bayes: heckprobit — Bayesian probit model with sample selection

Description

bayes: heckprobit fits a Bayesian sample-selection probit regression to a partially observed binary outcome; see [BAYES] bayes and [R] heckprobit for details.

Quick start

Bayesian sample-selection probit regression of y on x1 and x2, using z1 and z2 to model selection and using default normal priors for regression coefficients and atanh-correlation

\[
\text{bayes: heckprobit y x1 x2, select(z1 z2)}
\]

Use a standard deviation of 10 instead of 100 for the default normal priors

\[
\text{bayes, normalprior(10): heckprobit y x1 x2, select(z1 z2)}
\]

Use uniform priors for the slopes and a normal prior for the intercept of the main regression

\[
\text{bayes, prior({y: x1 x2}, uniform(-10,10)) ///}
\text{prior({y:_cons}, normal(0,10)): heckprobit y x1 x2, select(z1 z2)}
\]

Save simulation results to simdata.dta, and use a random-number seed for reproducibility

\[
\text{bayes, saving(simdata) rseed(123):, ///}
\text{heckprobit y x1 x2, select(z1 z2)}
\]

Specify 20,000 MCMC samples, set length of the burn-in period to 5,000, and request that a dot be displayed every 500 simulations

\[
\text{bayes, mcmcsize(20000) burnin(5000) dots(500):, ///}
\text{heckprobit y x1 x2, select(z1 z2)}
\]

In the above, request that the 90% HPD credible interval be displayed instead of the default 95% equal-tailed credible interval

\[
\text{bayes, clevel(90) hpd}
\]

Also see Quick start in [BAYES] bayes and Quick start in [R] heckprobit.

Menu

Statistics > Binary outcomes > Bayesian regression > Probit model with sample selection
Syntax

bayes [, bayesopts] : heckprobit depvar indepvars [if] [in] [weight],
                  select([depvars =] varlist [, noconstant offset(varname)]) [options]

options                Description

Model
*select()               specify selection equation: dependent and independent
                        variables; whether to have constant term and offset variable
noconstant              suppress constant term
offset(varname)         include varname in model with coefficient constrained to 1

Reporting
display_options         control spacing, line width, and base and empty cells
level(#)                set credible level; default is level(95)

*select() is required.

The full specification is select([depvars =] varlist [, noconstant offset(varname)])
 indepvars and varlist may contain factor variables; see [U] 11.4.3 Factor variables.
depvar, indepvars, varlist, and depvar may contain time-series operators; see [U] 11.4.4 Time-series varlists.
weights are allowed; see [U] 11.1.6 weight.
bayes: heckprobit, level() is equivalent to bayes, clevel(): heckprobit.
For a detailed description of options, see Options in [R] heckprobit.

bayesopts               Description

Priors
*normalprior(#)        specify standard deviation of default normal priors for regression
                        coefficients and atanh-correlation; default is normalprior(100)
prior(priorspec)        prior for model parameters; this option may be repeated
 dryrun                 show model summary without estimation

Simulation
nchains(#)              number of chains; default is to simulate one chain
mcmcsize(#)            MCMC sample size; default is mcmcsize(10000)
burnin(#)               burn-in period; default is burnin(2500)
thinning(#)            thinning interval; default is thinning(1)
rseed(#)               random-number seed
exclude(paramref)      specify model parameters to be excluded from the simulation results

Blocking
*blocksize(#)          maximum block size; default is blocksize(50)
block(paramref[, blockopts]) specify a block of model parameters; this option may be repeated
 blocksummary          display block summary
*noblocking            do not block parameters by default
Initialization

\texttt{initial(initspec)} specify initial values for model parameters with a single chain\newline\texttt{init#(initspec)} specify initial values for \texttt{#}th chain; requires \texttt{nchains()}
\texttt{initall(initspec)} specify initial values for all chains; requires \texttt{nchains()}
\texttt{nomleinit} suppress the use of maximum likelihood estimates as starting values\newline\texttt{initrandom} specify random initial values\newline\texttt{initsummary} display initial values used for simulation\newline*\texttt{nositely} display output from the estimation command during initialization

Adaptation

\texttt{adaptation(adaptopts)} control the adaptive MCMC procedure\newline\texttt{scale(#)} initial multiplier for scale factor; default is \texttt{scale(2.38)}\newline\texttt{covariance(cov)} initial proposal covariance; default is the identity matrix

Reporting

\texttt{clevel(#)} set credible interval level; default is \texttt{clevel(95)}\newline\texttt{hpd} display HPD credible intervals instead of the default equal-tailed credible intervals\newline\texttt{eform[(string)]} report exponentiated coefficients and, optionally, label as \texttt{string}\newline\texttt{batch(#)} specify length of block for batch-means calculations; default is \texttt{batch(0)}\newline\texttt{saving([filename[, replace]])} save simulation results to \texttt{filename.dta}\newline\texttt{nomodelsummary} suppress model summary\newline\texttt{chainedetail} display detailed simulation summary for each chain\newline\texttt{[no] dots} suppress dots or display dots every 100 iterations and iteration numbers every 1,000 iterations; default is \texttt{nodots}\newline\texttt{dots(#[, every(#)])} display dots as simulation is performed\newline\texttt{[no] show(paramref)} specify model parameters to be excluded from or included in the output\newline\texttt{notable} suppress estimation table\newline\texttt{noheader} suppress output header\newline\texttt{title(string)} display \texttt{string} as title above the table of parameter estimates\newline\texttt{display_options} control spacing, line width, and base and empty cells

Advanced

\texttt{search(search_options)} control the search for feasible initial values\newline\texttt{corrlag(#)} specify maximum autocorrelation lag; default varies\newline\texttt{corrtol(#)} specify autocorrelation tolerance; default is \texttt{corrtol(0.01)}

*Starred options are specific to the \texttt{bayes} prefix; other options are common between \texttt{bayes} and \texttt{bayesmh}.
Options \texttt{prior()} and \texttt{block()} may be repeated.
\texttt{priorspec} and \texttt{paramref} are defined in \texttt{[BAYES] bayesmh}.
\texttt{paramref} may contain factor variables; see [U] 11.4.3 Factor variables.
See [U] 20 Estimation and postestimation commands for more capabilities of estimation commands.
Model parameters are regression coefficients  \{\texttt{depvar:indepvars}\} for the main regression and \{\texttt{select:varlist\_}\} for the selection equation, and atanh-transformed correlation \{\texttt{athrho}\}. Use the \texttt{dryrun} option to see the definitions of model parameters prior to estimation.
For a detailed description of \texttt{bayesopts}, see Options in \texttt{[BAYES] bayes}.\newline
Remarks and examples

For a general introduction to Bayesian analysis, see [BAYES] Intro. For a general introduction to Bayesian estimation using an adaptive Metropolis–Hastings algorithm, see [BAYES] bayesmh. For remarks and examples specific to the bayes prefix, see [BAYES] bayes. For details about the estimation command, see [R] heckprobit.

For a simple example of the bayes prefix, see Introductory example in [BAYES] bayes. Also see Heckman selection model in [BAYES] bayes.

Stored results

See Stored results in [BAYES] bayes.

Methods and formulas

See Methods and formulas in [BAYES] bayesmh.

Also see

[BAYES] bayes — Bayesian regression models using the bayes prefix
[R] heckprobit — Probit model with sample selection
[BAYES] Bayesian postestimation — Postestimation tools for bayesmh and the bayes prefix
[BAYES] Bayesian estimation — Bayesian estimation commands
[BAYES] Bayesian commands — Introduction to commands for Bayesian analysis
[BAYES] Intro — Introduction to Bayesian analysis
[BAYES] Glossary