



# Index

## A

Acceptance, of hypothesis. *See* Hypothesis testing

Additive property, of chi-square random variables, 264

Algebra of events, 56–57, 56f, 57f

Analysis, of residuals, 384–386, 385–386f

Analysis of variance (ANOVA), 445–476

- introduction, 445–446
- one-way, 446, 448–459, 452t, 453f, 456t
  - multiple comparisons of sample means, 456–458
  - with unequal sample sizes, 458–459
- overview, 446–448
- two-way, 446
  - hypothesis testing for, 464–468, 467t, 468f
  - interaction, 446, 469–476, 475t, 476f
  - introduction and parameter estimation, 460–463

ANOVA. *See* Analysis of variance

Approximately normal, 31, 31f

Assessment, of models, 384–386, 385–386f

Assignable cause, 553

Association, causation v., 39

Associative law, 56–57

Averaging notation, in two-way analysis of variance, 461–462

## B

Balanced case, in one-way analysis of variance, 459

Bar graph, 10, 10t, 11f

Bayes estimator, 236, 279–284

Bayes' formula, 68–75, 69f

Bayesian approach, 236, 604–606

Behrens-Fisher problem, 324–325

Bernoulli density function, 419

Bernoulli, Jacob, 5–6

Bernoulli, James, 141

Bernoulli parameter, maximum likelihood estimator of, 237–240

Bernoulli populations
 

- hypothesis testing in, 329–336
- testing equality of parameters in two of, 333–336

Bernoulli random variables, 141–148, 143f, 148f
 

- confidence interval for mean of, 268–272, 272t
- generation of, 639
- hypergeometric random variables and, 157–158
- sign test and, 522
- signed rank test and, 526–527

Beta distribution, 282

Between samples sum of squares, 451–452, 454, 456t, 459

Bias, of estimators, 274

Bimodal data set, 33, 34f

Binary output data, logistic regression models for, 418–421, 418f

Binomial distribution function, computation of, 147–148, 148f

Binomial distribution, hypothesis testing in, 329–336

Binomial random variables, 141–148, 143f, 148f
 

- binomial distribution function, 147–148, 148f
- central limit theorem and, 212, 215–216, 216f
- generation of, 639–640
- hypergeometric random variables and, 158–160, 225–226

Binomial random variables (*continued*)  
 Poisson random variables and, 150–153  
 sign test and, 522  
 Bootstrap method, 623–630  
 mean square error and, 627–628  
 population mean and, 623–624, 628–630  
 population variance and, 623–625  
 probability and, 629–630  
 random variable generation, 627  
 Box plot, 27, 27f  
 Box-Muller method, 642

## C

Categorical data analysis, goodness of fit tests  
 and, 489–514  
 critical region determination by simulation,  
 496–499, 498f  
 dummy variables for, 416–418  
 introduction, 489–490  
 Kolmogorov–Smirnov goodness of fit test for  
 continuous data, 510–514, 511f  
 tests of independence in contingency tables,  
 501–505  
 tests of independence in contingency tables  
 having fixed marginal totals, 505–510,  
 506f  
 tests when all parameters are specified,  
 489–497, 493t, 498f  
 tests when some parameters are unspecified,  
 499–501  
 Causation, association v., 39  
 Central limit theorem, 210–219, 212–216f,  
 approximate distribution of sample mean,  
 216–218  
 binomial random variables and, 212,  
 215–216, 216f  
 defined, 210  
 independent random variables, 226–227  
 Monte Carlo study and, 643  
 probability mass function and, 212–215f  
 sample size needed, 218–219, 219f  
 Chance variation, 553  
 Channel noise disturbance, 173  
 Chebyshev's inequality, 27–30, 29t  
 one-sided, 29–30  
 random variables and, 128–131  
 weak law of large numbers, 129  
 Chi-square approximation, 497–498  
 Chi-square density function, 188–193,  
 190f, 192f  
 Chi-square distribution, 188–193, 190f, 192f,  
 327–328, 607  
 gamma random variables and, 191–192, 192f  
*t*-distribution and, 193–195, 193f, 194f  
 Chi-square goodness of fit tests, history of, 6  
 Chi-square random variables, 189, 190f  
 additive property of, 264  
 estimators of variance from, 447–448  
*F*-distribution and, 195–196, 195f  
 mean and variance of, 192  
 sample mean and variance distribution with,  
 222  
 Choosing, of normal prior, 283–284  
 Class boundaries, 15, 15f, 15t  
 Class intervals, 14–15, 15f, 15t  
 Classical approximation, for rank sum test,  
 535–537  
 Classical simulation, for rank sum test,  
 537–539, 538f  
 Coefficient of determination, sample correlation  
 coefficient and, 382–384  
 Coefficient of multiple determination, 411  
 Column factors  
 hypothesis testing for, 464–468, 467t, 468f  
 in two-way analysis of variance, 460  
 deviation from grand mean due to, 462  
 row factor interaction with, 446, 469–476,  
 475t, 476f  
 Column sum of squares, 467t  
 Combinations, permutations and, 63  
 Combining, of unbiased estimators, 275–277  
 Common density function, independent  
 variables and, 102–103  
 Commutative law, 56–57, 56f  
 Comparison, of sample means, 456–458  
 Complement  
 in sample space, 55  
 in Venn diagram, 56–57, 56f, 57f  
 Composite hypothesis, 298  
 Computational identity, of sum of squares of  
 residuals, 366  
 Conditional densities, 106–107  
 Conditional distributions, 104–107,  
 337–339  
 random numbers and, 164–166, 166f  
 Conditional probability, 65–68, 66f, 106  
 independent events and, 75–79  
 Poisson random variables and, 153–154  
 Conditional probability density function,  
 280–281

- Conditional probability mass function, 105–106
  - Confidence, 235
  - Confidence interval. *See also specific intervals*, e.g. 95 Percent confidence interval
    - for difference in means of two normal distributions, 260–268, 262–263f, 266–267f, 260t, 268t
    - for difference in population means, 456–458
    - for estimating unknown mean, 247–252
    - for exponential distribution in life testing, 594–596
    - for mean of Bernoulli random variable, 268–272, 272t
    - for mean of exponential distribution, 272–273
    - for normal mean with unknown variance, 252–258, 256f
    - one-sided lower, 249–250, 255–257, 256f, 260t, 261–265, 261–263f, 268t
    - one-sided upper, 249–250, 254–255, 260t, 261, 265–268, 266–267f
    - prediction interval  $v$ ., 381
    - for regression parameters
      - $\alpha$ , 376–377
      - $\beta$ , 371–372
    - in sequential testing for exponential distribution in life testing, 599–601
    - two-sided, 248, 251, 255–257, 256f, 261–268, 262–263f, 266–267f
    - for unknown probability, 268–272, 272t
    - for variance of normal distribution, 259–260, 260t
  - Confidence interval estimators
    - of difference in means of two normal distributions, 260–265, 262–263f
    - for mean of exponential distribution, 272
    - of mean response, 377–379, 411–413
  - Contingency tables
    - with fixed marginal totals, tests of independence in, 505–510, 506f
    - tests of independence in, 501–505
  - Continuous data, Kolmogorov–Smirnov goodness of fit test for, 510–514, 511f
  - Continuous random variables, 93–94, 640–643
  - Control charts, 553
    - fraction defective, 565–567
    - mean, 554–562, 556f, 564f
      - case of unknown, 557–562, 559t
    - number of defect, 567–570, 569t
    - for population mean, 571–581
      - moving-average, 571–573, 573t, 574f
    - variance, 562–565, 564f
  - Control group, 164
  - Control limits
    - for fraction defective, 566–567
    - for mean control charts, 555–557, 556f
    - for moving-average, 572–573, 573t, 574f
    - for variance control charts, 563–564, 564f
  - Controls, 336
  - Correlation analysis, history of, 6
  - Counting
    - basic principles of, 60–63
    - notation and terminology, 63–65
  - Covariance, 121–122
  - Critical region, 298, 300
    - for goodness of fit tests when all parameters are specified, 490–491
    - one-sided, 307
    - simulation for determination of, 496–499, 498f
  - Critical value, for determining independence of characteristics of population member, 504
  - Cumulative distribution function, 91–93, 93f, 96
    - exponential random variables and, 177
    - probability density function and, 94–95, 94f
  - Cumulative sum control charts, 579–581
- D**
- Data collection, descriptive statistics and, 1–2
  - Data sets
    - describing, 9–17
      - frequency tables and graphs, 10, 10t, 11f, 12f
      - grouped data, histograms, ogives, and stem and leaf plots, 14–17, 14t, 15f, 15t, 16f, 18t
      - relative frequency tables and graphs, 10–14, 13–14f, 13t
    - normal and skewed, 31–33, 31f, 32f, 34f
    - summarizing, 17–27
      - sample mean, sample median, and sample mode, 17–22
      - sample percentiles and box plots, 24–27, 26t, 27f
      - sample variance and sample standard deviation, 22–24
  - Defects, probability of, 329–337
  - DeMorgan's laws, 57

- Density function. *See also* Probability density function  
 Bernoulli, 419  
 chi-square, 188–193, 190f, 192f  
 common, 102–103  
 conditional probability, 280–281  
 $F$ , 195–196, 195f  
 gamma distribution and, 186–188  
 joint, 236–237, 242, 244  
 joint probability, 99–101  
 of logistics distribution, 197  
 normal, 168, 168f, 188, 189f  
 posterior, 280, 284  
 random variables and, 112  
 Rayleigh, 591  
 $t$ , 193, 193f, 252, 253t  
 Weibull, 608, 609f
- Density, mode of, 284
- Dependent events, 75–79, 78f
- Dependent variable. *See* Response variable
- DES. *See* Diethylstilbestrol
- Descriptive statistics  
 Chebyshev's inequality, 27–30, 29t  
 data collection and, 1–2  
 describing data sets, 9–17  
     frequency tables and graphs, 10, 10t, 11f, 12f  
     grouped data, histograms, ogives, and stem and leaf plots, 14–17, 14t, 15f, 15t, 16f, 18t  
     relative frequency tables and graphs, 10–14, 13–14f, 13t  
 history of, 7  
 normal data sets, 31–34, 32f, 34f  
 paired data sets and the sample correlation coefficient, 34–41, 35t, 36f, 39f  
 summarizing data sets, 17–27  
     sample mean, sample median, and sample mode, 17–22  
     sample percentiles and box plots, 24–27, 26t, 27f  
     sample variance and sample standard deviation, 22–24
- Deviation from grand mean due to column  $j$ , 462
- Deviation from grand mean due to row  $i$ , 462
- Diethylstilbestrol (DES), 336
- Difference, in means of two normal distributions, 260–268, 262–263f, 266–267f, 268t
- Discrete inverse transform method, 639–640
- Discrete random variables, 91–92  
 expectation and, 111, 111f  
 generation of, 638–640  
 probability mass function and, 92–93, 638–640
- Dispersion parameter, 197
- Distribution  
 binomial, hypothesis testing in, 329–336  
 chi-square, 188–193, 190f, 192f, 327–328, 607  
 conditional, 337–339  
 exponential, confidence interval for mean of, 272–273  
 $F$ , 328–329  
 gamma, 595, 604–607  
 hypothesis testing for determining equality of  $m$  population distributions, 508–509  
 of least squares estimators, 361–367, 368f, 369f  
 life, 245–247  
 multivariate normal, 404  
 normal  
     confidence interval for variance of, 259–260, 260t  
     estimation of difference in means of, 260–268, 262–263f, 266–267f, 268t  
 Poisson  
     goodness of fit tests for, 499–501  
     hypothesis testing concerning mean of, 336–339  
     variance in, 395–396  
 prior, 279–284, 604–606  
 probability, of estimator of mean response, 377–378  
 rate of, 590  
 of sample, goodness of fit tests for, 489–499, 493t, 498f  
 uniform, estimating mean of, 244–245
- Distribution function. *See also* Cumulative distribution function; Probability distribution function  
 binomial, 147–148, 148f  
 of continuous random variable, 640–641  
 empirical, 624–625  
 moment generating function and, 128  
 of normal random variables, 170–171  
 Poisson  
     computation of, 155–156  
     number of defects and, 567–570, 569t  
 probability and, 91–92  
 random variables and, 91–92, 624

- of rank sum test, 531
- signed rank test for, 525, 526f
- two-sample problem and, 531
- Distribution results, summary of, 381–382
- Distributive law, 56–57, 57f
- Doll, R., 17
- Dot notation, in two-way analysis of variance, 461–462
- Double-blind test, 164
- Dummy variables, for categorical data, 416–418

## E

- Effect of column  $j$ , 470
- Effect of row  $i$ , 470
- Empirical distribution function, 624–625
- Empirical rule, 33–34
- Entropy, 110
- Equal variance, testing equality of means of two normal populations with, 324–325, 325t
- Equality
  - of  $m$  population distributions, hypothesis testing for, 508–509
  - of means of two normal populations, 318–326, 319t, 321f, 323f, 325t
    - case of known variance, 318–320, 319t
    - case of unknown and equal variance, 324–325, 325t
    - case of unknown variance, 320–324, 321f, 323f
    - hypothesis testing of, 318–326, 319t, 321f, 323f, 325t
    - paired  $t$ -test, 325–326
  - of parameters in two Bernoulli populations, 333–336
  - of population means, hypothesis testing of, 446, 448–459, 452t, 453f, 456t
  - of variance, of two normal populations, 328–329
- Error
  - mean square error of point estimators, 273–279
  - type I, 298, 300
  - type II, 298, 302–305, 304f
- Error sum of squares, 465, 467t
- Estimated regression line, 360
- Estimates
  - defined, 236
  - interval, 235, 247–260
    - confidence interval for normal mean with unknown variance, 252–258, 256f

- confidence interval for variance of normal distribution, 259–260, 260t
- hypothesis testing v., 310
- prediction, 257–258
- for unknown mean, 247–252
- Estimation
  - of life distributions, 245–247
  - of mean of uniform distribution, 244–245
  - of mean response, 411–413
  - of parameters, 235–284
    - approximate confidence interval for mean of Bernoulli random variable, 268–272, 272t
    - Bayes estimator, 236, 279–284
    - confidence interval for mean of exponential distribution, 272–273
    - of difference in means of two normal distributions, 260–268, 262–263f, 266–267f, 268t
    - interval estimates, 235, 247–260, 310
    - introduction, 235–236
    - of life distributions, 245–247
    - maximum likelihood estimators, 235–247, 260, 284, 500
    - point estimator evaluation, 273–279
    - for two-way analysis of variance, 460–463
- Estimators
  - Bayes, 236, 279–284
  - bias of, 274
  - confidence interval
    - of difference in means of two normal distributions, 260–265, 262–263f
    - for mean of exponential distribution, 272
    - of mean response, 377–379, 411–413
  - defined, 236
  - of deviance from grand mean, 462–463
  - of grand mean, 462–463
  - least squares, 385
    - distribution of, 361–367, 368f, 369f
    - in multiple linear regressions, 400–408, 413–414
    - in polynomial regression, 397–399
    - of regression parameters, 359–361, 361f, 362f
    - for Weibull distribution in life testing, 610–612
  - maximum likelihood, 235–247, 279
  - of Bernoulli parameter, 237–240
  - of difference in means of two normal distributions, 260
  - evaluation of, 277–278

Estimators (*continued*)

- for exponential distribution in life testing, 593–594, 602–604
  - least squares estimators as, 365–366
  - for life distributions, 245–247
  - in logistic regression models, 418
  - for mean of exponential distribution, 272
  - of normal population, 242–244
  - of Poisson parameter, 240–241
  - in sequential testing for exponential distribution in life testing, 599, 601
  - for Weibull distribution in life testing, 608–610
  - weighted least squares estimators as, 392
  - point
    - evaluation of, 273–279
    - for hypothesis testing, 299–300
    - of mean response, 377, 411
  - pooled, 261, 321
  - unbiased, 274–279
  - of variance, 447–448
    - for one-way analysis of variance, 446, 448–459, 452t, 453f, 456t
    - for two-way analysis of variance, 464–467
    - for two-way analysis of variance with interaction, 469–476, 475t, 476f
    - weighted least squares, 390–396, 395f
- Evaluation, of point estimator, 273–279
- Events, 54–55
- algebra of, 56–57, 56f, 57f
  - independent, 75–79, 78f
  - odds of, 59
- Expectation, 107–111, 111f
- properties of, 111–118
  - of a random variable function, 113–115
  - of sums of random variables, 115–118
- Expected value. *See* Expectation
- Exponential distribution
- confidence interval for mean of, 272–273
  - gamma distribution and, 188
  - in life testing, 592–606, 598f
    - Bayesian approach, 604–606
    - sequential testing, 598–602, 598f
    - simulation testing with stopping at  $r$ th failure, 592–598
    - simulation testing with stopping by fixed time, 602–604
  - Poisson process and, 181–182
- Exponential random variables, 177–185, 181f
- generation of, 641–642
  - memoryless, 178

moment generating functions and, 177

Poisson process, 181–183, 181f

probability and, 179–180

sample means for, 218–219, 219f

Exponentially weighted moving-average control charts, 573–578, 575f, 578f

**F**

Failure rate. *See* Hazard rate

$F$ -density function, 195–196, 195f

$F$ -distribution, 195–196, 195f, 328–329

Finite populations, sampling distributions from, 223–227

First quartile, 25–27

Fisher, Ronald A., 6–7

Fisher-Irwin test, 335

Fixed margins, contingency tables with, tests of independence in, 505–510, 506f

Fraction defective control charts, 565–567

Frequency interpretation

- of expectations, 108

- probability, 53

Frequency tables and graphs, 10, 10t, 11f, 12f

- frequency histogram, 16

- frequency polygon, 10, 10t, 12f

- relative, 10–14, 13–14f, 13t

- sample mean and, 19–20, 22

- sample median and, 20–22

- sample mode and, 22

$F$ -statistic, in two-way analysis of variance with interaction, 474, 475t

Future response, prediction interval of, 379–381

- in multiple linear regression, 411–416, 413t, 414f, 415f

**G**

Galton, Francis, 6, 372–373

Gamma density, 188, 189f

Gamma distribution, 595, 604–607

Gamma function, 186–187

Gamma random variables, 187–188, 189f

- chi-square distribution and, 191–192, 192f

Gauss, Karl Friedrich, 5–6

Generation

- of random numbers, 620–622

- of random variables, 497, 627, 638–643

Goodness of fit tests, 489–514

- critical region determination by simulation, 496–499, 498f

- introduction, 489–490
  - Kolmogorov–Smirnov goodness of fit test for continuous data, 510–514, 511f
  - tests of independence in contingency tables, 501–505
  - tests of independence in contingency tables having fixed marginal totals, 505–510, 506f
  - tests when all parameters are specified, 489–497, 493t, 498f
  - tests when some parameters are unspecified, 499–501
  - Gosset, W.S., 6
  - Grand mean, 462–463, 469
  - Graunt, John, 4–5, 4t, 5t
  - Grouped data, 14–17, 14t, 15f, 15t, 16f, 18t
- H**
- Halley, Edmund, 5
  - Hardy's lemma, 36
  - Hazard rate, 245, 589
  - Hazard rate functions, 589–592
  - Hill, A.B., 17
  - Histograms, 14t, 15f, 16–17, 18t
    - normal, 31–33, 31f, 32f, 34f
  - Hypergeometric random variables, 156–160
    - Bernoulli random variables and, 157–158
    - binomial random variables and, 158–160, 225–226
    - mean and variance of, 157–158
  - Hypothesis testing, 297–339
    - in Bernoulli populations, 329–336
    - of equality of  $m$  population distributions, 508–509
    - of equality of means of two normal populations, 318–326, 319t, 321f, 323f, 325t
      - case of known variance, 318–320, 319t
      - case of unknown and equal variance, 324–325, 325t
      - case of unknown variance, 320–324, 321f, 323f
    - paired  $t$ -test, 325–326
    - of equality of population means, 446, 448–459, 452t, 453f, 456t
    - of equality of variance of two normal populations, 328–329
    - of independence in contingency tables, 501–505
    - of independence in contingency tables having fixed marginal totals, 505–510, 506f
    - of independence of characteristics of population member, 501–505
    - interval estimates v., 310
    - introduction, 297
    - of mean of normal population, 299–317, 301f, 304f, 311t, 313f, 315f, 317t
      - case of known variance, 299–311, 301f, 304f, 311t
      - case of unknown variance, 311–317, 313f, 316f, 317t
    - for mean of Poisson distribution, 336–339
    - multiple linear regression and, 409–411, 410t
    - of multiple population means, 446–447
    - of probability distribution of sample, 489–499, 493t, 498f
    - of regression parameters
      - $\alpha$ , 376–377
      - $\beta$ , 370–371
    - of regression to mean, 373–374
    - robustness of, 311
    - of row and column interaction, 469–476, 475t, 476f
    - significance levels, 298–299
    - for two-way analysis of variance, 464–468, 467t, 468f
    - for variance of normal population, 327–329
- I**
- Independence, tests of
    - in contingency tables, 501–505
    - in contingency tables having fixed marginal totals, 505–510, 506f
  - Independent events, 75–79, 78f
  - Independent increment assumption, 181
  - Independent random variables, 101–104
    - central limit theorem for, 210–219, 226–227
    - moment generating functions of, 127–128
    - sample mean and variance distribution with, 222
    - sample mean distribution with, 2, 15
    - signed rank test and, 526–527
  - Independent variable. *See* Input variable
  - Indicator random variable, 90–91
    - covariance of, 125–126
    - expectation for, 109
    - variance of, 119–120
  - Individual moment generating functions, 127–128

Individual probability mass function, joint and, 96–99, 98t

Inferential statistics  
   history of, 6–7  
   probability models and, 2–3

Inheritance, regression to mean and, 372–375, 373f, 374f

Input variable, 357–358  
   variation in response to, 382–384, 390–396, 395f

Interaction, two-way analysis of variance with, 446, 469–476, 475t, 476f

Intersection  
   of sample space, 55  
   in Venn diagram, 56–57, 56f, 57f

Interval estimates, 235, 247–260  
   confidence interval for normal mean with unknown variance, 252–258, 256f  
   confidence interval for variance of normal distribution, 259–260, 260t  
   hypothesis testing v., 310  
   prediction, 257–258  
   for unknown mean, 247–252

Inverse transformation method, 639–641

$i$ th order statistic, 594

## J

Joint cumulative probability distribution function, 96, 103–104

Joint density  
   conditional densities and, 107  
   random numbers and, 166–168

Joint density function, 236–237, 242, 244

Joint probability density function, 99–101

Joint probability mass function  
   conditional probability mass function and, 106  
   individual and, 96–99, 98t, 99t

Jointly continuous, 99, 102–103

Jointly distributed random variables, 95–107, 98t, 99t  
   conditional distributions, 104–107  
   independent, 101–104

## K

Kolmogorov's law of fragmentation, 243

Kolmogorov–Smirnov goodness of fit test, for continuous data, 510–514, 511f

Kolmogorov–Smirnov test statistic, 510–514, 511f

Kruskal–Wallis test, 540

## L

Laplace, Pierre-Simon, 5–6

Least squares estimators, 385  
   distribution of, 361–367, 368f, 369f  
   in multiple linear regression, 400–408, 413–414  
   in polynomial regression, 397–399  
   of regression parameters, 359–361, 361f, 362f  
   for Weibull distribution in life testing, 610–612  
   weighted, 390–396, 395f

Left-end inclusion convention, 15

Level of significance. *See* Significance level

Levels, in two-way analysis of variance, 460

Life distributions, estimation of, 245–247

Life testing, 589–612  
   exponential distribution in, 592–606, 598f  
     Bayesian approach, 604–606  
     sequential testing, 598–602, 604f  
     simulation testing with stopping at  $r$ th failure, 592–598  
     simulation testing with stopping by fixed time, 602–604  
   hazard rate functions, 589–592  
   introduction, 589  
   two-sample problem, 606–608  
   Weibull distribution in, 608–612, 609f

Likelihood function, 236–237

Line graph, 10, 10t, 11f

Linear regression equation, 357–358. *See also* Multiple linear regression  
   assessment of, 384–386, 385–386f

Linearity, transforming to, 387–390, 388f, 389f, 389t, 390t

Logarithms, for transforming to linearity, 387–390, 388f, 389f, 389t, 390t

Logistic regression function, 416, 418f

Logistic regression models, for binary output data, 416–419, 418f

Logistics distribution, of random variables, 196–197

Logistics random variable, 197

Logit, 419

Lognormal distribution, 243

Lower confidence interval  
   for difference in means of two normal distributions, 260–264, 261–262f, 267t  
   for normal mean with unknown variance, 255–257, 256f  
   for unknown mean, 249–251



- for unknown probability, 272t
- for variance of normal distribution, 260t
- Lower control limits
  - for exponentially weighted moving-average, 576–578, 578f
  - for fraction defective, 566–567
  - for mean control charts, 555–557, 556f
  - for moving-average, 572–573, 573t, 574f
  - for number of defects, 568–570
  - for variance control charts, 563–564, 564f
- M**
- Mann-Whitney test. *See* Rank sum test
- Marginal probability mass function, 98
- Markov's inequality, random variables and, 128–131
- Mass function. *See* Probability mass function
- Matrix notation
  - for multiple linear regression, 401–403
  - for polynomial regression, 400
- Maximum likelihood estimators, 235–247, 284, 500
  - of Bernoulli parameter, 237–240
  - of difference in means of two normal distributions, 260
  - evaluation of, 277–278
  - for exponential distribution in life testing, 593–594, 602–604
  - least squares estimators as, 365–366
  - for life distributions, 245–247
  - in logistic regression models, 418
  - for mean of exponential distribution, 272
  - of normal population, 242–244
  - of Poisson parameter, 240–241
  - in sequential testing for exponential distribution in life testing, 599, 601
  - for Weibull distribution in life testing, 608–610
  - weighted least squares estimators as, 392
- Mean. *See also* Population means; Sample mean
  - of Bernoulli random variable, confidence interval for, 268–272, 272t
  - of chi-square random variable, 192
  - confidence interval estimators of mean response, 377–379, 411–413
  - estimation of difference in means of two normal distributions, 260–268, 262–263f, 266–267f, 268t
  - of exponential distribution, confidence interval for, 272–273
  - for exponentially weighted moving-average, 575
  - grand, 462–463, 469
  - of hypergeometric random variables, 157–158
  - of least squares estimators, 362–364
  - for moving-average, 572
  - normal, confidence intervals for, 252–258, 256f
  - of normal population, hypothesis testing concerning, 299–317, 301f, 301f, 311t, 313f, 316f, 317t
    - case of known variance, 299–311, 301f, 304f, 311t
    - case of unknown variance, 311–317, 313f, 316f, 317t
  - of normal random variables, 169–170
  - permutation tests and, 634
  - of Poisson distribution, hypothesis testing for, 336–339
  - Poisson distribution with unknown value of, goodness of fit tests for, 499–501
  - population, 208–209, 209f
  - regression to, 372–376, 373f, 374f, 375t, 376f
  - testing equality of means of two normal populations, 318–326, 319t, 321f, 323f, 325t
    - case of known variance, 318–320, 319t
    - case of unknown and equal variance, 324–325, 325t
    - case of unknown variance, 320–324, 321f, 323f
    - paired *t*-test, 325–326
  - of uniform distribution, 244–245
  - of uniform random variables, 162
  - unknown
    - confidence intervals for normal mean with unknown variance, 252–258, 256f
    - control charts for, 557–562, 559t
    - estimates of, 247–252
- Mean control chart, 554–562, 556f, 564f
  - case of unknown, 557–562, 559t
- Mean life, maximum likelihood estimator of, 603–604
- Mean response
  - estimation of, 411–413
  - statistical inferences concerning, 377–379
- Mean square error
  - bootstrap method and, 627–628
  - of point estimators, 273–279

Median, sign test for, 523–524, 524f  
 Memoryless, exponential random variables, 178  
 Modal values, 22  
 Mode, of density, 284  
 Models, assessment of, 384–386, 385–386f  
 Moment generating functions  
   chi-square distribution, 188  
   chi-square random variable, 188–190, 190f  
   exponential random variables and, 177  
   gamma distribution and, 186–188  
   normal random variables and, 175  
   of Poisson random variables, 149–150  
   of random variables, 127–126  
 Monte Carlo simulation, 255–257, 256f,  
   622–623  
   determining runs in, 643–644  
 Moving-average control charts, 571–573, 573t,  
   574f  
   exponentially weighted, 573–578, 575f, 578f  
 Multidimensional integrals, simulation of,  
   255–257, 256f  
 Multiple comparisons, of sample means,  
   456–458  
 Multiple linear regression, 400–416, 403t, 404f,  
   405f, 406f, 410t, 413t, 414f, 415f  
 Multiple probability distributions, equality  
   testing of, 539–541  
 Multiple regression equation, 358  
 Multivariate normal distribution, 404  
 Mutually exclusive, in sample space, 55

## N

*Natural and Political Observations Made upon  
 the Bills of Mortality*, 4–5, 4t, 5t  
 Negatively correlated, 37  
 Neyman, Jerzy, 7  
 95 Percent confidence interval  
   of difference in means of two normal  
     distributions, 260–268, 252–253f  
   for estimating unknown mean, 247–252  
   for mean of exponential distribution,  
     272–273  
   of mean response, 378  
   for normal mean with unknown variance,  
     253–257, 254f  
   for regression parameters, 372  
   for unknown probability, 269–271  
 95 Percent prediction interval, 416  
 99 Percent confidence interval  
   for estimating unknown mean, 250–251  
   for unknown probability, 271  
 90 Percent confidence interval  
   of difference in means of two normal  
     distributions, 265–268, 266–267f  
   for variance of normal distribution, 259–260  
 Nonparametric hypothesis tests, 521–545  
   introduction to, 521  
   runs test for randomness, 541–545, 543f  
   sign test, 521–525, 524f  
   signed rank test, 525–531, 526f  
   two-sample problem, 531–535, 534f, 538f  
     classical approximation and simulation,  
       535–539  
     equality of multiple probability  
       distributions, 539–541  
 Nonparametric inference, 207–208  
 Nonrandom sample, 3  
 Normal approximations, in permutation tests,  
   633–637  
 Normal data sets, 34, 31f, 32f, 34f  
 Normal density function, 168, 168f, 188, 189f  
 Normal distribution  
   confidence interval for variance of, 259–260,  
     260t  
   estimation of difference in means of,  
     260–268, 262–263f, 266–267f, 268t  
 Normal equations  
   in multiple linear regression, 401–403  
   in polynomial regression, 397  
   of regressions, 359–360  
 Normal histograms, 33, 31f, 32f, 34f  
 Normal mean, with unknown variance,  
   confidence intervals for, 252–258, 256f  
 Normal populations  
   maximum likelihood estimator of, 242–244  
   mean of  
     hypothesis testing concerning, 299–317,  
       301f, 304f, 311t, 313f, 316f, 317t  
     testing equality of means of two normal  
       populations, 318–326, 319t, 321f, 323f,  
       325t  
   sampling distributions from, 220–223  
     joint distribution, 221–223  
     sample mean distribution, 221  
   variance of, hypothesis testing for, 327–329  
 Normal prior, choosing of, 283–284  
 Normal random variables, 168–177, 172f, 176f  
   chi-square distribution, 188–193, 190f, 192f  
   *F*-distribution, 195–196, 195f  
   generation of, 642–643  
   mean and variance of, 169–170  
   normal density function, 168, 168f

- standard normal distribution and, 171–172, 172f, 175–177, 176f
- sums of, 175
- $t$ -distribution, 193–195, 193f, 195f
- Notation
  - dot, in two-way analysis of variance, 461–462
  - for least squares estimators, 366
  - matrix
    - in multiple linear regression, 401–403
    - for polynomial regression, 400
- Null hypothesis, 298
  - permutation tests and, 631–633
- Number of defect control charts, 567–570, 569t
- O**
  - Observational study, 336
  - OC curve. *See* Operating characteristic curve
  - Odds for success, 419
  - Odds of event, 59
  - Ogives, 14t, 16–17, 16f, 18t
  - $100(1 - \alpha)$  Percent confidence interval of
    - difference in means of two normal distributions, 261–265, 268t
    - for estimating unknown mean, 250–251
    - for exponential distribution in life testing, 596
    - for mean of exponential distribution, 272
    - of mean response, 378
    - for normal mean with unknown variance, 252–258
    - for regression parameters
      - $\alpha$ , 377
      - $\beta$ , 371–372
    - in sequential testing for exponential distribution in life testing, 599–600
    - for unknown probability, 269–271
    - for variance of normal distribution, 260t
  - $100(1 - \alpha)$  Percent confidence region, 263
  - $100(1 - \alpha)$  Percent prediction interval, 381, 416
  - One-sided Chebyshev's inequality, 29–30
  - One-sided critical region, 307
  - One-sided hypothesis tests
    - for mean of normal population, case of known variance, 306–310
    - for testing equality of means of two normal populations, 321
  - One-sided lower confidence interval
    - of difference in means of two normal distributions, 260–264, 261–262f, 267t
    - for normal mean with unknown variance, 255–257, 256f
    - for unknown mean, 249–251
    - for unknown probability, 272t
    - for variance of normal distribution, 260t
  - One-sided null hypothesis, sign test and, 524–525
  - One-sided  $t$ -tests, for mean of normal population with unknown variance, 314–317, 316f
  - One-sided upper confidence interval
    - for difference in means of two normal distributions, 261, 265–268, 266–267f
    - for normal mean with unknown variance, 254–255
    - for unknown mean, 249–251
    - for unknown probability, 272t
    - for variance of normal distribution, 260t
  - One-way analysis of variance, 446, 448–459, 452t, 453f, 456t
    - multiple comparisons of sample means, 456–458
    - with unequal sample sizes, 458–459
  - Operating characteristic (OC) curve, 303–304, 304f
    - for one-sided hypothesis testing for mean of normal population, 307–308
  - Out of control, 553, 555–557, 556f
  - Overlook probabilities, 74
- P**
  - Paired data sets, 34–36, 34t, 35f
    - sample correlation coefficient and, 37–40, 39f
  - Paired  $t$ -test, 325–326, 523
  - Parameter estimation, 235–284
    - approximate confidence interval for mean of Bernoulli random variable, 268–272, 272t
    - Bayes estimator, 236, 279–284
    - confidence interval for mean of exponential distribution, 272–273
    - of difference in means of two normal distributions, 260–268, 262–263f, 266–267f, 268t
    - interval estimates, 235, 247–260, 310
    - introduction, 235–236
    - of life distributions, 245–247
    - maximum likelihood estimators, 235–247, 260, 284
    - point estimator evaluation, 273–279

- Parameter estimation (*continued*)
  - for two-way analysis of variance, 460–463
  - for Weibull distribution in life testing, 610–612
- Parametric inference, 207–208
- Pearson, Egon, 7
  - The Pareto Distribution, 183–185
- Pearson, Karl, 6, 373, 496
- Permutation, 61–63
- Permutation tests, 630–638
  - implementation of, 631–632
  - normal approximations in, 633–637
  - null hypothesis and, 631–633
  - two sample, 637–638
- Pie chart, 12, 13–14f
- Point estimates, 235
- Point estimators
  - evaluation of, 273–279
  - for hypothesis testing, 299–300
  - of mean response, 377, 411
- Point prediction, 413
- Poisson distribution
  - hypothesis testing concerning mean of, 336–339
  - with unknown mean, goodness of fit tests for, 499–501
  - variance in, 395–396
- Poisson distribution function
  - computation of, 155–156
  - number of defects and, 567–570, 569t
- Poisson parameters
  - maximum likelihood estimator of, 240–241
  - testing of relationship between, 337–339
- Poisson probability mass function, 148–150, 149f, 154–155
- Poisson process, exponential random variables and, 181–183, 181f
- Poisson random variables, 148–156, 149f
  - binomial random variables and, 150–153
  - conditional probability and, 153–154
  - moment generating functions of, 149–150
  - probability mass function and, 148–150, 149f, 154–155
- Poisson, S.D., 148
- Polynomial regression, 397–400, 398f, 399f
- Pooled estimator, 266, 321
- Population distributions
  - empirical distribution and, 624
  - equality of, hypothesis testing for, 508–509
  - signed rank test for, 529–531
- Population means, 208–209, 209f
  - bootstrap method and, 623–624, 628–630
  - confidence interval for difference in, 456–458
  - control charts for, 571–581
  - cumulative sum, 579–581
  - exponentially weighted moving-average, 573–578, 575f, 578f
  - moving-average, 571–573, 573t, 574f
  - hypothesis testing of equality of, 446, 448–459, 452t, 453f, 456t
  - multiple, hypothesis testing of, 446–447
- Population median, sign test for, 523–524, 524f
- Population variance, 208–209
  - bootstrap method and, 623–625
- Populations
  - definition of, 207
  - samples and, 3
  - sampling distributions from
    - finite, 223–227
    - normal, 220–223
- Positively correlated, 37
- Posterior density function, 280, 284
- Power-function, of hypothesis test, 304
- Prediction interval
  - confidence interval v., 381
  - of future response, 379–381
  - of response at input level  $x_0$ , 381
  - of response in multiple linear regression, 411–416, 413t, 414f, 415f
- Prior distributions, 279–284, 604–606
- Probability, 53–79
  - axioms of, 57–59, 59f
  - Bayes' formula, 68–75, 69f
  - Bernoulli random variables, 141–148
  - binomial random variables, 143–147
  - bootstrap method and, 629–630
  - central limit theorem, 210–219, 212–215f, 216f
  - chi-square distribution and, 187
  - conditional, 65–68, 66f, 106
  - continuous random variable and, 94
  - counting and, 60–65
  - of defects, 329–337
  - distribution function and, 91–92
  - events, 54–55
    - independent, 75–79, 78f
  - expectation, 107–111, 111f
  - exponential random variables and, 179–180
  - fraction defective, 565–567
  - introduction to, 53–54
  - overlook, 74

- Poisson random variables and, 148–156, 149f
  - of random variables, 89–90
  - rank sum tests, 532–533
  - sample space, 54–55
    - with equally likely outcomes, 59–65
  - signed rank test and, 527–529
  - of uniform random variables, 160–161, 161f
  - unknown, confidence interval for, 268–272, 272t
  - Venn diagram and algebra of events, 56–57, 56f, 57f
  - Probability density function, 93–94
    - cumulative distribution function and, 94–95, 94f
    - exponential random variables and, 177
    - joint, 99–101
    - of uniform random variables, 160–161, 161f
    - updated, 280
  - Probability distribution
    - of estimator of mean response, 377–378
    - of sample, goodness of fit tests for, 489–499, 493t, 498f
  - Probability distribution function
    - joint cumulative, 96, 103–104
    - Poisson, 148–156
    - of populations, 207
    - random variable and expectation, 111–113
    - signed rank test for, 529–531
  - Probability mass function, 92–93, 93f, 245–247
    - Bernoulli random variables, 142–144, 143f
    - binomial random variables, 142–144, 143f
    - central limit theorem and, 212–215f
    - conditional, 105–106
    - discrete random variables, 92–93, 638–640
    - expectation of, 107–108
    - hypergeometric random variables, 156–157
    - individual and joint, 96–99, 98t
    - marginal, 98
    - Poisson, 148–150, 149f, 154–155
    - Poisson random variables, 148–150, 149f, 154–155
  - Probability models, inferential statistics and, 2–3
  - Probability theory, statistics and, 5–6
  - Probit model, 419
  - Pseudo random numbers, 257, 620
  - p*-value
    - for determining independence of characteristics of population member, 504–505
    - for goodness of fit tests when all parameters are specified, 491, 494–499, 498f
    - for goodness of fit tests when some parameters are unspecified, 500–501
    - for hypothesis testing
      - in Bernoulli populations, 330–333, 335
      - of equality of population means, 452, 453f
      - of mean of normal population, 302, 307–310, 311t, 313, 315–317, 316f, 317t
      - of mean of Poisson distribution, 337–338
      - with multiple linear regression, 411
      - of regression parameters,  $\beta$ , 370
      - of regression to mean, 374
      - of variance of normal population, 327, 329
    - for Kolmogorov–Smirnov goodness of fit test, 513
    - for one-sided hypothesis testing for mean of normal population, 307–310
    - permutation tests for, 630–638
      - normal approximations in, 633–637
      - two sample, 637–638
    - rank sum test and, 533–535, 534f
    - in sequential testing for exponential distribution in life testing, 601–602
    - signed rank test for, 527–529
    - simulation for approximation of, 496–499, 498f
    - for testing equality of means of two normal populations, 321–324, 323f, 325t, 326
    - in two-way analysis of variance, 467, 467t, 468f, 473–474, 475t, 476f
- Q**
- Quality control, 553–581
    - fraction defective control charts, 565–567
    - introduction to, 553–554
    - mean control chart, 554–562, 556f, 564f
    - number of defect control charts, 567–570, 569t
    - population mean control charts, 571–581
      - cumulative sum, 579–581
      - exponentially weighted moving-average, 573–578, 575f, 578f
      - moving-average, 571–573, 573t, 574f
    - variance control chart, 562–565, 564f

**R**

- Random error, in response to input variable, 357–358, 361
- Random numbers, 620–623
  - definition of, 163, 163t
  - generation of, 620–622
  - Monte Carlo simulation approach, 622–623
  - pseudo, 620
  - use of, 164–166, 166f
- Random sample, 3, 207, 223
  - runs test for, 541–545, 543f
- Random variables, 89–92. *See also specific random variables*
  - Bernoulli and binomial, 141–148, 143f, 148f
  - central limit theorem, 210–219, 212–215f
  - Chebyshev's inequality, 128–131
  - continuous, 93–94, 640–643
  - density function and, 112
  - discrete, 91–92, 638–640
  - distribution function and, 91–92, 624
  - entropy of, 110
  - expectation of function of, corollary of, 114–115
  - expected value of sums of, 115–118
  - exponential, 177–182, 181f, 641–642
  - gamma distribution of, 186–188, 189f
  - generation of, 497, 627, 638–643
  - hypergeometric, 156–160
  - indicator, 90–91
  - jointly distributed, 95–107, 98t
    - conditional distributions, 104–107
    - independent, 101–104
  - logistics distribution, 196–197
  - Markov's inequality, 128–130
  - moment generating functions, 127–128
  - normal, 168–177, 172f, 176f, 642–643
    - chi-square distribution, 188–193, 190f, 192f
    - $F$ -distribution, 195–196, 195f
    - $t$ -distribution, 193–195, 193f, 194f
  - Poisson, 148–156, 149f
  - probability distribution function and expectation, 111–113
  - types of, 92–95, 93f, 94f
  - uniform, 160–168, 161f, 163t, 166f
  - variance of, 118–120, 162, 169–170, 192, 222, 447–448
  - variance of a sum of, 123–126
  - weak law of large numbers, 130
- Rank sum test, 521, 531–541, 534f, 538f
  - distribution function of, 531
  - Kruskal–Wallis* test, 540
  - probability and, 532–533
  - $p$ -value and, 533–535, 534f
    - classical approximation and simulation, 535–539, 538f
- Rate of distribution, 590
- Rayleigh density function, 591
- Recursive formula, mean control chart and, 559–560, 559t
- Referents, 334
- Regression, 357–420
  - analysis of residuals and assessing models, 384–386, 385–386f
  - coefficient of determination and sample correlation coefficient, 382–384
  - distribution of least squares estimators, 361–367, 368f, 369f
  - history of, 6
  - introduction, 357–358, 358f
  - least squares estimators of regression parameters, 359–361, 361f, 362f
  - logistic regression models for binary output data, 416–420, 418f
  - to mean, 372–376, 373f, 374f, 375t, 376f
  - multiple linear, 400–416, 402t, 404f, 405f, 406f, 410t, 414f, 415f
    - predicting future responses, 411–416, 413t, 414f, 415f
  - polynomial, 397–400, 398f, 399f
  - statistical inferences about regression parameters, 367–382, 373f, 374f, 375t, 376f
    - $\alpha$ , 376–377
    - $\beta$ , 368–376, 373f, 374f, 375t, 376f
    - mean response, 377–379
    - prediction interval of future response, 379–381
    - summary of distribution results, 381–382
  - transforming to linearity, 387–390, 388f, 389f, 389t, 390t
  - use of dummy variables, 416–418
  - weighted least squares, 390–396, 395f
- Regression coefficients, 358, 397
- Regression fallacy, 376
- Regression parameters
  - least squares estimators of, 359–361, 361f, 362f
  - statistical inferences about, 367–382, 373f, 374f, 375t, 376f
    - $\alpha$ , 376–377
    - $\beta$ , 368–376, 373f, 374f, 375t, 376f
    - mean response, 377–379

prediction interval of future response,  
     379–381  
 summary of distribution results, 381–382  
 Rejection, of hypothesis. *See* Hypothesis testing  
 Relative frequency tables and graphs, 10–14,  
     13–14f, 13t, 16  
 Residuals, 364–366  
     analysis of, 384–386, 385–386f  
     in multiple linear regression, 408–409, 411  
     standardized, 385–386, 385–386f  
 Response variable, 357–358  
     prediction interval of future response,  
         379–381  
     in multiple linear regression, 411–416,  
         413t, 414f, 415f  
     variation in, 382–384  
         with input variable, 390–396, 395f  
 Robustness, of hypothesis test, 311  
 Row factors  
     hypothesis testing for, 464–468, 467t, 468f  
     in two-way analysis of variance, 460  
         column factor interaction with, 446,  
         469–476, 475t, 476f  
         deviation from grand mean due to, 462  
 Row sum of squares, 466, 467t  
 Run, 541  
 Runs test for randomness, 521, 541–542, 543f

## S

Sample  
     definition of, 207  
     populations and, 3  
 Sample 100 $p$  percentile, 24–25  
 Sample correlation coefficient, 37–41, 39f  
     association v. causation, 40–41  
     coefficient of determination and, 382–384  
     properties of, 37–38, 41  
 Sample mean, 17–20, 22  
     central limit theorem for, 216–218  
     distribution of, with chi-square random  
         variables, 222  
     for exponential random variables, 218–219,  
         219f  
     for independent random variables, 2, 15  
     multiple comparisons of, 456–458  
     of normal data set, 31  
     for normal population, 220–221  
     population, 208–209, 209f  
     sample variance distribution with, 221–223  
 Sample median, 20–22, 31  
 Sample mode, 22  
 Sample percentiles, 24–25  
 Sample quartiles, 25–27, 27f  
 Sample size, one-way analysis of variance with  
     unequal sample sizes, 458–459  
 Sample spaces, 54–55  
     having equally likely outcomes, 59–65  
 Sample standard deviation, 24, 219–220  
 Sample variance, 22–24, 219–220  
     for normal population, 220  
     sample mean distribution with, 221–223  
 Sampling, 207  
 Sampling distributions  
     form finite populations, 223–227  
     form normal populations, 220–223  
         joint distribution, 221–223  
         sample mean distribution, 221  
 Scatter diagram, 34–36, 36f, 358, 358f, 373,  
     385–386, 385–386f, 397  
 Second quartile, 25–27  
 Selection, of normal prior, 283–284  
 Sequence of interarrival times, 182  
 Sequential testing, for exponential distribution  
     in life testing, 598–602, 598f  
 Sign test, 521–525, 524f  
     Bernoulli random variables, 522  
     binomial random variables, 522  
     one-sided null hypothesis and, 524–525  
     paired  $t$ -test v., 523  
     for population median, 523–524, 524f  
 Signed rank test, 521, 525–531, 526f  
     for distribution function, 525, 526f  
     for probability distribution function,  
         529–531  
     for  $p$ -value, 527–529  
 Significance level, 298–299  
 Significance level  $\alpha$  test  
     for determining independence of  
         characteristics of population member,  
         503–505  
     for goodness of fit tests when all parameters  
         are specified, 496  
     for hypothesis testing  
         in Bernoulli populations, 329–333  
         of equality of population means, 452, 459  
         of mean of normal population, 300–302,  
         307–309, 311t, 312, 314–316, 317t  
         of mean of Poisson distribution, 336–337  
         of regression to mean, 374  
         of variance of normal population, 329  
     for Kolmogorov–Smirnov goodness of fit test,  
         513–514

Significance level  $\alpha$  test (*continued*)  
   for testing equality of means of two normal populations, 318–321, 324–325, 325t  
   in two-way analysis of variance, 467, 467t, 472–474, 475t  
 Simple hypothesis, 298  
 Simple regression equation, 358, 362f, 369f  
   assessment of, 384–386, 385–386f  
 Simulation  
   for determination of critical region, 496–499, 498f  
   of single and multidimensional integrals, 251–253, 256f  
 Simulation run, 623  
   in Monte Carlo study, 643–644  
 Simulation testing, for exponential distribution  
   in life testing, 592–598, 602–604  
 Single integrals, simulation of, 255–257, 256f  
 Skewed data set, 31, 32f  
 Skewed random variables, 142, 143f  
 Standard deviation  
   definition of, 121  
   mean control chart and, 560–561  
   variance control chart, 562–563  
 Standard logistic, 197  
 Standard normal distribution, 171–172, 172f, 175–177, 176f  
   central limit theorem and, 217  
   of mean control chart, 555  
    $t$ -distribution and, 192–195, 193f, 194f  
 Standard normal random variable, 221  
   central limit theorem and, 210, 216–217  
 Standardized residuals, 385–386, 385–386f  
 Stationary increment assumption, 181–182  
 Statistical analysis, 1  
 Statistical inferences, about regression  
   parameters, 367–382, 373f, 374f, 375t, 376f  
    $\alpha$ , 376–377  
    $\beta$ , 368–376, 373f, 374f, 375t, 376f  
   mean response, 377–379  
   prediction interval of future response, 379–381  
   summary of distribution results, 381–382  
 Statistical theory, 1  
 Statistics  
   application of, 6–7  
   definition of, 1, 6–7, 6t, 207–208  
   descriptive, 1–2  
   history of, 3–7, 4t, 5t, 6t

  inferential, 2–3  
   introduction to, 1–7  
 Stem and leaf plots, 16–17, 18t  
   of normal data set, 33  
   sample mean and, 21  
   sample median and, 21  
 Subjective interpretation, probability, 53  
 Success, odds for, 419  
 Sum of squares  
   column, 467t  
   error, 465, 467t  
   row, 466, 467t  
   between samples, 451–452, 454, 456t, 459  
   within samples, 450, 454, 456t, 459  
   in two-way analysis of variance with  
     interaction, 471–474, 475t  
 Sum of squares identity, 453–454  
 Sum of squares of residuals, 364–366, 408–409, 411  
 Survival rate, 245–246

## T

$t$ -density function, 193, 193f, 252, 253f  
 $t$ -distribution, 193–195, 193f, 194f, 223  
 Test statistic  
   for determining independence of  
     characteristics of population member, 503–505  
   for goodness of fit tests when all parameters are specified, 490–492, 494–496  
   for goodness of fit tests when some parameters are unspecified, 500–501  
   for hypothesis testing  
     in Bernoulli populations, 330  
     of equality of population means, 451–452, 455, 456t  
     of mean of normal population, 302, 307, 309, 315t, 312–315, 317t  
     of regression parameters, 370–371  
     of regression to mean, 374  
     of variance of normal population, 327, 329  
 Kolmogorov–Smirnov, 510–514, 511f  
   for one-sided hypothesis testing for mean of normal population, 307, 309  
   for testing equality of means of two normal populations, 320–322, 324, 325t, 326  
   for testing independence in contingency tables, 506–507, 506f  
   in two-way analysis of variance, 467, 467t  
 Testing. *See* Goodness of fit tests; Hypothesis testing; Life testing



- Tests of independence
    - in contingency tables, 501–505
    - in contingency tables having fixed marginal totals, 505–510, 506f
  - Third quartile, 25–27
  - Threshold model, 420
  - Ties
    - rank sum test and, 531
    - signed rank test and, 530–531
  - T*-method, 456–458
  - Total-time-on-test statistic, 594–595, 604
  - T*-random variable, 264
  - Transformation, to linearity, 387–390, 388f, 389f, 389t, 390t
  - Treatment group, 164
  - Tree diagram, random numbers and, 166, 166f
  - t*-tests, 311–317, 313f, 316f, 317t
    - one-sided, 314–317, 316f
    - paired, 325–326
    - p*-value of two-sample, 323f
    - two-sided, 311–313, 313f
  - Two sample permutation tests, 637–638
  - Two-sample problem, 521, 531–539, 534f, 538f
    - distribution function of, 531
    - in life testing, 606–608
    - probability and, 532–533
    - p*-value and, 533–535, 534f
      - classical approximation and simulation, 535–539, 538f
  - Two-sided confidence interval, 248, 251
    - of difference in means of two normal distributions, 261–268, 262–263f, 266–267f
    - for normal mean with unknown variance, 255–257, 256f
    - for unknown probability, 272t
  - Two-sided *t*-tests, for mean of normal population with unknown variance, 311–313, 313f
  - Two-way analysis of variance, 446
    - hypothesis testing for, 464–468, 467t, 468f
    - with interaction, 446, 469–476, 475t, 476f
    - introduction and parameter estimation, 460–463
  - Type I errors, 298, 300
  - Type II errors, 298, 302–305, 304f
- U**
- Unbalanced case, in one-way analysis of variance, 459
  - Unbiased estimators, 274–279
  - Uniform distribution, estimating mean of, 244–245
  - Uniform random variables, 160–168, 161f, 163t, 166f
    - mean and variance of, 162
    - probability density function of, 160–161, 161f
    - random numbers, 166–168
  - Union
    - of sample space, 55
    - in Venn diagram, 56–57, 56f, 57f
  - Unit normal distribution. *See* Standard normal distribution
  - Unknown mean
    - confidence intervals for normal mean with unknown variance, 252–257, 256f
    - estimates of, 247–252
  - Unknown parameters. *See* Parameter estimation
  - Unknown probability, confidence interval for, 268–272, 272t
  - Unknown variance
    - confidence intervals for normal mean with, 252–257, 256f
    - hypothesis testing for mean of normal population with, 311–317, 313f, 316f, 317t
    - testing equality of means of two normal populations with, 320–325, 321f, 323f, 325t
  - Updated probability density function, 280
  - Upper confidence interval
    - for difference in means of two normal distributions, 255, 265–266, 266–267f
    - for normal mean with unknown variance, 254–255
    - for unknown mean, 249–251
    - for unknown probability, 272t
    - for variance of normal distribution, 260t
  - Upper control limits
    - for exponentially weighted moving-average, 576–578, 578f
    - for fraction defective, 566–567
    - for mean control charts, 555–557, 556f
    - for moving-average, 572–573, 573t, 574f
    - for number of defects, 568–570
    - for variance control charts, 563–564, 564f

## V

Variance, 118–121. *See also* Analysis of variance;  
     Population variance; Sample variance  
   of chi-square random variable, 192, 222,  
     447–448  
   covariance, 121–122  
   definition of, 119  
   distribution of, with chi-square random  
     variables, 222  
   estimators of, 447–448  
     for one-way analysis of variance, 446,  
       448–459, 452t, 453f, 456t  
     in two-way analysis of variance, 464–467  
     in two-way analysis of variance with  
       interaction, 469–476, 475t, 476f  
   for exponentially weighted moving-average,  
     575–576  
   of hypergeometric random variables,  
     157–158  
   of independent random variables, 222  
   of indicator random variable, 119–121  
   known  
     equality of means of two normal  
       populations with, 318–320, 319t  
     hypothesis testing for mean of normal  
       population with, 299–311, 301f, 304f,  
       311t  
   of least squares estimators, 362–364,  
     405–407  
   for moving-average, 572  
   of normal distribution, confidence interval  
     for, 259–260, 259t  
   of normal population, hypothesis testing for,  
     327–329  
   of normal random variables, 169–170

    permutation tests and, 634–635  
     population, 208–209  
   of random variables, 118–121, 123–126,  
     162, 169–170, 193, 222, 447–448  
   in response to input variable, 382–384,  
     390–396, 395f  
   sample, 22–24, 219–220  
   of a sum of random variables, 124–126  
   of uniform random variables, 162  
   unknown  
     confidence intervals for normal mean with,  
       252–257, 256f  
     equality of means of two normal  
       populations with, 320–324, 321f, 323f  
     hypothesis testing for mean of normal  
       population with, 311–317, 313f, 316f,  
       317t  
     unknown and equal, testing equality of  
       means of two normal populations with,  
       324–325, 325t  
 Variance control chart, 562–565, 564f  
 Venn diagram, 56–57, 56f, 57f  
   probability axioms and, 58–59, 59f

## W

Weak law of large numbers, 130  
 Weibull density function, 608, 609f  
 Weibull distribution, in life testing, 608–612,  
   609f  
 Weighted average, 19–20  
 Weighted least squares estimators, 390–396,  
   395f  
 Wilcoxon test. *See* Rank sum test  
 Within samples sum of squares, 450, 454, 456t,  
   459