Two-Way ANOVA
Part 1: Definitions and Models

STAT 705: Regression and Analysis of Variance

Two-Way ANOVA

- Two factors, each with multiple levels
- We consider factorial treatment structures
  - Each level of one factor is combined with each level of the other factor
  - Treatments are all possible combinations of the factor levels
- Example
  - Bake cookies from 3 different recipes (R1, R2, R3) at 2 different oven temperatures (T1, T2)
  - Six treatments total
    - (T1, R1), (T1, R2), (T1, R3)
    - (T2, R1), (T2, R2), (T2, R3)

Descriptive Analysis

- A descriptive statistical analysis of a two factor study includes the following steps
  1. Put the data in an appropriate (spreadsheet) format
  2. Obtain the means of the treatments
  3. Arrange the means in a two-way table according to the two factors
  4. Obtain a two-way plot of the means
  5. Interpret the results
- Calculations are usually carried out using statistical software

Example

An experiment was conducted to study the effects of treating fabric with inorganic salts on the flammability of fabric. Two application levels (concentrations) and three salts were used, and a vertical burn test was used on three specimens of cloth for each salt and level combination. The response variable is the temperature at which the fabric specimen ignites.

	Salt		
Level	**Untreated**	**$CaCO_3$**	**$CaCl_2$**
1	812, 827, 876	733, 728, 720	725, 727,719
2	945, 881, 919	786, 771, 779	756, 781, 814

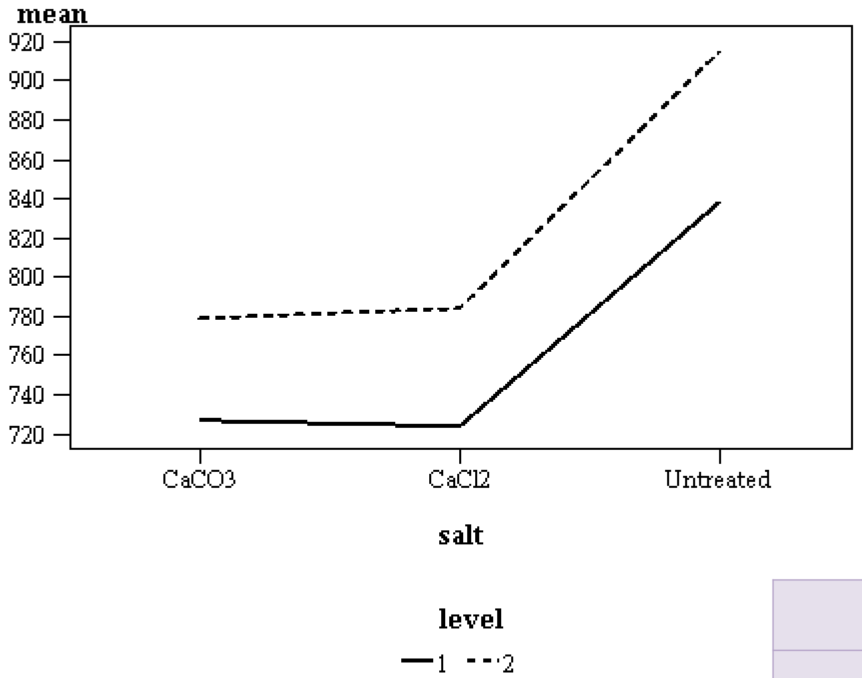
Source: Hsieh and Hardin, "Effects of Selected Inorganic Salts on Cotton Flammability", *Textile Research Journal* , Vol. 54, No. 3, 1984, pp. 171-179.

Data Spreadsheet & Find Means

Level	Salt	Temperature
1	Untreated	812
1	Untreated	827
1	Untreated	876
2	Untreated	945
2	Untreated	881
2	Untreated	919
1	$CaCO_3$	733
1	$CaCO_3$	728
1	$CaCO_3$	720
2	$CaCO_3$	786
2	$CaCO_3$	771
2	$CaCO_3$	779
1	$CaCl_2$	725
1	$CaCl_2$	727
1	$CaCl_2$	719
2	$CaCl_2$	756
2	$CaCl_2$	781
2	$CaCl_2$	814

Treatment Means			
	Salt		
Level	Untreated	$CaCO_3$	$CaCl_2$
1	838.33	727.00	723.67
2	915.00	778.67	783.67

Mean Profile Plot



Put the means in a two-way table and obtain a two-way plot of the means.

Mean Temperature To Ignite			
	Salt		
Level	Untreated	$CaCO_3$	$CaCl_2$
1	838.33	727.00	723.67
2	915.00	778.67	783.67

Interpret Results

- Fabric treated with $CaCO_3$ or $CaCl_2$ ignite at lower temperatures than Untreated Fabric.
- Fabric treated at concentration level 1 ignite at lower temperatures than fabric treated at concentration level 2.

Notation for Two-Way ANOVA

- One factor is A, with levels i = 1, 2, … , a
- Other factor is B, with levels j = 1, 2, … , b
- μ_{ij} is the population mean for the treatment defined by the i^{th} level of A and j^{th} level of B
- Y_{ijk} is the observed response for the k^{th} subject in the i^{th} level of A and j^{th} level of B
- ε_{ijk} is the random ‘noise’ for the k^{th} subject in the i^{th} level of A and j^{th} level of B

Marginal Means and Overall Mean

- μ_{ij} is the population mean for the cell (i.e. treatment) for Ai and Bj
- The average of population means for a given level of a factor is called the **marginal mean** for this factor
  - We call it a marginal mean because it may be displayed in the margins of the two way table of population means
- For instance, the marginal mean for A1 is denoted as
  $\mu_{A1} = (\mu_{11} + \mu_{12} + \mu_{13})/3$
- The **overall mean** is the average of all the treatment means, and we denote it as μ (with no subscripts)

Notation in Tabular Form

The symbols used to denote the population means for the treatments, the marginal means, and the overall mean are shown in the table below

Factor levels	B1	B2	B3	Marginal Means
A1	μ_{11}	μ_{12}	μ_{13}	μ_{A1}
A2	μ_{21}	μ_{22}	μ_{23}	μ_{A2}
Marginal Means	μ_{B1}	μ_{B2}	μ_{B3}	μ

overall mean

e.g. $\mu_{B1} = (\mu_{11} + \mu_{21})/2$

Means for Fabric Example

- We estimate the *population* means (cell means, marginal means and overall mean) with the respective *sample* means
- These estimates are least squares estimates

Level	Treatment Means			Marginal Means
	Untreated	$CaCO_3$	$CaCl_2$	
1	838.33	727.00	723.67	763.00
2	915.00	778.67	783.67	825.78
Marginal Means	876.67	752.84	753.67	794.39

Population Main Effects

- The main effects for factor A are the differences between the marginal means for factor A and the overall mean
  - main effect of level Ai = $\mu_{Ai} - \mu$
- Similarly, the main effects for factor B are the differences between the marginal means for factor B and the overall mean
  - main effect of level Bj = $\mu_{Bj} - \mu$

Main Effects for Fabric Data

Level	Treatment Means			Marginal Means
	Untreated	$CaCO_3$	$CaCl_2$	
1	838.33	727.00	723.67	763.00
2	915.00	778.67	783.67	825.78
Marginal Means	876.67	752.84	753.67	794.39

- Main effect for ‘Level 1’ (A1): 763.00 – 794.39 = -31.39
  (negative because this marginal mean is below average)
- Main effect for ‘Untreated’ (B1): 876.67 – 794.39 = 82.28
  (positive because this marginal mean is above average)

Additive Effects

- The effects of factors A and B are additive if, for all treatment combinations (Ai, Bj), the population mean can be expressed as

$$\mu_{ij} = \mu + (\mu_{Ai} - \mu) + (\mu_{Bj} - \mu)$$

- This is usually written

$$\mu_{ij} = \mu + \alpha_i + \beta_j$$

overall mean

main effect of Ai

main effect of Bj

Additive Model

- Two-way ANOVA model can be written

  $Y_{ijk} = \mu_{ij} + \varepsilon_{ijk}$

- For an additive model, $\mu_{ij} = \mu + \alpha_i + \beta_j$

- **Additive Model** $Y_{ijk} = \mu + \alpha_i + \beta_j + \varepsilon_{ijk}$

- Assume $\varepsilon_{ijk} \sim NIID(0, \sigma^2)$

Interaction Effects

- If one or more treatments have population means that cannot be expressed additively, then we say there is an interaction
- For treatment (Ai, Bj), the interaction effect
  - is denoted as $(\alpha\beta)_{ij}$
  - is defined as the difference between the treatment mean and the additive effect
  - $(\alpha\beta)_{ij} = \mu_{ij} - (\mu + \alpha_i + \beta_j)$

Interaction Model

- Two-way ANOVA model can be written

  $Y_{ijk} = \mu_{ij} + \varepsilon_{ijk}$

- For an interaction model, $\mu_{ij} = \mu + \alpha_i + \beta_j + (\alpha\beta)_{ij}$

- **Interaction Model** $Y_{ijk} = \mu + \alpha_i + \beta_j + (\alpha\beta)_{ij} + \varepsilon_{ijk}$

- Assume $\varepsilon_{ijk} \sim \text{NIID}(0, \sigma^2)$
- If all of the $(\alpha\beta)_{ij}$'s are 0, then the factors are additive

Interactions in Fabric Data

Level	Treatment Means			Marginal Means
	Untreated	$CaCO_3$	$CaCl_2$	
1	838.33	727.00	723.67	763.00
2	915.00	778.67	783.67	825.78
Marginal Means	876.67	752.84	753.67	794.39

Examine the means for $CaCl_2$, Level 1

Does $\hat{\mu}_{ij} = \hat{\mu} + \hat{\alpha}_i + \hat{\beta}_j$?

$\hat{\alpha}_i = 753.67 - 794.39 = -40.72$

$\hat{\beta}_j = 763.00 - 794.39 = -31.39$

Does $723.67 = 794.39 - 40.72 - 31.39$?

No… $723.28 \neq 722.28$

Means are not additive

This combination is NOT additive, so there is an interaction. It is not necessary to check the other combinations.

~ ~ HOWEVER ~ ~

The difference (between 723.28 and 722.28) seems small , so the interaction may not be statistically significant.

What You Should Know

- Two-way ANOVA data can be summarized
  - In a two-way tables of means
  - With a mean profile plot
- We have considered two models
  - Additive model: $Y_{ijk} = \mu + \alpha_i + \beta_j + \varepsilon_{ijk}$
  - Interaction model: $Y_{ijk} = \mu + \alpha_i + \beta_j + (\alpha\beta)_{ij} + \varepsilon_{ijk}$
- For both models, we assume $\varepsilon_{ijk} \sim NIID(0, \sigma^2)$
- If interaction exists, it may not be significant.

  (We will have a formal test for this.)