



PhUSE 2010
Paper SPo6



Quanticate

CONFIDENCE COMES FROM QUALITY

Useful Tips for Analysis of Variance (ANOVA) in Multicenter Placebo Controlled Clinical Trials

Joanna Romaniuk
Quanticate, Warsaw, Poland

Plan of the presentation



- **What is Analysis of Variance (ANOVA)?**
- **Fundamentals of the ANOVA**
- **Balanced and Unbalanced Data**
- **Model Assumptions**
- **Conclusions**

What is Analysis of Variance (ANOVA)?

- **ANOVA** is a statistical tool used to identify differences between experimental group means.



- **Analysis of Variance (ANOVA)** is commonly performed on the data coming from multicenter placebo controlled clinical trials in order to evaluate the size of the difference in efficacy between the **study medication** and **placebo**.



What is Analysis of Variance (ANOVA)?

When the difference in efficacy between the **study medication** and **placebo** is significant it can be assumed that:

Study medication is more effective than placebo.

Fundamentals of the ANOVA

- **ANOVA** method seeks to detect **sources of variation** in the values of **dependent variable** and divide the total variability into components associated with **each source**.



- The **total variability** is the **sum of squared deviations of each measurement from the overall mean** and can be decomposed into a sum of squares (SS) due to suspected sources of variation (**model sum of squares**) and a **sum of squares (SS) resulting from the error**:

$$SS(Total) = SS(Model) + SS(Error)$$

Fundamentals of the ANOVA

ANOVA Table:

<i>Source of variation</i>	<i>Sum of squares</i>	<i>DF</i>	<i>Mean Square</i>	<i>F Statistic</i>
<i>Model</i>	$SS(Model)$	$k - 1$	$SS(Model)/(k - 1)$	$F = MS(Model)/MS(Error)$
<i>Error</i>	$SS(Error)$	$N - k$	$SS(Error)/(N - k)$	
<i>Total</i>	$SS(Total)$	$N - 1$		

k - number of categories (groups) of the factor
(independent variable),

N - number of observations from all groups.



Balanced and unbalanced data

- **Balanced design** - all cells sizes are exactly equal.

An example of balanced data design:

<i>Table of Treatment by Center</i>							
<i>Treatment</i>	<i>Center</i>						<i>Total</i>
<i>Frequency</i>	<i>1</i>	<i>2</i>	<i>3</i>	<i>4</i>	<i>5</i>	<i>6</i>	
<i>A</i>	6	6	6	6	6	6	36
<i>B</i>	6	6	6	6	6	6	36
<i>Placebo</i>	6	6	6	6	6	6	36
<i>Total</i>	18	18	18	18	18	18	108

Balanced and unbalanced data

- **Unbalanced design** - one in which the cells sizes are not exactly equal or/and some data is missing.

<i>Table of Treatment by Center</i>							
<i>Treatment</i>	<i>Center</i>						<i>Total</i>
<i>Frequency</i>	<i>1</i>	<i>2</i>	<i>3</i>	<i>4</i>	<i>5</i>	<i>6</i>	
<i>A</i>	10	5	6	0	9	18	48
<i>B</i>	10	5	0	7	10	18	50
<i>Placebo</i>	0	5	7	7	9	18	46
<i>Total</i>	20	15	13	14	28	54	144

When data design is unbalanced the use of simple ANOVA statistical procedures is not appropriate!

Balanced and unbalanced data



Solution to the problem of unbalanced data:

choose the **appropriate** Sum of Squares Test out of four tests available in SAS®.

- **SAS® Type I sums of squares** → Each term is adjusted for all terms previously fit in the model. **Type I Test is suitable only for balanced designs.**
- **Type II sums of squares** → Main effects are adjusted for the other, ignoring the interaction effects. **Type II sums of squares are inappropriate if the interaction term cannot be assumed to be zero.**

Balanced and unbalanced data

- **Type III sums of squares** (**recommended for general use in the ANOVA**) → Every effect is adjusted for all other effects listed in the model statement.
- **Type IV sums of squares** are preferred if any cell size equals zero.

Balanced and unbalanced data

Unbalanced data requires Type II, III or IV sums of squares.

Sums of squares for unbalanced data are computed with the use of **least squares means** (the estimates for group means obtained from the ANOVA model).

Balanced and unbalanced data

Assume analyzing data from multicenter placebo-controlled clinical trial with three treatment groups (A, B and Placebo) performed in 6 sites (1, 2, 3, 4, 5, 6). The primary endpoint is the worst possible pain score rated by patients within 24 hour post surgery. Data extract can be seen below:

<i>Obs</i>	<i>Subject</i>	<i>Center</i>	<i>Race</i>	<i>Treatment</i>	<i>Pain</i>
1	1001	1	Black	A	1
2	1002	1	Black	B	2
3	1003	1	Black	Placebo	10
4	1004	1	Black	A	2
5	1005	1	Black	B	7
6	1006	1	Black	Placebo	9
7	1007	1	Black	A	3
8	1008	1	Black	B	8
9	1009	1	Black	Placebo	8
10	1010	1	Black	A	3
...	...	...	...	...	...

Balanced and unbalanced data

- In order to investigate the design of the data the PROC FREQ procedure has to be performed:

```
proc freq data=pain;  
    tables treatment*center/ norow nocol nopercnt;  
run;
```

Balanced and unbalanced data

- The procedure generates cross-table by treatment and center.

<i>Table of Treatment by Center</i>							
<i>Treatment</i>	<i>Center</i>						<i>Total</i>
<i>Frequency</i>	<i>1</i>	<i>2</i>	<i>3</i>	<i>4</i>	<i>5</i>	<i>6</i>	
<i>A</i>	10	5	6	0	9	18	48
<i>B</i>	10	5	0	7	10	18	50
<i>Placebo</i>	0	5	7	7	9	18	46
<i>Total</i>	20	15	13	14	28	54	144

Balanced and unbalanced data



- The PROC GLM procedure generates different types of sums of squares :

```
proc glm data=pain;  
class treatment center;  
model pain=treatment center treatment*center / ss1 ss2 ss3 ss4;  
lsmeans treatment / pdiff;  
output out=predicted r=Residual predicted=Predicted  
rstudent=Studentiz;  
run;
```

Balanced and unbalanced data

- Different sums of squares :

Source	DF	Type I SS	Mean Square	F Value	Pr > F
Treatment	2	19.44379	9.72189	3.06	0.0502
Center	5	46.28160	9.25632	2.92	0.0158
Treatment*Center	7	195.46173	27.92310	8.80	<.0001

Source	DF	Type II SS	Mean Square	F Value	Pr > F
Treatment	2	28.74791	14.37395	4.53	0.0126
Center	5	46.28160	9.25632	2.92	0.0158
Treatment*Center	7	195.46173	27.92310	8.80	<.0001

Source	DF	Type III SS	Mean Square	F Value	Pr > F
Treatment	2	35.60088	17.80044	5.61	0.0046
Center	5	45.97286	9.19457	2.90	0.0164
Treatment*Center	7	195.46173	27.92310	8.80	<.0001

Source	DF	Type IV SS	Mean Square	F Value	Pr > F
Treatment	2	27.21133	13.60566	4.29	0.0158
Center	5	40.85383	8.17076	2.57	0.0296
Treatment*Center	7	195.46173	27.92310	8.80	<.0001



Model assumptions

Error components associated with the scores of the dependent variable should be:

- independent of each other,
- normally distributed with zero mean and an unknown but fixed variance.

Model assumptions

Verification of model assumptions:

- (1) independent error terms → scatter plot between the predicted values and the residuals
(a residual plot should have a **random distribution**).
- (2) homogeneity → box plots by treatments.
- (3) normality → normal probability plot.

Model assumptions

The example of SAS® code that might be useful in the verification of model assumptions is presented below:

```
proc glm data=pain;
  class treatment center;
  model pain=treatment center treatment*center ;
  lsmeans treatment / pdiff;
  output out=predicted r=Residual predicted=Predicted
  rstudent=Studentiz;
run;

goptions reset = all;
libname gfont0 (sashelp);
goptions keymap=win1250;
goptions device=win target=winprtm FTEXT=centxi htext=3 HSIZE=4 VSIZE=3;
symbol1 color=black value=dot;
proc univariate data=pred1 normal;
  var residual;
  histogram residual / midpoints= 0 1 2 3 4 nrows=4 normal;
  qqplot residual;
run;

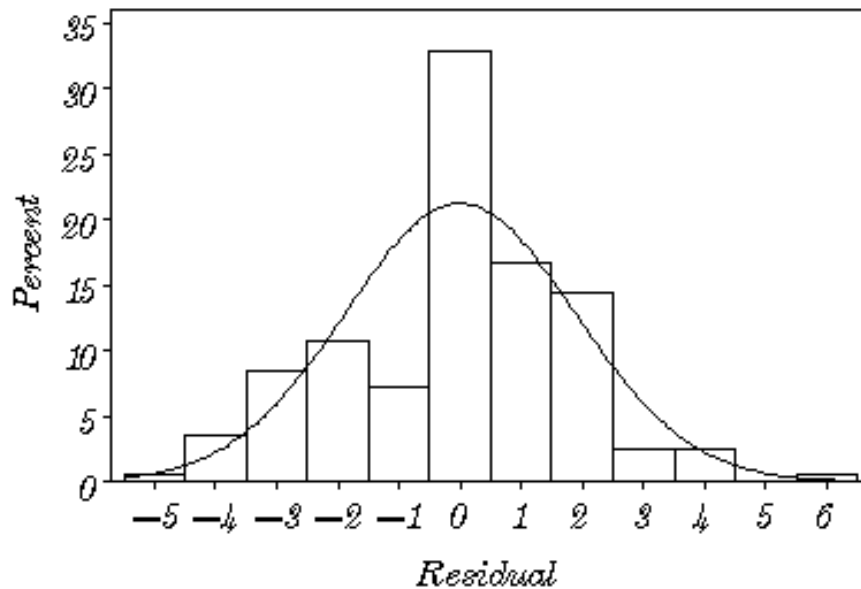
proc gplot data=pred1;
  plot Residual*Predicted Residual*Observation Studentiz*Predicted;
run;
quit;

SYMBOL1 INTERPOL=BOX VALUE=NONE c=blue width = 1 bwidth = 1;
goptions hsize=4 vsize=3 htext=1;

proc gplot data=pred1 ;
  plot residual*treatment / hminor=0;
run;
quit;
```

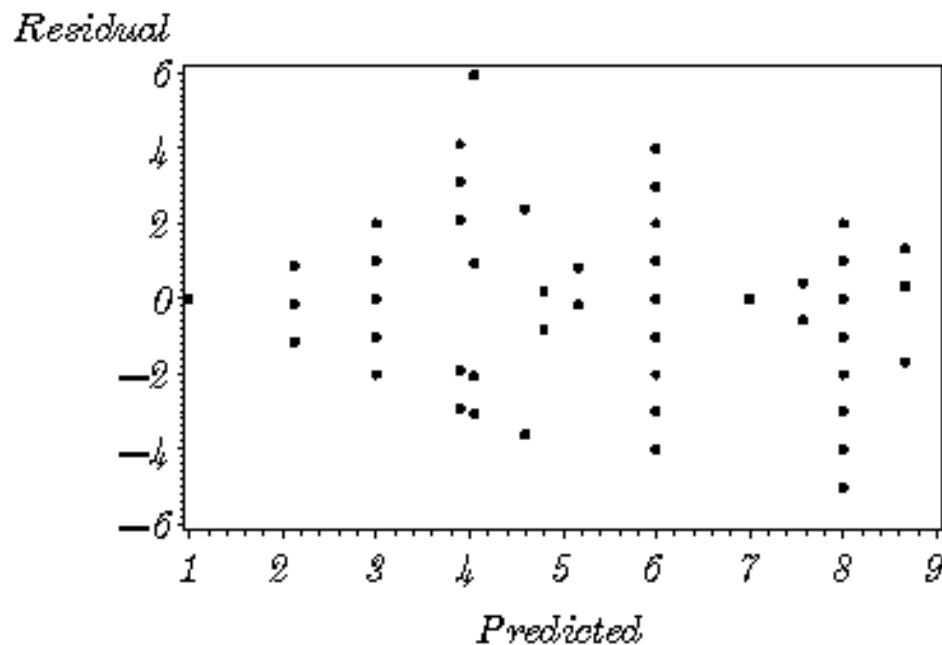
Model assumptions

- Histogram of residuals indicates non-normality:



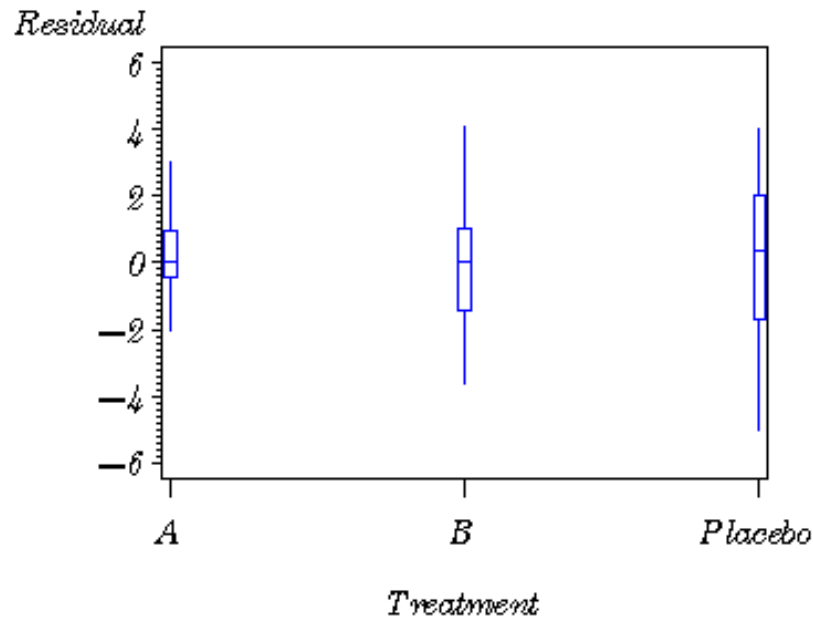
Model assumptions

- Residual vs Predicted values scatter plot does not show any systematic unexplained or cyclic pattern.



Model assumptions

- Box plots generated for residuals for each treatment group show unequal variances.



Model assumptions

When the data seriously violates ANOVA assumptions, researchers have a few options:

- detect outliers,
- apply a transformation to the response variable,
- use a non-parametric (rank based) test,
- fit a different model, one that requires different distributional assumptions.

Model assumptions

- **Detection of outliers,**
- **Data transformations.**

Outliers → cases with unusual or extreme values on a particular variable.

Outliers detection → by plotting the standardized residuals against predicted values.

Absolute value of the standardized residual greater than 2.5 → OUTLIER.

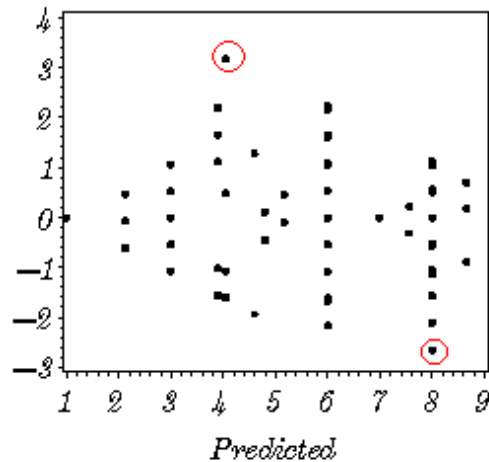
Always verify whether outliers result from the experimental error and if so, they should be eliminated from the analyses or adequately adjusted to the distribution of the empirical data.

Model assumptions

Detection of outliers:

```
proc glm data=pain;
  class treatment center;
  model pain=treatment center treatment*center ;
  lsmeans treatment / pdiff;
  output out=predicted r=Residual predicted=Predicted
  rstudent=Studentiz; ←
run;
```

Studentized residual



Model assumptions

Data can be used to estimate the appropriate transformation.
Box and Cox proposed the power transformation where:

$$Y_t = \begin{cases} y^\lambda & \text{when } \lambda \neq 0 \\ \log y & \text{when } \lambda = 0 \end{cases}$$

Y_t is the transformed response

λ is the integer varying over the range of -3 to 3.

Model assumptions

- The most appropriate transformation can be easily determined by the SAS® system using the PROC TRANSREG procedure:

```
data pain1;  
    set pred1;  
    pain=pain+1;  
    if (studentiz gt 2.5 or studentiz lt -2.5) then delete;  
run;  
  
proc transreg data=pain1;  
    model boxcox(pain)=class(treatment center treatment*center);  
run;
```

Model assumptions

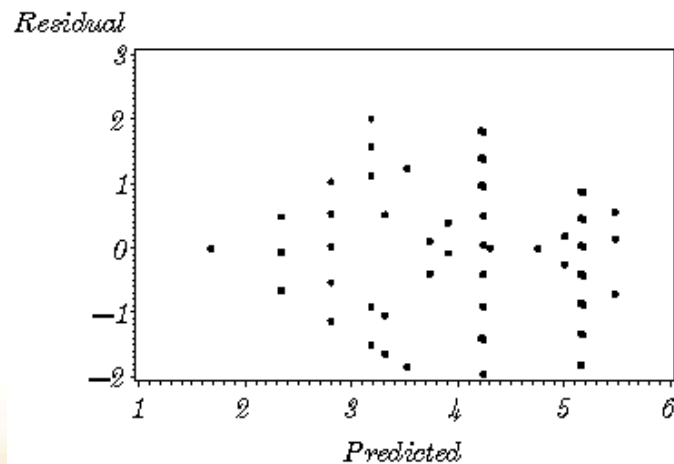
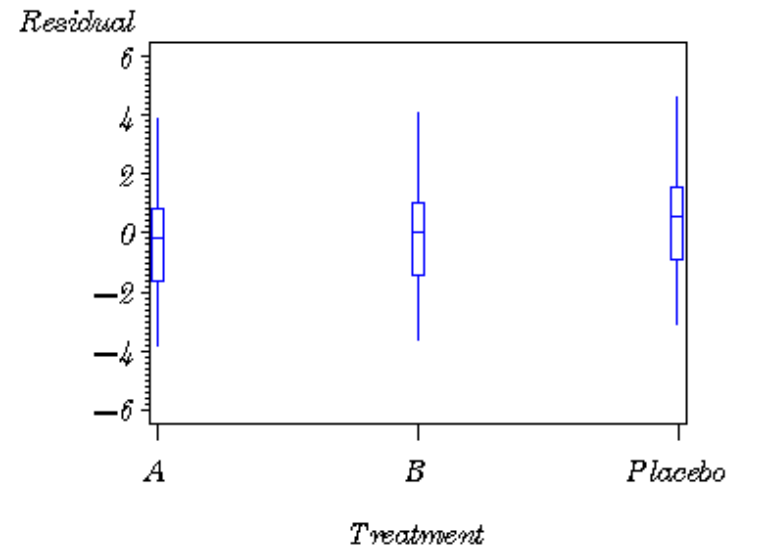
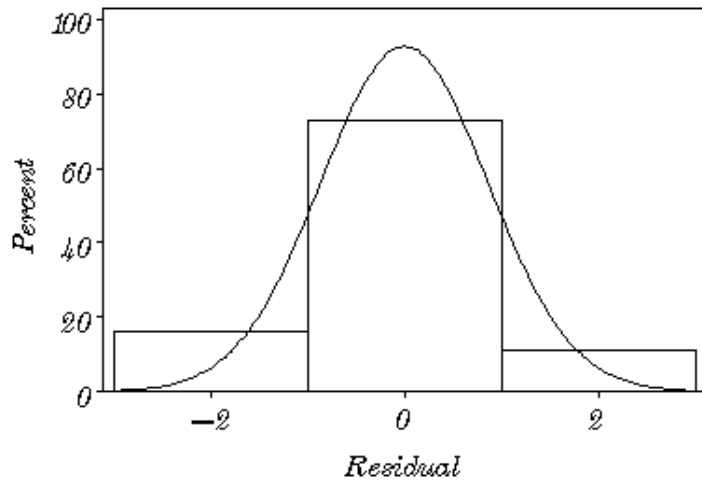
- Results of the PROC TRANSREG:
Best transformation: with $\text{Lambda}=0.75$.

<i>Transformation Information for BoxCox(Pain)</i>				
<i>Lambda</i>		<i>R-Square</i>	<i>Log Like</i>	
-3.00		0.59	-391.519	
-2.00		0.59	-273.629	
-1.00		0.60	-180.807	
0.50		0.59	-114.446	*
0.75		0.59	-113.141	<
1.00	+	0.58	-114.409	*
2.00		0.54	-140.730	
3.00		0.50	-190.481	
< - Best Lambda * - Confidence Interval + - Convenient Lambda				

```
data pain2;
  set pain1;
  pain1=pain** .75;
run;
```

Model assumptions

Verification of ANOVA assumptions for the transformed data:



Model assumptions

Results obtained from ANOVA model for transformed data:

<i>Source</i>	<i>DF</i>	<i>Sum of Squares</i>	<i>Mean Square</i>	<i>F Value</i>	<i>Pr > F</i>
<i>Model</i>	17	181.94724	10.70277	14.18	<.0001
<i>Error</i>	147	110.91501	0.75452		
<i>Corrected Total</i>	164	292.86226			

<i>Source</i>	<i>DF</i>	<i>Type III SS</i>	<i>Mean Square</i>	<i>F Value</i>	<i>Pr > F</i>
<i>Treatment</i>	2	80.621579	40.310789	53.43	<.0001
<i>Center</i>	5	30.052794	6.010558	7.97	<.0001
<i>Treatment *Center</i>	10	40.867448	4.086744	5.42	<.0001

Model assumptions

Post-hoc test adequate for unbalanced data:



<i>Treatment</i>	<i>Pain1 LSMEAN</i>	<i>LSMEAN Number</i>
<i>A</i>	3.510432	1
<i>B</i>	4.612522	2
<i>Placebo</i>	5.347940	3

<i>Least Squares Means for effect Treatment</i> <i>Pr > t for H0: LSMean(i)=LSMean(j)</i> <i>Dependent Variable: pain1</i>			
<i>i/j</i>	<i>1</i>	<i>2</i>	<i>3</i>
<i>1</i>		<.0001	<.0001
<i>2</i>	<.0001		<.0001
<i>3</i>	<.0001	<.0001	

Conclusions



In order to properly conduct ANOVA, the analyst should:

- (1) understand how an unbalanced data set differs from a balanced one;**
- (2) know what sums of squares can be computed in SAS® and how to choose the best one for the given data design;**
- (3) check for the existence of the outliers;**
- (4) always verify model assumptions and, if they are not fulfilled, apply an adequate transformation to the response variable or use a non-parametric test or fit a different model, one that requires different distributional assumptions.**

Thank you!

Contact Information

Joanna Romaniuk

Quanticate Polska Sp. z o.o.

Hankiewiczza 2

02-103 Warsaw

Poland

Tel: +48(0) 22 576 21 40

Fax: +48(0) 22 576 21 59

E-mail: joanna.romaniuk@quanticate.com

Brand and product names are trademarks of their respective companies.