

Note: Page numbers followed by *f* indicate figures, and *t* indicate tables.

## A

Acceptance-rejection method for random variates, 135–136  
 Accuracy, 166–168, 235, 241, 294–295, 298, 300–301, 313, 324, 325*f*, 372–373, 376–378, 732  
 AC. *See* Anomaly correlation (AC)  
 Activation function, 291–292  
 Additive law of probability, 12  
 Affine kernel dressing, 343  
 Agglomerative methods, 722–725, 732–734  
 Akaike Information Criterion, 277, 498–499  
 Aliasing, 534–537  
 Alternative hypothesis, 144, 146–151, 153, 165, 168–170, 499  
 Amplitude, 452, 462, 519–525, 527–531, 533–536, 540–546, 572, 626–627, 638, 662–663  
   definition of, 519  
   element of PCA, 626–627  
 Amplitude error, 452, 462–463  
 Analysis equation, 525, 630, 672  
 Analysis of deviance, 276  
 Analysis formula, 619–620, 623–628, 639–640  
 Analysis of variance (ANOVA), 199–206  
   one-way, 199–201  
   two-way, 202–205  
   in regression, 178, 180  
 Angular frequency, 519, 530, 535  
 Annual maximum data, 113, 117–118  
 Anomaly, 50, 96–97, 232, 486, 504, 597, 660, 692  
 Anomaly correlation (AC), 452–455, 722  
   centered, 452–455  
   uncentered, 452–455  
 Anomaly matrix, 577, 659–660  
 Anscombe's quartet, 54, 54*f*  
 AR(1) model, 500, 501*f*, 502–507, 509–510, 513–517, 537–538, 538*f*, 544, 598, 600*f*, 608  
 AR(2) model, 504–510, 507*f*, 514, 516–517, 539–540, 539–540*f*, 663  
 Area under ROC curve, 413, 473–475

ARMA (autoregressive moving average) model, 513, 515–517, 538  
 Artificial neural network, 290–293  
 Artificial skill, 265, 526–528  
 Asymptotic equitability, 383  
 Attractor, 315–318, 315*f*, 317*f*  
 Attributes diagram, 408–410, 409*f*  
 Autocorrelation, 61–62, 156, 158–160, 180–181, 190, 397–398, 470, 491, 493, 503, 511–513, 598, 643, 649–650  
 Autocorrelation function, 62–63, 158, 487, 494*f*, 502, 506–508, 507*f*, 511, 513, 517, 539*f*  
 Autocovariance function, 63, 485, 661–662, 664  
 Autoregression, 157–158, 500–504, 598–599, 608  
 Autoregressive model, 500–501, 504, 506, 508*f*, 509, 514, 516, 537–540, 599  
   forecasting with, 516  
   order selection among, 509  
   spectra of, 537–540  
   statistical simulation with, 514  
 Autoregressive-moving average model, 513, 598  
 Average-linkage clustering, 723  
 Average-rank histogram, 442–443  
 Axioms of probability, 9

## B

Backward elimination, 262–263, 691  
 Bagging, 290, 718  
 Band-depth rank histogram (BDH), 442  
 Base rate, 376, 379, 383–385, 410  
 Bayesian Information Criterion (BIC), 277, 498–499, 509–510  
 Bayesian interpretation of probability, 9–11, 17–18, 125, 235, 305, 709  
 Bayesian model averaging (BMA), 338–343  
 Bayes' Theorem, 17–18, 209–213, 215–218, 225, 344, 709  
 BC<sub>a</sub> confidence interval, 189  
 Bernoulli distribution, 81, 275, 405, 490–491, 498–499

Bernoulli trial, 81, 83–84, 117, 209, 211–212, 217, 220–221  
 Beta-binomial distribution, 218–220  
 Beta distribution, 43, 111–113, 211–212, 217–219, 342, 593  
   mode, 217–218  
 Beta-Pascal distribution, 220–221  
 Between-groups covariance matrix, 704–706, 726  
 BHF representation, 383–384  
 Bias, 65–66, 189–190, 192, 268–270, 298–300, 299*f*, 308, 327–328, 330–331, 338–339, 348, 350, 352, 354, 372, 378, 387–389, 396–398, 403–404, 406–407, 434–435, 543, 589, 646, 660, 733  
 Bias ratio, 378, 382–384, 385*f*, 387  
 Binary variable, 129, 254–255, 259, 425, 514  
 Binned spread-error diagram, 438–439, 438–439*f*  
 Binomial distribution, 79–87, 89, 148–149, 197, 209–211, 220, 471, 488, 490–491  
 Bi-normal model, 416  
 Biplot, 666–667  
 Bivariate autoregression, 599  
 Bivariate normal distribution, 97–100, 252, 352, 587–588, 595  
 Biweight kernel, 36*r*  
 Block maximum data, 113–118  
 Blue noise, 538, 538*f*  
 Bonferroni method, 201, 473–475, 477, 544, 609–611  
 Bootstrap  
   approximations to sampling distributions, 650  
   circular block, 192  
   confidence intervals, 148–151  
   moving-blocks, 191–192, 191*f*, 479–480, 650  
   nearest-neighbor, 191, 282–283, 713  
   for verification statistic, 479–480  
 Box-Cox transformation, 44–45, 47–50  
 Box-Jenkins model, 500, 513  
 Box-Muller method for Gaussian variates, 136–137

- Boxplot. *See* Box-and-whisker plot  
 Box-and-whisker plot, 30–31, 363  
 Breeding method, 322  
 Brier score (BS), 401–410, 416–419, 467, 475–476  
 Brier skill score (BSS), 402–403, 408–410, 421–422, 458, 467, 469, 475–476  
 Broken stick model, 642  
 Brushing, 71–72  
 Buell patterns, 638  
 BUGS (Bayesian inference using Gibbs sampling), 232–233  
 Burn-in, 228, 232
- C**  
 Calibration, 372, 435  
 Calibration function, 405–408, 406*f*, 410, 476  
 Calibration-refinement factorization, 370–373, 375–376, 383–386, 398–399, 400*t*, 402, 404–405, 411, 422  
 Calibration simplex, 422–424  
 Canandaigua temperature data, 100, 156, 244–245, 285, 595, 599, 620, 673, 674*t*  
 Canonical correlation analysis (CCA), 670–687  
   applied to fields, 676–678  
   canonical variates, canonical vectors, and canonical correlations, 669–670  
   computational considerations, 683–687  
   operational forecast system, 680  
   properties of, 670–676  
 Canonical pattern, 677  
 Canonical variate, 670–676, 679–682  
 Canonical vector, 670–676, 679, 682*f*, 683–687, 690  
 Censored Gaussian distribution, 101–103  
 Central credible interval (CCI), 213–214, 306–307, 431–432  
 Central Limit Theorem, 94, 113–114, 119, 149, 151, 226, 243, 470, 541, 587, 591, 683, 700  
   multivariate, 601–602  
   univariate, 98  
 Centroid clustering, 723–724  
 Certainty equivalent, 308–310  
 Chain-dependent process, 494–495  
 Chaining, 724–725, 729, 730*f*  
 Chaos, 5, 313  
 Characteristic value, 571  
 Characteristic vector, 571.  
   *See also* Eigenvector  
 Chartjunk, 63–64  
 Check function, 286, 431  
 Chi-square ( $\chi^2$ ) distribution, 110, 478, 543  
 Chi-square ( $\chi^2$ ) test, 161–162, 166, 478–479  
 Cholesky decomposition, 574–575, 684  
 City-block distance, 722  
 Classical nonparametric tests, 171–177  
 Classical statistical forecasting, 294–296  
 Classification and regression trees (CART), 716–717  
 Clayton skill score (CSS), 381, 467–468  
 Cluster analysis  
   *vs.* discriminant analysis, 721  
   hierarchical clustering, 722–732  
     agglomerative methods using distance matrix, 722–725  
     dendrogram, or tree diagram, 725–726  
     divisive methods, 732  
     number of clusters, 726–732  
     Ward's minimum variance method, 725  
   nonhierarchical clustering, 732–734  
     clustering using mixture distributions, 734  
     K-means method, 732–733  
     nucleated agglomerative clustering, 733–734  
 Coefficient of determination ( $R^2$ ), 242, 244–245, 254, 256  
 Coefficient profile graph, 272–274, 273*f*  
 Cohen's kappa, 380  
 Co-inertia analysis, 689  
 Column binding, 554  
 Column vector, 559–561, 563, 566–567  
 Combined PCA, 632, 678  
 Complements, 11–13  
 Complete-linkage clustering, 727–729  
 Computational form, 56–59, 62, 90, 238–240  
   correlation coefficient, 57, 59–60, 68–69, 238  
   regression coefficient, 243–245, 262  
   skewness coefficient, 57–58  
   standard deviation, 26–27, 50–51, 159–160  
 Concordant, 60  
 Conditional bias, 348, 372, 396–398, 403–407, 436  
 Conditional climatology, 16, 294  
 Conditional distribution, 98*f*, 99–100, 100*f*, 232, 238–239, 294, 344–345, 370–372, 378, 399, 401, 403, 405, 591–593, 711–712  
 Conditional probability, 13–15, 14*f*, 18, 53–54, 99, 346–347, 370–371, 386, 401, 405, 488, 490–491, 496, 701  
 Conditional quantile plot, 398–399, 399*f*  
 Conditioning event, 13–15  
 Confidence interval, 148–151, 160, 172, 186, 188–189, 201, 243, 251–253, 470–475, 477, 516*f*, 541–542, 608–610, 612  
   regression, 251  
   and ROC diagram, 473–475  
   simultaneous joint, 473–474  
 Conjugate distribution, 216–227  
 Consecutive integers  
   mean of, 172  
   variance of, 172  
 Consensus forecasting, 309  
 Consistent scoring rule, 417  
 Contingency table, 374–383, 470–473  
 Continuity correction, 149–150, 180  
 Continuous distributions, 91–122  
   beta distributions, 111–113  
   distribution functions and expected values, 91–93  
   extreme-value distributions, 113–118  
   gamma distributions, 104–111  
     evaluating gamma distribution probabilities, 107  
     gamma distribution in operational climatology, 108  
   Gaussian distributions, 93–100  
     bivariate normal distribution and conditional probability, 97–100  
     evaluating Gaussian probabilities, 96  
     mixture distributions, 119–122  
   *vs.* discrete distributions, 78–79  
 Continuous random variable, 78–79, 92–93, 161, 487, 501  
 Continuous ranked probability score (CRPS), 330–333, 335, 425–427, 426*f*, 469  
 Contrast, 24, 49–50, 75, 78, 91, 107, 158–160, 209, 265, 300–301, 316–317, 322–324, 346, 373, 459, 553, 613–614, 635, 732  
 Copula, 353–358  
 Correct rejection, 375, 544  
 Correlation  
   auto-, 61–63, 156, 158–160, 168, 190, 267, 397–398, 470, 478, 487, 491–493, 494*f*, 502, 504, 506–507, 508*f*, 511–513, 537–538, 539*f*, 643, 649–650  
   heterogeneous, 672, 675, 677, 688–689  
   Kendall's  $\tau$ , 59–61  
   Lagged, 61–62, 513, 608, 661  
   Pearson product-moment, 54–55, 59, 61, 69, 398, 453  
   Spearman rank, 59–61, 354–355  
 Correlation maps, 72–74, 677*f*, 689*f*  
 Correlation matrix, 68–70, 72–73, 270, 555, 565–567, 619, 623–626, 630–632, 632*f*, 636, 644, 660, 669, 673–674, 700  
 Cosine function, 517–518  
   wave, 519  
 Cost/loss ratio, 463–468  
 Counting norm statistic, 193  
 Covariance, 54–56, 131, 485, 555, 558, 565, 571, 578–580, 610–611, 631–632, 660, 663–664, 673–674, 690  
 Covariance matrix, 318, 429–430, 445, 555, 561, 565, 571–572, 576–581, 587–591, 597, 602–604, 606–607, 618, 620, 625*t*, 628, 647, 657–659, 664, 676, 691, 697, 700  
   eigenvalues and eigenvectors of, 570–574  
   for linear combinations, 580–582  
   for regression parameters, 567, 665–666  
   *vs.* correlation matrix, 623–626

- Covariance stationarity, 485
- Credible interval forecast
  - central, 306–307
  - fixed-width, 306–307
  - operational, 307
- Critical region, 146, 165
- Critical success index (CSI), 377
- Critical value, 146, 150–151, 163–168, 249*f*, 478*t*, 595–596, 640, 643
- Cross-validation, 38–39, 265–268, 734
- CRPS (continuous ranked probability score), 330–333, 335, 425–427, 426*f*, 469
- CSI (critical success index), 377
- CSS (Clayton skill score), 381, 467–468
- Cumulative distribution function, 43–44, 85, 163, 165–166, 330
- Cumulative frequency distribution, 40–44, 42*f*
- Cumulative probability, 40–44, 83*f*, 85, 92, 97, 107, 111, 117, 124, 165, 248, 309, 328, 337, 346–347, 430, 432, 588–591, 595–596, 607
- Cyclostationarity, 486
- D**
- Data matrix, 554, 567, 631–632, 655, 659–660, 666, 721
- Dawid-Sebastiani score, 428–430
- Deciles, 25, 337–338
- Decustering, 119
- Degrees of freedom, 110, 152, 161–162, 169–171, 200–201, 203–205, 276, 371, 391, 478, 492–493, 496, 543, 588–591, 594–596, 602–603
- Delay window, 661, 662*f*, 663, 665
- Democratic voting method, 328
- DeMorgan's Laws, 13
- Dendrogram, 725–726
- Dependence template, 354–355
- Dependent sample, 258
- Dependent variable, 15–16, 236
- Derived predictor variables, 254–256
- Descriptive statistics, 3–4
- Developmental sample, 258, 260, 264, 267, 298
- Deviance, 169, 276
- Diagnostic verification, 370–371, 384–385, 399
- Diagonal matrix, 562, 565–567, 571–572, 574–575, 577, 619, 656–657, 659–660, 670–671, 685–688, 692
- Dichotomous event, 82, 274–275, 308, 363–364, 400–402, 400*r*, 410, 413, 416–417, 427, 435–436, 465–466, 490–491
- Diebold–Mariano test, 158–160, 511–513
- Diffuse prior, 215, 222–223, 225
- Discordant, 60
- Discrete distributions, 79–88
  - binomial, 79–82
  - geometric, 82–83
  - negative binomial, 83–86
  - Poisson, 86–88
  - vs. continuous, 78–79
- Discrete Fourier transform, 529–530, 531*t*, 534, 542
- Discrete random variable, 78–79, 88–89, 92–93, 161, 487–488
- Discriminant analysis, 695–696, 699–700, 703–718, 721, 727
- Discriminant coordinates, 705
- Discriminant function, 695–698, 699*f*, 704, 710–711
- Discriminant space, 698, 705–708
- Discrimination, 372, 379, 410–412, 695
- Discrimination and classification
  - alternatives to classical discriminant analysis, 711–714
  - discrimination vs. classification, 695
  - forecasting with discriminant analysis, 710–711
  - multiple discriminant analysis (MDA), 704–710
  - Fisher's procedure for more than two groups, 704–708
  - minimizing expected cost of misclassification, 708
  - probabilistic classification, 709–710
  - separating two populations, 695–703
  - equal covariance structure, Fisher's linear discriminant, 695–700
  - Fisher's linear discriminant for multivariate normal data, 700
  - minimizing expected cost of misclassification, 701–702
  - unequal covariances, quadratic discrimination, 702–703
- Discrimination diagram, 410–412
- Discrimination distance, 411, 411*f*
- Dispersion, 25–26, 44, 46, 120, 183–185, 286, 321, 323–324, 326–329, 339, 354, 360–362, 370–371, 407, 557, 582–584, 588–591, 602, 622, 700
- Dispersion matrix, 555, 639, 642, 722
- Distance
  - city-block, 722
  - Euclidean, 428, 444–445, 556–557, 582–583, 705, 721–722
  - Karl-Pearson, 722, 727, 729, 733
  - Mahalanobis, 557–558, 582–584, 587–588, 592, 595–596, 596*f*, 602–603, 608–609, 611*f*, 628, 697, 722
  - Minkowski, 722
- Distance matrix, 721–722
- Distribution-free tests, 144
- Distribution parameters, 77–78, 85–86, 89*t*, 93, 95, 114, 117–118, 125, 127, 130, 143, 163–164, 166–168, 335, 425, 734
- Distributions-oriented verification, 371
- Divisive clustering, 732
- Domain expertise, 259
- Domain size effects, 638
- Dominant variance selection rules, 651
- Dotplot, 34*f*, 35, 37
- Dot product, 560–561, 563–564, 571, 573, 620–622, 630, 666, 674–675, 697, 699
- Double penalty problem, 455–456
- Drop-size distribution, 109
- Dummy variable, 240–241, 254–255, 274–275
- Durbin-Watson test, 248–250
- Dyadic product, 564
- E**
- Economic value
  - value score, 466–467
  - verification based on, 463–468
  - connections with other verification statistics, 467–468
  - optimal decision making and cost/loss ratio problem, 463–466
- EDA (exploratory data analysis), 23–75, 721
- Eddy covariance, 54–55
- Effective multiplet, 647–648, 650, 652, 657
- Effective sample size, 157–160, 476, 511–513, 643, 649–650
- Eigendecomposition, 684–687, 705–708
- Eigenspace, 571
- Eigenvalue
  - of  $(2 \times 2)$  symmetric matrix, 573
  - of covariance matrix, 572
  - direct extraction of, from sample covariance matrix, 658–659
  - rules based on size of last retained, 641–642
  - sampling properties of, 645–650
  - of square matrix, 570–574
- Eigenvalue spectrum, 640, 642, 651
- Eigenvector, 570–574
  - of  $(2 \times 2)$  symmetric matrix, 573
  - of covariance matrix, 572
  - direct extraction of, from sample covariance matrix, 658–659
  - PCA, 627–628, 632–634
  - rotation of, 650–658
  - sampling properties of, 645–650
- Eigenvector scaling, 627–628, 627*t*
- sensitivity of orthogonal rotation to initial, 655–657
- Eighth, 25
- El Niño, 35, 52, 67–68, 119, 232, 327, 680, 713, 713*f*
- Embedding dimension, 661, 663–665
- EM (expectation-maximization) algorithm, 120, 128–131, 130*r*, 339, 734
- Empirical copula, 353–356, 358
- Empirical copula coupling (ECC), 354, 356, 357*f*
- Empirical cumulative distribution function, 40, 85, 165–166
- Empirical distributions
  - exploratory techniques for paired data, 52–63
  - autocorrelation function, 62–63

- Empirical distributions (*Continued*)
- Pearson (ordinary) correlation, 54–59
    - scatterplots, 53–54
    - serial correlation, 61–62
    - Spearman rank correlation and Kendall's  $\tau$ , 59–61
  - graphical summary techniques, 29–44
    - boxplots, 30–31
    - cumulative frequency distributions, 40–44
    - histograms, 34–35
    - kernel density smoothing, 35–40
    - other boxplot variants, 33–34
    - schematic plots, 31–33
    - stem-and-leaf display, 29–30
  - numerical summary measures, 25–29
  - reexpression, 44–52
    - power transformations, 44–48
    - standardized anomalies, 50–52
  - vs. parametric, 77–78
  - visualization for higher-dimensional data, 63–75
    - correlation maps, 72–74
    - correlation matrix, 68–69
    - glyph scatterplot, 65–66
    - rotating scatterplot, 67–68
    - scatterplot matrix, 70–72
    - star plot, 64
- Empirical orthogonal function (EOF) analysis, 617–667
- Empirical orthogonal variable, 626–627
- Empirical orthogonal weights, 626
- Energy score, 426, 428, 444–445
- Ensemble average, 323–324, 342, 344–345, 350, 437–438
- Ensemble consistency, 432–434, 437–438, 440–441
- Ensemble dressing, 338–343
- Ensemble energy score, 444–445
- Ensemble forecasting
  - choosing initial ensemble members, 321–323
  - effects of model errors, 325–327
  - ensemble average and ensemble dispersion, 323–324
  - ensemble forecasts, 319–327
  - graphical display of ensemble forecast information, 359–366
  - statistical postprocessing, ensemble MOS, 327–328
  - stochastic dynamical systems in phase space, 313–317
- Ensemble forecasts, verification of, 432–445
  - minimum spanning tree (MST) histogram, 439–441
  - verification rank histogram, 433–437
- Ensemble Kalman filter (EnKF), 322
- Ensemble members, 319–323, 320*f*, 326–330, 335, 337–342, 345, 347–348, 432–434, 440–441, 440*f*, 731–732
- Ensemble meteogram, 362–363, 365*f*
- Ensemble MOS, 327–351
- Ensemble postprocessing, 327–359
- Ensemble ranked probability score, 420–422
- Entropy, 417, 436, 717
- Epanechnikov kernel, 36*t*
- Error function, 95–96
- Equitability, 383, 391
- Equitable threat score (ETS), 381, 383
- Equivalence of events, 377
- Equivalent number of independent samples, 157. *See also* Effective sample size
- Erlang distribution, 109
- Error bars, 148, 150, 438*f*
- Euclidean distance, 428, 444–445, 556–557, 557*f*, 583, 705, 721–722
- Euclidean norm, 560, 685
- Euler exponential notation, 534
- Euler's constant, 93*t*, 115
- Event
  - compound event, 7–8
  - elementary event, 7–8
- Excess kurtosis, 28–29
- Exchangeability principle, 182–183, 185
- Expansion coefficient, 626–627
- Expected cost of misclassification, 701–702, 708
- Expected payoff, 308–309
- Expected value, 88–93, 252, 386–387, 469–470, 504–505, 514, 577–578, 627–628
- Experimental design
  - completely randomized, 199–201
  - randomized block, 202–205
- Exploratory data analysis (EDA), 23–75, 721
- Exponential distribution, 104–105, 109, 120–122, 134, 137, 170, 494–495, 540, 544–545, 607
- Extended EOF (EEOF) analysis, 632, 634
- Extended logistic regression (XLR), 97, 336–337
- Extremal dependence index (EDI), 379
- Extremal Types Theorem, 113–114
- Extreme event, 113, 117, 322–323
- Extreme-value distribution, 103, 113–119, 190
- F**
- F* distribution, 184–185, 602
- F*-ratio, 264
- F*-test, 184–185, 242–243
- Factor analysis, 617
- Factorial function, 84. *See also* Gamma function
- False alarm, 375, 377–379
- False alarm rate, 378–381, 383–384, 412–415, 470–475
- False alarm ratio, 378, 382–384, 387–388, 470–471
- False discovery rate, 194–197
- False-positive fraction, 379
- False rejection, 146, 197
- Fast Fourier transform (FFT), 534
- Fences, 31–33, 32*f*
- Field forecasts, 445–451, 456, 459–462, 665
- Field significance, 192–199
- Filliben test for normality, 166–168, 594
- Finley tornado forecasts, 381, 382*t*, 472–475
- First-order autoregression, 500–504
  - in choosing block length, 191–192
  - in estimating effective sample size, 156
- Fisher's linear discriminant function, 696–697
- Fisher-Tippett distribution. *See* Gumbel distribution
- Fisher Z-transformation, 49
- Flat prior, 215
- Folding frequency, 535–536
- Forecasting
  - classical, 294–296
  - ensemble, 313–366
  - MOS, 296–301, 327–328
  - perfect prog, 296–301
  - subjective, 304–305
  - with autoregressive model, 516
  - with CCA, 676–678
  - with discriminant analysis, 710–711
  - with MCA, 691
  - with RDA, 692
- Forecasting forecast skill, 324
- Forecast quality, 369, 372–373, 403–405, 424, 427, 446–447, 466
- Forecast skill, 324, 331, 373–374, 410, 468, 495, 692
- Forecast variance, 318
- Forecast verification
  - nonprobabilistic forecasts of continuous predictands
    - conditional quantile plots, 398–399
    - scalar accuracy measures, 394–397
    - skill scores, 397–398, 516
  - nonprobabilistic forecasts of discrete predictands, 374–394
    - 2×2 contingency table, 374–376
    - conversion of probabilistic to nonprobabilistic forecasts, 386–388
    - extensions for multicategory discrete predictands, 388–394
    - scalar attributes characterizing 2×2 contingency tables, 376–379
    - skill scores for 2×2 contingency tables, 379–383
    - which score, 383–385
  - nonprobabilistic forecasts of fields, 445–463
    - anomaly correlation, 452–455
    - field verification based on spatial structure, 455–463
    - general considerations for field forecasts, 445–447
    - mean squared error, 448–452
    - S1 score, 447–448

- probability forecasts for continuous predictands, 425–431
  - probability forecasts of discrete predictands, 400–424
    - algebraic decomposition of Brier score, 402–404
    - Brier score, 401–402
    - discrimination diagram, 410–412
    - hedging, and strictly proper scoring rules, 418–419
    - joint distribution for dichotomous events, 400–401
    - probability forecasts for multiple-category events, 419–424
    - reliability diagram, 404–410
    - ROC diagram, 412–416
  - sampling and inference for verification statistics, 470–480
    - reliability diagram sampling characteristics, 476–478
    - resampling verification statistics, 479–480
    - ROC diagram sampling characteristics, 473–475
    - sampling characteristics of contingency table statistics, 470–473
  - verification based on economic value, 463–468
  - verification of ensemble forecasts, 432–445
    - minimum spanning tree (MST) histogram, 439–441
    - verification rank histogram, 433–437
  - Forward selection, 260–261, 263–264, 263*f*, 267, 271, 273–274, 296
  - Fourier analysis, 548, 662–663
  - Fourier line spectrum, 529–533
  - Fourth-spread, 26
  - Fractiles, 24
  - Fractions skill score (FSS), 457–458
  - Frechet distribution, 115
  - Frequency domain analysis, 486–487, 528–529
  - Frequency interpretation of probability, 10
  - Full enumeration, 182–183
  - Full rank, 564, 574–575, 587
  - Fundamental frequency, 517, 519, 524, 531
- G**
- Gamma distribution, 104–111
    - algorithm for maximum likelihood estimates of parameters, 127
    - evaluating gamma distribution probabilities, 107
    - gamma distribution in operational climatology, 108
    - generalized, 110–111
    - mode, 223
  - Gamma function, 83–84, 84*t*, 114, 212, 223, 334–335, 588–591
  - Gandin-Murphy skill score (GMSS), 391–393
  - Gaussian approximation to the binomial, 149–150, 150*f*, 471, 474–475
  - Gaussian copula, 358
  - Gaussian distribution, 93–100
    - bivariate normal distribution and conditional probability, 97–100
    - censored, 102–103
    - circular, 192
    - evaluating Gaussian probabilities, 96
    - lognormal, 101–102
    - truncated, 102
  - Gaussian ensemble dressing (GED), 342
  - Gaussian kernel, 36–38, 40, 233, 342
  - Gauss-Markov theorem, 268–269
  - Generalized extreme value (GEV) distribution, 103, 114, 333
    - Type I (Gumbel), 114–115
    - Type II (Frechet), 115
    - Type III (Weibull), 115–116
  - Generalized linear model, 274, 292
  - Generalized Pareto distribution, 118–119, 122, 336
  - Generalized variance, 588–591
  - Geometric distribution, 82–83, 117, 220–221, 491
  - Gerrity skill score, 391–393
  - Gibbs sampler, 231–234
  - Gilbert score, 377
  - Gilbert skill score (GSS), 377, 381
  - Gini Index, 717
  - Global significance, 192
  - Glyph scatter plot, 65–66, 65*f*, 399
  - Goodness-of-fit tests, 160–168
    - Chi-square, 161–162
    - Filliben Q–Q correlation, 166–168
    - Kolmogorov–Smirnov (K–S), 160–168
    - Lilliefors, 162–165
    - Shapiro–Wilk, 166
  - Graphical summary techniques, 29–44
    - boxplots, 30–31
    - boxplot variants, 33–34
    - cumulative frequency distributions, 40–44
    - histograms, 34–35
    - kernel density smoothing, 35–40
    - schematic plots, 31–33
    - stem-and-leaf display, 29–30
  - Gringorten's skill score (GSS), 380–381
  - Gumbel distribution, 114–115, 168
- H**
- Half-Brier score, 402
  - Hanssen-Kuipers discriminant, 380–381
  - Harmonic analysis, 517–527
  - Harmonic function
    - representing simple time series with, 518–520
    - as uncorrelated regression predictors, 528–529
  - Harmonic predictors, 256, 529
  - Harmonics, higher, 524–527
  - Heatmap, 70
  - Hedging, 418–419
  - Heidke skill score (HSS), 380, 389–390
  - Heterogeneous correlations, 672, 675, 677, 688–689
  - Heteroscedasticity, 245–247
  - Hexbin plot, 65–66
  - H–F diagram, 384–385, 385*f*
  - Hierarchical clustering
    - agglomerative methods using distance matrix, 722–725
    - dendrogram, or tree diagram, 725–726
    - divisive methods, 732
    - number of clusters, 726–732
    - Ward's minimum variance method, 725
  - Hierarchical models, 232–234
  - Higher harmonics, 517, 524–527
  - Highest posterior density (HPD) interval, 214
  - Hinges, 25
  - Hinkley  $d_h$  statistic, 46–48, 49*f*
  - Histogram
    - bivariate, 65–66, 399
    - minimum spanning tree (MST), 439–441
    - superposition of PDFs onto, 123
    - verification rank, 346, 430–431, 433–437
  - Hit, 377–378
  - Hit rate, 377, 379–384, 387–388, 412–415, 470–471, 473–474
  - Homogeneous correlations, 671–672, 675, 677, 689–691
  - Homoscedasticity, 246–247
  - Honest significant difference, 201
  - Hotelling  $T^2$ , 602–608, 700
  - Hyperparameter, 213–214, 216–224, 232–234
  - Hyperprior, 232–234
  - Hypothesis test
    - Bonferroni method, 201, 473–475, 477, 544, 609–611
    - counting norm, 193
    - false discovery rate (FDR), 194–197
    - multiplicity problem for independent tests, 192–194
    - nonparametric tests, 171–192
      - bootstrap, 185–192
      - classical nonparametric tests for location, 171–177
      - field significance given spatial correlation, 197–199
      - permutation tests, 182–185
      - resampling tests, 181–182
    - parametric tests, 151–171
      - goodness-of-fit tests. (*see* Goodness-of-fit tests)
      - Hotelling  $T^2$ , 602–608
      - likelihood ratio test, 168–171
      - one-sample  $t$  test, 151–153

Hypothesis test (*Continued*)

- test for differences of mean under serial dependence, 156–160
- tests for differences of mean under independence, 153–155
- tests for differences of mean for paired samples, 155–156
- PCA rules based on, 642–644

**I**

- Ideal user, 467–468
- Identity matrix, 561–562, 564
- Ignorance score, 331, 416–418, 424, 428–429, 444
- Improper prior, 215, 222
- Incomplete beta function, 113
- Incomplete gamma function, 107–109
- Independence, 14–16, 86–87, 153–160, 229–231, 279–281, 472–473, 476, 479, 491–493, 610
- Independent tests, 192–194, 268
- Independent variable, 125–126, 236, 284, 592–593
- Indicator variable, 129–130, 254–255, 705
- Inferential statistics, 3–4
- Infinite monkeys theorem, 192–193
- Information matrix, 131, 279
- Inner product, 560, 564
- Innovation variance, 502
- Interquartile range (IQR), 26, 32, 37–38, 365f
- Intersection, 12–14, 214, 377, 415, 564
- Interval score, 432
- Intrinsic EOFs, 635
- Invariance, 182
- Inversion method for random variates, 134–135
- Inverting hypothesis tests, 148–151

**J**

- Jaccard coefficient, 377
- Jackknife, 189–190, 268
- JAGS (Just Another Gibbs Sampler), 232
- January thaw, 51
- Joint distribution
  - for dichotomous events, 400–401
  - of forecasts and observations, 370–371
  - multivariate normal, 587
- Joint probability, 12–14, 375, 492, 609
- Jordan decomposition, 571–572

**K**

- K-means clustering, 732–733
- Kaiser's rule, 641
- Kalman filter, 301
- Karl Pearson distance, 722, 727, 729
- Kendall's  $\tau$ , 59–61, 179–180, 283–284

- kernel density estimate, 37
  - discrimination and classification using, 713–714
  - simulation from, 137–138
  - smoothing, 35–40
- Kolmogorov-Smirnov (K-S) test, 162–165
- Kuipers' performance index, 380–381
- Kurtosis, 28–29

**L**

- L1 regularization, 272, 678
- Lasso, 272–274
- L-moments, 114–117
- L-scale, 183, 187
- Ladder of powers, 45
- Lag-1 autocorrelation, 61–62, 61f, 156, 158–160, 180–181, 267, 397–398, 476, 478, 491, 501–502, 544, 649–650
- Lagged correlation, 61–62, 513, 608, 661
- Lagged covariance, 598–600
- Latent value, 571
- Latent vector, 571. *See also* Eigenvector
- Law of cosines, 451
- Law of large numbers, 10
- Law of total probability, 16–17, 406–407
- Least-absolute-deviation (LAD) regression, 247–248, 286, 293
- Least-squares regression, 178, 235–236, 236f, 242–246, 270, 284–286, 295–296, 347, 350, 398, 408, 504, 620, 658
- Leave-one-out cross validation, 266–268, 710
- Leptokurtic, 28–29
- Level of a test, 145–146
- Likelihood, 18, 47–49, 125, 169, 210–213, 215–217, 219, 221, 225, 232–234, 276, 305, 344, 346, 411–413, 702, 709–710
- Likelihood-base rate factorization, 371, 376, 380–381, 383–384, 400f, 404, 410
- Likelihood function, 125–127, 169, 227, 509, 734
- Likelihood ratio test, 168–171, 180, 276, 278, 499
- Lilliefors test, 162–165
- Linear algebra, 69, 558
- Linear combinations, 153, 291, 322, 329, 564, 580–582, 591, 596–597, 607, 612–614, 617–618, 620–623, 628, 635, 662, 669, 674, 678, 688, 690, 692, 695–696, 702
- Linear congruential generator, 132–134
- Linear correlation, 54, 58
- Linear discriminant analysis, 699f, 703
- Linear regression
  - analysis of variance table, 240–241
  - derived predictor variables in multiple regression, 254–256
  - distribution of residuals, 238–240
  - examining residuals, 245–250
  - goodness-of-fit measures, 241–243
  - multiple linear regression, 253–256

- prediction intervals, 250–253
- principal components, 619
- sampling distributions of regression coefficients, 243–245
- simple linear regression, 244
- Link function, 274, 279
- Loading, 626
- Locality, 424
- Local test, 192–197, 198f
- Location, 26, 171–177, 427f, 698f
- Loess smoothing, 281
- Logarithmic score, 416–418
- Logarithmic transformation, 45, 47–48, 49f, 167–168, 233, 248
- Log-eigenvalue (LEV) diagram, 640–641
- Logistic distribution, 103–104
- Logistic function, 275–276, 291–292, 711–712
- Logistic regression, 274–279, 336–338, 340–341, 404, 711–712, 712f
- Logit function, 103–104
- Logit transformation, 275
- Log-hyperbolic sine transformation, 49
- Log-likelihood, 46–47, 126–128, 130–131, 169–170, 169f, 276–278, 280–281, 331, 429–430, 498–499
- Lognormal distribution, 101–102, 336
- Log-odds, 49, 275, 336–337
- Log-Pearson III distribution, 110–111
- Loss, 182, 300–301, 308–309, 386, 463–466

**M**

- Madden-Julian oscillation, 634
- Mahalanobis distance, 557–558, 582–584, 587, 595–596, 603, 611f, 628, 697, 722
- Mahalanobis transformation, 597, 628, 685–688
- Mann-Kendall trend test, 178–181, 283–284
- Marginal distribution, 53–54, 98, 352, 354–356, 370–371, 375–376, 380–381, 396, 403, 591, 594
- Marginal total, 375, 492
- Markov chain, 79–80, 227, 487–492, 494–499
- Markov chain Monte Carlo (MCMC), 227–228
- Markovian property, 487–488, 490–492, 497, 502
- Markov process, 83, 493, 501
- Marshall-Palmer distribution, 109
- Matrices
  - addition of, 561–570
  - computation of covariance and correlation matrices, 565
  - correlation. (*see* Correlation matrix)
  - covariance. (*see* Covariance matrix)
  - data. (*see* Data matrix)
  - determinant, 564
  - diagonalization, 572
  - dimension, 561

- distance, 721–722
  - element, 561–570
  - identity, 561–570
  - inverse, 131, 565, 676
  - invertable, 591, 601, 679–681
  - lower-triangular, 574
  - multiple linear regression expressed in
    - matrix notation, 567
  - multiplication of, 562
  - nonsingular, 564
  - orthogonal. (*see* Orthogonal, matrix)
  - partitioned, 579, 669
  - positive definite, 572, 584, 587
  - random vectors and, 577–584
    - (*see also* Random vectors and matrices)
  - square, 561–562, 569
  - square root. (*see* Square-root matrix)
  - subtraction of, 575
  - symmetric, 573
  - Toeplitz, 661–662, 664
  - trace, 564
  - transpose, 562
  - triangular, 69, 135–136, 138, 506, 508, 574
  - unitary, 569
  - Matrix algebra
    - eigenvalues and eigenvectors of square matrix, 570–574
    - matrices. (*see* Matrices)
    - multivariate distance, 556–558
    - random vectors and matrices, 577–584
      - expectations and other extensions of univariate concepts, 577–579
      - linear combinations, 580–582
      - Mahalanobis distance, 582–584
      - partitioning vectors and matrices, 579–580
    - singular-value decomposition (SVD), 576–577
    - square roots of symmetric matrix, 574–576
    - vectors, 577–584
  - Maximum covariance analysis (MCA), 577, 687–691
  - Maximum-distance clustering, 723
  - Maximum likelihood, 47–49, 102, 106–107, 114, 116, 120, 125–131, 170, 213, 219–220, 276, 278f, 331, 349
  - Maximum likelihood estimator, 106, 161–162, 490, 589
  - Maximum margin classifier, 715–716
  - Mean, 26
    - test for differences of
      - under independence, 153–155
      - for paired samples, 155–156
    - multivariate (vector), 601–602
    - under serial dependence, 156–160
  - Mean absolute error (MAE), 394–395, 397
  - Mean error (ME). *See* Bias
  - Mean squared error (MSE)
    - to determine stopping criterion, 264
    - for non-probabilistic forecasts, 395
    - regression, 241
  - Mean vector, 318, 352, 577–578, 581, 591, 596–597, 600–614, 669, 679–681, 683, 696–697, 700, 709
  - Measures-oriented verification, 371
  - Median, 25–27, 30–33, 39–40, 43–44, 46, 50, 107–108, 117, 181, 189, 230–231, 284, 286, 293, 309, 334, 395, 642
  - Median absolute deviation, 27
  - Median-slope regression, 283–285
  - Member-by-member postprocessing, 358
  - Mersenne twister, 133–134
  - Metaverification, 383
  - Method of maximum likelihood, 106, 114, 125, 168–169, 276
  - Method of moments, 85–88, 95, 112, 114, 116, 350
  - Method of successive subdivision, 309
  - Metropolis–Hastings algorithm, 228–231
  - Minimax linkage, 723, 726
  - Minimum-distance clustering, 723
  - Minimum spanning tree (MST) histogram, 439–441
  - Minimum volatility methods, 191–192
  - Minkowski metric, 722
  - Misclassification, 701–702, 708, 728–729
  - Miss, 432
  - Mixed exponential distribution, 120–121, 137, 494–495
  - Mixture distributions, 119–122, 137–138, 338–339, 343, 734
  - Mixture probability density, 119–122
  - MODE, 459–461
  - Model errors, 325–328
  - Model output statistics (MOS), 296–301, 386, 399, 678–679
  - Modes of variation, 626
  - Monte-Carlo test, 181
  - Moving-average process, 513
  - Moving-blocks bootstrap, 191–192, 479–480, 650
  - Moving-blocks cross validation, 266–268
  - Multichannel SSA (MSSA), 665
  - Multicollinearity, 665
  - Multimodel ensemble, 326, 329–330
  - Multinomial distribution, 86
  - Multiple correlation, 270–271
  - Multiple discriminant analysis (MDA)
    - Fisher's procedure for more than two groups, 704–708
    - minimizing expected cost of misclassification, 708
    - probabilistic classification, 709–710
  - Multiple linear regression, 236, 253–256, 274, 298, 567, 580, 665
  - Multiple-state Markov chains, 495–496
  - Multiplicative law of probability, 14, 16–17, 80–83, 387–388
  - Multiplicity, 192–199, 242–243, 277, 498, 544–545, 609
  - Multivariate autoregression, 598–599
  - Multivariate central limit theorem, 587, 601–602
  - Multivariate distance, 556–558
  - Multivariate kurtosis, 594
  - Multivariate multiple regression, 303–304, 691
  - Multivariate normal (MVN) distribution
    - assessing multinormality, 593–596
    - definition of, 587–590
    - four handy properties of, 591–593
    - Hotelling's  $T^2$ , 602–608
    - inferences about multinormal mean vector, 601–614
    - interpretation of multivariate statistical significance, 611–614
    - multivariate central limit theorem, 601–602
    - probability ellipses, 588
    - simulation from multivariate normal distribution, 596–601
    - simultaneous confidence statements, 608–611
  - Multivariate outlier, 583–584, 594–596, 628–629, 629f
  - Multivariate rank histogram, 441–443
  - Multivariate skewness, 594
  - Multivariate time series, 597–601, 661
  - Mutually exclusive and collectively exhaustive (MECE), 79–80, 86, 161, 305–306, 374, 388, 401, 490, 695
- ## N
- Nadaraya-Watson kernel-weighted average, 40
  - Nash-Sutcliffe efficiency, 398
  - Nearest-neighbor bootstrap, 191
  - Negative binomial distribution, 83–86
  - Neural network, 716–718
  - Neural network classifier, 716–718
  - Newton–Raphson method, 127–128, 277–281
  - Nominal predictand, 419
  - Nonhierarchical clustering
    - clustering using mixture distributions, 734
    - $K$ -means method, 732–733
    - nucleated agglomerative clustering, 733–734
  - Nonhomogeneous Gaussian regression (NGR), 329–331
  - Nonhomogeneous regression, 328–336
    - censored, 334–336
    - Gaussian, 329–331
    - lognormal, 331–332
    - truncated, 333–334
  - Noninformative prior, 215

- Nonlinear regression
  - logistic regression, 274–279
  - Poisson regression, 279–281
  - probit regression, 275–276
- Nonparametric regression
  - local polynomial, 281–283
  - median-slope (Theil-Sen), 283–285
  - quantile, 283–285
- Nonparametric test, 171–192
  - bootstrap, 185–192
  - classical nonparametric tests for location, 171–177
  - permutation tests, 182–185
  - vs. parametric tests, 143–144
  - resampling tests, 181–182
- Nonprobabilistic forecast
  - of discrete predictands, 374–394
    - 2×2 contingency table, 374–376
    - conditional quantile plots, 398–399
    - conversion of probabilistic to nonprobabilistic forecasts, 386–388
    - extensions for multicategory discrete predictands, 388–394
    - scalar accuracy measures, 394–397
    - scalar attributes characterizing 2×2 contingency tables, 376–379
    - scores, 383–385
    - skill scores, 379–383, 397–398
  - of fields, 445–463
    - anomaly correlation, 452–455
    - field verification based on spatial structure, 455–463
    - general considerations for field forecasts, 445–447
    - mean squared error, 448–452
    - S1 score, 447–448
- Nonuniform random number generation, 134–136
- Normal distribution, 93–94, 97–100, 252, 587–588, 596–601, 628–629, 700, 709
- Normal equations, 237–238, 244, 568
- Normalization transformation, 50
- Normal probability plot, 124–125
- Normal quantile transform, 111, 189
- North Atlantic Oscillation (NAO), 232
- North *et al.* rule of thumb, 648–650
- Nowcasting, 294, 710
- Nucleated agglomerative clustering, 733–734
- Null distribution, 144–147, 149–150, 150f, 152, 155–156, 165, 170–173, 176–177, 180, 182–185, 184f, 194, 203–205, 493–494, 596, 603
- Null event, 8
- Null hypothesis, 144–154, 158–172, 193–194, 199, 201, 203–204, 242–245, 472–475, 492, 543, 594, 612–613, 642, 644–645, 678, 700
- Numerical summary measures, 25–29
- Nyquist frequency, 530–531, 533, 535–537, 535–536f
- O**
  - Objective forecast, 235
    - classical statistical forecasting, 294–296
    - operational MOS forecasts, 301–304
    - perfect prog and MOS, 296–301
  - Oblique rotation, 655
  - Observed Fisher information, 131, 589
  - Odds ratio, 49, 276, 377–378, 381, 391, 471–473
  - Odds ratio skill score (ORSS), 381, 385, 385f
  - One-point correlation map, 72–74
  - One-s.e. rule, 267f, 268, 271–274, 277–279
  - One-sample *t* test, 151–153, 185, 201, 244, 472, 604
  - One-sided test, 147, 165
  - One-tailed test, 147, 153, 158–159
  - Optical flow, 462–463
  - Optimal decision making, 463–466
  - Orbit, 313–314
  - Order selection criteria, 509–510
  - Order statistics, 24, 41–43, 195, 282
  - Ordinal predictand, 419–420
  - Ordinary least squares (OLS), 236, 277, 285, 293, 295–296, 347
  - Orthogonal
    - matrix, 569–570, 618
    - rotation, 570f, 651–652, 655–657
    - transformation, 569–571
    - vectors, 571
  - Orthogonality, 528, 571–572, 618–620, 630, 650–651, 654–656
  - Orthonormal, 571, 576, 671, 687–688
  - Outer product, 564, 566–567, 572, 598, 676
  - Outliers, 24, 26, 28, 53, 69, 176, 184, 247–248, 285–286, 395–396, 434, 583–584, 594–596, 628–629, 629f, 722
  - Overconfidence, 156, 308, 327, 407, 437
  - Overfitting, 258f, 259, 264, 266, 287–288, 509–510, 525–528, 710
  - Overforecasting, 65–66, 378, 399, 405–407, 434–436, 448
  - Overlap method, 154–155, 160
- P**
  - Pacific North America (PNA) pattern, 74, 630–631, 653, 677, 677f
  - Paired data
    - correlation matrix, 68–69
    - glyph scatterplot, 65–66
    - Pearson (ordinary) correlation, 54–59
    - rotating scatterplot, 67–68
    - scatterplot matrix, 70–72
    - scatterplots, 53–54
    - Spearman rank correlation and Kendall's *t*, 59–61
    - star plot, 64
  - Paired *t* test, 604
  - Parameter fitting
    - EM algorithm, 128–130
    - likelihood function, 125–127
    - method of moments, 85–88, 95, 105–106, 111–112, 114–115, 125, 350
    - Newton–Raphson method, 127–128, 279–281
  - Parameterization, 325–327
  - Parameters vs. statistics, 78
  - Parametric bootstrap, 190
  - Parametric distribution
    - continuous distributions, 91–122
      - beta distributions, 111–113
      - distribution functions and expected values, 91–93
      - extreme-value distributions, 113–118
      - gamma distributions. (*see* Gamma distribution)
      - Gaussian distributions. (*see* Gaussian distribution)
      - mixture distributions, 119–122
    - discrete distributions, 79–88
      - binomial distribution, 79–82
      - geometric distribution, 82–83
      - negative binomial distribution, 83–86
  - parameter fitting using maximum likelihood, 125–131
    - EM algorithm, 128–130
    - likelihood function, 125–127
    - Newton–Raphson method, 127–128
    - sampling distribution of maximum-likelihood estimates, 130–131
  - Parametric tests, 143, 151–171
    - goodness-of-fit tests, 160–168
      - comparing Gaussian and gamma distribution fits using  $\chi^2$  test, 161
      - comparing Gaussian and gamma fits using K-S test, 164
    - Filliben Q–Q correlation test or Gaussian distribution, 166
    - Hotelling  $T^2$  test, 602–608
    - likelihood ratio test, 168–171
    - one-sample *t* test, 151–153
    - for paired samples, 155–156
    - under serial dependence, 156–160
    - tests for differences of mean under independence, 153–155
    - vs. nonparametric tests, 143–144
  - Partial duration data, 118
  - Partial whitening, 693
  - Partitioning vectors and matrices, 579–580
  - Pascal distribution, 83
  - Pattern coefficients, 626
  - Pattern correlation, 453
  - Pattern significance, 192
  - Pattern vector, 626
  - Peaks-over-threshold (POT), 118–119
  - Pearson correlation, 54–59, 69, 72–73, 100, 238, 244–245, 254, 398, 528, 620
  - Pearson III distribution, 103f, 110, 168, 643–644



- Pearce skill score (PSS), 380–381, 383–384, 386–387, 389–391, 472–473
- Percentile, 24–25, 108, 109*f*, 147, 170, 186, 187–188*f*, 188–189, 244–245, 347, 461, 642
- Percent (or proportion) correct, 377, 380, 382, 388–390
- Perceptrons, 290–291
- Perfect prog, 297–301, 299*f*
- Performance diagram, 384, 384*f*
- Period, 117, 132, 169–170, 487, 507, 518–519, 530–533, 663
- Periodogram. *See* Fourier line spectrum
- Permutation test, 182–185, 187, 479–480
- Persistence, 61–62, 156–157, 295–296, 373, 397, 449, 453–455, 462, 488–489, 491, 493, 495, 499–500, 515
- Persistence parameter, 491, 493
- p*-hacking, 145–146
- Phase angle, 519–520, 520*f*, 522–525, 529
- Phase association, 452
- Phase shift, 519–520
- Phase space, 313–319, 321, 326, 432–433, 634
- PIT, 134
- PIT histogram, 430–431
- Pivotal statistics, 182
- Platykurtic, 28–29
- Plotting position, 41–43, 42*r*, 166, 417, 595–596
- Plug-in principle, 185–188
- Plume graph, 363, 366*f*
- Poisson distribution, 86–88
- Poisson process, 86–87
- Poisson regression, 233, 279–281
- Polya distribution, 218–220
- Pooled estimate of variance, 154, 186–188, 697–698, 700, 704–706, 709
- Posterior distribution, 210–211, 213–234, 322, 345
- Posterior predictive distribution, 215–216
- Posterior probability, 18, 713–714
- Positive predictive value, 378
- Power function, 147, 150–151, 151*f*
- Power method, 660
- Power spectrum, 529–530, 640
- Power of a test, 146–147
- Power transformation, 44–48, 101, 110–111, 246–247, 593
- Predictand, 236–237, 239–242, 244–248, 248*f*, 250–253, 256, 259–264, 266–267, 271, 274–275, 279, 285, 288, 292, 294, 300–301, 331, 333, 337–338, 346–347, 362–363, 369, 373, 397, 425, 468, 500, 679–680, 691, 710
- Prediction interval, 250–253
- Prediction markets, 309
- Predictive distribution, 215–216
- Predictor selection  
    cross validation, 265–268
- importance of careful predictor selection, 257–260
- screening predictors, 260–263
- stopping rules, 263–265
- Predictor variable, 236, 239*f*, 240, 244, 249*f*, 253–256, 260, 262–265, 270–271, 280, 286, 288, 290, 293, 297, 347
- Pre-rank, 441–443
- Principal component, 64, 595–596, 617–667
- Principal component analysis (PCA)  
    application of to geophysical fields, 629–638
- basics of principal component analysis, 617–629
- connections to multivariate normal distribution, 628–629
- definition of, 617–622
- scaling conventions in, 627–628
- varied terminology of, 626–627
- CCA applied to fields, 676–678
- computational considerations, 658–661
- for multiple fields, 632–634
- rotation of eigenvectors, 650–658
- rotation mechanics, 651–655
- sensitivity of orthogonal rotation to initial eigenvector scaling, 655–657
- sampling properties of eigenvalues and eigenvectors, 658–659
- for single field, 629–632
- S-mode, 631–632
- T-mode, 631–632
- truncation of principal components, 639–645
- in two dimensions, 620
- uses of, 661–667
- via SVD, 659–660
- Principal component selection rules, 640–641
- Principal components regression, 665–666
- Prior distribution, 210–212, 214–215, 217, 222, 226, 229–232, 344
- Prior probability, 18, 217, 701
- Probabilistic classification, 709–710
- Probability  
    axioms of, 7
- definition of, 7
- elements of, 7–9
- frequency interpretation of, 10
- multiplicative law of, 14, 16–17
- properties of, 11–18
- Bayes' Theorem, 17–18
- conditional probability, 13–14
- DeMorgan's Laws, 13
- domain, subsets, complements, and unions, 11–13
- independence, 14–16
- Law of Total Probability, 16–17
- Probability density function (PDF), 35–37, 91, 93*r*, 120*f*, 146, 149–150, 181, 209–210, 307*f*, 317–318, 340, 467, 557, 587
- Probability of detection (POD), 379
- Probability distribution function, 40, 80, 82–83, 83*f*, 85–86, 89*r*, 121, 123, 194, 210, 215–216, 220–221, 370, 491
- Probability of false detection (POFD), 379
- Probability forecasts  
    for continuous predictands, 425–431
- for discrete predictands, 400–424
- algebraic decomposition of Brier score, 402–404
- Brier score, 401–402
- discrimination diagram, 410–412
- joint distribution for dichotomous events, 400–401
- for multiple-category events, 419–424
- reliability diagram, 404–410
- ROC diagram, 412–416
- Probability integral transform (PIT)  
    histogram, 134, 430
- Probability mass function, 80
- Probability of precipitation (PoP) forecast, 15, 18, 277, 305–306, 386, 387*f*, 411, 481*t*
- Probability plot, 85, 124–125
- Probability tables, 99, 101
- Probability wheel, 308
- Probit function, 103–104
- Probit regression, 275–276, 346–347
- Projection of a vector, 687–688, 692, 697
- Proper scoring rule, 418–419
- Proper value, 571
- Proper vector, 571. *See also* Eigenvector
- Proportional odds logistic regression, 337–338
- Proportion correct, 377, 380, 382, 386–391
- Pseudoperiodicity, 664
- Pseudo-random-number generator, 131–132
- p*-value, 149, 159–160, 607
- ## Q
- Quadratic discrimination, 702–703
- Quadratic form, 583–584, 587
- Quadratic kernel, 36
- Quantile, 24–25, 45, 107–109, 115, 117–118, 161–162, 188–190, 204–205, 213–214, 222–223, 328, 394, 431–432, 610–611
- Quantile function, 92, 97, 104, 109, 114, 116, 135–137, 166, 356, 593
- Quantile mapping, 111
- Quantile-quantile (Q-Q) plot, 124–125
- Quantile regression, 285–287
- Quantile regression forests, 290
- Quantile score, 431
- Quartic kernel, 36, 38*f*, 138
- Quartile, 25–26, 30, 31*f*, 32–33, 45, 309–310, 463*f*
- Quintile, 25
- ## R
- Randomization tests, 181
- Random forest, 288–290
- Random matrix, 577

- Random number generator, 131–134, 186, 599
  - Random variable, 40, 78–80, 88–93, 95–96, 110–111, 124, 161, 238, 286, 430, 487–488, 491, 501, 577, 580–581
  - Random vectors and matrices
    - expectations and other extensions of univariate concepts, 577–579
    - linear combinations, 580–582
    - Mahalanobis distance, 582–584
    - partitioning vectors and matrices, 579–580
  - Rank correlation, 59–61
  - Ranked data, 24
  - Ranked probability score
    - continuous (CRPS), 330, 425–427
    - discrete (RPS), 425
    - ensemble continuous (eCRPS), 444, 467
  - Rank histogram, 433–437
  - Rank-sum test, 172–176
  - Ratio of success, 381
  - Ratio of verification, 377–378
  - Rayleigh distribution, 137
  - Recalibration, 346, 410
  - Receiver operating characteristic (ROC), 412–416
  - Rectangular distribution, 112
  - Red noise, 501, 537–538, 544–545
  - Reduction of variance (RV), 398, 449–450
  - Redundancy analysis, 692
  - Re-expression
    - power transformations, 44–48
    - standardized anomalies, 50–52
  - Reference forecast, 373–374, 390, 402, 418, 495
  - Refinement, 31, 373, 401, 403, 407
  - Refinement distribution, 370–371, 375–376, 405, 407–409, 437
  - Reforecasting, 301
  - Regression. *See* Linear regression; Logistic regression; Poisson regression;
  - Regression coefficients, sampling distribution, 243–245
  - Regression constant, 253, 329
  - Regression equation, 237, 239, 244–245, 251, 254, 258–259, 259f, 261, 262f, 263–264, 267, 276, 286, 293–295, 297–298, 398, 504, 521–522, 529
  - Regression estimation of event probabilities (REEP), 275, 277
  - Regression parameters, 238, 243–246, 251–253, 263, 270–272, 274, 276–277, 279–282, 329–330, 333–334, 338–339, 350, 568, 665
  - Regression sum of squares, 240, 529
  - Regression tree, 288–289
  - Regularization, 268–274
  - Rejection level, 145
  - Rejection method for random variates, 135–136
  - Rejection region, 146–147
  - Relative (or reduction of) ignorance, 424
  - Relative operating characteristic (ROC), 412–416
  - Reliability, 372–373, 378, 398, 403–410, 436, 476–478
  - Reliability diagram, 404–410, 422, 436–437, 476–478
  - Reliability index, 436
  - Replacement, 182–183, 185–188, 650
  - Representativeness error, 468–469
  - Re-randomization tests, 181
  - Resampling test, 144, 171, 181–182, 479–480, 678
  - Residual, 199, 236–240, 245–250
  - Residual plot, 245–248, 246f, 254, 256
  - Residual variance, 239–240, 245–249, 502, 509–510, 692
  - Resistance, 23–24, 306
  - Resistant statistics, 26–27, 247–248
  - Resolution
    - and Brier score, 403
    - definition of, 372
  - Return period, 116–118
  - Ridge regression, 270–271
  - Ridge trace, 271–273, 272f
  - Robust statistic, 23–25
  - Root-mean-squared error (RMSE), 324, 395–397, 449, 451–452, 462–463
  - Rotated principal components, 651, 654–657
  - Rotating scatterplot, 67–68, 594–595
  - Row binding, 554
  - Rug plot, 53–54
  - Rule N, 642–644
- S**
- SAL (structure-amplitude-location), 461–462, 463f
  - S1 score, 447–448
  - Sample climatological distribution, 371, 392, 401
  - Sample space, 80, 86–87, 305–306, 487, 490
  - Sample statistics, 78, 95–96, 106, 130–131, 144, 148, 155, 190, 268, 376, 554–555, 599, 608, 650
  - Sampling distribution
    - cumulative probability, 43
    - of maximum-likelihood estimates, 130–131
    - test statistic, 143–145, 147–150, 152–156, 159, 161, 163, 165–166, 169–173, 177, 179, 181–183, 186–188, 204–205, 249–250, 478, 492–493, 594, 606
  - Scale parameter, 102, 104–105, 107, 110, 118–119, 333–335
  - Scaling transformation, 566–567, 597, 623
  - Scatterplot, 53–54
    - glyph, 65–66
    - matrix, 70–72
    - rotating, 67–68
  - Schaake shuffle, 356–357
  - Schematic plot, 31–33, 94–95
  - SCoTLASS, 658
  - Scree graph, 640, 641f
  - Screening regression, 260
  - Seed, for random-number generation, 132, 135
  - Selection rules, 640
  - Self-organizing maps, 734–738
  - Sensitive to distance, 391, 420, 424
  - Sensitivity, 24, 26, 194–195, 197, 294, 313, 319, 379, 429, 655–657
  - Serial correlation, 61–62, 119, 148, 156, 158–160, 172, 180–181, 191–192, 204r, 228–229, 248–250, 256, 477–478, 488–489, 491, 493, 495, 499–501, 503, 510, 607, 650
  - Serial dependence, 15–16, 156–160, 227, 492–493, 608
  - Shannon information, 417
  - Shape parameter, 104–107, 109–110, 114, 118–119, 164, 170, 223, 345
  - Shapiro–Wilk test for normality, 166
  - Sharpness, 370–371, 373, 417, 428–429, 437
  - Shifted gamma distribution, 110
  - Shrinkage, 268–274
  - Sieve bootstrap, 191
  - Signal detection theory, 412
  - Signed rank test, 171–172, 176–177, 178r
  - Significance testing. *See* Hypothesis test
  - Simple linear regression, 236–238, 253–254, 500, 679
  - Simple structure, 651–652, 658
  - Simultaneous confidence statements, 608–611
  - Sine function, 517–520
  - Single-linkage clustering, 724–725
  - Singular spectrum analysis (SSA), 661–665
    - for AR(2) series, 663
    - multichannel SSA (MSSA), 665
  - Singular systems analysis (SSA), 661–665
  - Singular value decomposition (SVD), 576–577, 684–685, 688–690
    - calculating CCA through, 684–685
    - maximum covariance analysis, 687–691
    - PCA via, 659–660
    - redundancy analysis, 692
  - Singular values, 576–577, 659–660, 688
  - Singular vectors, 322, 576–577, 659, 688
  - Skewness, 28, 47–50, 69, 94–95, 107–108, 189, 594
  - Skewness coefficient, 27–28, 57–58
  - Skill, 373–374, 383–385, 444–445
  - Skill score, 379–383, 392, 397–398, 475–476
  - Slack variable, 714f, 715
  - Smirnov test, 162–165
  - Softmax function, 292
  - Southern oscillation index (SOI), 51–52, 665
  - Spaghetti plot, 360–362, 362f
  - Spatial correlation, 181, 192, 197–199, 204, 358–359, 469, 479–480, 612–613
  - Spearman rank correlation, 59–61, 354–355

- Specificity, 379
  - Spectral analysis, 528–545
  - Spectral decomposition of a matrix, 571–573, 576, 647–648, 659–660, 673–676
  - Spectral estimates, 530, 540–545
  - Spectrum, 45, 529–533, 532*f*, 536, 542–543, 640, 661–665
  - Spread, 25–27, 46–48, 102, 174, 226, 337, 370–371, 622
  - Spread-skill relationship, 324, 337, 345
  - Square-root matrix, 574–575, 597
  - Stamp map, 360, 361*f*
  - Standard deviation, 26–27, 37–38, 50–52, 54–55, 57, 59–60, 78, 95–96, 95*f*, 98–101, 106, 120, 130, 150, 153–154, 159–160, 170–171, 182, 186, 187*f*, 188–189, 226, 231, 238, 244–245, 324, 340, 349, 360–362, 396, 398, 451, 471, 486, 511–512, 557, 557*f*, 582–583, 620, 671, 696, 727
  - Standard gamma distribution, 107–109, 341
  - Standard Gaussian distribution, 95–96, 103–104, 152, 154, 189, 330, 471, 593, 610
  - Standardized anomaly, 50–52, 96–97
  - Standardized precipitation index (SPI), 110
  - Star plot, 64
  - Stationarity
    - covariance, 485
    - strict, 485
  - Stationary probability, 491, 494
  - Statistical distance. *See* Mahalanobis distance
  - Statistical expectation, 88–91, 426, 467
  - Statistical forecasting
    - classical, 294–296
    - ensemble forecasting
      - choosing initial ensemble members, 321–323
      - effects of model errors, 325–327
      - ensemble mean and ensemble dispersion, 323–324
      - ensemble forecasts, 319–327
      - graphical display of ensemble forecast information, 359–366
      - statistical postprocessing, ensemble MOS, 327–328
      - stochastic dynamical systems in phase space, 313–317
  - linear regression, 235–253
    - analysis of variance table, 240–241
    - derived predictor variables in multiple regression, 254–256
    - distribution of residuals, 238–240
    - examining residuals, 245–250
    - goodness-of-fit measures, 241–243
    - multiple linear regression, 253–256
    - prediction intervals, 250–253
    - sampling distributions of regression coefficients, 243–245
    - simple linear regression, 236–238
    - nonlinear regression, 274–281
    - objective forecasts using traditional statistical methods, 294–304
  - predictor selection, 257–268
    - cross validation, 265–268
    - importance of careful predictor selection, 257–260
    - screening predictors, 260–263
    - stopping rules, 263–265
  - subjective probability forecasts, 304–310
    - assessing continuous distributions, 309–310
    - assessing discrete probabilities, 308–309
    - central credible interval forecasts, 306–307
    - nature of subjective forecasts, 304–305
    - subjective distribution, 305–306
  - Statistical inference, 15, 110, 143–206
  - Statistical learning, 287–288
  - Statistical postprocessing, 296–297, 342
  - Statistical significance, 195–197, 470, 474, 492, 499, 502, 544, 607, 611–614
  - Statistical simulation
    - with autoregressive model, 514
    - Box-Muller method for Gaussian random number generation, 136–137
    - multivariate, 596–601
    - nonuniform random number generation by inversion, 134–135
    - nonuniform random number generation by rejection, 135–136
    - simulating from mixture distributions and kernel density estimates, 137–138
    - uniform random number generators, 132–134
    - vector time series, 597–601
  - Statistics vs. parameters, 78
  - Stem-and-leaf display, 29–30
  - Stepwise regression, 260
  - Stippling method, 194, 197–199
  - Stochastic dynamic prediction, 318
  - Stochastic physics, 326–327
  - Stochastic process, 79–80, 487, 489
  - Stopping criterion, 263–264, 288, 732
  - Stopping rules
    - for clusters, 726–732
    - for forward selection, 263
  - Stratification, 295, 301, 721
  - Strictly proper scoring rule, 418–419, 427
  - Strict stationarity, 485
  - Structure function, 445
  - Student's *t*, 152
  - Subjective distribution, 229, 305–306, 307*f*, 394
  - Subjective forecast
    - assessing continuous distributions, 309–310
    - assessing discrete probabilities, 308–309
    - central credible interval forecasts, 306–307
    - nature of subjective forecasts, 304–305
    - subjective distribution, 305–306
  - Subjective interpretation of probability, 10–11, 271
  - Subsample relative frequency, 403–404
  - Subset, 11–13, 62, 97, 257–259, 266, 315–316, 374, 409*f*, 438, 553, 591–593, 651
  - Subspace, 554, 705–708
  - Successive subdivision, 309
  - Sufficiency, 182, 370, 416
  - Support, 36–37, 91, 113, 125, 199, 229
  - Support vector regression, 293
  - Support vector classifier, 714*f*, 715–716
  - Support vector machine, 714–716
  - Survivorship bias, 192–199, 201
  - Symmetric matrix, 561–562
    - eigenvalues and eigenvectors of, 573
    - square roots of, 574–576
  - Symmetry, 25–28, 44–48, 49*f*, 64*f*, 167–168, 593–594, 675
  - Synthesis equation, 524, 622, 630, 675–676
  - Synthesis formula, 619–620, 623–625, 639–640
  - Synthetic weather generator, 598
  - Systematic bias, 299*f*, 372
- ## T
- t* distribution, 152–154, 603, 609
  - t*-ratio, 244, 262, 280–281
  - t* test, 151–153, 155–156, 158, 171–192, 244, 502, 603
  - Talagrand diagram, 433–434
  - Taylor diagram, 396–397, 451–452, 452*f*
  - Teleconnection, 73–74, 653, 654*f*
  - Teleconnectivity, 73, 74*f*, 232
  - Temporal autocorrelation, 61, 73, 662
  - Tensor product, 564
  - Tercile, 25, 336, 407, 408*f*
  - Ternary reliability diagram, 424
  - Test level, 145–146, 193, 264, 544–545
  - Test statistic, 143–145, 147–150, 152–156, 159, 161, 163, 165–166, 169–173, 177, 179, 181–183, 186–188, 204–205, 249–250, 478, 492–493, 594, 606
  - Thinning, 228, 232
  - Theil-Sen regression, 285
  - Thom estimator, 106
  - Threat score (TS), 377, 381–382, 387, 387*f*
  - Tikhonov regularization, 270
  - Time domain, 132, 486–487
  - Time between effectively independent samples, 158–160, 228
  - Time-frequency analysis, 545–549
  - Time lag, 294, 297, 678–679
  - Time series
    - harmonic analysis, 517–527
    - cosine and sine functions, 517–518
    - estimation of amplitude and phase of single harmonic, 520–524
    - higher harmonics, 524–527

Time series (*Continued*)

representing simple time series with harmonic function, 518–520

spectral analysis, 528–545

- aliasing, 534–537
- computing spectra, 533–534
- harmonic functions as uncorrelated regression predictors, 528–529
- periodogram, or Fourier line spectrum, 529–533
- sampling properties of spectral estimates, 540–545
- theoretical spectra of autoregressive models, 537–540

time domain—discrete data, 487–499

- deciding among alternative orders of Markov chains, 498–499
- higher-order Markov chains, 497–498
- Markov chains, 487–488
- multiple-state Markov chains, 495–496
- some applications of two-state Markov chains, 494–495
- test for independence vs. first-order serial dependence, 492–493
- two-state, first-order Markov chains, 488–492

time domain—continuous data, 500–517

- AR(2) model, 505–509
- autoregressive-moving average models, 513
- first-order autoregression, 500–504
- higher-order autoregressions, 504–505
- order selection criteria, 509–510
- simulation and forecasting with continuous time-domain models, 513–517
- variance of time average, 510–512

T-M diagram, 385*f*

Toeplitz matrix, 661–662, 664

Training data, 247, 269–272, 280, 282–283, 286–288, 290, 301, 330, 333–334, 338–339, 344, 347–348, 353*f*, 467, 524, 589, 680, 695, 699*f*, 712*f*, 715, 721, 734

Training sample, 258–259, 330–331, 344, 691, 695–696, 710–711, 713, 715

Trajectory, 313–314, 316–317

Transition probability, 494, 496–498

Transpose (vector or matrix), 559, 562, 566–567, 569–570, 577

Tree diagram. *See* Dendrogram

Triangular kernel, 37

Tricube kernel, 282

Trimean, 26

Trimmed mean, 26

Trimmed variance, 27

True positive fraction, 379

True skill statistic (TSS), 380–381

Truncated Gaussian distribution, 102

Tukey's ladder, 45

Two-sample *t* test, 154–155, 158–159, 199, 202, 204–205, 603

Two-sided test, 147

Two-tailed test, 147, 150

Type I error, 146–147

Type II error, 146–147

## U

U-statistic, 173, 176, 373, 474

Uncertainty, 3–7, 99–100, 131, 155–156, 209–211, 214–216, 218–219, 222–227, 239–240, 250–251, 304, 306, 313, 319, 324, 327, 329, 342, 363, 404, 427, 432–433, 469, 502–503, 515, 607, 647

Unconditional bias, 372, 396, 398, 434–435, 451, 453

Unconditional distribution, 100, 100*f*, 239, 241, 371, 373, 399*f*

Underforecast, 378, 406–408, 435–436, 443*f*

Undersample, 192, 535

Uniform distribution, 112–113, 132, 135, 168, 217–218, 430, 434

Uniform random number generator, 132–134, 494

Union, 11–13, 146–147, 389

## V

Vague prior, 215, 233

Validity, 372, 645

Value score, 466–467

Variance, 27, 44, 54–55, 89, 119–120, 152, 240–241, 502, 635–638

Variance-covariance matrix, 555

Variance inflation factor (VIF), 158–160, 180–181, 228, 250, 252, 510–513

Variance reduction, 135

Variance-stabilizing transformation, 44

Variance of a time average, 156, 158, 510–512, 537

Varimax rotation, 651–652

Variogram score, 445

Vector, 559–561

- addition, 559
- column, 559, 561
- forecast, 419–424, 439–441
- multiplication, 560

- observation, 420
- outer product of, 564
- regression parameters, 568
- row, 554, 559–560, 563
- subtraction, 559

Venn diagram, 8–9, 9*f*, 11, 12*f*, 13–14

Verification data, 259, 266, 369, 371, 375–376, 379, 382, 386*f*, 394–395, 438, 473, 481*t*, 719*t*

Verification rank histogram, 433–437

Violin plot, 39–40, 39*f*

Von Mises distribution, 427*t*

## W

Waiting distribution, 83–85

Walker test, 544

Ward's minimum variance clustering, 725

Wavelets, 459, 548–549

Weak stationarity, 485

Weather generator, 598

Weibull distribution, 115–116, 116*f*, 229–230

Whitening transformation, 685–687, 693

White noise, 469–470, 508–509, 538, 538*f*, 543

White-noise variance, 502–503, 506, 509–510, 513, 516

Wilcoxon–Mann–Whitney test, 171–172, 174, 182–183

Wilcoxon signed-rank test, 171–172, 176–177, 178*t*

Windowed Fourier transform, 546–547

Wind power density, 226

Winkler's score, 432

Within-groups covariance matrix, 705, 726

## X

X–y plot, 53

## Y

Youden's index, 380–381

Yule–Kendall index, 28

Yule's Q, 381

Yule–Walker equations, 504–505, 509, 511–513, 599

## Z

Z-scores, 557–558

Z-transformation for the correlation coefficient, 49