EXAMPLES INDEX

COPYRIGHTED MATERIAL

Abortion opinions, 7, 16, 254–255, 257–260, 281–283
  student survey, 315–318
AIDS and AZT use, 116, 153
Alligator food choice, 160–163, 187
Assisted living enrollment, 322
Astrology and ghosts, 59
Auto accidents and seat-belts, 58
Automobile accidents, 192, 204–210, 222

Baseball complete games, 84
Basketball free throws, 279–281
Belief in afterlife, 25, 164–166, 221
Birth control and premarital sex, 214–217
Birth control and religion, 223, 225
Birth weight and smoking case-control study, 234–235
Breast cancer and mammograms, 28
Breast cancer and prednisolone, 61
Breast cancer and tamoxifen, 57
Buchanan votes for President, 85

Cancer and aspirin, 59
Cancer and smoking case-control study, 35
Cancer of larynx treatments, 61
Cancer remission, 113, 156
Carcinoma diagnoses by pathologists, 292–295
Carcinoma diagnosis agreement, 243–246
Cholesterol and cereal, 191
Clinical trial for fungal infections, 156
Clinical trial with Bayesian inference, 51
Close friends, 87
Coffee brand market share, 235–239
Coke tasting, 252
Crossover clinical trial, 270, 297

Death penalty and race, 53–56, 62, 117, 222
Depression clinical trial, 269–270, 297
Developmental toxicity, 192
Diagnostic testing, 26–28, 57

Endometrial cancer grade, 139–144
Environment opinions, 22, 227, 231, 240, 242, 275–276
Environmental threats exaggerated, 190
Evolution and political ideology, 78–82, 146

Fish hatching, 297

Global warming, 63, 251
Gore/Bush election and Buchanan vote, 84
Government spending, 224
Graduate admissions and gender, Berkeley, 154

An Introduction to Categorical Data Analysis, Third Edition. Alan Agresti.
© 2019 John Wiley & Sons, Inc. Published 2019 by John Wiley & Sons, Inc.
Companion Website: www.wiley.com/go/Agresti/CDA_3e

JWST911-ExInd JWST911-Agresti September 12, 2018 11:12 Printer Name: Trim: 254mm × 178mm

Graduate admissions and gender, Florida, 132–133, 297
Gunshot wounds and intent, 58

Happiness and family income, 59, 172–174
Happiness and heaven, 195–197
Heart attacks and aspirin, 30–34
Heart catheterization and race, 58
Heart disease and blood pressure, 135
Heart disease and cholesterol, 151
Heart disease and snoring, 69–72, 86, 148–149
Heaven and hell, 249
Heaven, belief in, 22
Homicide rates in Britain and US, 57
Horseshoe crabs
  classification table, 110–111
  classification tree, 303–305
  color and width predictors, 102–107, 127
  count response, 73–75, 220–221
  discriminant analysis, 301
  generalized additive model, 311
  marginal effects, 108–109
  model selection, 124–130
  monandrous or polyandrous, 152
  multiple predictors, 124–130, 152
  predictive power, 111–113
  ROC curve, 111
  width predictor, 91–93, 95–97

Infant malformation and alcohol use, 43–45, 84
Insomnia clinical trial, 260–263, 265, 284–285
Iris flower data, 321

Job satisfaction, 188
  by income, 61

Kyphosis risk factors, 115, 312–313

Lung cancer and gender, 57
Lung cancer and smoking in China, 155
Lung cancer clinical trial, 189
Lung cancer survival, 218–219

Mammograms, 28
Marijuana use, 99–100, 131
Marital happiness and family income, 189
Market basket data, 306
MBTI, 116–117, 120, 153
Mental impairment, 178–181, 183–184, 190
Movie ratings, 252
Multiple sclerosis ratings, 252
Murder rates, race and gender, 62
Murderer and victim race, 59
Myocardial infarction case-control study, 249

Nonmetastatic osteosarcoma and lymphocytic infiltration, 155

Obesity in children, 271

Pig farmers veterinary source, 270
Political affilation by race, 60
Political affiliation and global warming, 63
Political affiliation gender gap, 39–42
Political ideology and evolution, 78–82
Political ideology and party affiliation, 86, 169–178, 185
Political ideology and religion, 189
Premarital sex and birth control, 214–217
Presidential voting, 58
  election clustering, 308
Prostate cancer and PSA test, 57

Region of residence mobility, 250
Religion and birth control, 223, 225
Religious affiliation mobility, 250
Religious beliefs and education, 61
Respiratory illness and maternal smoking, 263–265
Russian roulette, 22

Same-sex marriage and political party, 58
Seat belts and injury, 58
Sex, extramarital and premarital, 251, 297
Sexual intercourse and gender, 226
Shopping destination choice, 167
Silicon wafer imperfections, 88
Simpson's paradox
  death penalty, 54, 62
  graduate admissions, 133
  statewide death rates, 62
Siskel and Ebert movie ratings, 252
Smoking and heart disease, 58
Smoking and lung cancer, 35, 58
Smoking prevention study, 289–291, 298
Student performance, 288–289
Student substance use, 198–203, 212–213, 268, 296
Student survey, 152, 315–318
Surgery meta-analysis, 285–288

JWST911-ExInd JWST911-Agresti September 12, 2018 11:12 Printer Name: Trim: 254mm × 178mm

Tea tasting experiment, 47–50
Tennis player ranking, 247–248, 252
Teratology experiment, 295
Titanic survival and gender, 58
Tonsil size and streptococcus, 186–187
Toy example of complete separation, 137
Travel credit card and income, 87

Vegetarianism, 23

World Cup odds, 57

SUBJECT INDEX

COPYRIGHTED MATERIAL

Adjacent-categories logit model, 184–185
Agreement, 243–246, 292–295
Agresti–Coull confidence interval, 9, 19, 332, 343
AIC, 128–130, 206, 293, 319
Artofstat website, 5, 10
Association, 25–63, 325–326
Autoregressive correlations, 257
Average marginal effect, 109, 146, 147, 178

Backward elimination, 125, 319
Baseline-category logit model, 160–167
  Bayesian fitting, 183
  goodness of fit, 164
  marginal distributions, 235–237
Bayesian inference
  comparing proportions, 50
  cumulative logit model, 183–184
  equal-tail interval, 14
  generalized linear models, 83
  HPD interval, 50
  introduction, 13–15
  large *p*, 321
  logistic regression, 140–143
  loglinear models, 203
  posterior interval, 14
  R software, 142
  SAS software, 336
  Stata software, 343–345
  two-way tables, 50
Bernoulli trial, 3
Beta distribution, 15–16, 50
Bias reduction in logistic regression, 143–144
Bias/variance tradeoff, 128–130, 311–312
  classification trees, 304, 306
  penalized likelihood, 314–315, 318
BIC, 129
Binary data, 2
  generalized linear models, 68–72
  grouped versus ungrouped, 72, 131
  logistic regression, 89–157
Binomial distribution, 3–5
  Bayesian inference, 15–16
  small-area estimation, 278
  small-sample test, 12–13
Binomial sampling, 28
Bonferroni method, 319
Bradley–Terry model, 247–248

Canonical link function, 67
Case-control study, 35–36, 234
  logistic regression, 93
  odds ratio estimation, 35–36
Categorical data, 1–348

An Introduction to Categorical Data Analysis, Third Edition. Alan Agresti.
© 2019 John Wiley & Sons, Inc. Published 2019 by John Wiley & Sons, Inc.
Companion Website: www.wiley.com/go/Agresti/CDA_3e

JWST911-SubInd JWST911-Agresti September 19, 2018 8:28 Printer Name: Trim: 254mm × 178mm

Chi-squared distribution, 37
  partitioning, 41–42
Chi-squared tests of independence, 36–46, 216
Classification
  discriminant analysis, 300–302
  multiple categories, 302
  tree-based, 302–306
Classification tables, 110–112
  discriminant analysis, 300–302
  tree-based prediction, 304
Classification tree, 302–306
  R software, 305
  SAS software, 342
  versus logistic regression, 305–306
Clinical trial, 36
Cluster analysis, 306–310
  R software, 309
  SAS software, 342
  Stata software, 346
Clustered categorical data, 253–298
  SAS software, 341
  SPSS software, 348
  Stata software, 346
Cochran–Armitage trend test, 46
Cochran–Mantel–Haenszel test, 101, 237
Cohen's kappa, 246
Collapsibility conditions, 211–213
Collapsing categorical scales, 176, 211–213
Complete separation, 136–140
Concordance index, 111
Conditional associations, 53, 197
Conditional distribution, 26
Conditional independence, 56, 101, 107
  graphs, 291
  logistic model, 101, 107
  loglinear model, 197, 201, 210–213
  significance tests, 101
Conditional logistic regression, 144–145, 233–235
Conditional maximum likelihood, 144, 233
Conservative inference, 13, 18, 49
Contingency table, 26, 325
Continuation-ratio logits, 186
Continuity correction, 18, 229
Cook's distance, 135
Correlation
  clustered data, 257, 275
  predictive power, 112–113, 182
  testing independence for ordinal variables, 43, 217
Credible interval, 14
Cross-product ratio, 33
Cross-validation, 111, 300
  k-fold, 315
Cumulative distribution function, 149, 168
Cumulative link model, 175–184
  Bayesian fitting, 183
Cumulative logit model, 167–184
  Bayesian fitting, 183–184
  goodness of fit, 176–178
  interpretations using probabilities, 178–181
  marginal distributions, 240
  marginal model, 260–263, 266
  non-proportional odds, 176–178
  random effects model, 284–285
  transitional model, 265
Cumulative odds ratio, 169
Cumulative probability, 149, 167, 171, 181, 240
  R, 10, 20
Cumulative probit model, 175, 181–182

Degrees of freedom interpretation, 38
Dendrogram, 307
Dependent proportions, 228–230
Dependent samples, 227
Deviance, 80–81
  analysis of, 80
  conservative for sparse data, 293
  goodness of fit, 130–132, 164, 200
  model comparison, 80
  residuals, 81, 134, 154
Diagnostic testing, 134–136
Difference of proportions, 29–30
  dependent samples, 230
Dimension reduction, 318–320
Dirichlet distribution, 16, 321
Discrete choice model, 166–167
Discriminant analysis, 300–302
  R software, 301
  SAS software, 342
  Stata software, 346
  versus logistic regression, 302
Dispersion parameter, 220
Dissimilarity index, 207
Dissimilarity, clustering, 306–307
Dummy variable, 98

Empty cell, 138
Exchangeable correlations, 257
Experimental study, 36
Explanatory variable, 2

Factors, 98
False discovery rate, 319–320
Firth penalized likelihood, 143

JWST911-SubInd JWST911-Agresti September 19, 2018 8:28 Printer Name: Trim: 254mm × 178mm

Fisher scoring, 82–83, 169, 328
Fisher's exact test, 46–50, 145
  $r \times c$ table, 49
Fixed effects, 273–274
Forward selection, 126, 153

Gauss–Hermite quadrature, 277–278
GEE method, 231, 236, 256–263
Generalized additive model, 310–313
  binary data, 91, 311–313
  count data, 74, 311
Generalized estimating equations (GEE) methods, 256–263
  binary matched pairs, 231
  multinomial responses, 260–263
  nominal matched pairs, 236
  ordinal matched pairs, 240
Generalized linear mixed models, 274–298
  logistic-normal, 275
  multilevel, 288–291
  ordinal, 284–285
  random intercept, 274
  random intercept and slope, 287
Generalized linear model, 65–88
  binary data, 68–72, 89–157
  components, 66–67
  counts, 72–76, 217–221
  fitting, 82–83
  inference, 76–82
  R software, 71
  SAS software, 334
  SPSS software, 347
  Stata software, 344
Geometric distribution, 22
GLM, 65
Graphical methods, 41, 210

Hierarchical clustering, 307–310
Hierarchical model, 288–291
High-dimensional data, 307, 313–321
  Bayesian methods, 321
Highest posterior density interval, 50
Homogeneous association, 56
  logistic model, 99, 101
  loglinear model, 197–200, 205, 207
  significance test, 101
Hosmer–Lemeshow test, 132
Hypergeometric distribution, 47, 145

Identity link function, 67, 146
Independence
  chi-squared tests, 36–46, 172–174, 216
  loglinear model, 194–195
Independence from irrelevant alternatives, 167
Independence graphs, 210–213
Indicator variable, 98
  coding, 98–101
Infinite estimates
  Bayesian shrinkage, 142
  finite with penalized likelihood, 144
  logistic regression, 136–140, 144
  multinomial models, 178
Influence diagnostics, 134–136
Interaction
  cumulative logit model, 260, 284
  loglinear model, 196
  random effects model, 287
  three-way contingency table, 197
Item response models, 283, 328
Iteratively reweighted least squares, 83, 327

Jeffreys prior distribution, 15, 50
  binomial parameter, 15
Joint distribution, 26

Kappa agreement measure, 246

Laplace approximation, 278
Lasso, 314–319
  R software, 316
Latent class models, 291–295
Latent variable, 147, 273
Latent variable model
  binary data, 147–150
  latent class models, 291–295
  ordinal data, 174–182
  probability comparison of groups, 180
Leverage, 81
Likelihood equations, 23, 82, 255
Likelihood function, 6, 77
Likelihood-ratio statistic, 11, 77
  contingency tables, 38
  deviance difference, 80
Likelihood-ratio test, 11
  GLM, 77
  GLMM, 278
  independence, 38
  logistic regression, 125, 130, 138, 140
  proportion, 12
Linear discriminant analysis, 300–302
Linear predictor (GLM), 66
Linear probability model, 68–69, 71, 146–147, 300, 301
Linear trend test of independence, 43, 216
Linear-by-linear association model, 215–217

JWST911-SubInd JWST911-Agresti September 19, 2018 8:28 Printer Name: Trim: 254mm × 178mm

Link function, 66
identity, 67, 146, 230
log, 67, 194
logit, 67, 89, 168, 231
probit, 147, 181
Local odds ratio, 185, 215
Logistic distribution, 149–150, 175
Logistic regression, 67–71, 89–157
ANOVA representation of factors, 100
Bayesian fitting, 140–143
case-control studies, 93, 234–235
categorical predictors, 98–102
clustered data, 254–263
collapsibility conditions, 213
conditional, 144–145, 233–234
confidence intervals for effects, 94
confidence intervals for probabilities, 96–98
effects, 90, 107–109
grouped versus ungrouped data file, 72
inference, 94–98
infinite estimates, 136–140, 144, 178
interaction, 106
linear approximation, 90, 108–109, 146
marginal effects, 108–109
marginal models, 231, 254–263
Markov, 263–266
matched pairs, 231, 233–235, 274–276
model checking, 130–136
model comparison, 104
model selection, 113, 123
multinomial, 159–192
multiple, 102–107
normal distribution implication, 94
penalized likelihood, 143–144
probability interpretations, 107–109
R software, 91
residuals, 132–134
retrospective studies, 93, 234
SAS software, 334
significance tests, 95
standardized interpretations, 109
Stata software, 344
subject-specific, 233, 255, 274–291
uncorrelated explanatory variables, 107, 213
Logistic-normal model, 275
Logit link function, 67
Loglinear models, 67, 193–226
Bayesian fitting, 203
collapsibility conditions, 211–213
conditional independence, 197
count response data, 217–221
goodness of fit tests, 200
homogeneous association, 197–200, 203, 205, 207
independence, 194–195
logistic model connection, 207–210
mutual independence, 197
ordinal variables, 214–217
quasi independence, 244
residuals, 201
saturated, 196–197
three-factor interaction, 205–206
LogXact, 145, 328

Machine learning, 302
Mantel–Haenszel statistic, 237
Marginal distribution, 26, 228
Marginal effects, 108–109, 178
Marginal homogeneity
T dimensional, 260, 296
c categories, 235–242
2 × 2 tables, 228–229
ordinal variables, 241
test for nominal variables, 235–239
test for ordinal variables, 240–242
Marginal models, 254–263
compared with GLMM, 276, 277, 282, 284–285
matched pairs, 230–232, 276
ordinal, 260–263, 284
versus GLMMs, 283
Markov chain Monte Carlo, 14
Markov logistic regression, 263–266
Matched pairs, 227–252, 274–276
Maximum likelihood estimate, 6
conditional, 144
infinite, 137
McNemar test, 228–229, 237, 249
Median effective level, 90
Meta-analysis, 285–288
Mid *P*-value, 13
confidence interval for odds ratio, 50
Fisher's exact test, 49
Missing at random, 267
Missing completely at random, 266
Missing data, 266–268
Mixed-membership model, 295
Mixture models, 291
latent class, 291–295
ML estimate, 6
Model averaging, 129
Model comparison
deviance, 80
logistic regression, 104

JWST911-SubInd JWST911-Agresti September 19, 2018 8:28 Printer Name: Trim: 254mm × 178mm

Model selection
  high-dimensional, 318–320
  logistic, 123–130
Model smoothing, 311
Monte Carlo methods, 14, 49, 278
Mosaic plot, 41
Multicategory logit models, 159–192
Multicollinearity, 143, 314
Multilevel model, 288–291
Multinomial distribution, 5, 159, 194, 321
Multinomial models, 159–192
  GLMMs, 284–285
  large *p*, 321
  marginal models, 260–263
  SAS software, 336
  SPSS software, 348
  Stata software, 345
Multinomial sampling, 28, 194
Multiple comparisons, 320
  false discovery rate, 319–320
Multiple correlation, 112
  binary GLM, 112
  ordinal model, 182
Multiple imputation, 267–268

Natural parameter, 67
Negative binomial regression, 220–221
Newton–Raphson algorithm, 83, 292
Nominal variables, 2
  loglinear models, 194–213
  multicategory logit models, 159–167, 235
Null deviance, 80
Null model, 77, 80, 96

Observational study, 36
Odds ratio, 31
  agreement, 245
  case-control studies, 35
  conditional and marginal, 55
  confidence interval, 33–34, 94
  confidence interval for small *n*, 50
  local, 185, 215
  logistic regression, 90, 143
  loglinear parameters, 196, 198, 199, 202, 205–206
  relative risk approximation, 36
Offset, 217
One-standard-error rule, 305, 315
Ordinal quasi symmetry, 241–242
Ordinal tests
  greater power, 45, 172–174
Ordinal variables, 2
  GLMMs, 284–285
  loglinear models, 214–217
  marginal models, 260–263
  multicategory logit models, 167–184
  test of independence, 42–46, 216–217
Ordinary least squares
  linear discriminant analysis, 300
  linear probability model, 146
Overdispersion, 75–76, 217, 220–221
  quasi-likelihood methods, 256

Parsimony, 128
  model smoothing, 128–130, 311, 318
Partial tables, 53
Partitioning chi-squared, 41–42
Pearson chi-squared statistic, 37, 130
  Pearson residual contributions, 81
  score test, 78
  standardized residuals, 134
Pearson residual, 81, 134, 154
Penalized likelihood, 143–144, 314
Perfect discrimination, 136
Poisson distribution, 72, 193, 194
  loglinear models, 226
  overdispersion, 75–76, 217
Poisson GLM, 73, 194–221
Poisson loglinear model, 73–75, 194–226
Poisson regression, 73–76, 220
Population-averaged effects, 233, 255, 276
Posterior distribution, 14
Posterior interval
  2 × 2 tables, 50–52
  highest posterior density, 50
  logistic parameter, 142–143
  odds ratio, 51, 203
  proportion, 14
Power and sample size determination, 150
Predictive power, 110–113, 182
Principal component analysis, 319
Prior distribution, 14, 16, 51
  beta, 15, 50
  binary response probabilities, 15–16
  conjugate, 15, 50
  Dirichlet, 321
  Jeffreys, 15
  normal, 83, 141
  spike-and-slab, 321
  uniform, 15, 321
Probit model, 147–150, 175, 181–182, 327
Profile likelihood confidence interval, 78, 79, 95, 144

Proportion
  confidence intervals, 8–9
  difference, 29–30
  difference for dependent samples, 230
  significance tests, 7, 13
Proportional hazards model, 219
Proportional odds, 169
  adjacent-categories logits, 184–185
  cumulative logits, 168–184
  sequential logits, 186–187
Pruning classification tree, 304
Purposeful selection, 126–127

Quadrature points, 277
Qualitative variable, 1
Quantitative variable, 1
Quasi independence, 244, 246
Quasi symmetry, 238–239
  agreement modeling, 245
  Bradley–Terry model, 247
  ordinal, 241–242
  testing marginal homogeneity, 238, 241
Quasi-complete separation, 136–140, 144, 177
Quasi-likelihood, 255–257, 259

R (software), 331
  Anova function, 78, 95, 100, 104, 124, 202, 216
  CMHtest function, 44
  MCMCpack package, 142, 203
  PropCIs package, 20, 29, 31, 34, 51, 229, 230
  VGAM package, 91, 161, 170, 173, 174, 185
  anova function, 105, 260, 313
  bestglm function, 129
  binom package, 19
  brglm2 package, 137, 140
  car package, 95, 100, 104
  chisq.test function, 40
  clmm function, 284
  cond package, 141
  confint function, 78
  diffci.bayes function, 51
  dist function, 309
  epitools package, 34, 49, 50
  exact2x2 package, 49
  exactci package, 20
  fisher.test function, 49
  gam function, 74, 313
  gam package, 313
  geeglm function, 260
  geepack package, 260
  gee function, 231, 258
  glm.scoretest function, 78
  glmer function, 275, 279, 281, 286, 287, 290
  glmnet package, 316
  glm function, 71, 74, 78, 91, 95, 96, 99, 102, 104–106, 110, 112, 113, 116, 124, 128, 131, 134, 137, 139, 141, 146, 148, 195, 196, 199, 201, 202, 208, 216, 219, 239, 242, 244, 247, 257, 264, 281, 291, 316, 318
  hclust function, 309
  lda function, 301
  lme4 package, 275, 279, 281, 286, 287, 290
  logistf package, 144
  logitmfx function, 109
  loglm function, 195
  lrtest function, 162, 173, 174
  mcnemar.test function, 229
  mfx package, 109
  multgee package, 236, 240, 261, 262
  ocAME function, 180
  orci.bayes function, 51
  ordinal package, 284
  p.adjust function, 320
  pROC package, 112
  poLCA function, 294
  polr function, 179, 181, 182
  predict function, 179
  prop.test function, 18, 31
  prune function, 305
  psych package, 246
  read.table command, 17
  rpart package, 305
  statmod package, 78
  stepAIC function, 129
  vcdExtra package, 44
  vglm function, 161, 162, 165, 170, 171, 173, 174, 177, 185, 186
  introduction, 17–18
R-squared
  binary data, 112
  ordinal model, 182
Random component (GLM), 66
Random effects models, 273–298
Random intercept, 274, 285
Random slope, 285–288
Rasch model, 283, 328
Rater agreement, 243–246, 292–295
Rates, modeling, 217–219
Regularization methods, 143–144, 313, 321
Relative risk, 30–31
  odds ratio approximation, 36
Repeated measurement, 253–298

JWST911-SubInd JWST911-Agresti September 19, 2018 8:28 Printer Name: Trim: 254mm × 178mm

Residual deviance, 80
Residuals
  contingency table, 39
  deviance, 81, 134, 154
  GLM, 81–82
  logistic regression, 132–135
  loglinear models, 201
  Pearson, 81, 134, 154
  standardized, 39, 81–82, 132–134, 201
Response variable, 2
Retrospective study, 35
  logistic regression, 93
Ridge regression, 314
ROC curves, 111–112

Sample size determination
  comparing proportions, 150
  logistic regression, 150–151
Sampling zero, 138
Sandwich standard errors, 257
SAS, 332–342
Saturated model, 80
  logistic, 130
  loglinear, 196–197
Score confidence interval
  dependent proportions, 230
  difference of proportions, 29, 31
  odds ratio, 34
  proportion, 9, 18, 19
  relative risk, 31
Score test, 11
  correlation, 43, 217
  GLM, 78
  independence, 38
  marginal homogeneity, 237
  proportion, 12
Scores, choice of, 45, 87, 106, 216
Sensitivity and specificity, 26–28
Separation
  data, 136–140
  variables, 210–212
Sequential logit model, 186–187
Shrinkage estimator, 143, 279, 314, 315, 318
  Bayesian, 321
Sign test, 22
Simpson's paradox, 54, 133, 326
Small-area estimation, 278–281
Small-sample inference
  $r \times c$ table, 49
  2×2 table, 46–50
  binomial, 12–13
  logistic regression, 145
Smoothing
  binary data, 311
  generalized additive model, 310
  penalized likelihood, 314–318
Software summaries, 331–348
Sparse data, 138, 143, 164
  deviance is conservative, 293
Sparse structure, 318
Spline functions, 311
SPSS, 346–348
Square contingency tables, 227–252
Standardized coefficients, 109
Standardized residual, 39
  generalized linear model, 81–82
  independence, 39, 243
  logistic regression, 132–134, 154
  loglinear models, 201
  symmetry in square table, 238
Standardized variables, 109, 141, 314
Stata (software), 342–346
Statistical independence, 28
StatXact, 145, 328, 331
Stepwise variable selection, 125–127
Subject-specific effects, 232–234, 255, 276
Survival model, 218–219
Symmetry, 237–239, 245

Transitional models, 263–266
Tree-based classification, 302–306
Trend tests of independence, 42–46, 216

Uniform association, 215
Uniform distribution, 15, 321

Variable selection, 123–130, 318–320
  high-dimensional, 313–321
Variance component, 274, 277

Wald confidence interval, 11, 78
  difference of proportions, 29, 31
  GLM parameter, 78
  logistic parameter, 94
  odds ratio, 33–34
  proportion, 8, 11
Wald test, 10, 138
  GLM, 77
  infinite logistic estimate, 138
  logistic parameter, 95
  proportion, 12
Weighted kappa, 246
Weighted least squares, 83
Wilson confidence interval, 9, 332, 343