11 Tests for the Effects of Two Factors (Cont'd)

MTH 3240 Environmental Statistics

Spring 2020

Notes

Objectives

Objectives:

- Write out the group means and treatment effects versions of the two-factor ANOVA model, including any assumptions about the random error term ϵ. (**Optional for Spring 2020**)
- Interpret sums of squares, degrees of freedom, and mean squares in two-factor ANOVA.
- Carry out two-factor ANOVA F tests for the effects of two factors and their interaction effect.
- Interpret the interaction effect in a two-factor study.
- Obtain and interpret fitted values and residuals in two-factor ANOVA.
- Use plots to check the normality and common population standard deviation assumptions required by the ANOVA F tests.

Notes

Two-Factor ANOVA Models (Optional for Spring 2020)

- For **two-factor ANOVA**, we'll suppose each of the ab groups is a sample from a **normal** population.

  The mean of the population corresponding to the i, jth group is $\boldsymbol{\mu_{ij}}$.

  The population standard deviations are required to all the same, $\boldsymbol{\sigma}$.

Notes

(Optional for Spring 2020)

Factor B

		Level $j = 1$	Level $j = 2$	$\cdots$	Level $j = b$	
Factor A	Level $i = 1$	$\left.\begin{matrix} Y_{111} \\ Y_{112} \\ \vdots \\ Y_{11n} \end{matrix}\right\} \sim N(\mu_{11}, \sigma)$	$\left.\begin{matrix} Y_{121} \\ Y_{122} \\ \vdots \\ Y_{12n} \end{matrix}\right\} \sim N(\mu_{12}, \sigma)$	$\cdots$	$\left.\begin{matrix} Y_{1b1} \\ Y_{1b2} \\ \vdots \\ Y_{1bn} \end{matrix}\right\} \sim N(\mu_{1b}, \sigma)$	$\mu_{1\cdot}$
	Level $i = 2$	$\left.\begin{matrix} Y_{211} \\ Y_{212} \\ \vdots \\ Y_{21n} \end{matrix}\right\} \sim N(\mu_{21}, \sigma)$	$\left.\begin{matrix} Y_{221} \\ Y_{222} \\ \vdots \\ Y_{22n} \end{matrix}\right\} \sim N(\mu_{22}, \sigma)$	$\cdots$	$\left.\begin{matrix} Y_{2b1} \\ Y_{2b2} \\ \vdots \\ Y_{2bn} \end{matrix}\right\} \sim N(\mu_{2b}, \sigma)$	$\mu_{2\cdot}$
	$\vdots$	$\vdots$	$\vdots$	$\cdots$	$\vdots$	$\vdots$
	Level $i = a$	$\left.\begin{matrix} Y_{a11} \\ Y_{a12} \\ \vdots \\ Y_{a1n} \end{matrix}\right\} \sim N(\mu_{a1}, \sigma)$	$\left.\begin{matrix} Y_{a21} \\ Y_{a22} \\ \vdots \\ Y_{a2n} \end{matrix}\right\} \sim N(\mu_{a2}, \sigma)$	$\cdots$	$\left.\begin{matrix} Y_{ab1} \\ Y_{ab2} \\ \vdots \\ Y_{abn} \end{matrix}\right\} \sim N(\mu_{ab}, \sigma)$	$\mu_{a\cdot}$
		$\mu_{\cdot 1}$	$\mu_{\cdot 2}$	$\cdots$	$\mu_{\cdot b}$	μ

Notes

(Optional for Spring 2020)

- The value of μ_{ij} will depend on the level i of factor A and the level j of factor B.
- We can describe the data using the **statistical model** called the ***additive effects two-factor ANOVA model***.

  In the model, each μ_{ij} is written in terms of an **effect** of level i of factor A and an **effect** of level j of factor B.

Notes

Additive Two-Factor ANOVA Model (Optional for Spring 2020)

Additive Two-Factor ANOVA Model: One statistical model for describing data in a two-factor study is:

$$Y_{ijk} = \underbrace{\mu + \alpha_i + \beta_j}_{\text{This is } \mu_{ij}} + \epsilon_{ijk}\,,$$

where

Y_{ijk} is the kth observation at the ith level of factor A and jth level of factor B.
μ is a constant called the ***overall true mean***.
α_i is the ***effect*** of the ith level of factor A.
β_j is the ***effect*** of the jth level of factor B.
ϵ_{ijk} is a N(0, σ) ***random error*** term.

Notes

(Optional for Spring 2020)

Example

For the soil phosphorus study, the **additive effects model** for describing a phosphorus concentration Y is of the form

$$Y = \text{Overall Mean} + \text{Soil Type Effect} + \text{Topography Effect} + \text{Error}$$

- But the **additive effects model** requires that the **effects** of the two factors be ***additive***.
- They're ***additive*** when the effect of **factor** $\boldsymbol{A}$ is the **same regardless** of the **level** of **factor** $\boldsymbol{B}$, and the effect of factor B is the **same regardless** of the **level** of **factor** $\boldsymbol{A}$.

Notes

Interactions

- When the effects of two factors **aren't** additive, we say that there's an ***interaction effect*** between them.
- An ***interaction effect*** exists when the effect of **factor** $\boldsymbol{A}$ is **different** depending on the **level** of **factor** $\boldsymbol{B}$, and the effect of **factor** $\boldsymbol{B}$ is **different** depending on the **level** of **factor** $\boldsymbol{A}$.
- We can check for an **interaction effect** by graphing the **group means** either in a **bar plot** or in an ***interaction plot***.

Notes

Example

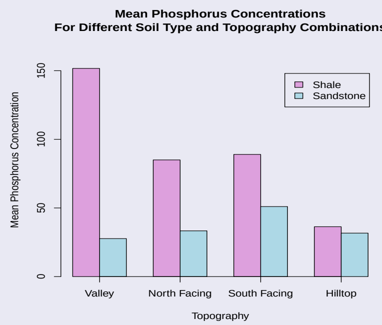
For the soil phosphorus study, the eight **group means** are shown in the **body** of the table below.

Factor B: Topography

		Valley (j=1)	North-Facing (j=2)	South-Facing (j=3)	Hilltop (j=4)	
Factor A: Soil Type	Shale (i=1)	$\bar{Y}_{11} = 151.7$	$\bar{Y}_{12} = 85.0$	$\bar{Y}_{13} = 89.0$	$\bar{Y}_{14} = 36.3$	$\bar{Y}_{1\cdot} = 90.5$
	Sandstone (i=2)	$\bar{Y}_{21} = 27.7$	$\bar{Y}_{22} = 33.3$	$\bar{Y}_{23} = 51.0$	$\bar{Y}_{24} = 31.7$	$\bar{Y}_{2\cdot} = 35.9$
		$\bar{Y}_{\cdot 1} = 89.7$	$\bar{Y}_{\cdot 2} = 59.2$	$\bar{Y}_{\cdot 3} = 70.0$	$\bar{Y}_{\cdot 4} = 34.0$	$\bar{Y} = 63.2$

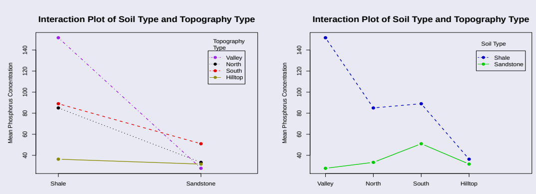
Notes

A **bar plot** of the **group means** is shown (again) below.



Notes

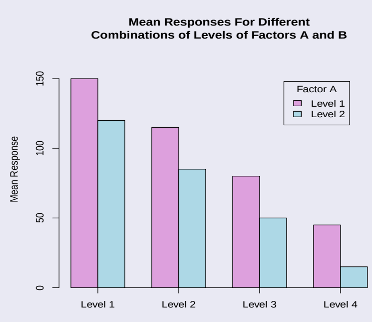
Here are **interaction plots** of the **group means**.



Notes

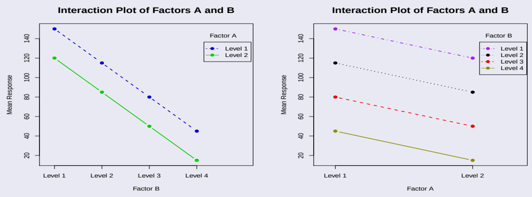
Example

Here's **bar plot** showing group means for two factors whose effects are **additive**.



Notes

Here are **interaction plots** of those same **group means** (for two factors whose effects are **additive**).



Notes

- When the lines in an **interaction plot** are **parallel**, the effects of the two factors are **additive**.

  When the lines **aren't parallel**, the effects **aren't additive**, i.e. there's an **interaction effect**.
- In practice:
  - We only have to look at one of the **interaction plots** or the other (not both).
  - The lines in an **interaction plot** won't ever be **exactly** parallel because of random ***sampling error***.

Notes

Two-Factor ANOVA Model With Interaction (Optional for Spring 2020)

- The model we'll use when the effects of two factors **aren't** additive is the ***two-factor ANOVA model with interaction effect***.

Two-Factor ANOVA Model With Interaction Effect: Another statistical model for describing data in a two-factor study is:

$$Y_{ijk} = \underbrace{\mu + \alpha_i + \beta_j + (\alpha\beta)_{ij}}_{\text{This is } \mu_{ij}} + \epsilon_{ijk} \, ,$$

Notes

(Optional for Spring 2020)

where

$\boldsymbol{Y_{ijk}}$ is the kth observation at the ith level of factor A and jth level of factor B.
$\boldsymbol{\mu}$ is a constant called the ***overall true mean***.
$\boldsymbol{\alpha_i}$ is the ***main effect*** of the ith level of factor A.
$\boldsymbol{\beta_j}$ is the ***main effect*** of the jth level of factor B.
$\boldsymbol{(\alpha\beta)_{ij}}$ is an ***interaction effect*** representing **non-additivity** in the effects of the two factors.
$\boldsymbol{\epsilon_{ijk}}$ is a N$(0, \sigma)$ ***random error*** term.

Notes

- We can fit **either ANOVA model** (the model **with** the **interaction** effect or the **additive** effects model) to data (using software), ...

  ... but the **model *with*** the **interaction** effect is **more widely applicable** because it **doesn't *require*** an assumption that the effects be **additive**, ...

  ... and it can still be used even if they ***are*** **additive**.

  So it's **"safer"** to **include** the **interaction effect** in the **model** when carrying out a **two-factor ANOVA**.

  From here on, we'll ***always*** **include** the **interaction effect** in the model.

Notes

Example

For the soil phosphorus study, the interaction plots suggest that the effects of **soil type** and **topography *aren't*** **additive**, so the we'd definitely use the **ANOVA model *with*** the **interaction effect**.

For a phosphorus concentration Y, the model has the form:

$$Y = \text{Overall Mean} + \text{Soil Type Effect} + \text{Topography Effect} + \text{Interaction Effect} + \text{Error}$$

Notes

Fitted Values and Residuals

- The **group means** $\bar{Y}_{11}, \bar{Y}_{12}, \ldots, \bar{Y}_{ab}$ are sometimes called ***fitted values***.

Fitted Values:

$$\text{Fitted Value for } ij\text{th Group} = \bar{Y}_{ij}$$

- **Comments**:
  - Statistical software reports n **duplicates** of the **fitted value** for **each** of the ab **groups**, one duplicate for each of the n individuals in the group.

Notes

Residuals

- A ***residual***, denoted e_{ijk}, is the **deviation** of an individual's observed Y_{ijk} value away from the **fitted value** for that individual.

Residuals:

$$e_{ijk} = Y_{ijk} - \bar{Y}_{ij}$$

- Statistical software reports the values of all N **residuals**, one for each individual in the study.

Notes

Sums of Squares

- Recall that we'll test for a **factor A main effect** by comparing the **between-rows variation** to **within-groups** variation, ...

  and we'll test for a **factor B main effect** by comparing the **between-columns variation** to **within-groups** variation.

  (We'll also test for an **interaction effect** by comparing **variation due to the interaction** to **within-groups** variation.)
- We measure the different types of variation using ***sums of squares*** (and later ***mean squares***).

Notes

- We measure the **between-rows variation** by the ***factor A sum of squares***, denoted **SSA**.

Factor A Sum of Squares:

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

SSA will be **large** when there's substantial variation among **row means** $\bar{Y}_{1\cdot}, \bar{Y}_{2\cdot}, \ldots, \bar{Y}_{a\cdot}$, which would suggest **factor A** has an **effect**.

Notes

- We measure the **between-columns variation** by the ***factor B sum of squares***, denoted **SSB**.

Factor B Sum of Squares:

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

SSB will be **large** when there's substantial variation among **column means** $\bar{Y}_{\cdot 1}, \bar{Y}_{\cdot 2}, \ldots, \bar{Y}_{\cdot b}$, which would suggest **factor B** has an **effect**.

Notes

- We measure **variation** due to an **interaction effect** by the ***AB interaction sum of squares***, denoted **SSAB**.

AB Interaction Sum of Squares:

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

It can be shown that SSAB will be **large** when the **group means** $\bar{Y}_{11}, \bar{Y}_{12}, \ldots, \bar{Y}_{ab}$ aren't consistent with an additive effects model, which would suggest there's an **interaction effect** between **factors A** and **B** (i.e. their effects aren't additive).

Notes

- We measure **within-groups variation** by the ***error sum of squares***, denoted **SSE**.

Error Sum of Squares:

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

The **error sum of squares is just the *sum of squared residuals***, and reflects variation due to **random error**.

SSE will be **large** if there's **substantial** variation among individual observations (Y_{ijk}'s) **within** groups, and **small** otherwise.

Notes

The ANOVA Partition

- The ***total sum of squares***, denoted **SSTo**, measures **total variation** in the data.

Total Sum of Squares:

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

SSTo reflects **total variation** due both to effects of the factors (if they have any effects) **and** random variation among individual observations within groups.

Notes

- It can be shown that

Two-Factor ANOVA Partition:

$$\text{SSTo} = \text{SSA} + \text{SSB} + \text{SSAB} + \text{SSE}.$$

This splits the total variation in the data as:

$$\begin{aligned} \textbf{Total Variation} = \; & \textbf{Between-Rows Variation} \\ & + \textbf{Between-Columns Variation} \\ & + \textbf{Variation Due to Interaction} \\ & + \textbf{Within-Groups Variation} \end{aligned}$$

Notes

Example

For the soil phosphorus study, statistical software reports the following **sums of squares**.

SSTo	=	51406.0	(Total variation)
SSA	=	17876.0	(Variation due to soil type)
SSB	=	9693.8	(Variation due to topography)
SSAB	=	11390.8	(Variation due to interaction)
SSE	=	12445.3	(Variation due to random error)

We see that the **two-factor ANOVA partition** holds since

$$51406.0 = 17876.0 + 9693.8 + 11390.8 + 12445.3.$$

Notes

This indicates that a large portion of the total variation in phosphorus measurements (**17876.0** out of **51406.0**, or **35%**) is due to the difference between the row means (soil types).

Notes

Degrees of Freedom

- Here are the ***degrees of freedom*** associated with the sums of squares in two-factor ANOVA. (These will determine which F distributions our p-values come from.)

Degrees of Freedom: For two-factor ANOVA, the degrees of freedom are:

$$\begin{aligned} df \text{ for SSTo} &= N-1 \\ df \text{ for SSA} &= a-1 \\ df \text{ for SSB} &= b-1 \\ df \text{ for SSAB} &= (a-1)(b-1) \\ df \text{ for SSE} &= ab(n-1) = N-ab \end{aligned}$$

Notes

- The degrees of freedom, like the associated sums of squares, are additive in the following sense.

Additivity of Degrees of Freedom:

df for SSTo $=$ df for SSA$+df$ for SSB$+df$ for SSAB$+df$ for SSE.

Notes

Example

For the soil phosphorus study, we have $a = 2$ soil types, $b = 4$ topographies, and $n = 3$ phosphorus observations per group. Thus the total number of phosphorus observations (overall sample size) is $N = 24$, and

$$\begin{aligned} df \text{ for SSTo} &= 23 \\ df \text{ for SSA} &= 1 \\ df \text{ for SSB} &= 3 \\ df \text{ for SSAB} &= 3 \\ df \text{ for SSE} &= 16. \end{aligned}$$

As expected,

$$23 = 1 + 3 + 3 + 16.$$

Notes

Mean Squares

- A ***mean square*** is a **sum of squares** divided by its **degrees of freedom**.
- The ***factor A mean square***, ***factor B mean square***, AB ***interaction mean square***, and ***mean squared error*** are below.

Mean Squares: For two-factor ANOVA, the mean squares are

$$\text{MSA} = \frac{\text{SSA}}{a-1} \qquad \text{MSB} = \frac{\text{SSB}}{b-1}$$

$$\text{MSAB} = \frac{\text{SSAB}}{(a-1)(b-1)} \qquad \text{MSE} = \frac{\text{SSE}}{ab(n-1)}$$

Notes

The ANOVA F Tests

- There are **three** sets of hypotheses for the ***two-factor ANOVA F tests*** tests.

  In each case, H_0 says there's **no effect**, and H_a say's there's **an effect**.

Notes

1. Test for a ***factor A main effect***:

   H_{OA} : There's no factor A effect
   H_{aA} : There is a factor A effect

2. Test for a ***factor B main effect***:

   H_{OB} : There's no factor B effect
   H_{aB} : There is a factor B effect

3. Test for an AB ***interaction effect***:

   H_{OAB} : There's no factor A and B interaction effect
   H_{aAB} : There is a factor A and B interaction effect

Notes

(Optional for Spring 2020)

In terms of the **ANOVA model parameters** these are written as:

1. Test for a ***factor A main effect***:

   H_{OA} : $\alpha_1 = \alpha_2 = \cdots = \alpha_a = 0$
   H_{aA} : The α_i's don't all equal 0

2. Test for a ***factor B main effect***:

   H_{OB} : $\beta_1 = \beta_2 = \cdots = \beta_b = 0$
   H_{aB} : The β_j's don't all equal 0

3. Test for an AB ***interaction effect***:

   H_{OAB} : $(\alpha\beta)_{11} = (\alpha\beta)_{12} = \cdots = (\alpha\beta)_{ab} = 0$
   H_{aAB} : The $(\alpha\beta)_{ij}$'s don't all equal 0

Notes

- Here are the corresponding ***two-factor ANOVA F test statistics***:

Two-Factor ANOVA F Test Statistics:

$$F_A = \frac{\text{MSA}}{\text{MSE}}$$

$$F_B = \frac{\text{MSB}}{\text{MSE}}$$

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

Notes

- Not that we can think of

$$F_A = \frac{\textbf{Between-Rows Variation}}{\textbf{Within-Groups Variation}}.$$

$$F_B = \frac{\textbf{Between-Columns Variation}}{\textbf{Within-Groups Variation}}.$$

$$F_{AB} = \frac{\textbf{Variation Due To Interaction}}{\textbf{Within-Groups Variation}}.$$

Notes

- In each case, **if H_0 was true**, it can be shown, F would be **approximately equal to 1**.
- But if **if H_0 was true** was true, F would be **greater than 1**.

In each case, *large* values of F (larger than about 1) provide evidence in favor of H_a.

Notes

- Now suppose the ab groups are samples from **normal** populations that all have the **same standard deviation σ** (or that they have the **same σ** and the common sample size n is **large**).

  In this case, the **null distributions** are as follows.

Notes

Sampling Distribution of F Under H_0: If F_A, F_B, and F_{AB} are the two-factor ANOVA F test statistics, then:

1. When H_{OA} is true,

$$F_A = \frac{\text{MSA}}{\text{MSE}} \sim F\,(a-1,\; N-ab)\,.$$

2. When H_{OB} is true,

$$F_B = \frac{\text{MSB}}{\text{MSE}} \sim F\,(b-1,\; N-ab)\,.$$

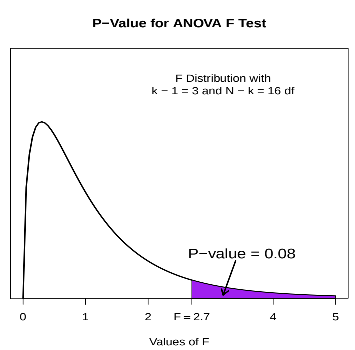
3. When H_{OAB} is true,

$$F_{AB} = \frac{\text{MSAB}}{\text{MSE}} \sim F\,((a-1)(b-1),\; N-ab)\,.$$

Notes

- **P-values** and **rejection regions** are obtained from the *upper tail* of the **appropriate F distribution**.
- The next slide shows the **p-value** when the observed test statistic value is $F = 2.7$.

Notes



Notes

The ANOVA Table

- The results of an **analysis of variance** are summarized in a ***two-factor ANOVA table*** having the form shown below.

Two-Factor ANOVA Table:

Source	DF	SS	MS	F	P-value
Factor A	$a-1$	SSA	MSA = SSA/$(a-1)$	MSA/MSE	p
Factor B	$b-1$	SSB	MSB = SSB/$(b-1)$	MSB/MSE	p
Interaction	$(a-1)(b-1)$	SSAB	MSAB = SSAB/$((a-1)(b-1))$	MSAB/MSE	p
Error	$ab(n-1)$	SSE	MSE = SSE/$(ab(n-1))$		
Total	$N-1$	SSTo			

Notes

Procedure for Assessing Significance in Two-Factor ANOVA:

- **First** look at the results for the AB **interaction effect**.
  1. If it's **statistically significant**, there's **no need proceed** to the results of the tests for main effects since we already know that both factors have effects.

     (But the effect of each factor is different depending on the level of the other factor.)
  2. If the AB interaction ***isn't* significant**, **proceed** to the results of the tests for **factor** A and B **main effects**.

Notes

Example

For the soil phosphorus study, the **ANOVA table** (obtained using software) is below.

Source of Variation	df	Sum of Squares	Mean Square	f	P-value
Soil Type	1	17876.0	17876.0	22.98	0.000
Topography Type	3	9693.8	3231.3	4.15	0.024
Interaction	3	11390.8	3796.9	4.88	0.013
Error	16	12445.3	777.8		
Total	23	51406.0			

What **conclusions** are appropriate (using $\alpha = 0.05$)?

Notes

We **first** assess significance of the **interaction effect**.

The p-value for the interaction is **0.013**, so we conclude there's a **statistically significant** interaction effect between **soil type** and **topography**.

In this case, there's **no need to proceed** to look at the p-values for the **soil type** and **topography main effects**.

We **already know** (because the interaction is significant) that each has an effect on phosphorus, **regardless** of what their p-values are.

Each factor has a **different effect** depending on the **level** of the **other factor**.

Notes

Checking the ANOVA Assumptions

- The **ANOVA** F **tests** require that the ab groups (samples) are from **normal** populations (or that their sample sizes are **large**) whose **standard deviations** are all **equal**.

Notes

- Two ways to check the **normality assumption**:
  - Make ab *separate* **histograms** or **normal probability plots**, one for each of the ab **groups**.
  - Make a *single* **histogram** or **normal probability plot** plot of the N **residuals** e_{ijk}.

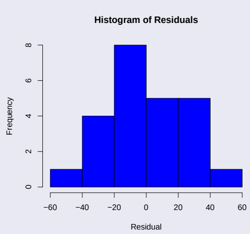
Notes

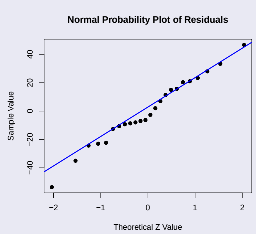
- A few ways to check the **equal population standard deviation** assumption:
  - An ***individual value plot*** of the ab groups.
  - A plot of the ***residuals*** (y-axis) versus **fitted values** (group means, x-axis).
- In both plots, we look for roughly **equal amounts of within-group (vertical) spread** across the ab groups.

Notes

Example

For the soil phosphorus study, a **histogram** and **normal probability plot** of the **residuals** are below.

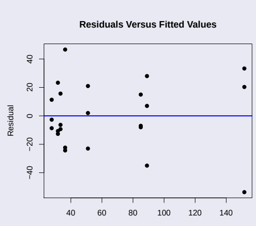




Notes

The plots show that the **normality assumption** appears to be met.

A plot the **residuals** versus the **fitted values** (**group means**) is below.



Notes

The amount of (vertical) spread of the points is roughly the same from left to right, so the **equal standard deviation assumption** appears to be met.

Because the **normality** and **equal standard deviation assumptions** are met, the results of the F **tests** are **valid**.

Notes

Notes

Notes

Notes