

## CHAPTER 19

### Two Factor Studies, Balanced case

We'll consider experimental studies first.

There are pros and cons to including two factors in the same study, rather than taking the "one factor at a time approach" (OFAAT).

One pro is efficiency; you get two experiments in one.

A con is that it's a more complicated approach, therefore more prone to errors and to missing data. Missing data cause more problems with the analysis than in the one-factor design.

106

#### Two-way ANOVA: Meaning of Model Elements

*Ex. Corn Yield*

A field was divided into 20 equal-size plots, each with the same base fertility level, and same exposure to sun and rain; each was planted with the same amount of corn seed in the same manner.

Fertilizer and manure were applied to each plot. At the end of the growing season, the corn was harvested and the yields of the twenty plots were measured.

Goal: Assess how corn yield depends on level of nitrogen-based fertilizer and level of manure used.

In this experiment, two levels of each factor were used (low, high).

$Y$  = yield of corn (metric tons)

Factor 1: Fertilizer (low is 45 kg per hectare, high is 135 kg per hectare)

Factor 2: Manure (low is 84 kg per hectare, high is 168 kg per hectare)

107

*Example: Corn Yield*  $n = 5$  plots per “cell” (A “cell” is a particular combination of the levels of the two factors, e.g. “low, low” means “low fertilizer” and “low manure”). The summary statistics are given in the two tables below.

Cell Means

		Fertilizer	
Manure		Low	High
Low		11.3	13.9
High		14.0	15.1

Cell SDs

		Fertilizer	
Manure		Low	High
Low		1.9	1.7
High		1.8	1.3

**General framework for balanced two-way ANOVA**

Our approach to analysis will allow us to easily accommodate more than two levels per factor.

To start with, for discussion of the model parameters, suppose we know the true mean yields for every treatment (factor-level combination, or “cell” of the design).

### Two-way ANOVA Model

Table of cell/treatment means, factor level means, and main effects:

Factor A	Factor B		Row Average	Main (Row) Effects
	$j = 1$	$j = 2$		
$i = 1$	$\mu_{11}$	$\mu_{12}$	$\mu_{1.}$	$\mu_{1.} - \mu_{..}$
$i = 2$	$\mu_{21}$	$\mu_{22}$	$\mu_{2.}$	$\mu_{2.} - \mu_{..}$
Column average	$\mu_{.1}$	$\mu_{.2}$	$\mu_{..}$	
Main (col) effects	$\mu_{.1} - \mu_{..}$	$\mu_{.2} - \mu_{..}$		

The mean response for a given treatment in a two-factor study is denoted by  $\mu_{ij}$ , where  $i$  is the level of factor A ( $i = 1, \dots, a$ ) and  $j$  is the level of factor B ( $j = 1, \dots, b$ ).

Row averages, column averages, and overall average:

$$\mu_{.i} = \frac{\sum_{j=1}^b \mu_{ij}}{b}, \mu_{.j} = \frac{\sum_{i=1}^a \mu_{ij}}{a}, \mu_{..} = \frac{\sum_{i=1}^a \sum_{j=1}^b \mu_{ij}}{ab}$$

The main row effects are denoted  $\alpha_i$ ,  $i = 1, \dots, a$

The main column effects are denoted  $\beta_j$ ,  $j = 1, \dots, b$

110

*Example 1* Find the factor level means and main effects for the following set of **true** cell means:

Manure	Fertilizer	
	Low	High
Low	11.0	14.0
High	14.0	15.0

111

Example 2 Suppose the true cell means were the following:

Manure	Fertilizer		Row		Main row effect
	Low	High	Average	Average	
Low	12.0	14.0	13.0		-2.0
High	16.0	18.0	17.0		2.0
Column average	14.0	16.0	15.0		
Main col. effect	-1.0	1.0			

In this special case, the factor effects are **additive**. This means, for instance, that each cell mean can be found by adding the respective row and column effects to the overall mean  $\mu_{..}$ .

$$\begin{aligned}\mu_{11} &= \mu_{..} + \alpha_1 + \beta_1 = 15 + (-2) + (-1) \\ \mu_{12} &= \mu_{..} + \alpha_1 + \beta_2 = 15 + (-2) + 1 \\ &\vdots\end{aligned}$$

This also means that  $\mu_{ij} = \mu_{i.} + \mu_{.j} - \mu_{..}$ .

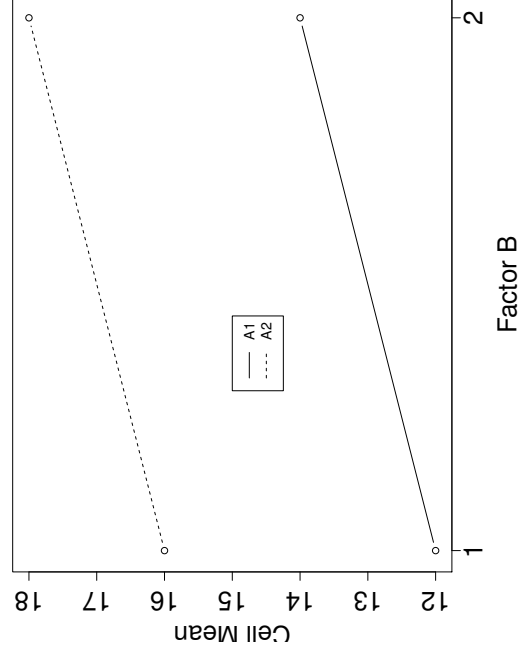
This is an example of an **additive** model. We also say there is **no interaction** between the two factors.

112

### Interaction plot (true cell means)

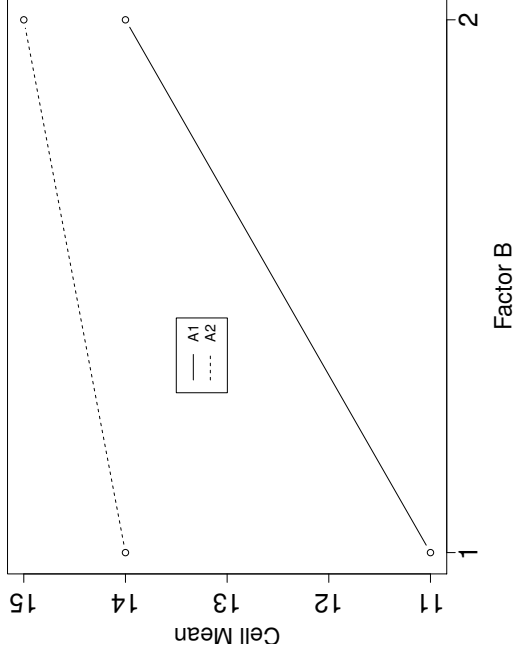
Interaction Plot is of factor level means with levels of one factor represented on the horizontal axis, and means for the levels of the other factor connected by line segments.

Interaction plot, additive model (Example 2):



113

Interaction plot, model with interaction (Example 1):



114

Let  $\alpha_i \equiv \mu_{i.} - \mu_{..}$  represent the  $i^{\text{th}}$  row effect, and  $\beta_j \equiv \mu_{.j} - \mu_{..}$  represent the  $j^{\text{th}}$  column effect:

If for all  $i, j$ , the cell mean  $\mu_{ij} = \mu_{..} + \alpha_i + \beta_j$ , then the factor effects are additive. Otherwise, the factor effects are interacting.

*Interaction of  $i^{\text{th}}$  level of Factor A with  $j^{\text{th}}$  level of Factor B is defined to be*

$$(\alpha\beta)_{ij} = \mu_{ij} - (\mu_{..} + \alpha_i + \beta_j)$$

*Example (p. ??)*

Manure	Fertilizer		Row		Main row effect
	Low	High	Average		
Low	11.0	14.0	12.5		-1.0
High	14.0	15.0	14.5		1.0
Column avg.	12.5	14.5	$\mu_{..} = 13.5$		
Main col. effects	-1.0	1.0			

$$(\alpha\beta)_{11} = 11.0 - (13.5 + (-1.0) + (-1.0)) = 11.0 - 11.5 = -.5$$

$$(\alpha\beta)_{12} = 14.0 - (13.5 + (-1.0) + (1.0)) = 14.0 - 13.5 = .5$$

⋮

115

There are several ways to recognize whether or not interaction is present. All of the following four ways are equivalent:

1. By examining whether all  $\mu_{ij}$  can be expressed as the sums  $\mu_{..} + \alpha_i + \beta_j$ .
2. By examining whether the difference between the mean responses for any two levels of factor B is the same for all levels of factor A.
3. By examining whether the difference between the mean responses for any two levels of factor A is the same for all levels of factor B.
4. By examining whether the treatment means curves for the different factor levels in an interaction plot are parallel.

If any one of the above four conditions holds, then they all hold, and there is no interaction.

116

### **Important vs. Unimportant Interactions**

When interaction is present, it is generally the case that factor level means are not useful. Rather, the effects of Factor B need to be reported separately for each level of Factor A (and vice versa, the effects of Factor A are reported separately for each level of Factor B).

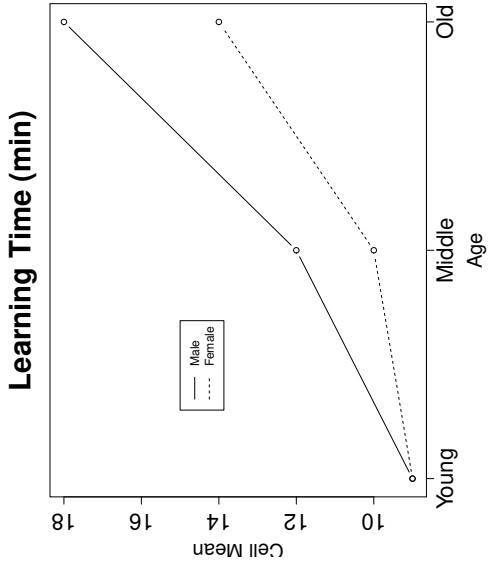
Sometimes, the observed interaction is so small that it is still meaningful to report main effects. We say the interaction is “unimportant” in this case.

*Example* Learning time for Males and Females in Three Age Groups (Section 19.2)

Consider a two-factor observational study, in which the effects of gender (male, female) and age (young, middle, old) on learning time (measured in minutes) are of interest.

117

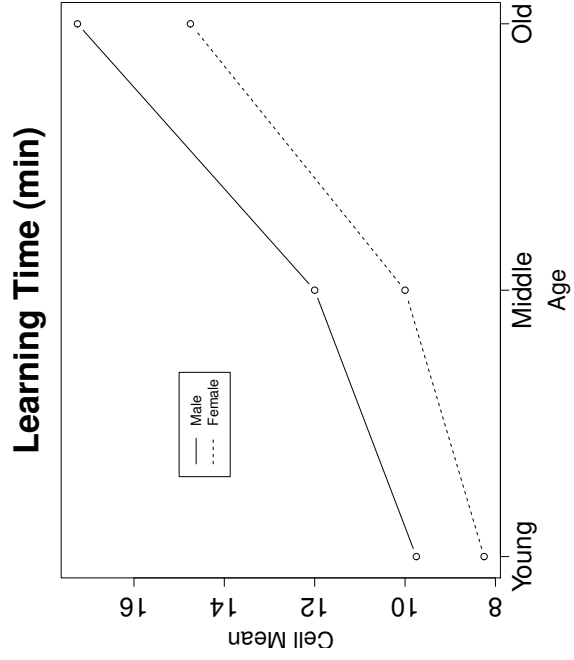
Example Interaction in Learning Time study, Case 1



Discuss: Is interaction important here? Why or why not?

118

Example Interaction in Learning Time study, Case 2

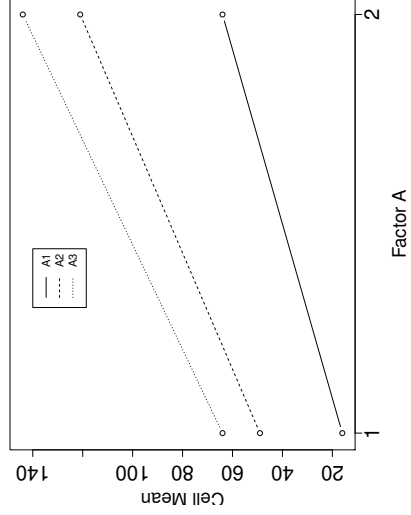


119

### Transformable Interaction

Sometimes, in two-way ANOVA, there is an important interaction, but it can be eliminated by a smooth transformation  $g(y)$  of the response variable  $Y$ . That is, if we fit the same two-way model, but to  $Y' = g(Y)$ , we no longer have important interaction.

*Example* This is a typical example of what a transformable interaction looks like.



120

### Interpretation of interactions

If important, untransformable interaction is present, we need to present the interpretation in as intuitive a manner as possible in the study context.

*Ex. Learning Time (p. ??)*

121



### Two-way ANOVA, factor effects model

The  $k^{\text{th}}$  observation in cell  $(i, j)$  is denoted  $Y_{ijk}$ ,  $i = 1, \dots, a$ ,  $j = 1, \dots, b$ ,  $k = 1, \dots, n$ , and the model for this observation says:

$$Y_{ijk} = \mu_{..} + \alpha_i + \beta_j + (\alpha\beta)_{ij} + \epsilon_{ijk}$$

where

$\mu_{..}$  is a constant,

$$\sum_{i=1}^a \alpha_i = 0$$

$$\sum_{j=1}^b \beta_j = 0$$

$$\sum_{j=1}^b (\alpha\beta)_{ij} = 0 \text{ for all } i = 1, \dots, a$$

$$\sum_{i=1}^a (\alpha\beta)_{ij} = 0 \text{ for all } j = 1, \dots, b$$

Assumptions on the  $\epsilon_{ijk}$ : independent, normally distributed, with mean 0 and constant variance  $\sigma^2$

122

### Notation

Let  $Y_{ij.} = \sum_{k=1}^n Y_{ijk}$ , be the sum of the observations for the treatment corresponding to the  $i^{\text{th}}$  level of factor A and the  $j^{\text{th}}$  level of factor B.

The corresponding mean is:

$$\bar{Y}_{ij.} = \frac{1}{n} Y_{ij.}$$

The total of all observations for the  $i^{\text{th}}$  level of factor A is:

$$Y_{i..} = \sum_{j=1}^b \sum_{k=1}^n Y_{ijk}$$

and the corresponding mean is:

$$\bar{Y}_{i..} = \frac{1}{bn} Y_{i..}$$

123

### Notation, con.

The total of all observations for the  $j^{\text{th}}$  level of factor B is:

$$Y_{.j} = \sum_{i=1}^a \sum_{k=1}^n Y_{ijk}$$

and the corresponding mean is:

$$\bar{Y}_{.j} = \frac{1}{an} Y_{.j}.$$

The sum of all observations in the study is:

$$Y_{...} = \sum_{i=1}^a \sum_{j=1}^b \sum_{k=1}^n Y_{ijk}$$

and the overall mean is

$$\bar{Y}_{...} = \frac{1}{abn} Y_{...}$$

124

### Least-squares estimates for the two-way ANOVA model

Factor effects model equation:

$$\mu_{ij} = \mu_{..} + \alpha_i + \beta_j + (\alpha\beta)_{ij}$$

Least-squares estimates of  $\mu_{ij}$  are the cell means  $\bar{Y}_{ij}$ , fitted values are

$$\hat{Y}_{ijk} = \bar{Y}_{ij}, \text{ and residuals are } e_{ijk} = Y_{ijk} - \hat{Y}_{ijk} = Y_{ijk} - \bar{Y}_{ij}.$$

Estimates of the model equation parameters:

$$\hat{\mu}_{..} = \bar{Y}_{...}$$

$$\hat{\alpha}_i = \bar{Y}_{i..} - \bar{Y}_{...}$$

$$\hat{\beta}_j = \bar{Y}_{.j.} - \bar{Y}_{...}$$

$$\widehat{(\alpha\beta)}_{ij} = \bar{Y}_{ij.} - \bar{Y}_{i..} - \bar{Y}_{.j.} + \bar{Y}_{...}$$

125

**Decomposition of deviations, sums of squares, and degrees of freedom**  
(first stage)

First, get the decompositions as if we were just doing a one-way ANOVA, where the single treatment factor is defined by all possible pairs of Factors A and B, and which has  $r = ab$  levels.

$$Y_{ijk} - \bar{Y}_{...} = (\bar{Y}_{ij.} - \bar{Y}_{...}) + (Y_{ijk} - \bar{Y}_{ij.})$$

$$\sum_{i=1}^a \sum_{j=1}^b \sum_{k=1}^n (Y_{ijk} - \bar{Y}_{...})^2 = \sum_{i=1}^a \sum_{j=1}^b n(\bar{Y}_{ij.} - \bar{Y}_{...})^2 + \sum_{i=1}^a \sum_{j=1}^b \sum_{k=1}^n (Y_{ijk} - \bar{Y}_{ij.})^2$$

$$\text{SSTO} = \text{SSTR} + \text{SSE}$$

$$nab - 1 = (ab - 1) + (n - 1)ab$$

In the second stage of the decomposition of SS, the treatment sum of squares is broken down into parts due to main effect of A, main effect of B, and interaction effect AB. This stage is crucial in two-way ANOVA.

**Decomposition of treatment deviations, SS, df** For  $i = 1, \dots, a;$   
 $j = 1, \dots, b; k = 1, \dots, n;$

$$\bar{Y}_{ij.} - \bar{Y}_{...} = (\bar{Y}_{i..} - \bar{Y}_{...}) + (\bar{Y}_{.j.} - \bar{Y}_{...}) + (\bar{Y}_{ij.} - \bar{Y}_{i..} - \bar{Y}_{.j.} + \bar{Y}_{...})$$

$$n \sum_{i=1}^a \sum_{j=1}^b (\bar{Y}_{ij.} - \bar{Y}_{...})^2 = nb \sum_{i=1}^a (\bar{Y}_{i..} - \bar{Y}_{...})^2 + na \sum_{j=1}^b (\bar{Y}_{.j.} - \bar{Y}_{...})^2$$

$$+ n \sum_{i=1}^a \sum_{j=1}^b (\bar{Y}_{ij.} - \bar{Y}_{i..} - \bar{Y}_{.j.} + \bar{Y}_{...})^2$$

$$ab - 1 = (a - 1) + (b - 1) + (a - 1)(b - 1)$$

## Mean Squares

$$MSE = \frac{SSE}{ab(n-1)}$$

$$MSA = \frac{SSA}{a-1}$$

$$MSB = \frac{SSB}{b-1}$$

$$MSAB = \frac{SSAB}{(a-1)(b-1)}$$

128

## Expected Mean Squares

$$E(MSE) = \sigma^2$$

$$E(MSA) = \sigma^2 + nb \frac{\sum \alpha_i^2}{a-1}$$

$$E(MSB) = \sigma^2 + na \frac{\sum \beta_j^2}{b-1}$$

$$E(MSAB) = \sigma^2 + n \frac{\sum \sum (\alpha_i \beta_j)^2}{(a-1)(b-1)}$$

129

### ANOVA table

Source of Variation	Sum of Squares	df	Mean Square
Factor A	SSA	$a - 1$	$\frac{SSA}{a-1}$
Factor B	SSB	$b - 1$	$\frac{SSB}{b-1}$
AB interaction	SSAB	$(a - 1)(b - 1)$	$\frac{SSAB}{(a-1)(b-1)}$
Error	SSE	$ab(n - 1)$	$\frac{SSE}{ab(n-1)}$
Total	SSTO	$nab - 1$	

where

$$SSA = nb \sum_{i=1}^a (\bar{Y}_{i..} - \bar{Y}_{...})^2$$

$$SSB = na \sum_{j=1}^b (\bar{Y}_{.j.} - \bar{Y}_{...})^2$$

$$SSAB = n \sum_{i=1}^a \sum_{j=1}^b (\bar{Y}_{ij.} - \bar{Y}_{i..} - \bar{Y}_{.j.} + \bar{Y}_{...})^2$$

$$SSE = \sum_{i=1}^a \sum_{j=1}^b \sum_{k=1}^n (Y_{ijk} - \bar{Y}_{ij.})^2$$

There are three  $F$  tests

**Question 1 for ANOVA Analysis** First question we examine in the statistical analysis is: Is there an interaction between Factor A and Factor B? That is, does the effect of Factor A depend on what is the level of Factor B?

*Ex. Corn Yield*

Cell Means

Manure	Fertilizer		Individual Measures of effect
	Low	High	
Low	11.3	13.9	$13.9 - 11.3 = 2.6$
High	14.0	15.1	$15.1 - 14.0 = 1.1$
Individual Measures of effect	$14.0 - 11.3 = 2.7$	$15.1 - 13.9 = 1.2$	

The individual measures of effect are not equal, for this data.

Question: Are the observed differences just due to chance?

### Example Corn Yield

Set up the data set in R:

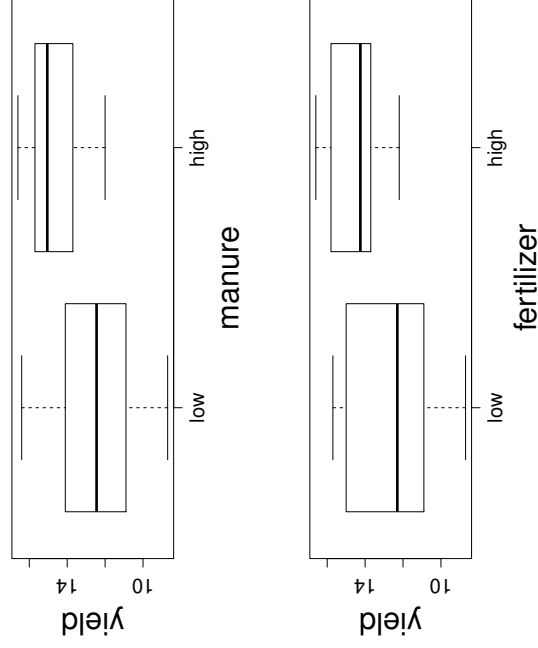
There are twenty observations in the file `corn-yield.txt`. Do the following to make the data frame:

```
> yield <- read.table("corn-yield.txt", header=TRUE)
> fertilizer <- rep(c("high", "low"), c(10,10))
> manure <- rep(rep(c("high", "low"), c(5,5)), 2)
> fertilizer <- factor(fertilizer, levels=c("low", "high"))
> yield.df <- data.frame(yield=yield,
+   fertilizer=fertilizer, manure=manure)
```

132

Get the boxplots of “high” and “low” levels, for each factor separately:

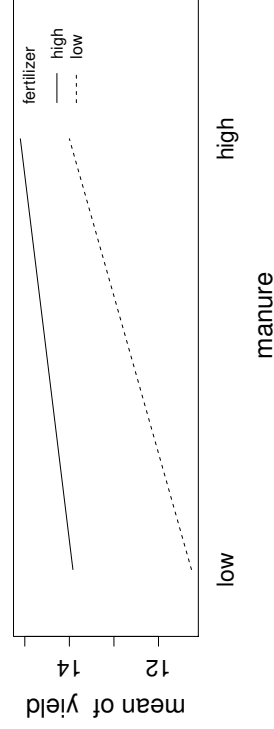
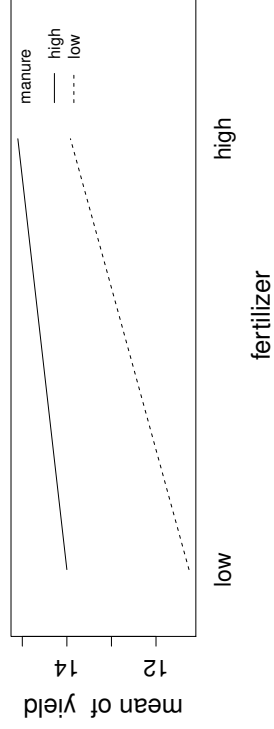
```
> plot(yield ~ fertilizer*manure, data=yield.df)
```



133

**Interaction Plot** We do this two possible ways; choose one.

```
> interaction.plot(fertilizer,manure, response=yield)
> interaction.plot(manure, fertilizer, response=yield)
```



134

```
> m1 <- aov(yield ~ manure*fertilizer, data=yield.df)
> anova(m1)
Analysis of Variance Table
```

```
Response: yield
          Df Sum Sq Mean Sq F value    Pr(>F)
manure     1 19.208  19.208   6.9218 0.01816 *
fertilizer  1 17.672  17.672   6.3683 0.02258 *
manure:fertilizer  1  3.042   3.042   1.0962 0.31066
Residuals 16 44.400   2.775
```

Three  $F$  tests are given in the table.

Each  $F$  statistic (labelled “ $F$  value”) is a ratio of mean squares:

$$F_{\text{obs}} = \frac{MS(\text{Factor})}{MS(\text{Residuals})}$$

135

### Test of interaction effect

$H_0$  : No interaction between the two factors

Or, more specifically:

$H_0$  : all  $(\alpha\beta)_{ij} = 0$

$H_a$  : not all  $(\alpha\beta)_{ij} = 0$

The test statistic is:

$$F_{\text{obs}} = \frac{\text{MSAB}}{\text{MSE}}$$

Under  $H_0$ , the distribution of  $F_{\text{obs}}$  is  $F_{(a-1)(b-1), (n-1)ab}$

Decision Rule:

136

```
Response: yield
      Df Sum Sq Mean Sq F value Pr(>F)
manure  1 19.208  19.208  6.9218 0.01816 *
fertilizer  1 17.672  17.672  6.3683 0.02258 *
manure:fertilizer  1  3.042   3.042  1.0962 0.31066
Residuals 16 44.400   2.775
```

$F_{\text{obs}} = \frac{3.042}{2.775} = 1.096$ ; the null distribution of the  $F$ -statistic is  $F_{1,16}$ ,  
 $P = .3107$

With this large a P-value we can “accept  $H_0$ ” and conclude there is no interaction effect.

In this case, we proceed to test factor main effects. If the interaction **had** been significant, we would not have gone on to test factor main effects.

137



### Tests for Factor Main Effects

The hypotheses for Factor A main effects:

$$H_0 : \alpha_1 = \alpha_2 = \dots = \alpha_a = 0$$

$$H_a : \text{Not all } \alpha_i = 0$$

The test statistic is:

$$F_{\text{obs}} = \frac{\text{MSA}}{\text{MSE}}$$

Under  $H_0$ , the distribution of  $F_{\text{obs}}$  is  $F_{a-1, (n-1)ab}$ .

Decision Rule: Reject  $H_0$  at level  $\alpha$  if  $F_{\text{obs}} > F_{1-\alpha, a-1, (n-1)ab}$

### Factor B main effects

Hypotheses, test statistic, null distribution, and decision rule:

$$H_0 : \beta_1 = \beta_2 = \dots = \beta_b = 0 \text{ vs. } H_a : \text{Not all } \beta_j = 0$$

138

### Estimation of Contrast of Factor Level Means

This is the followup analysis that is appropriate when:

1. Interaction effect is not significant, with P-value  $> .2$ , and
2. The factor for which we are analyzing contrasts of level means had a significant  $F$  test.

*Example. Corn Yield*  $2 \times 2$  factorial. Only one pairwise comparison of interest for Manure, and ditto for Fertilizer.

For  $a > 2$  or  $b > 2$ , we are usually interested in multiple pairwise comparisons, or multiple factor level contrasts. Then the multiple comparison procedures for one-way ANOVA can be used with only minor modifications for two-factor studies.

Ex.  $a = 3$ , estimate all three pairwise comparisons plus  $L_1 = \frac{\mu_1 + \mu_2}{2} - \mu_3$  using Scheffé method.

139

## Two-way ANOVA: Follow-up analysis

Consider the general case of  $a$  levels of Factor A,  $b$  levels of Factor B.

**What to do if factors do not interact** (Text: Section 19.8)

If the first step of analysis determines that interactions are insignificant and/or unimportant, then follow-up analysis is based on contrasts of factor level means. If only one comparison is being made, use individual  $t$  or  $F$  methods. If several comparisons are being made, each of the three multiple comparison methods (Bonferroni, Scheffe, Tukey) can be adapted to the two-way layout; the choice depends on the application.

140

Factor A contrasts and standard errors:

Let  $L = \sum_{i=1}^a c_i \mu_{i.}$ , with  $\sum c_i = 0$ ; estimate  $L$  by

$$\hat{L} = \sum_{i=1}^a c_i \bar{Y}_{i.}$$

In a balanced two-way design, each row mean  $\bar{Y}_{i.}$  is an average of  $bn$  observations, and so has variance  $\frac{\sigma^2}{bn}$ . So, the variance and estimated variance of  $\hat{L}$  are:

$$\begin{aligned} \text{Var}(\hat{L}) &= \frac{\sigma^2}{bn} \sum_{i=1}^a c_i^2 \\ \widehat{\text{Var}}(\hat{L}) &= \frac{\text{MSE}}{bn} \sum_{i=1}^a c_i^2 \end{aligned}$$

141

Factor B contrasts and standard errors:

Let  $L = \sum_{j=1}^b c_j \mu_{.j}$ , with  $\sum c_j = 0$ ; estimate  $L$  by

$$\hat{L} = \sum_{j=1}^b c_j \bar{Y}_{.j}.$$

In a balanced two-way design, each column mean  $\bar{Y}_{.j}$  is an average of  $an$  observations, and so has variance  $\frac{\sigma^2}{an}$ . So, the variance and estimated variance of  $\hat{L}$  are:

$$\begin{aligned} \text{Var}(\hat{L}) &= \frac{\sigma^2}{an} \sum_{j=1}^b c_j^2 \\ \widehat{\text{Var}}(\hat{L}) &= \frac{\text{MSE}}{an} \sum_{j=1}^b c_j^2 \end{aligned}$$

142

**What to do if factors do interact** (Text: Section 19.9)

If the first step of analysis determines that interaction is significant and/or practically important, then the follow-up analysis is based on contrasts of cell means.

If only one comparison is being made, use individual  $t$  or  $F$  methods. If several comparisons are being made, each of the three multiple comparison methods (Bonferroni, Scheffe, Tukey) can be adapted to the two-way layout, whichever is deemed appropriate.

143

Cell mean contrasts, estimates and standard errors:

Let  $L = \sum_{i=1}^a \sum_{j=1}^b c_{ij} \mu_{ij}$ , with  $\sum c_{ij} = 0$ ; estimate  $L$  by

$$\hat{L} = \sum_{i=1}^a \sum_{j=1}^b c_{ij} \bar{Y}_{ij}.$$

So, the variance and estimated variance of  $\hat{L}$  are:

$$\text{Var}(\hat{L}) = \frac{\sigma^2}{n} \sum_{i=1}^a \sum_{j=1}^b c_{ij}^2$$

$$\widehat{\text{Var}}(\hat{L}) = \frac{\text{MSE}}{n} \sum_{i=1}^a \sum_{j=1}^b c_{ij}^2$$

144

**Interaction Model: Scheffé method for estimating contrasts of cell means**

First recall the one-way ANOVA application of Scheffé method (p. 82 of these notes): The critical constant (in one-way ANOVA notation) was  $\sqrt{S^2}$ , where  $S^2$  is given by:

Let  $S^2 = (r - 1)F(1 - \alpha; r - 1, n_T - r)$  (Here the capital "S" refers to critical constant for Scheffé method)

Since the  $a \times b$  two-way ANOVA model with interaction effects is equivalent to the one-way ANOVA model with  $r = ab$ , the Scheffé method is applied exactly as before, and the critical constant is:

$$S = \sqrt{(ab - 1)F(1 - \alpha; ab - 1, abn - ab)}$$

145

### Section 19.10 Pooling Sums of Squares in Two-Factor Analysis of Variance

If interaction effect is statistically insignificant with a  $P$ -value well over .20, we can drop the interaction term. The advantages of this approach:

- ▶ We gain degree(s) of freedom for error.
- ▶ The model and followup analysis are logically consistent.

The model equation for the additive model (model w/out interaction):

$$E(Y_{ijk}) = \mu_{..} + \alpha_i + \beta_j + \epsilon_{ij}$$

Analysis Notes:

- 1 The sums of squares for A and B are the same as before.
- 2 The new error sum of squares is the sum of the error and interaction sums of squares from the interaction model.
- 3 Degrees of freedom for error is  $ab(n - 1) + (a - 1)(b - 1) = abn - a - b + 1$

146

### Ex Corn Yield

Fit the additive model. ((Additive model means the model without interaction.)

```
> m2 <- aov(yield ~ fertilizer + manure, data=yield.df)
> summary(m2)
          Df Sum Sq Mean Sq F value Pr(>F)
fertilizer  1  17.67   17.672    6.332 0.0222 *
manure      1   19.21   19.208    6.883 0.0178 *
Residuals  17  47.44    2.791
```

Do the tests for **main effects** of fertilizer and manure, exactly as in the interaction model. The mean squares for fertilizer and manure are exactly the same as before, but the  $F$  statistics and  $P$ -values are slightly different, because there is a different mean square for residuals, and more d.f. for the denominator of the  $F$  statistics.

147

### Example Corn yield

#### Follow-up analysis

Let's find the 95% CI for effect of fertilizer (High - Low).

The point estimate is:  $\bar{Y}_{\text{high}} - \bar{Y}_{\text{low}} = 1.88$

Estimate the model parameter  $\sigma^2$  by the MS(Residuals) = 2.7907.

There are 10 observations in each fertilizer level, so the SE of the effect estimate is  $\sqrt{2.7907(1/10 + 1/10)} = .7471$ .

Resulting 95% confidence interval is (.30, 3.46) (work not shown). This is based on a  $t$  distribution with 17 d.f., after pooling sums of squares. The CI does not include zero, so we conclude the main effect of fertilizer is significant at level  $\alpha = .05$ .

148

### Removable interaction

Sometimes the  $Y$  variable can be transformed so as to “remove the interaction.” So long as this does not cause unequal variance or other model violations, it is a good thing.

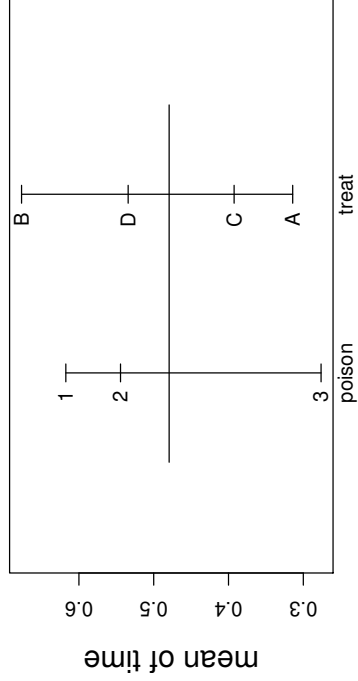
The following example is to illustrate an interaction removable by transformation.

149

*Example: Poisons* The data is  $T$  = survival time (in ten-hour time units) of groups of four animals randomly allocated to three poisons and four treatments. The experiment was part of an investigation to combat the effects of certain toxic agents.<sup>4</sup>

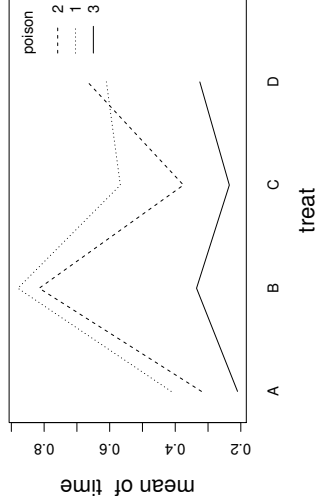
```
> library("boot")
> data(poisons)
> plot.design(poisons)
```

Plot of factor level means:



150

```
> interaction.plot(treat,poison,response=time)
```



```
> anova(toxic.aov)
```

Analysis of Variance Table

```
Response: time
      Df Sum Sq Mean Sq F value    Pr(>F)
poison  2  1.03301  0.51651  23.2217 3.331e-07 ***
treat   3  0.92121  0.30707  13.8056 3.777e-06 ***
poison:treat  6  0.25014  0.04169   1.8743  0.1123
Residuals  36  0.80072  0.02224
```

151

*Example: Poisons* Data were transformed to  $Y = 1/T =$  rate of dying because the transformed data satisfied model assumptions better than the original data.

Questions: Is there a difference in mean rate of dying among the three poisons? Is there a difference in mean rate of dying among the four treatments?

And first, we have to answer the question: Does effect of treatment depend upon the poison; that is, is there an interaction effect?

This is a  $3 \times 4$  factorial design with four observations per cell (four replications). Both of the factors, poisons and treatments, are of equal interest.

152

Cell means:

	A	B	C	D
I	2.49	1.16	1.86	1.69
II	3.27	1.39	2.71	1.70
III	4.80	3.03	4.26	3.09

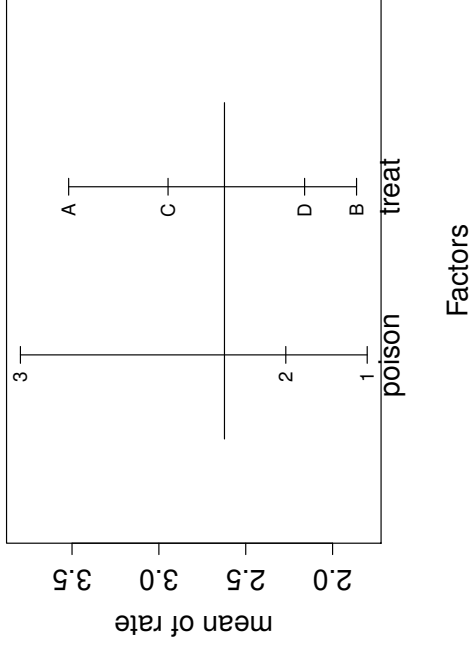
Cell SDs:

	A	B	C	D
I	0.50	0.20	0.49	0.36
II	0.82	0.55	0.42	0.70
III	0.53	0.42	0.23	0.24

153

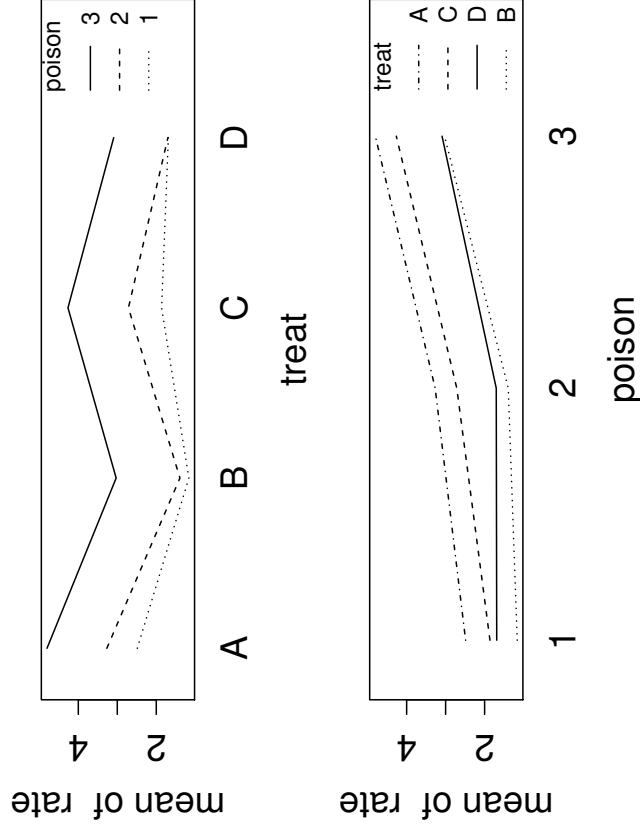


Design plot, showing poison and treatment averages (main effects):



154

Interaction plots:



155

### Check assumptions:

- ▶ Randomization—we were told treatments were allocated at random, so this assumption is satisfied.
- ▶ Normality — Normal quantile-quantile plot of studentized residuals looks more like a straight line than before transformation, so normality is better satisfied for the transformed than for the raw data. (Fill in some detail in class.)
- ▶ Constant variance—The plot of studentized residuals vs. fitted values shows no gross violation of the assumption.

And we would like to also assume no interaction. We will check this assumption with an  $F$  test.

156

**First step of analysis is to check the assumption of “no interaction”** We use a preliminary  $F$  test to do this.

Fit the model with interaction. Below is the resulting ANOVA table:

	Df	Sum Sq	Mean Sq	F value	Pr(>F)
poison	2	34.88	17.439	72.64	2.31e-13 ***
treat	3	20.41	6.805	28.34	1.38e-09 ***
poison:treat	6	1.57	0.262	1.09	0.387
Residuals	36	8.64	0.240		

$H_0$  : No interaction between poison and treatment, in effect on rate

$H_a$  : There is some interaction

$H_0 : (\alpha\beta)_{ij} = 0$  for  $i = 1, \dots, 3; j = 1, \dots, 4$

$H_a$  : Not all  $(\alpha\beta)_{ij} = 0$

$F = 1.09$ , null distr.  $F_{6,36}$ ,  $P = .387$ , and since  $P > .2$  we can accept  $H_0$  and proceed to fit the simpler, additive model.

157

**Fit additive model** Below is the resulting ANOVA table:

	Df	Sum Sq	Mean Sq	F value	Pr(>F)
poison	2	34.88	17.439	71.71	2.86e-14 ***
treat	3	20.41	6.805	27.98	4.19e-10 ***
Residuals	42	10.21	0.243		

Now test the hypotheses of primary interest, the tests of main effects of poisons and treatments:

►  $H_0$  : No effect due to poison,  $H_a$  : There is a difference among the means of the three poisons

OR:

$$H_0 : \alpha_1 = \alpha_2 = \alpha_3 = 0$$

$$H_a : \text{Not all } \alpha_i = 0$$

$F = 71.71$ ,  $P = 2.86 \times 10^{-14}$  (from  $F_{2,42}$ ); reject  $H_0$  and conclude there are differences in mean rate of dying among the three poisons.

►  $H_0 : \beta_1 = \beta_2 = \beta_3 = \beta_4 = 0$  (OR: No effect due to treatment)

*Example: Poisons* With response variable  $Y = \text{rate (of dying)}$ , we have seen that the additive model is appropriate.

And, both tests of main effects (Poison, Treatment) were significant.

For followup analysis, we will do all pairwise comparisons with Tukey's method, separately for Poison and for Treatment.

Tukey critical constants:

$$\text{Poison: } 1/\sqrt{2} \times q(1 - \alpha, a, nab - a - b + 1) = 1/\sqrt{2} \times q(.95, 3, 42) = 2.4295$$

Treatment:

$$1/\sqrt{2} \times q(1 - \alpha, b, nab - a - b + 1) = 1/\sqrt{2} \times q(.95, 4, 42) = 2.6750$$

Estimates of standard error of a pairwise comparison:

$$\text{Poison: } \sqrt{2\text{MSE}/bn} = \sqrt{2(.243)/16} = .1743$$

$$\text{Treatment: } \sqrt{2\text{MSE}/an} = \sqrt{2(.243)/12} = .2012$$

Allowances (Tukey's HSD):

$$\text{Poison: } 2.4295 \times .1743 = .4235$$

$$\text{Treatment: } 2.6750 \times .2012 = .5382$$

Cell means:

	A	B	C	D	Poison Means
I	2.49	1.16	1.86	1.69	1.80
II	3.27	1.39	2.71	1.70	2.27
III	4.80	3.03	4.26	3.09	3.80
<hr/>					
Treatment Means	3.52	1.86	2.95	2.16	2.62

Poison comparisons:

$$D1 = \mu_{1.} - \mu_{2.}, \hat{D}1 = 1.80 - 2.27 = -.47$$

$$D2 = \mu_{1.} - \mu_{3.}, \hat{D}2 = 1.80 - 3.80 = -2.0$$

$$D3 = \mu_{2.} - \mu_{3.}, \hat{D}3 = 2.27 - 3.80 = -1.53$$

95% confidence intervals by Tukey's HSD method,  $\hat{D} \pm .4235$ :

$$D1: (-.89, -.05)$$

$$D2: (-2.42, -1.58)$$

$$D3: (-1.95, -1.11)$$

160

Treatment comparisons:

$$D1 = \mu_{.1} - \mu_{.2}, \hat{D}1 = 3.52 - 1.86 = 1.66$$

$$D2 = \mu_{.1} - \mu_{.3}, \hat{D}2 = 3.52 - 2.95 = .57$$

$$D3 = \mu_{.1} - \mu_{.4}, \hat{D}3 = 3.52 - 2.16 = 1.36$$

D4

D5

D6

95% confidence intervals by Tukey's HSD method,  $\hat{D} \pm .5382$

$$D1: 3.52 - 1.86 \pm .5382$$

$$D2: 3.52 - 2.95 \pm .5382$$

$$D3: 3.52 - 2.16 \pm .5382$$

$$D4: 1.86 - 2.95 \pm .5382$$

161

## CHAPTER 21 Randomized Block Design

This is a very commonly-used design. Call the response variable  $Y$ ; let  $r$  be the number of treatments. Suppose there is a nuisance variable which is known to affect the variance of  $Y$ . In this design we group experimental units into several blocks; each block contains  $r$  units that have equal or similar values of the nuisance variable. Then apply each treatment to one unit in each block. Thus each block is a *replicate* of the base experiment.

- ▶ For this to be a randomized block design, there must be random assignment of treatments within the blocks.
- ▶ If blocking is done well, the experimental error variance can be greatly reduced in comparison to what would be obtained from a completely randomized one-way design.

*Example 1: Crop yield* Suppose the field for the experiment has a fertility gradient, and suppose you have  $r = 4$  treatments as in the corn yield (fertility, manure) example. Divide the field up into five blocks of four plots each from High (Block 1) to Low fertility (Block 5). Randomly assign the four treatments within each of the five blocks. This makes sense, to be sure each treatment gets a fair representation in the range of available fertility levels.

164

*Example 2: Risk premium* (text, p. 895/901). In an experiment on decision making, executives were exposed to one of three methods of quantifying the maximum risk premium they would be willing to pay to avoid uncertainty in a business decision.

Treatment factor: Method (utility, worry, comparison)

Response variable:  $Y$  = degree of confidence in the method, on a scale from 0 (no confidence) to 20 (highest confidence).

Experimental units: Fifteen executives served as subjects to whom the levels of the trt factor would be assigned at random within blocks.

Blocking factor: Age, five levels from oldest to youngest; three executives per block.

165

Main reasons for blocking:

- ▶ Reduction of confounding: In small experiments especially, in the completely randomized design, the randomization may not “work” to perfectly balance all confounding variables among the treatment groups. With the RB design, you are guaranteed to get fair comparisons of all treatments with respect to the blocking variable.
- ▶ By controlling for block factor in this way, you reduce the experimental error variance, compared to the one-way completely randomized design, sometimes by a lot.

166

### *Randomized blocks, randomization in R*

```
> set.seed(16)
> for (i in 1:5)
> {
>   print(sample(1:3, replace=FALSE))
> }
#
[1] 1 2 3
[1] 1 3 2
[1] 3 2 1
[1] 2 3 1
[1] 3 1 2
```

167

## Randomization chart

Give ID's 1, 2, 3 to the three exec's within a block. Our randomization:

```
[1] 1 2 3  
[1] 1 3 2  
[1] 3 2 1  
[1] 2 3 1  
[1] 3 1 2
```

Blocks	Treatments		
1	Exec. 1	2	3
2	Exec. 1	Exec. 2	Exec. 3
3	3	Exec. 3	Exec. 2
4	2	2	1
5	3	3	1
		1	2

168

## Model for randomized block design with one observation per cell

We assume the additive two-way model, the same as the additive model discussed for the two-way factorial design. However, different notation is used to emphasize the nature of the randomized block design, which is different from the nature of the factorial experiment.

169



### Model for randomized blocks design with fixed effects for blocks

Rows represent blocks; columns represent treatments. Let  $n_b$  = no. of blocks = no. of rows in the two-way layout. Let  $r$  = no. of treatments = no. of columns in the two-way layout.

Model equation for observation  $Y_{ij}$ , which is the response for the  $j^{\text{th}}$  treatment in the  $i^{\text{th}}$  block:

$$Y_{ij} = \mu_{..} + \rho_i + \tau_j + \epsilon_{ij}$$

where

$\mu_{..}$  is a constant,

$\rho_i$  are constants for the block (row) effects;  $\sum_{i=1}^{n_b} \rho_i = 0$

$\tau_j$  are constants for the treatment effects;  $\sum_{j=1}^r \tau_j = 0$

Assumptions on the  $\epsilon_{ij}$ : independent, normally distributed, with mean 0 and constant variance  $\sigma^2$

170

### Notation for observed means

Let  $\bar{Y}_i = \frac{\sum_{j=1}^r Y_{ij}}{r}$  be the mean of observations in the  $i^{\text{th}}$  block,

let  $\bar{Y}_{.j} = \frac{\sum_{i=1}^{n_b} Y_{ij}}{n_b}$  be the mean of observations in the  $j^{\text{th}}$  treatment,

and let  $\bar{Y}_{..} = \frac{1}{n_b r} \sum_{i=1}^{n_b} \sum_{j=1}^r Y_{ij}$ , be the grand mean.

Further,

let  $\hat{\rho}_i = \bar{Y}_i - \bar{Y}_{..}$ .

and  $\hat{\tau}_j = \bar{Y}_{.j} - \bar{Y}_{..}$ .

171

### Fitted values:

$$\begin{aligned}\hat{Y}_{ij} &= \bar{Y}_{..} + (\bar{Y}_{i.} - \bar{Y}_{..}) + (\bar{Y}_{.j} - \bar{Y}_{..}) \\ &= \hat{\mu}_{..} + \hat{\rho}_i + \hat{\tau}_j\end{aligned}$$

Note:  $\hat{Y}_{ij}$  simplifies to  $\bar{Y}_{i.} + \bar{Y}_{.j} - \bar{Y}_{..}$ .

### Residuals:

$$e_{ij} = Y_{ij} - \hat{Y}_{ij} = Y_{ij} - \bar{Y}_{i.} - \bar{Y}_{.j} + \bar{Y}_{..}$$

172

### Sums of squares

There are three sums of squares, for Blocks, Treatments, and Block  $\times$  Treatment interaction. Below are the formulas for “by-hand” calculation of the sums of squares.

$$SSBL = \sum_{i=1}^{n_b} r(\bar{Y}_{i.} - \bar{Y}_{..})^2$$

$$SSSTR = \sum_{j=1}^r n_b(\bar{Y}_{.j} - \bar{Y}_{..})^2$$

$$SSBL.TR = \sum_{i=1}^{n_b} \sum_{j=1}^r (Y_{ij} - \bar{Y}_{i.} - \bar{Y}_{.j} + \bar{Y}_{..})^2$$

**IMPORTANT:** With only one observation per cell, the interaction sum of squares is used as the error sum of squares.

173

Degrees of freedom:

$$\text{df(Blocks)} = n_b - 1$$

$$\text{df(Treatment)} = r - 1$$

$$\text{df(BI.Tr)} = (n_b - 1)(r - 1)$$

Recall: A Mean Square is a sum of squares, divided by its degrees of freedom. A mean square is a statistic and has a sampling distribution.

Mean Squares in RCB Design, formulas

$$\text{MSBL} = \frac{\text{SSBL}}{n_b - 1}$$

$$\text{MSTR} = \frac{\text{SSTR}}{r - 1}$$

$$\text{MSBL.TR} = \frac{\text{SSBL.TR}}{(n_b - 1)(r - 1)}$$

174

Expected Mean Squares

$$E(\text{MSBL}) = \sigma^2 + r \frac{\sum \rho_i^2}{n_b - 1}$$

$$E(\text{MSTR}) = \sigma^2 + n_b \frac{\sum \tau_j^2}{r - 1}$$

$$E(\text{MSBL.TR}) = \sigma^2$$

ANOVA table

Source of Variation	Sum of Squares	df	Mean Square	Expected Mean Square
Blocks	SSBL	$n_b - 1$	MSBL	$\sigma^2 + r \frac{\sum \rho_i^2}{n_b - 1}$
Treatments	SSTR	$r - 1$	MSTR	$\sigma^2 + n_b \frac{\sum \tau_j^2}{r - 1}$
Error	SSBL.TR	$(n_b - 1)(r - 1)$	MSBL.TR	$\sigma^2$

175

## F Test for Treatment Factor

*Hypotheses to be tested:*

$H_0$ :  $\tau_1 = \tau_2 = \dots = \tau_r = 0$  vs.

$H_a$ : At least one  $\tau_i$  is not zero.

*Test Statistic:*

$$F_{\text{obs}} = \frac{\text{MSTR}}{\text{MSBL.TR}}$$

*Null distribution:*  $F_{r-1, (n_b-1)(r-1)}$

*P-value:* area to the right of the observed  $F$  statistic, under the curve of the null distribution

*Decision rule:* Reject  $H_0$  at level  $\alpha$ , if  $P < \alpha$ . If you do reject  $H_0$ , go on to do followup analysis of treatment (column) means, using whichever of Bonferroni, Scheffé, or Tukey methods is appropriate.

176

## F Test for Blocks

Since Block is a “nuisance factor,” we are not too interested in testing for block effects.

However, we would like to see a significant F test for blocks, because this gives an indication that this factor was useful for blocking.

The main statistical reason for forming blocks is to **reduce error variability** by comparing every treatment within each group of units in a block.

If the F test for Block is significant, this indicates that it was valuable to “take out” block effects from the error.

177

*Example of randomized blocks design: Risk premium*

**Data**

Block —	Method ( <i>j</i> )			Average
	Utility	Worry	Comparison	
1 (oldest)	1	5	8	4.7
2	2	8	14	8.0
3	7	9	16	10.7
4	6	13	18	12.3
5 (youngest)	12	14	17	14.3
Average	5.6	9.8	14.6	10.0

$$\hat{Y}_{11} = \bar{Y}_{1.} + \bar{Y}_{.1} - \bar{Y}_{..} = 4.7 + 5.6 - 10.0 = .3$$

$$e_{11} = Y_{11} - \hat{Y}_{11} = 1 - .3 = .7$$

**Ex: Risk Premium Outline of analysis “by hand”**

**Fitted Values:**

Block	Method ( <i>j</i> )		
	Utility	Worry	Comparison
1	.3	4.5	9.3
2	3.6	7.8	12.6
3	6.3	10.5	15.3
4	7.9	12.1	16.9
5			

**Residuals:**

Block	Method ( <i>j</i> )		
	Utility	Worry	Comparison
1	.7	.5	-1.3
2	-1.6	.2	1.4
3	.7	-1.5	.7
4	-1.9	.9	1.1
5			

*Risk premium, RB analysis continued* Refer to the tables of summary statistics to calculate the following items for the randomized blocks analysis:

1  $SSTR = n_b \sum_{j=1}^r (\bar{Y}_{.j} - \bar{Y}_{..})^2 =$

2  $df(\text{Treatment}) =$

3  $SSBL = r \sum_{i=1}^{n_b} (\bar{Y}_{i.} - \bar{Y}_{..})^2 =$

4  $df(\text{Blocks}) =$

5  $SSBL.TR = \sum_{i=1}^{n_b} \sum_{j=1}^r (Y_{ij} - \bar{Y}_{i.} - \bar{Y}_{.j} + \bar{Y}_{..})^2 =$

6  $df(\text{Error}) =$

**ANOVA table**

Source of Variation	Sum of Squares	df	Mean Square	F	P
Blocks					
Treatments					
Error					

We will now get the analysis in R.

**Ex: Risk Premium** Create the dataframe, fit the randomized blocks model, and get the ANOVA table.

```
> confidence <- c(1,2,7,6,12,5,8,9,13,14,8,14,16,18,17)
> method <- factor(rep(c("utility", "worry", "comp"), c(5,5,5)))
> block <- factor(rep(1:5,3))
> riskprem <- data.frame(confidence=confidence,
+ method=method,block=block)
> riskprem.aov <- aov(confidence ~ block + method,
+ data=riskprem)
> anova(riskprem.aov)
```

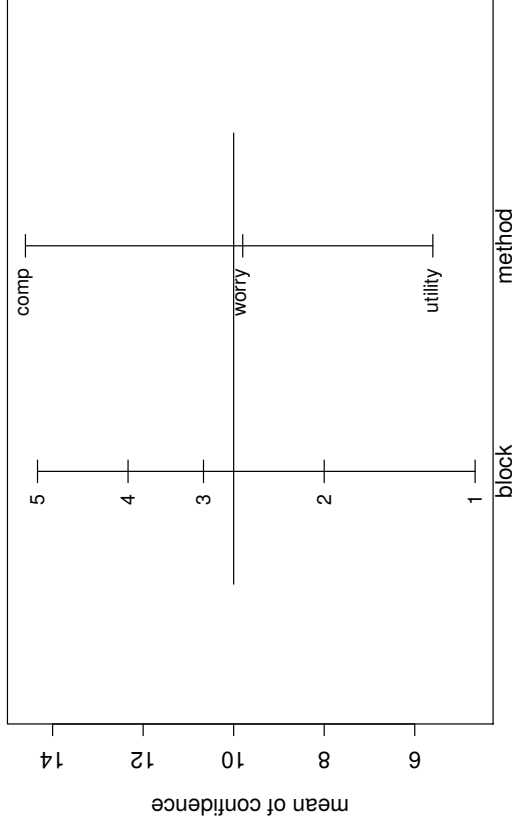
```
      Df Sum Sq Mean Sq F value    Pr(>F)
block  4 171.333  42.833  14.357 0.0010081 **
method  2 202.800 101.400  33.989 0.0001229 ***
Residuals  8  23.867   2.983
```

**We should look at relevant plots, and check assumptions before drawing conclusions from the F tests.**

182

**Get the plots of factor level means for both blocks and treatment.**

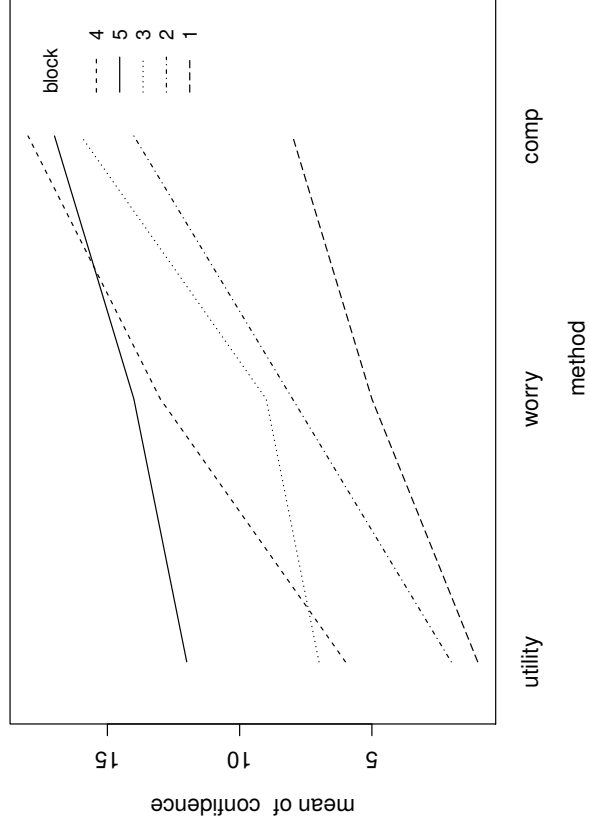
```
> plot.design(confidence ~ block + method)
```



Factors

183

```
> interaction.plot(method, block, response=confidence)
```



184

### Tukey's 1 df test for transformable nonadditivity

Suppose that the additive model does not hold for the original response  $Y_{ij}$  but does hold for a nonlinear transformation of  $Y$  such as  $\log(Y)$  or  $\sqrt{Y}$  (Call the transformed variable  $Y'$ ).

Claim: In this situation, the plot of residuals vs. fitted values from the additive model fit to  $Y$  will have a quadratic appearance.

185



A formal test for transformable nonadditivity based on the detection of a curvilinear relationship between  $Y - \hat{Y}$  and  $\hat{Y}$  is due to J. W. Tukey<sup>5</sup>

**Procedure**

- 1 Fit the additive two-way model to the  $Y_{ij}$ ; call the fitted values  $\hat{Y}_{ij}$  and the residuals  $e_{ij}$ .
- 2 Let  $q_{ij} = \hat{Y}_{ij}^2$ — the squares of the fitted values from Step 1.
- 3 Fit the Step 1 additive two-way model to the  $q_{ij}$ , call the residuals  $e_{ij}^q$ .
- 4 Let  $P$  be the sum of products of the two sets of residuals, the  $e_{ij}$  and the  $e_{ij}^q$ .

$$P = \sum_{i=1}^{n_b} \sum_{j=1}^r e_{ij} e_{ij}^q$$

- 5 Let  $Q$  be the sum of squares of the residuals  $e_{ij}^q$

$$Q = \sum_{i=1}^{n_b} \sum_{j=1}^r (e_{ij}^q)^2$$

---

<sup>5</sup>This explanation is taken from Box, Hunter and Hunter, *Statistics for Experimenters*, p. 222

**Tukey's 1 df test procedure (con.)**

The rest of the procedure uses a decomposition of the original residual sum of squares from Step 1 (SSE), into a part for transformable nonadditivity and the remainder:

- 5 Form the sum of squares for transformable nonadditivity having one degree of freedom,

$$SS_{na} = \frac{P^2}{Q}$$

- 6 Let  $SS_{rem} = SSE - SS_{na}$ , with degrees of freedom  $(n_b - 1)(r - 1) - 1$
- 7 Form the  $F$  ratio

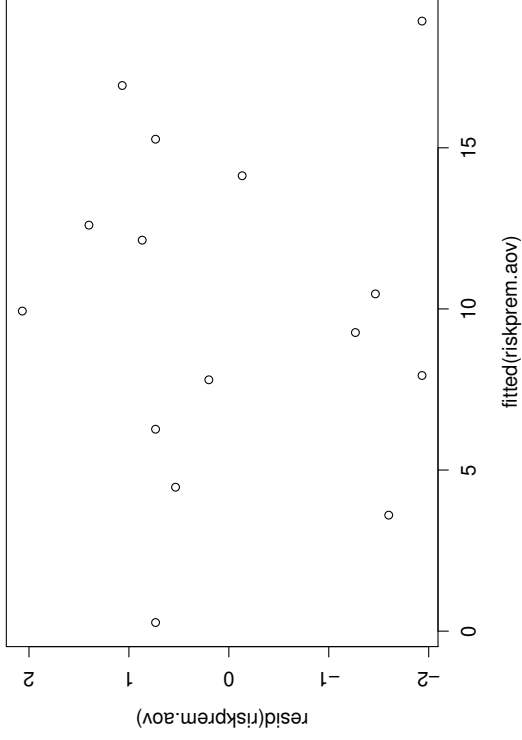
$$F_{na} = \frac{SS_{na}/1}{SS_{rem}/(n_b(r) - n_b - r)}$$

Under the null hypothesis, this statistic has the  $F$  distribution with 1 and  $(n_b - 1)(r - 1) - 1 = n_b r - n_b - r$  df.

A large  $F$  indicates there **is** transformable nonadditivity.

Ex. *Risk premium* The plot of residuals vs. fitted values doesn't indicate existence of transformable nonadditivity:

```
> plot(fitted(riskprem.aov), resid(riskprem.aov))
```



188

Ex *Risk Premium* Below is R code and output, to carry out the Tukey 1 df test. Note you need the SSE from the original ANOVA; this is in the ANOVA table p. ??, SSE = 23.867.

```
> fit <- riskprem.aov
> e <- residuals(fit)
> qij <- fitted(fit)^2
> fit.qij <- aov(qij ~ block + method)
> qij.e <- residuals(fit.qij)
> P <- sum(e*qij.e)
> Q <- qij.e %**% qij.e
> Sna <- P^2/Q; Sna
[1,]
[1,] 0.2626651
> Srem <- 23.867 - Sna
> Fstar <- Sna/(Srem/7); Fstar; 1-pf(Fstar,df1=1,df2=7)
[1,] 0.07789483
[1,]
[1,] 0.7882366
```

The value of the test statistic is  $F^* = .0779$  ( $P = .79$ ), so we conclude there is no evidence of transformable nonadditivity.

189