

The Probabilistic Method

Third Edition

Noga Alon

*School of Mathematics
Raymond and Beverly Sackler Faculty of Exact Sciences
Tel Aviv University*

Joel H. Spencer

*Courant Institute of Mathematical Sciences
New York University*



WILEY

A JOHN WILEY & SONS, INC., PUBLICATION

**WILEY-INTERSCIENCE
SERIES IN DISCRETE MATHEMATICS AND OPTIMIZATION**

A complete list of titles in this series appears at the end of this volume.

Contents

xiii
xv

PREFACE

ACKNOWLEDGMENTS

PART I METHODS

1 The Basic Method	1
1.1 The Probabilistic Method	1
1.2 Graph Theory	1
1.3 Combinatorics	3
1.4 Combinatorial Number Theory	9
1.5 Disjoint Pairs	10
1.6 Exercises	11
<i>The Probabilistic Lens: The Erdős–Ko–Rado Theorem</i>	<i>13</i>
2 Linearity of Expectation	15
2.1 Basics	15
2.2 Splitting Graphs	16
2.3 Two Quickies	18
2.4 Balancing Vectors	19
2.5 Unbalancing Lights	21
2.6 Without Coin Flips	22
2.7 Exercises	23
<i>The Probabilistic Lens: Brégman's Theorem</i>	<i>24</i>

3 Alterations	27
3.1 Ramsey Numbers	27
3.2 Independent Sets	29
3.3 Combinatorial Geometry	30
3.4 Packing	31
3.5 Recoloring	32
3.6 Continuous Time	35
3.7 Exercises	39
<i>The Probabilistic Lens: High Girth and High Chromatic Number</i>	41
4 The Second Moment	43
4.1 Basics	43
4.2 Number Theory	44
4.3 More Basics	47
4.4 Random Graphs	49
4.5 Clique Number	53
4.6 Distinct Sums	54
4.7 The Rödl Nibble	56
4.8 Exercises	61
63	63
<i>The Probabilistic Lens: Hamiltonian Paths</i>	67
5 The Local Lemma	67
5.1 The Lemma	70
5.2 Property B and Multicolored Sets of Real Numbers	71
5.3 Lower Bounds for Ramsey Numbers	73
5.4 A Geometric Result	74
5.5 The Linear Arboricity of Graphs	78
5.6 Latin Transversals	79
5.7 The Algorithmic Aspect	82
5.8 Exercises	83
<i>The Probabilistic Lens: Directed Cycles</i>	85
6 Correlation Inequalities	86
6.1 The Four Functions Theorem of Ahlswede and Daykin	89
6.2 The FKG Inequality	90
6.3 Monotone Properties	92
6.4 Linear Extensions of Partially Ordered Sets	94
6.5 Exercises	95
<i>The Probabilistic Lens: Turán's Theorem</i>	

7 Martingales and Tight Concentration	97
7.1 Definitions	97
7.2 Large Deviations	99
7.3 Chromatic Number	101
7.4 Two General Settings	103
7.5 Four Illustrations	107
7.6 Talagrand's Inequality	109
7.7 Applications of Talagrand's Inequality	113
7.8 Kim–Vu Polynomial Concentration	115
7.9 Exercises	116
<i>The Probabilistic Lens: Weierstrass Approximation Theorem</i>	117
8 The Poisson Paradigm	119
8.1 The Janson Inequalities	119
8.2 The Proofs	121
8.3 Brun's Sieve	124
8.4 Large Deviations	127
8.5 Counting Extensions	129
8.6 Counting Representations	130
8.7 Further Inequalities	133
8.8 Exercises	135
<i>The Probabilistic Lens: Local Coloring</i>	136
9 Pseudorandomness	139
9.1 The Quadratic Residue Tournaments	140
9.2 Eigenvalues and Expanders	143
9.3 Quasirandom Graphs	149
9.4 Exercises	156
<i>The Probabilistic Lens: Random Walks</i>	157
PART II TOPICS	
10 Random Graphs	161
10.1 Subgraphs	162
10.2 Clique Number	164
10.3 Chromatic Number	166
10.4 Zero-One Laws	167
10.5 Exercises	175
<i>The Probabilistic Lens: Counting Subgraphs</i>	177

11 The Erdős–Rényi Phase Transition 179

- 11.1 An Overview 180
- 11.2 Three Processes 182
- 11.3 The Galton–Watson Branching Process 183
- 11.4 Analysis of the Poisson Branching Process 184
- 11.5 The Graph Branching Model 186
- 11.6 The Graph and Poisson Processes Compared 187
- 11.7 The Parametrization Explained 190
- 11.8 The Subcritical Regimes 190
- 11.9 The Supercritical Regimes 191
- 11.10 The Critical Window 194
- 11.11 Analogies to Classical Percolation Theory 197
- 11.12 Exercises 201

The Probabilistic Lens: The Rich Get Richer 203

12 Circuit Complexity 205

- 12.1 Preliminaries 205
- 12.2 Random Restrictions and Bounded-Depth Circuits 207
- 12.3 More on Bounded-Depth Circuits 211
- 12.4 Monotone Circuits 214
- 12.5 Formulae 217
- 12.6 Exercises 218

The Probabilistic Lens: Maximal Antichains 219

13 Discrepancy 221

- 13.1 Basics 221
- 13.2 Six Standard Deviations Suffice 223
- 13.3 Linear and Hereditary Discrepancy 226
- 13.4 Lower Bounds 229
- 13.5 The Beck–Fiala Theorem 231
- 13.6 Exercises 232

The Probabilistic Lens: Unbalancing Lights 234

14 Geometry 237

- 14.1 The Greatest Angle Among Points in Euclidean Spaces 238
- 14.2 Empty Triangles Determined by Points in the Plane 239
- 14.3 Geometrical Realizations of Sign Matrices 241
- 14.4 e -Nets and VC-Dimensions of Range Spaces 243

- 14.5 Dual Shatter Functions and Discrepancy 248
- 14.6 Exercises 251

The Probabilistic Lens: Efficient Packing 252

15 Codes, Games and Entropy 255

- 15.1 Codes 255
- 15.2 Liar Game 258
- 15.3 Tenure Game 260
- 15.4 Balancing Vector Game 261
- 15.5 Nonadaptive Algorithms 264
- 15.6 Half Liar Game 264
- 15.7 Entropy 266
- 15.8 Exercises 271

The Probabilistic Lens: An Extremal Graph 273

16 Derandomization 275

- 16.1 The Method of Conditional Probabilities 275
- 16.2 d -Wise Independent Random Variables in Small Sample Spaces 280
- 16.3 Exercises 284

The Probabilistic Lens: Crossing Numbers, Incidences, Sums and Products 285

17 Graph Property Testing 289

- 17.1 Property Testing 289
- 17.2 Testing Colorability 290
- 17.3 Szemerédi’s Regularity Lemma 294
- 17.4 Testing Triangle-Freeness 298
- 17.5 Characterizing the Testable Graph Properties 300
- 17.6 Exercises 302

The Probabilistic Lens: Turán Numbers and Dependent Random Choice 303

Appendix A: Bounding of Large Deviations 307

- A.1 Chernoff Bounds 307
- A.2 Lower Bounds 315
- A.3 Exercises 320

The Probabilistic Lens: Triangle-Free Graphs Have Large Independence Numbers 321

Appendix B: Paul Erdős 323

B.1 Papers 323

B.2 Conjectures 325

B.3 On Erdős 326

B.4 Uncle Paul 327

References 331

Author Index 345

Subject Index 349

Preface

The Probabilistic Method is one of the most powerful and widely used tools applied in combinatorics. One of the major reasons for its rapid development is the important role of randomness in theoretical computer science and in statistical physics.

The interplay between discrete mathematics and computer science suggests an algorithmic point of view in the study of the probabilistic method in combinatorics and this is the approach we tried to adopt in this book. The book thus includes a discussion of algorithmic techniques together with a study of the classical method as well as the modern tools applied in it. The first part of the book contains a description of the tools applied in probabilistic arguments, including the basic techniques that use expectation and variance, as well as the more recent applications of martingales and correlation inequalities. The second part includes a study of various topics in which probabilistic techniques have been successful. This part contains chapters on discrepancy and random graphs, as well as on several areas in theoretical computer science: circuit complexity, computational geometry, and derandomization of randomized algorithms. Scattered between the chapters are gems described under the heading *The Probabilistic Lens*. These are elegant proofs that are not necessarily related to the chapters after which they appear and can usually be read separately.

The basic Probabilistic Method can be described as follows: In order to prove the existence of a combinatorial structure with certain properties, we construct an appropriate probability space and show that a randomly chosen element in this space has the desired properties with positive probability. This method was initiated by Paul Erdős, who contributed so much to its development over a fifty year period, that it seems appropriate to call it "The Erdős Method." His contribution can be measured not only by his numerous deep results in the subject, but also by his many intriguing problems and conjectures that stimulated a big portion of the research in the area.

It seems impossible to write an encyclopedic book on the Probabilistic Method; too many recent interesting results apply probabilistic arguments, and we do not even try to mention all of them. Our emphasis is on methodology, and we thus try to describe the ideas, and not always to give the best possible results if these are too

technical to allow a clear presentation. Many of the results are asymptotic, and we use the standard asymptotic notation: for two functions f and g , we write $f = O(g)$ if $f \leq cg$ for all sufficiently large values of the variables of the two functions, where c is an absolute positive constant. We write $f = \Omega(g)$ if $g = O(f)$ and $f = \Theta(g)$ if $f = O(g)$ and $f = \Omega(g)$. If the limit of the ratio f/g tends to zero as the variables of the functions tend to infinity we write $f = o(g)$. Finally, $f \sim g$ denotes that $f = (1 + o(1))g$, that is, f/g tends to 1 when the variables tend to infinity. Each chapter ends with a list of exercises. The more difficult ones are marked by (*). The exercises enable readers to check their understanding of the material and also provide the possibility of using the book as a textbook.

This is the third edition of the book; it contains several improved results and covers various additional topics that developed extensively during the last few years. The additions include a modern treatment of the Erdős-Rényi phase transition discussed in Chapter 11, focusing on the behavior of the random graph near the emergence of the giant component and briefly exploring its connection to classical percolation theory. Another addition is Chapter 17, Graph Property Testing—a recent topic that combines combinatorial, probabilistic and algorithmic techniques. This chapter also includes a proof of the Regularity Lemma of Szemerédi (described in a probabilistic language) and a presentation of some of its applications in the area. Further additions are two new *Probabilistic Lenses*, several additional exercises, and a new part in Appendix A focused on lower bounds.

It is a special pleasure to thank our wives, Nurit and Mary Ann. Their patience, understanding and encouragement have been key ingredients in the success of this enterprise.

NOGA ALON
JOEL H. SPENCER

Acknowledgments

We are very grateful to all our students and colleagues who contributed to the creation of this third edition by joint research, helpful discussions and useful comments. These include Miklos Bona, Andrzej Dudek, Mathieu Dutour, Juliana Freire, Sarel Har-Peled, Johan Hastad, Rani Hod, Mihyun Kang, Michael Krivelevich, Eyal Lubetzky, Russell Lyons, Nabil Mustafa, Nathan Linial, Yuval Peres, Xue Rui, Alexander Sapozhenko, Asaf Shapira, Aravind Srinivasan, Benny Sudakov, Prasad Tetali and William Wu, who pointed out various inaccuracies and misprints, and suggested improvements in the presentation as well as in the results. Needless to say, the responsibility for the remaining mistakes, as well as the responsibility for the (hopefully not many) new ones, is solely ours.

It is a pleasure to thank Rani Hod and Eyal Lubetzky for their great technical help in the preparation of the final manuscript for this book.