

Modern Discrete Probability

Providing a graduate-level introduction to discrete probability and its applications, this book develops a toolkit of essential techniques for analyzing stochastic processes on graphs, other random discrete structures, and algorithms.

The topics covered in this book include the first and second moment methods, concentration inequalities, coupling and stochastic domination, martingales and potential theory, spectral methods, and branching processes. Each chapter expands on a fundamental technique, outlining common uses and showing them in action on simple examples and more substantial classical results. The focus is predominantly on non-asymptotic methods and results.

All chapters provide a detailed background review section, plus exercises and signposts to the wider literature. Readers are assumed to have undergraduate-level linear algebra and basic real analysis, while prior exposure to graduate-level probability is recommended.

This much-needed broad overview of discrete probability could serve as a textbook or as a reference for researchers in mathematics, statistics, data science, computer science, and engineering.

SÉBASTIEN ROCH is Professor of Mathematics at the University of Wisconsin–Madison. He has received an NSF CAREER award, an Alfred P. Sloan Fellowship, and a Simons Fellowship in Mathematics, and is a Fellow of the Institute for Mathematical Statistics.

CAMBRIDGE SERIES IN STATISTICAL AND
 PROBABILISTIC MATHEMATICS

Editorial Board

- Z. Ghahramani (Department of Engineering, University of Cambridge)
 R. Gill (Mathematical Institute, Leiden University)
 F. P. Kelly (Department of Pure Mathematics and Mathematical Statistics,
 University of Cambridge)
 B. D. Ripley (Department of Statistics, University of Oxford)
 S. Ross (Department of Industrial and Systems Engineering,
 University of Southern California)
 M. Stein (Department of Statistics, Rutgers University & Department of Statistics,
 University of Chicago)

This series of high-quality upper-division textbooks and expository monographs covers all aspects of stochastic applicable mathematics. The topics range from pure and applied statistics to probability theory, operations research, optimization, and mathematical programming. The books contain clear presentations of new developments in the field and also of the state of the art in classical methods. While emphasizing rigorous treatment of theoretical methods, the books also contain applications and discussions of new techniques made possible by advances in computational practice.

A complete list of books in the series can be found at www.cambridge.org/statistics.
 Recent titles include the following:

34. *Regression for Categorical Data*, by Gerhard Tutz
35. *Exercises in Probability (Second Edition)*, by Loïc Chaumont and Marc Yor
36. *Statistical Principles for the Design of Experiments*, by R. Mead, S. G. Gilmour and A. Mead
37. *Quantum Stochastics*, by Mou-Hsiung Chang
38. *Nonparametric Estimation under Shape Constraints*, by Piet Groeneboom and Geurt Jongbloed
39. *Large Sample Covariance Matrices and High-Dimensional Data Analysis*, by Jianfeng Yao, Shurong Zheng and Zhidong Bai
40. *Mathematical Foundations of Infinite-Dimensional Statistical Models*, by Evarist Giné and Richard Nickl
41. *Confidence, Likelihood, Probability*, by Tore Schweder and Nils Lid Hjort
42. *Probability on Trees and Networks*, by Russell Lyons and Yuval Peres
43. *Random Graphs and Complex Networks (Volume 1)*, by Remco van der Hofstad
44. *Fundamentals of Nonparametric Bayesian Inference*, by Subhashis Ghosal and Aad van der Vaart
45. *Long-Range Dependence and Self-Similarity*, by Vlasov Pipiras and Murad S. Taqqu
46. *Predictive Statistics*, by Bertrand S. Clarke and Jennifer L. Clarke
47. *High-Dimensional Probability*, by Roman Vershynin
48. *High-Dimensional Statistics*, by Martin J. Wainwright
49. *Probability: Theory and Examples (Fifth Edition)*, by Rick Durrett
50. *Model-based Clustering and Classification for Data Science*, by Charles Bouveyron *et al.*
51. *Spectral Analysis for Univariate Time Series*, by Donald B. Percival and Andrew T. Walden
52. *Statistical Hypothesis Testing in Context*, by Michael P. Fay and Erica H. Brittain
53. *The Fundamentals of Heavy Tails*, by Jayakrishnan Nair, Adam Wierman and Bert Zwart
54. *Random Graphs and Complex Networks (Volume 2)*, by Remco van der Hofstad
55. *Modern Discrete Probability*, by Sébastien Roch
56. *Generalized Additive Models for Location, Scale and Shape*, by Mikis D. Stasinopoulos *et al.*

Modern Discrete Probability

An Essential Toolkit

Sébastien Roch
University of Wisconsin–Madison





Shaftesbury Road, Cambridge CB2 8EA, United Kingdom
One Liberty Plaza, 20th Floor, New York, NY 10006, USA
477 Williamstown Road, Port Melbourne, VIC 3207, Australia
314–321, 3rd Floor, Plot 3, Splendor Forum, Jasola District Centre,
New Delhi – 110025, India
103 Penang Road, #05–06/07, Visioncrest Commercial, Singapore 238467

Cambridge University Press is part of Cambridge University Press & Assessment,
a department of the University of Cambridge.

We share the University's mission to contribute to society through the pursuit of
education, learning and research at the highest international levels of excellence.

www.cambridge.org

Information on this title: www.cambridge.org/9781009305112

DOI: 10.1017/9781009305129

© Sébastien Roch 2024

This publication is in copyright. Subject to statutory exception and to the provisions
of relevant collective licensing agreements, no reproduction of any part may take
place without the written permission of Cambridge University Press & Assessment.

First published 2024

A catalogue record for this publication is available from the British Library

Library of Congress Cataloging-in-Publication Data

Names: Roch, Sébastien, author.

Title: Modern discrete probability : an essential toolkit / Sébastien
Roch, University of Wisconsin, Madison.

Description: Cambridge, United Kingdom ; New York, NY : Cambridge
University Press, 2024. | Series: Cambridge series in statistical and
probabilistic mathematics ; 55 | Includes bibliographical references and index.

Identifiers: LCCN 2023044218 | ISBN 9781009305112 (hardback) | ISBN
9781009305129 (ebook)

Subjects: LCSH: Stochastic processes. | Graph theory.

Classification: LCC QA274.2 .R63 2024 | DDC 519.2/3–dc23/eng/20231023
LC record available at <https://lccn.loc.gov/2023044218>

ISBN 978-1-009-30511-2 Hardback

Cambridge University Press & Assessment has no responsibility for the persistence
or accuracy of URLs for external or third-party internet websites referred to in this
publication and does not guarantee that any content on such websites is, or will remain,
accurate or appropriate.

To Betsy

Contents

<i>Preface</i>	xi
<i>Notation</i>	xv
1 Introduction	1
1.1 Background	1
1.1.1 Review of Graph Theory	1
1.1.2 Review of Markov Chain Theory	7
1.2 Some Discrete Probability Models	14
Exercises	17
2 Moments and Tails	21
2.1 Background	21
2.1.1 Definitions	21
2.1.2 Basic Inequalities	22
2.2 First Moment Method	25
2.2.1 The Probabilistic Method	25
2.2.2 Boole's Inequality	28
2.2.3 ▷ Random Permutations: Longest Increasing Subsequence	30
2.2.4 ▷ Percolation: Existence of a Non-trivial Threshold on \mathbb{Z}^2	31
2.3 Second Moment Method	35
2.3.1 Paley–Zygmund Inequality	35
2.3.2 ▷ Random Graphs: Subgraph Containment and Connectivity in the Erdős–Rényi Model	37
2.3.3 ▷ Percolation: Critical Value on Trees and Branching Number	43
2.4 Chernoff–Cramér Method	49
2.4.1 Tail Bounds Via the Moment-Generating Function	49
2.4.2 Sub-Gaussian and Sub-Exponential Random Variables	54
2.4.3 ▷ Probabilistic Analysis of Algorithms: Knapsack Problem	61
2.4.4 Epsilon-Nets and Chaining	65
2.4.5 ▷ Data Science: Johnson–Lindenstrauss Lemma and Application to Compressed Sensing	72
2.4.6 ▷ Data Science: Classification, Empirical Risk Minimization, and VC Dimension	79
Exercises	90

3	Martingales and Potentials	95
3.1	Background	95
3.1.1	Stopping Times	95
3.1.2	▷ Markov Chains: Exponential Tail of Hitting Times and Some Cover Time Bounds	101
3.1.3	Martingales	103
3.1.4	▷ Percolation: Critical Regime on Infinite d -regular Tree	113
3.2	Concentration for Martingales and Applications	114
3.2.1	Azuma–Hoeffding Inequality	114
3.2.2	Method of Bounded Differences	115
3.2.3	▷ Random Graphs: Exposure Martingale and Application to the Chromatic Number in Erdős–Rényi Model	124
3.2.4	▷ Random Graphs: Degree Sequence of Preferential Attachment Graphs	128
3.2.5	▷ Data Science: Stochastic Bandits and the Slicing Method	133
3.2.6	Coda: Talagrand’s Inequality	139
3.3	Potential Theory and Electrical Networks	141
3.3.1	Martingales, the Dirichlet Problem and Lyapounov Functions	142
3.3.2	Basic Electrical Network Theory	150
3.3.3	Bounding the Effective Resistance via Variational Principles	158
3.3.4	▷ Random Walks: Pólya’s Theorem, Two Ways	169
3.3.5	▷ Randomized Algorithms: Wilson’s Method for Generating Uniform Spanning Trees	170
	Exercises	176
4	Coupling	182
4.1	Background	182
4.1.1	Basic Definitions	182
4.1.2	▷ Random Walks: Harmonic Functions on Lattices and Infinite d -regular Trees	184
4.1.3	Total Variation Distance and Coupling Inequality	186
4.1.4	▷ Random Graphs: Degree Sequence in Erdős–Rényi Model	192
4.2	Stochastic Domination	194
4.2.1	Definitions	195
4.2.2	Ising Model: Boundary Conditions	202
4.2.3	Correlation Inequalities: FKG and Holley’s Inequalities	205
4.2.4	▷ Random Graphs: Janson’s Inequality and Application to the Clique Number in the Erdős–Rényi Model	210
4.2.5	▷ Percolation: RSW Theory and a Proof of Harris’ Theorem	213
4.3	Coupling of Markov Chains and Application to Mixing	219
4.3.1	Bounding the Mixing Time via Coupling	219
4.3.2	▷ Random Walks: Mixing on Cycles, Hypercubes, and Trees	222
4.3.3	Path Coupling	229
4.3.4	▷ Ising Model: Glauber Dynamics at High Temperature	231
4.4	Chen–Stein Method	234

Contents

ix

4.4.1	Main Bounds and Examples	235
4.4.2	Some Motivation and Proof	241
4.4.3	▷ Random Graphs: Clique Number at the Threshold in the Erdős–Rényi Model	249
	Exercises	250
5	Spectral Methods	256
5.1	Background	256
5.1.1	Eigenvalues and Their Variational Characterization	256
5.1.2	Elements of Spectral Graph Theory	261
5.1.3	Perturbation Results	268
5.1.4	▷ Data Science: Community Recovery	270
5.2	Spectral Techniques for Reversible Markov Chains	276
5.2.1	Spectral Gap	278
5.2.2	▷ Random Walks: A Spectral Look at Cycles and Hypercubes	282
5.2.3	▷ Markov Chains: Varopoulos–Carne and Diameter-Based Bounds on the Mixing Time	287
5.2.4	▷ Randomized Algorithms: Markov Chain Monte Carlo and a Quantitative Ergodic Theorem	292
5.2.5	Spectral Radius	296
5.3	Geometric Bounds	300
5.3.1	Cheeger’s Inequality	304
5.3.2	▷ Random Walks: Trees, Cycles, and Hypercubes Revisited	308
5.3.3	▷ Random Graphs: Existence of an Expander Family and Application to Mixing	310
5.3.4	▷ Ising Model: Glauber Dynamics on Complete Graphs and Expanders	315
5.3.5	Congestion Ratio	320
	Exercises	323
6	Branching Processes	327
6.1	Background	327
6.1.1	Basic Definitions	327
6.1.2	Extinction	328
6.1.3	▷ Percolation: Galton–Watson Trees	332
6.1.4	Multitype Branching Processes	333
6.2	Random-Walk Representation	338
6.2.1	Exploration Process	338
6.2.2	Duality Principle	340
6.2.3	Hitting-Time Theorem	341
6.2.4	▷ Percolation: Critical Exponents on the Infinite b -ary Tree	345
6.3	Applications	348
6.3.1	▷ Probabilistic Analysis of Algorithms: Binary Search Tree	348
6.3.2	▷ Data Science: The Reconstruction Problem, the Kesten–Stigum Bound and a Phase Transition in Phylogenetics	356

6.4	▷ Finale: The Phase Transition of the Erdős–Rényi Model	368
6.4.1	Statement and Proof Sketch	368
6.4.2	Bounding Cluster Size: Domination by Branching Processes	370
6.4.3	Concentration of Cluster Size: Second Moment Bounds	379
6.4.4	Critical Case via Martingales	384
6.4.5	▷ Encore: Random Walk on the Erdős–Rényi Graph	388
	Exercises	391
Appendix A	Useful Combinatorial Formulas	397
Appendix B	Measure-Theoretic Foundations	398
B.1	Probability Spaces	398
B.2	Random Variables	400
B.3	Independence	403
B.4	Expectation	406
B.5	Fubini's Theorem	411
B.6	Conditional Expectation	413
B.7	Filtered Spaces	418
	<i>Bibliography</i>	419
	<i>Index</i>	429

Preface

This book arose from a set of lecture notes prepared for a one-semester topics course I taught at the University of Wisconsin–Madison in 2014, 2017, 2020, and 2023, which attracted a wide spectrum of students in mathematics, computer sciences, engineering, and statistics.

What Is It About?

The purpose of the book is to provide a graduate-level introduction to discrete probability. Topics covered are drawn primarily from stochastic processes on graphs: percolation, random graphs, Markov random fields, random walks on graphs, and so on. No attempt is made at covering these broad areas in depth. Rather, the emphasis is on illustrating important techniques used to analyze such processes. Along the way, many standard results regarding discrete probability models are worked out.

The “modern” in the title refers to the (nonexclusive) focus on non-asymptotic methods and results, reflecting the impact of the theoretical computer science literature on the trajectory of this field. In particular, several applications in randomized algorithms, probabilistic analysis of algorithms, and theoretical machine learning are used throughout to motivate the techniques described (although, again, these areas are not covered exhaustively).

Of course, the selection of topics is somewhat arbitrary and driven in part by personal interests. But the choice was guided by a desire to introduce techniques that are widely used across discrete probability and its applications. The material discussed here is developed in much greater depth in the following (incomplete list of) excellent textbooks and expository monographs, many of which influenced various sections of this book:

- Agarwal, Jiang, Kakade, Sun. *Reinforcement Learning: Theory and Algorithms*. [AJKS22]
- Aldous, Fill. *Reversible Markov Chains and Random Walks on Graphs*. [AF]
- Alon, Spencer. *The Probabilistic Method*. [AS11]
- B. Bollobás. *Random graphs*. [Bol01]
- Boucheron, Lugosi, Massart. *Concentration Inequalities: A Nonasymptotic Theory of Independence*. [BLM13]
- Chung, Lu. *Complex Graphs and Networks*. [CL06]
- Durrett. *Random Graph Dynamics*. [Dur06]
- Frieze and Karoński. *Introduction to Random Graphs*. [FK16]
- Grimmett. *Percolation*. [Gri10b]
- Janson, Luczak, Rucinski. *Random Graphs*. [JLR11]
- Lattimore, Szepesvári. *Bandit Algorithms*. [LS20]

- Levin, Peres, Wilmer. *Markov Chains and Mixing Times*. [LPW06]
- Lyons, Peres. *Probability on Trees and Networks*. [LP16]
- Mitzenmacher, Upfal. *Probability and Computing: Randomized Algorithms and Probabilistic Analysis*. [MU05]
- Motwani, Raghavan. *Randomized Algorithms*. [MR95]
- Rassoul-Agha, Seppäläinen. *A Course on Large Deviations with an Introduction to Gibbs Measures*. [RAS15]
- S. Shalev-Shwartz and S. Ben-David. *Understanding Machine Learning: From Theory to Algorithms*. [SSBD14]
- van Handel. *Probability in High Dimension*. [vH16]
- van der Hofstad. *Random Graphs and Complex Networks. Vol. 1*. [vdH17]
- Vershynin. *High-Dimensional Probability: An Introduction with Applications in Data Science*. [Ver18]

In fact, the book is meant as a first foray into the basic results and/or toolkits detailed in these more specialized references. My hope is that, by the end, the reader will have picked up sufficient fundamental background to learn advanced material on their own with some ease. I should add that I used many additional helpful sources; they are acknowledged in the “Bibliographic Remarks” at the end of each chapter. It is impossible to cover everything. Some notable omissions include, for example, graph limits [Lov12], influence [KS05], and group-theoretic methods [Dia88], among others. Much of the material covered here (and more) can also be found in [HMRAR98], [Gri10a], and [Bre17] with a different emphasis and scope.

Prerequisites

It is assumed throughout that the reader is fluent in undergraduate linear algebra, for example, at the level of [Ax15], and basic real analysis, for example, at the level of [Mor05]. In addition, it is recommended that the reader has taken at least one semester of graduate probability at the level of [Dur10]. I am also particularly fond of [Wil91], which heavily influenced Appendix B, where measure-theoretic background is reviewed. Some familiarity with countable Markov chain theory is necessary, as covered for instance in [Dur10, chapter 6]. An advanced undergraduate or Master’s level treatment such as [Dur12], [Nor98], [GS20], [Law06], or [Bre20] will suffice, however.

Organization of the Book

The book is organized around five major “tools.” The reader will have likely encountered those tools in prior probability courses. The goal here is to develop them further, specifically with their application to discrete random structures in mind, and to illustrate them in this setting on a variety of major classical results and applications.

In the interest of keeping the book relatively self-contained and serving the widest spectrum of readers, each chapter begins with a “background” section that reviews the basic

material on which the rest of the chapter builds. The remaining sections then proceed to expand on two or three important specializations of the tools. While the chapters are meant to be somewhat modular, results from previous chapters do occasionally make an appearance.

The techniques are illustrated throughout with simple examples first, and then with more substantial ones in separate sections marked with the symbol \triangleright . I have attempted to provide applications from many areas of discrete probability and theoretical computer science, although some techniques are better suited for certain types of models or questions. The examples and applications are important: many of the tools are quite straightforward (or even elementary), and it is only when seen in action that their full power can be appreciated. Moreover, the \triangleright sections serve as an excuse to introduce the reader to classical results and important applications – beyond their reliance on specific tools.

Chapter 1 introduces some of the main probability on graph models we come back to repeatedly throughout the book. It begins with a brief review of graph theory and Markov chain theory.

Chapter 2 starts out with the probabilistic method, including the first moment principle and second moment method, and then it moves on to concentration inequalities for sums of independent random variables, mostly sub-Gaussian and sub-exponential variables. It also discusses techniques to analyze the suprema of random processes.

Chapter 3 turns to martingales. The first main topic there is the Azuma–Hoeffding inequality and the method of bounded differences with applications to random graphs and stochastic bandit problems. The second main topic is electrical network theory for random walks on graphs.

Chapter 4 introduces coupling. It covers stochastic domination and correlation inequalities as well as couplings of Markov chains with applications to mixing. It also discusses the Chen–Stein method for Poisson approximation.

Chapter 5 is concerned with spectral methods. A major topic there is the use of the spectral theorem and geometric bounds on the spectral gap to control the mixing time of a reversible Markov chain. The chapter also introduces spectral methods for community recovery in network analysis.

Chapter 6 ends the book with applications of branching processes. Among other applications, an introduction to the reconstruction problem on trees is provided. The final section gives a detailed analysis of the phase transition of the Erdős–Rényi graph, where techniques from all chapters of the book are brought to bear.

Acknowledgments

The lecture notes on which this book is based were influenced by graduate courses of David Aldous, Steve Evans, Elchanan Mossel, Yuval Peres, Alistair Sinclair, and Luca Trevisan at UC Berkeley, where I learned much of this material. In particular, scribe notes for some of these courses helped shape early iterations of this book.

I have also learned a lot over the years from my collaborators and mentors as well as my former and current Ph.D. students and postdocs. I am particularly grateful to Elchanan Mossel and Allan Sly for encouragements to finish this project and to the UW–Madison students who have taken the various iterations of the course that inspired the book for their invaluable feedback.

Warm thanks to everyone in the departments of mathematics at UCLA and UW–Madison who have provided the stimulating environments that made this project possible. Beyond my current departments, I am particularly indebted to my colleagues in the NSF-funded Institute for Foundations of Data Science (IFDS) who have significantly expanded my knowledge of applications of this material in machine learning and statistics.

This book would have not have been possible without the hard work and advice of various people at Cambridge University Press and Integra, including Diana Gillooly, Natalie Tomlinson, Anna Scriven, Rebecca Grainger, Clare Dennison, Bhavani Vijayamani, and Sajukrishnan Balakrishnan, as well as several anonymous reviewers.

Finally, I thank my parents, my wife, and my son for their love, patience, and support.

Notation

Throughout the book, we will use the following notation.

- The real numbers are denoted by \mathbb{R} , the non-negative reals are denoted by \mathbb{R}_+ , the integers are denoted by \mathbb{Z} , the non-negative integers are denoted by \mathbb{Z}_+ , the natural numbers (i.e., positive integers) are denoted by \mathbb{N} , and the rational numbers are denoted by \mathbb{Q} . We will also use the notation $\overline{\mathbb{Z}}_+ := \{0, 1, \dots, +\infty\}$.
- For two reals $a, b \in \mathbb{R}$,

$$a \wedge b := \min\{a, b\}, \quad a \vee b := \max\{a, b\},$$

and

$$a^+ = 0 \vee a, \quad a^- = 0 \vee (-a).$$

- For a real a , $\lfloor a \rfloor$ is the largest integer that is smaller than or equal to a and $\lceil a \rceil$ is the smallest integer that is larger than or equal to a .
- For $x \in \mathbb{R}$, the natural (i.e., base e) logarithm of x is denoted by $\log x$. We also let $\exp(x) = e^x$. NATURAL LOGARITHM
- For a positive integer $n \in \mathbb{N}$, we let

$$[n] := \{1, \dots, n\}.$$

- The cardinality of a set A is denoted by $|A|$. The powerset of A is denoted by 2^A .
- For two sets A, B , their Cartesian product is denoted by $A \times B$.
- We will use the following notation for standard vectors: $\mathbf{0}$ is the all-zero vector, $\mathbf{1}$ is the all-one vector, and \mathbf{e}_i is the standard basis vector with a 1 in coordinate i and 0 elsewhere. In each case, the dimension is implicit, as well as whether it is a row or column vector.
- For a vector $\mathbf{u} = (u_1, \dots, u_n) \in \mathbb{R}^n$ and real $p > 0$, its p -norm (or ℓ^p -norm) is p -NORM

$$\|\mathbf{u}\|_p := \left(\sum_{i=1}^n |u_i|^p \right)^{1/p}.$$

When $p = +\infty$, we have

$$\|\mathbf{u}\|_\infty := \max_i |u_i|.$$

We also use the notation $\|\mathbf{u}\|_0$ to denote the number of non-zero coordinates of \mathbf{u} (although it is not a norm; see Exercise 1.1). For two vectors $\mathbf{u} = (u_1, \dots, u_n), \mathbf{v} = (v_1, \dots, v_n) \in \mathbb{R}^n$, their *inner product* is INNER PRODUCT

$$\langle \mathbf{u}, \mathbf{v} \rangle := \sum_{i=1}^n u_i v_i.$$

The same notations apply to row vectors.

- For a matrix A , we denote the entries of A either by $A(i, j)$ or by $A_{i,j}$ (unless otherwise specified). The i th row of A is denoted by $A(i, \cdot)$ or $A_{i,\cdot}$. The j th column of A is denoted by $A(\cdot, j)$ or $A_{\cdot,j}$. The transpose of A is A^T .
- For a vector $\mathbf{z} = (z_1, \dots, z_d)$, we let $\text{diag}(\mathbf{z})$ be the diagonal matrix with diagonal entries z_1, \dots, z_d .
- The *binomial coefficients* are defined as

$$\binom{n}{k} = \frac{n!}{k!(n-k)!},$$

where $k, n \in \mathbb{N}$ with $k \leq n$ and $n! = 1 \times 2 \times \dots \times n$ is the factorial of n . Some standard approximations for $\binom{n}{k}$ and $n!$ are listed in Appendix A. See also Exercises 1.2, 1.3, and 1.4.

- We use the abbreviation “a.s.” for “almost surely,” that is, with probability 1. We use “w.p.” for “with probability.”
- Convergence in probability is denoted as \rightarrow_p . Convergence in distribution is denoted as \xrightarrow{d} .
- For a random variable X and a probability distribution μ , we write $X \sim \mu$ to indicate that X has distribution μ . We write $X \stackrel{d}{=} Y$ if the random variables X and Y have the same distribution.
- For an event A , the random variable $\mathbf{1}_A$ is the indicator of A , that is, it is 1 if A occurs and 0 otherwise. We also use $\mathbf{1}\{A\}$.
- For probability measures μ, ν on a countable set S , their *total variation distance* is

$$\|\mu - \nu\|_{\text{TV}} := \sup_{A \subseteq S} |\mu(A) - \nu(A)|.$$

- For non-negative functions $f(n), g(n)$ of $n \in \mathbb{Z}_+$, we write $f(n) = O(g(n))$ if there exists a positive constant $C > 0$ such that $f(n) \leq Cg(n)$ for all n large enough. Similarly, $f(n) = \Omega(g(n))$ means that $f(n) \geq cg(n)$ for some constant $c > 0$ for all n large enough. The notation $f(n) = \Theta(g(n))$ indicates that both $f(n) = O(g(n))$ and $f(n) = \Omega(g(n))$ hold. We also write $f(n) = o(g(n))$ or $g(n) = \omega(f(n))$ or $f(n) \ll g(n)$ or $g(n) \gg f(n)$ if $f(n)/g(n) \rightarrow 0$ as $n \rightarrow +\infty$. If $f(n)/g(n) \rightarrow 1$ we write $f(n) \sim g(n)$. The same notations are used for functions of a real variable x as $x \rightarrow +\infty$.

BINOMIAL
COEFFICIENTS

TOTAL
VARIATION
DISTANCE