

Index

- α , type I error, 93–94, 100, 106
- α_i , treatment effect, 109–110, 123, 135, 209
- α parametric Y -intercept, *see* Y -intercept
- $(\alpha\beta)_{ij}$, interaction effect, 135, 241
- a sample Y -intercept, *see* Y -intercept
- acceptance region, 92
- accuracy of data, 6
- alternative hypothesis, 90–91
- analysis of variance
 - assumptions of, 131–132
 - introduction and nomenclature, 123
 - model I, 122, 124
 - model II, 122–123
 - nomenclature, 123
 - partitioning of total sum of squares, 125–127
 - table, 128
 - two-way
 - mixed model, 143
 - model I, 135–143
 - nonparametric methods in lieu of, 153
 - with unequal sample sizes, 127
- ANOVA. *See* Analysis of variance
- a posteriori comparisons, 133–135
- approximate t -test of equality of means assuming unequal variances, 115
- area under curve, 61–64
- assumptions of statistical tests, 101
- attributes, 4
- average, 25–26
- b , statistical regression coefficient (slope), 211, 212, 215
- β , parametric value for regression coefficient (slope), 213
- β , type II error, 93–96, 106
- β_i , treatment effect, 135
- bar graph, 19
- Bayes’ theorem, 53–57
- bimodal distribution, 29, 79
- binomial distribution, 46–47
- bivariate
 - normal distribution, 195
 - scattergram, 195
- box plots, 21
- χ^2 , chi-square, 166, 169
- $\chi^2_{\alpha[v]}$, critical chi-square at probability level α , and degrees of freedom, 168
- causation, 193, 210
- central limit theorem, 75
- central tendency measure of, 26
- character variables, 4
- chi-square distribution, 166, 169
- chi-square test, 166
 - for goodness of fit, 166–169
 - degrees of freedom, 167
- clumping,
 - departure from Poisson distribution, 49
- coefficient
 - Kendall’s, 202–203
 - Pearson’s, 194, 196
 - Spearman’s, 199–200
 - regression, *see* regression.
- collinearity, *see* multicollinearity
- comparison of
 - a single observation with the mean of a sample, 116
 - means
 - planned, 133
 - unplanned, 132
- comparisons
 - Paired, 117, 159
- confidence
 - interval for predicted value in regression analysis, 218
 - limits of an estimated mean, 86
- continuous variables, 5
 - frequency distributions of, 57–63
- correction
 - Edward’s, 184
 - Williams’, 178
 - Yates’, for continuity, 175

- correlation
 - applications of, 199
 - matrix analysis, 235, 240, 244
 - and regression, contrasted, 209–210
 - significance tests in, 197–198
 - spurious, 193
- critical
 - region, 92
 - value, 104
- cumulative
 - normal curve, 57–63
- curve
 - area under, 64
 - curvilinear regression, 211, 221
- df*, degrees of freedom, 33
- data
 - accuracy of, 6
 - precision of, 6
- datum vs. data, 4
- degrees of freedom, 33
- dependent variable, 6
- descriptive statistics, 8
- Design
 - paired comparisons, 117, 159
 - determination, coefficient of, 212, 216
- deviate
 - standard normal, 63
- difference
 - between two means, 108
 - from the mean, sum of, 32
 - significance of, 110
 - between two variances, testing significance of, 115
- discrete variables, 3, 45
- distribution
 - bimodal, 29, 79
 - binomial, 46
 - bivariate normal, 195
 - chi-square, 166, 169
 - clumped, 49
 - contagious, 49
 - continuous numeric variables, of, 57
 - discontinuous numeric variables, of, 44
 - frequency *See* frequency distribution
 - kurtosis, 77, 114
 - leptokurtic, 77
 - normal, 59–60
 - platykurtic, 77
 - Poisson, 48
 - probability, 43
 - qualitative variable, of, 44
 - sample means, of, 73
 - skewness, 77, 114
 - Standard normal, 59, 63
 - Student's *t*-, 86–87, 103–104
- distribution-free methods *See* nonparametric methods
- ϵ_{ij} , random deviation of the *j*th individual of group, 21, 109, 123
- effects
 - random group, 123
 - treatment, 123
- equality of
 - two means assuming unequal variances, 115
 - variances, assumption, 115, 132
- equation
 - multiple regression, algebraic computation, 236–240
 - multiple regression, sample and parametric, 235
 - regression (sample), 211–212
 - regression (parametric), 213
- error
 - rate, experimentwise, 132–133
 - standard, *see* Standard error
 - type I, 94, 106
 - type II, 94–95, 106
- estimation, point, 89
- exact test, Fisher's, 178–183
- exhaustiveness, 4
- expected
 - binomial frequencies, 46–48
 - normal frequencies, 60–63
 - Poisson frequencies, 46–52
 - value
 - for *Y*, given *X*, 217, 218
- experimentwise error rate, 132–133
- f_e , expected frequency, 167
- f_o , observed frequency, 167
- F*, variance ratio, 125–126
- $F_{\max}$ -test, 115, 132
- F*-distribution, 126
- Fisher's exact test, 178–183
- Fmax test of homogeneity of variances, 115, 132
- frequencies
 - expected, 166–167
 - observed, 166–167
- frequency distribution
 - of continuous variables, 15–17
 - normal, 23
 - qualitative, 13–15
 - quantitative discontinuous, 13–15
- g_1 , measure of skewness, 114
- g_2 , measure of kurtosis, 114
- y_1 , parametric measure of skewness, 114
- y_2 , parametric measure of kurtosis, 114

- G-test
 - introduction and rationale, 176
 - for goodness of fit, single classification, 177
 - of independence, 177
 - Williams' correction, 178
- goodness of fit tests
 - chi-square, 166–169
 - G-test, 176
 - Kolmogorov–Smirnov test, 170–172
- H*, statistic for Kruskal–Wallis test
 - with no ties, 153
 - with ties, 154
- H*₀, null hypothesis, 90
- H*₁, alternative hypothesis, 90
- heteroscedasticity, 115
- histogram, 21
- homogeneity
 - of variances assumption in analysis of variance, 115, 132
- homoscedasticity, 115, 132
- hypothesis
 - alternative, 90
 - null, 90
 - testing, 90
- independence
 - assumption in analysis of variance, 113, 132
 - exact test of, *see* Fisher's exact test
- independent variable, 6
- inferential statistics, 8
- interaction, 135, 241
 - synergy, 136
 - interference, 136
- Kendall's coefficient of rank correlation, 202–207
- Kolmogorov–Smirnov one-sample test, 170–172
- Kruskal–Wallis test, 153–159
- kurtosis, 114
- L*, likelihood ratio test, 176
- least squares, explained, 221
- leptokurtic curve, 78
- level, significance, 83–84, 86
- likelihood ratio test, *see* G test
- limits
 - confidence, *see* confidence limits
- linear regression. *See* regression
- logistic regression, 251–254
- logit, 252
- log-odds, 252
- MS*, mean square, 125
 - among groups, 125–126
 - error, 125–126
 - regression (explained), 215–217
 - total, 125–126
 - unexplained, 215–217
- Mallow's Cp, *see* multiple regression
- Man–Whitney *U*-test, 148–153
- Matrix, correlation, 240, 243, 244
- maximum likelihood, 84
- McNemar test for a matched design, 183–184
- mean
 - sample, 26
 - parametric, 26
 - of Poisson distribution, 50
 - of a sample, comparison of a single observation with, 116–117
 - square, *see MS*
- means
 - distribution of, 73–77
 - t*-test of the difference between two, 108–112
- measure of central tendency, 26
- median, 27–28
- mixed model
 - two-way ANOVA, 143
- mode, 28–29
- Model I
 - ANOVA, 122
- Model II
 - ANOVA, 123
- multiple regression
 - algebraic computation of equation, 236
 - backward elimination procedure, 242
 - equation, 235
 - forward selection procedure, 242
 - F to enter and F to remove, 242
 - Mallow's Cp, 240–241
 - parsimony, principle of, 240
 - significance tests, 240
 - stepwise procedure, 242
- multicollinearity, 235–236
- mutual exclusiveness, 4
- N*, population size, 8
- n*, sample size, 8
- nonparametric methods
 - in lieu of ANOVA, 146
 - power of, 146
 - see* Kruskal–Wallis
 - see* Mann–Whitney
 - see* Wilcoxon signed-ranks test
- normal
 - curve
 - areas of, 61
 - height of ordinate, 60
 - deviates
 - standard, 63
 - distribution
 - standard, 63

- normality
 - tests of the assumption of, 102, 113, 132
- null hypothesis, 90
- observations, individual, 4
- odds ratio, 185–188
- one-tailed tests
 - compared with two-tailed test, 100–101
- p , probability, 105
- p , binomial probability, 47
- paired comparisons
 - nonparametric tests, 159–163
 - t -test for, 117–119
- parameters, 8
- parametric
 - mean, 8, 33
 - product-moment correlation coefficient, 193
 - regression coefficient, 212
 - variance, 22
- percentiles, 71
- pie charts, 19
- platykurtic curve, 78
- point estimation, 89
- Poisson distribution, 48–53
 - calculation of expected frequencies, 50–51
 - clumping, 49
 - mean, 50
 - repulsion, 49
 - variance, 50
- polygon, 21
- population
 - statistics, *see* parameters
- power of a test, 95–98
- predictor variable, 209
- probability
 - distribution
 - qualitative and discontinuous numeric variables, of, 44–46
 - continuous numeric variables, of, 57–63
 - random sampling, and, 43–44
- product-moment correlation coefficient
 - formula for 196
- q , binomial probability, 47
- qualitative frequency distribution, 13–15
- quantitative frequency distribution, 15–17
- r_s , Spearman's coefficient of correlation, 199–202
- r , product-moment correlation coefficient, *see* product-moment correlation coefficient
- random sampling, 43
- range, 30
 - mean
- rank correlation
 - Kendall's coefficient of, 202–207
- ranking data, 147
- region
 - acceptance, 92
 - critical, 92
 - rejection, 92
- regression
 - analysis of variance, as, 215–217
 - assumptions, 214–215
 - coefficient
 - parametric value for, 217
 - standard error of, 217
 - test of significance of, 217
 - contrasted with correlation, 19
 - curvilinear
 - equation
 - simple, 212
 - multiple, 236–240
 - introduction to, 210
 - lack-of-fit, 213–214, 219–221
 - linear
 - logistic, 251–254
 - more than one value of Y per X , 213–214
 - multicollinearity, 235–236
 - multiple, 234–235
 - F to enter and F to remove, 242
 - Mallow's Cp, 240–241
 - model-building, 240–243
 - prediction of values of Y , 217
 - dummy independent variables, 247
 - residual analysis, 221–222
 - slope, 212
 - transformations in, 225–232
 - writing up results, 255
- rejection region, 92
- relative risk, 188–189
- residuals, examination of, *see* regression
- response variable, 210
- s , sample standard deviation, 33–37
- s^2 , sample variance, 33–37
- s_b , standard error of regression coefficient, 217
- $S_{\bar{y}}$, standard error of the mean, 75, 76, 86
- SS, sum of squares, 31, 36
- σ^2 , parametric variance
- sample
 - mean, 26
 - representative, 9
 - size, 9
 - statistics, 26
 - standard deviation and variance, 33–37
- sampling strategies, 8–9, 44
- Scheffé test, 133–135
- signed-ranks test, Wilcoxon's, 159–163

- significance
 - classical approach to, 83–43
 - levels, 94
 - statistical, 91
- size, sample, 9
- skewness, 77, 114
- slope, *see* regression
- spearman’s coefficient of rank correlation, 199–202
- spurious correlations, 194
- square, mean, 125–126
- square root transformation, *see* transformations in regression
- squares
 - least, 221
 - sums of (in ANOVA), 125–127
- standard
 - Deviate, 63
 - deviation, *see* *s*, standard deviation
 - error
 - of mean differences, 110
- standardization, 63
- standardized deviate, 63
- statistical
 - significance, *see* significance
 - thinking, 42
- statistics
 - descriptive, 8
 - of dispersion, 29–38
 - of location, 26–29
- stem-and-leaf display, 17
- stepwise multiple regression analysis, *see*, regression, multiple, model-building.
- Student’s *t*-distribution, 86–87, 103–104
- sum of squares
 - among groups, 124
 - explained, 212
 - interaction, 241
 - total, 124
 - unexplained, 212
- summation sign
- synergism, 136
- $t_{\alpha[v]}$, critical values of Student’s distribution for *v* degrees of freedom, 87
- T*, critical value of rank sum of Wilcoxon’s signed ranks test, 160
- τ , Kendall’s coefficient of rank correlation, 203
- t*-distribution, Student’s, 86–87, 103–104
- t*-test for
 - difference between a single specimen and a sample, 116
 - difference between two means, 112
 - paired comparisons, 117
- table
 - two-by-two frequency, 172, 173
- testing, hypothesis, *see* hypothesis testing
- tolerance, 236
- total sum of squares, 124
- transformations, *see* regression, transformations
- treatment effect, 7, 90, 97
- two-by-two table, 172, 173
- two-tailed tests, 100
- two-way
 - analysis of variance, *see* analysis of variance
- Type I error, 94
- Type II error, 94–95
- U*-test, Mann–Whitney, 148–153
- unbiased estimators, 33
- variable
 - continuous numeric, 5
 - dependent, 6
 - discontinuous numeric, 5
 - independent, 6
 - ordered, 4
 - qualitative, 4
 - ranked, 4
- variance
 - of a Poisson distribution, 50
 - parametric, 31–33
 - sample, 33–37
- variance inflation factor, 236
- variances
 - tests of homogeneity of, 115, 132
- variate, 4
- Wilcoxon’s signed-ranks test, 159–163
- X axis, 18
- χ^2 , sample statistic of chi-square distribution, *see* chi-square test
- χ^2 -test, *see* chi-square test
- $\hat{Y}$, predicted value of *Y*, 211–212
- $\bar{Y}$, mean of a sample, 26
- Y axis, 18
- Y-intercept, 212
- Yates’ correction for continuity, 175
- z scores, 98