Contents

Table of Contents **i**

Comprehensive List of Definitions, Lemmas, Propositions, Theorems, Corollaries, Examples and Exercises **xxiii**

Preface **1**

I Stochastic Processes in General **2**

1 Basics **3**
1.1 So, What Is a Stochastic Process? 3
1.2 Random Functions 5
1.3 Exercises 7

2 Building Processes **15**
2.1 Finite-Dimensional Distributions 15
2.2 Consistency and Extension 16

3 Building Processes by Conditioning **21**
3.1 Probability Kernels 21
3.2 Extension via Recursive Conditioning 22
3.3 Exercises 25

II One-Parameter Processes in General **26**

4 One-Parameter Processes **27**
4.1 One-Parameter Processes 27
4.2 Operator Representations of One-Parameter Processes 31
4.3 Exercises 31

5 Stationary Processes **34**
5.1 Kinds of Stationarity 34

5.2 Strictly Stationary Processes and Measure-Preserving Transformations . . . 35
5.3 Exercises . . . 37

6 Random Times **38**
6.1 Reminders about Filtrations and Stopping Times . . . 38
6.2 Waiting Times . . . 39
6.3 Kac's Recurrence Theorem . . . 41
6.4 Exercises . . . 45

7 Continuity **46**
7.1 Kinds of Continuity for Processes . . . 46
7.2 Why Continuity Is an Issue . . . 48
7.3 Separable Random Functions . . . 50
7.4 Exercises . . . 51

8 More on Continuity **52**
8.1 Separable Versions . . . 52
8.2 Measurable Versions . . . 57
8.3 Cadlag Versions . . . 58
8.4 Continuous Modifications . . . 59

III Markov Processes **60**

9 Markov Processes **61**
9.1 The Correct Line on the Markov Property . . . 61
9.2 Transition Probability Kernels . . . 62
9.3 The Markov Property Under Multiple Filtrations . . . 65
9.4 Exercises . . . 67

10 Markov Characterizations **70**
10.1 Markov Sequences as Transduced Noise . . . 70
10.2 Time-Evolution (Markov) Operators . . . 72
10.3 Exercises . . . 76

11 Markov Examples **78**
11.1 Probability Densities in the Logistic Map . . . 78
11.2 Transition Kernels and Evolution Operators for the Wiener Process 80
11.3 Lévy Processes and Limit Laws . . . 82
11.4 Exercises . . . 86

12 Generators **88**
12.1 Exercises . . . 93

13 Strong Markov, Martingales **95**
13.1 The Strong Markov Property 95
13.2 Martingale Problems 96
13.3 Exercises 98

14 Feller Processes **99**
14.1 Markov Families 99
14.2 Feller Processes 100
14.3 Exercises 106

15 Convergence of Feller Processes **107**
15.1 Weak Convergence of Processes with Cadlag Paths (The Skorokhod Topology) 107
15.2 Convergence of Feller Processes 109
15.3 Approximation of Ordinary Differential Equations by Markov Processes 111
15.4 Exercises 113

16 Convergence of Random Walks **114**
16.1 The Wiener Process is Feller 114
16.2 Convergence of Random Walks 116
16.2.1 Approach Through Feller Processes 117
16.2.2 Direct Approach 119
16.2.3 Consequences of the Functional Central Limit Theorem . 120
16.3 Exercises 121

IV Diffusions and Stochastic Calculus **123**

17 Diffusions and the Wiener Process **124**
17.1 Diffusions and Stochastic Calculus 124
17.2 Once More with the Wiener Process and Its Properties 126
17.3 Wiener Measure; Most Continuous Curves Are Not Differentiable 127

18 Stochastic Integrals Preview **130**
18.1 Martingale Characterization of the Wiener Process 130
18.2 A Heuristic Introduction to Stochastic Integrals 131

19 Stochastic Integrals and SDEs **133**
19.1 Integrals with Respect to the Wiener Process 133
19.2 Some Easy Stochastic Integrals, with a Moral 138
19.2.1 $\int dW$ 138
19.2.2 $\int W dW$ 139
19.3 Itô's Formula 141
19.3.1 Stratonovich Integrals 146
19.3.2 Martingale Representation 146
19.4 Stochastic Differential Equations 147

19.5 Brownian Motion, the Langevin Equation, and Ornstein-Uhlenbeck Processes . . . 153
19.6 Exercises . . . 155

20 More on SDEs **158**
20.1 Solutions of SDEs are Diffusions . . . 158
20.2 Forward and Backward Equations . . . 160
20.3 Exercises . . . 163

21 Small-Noise SDEs **164**
21.1 Convergence in Probability of SDEs to ODEs . . . 165
21.2 Rate of Convergence; Probability of Large Deviations . . . 166

22 Spectral Analysis and L_2 Ergodicity **170**
22.1 White Noise . . . 171
22.2 Spectral Representation of Weakly Stationary Procesess . . . 174
22.2.1 How the White Noise Lost Its Color . . . 180
22.3 The Mean-Square Ergodic Theorem . . . 181
22.3.1 Mean-Square Ergodicity Based on the Autocovariance . . . 181
22.3.2 Mean-Square Ergodicity Based on the Spectrum . . . 183
22.4 Exercises . . . 184

V Ergodic Theory **186**

23 Ergodic Properties and Ergodic Limits **187**
23.1 General Remarks . . . 187
23.2 Dynamical Systems and Their Invariants . . . 188
23.3 Time Averages and Ergodic Properties . . . 191
23.4 Asymptotic Mean Stationarity . . . 194
23.5 Exercises . . . 197

24 The Almost-Sure Ergodic Theorem **198**

25 Ergodicity and Metric Transitivity **204**
25.1 Metric Transitivity . . . 204
25.2 Examples of Ergodicity . . . 206
25.3 Consequences of Ergodicity . . . 207
25.3.1 Deterministic Limits for Time Averages . . . 208
25.3.2 Ergodicity and the approach to independence . . . 208
25.4 Exercises . . . 209

26 Ergodic Decomposition **210**
26.1 Preliminaries to Ergodic Decompositions . . . 210
26.2 Construction of the Ergodic Decomposition . . . 212
26.3 Statistical Aspects . . . 216
26.3.1 Ergodic Components as Minimal Sufficient Statistics . . . 216

26.3.2 Testing Ergodic Hypotheses 218
26.4 Exercises 219

27 Mixing **220**
27.1 Definition and Measurement of Mixing 221
27.2 Examples of Mixing Processes 223
27.3 Convergence of Distributions Under Mixing 223
27.4 A Central Limit Theorem for Mixing Sequences 225

VI Information Theory **227**

28 Entropy and Divergence **228**
28.1 Shannon Entropy 229
28.2 Relative Entropy or Kullback-Leibler Divergence 231
28.2.1 Statistical Aspects of Relative Entropy 232
28.3 Mutual Information 234
28.3.1 Mutual Information Function 234

29 Rates and Equipartition **236**
29.1 Information-Theoretic Rates 236
29.2 Asymptotic Equipartition 239
29.2.1 Typical Sequences 242
29.3 Asymptotic Likelihood 243
29.3.1 Asymptotic Equipartition for Divergence 243
29.3.2 Likelihood Results 243
29.4 Exercises 243

VII Large Deviations **245**

30 Large Deviations: Basics **246**
30.1 Large Deviation Principles: Main Definitions and Generalities . . 246
30.2 Breeding Large Deviations 250

31 IID Large Deviations **256**
31.1 Cumulant Generating Functions and Relative Entropy 257
31.2 Large Deviations of the Empirical Mean in $\mathbb{R}^d$ 260
31.3 Large Deviations of the Empirical Measure in Polish Spaces . . . 263
31.4 Large Deviations of the Empirical Process in Polish Spaces . . . 264

32 Large Deviations for Markov Sequences **266**
32.1 Large Deviations for Pair Measure of Markov Sequences 266
32.2 Higher LDPs for Markov Sequences 270

33 The Gärtner-Ellis Theorem **271**
33.1 The Gärtner-Ellis Theorem 271
33.2 Exercises 275

34 Freidlin-Wentzell Theory **276**
34.1 Large Deviations of the Wiener Process 277
34.2 Large Deviations for SDEs with State-Independent Noise 282
34.3 Large Deviations for State-Dependent Noise 283
34.4 Exercises 284

Bibliography **284**

Definitions, Lemmas, Propositions, Theorems, Corollaries, Examples and Exercises

Chapter 1 Basic Definitions: Indexed Collections and Random Functions **3**

Definition 1 A Stochastic Process Is a Collection of Random Variables . . . 3
Example 2 Random variables . . . 4
Example 3 Random vector . . . 4
Example 4 One-sided random sequences . . . 4
Example 5 Two-sided random sequences . . . 4
Example 6 Spatially-discrete random fields . . . 4
Example 7 Continuous-time random processes . . . 4
Example 8 Random set functions . . . 4
Example 9 One-sided random sequences of set functions . . . 4
Example 10 Empirical distributions . . . 5
Definition 11 Cylinder Set . . . 5
Definition 12 Product σ-field . . . 6
Definition 13 Random Function; Sample Path . . . 6
Definition 14 Functional of the Sample Path . . . 6
Definition 15 Projection Operator, Coordinate Map . . . 6
Theorem 16 Product σ-field-measurability is equvialent to measurability of all coordinates . . . 6
Definition 17 A Stochastic Process Is a Random Function . . . 7
Corollary 18 Measurability of constrained sample paths . . . 7
Example 19 Random Measures . . . 7
Example 20 Point Processes . . . 7
Example 21 Continuous random processes . . . 7
Exercise 1.1 The product σ-field answers countable questions . . . 7

Exercise 1.2 The product σ-field constrained to a given set of paths 7

Chapter 2 Building Infinite Processes from Finite-Dimensional Distributions 15

Definition 22 Finite-dimensional distributions 15
Theorem 23 Finite-dimensional distributions determine process distributions 16
Definition 24 Projective Family of Distributions 17
Lemma 25 FDDs Form Projective Families 17
Proposition 26 Randomization, transfer 18
Theorem 27 Daniell Extension Theorem 18
Proposition 28 Carathéodory Extension Theorem 19
Theorem 29 Kolmogorov Extension Theorem 19

Chapter 3 Building Infinite Processes from Regular Conditional Probability Distributions 21

Definition 30 Probability Kernel 21
Definition 31 Composition of probability kernels 22
Proposition 32 Set functions continuous at $\emptyset$ 23
Theorem 33 Ionescu Tulcea Extension Theorem 23
Exercise 3.1 Łomnick-Ulam Theorem on infinite product measures 25
Exercise 3.2 Measures of cylinder sets 25

Chapter 4 One-Parameter Processes, Usually Functions of Time 27

Definition 34 One-Parameter Process 28
Example 35 Bernoulli process 28
Example 36 Markov models 28
Example 37 "White Noise" (Not Really) 28
Example 38 Wiener Process 29
Example 39 Logistic Map 29
Example 40 Symbolic Dynamics of the Logistic Map 29
Example 41 IID Samples 30
Example 42 Non-IID Samples 30
Example 43 Estimating Distributions 30
Example 44 Doob's Martingale 30
Example 45 The One-Dimensional Ising Model 30
Example 46 Text 31
Example 47 Polymer Sequences 31
Definition 48 Shift Operators 31
Exercise 4.1 Existence of proto-Wiener processes 32
Exercise 4.2 Time-Evolution Semi-Group 32

Chapter 5 Stationary One-Parameter Processes 34

Definition 49 Strong Stationarity 34
Definition 50 Weak Stationarity 35
Definition 51 Conditional (Strong) Stationarity 35

Theorem 52 Stationarity is Shift-Invariance . . . 35
Definition 53 Measure-Preserving Transformation . . . 36
Corollary 54 Measure-preservation implies stationarity . . . 36
Exercise 5.1 Functions of Stationary Processes . . . 37
Exercise 5.2 Continuous Measure-Preserving Families of Transformations . . . 37
Exercise 5.3 The Logistic Map as a Measure-Preserving Transformation . . . 37

Chapter 6 Random Times and Their Properties 38
Definition 55 Filtration . . . 38
Definition 56 Adapted Process . . . 39
Definition 57 Stopping Time, Optional Time . . . 39
Definition 58 $\mathcal{F}_\tau$ for a Stopping Time τ . . . 39
Definition 59 Hitting Time . . . 40
Example 60 Fixation through Genetic Drift . . . 40
Example 61 Stock Options . . . 40
Definition 62 First Passage Time . . . 40
Definition 63 Return Time, Recurrence Time . . . 41
Proposition 64 Some Sufficient Conditions for Waiting Times to be Weakly Optional . . . 41
Lemma 65 Some Recurrence Relations for Kac's Theorem . . . 42
Theorem 66 Recurrence in Stationary Processes . . . 42
Corollary 67 Poincaré Recurrence Theorem . . . 43
Corollary 68 "Nietzsche" . . . 43
Theorem 69 Kac's Recurrence Theorem . . . 44
Example 70 Counter-example for Kac's Recurrence Theorem . . . 45
Exercise 6.1 Weakly Optional Times and Right-Continuous Filtrations . . . 45
Exercise 6.2 Discrete Stopping Times and Their σ-Algebras . . . 45
Exercise 6.3 Kac's Theorem for the Logistic Map . . . 45

Chapter 7 Continuity of Stochastic Processes 46
Definition 71 Continuity in Mean . . . 46
Definition 72 Continuity in Probability, Stochastic Continuity . . . 47
Definition 73 Continuous Sample Paths . . . 47
Definition 74 Cadlag . . . 47
Definition 75 Versions of a Stochastic Process, Stochastically Equivalent Processes . . . 47
Lemma 76 Versions Have Identical FDDs . . . 47
Definition 77 Indistinguishable Processes . . . 48
Proposition 78 Continuous Sample Paths Have Measurable Extrema 48
Example 79 A Horrible Version of the proto-Wiener Process . . . 49
Definition 80 Separable Function . . . 50
Lemma 81 Some Sufficient Conditions for Separability . . . 50
Definition 82 Separable Process . . . 51

Exercise 7.1 Piecewise Linear Paths Not in the Product σ-Field . 51
Exercise 7.2 Indistinguishability of Right-Continuous Versions . . 51

Chapter 8 More on Continuity 52
Definition 83 Compactness, Compactification 53
Proposition 84 Compact Spaces are Separable 53
Proposition 85 Compactification is Possible 53
Example 86 Compactifying the Reals 53
Lemma 87 Alternative Characterization of Separability 54
Lemma 88 Confining Bad Behavior to a Measure Zero Event . . 54
Lemma 89 A Null Exceptional Set 55
Theorem 90 Separable Versions, Separable Modification 56
Corollary 91 Separable Modifications in Compactified Spaces . . 57
Corollary 92 Separable Versions with Prescribed Distributions Exist 57
Definition 93 Measurable random function 58
Theorem 94 Exchanging Expectations and Time Integrals 58
Theorem 95 Measurable Separable Modifications 58
Theorem 96 Cadlag Versions 58
Theorem 97 Continuous Versions 59
Definition 98 Modulus of continuity 59
Lemma 99 Modulus of continuity and uniform continuity 59
Definition 100 Hölder continuity 59
Theorem 101 Hölder-continuous versions 59

Chapter 9 Markov Processes 61
Definition 102 Markov Property 61
Lemma 103 The Markov Property Extends to the Whole Future 61
Definition 104 Product of Probability Kernels 62
Definition 105 Transition Semi-Group 63
Theorem 106 Existence of Markov Process with Given Transition Kernels 63
Definition 107 Invariant Distribution 64
Theorem 108 Stationarity and Invariance for Homogeneous Markov Processes 65
Definition 109 Natural Filtration 65
Definition 110 Comparison of Filtrations 65
Lemma 111 The Natural Filtration Is the Coarsest One to Which a Process Is Adapted 65
Theorem 112 Markovianity Is Preserved Under Coarsening 66
Example 113 The Logistic Map Shows That Markovianity Is Not Preserved Under Refinement 66
Exercise 9.1 Extension of the Markov Property to the Whole Future 67
Exercise 9.2 Futures of Markov Processes Are One-Sided Markov Processes 67
Exercise 9.3 Discrete-Time Sampling of Continuous-Time Markov Processes 67

Exercise 9.4 Stationarity and Invariance for Homogeneous Markov Processes . . . 67
Exercise 9.5 Rudiments of Likelihood-Based Inference for Markov Chains . . . 68
Exercise 9.6 Implementing the MLE for a Simple Markov Chain . 68
Exercise 9.7 The Markov Property and Conditional Independence from the Immediate Past . . . 68
Exercise 9.8 Higher-Order Markov Processes . . . 69
Exercise 9.9 AR(1) Models . . . 69

Chapter 10 Alternative Characterizations of Markov Processes 70
Theorem 114 Markov Sequences as Transduced Noise . . . 70
Definition 115 Transducer . . . 71
Definition 116 Markov Operator on Measures . . . 72
Definition 117 Markov Operator on Densities . . . 72
Lemma 118 Markov Operators on Measures Induce Those on Densities . . . 72
Definition 119 Transition Operators . . . 73
Definition 120 L_∞ Transition Operators . . . 73
Lemma 121 Kernels and Operators . . . 74
Definition 122 Functional . . . 74
Definition 123 Conjugate or Adjoint Space . . . 74
Proposition 124 Conjugate Spaces are Vector Spaces . . . 74
Proposition 125 Inner Product is Bilinear . . . 74
Example 126 Vectors in $\mathbb{R}^n$ . . . 74
Example 127 Row and Column Vectors . . . 74
Example 128 L_p spaces . . . 75
Example 129 Measures and Functions . . . 75
Definition 130 Adjoint Operator . . . 75
Proposition 131 Adjoint of a Linear Operator . . . 75
Lemma 132 Markov Operators on Densities and L_∞ Transition Operators . . . 75
Theorem 133 Transition operator semi-groups and Markov processes . . . 76
Lemma 134 Markov Operators are Contractions . . . 76
Lemma 135 Markov Operators Bring Distributions Closer . . . 76
Theorem 136 Invariant Measures Are Fixed Points . . . 76
Exercise 10.1 Kernels and Operators . . . 76
Exercise 10.2 L_1 and L_∞ . . . 77
Exercise 10.3 Operators and Expectations . . . 77
Exercise 10.4 Bayesian Updating as a Markov Process . . . 77
Exercise 10.5 More on Bayesian Updating . . . 77

Chapter 11 Examples of Markov Processes **78**
Definition 137 Processes with Stationary and Independent Increments . . . 82
Definition 138 Lévy Processes . . . 82
Example 139 Wiener Process is Lévy . . . 82
Example 140 Poisson Counting Process . . . 82
Theorem 141 Processes with Stationary Independent Increments are Markovian . . . 83
Theorem 142 Time-Evolution Operators of Processes with Stationary, Independent Increments . . . 83
Definition 143 Infinitely-Divisible Distributions and Random Variables . . . 84
Proposition 144 Limiting Distributions Are Infinitely Divisible . 84
Theorem 145 Infinitely Divisible Distributions and Stationary Independent Increments . . . 84
Corollary 146 Infinitely Divisible Distributions and Lévy Processes 84
Definition 147 Self-similarity . . . 85
Definition 148 Stable Distributions . . . 86
Theorem 149 Scaling in Stable Lévy Processes . . . 86
Exercise 11.1 Wiener Process with Constant Drift . . . 86
Exercise 11.2 Perron-Frobenius Operators . . . 86
Exercise 11.3 Continuity of the Wiener Process . . . 86
Exercise 11.4 Independent Increments with Respect to a Filtration 86
Exercise 11.5 Poisson Counting Process . . . 86
Exercise 11.6 Poisson Distribution is Infinitely Divisible . . . 87
Exercise 11.7 Self-Similarity in Lévy Processes . . . 87
Exercise 11.8 Gaussian Stability . . . 87
Exercise 11.9 Poissonian Stability? . . . 87
Exercise 11.10 Lamperti Transformation . . . 87

Chapter 12 Generators of Markov Processes **88**
Definition 150 Infinitessimal Generator . . . 89
Definition 151 Limit in the L-norm sense . . . 89
Lemma 152 Generators are Linear . . . 89
Lemma 153 Invariant Distributions of a Semi-group Belong to the Null Space of Its Generator . . . 89
Lemma 154 Invariant Distributions and the Generator of the Time-Evolution Semigroup . . . 89
Lemma 155 Operators in a Semi-group Commute with Its Generator 90
Definition 156 Time Derivative in Function Space . . . 90
Lemma 157 Generators and Derivatives at Zero . . . 90
Theorem 158 The Derivative of a Function Evolved by a Semi-Group . . . 90
Corollary 159 Initial Value Problems in Function Space . . . 91
Corollary 160 Derivative of Conditional Expectations of a Markov Process . . . 91

Definition 161 Resolvents 92
Definition 162 Yosida Approximation of Operators 92
Theorem 163 Hille-Yosida Theorem 93
Corollary 164 Stochastic Approximation of Initial Value Problems 93
Exercise 12.1 Generators are Linear 93
Exercise 12.2 Semi-Groups Commute with Their Generators . . . 94
Exercise 12.3 Generator of the Poisson Counting Process 94

Chapter 13 The Strong Markov Property and Martingale Problems 95
Definition 165 Strongly Markovian at a Random Time 96
Definition 166 Strong Markov Property 96
Example 167 A Markov Process Which Is Not Strongly Markovian 96
Definition 168 Martingale Problem 97
Proposition 169 Cadlag Nature of Functions in Martingale Problems 97
Lemma 170 Alternate Formulation of Martingale Problem 97
Theorem 171 Markov Processes Solve Martingale Problems . . . 97
Theorem 172 Solutions to the Martingale Problem are Strongly Markovian 97
Exercise 13.1 Strongly Markov at Discrete Times 98
Exercise 13.2 Markovian Solutions of the Martingale Problem . . 98
Exercise 13.3 Martingale Solutions are Strongly Markovian . . . 98

Chapter 14 Feller Processes 99
Definition 173 Initial Distribution, Initial State 99
Lemma 174 Kernel from Initial States to Paths 100
Definition 175 Markov Family 100
Definition 176 Mixtures of Path Distributions (Mixed States) . . 100
Definition 177 Feller Process 101
Definition 178 Contraction Operator 101
Definition 179 Strongly Continuous Semigroup 101
Definition 180 Positive Operator 101
Definition 181 Conservative Operator 102
Lemma 182 Continuous semi-groups produce continuous paths in function space 102
Definition 183 Continuous Functions Vanishing at Infinity 102
Definition 184 Feller Semigroup 102
Lemma 185 The First Pair of Feller Properties 103
Lemma 186 The Second Pair of Feller Properties 103
Theorem 187 Feller Processes and Feller Semigroups 103
Theorem 188 Generator of a Feller Semigroup 103
Theorem 189 Feller Semigroups are Strongly Continuous 103
Proposition 190 Cadlag Modifications Implied by a Kind of Modulus of Continuity 104
Lemma 191 Markov Processes Have Cadlag Versions When They Don't Move Too Fast (in Probability) 104

Lemma 192 Markov Processes Have Cadlag Versions If They Don't Move Too Fast (in Expectation) 104
Theorem 193 Feller Implies Cadlag 105
Theorem 194 Feller Processes are Strongly Markovian 105
Theorem 195 Dynkin's Formula 106
Exercise 14.1 Yet Another Interpretation of the Resolvents 106
Exercise 14.2 The First Pair of Feller Properties 106
Exercise 14.3 The Second Pair of Feller Properties 106
Exercise 14.4 Dynkin's Formula 106
Exercise 14.5 Lévy and Feller Processes 106

Chapter 15 Convergence of Feller Processes 107
Definition 196 Convergence in Finite-Dimensional Distribution . 107
Lemma 197 Finite and Infinite Dimensional Distributional Convergence 108
Definition 198 The Space **D** 108
Definition 199 Modified Modulus of Continuity 108
Proposition 200 Weak Convergence in $\mathbf{D}(\mathbb{R}^+, \Xi)$ 108
Proposition 201 Sufficient Condition for Weak Convergence . . . 109
Definition 202 Closed and Closable Generators, Closures 109
Definition 203 Core of an Operator 109
Lemma 204 Feller Generators Are Closed 109
Theorem 205 Convergence of Feller Processes 110
Corollary 206 Convergence of Discret-Time Markov Processes on Feller Processes 111
Definition 207 Pure Jump Markov Process 111
Proposition 208 Pure-Jump Markov Processea and ODEs 112
Exercise 15.1 Poisson Counting Process 113
Exercise 15.2 Exponential Holding Times in Pure-Jump Processes 113
Exercise 15.3 Solutions of ODEs are Feller Processes 113

Chapter 16 Convergence of Random Walks 114
Definition 209 Continuous-Time Random Walk (Cadlag) 117
Lemma 210 Increments of Random Walks 117
Lemma 211 Continuous-Time Random Walks are Pseudo-Feller . 117
Lemma 212 Evolution Operators of Random Walks 118
Theorem 213 Functional Central Limit Theorem (I) 118
Lemma 214 Convergence of Random Walks in Finite-Dimensional Distribution 119
Theorem 215 Functional Central Limit Theorem (II) 119
Corollary 216 The Invariance Principle 120
Corollary 217 Donsker's Theorem 121
Exercise 16.1 Example 167 Revisited 121
Exercise 16.2 Generator of the d-dimensional Wiener Process . . 121
Exercise 16.3 Continuous-time random walks are Markovian . . . 121
Exercise 16.4 Donsker's Theorem 121

Exercise 16.5 Diffusion equation 122
Exercise 16.6 Functional CLT for Dependent Variables 122

Chapter 17 Diffusions and the Wiener Process 124
Definition 218 Diffusion 124
Lemma 219 The Wiener Process Is a Martingale 126
Definition 220 Wiener Processes with Respect to Filtrations . . . 127
Definition 221 Gaussian Process 127
Lemma 222 Wiener Process Is Gaussian 127
Theorem 223 Almost All Continuous Curves Are Non-Differentiable 128

Chapter 18 Stochastic Integrals with the Wiener Process 130
Theorem 224 Martingale Characterization of the Wiener Process 131

Chapter 19 Stochastic Differential Equations 133
Definition 225 Progressive Process 133
Definition 226 Non-anticipating filtrations, processes 134
Definition 227 Elementary process 134
Definition 228 Mean square integrable 134
Definition 229 $\mathcal{S}_2$ norm 134
Proposition 230 $\|\cdot\|_{\mathcal{S}_2}$ is a norm 134
Definition 231 Itô integral of an elementary process 135
Lemma 232 Approximation of Bounded, Continuous Processes by Elementary Processes 135
Lemma 233 Approximation by of Bounded Processes by Bounded, Continuous Processes 136
Lemma 234 Approximation of Square-Integrable Processes by Bounded Processes 136
Lemma 235 Approximation of Square-Integrable Processes by Elementary Processes 136
Lemma 236 Itô Isometry for Elementary Processes 136
Theorem 237 Itô Integrals of Approximating Elementary Processes Converge 137
Definition 238 Itô integral 138
Corollary 239 The Itô isometry 138
Definition 240 Itô Process 141
Lemma 241 Itô processes are non-anticipating 141
Theorem 242 Itô's Formula in One Dimension 141
Example 243 Section 19.2.2 summarized 145
Definition 244 Multidimensional Itô Process 145
Theorem 245 Itô's Formula in Multiple Dimensions 146
Theorem 246 Representation of Martingales as Stochastic Integrals (Martingale Representation Theorem) 147
Theorem 247 Martingale Characterization of the Wiener Process 147
Definition 248 Stochastic Differential Equation, Solutions 147
Lemma 249 A Quadratic Inequality 148

Definition 250 Maximum Process 148
Definition 251 The Space $\mathcal{QM}(T)$ 149
Lemma 252 Completeness of $\mathcal{QM}(T)$ 149
Proposition 253 Doob's Martingale Inequalities 149
Lemma 254 A Maximal Inequality for Itô Processes 149
Definition 255 Picard operator 150
Lemma 256 Solutions are fixed points of the Picard operator . . 150
Lemma 257 A maximal inequality for Picard iterates 150
Lemma 258 Gronwall's Inequality 151
Theorem 259 Existence and Uniquness of Solutions to SDEs in One Dimension 151
Theorem 260 Existence and Uniqueness for Multidimensional SDEs 153
Exercise 19.1 Basic Properties of the Itô Integral 155
Exercise 19.2 Martingale Properties of the Itô Integral 156
Exercise 19.3 Continuity of the Itô Integral 156
Exercise 19.4 "The square of dW" 156
Exercise 19.5 Itô integrals of elementary processes do not depend on the break-points 156
Exercise 19.6 Itô integrals are Gaussian processes 156
Exercise 19.7 A Solvable SDE 156
Exercise 19.8 Building Martingales from SDEs 157
Exercise 19.9 Brownian Motion and the Ornstein-Uhlenbeck Process 157
Exercise 19.10 Again with the martingale characterization of the Wiener process 157

Chapter 20 More on Stochastic Differential Equations 158
Theorem 261 Solutions of SDEs are Non-Anticipating and Continuous 158
Theorem 262 Solutions of SDEs are Strongly Markov 159
Theorem 263 SDEs and Feller Diffusions 159
Corollary 264 Convergence of Initial Conditions and of Processes 160
Example 265 Wiener process, heat equation 162
Example 266 Ornstein-Uhlenbeck process 163
Exercise 20.1 Langevin equation with a conservative force 163

Chapter 21 Large Deviations for Small-Noise Stochastic Differential Equations 164
Theorem 267 Small-Noise SDEs Converge in Probability on No-Noise ODEs 166
Lemma 268 A Maximal Inequality for the Wiener Process 167
Lemma 269 A Tail Bound for Maxwell-Boltzmann Distributions . 167
Theorem 270 Upper Bound on the Rate of Convergence of Small-Noise SDEs 168

Chapter 22 Spectral Analysis and L_2 Ergodicity **170**
Proposition 271 Linearity of White Noise Integrals 172
Proposition 272 White Noise Has Mean Zero 172
Proposition 273 White Noise and Itô Integrals 173
Proposition 274 White Noise is Uncorrelated 173
Proposition 275 White Noise is Gaussian and Stationary 173
Definition 276 Autocovariance Function 174
Lemma 277 Autocovariance and Time Reversal 174
Definition 278 Second-Order Process 174
Definition 279 Spectral Representation, Power Spectrum 174
Lemma 280 Regularity of the Spectral Process 175
Definition 281 Jump of the Spectral Process 175
Proposition 282 Spectral Representations of Weakly Stationary Processes 176
Definition 283 Orthogonal Increments 176
Lemma 284 Orthogonal Spectral Increments and Weak Stationarity 176
Definition 285 Spectral Function, Spectral Density 177
Theorem 286 Weakly Stationary Processes Have Spectral Functions 177
Theorem 287 Existence of Weakly Stationary Processes with Given Spectral Functions 179
Definition 288 Jump of the Spectral Function 179
Lemma 289 Spectral Function Has Non-Negative Jumps 179
Theorem 290 Wiener-Khinchin Theorem 179
Definition 291 Time Averages 181
Definition 292 Integral Time Scale 182
Theorem 293 Mean-Square Ergodic Theorem (Finite Autocovariance Time) 182
Corollary 294 Convergence Rate in the Mean-Square Ergodic Theorem 183
Lemma 295 Mean-Square Jumps in the Spectral Process 183
Lemma 296 The Jump in the Spectral Function 183
Lemma 297 Existence of L_2 Limits for Time Averages 183
Lemma 298 The Mean-Square Ergodic Theorem 184
Exercise 22.1 Mean-Square Ergodicity in Discrete Time 184
Exercise 22.2 Mean-Square Ergodicity with Non-Zero Mean . . . 185
Exercise 22.3 Functions of Weakly Stationary Processes 185
Exercise 22.4 Ergodicity of the Ornstein-Uhlenbeck Process? . . . 185

Chapter 23 Ergodic Properties and Ergodic Limits **187**
Definition 299 Dynamical System 189
Lemma 300 Dynamical Systems are Markov Processes 189
Definition 301 Observable 189
Definition 302 Invariant Functions, Sets and Measures 189
Lemma 303 Invariant Sets are a σ-Algebra 189
Lemma 304 Invariant Sets and Observables 190
Definition 305 Infinitely Often, i.o. 190

Lemma 306 "Infinitely often" implies invariance 190
Definition 307 Invariance Almost Everywhere 190
Lemma 308 Almost-invariant sets form a σ-algebra 190
Lemma 309 Invariance for simple functions 190
Definition 310 Time Averages 191
Lemma 311 Time averages are observables 191
Definition 312 Ergodic Property 191
Definition 313 Ergodic Limit 191
Lemma 314 Linearity of Ergodic Limits 192
Lemma 315 Non-negativity of ergodic limits 192
Lemma 316 Constant functions have the ergodic property 192
Lemma 317 Ergodic limits are invariantifying 192
Lemma 318 Ergodic limits are invariant functions 193
Lemma 319 Ergodic limits and invariant indicator functions . . . 193
Lemma 320 Ergodic properties of sets and observables 193
Lemma 321 Expectations of ergodic limits 193
Lemma 322 Convergence of Ergodic Limits 194
Lemma 323 Cesàro Mean of Expectations 194
Corollary 324 Replacing Boundedness with Uniform Integrability 194
Definition 325 Asymptotically Mean Stationary 195
Lemma 326 Stationary Implies Asymptotically Mean Stationary 195
Proposition 327 Vitali-Hahn Theorem 195
Theorem 328 Stationary Means are Invariant Measures 195
Lemma 329 Expectations of Almost-Invariant Functions 196
Lemma 330 Limit of Cesàro Means of Expectations 196
Lemma 331 Expectation of the Ergodic Limit is the AMS Expectation 196
Corollary 332 Replacing Boundedness with Uniform Integrability 197
Theorem 333 Ergodic Limits of Bounded Observables are Conditional Expectations 197
Corollary 334 Ergodic Limits of Integrable Observables are Conditional Expectations 197
Exercise 23.1 "Infinitely often" implies invariant 197
Exercise 23.2 Invariant simple functions 197
Exercise 23.3 Ergodic limits of integrable observables 197

Chapter 24 The Almost-Sure Ergodic Theorem 198
Definition 335 Lower and Upper Limiting Time Averages 199
Lemma 336 Lower and Upper Limiting Time Averages are Invariant 199
Lemma 337 "Limits Coincide" is an Invariant Event 199
Lemma 338 Ergodic properties under AMS measures 199
Theorem 339 Birkhoff's Almost-Sure Ergodic Theorem 199
Corollary 340 Birkhoff's Ergodic Theorem for Integrable Observables 203

Chapter 25 Ergodicity and Metric Transitivity 204
Definition 341 Ergodic Systems, Processes, Measures and Transformations . . . 204
Definition 342 Metric Transitivity . . . 205
Lemma 343 Metric transitivity implies ergodicity . . . 205
Lemma 344 Stationary Ergodic Systems are Metrically Transitive 205
Example 345 IID Sequences, Strong Law of Large Numbers . . . 206
Example 346 Markov Chains . . . 206
Example 347 Deterministic Ergodicity: The Logistic Map . . . 206
Example 348 Invertible Ergodicity: Rotations . . . 207
Example 349 Ergodicity when the Distribution Does Not Converge 207
Theorem 350 Ergodicity and the Triviality of Invariant Functions 207
Theorem 351 The Ergodic Theorem for Ergodic Processes . . . 208
Lemma 352 Ergodicity Implies Approach to Independence . . . 208
Theorem 353 Approach to Independence Implies Ergodicity . . . 208
Exercise 25.1 Ergodicity implies an approach to independence . . 209
Exercise 25.2 Approach to independence implies ergodicity . . . 209
Exercise 25.3 Invariant events and tail events . . . 209
Exercise 25.4 Ergodicity of ergodic Markov chains . . . 209

Chapter 26 Decomposition of Stationary Processes into Ergodic Components 210
Proposition 354 . . . 211
Proposition 355 . . . 211
Proposition 356 . . . 211
Definition 357 . . . 212
Proposition 358 . . . 212
Proposition 359 . . . 212
Proposition 360 . . . 213
Proposition 361 . . . 213
Proposition 362 . . . 213
Definition 363 . . . 213
Proposition 364 . . . 213
Definition 365 . . . 214
Proposition 366 . . . 214
Proposition 367 . . . 214
Proposition 368 . . . 215
Lemma 369 . . . 215
Lemma 370 . . . 215
Lemma 371 . . . 215
Theorem 372 . . . 216
Definition 373 . . . 216
Lemma 374 . . . 217
Theorem 375 . . . 217
Theorem 376 . . . 218
Corollary 377 . . . 218

Corollary 378 . . . 218
Exericse 26.1 . . . 219

Chapter 27 Mixing 220
Definition 379 . . . 221
Lemma 380 . . . 221
Theorem 381 . . . 221
Definition 382 . . . 221
Lemma 383 . . . 221
Theorem 384 . . . 222
Example 385 . . . 223
Example 386 . . . 223
Example 387 . . . 223
Example 388 . . . 223
Lemma 389 . . . 224
Theorem 390 . . . 224
Theorem 391 . . . 224
Definition 392 . . . 225
Lemma 393 . . . 225
Definition 394 . . . 226
Definition 395 . . . 226
Theorem 396 . . . 226

Chapter 28 Shannon Entropy and Kullback-Leibler Divergence 228
Definition 397 . . . 229
Lemma 398 . . . 230
Definition 399 . . . 231
Definition 400 . . . 231
Lemma 401 . . . 231
Lemma 402 . . . 231
Definition 403 . . . 231
Lemma 404 . . . 232
Lemma 405 . . . 232
Definition 406 . . . 232
Lemma 407 . . . 233
Lemma 408 . . . 233
Corollary 409 . . . 233
Definition 410 . . . 233
Corollary 411 . . . 234
Definition 412 . . . 234
Proposition 413 . . . 234
Proposition 414 . . . 234
Definition 415 . . . 235
Theorem 416 . . . 235

Chapter 29 Entropy Rates and Asymptotic Equipartition 236
Definition 417 . . . 236
Definition 418 . . . 236
Lemma 419 . . . 237
Theorem 420 . . . 237
Theorem 421 . . . 237
Lemma 422 . . . 238
Example 423 . . . 238
Example 424 . . . 238
Definition 425 . . . 238
Definition 426 . . . 239
Lemma 427 . . . 240
Lemma 428 . . . 240
Lemma 429 . . . 240
Definition 430 . . . 241
Lemma 431 . . . 241
Lemma 432 . . . 241
Lemma 433 . . . 241
Lemma 434 . . . 241
Theorem 435 . . . 242
Theorem 436 . . . 243
Theorem 437 . . . 243
Exercise 29.1 . . . 244
Exercise 29.2 . . . 244

Chapter 30 General Theory of Large Deviations 246
Definition 438 . . . 246
Definition 439 . . . 246
Lemma 440 . . . 247
Lemma 441 . . . 247
Definition 442 . . . 247
Lemma 443 . . . 247
Definition 444 . . . 248
Lemma 445 . . . 248
Lemma 446 . . . 248
Definition 447 . . . 248
Lemma 448 . . . 248
Lemma 449 . . . 249
Theorem 450 . . . 249
Theorem 451 . . . 250
Definition 452 . . . 251
Definition 453 . . . 251
Definition 454 . . . 252
Definition 455 Empirical Process Distribution . . . 252
Corollary 456 . . . 252
Corollary 457 . . . 253

Definition 458 . . . 253
Theorem 459 . . . 253
Theorem 460 . . . 254
Theorem 461 . . . 254
Definition 462 Exponentially Equivalent Random Variables . . . 254
Lemma 463 . . . 254

Chapter 31 Large Deviations for IID Sequences: The Return of Relative Entropy 256
Definition 464 . . . 257
Definition 465 . . . 257
Lemma 466 . . . 258
Definition 467 . . . 258
Definition 468 . . . 258
Lemma 469 . . . 259
Theorem 470 . . . 260
Proposition 471 . . . 263
Proposition 472 . . . 263
Lemma 473 . . . 263
Lemma 474 . . . 263
Theorem 475 . . . 264
Corollary 476 . . . 265
Theorem 477 . . . 265

Chapter 32 Large Deviations for Markov Sequences 266
Theorem 478 . . . 267
Corollary 479 . . . 269
Corollary 480 . . . 269
Corollary 481 . . . 269
Corollary 482 . . . 270
Theorem 483 . . . 270
Theorem 484 . . . 270

Chapter 33 Large Deviations for Weakly Dependent Sequences via the Gärtner-Ellis Theorem 271
Definition 485 . . . 272
Lemma 486 . . . 272
Lemma 487 Upper Bound in the Gärtner-Ellis Theorem for Compact Sets . . . 272
Lemma 488 Upper Bound in the Gärtner-Ellis Theorem for Closed Sets . . . 272
Definition 489 . . . 273
Lemma 490 . . . 273
Lemma 491 . . . 273
Definition 492 Exposed Point . . . 273
Definition 493 . . . 273

Lemma 494 Lower Bound in the Gärtner-Ellis Theorem for Open Sets 274
Theorem 495 Abstract Gärtner-Ellis Theorem 274
Definition 496 Relative Interior 274
Definition 497 Essentially Smooth 274
Proposition 498 274
Theorem 499 Euclidean Gärtner-Ellis Theorem 274
Exercise 33.1 275
Exercise 33.2 275
Exercise 33.3 275
Exercise 33.4 275
Exercise 33.5 275

Chapter 34 Large Deviations for Stochastic Differential Equations 276
Definition 500 Cameron-Martin Spaces 277
Lemma 501 Cameron-Martin Spaces are Hilbert 277
Definition 502 Effective Wiener Action 278
Proposition 503 Girsanov Formula for Deterministic Drift 278
Lemma 504 Exponential Bound on the Probability of Tubes Around Given Trajectories 278
Lemma 505 Trajectories are Rarely Far from Action Minima . . 279
Proposition 506 Compact Level Sets of the Wiener Action 280
Theorem 507 Schilder's Theorem on Large Deviations of the Wiener Process on the Unit Interval 280
Corollary 508 Extension of Schilder's Theorem to $[0, T]$ 281
Corollary 509 Schilder's Theorem on $\mathbb{R}^+$ 281
Definition 510 SDE with Small State-Independent Noise 282
Definition 511 Effective Action under State-Independent Noise . 282
Lemma 512 Continuous Mapping from Wiener Process to SDE Solutions 282
Lemma 513 Mapping Cameron-Martin Spaces Into Themselves . 283
Theorem 514 Freidlin-Wentzell Theorem with State-Independent Noise 283
Definition 515 SDE with Small State-Dependent Noise 283
Definition 516 Effective Action under State-Dependent Noise . . 284
Theorem 517 Freidlin-Wentzell Theorem for State-Dependent Noise 284
Exercise 34.1 Cameron-Martin Spaces are Hilbert Spaces 284
Exercise 34.2 Lower Bound in Schilder's Theorem 284
Exercise 34.3 Mapping Cameron-Martin Spaces Into Themselves 284