Photo: Vladimir Batagelj, ***UNI-LJ***

Introduction to Multivariate Analysis

Anuška Ferligoj

UL, Ljubljana, Slovenia

NRU HSE, Moscow, Russia

Content

All the materials for the course are on:

http://vladowiki.fmf.uni-lj.si/doku.php?id=ru:hse:mva

Types of statistical analysis

Univariate analysis

Analysis of one variable

- frequency distribution (e.g., table, histogram)
- mean values (e.g., mode, median, aritmetic mean)
- measures of variation (e.g., range, mean absolute deviation, variance)
- measures of skewness
- measures of kurtosis

Bivariate analysis

Analysis of two variables

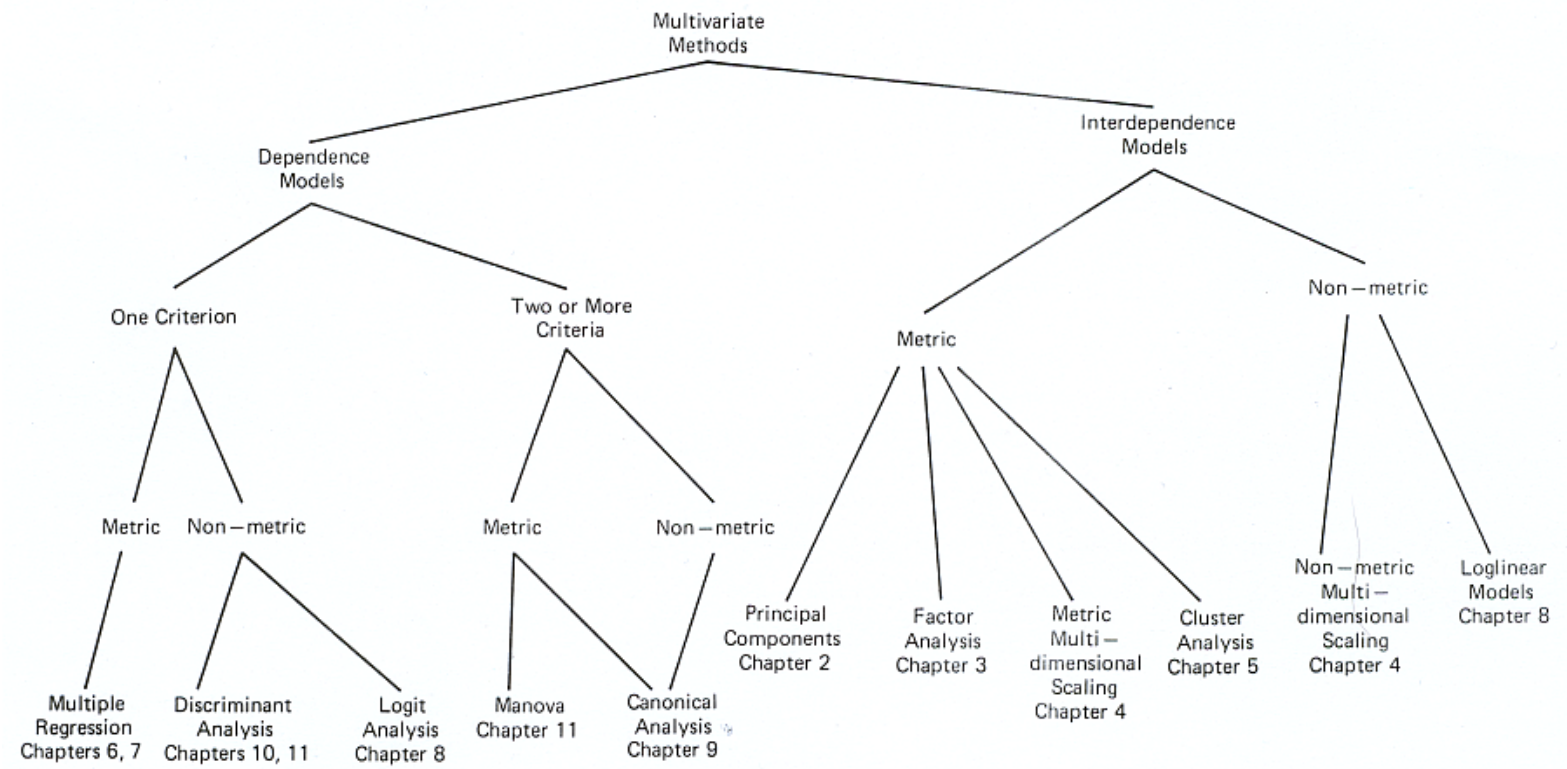
- for interval and ratio variables: Pearson's correlation coefficient
- for ordinal variables: Spearman's rank correlation coefficient
- for nominal variables: χ^2, contingency coefficients

Multivariate analysis

Analysis of several variables

- graphical presentation
- cluster analysis
- principal component analysis
- multiple regression
- factor analysis
- canonical correlation analysis
- discriminant analysis
- ...

Overview of multivariate analysis



Matrix notation

Data matrix:

$$X = \begin{bmatrix} x_{11} & x_{12} & \dots & x_{1m} \\ x_{21} & x_{22} & \dots & x_{2m} \\ \vdots & \vdots & \ddots & \vdots \\ x_{n1} & x_{n2} & \dots & x_{nm} \end{bmatrix}$$

n units

m variables

Mean:

$$\mu = \begin{bmatrix} \mu_1 \\ \vdots \\ \mu_m \end{bmatrix}$$

Variance–covariance matrix:

$$\Sigma = \begin{bmatrix} \sigma_{11} & \sigma_{12} & \dots & \sigma_{1m} \\ \sigma_{21} & \sigma_{22} & \dots & \sigma_{2m} \\ \vdots & \vdots & \ddots & \vdots \\ \sigma_{m1} & \sigma_{m2} & \dots & \sigma_{mm} \end{bmatrix}$$

Linear combination

Linear combination of m variables:

$$Y = a_1X_1 + a_2X_2 + ... + a_mX_m = Xa$$

$$a = \begin{bmatrix} a_1 \\ \vdots \\ a_m \end{bmatrix}$$

$$\mu_Y = a'\mu$$

$$\text{var}(Y) = \text{var}(Xa) = a'\Sigma a$$