Bayes: glm — Bayesian generalized linear models

Description
bayes: glm fits a Bayesian generalized linear model to outcomes of different types such as continuous, binary, count, and so on; see [BAYES] bayes and [R] glm for details.

Quick start
Bayesian generalized linear model of \( y \) on \( x_1 \) and \( x_2 \), using the Gaussian family and log link and using default normal priors for regression coefficients
\[
\text{bayes: glm } y \ x_1 \ x_2, \ family(\text{gaussian}) \ link(\text{log})
\]
Use a standard deviation of 10 instead of 100 for the default normal priors
\[
\text{bayes, normalprior}(10): \ glm \ y \ x_1 \ x_2, \ family(\text{gaussian}) \ link(\text{log})
\]
Use uniform priors for the slopes and a normal prior for the intercept
\[
\text{bayes, prior(} \{y: \ x_1 \ x_2\}, \ uniform(-10,10)) /// \\
\text{prior(} \{y:_cons\}, \ normal(0,10)): /// \\
\text{glm } y \ x_1 \ x_2, \ family(\text{gaussian}) \ link(\text{log})
\]
Save simulation results to simdata.dta, and use a random-number seed for reproducibility
\[
\text{bayes, saving(simdata) rseed(123): ///} \\
\text{glm } y \ x_1 \ x_2, \ family(\text{gaussian}) \ link(\text{log})
\]
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
\[
\text{bayes, mcmcsize(20000) burnin(5000) dots(500): ///} \\
\text{glm } y \ x_1 \ x_2, \ family(\text{gaussian}) \ link(\text{log})
\]
In the above, request that the 90\% HPD credible interval be displayed instead of the default 95\% equal-tailed credible interval
\[
\text{bayes, clevel(90) hpd}
\]
Fit a logit model and display results as odds ratios
\[
\text{bayes: glm } z \ x_1 \ x_2, \ family(\text{binomial}) \ eform
\]
Display odds ratios on replay
\[
\text{bayes, eform}
\]
Also see Quick start in [BAYES] bayes and Quick start in [R] glm.

Menu
Statistics > Generalized linear models > Bayesian generalized linear models (GLM)
Syntax

```
bayes [, bayesopts] : glm depvar [ indepvars ] [ if ] [ in ] [ weight ] [, options]
```

<table>
<thead>
<tr>
<th>options</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Model</td>
<td></td>
</tr>
<tr>
<td><code>family(familyname)</code></td>
<td>distribution of <code>depvar</code>; default is <code>family(gaussian)</code></td>
</tr>
<tr>
<td><code>link(linkname)</code></td>
<td>link function; default is canonical link for <code>family()</code> specified</td>
</tr>
<tr>
<td>Model 2</td>
<td></td>
</tr>
<tr>
<td><code>noconstant</code></td>
<td>suppress constant term</td>
</tr>
<tr>
<td><code>exposure(varname)</code></td>
<td>include ln(<code>varname</code>) in model with coefficient constrained to 1</td>
</tr>
<tr>
<td><code>offset(varname)</code></td>
<td>include <code>varname</code> in model with coefficient constrained to 1</td>
</tr>
<tr>
<td><code>asis</code></td>
<td>retain perfect predictor variables</td>
</tr>
<tr>
<td><code>mu(varname)</code></td>
<td>use <code>varname</code> as the initial estimate for the mean of <code>depvar</code></td>
</tr>
<tr>
<td><code>init(varname)</code></td>
<td>synonym for <code>mu(varname)</code></td>
</tr>
<tr>
<td>Reporting</td>
<td></td>
</tr>
<tr>
<td><code>eform</code></td>
<td>report exponentiated coefficients</td>
</tr>
<tr>
<td><code>display_options</code></td>
<td>control spacing, line width, and base and empty cells</td>
</tr>
<tr>
<td><code>level(#)</code></td>
<td>set credible level; default is <code>level(95)</code></td>
</tr>
</tbody>
</table>

`indepvars` may contain factor variables; see [U] 11.4.3 Factor variables.
`depvar` and `indepvars` may contain time-series operators; see [U] 11.4.4 Time-series varlists.
Fweights are allowed; see [U] 11.1.6 weight.
bayes: glm, level() is equivalent to `bayes, clevel(): glm`.
For a detailed description of `options`, see Options in [R] glm.

bayesopts

<table>
<thead>
<tr>
<th>Priors</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td><code>normalprior(#)</code></td>
<td>specify standard deviation of default normal priors for regression coefficients; default is <code>normalprior(100)</code></td>
</tr>
<tr>
<td><code>prior(priorspec)</code></td>
<td>prior for model parameters; this option may be repeated show model summary without estimation</td>
</tr>
<tr>
<td><code>dryrun</code></td>
<td></td>
</tr>
</tbody>
</table>

Simulation

<table>
<thead>
<tr>
<th>Simulation</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td><code>nchains(#)</code></td>
<td>number of chains; default is to simulate one chain</td>
</tr>
<tr>
<td><code>mcmcsize(#)</code></td>
<td>MCMC sample size; default is <code>mcmcsize(10000)</code></td>
</tr>
<tr>
<td><code>burnin(#)</code></td>
<td>burn-in period; default is <code>burnin(2500)</code></td>
</tr>
<tr>
<td><code>thinning(#)</code></td>
<td>thinning interval; default is <code>thinning(1)</code></td>
</tr>
<tr>
<td><code>rseed(#)</code></td>
<td>random-number seed</td>
</tr>
<tr>
<td><code>exclude(paramref)</code></td>
<td>specify model parameters to be excluded from the simulation results</td>
</tr>
</tbody>
</table>

Blocking

<table>
<thead>
<tr>
<th>Blocking</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td><code>blocksize(#)</code></td>
<td>maximum block size; default is <code>blocksize(50)</code></td>
</tr>
<tr>
<td><code>block(paramref[, blockopts])</code></td>
<td>specify a block of model parameters; this option may be repeated display block summary</td>
</tr>
<tr>
<td><code>noblocking</code></td>
<td>do not block parameters by default</td>
</tr>
</tbody>
</table>
Initialization

\textbf{\texttt{initial}}(\texttt{initspec}) specify initial values for model parameters with a single chain
\texttt{init\#(initspec)} specify initial values for \#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
\texttt{initsummary} display initial values used for simulation
\texttt{*noisily} display output from the estimation command during initialization

Adaptation

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

Reporting

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

Advanced

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

\dag 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 \texttt{[U] 11.4.3 Factor variables}.
See \texttt{[U]} \texttt{20 Estimation and postestimation commands} for more capabilities of estimation commands.
Model parameters are regression coefficients \{\texttt{depvar:indepvars}\}. Use the \texttt{dryrun} option to see the definitions of model parameters prior to estimation.
For a detailed description of \texttt{bayesopts}, see \texttt{Options} in \texttt{[BAYES] bayes}.
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] glm.

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

bayes: glm does not estimate the scale parameter but uses a fixed value as provided by the glm command. If you want to fit a GLM and estimate the scale parameter, use bayes: meglm without specifying random effects.

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] glm — Generalized linear models
[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