
Contents

<i>Expanded Contents</i>	<i>page</i> viii
<i>Glossary of Symbols</i>	xvi
<i>Preface</i>	xxi
1 Introduction	1
2 Priors on Function Spaces	10
3 Priors on Spaces of Probability Measures	25
4 Dirichlet Processes	59
5 Dirichlet Process Mixtures	102
6 Consistency: General Theory	123
7 Consistency: Examples	165
8 Contraction Rates: General Theory	192
9 Contraction Rates: Examples	233
10 Adaptation and Model Selection	270
11 Gaussian Process Priors	310
12 Infinite-Dimensional Bernstein–von Mises Theorem	361
13 Survival Analysis	391
14 Discrete Random Structures	437
<i>Appendices</i>	507
<i>References</i>	623
<i>Author Index</i>	638
<i>Subject Index</i>	642

Expanded Contents

<i>Glossary of Symbols</i>	xvi
<i>Preface</i>	xxi
1 Introduction	1
1.1 Motivation	1
1.1.1 Classical versus Bayesian Nonparametrics	1
1.1.2 Parametric versus Nonparametric Bayes	3
1.2 Challenges of Bayesian Nonparametrics	3
1.2.1 Prior Construction	3
1.2.2 Computation	4
1.2.3 Asymptotic Behavior	5
1.3 Priors, Posteriors and Bayes's Rule	5
1.3.1 Absolute Continuity	7
1.4 Historical Notes	8
2 Priors on Function Spaces	10
2.1 Random Basis Expansion	10
2.2 Stochastic Processes	13
2.2.1 Gaussian Processes	13
2.2.2 Increasing Processes	14
2.3 Probability Densities	15
2.3.1 Exponential Link Function	16
2.3.2 Construction through Binning	17
2.3.3 Mixtures	17
2.3.4 Feller Approximation	18
2.4 Nonparametric Normal Regression	20
2.5 Nonparametric Binary Regression	20
2.6 Nonparametric Poisson Regression	22
2.7 Historical Notes	23
Problems	23
3 Priors on Spaces of Probability Measures	25
3.1 Random Measures	25
3.1.1 Other Topologies	26
3.2 Construction through a Stochastic Process	26

3.3	Countable Sample Spaces	29
3.3.1	Construction through Normalization	29
3.3.2	Construction through Stick Breaking	30
3.3.3	Countable Dirichlet Process	31
3.4	Construction through Structural Definitions	33
3.4.1	Construction through a Distribution on a Dense Subset	33
3.4.2	Construction through a Randomly Selected Discrete Set	33
3.4.3	Construction through Random Rectangular Partitions	34
3.4.4	Construction through Moments	36
3.4.5	Construction through Quantiles	36
3.4.6	Construction by Normalization	36
3.5	Construction through a Tree	37
3.6	Tail-Free Processes	39
3.7	Pólya Tree Processes	48
3.7.1	Relation with the Pólya Urn Scheme	51
3.7.2	Mixtures of Pólya Tree Processes	53
3.7.3	Partially Specified Pólya Tree	55
3.7.4	Evenly Split Pólya Tree	55
3.8	Historical Notes	56
	Problems	56
4	Dirichlet Processes	59
4.1	Definition and Basic Properties	59
4.1.1	Expectations, Variances and Co-Variances	60
4.1.2	Self-Similarity	61
4.1.3	Conjugacy	62
4.1.4	Marginal and Conditional Distributions	64
4.1.5	Number of Distinct Values	65
4.2	Constructions	69
4.2.1	Construction via a Stochastic Process	69
4.2.2	Construction through Distribution Function	70
4.2.3	Construction through a Gamma Process	71
4.2.4	Construction through Pólya Urn Scheme	71
4.2.5	Stick-Breaking Representation	72
4.3	Further Properties	73
4.3.1	Discreteness and Support	73
4.3.2	Convergence	74
4.3.3	Approximations	76
4.3.4	Mutual Singularity of Dirichlet Processes	79
4.3.5	Tails of a Dirichlet Process	80
4.3.6	Distribution of Median	82
4.3.7	Distribution of Mean	82
4.4	Characterizations	84
4.5	Mixtures of Dirichlet Processes	87
4.6	Modifications	90

4.6.1	Invariant Dirichlet Process	90
4.6.2	Constrained Dirichlet Process	93
4.6.3	Penalized Dirichlet Process	94
4.7	Bayesian Bootstrap	95
4.8	Historical Notes	96
	Problems	97
5	Dirichlet Process Mixtures	102
5.1	Dirichlet Process Mixtures	102
5.2	MCMC Methods	105
5.3	Variational Algorithm	109
5.4	Predictive Recursion Deconvolution Algorithm	112
5.5	Examples of Kernels	115
5.6	Historical Notes	121
	Problems	121
6	Consistency: General Theory	123
6.1	Consistency and Its Implications	123
6.2	Doob's Theorem	129
6.3	Inconsistency	131
6.4	Schwartz's Theorem	137
6.5	Tail-Free Priors	144
6.6	Permanence of the Kullback-Leibler Property	144
6.7	General Observations	147
6.7.1	Independent Observations	150
6.7.2	Markov Processes	151
6.8	Alternative Approaches	152
6.8.1	Separation	152
6.8.2	Le Cam's Inequality	153
6.8.3	Predictive Consistency	155
6.8.4	Martingale Approach	157
6.8.5	α -Posterior	159
6.9	Historical Notes	160
	Problems	161
7	Consistency: Examples	165
7.1	Priors with the Kullback-Leibler Property	165
7.1.1	Pólya Trees	165
7.1.2	Kernel Mixtures	166
7.1.3	Exponential Densities	170
7.2	Density Estimation	171
7.2.1	Normal Mixtures	172
7.2.2	Dirichlet Process Mixtures of a General Kernel	174
7.2.3	Pólya Tree Process	175
7.2.4	Exponential Densities	176

7.3	Other Nonparametric Models	177
7.3.1	Nonparametric Binary Regression	177
7.3.2	Nonparametric Regression with Normal Errors	179
7.3.3	Spectral Density Estimation	179
7.4	Semiparametric Models	181
7.4.1	Location Problem	181
7.4.2	Linear Regression with Unknown Error Density	182
7.4.3	Binary Nonparametric Monotone Regression	184
7.5	Historical Notes	185
	Problems	186
8	Contraction Rates: General Theory	192
8.1	Introduction	192
8.2	Independent Identically Distributed Observations	196
8.2.1	Further Refinements	202
8.2.2	Priors Based on Finite Approximating Sets	204
8.3	General Observations	206
8.3.1	Independent Observations	208
8.3.2	Gaussian Regression with Fixed Design	211
8.3.3	Markov Chains	213
8.3.4	White Noise Model	215
8.3.5	Gaussian Time Series	216
8.4	Lower Bounds	218
8.5	Misspecification	219
8.5.1	Convex Models	223
8.5.2	Nonparametric Regression	223
8.6	α -Posterior	227
8.7	Historical Notes	229
	Problems	229
9	Contraction Rates: Examples	233
9.1	Log-Spline Priors	233
9.2	Priors Based on Dirichlet Processes	236
9.3	Bernstein Polynomials	238
9.4	Dirichlet Process Mixtures of Normal Kernel	240
9.4.1	Approximation	242
9.4.2	Prior Concentration	246
9.4.3	Entropy Estimate and Controlling Complexity	249
9.4.4	Proof of Theorem 9.9	250
9.4.5	Wishart Prior	251
9.5	Non-i.i.d. Models	252
9.5.1	Finite Sieves	252
9.5.2	Whittle Estimation of a Spectral Density	254
9.5.3	Nonlinear Autoregression	255
9.5.4	White Noise with Conjugate Priors	257

9.5.5	Nonparametric Regression Using Splines	259
9.5.6	Binary Nonparametric Regression with a Dirichlet Process Prior	261
9.5.7	Interval Censoring Using a Dirichlet Process Prior	262
9.6	Historical Notes	263
	Problems	264
10	Adaptation and Model Selection	270
10.1	Introduction	270
10.2	Independent Identically Distributed Observations	272
10.2.1	Universal Weights	277
10.2.2	Parametric Rate	277
10.2.3	Two Models	278
10.3	Examples	278
10.3.1	Priors Based on Finite Approximating Sets	278
10.3.2	White Noise Model	280
10.3.3	Finite-Dimensional Approximations	282
10.3.4	Log-Spline Models	284
10.4	Finite Random Series	289
10.4.1	Density Estimation	294
10.4.2	Nonparametric Normal Regression	294
10.4.3	Nonparametric Binary Regression	296
10.4.4	Nonparametric Poisson Regression	296
10.4.5	Functional Regression	297
10.4.6	Whittle Estimation of a Spectral Density	298
10.5	Model Selection Consistency	299
10.5.1	Testing a Point Null	300
10.5.2	General Case	300
10.5.3	Testing Parametric versus Nonparametric Models	302
10.6	Historical Notes	303
	Problems	305
11	Gaussian Process Priors	310
11.1	Definition and Examples	310
11.2	Reproducing Kernel Hilbert Space	314
11.3	Posterior Contraction Rates	319
11.3.1	Density Estimation	321
11.3.2	Nonparametric Binary Regression	322
11.3.3	Nonparametric Normal Regression	322
11.3.4	White Noise Model	323
11.4	Specific Gaussian Processes as Priors	323
11.4.1	Brownian Motion and Its Primitives	324
11.4.2	Riemann-Liouville Process	327
11.4.3	Fractional Brownian Motion	328
11.4.4	Stationary Processes	329
11.4.5	Series Priors	334

11.5	Rescaled Gaussian Processes	340
11.5.1	Self-Similar Processes	341
11.5.2	Stationary Gaussian Processes	342
11.6	Adaptation	344
11.7	Computation	348
11.7.1	Kernel Methods and the Posterior Mode	349
11.7.2	Density Estimation	351
11.7.3	Nonparametric Binary Regression	352
11.7.4	Expectation Propagation	353
11.7.5	Laplace Approximation	356
11.8	Historical Notes	357
	Problems	358
12	Infinite-Dimensional Bernstein–von Mises Theorem	361
12.1	Introduction	361
12.2	Dirichlet Process	363
12.2.1	Strong Approximation	365
12.3	Semiparametric Models	368
12.3.1	Functionals	370
12.3.2	Strict Semiparametric Model	375
12.3.3	Cox Proportional Hazard Model	377
12.4	White Noise Model	383
12.4.1	Full Parameter	383
12.4.2	Linear Functionals	385
12.5	Historical Notes	388
	Problems	389
13	Survival Analysis	391
13.1	Introduction	391
13.2	Dirichlet Process Prior	394
13.3	Beta Process Prior	397
13.3.1	Discrete Time	397
13.3.2	Continuous Time	399
13.3.3	Sample Path Generation	401
13.3.4	Mixtures of Beta Processes	403
13.4	Neutral to the Right and Independent Increment Processes	403
13.4.1	Consistency	413
13.4.2	Bernstein–von Mises Theorem	416
13.5	Smooth Hazard Processes	419
13.6	Proportional Hazard Model	423
13.6.1	Posterior Distribution	423
13.6.2	Bernstein–von Mises Theorem	427
13.7	The Bayesian Bootstrap for Censored Data	428
13.7.1	Survival Data without Covariates	428
13.7.2	Cox Proportional Hazard Model	429

13.8	Historical Notes	432
	Problems	433
14	Discrete Random Structures	437
14.1	Exchangeable Partitions	437
14.1.1	The Chinese Restaurant Process	443
14.1.2	The Chinese Restaurant Franchise Process	444
14.2	Species Sampling Processes	445
14.2.1	Posterior Distribution	452
14.2.2	Species Sampling Process Mixtures	456
14.3	Gibbs Processes	457
14.4	Pitman-Yor Process	464
14.5	Poisson-Kingman Processes	472
14.6	Normalized Inverse-Gaussian Process	480
14.7	Normalized Completely Random Measures	482
14.8	Relations between Classes of Discrete Random Probability Measures	488
14.9	Dependent Random Discrete Distributions	490
14.9.1	Kernel Stick-Breaking Process	491
14.9.2	Local Dirichlet Process	492
14.9.3	Probit Stick-Breaking Process	493
14.9.4	Ordering Dependent Stick-Breaking Processes	493
14.9.5	Nested Dirichlet Processes	494
14.10	The Indian Buffet Process	495
14.11	Historical Notes	502
	Problems	504
	<i>Appendices</i>	507
A	Space of Probability Measures	507
B	Space of Probability Densities	516
C	Packing, Covering, Bracketing and Entropy Numbers	528
D	Hypothesis Tests	533
E	Polynomials, Splines and Wavelets	546
F	Elements of Empirical Processes	559
G	Finite-Dimensional Dirichlet Distribution	562
H	Inverse-Gaussian Distribution	571
I	Gaussian Processes	573
J	Completely Random Measures	592

K	Inequalities and Estimates	609
L	Miscellaneous Results	614
M	Elements of Markov Chain Monte Carlo	619
	<i>References</i>	623
	<i>Author Index</i>	638
	<i>Subject Index</i>	642