

# Package: PartCensReg (via r-universe)

September 1, 2024

**Type** Package

**Title** Estimation and Diagnostics for Partially Linear Censored Regression Models Based on Heavy-Tailed Distributions

**Version** 1.39

**Author** Marcela Nunez Lemus, Christian E. Galarza, Larissa Avila Matos, Victor H Lachos

**Maintainer** Marcela Nunez Lemus <marcela.nunez.lemus@gmail.com>

**Imports** ssym, optimx, Matrix

**Suggests** SMNCensReg, AER

**Description** It estimates the parameters of a partially linear regression censored model via maximum penalized likelihood through of ECME algorithm. The model belong to the semiparametric class, that including a parametric and nonparametric component. The error term considered belongs to the scale-mixture of normal (SMN) distribution, that includes well-known heavy tails distributions as the Student-t distribution, among others. To examine the performance of the fitted model, case-deletion and local influence techniques are provided to show its robust aspect against outlying and influential observations. This work is based in Ferreira, C. S., & Paula, G. A. (2017) <doi:10.1080/02664763.2016.1267124> but considering the SMN family.

**License** GPL (>= 2)

**NeedsCompilation** no

**Date/Publication** 2018-03-08 23:03:05 UTC

**Repository** <https://mnunez5.r-universe.dev>

**RemoteUrl** <https://github.com/cran/PartCensReg>

**RemoteRef** HEAD

**RemoteSha** e0752c6876792ff3bb3d1febcb2fbd11901b0c35

## Contents

PartCensReg-package . . . . .	2
Cens.SMN.PCR . . . . .	3
<b>Index</b>	<b>6</b>

---

PartCensReg-package     *Estimation and Diagnostics for Partially Linear Censored Regression Models Based on Heavy-Tailed Distributions*

---

## Description

It estimates the parameters of a partially linear regression censored model via maximum penalized likelihood through of ECME algorithm. The model belong to the semiparametric class, that including a parametric and nonparametric component. The error term considered belongs to the scale-mixture of normal (SMN) distribution, that includes well-known heavy tails distributions as the Student-t distribution, among others. To examine the performance of the fitted model, case-deletion and local influence techniques are provided to show its robust aspect against outlying and influential observations. This work is based in Ferreira, C. S., & Paula, G. A. (2017) <doi:10.1080/02664763.2016.1267124> but considering the SMN family.

## References

- Ferreira, C. S., & Paula, G. A. (2017). Estimation and diagnostic for skew-normal partially linear models. *Journal of Applied Statistics*, 44(16), 3033-3053.
- Ibache-Pulgar, G., Paula, G. A., & Cysneiros, F. J. A. (2013). Semiparametric additive models under symmetric distributions. *Test*, 22(1), 103-121.
- Ibache-Pulgar, G., & Paula, G. A. (2011). Local influence for Student-t partially linear models. *Computational Statistics & Data Analysis*, 55(3), 1462-1478.

## See Also

[CensReg.SMN](#)

## Examples

```
dtawage = get(data(PSID1976,package = "AER"))
y = dtawage$wage
cc = c(rep(0,428),rep(1,325))
tt = dtawage$exper
x = cbind(dtawage$education,dtawage$age, dtawage$hhours, dtawage$h wage, dtawage$tax,
dtawage$youngkids, dtawage$oldkids)

#Normal case by default with only 10 iterations
PCR.default1 = Cens.SMN.PCR(x=x, y=y, c=cc, cens="left",tt =tt,iter.max = 10,Diagnostic = FALSE)

## Not run:
#This may take few minutes
```

```
#Normal case by default with full (200) iterations
PCR.default2 = Cens.SMN.PCR(x=x, y=y, c=cc, cens="left",tt =tt)

#contaminated normal case
PCR.CN = Cens.SMN.PCR(x=x, y=y, c=cc, cens="left",tt =tt,type="NormalC",
nu = c(0.1,0.1),iter.max = 100)

## End(Not run)
```

---

Cens.SMN.PCR                      *Estimation and diagnostics for partially linear censored regression models*

---

## Description

Return the MPL estimates obtained through of ECME algorithm for partially linear regression models with censored data under scale-mixture of normal (SMN) distributions (some members are the normal, Student-t, slash and contaminated normal distribution). The types of censoring considered are left and right. Graphics for diagnostic analysis such as case-deletion and local influence techniques are provided to show its robust aspect against outlying and influential observations.

## Usage

```
Cens.SMN.PCR(x, y, c, cens = "left", tt, nu = NULL, error = 10^-6, iter.max = 200,
type = "Normal", alpha.FIX = TRUE, nu.FIX = TRUE, alpha.in = 10^-3, k = 1,
Diagnostic = TRUE, a = 2)
```

## Arguments

x	Matrix or vector of covariates.
y	Vector of responses.
c	Vector of censoring indicators. For each observation: 1 if censored and 0 if non-censored.
cens	'left' for left censoring and 'right' for righth censoring.
tt	Vector of values of a continuous covariate for the nonparametric component of the model.
nu	Initial value of the parameter of the SMN family. In the case of the Student-t and slash is a scalar, in the contaminated normal is a vector bidimensional.
error	The convergence maximum error. By default = $10^{-6}$ .
iter.max	The maximum number of iterations of the ECME algorithm. By default = 200.
type	Represents the type of distribution to be used in fitting: 'Normal' for normal, 'T' for Student-t, 'Slash' for slash and 'NormalC' for contaminated normal distribution respectively. By default = 'Normal'
alpha.FIX	TRUE or FALSE. Indicate if smoothing parameter will be estimated. By default = TRUE.

nu.FIX	TRUE or FALSE. Indicate if $\nu$ will be estimated. By default = TRUE.
alpha.in	Initial value of smoothing parameter.
k	For the local influence in explanatory variable perturbation, indicates the $k$ -th explanatory variable (assumed continuous) of the design matrix $X$ to be perturbed.
Diagnostic	TRUE or FALSE. Indicates if diagnostic graph should be built for the fitted model (index plot in local influence). By default = TRUE.
a	The value for $a$ considered in the benchmark value for the index plot in local influence: $M(0)_l > \bar{M}(0) + a * SM(0)$ .

### Details

We consider a partial linear model which belongs to the class of semiparametric regression models with vector of response  $Y = (Y_1, \dots, Y_n)$  and with errors  $\epsilon_i$  which are independent and identically distributed according to a SMN distribution. To be more precise,

$$Y_i = x_i^T \beta + n_i^T f + \epsilon_i,$$

for  $i = 1, \dots, n$ , where  $f = (f(t_1^0), \dots, f(t_r^n))^T$  is an  $r \times 1$  vector with  $t_1^0, \dots, t_r^n$  being the distinct and ordered values of  $t_i$ ;  $n_i$  is a  $r \times 1$  vector of incidence whose  $s$ -th element equals the indicator function  $I(t_i = t_s^0)$  for  $s = 1, \dots, r$ .

### Value

beta	ECME estimates for the parametric component.
sigma2	ECME estimates for the scale parameter.
Alpha	If alpha.FIX = FALSE, it returns the estimated value of the smoothing parameter, else returns the initial value assigned in alpha.in.
AIC	AIC criteria for model selection.
ff	ECME estimates for the nonparametric component.
yest	Predicted values of the model.
loglik	Value of the log-likelihood under the fitted model.
iter	Number of iterations of the ECME algorithm.
nu	If nu.FIX = FALSE, it returns the estimated value of $\nu$ parameter, else returns the initial value assigned in nu.
MI	Observed information matrix.
D	A list of objects for diagnostic analysis that contains: the Hessian matrix (Hessian), values for generalized Cook's distance (GD1) and the values of the conformal normal curvature for the following perturbation schemes: Case-weight (Curvature_W), scale (Curvature_S), explanatory variable (Curvature_E) and response variable (Curvature_R).

### Warning

For the contaminated normal case, if nu parameters were close to the bounds, i.e., close to 0 or 1, computational problems could arise.

**Note**

When `alpha.FIX = FALSE` the algorithm may take a long time to converge. The package estimates the value  $\nu$  in each iteration taking as an estimate the argument that maximizes the actual marginal log-likelihood function, already evaluated in the estimates of  $\beta$  and  $\sigma^2$ . The diagnostic analysis is performed considering the estimated final value of  $\theta$  obtained in the last iteration of the ECME algorithm.

**Author(s)**

Marcela Nunez Lemus, Christian E. Galarza, Larissa Avila Matos and Victor H. Lachos.

**References**

- Ferreira, C. S., & Paula, G. A. (2017). Estimation and diagnostic for skew-normal partially linear models. *Journal of Applied Statistics*, 44(16), 3033-3053.
- Ibacache-Pulgar, G., Paula, G. A., & Cysneiros, F. J. A. (2013). Semiparametric additive models under symmetric distributions. *Test*, 22(1), 103-121.
- Ibacache-Pulgar, G., & Paula, G. A. (2011). Local influence for Student-t partially linear models. *Computational Statistics & Data Analysis*, 55(3), 1462-1478.

**See Also**

[CensReg.SMN](#)

**Examples**

```
dtawage = get(data(PSID1976,package = "AER"))
y = dtawage$wage
cc = c(rep(0,428),rep(1,325))
tt = dtawage$exper
x = cbind(dtawage$education,dtawage$age, dtawage$hhours, dtawage$h wage, dtawage$tax,
dtawage$youngkids, dtawage$oldkids)

#Normal case by default with only 10 iterations
PCR.default1 = Cens.SMN.PCR(x=x, y=y, c=cc, cens="left",tt =tt,iter.max = 10,Diagnostic = FALSE)

## Not run:
#This may take few minutes
#Normal case by default with full (200) iterations
PCR.default2 = Cens.SMN.PCR(x=x, y=y, c=cc, cens="left",tt =tt)

#contaminated normal case
PCR.CN = Cens.SMN.PCR(x=x, y=y, c=cc, cens="left",tt =tt,type="NormalC",
nu = c(0.1,0.1),iter.max = 100)

## End(Not run)
```

# Index

- \* **censored**

- Cens.SMN.PCR, [3](#)

- PartCensReg-package, [2](#)

- \* **censoring**

- Cens.SMN.PCR, [3](#)

- PartCensReg-package, [2](#)

- \* **diagnostic**

- Cens.SMN.PCR, [3](#)

- PartCensReg-package, [2](#)

- \* **package**

- Cens.SMN.PCR, [3](#)

- PartCensReg-package, [2](#)

- \* **partially**

- Cens.SMN.PCR, [3](#)

- PartCensReg-package, [2](#)

- \* **partial**

- Cens.SMN.PCR, [3](#)

- PartCensReg-package, [2](#)

Cens.SMN.PCR, [3](#)

CensReg.SMN, [2](#), [5](#)

PartCensReg-package, [2](#)