Package ‘zic’

August 22, 2017

Version 0.9.1
Date 2017-08-22
Title Bayesian Inference for Zero-Inflated Count Models
Author Markus Jochmann <markus.jochmann@ncl.ac.uk>
Maintainer Markus Jochmann <markus.jochmann@ncl.ac.uk>
Description Provides MCMC algorithms for the analysis of zero-inflated count models. The case of stochastic search variable selection (SVS) is also considered. All MCMC samplers are coded in C++ for improved efficiency. A data set considering the demand for health care is provided.
License GPL (>= 2)
Depends R (>= 3.0.2)
Imports Rcpp (>= 0.11.0), coda (>= 0.14-2)
LinkingTo Rcpp, RcppArmadillo
NeedsCompilation yes
Repository CRAN
Date/Publication 2017-08-22 12:26:54 UTC

R topics documented:

  docvisits .................................................. 2
  zic ............................................................ 3
  zic.svs ....................................................... 5

Index 8
Description

This data set gives the number of doctor visits in the last three months for a sample of German male individuals in 1994. The data set is taken from Riphahn et al. (2003) and is a subsample of the German Socioeconomic Panel (SOEP). In contrast to Riphahn et al. (2003) only male individuals from the last wave are considered. See Jochmann (2013) for further details.

Usage

data(docvisits)

Format

This data frame contains 1812 observations on the following 22 variables:

- **docvisits**: number of doctor visits in last 3 months
- **age**: age
- **agesq**: age squared / 1000
- **age30**: 1 if age >= 30
- **age35**: 1 if age >= 35
- **age40**: 1 if age >= 40
- **age45**: 1 if age >= 45
- **age50**: 1 if age >= 50
- **age55**: 1 if age >= 55
- **age60**: 1 if age >= 60
- **health**: health satisfaction, 0 (low) - 10 (high)
- **handicap**: 1 if handicapped, 0 otherwise
- **hdegree**: degree of handicap in percentage points
- **married**: 1 if married, 0 otherwise
- **schooling**: years of schooling
- **hhincome**: household monthly net income, in German marks / 1000
- **children**: 1 if children under 16 in the household, 0 otherwise
- **self**: 1 if self employed, 0 otherwise
- **civil**: 1 if civil servant, 0 otherwise
- **bluec**: 1 if blue collar employee, 0 otherwise
- **employed**: 1 if employed, 0 otherwise
- **public**: 1 if public health insurance, 0 otherwise
- **addon**: 1 if add-on insurance, 0 otherwise
References


---

**Description**

*zic* fits zero-inflated count models via Markov chain Monte Carlo methods.

**Usage**

```r
zic(formula, data, a0, b0, c0, d0, e0, f0,
    nNburnin, nNmcmc, nNthin, tune = 1.0, scale = TRUE)
```

**Arguments**

- `formula`: A symbolic description of the model to be fit specifying the response variable and covariates.
- `data`: A data frame in which to interpret the variables in `formula`.
- `a0`: The prior variance of $\alpha$.
- `b0`: The prior variance of $\beta_j$.
- `c0`: The prior variance of $\gamma$.
- `d0`: The prior variance of $\delta_j$.
- `e0`: The shape parameter for the inverse gamma prior on $\sigma^2$.
- `f0`: The inverse scale parameter the inverse gamma prior on $\sigma^2$.
- `nNburnin`: Number of burn-in iterations of the sampler.
- `nNmcmc`: Number of iterations of the sampler.
- `nNthin`: Thinning interval.
- `tune`: Tuning parameter of Metropolis-Hastings step.
- `scale`: If true, all covariates (except binary variables) are rescaled by dividing by their respective standard errors.
The considered zero-inflated count model is given by

\[
y_i^* \sim \text{Poisson}(\exp(\eta_i^*)), \\
\eta_i^* = \alpha + x_i' \beta + \varepsilon_i, \ \varepsilon_i \sim N(0, \sigma^2), \\
d_i^* = \gamma + x_i' \delta + \nu_i, \ \nu_i \sim N(0, 1), \\
y_i = 1(d_i^* > 0) y_i^*,
\]

where \(y_i\) and \(x_i\) are observed. The assumed prior distributions are

\[
\alpha \sim N(0, a_0), \\
\beta_k \sim N(0, b_0), \quad k = 1, \ldots, K, \\
\gamma \sim N(0, c_0), \\
\delta_k \sim N(0, d_0), \quad k = 1, \ldots, K, \\
\sigma^2 \sim \text{Inv-Gamma}(e_0, f_0).
\]

The sampling algorithm described in Jochmann (2013) is used.

**Value**

A list containing the following elements:

- `alpha`: Posterior draws of \(\alpha\) (coda mcmc object).
- `beta`: Posterior draws of \(\beta\) (coda mcmc object).
- `gamma`: Posterior draws of \(\gamma\) (coda mcmc object).
- `delta`: Posterior draws of \(\delta\) (coda mcmc object).
- `sigma2`: Posterior draws of \(\sigma^2\) (coda mcmc object).
- `acc`: Acceptance rate of the Metropolis-Hastings step.

**References**


**Examples**

```r
## not run:
data( docvisits )
mdl <- docvisits ~ age + agesq + health + handicap + hdegree + married + schooling + hhincome + children + self + civil + bluc + employed + public + addon
post <- zic( f, docvisits, 10N0, 10N0, 10N0, 10N0, 1N0, 1N0, 1000, 10000, 10, 1.0, TRUE )
## end(not run)```
Description

zic.svs applies SVS to zero-inflated count models.

Usage

zic.svs(formula, data,
    a0, g0.beta, h0.beta, nu0.beta, r0.beta, s0.beta, e0, f0,
    c0, g0.delta, h0.delta, nu0.delta, r0.delta, s0.delta,
    n.burnin, n.mcmc, n.thin, tune = 1.0, scale = TRUE)

Arguments

formula A symbolic description of the model to be fit specifying the response variable
    and covariates.
data A data frame in which to interpret the variables in formula.
a0 The prior variance of α.
g0.beta The shape parameter for the inverse gamma prior on κ^β_k.
h0.beta The inverse scale parameter for the inverse gamma prior on κ^β_k.
u0.beta Prior parameter for the spike of the hypervariances for the β_k.
r0.beta Prior parameter of ω^β.
s0.beta Prior parameter of ω^β.
e0 The shape parameter for the inverse gamma prior on σ^2.
f0 The inverse scale parameter the inverse gamma prior on σ^2.
c0 The prior variance of γ.
g0.delta The shape parameter for the inverse gamma prior on κ^δ_k.
h0.delta The inverse scale parameter for the inverse gamma prior on κ^δ_k.
u0.delta Prior parameter for the spike of the hypervariances for the δ_k.
r0.delta Prior parameter of ω^δ.
s0.delta Prior parameter of ω^δ.
n.burnin Number of burn-in iterations of the sampler.
n.mcmc Number of iterations of the sampler.
n.thin Thinning interval.
tune Tuning parameter of Metropolis-Hastings step.
scale If true, all covariates (except binary variables) are rescaled by dividing by their
    respective standard errors.
Details

The considered zero-inflated count model is given by

\[ y_i^* \sim \text{Poisson}[\exp(\eta_i^*)], \]
\[ \eta_i^* = \alpha + x_i' \beta + \epsilon_i, \quad \epsilon_i \sim N(0, \sigma^2), \]
\[ d_i^* = \gamma + x_i' \delta + \nu_i, \quad \nu_i \sim N(0, 1), \]
\[ y_i = 1(d_i^* > 0)y_i^*, \]

where \( y_i \) and \( x_i \) are observed. The assumed prior distributions are

\[ \alpha \sim N(0, a_0), \]
\[ \beta_k \sim N(0, \tau_k^\beta \kappa_k^\beta), \quad k = 1, \ldots, K, \]
\[ \kappa_j^\beta \sim \text{Inv-Gamma}(g_{0}^\beta, h_{0}^\beta), \]
\[ \tau_k^\beta \sim (1 - \omega^\beta)\delta_{\nu, \delta}^\beta + \omega^\beta \delta_1, \]
\[ \omega^\beta \sim \text{Beta}(r_{\beta}^\beta, s_{\beta}^\beta), \]
\[ \gamma \sim N(0, c_0), \]
\[ \delta_k \sim N(0, \tau_k^\delta \kappa_k^\delta), \quad k = 1, \ldots, K, \]
\[ \kappa_k^\delta \sim \text{Inv-Gamma}(g_{0}^\delta, h_{0}^\delta), \]
\[ \tau_k^\delta \sim (1 - \omega^\delta)\delta_{\nu, \delta}^\delta + \omega^\delta \delta_1, \]
\[ \omega^\delta \sim \text{Beta}(r_{\delta}^\delta, s_{\delta}^\delta), \]
\[ \sigma^2 \sim \text{Inv-Gamma}(e_0, f_0). \]

The sampling algorithm described in Jochmann (2013) is used.

Value

A list containing the following elements:

- alpha: Posterior draws of \( \alpha \) (coda mcmc object).
- beta: Posterior draws of \( \beta \) (coda mcmc object).
- gamma: Posterior draws of \( \gamma \) (coda mcmc object).
- delta: Posterior draws of \( \delta \) (coda mcmc object).
- sigma2: Posterior draws of \( \sigma^2 \) (coda mcmc object).
- I.beta: Posterior draws of indicator whether \( \tau_j^\beta \) is one (coda mcmc object).
- I.delta: Posterior draws of indicator whether \( \tau_j^\delta \) is one (coda mcmc object).
- omega.beta: Posterior draws of \( \omega^\beta \) (coda mcmc object).
- omega.delta: Posterior draws of \( \omega^\delta \) (coda mcmc object).
- acc: Acceptance rate of the Metropolis-Hastings step.
References


Examples

```r
## not run:
data( docvisits )
mdl <- docvisits ~ age + agesq + health + handicap + hdegree + married + schooling + hhincome + children + self + civil + bluec + employed + public + addon
post <- zic.svs( mdl, docvisits, 10.0, 5.0, 5.0, 1.0e-04, 2.0, 2.0, 1.0, 1.0, 10.0, 5.0, 5.0, 1.0e-04, 2.0, 2.0, 1000, 10000, 10, 1.0, TRUE )
## end(not run)
```
Index

*Topic datasets
  docvisits, 2

docvisits, 2

zic, 3
zic.svs, 5