Jan 16th, 10:30 AM - 12:00 PM

SPSS Workshop

Pamela Russell
Bridgewater State College, prussell@bridgew.edu

Workshop: Using the Statistical Package for the Social Sciences

Pam Russell, Ph.D.
MAHPLS Department
CART Research Fellow
prussell@bridgew.edu
x2059

Data Entry

- Click DO NOT install Citrix
- Select type in data OR navigate to existing data file
  - If open exiting file – simple close output file
- Organization
  - One row for each subject
  - One column for each variable
    - Code grouping data (categorical = nominal level)
  - Two views
    - Data view (enter numbers)
    - Variable view (variable definitions)
      - Accessible by double-clicking gray "var" column headers

Overview

- SPSS Access
- Data entry
- Parametric statistics assumptions
- Descriptive analyses
  - Checking assumptions
  - Frequencies
  - Descriptives
  - Explore
- T-tests
  - When to use
  - How to use
  - Output Interpretation
  - Expression of results
- Correlation
  - When to use
  - How to use
  - Output Interpretation
  - Expression of results

Defining Variables

- Variable view
- Name variable
  - Not case sensitive
  - 8 characters or less
  - No blank spaces
- Type
  - Nominal (naming)
  - Ordinal (ranking)
  - Ratio/interval (equal intervals)
- Labels (nominal)
  - Define labels
    - Meaningful descriptive name
    - Case sensitive
    - Used in output
  - Input value: label: ADD
- Types
  - Mostly numeric
    - Change format
  - String (no analysis)

SPSS Access

- Login: Citrix Server:
  - On campus: http://citrix.campus.bridgew.edu
  - From BSC homepage: citrix1
  - Off campus: https://citrix.bridgew.edu
- Retrieve data from:
  - Public drive
  - Your Web host (username on 'Webhost' (W:))
  - Your Computer
    - CITRIX – C$ on client (V:) → Documents and Settings → username → Administrators documents → (folders resident your computer)
- Save data same above locations

Defining Variables

- Missing values
  - Best if have none !
  - Choose discrete numbers
    - Not present in the data set
    - Several numbers allow track reason for missing
      - "8" = don't know
      - "9" = not applicable
      - "99" = subject failed to respond
  - Select a range
  - Combine range with single discrete value

Data Import

- Input from other file formats
  - ASCII
  - Excel
- Use from another file (when open)
- Use cut/copy and paste from Excel directly into SPSS data editor

Descriptive Analyses

- Click on STATISTICS box
  - Select appropriate measures of
    - Central tendency (mean, median, mode)
    - Dispersion (SD, variance, range, min, max, SEM)
    - Distribution (skewness, kurtosis)
- Click on CHARTS box
  - Select HISTOGRAM w/normal curve
  - Many other options available
- CONTINUE » OK

Parametric Statistics Assumptions

- Normal distribution
  - Approximates the bell curve – eyeball data
- Homogeneity of Variance
  - SPSS tests differently for each analysis
- Interval Data
  - Distance between individual points on scale are equal ALL along scale
- Independence
  - Data from all subjects is independent of each other

Descriptive Analyses Output

- Statistics for each variable
- Skewness and kurtosis values
  - Calculate z scores to interpret
- Tabulated frequency distributions
- Normal curves overlaid on histogram
- May want to print these out

Descriptive Analyses

- Frequencies
  - Use this to look at the data
  - View frequency distribution
  - Eyeball normality
- ANAYLYZE » DESCRIPTIVE STATISTICS » FREQUENCIES
  - Transfer variables in with arrow key
  - Check off frequency tables

Descriptive Analyses

- To actually test normality-use EXPLORE
- Also allows access to descriptive stats
  - Central tendency and dispersion
- ANAYLYZE » DESCRIPTIVE STATISTICS » EXPLORE
  - Transfer variables into dependent list with arrow key

Descriptive Analyses

- Click on STATISTICS box
  - Select appropriate measures (descriptives)
- Click on PLOTS box
  - Check normality plots with tests
  - Under box plots
    - Select factor levels together
  - Under descriptive
    - Select stem and leaf plots

T-Test Output

- Three tables
  - Paired sample statistics
  - Pearson's correlation w/ 2-tailed signif
  - Paired samples test
    - Stats for pairs, CI for 95% of scores, t and dof
    - Sig for 2-tailed (divide by 2 to get 1-tailed)
- Writing up results
  - Include dof, value of t, and p level/value
  - EXAMPLE: (t(dof) = t value, p < .05)

Descriptive Analyses Output

- Check significance - Kolmogorov-Smirnov test
  - Sig > .05 – data normally distributed
  - Sig < .05 – data NOT normal !
- Stem and leaf plots
  - Show observed data overlaid on expected normal (straight line)
- Without normal distribution move to non-parametric statistics

T-Test

- ANALYZE » COMPARE MEANS » INDEPENDENT SAMPLES T TEST
  - Move dependent variable in with arrow
  - Select grouping variable
    - DEFINE GROUPS
      - Enter codes (1 = female, 2 = male)
      - Could use > or < as cut points
  - OPTIONS – select confidence interval (95%)
    - Exclude cases by analysis for missing values

T-Test

- Assess differences on one dependent variable (outcome measure)
  - One sample versus a single score (criterion)
  - Dependent samples (paired t-test)
    - Two conditions different subjects for each
  - Independent samples
    - Two conditions different subjects for each
- ANALYZE » COMPARE MEANS » PAIRED SAMPLES T TEST
  - Move pair of variables in with arrow
  - OPTIONS – select confidence interval (95%)
    - Exclude cases by analysis for missing values

T-Test Output

- Two tables
  - Group statistics
  - Independent samples t-test
    - Check equality of variances
      - If Levene's p<.05 variances are significantly different
      - Read t-test results from appropriate row
    - Stats, CI for 95% of scores, t and dof
    - Sig for 2-tailed (divide by 2 to get 1-tailed)
- Writing up results
  - Include dof, value of t, and p level/value
  - EXAMPLE: (t(dof) = t value, p < .05)
  - State means and SD
  - Indicate greater/smaller mean

Correlation

- Use a correlation to assess degree of linear relationship between two variables
- Use scatterplot first
  - Check general trend of data
  - Look for outliers
  - GRAPHS » INTERACTIVE » SCATTER

Correlation

- Specify one OR two- tailed test
  - Match to hypothesis
  - Check Flag of significant correlations
- Options box
  - Select statistics
  - Choose how to deal with missing values

Scatterplot

- Define a SIMPLE scatter plot (2 variables)
- Set up scatterplot axes
  - Y axis = usually dependent variable
  - X axis = independent variable
  - Panel variables
    - Use a grouping variable
  - Label cases by
    - Names individual cases
    - Helpful for outliers
    - Difficult with large data sets
- Check output for trends and outliers

Correlation Output

- Correlation matrix
  - More than 2 variables
  - Ones on the diagonal
  - Lower left triangle = upper right triangle
- Look for correlation value, n, & sig
- Writing your results
  - State strength and direction of correlation
  - Provide (r(n) = value of r, $p < .05$)
  - State direction of relationship using correlated variables

Correlation

- Types
  - BIVARIATE – 2 variables
  - PARTIAL – 2 variables & control for third
- ANALYZE » CORRELATE » BIVARIATE
  - Move two variables into variable list
  - Select correlation type
    - Pearson's r (interval or ratio data)
    - Spearman (non-parametric statistic)
    - Kendall's (non-parametric statistic)

Sources of Support

- Current CART Research Fellow
  prussell@bridgew.edu
- Statistics Workshops
  - Given this spring semester
  - Hank Vandenburgh, Walter Caroll
- SPSS tutorials and help within software
- University of Texas, Statistical Support Website SPSS for Windows
  - http://www.utexas.edu/its/rc/tutorials/