
Table of Contents

Acknowledgments	xii
PART I: Preliminaries	1
Chapter 1	
Introduction	3
1.1 What is Inter-Rater Reliability?	4
1.2 Defining Experimental Parameters	10
1.3 Formulation of Agreement Coefficients	14
1.4 Different Reliability Types	17
1.5 Statistical Inference	20
1.6 Book's Structure	21
1.7 Choosing the Right Method	23
PART II: Chance-Corrected Agreement Coefficients	25
Chapter 2	
Agreement Coefficients for Nominal Ratings: A Review	27
2.1 The Problem	28
2.2 Agreement for two Raters and two Categories	31
2.3 Agreement for two Raters and a Multiple-Level Scale	41
2.4 Agreement for Multiple Raters on a Multiple-Level Scale	48
2.5 The Kappa Coefficient and Its Paradoxes	57
2.6 Weighting of the Kappa Coefficient	62
2.7 More Alternative Agreement Coefficients	65
2.8 Concluding Remarks	69
Chapter 3	
Agreement Coefficients for Ordinal, Interval, and Ratio Data	73
3.1 Overview	74
3.2 Generalizing Kappa in the Context of two Raters	75
3.3 Agreement Coefficients for Interval Data: the Case of two Raters ..	82

3.4 Agreement Coefficients for Interval Data: the Case of three Raters or More	84
3.5 More Weighting Options for Agreement Coefficients	91
3.6 Concluding Remarks	98

Chapter 4

Constructing Agreement Coefficients: AC_1 and Aickin's α	101
4.1 Overview	102
4.2 Gwet's AC_1 and Aickin's α for two Raters	104
4.3 Aickin's Theory	108
4.4 Gwet's Theory	112
4.5 Calculating AC_1 for three Raters or More	118
4.6 AC_2 : the AC_1 Coefficient for Ordinal and Interval Data	121
4.7 Concluding Remarks	127

Chapter 5

Agreement Coefficients and Statistical Inference	129
5.1 The problem	130
5.2 Finite Population Inference in Inter-Rater Reliability Analysis ...	133
5.3 Conditional Inference	138
5.4 Unconditional Inference	155
5.5 Sample Size Estimation	158
5.6 Concluding Remarks	161

Chapter 6

Benchmarking Inter-Rater Reliability Coefficients	163
6.1 Overview	164
6.2 Benchmarking the Agreement Coefficient	165
6.3 The Proposed Benchmarking Method	173
6.4 Concluding Remarks	180

PART III: Intraclass Correlation Coefficients 183

Chapter 7

Intraclass Correlation: A Measure of Raters' Agreement	185
7.1 Introduction	186
7.2 Statistical Models	186
7.3 The Bland-Altman Plot	189
7.4 Sample Size Calculations	192

Chapter 8

Intraclass Correlation in One-Factor Studies.....195

8.1 Intraclass Correlation under Model 1A196

8.2 Intraclass Correlation under Model 1B201

8.3 Statistical Inference about ICC under Models 1A and 1B206

8.4 Concluding Remarks223

Chapter 9

Intraclass Correlations under the Random Factorial Design225

9.1 The Issues226

9.2 The Intraclass Correlation Coefficients228

9.3 Statistical Inference about the ICC238

9.4 Sample Size Calculations247

9.5 Special Topics256

Chapter 10

Intraclass Correlations under the Mixed Factorial Design269

10.1 The Problem270

10.2 Intraclass Correlation Coefficient271

10.3 Statistical Inference About the ICC281

10.4 Sample Size Calculations290

10.5 Special Topics302

**PART IV: MISCELLANEOUS TOPICS ON THE ANALYSIS
OF INTER-RATER RELIABILITY
EXPERIMENTS309**

Chapter 11

Inter-Rater Reliability: Conditional Analysis311

11.1 Overview312

11.2 Conditional Agreement Coefficient for two Raters in
ACM Reliability Studies314

11.3 Validity and Conditional Coefficients for three Raters
or More in ACM Studies326

11.4 Conditional Agreement Coefficient for two Raters in
RCM Reliability Studies334

11.5 Concluding Remarks341

Chapter 12

Measures of Association and Item Analysis 343

12.1 Overview 344

12.2 Cronbach’s Alpha 344

12.3 Pearson & Spearman Correlation Coefficients 351

12.4 Kendall’s Tau 356

12.5 Kendall’s Coefficient of Concordance (KCC) 360

12.6 Concluding Remarks 364

PART V: Appendices 367

Appendix A: Data Tables 369

Appendix B: Software Solutions 375

B.1 The R Software 375

B.2 AgreeStat for Excel 386

B.3 Online Calculators 387

B.4 SAS Software 388

B.5 SPSS & STATA 388

B.6 Concluding Remarks 389

Bibliography 391

List of Notations 399

Author index 404

Subject index 407