

Index

- $\mathbb{Z} := \{0, 1, 2, \dots\}$
 $\mathbb{N} := \{1, 2, 3, \dots\}$
 $A \otimes B$ for matrices, 325, 460
 $B(K, L)$, 209
 $C_X(\alpha)$, 153
 F_X , 50
 I_F , 49
 $L_X(\alpha)$, 155
 $M_X(\alpha)$, 146
 P_U^\perp , 299
 P_U , 299
 $R(\theta, \mathbf{A})$, 475
 $U \otimes V$ for subspaces, 326
 U^\perp , 299
 A^\dagger , 444
 $\text{Cov}(X, Y)$, 68
 $\mathbb{E}(X)$, 56, 61
 $\Gamma(K)$, 148
 $\Omega, \omega, \omega^{\text{act}}$, 36
 $\mathbb{P}(B | A)$, 7, 73
 $\mathbb{R}^J \otimes \mathbb{R}^K$, 324
 $\text{Var}(X)$, 66
 $\mathbf{Y}, \mathbf{y}^{\text{obs}}, \bar{\mathbf{y}}_{\text{obs}}$, 169
 $\mathbf{x} \otimes \mathbf{y}$ for vectors, 325, 460
 \mathbf{v}^\dagger , 448
 dev, Dev , 226
 $\ell(\theta; \mathbf{y})$, 182
 $\varphi(x), \Phi(x)$, 146
 $\varphi_X(\alpha)$, 166
 lhd , 171
 lr, LR , 224
 $\langle \mathbf{x}, \mathbf{y} \rangle$, 296
 complex case, 449
 μ_X , 56
 $\pi(\theta | \mathbf{y}^{\text{obs}})$, 200
 $\rho(X, Y)$, 68
 \mathcal{F} , 43
 \mathcal{L}^1 , 61
 \mathcal{L}^2 , 65
 $\mathcal{R}(\frac{1}{2}\theta, \mathcal{A})$, 475
 $\sigma_x, \sigma_y, \sigma_z$, 474
 $s_{\text{obs}}, S_n, s_n, \hat{\sigma}_{n-1}$, 170
 $\{F, G\}$ anticommutator, 479
 $c + [a, b], c[a, b]$, etc, 305
 f_X , 53
 $g_X(\alpha)$, 143
 p_X , 51
 $q(\mathbf{x}, \mathbf{x}), Q(\mathbf{X}, \mathbf{Y})$, etc, 304
 $\text{App}(f \leftarrow g)$, 197
 $\text{BVN}(\mu_x, \mu_y; \sigma_x^2, \sigma_y^2; \rho)$, 369
 $\text{Ent}(f_Z)$, 197
 $F_{r,s}^*(\eta)$, 305
 $\text{MVN}_n(\boldsymbol{\mu}, V)$, 367
 $\text{SN}(U)$, 297
 $\text{SO}(3)$, 475
 $\text{SRW}(p)$, 117
 $\text{SU}(2)$, 473
 $\text{power}(\theta)$, 225

 Actual outcome ω^{act} , 36, 48, 169
 Actual statistics, 48
 Addition rule, 6, 39, 42
 Adjoint, 323
 Hermitian, 444
 Akaike Information Criterion (AIC),
 236, 303, 378
 Aliasing problem, 333
 Almost surely, 44
 Ancillary statistics, 186, 263, 279
 ANCOVA example, 336
 ANOVA
 a trailer, 289
 introduction, 29
 modifications, 358

- on WinBUGS, 333
 - the Bayesian view, 333
 - theory, 326
 - with interaction and replication, 331
- ANOVA table, 307, 312
- Aspect experiment, 483
- Asymptotic normality of MLEs, 193
- Auto-regressive model
 - mentioned, 513
- Axiom of Choice, 498
- Axiomatization
 - full, 43
- Ballot theorem, 126
 - martingale proof, 414
- Banach–Tarski Paradox, 43
- Barron’s identity, 199
- Basis, 296
- Bayes’ formula, 77
 - for events, 75
 - in probability, 259
 - in Statistics
 - multi-parameter, 266
 - one-parameter, 200
 - simple case, 8
- Bayes’ Theorem, 75
- Bayesian filter
 - change-point detection, 84
- Bayesian hypothesis testing
 - see* Bayesian significance testing 234
- Bayesian posterior density, 200
- Bayesian priors, 200
 - multiparameter reference
 - difficulties with, 381
 - philosophy, 219
 - reference, 204, 436
 - important warning, 276
 - multiparameter, 379
 - use of imaginary results, 217
 - Your choice, 215
- Bayesian significance testing, 214, 219
 - sharp hypotheses, 234
- Behrens–Fisher problem, 308
- Bell inequalities, 483
- Berry–Esseen Theorem, 163
- Binomial coefficients, 9, 20–21
- Birthdays problem, 13
 - true story, 26
- Black–Scholes formula, 421
- Borel sets and functions, 45
- Borel–Cantelli Lemma
 - first, 45
 - proof, 60
 - second, 97
- Bounded-convergence theorem, 65
- Boys-and-Girls Principle, 432, 435
- Boys-and-Girls problem, 100
- Branching process, 394
 - martingale, 418
 - typed, 398
- Buffon’s needle, 56
- Canonical Commutation Relation (CCR), 455
 - Weyl’s formulation, 458
- Car convoys, 103
 - discussion, 507
- Car parking, 76
- Car-and-Goats problem, 15, 73
- Cauchy–Schwarz inequality, 66
- Central Limit Theorem (CLT), 156
 - and entropy, 198
 - for medians, 165
 - Multivariate, 374
- Change-pdf trick, 147
- Change-point-detection filter, 84
- Characteristic functions, 166
- Chi-Squared Principle (ChiSqP), 345
 - for discrete variables, 345
 - proof, 377
- Choleski decomposition, 366
- Classical Probability
 - as commutative Quantum Probability, 461
- Clifford algebra, 490
- CNOT gate, 463
- Coin tossing
 - rigorous model, 110
- Conditional expectation
 - $\mathbb{E}(X | A)$, 385

- $\mathbb{E}(X | Y)$, 403
- $\mathbb{E}(X | Y)$, Y discrete, 388
- $\mathbb{E}(X | \mathcal{G})$, 405
 - and independence, 402
 - fundamental property, 390
- Conditional Mean Formula, 390
- Conditional pdfs, 258
- Conditional probability
 - definition, 73
 - motivation, 6
 - recursive calculation, 82–87
- Conditional uniformity
 - and sufficiency, 435
 - for Poisson processes, 435
- Conditional variance, 391
- Conditional Variance Formula, 392
- Conditioning
 - need for care with, 260
- Confidence Intervals (CIs)
 - Bayesian, 200
 - examples, 175–181
 - Frequentist Definition, 173
 - Frequentist Interpretation, 174
- Confidence Regions
 - and hypothesis tests, 226
- Conformal field theory, 439
- Confounding, 179, 357
- Conjugate priors, 207
- Constant-risk principle, 411
- Continued fractions, 497
- Continuous RV (so called), 53
- Contours for BVN, 371
- Convex function, 64
- Convolution, 251
- Correlation coefficient, 68, 371
 - as measure of association, 371
 - Statistics of, 372
- Countable sets, 496
- Coupon collecting, 432
- Covariance, 68
- Cramér–Rao Theorem, 190
- Credible interval, 172
- Cromwell’s dictum, 216
- Crystallization, 36, 48
- Cumulant generating function, 153
- Cumulants, 153
- Cumulative Distribution Function, 50
- de Finetti’s Theorem, 220
 - martingale proof, 419
- Decoherence, 492
- Degrees of freedom, 306, 312, 327
- Density matrices in Quantum Theory, 452
 - for subsystems, 467
- Deutsch’s problem, 442, 446
- Deviance, 226
 - chi-squared results for, 230, 345
- Deviance Information Criterion
 - see also* Akaike Information Criterion
 - Bayesian, 378
- Difference equations, 118, 387
- Differential Geometry
 - in Statistics, 379
- Dimension, 296
- Dimerization, 76
- Dirac equation, 489
- Discrete Random Variable, 51
- Distribution Function F_X of X , 50
- Distribution-free tests, 317
- Edgeworth expansions, 163
- Effective prior sample size, 202
- Efficiency, 188
- Einstein–Podolsky–Rosen paradox, 481
- Empirical Distribution Function (EDF), 116
- Entanglement, 465, 480
- Entangling circuits, 465
- Entropy, 196
 - and Fisher information, 198
 - and the CLT, 198
- EPR–Bohm paradox, 481
- Ergodic property, 269
- Estimators and estimates, 187
- Euler’s formula, 101
- European option, 422
- Events, 36–39
- Exchangeability, 220
- Exclusion principle
 - mentioned, 473

- Expectation, μ_X or $\mathbb{E}(X)$, of X
 general case, 59
- Exponential distribution
 lack-of-memory property, 82
- Exponential family of distributions, 184
- Extension-to-Basis Principle, 296, 297
- F-test, 301
 and choice of model, 303
 sharp-hypothesis nature, 302
- Fiducial school
 mentioned, 311
- Fisher and Mendel, 346
- Fisher information, 189
 and entropy, 198
 and the Uncertainty Principle, 459
- Fisher's z -transform, 373
- Fisher–Behrens problem, 308
- Fisher-information metric, 380
- Five-nations problem, 97
- Four-liars problem, 19
 solution, 79
- Fourier inversion formula, 457
- Fourier transform, 457
- Frequentist prior, 172
- Fundamental model, 44
 and coin tossing, 110
 universality, 112
- Fundamental Theorem of Statistics, 116
- Galton–Watson branching process, 394
- Gambler's ruin, 117, 413
- Gamma function $\Gamma(\cdot)$, 148
- Gaussian integral, 146
- General Linear Model, 336
- Generating functions
 general comments on, 142
- Genetics, 87–96
- Gibbs sampler, 268
see also WinBUGS
 Cauchy case, 273
 for $N(\mu, \text{prec } \rho)$, 268
 hierarchical t model, 351
 sticking phenomenon, 353
- Glivenko–Cantelli Theorem, 116
- Goodness of fit, 339–348
- Grover's algorithm, 446
 circuitry, 467
- Hadamard gate, 463
- Hamiltonian, 445
- Hardy–Weinberg law, 89
- Hat-matching problem, 40, 70
- Hazard functions, 80–82
- Hedging strategy, 422
- Heisenberg picture, 486
- Heisenberg Uncertainty Principle *see*
 Uncertainty Principle
- Hidden variables, 483
- Hierarchical models, 215
 t_ν, ν random, 348
- History \mathcal{F}_n , 128
- History \mathcal{F}_T , 129
- Hitting probabilities, 118
- Hypothesis testing, 222
 sequential, 416
- Hypothesis tests
 and confidence regions, 226
- IID Random Variables, 103
- Inclusion-exclusion principle
 general, 40
 proof, 58
- Independence, 8, 96–103, 241
- Independence means multiply
 for expectations, 101
 for MGFs, 152
 for pgfs, 144
 for probabilities, 96
- Indicator function, 49
- Individual Confidence Intervals, 328
- Interaction, 331, 360
- Interarrival times, 428
- Interchange circuit, 464
- Intervals
 notation for, 305
- Jacobian, 243–246
- Jacobian Theorem, 246
- Jensen's inequality, 64
- Joint pdf, 246
- Joint pmf, 240

- Kolmogorov Statistic, 339
 - distribution, 342
- Kullback–Leibler relative entropy, 197
- Kurtosis, 154

- Lack-of-memory property, 82
- Laplace transform, 155
- Law of the iterated logarithm, 116
- Least-squares methods, 69
 - see also* Linear Regression
- Leverage, 316
- Likelihood function, 181
 - examples, 182
- Likelihood-ratio martingale, 409
- Likelihood-ratio test, 223
- Linear Algebra, 295–297
- Linear Model
 - General, 336
- Linear regression
 - a trailer, 288
 - Bayesian view, 332
 - goodness of fit, 355
 - in Probability, 69
 - in Statistics, 312
 - introduction, 29
 - on WinBUGS, 332
- Linear transformation, 323
- Location parameter
 - Bayesian theory, 202
 - CIs for, 186, 264
 - special case, 263
- Log-likelihood function, 182
 - examples, 182
- Logic gates, 445
- Logistic regression, 357
- Long-term-relative-frequency idea, 5
 - logical flaws in, 25
 - resolution, 115
- Lung cancer and smoking, 179

- Marginal pmf, 241
- Markov Chain Monte Carlo (MCMC)
 - see* Gibbs sampler
 - see also* WinBUGS
- Markov's inequality, 106
- Martingale
 - 'Markovian', 409
 - 'product', 408
 - 'sum', 408
 - definition, 406
 - Likelihood-ratio, 409
- Martingale-convergence theorem, 418
- Marx brothers, 41
- Maximum absolute deviation, 330
- Maximum-entropy distributions, 198
- Maximum-Likelihood Estimators (MLEs), 192
 - asymptotic normality, 193, 375
- Mean *see* Expectation
- Mean-square-error, 69
- Measure theory, 42–46
 - policy on, 23
- Median, 72
- Minimal-sufficient statistic, 262
- Minimum-Variance Bound, 190
- Mixtures of conjugates, 213
- Mode, 210
- Model choice, 378
 - in polynomial regression, 320
 - introduction to AIC, etc, 236
 - references, 522
- Moment Generating Function (MGF), 146
- Moment problem
 - non-uniqueness, 500
- Moments, 151
- Monotone-convergence theorem, 60
- Monster group, 439
- Moonshine, 439
- Multilinear regression
 - remarks on, 322
- Multiplication rule
 - general, 7
- Multiply means independence
 - in Quantum Theory, 461
- Multivariate normal
 - conditioning, 369
 - pdf, 367

- National Lottery, 10, 18
- Nelder's dictum, 319
- No-cloning theorem, 471

- Non-Borel set, 499
- Non-linear regression, 322
- Non-locality, 481
- Non-orthogonality, 316
- Non-parametric Statistics, 317
- Normal integral, 146
- Normal-sampling theorem, 305
 - a trailer, 286
- NOT gate, 463
- Notational conflicts, 170
- Null event, 44

- $O(\cdot)$ notation, 495
- $o(\cdot)$ notation, 496
- Observable, 444
- OK Corral, 410
- Option pricing, 421
- Order statistics, 107
- Oriented parallelogram, 243
- Orthogonality implies independence
 - within MVN, 368
- Orthonormal basis, 296
- Orthonormality principle, 297
 - 2-dimensional story, 283
 - for projections, 300
- Outliers, 340

- p-value of data, 224
- Parallel-axis result, 67
- Parseval's formula, 457
- Parsimonious models, 321
- Partitioned matrices, 368
- Pascal's triangle, 20
- Paths
 - counting, 125
- Patterns
 - waiting for, 19, 21, 22, 413
- Pauli Exclusion Principle
 - mentioned, 473
- Pauli matrices $\sigma_x, \sigma_y, \sigma_z$, 474
- Pdf of a sum, 251
- Pearson Statistic, 346, 374
- Perfect simulation
 - mentioned, 275
- Permutations, 9
- Perpendicular projections, 299
 - characterization, 324
- Phase, 445
- Point Estimation, 187
- Poisson Line Process (PLP), 437
- Poisson Point Process (PPP), 434
 - rigorous construction, 436
- Poisson process
 - definitions, 427
 - large gaps in, 22
 - martingale characterization, 433
- Polarization, 482
- Polya's urn, 76
 - and martingale convergence, 418
 - Bayesian isomorphism, 79
 - main martingale, 409
 - other martingales, 410
- Polynomial regression, 318
- Positive-definite symmetric matrices, 366
- Possible outcome ω , 36
- Power function for test, 225
- Pre-experiments, 202
- Pre-priors, 202
- Pre-Statistics, 48
- Predicting new observations, 211, 314
- Predictive probabilities, 211
- Principle of parsimony, 236, 303, 321
- Probability
 - addition rule
 - extended, 39
 - full, 42
 - simple, 39
 - monotone-convergence properties, 43
- Probability and Statistics
 - notational conflicts, 170
 - their interrelation, 27
- Probability density function (pdf), 53
 - transformation rules, 55
- Probability generating function (pgf), 143
- Probability mass function (pmf), 51
- Pure-product states, 461

- quant-ph, 523
- Quantile-quantile plots, 340

- Quantum circuit, 463
- Quantum computers
 - NMR implementation, 470
- Quantum gate, 463
- Quantum Probability
 - and 'Realism', 491
 - and Classical Probability, 461
 - fundamental postulate, 450
- Quaternions, 474
- Qubits, 444

- Random effects
 - in ANOVA, 358
- Random sum
 - generating functions, 393
 - variance, 392
- Random Variable
 - definition, 47
 - fractal, 123
 - need for measurability, 48
- Random Walk
 - expected number of visits, 120
 - on \mathbb{Z}^2 , 121
 - on fractal, 122
 - on free group, 122
 - Simple Random Walk SRW, 117
- Random-effects model, 358
- Random-number generator, 130
 - C code, 134
- Range of a transformation, 323
- Realization, 36
- Record problem, 103
- Reference priors
 - see under* Bayesian priors
- Reflection principle, 125
- Regular pdfs, 172
- Rejection sampling, 137
- Relativity, 488
- Replication, 331
- Residual, 28
- Residuals, 285, 286, 288, 291, 305
 - looking at, 356, 360
- Return probabilities, 118
- Reverse martingales, 419
- Riemann zeta-function, 100
- Robustness of methods, 354

- Rotations, 475
- RSS, 305, 312, 327
- rss 28,
 - see also* RSS
- Rules for expectations, 62, 70

- Saddle-point method
 - some references, 164
- Sample Mean
 - mean and variance, 104
- Sample Median, 108
 - CLT for, 165
 - use of, 109
- Sampling with replacement, 99
- Sampling without replacement, 71
- Scale parameter, 206, 227
- Scheffé's Lemma, 161
- Schor's algorithm
 - mentioned, 442
- Schrödinger picture, 486
- Schrödinger's equation, 445
- score, 189
- SD line, 371
- Self-adjoint, 324
 - Hermitian case, 444
- Sequential testing, 416
- Sharp hypotheses, 233, 234
- Sharp hypothesis
 - F-test and, 302
- Shrinkage (mentioned), 284
- Signed area, 243
- Significance level of data, 224
- Significance level of test, 224
- Simpson's paradox, 179
- Simulating from
 - Beta distributions, 250
 - exponential distributions, 136
 - Gamma distributions, 256
 - normal distributions
 - polar method, 250
 - rejection sampling, 137
 - t distributions, 254, 256
 - two-sided exponential
 - distributions, 136
- Simulation
 - example, 32

- F-inverse principle, 50
 - ratio method, 255
- Simultaneous Confidence Intervals, 277, 328
- Size of test, 224
- Skewness, 154
- Smoking and lung cancer, 179
- Span, spanning set, 295
- Spearman's rank-correlation coefficient, 317
- Special distributions
 - Bernoulli, 51
 - Beta(K, L), 209
 - binomial, 10
 - Bivariate Normal, 369
 - Cauchy, 55
 - chi-squared, 152
 - non-central, 153
 - exponential E(rate λ) or E(λ), 54
 - $F_{r,s}$, 252
 - Gamma(K , mean μ), 149
 - Gamma(K , rate λ), 149
 - geometric, 53
 - hypergeometric, 17–18
 - log-normal, 148
 - Multivariate Normal, 367
 - negative binomial, 53
 - normal $N(\mu, \sigma^2)$, 54
 - normal $N(\theta, \text{prec } \rho)$, 207
 - Poisson, 51
 - standard normal $SN(U)$ on subspace U , 298
 - Standardized t, $StandT_\nu$, 253
 - Student's t_ν , 253
 - uniform $U[a, b]$, 54
 - Weibull, 82
- Special Relativity, 488
- Spectral Theorem, 449
 - for symmetric matrices, 365
- Spin, 477
 - as intrinsic angular momentum, 490
 - spinor representation, 478
- Standard deviation
 - definition, 68
- Standardized form of an RV, 156
- Static variables in C, 499
- Statistical Mechanics
 - some references, 439
- Statistics and Probability
 - their interrelation, 27
- Sticking phenomenon for Gibbs sampler, 353
- Stirling's formula, 11
 - accurate version, 148
- Stochastic process
 - discrete-time, 406
- Stopping time, 128
- Stopping-time principle, 412
- Strong Law of Large Numbers (SLLN), 113
 - for coin tossing, 111
 - martingale proof, 420
 - proof of special case, 113
- Strong Markov principle, 129
- Studentized range, 329
- Sufficient statistic
 - Bayesian version, 203
 - first definition, 183
 - Frequentist interpretation, 183–186
 - minimal, 262
- Sums of independent RVs
 - Gamma, 152
 - normal, 152
 - Poisson, 144
- Superposition, 444
 - versus probabilistic mixture, 452
- Tchebychev's inequality, 105
- Tensor products, 324, 460
- Test for disease, 3–4, 7, 8
- The Chi-Squared Principle (ChiSqP)
 - for goodness of fit, 347
- Time series
 - some references, 513
- Tower property, 402
- Transforming observations, 148
- Two-Envelopes 'Paradox', 16, 399, 404, 501
- Two-sample problem
 - free, 308

- paired, 308
- pooled, 311
- Type I and Type II errors, 226

- Unbiased Estimator, 188
 - of variance, 104
- Uncertainty Principle
 - and Fisher information, 459
 - as Mathematics
 - general case, 453
 - as Physics, 454
 - for position and momentum, 454
- Uncountable sets, 496
- Unitary matrices, 445
- Universality for quantum gates, 469
- Upper percentage points, 304

- Variance
 - as measure of spread, 67
 - definition, 66
 - of sum of independent RVs, 102

- Waiting for patterns, 19, 21, 22, 413
- Waiting-time paradox, 434
- Weak Law of Large Numbers (WLLN),
107
- Weierstrass approximation theorem
 - proof of, 109
- Welch statistic
 - mentioned, 311
- Wilks' Theorem, 345
 - proof, 377
- WinBUGS
 - for ANOVA, 333
 - for Cauchy, 272
 - for 'hierarchical t' model, 352
 - for modified ANOVA models, 358
 - for normal, 270
 - for regression, 332

- Zitterbewegung, 491