

Probability tools, tricks, and miracles

©David Pollard

1	What and why	1
1.1	How I chose material for this book	1
1.2	Naming and notation	4
1.3	Brief description of Chapters	5
1.4	Notes	7
2	It's just calculus and convexity	9
2.1	Introduction	9
2.2	Polynomial approximations	10
2.3	Convexity of the log MGF	12
2.4	Conjugate convex functions	14
2.5	Approximation of factorials	16
2.6	Distances between probability measures	20
2.7	Bounding sums by integrals	22
2.8	Problems	22
2.9	Notes	25
3	The moment generating function method	29
3.1	Tail bounds from the moment generating function	30
3.2	Behavior of the tail bound near the origin	31
3.3	Normal	33
3.4	Global behavior of the normal hazard rate	38
3.5	Poisson	44
3.6	Gamma and chi-squared	46
3.7	Binomial	49
3.8	Sampling and the hypergeometric	53
3.9	Problems	55
3.10	Notes	60
4	Poisson-Binomial	66
4.1	The extreme Binomial	66
4.2	The PGF of the Poisson-Binomial	67
4.3	Polynomials with real roots	69
4.4	Sampling and the Poisson-Binomial	72
4.5	Ratios and mode(s)	76
4.6	Mode(s) of the Poisson-Binomial	77
	4.6.1 Zeros and ones	80
4.7	Binomial tails versus Poisson-Binomial tails	82
4.8	The median of the Poisson-Binomial	87
4.9	Problems	88
4.10	Notes	90

5	Orlicz spaces	93
5.1	What is an orlicz space?	93
5.2	Facts about orlicz functions	95
5.3	Orlicz functions that grow exponentially fast	98
5.4	Norms of maxima of finitely many variables	101
5.5	Conjugacy and Young's inequality	103
5.6	Comparison of $\mathcal{L}^p$ and $\mathcal{L}^\Psi$	107
5.7	Problems	109
5.8	Notes	113
6	Multivariate normal	116
6.1	Introduction	116
6.1.1	Subgaussian concentration	117
6.1.2	Comparison inequalities	119
6.2	The path method	120
6.3	Concentration of LIPSCHITZ functionals	122
6.3.1	The PISIER-MAUREY method	123
6.3.2	The smart path method	124
6.3.3	The stochastic calculus method	126
6.4	The gaussian isoperimetric inequality	130
6.5	Tail bound comparisons	131
6.6	An application of the GORDON inequality	135
6.7	A generalized FERNIQUE inequality	138
6.8	Problems	140
6.9	Notes	143
7	Subgaussian distributions	149
7.1	Definition of subgaussian	149
7.2	Characterizations of subgaussianity	151
7.3	The Hoeffding inequality for independent summands	153
7.4	From independence to martingales	155
7.5	A cautionary example	160
7.6	Problems	162
7.7	Notes	163
8	Between subgaussian and subexponential	165
8.1	Introduction	165
8.2	A BENNETT inequality for independent summands	166
8.3	A BERNSTEIN inequality	169
8.4	Quadratics in independent subgaussians	172
8.5	A BENNETT inequality for martingales	178
8.6	Concentration of random polynomials	180
8.7	Problems	188
8.8	Notes	192

9	Measurability woes	195
9.1	The difficulty	195
9.2	Essential supremum and infimum	196
9.3	Outer integrals and measurable cover functions	197
9.4	Separable versions	198
9.5	Problems	201
9.6	Notes	201
10	Chaining	204
10.1	Maximal inequalities	204
10.2	A reminder about ORLICZ norms	206
10.3	What is chaining?	208
10.4	Covering and packing numbers	213
10.5	Chaining with packing numbers	215
10.5.1	Expected values	216
10.5.2	Orlicz norms	217
10.5.3	Conditional expected values	217
10.5.4	Tail probabilities	218
10.6	Oscillation and continuity of sample paths	219
10.7	An example of a classical chaining argument	221
10.8	Problems	223
10.9	Notes	225
11	Chaining via partitions	229
11.1	Overview	229
11.2	Partitions, nested and not nested	232
11.2.1	Case: Nested partitions	232
11.2.2	Case: Non-nested partitions	232
11.2.3	Chaining bounds	233
11.3	Fernique and majorizing measures	234
11.4	Majorizing measures as smoothers	236
11.5	Partitions with weights	238
11.6	From majorizing measures to weighted frameworks	241
11.7	Talagrand and admissible partitions	244
11.8	From majorizing measures to admissible partitions	247
11.9	An illuminating example	250
11.9.1	Paths bounded but not continuous	251
11.9.2	Failure of the packing argument	251
11.9.3	Majorizing measure	251
11.9.4	Weighted partitions	252
11.9.5	Admissible partitions	252
11.10	A comparison of chaining methods	253
11.11	Problems	254
11.12	Notes	257

12 Majorizing measures for gaussian processes	260
12.1 Introduction	260
12.2 A gaussian growth property	261
12.3 From functional to nested partitions with weights	263
12.4 Uniformly continuous gaussian sample paths	266
12.5 From functional to admissible partitions	268
12.6 Beyond gaussian	273
12.7 Problems	273
12.8 Notes	274
 13 Symmetrization	 276
13.1 Symmetry?	276
13.2 Stochastic processes as random objects	284
13.3 A generic tail bound by symmetrization	287
13.4 Empirical processes	287
13.5 Expected supremum of an empirical process	289
13.6 Symmetrization of independent processes	293
13.7 Maximal tail bounds for an empirical process	294
13.8 Oscillation control for an empirical process	298
13.9 U-processes	301
13.10 Problems	304
13.11 Notes	307
 14 VC Sets	 311
14.1 An introductory example	311
14.2 VC-dimension, shatter-dimension	314
14.3 Proof of the shatter theorem	319
14.4 Subsets of the discrete hypercube as graphs	322
14.5 Haussler's improvement of the packing bound	325
14.6 Problems	329
14.7 Notes	331
 15 Fat shattering	 334
15.1 VC-subgraphs	334
15.2 Fat shattering	338
15.3 Packing for $[0, 1]$ -valued functions	342
15.4 Packing for subsets of an $\mathcal{L}^q$ ball	344
15.5 Chaining bounds	347
15.6 Lattice cubes surrounded by well separated vectors	348
15.7 Problems	351
15.8 Notes	355
 16 Convergence in distribution	 358
16.1 Difficulties with the uniform metric	358
16.2 Convergence in distribution, new style	362
16.3 Approximation versus limit theorem	366
16.4 Convergence of the uniform empirical process	368

16.5	Problems	371
16.6	Notes	372
17	Concentration: Talagrand's inductive method	374
17.1	Introduction	374
17.2	Examples	376
17.3	Convexity and convex hulls	379
17.4	Control by a single point	381
17.5	Proof of the convex hull theorem	384
17.6	Problems	389
17.7	Notes	392
18	Concentration: the entropy method	394
18.1	Introduction	395
18.2	Tensorization	396
18.3	Marginal calculations	398
18.4	Massart et al	403
18.5	Concentration for empirical processes	407
18.6	Problems	407
18.7	Notes	408
19	Bracketing methods	409
19.1	What is bracketing?	410
19.2	Bounded independent summands	414
19.3	Independent summands with truncation	418
19.4	A generic bracketing bound	423
19.5	Phi mixing	428
19.6	Absolute regularity	431
19.7	Strong mixing	433
19.8	Tail probability bounds	434
19.9	Problems	434
19.10	Notes	435