

Contents

Preface	xiii
1 Introduction	1
Review of Simple Cases of Null-Hypothesis Significance Testing	1
Statistically Signifying and Practical Significance	3
Definition of Effect Size	4
Controversy About Null-Hypothesis Significance Testing	4
The Purpose of This Book and the Need for a Broad Approach	6
Power Analysis	7
Meta-Analysis	8
Assumptions of Test Statistics and Effect Sizes	9
Violation of Assumptions in Real Data	10
Exploring the Data for a Possible Effect of a Treatment on Variability	14
Worked Examples of Measures of Variability	19
Questions	21
2 Confidence Intervals for Comparing the Averages of Two Groups	23
Introduction	23
Confidence Intervals for $\mu_a - \mu_b$: Independent Groups	24
Worked Example for Independent Groups	29
Further Discussions and Methods	31
Solutions to Violations of Assumptions: Welch's Approximate Method	32
Worked Example of the Welch Method	34

	Yuen's Confidence Interval for the Difference Between Two Trimmed Means	36
	Other Methods for Independent Groups	40
	Dependent Groups	43
	Questions	46
3	The Standardized Difference Between Means	48
	Unfamiliar and Incomparable Scales	48
	Standardized Difference Between Means:	
	Assuming Normality and a Control Group	49
	Equal or Unequal Variances	53
	Tentative Recommendations	55
	Additional Standardized-Difference Effect Sizes	
	When There Are Outliers	57
	Technical Note 3.1: A Nonparametric Estimator of Standardized-Difference Effect Sizes	58
	Confidence Intervals for a Standardized-Difference Effect Size	59
	Confidence Intervals Using Noncentral Distributions	64
	The Counternull Effect Size	65
	Dependent Groups	67
	Questions	68
4	Correlational Effect Sizes for Comparing Two Groups	70
	The Point-Biserial Correlation	70
	Example of r_{pb}	71
	Confidence Intervals and Null-Counternull Intervals for r_{pop}	72
	Assumptions of r and r_{pb}	73
	Unequal Sample Sizes	76
	Unreliability	76
	Restricted Range	81
	Small, Medium, and Large Effect Size Values	85
	Binomial Effect Size Display	87
	Limitations of the BESD	89
	The Coefficient of Determination	91
	Questions	95
5	Effect Size Measures That Go Beyond Comparing Two Centers	98
	The Probability of Superiority:	
	Independent Groups	98

Example of the <i>PS</i>	101	
A Related Measure of Effect Size	103	
Assumptions	103	
The Common Language Effect Size Statistic	105	
Technical Note 5.1: The <i>PS</i> and its Estimators	106	
Introduction to Overlap	106	
The Dominance Measure	107	
Cohen's U_3	108	
Relationships Among Measures of Effect Size	109	
Application to Cultural Effect Size	110	
Technical Note 5.2: Estimating Effect Sizes		
Throughout a Distribution	111	
Hedges–Friedman Method	112	
Shift–Function Method	112	
Other Graphical Estimators of Effect Sizes	113	
Dependent Groups	114	
Questions	115	
6 Effect Sizes for One-Way ANOVA Designs		117
Introduction	117	
ANOVA Results for This Chapter	117	
A Standardized–Difference Measure of Overall		
Effect Size	118	
A Standardized Overall Effect Size		
Using All Means	119	
Strength of Association	120	
Eta Squared (η^2)	121	
Epsilon Squared (ϵ^2) and Omega Squared (ω^2)	121	
Strength of Association for Specific Comparisons	123	
Evaluation of Criticisms of Estimators of Strength		
of Association	124	
Standardized–Difference Effect Sizes for Two		
of k Means at a Time	127	
Worked Examples	128	
Statistical Significance, Confidence Intervals,		
and Robustness	129	
Within–Groups Designs and Further Reading	134	
Questions	137	
7 Effect Sizes for Factorial Designs		139
Introduction	139	
Strength of Association: Proportion of Variance		
Explained	140	

Partial $\hat{\omega}^2$ 141
 Comparing Values of $\hat{\omega}^2$ 142
 Ratios of Estimates of Effect Size 143
 Designs and Results for This Chapter 144
 Manipulated Factors Only 146
 Manipulated Targeted Factor and Intrinsic
 Peripheral Factor 148
 Illustrative Worked Examples 150
 Comparisons of Levels of a Manipulated Factor
 at One Level of a Peripheral Factor 153
 Targeted Classificatory Factor and Extrinsic
 Peripheral Factor 155
 Classificatory Factors Only 156
 Statistical Inference and Further Reading 160
 Within-Groups Factorial Designs 162
 Additional Designs and Measures 165
 Limitations and Recommendations 166
 Questions 167

8 Effect Sizes for Categorical Variables 170

Background Review 170
 Chi-Square Test and *Phi* 173
 Null-Counter null Interval for Φ_{pop} 176
 The Difference Between Two Proportions 177
 Approximate Confidence Interval for $P_1 - P_2$ 182
 Relative Risk and the Number Needed to Treat 183
 The Odds Ratio 188
 Construction of Confidence Intervals for OR_{pop} 191
 Tables Larger Than 2×2 193
 Odds Ratios for Large $r \times c$ Tables 195
 Multiway Tables 196
 Recommendations 196
 Questions 198

9 Effect Sizes for Ordinal Categorical Variables 200

Introduction 200
 The Point-Biserial r Applied to Ordinal Categorical
 Data 202
 Confidence Interval and Null-Counter null Interval
 for r_{pop} 203
 Limitations of r_{pb} for Ordinal Categorical Data 203
 The Probability of Superiority Applied to Ordinal
 Data 205

Worked Example of Estimating the <i>PS</i> From Ordinal Data	206
The Dominance Measure and Somers' <i>D</i>	211
Worked Example of the <i>ds</i>	213
Generalized Odds Ratio	213
Cumulative Odds Ratio	214
The <i>Phi</i> Coefficient	216
A Caution	216
References for Further Discussion of Ordinal Categorical Methods	217
Questions	217
References	219
Author Index	237
Subject Index	245