



VOLUME TWO

**ADVANCED
MATHEMATICAL TOOLS
FOR AUTOMATIC
CONTROL ENGINEERS**

Stochastic Techniques

ALEXANDER S. POZNYAK

Contents

Preface	xv
Notations and Symbols	xxi
List of Figures	xxvii
List of Tables	xxix

PART I BASICS OF PROBABILITY 1

Chapter 1 Probability Space 3

1.1	Set operations, algebras and sigma-algebras	3
1.1.1	Set operations, set limits and collections of sets	3
1.1.2	Algebras and sigma-algebras	6
1.2	Measurable and probability spaces	9
1.2.1	Measurable spaces and finite additive measures	9
1.2.2	The Kolmogorov axioms and the probability space	10
1.2.3	Sets of a null measure and completeness	15
1.3	Borel algebra and probability measures	15
1.3.1	The measurable space $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$	15
1.3.2	The measurable space $(\mathbb{R}^N, \mathcal{B}(\mathbb{R}^N))$	24
1.3.3	The measurable space $(\mathbb{R}^T, \mathcal{B}(\mathbb{R}^T))$	25
1.3.4	Wiener measure on $(\mathbb{R}^{[0, \infty]}, \mathcal{B}(\mathbb{R}^{[0, \infty]}))$	26
1.4	Independence and conditional probability	26
1.4.1	Independence	26
1.4.2	Pair-wise independence	29
1.4.3	Definition of conditional probability	29
1.4.4	Bayes's formula	30

Chapter 2 Random Variables 33

2.1	Measurable functions and random variables	33
2.1.1	Measurable functions	33
2.1.2	Borel functions and multidimensional random variables	34
2.1.3	Indicators and discrete variables	36

2.2	Transformation of distributions	37
2.2.1	Functionally connected random variables	37
2.2.2	Transformation of densities	39
2.3	Continuous random variables	42
2.3.1	Continuous variables	43
2.3.2	Theorem on monotone approximation	43
Chapter 3 Mathematical Expectation		47
3.1	Definition of mathematical expectation	47
3.1.1	Whittle axioms	47
3.1.2	Mathematical expectation as the Lebesgue integral	49
3.1.3	Moments, mean, variance, median and α -quantile	51
3.2	Calculation of mathematical expectation	52
3.2.1	Discrete variables	52
3.2.2	Continuous variables	53
3.2.3	Absolutely continuous variables	54
3.3	Covariance, correlation and independence	60
3.3.1	Covariance	60
3.3.2	Correlation	60
3.3.3	Relation with independence	61
Chapter 4 Basic Probabilistic Inequalities		63
4.1	Moment-type inequalities	63
4.1.1	Generalized Chebyshev inequality	63
4.1.2	Hölder inequality	65
4.1.3	Cauchy–Bounyakovski–Schwartz inequality	66
4.1.4	Jensen inequality	66
4.1.5	Lyapunov inequality	67
4.1.6	Kulbac inequality	68
4.1.7	Minkowski inequality	68
4.1.8	r -Moment inequality	69
4.1.9	Exponential inequalities	70
4.2	Probability inequalities for maxima of partial sums	72
4.2.1	Classical Kolmogorov-type inequalities	72
4.2.2	Hájek–Renyi inequality	73
4.2.3	Relation of a maxima of partial sum probabilities with a distribution of the last partial sum	73
4.3	Inequalities between moments of sums and summands	80

Chapter 5	Characteristic Functions	83
5.1	Definitions and examples	83
5.1.1	Some examples of characteristic functions	84
5.1.2	A linear transformed random vector and the sum of independent random vectors	87
5.2	Basic properties of characteristic functions	88
5.2.1	Simple properties	88
5.2.2	Relations with moments	90
5.3	Uniqueness and inversion	94
5.3.1	Uniqueness	94
5.3.2	Inversion formula	95
5.3.3	Parseval's-type relation	98
PART II DISCRETE TIME PROCESSES		101
Chapter 6	Random Sequences	103
6.1	Random process in discrete and continuous time	103
6.2	Infinitely often events	104
6.2.1	Main definition	104
6.2.2	Tail events and the Kolmogorov zero-one law	104
6.2.3	The Borel-Cantelli lemma	107
6.3	Properties of Lebesgue integral with probabilistic measure	110
6.3.1	Lemma on monotone convergence	110
6.3.2	Fatou's lemma	112
6.3.3	The Lebesgue dominated convergence theorem	113
6.3.4	Uniform integrability	114
6.4	Convergence	117
6.4.1	Various modes of convergence	117
6.4.2	Fundamental sequences	119
6.4.3	Distributional convergence	122
6.4.4	Relations between convergence concepts	123
6.4.5	Some converses	125
Chapter 7	Martingales	133
7.1	Conditional expectation relative to a sigma-algebra	133
7.1.1	Main definition	133
7.1.2	Some properties of conditional expectation	134

7.2	Martingales and related concepts	138
7.2.1	Martingales, submartingales and supermartingales	138
7.2.2	Some examples	140
7.2.3	Decompositions of submartingales and quadratic variation	143
7.2.4	Markov and stopping times	146
7.3	Main martingale inequalities	156
7.3.1	Doob's inequality of the Kolmogorov type	157
7.3.2	Doob's moment inequalities	158
7.4	Convergence	161
7.4.1	'Up-down' crossing lemma	161
7.4.2	Doob's theorem on sub (super) martingale almost sure convergence	161
7.4.3	Martingale decomposition and almost sure convergence	162
7.4.4	Robbins–Siegmund theorem and its corollaries	164
Chapter 8 Limit Theorems as Invariant Laws		175
8.1	Characteristics of dependence	176
8.1.1	Main concepts of dependence	176
8.1.2	Some inequalities for the covariance and mixing coefficients	182
8.1.3	Analog of Doob's inequality for mixingales	186
8.2	Law of large numbers	189
8.2.1	Weak law of large numbers	189
8.2.2	Strong law of large numbers	191
8.3	Central limit theorem	209
8.3.1	The i.i.d. case	210
8.3.2	General case of independent random variables	211
8.3.3	CLT for martingale arrays (or double sequences)	214
8.4	Logarithmic iterative law	225
8.4.1	Brief survey	225
8.4.2	Main result on the relative compactness	226
8.4.3	The classical form of LIL	235

PART III CONTINUOUS TIME PROCESSES 237

Chapter 9 Basic Properties of Continuous Time Processes		239
9.1	Main definitions	239
9.1.1	Sample path or trajectory	239
9.1.2	Finite-joint distributions	240

9.2	Second-order processes	241
9.2.1	Quadratic-mean continuity	241
9.2.2	Separable stochastic processes	242
9.2.3	Criterion of mean-square continuity	243
9.3	Processes with orthogonal and independent increments	244
9.3.1	Processes with orthogonal increments	244
9.3.2	Processes with stationary orthogonal increments	246
9.3.3	Processes with independent increments	247
9.3.4	Poisson process	248
9.3.5	Wiener process or Brownian motion	254
9.3.6	An invariance principle and LIL for Brownian motion	255
9.3.7	'White noise' and its interpretation	259
Chapter 10 Markov Processes		263
10.1	Definition of Markov property	263
10.1.1	Main definition	263
10.1.2	Criterion for a process to have the Markov property	264
10.2	Chapman–Kolmogorov equation and transition function	267
10.2.1	Transition probability and its four main properties	267
10.2.2	Two-step interpretation of the Chapman–Kolmogorov equation	267
10.2.3	Homogeneous Markov processes	268
10.2.4	Process with independent increments as MP	269
10.2.5	Strong Markov property	271
10.3	Diffusion processes	271
10.3.1	Main definition	271
10.3.2	Kolmogorov's backward and forward differential equation	272
10.4	Markov chains	277
10.4.1	Main definitions	277
10.4.2	Expectation time before changing a state	278
10.4.3	Ergodic theorem	281
Chapter 11 Stochastic Integrals		287
11.1	Time-integral of a sample-path	288
11.1.1	Integration of a simple stochastic trajectory	288
11.1.2	Integration of the product of a random process and a deterministic function	290
11.2	λ -stochastic integrals	292
11.2.1	General discussion	292
11.2.2	Variation of the sample path of a standard one-dimensional Wiener process	293

11.2.3	Mean square λ -approximation	295
11.2.4	Itô (non-anticipating) case ($\lambda = 0$)	297
11.2.5	Stratonovich case ($\lambda = 1/2$)	299
11.3	The Itô stochastic integral	299
11.3.1	The class of quadratically integrable random non-anticipating step functions	299
11.3.2	The Itô stochastic integral as the function of the upper limit	304
11.3.3	The Itô formula	306
11.4	The Stratonovich stochastic integral	317
11.4.1	Main property of λ -stochastic integrals	317
11.4.2	The Stratonovich differential	320
11.4.3	Multidimensional case	322

Chapter 12 Stochastic Differential Equations **323**

12.1	Solution as a stochastic process	323
12.1.1	Definition of a solution	323
12.1.2	Existence and uniqueness	324
12.1.3	Dependence on parameters and on initial conditions	332
12.1.4	Moments of solutions	335
12.2	Solutions as diffusion processes	338
12.2.1	General Markov property	338
12.2.2	Solution as a diffusion process	339
12.3	Reducing by change of variables	342
12.3.1	General description of the method	342
12.3.2	Scalar stochastic differential equations	343
12.4	Linear stochastic differential equations	346
12.4.1	Fundamental matrix	346
12.4.2	General solution	349

PART IV APPLICATIONS **355**

Chapter 13 Parametric Identification **357**

13.1	Introduction	357
13.1.1	Parameters estimation as a component of identification theory	357
13.1.2	Problem formulation	358
13.2	Some models of dynamic processes	359
13.2.1	Autoregression (AR) model	359
13.2.2	Regression (R) model	359
13.2.3	Regression–autoregression (RAR) model	360

13.2.4	Autoregression–moving average (ARMA) model	360
13.2.5	Regression–autoregression–moving average (RARMA or ARMAX) model	361
13.2.6	Nonlinear regression–autoregression–moving average (NRARMAX) model	362
13.3	LSM estimating	363
13.3.1	LSM deriving	363
13.3.2	Recurrent matrix version of LSM	365
13.4	Convergence analysis	367
13.4.1	Unbiased estimates	367
13.4.2	Asymptotic consistency	368
13.5	Information bounds for identification methods	381
13.5.1	Cauchy–Schwartz inequality for stochastic matrices	381
13.5.2	Fisher information matrix	383
13.5.3	Cramér–Rao inequality	388
13.6	Efficient estimates	389
13.6.1	Efficient and asymptotic efficient estimates	389
13.6.2	Recurrent LSM with a nonlinear residual transformation	395
13.6.3	The best selection of nonlinear residual transformation	403
13.6.4	Asymptotically efficient procedure under correlated noise	405
13.7	Robustification of identification procedures	406
13.7.1	The Huber robustness	406
13.7.2	Robust identification of static (regression) models	408
13.7.3	Robust identification of dynamic (autoregression) models	414
Chapter 14 Filtering, Prediction and Smoothing		417
14.1	Estimation of random vectors	417
14.1.1	Problem formulation	417
14.1.2	Gauss–Markov theorem	418
14.1.3	Linear unbiased estimates	419
14.1.4	Lemma on normal correlation	421
14.2	State-estimating of linear discrete-time processes	422
14.2.1	Description of linear stochastic models	422
14.2.2	Discrete-time Kalman filtering	422
14.2.3	Discrete-time prediction and smoothing	425
14.3	State-estimating of linear continuous-time processes	427
14.3.1	Structure of an observer for linear stochastic processes	427
14.3.2	Continuous-time linear filtering	431
14.3.3	Continuous-time prediction and smoothing	436

Chapter 15 Stochastic Approximation	439
15.1 Outline of chapter	439
15.2 Stochastic nonlinear regression	440
15.2.1 Nonlinear regression problem	440
15.2.2 Robbins–Monro procedure and convergence with probability one	441
15.2.3 Asymptotic normality	444
15.2.4 Logarithmic iterative law	449
15.2.5 Step-size optimization	449
15.2.6 Ruppert–Polyak version with averaging	452
15.2.7 Convergence under dependent noises	456
15.3 Stochastic optimization	462
15.3.1 Stochastic gradient method	462
15.3.2 Kiefer–Wolfowitz procedure	465
15.3.3 Random search method	468
Chapter 16 Robust Stochastic Control	471
16.1 Introduction	471
16.2 Problem setting	472
16.2.1 Stochastic uncertain systems	472
16.2.2 Terminal condition, feasible and admissible controls	475
16.2.3 Highest cost function and robust optimal control	476
16.3 Robust stochastic maximum principle	477
16.3.1 First- and second-order adjoint processes	477
16.3.2 Main result	479
16.4 Proof of Theorem 16.1	480
16.4.1 Formalism	480
16.4.2 Proof of Properties 1–3	482
16.4.3 Proof of Property 4 (maximality condition)	486
16.5 Discussion	492
16.5.1 The important comment on Hamiltonian structure	492
16.5.2 RSMP for the control-independent diffusion term	492
16.5.3 The standard stochastic maximum principle	493
16.5.4 Deterministic systems	493
16.5.5 Comment on possible non-fixed horizon extension	493
16.5.6 The case of absolutely continuous measures for uncertainty set	494
16.5.7 Uniform density case	495
16.5.8 Can the complementary slackness inequalities be replaced by equalities?	495
16.6 Finite uncertainty set	495
16.6.1 Main result	496

16.6.2	Min-max production planning	497
16.6.3	Min-max reinsurance–dividend management	503
16.7	Min-Max LQ-control	508
16.7.1	Stochastic uncertain linear system	508
16.7.2	Feasible and admissible control	509
16.7.3	Min-max stochastic control problem setting	509
16.7.4	The problem presentation in Mayer form	510
16.7.5	Adjoint equations	511
16.7.6	Hamiltonian form	512
16.7.7	Basic theorem	512
16.7.8	Normalized form for the adjoint equations	515
16.7.9	The extended form for the closed-loop system	516
16.7.10	Riccati equation and robust optimal control	518
16.7.11	Linear stationary systems with infinite horizon	523
16.7.12	Numerical example	526
16.8	Conclusion	527
	Bibliography	529
	Index	535