4. Analysis of Two-Factor Experiments Based on Cell-Means Models

An Example Two-Factor Experiment

Researchers were interested in studying the effects of 2 diets (low fiber, high fiber) and 3 drugs (D1, D2, D3) on weight gained by Yorkshire pigs. A total of 12 pigs were assigned to the 6 diet $\times$ drug combinations using a balanced and completely randomized experimental design. Pigs were housed in individual pens, injected with their assigned drugs once per week, and fed their assigned diets for a 6-week period. The amount of weight gained during the 6-week period was recorded for each pig.

Factors and Levels

- This experiment involves 2 *factors*: Diet and Drug.
- The factor Diet has 2 *levels*: low fiber and high fiber.
- The factor Drug has 3 *levels*: D1, D2, and D3.

Treatment Design vs. Experimental Design

- A combination of one level from each factor forms a *treatment*.
- The *treatment design* used in this experiment is known as a *full-factorial treatment design* because each possible combination of one level from each factor was applied to at least one experimental unit.
- The *experimental design* is a balanced *completely randomized design* (CRD) because all possible balanced assignments of the 12 pigs to the 6 treatment groups were equally likely.

Table of Treatments

Treatment	Diet	Drug
1	1	1
2	1	2
3	1	3
4	2	1
5	2	2
6	2	3

Diet 1 = Low Fiber, Diet 2 = High Fiber

Drug 1 = D1, Drug 2 = D2, Drug 3 = D3

The Cell-Means Model

For $i = 1, 2$; $j = 1, 2, 3$; and $k = 1, 2$; let y_{ijk} denote the weight gain of the k^{th} pig that received diet i and drug j, and suppose

$$\begin{aligned} y_{ijk} &= \mu_{ij} + \epsilon_{ijk}, \\ \epsilon_{ijk} &\overset{iid}{\sim} \mathcal{N}(0, \sigma^2), \end{aligned}$$

where $\mu_{11}, \mu_{12}, \mu_{13}, \mu_{21}, \mu_{22}, \mu_{23} \in \mathbb{R}$ and $\sigma^2 \in \mathbb{R}^+$ are unknown parameters.

Table of Treatments and Means for Cell-Means Model

Treatment	Diet	Drug	Mean
1	1	1	μ_{11}
2	1	2	μ_{12}
3	1	3	μ_{13}
4	2	1	μ_{21}
5	2	2	μ_{22}
6	2	3	μ_{23}

Diet 1 = Low Fiber,
Diet 2 = High Fiber

Drug 1 = D1, Drug 2 = D2,
Drug 3 = D3

	Drug 1	Drug 2	Drug 3
Diet 1	μ_{11}	μ_{12}	μ_{13}
Diet 2	μ_{21}	μ_{22}	μ_{23}

Cell-Means Model in Matrix and Vector Form

$$\begin{bmatrix} y_{111} \\ y_{112} \\ y_{121} \\ y_{122} \\ y_{131} \\ y_{132} \\ y_{211} \\ y_{212} \\ y_{221} \\ y_{222} \\ y_{231} \\ y_{232} \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} \mu_{11} \\ \mu_{12} \\ \mu_{13} \\ \mu_{21} \\ \mu_{22} \\ \mu_{23} \end{bmatrix} + \begin{bmatrix} \epsilon_{111} \\ \epsilon_{112} \\ \epsilon_{121} \\ \epsilon_{122} \\ \epsilon_{131} \\ \epsilon_{132} \\ \epsilon_{211} \\ \epsilon_{212} \\ \epsilon_{221} \\ \epsilon_{222} \\ \epsilon_{231} \\ \epsilon_{232} \end{bmatrix}$$

$$\boldsymbol{y} = \boldsymbol{X}\boldsymbol{\beta} + \boldsymbol{\epsilon}, \quad \boldsymbol{\epsilon} \sim \mathcal{N}(\boldsymbol{0}, \sigma^2 \boldsymbol{I})$$

Each Cell Mean is Estimable

$$\mathrm{E}(\boldsymbol{y}) = \boldsymbol{X}\boldsymbol{\beta} = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} \mu_{11} \\ \mu_{12} \\ \mu_{13} \\ \mu_{21} \\ \mu_{22} \\ \mu_{23} \end{bmatrix} = \begin{bmatrix} \mu_{11} \\ \mu_{11} \\ \mu_{12} \\ \mu_{12} \\ \mu_{13} \\ \mu_{13} \\ \mu_{21} \\ \mu_{21} \\ \mu_{22} \\ \mu_{22} \\ \mu_{23} \\ \mu_{23} \end{bmatrix}$$

$\boldsymbol{\beta}$ is Estimable When $\text{rank}(\boldsymbol{X}) = p$

For the General Linear Model, the parameter vector $\boldsymbol{\beta}$ is estimable whenever $\boldsymbol{X}$ has full-column rank, i.e., whenever $\text{rank}(\boldsymbol{X}) = p$.

Finding the OLS Estimator of β

$$\boldsymbol{X} = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix} \implies \boldsymbol{X}^\top \boldsymbol{X} = \begin{bmatrix} 2 & 0 & 0 & 0 & 0 & 0 \\ 0 & 2 & 0 & 0 & 0 & 0 \\ 0 & 0 & 2 & 0 & 0 & 0 \\ 0 & 0 & 0 & 2 & 0 & 0 \\ 0 & 0 & 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 0 & 0 & 2 \end{bmatrix}$$

$$\implies \boldsymbol{X}^\top \boldsymbol{X} = 2\boldsymbol{I}_{6\times 6}$$

$$\implies (\boldsymbol{X}^\top \boldsymbol{X})^{-1} = \frac{1}{2}\boldsymbol{I}_{6\times 6}$$

$$\boldsymbol{X} = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix}, \boldsymbol{y} = \begin{bmatrix} y_{111} \\ y_{112} \\ y_{121} \\ y_{122} \\ y_{131} \\ y_{132} \\ y_{211} \\ y_{212} \\ y_{221} \\ y_{222} \\ y_{231} \\ y_{232} \end{bmatrix} \Longrightarrow \boldsymbol{X}^\top \boldsymbol{y} = \begin{bmatrix} y_{11\cdot} \\ y_{12\cdot} \\ y_{13\cdot} \\ y_{21\cdot} \\ y_{22\cdot} \\ y_{23\cdot} \end{bmatrix}$$

$$\therefore \widehat{\boldsymbol{\beta}} = (\boldsymbol{X}^\top \boldsymbol{X})^{-1} \boldsymbol{X}^\top \boldsymbol{y} = \frac{1}{2} \boldsymbol{I} \boldsymbol{X}^\top \boldsymbol{y} = \frac{1}{2} \boldsymbol{X}^\top \boldsymbol{y} = \begin{bmatrix} \bar{y}_{11\cdot} \\ \bar{y}_{12\cdot} \\ \bar{y}_{13\cdot} \\ \bar{y}_{21\cdot} \\ \bar{y}_{22\cdot} \\ \bar{y}_{23\cdot} \end{bmatrix}$$

Table of Cell Means

	Drug 1	Drug 2	Drug 3
Diet 1	μ_{11}	μ_{12}	μ_{13}
Diet 2	μ_{21}	μ_{22}	μ_{23}

Marginal Means

	Drug 1	Drug 2	Drug 3	
Diet 1	μ_{11}	μ_{12}	μ_{13}	$\bar{\mu}_{1\cdot}$
Diet 2	μ_{21}	μ_{22}	μ_{23}	$\bar{\mu}_{2\cdot}$
	$\bar{\mu}_{\cdot 1}$	$\bar{\mu}_{\cdot 2}$	$\bar{\mu}_{\cdot 3}$	$\bar{\mu}_{\cdot\cdot}$

$$\bar{\mu}_{i\cdot} = \frac{1}{3}\sum_{j=1}^{3} \mu_{ij}$$

$$\bar{\mu}_{\cdot j} = \frac{1}{2}\sum_{i=1}^{2} \mu_{ij}$$

$$\bar{\mu}_{\cdot\cdot} = \frac{1}{(2)(3)}\sum_{i=1}^{2}\sum_{j=1}^{3} \mu_{ij}$$

Least Squares Means (LSMEANS) in SAS

SAS can be used to compute LSMEANS.

LSMEANS are simply OLS estimators of cell or marginal means.

Each LSMEAN has the form $\boldsymbol{c}^\top \widehat{\boldsymbol{\beta}}$ for an appropriate vector $\boldsymbol{c}$.

For example, the LSMEAN for diet 1 is $\boldsymbol{c}^\top \widehat{\boldsymbol{\beta}}$ with

$$\boldsymbol{c}^\top = \left[\frac{1}{3}, \frac{1}{3}, \frac{1}{3}, 0, 0, 0\right] \text{ and } \widehat{\boldsymbol{\beta}} = [\bar{y}_{11\cdot}, \bar{y}_{12\cdot}, \bar{y}_{13\cdot}, \bar{y}_{21\cdot}, \bar{y}_{22\cdot}, \bar{y}_{23\cdot}]^\top.$$

LSMEANS (continued)

Note that the LSMEAN for Diet 1 is simply an average of the estimated means for treatments involving Diet 1.

When data are balanced, the LSMEAN for Diet 1 is also just the average of responses for all pigs that were fed Diet 1.

When data are unbalanced, the LSMEAN for Diet 1 may not equal the average of responses for all pigs that were fed Diet 1.

Standard Error

A *standard error* is the estimated standard deviation of a statistic.

A *standard error* is usually found by estimating the variance of a statistic and then taking the square root of the estimate.

Because each LSMEAN has the form $\boldsymbol{c}^\top \widehat{\boldsymbol{\beta}}$ for an appropriate vector $\boldsymbol{c}$, the standard error for an LSMEAN is given by

$$\sqrt{\widehat{\mathrm{Var}}(\boldsymbol{c}^\top \widehat{\boldsymbol{\beta}})} = \sqrt{\widehat{\sigma}^2 \boldsymbol{c}^\top (\boldsymbol{X}^\top \boldsymbol{X})^- \boldsymbol{c}}.$$

Effects we can estimate

- Simple Effects
- Main Effects
- Interactions

Simple Effects

A *simple effect* is the difference between cell means that differ in level for only one factor.

In our two-factor example, *simple effects* are differences between cell means within any row or within any column.

	Drug 1	Drug 2	Drug 3
Diet 1	μ_{11}	μ_{12}	μ_{13}
Diet 2	μ_{21}	μ_{22}	μ_{23}

Simple Effects (continued)

The *simple effect* of Diet for Drug 1 is $\mu_{11} - \mu_{21}$.

The *simple effect* of Drug 2 vs. Drug 3 for Diet 2 is $\mu_{22} - \mu_{23}$.

	Drug 1	Drug 2	Drug 3
Diet 1	μ_{11}	μ_{12}	μ_{13}
Diet 2	μ_{21}	μ_{22}	μ_{23}

Main Effects

A *main effect* is the difference between marginal means associated with two levels of a factor.

In our two-factor example, the *main effect* of Diet is $\bar{\mu}_{1\cdot} - \bar{\mu}_{2\cdot}$.

	Drug 1	Drug 2	Drug 3	
Diet 1	μ_{11}	μ_{12}	μ_{13}	$\bar{\mu}_{1\cdot}$
Diet 2	μ_{21}	μ_{22}	μ_{23}	$\bar{\mu}_{2\cdot}$
	$\bar{\mu}_{\cdot 1}$	$\bar{\mu}_{\cdot 2}$	$\bar{\mu}_{\cdot 3}$	$\bar{\mu}_{\cdot\cdot}$

Main Effects (continued)

In our two-factor example, the *main effects* of Drug involve the differences $\bar{\mu}_{.1} - \bar{\mu}_{.2}$, $\bar{\mu}_{.1} - \bar{\mu}_{.3}$, and $\bar{\mu}_{.2} - \bar{\mu}_{.3}$.

	Drug 1	Drug 2	Drug 3	
Diet 1	μ_{11}	μ_{12}	μ_{13}	$\bar{\mu}_{1.}$
Diet 2	μ_{21}	μ_{22}	μ_{23}	$\bar{\mu}_{2.}$
	$\bar{\mu}_{.1}$	$\bar{\mu}_{.2}$	$\bar{\mu}_{.3}$	$\bar{\mu}_{..}$

Main Effects (continued)

If $\bar{\mu}_{1\cdot} = \bar{\mu}_{2\cdot}$, it would be customary to say, "There is no Diet main effect."

If $\bar{\mu}_{\cdot 1} = \bar{\mu}_{\cdot 2} = \bar{\mu}_{\cdot 3}$, it would be customary to say, "There are no Drug main effects."

	Drug 1	Drug 2	Drug 3	
Diet 1	μ_{11}	μ_{12}	μ_{13}	$\bar{\mu}_{1\cdot}$
Diet 2	μ_{21}	μ_{22}	μ_{23}	$\bar{\mu}_{2\cdot}$
	$\bar{\mu}_{\cdot 1}$	$\bar{\mu}_{\cdot 2}$	$\bar{\mu}_{\cdot 3}$	$\bar{\mu}_{\cdot\cdot}$

Interaction Effects

The linear combination $\mu_{ij} - \mu_{ij'} - \mu_{i'j} + \mu_{i'j'}$ for $i \neq i'$ and $j \neq j'$ is an *interaction effect*.

For example,

$$\mu_{11}-\mu_{12}-\mu_{21}+\mu_{22} = (\mu_{11}-\mu_{12})-(\mu_{21}-\mu_{22}) = (\mu_{11}-\mu_{21})-(\mu_{12}-\mu_{22})$$

is an interaction effect.

	Drug 1	Drug 2	Drug 3
Diet 1	μ_{11}	μ_{12}	μ_{13}
Diet 2	μ_{21}	μ_{22}	μ_{23}

Interaction Effects (continued)

When all interaction effects are zero, we may say there are “no interactions” between the factors or that the two factors do not “interact.”

When there are no interactions between factors, the simple effects of either factor are the same across all levels of the other factor.

For example, when there are no interactions between the factors Diet and Drug, the simple effect of Diet is the same for each level of Drug. Likewise, any simple effect of Drug is the same for both diets.

Testing for Non-Zero Effects

We can test whether simple effects, main effects, or interaction effects are zero vs. non-zero using tests of the form

$$\mathrm{H}_0 : \boldsymbol{C\beta} = \mathbf{0} \text{ vs. } \mathrm{H}_A : \boldsymbol{C\beta} \neq \mathbf{0}.$$

The following slides give appropriate C matrices for several examples.

Testing for Non-Zero Effects

H_0: No simple effect of Diet for Drug 1 $(\mu_{11} = \mu_{21})$

$$\begin{bmatrix} 1 & 0 & 0 & -1 & 0 & 0 \end{bmatrix} \begin{bmatrix} \mu_{11} \\ \mu_{12} \\ \mu_{13} \\ \mu_{21} \\ \mu_{22} \\ \mu_{23} \end{bmatrix} = \begin{bmatrix} 0 \end{bmatrix}$$

H_0: No simple effect of Drug 2 vs. Drug 3 for Diet 2 $(\mu_{22} = \mu_{23})$

$$\begin{bmatrix} 0 & 0 & 0 & 0 & 1 & -1 \end{bmatrix} \begin{bmatrix} \mu_{11} \\ \mu_{12} \\ \mu_{13} \\ \mu_{21} \\ \mu_{22} \\ \mu_{23} \end{bmatrix} = \begin{bmatrix} 0 \end{bmatrix}$$

Testing for Non-Zero Effects

H_0: No Diet Main Effect $(\bar{\mu}_{1\cdot} = \bar{\mu}_{2\cdot})$

$$\begin{bmatrix} \frac{1}{3} & \frac{1}{3} & \frac{1}{3} & -\frac{1}{3} & -\frac{1}{3} & -\frac{1}{3} \end{bmatrix} \begin{bmatrix} \mu_{11} \\ \mu_{12} \\ \mu_{13} \\ \mu_{21} \\ \mu_{22} \\ \mu_{23} \end{bmatrix} = \begin{bmatrix} 0 \end{bmatrix}$$

Testing for Non-Zero Effects

H_0: No Drug Main Effects $(\bar{\mu}_{.1} = \bar{\mu}_{.2} = \bar{\mu}_{.3})$

$$\begin{bmatrix} \frac{1}{2} & -\frac{1}{2} & 0 & \frac{1}{2} & -\frac{1}{2} & 0 \\ \frac{1}{2} & 0 & -\frac{1}{2} & \frac{1}{2} & 0 & -\frac{1}{2} \end{bmatrix} \begin{bmatrix} \mu_{11} \\ \mu_{12} \\ \mu_{13} \\ \mu_{21} \\ \mu_{22} \\ \mu_{23} \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

Testing for Non-Zero Effects

H_0: No Drug Main Effects $(\bar{\mu}_{.1} = \bar{\mu}_{.2} = \bar{\mu}_{.3})$

Here is an alternative specification that will yield the same test.

$$\begin{bmatrix} \frac{1}{2} & -\frac{1}{2} & 0 & \frac{1}{2} & -\frac{1}{2} & 0 \\ 0 & \frac{1}{2} & -\frac{1}{2} & 0 & \frac{1}{2} & -\frac{1}{2} \end{bmatrix} \begin{bmatrix} \mu_{11} \\ \mu_{12} \\ \mu_{13} \\ \mu_{21} \\ \mu_{22} \\ \mu_{23} \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

Testing for Non-Zero Effects

H_0: No Drug Main Effects $(\bar{\mu}_{.1} = \bar{\mu}_{.2} = \bar{\mu}_{.3})$

Here is yet another alternative specification that will yield the same test.

$$\begin{bmatrix} \frac{1}{2} & -\frac{1}{2} & 0 & \frac{1}{2} & -\frac{1}{2} & 0 \\ \\ \frac{1}{4} & \frac{1}{4} & -\frac{1}{2} & \frac{1}{4} & \frac{1}{4} & -\frac{1}{2} \end{bmatrix} \begin{bmatrix} \mu_{11} \\ \mu_{12} \\ \mu_{13} \\ \mu_{21} \\ \mu_{22} \\ \mu_{23} \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

Testing for Non-Zero Effects

H_0: No Diet-by-Drug Interactions:
($\mu_{ij} - \mu_{ij'} - \mu_{i'j} + \mu_{i'j'} = 0$ for all $i \neq i'$ and $j \neq j'$)

$$\begin{bmatrix} 1 & -1 & 0 & -1 & 1 & 0 \\ 1 & 0 & -1 & -1 & 0 & 1 \end{bmatrix} \begin{bmatrix} \mu_{11} \\ \mu_{12} \\ \mu_{13} \\ \mu_{21} \\ \mu_{22} \\ \mu_{23} \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

R/SAS Code and Output

- The R/SAS code and output for the above sample are given as separate handouts.
- We will discuss alternative parameterizations of the cell-means model as part of the implementation in SAS and R

