

Subject index

A

abbreviations 42
adjust 192, 254
ado-files 270–276
ado-path 5
adoupdate 23
analysis of variance 115
aorder 89
append 89–90
archiving data 302–303
arithmetic operators 75–76
ASCII data 62–63
audit trail 287–288
axes (graphs) 134–139
axis labels (graphs) 134–137
axis ticks (graphs) 134–137

B

backing up data 301–302
baplot 245
bar graphs 152–156
binomial regression 173–174
binreg 173–174
bkup 301
Bland–Altman plot 245–246
box-and-whisker plot 150–151
browse 102
by-graphs 167–168
by: (prefix) 40

C

c() (returned settings and constants) .. 266
capture 281
case–control data 175–184
 matched 182–184, 203–204
categorical predictors 192–195
cc 175–177

cci 122
cd 58–59
chi-squared test 104–105
ci 118–119
cii 119–120
classification of diseases 260
clogit 203–204
clustered observations 204–207
cmdlog 20
codebook 288–290
codebook (command) 98–99
cohort data without censorings ... 171–174
collapse 93–94
colors (graphs) 145
combining files 89–92
combining graphs 168–169
command log 20
Command window 8
comments in do-files and ado-files 41
comparing measurements 243–247
competing events (Cox regression) .. 236–
 237
complex strings 53–54
compress 54–55
confidence intervals 118–120
confidentiality 304–305
continue 281
contract 92–93
copying graphs to documents 170
copying tables to documents 19
correcting errors 296
correlate 241
correlation coefficient 240–241
Cox regression 218–223
creturn list 266
cs 171–174
csi 122

cut() (egen function) 79–80

D

Data Browser 11
 Data Editor 11
 data entry 60–61, 292–294
 data protection 304–305
 dataset label 67
 date formats 48
 date variables 48–50
 db 12
 debugging programs 275–276
 decimal periods and commas 19, 62
 delayed entry (Cox regression) 220–223
 #delimit 41
 describe 98
 destring 52
 diagnostic tests 247–255
 combining information 252–255
 diagnostics, linear regression 187–189
 diagt 249–250
 dialogs 12
 direct standardization 232–235
 disease classification 260
 display 122
 do 13
 do-file 12–13
 Do-file Editor 11
 dotplot (graph command) 151–152
 double quotes 40–41
 drop 87–88
 droplines (graphs) 164–165
 dstdize 233–235
 dummy variables 192–195

E

e() (returned results) 264–265
 effect modification 195–197
 egen 77–80
 egenmore (egen functions) 80
 encode 52
 encryption 305
 entering data 60–61, 292–294
 EpiData (program for entering data) 61
 epitab (family of commands) 171

ereturn list 264–265
 error correction 296
 error finding 295–296
 error messages 25–27
 estat gof 201
 estat phtest 219
 Excel (Microsoft) 64–65
 exit 6
 expand 93
 explicit subscripting 82–85
 extended functions 77–80

F

FAQs 24
 fdasave 65
 fdause 65
 file path 58–59
 file types 29–30
 findit 23–24
 Fisher’s exact test 104–105
 folder structure (Windows) 311–313
 foreach 277–278
 format 43–44
 formats
 date 48
 numeric 43–44
 string 51
 forvalues 278–281
 function plots 165–166
 functions
 extended 77–80
 mathematical 76–77
 statistical 77

G

generalized linear models 202
 generate 73–75
 gladder 110–112
 gold standard 247
 goodness-of-fit test (logistic regression) 201
 graph area 124
 graph bar (graph command) 152–156
 graph box (graph command) 150–151

graph combine (graph command) .. 168–169
 graph command syntax 125–128
 graph display (graph command) ... 170
 graph export (graph command) ... 170
 graph matrix (graph command) ... 166
 graph options 126, 134–144
 graph save (graph command) 169
 graph schemes 131–134
 graph size 128–131
 graph use (graph command) 169
 graphs 123–170
 grid lines (graphs) 134–136
 group() (egen function) 79
 groups 107–108
 gsort 89

H

help 22
 histogram (graph command) ... 148–150

I

icd9 260
 if (command) 277
 if (qualifier) 36–38
 immediate commands 120–122
 in (qualifier) 38
 incidence rates 209–211
 incidence-rate ratio ... 211–213, 225–226
 indicator variables 192–195
 indirect standardization 227–232
 infile 63
 infix 63
 input 61
 insheet 62–63
 installation 1
 interaction 195–197
 interrater agreement 239–240
 intraclass correlation coefficient 243
 ir 212
 iri 121

K

kap 239–240
 Kaplan–Meier survival function .. 215–217

kappa statistic 239–240
 keep 87–88
 kernel density curve 149–150

L

label data 67
 label define 67–68
 label list 69
 label values 67–68
 label variable 67
 labels 67–71
 ladder 110–112
 legend (graph option) 140–142
 level() (option) 118
 life table survival analysis 213
 lincom 191, 199–202
 line patterns (graphs) 145–146
 line plots 158–162
 linear regression 185–197
 diagnostics 187–189
 list 100–102
 log 16–17
 log book 290–292
 log file 16–18
 logical expressions 74–75
 logical operators 37
 logistic (command) 198–202
 logistic regression 198–202
 logit 198–199
 log-rank test 217
 log-scaled axes (graphs) 137–138
 loneway 243
 long command lines 41–42
 lroc 253–254
 ltable 213

M

macros 268–270
 manuals 21, 307–308
 marker symbols (graphs) 146–147
 master do-file 292
 matched case–control data 182–184, 203–204
 matching datasets 90–92
 mathematical functions 76–77

- matrix graphs 166
- mcc 182–184
- mean 112
- measurement comparison 243–247
- measurement, reproducibility 239–243
- measurement variation 240–243
- memory considerations 54–55
- merge 90–92
- merging datasets 90–92
- meta-analysis 260
- mhodds 179–181
- missing (option) 40
- missing values 45–46, 297
- mkdir 58–59
- more 9
- mortality rates 209–211
- multinomial logistic regression 202
- multiple axes (graphs) 138–139
- multiple-failure data (Cox regression) 236
- mvdecode 45
- mvencode 45

- N**
- _N 83
- _n 83
- NetCourses 24
- no label (option) 40
- nonparametric tests 117–118
- normal curve 149–150
- notes 69
- NoteTab Light 18
- numeric
 - formats 43–44
 - lists 35–36
 - ranges 36
 - variables 43–50
- numlabel 68

- O**
- ODBC 65
- odbc (command) 65
- oneway 115
- online help 21–24

- operators
 - arithmetic 75–76
 - logical 37
 - relational 37
- options 39–40
 - graph 126
- order 88–89
- ordered logit estimation 202
- outfile 64
- output log 16–18
- outsheet 64

- P**
- panel data 206
- parallel lists 279–281
- pharmacokinetic data 260
- pk (family of commands) 260
- plot area 124
- poisson (command) 225–226
- Poisson regression 225–226
- polytomous logistic regression 202
- postestimation commands 263
- power estimation 259–260
- precision 46–48
- predict
 - after linear regression 187–188
 - after logistic regression 200–201
- predictive values 248–250
- preferences, window 7–11
- preserve 87
- profile.do 17–18, 281–282
- programs 270–276
- proportion 103–104
- proportional hazards regression 218–220
- protecting data 304–305

- Q**
- qnorm 109
- Q–Q plot 109
- qualifiers 36–38
- questionnaires 288
- quotes
 - double 40–41
 - single 268–269

R

r () (returned results) 263–264
 random-effects models 207
 random numbers 257
 random sampling 88, 257–258
 range plots 162–163
 ranksum 118
 receiver operating characteristic 250–252
 recode 80–82
 reference lines (graphs) 163
 regress 185–197
 regression analysis 185–207
 regression diagnostics 187–189
 relational operators 37
 rename 88
 reordering variables 88–89
 replace 73–75
 reproducibility of measurements 239–243
 reshape 94–95
 reshaping data 92–95
 residual-versus-fitted plot 188–189
 restore 87
 Results window 8–9
 return list 263–264
 Review window 9
 robust variance estimates 204–207
 ROC analysis 250–252
 roccomp 252
 roctab 250–251
 rvfplot 188–189

S

sample 88, 257
 sample size estimation 259–260
 sampsi 259–260
 SAS 65
 save 57–60
 saved results 263–268
 saving graphs 169
 scatterplots 156–158
 schemes (graphs) 131–134
 sdpair 246–247
 sdtest 116–117
 search 22–23

selecting observations 88
 selecting variables 87
 sensitivity 248–250
 set trace 276
 shell 282
 simulations 258–259
 single quotes 268–269
 slist 102
 SMCL format 17–18
 SMR 227–232
 sort 89
 specificity 248–250
 spreadsheets 64–65
 SSC archive 23
 standard populations 232–235
 standardization 227–235
 direct 232–235
 indirect 227–232
 standardized mortality ratio 227–232
 Stata manuals 21, 307–308
 Statalist 25
 statistical functions 77
 statsby: (prefix) 266–268
 Stat/Transfer 64
 stcox 218–223
 stcurve 219–220
 stir 212
 storage types 46–48
 stphplot 218–219
 stptime 210–211
 strate 229
 stratified analysis 171–184, 211–213
 case–control studies 175–184
 cohort studies 171–174
 incidence rates 211–213
 string formats 51
 string variables 50–54
 sts graph 216
 sts list 215
 sts test 217
 stset 214–215
 stsplot 223–225
 subscripting 82–85
 summarize 99–100
 survey data 205–206, 261

survival analysis 213–217
 svy: (prefix) 205–206
 svyset 205–206
 swilk 109–110
 syntax (command) 270–271
 syntax diagrams 31–35
 sysdir 5
 sysuse 60

T

tab1 103
 tab2 104–106
 tabi 120–121
 table 106, 114–115
 tabm 106–107
 tabodds 177–179
 tabstat 112–114
 tabulate 102
 tabulate, summarize 114
 technical support 25
 test 194–195
 testparm 194–195
 text editors 18
 text (graphs) 140–144
 time-dependent covariates (Cox regression)
 235
 time-series data 261
 title (graphs) 140
 tostring 52
 transformations 110–112
 translation programs 64
 ttest 115–116
 ttesti 122
 twoway connected (graph command)
 160–162
 twoway dropline (graph command)
 164
 twoway function (graph command)
 165–166
 twoway line (graph command) 158–
 160
 twoway rcap (graph command) 162–
 163
 twoway rspike (graph command) 162–
 163

twoway scatter (graph command)
 156–158

U

update 2–5
 updating Stata 2–5
 use 57–60

V

value labels 67–68
 variable labels 67
 variable lists 35
 variables 43–55
 dates 48–50
 numeric 43–50
 string 50–54
 Variables window 10
 variation of measurements 240–243
 view 18
 Viewer window 10

W

webuse 60
 weights 38–39
 window preferences 7–11
 Windows (Microsoft) 311–315
 WinZip 301

X

xi: (prefix) 194–196
 XML format 64–65
 xmlsave 64
 xmluse 65
 xpose 95
 xt (family of commands) 206–207
 xtlogit 206–207