

Package ‘MEMSS’

January 2, 2012

Version 0.9-0

Date 2011-02-15

Title Data sets from Mixed-effects Models in S

Author Douglas Bates <bates@stat.wisc.edu>, Martin Maechler
<maechler@R-project.org> and Ben Bolker <bbolker@gmail.com>

Maintainer <lme4-authors@R-forge.wu-wien.ac.at>

Description Data sets and sample analyses from Pinheiro and Bates, “Mixed-effects Models in S and S-PLUS” (Springer, 2000).

Depends R(>= 2.12.0), lme4 (>= 0.999375-36)

LazyData yes

License GPL (>= 2)

Repository CRAN

Repository/R-Forge/Project lme4

Repository/R-Forge/Revision 1265

Date/Publication 2011-02-17 13:06:45

R topics documented:

Alfalfa	2
Assay	3
BodyWeight	4
Cefamandole	5
CO2	6
Dialyzer	7
Earthquake	8
ergoStool	9
Fatigue	9

Gasoline	10
Glucose	11
Glucose2	12
Gun	13
IGF	13
Machines	14
MathAchieve	15
MathAchSchool	15
Meat	16
Milk	17
Muscle	17
Nitrendipene	18
Oats	19
Orange	20
Orthodont	21
Ovary	22
Oxide	22
PBG	23
Phenobarb	24
Pixel	25
Quinidine	26
Rail	27
RatPupWeight	28
Relaxin	28
Remifentanil	29
Soybean	30
Spruce	31
Tetracycline1	31
Tetracycline2	32
Theoph	32
Wafer	34
Wheat	34
Wheat2	35
Index	36

Alfalfa

Split-Plot Experiment on Varieties of Alfalfa

Description

The Alfalfa data frame has 72 rows and 4 columns.

Format

This data frame contains the following columns:

Variety a factor with levels Cossack, Ladak, and Ranger

Date a factor with levels None S1 S20 O7

Block a factor with levels A to F

Yield a numeric vector

Details

These data are described in Snedecor and Cochran (1980) as an example of a split-plot design. The treatment structure used in the experiment was a 3×4 full factorial, with three varieties of alfalfa and four dates of third cutting in 1943. The experimental units were arranged into six blocks, each subdivided into four plots. The varieties of alfalfa (*Cossac*, *Ladak*, and *Ranger*) were assigned randomly to the blocks and the dates of third cutting (*None*, *S1*—September 1, *S20*—September 20, and *O7*—October 7) were randomly assigned to the plots. All four dates were used on each block.

Source

Pinheiro, J. C. and Bates, D. M. (2000), *Mixed-Effects Models in S and S-PLUS*, Springer, New York. (Appendix A.1)

Snedecor, G. W. and Cochran, W. G. (1980), *Statistical Methods (7th ed)*, Iowa State University Press, Ames, IA

Examples

```
str(Alfalfa)
(m1 <- lmer(Yield ~ Variety * Date + (1|Block), Alfalfa, verbose = TRUE))
```

Assay

Bioassay on Cell Culture Plate

Description

The Assay data frame has 60 rows and 4 columns.

Format

This data frame contains the following columns:

Block an factor with levels A and B identifying the block of the well

sample a factor with levels a to f identifying the sample corresponding to the well

dilut an ordered factor with levels 1 to 5 indicating the dilution applied to the well

logDens a numeric vector of the log-optical density

Details

These data, courtesy of Rich Wolfe and David Lansky from Searle, Inc., come from a bioassay run on a 96-well cell culture plate. The assay is performed using a split-block design. The 8 rows on the plate are labeled A–H from top to bottom and the 12 columns on the plate are labeled 1–12 from left to right. Only the central 60 wells of the plate are used for the bioassay (the intersection of rows B–G and columns 2–11). There are two blocks in the design: Block A contains columns 2–6 and Block B contains columns 7–11. Within each block, six samples are assigned randomly to rows and five (serial) dilutions are assigned randomly to columns. The response variable is the logarithm of the optical density. The cells are treated with a compound that they metabolize to produce the stain. Only live cells can make the stain, so the optical density is a measure of the number of cells that are alive and healthy.

Source

Pinheiro, J. C. and Bates, D. M. (2000), *Mixed-Effects Models in S and S-PLUS*, Springer, New York. (Appendix A.2)

Examples

```
str(Assay)
m1 <- lmer(logDens ~ sample * dilut + (1|Block) + (1|Block:sample) +
           (1|Block:dilut), Assay, verbose = TRUE)
print(m1, corr = FALSE)
anova(m1)
m2 <- lmer(logDens ~ sample + dilut + (1|Block) + (1|Block:sample) +
           (1|Block:dilut), Assay, verbose = TRUE)
print(m2, corr = FALSE)
anova(m2)
m3 <- lmer(logDens ~ sample + dilut + (1|Block) + (1|Block:sample),
           Assay, verbose = TRUE)
print(m3, corr = FALSE)
anova(m3)
anova(m2, m3)
```

BodyWeight

Rat weight over time for different diets

Description

The BodyWeight data frame has 176 rows and 4 columns.

Format

This data frame contains the following columns:

weight a numeric vector giving the body weight of the rat (grams).

Time a numeric vector giving the time at which the measurement is made (days).

Rat an factor with levels A to P identifying the rat whose weight is measured.

Diet a factor with levels a to c indicating the diet that the rat receives.

Details

Hand and Crowder (1996) describe data on the body weights of rats measured over 64 days. These data also appear in Table 2.4 of Crowder and Hand (1990). The body weights of the rats (in grams) are measured on day 1 and every seven days thereafter until day 64, with an extra measurement on day 44. The experiment started several weeks before “day 1.” There are three groups of rats, each on a different diet.

Source

Pinheiro, J. C. and Bates, D. M. (2000), *Mixed-Effects Models in S and S-PLUS*, Springer, New York. (Appendix A.3)

Crowder, M. and Hand, D. (1990), *Analysis of Repeated Measures*, Chapman and Hall, London.

Hand, D. and Crowder, M. (1996), *Practical Longitudinal Data Analysis*, Chapman and Hall, London.

Examples

```
str(BodyWeight)
```

Cefamandole

Pharmacokinetics of Cefamandole

Description

The Cefamandole data frame has 84 rows and 3 columns.

Format

This data frame contains the following columns:

Subject a factor giving the subject from which the sample was drawn.

Time a numeric vector giving the time at which the sample was drawn (minutes post-injection).

conc a numeric vector giving the observed plasma concentration of cefamandole (mcg/ml).

Details

Davidian and Giltinan (1995, 1.1, p. 2) describe data obtained during a pilot study to investigate the pharmacokinetics of the drug cefamandole. Plasma concentrations of the drug were measured on six healthy volunteers at 14 time points following an intravenous dose of 15 mg/kg body weight of cefamandole.

Source

Pinheiro, J. C. and Bates, D. M. (2000), *Mixed-Effects Models in S and S-PLUS*, Springer, New York. (Appendix A.4)

Davidian, M. and Giltinan, D. M. (1995), *Nonlinear Models for Repeated Measurement Data*, Chapman and Hall, London.

Examples

```
str(Cefamandole)
xyplot(conc ~ Time, Cefamandole, groups = Subject, type = c("g", "b"),
       aspect = 'xy', scales = list(y = list(log = 2)),
       auto.key = list(space = "right", lines= TRUE))
xyplot(conc ~ Time|Subject, Cefamandole, type = c("g", "b"),
       index.cond = function(x,y) min(y), aspect = 'xy',
       scales = list(y = list(log = 2)))
#fm1 <- nlsList(SSbiexp, data = Cefamandole)
```

CO2

Carbon Dioxide uptake in grass plants

Description

The CO2 data frame has 84 rows and 5 columns of data from an experiment on the cold tolerance of the grass species *Echinochloa crus-galli*.

Usage

CO2

Format

This data frame contains the following columns:

Plant an factor giving a unique identifier for each plant.

Type a factor with levels Quebec Mississippi giving the origin of the plant

Treatment a factor with levels nonchilled chilled

conc a numeric vector of ambient carbon dioxide concentrations (mL/L).

uptake a numeric vector of carbon dioxide uptake rates ($\mu\text{mol}/m^2 \text{ sec}$).

Details

The CO_2 uptake of six plants from Quebec and six plants from Mississippi was measured at several levels of ambient CO_2 concentration. Half the plants of each type were chilled overnight before the experiment was conducted.

Source

Potvin, C., Lechowicz, M. J. and Tardif, S. (1990) "The statistical analysis of ecophysiological response curves obtained from experiments involving repeated measures", *Ecology*, **71**, 1389–1400.

Pinheiro, J. C. and Bates, D. M. (2000) *Mixed-effects Models in S and S-PLUS*, Springer.

Examples

```

require(stats); require(graphics)
coplot(uptake ~ conc | Plant, data = C02, show.given = FALSE, type = "b")
## fit the data for the first plant
fm1 <- nls(uptake ~ SSasym(conc, Asym, lrc, c0),
  data = C02, subset = Plant == 'Qn1')
summary(fm1)
## fit each plant separately
fm1list <- list()
for (pp in levels(C02$Plant)) {
  fm1list[[pp]] <- nls(uptake ~ SSasym(conc, Asym, lrc, c0),
    data = C02, subset = Plant == pp)
}
## check the coefficients by plant
sapply(fm1list, coef)

```

Dialyzer

*High-Flux Hemodialyzer***Description**

The Dialyzer data frame has 140 rows and 5 columns.

Format

This data frame contains the following columns:

Subject a factor with levels A to T

QB a factor with levels 200 and 300 giving the bovine blood flow rate (dL/min).

pressure the transmembrane pressure (dmHg).

rate the hemodialyzer ultrafiltration rate (mL/hr).

index index of observation within subject—1 through 7.

Details

Vonesh and Carter (1992) describe data measured on high-flux hemodialyzers to assess their *in vivo* ultrafiltration characteristics. The ultrafiltration rates (in mL/hr) of 20 high-flux dialyzers were measured at seven different transmembrane pressures (in dmHg). The *in vitro* evaluation of the dialyzers used bovine blood at flow rates of either 200~dl/min or 300~dl/min. The data, are also analyzed in Littell, Milliken, Stroup, and Wolfinger (1996).

Source

Pinheiro, J. C. and Bates, D. M. (2000), *Mixed-Effects Models in S and S-PLUS*, Springer, New York. (Appendix A.6)

Vonesh, E. F. and Carter, R. L. (1992), Mixed-effects nonlinear regression for unbalanced repeated measures, *Biometrics*, **48**, 1-18.

Littell, R. C., Milliken, G. A., Stroup, W. W. and Wolfinger, R. D. (1996), *SAS System for Mixed Models*, SAS Institute, Cary, NC.

Examples

```
str(Dialyzer)
```

 Earthquake

Earthquake Intensity

Description

The Earthquake data frame has 182 rows and 5 columns.

Format

This data frame contains the following columns:

Quake a factor with levels A to U

Richter the intensity of the earthquake on the Richter scale

distance the distance from the seismological measuring station to the epicenter of the earthquake (km)

soil a factor with levels S (soil) and R (rock) giving the soil condition at the measuring station

accel maximum horizontal acceleration observed (g).

Details

Measurements recorded at available seismometer locations for 23 large earthquakes in western North America between 1940 and 1980. They were originally given in Joyner and Boore (1981); are mentioned in Brillinger (1987); and are analyzed in Davidian and Giltinan (1995).

Source

Pinheiro, J. C. and Bates, D. M. (2000), *Mixed-Effects Models in S and S-PLUS*, Springer, New York. (Appendix A.8)

Davidian, M. and Giltinan, D. M. (1995), *Nonlinear Models for Repeated Measurement Data*, Chapman and Hall, London.

Joyner and Boor (1981), Peak horizontal acceleration and velocity from strong-motion records including records from the 1979 Imperial Valley, California, earthquake, *Bulletin of the Seismological Society of America*, **71**, 2011-2038.

Brillinger, D. (1987), Comment on a paper by C. R. Rao, *Statistical Science*, **2**, 448-450.

Examples

```
str(Earthquake)
```

ergoStool

Ergometrics experiment with stool types

Description

The ergoStool data frame has 36 rows and 3 columns.

Format

This data frame contains the following columns:

effort a numeric vector giving the effort (Borg scale) required to arise from a stool

Type a factor with levels T1, T2, T3, and T4 giving the stool type

Subject a factor with levels A to I

Details

Devore (2000) cites data from an article in *Ergometrics* (1993, pp. 519-535) on “The Effects of a Pneumatic Stool and a One-Legged Stool on Lower Limb Joint Load and Muscular Activity.”

Source

Pinheiro, J. C. and Bates, D. M. (2000), *Mixed-Effects Models in S and S-PLUS*, Springer, New York. (Appendix A.9)

Devore, J. L. (2000), *Probability and Statistics for Engineering and the Sciences (5th ed)*, Duxbury, Boston, MA.

Examples

```
options(show.signif.stars = FALSE)
str(ergoStool)
print(m1 <- lmer(effort ~ Type + (1|Subject), ergoStool), corr = FALSE)
anova(m1)
```

Fatigue

Cracks caused by metal fatigue

Description

The Fatigue data frame has 262 rows and 3 columns.

Format

This data frame contains the following columns:

Path the test path (or test unit) identifier - a factor with levels A to U.

cycles number of test cycles at which the measurement is made (millions of cycles).

relLength relative crack length (dimensionless).

Details

These data are given in Lu and Meeker (1993) where they state “We obtained the data in Table 1 visually from figure 4.5.2 on page 242 of Bogdanoff and Kozin (1985).” The data represent the growth of cracks in metal for 21 test units. An initial notch of length 0.90 inches was made on each unit which then was subjected to several thousand test cycles. After every 10,000 test cycles the crack length was measured. Testing was stopped if the crack length exceeded 1.60 inches, defined as a failure, or at 120,000 cycles.

Source

Lu, C. Joseph , and Meeker, William Q. (1993), Using degradation measures to estimate a time-to-failure distribution, *Technometrics*, **35**, 161-174

Examples

```
str(Fatigue)
xyplot(relLength ~ cycles | Path, Fatigue, type = c("g", "b"),
       aspect = 'xy', xlab = "Number of test cycles (millions)",
       ylab = "Relative crack length (dimensionless)",
       layout = c(7,3))
```

Gasoline

Refinery yield of gasoline

Description

The Gasoline data frame has 32 rows and 6 columns.

Format

This data frame contains the following columns:

yield a numeric vector giving the percentage of crude oil converted to gasoline after distillation and fractionation

endpoint a numeric vector giving the temperature (degrees F) at which all the gasoline is vaporized

Sample the inferred crude oil sample number - a factor with levels A to J

API a numeric vector giving the crude oil gravity (degrees API)

vapor a numeric vector giving the vapor pressure of the crude oil (lbf/in²)

ASTM a numeric vector giving the crude oil 10% point ASTM—the temperature at which 10% of the crude oil has become vapor.

Details

Prater (1955) provides data on crude oil properties and gasoline yields. Atkinson (1985) uses these data to illustrate the use of diagnostics in multiple regression analysis. Three of the covariates—API, vapor, and ASTM—measure characteristics of the crude oil used to produce the gasoline. The other covariate — endpoint—is a characteristic of the refining process. Daniel and Wood (1980) notice that the covariates characterizing the crude oil occur in only ten distinct groups and conclude that the data represent responses measured on ten different crude oil samples.

Source

Prater, N. H. (1955), Estimate gasoline yields from crudes, *Petroleum Refiner*, **35** (5).
 Atkinson, A. C. (1985), *Plots, Transformations, and Regression*, Oxford Press, New York.
 Daniel, C. and Wood, F. S. (1980), *Fitting Equations to Data*, Wiley, New York
 Venables, W. N. and Ripley, B. D. (1999) *Modern Applied Statistics with S-PLUS (3rd ed)*, Springer, New York.

Examples

```
str(Gasoline)
xyplot(yield ~ endpoint | Sample, Gasoline, aspect = 'xy',
       main = "Gasoline data", xlab = "Endpoint (degrees F)",
       ylab = "Percentage yield",
       type = c("g", "p", "r"),
       index.cond = function(x,y) coef(lm(y~x))[2],
       layout = c(5,2))
print(m1 <- lmer(yield ~ endpoint + (1|Sample), Gasoline), corr = FALSE)
m2 <- lmer(yield ~ endpoint + (endpoint|Sample), Gasoline, verbose = 1)
print(m2)
Gasoline$endptC <- with(Gasoline, endpoint - mean(endpoint))
m3 <- lmer(yield ~ endpoint + (endptC|Sample), Gasoline, verbose = 1)
print(m3)
xyplot(endptC ~ '(Intercept)', ranef(m3)[[1]], type = c("g", "p", "r"),
       aspect = 1)
```

 Glucose

Glucose levels over time

Description

The Glucose data frame has 378 rows and 4 columns.

Format

This data frame contains the following columns:

Subject a factor with levels A to F

Time a numeric vector

conc a numeric vector of glucose levels

Meal an ordered factor with levels 2am < 6am < 10am < 2pm < 6pm < 10pm

Source

Hand, D. and Crowder, M. (1996), *Practical Longitudinal Data Analysis*, Chapman and Hall, London.

Examples

```
str(Glucose)
xyplot(conc ~ Time | Meal * Subject, Glucose)
```

 Glucose2

Glucose Levels Following Alcohol Ingestion

Description

The Glucose2 data frame has 196 rows and 4 columns.

Format

This data frame contains the following columns:

Subject a factor with levels A to G

Date a factor with levels 1 2 indicating the occasion in which the experiment was conducted.

Time a numeric vector giving the time since alcohol ingestion (in min/10).

glucose a numeric vector giving the blood glucose level (in mg/dl).

Details

Hand and Crowder (Table A.14, pp. 180-181, 1996) describe data on the blood glucose levels measured at 14 time points over 5 hours for 7 volunteers who took alcohol at time 0. The same experiment was repeated on a second date with the same subjects but with a dietary additive used for all subjects.

Source

Pinheiro, J. C. and Bates, D. M. (2000), *Mixed-Effects Models in S and S-PLUS*, Springer, New York. (Appendix A.10)

Hand, D. and Crowder, M. (1996), *Practical Longitudinal Data Analysis*, Chapman and Hall, London.

Examples

```
str(Glucose2)
xyplot(glucose ~ Time | Subject, Glucose2, type = c("g", "b"),
       groups = Date, aspect = 'xy', layout = c(4,2),
       index.cond = function(x,y) max(y))
```

Gun *Methods for firing naval guns*

Description

The Gun data frame has 36 rows and 4 columns.

Format

This data frame contains the following columns:

rounds a numeric vector

Method a factor with levels M1 and M2

Team an ordered factor with levels T1S < T3S < T2S < T1A < T2A < T3A < T1H < T3H < T2H

Physique an ordered factor with levels Slight < Average < Heavy

Details

Hicks (p.180, 1993) reports data from an experiment on methods for firing naval guns. Gunners of three different physiques (slight, average, and heavy) tested two firing methods. Both methods were tested twice by each of nine teams of three gunners with identical physique. The response was the number of rounds fired per minute.

Source

Hicks, C. R. (1993), *Fundamental Concepts in the Design of Experiments (4th ed)*, Harcourt Brace, New York.

Examples

```
str(Gun)
```

IGF *Radioimmunoassay of IGF-I Protein*

Description

The IGF data frame has 237 rows and 3 columns.

Format

This data frame contains the following columns:

Lot an ordered factor giving the radioactive tracer lot.

age a numeric vector giving the age (in days) of the radioactive tracer.

conc a numeric vector giving the estimated concentration of IGF-I protein (ng/ml)

Details

Davidian and Giltinan (1995) describe data obtained during quality control radioimmunoassays for ten different lots of radioactive tracer used to calibrate the Insulin-like Growth Factor (IGF-I) protein concentration measurements.

Source

Davidian, M. and Giltinan, D. M. (1995), *Nonlinear Models for Repeated Measurement Data*, Chapman and Hall, London.

Pinheiro, J. C. and Bates, D. M. (2000), *Mixed-Effects Models in S and S-PLUS*, Springer, New York. (Appendix A.11)

Examples

```
str(IGF)
```

Machines

Productivity Scores for Machines and Workers

Description

The Machines data frame has 54 rows and 3 columns.

Format

This data frame contains the following columns:

Worker an ordered factor giving the unique identifier for the worker.

Machine a factor with levels A, B, and C identifying the machine brand.

score a productivity score.

Details

Data on an experiment to compare three brands of machines used in an industrial process are presented in Milliken and Johnson (p. 285, 1992). Six workers were chosen randomly among the employees of a factory to operate each machine three times. The response is an overall productivity score taking into account the number and quality of components produced.

Source

Pinheiro, J. C. and Bates, D. M. (2000), *Mixed-Effects Models in S and S-PLUS*, Springer, New York. (Appendix A.14)

Milliken, G. A. and Johnson, D. E. (1992), *Analysis of Messy Data, Volume I: Designed Experiments*, Chapman and Hall, London.

Examples

```
str(Machines)
```

MathAchieve	<i>Mathematics achievement scores</i>
-------------	---------------------------------------

Description

The MathAchieve data frame has 7185 rows and 6 columns.

Format

This data frame contains the following columns:

School an ordered factor identifying the school that the student attends

Minority a factor with levels No Yes indicating if the student is a member of a minority racial group.

Sex a factor with levels Male Female

SES a numeric vector of socio-economic status.

MathAch a numeric vector of mathematics achievement scores.

MEANSES a numeric vector of the mean SES for the school.

Details

Each row in this data frame contains the data for one student.

Examples

```
str(MathAchieve)
```

MathAchSchool	<i>School demographic data for MathAchieve</i>
---------------	--

Description

The MathAchSchool data frame has 160 rows and 7 columns.

Format

This data frame contains the following columns:

School a factor giving the school on which the measurement is made.

Size a numeric vector giving the number of students in the school

Sector a factor with levels Public Catholic

PRACAD a numeric vector giving the percentage of students on the academic track

DISCLIM a numeric vector measuring the discrimination climate

HIMINTY a factor with levels 0 1

MEANSES a numeric vector giving the mean SES score.

Details

These variables give the school-level demographic data to accompany the MathAchieve data.

Examples

```
str(MathAchSchool)
```

Meat

Tenderness of meat

Description

The Meat data frame has 30 rows and 4 columns.

Format

This data frame contains the following columns:

Storage an ordered factor specifying the storage treatment - 1 (0 days), 2 (1 day), 3 (2 days), 4 (4 days), 5 (9 days), and 6 (18 days)

score a numeric vector giving the tenderness score of beef roast.

Block an ordered factor identifying the muscle from which the roast was extracted with levels II < V < I < III < IV

Pair an ordered factor giving the unique identifier for each pair of beef roasts with levels II-1 < ... < IV-1

Details

Cochran and Cox (section 11.51, 1957) describe data from an experiment conducted at Iowa State College (Paul, 1943) to compare the effects of length of cold storage on the tenderness of beef roasts. Six storage periods ranging from 0 to 18 days were used. Thirty roasts were scored by four judges on a scale from 0 to 10, with the score increasing with tenderness. The response was the sum of all four scores. Left and right roasts from the same animal were grouped into pairs, which were further grouped into five blocks, according to the muscle from which they were extracted. Different storage periods were applied to each roast within a pair according to a balanced incomplete block design.

Source

Cochran, W. G. and Cox, G. M. (1957), *Experimental Designs*, Wiley, New York.

Examples

```
str(Meat)
```

Milk	<i>Protein content of cows' milk</i>
------	--------------------------------------

Description

The Milk data frame has 1337 rows and 4 columns.

Format

This data frame contains the following columns:

protein a numeric vector giving the protein content of the milk.

Time a numeric vector giving the time since calving (weeks).

Cow an ordered factor giving a unique identifier for each cow.

Diet a factor with levels barley, barley+lupins, and lupins identifying the diet for each cow.

Details

Diggle, Liang, and Zeger (1994) describe data on the protein content of cows' milk in the weeks following calving. The cattle are grouped according to whether they are fed a diet with barley alone, with barley and lupins, or with lupins alone.

Source

Diggle, Peter J., Liang, Kung-Yee and Zeger, Scott L. (1994), *Analysis of longitudinal data*, Oxford University Press, Oxford.

Examples

```
str(Milk)
```

Muscle	<i>Contraction of heart muscle sections</i>
--------	---

Description

The Muscle data frame has 60 rows and 3 columns.

Format

This data frame contains the following columns:

Strip an ordered factor indicating the strip of muscle being measured.

conc a numeric vector giving the concentration of CaCl₂

length a numeric vector giving the shortening of the heart muscle strip.

Details

Baumann and Waldvogel (1963) describe data on the shortening of heart muscle strips dipped in a CaCl_2 solution. The muscle strips are taken from the left auricle of a rat's heart.

Source

Baumann, F. and Waldvogel, F. (1963), La restitution pastsystolique de la contraction de l'oreillette gauche du rat. Effets de divers ions et de l'acetylcholine, *Helvetica Physiologica Acta*, **21**.

Examples

```
str(Muscle)
```

Nitrendipene

Assay of nitrendipene

Description

The Nitrendipene data frame has 89 rows and 4 columns.

Format

This data frame contains the following columns:

activity a numeric vector

NIF a numeric vector

Tissue an ordered factor with levels 2 < 1 < 3 < 4

log.NIF a numeric vector

Source

Bates, D. M. and Watts, D. G. (1988), *Nonlinear Regression Analysis and Its Applications*, Wiley, New York.

Examples

```
str(Nitrendipene)
```

Oats

Split-plot Experiment on Varieties of Oats

Description

The Oats data frame has 72 rows and 4 columns.

Format

This data frame contains the following columns:

Block an ordered factor with levels VI < V < III < IV < II < I

Variety a factor with levels Golden Rain Marvellous Victory

nitro a numeric vector

yield a numeric vector

Details

These data have been introduced by Yates (1935) as an example of a split-plot design. The treatment structure used in the experiment was a 3×4 full factorial, with three varieties of oats and four concentrations of nitrogen. The experimental units were arranged into six blocks, each with three whole-plots subdivided into four subplots. The varieties of oats were assigned randomly to the whole-plots and the concentrations of nitrogen to the subplots. All four concentrations of nitrogen were used on each whole-plot.

Source

Pinheiro, J. C. and Bates, D. M. (2000), *Mixed-Effects Models in S and S-PLUS*, Springer, New York. (Appendix A.15)

Venables, W. N. and Ripley, B. D. (1999) *Modern Applied Statistics with S-PLUS (3rd ed)*, Springer, New York.

Examples

```
str(Oats)
```

 Orange

Growth of orange trees

Description

The Orange data frame has 35 rows and 3 columns of records of the growth of orange trees.

Usage

Orange

Format

This data frame contains the following columns:

Tree a factor indicating the tree on which the measurement is made.

age a numeric vector giving the age of the tree (days since 1968/12/31)

circumference a numeric vector of trunk circumferences (mm). This is probably “circumference at breast height”, a standard measurement in forestry.

Source

Draper, N. R. and Smith, H. (1998), *Applied Regression Analysis (3rd ed)*, Wiley (exercise 24.N).

Pinheiro, J. C. and Bates, D. M. (2000) *Mixed-effects Models in S and S-PLUS*, Springer.

Examples

```
xyplot(circumference ~ age, Orange, groups = Tree, type = c("g", "b"),
       auto.key = list(space = "right", lines = TRUE), aspect = "xy",
       xlab = "Age (days since 1968/12/31)", ylab = "Circumference (mm)")
## Not run:
m1 <- nlmer(circumference ~ SSlogis(age, Asym, xmid, scal) ~ Asym|Tree,
           Orange, verbose = TRUE,
           start = c(Asym = 190, xmid = 730, scal = 350))
.Call("mer_optimize", m1, 1L, 1L, PACKAGE = "lme4")
print(m1)
ranef(m1)

## End(Not run)
```

Orthodont

Growth curve data on an orthodontic measurement

Description

The Orthodont data frame has 108 rows and 4 columns of the change in an orthodontic measurement over time for several young subjects.

Format

This data frame contains the following columns:

distance a numeric vector of distances from the pituitary to the pterygomaxillary fissure (mm). These distances are measured on x-ray images of the skull.

age a numeric vector of ages of the subject (yr).

Subject an ordered factor indicating the subject on which the measurement was made. The levels are labelled M01 to M16 for the males and F01 to F13 for the females. The ordering is by increasing average distance within sex.

Sex a factor with levels Male and Female

Details

Investigators at the University of North Carolina Dental School followed the growth of 27 children (16 males, 11 females) from age 8 until age 14. Every two years they measured the distance between the pituitary and the pterygomaxillary fissure, two points that are easily identified on x-ray exposures of the side of the head.

Source

Pinheiro, J. C. and Bates, D. M. (2000), *Mixed-Effects Models in S and S-PLUS*, Springer, New York. (Appendix A.17)

Potthoff, R. F. and Roy, S. N. (1964), "A generalized multivariate analysis of variance model useful especially for growth curve problems", *Biometrika*, **51**, 313–326.

Examples

```
str(Orthodont)
```

Ovary

Counts of Ovarian Follicles

Description

The Ovary data frame has 308 rows and 3 columns.

Format

This data frame contains the following columns:

Mare an ordered factor indicating the mare on which the measurement is made.

Time time in the estrus cycle. The data were recorded daily from 3 days before ovulation until 3 days after the next ovulation. The measurement times for each mare are scaled so that the ovulations for each mare occur at times 0 and 1.

follicles the number of ovarian follicles greater than 10 mm in diameter.

Details

Pierson and Ginther (1987) report on a study of the number of large ovarian follicles detected in different mares at several times in their estrus cycles.

Source

Pinheiro, J. C. and Bates, D. M. (2000), *Mixed-Effects Models in S and S-PLUS*, Springer, New York. (Appendix A.18)

Pierson, R. A. and Ginther, O. J. (1987), Follicular population dynamics during the estrus cycle of the mare, *Animal Reproduction Science*, **14**, 219-231.

Examples

```
str(Ovary)
```

Oxide

Variability in Semiconductor Manufacturing

Description

The Oxide data frame has 72 rows and 5 columns.

Format

This data frame contains the following columns:

Source a factor with levels 1 and 2

Lot a factor giving a unique identifier for each lot.

Wafer a factor giving a unique identifier for each wafer within a lot.

Site a factor with levels 1, 2, and 3

Thickness a numeric vector giving the thickness of the oxide layer.

Details

These data are described in Littell et al. (1996, p. 155) as coming “from a passive data collection study in the semiconductor industry where the objective is to estimate the variance components to determine the assignable causes of the observed variability.” The observed response is the thickness of the oxide layer on silicon wafers, measured at three different sites of each of three wafers selected from each of eight lots sampled from the population of lots.

Source

Pinheiro, J. C. and Bates, D. M. (2000), *Mixed-Effects Models in S and S-PLUS*, Springer, New York. (Appendix A.20)

Littell, R. C., Milliken, G. A., Stroup, W. W. and Wolfinger, R. D. (1996), *SAS System for Mixed Models*, SAS Institute, Cary, NC.

Examples

```
str(Oxide)
```

 PBG

Effect of Phenylbiguanide on Blood Pressure

Description

The PBG data frame has 60 rows and 5 columns.

Format

This data frame contains the following columns:

deltaBP a numeric vector

dose a numeric vector

Run an ordered factor with levels T5 < T4 < T3 < T2 < T1 < P5 < P3 < P2 < P4 < P1

Treatment a factor with levels MDL 72222 P1acebo

Rabbit an ordered factor with levels 5 < 3 < 2 < 4 < 1

Details

Data on an experiment to examine the effect of a antagonist MDL 72222 on the change in blood pressure experienced with increasing dosage of phenylbiguanide are described in Ludbrook (1994) and analyzed in Venables and Ripley (1999, section 8.8). Each of five rabbits was exposed to increasing doses of phenylbiguanide after having either a placebo or the HD5-antagonist MDL 72222 administered.

Source

Pinheiro, J. C. and Bates, D. M. (2000), *Mixed-Effects Models in S and S-PLUS*, Springer, New York. (Appendix A.21)

Venables, W. N. and Ripley, B. D. (1999) *Modern Applied Statistics with S-PLUS (3rd ed)*, Springer, New York.

Ludbrook, J. (1994), Repeated measurements and multiple comparisons in cardiovascular research, *Cardiovascular Research*, **28**, 303-311.

Examples

```
str(PBG)
```

Phenobarb

Phenobarbital Kinetics

Description

The Phenobarb data frame has 744 rows and 7 columns.

Format

This data frame contains the following columns:

Subject an ordered factor identifying the infant.

Wt a numeric vector giving the birth weight of the infant (kg).

Apgar an ordered factor giving the the 5-minute Apgar score for the infant. This is an indication of health of the newborn infant.

ApgarInd a factor indicating whether the 5-minute Apgar score is < 5 or >= 5.

time a numeric vector giving the time when the sample is drawn or drug administered (hr).

dose a numeric vector giving the dose of drug administered (ug/kg).

conc a numeric vector giving the phenobarbital concentration in the serum (ug/L).

Details

Data from a pharmacokinetics study of phenobarbital in neonatal infants. During the first few days of life the infants receive multiple doses of phenobarbital for prevention of seizures. At irregular intervals blood samples are drawn and serum phenobarbital concentrations are determined. The data were originally given in Grasela and Donn(1985) and are analyzed in Boeckmann, Sheiner and Beal (1994), in Davidian and Giltinan (1995), and in Littell et al. (1996).

Source

Pinheiro, J. C. and Bates, D. M. (2000), *Mixed-Effects Models in S and S-PLUS*, Springer, New York. (Appendix A.23)

Davidian, M. and Giltinan, D. M. (1995), *Nonlinear Models for Repeated Measurement Data*, Chapman and Hall, London. (section 6.6)

Grasela and Donn (1985), Neonatal population pharmacokinetics of phenobarbital derived from routine clinical data, *Developmental Pharmacology and Therapeutics*, **8**, 374-383.

Boeckmann, A. J., Sheiner, L. B., and Beal, S. L. (1994), *NONMEM Users Guide: Part V*, University of California, San Francisco.

Littell, R. C., Milliken, G. A., Stroup, W. W. and Wolfinger, R. D. (1996), *SAS System for Mixed Models*, SAS Institute, Cary, NC.

Examples

```
str(Phenobarb)
```

Pixel	<i>X-ray pixel intensities over time</i>
-------	--

Description

The Pixel data frame has 102 rows and 4 columns of data on the pixel intensities of CT scans of dogs over time

Format

This data frame contains the following columns:

Dog a factor with levels A to J designating the dog on which the scan was made

Side a factor with levels L and R designating the side of the dog being scanned

day a numeric vector giving the day post injection of the contrast on which the scan was made

pixel a numeric vector of pixel intensities

Source

Pinheiro, J. C. and Bates, D. M. (2000) *Mixed-effects Models in S and S-PLUS*, Springer.

Examples

```
options(show.signif.stars = FALSE)
str(Pixel)
summary(Pixel)
(fm1 <- lmer(pixel ~ day + I(day^2) + (1|Dog:Side) + (day|Dog), Pixel))
```

Quinidine

Quinidine Kinetics

Description

The Quinidine data frame has 1471 rows and 14 columns.

Format

This data frame contains the following columns:

Subject a factor identifying the patient on whom the data were collected.

time a numeric vector giving the time (hr) at which the drug was administered or the blood sample drawn. This is measured from the time the patient entered the study.

conc a numeric vector giving the serum quinidine concentration (mg/L).

dose a numeric vector giving the dose of drug administered (mg). Although there were two different forms of quinidine administered, the doses were adjusted for differences in salt content by conversion to milligrams of quinidine base.

interval a numeric vector giving the when the drug has been given at regular intervals for a sufficiently long period of time to assume steady state behavior, the interval is recorded.

Age a numeric vector giving the age of the subject on entry to the study (yr).

Height a numeric vector giving the height of the subject on entry to the study (in.).

Weight a numeric vector giving the body weight of the subject (kg).

Race a factor with levels *Caucasian*, *Latin*, and *Black* identifying the race of the subject.

Smoke a factor with levels *no* and *yes* giving smoking status at the time of the measurement.

Ethanol a factor with levels *none*, *current*, *former* giving ethanol (alcohol) abuse status at the time of the measurement.

Heart a factor with levels *No/Mild*, *Moderate*, and *Severe* indicating congestive heart failure for the subject.

Creatinine an ordered factor with levels *< 50* *<=> 50* indicating the creatine clearance (mg/min).

glyco a numeric vector giving the alpha-1 acid glycoprotein concentration (mg/dL). Often measured at the same time as the quinidine concentration.

Details

Verme et al. (1992) analyze routine clinical data on patients receiving the drug quinidine as a treatment for cardiac arrhythmia (atrial fibrillation of ventricular arrhythmias). All patients were receiving oral quinidine doses. At irregular intervals blood samples were drawn and serum concentrations of quinidine were determined. These data are analyzed in several publications, including Davidian and Giltinan (1995, section 9.3).

Source

Pinheiro, J. C. and Bates, D. M. (2000), *Mixed-Effects Models in S and S-PLUS*, Springer, New York. (Appendix A.25)

Davidian, M. and Giltinan, D. M. (1995), *Nonlinear Models for Repeated Measurement Data*, Chapman and Hall, London.

Verme, C. N., Ludden, T. M., Clementi, W. A. and Harris, S. C. (1992), Pharmacokinetics of quinine in male patients: A population analysis, *Clinical Pharmacokinetics*, **22**, 468-480.

Examples

```
str(Quinidine)
```

 Rail

Evaluation of Stress in Railway Rails

Description

The Rail data frame has 18 rows and 2 columns.

Format

This data frame contains the following columns:

Rail an ordered factor identifying the rail on which the measurement was made.

travel a numeric vector giving the travel time for ultrasonic head-waves in the rail (nanoseconds). The value given is the original travel time minus 36,100 nanoseconds.

Details

Devore (2000, Example 10.10, p. 427) cites data from an article in *Materials Evaluation* on “a study of travel time for a certain type of wave that results from longitudinal stress of rails used for railroad track.”

Source

Pinheiro, J. C. and Bates, D. M. (2000), *Mixed-Effects Models in S and S-PLUS*, Springer, New York. (Appendix A.26)

Devore, J. L. (2000), *Probability and Statistics for Engineering and the Sciences (5th ed)*, Duxbury, Boston, MA.

Examples

```
str(Rail)
(fm1 <- lmer(travel ~ 1 | Rail, Rail))
```

RatPupWeight	<i>The weight of rat pups</i>
--------------	-------------------------------

Description

The RatPupWeight data frame has 322 rows and 5 columns.

Format

This data frame contains the following columns:

weight a numeric vector

sex a factor with levels Male Female

Litter a factor, the litter number

Lsize a numeric vector

Treatment an ordered factor with levels Control < Low < High

Source

Pinheiro, J. C. and Bates, D. M. (2000), *Mixed-Effects Models in S and S-PLUS*, Springer, New York.

Examples

```
str(RatPupWeight)
```

Relaxin	<i>Assay for Relaxin</i>
---------	--------------------------

Description

The Relaxin data frame has 198 rows and 3 columns.

Format

This data frame contains the following columns:

Run an ordered factor with levels 5 < 8 < 9 < 3 < 4 < 2 < 7 < 1 < 6

conc a numeric vector

cAMP a numeric vector

Source

Pinheiro, J. C. and Bates, D. M. (2000), *Mixed-Effects Models in S and S-PLUS*, Springer, New York.

Examples

```
str(Relaxin)
```

Remifentanil

Pharmacokinetics of remifentanil

Description

The Remifentanil data frame has 2107 rows and 12 columns.

Format

This data frame contains the following columns:

ID a numeric vector

Subject an ordered factor

Time a numeric vector

conc a numeric vector

Rate a numeric vector

Amt a numeric vector

Age a numeric vector

Sex a factor with levels Female Male

Ht a numeric vector

Wt a numeric vector

BSA a numeric vector

LBM a numeric vector

Source

Pinheiro, J. C. and Bates, D. M. (2000), *Mixed-Effects Models in S and S-PLUS*, Springer, New York.

Examples

```
str(Remifentanil)
```

Soybean

Growth of soybean plants

Description

The Soybean data frame has 412 rows and 5 columns.

Format

This data frame contains the following columns:

Plot a factor giving a unique identifier for each plot.

Variety a factor indicating the variety; Forrest (F) or Plant Introduction \#416937 (P).

Year a factor indicating the year the plot was planted.

Time a numeric vector giving the time the sample was taken (days after planting).

weight a numeric vector giving the average leaf weight per plant (g).

Details

These data are described in Davidian and Giltinan (1995, 1.1.3, p.7) as “Data from an experiment to compare growth patterns of two genotypes of soybeans: Plant Introduction \#416937 (P), an experimental strain, and Forrest (F), a commercial variety.”

Source

Pinheiro, J. C. and Bates, D. M. (2000), *Mixed-Effects Models in S and S-PLUS*, Springer, New York. (Appendix A.27)

Davidian, M. and Giltinan, D. M. (1995), *Nonlinear Models for Repeated Measurement Data*, Chapman and Hall, London.

Examples

```
str(Soybean)
#summary(fm1 <- nlsList(SSlogis, data = Soybean))
```


Source

Pinheiro, J. C. and Bates, D. M. (2000), *Mixed-Effects Models in S and S-PLUS*, Springer, New York.

Examples

```
str(Tetracycline1)
```

Tetracycline2	<i>Pharmacokinetics of tetracycline</i>
---------------	---

Description

The Tetracycline2 data frame has 40 rows and 4 columns.

Format

This data frame contains the following columns:

conc a numeric vector

Time a numeric vector

Subject an ordered factor with levels 4 < 5 < 2 < 1 < 3

Formulation a factor with levels Berkmycin tetramycin

Source

Pinheiro, J. C. and Bates, D. M. (2000), *Mixed-Effects Models in S and S-PLUS*, Springer, New York.

Examples

```
str(Tetracycline2)
```

Theoph	<i>Pharmacokinetics of theophylline</i>
--------	---

Description

The Theoph data frame has 132 rows and 5 columns of data from an experiment on the pharmacokinetics of theophylline.

Usage

Theoph

Format

This data frame contains the following columns:

Subject a factor with levels A, . . . , L identifying the subject on whom the observation was made.

Wt weight of the subject (kg).

Dose dose of theophylline administered orally to the subject (mg/kg).

Time time since drug administration when the sample was drawn (hr).

conc theophylline concentration in the sample (mg/L).

Details

Boeckmann, Sheiner and Beal (1994) report data from a study by Dr. Robert Upton of the kinetics of the anti-asthmatic drug theophylline. Twelve subjects were given oral doses of theophylline then serum concentrations were measured at 11 time points over the next 25 hours.

These data are analyzed in Davidian and Giltinan (1995) and Pinheiro and Bates (2000) using a two-compartment open pharmacokinetic model, for which a self-starting model function, `SSfol`, is available.

Source

Boeckmann, A. J., Sheiner, L. B. and Beal, S. L. (1994), *NONMEM Users Guide: Part V*, NONMEM Project Group, University of California, San Francisco.

Davidian, M. and Giltinan, D. M. (1995) *Nonlinear Models for Repeated Measurement Data*, Chapman & Hall (section 5.5, p. 145 and section 6.6, p. 176)

Pinheiro, J. C. and Bates, D. M. (2000) *Mixed-effects Models in S and S-PLUS*, Springer (Appendix A.29)

See Also

[SSfol](#)

Examples

```
xyplot(conc ~ Time | Subject, Theoph, aspect = 'xy',
       xlab = "Time since drug administration (hr)",
       ylab = "Theophylline concentration (mg/L)")
Theoph.D <- subset(Theoph, Subject == "D")
fm1 <- nls(conc ~ SSfol(Dose, Time, lKe, lKa, lCl),
          data = Theoph.D)
summary(fm1)
plot(conc ~ Time, data = Theoph.D,
     xlab = "Time since drug administration (hr)",
     ylab = "Theophylline concentration (mg/L)",
     main = "Observed concentrations and fitted model",
     sub = "Theophylline data - Subject 4 only",
     las = 1, col = 4)
xvals <- seq(0, par("usr")[2], len = 55)
lines(xvals, predict(fm1, newdata = list(Time = xvals)),
     col = 4)
```

Wafer *Modeling of Analog MOS Circuits*

Description

The Wafer data frame has 400 rows and 4 columns.

Format

This data frame contains the following columns:

Wafer a factor with levels 1 2 3 4 5 6 7 8 9 10

Site a factor with levels 1 2 3 4 5 6 7 8

voltage a numeric vector

current a numeric vector

Source

Pinheiro, J. C. and Bates, D. M. (2000), *Mixed-Effects Models in S and S-PLUS*, Springer, New York.

Examples

```
str(Wafer)
```

Wheat *Yields by growing conditions*

Description

The Wheat data frame has 48 rows and 4 columns.

Format

This data frame contains the following columns:

Tray an ordered factor with levels 3 < 1 < 2 < 4 < 5 < 6 < 8 < 9 < 7 < 12 < 11 < 10

Moisture a numeric vector

fertilizer a numeric vector

DryMatter a numeric vector

Source

Pinheiro, J. C. and Bates, D. M. (2000), *Mixed-Effects Models in S and S-PLUS*, Springer, New York.

Examples

```
str(Wheat)
```

Wheat2

Wheat Yield Trials

Description

The Wheat2 data frame has 224 rows and 5 columns.

Format

This data frame contains the following columns:

Block an ordered factor with levels 4 < 2 < 3 < 1

variety a factor with levels ARAPAHOE BRULE BUCKSKIN CENTURA CENTURK78 CHEYENNE CODY COLT GAGE HOMESTEAD KS831374 LANCER LANCOTA NE83404 NE83406 NE83407 NE83432 NE83498 NE83T12 NE84557 NE85556 NE85623 NE86482 NE86501 NE86503 NE86507 NE86509 NE86527 NE86582 NE86606 NE86607 NE86T666 NE87403 NE87408 NE87409 NE87446 NE87451 NE87457 NE87463 NE87499 NE87512 NE87513 NE87522 NE87612 NE87613 NE87615 NE87619 NE87627 NORKAN REDLAND ROUGHRIDER SCOUT66 SIOUXLAND TAM107 TAM200 VONA

yield a numeric vector

latitude a numeric vector

longitude a numeric vector

Source

Pinheiro, J. C. and Bates, D. M. (2000), *Mixed-Effects Models in S and S-PLUS*, Springer, New York.

Examples

```
str(Wheat2)
```

Index

*Topic **datasets**

Alfalfa, 2
Assay, 3
BodyWeight, 4
Cefamandole, 5
CO2, 6
Dialyzer, 7
Earthquake, 8
ergoStool, 9
Fatigue, 9
Gasoline, 10
Glucose, 11
Glucose2, 12
Gun, 13
IGF, 13
Machines, 14
MathAchieve, 15
MathAchSchool, 15
Meat, 16
Milk, 17
Muscle, 17
Nitrendipene, 18
Oats, 19
Orange, 20
Orthodont, 21
Ovary, 22
Oxide, 22
PBG, 23
Phenobarb, 24
Pixel, 25
Quinidine, 26
Rail, 27
RatPupWeight, 28
Relaxin, 28
Remifentanil, 29
Soybean, 30
Spruce, 31
Tetracycline1, 31
Tetracycline2, 32
Theoph, 32
Wafer, 34
Wheat, 34
Wheat2, 35

Alfalfa, 2
Assay, 3

BodyWeight, 4

Cefamandole, 5
CO2, 6

Dialyzer, 7

Earthquake, 8
ergoStool, 9

Fatigue, 9

Gasoline, 10
Glucose, 11
Glucose2, 12
Gun, 13

IGF, 13

Machines, 14
MathAchieve, 15
MathAchSchool, 15
Meat, 16
Milk, 17
Muscle, 17

Nitrendipene, 18

Oats, 19
Orange, 20
Orthodont, 21
Ovary, 22
Oxide, 22

PBG, [23](#)
Phenobarb, [24](#)
Pixel, [25](#)

Quinidine, [26](#)

Rail, [27](#)
RatPupWeight, [28](#)
Relaxin, [28](#)
Remifentanyl, [29](#)

Soybean, [30](#)
Spruce, [31](#)
SSfol, [33](#)

Tetracycline1, [31](#)
Tetracycline2, [32](#)
Theoph, [32](#)

Wafer, [34](#)
Wheat, [34](#)
Wheat2, [35](#)