Package: brtobit (via r-universe)

September 16, 2024

Title Bias-Reduced Tobit Regression

Version 0.1-2

Date 2024-05-12

Depends R (>= 3.6.0)

Imports stats, crch, sandwich

Suggests distributions3, memisc

Description Tobit models are regression models with a Gaussian response variable left-censored at zero, constant latent variance, and a latent mean that depends on covariates through a linear predictor. As an alternative to plain maximum likelihood estimation, the adjusted score equations of Kosmidis and Firth (2010) <doi:10.1214/10-ejs579> are utilized to obtain bias-reduced estimates of the model parameters.

License GPL-2 | GPL-3

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

RemoteUrl https://github.com/r-forge/topmodels

RemoteRef HEAD

RemoteSha 06b70d6fea89fc0d7f1e153e7fe9a5f80cca0aee

Contents

`brtobit` *Bias-Reduced Tobit Regression*

Description

Fitting tobit regression models with bias-reduced estimation (rather than plain maximum likelihood).

Usage

```
brtobit(formula, data, subset, na.action,
  model = TRUE, y = TRUE, x = FALSE,
  control = brtobit_control(...), ...)

brtobit_fit(x, y, control = brtobit_control())

brtobit_control(fsmaxit = 100, start = NULL, epsilon = 1e-08, type = "BR", ...)
```

Arguments

formula	a formula expression of the form y ~ x1 + x2 where y is the response and x1 and x2 are regressor variables for the location of the latent Gaussian distribution.
data	an optional data frame containing the variables occurring in the formulas.
subset	an optional vector specifying a subset of observations to be used for fitting.
na.action	a function which indicates what should happen when the data contain NAs.
model	logical. If TRUE *model frame* is included as a component of the returned value.
x, y	for brtobit: logical. If TRUE the model matrix and response vector used for fitting are returned as components of the returned value. For brtobit_fit: x is a design matrix with regressors for the location and y is a vector of observations.
...	arguments to be used to form the default control argument if it is not supplied directly.
control, fsmaxit, start, epsilon	a list of control parameters passed for the Fisher scoring optimization.
type	character. Should bias-reduced (BR) or plain maximum likelihood (ML) estimation be used?

Details

brtobit fits tobit regression models with bias-reduced (BR) estimation as introduced by Köll et al. (2021). The model assumes an underlying latent Gaussian variable:

$$y_i^* \sim \mathcal{N}(\mu_i, \sigma^2)$$

which is only observed if positive and zero otherwise: $y_i = \max(0, y_i^*)$. The latent mean μ_i is linked to a linear predictor

$$\mu_i = x_i^\top \beta$$

and the latent variance σ^2 is assumed to be constant.

brtobit_fit is the lower level function where the actual fitting takes place.

A set of standard extractor functions for fitted model objects is available for objects of class "brtobit", including methods to the generic functions print, summary, coef, vcov, logLik, predict, model.frame, model.matrix, bread (from the **sandwich** package), getSummary (from the **memisc** package, enabling mtable), and prodist (from the **distributions3** package, enabling various methods and graphics from the **topmodels** packages).

In the future we intend to extend the implementation to heteroscedastic tobit models in the crch package (Messner, Mayr, Zeileis 2016).

Value

brtobit returns an object of class "brtobit", i.e., a list with components as follows. brtobit_fit returns an unclassed list with components up to converged.

coefficients	vector of estimated regression coefficients (plus the variance),
bias	bias estimate,
vcov	covariance matrix of all parameters in the model,
loglik	the log-likelihood of the fitted model,
df	number of estimated parameters,
nobs	number of observations,
grad	gradient vector,
control	list of control parameters,
iterations	number of iterations,
converged	logical indicating whether the Fisher scoring optimization converged,
call	the original function call,
formula	the original formula,
terms	terms objects for the model,
levels	levels of the categorical regressors,
contrasts	contrasts corresponding to levels from the respective models,
model	the full model frame (if model = TRUE),
y	the numeric response vector (if y = TRUE),
x	model matrix (if x = TRUE).

References

Köll S, Kosmidis I, Kleiber C, Zeileis A (2021). "Bias Reduction as a Remedy to the Consequences of Infinite Estimates in Poisson and Tobit Regression." arXiv:2101.07141, arXiv.org E-Print Archive. https://arxiv.org/abs/2101.07141

Messner JW, Mayr GJ, Zeileis A (2016). Heteroscedastic Censored and Truncated Regression with crch. *The R Journal*, **8**(1), 173–181. https://journal.R-project.org/archive/2016-1/messner-mayr-zeileis.pdf.

See Also

crch

Examples

```
## artificial data generating process from Koell et al. (2021)
dgp <- function(n = 100, coef = c(1, 1, -10, 2), prob = 0.25) {
  x2 <- runif(n, -1, 1)
  x3 <- rbinom(n, size = 1, prob = ifelse(x2 > 0, prob, 1 - prob))
  y <- rnorm(n, mean = coef[1] + coef[2] * x2 + coef[3] * x3, sd = sqrt(coef[4]))
  y[y <= 0] <- 0
  data.frame(y, x2, x3)
}

set.seed(2020-10-29)
d <- dgp()

## models
m22_ml <- brtobit(y ~ x2 + x3, data = d, type = "ML", fsmaxit = 28)
m22_br <- brtobit(y ~ x2 + x3, data = d, type = "BR")
m2_all <- brtobit(y ~ x2, data = d, type = "ML")
m2_sub <- update(m2_all, subset = x3 == 0)

if(require("memisc")) {

## Table 2
mtable("ML" = m22_ml, "BR" = m22_br, "ML/sub" = m2_sub, "ML/SST" = m2_all,
  summary.stats = c("Log-likelihood", "N"))

}
```

Index