

MA50259: Statistical Design of Investigations

Dr. Sandipan Roy

Lecture 3

Experimental designs with multiple factors

- ▶ Completely randomized designs (CRD) with one factor are useful only when one factor is under study
- ▶ **One-at-a-time design** study each factor separately while holding all others constant.
- ▶ **Factorial Design** consider all possible combinations of the levels of all the factors under study
- ▶ **Interaction** when the effect of one factor changes depending on the level of another factor
- ▶ Interactions are more likely to be seen factorial rather than a one-at-a-time design

Factorial experimental designs: Example

- ▶ Design with two factors of four levels each
- ▶ **One-at-a-time design** eight replicates at each combination level. Total of 56 runs.
- ▶ **Factorial Design** two replicates of each cell. eight replicates of each level of each factor so is equivalent to the one-at-a-time design. Total of $2 \times 16 = 32$ runs
- ▶ Factorial Design is more efficient!

Factorial experimental designs: Example

- ▶ Assume objective is to find the factor combination that maximises the response
- ▶ **One-at-a-time design:** Complete the experiments with one factor first. Next calculate the cell means and then select the level with the highest mean. The second factor would be varied while holding the first constant at its optimal level.
- ▶ The optimal level of one factor may depend upon the level of the other factor. Therefore, by varying one factor at a time, the overall optimum may be missed!

Factorial experimental designs: Example

- ▶ **Factorial design** experiments are run at all combinations of the levels of the two factors
- ▶ In the factorial design (FD), the 32 treatment combinations would be randomised to the experimental units, thus preventing biases from unknown sources.
- ▶ In a FD the number of replicates of each factor level is equal to the number of replicates per cell times the product of the levels of all other factors in the design: $2 \times 4 = 8$

- this is called hidden replication!

Factorial experimental designs: Example 2

- ▶ Four factors where each factor has only two levels
- ▶ A factorial design would require all combinations of four factors at two levels, or $2^4 = 16$ cells
- ▶ Two replicates for each cell for a total of $2 \times 16 = 32$ experiments.
- ▶ To examine the effect of any one of the four factors, half the runs (or $2 \times 2^3 = 16$ due to the hidden replication) would be at one level of the factor and half at the other level.
- ▶ Thus the treatment effect would consist of a difference of two averages of 16. Results from the same 32 experiments can be used to calculate the treatment effect for each of the four factors.
- ▶ In a one-at-a-time plan, 32 runs would be required for comparing the levels of each factor while holding the others constant. This would result in $4 \times 16 + 16 = 80$ experiments!

Interactions

- ▶ **Interaction** (or joint effect between two factors): the effect of one factor upon the response will differ depending on the level of the other factor
- ▶ Interactions are common in practice, using the one-at-a-time strategy of experimentation tacitly assumes that interactions do not exist!
- ▶ Interaction plots: describe the effect of one factor upon the response, and then contrasting or comparing how that effect changes depending on the level of the other factor

Analysis of a Two-Factor Factorial

Means model

$$y_{ijk} = \mu_{ij} + \epsilon_{ijk}$$

- i represents the level of the first factor
 - ▶ j represents the level of the second factor
 - ▶ k represents the replicate number
 - ▶ μ_{ij} represents the expected response in the (i, j) th cell.

effects model

$$y_{ijk} = \mu + \alpha_i + \beta_j + \gamma_{ij} + \epsilon_{ijk}$$

- ▶ α_i, β_j are the main treatment effects
- ▶ γ_{ij} are the interaction effects represent the difference between the cell mean, μ_{ij} and $\mu + \alpha_i + \beta_j$
- ▶ ϵ_{ij} is the experimental error such that $E[\epsilon_{ij}] = 0$

Statistical analysis, the treatment effects model

first factor has two levels, the second factor has three levels, and there are two replicates per cell:

$$\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\epsilon}$$

where $\boldsymbol{\epsilon} \sim MVN(\mathbf{0}, \sigma^2 \mathbf{I})$ and

$$\begin{pmatrix} y_{111} \\ y_{112} \\ y_{211} \\ y_{212} \\ y_{121} \\ y_{122} \\ y_{221} \\ y_{222} \\ y_{131} \\ y_{132} \\ y_{231} \end{pmatrix} = \begin{pmatrix} 1 & 1 & 0 & 1 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 1 & 1 & 0 & 1 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 1 & 1 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 1 & 0 & 1 & 1 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 & 1 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 & 1 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 1 & 0 & 1 & 0 & 1 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 1 & 0 & 1 & 0 & 1 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 1 & 1 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 1 & 0 \\ 1 & 1 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 1 & 0 \\ 1 & 0 & 1 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} \mu \\ \alpha_1 \\ \alpha_2 \\ \beta_1 \\ \beta_2 \\ \beta_3 \\ \gamma_{11} \\ \gamma_{21} \\ \gamma_{12} \\ \gamma_{22} \\ \gamma_{13} \end{pmatrix} + \begin{pmatrix} \epsilon_1 \\ \epsilon_2 \\ \epsilon_3 \\ \epsilon_4 \\ \epsilon_5 \\ \epsilon_6 \\ \epsilon_7 \\ \epsilon_8 \\ \epsilon_9 \\ \epsilon_{10} \\ \epsilon_{11} \end{pmatrix}$$

Statistical analysis, the treatment effects model

$\mathbf{X}^T \mathbf{X}$ is singular and to solve the normal equations the R function `lm` drops the indicators for the first level of each factor in the main effect columns, and creates the columns for the interaction as all possible products of the main effect columns. This makes the new $\mathbf{X}^T \mathbf{X}$ matrix full rank

$$\begin{pmatrix} y_{111} \\ y_{112} \\ y_{211} \\ y_{212} \\ y_{121} \\ y_{122} \\ y_{221} \\ y_{222} \\ y_{131} \\ y_{132} \\ y_{231} \\ y_{232} \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 & 0 & 0 \\ 1 & 0 & 1 & 0 & 0 & 0 \\ 1 & 0 & 1 & 0 & 0 & 0 \\ 1 & 1 & 1 & 0 & 1 & 0 \\ 1 & 1 & 1 & 0 & 1 & 0 \\ 1 & 0 & 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 1 & 0 & 0 \\ 1 & 1 & 0 & 1 & 0 & 1 \\ 1 & 1 & 0 & 1 & 0 & 1 \end{pmatrix} \begin{pmatrix} \mu \\ \alpha_2 \\ \beta_2 \\ \beta_3 \\ \gamma_{22} \\ \gamma_{23} \end{pmatrix} + \begin{pmatrix} \epsilon_{111} \\ \epsilon_{112} \\ \epsilon_{211} \\ \epsilon_{212} \\ \epsilon_{121} \\ \epsilon_{122} \\ \epsilon_{221} \\ \epsilon_{222} \\ \epsilon_{131} \\ \epsilon_{132} \\ \epsilon_{231} \\ \epsilon_{232} \end{pmatrix}$$