

DigiSem
Wir beschaffen und
digitalisieren



^b
UNIVERSITÄT
BERN
Universitätsbibliothek Bern

Dieses Dokument steht Ihnen online zur Verfügung
dank DigiSem, einer Dienstleistung der
Universitätsbibliothek Bern.

Kontakt: Gabriela Scherrer
Koordinatorin digitale Semesterapparate
E-Mail digisem@ub.unibe.ch, Telefon 031 631 93 26

José C. Pinheiro
Douglas M. Bates

Mixed-Effects Models in S and S-PLUS

With 172 Illustrations



Springer



José C. Pinheiro
Department of Biostatistics
Novartis Pharmaceuticals
One Health Plaza
East Hanover, NJ 07936-1080
USA
jose.pinheiro@pharma.novartis.com

Douglas M. Bates
Department of Statistics
University of Wisconsin
Madison, WI 53706-1685
USA
bates@stat.wisc.edu

Series Editors:

J. Chambers
Bell Labs, Lucent
Technologies
600 Mountain Ave.
Murray Hill, NJ 07974
USA

W. Eddy
Department of Statistics
Carnegie Mellon University
Pittsburgh, PA 15213
USA

W. Härdle
Institut für Statistik und
Ökonometrie
Humboldt-Universität zu Berlin
Spandauer Str. 1
D-10178 Berlin
Germany

S. Sheather
Australian Graduate School
of Management
University of New South
Wales
Sydney NSW 2052
Australia

L. Tierney
School of Statistics
University of Minnesota
Vincent Hall
Minneapolis, MN 55455
USA

Library of Congress Cataloging-in-Publication Data
Pinheiro, José C.

Mixed-effects models in S and S-PLUS / José C. Pinheiro, Douglas M. Bates
p. cm. — (Statistics and computing)
Includes bibliographical references and index.
ISBN 0-387-98957-9 (alk. paper)
I. Bates, Douglas M. II. Title. III. Series.
QA76.73.S15P56 2000
005.13'3—dc21

99-053566

Printed on acid-free paper.

© 2000 Springer Verlag New York, LLC

All rights reserved. This work may not be translated or copied in whole or in part without the written permission of the publisher (Springer-Verlag New York, LLC, 175 Fifth Avenue, New York, NY 10010, USA), except for brief excerpts in connection with reviews or scholarly analysis. Use in connection with any form of information storage and retrieval, electronic adaptation, computer software, or by similar or dissimilar methodology now known or hereafter developed is forbidden. The use of general descriptive names, trade names, trademarks, etc., in this publication, even if the former are not especially identified, is not to be taken as a sign that such names, as understood by the Trade Marks and Merchandise Marks Act, may accordingly be used freely by anyone.

Printed in the United States of America. (HAM)

9 8 7 6 5

SPIN 10995662

Springer Verlag is a part of *Springer Science+Business Media*

springeronline.com

Index

- ACF, *see* autocorrelation function
ACF S function, 241, 257, 397, 451
 lme objects, 452
AIC, 10, 83
Akaike Information Criterion, *see* AIC
Alfalfa, *see* datasets
alternating algorithm, *see* Lindstrom
 and Bates algorithm
analysis of covariance model, *see*
 ANCOVA
ANCOVA, 30–40
anova, 12, 24, 453
 with single argument, 90
apparent volume of distribution, 379
AR, *see* autoregressive model
AR(1), 229
ARMA, *see* autoregressive–moving
 average model
Assay, *see* datasets
asymptotic regression model, 301, 511
 through the origin, 513
 with an offset, 368, 512
augPred, 39, 361
autocorrelation function, 227
autoregressive model, 228
autoregressive–moving average model,
 230
autoregressive–moving average model,
 228
Bayesian Information Criterion, *see*
 BIC
best linear unbiased predictor, *see*
 predictions
BIC, 10, 83
biexponential model, 278, 514
block-diagonal matrix, 162
BLUP, *see* predictions
BodyWeight, *see* datasets
CAR(1), *see* continuous AR(1)
Cefamandole, *see* datasets
clearance, 295, 351
CO2, *see* datasets
coef
 lmList objects, 457
 lme objects, 455
compartment model
 first-order open-compartment, 351,
 516
 one-compartment, 295
 one-compartment open with first-
 order absorption, 378
 two-compartment, 278

- compound symmetry, 161, 227
- confidence intervals, 92
 - correlation coefficient, 93
 - fixed effects, 92
 - standard deviation, 93
 - variance-covariance components, 93
- continuous AR(1), 229
- contrasts, 14
 - Helmert, 14
 - polynomial, 46
 - treatment, 16
- correlation structures, 205, 226
 - AR, 228
 - ARMA, 228
 - compound symmetry, 227
 - general, 228
 - MA, 229
 - spatial, 230
 - with gls, 251
 - with lme, 239
 - with nlme, 395
- corStruct, 232-249
 - classes, 234
 - corAR1, 235, 397
 - corARMA, 236, 397
 - corCAR1, 236
 - corCompSymm, 233
 - corExp, 238
 - corRatio, 263
 - corSpher, 262
 - corSymm, 234
 - spatial correlation classes, 237
- datasets
 - Alfalfa, 55, 425
 - Assay, 163, 425
 - BodyWeight, 105, 221, 244, 427
 - Cefamandole, 302, 427
 - CO2, 111, 130, 368, 428
 - Dialyzer, 214, 255, 401, 429
 - DNase, 131, 410, 429
 - Earthquake, 303, 430
 - ergoStool, 12, 431
 - Glucose2, 412, 432
 - IGF, 144, 155, 433
 - Indometh, 277, 433
 - Loblolly, 301, 434
 - Machines, 21, 435
 - Oats, 45, 160, 435
 - Orange, 338, 356, 436
 - Orthodont, 30, 147, 250, 436
 - Ovary, 239, 395, 437
 - Oxboys, 53, 100, 437
 - Oxide, 167, 437
 - PBG, 108, 411, 438
 - PBIB, 53, 439
 - Phenobarb, 294, 440
 - Pixel, 40, 440
 - Quinidine, 123, 378, 393, 441
 - Rail, 4, 443
 - Soybean, 287, 443
 - Spruce, 112, 444
 - Theoph, 120, 350, 363, 392, 413, 444
 - Wafer, 116, 170, 386, 448
 - Wheat2, 260, 448
- deriv, 339
- design plot, 13
- diagnostic plots
 - lme objects, 11
 - nlme objects, 361
 - random effects, 188-196
 - within-group error, 174-187
- Dialyzer, *see* datasets
- DNase, *see* datasets
- dotplot, 372
- Earthquake, *see* datasets
- EM algorithm, 79
- empirical model, 274
- ergoStool, *see* datasets
- extended nonlinear regression model, 332
- first-order open-compartment model, 351, 516
- fitted, 458
- fixed effects, 58
 - degrees-of-freedom, 91
 - hypothesis tests, 87
- fixef, 459
- four-parameter logistic model, 410, 517
- gapply, 120, 460
- Gauss-Newton algorithm, 325, 331
- Gauss-Seidel algorithm, 334
- generalized least squares, 201

- generalized nonlinear least squares, 333
- getGroups, 100, 461
- getInitial, 345
- gls, 205, 249–267, 462
 - methods, 250
 - with correlation structures, 251
 - with variance functions, 251
- GLS model, 203–205
- Glucose2, *see* datasets
- gnls, 332, 401–409, 464
 - methods, 402
- GNLS model, 333
 - approximate distributions of estimates, 335
- gradient attribute, 339
- grouped data, 97
 - balanced, 99
- groupedData
 - balancedGrouped, 109
 - constructor, 101, 108
 - display formula, 98
 - inner factor, 107
 - outer factor, 104
- groupedData, 466
- growth curve data, 30
- gsummary, 106, 121, 469
- heteroscedasticity, 178, 201, 291
- IGF, *see* datasets
- Indometh, *see* datasets
- information matrix, 82, 323
- initial attribute, 344
- intervals, 471
 - gnls objects, 403
 - lmList objects, 142, 473
 - lme objects, 156, 471
 - nlme objects, 363
 - nlsList objects, 281, 350
- isotropic correlation, 226, 231
- Laird-Ware model, *see* linear mixed-effects model
- likelihood
 - components of, 71
 - extended LME model, 203
 - extended NLME model
 - adaptive Gaussian approximation, 332
 - Laplacian approximation, 331
 - LME approximation, 331
 - GLS model, 204
 - multilevel LME model, 77
 - multilevel NLME model
 - adaptive Gaussian approximation, 321
 - Laplacian approximation, 318
 - LME approximation, 314
 - optimization, 79
 - profiled, 65
 - pseudo, 207
 - QR decompositions with, 68
 - restricted, 75
 - single-level LME model, 62
 - single-level NLME model
 - adaptive Gaussian approximation, 321
 - Laplacian approximation, 315
 - LME approximation, 313
- likelihood ratio tests, 83
- Lindstrom and Bates algorithm, 313, 330
 - LME step, 313
 - PNLS step, 313
- Lindstrom and Bates model, *see* non-linear mixed-effects model
- linear mixed-effects model
 - extended, 202–203
 - multilevel, 40, 60
 - single-level, 28, 58
- lm, 5, 134–138
- lme, 8, 146–174, 474
 - fixed and random, 146
 - confidence intervals, 156
 - maximum likelihood estimation, 150
 - methods, 147
 - multilevel, 167–174
 - coefficients, 170
 - predictions, 174
 - REML estimation, 150
 - single-level, 146–166
 - split-plot analysis, 160
 - with lmList object, 147
 - with variance functions, 214
- LME model, *see* linear mixed-effects model

- lmeControl**, 476
- lmList**, 32, 139–146, 478
 - confidence intervals, 142
 - methods, 140
- Loblolly**, *see* datasets
- logistic model**, 274, 338, 519
- logLik**, 479
- MA**, *see* moving average model
- Machines**, *see* datasets
- Manhattan distance**, 230
- matrix logarithm**, 78
- maximum likelihood estimators**
 - approximate distribution, 81
 - LME model, 66
- mechanistic model**, 274
- Michaelis–Menten model**, 520
- MLE**, *see* maximum likelihood estimators
- model**
 - asymptotic regression, 301, 511
 - with an offset, 368
 - asymptotic regression through the origin, 513
 - asymptotic regression with an offset, 512
 - biexponential, 278, 514
 - empirical, 274
 - first-order open-compartment, 351, 516
 - four-parameter logistic, 410, 517
 - logistic, 274, 289, 338
 - logistic regression, 519
 - mechanistic, 274
 - Michaelis–Menten, 520
 - one-compartment, 295
 - one-compartment open with first-order absorption, 378
 - two-compartment, 278
- moving average model**, 229
- multilevel model**
 - likelihood, 77
 - linear mixed-effects, 40, 60
 - lme** fit, 167
 - nonlinear mixed-effects, 309
- naPattern** argument to **nlme**, 298, 380
- Newton–Raphson algorithm**, 79
- nlme** library
 - obtaining, viii
- nlme**, 283, 479
 - fixed and random, 355
 - covariate modeling, 365–385
 - extended NLME model, 391
 - maximum likelihood estimation, 358
 - methods, 357
 - multilevel, 385–391
 - REML estimation, 387
 - single-level, 354–365
 - with **nlsList** object, 357
 - with variance functions, 391
- NLME model**, *see* nonlinear mixed-effects model
- nlmeControl**, 483
- nls**, 279, 338–342
- nlsList**, 280, 347–354, 485
 - methods, 349
- nonidentifiability**, 204
- nonlinear least squares**, 278
- nonlinear mixed-effects model**, 282
 - approximate distributions of estimates, 322
 - Bayesian hierarchical, 311
 - compared to LME model, 273–277
 - extended, 328–332
 - likelihood estimation, 312
 - multilevel, 309–310
 - nonparametric maximum likelihood, 311
 - single-level, 306–309
- nonlinear regression model**, 278
- NONMEM** software, 310
- normal plot**
 - of random effects, 188
 - of residuals, 179, 180
- nugget effect**, 231
- Oats**, *see* datasets
- one-compartment open model** with first-order absorption, 378
- one-way ANOVA**, 4–11
 - fixed-effects model, 6
 - random-effects model, 7
- Orange**, *see* datasets
- Orthodont**, *see* datasets
- orthogonal-triangular decomposition**, *see* QR decomposition

- Ovary, *see* datasets
- Oxboys, *see* datasets
- Oxide, *see* datasets
- pairs
 - lmList objects, 141
 - lme objects, 188, 190, 486
 - nlme objects, 359
- partially linear models, 342
- PBG, *see* datasets
- PBIB, *see* datasets
- pdMat, 157
 - classes, 158
 - pdBlocked, 162
 - pdCompSymm, 161
 - pdDiag, 158, 283, 364
 - pdIdent, 164
- peeling, 278
- penalized nonlinear least squares, 313
- Phenobarb, *see* datasets
- Pixel, *see* datasets
- plot
 - Variogram objects, 494
 - gnls objects, 404
 - groupedData object, 492
 - groupedData objects, 105, 490
 - lme objects, 175, 488
 - lm objects, 135
 - nls objects, 341
 - nlsList objects, 350
- positive-definite matrix, *see* variance-covariance
- precision factor, 313
- predictions
 - lme objects, 150, 495
 - augmented, 39, 361
 - BLUP, 37, 71, 94
 - multilevel model, 174
 - NLME model, 323
 - random effects, 37
 - response, 37, 94
- pseudo-likelihood, 207
- qqnorm
 - gls objects, 253
 - lme objects, 179, 180, 497
 - nlme objects, 361
 - random effects, 188
- QR decomposition, 66, 326
- Quasi-Newton algorithm, 79
- Quinidine, *see* datasets
- Rail, *see* datasets
- random effects
 - crossed, 163
 - multilevel, 60
 - overparameterization, 156
 - single-level, 58
- randomized block design, 12–21
- ranef, 498
 - lmList objects, 501
 - lme objects, 499
- rate constant, 278, 351, 379
- relative precision factor, 59
 - parameterization for, 78
- REML, *see* restricted maximum likelihood
- residuals
 - normalized, 239
 - Pearson, 149
 - response, 149
- residuals S function, 503
- restricted maximum likelihood
 - LME model, 75
 - NLME model, 314
- SBC, *see* BIC
- scatter-plot matrix, 359
- Schwarz's Bayesian Criterion, *see* BIC
- self-starting models, 342–347
 - available in nlme library, 346
 - SSasymp, 301, 511
 - SSasympOff, 369, 512
 - SSasympOrig, 513
 - SSbiexp, 279, 514
 - SSf01, 352, 516
 - SSfpl, 410, 517
 - SSlogis, 288, 347, 519
 - SSmicmen, 520
- selfStart, 343
 - constructor, 504, 505
 - formula objects, 506
 - functions, 346
- semivariogram, 230
 - robust, 231
- serial correlation, 226–230
- shrinkage estimates, 152
- Soybean, *see* datasets

spatial correlation, 230–232

exponential, 232

Gaussian, 232

linear, 232

rational quadratic, 232

spherical, 232

spatial data, 226

split-plot, 45–52

Spruce, *see* datasets

starting estimates, 340

Theoph, *see* datasets

time-series data, 226

time-varying covariate, 378

trellis

aspect ratio, 111

display of grouped data, 110

display of multilevel data, 116

panel function, 114

plot layout, 110

two-stage model, 309, 333

unbalanced data, 25

varFunc, 208–225

classes, 208

varComb, 213

varConstPower, 212, 220, 392

varExp, 211

varFixed, 208

varIdent, 209

varPower, 210, 217, 290

varReg, 268

variance covariate, 206

variance functions, 206–225

with gls, 251

with lme, 214

with nlme, 391

variance weights, 208

variance-covariance

components, 93

of random effects, 58

of response, 66

of within-group error, 202

pdMat classes, 157

Variogram, 245, 264, 507

lme objects, 508

varWeights, 208

volume of distribution, 295

Wafer, *see* datasets

Wheat2, *see* datasets

within-group error

assumptions, 58

correlation, 202

heteroscedasticity, 202