

Copyrighted Material

# Introduction to Probability for Computing

**Mor Harchol-Balter**



Copyrighted Material

# Introduction to Probability for Computing

MOR HARCHOL-BALTER

*Carnegie Mellon University, Pennsylvania*



# Contents

<b>Preface</b>	<i>page xv</i>
----------------	----------------

<b>Acknowledgments</b>	xxii
------------------------	------

## Part I Fundamentals and Probability on Events

<b>1 Before We Start ... Some Mathematical Basics</b>	2
1.1 Review of Simple Series	2
1.2 Review of Double Integrals and Sums	4
1.3 Fundamental Theorem of Calculus	7
1.4 Review of Taylor Series and Other Limits	8
1.5 A Little Combinatorics	11
1.6 Review of Asymptotic Notation	15
1.6.1 Big-O and Little-o	15
1.6.2 Big-Omega and Little-omega	17
1.6.3 Big-Theta	18
1.7 Exercises	18
<b>2 Probability on Events</b>	21
2.1 Sample Space and Events	21
2.2 Probability Defined on Events	22
2.3 Conditional Probabilities on Events	24
2.4 Independent Events	27
2.5 Law of Total Probability	30
2.6 Bayes' Law	32
2.7 Exercises	34

## Part II Discrete Random Variables

<b>3 Common Discrete Random Variables</b>	44
3.1 Random Variables	44
3.2 Common Discrete Random Variables	45
3.2.1 The Bernoulli( $p$ ) Random Variable	46
3.2.2 The Binomial( $n, p$ ) Random Variable	47
3.2.3 The Geometric( $p$ ) Random Variable	48

3.2.4	The Poisson( $\lambda$ ) Random Variable	49
3.3	Multiple Random Variables and Joint Probabilities	50
3.4	Exercises	54
<b>4</b>	<b>Expectation</b>	<b>58</b>
4.1	Expectation of a Discrete Random Variable	58
4.2	Linearity of Expectation	63
4.3	Conditional Expectation	67
4.4	Computing Expectations via Conditioning	72
4.5	Simpson's Paradox	74
4.6	Exercises	76
<b>5</b>	<b>Variance, Higher Moments, and Random Sums</b>	<b>83</b>
5.1	Higher Moments	83
5.2	Variance	85
5.3	Alternative Definitions of Variance	86
5.4	Properties of Variance	88
5.5	Summary Table for Discrete Distributions	91
5.6	Covariance	91
5.7	Central Moments	92
5.8	Sum of a Random Number of Random Variables	93
5.9	Tails	97
5.9.1	Simple Tail Bounds	98
5.9.2	Stochastic Dominance	99
5.10	Jensen's Inequality	102
5.11	Inspection Paradox	104
5.12	Exercises	107
<b>6</b>	<b>z-Transforms</b>	<b>116</b>
6.1	Motivating Examples	116
6.2	The Transform as an Onion	117
6.3	Creating the Transform: Onion Building	118
6.4	Getting Moments: Onion Peeling	120
6.5	Linearity of Transforms	121
6.6	Conditioning	123
6.7	Using z-Transforms to Solve Recurrence Relations	124
6.8	Exercises	128
<b>Part III Continuous Random Variables</b>		
<b>7</b>	<b>Continuous Random Variables: Single Distribution</b>	<b>134</b>
7.1	Probability Density Functions	134
7.2	Common Continuous Distributions	137
7.3	Expectation, Variance, and Higher Moments	141

7.4	Computing Probabilities by Conditioning on a R.V.	143
7.5	Conditional Expectation and the Conditional Density	146
7.6	Exercises	150
<b>8</b>	<b>Continuous Random Variables: Joint Distributions</b>	<b>153</b>
8.1	Joint Densities	153
8.2	Probability Involving Multiple Random Variables	156
8.3	Pop Quiz	160
8.4	Conditional Expectation for Multiple Random Variables	161
8.5	Linearity and Other Properties	163
8.6	Exercises	163
<b>9</b>	<b>Normal Distribution</b>	<b>170</b>
9.1	Definition	170
9.2	Linear Transformation Property	172
9.3	The Cumulative Distribution Function	173
9.4	Central Limit Theorem	176
9.5	Exercises	178
<b>10</b>	<b>Heavy Tails: The Distributions of Computing</b>	<b>181</b>
10.1	Tales of Tails	181
10.2	Increasing versus Decreasing Failure Rate	183
10.3	UNIX Process Lifetime Measurements	186
10.4	Properties of the Pareto Distribution	187
10.5	The Bounded-Pareto Distribution	189
10.6	Heavy Tails	189
10.7	The Benefits of Active Process Migration	190
10.8	From the 1990s to the 2020s	191
10.9	Pareto Distributions Are Everywhere	192
10.10	Summary Table for Continuous Distributions	194
10.11	Exercises	194
<b>11</b>	<b>Laplace Transforms</b>	<b>198</b>
11.1	Motivating Example	198
11.2	The Transform as an Onion	198
11.3	Creating the Transform: Onion Building	200
11.4	Getting Moments: Onion Peeling	201
11.5	Linearity of Transforms	203
11.6	Conditioning	203
11.7	Combining Laplace and z-Transforms	204
11.8	One Final Result on Transforms	205
11.9	Exercises	206



## Part IV Computer Systems Modeling and Simulation

<b>12</b>	<b>The Poisson Process</b>	210
12.1	Review of the Exponential Distribution	210
12.2	Relating the Exponential Distribution to the Geometric	211
12.3	More Properties of the Exponential	213
12.4	The Celebrated Poisson Process	216
12.5	Number of Poisson Arrivals during a Random Time	219
12.6	Merging Independent Poisson Processes	220
12.7	Poisson Splitting	221
12.8	Uniformity	224
12.9	Exercises	225
<b>13</b>	<b>Generating Random Variables for Simulation</b>	229
13.1	Inverse Transform Method	229
13.1.1	The Continuous Case	230
13.1.2	The Discrete Case	231
13.2	Accept–Reject Method	232
13.2.1	Discrete Case	233
13.2.2	Continuous Case	234
13.2.3	A Harder Problem	238
13.3	Readings	238
13.4	Exercises	238
<b>14</b>	<b>Event-Driven Simulation</b>	240
14.1	Some Queueing Definitions	240
14.2	How to Run a Simulation	242
14.3	How to Get Performance Metrics from Your Simulation	244
14.4	More Complex Examples	247
14.5	Exercises	249

## Part V Statistical Inference

<b>15</b>	<b>Estimators for Mean and Variance</b>	255
15.1	Point Estimation	255
15.2	Sample Mean	256
15.3	Desirable Properties of a Point Estimator	256
15.4	An Estimator for Variance	259
15.4.1	Estimating the Variance when the Mean is Known	259
15.4.2	Estimating the Variance when the Mean is Unknown	259
15.5	Estimators Based on the Sample Mean	261
15.6	Exercises	263
15.7	Acknowledgment	264

<b>16</b>	<b>Classical Statistical Inference</b>	<b>265</b>
16.1	Towards More General Estimators	265
16.2	Maximum Likelihood Estimation	267
16.3	More Examples of ML Estimators	270
16.4	Log Likelihood	271
16.5	MLE with Data Modeled by Continuous Random Variables	273
16.6	When Estimating More than One Parameter	276
16.7	Linear Regression	277
16.8	Exercises	283
16.9	Acknowledgment	284
<b>17</b>	<b>Bayesian Statistical Inference</b>	<b>285</b>
17.1	A Motivating Example	285
17.2	The MAP Estimator	287
17.3	More Examples of MAP Estimators	290
17.4	Minimum Mean Square Error Estimator	294
17.5	Measuring Accuracy in Bayesian Estimators	299
17.6	Exercises	301
17.7	Acknowledgment	304
<b>Part VI Tail Bounds and Applications</b>		
<b>18</b>	<b>Tail Bounds</b>	<b>306</b>
18.1	Markov's Inequality	307
18.2	Chebyshev's Inequality	308
18.3	Chernoff Bound	309
18.4	Chernoff Bound for Poisson Tail	311
18.5	Chernoff Bound for Binomial	312
18.6	Comparing the Different Bounds and Approximations	313
18.7	Proof of Chernoff Bound for Binomial: Theorem 18.4	315
18.8	A (Sometimes) Stronger Chernoff Bound for Binomial	316
18.9	Other Tail Bounds	318
18.10	Appendix: Proof of Lemma 18.5	319
18.11	Exercises	320
<b>19</b>	<b>Applications of Tail Bounds: Confidence Intervals and Balls and Bins</b>	<b>327</b>
19.1	Interval Estimation	327
19.2	Exact Confidence Intervals	328
19.2.1	Using Chernoff Bounds to Get Exact Confidence Intervals	328
19.2.2	Using Chebyshev Bounds to Get Exact Confidence Intervals	331
19.2.3	Using Tail Bounds to Get Exact Confidence Intervals in General Settings	332
19.3	Approximate Confidence Intervals	334

19.4	Balls and Bins	337
19.5	Remarks on Balls and Bins	341
19.6	Exercises	341
<b>20</b>	<b>Hashing Algorithms</b>	<b>346</b>
20.1	What is Hashing?	346
20.2	Simple Uniform Hashing Assumption	348
20.3	Bucket Hashing with Separate Chaining	349
20.4	Linear Probing and Open Addressing	352
20.5	Cryptographic Signature Hashing	355
20.6	Remarks	360
20.7	Exercises	360
<b>Part VII Randomized Algorithms</b>		
<b>21</b>	<b>Las Vegas Randomized Algorithms</b>	<b>364</b>
21.1	Randomized versus Deterministic Algorithms	364
21.2	Las Vegas versus Monte Carlo	366
21.3	Review of Deterministic Quicksort	367
21.4	Randomized Quicksort	368
21.5	Randomized Selection and Median-Finding	370
21.6	Exercises	373
<b>22</b>	<b>Monte Carlo Randomized Algorithms</b>	<b>383</b>
22.1	Randomized Matrix-Multiplication Checking	383
22.2	Randomized Polynomial Checking	387
22.3	Randomized Min-Cut	389
22.4	Related Readings	394
22.5	Exercises	394
<b>23</b>	<b>Primality Testing</b>	<b>403</b>
23.1	Naive Algorithms	403
23.2	Fermat's Little Theorem	404
23.3	Fermat Primality Test	408
23.4	Miller–Rabin Primality Test	410
23.4.1	A New Witness of Compositeness	410
23.4.2	Logic Behind the Miller–Rabin Test	411
23.4.3	Miller–Rabin Primality Test	413
23.5	Readings	415
23.6	Appendix: Proof of Theorem 23.9	415
23.7	Exercises	417

## Part VIII Discrete-Time Markov Chains



<b>24</b>	<b>Discrete-Time Markov Chains: Finite-State</b>	<b>420</b>
24.1	Our First Discrete-Time Markov Chain	420
24.2	Formal Definition of a DTMC	421
24.3	Examples of Finite-State DTMCs	422
24.3.1	Repair Facility Problem	422
24.3.2	Umbrella Problem	423
24.3.3	Program Analysis Problem	424
24.4	Powers of $\mathbf{P}$ : $n$ -Step Transition Probabilities	425
24.5	Limiting Probabilities	426
24.6	Stationary Equations	428
24.7	The Stationary Distribution Equals the Limiting Distribution	429
24.8	Examples of Solving Stationary Equations	432
24.9	Exercises	433
<b>25</b>	<b>Ergodicity for Finite-State Discrete-Time Markov Chains</b>	<b>438</b>
25.1	Some Examples on Whether the Limiting Distribution Exists	439
25.2	Aperiodicity	441
25.3	Irreducibility	442
25.4	Aperiodicity plus Irreducibility Implies Limiting Distribution	443
25.5	Mean Time Between Visits to a State	448
25.6	Long-Run Time Averages	450
25.6.1	Strong Law of Large Numbers	452
25.6.2	A Bit of Renewal Theory	454
25.6.3	Equality of the Time Average and Ensemble Average	455
25.7	Summary of Results for Ergodic Finite-State DTMCs	456
25.8	What If My DTMC Is Irreducible but Periodic?	456
25.9	When the DTMC Is Not Irreducible	457
25.10	An Application: PageRank	458
25.10.1	Problems with Real Web Graphs	461
25.10.2	Google's Solution to Dead Ends and Spider Traps	462
25.10.3	Evaluation of the PageRank Algorithm and Practical Considerations	463
25.11	From Stationary Equations to Time-Reversibility Equations	464
25.12	Exercises	469
<b>26</b>	<b>Discrete-Time Markov Chains: Infinite-State</b>	<b>479</b>
26.1	Stationary = Limiting	479
26.2	Solving Stationary Equations in Infinite-State DTMCs	480
26.3	A Harder Example of Solving Stationary Equations in Infinite-State DTMCs	483
26.4	Ergodicity Questions	484
26.5	Recurrent versus Transient: Will the Fish Return to Shore?	487
26.6	Infinite Random Walk Example	490
26.7	Back to the Three Chains and the Ergodicity Question	492

---

26.7.1	Figure 26.8(a) is Recurrent	492
26.7.2	Figure 26.8(b) is Transient	492
26.7.3	Figure 26.8(c) is Recurrent	494
26.8	Why Recurrence Is Not Enough	494
26.9	Ergodicity for Infinite-State Chains	496
26.10	Exercises	498
<b>27</b>	<b>A Little Bit of Queueing Theory</b>	<b>510</b>
27.1	What Is Queueing Theory?	510
27.2	A Single-Server Queue	511
27.3	Kendall Notation	513
27.4	Common Performance Metrics	514
27.4.1	Immediate Observations about the Single-Server Queue	515
27.5	Another Metric: Throughput	516
27.5.1	Throughput for $M/G/k$	517
27.5.2	Throughput for Network of Queues with Probabilistic Routing	518
27.5.3	Throughput for Network of Queues with Deterministic Routing	519
27.5.4	Throughput for Finite Buffer	520
27.6	Utilization	520
27.7	Introduction to Little's Law	521
27.8	Intuitions for Little's Law	522
27.9	Statement of Little's Law	524
27.10	Proof of Little's Law	525
27.11	Important Corollaries of Little's Law	527
27.12	Exercises	531
	<b>References</b>	<b>539</b>
	<b>Index</b>	<b>544</b>