Randomized Block Design with Sampling

- Sometimes we may have more than one observation per treatment-block combination.

- **Within each block**, we have a sample of $n \geq 2$ observations having the same treatment.

- Model equation for RBD with sampling:

  - $\varepsilon_{ij}$ was **experimental error** → measures variation among units having the same treatment (across the collection of blocks) \[\text{var}(\varepsilon_{ij}) = \sigma^2\]

  - $\delta_{ijk}$ is **sampling error** → measures variation among units having the same treatment within the same block \[\text{var}(\delta_{ijk}) = \sigma_{\delta}^2\]

- In this situation, we must look carefully at the Expected MS to choose the appropriate denominator for our F-statistic.

- Assuming treatment effects are fixed and block effects are random:
Testing for treatment effects:

Recall $H_0$:

- If $H_0$ is true, then which two Mean Squares have the same expected value?

- Appropriate test statistic is:

$$ F^* = \quad \text{Reject } H_0 \text{ if:} $$

- What is the test statistic for testing $H_0$: $\sigma^2 = 0$?

$$ F^* = \quad \text{Reject } H_0 \text{ if:} $$

- What is the test statistic for testing $H_0$: $\sigma^2_{\beta} = 0$?

$$ F^* = \quad \text{Reject } H_0 \text{ if:} $$

- What is the test statistic for testing $H_0$: $\sigma^2 = 0$?

$$ F^* = \quad \text{Reject } H_0 \text{ if:} $$
**Example:** Experiment on stretching ability (Table 10.6, p. 474)

Response = stretching ability of rubber material
Treatments = 7 materials (A, B, C, D, E, F, G)
Blocks = 13 lab sites

• At each lab, there were \( n = 4 \) units for each type of material.

\( n = 4, \, t = 7, \, b = 13 \rightarrow \) total of observations overall.

• Is there a significant difference in mean stretching ability among the seven materials?

• We test:

\[ F^* = \]

Compare to

Software gives P-value:

• Reject \( H_0 \) and conclude there is a significant difference in mean stretching ability among the seven materials.
● Which of the materials are significantly different in terms of mean stretching ability?

● Can use Tukey multiple comparisons procedure (experimentwise error rate $\alpha = 0.05$).

Results from software:

Latin Square Designs

● Sometimes we may have two blocking factors.

**Example:** Suppose we are comparing tire performance across four tire brands (label them A, B, C, D).
• The blocking factors are Car (1, 2, 3, 4) and Tire Position (1, 2, 3, 4).

• If we make each car/position combination a block, we have 16 blocks → we need 64 tires (inefficient and costly!)

• What if we only have 16 tires for the experiment?

A Poor Arrangement:

• Here, the value of car as a blocking factor is lost.

• Each car has only one brand of tire.
A Better Arrangement:

- Now each car gets each brand of tire and each position gets each brand of tire.

- This design is called a **Latin Square**.

- Each row and each column contains each treatment **once and only once**.

- A $t \times t$ Latin Square is used for an experiment for $t$ treatments and **two** blocking factors:
  - Row factor with $t$ levels
  - Column factor with $t$ levels
Formal Linear Model for Latin Square:

**Note:** In a Latin Square design, there is assumed to be no interaction!

**Example** (Table 10.4): Experiment to study the effect of music type on employee productivity

- **Treatments:** A = rock & roll, B = country, C = easy listening, D = classical, E = none.

- **Row factor levels:** 5 times of day (9-10, 10-11, 11-12, 1-2, 2-3)

- **Column factor levels:** 5 days of week (Mon, Tue, Wed, Thu, Fri)
A $5 \times 5$ Latin Square is:

- Each music type appears once on each day and once at each time of day.

- Testing for a significant effect of music type on mean productivity:

  \[ F^* = \]

  - There is a significant difference in mean productivity among the five music types.

  - **Note:** There is also a significant row effect (time of day) and a significant column effect (day of week).
• Specifically, which music types are significantly different?

• Using Tukey’s procedure, we see:

Summary:
• Main advantage of a Latin Square design: Efficiency – can perform useful tests with relatively few experimental units.

• Main disadvantage: cannot test for any interaction.