

# Package ‘tram’

November 2, 2022

**Title** Transformation Models

**Version** 0.8-0

**Date** 2022-11-02

**Description** Formula-based user-interfaces to specific transformation models implemented in package 'mlt'. Available models include Cox models, some parametric survival models (Weibull, etc.), models for ordered categorical variables, normal and non-normal (Box-Cox type) linear models, and continuous outcome logistic regression (Lohse et al., 2017, <[DOI:10.12688/f1000research.12934.1](https://doi.org/10.12688/f1000research.12934.1)>). The underlying theory is described in Hothorn et al. (2018) <[DOI:10.1111/sjos.12291](https://doi.org/10.1111/sjos.12291)>. An extension to transformation models for clustered data is provided (Barbanti and Hothorn, 2022, <[arxiv:1910.09219](https://arxiv.org/abs/1910.09219)>). Multivariate conditional transformation models (Klein et al, 2022, <[DOI:10.1111/sjos.12501](https://doi.org/10.1111/sjos.12501)>) can be fitted as well.

**Depends** R (>= 3.5.0), mlt (>= 1.4-3)

**Imports** Formula, multcomp, variables (>= 1.0-4), basefun (>= 1.1-2), sandwich, stats, survival, graphics, Matrix, methods, mvtnorm

**Suggests** MASS, TH.data, trtf (>= 0.3-3), mlbench, knitr, quantreg, colorspace, ATR, lme4, merDeriv, SparseGrid, alabama, numDeriv, gridExtra, lattice, latticeExtra, HSAUR3, ordinalCont, coxme, mlt.docreg, ordinal, coin, asht, gamlss

**VignetteBuilder** knitr

**URL** <http://ctm.R-forge.R-project.org>

**Encoding** UTF-8

**License** GPL-2

**NeedsCompilation** yes

**Author** Torsten Hothorn [aut, cre] (<<https://orcid.org/0000-0001-8301-0471>>),  
Luisa Barbanti [aut] (<<https://orcid.org/0000-0001-5352-5802>>),  
Sandra Siegfried [aut] (<<https://orcid.org/0000-0002-7312-1001>>),  
Brian Ripley [ctb],  
Bill Venables [ctb],  
Douglas M. Bates [ctb],  
Nadja Klein [ctb]

**Maintainer** Torsten Hothorn <Torsten.Hothorn@R-project.org>

**Repository** CRAN

**Date/Publication** 2022-11-02 09:20:02 UTC

## R topics documented:

Aareg . . . . .	2
BoxCox . . . . .	4
Colr . . . . .	5
Coxph . . . . .	7
Lehmann . . . . .	8
Lm . . . . .	10
ltmatrices . . . . .	11
mm1t . . . . .	12
mtram . . . . .	15
perm_test . . . . .	16
Polr . . . . .	18
score_test . . . . .	20
Survreg . . . . .	21
tram . . . . .	22
tram-methods . . . . .	25
<b>Index</b>	<b>31</b>

---

Aareg	<i>Aalen Additive Hazards Model</i>
-------	-------------------------------------

---

### Description

Aalen model with fully parameterised hazard function

### Usage

```
Aareg(formula, data, subset, weights, offset, cluster, na.action = na.omit, ...)
```

### Arguments

formula	an object of class "formula": a symbolic description of the model structure to be fitted. The details of model specification are given under <a href="#">tram</a> and in the package vignette.
data	an optional data frame, list or environment (or object coercible by <code>as.data.frame</code> to a data frame) containing the variables in the model. If not found in data, the variables are taken from <code>environment(formula)</code> .
subset	an optional vector specifying a subset of observations to be used in the fitting process.

weights	an optional vector of weights to be used in the fitting process. Should be NULL or a numeric vector. If present, the weighted log-likelihood is maximised.
offset	this can be used to specify an <code>_a priori_</code> known component to be included in the linear predictor during fitting. This should be NULL or a numeric vector of length equal to the number of cases.
cluster	optional factor with a cluster ID employed for computing clustered covariances.
na.action	a function which indicates what should happen when the data contain NAs. The default is set to <code>na.omit</code> .
...	additional arguments to <code>tram</code> .

### Details

This function allows simultaneous estimation of the cumulative hazard parameterised by a Bernstein polynomial. The model is typically fitted with time-varying coefficients, all types of random censoring and truncation are allowed.

The responses is bounded (`bounds = c(0, Inf)`) when specified as a `Surv` object. Otherwise, bounds can be specified via ...

### Value

An object of class `Aareg`, with corresponding `coef`, `vcov`, `logLik`, `estfun`, `summary`, `print`, `plot` and `predict` methods.

### References

Torsten Hothorn, Lisa Moest, Peter Buehlmann (2018), Most Likely Transformations, *Scandinavian Journal of Statistics*, **45**(1), 110–134, doi:[10.1111/sjos.12291](https://doi.org/10.1111/sjos.12291).

### Examples

```
data("GBSG2", package = "TH.data")
library("survival")
GBSG2$time <- as.numeric(GBSG2$time)
GBSG2$y <- with(GBSG2, Surv(time, cens))

### Cox proportional hazards model
m1 <- Coxph(y ~ horTh, data = GBSG2, support = c(1, 1500))
logLik(m1)

### Aalen additive hazards model with time-varying effects
m2 <- Aareg(y | horTh ~ 1, data = GBSG2, support = c(1, 1500))
logLik(m2)

### compare the hazard functions
nd <- data.frame(horTh = unique(GBSG2$horTh))
col <- 1:2
lty <- 1:2
plot(as.mlt(m1), newdata = nd, type = "hazard",
```

```

col = col, lty = lty[1], xlab = "time")
plot(as.mlt(m2), newdata = nd, type = "hazard",
     col = col, lty = 2, add = TRUE)
legend("topright", col = rep(col, each = 2),
      lty = rep(1:2), bty = "n",
      legend = paste(rep(paste("horTh:",
                              levels(nd$horTh)), each = 2),
                    rep(c("Cox", "Aalen"), 2)))

```

---

BoxCox

*(Similar to) Box-Cox Models*


---

### Description

Non-normal linear regression inspired by Box-Cox models

### Usage

```
BoxCox(formula, data, subset, weights, offset, cluster, na.action = na.omit, ...)
```

### Arguments

formula	an object of class "formula": a symbolic description of the model structure to be fitted. The details of model specification are given under <a href="#">tram</a> and in the package vignette.
data	an optional data frame, list or environment (or object coercible by <code>as.data.frame</code> to a data frame) containing the variables in the model. If not found in data, the variables are taken from <code>environment(formula)</code> .
subset	an optional vector specifying a subset of observations to be used in the fitting process.
weights	an optional vector of weights to be used in the fitting process. Should be <code>NULL</code> or a numeric vector. If present, the weighted log-likelihood is maximised.
offset	this can be used to specify an <code>_a priori_</code> known component to be included in the linear predictor during fitting. This should be <code>NULL</code> or a numeric vector of length equal to the number of cases.
cluster	optional factor with a cluster ID employed for computing clustered covariances.
na.action	a function which indicates what should happen when the data contain NAs. The default is set to <code>na.omit</code> .
...	additional arguments to <a href="#">tram</a> .

### Details

A normal model for transformed responses, where the transformation is estimated from the data simultaneously with the regression coefficients. This is similar to a Box-Cox transformation, but the technical details differ. Examples can be found in the package vignette.

The model is defined with a negative shift term. Large values of the linear predictor correspond to large values of the conditional expectation response (but this relationship is potentially nonlinear).

**Value**

An object of class `BoxCox`, with corresponding `coef`, `vcov`, `logLik`, `estfun`, `summary`, `print`, `plot` and `predict` methods.

**References**

Torsten Hothorn, Lisa Moest, Peter Buehlmann (2018), Most Likely Transformations, *Scandinavian Journal of Statistics*, **45**(1), 110–134, doi:[10.1111/sjos.12291](https://doi.org/10.1111/sjos.12291).

**Examples**

```
data("BostonHousing2", package = "mlbench")

lm(cmedv ~ crim + zn + indus + chas + nox + rm + age + dis +
    rad + tax + ptratio + b + lstat, data = BostonHousing2)

BoxCox(cmedv ~ chas + crim + zn + indus + nox +
    rm + age + dis + rad + tax + ptratio + b + lstat,
    data = BostonHousing2)
```

---

 Colr

*Continuous Outcome Logistic Regression*


---

**Description**

A proportional-odds model for continuous variables

**Usage**

```
Colr(formula, data, subset, weights, offset, cluster, na.action = na.omit, ...)
```

**Arguments**

<code>formula</code>	an object of class "formula": a symbolic description of the model structure to be fitted. The details of model specification are given under <code>tram</code> and in the package vignette.
<code>data</code>	an optional data frame, list or environment (or object coercible by <code>as.data.frame</code> to a data frame) containing the variables in the model. If not found in <code>data</code> , the variables are taken from <code>environment(formula)</code> .
<code>subset</code>	an optional vector specifying a subset of observations to be used in the fitting process.
<code>weights</code>	an optional vector of weights to be used in the fitting process. Should be <code>NULL</code> or a numeric vector. If present, the weighted log-likelihood is maximised.
<code>offset</code>	this can be used to specify an <code>_a priori_</code> known component to be included in the linear predictor during fitting. This should be <code>NULL</code> or a numeric vector of length equal to the number of cases.

cluster	optional factor with a cluster ID employed for computing clustered covariances.
na.action	a function which indicates what should happen when the data contain NAs. The default is set by the na.action setting of options, and is na.fail if that is unset.
...	additional arguments to <code>tram</code> .

## Details

Simultaneous estimation of all possible binary logistic models obtained by dichotomisation of a continuous response. The regression coefficients can be constant allowing for an interpretation as log-odds ratios.

The model is defined with a positive shift term, thus  $\exp(\text{coef}())$  is the multiplicative change of the odds ratio (conditional odds of treatment or for a one unit increase in a numeric variable divided by conditional odds of reference). Large values of the linear predictor correspond to small values of the conditional expectation response (but this relationship is nonlinear).

## Value

An object of class `Colr`, with corresponding `coef`, `vcov`, `logLik`, `estfun`, `summary`, `print`, `plot` and `predict` methods.

## References

Tina Lohse, Sabine Rohrmann, David Faeh and Torsten Hothorn (2017), Continuous Outcome Logistic Regression for Analyzing Body Mass Index Distributions, *F1000Research*, **6**(1933), doi:10.12688/f1000research.12934.1.

Torsten Hothorn, Lisa Moest, Peter Buehlmann (2018), Most Likely Transformations, *Scandinavian Journal of Statistics*, **45**(1), 110–134, doi:10.1111/sjos.12291.

## Examples

```
data("BostonHousing2", package = "mlbench")

lm(cmedv ~ crim + zn + indus + chas + nox + rm + age + dis +
    rad + tax + ptratio + b + lstat, data = BostonHousing2)

Colr(cmedv ~ chas + crim + zn + indus + nox +
    rm + age + dis + rad + tax + ptratio + b + lstat,
    data = BostonHousing2)
```

Coxph

*Cox Proportional Hazards Model***Description**

Cox model with fully parameterised baseline hazard function

**Usage**

```
Coxph(formula, data, subset, weights, offset, cluster, na.action = na.omit, ...)
```

**Arguments**

formula	an object of class "formula": a symbolic description of the model structure to be fitted. The details of model specification are given under <a href="#">tram</a> and in the package vignette.
data	an optional data frame, list or environment (or object coercible by <code>as.data.frame</code> to a data frame) containing the variables in the model. If not found in data, the variables are taken from <code>environment(formula)</code> .
subset	an optional vector specifying a subset of observations to be used in the fitting process.
weights	an optional vector of weights to be used in the fitting process. Should be NULL or a numeric vector. If present, the weighted log-likelihood is maximised.
offset	this can be used to specify an <i>a priori</i> known component to be included in the linear predictor during fitting. This should be NULL or a numeric vector of length equal to the number of cases.
cluster	optional factor with a cluster ID employed for computing clustered covariances.
na.action	a function which indicates what should happen when the data contain NAs. The default is set to <code>na.omit</code> .
...	additional arguments to <a href="#">tram</a> .

**Details**

The original implementation of Cox models via the partial likelihood, treating the baseline hazard function as a nuisance parameter, is available in [coxph](#). This function allows simultaneous estimation of the log-hazard ratios and the log-cumulative baseline hazard, the latter parameterised by a Bernstein polynomial. The model can be fitted under stratification (time-varying coefficients), all types of random censoring and truncation. An early reference to this parameterisation is McLain and Ghosh (2013).

The response is bounded (`bounds = c(0, Inf)`) when specified as a `Surv` object. Otherwise, bounds can be specified via ...

Parameters are log-hazard ratios comparing treatment (or a one unit increase in a numeric variable) with a reference.

**Value**

An object of class `Coxph`, with corresponding `coef`, `vcov`, `logLik`, `estfun`, `summary`, `print`, `plot` and `predict` methods.

**References**

Alexander C. McLain and Sujit K. Ghosh (2013). Efficient Sieve Maximum Likelihood Estimation of Time-Transformation Models, *Journal of Statistical Theory and Practice*, 7(2), 285–303, doi:10.1080/15598608.2013.772835.

Torsten Hothorn, Lisa Moest, Peter Buehlmann (2018), Most Likely Transformations, *Scandinavian Journal of Statistics*, 45(1), 110–134, doi:10.1111/sjos.12291.

**Examples**

```
data("GBSG2", package = "TH.data")

library("survival")
(m1 <- coxph(Surv(time, cens) ~ horTh, data = GBSG2))

(m2 <- Coxph(Surv(time, cens) ~ horTh, data = GBSG2))

### Wald intervals
confint(m1)
confint(m2)
### profile likelihood interval
confint(profile(m2))
### score interval
confint(score_test(m2))
### permutation score interval; uses permutation distribution
### see coin::independence_test
## Not run: confint(perm_test(m2))
```

---

Lehmann

---

*Proportional Reverse Time Hazards Linear Regression*


---

**Description**

Non-normal linear regression for Lehmann-alternatives

**Usage**

```
Lehmann(formula, data, subset, weights, offset, cluster, na.action = na.omit, ...)
```

**Arguments**

formula	an object of class "formula": a symbolic description of the model structure to be fitted. The details of model specification are given under <code>tram</code> and in the package vignette.
data	an optional data frame, list or environment (or object coercible by <code>as.data.frame</code> to a data frame) containing the variables in the model. If not found in data, the variables are taken from <code>environment(formula)</code> .
subset	an optional vector specifying a subset of observations to be used in the fitting process.
weights	an optional vector of weights to be used in the fitting process. Should be <code>NULL</code> or a numeric vector. If present, the weighted log-likelihood is maximised.
offset	this can be used to specify an <code>_a priori_</code> known component to be included in the linear predictor during fitting. This should be <code>NULL</code> or a numeric vector of length equal to the number of cases.
cluster	optional factor with a cluster ID employed for computing clustered covariances.
na.action	a function which indicates what should happen when the data contain NAs. The default is set to <code>na.omit</code> .
...	additional arguments to <code>tram</code> .

**Details**

This transformation model uses the cumulative distribution function for the standard Gumbel maximum extreme value distribution to map the shifted transformation function into probabilities. The exponential of the shift parameter can be interpreted as a Lehmann-alternative or reverse time hazard ratio.

**Value**

An object of class `Lehmann`, with corresponding `coef`, `vcov`, `logLik`, `estfun`, `summary`, `print`, `plot` and `predict` methods.

**References**

Erich L. Lehmann (1953), The Power of Rank Tests, *The Annals of Mathematical Statistics*, **24**(1), 23-43.

Torsten Hothorn, Lisa Moest, Peter Buehlmann (2018), Most Likely Transformations, *Scandinavian Journal of Statistics*, **45**(1), 110–134, doi:[10.1111/sjos.12291](https://doi.org/10.1111/sjos.12291).

**Examples**

```
data("BostonHousing2", package = "mlbench")

lm(cmedv ~ crim + zn + indus + chas + nox + rm + age + dis +
    rad + tax + ptratio + b + lstat, data = BostonHousing2)

Lehmann(cmedv ~ chas + crim + zn + indus + nox +
```

```
rm + age + dis + rad + tax + ptratio + b + lstat,
data = BostonHousing2)
```

Lm

*Normal Linear Model***Description**

Normal linear model with benefits

**Usage**

```
Lm(formula, data, subset, weights, offset, cluster, na.action = na.omit, ...)
```

**Arguments**

formula	an object of class "formula": a symbolic description of the model structure to be fitted. The details of model specification are given under <a href="#">tram</a> and in the package vignette.
data	an optional data frame, list or environment (or object coercible by <code>as.data.frame</code> to a data frame) containing the variables in the model. If not found in data, the variables are taken from <code>environment(formula)</code> .
subset	an optional vector specifying a subset of observations to be used in the fitting process.
weights	an optional vector of weights to be used in the fitting process. Should be NULL or a numeric vector. If present, the weighted log-likelihood is maximised.
offset	this can be used to specify an <code>_a priori_</code> known component to be included in the linear predictor during fitting. This should be NULL or a numeric vector of length equal to the number of cases.
cluster	optional factor with a cluster ID employed for computing clustered covariances.
na.action	a function which indicates what should happen when the data contain NAs. The default is set to <code>na.omit</code> .
...	additional arguments to <a href="#">tram</a> .

**Details**

A normal linear model with simultaneous estimation of regression coefficients and scale parameter(s). This function also allows for stratum-specific intercepts and variances as well as censoring and truncation in the response.

Note that the scale of the parameters is different from what is reported by [lm](#); the discrepancies are explained in the package vignette.

The model is defined with a negative shift term. Large values of the linear predictor correspond to large values of the conditional expectation response.

**Value**

An object of class `Lm`, with corresponding `coef`, `vcov`, `logLik`, `estfun`, `summary`, `print`, `plot` and `predict` methods.

**References**

Torsten Hothorn, Lisa Moest, Peter Buehlmann (2018), Most Likely Transformations, *Scandinavian Journal of Statistics*, **45**(1), 110–134, doi:[10.1111/sjos.12291](https://doi.org/10.1111/sjos.12291).

**Examples**

```
data("BostonHousing2", package = "mlbench")

lm(cmedv ~ crim + zn + indus + chas + nox + rm + age + dis +
    rad + tax + ptratio + b + lstat, data = BostonHousing2)

Lm(cmedv ~ chas + crim + zn + indus + nox +
    rm + age + dis + rad + tax + ptratio + b + lstat,
    data = BostonHousing2)
```

---

ltmatrices

*Lower Triangular Matrices*


---

**Description**

Storage of and computations on multiple lower triangular matrices

**Usage**

```
ltmatrices(x, diag = FALSE, byrow = FALSE, names = TRUE)
diagonals(x, ...)
```

**Arguments**

<code>x</code>	a numeric matrix, each row is interpreted as the elements of a lower triangular matrix in row order (byrow being TRUE) or column order (otherwise). The matrices may contain non-unit diagonals (when <code>diag</code> is TRUE).
<code>diag</code>	logical; <code>x</code> contains non-unit diagonals if TRUE.
<code>byrow</code>	logical; <code>x</code> is in row order if TRUE.
<code>names</code>	an optional vector of names, autogenerated if TRUE.
<code>...</code>	additional arguments.

**Details**

Currently only for internal use.

**Value**

An object of class `ltmatrices`

**Examples**

```
x <- matrix(runif(30), ncol = 3)
ltmatrices(x, diag = FALSE, byrow = TRUE)
```

---

mmlt

*Multivariate Conditional Transformation Models*


---

**Description**

A proof-of-concept implementation of multivariate conditional transformation models

**Usage**

```
mmlt(..., formula = ~ 1, data, conditional = GAUSSIAN, theta = NULL,
      control.outer = list(trace = FALSE), scale = FALSE, dofit = TRUE)
## S3 method for class 'mmlt'
coef(object, newdata,
      type = c("all", "marginal", "Lambda", "Lambdainv",
              "Sigma", "Corr", "Spearman"),
      ...)
## S3 method for class 'mmlt'
predict(object, newdata, margins = 1:J,
        type = c("trafo", "distribution", "density"),
        log = FALSE, ...)
```

**Arguments**

<code>...</code>	marginal transformation models, one for each response, for <code>mmlt</code> . Additional arguments for the methods.
<code>formula</code>	a model formula describing a model for the dependency structure via the lambda parameters. The default is set to <code>~ 1</code> for constant lambdas.
<code>data</code>	a <code>data.frame</code> .
<code>conditional</code>	logical; parameters are defined conditionally (only possible when all models are probit models). This is the default as described by Klein et al. (2022). If <code>FALSE</code> , parameters can be directly interpreted marginally, this is explained in Section 2.6 by Klein et al. (2022). Using <code>conditional = FALSE</code> with probit-only models gives the same likelihood but different parameter estimates.
<code>theta</code>	an optional vector of starting values.
<code>control.outer</code>	a list controlling <a href="#">auglag</a>
<code>scale</code>	logical; parameters are not scaled prior to optimisation by default.

dofit	logical; parameters are fitted by default, otherwise a list with log-likelihood and score function is returned.
object	an object of class mmlt.
newdata	an optional data.frame coefficients and predictions shall be computed for.
type	type of coefficient or prediction to be returned.
margins	indices defining marginal models to be evaluated. Can be single integers giving the marginal distribution of the corresponding variable, or multiple integers (currently only 1:j implemented).
log	logical; return log-probabilities or log-densities of TRUE.

### Details

The function implements multivariate conditional transformation models as described by Klein et al (2020). The response is assumed absolutely continuous at the moment, discrete versions will be added later.

Below is a simple example for an unconditional bivariate distribution. See `demo("undernutrition", package = "tram")` for a conditional three-variate example.

### Value

An object of class `mmlt` with `coef` and `predict` methods.

### References

Nadja Klein, Torsten Hothorn, Luisa Barbanti, Thomas Kneib (2020), Multivariate Conditional Transformation Models. *Scandinavian Journal of Statistics*, doi:10.1111/sjos.12501.

### Examples

```
data("cars")

### fit unconditional bivariate distribution of speed and distance to stop
## fit unconditional marginal transformation models
m_speed <- BoxCox(speed ~ 1, data = cars, support = ss <- c(4, 25),
  add = c(-5, 5))
m_dist <- BoxCox(dist ~ 1, data = cars, support = sd <- c(0, 120),
  add = c(-5, 5))

## fit multivariate unconditional transformation model
m_speed_dist <- mmlt(m_speed, m_dist, formula = ~ 1, data = cars)

## log-likelihood
logLik(m_speed_dist)
sum(predict(m_speed_dist, newdata = cars, type = "density", log = TRUE))

## lambda defining the Cholesky of the precision matrix,
## with standard error
coef(m_speed_dist, type = "Lambda")
```

```

sqrt(vcov(m_speed_dist)["dist.sped.(Intercept)",
                        "dist.sped.(Intercept)"])

## simpler: Wald test of independence of speed and dist (the "dist.sped.(Intercept)"
## coefficient)
summary(m_speed_dist)

## linear correlation, ie Pearson correlation of speed and dist after
## transformation to bivariate normality
(r <- coef(m_speed_dist, type = "Corr"))

## Spearman's rho (rank correlation) of speed and dist on original scale
(rs <- coef(m_speed_dist, type = "Spearman"))

## evaluate joint and marginal densities (needs to be more user-friendly)
nd <- expand.grid(c(nd_s <- mkgrid(m_speed, 100), nd_d <- mkgrid(m_dist, 100)))
nd$d <- predict(m_speed_dist, newdata = nd, type = "density")

## compute marginal densities
nd_s <- as.data.frame(nd_s)
nd_s$d <- predict(m_speed_dist, newdata = nd_s, margins = 1L,
                type = "density")
nd_d <- as.data.frame(nd_d)
nd_d$d <- predict(m_speed_dist, newdata = nd_d, margins = 2L,
                type = "density")

## plot bivariate and marginal distribution
col1 <- rgb(.1, .1, .1, .9)
col2 <- rgb(.1, .1, .1, .5)
w <- c(.8, .2)
layout(matrix(c(2, 1, 4, 3), nrow = 2), width = w, height = rev(w))
par(mai = c(1, 1, 0, 0) * par("mai"))
sp <- unique(nd$speed)
di <- unique(nd$dist)
d <- matrix(nd$d, nrow = length(sp))
contour(sp, di, d, xlab = "Speed (in mph)", ylab = "Distance (in ft)", xlim = ss, ylim = sd,
        col = col1)
points(cars$speed, cars$dist, pch = 19, col = col2)
mai <- par("mai")
par(mai = c(0, 1, 0, 1) * mai)
plot(d ~ speed, data = nd_s, xlim = ss, type = "n", axes = FALSE,
     xlab = "", ylab = "")
polygon(nd_s$speed, nd_s$d, col = col2, border = FALSE)
par(mai = c(1, 0, 1, 0) * mai)
plot(dist ~ d, data = nd_d, ylim = sd, type = "n", axes = FALSE,
     xlab = "", ylab = "")
polygon(nd_d$d, nd_d$dist, col = col2, border = FALSE)

### NOTE: marginal densities are NOT normal, nor is the joint
### distribution. The non-normal shape comes from the data-driven
### transformation of both variables to joint normality in this model.

```

---

mtram

*Transformation Models for Clustered Data*


---

## Description

Marginally interpretable transformation models for clustered data.

## Usage

```
mtram(object, formula, data,
      grd = SparseGrid::createSparseGrid(type = "KPU",
        dimension = length(rt$cnms[[1]]), k = 10),
      Hessian = FALSE, tol = .Machine$double.eps, ...)
```

## Arguments

object	A tram object.
formula	A formula specifying the random effects.
data	A data frame.
grd	A sparse grid used for numerical integration to get the likelihood.
Hessian	A logical, if TRUE, the hessian is computed and returned.
tol	numerical tolerance.
...	Additional argument.

## Details

A Gaussian copula with a correlation structure obtained from a random intercept or random intercept / random slope model (that is, clustered or longitudinal data can be modelled only) is used to capture the correlations whereas the marginal distributions are described by a transformation model. The methodology is described in Barbanti and Hothorn (2022) and examples are given in the `mtram` package vignette.

This is a proof-of-concept implementation. Only `coef()` and `logLik()` methods are available at the moment.

## Value

An object of class `tram` with `coef()` and `logLik()` methods.

## References

Luisa Barbanti and Torsten Hothorn (2022). A Transformation Perspective on Marginal and Conditional Models, [doi:10.48550/arXiv.1910.09219](https://doi.org/10.48550/arXiv.1910.09219).

## Examples

```
### For illustrations see
## vignette("mtram", package = "tram")
## or
## demo("mtram", package = "tram")
```

---

perm\_test

*Permutation Transformation Tests*

---

## Description

P-values for a parameter in a linear transformation model and corresponding confidence intervals obtained from by the permutation principle

## Usage

```
perm_test(object, ...)
## S3 method for class 'tram'
perm_test(object, parm = names(coef(object)),
  statistic = c("Score", "Likelihood", "Wald"),
  alternative = c("two.sided", "less", "greater"),
  nullvalue = 0, confint = TRUE, level = .95,
  Taylor = FALSE, block_permutation = TRUE, maxsteps = 25, ...)
```

## Arguments

object	an object of class <code>tram</code>
parm	a vector of names of parameters to be tested. These parameters must be present in object.
statistic	a character string specifying the statistic to be permuted. The default <code>Score</code> is the classical permutation test for the esiduals of a model excluding the parameter <code>parm</code> . Only available for <code>nullvalue = 0</code> , confidence intervals are not available. Permuting the likelihood or the model coefficients under the <code>nullvalue</code> is highly experimental as are the corresponding confidence intervals.
alternative	a character string specifying the alternative hypothesis, must be one of <code>"two.sided"</code> (default), <code>"greater"</code> or <code>"less"</code> .
nullvalue	a number specifying an optional parameter used to form the null hypothesis.
confint	a logical indicating whether a confidence interval should be computed. Score confidence intervals are computed by default. A 1st order Taylor approximation to the <code>Score</code> statistic is used with <code>Taylor = TRUE</code> (in case numerical inversion of the score statistic fails, Wald-type confidence intervals relying from this approximation are returned) . For the remaining likelihood and Wald statistics, confidence intervals are highly experimental (and probably not worth looking at).

level	the confidence level.
block_permutation	a logical indicating wheather stratifying variables shall be interpreted as blocks defining admissible permutations.
Taylor	a logical requesting the use of a 1st order Taylor approximation when inverting the score statistic.
maxsteps	number of function evaluations when inverting the score statistic for computing confidence intervals.
...	additional arguments to <a href="#">independence_test</a> .

### Details

Permutation test for one single parameters in the linear predictor of object is computed. This parameters must be present in object. This is somewhat experimental and not recommended for serious practical use (yet!).

### Value

An object of class `htest` or a list thereof. See [Coxph](#) for an example.

### Examples

```
## Tritiated Water Diffusion Across Human Chorioamnion
## Hollander and Wolfe (1999, p. 110, Tab. 4.1)
diffusion <- data.frame(
  pd = c(0.80, 0.83, 1.89, 1.04, 1.45, 1.38, 1.91, 1.64, 0.73, 1.46,
        1.15, 0.88, 0.90, 0.74, 1.21),
  age = factor(rep(c("At term", "12-26 Weeks"), c(10, 5)))
)

### plot the two quantile functions
boxplot(pd ~ age, data = diffusion)

### the Wilcoxon rank sum test, with a confidence interval
### for a median shift
wilcox.test(pd ~ age, data = diffusion, conf.int = TRUE, exact = TRUE)

### a corresponding parametric transformation model with a log-odds ratio
### difference parameter, ie a difference on the log-odds scale
md <- Colr(pd ~ age, data = diffusion)

### assess model fit by plotting estimated distribution fcts
agef <- sort(unique(diffusion$age))
col <- c("black", "darkred")
plot(as.mlt(md), newdata = data.frame(age = agef),
     type = "distribution", col = col)
legend("bottomright", col = col, lty = 1, legend = levels(agef),
     bty = "n", pch = 19)
## compare with ECDFs: not too bad (but not good, either)
npfit <- with(diffusion, tapply(pd, age, ecdf))
```

```

lines(npfit[[1]], col = col[1])
lines(npfit[[2]], col = col[2])

### Wald confidence interval
confint(md)

### Likelihood confidence interval
confint(profile(md))

### Score confidence interval
confint(score_test(md))
confint(score_test(md, Taylor = TRUE))

### exact permutation score test
(pt <- perm_test(md, confint = TRUE, distribution = "exact"))
(pt <- perm_test(md, confint = TRUE, distribution = "exact",
                 Taylor = TRUE))

### compare with probabilistic indices obtained from asht::wmwTest
if (require("asht", warn.conflicts = FALSE)) {
  print(wt2 <- wmwTest(pd ~ I(relevel(age, "At term")),
                     data = diffusion, method = "exact.ce"))
  ### as log-odds ratios
  print(PI(prob = wt2$conf.int))
  print(PI(prob = wt2$estimate))
}

```

---

 Polr

*Ordered Categorical Regression*


---

## Description

Some regression models for ordered categorical responses

## Usage

```
Polr(formula, data, subset, weights, offset, cluster, na.action = na.omit,
     method = c("logistic", "probit", "loglog", "cloglog", "cauchit"), ...)
```

## Arguments

formula	an object of class "formula": a symbolic description of the model structure to be fitted. The details of model specification are given under <a href="#">tram</a> and in the package vignette.
data	an optional data frame, list or environment (or object coercible by <code>as.data.frame</code> to a data frame) containing the variables in the model. If not found in data, the variables are taken from <code>environment(formula)</code> .
subset	an optional vector specifying a subset of observations to be used in the fitting process.

weights	an optional vector of weights to be used in the fitting process. Should be NULL or a numeric vector. If present, the weighted log-likelihood is maximised.
offset	this can be used to specify an <code>_a priori_</code> known component to be included in the linear predictor during fitting. This should be NULL or a numeric vector of length equal to the number of cases.
cluster	optional factor with a cluster ID employed for computing clustered covariances.
na.action	a function which indicates what should happen when the data contain NAs. The default is set by the <code>na.action</code> setting of <code>options</code> , and is <code>na.fail</code> if that is unset.
method	a character describing the link function.
...	additional arguments to <code>tram</code> .

## Details

Models for ordered categorical responses reusing the interface of `polr`. Allows for stratification, censoring and truncation.

The model is defined with a negative shift term, thus `exp(coef())` is the multiplicative change of the odds ratio (conditional odds for reference divided by conditional odds of treatment or for a one unit increase in a numeric variable). Large values of the linear predictor correspond to large values of the conditional expectation response (but this relationship is nonlinear).

## Value

An object of class `Polr`, with corresponding `coef`, `vcov`, `logLik`, `estfun`, `summary`, `print`, `plot` and `predict` methods.

## References

Torsten Hothorn, Lisa Moest, Peter Buehlmann (2018), Most Likely Transformations, *Scandinavian Journal of Statistics*, **45**(1), 110–134, doi:[10.1111/sjos.12291](https://doi.org/10.1111/sjos.12291).

## Examples

```
data("wine", package = "ordinal")

library("MASS")
polr(rating ~ temp + contact, data = wine)

Polr(rating ~ temp + contact, data = wine)
```

score\_test

*Transformation Score Tests and Confidence Intervals***Description**

P-values and confidence intervals for parameters in linear transformation models obtained from by the score test principle

**Usage**

```
score_test(object, ...)
## S3 method for class 'tram'
score_test(object, parm = names(coef(object)),
           alternative = c("two.sided", "less", "greater"), nullvalue = 0,
           confint = TRUE, level = .95, Taylor = FALSE, maxsteps = 25, ...)
```

**Arguments**

object	an object of class <code>tram</code>
parm	a vector of names of parameters to be tested. These parameters must be present in object.
alternative	a character string specifying the alternative hypothesis, must be one of "two.sided" (default), "greater" or "less".
nullvalue	a number specifying an optional parameter used to form the null hypothesis.
confint	a logical indicating whether a confidence interval should be computed. Score confidence intervals are computed by default. A 1st order Taylor approximation to the Score statistic is used with <code>Taylor = TRUE</code> (in case numerical inversion of the score statistic fails, Wald confidence intervals relying from this approximation are returned).
level	the confidence level.
Taylor	a logical requesting the use of a 1st order Taylor approximation when inverting the score statistic.
maxsteps	number of function evaluations when inverting the score statistic for computing confidence intervals.
...	additional arguments, currently ignored.

**Details**

Score tests and confidence intervals for the parameters in the linear predictor of object are computed. These parameters must be present in object.

**Value**

An object of class `htest` or a list thereof. See [Coxph](#) for an example. A corresponding permutation test for parameters in a transformation models is available in [perm\\_test](#).

**Description**

Weibull, log-normal, log-logistic and other parametric models (not exclusively) for survival analysis

**Usage**

```
Survreg(formula, data, subset, weights, offset, cluster, na.action = na.omit,
        dist = c("weibull", "logistic", "gaussian", "exponential", "rayleigh",
                "loggaussian", "lognormal", "loglogistic"), scale = 0, ...)
```

**Arguments**

formula	an object of class "formula": a symbolic description of the model structure to be fitted. The details of model specification are given under <a href="#">tram</a> and in the package vignette.
data	an optional data frame, list or environment (or object coercible by <code>as.data.frame</code> to a data frame) containing the variables in the model. If not found in data, the variables are taken from <code>environment(formula)</code> .
subset	an optional vector specifying a subset of observations to be used in the fitting process.
weights	an optional vector of weights to be used in the fitting process. Should be NULL or a numeric vector. If present, the weighted log-likelihood is maximised.
offset	this can be used to specify an <code>_a priori_</code> known component to be included in the linear predictor during fitting. This should be NULL or a numeric vector of length equal to the number of cases.
cluster	optional factor with a cluster ID employed for computing clustered covariances.
na.action	a function which indicates what should happen when the data contain NAs. The default is set by the <code>na.action</code> setting of options, and is <code>na.fail</code> if that is unset.
dist	character defining the conditional distribution of the (not necessarily positive) response, current choices include Weibull, logistic, normal, exponential, Rayleigh, log-normal (same as log-gaussian), or log-logistic.
scale	a fixed value for the scale parameter(s).
...	additional arguments to <a href="#">tram</a> .

**Details**

Parametric survival models reusing the interface of [survreg](#). The parameterisation is, however, a little different, see the package vignette.

The model is defined with a negative shift term. Large values of the linear predictor correspond to large values of the conditional expectation response (but this relationship is nonlinear). Parameters are log-hazard ratios comparing a reference with treatment (or a one unit increase in a numeric variable).

**Value**

An object of class `Survreg`, with corresponding `coef`, `vcov`, `logLik`, `estfun`, `summary`, `print`, `plot` and `predict` methods.

**References**

Torsten Hothorn, Lisa Moest, Peter Buehlmann (2018), Most Likely Transformations, *Scandinavian Journal of Statistics*, **45**(1), 110–134, doi:[10.1111/sjos.12291](https://doi.org/10.1111/sjos.12291).

**Examples**

```
data("GBSG2", package = "TH.data")

library("survival")
survreg(Surv(time, cens) ~ horTh, data = GBSG2)

Survreg(Surv(time, cens) ~ horTh, data = GBSG2)
```

---

tram

---

*Stratified Linear Transformation Models*


---

**Description**

Likelihood-inference for stratified linear transformation models

**Usage**

```
tram(formula, data, subset, weights, offset, cluster, na.action = na.omit,
      distribution = c("Normal", "Logistic", "MinExtrVal", "MaxExtrVal",
                      "Exponential", "Cauchy", "Laplace"),
      transformation = c("discrete", "linear", "logarithmic", "smooth"),
      LRtest = TRUE, prob = c(0.1, 0.9), support = NULL,
      bounds = NULL, add = c(0, 0), order = 6,
      negative = TRUE, scale = TRUE, scale_shift = FALSE, extrapolate = FALSE,
      log_first = FALSE, sparse_nlevels = Inf,
      model_only = FALSE, constraints = NULL, ...)
tram_data(formula, data, subset, weights, offset, cluster, na.action = na.omit)
```

**Arguments**

formula	an object of class "formula": a symbolic description of the model structure to be fitted. The details of model specification are given under Details and in the package vignette.
data	an optional data frame, list or environment (or object coercible by <code>as.data.frame</code> to a data frame) containing the variables in the model. If not found in data, the variables are taken from <code>environment(formula)</code> .

subset	an optional vector specifying a subset of observations to be used in the fitting process.
weights	an optional vector of case weights to be used in the fitting process. Should be NULL or a numeric vector. If present, the weighted log-likelihood is maximised.
offset	this can be used to specify an <code>_a priori_</code> known component to be included in the linear predictor during fitting. This should be NULL or a numeric vector of length equal to the number of cases.
cluster	optional factor with a cluster ID employed for computing clustered covariances.
na.action	a function which indicates what should happen when the data contain NAs. The default is set to <code>na.omit</code> .
distribution	character specifying how the transformation function is mapped into probabilities. Available choices include the cumulative distribution functions of the standard normal, the standard logistic and the standard minimum extreme value distribution.
transformation	character specifying the complexity of the response-transformation. For discrete responses, one parameter is assigned to each level (except the last one), for continuous responses linear, log-linear and smooth (parameterised as a Bernstein polynomial) function are implemented.
LRtest	logical specifying if a likelihood-ratio test for the null of all coefficients in the linear predictor being zero shall be performed.
prob	two probabilities giving quantiles of the response defining the support of a smooth Bernstein polynomial (if <code>transformation = "smooth"</code> ).
support	a vector of two elements; the support of a smooth Bernstein polynomial (if <code>transformation = "smooth"</code> ).
bounds	an interval defining the bounds of a real sample space.
add	add these values to the support before generating a grid via <code>mkgrid</code> .
order	integer $\geq 1$ defining the order of the Bernstein polynomial (if <code>transformation = "smooth"</code> ).
negative	logical defining the sign of the linear predictor.
scale	logical defining if variables in the linear predictor shall be scaled. Scaling is internally used for model estimation, rescaled coefficients are reported in model output.
scale_shift	a logical choosing between two different model types in the presence of a scaling term, see <code>ctm</code> .
extrapolate	logical defining the behaviour of the Bernstein transformation function outside support. The default FALSE is to extrapolate linearly without requiring the second derivative of the transformation function to be zero at support. If TRUE, this additional constraint is respected.
sparse_nlevels	integer; use a sparse model matrix if the number of levels of an ordered factor is at least as large as <code>sparse_nlevels</code> .
log_first	logical; if TRUE, a Bernstein polynomial is defined on the log-scale.
model_only	logical, if TRUE the unfitted model is returned.

constraints additional constraints on regression coefficients in the linear predictor of the form `lhs %*% coef(object) >= rhs`, where `lhs` and `rhs` can be specified as a character (as in `glht`) or by a matrix `lhs` (assuming `rhs = 0`), or as a list containing the two elements `lhs` and `rhs`.

... additional arguments.

## Details

The model formula is of the form `y | s ~ x | z` where `y` is an at least ordered response variable, `s` are the variables defining strata and `x` defines the linear predictor. Optionally, `z` defines a scaling term (see `ctm`). `y ~ x` defines a model without strata (but response-varying intercept function) and `y | s ~ 0` sets-up response-varying coefficients for all variables in `s`.

The two functions `tram` and `tram_data` are not intended to be called directly by users. Instead, functions `Coxph` (Cox proportional hazards models), `Survreg` (parametric survival models), `Polr` (models for ordered categorical responses), `Lm` (normal linear models), `BoxCox` (non-normal linear models) or `Colr` (continuous outcome logistic regression) allow direct access to the corresponding models.

The model class and the specific models implemented in **tram** are explained in the package vignette of package **tram**. The underlying theory of most likely transformations is presented in Hothorn et al. (2018), computational and modelling aspects in more complex situations are discussed by Hothorn (2018).

## Value

An object of class `tram` inheriting from `mlt`.

## References

Torsten Hothorn, Lisa Moest, Peter Buehlmann (2018), Most Likely Transformations, *Scandinavian Journal of Statistics*, **45**(1), 110–134, doi:10.1111/sjos.12291.

Torsten Hothorn (2018), Most Likely Transformations: The `mlt` Package, *Journal of Statistical Software*, forthcoming. URL: <https://cran.r-project.org/package=mlt.docreg>

## Examples

```
data("BostonHousing2", package = "mlbench")

### unconstrained regression coefficients
### BoxCox calls tram internally
m1 <- BoxCox(cmedv ~ chas + crim + zn + indus + nox +
             rm + age + dis + rad + tax + ptratio + b + lstat,
             data = BostonHousing2)

### now with two constraints on regression coefficients
m2 <- BoxCox(cmedv ~ chas + crim + zn + indus + nox +
             rm + age + dis + rad + tax + ptratio + b + lstat,
             data = BostonHousing2,
             constraints = c("crim >= 0", "chas1 + rm >= 1.5"))
```

```

coef(m1)
coef(m2)

K <- matrix(0, nrow = 2, ncol = length(coef(m2)))
colnames(K) <- names(coef(m2))
K[1, "crim"] <- 1
K[2, c("chas1", "rm")] <- 1
m3 <- BoxCox(cmedv ~ chas + crim + zn + indus + nox +
             rm + age + dis + rad + tax + ptratio + b + lstat,
             data = BostonHousing2,
             constraints = list(K, c(0, 1.5)))
all.equal(coef(m2), coef(m3))

```

**Description**

Methods for objects inheriting from class tram

**Usage**

```

## S3 method for class 'tram'
as.mlt(object)
## S3 method for class 'tram'
model.frame(formula, ...)
## S3 method for class 'tram'
model.matrix(object, data = object$data, with_baseline = FALSE, ...)
## S3 method for class 'stram'
model.matrix(object, data = object$data, with_baseline = FALSE,
             what = c("shifting", "scaling"), ...)
## S3 method for class 'tram'
coef(object, with_baseline = FALSE, ...)
## S3 method for class 'Lm'
coef(object, as.lm = FALSE, ...)
## S3 method for class 'Survreg'
coef(object, as.survreg = FALSE, ...)
## S3 method for class 'tram'
vcov(object, with_baseline = FALSE, complete = FALSE, ...)
## S3 method for class 'tram'
logLik(object, parm = coef(as.mlt(object), fixed = FALSE), ...)
## S3 method for class 'tram'
estfun(object, parm = coef(as.mlt(object), fixed = FALSE), ...)
## S3 method for class 'tram'
predict(object, newdata = model.frame(object),
        type = c("lp", "trafo", "distribution", "survivor", "density",
                 "logdensity", "hazard", "loghazard", "cumhazard",

```

```

        "quantile"),
    ...)
## S3 method for class 'stram'
predict(object, newdata = model.frame(object),
        type = c("lp", "trafo", "distribution", "survivor", "density",
                "logdensity", "hazard", "loghazard", "cumhazard",
                "quantile"), what = c("shifting", "scaling"),
    ...)
## S3 method for class 'tram'
plot(x, newdata = model.frame(x),
     which = c("QQ-PIT", "baseline only", "distribution"),
     confidence = c("none", "interval", "band"), level = 0.95,
     K = 50, cheat = K, col = "black", fill = "lightgrey", lwd = 1, ...)
## S3 method for class 'tram'
residuals(object, ...)
## S3 method for class 'tram'
PI(object, newdata = model.frame(object), reference = 0,
    one2one = FALSE, ...)
## Default S3 method:
PI(object, prob, link = "logistic", ...)
## S3 method for class 'tram'
OVL(object, newdata = model.frame(object), reference = 0,
    one2one = FALSE, ...)
## Default S3 method:
OVL(object, link = "logistic", ...)
## S3 method for class 'tram'
TV(object, newdata = model.frame(object), reference = 0,
    one2one = FALSE, ...)
## Default S3 method:
TV(object, link = "logistic", ...)
## S3 method for class 'tram'
L1(object, newdata = model.frame(object), reference = 0,
    one2one = FALSE, ...)
## Default S3 method:
L1(object, link = "logistic", ...)
## S3 method for class 'tram'
ROC(object, newdata = model.frame(object), reference = 0,
     prob = 1:99 / 100, one2one = FALSE, ...)
## Default S3 method:
ROC(object, prob = 1:99 / 100, link = "logistic", ...)
## S3 method for class 'ROCtram'
plot(x, lty = 1:ncol(x), col = "black",
     fill = "lightgrey", lwd = 1, ...)

```

## Arguments

object, formula, x

a fitted stratified linear transformation model inheriting from class tram. PI also takes a numeric vector in the default method.

<code>data</code>	an optional data frame.
<code>with_baseline</code>	logical, if TRUE all model parameters are returned, otherwise parameters describing the baseline transformation are ignored.
<code>as.lm</code>	logical, return parameters in the <code>lm</code> parameterisation if TRUE.
<code>as.survreg</code>	logical, return parameters in the <code>survreg</code> parameterisation in TRUE.
<code>parm</code>	model parameters, including baseline parameters.
<code>complete</code>	currently ignored
<code>newdata</code>	an optional data frame of new observations.
<code>reference</code>	an optional data frame of reference observations, or a numeric vector of reference values.
<code>type</code>	type of prediction, current options include linear predictors (" <code>lp</code> ", of $x$ variables in the formula $y   s \sim x$ ), transformation functions (" <code>trafo</code> ") or distribution functions on the scale of the cdf (" <code>distribution</code> "), survivor function, density function, log-density function, hazard function, log-hazard function, cumulative hazard function or quantile function.
<code>which</code>	type of plot, either a QQ plot of the probability-integral transformed observations (" <code>QQ-PIT</code> "), of the baseline transformation of the whole distribution.
<code>what</code>	type of model matrix / linear predictor: <code>shifting</code> returns model model matrix / linear predictor for shift term, <code>scaling</code> for the scale term.
<code>confidence</code>	type of uncertainty assessment.
<code>level</code>	confidence level.
<code>K</code>	number of grid points in the response, see <code>plot.ctm</code> .
<code>cheat</code>	reduced number of grid points for the computation of confidence bands, see <code>confband</code> .
<code>col</code>	line color.
<code>fill</code>	fill color.
<code>lwd</code>	line width.
<code>lty</code>	line type.
<code>prob</code>	a numeric vector of probabilities..
<code>link</code>	a character identifying a link function.
<code>one2one</code>	logical, compute the ROC curve (and derived measures) comparing each row in <code>newdata</code> with each row in <code>reference</code> (FALSE, the default), or compare observations rowwise (TRUE).
<code>...</code>	additional arguments to the underlying methods for class <code>mlt</code> , see <code>mlt-methods</code> .

## Details

`coef` can be used to get (and set) model parameters, `logLik` evaluates the log-likelihood (also for parameters other than the maximum likelihood estimate); `vcov` returns the estimated variance-covariance matrix (possibly taking `cluster` into account) and `estfun` gives the score contribution by each observation. `predict` and `plot` can be used to inspect the model on different scales.

PI computes the probabilistic index (or concordance probability or AUC) for all observations in newdata, relative to reference, ie the probability

$$P(Y_1 \leq Y_0 \mid x_0, x_1)$$

of observing a smaller value of a randomly sampled observation conditional on  $x_1$  compared to a randomly sampled reference observation, which is conditional on  $x_0$ . This is equivalent to the area under the receiver operating curve (ROC). The probability only applies within strata, response-varying coefficients are not allowed.

Under the same setup, OVL gives the overlap coefficient, which is one minus the total variation and one minus half the  $L_1$  distance between the two conditional densities. The overlap coefficient is identical to the Youden index and the Smirnov statistic.

PI and friends also accept an argument `conf.level` which triggers computation of simultaneous Wald confidence intervals for these measures. Arguments in ... are forwarded to [glht](#).

## References

Torsten Hothorn, Lisa Moest, Peter Buehlmann (2018), Most Likely Transformations, *Scandinavian Journal of Statistics*, **45**(1), 110–134, doi:10.1111/sjos.12291.

## See Also

[mlt-methods](#), [plot.ctm](#)

## Examples

```
data("BostonHousing2", package = "mlbench")

### fit non-normal Box-Cox type linear model with two
### baseline functions (for houses near and off Charles River)
BC_BH_2 <- BoxCox(cmedv | 0 + chas ~ crim + zn + indus + nox +
                  rm + age + dis + rad + tax + ptratio + b + lstat,
                  data = BostonHousing2)
logLik(BC_BH_2)

### classical likelihood inference
summary(BC_BH_2)

### coefficients of the linear predictor
coef(BC_BH_2)

### plot linear predictor (mean of _transformed_ response)
### vs. observed values
plot(predict(BC_BH_2, type = "lp"), BostonHousing2$cmedv)

### all coefficients
coef(BC_BH_2, with_baseline = TRUE)

### compute predicted median along with 10% and 90% quantile for the first
### observations
```

```

predict(BC_BH_2, newdata = BostonHousing2[1:3,], type = "quantile",
       prob = c(.1, .5, .9))

### plot the predicted density for these observations
plot(BC_BH_2, newdata = BostonHousing2[1:3, -1],
     which = "distribution", type = "density", K = 1000)

### evaluate the two baseline transformations, with confidence intervals
nd <- model.frame(BC_BH_2)[1:2, -1]
nd$chas <- factor(c("0", "1"))
library("colorspace")
col <- diverge_hcl(2, h = c(246, 40), c = 96, l = c(65, 90))
fill <- diverge_hcl(2, h = c(246, 40), c = 96, l = c(65, 90), alpha = .3)
plot(BC_BH_2, which = "baseline only", newdata = nd, col = col,
     confidence = "interval", fill = fill, lwd = 2,
     xlab = "Median Value", ylab = expression(h[Y]))
legend("bottomright", lty = 1, col = col,
      title = "Near Charles River", legend = c("no", "yes"), bty = "n")

### cars data; with quantile functions
plot(dist ~ speed, data = cars)
m <- Colr(dist ~ speed, data = cars)
q <- predict(as.mlt(m), newdata = data.frame(speed = s <- 6:25),
            type = "quantile", prob = c(1, 5, 9) / 10)
lines(s, q[1,])
lines(s, q[2,])
lines(s, q[3,])

nd <- data.frame(speed = s <- as.double(1:5 * 5))

# Prob(dist at speed s > dist at speed 0)
# speed 0 is reference, not a good choice here
PI(m, newdata = nd)

# Prob(dist at speed s > dist at speed 15)
lp15 <- c(predict(m, newdata = data.frame(speed = 15)))
PI(m, newdata = nd, reference = lp15)
PI(m, newdata = nd, reference = nd[,3], drop = FALSE))

# Prob(dist at speed s' > dist at speed s)
PI(m, newdata = nd, reference = nd)
# essentially:
lp <- predict(m, newdata = nd)
PI(object = dist(lp))
# same, with simultaneous confidence intervals
PI(m, newdata = nd, reference = nd, conf.level = .95)

# plot ROC curves + confidence bands
# compare speed 20 and 25 to speed 15
plot(ROC(m, newdata = nd[4:5,], drop = FALSE],
     reference = nd[3,], drop = FALSE],
     conf.level = 0.95))

```

```

# Overlap of conditional densities at speed s' and s
OVL(m, newdata = nd, reference = nd)

### ROC analysis (takes too long for CRAN Windows)
if (require("mlbench") && .Platform$OS.type != "windows") {

  layout(matrix(1:4, nrow = 2))
  data("PimaIndiansDiabetes2", package = "mlbench")
  dia <- sort(unique(PimaIndiansDiabetes2$diabetes))
  nd <- data.frame(diabetes = dia,
                  age = 29, mass = 32) ### median values

  ### unconditional ROC analysis: glucose tolerance test
  m0 <- Colr(glucose ~ diabetes, data = PimaIndiansDiabetes2)
  # ROC curve + confidence band
  plot(ROC(m0, newdata = nd[2,,drop = FALSE], conf.level = .95))
  # Wald interval for AUC
  PI(m0, newdata = nd[2,,drop = FALSE], conf.level = .95)
  # score interval for AUC
  PI(-c(coef(m0), score_test(m0)$conf.int[2:1]))

  ### adjusted ROC analysis for age and mass
  m1 <- Colr(glucose ~ diabetes + age + mass, data = PimaIndiansDiabetes2)
  # ROC curve + confidence band (this is the same for all ages /
  # masses)
  plot(ROC(m1, newdata = nd[2,,drop = FALSE],
          reference = nd[1,,drop = FALSE],
          conf.level = .95))
  # Wald interval for adjusted AUC
  PI(m1, newdata = nd[2,,drop = FALSE], reference = nd[1,,drop = FALSE],
    conf.level = .95)
  # Score interval for adjusted AUC
  PI(-c(coef(m1)[1], score_test(m1, names(coef(m1))[1])$conf.int[2:1]))

  ### conditional ROC analysis: AUC regression ~ age + mass
  m2 <- Colr(glucose ~ diabetes * (age + mass), data = PimaIndiansDiabetes2)
  # ROC curve for a person with age = 29 and mass = 32
  plot(ROC(m2, newdata = nd[2,,drop = FALSE],
          reference = nd[1,,drop = FALSE],
          conf.level = .95))
  # AUC for persons ages 21:81, all with mass = 32
  nd1 <- data.frame(diabetes = nd[1,"diabetes"], age = 21:81, mass = 32)
  nd2 <- data.frame(diabetes = nd[2,"diabetes"], age = 21:81, mass = 32)
  auc <- PI(m2, newdata = nd2, reference = nd1, one2one = TRUE,
            conf.level = 0.95)
  plot(nd1$age, auc[, "Estimate"], xlab = "Age (in years)", ylab =
        "AUC", ylim = c(0, 1), type = "l")
  lines(nd1$age, auc[, "lwr"], lty = 3)
  lines(nd1$age, auc[, "upr"], lty = 3)
}

```

# Index

- \* **matrix**
    - ltmatrices, 11
  - \* **models**
    - Aareg, 2
    - BoxCox, 4
    - Colr, 5
    - Coxph, 7
    - Lehmann, 8
    - Lm, 10
    - mmlt, 12
    - mtram, 15
    - Polr, 18
    - Survreg, 21
    - tram, 22
  - \* **regression**
    - Aareg, 2
    - BoxCox, 4
    - Colr, 5
    - Coxph, 7
    - Lehmann, 8
    - Lm, 10
    - Polr, 18
    - Survreg, 21
    - tram, 22
  - \* **smooth**
    - Aareg, 2
    - BoxCox, 4
    - Colr, 5
    - Coxph, 7
    - Lehmann, 8
    - tram, 22
  - \* **survival**
    - Aareg, 2
    - Coxph, 7
    - Survreg, 21
    - tram, 22
- Aareg, 2  
as.mlt.tram (tram-methods), 25  
auglag, 12  
BoxCox, 4, 24  
coef.Lm (tram-methods), 25  
coef.mmlt (mmlt), 12  
coef.Survreg (tram-methods), 25  
coef.tram (tram-methods), 25  
Colr, 5, 24  
confband, 27  
Coxph, 7, 17, 20, 24  
coxph, 7  
ctm, 23, 24  
diagonals (ltmatrices), 11  
estfun.tram (tram-methods), 25  
glht, 24, 28  
independence\_test, 17  
L1 (tram-methods), 25  
Lehmann, 8  
Lm, 10, 24  
lm, 10, 27  
logLik.tram (tram-methods), 25  
ltmatrices, 11  
mkgrid, 23  
mmlt, 12  
model.frame.tram (tram-methods), 25  
model.matrix.stram (tram-methods), 25  
model.matrix.tram (tram-methods), 25  
mtram, 15  
OVL (tram-methods), 25  
perm\_test, 16, 20  
PI (tram-methods), 25  
plot.ctm, 27, 28  
plot.ROCtram (tram-methods), 25  
plot.tram (tram-methods), 25

Polr, [18](#), [24](#)  
polr, [19](#)  
predict.mmlt (mmlt), [12](#)  
predict.stram (tram-methods), [25](#)  
predict.tram (tram-methods), [25](#)  
  
residuals.tram (tram-methods), [25](#)  
ROC (tram-methods), [25](#)  
  
score\_test, [20](#)  
Survreg, [21](#), [24](#)  
survreg, [21](#), [27](#)  
  
tram, [2-7](#), [9](#), [10](#), [16](#), [18-21](#), [22](#)  
tram-methods, [25](#)  
tram\_data (tram), [22](#)  
TV (tram-methods), [25](#)  
  
vcov.tram (tram-methods), [25](#)