Statistics for Psychosocial Research

Lecturer: William Eaton

1. First term of a two-term course
2. This term focuses on the evaluation of psychosocial measures
3. Second term builds on the first and focuses on ‘latent’ constructs -- measures that cannot be directly observed.

Role of Statistics In This Course

1. We will present basic formulas, and expect you to be able to interpret them and tell us why they Are important
2. We do not expect you to memorize formulas, unless we state otherwise
3. Do not need to know calculus
4. Do not need to be able to derive formulas

Session 1. Introduction

1. Introduction
2. Purpose of course
3. Review of syllabus
4. Classical test theory
5. Introduction to reliability, examples
6. After this class students will be able to:
   a. briefly describe the concept of reliability in both intuitive and statistical terms
   b. identify the key assumptions of classical test theory.

Introduction to Reliability

1. Consistency between two measures of the same thing

2. Ratio of true to total variance

Example of Test-Retest Reliability

1. Elderly self-report of the extent to which arthritis affects their lives.
2. Elderly self-report of their ability to perform basic tasks such as eating, dressing, and walking.

Example of Test-Retest Reliability

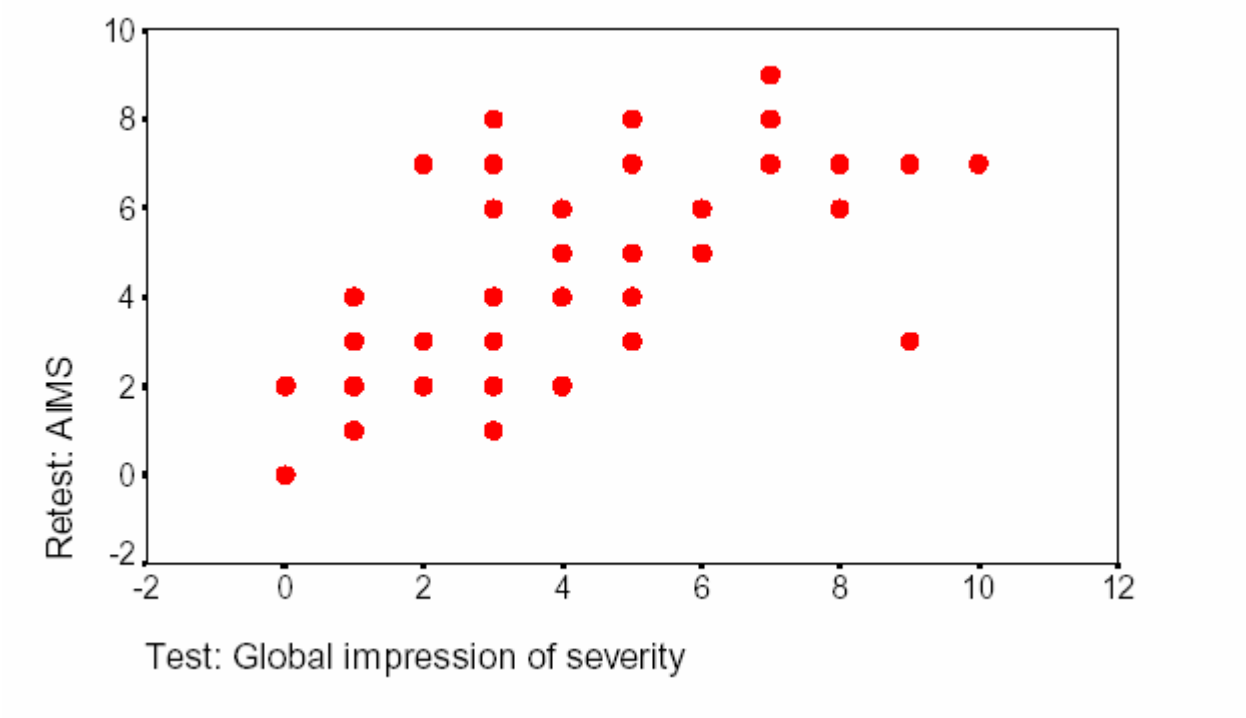


Figure 1: Test-retest scatterplot of 'global impression of severity,' which indexes respondents' global assessment of the extent to which arthritis affects them (0= not at all, 11=extreme). Correlation above is .66. (Mail questionnaire, test-retest time is two weeks n=51).

Source: chpe.buseco.monash.edu.au/pubs/wp116.pdf

Example of Test-Retest Reliability

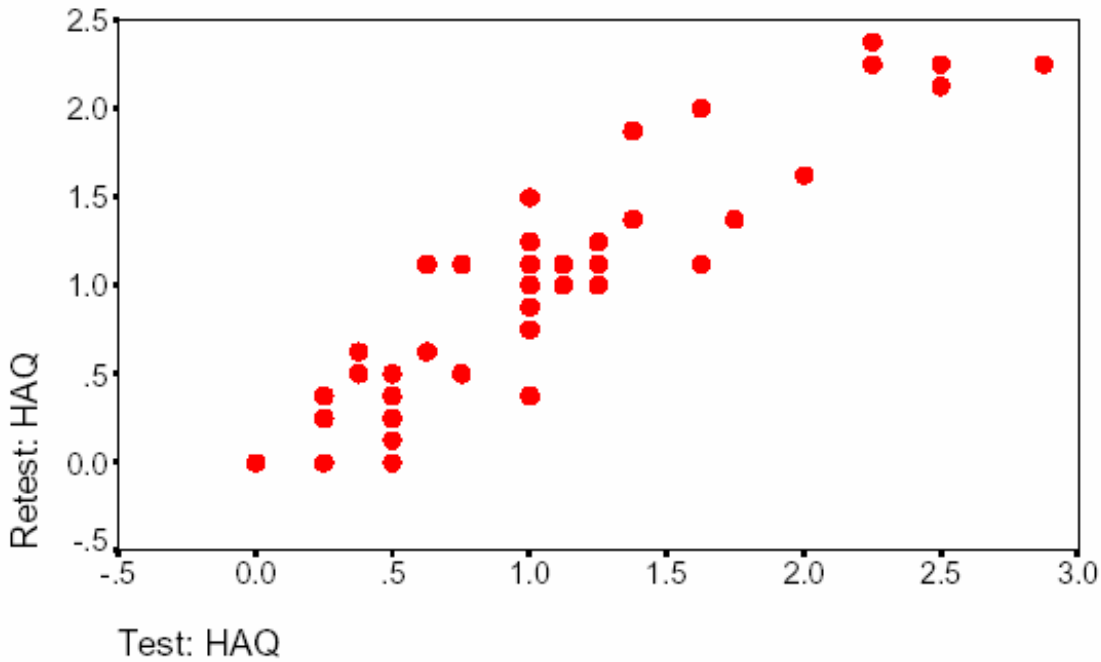


Figure 2: Test-retest scatterplot of the HAQ disability index, which comprised of eight weighted scales indexing ability to perform basic tasks such as eating, dressing, and walking (0=no difficulty, 3=extreme difficulty). Correlation above is .93

Source: chpe.buseco.monash.edu.au/pubs/wp116.pdf

Example of Inter-Rater Reliability

1. Correspondence between mother and father's report of their child's impulsivity.
2. Correspondence between mother and child's report of child's impulsivity.

Example of Inter-Rater Reliability

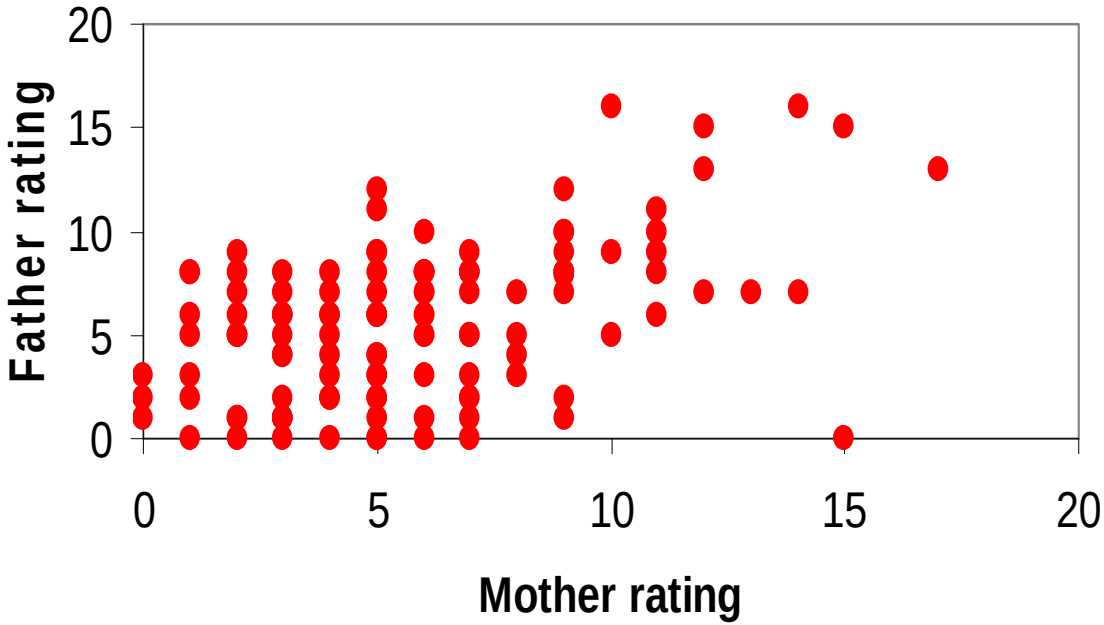


Figure 3: Inter-rater reliability of child's impulsivity, a scale that is comprised of twenty items (0=low and 20=high). Correlation above is .56.

Source: chpe.buseco.monash.edu.au/pubs/wp116.pdf

Example of Inter-Rater Reliability

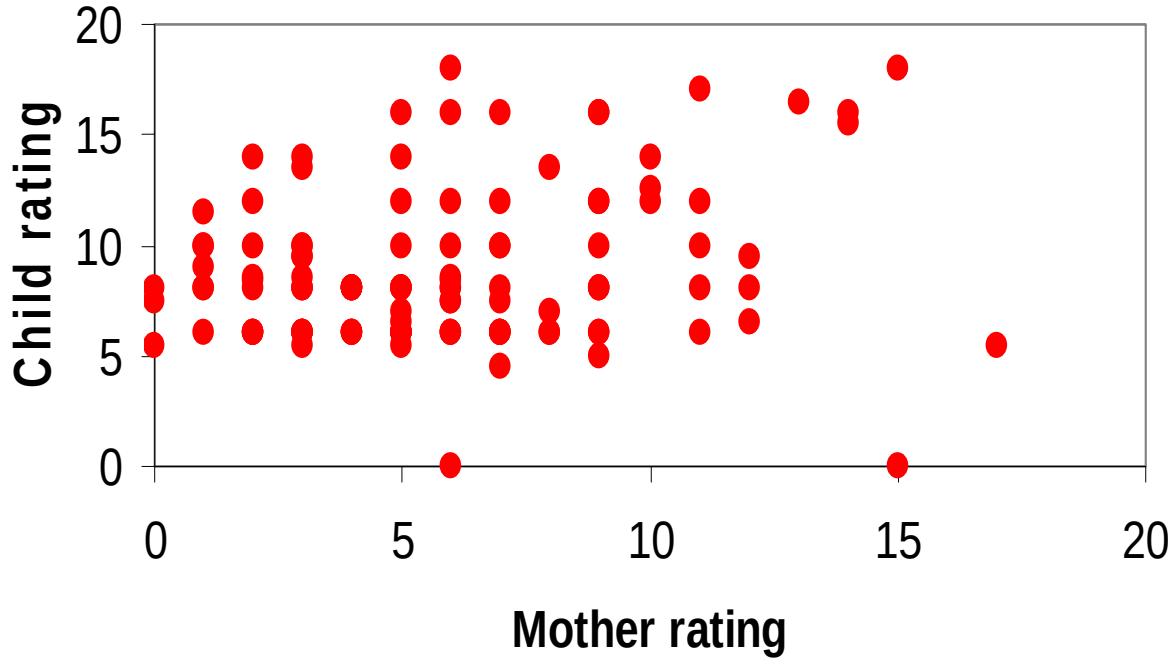


Figure 4: Inter-rater reliability of child's impulsivity, a scale that is comprised of twenty items (0=low and 20=high). Correlation above is .30.

Source: chpe.buseco.monash.edu.au/pubs/wp116.pdf

Internal Consistency Reliability

The degree to which items measuring the same construct are associated with each other

Strongly agree, agree somewhat, disagree somewhat, strongly disagree:

1. “I am depressed”
2. “I feel sad”
3. “I am blue”
4. “I feel happy”
5. “I am content”

Classical Test Theory

$x = T_x + e$

Assumptions:

1) $E(e) = 0$
2) $cov(T_x,e) = 0$
3) $cov(e_i,e_j) = 0$

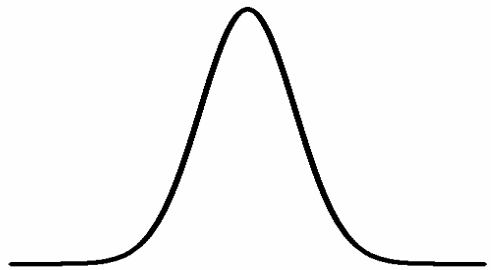
N.B.:

$Var\ (X) = Var\ (T_x + e)$

$= Var\ (T_x) + 2\ COV\ (T_x,e) + Var\ (e)$

$= Var\ (T_x) + Var\ (e)$

T



\+

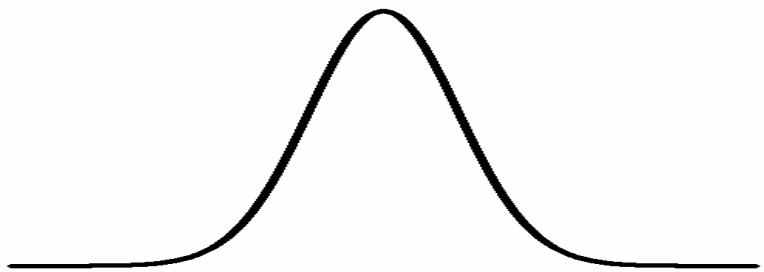
\+

e

=

=

X



Reliability is the Consistency of Measurement

1. The correlation between parallel measures

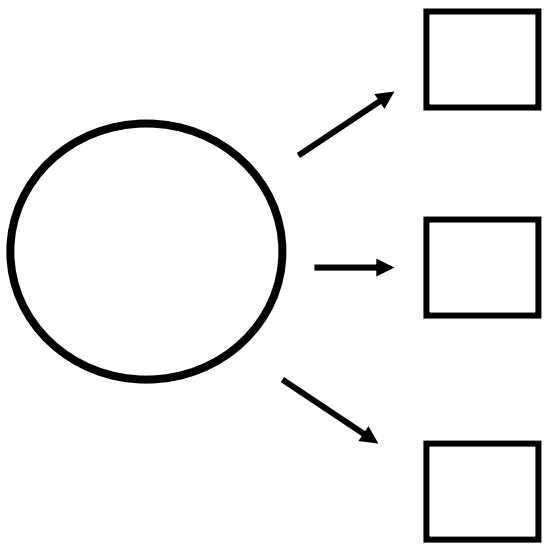
$$\rho_{xx} = r_{x1x2}$$

2. The ratio of True score to Total score variance

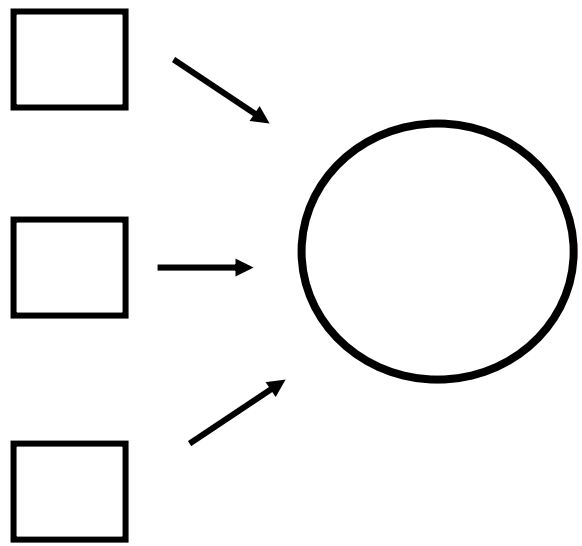
$$\rho_{xx} = \frac{V(T_x)}{V(O_x)} = \frac{V(O_x) - V(e_x)}{V(O_x)}$$

Scale versus Index

Scale



Index



Scale versus Index

1. Scales are often unidimensional
2. Internal consistency not expected for index
3. Examples of Scales
    a. Distress
    b. Self-esteem
    c. Attitudes toward Abortion
4. Examples of Indices
    a. Life event scales
    b. Socioeconomic status
    c. Gross Domestic Product

Modalities of Measurement

1. Clinical Rating
2. Examination
3. Self-report-- Structured Interview
4. Telephone Interview
5. Computer-assisted Interview
6. Paper and pencil
7. Informant Interview