

This work is licensed under a [Creative Commons Attribution-NonCommercial-ShareAlike License](https://creativecommons.org/licenses/by-nc-sa/4.0/). Your use of this material constitutes acceptance of that license and the conditions of use of materials on this site.



Copyright 2006, The Johns Hopkins University and Elizabeth Garrett-Mayer. All rights reserved. Use of these materials permitted only in accordance with license rights granted. Materials provided "AS IS"; no representations or warranties provided. User assumes all responsibility for use, and all liability related thereto, and must independently review all materials for accuracy and efficacy. May contain materials owned by others. User is responsible for obtaining permissions for use from third parties as needed.

Statistics in Psychosocial
Research
Lecture 9
Factor Analysis II

Lecturer: Elizabeth Garrett-Mayer

Uniqueness Issues

- One covariance structure can be produced by the same number of common factors with a different configuration
- One covariance structure can be produced by factor models with different numbers of common factors
- One covariance structure can be produced by a factor model and also by a non-factor analytic model.

Methods for Extracting Factors

- Principal Components (already discussed)
 - Principal Factor Method
 - Iterated Principal Factor / Least Squares
 - Maximum Likelihood (ML)
- ★ Most common(?): ML and Least Squares

Principal Factor Analysis

Uses communalities to estimate and assumes true communalities are the squared multiple correlation coefficients (as described in one of previous slides).

- Uniqueness(X) \approx $\text{Var}(e)$
- $\text{Var}(X) = \text{Comm}(X) + \text{Uniqueness}(X)$.

(1) Estimate uniqueness (i.e., $\text{var}(e_j)$) using $1 - R^2$ where R^2 is for a regression of X_j on all other X 's.

This assumes that the “communality” of X is the same as the amount of X described by the other X 's.

(2) (Correlation) Perform principal components on:

$$\text{Corr}(X) - \text{Var}(e) \cong \text{Communality}$$

$$\begin{pmatrix} 1 & \dots & \text{Corr}(X_1, X_n) \\ \dots & \dots & \dots \\ \text{Corr}(X_n, X_1) & \dots & 1 \end{pmatrix} - \begin{pmatrix} \text{Var}(e_1) & \dots & 0 \\ \dots & \dots & \dots \\ 0 & \dots & \text{Var}(e_n) \end{pmatrix} = \begin{pmatrix} 1 - \text{Var}(e_1) & \dots & \text{Corr}(X_1, X_n) \\ \dots & \dots & \dots \\ \text{Corr}(X_n, X_1) & \dots & 1 - \text{Var}(e_n) \end{pmatrix}$$

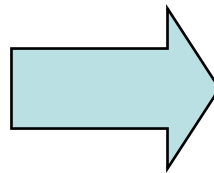
Principal Factor Analysis

- Simplified explanation
- Steps:
 1. Get initial estimates of communalities
 - squared multiple correlations
 - highest absolute correlation in row
 2. Take correlation matrix and replace diagonal elements by communalities. We call this the “adjusted” correlation matrix.
 3. Apply principal components analysis

Principal Factor Analysis

1. Obtain correlation (covariance) matrix

1	r_{12}	r_{13}	r_{14}	r_{15}	r_{16}	r_{17}
r_{21}	1	r_{23}	r_{24}	r_{25}	r_{26}	r_{27}
r_{31}	r_{32}	1	r_{34}	r_{35}	r_{36}	r_{37}
r_{41}	r_{42}	r_{43}	1	r_{45}	r_{46}	r_{47}
r_{51}	r_{52}	r_{53}	r_{55}	1	r_{56}	r_{57}
r_{61}	r_{62}	r_{63}	r_{64}	r_{65}	1	r_{67}
r_{71}	r_{72}	r_{73}	r_{74}	r_{75}	r_{76}	1



2. Replace 1's (variances) with estimate of communality

h_1^2	r_{12}	r_{13}	r_{14}	r_{15}	r_{16}	r_{17}
r_{21}	h_2^2	r_{23}	r_{24}	r_{25}	r_{26}	r_{27}
r_{31}	r_{32}	h_3^2	r_{34}	r_{35}	r_{36}	r_{37}
r_{41}	r_{42}	r_{43}	h_4^2	r_{45}	r_{46}	r_{47}
r_{51}	r_{52}	r_{53}	r_{55}	h_5^2	r_{56}	r_{57}
r_{61}	r_{62}	r_{63}	r_{64}	r_{65}	h_6^2	r_{67}
r_{71}	r_{72}	r_{73}	r_{74}	r_{75}	r_{76}	h_7^2

3. Apply principal components to "adjusted" correlation matrix and use results.

Iterative Principal Factor / Least Squares

1. Perform Principal Factor as described above.
2. Instead of stopping after principal components, reestimate communalities based on loadings.
3. Repeat until convergence

Better than without iterating!

Iterated Principal Factors / Least Squares

Standard Least Squares approach

Minimize:

$$\sum_j \left(1 - \text{Comm}(X_j) - \text{Var}(e_j)\right)^2$$

Maximum Likelihood Method

- Assume F's are normal
- Use likelihood function
- Maximize parameters
- Iterative procedure
- Notes:
 - normality matters!
 - Estimates can get “stuck” at boundary (e.g. communality of 0 or 1).
 - Software choice matters (e.g. SPSS vs. Stata)
 - Must rotate for interpretable solution

Choice of Method

- Give different results because they
 - use different procedures
 - use different restrictions
 - make different assumptions
- Benefit of ML
 - Can get statistics which allow you to compare factor analytic models

Which Method Should You Use?

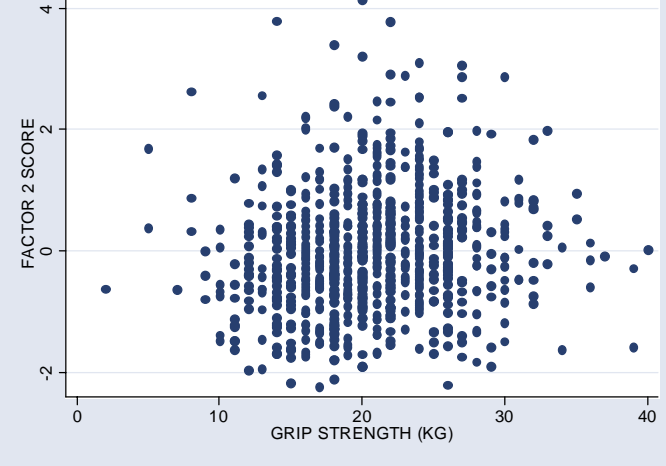
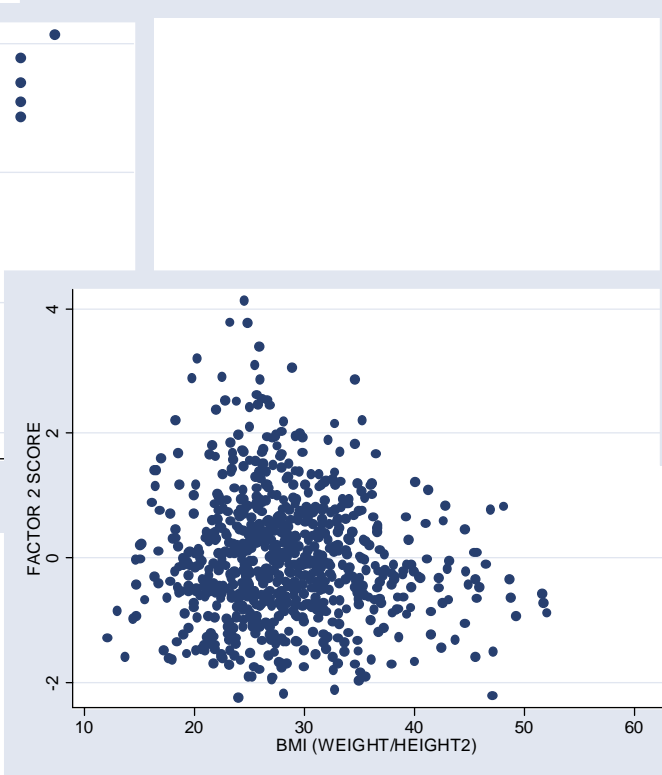
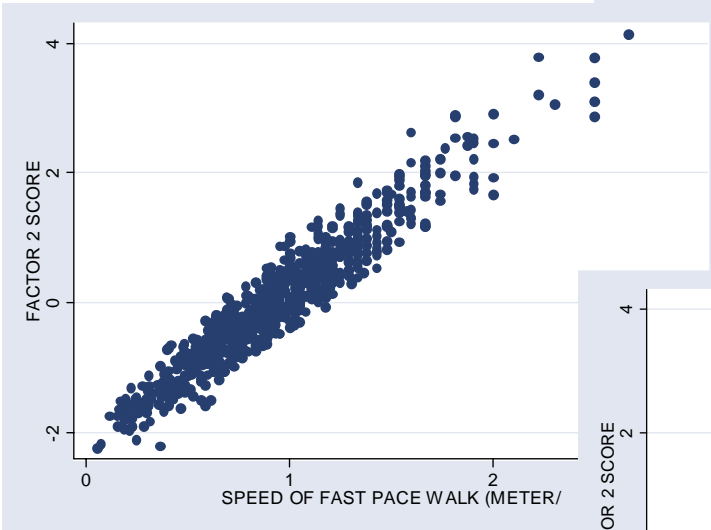
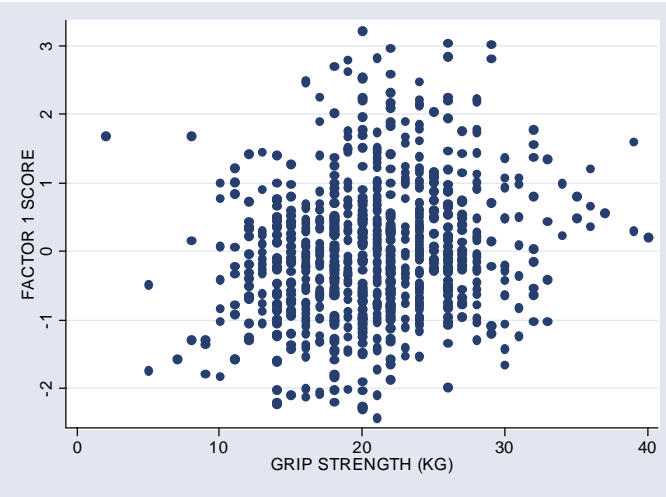
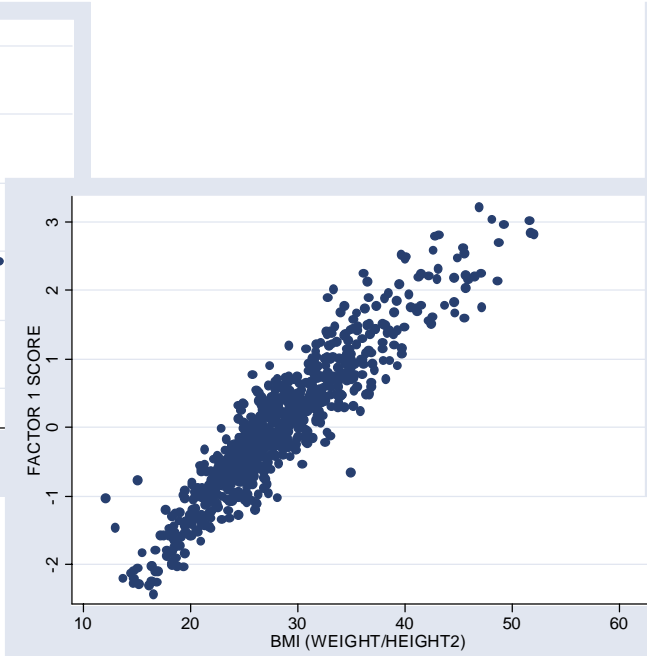
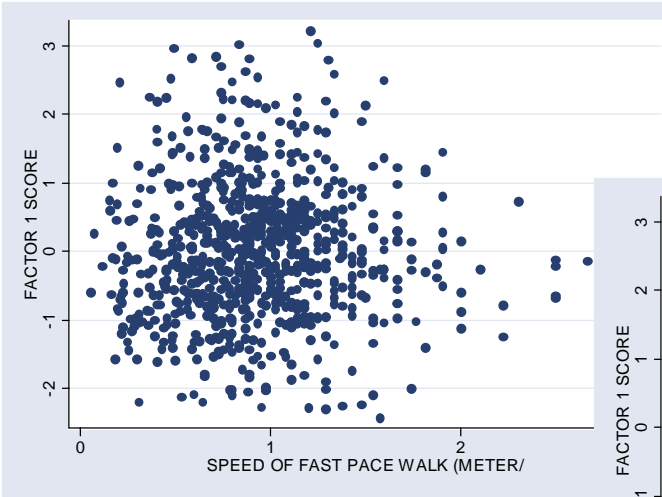
- Statisticians: PC and ML
- Social Sciences: LS and Principal Factor
- Stata:
 - ‘pca’ command (principal components) in Stata 8
 - ‘factor, pc’ command (principal components) in Stata7
 - ‘factor, pf’ (principal factor)
 - ‘factor, ipf’ (iterated principal factor)
 - ‘factor, ml’ (maximum likelihood)
- ★ Caution! *ipf* and *ml* may not converge to the right answer! Look for uniqueness of 0 or 1. Problem of “identifiability” or getting “stuck.” Would be nice if it would tell you....
- ★ For this class? I like IPF or ML, but don’t always converge.

Correlation vs. Covariance?

- Correlation MUCH more commonly seen.
- If using covariance, want measures in comparable units.
- Correlation for validation, covariance for summary
 - if summary is the goal, relative variation in X's matter.
- If using covariance, do not use “number of eigenvalues > 1” or “scree plot” for determining number of factors! Nonsensical!
- Stata: all factor analyses are based on correlations. Only can use covariance in PC.

Factor Scores and Scales

- Each object (e.g. each woman) gets a factor score for each factor.
- Old data vs. New data
- The factors themselves are variables
- “Object’s” score is weighted combination of scores on input variables
- These weights are NOT the factor loadings!
- Loadings and weights determined simultaneously so that there is no correlation between resulting factors.
- We won’t bother here with the mathematics....



Interpretation

- Naming of Factors
- Wrong Interpretation: Factors represent separate groups of people.
- Right Interpretation: Each factor represents a continuum along which people vary (and dimensions are orthogonal if orthogonal)

Exploratory versus Confirmatory

- Exploratory:
 - summarize data
 - describe correlation structure between variables
 - generate hypotheses
- Confirmatory (more next term)
 - testing consistency with a preconceived theory
 - A kind of structural equation modeling
 - Usually force certain loadings to zero.
 - More useful when looking for associations between factors or between factors and other observed variables.

Test Based Inference for Factor Analysis

- Likelihood ratio test:
 - Goodness of fit: compares model prediction to observed data. Sensitive to sample size.
 - Comparing models: Compare deviance statistics from different models
- Information Criteria:
 - Choose model with highest IC. Penalize for number of parameters (principle of parsimony)
 - BIC (Schwarz):
 $\log(L) - \ln(N/2) * (\text{number of parameters})$
 - AIC (Akaike):
 $\log(L) - (\text{number of parameters in model})$

Factor Analysis with Categorical Observed Variables

- Factor analysis hinges on the correlation matrix
- As long as you can get an interpretable correlation matrix, you can perform factor analysis
- Binary items?
 - Tetrachoric correlation
 - Expect attenuation!
- Mixture of items?
 - Mixture of measures
 - All must be on comparable scale.

Criticisms of Factor Analysis

- Labels of factors can be arbitrary or lack scientific basis
- Derived factors often very obvious
 - defense: but we get a quantification
- “Garbage in, garbage out”
 - really a criticism of input variables
 - factor analysis reorganizes input matrix
- Too many steps that could affect results
- Too complicated
- Correlation matrix is often poor measure of association of input variables.

Stata Commands

factor y1 y2 y3,

Options:

Estimation method: *pcf, pf, ipf, and ml* (default is *pf*)

Number of factors to keep: *factor(#)*

Use covariance instead of correlation matrix: *cov*

* For Stata 8, to do principal components, must use '*pca*' command!

Post-factor commands:

rotation: *rotate* or *rotate, promax*

screepplot: *greigen*

generate factor scores: *score f1 f2 f3...*