### How to measure? Which scales exist?

<table>
<thead>
<tr>
<th>Scale name</th>
<th>Mathematical operations</th>
</tr>
</thead>
<tbody>
<tr>
<td>N&lt;sub&gt;x&lt;/sub&gt;&lt;sup&gt;c&lt;/sup&gt;</td>
<td>Nominal scale: =, ≠</td>
</tr>
<tr>
<td>O&lt;sub&gt;x&lt;/sub&gt;</td>
<td>Ordinal scale: =, ≠, &gt;, &lt;</td>
</tr>
<tr>
<td>I&lt;sub&gt;x&lt;/sub&gt;</td>
<td>Interval scale: =, ≠, &gt;, &lt;, -, +</td>
</tr>
<tr>
<td>R&lt;sub&gt;x&lt;/sub&gt;</td>
<td>Rational scale: =, ≠, &gt;, &lt;, -, +, ×, ÷</td>
</tr>
</tbody>
</table>

**S<sub>x</sub><sup>c</sup>** = scale name [N, O, I, R]; x = number of this scale >1; c = number of categories for N-scale

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: $M = \frac{\sum X}{n}$
  - Population mean: $\mu = \frac{\sum X}{N}$
  - Takes into account all values
  - Easily distorted by extreme values
Differences in means for $N_2$-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

<table>
<thead>
<tr>
<th>Scales</th>
<th>Appropriate Inference Method</th>
</tr>
</thead>
<tbody>
<tr>
<td>$N^2 \times N^2$</td>
<td>Fisher’s exact test; Odds Ratio</td>
</tr>
<tr>
<td>$N^c \times N^d$</td>
<td>$\text{CHI}^2$ (with $c&gt;2$ and/or $d&gt;2$)</td>
</tr>
<tr>
<td>$N^2 \times O$</td>
<td>Mann-Whitney-U-test</td>
</tr>
<tr>
<td>$N^2 \times I$</td>
<td>T-test</td>
</tr>
<tr>
<td>$N_x^c \times I$</td>
<td>$[M]\text{Anova}$ (with $x&gt;1$ and/or $c&gt;2$)</td>
</tr>
<tr>
<td>$I_x \times N^c$</td>
<td>Discriminant analysis (with $x&gt;1$ and $c&gt;1$)</td>
</tr>
<tr>
<td>$O \times O$</td>
<td>Spearman/Kendall rank correlation</td>
</tr>
<tr>
<td>$I \times I$</td>
<td>Pearson correlation</td>
</tr>
<tr>
<td>$N_x$</td>
<td>Cluster analysis (with $x&gt;2$)</td>
</tr>
<tr>
<td>$O_x$</td>
<td>Multidimensional scaling (with $x&gt;2$)</td>
</tr>
<tr>
<td>$I_x$</td>
<td>Factor analysis (with $x&gt;2$)</td>
</tr>
</tbody>
</table>
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

by Alan Agresti (Author), Chris Franklin (Author)
Softcover: 848 pages
Publisher: Prentice Hall; 2 edition (2009)
Language: English
ISBN-10: 0-13-135746-8

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