

LAB 5 INSTRUCTIONS

TWO-WAY ANOVA

The General Linear Model (GLM) provides a wide range of statistical tools to describe the relationship between a response variable and a set of explanatory variables. It covers a variety of linear models including linear regression and ANOVA. Here we will discuss a special case of GLM, the Univariate GLM which compares the means of several groups for one response variable.

The *Univariate GLM* in SPSS can run regression models. Indeed, it is enough to enter the explanatory quantitative variables as covariates and any explanatory categorical variables should be entered as fixed factors; there is no need to create dummy variables. The estimates of the regression coefficients obtained with the Univariate GLM procedure in SPSS will be identical to those obtained with *Linear Regression* tool. However, in GLM the estimates are provided on request only (the "Parameter estimates" under the *Options* button in the GLM dialog box should be selected), whereas the estimates are provided automatically in *Linear Regression* procedure. One of the advantages of GLM over *Linear Regression* tool in SPSS is the possibility of testing directly for interactions between the explanatory variables (regression models cannot handle interaction unless explicit cross-product interaction terms are added).

Though regression models may be run without difficulty in GLM, in practice univariate GLM is used mainly to run analysis of variance (ANOVA) and analysis of covariance (ANCOVA). As we discussed in *Lab 2 Instructions*, analysis of variance uses two types of variables: variables whose group means are compared (response variables) and variables that determine the grouping (factors). In the Univariate GLM only one response variable is used. Analysis of variance is used to examine the main and interaction effects of categorical explanatory variables (factors) on a quantitative response variable. A "main effect" is the direct effect of an explanatory variable on the response variable. An "interaction effect" is the joint effect of two or more explanatory variables on the response variable.

In this lab you will learn how to use the GLM tool in SPSS to examine the association between the score on an essay and two factors: typing ability and writing method. Moreover, you will explore the interaction between the two factors.

We will demonstrate these features of SPSS using a simple example with three explanatory variables. The data come from "Multivariate Statistics: Concepts, Models and Applications" by David W. Stockburger.

Example: Suppose an instructor is giving an essay to his class. He randomly divides the class into two groups such that half the class writes the essay with a blue-book and the other half with notebook computers. In addition, the students are divided into three groups, according to their typing skills. Answers written in blue-books will be transcribed to word processors and scoring will be done blindly. More precisely, the instructor will not know the method or skill level of the student when scoring the essay. The response variable will be the score on the essay.

<u>Column</u>	<u>Variable Name</u>	<u>Description of Variable</u>
1	ability	Typing ability (0=none, 1=some, 2=lots),
2	method	Writing method (0=blue-book, 1=notebook computer),
3	score	Score on the essay.

The data can be downloaded by clicking on the link below:

[DOWNLOAD DATA](#)

Before we apply the GLM tool in SPSS to our data, we will explore the differences in average scores across both factors with graphs of marginal means, clustered boxplots and summary univariate statistics.

1. Table of Marginal Means

The table of marginal means (means of levels of one factor collapsed over levels of the other factor) may provide an important insight into the relationship between a quantitative response variable and the categorical variables (factors). Though the table can be requested in the dialog box for GLM (*Descriptive Statistics* in *Univariate: Options* dialog box), it is recommended to obtain the table at the exploration stage of data analysis.

Let us consider two factors A and B, each at three levels A1, A2, A3 and B1, B2 and B3, respectively. Suppose the means for each combination of the levels of the two factors are given in the following table:

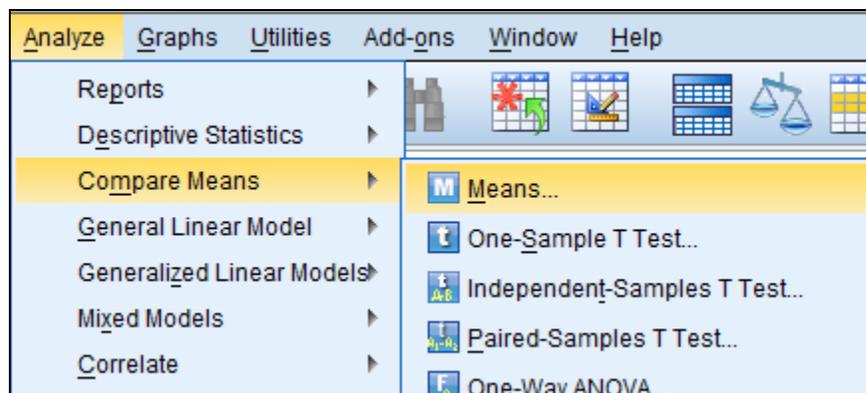
A/B	B1	B2	B3	Average
A1	●	●	●	●
A2	●	●	●	●
A3	●	●	●	●
Average	●	●	●	●

The above table is called a *table of marginal means*. The means inside the boxes are called *cell means*, the means in the margins are called *marginal means*, and the number on the bottom right-hand corner is called the *grand mean*.

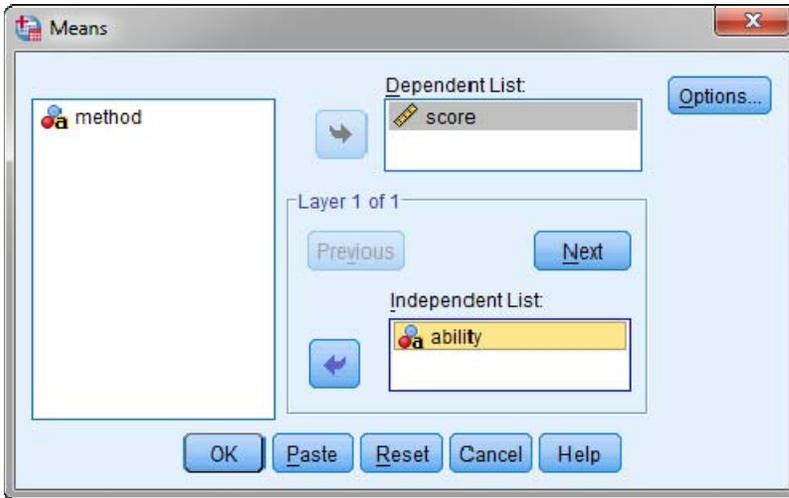
The table of marginal means can be used to assess the main and interaction effects. With two factors A and B, a main effect for a factor A can be assessed by taking the means for each level of A and averaging them across the levels of B. If these marginal means for the levels of A are approximately the same, this would indicate no main effect for factor A. On the other hand, a main effect for the factor B can be examined by taking the means for each level of B and averaging them across the levels of A. If these marginal means for the levels of B are approximately the same, this would indicate no main effect for factor B. An interaction effect between factors A and B is determined by whether the difference between the means of any two levels of factor B is constant at each level of factor A (or vice versa).

This is easily viewed in a plot of the marginal means (also called a profile plot) by checking to see whether or not the lines corresponding to the levels of one of the factors are parallel. Parallel lines indicate no interaction, whereas nonparallel lines suggest that an interaction may exist. The profile plots will be discussed later.

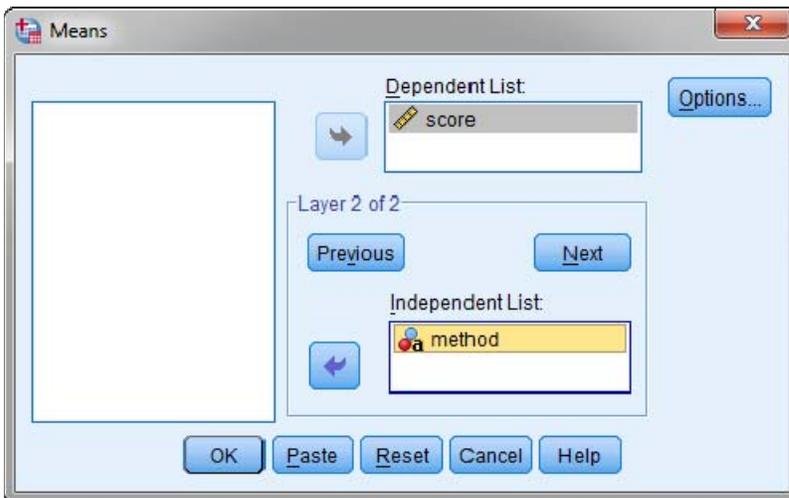
To obtain the table of the marginal means of the variable *Score* for all combinations of the two factors, click on *Analyze* from the main menu bar, then on *Compare Means* from the drop-down menu, and finally on *Means*.



This opens the *Means* dialog box. Now select and move the *score* variable into the *Dependent List* and the *ability* variable into the *Independent List* box.



Click *Next* button and select and move the *method* variable into the *Independent List* box as shown below:



Click *OK* to run the procedure. The table is displayed on the next page.

Report

score		Mean	N	Std. Deviation
ability	method			
none	blue-book	26.67	3	4.726
	notebook computer	28.00	3	3.464
	Total	27.33	6	3.777
some	blue-book	31.00	3	2.646
	notebook computer	36.67	3	3.786
	Total	33.83	6	4.262
lots	blue-book	33.33	3	2.517
	notebook computer	27.00	3	4.583
	Total	30.17	6	4.792
Total	blue-book	30.33	9	4.183
	notebook computer	30.56	9	5.747
	Total	30.44	18	4.878

The above table of means may be rearranged and presented as follows:

Method/Ability	None	Some	Lots	Average
Blue-book	26.67	31.00	33.33	30.33
Notebook computer	28.00	36.67	27.00	30.56
Average	27.33	33.83	30.17	30.44

An analysis of these means reveals that there is very little difference between the marginal means for the different levels of *method* across the levels of *ability* (30.31 vs. 30.56). As these marginal means are almost the same, we conclude that there is no main effect of *method*.

The marginal means of *ability* over levels of *method* vary in a much wider range, given the magnitude of the standard deviations. More precisely, the marginal means are 27.33, 33.83, and 30.17 with the mean for "Some" being the highest. Thus there is strong main effect of *ability*.

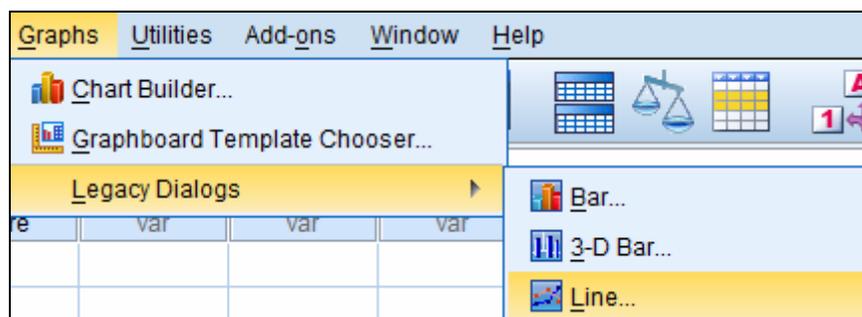
The cell means show an increasing pattern for levels of *ability* using a blue-book (26.67 vs. 31.00 vs. 33.33) and a different pattern for levels of *ability* using a computer (28.00 vs. 36.67 vs. 27.00). There is an evidence of an interaction between *ability* and *method*.

2. Plots of Marginal Means

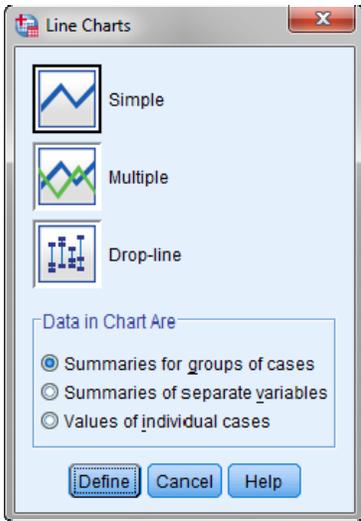
The information contained in the table of marginal means can be visualized with a plot of marginal means (profile plot). When two factors are involved, the levels of one of the factors are marked on the horizontal axis and separate lines are drawn at the height of the mean for each level of the other factor. The graph of marginal means is a component of the *Univariate GLM* default output (*Plots*).

In this section we will demonstrate how to use line charts in SPSS to examine the relationship between the response variable *score* and the two factors: *method* and *ability*. More precisely, we will obtain the plot of mean score versus typing ability with the lines representing two levels of writing method. As we noticed in the previous section, parallel lines in the plot would indicate no interaction, whereas nonparallel lines would suggest that an interaction may exist.

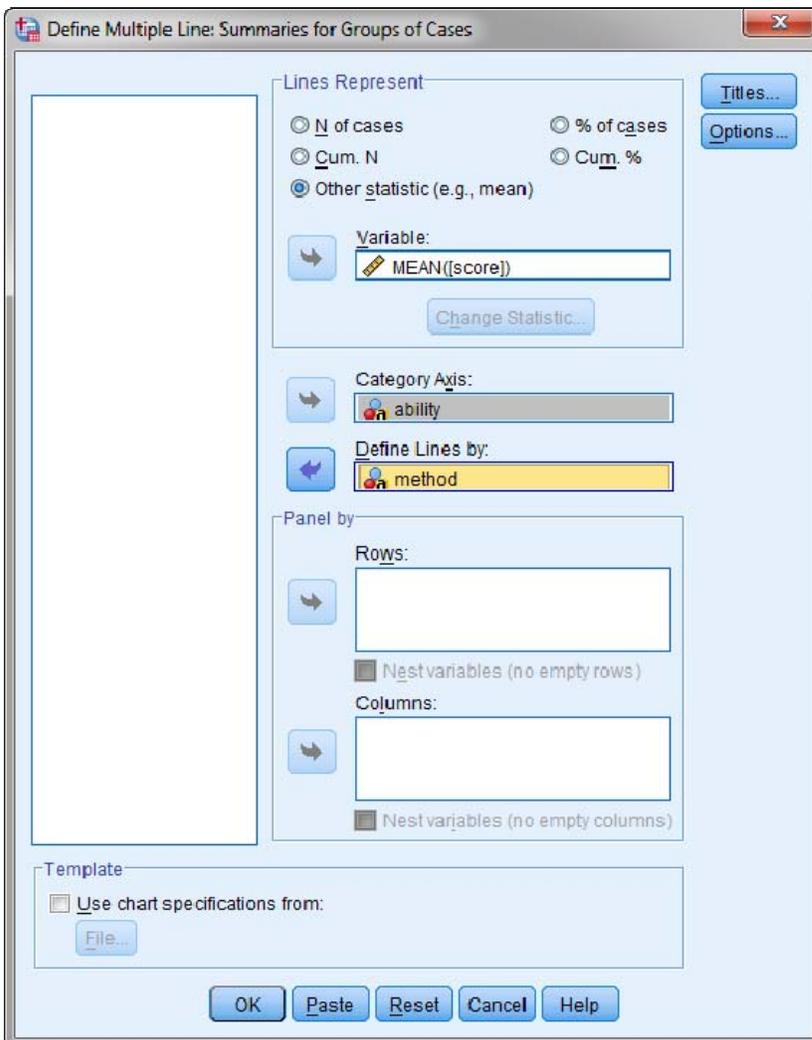
In order to obtain the line chart of the average score versus typing ability for the two methods, select *Legacy Dialogs*, then *Line* in the *Graphs* menu.



The *Line Charts* dialog box opens as follows:

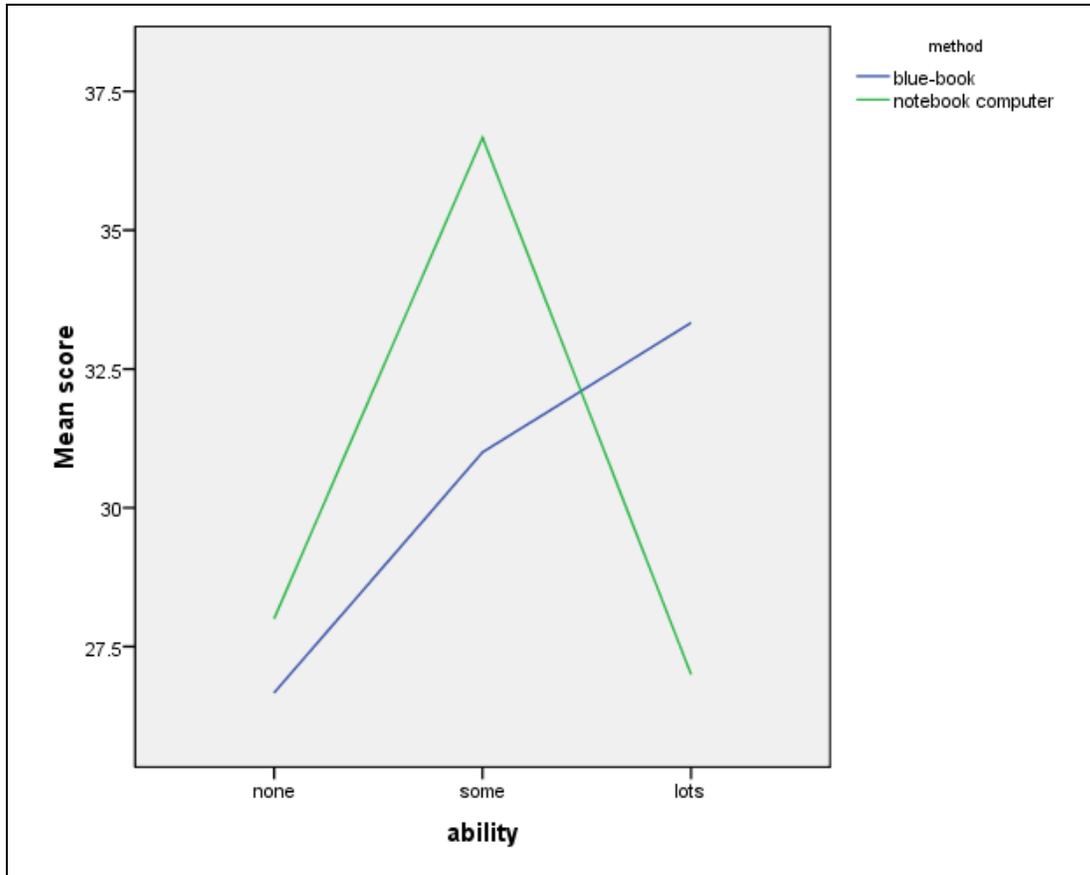


Select the *Multiple* icon in the above dialog box and make sure that *Summaries for groups of cases* radio button is selected (default option). Click *Define* button. This opens the *Define Multiple Line: Summaries for Groups of Cases* dialog box.



In the *Lines Represent* section, check the *Other statistic* radio button, and then select and move the *score* variable into the *Variable* box. Then select and move the *ability* variable into the *Category Axis* box and the *method* variable into the *Define Lines by* box. The dialog box should look like the one shown in the picture above. Finally click *OK*.

The following line chart will be displayed.



It is worthy to notice that the line chart with the factor *method* displayed on the horizontal axis and the lines defined by *ability* would be equally correct and provide the same information as the graph obtained below. Nevertheless one of them may be easier to interpret than the other one. It is recommended that both versions of the graphs be attempted and the one which is easier to interpret be included in the final statistical report.

The lines in the graph are obtained by connecting the factor-level means for the three ability levels.

The average score increases as typing ability increases for students writing the essay with blue-book. Notice in this case a substantial change in the average score between students with no typing ability and students with some typing ability and moderate increase in the average score for students with the highest typing ability.

For students using notebook computers, the average score of students with some typing ability is much higher than the average score of students with no typing ability. However, the average score of students with the highest typing ability is much lower than the average score of students with some typing ability and even lower than the average score of students with no typing ability.

For students with none or some typing ability, the average score of students using notebook computers is higher than the average scores of students using blue-book. The effects do not seem to be additive; notice the much larger difference in the average score between the two writing methods (blue-book and notebook computer) for students with some typing ability and those with none typing ability. For students with the highest typing ability, the average score of students using blue-book is much higher than the average score of students using notebook computers. This part of the graph displays the greatest degree of non-additivity and may indicate the presence of interaction.

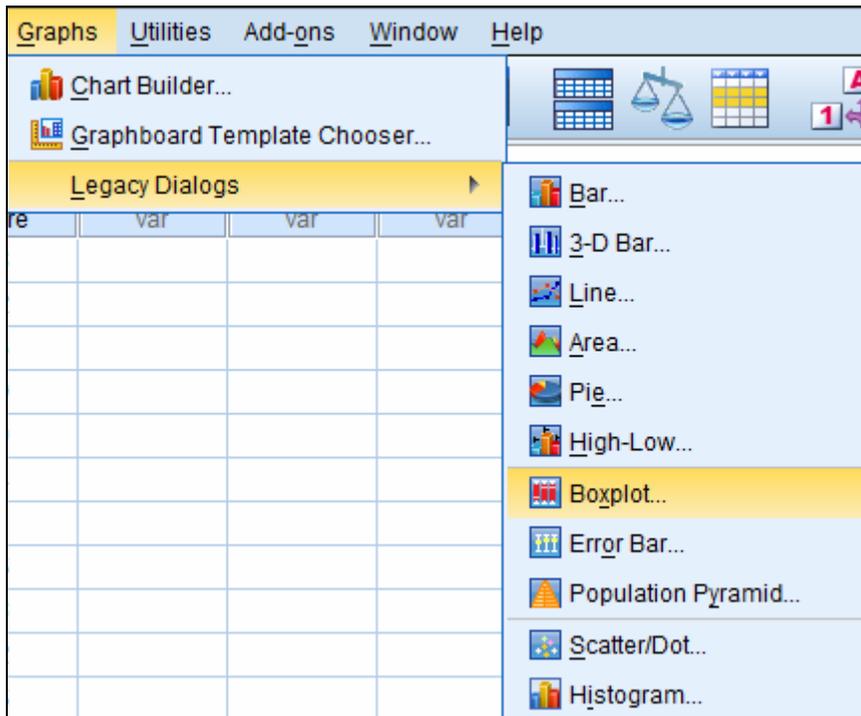
Summarizing the above graph confirms the conclusions reached in Section 1. There is strong main effect of *ability*, no main effect of *method* and an interaction may exist between *ability* and *method*.

Notice that the average scores are based on very small groups, each consisting of just three observations. Thus the observed patterns may not hold in general (may not be generalized to the whole population).

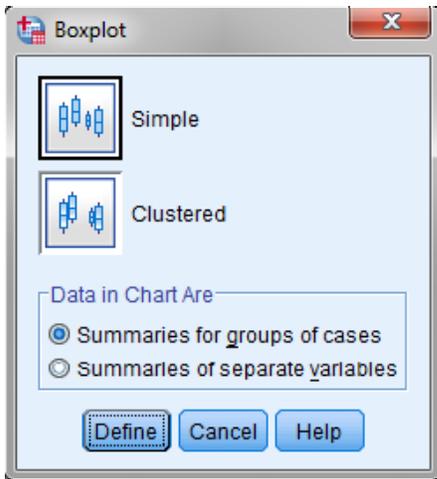
3. Clustered boxplot

The boxplots of all combinations of the two factors levels may be very useful to examine the assumptions of normality and equal variances in two-way analysis of variance. You can display boxplots for two factors together by using *Clustered Boxplots* option in SPSS.

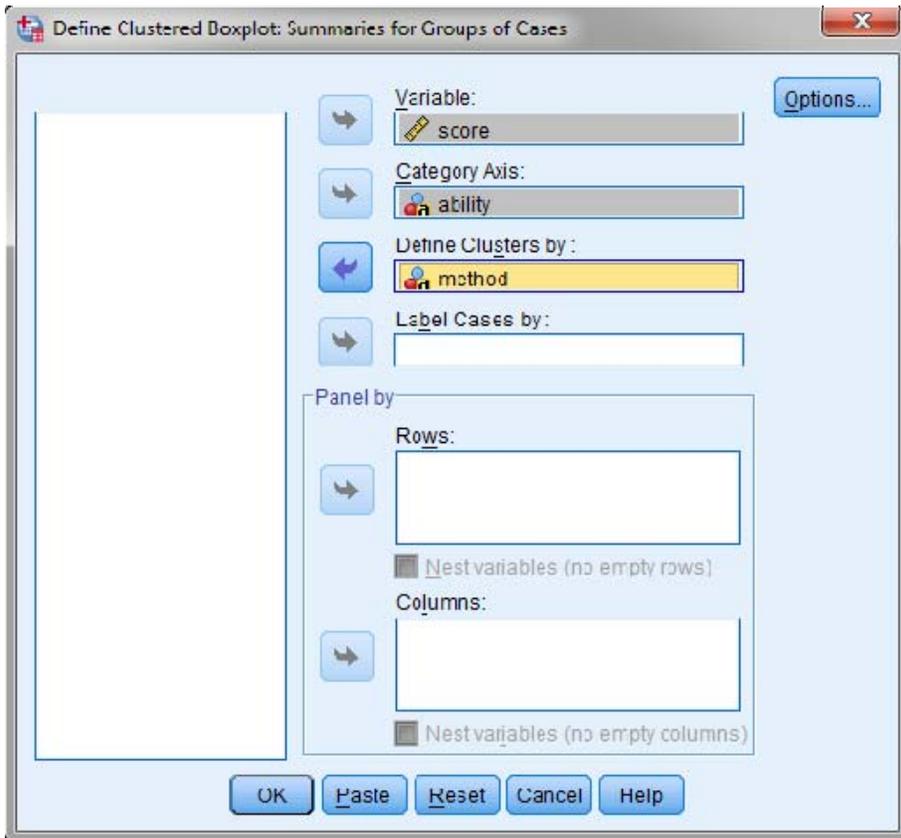
To obtain a clustered boxplot of the mean score with typing ability on the category axis and clusters defined by writing method we proceed as follows. Click the *Graphs* option in the main menu bar, and then on *Boxplot*.



This opens the *Boxplot* dialog box shown below:



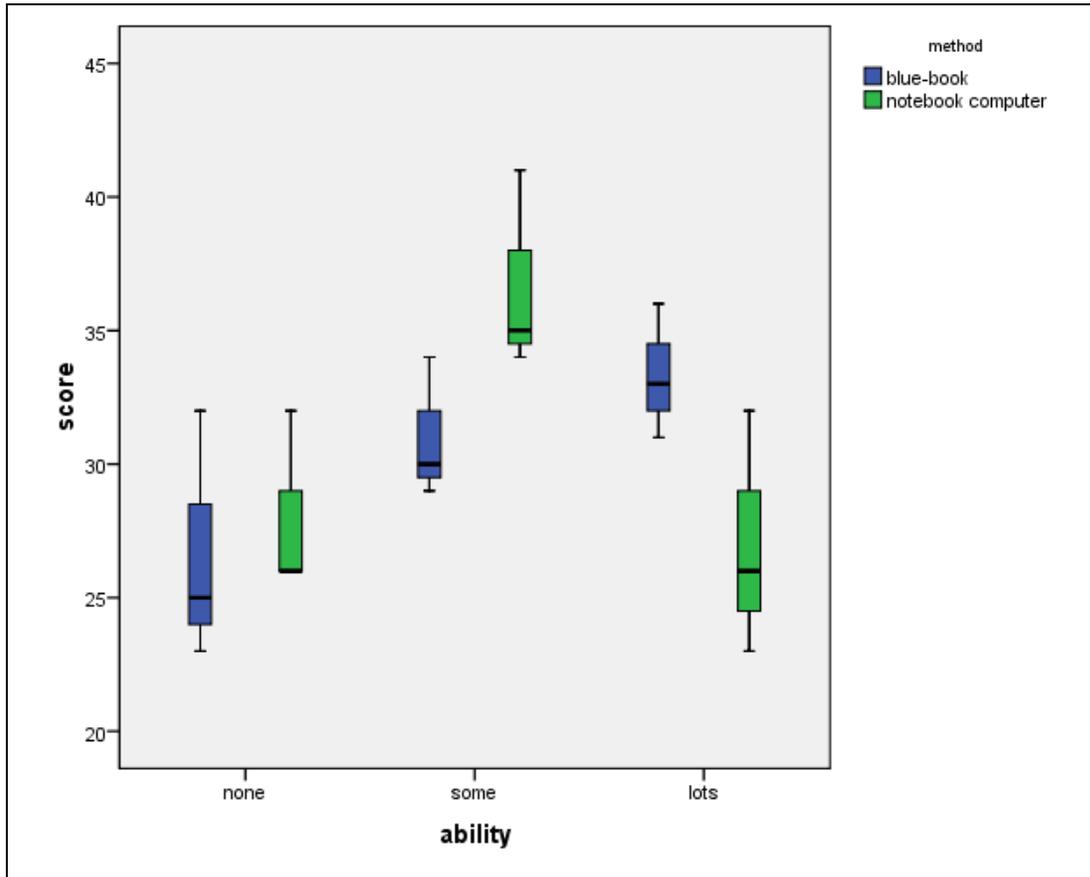
Select the *Clustered* button, and then click on *Define* button. The *Define Clustered Boxplot: Summaries for Groups of Cases* dialog box will be displayed as below.



Select and move the *score* variable into the *Variable* box, the *ability* variable to the *Category Axis* box and the *method* variable into the *Define Clusters by* box. Then click *OK*.

In this case the boxplots of scores for each of the two methods (blue-book and notebook computer) are displayed for the three levels of the variable ability marked on the category axis. In this way, we will be able to compare the distributions of scores for students using blue-book and notebook computers as the typing ability increases.

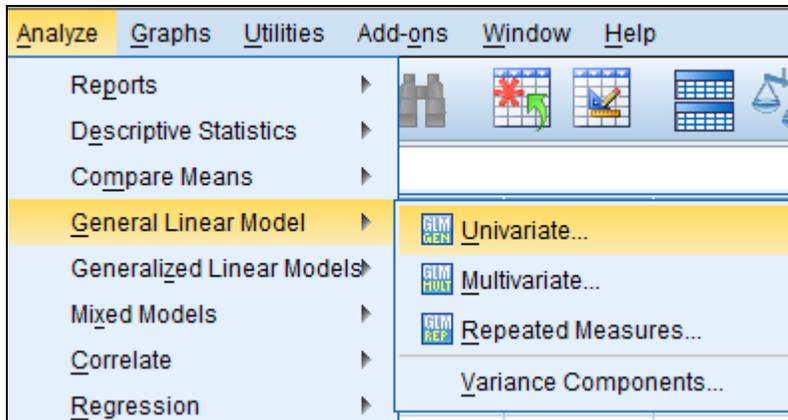
The clustered boxplot is displayed on the next page. The boxplot shows how the response variable (here, score on the essay) varies for each level of the two factors (typing ability and writing method). The clustered boxplot shows that students using notebook computer did better (as measured by the median score) than their counterparts using the bluebook for the two ability levels, no typing ability and some typing ability. However, students using blue book did much better than students using notebook computers for the high typing ability group. Notice approximately the same spread of the 6 distributions as measured by their interquartile ranges.



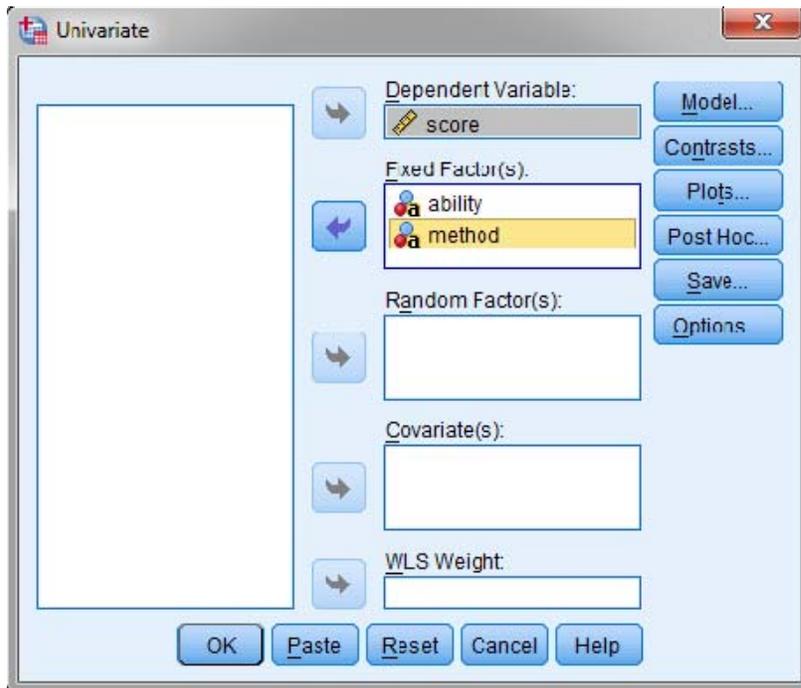
4. General Linear Model (GLM)

Univariate GLM is the general linear model now often used to implement such essential statistical procedures as regression and ANOVA. In particular, the procedure can be used to carry out two-way analysis of variance. In this particular case, GLM can be used to estimate the main and interaction effects of categorical explanatory variables (factors) on the response variable and test hypotheses about the effects. Both balanced and unbalanced models can be tested. A design is balanced if each cell in the model contains the same number of cases.

In order to apply GLM to our data, click on *Analyze* in the main menu bar, then on *General Linear Model* from the pull-down menu, and finally *Univariate*.



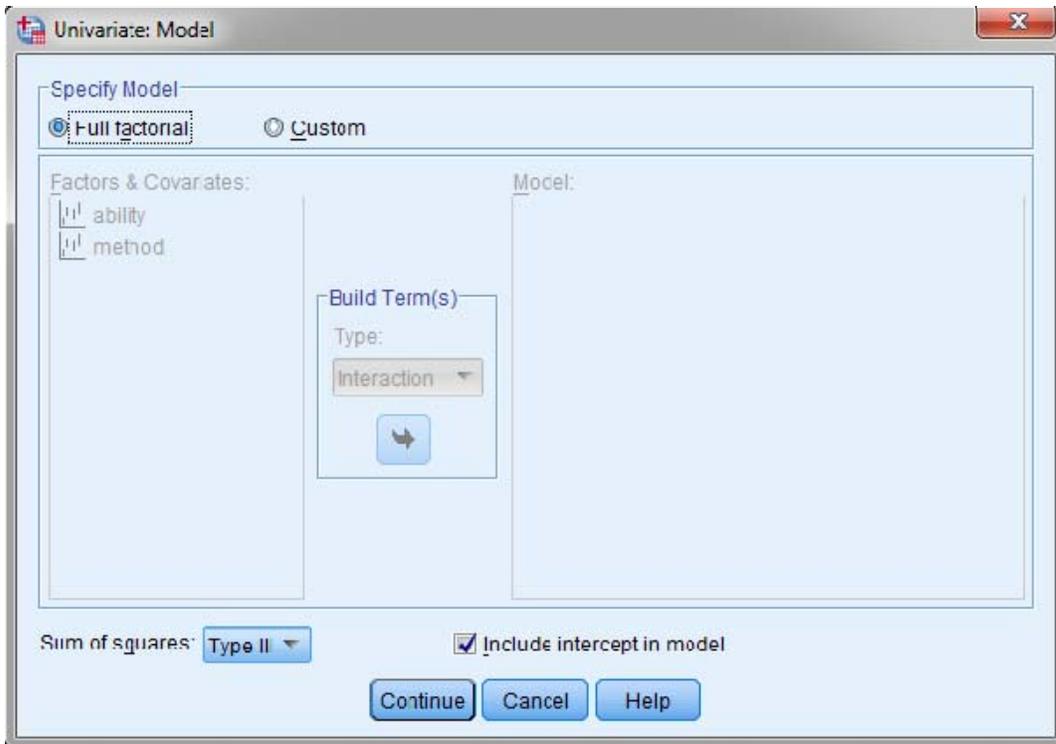
This opens the *Univariate* dialog box. Select and move the *score* variable into the *Dependent Variable* box, and both *ability* and *method* variables to the *Fixed Factor(s)* box. If the response variable also depends on another interval quantitative variable, this variable is called a covariate. You can control for the covariate by selecting and moving the variable into the *Covariate(s)* box. There are no known covariates in our example.



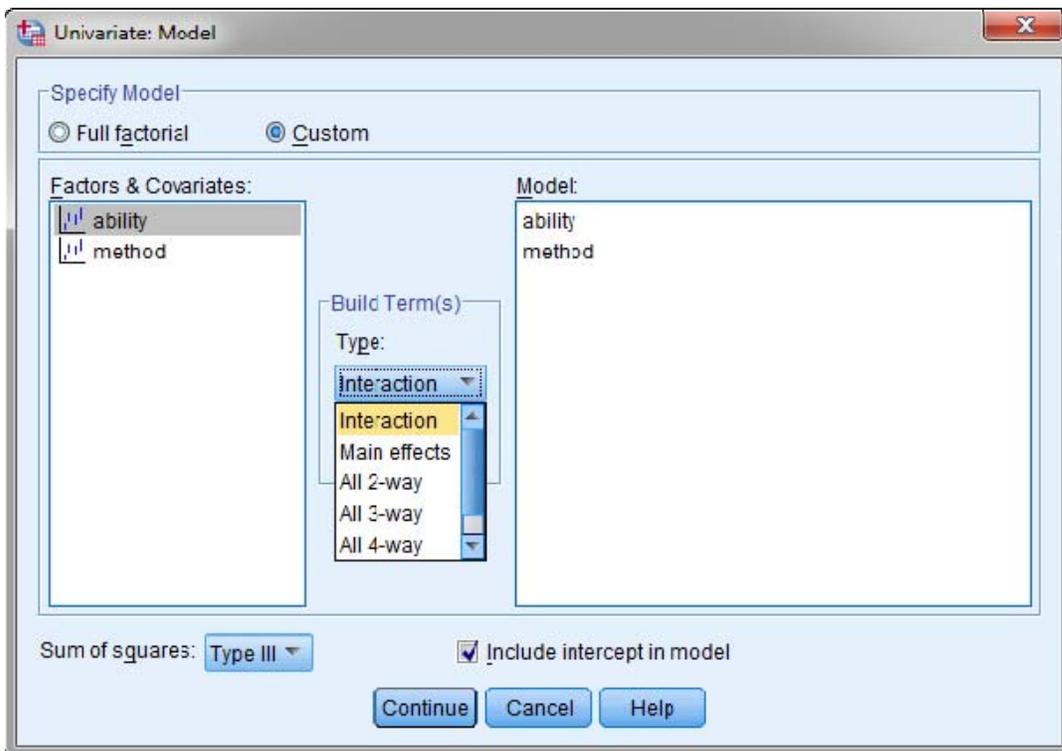
The *Univariate* dialog box contains several options to custom the analysis for your data. The options are available by clicking on the buttons in the right section of the dialog box. We briefly discuss each of the buttons below.

4.1 Model

Click on the *Model:* button in the right section of the *Univariate* dialog box. The *Univariate: Model* dialog box will be displayed. It allows you to customize the list of explanatory variables that you wish to test out. The default model is a “Full factorial” model. The full factorial model contains all factor main effects, all covariate main effects, and all factor-by-factor interactions. It does not contain covariate interactions.



If you check the *Custom* radio button, you can specify only those factors and covariates whose main effects and interactions you wish to explore. The variables of interest should be selected in the left-hand list and moved into the *Model* list in the right-hand section with the arrow in the middle of the dialog box.



Build Term(s): In the *Build Term(s)* section you can select the effects of interest from the drop-down menu. The default option is *Interaction* which creates the highest-level interaction term of all selected variables. In order to include all interaction terms of a particular order for a set of factors, select the factors together and from the *Build Term(s)* drop down list, select the desired order number. For example, *All 2-way* option creates all possible two-way interactions of the selected variables. You may find out which interactions should be examined by obtaining profile plots for the factors. Profile plots will be discussed later.

Sum of Squares: SPSS uses four methods to calculate sum of squares labelled Type I, Type II, Type III, and Type IV. You can request any of the four types of sum of squares to be used in your analysis by selecting the appropriate method in the *Sum of Squares* drop down menu. Type III is the default and it is the most commonly used option for balanced or unbalanced models with no missing cells. When missing cells are present, Type IV sum of squares method should be used. As the data in our example is balanced, we will use the default option in the analysis. The choice of sum of squares method in analysis is important; invalid conclusions might be reached by using an inappropriate sum of squares method.

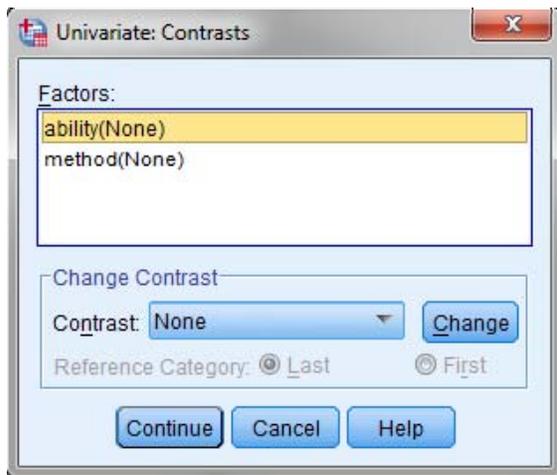
Include intercept in model: This is the default as the intercept is usually included in the model. If you can assume that the data pass through the origin, you can exclude the intercept.

Click the *Continue* button to close the *Univariate: Model* dialog box.

4.2 Contrasts

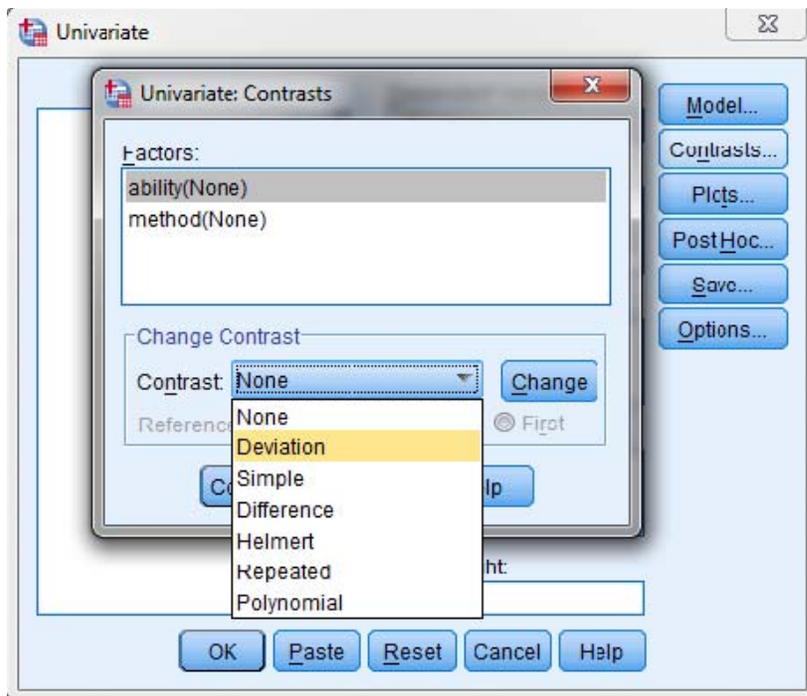
Some important questions of interest can be answered by a comparison of means among some or all of the levels of a factor. This comparison can be implemented by using the linear combinations of the level means and testing hypotheses about the combinations. The linear combination of the level means is called a *contrast*. You can specify a contrast for each factor in the model.

Click the *Contrasts* button in the *Univariate* dialog box. The *Contrasts* dialog box will be displayed.



Testing hypotheses about contrasts is based on the general null hypothesis $LB = 0$, where L is the contrast coefficients matrix and B is the parameter vector. When a contrast is specified, an L matrix is created. The columns of the L matrix corresponding to the factor match the contrast. For example, if we wish to compare the average results obtained for levels 1 and 3 with the average results for levels 2 and 4 of a factor, we can consider the contrast $C = (\mu_1 + \mu_3)/2 - (\mu_2 + \mu_4)/2$ and test the hypothesis that the contrast is zero. The corresponding column in the matrix L would be $[1/2, -1/2, 1/2, -1/2]$ and the parameter vector $B = [\mu_1, \mu_2, \mu_3, \mu_4]$.

In order to define the type of contrast for a factor, select the factor in the *Factors* list in the *Univariate: Contrasts* dialog box. Then click the arrow in the *Contrast* button and choose the appropriate contrast for the factor from the drop-down menu.



The following contrast types can be selected: *Deviation* (compares the mean of each level, except a reference category, to the mean of all of the levels, grand mean), *Simple* (compares the mean of each level to the mean of a specified level; this type of contrast is especially useful when there is a control group), *Difference* (compares the mean of each level, except the first, to the mean of previous levels), *Helmert* (compares the mean of each level of the factor, except the last, to the mean of subsequent levels), *Repeated* (compares the mean of each level, except the last, to the mean of the subsequent level), and *Polynomial* (compares the linear effect, quadratic effect, cubic effect, and so on).

The output includes an F statistic for each set of contrasts. Also displayed for the contrast differences are Bonferroni-type simultaneous confidence intervals based on Student's t distribution.

Click the *Continue* button to close the *Univariate: Contrasts* dialog box.

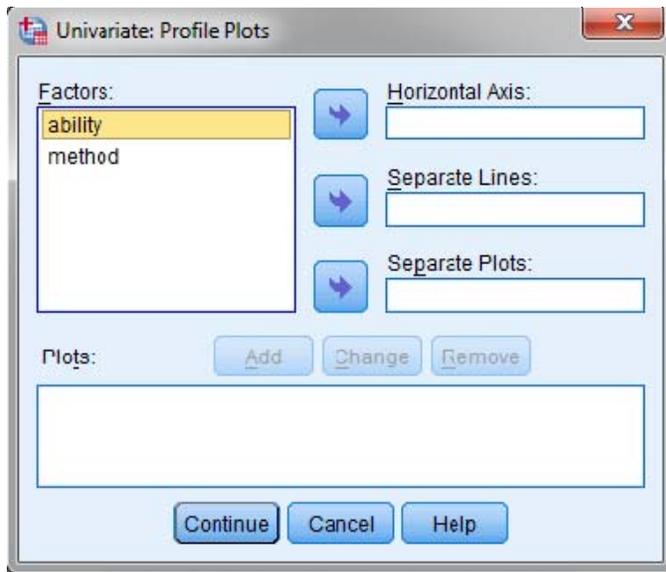
4.3 Profile Plots

Plots of marginal means also called profile plots were already discussed in Section 2 as they constitute a very important component of the exploratory stage of data analysis. A plot for two factors A and B is obtained by placing the levels of factor A (ability) on the horizontal axis, and the cell means of the response variable (score) on the vertical axis. The lines in the body of the plot represent the levels of factor B (method). As we noticed earlier the specific placement of the two factors in the plot is arbitrary; alternatively factor B could be plotted on the horizontal axis and factor A as the lines).

A profile plots provides information about the possible existence of a main effect for A, a main effect for B, and/or interaction effect. It shows whether the estimated marginal means are increasing or decreasing across levels. For two or more factors, parallel lines indicate that there is no interaction between factors, which means that you can examine the levels of only one factor. Lines that cross each other indicate an

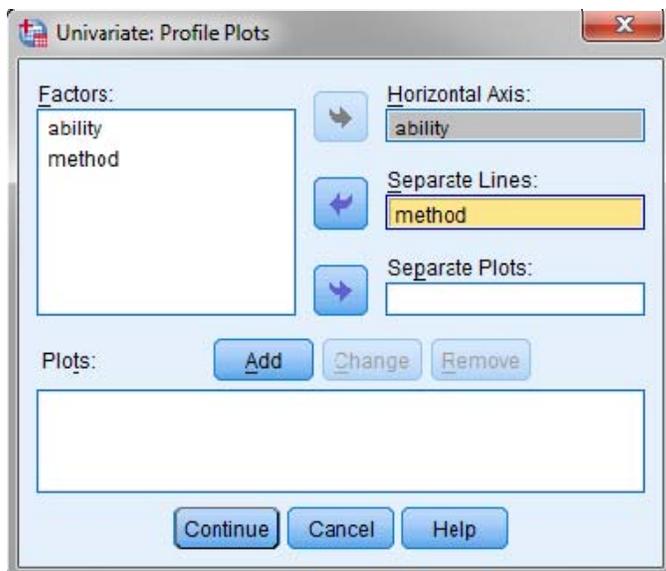
interaction. Remember that statistical significance of the main and interaction effects can be only determined by the F test. Nevertheless, profile plots can provide you with a rough idea as to the possible existence of the main and interaction effects. In particular, lines in the profile plot that are nearly parallel may not necessarily indicate lack of interaction.

Click the *Plots...* button in the *Univariate* dialog box. The *Plots* dialog box will be displayed.



To plot the means for the values of a single factor, move that factor into the *Horizontal Axis* box and click *Add*. If you want to display the means for all combinations of values of the two factors, move the second factor into the *Separate Lines* box and click *Add*. The horizontal axis factor means will be plotted separately for each value of the second factor.

In order to display the interaction between *ability* and *method*, select the variable *ability* and move it into the *Horizontal Axis* box and the variable *method* into the *Separate Lines* box. Click *Add*



Click the *Continue* button.

4.4 Post Hoc

Once you have determined that differences exist among the means, post hoc range tests and pairwise multiple comparisons can determine which means differ. With the button *Post Hoc* in the *Univariate* dialog box you can perform several tests for this purpose. These tests are called multiple comparison tests.

The dialog box of *Univariate: Post Hoc Multiple Comparisons for Observed Means* is displayed on the next page. The dialog box consists of four sections: (i) the *Factors and Post Hoc Tests for* section allows you to specify the factors to be tested, (ii) multiple comparison tests that assume equal variances and (iii) multiple comparison tests that do not make this assumption.

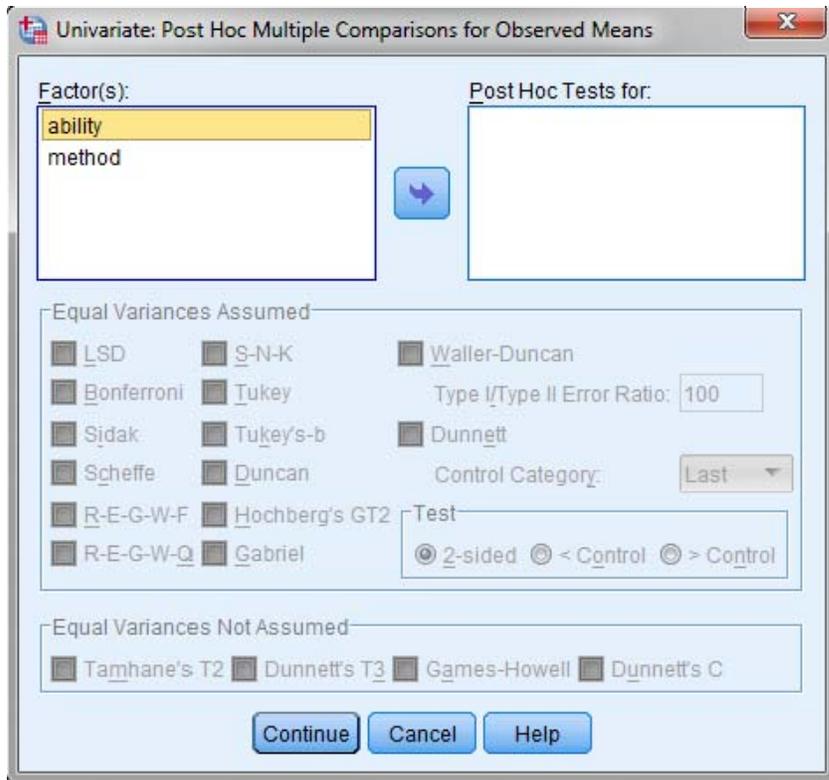
In order to carry out one of the selected tests, select the factors to be tested in the *Factors* list (for example, *ability* and *method*) and move them into the *Post Hoc Tests for* list.

Equal variances tests: SPSS offers a large variety of multiple comparison procedures that you can use, assuming equal variances. We discuss the four most popular choices:

- **LSD:** Least Significant Difference procedure performs individual t-tests between all pairs of groups. However, the disadvantage of the test is that the observed level of significance is not adjusted for the number of multiple comparisons made. Indeed, as the number of groups becomes larger, the number of comparisons rapidly increases, which increases the chance of falsely finding significant differences between groups. It does not require equal sample sizes. LSD is the most liberal of the post-hoc tests (it is most likely to reject the null hypothesis in favor of finding that groups do differ).
- **Bonferroni:** The test ensures that the probability of false rejection of the null hypothesis of equal means for any two groups among all comparisons does not exceed 0.05. The LSD method uses a probability of 0.05 for each individual comparison of two groups. The Bonferroni test controls by dividing the significance level by the number of comparisons. The Bonferroni test is very conservative (likely to accept the null hypothesis of no group differences) when a large number of group means are being compared.
- **Tukey:** Tukey's honestly significant difference (HSD) test uses the Studentized range statistic to make all pairwise comparisons between groups. The Tukey test can only be used when the sample sizes are the same. When testing a large number of pairs of means (pairwise comparisons), The Tukey method is preferred as it is a very conservative test (large number of groups tends to inflate Type I errors so the conservative approach is justified in this case). Tukey's honestly significant difference test is more powerful than the Bonferroni test. For a small number of pairs, Bonferroni is more powerful.

Note that *Tukey's-b* also listed in the *Univariate: Post Hoc* dialog box is a less conservative version of Tukey's HSD test.

- **Scheffe:** Scheffé's method applies to all possible contrasts among the factor level means, not just the pairwise differences considered by Tukey's method. This is the most conservative of all multiple comparison procedures which means that this test requires the difference between two means be large before the difference is considered significant. The Scheffé's test is very effective in controlling Type I error in post hoc comparisons but it does so at the cost of a loss in statistical power (more Type II errors may be made – making decision that you do not have a relationship when in fact you do). It is not preferred for particular comparisons, but it is recommended for a large number of comparisons. If only pairwise comparisons are to be made, the Tukey method will result in narrower confidence intervals.

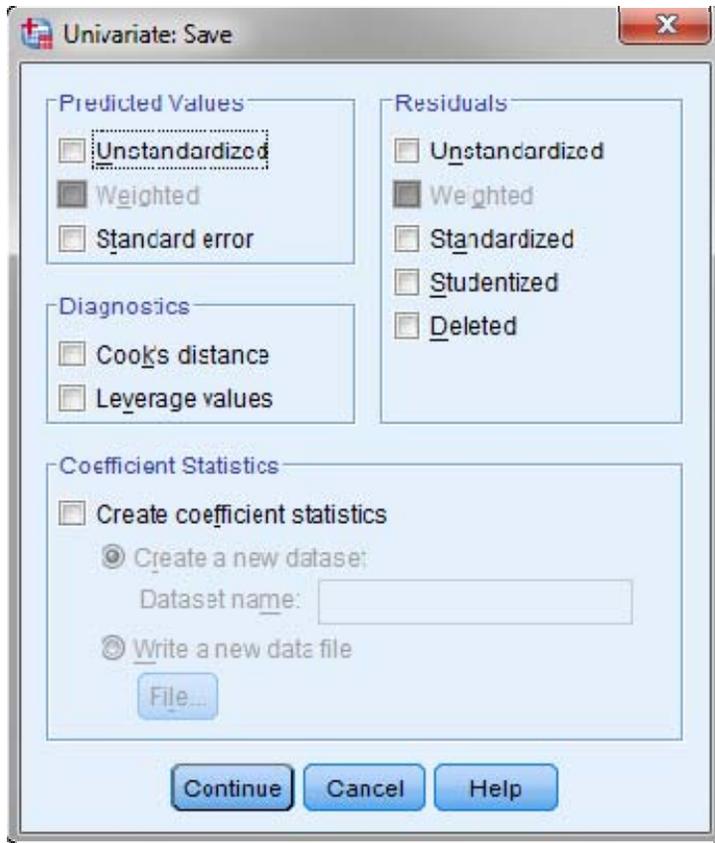


Equal variances not assumed tests: You can choose among many multiple comparison tests: Tamhane's T2, Dunnett's T3, Games Howell and Dunnett's C. For information about a test, click Help button.

The Post Hoc procedure in SPSS also outputs a **homogenous subsets table**. The table summarizes the post hoc tests in a more easily interpretable format. In each subset column, the factor levels are displayed which are not significantly different from each other. The Homogenous Post Hoc tests table is not available for all post hoc tests.

4.5 Save

The values predicted by the model, various residuals, and a number of diagnostic statistics can be saved as variables in the *Data Editor* window in the Univariate GLM procedure. To do this, click *Save* in the *Univariate* dialog box and then specify which variables you want to save. You can specify that the variables have to be included in a new or opened Data Editor window ("create a new dataset") or saved in a new file on disk ("write a new data file"). Some of these variables can be used for examining assumptions about the data.



4.6 Options

The *Univariate: Options* dialog box allows you to obtain the estimates of the population marginal means in the cells (*Estimated Marginal Means* pane) and obtain several very useful statistics, plots and tests (*Display* pane).

First you have to select the factors and interactions for which you want estimates of the population marginal means in the cells. Then you can select useful statistics, plots and tests. Among others:

Descriptive statistics: to obtain observed means, standard deviations, and counts for all cells.

Homogeneity tests: To calculate the Levene statistic (discussed in Lab 1 Instructions) for testing equality of variance for all cells.

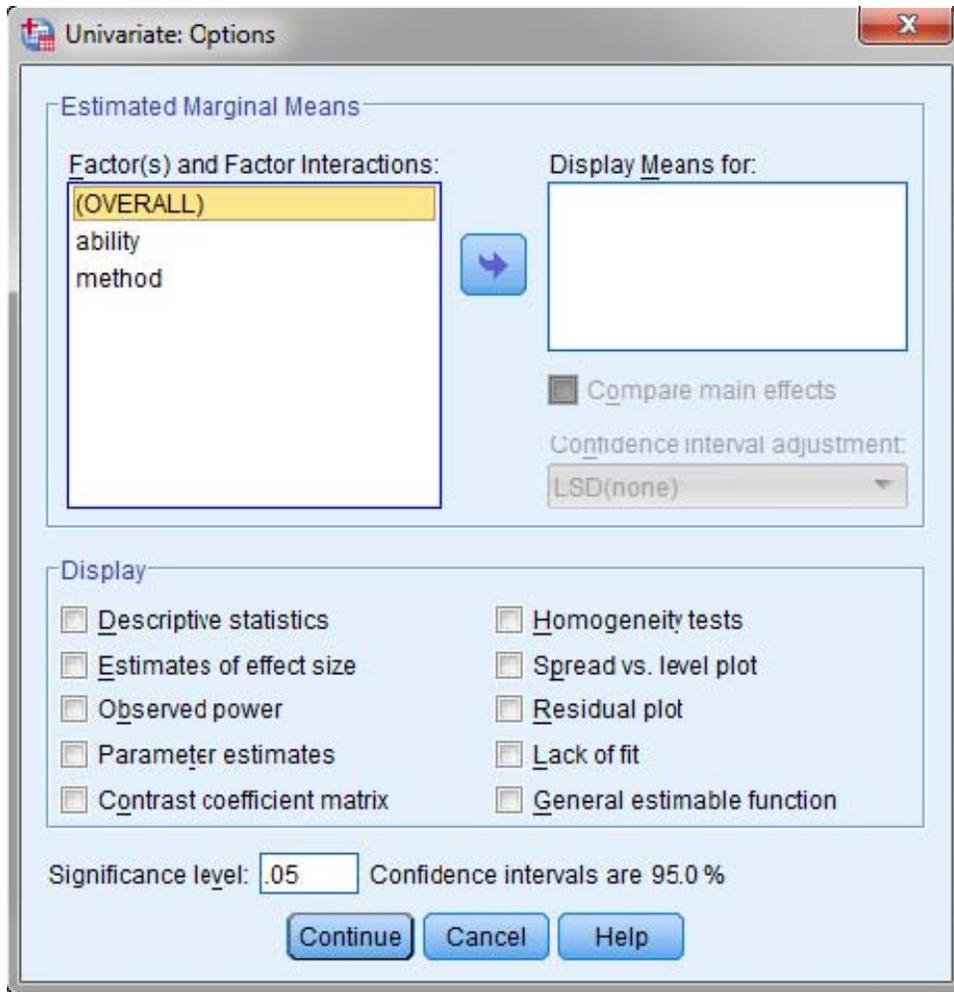
Spread vs. level plots: To obtain plots of cell variances and cell standard deviations against the cell means.

Residual plot: To obtain plots of observed values, predicted values, and residuals.

The spread-versus-level and residual plots options are useful for checking assumptions about the data.

Observed power: To obtain the power of the test when the alternative hypothesis is set based on the observed value.

Lack of fit: To check if the relationship between the response variable and the explanatory variables can be adequately described by the model.



5. GLM for Example Data

In this section we will apply GLM Univariate procedure to our data in Section 1. At the exploratory stage we found strong main effect of ability, no main effect of method and an interaction between ability and method. In order to find whether the observed effects are significant, GLM Univariate procedure will be applied to the data.

The analysis is done by selecting Analyze in the main menu, General Linear Model, and Univariate in the drop-down menu. The *Full Factorial* model will be specified, and Tukey and Scheffe's post hoc tests will be requested. The only *Options* that will be selected in this example is the *Descriptive Statistics* option under Display.

We will provide the GLM output, explain the entries in the output and also interpret the output. The output of GLM procedure for our data is displayed below.

Levene's Test of Equality of Error Variances ^a			
Dependent Variable: score			
F	df1	df2	Sig.
.672	5	12	.652

Tests the null hypothesis that the error variance of the dependent variable is equal across groups.

a. Design: Intercept + ability + method + ability * method

The Levene test with the p-value of 0.652 indicates that there is no evidence that the assumption of equal standard deviations is violated. Notice that also the values of the standard deviations for the groups obtained in Section 1 did not indicate any considerable differences.

Analysis of variance allows us to test the null hypothesis that there is no association between the mean final exam score and typing ability and writing method. There are three null hypotheses that have been tested in the table given below:

1. H_{01} : No association between mean score and typing ability.
2. H_{02} : No association between mean score and writing method.
3. H_{03} : No interaction between typing ability and writing method.

Tests of Between-Subjects Effects					
Dependent Variable: score					
Source	Type III Sum of Squares	df	Mean Square	F	Sig.
Corrected Model	238.444 ^a	5	47.689	3.447	.037
Intercept	16683.556	1	16683.556	1206.040	.000
ability	127.444	2	63.722	4.606	.033
method	.222	1	.222	.016	.901
ability * method	110.778	2	55.389	4.004	.047
Error	166.000	12	13.833		
Total	17088.000	18			
Corrected Total	404.444	17			

a. R Squared = .590 (Adjusted R Squared = .419)

P-values for H_{01} , H_{02} and H_{03} , respectively

In the first column of the above output you find sums of squares (based on Type III Sum of Squares). The total sum of squares (Corrected Total, 404.444) consists of two components, namely the part accounted for (Corrected Model, 238.444) and the part not accounted for (Error, 166.00). The sum of squares accounted for consists of three components: the two factors (ability and method) and the interaction term. The interaction term refers to the joint effect of the two factors on the response (score) and is indicated in the table as ability*method. The displayed sum of squares accounted for (Corrected Model) 238.444 is equal the sum of the three components, 127.444+0.222+110.778. In general, the sum for Corrected Model is always equal to the sum of the three components if the number of cases is the same for all cells. Otherwise, SPSS adjusts the sum on the basis of the number of cases.

The second column shows the degrees of freedom (DF) for the various effects. They are computed as follows:

- DF for each factor (ability, method): the number of groups minus 1, $DF(\text{ability})=3-1=2$,
 $DF(\text{method})=2-1=1$.
 DF for the interaction term: $DF(\text{ability}) * DF(\text{method})=2 * 1=2$

DF for the full model (Corrected Model): The sum of degrees of freedom for the two factors and the interaction term: $2+1+2=5$,

DF for the Error term: the number of cases minus 1 minus the degrees of freedom for the full model:
 $(18-1)-5=12$,

DF for the total sum of squares (Corrected Total): the number of cases minus 1 = $18-1=17$.

The next column contains the mean sum of squares. For each category this is the sum of squares divided by the corresponding degrees of freedom. The column “F” contains the value of the f statistic for each effect. The value of F is computed as follows:

$$F\text{-value} = \frac{\text{Mean Sum of Squares for the effect}}{\text{Mean Sum of Squares for the residuals}}$$

The F statistic follows an F distribution with the number of degrees of freedom for the numerator equal to the number of degrees of freedom for the effect and the number of degrees of freedom for the denominator equal to the number of degrees of freedom for the residuals (Error).

For each F value and the corresponding degrees of freedom, SPSS computes the p-value. They are given in the last column. If we use the significance level $\alpha=0.05$, we conclude that both ability and its interaction with method are significant. The effect of method is not significant. The conclusions are very consistent with the results of the preliminary analysis in Sections 1 and 2.

The Tukey HSD and Scheffe post hoc procedures produce the same significant results (significant difference between “none” and “some” ability).

Post Hoc Tests							
ability							
Multiple Comparisons							
Dependent Variable: score							
	(I) ability	(J) ability	Mean Difference (I-J)	Std. Error	Sig.	95% Confidence Interval	
						Lower Bound	Upper Bound
Tukey HSD	none	some	-6.50*	2.147	.026	-12.23	-.77
		lots	-2.83	2.147	.412	-8.56	2.90
	some	none	6.50*	2.147	.026	.77	12.23
		lots	3.67	2.147	.242	-2.06	9.40
	lots	none	2.83	2.147	.412	-2.90	8.56
		some	-3.67	2.147	.242	-9.40	2.06
Scheffe	none	some	-6.50*	2.147	.033	-12.49	-.51
		lots	-2.83	2.147	.444	-8.82	3.15
	some	none	6.50*	2.147	.033	.51	12.49
		lots	3.67	2.147	.271	-2.32	9.65
	lots	none	2.83	2.147	.444	-3.15	8.82
		some	-3.67	2.147	.271	-9.65	2.32

Based on observed means.
 The error term is Mean Square(Error) = 13.833.
 *. The mean difference is significant at the .05 level.

Students are strongly encouraged to carry out additional analysis for the data, for example specify some useful contrasts and examine the residual plots. Notice that it is very difficult to verify distributional assumptions in this case given the small sample sizes in the groups.

Henryk Kolacz
University of Alberta
September 2012