

# SPSS ADVANCED ANALYSIS



**WENDIANN SETHI**  
**SPRING 2011**

# Statistical techniques to be covered

- **Explore relationships among variables**

- Correlation
- Regression/Multiple regression
- Logistic regression
- Factor analysis

- **Compare groups**

- Non-parametric statistics
- T-tests
- One-way analysis of variance ANOVA
- Two-way between groups ANOVA
- Multivariate analysis of variance MANOVA

# Correlation



Aim: find out whether a relationship exists and determine its magnitude and direction

Correlation coefficients:

Pearson product moment correlation coefficient

Spearman rank order correlation coefficient

Assumptions:

relationship is linear

Homoscedasticity: variability of DV should remain constant at all values of IV

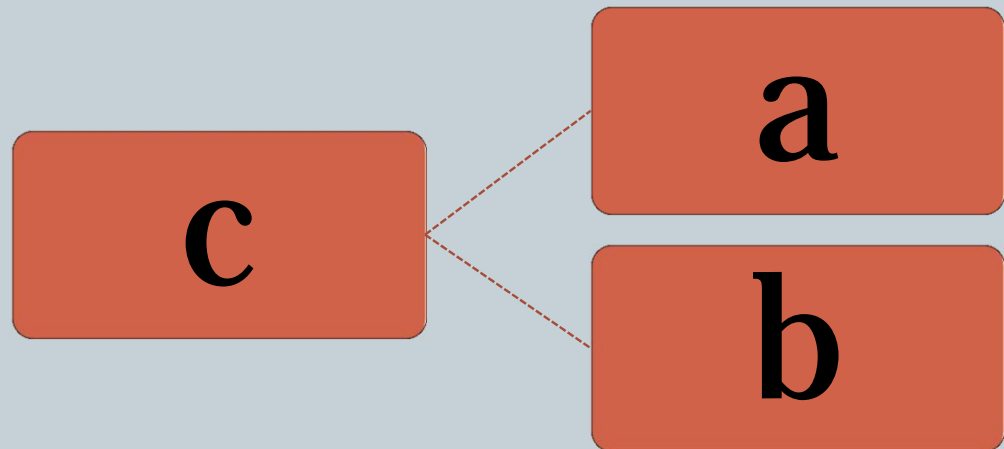
# Partial correlation



Aim: to explore the relationship between two variables while statistically controlling for the effect of another variable that may be influencing the relationship

Assumptions:

same as correlation



# Regression



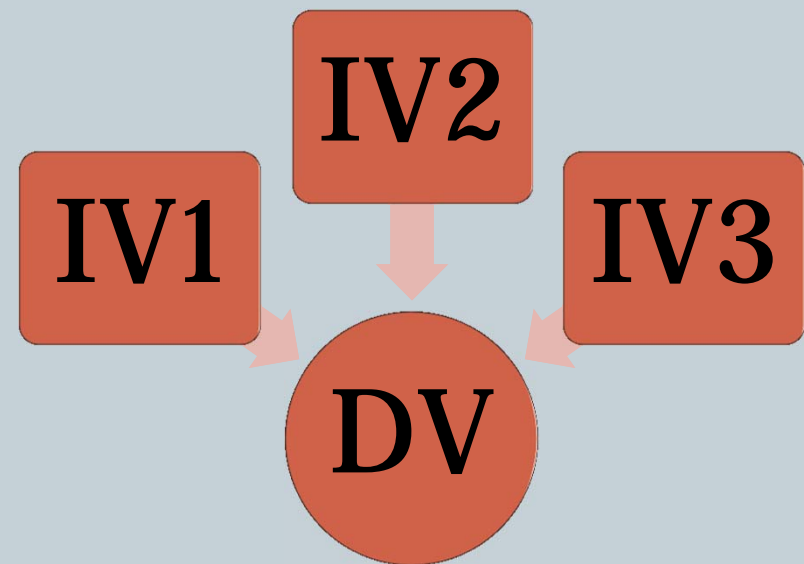
Aim: use after there is a significant correlation to find the appropriate linear model to predict DV (scale or ordinal) from one or more IV (scale or ordinal)

Assumptions:

- sample size needs to be large enough
- multicollinearity and singularity
- outliers
- normality
- linearity
- homoscedasticity

Types:

- standard
- hierarchical
- stepwise



# Logistic regression



**Aim:** create a model to predict DV (categorical – 2 or more categories) given one or more IV (categorical or numerical/scale)

**Assumptions:**

sample size large enough

multicollinearity

outliers

**Procedure note:**

use Binary Logistic for DV of 2 categories (coding 0/1)

use Multinomial Logistic for DV for more than 2 categories

# Factor analysis



Aim: to find what items (variables) clump together.  
Usually used to create subscales. Data reduction.

Factor analysis:

exploratory

confirmatory

SPSS -> Principal component analysis

# Three steps of factor analysis



- 1. Assessment of the suitability of the data for factor analysis**
- 2. Factor extraction**
- 3. Factor rotation and interpretation**

# 1. Assessment of the suitability



1. Sample size: 10 to 1 ratio
2. Strength of the relationship among variables (items)

## Step 2. Factor extraction



1. Commonly used technique principal components analysis
2. Kaiser's criterion: only factors with eigenvalues of 1.0 or more are retained – may give too many factors
3. Scree test: plot of the eigenvalues, retain all the factors above the “elbow”
4. Parallel analysis: compares the size of the eigenvalues with those obtained from randomly generated data set of the same size

# Step 3: factor rotation and interpretation



## 1. Orthogonal rotation

1. uncorrelated
2. Easier to interpret
3. Varimax

## 2. Oblique rotation

1. Correlated
2. Harder to interpret
3. Direct Oblimin

# Statistical techniques to be covered

- **Explore relationships among variables**

- Correlation
- Regression/Multiple regression
- Logistic regression
- Factor analysis

- **Compare groups**

- Non-parametric statistics
- T-tests
- One-way analysis of variance ANOVA
- Two-way between groups ANOVA
- Multivariate analysis of variance MANOVA

# Nonparametric tests



<b>Non-parametric techniques</b>	<b>Parametric techniques</b>
Chi-square test for goodness of fit	None
Chi-square test for independence	None
Kappa measure of agreement	None
Mann-Whitney U Test	Independent samples t-test
Wilcoxon Signed Rank Test	Paired samples t-test
Kruskal-Wallis Test	One-way between groups ANOVA
Friedman Test	One-way repeated measures ANOVA

# T-test for independent groups



- ▶ **Aim:** Testing the differences between the means of two independent samples or groups
- ▶ **Requirements:**
  - Only one independent (grouping) variable IV (ex. Gender)
  - Only two levels for that IV (ex. Male or Female)
  - Only one dependent variable (DV)
- ▶ **Assumptions:**
  - Sampling distribution of the difference between the means is normally distributed
  - Homogeneity of variances – Tested by Levene’s Test for Equality of Variances
- ▶ **Procedure:**
  - ANALYZE>COMPARE MEANS>INDEPENDENT SAMPLES T-TEST
  - Test variable – DV
  - Grouping variable – IV
  - DEFINE GROUPS (need to remember your coding of the IV)
  - Can also divide a range by using a cut point

# Paired Samples T-test



- ▶ **Aim:** used in repeated measures or correlated groups designs, each subject is tested twice on the same variable, also matched pairs
- ▶ **Requirements:**
  - Looking at two sets of data – (ex. pre-test vs. post-test)
  - Two sets of data must be obtained from the same subjects or from two matched groups of subjects
- ▶ **Assumptions:**
  - Sampling distribution of the means is normally distributed
  - Sampling distribution of the difference scores should be normally distributed
- ▶ **Procedure:**
  - ANALYZE>COMPARE MEANS>PAIRED SAMPLES T-TEST

# One-way Analysis of Variance



- ▶ **Aim:** looks at the means from several independent groups, extension of the independent sample t-test
- ▶ **Requirements:**
  - Only one IV (categorical)
  - More than two levels for that IV
  - Only one DV (numerical)
- ▶ **Assumptions:**
  - The populations that the sample are drawn are normally distributed
  - Homogeneity of variances
  - Observations are all independent of one another
- ▶ **Procedure:**
  - ANALYZE>COMPARE MEANS>One-Way ANOVA
    - Dependent List – DV
    - Factor – IV

# Two-way Analysis of Variance



▶ **Aim:** test for main effect and interaction effects on the DV

▶ **Requirements:**

- Two IV (categorical variables)
- Only one DV (continuous variable)

▶ **Procedure:**

ANALYZE>General Linear Model>Univariate

- Dependent List – DV
- Fixed Factor – IVs

# MANOVA



**Aim: extension of ANOVA when there is more than one DV (should be related)**

**Assumptions:**

sample size

normality

outliers

linearity

homogeneity of regression

multicollinearity and singularity

homogeneity of variance-covariance matrices