An introduction to Psychometric Theory
Theory of Data, Issues in Scaling

William Revelle

Department of Psychology
Northwestern University
Evanston, Illinois USA

April, 2023
Outline

Science as Model fitting
   Data and scaling
      Assigning Numbers to Observations
Coomb’s Theory of Data
Ordering people,
   Fits
      Proximity rather than order
Ordering objects
   Importance of non-linearity
Thurstonian scaling
   Thurstonian Scaling
MDS
Unfolding
   Preferential Choice
Types of scales and how to describe data
   Describing data graphically
   Central Tendency
More scaling examples
Data = Model + Residual

• The fundamental equations of statistics are that
  • Data = Model + Residual
  • Residual = Data - Model
• The problem is to specify the model and then evaluate the fit of the model to the data as compared to other models
  • Fit = f(Data, Residual)
  • Typically: \( \text{Fit} = f \left( 1 - \frac{\text{Residual}^2}{\text{Data}^2} \right) \)
  • \( \text{Fit} = f \left( 1 - \frac{(\text{Data} - \text{Model})^2}{\text{Data}^2} \right) \)
• Even for something as simple as the mean is a model of the data. The residual left over after we remove the mean is the variance.
• This is a course in developing, evaluating, and comparing models of data.
Psychometrics as model estimation and model fitting

We will explore a number of models

1. Modeling the process of data collection and of scaling
   • $X = f(\theta)$
   • How to measure $X$, properties of the function $f$.

2. Correlation and Regression
   • $Y = \beta X$
   • $R_{xy} = \frac{\sigma_{xy}}{\sigma_x \sigma_y}$

3. Factor Analysis and Principal Components Analysis
   • $R = FF' + U^2$  $R = CC'$

4. Reliability $\rho_{xx} = \frac{\sigma_\theta^2}{\sigma_X^2}$

5. Item Response Theory
   • $p(X|\theta, \delta) = f(\theta - \delta)$

6. Structural Equation Modeling
   • $\rho_{yy} Y = \beta \rho_{xx} X$
A theory of data and fundamentals of scaling

\[
\begin{align*}
\chi_1 & \quad \chi_2 & \quad \chi_3 \\
\delta_1 & \rightarrow & \rightarrow & \rightarrow & \rightarrow \\
\delta_2 & \rightarrow & \rightarrow & \rightarrow & \rightarrow \\
\delta_3 & \rightarrow & \rightarrow & \rightarrow & \rightarrow \\
\delta_4 & \rightarrow & \rightarrow & \rightarrow & \rightarrow \\
\delta_5 & \rightarrow & \rightarrow & \rightarrow & \rightarrow \\
\delta_6 & \rightarrow & \rightarrow & \rightarrow & \rightarrow \\
\delta_7 & \rightarrow & \rightarrow & \rightarrow & \rightarrow \\
\delta_8 & \rightarrow & \rightarrow & \rightarrow & \rightarrow \\
\delta_9 & \rightarrow & \rightarrow & \rightarrow & \rightarrow \\
\end{align*}
\]

Error: \[\delta_1, \delta_2, \delta_3, \delta_4, \delta_5, \delta_6, \delta_7, \delta_8, \delta_9\]

Latent X: \[\chi_1, \chi_2, \chi_3\]

Latent Y: \[\xi_1, \xi_2\]

Y: \[Y_1, Y_2, Y_3, Y_4, Y_5, Y_6, Y_7, Y_8\]

Error: \[\epsilon_1, \epsilon_2, \epsilon_3, \epsilon_4, \epsilon_5, \epsilon_6, \epsilon_7, \epsilon_8\]
Consider the following numbers, what do they represent?

Table: Numbers without context are meaningless. What do these number represent? Which of these numbers represent the same thing?

<table>
<thead>
<tr>
<th>Number</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>2.71828</td>
<td>Natural logarithm</td>
</tr>
<tr>
<td>3.14159</td>
<td>Pi</td>
</tr>
<tr>
<td>24</td>
<td></td>
</tr>
<tr>
<td>86,400</td>
<td></td>
</tr>
<tr>
<td>37</td>
<td></td>
</tr>
<tr>
<td>98.7</td>
<td></td>
</tr>
<tr>
<td>365.25</td>
<td>Days in a year</td>
</tr>
<tr>
<td>365.2563</td>
<td></td>
</tr>
<tr>
<td>31,557,600</td>
<td></td>
</tr>
<tr>
<td>31,558,150</td>
<td></td>
</tr>
<tr>
<td>3,412.14</td>
<td></td>
</tr>
<tr>
<td>3,412.1416</td>
<td></td>
</tr>
<tr>
<td>299,792,458</td>
<td></td>
</tr>
<tr>
<td>6.022141 * 10^23</td>
<td></td>
</tr>
<tr>
<td>42</td>
<td></td>
</tr>
</tbody>
</table>

See also Thorndike (1904) for an amazing introduction to the problem of numbers and measurement.
Clyde Coombs and the Theory of Data

1. O = the set of objects
   - \( O = \{o_i, o_j \ldots o_n\} \)

2. S = the set of Individuals
   - \( S = \{s_i, s_j \ldots s_n\} \)

3. Two comparison operations
   - order ( \( x > y \) )
   - proximity ( \(|x - y| < \epsilon\) )

4. Two types of comparisons
   - Single dyads
     - \((s_i, s_j) (s_i, o_j) (o_i, o_j)\)
   - Pairs of dyads
     - \((s_i, s_j)(s_k, s_l) (s_i, o_j)(s_k, o_l) (o_i, o_j)(o_k, o_l)\)

Coombs (1964)
2 types of comparisons: Monotone ordering and single peak proximity
### Theory of Data and types of measures

**Table:** The theory of data provides a $3 \times 2 \times 2$ taxonomy for various types of measures

<table>
<thead>
<tr>
<th>Elements of Dyad</th>
<th>Number of Dyads</th>
<th>Comparison</th>
<th>Name</th>
</tr>
</thead>
<tbody>
<tr>
<td>People x People</td>
<td>1</td>
<td>Order</td>
<td>Tournament rankings</td>
</tr>
<tr>
<td>People x People</td>
<td>1</td>
<td>Proximity</td>
<td>Social Networks</td>
</tr>
<tr>
<td>Objects x Objects</td>
<td>1</td>
<td>Order</td>
<td>Scaling</td>
</tr>
<tr>
<td>Objects x Objects</td>
<td>1</td>
<td>Proximity</td>
<td>Similarities</td>
</tr>
<tr>
<td>People x Objects</td>
<td>1</td>
<td>Order</td>
<td>Ability Measurement</td>
</tr>
<tr>
<td>People x Objects</td>
<td>1</td>
<td>Proximity</td>
<td>Attitude Measurement</td>
</tr>
<tr>
<td>People x People</td>
<td>2</td>
<td>Order</td>
<td>Tournament rankings</td>
</tr>
<tr>
<td>People x People</td>
<td>2</td>
<td>Proximity</td>
<td>Social Networks</td>
</tr>
<tr>
<td>Objects x Objects</td>
<td>2</td>
<td>Order</td>
<td>Scaling</td>
</tr>
<tr>
<td>Objects x Objects</td>
<td>2</td>
<td>Proximity</td>
<td>Multidimensional scaling</td>
</tr>
<tr>
<td>People x Objects</td>
<td>2</td>
<td>Order</td>
<td>Ability Comparisons</td>
</tr>
<tr>
<td>People x Objects</td>
<td>2</td>
<td>Proximity</td>
<td>Preferential Choice</td>
</tr>
<tr>
<td>People x Objects x Objects</td>
<td>2</td>
<td>Proximity</td>
<td>Individual Differences in Multidimensional Scaling</td>
</tr>
</tbody>
</table>
Tournaments to order people (or teams)

1. Goal is to order the players by outcome to predict future outcomes

2. Complete Round Robin comparisons
   - Everyone plays everyone
   - Requires $N \times (N - 1)/2$ matches
   - How do you scale the results?

3. Partial Tournaments – Seeding and group play
   - World Cup
   - NCAA basketball
   - Is the winner really the best?
   - Can you predict other matches
Simulating a hypothetical chess game

1. Set the random seed to get the same results
2. Generate a sequence of latent values
3. Find the matrix sum of a column vector and row vector
4. Convert to probabilities (using a logit model)
5. Convert probabilities to outcomes
6. Show the results

```r
> set.seed(42)

> p <- seq(-1.5, 1.5, 0.2)
> n <- length(p)

> pdif <- -p %+% t( p )

> prob <- 1/(1 + exp(pdif))

> match <- matrix(rbinom(n*n,1,prob),n,n)

> tournament <- t(upper.tri(match) * (1-match)) + upper.tri(match)*match

> colnames (tournament) <- rownames(match) <- paste("P",1:n,sep="")

> diag(tournament) <- NA
> tournament
```
A hypothetical chess tournament

**Table:** Simulated wins and losses for 16 chess players. Entries reflect row beating column. Thus, P1 wins 4 matches, while P16 wins 14 matches.

<table>
<thead>
<tr>
<th></th>
<th>P1</th>
<th>P2</th>
<th>P3</th>
<th>P4</th>
<th>P5</th>
<th>P6</th>
<th>P7</th>
<th>P8</th>
<th>P9</th>
<th>P10</th>
<th>P11</th>
<th>P12</th>
<th>P13</th>
<th>P14</th>
<th>P15</th>
<th>P16</th>
</tr>
</thead>
<tbody>
<tr>
<td>P1</td>
<td>NA</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>P2</td>
<td>0</td>
<td>NA</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>P3</td>
<td>1</td>
<td>0</td>
<td>NA</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>P4</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>NA</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>P5</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>NA</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>P6</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>NA</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>P7</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>NA</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>P8</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>NA</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>P9</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>NA</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>P10</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>NA</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>P11</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>NA</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>P12</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>NA</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>P13</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>NA</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>P14</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>NA</td>
<td>1</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>P15</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>NA</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>P16</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>NA</td>
</tr>
</tbody>
</table>
The problem: How to scale the players

1. We want to assign numeric values to each player
2. What is best way to map from the values to the data?
3. How well do these values recreate the data?
4. Although players ranks can vary infinitely, pairwise competitions always are between 0 and 1
5. What kind of ranking can we use, what kind of choice model?
Multiple ways of ordering the results

1. Find the mean for each row
2. Express these as normal deviates
3. Express means as logit units
4. Organize all three and the original latent score into a data frame
5. Show the results
6. Graph the results

Note: these numbers have been revised because they were not correcting for the NAs
All three methods match the latent pretty well well
So why bother with normal or logistic modeling, why not just use total score?

1. How to predict wins and losses from the prior scores
   - What is the likelihood that player P16 will beat player 1 if they play again? We need some mapping function from scale to model of the data
2. \( P(A > B) = f(A - B) \) But what is the function?
   - Must map unlimited A and B into 0-1 space
3. Several classic rules
   - Bradley - Terry - Luce Choice rule
     \[ p(A > B | A, B) = \frac{p(A)}{p(A) + p(B)}. \] (1)
   - Thurston Normal deviation model
     \[ p(A > B | A, B) = \text{pnorm}(z_A - z_B) \] (2)
   - Elo/Rasch logistic model where \( \text{logit}_A = \log(p_A/(1 - p_A)) \)
     \[ p(A > B | A, B) = \frac{1}{1 + e^{(\text{logit}_B - \text{logit}_A)}} \] (3)

(Bradley & Terry, 1952; Luce, 1959, 1977; Joe, 1991; Elo, 1978; Elo, 1961)
How well do these various models fit the data?

1. Generate the model of wins and losses
   - Compare to the data
   - Find the residuals
   - Summarize these results

2. Can do it “by hand”
   - Take the scale model, model the data
   - Find residuals
   - Find a goodness of fit

3. Can use a *psych* function: `scaling.fits` to find the fit
   - Although we don’t need to know how the function works, it is possible to find out by using just the function name
   - To find out how to call a function, `?function`, e.g., `?scaling.fits`
   - To run a function, just say `function()` e.g. `scaling.fits(model, data)`
Bradly - Terry - Luce model based upon scores

\[
> \text{round(score}, 4) \\
\begin{array}{cccccccccccccccc}
0.2667 & 0.1333 & 0.2000 & 0.2667 & 0.3333 & 0.3333 & 0.4667 & 0.6000 & 0.4667 & 0.5333 & 0.8000 & 0.5333 & 0.6000 & 0.8000 & 0.7333 & 0.9333 \\
\end{array}
\]

\[
> btl <- \text{score}/(score + t(score)) \\
\]

\[
\text{round(btl, 2)} \\
\begin{array}{cccccccccccccccc}
0.50 & 0.67 & 0.57 & 0.50 & 0.44 & 0.44 & 0.36 & 0.31 & 0.36 & 0.33 & 0.25 & 0.33 & 0.25 & 0.27 & 0.22 \\
P2 & 0.33 & 0.50 & 0.40 & 0.33 & 0.29 & 0.29 & 0.22 & 0.18 & 0.22 & 0.20 & 0.14 & 0.20 & 0.18 & 0.14 & 0.15 & 0.12 \\
P3 & 0.43 & 0.60 & 0.50 & 0.43 & 0.38 & 0.38 & 0.30 & 0.25 & 0.30 & 0.27 & 0.20 & 0.27 & 0.25 & 0.20 & 0.21 & 0.18 \\
P4 & 0.50 & 0.67 & 0.57 & 0.50 & 0.44 & 0.44 & 0.36 & 0.31 & 0.36 & 0.33 & 0.25 & 0.33 & 0.25 & 0.27 & 0.22 \\
P5 & 0.56 & 0.71 & 0.62 & 0.56 & 0.50 & 0.50 & 0.42 & 0.36 & 0.42 & 0.38 & 0.29 & 0.38 & 0.36 & 0.29 & 0.31 & 0.26 \\
P6 & 0.56 & 0.71 & 0.62 & 0.56 & 0.50 & 0.42 & 0.36 & 0.42 & 0.38 & 0.29 & 0.38 & 0.36 & 0.29 & 0.31 & 0.26 \\
P7 & 0.64 & 0.78 & 0.70 & 0.64 & 0.58 & 0.58 & 0.50 & 0.44 & 0.50 & 0.47 & 0.37 & 0.47 & 0.44 & 0.37 & 0.39 & 0.33 \\
P8 & 0.69 & 0.82 & 0.75 & 0.69 & 0.64 & 0.64 & 0.56 & 0.50 & 0.56 & 0.53 & 0.43 & 0.53 & 0.50 & 0.43 & 0.45 & 0.39 \\
P9 & 0.64 & 0.78 & 0.70 & 0.64 & 0.58 & 0.58 & 0.50 & 0.44 & 0.50 & 0.47 & 0.37 & 0.47 & 0.44 & 0.37 & 0.39 & 0.33 \\
P10 & 0.67 & 0.80 & 0.73 & 0.67 & 0.62 & 0.62 & 0.53 & 0.47 & 0.53 & 0.50 & 0.40 & 0.50 & 0.47 & 0.40 & 0.42 & 0.36 \\
P11 & 0.75 & 0.86 & 0.80 & 0.75 & 0.71 & 0.71 & 0.63 & 0.57 & 0.63 & 0.60 & 0.50 & 0.60 & 0.57 & 0.50 & 0.52 & 0.46 \\
P12 & 0.67 & 0.80 & 0.73 & 0.67 & 0.62 & 0.62 & 0.53 & 0.47 & 0.53 & 0.50 & 0.40 & 0.50 & 0.47 & 0.40 & 0.42 & 0.36 \\
P13 & 0.69 & 0.82 & 0.75 & 0.69 & 0.64 & 0.64 & 0.56 & 0.50 & 0.56 & 0.53 & 0.43 & 0.53 & 0.50 & 0.43 & 0.45 & 0.39 \\
P14 & 0.75 & 0.86 & 0.80 & 0.75 & 0.71 & 0.71 & 0.63 & 0.57 & 0.63 & 0.60 & 0.50 & 0.60 & 0.57 & 0.50 & 0.52 & 0.46 \\
P15 & 0.73 & 0.85 & 0.79 & 0.73 & 0.69 & 0.69 & 0.61 & 0.55 & 0.61 & 0.58 & 0.48 & 0.58 & 0.55 & 0.48 & 0.50 & 0.44 \\
P16 & 0.78 & 0.88 & 0.82 & 0.78 & 0.74 & 0.74 & 0.67 & 0.61 & 0.67 & 0.64 & 0.54 & 0.64 & 0.61 & 0.54 & 0.56 & 0.50 \\
(\text{Bradley & Terry, 1952; Luce, 1959, 1977})
\]

\[
p(A > B|A, B) = \frac{p(A)}{p(A) + p(B)}. \tag{4}
\]
BTL Residuals are data - model

```r
> resid <- tournament - btl

> round(resid, 2)

P1  P2  P3  P4  P5  P6  P7  P8  P9  P10 P11 P12 P13 P14 P15 P16
P1  NA  0.33 -0.57  0.50  0.56 -0.44 -0.36 -0.31 -0.36 -0.33  0.75 -0.33 -0.25 -0.27 -0.22
P2 -0.33 NA  0.60  0.67 -0.29 -0.29 -0.22 -0.18 -0.22 -0.20 -0.14 -0.20 -0.18 -0.14 -0.15 -0.12
P3  0.57 -0.60 NA -0.43 -0.38 -0.38  0.70 -0.25 -0.27 -0.20 -0.27 -0.25 -0.20 -0.21 -0.18
P4 -0.50 -0.67  0.43 NA  0.56  0.56 -0.36 -0.31 -0.36  0.67 -0.25 -0.33 -0.31 -0.25 -0.27 -0.22
P5 -0.56  0.29  0.38 -0.56 NA  0.50  0.58 -0.36 -0.42 -0.38 -0.29  0.62 -0.36 -0.29 -0.31 -0.26
P6  0.44  0.29  0.38 -0.56 -0.50 NA -0.42 -0.36  0.58 -0.38 -0.29  0.62 -0.36 -0.29 -0.31 -0.26
P7  0.36  0.22 -0.70  0.36 -0.58  0.42 NA  0.56  0.50  0.53 -0.37 -0.47 -0.44 -0.37 -0.39 -0.33
P8  0.31  0.18  0.25  0.31  0.36  0.36 -0.56 NA  0.44 -0.53 -0.43 -0.53  0.50  0.57 -0.45 -0.39
P9  0.36  0.22 -0.70  0.36 -0.58 -0.50 -0.44 NA  0.53 -0.37  0.53  0.56 -0.37 -0.39 -0.33
P10 0.33  0.20  0.27 -0.67  0.38 -0.38 -0.53  0.53 -0.53 NA  0.40  0.50 -0.47 -0.40 -0.42  0.64
P11 -0.75  0.14  0.20  0.25  0.29  0.29  0.37  0.43  0.37  0.40 NA  0.40  0.43 -0.50  0.48 -0.46
P12 0.33  0.20  0.27  0.33 -0.62 -0.62  0.47  0.53 -0.53 -0.50 -0.40 NA  0.47  0.60  0.58 -0.36
P13 0.31  0.18  0.25  0.31  0.36  0.36  0.44 -0.50 -0.56 -0.47 -0.43 -0.47 NA  0.43 -0.45 -0.39
P14 0.25  0.14  0.20  0.25  0.29  0.29  0.37 -0.57  0.37  0.40  0.50 -0.60  0.43 NA  0.48 -0.46
P15 0.27  0.15  0.21  0.27  0.31  0.31  0.39  0.45  0.39  0.42 -0.48 -0.58  0.45 -0.48 NA -0.44
P16 0.22  0.12  0.18  0.22  0.26  0.26  0.33  0.39  0.33 -0.64  0.46  0.36  0.39  0.46  0.44 NA
```
Find Goodness of Fit “by hand”

1. Find model
2. Find Residual = Model - Data
3. Goodness of Fit is 
   \[1 - \frac{\text{Residual}^2}{\text{Data}^2}\]

Use a simple function:

```r
scaling.fits(score, tournament, test = "choice", rowwise = FALSE)
```

```
$GF
[1] 0.651827

$original
[1] 120

$resid
[1] 41.78075
```
Automate it by calling a function (scaling.fits) repeatedly for alternative models

These data may be analyzed using repeated calls to the scaling.fits function:

```r
> tests <- c("choice","logit","normal")
> fits <- matrix(NA, ncol = 3,nrow=4)
> for (i in 1:4) {
+ for (j in 1:3) {
+ fits[i, j] <- scaling.fits(chess.df[i], data = tournament,
+ test = tests[j],rowwise=FALSE)$GF[1] } }
> rownames(fits) <- c("latent","observed", "normed", "logistic")
> colnames(fits) <- c("choice", "logistic", "normal")
> round(fits, 2)
```

<table>
<thead>
<tr>
<th></th>
<th>choice</th>
<th>logistic</th>
<th>normal</th>
</tr>
</thead>
<tbody>
<tr>
<td>latent</td>
<td>0.63</td>
<td>0.67</td>
<td>0.65</td>
</tr>
<tr>
<td>observed</td>
<td>0.65</td>
<td>0.59</td>
<td>0.63</td>
</tr>
<tr>
<td>normed</td>
<td>0.66</td>
<td>0.68</td>
<td>0.70</td>
</tr>
<tr>
<td>logistic</td>
<td>0.66</td>
<td>0.70</td>
<td>0.70</td>
</tr>
</tbody>
</table>

Note how the scaled data fit the observed choices better than the actual observed orders fit.
Advanced: The scaling.fits function

```r
> scaling.fits <-
function (model, data, test = "logit", digits = 2, rowwise = TRUE) {
  model <- as.matrix(model)
  data <- as.matrix(data)
  if (test == "choice") {
    model <- as.vector(model)
    if (min(model) <= 0)
      model <- model - min(model)
    prob = model/(model %+% t(model))
  } else {
    pdif <- model %+% -t(model)
    if (test == "logit") {
      prob <- 1/(1 + exp(-pdif))
    } else {
      if (test == "normal") {
        prob <- pnorm(pdif)
      }
    }
  }
  if (rowwise) {
    prob = 1 - prob
  }
  error <- data - prob
  sum.error2 <- sum(error^2, na.rm = TRUE)
  sum.data2 <- sum(data^2, na.rm = TRUE)
  gof <- 1 - sum.error2/sum.data2
  fit <- list(GF = gof, original = sum.data2, resid = sum.error2,
              residual = round(error, digits))
  return(fit)
}
```
**Friendship as proximity**

1. Chess or football provides a ranking based upon an ordering relationship \( p_i > p_j \).

2. Alternatively, friendship groups are based upon closeness \( |p_i - p_j| < \delta \)
   - 2.1 Do you know person j?
   - 2.2 Do you like person j? or as an alternative:
   - 2.3 Please list all your friends in this class (and is j included on the list)
   - 2.4 Would you be interested in having a date with person j?
   - 2.5 Would you like to have sex with person j?
   - 2.6 Would you marry person j?

3. Typically such data will be a rectangular matrix for there are asymmetries in closeness.
Moh’s hardness scale provides rank orders of hardness

How hard is rock? The scratch test.

Table: Mohs’ scale of mineral hardness. An object is said to be harder than X if it scratches X. Also included are measures of relative hardness using a sclerometer (for the hardest of the planes if there is a anisotropy or variation between the planes) which shows the non-linearity of the Mohs scale (Burchard, 2004).

<table>
<thead>
<tr>
<th>Mohs Hardness</th>
<th>Mineral</th>
<th>Scratch hardness</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Talc</td>
<td>.59</td>
</tr>
<tr>
<td>2</td>
<td>Gypsum</td>
<td>.61</td>
</tr>
<tr>
<td>3</td>
<td>Calcite</td>
<td>3.44</td>
</tr>
<tr>
<td>4</td>
<td>Fluorite</td>
<td>3.05</td>
</tr>
<tr>
<td>5</td>
<td>Apapatite</td>
<td>5.2</td>
</tr>
<tr>
<td>6</td>
<td>Orthoclase Feldspar</td>
<td>37.2</td>
</tr>
<tr>
<td>7</td>
<td>Quartz</td>
<td>100</td>
</tr>
<tr>
<td>8</td>
<td>Topaz</td>
<td>121</td>
</tr>
<tr>
<td>9</td>
<td>Corundum</td>
<td>949</td>
</tr>
<tr>
<td>10</td>
<td>Diamond</td>
<td>85,300</td>
</tr>
</tbody>
</table>
Measuring Hardness – Scratch versus Mohs

![Graph depicting the comparison between Scratch Hardness and Mohs Hardness.](image)

The graphs show the relationship between Scratch Hardness and Mohs Hardness. The left graph plots Scratch Hardness against Mohs Hardness, while the right graph plots the logarithm of Scratch Hardness against Mohs Hardness. The data points are marked with circles, and the lines connect these points to illustrate the trend. The Mohs scale is indicated on the x-axis, ranging from 1 to 10, and the hardness values are on the y-axis. The logarithmic scale is used for the Scratch Hardness to better visualize the relationship across the Mohs range.
Why we should report effect sizes rather than $p$: Another example of non-linearity

Probaability is a bad estimate of effect size

Effect size

Probability
Ordering based upon external measures

Table: The Beaufort scale of wind intensity is an early example of a scale with roughly equal units that is observationally based. Although the units are roughly in equal steps of wind speed in nautical miles/hour (knots), the force of the wind is not linear with this scale, but rather varies as the square of the velocity.

<table>
<thead>
<tr>
<th>Force</th>
<th>Wind (Knots)</th>
<th>WMO Classification</th>
<th>Appearance of Wind Effects</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>Less than 1</td>
<td>Calm</td>
<td>Sea surface smooth and mirror-like</td>
</tr>
<tr>
<td>1</td>
<td>1-3</td>
<td>Light Air</td>
<td>Scaly ripples, no foam crests</td>
</tr>
<tr>
<td>2</td>
<td>4-6</td>
<td>Light Breeze</td>
<td>Small wavelets, crests glassy, no breaking</td>
</tr>
<tr>
<td>3</td>
<td>7-10</td>
<td>Gentle Breeze</td>
<td>Large wavelets, crests begin to break, scattered whitecaps</td>
</tr>
<tr>
<td>4</td>
<td>11-16</td>
<td>Moderate Breeze</td>
<td>Small waves 1-4 ft. becoming longer, numerous whitecaps</td>
</tr>
<tr>
<td>5</td>
<td>17-21</td>
<td>Fresh Breeze</td>
<td>Moderate waves 4-8 ft taking longer form, many whitecaps, some spray</td>
</tr>
<tr>
<td>6</td>
<td>22-27</td>
<td>Strong Breeze</td>
<td>Larger waves 8-13 ft, whitecaps common more spray</td>
</tr>
<tr>
<td>7</td>
<td>28-33</td>
<td>Near Gale</td>
<td>Sea heaps up, waves 13-20 ft, white foam streaks off breakers</td>
</tr>
<tr>
<td>8</td>
<td>34-40</td>
<td>Gale Moderately</td>
<td>High (13-20 ft) waves of greater length, edges of crests begin to break into spindrift, foam blown in streaks</td>
</tr>
<tr>
<td>9</td>
<td>41-47</td>
<td>Strong Gale</td>
<td>High waves (20 ft), sea begins to roll, dense streaks of foam, spray may reduce visibility</td>
</tr>
<tr>
<td>10</td>
<td>48-55</td>
<td>Storm</td>
<td>Very high waves (20-30 ft) with overhanging crests, sea white with densely blown foam, heavy rolling, lowered visibility</td>
</tr>
<tr>
<td>11</td>
<td>56-63</td>
<td>Violent Storm</td>
<td>Exceptionally high (30-45 ft) waves, foam patches cover sea, visibility more reduced</td>
</tr>
<tr>
<td>12</td>
<td>64+</td>
<td>Hurricane</td>
<td>Air filled with foam, waves over 45 ft, sea completely white with driving spray, visibility greatly reduced</td>
</tr>
</tbody>
</table>
The Beaufort scale is non-linear with force or probability of capsizing.
Models of scaling objects

1. Assume each object \((a, b, ...z)\) has a scale value \((A, B, ...Z)\) with some noise for each measurement.

2. Probability of \(A > B\) increases with difference between \(a\) and \(b\)

3. \(P(A > B) = f(a - b)\)

4. Can we find a function, \(f\), such that equal differences in the latent variable \((a, b, c)\) lead to equal differences in the observed variable?

5. Several alternatives
   - Direct scaling on some attribute dimension (simple but flawed)
   - Indirect scaling by paired comparisons (more complicated but probably better)
Scaling of Objects: $O \times O$ comparisons

1. Typical object scaling is concerned with order or location of objects

2. Subjects are assumed to be random replicates of each other, differing only as a source of noise

3. Absolute scaling techniques
   - Grant Proposals: 1 to 5
   - “On a scale from 1 to 10” this [object] is a X?
   - If A is 1 and B is 10, then what is C?
   - College rankings based upon selectivity
   - College rankings based upon "yield"
   - Zagat ratings of restaurants
   - A - F grading of papers
Absolute scaling: difficulties

1. "On a scale from 1 to 10" this [object] is a X?
   • sensitive to context effects
   • what if a new object appears?
   • Need unbounded scale

2. If A is 1 and B is 10, then what is C?
   • results will depend upon A, B
Absolute scaling: artifacts

1. College rankings based upon selectivity
   - accept/applied
   - encourage less able to apply

2. College rankings based upon “yield”
   - matriculate/accepted
   - early admissions guarantee matriculation
   - don’t accept students who will not attend

3. Proposed solution: college choice as a tournament
   - Consider all schools that accept a student
   - Which school does he/she choose?

Avery, Glickman, Hoxby & Metrick (2013)
A revealed preference ordering Avery et al. (2013)
A revealed preference ordering Avery et al. (2013)

<table>
<thead>
<tr>
<th>Rank Based on Matriculation (with Covariates)</th>
<th>College Name</th>
<th>Implied Prob. of “Winning” vs. College Listed...</th>
<th>Rank Based on Matriculation (no Covariates)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Harvard University</td>
<td>9.13 0.59 0.93</td>
<td>1</td>
</tr>
<tr>
<td>2</td>
<td>Caltech</td>
<td>8.77 0.56 0.92</td>
<td>3</td>
</tr>
<tr>
<td>3</td>
<td>Yale University</td>
<td>8.52 0.59 0.92</td>
<td>2</td>
</tr>
<tr>
<td>4</td>
<td>MIT</td>
<td>8.16 0.51 0.89</td>
<td>5</td>
</tr>
<tr>
<td>5</td>
<td>Stanford University</td>
<td>8.11 0.52 0.90</td>
<td>4</td>
</tr>
<tr>
<td>6</td>
<td>Princeton University</td>
<td>8.02 0.73 0.90</td>
<td>6</td>
</tr>
<tr>
<td>7</td>
<td>Brown University</td>
<td>7.01 0.56 0.78</td>
<td>7</td>
</tr>
<tr>
<td>8</td>
<td>Columbia University</td>
<td>6.77 0.54 0.73</td>
<td>8</td>
</tr>
<tr>
<td>9</td>
<td>Amherst College</td>
<td>6.61 0.51 0.71</td>
<td>9</td>
</tr>
<tr>
<td>10</td>
<td>Dartmouth</td>
<td>6.57 0.52 0.72</td>
<td>10</td>
</tr>
<tr>
<td>11</td>
<td>Wellesley College</td>
<td>6.51 0.53 0.71</td>
<td>12</td>
</tr>
<tr>
<td>12</td>
<td>University of Pennsylvania</td>
<td>6.39 0.56 0.71</td>
<td>11</td>
</tr>
</tbody>
</table>
Weber-Fechner Law and non-linearity of scales

1. Early studies of psychophysics by Weber (1834b,a) and subsequently Fechner (1860) demonstrated that the human perceptual system does not perceive stimulus intensity as a linear function of the physical input.

2. The basic paradigm was to compare one weight with another that differed by amount $\Delta$, e.g., compare a 10 gram weight with an 11, 12, and 13 gram weight, or a 10 kg weight with a 11, 12, or 13 kg weight.

3. What was the $\Delta$ that was just detectable? The finding was that the perceived intensity follows a logarithmic function.

4. Examining the magnitude of the “just noticeable difference” or $JND$, Weber (1834b) found that

$$JND = \frac{\Delta \text{Intensity}}{\text{Intensity}} = \text{constant}. \quad (5)$$
Weber-Fechner Law and non-linearity of scales

1. An example of a logarithmic scale of intensity is the decibel measure of sound intensity.

2. Sound Pressure Level expressed in decibels (dB) of the root mean square observed sound pressure, $P_o$ (in Pascals) is

$$L_p = 20\log_{10}\frac{P_o}{P_{ref}} \quad (6)$$

3. where the reference pressure, $P_{ref}$, in the air is $20\mu Pa$.

4. Just to make this confusing, the reference pressure for sound measured in the ocean is $1\mu Pa$. This means that sound intensities in the ocean are expressed in units that are 20 dB higher than those units used on land.
The Just Noticeable Difference in Person and risk perception

1. Although typically thought of as just relevant for the perceptual experiences of physical stimuli, Ozer (1993) suggested that the JND is useful in personality assessment as a way of understanding the accuracy and inter judge agreement of judgments about other people.

2. In addition, Sinn (2003) has argued that the logarithmic nature of the Weber-Fechner Law is of evolutionary significance for preference for risk and cites Bernoulli (1738) as suggesting that our general utility function is logarithmic.

3. The whole of Prospect Theory (Kahneman & Tversky, 1979; Kahneman, 2011) is based upon this non-linearity of utilities: Better to skip lunch than be someone’s dinner.
Money and non linearity

... the utility resulting from any small increase in wealth will be inversely proportionate to the quantity of goods already possessed .... if ... one has a fortune worth a hundred thousand ducats and another one a fortune worth same number of semi-ducats and if the former receives from it a yearly income of five thousand ducats while the latter obtains the same number of semi-ducats, it is quite clear that to the former a ducat has exactly the same significance as a semi-ducat to the latter (Bernoulli, 1738, p 25).

Implies a log function for utility.
Econs and Humans

1. Simple expected value theory \(\Rightarrow\) value = probability of event \(\times\) value of event

2. Bernoulli theory of expected utility came to dominate choice theory and is fundamental to economics

3. Studied by comparing gambles and showing utility is non-linear with value
   - Would you rather have $80 or a 80% chance of $100 + 20% of $10?
   - expected value is 80 versus .8 * 100 + .2 * 10 = 82

4. Bernoulli value (from Kahneman, 2011)

<table>
<thead>
<tr>
<th>Wealth (millions)</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
<th>10</th>
</tr>
</thead>
<tbody>
<tr>
<td>Utility units</td>
<td>10</td>
<td>30</td>
<td>48</td>
<td>60</td>
<td>70</td>
<td>78</td>
<td>84</td>
<td>90</td>
<td>96</td>
<td>100</td>
</tr>
</tbody>
</table>
Kahneman and Tversky: Prospect Theory

Losses are more painful than gains are pleasant

Kahneman & Tversky (1979)
Better to skip lunch than be someone’s dinner.
Thurstonian scaling: basic concept

1. Every object has a value
2. Rated strength of object is noisy with Gaussian noise
3. \( P(A > B) = f(z_a - z_b) \)
4. Assume equal variance for each item
5. Convert choice frequency to normal deviates
6. Scale is average normal deviates
Thurstone choice model

Figure: Thurstone's model of paired discrimination. Left panel: three items differ in their mean level as well as their variance. Right panel: choice between two items with equal variance reflects the relative strength of the two items. The shaded section represents choosing item 2 over item 1.
Thurstone’s Vegetable data as an example of one dimensional scaling

Table: Consider the likelihood of liking a vegetable. Numbers reflect probability that the column is preferred to the row. Can we turn this into a scale?

The veg data set from the psych package in R

<table>
<thead>
<tr>
<th>Variable</th>
<th>Turn</th>
<th>Cab</th>
<th>Beet</th>
<th>Asp</th>
<th>Car</th>
<th>Spin</th>
<th>S.Beans</th>
<th>Peas</th>
<th>Corn</th>
</tr>
</thead>
<tbody>
<tr>
<td>Turn</td>
<td>0.50</td>
<td>0.82</td>
<td>0.77</td>
<td>0.81</td>
<td>0.88</td>
<td>0.89</td>
<td>0.90</td>
<td>0.89</td>
<td>0.93</td>
</tr>
<tr>
<td>Cab</td>
<td>0.18</td>
<td>0.50</td>
<td>0.60</td>
<td>0.72</td>
<td>0.74</td>
<td>0.74</td>
<td>0.81</td>
<td>0.84</td>
<td>0.86</td>
</tr>
<tr>
<td>Beet</td>
<td>0.23</td>
<td>0.40</td>
<td>0.50</td>
<td>0.56</td>
<td>0.74</td>
<td>0.68</td>
<td>0.84</td>
<td>0.80</td>
<td>0.82</td>
</tr>
<tr>
<td>Asp</td>
<td>0.19</td>
<td>0.28</td>
<td>0.44</td>
<td>0.50</td>
<td>0.56</td>
<td>0.59</td>
<td>0.68</td>
<td>0.60</td>
<td>0.73</td>
</tr>
<tr>
<td>Car</td>
<td>0.12</td>
<td>0.26</td>
<td>0.26</td>
<td>0.44</td>
<td>0.50</td>
<td>0.49</td>
<td>0.57</td>
<td>0.71</td>
<td>0.76</td>
</tr>
<tr>
<td>Spin</td>
<td>0.11</td>
<td>0.26</td>
<td>0.32</td>
<td>0.41</td>
<td>0.51</td>
<td>0.50</td>
<td>0.63</td>
<td>0.68</td>
<td>0.63</td>
</tr>
<tr>
<td>S.Beans</td>
<td>0.10</td>
<td>0.19</td>
<td>0.16</td>
<td>0.32</td>
<td>0.43</td>
<td>0.37</td>
<td>0.50</td>
<td>0.53</td>
<td>0.64</td>
</tr>
<tr>
<td>Peas</td>
<td>0.11</td>
<td>0.16</td>
<td>0.20</td>
<td>0.40</td>
<td>0.29</td>
<td>0.32</td>
<td>0.47</td>
<td>0.50</td>
<td>0.63</td>
</tr>
<tr>
<td>Corn</td>
<td>0.07</td>
<td>0.14</td>
<td>0.18</td>
<td>0.27</td>
<td>0.24</td>
<td>0.37</td>
<td>0.36</td>
<td>0.37</td>
<td>0.50</td>
</tr>
</tbody>
</table>

#show the data from the veg data set from the psych package
veg

(Guilford, 1954)
Some simple R

> veg #shows the data

<table>
<thead>
<tr>
<th></th>
<th>Turn</th>
<th>Cab</th>
<th>Beet</th>
<th>Asp</th>
<th>Car</th>
<th>Spin</th>
<th>S.Beans</th>
<th>Peas</th>
<th>Corn</th>
</tr>
</thead>
<tbody>
<tr>
<td>Turn</td>
<td>0.500</td>
<td>0.318</td>
<td>0.770</td>
<td>0.811</td>
<td>0.878</td>
<td>0.892</td>
<td>0.899</td>
<td>0.892</td>
<td>0.926</td>
</tr>
<tr>
<td>Cab</td>
<td>0.182</td>
<td>0.500</td>
<td>0.601</td>
<td>0.723</td>
<td>0.743</td>
<td>0.736</td>
<td>0.811</td>
<td>0.845</td>
<td>0.858</td>
</tr>
<tr>
<td>Beet</td>
<td>0.230</td>
<td>0.399</td>
<td>0.500</td>
<td>0.561</td>
<td>0.736</td>
<td>0.676</td>
<td>0.845</td>
<td>0.797</td>
<td>0.818</td>
</tr>
<tr>
<td>Asp</td>
<td>0.189</td>
<td>0.277</td>
<td>0.439</td>
<td>0.500</td>
<td>0.561</td>
<td>0.588</td>
<td>0.676</td>
<td>0.601</td>
<td>0.730</td>
</tr>
<tr>
<td>Car</td>
<td>0.122</td>
<td>0.257</td>
<td>0.264</td>
<td>0.439</td>
<td>0.500</td>
<td>0.493</td>
<td>0.574</td>
<td>0.709</td>
<td>0.764</td>
</tr>
<tr>
<td>Spin</td>
<td>0.108</td>
<td>0.264</td>
<td>0.324</td>
<td>0.412</td>
<td>0.507</td>
<td>0.500</td>
<td>0.628</td>
<td>0.682</td>
<td>0.628</td>
</tr>
<tr>
<td>S.Beans</td>
<td>0.101</td>
<td>0.189</td>
<td>0.155</td>
<td>0.324</td>
<td>0.426</td>
<td>0.372</td>
<td>0.500</td>
<td>0.527</td>
<td>0.642</td>
</tr>
<tr>
<td>Peas</td>
<td>0.108</td>
<td>0.155</td>
<td>0.203</td>
<td>0.399</td>
<td>0.291</td>
<td>0.318</td>
<td>0.473</td>
<td>0.500</td>
<td>0.628</td>
</tr>
<tr>
<td>Corn</td>
<td>0.074</td>
<td>0.142</td>
<td>0.182</td>
<td>0.270</td>
<td>0.236</td>
<td>0.372</td>
<td>0.358</td>
<td>0.372</td>
<td>0.500</td>
</tr>
</tbody>
</table>

> colMeans(veg) #show the means (but too many decimals)

<table>
<thead>
<tr>
<th></th>
<th>Turn</th>
<th>Cab</th>
<th>Beet</th>
<th>Asp</th>
<th>Car</th>
<th>Spin</th>
<th>S.Beans</th>
<th>Peas</th>
<th>Corn</th>
</tr>
</thead>
<tbody>
<tr>
<td>Turn</td>
<td>0.179</td>
<td>0.333</td>
<td>0.382</td>
<td>0.493</td>
<td>0.542</td>
<td>0.549</td>
<td>0.640</td>
<td>0.658</td>
<td>0.721</td>
</tr>
</tbody>
</table>

> round(colMeans(veg)) #round off, but not enough decimals

<table>
<thead>
<tr>
<th></th>
<th>Turn</th>
<th>Cab</th>
<th>Beet</th>
<th>Asp</th>
<th>Car</th>
<th>Spin</th>
<th>S.Beans</th>
<th>Peas</th>
<th>Corn</th>
</tr>
</thead>
<tbody>
<tr>
<td>Turn</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
</tbody>
</table>

> round(colMeans(veg),2) #this looks pretty good

<table>
<thead>
<tr>
<th></th>
<th>Turn</th>
<th>Cab</th>
<th>Beet</th>
<th>Asp</th>
<th>Car</th>
<th>Spin</th>
<th>S.Beans</th>
<th>Peas</th>
<th>Corn</th>
</tr>
</thead>
<tbody>
<tr>
<td>Turn</td>
<td>0.18</td>
<td>0.33</td>
<td>0.38</td>
<td>0.49</td>
<td>0.54</td>
<td>0.55</td>
<td>0.64</td>
<td>0.66</td>
<td>0.72</td>
</tr>
</tbody>
</table>
Three ways of plotting the data

```r
op <- par(mfrow=c(1,3)) # I want to draw two graphs
plot(colMeans(veg), main="simple plot")  # the basic plot command
dotchart(colMeans(veg), main="A dot chart")  # dot charts are more informative
error.dots(veg, eyes=TRUE, main="Vegetables and Cats Eyes")
```

```r
op <- par(mfrow=c(1,1))  # set the plotting back to a single graph
```
And yet more ways of plotting the data

```
op <- par(mfrow=c(1,2)) # I want to draw two graphs
dotchart(colMeans(veg)) # dot charts are more informative
error.dots(veg) # add error bars to the plot
op <- par(mfrow=c(1,1)) # set the plotting back to a single graph
```
Alternatively, use the `error.bars` function from psych

```r
error.bars(veg, bars=TRUE, ylab="Preference", xlab="Vegetables", main="Mean and 95% confidence intervals")
```
Naive scaling

1. Show the data
2. Find the mean for each column. Round to 2 decimals
3. Subtract the mean for the first column from the means
4. But these are not really useful scale values.
Convert the vegetables data set to normal deviates

```r
> z.veg <- qnorm(as.matrix(veg))
> round(z.veg,2) #see table

<table>
<thead>
<tr>
<th></th>
<th>Turn</th>
<th>Cab</th>
<th>Beet</th>
<th>Asp</th>
<th>Car</th>
<th>Spin</th>
<th>S.Beans</th>
<th>Peas</th>
<th>Corn</th>
</tr>
</thead>
<tbody>
<tr>
<td>Turn</td>
<td>0.00</td>
<td>0.91</td>
<td>0.74</td>
<td>0.88</td>
<td>1.17</td>
<td>1.24</td>
<td>1.28</td>
<td>1.24</td>
<td>1.45</td>
</tr>
<tr>
<td>Cab</td>
<td>-0.91</td>
<td>0.00</td>
<td>0.26</td>
<td>0.59</td>
<td>0.65</td>
<td>0.63</td>
<td>0.88</td>
<td>1.02</td>
<td>1.07</td>
</tr>
<tr>
<td>Beet</td>
<td>-0.74</td>
<td>-0.26</td>
<td>0.00</td>
<td>0.15</td>
<td>0.63</td>
<td>0.46</td>
<td>1.02</td>
<td>0.83</td>
<td>0.91</td>
</tr>
<tr>
<td>Asp</td>
<td>-0.88</td>
<td>-0.59</td>
<td>-0.15</td>
<td>0.00</td>
<td>0.15</td>
<td>0.22</td>
<td>0.46</td>
<td>0.26</td>
<td>0.61</td>
</tr>
<tr>
<td>Car</td>
<td>-1.17</td>
<td>-0.65</td>
<td>-0.63</td>
<td>-0.15</td>
<td>0.00</td>
<td>-0.02</td>
<td>0.19</td>
<td>0.55</td>
<td>0.72</td>
</tr>
<tr>
<td>Spin</td>
<td>-1.24</td>
<td>-0.63</td>
<td>-0.46</td>
<td>-0.22</td>
<td>0.02</td>
<td>0.00</td>
<td>0.33</td>
<td>0.47</td>
<td>0.33</td>
</tr>
<tr>
<td>S.Beans</td>
<td>-1.28</td>
<td>-0.88</td>
<td>-1.02</td>
<td>-0.46</td>
<td>-0.19</td>
<td>-0.33</td>
<td>0.00</td>
<td>0.07</td>
<td>0.36</td>
</tr>
<tr>
<td>Peas</td>
<td>-1.24</td>
<td>-1.02</td>
<td>-0.83</td>
<td>-0.26</td>
<td>-0.55</td>
<td>-0.47</td>
<td>-0.07</td>
<td>0.00</td>
<td>0.33</td>
</tr>
<tr>
<td>Corn</td>
<td>-1.45</td>
<td>-1.07</td>
<td>-0.91</td>
<td>-0.61</td>
<td>-0.72</td>
<td>-0.33</td>
<td>-0.36</td>
<td>-0.33</td>
<td>0.00</td>
</tr>
</tbody>
</table>

> scaled.veg <- colMeans(z.veg)
> round(scaled.veg,2)

<table>
<thead>
<tr>
<th></th>
<th>Turn</th>
<th>Cab</th>
<th>Beet</th>
<th>Asp</th>
<th>Car</th>
<th>Spin</th>
<th>S.Beans</th>
<th>Peas</th>
<th>Corn</th>
</tr>
</thead>
<tbody>
<tr>
<td>Turn</td>
<td>-0.99</td>
<td>-0.47</td>
<td>-0.33</td>
<td>-0.01</td>
<td>0.13</td>
<td>0.16</td>
<td>0.41</td>
<td>0.46</td>
<td>0.64</td>
</tr>
</tbody>
</table>

> scaled <- scaled.veg - min(scaled.veg)
> round(scaled,2)

<table>
<thead>
<tr>
<th></th>
<th>Turn</th>
<th>Cab</th>
<th>Beet</th>
<th>Asp</th>
<th>Car</th>
<th>Spin</th>
<th>S.Beans</th>
<th>Peas</th>
<th>Corn</th>
</tr>
</thead>
<tbody>
<tr>
<td>Turn</td>
<td>0.00</td>
<td>0.52</td>
<td>0.65</td>
<td>0.98</td>
<td>1.12</td>
<td>1.14</td>
<td>1.40</td>
<td>1.44</td>
<td>1.63</td>
</tr>
</tbody>
</table>
```

1. Convert to normal deviates using the `norm` function. But that only works on matrices, so we need to convert the data.frame into a matrix.

2. Display the data

3. Find the column means and show them

4. subtract the smallest value to form a positive scale.
Form the model based upon these scale values

1. Subtract the column value from the row value using the `matrix.addition` function from `psych`.

2. Show the result

3. Convert the normal deviates into probabilities using the `pnorm` function.

```r
pdif <- scaled %+% t(scaled)
colnames(pdif) <- rownames(pdif) <- colnames(z.veg)
round(pdif, 2)

modeled <- pnorm(pdif)
round(modeled, 2)
```
Data = Model + Residual

1. What is the model?
   - Pref = Mean (preference)
   - \( p(A > B) = f(A, B) \)
   - what is f?

2. Possible functions
   - \( f = A - B \) (simple difference)
   - \( \frac{A}{A+B} \) Luce choice rule
   - Thurstonian scaling
   - logistic scaling

3. Evaluating functions – Goodness of fit
   - Residual = Model - Data
   - Minimize residual
   - Minimize residual^2
Examine the residuals

```r
> resid <- veg - modeled
> round(resid,2)

<table>
<thead>
<tr>
<th></th>
<th>Turn</th>
<th>Cab</th>
<th>Beet</th>
<th>Asp</th>
<th>Car</th>
<th>Spin</th>
<th>S.Beans</th>
<th>Peas</th>
<th>Corn</th>
</tr>
</thead>
<tbody>
<tr>
<td>Turn</td>
<td>0.00</td>
<td>0.12</td>
<td>0.03</td>
<td>-0.03</td>
<td>0.01</td>
<td>0.02</td>
<td>-0.02</td>
<td>-0.03</td>
<td>-0.02</td>
</tr>
<tr>
<td>Cab</td>
<td>-0.12</td>
<td>0.00</td>
<td>0.05</td>
<td>0.05</td>
<td>0.02</td>
<td>0.00</td>
<td>0.00</td>
<td>0.02</td>
<td>-0.01</td>
</tr>
<tr>
<td>Beet</td>
<td>-0.03</td>
<td>-0.05</td>
<td>0.00</td>
<td>-0.07</td>
<td>0.06</td>
<td>-0.01</td>
<td>0.07</td>
<td>0.01</td>
<td>-0.02</td>
</tr>
<tr>
<td>Asp</td>
<td>0.03</td>
<td>-0.05</td>
<td>0.07</td>
<td>0.00</td>
<td>0.01</td>
<td>0.02</td>
<td>0.01</td>
<td>-0.08</td>
<td>-0.01</td>
</tr>
<tr>
<td>Car</td>
<td>-0.01</td>
<td>-0.02</td>
<td>-0.06</td>
<td>-0.01</td>
<td>0.00</td>
<td>-0.02</td>
<td>-0.04</td>
<td>0.08</td>
<td>0.07</td>
</tr>
<tr>
<td>Spin</td>
<td>-0.02</td>
<td>0.00</td>
<td>0.01</td>
<td>-0.02</td>
<td>0.02</td>
<td>0.00</td>
<td>0.03</td>
<td>0.06</td>
<td>-0.06</td>
</tr>
<tr>
<td>S.Beans</td>
<td>0.02</td>
<td>0.00</td>
<td>-0.07</td>
<td>-0.01</td>
<td>0.04</td>
<td>-0.03</td>
<td>0.00</td>
<td>0.01</td>
<td>0.05</td>
</tr>
<tr>
<td>Peas</td>
<td>0.03</td>
<td>-0.02</td>
<td>-0.01</td>
<td>0.08</td>
<td>-0.08</td>
<td>-0.06</td>
<td>-0.01</td>
<td>0.00</td>
<td>0.05</td>
</tr>
<tr>
<td>Corn</td>
<td>0.02</td>
<td>0.01</td>
<td>0.02</td>
<td>0.01</td>
<td>-0.07</td>
<td>0.06</td>
<td>-0.05</td>
<td>-0.05</td>
<td>0.00</td>
</tr>
</tbody>
</table>

> sum(resid)
[1] 3.816392e-16
> sum(resid^2)
[1] 0.1416574
> sum(resid^2)/sum(veg^2)
[1] 0.005697482
1-sum(resid^2)/sum(veg^2)
[1] 0.9943025
```

1. Subtract the model from the data to find the residuals
2. Sum the residuals (equal 0)
3. Sum the squared residuals
4. Compare this to the original data (badness of fit)
5. Convert to a goodness of fit
Consider alternative scaling models

<table>
<thead>
<tr>
<th></th>
<th>constant</th>
<th>equal</th>
<th>squared</th>
<th>reversed</th>
<th>raw</th>
<th>thurstone</th>
</tr>
</thead>
<tbody>
<tr>
<td>Turn</td>
<td>0.5</td>
<td>1</td>
<td>1</td>
<td>9</td>
<td>0.00</td>
<td>0.00</td>
</tr>
<tr>
<td>Cab</td>
<td>0.5</td>
<td>2</td>
<td>4</td>
<td>8</td>
<td>0.15</td>
<td>0.52</td>
</tr>
<tr>
<td>Beet</td>
<td>0.5</td>
<td>3</td>
<td>9</td>
<td>7</td>
<td>0.20</td>
<td>0.65</td>
</tr>
<tr>
<td>Asp</td>
<td>0.5</td>
<td>4</td>
<td>16</td>
<td>6</td>
<td>0.31</td>
<td>0.98</td>
</tr>
<tr>
<td>Car</td>
<td>0.5</td>
<td>5</td>
<td>25</td>
<td>5</td>
<td>0.36</td>
<td>1.12</td>
</tr>
<tr>
<td>Spin</td>
<td>0.5</td>
<td>6</td>
<td>36</td>
<td>4</td>
<td>0.37</td>
<td>1.14</td>
</tr>
<tr>
<td>S.Bean</td>
<td>0.5</td>
<td>7</td>
<td>49</td>
<td>3</td>
<td>0.46</td>
<td>1.40</td>
</tr>
<tr>
<td>Peas</td>
<td>0.5</td>
<td>8</td>
<td>64</td>
<td>2</td>
<td>0.48</td>
<td>1.44</td>
</tr>
<tr>
<td>Corn</td>
<td>0.5</td>
<td>9</td>
<td>81</td>
<td>1</td>
<td>0.54</td>
<td>1.63</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>choice</th>
<th>logistic</th>
<th>normal</th>
</tr>
</thead>
<tbody>
<tr>
<td>Constant</td>
<td>0.81</td>
<td>0.81</td>
<td>0.81</td>
</tr>
<tr>
<td>Equal</td>
<td>0.99</td>
<td>0.88</td>
<td>0.81</td>
</tr>
<tr>
<td>Squared</td>
<td>0.98</td>
<td>0.74</td>
<td>0.74</td>
</tr>
<tr>
<td>Reversed</td>
<td>0.40</td>
<td>-0.27</td>
<td>-0.43</td>
</tr>
<tr>
<td>Raw</td>
<td>0.97</td>
<td>0.89</td>
<td>0.93</td>
</tr>
<tr>
<td>Thurstone</td>
<td>0.97</td>
<td>0.97</td>
<td>0.99</td>
</tr>
</tbody>
</table>

1. Constant says all items are equal
2. Equal implies the steps are all 1
3. Square the values of equal
4. Reverse the rank order!
5. Just the scale values based upon means
6. Thurstonian scaling
We don’t really care all that much about vegetables, but we do care about the process of model fitting.

1. Examine the data
2. Specify a model
3. Estimate the model
4. Compare the model to the data
5. Repeat until satisfied or exhausted
Multidimensional Scaling: \((|o_i - o_j| < |o_k - o_l|)\)

\[
\text{Distance}_{xy} = \sqrt{\sum_{i=1}^{n} (x_i - y_i)^2}.
\] (7)

Consider the cities data set of airline distances.

> cities

\[
\begin{array}{cccccccccccc}
\text{ATL} & \text{BOS} & \text{ORD} & \text{DCA} & \text{DEN} & \text{LAX} & \text{MIA} & \text{JFK} & \text{SEA} & \text{SFO} & \text{MSY} \\
\text{ATL} & 0 & 934 & 585 & 542 & 1209 & 1942 & 605 & 751 & 2181 & 2139 & 424 \\
\text{BOS} & 934 & 0 & 853 & 392 & 1769 & 2601 & 1252 & 183 & 2492 & 2700 & 1356 \\
\text{ORD} & 585 & 853 & 0 & 598 & 918 & 1748 & 1187 & 720 & 1736 & 1857 & 830 \\
\text{DCA} & 542 & 392 & 598 & 0 & 1493 & 2305 & 922 & 209 & 2328 & 2442 & 964 \\
\text{DEN} & 1209 & 1769 & 918 & 1493 & 0 & 836 & 1723 & 1636 & 1023 & 951 & 1079 \\
\text{LAX} & 1942 & 2601 & 1748 & 2305 & 836 & 0 & 2345 & 2461 & 957 & 341 & 1679 \\
\text{MIA} & 605 & 1252 & 1187 & 922 & 1723 & 2345 & 0 & 1092 & 2733 & 2594 & 669 \\
\text{JFK} & 751 & 183 & 720 & 209 & 1636 & 2461 & 1092 & 0 & 2412 & 2577 & 1173 \\
\text{SEA} & 2181 & 2492 & 1736 & 2328 & 1023 & 957 & 2733 & 2412 & 0 & 681 & 2101 \\
\text{SFO} & 2139 & 2700 & 1857 & 2442 & 951 & 341 & 2594 & 2577 & 681 & 0 & 1925 \\
\text{MSY} & 424 & 1356 & 830 & 964 & 1079 & 1679 & 669 & 1173 & 2101 & 1925 & 0 \\
\end{array}
\]
A two dimensional solution of the airline distances

> city.location <- cmdscale(cities, k=2)
> plot(city.location,type="n", xlab="Dimension 1", ylab="Dimension 2",main ="cmdscale(cities)"
> text(city.location,labels=names(cities))
> round(city.location,0)

[,1] [,2]
ATL -571  248
BOS -1061 -548
ORD  -264 -251
DCA  -861 -211
DEN   616  10
LAX  1370  376
MIA  -959  708
JFK  -970 -389
SEA  1438 -607
SFO  1563  88
MSY  -301  577

1. Use the cmdscale function to do multidimensional scaling, ask for a 2 dimensional solution
2. Plot the results (don’t actually show the points)
3. Add the names of the cities
4. Show the numeric results
Original solution for 11 US cities. What is wrong with this figure? Axes of solutions are not necessarily directly interpretable.
Revised solution for 11 US cities after making
\texttt{city.location <- -city.location} \textbf{and adding a US map.}

The correct locations of the cities are shown with circles. The MDS solution is the center of each label. The central cities (Chicago, Atlanta, and New Orleans are located very precisely, but Boston, New York and Washington, DC are north and west of their correct locations.

\begin{figure}[h]
\centering
\includegraphics[width=\textwidth]{multi-dimensional-scaling}
\caption{MultiDimensional Scaling of US cities}
\end{figure}
Preferential Choice: Unfolding Theory (\(|s_i - o_j| < |s_k - o_l|\))

1. “Do I like asparagus more than you like broccoli?” compares how far apart my ideal vegetable is to a particular vegetable (asparagus) with respect to how far your ideal vegetable is to another vegetable (broccoli).

2. More typical is the question of whether you like asparagus more than you like broccoli. This comparison is between your ideal point (on an attribute dimension) to two objects on that dimension.

3. Although the comparisons are ordinal, there is a surprising amount of metric information in the analysis.

4. This involves unfolding the individual preference orderings to find a joint scale of individuals and stimuli (Coombs, 1964, 1975).

5. Can now be done using multidimensional scaling of people and objects using proximity measures.
Some data collection

Please email me your answers to the these two set of questions

1. If you had complete choice in your life, how many children would you like to have?
2. If had complete choice, please rank order the number of children you would like to have
3. Part 2:
4. If you had complete choice in your life, how many children would you like to have? Call that A
5. If you could not have N children, would you rather have A - 1 or A + 1 (call that B)
6. If you could not have A or B children, would you rather have 1 + max(A,B) or min(A,B) - 1, call that C
7. If could not have A,B, or C children, would you rather have 1 + max(A,B,C) or min(A,B,C) - 1, call that D
8. If could not have A,B, C or D children, would you rather have 1 + max(A,B,C,D) or min(A,B,C,D) - 1, call that E

Please email revelle@northwestern.edu your answers.
**Two different J scales**

**Table:** Midpoint ordering gives some metric information. Left hand side: If the midpoint (2|3) comes after (to the right of) the midpoint (0|5) that implies that 3 is closer to 5 than 0 is to 2. Right hand side: The midpoint (2|3) comes before (0|5) and thus 2 is closer to 0 than 3 is to 5. Similarly, that 2|5 comes before 3|4 implies that 4 is closer to 5 than 2 is to 3.

<table>
<thead>
<tr>
<th></th>
<th>0</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>0</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
</tr>
<tr>
<td>0</td>
<td>0</td>
<td>0</td>
<td>5</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0</td>
<td>0</td>
<td>1</td>
<td>2</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0</td>
<td>0</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0</td>
<td>0</td>
<td></td>
<td></td>
<td></td>
<td>4</td>
<td>5</td>
</tr>
<tr>
<td>0</td>
<td>0</td>
<td></td>
<td></td>
<td></td>
<td>3</td>
<td>4</td>
</tr>
<tr>
<td>0</td>
<td>0</td>
<td></td>
<td></td>
<td></td>
<td>2</td>
<td>5</td>
</tr>
</tbody>
</table>
Measuring Abilities and Attitudes

1. Abilities and most models of personality assume an order relationship
   - The comparison is between the person and an item.
   - \( s_i > o_j \)
   - A measurement mode without error is the Guttman scale where \( \text{prob}(\text{correct}|\theta, \delta) = 1|\theta > \delta, \quad 0|\theta < \delta \)
   - With error, a prototypical example is the Rasch scale where \( \text{prob}(\text{correct}|\theta, \delta) = f(\theta - \delta) \)

2. Attitudes (and some personality models) assume a single peak (non-monotone) ordering
   - People endorse attitudes that they are close to, and reject more extreme items.
The Bogardus Social Distance scale as a Guttman scale

Table: The *Bogardus Social Distance Scale* is one example of items that can be made to a *Guttman scale*

“According to my first feeling reactions I would willingly admit members of each race (as a class, and not the best I have known, nor the worst member) to one or more of the classifications under which I have placed a cross (x).”

1. Would exclude from my country
2. As visitors only to my country
3. Citizenship in my country
4. To employment in my occupation in my country
5. To my street as neighbors
6. To my club as personal chums
7. To close kinship by marriage

*(Bogardus, 1925)*
Creating a Guttman scale

```r
> guttman <- matrix(rep(0,56),nrow=8)
> for (i in 1:7) { for (j in 1:i) {guttman[i+1,j] <- 1}}
> rownames(guttman) <- paste("S",1:8,sep="")
> colnames(guttman) <- paste("O",1:7,sep="")
> guttman
     O1 O2 O3 O4 O5 O6 O7
S1   0  0  0  0  0  0  0
S2   1  0  0  0  0  0  0
S3   1  1  0  0  0  0  0
S4   1  1  1  0  0  0  0
S5   1  1  1  1  0  0  0
S6   1  1  1  1  1  0  0
S7   1  1  1  1  1  1  0
S8   1  1  1  1  1  1  1
> rowSums(guttman)
     S1 S2 S3 S4 S5 S6 S7 S8
    0  1  2  3  4  5  6  7
```

1. Create a matrix of 0s
2. Add 1s below the diagonal
3. Give the rows and columns names
4. Show it
5. “score” it
A basic error model with parallel trace lines
Non-monotonic trace lines measure attitudes

Probability of endorsement vs latent attribute
Preferential choice as a comparison of an ideal point for the individual and the objects being chosen

1. The data are rank orders, but the analysis can produce quasi interval scales

2. Originally done by hand, now can be done using the *smacof* package (de Leeuw & Mair, 2009)
**Four types of scales and their associated statistics**

**Table:** Four types of scales and their associated statistics (Rossi, 2007; Stevens, 1946) The statistics listed for a scale are invariant for that type of transformation.

<table>
<thead>
<tr>
<th>Scale</th>
<th>Basic operations</th>
<th>Transformations</th>
<th>Invariant statistic</th>
<th>Examples</th>
</tr>
</thead>
<tbody>
<tr>
<td>Nominal</td>
<td>equality</td>
<td>Permutations</td>
<td>Counts</td>
<td>Detection</td>
</tr>
<tr>
<td></td>
<td>$x_i = x_j$</td>
<td></td>
<td>Mode</td>
<td>Species classification</td>
</tr>
<tr>
<td>Ordinal</td>
<td>order</td>
<td>Monotonic (homeomorphic)</td>
<td>Median</td>
<td>Mhos Hardness scale</td>
</tr>
<tr>
<td></td>
<td>$x_i &gt; x_j$</td>
<td>$x' = f(x)$ f is monotonic</td>
<td>Percentiles</td>
<td>Beaufort Wind (intensity)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Spearman correlations*</td>
<td>Richter earthquake scale</td>
</tr>
<tr>
<td>Interval</td>
<td>differences</td>
<td>Linear (Affine)</td>
<td>Mean ($\mu$)</td>
<td>Temperature ($^\circ F$, $^\circ C$)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>$x' = a + bx$</td>
<td>Standard Deviation ($\sigma$)</td>
<td>Beaufort Wind (velocity)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Pearson correlation ($r$)</td>
<td></td>
</tr>
<tr>
<td>Ratio</td>
<td>ratios</td>
<td>Multiplication (Similarity)</td>
<td>Coefficient of variation ($\frac{\sigma}{\mu}$)</td>
<td>Length, mass, time</td>
</tr>
<tr>
<td></td>
<td></td>
<td>$x' = bx$</td>
<td></td>
<td>Temperature ($^\circ K$)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Heating degree days</td>
</tr>
</tbody>
</table>

The Beaufort wind speed scale is interval with respect to the velocity of the wind, but only ordinal with respect to the effect of the wind. The Richter scale of earthquake intensity is a logarithmic scale of the energy released but...
Graphical and tabular summaries of data

1. The Tukey 5 number summary shows the important characteristics of a set of numbers
   - Maximum
   - 75th percentile
   - Median (50th percentile)
   - 25th percentile
   - Minimum

2. Graphically, this is the box plot
   - Variations on the box plot include confidence intervals for the median
The summary command gives the Tukey 5 numbers

```r
> summary(sat.act)

          gender  education  age  ACT  SATV
Min. :1.000 Min. :0.000 Min. :13.00 Min. : 3.00 Min. :200.0
1st Qu.:1.000 1st Qu.:3.000 1st Qu.:19.00 1st Qu.:25.00 1st Qu.:550.0
Median :2.000 Median :3.000 Median :22.00 Median :29.00 Median :620.0
Mean :1.647  Mean :3.164  Mean :25.59  Mean :28.55  Mean :612.2
3rd Qu.:2.000  3rd Qu.:4.000  3rd Qu.:29.00  3rd Qu.:32.00  3rd Qu.:700.0
Max. :2.000  Max. :5.000  Max. :65.00  Max. :36.00  Max. :800.0
```
A box plot of the first 4 sat.act variables

A Tukey Boxplot

```r
boxplot(sat.act[1:4], main = "A Tukey Boxplot")
```
A violin or density plot of the first 5 epi.bfi variables
The `describe` function gives more descriptive statistics

```r
> describe(sat.act)

vars  n  mean  sd  median  trimmed  mad  min  max  range  skew kurtosis  se
gender   1 700  1.65  0.48   2   1.68  0.00   1  2  1 -0.61   -1.62   0.02
education 2 700  3.16  1.43   3   3.31  1.48   0  5  5 -0.68   -0.07   0.05
age      3 700  25.59  9.50  22   23.86  5.93  13 65 52  1.64    2.42   0.36
ACT      4 700  28.55  4.82  29   28.84  4.45   3 36 33 -0.66    0.53   0.18
SATV     5 700  612.23 112.90 620  619.45 118.61 200 800 600 -0.64    0.33   4.27
SATQ     6 687  610.22 115.64 620  617.25 118.61 200 800 600 -0.59   -0.02   4.41
```
Multiple measures of central tendency

**mode**  The most frequent observation. Not a very stable measure, depends upon grouping. Can be used for categorical data.

**median**  The number with 50% above and 50% below. A powerful, if underused, measure. Not sensitive to transforms of the shape of the distribution, nor outliers. Appropriate for ordinal data, and useful for interval data.

**mean**  One of at least seven measures that assume interval properties of the data.
Multiple ways to estimate the mean

Arithmetic mean \( \bar{X} = \bar{X} = \left( \sum_{i=1}^{N} X_i \right) / N \) mean \( (x) \)

Trimmed mean throws away the top and bottom t% of observations. This follows the principle that all data are normal at the middle. \( \text{mean}(x, \text{trim}=.1) \)

Winsorized mean Find the arithmetic mean after replacing the n lowest observations with the nth value, and the N largest values with the Nth largest. \( \text{winsor}(x, \text{trim}=.2) \)

Geometric Mean \( \bar{X}_{\text{geometric}} = \sqrt[N]{\prod_{i=1}^{N} X_i} = e^{\sum \ln(x)} / N \) (The anti-log of the mean log score). \( \text{geometric.mean}(x) \)

Harmonic Mean \( \bar{X}_{\text{harmonic}} = \frac{N}{\sum_{i=1}^{N} 1/X_i} \) (The reciprocal of the mean reciprocal). \( \text{harmonic.mean}(x) \)

Circular Mean \( \bar{X}_{\text{circular}} = \tan^{-1} \left( \frac{\sum \cos(x)}{\sum \sin(x)} \right) \) circular.mean \( (x) \)

\( \text{circadian.mean} \) circular.mean \( (x) \) (where x is in hours)
Circular statistics

Table: Hypothetical mood data from six subjects for four mood variables. The values reflect the time of day that each scale achieves its maximum value for each subject. Each mood variable is just the previous one shifted by 5 hours. Note how this structure is preserved for the circular mean but not for the arithmetic mean.

<table>
<thead>
<tr>
<th>Subject</th>
<th>Energetic Arousal</th>
<th>Positive Affect</th>
<th>Tense Arousal</th>
<th>Negative Affect</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>9</td>
<td>14</td>
<td>19</td>
<td>24</td>
</tr>
<tr>
<td>2</td>
<td>11</td>
<td>16</td>
<td>21</td>
<td>2</td>
</tr>
<tr>
<td>3</td>
<td>13</td>
<td>18</td>
<td>23</td>
<td>4</td>
</tr>
<tr>
<td>4</td>
<td>15</td>
<td>20</td>
<td>1</td>
<td>6</td>
</tr>
<tr>
<td>5</td>
<td>17</td>
<td>22</td>
<td>3</td>
<td>8</td>
</tr>
<tr>
<td>6</td>
<td>19</td>
<td>24</td>
<td>5</td>
<td>10</td>
</tr>
<tr>
<td>Arithmetic Mean</td>
<td>14</td>
<td>19</td>
<td>12</td>
<td>9</td>
</tr>
<tr>
<td>Circular Mean</td>
<td>14</td>
<td>19</td>
<td>24</td>
<td>5</td>
</tr>
</tbody>
</table>
Some hypothetical data stored in a data.frame

<table>
<thead>
<tr>
<th>Participant</th>
<th>Name</th>
<th>Gender</th>
<th>θ</th>
<th>X</th>
<th>Y</th>
<th>Z</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Bob</td>
<td>Male</td>
<td>1</td>
<td>12</td>
<td>2</td>
<td>1</td>
</tr>
<tr>
<td>2</td>
<td>Debby</td>
<td>Female</td>
<td>3</td>
<td>14</td>
<td>6</td>
<td>4</td>
</tr>
<tr>
<td>3</td>
<td>Alice</td>
<td>Female</td>
<td>7</td>
<td>18</td>
<td>14</td>
<td>64</td>
</tr>
<tr>
<td>4</td>
<td>Gina</td>
<td>Female</td>
<td>6</td>
<td>17</td>
<td>12</td>
<td>32</td>
</tr>
<tr>
<td>5</td>
<td>Eric</td>
<td>Male</td>
<td>4</td>
<td>15</td>
<td>8</td>
<td>8</td>
</tr>
<tr>
<td>6</td>
<td>Fred</td>
<td>Male</td>
<td>5</td>
<td>16</td>
<td>10</td>
<td>16</td>
</tr>
<tr>
<td>7</td>
<td>Chuck</td>
<td>Male</td>
<td>2</td>
<td>13</td>
<td>4</td>
<td>2</td>
</tr>
</tbody>
</table>

> s.df <- read.clipboard()
> dim(s.df) #how many elements are in each dimension
[1] 7 7
> str(s.df) #show the structure
'data.frame': 7 obs. of 7 variables:
$ Participant: int 1 2 3 4 5 6 7
$ Name : Factor w/ 7 levels "Alice","Bob",...: 2 4 1 7 5 6 3
$ Gender : Factor w/ 2 levels "Female","Male": 2 1 1 1 2 2 2
$ theta : int 1 3 7 6 4 5 2
$ X : int 12 14 18 17 15 16 13
$ Y : num 2 6 14 12 8 10 4
$ Z : int 1 4 64 32 8 16 2

g
Saving the data.frame in a readable form

The previous slide is readable by humans, but harder to read by computer. PDFs are formatted in a rather weird way. We can share data on slides by using the dput function. Copy this output to your clipboard from the slide, and then get it into R directly.

```r
> dput(sf.df)

structure(list(ID = 1:7, Name = structure(c(2L, 4L, 1L, 7L, 5L, 6L, 3L), .Label = c("Alice", "Bob", "Chuck", "Debby", "Eric", "Fred", "Gina"), class = "factor"), gender = structure(c(2L, 1L, 1L, 2L, 2L, 2L), .Label = c("Female", "Male"), class = "factor"), theta = c(1L, 3L, 7L, 6L, 4L, 5L, 2L), X = c(12L, 14L, 18L, 17L, 15L, 16L, 13L), Y = c(2L, 6L, 14L, 12L, 8L, 10L, 4L), Z = c(1L, 4L, 64L, 32L, 8L, 16L, 2L)), .Names = c("ID", "Name", "gender", "theta", "X", "Y", "Z"), class = "data.frame", row.names = c(NA, -7L))
```

```r
my.data <- structure(list(ID = 1:7, Name = structure(c(2L, 4L, 1L, 7L, 5L, 6L, 3L), .Label = c("Alice", "Bob", "Chuck", "Debby", "Eric", "Fred", "Gina"), class = "factor"), gender = structure(c(2L, 1L, 1L, 2L, 2L, 2L), .Label = c("Female", "Male"), class = "factor"), theta = c(1L, 3L, 7L, 6L, 4L, 5L, 2L), X = c(12L, 14L, 18L, 17L, 15L, 16L, 13L), Y = c(2L, 6L, 14L, 12L, 8L, 10L, 4L), Z = c(1L, 4L, 64L, 32L, 8L, 16L, 2L)), .Names = c("ID", "Name", "gender", "theta", "X", "Y", "Z"), class = "data.frame", row.names = c(NA, -7L))
```
Sorting the data can display certain features

We use the order function applied to the “Names” column and then to the 4th column.

```r
> my.data.alpha <- my.data[order(my.data[,"Name"]),] > my.data.theta <- my.data[order(my.data[,4]),]
> my.data.alpha
 ID Name gender theta X Y Z
3  3 Alice Female  7 18 14 64
1  1 Bob Male  1 12 2 1
7  7 Chuck Male  2 13 4 2
2  2 Debby Female  3 14 6 4
5  5 Eric Male  4 15 8 8
6  6 Fred Male  5 16 10 16
4  4 Gina Female  6 17 12 32
> my.data.theta
 ID Name gender theta X Y Z
1  1 Bob Male  1 12 2 1
7  7 Chuck Male  2 13 4 2
2  2 Debby Female  3 14 6 4
5  5 Eric Male  4 15 8 8
6  6 Fred Male  5 16 10 16
4  4 Gina Female  6 17 12 32
3  3 Alice Female  7 18 14 64
```

It was harder to see the perfect relationship between \( \theta \) and \( X, Y, \) and \( Z \) with the original data.
Multiple estimates of the central tendency using the apply function

```r
> apply(my.data[4:7],2,mean)
  theta   X    Y    Z
4.00000 15.00000 8.00000 18.14286
> apply(my.data[4:7],2,mean,trim=.2)
  theta   X    Y    Z
 4.0 15.0 8.0 12.4
> apply(my.data[4:7],2,winsor.mean,trim=.2)
  theta   X    Y    Z
4.00000 15.00000 8.00000 12.91429
> apply(my.data[4:7],2,harmonic.mean)
  theta   X    Y    Z
2.699725 14.729687 5.399449 3.527559
> apply(my.data[4:7],2,geometric.mean)
  theta   X    Y    Z
3.380015 14.865151 6.760030 8.000000
```

1. The basic mean is applied to columns 4 - 7
2. Then do this, but trim the top and bottom 20%
3. Now, don’t trim, but winsorize
4. Compare with the harmonic mean
5. Compare with geometric mean.
**Effect of reciprocal transformation upon means**

**Table**: Hypothetical study of arousal using an exciting movie. The post test shows greater arousal if measured using skin conductance (higher skin conductance means more arousal), but less arousal if measured using skin resistance (higher skin conductance means less arousal)

<table>
<thead>
<tr>
<th>Condition</th>
<th>Subject</th>
<th>Skin Conductance</th>
<th>Skin Resistance</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pretest (Control)</td>
<td>1</td>
<td>2</td>
<td>.50</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>2</td>
<td>.50</td>
</tr>
<tr>
<td>Average</td>
<td></td>
<td>2</td>
<td>.50</td>
</tr>
<tr>
<td>Posttest (Movie)</td>
<td>1</td>
<td>1</td>
<td>1.00</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>4</td>
<td>.25</td>
</tr>
<tr>
<td>Average</td>
<td></td>
<td>2.5</td>
<td>.61</td>
</tr>
</tbody>
</table>
Non linearity can influence means if the variances differ.
What is the "average" class size?

Table: Average class size depends upon point of view. For the faculty members, the median of 10 is very appealing. From the Dean’s perspective, the faculty members teach an average of 50 students per calls. But what about the students?

<table>
<thead>
<tr>
<th>Faculty Member</th>
<th>Freshman/Sophomore</th>
<th>Junior</th>
<th>Senior</th>
<th>Graduate</th>
<th>Mean</th>
<th>Median</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>20</td>
<td>10</td>
<td>10</td>
<td>10</td>
<td>12.5</td>
<td>10</td>
</tr>
<tr>
<td>B</td>
<td>20</td>
<td>10</td>
<td>10</td>
<td>10</td>
<td>12.5</td>
<td>10</td>
</tr>
<tr>
<td>C</td>
<td>20</td>
<td>10</td>
<td>10</td>
<td>10</td>
<td>12.5</td>
<td>10</td>
</tr>
<tr>
<td>D</td>
<td>20</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>35.0</td>
<td>15</td>
</tr>
<tr>
<td>E</td>
<td>200</td>
<td>100</td>
<td>400</td>
<td>100</td>
<td>177.5</td>
<td>150</td>
</tr>
<tr>
<td>Total Mean</td>
<td>56</td>
<td>46</td>
<td>110</td>
<td>10</td>
<td>50.0</td>
<td>39</td>
</tr>
<tr>
<td>Median</td>
<td>20</td>
<td>10</td>
<td>10</td>
<td>10</td>
<td>12.5</td>
<td>10</td>
</tr>
</tbody>
</table>
Class size from the students’ point of view.

Table: Class size from the students’ point of view. Most students are in large classes; the median class size is 200 with a mean of 223.

<table>
<thead>
<tr>
<th>Class size</th>
<th>Number of classes</th>
<th>number of students</th>
</tr>
</thead>
<tbody>
<tr>
<td>10</td>
<td>12</td>
<td>120</td>
</tr>
<tr>
<td>20</td>
<td>4</td>
<td>80</td>
</tr>
<tr>
<td>100</td>
<td>2</td>
<td>200</td>
</tr>
<tr>
<td>200</td>
<td>1</td>
<td>200</td>
</tr>
<tr>
<td>400</td>
<td>1</td>
<td>400</td>
</tr>
</tbody>
</table>
A psychotherapist is asked what is the average length of time that a patient is in therapy. This seems to be an easy question, for of the 20 patients, 19 have been in therapy for between 6 and 18 months (with a median of 12) and one has just started. Thus, the median client is in therapy for 52 weeks with an average (in weeks) 
\[(1 \times 1 + 19 \times 52)/20 \text { or } 49.4.\]
However, a more careful analysis examines the case load over a year and discovers that indeed, 19 patients have a median time in treatment of 52 weeks, but that each week the therapist is also seeing a new client for just one session. That is, over the year, the therapist sees 52 patients for 1 week and 19 for a median of 52 weeks. Thus, the median client is in therapy for 1 week and the average client is in therapy of \[(52 \times 1 + 19 \times 52)/(52+19) \text { or } 14.6 \text{ weeks}.\]
Does teaching effect learning?

1. A leading research team in motivational and educational psychology was interested in the effect that different teaching techniques at various colleges and universities have upon their students. They were particularly interested in the effect upon writing performance of attending a very selective university, a less selective university, or a two year junior college.

2. A writing test was given to the entering students at three institutions in the Boston area. After one year, a similar writing test was given again. Although there was some attrition from each sample, the researchers report data only for those who finished one year. The pre and post test scores as well as the change scores were as shown below:
Types of teaching affect student outcomes?

**Table:** Three types of teaching and their effect on student outcomes

<table>
<thead>
<tr>
<th>School</th>
<th>Pretest</th>
<th>Posttest</th>
<th>Change</th>
</tr>
</thead>
<tbody>
<tr>
<td>Junior College</td>
<td>1</td>
<td>5</td>
<td>4</td>
</tr>
<tr>
<td>Non-selective university</td>
<td>5</td>
<td>27</td>
<td>22</td>
</tr>
<tr>
<td>Selective university</td>
<td>27</td>
<td>73</td>
<td>45</td>
</tr>
</tbody>
</table>

From these data, the researchers concluded that the quality of teaching at the selective university was much better than that of the less selective university or the junior college and that the students learned a great deal more. They proposed to study the techniques used there in order to apply them to the other institutions.
Teaching and math performance

Another research team in motivational and educational psychology was interested in the effect that different teaching at various colleges and universities affect math performance. They used the same schools as the previous example with the same design.

Table: Three types of teaching and their effect on student outcomes

<table>
<thead>
<tr>
<th>School</th>
<th>Pretest</th>
<th>Posttest</th>
<th>Change</th>
</tr>
</thead>
<tbody>
<tr>
<td>Junior College</td>
<td>27</td>
<td>73</td>
<td>45</td>
</tr>
<tr>
<td>Non-selective university</td>
<td>73</td>
<td>95</td>
<td>22</td>
</tr>
<tr>
<td>Selective university</td>
<td>95</td>
<td>99</td>
<td>4</td>
</tr>
</tbody>
</table>

They concluded that the teaching at the junior college was far superior to that of the select university. What is wrong with this conclusion?
Effect of teaching, effect of students, or just scaling?

**Writing**

- Ivy
- TC
- JC

**Math**

- Ivy
- TC
- JC

Performance vs. time in school.
The effect of scaling upon the latent variable - observed variable relationship

..../205.16/205scaling.pdf
The problem of scaling is ubiquitous

1. A leading cognitive developmentalist believed that there is a critical stage for learning spatial representations using maps. Children younger than this stage are not helped by maps, nor are children older than this stage.

2. He randomly assigned 3rd, 5th, and 7th grade students into two conditions (nested within grade), control, and map use. Performance was measures on a task of spatial recall (children were shown toys at particular locations in a set of rooms and then asked to find them again later.) Half the children were shown a map of the rooms before doing the task.

3. Their scores were

<table>
<thead>
<tr>
<th>Grade</th>
<th>No Map</th>
<th>Maps</th>
<th>Effect</th>
</tr>
</thead>
<tbody>
<tr>
<td>3rd</td>
<td>5</td>
<td>27</td>
<td>22 Too young</td>
</tr>
<tr>
<td>5th</td>
<td>27</td>
<td>73</td>
<td>46 Critical period</td>
</tr>
<tr>
<td>7th</td>
<td>73</td>
<td>95</td>
<td>22 Too old</td>
</tr>
</tbody>
</table>
Map use is most effective at a particular developmental stage

Recall varies by age and exposure to maps

![Graph showing recall varies by age and exposure to maps](image)
R code for the prior figure

```r
mapuse <- matrix(c(3,5,27,5,27,73,7,73,95),ncol=3,byrow=TRUE)
colnames(mapuse) <- c("grade","nomaps","maps")
rownames(mapuse) <- c("3rd","5th","7th")
maps.df <- data.frame(mapuse)
maps.df

with(maps.df,plot(maps~grade,ylab="Recall",ylim=c(0,100),
typ="b", main="Recall varies by age and exposure to maps"))
with(maps.df,points(nomaps~grade,ylab="Recall",
    ylim=c(0,100),typ="b",lty="dashed")
> text(5,80,"maps")  #add line labels
> text(5,15,"nomaps")
```

<table>
<thead>
<tr>
<th>grade</th>
<th>nomaps</th>
<th>maps</th>
</tr>
</thead>
<tbody>
<tr>
<td>3rd</td>
<td>3</td>
<td>5</td>
</tr>
<tr>
<td>5th</td>
<td>5</td>
<td>27</td>
</tr>
<tr>
<td>7th</td>
<td>7</td>
<td>73</td>
</tr>
</tbody>
</table>
Yet another developmentalist

Another cognitive developmentalist believed that there is a critical stage but that it appears earlier than previously thought. Children younger than this stage are not helped by maps, nor are children older than this stage. He randomly assigned 1st, 3rd, 5th and 7th grade students into two conditions (nested within grade), control and mapa use. Performance was measured on a task of spatial recall (children were shown toys at particular locations in a set of rooms and then asked to find them again later. Half the children were shown a map of the room before doing the task.

The scores were

<table>
<thead>
<tr>
<th>No Map</th>
<th>Maps</th>
<th>Effect</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>1st grade</td>
<td>2</td>
<td>12</td>
<td>10 Too young</td>
</tr>
<tr>
<td>3rd grade</td>
<td>12</td>
<td>50</td>
<td>38</td>
</tr>
<tr>
<td>5th grade</td>
<td>50</td>
<td>88</td>
<td>38 Critical period</td>
</tr>
<tr>
<td>7th grade</td>
<td>88</td>
<td>98</td>
<td>10 Too old</td>
</tr>
</tbody>
</table>
A critical period in developmental?

Recall varies by age and exposure to maps

Recall

maps

nomaps

grade
R code for the prior figure

```r
mapuse <- matrix(c(1,2,12,10,3,12,5,50,5,50,38,88,7,88,98,10),ncol=4,byrow=TRUE)
colnames(mapuse) <- c("grade","nomaps","maps","Diff")
rownames(mapuse) <- c("1st","3rd","5th","7th")
maps.df <- data.frame(mapuse)
maps.df
with(maps.df,plot(maps~grade,ylab="Recall",ylim=c(0,100),
typ="b",main="Recall varies by age and exposure to maps"))
with(maps.df,points(nomaps~grade,ylab="Recall",
    ylim=c(0,100),typ="b",lty="dashed"))
text(4,75,"maps")  #add line labels
text(4,20,"nomaps")

grade     nomaps maps
3rd       3 5 27
5th       5 27 73
7th       7 73 95
```
Traditional levels of measurement

**Nominal**  Categories: X, Y, W, V

**Ordinal**  Ranks \((X > Y > W > V)\)

**Interval**  Equal Differences \((X - Y > W - V)\)

**Ratio**  Equal intervals with a zero point \((X/Y > W/V)\)
Types of scales and types of inference

1. Nominal allow us to say whether groups differ in frequency
2. Ordinal allows to compare rank orders of the data, is one score greater than another score. Any monotonic transformation will preserve rank order.
3. Interval is the claim that we can compare the magnitude of intervals. Only linear transformations will preserve interval information (i.e. we can add and subtract the numbers and preserve interval information.
4. Ratio scales preserve absolute magnitude differences.
Ordinal scales

1. Any monotonic transformation will preserve order
2. Inferences from observed to latent variable are restricted to rank orders
3. Statistics: Medians, Quartiles, Percentiles
**Interval scales**

1. Possible to infer the magnitude of differences between points on the latent variable given differences on the observed variable? X is as much greater than Y as Z is from W.

2. Linear transformations preserve interval information.


4. Although our data are actually probably just ordinal, we tend to use interval assumptions.
Ratio Scales

1. Interval scales with a zero point
2. Possible to compare ratios of magnitudes \((X\) is twice as long as \(Y)\)
3. Are there any psychological examples?
The search for an appropriate scale

1. Is today colder than yesterday? (ranks) Is the amount that today is colder than yesterday more than the amount that yesterday was colder than the day before? (intervals)
   - $50F - 39F < 68F - 50F$
   - $10C - 4C < 20C - 10C$
   - $283K - 277K < 293K - 283K$

2. How much colder is today than yesterday?
   - (Degree days as measure of energy use) is almost ratio
   - $K$ as measure of molecular energy
Measurement confusions – arousal

1. Arousal is a fundamental concept in many psychological theories. It is thought to reflect basic levels of alertness and preparedness. Typical indices of arousal are measures of the amount of palmer sweating.

2. This may be indexed by the amount of electricity that is conducted by the fingertips.

3. Alternatively, it may be indexed (negatively) by the amount of skin resistance of the finger tips. The Galvanic Skin Response (GSR) reflects moment to moment changes, SC and SR reflect longer term, basal levels.

4. High skin conductance (low skin resistance) is thought to reflect high arousal.
Arousal and anxiety

1. Anxiety is thought to be related to arousal. The following data were collected by two different experimenters. One collected Resistance, conductance data. Low anxiety had:
   - Resistance: 1, 5
   - Conductance: 1, .2
High anxiety had:
   - Resistance: 2, 2
   - Conductance: .5, .5
The means were therefore:
   - Resistance: 3
   - Conductance: .6

2. That is, the low anxiety participants had higher skin resistance and thus were more relaxed, but they also had higher skin conductance, and thus were more aroused.

3. How can this be?
Multiple measures of dispersion

**Range** (highest - lowest) is sensitive to the number of observations, but is a very good way to detect errors in data entry.

**MAD** (Median Absolute Deviation from the Median) applied ordinal statistics to interval measures

**Variance** ($\sigma^2$) is the Mean Square deviation (implies interval data)

**Standard Deviation** ($\sigma$) is the Root Mean Square deviation.

**Coefficient of Variation** $\frac{\sigma_x}{\mu_x}$

**Average difference** $\sigma_x \sqrt{2}$
Normal and non-normal curves

Normal and non-normal

- \( N(0,1) \)
- Cauchy
- logistic
- \( N(0,2) \)
Three normal curves

![Three normal curves](image)

- N(0,1)
- N(1,.5)
- N(0,2)
Seriously contaminated data

Normal and contaminated data

![Graph showing normal and contaminated data distributions.](attachment:graph.png)
The normal curve and its frequent transforms

Alternative scalings of the normal curve

Probability of z

-3 -2 -1 0 1 2 3
0.0 0.1 0.2 0.3 0.4

percentile

IQ

SAT

stanine

Galton/Tukey

lower fence Q1 Median Q3 Upper Fence
Decision making and the benefit of extreme selection ratios

1. Typical traits are approximated by a normal distribution.
2. Small differences in means or variances can lead to large differences in relative odds at the tails.
3. Accuracy of decision/prediction is higher for extreme values.
4. Do we infer trait mean differences from observing differences of extreme values?
The effect of small mean differences at the tails of a distribution

**difference = .25**

**difference = .5**
The effect of small differences in variance at the tails of a distribution

- **sigma = 1.1**
  - Probability density functions (PDFs)
  - Cumulative distribution functions (CDFs)
- **sigma = 1.2**
  - PDFs
  - CDFs

Relative comparison of distributions with differing variances at their tails.
Table: Tukey’s ladder of transformations. One goes up and down the ladder until the relationships desired are roughly linear or the distribution is less skewed. The effect of taking powers of the numbers is to emphasize the larger numbers, the effect of taking roots, logs, or reciprocals is to emphasize the smaller numbers.

<table>
<thead>
<tr>
<th>Transformation</th>
<th>effect</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>$x^3$</td>
<td>emphasize large numbers</td>
<td>reduce negative skew</td>
</tr>
<tr>
<td>$x^2$</td>
<td>emphasize large numbers</td>
<td>reduce negative skew</td>
</tr>
<tr>
<td>$\sqrt{x}$</td>
<td>emphasize smaller numbers</td>
<td>reduce positive skew</td>
</tr>
<tr>
<td>$-\frac{1}{x}$</td>
<td>emphasize smaller numbers</td>
<td>reduce positive skew</td>
</tr>
<tr>
<td>log($x$)</td>
<td>emphasize smaller numbers</td>
<td>reduce positive skew</td>
</tr>
<tr>
<td>$-\frac{1}{x^2}$</td>
<td>emphasize smaller numbers</td>
<td>reduce positive skew</td>
</tr>
<tr>
<td>$-\frac{1}{x^3}$</td>
<td>emphasize smaller numbers</td>
<td>reduce positive skew</td>
</tr>
</tbody>
</table>
Tukey's ladder of transformations

Figure: Tukey (1977) suggested a number of transformations of data that allow relationships to be seen more easily. Ranging from the cube to the reciprocal of the cube, these transformations emphasize different parts of the distribution.
The best scale is the one that works best

1. Money is linear but negatively accelerated with utility.
2. Perceived intensity is a log function of physical intensity.
3. Probability of being correct is a logistic or cumulative normal function of ability.
4. Energy used to heat a house is linear function of outdoor temperature.
5. Time to fall a particular distance varies as the square root of the distance \( (s = at^2 \iff t = \sqrt{\frac{s}{a}}) \)
6. Gravitational attraction varies as \( 1/distance^2 \) \( (F = G \frac{m_1 m_2}{d^2}) \)
7. Hull speed of sailboat varies as square root of length of boat.
8. Sound intensity in db is \( \log(\text{observed}/\text{reference}) \)
9. pH of solutions is \( -\log(\text{concentration of hydrogen ions}) \)


Appleton-Century-Crofts Print.