

# Package ‘extendedFamily’

September 24, 2023

**Type** Package

**Title** Additional Families for Generalized Linear Models

**Version** 0.2.3

**Author** Greg McMahan

**Maintainer** Greg McMahan <gmcmacran@gmail.com>

**Description** Creates family objects identical to stats family but for new links.

**License** GPL-3

**Encoding** UTF-8

**LazyData** true

**RoxygenNote** 7.2.3

**Suggests** testthat, covr, numDeriv (>= 2016.8-1)

**Imports** stats (>= 3.5.3), assertthat (>= 0.2.1)

**Depends** R (>= 3.5.0)

**NeedsCompilation** no

**Repository** CRAN

**Date/Publication** 2023-09-24 16:50:02 UTC

## R topics documented:

binomialEF . . . . .	2
heart . . . . .	3

<b>Index</b>	<b>5</b>
--------------	----------

**Description**

Additional Binomial Links For Generalized Linear Models

**Usage**

```
binomialEF(link = "loglog", alpha = 1)
```

**Arguments**

link	name of link function. One of loglog, logc, identity, or odds-power (Default: loglog)
alpha	power for odds-power link. Not used otherwise. (Default: 1)

**Details**

family is a generic function with methods for classes "glm" and "lm".

The loglog link works well for many datasets. The range of the link is negative infinity to positive infinity. For all other links, this is not true. This can cause a failure to converge in R's glm function. If this happens, the link does not work well for the training data. Try another link.

**Value**

An object of class "family" (which has a concise print method). This is a list with elements

- family: character: the family name.
- link: character: the link name.
- linkfun: function: the link.
- linkinv: function: the inverse of the link function.
- variance: function: the variance as a function of the mean.
- dev.resids function giving the deviance for each observation as a function of (y, mu, wt), used by the **residuals** method when computing deviance residuals.
- aic: function giving the AIC value if appropriate (but NA for the quasi- families). More precisely, this function returns  $-2 \ell + 2 s$ , where  $\ell$  is the log-likelihood and  $s$  is the number of estimated scale parameters. Note that the penalty term for the location parameters (typically the "regression coefficients") is added elsewhere, e.g., in **glm.fit()** or **AIC()**, see the AIC example in **glm**. See **logLik** for the assumptions made about the dispersion parameter.
- initialize: expression. This needs to set up whatever data objects are needed for the family as well as n (needed for AIC in the binomial family) and mustart (See **glm**)
- validmu: logical function. Returns TRUE if a mean vector mu is within the domain of variance.

- `valideta`: logical function. Returns TRUE if a linear predictor  $\eta$  is within the domain of `linkinv`.
- `simulate`: (optional) function `simulate(object, nsim)` to be called by the "lm" method of `simulate`. It will normally return a matrix with `nsim` columns and one row for each fitted value, but it can also return a list of length `nsim`. Clearly this will be missing for 'quasi-' families.
- `dispersion`: (optional for R  $\geq$  4.3.0) numeric: value of the dispersion parameter, if fixed, or `NA_real_` if free.

### Examples

```
library(stats)
library(extendedFamily)

# loglog example
data(heart)
model <- glm(
  formula = death ~ anterior + hcabg +
    kk2 + kk3 + kk4 + age2 + age3 + age4,
  data = heart,
  family = binomialEF(link = "loglog")
)
```

---

heart

*Heart Attack Data*

---

### Description

A dataset containing 4,483 rows about heart attack victims.

### Usage

```
data(heart)
```

### Format

A data frame with 4,483 rows and 11 variables:

**death** 1 - Death within 48 hours of myocardial infarction onset. 0 - otherwise.

**anterior** 1 - anterior infarction. 0 - inferior infarction

**hcabg** 1 - Medical history of cardiac bypass surgery. 0 - otherwise

**kk1** 1 - Killip class 1. 0 - otherwise

**kk2** 1 - Killip class 2. 0 - otherwise

**kk3** 1 - Killip class 3. 0 - otherwise

**kk4** 1 - Killip class 4. 0 - otherwise

**age1** 1 - Patient age less than 60. 0 - otherwise

**age2** 1 - Patient age between 60 and 69 inclusive. 0 - otherwise

**age3** 1 - Patient age between 70 and 79 inclusive. 0 - otherwise

**age4** 1 - Patient age 80 or higher. 0 - otherwise

**Details**

The raw dataset is minimally modified in this package. Rows containing missing values were removed. The center variable was dropped. This data is referred to as "heart01" in Generalized Linear Models and Extensions book.

**Source**

Hardin, James William., and Joseph Hilbe. Generalized Linear Models and Extensions. Stata Press, 2012.

Dataset can be downloaded at <https://www.stata-press.com/data/glmext.html>

# Index

\* **datasets**

heart, 3

binomialEF, 2

heart, 3