bayes: xtmlogit — Bayesian random-effects multinomial logit model

Description

`bayes: xtmlogit` fits a Bayesian panel-data random-effects multinomial logit model to categorical outcomes; see [BAYES] **bayes** and [XT] **xtmlogit** for details.

Quick start

Bayesian random-effects multinomial logit model of `y` on `x1` and `x2` with random intercepts by `id` (after `xtset`ing on panel variable `id`), using default normal priors for regression coefficients and default inverse-gamma prior for the variance of random intercepts

```
bayes: xtmlogit y x1 x2
```

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

```
bayes, normalprior(10): xtmlogit y x1 x2
```

Use a shape of 1 and a scale of 2 instead of values of 0.01 for the default inverse-gamma prior

```
bayes, igammaprior(1 2): xtmlogit y x1 x2
```

Use uniform priors for the slopes and a normal prior for the intercept

```
bayes, prior({y: x1 x2}, uniform(-10,10)) ///
prior({y:_cons}, normal(0,10)): xtmlogit y x1 x2
```

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

```
bayes, saving(simdata) rseed(123): xtmlogit y x1 x2
```

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

```
bayes, mcmcsize(20000) burnin(5000) dots(500): xtmlogit y x1 x2
```

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

```
bayes, clevel(90) hpd
```

Bayesian random-effects multinomial logit model of `y` on `x1` and `x2`, with the second outcome as the base outcome

```
bayes: xtmlogit y x1 x2, baseoutcome(2)
```

As above, but report relative-risk ratios

```
bayes: xtmlogit y x1 x2, baseoutcome(2) rrr
```

As above, but using shared random-effects covariance between outcomes

```
bayes: xtmlogit y x1 x2, baseoutcome(2) covariance(shared) rrr
```

Also see *Quick start* in [BAYES] **bayes** and *Quick start* in [XT] **xtmlogit**.

Menu

Statistics > Longitudinal/panel data > Categorical outcomes > Bayesian regression > Multinomial logistic regression

Syntax

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

options	Description
Model	
`noconstant`	suppress constant term
`baseoutcome(#)`	value of *depvar* that will be the base outcome; default is the last outcome level
`covariance(`*vartype*`)`	variance–covariance structure of the random effects; default is `covariance(independent)`
Reporting	
`rrr`	report relative-risk ratios
display_options	control spacing, line width, and base and empty cells
`level(#)`	set credible level; default is `level(95)`

A panel variable must be specified; see [XT] **xtset**.

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**.

`fweight`s are allowed; see [U] **11.1.6 weight**.

`bayes: xtmlogit, level()` is equivalent to `bayes, clevel(): xtmlogit`.

For a detailed description of options, see *Options* in [XT] **xtmlogit**.

bayesopts	Description
Priors	
* `normalprior(#)`	specify standard deviation of default normal priors for regression coefficients; default is `normalprior(100)`
* `igammaprior(# #)`	specify shape and scale of default inverse-gamma prior for variance components; default is `igammaprior(0.01 0.01)`
* `iwishartprior(#` [...]`)`	specify degrees of freedom and, optionally, scale matrix of default inverse-Wishart prior for unstructured random-effects covariance
`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	
block(*paramref* [, *blockopts*])	specify a block of model parameters; this option may be repeated
blocksummary	display block summary
Initialization	
initial(*initspec*)	specify initial values for model parameters with a single chain
init#(*initspec*)	specify initial values for #th chain; requires nchains()
initall(*initspec*)	specify initial values for all chains; requires nchains()
nomleinitial	suppress the use of maximum likelihood estimates as starting values
initrandom	specify random initial values
initsummary	display initial values used for simulation
* noisily	display output from the estimation command during initialization
Adaptation	
adaptation(*adaptopts*)	control the adaptive MCMC procedure
scale(#)	initial multiplier for scale factor; default is scale(2.38)
covariance(*cov*)	initial proposal covariance; default is the identity matrix
Reporting	
clevel(#)	set credible interval level; default is clevel(95)
hpd	display HPD credible intervals instead of the default equal-tailed credible intervals
* rrr	report relative-risk ratios
eform[(*string*)]	report exponentiated coefficients and, optionally, label as *string*
remargl	compute log marginal-likelihood; suppressed by default
batch(#)	specify length of block for batch-means calculations; default is batch(0)
saving(*filename* [, replace])	save simulation results to *filename*.dta
nomodelsummary	suppress model summary
chainsdetail	display detailed simulation summary for each chain
[no]dots	suppress dots or display dots every 100 iterations and iteration numbers every 1,000 iterations; default is nodots
dots(#[, every(#)])	display dots as simulation is performed
[no]show(*paramref*)	specify model parameters to be excluded from or included in the output
showreffects[(*reref*)]	specify that all or a subset of random-effects parameters be included in the output
notable	suppress estimation table
noheader	suppress output header
title(*string*)	display *string* as title above the table of parameter estimates
display_options	control spacing, line width, and base and empty cells
Advanced	
search(*search_options*)	control the search for feasible initial values
corrlag(#)	specify maximum autocorrelation lag; default varies
corrtol(#)	specify autocorrelation tolerance; default is corrtol(0.01)

* Starred options are specific to the bayes prefix; other options are common between bayes and bayesmh.

The full specification of iwishartprior() is iwishartprior(# [*matname*] [, relevel(*levelvar*)]).

Options prior() and block() may be repeated.

priorspec and *paramref* are defined in [BAYES] **bayesmh**.

paramref may contain factor variables; see **[U] 11.4.3 Factor variables**.

collect is allowed; see **[U] 11.1.10 Prefix commands**.

See **[U] 20 Estimation and postestimation commands** for more capabilities of estimation commands.

Model parameters are regression coefficients {*outcome*$_1$:*indepvars*}, {*outcome*$_2$:*indepvars*}, and so on, where *outcome*$_\#$'s are the values of the dependent variable or the value labels of the dependent variable if they exist, random effects {U#[*panelvar*]} or simply {U#}, and random-effects variances {var_U#} or, if random effects are correlated, covariance {U:Sigma,m}; see *Methods and formulas* for a full list of parameters. Use the dryrun option to see the definitions of model parameters prior to estimation.

For a detailed description of *bayesopts*, see *Options* in [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 [XT] **xtmlogit**.

For a simple example of the bayes prefix, see *Introductory example* in [BAYES] **bayes**. Also see *Panel-data models* in [BAYES] **bayes**.

▷ Example 1

Let's revisit example 1 from [XT] **xtmlogit**. The example uses a fictional estatus dataset to model women employment status, estatus, as a function of various socioeconomic factors such as having children under 18 years of age, hhchild; age; household income, hhincome; having significant other, hhsigno; and whether the woman is the primary breadwinner, bwinner. The employment status falls into three categories: employed, unemployed, and out of labor force.

Women are identified by the id variable, which is declared as the panel variable.

```
. use https://www.stata-press.com/data/r19/estatus
(Fictional employment status data)

. xtset id

Panel variable: id (unbalanced)
```

Let's fit a Bayesian analog of the model from example 1 of [XT] **xtmlogit**. The dataset contains 800 random effects and a total of 4,761 observations. To speed up the execution, we reduce the MCMC sample size from the default of 10,000 to 1,000, and we specify the `rseed()` option for reproducibility.

```
. bayes, rseed(17) mcmcsize(1000): xtmlogit estatus i.hhchild age hhincome
> i.hhsigno i.bwinner
note: Gibbs sampling is used for variance components.

Burn-in 2500 aaaaaaaaa1000aaaaaaaaa2000aaaaa done
Simulation 1000 .........1000 done

Model summary
------------------------------------------------------------------------------

Likelihood:
  estatus ~ mlogit(xb_Out_of_labor_force,xb_Unemployed)

Priors:
  {Out_of_lab~e:1.hhchild} ~ normal(0,10000)                               (1)
        {Out_of_lab~e:age} ~ normal(0,10000)                               (1)
   {Out_of_lab~e:hhincome} ~ normal(0,10000)                               (1)
  {Out_of_lab~e:1.hhsigno} ~ normal(0,10000)                               (1)
  {Out_of_lab~e:1.bwinner} ~ normal(0,10000)                               (1)
      {Out_of_lab~e:_cons} ~ normal(0,10000)                               (1)
                  {U1[id]} ~ normal(0,{var_U1})                            (1)
    {Unemployed:1.hhchild} ~ normal(0,10000)                               (2)
          {Unemployed:age} ~ normal(0,10000)                               (2)
     {Unemployed:hhincome} ~ normal(0,10000)                               (2)
    {Unemployed:1.hhsigno} ~ normal(0,10000)                               (2)
    {Unemployed:1.bwinner} ~ normal(0,10000)                               (2)
        {Unemployed:_cons} ~ normal(0,10000)                               (2)
                  {U2[id]} ~ normal(0,{var_U2})                            (2)

Hyperprior:
  {var_U1 var_U2} ~ igamma(0.01,0.01)
------------------------------------------------------------------------------
(1) Parameters are elements of the linear form xb_Out_of_labor_force.
(2) Parameters are elements of the linear form xb_Unemployed.
```

```
Bayesian RE multinomial logistic regression      MCMC iterations  =      3,500
Metropolis-Hastings and Gibbs sampling           Burn-in          =      2,500
                                                 MCMC sample size =      1,000
Group variable: id                               Number of groups =        800
                                                 Obs per group:
                                                              min =          5
                                                              avg =        6.0
                                                              max =          7
Base outcome: Employed                           Number of obs    =      4,761
                                                 Acceptance rate  =       .462
                                                 Efficiency:  min =      .0067
                                                              avg =     .02054
Log marginal-likelihood                                       max =     .03473

------------------------------------------------------------------------------
             |                                                Equal-tailed
             |      Mean   Std. dev.     MCSE     Median  [95% cred. interval]
-------------+----------------------------------------------------------------
Out_of_lab~e |
     hhchild |
        Yes  |  .4577437   .0904496   .017864   .4640043   .2710479   .6218431
         age | -.002879   .0055965   .001219  -.0026383  -.0130767   .0085352
    hhincome | -.0042843   .0018489   .000402  -.0040465  -.0083297  -.0014658
             |
     hhsigno |
        Yes  |  .4691271   .0889745   .017166   .4582264   .3251738   .6559253
             |
     bwinner |
        Yes  | -.4503803   .0732228    .01895  -.4500302  -.5924365  -.3002816
          U1 |         1          0         0          1          1          1
       _cons | -.5534515   .2478516   .060647  -.5376768  -1.010935  -.0486457
-------------+----------------------------------------------------------------
Unemployed   |
     hhchild |
        Yes  | -.0519455   .1168531   .023891  -.0398437  -.2755692   .1858482
         age |  .0092687   .0075203   .001441   .0091324  -.0050356   .0250353
    hhincome | -.0293463   .0030542    .00118  -.0293997  -.0356738  -.0227989
             |
     hhsigno |
        Yes  |  .0412739    .114903   .021569   .0361712  -.1694103   .2494766
             |
     bwinner |
        Yes  | -.1812031   .1003491   .033786  -.1773746  -.3642266   .0072658
          U2 |         1          0         0          1          1          1
       _cons | -.3242398   .3382746   .100034  -.3894121  -.9363997    .335811
-------------+----------------------------------------------------------------
      var_U1 |  .8864246   .0884571    .01501   .8815608   .7235478   1.060998
      var_U2 |  .7769171   .1137603   .025427    .757853    .605616   1.036373
------------------------------------------------------------------------------
Note: Default priors are used for model parameters.
Note: Adaptation tolerance is not met in at least one of the blocks.
```

Because the `Employed` outcome level is selected as the base outcome, the results are reported only for the `Out_of_labor_force` and `Unemployed` outcome levels. The posterior mean estimates for regression coefficients and variances of random effects are similar to the maximum likelihood estimates from example 1 from [XT] **xtmlogit**.

The Bayesian model introduced one set of random intercepts for each outcome level except the base outcome: `{U1[id]}` and `{U2[id]}`. By default, the random effects are assigned independent normal priors with variances `{var_U1}` and `{var_U2}`, respectively.

Following the original example, we can obtain estimates of relative-risk ratios by specifying the `rrr` option with `bayes`.

```
. bayes, rrr noheader

----------------------------------------------------------------------------------
             |                                                Equal-tailed
             |       RRR   Std. dev.      MCSE     Median  [95% cred. interval]
-------------+--------------------------------------------------------------------
Out_of_lab~e |
     hhchild |
        Yes  |  1.586934   .1424381    .028002    1.59043   1.311338   1.862358
         age |  .9971407   .0055807    .001215   .9973652   .9870084   1.008572
    hhincome |  .9957265   .0018402      .0004   .9959617   .9917049   .9985352
             |
     hhsigno |
        Yes  |  1.605004   .1453829    .028037   1.581267   1.384271   1.926925
             |
     bwinner |
        Yes  |  .6390981   .0469746    .012111   .6376089   .5529823   .7406096
          U1 |  2.718282          0          0   2.718282   2.718282   2.718282
       _cons |  .5928508      .1479    .036194   .5841037   .3638785   .9525185
-------------+--------------------------------------------------------------------
Unemployed   |
     hhchild |
        Yes  |  .9559229   .1133981    .023185   .9609396   .7591399   1.204239
         age |   1.00934   .0075995    .001456   1.009174    .994977   1.025351
    hhincome |  .9710847   .0029661    .001146   .9710282    .964955    .977459
             |
     hhsigno |
        Yes  |  1.049005   .1200204    .022401   1.036833   .8441625   1.283353
             |
     bwinner |
        Yes  |  .8384893   .0847624    .028742    .837466   .6947338   1.007292
          U2 |  2.718282          0          0   2.718282   2.718282   2.718282
       _cons |  .7657877   .2650145    .074494    .677455   .3921156   1.399075
-------------+--------------------------------------------------------------------
      var_U1 |  .8864246   .0884571     .01501   .8815608   .7235478   1.060998
      var_U2 |  .7769171   .1137603    .025427    .757853    .605616   1.036373
----------------------------------------------------------------------------------
```

The original example also estimated marginal probabilities with respect to the `hhchild` variable using the `margins` command. Below, we demonstrate Bayesian estimation of these marginal probabilities using Bayesian predictions.

First, we save the simulation results produced by `bayes: xtmlogit` to a permanent Stata dataset.

```
. bayes, saving(xtmlogitsim, replace)
note: file xtmlogitsim.dta saved.
```

We then define a Stata program, margprob, that calculates the marginal probabilities based on the simulated outcomes. See *User-defined Stata programs* in [BAYES] **bayespredict** for details.

```
. program margprob
  1.         version 19.5      // (or version 19 if you do not have StataNow)
  2.         args sum ysim
  3.         local xvar $BAYESPR_extravars
  4.         local ylabel $BAYESPR_passthruopts
  5.         gettoken ylabel xlabel : ylabel
  6.         tempvar presid
  7.         generate byte `presid' = `ysim' == `ylabel' if `xvar' == `xlabel'
  8.         summarize `presid', meanonly
  9.         scalar `sum' = r(mean)
 10. end
```

In addition to the simulated outcome `ysim', the program uses the conditional variable `xvar', hhchild in our example, passed as an extra variable, and two indices `ylabel' and `xlabel' that specify the outcome category and the conditional variable category, respectively. `ylabel' takes values 1 , 2, and 3, and `xlabel' takes values 0 and 1. `ylabel' and `xlabel' values are specified in the passthruopts() options of bayespredict. To calculate all marginal probabilities, we need to call the program for all six combinations of ylabel and xlabel.

Given the size of the dataset, calculating the Bayesian marginal probabilities using a user-defined Stata program is time consuming and will take a couple of minutes. We specify the dots option with bayespredict to monitor the simulation progress.

```
. bayespredict
> (pr1childNo :@margprob {_ysim1}, extravars(hhchild) passthruopts(1 0))
> (pr1childYes:@margprob {_ysim1}, extravars(hhchild) passthruopts(1 1))
> (pr2childNo :@margprob {_ysim1}, extravars(hhchild) passthruopts(2 0))
> (pr2childYes:@margprob {_ysim1}, extravars(hhchild) passthruopts(2 1))
> (pr3childNo :@margprob {_ysim1}, extravars(hhchild) passthruopts(3 0))
> (pr3childYes:@margprob {_ysim1}, extravars(hhchild) passthruopts(3 1)),
>         saving(xtmlogitpred, replace) rseed(17) dots
Computing predictions 1000 .........1000 done

file xtmlogitpred.dta saved.
file xtmlogitpred.ster saved.
```

The posterior predicted marginal probabilities are saved as xtmlogitpred estimation results.

Finally, we use bayesstats summary to calculate posterior estimates of the marginal probabilities.

```
. bayesstats summary {pr1childNo} {pr1childYes}
> {pr2childNo} {pr2childYes}
> {pr3childNo} {pr3childYes} using xtmlogitpred

Posterior summary statistics                        MCMC sample size =     1,000
```

	Mean	Std. dev.	MCSE	Median	Equal-tailed [95% cred.	interval]
pr1childNo	.3001763	.0133316	.000946	.3004053	.2745694	.3259878
pr1childYes	.3909447	.0113229	.00092	.3914604	.368676	.4133477
pr2childNo	.1615598	.0114636	.001266	.1616008	.1377913	.1843972
pr2childYes	.1368543	.0088284	.000782	.1367061	.1205597	.1546466
pr3childNo	.5382639	.0150199	.001759	.5382472	.508612	.5678825
pr3childYes	.4722009	.0114689	.001325	.4721923	.4497668	.4949767

Because we used uninformative default priors, the reported posterior mean estimates are close to the marginal probabilities calculated by the `margins` command.

◁

Stored results

See *Stored results* in [BAYES] **bayes**. In addition, `bayes: xtmlogit` also stores the following results:

Macros

`e(ivar)`	variable denoting groups
`e(baseoutcome)`	base outcome
`e(redistrib)`	distribution of random effects
`e(covariance)`	random-effects covariance structure

Methods and formulas

Bayesian random-effects multinomial logit models are based on random-effects multinomial logit models described in *Methods and formulas* of [XT] **xtmlogit**.

A multinomial logit model for a dependent variable with J outcome levels has $J-1$ equations, ignoring the baseline outcome, each having its own set of random intercepts. The equation for the #th outcome level includes a random-effects parameter `{U#[`*panelvar*`]}`, where *panelvar* is the panel variable. You can also refer to the random-effects parameters simply as `{U#}`. Random effects `{U#}`'s can be independent, shared, or correlated.

Independent `{U#}`'s, `covariance(independent)`, are assigned independent normal priors with zero means and random-effects variances `{var_U#}`'s. The default prior for `{var_U#}` is an inverse-gamma distribution with shape and scale of 0.01. You can use the `igammaprior()` options to change the default shape and scale parameters.

For a shared covariance structure, `covariance(shared)`, there is one random-effects parameter, `{U[`*panelvar*`]}`, shared between the outcome-level equations.

For an identity covariance structure, `covariance(identity)`, the random effects `{U#[`*panelvar*`]}` are different but have the same prior variance `{var_U}`.

For an exchangeable covariance structure, `covariance(exchangeable)`, `{U#[`*panelvar*`]}`'s are assigned `mvn0exchangeable(`$J-1$`, {var_U}, {rho_U})` prior. The default prior for the correlation parameter `{rho_U}` is uniform on $(-1,1)$.

For an unstructured covariance, `covariance(unstructured)`, `{U#[`*panelvar*`]}`'s are assigned `mvn0(`$J-1$`, {U:Sigma,m})` prior. The default hyperprior for the variance–covariance matrix `{U:Sigma,m}` is inverse-Wishart with J degrees of freedom and the identity scale matrix. You can use the `iwishartprior()` option to change the default degrees of freedom and scale matrix.

See *Methods and formulas* in [BAYES] **bayesmh**.

Also see

[BAYES] **bayes** — Bayesian regression models using the bayes prefix

[XT] **xtmlogit** — Fixed-effects and random-effects multinomial logit models

[BAYES] **Bayesian postestimation** — Postestimation tools after Bayesian estimation

[BAYES] **Bayesian estimation** — Bayesian estimation commands

[BAYES] **Bayesian commands** — Introduction to commands for Bayesian analysis

[BAYES] **Intro** — Introduction to Bayesian analysis

[BAYES] **Glossary**

Stata, Stata Press, and Mata are registered trademarks of StataCorp LLC. Stata and Stata Press are registered trademarks with the World Intellectual Property Organization of the United Nations. StataNow and NetCourseNow are trademarks of StataCorp LLC. Other brand and product names are registered trademarks or trademarks of their respective companies. Copyright © 1985–2025 StataCorp LLC, College Station, TX, USA. All rights reserved.

For suggested citations, see the FAQ on citing Stata documentation.