

1	Introduction	1
1.1	A Brief Synopsis	1
1.2	An Informal Description	6
1.2.1	The Fisher Metric and the Amari–Chentsov Structure for Finite Sample Spaces	7
1.2.2	Infinite Sample Spaces and Functional Analysis	8
1.2.3	Parametric Statistics	10
1.2.4	Exponential and Mixture Families from the Perspective of Differential Geometry	14
1.2.5	Information Geometry and Information Theory	15
1.3	Historical Remarks	17
1.4	Organization of this Book	20
2	Finite Information Geometry	25
2.1	Manifolds of Finite Measures	25
2.2	The Fisher Metric	29
2.3	Gradient Fields	39
2.4	The m - and e -Connections	42
2.5	The Amari–Chentsov Tensor and the α -Connections	47
2.5.1	The Amari–Chentsov Tensor	47
2.5.2	The α -Connections	50
2.6	Congruent Families of Tensors	52
2.7	Divergences	68
2.7.1	Gradient-Based Approach	68
2.7.2	The Relative Entropy	70
2.7.3	The α -Divergence	73
2.7.4	The f -Divergence	76
2.7.5	The q -Generalization of the Relative Entropy	78

2.8	Exponential Families	79
2.8.1	Exponential Families as Affine Spaces	79
2.8.2	Implicit Description of Exponential Families	84
2.8.3	Information Projections	91
2.9	Hierarchical and Graphical Models	100
2.9.1	Interaction Spaces	101
2.9.2	Hierarchical Models	108
2.9.3	Graphical Models	112
3	Parametrized Measure Models	121
3.1	The Space of Probability Measures and the Fisher Metric	121
3.2	Parametrized Measure Models	135
3.2.1	The Structure of the Space of Measures	139
3.2.2	Tangent Fibration of Subsets of Banach Manifolds	140
3.2.3	Powers of Measures	143
3.2.4	Parametrized Measure Models and k -Integrability	150
3.2.5	Canonical n -Tensors of an n -Integrable Model	164
3.2.6	Signed Parametrized Measure Models	168
3.3	The Pistone–Sempi Structure	170
3.3.1	e -Convergence	170
3.3.2	Orlicz Spaces	172
3.3.3	Exponential Tangent Spaces	176
4	The Intrinsic Geometry of Statistical Models	185
4.1	Extrinsic Versus Intrinsic Geometric Structures	185
4.2	Connections and the Amari–Chentsov Structure	189
4.3	The Duality Between Exponential and Mixture Families	201
4.4	Canonical Divergences	210
4.4.1	Dual Structures via Divergences	210
4.4.2	A General Canonical Divergence	213
4.4.3	Recovering the Canonical Divergence of a Dually Flat Structure	215
4.4.4	Consistency with the Underlying Dualistic Structure	217
4.5	Statistical Manifolds and Statistical Models	219
4.5.1	Statistical Manifolds and Isostatistical Immersions	220
4.5.2	Monotone Invariants of Statistical Manifolds	223
4.5.3	Immersion of Compact Statistical Manifolds into Linear Statistical Manifolds	226
4.5.4	Proof of the Existence of Isostatistical Immersions	228
4.5.5	Existence of Statistical Embeddings	238

5	Information Geometry and Statistics	241
5.1	Congruent Embeddings and Sufficient Statistics	241
5.1.1	Statistics and Congruent Embeddings	244
5.1.2	Markov Kernels and Congruent Markov Embeddings	253
5.1.3	Fisher–Neyman Sufficient Statistics	261
5.1.4	Information Loss and Monotonicity	263
5.1.5	Chentsov’s Theorem and Its Generalization	268
5.2	Estimators and the Cramér–Rao Inequality	277
5.2.1	Estimators and Their Bias, Mean Square Error, Variance	277
5.2.2	A General Cramér–Rao Inequality	281
5.2.3	Classical Cramér–Rao Inequalities	286
5.2.4	Efficient Estimators and Consistent Estimators	287
6	Fields of Application of Information Geometry	295
6.1	Complexity of Composite Systems	295
6.1.1	A Geometric Approach to Complexity	296
6.1.2	The Information Distance from Hierarchical Models	298
6.1.3	The Weighted Information Distance	307
6.1.4	Complexity of Stochastic Processes	317
6.2	Evolutionary Dynamics	327
6.2.1	Natural Selection and Replicator Equations	328
6.2.2	Continuous Time Limits	333
6.2.3	Population Genetics	336
6.3	Monte Carlo Methods	348
6.3.1	Langevin Monte Carlo	350
6.3.2	Hamiltonian Monte Carlo	351
6.4	Infinite-Dimensional Gibbs Families	354
Appendix A Measure Theory		361
Appendix B Riemannian Geometry		367
Appendix C Banach Manifolds		381
References		387
Index		397
Nomenclature		403