

Contents

Preface	vii
Introduction	xvii
1 Information Sources	1
1.1 Probability Spaces and Random Variables	1
1.2 Random Processes and Dynamical Systems	5
1.3 Distributions	7
1.4 Standard Alphabets	12
1.5 Expectation	13
1.6 Asymptotic Mean Stationarity	16
1.7 Ergodic Properties	17
2 Pair Processes: Channels, Codes, and Couplings	21
2.1 Pair Processes	21
2.2 Channels	22
2.3 Stationarity Properties of Channels	25
2.4 Extremes: Noiseless and Completely Random Channels	29
2.5 Deterministic Channels and Sequence Coders	30
2.6 Stationary and Sliding-Block Codes	31
2.7 Block Codes	37
2.8 Random Punctuation Sequences	38
2.9 Memoryless Channels	42
2.10 Finite-Memory Channels	42
2.11 Output Mixing Channels	43
2.12 Block Independent Channels	45
2.13 Conditionally Block Independent Channels	46
2.14 Stationarizing Block Independent Channels	46
2.15 Primitive Channels	48
2.16 Additive Noise Channels	49
2.17 Markov Channels	49

2.18	Finite-State Channels and Codes	50
2.19	Cascade Channels	51
2.20	Communication Systems	52
2.21	Couplings	52
2.22	Block to Sliding-Block: The Rohlin-Kakutani Theorem	53
3	Entropy	61
3.1	Entropy and Entropy Rate	61
3.2	Divergence Inequality and Relative Entropy	65
3.3	Basic Properties of Entropy	69
3.4	Entropy Rate	78
3.5	Relative Entropy Rate	81
3.6	Conditional Entropy and Mutual Information	82
3.7	Entropy Rate Revisited	90
3.8	Markov Approximations	91
3.9	Relative Entropy Densities	93
4	The Entropy Ergodic Theorem	97
4.1	History	97
4.2	Stationary Ergodic Sources	100
4.3	Stationary Nonergodic Sources	106
4.4	AMS Sources	110
4.5	The Asymptotic Equipartition Property	114
5	Distortion and Approximation	117
5.1	Distortion Measures	117
5.2	Fidelity Criteria	120
5.3	Average Limiting Distortion	121
5.4	Communications Systems Performance	123
5.5	Optimal Performance	124
5.6	Code Approximation	124
5.7	Approximating Random Vectors and Processes	129
5.8	The Monge/Kantorovich/Vasershtein Distance	132
5.9	Variation and Distribution Distance	132
5.10	Coupling Discrete Spaces with the Hamming Distance	134
5.11	Process Distance and Approximation	135
5.12	Source Approximation and Codes	141
5.13	\bar{d} Continuous Channels	142
6	Distortion and Entropy	147
6.1	The Fano Inequality	147
6.2	Code Approximation and Entropy Rate	150
6.3	Pinsker's and Marton's Inequalities	152
6.4	Entropy and Isomorphism	156
6.5	Almost Lossless Source Coding	160
6.6	Asymptotically Optimal Almost Lossless Codes	168

6.7	Modeling and Simulation	169
7	Relative Entropy	173
7.1	Divergence	173
7.2	Conditional Relative Entropy	189
7.3	Limiting Entropy Densities	202
7.4	Information for General Alphabets	204
7.5	Convergence Results	216
8	Information Rates	219
8.1	Information Rates for Finite Alphabets	219
8.2	Information Rates for General Alphabets	221
8.3	A Mean Ergodic Theorem for Densities	225
8.4	Information Rates of Stationary Processes	227
8.5	The Data Processing Theorem	234
8.6	Memoryless Channels and Sources	235
9	Distortion and Information	237
9.1	The Shannon Distortion-Rate Function	237
9.2	Basic Properties	239
9.3	Process Definitions of the Distortion-Rate Function	242
9.4	The Distortion-Rate Function as a Lower Bound	250
9.5	Evaluating the Rate-Distortion Function	252
10	Relative Entropy Rates	265
10.1	Relative Entropy Densities and Rates	265
10.2	Markov Dominating Measures	268
10.3	Stationary Processes	272
10.4	Mean Ergodic Theorems	275
11	Ergodic Theorems for Densities	281
11.1	Stationary Ergodic Sources	281
11.2	Stationary Nonergodic Sources	286
11.3	AMS Sources	290
11.4	Ergodic Theorems for Information Densities	293
12	Source Coding Theorems	295
12.1	Source Coding and Channel Coding	295
12.2	Block Source Codes for AMS Sources	296
12.3	Block Source Code Mismatch	307
12.4	Block Coding Stationary Sources	310
12.5	Block Coding AMS Ergodic Sources	312
12.6	Subadditive Fidelity Criteria	319
12.7	Asynchronous Block Codes	321
12.8	Sliding-Block Source Codes	323
12.9	A Geometric Interpretation	333

13	Properties of Good Source Codes	335
13.1	Optimal and Asymptotically Optimal Codes	335
13.2	Block Codes	337
13.3	Sliding-Block Codes	343
14	Coding for Noisy Channels	359
14.1	Noisy Channels	359
14.2	Feinstein's Lemma	361
14.3	Feinstein's Theorem	364
14.4	Channel Capacity	367
14.5	Robust Block Codes	372
14.6	Block Coding Theorems for Noisy Channels	375
14.7	Joint Source and Channel Block Codes	377
14.8	Synchronizing Block Channel Codes	380
14.9	Sliding-block Source and Channel Coding	384
	References	395
	Index	405