

Index

- #N/A (Excel code), 89
- #NUM (Excel error message), 201
 - in RANDOM and NORMALRANDOM functions, 223
- aggregation, 44–50
- aggregation problem. *See* ecological correlation
- algebra of expectations, 264–277
 - and constants, 271
 - and Gauss–Markov theorem, 366–374
 - and Gauss–Markov theorem in univariate case, 347–350, 351–352
 - for random walk, 632
 - and rules for computing expectations, 367, 369
 - and SE, 379, 380–381
 - and variance and SD of a random variable, 271–276
- AlgebraOfExpectations.xls, 266–274
- alternative hypothesis, 257–258, 430 *See also* hypothesis tests
 - in cigarette demand example, 444
 - one-tailed, 258
 - two-tailed, 257–258
- annual rates and quarterly data, 617
- AnnualGDP.xls, 609, 652–657
- Anscombe.xls, 127
- Anscombe’s data, 126
- Answers folder, 2
- AR(1) model. *See also* AR(1) process
 - data generation process, 562–563
 - step-by-step implementation, 563–564
- AR(1) process,
 - graphs of error terms in, 564
- array functions in Excel, 88
- artificial sample. *See* bootstrap
- assumptions of classical econometric model. *See also* box models, comparing
 - average of errors is zero, 327–328, 330
 - errors independent of X ’s, 320, 323–324, 331
 - homoskedastic errors, 320–321, 331
 - lack of autocorrelation, 331
 - linear in parameters with additive error term, 330
 - no multicollinearity, 331
 - normal errors not assumed, 331
 - X ’s fixed in repeated sampling, 328–329, 330
- assumptions of measurement box model, 287, 289
- astronomers and measurement box model, 282
- asymptotic. *See* large-sample analysis
- AutoCorr.xls, 563–599
- autocorrelation, 296. *See also* AR(1) model;
 - Durbin–Watson test; estimated ρ test;
 - feasible generalized least squares; generalized least squares and autocorrelation
 - in cigarette example, 559–560
 - comparing auto- and unautocorrelated processes, 565
 - consequences for OLS, 567
 - defined, 561
 - detection via eyeballing residuals, 576
 - effect of autocorrelation in the X ’s, 572
 - effect on P -values, 572
 - estimated SEs incorrect, 570–571
 - first-order autocorrelation coefficient, 562
 - two basic strategies for handling, 560
 - versus autoregression, 562
- autoregression,
 - defined, 562
 - versus autocorrelation, 562
- autoregressive processes. *See also* random walk
 - algebra of expectations, 632
 - asymptotic correlations, 629–630
 - defined, 562
 - explosive, 630–631
 - initial conditions, 630
 - order of, 562
 - stable, 631
 - and stationarity, 628–633
 - and strong dependence, 634
 - unit root processes as special case, 634
- auxiliary regression,
 - and Breusch–Pagan test, 528
 - and computer use example, 500
 - and omitted variable rule, 197

Cambridge University Press

0521843197 - Introductory Econometrics: Using Monte Carlo Simulation with Microsoft Excel®

Humberto Barreto and Frank M. Howland

Index

[More information](#)

762

Index

- Basmann, Robert, 742, 747
- best linear unbiased estimator (BLUE). *See also* Gauss–Markov theorem; generalized least squares
- confusion with minimum variance estimator, 375
- Monte Carlo evidence for univariate case, 345–346
- bias. *See also* biased estimator; consistency; omitted variable bias
- in lagged dependent variable models, 633
- and measurement box model, 295
- biased estimator. *See also* simultaneous equations model (SEM)
- defined, 343
- example of, 344
- and omitted variable bias, 494
- binary response variable. *See* dummy dependent variable
- bootstrap. *See also* Bootstrap (Bootstrap.xls)
- add-in
- compared to Monte Carlo simulation, 710
- confidence interval, 717
- defined, 709–710
- free throw example, 711–712
- Monte Carlo evidence, 724
- paired XY, 716
- paired XY versus residuals, 726
- R^2 with residuals bootstrap, 723–724
- sampling distribution, 712
- versus conventional methods, 727
- Bootstrap (Bootstrap.xls) add-in,
- demonstration of bootstrap, 722
- installation, 718
- practice example, 718–720
- BootstrapR2.xls, 722–726
- box models, 239–277. *See also* coin flip box model, polling box model
- benefits of, 240
- comparing, 290, 321, 332
- conceptual areas associated with, 243–244, 293
- and Monte Carlo simulation, 244–249
- principles for constructing, 240–241
- recipe for building, 240–241
- SD of the box, 244
- BoxModel.xls, 244–249, 277
- BPSampDist.xls, 530–533
- Breusch–Pagan (BP) test. *See also* heteroskedastic errors
- bivariate example, 529
- used to detect heteroskedasticity, 531
- intuition behind, 528
- real world example of, 551
- sampling distribution of, 530
- Bricklin, Daniel, 4
- Brownstone, David, 712, 727, 732
- campaign contribution example, 669–694
- CampCont.xls, 670–679
- case resampling. *See* bootstrap
- Cauchy distribution, 267, 288
- central limit theorem, 258, 308, 312
- applied to OLS estimators, 416–417, 418, 419
- and free throw example, 229
- and test statistics, 435
- CEM. *See* classical econometric model
- chance error,
- in campaign contribution example, 670
- in classical econometric model, 320
- and difference between two sample averages, 307–308
- in digitalis example, 714
- in free throw example, 228
- in measurement box model, 285, 291
- in Raid example, 668
- in weight of adult US citizens example, 256
- versus estimation error, 401
- chi-square distribution,
- convergence to normal distribution, 460
- degrees of freedom of, 459, 460
- expected value and SD, 460
- and F -distribution, 461
- and normal distribution, 458–460
- recipe for creating, 455, 458–459, 460
- and SSR, 419, 460
- ChiSquareDist.xls, 459–460
- Cig.xls, 11–15
- cigarette example. *See also* autocorrelation; controlled experiment; observational study
- and aggregation, 45
- autocorrelation in, 559–560
- bivariate regression in, 21
- and classical econometric model, 443
- confounding in, 22
- data limitations, 24–25
- historical data, 11
- and hypothesis tests, 443–449
- hypothetical data, 15
- hypothetical demand, 13
- multivariate regression, 24
- state data, 18, 19
- CigDataInference.xls, 445
- classical econometric model (CEM), 319. *See also* assumptions of classical econometric model
- bivariate example, 353
- experimental setup for, 322
- implemented in LinestRandomVariables.xls, 412
- and measurement box model, 282, 301
- and multiple regression, 329
- origins, 334
- philosophical interpretation of omitted variables in, 330
- pictorial representation, 321, 329
- sampling in, 324
- univariate example, 336
- in words, 328–333
- cluster samples, 254. *See also* presidential heights example
- and CPS, 311–312
- CMF. *See* conditional mean function
- coal mine safety example,
- and dummy variables, 614–617
- and elasticities, 616
- and hypothesis tests, 616
- and practical importance of estimates, 616
- CoalMining.xls, 615–616
- coefficient of determination. *See* R^2

Cambridge University Press

0521843197 - Introductory Econometrics: Using Monte Carlo Simulation with Microsoft Excel®

Humberto Barreto and Frank M. Howland

Index

[More information](#)

Index

763

- cohort,
 - defined, 304
- coin flip box model, 242–250, 286
 - applied to free throw example, 258
- coin flip example,
 - and box model, 242–249
 - and hypothesis tests, 431
- computer use example,
 - and omitted variable bias, 498–501, 502
 - and random X 's data generation process, 502
- ComputerUse1997.xls, 499–501
- conditional mean function (CMF). *See also* regression line
 - defined, 69
 - as first compression of the data, 96–98
 - versus regression line, 101
- conditional median function, 69
- conditional probability function. *See also* linear probability model; nonlinear least squares
 - campaign contribution example, 675
 - defined, 673
 - ways to estimate, 675
- confidence interval game, 422
 - and uniform distribution, 422–424
- confidence level in, 423
- confidence intervals, 421–430. *See also* confidence interval game
 - and cigarette demand example, 447
 - confidence level in, 424
 - defined, 422
 - and hypothesis tests, 421–422
 - proper interpretation of, 450
 - for sample slope, 425–430
- ConfidenceIntervals.xls, 422–429
- confounding, 43. *See also* omitted variable bias; simultaneity (simultaneous equations) bias
 - in bivariate regression, 177
 - in cigarette example, 22
 - in education and wages two box example, 313–314
 - in mortgage discrimination example, 697, 706
 - and wage discrimination, 62, 206
- consistency, 262–264
 - of estimated rho test, 579–580
 - in free throw example, 234
 - lack of, in simultaneous equations model (SEM), 736
 - of Root Mean Squared Error (RMSE), 394
 - and weak dependence, 634
- Consistency.xls, 262–264, 277
- consistent estimator, example of, 262–264
- constants,
 - and expectations operator, 271
- controlled experiment. *See also* omitted variable bias
 - and cigarette example, 16–17
 - and omitted variable bias, 492
- correction factor (CF). *See also* sampling without replacement, correction factor
- correlated slope estimates. *See* joint sampling distribution
- correlated X 's,
 - and joint confidence region, 484–487
 - and joint sampling distribution, 481
 - and SEs of slopes, 482
- CorrelatedEstimates.xls, 478, 488
- correlation, 35–51
 - association vs. causation and, 43–44
 - common misunderstandings about, 40
 - computing value of, 36–37
 - and heteroskedasticity, 41
 - and nonlinearity, 41–42
 - as poor descriptive statistic (shortcomings of), 40–43
 - and regression line, 102–103
 - and trend DGP, 627
- Correlation.xls, 34–50
- Costa Rica example, 37–38
- Covariance. *See also* correlation
 - in trend DGP, 626–627
- CPIMZM.xls, 601
- CPS. *See* Current Population Survey
- CPS.doc, 48
- CPS90ExpWorkers.xls, 315
- CPS90Workers.xls, 311–313
- crime rate. *See also* simultaneous equations model (SEM)
 - Bureau of Justice data, 732
- cross-sectional data,
 - defined, 18
- cross tab (cross tabulation),
 - defined, 59
 - income as a function of education, 60
 - law school admissions grid, 70
- cumulative distribution function,
 - defined, 684, 687
- Current Population Survey (CPS),
 - and cluster samples, 311–312
 - and computer use example, 499
 - and earnings function data, 475
 - and education and wages example, 310–311
 - and example of ecological correlation, 48–50
 - and IndianaFTWorkers.xls, 53
 - and months in survey for households, 67
 - and unemployment data, 618
 - and wage discrimination example, 205
- Data Analysis: Regression add-in,
 - and hypothesis tests, 447, 451
 - finding OLS regression line using, 85–86
 - zero intercept bug, 86
- data generation process (DGP),
 - and box model, 240–277
 - defined, 238
- DDV Gauss Newton (DDVGN.xls) add-in. *See also* DDV (DDV.xls) add-in
 - described, 690
- degrees of freedom,
 - and adjusting SD of residuals, 393
 - defined, 393
 - and t -distribution, 439–441
- descriptive statistics,
 - univariate, 34–35
- DGP. *See* data generation process
- dice example,
 - and expectations operator, 270–274, 276

- dichotomous choice variable. *See* dummy dependent variable
- die rolls,
 - and box model, 239–241
- difference between sample averages,
 - as estimate of difference between population averages, 304
 - as random variable, 307–308
- differencing data,
 - and random walk model, 653
 - and strong dependence, 636–637
- digitalis example,
 - and bootstrapping, 713–720
- distributed lag model,
 - defined, 559
- distribution of error terms,
 - and OLS regression statistics, 413–420
- disturbances. *See* error terms
- DoubleCompression.xls, 96–99
- double-logarithmic functional form,
 - and cigarette demand example, 448–449
- draws from a box. *See* box model
- dummy dependent variable. *See also* dummy dependent variable (DDV) model
 - defined, 663
 - examples, 665
 - versus continuous variable, 664
- Dummy Dependent Variable (DDV) add-in. *See also* DDV Gauss Newton (DDVGN.xla) add-in
 - installation, 690
 - and mortgage discrimination example, 701
 - for NLLS estimation, 690
 - predicted probability table, 693
- dummy dependent variable (DDV) model. *See also* linear probability model; nonlinear least squares; threshold
 - defined, 665
 - intuition in campaign contribution example, 669–671
 - intuition in Raid example, 666–669
 - and role of threshold, 669
 - visual box model for, 672
- dummy variables,
 - base case, 201
 - defined, 199
 - how to create, in Excel, 199
 - interaction terms and, 209, 475–476, 477
 - as intercept shifter, 207
 - and perfect multicollinearity, 200
 - poorly named, 199
 - properties, 202
 - and seasonal adjustment, 621–622
 - and several categories, 200
 - in savings behavior example, 454
 - synonyms for “dummy”, 198
 - in time series models, 613–614
- Durbin, James, 558, 582
- Durbin–Watson statistic,
 - computed, 582
 - formula for, 582
 - Durbin–Watson test,
 - implementing formula, 582
 - inability to distinguish between types of autocorrelation, 587
 - and lagged dependent variables, 649–650, 651, 652
 - and printed tables, 585
 - Monte Carlo of sampling distribution, 583
 - via P Value Calculator add-in, 585–586
- earnings and education example, 99–101
 - conditional mean function, 99
- earnings function,
 - defined, 148
 - empirical estimates, 154
 - and heteroskedasticity, 549–554
 - and internal rate of return (IRR), 150
 - interpreting coefficient estimates, 154
 - justification for semilog, 149–151
 - present value, 150
- EastNorthCentralFTWorkers.xls, 65–69, 99–101, 131
- EcolCorr.xls, 46–47
- EcolCorrCPS.xls, 48
- ecological correlation, 44–50
 - defined, 44
 - hypothetical example, 45–47
 - real example, 48–50
- ecological fallacy,
 - defined, 44
- education and wages example,
 - and hypothesis tests, 450–451
 - linear model, 451
 - semilog model, 451
- education and wages two box example, 310–314
- Efron, Bradley, 709, 726, 728
- elasticity,
 - advantages of, 157
 - and cigarette demand example, 447, 448–449
 - computing, 15, 156–157
 - defined, 15, 156
 - and functional form, 158
 - local phenomenon, 157–158
- ellipse, confidence. *See* joint confidence region
- empirical histogram,
 - of Monte Carlo experiment, 248–249
 - of sample, 246–247
- endogeneity,
 - types of bias, 741
- enforcement spending. *See also* simultaneous equations model (SEM)
 - Bureau of Justice data, 732
- error terms. *See also* forecast error, distribution of error terms
 - composite, 323
 - in Hooke’s Law example, 298
 - interpretation of, 443, 454
 - interpretation of in CEM, 320
 - and luck, 320
 - and omitted variables, 320, 328
 - SD of, estimated by RMSE, 393
 - SD of, estimated by SD of residuals, 392
 - versus residuals, 341, 384

- estimated ρ test,
 - consistency of, 579–580
 - defined, 578–579
 - and lagged dependent variables, 649–651, 652
 - Monte Carlo evidence, 579–580
 - properties, 579
 - and weak dependence, 634
- estimated SE, 307, 327, 386, 437
 - and spurious regression, 613
- estimated slope as random variable,
 - in Hooke's Law example, 299–300
- estimates, 281–282
- EstimatingSDErrors.xls, 392–396
- estimation error versus chance error, 401
- estimator,
 - defined, 107–108, 336
- estimators. *See also* least absolute deviation (LAD) estimator, least median of squares (LMS) estimator, linear estimator, nonlinear least squares (NLLS), ordinary least squares (OLS)
 - 0.9Estimator, 339
 - diminishing weights, 339
 - random linear, 340
 - sample average, 338
- Excel,
 - add-ins, 2
 - array functions, 88
 - as teaching tool, 90
 - calculation, 223–224
 - Excel no-intercept bug, 457
 - file naming conventions, 1
 - foreign language, 5
 - history, 4
 - Macintosh compatibility, 5, 8
 - minimum requirements, 5
 - Q&A sheets, 2
 - security patches, 5
 - security settings, 5–7
 - trademark, 4
 - troubleshooting, 8
 - Update Links, 224
 - versions, 4–5
 - Visual Basic error message, 9
- Excel codes,
 - #N/A, 89
- Excel functions,
 - NORMALRANDOM, 223
 - NORMINV, 221
 - RAND, 221
 - RANDOM, 221
 - ROUND, 293, 323
- exercise and sports,
 - regression analyses of, 320, 334
- expectations operator, 265, 267–268, 272–275, 276. *See also* algebra of expectations
 - formula for, 267
- expected value (EV), 243
 - computing in box model, 244
 - defined, 267
 - in measurement box model, 289
 - as probability weighted sum, 267
 - of sample average, 293
 - of sums and averages of random variables, 274–276
- ExpGrowthModel.xls, 607
- explained sum of squares,
 - defined, 124
- explained variation. *See* regression sum of squares
- exponential distribution,
 - for errors, 467
- exponential trend. *See* trends, exponential
- ExponentialDist.xls, 467
- extrapolation,
 - dangers of, 26–27, 477
- F -distribution, 455, 461
 - and chi-square distribution, 461
 - degrees of freedom of, 461
 - recipe for creating, 461
- FDist.xls, 461
- FDistEarningsFn.xls, 476–477
- FDistGalileo.xls, 464–467
- FDistFoodStamps.xls, 469–475, 487–488
- feasible generalized least squares (FGLS), 596, 599
 - defined, 596
 - Monte Carlo evidence, 597–599
 - versus GLS, 596
 - versus OLS, 599
 - Prais–Winsten versus Cochrane–Orcutt, 596
- Female.xls, 199–201, 205–211
- first difference. *See* differencing data
- Fisher, Ronald, 455, 713–715
- food stamps example,
 - and F -statistics, 489
 - and hypothesis tests, 468–475
 - restricted and unrestricted models in, 473–474
- forecast error. *See also* standard error of the forecast error
 - components of, 400
 - defined, 398
 - estimation error versus chance error, 401
 - Monte Carlo simulation of, 385–407
- forecasted Y ,
 - defined, 398
 - Monte Carlo simulation, 385–407
- forecasting. *See also* interval forecast, point forecast, standard error of the forecast; standard error of the forecast error
 - and autocorrelation, 652–658
 - and cigarette example, 25–27
 - depends on model of DGP, 652–657
 - standard errors associated with, 398–408
- ForecastingGDP.xls, 657
- Frankston, Bob, 4
- FreeThrowAutoCorr.xls, 602
- free throw example, 225–229
 - autocorrelation, 561
 - autocorrelation exercise, 602
 - bootstrap example, 711–712
 - coin-flip box model applied to, 258

Cambridge University Press

0521843197 - Introductory Econometrics: Using Monte Carlo Simulation with Microsoft Excel®

Humberto Barreto and Frank M. Howland

Index

[More information](#)

766

free throw example (*cont.*)
 and hypothesis tests, 145–147, 257–260
 Monte Carlo simulation, 226–229
 streaks in free throws, 230–231
 Freedman, et al. *Statistics*, 240, 278
F-statistic,
 computing, 470
 defined, 462
 and error distributions, 467, 487
 and multicollinearity, 482
 and *t*-statistics, 475
F-test, 462. *See also* Breusch-Pagan (BP) test
 and estimated ρ test, 649
 implementing, 465–466, 467–468, 470–472
 and SSR, 462, 463
 in total personal income example, 476, 477
F-tests versus *t*-tests, 473–475
 FuncFormCatalog.xls, 162
 functional forms,
 catalog of, 161
 comparison of residuals, 147
 defined, 141
 double-log, 145, 646–647
 elasticity, 158
 nonlinear, 142–143
 reciprocal, 145
 semilog, 148, 154–155, 451
 for trends in time series models, 605–611
 fundamental question of inferential statistics,
 311
 and hypothesis tests, 257
 Galileo,
 and measurement box model, 302
 Galileo example,
 and *F*-tests, 463–467
 linear functional form, 139–140
 nonlinear functional form, 141–143
 regression through the origin, 144
 residual plot, 141, 143
 restricted and unrestricted models in, 463
 Galileo.xls, 139–144
 Galton, Francis, 51. *See also* regression line, there
 are two regression lines
 and regression, 103–104, 107, 136
 names regression, 104
 regression effect, 95
 GaltonHeights.xls,
 Gauss, Carl Friedrich, 281, 291. *See also* Gaussian
 errors
 and invention of ordinary least squares, 91
 and measurement box model, 286–287, 302
 Gaussian errors, 329
 Gauss–Markov theorem. *See also* best linear
 unbiased estimator (BLUE)
 and algebra of expectations, 366–374
 and distribution of the errors, 363–366
 bivariate case, Monte Carlo, 362
 confusion with min SSR, 375
 defined, 335, 346
 history, 376
 and Monte Carlo simulation, 366
 proof of, in univariate case, 347–350, 351–352
 proof, bivariate model, with 10 observations,
 366–374

Index

GaussMarkovBivariate.xls, 353–366, 372–374
 GaussMarkovUnivariate.xls, 337–340, 341–342,
 343–352, 370
 generalized least squares (GLS) and
 autocorrelation. *See also* feasible generalized
 least squares (FGLS)
 example, 591
 first observation, 590
 Monte Carlo evidence, 593
 transformation, 588
 versus OLS, 593
 generalized least squares (GLS) and
 heteroskedasticity,
 GLS SE is consistent, 547
 information requirement, 547–549
 intuition, 542
 Monte Carlo evidence, 547
 transformation, 544–546
 Gini coefficient. *See also* simultaneous equations
 model (SEM)
 country comparison, 727, 732
 in simultaneous equations example, 727,
 732
 Goldberger, Arthur,
 and random X 's data generation process, 507
 Gosset, William Sealy, 438–439, 442
 and Monte Carlo simulation, 451–452
 graph of averages, 96, 100–101. *See also*
 conditional mean function
 as conditional probability function, 675
 and double compression of the data, 97–99
 and PivotTable, 98
 and regression line, 98–99
 and two regression lines, 105–106
 Griliches, Zvi, 507
 Haavelmo, Tryvge, 316
 heating oil example, 168–188
 comparing four models, 170
 and multicollinearity, 185–186
 more homogenous groups, 181
 Het.xls, 513–515, 519–527
 heteroskedastic errors. *See also* Breusch-Pagan
 (BP) test; generalized least squares (GLS)
 and heteroskedasticity nonlinear least
 squares (NLLS); robust standard errors
 (robust SEs)
 bivariate example, 519–527
 bivariate Monte Carlo evidence, 519–521
 in bootstrap example, 715–716, 720
 defined, 508, 511
 and earnings function Monte Carlo simulation,
 551–554
 effect on OLS estimated SE, 519
 effect on OLS *P*-value, 521
 general implementation, 522–523
 general Monte Carlo evidence, 524
 and inconsistency of OLS estimated SE,
 542
 and semilog functional form, 554
 univariate exact SE, 518
 univariate example, 513–518
 univariate Monte Carlo evidence, 515
 visual explanation, 508–509
 visual summary of ideas, 555

- heteroskedasticity, 41, 129–133, 296, 331. *See also*
 - heteroskedastic errors
 - consequences for OLS, 555–556
 - defined, 130
 - two basic strategies for handling, 560
- HetGLS.xls, 543–544, 546–548
- HetRobustSE.xls,
 - comparing conventional and robust SEs, 537–540
- histogram,
 - in SAT example, 34–35
 - of residuals, 119–121, 133, 134
- Histogram (Histogram.xla) add-in,
 - installation, 55
- homogeneous groups,
 - and confounding, 178–181
 - and wage discrimination, 63
- homoskedastic errors,
 - defined, 508
 - visual explanation, 509
- homoskedasticity, and root mean square error (RMSE), 41
 - defined, 130
 - in measurement box model, 289
- Hooke, Robert, 296
- Hooke's Law example,
 - data generation process for, 297–298, 299
 - with heteroskedasticity, 519
 - and measurement box model, 296, 301
- HookesLaw.xls, 298–300
- hourly earnings and education. *See* wage and education
- HourlyEarnings.xls, 133
- HowtoPredictYfromLnY.doc, 608, 610
- hypothesis tests, 257–262, 430. *See also* *F*-tests
 - and coin flip example, 431
 - and confidence intervals, 421–422
 - in education and wages two box example, 310–313
 - how to report results of, 433
 - and multicollinearity, 480–487
 - normal distribution and, 433
 - test statistic, 435–436
- HypothesisTest.xls, 431–433
- hypothetical difference of wages example, 306–309
- hypothetical wage example,
 - and confidence intervals, 421–422
 - and hypothesis tests, 421–422
- ice cream sales example, 613
- identically distributed errors, 297–298
- identically distributed random variables, 275
- identification, 745
- IMRGDPFunForm.xls, 145–147, 257–260
- IMRGDPReg.xls, 128
- independence, 275. *See also* independent random variables
 - of error terms, 289
 - between error terms and *X* variables, 297–298
 - of observations in trend time series, 627
- independent random variables, variance and SD
 - of sum of, 275–276
- independent variables, in simultaneous equations model (SEM), 731
- India population example, 605–608
- IndianaFTWorkers.xls, 53–65
- IndiaPopulation.xls, 605–608
- individual null hypothesis versus joint null hypothesis, 482
- infant mortality example,
 - and nonlinearity, 128
- inference, 238
 - defined, 215
 - fundamental question, 238
 - how it can break down, 295–296
 - and mountain peaks example, 283–284
 - reasoning from sample to population, 281
 - steps in, 260–261
- instrumental variables,
 - defined, 742
- intercept in ordinary least squares, 113
- internal rate of return (IRR), 150
- interval forecast,
 - defined, 398–399
- iteration,
 - defined, 686
- joint confidence region, 484–487
 - and RMSE, 485
- joint null hypotheses, 479–482, 485
- joint sampling distribution, 481
 - and correlated *X*s, 482, 483
- kitchen sink regression, 455
- Krueger, Alan, 498
- LaggedDepVar.xls, 649
- lagged dependent variables, 638. *See also* partial adjustment
 - defined, 638
 - and estimated ρ test, 649–651, 652
 - and money demand example, 645–648
 - testing for autocorrelation in the presence of, 649–651
- lagging,
 - defined, 561
- large-sample analysis, 262, 416–419. *See also* central limit theorem
- least absolute deviation (LAD) estimator, 81
- least median of squares (LMS) estimator, 81
- least squares estimators,
 - sample average, 341–342
- Legendre, Adrian-Marie,
 - and invention of ordinary least squares, 91
- linear estimator,
 - averages slopes, 360
 - defined, 340
 - extreme points, 358
 - ordinary least squares slope estimator, 354–355
- linear model, in education and wages example, 451
- linear probability model (LPM),
 - campaign contribution example, 677
 - versus nonlinear least squares (NLS), 683
 - ordinary least squares in a DDV Model, 676
 - weaknesses, 678–679

Cambridge University Press

0521843197 - Introductory Econometrics: Using Monte Carlo Simulation with Microsoft Excel®

Humberto Barreto and Frank M. Howland

Index

[More information](#)

768

linear trend. *See* trends, linear

linearity,
 different meanings of, 340

LINEST (Excel function), 86–89, 412, 445
 multiple regression example, 173
 and perfect multicollinearity, 187
 in skiing example, 325
 statistics as random variables, 412
 and whole-model F -test, 455

LinestRandomVariables.xls, 412–420

lin-log. *See* functional forms, semilog

log normal distribution, 414
 and confidence intervals, 429
 for errors, 417–419
 in hypothetical difference of wages example, 306

logistic distribution, 732

logit,
 compared to probit, 680–681
 defined, 680
 terminology, 682

log-linear. *See* functional forms, semilog; earnings function

log-linear trend. *See* trends, exponential

long regression, 456–457. *See also* short regression
 computer use example, 499
 in skiing example, 496
 and spurious regression, 612–613, 635

LPM. *See* linear probability model (LPM)

LPMonteCarlo.xls, 679

Luteinizing.xls, 601

matrix algebra, 89

maximum likelihood,
 via DDV add-in, 675
 and nonlinear least squares (NLLS), 681
 terminology, 682

MCSim add-in. *See* Monte Carlo simulation (MCSim.xla) add-in

Measure.xls, 287, 291–292, 293

measurement box model, 286–291. *See also* assumptions of measurement box model
 and bivariate regression, 296–301
 and classical econometric model, 301
 versus classical econometric model, 282
 compared to other box models, 290
 data generation process, 289
 expected value of measurement in, 289
 in equations, 288
 pictorial representation of, 287–288
 precision of measuring device, 289, 292, 295
 unbiased measurement process, 289

measurement error. *See also* heteroskedastic errors
 defined, 284
 idea behind, 284–285

median,
 sum of squared residuals for, 110

Mincer, Jacob, 138, 161

minimizing SSR,
 and NLLS, 685
 not same as finding best estimator, 375

missing values, 89–90
 in CPS, 15, 67
 shown by a period, 562

Index

Misspecification.xls, 602

money demand example, 645–648
 Durbin–Watson test, 646–651
 elasticities, 647–648
 estimated ρ test, 646–651
 feasible generalized least squares (FGLS), 646
 and partial adjustment, 645–647
 theory, 645

MoneyDemand.xls, 645–647

Monte Carlo simulation, 661. *See also* box model;
 Monte Carlo simulation (MCSim.xla) add-in
 advantages of, 229–230
 and algebra of expectations, 276, 277
 and assuming a zero average error, 327
 approximate SE of OLS regression slope, 326–327
 approximation to expected value, 269
 and best linear unbiased estimator, 345–346
 and the box model, 244–249
 checking for stationarity and, 628–630
 of chi-square distribution, 459–460
 and confidence intervals, 424–425, 426, 427–429
 and consistency, 263–264
 convergence of empirical histogram toward probability histogram, 274
 and correlation of Training and Talent, 333
 defined, 216
 draws with replacement, 331
 and Durbin–Watson test, 650–651, 652
 empirical histogram as approximation to probability histogram, 248–249, 387
 estimated ρ test, 579–580, 650–651, 652
 and F -distribution, 461
 and feasible generalized least squares (FGLS), 597–599
 and forecast errors, 657
 and F -tests, 466–467, 472–473
 free throw example, 225–229
 and Gauss–Markov theorem, 366
 and heteroskedasticity, 520
 in Hooke’s Law example, 299–300
 and hypothesis tests, 437–441
 and hypothetical difference of wages example, 308–309
 and joint sampling distribution, 480
 and measurement box model, 291–292
 as method to approximate SE, 227, 387
 and mountain peaks example, 291–292
 not a method of estimating unknown parameters, 292, 293
 as practiced by Gosset, 451–452
 random walks and, 631–632, 633
 SD as approximation of SE, 249, 291, 299–300, 308, 327, 416
 for SE of sample average estimator, 349
 spurious regression and, 635–636, 637
 streaks in free throws, 230–231
 of trend DGP, 627
 X ’s fixed in repeated sampling in, 331

Monte Carlo simulation (MCSim.xla) add-in,
 versus built-in Monte Carlo, 232
 and consistency, 263–264
 installation, 232

Cambridge University Press

0521843197 - Introductory Econometrics: Using Monte Carlo Simulation with Microsoft Excel®

Humberto Barreto and Frank M. Howland

Index

[More information](#)

Index

769

- Monte Carlo simulation with Solver (MCSimSolver.xls) add-in
 - described, 235
- MonteCarlo.xls, 226–229
- MortDisc.xls, 697–702
- MortDiscMCSim.xls, 702–703
- mortgage discrimination example, 695
 - data, 695
 - data generation process, 697–700
 - drawing a sample, 700
 - implementing the DGP, 697–700
 - raw differential, 703–706
- mountain peaks example, 283–296
- and measurement box model, 287
- multicollinearity,
 - between age and experience, 188–189
 - in bivariate regression, 355
 - consequences of perfect, 185–186
 - defined, 184
 - effect on SE, 383
 - and hypothesis tests, 480–487
 - and LINEST, 187
 - near-perfect, 189
 - perfect, 187
- Multicollinearity.xls, 700
- multiple regression, 165
 - analytical derivation of coefficients, 194
 - and confounding, 183, 255
 - defined, 166
 - as double compression, 182
 - improving prediction, 173
 - interpreting coefficients, 183
 - LINEST output and, 89
- MultiReg.xls, 168–184
- MyMonteCarlo.xls, 488
- NLLSFit.xls, 682–688
- NLLSMCSim.xls, 689
- NoInterceptBug.xls, 457
- nominal vs. real variables, 609
- nonlinear estimator,
 - sample median as, 340
- nonlinear fitting. *See* nonlinear least squares (NLLS)
- nonlinear least squares (NLLS),
 - and consistency, 689
 - compared to maximum likelihood, 681
 - demonstration of fitting, 684, 687
 - displaying results of, 693
 - estimation with Solver, 687
 - and heteroskedasticity, 695
 - interpreting coefficient estimates of, 691–692
 - versus linear probability model, 683, 694
 - Monte Carlo evidence, 689
 - no analytic solution, 686
 - S shape of, 681
 - terminology for, 682
 - visualizing surface, 685
- nonlinear regression. *See* nonlinear least squares (NLLS)
- normal approximation,
 - and *P*-values, 445, 446
 - versus Monte Carlo simulation, 253–254
- normal distribution, 291, 414
 - and chi-square distribution, 458–460
 - and confidence interval game, 424–425
 - and Gauss–Markov theorem, 363–366
 - and test statistics, 435
 - versus *t* distribution
- NORMALRANDOM (Excel function), 293, 414, 568. *See also* normal distribution
- #NAME? error, 223
- to generate normally distributed random numbers, 223
- and heteroskedasticity, 513
- superior to NORMINV, 223
- NORMINV (Excel function),
 - explanation of, 222
 - NLLS example, 699
 - problems with, 223
 - using, to generate normally distributed random numbers, 221
- notation for estimates and parameters, 299
- null and alternative hypotheses,
 - null hypothesis, 257, 430, 464
 - in cigarette demand example, 444
 - in total personal income example, 476
- objective function, and ordinary least squares, 91–92
- observation, 73–74, 75, 82–83
- observational study. *See also* omitted variable bias
 - cigarette example, 17
 - and omitted variable bias, 491
 - types of, 17
- Ohtani, Kazuhiro, 721
- OLS. *See* ordinary least squares
- OLS Regression (OLSRegression.xls) add-in,
 - installation, 540
 - for robust SE, 540
- OLSFormula.xls, 108–113
- omitted variable,
 - but no bias, 495
 - defined, 490, 494
- omitted variable bias. *See also* confounding
 - basic lesson of, 505–506
 - and computer use example, 500–501
 - and controlled experiment, 492
 - in a hypothetical example, 494–497
 - in mortgage discrimination example, 697
 - and observational study, 491
 - and Random *X*'s DGP, 503, 504–505
 - and spurious regression, 612–613
 - violation of independence, 496
 - visual explanation, 494
- omitted variable rule,
 - applied to computer use example, 500–501
 - defined, 197
 - and spurious regression, 612–613
- one-tailed test,
 - in education and wages two box example, 310
- one-tailed versus two-tailed tests, 433
- in cigarette demand example, 444
- ordinary least squares (OLS), 72–94. *See also* sum of squared residuals, minimizing.
 - alternatives to, 81
 - computing, 82–84
 - deriving formulas for, 91–94

- ordinary least squares (*cont.*)
 - formulas for, 82–84, 111–112
 - versus 2SLS, 742–743, 744
 - weighted sum of Y values, 84, 93–94
- ordinary least squares (OLS) estimator. *See also* least squares estimator
 - computing expected values of, 368, 370–372
 - as unbiased estimator, 326, 368, 370–372
- ordinary least squares slope,
 - as unbiased estimator, 357–358
 - as weighted sum, 111–112, 354–357
- orthogonal regression, 81
- outliers, 81
 - and Anscombe data set, 127
- P Value Calculator (PValue.xls) add-in
 - installation, 259
 - and Durbin–Watson test, 585–586
- paired XY resampling. *See* bootstrap
- PairedXYBootstrap.xls, 715–716, 720
- parameter, 238
- partial adjustment, 638–647
 - equilibrium value for dependent variable, 643
 - long run, 643–645, 647
 - and money demand example, 645–647
 - short run, 641–643, 647
- PartialAdjustment.xls, 639–645
- Pearson, Karl, 51, 453, 455, 458
- PercentageBootstrap.xls, 711–712
- PivotTable (Excel), dynamic data exploration, 63
 - in education and wages two box example, 311
 - and graph of averages, 105
 - grouping, 63, 67
 - and heteroskedasticity, 133, 134
 - toolbar, 57
- point forecast,
 - defined, 398
- point of averages, 35–37. *See also* scatter diagram
 - in Costa Rica example, 38
 - and ordinary least squares intercept, 113
 - and ordinary least squares regression line, 84, 112, 113
- polling box model, 251–257, 286
- population, 238
 - and box model, 243
- predicted value of Y given X. *See* conditional mean function, predicted Y
- predicted Y, 74–75
 - and conditional mean function, 96–98
 - computing, 83
 - defined, 398
 - equation for, 83
 - and regression line, 98–99
 - and residuals, 114–116
 - and two regression lines, 106
 - and vertical strips, 96–98
- present value, 150
- presentation of regression results, 445–449
- presidential heights example, 251–254. *See also* cluster sample
 - consecutive presidents sampling scheme, 254
- PresidentialHeights.xls, 251–254
- probability density function (pdf), 265
- probability histogram of sample slope,
 - with log normally distributed errors, 416, 417
 - with normally distributed errors, 416
- probability histograms, 243, 266
 - approximated via Monte Carlo simulation, 248–249, 251–253
 - and convergence to normal curve, 249
 - defined, 247
 - and expected value of sample sum, 243–244
 - for F-statistic, 466–467
 - and SE of sample sum, 253
 - versus empirical histograms, 266
- probit,
 - compared to logit, 680–681
 - defined, 680
 - terminology, 682
- proxy variable, 25
- P-values,
 - for difference between population averages, 305
 - in education and wages two box example, 313
 - and F-statistic, 470
 - and F-tests, 463
 - and Monte Carlo simulation, 521, 525, 530, 531, 532, 533, 572, 574, 575, 593, 594, 635–636, 637, 650–651
 - proper interpretation of, 261
 - spurious regression and, 259, 431, 635–636, 637
- qualitative response variable. *See* dummy dependent variable
- R Squared. *See* R²
- R², 122–126
 - adjusted, 125, 607
 - bootstrap example, 723–724
 - computing, 124–125
 - defined, 123
 - and functional form comparison, 146
 - and kitchen sink regression, 455
 - logic of, 123–124
 - misuse of, 125
 - and regression through the origin, 457
 - and trends, 607
- Raid example, 666–669
- Raid.xls, 666–669
- RAND (Excel function), 218
 - criticized in literature, 237
 - performance and evaluation, 218
 - usage in a formula, 221
- RANDOM (Excel function), 413. *See also* uniform distribution
 - #NAME? error, 223
 - algorithm, 219
 - performance and evaluation, 219
 - superior to RAND, 223
 - usage in a formula, 221
- random number generation,
 - comparing random number generators, 218
 - criticism of Excel’s RAND, 237
 - Data Analysis add-in warning, 221
 - discussed in Basic Tools folder, 3
 - linear congruential generator, 217
 - in LinestRandomVariables.xls, 413
 - log normal, 414, 415

Cambridge University Press

0521843197 - Introductory Econometrics: Using Monte Carlo Simulation with Microsoft Excel®

Humberto Barreto and Frank M. Howland

Index

[More information](#)

Index

771

- multiple recursive generator, 219
- normal distribution, 221, 414, 415
- performance and evaluation, 217–218
- pseudo-randomness, 216–217
- RAND function, 218
- RANDOM function, 219
- RND function, 218
- uniform distribution, 221, 271, 413, 415
- random variables, 244, 265–276
 - defined, 265
 - discrete, 265
 - sums and averages of, 274–276
- random walk, 631–633
 - with drift, 632–633
- random X 's, 476. *See also* weak dependence
 - and spurious regression, 611
 - and time series, 633
- random X 's DGP, 611
 - compared to classic econometric model, 505–506
 - compared to polling box model, 502–504
 - implementation, 502
 - rationale, 503, 504–505
 - with heteroskedasticity, 550
- Real-Time Data Set for Macroeconomists (RTDSM), 648
- reduced form equation,
 - in SEM example, 733
- Reg SS. *See* regression sum of squares
- Reg.xls, 73–90
- regression. *See also* conditional mean function
 - and Average Y line, 76
 - earnings and education example, 99
 - flexibility, 146
 - four ways to run in Excel. *See* regression line, fitting in Excel.
 - limitations, 126–134
 - linear in the parameters, 138, 141, 145
 - multiple. *See* multiple regression
 - notation for coefficients, 83
 - origin of name, 104
 - through origin, 87, 125
- regression effect, 104, 107
- regression line,
 - “by eye,” 73–74, 77
 - computing. *See* ordinary least squares; sum of squared residuals, minimizing
 - versus conditional mean function, 101
 - and correlation, 102–103
 - as double compression, 95–103
 - defined, 72
 - equation for, 73–74, 75
 - fitting in Excel, 84–90
 - intercept, 73
 - versus SD line, 101–103
 - slope, 73, 74
 - there are two regression lines, 103–107
- regression plane, 166
 - equation for, 166
 - live visualization, 167–168
- regression statistics,
 - as random variables, 325, 412–420, 458
- regression sum of squares (Reg SS),
 - defined, 124
- regression through the origin, 457. *See also*
 - restricted regression; Data Analysis; Regression add-in
 - and estimated ρ test, 578
 - Galileo example, 144
- regression to the mean. *See* regression effect
- replication of prior results, 648, 662
- residuals. *See also* bootstrap
 - defined, in multiple regression, 149, 166
 - versus error terms, 341, 384
- residual plots, 75
 - and heteroskedasticity, 133
 - Galileo example, 141, 143
 - in infant mortality example, 128
 - and plots of y versus x , 115–116, 119–121
- ResidualPlot.xls, 114–116
- residuals, 75–81
 - average of, 76
 - computing, 75, 83, 115–116
 - defined, 75
 - versus errors, 117
 - equation for, 72–75, 83
 - as estimates of error terms, 384
 - histogram. *See* histogram, of residuals, 75
- restricted and unrestricted models,
 - in food stamps example, 468–469, 473–474
 - in Galileo example, 463
 - in total personal income example, 475–476
- restricted model. *See also* restricted regression
 - in savings behavior example, 454
- restricted regression, 456–458
 - examples of, 456–458
 - and F -tests, 462
 - and Solver, 464
 - and SSR, 464–465
- RMSE. *See* root mean square error (RMSE)
- RMSE.xls, 118–121
- RNGPractice.xls, 221–225
- RNGTheory.xls, 217–220
- robust regression, 81
- robust standard errors (robust SEs),
 - better than conventional standard errors, 537
 - biased, but consistent, 540
 - defined, 534
 - formula for, 537
 - versus generalized least squares (GLS), 534
 - in OLS Regression add-in, 540
 - Monte Carlo evidence, 538
- root mean square error (RMSE),
 - as adjusted SD of residuals, 393
 - as biased estimator of SD of errors, 394
 - computing, 117, 118
 - as consistent estimator of SD of errors, 394
 - defined generally, 116
 - and degrees of freedom, 118
 - as estimate of SD of errors, 384
 - and estimated SE, 386, 426–427
 - explanation of bias in estimating SD of errors, 396
 - and heteroskedasticity, 133
 - in Hooke's Law example, 299
 - other names for, 117
 - as overall, single measure of spread, 121
 - rule of thumb using, 118
 - and SD line, 117

Cambridge University Press

0521843197 - Introductory Econometrics: Using Monte Carlo Simulation with Microsoft Excel®

Humberto Barreto and Frank M. Howland

Index

[More information](#)

772

Index

- ROUND (Excel function), 293, 323
- RSquared.xls, 123–125
- SameRegLineDifferentData.xls, 130–131
- sample average,
 - computing expected value of, 276
 - computing variance and SD of, 276
 - with heteroskedasticity, 516
- as least squares estimator, 109–110, 136–137, 341–342
- as minimum variance estimator, 347–350
- in mountain peaks example, 291–292
- probability histogram of, 251–254, 291, 293
- as random variable, 305
- and regression slope, 113
- SE of, 379
- as weighted sum, 108–109, 113
- weight of adult US citizens example, 255–256
- sample percentage, 258
- sample SD,
 - as estimate of SD of box, 256, 294–295, 312–313
- sample sum, 243
- samples, 243
- sampling,
 - in classical econometric model, 324
 - alternatives to simple random sampling. *See* cluster samples
- sampling distribution. *See also* probability histogram
 - and bootstrap, 712
 - depends on distribution of errors, 412
 - depends on number of observations, 412
 - depends on recipe, 412
 - depends on values of the X s, 412
 - of difference of sample averages, 312
 - of NLLS coefficient, 702–703
 - for regression statistics, 419–420
 - of sample slope, 435–442
 - of test statistic, 437–442
- sampling without replacement. *See also* presidential heights example
 - correction factor (CF), 253, 256, 307
 - in hypothetical difference of wages example, 306
- SAT example,
 - and R^2 , 122–124
 - and correlation, 34–36
 - and double compression of the data, 96–99
 - and heteroskedasticity, 130–131
 - and residuals, 114–116
 - and root mean square error (RMSE), 118–122
 - and SD line versus regression line, 102
 - and two regression lines, 104–107
- savings behavior example,
 - and F -statistics, 488
 - and hypothesis tests for multiple parameters, 453–454
 - dummy variables in, 454
 - restricted model in, 454
 - unrestricted model in, 454
- scatter diagram, 35–50
 - versus histogram, 38
 - recreating a “mind’s eye”, 37–38, 42
 - as useful display, 42–43
- scatter plot,
 - and ordinary least squares, 72–77
 - and regression line, 72–75
- SD,
 - of population versus SD of sample, 253
 - and squared deviations, 77
 - versus standard error, 244
 - two SDs rule of thumb, 34
- SD line, 35–36, 37
 - and two regression lines, 106–107
 - versus regression line, 101–103
- seasonal adjustment, 617–624
 - comparing methods of, 623, 624
 - and dummy variables, 621–622
 - how to do it, 621–622
 - as practiced by the BLS, 623
 - purpose of, 620
 - seasonal index, 622
 - should one do it?, 623–624
 - and trends in time series, 623
- SeasonalPractice.xls, 623
- SeasonalTheory.xls, 620–623
- SEb1OLS.xls, 381, 384–386
- SEForecast.xls, 399–407
- semilog functional form. *See* functional forms, semilog
- SemiLogEarningsFn.xls, 149–155, 450–451
- separate null hypotheses versus joint null hypotheses, 478
- serial correlation. *See* autocorrelation
- short regression, 456–457. *See also* long regression
 - in computer use example, 499
 - in skiing example, 496
 - and spurious regression, 612–613, 635
- significance level. *See* P -value
- SimEq.xls, 733–744
- simple random sample, 254–255
 - in two box model, 304
- Simpson’s paradox, 44. *See also* confounding
- simulation in econometrics, 452
- simultaneous equations (simultaneity) bias, 736
 - as confounding, 740
 - corrected by two stage least squares (2SLS), 742–743
 - defined, 736
 - factors affecting, 739–740
 - lack of consistency of OLS, 736
 - source of, 737–739
- simultaneous equations model (SEM). *See also* simultaneous equations bias; two stage least squares (2SLS)
 - crime example, 732
 - defined, 730–731
- skiing example, 317–328
 - and classical econometric model, 319–328
 - measurement box DGP for, 317–319
 - modified version of more realistic model, 320
 - more realistic model of DGP, 319–320
- Skiing.xls, 322–328
- SkiingOVB.xls, 494–497
- slope estimates,
 - as random variables, 325–326

Cambridge University Press

0521843197 - Introductory Econometrics: Using Monte Carlo Simulation with Microsoft Excel®

Humberto Barreto and Frank M. Howland

Index

[More information](#)

Index

773

- small samples, 262
 - and estimated SE, 427–429
 - and test statistics, 435, 437–441
- Smith, Adam, 160
- Solver add-in, 77–78
 - and minimizing sum of squared residuals, 110
 - and NLLS, 687
 - using, to obtain OLS estimates, 78
 - obtaining, 77
- spurious regression,
 - defined, 612
 - and omitted variable bias, 612–613
 - and random walks, 634–636
 - and random X 's, 611, 635–636
 - and trends in time series, 611–613
- Spurious.xls, 611–613, 635
- square root law,
 - for SE of the difference of sample averages, 305
- standard deviation (SD) in SAT example, 34
- standard deviation (SD) of a random variable,
 - defined, 268
- standard error (SE), 243. *See also* estimated SE,
 - standard error of the forecast, standard error of the forecast error
 - and algebra of expectations, 369–370, 372–373, 379, 380–381, 516–518, 535
 - approximate (from Monte Carlo), 387
 - biased by heteroskedasticity, 521, 537
 - and correlated X 's, 482
 - determinants of, 387
 - of difference of sample averages, 305, 308
 - estimated, 536
 - estimated vs. exact, 307
 - exact, 300, 347–349, 381–383
 - and heteroskedastic errors, 536–537
 - of OLS slope, 326–327, 380–381
 - using RMSE to estimate, 386
 - of sample average, 253, 255, 291, 292, 293, 369–370, 379
 - of sample slope, 381
 - of the sample percentage, 258
 - and sample size, 388
 - and the SD of the errors, 388
 - and the SD of X , 388
- standard error (SE) for sample sum,
 - computing in box model, 244
- standard error (SE) of sample slope,
 - estimated, 426, 428
 - exact, 416–417, 419, 425–426, 428
- standard error (SE) of the forecast error,
 - and autocorrelation, 655–657
- standard error (SE) versus SD, 244, 273
- standard error of the estimate. *See* root mean square error (RMSE)
- standard error of the forecast,
 - depends on X , 402
 - formula, 402
 - how to compute, 408
 - intuition, 402
- standard error of the forecast error,
 - formula, 404
 - intuition, 404
- standard error of the regression. *See* root mean square error (RMSE)
- standard units, 431
- static model, defined, 558
- stationarity, 624
 - converting a time series to, 628
 - defined, 625
 - and trends, 625–628
- Stationarity.xls, 625
- statistical inference. *See* inference
- statistical significance, 259, 411, 431
 - meaning of, 315
 - versus practical importance, 261–264, 411, 447, 449–450, 476
- streaks,
 - in free throw shooting, 230–231
- Street, Picabo, 317
- strong dependence. *See* autoregressive processes,
 - and strong dependence
- structural equation,
 - in SEM example, 732
- Student. *See* Gosset, William Sealy
- sum of squared residuals (SSR), 77–80
 - and chi-square distribution, 460
 - computing, 77, 83
 - and hypothesis tests, 487
 - and kitchen sink regression, 455
 - as function of intercept and slope, 78–80
 - for median, 110
 - minimizing, 77–80
 - for sample average, 109–110
 - sampling distribution for, 419, 420
- summarizing a bivariate relationship, 122
- t -distribution, 261, 413, 439–442
 - and degrees of freedom, 439–441
 - and test statistics, 435–436
- tabulation. *See also* cross tab
 - history, 15
- TDIST (Excel function), 447
- test statistics, 258–259
- tests for equality of two parameters,
 - 468–475
- tests of significance. *See* hypothesis tests
- Theil, Henri, 742
- threshold,
 - campaign contribution example, 670
 - general role in DDV models, 669
 - mortgage discrimination example, 700
 - Raid example, 666
- Tibshirani, Robert, 726, 728
- time series,
 - defined, 558
- time trend term (t), 605
- TimeSeriesDummyVariables.xls, 613–614, 660
- total personal income example,
 - and F -statistics, 475–477
- total sum of squares,
 - defined, 124
- total variation. *See* total sum of squares
- transforming to obtain predicted values,
 - from logarithmic to linear, 608, 610–611
- trend stationary process,
 - defined, 628
- Trendline (Excel Chart option), 85
 - in skiing example, 325

trends, 605–613
 and autocorrelation, 653–657
 and spurious regression, 611–613
 as non-stationary processes, 625–628
 comparing, 608, 609–611
 exponential, 607–611
 linear, 605–609, 610
t-statistics, 434–442
 and cigarette demand example, 446–447
 in food stamps example, 475
 and *F*-statistics, 475
t-tests. *See also t*-statistics, test statistics
 versus *F*-tests, 473–475
two box model, 303–314
 for education and wages two box example, 311–313
 pictorial representation of, 304, 312
two stage least squares (2SLS),
 adjustment needed for SE, 744
 defined, 742–743, 744
 versus OLS, 742–743, 744
 in SEM example, 742–743
TwoBox.xls, 306–309
TwoRegressionLines.xls, 104–107

U.S. Real GDP example, 609–611
unadjusted data, 619
unbiased estimators, 262. *See also* biased estimator
 for bivariate slope, 357–361
 defined, 343
 OLS Slope estimator, 356
unemployment rates,
 and seasonal adjustment, 617–619, 620
uniform distribution, 413, 414
 confidence interval game, 422–424
 and confidence intervals, 429
 for errors, 467
 and expected value, 422
 and Gauss–Markov Theorem, 363–366
 and hypothesis tests, 432
units for coefficients, 322
univariate sample statistics. *See* sample average,
 sample SD, sample percentage

unrestricted model. *See also* restricted and
 unrestricted models; restricted regression;
 unrestricted regression
 in savings behavior example, 454
unrestricted regression, 456

Valleta, David, 712, 727, 732
variance and SD of a random variable,
 defined, 268
 rules under algebra of expecations, 271–276
vertical strips,
 and conditional mean function, 96
VisiCalc, 4
Visual Basic,
 error messages, 9
 macros, 89

wage and education,
 and heteroskedasticity, 131–133
wage discrimination,
 defined, 62
WagesOct97.xls, 549–554
Watson, Geoffrey S., 558, 582
weak dependence. *See also* differencing data
 and estimated ρ test, 634
 defined, 634
weight of adult U.S. citizens example, 254–256
weighted average, defined, 61
weights, in formula for ordinary least squares
 slope, 111–113
whole-model *F* test, 455, 457
Winter Olympics Games, 317
with replacement,
 drawing, in computer use example, 502–504
without replacement,
 drawing, in computer use example, 502–504
Wooldridge, Jeffrey, 559

ZandTTests.xls, 436–441
zero intercept. *See* regression through the origin
z-statistic, 258–259
 for difference between population averages, 305
 in education and wages two box example,
 312–313
z-test, 433