

Probability Theory and Statistical Inference

Doubt over the trustworthiness of published empirical results is not unwarranted and is often a result of statistical misspecification: invalid probabilistic assumptions imposed on data. Now in its second edition, this bestselling textbook offers a comprehensive course in empirical research methods, teaching the probabilistic and statistical foundations that enable the specification and validation of statistical models, providing the basis for an informed implementation of statistical procedure to secure the trustworthiness of evidence. Each chapter has been thoroughly updated, accounting for developments in the field and the author's own research. The comprehensive scope of the textbook has been expanded by the addition of a new chapter on the Linear Regression and related statistical models. This new edition is now more accessible to students of disciplines beyond economics and includes more pedagogical features, with an increased number of examples as well as review questions and exercises at the end of each chapter.

ARIS SPANOS is Wilson E. Schmidt Professor of Economics at Virginia Polytechnic Institute and State University. He is the author of *Statistical Foundations of Econometric Modelling* (Cambridge, 1986) and, with D. G. Mayo, *Error and Inference: Recent Exchanges on Experimental Reasoning, Reliability, and the Objectivity and Rationality of Science* (Cambridge, 2010).

Probability Theory and Statistical Inference

Empirical Modeling with
Observational Data

Second Edition

Aris Spanos

Virginia Tech

(Virginia Polytechnic Institute & State University)



CAMBRIDGE
UNIVERSITY PRESS

Cambridge University Press & Assessment
 978-1-107-18514-2 — Probability Theory and Statistical Inference
 Empirical Modeling with Observational Data 2nd Edition
 Aris Spanos
 Frontmatter
[More Information](#)



CAMBRIDGE
 UNIVERSITY PRESS

Shaftesbury Road, Cambridge CB2 8EA, United Kingdom

One Liberty Plaza, 20th Floor, New York, NY 10006, USA

477 Williamstown Road, Port Melbourne, VIC 3207, Australia

314–321, 3rd Floor, Plot 3, Splendor Forum, Jasola District Centre, New Delhi – 110025, India

103 Penang Road, #05–06/07, Visioncrest Commercial, Singapore 238467

Cambridge University Press is part of Cambridge University Press & Assessment,
 a department of the University of Cambridge.

We share the University's mission to contribute to society through the pursuit of
 education, learning and research at the highest international levels of excellence.

www.cambridge.org

Information on this title: www.cambridge.org/9781107185142

DOI: 10.1017/9781316882825

© Aris Spanos 2019

This publication is in copyright. Subject to statutory exception and to the provisions
 of relevant collective licensing agreements, no reproduction of any part may take
 place without the written permission of Cambridge University Press & Assessment.

First published 1999

Third printing 2007

Second edition 2019

A catalogue record for this publication is available from the British Library

Library of Congress Cataloging-in-Publication data

Names: Spanos, Aris, 1952– author.

Title: Probability theory and statistical inference : empirical modelling
 with observational data / Aris Spanos (Virginia College of Technology).

Description: Cambridge ; New York, NY : Cambridge University Press, 2019. |

Includes bibliographical references and index.

Identifiers: LCCN 2019008498 (print) | LCCN 2019016182 (ebook) | ISBN 9781107185142 | ISBN

Subjects: LCSH: Probabilities – Textbooks. | Mathematical
 statistics – Textbooks.

Classification: LCC QA273 (ebook) | LCC QA273 .S6875 2019 (print) | DDC 519.5–dc23

LC record available at <https://lcn.loc.gov/2019008498>

ISBN 978-1-107-18514-2 Hardback

ISBN 978-1-316-63637-4 Paperback

Cambridge University Press & Assessment has no responsibility for the persistence
 or accuracy of URLs for external or third-party internet websites referred to in this
 publication and does not guarantee that any content on such websites is, or will
 remain, accurate or appropriate.

Cambridge University Press & Assessment
978-1-107-18514-2 — Probability Theory and Statistical Inference
Empirical Modeling with Observational Data 2nd Edition
Aris Spanos
Frontmatter
[More Information](#)

To my grandchildren Nicholas, Jason, and Evie,
my daughters Stella, Marina, and Alexia, and my
wife Evie for their unconditional love and support

Contents

<i>Preface to the Second Edition</i>	<i>page xix</i>
1 An Introduction to Empirical Modeling	1
1.1 Introduction	1
1.2 Stochastic Phenomena: A Preliminary View	3
1.2.1 Chance Regularity Patterns	3
1.2.2 From Chance Regularities to Probabilities	7
1.2.3 Chance Regularity Patterns and Real-World Phenomena	11
1.3 Chance Regularities and Statistical Models	12
1.4 Observed Data and Empirical Modeling	14
1.4.1 Experimental vs. Observational Data	14
1.4.2 Observed Data and the Nature of a Statistical Model	15
1.4.3 Measurement Scales and Data	16
1.4.4 Measurement Scale and Statistical Analysis	18
1.4.5 Cross-Section vs. Time Series, is that the Question?	20
1.4.6 Limitations of Economic Data	22
1.5 Statistical Adequacy	23
1.6 Statistical vs. Substantive Information*	25
1.7 Looking Ahead	27
1.8 Questions and Exercises	28
2 Probability Theory as a Modeling Framework	30
2.1 Introduction	30
2.1.1 Primary Objective	30
2.1.2 Descriptive vs. Inferential Statistics	30
2.2 Simple Statistical Model: A Preliminary View	32
2.2.1 The Basic Structure of a Simple Statistical Model	33
2.2.2 The Notion of a Random Variable: A Naive View	34
2.2.3 Density Functions	35
2.2.4 A Random Sample: A Preliminary View	36
2.3 Probability Theory: An Introduction	40
2.3.1 Outlining the Early Milestones of Probability Theory	40
2.3.2 Probability Theory: A Modeling Perspective	42
2.4 A Simple Generic Stochastic Mechanism	42
2.4.1 The Notion of a Random Experiment	42
2.4.2 A Bird’s-Eye View of the Unfolding Story	44

viii	Contents	
2.5	Formalizing Condition [a]: The Outcomes Set	45
2.5.1	The Concept of a Set in Set Theory	45
2.5.2	The Outcomes Set	45
2.5.3	Special Types of Sets	46
2.6	Formalizing Condition [b]: Events and Probabilities	48
2.6.1	Set-Theoretic Operations	48
2.6.2	Events vs. Outcomes	51
2.6.3	Event Space	51
2.6.4	A Digression: What is a Function?	58
2.6.5	The Mathematical Notion of Probability	59
2.6.6	Probability Space $(\mathbf{S}, \mathfrak{I}, \mathbb{P}(.))$	63
2.6.7	Mathematical Deduction	64
2.7	Conditional Probability and Independence	65
2.7.1	Conditional Probability and its Properties	65
2.7.2	The Concept of Independence Among Events	69
2.8	Formalizing Condition [c]: Sampling Space	70
2.8.1	The Concept of Random Trials	70
2.8.2	The Concept of a Statistical Space	72
2.8.3	The Unfolding Story Ahead	74
2.9	Questions and Exercises	75
3	The Concept of a Probability Model	78
3.1	Introduction	78
3.1.1	The Story So Far and What Comes Next	78
3.2	The Concept of a Random Variable	79
3.2.1	The Case of a Finite Outcomes Set: $S = \{s_1, s_2, \dots, s_n\}$	80
3.2.2	Key Features of a Random Variable	81
3.2.3	The Case of a Countable Outcomes Set: $S = \{s_1, s_2, \dots, s_n, \dots\}$	85
3.3	The General Concept of a Random Variable	86
3.3.1	The Case of an Uncountable Outcomes Set S	86
3.4	Cumulative Distribution and Density Functions	89
3.4.1	The Concept of a Cumulative Distribution Function	89
3.4.2	The Concept of a Density Function	91
3.5	From a Probability Space to a Probability Model	95
3.5.1	Parameters and Moments	97
3.5.2	Functions of a Random Variable	97
3.5.3	Numerical Characteristics of Random Variables	99
3.5.4	Higher Moments	102
3.5.5	The Problem of Moments*	110
3.5.6	Other Numerical Characteristics	112
3.6	Summary	118
3.7	Questions and Exercises	119
	Appendix 3.A: Univariate Distributions	121

	Contents	ix
3.A.1	Discrete Univariate Distributions	121
3.A.2	Continuous Univariate Distributions	123
4	A Simple Statistical Model	130
4.1	Introduction	130
4.1.1	The Story So Far, a Summary	130
4.1.2	From Random Trials to a Random Sample: A First View	130
4.2	Joint Distributions of Random Variables	131
4.2.1	Joint Distributions of Discrete Random Variables	131
4.2.2	Joint Distributions of Continuous Random Variables	133
4.2.3	Joint Moments of Random Variables	136
4.2.4	The n Random Variables Joint Distribution	138
4.3	Marginal Distributions	139
4.4	Conditional Distributions	142
4.4.1	Conditional Probability	142
4.4.2	Conditional Density Functions	143
4.4.3	Continuous/Discrete Random Variables*	146
4.4.4	Conditional Moments	146
4.4.5	A Digression: Other Forms of Conditioning	148
4.4.6	Marginalization vs. Conditioning	150
4.4.7	Conditioning on Events vs. Random Variables	151
4.5	Independence	155
4.5.1	Independence in the Two Random Variable Case	155
4.5.2	Independence in the n Random Variable Case	156
4.6	Identical Distributions and Random Samples	158
4.6.1	Identically Distributed Random Variables	158
4.6.2	A Random Sample of Random Variables	160
4.7	Functions of Random Variables	161
4.7.1	Functions of One Random Variable	161
4.7.2	Functions of Several Random Variables	162
4.7.3	Ordered Sample and its Distributions*	165
4.8	A Simple Statistical Model	166
4.8.1	From a Random Experiment to a Simple Statistical Model	166
4.9	The Statistical Model in Empirical Modeling	167
4.9.1	The Concept of a Statistical Model: A Preliminary View	167
4.9.2	Statistical Identification of Parameters	168
4.9.3	The Unfolding Story Ahead	169
4.10	Questions and Exercises	170
	Appendix 4.A: Bivariate Distributions	171
4.A.1	Discrete Bivariate Distributions	171
4.A.2	Continuous Bivariate Distributions	172
5	Chance Regularities and Probabilistic Concepts	176
5.1	Introduction	176

x	Contents	
	5.1.1 Early Developments in Graphical Techniques	176
	5.1.2 Why Do We Care About Graphical Techniques?	177
5.2	The t-Plot and Independence	178
5.3	The t-Plot and Homogeneity	184
5.4	Assessing Distribution Assumptions	189
	5.4.1 Data that Exhibit Dependence/Heterogeneity	189
	5.4.2 Data that Exhibit Normal IID Chance Regularities	195
	5.4.3 Data that Exhibit Non-Normal IID Regularities	196
	5.4.4 The Histogram, the Density Function, and Smoothing	201
	5.4.5 Smoothed Histograms and Non-Random Samples	206
5.5	The Empirical CDF and Related Graphs*	206
	5.5.1 The Concept of the Empirical cdf (ecdf)	207
	5.5.2 Probability Plots	208
	5.5.3 Empirical Example: Exchange Rate Data	215
5.6	Summary	218
5.7	Questions and Exercises	219
	Appendix 5.A: Data – Log>Returns	220
6	Statistical Models and Dependence	222
6.1	Introduction	222
	6.1.1 Extending a Simple Statistical Model	222
6.2	Non-Random Sample: A Preliminary View	224
	6.2.1 Sequential Conditioning: Reducing the Dimensionality	225
	6.2.2 Keeping an Eye on the Forest!	227
6.3	Dependence and Joint Distributions	228
	6.3.1 Dependence Between Two Random Variables	228
6.4	Dependence and Moments	229
	6.4.1 Joint Moments and Dependence	229
	6.4.2 Conditional Moments and Dependence	232
6.5	Joint Distributions and Modeling Dependence	233
	6.5.1 Dependence and the Normal Distribution	234
	6.5.2 A Graphical Display: The Scatterplot	236
	6.5.3 Dependence and the Elliptically Symmetric Family	240
	6.5.4 Dependence and Skewed Distributions	245
	6.5.5 Dependence in the Presence of Heterogeneity	257
6.6	Modeling Dependence and Copulas*	258
6.7	Dependence for Categorical Variables	262
	6.7.1 Measurement Scales and Dependence	262
	6.7.2 Dependence and Ordinal Variables	263
	6.7.3 Dependence and Nominal Variables	266
6.8	Conditional Independence	268
	6.8.1 The Multivariate Normal Distribution	269
	6.8.2 The Multivariate Bernoulli Distribution	271
	6.8.3 Dependence in Mixed (Discrete/Continuous) Variables	272

	Contents	xi
6.9	What Comes Next?	273
6.10	Questions and Exercises	274
7	Regression Models	277
7.1	Introduction	277
7.2	Conditioning and Regression	279
7.2.1	Reduction and Conditional Moment Functions	279
7.2.2	Regression and Skedastic Functions	281
7.2.3	Selecting an Appropriate Regression Model	288
7.3	Weak Exogeneity and Stochastic Conditioning	292
7.3.1	The Concept of Weak Exogeneity	292
7.3.2	Conditioning on a σ -Field	295
7.3.3	Stochastic Conditional Expectation and its Properties	297
7.4	A Statistical Interpretation of Regression	301
7.4.1	The Statistical Generating Mechanism	301
7.4.2	Statistical vs. Substantive Models, Once Again	304
7.5	Regression Models and Heterogeneity	308
7.6	Summary and Conclusions	310
7.7	Questions and Exercises	312
8	Introduction to Stochastic Processes	315
8.1	Introduction	315
8.1.1	Random Variables and Orderings	316
8.2	The Concept of a Stochastic Process	318
8.2.1	Defining a Stochastic Process	318
8.2.2	Classifying Stochastic Processes; What a Mess!	320
8.2.3	Characterizing a Stochastic Process	322
8.2.4	Partial Sums and Associated Stochastic Processes	324
8.2.5	Gaussian (Normal) Process: A First View	328
8.3	Dependence Restrictions (Assumptions)	329
8.3.1	Distribution-Based Concepts of Dependence	329
8.3.2	Moment-Based Concepts of Dependence	330
8.4	Heterogeneity Restrictions (Assumptions)	331
8.4.1	Distribution-Based Heterogeneity Assumptions	331
8.4.2	Moment-Based Heterogeneity Assumptions	333
8.5	Building Block Stochastic Processes	335
8.5.1	IID Stochastic Processes	335
8.5.2	White-Noise Process	336
8.6	Markov and Related Stochastic Processes	336
8.6.1	Markov Process	336
8.6.2	Random Walk Processes	338
8.6.3	Martingale Processes	340
8.6.4	Martingale Difference Process	342
8.7	Gaussian Processes	345

8.7.1	AR(p) Process: Probabilistic Reduction Perspective	345
8.7.2	A Wiener Process and a Unit Root [UR(1)] Model	349
8.7.3	Moving Average [MA(q)] Process	352
8.7.4	Autoregressive vs. Moving Average Processes	353
8.7.5	The Brownian Motion Process*	354
8.8	Counting Processes*	360
8.8.1	The Poisson Process	361
8.8.2	Duration (Hazard-Based) Models	363
8.9	Summary and Conclusions	364
8.10	Questions and Exercises	367
	Appendix 8.A: Asymptotic Dependence and Heterogeneity Assumptions*	369
8.A.1	Mixing Conditions	369
8.A.2	Ergodicity	370
9	Limit Theorems in Probability	373
9.1	Introduction	373
9.1.1	Why Do We Care About Limit Theorems?	374
9.1.2	Terminology and Taxonomy	375
9.1.3	Popular Misconceptions About Limit Theorems	376
9.2	Tracing the Roots of Limit Theorems	377
9.2.1	Bernoulli's Law of Large Numbers: A First View	377
9.2.2	Early Steps Toward the Central Limit Theorem	378
9.2.3	The First SLLN	381
9.2.4	Probabilistic Convergence Modes: A First View	381
9.3	The Weak Law of Large Numbers	383
9.3.1	Bernoulli's WLLN	383
9.3.2	Poisson's WLLN	385
9.3.3	Chebyshev's WLLN	386
9.3.4	Markov's WLLN	387
9.3.5	Bernstein's WLLN	388
9.3.6	Khinchin's WLLN	389
9.4	The Strong Law of Large Numbers	390
9.4.1	Borel's (1909) SLLN	390
9.4.2	Kolmogorov's SLLN	391
9.4.3	SLLN for a Martingale	392
9.4.4	SLLN for a Stationary Process	394
9.4.5	The Law of Iterated Logarithm*	395
9.5	The Central Limit Theorem	396
9.5.1	De Moivre–Laplace CLT	397
9.5.2	Lyapunov's CLT	399
9.5.3	Lindeberg–Feller's CLT	399
9.5.4	Chebyshev's CLT	401
9.5.5	Hajek–Sidak CLT	401

	Contents	xiii
9.5.6	CLT for a Martingale	402
9.5.7	CLT for a Stationary Process	402
9.5.8	The Accuracy of the Normal Approximation	403
9.5.9	Stable and Other Limit Distributions*	404
9.6	Extending the Limit Theorems*	406
9.6.1	A Uniform SLLN*	409
9.7	Summary and Conclusions	409
9.8	Questions and Exercises	410
	Appendix 9.A: Probabilistic Inequalities	412
9.A.1	Probability	412
9.A.2	Expectation	413
	Appendix 9.B: Functional Central Limit Theorem	414
10	From Probability Theory to Statistical Inference	421
10.1	Introduction	421
10.2	Mathematical Probability: A Brief Summary	422
10.2.1	Kolmogorov’s Axiomatic Approach	422
10.2.2	Random Variables and Statistical Models	422
10.3	Frequentist Interpretation(s) of Probability	423
10.3.1	“Randomness” (Stochasticity) is a Feature of the Real World	423
10.3.2	Model-Based Frequentist Interpretation of Probability	424
10.3.3	Von Mises’ Frequentist Interpretation of Probability	426
10.3.4	Criticisms Leveled Against the Frequentist Interpretation	427
10.3.5	Kolmogorov Complexity: An Algorithmic Perspective	430
10.3.6	The Propensity Interpretation of Probability	431
10.4	Degree of Belief Interpretation(s) of Probability	432
10.4.1	“Randomness” is in the Mind of the Beholder	432
10.4.2	Degrees of Subjective Belief	432
10.4.3	Degrees of “Objective Belief”: Logical Probability	435
10.4.4	Which Interpretation of Probability?	436
10.5	Frequentist vs. Bayesian Statistical Inference	436
10.5.1	The Frequentist Approach to Statistical Inference	436
10.5.2	The Bayesian Approach to Statistical Inference	440
10.5.3	Cautionary Notes on Misleading Bayesian Claims	443
10.6	An Introduction to Frequentist Inference	444
10.6.1	Fisher and Neglected Aspects of Frequentist Statistics	444
10.6.2	Basic Frequentist Concepts and Distinctions	446
10.6.3	Estimation: Point and Interval	447
10.6.4	Hypothesis Testing: A First View	449
10.6.5	Prediction (Forecasting)	450
10.6.6	Probability vs. Frequencies: The Empirical CDF	450
10.7	Non-Parametric Inference	453
10.7.1	Parametric vs. Non-Parametric Inference	453

xiv Contents

10.7.2	Are Weaker Assumptions Preferable to Stronger Ones?	454
10.7.3	Induction vs. Deduction	457
10.7.4	Revisiting Generic Robustness Claims	458
10.7.5	Inference Based on Asymptotic Bounds	458
10.7.6	Whither Non-Parametric Modeling?	460
10.8	The Basic Bootstrap Method	461
10.8.1	Bootstrapping and Statistical Adequacy	462
10.9	Summary and Conclusions	464
10.10	Questions and Exercises	466
11	Estimation I: Properties of Estimators	469
11.1	Introduction	469
11.2	What is an Estimator?	469
11.3	Sampling Distributions of Estimators	472
11.4	Finite Sample Properties of Estimators	474
11.4.1	Unbiasedness	474
11.4.2	Efficiency: Relative vs. Full Efficiency	475
11.4.3	Sufficiency	480
11.4.4	Minimum MSE Estimators and Admissibility	485
11.5	Asymptotic Properties of Estimators	488
11.5.1	Consistency (Weak)	488
11.5.2	Consistency (Strong)	490
11.5.3	Asymptotic Normality	490
11.5.4	Asymptotic Efficiency	491
11.5.5	Properties of Estimators Beyond the First Two Moments	492
11.6	The Simple Normal Model: Estimation	493
11.7	Confidence Intervals (Interval Estimation)	498
11.7.1	Long-Run “Interpretation” of CIs	499
11.7.2	Constructing a Confidence Interval	499
11.7.3	Optimality of Confidence Intervals	501
11.8	Bayesian Estimation	502
11.8.1	Optimal Bayesian Rules	503
11.8.2	Bayesian Credible Intervals	504
11.9	Summary and Conclusions	505
11.10	Questions and Exercises	507
12	Estimation II: Methods of Estimation	510
12.1	Introduction	510
12.2	The Maximum Likelihood Method	511
12.2.1	The Likelihood Function	511
12.2.2	Maximum Likelihood Estimators	514
12.2.3	The Score Function	517
12.2.4	Two-Parameter Statistical Model	519
12.2.5	Properties of Maximum Likelihood Estimators	524

	Contents	xv
12.2.6	The Maximum Likelihood Method and its Critics	532
12.3	The Least-Squares Method	534
12.3.1	The Mathematical Principle of Least Squares	534
12.3.2	Least Squares as a Statistical Method	535
12.4	Moment Matching Principle	536
12.4.1	Sample Moments and their Properties	539
12.5	The Method of Moments	543
12.5.1	Karl Pearson's Method of Moments	543
12.5.2	The Parametric Method of Moments	544
12.5.3	Properties of PMM Estimators	546
12.6	Summary and Conclusions	547
12.7	Questions and Exercises	549
	Appendix 12.A: Karl Pearson's Approach	551
13	Hypothesis Testing	553
13.1	Introduction	553
13.1.1	Difficulties in Mastering Statistical Testing	553
13.2	Statistical Testing Before R. A. Fisher	555
13.2.1	Francis Edgeworth's Testing	555
13.2.2	Karl Pearson's Testing	556
13.3	Fisher's Significance Testing	558
13.3.1	A Closer Look at the p -value	561
13.3.2	R. A. Fisher and Experimental Design	563
13.3.3	Significance Testing: Empirical Examples	565
13.3.4	Summary of Fisher's Significance Testing	568
13.4	Neyman–Pearson Testing	569
13.4.1	N-P Objective: Improving Fisher's Significance Testing	569
13.4.2	Modifying Fisher's Testing Framing: A First View	570
13.4.3	A Historical Excursion	574
13.4.4	The Archetypal N-P Testing Framing	575
13.4.5	Significance Level α vs. the p -value	578
13.4.6	Optimality of a Neyman–Pearson Test	580
13.4.7	Constructing Optimal Tests: The N-P Lemma	586
13.4.8	Extending the Neyman–Pearson Lemma	588
13.4.9	Constructing Optimal Tests: Likelihood Ratio	591
13.4.10	Bayesian Testing Using the Bayes Factor	594
13.5	Error-Statistical Framing of Statistical Testing	596
13.5.1	N-P Testing Driven by Substantively Relevant Values	596
13.5.2	Foundational Issues Pertaining to Statistical Testing	598
13.5.3	Post-Data Severity Evaluation: An Evidential Account	600
13.5.4	Revisiting Issues Bedeviling Frequentist Testing	603
13.5.5	The Replication Crises and Severity	609
13.6	Confidence Intervals and their Optimality	610
13.6.1	Mathematical Duality Between Testing and CIs	610

xvi Contents

13.6.2	Uniformly Most Accurate CIs	612
13.6.3	Confidence Intervals vs. Hypothesis Testing	613
13.6.4	Observed Confidence Intervals and Severity	614
13.6.5	Fallacious Arguments for Using CIs	614
13.7	Summary and Conclusions	615
13.8	Questions and Exercises	617
	Appendix 13.A: Testing Differences Between Means	620
13.A.1	Testing the Difference Between Two Means	620
13.A.2	What Happens when $Var(X_{1t}) \neq Var(X_{2t})$?	621
13.A.3	Bivariate Normal Model: Paired Sample Tests	622
13.A.4	Testing the Difference Between Two Proportions	623
13.A.5	One-Way Analysis of Variance	624
14	Linear Regression and Related Models	625
14.1	Introduction	625
14.1.1	What is a Statistical Model?	625
14.2	Normal, Linear Regression Model	626
14.2.1	Specification	626
14.2.2	Estimation	628
14.2.3	Fitted Values and Residuals	633
14.2.4	Goodness-of-Fit Measures	635
14.2.5	Confidence Intervals and Hypothesis Testing	635
14.2.6	Normality and the LR Model	642
14.2.7	Testing a Substantive Model Against the Data	643
14.3	Linear Regression and Least Squares	648
14.3.1	Mathematical Approximation and Statistical Curve-Fitting	648
14.3.2	Gauss–Markov Theorem	651
14.3.3	Asymptotic Properties of OLS Estimators	653
14.4	Regression-Like Statistical Models	655
14.4.1	Gauss Linear Model	655
14.4.2	The Logit and Probit Models	655
14.4.3	The Poisson Regression-Like Model	657
14.4.4	Generalized Linear Models	657
14.4.5	The Gamma Regression-Like Model	658
14.5	Multiple Linear Regression Model	658
14.5.1	Estimation	660
14.5.2	Linear Regression: Matrix Formulation	661
14.5.3	Fitted Values and Residuals	662
14.5.4	OLS Estimators and their Sampling Distributions	665
14.6	The LR Model: Numerical Issues and Problems	666
14.6.1	The Problem of Near-Collinearity	666
14.6.2	The Hat Matrix and Influential Observations	673
14.6.3	Individual Observation Influence Measures	674

	Contents	xvii
14.7	Conclusions	675
14.8	Questions and Exercises	677
	Appendix 14.A: Generalized Linear Models	680
	14.A.1 Exponential Family of Distributions	680
	14.A.2 Common Features of Generalized Linear Models	681
	14.A.3 MLE and the Exponential Family	682
	Appendix 14.B: Data	683
15	Misspecification (M-S) Testing	685
15.1	Introduction	685
15.2	Misspecification and Inference: A First View	688
	15.2.1 Actual vs. Nominal Error Probabilities	688
	15.2.2 Reluctance to Test the Validity of Model Assumptions	691
15.3	Non-Parametric (Omnibus) M-S Tests	694
	15.3.1 The Runs M-S Test for the IID Assumptions [2]–[4]	694
	15.3.2 Kolmogorov’s M-S Test for Normality ([1])	695
15.4	Parametric (Directional) M-S Testing	697
	15.4.1 A Parametric M-S Test for Independence ([4])	697
	15.4.2 Testing Independence and Mean Constancy ([2] and [4])	698
	15.4.3 Testing Independence and Variance Constancy ([2] and [4])	700
	15.4.4 The Skewness–Kurtosis Test of Normality	700
	15.4.5 Simple Normal Model: A Summary of M-S Testing	701
15.5	Misspecification Testing: A Formalization	703
	15.5.1 Placing M-S Testing in a Proper Context	703
	15.5.2 Securing the Effectiveness/Reliability of M-S Testing	704
	15.5.3 M-S Testing and the Linear Regression Model	705
	15.5.4 The Multiple Testing (Comparisons) Issue	706
	15.5.5 Testing for t -Invariance of the Parameters	707
	15.5.6 Where do Auxiliary Regressions Come From?	707
	15.5.7 M-S Testing for Logit/Probit Models	710
	15.5.8 Revisiting Yule’s “Nonsense Correlations”	710
	15.5.9 Respecification	713
15.6	An Illustration of Empirical Modeling	716
	15.6.1 The Traditional Curve-Fitting Perspective	716
	15.6.2 Traditional ad hoc M-S Testing and Respecification	718
	15.6.3 The Probabilistic Reduction Approach	721
15.7	Summary and Conclusions	729
15.8	Questions and Exercises	731
	Appendix 15.A: Data	734
	<i>References</i>	736
	<i>Index</i>	752

Preface to the Second Edition

The original book, published 20 years ago, has been thoroughly revised with two objectives in mind. First, to make the discussion more compact and coherent by avoiding repetition and many digressions. Second, to improve the methodological coherence of the proposed empirical modeling framework by including material pertaining to foundational issues that has been published by the author over the last 20 years or so in journals on econometrics, statistics, and philosophy of science. In particular, this revised edition brings out more clearly several crucial distinctions that elucidate empirical modeling, including (a) the statistical vs. the substantive information/model, (b) the modeling vs. the inference facet of statistical analysis, (c) testing within and testing outside the boundary of a statistical model, and (d) pre-data vs. post-data error probabilities. These distinctions shed light on several foundational issues and suggest solutions. In addition, the comprehensiveness of the book has been improved by adding Chapter 14 on the linear regression and related models.

The current debates on the “*replication crises*” render the methodological framework articulated in this book especially relevant for today’s practitioner. A closer look at the debates (Mayo, 2018) reveals that the non-replicability of empirical evidence problem is, first and foremost, a problem of *untrustworthy evidence* routinely published in prestigious journals. The current focus of that literature on the abuse of significance testing is rather misplaced, because it is only a part of a much broader problem relating to the mechanical application of statistical methods without a real understanding of their assumptions, limitations, proper implementation, and interpretation of their results. The abuse and misinterpretation of the p -value is just symptomatic of the same uninformed implementation that contributes majorly to the problem of untrustworthy evidence. Indeed, the same uninformed implementation often ensures that untrustworthy evidence is routinely replicated, when the same mistakes are repeated by equally uninformed practitioners! In contrast to the current conventional wisdom, it is argued that a major contributor to the untrustworthy evidence problem is *statistical misspecification*: invalid probabilistic assumptions imposed on one’s data, another symptom of the same uninformed implementation. The primary objective of this book is to provide the necessary probabilistic foundation and the overarching modeling framework for an informed and thoughtful application of statistical methods, as well as the proper interpretation of their inferential results. The emphasis is placed less on the mechanics of the application of statistical methods, and more on understanding their assumptions, limitations, and proper interpretation.

xx Preface to the Second Edition

Key Features of the Book

- It offers a seamless integration of probability theory and statistical inference with a view to elucidating the interplay between deduction and induction in “learning from data” about observable phenomena of interest using statistical procedures.
- It develops frequentist modeling and inference from first principles by emphasizing the notion of a statistical model and its adequacy (the validity of its probabilistic assumptions vis-à-vis the particular data) as the cornerstone for reliable inductive inference and trustworthy evidence.
- It presents frequentist inference as well-grounded procedures whose optimality is assessed by their capacity to achieve genuine “learning from data.”
- It focuses primarily on the skills and the technical knowledge one needs to be able to begin with substantive questions of interest, select the relevant data carefully, and proceed to establish trustworthy evidence for or against hypotheses or claims relating to the questions of interest. These skills include understanding the statistical information conveyed by data plots, selecting appropriate statistical models, as well as validating them using misspecification testing before any inferences are drawn.
- It articulates reasoned responses to several charges leveled against several aspects of frequentist inference by addressing the underlying foundational issues, including the use and abuse of p -values and confidence intervals, Neyman–Pearson vs. Fisher testing, and inference results vs. evidence that have bedeviled frequentist inference since the 1930s. The book discusses several such foundational issues/problems and proposes ways to address them using an error statistical perspective grounded in the concept of severity. Methodological issues discussed in this book include rebuttals to widely used, ill-thought-out arguments for ignoring statistical misspecification, as well as principled responses to certain Bayesian criticisms of the frequentist approach.
- Its methodological perspective differs from the traditional textbook perspective by bringing out the perils of curve-fitting and focusing on the key question: How can empirical modeling lead to “learning from data” about phenomena of interest by giving rise to trustworthy evidence?

NOTE: All sections marked with an asterisk (*) can be skipped at first reading without any serious interruption in the flow of the discussion.

Acknowledgments

More than any other person, Deborah G. Mayo, my colleague and collaborator on many foundational issues in statistical inference, has helped to shape my views on several methodological issues addressed in this book; for that and the constant encouragement, I’m most grateful to her. I’m also thankful to Clark Glymour, the other philosopher of science with whom I had numerous elucidating and creative discussions on many philosophical issues discussed in the book. Thanks are also due to Sir David Cox for many discussions that helped me appreciate the different perspectives on frequentist inference. Special thanks are also due to my longtime collaborator, Anya McGuirk, who contributed majorly in puzzling out several thorny issues discussed in this book. I owe a special thanks to Julio Lopez for

his insightful comments, as well as his unwavering faith in the coherence and value of the proposed approach to empirical modeling. I’m also thankful to Jesse Bledsoe for helpful comments on chapter 13 and Mariusz Kamienski for invaluable help on the front cover design.

I owe special thanks to several of my former and current students over the last 20 years, who helped to improve the discussion in this book by commenting on earlier drafts and finding mistakes and typos. They include Elena Andreou, Andros Kourtellos, Carlos Elias, Maria Heracleous, Jason Bergtold, Ebere Akobundu, Andreas Koutris, Alfredo Romero, Niraj Pouydal, Michael Michaelides, Karo Solat, and Mohammad Banasaz.

Symbols

- \mathbb{N} – set of natural numbers $\mathbb{N}:=\{1, 2, ..., n, ... \}$
- \mathbb{R} – the set of real numbers; the real line $(-\infty, \infty)$
- \mathbb{R}^n $:= \overbrace{\mathbb{R} \times \mathbb{R} \times \cdots \times \mathbb{R}}^{n \text{ times}}$
- \mathbb{R}_+ – the set of positive real numbers; the half real line $(0, \infty)$
- $f(x; \theta)$ – density function of X with parameters θ
- $F(x; \theta)$ – cumulative distribution function of X with parameters θ
- $N(\mu, \sigma^2)$ – Normal distribution with mean μ and variance σ^2
- \mathcal{E} – Random Experiment (RE)
- S – outcomes set (sample space)
- \mathfrak{S} – event space (a σ –field)
- $\mathbb{P}(\cdot)$ – probability set function
- $\sigma(X)$ – minimal sigma-field generated by X

Acronyms

- AR(p) – Autoregressive model with p lags
- CAN – Consistent, Asymptotically Normal
- cdf – cumulative distribution function
- CLT – Central Limit Theorem
- ecdf – empirical cumulative distributrion function
- GM – Generating Mechanism
- IID – Indepedent and Identically Distributed
- LS – Least-Squares
- ML – Maximum Likelihood
- M-S – Mis-Specification
- N-P – Neyman-Pearson
- PMM – Parametric Method of Moments
- SLLN – Strong Law Large Numbers
- WLLN – Weak Law Large Numbers
- UMP – Uniformly Most Powerful