Statistics for Dummies

- from scales to test -

Matthias Rauterberg

Eindhoven University of Technology

2013

How to measure? Which scales exist?

S_x^c	**S**cale name:	mathematical operations
N_x^c	**N**ominal scale:	**= , ≠**
O_x	**O**rdinal scale:	**= , ≠ , > , <**
I_x	**I**nterval scale:	**= , ≠ , > , < , – , +**
R_x	**R**ational scale:	**= , ≠ , > , < , – , + , × , ÷**

S_x^c **S**=scale name [N, O, I, R]; x=number of this scale >1; c=number of categories for N-scale

See further at http://en.wikipedia.org/wiki/Scale_(measurement)
ftp://ftp.sas.com/pub/neural/measurement.html

An easy example: measure of central tendency

- A way of summarising the data using a single value that is in some way representative of the entire data set
  - It is not always possible to follow the same procedure in producing a central representative value: this changes with the shape of the distribution

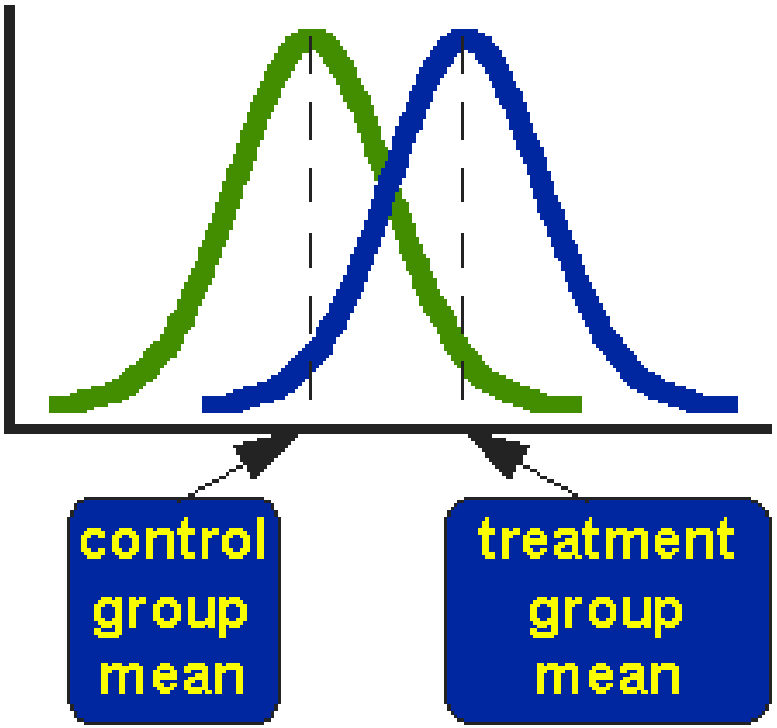
- **Mode** [recommended for **N-scale**]
  - Most frequent value
  - Does not take into account exact scores
  - Unaffected by extreme scores
  - Not useful when there are several values that occur equally often in a set

Measures of central tendency (cont'd)

- **Median** [recommended for **O-scale**]
  - The values that falls exactly in the midpoint of a ranked distribution
  - Does not take into account exact scores
  - Unaffected by extreme scores
  - In a small set it can be unrepresentative

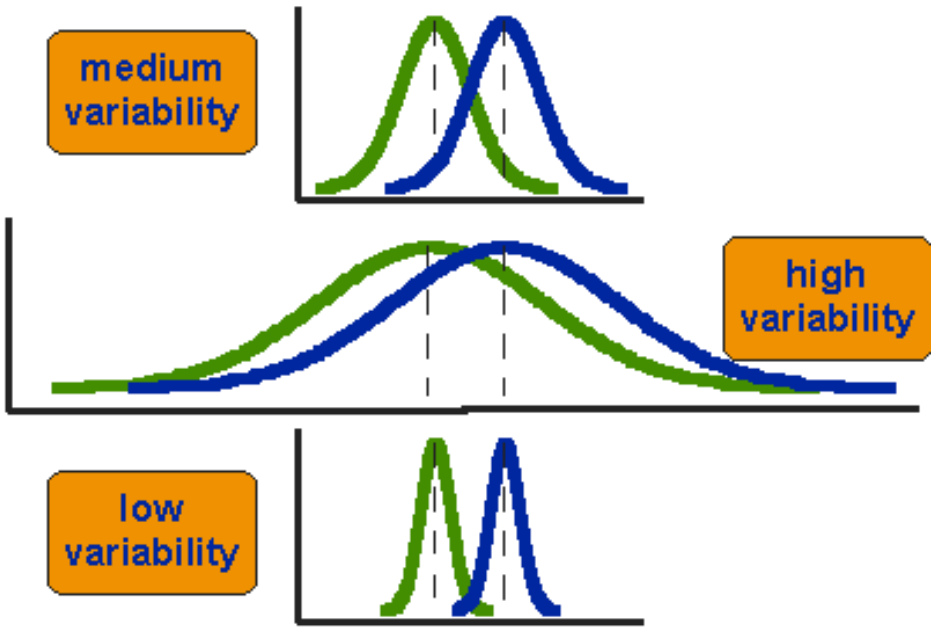
- **Mean** (Arithmetic average) [recommended for **I-scale**]
  - Sample mean: $M = \frac{\Sigma X}{n}$ Population mean: $\mu = \frac{\Sigma X}{N}$
  - Takes into account all values
  - Easily distorted by extreme values

Differences in means for N_1^2-scale and I_1-scale



- In order to know whether a difference between two means is important, we need to know how much the scores vary around the means.

Differences in means for N_1^2-scale and I_1-scale (cont'd)



- Holding the difference between the means constant
- With high variability the two groups nearly overlap
- With low variability the two groups show very little overlap

Scale combinations leads to inference methods

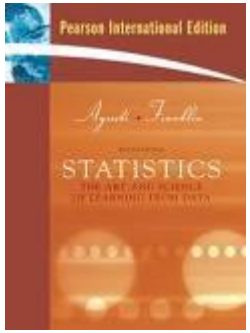
Scales	**Appropriate Inference Method**
$N^2 * N^2$	Fisher's exact test; Odds Ratio
$N^c * N^d$	CHI^2 (with c>2 and/or d>2)
$N^2 * O$	Mann-Whitney-U-test
$N^2 * I$	T-test
$N_x{}^c * I$	[M]Anova (with x>1 and/or c>2)
$I_x * N^c$	Discriminant analysis (with x>1 and c>1)
$O * O$	Spearman/Kendall rank correlation
$I * I$	Pearson correlation
N_x	Cluster analysis (with x>2)
O_x	Multidimensional scaling (with x>2)
I_x	Factor analysis (with x>2)

Choosing a significance level

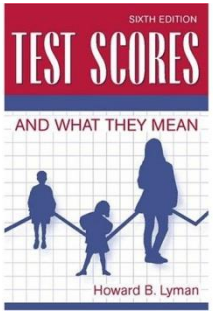
- In general
  - Pilot program and intervention evaluations use liberal significance levels (.2 - .1) to avoid discarding effective interventions.
  - Generally accepted is a significance level of .05
  - Pure research uses conservative significance levels (.01-.001) to avoid wide dissemination of erroneous results.

References

Research Methodology: A Step-By-Step for Beginners.
by Ranjit Kumar (Author)
Paperback: 332 pagina's
Publisher: Sage Publications; 2de Edition; mei 2005
ISBN10 141291194X
ISBN13 9781412911948



Statistics: The Art and Science of Learning from Data.
by Alan Agresti (Author), Chris Franklin (Author)
Softcover: 848 pages
Publisher: Prentice Hall; 2 edition (2009)
Language: English
ISBN-10: 0-13-135746-8
ISBN-13: 978-0-13-135746-4



Test Scores and What They Mean.
by Howard B. Lyman (Author)
Paperback: 190 pages
Publisher: Allyn & Bacon; 6 edition (November 6, 1997)
Language: English
ISBN-10: 0205175392
ISBN-13: 978-0205175390