Package 'vasicekreg'

October 12, 2022

Type Package

Title Regression Modeling Using Vasicek Distribution

Version 1.0.1

Date 2021-05-01

Author Josmar Mazucheli [aut, cre],
Bruna Alves [ctb]

Maintainer Josmar Mazucheli <jmazucheli@gmail.com>

Description Vasicek density, cumulative distribution, quantile functions and random deviate generation of Vasicek distribution. In addition, there are two functions for fitting the Generalized Additive Models for Location Scale and Shape introduced by Rigby and Stasinopoulos (2005, <doi:10.1111/j.1467-9876.2005.00510.x>). Some functions are written in C++ using 'Rcpp', developed by Eddelbuettel and Francois (2011, <doi:10.18637/jss.v040.i08>).

License MIT + file LICENSE

Encoding UTF-8

ByteCompile yes

LazyData true

LinkingTo Rcpp

Imports Rcpp, stats, gamlss, gamlss.dist, mvtnorm

Suggests testthat (>= 3.0.0)

Depends R (>= 3.6)

RoxygenNote 7.1.1

NeedsCompilation yes

Repository CRAN

Date/Publication 2021-05-27 17:40:02 UTC

R topics documented:

`vasicekreg-package` *Overview of the vasicekreg package*

Description

The **vasicekreg** package implements the probability density function, quantile function, cumulative distribution function and random number generation function for Vasicek distribution parameterized, either, as a function of its mean or its τ-th quantile, $0 < \tau < 1$. In addition, two gamlss frameworks for regression analysis are available. Some function are written in C++ using **Rcpp**.

Details

`bodyfat`: Body fat data set.

`VASIM`: For mean modeling (con/in)ditional on covariate(s).

`VASIQ`: For quantile modeling (con/in)ditional on covariate(s).

Author(s)

Josmar Mazucheli <jmazucheli@gmail.com>

Bruna Alves <pg402900@uem.br>

`bodyfat` *Percentage of body fat data set*

Description

The body fat percentage of individuals assisted in a public hospital in Curitiba, Paraná, Brazil.

Usage

```
data(bodyfat, package = "vasicekreg")
```

Format

A data-frame with 298 observations and 9 columns:

- `ARMS`: arms fat percentage.
- `LEGS`: legs fat percentage.
- `BODY`: body fat percentage.
- `ANDROID`: android fat percentage.
- `GYNECOID`: ginecoid fat percentage.
- `AGE`: age of individuals.
- `BMI`: body mass index.
- `SEX`: 1 for female, 2 for male.
- `IPAQ`: 0 for IPAQ = sedentary, 1 for IPAQ = insufficiently active and 2 for IPAQ = active.

Author(s)

Josmar Mazucheli <jmazucheli@gmail.com>

Bruna Alves <pg402900@uem.br>

Source

http://www.leg.ufpr.br/doku.php/publications:papercompanions:multquasibeta

References

Mazucheli, J., Leiva, V., Alves, B., and Menezes A. F. B., (2021). A new quantile regression for modeling bounded data under a unit Birnbaum-Saunders distribution with applications in medicine and politics. *Symmetry*, **13**(4) 1–21.

Petterle, R. R., Bonat, W. H., Scarpin, C. T., Jonasson, T., and Borba, V. Z. C., (2020). Multivariate quasi-beta regression models for continuous bounded data. *The International Journal of Biostatistics*, 1–15, (preprint).

Examples

```
data(bodyfat, package = "vasicekreg")

bodyfat$BMI <- bodyfat$BMI / 100
bodyfat$SEX <- as.factor(bodyfat$SEX)
bodyfat$IPAQ<- as.factor(bodyfat$IPAQ)

library(gamlss)

# mean fit
fitmean.logit <- gamlss(ARMS ~ AGE + BMI + SEX + IPAQ,
data = bodyfat, family = VASIM(mu.link = "logit", sigma.link = "logit"))

fitmean.probit <- gamlss(ARMS ~ AGE + BMI + SEX + IPAQ,
data = bodyfat, family = VASIM(mu.link = "probit", sigma.link = "logit"))

# quantile fit - tau = 0.5

tau <- 0.50
fitquant.logit <- gamlss(ARMS ~ AGE + BMI + SEX + IPAQ, data = bodyfat,
family = VASIQ(mu.link = "logit", sigma.link = "logit"))

fittaus <- lapply(c(0.10, 0.25, 0.50, 0.75, 0.90), function(Tau)
{
 tau <<- Tau;
 gamlss(ARMS ~ AGE + BMI + SEX + IPAQ, data = bodyfat,
 family = VASIQ(mu.link = "logit", sigma.link = "logit"))
})

sapply(fittaus, summary, USE.NAMES = TRUE)
```

VASIM *The Vasicek distribution - mean parameterization*

Description

The function VASIM() define the Vasicek distribution for a gamlss.family object to be used in GAMLSS fitting. VASIM() has mean equal to the parameter mu and sigma as shape parameter. The functions dVASIM, pVASIM, qVASIM and rVASIM define the density, distribution function, quantile function and random generation for Vasicek distribution.

Usage

```
dVASIM(x, mu, sigma, log = FALSE)

pVASIM(q, mu, sigma, lower.tail = TRUE, log.p = FALSE)

qVASIM(p, mu, sigma, lower.tail = TRUE, log.p = FALSE)

rVASIM(n, mu, sigma)

VASIM(mu.link = "logit", sigma.link = "logit")
```

Arguments

x, q	vector of quantiles on the (0,1) interval.
mu	vector of the mean parameter values.
sigma	vector of shape parameter values.
log, log.p	logical; If TRUE, probabilities p are given as log(p).
lower.tail	logical; If TRUE, (default), $P(X \leq x)$ are returned, otherwise $P(X > x)$.
p	vector of probabilities.
n	number of observations. If length(n) > 1, the length is taken to be the number required.
mu.link	the mu link function with default logit.
sigma.link	the sigma link function with default logit.

Details

Probability density function

$$f(x \mid \mu, \sigma) = \sqrt{\frac{1-\sigma}{\sigma}} \exp\left\{\frac{1}{2}\left[\Phi^{-1}(x)^2 - \left(\frac{\Phi^{-1}(x)\sqrt{1-\sigma} - \Phi^{-1}(\mu)}{\sqrt{\sigma}}\right)^2\right]\right\}$$

Cumulative distribution function

$$F(x \mid \mu, \sigma) = \Phi\left(\frac{\Phi^{-1}(x)\sqrt{1-\sigma} - \Phi^{-1}(\mu)}{\sqrt{\sigma}}\right)$$

Quantile function

$$Q(\tau \mid \mu, \sigma) = F^{-1}(\tau \mid \mu, \sigma) = \Phi\left(\frac{\Phi^{-1}(\mu) + \Phi^{-1}(\tau)\sqrt{\sigma}}{\sqrt{1-\sigma}}\right)$$

Expected value

$$E(X) = \mu$$

Variance

$$Var(X) = \Phi_2\left(\Phi^{-1}(\mu), \Phi^{-1}(\mu), \sigma\right) - \mu^2$$

where $0 < (x, \mu, \tau, \sigma) < 1$ and $\Phi_2(\cdot)$ is the probability distribution function for the standard bivariate normal distribution with correlation σ.

Value

`VASIM()` return a gamlss.family object which can be used to fit a Vasicek distribution by gamlss() function.

Note

Note that for `VASIQ()`, mu is the τ-th quantile and sigma a shape parameter. The `gamlss` function is used for parameters estimation.

Author(s)

Josmar Mazucheli <jmazucheli@gmail.com>

Bruna Alves <pg402900@uem.br>

References

Hastie, T. J. and Tibshirani, R. J. (1990). *Generalized Additive Models*. Chapman and Hall, London.

Mazucheli, J., Alves, B. and Korkmaz, M. C. (2021). The Vasicek quantile regression model. *(under review)*.

Rigby, R. A. and Stasinopoulos, D. M. (2005). Generalized additive models for location, scale and shape (with discussion). *Applied. Statistics*, **54**(3), 507–554.

Rigby, R. A., Stasinopoulos, D. M., Heller, G. Z. and De Bastiani, F. (2019). *Distributions for modeling location, scale, and shape: Using GAMLSS in R*. Chapman and Hall/CRC.

Stasinopoulos, D. M. and Rigby, R. A. (2007) Generalized additive models for location scale and shape (GAMLSS) in R. *Journal of Statistical Software*, **23**(7), 1–45.

Stasinopoulos, D. M., Rigby, R. A., Heller, G., Voudouris, V. and De Bastiani F. (2017) *Flexible Regression and Smoothing: Using GAMLSS in R*, Chapman and Hall/CRC.

Vasicek, O. A. (1987). Probability of loss on loan portfolio. *KMV Corporation*.

Vasicek, O. A. (2002). The distribution of loan portfolio value. *Risk*, **15**(12), 1–10.

See Also

`VASIQ`, `pmvnorm`.

Examples

```
set.seed(123)
x <- rVASIM(n = 1000, mu = 0.50, sigma = 0.69)
R <- range(x)
S <- seq(from = R[1], to = R[2], length.out = 1000)

hist(x, prob = TRUE, main = 'Vasicek')
lines(S, dVASIM(x = S, mu = 0.50, sigma = 0.69), col = 2)

plot(ecdf(x))
lines(S, pVASIM(q = S, mu = 0.50, sigma = 0.69), col = 2)

plot(quantile(x, probs = S), type = "l")
lines(qVASIM(p = S, mu = 0.50, sigma = 0.69), col = 2)

library(gamlss)
set.seed(123)
data <- data.frame(y =  rVASIM(n = 100, mu = 0.5, sigma = 0.69))

fit <- gamlss(y ~ 1, data = data, mu.link = 'logit', sigma.link = 'logit', family = VASIM)
1 /(1 + exp(-fit$mu.coefficients))
1 /(1 + exp(-fit$sigma.coefficients))

set.seed(123)
n <- 100
x <- rbinom(n, size = 1, prob = 0.5)
eta <- 0.5 + 1 * x;
mu <- 1 / (1 + exp(-eta));
sigma <- 0.1;
y <- rVASIM(n, mu, sigma)
data <- data.frame(y, x)

fit <- gamlss(y ~ x, data = data, family = VASIM, mu.link = 'logit', sigma.link = 'logit');
```

VASIQ *The Vasicek distribution - quantile parameterization*

Description

The function VASIQ() define the Vasicek distribution for a gamlss.family object to be used in GAMLSS fitting. VASIQ() has the τ-th quantile equal to the parameter mu and sigma as shape parameter. The functions dVASIQ, pVASIQ, qVASIQ and rVASIQ define the density, distribution function, quantile function and random generation for Vasicek distribution.

Usage

```
dVASIQ(x, mu, sigma, tau = 0.5, log = FALSE)
```

```
pVASIQ(q, mu, sigma, tau = 0.5, lower.tail = TRUE, log.p = FALSE)

qVASIQ(p, mu, sigma, tau = 0.5, lower.tail = TRUE, log.p = FALSE)

rVASIQ(n, mu, sigma, tau = 0.5)

VASIQ(mu.link = "logit", sigma.link = "logit")
```

Arguments

x, q	vector of quantiles on the (0,1) interval.
mu	vector of the τ-th quantile parameter values.
sigma	vector of shape parameter values.
tau	the τ-th fixed quantile in [d-p-q-r]-VASIQ function.
log, log.p	logical; If TRUE, probabilities p are given as log(p).
lower.tail	logical; If TRUE, (default), $P(X \leq x)$ are returned, otherwise $P(X > x)$.
p	vector of probabilities.
n	number of observations. If length(n) > 1, the length is taken to be the number required.
mu.link	the mu link function with default logit.
sigma.link	the sigma link function with default logit.

Details

Probability density function

$$f\left(x \mid \mu, \sigma, \tau\right) = \sqrt{\tfrac{1-\sigma}{\sigma}} \exp\left\{\frac{1}{2}\left[\Phi^{-1}\left(x\right)^2 - \left(\frac{\sqrt{1-\sigma}\left[\Phi^{-1}\left(x\right) - \Phi^{-1}\left(\mu\right)\right] - \sqrt{\sigma}\,\Phi^{-1}\left(\tau\right)}{\sqrt{\sigma}}\right)^2\right]\right\}$$

Cumulative distribution function

$$F\left(x \mid \mu, \sigma, \tau\right) = \Phi\left[\frac{\sqrt{1-\sigma}\left[\Phi^{-1}\left(x\right) - \Phi^{-1}\left(\mu\right)\right] - \sqrt{\sigma}\,\Phi^{-1}\left(\tau\right)}{\sqrt{\sigma}}\right]$$

where $0 < (x, \mu, \tau, \sigma) < 1$, μ is the τ-th quantile and σ is the shape parameter.

Value

VASIQ() return a gamlss.family object which can be used to fit a Vasicek distribution by gamlss() function.

Note

Note that for VASIQ(), mu is the τ-th quantile and sigma a shape parameter. The gamlss function is used for parameters estimation.

Author(s)

Josmar Mazucheli <jmazucheli@gmail.com>

Bruna Alves <pg402900@uem.br>

References

Hastie, T. J. and Tibshirani, R. J. (1990). *Generalized Additive Models*. Chapman and Hall, London.

Mazucheli, J., Alves, B. and Korkmaz, M. C. (2021). The Vasicek quantile regression model. *(under review)*.

Rigby, R. A. and Stasinopoulos, D. M. (2005). Generalized additive models for location, scale and shape (with discussion). *Applied. Statistics*, **54**(3), 507–554.

Rigby, R. A., Stasinopoulos, D. M., Heller, G. Z. and De Bastiani, F. (2019). *Distributions for modeling location, scale, and shape: Using GAMLSS in R*. Chapman and Hall/CRC.

Stasinopoulos, D. M. and Rigby, R. A. (2007) Generalized additive models for location scale and shape (GAMLSS) in R. *Journal of Statistical Software*, **23**(7), 1–45.

Stasinopoulos, D. M., Rigby, R. A., Heller, G., Voudouris, V. and De Bastiani F. (2017) *Flexible Regression and Smoothing: Using GAMLSS in R*, Chapman and Hall/CRC.

Vasicek, O. A. (1987). Probability of loss on loan portfolio. *KMV Corporation*.

Vasicek, O. A. (2002). The distribution of loan portfolio value. *Risk*, **15**(12), 1–10.

See Also

VASIM.

Examples

```
set.seed(123)
x <- rVASIQ(n = 1000, mu = 0.50, sigma = 0.69, tau = 0.50)
R <- range(x)
S <- seq(from = R[1], to = R[2], length.out = 1000)

hist(x, prob = TRUE, main = 'Vasicek')
lines(S, dVASIQ(x = S, mu = 0.50, sigma = 0.69, tau = 0.50), col = 2)

plot(ecdf(x))
lines(S, pVASIQ(q = S, mu = 0.50, sigma = 0.69, tau = 0.50), col = 2)

plot(quantile(x, probs = S), type = "l")
lines(qVASIQ(p = S, mu = 0.50, sigma = 0.69, tau = 0.50), col = 2)

library(gamlss)
set.seed(123)
data <- data.frame(y = rVASIQ(n = 100, mu = 0.50, sigma = 0.69, tau = 0.50))

tau <- 0.5
fit <- gamlss(y ~ 1, data = data, family = VASIQ(mu.link = 'logit', sigma.link = 'logit'))
1 /(1 + exp(-fit$mu.coefficients)); 1 /(1 + exp(-fit$sigma.coefficients))
```

```
set.seed(123)
n <- 100
x <- rbinom(n, size = 1, prob = 0.5)
eta <- 0.5 + 1 * x;
mu <- 1 / (1 + exp(-eta));
sigma <- 0.5;
y <- rVASIQ(n, mu, sigma, tau = 0.5)
data <- data.frame(y, x, tau = 0.5)

tau <- 0.5;
fit <- gamlss(y ~ x, data = data, family = VASIQ)

fittaus <- lapply(c(0.10, 0.25, 0.50, 0.75, 0.90), function(Tau)
{
 tau <<- Tau;
 gamlss(y ~ x, data = data, family = VASIQ)
})

sapply(fittaus, summary)
```

Index