bayes: xtpoisson — Bayesian random-effects Poisson model

Description

`bayes: xtpoisson` fits a Bayesian panel-data random-effects Poisson model to a nonnegative count outcome; see [BAYES] **bayes** and [XT] **xtpoisson** for details.

Quick start

Bayesian random-effects Poisson 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: xtpoisson y x1 x2
```

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

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

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

```
bayes, saving(simdata) rseed(123): xtpoisson 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): xtpoisson 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
```

Use a normal prior distribution for random effects instead of the default gamma prior

```
bayes: xtpoisson y x1 x2, normal
```

Display incidence-rate ratios instead of coefficients

```
bayes: xtpoisson y x1 x2, irr
```

Display incidence-rate ratios on replay

```
bayes, irr
```

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

Menu

Statistics > Longitudinal/panel data > Count outcomes > Bayesian regression > Poisson regression

Syntax

bayes [, *bayesopts*] : xtpoisson *depvar* [*indepvars*] [*if*] [*in*] [, *options*]

options	Description
Model	
noconstant	suppress constant term
exposure(*varname*$_e$)	include ln(*varname*$_e$) in model with coefficient constrained to 1
offset(*varname*$_o$)	include *varname*$_o$ in model with coefficient constrained to 1
normal	use a normal distribution for random effects instead of gamma
Reporting	
irr	report incidence-rate 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, *indepvars*, *varname*$_e$, and *varname*$_o$ may contain time-series operators; see **[U] 11.4.4 Time-series varlists**.

bayes: xtpoisson, level() is equivalent to bayes, clevel(): xtpoisson.

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

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)
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
* irr	report incidence-rate 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.

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 {*depvar*:*indepvars*}, random effects {U[*panelvar*]} or simply {U}, and parameter {alpha} with the gamma prior or random-effects variance {var_U} with the normal prior; also see *Methods and formulas*. 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] **xtpoisson**.

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] **xtpoisson**. The example models the number of ship accidents, accident, affected by the period of their construction and operation. The factor variables co_75_79, co_70_74, and co_65_69 mark consecutive construction periods of 5 years, and op_75_79 indicates the operating period between 1975 and 1979.

```
. use https://www.stata-press.com/data/r19/ships
. xtset
Panel variable: ship (balanced)
```

The number of accidents is modeled by a Poisson distribution with the number of months in service, service, as exposure. The ship variable identifies the individual ships and is set as the panel variable.

We use bayes: xtpoisson to fit the Bayesian analog of the model. We use the default priors for regression coefficients and random effects. The random effects are assigned an exponential gamma prior with a hyperparameter {alpha}. The latter is assigned an inverse-gamma hyperprior. To improve sampling efficiency, we double the burn-in period, burnin(5000). We also include the irr option to report incidence-rate ratios instead of regression coefficients.

```
. bayes, burnin(5000) rseed(17):
> xtpoisson accident op_75_79 co_65_69 co_70_74 co_75_79, exp(service) irr
Burn-in 5000 aaaaaaaaa1000aaaaaaaaa2000aaaaaaaaa3000aaaaaaaaa4000aaaaaaaaa5000
> done
Simulation 10000 .........1000.........2000.........3000.........4000.........
> 5000.........6000.........7000.........8000.........9000.........10000 done
Model summary
------------------------------------------------------------------------------
Likelihood:
  accident service ~ poissonreg(xb_accident)

Priors:
  {accident:op_75_79} ~ normal(0,10000)                                    (1)
  {accident:co_65_69} ~ normal(0,10000)                                    (1)
  {accident:co_70_74} ~ normal(0,10000)                                    (1)
  {accident:co_75_79} ~ normal(0,10000)                                    (1)
     {accident:_cons} ~ normal(0,10000)                                    (1)
            {U[ship]} ~ expgamma(1/{alpha},{alpha})                        (1)

Hyperprior:
  {alpha} ~ igamma(0.01,0.01)
------------------------------------------------------------------------------
(1) Parameters are elements of the linear form xb_accident.

Bayesian RE Poisson regression                   MCMC iterations  =     15,000
Random-walk Metropolis-Hastings sampling         Burn-in          =      5,000
                                                 MCMC sample size =     10,000
Group variable: ship                             Number of groups =          5
                                                 Obs per group:
                                                              min =          6
                                                              avg =        6.8
                                                              max =          7
                                                 Number of obs    =         34
                                                 Acceptance rate  =      .4103
                                                 Efficiency:  min =    .004533
                                                              avg =     .02627
Log marginal-likelihood                                       max =     .06637

------------------------------------------------------------------------------
             |                                                Equal-tailed
             |       IRR   Std. dev.     MCSE     Median  [95% cred. interval]
-------------+----------------------------------------------------------------
accident     |
    op_75_79 |  1.482028   .1872034   .012245   1.466002    1.15391   1.885356
    co_65_69 |  2.056534   .3147425   .012217   2.038204   1.516147   2.745889
    co_70_74 |  2.365398   .4163733   .027752    2.31289   1.673906   3.377834
    co_75_79 |  1.641278    .386874   .024248   1.610142   1.021594   2.514659
       _cons |  .0014965    .000378   .000056   .0014293   .0009432   .0024066
-------------+----------------------------------------------------------------
       alpha |   .182512    .149803   .012089   .1351156   .0271875    .606201
------------------------------------------------------------------------------
Note: Variable service is included in the model as the exposure.
Note: _cons estimates baseline incidence rate.
Note: Default priors are used for model parameters.
```

The posterior mean estimates for regression coefficients are similar to the maximum likelihood estimates reported in example 1. The posterior mean estimate for {alpha}, about 0.18, is greater than its maximum likelihood counterpart, 0.09, because its marginal posterior distribution is skewed.

We can use bayesstats summary to report posterior estimates for the random effects {U[ship]}.

```
. bayesstats summary {U[1/5]}

Posterior summary statistics                      MCMC sample size =    10,000
```

U[ship]	Mean	Std. dev.	MCSE	Median	Equal-tailed [95% cred.	interval]
1	.0603	.2287246	.028578	.0650362	-.4104326	.5109287
2	-.4250167	.2156037	.035458	-.4172961	-.8843667	-.0511939
3	-.422064	.3049497	.032655	-.3893351	-1.115965	.0824852
4	-.0106956	.2549407	.026575	-.0067523	-.5325561	.4791908
5	.3031554	.2326076	.025797	.3001452	-.1498397	.7672204

Next, we would like to assess the goodness of fit of the model by using bayespredict and bayesstats ppvalues to perform posterior predictive checks. But first, we need to save the current simulation results to a permanent Stata dataset.

```
. bayes, saving(xtpoissim1)
note: file xtpoissim1.dta saved.
```

Deviance is commonly used as a goodness-of-fit statistic for generalized linear models. We define a Mata function, deviance(), that computes the deviance, which will be used by bayespredict to compute the deviance based on the simulated outcome ysim and the mean vector mu.

```
. mata:
------------------------------------------------- mata (type end to exit) -----
: real scalar deviance(real colvector ysim, real colvector mu) {
>         return (2*sum(ysim:*ln(ysim:/mu):-ysim:+mu))
> }
: end
-------------------------------------------------------------------------------
```

Next, we call bayespredict to compute the deviance of outcomes simulated from the posterior predictive distribution and save the results in xtpoispred1.

```
. bayespredict (@deviance({_ysim1},{_mu1})), rseed(17) saving(xtpoispred1)
Computing predictions ...
file xtpoispred1.dta saved.
file xtpoispred1.ster saved.
```

Now, we can compute the posterior predictive p-value of the deviance statistics using the bayesstats ppvalues command.

```
. bayesstats ppvalues using xtpoispred1

Posterior predictive summary      MCMC sample size =    10,000
```

T	Mean	Std. dev.	E(T_obs)	P(T>=T_obs)
_ysim1_deviance	25.02129	7.157104	39.40344	.0523

```
Note: P(T>=T_obs) close to 0 or 1 indicates lack of fit.
```

The estimated p-value is only 0.05, but in the absence of a reference model, it is difficult to decide whether this indicates a lack of fit.

◁

Stored results

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

Macros

`e(ivar)`	variable denoting groups
`e(redistrib)`	distribution of random effects

Methods and formulas

Bayesian random-effects Poisson models are based on random-effects Poisson models described in *Methods and formulas* of [XT] **xtpoisson**.

Let y_{it} be the count for the tth observation in the ith group. We assume $y_{it} \mid u_i, \lambda_{it} \sim \text{Poisson}\{\exp(u_i)\lambda_{it}\}$, with $\lambda_{it} = \exp(\mathbf{x}_{it}\boldsymbol{\beta} + \text{offset}_{it})$ and u_i a parameter that varies randomly across groups. In `bayes: xtpoisson`, parameters u_i's are represented by `{U[`*panelvar*`]}`, where *panelvar* is the panel variable.

By default, random effects $\exp(u_i)$ are a priori independent and have a gamma prior distribution with mean 1 and variance α. u_i's are thus assigned an exponential gamma prior with shape $1/\alpha$ and scale α. The hyperparameter α, `{alpha}` in the output of `bayes: xtpoisson`, has an inverse-gamma prior with shape and scale of 0.01.

When the `normal` option is specified with `xtpoisson`, the random effects u_i's are assigned a normal prior distribution with mean 0 and variance σ_u^2, denoted as `{var_U}` in the output of `bayes: xtpoisson`. By default, σ_u^2 is assigned an inverse-gamma prior with shape and scale of 0.01.

You can use the `igammaprior()` option to change the shape and scale of the default inverse-gamma prior. See *Methods and formulas* in [BAYES] **bayesmh**.

Also see

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

[XT] **xtpoisson** — Fixed-effects, random-effects, and population-averaged Poisson 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**

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