Contents

Preface vii
Acknowledgments ix
List of Examples xix
List of Figures xxv

Chapter 1. Introduction 1
1. Role of Statistical Demography 1
2. Guide for the Reader 4
3. Statistical Notation and Preliminaries 4

Chapter 2. Sources of Demographic Data 9
1. Populations: Open and Closed 9
2. De Facto and De Jure Populations 11
3. Censuses and Population Registers 15
4. Lexis Diagram and Classification of Events 16
5. Register Data and Epidemiologic Studies 19
  5.1. Event Histories from Registers 19
  5.2. Cohort and Case-Control Studies 19
  5.3. Advantages and Disadvantages 20
  5.4. Confounding 22
6. Sampling in Censuses and Dual System Estimation 24
   Exercises and Complements 27

Chapter 3. Sampling Designs and Inference 31
1. Simple Random Sampling 32
2. Subgroups and Ratios 35
3. Stratified Sampling 36
   3.1. Introduction 36
   3.2. Stratified Simple Random Sampling 37
   3.3. Design Effect for Stratified Simple Random Sampling 38
   3.4. Poststratification 39
4. Sampling Weights 40
   4.1. Why Weight? 40

xi
xii  Contents

4.2. Forming Weights 41
4.3. Non-Response Adjustments 43
4.4. Effect of Weighting on Precision 45
5. Cluster Sampling 46
  5.1. Introduction 46
  5.2. Single Stage Sampling with Replacement 47
  5.3. Single Stage Sampling without Replacement 47
  5.4. Multi-Stage Sampling 49
  5.5. Stratified Samples 50
6. Systematic Sampling 52
7. Distribution Theory for Sampling 53
  7.1. Central Limit Theorems 53
  7.2. The Delta Method 55
  7.3. Estimating Equations 56
8. Replication Estimates of Variance 61
  8.1. Jackknife Estimates 61
  8.2. Bootstrap Estimates 62
  8.3. Replication Weights 63
Exercises and Complements 64

Chapter 4. Waiting Times and Their Statistical Estimation 71
  1. Exponential Distribution 71
  2. General Waiting Time 76
    2.1. Hazards and Survival Probabilities 76
    2.2. Life Expectancies and Stable Populations 79
      2.2.1. Life Expectancy 79
      2.2.2. Life Table Populations and Stable Populations 81
      2.2.3. Changing Mortality 82
      2.2.4. Basics of Pension Funding 84
      2.2.5. Effect of Heterogeneity 85
    2.3. Kaplan-Meier and Nelson-Aalen Estimators 85
    2.4. Estimation Based on Occurrence-Exposure Rates 88
  3. Estimating Survival Proportions 91
  4. Childbearing as a Repeatable Event 93
    4.1. Poisson Process Model of Childbearing 93
    4.2. Summary Measures of Fertility and Reproduction 96
    4.3. Period and Cohort Fertility 101
      4.3.1. Cohort Fertility is Smoother 101
      4.3.2. Adjusting for Timing 103
      4.3.3. Effect of Parity on Pure Period Measures 104
    4.4. Multiple Births and Effect of Pregnancy on Exposure Time 106
  5. Poisson Character of Demographic Events 107
  6. Simulation of Waiting Times and Counts 109
Exercises and Complements 110
## Chapter 5. Regression Models for Counts and Survival

1. Generalized Linear Models
   1.1. Exponential Family
   1.2. Use of Explanatory Variables
   1.3. Maximum Likelihood Estimation
   1.4. Numerical Solution
   1.5. Inferences
   1.6. Diagnostic Checks

2. Binary Regression
   2.1. Interpretation of Parameters and Goodness of Fit
   2.2. Examples of Logistic Regression
   2.3. Applicability in Case-Control Studies

3. Poisson Regression
   3.1. Interpretation of Parameters
   3.2. Examples of Poisson Regression
   3.3. Standardization
   3.4. Loglinear Models for Capture-Recapture Data

4. Overdispersion and Random Effects
   4.1. Direct Estimation of Overdispersion
   4.2. Marginal Models for Overdispersion
   4.3. Random Effect Models

5. Observable Heterogeneity in Capture-Recapture Studies

6. Bilinear Models

7. Proportional Hazards Models for Survival

8. Heterogeneity and Selection by Survival


10. Simulation of the Regression Models

### Exercises and Complements

## Chapter 6. Multistate Models and Cohort-Component Book-Keeping

1. Multistate Life-Tables
   1.1. Numerical Solution Using Runge-Kutta Algorithm
   1.2. Extension to Multistate Case
   1.3. Duration-Dependent Life-Tables
      1.3.1. Heterogeneity Attributable to Duration
      1.3.2. Forms of Duration-Dependence
      1.3.3. Aspects of Computer Implementation
      1.3.4. Policy Significance of Duration-Dependence
   1.4. Nonparametric Intensity Estimation
   1.5. Analysis of Nuptiality
   1.6. A Model for Disability Insurance

2. Linear Growth Model
   2.1. Matrix Formulation
Contents

2.2. Stable Populations 183
2.3. Weak Ergodicity 185
3. Open Populations and Parametrization of Migration 186
  3.1. Open Population Systems 186
  3.2. Parametric Models 186
    3.2.1. Migrant Pool Model 187
    3.2.2. Bilinear Models 187
4. Demographic Functionals 189
5. Elementwise Aspects of the Matrix Formulation 191
6. Markov Chain Models 191
Exercises and Complements 193

Chapter 7. Approaches to Forecasting Demographic Rates 198
1. Trends, Random Walks, and Volatility 198
2. Linear Stationary Processes 201
  2.1. Properties and Modeling 202
    2.1.1. Definition and Basic Properties 202
    2.1.2. ARIMA Models 203
    2.1.3. Practical Modeling 206
  2.2. Characterization of Predictions and Prediction Errors 210
    2.2.1. Stationary Processes 210
    2.2.2. Integrated Processes 211
    2.2.3. Cross-Correlations 216
3. Handling of Nonconstant Mean 216
  3.1. Differencing 216
  3.2. Regression 218
  3.3. Structural Models 219
4. Heteroscedastic Innovations 220
  4.1. Deterministic Models of Volatility 221
  4.2. Stochastic Volatility 222
Exercises and Complements 223

Chapter 8. Uncertainty in Demographic Forecasts: Concepts,
  Issues, and Evidence 226
1. Historical Aspects of Cohort-Component Forecasting 228
  1.1. Adoption of the Cohort-Component Approach 228
  1.2. Whelpton’s Legacy 228
  1.3. Do We Know Better Now? 231
2. Dimensionality Reduction for Mortality 234
  2.1. Age-Specific Mortality 234
  2.2. Cause-Specific Mortality 236
3. Conceptual Aspects of Error Analysis 238
  3.1. Expected Error and Empirical Error 238
  3.2. Decomposing Errors 238
    3.2.1. Error Classifications 238
    3.2.2. Alternative Decompositions 240
3. Formula-Based Allocations 346
   3.1. Theoretical Construction 346
      3.1.1. Apportionment of the U.S. House of Representatives 347
      3.1.2. Rationale Behind Allocation Formulas 348
   3.2. Effect of Inaccurate Demographic Statistics 349
   3.3. Beyond Accuracy 350
4. Decision Theory and Loss Functions 351
   4.1. Introduction 351
   4.2. Decision Theory for Statistical Agencies 353
   4.3. Loss Functions for Small Area Estimates 357
   4.4. Loss Functions for Apportionment and Redistricting 359
      4.1.1. Apportionment 359
      4.1.2. Redistricting 360
   4.5. Loss Functions and Allocation of Funds 361
      4.5.1. Effects of Over- and Under-Allocation 361
      4.5.2. Formula Nonoptimality 362
      4.5.3. Optimal Data Quality with Multiple Statistics and Uses 363
5. Comparing Risks of Adjusted and Unadjusted Census Estimates 363
   5.1. Accounting for Variances of Bias Estimates 364
   5.2. Effect of Unmeasured Biases on Comparisons of Accuracy 365
6. Decision Analysis of Adjustment for Census Undercount 365
7. Cost-Benefit Analysis of Demographic Data 367
   Exercises and Complements 368

References 371
Author Index 397
Subject Index 405