Index

ADJUST option, 465
agriculture, 26
Akaike’s information criterion, 240, 457
PROC MIXED, 457
amatadine, 305–308
analgesic trial, 303–305
analysis of covariance, 15, 16, 18
animal breeding, 26
animal disease, cluster sample, 384–385
animal feeding, 72–73
animal physiology trial (breathing), 374–379
ANOVA, 5, 8, 9, 26
repeated measures, 233
anti-anxiety agent, 402
apoptosis, 394, 395
approximation, and bias, 137
area under curve (AUC), 232, 335
asthma, 329, 391, 412
ASYCOV option, 458
autocorrelation plot, 83, 96
autoregression: see first-order autoregressive correlation
average bioequivalence, 397, 401

Balaaam’s design, 305–308
balance, 27, 29–31
continuous effect, 31
least squares, 51
multi-centre, 211
unbalanced design, 327

balanced incomplete block designs, 425–429, 431, 433
banded covariance, 239
baseline covariate, 143
cross-over, 294
fixed effects, 294
hypertension study, 15–16
linearity assumption, 15–16
pre-treatment, 264
Bayesian approach
classical statistics comparison, 57
‘cold feet’ analysis, 147
exact statistics, 59
GLMM significance testing, 141
historical, 27
hypertension trial analysis, 84–86
negative variance components, 72
parameter estimation, 61
pre-eclampsia, 220, 221
software (PROC MCMC), 483
software (PROC MIXED), 473
standard error bias, 74
see also posterior density; posterior distribution
Bernoulli distribution, 115
general exponential form, 123
variance matrix, 120
Bernoulli form, 133, 480
dispersion parameter, 139
PROC GLIMMIX, 220
pre-eclampsia, 218–230

© 2015 John Wiley & Sons, Ltd. Published 2015 by John Wiley & Sons, Ltd.
Companion Website: www.wiley.com/go/brown/applied_mixed
Index

beta distribution, 67
between-subject design
  animal physiology, 375
different variances, 370–373
bias
  of approximations, 137
Bayesian models, 60
cross-over, 327
dispersion parameter, 139–140
empirical estimator, 243
fixed effects, 50
GLMM, 139
mixed model, 26
ML estimates, 48
pseudo-likelihood, 131, 132
random effects, 133, 137
random effects prediction, 53
shrinkage, 136–137, 139
see also error; standard error
bilaterial data, 418–422
binary data
cross-over, 317–321
dispersion parameter, 120
extended binary form, 170, 177
matchet sets, 357
modelling, 115–117
odds ratio, 318, 326
sample size estimation, 215, 288
see also Bernoulli form; binomial form;
  parameterisation
binomial distribution, 116, 120
  exponential form, 123–124
  variance matrix, 120
binomial form, 133, 480
covariance pattern model, 477
dispersion parameter, 139
mortality estimate, 388
bioavailability, 397
bioequivalence, 329, 397
bioequivalence trials, 329
bivariate normal distribution, 22
block diagonal matrix, 41, 44
blocking effect, 31, 234, 482
covariance pattern, 31, 44–46
error stratum, 32
G matrix, 46, 279, 464
R matrix, 44–45
random coefficients model, 31
REPEATED statement, 372
blood pressure: see hypertension
breast screening, 414–415
Edinburgh randomised trial of, 414–415
breathing trial: see animal physiology trial
cancer
  children’s (herpes virus study), 270–286
  ovarian, 262
canonical link function: see under link
  function
cardiac output trial
  complete block design, 297–299
  covariance pattern model, 309–314
cardiology: see heart failure; hypertension
cardiology trial, dogs, 381–383
carry-over
  Balaam’s design, 305–308
critiqued, 305
  four-period, four-treatment, 298–299
  information recovery, 303
  Koch’s design, 301–305, 310
  model choice, 312
  standard error, 308
case control studies: see matched
  case–control studies
categorical data
cross-over trial, 321–323
  sample size, 217, 288
  unordered, 180
see also mixed ordinal logistic regression;
  ordinal logistic regression;
  unordered categorical data
categorical effects, 36
  X matrix, 36
categorical mixed model, 168–196
  ‘cold feet’ analysis, 183, 196
  epilepsy trial, 258
  GLMM comparison, 240
  model checking, 182, 360–362
  CD4 count, 267–269
centre effects
  model analysis, 198–202
  model checking, 93
random, 372–373
see also centre-treatment effects
centre-treatment effects, 199–201
  binary data, 198–202
  ‘cold feet’ analysis, 150
  DBP analysis, 90, 95
different variances, 372–373
model analysis, 198–202
negative variance components, 211
random effects model, 18–19
see also multi-centre trial; treatment effect estimates
centre-treatment interaction
normal mixed models, 86–88
plausibility, 209
centres
number of, 210
size, 210
cerebrovascular insufficiency, 317–321
change score, 317, 322
contingency table, 323
chemotherapy: see herpes virus cancer study
chi-squared test, 78, 224
CHISQ option, 470, 478
GLM, 140–141
Pearson, 224
test for trend (contingency table), 115, 168
Wald statistic, 220, 470
childhood cancer study, 270–286
CL option, 458, 459, 462, 467
cluster randomised trials, 411–415
asthma treatment, 412
‘cold feet’ analysis: see under hypertension study
cluster sample survey, 384–385
communicating results, 24–25, 58
complete block designs, 297–299
compound symmetry pattern, 46
conditional distribution, 68
conditional logistic regression, 125
confidence intervals
‘cold feet’ analysis, 149–150
ESTIMATE statement, 467
GLM(M), 141–142
LSMEANS statement, 465
MODEL statement, 457, 458
normal mixed model, 78
conjugacy, 67
containment, 28–29
ESTIMATE statement, 466–470
containment stratum, 31–33
contingency table analysis
dysmenorrhoea trial, 321–323
GLM superiority, 115
logistic regression, 168
contrast, 75–77, 470
CONTRAST statement, 470–471
convergence
Bayesian methods, 59
GEE, 130
GLM, 123
GLMM, 134
iterative methods, 54–55
Newton–Raphson, 54
PROC MIXED, 271, 456
pseudo-likelihood, 132
shrinkage, 176
uniform effect categories, 133–134, 139
CORR option, 462, 482
correlation
compound symmetry, 46
negative, 72–73
sample size, 286–288
cot death study, 357–370
binary variables, 362
count data, 116, 254
categorical mixed model, 260
GLMM, 260
see also under covariance pattern model:
Poisson distribution
covariance matrix: see V matrix
covariance parameters
categorical mixed model, 259
exponential decay, 245
fixed/mixed comparison, 19–20
nested, 240: see also nesting
R matrix, 182
simple model, 19–20
covariance pattern
compound symmetry model, 236–237
count data, 254–262
first-order autoregressive, 235–237, 241
general (unstructured), 235
heterogeneous, 237
SAS, 472, 479
simple, 235–237
Toeplitz, 236, 247
covariance pattern choice
categorical mixed models, 182
repeated measures, 239–241, 335
strategy for, 241, 335
and trial size, 247
covariance pattern model, 19–22,
175–177, 362
blocking variable, 31, 234–235
compound symmetry structure, 313
cross-over trial, 140, 290, 308–313
covariance pattern model (continued) 140

dispersion parameter, 140
error stratum, 32
first-order autoregressive, 333
four-way cross-over trial, 309–314
GLMM \( \mathbf{V} \) matrix, 127
historical, 26
mixed ordinal logistic regression, 175
model fitting, 245–247
multi-centre trial with repeated measures, 350–352
normal data, 244–254
model checking, 248–254
pattern selection, 245–247
number of parameters, 181
quasi-likelihood, 128
random coefficients model comparison, 343
repeated measures, 234–244
\texttt{REPEATED} statement, 472–473
repeated within visits, 330–335
see also reparameterisation

covariance structure
covariance pattern models, 44–46
likelihood ratio test, 335
random coefficients model, 42–44
random effects model, 39–42
SAS options, 464
covariate effects
parameters, 36
\( \mathbf{X} \) matrix, 36
Crohn’s disease, 394
cross-over trial
AB/BA, 290–296, 317–321
Balaam’s design, 305–308
balance, 30
binary data analysis, 317–321
categorical data, 321–323
complete block, 297–299
covariance pattern models, 140, 290, 308–313
defined, 289
error stratum, 31–33
fixed effects model, 35
four-way, 309–314
higher order complete block designs, 297–299
incomplete block designs, 302–305
Koch’s design, 303–305
matched studies similarity, 357
mixed models advantage, 1–3, 24, 290
multi-centre, 355
non-informative prior, 66, 76
optimal designs, 305–308
parallel group comparison, 294, 295, 325, 375
random effects model advantage, 326–328
simple model, 3–12
structured covariance pattern model, 308–313
by time, 309
by treatment, 309
trial design, 325
two-period, 303–305
cubic random coefficients model, 271, 273
cumulative probability, 169, 170
data transformation, 79
\texttt{DDFM} option, 457, 459
degrees of freedom (DF)
confidence intervals, 78
\( F \) tests, 76–77
negative variance, 70, 211
REML adjustment, 26
SAS default, 459
variance parameter accuracy, 49, 50, 75, 148
see also Satterthwaite DF
density function, 123, 124
canonical link function, 118
general exponential form, 117
likelihood function, 47
deprivation score, 359–360
design matrix: see \( \mathbf{X} \) matrix; \( \mathbf{Z} \) matrix
\texttt{DIFF} option, 465
dispersion parameter, 120–121
and bias, 139
cold feet’ analysis, 148–149
covariance pattern model, 140
fitting, 139–140
GLM, 139
GLMM, 127, 139–140
diuretic treatment
heart failure, 292–294
pre-eclampsia, 218–230
dose response studies: see repeated measures
within visits
dropout, 1–3, 91
cot death study, 363
DBP analysis, 95
hypertension trial, 13
repeated measures, 3, 242
drug registration, 313
dysmenorrhoea, 321–323

E option, 467
E3 option, 459, 460
eczema treatment, 391–393
effective sample size, 83, 100, 101, 152
efficiency, 292
electrocardiogram assessment, 317–321
EMPIRICAL option, 75, 244, 458
empirical variance, 75, 137, 243
PROC GENMOD, 318, 479
epilepsy trial, 114, 254, 258
categorical mixed model, 258
GLM analysis, 260
error
model choice, 7, 312
residual variance, 70
sample size, 327
variance parameters, 73
see also bias; residual; standard error
error stratum, 31
higher level, 32
mixed models information recovery, 32
multi-centre trial, 198
hypertension trial, REML, 85
ESTIMATE statement, 204, 345, 466–470
non-estimable, 468
non-estimable effects, 468
ethics
medical studies, 25
sample size calculation, 286
trial design, 375
event history analysis, 391–393
exact logistic regression, 134
experiment design, sample size, 375–379
explaining results, 24–25, 59
exponential family, 113, 117–118 see also GLM
exponential form, general, 117–118, 123–125
extended binary form, 177
F distribution, 70
F test, 5
CONTRAST statement, 470
contrasts, 76
denominator DF, 76–77
GLM(M), 140–141
PROC MIXED, 457
repeated measures, 242
factor effects: see categorical effects
farm, disease prevalence, 384
FDA, 397, 398
first-order autoregressive correlation, 235–237
Fisher scoring, 55
fixed effects estimates, 50–51
categorical mixed model, 260
covariance pattern models, 23
fixed effects matrix: see X matrix
fixed effects model, 35–37
assessing fixed effects, 247
assumptions, 7
balance, 29
cerebrovascular insufficiency, 317–321
complete block designs, 297–299
GEE, 130
incomplete data, 7, 290
matched sets, 357
mixed models comparison, 7
multi-centre trial, 198–199, 202–204
outliers, multi-centre trial, 198
random coefficients comparison, 263
random effects model, 7
repeated measures, 232–233
choice, 241–242
significance testing, 75–78
uniform categories, 133–134, 139
see also contingency table analysis: GLM
foetal scans, 380
4×4 factorial design
labouratory study of, 394–397
G matrix
covariance pattern model, 46
mixed ordinal logistic regression, 174
random coefficients model, 43–44
random effects model, 39–40
SAS, 462, 464
unordered categorical data, 180
gait, 418
general autoregressive model: see Toeplitz
covariance pattern model
generalisation of results, 210
generalised estimating equations (GEEs), 130–131, 481–483
Index

generalised least squares: see iterative
generalised least squares

generalised linear mixed models: see GLMM
generalised linear models: see GLM

Genstat, 453

Gewerke test, 111, 152

GLMMIX macro, 131, 158–480

GLM (generalised linear model), 113–125

V matrix, 119–120

confidence intervals, 141–142

contingency table, 114

defined, 118–121

dispersion parameter, 139

distributions, 105–107

fitting, 121–123

meta-analysis, 219–220

PROC GENMOD, 155, 481–483

significance testing, 140–141

GLMM (generalised linear mixed model), 125–132, 480

categorical model comparison, 240

confidence intervals, 141–142

defined, 126–127

dispersion parameter, 139–140

epilepsy trial, 260

fitting, 129–132

Bayesian approach, 132

GEE, 130–131, 481–483

pseudo-likelihood, 131–132

likelihood function, 127

meta-analysis, 218–230

model checking, 142

quasi-likelihood, 128–129

significance testing, 141

see also GLMMIX macro; mixed ordinal logistic regression; unordered categorical data

global estimates, 198, 355

centre and centre-treatment effects

random, 198

centre number, 198, 355

centre-treatment interaction, 198, 355

cluster sample surveys, 384

double hierarchical, 356–357

effect on standard error, 217

Goldstein, random effects fitting software, 177

grid search, 473

GROUP option, 372, 464

haemophilia study: see HIV

heart failure (diuretics) study, 292–294

Hedeker and Gibbons software, 177

herpes virus cancer study, 270–286

antibody level, 270–286

Hessian matrix, 54

hierarchical data, 23–24, 28–29

containment, 28–29

mixed model, 23–24

hierarchical model, 18

hierarchical structure
double, 356–357

multi-centre trial, 217, 356

highest posterior density interval (HPD), 61, 108

HIV

linear random coefficients model, 267–269

model fitting, 268

hypertension study

analysis models, 84–86

baseline covariate, 15–16

between-subject analysis, 371

categorical ‘cold feet’ analysis, 183

correlation parameter interpretation, 186

odds ratio, 185, 186

centre effects, 16–17

centre-treatment interaction, 17–18

‘cold feet’ analysis, 143–147, 162–163

data modelling, 13–14

fixed effects models, 157

Hedeker & Gibbons software, 183, 184

incomplete data, 90–95

introduction, 12–13

model checking, 90–95, 162–163

models, 143–144

multi-centre trial, 350–352

normal data, 244–254

random effects model, 147

repeated measures, 19–22, 350–352

results analysis, 86–88

sample size, 287

SAS code, 101–112

treatment-centre, 213–217
Index

503

ID statement, 474

identity function, as link function, 115, 125

incomplete block design, 302–305

incomplete data, 7

balance, 31

cross-over trial, 290–291

DBP analysis, 91

mixed model advantage, 1–3, 23–24

not at random, 234

randomness requirement, 81

repeated measures, 3, 231–234, 329–330

unbalanced design, 327

individual bioequivalence, 401–402

inference

fixed effects model, 202

model choice, 198

random effects model, 200, 209

see also generalisation of results

information matrix, 55

information recovery, 10, 32, 303, 304 see also incomplete data

integrated care pathways, 412

intention to treat, 84

inter/intra-observer variability foetal scan, 380

heart wall thickness, 381–383

invariance to time origin, 44, 278

inverse gamma distribution, 63, 66, 148

PRIOR statement, 473

iterative generalised least squares, 49–50

fixed effects estimation, 51

incomplete block design, 303

random effects prediction, 53

restricted, 49–50, 57

variance parameter estimation, 55–57

iterative methods, 53–55

GEE, 132

GLM fitting, 121–123

Newton–Raphson, 54–55, 121–123

pseudo-likelihood maximisation, 131–132

Jeffreys’ method, 66, 474

Kenward–Roger adjustment, 74–76, 88, 89, 148, 243, 265, 272, 312, 313, 327, 404, 459

knee angles, 420

Koch’s design, 303–305

L matrix, 470

last value carried forward, 14

least squares: see IGLS; ordinary least squares

least squares mean, 471, 490

leukaemia: see childhood cancer study

likelihood function

GLM, 121

GLMM, 127

infinite, 456

information criteria measures, 240

log likelihood, 47–50

GLM fitting, 121

GLMM, 128

ordinary/REML, 48–49

mixed ordinal logistic regression, 173

model fitting, 47–50

non-informative prior similarity, 62

ordinal logistic regression, 169

ordinary/REML, 48–49

standardised, 60

true likelihood, 47–50

see also maximum likelihood;

pseudo-likelihood;

quasi-likelihood; REML

likelihood ratio test, 240

covariance structure determination, 335–336

variance components significance, 77, 265

linear dependencies, X matrix, 37

linearised pseudo variable, 122, 131–132, 142

link function, 114–115

canonical, 115, 118–119, 123, 124

GLMM, 126

mixed ordinal logistic regression, 169

non-canonical, 119

see also log likelihood function; logit link function

Lipsitz macro

parameterisation, 177

source, 477

local estimates, 198, 355

centre effects random, 355

grid search, 456

meta-analysis, 217
Index

local maximum, PROC MIXED, 456
location parameter, 61, 117
log likelihood: see under likelihood function
log link function, 135–136
logistic regression, 125
conditional, 125
see also conditional logistic regression;
logit link function
logit link function, 123
Bernoulli, 123
binomial, 123–124
cumulative probability, 169
generalised, 178
see also link function; log link function
LSMEANS statement, 465–466
LSMEANSTIMATE statement, 454, 477,
481
lung function trial, 374–379
MAKE statement, 476
marginal methods, 49 see also REML
marginal posterior distribution, 61
Markov chain Monte Carlo methods, 68
matched case–control studies, 357
cot death study, 357–370
see also matched sets
matched sets
cot death study, 362–364
fixed effects model, 358
Matlab, 425, 426
matrix notation, 34
matrix, positive semi-definite, 264
maximum likelihood
Bayesian comparison, 58
fixed effects estimation, 50
historical context, 25
model fitting, 47
random effects prediction, 51–53
see also likelihood function; REML
mean response, 232–233
median, 61, 86
meta-analysis, 217–218, 356
data inclusion criteria, 221
example, 218–230
mixed model, 2
method of scoring: see Fisher scoring
METHOD option, 458
Metropolis algorithm 68–69, 86, 483
micturition frequency, 293
missing data: see incomplete data
mixed model, 476
advantages, 1–3, 23–24
defined, 2, 22, 37–39
disadvantages, 24
historical, 25–27
incomplete data, 234
multi-centre trial, 198
see also covariance pattern model; normal
mixed model; random coefficients
model; random effects model
mixed ordinal logistic regression
binary modelling, 169
covariance pattern model, 175
defined, 173–177
as GLMM, 169, 173–177
model fitting, 177, 183, 184
R matrix, 174–176
MLwiN, 453–454
mode, 61
model-based approach, covariance pattern
models, cross-over trials, 312
model building, 3–12
childhood cancer study, 270
non-linear temporal, 270
repeated measures, 335
model checking, 79–81
categorical mixed, 182
covariance pattern, 244, 248–254
DBP analysis
centre effects, 93, 95
centre-treatment effects, 95, 95
GLMM, 142
outliers, 90
residuals, 90–95
SAS code, 101
matched case–control, 360–362
Pearson residuals, 256
polynomial random coefficients, 270
random coefficients, 265, 273
random effects ‘cold feet’ analysis, 145,
162–163
see also residual plots
model choice
assumptions, 312
binary data, 326
covariance pattern model, 312
multi-centre trial, 198–202
non-normal data, 326
small samples, 410
standard error, 295, 312
model comparison, covariance pattern models, 240
model fitting, 46–57
Bayesian, 57–69
covariance pattern model
repeated measures, 240, 245–247
systolic blood pressure trial, 335–340
GLM, 121–123
GLMM problems, 139–140
likelihood function, 47–50
linear random coefficients, 267
manually, 11
measures of, 224, 240
mixed ordinal logistic regression, 177, 183, 184
multi-centre trial with repeated measures, 350–352
overfitting, 239–240
parameter estimation, 108–110
PROC MIXED, 457
statistical comparison, 240
unordered categorical data, 180
model selection: see model choice
MODEL statement, 458–462, 481
mortality estimates, 386–388
mouthwash trial, 295, 325
multi-centre trial
analysis considerations, 209–211
balance, 211
centre effects, 209–211
cross-over, 355
defined, 197
fixed effects model
analysis implications, 198–199
centre-treatment effects omitted, 201–202
outliers, 198
hierarchical, 217, 356
meta-analysis, 198–202, 217–218
data inclusion criteria, 221
example, 218–230
outliers, 217–218
mixed model, 198
number of centres, 210
random effects model, 220
repeated measurements, 349–352
sample size estimation, 211–217
multinomial correlation, R matrix, 172–173
multinomial distribution, mixed ordinal logistic regression, 172
multinomial probability, unordered categorical, 178
multiple contrast, 75–77
multivariate normal distribution
checking, 248–254, 266, 273
density function, likelihood, 47
SAS coding, 279
negative correlation, 72–73
nesting, 240–241, 247, 351
Newton–Raphson iteration, 54–55, 121–123
nicotine, 394, 395
NOBOUND option, 264, 272, 409
NOCLPRINT option, 458
NOINFO option, 458
NOITPRINT option, 458
non-comparative data, 265
non-estimable effects, 158, 468
non-normal data
distribution, 113, 114
likelihood, 128
model selection, 326
see also GLM
normal data
covariance pattern model, 244–254
sample size estimation, 211–217, 286
normal distribution
assumption checking, 79
bivariate, 22
link function, 125
residual, 4
transformation to, 267
see also multivariate normal distribution
normal mixed model, 34, 37–39
Bayesian approach, 57–69
fixed effects estimation, 50–51
likelihood function, 47–50
model fitting, 46–57
random effects estimation, 51–53
relation to GLM, 113
variance parameter estimation, 53–57
see also covariance pattern model; mixed model; random coefficients model; random effects model
normal probability plots, 79, 90, 142, 163, 360, 361
null hypothesis, 62, 76
numerical methods, 53–55, 67
observer variation
  binary data, 28, 325, 326
categorical data, 265, 321
‘cold feet’ analysis, 150
dog cardiology, 381–382
epilepsy, 260
foetal scans, 380
logit link function, 123
odds ratio, 135, 166, 185, 186, 318, 326
ODS OUTPUT statement, 475, 480, 482
ODS SELECT statement, 110, 164, 484
oedema status, 293, 296
offset, Poisson distribution, 116
one-parameter distributions, 115–117
order effects, 290
ordered categorical data: see under ordinal logistic regression
ordinal logistic regression, 168–173
binary data, 319
categorical data, 323, 357, 359, 363
see also mixed ordinary logistic regression
ordinary least squares, 51, 84
outliers, 79–81
GLMM model checking, 142
meta-analysis, 217–218
multi-centre analysis, 84, 87
Pearson residual plots, 257
polynomial random coefficients model, 244, 266
random coefficients model, 267
random effects model, 198
repeated measures analysis, 248–254
SOLUTION option, 464
ovarian cancer trial, 262
overfitting, 239–240
overinterpretation of data, 273
overparameterisation, 37
p-values, Bayesian, 59, 78
parallel group design, 294, 325, 375
parameter estimation, Bayesian, 61
parameterisation, 175, 181
Parkinsonism drug trial, 305–308
PARMS statement, 473
partitioning
  categorical data, 217
  extended binary form, 169
  proportional odds assumption, 181
patella tracking, 419, 420
patient: see subject
PDIFF option, 465
Poisson residuals, 156, 257
period effect, 293: see also time effect
pharmaceutical industry
  empirical variance, 244
  model choice, 313
physiological response, 232
plaque score, 295
PM option, 459, 462
Poisson data
  covariance pattern model fitting, SAS, 481
  dispersion parameter, 139
  shrinkage, 136–137
Poisson distribution
  general exponential form, 124
  model checking, 257
  offset, 116
  variance matrix, 120
Poisson regression: see log link function
polynomial model building, 270–272
pooled comparison, 291
population bioequivalence, 397, 401, 402
population averaged method: see marginal quasi-likelihood
post-operative complications, 389
posterior density, 59–60
determination, 59–60
posterior distribution, 58, 61
evaluation, 67–69
marginal, 61
SAS, 110
simulated, 68
power: see sample size estimation
pre-eclampsia, 218–230
pre-treatment: see under baseline covariate
predicted values, 482
Prescott’s test, 318, 322
prior distribution, 58
conjugate, 67
flat, 63, 66, 474
informative, 74
non-informative, 58, 66
‘cold feet’ analysis, 143, 149
prior similarity, 63
specification, 62–67
PRIOR statement, 473–474, 486–487
proper, 66–67
properties, 66–67
probability
cumulative, 169, 170
multinomial, 178
see also p-values
probability intervals, 60
probit function, 119
PROC GENMOD, 155, 480–483
PROC GLIMMIX, 137–138, 150, 185, 260, 313, 476
PROC MCMC, 483–488
PROC MIXED, 454–476
default fitting, 55
empirical variance, 75
negative variance, 70
statement options, 478: see also specific statement
variance parameter fixing, 74
proportional odds assumption, 181
protocol checking, multi-centre trial, 198
pseudo-likelihood, 131–132
‘cold feet’ analysis, 147
GLIMMIX macro, 476
pseudovariable, 122, 131–132
quadratic function, 53
quadratic random coefficients model, 273–278
qualitative variables: see categorical effects
quasi-likelihood, 128–129
categorical, 182
GLMM, 129
maximisation, 129–132
R matrix, 40
banded covariance, 239
compound symmetry structure, 238
covariance pattern model, 44–46
general structure, 238
GLMM, 126
mixed ordinal logistic regression, 174–176
random coefficients model, 43
random effects model, 40
repeated measures, 235, 238
REPEATED statement, 472–473
submatrices, 45
uncorrelated, 238
radiologist reliability, 380
random coefficients model, 23, 267
covariance pattern model comparison, 343
covariance structure, 42–44
cubic, 273
effect stratum, 33
examples, 267–286
fitting by GEE, 132
GLMM V matrix, 127
likelihood function, 127
linear, 263
model fitting, 263
repeated measures (HIV), 267–269
model checking, 273
non-linear, 270
polynomial (childhood cancer), 270, 281
quadratic, 273–278
random effects, 21
RANDOM statement, 279, 464
repeated measures, 21–22, 262
repeated measures within visits, 341–343
shrinkage of estimates, 265
significance testing, 75–78, 265
systolic blood pressure trial, 341–347
see also count data; covariance pattern model; event history analysis;
random coefficients model
random effects coefficient, 37, 51–53
random effects estimation, 12
random effects model, 22
assumptions, 7
balance, 30–31
‘cold feet’ analysis, 147
complete block designs, 297–299
compound symmetry, 236
covariance pattern comparison, 309–314
covariance structure, 39–42
cross-over trial, 143, 326–328
fitting by GEE, 132
fitting software, 181, 452–454
fixed models comparison, 7–9
GLMM, 126, 127
hierarchical, 18
likelihood, 127
model checking, 79–81
multi-centre trial, 200–202
negative variance, 69
pre-eclampsia trial, 218–230
R matrix, 176
random coefficient models, 21
random effects model (continued)  
significance testing, 75–78  
standard error bias, 148  
uniform effects, 133–134, 139  
V matrix, 127  
see also reparameterisation  
RANDOM statement, 279, 336–338, 462–465  
random vs. fixed effects modelling, 7  
randomisation, cross-over trial, 290  
ranking, 322, 387–388  
reciprocal distribution, 63  
recovery of information  
  binary data cross-over, 317–321  
  cross-over, 292, 294  
  incomplete block, 303  
  Koch’s design, 303  
  matched sets, 357  
  two-subject period cross-over, 303  
reference category, 158  
relative rate, 136, 257  
relative risk, 136  
see also log link function  
REML (residual maximum likelihood)  
  DBP analysis model, 85  
  definition, 48  
  fixed effects estimation, 50  
  historical, 26  
  log likelihood, PROC MIXED, 454  
  model fitting, 47  
  random effects prediction, 51–53  
  standard error, 87  
reparameterisation  
  case control, 362  
  GLMM, 176  
  random effects models as covariance  
    pattern model, 175–176  
repeated measures, 231–234  
  analyses by time point, 233  
  banded covariance, 239  
  containment stratum, 32  
  covariance pattern model, 234–244  
    count data, 254  
    normal data, 244–254  
  different covariances, 238  
  different variances, 237  
  error stratum, 32  
  event history analysis, 391  
  fixed effects model, 232–233  
  hypertension trial, 19–22, 244–254  
linear, 262, 267–269  
mixed model, 3, 234  
model choice, 241–242, 266  
non-comparative datasets, 265  
polynomial, 270–272  
sample size estimation, 287  
see also outliers  
repeated measures within visits, 329–347  
covariance structure, 332  
cross-over trial (SBP)  
  covariance pattern choice, 335–340, 343–347  
  random coefficients, 341–347  
  reps cross-over, 335–340, 343–347  
  defined, 329  
  multi-centre trial, 349–352  
  random coefficients model, 341–347  
  treatment-reps interaction, 329–330  
REPEATED statement, 336–338, 472–473, 482  
replicate cross-over designs, bioequivalence  
  studies with, 397–411  
residual, 4, 482  
error strata, 31  
full, 48  
homogeneity, 79  
normal distribution, 4  
Pearson, 142  
standardised, 79  
see also R matrix; residual plots  
residual maximum likelihood: see REML  
residual plots  
  covariance pattern model, 244  
  GLMM, 142  
  matched case–control, 359  
  normal data, 79–81, 90–92  
  polynomial random coefficients model, 257, 262  
  repeated measures, 248–254, 262  
SAS code, 155  
residual variance, 4–6  
response profile, 232  
restricted IGLS, 49–50, 57  
restricted maximum likelihood: see REML  
resurface, 420  
robust methods, 80  
run-in period, 290, 310  
safety trials, 263  
sample size
model selection, 389
shrinkage, 386–390
small, 326
sample size estimation
animal physiology trial, 375–379
categorical data, 288
multi-centre trial, 211–217
precision of, 217
repeated measures, 287
sampling cost
binary data, 215
normal data, 214
SAS
empirical standard error, 352
GLIMMIX macro, 480
Lipsitz macro source, 453
ordinal logistic regression macro, 323
PROC GENMOD, 158, 480–483
PROC MIXED, 454–476
base density, 75
default fitting, 55
empirical variance, 75
negative variance, 75
statement options, 478: see also specific statement
residual plots, 155
syntax, 454
variance parameter fixing, 75
uniform reference category, 155
V matrix, 462
SATTERTH option, 459, 460
Satterthwaite DF, 76–77
GLMM, 141
repeated measures data, 242
Schwarz’s information criterion, 241, 457
selected comparison design, 431–433
sensitivity analysis, 377
shrinkage, 12, 24
bias, 132, 136, 139
dispersion parameter as metric, 139
ESTIMATE statement, 469–470
GLMM, 136–137, 139
hierarchical multi-centre trial, 356
meta-analysis, 217–218
random coefficients estimates, 265
random effects, 52, 218, 220, 221
ranking, 386–388
raw data comparison, 386–388
small area mortality, 386, 387
surgeon performance, 390
treatment by centre, 203
side effects, 80, 263
significance testing, 75–78
Bayesian: see p-values ‘cold feet’ analysis, 143, 147
GLM, 140–141
GLMM, 141
negative variance components, 211
random coefficients model, 265
repeated measures, 242
simulation, 59, 74, 243
SLICE statement, 471
SLICEDIFF option,
small area estimates, 386–388
small sample bias, 312, 326
shrunken estimates, 386
social deprivation score, 359–360
SOLUTION option, 163, 459, 463–464, 480
sparse data, 141, 313
sphericity, test of, 236
S-plus, 453
SPSS, 453
standard error, 50, 72, 87–88, 134, 149, 243–244
Bayesian, 61, 86–88
carry-over, 298, 303, 308
compound symmetry structure, 313
covariance pattern models, 23, 312
cross-over, 326
different variances, 370, 371
empirical ‘cold feet’ analysis, 187
model-based comparison, 247, 257
fixed effects, 50, 243–244
heart failure cross-over trial, 294
hypertension study, 86–88
increase by mixed model, 198
linear random coefficients model, 267
model-based, compound symmetry structure, 312
model-based/empirical covariance pattern model (cross-over), 312, 337
epilepsy trial, 260
model choice, 10, 295, 313
multi-centre trial, 202–204, 217
random effects, 137
repeated measures fixed effects, 233
510  Index

standard error (continued)
  - simple model, 6
  - treatment-centre interaction, 86, 87
  - treatment, fixed effects model, 199
Stata, 454
statistics, explaining results, 24–25
STORE statement, 474
subject effects, 3
SUBJECT option, G matrix blocking, 464
subject-time effects, 21
submatrices, R matrix, 45
sudden infant death syndrome: see cot death study
summary statistics, 329
surgical audit, 2
syntax: see specific procedure or option
t test, 6, 76, 322, 388
Taylor series, 122, 131
test for trend: see chi-squared test; t-test
thinning, 83, 111, 167
time effect, 293
covariance patterns, 237
model choice, 271
non-linear models, 263
random coefficients model, 21–22, 265, 341–343
virus antibodies, 19, 270–278
visit time–response model, 341–342
time interval, 239
time slope, 21, 341
Toeplitz covariance pattern model, 236, 245
  - repeated measures, 332
  - SAS, 480, 482
toxicoology experiments, 119
trace plot, 69, 83, 96
treatment effect estimates, 198–199
  - binary data cross-over trial, 317–321
  - DBP study, 89
fixed effects, 198–199
  - local/global, multi-centre trial, 217
  - multi-centre trial with repeated measures, 349, 352
random effects model, 197, 203
relative rate, 257
separate analyses by time, 233
variance, 292
treatment-centre: see centre-treatment
treatment-rep effects, 344
treatment-time effect
  - hypertension trial, 19–20
  - linear random coefficients model, 265
  - omitted, 241–242
  - random coefficients models, 263
  - repeated measures, 247
  - weighting, 247
  - see also time effect
trial design
  - cross-over trial, 325
  - inclusion criteria, 221
  - optimisation, 375–379
  - sensitivity analysis, 377
  - variance components, 374–379
  - see also sample size estimation
  - trial-treatment, pre-eclampsia study, 218–221
type III tests: see Wald statistic
  - TYPE option, 279
ulcerative colitis, 394
ultrasound scans
  - dog cardiology, 381–382
  - foetal, 380
under-dispersion, 139
uniform distribution, non-informative prior, 63
uniform effect categories, 144
  - ‘cold feet’ analysis, 143
  - GLMM, 133–134
  - reparameterisation, 362
unordered categorical data, 180
V matrix, 37, 39, 462
  - Bernoulli distribution, 120
  - between-subject trial, 371
  - binomial distribution, 120
  - fitting methods, 126
  - GLM, 119–120
  - GLMM, 126
  - Poisson distribution, 120
  - random coefficients model, 44
  - random effects model, 40–42
  - repeated measures within visits, 330–335
  - SAS, 456, 462
standard error bias, 87–88
time origin invariance, 44, 278
Index

V option, 463

Variance
- canonical link function, 119
different variances, 237, 370–373
dispersion parameter, 120–121, 127
empirical, 75, 137, 244
empirical/model based (PROC GENMOD), 480–483
general exponential form, 117–118
negative, 363
non-convergence, 264
randomisation, cross-over trial, 291
sample size estimation, 215–216, 286
variance components, 7, 473
animal physiology trial, 374
ANOVA, 8
Bayesian, 85
bias, 149
DBP, 89
estimation, 61
negative, 69–73, 134, 143, 211
random coefficients, 272
negative:Bayesian, 72
number of centres, 210
variance matrix: see V matrix
variance parameters
accuracy, 73, 148
bias, 50
pseudo-likelihood, 131
shrinkage, 136–137
covariance, 55
defined, 46–47

GEE, 131
IGLS, 54–57
maximum likelihood methods, 53
significance testing, 77, 141

Wald statistic
GLM(M), 141
PROC MIXED (Type III test), 457
t test, 76
washout, 290
WEIGHT statement, 474
weighting, treatment-time effect, 247
Williams modification, dispersion parameter, 121
WinBUGS, 454
withdrawal: see dropout
within-subject design
animal physiology, 375–378
correlation, 20
cross-over trial, 291
see also cross-over trial
within-visit observations, correlation matrix, 330

X matrix, 36–37
mixed ordinal logistic regression, 171
unordered categorical data, 179

Z matrix, 38–39
random coefficients, 43
unordered categorical data, 179