

Author Index

- Abu-Mostafa, Y. 384, 469
 Achanta, R. 384, 469
 Akaike, H. 240, 469
 Anderson, T. W. 72, 469
 Bastien, F. 386, 469
 Becker, R. 134, 469
 Bellhouse, D. R. 37, 469
 Bengio, Y. 384, 386, 469, 474
 Benjamini, Y. 304, 431, 465, 469
 Berger, J. O. 37, 271, 469
 Bergeron, A. 386, 469
 Bergstra, J. 386, 469
 Berk, R. 333, 432, 469
 Berkson, J. 134, 470
 Bernardo, J. M. 271, 470
 Bibby, J. M. 38, 72, 475
 Bien, J. 332, 477
 Birch, M. W. 134, 470
 Bishop, C. 383, 470
 Bloomston, M. 93, 473
 Boos, D. D. 188, 470
 Boser, B. 402, 470
 Bouchard, N. 386, 469
 Breiman, L. 135, 359, 466, 470
 Breuleux, O. 386, 469
 Brown, L. 333, 432, 469
 Bühlmann, P. 333, 477
 Buja, A. 333, 432, 469
 Carlin, B. P. 93, 271, 470
 Carroll, R. J. 459, 476
 Casella, G. 433, 475
 Chambers, J. 134, 469
 Chen, Z. 384, 475
 Chia, D. 384, 475
 Cho, C. S. 93, 473
 Chopin, N. 271, 473
 Cleveland, W. S. 11, 470
 Cohen, A. 405, 474
 Corbet, A. 472
 Cortes, C. 384, 474
 Courville, A. 384, 469
 Cover, T. M. 54, 470
 Cox, D. R. 54, 134, 159, 160, 272, 470
 Crowley, J. 160, 470
 Cudkowicz, M. 360, 474
 de Finetti, B. 271, 470
 Dembo, A. 54, 470
 Dempster, A. P. 159, 470
 Desjardins, G. 386, 469
 Dezeure, R. 333, 477
 Diaconis, P. 272, 470
 DiCiccio, T. 212, 471
 Donnelly, P. 271, 475
 Donoho, D. L. 462, 471
 Edwards, A. W. F. 38, 471
 Efron, B. 11, 21, 38, 53, 54, 72, 93, 94,
 110, 111, 136, 159, 161, 185–187,
 212, 214, 215, 240, 241, 272, 273,
 278, 304–307, 332, 333, 359, 430,
 432, 433, 458, 459, 471–473, 477
 Ertaylan, G. 360, 474
 Eskin, E. 405, 474
 Fang, L. 360, 474
 Feldman, D. 430, 472
 Fields, R. C. 93, 473
 Finney, D. J. 272, 472
 Fisher, R. A. 192, 212, 464, 472
 Fithian, W. 333, 472, 474
 Freund, Y. 359, 466, 472, 476
 Friedman, J. 134, 135, 240, 331, 332,
 359–361, 383, 470, 472, 473, 477
 Geisser, S. 240, 473
 Gerber, M. 271, 473
 Gholami, S. 93, 473
 Good, I. 92, 473

- Goodfellow, I. J. 386, 469
 Gous, A. 272, 273, 472
 Grosse-Wentrup, M. 360, 474
 Guyon, I. 402, 470
 Haenszel, W. 159, 475
 Hall, P. 212, 473
 Hampel, F. R. 187, 473
 Hardiman, O. 360, 474
 Harford, T. 241, 473
 Hastie, T. 134, 240, 331–333, 359–361,
 383, 384, 404, 405, 469, 472–474,
 476–478
 Hawe, J. 360, 474
 Hinkley, D. V. 54, 72, 470, 472
 Hinton, G. 384, 474, 476
 Hochberg, Y. 304, 431, 465, 469
 Hoeffding, W. 135, 306, 473
 Hoerl, A. E. 110, 473
 Hothorn, T. 360, 474
 Huber, P. J. 187, 473
 Jaeckel, L. A. 186, 474
 James, W. 109, 474
 Jansen, L. 333, 474
 Janson, L. 93, 473
 Javanmard, A. 333, 474
 Jaynes, E. 271, 474
 Jeffreys, H. 271, 474
 Jin, L. X. 93, 473
 Johnson, N. L. 37, 474
 Johnstone, I. 240, 332, 333, 472
 Kaplan, E. L. 159, 474
 Kass, R. E. 271–273, 474
 Kennard, R. W. 110, 473
 Kent, J. T. 38, 72, 475
 Koh, P. W. 384, 475
 Kotz, S. 37, 474
 Krizhevsky, A. 384, 476
 Kuffner, R. 360, 474
 Laird, N. M. 159, 470
 Lamblin, P. 386, 469
 Le, Q. V. 384, 475
 LeCun, Y. 384, 474
 Lee, J. 333, 474
 Lehmann, E. L. 54, 474
 Leitner, M. L. 360, 474
 Leslie, C. 405, 474
 Levine, E. A. 93, 473
 Li, G. 360, 474
 Liang, P. S. 384, 385, 477
 Liaw, A. 359, 474
 Liberman, M. 462, 474
 Lin, Y. 403, 477
 Loader, C. 405, 473
 Lockhart, R. 333, 475
 Loftus, J. 333, 477
 Louis, T. A. 93, 271, 470
 Lynden-Bell, D. 157, 475
 MacDonald, J. M. 304, 476
 Macke, J. H. 360, 474
 Mackey, L. 360, 474
 Maitthel, S. K. 93, 473
 Mallows, C. L. 240, 475
 Mantel, N. 159, 475
 Mardia, K. V. 38, 72, 475
 McCullagh, P. 134, 332, 475
 Meier, P. 159, 474
 Metropolis, N. 271, 475
 Meyer, T. 360, 474
 Miller, R. G., Jr 185, 304, 431, 475
 Montanari, A. 333, 474
 Morris, C. 110, 472
 Nelder, J. 134, 332, 475
 Nesterov, Y. 384, 475
 Neyman, J. 21, 212, 464, 475
 Ng, A. 384, 475
 Ngiam, J. 384, 475
 Noble, W. S. 405, 474
 Norel, R. 360, 474
 Norton, J. A. 93, 473
 O'Hagan, A. 271, 475
 Olshen, R. A. 135, 359, 470
 Park, T. 433, 475
 Pascanu, R. 386, 469
 Pawlik, T. M. 93, 473
 Pearson, E. S. 464, 475
 Pearson, K. 464, 475
 Peers, H. W. 38, 215, 271, 478
 Pericchi, L. R. 271, 469
 Petrosian, V. 136, 472
 Popescu, B. 359, 473
 Poultides, G. A. 93, 473
 Pritchard, J. 271, 475
 Quenouille, M. H. 185, 475
 R Core Team 134, 475
 Raftery, A. E. 272, 474
 Reid, N. 272, 470
 Reid, S. 333, 477
 Ridgeway, G. 304, 359, 476
 Ripley, B. D. 383, 476
 Ritov, Y. 333, 477

- Robbins, H. 92, 109, 433, 476
 Ronchetti, E. M. 187, 473
 Rosenbluth, A. W. 271, 475
 Rosenbluth, M. N. 271, 475
 Rosset, S. 404, 476
 Rousseuw, P. J. 187, 473
 Rubin, D. B. 159, 187, 470, 476
 Salakhutdinov, R. 384, 476
 Savage, L. J. 271, 476
 Schapire, R. 359, 466, 472, 476
 Scheffé, H. 430, 476
 Schmidt, C. 93, 473
 Schoenfeld, D. 360, 474
 Schölkopf, B. 402, 476
 Schwarz, G. 273, 476
 Senn, S. 38, 476
 Serfling, R. J. 188, 470
 Sherman, A. 360, 474
 Siegmund, D. 304, 477
 Simon, N. 332, 477
 Singh, K. 53, 215, 478
 Smola, A. 402, 476
 Soric, B. 304, 476
 Spevack, M. 93, 476
 Spolverato, G. 93, 473
 Squires, I., Malcolm 93, 473
 Srivastava, N. 384, 476
 Stahel, W. A. 187, 473
 Stefanski, L. 459, 476
 Stein, C. 109, 111, 186, 241, 271, 472, 474, 476
 Stephens, M. 271, 475
 Stigler, S. M. 462, 476
 Stolovitzky, G. 360, 474
 Stone, C. J. 135, 359, 470
 Stone, M. 240, 477
 Storey, J. D. 304, 477
 Sun, D. 333, 472, 474
 Sun, Y. 333, 474
 Sutskever, I. 384, 476
 Tanner, M. A. 273, 477
 Taylor, J. 304, 332, 333, 472, 474, 475, 477
 Teller, A. H. 271, 475
 Teller, E. 271, 475
 Thisted, R. 93, 472, 477
 Thomas, J. A. 54, 470
 Tibshirani, R. 134, 187, 215, 240, 241, 271, 331–333, 359–361, 383, 404, 433, 465, 472, 473, 475, 477, 478
 Toulmin, G. 92, 473
 Tran, L. 360, 474
 Tran, T. B. 93, 473
 Tukey, J. W. 11, 185, 187, 465, 477
 Turian, J. 386, 469
 van de Geer, S. 333, 477
 van Ligteneberg, J. 360, 474
 Vapnik, V. 402, 470, 477
 Vaughan, R. 360, 474
 Vincent, P. 384, 469
 Votopoulos, K. I. 93, 473
 Wager, S. 359, 384, 385, 477
 Wahba, G. 403, 404, 477
 Wainwright, M. 331–333, 473
 Wald, A. 465, 477
 Wang, L. 360, 474
 Wang, S. 384, 385, 477
 Warde-Farley, D. 386, 469
 Wasserman, L. 271, 273, 474
 Weber, S. M. 93, 473
 Wedderburn, R. W. M. 134, 478
 Welch, B. L. 38, 215, 271, 478
 Westfall, P. 304, 431, 478
 Weston, J. 405, 474
 Wiener, M. 359, 474
 Wilks, A. 134, 469
 Williams, C. 472
 Wong, W. H. 273, 477
 Worhunsky, D. J. 93, 473
 Xie, M. 53, 215, 478
 Ye, J. 240, 478
 Yekutieli, D. 431, 469
 Ylvisaker, D. 272, 470
 Young, S. 304, 431, 478
 Zach, N. 360, 474
 Zhang, C.-H. 333, 478
 Zhang, H. 403, 477
 Zhang, K. 333, 432, 469
 Zhang, S. 333, 478
 Zhao, L. 333, 432, 469
 Zhu, J. 404, 476
 Zou, H. 240, 332, 478

Cambridge University Press
978-1-108-82341-8 — Computer Age Statistical Inference, Student Edition
Bradley Efron , Trevor Hastie
Index
[More Information](#)

Subject Index

- abc method, 202, 212
- Accelerated gradient descent, 371
- Acceleration, 200, 214
- Accuracy, 15
 - after model selection, 415–421
- Accurate but not correct, 415
- Activation function, 367, 373
 - leaky rectified linear, 374
 - rectified linear, 374
 - ReLU, 374
 - tanh, 374
- Active set, 311, 318
- adaboost** algorithm, 352–356, 462
- Adaboost.M1, 353
- Adaptation, 417
- Adaptive estimator, 417
- Adaptive rate control, 371
- Additive model, 335
 - adaptive, 357
- Adjusted compliance, 417
- Admixture modeling, 266–270
- AIC, *see* Akaike information criterion
- Akaike information criterion, 217, 227, 235, 240, 256, 278
- Allele frequency, 267
- American Statistical Association, 464
- Ancillary, 46, 48, 146
- Apparent error, 220, 222, 228
- arcsin transformation, 100
- Arthur Eddington, 462
- Asymptotics, xvi, 125, 126
- Autoencoder, 374–376
- Backfitting, 357
- Backpropagation, 368–370
- Bagged estimate, 417, 419
- Bagging, 235, 338, 419, 421, 432
- Balance equations, 266
- Barycentric plot, 269
- Basis expansion, 387
- Bayes
 - deconvolution, 435–438
 - factor, 254, 295
 - false-discovery rate, 289
 - posterior distribution, 264
 - posterior probability, 290
 - shrinkage, 221
 - t -statistic, 265
 - theorem, 23
- Bayes–frequentist estimation, 425–430
- Bayesian
 - inference, 23–38
 - information criterion, 256
 - lasso, 433
 - lasso prior, 428
 - model selection, 254
 - trees, 360
- Bayesian information criterion, 278
- Bayesianism, 3
- BCa
 - accuracy and correctness, 213
 - confidence density, 210, 215, 247, 252, 253
 - interval, 210
 - method, 200
- Benjamini and Hochberg, 286
- Benjamini–Yekutieli, 413
- Bernoulli, 349
- Best-approximating linear subspace, 375
- Best-subset selection, 309
- Beta
 - distribution, 57, 249
- BH $_q$, 286
- Bias, 15, 364
- Bias-corrected, 341
 - and accelerated, *see* BCa method
 - confidence intervals, 198–199
 - percentile method, 198

- Bias-correction value, 199
- Biased estimation, 331
- BIC, *see* Bayesian information criterion
- Big-data era, xv, 461
- Binomial, 115, 123
 - distribution, 57, 123, 249
 - log-likelihood, 392
 - standard deviation, 117
- Bioassay, 115
- Biometrika*, 464
- Bivariate normal, 190
- Bonferroni bound, 283
- Boole's inequality, 284
- Boosting, 330, 335, 344–361
- Bootstrap, 7, 163–188, 277, 338
 - Baron Munchausen, 185
 - Bayesian, 176, 187
 - cdf, 195
 - confidence intervals, 189–215
 - ideal estimate, 169, 187
 - jackknife after, 187
 - moving blocks, 176
 - multisample, 175
 - nonparametric, 167–171, 226
 - out of bootstrap, 241
 - packages, 186
 - parametric, 177–181, 232, 322, 443
 - probabilities, 172
 - replication, 167
 - sample, 167
 - sample size, 187, 213
 - smoothing, 235, 417, 419
 - t , 204
 - t intervals, 203–206
- Bound form, 315
- Bounding hyperplane, 411
- Burn-in, 270
- BY_q algorithm, 413
- Causal inference, xvi
- Censored
 - data, 141–146
 - not truncated, 157
- Centering, 112
- Central limit theorem, 125
- Chain rule for differentiation, 368
- Classic statistical inference, 3–77
- Classification, 130, 218
- Classification accuracy, 387
- Classification error, 218
- Classification tree, 359
- Cochran–Mantel–Haenszel test, 138
- Coherent behavior, 271
- Common task framework, 462
- Compliance, 407
- Computational bottleneck, 134
- Computer age, xv
- Computer-intensive, 133
 - inference, 197, 278
 - statistics, 167
- Conditional, 61
- Conditional distribution
 - full, 263
- Conditional inference, 47–50, 146, 149
 - lasso, 328
- Conditionality, 46
- Confidence
 - density, 208, 209, 245
 - distribution, 206–211
 - interval, 18
 - region, 410
- Conjugate, 263, 269
 - prior, 248
 - priors, 247
- Convex optimization, 314, 318, 331, 333, 389
- Convolution, 436, 459
 - filters, 380
 - layer, 379
- Corrected differences, 424
- Correlation effects, 305
- Covariance
 - formula, 322
 - penalty, 227–235
- Coverage, 189
- Coverage level, 284
- Coverage matching prior, 246–247
- Cox model, *see* proportional hazards model
- C_p , 226, 227, 230, 240, 278, 310, 407, 408, 416
- Cramér–Rao lower bound, 46
- Credible interval, 206, 430
- Cross-validation, 217–241, 278, 346
 - 10-fold, 337
 - estimate, 223
 - K -fold, 310
 - leave one out, 223, 240
- Cumulant generating function, 70
- Curse of dimensionality, 399
- Dark energy, 219, 240
- Data analysis, 465
- Data science, xvii, 465, 466

- Data sets
ALS, 345
AML, *see* leukemia
baseball, 99
butterfly, 82
cell infusion, 118
cholesterol, 408, 415, 416
CIFAR-100, 377
diabetes, 103, 218, 409, 427, 429
dose-response, 115
galaxy, 126
handwritten digits (MNIST), 365
head/neck cancer, 142
human ancestry, 267
insurance, 138
kidney function, 165, 231
leukemia, 184, 204, 389
NCOG, 141
nodes, 438, 441, 444, 452, 453, 456
pediatric cancer, 150
police, 297
prostate, 259, 282, 299, 421, 423, 437, 448–450
protein classification, 397
shakespeare, 85
spam, 119, 133, 218, 224, 310–312, 336
student score, 181, 189, 194, 210, 211
supernova, 219, 221, 226, 230, 233
vasoconstriction, 250, 251, 256, 262
- Data snooping, 411
 De Finetti, B., 36, 37, 261, 465
 De Finetti–Savage school, 261
 Debias, 328
 Decision rule, 285
 Decision theory, xvi
 Deconvolution, 436
 Deep learning, 363–386
 Definitional bias, 445
 Degrees of freedom, 230, 240, 322–323
 Delta method, 16, 427, 433
 Deviance, 118, 124, 125, 311
 Deviance residual, 129
 Diffusion tensor imaging, 301
 Direct evidence, 110, 115, 435
 Directional derivatives, 166
 Distribution
 beta, 57, 249
 binomial, 57, 123, 249
 gamma, 57, 123, 249
 Gaussian, 57
 normal, 57, 123, 249
 Poisson, 57, 123, 249
 Divide-and-conquer algorithm, 336
 Document retrieval, 308
 Dose–response, 115
 Dropout learning, 380, 384
 DTI, *see* diffusion tensor imaging
 Early computer-age, xvi, 79–279
 Early stopping, 374
 Effect size, 282, 298, 412, 421
 Efficiency, 46, 126
 Eigenratio, 170, 181, 202
 Elastic net, 326, 368
 Ellipsoid, 411
 EM algorithm, 153–157
 missing data, 277
 Empirical Bayes, 79–94, 98, 275
 estimation strategies, 435–459
 information, 457
 large-scale testing, 288–292
 Empirical null, 296
 estimation, 299–300
 maximum-likelihood estimation, 306
 Empirical probability distribution, 168
 Ensemble, 335, 345
 Ephemeral predictors, 236
 Epoch, 371
 Equilibrium distribution, 266
 Equivariant, 111
 Exact inferences, 125
 Expectation parameter, 124
 Experimental design, xvi
 Exponential family, 56–75, 234
 p -parameter, 123, 426, 438
 curved, 72
 one-parameter, 122
 F distribution, 410
 F tests, 407
 f -modeling, 438, 448, 454–458
 Fake-data principle, 155, 161, 277
 False coverage
 control, 412
 False discovery, 285
 control, 412
 control theorem, 304
 proportion, 285
 rate, 281–307
 False-discovery

- rate, 9
- Family of probability densities, 67
- Family-wise error rate, 284
- FDR, *see* false-discovery rate
- Feed-forward, 363
- Fiducial, 278
 - constructions, 207
 - density, 208
 - inference, 53
- Fisher, 83
- Fisher information, 30, 43, 62
 - bound, 43
 - matrix, 246, 441
- Fisherian correctness, 213
- Fisherian inference, 40–54, 245
- Fixed-knot regression splines, 356
- Flat prior, 245
- Forward pass, 369
- Forward-stagewise, 357
 - fitting, 330
- Forward-stepwise, 308–313
 - computations, 332
 - logistic regression, 332
 - regression, 310
- Fourier
 - method, 454
 - transform, 454
- Fraillities, 453
- Frequentism, 3, 13–23, 31, 36, 53, 153, 278
- Frequentist, 426
 - inference, 13–22
 - strongly, 227
- Fully connected layer, 380
- Functional gradient descent, 351
- FWER, *see* family-wise error rate
- g*-modeling, 437
- Gamma, 123
 - distribution, 57, 123, 249
- General estimating equations, xvi
- General information criterion, 258
- Generalized
 - linear mixed model, 451–454
 - linear model, 114–129, 277
 - ridge problem, 396
- Genome, 267
- Genome-wide association studies, 466
- Gibbs sampling, 261–270, 278, 427
- GLM, *see* generalized linear model
- GLMM, *see* generalized linear mixed model
- Google flu trends, 239, 241
- Gradient boosting, 349–352
- Gradient descent, 366, 368
- Gram matrix, 393
- Gram-Schmidt orthogonalization, 332
- Graphical lasso, 331
- Graphical models, xvi
- Greenwood's formula, 144, 158
- Group lasso, 331
- Hadamard product, 370
- Handwritten digits, 365
- Haplotype estimation, 271
- Hazard rate, 138–141
 - parametric estimate, 145
- Hidden layer, 363, 364, 366
- High-order interaction, 336
- Hinge loss, 392
- Hints
 - learning with, 381
- Hoeffding's lemma, 124
- Holm's procedure, 284, 304
- Homotopy path, 316
- Hypergeometric distribution, 148, 159
- Imputation, 156
- Inadmissible, 98
- Indirect evidence, 107, 115, 277, 300, 435, 454, 457
- Inductive inference, 126
- Inference, 3
- Inference after model selection, 407–433
- Inferential triangle, 461
- Infinitesimal forward stagewise, 330
- Infinitesimal jackknife, 175
 - estimate, 419
 - standard deviations, 420
- Influence function, 182–185
 - empirical, 183
- Influenza outbreaks, 239
- Input distortion, 381, 385
- Input layer, 367
- Insample error, 228
- Inverse chi-squared, 272
- Inverse gamma, 249, 272
- IRLS, *see* iteratively reweighted least squares
- Iteratively reweighted least squares, 311, 332
- Jackknife, 163–188, 277, 341
 - estimate of standard error, 164
 - standard error, 186

- James–Stein
 - estimation, 96–112, 292, 315, 423
 - ridge regression, 276
- Jeffreys
 - prior, 247
- Jeffreys’
 - prior, 29–31, 37, 206, 211, 246
 - prior, multiparameter, 252
 - scale, 295
- Jumpiness of estimator, 418
- Kaplan–Meier, 138, 141, 143, 144
 - estimate, 141–146, 277
- Karush–Kuhn–Tucker optimality
 - conditions, 318
- Kernel
 - function, 394
 - logistic regression, 398
 - method, 387–405
 - smoothing, 387, 399–402
 - SVM, 398
 - trick, 387, 393–395, 404
- Knots, 319
- Kullback–Leibler distance, 118
- ℓ_1 regularization, 331
- Lagrange
 - dual, 393
 - form, 315, 318
 - multiplier, 403
 - primal, 403
- Large-scale
 - hypothesis testing, 281–307
 - testing, 282–285
- Large-scale prediction algorithms, 461
- Lasso, 106, 219, 226, 231, 240, 308–333
 - modification, 322
 - path, 322
 - penalty, 368
- Learning from the experience of others,
 - 109, 290, 300, 435, 457
- Learning rate, 370
- Least squares, 103, 118, 309
- Least-angle regression, 319–323, 331
- Least-favorable family, 272
- Left-truncated, 157
- Lehmann alternative, 304
- Life table, 138–141
- Likelihood function, 40
 - concavity, 124
- Limited-translation rule, 303
- Lindsey’s method, 71, 179
- Linearly separable, 387
- Link function, 247, 351
- Local false-discovery rate, 290, 292–296
- Local regression, 399–402, 405
- Local translation invariance, 380
- Log polynomial regression, 423
- Log-rank statistic, 159
- Log-rank test, 138, 146–149, 159, 277
- Logic of inductive inference, 193, 213
- Logistic regression, 115–121, 146, 223,
 - 309, 387
 - multiclass, 367
- Logit, 115
- Loss plus penalty, 397
- Machine learning, 217, 278, 387
- Mallows’ C_p , *see* C_p
- Mantel–Haenzel test, 138
- MAP, 106
- MAP estimate, 433
- Margin, 388
- Marginal density, 422, 436
- Markov chain Monte Carlo, *see* MCMC
- Markov chain theory, 266
- Martingale theory, 304
- Matching prior, 206, 208
- Matlab, 281
- Matrix completion, 331
- Max pool layer, 378
- Maximized a-posteriori probability, *see* MAP
- Maximum likelihood, 309
- Maximum likelihood estimation, 40–54
- MCMC, 244, 261–270, 278, 427
- McNemar test, 352
- Mean absolute deviation, 462
- Median unbiased, 198
- Memory-based methods, 402
- Meter reader, 31
- Meter-reader, 38
- Microarrays, 236, 281
- Minitab, 281
- Misclassification error, 312
- Missing data, 153–157, 336
 - EM algorithm, 277
- Missing-species problem, 82–88
- Mixed features, 336
- Mixture density, 289
- Model averaging, 421
- Model selection, 253–260, 411
 - criteria, 260
- Monotone lasso, 330
- Monotonic increasing function, 192

- Multinomial
 - distribution, 64–67, 439
 - from Poisson, 66
- Multiple testing, 282
- Multivariate
 - analysis, 125
 - normal, 58–62
- n*-gram, 397
- N-P complete, 309
- Nadaraya–Watson estimator, 400
- Natural parameter, 122
- Natural spline model, 444
- NCOG, *see* Northern California Oncology Group
- Nested models, 309
- Neural Information Processing Systems, 384
- Neural network, 363–386
 - adaptive tuning, 372
 - number of hidden layers, 373
- Neurons, 363
- Neyman’s construction, 189, 191, 201, 212
- Neyman–Pearson, 19, 20, 303
- Non-null, 282
- Noncentral chi-square variable, 215
- Nonlinear transformations, 387
- Nonlinearity, 373
- Nonparameteric
 - regression, 387
- Nonparametric, 56, 133
 - MLE, 157, 168
 - percentile interval, 195
- Normal
 - correlation coefficient, 190
 - distribution, 57, 123, 249
 - multivariate, 58–62
 - regression model, 427
 - theory, 125
- Northern California Oncology Group, 141
- Nuclear norm, 331
- Nuisance parameters, 149, 207
- Objective Bayes, 37, 278
 - inference, 243–273
 - intervals, 206–211
 - prior distribution, 244–247
- OCR, *see* optical character recognition
- Offset, 360
- OLS
 - algorithm, 416
 - estimation, 408
 - predictor, 230
- One-sample nonparametric bootstrap, 169
- One-sample problems, 164
- OOB, *see* out-of-bag error
- Optical character recognition, 365
- Optimal separating hyperplane, 387–389
- Optimal-margin classifier, 388
- Optimality, 19
- Oracle, 285
- Orthogonal parameters, 272
- Out-of-bag error, 241, 338, 340–341
- Out-the-box learning algorithm, 335
- Output layer, 364
- Outsample error, 228
- Over parametrized, 308
- Overfitting, 314
- Overshrinks, 102
- p*-value, 9, 292
- Package/program
 - gbm**, 346, 359
 - glmnet**, 223, 325, 332, 359
 - h2o**, 384
 - lars**, 322, 330
 - liblinearR**, 393
 - locfdr**, 299–301, 306, 451
 - lowess**, 6, 231, 400
 - nlm**, 442
 - randomForest**, 338, 359
 - selectiveInference**, 333
- Pairwise inner products, 393
- Parameter space, 23, 30, 57, 65, 69
- Parametric bootstrap, 252
- Parametric family, 177
- Parametric models, 56–75
- Partial likelihood, 149, 152, 158, 160, 277, 352
- Partial logistic regression, 159
- Partial residual, 357
- Path-wise coordinate descent, 324
- Penalized
 - least squares, 106
 - likelihood, 106, 442
 - logistic regression, 368
 - maximum likelihood, 235, 317
- Percentile method, 193–198
 - central interval, 195
- Permutation null, 299, 306
- Permutation test, 51–53
- Phylogenetic tree, 271

- Piecewise
 - linear, 323
 - nonlinear, 324
- Pivotal
 - argument, 191
 - quantity, 204, 206
 - statistic, 17
- .632 rule, 241
- Poisson, 123, 201
 - distribution, 57, 123, 249
 - regression, 126–129, 259, 294, 305, 449
- Poisson regression, 179
- Polynomial kernel, 394, 404
- Positive-definite function, 394
- Post-selection inference, 327, 407–433
- Posterior density, 245, 248
- Posterior distribution, 429
- Postwar era, 275
- Prediction
 - errors, 225
 - rule, 217–222
- Predictors, 130, 217
- Principal components, 374
- Prior distribution, 244–253
 - beta, 249
 - conjugate, 247–253
 - coverage matching, 246–247
 - gamma, 249
 - normal, 249
 - objective Bayes, 244
 - proper, 249
- Probit analysis, 118, 126, 134
- Propagation of errors, 433
- Proper prior, 249
- Proportional hazards model, 138, 149–153, 277
- Proximal-Newton, 325
- q -value, 290
- QQ plot, 297
- QR decomposition, 321, 332
- Quadratic program, 389
- Quasilikelihood, 277
- Quetelet, Adolphe, 464
- R**, 186, 281
- Random forest, 218, 238, 335–343, 358–361
 - adaptive nearest-neighbor estimator, 339
 - leave-one-out cross-validated error, 340
 - Monte Carlo variance, 341
 - sampling variance, 341
 - standard error, 341–342
- Randomization, 51–53
- Rao–Blackwell, 236, 240
- Rate annealing, 372
- Rectified linear, 371
- Regression, 115
- Regression rule, 228
- Regression to the mean, 34
- Regression tree, 130–134, 277, 359
- Regularization, 106, 181, 308, 391, 442
 - path, 316
- Relevance, 300–303
- Relevance function, 303
- Relevance theory, 307
- Reproducing kernel Hilbert space, 387, 396, 404
- Resampling, 171
 - plans, 171–177
 - simplex, 172, 177
 - vector, 171
- Residual deviance, 293
- Response, 130, 217
- Ridge regression, 102–107, 218, 314, 338, 343, 384, 393
 - James–Stein, 276
- Ridge regularization, 380
 - logistic regression, 404
- Right-censored, 157
- Risk set, 151
- RKHS, *see* reproducing-kernel Hilbert space
- Robbins’ formula, 79, 81, 436, 454
- Robust estimation, 182–185
- Royal Statistical Society, 464
- S language, 281
- Sample correlation coefficient, 190
- Sample size coherency, 258
- Sampling distribution, 322
- SAS, 281
- Savage, L. J., 36, 37, 53, 207, 243, 261, 465
- Scale of evidence
 - Fisher, 255
 - Jeffreys, 255
- Scheffé
 - interval, 409, 410, 430
 - theorem, 411
- Score function, 44
- Score tests, 311

- Second-order accuracy, 200–203
- Selection bias, 34, 421–424
- Self-consistent, 156
- Separating hyperplane, 387
 - geometry, 402
- Seven-league boots, 463
- Shrinkage, 121, 326, 349
 - estimator, 62, 96, 99, 101, 423
- Sigmoid function, 364
- Significance level, 284
- Simulation, 163–215
- Simultaneous confidence intervals, 408–412
- Simultaneous inference, 304, 431
- Sinc kernel, 454, 459
- Single-nucleotide polymorphism, *see* SNP
- Smoothing operator, 357
- SNP, 267
- Soft-margin classifier, 390–391
- Soft-threshold, 325
- Softmax, 367
- Spam filter, 121
- Sparse
 - models, 308–333
 - principal components, 331
- Sparse matrix, 326
- Sparsity, 331
- Split-variable randomization, 338, 343
- SPSS, 281
- Squared error, 218
- Standard candles, 219, 240
- Standard error, 163
 - external, 421
 - internal, 421
- Standard interval, 189
- Stein's
 - paradox, 110
 - unbiased risk estimate, 227, 240
- Stepwise selection, 309
- Stochastic gradient descent, 370
- Stopping rule, 33, 426
- Stopping rules, 253
- String kernel, 397, 398
- Strong rules, 326, 332
- Structure, 271
- Structure matrix, 102, 438
- Student t
 - confidence interval, 409
 - distribution, 204, 282
 - statistic, 464
 - two-sample, 8, 282
- Studentized range, 431
- Subgradient
 - condition, 318
 - equation, 322, 325
- Subjective prior distribution, 243
- Subjective probability, 243
- Subjectivism, 36, 243, 253, 271
- Sufficiency, 46
- Sufficient
 - statistic, 69, 118, 122
 - vector, 69
- Supervised learning, 364
- Support
 - set, 389, 390
 - vector, 389
 - vector classifiers, 393
 - vector machine, 329, 387–405
- SURE, *see* Stein's unbiased risk estimate
- Survival analysis, 138–161, 277
- Survival curve, 144, 289
- SVM
 - Lagrange dual, 403
 - Lagrange primal, 403
 - loss function, 403
- Taylor series, 165, 433
- Theoretical null, 296
- Tied weights, 380
- Time series, xvi
- Training set, 217
- Transformation invariance, 191–193, 246
- Transient episodes, 237
- Trees
 - averaging, 359
 - best-first, 344
 - depth, 346
 - terminal node, 132
- Tricube kernel, 400, 401
- Trimmed mean, 183
- Triple-point, xv
- True error rate, 219
- True-discovery rates, 296
- Tukey, J. W., 431, 465
- Tukey, J. W., 431
- Tweedie's formula, 422, 432, 454
- Twenty-first-century methods, xvi, 281–461
- Two-groups model, 288
- Uncorrected differences, 424
- Uninformative prior, 29, 177, 243, 271
- Universal approximator, 363

Subject Index

491

- Unlabeled images, 377
- Unobserved covariates, 298
- Validation set, 222
- Variable-importance plot, 342–343, 347
- Variance, 15
- Variance reduction, 335
- Velocity vector, 372
- Voting, 344
- Warm starts, 324, 375
- Weak learner, 344, 353
- Weight
 - decay, 368
 - regularization, 373, 374
 - sharing, 364, 379
- Weighted exponential loss, 356
- Weighted least squares, 325
- Weighted majority vote, 352
- Weights, 364
- Wide data, 308, 331
- Wilks' likelihood ratio statistic, 256
- Winner's curse, 34, 421
- Winsorized mean, 183
- Working response, 325, 332
- $z^{(\alpha)}$, 196
- Zero set, 306