$ | $(a-1)(b-1)$ | $\frac{SS_E}{(a-1)(b-1)}$ | |
| Total | $SS_T$ | $N-1$ | | |

# Computing Sums of Squares in RCBD

- Sums of squares (SS) can be computed efficiently using spreadsheet software like Excel:
  - Each observation can be listed in a column.
  - Square and sum each column to calculate the respective sum of squares.

# Computing Sums of Squares in RCBD

Alternatively, SS formulas can be expressed in terms of treatment and block totals:

**Formulas:**

- Total sum of squares:

$$\text{SST} = \sum_{i=1}^{a}\sum_{j=1}^{b} y_{ij}^2 - \frac{y_{..}^2}{N}$$

- Treatment sum of squares:

$$\text{SSTreatments} = \frac{1}{b}\sum_{i=1}^{a} y_{i.}^2 - \frac{y_{..}^2}{N}$$

- Block sum of squares:

$$\text{SSBlocks} = \frac{1}{a}\sum_{j=1}^{b} y_{.j}^2 - \frac{y_{..}^2}{N}$$

# Computing Sums of Squares in RCBD

Error sum of squares:

$$\text{SSE} = \text{SST} - \text{SSTreatments} - \text{SSBlocks}$$

**Notes:**

- Here, $y_{i.}$ and $y_{.j}$ represent treatment and block totals, and $y_{..}$ is the grand total.
- $N = a \times b$ is the total number of observations.

# Example RCBD

■ **TABLE 4.3**

**Randomized Complete Block Design for the Vascular Graft Experiment**

| Extrusion Pressure (PSI) | Batch of Resin (Block) 1 | 2 | 3 | 4 | 5 | 6 | Treatment Total |
|---|---|---|---|---|---|---|---|
| 8500 | 90.3 | 89.2 | 98.2 | 93.9 | 87.4 | 97.9 | 556.9 |
| 8700 | 92.5 | 89.5 | 90.6 | 94.7 | 87.0 | 95.8 | 550.1 |
| 8900 | 85.5 | 90.8 | 89.6 | 86.2 | 88.0 | 93.4 | 533.5 |
| 9100 | 82.5 | 89.5 | 85.6 | 87.4 | 78.9 | 90.7 | 514.6 |
| Block Totals | 350.8 | 359.0 | 364.0 | 362.2 | 341.3 | 377.8 | $y_{..} = 2155.1$ |

# Performing ANOVA in RCBD

- To conduct the analysis of variance, calculate the following sums of squares.
- Total sum of squares (SST):

$$\text{SST} = \sum_{i=1}^{4}\sum_{j=1}^{6} y_{ij}^2 - \frac{(2155.1)^2}{24} = 480.31$$

- Treatment sum of squares (SSTreatments):

$$\text{SSTreatments} = \frac{1}{6}\left((556.9)^2 + (550.1)^2 + (533.5)^2 + (514.6)^2\right) - \frac{(2155.1)^2}{24} = 1$$

- Block sum of squares (SSBlocks):

$$\text{SSBlocks} = \frac{1}{4}\left((350.8)^2 + (359.0)^2 + \cdots + (377.8)^2\right) - \frac{(2155.1)^2}{24} = 192.25$$

- Error sum of squares (SSE):

$$\text{SSE} = \text{SST} - \text{SSTreatments} - \text{SSBlocks} = 480.31 - 178.17 - 192.25 = 109.89$$

# RCBD Results

**TABLE 4.4**

**Analysis of Variance for the Vascular Graft Experiment**

| Source of Variation | Sum of Squares | Degrees of Freedom | Mean Square | $F_0$ | $P$-Value |
|---|---|---|---|---|---|
| Treatments (extrusion pressure) | 178.17 | 3 | 59.39 | 8.11 | 0.0019 |
| Blocks (batches) | 192.25 | 5 | 38.45 | | |
| Error | 109.89 | 15 | 7.33 | | |
| Total | 480.31 | 23 | | | |