

Subject index

Symbols

β weights 278
 η^2 226
 ϕ 133
* comment 85
/* and */ comment 85

A

acquiring datasets 525–526
add, label define option 2
agreement, intraclass correlation ... 260
alpha reliability 377
ameans command 97
analysis of covariance see ANCOVA
analysis of variance see ANOVA
ANCOVA 237–248
ANOVA,
degrees of freedom 225
equal-variance test 225
one-way 219–220
one-way power analysis .. 232–234,
263–265
power analysis .. 232–234, 263–270
repeated-measures 255
repeated-measures power
analysis 267–269
two-way 249–255
two-way power analysis ... 265–267
ANOVA assumptions 220
ATS at UCLA 203

B

bar chart 108, 350
bar graph of means 231
Bartlett test of equal variances 225
beta weights 210, 278
limitation 313

binary variables 298
binscatter command 197–201, 317
block regression 296–305
blog see Stata Blog
Bonferroni multiple-comparison test ...
..... 207, 223
bookstore, Stata 522
bootstrap estimation of standard errors
..... 288
bootstrap regression 283
Boston College Stata site 520
boundary characteristic curve .. 505–506
box plot 118, 237

C

case, changing 48
casewise deletion 204
categorical covariates 239, 243
categorical predictors, regression ... 298
categorical variable 305–306
cause and effect 195
center command 310
centering 321–322
chi-squared
table 129
test 127, 353
chitable command 129–131
clear command 29
clonevar command 458
cls command 16
codebook command 44–45, 55–56,
138
codebook example 25–28
coding system 24–28
coefficient of variation 117
Cohen's d 170–171, 175–182
Cohen's f 232–234, 263

collinearity 292
 command structure 78
 Command window 8, 10
 compatibility 42
 confidence interval
 regression line 210
 slope 210
 constant 210
 continuous covariates 239–241
 continuous variable 317
 conventions used in the book 1
 copying, HTML format 19, 89
 copying results to word processor ... 19,
 88
correlate command 204–205
 correlation,
 interpreting 201
 limitation 313
 multiple comparison 207
 correlation ratio 226, 247–248
 count outcome variables 341
 Cramér’s *V*, measure of association ...
 133
 creating value labels 57–60
 criterion-related validity 384
 cross-tabulation 124
 curve fitting 315–321

D

data,
 long format 162, 221
 wide format 160, 221
 Data Editor 29–33, 40–41
 dataset contents 43
 dataset,
 acquisition 525–526
 c10interaction 307, 330
 c11barchart 350
 cancer 10, 11
 census 329
 censusfv 20
 chapter6_aspirin 136
 chapter13_missing 407
 chores 173

dataset, *continued*

 create 21–23
 depression 216
 descriptive_gss 100, 120, 121
 divorce 335
 download xxxv, 525–526
 environ 338
 firstsurvey 43
 firstsurvey_chapter4 80, 85,
 92
 flourishing_bmi 422, 423
 gss2002_and_2006_chapter12
 400
 gss2002_chapter6 148, 149
 gss2002_chapter7 157, 163, 164,
 166, 173, 187
 gss2002_chapter8 216
 gss2002_chapter9 271
 gss2002_chapter10 329, 330,
 448
 gss2002_chapter11 367
 gss2006_chapter6 124, 137, 142,
 145, 148
 gss2006_chapter6_10percent
 145
 gss2006_chapter8 190, 215, 216
 gss2006_chapter8_selected 207
 gss2006_chapter9 229, 238
 gss2006_chapter9_2way 249
 gss2006_chapter12 370, 377
 gss2006_chapter12_selected
 391
 intraclass 261
 kappa1 381
 kuder-richardson 379
 long 162
 longitudinal_exercise 479
 longitudinal_mixed 456
 nlsw88 331
 nlsy97_chapter7 182, 186
 nlsy97_chapter11 341, 353
 nlsy97_selected_variables
 271, 296
 ops2004 274, 420
 partyid 222, 235, 271

dataset, *continued*

positive..... 75
 regsmpl..... 315, 407, 449
 relate.....xxxv, 52, 75
 relate_small xxxv
 retest..... 375
 severity..... 366
 spearman..... 212, 216
 wide..... 159
 wide9..... 256
 degrees of freedom, 129
 ANOVA..... 225
 one-sample *t* test..... 166
 dependent *t* test..... 173
 dependent variable 126
 describe command..... 10, 46–47, 53
 dfbeta command..... 291–292
 dialog box,
 alpha..... 377
 anova..... 249
 codebook..... 44–45
 correlate..... 204
 describe..... 46–47
 egen..... 70–72
 generate..... 65, 67–68
 graph bar..... 143–144, 231
 graph box..... 237
 graph hbox..... 118
 graph pie..... 104
 histogram..... 14–15, 108–110
 kwallis..... 235
 lfit..... 211
 logistic..... 342–343
 margins..... 242, 244–245, 251
 nestreg..... 323
 oneway..... 222–223
 open..... 66
 pcorr..... 279–280
 power..... 178–179
 prtest..... 157–163
 recode..... 61–62, 297
 regress..... 208, 275
 rename..... 57
 rvplot..... 286
 scatter..... 191–196

dialog box, *continued*

sdtest..... 171
 sktest..... 113
 Submit vs. OK..... 17
 summarize..... 11
 tab1..... 100
 tabi..... 140–141
 table..... 142
 tabstat..... 117
 tabulate..... 68–69, 124–128, 138
 ttest..... 164–165, 168–170, 173
 dictionary file..... 52
 difference of means test..... 166
 difference of proportions test..... 159
 discontinuity graph 199–201
 display command..... 170–171, 348,
 439–440
 do-file,
 continuation line..... 167
 introduction..... 6
 Do-file Editor 83–87
 download datasets 525–526
 drop command..... 73
 dummy variables 298

E

effect size..... 170–171, 226, 247, 279
 η^2 226, 247–248
 egen command..... 65, 70–72, 371
 egen count command..... 371
 egen rowmean command..... 72, 371
 egen rowmiss command..... 71
 egenmore command..... 71
 entering data..... 29–33
 equal variance, Bartlett test..... 225
 esize command..... 170–171
 estat esize command..... 247–248
 estat report command..... 490–491
 estat vif postestimation command ..
 293–294
 estimates store command..... 353
 Excel
 exporting data..... 47–48
 importing data 47–48
 exit Stata..... 18, 43

exponentiation 348
 external validity 204

F

F ratio 225
F test of unequal variances 172
 Facebook see Stata on Facebook
 factor analysis, 386
 commonality 389
 eigenvalue 389, 394
 exploratory factor analysis 387
 extraction 389
 factor score 389, 398
 loading 389
 oblique rotation 389, 396
 orthogonal rotation 389, 395
 PCF 388
 PF 387
 postestimation 390
 principal component analysis 388
 principal-component factor analysis 388
 promax 396
 rotation 389
 scree plot 389, 394
 simple structure 389
 varimax 395
 factor variable 305–306
 fixed effects 453–454
 fonts, fixed 88
 format,
 numeric 31
 string 32
fre command 102–104
 frequency distributions 101
ftable command 225
 full mediation 443

G

gamma, measure of association 138
 generalized linear model 435–436
 generalized structural equation modeling 421–449
generate command 65–70
 geometric mean 97

glm command 435–436
 Goodman and Kruskal's gamma, measure of association 138
 graded response IRT model 502–509
 GradPlan 522
 graph,
 alternative scattergram 197–201
 bar chart 108, 144, 350
 box plot 118
 collinearity 292
 discontinuity 199–201
 hanging rootogram 281
 heteroskedasticity 285
 histogram 111, 281
 medians 237
 overlay two-way showing
 interaction effects 308
 pie chart 104–108
 residual versus fitted 286
 scattergram 190–197
graph bar command 143–144, 231, 350
graph box command 237
 Graph Editor 106–108
 graphics book 523
 growth curve 456
gsem command 421–449, 453
 logistic regression 433–436
 standardized solution 433
 GUI interface 8

H

harmonic mean 97
 help 6
 video 20
 web-based 20
help, listcoef option 348–350
help label command 43
 heteroskedasticity 285
 hierarchical linear models 451–480
 hierarchical regression 296–305, 323–325, 360–362
 histogram 111
histogram command 13–18, 281

HLM see hierarchical linear models
HTML format 89

I

ICC see item characteristic curve
imputation see multiple imputation
increment in R^2 279
independent variable 126
indicator variables 298
interaction term 306–309
interactive table 140
intercept 210
intercept: growth trajectory 456
interquartile range 117
interval-level variables 96
intraclass correlation 260, 262
IRT see item response theory
`irt 1pl` command 490–491
`irt 2pl` command 497–498
`irt grm` command 504–505
`irtgraph icc` command 491, 499,
505–506
`irtgraph iif` command 493, 500,
507
`irtgraph tif` command 494, 501,
508
item characteristic curve 485, 491,
499
item information function graph 493,
500, 506–507
item response theory 481–516

J

`jitter()`, `scatter` option 194

K

kappa 380
kappa, weighted 382
kappa with three raters 382
`keep` command 73
Kendall's tau, measure of association . .
. 138
Kruskal–Wallis test, ANOVA alternative
. 235

Kuder–Richardson coefficient of
reliability 379
kurtosis 99, 113, 283

L

`label variable` command 59
labeling values 33
labeling variables 23
likelihood-ratio chi-squared test . . . 353–
354
limitations of Stata 526
linear regression 421–423
`list` command 81–82, 222
`list` option, `nolabel` 222
`listcoef` command 348–350
listwise deletion 204
`log`, `.smc1` extension 89
`log` files 89
`log` files and graphs 90
`logistic` command 342–347
logistic regression 333–367, 441
bar chart 350
exponentiation 348
hypothesis tests 353
interpreting odds ratio 347
likelihood-ratio chi-squared test . . .
. 353
logits 340
McFadden pseudo- R^2 346
nested 360–362
nonlinear 336
odds ratio 338
percentage change 347
pseudo- R^2 346
 S -curve 336
vs. OLS regression 337
Wald chi-squared test 353
`logit` command 336, 342–347,
434–436
logits 340
long format 162, 169, 221, 256–258,
458
longitudinal data 452–453
`lrdrop1` command 353–354
`lrtest` command 353

M

MAR.....403–405
margins command..242, 244–246, 311, 357–359
marginsplot command.....252–255, 311–312
 maximum number of variables 526
 MCAR.....403–404
 McFadden pseudo- R^2 346
 mean squares..... 225
 measure of association,
 η^2 226, 247–248
 ϕ 133
 odds ratio..... 135
 V 133
median command..... 184
 median, graph box plot 237
 mediation 442–446
 correlated residuals 447
 full 443
 partial 443
 menu, open..... 66
mi estimate command 413–414
mi impute chained command.... 407
mi impute command.....411–412
mi impute mvn command....407, 412
mi register command 411
mi set command 411–412
mibeta command 414–416
 missing, count..... 371
 missing values.....27, 56, 67, 79, 401–420, 456
 types..... 55, 56
misstable command.....408–409
mixed command 453–474
 mixed models 451–480
more command..... 4, 44
 multicollinearity 292
 multilevel analysis 451–480
 multiple comparison,
 Bonferroni 223
 Scheffé..... 223
 Šidák 223
 multiple comparison and correlation...
 207

multiple correlation.....276
 multiple imputation..... 401–420
 multiple regression command..275–279
 multiple regression diagnostics..... 288
 multiple regression with interaction
 term..... 307–312
 multiple regression,
 block..... 296–305
 categorical predictors 298
 dummy variables 298
 hierarchical 296–305
 indicator variables 298
 influential case..... 289
 nested..... 296–305
 outlier 289
 residual 285
 weighted data 294
mvdecode command 56, 67, 167

N

naming variables.....25
 nested regression 296–305, 323–325, 360–362
nestreg command .. 302–305, 323–325, 360–362
 NetCourses.....525
no label, list option.....222
 nominal-level variables..... 96
 nonlinear 336
 nonlinear regression 313–325
 nonparametric ANOVA alternatives
 234
 nonparametric tests 182
 Mann–Whitney 182
 median 183
 rank sum 182
 normally distributed residuals 284
numlabel command..... 82, 104

O
 odds ratio 135, 338
 interpretation 347
 percentage change.....347
 OLS regression vs. logistic regression...
 337

- omega2 command..... 247
 - one-sample *t* test..... 164
 - one-parameter logistic IRT model.. 484–496
 - one-way ANOVA..... 219–220
 - oneway option
 - bonferroni..... 223
 - scheffe..... 223
 - sidak..... 223
 - open
 - existing dataset..... 9–11
 - Stata-installed dataset..... 10
 - optifact command..... 520
 - option,
 - add, label define..... 2
 - help, listcoef..... 348–350
 - jitter(), scatter..... 194
 - percent, listcoef..... 350
 - ordinal-level variables..... 96
 - outlier..... 289
- P**
- paired *t* test..... 173
 - panel data..... 452–453
 - part correlation..... 279
 - partial mediation..... 443
 - pasting,..... 89
 - reformatting..... 88
 - pasting results to word processor,.. 88
 - formatting..... 88
 - path analysis..... 442–446
 - pcorr command..... 279, 300
 - Pearson’s chi-squared..... 138
 - percent, listcoef option..... 350
 - pie chart..... 104–108
 - plot a confidence interval..... 211
 - Poisson regression..... 341
 - postestimation..... 526–527
 - postestimation command,
 - estat vif..... 293–294
 - margins..... 311
 - marginsplot..... 311–312
 - predict..... 289–291, 307
 - test..... 301–302
 - Postestimation Selector..... 526–527
 - power analysis..... 174–182
 - one-way ANOVA..... 232–234, 263–265
 - repeated-measures ANOVA ... 267–269
 - two means..... 179–182
 - two-way ANOVA..... 265–267
 - power command..... 178–182
 - power oneway command..... 232–234, 263–265
 - power repeated command.... 267–269
 - power twomeans command.... 181–182
 - power twoway command..... 265–267
 - powerreg command..... 325–328
 - predict command..... 473
 - predict postestimation command.....
 - 289–291, 307
 - predict postregression..... 289–291
 - predictive validity..... 384
 - probability tables..... 129
 - product term..... 309
 - project outline..... 52
 - prophecy formula..... 376
 - proportions,
 - one-sample test..... 157
 - two-sample test..... 159
 - prtest command..... 157–163
 - pseudo- R^2 346
 - pwcorr command..... 205–207
 - pweights..... 295–296
 - pwmean command..... 227–229
- Q**
- quadratic model,
 - centering..... 321–322
 - examining..... 323–325
 - fitting..... 315–321
 - qualifier,
 - if with missing values..... 79
 - in..... 80
- R**
- R^2 170–171, 276
 - change..... 279
 - random coefficients.. 454–456, 471–474

- random effects 454–456
 - random intercept 454–456, 461–470
 - random sample 191
 - how to draw 155
 - random sampling 153
 - random slope 454–456
 - randomization 153
 - alternative to 238
 - how to perform 155
 - ranksum** command 182
 - Rasch modeling 482
 - recode** command 61–63, 297
 - recoding 167
 - ranges 61
 - regress** command .. 208–211, 275–276, 299–301, 307–309
 - regression,
 - block 296–305
 - bootstrap 283
 - categorical predictors 298
 - dummy variables 298
 - hierarchical 296–305
 - indicator variables 298
 - influential case 289
 - nested 296–305
 - outlier 289
 - residual 285
 - robust 283
 - weighted data 294
 - regression diagnostics 288
 - regression line, plotting 196
 - regression with interaction term .. 307–312
 - reliability,
 - alpha 377
 - equivalent forms 376
 - IRT 509–512
 - kappa 380
 - kappa with three raters 382
 - Kuder–Richardson coefficient of reliability 379
 - prophecy formula 376
 - split-half 376
 - test–retest 375
 - weighted kappa 382
 - rename** command 57, 256
 - rename variable 256
 - repeated-measures ANOVA 255
 - repeated-measures design 267–269
 - repeated-measures *t* test 173
 - replace** command 65, 167
 - reshape** command 256–258, 458
 - reshape wide to long format ... 256–258
 - residual 284, 285
 - Results window 8
 - clear 16
 - more** 4
 - scroll size 4
 - reverse coding 371
 - reverse-code variables 60
 - Review window 8
 - robust estimation of standard errors...
 - 288
 - robust regression 283
 - root mean squared error 210
- S**
- sample, draw 191
 - sample** command 145, 191
 - SAS XPORT 48
 - save** command 41–42
 - saveold** command 42
 - scale construction 370
 - reverse coding 371
 - scale creation 70
 - scatter** command 191–198
 - scattergram 190–201
 - scattergram with confidence interval ..
 - 211
 - scattergram with **jitter()** option .. 194
 - scattergram, alternative 197–201
 - Scheffé, multiple comparison 223
 - schemes 15
 - scientific notation 278
 - scree plot 393–394
 - S*-curve 336
 - sdtest** command 171–172
 - search** command 6, 520
 - seed for starting a random sample .. 191
 - sem** advantages 421

- SEM Builder 423–430
 restore default settings 431
sem command 421–449, 453
sembuilder command 423
 semipartial correlation 279, 300
set more off command 4
set seed command 145, 154, 191
 Šidák, multiple comparison 223
 significance,
 statistical 203
 substantive 203
 skewness 99, 113, 283
sktest command 283
 slope 210
 SMCL... see Stata Markup and Control
 Language
spearman command 212
 Spearman's rho 212
 split-half reliability 376
 SPSS 48
ssc command 520
 standardized beta coefficient 210
 standardized beta weights 278
 standardized regression coefficients...
 278
 Stat/Transfer 48, 526
 Stata, limitations 526
 Stata Blog 521
 Stata bookstore 522
 Stata code for textbook examples.. 520
 Stata/IC limitations 526
Stata Journal 522
 Stata Markup and Control Language..
 89
 Stata NetCourses 525
 Stata on Facebook 521
 Stata on Twitter 521
 Stata Portal, UCLA 203, 520, 523
 Stata screen 7
 Stata tutorial 18, 525
 Statalist Forum 521
 statistical significance versus
 substantive significance... 203
 structural equation modeling.. 421–449
 sum of squares 225
summarize command.. 11–13, 113, 208,
 282
summarize(), **tabulate** option... 249–
 255
 summary of data 27
sunflower command 193
sysuse command 10
- T**
- t* test,
 one-sample 164
 power analysis 178–182
 two-sample 166
tab command 145
tab1 command 101, 204
tab2 command 298
tabdisp command 258
tabi command 140–141
 table 124
 probability 129
 summary statistics 235
 table calculator 140–141
table command 142–143
tabstat command 117, 235
tabulate command 64,
 68–70, 101, 124–128, 136, 138–
 140, 249–255
 tau, measure of association 138
 test information function graph... 494,
 501, 508
 test of significance,
 kurtosis 113
 skewness 113
test postestimation command... 301–
 302
 test–retest reliability 375
 tests,
 Bonferroni 207
 chi-squared 127
 dependent *t* test 173
 likelihood-ratio chi-squared... 353–
 354
 likelihood-ratio chi-squared with
 logistic regression 353
 logistic regression 353

tests, *continued*

- long format for portions.....162
- Mann–Whitney182
- median183
- multiple comparison with correlations207
- nonparametric182
- one-sample *t* test.....164
- paired *t* test173
- proportions157, 159
- rank sum182
- repeated-measures *t* test173
- skewness and kurtosis283
- two-sample *t* test.....166
- unequal variances171
- Wald chi-squared test353
- wide format for proportions...160
- z* test for proportions163
- z* test with logistic regression..353
- textbook examples using Stata
 - commands.....520
- three-parameter logistic IRT model....
 -487
- time-invariant covariate.....456
- time-varying covariate456
- tolerance.....293–294
- toolbar,
 - Stata for Mac9
 - Stata for Windows9
- `ttest` command.....164–170, 173
- tutorial.....525
- tutorials18
- Twitter see Stata on Twitter
- two-by-two table.....124
- two-parameter logistic IRT model..486–487, 496–501
- two-way ANOVA.....249–255
- `twoway` command196–197, 308
- `twoway lfit` command.....473

U

- UCLA Stata Portal20, 203, 520
 - reshaping data wide to long...256
- user-written commands520

V

- validity,
 - criterion related.....384
 - external.....204
 - predictive384
- value labels23, 33–40, 57–60, 82
- variable labels23
- variable name.....22, 25
 - reserved.....22
- Variables Manager...34–40, 57–59, 73
- Variables window8
- variance inflation factor293–294
- versions.....42
- video tutorials18
- VIF see variance inflation factor

W

- Wald chi-squared test.....353
- web search520
- weighted data.....294–296
- weighted kappa.....382
- weights, `pweights`.....295–296
- wide format160, 170, 221, 256–258, 457
- working directory9

X

- `xtreg` command.....262
- `xtset` command.....258

Z

- z* test,
 - one-sample proportion157
 - two-sample proportion159
- zero-inflated models341