Workshop on SPSS Syntax

1. Using SPSS Syntax to Run Univariate and Bivariate Analyses

Jonathan Zhu

Sept 3, 2014

香港城市大學 City University of Hong Kong
Department of 媒體與傳播系 Media and Communication
Web Mining Lab 互聯網挖掘實驗室

SPSS Windows

- Default:
  - Data editor (→ *.sav)
  - Output viewer (→ *.spv)
- The missing third window:
  - Syntax editor (→ *.sps)

What Can You Do with SPSS Syntax?

- Statistical Analysis
  - Univariate (covered here)
  - Bivariate (covered here)
  - Multivariate
- Variable Transformation
  - Create new variables
  - Change existing variables
  - Rename existing variables
  - etc.
- Data File Manipulation
  - Combine files
  - Split files
  - Aggregate individual cases to summarized cases
  - Transpose variables to cases or cases to variables
  - etc.

Read/Save SPSS Data File

Read SPSS Data:

get file="FILE ADDRESS".

Save to SPSS Data:

save out="FILE ADDRESS".

Elementary Statistical Analysis

- Univariate Analysis:
  - Mean, median, mode
  - Standard deviation, variance
  - Frequency, percentage, cumulated percentage

- Bivariate Analysis

<table>
<tr><td></td><td colspan="3">Independent Variable</td></tr>
<tr><td>Dependent Variable</td><td>Categorical (2 categories)</td><td>Categorical (3+ categ's)</td><td>Continuous</td></tr>
<tr><td>Categorical (2 categ's)</td><td colspan="2" rowspan="2">Crosstabulation (Chi-square test)</td><td>Logistic Regression</td></tr>
<tr><td>Categorical (3+ categ's)</td><td>Multinomial Logistic Reg (NOMREG)</td></tr>
<tr><td>Continuous</td><td>t-test</td><td>ANOVA (F-test)</td><td>Correlation /Regression</td></tr>
</table>

SPSS Syntax for Univariate Analysis

Y is a continuous variable:

descriptive var=***Y***/stat=all.

- "stat=all" generates mean, median, standard deviation, variance, minimum, maximum, kurtosis, skewness, and other descriptive statistics.

Y is a categorical variable:

frequencies var=***Y***
/histogram=normal.

- "histogram=normal" generates a histogram chart, compared with normal distribution.
- "stat=all" can be added to generate similar descriptive statistics as the left.

SPSS Syntax for Bivariate Analysis

Test	SPSS Syntax	Notes
t-test	t-test group=***X*(VALUE1,VALUE2)**/var=***Y***	IV has 2 groups.
ANOVA	means ***Y*** by ***X***/stat=anova.	IV has 3+ groups.
Crosstabulation	crosstabs ***Y*** by ***X***/cell=col/stat=chi.	Show column %.
Correlation	correlate ***Y X***.	Correlation matrix.
Regression	regression dep=***Y***/enter=***X***.	
Logistic Regression	logistic regression ***Y*** with ***X***.	IV has 2 groups.
Multinomial Logistic Regression	nomreg ***Y*** with ***X***.	IV has 3+ groups.

X is the independent variable; ***Y*** is the dependent variable.

Multivariate Analysis

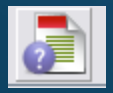
Dependent Variable	Independent Variables		
	Categorical	Mixed of Categorical and Continuous	Continuous
Categorical (2 categories)	Logistic Regression		
Categorical (3+ categories)	Multinomial Logistic Regression (NOMREG)		
Continuous	MANOVA, Multiple Regression	Multiple Regression	

SPSS Syntax for Multivariate Analysis

Test	SPSS Syntax	Notes
Logistic Regression	logistic reg ***Y*** with ***X1 X2*** ... by ***Z1 Z2*** …	***Z1, Z2*** … categorical variables
Multinomial Log Reg	nomreg ***Y*** with ***X1 X2*** ... by ***Z1 Z2*** …	***Z1, Z2*** … categorical variables
MANOVA	manova ***Y*** by ***X1 X2*** ...	
Multiple Regression	reg dep=***Y***/enter ***X1 X2 C1 C2*** ...	***C1, C2*** … dummy coded variables

X is the independent variable; ***Y*** is the dependent variable.

Where to Find Help

- Inside SPSS syntax window:
  - type the first word (e.g., t-test, regression, etc.) of a syntax statement, and click the "Syntax Help" icon  on the top panel
  - the help page of the relevant statement will show up in your web broswer
- Google "spss syntax tutorial" to find dozens of good sites, such as
  - Learning Syntax – Raynald's SPSS Tools
  - Getting Started with SPSS Syntax
  - Statistical Consulting Seminars: Introduction to SPSS Syntax
  - …