Package: Fahrmeir (via r-universe)

November 13, 2026

Version 2016.5.31.9001

Title Data from the Book ``Multivariate Statistical Modelling Based on Generalized Linear Models'', First Edition, by Ludwig Fahrmeir and Gerhard Tutz

Author Kjetil B Halvorsen [aut, com, cre]

Maintainer Kjetil B Halvorsen <kjetil1001@gmail.com>

Depends stats, R (>= 2.1.0)

Suggests MASS

LazyData TRUE

Description Data and functions for the book ``Multivariate Statistical Modelling Based on Generalized Linear Models'', first edition, by Ludwig Fahrmeir and Gerhard Tutz. Useful when using the book. Compatibility updates by Renova.

License GPL (>= 2)

NeedsCompilation no

Repository https://renova-r.r-universe.dev

Date/Publication 2026-11-13 16:11:42 UTC

RemoteUrl https://github.com/renova-r/Fahrmeir

RemoteRef renova-2016.5.31.9001

RemoteSha 73c72111a1543bc97d3de106804594523f44b463

Contents

breath 2
caesar 3
cells 4
credit 5
happy 6
headneck 7
ohio 7
Regensburg 8

rheuma 9
tonsil 10
visual 11
wine 11

Index **13**

`breath` *Breathing Test*

Description

Effects of age and smoking status on breathing test results for workers in industrial plants in Texas.

Usage

```
data(breath)
```

Format

A data frame with 18 observations on the following 4 variables.

Age a factor with levels `<40` `40-59`

n number of workers in group

Smoking.status a factor with levels `Current.smoker` `Former.smoker` `Never.smoked`

Breathing.test a factor with levels `Abnormal` `Borderline` `Normal`

Details

We consider the effects of age and smoking status upon breathing test results for workers in industrial plants in Texas. The test results are given on an ordered scale with categories "Abnormal", "Borderline" and "Normal". It is of interest how age and smoking status are connected to breathing test results.

Source

Ludwig Fahrmeir, Gerhard Tutz (1994): Multivariate Statistical Modelling Based on Generalized Linear Models. Springer Series in Statistics. Springer Verlag. New-York Berlin Heidelberg

Examples

```
str(breath)
breath$Breathing.test <- ordered(breath$Breathing.test)
library(MASS)
breath.polr1 <- polr(Breathing.test ~ Age*Smoking.status, weight=n,
                     data=breath)
breath.polr2 <- polr(Breathing.test ~ Age*Smoking.status, weight=n,
                     data=breath, method="cloglog")
summary(breath.polr1)
```

```
summary(breath.polr2)
# continuation ratio models (as of page 89) might be fitted with
# Design or VGAM package.
```

`caesar` *Caesarian Birth Study*

Description

Data on infection from births by Caesarian section

Usage

```
data(caesar)
```

Format

A data frame with 24 observations on the following 7 variables.

y a factor with levels `1` `2` `3`, the response

w number of patients in group

noplan a factor with levels `not planned`, was the caesarian planned?

factor a factor with levels `risk factors without`, was there risk factors?

antib a factor with levels `antibiotics without`

yl logistic response, 0=no infection

patco covariate pattern number

Details

Infection from birth by Caesarian section. The response variable, `y`, has levels 1=type I infection, 2=type II infection, 3=none infection. Where risk-factors (diabetes, overweight, others) present? Where antibiotics used as prophylaxis? Aim is to analyse effects on response by covariates.

Author(s)

Kjetil Halvorsen

Source

Ludwig Fahrmeir, Gerhard Tutz (1994): Multivariate Statistical Modelling Based on Generalized Linear Models. Springer Series in Statistics. Springer Verlag. New-York Berlin Heidelberg

Examples

```
summary(caesar)
caesar.glm1 <- glm(yl ~ noplan+factor+antib, data=caesar, weight=w,
                       family=binomial(link="logit"))
caesar.glm2 <- glm(yl ~ noplan+factor+antib, data=caesar, weight=w,
                    family=binomial(link="probit"))
summary(caesar.glm1)
summary(caesar.glm2)
```

`cells` *Cellular Differentiation*

Description

The effect of two agents of immuno-activating ability that may induce cell differentiation was investigated.

Usage

```
data(cells)
```

Format

A data frame with 16 observations on the following 3 variables.

y number of cells differentiating

TNF dose of TNF, U/ml

IFN dose of IFN, U/ml

Details

The effect of two agents of immuno-activating ability that may induce cell differentiation was investigated. As response variable the number of cells that exhibited markers after exposure was recorded. It is of interest if the agents TNF (tumor necrosis factor) and IFN (interferon) stimulate cell differentiation independently, or if there is a synergetic effect. 200 cells were examined at each dose combination.

Author(s)

Kjetil Halvorsen

Source

Ludwig Fahrmeir, Gerhard Tutz (1994): Multivariate Statistical Modelling Based on Generalized Linear Models. Springer Series in Statistics. Springer Verlag. New-York Berlin Heidelberg

Examples

```
str(cells)
cells.poisson <- glm(y~TNF+IFN+TNF:IFN, data=cells,
                     family=poisson)
summary(cells.poisson)
confint(cells.poisson)
# Now we follow the book, example 2.6, page 51:
# there seems to be overdispersion?
cells.quasi <- glm(y~TNF+IFN+TNF:IFN, data=cells,
                     family=quasipoisson)
summary(cells.quasi)
anova(cells.quasi)
confint(cells.quasi)
# We follow the book, example 2.7, page 56:
with(cells, tapply(y, factor(TNF), function(x) c(mean(x), var(x))))
# which might indicate the use of a negative binomial model
```

`credit` *Credit Score Data From a South German Bank*

Description

The `credit` data frame has 1000 rows and 8 columns. This are data for 1000 clients of a south german bank, 700 good payers and 300 bad payers. They are used to construct a credit scoring method.

Usage

```
data(credit)
```

Format

This data frame contains the following columns:

Y a factor with levels `buen` `mal`, the response variable. buen is the good payers.

Cuenta a factor with levels `no good running` `bad running`, quality of the credit clients bank account.

Mes a numeric vector, duration of loan in months.

Ppag a factor with levels `pre buen pagador` `pre mal pagador`, if the client previosly have been a good or bad payer.

Uso a factor with levels `privado` `profesional`, the use to which the loan is made.

DM a numeric vector, the size of loan in german marks.

Sexo a factor with levels `mujer` `hombre`, sex of the client.

Estc a factor with levels `no vive solo` `vive solo`, civil state of the client.

Source

Fahrmeier and Tutz, *Multivariate Generalized Linear Models*, Springer Verlag.

Examples

```
summary(credit)
```

---

happy *Reported Happiness*

Description

Relationship between sex, years in school, and reported happiness.

Usage

```
data(happy)
```

Format

A data frame with 24 observations on the following 4 variables.

Rep.happiness an ordered factor with levels `Not to happy` < `Pretty happy` < `Very happy`

School a factor with levels `<12` `>16` `12` `13-16`

Sex a factor with levels `Females` `Males`

n number of persons in group

Source

Ludwig Fahrmeir, Gerhard Tutz (1994): Multivariate Statistical Modelling Based on Generalized Linear Models. Springer Series in Statistics. Springer Verlag. New-York Berlin Heidelberg

Examples

```
str(happy)
table(happy)
```

headneck *Head and Neck Cancer data*

Description

Data from a head and neck cancer study where time was discretized by one-month intervals.

Usage

```
data(headneck)
```

Format

A data frame with 47 observations on the following 4 variables.

month a numeric vector

atrisk a numeric vector, number at risk

deaths a numeric vector

withdrawals a numeric vector

Source

Ludwig Fahrmeir, Gerhard Tutz (1994): Multivariate Statistical Modelling Based on Generalized Linear Models. Springer Series in Statistics. Springer Verlag. New-York Berlin Heidelberg

Examples

```
str(headneck)
summary(headneck)
with(headneck, {plot(month, atrisk, type="s");
                lines(month, deaths, type="s", col="red");
                lines(month, withdrawals, type="S", col="green")})
```

ohio *Air Pollution and Health*

Description

Air Pollution and Health, annual data on children 7 to ten years old in Ohio.

Usage

```
data(ohio)
```

Format

A data frame with 32 observations on the following 6 variables.

a7 Presence (1) or absence (0) of respiratory infection

a8 Presence (1) or absence (0) of respiratory infection

a9 Presence (1) or absence (0) of respiratory infection

a10 Presence (1) or absence (0) of respiratory infection

mother.smoke a factor with levels `no` `yes`

n number of children

Details

Within the harvard Study of Air Pollution and Health, 537 children were examined annually from age 7 to 10, on the presence or absence of respiratory infection. So there are four repeated measurements on each child, or "short time series". The only available covariate is mothers smoking status at start of study.

Source

Ludwig Fahrmeir, Gerhard Tutz (1994): Multivariate Statistical Modelling Based on Generalized Linear Models. Springer Series in Statistics. Springer Verlag. New-York Berlin Heidelberg

Examples

```
str(ohio)
summary(ohio)
```

Regensburg *Job Expectation*

Description

A sample of psychology students was asked if they expected to find adecuate employment after graduation.

Usage

```
data(Regensburg)
```

Format

A data frame with 30 observations on the following 4 variables.

y response categories

n number of students with this response in group

age age in years

lage natural log of age

Details

In a study on the perspectives of students, psychology students at the university of Regensburg have been asked if they expect to find an adequate employment after getting their degree. The response categories where ordered with respect to their expectation. Categories where "don't expect adequate employment" - 1, "not sure" - 2, "immediately after the degree" - 3.

Source

Ludwig Fahrmeir, Gerhard Tutz (1994): Multivariate Statistical Modelling Based on Generalized Linear Models. Springer Series in Statistics. Springer Verlag. New-York Berlin Heidelberg

Examples

```
str(Regensburg)
summary(Regensburg)
# Example 3.5 page 83 in book:
library(MASS)
Regensburg$y <- ordered(Regensburg$y)
Regensburg.polr <- polr(y~lage, data=Regensburg, weights = n)
summary(Regensburg.polr)
class(Regensburg.polr)
```

---

rheuma *Data from Patients with Acute Rheumatoid Arthritis*

Description

Data from patients with acute rheumathoid arthritis. A new agent was compared with an active control, and each patient was evaluated on a five-point assessment scale.

Usage

```
data(rheuma)
```

Format

A data frame with 10 observations on the following 3 variables.

Drug a factor with levels `Active.control` `New.agent`

Improvement an ordered factor with levels `Much.worse` < `Worse` < `No.change` < `Improved` < `Much.improved`

n number of patients in group

Details

The global assessment in this example may be subdivided in the coarse response "improvement", "no change" and "worse". On a higher level improvement is split into "much improved" and "improved", while the "worse" category is split into "worse" and "much worse".

Source

Ludwig Fahrmeir, Gerhard Tutz (1994): Multivariate Statistical Modelling Based on Generalized Linear Models. Springer Series in Statistics. Springer Verlag. New-York Berlin Heidelberg

Examples

```
str(rheuma)
summary(rheuma)
```

---

tonsil *Data Set of Tonsil Size in Children*

Description

Children have been classified according to their relative tonsil size and wheater or not they are carriers of Streptococcus pyogenes.

Usage

```
data(tonsil)
```

Format

A data frame with 6 observations on the following 3 variables.

Streptococcus.p a factor with levels `carriers` `noncarriers`

Size numeric, 1, 2 or 3, tonsil size

n number of children in group

Details

It may be assumed that tonsil size always starts in the normal state "present but not enlarged" (category 1). If the tonsils grow abnormally, they may become "enlarged" (category 2), if the process does not stop, they may become "greatly enlarged" (category 3).

Source

Ludwig Fahrmeir, Gerhard Tutz (1994): Multivariate Statistical Modelling Based on Generalized Linear Models. Springer Series in Statistics. Springer Verlag. New-York Berlin Heidelberg

Examples

```
str(tonsil)
summary(tonsil)
```

visual *Visual Impairment Data*

Description

For 5199 individuals bivariate binary responses were observed, indicating wheater or not an eye was visually impaired, with covariates. The main objective is to analyze the influence of age and race on visual impairment, controlling for education, a surrogate for socioeconomic status. Data are only given individually for right and left eye, the bivariate response is lost.

Usage

```
data(visual)
```

Format

The format is: List of 2 $ left :'data.frame': 16 obs. of 4 variables: ..$ left: Factor w/ 2 levels "no","yes": 2 1 2 1 2 1 2 1 2 1 ... ..$ race: Factor w/ 2 levels "black","white": 2 2 2 2 2 2 2 2 1 1 ... ..$ age : Factor w/ 4 levels "40-50","51-60",..: 1 1 2 2 3 3 4 4 1 1 ... ..$ n : int [1:16] 15 617 24 557 42 789 139 673 29 750 ... $ right:'data.frame': 16 obs. of 4 variables: ..$ right: Factor w/ 2 levels "no","yes": 2 1 2 1 2 1 2 1 2 1 ... ..$ race : Factor w/ 2 levels "black","white": 2 2 2 2 2 2 2 2 1 1 ... ..$ age : Factor w/ 4 levels "40-50","51-60",..: 1 1 2 2 3 3 4 4 1 1 ... ..$ n : int [1:16] 19 613 25 556 48 783 146 666 31 748 ...

Source

Ludwig Fahrmeir, Gerhard Tutz (1994): Multivariate Statistical Modelling Based on Generalized Linear Models. Springer Series in Statistics. Springer Verlag. New-York Berlin Heidelberg

Examples

```
str(visual)
summary(visual)
```

wine *Bitterness of White Wines*

Description

In a study on the bitterness of white wine it is of interest wheater treatments that can be controlleds during pressing the grapes influence the bitterness of wines. The two factors considered are the temperature and the admission of contact with skin when pressing the grapes.

Usage

```
data(wine)
```

Format

A data frame with 72 observations on the following 5 variables.

temp a factor, temperature, with levels `high` `low`

contact a factor with levels `no` `yes`

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

judge a factor with levels `1` `2` `3` `4` `5` `6` `7` `8` `9`

score numeric, ordinal score, from '1'=nonbitter to '5'=very bitter

Source

Ludwig Fahrmeir, Gerhard Tutz (1994): Multivariate Statistical Modelling Based on Generalized Linear Models. Springer Series in Statistics. Springer Verlag. New-York Berlin Heidelberg

Examples

```
str(wine)
summary(wine)
```

Index

∗ **bivariate logistic**
visual, 11
∗ **cells**
cells, 4
∗ **compound cumulative model**
rheuma, 9
∗ **continuation ratio**
breath, 2
∗ **credit scoring**
credit, 5
∗ **cumulative logistic model**
Regensburg, 8
∗ **datasets**
breath, 2
caesar, 3
cells, 4
credit, 5
happy, 6
headneck, 7
ohio, 7
Regensburg, 8
rheuma, 9
tonsil, 10
visual, 11
wine, 11
∗ **negative binomial**
cells, 4
∗ **ordered categorical response**
breath, 2
∗ **ordinal response**
Regensburg, 8
∗ **poisson**
cells, 4
∗ **proportional odds**
Regensburg, 8
∗ **repeated measurement**
ohio, 7
∗ **short time series**
ohio, 7
∗ **two-step cumulative model**
rheuma, 9

breath, 2

caesar, 3
cells, 4
credit, 5

happy, 6
headneck, 7

ohio, 7

Regensburg, 8
rheuma, 9

tonsil, 10

visual, 11

wine, 11