The Overall Two-Factor Analysis

ERSH 8310

Today's Class

- Two-way ANOVA
  - How to compute values.
  - How to test main effects statistically.
  - How to test interactions statistically.
- Doing all of this in SPSS.

The Overall Two-Factor Analysis

- In Chapter 2 the total sum of squares was partitioned into the between-group sum of squares (SSbetween) and the within-group sum of squares (SSwithin).

- In the analysis of the factorial design the SSbetween is to be divided into:
  - the sum of squares reflecting the main effect of factor A (SS_A),
  - the sum of squares reflecting the main effect of factor B (SS_B),
  - the sum of squares reflecting the A ×B interaction effect ($SS_{A \times B}$).

- The same number of subjects in each of the treatment conditions will be considered in the current chapter.

Component Deviations

- The notational system is summarized in Table 11.1.
- The basic observation or score in the two-way factorial design is denoted as Y_{ijk}.
- Where i = 1,…,n, j = 1,…,a, and k = 1,…,b.

Component Deviations

- The book breaks items down into component deviations for hand calculations.

- These are useful to understand where the sums of squares come from in the two-factor ANOVA.

$$AB_{jk} = \sum_{i=1}^{n} Y_{ijk}, \qquad (1)$$

$$A_j = \sum_{i=1}^{n} \sum_{k=1}^{b} Y_{ijk}, \qquad (2)$$

$$B_k = \sum_{i=1}^{n} \sum_{j=1}^{a} Y_{ijk}, \qquad (3)$$

$$T = \sum_{i=1}^{n} \sum_{j=1}^{a} \sum_{k=1}^{b} Y_{ijk}. \qquad (4)$$

The means are $[\ \bar{Y}]_{jk} = AB_{jk} / n$, $[\ \bar{Y}]_{A_j} = A_j / bn$, $[\ \bar{Y}]_{B_k} = B_k / an$, and $[\ \bar{Y}]_T = T / abn$.

Partitioning the Deviations

- In Chapter 2, it was shown that:

$$SS_T = SS_{between} + SS_{within}$$

- In the two-way factorial design,

$$SS_{between} = SS_A + SS_B + SS_{A\times B}$$

because

$$\bar{Y}_{jk} - \bar{Y}_T = (\bar{Y}_{A_j} - \bar{Y}_T) + (\bar{Y}_{B_k} - \bar{Y}_T) + (\bar{Y}_{jk} - \bar{Y}_{A_j} - \bar{Y}_{B_k} + \bar{Y}_T).$$

Note that $$Y_{ijk} - \bar{Y}_T = (\bar{Y}_{jk} - \bar{Y}_T) + (Y_{ijk} - \bar{Y}_{jk}).$$

Computations in the Two-Way Analysis

- Identifying the Source of Variance
- The tentative rule has two steps:
  - List all factors, including the within-group factor.
  - Form all possible interactions with these factors, omitting the within-group factor (i.e., obtain interaction).
- For the two factor design, step 1 results in A, B, and S/AB; and step 2 results in A ×B.

Degrees of Freedom

- The degrees of freedom are:
  - $df_A = a - 1$
  - $df_B = b - 1$
  - $df_{A \times B} = (a - 1)(b - 1)$
  - $df_{S/AB} = ab(n - 1)$
  - $df_T = abn - 1$
- Computational Formulas
  - See Table 11.5 for the computational formulas.

Forming the Bracket Terms

- Whenever you square some total for the numerator, you will divide the number of scores that went into the total.

- Note that (see Table 11.5)

- Formulae on the next slide…

$$[A] = \frac{\sum_{j=1}^{a} A_j^2}{bn}, \qquad (8)$$

$$[B] = \frac{\sum_{k=1}^{b} B_k^2}{an}, \qquad (9)$$

$$[AB] = \frac{\sum_{j=1}^{a} \sum_{k=1}^{b} (AB_{jk})^2}{n}, \qquad (10)$$

$$[Y] = \sum_{i=1}^{n} \sum_{j=1}^{a} \sum_{k=1}^{b} Y_{ijk}^2, \qquad (11)$$

$$[T] = \frac{T^2}{abn}. \qquad (12)$$

Completing the Analysis

- Note that:

$$SS_T = SS_A + SS_B + SS_{A \times B} + SS_{S/AB}$$

- And that:

$$df_T = df_A + df_B + df_{A \times B} + df_{S/AB}.$$

- Hence:

$$abn - 1 = (a - 1) + (b - 1) + (a - 1)(b - 1) + ab(n-1).$$

More About Degrees of Freedom

- The number of degrees of freedom for a sum of squares equals the number of different observations used to calculate it minus the number of constraints imposed on those observations.

Mean Squares

- The mean squares are:
  - $MS_A = SS_A / df_A$
  - $MS_B = SS_B / df_B$
  - $MS_{A \times B} = SS_{A \times B} / df_{A \times B}$
  - $MS_{S/AB} = SS_{S/AB} / df_{S/AB}$

- It can be noted that SS_A, SS_B and $SS_{A \times B}$ are mutually orthogonal (Keppel, 1982).

F Ratios

$$F_A = \frac{MS_A}{MS_{S/AB}},$$

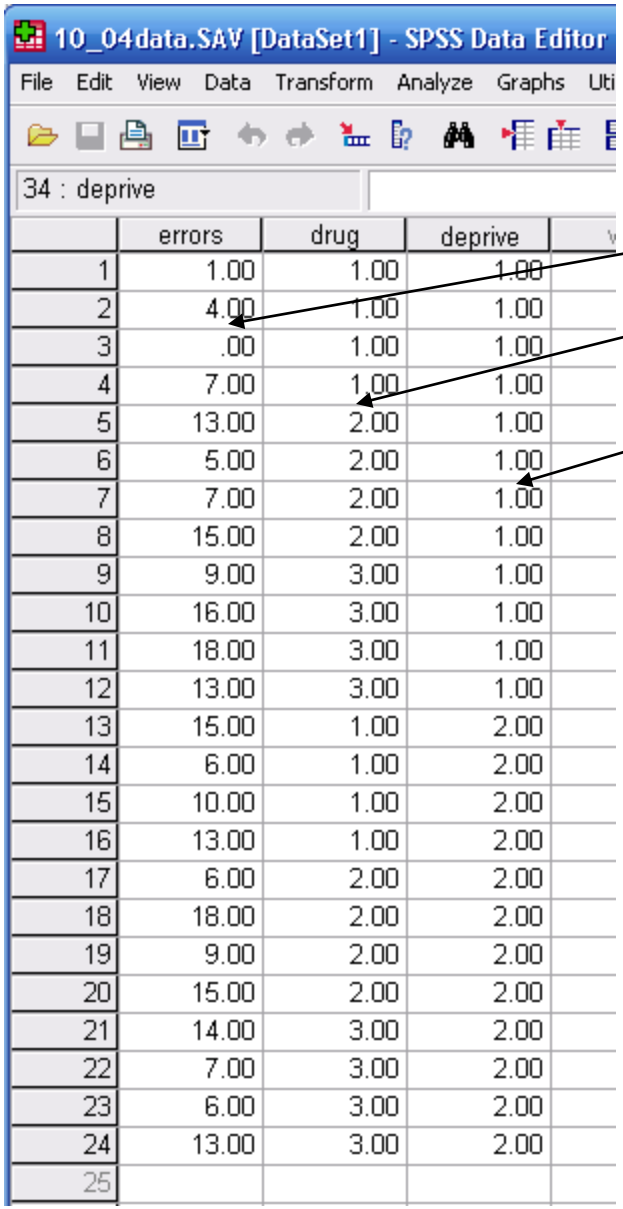
$$F_B = \frac{MS_B}{MS_{S/AB}},$$

$$F_{A \times B} = \frac{MS_{A \times B}}{MS_{S/AB}}.$$

A Numerical Example

- The example consists of a hypothetical investigation of the role of certain drugs [factor A-Control (a1), Drug X (a2), Drug Y (a3)] and drive level [factor B-1 hour of food deprivation (b1), 24 hour of food deprivation (b2)] on learning performance (Y) of monkeys.
- The animals are given a series of 20 ``oddity'' problems and the response measure Y is the number of errors in the 20 training trials.
- The design is a 3×2 factorial with a cell sample size of n = 4.

Data…



	errors	drug	deprive
1	1.00	1.00	1.00
2	4.00	1.00	1.00
3	.00	1.00	1.00
4	7.00	1.00	1.00
5	13.00	2.00	1.00
6	5.00	2.00	1.00
7	7.00	2.00	1.00
8	15.00	2.00	1.00
9	9.00	3.00	1.00
10	16.00	3.00	1.00
11	18.00	3.00	1.00
12	13.00	3.00	1.00
13	15.00	1.00	2.00
14	6.00	1.00	2.00
15	10.00	1.00	2.00
16	13.00	1.00	2.00
17	6.00	2.00	2.00
18	18.00	2.00	2.00
19	9.00	2.00	2.00
20	15.00	2.00	2.00
21	14.00	3.00	2.00
22	7.00	3.00	2.00
23	6.00	3.00	2.00
24	13.00	3.00	2.00
25			

Y has a column

Factor A has a column

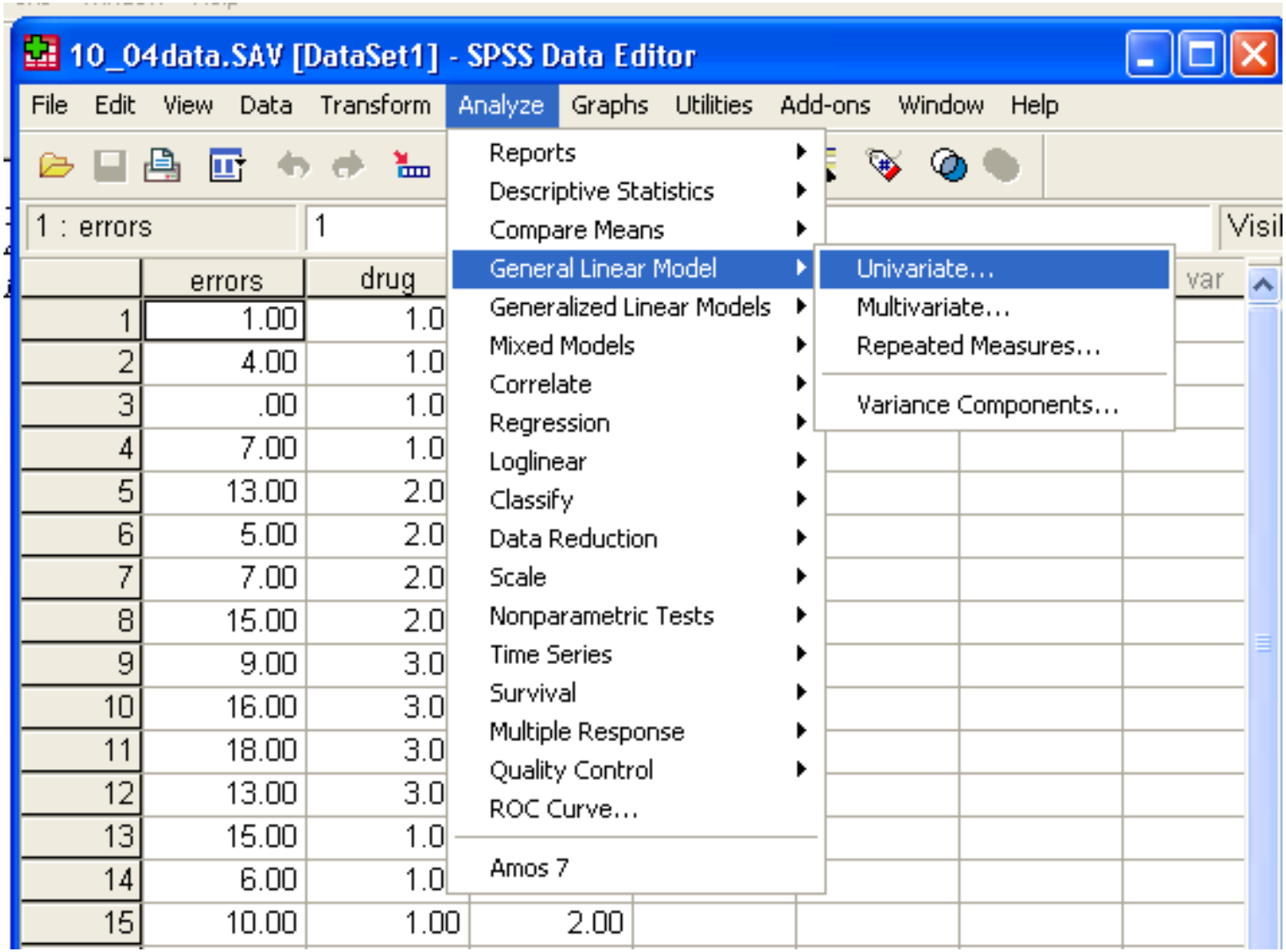
Factor B has a column

Multifactor ANOVA in SPSS

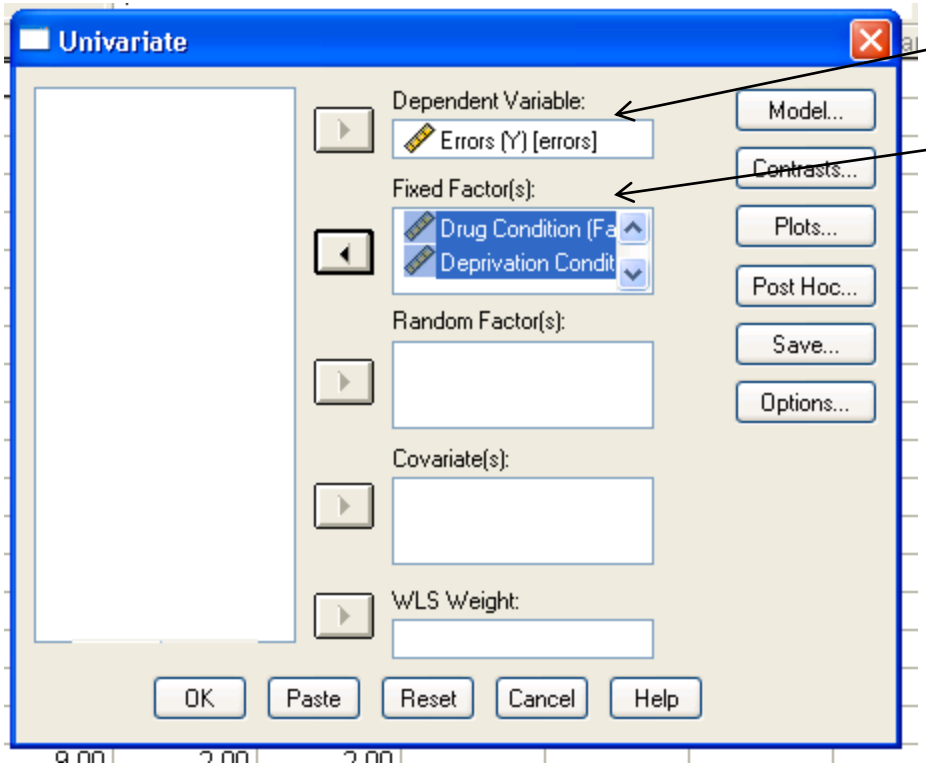
- As you may recall, up to this point we have only run analyses using the Analyze…Compare Means…One-Way ANOVA option.

- For a more general approach (with multiple factors), we must now use a slightly different option.
  - Analyze…General Linear Model…Univariate

GLM in SPSS


10_04data.SAV [DataSet1] - SPSS Data Editor
File Edit View Data Transform Analyze Graphs Utilities Add-ons Window Help
1 : errors
1
Reports
Descriptive Statistics
Compare Means
General Linear Model
Generalized Linear Models
Mixed Models
Correlate
Regression
Loglinear
Classify
Data Reduction
Scale
Nonparametric Tests
Time Series
Survival
Multiple Response
Quality Control
ROC Curve...
Amos 7
Univariate...
Multivariate...
Repeated Measures...
Variance Components...
errors drug
1 1.00 1.0
2 4.00 1.0
3 .00 1.0
4 7.00 1.0
5 13.00 2.0
6 5.00 2.0
7 7.00 2.0
8 15.00 2.0
9 9.00 3.0
10 16.00 3.0
11 18.00 3.0
12 13.00 3.0
13 15.00 1.0
14 6.00 1.0
15 10.00 1.00 2.00

Filling the GLM Boxes



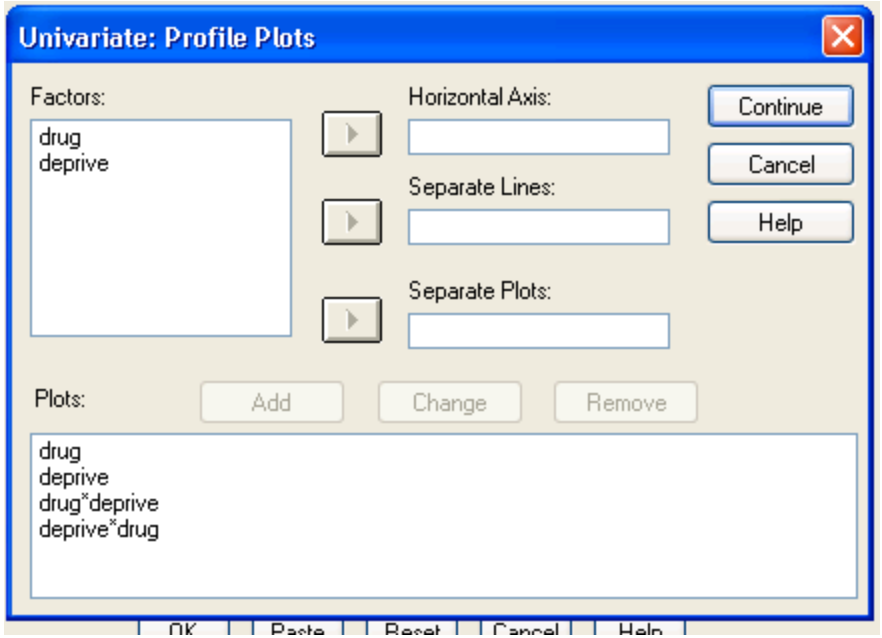
- Dependent Variable – Y
- Fixed Factors – Each Factor Variable
- The rest…we'll find out later…

Running SPSS

- The default state of SPSS is to run the full factorial analysis.

- However, we can get SPSS to create our mean plots for us (helpful in detecting interactions).

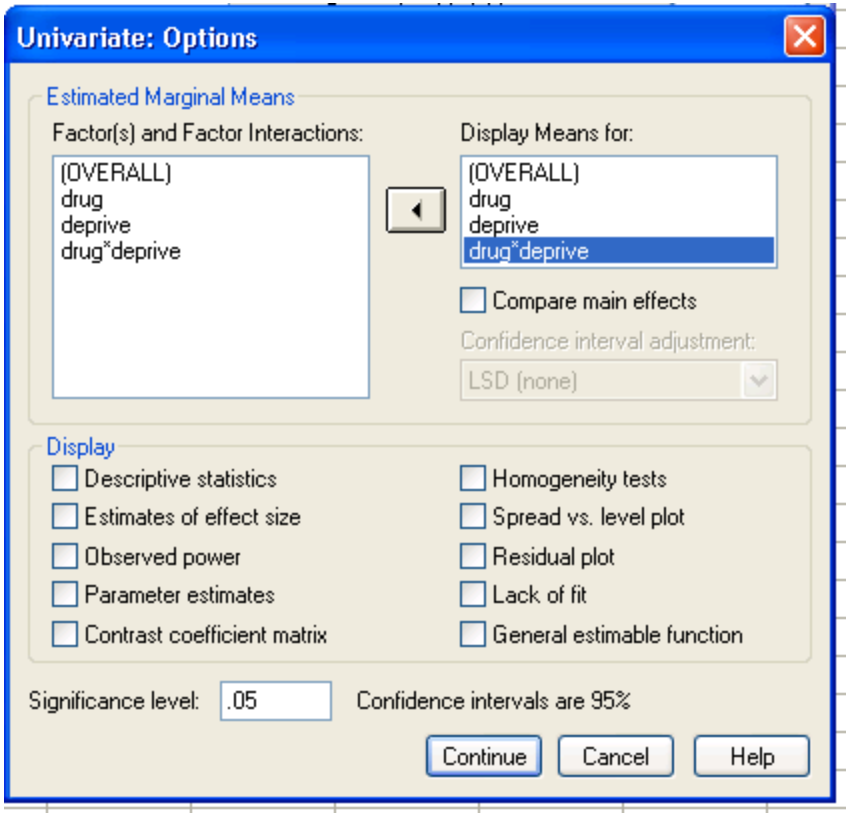
Plots Options…



Main effect plots: horizontal axis only.

Interaction plots: horizontal and separate lines.

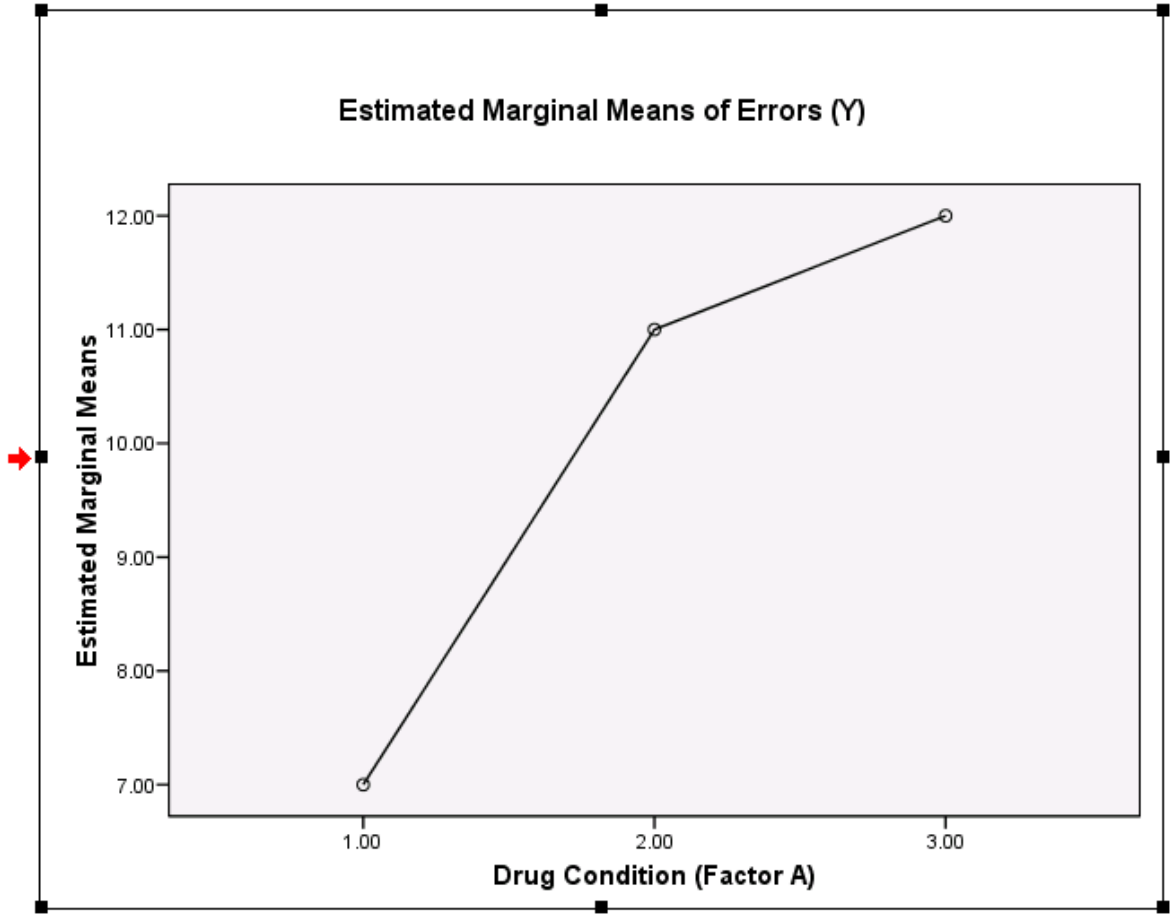
Options Box



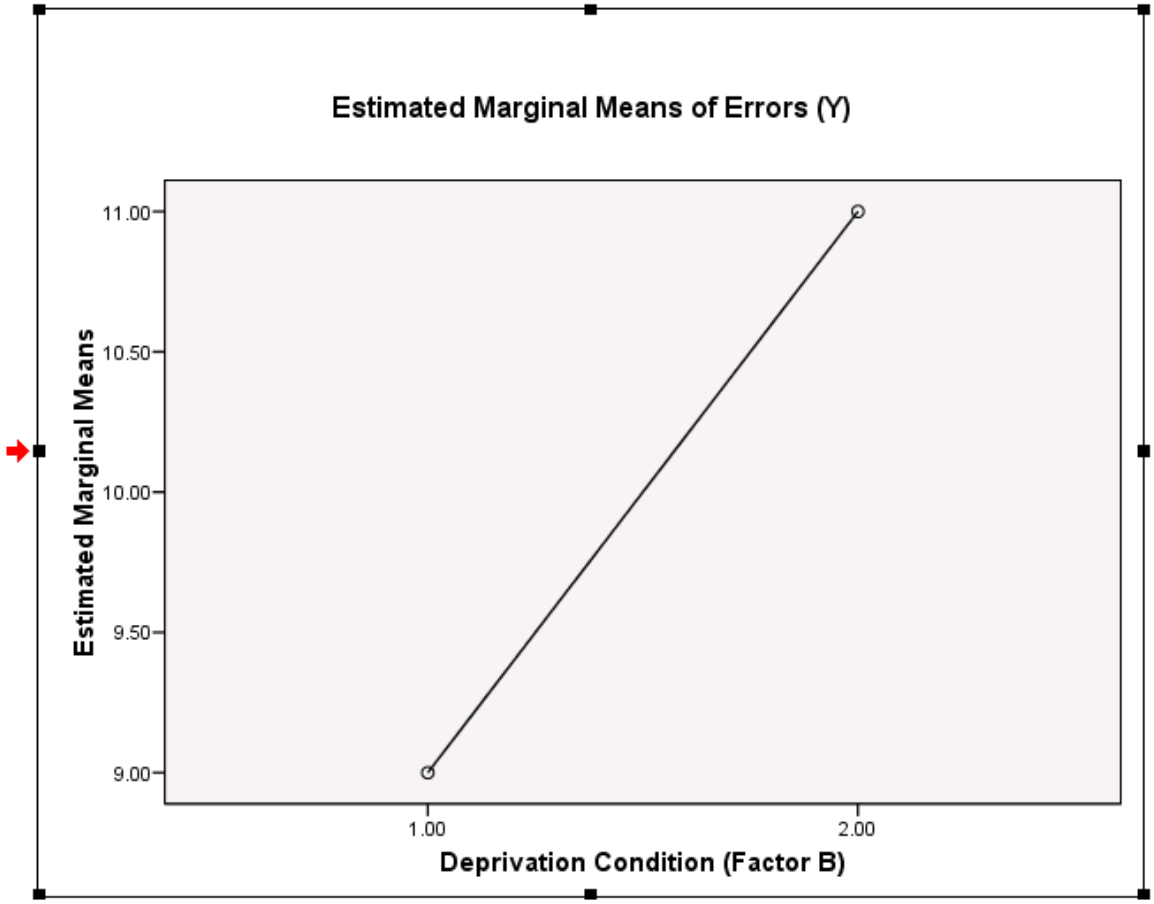
To get the means for each factor level and interaction, use the ‘Options’ box.

Analysis Output

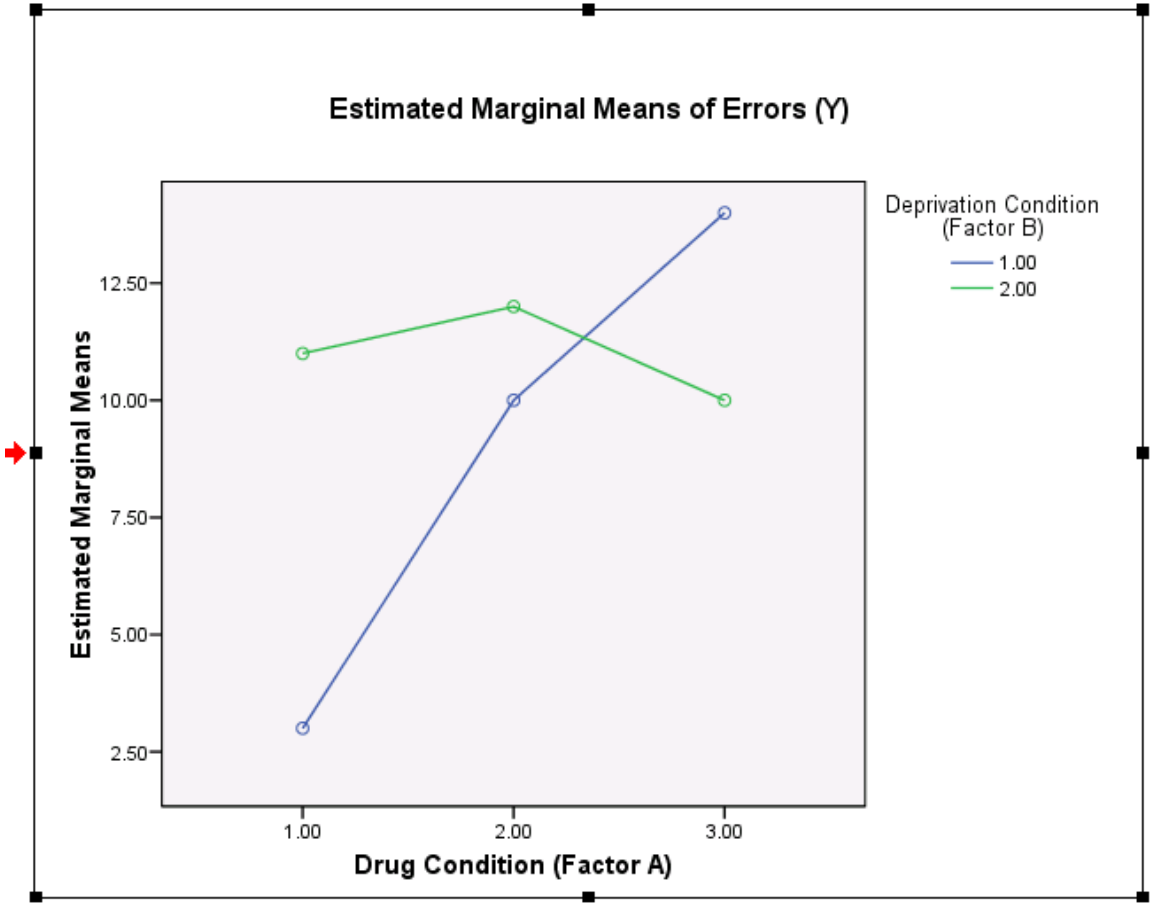
- First, lets look at some mean plots…



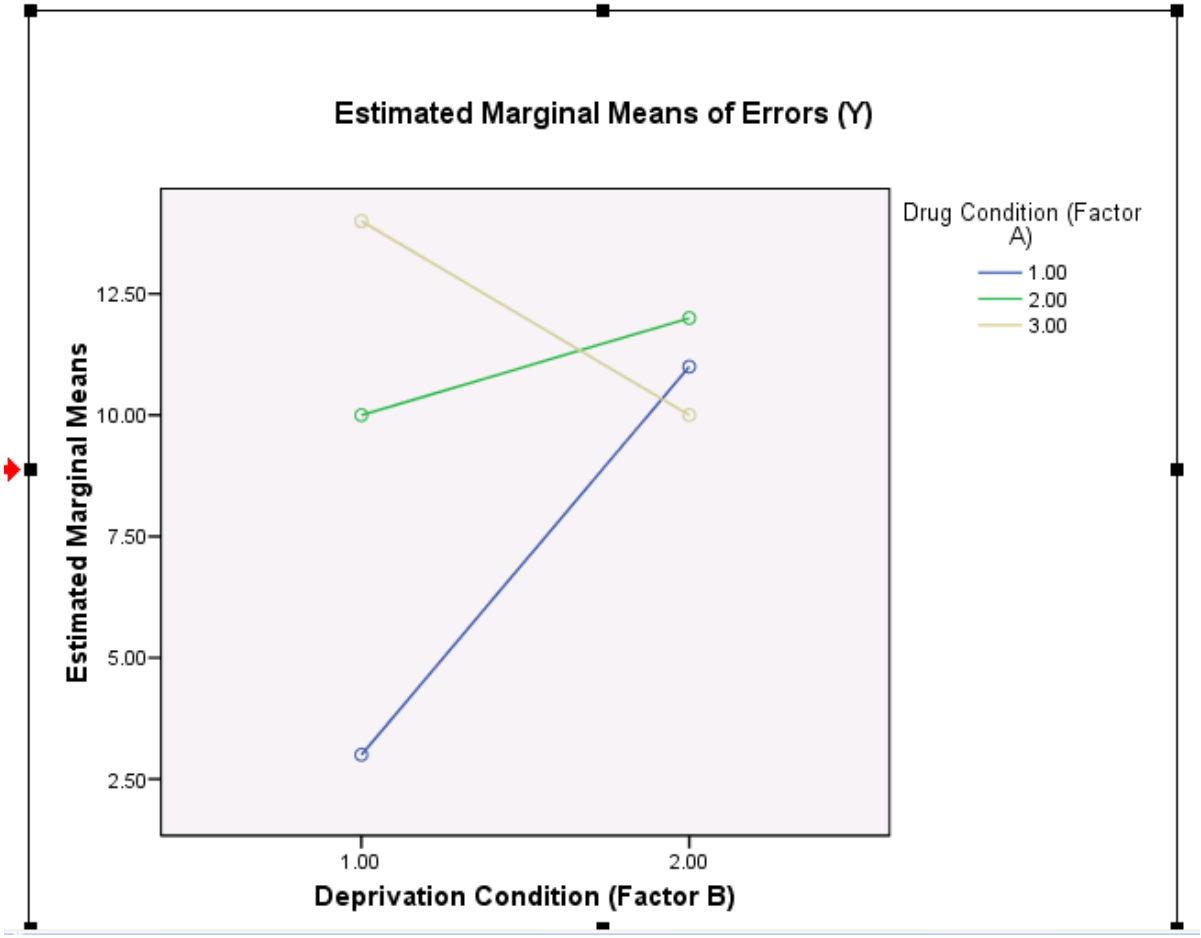
Factor B


Estimated Marginal Means of Errors (Y)
Estimated Marginal Means
11.00
10.50
10.00
9.50
9.00
1.00
2.00
Deprivation Condition (Factor B)

Interaction (Plot #1)



Interaction (Plot #2)



Statistical Output

Between-Subjects Factors

		N
Drug Condition (Factor A)	1.00	8
	2.00	8
	3.00	8
Deprivation Condition (Factor B)	1.00	12
	2.00	12

Main Effect Test for Factor A

Main Effect Test for Factor B

Interaction Test for AxB

Tests of Between-Subjects Effects

Dependent Variable: Errors (Y)

Source	Type III Sum of Squares	df	Mean Square	F	Sig.
Corrected Model	280.000[a]	5	56.000	3.055	.036
Intercept	2400.000	1	2400.000	130.909	.000
drug	112.000	2	56.000	3.055	.072
deprive	24.000	1	24.000	1.309	.268
drug * deprive	144.000	2	72.000	3.927	.038
Error	330.000	18	18.333		
Total	3010.000	24			
Corrected Total	610.000	23			

a. R Squared = .459 (Adjusted R Squared = .309)

Means .

Estimated Marginal Means

1. Grand Mean

Dependent Variable: Errors (Y)

Mean	Std. Error	95% Confidence Interval	
		Lower Bound	Upper Bound
10.000	.874	8.164	11.836

2. Drug Condition (Factor A)

Dependent Variable: Errors (Y)

Drug Condition (Factor A)	Mean	Std. Error	95% Confidence Interval	
			Lower Bound	Upper Bound
1.00	7.000	1.514	3.820	10.180
2.00	11.000	1.514	7.820	14.180
3.00	12.000	1.514	8.820	15.180

3. Deprivation Condition (Factor B)

Dependent Variable: Errors (Y)

Deprivation Condition (Factor B)	Mean	Std. Error	95% Confidence Interval	
			Lower Bound	Upper Bound
1.00	9.000	1.236	6.403	11.597
2.00	11.000	1.236	8.403	13.597

4. Drug Condition (Factor A) * Deprivation Condition (Factor B)

Dependent Variable: Errors (Y)

Drug Condition (Factor A)	Deprivation Condition (Factor B)	Mean	Std. Error	95% Confidence Interval	
				Lower Bound	Upper Bound
1.00	1.00	3.000	2.141	-1.498	7.498
	2.00	11.000	2.141	6.502	15.498
2.00	1.00	10.000	2.141	5.502	14.498
	2.00	12.000	2.141	7.502	16.498
3.00	1.00	14.000	2.141	9.502	18.498
	2.00	10.000	2.141	5.502	14.498

Final Thought

- The two-way ANOVA has more things to look at in the analysis:
  - Main effects for both factors.
    - These are like One-Way ANOVAs.
  - Interaction effects.

- Today we introduced how to obtain Two-Factor results by hand and by SPSS.

- Next time we will see how all of our one-way ANOVA features will apply to multi-factor ANOVA.

Next Class

- More from Chapter 11:
  - More Two-Way ANOVA.