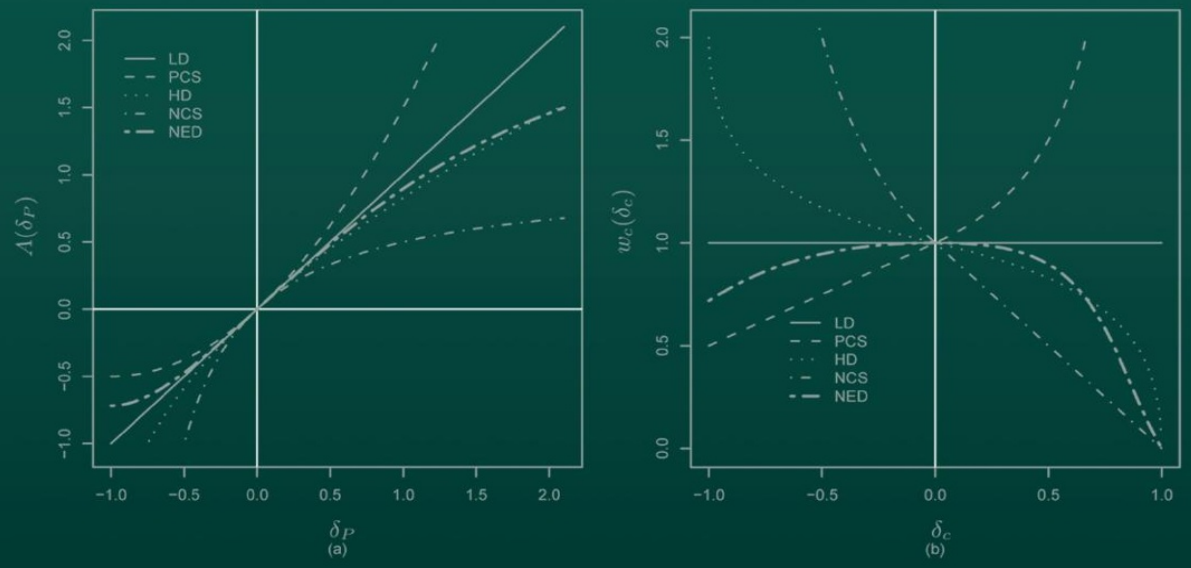
Monographs on Statistics and Applied Probability 120

Statistical Inference

The Minimum Distance Approach



Ayanendranath Basu
Hiroyuki Shioya
Chanseok Park

CRC Press
Taylor & Francis Group
A CHAPMAN & HALL BOOK

Monographs on Statistics and Applied Probability 120

Statistical Inference

The Minimum Distance Approach

Ayanendranath Basu
Indian Statistical Institute
Kolkata, India

Hiroyuki Shioya
Muroran Institute of Technology
Muroran, Japan

Chanseok Park
Clemson University
Clemson, South Carolina, USA

CRC Press
Taylor & Francis Group
Boca Raton London New York

CRC Press is an imprint of the
Taylor & Francis Group an **informa** business

A CHAPMAN & HALL BOOK

Contents

Preface **xv**

Acknowledgments **xix**

1 Introduction **1**
1.1 General Notation . . . 3
1.2 Illustrative Examples . . . 4
1.3 Some Background and Relevant Definitions . . . 7
1.3.1 Fisher Information . . . 7
1.3.2 First-Order Efficiency . . . 9
1.3.3 Second-Order Efficiency . . . 9
1.4 Parametric Inference Based on the Maximum Likelihood Method . . . 10
1.4.1 Hypothesis Testing by Likelihood Methods . . . 11
1.5 Statistical Functionals and Influence Function . . . 14
1.6 Outline of the Book . . . 18

2 Statistical Distances **21**
2.1 Introduction . . . 21
2.2 Distances Based on Distribution Functions . . . 22
2.3 Density-Based Distances . . . 25
2.3.1 The Distances in Discrete Models . . . 26
2.3.2 More on the Hellinger Distance . . . 33
2.3.3 The Minimum Distance Estimator and the Estimating Equations . . . 34
2.3.4 The Estimation Curvature . . . 38
2.3.5 Controlling of Inliers . . . 39
2.3.6 The Robustified Likelihood Disparity . . . 40
2.3.7 The Influence Function of the Minimum Distance Estimators . . . 43
2.3.8 ϕ-Divergences . . . 45
2.4 Minimum Hellinger Distance Estimation: Discrete Models . . . 46
2.4.1 Consistency of the Minimum Hellinger Distance Estimator . . . 47
2.4.2 Asymptotic Normality of the Minimum Hellinger Distance Estimator . . . 52

2.5 Minimum Distance Estimation Based on Disparities: Discrete Models . . . 55
2.6 Some Examples . . . 67

3 Continuous Models **73**
3.1 Introduction . . . 73
3.2 Minimum Hellinger Distance Estimation . . . 75
3.2.1 The Minimum Hellinger Distance Functional . . . 75
3.2.2 The Asymptotic Distribution of the Minimum Hellinger Distance Estimator . . . 78
3.3 Estimation of Multivariate Location and Covariance . . . 83
3.4 A General Structure . . . 87
3.4.1 Disparities in This Class . . . 93
3.5 The Basu–Lindsay Approach for Continuous Data . . . 94
3.5.1 Transparent Kernels . . . 98
3.5.2 The Influence Function of the Minimum Distance Estimators for the Basu–Lindsay Approach . . . 100
3.5.3 The Asymptotic Distribution of the Minimum Distance Estimators . . . 102
3.6 Examples . . . 107

4 Measures of Robustness and Computational Issues **115**
4.1 The Residual Adjustment Function . . . 116
4.2 The Graphical Interpretation of Robustness . . . 118
4.3 The Generalized Hellinger Distance . . . 126
4.3.1 Connection with Other Distances . . . 129
4.4 Higher Order Influence Analysis . . . 129
4.5 Higher Order Influence Analysis: Continuous Models . . . 136
4.6 Asymptotic Breakdown Properties . . . 137
4.6.1 Breakdown Point of the Minimum Hellinger Distance Estimator . . . 137
4.6.2 The Breakdown Point for the Power Divergence Family 139
4.6.3 A General Form of the Breakdown Point . . . 141
4.6.4 Breakdown Point for Multivariate Location and Covariance Estimation . . . 144
4.7 The $\boldsymbol{\alpha}$-Influence Function . . . 147
4.8 Outlier Stability of Minimum Distance Estimators . . . 149
4.8.1 Outlier Stability of the Estimating Functions . . . 152
4.8.2 Robustness of the Estimator . . . 153
4.9 Contamination Envelopes . . . 156
4.10 The Iteratively Reweighted Least Squares (IRLS) . . . 160
4.10.1 Development of the Algorithm . . . 160
4.10.2 The Standard IREE . . . 163
4.10.3 Optimally Weighted IREE . . . 164

4.10.4 Step by Step Implementation 166

5 The Hypothesis Testing Problem 167
5.1 Disparity Difference Test: Hellinger Distance Case 167
5.2 Disparity Difference Tests in Discrete Models 172
5.2.1 Second-Order Effects in Testing 175
5.3 Disparity Difference Tests: The Continuous Case 180
5.3.1 The Smoothed Model Approach 182
5.4 Power Breakdown of Disparity Difference Tests 184
5.5 Outlier Stability of Disparity Difference Tests 186
5.5.1 The GHD and the Chi-Square Inflation Factor 189
5.6 The Two-Sample Problem 191

6 Techniques for Inlier Modification 195
6.1 Minimum Distance Estimation: Inlier Correction in Small Samples 195
6.2 Penalized Distances 197
6.2.1 The Penalized Hellinger Distance 198
6.2.2 Minimum Penalized Distance Estimators 200
6.2.3 Asymptotic Distribution of the Minimum Penalized Distance Estimator 201
6.2.4 Penalized Disparity Difference Tests: Asymptotic Results 206
6.2.5 The Power Divergence Family versus the Blended Weight Hellinger Distance Family 207
6.3 Combined Distances 212
6.3.1 Asymptotic Distribution of the Minimum Combined Distance Estimators 216
6.4 ϵ-Combined Distances 222
6.5 Coupled Distances 225
6.6 The Inlier-Shrunk Distances 227
6.7 Numerical Simulations and Examples 230

7 Weighted Likelihood Estimation 235
7.1 The Discrete Case 236
7.1.1 The Disparity Weights 237
7.1.2 Influence Function and Standard Error 242
7.1.3 The Mean Downweighting Parameter 244
7.1.4 Examples 245
7.2 The Continuous Case 249
7.2.1 Influence Function and Standard Error: Continuous Case 251
7.2.2 The Mean Downweighting Parameter 252
7.2.3 A Bootstrap Root Search 253
7.2.4 Asymptotic Results 254

7.2.5 Robustness of Estimating Equations 255
7.3 Examples 256
7.4 Hypothesis Testing 261
7.5 Further Reading 263

8 Multinomial Goodness-of-Fit Testing 265
8.1 Introduction 265
8.1.1 Chi-Square Goodness-of-Fit Tests 266
8.2 Asymptotic Distribution of the Goodness-of-Fit Statistics . . 267
8.2.1 The Disparity Statistics 268
8.2.2 The Simple Null Hypothesis 268
8.2.3 The Composite Null Hypothesis 270
8.2.4 Minimum Distance Inference versus Multinomial Goodness-of-Fit 272
8.3 Exact Power Comparisons in Small Samples 273
8.4 Choosing a Disparity to Minimize the Correction Terms . . . 277
8.5 Small Sample Comparisons of the Test Statistics 280
8.5.1 The Power Divergence Family 280
8.5.2 The BWHD Family 282
8.5.3 The BWCS Family 283
8.5.4 Derivation of $F_S(y)$ for a General Disparity Statistic . 284
8.6 Inlier Modified Statistics 286
8.6.1 The Penalized Disparity Statistics 287
8.6.2 The Combined Disparity Statistics 288
8.6.3 Numerical Studies 290
8.7 An Application: Kappa Statistics 294

9 The Density Power Divergence 297
9.1 The Minimum L_2 Distance Estimator 298
9.2 The Minimum Density Power Divergence Estimator 300
9.2.1 Asymptotic Properties 303
9.2.2 Influence Function and Standard Error 308
9.2.3 Special Parametric Families 309
9.3 A Related Divergence Measure 311
9.3.1 The JHHB Divergence 311
9.3.2 Formulae for Variances 314
9.3.3 Numerical Comparisons of the Two Methods 316
9.3.4 Robustness 316
9.4 The Censored Survival Data Problem 317
9.4.1 A Real Data Example 318
9.5 The Normal Mixture Model Problem 322
9.6 Selection of Tuning Parameters 323
9.7 Other Applications of the Density Power Divergence 324

10 Other Applications **327**
10.1 Censored Data 327
10.1.1 Minimum Hellinger Distance Estimation in the Random Censorship Model 327
10.1.2 Minimum Hellinger Distance Estimation Based on Hazard Functions 329
10.1.3 Power Divergence Statistics for Grouped Survival Data 330
10.2 Minimum Hellinger Distance Methods in Mixture Models . . 331
10.3 Minimum Distance Estimation Based on Grouped Data . . . 332
10.4 Semiparametric Problems 335
10.4.1 Two-Component Mixture Model 335
10.4.2 Two-Sample Semiparametric Model 336
10.5 Other Miscellaneous Topics 337

11 Distance Measures in Information and Engineering **339**
11.1 Introduction 339
11.2 Entropies and Divergences 340
11.3 Csiszár's f-Divergence 341
11.3.1 Definition 341
11.3.2 Range of the f-Divergence 343
11.3.3 Inequalities Involving f-Divergences 345
11.3.4 Other Related Results 346
11.4 The Bregman Divergence 346
11.5 Extended f-Divergences 347
11.5.1 f-Divergences for Nonnegative Functions 347
11.5.2 Another Extension of the f-Divergence 351
11.6 Additional Remarks 352

12 Applications to Other Models **353**
12.1 Introduction 353
12.2 Preliminaries for Other Models 354
12.3 Neural Networks 356
12.3.1 Models and Previous Works 356
12.3.2 Feed-Forward Neural Networks 356
12.3.3 Training Feed-Forward Neural Networks 357
12.3.4 Numerical Examples 360
12.3.5 Related Works 360
12.4 Fuzzy Theory 361
12.4.1 Fundamental Elements of Fuzzy Sets 361
12.4.2 Measures of Fuzzy Sets 362
12.4.3 Generalized Fuzzy Divergence 364
12.5 Phase Retrieval 365
12.5.1 Diffractive Imaging 365
12.5.2 Algorithms for Phase Retrieval 367
12.5.3 Statistical-Distance-Based Phase Retrieval Algorithm 368

12.5.4 Numerical Example 369
12.6 Summary 371

Bibliography **373**

Index **403**