

## Index

- action, 1
  - space, 13, 54, 63, 71
- algorithm
  - Baum–Welch, 373, 378
  - coordinate ascent, 341
  - CuSum, 219, 222
  - dynamic programming, 374
  - expectation-maximization, 339, 378
  - forward-backward, 373, 375
  - gradient descent, 336
  - least mean squares, 350
  - recursive least squares, 350
  - SAGE, 341
  - stochastic gradient descent, 350
  - Viterbi, 373
- area under the curve, 52
- asymptotic equality, 196
  
- Bayes, 1, 12
- Benjamini–Hochberg procedure, 63, 65
- Berry–Esséen theorem, 198
- Bessel function, 132
- Bhattacharyya coefficient, 155, 162, 214
- biology, 340
- Bonferroni correction, 63
  
- Central Limit Theorem, 111, 186, 195, 198, 351
  - Lindeberg condition, 202
- change point, 217, 224, 371
- character recognition, 54
- characteristic function, 125
- Chernoff–Stein lemma, 189
- classification, 54, 65, 67, 90, 173
- communications, 340
- continuous mapping theorem, 325, 393
- convergence, 1, 365, 393
  - almost surely, 326, 330, 336
  - in probability, 251, 326, 350
  - mean-square, 239, 244, 245, 326
- correlation detector, 7
- correlation statistic, 117
- correlator detector, 108
- cost, 1, 7, 259
  - a posteriori, 12, 26, 55, 266
  - absolute-error, 267
  - additive, 266, 267, 270
  - function, 71
  - matrix, 54
  - nonuniform, 76
  - posterior, 89
  - squared-error, 267, 280, 297
  - uniform, 28, 38, 41, 56, 60, 73, 91, 129, 209, 263, 267
- covariance kernel, 239, 244, 252
- Cramér–Rao lower bound, 297, 300
  - in exponential family, 308
  - vector case, 307
- cumulant-generating function, 163, 272
- CuSum
  - algorithm, 219, 222
  - test statistic, 218
  
- data mining, 63
- decision boundary, 56
- decision region, 54, 56
- decision rule, 1, 208, 259
  - admissible, 60, 88
  - Bayesian, 10, 55, 89, 105, 130
  - complete class, 88
  - correlation, 112
  - deterministic, 17, 54, 71
  - equalizer, 15, 33, 62
  - linear, 121
  - locally most powerful, 82, 127, 138
  - MAP, 28, 57, 89, 116
  - maximum-correlation, 116
  - minimax, 10, 32, 43, 93, 105
  - ML, 28
  - MPE, 28, 116, 152
  - Neyman–Pearson, 62, 77, 93
  - non-dominated, 88
  - power, 43, 63, 78, 85
  - quadratic, 121, 122
  - randomized, 16, 17, 36, 37, 40, 58, 71, 88
  - sensitivity, 76
  - significance level, 63, 80
  - uniformly most powerful, 77, 92
- decision theory, 1
- delta theorem, 325, 393
- detection, 1

- m*-ary, 54
- average delay, 223
- coherent, 107
- quickest, 208
- quickest change, 217
- random process, 231
- sequential, 208
- singular, 238, 244, 246, 251
- detector
  - deflection-based, 136
  - locally optimum, 127
- digital communications, 28, 39, 44, 118
- dimensionality reduction, 157
- Dirac impulse, 239
- discriminant function, 56
- distance, 145
  - Ali–Silvey, 145, 152, 252
  - Bhattacharyya, 149, 171, 173, 182
  - Chernoff information, 150, 171
  - Euclidean, 118
  - Hamming, 7
  - Hellinger, 152
  - Itakura–Saito, 155, 252
  - Mahalanobis, 115, 119, 266
  - total variation, 145, 151
- divergence
  - binary Chernoff function, 238
  - binary KL, 160
  - Bregman, 155, 272
  - Chernoff, 145, 149, 168, 171, 233, 248
  - conditional KL, 236
  - f*-, 145, 152, 252
  - Jensen–Shannon, 152, 162
  - Kullback–Leibler, 1, 145, 161, 169, 233, 247, 272, 302, 309, 379
  - Pearson's  $\chi^2$ , 151
  - Rényi, 150
  - symmetrized, 152
- divergence rate
  - Chernoff, 231, 252
  - Kullback–Leibler, 231, 252
- dominated convergence theorem, 326, 331
- eigenfunction, 239, 241, 245
- eigenvalue, 114, 119, 123
- eigenvector, 113, 119, 123, 158
- entropy, 148, 379
  - cross-, 379
- epsilon contamination, 97
- equalizer rule, 15
- error probability
  - Bayesian, 152, 170
  - conditional, 41, 55, 117, 160
  - false alarm, 208, 223
  - family-wise rate, 63
  - lower bound, 134, 160, 162
  - pairwise, 173
  - upper bound, 134, 160, 172
  - Wald's bounds, 213
- estimation, 1
  - MAP, 373
- estimator, 7
  - absolute-error, 7, 262
  - asymptotically efficient, 326, 329, 331
  - asymptotically Gaussian, 326, 329, 331
  - asymptotically unbiased, 326, 329
  - asymptotics, 325
  - bias, 297
  - biased, 281, 312, 322
  - conditional-mean, 260
  - conditional-median, 262
  - conditional-mode, 263
  - consistent, 319, 326, 329
  - efficient, 300, 303, 308, 319
  - invariant, 259, 322
  - James–Stein, 355
  - least-squares, 338, 349
  - linear, 259, 265, 268
  - linear MMSE, 268, 346, 358
  - M*-, 338
  - MAP, 263, 267, 276, 319, 370, 373
  - median, 333
  - minimax, 280
  - ML, 12, 319
  - MMAE, 262
  - MMSE, 122, 260, 276, 361, 364, 370, 372
  - MVUE, 280, 281, 297, 305, 324
  - randomized, 354
  - recursive, 347
  - recursive ML, 347
  - sample-mean, 333
  - squared-error, 7, 260
  - unbiased, 259, 281, 284, 289, 297, 300, 305, 307
  - weighted least-squares, 339
- exponential family, 270, 303, 305, 341
  - canonical form, 271, 310
  - completeness, 289
  - mean-value parameter, 273, 324
  - ML estimation, 323
  - natural parameter space, 271
- exponential tilting, 194
- factorization theorem, 280, 283, 290
- false alarm rate, 221
- false discoveries, 63
- false discovery rate, 64
- family of distributions
  - complete, 286, 288
  - Gaussian, 291
  - Gaussian mixture, 336
  - general mixture, 337, 341
  - incomplete, 287

- location-scale, 315
- Poisson, 315
- reparameterization, 303, 309
- filter
  - causal, 358
  - Kalman, 358, 360
  - linear, 317, 358
  - nonlinear, 358, 367, 369
- Fisher information, 1, 297, 298, 341
  - matrix, 306, 311, 324, 333, 337
  - metric, 309
- game theory, 11, 15
- genomics, 63
- handwritten digit recognition, 54
- Hoeffding problem, 189
- hypothesis
  - alternative, 25, 101
  - composite, 72
  - dummy, 116
  - null, 25, 72
  - simple, 72
- hypothesis test, 6
  - acceptance region, 25
  - Bayesian, 25, 54, 73, 188, 245
  - binary, 6, 160, 167, 208, 233
  - Cauchy, 78, 83, 86, 110
  - composite, 7, 71, 105, 106, 220
  - Gaussian, 58, 68, 71, 77, 84, 101
  - Laplacian, 102, 108
  - locally most powerful, 71
  - $m$ -ary, 90, 105, 115, 231, 283
  - minimax, 25, 32, 54, 58
  - multiple, 54
  - Neyman–Pearson, 25, 40, 54, 189, 210
  - non-dominated, 71, 91
  - Poisson, 67, 100
  - power, 40
  - rejection region, 25
  - robust, 71, 92
  - sequential, 208
  - significance level, 40, 43
  - ternary, 56, 67, 91, 182
  - uniformly most powerful, 71
- imaging, 340
- inequality
  - Bonferroni, 176
  - Cauchy–Schwarz, 118, 137, 163, 299, 305, 307
  - Chebyshev, 164, 167
  - data processing, 145, 153, 160, 176, 252
  - Fano, 176
  - generalized Fano, 176
  - information, 297, 299, 311
  - Jensen, 145, 152, 286
  - Markov, 164, 194, 214
  - Pinsker, 156
  - Vajda, 156
- information theory, 1, 145, 176
- innovation sequence, 358
- integral
  - Riemann–Stieltjes, 239
  - stochastic, 239
- Jacobian, 273, 309, 323, 324, 393
- joint stochastic boundedness, 93
- Kalman filter, 358, 360
  - extended, 367
  - scalar case, 365
  - steady state, 365
- Karhunen–Loève decomposition, 239
- large deviations, 126, 160, 184
  - Bahadur–Rao theorem, 196
  - Cramér’s theorem, 185
  - Gärtner–Ellis theorem, 201
  - rate function, 165
  - refined, 194
  - upper bound, 167
- law of iterated expectations, 289
- likelihood equation, 320, 336, 348
- likelihood function, 12, 302, 319
  - unbounded, 336
- likelihood ratio, 56, 63, 87, 106, 107, 233, 250
  - clipped, 97
  - generalized, 84
  - generalized test, 71, 130, 220
  - monotone, 79, 154
  - test, 27, 77, 99, 107
- limiter
  - soft, 110
- loss, 7
  - KL divergence, 270
- loss function
  - surrogate, 145
  - zero-one, 7
- Lyapunov equation, 365
- machine learning, 1, 152, 275
- marginal distribution, 12
- matched filter, 105, 108
- matrix
  - circulant Toeplitz, 232
  - determinant, 232
  - DFT, 232
  - Hermitian transpose, 232
  - Hessian, 321, 336
  - trace, 137, 234, 297
- measure, 1, 4, 92
  - counting, 4

- Lebesgue, 4, 260
- Mercer's theorem, 243
- minimax, 1
- mismatched model, 338
- moment-generating function, 126, 163
- Monte-Carlo simulations, 298
- mutual information, 148, 178
  
- navigation, 360, 364
- noise
  - additive, 139
  - Cauchy, 110
  - colored, 246, 248
  - correlated, 105, 112, 118
  - Gaussian, 105
  - generalized Gaussian, 139
  - i.i.d., 105, 118
  - Laplacian, 108, 128
  - multiplicative, 139
  - non-Gaussian, 105, 111, 364
  - whitening, 105, 113
  
- observations, 1, 259, 304
  - asymptotic setup, 325
  - missing, 340, 378
  - sequence, 360
- optimization
  - complexity, 374
  - convex, 46, 165, 339
  - linear, 62
- orthogonality principle, 260, 268, 347, 358, 362
  
- $p$ -value, 63
- parameter, 71
  - estimation, 84
  - identifiability, 302, 309, 330, 331, 333
  - nonrandom, 87, 91, 280, 297
  - on boundary, 320, 328
  - orthogonality, 309
  - precision, 274
  - random, 72, 87, 90, 260, 297, 311
  - set, 259
  - space, 71
  - vector, 297, 306, 321, 333
- parameter estimation, 6
  - Bayesian, 259
  - for linear Gaussian models, 265
- payoff function, 11, 15
- performance metrics, 7
  - deflection criterion, 105, 135
- prior, 9
  - beta, 278
  - conjugate, 273
  - Dirichlet, 275
  - exponential, 276
  - gamma, 274, 277
  - least favorable, 14, 33, 62, 94
  - uniform, 118, 132, 152, 177, 277, 319
- probability
  - a posteriori, 26, 55
  - conditional, 2
  - density function, 2, 196
  - detection, 115, 131
  - false alarm, 40, 72, 111, 115, 128, 129, 168
  - mass function, 2
  - miss, 26, 40, 72, 129, 168
  - nominal distribution, 97
  - of detection, 40
  - space, 231
  - tail, 126
- probability distribution, 2
  - binomial, 287
  - exponential, 287
  - geometric mixture, 190
  - mixed, 278
  - multinomial, 275
  - nominal, 97
- probability measure, 2
  - absolutely continuous, 251
  - dominated, 251
  - equivalent, 251
  - orthogonal, 251
- pseudo-inverse, 339
  
- Q function
  - asymptotics, 117, 133, 134
- quantization, 157
  
- radar, 28, 40, 54, 72, 118, 129, 364
- Radon–Nikodym derivative, 231, 251
- random process
  - bandlimited, 240, 242
  - Brownian motion, 239, 242, 246, 248
  - continuous-time, 231, 238
  - counting process, 248
  - discrete-time, 231
  - dynamical system, 358, 367
  - Gauss–Markov, 240
  - Gaussian, 231, 234, 239, 248, 252, 317
  - hidden Markov model, 358, 369, 372
  - independent increments, 239
  - Markov, 231, 235, 369, 372
  - mean-square continuous, 243
  - Ornstein–Uhlenbeck, 240, 242
  - orthogonal increments, 243
  - periodic stationary, 232, 240, 241, 248
  - Poisson, 231, 248
  - power spectral mass, 242
  - spectral density, 234, 239, 252
  - spectral representation, 243
  - stationary, 231, 234, 252
  - weakly stationary, 239

- white Gaussian, 239
- random variable
  - continuous, 2, 196
  - discrete, 2
  - lattice type, 196, 197
  - orthogonal, 241
  - Poisson, 276
- randomization parameter, 105
- Rao–Blackwell theorem, 280, 284, 289
- Receiver Operating Characteristic, 43, 200
  - lower bound, 172
  - upper bound, 160
- regularity conditions, 298, 306, 319, 329
- relative entropy, 145
- reparameterization, 303, 317, 322, 324
- Riccati equation, 350, 363, 365
- risk, 8
  - Bayes, 9, 55, 209
  - conditional, 8, 25, 54, 88, 91, 259, 280, 297
  - curve, 34, 41
  - line, 34, 41
  - minimax, 40
  - posterior, 12, 28, 76, 260
  - set, 59, 88
  - vector, 59, 88
  - worst-case, 10, 58, 61
- robotics, 360
  
- saddlepoint, 15, 33, 94, 126, 320
- sample size, 208
- score function, 331, 333
- sequential test
  - Bayesian, 210
  - non-Bayesian, 210
  - SPRT, 210
- signal, 6
  - antipodal, 118
  - classification, 7, 105, 115
  - correlated, 105
  - detection, 38, 105, 176
  - deterministic, 106, 244
  - energy, 132
  - energy constraint, 118
  - estimation, 6, 361
  - fixed-lag smoothing, 364
  - Gaussian, 120, 246
  - impulse, 135
  - non-Gaussian, 136
  - Poisson, 139
  - prediction, 6, 361
  - pseudo-, 112
  - random, 106
  - reconstruction, 244
  - sampling, 238
  - selection, 105, 117, 145
  - sinusoid, 131, 138, 301, 316, 368
  - smoothing, 6
  - weak, 127
- signal to noise ratio, 117, 300, 301
  - generalized, 114
- Slutsky’s theorem, 393
- software testing, 110
- sonar, 54
- speech recognition, 54
- state, 1
  - nonrandom, 1
  - of Nature, 8
  - random, 1
  - sequence, 360
  - space, 13, 54, 71
  - trellis, 374
- stochastic ordering, 93
- stopping rule, 208
  - Shewhart, 219
- stopping time, 208
- sufficient statistic
  - complete, 334
- sufficient statistics, 1, 280, 281, 341
  - minimal, 283, 289
  - trivial, 282, 288
  
- target detection, 65
- Taylor series, 82, 248, 331, 348, 350, 364, 367
- template matching, 129, 135
- threshold, 27, 63
- trace of a matrix, 137
- tracking, 360, 364
- transform
  - convex conjugate, 165
  - discrete Fourier, 126, 232
  - Esscher, 194, 198
  - Fourier, 125, 126, 242
  - Karhunen–Loève, 240
  - Laplace, 126, 288, 290
  - Legendre–Fenchel, 165
  - unitary, 232
  
- utility function, 7