

---

## Index

---

- $\nabla f(\theta, \gamma)$ , 359
- $\Delta(x_t)$ , 497
- $\Delta^{-1}(x_t)$ , 500
  
- absolute return, 101
- acceptance/rejection method, 317
- accredited investor, 18
- ACF (autocorrelation function), 8, 100, 515
  - AR model, 541–546
  - AR(1), 538
  - difference equation, 544
  - MA model, 531
  - partial (PACF), 516, 547
  - sample (SACF), 100, 519
- adjusted price, 45–48
  - dividend, 46, 134
  - effect on option strike, 64
  - return of capital, 48
  - stock split, 45
- adjusted  $R^2$ , 423, 468
- after hours, 10, 37
- agent-based model, 357
- aggregate** (R function), 510
- AIC (Akaike information criterion), 445, 468
- AICc, 469
- Akaike information criterion (AIC), 445, 468
- algo (programmed trading), 126
- alpha, 63, 126
- alternative hypothesis, 382
- alternative trading system (ATS), 16, 35, 129
- analysis of variance (ANOVA), 415–421, 440
- Anderson-Darling test, 461
  
- annualized log return, 30
- annualized simple return, 29
- annualized volatility, 107, 491
- ANOVA (analysis of variance), 415–421, 440
- apply** (R function), 150
- APT (arbitrage pricing theory), 44
- AR model, 534–571
  - AR(1), 536–539
    - ACF, 538
    - PACF, 539
  - AR(2)
    - ACF, 543
  - AR( $p$ ), 534–548
    - ACF difference equation, 546
    - order determination, 548
  - autocorrelation, 538, 543, 547
  - causality, 539
  - explosive, 538
  - partial autocorrelation function (PACF), 547
  - stationarity, 535
- arbitrage, 38
  - statistical, 600
- arbitrage pricing theory (APT), 44
- ARCH (GARCH *see*), 576
  - ARCH effect, 578
  - tests for, 579
- Archimedean copula, 266
- ARFIMA model, 554
- arima** (R function), 561
- ARIMA model, 553–571
  - order determination, 560, 563–565
- arima.sim** (R function), 555
- ARMA model, 549–571
  - causality, 551

- invertibility, 552
- order determination, 557, 560, 563–565
- ARMAX model, 575
- asymptotic distribution, 270, 271, 308
- asymptotic inference, 378, 393
- ATS (alternative trading system), 16, 35, 129
- augmented Dickey-Fuller test, 595–596, 598
- autocorrelation, 7, 511–513, 518
  - returns, 99–101
  - tests, 526, 580
- autocorrelation function (ACF), 8, 100, 511–513, 515
  - AR model, 541–546
  - AR(1), 538
  - difference equation, 544
  - MA model, 531
  - partial (PACF), 516, 547
  - sample (SACF), 100, 519
- autocovariance function, 515
  - partial, 516
  - sample, 519
- autoregressive model, 338, 492, 534–571
  - autocorrelation, 538, 543, 547
  - stationarity, 535
  - vector autoregressive model, 535, 590
- autoregressive operator, 535
  - multivariate, 536
- backshift operator,  $B(\cdot)$ , 495, 531, 535
  - inverse,  $B^{-1}(\cdot)$ , 496
  - multivariate, 497, 531, 536
- bagplot, 227, 236
- Bank of Japan (BOJ), 22
- Basel Committee, 19, 119
  - Basel Accords, 119
  - Basel II, 119
- basis point, 22
- Bayes information criterion (BIC), 468
- Bayesian model, 346, 386–393
  - conjugate prior, 388
  - MCMC, 389
- Bayesian network, 284
- bear market, 126
- Bernoulli distribution, 283, 450
- beta, 61–63, 126
- beta distribution, 291
- bias, 347
  - bootstrap estimate, 395
- bias-variance tradeoff, 218, 220, 370, 379, 402, 463, 465
- BIC (Bayes information criterion), 468
- bid/ask, 35
  - bid-ask spread, 37
- big data, 139
- bill, Treasury, 22
- binary regression, 451
- binary regressor, 415, 419
- binary tree, 281
- binomial distribution, 283, 450
- binomial tree, 281, 283
- Black Monday (October 19, 1987), 99, 159
- black swan, 132
- Black Thursday (October 24, 1929), 159
- Black Tuesday (October 29, 1929), 159
- Black-Scholes-Merton differential equation, 280
- BlackRock iShares, 56
- BOJ (Bank of Japan), 22
- bond, 13, 20
  - coupon rate, 26
  - coupon stripping, 28
  - coupon yield to maturity, 28
  - par value, 21
  - price, 28
  - yield to maturity, 28
  - zero-coupon, 28
- bond, Treasury, 22

- book value, 38
- bootstrap, 393–396
  - bias estimate, 395
  - boot (R package), 396
  - bootstrap (R package), 396
  - confidence interval, 395
  - variance estimate, 394
- Bowley’s skewness, 96
- box-and-whisker plot, 224, 225
- Box-Cox transformation, 345, 422, 579
- Box-Ljung test, 526, 580
- Box-Pierce test, 526
- boxplot, 110, 224, 225
  - bivariate, 227, 236
- Brownian motion, 279
- BUGS (software), 389
- Bulletin Board (OTC), 16
- burn-in for a stochastic process, 357
- Burr distributions, 324
- buyback, stock, 33, 38, 43, 44
  
- $C_p$ , 424, 468
- CAC-40 (index), 59, 166
- calendar spread, 74
- call option, *see* option
- call spread, 74
- candlestick graphic, 33, 182
- CAPE (Shiller P/E ratio), 43
- capital asset pricing model (CAPM), 44
- capital market line, 72
- capitalization, market, 32
- CAPM (capital asset pricing model), 44
- Case-Shiller Index, 168
- cash settlement, 64
- Cauchy distribution, 293
- causality, 528, 537, 539, 551
- CBOE (Chicago Board Options Exchange), 17, 108, 113
- CBOE SKEW Index, 113
- CBOE Volatility Index, 108–111, 166
- CCF (cross-correlation function), 517
  - sample, 519
  
- CDF (cumulative distribution function), 211, 249
  - complementary, 258
- CDF-skewing, 302
- CDO (collateralized debt obligation), 20
- Center for Research in Security Prices (CRSP), 53, 171
- centered data, 208–210, 408
- central limit theorem, 272
- Chambers, John, 140
- change of variables technique, 255
- characteristic polynomial, 503, 531, 533, 535, 539, 543, 545, 546
  - AR model, 535, 543
  - discriminant, 504
  - MA model, 531, 533
- chi-squared distribution, 289, 292, 313, 383
- chi-squared test, 457
- Chicago Board Options Exchange (CBOE), 17, 108, 113
- Chicago Mercantile Exchange (CME), 17
- Cholesky factor, 269, 310, 320
- class (R object), 143
- classification model, 337, 449
  - logistic regression, 449
- Clayton copula, 266
- cleaning data, 183–186
  - missing data, 185
- CME (Chicago Mercantile Exchange), 17
- CME Group, 17
- coefficient of multiple correlation, 417
- coefficient of multiple determination, 417
- coherent measure, 118, 257
- cointegration, 599–602
  - Engle-Granger test, 600
  - Johansen test, 600
  - Phillips-Ouliaris test, 600
- collateralized debt obligation (CDO), 20

- COMEX, 17
- comma separated file (CSV), 162–163
  - header, 163
- Commodity Futures Trading
  - Commission (CFTC), 18
- complementary cumulative distribution function (CCDF), 258
- composite hypothesis, 381
- composition of functions, 495
- compound distribution, 307
- compounding, 26
  - continuous, 28
- computational inference, 378, 393
- conditional distribution, 6, 260, 310, 513
- conditional heteroscedasticity, 575–583
- conditional value at risk, 119, 404
- confidence interval, 379, 448
  - basic bootstrap, 395
- conjugate prior, 388
- consistency, 348
- consistent estimator, 348
- constrained optimization, 359, 362
  - feasible set, 362
- Consumer Price Index (CPI), 168
- continuous compounding, 28
- continuous data, 206
- continuous distribution, 252, 285
- contour plot, 235
- Cook’s distance, 433, 440, 443
- copula, 264–267, 310, 321, 334, 350
  - Archimedean, 266
  - copula (R package), 350
  - definition, 265
  - fitting, 350
  - Gaussian, 266
  - normal, 266
- $\text{Cor}(\cdot, \cdot)$  (correlation, *see*), 4
- Cornish-Fisher expansion, 406
- corporate bond, 25
  - seasoned, 26
- correction, 25
- correlation, 83, 85, 208, 210, 261–263, 584–591
  - among assets in financial crises, 127, 130, 276
  - autocorrelation, 100
  - distribution, 408–409, 584–591
  - Kendall’s tau, 98, 399
  - Pearson correlation coefficient, 83
  - sample, 83, 408
  - Spearman’s rank correlation, 97, 399
  - spurious, 584
  - tail dependence, 127, 130, 276, 402
  - test, 409
- correlogram, 519
- coupon bond, 21
- coupon rate, 26
- coupon stripping, 28
- $\text{Cov}(\cdot, \cdot)$  (covariance, *see*), 4
- covariance, 82, 208, 210, 261–263
- covered call, 74
- Cox-Ross-Rubinstein (CRR) model, 282
- Cramér-von Mises test, 461
- CRAN (Comprehensive R Archive Network), 146
- credit rating agency, 19
- cross-correlation function (CCF), 513, 517
  - sample, 519
- cross-correlogram, 519, 522
- cross-covariance function, 513
  - sample, 519
- cross-validation, 463–467
  - $k$ -fold, 465
  - leave out one (LOOCV), 464
- crossover, 122
- CRSP (Center for Research in Security Prices), 53, 171
  - database, 53, 171
- cryptocurrency, 14, 38
- CSV file (comma separated file), 162–163

- header, 163
- `cumsum` (R function), 154
- cumulative distribution function (CDF), 211, 249
  - complementary, 258
- currency exchange rates, 167, 168
- CVaR, 119, 404
- D'Agostino-Pearson test, 461
- DAG (directed acyclic graph), 388
- daily log return, 30
- dark pool, 16, 35, 129
- data, 2
  - continuous, 206
  - copyright, 164
  - derived, 158
  - discrete, 206
  - observed, 158
  - repository, 163
- data cleansing, 183–186
  - data quality, 197, 481
  - missing data, 185
- data frame, 152–153, 161
  - `stringsAsFactors`, 152, 163
  - converting to `xts` object, 174
- data-generating process, 2, 311
  - model, 311
- data reduction, 341
- data structure, 160
- data wrangling and munging, 183
- datasets**
  - CanadaDollar/USDollar daily, 601
  - Dow Jones daily, 50, 55, 85, 98, 127
  - Dow Jones daily 1987-2017, 83
  - GLD daily, 85
  - INTC daily, 9, 34, 39, 51, 70, 76, 77, 79, 80, 85, 86, 88, 98, 124, 127
  - INTC daily 2017, 61
  - INTC monthly, 9
  - INTC price adjustments, 47, 48
  - INTC weekly, 9
  - Nasdaq daily, 50, 55, 85, 86, 98, 127
  - Nasdaq daily 1987-2017, 83
  - S&P 500 daily, 50, 53, 55, 61, 83, 85, 89, 90, 92, 98–101, 108, 127, 569, 579
  - S&P 500 monthly, 93, 102, 105
  - S&P 500 Total Return daily, 53, 56
  - S&P 500 weekly, 93, 105
  - SKEW daily, 108
  - SPY daily, 56
  - T-Bill, 3-month monthly, 22
  - USDollar/Euro daily, 14, 601
  - VIX daily, 108, 110
  - yield curves, 24, 25
- date data, 156, 174
  - in data frames, 161
  - in `xts` objects, 174
  - ISO 8601, 157, 176
  - POSIX, 157, 176
- DAX (index), 59, 166
- death cross, 123
- deleted residual, 429, 464
- delta (Greek), 65, 75, 137
- density plot, 79, 226
  - bivariate, 227
- derivative, 13, 63–66
- descriptive statistics, 94, 205
- detrending, 505, 508
- deviance, 373, 451, 468
- deviance residual, 374, 451
- df (degrees of freedom), 289
- DFBETAS, 433, 440, 443
- DFFITs, 433, 440, 443
- Dickey-Fuller test, 595
  - augmented, 595–596
- `diff` (R function), 154, 174, 176, 499
  - `cumsum`, 154
  - `diffinv`, 154
  - on `xts` object, 155, 174
- difference equation, 501–505, 545, 546
- difference operator,  $\Delta(\cdot)$  or  $(1 - B)$ , 497

- fractional differencing, 498
- differential scaling, 302
- diffusion process, 279, 280, 490
- dimension reduction, 341
- directed acyclic graph (DAG), 388
- discount rate, 22
- discounted cash flow, 42
- discounted value, 28
- discrete data, 206
- discrete distribution, 248, 283
- discrete uniform distribution, 213, 323, 393
- dispersion, 207
- distribution family**
  - Bernoulli, 283, 450
  - beta, 286, 313
  - binomial, 283, 284, 313, 450
  - Burr distributions, 324
  - Cauchy, 286, 293, 313
  - chi-squared, 289, 292, 313, 383
  - Dagum, 301
  - Dirichlet, 287, 294, 313
  - discrete uniform, 213, 323, 393
  - double exponential, 286, 292, 299, 313
  - elliptical, 263, 295
  - exponential, 286, 292, 313
  - extreme value, 309
  - F, 289, 313, 383
  - gamma, 286, 313
  - Gaussian, 87, 286, 287, 312, 322
  - generalized error distribution, 299–300
  - generalized extreme value, 309
  - generalized lambda, 299, 317
  - generalized Laplace, 300
  - generalized Pareto, 301
  - geometric, 284, 313
  - Gumbel, 309
  - heavy-tailed, 273–277
  - hypergeometric, 284, 313
  - Johnson distributions, 324
  - Laplace, 286, 292, 299, 313
  - location-scale, 229, 298
  - logistic, 286, 313
  - lognormal, 286, 288, 313
  - Lorentz, 286
  - multinomial, 284, 294, 313
  - multivariate normal, 287, 295, 313
  - multivariate t, 295, 313
  - negative binomial, 284, 313
  - normal, 87, 286, 287, 312, 313, 322
    - skewed, 302
  - outlier-generating, 277
  - Pareto, 286, 292, 313
  - Pearson distributions, 324
  - Poisson, 284, 285, 312, 313, 322, 352
  - power law, 292
  - R functions for, 144, 313
  - skewed, 302
  - skewed generalized error distribution, 300
  - stable, 298, 313
  - t, 289, 296–298, 313, 383
    - location-scale t, 297, 298
    - noncentral t, 290
    - skewed, 302
    - “standardized” t, 298
  - uniform, 252, 286, 313, 315
  - variance gamma, 300
  - Weibull, 286, 292, 313
- diversification, 66, 114
- dividend, 45, 171
  - adjustment to prices, 46, 134
  - `getDividends` (R function), 171
  - reinvestment program, 45
  - stock, 45
- DJIA (Dow Jones Industrial Average), 49, 166, 479
  - divisor, 49
  - historical values, 51, 55
- double exponential distribution, 286, 292, 299, 313
- Dow Jones Industrial Average (DJIA), 49, 166, 479
  - divisor, 49
  - historical values, 51, 55

- Dow Jones Transportation Average, 50, 166
- Dow Jones Utility Average, 50, 166
- dp1yr (R package), 153, 497
- Durbin-Watson test, 434, 443, 526, 571
- DVaR, 119, 276
- $E(\cdot)$  (expectation, *see*), 4
- earnings, 41
- earnings per share (EPS), 41
  - forward earnings, 41
  - trailing earnings, 41
- EBITDA, 43
- ECB (European Central Bank), 22
- ECDF (empirical cumulative distribution function), 211–217, 276, 350
- folded, 214
- EDA (exploratory data analysis), 1, 205
- EDGAR, 170
- effective fed fund rate, 22, 25, 436, 438
- efficient frontier, 71
- efficient market, 120
- Efficient Market Hypothesis, 120
- 80/20 rule, 293
- elastic net, 370, 376
- Elliott-Rothenberg-Stock test, 598
- elliptical distribution, 263, 295
- emerging markets, 60
- empirical cumulative distribution function (ECDF), 211–217, 276, 350
- folded, 214
- empirical quantile, 95, 215–217
- endogenous/exogenous variable, 337
- Engle-Granger test, 600
- enterprise value (EV), 33
- equity asset, 13
- error (in a model), 246, 534
- compared to residual, 415
- error correction model (ECM), 601
- estimation, 311, 349–354
- least squares, 351, 363–370
  - maximum likelihood, 352–354, 371–375
  - method of moments, 349
  - Monte Carlo, 355–357
  - tuned inference, 354, 401, 402
- ETF (exchange-traded fund), 18, 56–59
- index, 56
  - inverse, 57–59
  - leveraged, 57–59
  - sector, 56
  - short, 59
  - tracking, 56
  - ultra, 57–59
- ETN (exchange-traded note), 18, 111
- Euromarket, 19, 22
- European Central Bank (ECB), 22
- EV (enterprise value), 33
- EV/EBITDA, 43
- evolutionary time series, 515
- EWI (S&P 500 Equal Weight Index), 50
- EWMA (exponentially weighted moving average), 39, 123
- model, 534
- exact inference, 393
- exceedance distribution, 232, 275
- q-q plot, 232
- exceedance threshold, 232, 275
- excess kurtosis, 82, 88, 251, 254
- excess return, 32, 60–63, 73
- alpha, 63
  - mixing log returns and interest rates, 32
- exchange, 16
- exchange-traded fund (ETF), 18, 56–59
- index, 56
  - inverse, 57–59
  - leveraged, 57–59
  - sector, 56
  - short, 59
  - ultra, 57–59
- exchange-traded note (ETN), 18, 111

- expectation, 250, 254
  - conditional, 6, 260, 513
  - linear operator, 256
- expected shortfall, 119, 404
- exploratory data analysis, 1, 205
- exponential distribution, 292
- exponential smoothing, 41, 534
- exponential tail, 274
- exponentially weighted moving
  - average (EWMA), 39, 123
  - model, 534
- externally studentized residual, 432, 433
- extreme value theory, 308
- F distribution, 289, 313, 383
- F test, 421
- face value, 21
- factor model, 44, 413
- fair market value, 32, 120
- FARIMA model, 554
- fat-tailed distribution, 273
- fed discount rate, 22
- fed fund rate, 22, 25, 438
  - effective, 22, 25, 436, 438
- Federal Open Market Committee (FOMC), 21, 25, 438
- filter, 40, 172, 495–500
- financial instrument, 13
- FINRA (Financial Industry Regulatory Authority), 18
- fiscal policy, 19
- Fisher scoring, 375
- Fitch Ratings, 19
- fitting matrix, 420
- fixed-income asset, 13
- FOMC (Federal Open Market Committee), 21, 25, 438
- `foreach` (R function), 396, 465
- forecasting in time series, 527, 561, 565–566
  - forecast horizon, 527
  - forecast origin, 527
- foreign exchange, 13, 167, 168
- forex, 13, 170
- formula in R, 454
- forward earnings, 41
- forward rate, 27
- forward selection, 425, 444
- forward shift operator,  $B^{-1}(\cdot)$ , 496
- fractional differencing, 498
- Frank copula, 267
- FRED (Federal Reserve Economic Data), 167–168
  - missing data, 189
- front-running, 129
- FTSE-100 (index), 59, 166
- fundamental analysis, 35
- future value, 27, 28
- futures, 13
- FX, 13, 170
- GAAP (Generally Accepted Accounting Principles), 43
- gamma distribution, 291
- GARCH model, 308, 580–583
- Gauss-Markov theorem, 419
- Gauss-Newton method, 367
- Gaussian copula, 266
- Gaussian distribution, 87, 287, 312, 322
- Gaussian white noise, 488
- GED (generalized error distribution), 299–300
- generalized error distribution, 299–300
  - skewed, 300
- generalized extreme value
  - distribution, 309
- generalized inverse (of a function), 250
- generalized lambda distribution, 299, 317
- generalized least squares, 369, 376
- generalized linear model, 376, 451
- generalized Pareto distribution, 301
- Generally Accepted Accounting Principles (GAAP), 43
- generating random numbers, 314–321
  - quasirandom numbers, 316, 321



- Gentleman, Robert, 140
- geometric Brownian motion, 279
- geometric distribution, 284
- geometric random walk, 492
- geometric series, convergence of, 500
- `getSymbols` (R function), 165, 168, 171–172
  - missing data, 165, 186
- `ggfortify` (R package), 180
- `ggplot2` (R package), 146, 156, 187
- GICS (Global Industry Classification Standard), 56
- Gini index, 214, 293
- GLD (ETF), 56, 86, 403
- Global Industry Classification Standard (GICS), 56
- GLS (generalized least squares), 369, 376
- golden cross, 123
- Gompertz model, 366
- goodness-of-fit test, 456
  - Anderson-Darling, 461
  - chi-squared, 457
  - Cramér-von Mises, 461
  - D’Agostino-Pearson, 461
  - Jarque-Bera, 461
  - Kolmogorov-Smirnov, 460
  - Lilliefors, 461
  - Shapiro-Francia, 461
  - Shapiro-Wilk, 461
  - test for normality, 461
- Google, 164
- government statistics, 164
- gradient,  $\nabla f(\theta, \gamma)$ , 359
- graphics, 221–238
  - bagplot, 227, 236
  - bivariate boxplot, 227, 236
  - bivariate density plot, 87, 227, 236
  - boxplot, 224
  - color, 156
  - density plot, 79, 226
  - ECDF plot, 212
  - folded ECDF plot, 214
  - half-normal plot, 233
  - histogram, 76, 115, 222
  - layer, 235, 237
  - log density plot, 90, 91, 227
  - mountain plot, 214
  - OHLC dataset, 34, 182
  - parametric line plot, 236
  - q-q plot, 91–93, 229–233, 276, 306
  - quantile-quantile plot, 91–93, 229–233, 276, 306
  - R, 234–238
  - scatterplot, 235
  - scatterplot matrix, 86, 235
  - starburst plot, 227
  - three-dimensional, 235
  - `xts` object, 179, 182
- Greeks, 65
- Group of Ten, 19
- Gumbel copula, 266
- Gumbel distribution, 309
- half-normal plot, 233
- Halton sequence, 316
- Hang Seng (index), 59, 166
- Hannan-Quinn information criterion, 471, 583
- hard-to-borrow, 66
- hat matrix, 420
- hazard function, 257
- head and shoulders, 124
- heavy-tailed distribution, 89, 273–277
- hedge, 73–75
  - cross hedge, 73
  - dynamic, 75
  - hedge ratio, 74
- hedge fund, 17, 75
- Hessian,  $H_f(\theta, x)$ , 359
- heteroscedasticity, 99, 414
  - conditional, 576
- high-frequency trading, 12, 132
- Hill estimator, 400–402
  - Hill plot, 402, 471
  - interval of stability, 402
- histogram, 76, 115, 217–218, 222

- Sturges's rule, 218
- historical volatility, 103
  - compared with implied volatility, 108
- `Hmisc` (R package), 146
- Holt-Winters method, 41, 534
- hypergeometric distribution, 284
- hypothesis testing, 311, 381–384
  - multiple hypotheses, 384, 525
  - one-sided test, 384
  - power, 382
  - simultaneous tests, 384, 525
  - type I and II errors, 382
- $I(d)$  process, 517, 533
- ICO (initial coin offering), 38
- identifiability, 532
- Ihaka, Ross, 140
- iid (independently and identically distributed), 5, 260
- IMA( $d, q$ ) model, 533
- implied volatility, 107
  - compared with historical volatility, 108
- income, 45
- independent random variables, 259
- index mutual fund, 18
- index, market, *see also* stock index, 49–55, 413, 479
- inference, 347, 523
  - asymptotic inference, 378, 525
  - computational inference, 378, 393
  - exact inference, 378
  - tuned inference, 218, 220, 354, 370, 379, 401, 402, 463
- information criteria, 471, 583
- initial coin offering (ICO), 38
- initial public offering (IPO), 37
- innovation, 534
- inputting data from the internet, 163, 165
  - `getSymbols`, 171
- insider information, 128
- institutional investor, 15, 125
- instrumental variable, 337
- integration (in time series), 500, 516
  - integrated ARMA model (ARIMA), 553
  - integrated moving average model, 533
  - integrated white noise, 516
  - integration operator,  $S(x_t)$ , 500
- interest, 20–28
  - compounded, 26
    - continuously, 28
- interest rate, 21–28
  - annualized, 21, 29
  - daily, 29
  - fed discount rate, 22
  - LIBOR, 22
  - prime rate, 22
  - risk-free interest rate, 22
  - structure, 24
- internally standardized residual, 433
- international markets, 59, 166
- interquartile range (IQR), 96, 208, 216, 225, 397
  - relationship to standard deviation, 96, 208
- intrinsic value, 64
- inverse CDF method, 317
- inverse ETF, 59, 111
- invertibility, 532, 534, 552
- investor, 15
  - accredited, 18
- IPO (initial public offering), 37
- IQR (interquartile range), 96, 208, 216, 225, 397
  - relationship to standard deviation, 96, 208
- iShares, 56
- ISLR (R package), 451
- ISO 8601 date format, 157, 176
- Jacobian (of a transformation), 268
- JAGS (software), 389
- Jarque-Bera test, 461
- Joe copula, 267
- Johansen test, 600

- Johnson  $S_U$  distributions, 324
- joint distribution, 258
- joint stationarity, 517–518
- Journal of Statistical Software*, 187
- jump, 77, 285
- Kaggle, 164
- Kendall's tau, 98, 399
- kernel density estimation, 79, 219–221, 226, 236
  - multivariate, 220, 227
- Knightian uncertainty, 127
- Kolmogorov-Smirnov test, 460
- KPSS (Kwiatkowski-Phillips-Schmidt-Shin) test, 507, 598
- KS test, 460
- kurtosis, 81, 251, 254
  - excess, 82, 88
- Kwiatkowski-Phillips-Schmidt-Shin (KPSS) test, 507, 598
- $L_1$  and  $L_2$  norms, 351
- $L_1$  estimation, 351, 376, 397, 429
- lag, 495, 497
  - backshift operator,  $B(\cdot)$ , 495
- lagging indicator, 514, 517
- Lagrange multiplier, 362
- Lagrangian function, 362
- Laplace distribution, 286, 292, 299, 313
- lasso regression, 369, 427, 447
- law of one price, 38
- leading indicator, 25, 514, 517
- least absolute values ( $L_1$ ) estimation, 351, 376, 397, 429
- least squares estimation, 351, 363–370, 375
  - Gauss-Markov theorem, 419
  - Gauss-Newton method, 367
  - generalized least squares, 369, 376
  - nonlinear least squares, 366–368, 376
  - nonnegative least squares, 370, 376
  - penalized, 369, 427, 447, 466
  - prediction, 448
  - properties (algebraic), 364
  - properties (statistical), 415, 418, 430
  - regularized, 369, 427, 447, 466
  - shrunk, 369, 427, 447, 466
  - weighted least squares, 369
- leave-out-one deletion, 433, 440
  - cross-validation, 464
- leptokurtic, 88
- leverage in regression, 429
  - Cook's distance, 433, 440
  - DFBETAS, 433, 440
  - DFFITS, 433, 440
- leveraged instrument, 13, 57–59
- LIBOR, 22
- likelihood function, 352–354, 371–375, 398
  - likelihood ratio, 385
- Lilliefors test, 461
- limit order, 35
- linear function, 495
- linear model, 375, 413–448, 493, 599
  - intercept (or not), 365, 419
- linear operator, 256, 495–505
  - expectation, 256
- linear process, 489
  - one-sided, 489, 529, 537
- linear transformation, 209, 229, 320, 495
- link function, 450
- Ljung-Box test, 526, 580
- location, 207
- location-scale family of distributions, 229, 298
- log density plot, 91, 227
- log return, 29–31, 491, 500
  - annualized, 30
- log-likelihood function, 373
- logarithmic scale, 54
- logistic function, 450
- logistic regression, 449
  - prediction, 453
- logit function, 450

- lognormal distribution, 288, 492
- lognormal geometric random walk, 492
- London Stock Exchange (LSE), 16
- long position, 65
- Lorenz curve, 214, 293
- LSE (London Stock Exchange), 16
- MA model, 529–571
  - ACF, 530
  - autocorrelation, 530
  - identifiability, 532
  - invertibility, 532
  - MA(1) invertibility, 533
  - order determination, 530
  - partial autocorrelation function (PACF), 532
- MACD, 123
- MAD (median absolute deviation), 96, 208, 397
  - relationship to standard deviation, 96, 208
- Mallows's  $C_p$ , 424, 468
- margin, 17
- margin call, 17
- marginal distribution, 259
- market, 16, 38
  - secondary, 38
- market capitalization, 32
- market index, *see also* stock index, 49–55
- market microstructure, 12
- market model, 60
- market order, 36
- market portfolio, 60
- market price, 34
- market risk, 113
- market timing, 121, 200
- market value, 32
- Markov chain, 279, 511
- Markov chain Monte Carlo (MCMC), 389
- Markovian property, 511
- married put, 74
- martingale, 279, 511
- MASS (R package), 146
- master limited partnership (MLP), 13
- MATLAB<sup>®</sup>, 146, 172
- maximum likelihood estimator (MLE), 352–354, 371–376
  - Fisher scoring, 375
- MCMC, 389
- mean, 3, 80, 250, 254, 272
  - of a random variable, 250
  - trimmed mean, 95, 397
  - Winsorized mean, 95, 397
- mean reversion, 122
- mean squared error (MSE), 348, 369, 463
- median, 95, 216
  - median( $\cdot$ ), 95
- median absolute deviation (MAD), 96, 208, 397
  - relationship to standard deviation, 96, 208
- mesokurtic, 88
- method of moments, 349
- methods (R package), 143
- Microsoft Excel<sup>®</sup>, 147, 162
- minimum variance portfolio, 69
- missing data, 142, 185–186, 189
- mixture distribution, 303, 307
- MLE (maximum likelihood estimator), 352–354, 371–376
  - Fisher scoring, 375
- MLP (master limited partnership), 13
- model, 114, 336–346
  - additive error, 339
  - autoregressive, 338, 492
  - Bayesian, 346, 386–393
  - classification, 337, 449
  - error, 246, 534
    - compared to residual, 415
  - factor, 44, 413
  - linear, 375, 413–448
  - model averaging, 472
  - multimodel inference, 472

- multiplicative error, 339
  - of relationship, 408
  - probability distribution, 245–323
  - regression, 337, 413
  - selection, 463–469, 472
    - time series, 560
  - smoothing, 346
  - time series, 346, 487–604
    - state-space, 494
- model averaging, 472
- model specification in R, 454
  - ARIMA, 562
- moment, 80, 81, 207, 208, 254
  - method of moments, 349
  - population, 254
  - sample, 80, 81, 207, 208
- momentum, 44
- monetary policy, 19
- moneyiness, 64
- Monte Carlo estimate, 355–357
- Monte Carlo simulation, 314–321, 355–357
- Moody’s Investors Service, 19, 26
  - Moody’s Seasoned Corporate Bonds, 26, 168, 436
- mortgage rates, 168
- mountain plot, 214
- moving average, 39–41, 123, 134, 172
  - convergence-divergence, 123
  - death cross, 123
  - golden cross, 123
  - MACD, 123
  - moving average model, *see*, 529
  - of white noise, 489
    - autocorrelation, 512
  - R functions, 172
  - rolling operation, 172
- moving average model (MA), 529–571
  - autocorrelation, 530
  - identifiability, 532
  - invertibility, 532
- moving average operator, 531
  - multivariate, 531
- MSE, mean squared error, 348, 369, 463
- MSE, mean squared residual (MSR), 364, 417, 463
- multicollinearity, 426
- multimodel inference, 472
- multivariate backshift operator,  $B(\cdot)$ , 497, 531, 536
- multivariate data, 82–87, 218, 258–269, 309
  - time series, 517
- multivariate distribution, 258–269, 294–295, 309
  - variance-covariance matrix, 5, 260, 264
- multivariate regression, 339, 344
- multivariate time series, 517
- munging data, 183
- municipal bond, 26
- mutual fund, 17
  - index, 18
- $N(\mu, \sigma^2)$ , 283
- NA (“Not Available”), 142, 185
- naked option, 66
- Nasdaq Composite Index, 50, 166
  - historical values, 51, 55
- Nasdaq National Market, 16
- national best bid and offer (NBBO), 35, 129
- NBBO (national best bid and offer), 35, 129
- Nelder-Mead method, 361, 375
- New York Mercantile Exchange (NYMEX), 17
- New York Stock Exchange (NYSE), 16
- Newton’s method, 360, 375
- Nikkei-225 (index), 59, 166
- noncentral t distribution, 290
- nonlinear least squares, 366–368, 376
- nonnegative definite matrix, 262
- nonnegative least squares, 370, 376
- nonparametric probability density estimation, 217–221, 226

- multivariate, 220, 227
- nonsynchronous trading, 11
- norm, 351, 363, 370
- normal copula, 266
- normal distribution, 87, 287, 312, 322
  - multivariate, 295
  - skewed, 302
  - test for, 461
  - visual assessment, 231, 457
- normal equations, 364
- note, Treasury, 22
- null hypothesis, 381
- numeraire, 2
- NYMEX (New York Mercantile Exchange), 17
- NYSE (New York Stock Exchange), 16
- NYSE Composite, 53
- Oanda, 170
- object, R, 143–144, 173–182
  - class, 143
  - mode, 144
  - type, 144
  - `xts`, 173–181
- OHLC dataset, 33, 166, 180–182
- OLS (ordinary least squares), 369
- omnibus hypothesis test, 383
- optimization, 358–376, 413
  - constrained, 359, 362, 413
  - direct method, 360
  - Gauss-Newton method, 367
  - iterative method, 360–362
  - least squares, 363–370
  - Levenberg-Marquardt method, 368
  - maximum likelihood, 352–354, 371–375
  - Nelder-Mead method, 361
  - Newton’s method, 360
  - quasi-Newton method, 360, 361, 368
    - Gauss-Newton method, 367
    - Levenberg-Marquardt method, 368
  - simulated annealing, 362
  - steepest descent, 360
- option, 63–66, 74–75, 171
  - adjustment of strike, 64
  - American, 64
  - calendar spread, 74
  - call spread, 74
  - cash settlement, 64
  - chain, 171
  - European, 64
  - `getOptionChain` (R function), 171
  - intrinsic value, 64
  - married put, 74
  - moneyness, 64
  - pricing, 65, 279, 280
  - put spread, 74
  - straddle, 75
  - strangle, 75
  - strap, 75
  - strike, 64
    - adjustment, 64
  - strip, 75
  - time value, 64
- order statistic, 95, 211, 269–270, 308
  - ECDF, 211
  - notation, 95
  - PDF, 270
  - quantile, 95
- orthogonal regression, 414, 471, 602
- oscillator, 124
- OTC Bulletin Board, 16
- outlier, 77, 93, 210, 277, 428
- outlier-generating distribution, 277
- overfitting, 417, 423, 428, 468, 556
  - regression, 423, 428
  - time series, 556
- p-value, 225, 239, 381, 385, 470
- PACF (partial autocorrelation function), 425, 516, 532, 547–548
  - AR model, 547
  - MA model, 532
  - sample, 548

- pairs trading, 75–76, 600
- par value, 21
- parameter space, 248
- parametric line plot, 236
- Pareto distribution, 292
  - generalized, 301
- Pareto tail, 275, 399
- partial autocorrelation function (PACF), 425, 516, 532, 547–548
  - AR model, 547
  - MA model, 532
  - sample, 548
- partial autocovariance function, 516
- partial F test, 425, 440
- partial regression, 424–426
- partitioning the variation, 306, 340, 415, 425
- pattern (technical)
  - crossover, 122
  - death cross, 123
  - golden cross, 123
  - head and shoulders, 124
  - MACD, 123
  - moving average
    - convergence-divergence, 123
    - death cross, 123
    - golden cross, 123
  - oscillator, 124
  - relative strength index, 124
  - resistance line, 126
  - stochastic oscillator, 124
  - support line, 123, 126
- payout ratio, 42
- PCA (principal components analysis), 409–413, 479
- PDF (probability density function), 217, 248, 253
  - estimation, 217–221, 399
  - kernel, 253
- PE, 41–43
  - CAPE (Shiller P/E ratio), 43
  - enterprise value to EBITDA, 43
  - forward, 42
  - trailing, 42
- Pearson correlation coefficient, 83
- Pearson distributions, 324
- percentile, 92
- $\Phi(\cdot)$ , 288
- $\phi(\cdot)$ , 288
- Phillips-Ouliaris test, 600
- Phillips-Perron test, 597, 598
- pivotal quantity, 380
- platykurtic, 88
- point estimate, 347
  - consistent, 348
  - unbiased, 347
- Poisson distribution, 285, 312, 322
  - jump, 285
- polynomial regression, 449
- polynomial tail, 275, 399
- portfolio, 60, 66–76, 407
  - efficient frontier, 71
  - efficient portfolio, 71
  - expected shortfall, 407
  - minimum variance portfolio, 69
  - optimal portfolio, 69, 371
  - rebalancing, 67, 407
  - return, 68, 407
  - risk measure, 407
  - Sharpe ratio, 72
  - tangency portfolio, 72
  - Treynor ratio, 72
  - VaR, 407
- portmanteau test, 526
- positive definite matrix, 262
- POSIX, 156, 174
- power curve, 383
- power law, 301
  - distribution, 292, 301
- power of test, 382
- power series, convergence of, 500
- `pracma` (R package), 146
- precision, 207, 388
- `predict` (R function), 448, 453, 565
- prediction, 348, 385, 448, 453
- prediction interval, 385, 448
- present value, 26–28
- price discovery, 37
- price-to-earnings ratio, 41–43

- CAPE (Shiller P/E ratio), 43
- enterprise value to EBITDA, 43
- forward, 42
- trailing, 42
- price-to-sales ratio (PSR), 41
- prime rate, 22
- principal components, 409–413, 479
  - constrained, 413
- probability density function (PDF),
  - 217, 248, 253
  - estimation, 217–221, 399
  - kernel, 253
- probability distribution, 247–314
  - nondegenerate, 247
  - R functions for, 144, 313
  - support, 248, 253
- product distribution, 310
- programmed trading, 126
- projection matrix, 420
- ProShares, 56
- pseudorandom number, 315, 321
- PSR (price-to-sales ratio), 41
- pump-and-dump, 128
- put option, *see* option
- put spread, 74
- PVaR, 115
- q-q plot, 91–93, 229–233, 276, 306
  - half-normal plot, 233
  - of exceedance distribution, 232
  - of mixtures, 306
  - two samples, 233
- quadratic programming, 69, 370, 413
- qualitative regressor, 415, 419
- Quandl, 169, 170
- quantile, 92, 95, 144, 215, 249, 253,
  - 259, 312, 404
  - Cornish-Fisher expansion, 406
  - estimation, 215, 356, 404–407
  - quantile function, 249, 250, 265,
    - 277
  - sample or empirical, 95,
    - 215–217, 405
- quantile-quantile plot, 91–93,
  - 229–233, 276, 306
  - of mixtures, 306
- quantmod** (R package), 165, 171, 187
- quartile, 96, 216
- quasirandom number, 316, 321
- R (software)**, 139–190
  - \$ (list extractor), 150, 152, 154,
    - 174
  - Bayesian methods, 389
  - caveats, 187
  - comma separated file (CSV), 163
    - header, 163
  - conditional, 143
  - data frame, 152–153
    - `stringsAsFactors`, 152, 163
    - merging, 153, 178
    - tibble, 153
  - data structure, 160
  - date data, 156, 174
  - environment, 145, 146
  - formula, 454, 455
  - function, 144
  - generic function, 143
  - `ggplot2`, 146, 187
  - graphics, 155, 234–238
    - color, 156
    - layout, 238
    - three-dimensional, 235
  - GUI app, 147
  - `help`, 141
    - `vignette`, 142
  - inputting data, 161, 171
    - from the internet, 163
  - linear algebra, 150
  - list, 149
  - log return, 154, 174, 176
  - logical operator, 142
  - logistic regression, 451–453
  - long vector, 148
  - matrix, 150
  - merging datasets, 153, 178, 186
  - `methods`, 143
  - missing value, 142, 185
  - model formula, 454, 455
  - object, 143



- optimization, 375–376
- package, 140, 145–147, 187
  - caveats, 147
  - lowest common denominator, 188
- parallel processing, 396, 465
- pracma, 146
- probability distributions, 311
- quadratic programming, 371
- quantmod, 171, 187
- random number generation, 314, 321–323
  - quasirandom number, 316, 321
  - randtoolbox, 316, 321
- regression, 375, 435–448
- ridge regression, 370
- rjags, 389
- RStudio, 140, 147
- Shiny, 147
- simulation, 314, 321–323, 555
  - of time series, 555
- task views, 141
- three-dimensional graphics, 235
- tibble, 153
- time series, 172–181, 496, 499, 561
- TTR, 172
- vector, 147
- vignette, 142
- web technologies, 172
- $R^2$ , 416
  - adjusted, 423, 468
- random number generation, 314–321
  - quasirandom number, 316, 321
- random variable, 2–5, 247–282
  - notation, 3
  - transformation, 251, 254, 267
- random walk, 132, 490–492, 505
  - autocovariance, 513
  - cumsum (R), 491
  - linear combination, 493, 599
  - multivariate, 493
  - with drift, 492, 505
- Random Walk Hypothesis, 120, 491
- rank transformation, 97, 398
- real estate investment trust (REIT), 13
- recession, 25
- regression, 351, 354, 363, 413–448, 571–575
  - ANOVA, 415
  - ARMA errors, 435, 571–575
  - assumptions, 414
  - binary, 451
  - Cook’s distance, 433, 440
  - deleted residual, 429, 464
  - DFBETAS, 433, 440
  - DFFITS, 433, 440
  - F test, 421
  - hat matrix, 420
  - intercept, 365, 419
  - leave-out-one deletion, 433, 440
  - leverage, 429
    - Cook’s distance, 433, 440
    - DFBETAS, 433, 440
    - DFFITS, 433, 440
  - logistic, 449
  - outlier, 428
  - partial F test, 425
  - polynomial, 449
  - qualitative regressor, 415, 419
  - residual, 424
  - residuals, 415, 417, 428–435, 440, 571
    - compared to errors, 415
    - deleted, 429, 464
    - serial correlation, 434, 571–575
    - studentized, 431–433
  - seemingly unrelated regressions, 339, 344
  - spurious, 584
  - t test, 421
  - variable selection, 425, 444
    - forward selection, 425, 444
  - variance inflation factor (VIF), 427, 446
- regression model, 337, 413
  - intercept (or not), 365, 419

- regulatory agency, 43
- REIT (real estate investment trust), 13
- relative strength index, 124
- resampling, 393–396
- residual, 415, 417, 424, 428–435, 440, 571
  - compared to error, 415
  - deleted residual, 429, 464
  - deviance, 374, 451
  - externally studentized, 432, 433
  - internally standardized, 433
  - serial correlation, 434, 571
  - standardized, 433
  - studentized, 431–433
  - studentized deleted, 433
- resistance line, 126
- return, 29–32, 67–73, 99–102
  - absolute, 101
  - aggregating over assets, 31
  - aggregating over time, 30
  - annualized, 29, 30
  - autocorrelations, 99–101
  - excess, 32, 60–63, 73
    - mixing log returns and interest rates, 32
  - expected, 68
  - log return, 29–31, 491, 500
    - annualized, 30
  - portfolio, 68, 407
  - R functions for, 155, 500
  - simple, 29, 500
    - annualized, 29
  - variance, 69
- return of capital, 45
  - adjustment to prices, 48
- revenue, 41
- reverse split, 45
- ridge regression, 369, 376, 427, 428, 447, 466
  - ridge trace, 369, 428, 447
- risk, 15, 66, 67, 113, 127
  - diversifiable, 114
  - market, 113
  - measure, 114, 118, 257, 407
  - coherent, 118, 257, 407
  - DVaR, 119
  - expected shortfall, 119
  - PVaR, 115
  - simple returns, 114, 404
  - tail loss, 119
  - VaR, 114
  - risk-free asset, 15, 22, 71
    - interest rate, 22
  - risky asset, 15
  - simple returns, 114, 404
  - specific, 114
  - systematic, 113
  - uncertainty, 127
- risk management, 113–120, 404
- risk neutral, 44
- robust method, 97, 396–399
- R-squared, 416
  - adjusted, 423, 468
- RSS, 363, 416
- RStudio, 140, 147
- Russell 2000, 53, 166
- Ryan, Jeffrey A., 164
- $S(\cdot)$  (summation operator), 500
- S&P 500 (Standard & Poor’s 500 Index), 50, 166
  - historical values, 51, 54, 55
  - total return, 53, 54, 166
- S&P 500 Equal Weight Index (EWI), 50
- SACF (sample ACF), 100, 519
- sample autocorrelation function (ACF or SACF), 100, 519
- sample autocovariance function, 519
- sample CCF (cross-correlation function), 519
- sample correlation, 83, 85, 208, 210, 408
- sample covariance, 82, 208, 210, 408
- sample cross-correlation function (CCF), 519
- sample cross-covariance function, 519
- sample mean, 80, 272
- sample moment, 80

- sample quantile, 95, 215–217
- sample standard deviation, 80
  - historical volatility, 104
- sample variance, 80, 289, 408
- SARIMA model, 555
- scale, 80, 207
- scatterplot, 235
  - three-dimensional, 235
- scatterplot matrix, 86, 235
- Schmidt-Phillips test, 598
- Schwarz information criterion (SIC), 468
- SDS (ETF), 58, 59
- seasonal adjustment, 508
- seasonal effects, 200
  - stock prices, 200
- seasonality, 41, 121, 508, 554, 555
  - ARIMA model, 555
  - ARMA model, 554
  - seasonal adjustment, 508
  - stock prices, 121
- seasoned security, 26, 38
- secondary market, 38
- sector, market, 55, 56
  - GICS, 56
- Securities and Exchange Commission (SEC), 17, 170
- seemingly unrelated regressions, 339, 344
- SH (ETF), 58
- shape, frequency distribution, 80–82, 207, 229–232
- Shapiro-Francia test, 461
- Shapiro-Wilk test, 461
- Sharpe ratio, 72
- Shibata's information criterion, 471, 583
- Shiller P/E ratio (CAPE), 43
- Shiny (R GUI apps), 145, 147
- short ETF, 59
- short interest, 44, 65
- short position, 65–66, 120
- short-and-distort, 128
- shrunk estimator, 325, 369
- SIC (Schwarz information criterion), 468
- significance, statistical, 381
  - p-value, 225, 239, 381, 470
- simple hypothesis, 381
- simple return, 29, 500
  - annualized, 29
- simulated annealing, 375
- simulating random numbers, 314–321
  - quasirandom number, 316, 321
- simulation, 355–357
  - in R, 321–323
  - of time series, 555
- simultaneous equations model, 339
- skew, 113
- SKEW Index, 113
- SkewDex<sup>®</sup>, 113
- skewed distribution
  - CDF-skewing, 302
  - differential scaling, 302
  - Lorenz curve, 214
- skewed generalized error
  - distribution, 300
- skewed normal distribution, 302
- skewed t distribution, 302
- skewness, 81, 251, 254
  - Bowley's, 96
  - Gini index, 214
  - Pearson's, 96
  - quantile-based measure, 96
- smoothing, 41, 346, 534
  - moving average, 41
- smoothing matrix, 221, 420
- smoothing parameter, 41, 218, 220, 346, 534
- Sobol' sequence, 316
- SPDR, 56
- Spearman's rank correlation, 97, 399
- specific risk, 114
- split, 45
  - adjustment to prices, 45
- spread, 207
- spreadsheet, 162
- spurious regression, 584
- SPY (ETF), 56, 57

- SSE (sum of squares of residuals), 363, 416
- SSE Composite (index), 59, 166
- SSO (ETF), 57
- stable distribution, 298, 313
- stale price, 37
- Standard & Poor's, 19
- Standard & Poor's 500 Index (S&P 500), 50, 166
  - historical values, 51, 54, 55
  - total return, 53, 54, 166
- standard deviation, 4, 80, 251, 254
  - historical volatility, 104
- standard error, 377, 567
- standardized data, 208–210, 408, 409, 411
- standardized residual, 433
- “standardized” t distribution, 298
- starburst plot, 227
- state space model, 494
- stationarity, 7, 100, 278, 514–518
  - joint, 517–518
  - strict, 514
  - weak, 514
- statistical arbitrage, 600
- statistical inference, 347, 376–379, 381
  - asymptotic inference, 378
  - computational inference, 378
  - confidence interval, 379
  - consistency, 348
  - exact inference, 378
  - hypothesis testing, 381–384
  - point estimate, 347
    - consistent, 348
    - unbiased, 347
  - tuned inference, 218, 220, 354, 370, 379, 401, 402, 463
- statistical model, 336–346
- statistical significance, 381
  - p-value, 225, 239, 381, 470
- Statistical Software, Journal of*, 187
- steepest descent, 360
- stochastic difference equation, 502
- stochastic differential equation (SDE), 279, 502
- stochastic oscillator, 124
- stochastic process, 278–282, 487
- stochastic volatility, 99, 414, 576
- stock, 13
- stock buyback, 33, 38, 43, 44
- stock dividend, 45
- stock index, 49–55, 413, 479
  - adjusted value, 53
  - capitalization weighted, 50
  - equally weighted, 50
  - market-value weighted, 50
  - price weighted, 49, 50
  - principal components, 413, 479
- stock split, 45
  - adjustment to indexes, 49
  - adjustment to option strikes, 64
  - adjustment to prices, 45
- stop order, 36
- stop-limit order, 36
- straddle, 75
- strangle, 75
- strap, 75
- strict stationarity, 514
- strike price, 63
- strip, 75
- stripping of coupons, 28
- STRIPS, 28
- Student's t distribution, *see also* t distribution, 289
- studentization, 289, 431
- studentized deleted residual, 433
- studentized residual, 431–433
- Sturges's rule, 218
- sum of squares, 341
- summation operator,  $S(x_t)$ , 500
- support, 248, 253
  - continuous distribution, 253
  - discrete distribution, 248
- support line, 123, 126
- survival function, 257
- SVXY (inverse VIX), 111
- swap, 13
- systematic risk, 113

- t distribution, 289, 296–298, 313, 383
  - location-scale, 295, 297, 298, 313
  - multivariate, 295, 313
  - noncentral, 290
  - skewed, 290, 302
  - “standardized”, 298
  - tail index, 296
- t test, 382, 421
- tail dependence, 127, 276, 402–403
  - coefficient, 277, 403
  - copula, 277
  - tail dependence function, 276
- tail index, 275, 399–402
  - estimation, 399–402, 471
    - Hill estimator, 400
  - exponential, 274
  - generalized Pareto, 402
  - of t distribution, 296
  - Pareto, 275, 399
  - polynomial, 275, 399
- tail loss, 119, 404
- tail risk index, 113
- tail threshold exceedance, 232, 275, 301
- tail weight, 274
- TailDex<sup>®</sup>, 113
- tangency portfolio, 72, 73
- technical analysis, 35, 122
- 10-K form, 169
- term structure, 24
- test statistic, 381
- three-dimensional graphics, 235
- threshold exceedance, 232, 275, 301
- tidyverse** (suite of R packages), 146
- time series, 5–12, 222, 346, 487–604
  - bivariate, 519
  - evolutionary, 515
  - model
    - AR, 534
    - ARCH, 576
    - ARIMA, 553
    - ARMA, 549
    - EWMA, 534
    - GARCH, 580
    - Gaussian white noise, 488
    - geometric random walk, 492
    - IMA, 533
    - linear process, 489
    - MA, 529
    - moving average of white noise, 489
    - multivariate random walk, 493
    - multivariate white noise, 493
    - random walk, 490
    - random walk with drift, 492
    - state-space, 494
    - white noise, 488
  - multivariate, 517
  - stationary, 514–518
  - xts** (R object), 173–181
- time value, 64
- Tokyo Stock Exchange (TSE), 16
- trader, 15
- trading friction, 37
- trading strategy, 73–76
  - crossoveer, 122
  - dynamic hedge, 75
  - hedge, 73–75
  - options, 74
  - pairs trading, 75–76, 600
- trailing earnings, 41
- tranch, 20
- transformation, 227, 251, 254, 267, 310, 315, 320, 344, 412, 421
  - Box-Cox, 345, 422, 579
  - linear, 209, 229, 268, 320, 421, 495
  - of a random variable, 251, 254, 267, 315, 320
    - change of variables technique, 255
  - of data, 208–210, 344, 412, 421
    - centering, 208, 209, 421
    - covariance transformation, 210
    - standardizing, 208, 209, 421
  - one-to-one, 267
  - power, 345
  - variance-stabilizing, 344, 422, 579

- Treasuries, 22–25
  - rates, 168
- Treasury Bill, 22
- trend (in time series), 505
  - test, 507
- Treynor ratio, 72
- trimmed mean, 95, 397
- TSE (Tokyo Stock Exchange), 16
- TTR (R package), 172
- tuned inference, 218, 220, 354–355, 370, 379, 401, 402, 463
- tuning parameter, 218, 220, 354, 370, 379, 401, 402
  
- U(0, 1), 252, 283, 286, 315
- Ulrich, Joshua, 172
- unbiased estimator, 347
- uncertainty, 127
- underlying, 63
- uniform distribution, 252, 286, 315
- unit circle, 504, 533, 544, 546
- unit root, 503, 505, 592–599
  - tests, 594–599
- UPRO (ETF), 57, 58
- US regulatory agency, 17–18
  - Commodity Futures Trading Commission (CFTC), 18
  - Securities and Exchange Commission (SEC), 17, 170
- US Treasuries, 22–25
  
- V(·) (variance, *see*), 4
- value at risk (VaR), 114–120, 232, 273, 276, 301, 309, 404–407
  - coherence, 118, 257, 407
  - CVaR, 119, 404
  - DVaR, 119, 276
  - estimation, 404–407
  - portfolio, 407
  - simple returns, 114, 404
- VAR (see vector autoregressive model), 535
- VAR(p), 535, 590
- variable selection, 425, 444
- variance, 4, 80, 251, 254, 289
  - bootstrap estimate, 394
  - of a linear combination, 4, 268–269
  - of a random variable, 4, 251
  - partitioning, 306, 340, 415
  - sample, 80, 289
- variance inflation factor (VIF), 427, 446
- variance-bias tradeoff, *see* bias-variance tradeoff
- variance-covariance matrix, 4, 5, 83, 260, 264, 269, 408
  - of a linear transformation, 268
  - sample, 83, 408
- variance-stabilizing transformation, 344, 422, 579
- vector autoregressive model, 339, 535, 590
- vector error correction model (VECM), 601
- VIX, 108–111, 166
  - historical values, 109–111
- volatility, 98–113
  - annualized, 107, 491
  - definition, 103
  - effect of length of time intervals, 93, 105
  - effect of outliers, 104
  - historical compared with implied, 108
  - historical volatility, 103
  - implied, 107
  - square-root rule, 107
  - stochastic, 99, 414
- volatility index, 108–112
- volatility smile, 112
- VolDex<sup>®</sup>, 112
  
- Watanabe-Akaike information criterion, 471
- weekend effect, 10
- Weibull distribution, 292
- weighted least squares, 369
- white noise, 488
  - linear combination, 493, 599

- multivariate, 493
- with drift, 488
- Wickham, Hadley, 146
- window size, 221
- Winsorized mean, 95, 397
- World Bank Open Data, 170
- wrangling data, 183
  
- XIV (inverse VIX), 111
- xts* object, 173–181
  - changing time period, 177
  - converting to data frame, 173
  - graphics, 179
  - merging, 178, 186
  - subsetting, 176
  
- Yahoo Finance, 164–166, 190
  - dividends, 182
  - international markets, 166
  - missing data, 186, 189
- yield curve, 24, 169
  - Nelson-Siegel, 24
  - spline, 24
  - Svensson, 24
- yield to maturity, 28
- Yule-Walker equations, 542, 558
  
- Zivot-Andrews test, 598