

Index

- 2-universal, 162
- 4-way independence, 168
- Affinity matrix, 201
- Algorithm
 - greedy k-clustering, 189
 - k-means, 185
 - singular value decomposition, 39
- Almost surely, 223
- Anchor Term, 281
- Aperiodic, 63
- Arithmetic mean, 369
- Bad pair, 225
- Bayes rule, 379
- Bayesian, 299
- Bayesian network, 299
- Belief Network, 299
- belief propagation, 298
- Bernoulli trials, 377
- Best fit, 29
- Bigoh, 360
- Binomial distribution, 218
 - approximated by Poisson, 377
- Boosting, 146
- Branching process, 241
- Breadth-first search, 230
- Cartesian coordinates, 9
- Cauchy-Schwartz inequality, 366, 368
- Central Limit Theorem, 375
- Characteristic equation, 388
- Characteristic function, 403
- Chebyshev's inequality, 5
- Chernoff bounds, 380
- Clustering, 182
 - k-center criterion, 189
 - k-means, 185
 - Social networks, 208
 - Sparse Cuts, 202
- CNF
 - CNF-sat, 248
- Cohesion, 204
- Combining expert advice, 143
- Commute time, 88
- Conditional probability, 372
- Conductance, 81
- Coordinates
 - Cartesian, 9
 - polar, 9
- Coupon collector problem, 90
- Cumulative distribution function, 372
- Current
 - probabilistic interpretation, 84
- Cycles, 236
 - emergence, 235
 - number of, 235
- Data streams
 - counting frequent elements, 165
 - frequency moments, 160
 - frequent element, 165
 - majority element, 165
 - number of distinct elements, 161
 - number of occurrences of an element, 164
 - second moment, 166
- Degree distribution, 218
 - power law, 218
- Depth-first search, 232
- Diagonalizable, 388
- Diameter of a graph, 225, 238
- Diameter two, 236
- dilation, 341
- Disappearance of isolated vertices, 236
- Discovery time, 85
- Distance
 - total variation, 68
- Distribution
 - vertex degree, 216
- Document ranking, 49

INDEX

- Effective resistance, 88
 Eigenvalue, 387
 Eigenvector, 42, 388
 Electrical network, 81
 Erdős Rényi, 215
 Error-correcting codes, 167
 Escape probability, 85
 Euler's constant, 91
 Event, 372
 Expander, 75
 Expected degree
 vertex, 215
 Expected value, 373
 Exponential generating function, 403
 Extinct families
 size, 245
 Extinction probability, 241, 243

 Finite fields, 405
 First-moment method, 223
 Fourier transform, 328, 403
 Frequency domain, 328

 $G(n,p)$, 215
 Gamma function, 10, 368
 Gaussian, 15, 376, 404
 fitting to data, 20
 tail, 371
 Gaussians
 separating, 18
 General tail bounds, 384
 Generating function, 241
 component size, 256
 for sum of two variables, 242
 Generating functions, 400
 Generating points in the unit ball, 13
 Geometric mean, 369
 Giant component, 216, 223, 227, 232, 236
 Gibbs sampling, 70
 Graph
 connectivity, 235
 resistance, 91
 Graphical model, 298
 Greedy
 k-clustering, 189
 Growth models, 254
 with preferential attachment, 260
 without preferential attachment, 254

 Haar wavelet, 342
 Harmonic function, 82
 Hash function
 universal, 162
 Heavy tail, 218
 Hidden Markov model, 295

 Hidden structure, 209
 Hitting time, 85, 96
 Hölder's inequality, 366, 369

 Immortality probability, 243
 Incoherent, 325, 328
 Increasing property, 223, 239
 unsatisfiability, 248
 Independence
 limited way, 167
 Independent, 372
 Indicator random variable, 226
 of triangle, 221
 Indicator variable, 374
 Ising model, 311
 Isolated vertices, 227, 236
 number of, 228

 Jensen's inequality, 370
 Johnson–Lindenstrauss lemma, 16, 17

 k-clustering, 189
 k-means clustering algorithm, 185
 Kernel methods, 201
 Kirchhoff's law, 83
 Kleinberg, 262

 Lagrange, 404
 Laplacian, 57
 Law of large numbers, 4, 6
 Learning, 109
 Linearity of expectation, 220, 373
 Lloyd's algorithm, 185
 Local algorithm, 262
 Long-term probabilities, 65

 m-fold, 240
 Markov chain, 63
 state, 67
 Markov Chain Monte Carlo, 64
 Markov random field, 301
 Markov's inequality, 5
 Matrix
 multiplication
 by sampling, 170
 diagonalizable, 388
 similar, 388
 Maximum-cut problem, 51
 Maximum likelihood estimation, 380
 Maximum likelihood estimator, 20
 Maximum principle, 82
 MCMC, 64
 Mean value theorem, 406
 Median, 375
 Metropolis–Hastings algorithm, 68

INDEX

- Mixing time, 65
- Model
 - random graph, 215
- Molloy Reed, 253
- Moment generating function, 403
- Mutually independent, 373

- Nearest neighbor problem, 18
- Nonuniform Random Graphs, 252
- Normalized conductance, 65, 73
- Number of triangles in $G(n, p)$, 221

- Ohm's law, 83
- Orthonormal, 394

- Page rank, 95
 - personalized, 98
- Persistent, 63
- Phase transition, 222
 - CNF-sat, 248
 - non-finite components, 258
- Poisson distribution, 377
- Polar coordinates, 9
- Polynomial interpolation, 167
- Positive semidefinite, 400
- Power iteration, 49
- Power law distribution, 218
- Power method, 39
- Power-law distribution, 252
- Principle component analysis, 44
- Probability density function, 372
- Probability distribution function, 372
- Psuedo random, 168
- Pure-literal heuristic, 249

- Queue, 249
 - arrival rate, 250

- Radon, 121
- Random graph, 215
- Random projection, 16
 - theorem, 17
- Random variable, 372
- Random walk
 - Eucleadean space, 92
 - in three dimensions, 93
 - in two dimensions, 92
 - on lattice, 92
 - undirected graph, 85
 - web, 95
- Rapid Mixing, 67
- Real spectral theorem, 389
- Replication, 240

- Resistance, 81, 91
 - effective, 84
- Restart, 95
 - value, 95
- Return time, 95

- Sample space, 372
- Sampling
 - length squared, 171
- Satisfying assignments
 - expected number of, 248
- Scale function, 342
- Scale vector, 342
- Second moment method, 221, 223, 224
- Sharp threshold, 223
- Similar matrices, 388
- Singular value decomposition, 29
- Singular vector, 31
 - first, 32
 - left, 34
 - right, 34
 - second, 32
- Six-degrees separation, 262
- Sketch
 - matrix, 173
- Sketches
 - documents, 177
- Small world, 261
- Smallest-clause heuristic, 249
- Spam, 97
- Spectral clustering, 190
- Stanley Milgram, 261
- State, 67
- Stirling approximation, 367
- Streaming model, 159
- Symmetric matrices, 389

- Tail bounds, 380, 384
- Tail of Gaussian, 371
- Taylor series, 362
- Threshold, 222
 - CNF-sat, 246
 - diameter $O(\ln n)$, 239
 - disappearance of isolated vertices, 227
 - emergence of cycles, 235
 - emergence of diameter two, 225
 - giant component plus isolated vertices, 237
- Time domain, 328
- Total variation distance, 68
- Trace, 397
- Triangle inequality, 366
- Triangles, 220

- Union bound, 373
- Unit clause heuristic, 249

INDEX

- Unitary matrix, 394
- Unsatisfiability, 248
- Variance, 374
- variational method, 367
- VC-dimension, 118
 - convex polygons, 121
 - finite sets, 121
 - half spaces, 121
 - intervals, 120
 - pairs of intervals, 121
- rectangles, 120
- spheres, 122
- Viterbi algorithm, 296
- Voltage
 - probabilistic interpretation, 83
- Wavelet, 341
- World Wide Web, 95
- Young's inequality, 366, 369