STAT 3119

Week6: 10/1/2019 @GWU

Outline

- Models for Two Factor Studies (ch 19.3)
- ANOVA : SS partition (ch 19.4)
- F-tests for main effects and interaction (ch 19.6)
- Model diagnostics (ch 19.5)

Two-factor (Two-way) ANOVA model (ch 19.3)

- The response variable Y is continuous
- There are two categorical explanatory variables (factors), called Factor A and Factor B.

- Factor A with levels $i = 1$ to a (a levels)
- Factor B with levels $j = 1$ to b (b levels)
- A particular combination of levels is called a **treatment** or a **cell**. There are **ab** treatments.
- Y_{ijk}: subscript index i denotes the level of the factor A, j denotes the level of the factor B, k denotes the kth observation for treatment (i, j), $k = 1$ to n, with $n \geq 1$:

In this chapter, we assume equal sample size in each treatment combination ($n_{ij} \equiv n$, $n > 1$; **total sample size** $n_T = abn$). This is called a **balanced** design. In later chapters, we will discuss (a) one case per treatment (b) **Unbalanced**: unequal sample sizes (more complicated).

Cell Means Models

Model assumptions: We assume that the response variable observations are independent, and normally distributed with a mean that may depend on the levels of the factors A and B, and a variance that does not (is constant).

The ANOVA model can be expressed in terms of cell (treatment) means μ_{ij}:

$$Y_{ijk} = {}_{ij} + \epsilon_{ijk}$$

where

- μ_{ij} is the population mean or expected value of all observations in cell (i, j).
- ϵ_{ijk} are iid $N(0, \sigma^2)$
- $i = 1, ..., a; j = 1, ..., b; k = 1, ...n$
- $Y_{ijk} \sim$ independent $N(\mu_{ij}, \sigma^2)$

- This ANOVA model is a special form of linear model and we can use an equivalent regression model to get the model parameters.
- **Estimation**: The best estimator for the cell means (least squares or maximum like-lihood estimators) is the sample mean of the observations in treatment (i, j), which is also the fitted values of the response variable

$$\hat{}_{ij} = \overline{Y}_{ij.} = \hat{Y}_{ijk}$$

- The residuals, defined as the difference between the observed and fitted values

$$e_{ijk} = Y_{ijk} - \hat{Y}_{ijk} = Y_{ijk} - \overline{Y}_{ij.}$$

and residuals are very useful for assessing the appropriateness of two-factor ANOVA model.

Factor Effects Model

Based on the definition of an interaction:

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

i.e.

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

Then we can replace each treatment mean μ_{ij} in terms of overall mean and **factor effects**, to obtain an equivalent version of cell means model as:

$$Y_{ijk} = \mu_{..} + \alpha_i + \beta_j + (\alpha\beta)_{ij} + \varepsilon_{ijk} \qquad \textbf{(19.23)}$$

where:

$$i = 1, \ldots, a; \; j = 1, \ldots, b; \; k = 1, \ldots, n$$

- $\mu_{..}$ is a constant (overall mean), $\mu_{..} = \dfrac{\sum_i \sum_j \mu_{ij}}{ab}$
- α_i are constants subject to the restriction $\sum \alpha_i = 0$

  (main effect α_i for factor A at the ith level. $\alpha_i = \mu_{i.} - \mu_{..}$)
- β_j are constants subject to the restriction $\sum \beta_j = 0$

  The main effect β_j for factor B at the jth level $\beta_j = \mu_{.j} - \mu_{..}$)
- $(\alpha\beta)_{ij}$ are constants subject to the restrictions:

  $\sum_i (\alpha\beta)_{ij} = 0 \qquad j = 1, \ldots, b$

  $\sum_j (\alpha\beta)_{ij} = 0 \qquad i = 1, \ldots, a$

  (The interaction effect $(\alpha\beta)_{ij}$ when factor A is at the ith level and factor B is at the jth level. $(\alpha\beta)_{ij} = \mu_{ij} - \mu_{i.} - \mu_{.j} + \mu_{..}$)
- ε_{ijk} are independent $N(0, \sigma^2)$

Estimates for Factor Effects Model (Ch 19.3)

- Y_{ijk}: subscript index i denotes the level of the factor A, j denotes the level of the factor B, k denotes the kth observation for treatment (i, j), $k = 1$ to n, with $n \geq 1$:

- A dot in the subscript indicates aggregation or summing over the variable represented by the index, the overline or bar indicates the average.
- Similar to the one-factor ANOVA models, the model parameter estimators (by LS or MLE) are based on sample means or differences of sample means.

	Parameter	Sum	Estimator (based on sample means)
Treatment (i,j) level mean	μ_{ij}	$Y_{ij.} = \sum_{k=1}^{n} Y_{ijk}$	$\bar{Y}_{ij.} = \frac{Y_{ij.}}{n}$
Factor A *i*th level mean	$\mu_{i.}$	$Y_{i..} = \sum_{j}^{b}\sum_{k}^{n} Y_{ijk}$	$\bar{Y}_{i..} = \frac{Y_{i..}}{bn}$
Factor B *j*th level mean	$\mu_{.j}$	$Y_{.j.} = \sum_{i}^{a}\sum_{k}^{n} Y_{ijk}$	$\bar{Y}_{.j.} = \frac{Y_{.j.}}{an}$
Overall mean	$\mu_{..}$	$Y_{...} = \sum_{i}^{a}\sum_{j}^{b}\sum_{k}^{n} Y_{ijk}$	$\bar{Y}_{...} = \frac{Y_{...}}{nab}$
Main effect for factor A at *i*th *level*	α_i		$\bar{Y}_{i..} - \bar{Y}_{...}$
Main effect for factor B at *j*th *level*	β_j		$\bar{Y}_{.j.} - \bar{Y}_{...}$
Interaction effect of A(*i*) and B(*j*)	$(\alpha\beta)_{ij}$		$\bar{Y}_{ij.} - \bar{Y}_{i..} - \bar{Y}_{.j.} + \bar{Y}_{...}$

Model Estimate: Example

Example (page 833):

- The Castle Bakery Company supplies wrapped Italian bread to a large number of supermarkets in a metropolitan area. An experimental study was made of the effects of **height of the shelf display (factor A: bottom, middle, top)** and the **width of the shelf display (factor B: regular, wide)** on **sales of this bakery's bread during the experimental period (Y, measured in cases)**.
- Twelve supermarkets, similar in terms of sales volume and customers, were utilized in the study. The six treatments were **assigned at random** to two stores each according to a completely randomized design, and the display of the bread in each store followed the **treatment** specifications for that store.

TABLE 19.7 **Sample Data and Notation for Two-Factor Study—Castle Bakery Example (sales in cases).**

Factor A (display height) i	Factor B (display width) j: B_1 (regular)	B_2 (wide)	Row Total	Display Height Average
A_1 (bottom)	47 (Y_{111}) 43 (Y_{112})	46 (Y_{121}) 40 (Y_{122})		
Total	90 ($Y_{11\cdot}$)	86 ($Y_{12\cdot}$)	176 ($Y_{1\cdot\cdot}$)	
Average	45 ($\bar{Y}_{11\cdot}$)	43 ($\bar{Y}_{12\cdot}$)		44 ($\bar{Y}_{1\cdot\cdot}$)
A_2 (middle)	62 (Y_{211}) 68 (Y_{212})	67 (Y_{221}) 71 (Y_{222})		
Total	130 ($Y_{21\cdot}$)	138 ($Y_{22\cdot}$)	268 ($Y_{2\cdot\cdot}$)	
Average	65 ($\bar{Y}_{21\cdot}$)	69 ($\bar{Y}_{22\cdot}$)		67 ($\bar{Y}_{2\cdot\cdot}$)
A_3 (top)	41 (Y_{311}) 39 (Y_{312})	42 (Y_{321}) 46 (Y_{322})		
Total	80 ($Y_{31\cdot}$)	88 ($Y_{32\cdot}$)	168 ($Y_{3\cdot\cdot}$)	
Average	40 ($\bar{Y}_{31\cdot}$)	44 ($\bar{Y}_{32\cdot}$)		42 ($\bar{Y}_{3\cdot\cdot}$)
Column total	300 ($Y_{\cdot 1\cdot}$)	312 ($Y_{\cdot 2\cdot}$)	612 ($Y_{\cdot\cdot\cdot}$)	
Display width average	50 ($\bar{Y}_{\cdot 1\cdot}$)	52 ($\bar{Y}_{\cdot 2\cdot}$)		51 ($\bar{Y}_{\cdot\cdot\cdot}$)

Model Estimate: Example in R (2)

1. read the data

```r
# read data from week5 folder online
Ex19  =read.table(
        url("https://raw.githubusercontent.com/npmldabook/Stat3119/master/Week6/CH19TA07.txt"))
head(Ex19, 3)
```

```
##   V1 V2 V3 V4
## 1 47  1  1  1
## 2 43  1  1  2
## 3 46  1  2  1
```

```r
names(Ex19) =  c("response", "height", "width", "units")

# make  categorical variables for factor A and B
Ex19$height =  as.factor(Ex19$height)
Ex19$width =  as.factor(Ex19$width)

str(Ex19)
```

```
## 'data.frame':    12 obs. of  4 variables:
##  $ response: int  47 43 46 40 62 68 67 71 41 39 ...
##  $ height  : Factor w/ 3 levels "1","2","3": 1 1 1 1 2 2 2 2 3 3 ...
##  $ width   : Factor w/ 2 levels "1","2": 1 1 2 2 1 1 2 2 1 1 ...
##  $ units   : int  1 2 1 2 1 2 1 2 1 2 ...
```

2. display the data and cell frequency nij=n

```
Ex19
```

```
##    response height width units
## 1        47      1     1     1
## 2        43      1     1     2
## 3        46      1     2     1
## 4        40      1     2     2
## 5        62      2     1     1
## 6        68      2     1     2
## 7        67      2     2     1
## 8        71      2     2     2
## 9        41      3     1     1
## 10       39      3     1     2
## 11       42      3     2     1
## 12       46      3     2     2
```

```
xtabs(~ height + width, data = Ex19)
```

```
##       width
## height 1 2
##      1 2 2
##      2 2 2
##      3 2 2
```

3. get the factor means, overall means and treatment means

```
(Overall.mean = mean(Ex19$response))
```

```
## [1] 51
```

```
(FactorA.mean = with(Ex19, by(response, height, mean )))
```

```
## height: 1
## [1] 44
## --------------------------------------------------------
## height: 2
## [1] 67
## --------------------------------------------------------
## height: 3
## [1] 42
```

```
(FactorB.mean = with(Ex19, by(response, width, mean )))
```

```
## width: 1
## [1] 50
## --------------------------------------------------------
## width: 2
## [1] 52
```

```
(treatment.mean = with(Ex19, by(response, list(width,height), mean )))
```

```
## : 1
## : 1
## [1] 45
## ---------------------------------------------------------
## : 2
## : 1
## [1] 43
## ---------------------------------------------------------
## : 1
## : 2
## [1] 65
## ---------------------------------------------------------
## : 2
## : 2
## [1] 69
## ---------------------------------------------------------
## : 1
## : 3
## [1] 40
## ---------------------------------------------------------
## : 2
## : 3
## [1] 44
```

4. get the main effects (A & B) and interaction (A:B)

```
Alpha.effect =  as.numeric(FactorA.mean -  Overall.mean)
Beta.effect =   as.numeric( FactorB.mean - Overall.mean)

list(Alpha=  Alpha.effect, Beta=Beta.effect)
```

```
## $Alpha
## [1] -7 16 -9
##
## $Beta
## [1] -1  1
```

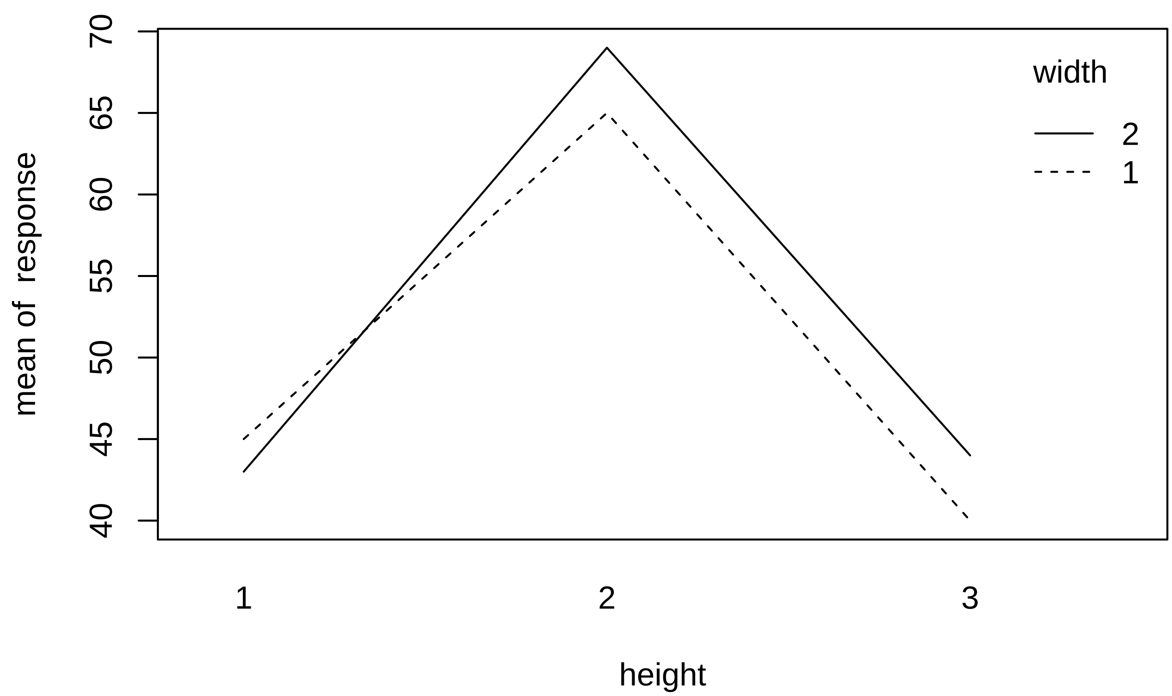
```
All   =  data.frame(height=rep(1:3,c(2,2,2)),width=rep(1:2,3), A.mean=rep(FactorA.mean,c(2,2,2)),
          B.mean=  rep(FactorB.mean, 3),  Trt.mean=as.numeric(treatment.mean), All.mean= Overall.mean)

Interaction =  as.numeric(treatment.mean  -All$A.mean
                - All$B.mean + Overall.mean)

(All= data.frame(All,Alpha=  rep(Alpha.effect,c(2,2,2)),  Beta=  rep(Beta.effect, 3), Interaction))
```

```
##   height width A.mean B.mean Trt.mean All.mean Alpha Beta Interaction
## 1      1     1     44     50       45       51    -7   -1           2
## 2      1     2     44     52       43       51    -7    1          -2
## 3      2     1     67     50       65       51    16   -1          -1
```

```
## 4      2     2     67     52       69       51     16    1              1
## 5      3     1     42     50       40       51     -9   -1             -1
## 6      3     2     42     52       44       51     -9    1              1
```

5. generating "the treatment means/interaction plot" (page 836)

```
with(Ex19, interaction.plot(x.factor = height, trace.factor = width, response = response))
```



What we can find from this plot:

- For both display widths, mean sales for the middle display height are substantially larger than those for the other two display heights.
- The effect of display width does not appear to be large.
- It seems that there were not much interaction between display height and display width in their effects on sales.

We can test the main effect and interaction using a two-way ANOVA analysis.

Partition of Total Sum of Squares, SS and MS

1. Partition of total deviation into deviation of treatment mean around overall mean and devation around treatment mean:

$$\underbrace{Y_{ijk} - \overline{Y}_{\cdots}}_{\substack{\text{Total}\\ \text{deviation}}} = \underbrace{\overline{Y}_{ij\cdot} - \overline{Y}_{\cdots}}_{\substack{\text{Deviation of estimated}\\ \text{treatment mean around}\\ \text{overall mean}}} + \underbrace{Y_{ijk} - \overline{Y}_{ij\cdot}}_{\substack{\text{Deviation}\\ \text{around estimated}\\ \text{treatment mean}}} \qquad \textbf{(19.36)}$$

Then we square both sides overall cases (i, j, k) (cross-product=0) to obtain

$$SSTO = SSTR + SSE$$

2. Further partition of Treatment Sum of Squares.

$$\underbrace{\overline{Y}_{ij\cdot} - \overline{Y}_{\cdots}}_{\substack{\text{Deviation of}\\ \text{estimated treatment}\\ \text{mean around}\\ \text{overall mean}}} = \underbrace{\overline{Y}_{i\cdot\cdot} - \overline{Y}_{\cdots}}_{\substack{A \text{ main}\\ \text{effect}}} + \underbrace{\overline{Y}_{\cdot j\cdot} - \overline{Y}_{\cdots}}_{\substack{B \text{ main}\\ \text{effect}}} + \underbrace{\overline{Y}_{ij\cdot} - \overline{Y}_{i\cdot\cdot} - \overline{Y}_{\cdot j\cdot} + \overline{Y}_{\cdots}}_{\substack{AB \text{ interaction}\\ \text{effect}}} \qquad \textbf{(19.38)}$$

Then we square both sides overall cases (i, j, k) (cross-product=0) to obtain

$$SSTR = SSA + SSB + SSAB$$

- where
  - SSA: factor A sum of squares, measures the variability of the estimated factor A level means
  - SSB: factor B sum of squares, measures the variability of the estimated factor B level means
  - $SSAB$: AB interaction sum of squares, measures the variability of the estimated interactions

3. Combined Partitioning in the first two steps:

$$SSTO = SSA + SSB + SSAB + SSE$$

Two-way ANOVA Table

Similar to one factor study, we divide the SS by its associated degrees of freedom to get the MS.

TABLE 19.8 ANOVA Table for Two-Factor Study with Fixed Factor Levels.

Source of Variation	SS	df	MS
Factor A	$SSA = nb\sum(\overline{Y}_{i\cdot\cdot} - \overline{Y}_{\cdots})^2$	$a-1$	$MSA = \frac{SSA}{a-1}$
Factor B	$SSB = na\sum(\overline{Y}_{\cdot j\cdot} - \overline{Y}_{\cdots})^2$	$b-1$	$MSB = \frac{SSB}{b-1}$
AB interactions	$SSAB = n\sum\sum(\overline{Y}_{ij\cdot} - \overline{Y}_{i\cdot\cdot} - \overline{Y}_{\cdot j\cdot} + \overline{Y}_{\cdots})^2$	$(a-1)(b-1)$	$MSAB = \frac{SSAB}{(a-1)(b-1)}$
Error	$SSE = \sum\sum\sum(Y_{ijk} - \overline{Y}_{ij\cdot})^2$	$ab(n-1)$	$MSE = \frac{SSE}{ab(n-1)}$
Total	$SSTO = \sum\sum\sum(Y_{ijk} - \overline{Y}_{\cdots})^2$	$nab-1$	

Then we can derive the expected values of these MS:

$$E\{MSE\} = \sigma^2 \tag{19.42a}$$

$$E\{MSA\} = \sigma^2 + nb\frac{\sum\alpha_i^2}{a-1} = \sigma^2 + nb\frac{\sum(\mu_{i.} - \mu_{..})^2}{a-1} \tag{19.42b}$$

$$E\{MSB\} = \sigma^2 + na\frac{\sum\beta_j^2}{b-1} = \sigma^2 + na\frac{\sum(\mu_{.j} - \mu_{..})^2}{b-1} \tag{19.42c}$$

$$E\{MSAB\} = \sigma^2 + n\frac{\sum\sum(\alpha\beta)_{ij}^2}{(a-1)(b-1)} \tag{19.42d}$$

$$= \sigma^2 + n\frac{\sum\sum(\mu_{ij} - \mu_{i.} - \mu_{.j} + \mu_{..})^2}{(a-1)(b-1)}$$

Which suggests that the test statistics based on the ratios of MSA , MSB, MSAB over MSE will provide information about the main effects and interactions of the two factors, with large values of the test statistics indicating the presence of these factor effects.

F-tests for Two-way ANOVA (Ch 19.6)

Test for Interaction Effect

$$\begin{aligned} &H_0\text{: all } (\alpha\beta)_{ij} = 0 \\ &H_a\text{: not all } (\alpha\beta)_{ij} \text{ equal zero} \end{aligned} \tag{19.43a}$$

As we noted from an examination of the expected mean squares in Table 19.8, the appropriate test statistic is:

$$F^* = \frac{MSAB}{MSE} \tag{19.44}$$

Large values of F^* indicate the existence of interactions. When H_0 holds, F^* is distributed as $F[(a-1)(b-1), (n-1)ab]$. Hence, the appropriate decision rule to control the Type I error at α is:

$$\begin{aligned} &\text{If } F^* \leq F[1-\alpha; (a-1)(b-1), (n-1)ab]\text{, conclude } H_0 \\ &\text{If } F^* > F[1-\alpha; (a-1)(b-1), (n-1)ab]\text{, conclude } H_a \end{aligned} \tag{19.45}$$

where $F[1-\alpha; (a-1)(b-1), (n-1)ab]$ is the $(1-\alpha)100$ percentile of the appropriate F distribution.

Test for Factor A main effects

$$H_0\colon \mu_{1\cdot} = \mu_{2\cdot} = \cdots = \mu_{a\cdot}$$
$$H_a\colon \text{not all } \mu_{i\cdot} \text{ are equal} \tag{19.46}$$

or equivalently:

$$H_0\colon \alpha_1 = \alpha_2 = \cdots = \alpha_a = 0$$
$$H_a\colon \text{not all } \alpha_i \text{ equal zero} \tag{19.46a}$$

we use the test statistic:

$$F^* = \frac{MSA}{MSE} \tag{19.47}$$

Again, large values of F^* indicate the existence of factor A main effects. Since F^* is distributed as $F[a-1, (n-1)ab]$ when H_0 holds, the appropriate decision rule for controlling the risk of making a Type I error at α is:

$$\text{If } F^* \leq F[1-\alpha; a-1, (n-1)ab], \text{ conclude } H_0$$
$$\text{If } F^* > F[1-\alpha; a-1, (n-1)ab], \text{ conclude } H_a \tag{19.48}$$

Test for Factor B main effects

$$H_0\colon \mu_{\cdot 1} = \mu_{\cdot 2} = \cdots = \mu_{\cdot b}$$
$$H_a\colon \text{not all } \mu_{\cdot j} \text{ are equal} \tag{19.49}$$

or equivalently:

$$H_0\colon \beta_1 = \beta_2 = \cdots = \beta_b = 0$$
$$H_a\colon \text{not all } \beta_j \text{ equal zero} \tag{19.49a}$$

The test statistic is:

$$F^* = \frac{MSB}{MSE} \tag{19.50}$$

and the appropriate decision rule for controlling the risk of a Type I error at α is:

$$\text{If } F^* \leq F[1-\alpha; b-1, (n-1)ab], \text{ conclude } H_0$$
$$\text{If } F^* > F[1-\alpha; b-1, (n-1)ab], \text{ conclude } H_a \tag{19.51}$$

Examples: two-factor ANOVA analysis

Two-Factor (two-way) ANOVA is very easy to implement in R. We specify the factor variable A and B, as $A * B$ or $A + B + A : B$ (exactly the same models).

```r
# fit = aov(response ~ height+ width +  height:width , data=Ex19 )

fit = aov(response~ height*width , data=Ex19 )
summary(fit)
```

```
##              Df Sum Sq Mean Sq F value   Pr(>F)    
## height        2   1544   772.0  74.710 5.75e-05 ***
## width         1     12    12.0   1.161    0.323    
## height:width  2     24    12.0   1.161    0.375    
## Residuals     6     62    10.3                     
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Results this example (page 845)

- The test F^*_{AB}= 1.161, with p-value=0.375, therefore, we can't reject the null hypothesis that display height and display width do not interact in their effects on sales
- The test F^*_{A}= 74.71, with p-value<0.001, therefore, we can reject the null hypothesis and conclude some definite effects associated with height of display level exist.
- The test F^*_{B}= 1.161, with p-value=0.323, therefore, we can't reject the null hypothesis that display width has no effect on sales.

What other additional analysis that could be done?

- We will need to do additional analyses to understand the height effect (factor A).
- There were three levels: bottom, middle and top. Based on the interaction picture, it appears the middle shelf increases sales.
- We could rerun the data with a one-way anova and use the methods we learned in the previous chapters to show this (e.g., Tukey procedure).

We can get the fitted value and residuals from model output.

```r
fit$fitted.values
```

```
##  1  2  3  4  5  6  7  8  9 10 11 12 
## 45 45 43 43 65 65 69 69 40 40 44 44
```

```r
fit$residuals
```

```
##  1  2  3  4  5  6  7  8  9 10 11 12 
##  2 -2  3 -3 -3  3 -2  2  1 -1 -2  2
```

We can also obtain the estimated main effects α_i, β_j and interactions $(\alpha\beta)_{ij}$ from model output.

But to set the zero-sum constraints, we need to specify the *contrasts* options before fitting the model. The *fit2* model has the same ANOVA table (SS partition and F-tests) but the main effects/interaction estimates are different compared to the previous *fit* model.

```r
options(contrasts = c("contr.sum", "contr.poly"))
fit2 <-  aov(response~ height*width , data=Ex19 )
summary(fit2)
```

```
##              Df Sum Sq Mean Sq F value   Pr(>F)
## height        2   1544   772.0  74.710 5.75e-05 ***
## width         1     12    12.0   1.161    0.323
## height:width  2     24    12.0   1.161    0.375
## Residuals     6     62    10.3
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

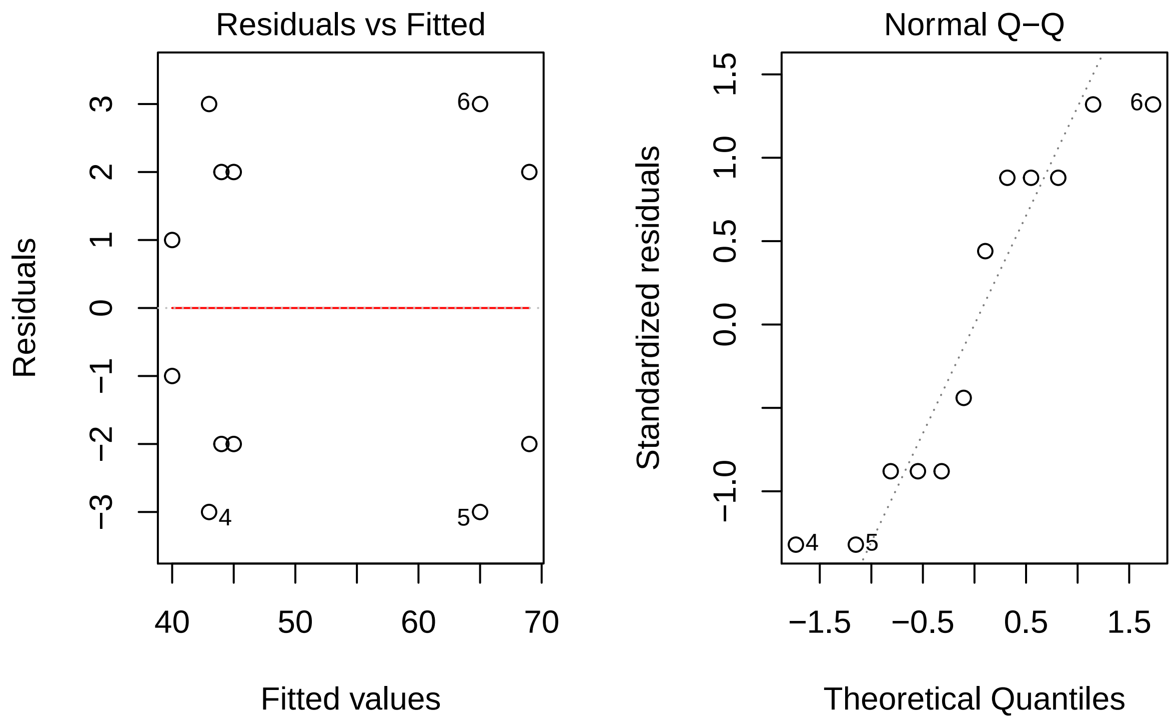
```r
dummy.coef(fit2)
```

```
## Full coefficients are
##
## (Intercept):      51
## height:            1   2   3
##                   -7  16  -9
## width:             1   2
##                   -1   1
## height:width:    1:1 2:1 3:1 1:2 2:2 3:2
##                    2  -1  -1  -2   1   1
```

Model diagnostics and Remedial measures (Ch 19.5)

- We still have similar diagnostic residual plots for two-way to test the constant variance and normal errors.

```r
par(mfrow=c(1,2))
plot(fit, 1)
plot(fit, 2)
```



- We can apply the normality test or tests for the constant variances that we studies in Chapter 18 and apply the remedial procedures such as weight least squares, Box-Cox transformation, and nonparametric tests if departure from the model assumptions are suspected or confirmed by the statistical tests.

Kimball Inequality

- If the test for interactions is conducted with level of significance α_1, that for factor A main effects with level of significance α_2, and that for factor B main effects with level of significance α_3, the level of significance α for the family of three tests is greater than the individual levels of significance.
- From the Bonferroni inequality (conservative), we can derive

$$\alpha \leq \alpha_1 + \alpha_2 + \alpha_3$$

- For the case considered here, a tighter inequality, **Kimball Inequality**, can be used, which utilizes the fact that the numerators of the three test statistics are independent and the denominator is the same in each case.

$$\alpha \leq 1 - (1 - \alpha_1)(1 - \alpha_2)(1 - \alpha_3)$$

- Example: When $\alpha_1 = \alpha_2 = \alpha_3 = 0.05$,

- Bonferroni inequality yields as the bound for the family level of significance:

$$\alpha \leq \alpha_1 + \alpha_2 + \alpha_3 = 0.15$$

- Kimball inequality yields the bound for the family level of significance:

$$\alpha \leq 1 - (1 - \alpha_1)(1 - \alpha_2)(1 - \alpha_3) = 1(.95)(.95)(.95) = .143$$

Summary this week

- Reading: Chapter 19.3-19.6
- HW: 19.10 [a-d, skip question (e)], 19.11, 19.30, 19.38 (due 10/10, next Thursday, by 6 pm before the class)