
Index

- Abbreviations, 738–739
 Almost surely, 37
 Anderson–Darling statistic, 619
 Asymptotic equivalence, 377, 619, 711, 730
 Asymptotic theory, 371–435
- Bachelor’s problem, *see* Secretary problem
 Bandwidth, 53, 553, 576, 580, 583
 Basu’s theorem, 444, 476, 477, 479, 480, 528, 529, 665, 671
 Bayes’ law, 6, 7, 556
 Bayesian estimation, 556–557, 583–589
 admissibility, 587
 asymptotics, 588
 Bayes risk, 557, 587, 588
 highest posterior density (HPD), 559, 599
 posterior, 556
 posterior risk, 557
 prior, 556
 conjugate, 584
 diffuse, 584
 exponential, 584
 improper, 586
 Jeffreys’, 585, 604
 normal, 583
 not invariant, 585
 quadratic loss, 586
 Benford’s law, 342, 368
 Binomial coefficient, 8, 707
 Birthday problem, 30
 Black–Scholes–Merton formula, 126
 Bochner’s theorem, 94
 Bolzano–Weierstrass theorem, 386, 705
 Bootstrap, 456, 482
- Borel–Cantelli lemmas, 378, 380–382, 385, 401, 402, 426
 Box–Müller transformation, 257, 258
 Brownian motion, 147, 434
- Capture–recapture, 109, 496, 560
 Cauchy
 criterion, 705
 principal value, 59
 sequence, 705
 Central limit theorem (CLT), 374
 conditions violated but limit is normal, 417–419
 for Student $t(2)$ with infinite variance, 418
 Lindeberg–Feller, 415, 685
 Lindeberg–Lévy, 403, 596
 multivariate, 404
 Characteristic function (c.f.), 62
 and symmetry, 80
 expansion, 84
 joint, 175
 of matrix variates, 197
 positive semidefinite property, 94
 Characterization theory, 323
 Cochran’s theorem, 321
 Collinear (random variables), 174
 Combination, 8, 27–31
 Commutative law, 3, 699
 failure for scalars, 705
 Completeness, 705
 Complex numbers, 700–702
 argument, 700
 complex conjugate, 701
 modulus, 700

- Complex numbers (*Cont.*)
 - polar form, 700
- Conditioning, 19–27, 171–225
 - implies partitioning, 466, 668, 694
 - sequential, 23
- Confidence interval, 557
 - asymptotic, 595
 - conservative, 598
 - for quantile, 594
 - for variance ratio, 590
 - from binomial, 597
 - from gamma, 589
 - from Laplace, 589
 - from normal, 590, 591, 594, 599
 - shortest, 591, 593
- Confidence region, 557, 591, 604
 - elliptical, 598
- Continuity theorem, 387
- Continuous mapping theorem (CMT), 389
 - of Mann and Wald, 433
- Contour plot, 177
 - normal, 178, 565
- Convergence
 - almost surely (a.s.), 373, 379
 - dominated, 721
 - equivalences, 389
 - in distribution, 371
 - continuity points, 378
 - implied by convergence in probability, 383
 - of sums and products, 392
 - to uniform, 377
 - in moments, 384
 - in probability, 372
 - but not a.s., 380
 - implied by a.s. convergence, 381
 - of sums and products, 391
 - uniform integrability, *see* Integrability
- Convexity, 721
 - of Lagrangian, 724
- Convolution, 230, 240–243
- Copula, 168, 278, 365
- Correlation
 - and (non)linear dependence, 189, 400
 - coefficient, 174
 - conditional, 176
 - partial, 225
 - GLR test statistic, 635
 - matrix, 195
 - multiple, 194
 - smaller than 1 in absolute value, 193
- Coverage, 557, 591–596, 599, 604
- Craig–Sakamoto lemma, 316
- Cramér’s deconvolution theorem, *see* Deconvolution
 - theorems
 - Cramér–Rao lower bound (CRLB), 490, 502–504, 516, 517, 523, 525–527, 547
 - and exponential family, 490
 - Cramér–von Mises statistic, 619
 - Cramér–Wold device, 372, 390, 404
 - Cross validation, 454, 482, 580
 - Cumulant, 61, 90, 92
 - location and scale, 92
 - Cumulative distribution function (c.d.f.), 33
 - joint, 150
 - marginal, 150
 - of absolute value, 238
- Data-generating process (DGP), 445
- De Morgan’s law, 4, 699
- Decision theory, 440, 696
- Decomposition
 - Frisch–Waugh, 549, 563
 - Jordan, 35, 40, 71
 - Lebesgue, 36, 720
- Deconvolution theorems, 322
 - Cramér, 283, 287, 313
 - Raikov, 111, 283
- Delta method, 374, 395
- Difference calculus, 714–716
- Distribution, 737–738
 - Bernoulli, 106, 232
 - maximum likelihood, 493
 - beta
 - bivariate, 156, 157, 208
 - c.d.f., 113
 - moments, 130
 - order statistics, 276
 - plots, 116
 - relation to F, 134
 - binomial, 106, 112, 403
 - c.d.f., 113
 - from Bernoulli, 232
 - moments through f.m.g.f., 87
 - Cauchy, 84, 277, 376
 - from normal, 137
 - multivariate, 177
 - plot, 116
 - relation to Laplace, 132
 - chi-bar-squared ($\bar{\chi}^2$)
 - sum of squared truncated normals, 651, 695
 - chi-squared (χ^2)
 - from gamma, 250
 - independence, 318
 - mixture of normals, 55
 - noncentral, 98, 99, 128
 - plots, 116

Distribution (*Cont.*)

- quadratic form, 307
- sum of squared normals, 234
- conditional, 171, 198–225
- continuous, 116–138
 - overview, 100
- cumulative, 33, 35
 - location and scale, 74
- Dirichlet, 158, 169, 208
- discrete, 105–116
 - overview, 96
- elliptical, 177, 212, 220, 290
 - c.f., 265
 - conditional expectation, 221
 - correlation coefficient, 336, 410
 - fourth-order moments, 299, 409
- exponential, 120, 155, 237, 251, 277, 479
 - as maximum entropy, 142
 - as no-memory distribution, 51, 115, 120, 147, 237, 251, 274
 - maximum likelihood, 506
- exponential family, 102, 138
 - c.g.f., 464, 483
 - c.g.f. and tilting, 90, 270
 - completeness of, 464
 - Cramér–Rao lower bound, 490
 - curved, 103, 505, 550, 591
 - multivariate, 464
 - natural parameterization, 102, 464, 495, 665, 666, 669, 671
 - regular, 103
- exponentially tilted, 90, 270, 422
- Fisher–Snedecor (F)
 - from chi-squared, 135
 - from normal sample, 345
 - from spherical, 265, 317, 336, 534, 571, 640, 674
 - moments, 136
 - ratio of quadratic forms, 317
 - relation to beta, 134
 - relation to Student, 135, 265
- gamma, 248, 477
 - asymptotics of mean, 406
 - c.d.f., 122
 - infinitely divisible, 145
 - marginal, 206
 - mean squared error, 451
 - mixture of Poisson, 55, 121
 - moments, 65, 130
 - not stable, 145
- generalized error, 98
- generalized extreme value, 98, 99, 131, 148, 376, 430
- generalized gamma, 98, 123
 - maximum likelihood, 504
- geometric, 108, 232
 - as no-memory distribution, 115, 116, 147
 - maximum likelihood, 495
 - moments through f.m.g.f., 87
- hypergeometric, 97, 108, 109
 - maximum likelihood, 496
 - method of moments, 560
 - multivariate, 154
- inverse Gaussian, 147
 - plot, 116
- joint, 149–169
 - bivariate continuous, 156–162
 - discrete, 152–156
 - trivariate continuous, 162–166
- Lévy–Khinchine, 105
- Laplace, 49, 50, 132
 - maximum likelihood, 510
 - method of moments, 559
 - relation to Cauchy, 132
- limiting, 371
 - degenerate, 372
- log-normal, 82, 126
 - plot, 116
 - product, 255
- logistic, 133, 494
- maximum-entropy, 141, 142
- McKay’s bivariate gamma, 161
- multinomial, 153–155, 204, 273, 618
- negative binomial, 55, 97, 121
 - from geometric, 232
 - relation to binomial, 107
- negative hypergeometric, 109
- normal, *see* Normal distribution
- of functions of r.v.s., 45–51, 227–279
- of maximum (minimum), 45, 46, 63
- of ratio, 256
- Pareto, 80, 154, 375, 435, 477, 478
 - maximum likelihood, 502
 - method of moments, 562
- Poisson, 111, 419
 - arrival time, 343
 - as limit of binomial, 112
 - c.d.f., 121
 - compound, 104, 429
 - cumulant, 98
 - maximum likelihood, 497
 - mean squared error, 497
 - mixture, 54, 104
 - moments through f.m.g.f., 87
 - relation to gamma, 120
- Rademacher, 34

Distribution (*Cont.*)

- reproductive, 234
- spherical, 177, 264, 265
- stable, 105, 145, 375
- standardized, 99
- Student (*t*), 55
 - c.d.f., 113
 - from normal sample, 345
 - moments, 136
 - multivariate, 177, 212, 220
 - relation to *F*, 135
- symmetric, 38
- tails, 39
- trivariate, 210, 211
- uniform, 48
 - as maximum entropy, 141
 - completeness of, 460
 - from exponential, 121
 - maximum, 429
 - maximum likelihood, 499, 501, 560
 - mean squared error, 452, 499, 501, 560
 - method of moments, 560, 561
 - moments, 66, 105
 - on unit sphere, 264
 - over unit disk, 209
 - representation, 238, 257, 258
- Weibull
 - maximum likelihood, 498
- Wishart, 329
 - c.f., 350
 - conditional, 355
 - density, 368
 - inverse, 355
 - of sample variance, 351
 - reproductive, 353

Distribution-free methods, 594, 618, 680–698

Distributive laws, 3

Divergence

- Cressie–Read, 698
- Jeffreys, 143, 655
- Kullback–Leibler, 104

EM algorithm, 548

Empirical distribution function (EDF), 330, 361, 510

- asymptotics, 404

Entropy, 103, 139–144, 269, 510, 654, 697

- and log-likelihood, 490, 521, 548
- dependence on variance, 140
- for normal distribution, 140

Error term, 215

Estimator, 437

- admissible, 440

- Bayes, 587
 - sample mean is not admissible, 476
- best, 438
- best asymptotically normal (BAN), 492
- best linear unbiased (BLUE), 447, 552, 567
- bias, 438
 - of correlation, 470
- consistent, 439
 - rate, 439
 - strong, 439
- efficient
 - asymptotically, 492
 - relative, 438, 448
 - superefficient, 547
- invariant, 440
- jackknife, 454
- Kaplan–Meier, 511
- kernel density, 553, 574, 576
- linear, 447
- maximum likelihood, *see* Maximum likelihood
- minimax, 440
- Nadaraya–Watson, 555, 582, 583
- nonparametric density, 553
- nonparametric regression, 555
- pooled, 447, 590, 629
- robust, 439
- shrinkage, 475, 483, 565, 601–604
- unbiased, 437
 - best (BUE), 438, 468, 471, 473, 474
 - median, 438
 - vs. estimate, 437
- Euler’s formula, 700, 701, 709
- Euler’s gamma, 99
- Event, 4, 9–19
- Expectation, 57
 - as linear operator, 59
 - existence, *see* Moment, existence
 - symbol *E*, 57
- Extreme value theory (EVT), 274, 376, 430
- Factorial, 27–31
 - generalized, 714–715
- Failure rate, 51
- Feller’s condition, 416
- Fisher’s *z* transformation, 413, 471
- Fisher’s information matrix, 489
- Fourier transform, 62, 148, 175
- Function
 - beta, 8
 - incomplete, 113
 - expansion, *see* Series expansion
 - factorial, 7

- Function (*Cont.*)
- gamma, 7
 - incomplete, 102, 123, 420
 - Legendre's duplication, 358
 - hypergeometric, 123, 147, 367, 725
 - Gauss, 97, 115
 - semimetric, 143
 - slowly varying, 376
- Fundamental theorem of algebra, 703
- Galton's fallacy, 564, 600
- Gauss–Markov theorem, 552, 567, 601
- Generating function
- and moments, 79–93
 - cumulant- (c.g.f.), 61
 - factorial-moment- (f.m.g.f.), 61, 715
 - moment- (m.g.f.), 60
 - (non)existence, 62, 82, 94
 - inverse moments, 85
 - joint, 175
 - probability- (p.g.f.), 61
- Gnedenko's theorem, 430
- Gurland–Geary formula, 256
- Hamel's equation, 182, 431, 704
- Hardy's convergence theorem, 707
- Hazard rate, 50, 115, 213, 421, 511, 576
- Heine–Borel theorem, 705
- Heisenberg's uncertainty principle, 148
- Hellinger distance, 144
- Helly's selection lemma, 386
- Helly–Bray theorem, 385
- Helmert's transformation, 245, 278, 351
- Highest posterior density (HPD), 559, 599
- Hotelling's T^2 , 329, 357, 613, 621, 634, 660, 672
- Hypothesis
- alternative (H_1), 605
 - local, 637, 641, 677, 696
 - one-sided, 606
 - two-sided, 606
 - composite, 605
 - null (H_0), 605
 - simple, 605
- Hypothesis testing, 605–698
- power, 607, 608
 - power function, 610, 621, 660, 663, 677, 696
 - monotonicity, 621, 658, 659
 - significance level, 607
 - size, 606
 - Type I error, 607
 - Type II error, 607
- i.i.d., 97
- and memory, 401
- Incidental-parameters problem, 509, 543
- Independence, 5, 19–27
- joint, 166, 223
 - vs. mutually exclusive, 5, 22, 23
 - vs. pairwise, 5, 22, 167
 - not transitive, 22
 - subset, 167
- Indicator function, 36
- Inequality
- arithmetic–geometric mean, 73, 196, 331
 - Bonferroni, 18, 598, 604, 695
 - Boole, 18, 344, 598
 - Cauchy–Schwarz, 193, 194
 - Chebyshev, 71, 233
 - Chernoff, 71
 - exponential, 21
 - Jensen, 59, 60, 73, 77, 93, 287, 392
 - Markov, 71
 - median within a standard deviation from mean, 75
 - mode–median–mean not valid, 78
 - triangle, 67, 148, 391, 395
 - $W \geq LR \geq LM$, 637
- Infinite divisibility, 104, 145, 230, 429
- Infinitely often (i.o.), 379
- Integrability
- absolute, 720
 - uniform, 433, 721
- Interval estimation, 551, 589–600
- Invariance principle, 374
- Jacobian
- factor, 229
 - matrix, 229
 - of transformation, 229
- Karlin–Rubin theorem, 617, 657–661, 675, 693
- Kendall's τ , 363, 365, 366, 369, 683
- Kernel, 53, 553
- Epanechnikov, 554, 578, 580
 - normal, 554, 555, 578, 579
- Kolmogorov–Smirnov statistic, 434, 619
- Kuhn–Tucker conditions, 651, 724
- Kullback–Leibler information criterion (KLIC), 104, 143, 202, 291, 518, 620, 654, 695, 697
- invariance, 269
- Lagrange
- convexity, 724
 - function, 599, 722
 - multiplier method, 721–725
- Large-deviation theorems, 422

- LASSO, 601
- Law of iterated expectations (LIE), 176, 191, 199, 201, 203, 205, 442
- Law of large numbers
 strong (SLLN), 373
 Kolmogorov–Etemadi, 401
 weak (WLLN), 373
 Chebyshev, 400
 Khinchine, 399
- Law of the iterated logarithm (LIL), 375, 425
- Least absolute deviations (LAD), 75, 600
 quantile regression, 601
- Least squares (LS), 75, 552, 562–574
 conditional and marginal effects, 566
 constrained (CLS), 570
 generalized (GLS), 569
 not invariant, 564
- Legendre’s duplication formula, 358
- Lehmann–Scheffé theorem, 443, 468
- Leibniz’ rule, 713
- l’Hôpital’s rule, 712
- Lindeberg condition, 415
- Linear conditional expectation, 216, 217
- Linear difference equation, 717
- Linear model, 444
 generalized (GLM), 445
 joint sufficiency, 458
 orthogonal parameterization, 562
- Loss function, 440, 475
- Mahalanobis distance, 177, 292, 615, 620
- Marsaglia–Bray transformation, 258
- Matrix inverse
 generalized, 150, 525, 567, 645
 partitioned, 714
- Maximum likelihood (ML), 485–550
 asymptotic distribution, 495, 501
 best asymptotically normal (BAN), 523
See also density
 asymptotic efficiency, 492
 asymptotic linearity in score, 521
 asymptotic sufficiency, 521
 concentrated, *see* profile
 consistency, 495, 520
 inconsistent, 508
 rate, 501
 corner solution, 495
 density, 548, 693
 estimator, 485
 (not) unique, 488, 498, 499, 506, 521
 from Bernoulli, 493
 from exponential, 506
 from generalized gamma, 504
 from geometric, 495
 from hypergeometric, 496
 from Laplace, 510
 from normal, 486–487, 505, 508, 512
 from Pareto, 502
 from Poisson, 497
 from uniform, 499, 501, 560
 from Weibull, 498
 existence, 485, 497, 514, 521
 Fisher’s information matrix, 489
 for normal linear model, 524–540
 estimators, 524
 Hessian, 524
 log-likelihood, 524
 partitioned model, 537
 prediction, 536
 properties of $\hat{\sigma}^2$, 525
 properties of $\hat{\beta}$, 525
 score (vector), 524
 variance decompositions, 532
 with constant term, 530
 Hessian matrix, 488, 521, 541, 544
 identification, 489, 520, 547
 information-matrix equivalence, 489, 515, 694
 invariance, 493, 497, 503, 512, 515
 likelihood, 485
 empirical, 548, 698
 nonparametric, 510, 511
 log-likelihood, 487
 as empirical entropy, 490, 521, 548
 partial, 489, 514, 542, 548, 602
 profile, 488, 514, 540
 modified, 542, 548, 586
 score (vector), 487, 488, 600
 expectation not zero, 541, 543
 expectation zero, 515
 locally sufficient, 520
 martingale representation, 516
 not zero, 495
 time series, 544, 648
- Mean squared error (MSE), 438, 449
 gamma, 451
 integrated (IMSE), 554
 of sample correlation, 471
 of sample variance, 449, 567
 Poisson, 497
 uniform, 452, 499, 501, 560
- Mean teacher, 561
- Mean-independence, 218, 219, 221
- Mean-value theorem, 711, 714
- Median, 39, 43
- Method of moments (MM), 551–552, 559–562

- Method of moments (*Cont.*)
 estimator
 from hypergeometric, 560
 from Laplace, 559
 from normal, 551, 552
 from Pareto, 562
 from uniform, 560, 561
 generalized, 600
 Mills' ratio, 213, 421
 Minimum
 under constraints, 721–725
 Mixing, 40, 52–55, 237
 and symmetrization, 53
 Mode, 39, 44–45
 Moment, 57, 63–66
 asymptotic, 372, 384, 391, 547
 central, 57, 79
 coefficient of variation, 58
 covariance, 173, 188, 189, 201
 with differentiable function, 190
 with even function, 190
 existence, 58, 72, 81, 84
 from integrating c.d.f., 68
 from quantile function, 70
 kurtosis, 58, 91, 185, 299, 322, 333
 excess, 58
 mean, 57
 absolute, 67, 191
 and symmetry, 77
 of inverse, 65
 of ratio, 413
 multivariate, 172, 184
 of function
 expansion, 394
 of inverse and ratio, 85, 256
 of linear transformation, 73, 185
 of quadratic transformation, 185
 order of, 57
 precision, 58
 skewness, 58
 and kurtosis (Hamburger's problem), 94
 and symmetry, 77
 standard deviation, 58
 variance, 58
 absolute, 67
 generalized, 196, 635
 of linear combination, 183
 total, 196, 635
 Monotone likelihood ratio (MLR), 617, 657, 662, 666, 668, 673, 674
 from exponential, 658
 from F, 659
 from uniform, 658
 through sufficient statistic, 657
 Monty Hall problem, 25
 Multinomial coefficient, 9
 Neyman structure, 617, 618, 663–666, 669
 Neyman's factorization theorem, 441, 443, 458, 463, 487, 520, 522, 556, 617, 657
 Neyman–Pearson lemma, 613, 614, 616, 617, 622, 623, 657, 658, 662–664, 666, 668, 672–674, 693, 696
 generalized, 665–668
 Neyman–Scott problem, 508, 543, 549
 Nonparametric estimation, 445, 552–556, 574–583
 bandwidth, 553, 576, 580, 583
 kernel density estimator, 553, 574, 576
 kernel regression estimator, 582, 583
 Normal distribution, 55, 99, 178–183
 as limit of Student's t , 138
 as maximum entropy, 142
 as special case of elliptical, 178, 282
 c.d.f., 124, 420
 c.f., 282
 censored, 213
 conditional, 292–295, 301
 on square, 295
 correlation
 and independence, 282, 289, 290
 and quadrant probability, 301
 of squares, 296
 coverage probability, 446
 density, 178, 243, 282
 independence
 and correlation, 282, 289, 290
 equivalence from sample, 312
 joint vs. pairwise, 291
 of \bar{x} and S , 349, 351
 of \bar{x} and s^2 , 245, 312, 346
 integrates to 1, 259
 joint sufficiency, 457
 linear forms, 281, 288
 (in)dependence of quadratic form, 308, 309
 independence, 244, 288
 moments, 286
 m.g.f., 72, 82, 89, 286
 m.g.f.'s convexity, 287, 483
 method of moments, 551, 552
 ML estimator, 486–487, 505
 moments, 127
 fourth-order, 296
 multivariate, 178, 281–323
 definition, 281
 nonnormal sum of normals, 281
 plots, 116, 178

- Normal distribution (*Cont.*)
 product, 253, 303
 m.g.f., 306
 quadratic forms, 283, 303–323
 (in)dependence of linear form, 308, 309
 chi-squared, 307, 318, 320
 covariance, 314
 cumulant, 304
 independence, 316, 318, 320
 m.g.f., 304, 306
 ratio, 317
 sum of products, 306
 sample from, 345–361
 joint density, 348
 truncated, 213
- Notation, 727–739
 O , o , 343, 710–711, 730
 O_p , o_p , 374
 vector derivative, 736
- Odds ratio, 493
- Omitted-variables bias, 481, 539
- Order statistics, 231, 271, 273, 276, 329, 362, 452
 and ranks, 363
 asymptotics, 404, 429
- Orthogonal (random variables), 174
- Orthogonal parameterization, 489, 526, 542, 549, 562
- Outlier, 439
- Paradox of Yule and Simpson, 224, 567
- Parameter
 nuisance, 442, 606
 of interest, 442, 606
 space, 95
- Partial sum, 424, 704, 706
- Partition, 6, 231
- Pascal's triangle, 28
- Permutation, 8, 27–31
 generalized, 714–715
- Pitman's asymptotic relative efficiency (ARE), 677
 and Bahadur's, 696
- Pivot, 558, 589–591, 604, 613
- Poisson limit theorem, 434
- Polar coordinates, 258, 260, 262, 264
- Polar decomposition of a matrix, 700
- Polynomial, 702–704
 monic, 702
- P–P plot, 369
- Predictor
 best, 217
 best linear, 217
- Probability
 and indicator function, 61, 330
 conditional, 6, 19–27
 definition, 5
 density function (p.d.f.), 36
 joint, 151
 marginal, 151
 integral transform (PIT), 232, 238, 276, 277, 279, 434, 604, 692
 limit, 372
 space, 5
 complete, 169
 joint, 169
 unconditional, 7, 40
 with probability 1, 37
- Proof
 by contradiction, 700
 by contrapositive, 700
 by deduction, 700
 by induction, 700
- Q–Q plot, 369
- Quantile, 39, 43, 231
 function, 39, 601
 and moments, 70
 not additive, 236
 value at risk (VaR), 237
- R^2 , 337, 532, 535, 536, 538
 density, 337, 367
 irrelevant variable, 540, 572, 662
- Radon–Nikodym derivative, 38
- Random variable, 33
 degenerate, 58, 384, 391
- Random walk, 31, 147, 276, 424, 434, 516
- Rao–Blackwell theorem, 443
- Reflection and symmetry, 44, 77, 190, 221, 338
- Risk, 440
 Bayes, 557, 587, 588
 posterior, 557
- Sample, 327–369
 correlation, 328, 335
 density, 358, 368
 multiple, 336, 532, 538
 smaller than 1 in absolute value, 328
 covariance, 328
 from urn, 9, 22, 340, 341
 harmonic mean, 331, 448, 584
 mean, 230, 327, 328
 inadmissible, 476
 moments of, 331, 334

- Sample (*Cont.*)
- median, 328
 - midrange, 231
 - mode, 328
 - moment
 - central, 327
 - raw, 327
 - order statistics, *see* Order statistics
 - random, 97, 327
 - range, 231, 273
 - space, 4
 - variance, 231, 327, 328
 - asymptotics, 406, 408, 409
 - distribution of, 346
 - moments of, 332, 334
 - nonsingularity, 352
- Scatter plot, 445, 492, 531, 537, 556, 564
- Schur complement, 225, 282, 368, 532, 538, 714, 728
- Secretary problem, 26
- Series expansion, 704–712
- absolutely convergent, 705
 - asymptotic, 710
 - binomial
 - definition, 708
 - radius of convergence, 708
 - conditionally convergent, 705
 - exponential function
 - definition, 700
 - radius of convergence, 707
 - generalized hypergeometric, *see* Function, hypergeometric
 - geometric, *see* binomial
 - hyperbolic functions, 708–709
 - logarithmic function
 - explicit representation, 705, 708
 - implicit definition, 708
 - radius of convergence, 705
 - nonconvergent, 705
 - not unique, 102
 - radius of convergence, 706
 - summable, 706–708, 725
 - trigonometric functions, 708–709
- Set, 3–4, 9–19, 704
- compact, 704
 - complement, 3
 - dense, 38, 377, 386, 430, 705
 - independent, 5
 - intersection, 3
 - mutually exclusive, 5
 - mutually independent, 5
 - pairwise independent, 5
 - power, 4
 - union, 3
- Sigma-algebra, 4, 31
- Skorokhod's representation theorem, 433
- Slutsky's lemma, 393
- Slutsky's theorem, 433
- Smoothing parameter, 53, 553
- Spearman's ρ , 365, 369, 680
- Spectral radius, 728
- St. Petersburg paradox, 64
- Stable limit theorem (SLT), 375, 427
 - and portfolio diversification, 428
- Standard error, 438
 - vs. standard deviation, 438
- Statistic, 327
- ancillary, 442
 - complete, 443
 - bounded, 443, 444
 - for exponential family, 464
 - for uniform sample, 460
 - generalized likelihood ratio (GLR), 616, 625–637
 - from exponential, 626
 - from gamma, 625
 - from normal, 627–637
 - in linear model, 630, 631
 - normalized, 373
 - standardized, 99, 373
 - sufficient, 441
 - conditionally, 442
 - in normal linear model, 458
 - minimal, 442, 466, 467
 - normal distribution, 457
- Stieltjes integral, 38, 717–721
- Stirling's approximation, 419, 435, 449
- Stirling's series, 711
- Stochastic dominance, 69
- Support, 34
- t -ratio, 329
 - asymptotics, 406
- t -value, 329
- Taylor series, 711
- Taylor's theorem, 711, 713
- Test
- admissible, 610, 612, 695
 - consistent, 611, 621
 - containing Hessian vs. information, 649, 693
 - Cox, 654, 695
 - critical region, 606
 - most powerful region (MPR), 609
 - critical value, 607
 - Durbin–Wu–Hausman, 653, 694
 - goodness-of-fit, 618, 690

Test (*Cont.*)

- invariant, 611
- Lagrange multiplier (LM), 616
 - derivation, 641
 - in linear model, 637
 - LM_λ vs. LM_q , 643, 645, 652, 694
 - nonlinear restrictions, 643
 - one-sided vector restrictions, 651
- likelihood ratio (LR), 614, 616
 - for nonnested models, 654, 695
 - in linear model, 637
 - nonlinear restrictions, 643
 - one-sided vector restrictions, 651
- Neyman structure, 617, 618, 664–666, 669
 - and similarity, 663
- nonparametric, 618
 - equality of two distributions, 686–690
 - independence, 680–686
- p-value, 620
- power comparisons, 677, 696
- rank-sum (Wilcoxon, Mann and Whitney), 686, 688, 697
- score, 614, 616, 643, 645, 652, 694
- sign, 686
- similar, 612
 - and Neyman structure, 663
- statistic, 606
 - from binomial, 619
 - from normal, 606, 620
 - significance of, 606
- unbiased, 609, 621
- uniformly most powerful (UMP), 609
 - for composite hypotheses, 658–659
 - from multivariate normal not possible, 672
 - t -ratio, 664
 - through sufficient statistic, 657
- uniformly most powerful invariant (UMPI), 612
 - from multivariate normal, 672
 - sample correlation, 674
- uniformly most powerful unbiased (UMPU), 610
 - from exponential family, 665, 668
 - from normal, 662
 - sample correlation, 671
 - t -ratio, 669
- Vuong, 695
- Wald (W), 614, 615
 - in linear model, 637
 - nonlinear restrictions, 643, 646
 - not invariant, 649
 - one-sided vector restrictions, 651

Tightness, 390

Tower property, 225

Transformation theorem, 229

proof by conditioning, 267

Truncation, 50, 213, 397

 U -statistic, 369

Uniform asymptotic negligibility (u.a.n.), 415, 733

Unit circle, 702

Unit disk, 209, 258

Unit sphere, 262

Variance stabilization, 412