

Subject Index

Symbols

β -binomial, **104**
 ϕ coefficient, 222
: (sequence operator), 33
80–20 rule, **26**

A

`abline()`, 269
`add1()`, 472, 502
added-variable plot, **312, 314**
 properties, 316
`addmargins()`, 45, 57
adjusted inertia, **246**
adjusted means, **278**
adjusted residual, **356**
AER package, 449, 468, 502, 503
`aes()`, 367
`aggregate()`, 51, 447
agreement, 146–153, 170
 Cohen's κ , 148–149
 intraclass correlation, 148
 observer agreement chart, 150
 observer bias, 152–153
 partial, 150–151
agreement chart, **17**
`agreementplot()`, 151, 152, 159
agridat package, 19, 217
AIC, 295, 355, 380, 390, 395, 434
`AIC()`, 356, 434, 462
Akaike Information Criterion, **355**
`allEffects()`, 279, 280, 294, 466, 474
alpha-blending, **20**
analysis of deviance, **434**
analysis of variance, 7, 127, 178, 262,
 350–351, 376
animation package, 21
`Anova()`, 287, 288, 292, 321, 322, 327,
 339, 347, 360, 366, 434, 501, 502

`anova()`, 183, 267, 271, 300, 322, 328,
 359, 360, 372, 380, 391, 411, 412,
 434, 462, 473, 477
`aperm()`, 44
`apply()`, 48, 51, 52, 79, 192, 195, 487
`array()`, 36
arrays, 32, 35–37
 creating, 58–59
`as.array()`, 212
`as.data.frame()`, 54, 63, 212, 213, 359
`as.data.frame.table()`, 54
`as.matrix()`, 63, 234, 256
`as.numeric()`, 54, 377, 379
`as.table()`, 42
aspect ratio, 27, **227**
assignment operator, 32
`assoc()`, 168, 357
association, **8**
 partial, 188–197
association graph, **180**
association ordering, **18**
association parameters, 351
association plot, **116, 145**, 145–146, 222
`assocstats()`, 9, 124, 125, 128,
 157–159
asymmetric CA plot, **249**
asymmetric map, **223, 224**
`avPlot()`, 315
`avPlots()`, 315–317
axes, 26, 242
 equating, 227

B

badness-of-fit, 355
bar graph, 75, 170
bar plot, 201, 436, 453
bar plots, 80, 117–118
`barplot()`, 436

baseline models, **179**, 341–345, 365
 Bayesian Information Criterion, **356**
 BIC, 295, 356, 380, 390, 395, 434
 BIC(), 434, 462
 binary, 116, 325, 363, 405–406
 binary events, 66
 binary response, **262**
 binary tree, **22**
 binary variables, **4**, 66–69, 74–76, 121, 209, 241, 262
 binom.or(), 421
 binom2.or(), 406, 411
 binomial, 490
 binomial distribution, 66–69, 74–76, 98, 99, 102–104, 262
 calculation of, 75
 examples, 66–69, 90–91
 moments of, 74
 plotting, 76, 94, 102–104
 binomial samples, 66–69, 74–76, **121**, 431
 binreg_plot(), 276, 301, 302
 biplot, **221**, **246**, **248**, 248–251, 254, 388
 biadditive, 252–254
 biplot(), 224, 253
 biplots, 27, 389
 bivariate, 405–406
 bivariate logistic model, **405**
 bivariate loglinear model, **406**
 blogits(), 407, 414, 419, 423
 Bonferroni correction, 481
 boxplot, 201, 437
 bs(), 411
 Burt matrix, **241**, **243**

C

c(), 32, 40, 50
 CA(), 225
 ca package, xvii, 224, 225, 244, 249
 ca(), 225, 226, 234, 238, 239, 241
 cabipl(), 250
 canonical analysis of categorical data, **222**
 canonical correlation, **221**
 car package, xvii, 287, 292, 306–308, 311, 313, 315, 321, 322, 327, 336, 347, 360, 366, 433, 434, 482, 484, 486, 496, 501, 502
 case form, **5**, 31, 39–40, 270
 converting, 53
 categorical data, 31, 39–44
 count, 69–72, 76–86
 definition, 4

 dichotomous, 66–69, 74–76
 fitting, 87–94
 statistical methods, 4
 stratifying, 127–129
 type-token, 72–73
 categorical frequency distributions
 diagnosing, 95–99
 categorical variable, **4**, 323
 causal models, 186–188
 cbind(), 37, 497
 cd_plot(), 455
 cdplot(), 455
 chisq.test(), 9, 165, 174
 Cicchetti Allison weights, 149
 classification, 146
 clustered data, 430
 CMHtest(), 125–128, 157, 158
 Cochran–Mantel–Haenszel tests, **125**
 general association, 125–126
 linear association, 126–127
 row means differ, 125
 stratified analysis, 127–129
 coef(), 266, 328, 357, 413, 497
 coeftest(), 266
 Cohen's κ , 148–149
 coindep_test(), 173, 175, 194
 collapse.table(), 53, 60
 color palettes, 27
 Hue–Chroma–Luminance, 27
 Hue–Saturation–Value, 27
 colorspace package, 171
 column effects model, **377**
 complete independence, 180
 complete separation, **276**
 component-plus-residual plot, **312**
 compositional data, **153**
 conditional(), 184
 conditional association, 195, 352
 conditional density plot, **455**
 conditional distributions, 118, **120**, 164, 326
 conditional independence, 181, 189, 191, 194, 207, 358, 360
 conditional plots, **275**, 313, 415
 conditional probabilities, 222
 conditioned plot, 143
 conditioning plot, **141**
 confidence bands, 268
 confidence ellipses, 389
 confint(), 123, 133, 274
 confint.Kappa(), 149
 constraints

zero-sum, 353
 constraints(), 496, 497
 constructed variable plot, **312**
 contingency ratios, **249**
 contingency table, 7, **115**, 116, 119, 147,
 188, 205–209, 221, 240, 248–251,
 270, 352
 contingency tables, 352
 plotting, 138
 contour(), 479
 contr.sum(), 351
 contr.treatment(), 351
 contrasts, 127, 273, 353, 376
 sum-to-zero, 351
 treatment, 351
 controlled comparison, **135**
 Cook's distance, 303, **305**, 482, 484
 cooks.distance(), 306, 482
 coplot, **141**, 143, **190**
 correlated data, 430
 correlation, 126, 221, 349, 404, 490
 corresp(), 224, 225
 correspondence analysis, 221–254, 350, 382
 asymmetric map, 224
 interactive coding, 232
 principal coordinates, 223
 properties, 224
 stacking, 232–238
 standard coordinates, 223
 supplementary variables, 238–240
 symmetric map, 224
 two-way tables, 222–231
 vs. mosaic displays, 231–232
 correspondence matrix, **223**, 248
 corrgram, 201
 corrrplot package, 378
 cotabplot(), 152, 168, 190, 195, 196,
 218, 417
 count, **66**
 count data, **6**, 69–86, 109, 118–119, 175,
 430, 433, 446, 490
 distributions, 73–109
 plotting, 80–82, 117–118
 count metameter, **99**, **100**
 countreg package, 440, 450, 453, 454, 457,
 471
 covariate, **107**
 Cramer's V, **9**, 128
 criterion variables, *see* response variables
 cross-classification, 121
 cross-sectional study, **121**

Crossings(), 396
 crossings model, **396**
 CrossTable(), 120
 crPlot(), 313
 crPlots(), 313
 cumulative logit, **324**
 cut(), 442
 cutfac(), 442, 447, 469
 cutq(), 447

D

data
 case form, 5, 7, 39–40, 270
 frequency form, 6, 7, 40–41, 270
 frequency vs. count, 6
 table form, 41–47
 data analysis
 hypothesis testing, 10
 hypothesis testing approach, 8
 model building, 10–13
 data ellipse, 201, **490**
 data frame, **32**, 37–38, 105
 converting, 53
 creating, 58–59
 importing, 38
 multinomial, 121
 subsetting, 48, 50, 60
 data plots, 25
 data reduction, 221
 data sets
 Abortion, 61, 158
 Accident, 219, 257, 322
 AirCrash, 216, 255, 502
 Arbuthnot, 66, 109, 320
 Arrests, 291
 Arthritis, 5, 7, 39, 55, 56, 119, 124,
 127, 174, 202, 263, 266–268, 273,
 275, 326, 336
 arthritis treatment, 124, 128
 barley, 19
 Bartlett, 197, 204, 217
 Berkeley admissions, 134
 birthwt, 321
 Bundesliga, 111
 Butterfly, 96
 Caesar, 61, 321, 372
 caith, 218, 255
 case2201, 501
 CoalMiners, 135, 407, 409
 CodParasites, 457, 503
 Cormorants, 502

- CrabSatellites, 440, 446, 451, 455
- criminal, 216, 255
- CyclingDeaths, 71, 111
- DanishWelfare, 61
- DaytonSurvey, 51, 52, 61, 372
- death by horse kick, 97, 102
- Depends, 112
- Detergent, 373
- Donner, 282, 287, 320
- Employment, 190, 192
- Federalist, 98, 110
- Federalist Papers, 98
- Geissler, 62, 105, 111
- Gilby, 255
- gss8590, 426
- HairEye, 8
- HairEyeColor, 41, 43, 48, 116, 139, 146, 179, 182, 225
- HairEyePlace, 218, 255
- HallOfFame, 112
- Hauser79, 392, 396
- Health, 369
- HorseKicks, 54, 56, 57, 70, 88, 91, 101
- Hospital, 158
- housing, 346
- Hoyt, 61
- ICU, 296, 308, 317, 320
- iris, 20
- jansen.strawberry, 217
- JobSat, 43, 158, 254
- Lifeboats, 155, 159
- Mammograms, 151, 159
- Master, 217
- Mental, 118, 228, 377, 379, 381, 383, 387
- minnesota.barley.yield, 19
- MSPatients, 147, 152
- NMES1988, 468, 491, 503
- PhdPubs, 108, 438, 445, 447, 450, 451, 484, 503
- PreSex, 186, 198, 244, 256
- Punishment, 193
- quine, 501
- RepVict, 229
- Saxony, 62, 103, 105, 107
- SexualFun, 147, 149
- Space shuttle disaster, 10–11
- SpaceShuttle, 270
- struc, 207
- Suicide, 235
- Titanic, 155, 210, 214, 220, 246
- Toxaemia, 415, 421, 423–425
- TV, 231, 256, 347
- UCBAdmissions, 61, 63, 116, 129, 201, 210, 257, 270, 356, 357, 359, 364, 367
- UKSoccer, 62, 78, 79, 111, 175, 253
- Vietnam, 219, 256, 347
- Vision, 389
- visual acuity, 141
- VisualAcuity, 63, 139, 141, 159, 390, 400, 426
- WeldonDice, 69, 75, 90
- Women1f, 336, 339, 346
- WomenQueue, 98, 110
- data visualization, **3**
- data+model plots, 26
- `data.frame()`, 37, 40, 54
- datasets** package, 41, 116
- `datasets()`, 61
- `dbinom()`, 74, 76
- `ddoublebinom()`, 105
- `ddply()`, 52
- `density()`, 455
- density plot, 201
- dependent variables, *see* response variables
- `devar()`, 413
- deviance, 267, 304, 354, **354**, 433, 444
- deviance residual, **304, 356, 481**
- deviance test statistics, 179
- DFBETA, 303, **305**
- DFFITS, **305**, 482
- `dffits()`, 306
- `dgeom()`, 74
- `Diag()`, 390, 396, 401
- diagnostic plots, 480–489
- dichotomous, 116, **263**
- dichotomous variables, **4**, 66, 116, 121, 209, 262, 405
- dictotomous variables, 406
- `dim()`, 34
- `dimnames()`, 34
- direct labels, 27, 302
- `direct.label()`, 81
- directlabels** package, 27, 80, 331
- discrete frequency distributions, 65–109
 - characteristics of, 73–74
 - plotting, 80–82, 102–104
- dispersion parameter, **83, 431**, 432, 433
- `dispersiontest()`, 449, 502, 503
- `distplot()`, 101, 103, 109, 111, 112
- distributions, 73–109

- binomial, 262, 430
- diagnosing, 95–99
- double binomial, 105
- fitting, 87–94
- gamma, 430, 431
- geometric, 446
- hypergeometric, 124
- normal, 430
- Poisson, 350, 354, 359, 430–432, 438
- `dlogseries()`, 74
- `dnbinom()`, 74, 84
- `dotchart()`, 385
- double binomial distribution, **105**
- `doubledecker()`, 168, 209, 373
- doubledecker plot, 201, 209–211
- doubledecker plots, **167**
- `dpois()`, 74, 80
- dual scaling, **222**, *see* correspondence analysis
- dummy coding, 273
- dynamic graphics, **21**, 203–204

E

- `Effect()`, 279–281, 333
- effect displays, **278**
- effect ordering, 18–20, **166**, 232
- effect plot, 26, 205, 278–281, 439
 - count data models, 465–467, 473–476
 - proportional odds model, 332–333
- effect-order sorting, **18**
- effects package, xvii, 278–280, 291, 294, 332, 342, 345, 419, 465–467, 499
- empirical logits, **407**
- `eqsplot()`, 27
- events/trials form, **270**
- excess zeros, 451–455
- `exp()`, 274, 413
- `expand.dft()`, 56, 79
- `expand.grid()`, 40, 80, 84
- explanatory, **116**
- explanatory model, 186
- explanatory variables, **7**, 120–121
- exponential family, **104**, **430**
- extracat package, 118

F

- `facet_grid()`, 275, 331, 460
- `facet_wrap()`, 275
- facets, **275**
- FactoMineR package, 225
- factor, **37**, 43–44, 51–447

- contrasts, 353
- ordinal, 400
- factor analysis, 325
- Fisher's exact test, 124
- `fisher.test()`, 124
- fit statistics, 193
- `fitdistr()`, 89
- `fitted()`, 206, 358, 381, 393, 413, 447
- fixed zeros, *see* structural zeros
- Fleiss-Cohen weights, 149
- `fluctile()`, 118
- fluctuation diagram, **118**
- foreign package, 38
- `fourfold()`, 135, 157, 158, 372
- fourfold display, **17**, 130–135
 - confidence rings, 133
- `fractions()`, 264
- frequencies
 - joint, 121
 - marginal, 121, 163
- frequency, **66**
- frequency data, **6**
- frequency form, **6**, 31, 40–41, 51–52, 115, 270, 359, 365, 392
 - converting, 53
- `ftable()`, 42, 46, 47, 61, 63, 233, 234

G

- `gam()`, 478, 479
- gamma distribution, 431
- gdata package, 38
- generalized additive model, **478**
- generalized linear mixed models, **430**
- generalized linear model, **262**, **264**, **430**, 429–499
 - assumptions, 490
 - components, 430–431
 - count data, 430, 435–440
 - diagnostic plots, 480–489
 - excess zeros, 451–455
 - goodness-of-fit, 433–434, 450–451
 - hypothesis tests, 432–433
 - leverage, 481
 - multivariate, 349
 - multivariate response, 489–499
 - negative binomial, 446
 - non-nested models, 434–435
 - overdispersion, 444–446
 - plotting, 447
 - quasi-Poisson, 445–446
 - variance functions, 431–432

generalized linear models, 104–109
 generalized logit model, **324**, 341–345
 plotting, 344
 generalized pairs plot, **197**
 geom_boxplot(), 460
 geom_density2d(), 503
 geom_jitter(), 460
 geom_point(), 269, 271, 275
 geom_smooth(), 271
 geometric distribution, 85–86, 446
 calculation of, 86
 moments of, 85–86
 getContrasts(), 384
 GGally package, 202
 ggplot(), 269, 275, 285, 331
 ggplot2 package, 13, 27, 80–82, 155, 202,
 268, 269, 271, 275, 276, 286, 340,
 367, 424, 437, 460, 461, 494, 503
 ggtern package, 155
 ggtern(), 155
 ggvis package, 21
 GLM, *see* generalized linear model
 glm(), 7, 104–106, 262, 265, 266, 270,
 271, 273, 276, 279, 297, 304, 321,
 322, 350, 351, 353, 354, 359–361,
 365, 368, 369, 371, 377, 379, 383,
 409–412, 415, 419, 426, 442, 445,
 446, 454, 499
 glm.nb(), 446, 467, 479, 482, 491, 497,
 499
 glmlist(), 398
 GLMM, *see* generalized linear mixed model
 gls(), 279
 gnm package, xvii, 383, 384, 390
 gnm(), 383, 387, 395, 426
 goodfit(), 89–91, 94, 103, 109–112
 goodness-of-fit, **87**, 295, 380, 390, 433,
 450–451
 loglinear models, 354
 googleVis package, 21
 gpairs package, xvii, 197, 202
 gpairs(), 202, 321, 441
 grapcon functions, **168**
 graphical elements, 167
 annotations, 27
 attributes, 26
 axes, 82, 277
 axis labels, 26
 color palettes, 27
 controls, 21
 labeling, 230

 legends, 26, 80, 82
 lines, 268
 overplotting, 237
 shading, 166, 169–175
 title, 26
 graphics
 80–20 rule, 26
 analysis graphics, 15–16
 aspect ratio, 27
 basic functions, 14
 color, 27
 data plots, 25
 data+model plots, 26
 design principals, 23
 design principles, 13–16, 135
 effect plots, 26
 interactive, 21–22, 203–204, 207
 model plots, 25
 presentation graphics, 16
 rendering, 18
 graphics package, 47, 455
 grid package, 170, 202
 grouped barchart, **117**

H

half-normal plot, **486**, 485–486
 half-normal plots, 489
 hanging rootogram, **93**
 hatvalues(), 304
 heplot(), 491
 heplot3d(), 491
 heplots package, 491
 heterogeneous uniform association, **400**
 hierarchical models, **178**, 351, **351**, 430
 high-order terms, **180**
 HistData package, 67, 109
 histograms, 80, 201
 hmmm package, 404
 homogeneity, 121, 141
 homogeneity analysis, **222**, *see*
 correspondence analysis
 homogeneity of association, **8**, 129–130
 homogeneous association, 354, 358, 360,
 363, 400
 homogeneous association model, **352**
 hurdle(), 454
 hurdle model, **454**, 454
 hypergeometric distribution, 124
 hypothesis
 independence vs. homogeneity, 121
 hypothesis matrix, 433

hypothesis testing, 8–10
hypothesis-error plot, **490**, 491–492

I

ICC(), 148
identify(), 27
image(), 479
importing data, 38
independence, 121, 140, 143, 145, 176–180, 350, 382, 393, 404
 classes of, 180
 in square tables, 389
independence_table(), 138, 164
independent, 175
independent variables, *see* explanatory variables
index of dispersion, **79**
indicator matrix, 240–246
indicator variables, 240
inertia, **222**
influence, **303**, 482–485
 joint, 315
 plotting, 306
influence.measures(), 306, 310
influenceIndexPlot(), 307
influencePlot(), 306, 307, 484
inter-rater agreement, **116**, *see* agreement
interaction(), 63, 234
interactions, 178, 294
interactive coding, 232, **232**
interactive graphics, 21, 203–204, 207
interp(), 172
interpolation, 172
intraclass correlation, 148
iplots package, 21

J

jitter, 268, 437
joint(), 184
joint correspondence analysis, **246**
joint dependence, 404
joint distribution, **120**, 389, 403, 490
joint frequencies, 121
joint independence, **145**, 176, 179, 181, 184, 186, 187, 189, 207–209
joint influence, **315**

K

Kappa(), 149
kde2d(), 503
KernSmooth package, 286

knitr package, xvii, 28, 523
Kruskal-Wallis test, 125
Kway(), 400, 401

L

labeling_border(), 169
labeling_cboxed(), 169
labeling_cells(), 169
labeling_conditional(), 169
labeling_doubledecker(), 169
labeling_lboxed(), 169
labeling_left(), 169
labeling_left2(), 169
labeling_list(), 169
labeling_residuals(), 169
labeling_value(), 169
Lahman package, 112, 217
latent class analysis, 325
latent variable, 325, 452
LaTeX, 56–58
lattice package, 19, 27, 80
least squares means, **278**
legend(), 230
legend_fixed(), 169
legend_resbased(), 169
legends, 27
leveled plot, 100, 102
levels(), 283, 345
levels model, **394**
leverage, 303, 304, 481, **481**
likelihood ratio test, 88, 125, 183, 186, 327, 354, 370, 433, 444
line graph, 80
line plot, 367
linear hypothesis, **433**
linear logistic regression model, **264**
linear logit model, **264**
linear predictor, **273**, **430**
linear probability model, **263**
linear probit regression, **264**
linear regression, 263–264
linear-by-linear model, **376**, 377
linearHypothesis(), 433, 434, 496, 498
lines(), 269, 447
link function, **430**, 432
list, 32, 89
lm(), 7, 266, 279, 304, 351, 491
lme(), 279
lme4 package, 279
lmer(), 279



lmtree package, 266, 422, 462
lmtest (), 471
 local odds ratio, **211**
loddsratio (), 123, 136, 212, 377, 381, 392, 418, 493
loess, 281, 313, 482
 log odds, **10**, **122**, 264, **264**
 log odds ratio, **122**, 252, 406
 local, 377
 log-multiplicative models, 382–389
 plotting, 385–389
 logarithmic series distribution, 86, 96, 98
 moments of, 86
logi.hist.plot (), 265
 logistic regression, 10–13, 122, 262–270, 325, 341–345, 404, 454
 fitting, 265–267
 influence diagnostics, 304–306
 influential cases, 317
 interpretation, 264–265
 leverage, 304
 model tests, 267
 multiple, 272
 multivariate, 349
 nested models, 335–341
 plotting, 265–275
 plotting, 265
 residuals, 303–304
logit, **122**, **264**
 logit function, **10**
 logit models, 363–367, 404
 interpretation, 365–366
 plotting, 366–367
logLik (), 434, 462
loglin (), 178, 184, 353, 354, 368
loglin2formula (), 184, 220
loglin2string (), 184
loglinb2 (), 406
 loglinear independence model, **351**
 loglinear model, **177**
 multivariate responses, 405–406
 loglinear models, 130, 177–179, 205–209, 349–363
 as GLMs, 352–353
 extensions, 375–425
 fitting, 179–183, 353–354, 357
 goodness-of-fit, 354–356
 independence model, 351
 multivariate responses, 375, 403–414
 ordinal variables, 376–389, 396–399
 plotting, 360–363, 381–382

 residuals, 356–357
 square tables, 389–399
loglm (), 130, 143, 178–180, 183, 184, 190, 192, 194, 206, 216, 218, 219, 350, 354, 357, 360, 368, 371–373, 426
logmult package, xvii, 216, 255, 383, 386, 387, 426
lowess (), 269
lrm (), 297, 301, 324
LRstats (), 183, 267, 288, 372, 399, 434
LRtest (), 297, 338
lrtest (), 328, 422, 462, 463
lsmeans package, 278

M

main effect, 351
 main-effect ordering, **18**
manipulate package, 21
margin.table (), 44, 45, 51, 52
 marginal distributions, **120**, 403
 marginal frequencies, **120**, 121, 163
 marginal homogeneity, 148, **152**, **389**, 391
 marginal model, **238**
 marginal relationships, 199, 415
 Marimekko chart, **170**
markov (), 184
 masking, **315**
MASS package, 27, 56, 89, 130, 178, 183, 218, 224, 225, 255, 264, 279, 298, 324, 326, 346, 354, 357, 446, 453, 501, 502
matplot (), 76, 407, 413
matpoints (), 413
 matrices, 32–35, 222, 279, 489
 binding, 35
 transposing, 35
matrix (), 34
mca (), 55, 224
mcja (), 242
mcnemar.test (), 159
 mean function, **431**
 mean-square contingency coefficient, 222
melt (), 331, 340
 method-of-moments, 445
mgcv package, 478
mjca (), 225, 241, 244, 246, 247, 256, 257
mlogit package, 342
 model building, 10–13, 261, 434
 model comparison plot, **399**
 model comparisons, 183, 186

model matrix, **353**, 489
 model object, **262**
 model plots, 25
`model.matrix()`, 240
`mosaic()`, 159, 168, 170, 171, 174, 179,
 184, 190, 193, 196, 202, 216, 218,
 219, 357, 361, 370, 372, 416
 mosaic display, **9**, 15, 17, 47, **161**, 161–215,
 222, 237, 244, 350, 370, 376, 379,
 402, 415
 3D, 203–204, 207
 conditional, 196
 generalized matrix, 201–203
 interpretation, 166
 matrix, 197
 shading, 166, 169–175
 three-way tables, 193
 two-way tables, 162–166
 vs. correspondence analysis, 231
 mosaic displays
 vs. correspondence analysis, 232
 mosaic matrix, **197**, 240, 244
`mosaic3d()`, 204, 208, 357
`mosaicplot()`, 47, 169
`Mult()`, 383
 multi-way table, 133, 141, 161, 203–204,
 232–238, 240, 354, 400–403
 fitting, 179
 plotting, 176
`multinom()`, 279, 324, 342, 347
 multinomial, 323
 multinomial logit model, **324**, 341–345
 plotting, 344
 multinomial sample, **121**
 multiple correspondence analysis, 240–242
 bivariate, 240–242
 multivariate, 243–246
 multiple testing, 481
 multiplicity, 481
 multivariate data, 243–246
 multivariate linear model, 404, **489**
 multivariate responses, 403–406, 414
`mutual()`, 184
 mutual independence, 184, 206–208, 353,
 357, 404

N

natural spline, **287**
 negative binomial distribution, 82–85, 98,
 99, 108
 calculation of, 84

examples, 92
 model, 446
 moments of, 83
 plotting, 84
 nested dichotomies, **324**, **335**, 335–341
 plotting, 340–341
 nested models, 335–341, **355**, 355, 359, 377,
 391, 434
 nlme package, 279
 nnet package, 279, 324, 342
`nobs()`, 434
 nominal variables, **4**, 66, 323
 response, 335–341
 nomogram, **301**
`nomogram()`, 301
 non-nested models, 434–435
 non-parametric regression, 268
 nonlinearity, 294
 nonnest2 package, 435
 normal QQ plots, **486**
 normal quantile plots, **486**
`ns()`, 287, 411, 414
 null model, 140

O

observer agreement, **146**, 146–153
 observer agreement chart, 150
 odds, **121**
 odds ratio, **23**, **121**, **122**, 121–124, 129–132,
 135, 137, 274, 377, 392, 406
 conditional, 352
 confidence interval, 123
 generalized, 211–215
`oddsratio()`, 158
 optimal scaling, **222**, *see* correspondence
 analysis
 Ord plot, 95–99
 limitations, 99
`Ord_plot()`, 96, 97, 99, 101, 109, 112
`ordered()`, 39, 43, 326, 342, 345
 ordinal variables, **4**, 66, 125–127, 146, 175,
 211, 323, 353, 376–381, 389,
 396–399
 response, 324–333
 three-way tables, 400–403
`outer()`, 76
`outlierTest()`, 484
 overdispersion, **82**, **431**, 444–446, 449–450
 overplotting, 237

P`p.adjust()`, 135

package

AER, 449, 468, 502, 503

agridat, 19, 217

animation, 21

ca, xvii, 224, 225, 244, 249

car, xvii, 287, 292, 306–308, 311, 313,
315, 321, 322, 327, 336, 347, 360,
366, 433, 434, 482, 484, 486, 496,
501, 502

colorspace, 171

corrplot, 378

countreg, 440, 450, 453, 454, 457, 471

datasets, 41, 116

directlabels, 27, 80, 331

effects, xvii, 278–280, 291, 294, 332,
342, 345, 419, 465–467, 499

extracat, 118

FactoMineR, 225

foreign, 38

gdata, 38

GGally, 202

ggplot2, 13, 27, 80–82, 155, 202, 268,
269, 271, 275, 276, 286, 340, 367,
424, 437, 460, 461, 494, 503

ggtern, 155

ggvis, 21

gnm, xvii, 383, 384, 390

googleVis, 21

gpairs, xvii, 197, 202

graphics, 47, 455

grid, 170, 202

heplots, 491

HistData, 67, 109

hmmm, 404

iplots, 21

KernSmooth, 286

knitr, xvii, 28, 523

Lahman, 112, 217

lattice, 19, 27, 80

lme4, 279

lmtest, 266, 422, 462

logmult, xvii, 216, 255, 383, 386, 387,
426

lsmeans, 278

manipulate, 21

MASS, 27, 56, 89, 130, 178, 183, 218,
224, 225, 255, 264, 279, 298, 324,
326, 346, 354, 357, 446, 453, 501,
502

mgcv, 478

mlogit, 342

nlme, 279

nnet, 279, 324, 342

nonnest2, 435

plyr, 52

poLCA, 279

popbio, 265

pscl, 435, 453

psych, 148

rCharts, 21, 22

reshape2, 331, 340

rggobi, 21

rgl, 203, 204, 225

rms, 301, 324, 329

rsm, 297, 479

sandwich, 445

shiny, 21

Sleuth2, 501

splines, 287, 503

stats, 47, 74, 178, 353

TeachingDemos, 155

texreg, 56

UBbipl, 250, 251

vcd, xvi, xvii, 5, 39, 47, 48, 60–63, 69,
70, 78, 89, 94, 96, 98, 101, 118,
119, 123, 129, 135, 136, 138, 139,
141, 147, 149, 155, 158, 167, 168,
170–173, 186, 190, 193, 194, 197,
202, 209, 212, 229, 233–235, 270,
276, 357, 377, 418, 450, 455, 517vcdExtra, xvi, xvii, 51, 53, 56, 58,
60–62, 71, 86, 105, 108, 111, 112,
118, 125, 151, 158, 159, 183, 184,
197, 204, 216, 218, 219, 254–257,
267, 282, 288, 296, 321, 347, 357,
361, 372, 373, 392, 396, 400, 407,
415, 434, 442, 502VGAM, 324, 328, 342, 404, 406, 407,
411, 413, 453, 490, 496, 499

XLConnect, 38

xlsx, 38

xtable, 56

`pairs()`, 169, 197, 200, 207, 491

pairs plot, 201–203

`pairs.table()`, 201, 209`palette()`, 171panel functions, **201**parallel coordinate plot, **20**Pareto chart, **26**Pareto distribution, **26**

- Pareto principle, **26**
- parquet diagram, **139**, *see* sieve diagram
- partial association, **189**, 192, 352
- partial proportional odds model, **327**
- partial residual plot, **312**, 313
- partial residuals, **279**
- partial-regression plot, **314**
- Pascal distribution, **82**
- `paste()`, 43, 234, 235
- `pbinom()`, 74
- `pchisq()`, 88
- `pdoublebinom()`, 105
- Pearson residual, 145, 163, 231, **304**, **356**, **481**
- Pearson residuals, 179, 222
- `persp()`, 479
- `persp.lm()`, 479
- `pgeom()`, 74
- ϕ coefficient, 222
- `pickCoef()`, 386, 395
- pie chart, 130
- `plogseries()`, 74
- `plot()`, 48, 94, 137, 159, 227, 244, 247, 249, 262, 279, 294, 297, 298, 306, 388, 389, 418, 450, 475, 482, 484, 486
- `plot.ca()`, 225, 226
- `plot.goodfit()`, 93
- `plot.gootfit()`, 110
- `plot.rc()`, 387
- `plot.xmean.ordinaly()`, 330
- `plot3d.ca()`, 225
- plotting, *see also* graphics
 - stratified displays, 134
- plyr package, 52
- `pnbinom()`, 74
- Poisson distribution, 76–82, 97, 99, 108, 350, 354, 359, 431, 432, 438
 - calculation of, 80
 - examples, 69–72, 78–80, 91–92
 - moments of, 77
 - plotting, 80–82, 93, 94, 99–102
 - violated assumptions, 83
- Poisson regression, **350**
- Poisson samples, 69–72, 76–82
- Poissonness plot, **99**, 99–102
- polar area chart, 130
- poLCA package, 279
- `polr()`, 279, 324, 326, 328, 332, 343, 347
- `poly()`, 411, 414
- Polya distribution, **83**
- `polygon()`, 269
- polytomous, **263**, 323, 335–341
- polytomous events, 66
- polytomous response, **262**, 323–345
- polytomous variables, **4**, 66
- popbio package, 265
- population marginal means, **278**
- `position_jitter()`, 269
- power series distributions, 86–87
- `ppois()`, 74, 79
- `prcomp()`, 253
- `predict()`, 262, 269, 279, 331, 339, 479
- predictor variables, *see* explanatory variables
- principal component analysis, **221**, 240, 404
- principal coordinates, 223
- principal inertia, **223**
- `princomp()`, 253
- `print()`, 58, 89, 262, 266, 279, 297
- `print.goodfit()`, 89
- probit models, 264, 325
- profile contrasts, 211
- `prop.table()`, 44, 45
- proportional odds model, **324**, **324**, 324–333, 341
 - assumptions, 327–329
 - effect plot, 332–333
 - fitting, 326–327
 - latent variable interpretation, 325–326
 - plotting, 329–333
- proportions, 271
- pscl package, 435, 453
- psych package, 148

Q

- `qbinom()`, 74
- `qdoublebinom()`, 105
- `qgeom()`, 74
- `qlogis()`, 413
- `qlogseries()`, 74
- `qnbinom()`, 74
- `qnorm()`, 486
- `qppois()`, 74
- `qqPlot()`, 486
- quadratic model, 137
- quantile–quantile (QQ) plots, 485–489
- quantile–quantile plots, 93
- quasi-independence, **141**, 396, 401
- quasi-independence model, **389**, 393
- quasi-Poisson, **431**, 445–446
- quasi-Poisson model, **445**
- quasi-symmetry, **389**, 393, 401

R

radial diagram, **130**
 rainbow(), 479
 rainbow_hcl(), 171
 random zeros, *see* sampling zeros
 raw residual, **356, 481**
 rbinom(), 74
 rc(), xvii, 383, 387, 427
 rCharts package, 21, 22
 rcL(), 387
 rdoublebinom(), 105
 read.csv(), 38
 read.delim(), 38
 read.table(), 38, 58, 59
 reciprocal averaging, **222**, *see*
 correspondence analysis
 recode(), 336
 reference category, 365
 regression, 262, 350, 352, 430, 489
 influence, 303
 leverage, 304
 linear, 263–264
 logistic, 263–270, 278, 341–345
 multiple logistic, 272
 multivariate, 349
 probit, 264
 regression spline, **287**
 releval(), 273, 342
 reliability, 146
 rendering, **18**
 rep(), 33
 reshape2 package, 331, 340
 residual deviance, **433**
 residualPlot(), 482
 residuals, 481
 deviance, 304, 445, 481
 loglinear models, 356–357
 partial, 279–281
 Pearson, 145, 163, 179, 222, 231, 304,
 481
 response, 481
 shading, 171–175
 standardized, 304, 481
 studentized, 304, 481, 484
 residuals(), 304, 357
 response, **116**
 response residual, **481**
 response variables, **7**, 120–121, 190
 discrete, 263
 rgeom(), 74
 rggobi package, 21

rgl package, 203, 204, 225
 rlogseries(), 74
 rms package, 301, 324, 329
 rnbinom(), 74
 rnegbin(), 453, 487
 rnorm(), 37, 487
 rootogram, **93**, 104, 107, 108, 436, 450
 rootogram(), 94, 109, 112, 450
 row effects model, **377**
 row plus column effects model, **377**
 row-and-column effects model, **382**
 rownames(), 57
 rpois(), 74, 256, 453, 487
 rsm package, 297, 479
 rstandard(), 304, 482
 rstudent(), 304, 482
 rzinegbin(), 453
 rzipois(), 453

S

s(), 480
 sample(), 37
 sample odds ratio, **122**
 sampling zeros, **368**
 sandwich package, 445
 sandwich(), 445, 448
 saturated model, 140, 178, **178, 180, 267**,
 351, **351**, 353, 404
 scale parameter, **431**
 scale_y_log10(), 437
 scaling, 325
 scatterplot, 17, 201, 437
 scatterplot matrix, 197, **197**, 240
 scatterplotMatrix(), 311
 score model matrix, 279
 scree plot, **225**, 239
 segments(), 230
 seq(), 33
 seq_loglm(), 186
 sequential models, 434
 set.seed(), 37, 173
 shading_binary(), 169
 shading_diagonal(), 171
 shading_Friendly(), 169
 shading_Friendly2(), 173
 shading_hcl(), 169
 shading_hsv(), 169
 shading_Marimekko(), 170
 shading_max(), 169, 173, 175
 shading_sieve(), 169
 shiny package, 21

- `sieve()`, 168, 357
 - sieve diagram, 116, 126, **139**, 138–145, 201, 222
 - simple effects, **189**
 - Simpson's paradox, 135, **135**
 - `simulate()`, 487
 - singular value decomposition, **222**, 223, 248–249
 - Sleuth2** package, 501
 - `sort()`, 43
 - `spacing_conditional()`, 169
 - `spacing_dimequal()`, 169
 - `spacing_equal()`, 169
 - `spacing_highlighting()`, 169
 - `spacing_increase()`, 169
 - spaghetti plot, *see* parallel coordinates plot
 - spineplot, **118**, **283**, **454**
 - `spineplot()`, 159, 455, 470
 - spinogram, 454, **454**
 - spline, 93
 - splines** package, 287, 503
 - square tables, 375, 389–399
 - models for, 390
 - ordinal variables, 396–399
 - three-way, 400–403
 - stacking, **232**
 - standard coordinates, 223
 - standardized residuals, **304**, **481**
 - `stat_smooth()`, 269, 275, 285
 - statistical analyses
 - Cramer's V, 9, 128
 - deviance tests, 179
 - exact vs. asymptotic tests, 124
 - goodness-of-fit, 87
 - likelihood ratio test, 88, 183
 - logistic regression, 10–13
 - Wald test, 335
 - Woolf's test, 129
 - stats** package, 47, 74, 178, 353
 - `stepAIC()`, 298, 347, 502
 - `str()`, 34, 43
 - stratified analysis, **8**, **128**, 127–129, 134, 141, 275, 415
 - stratifying variable, **115**
 - stripplot, 201
 - `struc_assoc()`, 169
 - `struc_mosaic()`, 169
 - `struc_sieve()`, 169
 - `strucplot()`, 168
 - strucplot framework, **141**, 141–146, **162**, **167**, 167–175
 - labeling, 179
 - `structable()`, 42, 46, 47, 52, 61, 63, 167, 233, 234, 372
 - structable framework, 190
 - structable objects, 47–48, 392
 - subsetting, 49–50
 - structural zeros, **214**, **368**, 452
 - studentized residual, **304**, **481**
 - `subset()`, 50, 62
 - subsetting, 48–50, 60
 - successive differences, 211
 - `sum()`, 52
 - `summarise()`, 52
 - `summary()`, 89, 123, 244, 262, 266, 279, 287, 297, 298, 306, 321, 327, 343, 360, 365, 438, 445, 501
 - `summary.goodfit()`, 89
 - `summary.Kappa()`, 149
 - supplementary variables, **238**
 - `Symm()`, 390, 401
 - symmetric map, **223**, **224**
 - symmetry, 140, **141**
 - symmetry model, **389**
- T**
- `t()`, 35, 48, 57
 - `table()`, 36, 42, 44, 46, 48, 55, 436
 - table form, **6**, 31, 41–47, 50, 52–58, 60, 115, 119, 121, 178, 354, 377
 - collapsing, 51
 - converting, 53
 - converting to, 44–47
 - decomposition, 184
 - publishing, 56–58
 - reordering, 43–44
 - subsetting, 48–49, 60
 - TeachingDemos** package, 155
 - ternary plot, **116**
 - test of independence, 130
 - tests for linear trend, 127
 - tests of association, 119–121, 124–125, 127
 - texreg** package, 56
 - `theme()`, 82
 - three-way table, 119, 130, 141, 146, 147, 232–238, 352, 354, 363, 400–403
 - fitting, 179
 - plotting, 176
 - thresholds, 325
 - `tile()`, 118, 159, 168
 - tile plot, **118**
 - `toeplitz()`, 170

`Topo()`, 390, 394, 395
 topological model, **394**
 treatment contracts, 273
 treemap, **22**
 trends, 376
 trilinear plot, **116**, **153**, 153–156
`tripplot()`, 155
 two-way table, 116–118, 121–130, 138–139,
 145, 162–166, 178, 221–231, 351,
 376, 389
 tests of association, 119–121, 124–125,
 127
 type-token, **72**, 72–73

U

UBbipl package, 250, 251
 uniform association, 400
 uniform association model, 377, **377**
 uniform interaction, **400**
`update()`, 297, 300, 345, 379, 390, 476

V

variable, 273
 binary, 4, 121, 209, 241, 262
 dichotomous, 4, 66, 121, 405–406
 discrete, 18
 effect ordering, 18–20
 explanatory, 7–8, 120–121
 nominal, 5, 66, 335–341
 number of, 7
 ordered, 39, 43–44
 ordinal, 4, 66, 125–333, 353, 376–381
 polytomous, 4, 66
 response, 7–8, 120–121
 stratifying, 127–129
 variable
 nominal, 323
 vcd package, xvi, xvii, 5, 39, 47, 48, 60–63,
 69, 70, 78, 89, 94, 96, 98, 101,
 118, 119, 123, 129, 135, 136, 138,
 139, 141, 147, 149, 155, 158, 167,
 168, 170–173, 186, 190, 193, 194,
 197, 202, 209, 212, 229, 233–235,
 270, 276, 357, 377, 418, 450, 455,
 517
 vcdExtra package, xvi, xvii, 51, 53, 56, 58,
 60–62, 71, 86, 105, 108, 111, 112,
 118, 125, 151, 158, 159, 183, 184,
 197, 204, 216, 218, 219, 254–257,
 267, 282, 288, 296, 321, 347, 357,

 361, 372, 373, 392, 396, 400, 407,
 415, 434, 442, 502
`vcov()`, 279, 448
 vector, **32**, 32–33, 222
 binding, 35
 transposing, 35
 VGAM package, 324, 328, 342, 404, 406,
 407, 411, 413, 453, 490, 496, 499
`vglm()`, 324, 328, 329, 404, 406, 411, 412,
 415, 420, 421, 496, 497
 visual impact, **135**
 visualization, 23
`vuong()`, 435
 Vuong's test, **434**
`vuongtest()`, 435

W

Wald test, 433, 444
`weighted.mean()`, 55, 88
 weights, 126, 149, 365
`with()`, 44
`within()`, 59, 79
 Woolf's test, 129
`woolf_test()`, 129, 130, 159

X

XLConnect package, 38
 xlsx package, 38
 xtable package, 56
`xtable()`, 56, 57, 63
`xtabs()`, 36, 42, 44, 46, 48, 52, 54, 59, 62,
 63, 124, 234, 235, 238, 283, 493
`xyplot()`, 80, 81, 84, 85, 110

Z

zero frequencies, 368–371
 sampling zeros, 368
 structural zeros, 368
 zero-altered model, **454**, 454
 zero-frequencies
 plotting, 454–455
 zero-inflated Poisson, **452**
 plotting, 453
 zero-sum constraints, 353
 zero-truncated distribution, **112**
`zeroinfl()`, 453, 454
 zeros
 excess, 451–455
 ZIP, *see* zero-inflated Poisson