

Copyright Cambridge University Press 2003. On-screen viewing permitted. Printing not permitted. <http://www.cambridge.org/0521642981>
You can buy this book for 30 pounds or \$50. See <http://www.inference.phy.cam.ac.uk/mackay/itila/> for links.

Information Theory, Inference, and Learning Algorithms

David J.C. MacKay

Information Theory, Inference, and Learning Algorithms

David J.C. MacKay
mackay@mrao.cam.ac.uk

©1995, 1996, 1997, 1998, 1999, 2000, 2001, 2002, 2003, 2004, 2005

©Cambridge University Press 2003

Version 7.2 (fourth printing) March 28, 2005

Please send feedback on this book via
<http://www.inference.phy.cam.ac.uk/mackay/itila/>

Version 6.0 of this book was published by C.U.P. in September 2003. It will remain viewable on-screen on the above website, in postscript, djvu, and pdf formats.

In the second printing (version 6.6) minor typos were corrected, and the book design was slightly altered to modify the placement of section numbers.

In the third printing (version 7.0) minor typos were corrected, and chapter 8 was renamed 'Dependent random variables' (instead of 'Correlated').

In the fourth printing (version 7.2) minor typos were corrected.

(C.U.P. replace this page with their own page ii.)

Contents

	Preface	v
1	Introduction to Information Theory	3
2	Probability, Entropy, and Inference	22
3	More about Inference	48
I	Data Compression	65
4	The Source Coding Theorem	67
5	Symbol Codes	91
6	Stream Codes	110
7	Codes for Integers	132
II	Noisy-Channel Coding	137
8	Dependent Random Variables	138
9	Communication over a Noisy Channel	146
10	The Noisy-Channel Coding Theorem	162
11	Error-Correcting Codes and Real Channels	177
III	Further Topics in Information Theory	191
12	Hash Codes: Codes for Efficient Information Retrieval	193
13	Binary Codes	206
14	Very Good Linear Codes Exist	229
15	Further Exercises on Information Theory	233
16	Message Passing	241
17	Communication over Constrained Noiseless Channels	248
18	Crosswords and Codebreaking	260
19	Why have Sex? Information Acquisition and Evolution	269
IV	Probabilities and Inference	281
20	An Example Inference Task: Clustering	284
21	Exact Inference by Complete Enumeration	293
22	Maximum Likelihood and Clustering	300
23	Useful Probability Distributions	311
24	Exact Marginalization	319
25	Exact Marginalization in Trellises	324
26	Exact Marginalization in Graphs	334
27	Laplace's Method	341

28	Model Comparison and Occam's Razor	343
29	Monte Carlo Methods	357
30	Efficient Monte Carlo Methods	387
31	Ising Models	400
32	Exact Monte Carlo Sampling	413
33	Variational Methods	422
34	Independent Component Analysis and Latent Variable Modelling	437
35	Random Inference Topics	445
36	Decision Theory	451
37	Bayesian Inference and Sampling Theory	457
V	Neural networks	467
38	Introduction to Neural Networks	468
39	The Single Neuron as a Classifier	471
40	Capacity of a Single Neuron	483
41	Learning as Inference	492
42	Hopfield Networks	505
43	Boltzmann Machines	522
44	Supervised Learning in Multilayer Networks	527
45	Gaussian Processes	535
46	Deconvolution	549
VI	Sparse Graph Codes	555
47	Low-Density Parity-Check Codes	557
48	Convolutional Codes and Turbo Codes	574
49	Repeat-Accumulate Codes	582
50	Digital Fountain Codes	589
VII	Appendices	597
A	Notation	598
B	Some Physics	601
C	Some Mathematics	605
	Bibliography	613
	Index	620