Package ‘endogMNP’

February 19, 2015

Version 0.2-1

Date 2010-03-30

Title R Package for Fitting Multinomial Probit Models with Endogenous Selection

Author Lane F. Burgette <`lb131@stat.duke.edu`>

Maintainer Lane F. Burgette <`lb131@stat.duke.edu`>

Depends R (>= 2.1), utils

Description endogMNP is an R package that fits a Bayesian multinomial probit model with endogenous selection, which is sometimes called an endogenous switching model. This can be used to model discrete choice data when respondents select themselves into one of several groups. This package is based on the MNP package by Kosuke Imai and David A. van Dyk. This package modifies their code.

URL `http://www.r-project.org`, `http://fds.duke.edu/db/aas/stat/faculty/lb131`

LazyLoad yes

Repository CRAN

Date/Publication 2010-04-02 17:08:05

NeedsCompilation yes

R topics documented:

`endogMNP` *Fitting the Multinomial Probit Model with Endogenous Selection via Markov chain Monte Carlo*

Description

`endogMNP` is used to fit a Bayesian multinomial probit model with endogenous selection or switching via Markov chain Monte Carlo. The computation uses the efficient partial marginal data augmentation algorithm that is developed by Burgette and Nordheim (2009), which is an extension of the sampler of Imai and van Dyk (2005).

Usage

```
endogMNP(selForm, outForm, dataSet = parent.frame(),
                selBase = NULL, outBase = NULL, latent=FALSE,
                invcdf = FALSE, n.draws = 5000, p.var = "Inf",
                p.df = n.dim+1, p.scale = 1, coef.start = 0,
                cov.start = 1, burnin = 0, thin = 0, verbose = FALSE,
minConst=TRUE, trace=TRUE)
```

Arguments

`selForm`	A symbolic description of the selection model portion of the model where the left-hand side indicates the category into which each observation has been selected.
`outForm`	A symbolic description of the outcome equation. The left-hand side is the response of interest.
`dataSet`	An optional data frame in which to interpret the variables in `selForm` and `outForm`. The default is the environment in which `endogMNP` is called.
`selBase`	The name of the base category for the selection equation. The default is the lowest level of the selection variable.
`outBase`	The name of the base category for the outcome equation. The default is the lowest level of the response variable.
`latent`	Logical. Store latent vectors?
`invcdf`	Logical. If `TRUE`, then the inverse cdf method is used for truncated normal sampling. If `FALSE`, then the rejection sampling method is used. The default is `FALSE`.
`n.draws`	A positive integer. The number of MCMC draws. The default is `5000`.
`p.var`	A positive definite matrix. The prior variance of the coefficients. A scalar input can set the prior variance to the diagonal matrix whose diagonal element is equal to that value. The default is `"Inf"`, which represents an improper noninformative prior distribution on the coefficients.
`p.df`	A positive integer greater than the dimension of the implied covariance matrix. The prior degrees of freedom parameter for the covariance matrix. The default is the dimension of the covariance matrix plus one.

`p.scale`	A block-diagonal, positive definite matrix whose leading diagonal elements are set to `1`. The first block has the size of the selection covariance matrix, and the other blocks are of the size of the outcome covariance matrices. This is the prior scale matrix for the covariance matrix. A scalar input can be used to set the scale matrix to a diagonal matrix with diagonal elements equal to the scalar input value, except that the first element in each block is set to one. The default is `1`.
`coef.start`	A vector. The starting values for the coefficients. A scalar input sets the starting values for all the coefficients equal to that value. The default is `0`.
`cov.start`	A positive definite matrix. The first elements in the blocks determined by the selection and outcome sizes should be set to `1`. This is the starting value for the covariance matrix. The leading diagonal elements are set to 1 if they are not already. A scalar input can be used to set the starting value to a diagonal matrix with diagonal elements equal to the scalar input value, except those of the leading diagonal elements are set to one. The default is `1`.
`burnin`	A positive integer. The burn-in interval for the Markov chain. It is the number of initial Gibbs draws that should not be stored. The default is `0`.
`thin`	A positive integer. The thinning interval for the Markov chain. This is the number of Gibbs draws between the recorded values that are skipped. The default is `0`.
`verbose`	Logical. If `TRUE`, messages along with a progress report of the Gibbs sampling are printed. The default is `FALSE`.
`minConst`	Logical. If `TRUE`, the covariance matrix is minimally constrained, which is to say only enough of the elements of the covariance matrix are constrained to be 1 in order to make the model identifiable. If `FALSE`, all of the diagonal elements of the covariance matrix are constrained to be 1. This can improve convergence in some cases.
`trace`	Logical. If `TRUE`, the trace restriction will be used to identify the model. This argument is ignored if `minConst=FALSE`.

Details

To fit the multinomial probit model when only the most preferred choice is observed, use the syntax for the formula, `outForm = y1 ~ x1 + x2`, where `y` is a factor variable indicating the most preferred choice and `x1` and `x2` are individual-specific covariates. The selection process is modeled by `selForm = y2 ~ x2 + x3` where `y2` contains the selection category. The y1 and y2 variables may contain missing values (coded as `NA`), however the x variables must be fully observed. Further, all but one of the selection categories must have at least some observed outcomes. (I.e., for a selection model we should observe the outcome for all selection groups except one.)

Value

An object of class `endogMNP` containing the following elements:

`call`	The matched call.
`param`	A matrix of the Gibbs draws for each parameter; i.e., the coefficients and covariance matrix. For the covariance matrix, the elements on or above the diagonal are returned.

`x`	The matrix of covariates.
`y`	The vector matrix of the selection and response variables.
`n.dim`	The number of columns in the covariance matrix.
`n.obs`	The number of observations.
`coefnames`	The names of the coefficients.
`W`	The three dimensional array of the latent variable, W. The first dimension represents the alternatives, and the second dimension indexes the observations. The third dimension represents the Gibbs draws. Note that the latent variables for the base categories are set to 0, and therefore omitted from the output.
`p.scale`	The prior scale of the covariance.
`n.cov`	The number of covariates.
`nu0`	The prior degrees of freedom.
`p.var`	The prior variance.
`n.param`	The number of parameters in the fit model.
`minConst`	Indicator of whether the covariance matrix was minimally constrained.
`n.dim1`	The number of dimensions for the selection equation.
`n.dim2`	The number of dimensions of each outcome equation.
`n.rep`	The number of stored Gibbs iterations.
`selForm`	The symbolic selection equation formula.
`outForm`	The symbolic outcome equation formula.
`dataSet`	The data-set.
`selBase`	The base category for the selection model.
`outBase`	The base category for the outcome.

Author(s)

Lane F. Burgette, Department Statistical Science, Duke University, <`lb131@stat.duke.edu`>.

References

Burgette, Lane F. and Erik V. Nordheim (2009). “A Full Gibbs Sampler for a Multinomial Probit Model with Endogeneity.” Available on request from the first author.

Imai, Kosuke and David A. van Dyk (2005). “A Bayesian Analysis of the Multinomial Probit Model Using Data Augmentation.” Journal of Econometrics. 124(2): 311-34.

`summary.endogMNP` *Summarizing the results for the Multinomial Probit Models with Endogeneity*

Description

`summary` method for class `endogMNP`.

Usage

```
  ## S3 method for class 'endogMNP'
summary(object, CI=c(2.5, 97.5), discard=1, ...)

  ## S3 method for class 'summary.endogMNP'
print(x, digits = max(3, getOption("digits") - 3), ...)
```

Arguments

`object`	An output object from `endogMNP`.
`CI`	A 2 dimensional vector of lower and upper bounds for the credible intervals used to summarize the results. The default is the equal tail 95 percent credible interval.
`discard`	Discard the observations before `discard` as burn-in.
`x`	An object of class `summary.endogMNP`.
`digits`	the number of significant digits to use when printing.
`...`	further arguments passed to or from other methods.

Value

`summary.endogMNP` yields an object of class `summary.endogMNP` containing the following elements:

`call`	The call from `endogMNP`.
`selBase`	The base category for the selection model.
`outBase`	The base category for the outcome model.
`n.obs`	The number of observations.
`n.param`	The number of estimated parameters (including fixed elements of the covariance).
`n.draws`	The number of Gibbs draws stored.
`coef.table`	The summary of the posterior distribution of the coefficients.
`cov.table`	The summary of the posterior distribution of the covariance matrix.

This object can be printed by `print.summary.enodgMNP`

Author(s)

Lane Burgette, Department of Statistical Science, Duke University, <`lb131@stat.duke.edu`>

See Also

`endogMNP`

Index