

Index

- Across-groups (*see* Among-groups)
- Across-subjects sum of squares, 555
- Across-treatments
 - mean squares, 556, 557 (Box 13.2), 559
 - sum of squares, 555, 557 (Box 13.2)
- Addition theorem, 161
 - for mutually exclusive events, 162
- Advanced Dungeons & Dragons*[®] (*AD&D*[®])
 - described, 164 (Exercise 5.1.11)
- Algebra of Expectations, 192–193, 696–699
- Algebra of Summations, 23–24, 682–684
- Algebra of Variances, 196, 699–702
- Alternative hypothesis, 328–329
 - analysis of variance, 490, 535, 568
 - for regression, 597, 615
 - composite, 332–334
 - directional, 333
 - effects on β and power, 364–366, 368, 393, 395
 - as “experimental” hypothesis, 333*n*
 - as hypothesis favored by experimenter, 333, 366
 - nondirectional, 333
 - for post hoc comparisons, 525, 535
 - simple, 333
 - synthetic, 368, 395, 437
 - for test of
 - entire distribution (Pearson’s *chi*-square), 619, 620, 637, 645
 - one correlation, 584, 585, 587, 614
 - one mean, 384, 391, 395
 - one proportion, 332–333
 - one variance, 462, 480
 - two correlations, 589, 614
 - two means 418, 438 (correlated groups)
 - two variances, 471, 480
- Among-groups (between-groups), 492, 501–502
 - mean squares, 510 (Box 12.2), 511–512, 559
 - sum of squares, 506–507 (Box 12.1), 507–509, 512, 555
- Analysis of variance
 - alternative hypothesis for, 490, 535, 568
 - assumptions underlying *F* test, 498–500, 504, 535
 - linear model for, 500–504, 543
 - one-way, 539; *see also specific topics in Table of Contents, Chapter 12*
 - regression, 598–607
 - repeated measures, 539, 554–559
 - simple, 539, 554
 - test hypothesis for, 382, 490, 535, 554, 556, 568
 - test statistic(s) for, 490–495, 513, 535, 547–549, 552–553, 556, 568
 - two-way, 539–554
 - fixed-effects, 544
 - interaction in, 542–543
 - linear model for, 543
 - mixed-effects, 544, 553
 - random-effects, 544, 551–553
- Analysis of variance summary table, 515–516, 551, 553–554
- ANOVA (*see* Analysis of variance)
- Association, tests of, 635–648, 679
 - degrees of freedom for, 642–643, 645
 - expected frequencies in, 638–640, 640–642, 645–647
 - test statistic, 640, 643, 645, 679
 - see also* Homogeneity of proportions, Pearson’s *chi*-square test of; Independence, Pearson’s *chi*-square test of
- Bar graph, 11–12; *see also* Histogram
- Between-groups (*see* Among-groups)
- Biased estimator (of variance), 302–303 (Box 8.2)
- Binomial coefficient, 207

- Binomial distribution, 203–214
 - normal approximation to, 251–256
 - parameters of, 220
 - Poisson approximation to, 745
 - table of probabilities, use of, 217
- Binomial expansion, 207–208
- Binomial experiment, 204–205
 - characteristics of, 204
- Binomial probability, 205–209
- Binomial proportion, 212–214
 - normal approximation to distribution of, 257–260
- Binomial random variable, 205
 - expected value of, 210
 - standardized, 213
 - table of probabilities, use of, 217
 - variance of, 211
- Binomial theorem, 208
- Bivariate distributions, 572
 - normal, 578–580
 - see also* Joint distributions
- Calculating class limits, 7–9
- Central Limit Theorem, 284
 - and binomial distribution, 285 (Box 8.1)
 - and *chi*-square distribution(s), 285 (Box 8.1), 459
 - and sampling distribution of mean, 283–287
 - and Student's *t* distribution(s), 405
- Central tendency (location), 26–27
- Chi*-square distribution(s), 455–461
 - and normal distribution, 459
 - parameter(s) of, 456
 - table of cumulative probabilities, use of, 459–460
- Chi*-square random variables, 455–456
 - addition and subtraction of, 460–461
 - expectation and variance of, 456–457
- Chi*-square tests, Pearson's
 - assumptions underlying, 656–657
 - degrees of freedom for, 624, 642–643
 - see also* Pearson's *chi*-square tests
- Class limits, calculating, 7–9
- Class width (interval width), 7
- Coefficient(s)
 - binomial, 207
 - confidence, 308
 - correlation, 110–117
 - determination, 120–121, 591, 597
 - kurtosis, 68
 - regression, 128, 596; *see also* Appendix III, 688–691
 - skewness (*Sk*), 65–66
 - Pearson's (*Sk_p*), 64–65
 - variation, 58–59
- Combinations, 173–175
 - defined, 173
 - formula for number of, 174
- Composite hypothesis, 332–334
- Conditional distributions
 - frequency, 118–119, 131–134
 - mean(s), 118–119, 131
 - range, 119
 - variance(s), 119–121, 132–134
 - probability, 573–574, 578–580
- Conditional probability, 157, 573–574
- Confidence
 - cumulative probability and, 308
 - defined, 307
 - statistical significance and, 340–341
- Confidence coefficient, 308–309
- Confidence interval(s)
 - defined, 305
 - mean
 - using normal distribution, 306–314
 - using Student's *t*, 311*n*, 413 (Box 10.3)
 - proportion (*p*), 305, 314–318
 - variance, 467 (Box 11.2)
 - see also* Estimation, interval
- Consistency, 298–299
 - of mean, 300–301
 - of relative frequency (sample proportion), 301
 - of variance (*S*²), 305
- Continuity, correction for
 - binomial, normal approximation to, 370–372
 - Pearson's *chi*-square, 648*n*
- Continuous random variables, 188, 225, 228, 577–580
 - representations of probability distributions, 268–269, 577–578
 - see also* Discrete random variable(s)
- Contrast, 524; *see also* Scheffé's Multiple Comparison Test
- Correlation
 - and causation, 115, 592
 - fourfold point, 664–666, 680
 - Pearson product-moment
 - computational formula, 113
 - data, 110–117
 - perfect, 115, 116–117 (Box 4.2)
 - population, 583–593
 - see also* Hypothesis tests, one correlation; Hypothesis tests, two correlations

842 Index

- Correlation coefficient (*see* Correlation, Pearson product-moment)
- Correlation ratio, 119–121, 590–592, 597
- Counting, Principle of Multiplication in, 165–166
- Covariance, 107–111, 115, 197
- Covariation, linear versus nonlinear, 117, 120–121, 593
- Critical region (*see* Rejection region)
- Critical value, 330
 and adjustments for continuity, 370–372
 and p -value, 357, 359
- Cumulative frequency
 defined, 5
 distribution(s), 5–6, 12–13
- Cumulative probability
 defined, 231
 tables, use of
 chi-square, 459–460
 Fisher's F , 470–471
 standard normal random variable, 244–250
 Student's t , 406–408
- Data, 3–4
 considerations in collection of, 352–356
 experimental versus observational, 352–353, 704–705
 exploratory analysis of, 147, 331, 353
 graphic representation of, 10–13
 tabular representation of, 4–10
- Decision theory, 350–352
- Degrees of freedom (df), 403–404
 chi-square, 455–456
 additive properties of, 461
 Fisher's F , 470
 analysis of variance (*see* Degrees of freedom, sums of squares)
 hypotheses about two variances, 472
 hypotheses about one variance, 463
 linear regression (*see* Degrees of freedom, sums of squares)
 number of free observations and, 404
 Pearson's *chi*-square tests
 association, 642–643, 645
 goodness-of-fit, 624
 Scheffé's Multiple Comparison Test, 526, 535
- Student's t , 404–405, 431, 433
 Behrens-Fisher estimate for unequal sample sizes, 433–434
 Welch's correction for unequal sample sizes, 434
- sums of squares
 across subjects, 555
 across treatments (in repeated measures ANOVA), 555
 among groups, 509
 columns, 546
 deviations from linearity, 602
 error (in two-way ANOVA), 545
 error (in ANOVA for regression), 601
 interaction, 546–547
 linear, 602
 pooled error, 553, 568
 residual (in repeated measures ANOVA), 555
 rows, 546
 within groups, 509
 within subjects, 555
 unconstrained cells and, 624, 642–643
- Density function (*see* Function[s], probability density)
- Dependent variable (in experimental design), 704
 predicting value of (in regression), 126–131
- Descriptive statistics, 19–24
 defined, 19
 desirable properties of, 20–21
 notational conventions for, 277
 see also specific descriptive statistics, e.g., Mean, and descriptive properties, e.g., Location
- Determination, coefficient of, 120–121, 591, 597
- Deviations (deviation scores), 44, 90–91, 107–109
 absolute, 51–52
- Deviations from linear regression, 599–600, 602
- df (*see* Degrees of freedom)
- Directional hypothesis, 333–334
- Discrete random variable(s), 187–188
 representation of probability distributions, 217–220, 575–577
 see also Continuous random variables
- Dispersion (variability), 48–49
 captured by Pearson's *chi*-square tests, 617
 probability distributions, 189, 194, 242
 statistics, 48–59; *see also specific statistics, e.g., Standard deviation*
- Distribution(s)
 bivariate, 572; *see also* Joint distributions
 center of mass captured by mean, 42–44

- conditional (*see* Conditional distributions)
 frequency, 5
 cumulative, 5–6, 12–13
 grouped, 6–9, 12
 relative, 5, 12
 joint (*see* Joint distributions)
 marginal (*see* Marginal distributions)
 population versus sample, 275–278
 probability (*see* Probability distributions and specific distributions, *e.g.*, Binomial distribution)
 properties of
 captured by Pearson's *chi*-square tests, 617
 see also specific properties, e.g., Location
 sample versus population, 275–277
 sampling, 281–282
 of mean, 282–291
 shape of (skew and kurtosis), 62–71
 Drawing random samples (*see* Random sampling)
- Effects, treatment (in analysis of variance), 502
 fixed, 544
 main, 543, 553
 mixed, 544, 553
 random, 551–553
- Efficiency, 299–300
- Element (of a set), 153
- Empty set, 154
- Equation, regression (linear or least squares), 127–131, 596, 688–691
 defined, 128
- Error
 in analysis of variance, 500, 502–503
 of prediction (departure from linearity), 127–128
 sums and means of squares in ANOVA, 511, 512, 709–715
 see also Type I error; Type II error
- Error variance, 492–493, 503–504
 pooled estimate of, in two-way ANOVA, 552–553
- Estimate, variance of, 128
- Estimation, 296–318
 interval, 296, 305–318; *see also* Confidence interval(s)
 point, 296–305
- Estimator(s), point, 296
 desirable properties of, 297–300
 of mean, 300–301
 of proportion, 301
 of standard deviation, 305
 of variance, unbiased, 303–304
- Event(s)
 certain, 154, 183
 defined, 153
 impossible, 154, 183
 independent, 158–160
 intersection of, 156
 mutually exclusive, 161–162
 union of, 160
- Expectation (expected value)
 of binomial random variable, 210
 of *chi*-square random variable, 456
 of continuous random variable, 242
 of difference, 419
 of difference between means, 421
 of discrete random variable, 189–193
 defined, 191
 of Fisher's *F*, 470
 of mean, 287–288
 of mean squares
 among groups, 511–512
 columns, 551–552
 deviations from linearity, 602
 interaction, 548
 linear regression, 602
 residual (in repeated measures ANOVA), 555
 rows, 547, 551–552
 within groups, 510–511
 of standardized random variable, 198–199
 of Student's *t*, 405
 of sums of squares, 511, 512, 709–715
 of variance (S^2), 302–303 (Box 8.2)
- Expectations, Algebra of, 192–193, 696–699
- Expected frequencies for Pearson's *chi*-square tests
 association
 homogeneity of proportions, 638–640, 640–642
 independence, 645–647
 goodness of fit, 619–620, 627–628
 Expected value (*see* Expectation)
- Experiment(s)
 binomial, 204–205
 random, 152
 sampling as, 278–279
- Experimental design, 538, 704–707
 correlated groups, 437–439
 repeated measures, 554

844 Index

- Experimental hypothesis 333*n*; *see also*
 Alternative hypothesis
- Exploratory data analysis, 147, 331, 353
- F* test (*F* ratio)
 hypotheses about two variances, 471–474
 assumptions underlying, 471–472, 480
 one-way ANOVA, 493–494, 513
 assumptions underlying, 498–500, 504, 535
 regression, 602–603, 615
 repeated measures ANOVA, 556, 557 (Box 13.2)
 Scheffé's Multiple Comparison Test, 524, 526
 two-way ANOVA, 544, 547–548, 549, 552–553, 568
- Factorial notation, 168
- Failure in binomial experiment, 204
- Fisher's *F* distribution (*and* random variable), 470–471
 approximation of values, procedure for, 729
 expected value, 470
 parameters, 470
 table of cumulative probabilities, use of, 470–471
- Fisher's *r*-to-*z*, 584–587, 588–589
- Fixed effects in ANOVA, 544
- Fourfold point correlation, 664–666, 680
- Fourfold table, 647–648, 679
- Freedom, degrees of (*see* Degrees of freedom)
- Frequencies
 expected and observed, 619–620, 622–623
 joint, 99
 type of data, 3–4
- Frequency
 cumulative (*see* Cumulative frequency)
 relative (*see* Relative frequency)
- Frequency distribution(s) (*see* Conditional distributions; Distribution[s], frequency; Joint distributions; Marginal distribution[s])
- Frequency polygon, 10–11
- Frequency table, 4–10
- Function(s)
 counting as, 209
 defined, 155
 linear, 126, 591
- probability, 220
 joint, 572
see also Probability, mathematical
 definition of
- probability density, 228–229, 240, 269
 bivariate, 576–580
 normal, 261–262
 standard normal, 240
 Student's *t*, 403, 404–405
- random variables as, 155
- standardization as, 198
- Genovese, Kitty, 128–129
- Goodness-of-fit test (Pearson's *chi*-square), 619–628, 679
 degrees of freedom for, 624
 expected frequencies for, 619–620, 627–628
 for normality, 657–661, 680
 observed frequencies for, 620
 test hypothesis for, 619, 620, 679
 test statistic for, 620–623, 625, 679
- Gosset, William S. ("Student"), 402
- Grand mean, 489, 502, 503, 504, 545
- Grand median, 661, 663
- Grand sum, 485
- Graph
 bar, 11
 line, 11, 576
 point, 10
 representation of bivariate distributions, 575–579
 representation of continuous distributions, 268–269, 577–579
 representation of discrete distributions, 218–220, 575–577
see also Histogram
- Graphic representation of data, 10–13
- Grouped frequency distributions, 6–9, 12
- Groups, among and within, 491–494, 501–503
- Haldane's statistic (*chi*-square test of association with small expected frequencies), 668–669, 680
- Histogram, 11–13, 575–577
 binomial distribution, effects of sample size on, 234–235
 as representation of discrete probability distribution, 218–220
 standardized binomial distribution, effects of sample size on, 235–240

- Homogeneity of proportions, Pearson *chi*-square test of, 636–643, 679
 degrees of freedom for, 642–643
 expected frequencies for, 638–640, 640–642
 test hypothesis for, 637, 679
 test statistic for, 640, 643, 679
- Homogeneity of variance, 499–500, 504, 607
- Homoscedasticity, 499; *see also* Homogeneity of variance
- Hypothesis
 alternative (*see* Alternative hypothesis)
 experimental, 333*n*; *see also* Alternative hypothesis
 null, 332, 382*n*, 418; *see also* Test hypothesis
 test (*see* Test hypothesis)
- Hypothesis tests
 association, 635–648, 679
 homogeneity of proportions, 636–643, 679
 independence, 644–647, 679
 contrasts, 524–527
 correlated proportions, 667–668, 680
 fourfold point correlation, 664–666, 680
 goodness-of-fit, 619–628, 679
 lower-tailed, 369; *see also* Rejection region
 many means (*see* Analysis of variance)
 many proportions (*see* Homogeneity of proportions)
 medians, 661–663
 one correlation (Pearson product-moment), 583–588
 alternative hypothesis for, 584, 585, 587, 614
 assumptions underlying, 583–584, 614
 test hypothesis for, 584, 586, 587, 614
 test statistics for, 585, 586–587, 587–588, 614
 one mean, 382–413
 alternative hypothesis for, 384, 391, 395
 computation of sample size for, 395–398
 decision tree, 388, 389 (Box 10.1)
 large sample, 389–398
 rejection region, 386, 392–393, 411–412
 small sample, 401–413
 test hypothesis for, 382, 384, 391
 test statistics for, 384–386, 389 (Box 10.1), 390, 392, 411, 451
 one proportion
 alternative hypothesis for, 332–333
 test hypothesis for, 332
 test statistics for, 334–336, 378
 one variance, 461–467
 alternative hypothesis for, 462, 480
 assumptions underlying, 463, 480
 rejection region, 465–466
 test hypothesis for, 461, 480
 test statistic for, 462–465, 480
- regression
 test hypotheses for, 597, 602, 615
 test statistics for, 602–603, 615
- two correlations, 588–590
 alternative hypothesis for, 589, 614
 assumptions underlying, 588–589, 614
 test hypothesis for, 589, 614
 test statistic for, 589, 614
- two means, 417–439
 alternative hypothesis for, 418, 438 (correlated groups)
 computation of sample size for, 436–437
 decision tree, 419, 420 (Box 10.4)
 large sample, 423
 rejection region, 423, 435
 small sample, 428–436
 test hypothesis for, 417, 418, 438 (correlated groups)
 test statistics for, 420 (Box 10.4), 422, 423, 431, 433, 434, 439 (correlated groups), 451
- two proportions, 439–440, 451
- two-tailed, 369–370; *see also* Rejection region
- two variances, 471–474
 alternative hypothesis for, 471, 480
 assumptions underlying, 480
 rejection region, 473–474
 test hypothesis for, 471, 480
 test statistic for, 471–473, 480
see also *F* test, hypotheses about two variances
- upper-tailed, 369; *see also* Rejection region
- Independence, Pearson *chi*-square test of, 644–647, 679
 degrees of freedom for, 645
 expected frequencies for, 645–647

846 Index

- test hypothesis for, 645, 679
- test statistic for, 645, 679
- Independence of mean and standard deviation, 499
- Independent events, 158–160
 - defined, 158
 - multiplication theorem for, 159
- Independent random variables, 197, 575, 578
 - covariation of, 197, 699
- Independent variables (in design of experiments), 704
 - regression, 598
- Indicator random variables, 209, 285 (Box 8.1), 301
- Inferential statistics, 275
- Integral, 240–242
- Interaction, 542–543
 - mean squares as denominator in F test for random- and mixed-effects ANOVA, 551–552
 - sum of squares pooled with sum of squares error, 552–553
- Interpolation, linear, 30, 80, 686–687
- Interquartile range, 142 n
- Intersection of events, 156
- Interval estimates (*see* Confidence intervals)
- Interval scales, 617, 618 (Box 15.1), 694–695
- Interval width (class width), 7
- Intervals, confidence (*see* Confidence interval[s])
- Joint distributions
 - frequency, 99–101, 117–118
 - probability, 571–580
- Kirk, J. T. (Captain, USS Enterprise, NCC-1701), 62
- Kurtosis, 68–71
 - captured by Pearson's *chi*-square tests, 617
 - coefficient of (Kur), 68
- “Law of Large Numbers,” 182
- Least squares, 128
- Levels of measurement, 617, 618 (Box 15.1), 692
- Limits
 - integration, 241
 - summation, 21–22
- Line, (least squares) regression, 128 n , 131–134, 596
- Line graphs, 11, 576
- Linear covariation, 117, 120–121, 590–592
- Linear interpolation, 30, 80, 686–687
- Linear model for analysis of variance, 500–504, 543
- Linear regression, 125–134, 596–607
 - deviations from, 599–600, 602; *see also* Error of prediction; Hypothesis tests, regression
 - least squares equation, 127–131, 596
- Location, 26–27
 - captured by Pearson's *chi*-square tests, 617
 - probability distributions, 189–192
 - statistics, 27–37, 39–44
 - see also specific location statistics, e.g., Mean*
- Logarithm, natural (*see* Natural logarithm)
- Lower-tailed tests, 369; *see also* Rejection region
- Main effects in ANOVA, 543, 553
- Marginal distribution(s)
 - frequency, 99, 117–119, 132–134
 - mean(s), 132
 - range, 101–103, 119
 - variance(s), 119–121, 134
 - probability, 572, 573
- Matching, 437, 706
- Mathematical definition of probability, 184
- Mean
 - center of mass of distribution, 42–44
 - conditional, 118–119, 131, 578
 - data collection (sample mean), 39–44
 - as estimator of population mean, 300–301
 - expected value of, 287–288
 - grand, 489, 502, 503, 504, 545
 - grouped data, 41–42
 - marginal, 132
 - point about which sum of deviations is zero, 44
 - population, estimation of
 - interval, 306–314, 413 (Box 10.3)
 - point, 300–301
 - sampling distribution of, 282–291
 - standard error of, 289
 - value about which sum of squared deviations is smallest, 599, 600 (Box 14.1)

- variance of, 288–289
- see also* Expectation; Hypothesis tests, one mean; Hypothesis tests, two means
- Mean (absolute) deviation from the mean, 51–52
- Mean squares
 - across groups (*see* Mean squares, among groups)
 - across treatments (in repeated measures ANOVA), 556, 559
 - computational formula for, 557 (Box 13.2)
 - among groups
 - computational formula for, 510 (Box 12.2)
 - expected value of, 511–512
 - and mean squares across treatments in repeated measures ANOVA, 559
 - between groups (*see* Mean squares, among groups)
 - columns, 546, 569
 - computational formula for, 548 (Box 13.1), 569
 - expected value of, 551–552
 - deviations from linearity, 602, 615
 - error (in ANOVA for regression), 602, 615
 - computational formula for, 603, 615
 - error (in two-way ANOVA), 545
 - computational formula for, 548 (Box 13.1)
 - interaction, 546, 569
 - computational formula for, 548 (Box 13.1), 569
 - expected value of, 548
 - linear regression, 602, 615
 - computational formula for, 604, 615
 - pooled error, 552–553, 568, 569
 - computational formula for, 569
 - residual (in repeated measures ANOVA), 555
 - computational formula for, 557 (Box 13.2)
 - and mean squares within groups in one-way ANOVA, 559
 - rows, 546, 569
 - computational formula for, 548 (Box 13.1), 569
 - expected value of, 547, 551–552
 - within groups
 - computational formula for, 510 (Box 12.2)
 - estimate of error variance, 512
 - expected value of, 510–511
 - and mean squares residual in repeated measures ANOVA, 559
- Measurement(s)
 - levels of, 617, 618 (Box 15.1), 692
 - precision of, 7–8
 - requirements for comparisons among, 86, 695
 - type of data (scores), 3–4
 - unobserved (population), as values of random variable, 277
- Measures, repeated (ANOVA), 554–559
- Median
 - as 50th percentile, 82–83
 - grand, 661, 663
 - grouped data, 29–36
 - interval, 31
 - relation to mean with skewed data, 37, 63–64
 - test, 661–663, 680
 - ungrouped data, 28–29
- Minimax loss criterion, 351–352
- Mixed effects in ANOVA, 544, 553
- Modality, 71; *see also* Kurtosis
- Mode, 27–28
- Models
 - probability (distributions) as, 151, 187, 225
 - statistical, in hypothesis testing, 331–332
- Multiplication theorem, 158
 - independent events, 159
- Mutual exclusivity, 153–154, 161–162
- Natural logarithm, 240, 584*n*
 - in Fisher's *r*-to-*z* transformation, 584
- Nominal scales, 617, 618 (Box 15.1), 692–693
- Nondirectional hypothesis, 333–334
- Nonlinear (curvilinear) covariation, 117, 120–121, 593
- Nonparametric statistics, 402, 619
- Normal approximation to binomial, 251–256
- Normal curve (*see* Normal random variable)
- Normal distribution, 232–264
 - parameters of, 260–261
 - see also* Normal random variable; Standard normal random variable
- Normal random variable, 260–264
 - probability density function for, 261–262
 - see also* Standard normal random variable
- Normality, goodness-of-fit test for, 657–661, 680

848 Index

- Null hypothesis, 332, 382*n*, 418; *see also* Test hypothesis
- Observed frequencies (in Pearson's *chi*-square), 620
- One-way analysis of variance, 539; *see also specific topics in Table of Contents, Chapter 12*
- Ordinal scales, 617, 618 (Box 15.1), 693–694
- Outlier(s), 142–143, 749 (Ans. Exercise 1.1.11)
- p*-value (*see* Significance probability)
- Parameter(s), 220
 binomial distribution, 220
 chi-square distribution, 456
 Fisher's *F* distribution, 470
 models of population values, 277
 normal distribution, 260–261
 notational convention, 220, 277
 specification of values, 220
 statistics and, 277
 Student's *t* distribution, 404–405
- Partitioning sum of squares, 505–508, 545–546, 554–555, 559
- Paull's Rule, 552–553
- Pearson product-moment correlation coefficient (*see* Correlation, Pearson product-moment)
- Pearson's *chi*-square tests
 association, 635–648, 679
 homogeneity of proportions, 636–643, 679
 independence, 644–647, 679
 with small samples (Haldane's statistic), 668–669, 680
 assumptions underlying, 656–657
 correlated proportions, 667–668, 680
 correlation, fourfold point, 664–666, 680
 goodness of fit, 619–628, 679
 for normality, 657–661, 680
 median test, 661–663
 properties of distributions captured by, 617
 sample sizes required for, 656, 663
- Pearson's skewness coefficient (Sk_P), 64–65
- Percentile(s), 81–83
 ungrouped data, 83, 760 (Ans. Exercise 3.1.3)
- Percentile rank, 79–81
- Permutations, 170–173
 defined, 170
 formula for number of, 172
- Point estimation, 296–305
- Point estimator(s)
 defined, 296
 desirable properties of, 297–300
 of population mean, 300–301
 of population proportion, 301
 of population variance, 301–304
- Point graphs, 10
- Poisson distribution, 745
 binomial approximation to, 745
 table of probabilities, use of, 746–747
- Pooled estimate of error variance (in two-way ANOVA), 552–553
- Pooled estimate of variance (in *t* test), 429–431
- Population(s), 276
 sample(s) and, 275–278
- Population coefficient of determination, 591, 597
- Population correlation coefficient, 583–593
 defined, 583
 testing hypotheses about, 583–590
- Population correlation ratio, 590–592, 597
- Population mean (*see* Confidence interval[s], mean; Expectation; Hypothesis tests, one mean; Hypothesis tests, two means; Point estimator[s], of population mean)
- Population proportion (*see* Proportion[s], population)
- Population standard deviation, point estimator of, 305
- Population value(s), 277
 notational convention, 277–278
 see also Parameter(s)
- Population variance (*see* Confidence interval[s], variance; Hypothesis tests, one variance; Hypothesis tests, two variances; Point estimator[s], of population variance)
- Post hoc tests, 524–527
- Power, 363–368, 387, 412–413, 436–437
 alternative hypothesis and, 364–366, 368, 393, 395
 defined, 364
 one-tailed and two-tailed test, 393, 395, 396
 sample size and, 364, 365, 395–398, 412, 436–437
 scientific significance and, 366–368, 395, 413, 436
 significance level and, 364, 365
 see also Type II error, probability of

- Precision of measurement, 7–8
 Prediction, 126–134
 Principle of Multiplication in Counting, 165–166
 Probability
 classical definition of, 153
 conditional, 157–158
 cumulative, 231
 distributions (*see* Probability distributions)
 mathematical definition of, 184
 model of relative frequency, 151, 187, 225
 relative frequency approach to, 180–182
 significance (*see* Significance probability)
 two events
 intersection, 156
 union, 160
 Type I error (α) (*see* Type I error, probability of)
 Type II error (β) (*see* Type II error, probability of)
 Probability density, 228–229, 240
 Probability distributions, 187–188
 bivariate, 572
 conditional, 573–574
 describing properties of, 189–198, 242
 joint, 571–580
 marginal, 572
 models of relative frequency
 distributions, 187, 225, 276–277
 representations of
 continuous, 268–269, 577–578
 discrete, 217–220, 575–577
 specification of, 220
 see also specific distributions, e.g.,
 Binomial distribution
 Probability function, 220
 Proportion(s)
 binomial, 212–214, 256–257
 population
 confidence intervals for, 314–318
 relative frequency as estimator of, 301, 315, 316
 sample (*see* Proportion, population, relative frequency as estimator of)
 see also
 Confidence interval(s), proportion;
 Hypothesis tests, association,
 homogeneity of proportions;
 Hypothesis tests, one proportion;
 Hypothesis tests, two proportions;
 r-to-*z*, 584–587, 588–589
 Random effects in ANOVA, 544, 551–553
 Random experiment(s), 152, 278–279
 Random sampling
 defined, 279
 fundamental assumption in data
 collection, 352–353
 as random experiment, 278–279
 Random variable(s)
 continuous, 188, 225, 228
 covariance of, 197
 defined, 152, 155
 discrete, 187–188
 expected value of, 191
 variance of, 194
 independence of, 197, 575, 578
 indicator, 209, 285 (Box 8.1), 301
 notational conventions for, xvii
 standard deviation of, 198
 standardized, 198–200
 statistics as, 279–280; *see also*
 Distribution(s), sampling;
 see also specific distributions, e.g.,
 Normal distribution, and *specific*
 random variables, e.g., Binomial
 random variable
 Random variables as functions, 155 (Box 5.1)
 Randomization (in experimental design), 705–706
 Range, 49–50
 conditional, 119
 interquartile, 142*n*
 marginal, 101–103, 119
 Rank, percentile, 79–81
 Ratio, correlation, 119–121, 590–592, 597
 Ratio scales, 617, 618 (Box 15.1), 695
 Regression, 125–134, 596–607, 688–691
 least-squares (linear) prediction,
 127–131, 606
 see also Hypothesis tests, regression
 Regression equation, linear (*or* least squares), 127–131, 596, 688–691
 Regression line (least squares), 128*n*, 131–134, 606
 Rejection of hypotheses versus acceptance of hypotheses, 356–357
 Rejection region, 329
 in analysis of variance, 495, 513
 determination by
 alternative hypothesis, 369–370
 significance level, 336–339
 in hypothesis tests about contrasts, 526

850 Index

- in hypothesis tests about correlation, 585, 587
- in hypothesis tests about means, 386, 392–393, 410–411, 423, 435
 - one-tailed and two-tailed, 393, 394, 395–396, 411
- in hypothesis tests about proportions, 336–339
 - one-tailed and two-tailed, 369–370
- in hypothesis tests about variances, 465–466, 473–474
- in Pearson's *chi*-square tests, 625
- Rejection rule, 336
 - minimax loss, 351–352
 - see also* Rejection region
- Relatedness, statistical (*see* Statistical relationship)
- Relative frequency
 - approach to (definition of) probability, 180–182
 - cumulative, 6
 - defined, 5
 - distribution(s), 5, 12
 - modeled by probability distributions, 187
 - as estimate of probability, 180–182
 - as estimator of proportion (p), 301, 315, 316
 - probability as model of, 151
- Repeated measures (ANOVA), 554–559
- Residual, mean squares (in repeated measures ANOVA), 555
 - computational formula for, 557 (Box 13.2)
 - and mean squares within groups in one-way ANOVA, 559
- Results
 - counting, 165–177
 - mutual exclusivity of, 153–154
 - number of, in classical probability, 153
- Sample mean (*see* Mean, data collection)
- Sample proportion (*see* Relative frequency as estimator of proportion)
- Sample size
 - calculating for
 - confidence interval of specified width, 311 (Example 8.10), 316–317
 - test of one mean, 395–398
 - test of two means, 436–437
 - effects on
 - histogram of binomial random variable, 234–235
 - histogram of standardized binomial random variable, 235–240
 - power, 364–365, 395–396
 - probability of Type II error, 353–356
 - statistical and scientific significance, 367–368, 395, 436
- equality of, and assumption of
 - homogeneity of variance, 432–433, 499–500
- parameter of binomial probability function, 220
- requirements for
 - Fisher's r -to- z , 583–584
 - normal approximation in testing
 - hypotheses about means, 389–390, 423, 452
 - normal approximation to binomial in testing hypotheses about
 - proportions, 335, 378, 440, 451, 783 (Ans. Exercise 9.3.7)
 - Pearson's *chi*-square tests, 656, 663
 - Poisson approximation to binomial, 745
- Sample space, 152
- Sample variance (*see* Unbiased estimator of population variance)
- Samples
 - drawing random (*see* Random sampling)
 - independent versus correlated groups, 437–439
 - and populations, 275–278
- Sampling
 - distribution(s) (*see* Distribution[s], sampling)
 - see also* Random sampling
- Scale(s) of measurement, 86, 617, 618 (Box 15.1), 692–695
- Scatter plot (scatter diagram), 99–104
- Scheffé's Multiple Comparison Test, 526–527
- Scientific significance
 - composite alternative hypotheses and, 368
 - effects of sample size on, 367, 395, 436
 - statistical significance and, 366–368, 395
- Score(s), 3, 79
 - deviation, 44, 90, 107–109
 - standard, 92–94
- Set, 153
 - empty, 154
- Shape of distribution (skew and kurtosis), 62–71
- Significance
 - level (α), 336
 - confidence and, 340–341
 - effects on, of multiple tests, 490–491

- effects on, of normal approximation to binomial, 370–372
- as probability of Type I error, 347
- probability of Type II error (β) and, 347–350
- and rejection region, determination of, 336–339
- scientific and statistical, 366–368, 395
- tests of (*see* Hypothesis tests)
- as unexpectedness, 339–340
- Significance probability (p -value), 357–360
 - defined, 357
 - reporting conventions, 358 (Box 9.1), 394
 - significance level (α) and, 357, 359–360, 394–395
- Simple hypothesis, 332–334
- Skew (*and* skewness), 63–68
 - captured by Pearson’s *chi*-square tests, 617
 - coefficient of (Sk), 65–66
 - Pearson’s coefficient of (Sk_P), 64–65
- Specification statement, 220
- Standard deviation
 - data collection (s^2), 55–57
 - computational formulae for, 56–57
 - unit of standardized collection of scores, 91, 93–94
 - distribution of means (*see* Standard error of the mean)
 - population, point estimator of, 305
 - random variable, 198
- Standard error of the mean, 289
- Standard normal random variable, 240–250
 - expectation and variance of, 242
 - limit of binomial as sample size increases, 240
 - demonstration of, 235–239
 - probability density function for, 240
 - properties of, 243–244
 - table of cumulative probabilities, use of, 244–250
- Standard scores, 92–94
- Standardization, 88–92, 198–200
- Standardized binomial proportion, 256–257
- Standardized binomial random variable, 214, 235–239
- Statistical relationship, 101–104, 590–591
 - linear, 119, 590
 - negative, 103
 - positive, 103
- Statistics
 - descriptive, 19–24
 - defined, 19
 - desirable properties of, 20–21
 - see also specific descriptive statistics, e.g., Mean, and descriptive properties, e.g., Location*
 - estimators, point, 296
 - desirable properties of, 297–300
 - interential, defined, 275
 - test (*see* Test statistic[s])
 - “Student” (William Sealy Gosset), 402
 - Student’s *t* (and Student’s *t* distribution), 402–408
 - assumptions underlying tests of means, 389 (Box 10.1), 403, 420 (Box 10.4), 437
 - expectation and variance of, 405
 - and normal distribution, 405–406
 - parameter of, 404–405
 - probability density function for, 403, 404–405
 - table of cumulative probabilities, use of, 407–408
 - test of one correlation, 587–588
- Subjects as own controls, 437, 438, 706–707
- Subsets, 153
- Success in binomial experiment, 204
- Sum(s) of squares, 505
 - across groups (*see* Sums of squares, among groups)
 - across subjects, 555
 - across treatments (in repeated measures ANOVA), 555
 - computational formula for, 557 (Box 13.2)
 - among groups, 506–507 (Box 12.1), 507–509
 - computational formula for, 508–509
 - expected value of, 512
 - between groups (*see* Sums of squares, among groups)
- columns, 546, 569
 - computational formula for, 548 (Box 13.1), 569
- deviations from linearity, 599–600, 615
- error (in ANOVA for regression), 599, 615
 - computational formula for, 603, 615
- error (in two-way ANOVA), 545, 569
 - computational formula for, 548 (Box 13.1), 569

852 Index

- interaction, 546, 569
 - computational formula for, 548 (Box 13.1), 569
- linear regression, 600–601, 615
 - computational formula for, 603–604, 615
- pooled error, 553, 569
 - computational formula for, 569
- residual (in repeated measures ANOVA), 555
 - computational formula for, 557 (Box 13.2)
- rows, 546, 569
 - computational formula for, 548 (Box 13.1)
- total
 - in ANOVA for regression, 599, 601
 - in one-way ANOVA, 505, 506–507 (Box 12.1), 507–509
 - in repeated measures ANOVA, 554–555
 - in two-way ANOVA, 545
- within groups, 506–507 (Box 12.1), 507–509, 511, 599, 601
 - computational formula for, 508–509
 - expected value of, 511
- within subjects, 555
 - computational formula for, 557 (Box 13.2)
- see also* Appendix VII, 708–715
- Summation(s)
 - Algebra of, 23–24, 682–684
 - double, 485–487
 - notation, 21–23
- Synthetic alternative hypotheses, 368, 395
- t* test (and *t* distribution) (*see* Student's *t*)
- Tables of (cumulative) probabilities, use of (*see specific distributions, e.g., Binomial distribution*)
- Test(s) (*see* Hypothesis tests; *specific entries, e.g., Goodness-of-fit-test*)
- Test hypothesis, 328–329, 332
 - for analysis of variance
 - one-way, 382, 490, 535
 - post hoc comparisons, 525, 535
 - for regression, 597, 602, 615
 - repeated measures, 556
 - two-way, 544, 568
- for test of
 - goodness of fit, 619, 620, 679
 - homogeneity of proportions, 637, 679
 - independence, 645, 679
 - medians, 661
 - one correlation (Pearson product-moment), 584, 586, 587, 614
 - one mean, 382, 384, 391
 - one proportion, 332
 - one variance, 461, 480
 - two correlations (Pearson product-moment), 586, 614
 - two means, 417, 418, 438 (correlated groups)
 - two proportions, 440
 - two variances, 471, 480
- Test statistic(s), 327, 329, 334–336
 - analysis of variance
 - one-way, 490–495, 513, 535
 - post hoc comparisons, 526, 535
 - regression, 602–603, 615
 - repeated measures, 556, 557 (Box 13.2)
 - two-way fixed effects, 547–549, 568
 - two-way mixed effects, 553
 - two-way random effects, 552–553, 568
 - correlated proportions (Pearson's *chi-square*), 667
 - defined, 329
 - desirable properties of, 334
 - goodness of fit, 620–623, 625, 679
 - Haldane's (*chi-square* test of association with small expected frequencies), 668, 669
 - homogeneity of proportions (Pearson's *chi-square*), 640, 643, 679
 - independence (Pearson's *chi-square*), 645, 679
 - one correlation (Pearson product-moment), 585, 586–587, 587–588, 614
 - one mean, 385, 389 (Box 10.1), 390, 410, 451
 - one proportion, 329, 334–335, 368–369
 - one variance, 462–465, 480
 - two correlations (Pearson product-moment), 588–589, 614
 - two means, 420 (Box 10.4), 422, 423, 431, 433, 434, 452
 - correlated groups, 439
 - two proportions, 440, 451
 - two variances, 471–474, 480
- Theory in hypothesis testing, 330–331
- Treatment(s), 500, 503, 555
 - effects, 502
 - groups, 705–707
 - mean squares across (in repeated measures ANOVA), 556, 559
- Treatment variance, 492–493, 503–504
- T*-scores, 96 (Exercise 3.2.16)

- Two-tailed tests, 369–370
 see also Rejection region
- Two-way analysis of variance, 539–554
 fixed-effects, 544
 interaction in, 542–543, 551–553
 main effects in, 543, 552, 553
 mixed-effects, 544, 553
 random-effects, 544, 551–553
- Type I error, 346–347
 costs associated with, 349–352
 probability of, 347
 with multiple t tests, 490–491
 see also Significance, level (α);
 Significance probability (p -value)
- Type II error, 347
 costs associated with, 351–352
 probability of, 347–350
 factors affecting, 347, 349–350,
 353–356
 power and, 363–364
- Unbiased estimator of population variance
 ($\hat{\sigma}^2$), 303–304
- Unbiasedness, 297
- Uncertainty, 101–102
 and reduction of marginal variability,
 119–121, 590–592
 surrounding statistical inferences, 327
- Ungrouped frequency distributions
 K th percentile, 83
 median, 28–29
- Uniform distribution, 188, 189–200 (*see esp.* 189), 225–231 (*see esp.* 225, 228)
- Union of events, 160
- Upper-tailed tests, 369; *see also* Rejection region
- Value set, 152
- Variability (dispersion), 48–49
- Variable(s) (*see* Independent variable[s];
 Random variable[s])
- Variance
 analysis of (*see* Analysis of variance)
 collection of data (s^2), 52–58
 biased estimator of population variance,
 302–303
 computational formulae for, 56–58
 conditional, 119–121, 132–134, 578, 592
 error, 492–493, 503–504; *see also*
 Expectation of mean squares
 grouped data, 54–55
 homogeneity of, 499–500, 504, 607
 marginal, 119–121, 134
 mean, sampling distribution of, 288–289
 pooled estimate of, 429–431
 population (σ^2)
 confidence intervals, 467 (Box 11.2)
 unbiased estimator of ($\hat{\sigma}^2$), 303–304
 random variable (expected variance),
 194–198
 computational formulae, 194–196
 continuous, 242
 discrete, 194
 standardized, 198–200
 see also specific random variables, e.g.,
 Binomial random variable
 sample (*see* Variance, population,
 unbiased estimator of)
 treatment, 492–493, 503–504
 see also Hypothesis tests, one variance;
 Hypothesis tests, two variances
- Variance of estimate, 128
- Variances, Algebra of, 196–198, 699–702
- Variation, coefficient of, 58–59
- Within-groups, 493
 mean squares, 510 (Box 12.2), 510–511,
 512, 559
 sum of squares, 506–507 (Box 12.1),
 507–509, 511, 599, 601
- Yates' correction for continuity, 648*n*