8.1-8.6 Two-Way Factorial Designs

Yibi Huang

Problem 8.1 — Sprouting Barley (p.166 in Oehlert)

Brewer's malt is produced from germinating barley, so brewers like to know under what conditions they should germinate their barley. The following is part of an experiment on barley germination.

- 30 lots of barley seeds, 100 seeds per lot, are randomly divided into 10 groups of 3 lots
- Each group receives a treatment according to
  - water amount used in germination — 4 ml or 8 ml
  - age of seeds in weeks after harvest — 1, 3, 6, 9, or 12
- Response: # of seeds germinating

	Age of Seeds (weeks)				
water	1	3	6	9	12
4(ml)	11	7	9	13	20
	9	16	19	35	37
	6	17	35	28	45
8(ml)	8	1	5	1	11
	3	7	9	10	15
	3	3	9	9	25

Basic Terminology

The sprouting barley experiment has 10 treatments. The 10 treatments has a **factorial structure**.

- A *factor* is an experimentally adjustable variable,
  e.g. water amount used in germination,
  age of seeds in weeks after harvest, ...
- Factors have *levels*, e.g.
  water amount is a factor with 2 levels (4 ml or 8 ml)
  age of seeds is a factor with 5 levels (1, 3, 6, 9, 12 weeks)
- A treatment is a *combination of factors*.
  In the barley experiment, the treatments are the 2×5 combinations of the possible levels of the two factors

(4ml, 1 wk)	(4ml, 3 wks)	(4ml, 6 wks)	(4ml, 9 wks)	(4ml, 12 wks)
(8ml, 1 wk)	(8ml, 3 wks)	(8ml, 6 wks)	(8ml, 9 wks)	(8ml, 12 wks)

Full k-Way Factorial Design

- Consider k factors with respectively L_1, $L_2, \ldots,$ L_k levels, a **full k-way factorial design** include all the $L_1 \times L_2 \times \cdots \times L_k$ combination of the k factors as treatments.

- A factorial design is said to be *balanced* if all the treatment groups have the same number of *replicates*. Otherwise, the design is *unbalanced*.
  - Question: How many units are there in a 3×2 design with 4 replicates?

- Balanced designs have many advantages, but not always necessary — sometimes if a unit fails (ex, a test tube gets dropped) we might end up with unbalanced results even if the original design was balanced

Data for a Two-Way $a \times b$ Design with n Replicates

	B-level 1	B-level 2		B-level b
A-level 1	y_{111} y_{112} $\vdots$ y_{11n}	y_{121} y_{122} $\vdots$ y_{12n}	$\cdots$ $\cdots$	y_{1b1} y_{1b2} $\vdots$ y_{1bn}
A-level 2	y_{211} y_{212} $\vdots$ y_{21n}	y_{221} y_{222} $\vdots$ y_{22n}	$\cdots$ $\cdots$	y_{2b1} y_{2b2} $\vdots$ y_{2bn}
$\vdots$ $\vdots$	$\vdots$ $\vdots$	$\vdots$ $\vdots$	$\ddots$ $\ddots$	$\vdots$ $\vdots$
A-level a	y_{a11} y_{a12} $\vdots$ y_{a1n}	y_{a21} y_{a22} $\vdots$ y_{a2n}	$\cdots$ $\cdots$	y_{ab1} y_{ab2} $\vdots$ y_{abn}

Means Model for a Two-Way Factorial Design

For a $a \times b$ two-way factorial experiment with n replicates

$$\text{means model}: y_{ijk} = \mu_{ij} + \varepsilon_{ijk} \quad \text{for} \begin{cases} i = 1, \dots, a, \\ j = 1, \dots, b, \\ k = 1, \dots, n. \end{cases}$$

- y_{ijk} = the kth replicate in the treatment formed from the ith level of factor A and jth level of factor B
- ε_{ijk}'s are i.i.d. $N(0, \sigma^2)$
- μ_{ij} = the mean response in the treatment formed from the ith level of factor A and jth level of factor B
- The means model regards the 2-way factorial design as a CRD with $a \times b$ treatments, ignoring the factorial structure of the treatments.

Two-Way Interaction Contrast

The two-way interaction contrast between level (i_1, i_2) of factor A and level (j_1, j_2) of factor B is defined as

$$C = \mu_{i_1 j_1} - \mu_{i_1 j_2} - \mu_{i_2 j_1} + \mu_{i_2 j_2},$$

which has two interpretations.

$$\begin{aligned} C &= \mu_{i_1 j_1} - \mu_{i_1 j_2} - \mu_{i_2 j_1} + \mu_{i_2 j_2} \\ &= \underbrace{(\mu_{i_1 j_1} - \mu_{i_1 j_2})}_{\substack{\text{effect of changing B from } j_1 \text{ to } j_2 \\ \text{when A is fixed at } i_1}} - \underbrace{(\mu_{i_2 j_1} - \mu_{i_2 j_2})}_{\substack{\text{effect of changing B from } j_1 \text{ to } j_2 \\ \text{when A is fixed at } i_2}} \\ &= \underbrace{(\mu_{i_1 j_1} - \mu_{i_2 j_1})}_{\substack{\text{effect of changing A from } i_1 \text{ to } i_2 \\ \text{when B is fixed at } j_1}} - \underbrace{(\mu_{i_1 j_2} - \mu_{i_2 j_2})}_{\substack{\text{effect of changing A from } i_1 \text{ to } i_2 \\ \text{when B is fixed at } j_2}} \end{aligned}$$

Two-Way Interaction

We say factor A and factor B have **no interaction** if and only if the two-way interaction contrast between any two levels of A and any two levels of B is 0, i.e.,

$$\mu_{i_1 j_1} - \mu_{i_1 j_2} - \mu_{i_2 j_1} + \mu_{i_2 j_2} = 0 \quad \text{for all } i_1, i_2, j_1, j_2,$$

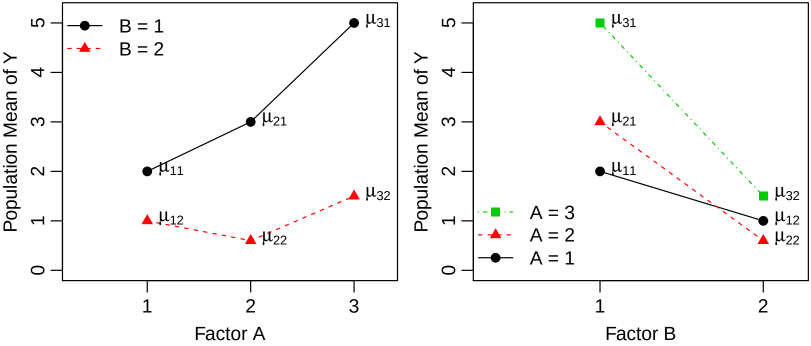
which has two interpretations:

- effect of A on Y doesn't change with the levels of B, and
- effect of B on Y doesn't change with the levels of A

Conversely, two factors A and B are said to have **two-way interactions** if the effect of A on Y changes with the levels of B, or the effect of B on Y changes with the levels of A.

Interaction Plots

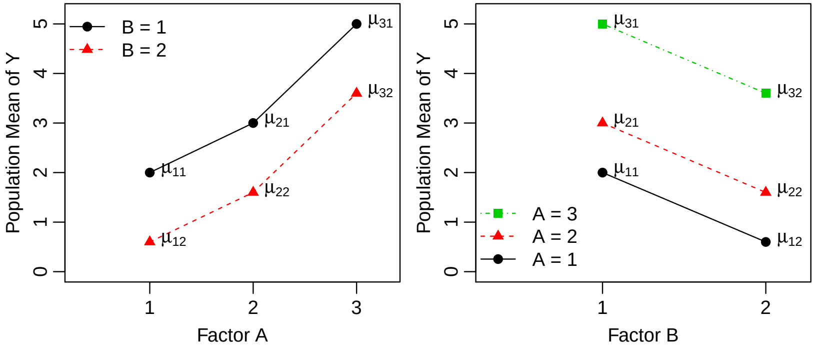
Plotting cell means (μ_{ij}) against levels of one factor (A or B), with different lines for the other factor (B or A)



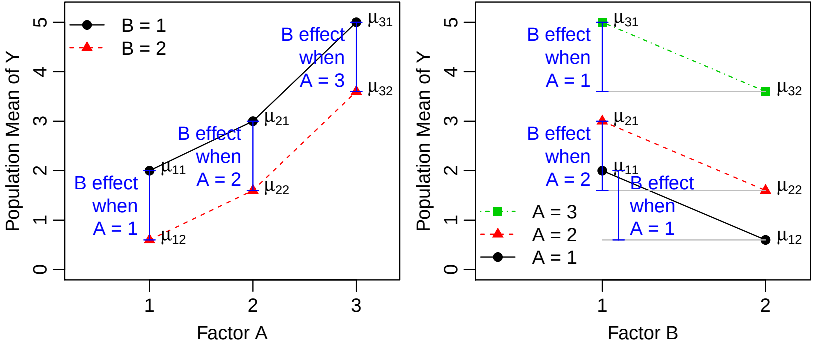
A is on the *x*-axis	*B* is on the *x*-axis
B is the trace factor	*A* is the trace factor

The two interaction plots convey the same information.

Parallel Lines Indicate No Interaction (1)

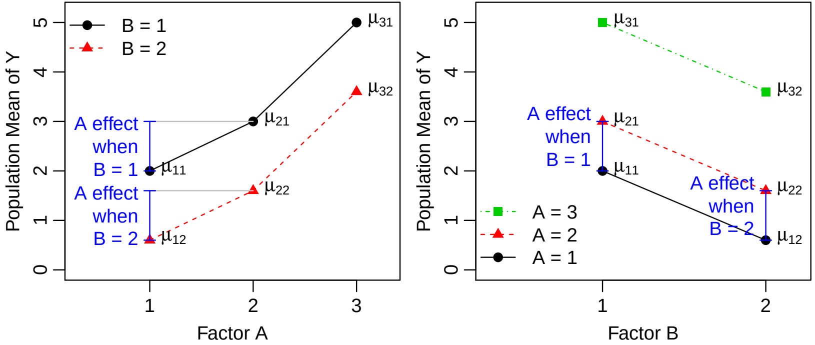


Parallel Lines Indicate No Interaction (2)



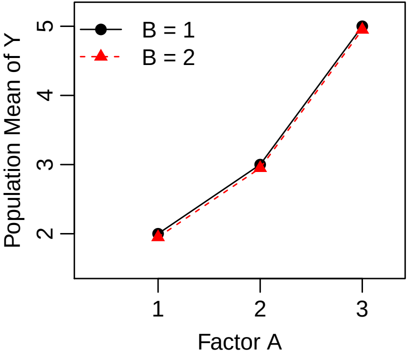
Effect of B on Y doesn't change with levels of A

Parallel Lines Indicate No Interaction (3)



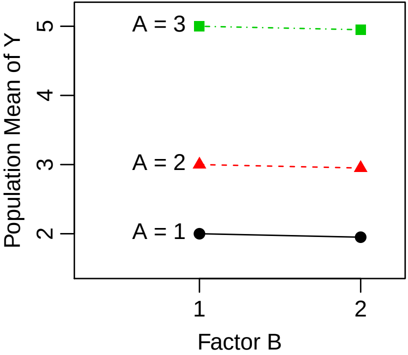
Effect of A on Y doesn't change with levels of B

What does the interaction plot below tell us? (1)



- No AB interaction
- B has no effect on Y since there is **no gap** between lines
- A has some effect on Y since the lines are **not horizontal**

What does the interaction plot below tell us? (2)



- No AB interaction
- B has no effect on Y since the lines are horizontal
- A has some effect on Y since there are gaps between lines

Interaction Plots for the Sprouting Barley Study

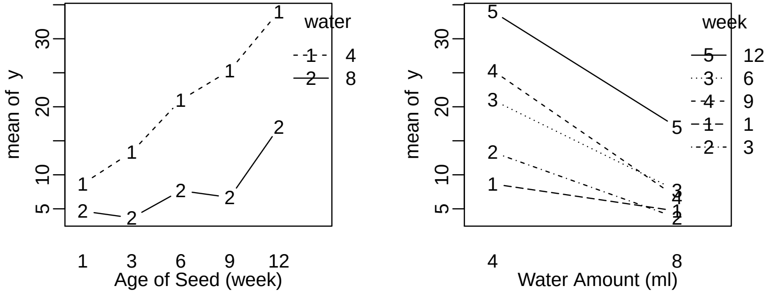
In reality, the population means μ_{ij} are not observable. Interaction plots are made using sample means $\overline{y}_{ij\bullet}$ rather than population means μ_{ij}.

y_{ijk}	Age of Seeds (weeks)				
	1	3	6	9	12
water 4(ml)	11	7	9	13	20
	9	16	19	35	37
	6	17	35	28	45
water 8(ml)	8	1	5	1	11
	3	7	9	10	15
	3	3	9	9	25

sample means $\overline{y}_{ij\bullet}$	Age of Seeds (weeks)				
	1	3	6	9	12
water 4(ml)	8.67	13.33	21.00	25.33	34.00
water 8(ml)	4.67	3.67	7.67	6.67	17.00

Interaction Plots in R

```
barley = read.table(
    "http://www.stat.uchicago.edu/~yibi/s222/SproutingBarley.txt",h=T)
with(barley, interaction.plot(week,water,y,type="b",
     xlab="Age of Seed (week)"))
with(barley, interaction.plot(water,week,y,type="b",
     xlab="Water Amount (ml)"))
```

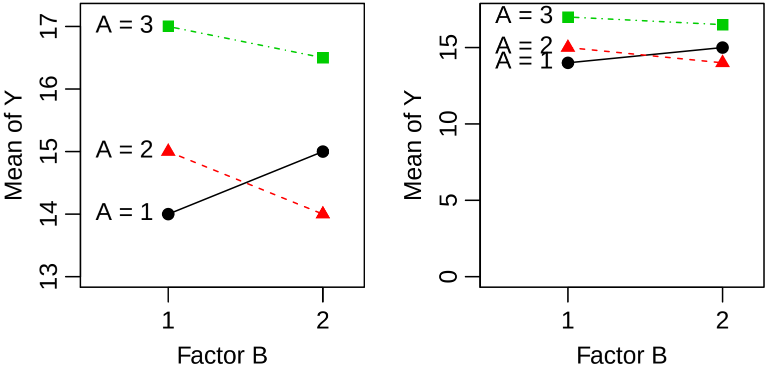


Note that lines in the interaction plots might not be exactly parallel even if the two factors have no interaction since $\bar{y}_{ij\bullet} \neq \mu_{ij}$.

The less parallel the lines, the stronger the evidence of interactions.

"Parallel" Or Not Is Affected by The Y-Scale

Please note that the y-scale might affect your perception of whether the lines are "parallel" or not.



Check the Y-scale and see if the change in the slopes is big enough to be important.

Additive Model

An **additive model** or **main-effect model** for two-way factorial data is as follows

$$y_{ijk} = \mu + \underbrace{\alpha_i}_{\text{A main effect}} + \underbrace{\beta_j}_{\text{B main effect}} + \varepsilon_{ijk} \quad \text{for} \begin{cases} i = 1, \ldots, a, \\ j = 1, \ldots, b, \\ k = 1, \ldots, n. \end{cases}$$

- The additive model takes the factorial structure of the $a \times b$ treatments into account
- The additive model is nested in the means model $y_{ijk} = \mu_{ij} + \varepsilon_{ijk}$ since the means model will become the additive model if

$$\mu_{ij} = \mu + \alpha_i + \beta_j \quad \text{for all } i, j.$$

Additive Model Assumes No Interactions

If the additive model $y_{ijk} = \mu + \alpha_i + \beta_j + \varepsilon_{ijk}$ is true, then

$$\mu_{ij} = \mu + \alpha_i + \beta_j \quad \text{for all } i, j,$$

we have

$$\begin{aligned} & \mu_{i_1 j_1} - \mu_{i_1 j_2} - \mu_{i_2 j_1} + \mu_{i_2 j_2} \\ = & (\mu + \alpha_{i_1} + \beta_{j_1}) - (\mu + \alpha_{i_1} + \beta_{j_2}) \\ & - (\mu + \alpha_{i_2} + \beta_{j_1}) + (\mu + \alpha_{i_2} + \beta_{j_2}) \\ = & 0 \end{aligned}$$

for all i_1, i_2, j_1, j_2. Thus the two factors have no interaction.

However, under the means model, the two factors might have interactions.

$$\mu_{i_1 j_1} - \mu_{i_1 j_2} - \mu_{i_2 j_1} + \mu_{i_2 j_2} \text{ might not be 0.}$$

Main-Effect-Interaction Model for 2-Way Factorial Designs

The main-effect-interaction model is an extension of the additive model that allows interactions

$$y_{ijk} = \mu + \alpha_i + \beta_j + \alpha\beta_{ij} + \varepsilon_{ijk} \quad \text{for} \begin{cases} i = 1, \dots, a, \\ j = 1, \dots, b, \\ k = 1, \dots, n. \end{cases}$$

- $\alpha\beta_{ij}$ is a parameter by itself. $\alpha\beta_{ij} \neq \alpha_i \times \beta_j$; $\alpha\beta_{ij} \neq \alpha \times \beta_{ij}$

$$\begin{array}{|c|c|c|c|} \hline \mu_{11} & \mu_{12} & \cdots & \mu_{1b} \\ \hline \mu_{21} & \mu_{22} & \cdots & \mu_{2b} \\ \hline \vdots & \vdots & \ddots & \vdots \\ \hline \mu_{a1} & \mu_{a2} & \cdots & \mu_{ab} \\ \hline \end{array} = \begin{array}{|c|} \hline \\ \mu \\ \\ \hline \end{array} + \begin{array}{|c|} \hline \alpha_1 \\ \hline \alpha_2 \\ \hline \vdots \\ \hline \alpha_a \\ \hline \end{array} + \begin{array}{|c|c|c|c|} \hline \beta_1 & \beta_2 & \cdots\cdots & \beta_b \\ \hline \end{array}$$

$$+ \begin{array}{|c|c|c|c|} \hline \alpha\beta_{11} & \alpha\beta_{12} & \cdots & \alpha\beta_{1b} \\ \hline \alpha\beta_{21} & \alpha\beta_{22} & \cdots & \alpha\beta_{2b} \\ \hline \vdots & \vdots & \ddots & \vdots \\ \hline \alpha\beta_{a1} & \alpha\beta_{a2} & \cdots & \alpha\beta_{ab} \\ \hline \end{array}$$

Main-Effect-Interaction Model Is Overparameterized

- The main-effect-interaction model $y_{ijk} = \mu + \alpha_i + \beta_j + \alpha\beta_{ij} + \varepsilon_{ijk}$ is equivalent to the means model $y_{ijk} = \mu_{ij} + \varepsilon_{ijk}$.
  They have identical predicted values, residuals, and SSE.
- For a two-way $a \times b$ design, the means model has ab parameters; the main-effect-interaction model has $1 + a + b + ab$ parameters

μ_{11}	μ_{12}	$\cdots$	μ_{1b}
μ_{21}	μ_{22}	$\cdots$	μ_{2b}
$\vdots$	$\vdots$		$\vdots$
μ_{a1}	μ_{a2}	$\cdots$	μ_{ab}

  - 1 parameter μ
  - a parameters for A main effects: $\alpha_1, \alpha_2, \ldots, \alpha_a$
  - b parameters for B main effects: $\beta_1, \beta_2, \ldots, \beta_b$
  - ab parameters for AB interactions:

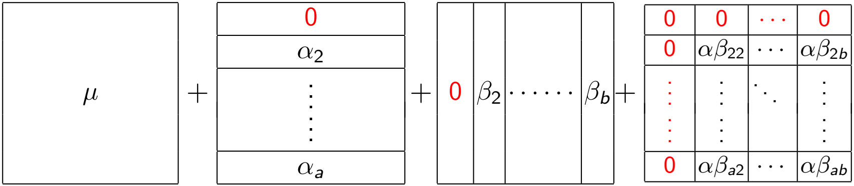
$\alpha\beta_{11}$	$\alpha\beta_{12}$	$\cdots$	$\alpha\beta_{1b}$
$\alpha\beta_{21}$	$\alpha\beta_{22}$	$\cdots$	$\alpha\beta_{2b}$
$\vdots$	$\vdots$	$\ddots$	$\vdots$
$\alpha\beta_{a1}$	$\alpha\beta_{a2}$	$\cdots$	$\alpha\beta_{ab}$

- Two equivalent models should have **identical numbers of parameters**. The main-effect-interaction model is overparameterized, meaning its parameters cannot be uniquely determined unless we set **constraints** on them.

Baseline Constraints (1)

R by default use the baseline constraints by setting all the parameters for the first level of a factor/interaction to 0,

$$\alpha_1 = 0, \quad \beta_1 = 0, \quad \alpha\beta_{1j} = \alpha\beta_{i1} = 0 \quad \text{for all } i, j.$$



So effectively, there are

- 1 parameter μ,
- $a - 1$ parameters for A main effects,
- $b - 1$ parameters for B main effects,
- $(a-1)(b-1)$ parameters for AB interactions.

In total, there are $1 + (a-1) + (b-1) + (a-1)(b-1) = ab$ parameters, same as the means model.

Baseline Constraints (2)

Under the baseline constraint: $\alpha_1 = 0$, $\beta_1 = 0$, $\alpha\beta_{1j} = \alpha\beta_{i1} = 0$

- $\mu_{11} = \mu + \underbrace{\alpha_1}_{=0} + \underbrace{\beta_1}_{=0} + \underbrace{\alpha\beta_{11}}_{=0} \quad \Rightarrow \quad \mu = \mu_{11}$
- $\mu_{i1} = \mu + \alpha_i + \underbrace{\beta_1}_{=0} + \underbrace{\alpha\beta_{i1}}_{=0} \quad \Rightarrow \quad \alpha_i = \mu_{i1} - \mu = \mu_{i1} - \mu_{11}$
  - So α_i = effect of changing factor A from level 1 to level i on the mean of y, when factor B is fixed at level 1
  - Level 1 is the baseline level of factor A
- $\mu_{1j} = \mu + \underbrace{\alpha_1}_{=0} + \beta_j + \underbrace{\alpha\beta_{1j}}_{=0} \quad \Rightarrow \quad \beta_j = \mu_{1j} - \mu = \mu_{1j} - \mu_{11}$
  - So β_j = effect of changing factor B from level 1 to level j on the mean of y, when factor A is fixed at level 1
  - level 1 is the baseline level of factor B
- $$\begin{aligned} \alpha\beta_{ij} &= \mu_{ij} - \mu - \alpha_i - \beta_j \\ &= \mu_{ij} - \mu_{11} - (\mu_{i1} - \mu_{11}) - (\mu_{1j} - \mu_{11}) \\ &= \mu_{ij} - \mu_{i1} - \mu_{1j} + \mu_{11} \end{aligned}$$

Zero-Sum Constraints (1)

For factorial data, the more commonly used constraints are the zero-sum constraints:

$$\sum_{i=1}^{a} \alpha_i = 0, \quad \sum_{j=1}^{b} \beta_j = 0, \quad \sum_{i=1}^{a} \alpha\beta_{ij} = 0 \text{ for all } j, \text{ and } \sum_{j=1}^{b} \alpha\beta_{ij} = 0 \text{ for all } i.$$

I.e., the row sums and column sums of the array $\{\alpha\beta_{ij}\}$ are all 0.

					sum
	$\alpha\beta_{11}$	$\alpha\beta_{12}$	$\cdots$	$\alpha\beta_{1b}$	0
	$\alpha\beta_{21}$	$\alpha\beta_{22}$	$\cdots$	$\alpha\beta_{2b}$	0
	$\vdots$	$\vdots$	$\ddots$	$\vdots$	$\vdots$
	$\alpha\beta_{a1}$	$\alpha\beta_{a2}$	$\cdots$	$\alpha\beta_{ab}$	0
sum	0	0	$\cdots$	0	

So effectively, there are

- 1 parameter μ,
- $a-1$ parameters for A main effects since $\alpha_a = -\sum_{i=1}^{a-1} \alpha_i$,
- $b-1$ parameters for B main effects since $\beta_b = -\sum_{j=1}^{b-1} \beta_j$,
- $(a-1)(b-1)$ parameters for AB interactions since the last row and the last column of the $\{\alpha\beta_{ij}\}$ array can be determined from the zero-sum constraint.

In total, there are $1 + (a-1) + (b-1) + (a-1)(b-1) = ab$ parameters, same as the means model.

Zero-Sum Constraints (2)

Since $\mu_{ij} = \mu + \alpha_i + \beta_j + \alpha\beta_{ij}$, summing them over i, we get

$$\sum_{i=1}^{a} \mu_{ij} = a\mu + \underbrace{\sum\nolimits_{i=1}^{a} \alpha_i}_{=0} + a\beta_j + \underbrace{\sum\nolimits_{i=1}^{a} \alpha\beta_{ij}}_{=0}, \Rightarrow \ \mu + \beta_j = \frac{1}{a}\sum_{i=1}^{a} \mu_{ij} = \overline{\mu}_{\bullet j}$$

Likewise, summing $\mu_{ij} = \mu + \alpha_i + \beta_j + \alpha\beta_{ij}$ over j, we get

$$\sum_{j=1}^{b} \mu_{ij} = b\mu + b\alpha_i + \underbrace{\sum\nolimits_{j=1}^{b} \beta_j}_{=0} + \underbrace{\sum\nolimits_{j=1}^{b} \alpha\beta_{ij}}_{=0}, \Rightarrow \ \mu + \alpha_i = \frac{1}{b}\sum_{j=1}^{b} \mu_{ij} = \overline{\mu}_{i\bullet}$$

Summing $\mu + \alpha_i = \overline{\mu}_{\bullet j}$ over i, we get

$$a\mu + \underbrace{\sum\nolimits_{i=1}^{a} \alpha_i}_{=0} = \sum_{i=1}^{a} \overline{\mu}_{\bullet j} = \sum_{i=1}^{a} \frac{1}{b} \sum_{j=1}^{b} \mu_{ij} \ \Rightarrow \ \mu = \frac{1}{ab}\sum_{i=1}^{a}\sum_{j=1}^{b} \mu_{ij} = \overline{\mu}_{\bullet\bullet}$$

Zero-Sum Constraints (3)

Under the zero-sum constraint, parameters in the means model and the main-effect-interaction model are related as follows

$$
\begin{aligned}
\mu &= \overline{\mu}_{\bullet\bullet} = \text{overall mean} \\
\alpha_i &= \overline{\mu}_{i\bullet} - \overline{\mu}_{\bullet\bullet} = \text{row mean} - \text{overall mean} \\
\beta_j &= \overline{\mu}_{\bullet j} - \overline{\mu}_{\bullet\bullet} = \text{column mean} - \text{overall mean} \\
\alpha\beta_{ij} &= \mu_{ij} - \mu - \alpha_i - \beta_j \\
&= \mu_{ij} - \overline{\mu}_{\bullet\bullet} - (\overline{\mu}_{i\bullet} - \overline{\mu}_{\bullet\bullet}) - (\overline{\mu}_{\bullet j} - \overline{\mu}_{\bullet\bullet}) \\
&= \mu_{ij} - \overline{\mu}_{i\bullet} - \overline{\mu}_{\bullet j} + \overline{\mu}_{\bullet\bullet} \\
&= \text{cell mean} - \text{row mean} - \text{column mean} + \text{overall mean}
\end{aligned}
$$

					row mean
	μ_{11}	μ_{12}	$\cdots$	μ_{1b}	$\overline{\mu}_{1\bullet}$
	μ_{21}	μ_{22}	$\cdots$	μ_{2b}	$\overline{\mu}_{2\bullet}$
	$\vdots$	$\vdots$		$\vdots$	$\vdots$
	μ_{a1}	μ_{a2}	$\cdots$	μ_{ab}	$\overline{\mu}_{a\bullet}$
column mean	$\overline{\mu}_{\bullet 1}$	$\overline{\mu}_{\bullet 2}$	$\cdots$	$\overline{\mu}_{\bullet b}$	$\overline{\mu}_{\bullet\bullet}$ = overall mean

Estimation of Parameters under the Zero-Sum Constraint (1)

Parameter estimation in a balanced factorial design under the Zero-Sum constraint is straightforward. For

- $y_{ijk} = \mu_{ij} + \varepsilon_{ijk}$ .... (means model)
- $y_{ijk} = \mu + \alpha_i + \beta_j + \alpha\beta_{ij} + \varepsilon_{ijk}$ (main-effect-interaction model)

the parameter estimates are

$$\begin{aligned}
\widehat{\mu}_{ij} &= \overline{y}_{ij\bullet} \\
\widehat{\mu} &= \overline{y}_{\bullet\bullet\bullet}, \\
\widehat{\alpha}_i &= \overline{y}_{i\bullet\bullet} - \overline{y}_{\bullet\bullet\bullet}, \\
\widehat{\beta}_j &= \overline{y}_{\bullet j\bullet} - \overline{y}_{\bullet\bullet\bullet} \\
\widehat{\alpha\beta}_{ij} &= \overline{y}_{ij\bullet} - \overline{y}_{i\bullet\bullet} - \overline{y}_{\bullet j\bullet} + \overline{y}_{\bullet\bullet\bullet}
\end{aligned}$$

Observe the estimates satisfy the zero-sum constraints:

$$\sum_{i=1}^{a} \widehat{\alpha}_i = \sum_{j=1}^{b} \widehat{\beta}_j = \sum_{i=1}^{a} \widehat{\alpha\beta}_{ij} = \sum_{j=1}^{a} \widehat{\alpha\beta}_{ij} = 0, \quad \text{for all } i, j.$$

Estimation of Parameters under the Zero-Sum Constraint (2)

Since the design is balanced, for any of the reduced models below,

- $y_{ijk} = \mu + \varepsilon_{ijk}$ ....(no main effects, no interaction)
- $y_{ijk} = \mu + \alpha_i + \varepsilon_{ijk}$ ....(main effects of A only)
- $y_{ijk} = \mu + \beta_j + \varepsilon_{ijk}$ ....(main effects of B only)
- $y_{ijk} = \mu + \alpha_i + \beta_j + \varepsilon_{ijk}$ ....(additive model)

the estimates of μ, α_i's, and β_j's under the zero-sum constraints are **identical** with those for the main-effects-interaction model:

$$\widehat{\mu} = \overline{y}_{\bullet\bullet\bullet}, \quad \widehat{\alpha}_i = \overline{y}_{i\bullet\bullet} - \overline{y}_{\bullet\bullet\bullet}, \quad \widehat{\beta}_j = \overline{y}_{\bullet j\bullet} - \overline{y}_{\bullet\bullet\bullet}$$

If NOT balanced, the estimates will change with the model.

Recall in a regression model, the estimate of a coefficient will change with the presence of other covariates in the model

Parameter Estimates Under the Baseline Constraints (May Skip)

Under the baseline constraints,

$$\alpha_1 = 0, \quad \beta_1 = 0, \quad \alpha\beta_{1j} = \alpha\beta_{i1} = 0 \quad \text{for all } i, j.$$

the least-square estimates for parameters in the 5 models are different (see below), even if the data is balanced.

Model Formula	Parameter Estimates			
	$\widehat{\mu}$	$\widehat{\alpha}_i$	$\widehat{\beta}_j$	$\widehat{\alpha\beta}_{ij}$
$\mu + \alpha_i + \beta_j + \alpha\beta_{ij}$	$\overline{y}_{11\bullet}$	$\overline{y}_{i1\bullet} - \overline{y}_{11\bullet}$	$\overline{y}_{1j\bullet} - \overline{y}_{11\bullet}$	$\overline{y}_{ij\bullet} - \overline{y}_{i1\bullet} - \overline{y}_{1j\bullet} + \overline{y}_{11\bullet}$
$\mu + \alpha_i + \beta_j$	$\overline{y}_{1\bullet\bullet} + \overline{y}_{\bullet 1\bullet} - \overline{y}_{\bullet\bullet\bullet}$	$\overline{y}_{i\bullet\bullet} - \overline{y}_{1\bullet\bullet}$	$\overline{y}_{\bullet j\bullet} - \overline{y}_{\bullet 1\bullet}$	–
$\mu + \alpha_i$	$\overline{y}_{1\bullet\bullet}$	$\overline{y}_{i\bullet\bullet} - \overline{y}_{1\bullet\bullet}$	–	–
$\mu + \beta_j$	$\overline{y}_{\bullet 1\bullet}$	–	$\overline{y}_{\bullet j\bullet} - \overline{y}_{\bullet 1\bullet}$	–
μ	$\overline{y}_{\bullet\bullet\bullet}$	–	–	–

- Simplicity in the formulas of parameter estimates is the primary reason we mostly use the zero-sum constraints for factorial data, even though R uses the baseline constraints
- **Don't memorize the formulas for the baseline constraints!**
- Models are not affected by the constraints imposed. The fitted values, residuals, df, SSE are not affected.

Fitted Values for a Main-Effect-Interaction Model

For a main-effect-interaction model, the fitted value for y_{ijk} under the zero-sum constraints is

$$\begin{aligned}
\widehat{y}_{ijk} &= \widehat{\mu} + \widehat{\alpha}_i + \widehat{\beta}_j + \widehat{\alpha\beta}_{ij} \\
&= \overline{y}_{\bullet\bullet\bullet} + (\overline{y}_{i\bullet\bullet} - \overline{y}_{\bullet\bullet\bullet}) + (\overline{y}_{\bullet j\bullet} - \overline{y}_{\bullet\bullet\bullet}) \\
&\quad + (\overline{y}_{ij\bullet} - \overline{y}_{i\bullet\bullet} - \overline{y}_{\bullet j\bullet} + \overline{y}_{\bullet\bullet\bullet}) \\
&= \overline{y}_{ij\bullet} = \text{cell mean}
\end{aligned}$$

which is equal to the fitted value under the baseline constraints:

$$\begin{aligned}
\widehat{y}_{ijk} &= \widehat{\mu} + \widehat{\alpha}_i + \widehat{\beta}_j + \widehat{\alpha\beta}_{ij} \\
&= \overline{y}_{11\bullet} + (\overline{y}_{i1\bullet} - \overline{y}_{11\bullet}) + (\overline{y}_{1j\bullet} - \overline{y}_{11\bullet}) \\
&\quad + \overline{y}_{ij\bullet} - \overline{y}_{i1\bullet} - \overline{y}_{1j\bullet} + \overline{y}_{11\bullet} \\
&= \overline{y}_{ij\bullet} = \text{cell mean.}
\end{aligned}$$

Fitted Values for an Additive Model

For an additive model (**no interaction**), the fitted value for y_{ijk} under the zero-sum constraints is

$$\begin{aligned}
\widehat{y}_{ijk} &= \widehat{\mu} + \widehat{\alpha}_i + \widehat{\beta}_j \\
&= \overline{y}_{\bullet\bullet\bullet} + (\overline{y}_{i\bullet\bullet} - \overline{y}_{\bullet\bullet\bullet}) + (\overline{y}_{\bullet j\bullet} - \overline{y}_{\bullet\bullet\bullet}) \\
&= \overline{y}_{i\bullet\bullet} + \overline{y}_{\bullet j\bullet} - \overline{y}_{\bullet\bullet\bullet} \\
&= \text{row mean} + \text{column mean} - \text{overall mean}
\end{aligned}$$

which is equal to the fitted value under the baseline constraints:

$$\begin{aligned}
\widehat{y}_{ijk} &= \widehat{\mu} + \widehat{\alpha}_i + \widehat{\beta}_j \\
&= (\overline{y}_{1\bullet\bullet} + \overline{y}_{\bullet 1\bullet} - \overline{y}_{\bullet\bullet\bullet}) + (\overline{y}_{i\bullet\bullet} - \overline{y}_{1\bullet\bullet}) + (\overline{y}_{\bullet j\bullet} - \overline{y}_{\bullet 1\bullet}) \\
&= \overline{y}_{i\bullet\bullet} + \overline{y}_{\bullet j\bullet} - \overline{y}_{\bullet\bullet\bullet} \\
&= \text{row mean} + \text{column mean} - \text{overall mean}
\end{aligned}$$

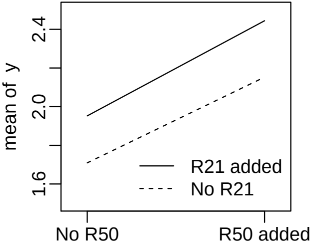
Example 8.6 Bacteria in Cheese (p.178 in Oehlert)

- Factor A: Bacteria R50#10, added or not
- Factor B: Bacteria R21#2, added or not
- 3 replicates
- Response: total free amino acids in cheddar cheese after 56 days of ripening.

	No R21	R21 added
No R50	1.697 1.601 1.830	2.211 1.673 1.973
R50 added	2.032 2.017 2.409	2.091 2.255 2.987

⇒

	No R21	R21 added
No R50	$\overline{y}_{11\bullet} = 1.709$	$\overline{y}_{12\bullet} = 1.952$
R50 added	$\overline{y}_{21\bullet} = 2.153$	$\overline{y}_{22\bullet} = 2.444$



Is there interaction?

Example 8.6 Bacteria in Cheese (p.178 in Oehlert)

	B-level 1	B-level 2	row mean
A-level 1	$\overline{y}_{11\bullet} = 1.709$	$\overline{y}_{12\bullet} = 1.952$	$\overline{y}_{1\bullet\bullet} = 1.831$
A-level 2	$\overline{y}_{21\bullet} = 2.153$	$\overline{y}_{22\bullet} = 2.444$	$\overline{y}_{2\bullet\bullet} = 2.299$
column mean	$\overline{y}_{\bullet 1\bullet} = 1.931$	$\overline{y}_{\bullet 2\bullet} = 2.198$	$\overline{y}_{\bullet\bullet\bullet} = 2.065$

$$\begin{aligned}
\widehat{\mu} &= \overline{y}_{\bullet\bullet\bullet} = 2.065 \\
\widehat{\alpha}_1 &= \overline{y}_{1\bullet\bullet} - \overline{y}_{\bullet\bullet\bullet} = 1.831 - 2.065 = -0.234 \\
\widehat{\beta}_1 &= \overline{y}_{\bullet 1\bullet} - \overline{y}_{\bullet\bullet\bullet} = 1.931 - 2.065 = -0.134 \\
\widehat{\alpha\beta}_{11} &= \overline{y}_{11\bullet} - \overline{y}_{1\bullet\bullet} - \overline{y}_{\bullet 1\bullet} + \overline{y}_{\bullet\bullet\bullet} \\
&= 1.709 - 1.831 - 1.931 + 2.065 = 0.012
\end{aligned}$$

The estimates of all other parameters can be computed using the zero-sum constraints.

$$\begin{aligned}
\widehat{\alpha}_1 + \widehat{\alpha}_2 = 0 \quad &\Rightarrow \quad \widehat{\alpha}_2 = -\widehat{\alpha}_1 = 0.234 \\
\widehat{\beta}_1 + \widehat{\beta}_2 = 0 \quad &\Rightarrow \quad \widehat{\beta}_2 = -\widehat{\beta}_1 = 0.134 \\
\widehat{\alpha\beta}_{11} + \widehat{\alpha\beta}_{12} = 0 \quad &\Rightarrow \quad \widehat{\alpha\beta}_{12} = -\widehat{\alpha\beta}_{11} = -0.012 \\
\widehat{\alpha\beta}_{11} + \widehat{\alpha\beta}_{21} = 0 \quad &\Rightarrow \quad \widehat{\alpha\beta}_{21} = -\widehat{\alpha\beta}_{11} = -0.012 \\
\widehat{\alpha\beta}_{12} + \widehat{\alpha\beta}_{22} = 0 \quad &\Rightarrow \quad \widehat{\alpha\beta}_{22} = -\widehat{\alpha\beta}_{12} = 0.012
\end{aligned}$$

Finding Parameter Estimates in R

Note that R finding parameter estimates using the **baseline constraints** by default.

```
> cheese = read.table(
    "http://users.stat.umn.edu/~gary/book/fcdae.data/exmpl8.6",h=T)
> cheese$r50 = as.factor(cheese$r50)
> cheese$r21 = as.factor(cheese$r21)
> lmcheese = lm(y ~ r50 + r21 + r50*r21, data=cheese)
> lmcheese$coef
(Intercept)        r502        r212   r502:r212
 1.70933333  0.44333333  0.24300000  0.04866667
```

$$\widehat{\mu} = \overline{y}_{11\bullet} \approx 1.709$$

$$\widehat{\alpha}_2 = \overline{y}_{21\bullet} - \overline{y}_{11\bullet} \approx 2.153 - 1.709 = 0.444$$

$$\widehat{\beta}_2 = \overline{y}_{12\bullet} - \overline{y}_{11\bullet} \approx 1.952 - 1.709 = 0.243$$

$$\widehat{\alpha}_{22} = \overline{y}_{22\bullet} - \overline{y}_{21\bullet} - \overline{y}_{12\bullet} + \overline{y}_{11\bullet}$$

$$\approx 2.444 - 2.153 - 1.952 + 1.709 = 0.048$$

	R21	
R50	No (B=1)	added (B=2)
No (A=1)	$\overline{y}_{11\bullet} = 1.709$	$\overline{y}_{12\bullet} = 1.952$
added (A=2)	$\overline{y}_{21\bullet} = 2.153$	$\overline{y}_{22\bullet} = 2.444$

and $\widehat{\alpha}_1 = \widehat{\beta}_1 = \widehat{\alpha\beta}_{11} = \widehat{\alpha\beta}_{12} = \widehat{\alpha\beta}_{21} = 0$.

How to Force R Using the Zero-Sum Constraints?

To force R using the zero-sum constraints, one needs to set the following

```
contrasts(cheese$r50) = contr.sum(2)
contrasts(cheese$r21) = contr.sum(2)
```

where the number 2 inside `contr.sum(2)` is the number of levels for the factor.

Next, one need to fit the `lm()` model again to update the coefficient.

```
> lmcheese = lm(y ~ r50 + r21 + r50*r21, data=cheese)
> lmcheese$coef
(Intercept)        r501        r211   r501:r211
 2.06466667 -0.23383333 -0.13366667  0.01216667
```

We get $\widehat{\mu} \approx 2.065$, $\widehat{\alpha}_1 \approx -0.234$, $\widehat{\beta}_1 \approx -0.134$, $\widehat{\alpha\beta}_{11} \approx 0.012$ which match our calculations. Estimates for other parameters can be determined by the zero-sum constraints.

Sum of Squares for Balanced 2-Way Factorial Designs (1)

An balanced $a \times b$ two-way factorial design with n replicates is also a CRD with ab treatments, so the sum of squares identity is still valid.

$$SST = SS_{trt} + SSE$$

where

$$SST = \sum_{i=1}^{a} \sum_{j=1}^{b} \sum_{k=1}^{n} (y_{ijk} - \overline{y}_{\bullet\bullet\bullet})^2 \quad \text{and}$$

$$SS_{trt} = n \sum_{i=1}^{a} \sum_{j=1}^{b} (\overline{y}_{ij\bullet} - \overline{y}_{\bullet\bullet\bullet})^2, \quad SSE = \sum_{i=1}^{a} \sum_{j=1}^{b} \sum_{k=1}^{n} (y_{ijk} - \overline{y}_{ij\bullet})^2$$

$$\begin{aligned}
\text{d.f. for SST} &= \text{total \# of observations} - 1 = abn - 1 \\
\text{d.f. for SS}_{trt} &= \text{\# of treatments} - 1 = ab - 1 \\
\text{d.f. for SSE} &= \text{total \# of observations} - \text{\# of treatments} \\
&= abn - ab = ab(n-1)
\end{aligned}$$

Sum of Squares for Balanced 2-Way Factorial Designs (2)

As the *ab* treatments have a factorial structure, SS_{trt} can be decomposed further as

$$SS_{trt} = SS_A + SS_B + SS_{AB}$$

in which

SS	formula	d.f.
SS_A	$n\sum_{i=1}^{a}\sum_{j=1}^{b}(\bar{y}_{i\bullet\bullet} - \bar{y}_{\bullet\bullet\bullet})^2$	$a-1$
SS_B	$n\sum_{i=1}^{a}\sum_{j=1}^{b}(\bar{y}_{\bullet j\bullet} - \bar{y}_{\bullet\bullet\bullet})^2$	$b-1$
SS_{AB}	$n\sum_{i=1}^{a}\sum_{j=1}^{b}(\bar{y}_{ij\bullet} - \bar{y}_{i\bullet\bullet} - \bar{y}_{\bullet j\bullet} + \bar{y}_{\bullet\bullet\bullet})^2$	$(a-1)(b-1)$
SS_{trt}	$n\sum_{i=1}^{a}\sum_{j=1}^{b}(\bar{y}_{ij\bullet} - \bar{y}_{\bullet\bullet\bullet})^2$	$ab-1$

Observe all the d.f.s for the SS of the main effects or interactions equal (number of parameters) − (number of constraint(s))

Sum of Squares for Balanced 2-Way Factorial Designs (3)

In summary

$$SST = SS_A + SS_B + SS_{AB} + SSE$$

$$SST = \sum_{i=1}^{a}\sum_{j=1}^{b}\sum_{k=1}^{n}(y_{ijk} - \overline{y}_{\bullet\bullet\bullet})^2$$

$$SS_A = \sum_{i=1}^{a}\sum_{j=1}^{b}\sum_{k=1}^{n}(\underbrace{\overline{y}_{i\bullet\bullet} - \overline{y}_{\bullet\bullet\bullet}}_{\widehat{\alpha}_i})^2 = bn\sum_{i=1}^{a}\widehat{\alpha}_i^2$$

$$SS_B = \sum_{i=1}^{a}\sum_{j=1}^{b}\sum_{k=1}^{n}(\underbrace{\overline{y}_{\bullet j\bullet} - \overline{y}_{\bullet\bullet\bullet}}_{\widehat{\beta}_j})^2 = an\sum_{j=1}^{b}\widehat{\beta}_j^2$$

$$SS_{AB} = \sum_{i=1}^{a}\sum_{j=1}^{b}\sum_{k=1}^{n}(\underbrace{\overline{y}_{ij\bullet} - \overline{y}_{i\bullet\bullet} - \overline{y}_{\bullet j\bullet} + \overline{y}_{\bullet\bullet\bullet}}_{\widehat{\alpha\beta}_{ij}})^2 = n\sum_{i=1}^{a}\sum_{j=1}^{b}\widehat{\alpha\beta}_{ij}^2$$

$$SSE = \sum_{i=1}^{a}\sum_{j=1}^{b}\sum_{k=1}^{n}(y_{ijk} - \overline{y}_{ij\bullet})^2$$

ANOVA Table for Balanced Two-Way Factorial Designs

Source	d.f.	SS	MS	F
Factor A	$a-1$	SS_A	$MS_A = \frac{SS_A}{a-1}$	$F_A = \frac{MS_A}{MSE}$
Factor B	$b-1$	SS_B	$MS_B = \frac{SS_B}{b-1}$	$F_B = \frac{MS_B}{MSE}$
AB Interaction	$(a-1)(b-1)$	SS_{AB}	$MS_{AB} = \frac{SS_{AB}}{(a-1)(b-1)}$	$F_{AB} = \frac{MS_{AB}}{MSE}$
Error	$ab(n-1)$	SSE	$MSE = \frac{SSE}{ab(n-1)}$	
Total	$abn-1$	SST		

Questions of Interest in a 2-Way Factorial Design

1. Does factor A has an effect on the response?
   E.g. does the age of seeds has an effect on germination?

$$\begin{cases} H_0: \alpha_1 = \cdots = \alpha_a = 0 \\ H_a: \text{not all } \alpha_i\text{'s} = 0, \end{cases} \Rightarrow F_A = \frac{\text{MS}_A}{\text{MSE}} \sim F_{a-1,\, ab(n-1)} \text{ under } \text{H}_0.$$

2. Does factor B has an effect on the response?
   E.g. does the water amount has an effect on germination?

$$\begin{cases} H_0: \beta_1 = \cdots = \beta_b = 0 \\ H_a: \text{not all } \beta_i\text{'s} = 0, \end{cases} \Rightarrow F_B = \frac{\text{MS}_B}{\text{MSE}} \sim F_{b-1,\, ab(n-1)} \text{ under } \text{H}_0.$$

3. Does the effect of factor A interact with that of factor B?
   E.g., does the effect of age change with water amount?

$$\begin{cases} H_0: \alpha\beta_{ij} = 0 \text{ for all } i, j \\ H_a: \alpha\beta_{ij} \neq 0 \text{ for some } i, j \end{cases} \Rightarrow F_{AB} = \frac{\text{MS}_{AB}}{\text{MSE}} \sim F_{(a-1)(b-1),\, ab(n-1)} \text{ under } \text{H}_0.$$

Example 8.6 Bacteria in Cheese (p.178 in Oehlert)

$$SS_A = bn \sum_{i=1}^{a} \widehat{\alpha}_i^2 = 2 \times 3 \times [(-0.234)^2 + 0.234^2] = 0.656$$

$$SS_B = an \sum_{j=1}^{b} \widehat{\beta}_j^2 = 2 \times 3 \times [(-0.134)^2 + 0.134^2] = 0.214$$

$$SS_{AB} = n \sum_{i=1}^{a} \sum_{j=1}^{b} \widehat{\alpha\beta}_{ij}^2 = 3 \times [0.012^2 \times 4] \approx 0.0017$$

Computing SSE needs more work. It is easier to compute the SST:

$$\begin{aligned} SST &= \sum_{i=1}^{a} \sum_{j=1}^{b} \sum_{k=1}^{n} (y_{ijk} - \overline{y}_{\bullet\bullet\bullet})^2 \\ &= (1.697 - 2.065)^2 + (1.601 - 2.065)^2 + (1.830 - 2.065)^2 \\ &\quad + \cdots + (2.987 - 2.065)^2 = 1.598 \end{aligned}$$

Then we can get

$$\begin{aligned} SSE &= SST - SS_A - SS_B - SS_{AB} \\ &= 1.598 - 0.656 - 0.214 - 0.0018 = 0.726. \end{aligned}$$

Example 8.6 Bacteria in Cheese — ANOVA table

Source	d.f.	SS	MS	F-value	P-value
A(R50)	1	0.656	0.656	7.23	0.028
B(R21)	1	0.214	0.214	2.36	0.16
AB interaction	1	0.0017	0.0017	0.019	0.89
Error	8	0.726	0.091		
Total	11	1.598			

Only main effect A (Bacteria R50) is moderately significant.
Main effect B and interaction are not.

One can also get the ANOVA table in R as follows.

```
> lmcheese = lm(y ~ r50 + r21 + r50*r21, data=cheese)
> anova(lmcheese)
Analysis of Variance Table

Response: y
          Df  Sum Sq Mean Sq F value  Pr(>F)
r50        1 0.65614 0.65614  7.2335 0.02752 *
r21        1 0.21440 0.21440  2.3636 0.16275
r50:r21    1 0.00178 0.00178  0.0196 0.89217
Residuals  8 0.72566 0.09071
```

Display of Data from Two Way Factorial Designs

y_{ijk}	Age of Seeds (weeks) 1	3	6	9	12
water 4(ml)	11	7	9	13	20
	9	16	19	35	37
	6	17	35	28	45
water 8(ml)	8	1	5	1	11
	3	7	9	10	15
	3	3	9	9	25

Cell means $\overline{y}_{ij\bullet}$	Age of Seeds (weeks) 1	3	6	9	12	Row means $\overline{y}_{i\bullet\bullet}$
water 4(ml)	8.67	13.33	21.00	25.33	34.00	20.47
water 8(ml)	4.67	3.67	7.67	6.67	17.00	7.93
Column means $\overline{y}_{\bullet j\bullet}$	6.67	8.50	14.33	16.00	25.50	$\overline{y}_{\bullet\bullet\bullet} = 14.2$ overall mean

Does water have an effect on gemination?

Does the age of seeds have an effect?

Finding Row Means, Column Means, Cell Means in R

Overall mean $\overline{y}_{\bullet\bullet\bullet} = \widehat{\mu}$:

```
> library(mosaic)
> mean(~y, data=barley)
[1] 14.2
```

Row means $\overline{y}_{i\bullet\bullet}$:

```
> mean(y ~ water, data=barley)
        4         8
20.466667  7.933333
```

Column means $\overline{y}_{\bullet j\bullet}$:

```
> mean(y ~ week, data=barley)
        1         3         6         9        12
 6.666667  8.500000 14.333333 16.000000 25.500000
```

Cell means ($\overline{y}_{ij\bullet}$, average of the 3 values in each cell):

```
> mean(y ~ week+water, data=barley)
      1.4       3.4       6.4       9.4      12.4       1.8       3.8
 8.666667 13.333333 21.000000 25.333333 34.000000  4.666667  3.666667
      6.8       9.8      12.8
 7.666667  6.666667 17.000000
```

$\widehat{\alpha}_i = \overline{y}_{i\bullet\bullet} - \overline{y}_{\bullet\bullet\bullet}$:

```
> mean(y ~ water, data=barley)-mean(~y, data=barley)
        4         8
 6.266667 -6.266667
```

$\widehat{\beta}_j = \overline{y}_{\bullet j\bullet} - \overline{y}_{\bullet\bullet\bullet}$:

```
> mean(y ~ week, data=barley)-mean(~y, data=barley)
         1          3          6          9         12
-7.5333333 -5.7000000  0.1333333  1.8000000 11.3000000
```

$\widehat{\alpha\beta}_{ij} = \overline{y}_{ij\bullet} - \overline{y}_{i\bullet\bullet} - \overline{y}_{\bullet j\bullet} + \overline{y}_{\bullet\bullet\bullet}$:

```
> cell.mean = matrix(mean(y ~ week+water, data=barley),nrow=2,byrow=T)
> cell.mean
         [,1]      [,2]      [,3]      [,4] [,5]
[1,] 8.666667 13.333333 21.000000 25.333333   34
[2,] 4.666667  3.666667  7.666667  6.666667   17
> row.mean = mean(y ~ water, data=barley)%o%rep(1,5); row.mean
       [,1]      [,2]      [,3]      [,4]      [,5]
4 20.466667 20.466667 20.466667 20.466667 20.466667
8  7.933333  7.933333  7.933333  7.933333  7.933333
> column.mean = rep(1,2)%o%mean(y ~ week, data=barley);column.mean
            1   3        6  9   12
[1,] 6.666667 8.5 14.33333 16 25.5
[2,] 6.666667 8.5 14.33333 16 25.5
> overall.mean = mean(~y, data=barley)
> cell.mean - row.mean - column.mean + overall.mean
       [,1]      [,2] [,3]      [,4]      [,5]
4 -4.266667 -1.433333  0.4  3.066667  2.233333
8  4.266667  1.433333 -0.4 -3.066667 -2.233333
```

```
> barley$weekfac = as.factor(barley$week)
> barley$waterfac = as.factor(barley$water)
> contrasts(barley$weekfac) = contr.sum(5)
> contrasts(barley$waterfac) = contr.sum(2)
> lmbarley = lm(y ~ waterfac + weekfac + waterfac*weekfac, data=barley)
> lmbarley$coef
      (Intercept)         waterfac1          weekfac1
       14.2000000         6.2666667        -7.5333333
         weekfac2          weekfac3          weekfac4
       -5.7000000         0.1333333         1.8000000
waterfac1:weekfac1 waterfac1:weekfac2 waterfac1:weekfac3
       -4.2666667        -1.4333333         0.4000000
waterfac1:weekfac4
        3.0666667
```

Observe that we get

$$\begin{array}{llll} \widehat{\mu} = 14.2, & \widehat{\alpha}_1 \approx 6.267, & & \\ \widehat{\beta}_1 \approx -7.533, & \widehat{\beta}_2 = -5.7, & \widehat{\beta}_3 \approx 0.13,3 & \widehat{\beta}_4 = 1.8, \\ \widehat{\alpha\beta}_{11} \approx -4.266, & \widehat{\alpha\beta}_{12} \approx -1.433, & \widehat{\alpha\beta}_{13} = 0.4 & \widehat{\alpha\beta}_{14} \approx 3.067 \end{array}$$

which match our calculations earlier. Estimates for other parameters can be determined by the zero-sum constraints.

Problem 8.1 — Sprouting Barley — ANOVA Table

```
> anova(lmbarley)
Analysis of Variance Table

Response: y
                Df  Sum Sq Mean Sq F value   Pr(>F)
waterfac         1 1178.13 1178.13 19.7232 0.000251 ***
weekfac          4 1321.13  330.28  5.5293 0.003645 **
waterfac:weekfac  4  208.87   52.22  0.8742 0.496726
Residuals       20 1194.67   59.73
```

Conclusion:

- It looks like both water and week main effects are significant, but their interactions are not
- Wait! Need to check model assumptions before making conclusions.

Factorial Designs v.s. One-At-a-Time Designs

When there are two factors A and B of interest, we could conduct two separate experiments and change only one factor at a time,

Experiment #1

A = 1	A = 2	A = 3	A = 4

Experiment #2

B = 1	B = 2	B = 3

rather than a two-way factorial design

	A = 1	A = 2	A = 3	A = 4
B= 1				
B= 2				
B= 3				

Advantage of Factorial Designs

Factorial designs are superior to one-at-a-time designs that change only one factor at a time because factorial design can

- test the effects of both factors at once — more *efficient* than one-at-a-time design, taking fewer experimental units to attain the same goal;
- investigate *interaction* of factors, but one-at-a-time designs cannot.
- broaden the inductive base for generalizing our results by trying a wide range of treatments