

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@nc1.ac.uk>

Maintainer Markus Jochmann <markus.jochmann@nc1.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

 docvisits

Demand for Health Care Data

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 \geq 30

age35 1 if age \geq 35

age40 1 if age \geq 40

age45 1 if age \geq 45

age50 1 if age \geq 50

age55 1 if age \geq 55

age60 1 if age \geq 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

- Jochmann, M. (2013). “What Belongs Where? Variable Selection for Zero-Inflated Count Models with an Application to the Demand for Health Care”, *Computational Statistics*, 28, 1947–1964.
- Riphahn, R. T., Wambach, A., Million, A. (2003). “Incentive Effects in the Demand for Health Care: A Bivariate Panel Count Data Estimation”, *Journal of Applied Econometrics*, 18, 387–405.
- Wagner, G. G., Frick, J. R., Schupp, J. (2007). “The German Socio-Economic Panel Study (SOEP) – Scope, Evolution and Enhancements”, *Schmollers Jahrbuch*, 127, 139–169.

zic

Bayesian Inference for Zero-Inflated Count Models

Description

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

Usage

```
zic(formula, data, a0, b0, c0, d0, e0, f0,
    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 .
b0	The prior variance of β_j .
c0	The prior variance of γ .
d0	The prior variance of δ_j .
e0	The shape parameter for the inverse gamma prior on σ^2 .
f0	The inverse scale parameter the inverse gamma prior on σ^2 .
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

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

where y_i and x_i are observed. The assumed prior distributions are

$$\begin{aligned} \alpha &\sim N(0, a_0), \\ \beta_k &\sim N(0, b_0), \quad k = 1, \dots, K, \\ \gamma &\sim N(0, c_0), \\ \delta_k &\sim N(0, d_0), \quad k = 1, \dots, K, \\ \sigma^2 &\sim \text{Inv-Gamma}(e_0, f_0). \end{aligned}$$

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

Value

A list containing the following elements:

alpha	Posterior draws of α (coda mcmc object).
beta	Posterior draws of β (coda mcmc object).
gamma	Posterior draws of γ (coda mcmc object).
delta	Posterior draws of δ (coda mcmc object).
sigma2	Posterior draws of σ^2 (coda mcmc object).
acc	Acceptance rate of the Metropolis-Hastings step.

References

Jochmann, M. (2013). “What Belongs Where? Variable Selection for Zero-Inflated Count Models with an Application to the Demand for Health Care”, *Computational Statistics*, 28, 1947–1964.

Examples

```
## Not run:
data( docvisits )
mdl <- docvisits ~ age + agesq + health + handicap + hdegree + married + schooling +
             hhincome + children + self + civil + bluec + employed + public + addon
post <- zic( f, docvisits, 10.0, 10.0, 10.0, 10.0, 1.0, 1.0, 1000, 10000, 10, 1.0, TRUE )
## End(Not run)
```

zic.svs

*SVS for Zero-Inflated Count Models***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_k^β .
h0.beta	The inverse scale parameter for the inverse gamma prior on κ_k^β .
nu0.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_k^δ .
h0.delta	The inverse scale parameter for the inverse gamma prior on κ_k^δ .
nu0.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

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

where y_i and x_i are observed. The assumed prior distributions are

$$\begin{aligned} \alpha &\sim \text{N}(0, a_0), \\ \beta_k &\sim \text{N}(0, \tau_k^\beta \kappa_k^\beta), \quad k = 1, \dots, K, \\ \kappa_j^\beta &\sim \text{Inv-Gamma}(g_0^\beta, h_0^\beta), \\ \tau_k^\beta &\sim (1 - \omega^\beta) \delta_{\nu_0^\beta} + \omega^\beta \delta_1, \\ \omega^\beta &\sim \text{Beta}(r_0^\beta, s_0^\beta), \\ \gamma &\sim \text{N}(0, e_0), \\ \delta_k &\sim \text{N}(0, \tau_k^\delta \kappa_k^\delta), \quad k = 1, \dots, K, \\ \kappa_k^\delta &\sim \text{Inv-Gamma}(g_0^\delta, h_0^\delta), \\ \tau_k^\delta &\sim (1 - \omega^\delta) \delta_{\nu_0^\delta} + \omega^\delta \delta_1, \\ \omega^\delta &\sim \text{Beta}(r_0^\delta, s_0^\delta), \\ \sigma^2 &\sim \text{Inv-Gamma}(e_0, f_0). \end{aligned}$$

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

Value

A list containing the following elements:

alpha	Posterior draws of α (coda mcmc object).
beta	Posterior draws of β (coda mcmc object).
gamma	Posterior draws of γ (coda mcmc object).
delta	Posterior draws of δ (coda mcmc object).
sigma2	Posterior draws of σ^2 (coda mcmc object).
I.beta	Posterior draws of indicator whether τ_j^β is one (coda mcmc object).
I.delta	Posterior draws of indicator whether τ_j^δ is one (coda mcmc object).
omega.beta	Posterior draws of ω^β (coda mcmc object).
omega.delta	Posterior draws of ω^δ (coda mcmc object).
acc	Acceptance rate of the Metropolis-Hastings step.

References

Jochmann, M. (2013). “What Belongs Where? Variable Selection for Zero-Inflated Count Models with an Application to the Demand for Health Care”, *Computational Statistics*, 28, 1947–1964.

Examples

```
## Not run:
data( docvisits )
mdl <- docvisits ~ age + agesq + health + handicap + hdegree + married + schooling +
  hhincome + children + self + civil + bluec + employed + public + addon
post <- zic.ssvs( 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](#)