

Contents

Part I Background to Psychological Measurement

1	Introduction	5
1.1	Constructs and measures (Chapter 2)	6
2	A Theory of Data	15
2.1	A Theory of Data: Objects, People, and Comparisons	15
2.1.1	Modeling the comparison process	16
2.2	Models and model fitting	16
2.3	A brief diversion: Functions in R	18
2.4	Tournaments: Ordering people ($p_i > p_j$)	19
2.4.1	Scaling of People	20
2.4.2	Alternative approaches to scaling people	22
2.4.3	Assigning numbers to people – the problem of rewarding merit	23
2.5	Social Networks: Proximities of People ($ p_i - p_j < \delta$)	23
2.5.1	Rectangular data arrays of similarity	23
2.5.2	Square arrays of similarity	24
2.6	The Scaling of Objects ($o_i < o_j$)	25
2.6.1	Weber-Fechner scales of subjective experience	27
2.6.2	Thurstonian Scaling	28
2.6.3	Alternative solutions to the ranking of objects	32
2.6.4	Why emphasize Thurstonian scaling?	34
2.7	Multiple Dimensional Scaling: Distances between Objects ($ o_i - o_j < o_k - o_l $)	34
2.8	Preferential Choice: Unfolding Theory ($ s_i - o_j < s_k - o_l $)	36
2.8.1	Individual Preferences – the I scale	36
2.8.2	Joint Preferences – the J scale	38
2.8.3	Partially ordered metrics	38
2.8.4	Multidimensional Unfolding	39
2.9	Measurement of Attitudes and Abilities (comparing s_i, o_j)	40
2.9.1	Measurement of abilities ($s_i > o_j$)	40
2.9.2	Measurement of attitudes ($ s_i - o_j < \delta$)	43
2.10	Theory of Data: some final comments	45

3	The problem of scale	47
3.1	Four broad classes of scales	49
3.1.1	Factor levels as Nominal values	51
3.1.2	Integers and Reals: Ordinal or Metric values?	51
3.2	Graphical and numeric summaries of the data	51
3.2.1	Sorting data as a summary technique	52
3.3	Numerical estimates of central tendency	52
3.3.1	Mode: the most frequent	53
3.3.2	Median: the middle observation	55
3.3.3	3 forms of the mean	57
3.3.4	Comparing variables or groups by their central tendency	59
3.4	The effect of non-linearity on estimates of central tendency	59
3.4.1	Circular Means	60
3.5	Whose mean? The problem of point of view	62
3.5.1	Average length of time in psychotherapy	62
3.5.2	Average class size	63
3.6	Non-linearity and interpretation of experimental effects	63
3.6.1	Linearity, non-linearity and the properties of measurement	66
3.7	Measures of dispersion	66
3.7.1	Measures of range	66
3.7.2	Average distance from the central tendency	67
3.7.3	Deviation scores and the standard deviation	68
3.7.4	Coefficient of variation	69
3.8	Geometric interpretations of Variance and Covariance	70
3.9	Variance, Covariance, and Distance	71
3.10	Standard scores as unit free measures	72
3.11	Assessing the higher order moments of the normal and other distributions	72
3.12	Generating commonly observed distributions	77
3.13	Mixed distributions	77
3.14	Robust measures of dispersion	78
3.15	Monotonic transformations of data and “Tukey’s ladder”	80
3.16	What is the fundamental scale?	82
4	Covariance, Regression, and Correlation	85
4.1	Correlation as the geometric mean of regressions	88
4.2	Regression and prediction	90
4.3	A geometric interpretation of covariance and correlation	91
4.4	The bivariate normal distribution	93
4.4.1	Confidence intervals of correlations	93
4.4.2	Testing whether correlations differ from zero	94
4.4.3	Testing the difference between correlations	95
4.5	Other estimates of association	98
4.5.1	Pearson correlation equivalents	99
4.6	Other measures of association	109
4.6.1	Naturally dichotomous data	110
4.6.2	Measures of association for categorical data	114
4.6.3	Intraclass Correlation	115
4.6.4	Quantile Regression	117

4.6.5	Kendall's Tau	117
4.6.6	Circular-circular and circular-linear correlations	118
4.7	Alternative estimates of effect size	118
4.8	Sources of confusion	119
4.8.1	Restriction of range	120
4.8.2	Spurious correlations	121
4.8.3	Non linearity, outliers and other problems: the importance of graphics	125
5	Multiple correlation and multiple regression	127
5.1	The variance of composites	127
5.2	Multiple regression	128
5.2.1	Direct and indirect effects, suppression and other surprises	130
5.2.2	Interactions and product terms: the need to center the data	130
5.2.3	Confidence intervals of the regression and regression weights	132
5.2.4	Multiple regression from the covariance/correlation matrix	134
5.2.5	The robust beauty of linear models	134
5.3	Partial and semi-partial correlation	134
5.3.1	Alternative interpretations of the partial correlation	136
5.4	Alternative regression techniques	136
5.4.1	Logistic regression	136
5.4.2	Poisson regression, quasi-Poisson regression, and negative-binomial regression	137
5.4.3	Using multiple regression for circular data	139
5.4.4	Robust regression using M estimators	141

Part II Latent Variable Models

6	Constructs, Components, and Factor models	145
6.1	Principal Components: an observed variable model	147
6.1.1	Eigenvalues and Eigenvectors	147
6.1.2	Principal components	149
6.2	Exploratory Factor Analysis: a latent variable model	150
6.2.1	Principal Axes Factor Analysis as an eigenvalue decomposition of a reduced matrix	152
6.2.2	Maximum Likelihood Factor Analysis and its alternatives	153
6.2.3	Comparing extraction techniques	156
6.2.4	Exploratory analysis with more than one factor/component	157
6.2.5	Comparing factors and components- part 1	158
6.3	Rotations and Transformations	161
6.3.1	Orthogonal rotations	164
6.3.2	Oblique transformations	164
6.3.3	Non-Simple Structure Solutions: The Simplex and Circumplex	167
6.3.4	Hierarchical and higher order models	171
6.3.5	Comparing factor solutions	171
6.4	The number of factors/components problem	175
6.5	The number of subjects problem	179
6.6	The problem of factoring items	180
6.7	Confirmatory Factor Analysis	180

6.8	Alternative procedures for reducing the complexity of the data	184
6.8.1	MDS solutions remove the general factor	185
6.8.2	Cluster analysis – poor man’s factor analysis?	186
6.9	Estimating factor scores, finding component and cluster scores	195
6.9.1	Extending a factor solution to new variables	198
6.9.2	Comparing factors and components – part 2	200
 Part III Classical and Modern Test Theory		
7	Classical Test Theory and the Measurement of Reliability	205
7.1	Reliability and True Scores	207
7.1.1	Parallel Tests, Reliability, and Corrections for Attenuation	208
7.1.2	Tau equivalent and congeneric tests	211
7.2	Reliability and internal structure	213
7.2.1	Split half reliability	213
7.2.2	Domain sampling	215
7.2.3	The internal structure of a test. Part 1: coefficient α	217
7.2.4	The internal structure of a test. Part 2: Guttman’s lower bounds of reliability	219
7.2.5	The internal structure of a test. Part 3: coefficients α , β , ω_h and ω_t	224
7.3	A comparison of internal consistency estimates of reliability	227
7.4	Estimation of reliability	230
7.4.1	Test-retest reliability: Stability across time	231
7.4.2	Intraclass correlations and the reliability of ratings across judges	232
7.4.3	Generalizability theory: reliability over facets	236
7.4.4	Reliability of a composite test	236
7.4.5	Reliability of a difference score	237
7.5	Using reliability to estimate true scores	238
8	The “New Psychometrics” – Item Response Theory	241
8.1	Dichotomous items and monotonic trace lines: the measurement of ability	242
8.1.1	Rasch Modeling - one parameter IRT	242
8.1.2	The normal ogive – another one parameter models	248
8.1.3	Parameter estimation	249
8.1.4	Item information	250
8.1.5	Two parameter models	251
8.1.6	Three parameter models	252
8.1.7	Four parameter models	253
8.2	Polytomous items	256
8.2.1	Ordered response categories	256
8.2.2	Multiple choice ability items	257
8.3	IRT and factor analysis of items	258
8.4	Test bias and Differential Item Functioning	260
8.5	Non-monotone trace lines – the measurement of attitudes	261
8.5.1	Unfolding theory revisited	261
8.5.2	Political choice	261
8.6	IRT and adaptive testing	261
8.7	Item banking and item comparison	263